Arctic-Yukon-Kuskokwim salmon reference database:

User's guide for full ProCite database

by

Matthew J. Nemeth and Megan K. Blees

LGL ALASKA RESEARCH ASSOCIATES, INC.
1101 East 76th Avenue, Suite B
Anchorage, Alaska 99518
Email: mnemeth@lgl.com
Phone: (907) 562-3339

Final report prepared for the

Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative
Scientific Technical Committee

November 2005

EXECUTIVE SUMMARY

References from salmon-related studies in the Arctic-Yukon-Kuskokwim (AYK) region of Alaska were collected from a variety of compilations generated by management programs, research initiatives, and commercially available bibliographic indexes. The references were placed into a bibliographic database to provide a searchable tool to help researchers, managers, and planners access reported salmon information from throughout the region. Over 5500 references were added to the database, primarily from existing lists held by various salmon-related programs in Alaska. Of these 5500, approximately 1600 are related to salmon specifically in the AYK region; the remaining references are from other areas, but may have transportable value to the AYK region. Reports (59%), journal articles (26%), and book sections (9%) comprised the majority of the references imported into the database. References were entered in ProCite (version 5), a commercial reference database that allows searching by multiple fields. For this initial version, the search was restricted to "low-hanging fruit" - sources and lists that were likely to provide a high number of references with a minimum amount of effort to import into the database. Future updates to the database may need to pursue references available individually or in small collections. A subset of the references are in bibliographic format on the AYK SSI web site at http://www.aykssi.org. The full ProCite database is available from:

Karen Gillis or Joe Spaeder
AYK Sustainable Salmon Initiative
c/o Bering Sea Fishermen's Assoc.
705 Christensen Dr.
Anchorage, AK 99501
jjspaeder@earthlink.net
907-235-0531 http://www.aykssi.org

Suggested citation:
Nemeth, M. J., and M. K. Blees. 2005. Arctic-Yukon-Kuskokwim salmon reference database: user's guide for full ProCite database. Final report for the AYK SSI Scientific Technical Committee by LGL Alaska Research Associates, Inc., Anchorage, Alaska.

BACKGROUND AND PURPOSE

This report describes the compilation of a database of published and unpublished salmon references from the Arctic-Yukon-Kuskokwim (AYK) region of Alaska. The project was conducted because most references for salmon from this area are spread across many collections, because most holdings are not in a searchable database, and because a large proportion of the salmon studies from the region are described in unpublished documents that are relatively inaccessible. Although unpublished, these reports contain some of the most current and important information on salmon science and management from the region. The objectives of this project were thus to 1) identify both published and unpublished information on salmon from the AYK region, with special emphasis on 'grey literature" reports that are often difficult to find, and 2) to place these references into a searchable bibliographic database that can easily be updated with new references in the future.

The AYK Sustainable Salmon Initiative (AYK SSI) funded the development of this AYK salmon database in the belief that the database would be a useful resource for a number of salmon initiatives and programs in the region. As of 2005, at least five major programs were scheduled to disburse funds for salmon assessment or research in the AYK region (AYK SSI 2005). These programs can refer to the salmon reference database when setting priorities for program goals and when developing requests for proposals. Individual scientists that respond to these requests can use the database when designing studies and developing proposals.

The database was developed using ProCite bibliographic software (ISI ResearchSoft, Berkeley CA), one of the better-known commercially available software programs. ProCite is fully searchable, allows references to be grouped into similar topics, and is compatible with word processing software. Many online journals and databases now allow references to be automatically exported into ProCite format. The 2005 version of the database was intended to capture both published and unpublished references. Many (but not all) of the individual references were edited to meet the publication style of the American Fisheries Society when converted into bibliographies. The search was extended only back to the 1980's so that more time would be devoted to the most recent references that were presumably the most relevant to present salmon populations and management.

The collection of all salmon-related references from the AYK region would take a tremendous amount of energy. To maximize efficiency, our emphasis in 2005 was to capture those references that were available in large series or groups of references known to be produced by organizations in the AYK region. We assumed that these large series or groups would contain about 90% of the recent references, and that it would therefore take about as much time to collect these as it would to collect the 10% that were not. This strategy was often referred to as the "90% strategy". Uncollected references (the "missing 10%") should be added opportunistically by AYK SSI staff in the future.

