

Durability of environment–recruitment relationships in aquatic ecosystems: insights from long-term monitoring in a highly modified estuary and implications for management

N. Tamburello,^{1*} B. M. Connors,² D. Fullerton,³ C. C. Phillis³

¹ESSA Technologies Ltd., Vancouver, British Columbia, Canada

²Quantitative Assessment Methods Section, Institute of Ocean Sciences, Fisheries and Oceans Canada, Sidney, British Columbia, Canada

³Bay-Delta Initiatives, Metropolitan Water District of Southern California, Sacramento, California

Abstract

The environment can strongly influence the survival and population dynamics of aquatic organisms. Our understanding of these relationships, typically based on simple linear regression, underpins many contemporary resource management decisions. However, such relationships can break down over time as ecosystems change. Even when durable, relationships may not be very useful for management if they exhibit high variability, context dependency, or nonstationarity. Here, we systematically review the literature to identify trends across environment–recruitment relationships for aquatic taxa from California’s San Francisco Bay and Sacramento–San Joaquin Delta Estuary. This delta is one of the most heavily modified aquatic ecosystems in North America, and home to numerous species of concern whose relationships with the environment inform regulatory actions and constraints. We retested 23 of these relationships spanning nine species using data that have accumulated in the years since they were first published (9–40 additional years) to determine their durability. Most relationships remained the same or stronger in direction and magnitude but showed declining predictive power with the addition of new data, particularly for older relationships that had not adjusted for recognized regime shifts in the system through the use of step changes or data splitting. Constantly refining these relationships may give the appearance of durability, but limit their practical value as policy tools when the present or future state of the ecosystem is unknown. We conclude by synthesizing emerging insights from the literature on best practices for the analysis, use, and refinement of environment–recruitment relationships to inform better decision making in dynamic ecosystems.

The environment can have a profound and complex influence on aquatic organisms and their population dynamics (e.g., Szuwalski et al. 2015). Understanding when and how the environment influences the survival, abundance, and

recruitment of fishes and other aquatic organisms has long fascinated and perplexed fish and fisheries scientists and managers (Hjort 1914; Cushing 1995; Jacobson and MacCall 1995). Knowing how the environment influences the population dynamics of fishes is of general ecological interest because of the light it can shed on, for example, the relative influence of bottom-up and top-down control in ecosystems. Quantifying how the environment influences recruitment can also, in theory, help inform fisheries and improve management. Given the now pervasive influence of humans over the world’s aquatic ecosystems (e.g., Halpern et al. 2015), understanding when and how the environment—and human influences on it—affects the dynamics and abundances of aquatic organisms is critical to many decisions in natural resource management.

Relationships between the environment and recruitment, defined here as any relationship between the number of individuals in a population (or their survival rate) and their

*Correspondence: ntamburello@essa.com

Nataschia Tamburello and Brendan M. Connors contributed equally to this work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Additional Supporting Information may be found in the online version of this article.

Special Issue: Long-term Perspectives in Aquatic Research.
Edited by: Stephanie Hampton, Matthew Church, John Melack and Mark Scheuerell.

environment (e.g., river flow, pesticide concentration, and water temperature), may break down over time. In a now classic review of fisheries literature, Myers (1998) found that only 22 of 77 environment–recruitment relationships still held after being re-examined with new data. The relationships that were most likely to stand the test of time were those with temperature at the limit of a species' range, where the influence of physiological tolerance thresholds outweighs that of more complex ecological interactions. Even when such relationships are reliable, they may not be very useful for informing management if the described relationship is characterized by high variability, context dependency, or nonstationarity as is often the case with recruitment data. Nonetheless, relationships between the environment (which we define broadly as both natural and those aspects under human control) and fish recruitment are commonly used to inform contemporary resource management decision-making and the conservation of aquatic species. Considering this, and the potential for environment–recruitment relationships to break down over time, there is a pressing need for guidance on best practices for the analysis, use, and refinement of environment–recruitment relationships to inform decision making in natural resource management.

