A programmable information system for management and analysis of aquatic species range data in California

Nicholas R. Santos [a,*], Jacob V.E. Katz [a,b], Peter B. Moyle [a,c], Joshua H. Viers [a,d]

[a] Center for Watershed Sciences, University of California, Davis, United States
[b] California Trout, United States
[c] Department of Wildlife, Fisheries, & Conservation Biology, University of California, Davis, United States
[d] School of Engineering, University of California, Merced, United States

A R T I C L E   I N F O

A B S T R A C T

The decline of species worldwide is both alarming and difficult to document due to a lack of reliable information on the geospatial extent and corresponding status of a given taxon. Freshwater habitats are disproportionately degraded globally with resultant declines in populations in freshwater fishes and subsequent retractions in biogeographic ranges. Conservation challenges in freshwater are compounded because aquatic taxa are inherently difficult to map. We addressed this problem for California freshwater fishes by developing the software and underlying database. The software consists of a Python program, database, and suite of tools using ESRI ArcGIS scripting interfaces to translate species range data into an electronic record set of occurrences housed in Microsoft Access. The system was designed to capture, store, map, and report on the spatial and temporal dynamics of targeted species by using standard spatial units as primary indexing objects to meet current natural resource management objectives. However, the software not only tracks the provenance of underlying empirical records through space and time, but also is robust to inferential modeling results and expert knowledge, which allows for future empirical discovery and validation. After importing and standardizing 274,555 records from 154 data layers, we found that most existing records are highly concentrated spatially, representing only 39% of the mapping domain. We also determined that most empirical records are skewed toward recreational fisheries, with few records documenting the range of native species found in California. Future biogeographic mapping efforts will be aided by the baseline data and updated range maps contained in the database. Although the system is currently used for the inventory and mapping of native freshwater fish species in California, the underlying informatics framework is agnostic to biological taxonomy or spatial realm allowing other to adapt the computer code and database for their own needs.

Software availability

1. Introduction

1.1. Background

Global environmental change from human activities is both widespread and rapid (Millenium Ecosystem Assessment, 2005). This is especially true in freshwater ecosystems, which are disproportionately affected by human activities (Dudgeon et al., 2006). Freshwater ecosystems are affected by pollution, water diversion and flood protection infrastructure, habitat alteration, and invasive species, among many stressors. Further, freshwater ecosystems and freshwater supplies are increasingly vulnerable to

the combined effects of human abstraction and hydroclimatic alteration (Vörösmarty et al., 2000).

Efforts to minimize the negative effects of anthropogenic change through conservation often rely on maps to depict subjects of interest (Wilson et al., 2007), such as the location of a species or a collection of species, or alternately places of refuge, such as a park or preserve. To initiate a process or program to conserve a species or a suite of species, it is necessary to know where to act. While this need is common in resource management, few systematic tools have been developed to collect, store, and map aquatic biodiversity through space and time. This may be due to the inherent challenges of mapping ranges of aquatic species including: animals in aqueous environments are often difficult to detect; methods of detection are comparatively poor; field observations are highly localized; present day distributions are confounded by human introductions; and human activities have extirpated species from suitable habitats (Naiman and Dudgeon, 2011; Nel et al., 2009; Pittock et al., 2008; Strayer and Dudgeon, 2010; Thieme et al., 2007; Viers and Rheinheimer, 2011; Vorosmarty et al., 2010). This paper introduces a software system designed to function within these constraints to produce high resolution, easily updated range maps for resource managers.

1.2. California

Like other regions of the Mediterranean Biome, California is disproportionately impacted by human activities (Grantham et al., 2010; Underwood et al., 2009). The state has the combined challenges of being ecologically unique − 82% of fish taxa are regionally endemic (Moyle et al., 2011) − and containing 38 million people. Freshwater ecosystems of California are threatened by several compounding factors, including but not limited to river regulation, pollution, and invasive species (Marchetti et al., 2004). The number of the state's native fish taxa in decline is worrisome. Of 129 native fish taxa 83% of are in sharp decline or have already gone extinct (Moyle et al., 2011). Indicative of the critical nature of this problem is the impending extinction of several California runs of anadromous salmonid (Katz et al., 2012).

The United States Department of Agriculture Forest Service (FS) is California's largest land holder, administering 9,526,000 ha across 18 National Forests, accounting for approximately 20% of the state's land surface. Management objectives, coupled with regulatory mandates, such as the federal Endangered Species Act, obligate the FS to identify, catalog, monitor and account for aquatic species. To date, FS and other resource managers have relied on digital maps generated as part of a statewide mapping effort of California fishes conducted in 1998 (Moyle and Randall, 1998). While these maps have proved helpful in better understanding the distribution of fishes in California, they were developed with early geographic information system (GIS) technologies at coarse scales and did not portray empirical data but instead were created using expert knowledge. Furthermore, these maps could not be readily updated as new information became available or combined with other mapping approaches, such as inferential spatial modeling − issues that afflict many similar efforts (Graham and Hijmans, 2006).

1.3. PISCES

To more effectively understand and manage the sensitive fishes of California's National Forests, we developed a GIS dependent database and accompanying software. The software, termed PISCES, was intended to robustly manage mapping efforts of sensitive fish species distributions across forests in ways that previous efforts could not. Designed as a decision support system for resource managers, PISCES incorporates and catalogs disparate data types of empirical and inferred species observations. Subsequent encoding of these observations relies on a standardized, yet generic, data framework that overcomes issues of spatial scale, temporal discontinuity, data format discrepancies, and regional context.

Standard outputs from PISCES map or tabulate phenomena that allow managers to identify species ranges, patterns of biodiversity, and areas of where biological data is lacking. Although intended to solve acute management problems associated with California's sensitive fish species on FS lands, the PISCES platform, with its enhanced GIS database and data provenance and transaction capabilities, can be applied to other natural resource management systems.

PISCES fits in with a general class of research software that attempts to organize, store, and present disparate sets of data (See (Guzman et al., 2013; Horsburgh et al., 2009; Rangel et al., 2010; Souza Muñoz et al., 2011; Villa et al., 2009)) − a common need as organizations gather overlapping data in different forms. This very basic question of "*where*" is foundational to much other work in ecology and environmental science and is the reason we have GIS. Systems such as PISCES aid in that question of where by translating, organizing, and collecting data into functional units with problem-domain specific tools. In fact, while our instance of PISCES is California focused, the software and database are a generalized system for tracking and analyzing species range *anywhere* using a predefined set of zones.

1.4. Design requirements

We imposed several formal design requirements on the development of PISCES that were intended to address issues of species data collection, storage, and analysis within existing desktop GIS software.

1. **Resource managers are often in a standalone computing environment with "stock" desktop software; thus, outcome must be serverless and portable and not require additional investment in hardware or software.** PISCES uses Python 2.6, ArcGIS and a Microsoft Access geodatabase for data processing and storage.
2. **Resource managers collect new information regularly; thus, data and results must be dynamic.** PISCES can process and import many new datasets and will update maps and tables as data are updated.
3. **Resource managers are not database engineers; thus, data must be easily updateable and flexible to unknown and non-standard data formats.** Data import occurs via extensible classes that can adapt to datasets regardless of format, design, scale, quality, or quantity. New classes can be added as needed to handle foreign data.
4. **Resource managers are faced with ever-changing mandates and management challenges; thus, resultant frameworks must be compatible with future questions and needs.** Flexible configuration and callback functions allow for runtime customization of maps to meet future needs. Stock functions and map configurations handle major outputs of interest and new functions can be added.
5. **Resource managers have inherent, place-based knowledge; thus, software must be able to capture verbal and written accounts.** PISCES includes tools that support interactive range viewing, input, and editing with experts in ArcMap 10 to capture data not currently in any file.
6. **Resource managers are human and mistakes can happen; thus, all data and changes must be traceable.** PISCES logs all transactions and stores all inputs unmodified so that changes can be traced. In many cases, a change can be reverted.

We evaluated the suitability of existing software and databases to meet our needs and found no ready solutions. Existing global methods for storing and containing aquatic species range data either were not domain specific, not high enough resolution, or did not cover the entire domain extent (Kaschner et al., 2010). Existing California-oriented systems frequently were not reproducible, did not have a mechanism to add data where empirical data is sparse or missing, were missing appropriate analysis tools, or relied on generalized GIS workflows rather than a unified system of range tracking, provenance, and mapping (Biogeographic Data Branch, 2011; BIOS, 2011a; Hijmans et al., 2000; Olson, 1997). Finally, existing systems frequently do not allow for true integration of empirical and expert data or were focused primarily on inferential modeling based on existing data instead of development and storage of range data.

For almost as long as biologists have had access to computers, we have been developing software systems to understand species range — GIS is, at its basic level, designed to help us understand this basic question of "where" some phenomenon occurs (Kaschner et al., 2010; Scott et al., 1987). Each system involves a series of tradeoffs regarding what can be represented with the available resources (Hurlbert and White, 2005; Jetz et al., 2012). Our software is the next iteration in such a series of databases and tools and reflects our set of design requirements.

