Exxon Valdez Oil Spill
State/Federal Natural Resource Damage Assessment Final Report

Database Management

Fish/Shellfish Study Number 30
Final Report

Carmine DiCostanzo
Bruce P. Simonson

Alaska Department of Fish and Game
Division of Commercial Fisheries
Computer Services Section
1255 W. 8th Street
Juneau, Alaska 99802-5526

June 1993

Exxon Valdez Oil Spill
State/Federal Natural Resource Damage Assessment Final Report

Database Management

Fish/Shellfish Study Number 30
Final Report

Carmine DiCostanzo
Bruce P. Simonson

Alaska Department of Fish and Game
Division of Commercial Fisheries
Computer Services Section
1255 W. 8th Street
Juneau, Alaska 99802-5526

June 1993

Database Management

Fish/Shellfish Study Number 30
Final Report

Study History: Fish/Shellfish Study Number 30 was initiated as
part of a 1992 detailed study plan (Database Management) in
recognition of the need to catalog and maintain electronic copies
of raw data generated by natural resource damage assessment
projects.

Abstract: In order to evaluate the extent of damage and injury
to natural resources as a result of the *Exxon Valdez* oil spill,
the following types of data would need to be synthesized and
carefully managed: 1) historical baseline data to establish
conditions prior to the spill; 2) data generated by damage
assessment projects to establish damage and extent of injury; and
3) data generated by ongoing monitoring projects to determine the
effectiveness of restoration programs. This project was
implemented to facilitate access to historical datasets
maintained by the Alaska Department of Fish and Game, and to
address the need to preserve the raw data generated by selected
damage assessment projects associated with the *Exxon Valdez* oil
spill. Procedures for developing data schemata and data
dictionaries were implemented to support the possibly of
integration of independently generated databases into geographic
information systems (GIS) and other consolidated databases. As
this project closed out, final project data with spatial and
temporal coverage descriptions were solicited from principal
investigators to facilitate future integration of databases, and
to allow for ecosystem views of affected habitat and injured
resources.

Key Words: Damage assessment, database management, databases,
Exxon Valdez, geographic information systems, historical baseline
data, raw data.

Citation:
DiCostanzo, C., and B.P. Simonson. 1993. Database management,
    Exxon Valdez Oil Spill State/Federal Natural Resources
    Damage Assessment Final Report (Fish/Shellfish Study Number
    30), Alaska Department of Fish and Game, Division of
    Commercial Fisheries, Juneau, Alaska.

Table of Contents

List of Figures

List of Appendices

Executive Summary

Data play a fundamental role in determining the extent of damage and injury to natural resources as the result of human error. In the case of the oil spill caused by the *T/V Exxon Valdez* in March of 1989, it was recognized that the following types of data would need to be synthesized and carefully managed:

historical baseline data –
    to establish conditions prior to the spill

data generated by damage assessment projects –
    to establish damage and extent of injury

data generated by ongoing monitoring projects –
    to determine the effectiveness of restoration programs

This project was implemented to facilitate access to historical datasets maintained by the Alaska Department of Fish and Game, and to address the need to preserve the raw data generated by selected damage assessment projects associated with the *Exxon Valdez* oil spill.

Although the integration of independently generated databases into geographic information systems (GIS) and other consolidated databases was not part of this project's objectives, procedures for developing data schemata and data dictionaries were implemented to support this possibility. As this project closed out, final project data with spatial and temporal coverage descriptions were solicited from principal investigators to facilitate future integration of databases, and to allow for ecosystem views of affected habitat and injured resources.

Introduction

The Alaska Department of Fish and Game (ADF&G) maintains several extensive historical and active databases to support its ongoing research and management activities. These datasets are an important component of the baseline data required in the damage assessment effort related to the *Exxon Valdez* oil spill.

Because of the importance of new data generated by damage assessment projects, special measures were taken to ensure their protection. Existing computer and telecommunications facilities within the ADF&G provided the backbone for these efforts, with additional resources applied as needed by oil spill projects.

In addition, this project provided technical assistance to principal investigators of damage assessment projects in the form of computer networking, data archival, database definition and documentation, report preparation, and software support.

1

Objectives

This database management project was implemented with the following objectives:

1) to facilitate access to selected historical and active datasets maintained by the Alaska Department of Fish and Game (ADF&G) in a variety of electronic formats,

2) to ensure that new data generated by selected damage assessment projects would be protected from accidental loss and would be suitably archived, and to help ensure the longevity of these data after projects close-out,

3) to coordinate cost effective access to ADF&G's existing computer and telecommunications facilities for oil spill related projects and activities, and

4) to provide technical assistance to principal investigators where possible and appropriate.

Methods

Database Management - Historical Datasets

Population levels of injured species in the area affected by the oil spill form an important component of the baseline information required for damage assessment. Several databases routinely maintained by ADF&G that relate to population levels were identified, including the department's historical harvest database (multiple species), and species specific databases of historical escapement and age/weight/length data.

The importance of the department's historical harvest database was immediately recognized. This database is maintained by the division of commercial fisheries, and contains over twenty years of data on some 200 species of fish and shellfish. Because harvest data play a key role in estimating population levels of many species in the area of the oil spill, many damage assessment projects required access to this database. In addition, it was anticipated that there would be an increased demand by the public and other agencies for information contained in this database.

This project developed a system to address these new demands on this historical data according to the following priorities: (1) completeness and accuracy of the contents of the database, (2) speed and ease of retrieval, and (3) support for a variety of software formats.

2

Other historical datasets maintained by ADF&G were used on a per-project basis. Access to these datasets was facilitated through technical support provided by this project.

Database Management - Damage Assessment Datasets

This database management project was initiated during the second year after the oil spill. Damage assessment projects were already in place at the time, and principal investigators had assumed responsibility for developing and maintaining datasets for their own projects. This approach was not replaced when this database project was created; it was recognized that the design and care of databases is best managed by those responsible for the acquisition, synthesis, and interpretation of data.

Nevertheless, proper archival and detailed documentation of datasets is a technical issue that may impede the stated purpose of specific damage assessment work. Furthermore, when separately conceived and executed projects must be archived with an eye for future use and integration, it is cost-effective to dedicate a single project to coordinate this task.

