

Open-Access Databases as Unprecedented Resources and Drivers of Cultural Change in Fisheries Science

Ryan A. McManamay

Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN

Ryan M. Utz

The National Ecological Observatory Network, 1685 38th St., Boulder, CO 80301. E-mail: utz.ryan@gmail.com

ABSTRACT: *Open-access databases with utility in fisheries science have grown exponentially in quantity and scope over the past decade, with profound impacts to our discipline. The management, distillation, and sharing of an exponentially growing stream of open-access data represents several fundamental challenges in fisheries science. Many of the currently available open-access resources may not be universally known among fisheries scientists. We therefore introduce many national- and global-scale open-access databases with applications in fisheries science and provide an example of how they can be harnessed to perform valuable analyses without additional field efforts. We also discuss how the development, maintenance, and utilization of open-access data are likely to pose technical, financial, and educational challenges to fisheries scientists. Such cultural implications that will coincide with the rapidly increasing availability of free data should compel the American Fisheries Society to actively address these problems now to help ease the forthcoming cultural transition.*

INTRODUCTION

The management, distillation, and sharing of an exponentially growing stream of data represents a fundamental challenge to fisheries science. Data across all subdisciplines of ecology are becoming available in unprecedented volumes due to advancements in computational technology and the rapid growth of resources with the explicit purpose of housing and providing data to all scientists (open access). Yet despite such trends, a lack of needed information related to fisheries management continues to be cited as a challenge (Pauly 1995; Crone and Tolstoy 2010; Olascoaga and Haller 2012). How would the fisheries science culture benefit if data sets behind all published articles or publicly funded grants were archived and maintained in accessible, online data warehouses? Such a theoretically attainable future would both benefit and pose challenges to our field.

Although many scientific problems benefit from additional data, the disparity between the growth in data availability and continued calls for more information could reflect several phenomena that the culture of fisheries science needs to address. Manipulating and managing large, complex databases requires

Bases de datos de acceso abierto como un recurso sin precedente y causante de cambio cultural en la ciencia pesquera

RESUMEN: *en la última década, el número de bases de datos de acceso abierto con utilidad para la ciencia pesquera ha crecido exponencialmente en cantidad y alcance y su impacto ha sido considerado como muy importante en esta disciplina. El manejo, depuración e intercambio de datos de acceso abierto representa retos fundamentales en la ciencia pesquera. Muchos de los recursos actualmente disponibles de acceso abierto pueden no ser conocidos por los científicos pesqueros. Por lo tanto, aquí se presentan varias bases de datos a nivel nacional e internacional de libre acceso con aplicación en las ciencias pesqueras y se da un ejemplo de cómo pueden ser aprovechadas para realizar valiosos análisis sin hacer esfuerzos adicionales de trabajo de campo. También se discute cómo el desarrollo, mantenimiento y uso de las base de datos de libre acceso muy posiblemente representarán retos importantes para los científicos de la pesca en cuanto a las dimensiones técnica, financiera y educativa. Tales implicaciones culturales, que coincidirán con la disponibilidad cada vez mayor de datos gratuitos, debieran servir de impulso a la Sociedad Americana de Pesquerías a que volcara activamente su atención sobre estos problemas con el fin de facilitar la transición cultural que se avecina.*

advanced technological skills that are beyond the capabilities of most fisheries scientists (Lynch 2008; Cukier 2010; Kolb et al. 2013), because structuring databases, queries, and exploration capabilities requires very specialized training (Fox and Hendler 2011), adequate funding (Tenopir et al. 2011), and dedicated staff (Kolb et al. 2013). In many cases, researchers may simply not know about data resources pertinent to their lines of inquiry. Perhaps most critical, many are uncomfortable with the concept of sharing data, even after projects are finished. Understandably, apprehension may exist due to fear of unacknowledged work, compromised intellectual property, or stolen research (Silver 2003). Such unease about contributing to open-access data inherently limits legitimate findings that may be drawn from those data sets. Ultimately, technological advancements and cultural evolution within the ecological sciences will substantially propel data sharing. Fisheries science must adapt accordingly as well.

Despite the increasing awareness of these needs and challenges arising from open-access databases (Silver 2003; Lynch 2008; Reichman et al. 2011), we have observed few

formal discussions on the matter among fisheries professionals. Symposia and data summits documenting the benefits and problems of open-access databases have occurred within the fisheries science community as early as 1980 (Pacific Marine Fisheries Commission 1980), and reports have highlighted the need for continual development of these resources, along with the associated challenges (Austen et al. 1998; Allen et al. 2006). Yet to our knowledge, a recently published article (Kolb et al. 2013) is one of the only contemporary descriptions of database management, standards, maintenance, and documentation in the fisheries-related peer-reviewed literature. During the 2012 Annual American Fisheries Society (AFS) Meeting in St. Paul, Minnesota, we organized a symposium entitled “Free Data: Opportunities in Open-Access Network Databases to Advance Spatiotemporal Scales of Inquiry in Fisheries Science.” The symposium attempted to provide a forum to acquaint the fisheries science community with open-access data systems. Presenters in the symposium exhibited programs offering unprecedented, nationwide fisheries data resources, many of which have already produced novel scientific discoveries and nearly all of which are rapidly expanding (see Table 1). However, as we have observed in the peer-reviewed literature, very little discussion involved the technical, financial, educational, or cultural obstacles to open-access data.

Open-access databases have and will continue to be developed and maintained by multiple institutions within the fisheries science community (e.g., Beard et al. 1998; Seeb et al. 2007; Frimpong and Angermeier 2009; Wang et al. 2011; Hamm 2012). Many fisheries professions on the frontier of database management have been well aware of these issues for some time (Geoghegan 1996; Kolb et al. 2013). However, the cultural problems associated with an increasingly open-access scientific community should be more rigorously addressed through formal discussions within the community to ease the burden of cultural transition. Below we discuss how open-access databases have already changed fisheries science and how they may continue to do so. We also provide examples of national- and international-scale open-access databases that many may not be aware of despite their ambitious scope and valuable data offered. Finally, we present some principle cultural implications that arise with the increasing availability of free data and challenge the AFS community to proactively address these problems.

HOW HAVE OPEN-ACCESS DATABASES CHANGED FISHERIES SCIENCE?