2. Methods

PISCES is a system to collect, process, store, and output species presence data. It is designed to aggregate and convert diverse spatial data for California native fish species to United States Geological Survey Hydrologic Unit Code 12 (HUC12) subwatersheds (henceforth referred to as mapping units, see http://water.usgs.gov/GIS/huc.html for more information), but can be adapted for other areal units. For geospatial operations PISCES uses the arcpy site package (library) for Python 2.6.5 and Microsoft Access 2010 via the pyodbc package for database interfacing.

The software was developed to transform, store, and output geospatial data imported from third party sources. The database was seeded by capturing expert knowledge data with purpose-built tools. All imported and transformed data was quality controlled by species experts. Range maps for 67 of California's 129 native fish taxa were produced. Additionally, summary map outputs allow visual representation of data analysis and spatial patterns in aquatic biodiversity. Fig. 1 shows the flow of data within and through PISCES.

2.1. Implementation

2.1.1. Storage

PISCES tracks species ranges through individual presence records for species in each watershed. Each presence record in PISCES includes foreign keys to the mapping unit, taxon, origin dataset, and a presence type along with any other relevant metadata specific to that record. PISCES presence classifications include whether the data are historic or current, for native or translocated range, and whether the data source is empirical, model-generated or expert knowledge. PISCES uses these attributes to layer and separate data upon output. Table 1 describes the fields PISCES stores for each presence record.

Presence records can be assigned to groups, such as our "quality controlled" dataset — a set of records that has been manually reviewed by species experts and is considered authoritative. PISCES also groups taxa, such as native taxa, or taxa of interest for specific analyses (e.g., by taxonomic family, or geographic region). Using these metadata, users can filter PISCES outputs to only include selected taxa and/or data types.

2.1.2. Inputs and processing

PISCES can natively import and store spatial or tabular data that describes the presence of a tracked taxon. By using a built-in extract, transform, and load (ETL) mechanism, PISCES is extensible to handle datasets regardless of design, scale, format, origin, or reliability through the use of a set of Python classes that handle increasing complexity and differing design. Core classes called "input filters" handle the most common data transformations and tracking, such as translation of spatial data to mapping unit identifiers, field mapping for important metadata, and pre-conversion of inputs to match expected formats. The core input filter classes include multiple extension points, including subclassing and custom method handlers for fields.
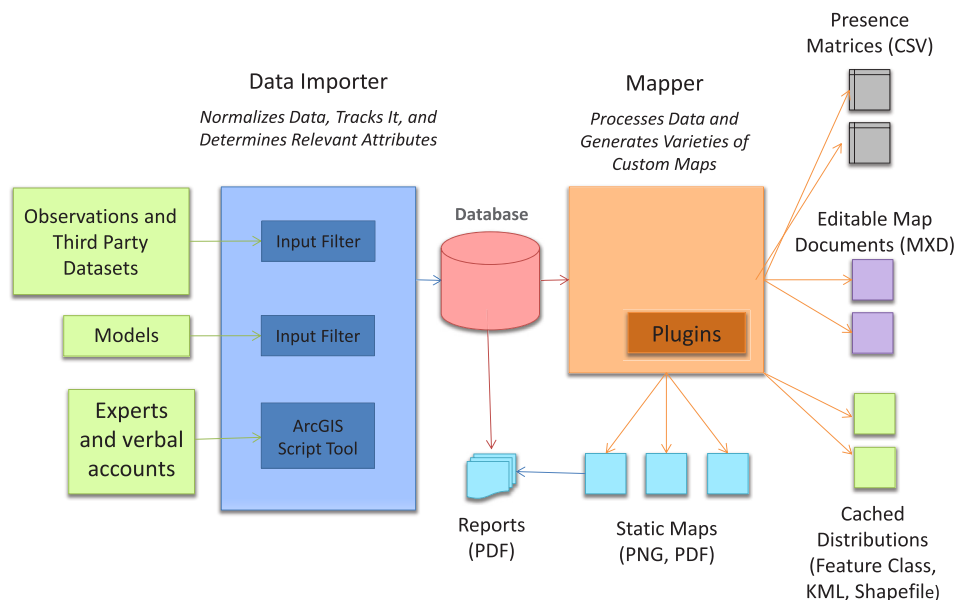

Fig. 1. PISCES schematic. Data flows in through multiple pathways, is standardized to the database, and is output via multiple pathways for analysis and mapping.

Table 1
Schema for a PISCES *Observation*. *Observation*s relate species, mapping units and presence types and include appropriate metadata.

Field name	MS access data type	Description
OBJECTID	AutoNumber	Primary key
Set_ID	Number	The Observation_Set that this Observation is associated with. Many Observations to one Observation_Set
Species_ID	Text	The Species FID that this record describes
Zone_ID	Text	Mapping Unit (HUC12) ID. Many Observations to one Zone_ID
Presence_Type	Number	The type of presence (observed, expert knowledge, modeled, current, historic, etc.) of a species
IF_Method	Number	The code path that processed this observation from source data
Longitude	Text	Optional. Longitude if this observation came from point data
Latitude	Text	Optional. Latitude if this observation came from point data
Survey_Method	Text	Optional. Type of surveying used to capture and identify the species
Notes	Memo	Additional information about the record added by operator
Observation_Date	Text	Date of the observation, as noted in the source data
Date_Added	Date/Time	The date/time that this record was added to PISCES
Other_Data	Memo	Semicolon separated other fields of interest captured by input filter

Table 2
Summary of PISCES records by source.

Group	# QC records	# Datasets included	# Records
United States Forest Service	24,398	36	148,957
Expert data	14,201	1	19,963
Moyle and Randall (1998)	8140	88	92,347
Federal Energy Regulatory Commission (FERC) Surveys	1008	6	5601
Trout unlimited[a]	572	5	872
California Department of Fish and Wildlife, California Natural Diversity Database (CNDDB)[b]	511	1	1498
California Department of Fish and Wildlife, Biogeographic Information & Observation System (BIOS)[c]	241	2	2106
US National Marine Fisheries Service (NMFS) Salmonid data[d]	221	6	1159
US Environmental Protection Agency Environmental Monitoring and Assessment Program (EMAP) data[e]	62	1	464
Various supporting data[f]	6	6	12

[a] (TU, 2011).
[b] (Biogeographic Data Branch, 2011).
[c] (BIOS, 2011a,b,c).
[d] (Schick et al., 2005).
[e] (EMAP, 2006).
[f] Datasets and reports from numerous agencies. For a full listing, see observations.gdb in Supplementary data.

In order to generate range maps from discrete records of observation and provide an appropriate level of uncertainty to data whose validity changes with time, input filters transform data, regardless of geometry, to a collection of mapping units. This method affords PISCES additional power, but also places limitations on the results. All accessible PISCES range data are in the form of these mapping unit presence records.

PISCES is able to track all changes to data, from import of new data to modifications of existing data via metadata attributes and a built-in transaction system. Data changes made using the built-in tools can be retroactively reverted to prior states if desired. Table 2 contains a summary of the data we have imported into PISCES, including data covering all National Forests within California and multiple statewide datasets.

2.1.3. Expert-verified range maps

Input filters allow for high volume data translation from existing datasets into PISCES, but have little capacity to validate, expand, or infer ranges as an expert can. Depending on source data design and method of translation to our mapping unit, stock input filters often produce overestimates of range in some places and incomplete records in others. To control for this problem, we developed a set of tools that provide a viewing and editing interface for experts within ArcMap 10. The editing workflow most typically begins with loading the PISCES export (see next section) of a taxon's range into ArcMap along with other statewide geographic data. Using the built-in selection tools users select mapping units to add, remove or query, which are then passed into PISCES tools to make modifications to underlying data records. Upon completion of a given taxon, records represented by that defined range are added to a Quality Controlled (referred to as QC) dataset in PISCES that contains expert validated range data.

At the time of this writing, the primary experts involved in the PISCES data validation were Jacob Katz and Peter Moyle, with input

from Ron Yoshiyama, Molly Stevens, Bjorn Erickson, and Joe Kiernan for specific taxa. Certain regions of California were verified by United States Forest Service Field Office Staff when possible. Many of the imported datasets had some or all of their data controlled by experts as well, but because this too was subjected to the initial editing process, PISCES was agnostic to those edits.

Validation by experts was based on numerous factors, depending on the taxon. In many cases, experts had personal experience with taxa in regions that computer-accessible datasets do not currently track. In other cases, life history characteristics of a taxon were used and compared with regional knowledge of streams. Information on the biology and status of each species was derived from detailed reviews in (Moyle et al., 2008, 2011; Moyle, 2002; Moyle et al., 2013; Moyle et al., 1995), collection of scientific literature and agency reports, and by personal communications with biologists.

For a full accounting of importing and editing data in PISCES, please see Appendix A: Data Flow through PISCES.

2.1.4. Mapping

The mapping code contained in PISCES is designed as a flexible component to allow for outputs supporting differing management and mapping needs. For example, PISCES does not contain code to generate a single, specific map. Rather, the code contains classes and supporting functions that expand a single "mapping configuration" into as many maps as necessary. Users create maps via configurations that specify general parameters such as title, layer names, and template ESRI Map Documents. These parameters are translated into an instance of the *fish_map* class (see Fig. 2). *fish_map*s have a one-to-many relationship with the *map_layer* class, which stores the layer-specific data. PISCES uses this relationship to display all data types available for a selected taxon in a single map in the default mapping configuration.