For these reasons, personnel experienced with computer networks, data backup and archival, database documentation, and other technical issues were made available to the principal investigators of damage assessment projects. Guidelines for documenting datasets, along with examples and associated software tools, were distributed to project leaders and key staff.

Technical Support

In addition to database work, this project provided computer technical support for selected oil spill projects in ADF&G. This support was provided at three levels: (1) computer wide area network (WAN) interconnection between project administration and project staff using existing telecommunications facilities within the ADF&G, (2) computer local area network (LAN) support at regional and field offices where oil spill projects were undertaken, and (3) detailed technical support to principal investigators of oil spill projects on specific issues.

Support on ADF&G CF's WAN was implemented to ensure efficient and cost-effective transport of data and project documentation between project personnel and oil spill administration. At the LAN level, additional disk capacity was allocated on ADF&G computer networks to store new project data. Network wiring dovetailed with existing token-ring and ethernet backbones on ADF&G CF's LANs. Additional backup and archival procedures were implemented on existing file servers, with the express intent of reducing the risk of accidental loss of assessment data as it was acquired. Support for principal investigators included development of database management

3

procedures, data export facilities, base level GIS, assistance
with document production and preparation of presentation
materials, and support for hardware and other technical issues.

Results

Databases

The primary work products of this project relate to the
databases used or created by selected ADF&G damage assessment
projects. The extent and availability of documentation of these
datasets is a function of the involvement of principal
investigators in the data management process. In some cases, the
detail of documentation which is available includes complete
database schema and data dictionaries, but more typically this
documentation is at the level of general database descriptions
and lists of electronic files.

Historical Datasets

A complete work-up of the harvest database was conducted,
including a full schema and detailed data dictionary, updated
internal code lists, and procedures for data maintenance and
version control. The database was ported to a UNIX platform
(funded by the ADF&G independent of the oil spill project).
Software was written to provide access to the database over the
ADF&G commercial fisheries WAN; backend procedures were
implemented to provide reports and electronic data extracts in
numerous database, spreadsheet, and statistical formats.
Requests for information are automatically logged in a relational
SQL database for resubmittal with or without modification of
report parameters.

During the development of this system, access to historical
harvest data was provided by project personnel on a case-by-case
basis. At the time of this writing, the automated system is in
the last phases of beta-testing, with final implementation and
documentation due in October of 1993. The system will provide
unattended custom reporting and data extracts in a wide variety
of formats on the ADF&G commercial fisheries WAN.

A comprehensive review of the contents and correctness of
data contained in this harvest database was implemented in the
initial phases of this project. A report detailing the changes
made to the historic data was published, along with periodic
reports relating to subsequent version control.

Daily harvest records from 1969 through the present are
represented in this database. Locations are coded by ADF&G
statistical area, the primary means of defining GIS parameters
used by management and research at the ADF&G. The main fields of
this database that are useful to assessment work include species,

catch in numbers and pounds, gear type, harvest type, and data on specific permits used to land the catch. Distribution of information in this database is constrained by Alaska law (in order to protect the fishing interests of private individuals); the system has been designed to take this into account.

Output can be provided in a wide variety of electronic formats suitable for databases, spreadsheets, and statistical packages. A large number of printed reports are also supported by the system.

A portion of the database documentation for this historical database has been extracted and included in Appendix A. This is representative of the documentation available for this database.

Damage Assessment Datasets

Individual damage assessment projects on the ADF&G CF WAN were allocated disk space on file servers in local ADF&G offices. These file servers have regularly scheduled archive and backup procedures, which were used to preserve copies of historical databases and new data collected by damage assessment projects.

Instructions for the detailed documentation of databases were supplied to principal investigators, along with the fully completed documentation of the historical harvest database. Recommendations for version control of evolving datasets were made, as well as proposed procedures for the orderly release of project data as they came available.

As this project and other damage assessment projects were in the final phases of completion, requests for electronic copies of project data increased. This project was asked to facilitate consolidation of project data and databases for selected ADF&G assessment work. A publication is in preparation which will summarize the results of this effort.

Location Specific Activities

Primary support for this project was coordinated out of the computer services section of the division of commercial fisheries of the ADF&G, headquartered in Juneau. Through the ADF&G CF WAN, network resources and databases were monitored and managed from the Juneau office.

Multiple offices of the ADF&G required the support of this project to ensure that the data associated with damage assessment projects were secure and safe from accidental loss. In general, the support provided took advantage of systems already in place; these systems were enhanced where needed by damage assessment projects.

The following summarizes the major activities which were required in local offices of the ADF&G, and indicates the level of support needed to coordinate damage assessment work of multiple projects, and to ensure that critical datasets are protected from accidental loss.

Juneau

Work on the historic harvest database was done in Juneau, which required the installation of additional hardware and software architectures, partially funded out of this oil spill project's budget. Additional disk capacity was installed, an SQL database of job requests was designed and installed, a UNIX workstation was commissioned to handle reporting and data extracts, and telecommunications facilities were established which were used to monitor WAN availability for oil spill projects.

Network support was coordinated out of this office; documentation and guidelines detailing the use of the ADF&G CF WAN were provided for oil spill projects by personnel from the Juneau office. In addition, support for the preparation of reports and presentations was headquartered in this office. Along these lines, the most important items were the installation of a simple baseline GIS package, and procedures for the cost effective production of 35 mm slides needed for formal presentations.

The Juneau office also assisted with coordinating efforts of the NOAA Auke Bay Lab (ABL) and the principal investigators of oil spill projects. ABL was provided with a set of the database documentation tools developed by this project. Preliminary steps at integrating the ABL hydrocarbon database with ADF&G historical and damage assessment datasets were completed.

Cordova

The Cordova office required substantial support. While a direct telecommunications connection to this office had been planned by the ADF&G for some time, the oil spill made this data link critical. Through the state agency responsible for the management of telecommunications, a 19.2 KB data path was installed that connected the Cordova office with the rest of the ADF&G CF's WAN.

The Cordova LAN file server was upgraded to Novell Netware 3.11, and additional disk capacity was made available for storing oil spill datasets.

An additional complication at the Cordova office arose through the temporary lease of a building dedicated to oil

spill projects, which was not easily connected to the WAN. This building, with its significant datasets, was connected to the WAN through the installation of a spread spectrum ethernet bridge. With the completion of a thinwire ethernet backbone, the datasets in this building were incorporated into the regular archive and backup cycles of the Cordova office LAN.

