

AQUATIC DATA ANALYSIS

Federal Aid Project F-239R-15

Harry Vermillion
Information Technology Professional II



Tom Remington, Director

Federal Aid in Fish and Wildlife Restoration

Job Progress Report

Colorado Division of Wildlife

Aquatic Wildlife Research Section

Fort Collins, Colorado

July 2008

STATE OF COLORADO

Bill Ritter, Governor

COLORADO DEPARTMENT OF NATURAL RESOURCES

Sherman Harris, Executive Director

COLORADO DIVISION OF WILDLIFE

Thomas E. Remington, Director

WILDLIFE COMMISSION

Robert Bray, Chair	Brad Coors, Vice Chair
Tim Glenn, Secretary	Dennis G. Buechler
Jeffrey A. Crawford	Dorothea Farris
Roy McAnally	Richard Ray
Robert G. Streeter	Harris Sherman

John Stulp, Department of Agriculture

AQUATIC RESEARCH STAFF

Mark S. Jones, General Professional VI, Aquatic Wildlife Research Leader
Arturo Avalos, Technician III, Research Hatchery
Rosemary Black, Program Assistant I
Stephen Brinkman, General Professional IV, F-243, Water Pollution Studies
Harry Crockett, General Professional IV, Eastern Plains Native Fishes
Matt Kondratieff, General Professional IV, Stream Habitat Restoration
Patrick Martinez, General Professional V, F-242, Coldwater Reservoir Ecology &
GOCO - Westslope Warmwater
R. Barry Nehring, General Professional V, F-237, Stream Fisheries Investigations
Kevin Rogers, General Professional IV, GOCO - Colorado Cutthroat Studies
Phil Schler, Hatchery Technician V, Research Hatchery
George Schisler, General Professional IV, F-394, Salmonid Disease Investigations
Kevin Thompson, General Professional IV, F-427, Whirling Disease Habitat Interactions and
GOCO – Boreal Toad
Harry Vermillion, Scientific Programmer/Analyst, F-239, Aquatic Data Analysis
Nicole Vieira, Physical Scientist III, Water Quality Studies
Paula Nichols, Federal Aid Coordinator
Kay Knudsen, Librarian

Prepared by: _____
Harry Vermillion, Information Technology Professional II

Approved by: _____
Mark S. Jones, Aquatic Wildlife Research Leader

Date: _____

The results of the research investigations contained in this report represent work of the authors and may or may not have been implemented as Division of Wildlife policy by the Director or the Wildlife Commission.

Table of Contents

Signature Page	ii
Study Objectives	1
Job 1. Aquatic Data Management System (ADAMAS).....	2
Database Management and Maintenance	2
The Application	4
Data Requests	5
Job 2. Technical Assistance.....	7
Appendix A: CDOW Aquatic Data Request Form.....	8

State: Colorado

Project No. F-239R

Title: Aquatic Data Analysis

Period Covered: July 1, 2003 to June 30, 2008

Study Objective: To develop analysis of aquatic biological data that accurately describes and/or predicts the status of fish communities and the results of management actions on these communities.

Study Objectives:

Job 1. Aquatic Data Management System (ADAMAS)

Objective: To continue to develop and maintain a computer based, statewide aquatic data management system which will facilitate standardized entry of survey data across the state and access to information regarding all aspects of aquatic data including stream and lake inventories, Scientific Collections (SCICOLL) reports and creel surveys. Active links between ADAMAS and the Aquatic Animal Health (AAHL) database as well as between those two databases and the Division Hatcheries database (TRANS5) will be established and maintained. This job includes aspects of the aquatic portion of the Colorado Vertebrate Ranking System (COVERS).

Job 2. Technical Assistance

Objective: To provide technical assistance to researchers, field biologists, and staff on a variety of aquatic data analysis topics. Topics to include creel survey, inventory survey, management categorization, recording of accurate location data through the use of Global Positioning Systems (GPS), hardware and software review, application development and other computer related data analysis needs.

Job 1. Aquatic Data Management System (ADAMAS)

Database Management and Maintenance

The effort to collect and enter current and historic fisheries survey data stored at various Division offices continues. At the beginning of this reporting period, the database held 13,681 surveys at 9,562 locations, with 356,588 fish sample records, representing 1,909,434 fish. Currently, the database includes 21,571 surveys at 12,376 locations, 902,939 sample records, representing 2,830,297 fish. The following table shows survey entry totals with sampling records and representative fish added for each year in the reporting period.