The key parameters behind each layer are the queries and callback functions, implemented as *map_layer* and *custom_query* classes, which allow significant flexibility in data display within the confines of standard ArcGIS capabilities. *map_layer*s contain Microsoft Access compatible SQL queries that return a set of
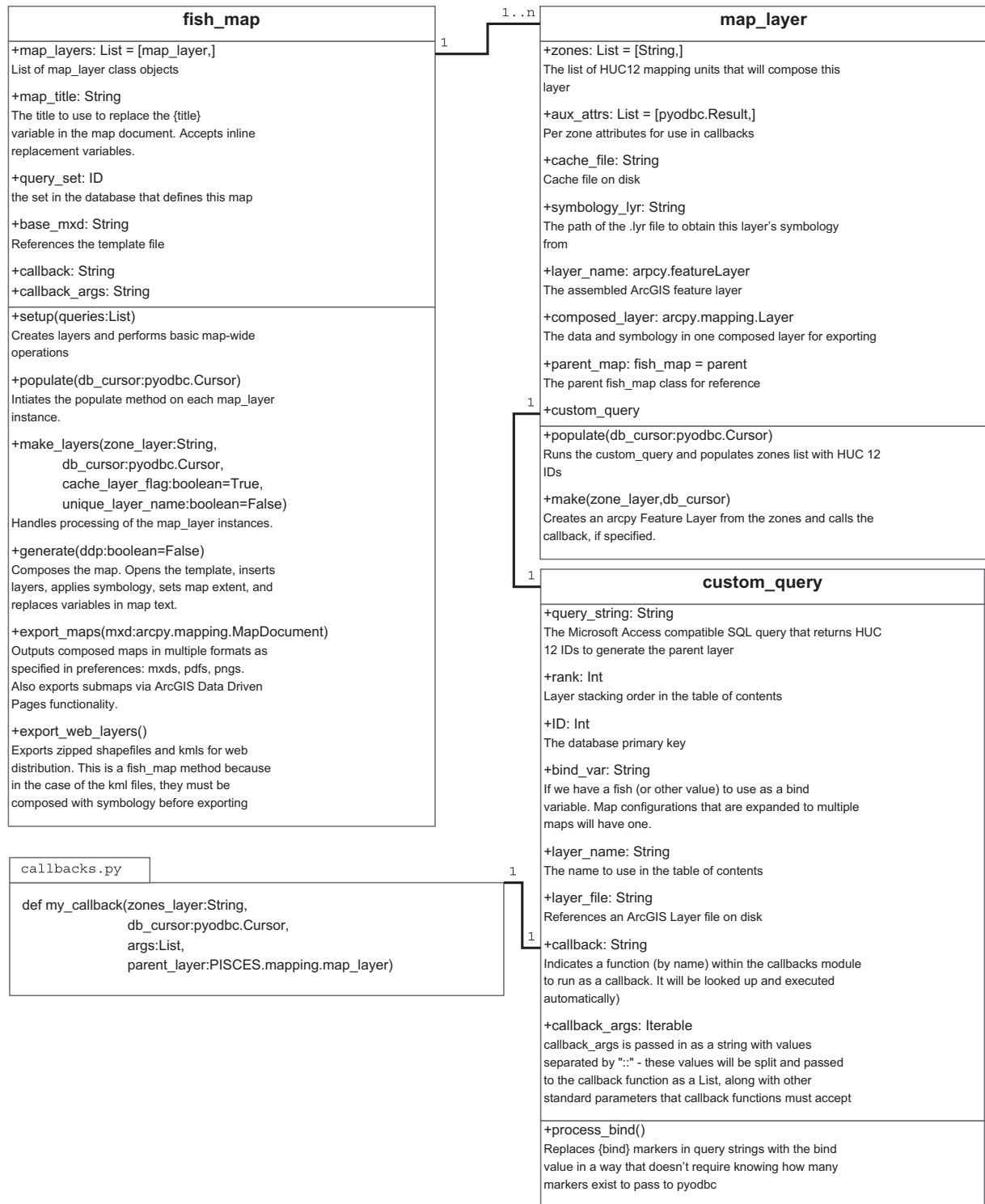
fish_map

+map_layers: List = [map_layer,]
List of map_layer class objects

+map_title: String
The title to use to replace the {title} variable in the map document. Accepts inline replacement variables.

+query_set: ID
the set in the database that defines this map

+base_mxd: String
References the template file

+callback: String

+callback_args: String

+setup(queries:List)
Creates layers and performs basic map-wide operations

+populate(db_cursor:pyodbc.Cursor)
Intiates the populate method on each map_layer instance.

+make_layers(zone_layer:String,
      db_cursor:pyodbc.Cursor,
      cache_layer_flag:boolean=True,
      unique_layer_name:boolean=False)
Handles processing of the map_layer instances.

+generate(ddp:boolean=False)
Composes the map. Opens the template, inserts layers, applies symbology, sets map extent, and replaces variables in map text.

+export_maps(mxd:arcpy.mapping.MapDocument)
Outputs composed maps in multiple formats as specified in preferences: mxds, pdfs, pngs. Also exports submaps via ArcGIS Data Driven Pages functionality.

+export_web_layers()
Exports zipped shapefiles and kmls for web distribution. This is a fish_map method because in the case of the kml files, they must be composed with symbology before exporting

---

`callbacks.py`

def my_callback(zones_layer:String,
      db_cursor:pyodbc.Cursor,
      args:List,
      parent_layer:PISCES.mapping.map_layer)

---

map_layer

+zones: List = [String,]
The list of HUC12 mapping units that will compose this layer

+aux_attrs: List = [pyodbc.Result,]
Per zone attributes for use in callbacks

+cache_file: String
Cache file on disk

+symbology_lyr: String
The path of the .lyr file to obtain this layer's symbology from

+layer_name: arpcy.featureLayer
The assembled ArcGIS feature layer

+composed_layer: arcpy.mapping.Layer
The data and symbology in one composed layer for exporting

+parent_map: fish_map = parent
The parent fish_map class for reference

+custom_query

+populate(db_cursor:pyodbc.Cursor)
Runs the custom_query and populates zones list with HUC 12 IDs

+make(zone_layer,db_cursor)
Creates an arcpy Feature Layer from the zones and calls the callback, if specified.

---

custom_query

+query_string: String
The Microsoft Access compatible SQL query that returns HUC 12 IDs to generate the parent layer

+rank: Int
Layer stacking order in the table of contents

+ID: Int
The database primary key

+bind_var: String
If we have a fish (or other value) to use as a bind variable. Map configurations that are expanded to multiple maps will have one.

+layer_name: String
The name to use in the table of contents

+layer_file: String
References an ArcGIS Layer file on disk

+callback: String
Indicates a function (by name) within the callbacks module to run as a callback. It will be looked up and executed automatically)

+callback_args: Iterable
callback_args is passed in as a string with values separated by "::" - these values will be split and passed to the callback function as a List, along with other standard parameters that callback functions must accept

+process_bind()
Replaces {bind} markers in query strings with the bind value in a way that doesn't require knowing how many markers exist to pass to pyodbc

Fig. 2. Simplified UML diagram of the mapping classes. Noncritical class elements are not shown.

mapping unit identifiers. *map_layer* methods translate these IDs into a composed spatial layer of those mapping units and symbolize it according to the parameters in the map configuration. In the case of species range maps, SQL queries of the following structure are most common:

select distinct Zone_ID from Observations where Species_ID = ? and Presence_Type = 3.

The "?" indicates a bind value − a runtime parameter for the query that allows us to pass in the species id for each map. In this case a *Presence_Type* of 3 limits the resulting layer to expert knowledge data, but identifiers of other data types are used in each layer in the default mapping configurations. PISCES composes a complete range map by combining layers representing each data type of interest for a species into a single map.

When a callback function is specified on a *map_layer*, the generated spatial layer is passed to the callback function, which returns the final layer to be added to the map. As a result, the callback has full control over the final layer and can change the data type or features, add fields, or otherwise modify the contents. Common implementations of callbacks include generating species diversity and assemblage information by mapping unit. Further, after the callback, the *fish_map* no longer assumes the layer contains mapping units, so once the layer is passed off, callbacks have full flexibility to change outputs − even to other spatial units.

2.1.5. Outputs

Primary outputs from PISCES are static digital maps, but PISCES can also output locational matrices of presence/absence, composed spatial layers, and reports. To accomplish map output functions, PISCES uses ArcGIS' native mapping capabilities as per the *arcpy.mapping* Python module, and composes maps from stored templates. Outputs include the composed map as an ArcGIS Map Document (MXD), static maps at multiple scales in PNG and PDF formats, and web-ready versions of the underlying layers as symbolized KMLs and zipped ESRI Shapefiles. By default, all generated features are cached locally for use in other analyses by researchers and managers.