Anchorage

The Anchorage office of the ADF&G served as a focal point for the interaction of principal investigators and managers of oil spill activities. Most of the correspondence relating to damage assessment projects, interim status reports, final reports, proposals for ongoing work, and budget management passed through this office. This project provided direct support for this communication, through the use of electronic mail on the WAN, and additional storage capacity for the electronic documents and datasets involved.

In addition, ADF&G upgraded the file server in Anchorage to handle the increased load induced by oil spill activities. Electronic storage was allocated in a data repository, which was made available to projects upon close-out. Additional archive and data backup procedures were implemented to protect the data stored in this repository.

Kodiak

The technical support provided by this project to the ADF&G Kodiak office included upgrades to electronic mail services, and an extensive reworking of network design and software by ADF&G's commercial fisheries computer services staff. The local computer staff in Kodiak assumed responsibility for data archival and backup procedures.

Kenai Peninsula

The support for offices of the ADF&G located on the Kenai peninsula is representative of the difficulty inherent in managing major damage assessment work in remote areas of Alaska. Like many of the isolated offices of the ADF&G, these sites are not connected to the ADF&G CF WAN, yet there is need to archive and backup data, facilitate communication between researchers, staff, biometric support, and project management, to assist with version control of evolving datasets, and provide technical support for computer hardware and software.

The oil spill exacerbated this problem in the Kenai peninsula offices; in particular, while the additional demands on local resources were successfully met, it was not

without exceptional effort by the principal investigators involved with damage assessment work.

This project worked with ADF&G staff and computer services to implement access to the electronic mail facility on the WAN. This important data link provided communication between project personnel and oil spill management in Anchorage and Juneau.

The raw data acquired by the projects out of these offices were somewhat atypical for damage assessment work; in particular, large volumes of data were obtained and analyzed using specialized equipment. Because the interpretation of these data are dependent on dedicated equipment, data for these projects were kept on-site. The local ADF&G offices assumed responsibility for archive and backup of these data.

Discussion

Data play a vital role in determining damages and the extent of injury to natural resources. In addition, demands on data are often made that exceed their original scope and purpose. For example, within the scientific arena, integration of separately acquired data is the key to rational views of ecosystems and realistic definition of natural habitat; in the legal and political arena, data can be used to justify litigation claims and to drive management policies.

The *Exxon Valdez* oil spill resulted in an unprecedented program of data acquisition and scientific effort to determine injury to natural resources. Millions of dollars were spent acquiring data by damage assessment projects, which, for the most part, were separately conceived and executed.

It is inherent in disasters of this sort that there is no lead time for creating a centralized effort with standards for data formats, frequency and extent of sampling, and systems for analysis and synthesis. The best that can be done is to act quickly, with the best talent available, to establish good projects with demonstrable utility. The approach might be different, if, for example, oil spills and other disasters were scheduled events, or alternatively, if it were assumed that there is a non-zero probability that oil spills and other disasters will occur in certain areas in specified time frames.

This project made clear the efficacy of managing project data in two separate modes: the "work in progress" mode, where principal investigators are actively building and analyzing new datasets, and the "close-out" mode, where the data associated with a project are archived and suitably documented to ensure longevity after the project ends, and the possibility for use in new contexts.

8

Support for Work in Progress

During the "work in progress" mode, principal investigators must be given support as they build data structures and systems for data entry, version control, archival, and analysis. It is clearly desirable to minimize the impact of such technical issues on an investigator's time. In Alaska, where an adequate telecommunications network is available for monitoring and managing electronic copies of databases (within the ADF&G), it appears the most cost effective means of accomplishing this is through centralized support.

It is essential that principal investigators maintain complete control and ownership of data while a project is underway. This ensures the flexibility needed for data development, new views of project data, ease of analysis, and ultimately, the best science possible.

It is also essential that simple and effective tools for proofing data points, data analysis, and data archival be made available to principal investigators "at their desks". Training must be provided in the use of these tools, so that investigators can be active participants in procedures designed to provide integrated views of data from multiple projects.

Specific areas for which clear needs were identified during the scope of this project include a simple PC-based GIS, methods for version control of evolving datasets, and suitable software for accessing and analyzing large databases. Investigators also were in need of technical assistance with hardware and software installation, and with the preparation of presentation materials and final reports. In some cases, it makes sense to centralize at least part of these functions, to realize economies of scale.

Support for Projects upon Completion

As projects close out, it is important to implement specific data management procedures that document datasets for possible use in other contexts. This is often an area where investigators have little experience, as it depends on principles of data abstraction and other esoteric concepts about internally consistent views of data entities and their attributes.

Nevertheless, properly conducted data documentation is heavily dependent on the investigator who acquired the data, as datasets very typically include internal codes, assumed units of measurement, methods for dealing with missing values and outliers, and other project specific idiosyncracies.

It is most reasonable to develop this documentation as projects close-out, since this is when conclusions based on data are finalized, and the content of a database is essentially static. However, it is not unusual for a database to continue to evolve after final reports are published, especially if other programs assume responsibility for future use and updates to the data. In such cases, it is important that procedures for version control and mechanisms for providing "snapshots" of databases be carefully implemented.

An overview of detailed database documentation is included in Appendix A, with examples extracted from various databases tracked by this project.

Conclusions and Recommendations

Baseline and assessment data are the key to determining damages and the extent of injury to resources. Similarly, data acquired through monitoring projects provide the means to prove the effectiveness of restoration programs.

Data management is not a trivial exercise, particularly if one takes seriously the need to ensure the longevity of data beyond the scope of a particular project. Though it is difficult, a balance must be struck between burdening an investigator with extensive data management procedures, and ensuring that a project's data can have utility in broader contexts after the project closes out.

Specific suggestions for improved data management are found in the discussion section of this paper. The main points are summarized here.

- Do recognize the importance of baseline data, newly acquired assessment data, and data obtained by future monitoring work in the restoration program.

- Do insist that investigators realize that the data which they collect may have utility beyond the scope of the project that originated it.

- Do take advantage of existing computer networks and data management facilities where possible.