REFERENCE SOURCES

Salmon references from the AYK region were obtained from three major categories of sources (Table 1). The first source was agencies or organizations with a historic research or management presence in Alaska. Knowledgeable personnel from the Alaska Department of Fish and Game (ADF&G) and from the U.S. Fish and Wildlife Service (USFWS) were consulted to obtain lists of references from internally-produced reports. The second source was salmon initiatives at the subregional level known to have already compiled some references. References from these first two sources typically existed in lists produced in word processing software, usually Microsoft Word. Such lists usually ranged from 100 to 600 references. The third source was large databases or collections that could be searched to identify a subset of references relevant to AYK salmon. Filters were set up to import these lists into ProCite. When references were not in identical format, multiple filters were needed and an attempt was thus made screen references based on geography, subject, or date (Table 2).

Approximately 5500 individual references were imported from 11 major sources (Figure 1). Approximately 2000 references were from peer-reviewed publications (journal articles or books), and another 3300 were from unpublished reports. The remaining references were maps, web pages, conference proceedings, graduate thesis, and so forth (Figure 2). Each source was given a "Group" label in ProCite, and all records were placed into the Group from which they originated. The origin source was also identified in field #44 ("Call Number") of each record. Database users should thus be able to determine the source of each reference by either looking at field #44 in any given record, or by calling up an entire group while in the ProCite database.

Duplicate records were identified by searching for records with similar titles, publication dates, and authors. Of the 6900 records initially compiled, approximately 1300 were duplicates. The greatest source of duplication was in ADF&G publications (primarily unpublished reports), which had been imported from multiple ADF&G sources and from the Norton Sound database. Duplicate records were reviewed individually, and the version with the most amount of peripheral information was kept. When this was a version from the FFW or the EVOS databases, the entire record was proofed for errors (unlike the rest of the FFW and EVOS records).

An overview of references by source is as follows:

Agencies and organizations

Three different reference series were imported from the ADF&G and another three were imported from the USFWS. The number of references included (in parentheses) and a description is as follows:

- ADF&G Regional Information Report (RIR) series (730 references). Produced by ADF&G, Commercial Fisheries Division (ADF&G CF). This series was listed in Microsoft Word, so a special filter was set up to import the list. All reports from Region III (the AYK region) were included; these reports range in date from 1988 to 2004. Some

reports were eventually published in peer-reviewed literature, but most remain as unpublished works. Of the 730 RIRs imported into the database, 513 of which were came from the RIR lists provided by ADF&G and are specific to the AYK region.

- ADF&G Sport Fish Technical Report series (1705 references). Produced by ADF&G Division of Sport Fish (ADF&G SF). This series was listed in EndNote bibliographic software. Because EndNote is easily imported into ProCite, all statewide Technical Reports were imported and included in our database so they can be evaluated at the user's discretion. These reports range in date from the 1960's through 2004. Some of these reports were eventually published in peer-review literature. Many of the 1705 references imported into the database are from outside the AYK region.

- ADF&G Technical Paper series (42 references). Produced by ADF&G Division of Subsistence (ADF&G DS). This series was listed in Adobe Acrobat PDF format. All reports going back to 1980 were included in our database. Some of these reports were eventually published in peer-review literature.

- USFWS Data Series (56 references). Report list was provided in Microsoft Word. Includes unpublished data series from projects conducted on USFWS Wildlife Refuges and adjacent areas. All reports published to date from Region 7 (Alaska) were included. Reports in this series were never published in peer-reviewed literature.

- USFWS Progress Report series (25 references). Report list was provided in Microsoft Word. Includes unpublished data series from projects conducted on USFWS Wildlife Refuges and adjacent areas. All reports published to date from Region 7 (Alaska) were included. Reports in this series were never published in peer-reviewed literature.

- USFWS Technical Report series (78 references). Report list was provided in Microsoft Word. Includes unpublished data series from projects conducted on USFWS Wildlife Refuges and adjacent areas. All reports published to date from Region 7 (Alaska) were included. Some of these reports were eventually published in peer-review literature.

- Minerals Management Service (35 references). These include references for the Northeastern Bering Sea, and were available in ProCite. The records are not specific to salmon, but included studies of marine resources, environmental quality, and oceanography. The references were filtered for relevance before including in this database. Those included in this database date back to 1987, and include both published and unpublished works.