2.1.6. Version control and data management

The PISCES software and data are managed within a Mercurial repository for complete versioning and data management. Internally, PISCES maintains its own simplified transaction system that stores input commands that would change data, messages provided by the editor of the data justifying the change, and copies of deleted or changed data. This transaction system prevents data loss between commits to Mercurial and provides a partial history of changes for each taxon.

2.2. Methods for key callbacks

2.2.1. Richness functions

Aside from species range maps, a common data need from PISCES is mapping unit level diversity information. PISCES addresses this need through a generic callback that takes a SQL query, an output datatype and a field name as parameters. It creates a field of the specified output datatype on the spatial data layer and runs that passed in query for every individual mapping unit in the *map_layer*.

The most common form of this query is:

select count(*) as col_value from (select distinct Observations.Species_ID from Observations where Observations.Presence_Type in (1,3,6,7,9) and Observations.Zone_ID = ?)

This query returns the number of taxa for which PISCES contains data indicating current presence in the mapping unit being processed. When using this callback, it will be run with the mapping unit identifier as the bind variable, replacing the "?," for every mapping unit in California. In this particular query, *Presence_Type* of 1,3,6,7, and 9 indicate current presence in that mapping unit. Additional limiting clauses are often added to queries, such as to impart designation for particular species groups or originating datasets. For more complicated measures of biodiversity, such as Jaccard Distance (a mathematical measure of similarity), we created new callbacks, and hooks are available in the code for new analysis callbacks as future needs arise. See Supplementary documents for the code for all callbacks.

2.3. Data assessment

For the purposes of this paper and assessing the current state of fish taxa range data, we developed a separate script to compare the existing Moyle and Randall (1998) data with our new refined range map. We also created a PISCES map set that shows the count of empirical observations in each mapping unit, which highlights gaps in sampling and conversion of historical records to digital formats.

2.3.1. Comparison of PISCES to Moyle and Randall, 1998

To assess potential changes in distributional data between Moyle and Randall (1998) and current PISCES ranges, we examined the Moyle and Randall ranges for which we also had quality-controlled PISCES data ($n = 19$).

For each taxon, we executed queries and overlap functions to determine:

1. quality controlled area of mapping units that make up the Moyle and Randall polygon;
2. percent overlap of the final PISCES range with Moyle and Randall polygon
3. percent overlap from the perspective of the PISCES range
4. percent overlap from the perspective of the Moyle and Randall polygon
5. change in location of range centroid from Moyle and Randall polygon to PISCES polygon

Using these data, we calculated summary statistics for change in range, change in centroid location, and percent overlap of the Moyle and Randall and PISCES ranges. Further, we plotted $\log_{10}$ change in centroid location versus $\log_{10}$ change in range area to visualize the overall change in current species range between these two datasets.

2.3.2. Data gaps

To identify potential data gaps, we mapped the distribution of observations throughout California. This map contained a single layer with the following query:

select distinct Zone_ID from Observations where Presence_Type in (1,7,9,10).

Presence_Types of 1, 7, 9, and 10 indicate records with empirical observations contained in PISCES to back them up, so this layer will only contain mapping units with at least one empirical observation. This map also configures this layer to use the *Postprocess_Zones* callback (described previously) to execute the query.

select count(*) as col_value from Observations where Observations.Zone_ID = ? and Observations.Presence_Type in (1,7,9,10).

This callback with this query will add an integer field to the generated layer with the number of empirical records for those mapping units, allowing us to symbolize the layer based upon record volume and assess the distribution of empirical observations of species.

3. Results

3.1. PISCES outputs and maps

PISCES contains data for 125 of 129 fish species native to California. Of 274,224 total records imported to date, 49,298 have been quality controlled. Expert verified maps of current California distribution were developed (see Fig. 3) for 67 fish taxa of conservation concern. Maps of historic distribution of were also developed for 42 of these 67 taxa. Using the same data underlying the maps, we have developed alpha diversity per mapping unit − species richness − maps (see Fig. 4a). This map uses the *postprocess_zones* callback. We also developed and ran a separate callback that determines the Jaccard Distance for each mapping unit (see Fig. 4b) as
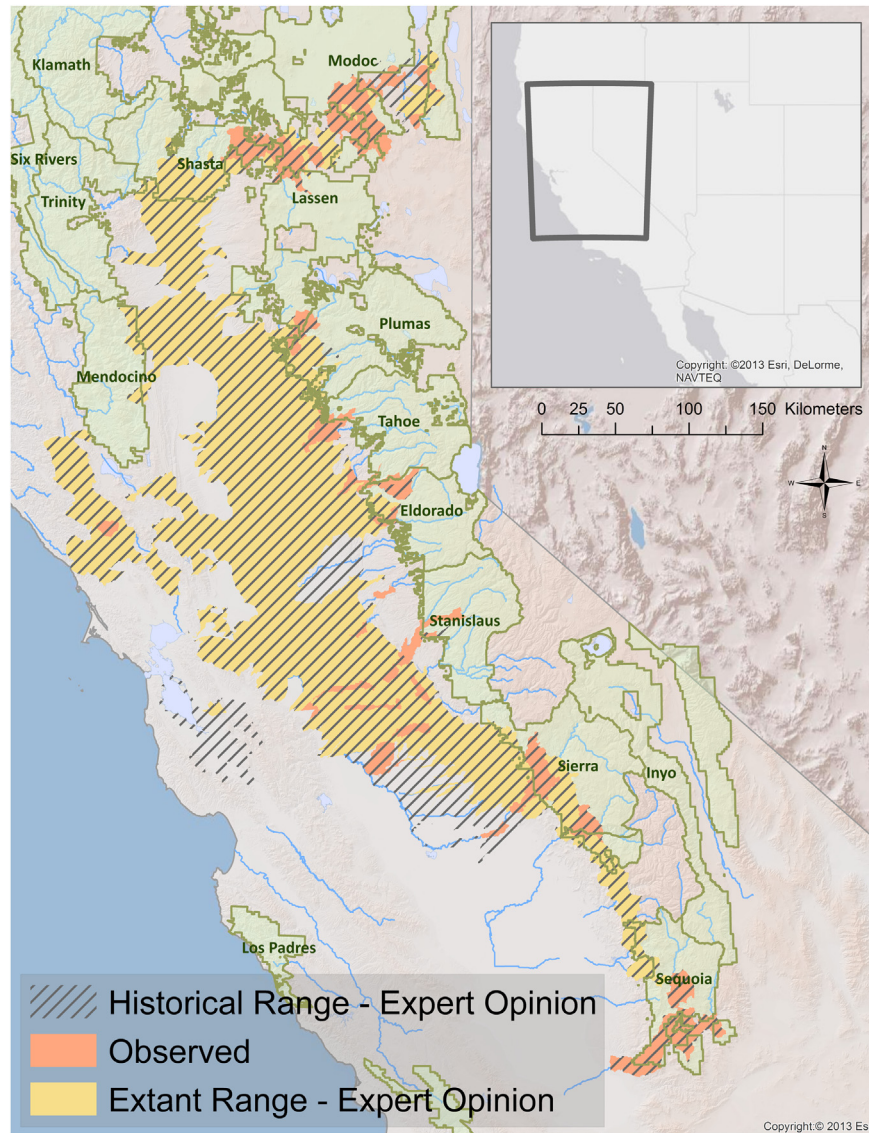
Fig. 3. Complete, expert-verified range map for Hardhead showing layering of multiple presence types.

a demonstration of potential analyses that can be run using the mapping system and data. Jaccard Distance measures similarity of adjacent populations, allowing us to visually and quantitatively assess hotspots of native species diversity and endemism.

3.2. Data comparison with Moyle and Randall

California fish species range maps from Moyle and Randall (1998) were compared with PISCES output range maps for 19 taxa. Table 3 and Fig. 5 contain the full results. The mean change in range area between Moyle and Randall polygons and PISCES generated range maps was 175%, with a median change of 53%. The minimum change was 3.5% and the maximum was 1308% for a taxon with a small mapped range. The first and third quartiles were 23% and 157%, respectively. The average percent overlap between the Moyle and Randall range and the PISCES range was 42% with a sample standard deviation of 23% and a median overlap of 46%. The minimum overlap was 0.5% and the maximum was 75% with a first quartile of 29% and a third quartile of 645%. Average centroid change (i.e., geographic distance between range centers) was

48.5 km with first and third quartiles of 7.9 km and 47.5 km respectively.

3.3. Empirical data availability

We also assessed the distribution of empirical data for native fish taxa in California. The state of empirical knowledge is poor, with only 37.6% (1675) of California HUC12s containing one or more empirical observations of native fish taxa in our system. This analysis is limited to the data available to us for PISCES. Within those 1675 HUC12s, the distribution of empirical observations was skewed to the low end. Fig. 6 shows the number of empirical records per HUC12 in PISCES.