Over the past two decades, open-access databases have already significantly changed fisheries science as a discipline. Because of publicly available data, global-scale marine stock assessments are now commonplace (e.g., Costello et al. 2012; Ricard et al. 2012) thanks to open-access catch data (Sea Around Us Project 2013) and published marine ecosystem models. Costello et al. (2012) developed a novel approach to discern declining trends in fisheries lacking any formal assessment using publicly available data, including marine stock assessments (Ricard et al. 2012), trends in catch (Food and Agriculture Organization of the United Nations 2011), and fish life histories

(Froese and Pauly 2012). Likewise, open-access databases for freshwater fish have provided opportunities to assess large-scale (e.g., continental) patterns in fish ecology. For example, open-access riverine fish assemblage data (e.g., U.S. Geological Survey [USGS] 2013) and geospatial landscape coverage (e.g., U.S. Environmental Protection Agency 2012; Multi-resolution Land Characteristics Consortium 2013) provided resources to establish relationships between landscape predictive frameworks and fish communities (e.g., Frimpong and Angermeier 2010). Similarly, Mims et al. (2010) mapped the frequency of different fish life histories across North America using publicly disseminated fish distribution data (NatureServe 2004). Using freely available information on aquatic resources, Loftus and Flather (2012) examined emerging trends in aquatic habitat, fish populations, and both recreational and commercial fisheries across the United States to isolate regions requiring more intensive natural resource management by the U.S. Forest Service.

Until recently, only individuals who possessed large data sets could explicitly test such broad-scale questions. Modern open-access data repositories provide the prospect of large-scale, high-resolution research for everyone. Though extensive databases have provided the means to address big questions, they also have expanded the conceptual frameworks of scientific questions. Influxes of data can change (1) how scientists view natural phenomena (Nelson 2008), (2) the analytical approaches and predictive output of research (Luo et al. 2011), and (3) the speed and nature of hypotheses generation and testing (Luo et al. 2011). In addition, scientists taking advantage of open-access data need familiarity with statistical/database programs that best handle larger data sets and increase data mining efficacy (Reichman et al. 2011). Furthermore, big data, processes, and results are now being packaged as research products to promote future meta-analyses and support evidence-based research (Reichman et al. 2011).

Open-access databases are also facilitating scientific debate through unprecedented means. Data transparency allows findings to be validated or disputed repeatedly by different researchers to eventually arrive at consensus. Global-scale marine fisheries stock assessments offer a classic example of such debate. Worm et al. (2006) famously derived quantitative models to conclude that by 2048 marine fisheries resources would disappear. Because the authors applied open-access resources to arrive at this conclusion, others could access the same data sources but render different conclusions (Murawski et al. 2007). Some contend that the first assessment misapplied open-access resources through poor understanding of the data, but both studies were peer-reviewed in top-tier journals. We argue that scientific debates spurring from use of open-access resources is a positive trend, because the consensus of conclusions derived from the same data source will be strongest when subjected to validation by multiple researchers.

The increasing prevalence and awareness of open-access data has also changed the roles of fisheries professionals. For example, data repositories are increasingly developed and maintained by universities and smaller agencies with varying

Table 1. Examples of open-access databases pertinent to fisheries science. Listed examples are limited to established regional-, national-, and international-scale efforts.

Database	Description	Website
Biodiversity Information Serving Our Nation (BISON)	Synthesized and permanent repository of biological occurrence data for the United States from numerous distributed systems and formats. Supported by the Core Science Analytics and Synthesis (CSAS) program within the USGS.	bison.usgs.ornl.gov
Data Observation Network for Earth (DataONE)	NSF-supported cyber infrastructure for the preservation, access, use and reuse of multiscale, multidiscipline, and multinational environmental science data.	dataone.org
Dryad	International repository of data underlying peer-reviewed biosciences publications. Allows authors to upload data from their accepted work.	datadryad.org
FishBase	Relational database of 28,500 marine and freshwater fish species, including distribution, phenological characteristics, habitat preferences, physiological attributes, International Union for Conservation of Nature Red List status, and taxonomic information.	fishbase.org
FishTraits	Trophic attributes, reproductive ecology, habitat associations, and salinity/temperature tolerances for 809 native and exotic North American freshwater fish taxa.	fishtraits.info
Global Lake Ecological Observatory Network (GLEON)	Physical, ecological, and biogeochemical data on a global network of lake ecosystems supported by a grassroots network of scientists.	gleon.org
Long Term Ecological Research Network (LTER)	NSF-supported network of long-term ecological studies, including sites and programs in stream, lake, and marine ecosystems throughout North America and the South Pacific. Includes heterogeneous variables across sites.	lternet.org
Multistate Aquatic Resources Information System (MARIS)	Population estimate, total catch, total weight, and water quality records collected by state agencies. Currently includes data for nearly 600 fish species collected from >16,000 sites across 16 states.	marisdata.org
National Ecological Observatory Network (NEON)	NSF-supported continental-scale observatory planning to collect 30 years of data to gauge the effects of climate change, land use change, and invasive species on natural resources and biodiversity.	neoninc.org
National Fish Habitat Action Plan (NFHAP)	Nationwide database of fish habitat quality delineated by National Hydrography Data (NHD) plus catchments. Includes land use, dams, road crossings, and habitat quality metrics.	fishhabitat.org
National Gap Analysis Program	Data sets used to determine how much of an ecosystem type or a target species' habitat is currently in conservation areas. Data include land cover, predicted distributions of vertebrate species, and stewardship layers.	gapanalysis.usgs.gov
NatureServe	Nonprofit conservation organization initiated to provide scientific resources for effective conservation. Hosts a freshwater fish distribution database linked to HUC-8 USGS watershed codes.	natureserve.org
Ocean Biogeographic Information System (OBIS)	Integrated marine species presence/absence data sets from around the world. Currently offers 33.6 million records.	obis.org
Ocean Observatories Initiative (OOI)	NSF-supported observation platform planning to collect climate variability, ocean circulation, area-sea exchange, and seafloor process data in coastal and deep sea ecosystems for 25–30 years.	oceanobservatories.org
Pacific Fisheries Information Network (PacFin)	Provides detailed marine fisheries data, including trawl survey data, bycatch estimates, and age structure of target species, from fisheries offshore of Oregon, Washington, British Columbia, and Alaska. Supported by the National Marine Fisheries Service.	pacfin.psmfc.org
Dr. Ransom A. Myers (RAM) Legacy Stock Assessment Database	Compilation of stock assessment results for >200 commercially exploited populations of marine organisms from around the world.	ramlegacy.marinebiodiversity.ca/ram-legacy-stock-assessment-database
Standard Methods for Sampling North American Freshwater Fishes	Freshwater fish data collected using standardized sampling techniques. Allows users to compare their data with those collected using standardized methods.	fisheriesstandardsampling.org
StreamNet	Provides a wealth of biological and physicochemical data related to fisheries management in the Pacific Northwest, with emphasis on the Columbia River basin. Maintained by the Pacific States Marine Fisheries Commission.	streamnet.org
USGS BioData	Provides access to USGS-collected aquatic bioassessment data. Includes fish, macroinvertebrate, and algal community data, as well as physical habitat survey data from across the United States.	aquatic.biodata.usgs.gov