- Do establish centralized standard basemaps and other global resources; do provide minimal guidelines for data documentation and data management; and do provide principal investigators with the support and tools to meet and utilize these centralized standards and guidelines.

- Do recognize and support the two arenas in which data must be managed: the period of time when a project is in progress and its databases are under development, and when a project closes out, and its databases must be archived and documented for possible future applications.

- Do not attempt to centralize data before its time; specifically, do not solicit and centralize data while a project is in progress.

- Do not attempt to centralize data, particularly to feed centralized projects, without providing immediate utility to the investigator who acquired the data.

- Do not expect investigators to manage their own data without providing investigators and their staff with suitable technical, hardware, and software support.

- Do not allow projects with significant datasets to close out before ensuring that adequate documentation and archival of data has taken place.

In conclusion, it is cost-effective to spend the small amount necessary to ensure that data acquired through millions of dollars worth of effort will have longevity and can be used in broader contexts. The data obtained today will very likely be the baseline data of tomorrow.

References

Cressie, N., *Statistics for Spatial Analysis*, John Wiley & Sons, New York, (1991).

Dukta, A. F. and Hanson, H. H., *Fundamentals of Data Normalization*, Addison-Wesley, Reading, MA (1989).

Fleming, C. C., and von Halle, B., *Handbook of Relational Database Design*, Addison-Wesley, Reading, MA (1989).

Kusler, J. A., and Kentula, Mary E., editors, *Wetland Creation and Restoration, the status of the science*, Island Press, Washington, D. C., (1990).

Simonson, B. P., *Exxon Valdez Oil Spill and related ADF&G Historical Databases: Data Schemata, Dictionaries, and GIS Coverages*, ADF&G publication (in preparation).

APPENDIX A - Database Documentation: Samples and Extracts

A comprehensive set of database documentation should, wherever possible, include each of the following items. Samples from documentation created by this project for the historical harvest database and selected assessment projects are attached in this appendix. Complete sets of this documentation are in preparation for publication in a separate report.

A single page description of the database, providing an overview of its contents (spatial and temporal extents) and data format. *(Figure 1)*

A visual representation of the database's content, indicating the relationship of keys, records, and fields. *(Figure 2)*

Comprehensive lists of records and fields, detailed descriptions of record layout and field attributes, and a complete set of internal codes used in the database. *(Figures 3, 4, 5, 6, and 7)*

At a minimum, a simple location map indicating areas for which data is available. *(Figure 8)*

Miscellaneous

Information on how data was sampled, including copies of data forms and procedures for acquisition and entry of data, information on update cycles and anticipated continued use of the database, and known written reports known to have made use of the data.

Figure 1.   Database Cover Sheet · (Sample)

This document provides a general description of the database,
including size, number of records, primary fields, format,
spatial and temporal extents, individuals responsible for
access and content, related databases, update/modification
schedules, and an indication of the extent to which internal
codes are used in the database.
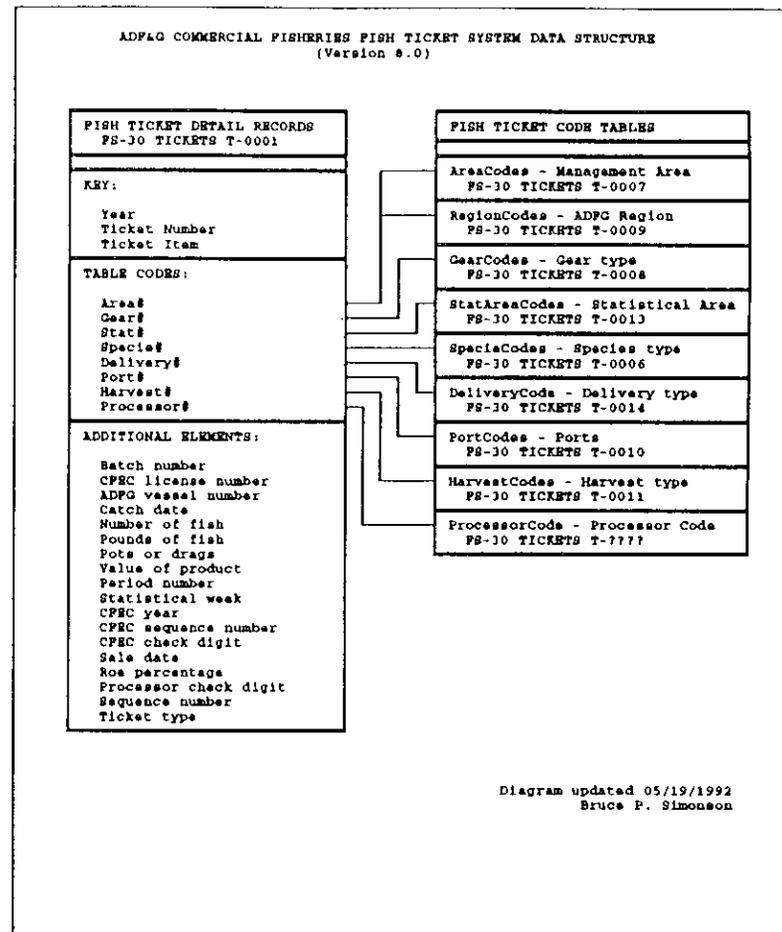
Figure 2.   Database Schema · (Sample)

This document provides a visual representation of the content
of the database, indicating the relationship between the tables
and columns in the database, with the primary keys and
cardinality associated with table linkages.

A.2

Figure 3. Comprehensive Record (Table) List · (Sample)

This document is a listing of all of the tables contained in the database, and includes references to last updates to the tables and the number of records found in each table.

Figure 4. Comprehensive Field (Column) List · (Sample)

This document lists all of the fields found in the database, along with fundamental attributes (internal storage format) and last update status.

A.3

NRDA/R PROJECT: PS-30 : Database Management

DATABASE NAME: ADFG CF Fish Ticket database; skeleton RBASE data structure

----

FILE NAME: TICKETS

TABLE / RECORD ID: Tickets (PS-30 TICKETS T-0001)

BRIEF DESCRIPTION: ADFG CF Fish Tickets

DESCRIPTION: This table contains the detail records of fish ticket detail
                       items. The unique key for a fish ticket item record is defined
                       as the combined Year, Ticket Number, and Ticket Item Number.

