CREATING A NORTH AMERICAN STURGEON INFORMATION INFRASTRUCTURE: IMPLICATIONS FOR COMPOSITE DATABASES AS A MULTIJURISDICTIONAL MANAGEMENT TOOL

By

Tracy L. Kolb

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Fisheries and Wildlife

2008

ABSTRACT

CREATING A NORTH AMERICAN STURGEON INFORMATION INFRASTRUCTURE: IMPLICATIONS FOR COMPOSITE DATABASES AS A MULTIJURISDICTIONAL MANAGEMENT TOOL

By

Tracy L. Kolb

Concerns about the decline of fisheries resources is becoming a mainstream policy issue. At the core of addressing fisheries decline is the need to maintain and publish knowledge about fish species, their distributions, and relative abundances. Compiling this knowledge entails a need for sharing information across multiple jurisdictions and disciplines. In order to address this need, I developed a nationwide pilot status and trends information system for lake sturgeon, called the Sturgeon Information Infrastructure (SII).

SII combines historical and current sturgeon status and trends data from state federal, and tribal agencies, academic institutions, and private organizations. Data were collected, standardized and entered into an online relational database, where they are searchable and mappable. Constraints on creating SII include lack of a standardization and classification system for lake sturgeon status and trends, lack of available and standardized georeferenced hydrography information, reluctance to share data amongst data providers, and lack of historical datasets about distribution and population abundances of lake sturgeon.

ACKNOWLEDGEMENTS

First and foremost I want to acknowledge Dr. William Taylor for accepting me in the Fisheries and Wildlife family at Michigan State. His confidence in my abilities has never wavered and he has continually pushed me to be a better and more balanced person. A project of this magnitude does not happen without the collaboration of many individuals. I am indebted to Doug Beard, Jack Liu, Scott McGill, Andrea Ostroff and Andy Loftus for their guidance and support throughout this project.

Thanks to Amy Schueller, Dan Hayes, Dana Infante, and Jon Hansen. Without their support, encouragement, and love I could not have completed this journey. I also owe a great deal of gratitude to Tim Boyd and Ed Bissell for stoically enduring my initial attempts at project management and having limitless patience during the design and redesign of the website.

Thanks also to the numerous individuals who gave me support: my lab mates, the Lake Huron Technical Committee, Jim Johnson, Jim Boase, Andrew McAnich, Chris Geddes, Tracy Hill, Nick Mandrak, Robert McLaughlin, Jeff Schaefer, Scott Nelson, Bruce Manny, Emily Zollweg, Gary Whelan, Randy Claramunt, Tim Haxton and Dave Clapp, my steering committee, and all of the participants that took the time and effort to take my data sharing survey.

Lastly, I want to thank my parents David and Maxine Kolb, my sisters Rae Beth and Natalie, and Kendra, the love of my life. You are my past, present and future. I love you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Research directed at addressing global issues such as bio-complexity, sustainability, and ecosystem change, while supporting regional and national policy and management decisions is needed in order to address emerging threats to natural resource conservation (Porter 2004; Hale et al. 2003; Brunt et al. 2002). These types of research transcend traditional disciplinary, jurisdictional, spatial, and/or temporal boundaries, deal with issues that require the input of more then just one or a few individuals, and require synthesizing and combining data that are collected at finer ecological scales (Baker et al. 2005; Han et al. 2002). As such, integrating datasets from multiple sources can be a powerful scientific tool for using diverse information in ways that support the decision making process or address broad scale issues (Porter 2004; Hale et al. 2003).

A common tool for integrating diverse datasets is the composite database (McLaughlin et al. 2001). A database is a collection of data or records that have structure imposed on them. I have coined the term “composite” database to refer to databases that standardize and integrate data from multiple sources. Composite databases are used because they imbue a project with several advantages. First, data cost less to reuse than to re-collect, reducing the need for both personnel and fiscal resources (Porter 1998). Second, many types of historical datasets can’t be recreated. Hence, there is utility to preserving historical datasets as they provide valuable information on past environmental conditions that can never be re-collected. Third, composite databases prevent or slow “data decay”. Data decay refers to data that have been traditionally collected and analyzed by a single individual or small group, and over time, our ability to locate and interpret those data has been diminished or lost through lack of documentation (Porter and Ramsey 2005). Most importantly, composite databases can be used for a wide variety of scientific inquiry, including: long-term studies, which use databases to retain project history; syntheses, which combine data for purposes then they were otherwise intended; and integrated multidisciplinary projects (McLaughlin et al. 2001; Porter 1998).

Composite databases and fisheries management

For North American fisheries resources, one essential need that a composite database project can accomplish is to map the spatial distribution of fish populations. Understanding current and historical distribution patterns is an essential prerequisite for determining causal relationships between ecosystem change and aquatic processes (Watson et al. 2004). Any plan that attempts to address management strategies for mitigating fisheries loss must have insight into historical as well as current species distribution, so scientists can determine where species are and how they are doing relative to past conditions and provide policy makers with needed general characterizations of a resource across time and space.

The United States Geologic Survey (USGS) and the United States Fish and Wildlife Services (USFWS) have mission statements that require them to preserve and protect natural resources at a national level to ensure the continued availability of those resources for ecosystem services and human uses. In order to fulfill those mission statements natural resources need to be inventoried and characterized so that threats can be identified, mitigated, and potential benefits optimized. There are a number of programs to help federal natural resources agencies meet their missions. One of these programs is the Fisheries and Aquatic Resources Node (FAR), initiated in 2001. The goals of this program are to provide an integrated, web-based resource that will: coordinate and link to fishery databases across the United States, as well as provide data on fisheries distribution and trends through time.

In 2004 the USGS in concert with FAR developed a 5-year strategic plan to help meet agency missions. They solicited input from peers within and outside of the USGS, including other Department of the Interior (DOI) bureaus, federal and state agencies, and non-governmental organizations. They found there was an immediate need to assess the status and trends of the nation's biological resources by moving beyond a "large collection of projects", towards an integrated effort to maintain and publish knowledge about species, their distributions, and relative abundances, that is, their status and trends.

Building a sturgeon information infrastructure (SII)

To address these issues the USGS solicited a prototype, internet-based information system that allows for assessment of status and trends of an aquatic species across its entire range. As completed, the information system was meant to be used as a template for development of additional information systems as well as could be used as a tool to manage inland freshwater aquatic resources. This system was developed in order to build a model for tracking status and trends of a species across its range, not to conduct a complete status report on the template species itself.

Sturgeons were chosen as the template species because they occur in the continental United States, and are a species that the USGS is interested in characterizing. The USGS is interested in sturgeons because there is significant public interest in collecting information and maintaining information about their status and trends.

Additionally, the majority of sturgeon biologists are nearing retirement age, so there appeared to be a need to provide for the long-term care of their data, making “data decay” a looming concern, (D. Beard, USGS, personal communication).

Lake sturgeon as a template species

The first step in creating the prototype database was to convene a meeting of “sturgeon experts” from around the country. The experts raised serious logistical concerns about data availability and personnel, with regards to integrating information from all 8 species of North American sturgeon. Based on the validity of these concerns the project was further streamlined to include only the single species lake sturgeon (*Acipenser fulvescens*). The lake sturgeon information system could still serve as a prototype by combining multiple datasets from diverse sources across the Great Lakes basin. This prototype system built for lake sturgeon is known as the Sturgeon Information System (SII).

The lake sturgeon is a late-maturing, slow-growing, long-lived fish (Cook et al. 1987). Lake sturgeon are found in many large rivers and lakes in North America. While there are some remnant or introduced populations in the southern and central United States, most sturgeon populations are in the Great Lakes basin including the Canadian provinces of Ontario and Quebec (Auer 1999, Baker 1980). Lake sturgeons travel within a home range and return to spawn in natal tributaries in spring (Sandilands 1987, Dumont et al. 1987, Priegel and Wirth 1971). Lake sturgeons are an ideal candidate for establishing this prototype system, because in 2003 scientists agreed on a common scheme for classifying lake sturgeon relative abundances (Zollweg et al. 2003).

SII Goals and objectives

The lack of standardized biological information on aquatic species available at national scales makes it essential that we build research and management tools with the ability to integrate what local or regional data already exist. Therefore the overarching goals for this project were to develop an internet-based information system for the scientists at USGS, that combined lake sturgeon status and trends information from multiple sources across the Great Lakes basin. Specific objectives were to build a composite, geospatial, database, for lake sturgeon with the ability to store existing data and integrate new data as they become available for: current and historical distributions of lake sturgeon, current and historical changes in lake sturgeon abundance, and types of locations of current and historical research projects related to lake sturgeon.

METHODS

Status and trends definition

The first step in creating the SII was to define a unit of measurement for lake sturgeon. Lake sturgeon are potadromous, and home to their natal streams during reproduction (Baker 1980; Harkness and Dymond 1961). Therefore the most natural unit of measurement of status and trends is at the level of populations. All SII status and trends information reflects the health of geographically distinct populations, defined by natal spawning tributaries.

For the purpose of measuring status and trends the following definitions were used: Status information is delineated by three general scales of information. At a coarse scale, status is defined as simply the presence or absence of a lake sturgeon population. At a finer scale, status is classified by categories denoting the relative abundances of lake sturgeon populations. At the finest scale, status is recorded as a population estimate in absolute numbers or ranges of numbers. Status was never inferred from a course scale to a finer scale, but if status was available at the abundance level, the status and presence of the record was inferred.

Trends were defined by changes in relative abundances through time, which at the coarsest scale can be characterized as a change from present to absent or vice versa, at a finer scale as a change from one status classification to another, and at the finest scale as a change in numbers of sturgeon present in a given population.

Information on sturgeon research was classified into broad categories representing the most common types of research. Research types were categorized as follows. First basic biological data collection, where researchers collected characteristics of sturgeon during a sampling period such as age, gender, length, and weight. Population estimates, where researchers actively used quantitative methodology to determine a population estimate for sturgeon during a sampling period. Telemetry studies, where researchers implanted telemetric tags to study the movements of sturgeon during the course of a given sampling period. Tagging studies, where researchers implanted tags in or on sturgeon, tracking their spatial movement over long time periods. Genetics, where researchers took genetic samples from the sturgeon populations. Contaminant studies, where researchers collected information about the amount or source of different contaminants in a given area during a sampling period. And finally, other scientific studies, which served as a bin for organizing research that didn't fit into the other categories such as sampling environmental conditions, habitat characteristics, other species present, etc.

Data collection

Information on available data was collected systematically by sending 25 surveys to all federal and state agency researchers across the Great Lakes basin inquiring as to the types of sturgeon data they collected and if they were willing to make that information available. Researchers were identified by the Great Lakes Sturgeon Website. Federal and state agency researchers were chosen because they were thought to have previously complied large data sets on lake sturgeon.

Data were also collected opportunistically by speaking to lake sturgeon scientists with the Michigan Department of Natural Resources, United States Geological Survey and the United States Fisheries and Wildlife Service, attending lake sturgeon coordination and American Fisheries Society meetings, and combing through peer-reviewed literature, government reports, online databases and unpublished datasets from the states of Minnesota, Wisconsin, Indiana, Illinois, Michigan, Ohio, Pennsylvania, New York and Vermont, the provinces of Ontario and Quebec, the federal governments of Canada and the United States, Michigan State, Michigan Technical, Central Michigan, State University of New York, Cornell, Wisconsin, St. Mary's, and Purdue Universities, as well as various Great Lakes tribal authorities and private entities (Baker 2006; Zollweg et al 2003; Auer and Baker 2002; Auer 1999; Bruch 1999; Dumont et al. 1987; Baker 1980; Harkness and Dymond 1961) .

Types of data sought included information about historical and current presence. Information pertaining to an existing classification schemes denoting relative abundances of current and historical populations of lake sturgeon, and information pertaining to current and historical research projects. Additionally, where available, citations or methodology used to generate data were sought and compiled. Additional types of information included in the SII were referenced to hatchery-reared populations, endemic populations, successfully reproducing populations, and if sturgeon were present, which life-history stages had been observed: eggs, larvae, juveniles, sub adults or adults. Lastly, information was compiled on which types of harvest could take place at the spawning tributaries. These auxiliary data were determined to be useful additions to the status and trends database, because they allow researchers subset criteria when querying the database for more refined synthesis and analysis (Michael Parsley, USGS, and Mark Collins, South Carolina DNR, personal communication).