Salmon initiatives at the subregional level

The database included references compiled by the Norton Sound Research and Restoration Plan (2002), by the AYK SSI Research and Restoration Plan (2005 draft), and by the Exxon Valdez Oil Spill Council (EVOS Council). We also examined the references used by the National Research Council for their report on AYK region salmon (NRC 2004), but could not import these references into ProCite due to formatting incompatibility.

- Norton Sound Research and Restoration Plan (NSRRP; 535 references). These references were already compiled in ProCite form, so all existing references were imported into our database. The NSRRP database includes all salmon-related references from the Norton Sound region. It also contains other references deemed useful to salmon research in Norton Sound, either because the references were from topic relevant to salmon (e.g., water quality), were from adjoining geographic areas (e.g., the Yukon River region), or were on a topic that could be useful to developing new research in Norton Sound (e.g., food chain studies in the Columbia River estuary). These references date back to 1980, and include both published and unpublished works (Nemeth 2002).

- AYK SSI Research and Restoration Plan (AYK SSI RRP) 2005 draft (110 references). These references were listed in Microsoft Word format. They were all imported into our ProCite database, usually individually since they were not all listed in a standard format. References date back to 1982, and are primarily works published in peer-reviewed literature.

- National Research Council AYK Salmon Report (NRC 2004; 420 references). These references were provided in a Microsoft Excel spreadsheet, or in Microsoft Word. Each reference was contained entirely in one cell in the spreadsheet. We were unable to devise a filter for importing these records into ProCite, even after consulting by telephone with the technical staff at ISI ResearchSoft (which distributes ProCite). Many of these references were already captured by the various other searches and imported into our database; those that were not are included in the "difficult 10%" that will need to be imported individually in the future.

- Exxon Valdez Oil Spill Council (EVOS; 2326 references). These references were provided in ProCite format as part of the GEM and TC databases. These databases include a wide range of references that the EVOS Council had compiled as part of its work in Prince William Sound, Alaska. Many of these references were not relevant to salmon or to the AYK Region; however, they were all imported into our database because many of them contained full abstracts (unlike references from the rest of our sources). These references date back to 1930, and include both published and unpublished works. Many of the references were originally added to the GEM and TC databases over time, and there is substantial variation in the way the references are formatted. This variation made it impossible to standardize in our database in way that would allow the citation to be output in American Fisheries Society (AFS) style. Nearly all incompatibilities with AFS style guidelines come from references originating from these two EVOS databases.

Large databases or collections

- Several commercially-distributed databases are available at the Alaska Resources Library and Information Service (ARLIS). The most relevant one was deemed to be Fish and Fisheries Worldwide (FFW), which primarily indexes peer-reviewed works. A number of searches (Table 2) were conducted to winnow references down to those relevant to salmon in the AYK region. A few initial searches on other databases (e.g., Biosis, Agricola) yielded

redundant references, so FFW was used exclusively thereafter. FFW references were able to be imported into ProCite fairly easily. Like the EVOS references, however, many of the FFW references contain severe formatting inconsistencies that prevent proper output in AFS style. Some references were re-formatted in our database using the Global Edit menu in ProCite and are now compatible with AFS style. The remainder will need to be corrected individually. A total of 233 references from FFW were imported into the database.

Quality control and error checking

Once imported into ProCite, an attempt was made to edit the references for output into the publication style used by the American Fisheries Society (AFS 2005). The primary fields used for in-text citations and bibliographies were screened for errors. In most cases, errors followed a pattern that could be detected and corrected using the Global Edit feature in ProCite. When errors did not follow a pattern, references were always corrected if they were in a primary field (such as the Author field), but often uncorrected if they were present in a secondary field. Few errors were corrected from the FFW and the EVOS databases, which simply had too much variability in their fields to be able to proof in the time allowed.

Typographical errors were considered the lowest priority corrections because they can be identified and corrected with the spell-checker functions of the word-processing programs that will presumably be used by end-users of the database. Typographical errors were corrected opportunistically, however, in the course of editing other aspects of individual records.