              AS OF: 04/22/1992 15:43:10 NUMBER OF RECORDS: 10982121

TABLE CREATED IN DD: 04/16/1992 13:21:47 READ/MODIFY PSWDS: Y / Y

  LAST ACTION IN DD: 04/22/1992 15:43:19 Update

ELEM	ELEMENT NAME AND ID#	ELEMENT DESCRIPTION	TYPE		IDX
1	Area# PS-30 TICKETS E-0002	Management Area code	TEXT	1	*
2	Batch PS-30 TICKETS E-0004	Batch number of ticket	TEXT	3	
3	Year PS-30 TICKETS E-0061	Year of ticket	TEXT	2	*
4	Ticket# PS-30 TICKETS E-0057	Ticket number	TEXT	6	*
5	TicketItem PS-30 TICKETS E-0058	Item number of record on ticket	TEXT	3	
6	CFEC PS-30 TICKETS E-0006	CFEC license number	TEXT	10	
7	ADFG PS-30 TICKETS E-0001	ADFG vessel number	TEXT	5	
8	Processor# PS-30 TICKETS E-0029	Processor code	TEXT	5	
9	Gear# PS-30 TICKETS E-0016	Fish Ticket Gear code	TEXT	2	*
10	CatchDate PS-30 TICKETS E-0005	Catch date	TEXT	4	
11	Stat# PS-30 TICKETS E-0038	AK CF Statistical Area	TEXT	6	
12	Filler03 PS-30 TICKETS E-0013	Filler (no data)	TEXT	3	
13	Special# PS-30 TICKETS E-0035	Fish Ticket Species Code	TEXT	3	*
14	Number PS-30 TICKETS E-0022	Number of fish/shellfish delivered	NUM	6	0
15	Pounds PS-30 TICKETS E-0028	Pounds of fish/shellfish delivered	NUM	7	0

ELEM	ELEMENT NAME AND ID#	ELEMENT DESCRIPTION	TYPE	IDX
16	Pots PS-30 TICKETS E-0027	Number of pots/drags used/made	NUM	4 0
17	Value PS-30 TICKETS E-0060	Value of product (whole dollars)	TEXT	7
18	ErrorFlags PS-30 TICKETS E-0012	Internal system error flags	TEXT	10
19	Period# PS-30 TICKETS E-0023	Period code	TEXT	2
20	StatWeek PS-30 TICKETS E-0039	Statistical week	TEXT	2
21	CFECYear PS-30 TICKETS E-0008	CFEC Year	TEXT	2
22	CFECSequenceNumber PS-30 TICKETS E-0007	CFEC sequence number	TEXT	2
23	CheckCFEC PS-30 TICKETS E-0009	CFEC check digit	TEXT	1
24	SaleDate PS-30 TICKETS E-0033	Sale date	TEXT	4
25	Delivery# PS-30 TICKETS E-0011	Delivery code	TEXT	2
26	RoePercent PS-30 TICKETS E-0032	Roe percentage	TEXT	3
27	Port# PS-30 TICKETS E-0024	Port code designating delivery site	TEXT	3
28	CheckProcessor PS-30 TICKETS E-0010	Processor check digit	TEXT	1
29	SeqNumber PS-30 TICKETS E-0034	Sequence number	TEXT	2
30	Harvest# PS-30 TICKETS E-0018	Fish Ticket Harvest code	TEXT	2
31	TicketType PS-30 TICKETS E-0059	Ticket type	TEXT	1

PS-30 DATABASE MANAGEMENT - NRDA/R DATA DICTIONARY TABLE -   PS-30 TICKETS T-0001

Figure 5. Detailed Record (Table) Layout · (Sample)

This document provides a detailed view of a record in the
database. There should be one of these documents for each
table in the database. In addition to showing the constituent
fields in a record, this document indicates record counts and
information on update / modification activity.

NRDA/R PROJECT: FS-30 : Database Management

DATABASE NAME: ADFG CF Fish Ticket database; skeleton RBASE data structure

---

FILE NAME: TICKETS

COLUMN/ELEMENT ID: Fishery# (FS-30 TICKETS E-0014)

BRIEF DESCRIPTION: Fisheries type

DESCRIPTION: This code provides a super-grouping of species (combining, for
                       example, all the codes corresponding to salmon). This field is
                       used during data entry as part of the data validation process,
                       and can be used during data reporting to select particular sets
                       of species.

       LAST EXTRACT: 04/22/1992 15:43:10

COLUMN ADDED TO DD: 04/16/1992 12:21:47 LAST ACTION IN DD: 05/22/1992 15:22:31 Update

       TYPE / SIZE: TEXT   1   UNITS: none

          FORMAT: 9   ALIGN: L   FILL: none

                                NULLS: not allowed

       DATA VALIDATION: The value in this field must exist in the Fishery Code table.

       DATA SOURCE: Fish ticket

TABLE NAME AND ID#	TABLE DESCRIPTION	X IDX
FisheryCodes FS-30 TICKETS T-0005	ADFG CF Fish Ticket Fishery Codes	U •
SpecieCodes FS-30 TICKETS T-0006	ADFG CF Species codes	U

FS-30 DATABASE MANAGEMENT - NRDA/R DATA ELEMENT -   FS-30 TICKETS E-0014

Figure 6.   Detailed Field (Column) Description - (Sample)

This document provides a detailed description of a field
(column) in the database. There should be one of these
documents for each field in the database. This document
provides detailed information about internal storage of this
data element, units assigned, null values, validation rules,
and information on the update / modification cycle.