Data standardization

After data were collected they were standardized for the SII database. The standardization process included two steps: creating a common naming format and creating a common geo-referencing data scheme. In order to create a common naming scheme, data from different source datasets with the same name but containing the different types information were streamlined and separated (Fig 1). Data with different names that had the same information were combined (Fig 2).

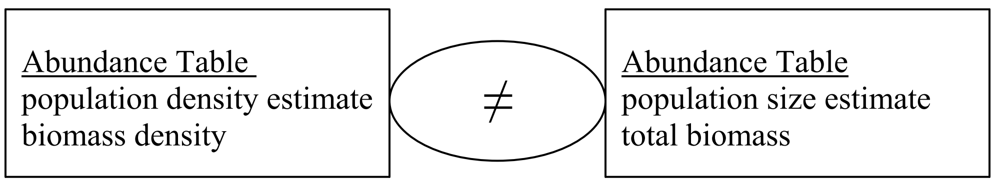


Figure 1. Data in tables that have the name but different meanings are separated during the database standardization process. In this example both tables have data called abundance, but the data within the tables were collected using different methodology.

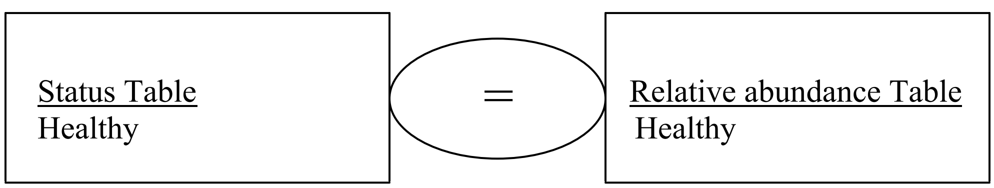


Figure 2: Data in tables with different names having the same definitions are combined during the database standardization process. In this example tables are named differently but hold the same information.

When creating a common geo-classification scheme, all data were geo-referenced using the USGS National Hydrographic Dataset (NHD). NHD is a comprehensive set of digital spatial data containing information about surface water features such as lakes, ponds, streams, rivers, springs and wells. Within the NHD, surface water features are combined to form "reaches," which provide the framework for linking water-related data to the NHD surface water drainage network. Reaches delineate sections of rivers that join at a confluence. Each reach has a unique 14 digit identification number called the reachcode. Each record in the SII database is linked to one or more reachcodes. SII geospatial information using NHD was based on 1:100,000-scale data.

Because data from Canada is not available in NHD form, a shapefile containing Canadian waterbodies, the National Scale Frameworks Hydrology (NSFH), was appended to the NHD. Canadian waterbodies were georeferenced by hand and assigned special case names. Special case names were then manually linked to line features within the NSFH. Canadian hydrography was provided by the Ontario Ministry of Natural Resources (Tim Haxton, personal communication).

Development of SII database

The SII database was developed in Microsoft Access, using relational database rules in order to maximize database integrity (Hernandez 2003). Relational database rules include: 1) tables that are constructed properly and efficiently, i.e.: each table represents a single object, is comprised of distinct fields, keeps redundant data to a minimum and is identified throughout the database by a field with unique values; 2) data integrity that is imposed at the field, table, and relationship level. Information and a schematic of the design is provided in Figure 3, information included in the database is listed in Table 1.

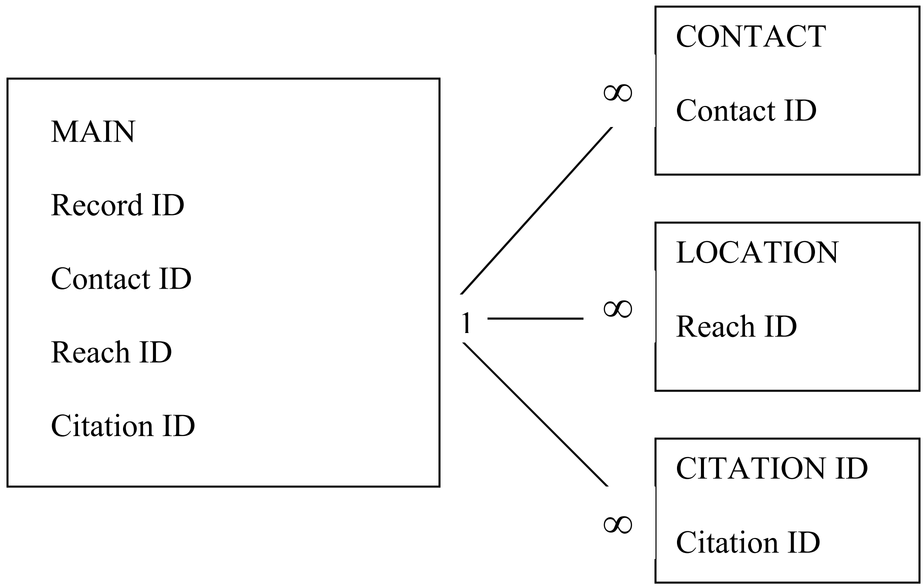


Figure 3: Design of the SII database indicating the data tables, relationships between tables, and fields used as the keys in the relationships.

Table 1. Synopsis of the information contained in each of the tables comprising the SII database.

1. Main
   - start date
   - end date
   - presence/absence
   - status type
   - abundance
   - life history
   - research type
   - endemism
   - reproduction successful
   - metadata

2. Citation
   - citation reference
   - year

3. Location
   - waterbody name
   - basin name
   - state/province name

4. Contact
   - contact name
   - agency
   - address
   - phone
   - email

Data quality was checked using several different methods. All records were verified with the original data. Also as mentioned previously, data were standardized by ensuring that fields with the same names in different source datasets contained the same information in the same format, and fields with different names but containing the same data were unified, because using existing structure from inherited datasets is not considered a "best practice" and should be avoided if possible (McLaughlin et al. 2001, Hernandez 1997). Many problems including poor design (tables that aren't properly linked), and insufficient data integrity (redundancy) that can arise from creating a single database can be compounded when combining multiple datasets. Additionally where possible, information regarding how the record was sampled, including any citations was included.

Development of SII website application

After the SII was created in Access (part of Microsoft Office Professional Edition, 2003, Microsoft Corporation), it was copied to Microsoft SQL Server (2003, Microsoft Corporation), an application that allows a database to be searchable online. A website was created to serve database information. The website also included sections on funding agencies, data providers, information about lake sturgeon, links to other lake sturgeon websites, and a report on the status and trends of lake sturgeon according to the SII database.

The main function of the webpage however, is to allow users to search the SII database online. The query page has fields that allow users to search by location: waterbody name, basin, state or province, status type: presence/absence, status classification, or abundance, research type. All search fields can be narrowed by time period, endemic populations only, successfully reproducing populations only, or by specific life history stages observed. Once the database is queried, all records which meet the criteria are returned in tabular form, with each record hyperlinked to its citations and/or metadata. The records can also be mapped using an interactive mapping site (IMS), created by the MSU Geography department for this project.

Essentially IMS is an ASP.NET website built on the ESRI ArcGIS Web ADF (Application Developer Framework) that uses ArcGIS Server to publish GIS maps to the internet. The IMS sits on a separate server which contains a copy of the NHD and the NSFH with the special case names as fields within the file. IMS is housed at MSU Remote Sensing & Geographic Information Science & Outreach Services (RS&GIS). The IMS can be thought of as the front-end or user interface that allows a user to interact with the geographic data by supplying search criteria that are then processed by IMS.

When a query is performed, a querystring in the form of a unique URL is passed to IMS server through the URL specifying the criteria the user has chosen. The querystring identifies locations to display within a specified geographic area. Applicable geographic areas include standard geolocators such as states, provinces, basins, HUCs (Hydrologic Cataloging Units), Component Basins (Canadian hydrologic units) and waterbodies (named water features). The smallest unit that locations are identified to is the reach level. The smallest unit that users are able to search on is the waterbody which consists of one or many reaches. Once the user chooses to map the results of a query, IMS selects the location criteria passed through the URL and it appears highlighted on the screen. The user then has the option to pan around, or zoom in and out, of the map screen.

Evaluation of SII

To evaluate the SII a steering committee of sturgeon biologists and policy experts was re-convened after development (Table 2). The steering committee was asked to evaluate the system by addressing the questions: what is useful about SII? What needs improvement? What is unnecessary? Is there information or functionality that is missing? Is it useful for biologists, policy makers, public, other scientists, NGO's or industry? Is appropriate access provided to the various groups that you would expect to use the system? Would reports containing status and trends of sturgeon be useful? Are the useful characteristics of status and trends reports that are currently being used in research and decision making reflected in capabilities of the SII? Do you think a status and trends information system such as this would be useful on a regional or national scale?

Table 2. Names, affiliations and areas of expertise of steering committee, convened in January 2008 to evaluate project progress and future needs.

Names	Affiliation	Expertise
William Taylor	Michigan State University	multijurdisctional management
Andrea Ostroff	USGS	information management
Douglas Beard	USGS	policy, information management
Michael Parsley	USGS	white sturgeon
Patrick Braaten	USGS	pallid and shovelnose sturgeon
Vaughn Paragamian	Idaho Fish and Game	white sturgeon
Andy Loftus	Loftus Consulting	information management
Jarrod Kosa	USFWS	multijurdisctional management

RESULTS

Amount, types and sources of data collected

Out of the 25 surveys that were sent to great lakes lake sturgeon scientists, 22 responded. When asked what lake sturgeon data had been or was being collected, surveys showed that information on both distribution and abundance were available through USFWS, USGS, Michigan (MDNR), Wisconsin (WDNR), and Ohio Departments of Natural Resources (ODNR), and New York State Department of Environmental Conservation (NYDEC) (Table 3). However, fewer entities collected abundance information than distribution information.

Table 3. Agency name and type of information collected, from surveys sent to US Fish and Wildlife Service (USFWS), United States Geological Survey (USGS), Michigan (MDNR), Wisconsin (WNDR), and Ohio Department of Natural Resources (ODNR), and New York State Department of Environmental Conservation (NYDEC).

	Information Type	
Agency Name	Distribution	Population Estimate
USFWS	x	x
WNDR	x	x
MDNR	x	x
NYDEC	x	x
OMNR	x	x
USGS	x	

Of great importance to this study we found that, only 10 of 22 scientists were willing to contribute their data to the SII (Figure 4). Reluctance to share data became a bottleneck in the data collection process and represents a critical obstacle to building composite fisheries databases. When asked why they were unwilling to share, the majority of scientists cited issues such as lack of time and fear of data misuse (Figure 5).

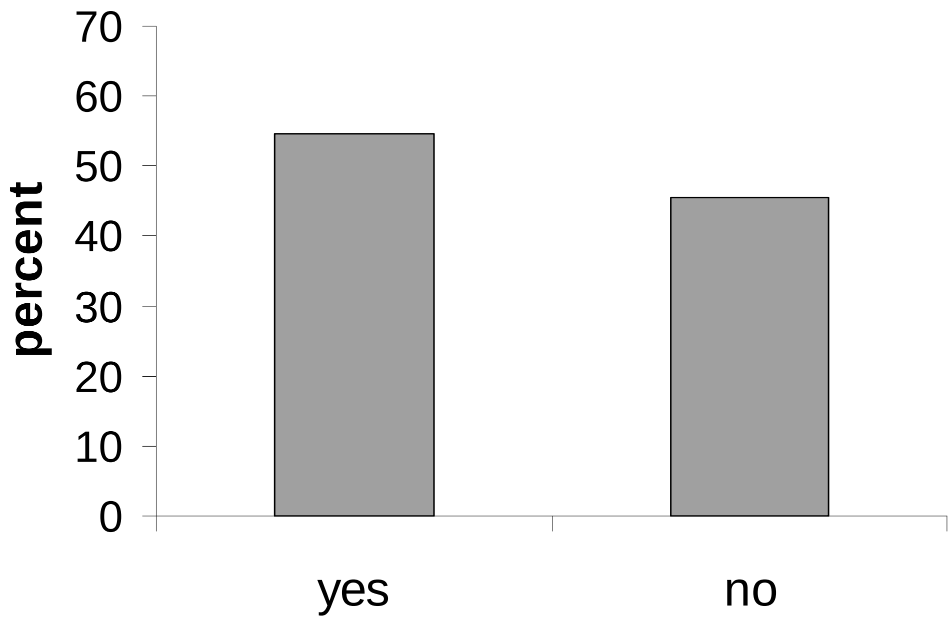


Figure 4. Willingness to share data from 22 surveys, sent to Great Lakes biologists and supervisors, asking yes or no are you willing to contribute lake sturgeon data that your agency collects to a composite database.