4. Discussion

4.1. Effectiveness of approach − comparison to Moyle and Randall

PISCES allows creation of species distributions based upon the most complete collection of fish data in the state. Moyle and
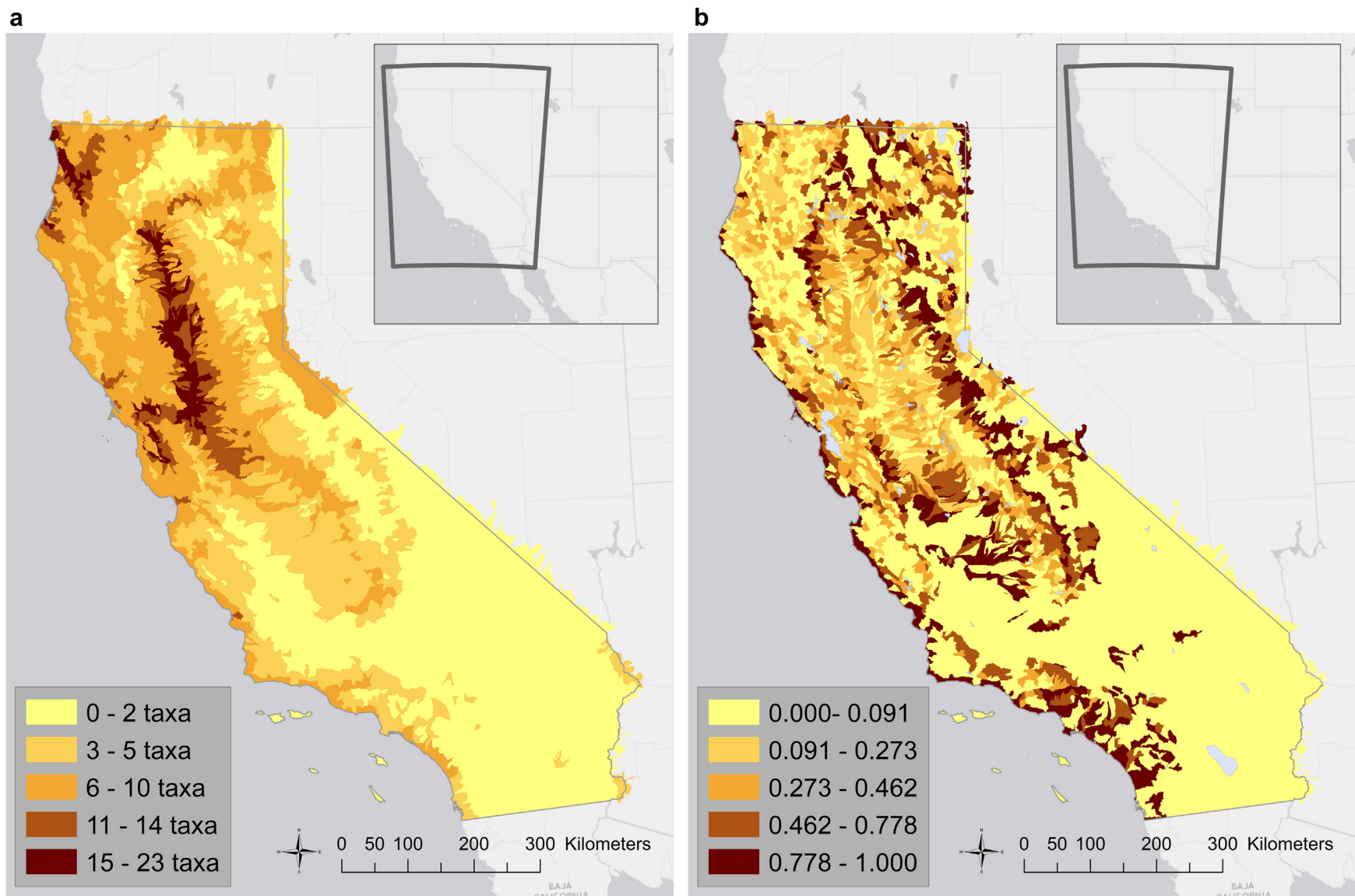
a

b

| 0 - 2 taxa |
| 3 - 5 taxa |
| 6 - 10 taxa |
| 11 - 14 taxa |
| 15 - 23 taxa |

0   50 100   200   300 Kilometers

| 0.000- 0.091 |
| 0.091 - 0.273 |
| 0.273 - 0.462 |
| 0.462 - 0.778 |
| 0.778 - 1.000 |

0   50 100   200   300 Kilometers

Fig. 4. a: Alpha richness (species count) for native fish in California by HUC12. b: Jaccard Distance − a measure of similarity − for native fish in California by HUC12.

Table 3
Metrics of change in range between Moyle and Randall and PISCES for 19 native California taxa.

Common name	Scientific name	Area change (sq. km)	Percent change	Percent overlap	Overlap: Moyle and Randall perspective (%)	Overlap: PISCES perspective (%)	Centroid change (%)	Absolute value area change (sq. km)
Arroyo chub	*Gila orcutti*	−7135.4	−26.0	48.7	57.0	77.0	59.6	7135.4
Central Coast coho salmon	*Oncorhynchus kisutch*	−7710.9	−53.6	27.5	31.6	68.1	116.3	7710.9
Chum salmon	*Oncorhynchus keta*	2396.1	769.2	11.4	99.2	11.4	6.4	2396.1
Goose Lake sucker	*Catostomus occidentalis lacusanserinus*	−134176.7	−99.1	0.5	0.5	56.7	414.0	134176.7
Hardhead	*Mylopharodon conocephalus*	3688.1	7.6	61.5	79.0	73.4	17.5	3688.1
Klamath largescale sucker	*Catostomus snyderi*	1965.9	86.4	47.3	92.0	49.4	19.6	1965.9
Lahontan redside	*Richardsonius egregius*	11448.7	163.8	32.8	89.8	34.0	10.5	11448.7
Mountain sucker	*Catostomus platyrhynchus*	6314.6	150.5	39.7	99.6	39.8	5.7	6314.6
Mountain whitefish	*Prosopium williamsoni*	3891.4	114.2	46.7	100.0	46.7	9.3	3891.4
Northern green sturgeon	*Acipenser medirostris*	4934.6	295.1	12.6	20.1	25.3	2.4	4934.6
Owens sucker	*Cyprinodon radiosus*	6254.5	223.4	30.7	99.4	30.7	35.5	6254.5
Pacific lamprey	*Entosphenus tridentata*	28693.0	32.4	75.3	99.8	75.4	32.0	28693.0
Reticulate sculpin	*Cottus perplexus*	891.8	1307.7	7.1	99.9	7.1	12.2	891.8
Riffle sculpin	*Cottus gulosus*	19688.1	40.8	65.0	94.8	67.4	84.5	19688.1
River lamprey	*Lampetra ayresi*	2894.6	21.2	64.4	86.7	71.5	6.2	2894.6
Sacramento pikeminnow	*Ptychocheilus grandis*	7182.1	5.2	66.1	76.2	83.4	62.4	7182.1
Sacramento tule perch	*Hysterocarpus traskii traskii*	−959.5	−3.6	67.8	79.4	82.3	10.8	959.5
Western brook lamprey	*Lampetra richardsoni*	15842.5	37.4	71.7	99.1	72.1	11.9	15842.5
White sturgeon	*Acipenser transmontanus*	13939.2	167.1	30.9	86.7	32.5	4.8	13939.2
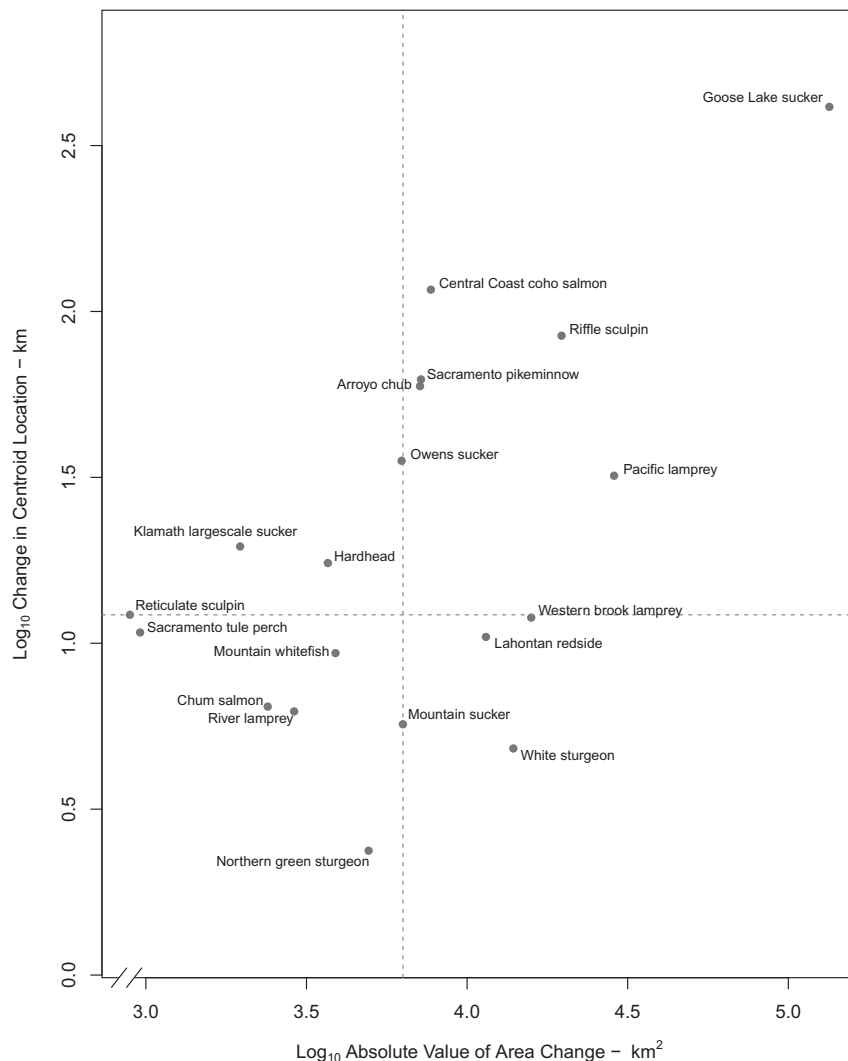
Fig. 5. Change in range between Moyle and Randall and PISCES. *Y* axis is log10 change in range centroid location (km) and *X* axis is log10 change in range area (m).
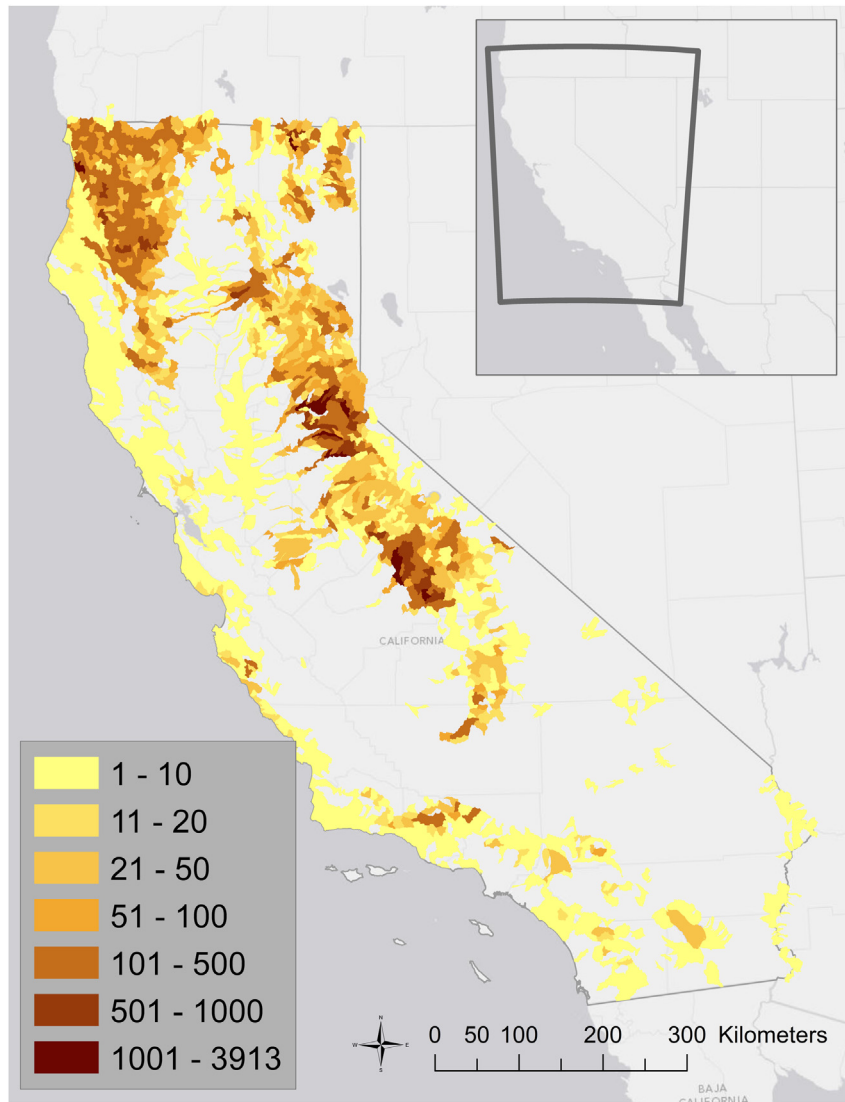
Fig. 6. Distribution and quantity of empirical records in PISCES.