degrees of multidisciplinary services (Lynch 2008; Kolb et al. 2013). Databases, rather than analytical results and interpretation, are being funded as deliverables (Lynch 2008; Kolb et al. 2013) that have the potential to move research into new directions. The degree to which the availability of open-access databases has increased the efficiency and productivity of institutions is unclear, because the annual global growth rate of

publications has remained steady (Larsen and vonIns 2010). However, one downside of increasing efficiency may be elevated expectations of institutions on research staff productivity. Collaborations have been on the rise, with the mean number of authors per paper in the sciences more than doubling between 1954 and 1998 (Larsen and von Ins 2010). Data sharing very likely has provided collaborative opportunities within and among scientific disciplines.

OPEN-ACCESS DATABASES IN FISHERIES SCIENCE

Freely available databases applicable to fisheries science have grown exponentially in quantity and scope over the past decade. Table 1 lists a number of regional-, national-, international-scale open-access data resources and illustrates the diversity of accessible information. Many examples listed in Table 1 provide site-specific collection information with varying degrees of detail. For example, the Ocean Biogeographic Information System provides records on tens of thousands of marine species from around the world but is limited to presence/absence data, and the Multistate Aquatic Resources Information System (MARIS) posts state agency-derived data on freshwater fish collections, many of which include abundance, length, and weight (Figure 1). Similarly, USGS BioData provides data on fish, invertebrate, and algal community collections and physical habitat surveys across the United States (Figure 1). Other databases offer organism-specific information: FishTraits provides biological, ecological, and environmental tolerance parameters for more than 800 North American freshwater fish, whereas FishBase houses physiological, phenotypic, and distributional information on thousands of marine and freshwater fishes. Many state agencies and educational entities have begun to host state-specific data sharing portals as well. For instance, the Fishes of Texas Project (supported by the University of Texas at Austin; fishesoftexas.org) hosts thousands of records from throughout the state, some dating back to the mid-1800s, and the Iowa Department of Natural Resources (iowadnr.gov) posts databases on a wealth of aquatic ecosystems and assemblages.

An Example of Open-Access Database Utility

To illustrate how open-access data may facilitate novel approaches to detecting trends, we provide an example of mapping fish traits using a combination of spatial (e.g., GIS) and life history data derived entirely from open-access sources. Maps of fish traits across watersheds can be useful for establishing links between landscape properties and fish life histories (e.g., see Olden and Kennard 2010), spatially predicting potential ecological responses to landscape development, or prioritizing conservation efforts. Digitized maps of 865 freshwater fish distributions within eight-digit hydrologic catalog units (HUC-8) were assembled from NatureServe (NatureServe 2004). We compiled lists of all native fish species ($n = 731$) currently existing (within the last two decades of sampling) within each HUC-8. Fish trait information was accessed through the FishTraits database (Frimpong and Angermeier 2009).

For the sake of brevity, we focused on only two traits: potadromy/anadromy and nest-guarding spawning behavior. Potadromous and anadromous fish are species that migrate entirely within freshwater or migrate from saltwater, respectively, to complete their life history requirements (Moyle and Cech 2004). Nest-guarding fishes construct a cavity or pit in which eggs are laid, fertilized, and guarded until embryos hatch or larval stages are reached (Balon 1975). Because trait information for all species was incomplete due to insufficient biological

information on highly endemic and/or not formally described species, we accessed NatureServe Explorer, FishBase, literature searches, and general searches to update missing traits with new information or find the closest phylogenetic relative as a substitute. Closest phylogenetic relatives were either the nearest parental clade (subgenus), species of potential hybridization, or species commonly misidentified as the species of interest (in that order of preference). Within a GIS, we summarized the number of potadromous/anadromous fish species and the proportion of nest-guarders currently occurring within each HUC-8 and mapped the distribution of traits (Figure 2).

Potadromous/anadromous fish were more numerous in the Pacific Northwest, Great Lakes Region, Ohio, and Tennessee basins and several watersheds in the Northeast (Figure 2). The proportion of nest-guarding fishes per watershed was higher in the Midwest and showed increasing prevalence with decreasing latitude (Figure 2). Fish traits are advantageous in that they consolidate information across many species into concise groups that can be used to infer convergent adaptive strategies and common responses to disturbance (Frimpong and Angermeier 2010). Maps of trait frequencies can provide a geographical base for prioritizing restoration or preventative management actions. For example, watersheds with many migratory fish may be prioritized for fish passage enhancement, whereas those with higher nest-guarding frequencies should effectively maintain sensitive populations by limiting anthropogenic flow fluctuations. Our brief analysis shows that the availability of large databases can quickly and efficiently produce scientific findings.

CULTURAL IMPLICATIONS

Although open-access databases will continue to create unprecedented opportunities in fisheries science, challenges also accompany their promulgation and use. Many researchers have based their careers on relatively small spatiotemporal-scale projects constrained by the limitations of fieldwork. Consequently, the concept of open-access data remains foreign to many and this can lead to multiple problems. For instance, researchers remain reluctant to share their own data, open-access sources are often inadequately acknowledged or incorrectly cited, and resources required to maintain these systems often proves scant (Allen et al. 2006). Yet to address the broad-scale environmental problems impacting contemporary aquatic and marine ecosystems, future researchers will inevitably rely on data that they did not collect. A cultural shift that includes cognizance of how open-access data systems should be ethically used, supported, and expanded must ensue.