Given the reluctance of many professionals to share data, access to large unpublished datasets was limited. However, I was able to access two major sources of data, the Great Lakes lake Sturgeon Tributary Database and Geographic Information System compiled in 2002 by the USFWS, and an unpublished summary report that defined status classifications for relative abundances of Lake sturgeon (Zollweg et al 2003). Because these sources were large in volume they were ideal for the SII. From those sources, I compiled 364 records from 79 scientists, representing 41 agencies, organizations, universities, or tribes. Of the 364 records, 242 record historical presence or absence, 180 record status, 32 record population abundance or estimates and 364 record historical and current research efforts (Table 4).

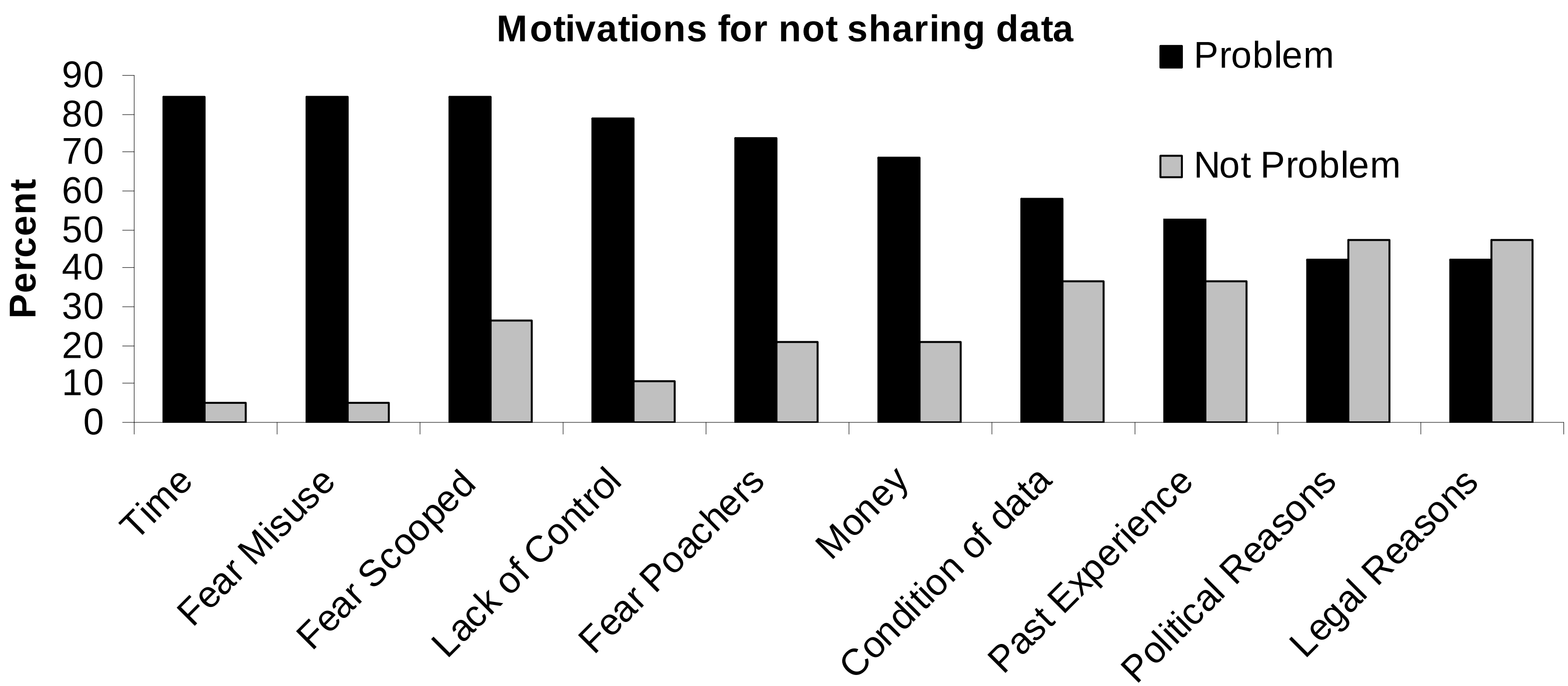


Figure 5. List of reasons that 22 Great Lakes sturgeon technicians, biologists, or supervisors were unwilling to share data. Reasons are: they lack the time, fear data misuse, fear being scooped by other scientists (fear that other scientists will publish on their data before they can), fear they will lose control of their data, fear sturgeon poachers will access location data and use it to illegally poach sturgeon, lack the money, fear that their data are in poor condition, and/or have political or legal reasons for refraining from contributing data. Scientists were able to choose or not respond to as many reasons as they wished.

Table 4. Amount, sources and types of data collected for the SII. Numbers number of sources from each agency and total records contributed for Presence/Absence (P/A), status, population estimates (PE) and research. Sources are Department of Fisheries and Oceans Canada (DFO), Michigan Department of Natural Resources (MDNR), New York State Department of Conservation (NYDEC), Ontario Ministry of Natural Resources (OMNR), United States Fish and Wildlife Agency (USFWS), United States Geological Survey (USGS), Wisconsin Department of Natural Resources (WDNR), Great Lakes tribes, Great Lakes universities and various state, federal, international, and private interests Various are grouped because they contributed an average of 1 record each. Incomplete represents records in the database that cannot be traced back to the original source because of lack of metadata.

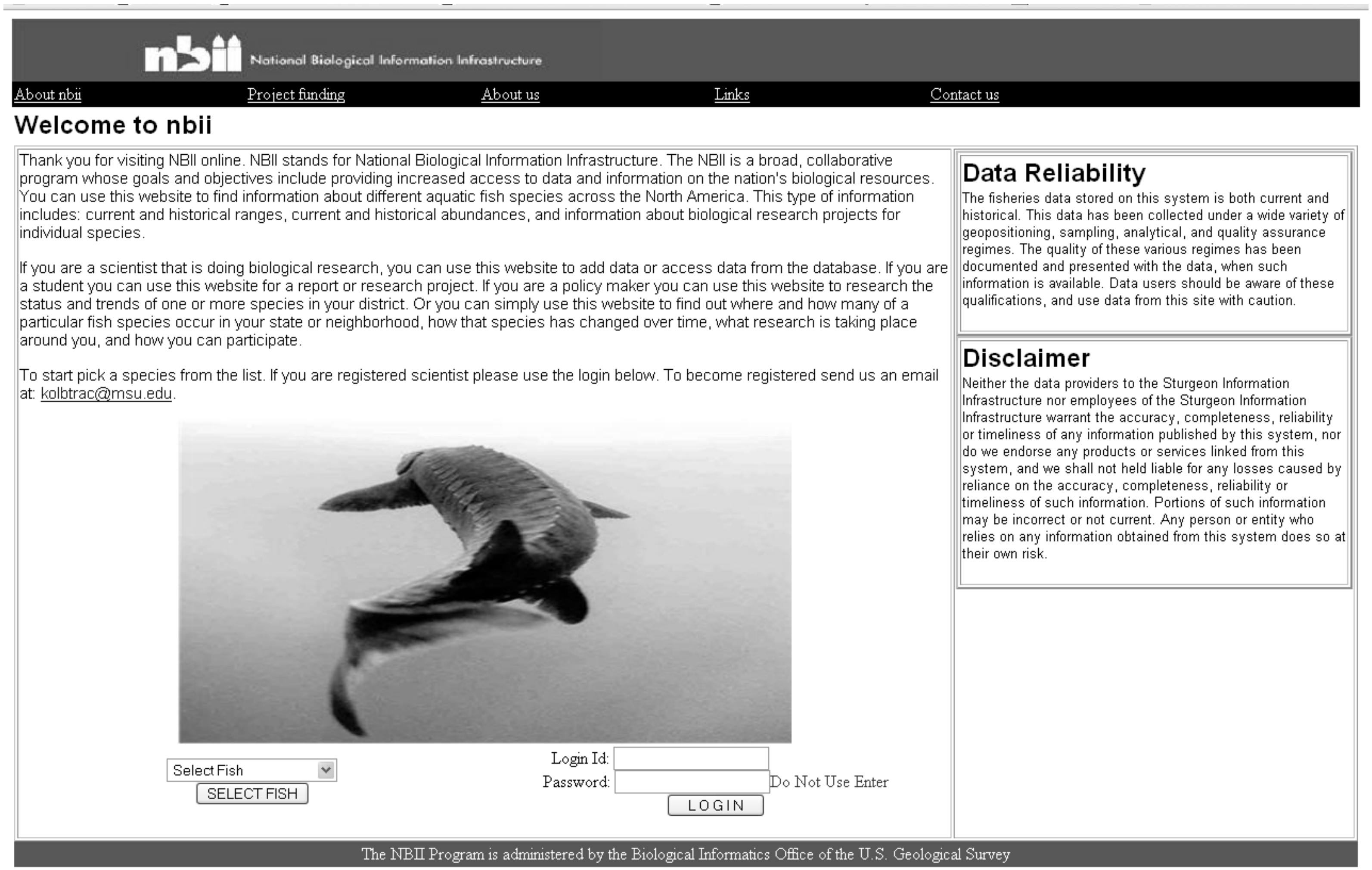
Agency	No. sources	P/A	Status	PE	Research
DFO	3	32	2	1	41
MDNR	5	29	12	3	44
NYDEC	7	9	9	1	13
OMNR	5	50	15	7	50
USFWS	11	28	20	5	38
USGS	3	6	4	0	6
WDNR	8	9	20	4	19
Tribes	11	11	13	0	18
Universities	16	19	24	4	25
Various	10	22	20	7	29
Incomplete	n/a	27	41	0	81

In 2003 lake sturgeon researchers from across the Great Lakes basin developed a classification scheme for lake sturgeon populations denoting their relative abundance (Zollweg et al. 2003). This was the classification scheme that SII used for defining status. There are four major classification types, based on observations of adult sturgeon numbers as they entered tributaries to spawn in spring. Healthy: denotes populations of 1,000 – 10,000 adult spawners. Remnant denotes populations of 10-1,000 adults spawners. Extirpated denotes populations of less than 10 adult spawners, and unknown denotes populations that have unknown amounts of adult spawners. Because this classification scheme wasn't developed until 2003, it couldn't be applied to records from before 2003, unless those records had population abundance estimates.

Description of SII

The SII is available at http://ntweb11.ais.msu.edu/sturgeon/SturgeonLogin.asp. The main page serves has several functions. It serves as a pathway for accessing specific species pages, links to more information about SII, and allows data providers to add or enter data by logging into the system. (Figure 6). Once lake sturgeon is selected. The main lake sturgeon page of the SII allows a user to view lake sturgeon data contributors, link to important lake sturgeon websites, query the SII database, view a report on the SII status and trends database, and learn more about lake sturgeon (Figure 7).

The query page allows users to search by location, status type, population estimate, and research type. All of these search fields can be narrowed by time period, endemic populations only, successfully reproducing populations only, specific life history stages sampled, management type, and by introduced, reintroduced or supplemented populations only. A definition of the field appears when the user moves the cursor over that field name (Figure 8). Once a record is queried. For example Basin = Michigan, Status = remnant a table of data meeting the criteria is returned to the user (Fig 9). A user can then click on the ID number of a specific record to obtain its metadata, citations associated with the record, and if available, how population estimates were obtained. The user can also map the record by clicking on the Map it! tool.


National Biological Information Infrastructure

About nbii | Project funding | About us | Links | Contact us

Welcome to nbii

Thank you for visiting NBII online. NBII stands for National Biological Information Infrastructure. The NBII is a broad, collaborative program whose goals and objectives include providing increased access to data and information on the nation's biological resources. You can use this website to find information about different aquatic fish species across the North America. This type of information includes: current and historical ranges, current and historical abundances, and information about biological research projects for individual species.