COMMON NAME -- SPECIE# -- SCIENTIFIC NAME

ABALONE	860 HALIOTOS KAMTSCHATKANA
ARCTIC CHAR (ANADROMOUS)	521 SALVELINUS ALPINUS
ARCTIC CHAR (GENERAL)	520 SALVELINUS ALPINUS
ARCTIC CHAR (NONANADROMOUS)	522 SALVELINUS ALPINUS
BLACKFISH	630 DALLIA PECTORALIS
BLUE MUSSEL	855 MYTILUS EDULIS
BURBOT	590 LOTA LOTA
CLAM, BUTTER	810 SAXIDOMUS GIGANTEUS
CLAM, COCKLE	820 CLINOCARDIUM NUTTALLI
CLAM, EASTERN SOFTSHELL	842
CLAM, GENERAL (DEADLOSS)	849
CLAM, GEODUCK	815 PAROPE GENEROSA
CLAM, LITTLE-NECK	840 PROTOTHACA STAMINEA
CLAM, RAZOR	830 SILIQUA PATULA
CLAM, SURF	812 SPISULA SOLIDISSIMA
COD, LONGFIN	260 ANTIMORA ROSTRATA
COD, PACIFIC (GRAY)	110 GADUS MACROCEPHALUS
CORAL	899 GENUS CORALLIUM
CRAB, BLUE KING	922 PARALITHODES PLATYPUS
CRAB, BLUE KING (DEADLOSS)	927 PARALITHODES PLATYPUS
CRAB, BOX	900 LOPHOLITHODES MANDTII & LOPHOLITHODES FORAMINATUS
CRAB, BOX (DEADLOSS)	909
CRAB, BROWN KING	923 LITHODES AEQUISPINA
CRAB, BROWN KING (DEADLOSS)	928 LITHODES AEQUISPINA
CRAB, DUNGENESS	910 CANCER MAGISTER
CRAB, DUNGENESS (DEADLOSS)	919 CANCER MAGISTER
CRAB, KING (GENERAL)	920 GENUS LITHODES AND GENUS PARALITHOD
CRAB, KING (GENERAL, DEADLOSS)	929 GENUS LITHODES AND GENUS PARALITHON
CRAB, KOREAN HORSEHAIR	940 ERIMACRUS ISENBECKII
CRAB, KOREAN HORSEHAIR (DEADLOSS)	949 ERIMACRUS ISENBECKII
CRAB, RED KING	921 PARALITHODES CAMTSCHATICA
CRAB, RED KING (DEADLOSS)	926 PARALITHODES CAMTSCHATICA
CRAB, SCARLET KING	924 LITHODES COUESI
CRAB, SCARLET KING (DEADLOSS)	925 LITHODES COUESI
CRAB, TANNER (BAIRDI)	931 CHIONOECTES BAIRDI
CRAB, TANNER (BAIRDI, DEADLOSS)	937 CHIONOECTES BAIRDI
CRAB, TANNER (GENERAL)	930 GENUS CHIONOECTES
CRAB, TANNER (GENERAL, DEADLOSS)	939 GENUS CHIONOECTES
CRAB, TANNER (OPILIO)	932 CHIONOECTES OPILIO
CRAB, TANNER (OPILIO, DEADLOSS)	938 CHIONOECTES OPILIO
DEEPWATER FLATFISH	118
DOLLY VARDEN (ANADROMOUS)	531 SALVELINUS MALMA
DOLLY VARDEN (GENERAL)	530 SALVELINUS MALMA
DOLLY VARDEN (NONANADROMOUS)	532 SALVELINUS MALMA
EELS OR EEL-LIKE FISH	210 MULTIPLE FAMILIES
FLOUNDER, ALASKA PLAICE	133 PLAICEPLEURONECTES QUADRITUBERCULATUS
FLOUNDER, ARROWTOOTH	121 ATHERESTHES STOMIAS
FLOUNDER, GENERAL	120 FAMILY PLEURONECTIDAE
FLOUNDER, STARRY	129 PLATICHTHYS STELLATUS
GRAYLING	610 THYMALLUS ARCTICUS
GREENLING, ATKA MACKEREL	193 PLEUROGRAMMUS MONOPYRTYGIUS
GREENLING, GENERAL	190 GENUS HEXAGRAMMOS
GREENLING, KELP	194 HEXAGRAMMUS DECAGRAMMUS

Figure 7.   Database Code Definition - (Sample)