Randall maps provided an important statewide overview but were too coarse to be of much utility at regional and watershed-level scales. Though geographically accurate at broad extents (i.e., statewide to continental — see WWF Freshwater Ecoregions in Abell et al., 2008), the Moyle and Randall maps were incapable of meeting many management needs including ease of revision, geospatial specificity, and multi-species regional analysis. HUC12 mapping units remain coarse for point-based queries, but allow for refined range data at a scale useful to local and regional managers. Further, because we keyed presence records for all species on a common mapping unit that generalizes multiple locations, PISCES runs functions and generates metrics that are infeasible to create or maintain with single-polygon ranges stored in separate files. This structure has the added benefit of maintaining the integrity and traceability of the source data for each record and allows multiple records to verify a species' presence. The mapping interface with SQL query layers and callback functions have allowed PISCES' outputs to be rapidly and repeatedly updated as new datasets became available, imported data was quality controlled, or new needs arose.

The data layering component has proven to be most useful in conveying species ranges. By separating out the empirical observations from expert knowledge, and historic from current or translocated, PISCES maps can at a glance convey current status, confidence levels in the data, and changes to distribution extent over time. By allowing expert knowledge data to be visualized on the same map as other data, PISCES can fill in gaps in empirical data resulting from sampling biases. Further, by obtaining source data from resource agencies or other partners that covers empirical observation and expert knowledge, PISCES refines the known ranges of species.

Each method for measuring change between Moyle and Randall and PISCES captured a distinct difference. Magnitude of range change was given by area change. The percent overlap and centroid change metrics helped us understand whether a range had shifted. A perfect overlap between PISCES maps and Moyle and Randall is impossible because the refined scale of HUC12 boundaries differ from coarse single polygon maps. Data translation appears to cap overlap for large ranges in the comparison set at approximately 75%. Smaller ranges are more susceptible to data translation effects, as can be seen with Reticulate Sculpin (*Cottus perplexus*, 7% overlap), whose range consists of only a few subwatersheds.

Outlier taxa of note are those affected by taxonomic discrepancies. For instance, while Goose Lake Sucker (*Catostomus occidentalis lacusanserinus*) is mapped as a distinct taxon in PISCES,

Moyle and Randall subsumed its range into that of the Sacramento Sucker (*Catostomus occidentalis occidentalis*). Similarly, the Central Coast Coho Salmon Evolutionarily Significant Unit (*Oncorhynchus kisutch*) is now managed as a distinct unit and is mapped in PISCES as such. As a result, the change metrics for these taxa are artificially high, but we chose to retain them in the analysis because they reflect the refinement capabilities of PISCES. A similar issue existed for Northern Green Sturgeon (*Acipenser medirostris*), but we were able to reliably separate the range from Southern Green Sturgeon in the Moyle and Randall data, so those range metrics remain accurate.

4.2. HUC12 advantages and limitations

The use of HUC12 subwatersheds as mapping units was a design requirement for Forest Service mapping needs. As the designated mapping unit for our implementation of PISCES, there were several advantages and limitations.

The primary advantage of HUC12s is the balance they strike between specificity of location and the need for generalized range information. PISCES imported empirical data as specific as GPS point locations from field surveys and as general as the complete range polygons of Moyle and Randall. This entire range of geographic data needed to be incorporated into coherent PISCES range maps. Transformation to HUC12s was a reasonable generalization of data for our purposes.

We also found HUC12s to be limiting. The HUC12 approach was inadequate and nonspecific for especially endemic species with ranges limited to one lake or stream reach. Similarly, exclusions based upon instream barriers, such as waterfalls, were difficult to convey. On the floor of California's Central Valley, where the concept of a watershed breaks down due to both lack of topography and frequent flow diversions, HUC12-based ranges are especially misleading. Finally, managers accustomed to streamline-based maps may find polygon ranges difficult to interpret.

4.3. Additional sources of error

We see three principal sources of error inherent in PISCES: translation errors, mapping unit specificity errors, and expert judgment errors when empirical data availability is low. We made every effort to either control or understand our error, but retained some error. In our data quality control process, we found translation errors to be the easiest to correct. We visually inspected the data and resulting records, removing overestimates resulting from intersecting input GIS datasets with our mapping unit or adding data for underestimates.

Errors related to our choice of mapping unit, the HUC12 watersheds, were harder to correct, depending on their particular manifestation. In Section 4.2, we discussed the limitations of HUC12s for accuracy purposes, but HUC12s can also be problematic for how they convey data. As a unit covering all terrain, not just aquatic resources, HUC12s are useful for conveying a species' full range. The 1:24,000 scale of HUC12 units limit inference and distinction between a given range and supporting specific occurrence. That is, in some cases, range maps can appear to as user as if the data are highly specific for the unit in its entirety, while in reality it conveys a likelihood that varies throughout the unit. While this limit is not inherently a data error, using range data for purposes of species location should be inferred with caution.

The final principal source of error, expert judgment, is most difficult to control. Error in expert judgment is often manifested as differences between the chosen spatial designation of an expert day to day, even in the presence of rules for range change. Future iterations of PISCES will have additional tools to counter check maps for internal consistency and calculate rates of error. Still, due to the data layering and separation of types in PISCES, the error related to this cause is somewhat minimized.