Full Participation

As researchers, we should recognize that the data we generate might prove valuable well beyond their originally intended use. The ability to share data has grown along with the expanding scope and number of open-access networks. Several resources mentioned in the preceding section (such as Dryad and the RAM Legacy Stock Assessment Database) offer

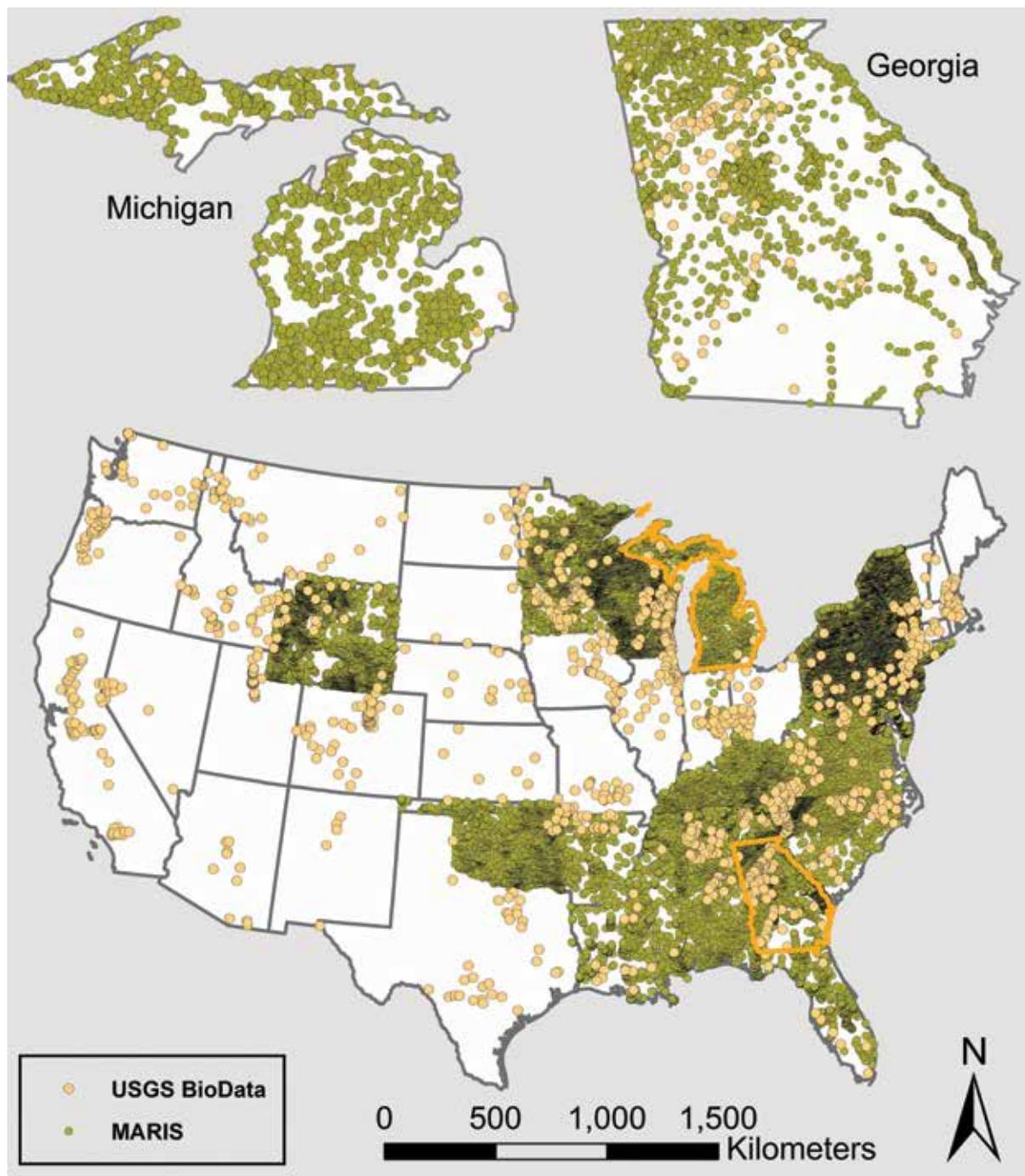


Figure 1. Distribution of fish sampling locations provided in USGS BioData and Multistate Aquatic Resources Information System (MARIS) open-access databases.

opportunities to upload data. Others, such as MARIS, offer a logistic framework for posting state agency-generated databases. Although many fisheries professionals remain understandably wary, the practice of data sharing should gain traction within our society. Most important, additional data improve the scope and inference capability of nearly all scientific endeavors and thus represent a substantial, fundamental value for the entire community. Yet despite the benefits of data sharing and avenues to help do so, an estimated 99% of ecological data remains inaccessible after publication (Reichman et al. 2011). Obviously, there are many cases in which data cannot or should not be shared, as in the case of sensitive information, ancillary data

not owned by the immediate authors, and data not supported by publication or documentation. However, for the most part, anxiety about data sharing should be allayed based on various reasons and awareness of incentives. Concerns that others may benefit from data at the personal expense of those who collected it can be easily preempted by retaining raw data from open-access sources until all planned publications have been accepted or by placing data in repositories requiring appropriate permission (Reichman et al. 2011). Institutions such as the National Science Foundation now expect greater data transparency as a condition for awarding grants and an increasing number of high impact publications (i.e., *Nature* and *PLoS Biology*) encourage

or require data sharing as a condition of manuscript acceptance. Finally, open-access data can facilitate novel means of professional dialog, such as online forums to debate findings. Such cultural evolution has been successfully implemented by several journals in the Public Library of Science (PLoS; plos.org) system.

Proper Citation, Acknowledgment, and Use

Ensuring that open-access data are consistently and properly acknowledged would significantly benefit both users and contributors. One major source of sensitivity toward sharing is the apprehension that data collectors may not be adequately recognized for their intellectual contribution (Silver 2003; Allen et al. 2006). Thus, open-access resources should not be disseminated unless supported by publication or technical documentation that provides proper acknowledgements. Additionally, proper data citation ensures that detailed sampling methodology can be tracked and understood without having to restate such information in studies that utilize the data. Nearly all open-access databases listed in Table 1 post citation and use guidelines to help those using data cite work appropriately. Authors, manuscript referees, and journal editors should consistently make certain that open-access sources are cited correctly. Best practices would also include a nod to the open-access database in the acknowledgement section of a manuscript.