This document provides a comprehensive definition of codes
contained in the database, and the rules for interpretting
these internal codes. A document of these codes should be
provided for each code table in the database, and mechanisms
should be provided for obtaining cross-references for these
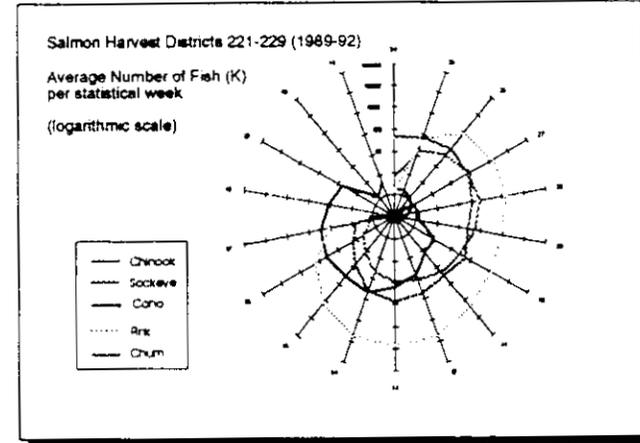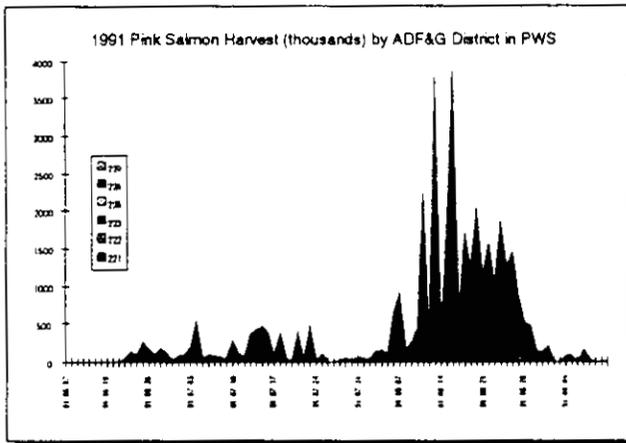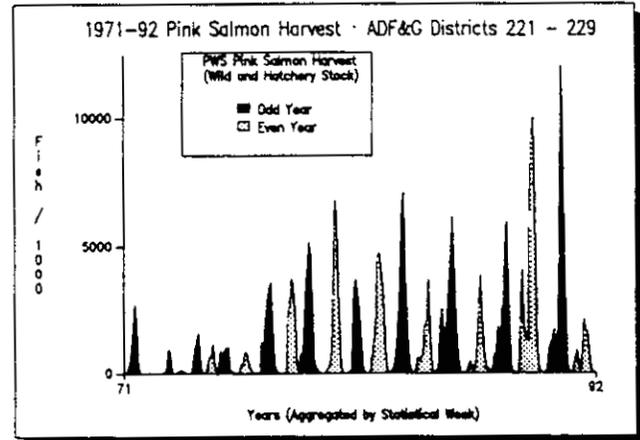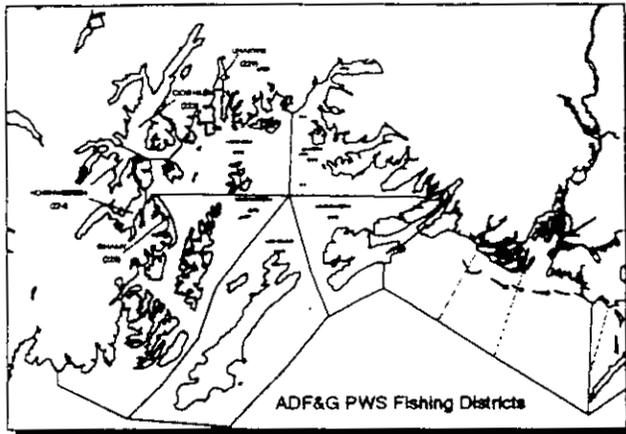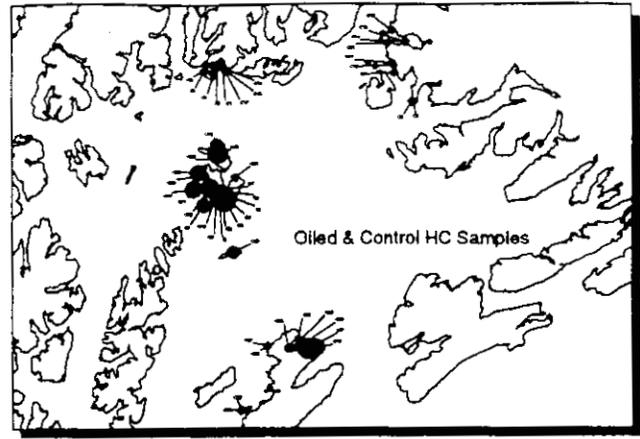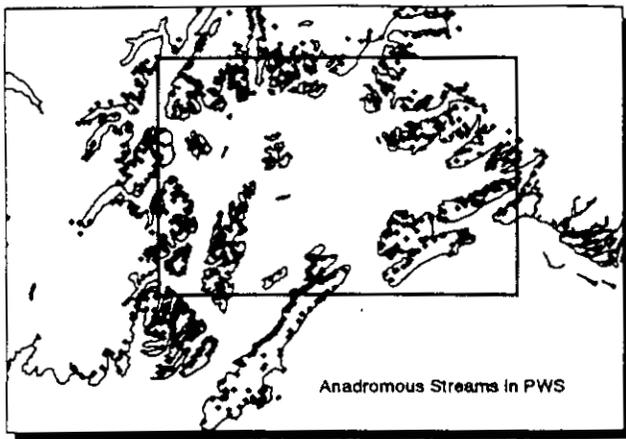codes (by code and/or by interpretation).

A.5

Figure 8. GIS Coverages and Associated Graphics · (Samples)

From left to right, top to bottom: (a) a plot of anadromous streams in PWS, taken directly from an electronic spreadsheet; (b) the same plot, zoomed in, with scaled dots indicating the hydrocarbon content in oiled and control samples (plotted directly from a PC database); (c) ADF&G districts in PWS, serving as a basemap; (d) historical harvest data, spanning multiple years, tied to ADF&G fishing districts in PWS; (e) a season's worth of historical data, by district; and (f) a harvest timing rose, showing average catch by statistical week in PWS.

These graphics are representative of mapping coverages (GIS) and software tools which should be available at the desktops of investigators as they acquire and proof data. It is important that the tools selected be affordable, and easy to use (which is not always the case with GIS technology). Investigators should be provided with standardized basemaps, so that data points in separate coverages can be proofed with the possibility for future integration.

APPENDIX B - ADF&G Historical Harvest Database

The ADF&G Division of Commercial Fisheries maintains a centralized database of historical harvest records for some 200 species of fish and shellfish. The twenty years of information in this database form a component of many ADF&G management and research activities, and are also important to many of the studies conducted as part of the *Exxon Valdez* oil spill.

Not all databases should be centralized; however, this harvest database and other historical and baseline databases are typical of the information for which centralized systems are appropriate. If a centralized database of information is to be of any real value, a number of basic issues must be addressed.

(1) Procedures must be implemented which ensure timely and accurate updates to centralized data.

(2) The data must be accessible to end-users, particularly those who are responsible for the acquisition and proofing of raw data.

(3) The database must be flexible, providing access to data in a variety of electronic and printed formats.

(4) Access to the data must be as transparent and efficient as possible; preferably access is automated and requires no intervention by technical support staff.

In the case of the harvest information, procedures for maintaining the centralized database have been in place for many years. Because of the increased demand for this data, the additional items in the above list received special attention and were enhanced as part of the database management project.

Statewide access to harvest data is possible over the commercial fisheries computer network that extends to all regional offices of the ADF&G. *(Figure 1)*

A special user interface allows access to harvest data across each of its dozen dimensions; output is provided in a variety of electronic and printed forms. *(Figure 2)*

Automated, unattended access to reports and data extracts of harvest data is available on the ADF&G Commercial Fisheries Computer Wide Area Network. *(Figure 3)*

ADF&G Historical Fishery Harvest Database

Alaska Department of Fish and Game Commercial Fisheries Wide Area Computer Network