If you are a scientist that is doing biological research, you can use this website to add data or access data from the database. If you are a student you can use this website for a report or research project. If you are a policy maker you can use this website to research the status and trends of one or more species in your district. Or you can simply use this website to find out where and how many of a particular fish species occur in your state or neighborhood, how that species has changed over time, what research is taking place around you, and how you can participate.

To start pick a species from the list. If you are registered scientist please use the login below. To become registered send us an email at: kolbtrac@msu.edu.

Select Fish

SELECT FISH

Login Id:

Password: Do Not Use Enter

LOGIN

Data Reliability

The fisheries data stored on this system is both current and historical. This data has been collected under a wide variety of geopositioning, sampling, analytical, and quality assurance regimes. The quality of these various regimes has been documented and presented with the data, when such information is available. Data users should be aware of these qualifications, and use data from this site with caution.

Disclaimer

Neither the data providers to the Sturgeon Information Infrastructure nor employees of the Sturgeon Information Infrastructure warrant the accuracy, completeness, reliability or timeliness of any information published by this system, nor do we endorse any products or services linked from this system, and we shall not held liable for any losses caused by reliance on the accuracy, completeness, reliability or timeliness of such information. Portions of such information may be incorrect or not current. Any person or entity who relies on any information obtained from this system does so at their own risk.

The NBII Program is administered by the Biological Informatics Office of the U.S. Geological Survey

Figure 6. SII sturgeon main page. Toolbars are from left to right: about nbii, project funding, about us, links, and contact us. Along the right side is a general disclaimer and a disclaimer about data reliability. In the bottom left corner is a drop down box for selecting a species, and in the bottom middle is a log in section for data providers.

Figure 7. Lake sturgeon main page. Toolbars are from left to right: participating agencies, about lake sturgeon, queries/reports, lake sturgeon links, and return to main.

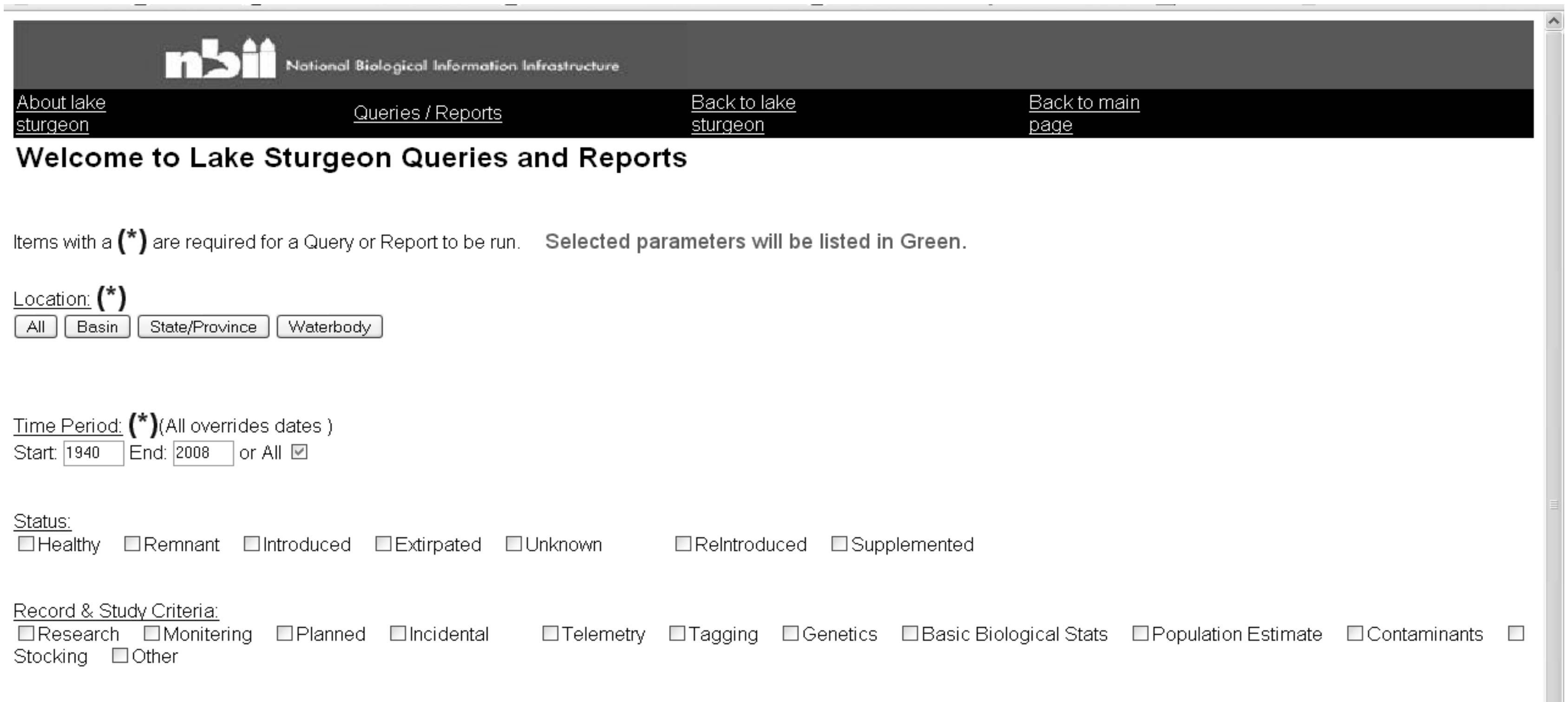


Figure 8. Lake sturgeon query page. At the top a user must select a location, which can be all locations, basin, state, province, or waterbody name. The user can specify a time period. The user can also select a status type, present, healthy, remnant, extirpated, unknown, and can narrow their search by reintroduced, introduced and supplemented populations. The user can also query by research type, life history, management, or additional record information.

Welcome to Display of Lake Sturgeon Queries and Reports

Map It ! | Export Data | Return to Query

ID	Waterbody	Status	Start	End	Research Type
255	Millecoquins River		2002		
256	Manistique River		2002		
257	Indian Lake		2002		
39	Fox River		1954	2003	
50	St. Joseph River		1995	2001	
16	Big Manistique Lake		1995	2001	
17	Manistique River		1995	2001	

Figure 9. Table with a partial listing of records that met criteria Basin = Michigan, Status = remnant. Note that each record ID is hyperlinked to its metadata. At the top of the table the user has the option to see his or her query returned as a map.

Lake sturgeon status and trends according to the SII

The SII database contains 242 records of presence information of which 34 (14%) are historical records. Historical information was defined as information collected prior to or during 1975 and current information was defined as after 1975 (J. Crossman, Michigan State University, personal communication). Prior to 1975, the SII contains 34 records of geographically distinct populations of lake sturgeon occurring in 16 tributaries to the Great Lakes basin (Figure 10, Table 5). Lake sturgeon populations in and around the Great Lakes region were estimated to number in the hundreds of thousands, if not millions (Tody 1974, Kinietz 1965, Slade and Auer 1997). Therefore it is important to note that the small number of tributaries reported by the SII is more reflective of a lack of historical information available to the database than the true amount of tributaries supporting historical sturgeon populations.

According to the SII post 1975 there are 208 records of geographically distinct populations of lake sturgeon occurring in 160 tributaries or landlocked bodies of water to the Great Lakes basin (Figure 11). In Lake Michigan there are 23 tributaries supporting populations of lake sturgeon, and in Lake Superior there are 13 tributaries or inland lakes that are currently supporting lake sturgeon populations (Table 6). In Lake Huron there are 19 tributaries currently supporting lake sturgeon populations. In both Lakes Erie and Ontario there are 5 tributaries or inland lakes that are currently supporting lake sturgeon populations. (Table7). In the St. Lawrence River there are 14 tributaries that are supporting lake sturgeon populations (Table 8).

Figure 10. Distribution of lake sturgeon from before 1975 according to records in the SII.

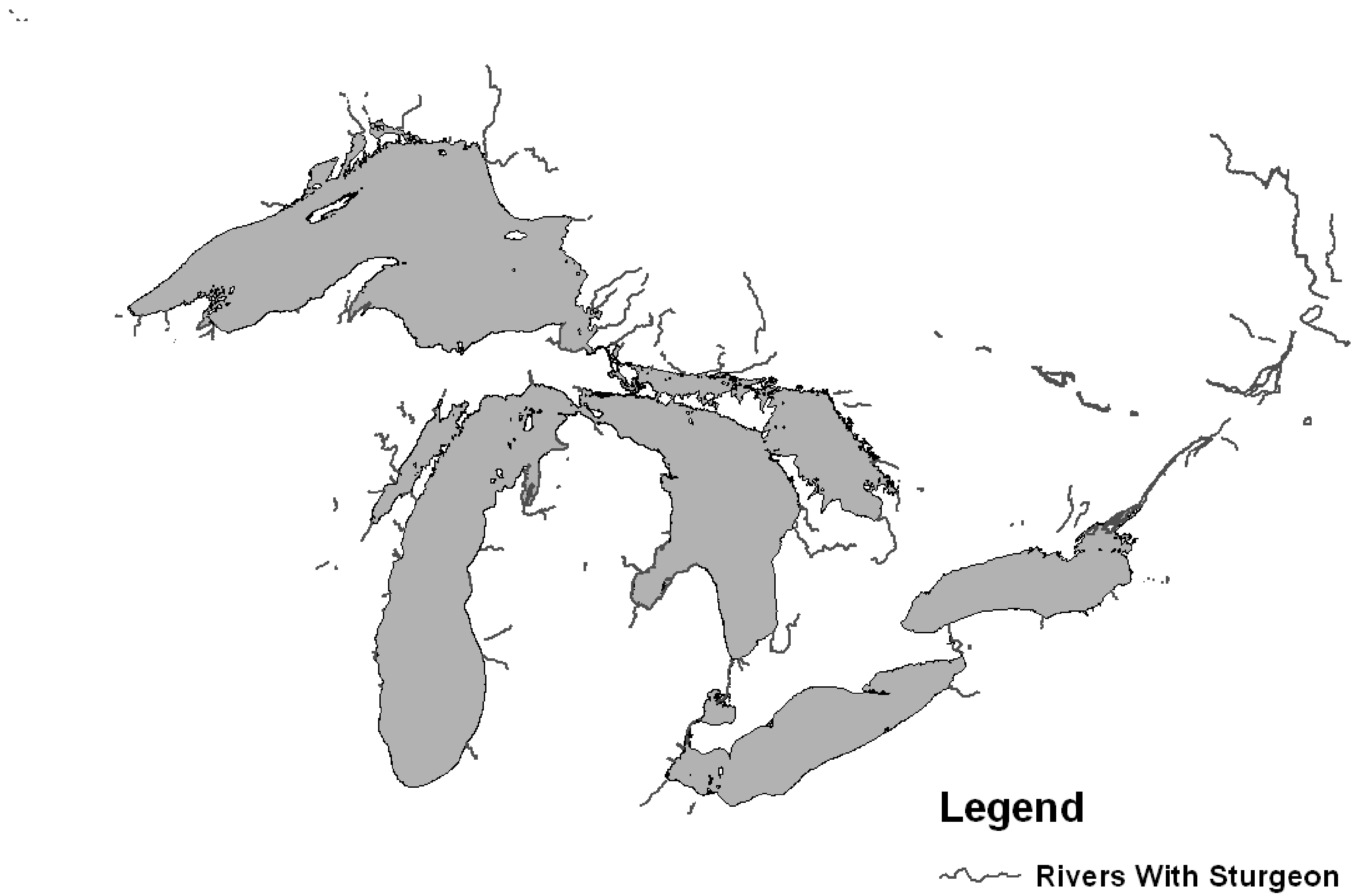


Figure 11. Current distribution of lake sturgeon according to records in the SII.

Table 5. Names of basins and waterbodies that had historical populations of lake sturgeon.

Basin	Waterbody name
Michigan	Fox River
	Muskegon River
	Menominee River
	Oconto River
	Peshtigo River
Huron	
	Garden River
	Mississagi River
	Thessalon River
Superior	
	Amnicon River
	Bad River
	Black Sturgeon River
	Goulais River
	Iron River
	Michipicoten River
	Ontonagon River
	Sturgeon River

Table 6. Names of basins and waterbodies in Lakes Michigan and Superior that have current populations of lake sturgeon.