Ensuring data quality and accuracy represents a major challenge associated with all open-access resources. Even if an investigator downloads the highest quality data possible, methodological misunderstandings could easily lead to spurious conclusions if the data were not suitable to address a particular question. One hypothetical example of data misuse could involve the MARIS data set (see Table 1). Though much of the data in the MARIS system are derived from agency sampling efforts for entire fish assemblages, many data sets within the system targeted select species, such as surveys for a sport fish of interest. Catches of nontarget species collected during such surveys can be reported in the data, even if the equipment and methodology used were not ideal for the nontarget species. An investigator interested in modeling the abundance or distribution of the hypothetical nontarget species would need to carefully consider whether or not to include such data. Analogous situations could be construed from any of the databases listed in Table 1.

Although it is the responsibility of investigators to understand the limitations of the data and apply it appropriately, the most commonly cited means to address the challenge is

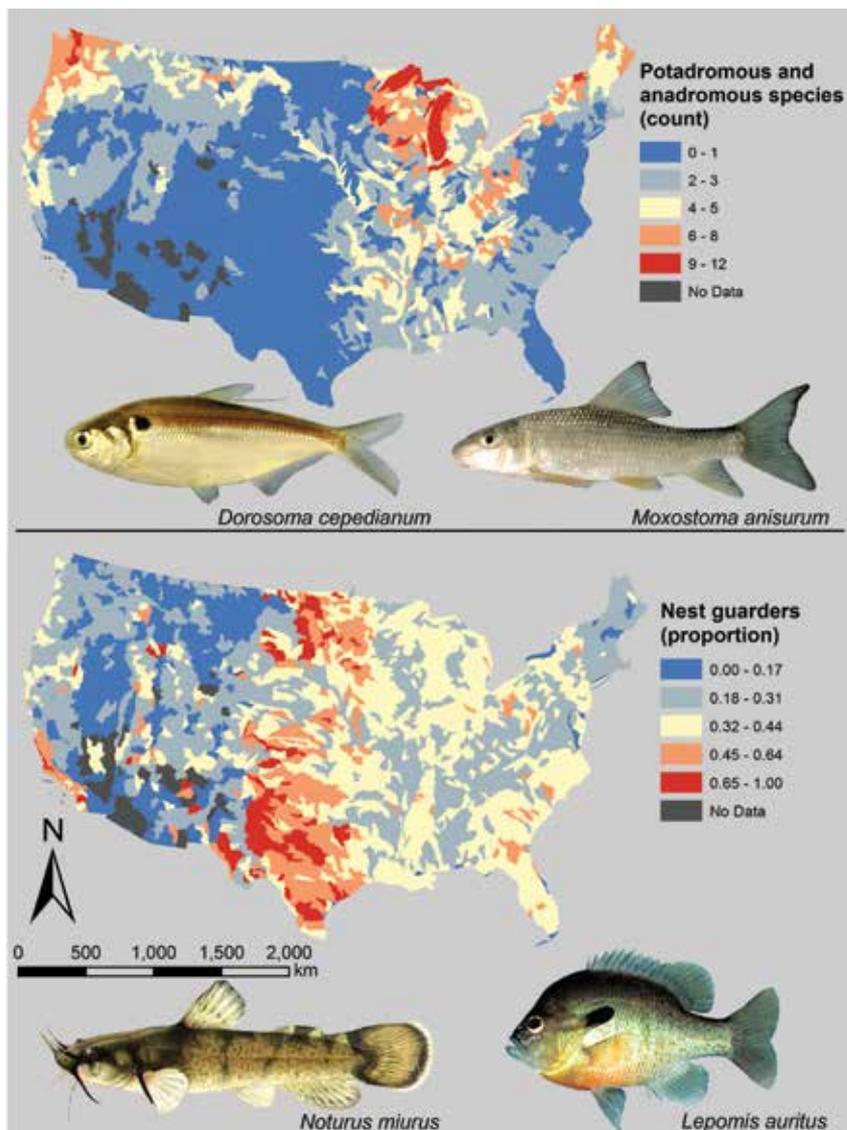


Figure 2. Example of the utility and opportunity provided by open-access data. Maps of fish trait (potadromy/anadromy and nest guarding) frequencies in watersheds across the United States were created using NatureServe fish distributions (NatureServe 2004), FishTraits (Frimpong and Angermeier 2009), and FishBase (Froese and Pauly 2012).

to provide high-quality and standardized metadata, defined as complementary information that describes all aspects of the data at hand. Metadata has received extensive attention in the ecological sciences literature and a common structure has been standardized by the Ecological Society of America: ecological metadata language (Fegraus et al. 2005). We will not delve into the metadata issue except to state that fisheries science should adopt similar standardized practices and incorporate the concept into educational programs. Every student graduating from a fisheries science program, either undergraduate or graduate, should be capable of understanding and applying metadata from commonly accessed data resources and also know how to document and structure data to make sure that it is used as it is intended in the future.

As a society, AFS would benefit from a greater awareness of deep ethical issues associated with proper data dissemination and use. Many fisheries scientists may never receive

personal acknowledgment for the data they helped generate. Though many such individuals are paid from publicly derived funds (i.e., taxes), the majority of fisheries researchers, directly or indirectly, also receive funding supported by the public. Many scientists possess and work with sensitive information (e.g., personally identifiable information) or data protected under copyrights, which if shared would constitute a breach of security or ethical violation within their respective organizations. Though many fisheries scientists are not placed in these difficult situations, the majority of fisheries professionals will likely face decisions regarding whether or not to share or accept data or when to extend coauthorship to data generators. Institutions or sponsors demanding open-access policies and 100% transparency in methods may require all raw data, including ancillary information, to be open access. Though the dissemination of final data products is typically encouraged, it is ethically problematic to pass along ancillary data owned by others, even if those data sets are open access. As another example, some scientists do not consider sharing unpublished data as grounds for coauthorship; however, each scientist has a personal responsibility to consider whether those who have shared data have also contributed to the publication by sharing ideas, such as methods for utilizing the data.