4.4. Empirical vs. expert knowledge data

Range maps were developed for a taxon's entire distribution within California. We attempted to gather all available geospatial data, but we obtained much of our data directly from Forest Service biologists. Despite importation of major statewide datasets from other state, federal, and resource agencies and organizations, the empirical data in PISCES shows a systematic sampling bias clustered on FS lands located in the mountainous regions of the state.

That large areas of the state known to be within the native range of numerous native taxa remain without empirical data points is a failing of sampling effort or of availability of extant data. We believe that many of these areas were sampled before advent of desktop GIS and, as a result, vast amounts of analog data remain in file cabinets of resource agency field offices throughout the state. Up to this point in time the effort and expense to bring historical data into the digital realm has seemed prohibitive. We hope that management tools like PISCES, designed to empower resource managers by making pertinent spatially explicit information available to them, will encourage efforts to convert analog data to digitally and publically accessible datasets.

Further research and efforts will be focused on completing range maps for the rest of California's native fish taxa and dissemination of that data. Current outlets include manual distribution to partners, the PISCES website, and the PISCES public Mercurial repository, but more effective dissemination for use by GIS professionals and state agencies is forthcoming.

5. Conclusion

Resource management and species conservation is hampered by insufficient knowledge of where species occur. We developed a geographic information system dependent database and accompanying software that combines expert knowledge with empirical data to create species range maps from all available data sources. Subsequent encoding of these records relies on a standardized, yet generic, data framework that overcomes issues of spatial scale, temporal discontinuity, data format differences, and regional contexts. As such, PISCES is a robust, easily updated system for tracking, management, and analysis of species ranges.

PISCES contains the most comprehensive collection of data to date on California's native fishes. Current data includes complete expert-verified ranges for 67 of 129 native fish taxa and synthesis maps that include alpha and beta diversities along with an empirical data distribution analysis. Most importantly, both the import and export mechanisms are adaptable so that as new data become available or new needs arise the system is able to incorporate diverse data sources and output the necessary spatial and tabular products. Although intended to solve acute management problems associated with California's sensitive fish species on Forest Service lands, PISCES – an enhanced GIS with data provenance and transaction capabilities – was easily adapts to the needs of other environmental phenomena, regions, and resource management issues.

Acknowledgments

bringing PISCES into being. We also thank Ron Yoshiyama, Molly Stevens, Bjorn Erickson, and Joe Kiernan at UC Davis, Kurt Fesenmeyer at Trout Unlimited, Steve Lindley and Sarah Paddock at the National Oceanographic and Atmospheric Administration, and United States Forest Service Region 5 field office staff.

Appendices

Supplementary data can be obtained from https://bitbucket.org/nickrsan/pisces/downloads/EMS_PISCES_Supplement.zip. Supplementary data includes:

1. A full snapshot of the PISCES software and data as well as data exports of the 19 species analyzed in this paper in PNG, PDF, and ArcMap Document format (for ArcGIS 10.0 or above).
2. Database relationships diagram.
3. Documentation of the database
4. Partial documentation for using the PISCES software (also available at http://pisces.ucdavis.edu/doc)
5. The code used for the analysis of change between Moyle and Randall and PISCES
6. ESRI feature classes of the diversity and empirical data used in Figs. 5 and 6

Appendix A: Data flow within and through PISCES

Data processing and flow in PISCES varies based upon source and type. Existing datasets will be processed by an input filter, which will perform the ETL, while expert knowledge is processed and entered by an operator via an ArcGIS toolbox tool. Additionally, since generated maps and matrices are highly customizable, the code execution varies. This section follows data additions and modifications from the perspective of the human running and interacting with PISCES (the Operator).

Dataset ETL

The first stage of dataset ETL requires human setup and configuration. First, the Operator copies the dataset to the geodatabase that PISCES checks for new data. The Operator performs any conversions required to make the dataset geodatabase compatible (e.g.: CSV to geodatabase table). The Operator then adds a record to the table in that database that tracks metadata regarding importation of new data (PISCES/inputs/new_data.mdb/NewData). This record requires specification of the taxon of the dataset, or the keyword "filter," indicating that the input filter will determine the taxon for each data record it encounters. "Filter" is the most common method used in PISCES as most datasets are multi-taxa.

The Operator also specifies the input filter for the new dataset, creating one if need be. Input filters are a combination of translation values for taxa ids and the code class that handles the import. The translation values are stored in table Alt_Codes in the main PISCES database, with a foreign key to the input filter ID. These values may be entered manually within Microsoft Access, or the Operator may use a tool in the PISCES ArcGIS toolbox that preloads values for the dataset, requiring the Operator to only enter the PISCES species IDs into the records. The Operator specifies the code class that handles the import in the input filter definition. This many input filters to many code classes structure allows for reuse of code classes and generic import filters that handle a broad set of translations common to species presence datasets.

Once the Operator has specified the input filter on the dataset's record in the NewData table, the final step before importing is to specify the field mapping for datasets containing multiple taxa. In new_data.mdb/FieldMapping, the Operator creates a record for each field to map, with a foreign key to the dataset record in the same database. Each primary field in a PISCES observation record may be mapped to a field in the dataset. The notes field allows a semicolon separated list of fields, indicating that PISCES should extract the data from all of those fields into a single string for storage. Each record can also be marked as required, indicating that PISCES should skip the record if it does not include the data. Finally, each field mapping can include a specification of a handler function to allow the Operator to tailor a generic input filter to a dataset with minor variations. The handler functions act as overrides that process the value for a field and can transform it, with many input filters having default handler functions for the primary PISCES fields.

After these configurations have been made, the operator opens a command prompt to the PISCES software directory (PISCES/scripts/PISCES) and runs the command "python main.py import." PISCES will then attempt to import data for all configured datasets that have not completed importation, retrying any that previously failed.

Common PISCES ETL operations

PISCES relies primarily on a few operations to transform datasets to HUC12s. For data that is already in a spatial format, PISCES conducts a spatial join to intersect the dataset with the HUC12 dataset and transfer the HUC12 attributes to the source records. PISCES can then read the HUC12 ID value from each record in the dataset and store that value as the species location.

Most translation steps involve variations on this method. Some datasets require field level translation in addition to the species ID translation. Others have multiple species per record, requiring the input filter to convert the dataset to a single-species per record format before the spatial join. Many datasets we processed were purely tabular data with coordinates, for which PISCES creates a spatial dataset from those coordinates to run the spatial join on.

Editing and quality control workflow

For textual data, reports, paper datasets, and expert knowledge, we used a different suite of tools to add information to PISCES. The PISCES toolbox (located at PISCES/tbx/PISCES.tbx) is a set of ArcGIS tools that enable the Operator to have various interactions with the database. The "Add or Modify Data" tool, shown in Fig. 7, is the primary method for range editing. To use this tool, the Operator opens an existing range map export for a single species and determines the HUC12s to modify. Using the ArcGIS selection tools, the Operator makes a selection on the appropriate HUC12 layer, then opens the Add or Modify Data tool, setting the name of the layer with the selection as the value of the input layer. The Operator can then specify the species these modifications pertain to, and the operation − add, remove, or transfer. Add and remove are straightforward: they add or remove the selected HUC12s from a species range (subject to caveats specified in another parameter). Transfer is meant to deal with situations where PISCES tracks subspecies, but an input dataset did not. It allows the Operator to move records from an incorrect taxon to the correct one. For example, an input dataset might specify the taxon as "Chinook," but PISCES tracks multiple evolutionarily significant units (ESUs) of Chinook, each with differing ranges. After reviewing the relevant information that enables assignment to a particular ESU or subspecies, the Operator can transfer the records to the correct taxon.

The species and selection are the required parameters, but from the dialog box the Operator can tweak the metadata and extra parameters for the changes. The dialog takes a message parameter for the Operator to indicate justification or sources for the edits and any other notes about the data. Additionally, a "subset" parameter lets an operator familiar with SQL *where* clauses tailor the impact of

Fig. 7. PISCES manual editing tool for expert additions, deletions, and modifications. From the PISCES toolbox in PISCES/tbx/PISCES.tbx/Modification/Add or Modify Data.

the changes by adding limiters, such as specific observation types, observation sets, and other constraints. Checkboxes are also available for directly limiting the changes to specific observation types, with the default limitation being that only expert opinion data is affected. Finally, the Operator can specify the input filter and observation set for new records to be assigned to and whether or not PISCES should return a new expert opinion layer into ArcMap so the changes can be viewed immediately.

After the Operator fills in the parameters and runs the tool, PISCES builds a database query to make the changes, with a placeholder (bind value) for the HUC12 IDs. It then loops through each HUC12 ID in the selection on the input layer and executes the query for each one, making the changes to the database. Finally, it adds a transaction record to the Transactions table in the database with the parameters to the tool and the result.

Mapping setup

Sections 2.1.4 and 2.2 detail the key internals of the mapping code. The operator can specify the inputs to this code in the tables *defs_Query_Sets*, *Map_Queries*, and *Query_Bind* in pisces.mdb.