Reporting Year	Surveys	Sample Records	Fish
pre-2003	13,681	356,588	1,909,434
2003-2004	1,313	27,999	48,073
2004-2005	1,735	147,711	177,646
2005-2006	2,146	174,621	351,194
2006-2007	1,130	44,332	113,202
2007-2008	1,566	151,688	230,672
Total	21,571	902,939	2,830,221

We continue to bring sampling surveys and creel surveys into the system from a variety of sources. Initially, the database was comprised of records from the CDOW Stream and Lake Databank (predecessor to ADAMAS), data files used to store entries for the Creel Survey Analysis Program (C-SAP), a database of historical sampling compiled by Kevin R Bestgen, Ph. D. to support the South Platte and Arkansas Basins' Eastern Plains Natives Fishes reports, CDOW surveys submitted by the biologists and SCICOLL reports. We have designed the database around the most common, basic data items collected in the field and have endeavored to standardize field data reporting formats based on that design.

Currently, data is reported by CDOW biologists and SCICOLL permit holders via an application written by CDOW researcher Kevin Rogers, Ph. D. - the "JakeOmatic" (JOM) - or standardized spreadsheet templates, but occasionally large groups of survey data are located in files, compiled and entered by database staff. As surveys are processed, sampling information is verified and compared to data from previously entered surveys. From time to time, historic survey reports with more detail and individual fish data are found to replace previously recorded summary information. A portion of the 7,000-plus surveys added during this reporting period were the result of a year-long effort to enter survey data from hardcopy files documenting sampling surveys on Colorado's West Slope waters. Hardcopy field data forms were digitally scanned for

future reference and the data were entered through the JOM. We were able to add 1,604 surveys with 35,954 sample records representing 130,167 fish to the database, with data from an additional 59 surveys and 4,461 sample records on hold for spatial referencing.

As the spatial reference is an integral part of the data model, survey location descriptions from the period prior to the Division's involvement in computer-based data storage (pre-1985) are seen to be less accurate, if they exist at all beyond the assigned water code of the reach sampled. The locations described by township, range and section or local landmarks sometimes reveal duplication in reporting as dates and sample data are brought into the comparison. The resulting duplicate survey and sample records are archived and culled from the database.

During this reporting period, several events affecting the ADAMAS database and CDOW aquatic data as a whole have taken place:

In 2005, at an Aquatic Section meeting in Brush, Colorado, the concept of "consolidating" the four, independent, Aquatic-themed databases utilized by the section to a single, centralized database with linkage to the Division's Geographic Information System (GIS) was first presented. The presentation noted that a consolidation would reduce redundancy in common information referenced by each individual system and duplication of effort in maintaining those data. Due to the interrelated and interdependent aspects of some of the processes of the Aquatic Section, it would be possible to introduce automation in some cases; hatchery disease certification, for instance. Consolidation of data would allow a refinement of control of, and access to all related aquatic data by authorized CDOW personnel.

In 2006, at an Aquatic Section Senior Staff meeting, a discussion concerning how aquatic data requests were being handled and some of the problems encountered during hearings due to the distribution of aquatic data from a variety of CDOW sources led to the development of a policy to deal with aquatic data requests centrally with comment by the Aquatic Senior Staff and other CDOW personnel privy to aquatic issues in the state.

In 2007, Governor Ritter came into office and quickly announced a multi-year information technology consolidation plan that folds state government's decentralized operations into the newly created Governor's Office of Information Technology (GOIT) with the mandate "To increase the effectiveness of government through the use of shared information and technology."

In the fall of 2007, after failed attempts to produce an accurate analysis/reporting portion of the ADAMAS application due to widely variable agency requirements and the anticipated loss of key CDOW Aquatic Section personnel to retirements, the decision to consolidate Aquatic Section databases to a single, centralized database was made by the Aquatic Section Senior Staff.

The resulting Aquatics Database design and its applications will meet criteria defined by GOIT and will be implemented with the full participation of the Division's Wildlife Technologies work unit. The database will include stocking and production data from TRANS5, aquatic disease data from the AAHL database, creel data from the database that currently supports C-SAP and the inventory sampling tables from ADAMAS as well as tables common to all four databases with spatial links to the Division's GIS. As of the end of this reporting period, the database has been designed and created on a CDOW SQL Server implementation and the data were migrated from the independent databases to the Aquatic Database on the server.