California's San Francisco Bay and Sacramento–San Joaquin Delta Estuary (hereafter “Bay Delta”) (Fig. 1) is an ideal system in which to examine the durability and usefulness (i.e., predictive power) of environment–recruitment relationships and their implications for decision making. The Bay Delta system has been continuously monitored in a systematic manner for long periods of time (e.g., since at least 1959 for fish populations, Stevens 1977), it is heavily altered and managed, and it is home

to numerous endangered species and associated regulatory actions and constraints including some which are based upon environment–recruitment relationships (reviewed in Kimmerer 2004). Numerous relationships have been described for taxa in the Bay Delta, but relatively few have been revisited to test whether these correlations are reliable in the face of new data.

Here, we systematically review the peer-reviewed and gray literature to identify environment–recruitment relationships for Bay Delta taxa. We then reanalyze a subset of these relationships where new data are available to quantify the extent to which the relationships still hold when confronted with new data. Finally, we synthesize emerging insights from the literature on best practices for the analysis, use, and refinement of environment–recruitment relationships to inform decision making in natural resource management.

Methods

Study area

The Bay Delta is made up of a large interior delta formed by the Sacramento and San Joaquin Rivers feeding into a series of basins separated by narrow, deep tidal channels, which flow into a seaward region and ultimately into the San Francisco and San Pablo Bays which are connected to the Pacific Ocean (Fig. 1A). The Bay Delta is one of the most heavily modified estuaries in the United States, and is strongly influenced by state and federal water project operations. Two pumping facilities export delta inflows to meet metropolitan and agricultural water needs. Water exports can affect fish directly through entrainment into the pumping facilities, and indirectly through the influence of reduced flows on a wide range of abiotic and biotic variables (Fig. 1B). Established quantitative relationships

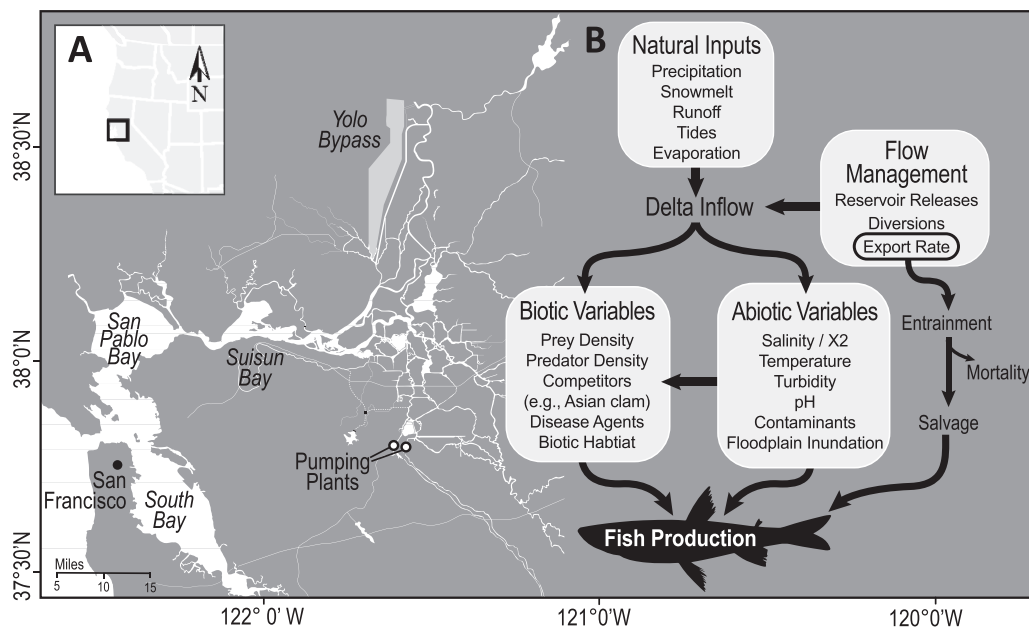


Fig. 1. Map of the Bay Delta region and key geographical features (A), along with an overlay of a conceptual model of key abiotic and biotic drivers known to influence fish production in this system (B). Adapted from Delta Independent Science Board (2015).

between flow characteristics and fish abundance, survival, and migration underpin regulatory decisions regarding levels of allowable exports of river flows and minimum outflow from the Bay Delta that maintain fish production and habitat within acceptable bounds (California Department of Fish and Wildlife [CDFW] 2016).