Resources for Databases

Open-access databases require financial and personnel support, a point often underappreciated by the communities that depend on them (Allen et al. 2006). To offer accessibility, databases must establish cyber-infrastructure capabilities and host the system on a proficient server. Ensuring data quality requires some degree of direct review, systematic digital checks, and creation of metadata, all of which involve dedication of time from personnel (Kolb et al. 2013). Additionally, databases housing sensitive information, such as data related to endangered species or highly valuable exploited commercial stocks, must be adequately protected from malicious intent. As databases proliferate in number and size, the need for committed resources will grow (Allen et al. 2006). Key major funding agencies, such as the National Science Foundation (NSF), have begun to offer programs to fund database creation and upkeep. Additionally, several nonprofit initiatives, including the Global Earth Observation System of Systems (earthobservations.org) and the Data Conservancy (dataconservancy.org) are designed to aid in the organization, integration, and distribution of complex environmental databases. However, to effectively sustain open-access databases, our scientific community as a culture must fully recognize their value and need for resources to maintain them (Lynch 2008).

Duplication of Effort

By providing common shared resources, open-access data have the potential to eliminate unnecessary duplication in compiling and curating information. As one example, the Core Science Analytics and Synthesis program of the USGS developed Biodiversity Information Serving Our Nation, a repository synthesizing biological occurrence data from a multitude of

resources including federal and state programs, universities, and publications (Table 1). Frequent use of the same underlying resource may increase scientific rigor by standardizing methodology across many different investigations. However, common resources may also result in considerable overlap in scientific queries, increased competition among individuals and teams, and decreased likelihood of ethical give-and-take in a future open-access society. Duplication in scientific efforts occurs not only by utilizing the same resources but also in the race to create them. To our knowledge, at least three independent concurrent efforts were executed to link the National Inventory of Dams with the National Hydrography Dataset Plus version 1 (Martin and Apse 2011; Hadjerioua et al. 2012; Ostroff et al. 2013). The technical and financial resources invested for each effort would have likely benefited from shared resources or at least shared knowledge. Though some duplication of effort is unavoidable due to research deadlines, disparate disciplines, or unwillingness to share recognition among multiple entities, open lines of communication within and among members of our society are needed. Ultimately, such dialog will increase collaboration, data creation efficiency, and more useful products that advance our science.

Shifting Patterns of Professional Experience

Analyzing data without setting foot in the field carries numerous potential consequences for fisheries scientists. Many of us entered fisheries driven by a fascination with aquatic environments resulting from experience outdoors. Given the attractiveness of database management in terms of funding support and advantages of data sharing for collaboration, we question how field-based studies and outreach in fisheries will be valued in the future. Publishing case studies is becoming increasingly difficult despite the value of publishing all findings (Clapham 2005). We foresee the possibility of diminishing incentives for field collections accompanied by heavier burdens on those who continue field activities. If our profession continues to largely shift away from fieldwork toward time spent in front of computers, will our profession remain attractive or even available to new scientists? In addition, as analyses harnessing open-access information continue to grow in scope and spatial scale, cognitive awareness and familiarity with local systems could potentially decline. Thus, we question whether the accumulation of information will be applicable to management at smaller scales. Will we be required to mandate field components in theses or dissertations?

Promulgating data sharing also increases the potential for cross-disciplinary research, which is increasingly regarded as critical to address contemporary environmental challenges (Pennington et al. 2013). Perhaps because many ecological problems involve multiple physical and biological processes operating at widely varying scales that require diverse expertise, interdisciplinary studies have often proven more impactful (Porter et al. 2012). Specifically to fisheries, high-quality hydrologic, oceanographic, and atmospheric data will allow scientists to investigate problems with resources they would never be capable of collecting within their own labs. Yet many

career assessment metrics within ecological and fisheries science rely on the evidence of scientific productivity solely within our discipline. For instance, interdisciplinary efforts may lead to reduced citation rates for researchers within the biological sciences (Larivière and Gringras 2010), which may represent a problem for career advancement in some areas of the fisheries profession, particularly academia. A full discussion on how to correct this cultural problem would be beyond the scope of our commentary. But the role of open-access databases on the growth of interdisciplinary science represents another reason why the field of fisheries must proactively adapt.

Changing Climate of Scientific Publishing

For an increasing proportion of scientists, participation in open-access data provision is not optional but mandatory. In February 2013, the executive branch of the U.S. federal government released a memorandum requiring federal agencies with greater than US\$100 million in research and development to provide public access to publications and published data generated by federally funded research (Holdren 2013). Entirely open-access journals have experienced rapid growth, and articles from these publications now comprise up to 12% of scientific work published annually (Laakso and Björk 2012). Many researchers favor open-access journals because of the accessibility. However, the debate over open-access journal policies remains equivocal. Widely varying publishing fees cause many to question not only editorial quality of open-access publications but also the existence of academic publishing in general (Van Noorden 2013). Other potential problems include the increasing fragmentation of information sources and the changing role or decreased justification of academic libraries (Monastersky 2013).

CONCLUSIONS

Open-access databases will increasingly influence all scientific disciplines in the coming decades, including fisheries science. In the face of such cultural evolution, expanding the lines of communication addressing the creation, participation, education, use, and maintenance of shared data resources among fisheries scientists is essential for our society to adapt accordingly. To ensure that our subdiscipline evolves in pace, we contend that the American Fisheries Society should take actions to encourage the evolution of an informed and data-transparent scientific community. The society could actively help realize this goal through several pathways. Open-access databases could be hosted by the society website, or at the very least dedicated web space could be provided to help scientists locate data (i.e., an active, web-based version of Table 1). Society journals could, and we argue that they should, (1) begin to offer the data used to generate articles if authors agree to have them posted; (2) strongly encourage or require data posting with accompanying metadata, similar to the PLoS ONE system; (3) require explicit instructions on how to acquire replicate data sets when articles assess open-access data; and (4) require authors to demonstrate steps made toward ensuring that they understand the structure of and have properly used open-access

data when publishing using such resources. The Electronic Services Advisory Board (ESAB) of AFS has been actively addressing these problems for over a decade and is well poised to advocate for the advancement of such concepts. Any number of other actions could be employed by the larger society to further data access and transparency, such as conference workshops, educational initiatives, and amendments to societal missions. Whatever actions, if any, are taken, our science will continue to evolve toward an open-access data society and our community must adapt as well as it can.