Basin	Waterbody name
Michigan	Bear Creek
	Big Manistique Lake
	Cedar River
	Fox River
	Grand River
	Indian Lake
	Kalamazoo River
	Green Bay
	Ludington Shoal
	Manistee River
	Manistique River
	Menominee River
	Millecoquins River
	Muskegon River
	Oconto River
	Pere Marquette River
	Peshtigo River
	Pike River
	Sheboygan River
	St. Joseph River
	St. Joseph Shoal
	Sturgeon Bay
	Wolf River
Superior	
	Bad River
	Batchawana River
	Black Sturgeon River
	Chippewa River
	Goulais River
	Kaministiquia River
	Nipigon River
	Ontonagon River
	Pic River
	Pigeon River
	St Louis River
	Sturgeon River
	White River

Table 7. Names of basins and waterbodies for Lakes Huron, Erie and Ontario that have current populations of lake sturgeon.

Basin	Waterbody name
Huron	Black Lake
	Blue Point
	Burt Lake
	Carp River
	Cheboygan River
	French River
	Garden River
	Magnetawan River
	Mississagi River
	Moon River
	Mullett Lake
	Naiscoot River,
	Nottawasaga River
	Rifle River
	Saginaw River
	Sauble River
	Severn River
	St. Marys River
	Thessalon River.
Erie	
	Detroit River
	Eastern Basin
	Lake St. Clair
	St. Clair River
	Upper Niagara River
Ontario	
	Black River
	Genesee River
	Niagara River
	Oneida/Cayuga Lakes
	Trent River

Table 8. Names of basins and waterbodies in the St. Lawrence River that have current populations of lake sturgeon.

Basin	Waterbody name
St. Lawrence River	Batiscan River
	Black Lake
	Des Milles Iles River
	Des Prairies River
	Detroit River
	Grasse River
	Lake Champlain
	L'Assomption River
	Oswegatchie River
	Ottawa River
	Raquette River
	St. Francois River
	St. Lawrence River
	St. Maurice River

Status – classification

SII provides a current snapshot of the status of sturgeon populations in the Great Lakes region (Figure 12). There are 8 tributaries recorded in the SII that have healthy populations of lake sturgeon: Wolf River, in the Lake Winnebago watershed, the Des Prairies, St. Maurice, and St. Lawrence Rivers, and Ottawa River's Allumette Lake and Lac Coulonge. There are 63 tributaries with remnant populations of lake sturgeon (Tables 9 and 10). Lake sturgeon have been extirpated from 43 tributaries (Tables 11 and 12). Additionally, there are 17 tributaries where lake sturgeon are present but their status is unknown (Table 13).

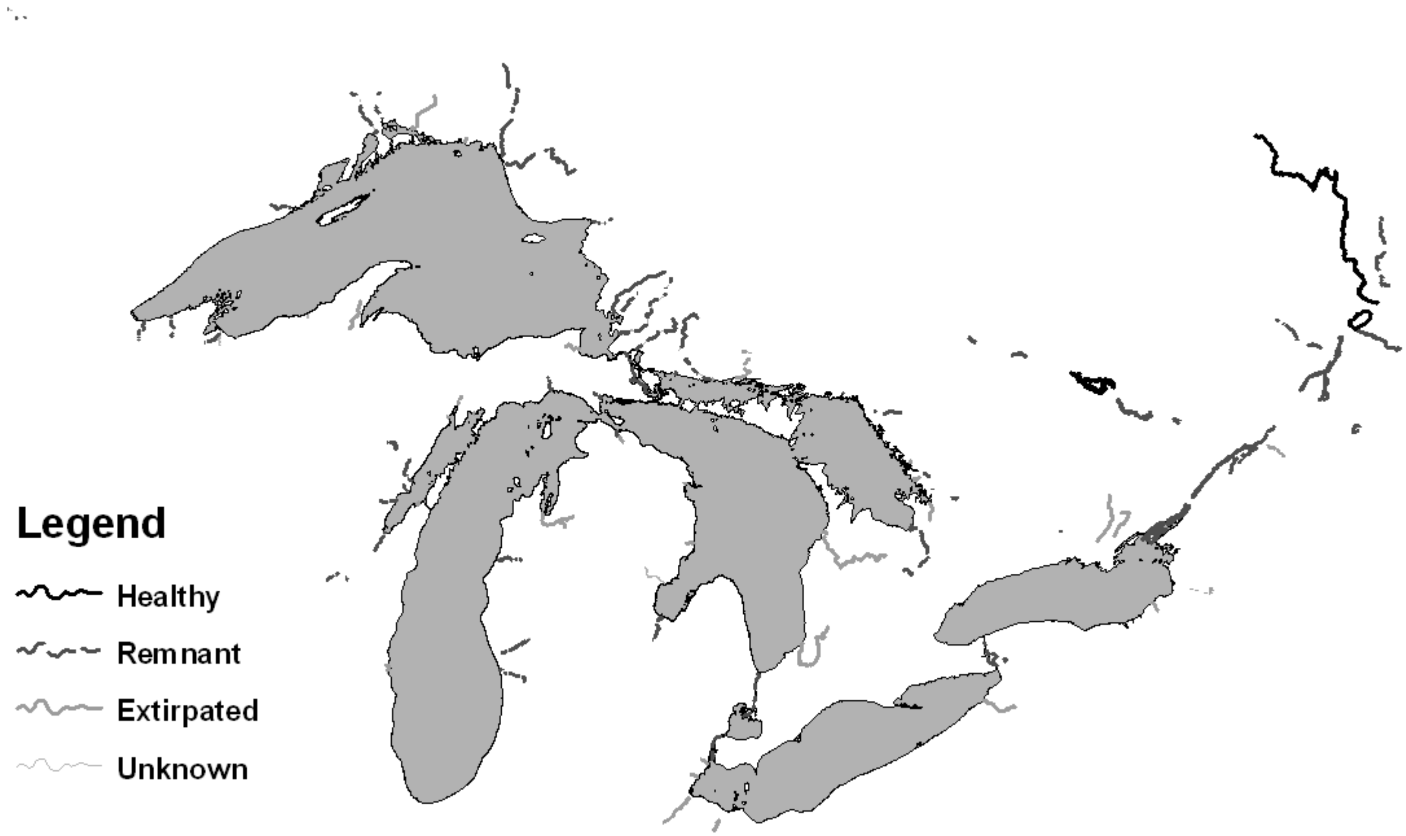


Figure 12. Status of lake sturgeon populations in the Great Lakes basin. Healthy populations are defined as having spawning populations of adults that range from 1,000 – 10,000 individuals. Remnant populations are defined as having spawning populations of adults that range from 10 – 1,000 individuals. Extirpated populations are defined as having spawning populations of less than 10 individuals. Unknown populations have unknown amounts of spawning adults.

Table 9. Names of basins and waterbodies with remnant populations of lake sturgeon for Lakes Michigan and Superior. Remnant populations are defined as having spawning populations of adults that range from 10 – 1,000 individuals.

Basin	Waterbody name
Michigan	Big Manistique Lake
	Black Lake
	Fox River
	Grand River
	Indian Lake
	Kalamazoo River
	Green Bay
	Manistee River
	Manistique River
	Menominee River
	Millecoquins River
	Muskegon River
	Oconto River
	Peshtigo River
	Pike River
	St. Joseph River
Superior	
	Bad River
	Batchawana River
	Black Sturgeon River
	Chippewa River
	Goulais River
	Kaministiquia River
	Michipicoten River
	Nipigon River
	Pic River
	Pigeon River
	St. Louis River
	Sturgeon River
	White River

Table 10. Names of basins and waterbodies with remnant populations of lake sturgeon for Lakes Huron, Erie, and Ontario and the St. Lawrence River. Remnant populations are defined as having spawning populations of adults that range from 10 – 1,000 individuals.

Basin	Waterbody name
Huron	Black Lake
	Burt Lake
	Carp River
	Cheboygan River
	French River
	Garden River
	Magnetawan River
	Mississagi River
	Moon River
	Mullett Lake
	Naiscoot River
	Nottawasaga River
	Saginaw River
	Severn River
	St. Marys River
	Thessalon River
Erie	
	Detroit River
	Eastern basin
	Lake St. Clair
	St. Clair River
	Upper Niagara River
Ontario	
	Black River
	Niagara River
	Trent River
St. Lawrence River	
	Batiscan River
	Des Milles Iles River
	Detroit River
	Grasse River
	Lake Champlain
	L'Assomption River
	Ottawa River
	Raquette River
	St. Francois River
	St. Lawrence River
	St. Regis River

Table 11. Names of basins and waterbodies with extirpated populations of lake sturgeon for Lakes Michigan, Superior and Huron. Extirpated populations are defined as having spawning populations of adults that are less than 10 individuals.

Basin	Waterbody name
Michigan	Barr Creek
	Boardman River
	Chicago Reef complex
	East/West Twin Rivers
	Escanaba River
	Kewaunee River
	Manitowoc River
	Menominee River
	Milwaukee River
	Root River
	Sturgeon Bay
	Sturgeon River
	Whitefish River
	Wolf River
Superior	
	Gravel River
	Harmony River
	Montreal River
	Ontonagon River
	Prairie River
	Stokely Creek
	Tahquamenon River
	Wolf River
Huron	
	Ausable River
	Black River
	Blind River
	Go Home River
	Manitou River
	Root River
	Saugeen River
	Seguin River
	Serpent River
	Sturgeon River
	Thunder Bay River

Table 12. Names of basins and waterbodies with extirpated populations of lake sturgeon for Lakes Erie, and Ontario and the St. Lawrence River. Extirpated populations are defined as having spawning populations of adults that are less than 10 individuals.

Basin	Waterbody name
Erie	Cattaraugus Creek
	Huron River
	Maumee River
	Raisin River
	Sandusky River
Ontario	
	Genesee River
	Napanee River
	Oswego River
	Salmon River
St. Lawrence	
	Salmon River

Table 13. Names of basins and waterbodies with unknown populations of lake sturgeon for Lakes Michigan, Superior, Huron, Erie, and Ontario and the St. Lawrence River.

Basin	Waterbody name
Michigan	Bear Creek
	Cedar River
	Ludington Shoal
	Manitowoc River
	Pere Marquette River
	Sheboygan River
	St. Joseph Shoal
	White River
Huron	
	Blue Point
	Echo River
	Key River
	Koshkawong River
	Rifle River
	Sauble River
	Spanish River
Ontario	
	Amherst Island Shoal
St. Lawrence	
	Oswegatchie River

Status – numbers

SII also provides a population estimate or range of estimates in absolute numbers for lake sturgeon. The SII records 31 estimates of population abundances for 22 tributaries (Table 14) and 5 lakes (Table 15). All population estimates are from 2002 or later. Generally there were only one set of population estimates that exist for each tributary, however in the Fox and Menominee Rivers, state and federal agencies are both measuring population abundances.

The SII was unable to detect trends in lake sturgeon using historical information. However, the status classification scheme developed by Zollweg et al. had time built in, so SII could detect trends by using the extirpated category. Using the extirpated category SII inferred that at least 131 tributaries to the Great Lakes basin did at one time support lake sturgeon populations, and that currently at least 43 (32%) of those tributaries have become extirpated. Furthermore, of the 88 remaining tributaries supporting lake sturgeon populations only 8 (9%) tributaries have populations of greater than 1000 adult spawners, 63 (72%) have populations of less then 1000 adult spawners, and the remaining 17 (19%), are unknown. Using a large scale composite system to track status and trends requires some knowledge of what research is and has been done. However, tracking down that information requires resources. Therefore, the SII also contains information on past research projects (Table 16) and ongoing research projects (Table 17).

Table 14. Population estimates or ranges of estimates for Great Lakes basin from United States Fish and Wildlife Service (USFWS), Enviro-Science Inc., Wisconsin Department of Natural Resources (WDNR), New York State Department of Environmental Conservation (NYDEC), Ontario Ministry of Natural Resources (OMNR), Central Michigan University (CMU), Michigan Technical University, Michigan Department of Natural Resources (MNDR), Consumer's Energy Environmental Department, University of Georgia, Purdue University, and unknown sources.