USER TIPS

ProCite contains several built-in features useful for managing the references in this database. In the Search menu (3rd tab on main ProCite screen), users can search for the references at several levels of detail. Examples of this include:

- Search all records for words found anywhere within a record;
- Search for words found only within specific fields;
- Search for categories of records ("workforms") such as journals, conference proceedings, etc.
- Conduct nested searches by searching within results of prior searches;

Database users can also organize records into new Groups based on origin, date, keywords, locations, and so forth. This function is often used to mark all references that appear in a given publication, or that apply to a given topic, species, or area.

Several aspects of the database have been designed to assist with searching and grouping functions. Keywords have been included with most references (Field #45). As mentioned earlier, the source of each reference was also listed in the Call Number field (#44).

The filters that were created to import records from a number of sources have been saved and are available with this database. Users that wish to import similar references from similar lists can adapt these filters to simplify the process.

One of the fundamental benefits of bibliographic databases is that they can be linked with word processing software to assist with document production. Using the Tools menu in Microsoft Word, references from ProCite can be inserted into the document as in-text citations (e.g., Jefferson and Franklin 1776). Using the Tools menu again, these in-text citations can then be converted into a Bibliography at the end of the document. Users can choose from dozens of pre-loaded bibliographic styles; the default setting on the database is currently set to "Transactions of the American Fisheries Society_mjn." This style is an adaptation of the AFS publication style that is customized to provide better integration with the references in this database.

RECOMMENDATIONS FOR FUTURE WORK

Database management

A master copy of the database should be used for all future edits and additions to the database; as appropriate, this master copy can be distributed to potential users as "update versions" of the database. One fundamental key to the management of the database is that all initiatives to edit and update the master copy are approved by a single individual with a good working knowledge of ProCite. This is especially important for the additions of new references added across relatively large time ranges; such additions are subject to "formatting creep" in which the entries into the database fields become inconsistent and thus difficult or impossible to export into a properly formatted bibliography. Such variation is evident in records imported from the FFW and EVOS databases. One way to reduce variation is for the manager of the database to create cheat sheets that describe the proper formatting of entries into different fields (e.g., authors are separated by double backward slashes in Field #1).

Correction of current reference entries

The biggest source of remaining headaches is from inconsistently-formatted records imported from FFW and from the EVOS databases. These formatting problems should not affect the search aspects of the database and are really only a concern when attempting to print out bibliographies. For this reason, references from these sources are not included in the bibliographic format available on the web (www.aykssi.org), but are instead distributed with the database version (in ProCite). Remaining typographic errors and formatting problems should be corrected on a case by case basis.

Addition of new references

New references from other relevant work will need to be added either individually or in groups from smaller compilations that those already found. Universities, government agencies, and local or regional organizations should be contacted for relevant lists of publications. The best format for such lists is Word or a text file. New journal articles can be searched for annually.

Most commercially-available search software now exports into ProCite with a two-step process, allowing relatively rapid updating with these new references.

REFERENCES

American Fisheries Society (AFS). 2005. AFS publication style guide. Available on the world wide web at: http://www.fisheries.org/html/publications/styleguide/styleguidetoc.shtml. Accessed April 2005.

Arctic-Yukon-Kuskokwim Sustainable Salmon Initiative (AYK SSI). 2005. Available on the world wide web at: http://www.aykssi.org/. Accessed April 2005.

ISI ResearchSoft. 2003. ProCite 5 users guide. Institute for Scientific Information, Berkeley, California.

Nemeth, M. J. 2002. Norton Sound salmon bibliography: a reference database to accompany the Norton Sound Research and Restoration Plan. Interim report for the Norton Sound Science and Technical Committee by LGL Alaska Research Associates, Inc., Anchorage, Alaska. 72p.

NRC (National Research Council). 2004. Developing a research and restoration plan for Arctic-Yukon-Kuskokwim (western Alaska) salmon. National Academy Press, Washington, D.C.

Table 1. Summary of sources, contacts, and treatments of references considered for entry into AYK salmon database.