ACKNOWLEDGMENTS

We thank P. Ruhl, J. Kennan, S. Hetrick, M. Davis, M. Bevelhimer, A. Loftus, S. Bonar, P. Neubauer, P. Goldstein, D. Okamoto, J. Parham, D. Wiefelich, D. Hendrickson, E. Frimpong, and J. Read for presenting their development and research on open-access databases at the 2012 Annual American Fisheries Society Meeting in St. Paul, Minnesota. We thank Glen Cada, Jeff Schaeffer, and five anonymous reviewers for providing helpful comments on earlier versions of this article. The authors contributed equally to this effort.

FUNDING

This research was sponsored by the United States Department of Energy's (DOE) Office of Energy Efficiency and Renewable Energy, Wind and Water Power Technologies Program. This article has been authored by an employee of Oak Ridge National Laboratory, managed by UT Battelle, LLC, under contract DE-AC05-00OR22725 with the U.S. Department of Energy. R. Utz is supported by National Science Foundation cooperative agreement #EF1138160.

REFERENCES

- Allen, S., D. Beard, W. Fisher, L. Hartsell, A. Iles, F. Janssen, A. Loftus, M. McGuire, D. Miller, A. Ostroff, S. Shipman, S.-M. Stedman, and J. Waldon. 2006. Proceedings of the National Fisheries Data Summit. Focusing on applications to the National Fish Habitat Initiative. Report from the Computer User Section of the American Fisheries Society, Bethesda, Maryland. Available: www.fishdata.org/content/reports. (November 2013).
- Austen, D., D. Beard, H. Drewes, P. Haverland, D. Johnston, A. Loftus, D. Miller, S. Sobaski, J. Waldon, and L. Borge-Willis. 1998. Proceedings of the National Freshwater Fisheries Database Summit. Report from the Computer User Section of the American Fisheries Society, Bethesda, Maryland. Available: www.fishdata.org/content/reports. (November 2013).
- Balon, E. K. 1975. Reproductive guilds in fishes: a proposal and definition. *Journal of the Fisheries Research Board of Canada* 32:821–864.
- Beard, T. D., D. Austen, S. J. Brady, M. E. Costello, H. G. Drewes, C. H. Young-Dubovsky, C. H. Flather, T. W. Gengerke, C. Larson, A. J. Loftus, and M. J. Mac. 1998. The Multi-state Aquatic Resources Information System: an Internet system to access fisheries information in the upper Midwestern United States. *Fisheries* 23(5):14–18.
- Clapham, P. 2005. Publish or perish. *Bioscience* 55:390–391.
- Costello, C., D. Ovando, R. Hilborn, S. D. Gaines, O. Deschenes, and S. E. Lester. 2012. Status and solutions for the world's unassessed fisheries. *Science* 338:517–520.
- Crone, T. J., and M. Tolstoy. 2010. Magnitude of the 2010 Gulf of Mexico oil leak. *Science* 330:634.
- Cukier, K. 2010. Data, data everywhere. *The Economist*. Available: www.economist.com/node/15557443?story_id=15557443. (February 2010).
- Fegraus, E. H., S. Andelman, M. B. Jones, and M. Schildhauer. 2005. Maximizing the value of ecological data with structured metadata: an introduction to ecological metadata language (EML) and principles for metadata creation. *Bulletin of the Ecological Society of America* 86:158–168.
- Food and Agriculture Organization of the United Nations. 2011. FishStat Plus—universal software for fishery statistical time series. Available: www.fao.org/fishery/statistics/en. (March 2013).