River Name	Population Estimate	Source
Bad	250	USFWS
Des Prairies	7000	Enviro-Science Inc.
Detroit	50-150	Enviro-Science Inc.
Fox	200-300	WDNR
Fox	100-200	USFWS
Grasse	10-20	NYDEC
Kaministiquia	140-175	ONMR
L'Assomption	50-150	Enviro-Science Inc.
Manistee	1-50	CMU
Manistique	1-50	Michigan Tech.
Menominee	200	WDNR
Menominee	200-1000	USFWS
Millecoquins	< 10	MDNR
Mississagi	150	Consumer's Energy
Muskegon	1-25	University of Georgia
Oconto	1-50	USFWS
Peshtigo	1-200	Purdue University
St. Francois	100	Enviro-Science Inc.
St. Lawrence	100-200	OMNR
St. Maurice	1250	Enviro-Science Inc.
St. Regis	1-100	Unknown
Sturgeon	200	Michigan Tech
White	15-1000	USFWS
Wolf	22000	WDNR

Table 15. Population estimates or ranges of estimates for Great Lakes basin from Ontario Ministry of Natural Resources (ONMR), Michigan Department of Natural Resources (MDNR), and the Service de l'amenagement et de l'exploitation de la faune.

Lake Name	Population Estimate	Source
Allumette	10,000	OMNR
Black	60	MNDR
Lac Coulonge	10,000	OMNR
Lac St. Pierre	10,000	Service de l'amengement

Table 16. Completed research projects carried out by agencies across the Great Lakes basin. Research categories include: Basic biological data collection (BB), population estimate (PE), telemetry (TL), tagging (TG), genetics (G), contaminant studies (C), and other scientific studies (SS). DNR stands for Department of Natural Resources and USFWS stands for United States Fish and Wildlife Service.

Agency Name	Research Type
Central Michigan University	BB, PE, TL, TG, G, O
Cornell University	BB, TG, O
Department Fisheries Oceans- Canada	BB, PE, O
Enviro-Science	BB, PE
Great Lakes Indian Fish and Wildlife Commission	BB, TL, O
Little River Band of Ottawa Indians	PE, TL, TG, C, O
Michigan DNR – Marquette Fisheries Station	BB, PE, TL, C
Michigan DNR – Mt. Clemens Fisheries Station	BB, PE, TL, TG, G, C
Michigan Technological University	BB, PE, TL, TG, G, O
New York State Department of Conservation	BB, PE, TL, TG, G, O
Ontario Ministry of Natural Resources	BB, PE, TL, TG, G, C, O
Purdue University	BB, PE, TL, TG, G, C, O
SUNY College of New York	BB, PE, TL, TG, G, O
USFWS – Alpena Fisheries Resource Office	BB, PE, TL, TG, G, C
USFWS – Ashland Fisheries Resources Office	BB, PE, TL, TG, G, C, O
USFWS- Marquette Biological Station	BB, PE, TG, G, O
USFWS – Green Bay Fishery Resource Office	BB, PE, TL, TG, G, C, O
University of Georgia	BB, PE, TL, TG, G, O
United States Army Corps of Engineers	BB, TL, TG, O
United States Geological Survey	BB, PE, TG, C, O
Vermont Department of Fisheries and Wildlife	BB, O
Wisconsin DNR	BB, PE, TL, TG, G, O

Table 17. Ongoing research projects carried out by agencies across the Great Lakes basin. Research categories include: Basic biological data collection (BB), population estimate (PE), telemetry (TL), tagging (TG), genetics (G), contaminant studies (C), and other scientific studies (SS). DNR stands for Department of Natural Resources and USFWS stands for United States Fish and Wildlife Service.

Agency Name	Research Type
Central Michigan University	BB, TL, TG, G, O
Department Fisheries Oceans- Canada	BB, PE
Enviro-Science	BB, PE
Fon Du Lac Band	PE
Grand Portage Band	PE
Grand Portage Chippewa Resource Program	PE
Great Lakes Indian Fish and Wildlife Commission	PE
Little River Band of Ottawa Indians	TL, TG, G, C, O
Michigan DNR – Marquette Fisheries Station	BB, PE, TL, G, C, O
Michigan DNR – Mt. Clemens Fisheries Station	BB, TG
Michigan State University	G
Michigan Technological University	BB, PE, TG, G, O
New York State Department of Conservation	BB, PE, TG, G, O
Ontario Ministry of Natural Resources	BB, PE, TG, G, C
Service de l'amenagement et de l'exploitation de la faune	PE
SUNY College of New York	BB
USFWS – Alpena Fisheries Resource Office	BB, TG
USFWS – Ashland Fisheries Resources Office	BB, PE, TG
USFWS – Green Bay Fishery Resource Office	BB, PE, TG, G, C, O
USFWS- Marquette Biological Station	BB, PE, TG
United States Geological Survey	BB, PE, TG, C, O
Vermont Department of Fisheries and Wildlife	BB, PE, G, O
Wisconsin DNR	BB, PE, TL, TG, G, O

Results of SII evaluation

The steering committee agreed the SII was useful because of the following reasons: it provide a quick synopsis of status, saving users time on searching for that information and it provided a centralized clearinghouse for organizing and maintaining data on status and trends. The mapping application, with its visual display of status, is especially attractive and useful to many types of users. Also, SII helps to identify who is doing what research, and where research is taking place, which means less time searching for available datasets. SII provided a meaningful way to discuss conservation and status of individual species by standardizing available information into a common reference scheme, while creating options for tracking the products of a specific agency, as SII tracks and compiles data from a variety of sources, not just a single source.

The development of the SII also provided valuable experience that can be used towards other data compilation or sharing projects, such as documented awareness that there unwillingness amongst scientists to contribute data to a composite database, uncertainty amongst database administrators on how to grant access to composite datasets that have already been compiled, awareness of the extent of which there is a lack of documented historical information about status and trends as well as a lack of metadata about many biological records. The SII can also be used as model when developing similar products as it helps developers visualize how status and trends information can be displayed and organized for management purposes.

The committee suggested that the SII needed the following improvements, most of which have already been implemented: species management plans should be added to site content, species information should be deep-linked to FishBase – an online database that provides a lot of basic biological information about fish. The steering committee also agreed that the SII also needs to ultimately provide the ability to perform mapping by multiple layers instead of by a single layer and that metadata and links to original records of sampling should be included where possible. Additional comments included: methods for deriving status should be made clear to the user, users should have the choice of querying by a single year or by a specific time period, locations where harvest is taking place should be identified and harvest should be delineated by type. The committee felt as much information as possible should be included in the query results, including agency and quality of data, and users should ultimately be able to query by map instead of data fields.

The committee suggested that a way to capture peer-reviewed literature that is not status and trends information was missing. They agreed that a method for communicating to users where to get more information; e.g., link to Google Scholar with text query, and other sources would be valuable as an information tool. The committee, also wanted to see links to data citations from data results that direct people to the original data source and a way to query by congressional districts.

The committee concluded that a status report using information from the SII would be of limited use to scientists because of the lack of detail in the data, but that the citation and research information would be very helpful. They recommended that SII could be very helpful by allowing scientists to track the status of single rivers through time; e.g., watching an extirpated river become healthy after stocking. The committee recommended that SII would help a scientist when addressing public information queries, and would be useful to new scientists taking over data in positions where a predecessor has retired or left.

The committee agreed that SII information would be useful to policy makers as a decision making- tool to see where management efforts have been successful, as an important communication tool for answering questions from Congress, and as an "expose" or transparency to the decision making process. The committee agreed that SII would be useful to the public in terms of outreach, enabling the public to search research, conservation and management inquires. Finally the committee recommended SII would be especially helpful if there were ways to capture threats and/or provide “public knowledge” of threats.

The committee noted that SII, in its current form, didn’t provide appropriate access to data. They suggested that having limited access would only cause trouble when restricting access to data, and that at its current level of data, the SII should be accessible to everyone. The committee agreed that the useful characteristics of status and trends reports that are currently being used in research and decision making are reflected in capabilities of the SII in a coarse way, but suggested other metrics:– genetics, successful reproduction, etc as proxies for measuring status and trends as well. They also recommended that SII should include management goals –e.g., number of fish to reach “healthy” status; or a probability of extinction statement in order to provide a framework for assessing the meaning of status and trends. Ultimately the committee concluded that SII was a very useful prototype. Additional comments and suggestions were that the system may be better for short-lived species where status and trends were more evident over a shorter time frame, and that it must be both field tested, and presented to other species researchers in order to be more fully developed.

DISCUSSION

Despite spending hundreds of millions of dollars on environmental monitoring and research each year, the United States does not know the full extent or condition of our natural resources on a national basis, or how they are changing (Bricker and Ruggiero 1998). Based on SII's ability to say little about the status and trends of sturgeon, despite collecting and piecing together hundreds of research records from across the Great Lakes, a fundamental change in our approach to environmental monitoring and reporting is needed if we are to meet the challenges facing us in assuring conservation of our fish populations into the future. Based on SII development it appears that current federal programs are too piecemeal, intermittent, and short term to provide the long-term information needs for understanding status and trends of species across an entire range.

Organization is an emergent property for any complex system, and efforts like the SII are necessary in order to develop that organization. Creating the SII was far from simple and during its evaluation, while it was agreed that it a worthwhile endeavor, it's real value was that its development provided a roadmap for the types of obstacles a similar future project would face. While creating the SII, I faced two types of general obstacles: technical and human-related (Hale et al. 2003; Pinkerton 1999). Technical barriers were things such as lack of consistency in resources, technology, data collection methods, experimental design, data quality standards, and laboratory procedures that hindered efforts to collect and integrate data. Human obstacles included difficulty finding unpublished data and reluctance to share data by data generators.

Specific technical challenges faced when developing the SII included: a lack of historical datasets about distribution and population abundances of lake sturgeon, inconsistency in methodology for determining status and trends, a lack of status and trends data for populations of lake sturgeon outside the Great Lakes basin with a dearth of metadata or citations for data that was available, and a lack of available and standardized georeferenced hydrography information.

The lack of available historical data impeded the SII's ability to detect trends in populations because time is inherent in trends data. SII has only 7 records of sturgeon information from pre-1960 and 24 records from pre-1975. However, based on a few publications and anecdotal evidence, there is general consensus that sturgeon populations have declined much from their historic levels (Baker 2006). However, it is fundamental to the missions of natural resource agencies that products like the SII have the capability to capture historic data in order to effectively characterize trends. Unfortunately, finding that historical data is difficult, because, based on personal data searching experience, historical information is not available electronically and is rarely associated with metadata or citations.

In the Great Lakes Region, sturgeon researchers are currently compiling historic records for a rewrite of the Michigan Lake Sturgeon Rehabilitation Plan (Gary Whelan, Michigan Department of Natural Resources, personal communication). To compile these records they are combing through county records and noting all observations of sturgeon, even anecdotal evidence. While, this type of detailed search is an inappropriate use of time and resources for the current version of the SII, I do recommend that future species composite database projects make sure to incorporate the information reflected in species management plans, because they are a large source of historical information. Clearly, the general lack of ability to find historical data is evidence that a system like the SII is needed and can be used to contain and preserve current data sets to slow or prevent further data decay.

A lack of consistency in methodology needed for the determination of status and trends was the biggest constraint for the creation of the SII. For systems such as the SII to function smoothly and reliably there needs to be an integrated effort amongst data contributors towards common data standards, mutually supporting data collection, and common reporting/distribution formats. Unfortunately commonality is the exception among datasets used for the SII. For example, lake sturgeon population estimates in the SII reflect 6 types of different methodologies: survey catches, observation, harvest, occasional angler catch, and mark-recapture through electrofishing surveys or gillnet survey captures. Furthermore, the majority of population estimate information doesn't have a corresponding citation. Therefore while the SII can display these different population estimates, it can't guide users on how to compare population estimates derived from different methodologies. Ultimately, composite database projects can create common organization schemes that group data for searching and displaying, but they can't fundamentally *change* data so that they can be universally compared, because of the different assumptions and methodologies used to collect the original data.