Organization	Division or Program	Type of Reference	Region of coverage	Contact person	Format	Imported into Database?	Formatted for AFS style?	# References	Years (range)
USFWS	Subsistence Fisheries	Data Series	National Wildlife Refuges, adjacent areas, in AK	Rod Simmons	MS Word	Y	Y	56	All reports published to date
USFWS	Subsistence Fisheries	Progress Reports	National Wildlife Refuges, adjacent areas, in AK	Rod Simmons	MS Word	Y	Y	25	All reports published to date
USFWS	Subsistence Fisheries	Technical Reports	National Wildlife Refuges, adjacent areas, in AK	Rod Simmons	MS Word	Y	Y	78	All reports published to date
ADFG	Commercial Fisheries	Regional Information Reports	ADFG CF Region III	Susan McNeil	MS Word	Y	Y	513	1988-2004
ADFG	Sport Fish	SF Technical Series, some CF reports, & outside pubs	Statewide (all AK regions)	Joanne MacClellan	Endnote	Y	Y	1705	1960's-2004(?)
ADFG	Subsistence Fisheries	DS Technical Paper Series	Alaska	Jim Simon	PDF	Y	Y	41	1980-2004
AYK SSI	AYK SSI	References for Research and Restoration Plan	Arctic-Yukon-Kuskokwim	Joe Spaeder	MS Word	Y	Y	110	1982-2005
National Research Council[1]	-	References for AYK Salmon report	Arctic-Yukon-Kuskokwim	Joe Spaeder	PDF	N	N	~420	-
EVOS[2, 4]	Technical Committee	GEM BIB	Statewide, PWS, GOA, Atlantic and Pacifc coast	R. Bochenek & C. Womac	ProCite	Y	N	2326	1930-2004
EVOS[2, 4]	Technical Committee	TC Bib	Prince William Sound (oil spill area)	R. Bochenek & C. Womac	ProCite	Y	N	Inc. in above	?
Fish & Fisheries Worldwide[3, 4]	Database	Peer-review	Arctic-Yukon-Kuskokwim	ARLIS	ProCite	Y	N	233	1922-2004
MMS	Outer Continental Shelf	Report Bibliography	Northeastern Bering Sea		ProCite	Y	Y	35	1980-current
Norton Sound Steering Committee	Fisheries Disaster Relief Fund	Database	Norton Sound, Alaska	Matt Nemeth	ProCite	Y	Y	535	1980-2004

[1] = References unable to be converted into ProCite.

[2] = References imported into AYK salmon database without filtering with terms in Table 2.

[3] = References filtered using terms in Table 2 before entering into AYK salmon database.

[4] = Formatting errors corrected in major fields used in bibliographies (title, authors, etc.). Remaining fields may be incorrectly formatted; all fields may contain typographical errors.

Table 2. Search terms used independently or in combination when searching large databases (e.g., Fish and Fisheries Worldwide) to identify references relevant to the AYK salmon database.

Geographic	Biological	Organizational
Alaska	Salmon	ADF&G
Yukon	Chum	USFWS
Arctic	Coho	USGS
Kuskokwim	Chinook	University
Norton Sound	Sockeye	MMS
Chukchi	Pink	
Bering Sea		
Kotzebue		
Northeast Pacific		