56 KB dedicated line

19.2 KB dedicated line
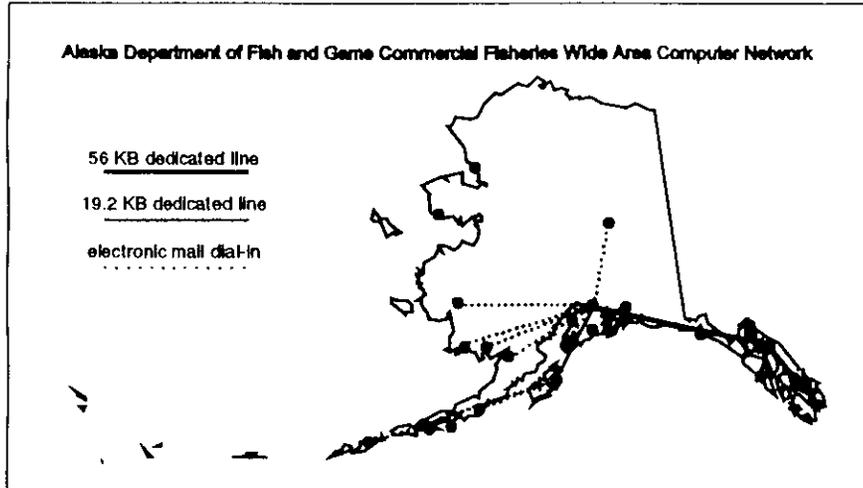
electronic mail dial-in

Figure 1. ADF&G Commercial Fisheries Statewide Computer Network

ADF&G's Division of Commercial Fisheries maintains a wide area computer network that connects the division's computer resources through electronic mail, client-server database applications, and local area file servers. This network is the backbone for electronic transfer of computer files and data within the division.
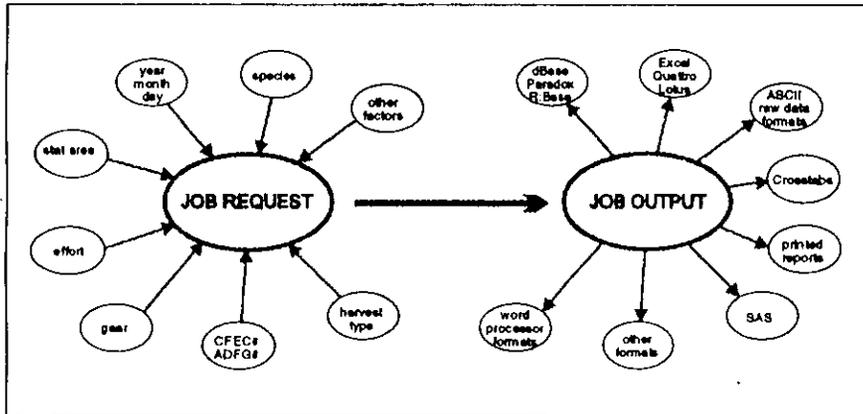
Figure 2. ADF&G Historical Fishery Harvest Database

Twenty years of fishery harvest information is stored in this centralized database. Users may request data extracts and reports that incorporate a large number of parameters, and obtain output customized for a variety of software packages and printers.
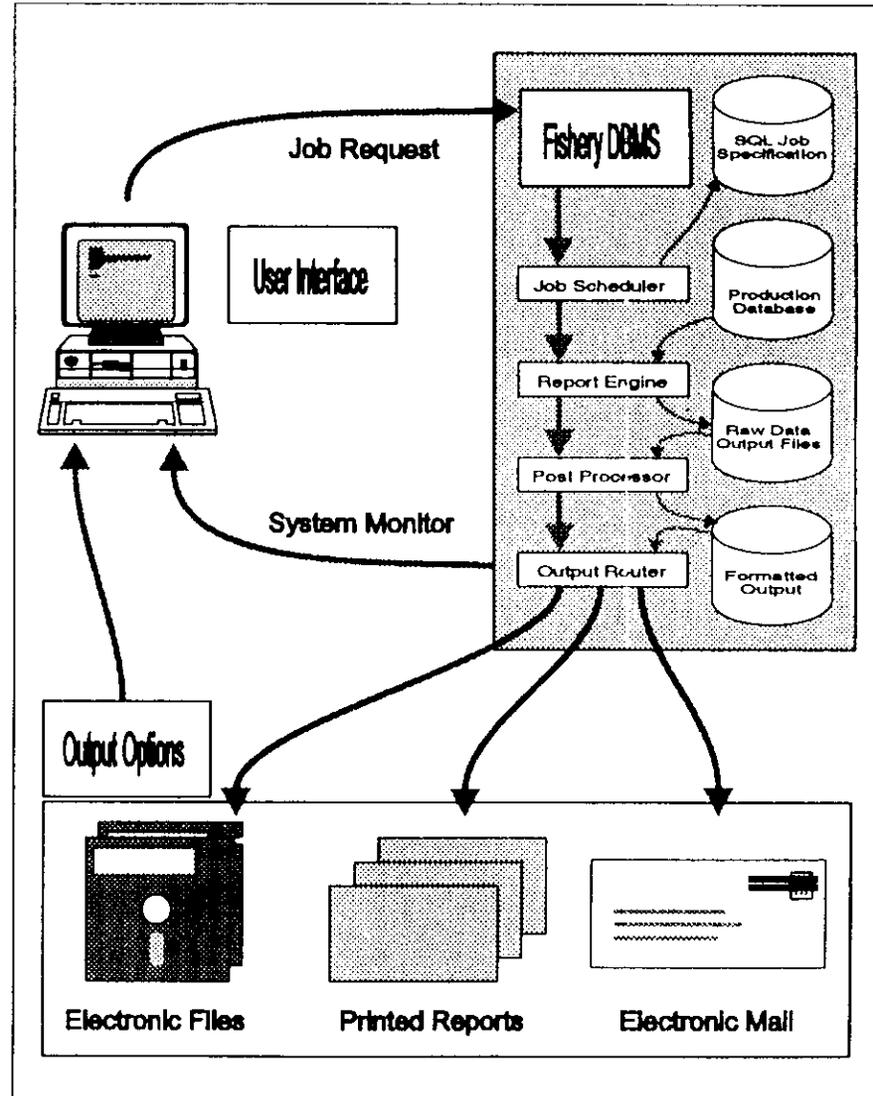
Job Request

Fishery DBMS

SQL Job Specification

User Interface

Job Scheduler

Production Database

Report Engine

Raw Data Output Files

Post Processor

System Monitor

Output Router

Formatted Output

Output Options

Electronic Files   Printed Reports   Electronic Mail

Figure 3. Automated Access to Historical Fishery Data

This user interface has been designed to allow remote access to fishery harvest data; turnaround is automatic, and generally occurs in a matter of minutes. Users may track the status of requests in progress, and resubmit previous report requests as required.

B.2