The Application

Standardization of inventory sampling data recording, entry, analysis and reporting was the primary target of the ADAMAS application. A committee of Aquatic Section field biologists was formed for the design and tracking of the requested, field-user features of the Graphic User Interface (GUI), monitoring standardization of report formats, and reviewing calculations used in the standard analyses available within the program. During application development, the JOM was designated by the Aquatic Section Senior Staff to be used by Division aquatic biologists and researchers as well as interested SCICOLL permit holders to enter and submit the results of field inventory sampling surveys for inclusion in the ADAMAS database. The JOM's data entry feature was designed around the committee's minimum reporting requirements and its analysis features are used as the basis for the ADAMAS application's calculations. SCICOLL users not inclined to use the JOM had the option of entering their data into spreadsheet templates provided by the Division and designed around the format of JOM data files.

In 2003, the Division contracted with Gnomon, Inc. of Carson City, Nevada to provide coding of the application, in order to take advantage of their employees' experience with Division's aquatic data as well as Microsoft's network and database management software packages.

During this reporting period, Gnomon had delivered seven test builds or modifications to the application. Each had been tested and commented on for problems, as requested, by the database staff and each of the biologists on the design committee. As time progressed, test versions of the application achieved success in the entry and edit portion of the program, but the reporting and analysis portion failed to provide the desired analysis results from a standardized set of test data, as well as actual sample survey data entered via the application.

Through the length of this project, the availability and acquisition of better programming and database management environments within the Division and the frustrations of all of the Division's Aquatic biologists and the database manager led to the Aquatic Research Leader's decision to terminate the contract.

As we have described in previous reports, the application's design was set up in two phases. The first phase was to get the application into the field with standardized entry and reporting intact. The second phase, inclusion of an updated, Windows-based version of the Creel Survey Analysis Program (C-SAP).

In the case of the second phase, George Schisler, Ph. D., a CDOW researcher, has utilized Colorado State University students to translate the C-SAP program from the previous DOS-based program into the Windows-based application for use in Dr. Schisler's research studies.

The previous version of C-SAP was written in Turbo Pascal and the data were stored in individual binary files, specific to each creel survey, in a format native to that language. The program had an export feature that would write the data out to dBase III+ files in a relational manner that could be used to store the data and re-import to the program at a future date. In 2001, I hired a temporary employee to run all available C-SAP binary files through the program for the express purpose of compiling all the entered data in a consistent manner, in the dBase structures.

During the translation, it was decided to retain the structure of the dBase export as the basis for data storage in the new version, with expansion as needed. This will allow import of historic creel data for use in the new version and comparison to new surveys on the same waters. Prior to the decision to include the Aquatic Section data consolidation, we migrated the dBase tables to MS Access tables, and the translated version of C-SAP stores data in a relational database, rather than binary files for each survey.

Testing the application has moved forward to use by Division aquatic biologists to enter and analyze actual field data with very good results. A comparison of converted creel survey records in the ADAMAS system's structures and hardcopy results stored at the Fort Collins Wildlife Research Center over the years has revealed a number of creel surveys either not yet in electronic form or not available to the database for consolidation. This has led to a search for any electronic data files held by the Aquatic Research Group, as well as the individual biologists. Files found will be converted to the new format for a complete, consistent set of creel surveys performed since the late 1980's.

Data Requests

Requests for data from the database continue to be filled in a timely manner, formatted as requested with priority given to support Division research and management needs. External government agencies, consultants, contractors and educational researchers are accommodated as expeditiously as possible. Angler requests are referred to Aquatic Area biologists.

This remains a manual process for the most part, although testing of the Gnomon application resulted in a summarization process originally used to check the results of the application's analyses. The resulting summary table has proved valuable as a consistent format for providing requestors with information about sample inventories without having to provide "raw" data to requestors who the Aquatic Data Request Group (described below) have determined not to need that level of resolution in the data provided.

The centralized process for review of requests by the Division's biologists prior to release of data continues to be revised. At this point in time, a formal request is made via email with the CDOW Aquatic Data Request Form (Appendix A). The form is meant to define the requestor's area of interest, the resolution of the data requested and advise the requestor of the provisional status of the data and their responsibilities as to redistribution of the data.