References

Abell, R., Thieme, M.L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., Coad, B., Mandrak, N., Balderas, S.C., Bussing, W., Stiassny, M.L.J., Skelton, P., Allen, G.R., Unmack, P., Naseka, A., Ng, R., Sindorf, N., Robertson, J., Armijo, E., Higgins, J.V., Heibel, T.J., Wikramanayake, E., Olson, D., Lopez, H.L., Reis, R.E., Lundberg, J.G., Perez, M.H.S., Petry, P., 2008. Freshwater ecoregions of the world: a new map of biogeographic units for freshwater biodiversity conservation. Bioscience 58 (5), 403–414.

Biogeographic Data Branch, C.D.o.F.a.G., 2011. In: Biogeographic Data Branch, C.D.o.F.a.G. (Ed.), California Natural Diversity Database (CNDDB). August 2011 ed: http://www.dfg.ca.gov/biogeodata/cnddb/.

BIOS, 2011a. In: (BIOS), C.D.o.F.a.G.B.I.a.O.S. (Ed.), Aquatic Organisms – Osteichthyes [ds193]. http://bios.dfg.ca.gov.

BIOS, 2011b. In: (BIOS), C.D.o.F.a.G.B.I.a.O.S. (Ed.), Inland Fisheries [ds192]. http://bios.dfg.ca.gov.

BIOS, 2011c. In: (BIOS), C.D.o.F.a.G.B.I.a.O.S. (Ed.), Inland Fisheries [ds195]. http://bios.dfg.ca.gov.

Dudgeon, D., Arthington, A.H., Gessner, M.O., Kawabata, Z.-I., Knowler, D.J., Lévêque, C., Naiman, R.J., Prieur-Richard, A.-H.l.n., Soto, D., Stiassny, M.L.J., Sullivan, C.A., 2006. Freshwater biodiversity: importance, threats, status and conservation challenges. Biol. Rev. 81, 163–182.

EMAP, 2006. Vertebrates data. In: U.S.E.P.D.E (Ed.), Environmental Monitoring and Assessment Program (EMAP). http://www.epa.gov/emap/html/data/index.html.

Graham, C.H., Hijmans, R.J., 2006. A comparison of methods for mapping species ranges and species richness. Glob. Ecol. Biogeogr. 15 (6), 578–587.

Grantham, T.E., Merenlender, A.M., Resh, V.H., 2010. Climatic influences and anthropogenic stressors: an integrated framework for streamflow management in Mediterranean-climate California, U.S.A. Freshw. Biol. 55, 188–204.

Guzman, J.A., Moriasi, D.N., Chu, M.L., Starks, P.J., Steiner, J.L., Gowda, P.H., 2013. A tool for mapping and spatio-temporal analysis of hydrological data. Environ. Model. Softw. 48 (0), 163–170.

Hijmans, R.J., Garrett, K.A., Huaman, Z., Zhang, D.P., Schreuder, M., Bonierbale, M., 2000. Assessing the geographic representativeness of genebank collections: the case of Bolivian wild potatoes. Conserv. Biol. 14 (6), 1755–1765.

Horsburgh, J.S., Tarboton, D.G., Piasecki, M., Maidment, D.R., Zaslavsky, I., Valentine, D., Whitenack, T., 2009. An integrated system for publishing environmental observations data. Environ. Model. Softw. 24 (8), 879–888.

Hurlbert, A.H., White, E.P., 2005. Disparity between range map- and survey-based analyses of species richness: patterns, processes and implications. Ecol. Lett. 8 (3), 319–327.

Jetz, W., McPherson, J.M., Guralnick, R.P., 2012. Integrating biodiversity distribution knowledge: toward a global map of life. Trends Ecol. Evol. 27 (3), 151–159.

Kaschner, K., Rius-Barile, J., Kesner-Reyes, K., Garilao, C., Kullander, S.O., Rees, T., Froese, R., 2010. AquaMaps: Predicted Range Maps for Aquatic Species. World Wide Web Electronic Publication.

Katz, J., Moyle, P., Quiñones, R., Israel, J., Purdy, S., 2012. Impending extinction of salmon, steelhead, and trout (Salmonidae) in California. Environ. Biol. Fishes, 1–18.

Marchetti, M.P., Light, T., Moyle, P.B., Viers, J.H., 2004. Fish invasions in California watersheds: testing hypotheses using landscape patterns. Ecol. Appl. 14 (5), 1507–1525.

Millenium Ecosystem Assessment, 2005. Ecosystems and Human Well-being: Synthesis. Island Press, Washington, DC.

Moyle, P., Israel, J.A., Purdy, S.E., 2008. Salmon, Steelhead and Trout in California: Status of an Emblematic Fauna. UC Davis Center for Watershed Sciences, Davis, CA.

Moyle, P.B., 2002. Inland Fishes of California. University of California Press.

Moyle, P.B., Katz, J.V.E., Quiñones, R.M., 2011. Rapid decline of California's native inland fishes: a status assessment. Biol. Conserv. 144 (10), 2414–2423.

Moyle, P.B., Katz, J.V.E., Quiñones, R.M., 2013. Fish Species of Special Concern. California Department of Fish and Wildlife (in press).

Moyle, P.B., Randall, P.J., 1998. In: Environment, U.D.I.C.f.t. (Ed.), California Native Fish Distributions.

Moyle, P.B.Y.,R.M., Williams, J.E., Wikramanayake, E.D., 1995. Fish Species of Special Concern in California. California Department of Fish and Game.

Naiman, R.J., Dudgeon, D., 2011. Global alteration of freshwaters: influences on human and environmental well-being. Ecol. Res. 26 (5), 865–873.

Nel, J.L., Roux, D.J., Abell, R., Ashton, P.J., Cowling, R.M., Higgins, J.V., Thieme, M., Viers, J.H., 2009. Progress and challenges in freshwater conservation planning. Aquat. Conserv. Marine Freshw. Ecosyst. 19 (4), 474–485.

Olson, A., 1997. Fish Species Range.

Pittock, J., Hansen, L.J., Abell, R., 2008. Running dry: freshwater biodiversity, protected areas and climate change. Biodiversity 9 (3–4), 30–38.

Rangel, T.F., Diniz-Filho, J.A.F., Bini, L.M., 2010. SAM: a comprehensive application for spatial analysis in macroecology. Ecography 33 (1), 46–50.

Schick, R.S., Edsall, A.L., Lindley, S.T., 2005. Historical and Current Distribution of Pacific Salmonids in the Central Valley, CA. NOAA Technical Memorandum NMFS-SWFSC-369 25.

Scott, J.M., Csuti, B., Jacobi, J.D., Estes, J.E., 1987. Species richness. BioScience 37 (11), 782–788.

Souza Muñoz, M., Giovanni, R., Siqueira, M., Sutton, T., Brewer, P., Pereira, R., Canhos, D., Canhos, V., 2011. openModeller: a generic approach to species' potential distribution modelling. GeoInformatica 15 (1), 111–135.

Strayer, D.L., Dudgeon, D., 2010. Freshwater biodiversity conservation: recent progress and future challenges. J. N. Am. Benthol. Soc. 29 (1), 344–358.

Thieme, M., Lehner, B., Abell, R., Hamilton, S.K., Kellndorfer, J., Powell, G., Riveros, J.C., 2007. Freshwater conservation planning in data-poor areas: an example from a remote Amazonian basin (Madre de Dios River, Peru and Bolivia). Biol. Conserv. 135 (4), 484–501.

TU, 2011. In: Unlimited, T. (Ed.), Conservation Success Index. http://www.tu.org/science/conservation-success-index.

Underwood, E.C., Viers, J.H., Klausmeyer, K.R., Cox, R.L., Shaw, M.R., 2009. Threats and biodiversity in the mediterranean biome. Divers. Distrib. 15 (2), 188–197.

Viers, J.H., Rheinheimer, D.E., 2011. Freshwater conservation options for a changing climate in California's Sierra Nevada. Mar. Freshw. Res. 62, 266–278.

Villa, F., Athanasiadis, I.N., Rizzoli, A.E., 2009. Modelling with knowledge: a review of emerging semantic approaches to environmental modelling. Environ. Model. Softw. 24 (5), 577–587.

Vörösmarty, C.J., Green, P., Salisbury, J.S., Lammers, R.B., 2000. Global water resources: vulnerability from climate change and population growth. Science 289 (5477), 284–288.

Vorosmarty, C.J., McIntyre, P.B., Gessner, M.O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S.E., Sullivan, C.A., Liermann, C.R., Davies, P.M., 2010. Global threats to human water security and river biodiversity. Nature 467 (7315), 555–561.

Wilson, K.A., Underwood, E.C., Morrison, S.A., Klausmeyer, K.R., Murdoch, W.W., Reyers, B., Wardell-Johnson, G., Marquet, P.A., Rundel, P.W., McBride, M.F., Pressey, R.L., Bode, M., Hoekstra, J.M., Andelman, S., Looker, M., Rondinini, C., Kareiva, P., Shaw, M.R., Possingham, H.P., 2007. Conserving biodiversity efficiently: what to do, where, and when. Plos Biol. 5 (9), 1850–1861.