Lake sturgeon trends could not be assessed using SII data, because the status classification used in the SII wasn't created until 2003, and is only available for lake sturgeon in the Great Lakes basin. Lake sturgeon in the Mississippi River, northern Ontario, and Quebec were not included in the classification scheme. Because the SII isn't a decision making tool, it can't classify lake sturgeon populations that haven't already been classified by biologists. Additionally lake sturgeon populations haven't been reclassified since the original classification scheme was developed, making detecting trends impossible. The scientist that spearheaded classifying lake sturgeon populations has changed positions and there are no immediate plans to update the scheme (Emily Zollweg, New York Department of Environmental Conservation, personal communication). If the SII is to stay relevant then once a classification scheme has been adopted, plans need to be put in place to keep it current. Without regular data updates, SII cannot fulfill USGS goals and objectives.

Another difficulty in creating a system based primarily on large unpublished datasets is that metadata documenting data collection is rare. In the SII's approximate 300 records only about 15% have citations associated with them. While state and federal agencies are increasingly requiring their data to be referenced, this practice isn't universal and the standards for documentation can vary largely between agencies. Therefore it is incumbent upon a system such as the SII to determine which and if documentation is appropriate and to ensure that each record incorporated meets those standards.

Lastly, the lack of available and standardized georeferenced hydrography information made the construction of the SII very cumbersome, because data layers had to manually appended. Because standardized datasets for North American are non-existent, this will be a problem for any composite database that seeks to track the status and trends of an aquatic species across its entire range, when that range extends outside of the United States. There are three potential solutions to this issue: the first is to simply make these projects United States based only. This option is reasonable because the agencies themselves are not international in nature and their mission statements declare their responsibility with the United States only. The second option is to collect what spatial datasets do exist internationally (sometimes at cost) and manually geo-reference and append spatial data to the overall geo-classification scheme. The last option would be to engage international governments to create a common geo-classification scheme.

Human-related barriers included identifying pertinent status and trends data and collecting those from diverse sources. Unfortunately, the process of data discovery and data collection is challenging and poorly understood (Hale et al. 2003). Identifying existing data sets is very difficult. No library of data sets exists where a researcher can simply look through a card catalog and pick the data sets that are most appropriate to address an analytical question. Most of these data sets exist at diverse locations and are not public knowledge (Hale et al. 2003). For a scientist to be able to decide if a particular dataset is appropriate for answering a research topic, the researcher must address such questions as: what are the data, who has permission to release these data, who maintains these data, how were these data collected, and what is the spatial and temporal extent of these data? These questions can be hard to get definitive answers to because even if data can be found many data generators are also reluctant to share information about the state of their datasets (Porter 2004).

Identifying available data sets to incorporate into the SII was a very time consuming process, it took approximately one year to identify pertinent datasets. That time was spent searching through peer-reviewed literature, grey literature, government reports, online databases and unpublished datasets, speaking to scientists, attending meetings and by sending surveys to fishery professionals across the Great Lakes basin. Finding out information on who was performing research and how the research was being done was next to impossible when that research remained unpublished. In fact, only by word of mouth did that information become accessible. However, there is a need to bridge the gap between what research takes place and what research is reported. If work is to further take place on standardizing the sampling and analysis of aquatic populations, then laying the groundwork for what research takes place to begin with is absolutely fundamental. In recognition that the search process is long and cumbersome, the SII includes information about historical and ongoing research projects, because they serve as a proxy for identifying pertinent datasets. Integrating that information alone makes SII a very valuable asset for tracking and identifying relevant datasets.

Once a researcher has identified a dataset, access can be difficult because the data generator often will not share the data (Porter 2004; Hale et al. 2003; Pinkerton 1999). The results of the surveys sent to sturgeon researchers across the Great Lakes basin indicate that at least 40% of scientists were unwilling to contribute their data to a common database. However, not all unwillingness to share data is equal, reluctance to share data due to fears of data misuse or being “scooped” by other scientists require redress differently than reasons such as limited time or money. In one conversation with a survey participant, he informed me that at his annual review he was not rewarded for sharing data, he was rewarded for publishing papers, clearly a disincentive for scientists to collaborate to a common database.

If scientists are rewarded for publishing individual manuscripts, and furthermore perceive that by contributing data to a composite database their research might get scooped, then sharing data presents a critical obstacle to building composite databases. A tool that can be used to address these issues is the data sharing agreement, which outlines how the data will be managed and accessible to others, before any data is ever contributed. The fact that the federal agencies are spearheading attempts to consolidate and standardize data, but that scientists at the ground are unwilling to share data, suggests improved communication from top-down federal agency administration as one possible intervention point. Another obvious solution is for federal agencies to consider moving beyond the traditional reward system for publishing papers towards a reward system that includes acknowledgement to scientists for contributing to composite databases.

Developing the SII, despite its obstacles, was ultimately a meaningful endeavor, because the topic of status and trends is both relevant and topical. Attempts to consolidate and display information about the status of a species gets at the core of a very simple, intuitive, and relevant public concern about biodiversity and the state of our natural resources. Most people outside the science world do not have the ability to generate or gather this information, but because natural resources are communally owned, it is the responsibility of natural resources agencies to make basic biological information about resources accessible to any member of the community. Additionally, understanding the state and change of our nations fishery assets is imperative if management is to address changes happening beyond a local or regional scale.

Development of the SII also probes at a number of underlying systemic issues relating to natural resource management within the United States. In particular fundamental issues associated with addressing large scale research questions which included: resistance to sharing data, how and if agencies cooperate and communicate with each other and with their own employees, how to preserve and use historical datasets, the general lack of biological standardization, and assessing if the creation of these databases yields returns enough to justify their investment in resources.

Given the worth of creating this type of database, my recommendation would be to maintain its content. There are few large-scale successes in the fisheries world that that can be pointed to as an example for the worth of databases, but this project has the potential to be the success by which others are defined. I would recommend that more than one individual be assigned to the project. By their very nature, these projects are collaborative and having a sole individual responsible for the entire process is unrealistic. I would also recommend that if further development is to take place, then understanding the research for a species should take place prior to compiling information on the status of a species. I think that by understanding the research one begins to understand what the research is able to tell us about the species.

USGS is the appropriate agency to develop and manage a national fisheries database, and on-line library of fisheries datasets, because USGS has dual role of science generator and user. But USGS needs to examine ways to improve efficiency (reduce fragmentation) in federal fisheries research, reduce the duplication of effort in data collection, and spearhead efforts to standardize fishery data standards at a national level in order to adequately transfer scientific information to all of its stakeholders and policy makers. Fish don't acknowledge management boundaries and if we manage them by our own jurisdictional narrow goals, rather than by their natural distributions, then we can never expect to fully characterize their status or understand their trends. This can only be accomplished if natural resource agencies must move towards a common goal of sharing, documenting and ultimately standardizing how fisheries data is collected and reported.

APPENDIX A

APPENDIX A

UCRIHS- approved study consent form

MICHIGAN STATE
UNIVERSITY

OFFICE OF REGULATORY AFFAIRS
Human Research Protection Programs

BIOMEDICAL & HEALTH INSTITUTIONAL REVIEW BOARD (BIRB)

COMMUNITY RESEARCH INSTITUTIONAL REVIEW BOARD (CRIRB)

SOCIAL SCIENCE/ BEHAVIORAL / EDUCATION INSTITUTIONAL REVIEW BOARD (SIRB)

202 Olds Hall
East Lansing, Michigan
48824-1046
517-355-2180
Fax: 517-432-4503

www.humanresearch.msu.edu
SIRB & BIRB: IRB@msu.edu
CRIRB: crirb@msu.edu

February 14, 2008

Tracy Kolb
41 Natural Resources

Dear Ms. Kolb:

In May 2007, you submitted an initial IRB application for your study titled "Developing an Information Infrastructure for North American Sturgeon." At that time, it was determined that your study did not need IRB review and approval and an e-mail was sent to you and Dr. Taylor indicating this determination. Recently, you contacted our office and requested a letter to present to the Graduate School. Thus, I am writing to confirm that it has been determined that your project does not meet the definition of "human subject" as defined by the DHHS Federal Regulations. "Human Subject" means "a living individual about whom an investigator (whether professional or student) conducting research obtains: (1) Data through intervention or interaction with the individual, or (2) Identifiable private information." [45 CFR 46.102(f)]. Accordingly, the federal regulations for the protection of human participants would not apply to your project and you do not need IRB approval to proceed.

While you are collecting data through interaction with the individual, you are not collecting information about that individual's personal attitudes, behaviors, or thoughts. You are collecting information through these individuals to learn more about the agency.

Therefore, your project does not need review by the Social Science/Behavioral/Education Institutional Review Board (IRB). Thank you for erring on the side of caution.

If you have any further questions regarding these comments, please do not hesitate to call me at 517-355-2180 x 228.

Sincerely,

Karalyn Burt
Administrator, Social Science/Behavioral/Education Institution Review Board (SIRB)
Human Research Protection Program

Cc: William Taylor
7 Natural Resources

APPENDIX B

APPENDIX B

Replication of 2004-2005 survey instrument

You are being asked to participate in a study on the state of research pertaining to North American sturgeon. The information that you provide is to assist the United States Geological Survey (USGS) in understanding how sturgeon status and trends data are managed and maintained.

Your completion of the survey is completely voluntary. You are free to not answer any question or to stop participating at any time. There are no risks or individual benefits associated with taking this survey. By completing this survey, you indicate your voluntary consent to participate in this study and have your answers included in the project data set.

If you have any questions about this study you may call or email me, Tracy Kolb, at 517.281.8722, kolbtrac@msu.edu, or my advisor William W. Taylor, at 517.353.0647, taylorw@msu.edu. If you have any questions or concerns regarding your rights as a study participant, you may contact Peter Vasilenko, PhD, Director of Human Subject Protection Programs at Michigan State University (517.355.2180, Fax: 517.432.4503, irb@msu.edu , 202 Olds Hall, East Lansing 48824).

Please note that the purpose of this questionnaire is only to determine the availability and extent of data for North American sturgeon. Completion of this questionnaire does not obligate the respondent in any way.

1A. Name:

1B. Agency:

1C. Address and telephone:

1D. E-mail:

1E. How long have you been working for this agency?

1F. How long have you been working in fisheries?

1G. What is your position at the agency (circle/bold one):

Technician Biologist Supervisor Other (please specify

2A. What species of sturgeon do you collect information on?

2B. **Do you collect information on the following? If so, is this information publicly available at either a detailed level or as summarized information? Please circle/bold your response.**

Data type	Collect?		Available?		
1. Distribution	Yes	No	Detail	Summary	Not Available
2. Population Abundance	Yes	No	Detail	Summary	Not Available
3. Age composition	Yes	No	Detail	Summary	Not Available
4. Size composition	Yes	No	Detail	Summary	Not Available
5. Genetic information	Yes	No	Detail	Summary	Not Available
6. Telemetry information	Yes	No	Detail	Summary	Not Available
7. Specific catch locations	Yes	No	Detail	Summary	Not Available
8. Other	Yes	No	Detail	Summary	Not Available

If other, please describe:

Please use the following definition: a composite database stores multiple datasets that have overlapping data content, but whose original database structures are incompatible with one another. The composite system integrates and standardizes these data sets so that all of the data can be analyzed together.

3A. Have you been asked to contribute to a composite database before? (please circle/bold)

yes no

3B. If so, what were your expectations for the composite database?

3C. Were you satisfied with the result of your contribution? (Please circle/bold)

Very satisfied Neither satisfied nor dissatisfied Dissatisfied

4A. Would you be willing to contribute any of the sturgeon information listed above to a composite database (please specify)?

4B. What would you expect from your contribution to the composite database?

5A. Do you know of a composite sturgeon database? (please circle/bold) yes no
5B Do you use it? (please circle/bold) yes no
5C. Would you use it? (please circle/bold) yes no

6A. **Please rank the following as impediments to sharing the data mentioned above: Please circle/bold your answer.**

1. Time	Critical impediment	Minor Impediment	No Impediment
2. Money	Critical impediment	Minor Impediment	No Impediment
3. Fear of data misuse	Critical impediment	Minor Impediment	No Impediment
4. Fear of being "scooped"	Critical impediment	Minor Impediment	No Impediment
5. Fear of poachers accessing data	Critical impediment	Minor Impediment	No Impediment
6. Past experience with data contribution project	Critical impediment	Minor Impediment	No Impediment
7. Lack of control over data that you contribute	Critical impediment	Minor Impediment	No Impediment
8. Fear that your data are in poor condition	Critical impediment	Minor Impediment	No Impediment
9. Political reasons	Critical impediment	Minor Impediment	No Impediment
10. Legal reasons	Critical impediment	Minor Impediment	No Impediment
11. Other	Critical impediment	Minor Impediment	No Impediment

6B. If other, please describe:

6C. Do you have any suggestions for overcoming any of the impediments listed above?