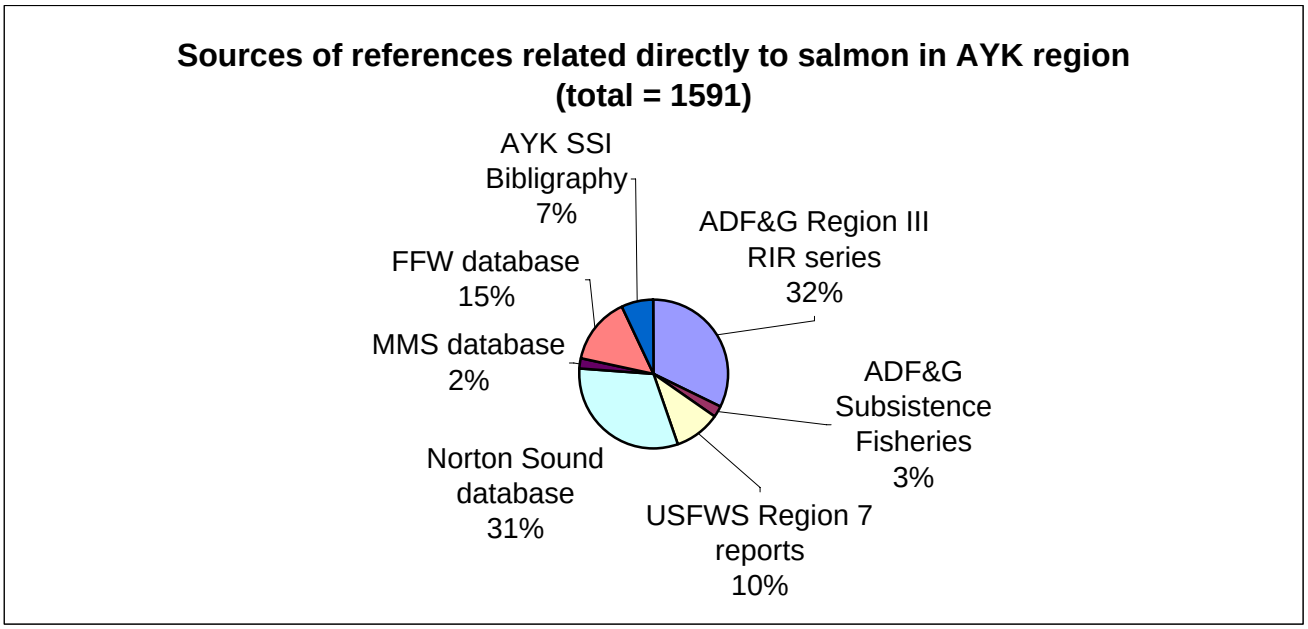

Figure 1. Origins of the 1591 references that came from sources directly within the AYK region.

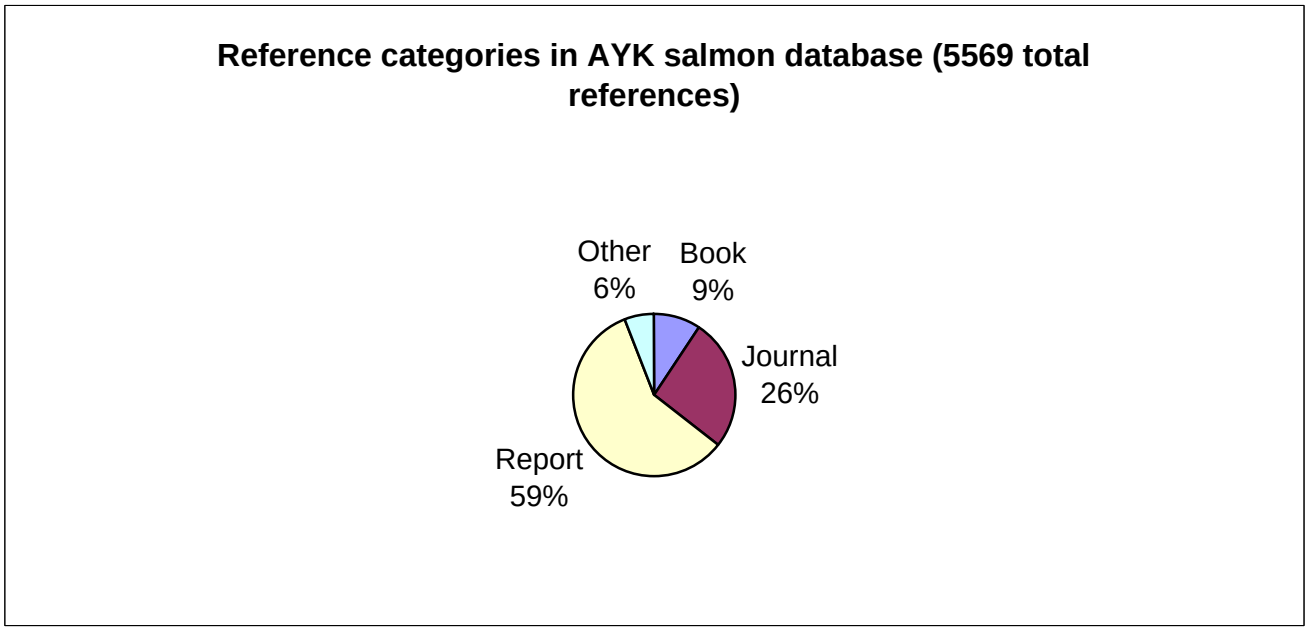

Figure 2. Categories of references contained in the AYK salmon database. Journal articles and books were assumed to be peer-reviewed, while reports were not. Other includes maps, web pages, conference proceedings, and databases.