The request, and sometimes the data requested, is distributed to the 17-member Aquatic Data Request Group via Division email for review and comment. The members include the Aquatic Research Leader, the Water Unit Manager, the four Senior (regional) Aquatic Biologists, the four regional Senior Wildlife Species Conservation biologists, the four regional Aquatic or Water Quality Wildlife Species Conservation biologists, the Aquatic Toxicology Researcher, the Aquatic GIS Specialist and the Aquatic Database Manager. The members of this group are aware of aquatic issues statewide and are all in contact with Aquatic Area biologists responsible for the management of waters in the requestor's area of interest. Discussions have taken place among the members via email to determine how the request is to be filled. Once everyone is in agreement or has bowed out of the discussion, the request is filled electronically via email and the request "package", including the data supplied as well as the request form and a copy of the email discussion, is archived for future reference, distribution to other parties involved in the issue (on request) and possible comparison should there be a question of changes to the data.

Overall, the process hasn't reduced the number of requests as expected - 37 formal requests, so far in calendar year 2008, as compared to 42 filled requests in 2007 and 30 formal requests in 2006 (prior to the development of the request process). The process has resulted in an improved method of communication between requestors and the Division, as well as a reduction in concerns for data re-distributed or possibly changed by the requestor. As the request process improves, some of the return requestors are beginning to attach GIS shapefiles defining their project boundaries, which, in turn, allows us to pull the requested data by a simple spatial query, speeding up the process immensely.

Job 2. Technical Assistance

The primary activities on Job 2 during this reporting period were:

- 1) To advise researchers concerning additional components and upgrades to desktop and laptop computers
- 2) Perform service-oriented tasks supporting the researchers' projects such as scanning aerial photography for analyses and photographs for use in presentations to public or professional groups
- 3) To assist researchers with programming needs, as in the development of an Access database used as a virtual, intermediate work area to process water quality data between the instrument that conducts the analyses and records the results and the server-based database used by our parent organization, Department of Natural Resources, to store and serve the results across the internet.

The changes in available data storage devices and management software - moving from a PC-based database backed up to tape systems, CD writers and DVD writers to a server-based Relational Database Management System (RDBMS) on the network that is routinely backed up, mirrored and maintained by the Wildlife Technologies staff - has improved the database' reliability. This combined with less expensive storage costs, has made the concept of archiving scanned images of hard copy reports and photographs more desirable as those documents and images become readily available as referential support to on-going projects.

In the last two years, we have been scanning and cataloging a library of photographic slides made during research efforts over the last 30 years in an effort to reduce storage space, retain the images and make them available for future reference and presentations. This effort has proved valuable to Division researchers and scanned images have been included in request packages from time to time.

In addition, we have begun to copy the Aquatic Research Group's variety of annual Federal Aid Reports, Technical Reports, White Papers, Special Reports and the researcher's individual publications to the Adobe portable data format (pdf) for distribution via the Internet and to reduce printing and shipping costs. This has led to a method of scanning past reports from hardcopy for re-distribution as pdfs, on a request basis.

Since the standardization of operating systems and the basic office suite of programs to Windows 2000/Windows XP operating systems and the XP Office suite, the resulting level of "peer support" continues to develop within the Division and the Aquatic Research Group, redefining the group's technology support needs. We will continue to adapt to the situation, providing what informal support is required.

APPENDIX A

CDOW Aquatic Data Request Form

REQUEST FORM FOR COLORADO DIVISION OF WILDLIFE DATA

1. (a) Name (s) of persons requesting data:

1. (b) Organization/Company/Agency Name (s):

1. (c) Organization/Company Agency Contact Information:

PHONE: _____ FAX: _____ email: _____

(Email address is where electronic data files would be sent)

2. (a) We are requesting data for the following water bodies/geographic area:

(Note that CDOW does not typically distribute point-sample locations or generate GIS maps)

2. (b) Describe the data you are requesting (fish species distributions? Water quality parameters?):

3. Please describe your intended use for this data:

4. You are advised of the following regarding the requested data:

(a) the data may be exempt from the Colorado Open Records Act, in which case, CDOW may deny your request (refer to CORA for exemptions)

(b) the data may be in provisional status (i.e., error check still in progress)

(c) raw data values should not be changed. If you have original or copies of data sheets or previous exports with differences in the data you receive, please call or email for possible corrections.

(d) Do not redistribute this data to parties not listed above. Other parties must submit a formal request to CDOW to insure that they receive the most updated version of the data available.

Name of CDOW Contact: Harry Vermillion
EMAIL: harry.vermillion@state.co.us
PHONE: 970-472-4314
FAX: 970-472-4457
Date data sent to email address listed in 1 (c) :