- Fox, P., and J. Hender. 2011. Changing the equation on scientific data visualization. *Science* 331:705–708.
- Frimpong, E. A., and P. L. Angermeier. 2009. Fish Traits: a database of ecological and life history traits of freshwater fishes of the United States. *Fisheries* 34:487–495.
- . 2010. Trait based approaches in the analysis of stream fish communities. Pages 109–136 in K. B. Gido and D. A. Jackson, editors. *Community ecology of stream fishes: concepts, approaches, and techniques*. American Fisheries Society, Symposium 73, Bethesda, Maryland.
- Froese, R., and D. Pauly, editors. 2012. *FishBase*. World Wide Web electronic publication. Available: www.fishbase.org. (February 2013).
- Geoghegan, P. 1996. The management of quality control and quality assurance systems in fisheries science. *Fisheries* 21(8):14–18.
- Hadjerioua, B., Y. Wei, and S.-C. Kao. 2012. An assessment of energy potential at non-powered dams in the United States. Department of Energy, GPO DOE/EE-0711, Washington, D.C. Available: http://nhaap.ornl.gov/system/files/NHAAP_NPD_FY11_Final_Report.pdf. (January 2014).
- Hamm, D. E. 2012. Development and evaluation of a data dictionary to standardize salmonid habitat assessments in the Pacific Northwest. *Fisheries* 37:6–18.
- Holdren, J. 2013. Increasing access to the results of federally funded scientific research. Memorandum for the heads of executive departments and agencies. Office of Science and Technology Policy, Executive Office of the President, Washington, D.C. Available: www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf. (August 2013).
- Kolb, T. L., E. A. Blukacz-Richards, A. M. Muir, R. M. Claramunt, M. A. Koops, W. W. Taylor, T. M. Sutton, M. T. Arts, and E. Bissel. 2013. How to manage data to enhance their potential for synthesis, preservation, sharing, and reuse—a Great Lakes case study. *Fisheries* 38:52–64.
- Laakso, M., and B.-C. Björk. 2012. Anatomy of open access publishing: a study of longitudinal development and internal structure. *BMC Medicine* 10:1–9.
- Lariviere, V., and Y. Gingras. 2010. On the relationship between interdisciplinarity and scientific impact. *Journal of the American Society for Information Science and Technology* 61:126–131.
- Larsen, P. O., and M. von Ins. 2010. The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics* 84:575–603.
- Loftus, A. J., and C. H. Flather. 2012. Fish and other aquatic resource trends in the United States: a technical document supporting the Forest Service 2010 RPA Assessment. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, General Technical Report RMRS-GTR-283, Fort Collins, Colorado.
- Luo, Y., K. Ogle, C. Tucker, S. Fei, C. Gao, S. Ladeau, J. S. Clark, and D. S. Schimel. 2011. Ecological forecasting and data assimilation in a data-rich era. *Ecological Applications* 25:1429–1442.
- Lynch, C. 2008. How do your data grow? *Nature* 455:28–29.
- Martin, E. H., and C. D. Apse. 2011. Northeast aquatic connectivity: an assessment of dams on Northeastern Rivers. The Nature Conservancy, Brunswick, ME. Available: http://static.rcngrants.org/sites/default/files/final_reports/NEAquaticConnectivity_Report.pdf. (January 2014).
- Mims, M. C., J. D. Olden, Z. R. Shattuck, and N. L. Poff. 2010. Life history trait diversity of native freshwater fishes in North America. *Ecology of Freshwater Fish* 19:390–400.
- Monastersky, R. 2013. Publishing frontiers: the library reboot. *Nature* 495:430–432.
- Moyle, P. B., and J. J. Cech. 2004. *Fishes. An introduction to ichthyology*, 5th ed. Prentice Hall, Upper Saddle River, New Jersey.
- Multi-resolution Land Characteristics Consortium. 2013. National Land Cover Database. Available: www.mrlc.gov/. (January 2013).
- Murawski, S., R. Methot, and G. Tromble. 2007. Biodiversity loss in the ocean: how bad is it? *Science* 316:1281–1284.
- NatureServe. 2004. Downloadable animal datasets. NatureServe Central Databases. Available: www.natureserve.org/getData/dataSets/watershedHucs/index.jsp. (November 2011).
- Nelson, S. 2008. Big data: the Harvard computers. *Nature* 455:36–37.
- Olascoaga, M. J., and G. Haller. 2012. Forecasting sudden changes in environmental pollution patterns. *Proceedings of the National Academy of Sciences* 109:4738–4743.
- Olden, J. D., and M. J. Kennard. 2010. Intercontinental comparison of fish life history strategies along a gradient of hydrologic variability. Pages 109–136 in K. B. Gido and D. A. Jackson, editors. *Community ecology of stream fishes: concepts, approaches, and techniques*. American Fisheries Society, Symposium 73, Bethesda, Maryland.
- Ostroff, A., D. Wiefelich, A. Cooper, and D. Infante. 2013. 2012 National Anthropogenic Barrier Dataset (NABD). USGS Science Base Catalog. Available: www.sciencebase.gov/catalog/item/512cf142e4b0855fde669828. (January 2014).
- Pacific Marine Fisheries Commission. 1980. 33rd Annual report of the Pacific Marine Fisheries Commission for the year 1980. Pacific Marine Fisheries Commission, Portland, Oregon. Available: www.psmfc.org/resources/publications-maps-2/psmfc-annual-reports. (November 2013).
- Pauly, D. 1995. Anecdotes and the shifting baseline syndrome in fisheries. *Trends in Ecology and Evolution* 10:430.
- Pennington, D. D., G. L. Simpson, M. S. McConnell, J. M. Fair, and R. J. Baker. 2013. Transdisciplinary research, transformative learning, and transformative science. *BioScience* 63:564–573.
- Porter, A. L., J. Garner, and T. Crowl. 2012. Research coordination networks: evidence of the relationship between funded interdisciplinary networking and scholarly impact. *BioScience* 62:282–288.
- Reichman, O. J., M. B. Jones, and M. P. Schildhauer. 2011. Challenges and opportunities of open data in ecology. *Science* 331:703–705.
- Ricard, D., C. Minto, O. Jensen, and J. Baum. 2012. Examining the knowledge base and status of commercially exploited marine species with the RAM Legacy Stock Assessment Database. *Fish and Fisheries* 13:380–398.
- Sea Around Us Project. 2013. Sea Around Us Project. Fisheries, Ecosystems and Biodiversity. Available: www.seaaroundus.org/. (March 2013).
- Seeb, L. W., A. Antonovich, M. A. Banks, T. D. Beacham, M. R. Bellinger, S. M. Blakenhip, M. R. Campbell, N. A. Decovich, J. C. Garza, C. M. Guthrie, T. A. Lundrigan, P. Moran, S. R. Narum, J. J. Stephenson, K. J. Supernault, D. J. Teel, W. D. Templin, J. K. Wenburg, S. F. Young, and C. T. Smith. 2007. Development of a standardized DNA database for Chinook Salmon. *Fisheries* 32:540–552.
- Silver, S. 2003. A new slant on data access. *Frontiers in Ecology and the Environment* 1:343.
- Tenopir, C., S. Allard, K. Douglas, A. U. Aydinoglu, L. Wu, E. Read, M. Manoff, and M. Frame. 2011. Data sharing by scientists: practices and perceptions. *PLoS ONE* 6:1–21.
- U.S. Environmental Protection Agency. 2012. Ecoregions of North America. Available: www.epa.gov/wed/pages/ecoregions.htm. (January 2013).
- USGS (U.S. Geological Survey). 2013. BioData—aquatic bioassessment data for the nation. Available: <https://aquatic.biodata.usgs.gov/landing.action>. (January 2013).
- Van Noorden, R. 2013. Open-access: the true cost of scientific publishing. *Nature* 495:426–429.
- Wang, L., D. Infante, P. Esselman, A. Cooper, D. Wu, W. Taylor, D. Beard, G. Whelan, and A. Ostroff. 2011. A hierarchical spatial framework and database for the National River Fish Habitat Condition Assessment. *Fisheries* 36:436–449.
- Worm, B., E. B. Barbier, N. Beaumont, J. E. Duffy, C. Folke, B. S. Halpern, J. B. C. Jackson, H. K. Lotze, F. Micheli, S. R. Palumbi, E. Sala, K. A. Selkoe, J. J. Stachowicz, and R. Watson. 2006. Impacts of biodiversity loss on ocean ecosystem services. *Science* 314:787–790. 



Innovative tracking solutions for
fish and wildlife since 2003



- High performance HDX and FDX PIT tags
- Glass and food-safe types
- ISO 11784/11785 compliant
- Long range and proximity readers
- Affordable monitoring stations
- Easy to install
- Antenna design tools
- Tag implantation equipment
- Expert technical support

(866) 484-3174 toll free
(503) 788-4380 international
orfid-pdx Skype
sales@oregonrfid.com

Visit our online store at oregonrfid.com