APPENDIX C

APPENDIX C

Waterbody names in the SII

Amherst Island Shoal
Amnicon River
Ausable River
Bad River
Barr Creek
Batchawana River
Batiscan River
Bear Creek
Big Manistique Lake
Black Lake
Black River
Black Sturgeon River
Blind River
Blue Point
Boardman River
Bronte Creek
Burt Lake
Carp River
Cattaraugus Creek
Cayuga Lake
Cedar River
Cheboygan River
Chequamegon Bay
Chicago Reef complex
Chippewa River
Des Milles Iles River
Des Prairies River
Detroit River
East/West Twin Rivers
Eastern Basin
Echo River
Escanaba River
Fox River
French River
Garden River
Genesee River
Go Home River
Goulais River
Grand River
Grasse River
Gravel River
Great Lakes Basin
Lake Michigan -Green Bay
Harmony River
Huron River
Indian Lake
Indiana shoreline
Iron River
Kalamazoo River
Kaministiquia River
Kewaunee River
Key River
Koshkawong River
Lake Champlain
Lake Huron- Georgian Bay
Lake Huron- Main Basin
Lake Huron- North Channel
Lake Michigan -Grand Traverse Bay
Lake Michigan -Little Traverse Bay
Lake St. Clair
Lake Winnebago
L'Assomption River
Little Sturgeon Bay
Ludington Shoal
Magnetawan River
Manistee River
Manistique River
Manitou River
Manitowoc River
Maumee River
Menominee River
Michigan shoreline
Michipicoten River
Millecoquins River
Milwaukee River
Mississagi River
Montreal River
Moon River
Mullett Lake

Muskegon River
Naiscoot River
Napanee River
Niagara River
Nipigon River
Nottawasaga River
Oconto River
Oneida Lake
Oneida/Cayuga Lakes
Ontonagon River
Oswegatchie River
Oswego River
Otsego Lake
Ottawa River
Ottawa River -Allumette Lake
Ottawa River -Holden Lake
Ottawa River -Lac Coulonge
Ottawa River -Lac des Chats
Ottawa River -Lac des Deux Montagnes
Ottawa River -Lac Deschenes
Ottawa River -Lac Dollard des Ormeaux
Ottawa River -Lac Deux Rocher Fendu
Ottawa River -Lac la Cave
Ouareau River
Pere Marquette River
Peshtigo River
Pic River
Pigeon River
Pike River
Portage Lake
Prairie River
Raisin River
Raquette River
Rifle River
Root River
Saginaw Bay
Saginaw River
Salmon River
Sandusky River
Sauble River
Saugeen River
Seguin River
Serpent River
Severn River
Sheboygan River
Spanish River
St Louis River
St. Clair River
St. Francois River
St. Joseph River
St. Joseph Shoal
St. Lawrence River
St. Lawrence River- La Prairie Basin
St. Lawrence River -Lac St. Louis
St. Lawrence River -Lac St. Pierre
St. Lawrence River -Lake St. Francis
St. Lawrence River -Lake St. Lawrence
St. Lawrence River -Thousand Islands
St. Louis River
St. Marys River
St. Maurice River
St. Regis River
Stokely Creek
Sturgeon Bay
Sturgeon Bay area shoals
Sturgeon River
Tahquamenon River
Thessalon River
Thunder Bay River
Trent River
Upper Niagara River
western Keweenaw penninsula
western Wisconsin waters
White River
Whitefish Bay
Whitefish River
Wisconsin shoreline
Wolf River
Keweenaw Bay
Mississagi River (landlocked)

Standardized Data Names in the SII

Status - data type: Boolean

Present: documented records of lake sturgeon
Healthy: spawning populations from 1,000-10,000 individuals
Remnant: spawning populations from 10-1,000 individuals
Extirpated: spawning populations of less than 10 individuals
Unknown: unknown amounts of adults in spawning populations

Status – data type: number

Population estimate lower bound: estimated lower range of adult spawners
Population estimate upper bound: estimated upper range of adult spawners

Record and study criteria – data type Boolean

Research: record of research
Monitoring: record of monitoring
Planned: observation was planned
Incidental: observation was incidental
Telemetry: researchers placed telemetric tags on lake sturgeon to monitor movement
Tagging: researchers placed on lake sturgeon tags to monitor movement
Genetics: researchers are studying lake sturgeon genetics
Basic biological stats: researchers collected lake sturgeon morphological information
Population estimate: researchers determined a lake sturgeon population estimate
Contaminant: researchers determined the amount or source of different contaminants
Stocking: researchers are doing lake sturgeon stocking studies
Other: researchers collected other information such as associated species, habitat information etc.

Life Stages Observed – data type Boolean

Spawning: spawning lake sturgeon were observed
Larvae: larval lake sturgeon were observed
Juveniles: juvenile lake sturgeon were observed (0-5 yrs of age)
Subadult: subadult lake sturgeon were observed (5-15 yrs of age)
Juvenile lake sturgeon were observed (0-5 yrs of age**)**
Adult: adult lake sturgeon were observed (> 15 yrs of age)

Management – data type Boolean
Tribal harvest: harvest of lake sturgeon allocated to tribes at study location
Recreational harvest: lake sturgeon are harvested recreationally at study location
Commercial harvest: lake sturgeon are harvested by a commercial fishery at study location

Search modifiers – data type Boolean

Reintroduced: lake sturgeon populations have been stocked at study site

Reproductive success: lake sturgeon populations are self-sustaining at study site

Endemic: lake sturgeon populations are endemic to study site

WORKS CITED

Auer, N. A. 1999. Population characteristics and movements of lake sturgeon in the Sturgeon River and Lake Superior. Journal of Great Lakes Research 25:282–293.

Baker, E. A. 2006. Lake sturgeon distribution and status in Michigan, 1996-2005. Michigan Department of Natural Resources, Fisheries Technical Report 2006-4, Ann Arbor.

Baker, J.P. 1980. The distribution, ecology, and management of the lake sturgeon (Acipenser fulvescens Rafinesque) in Michigan. Michigan Department of Natural Resources, Fisheries Research Report No. 1883, Ann Arbor, Michigan.

Baker, K.S., S. J. Jackson, and J. R. Wanetick. 2005. Strategies supporting heterogeneous data and interdisciplinary collaboration: towards an ocean informatics environment. Proceedings of the 38th International Conference on System Sciences, pp 1-9.

Baker, K.S., B. J. Benson, D. L. Henshaw, D. Blodgett, J. H. Porter, and S. G. Stafford. 2000. Evolution of a multisite network information system: the LTER Information Management Paradigm. BioScience 50: 963-978.

Baker, K. S. 1998. Palmer LTER information management. In Michener, W.K., J. H. Porter, and S. G. Stafford (Editors). Data and information management in the ecological sciences: A resource guide. Long Term Ecological Research Network Office, University of New Mexico, Albuquerque, New Mexico, U.S.A.

Bricker, O.P. and M.A. Ruggiero. 1998. Toward a national program for monitoring environmental resources. Ecological Applications. 8(2):326-329.

Bruch, R.M. 1999. Management of lake sturgeon on the Winnebago system – long term impacts of harvest and regulations on population structure. Journal of Applied Ichthyology 15:142-152.

Brunt. J. W., P. McCartney, K. Baker, and S. G. Stafford. 2002. The future of ecoinformatics in long-term ecological research. Proceedings of the 6th World Conference on Systematics, Cybernetics, and Informatics. July 14th- 18th. Orlando, Florida.

Cook, J. G., R. P. Lanka, J. L. Henszey, M. L. Neighbours, K. H. Dueholm, K. D. Kozie and S. H. Anderson. 1987. Species Abstracts for Federal Rare Animals. Section 5, Appedix 2 in Literature reviews of rare species known to occur on National Park Service lands, Midwest region. National Park Service, Midwest Regional Office, Omaha, NE.

Dumont, P. R. Fortin, G. Desjardins, and M. Bernard. 1987. Biology and exploitation of lake sturgeon (Acipenser fulvescens) in the Quebec waters of the Saint-Laurent River. Pages 57-76 in Proceedings of a workshop on the lake sturgeon (Acipenser fulvescens). Ontario Ministry of Natural Resources Fisheries Technical Report Series No. 23, Toronto.

Hale, S. S., A. H. Miglarese, M. P. Bradley, T. J. Belton, L. D. Cooper, M. T. Frame, C. A. Friel, L. M. Harwell, R. E. King, W. K. Michner, D. T. Nicolson, and B. G. Peterjohn. 2003. Managing troubled data: coastal data partnerships smooth data integration. Environmental Monitoring and Assessment 81: 133-148.

Han, J., R. B. Altman, V. Kumar, H. Mannila, and D. Pregibon. 2002. Emerging scientific applications in data mining. Communications of the ACM 45: 54-58.

Harkness, W.J.K. and J.R. Dymond. 1961. The lake sturgeon. Ontario Department of Lands and Forests, Toronto, Ontario.

Hernandez, M.J. 1997. Database design for mere mortals: a hands on guide to relational database design. Addison-Wesley, New York.

Kinietz, W.V. 1965. The Indians of the western Great Lakes 1615-1760. Ann Arbor Paperbacks, The University of Michigan Press, Ann Arbor, Michigan. 427 pp.

McLaughlin, R. L., L. M. Carl, T. Middel, M. Ross, D. L. G. Noakes, D. Hayes and J. R. Baylis. 2001. Potentials and pitfalls of integrating data from diverse sources: lessons from a historical database. Fisheries 26: 14-23.

Naimen, R. J., J. J. Magnuson, D. M. McKnight, and J. A. Stanford. 1995. The freshwater imperative: a research agenda. Island Press, Washington, DC.

Pinkerton, E. 1999. Factors in overcoming barriers to implementing co-management in British Columbia salmon fisheries. Conservation Ecology 3: 1-13.

Porter J. H. 2004. A brief history of data sharing in the U.S. Long Term Research Network. DRAFT (don't have citation information)

Porter J. H. 1998. Scientific databases for environmental research. In Michener, W.K., J. H. Porter, and S. G. Stafford (Editors). Data and information management in the ecological sciences: A resource guide, pp. 41- 46. Long Term Ecological Research Network Office, University of New Mexico, Albuquerque, New Mexico, U.S.A.

Porter, J.H. and K. W. Ramsey. 2003. Integrating Ecological Data: Tools and Techniques.

Priegel, G.R., and T.L. Wirth. 1971. The lake sturgeon. Its life history, ecology, and management. Wisconsin Department of Natural Resources Publication No. 240. Madison.

Sandilands, A.P. 1987. Biology of the lake sturgeon (Acipenser fulvescens) in the Kenogami River, Ontario. Pages 33-46 in Proceedings of a workshop on the lake sturgeon (Acipenser fulvescens). Ontario Fisheries Technical Report Series No. 23, Toronto.

Slade, J.W. and N.A. Auer, editors. 1997. Status of lake sturgeon in Lake Superior. A report prepared for the Lake Superior Technical Committee. USFWS, Fishery Resources Office, Ashland, Wisconsin. 45 pp.

Tody, W. H. 1974. Whitefish, sturgeon, and the early Michigan commercial fishery. Pages 45-60 in Michigan Fisheries Centennial Report 1873-1973. Michigan Department of Natural Resources. Lansing, Michigan.

Watson, R., Kitchingman, A., Gelchu, A., and D. Pauly (2004) Mapping global fisheries: sharpening our focus Fish and Fisheries 5 (2) , 168–177.

Zollweg, E. C., editor, R. F. Elliott, T. D. Hill, H. R. Quinlan, E. Trometer, J.W. Weisser, N. A. Auer, E. A. Baker, D. Carlson, M. Holtgren, and M. Thomas. 2003. Great Lakes Lake Sturgeon Coordination Meeting: Proceedings of the December 11-12, 2002 Workshop, Sault Ste. Marie, Michigan. U.S. Fish and Wildlife Service. Amherst, New York. 130 pp.