Literature review

We carried out a literature search of peer-reviewed publications, gray literature, and government agency reports to identify published examples of relationships between environmental variables and any population parameters of aquatic Bay Delta species. Our initial search was carried out using Google Scholar and various combinations of search terms based on variables expected to influence organism abundance as described in comprehensive reviews of Bay Delta ecology (e.g., Kimmerer 2004; California Department of Water Resources and U.S. Bureau of Reclamation 2016). These search terms included “fish,” “invertebrate,” “abundance,” “survival,” “entrainment,” “migration,” “environment–recruitment,” “environmental variable,” “flow,” “conductivity,” “turbidity,” “prey density,” and “X2” (the latter being distance in kilometers from the Golden Gate up the axis of the estuary to where the tidally averaged near-bottom salinity is 2 ppt or 2×10^{-9} mg L⁻¹; Jassby et al. 1995). The resulting set of publications and reports was then supplemented by consulting regulatory documents to identify additional relationships that underpin contemporary management decisions.

We created a catalogue of all identified publications to document their various characteristics, including publication year, focal species, the number of relationships examined, and whether the publication is cited in regulatory documents as informing management decisions (included here as Supporting Information). Because most of the government agency and gray literature reports were review documents that reproduced results from primary literature, we chose to focus further investigations only on the peer-reviewed literature. For each peer-reviewed study in our catalogue, we extracted each individually reported relationship into a second catalogue and documented characteristics including the focal species, predictor and response variables, type of analysis, timeframe, primary data source, and reported model outputs including intercept and slope parameters, R^2 , p -value, Akaike information criterion (AIC) score, and others. This catalogue of relationships was used both to carry out a qualitative analysis of trends in the study of such relationships in the Bay Delta, and to select a subset of environment–recruitment relationships suitable to retesting to quantify the extent to which the relationships still hold when confronted with new data.

Relationship selection criteria

We developed three tiers of criteria used to screen the full set of published environment–recruitment relationships to identify those suitable for reanalysis.

Tier 1 criteria excluded relationships for which reanalysis would be impractical because of (1) inability to obtain new data for reanalysis because of reliance on either one-time experiments (e.g., paired releases of tagged fish), data collection programs that have since ended, or second-order variables derived via complex integration of many other environmental variables; (2) use of analytical methods that would be impractical to replicate for a review study of this scale (e.g., whole ecosystem simulation models) or which make it difficult to compare strength and statistical support across relationships (e.g., nonlinear correlations such as GAMs and rank analyses); or (3) because the published relationship is so recent there would be few additional data points available (i.e., we excluded relationships if the number of years of new data were less than the number in the original time series or less than 10 years, whichever was less). In practice, this resulted in retention of linear regressions based on data from long-term monitoring programs with publicly available data sets, which formed a large majority (about 75%) of all the published relationships in our dataset examined prior to the application of screening criteria.

Tier 2 criteria screened out relationships with little or no statistical support in the initial analysis, given that the objective of our work was to evaluate the extent to which established relationships hold over time. For this purpose, we defined little or no statistical support as $p > 0.1$, or $\Delta\text{AIC} > 2$ from the top model or, where neither is reported in older publications, the absence of a fitted regression line and equation on a scatter plot of the relationship. While we recognize that selecting relationships previously identified as significant introduces a bias toward durability, the rarity of studies publishing nonsignificant relationships or full multivariate model selection approaches in this system makes it difficult to systematically test the opposite case, where previously nonsignificant relationships become significant with the addition of more data.

Lastly, Tier 3 criteria were used to reduce redundancy in the remaining set of relationships by selecting only a single relationship for each unique species–variable combination for reanalysis. When more than one relationship existed for the same species–variable combination relying on the same source data (e.g., delta smelt and X2), the following subcriteria were used to select a single relationship that was (1) most recently updated among the set to avoid reproducing past retests; (2), the longest time-series among the set for greater statistical power; (3) based on indices of abundance rather than extrapolated abundance estimates to reduce propagation of error, and (4) where all relationships in the set are statistically significant, the relationship with the strongest support based on variation explained (e.g., via R^2) or other rationale provided by the authors. We considered similar relationships using different source data (e.g., delta smelt_(fall midwater trawl) ~ X2 and delta smelt_(Bay Study midwater trawl) ~ X2) distinct and therefore were retained because of known differences in methodology,

target life stage, and conclusions that can be drawn from alternative source surveys.

The full catalogue of original and screened publications is available in the Supporting Information.

Data sources

For each relationship retained for reanalysis, we sought to obtain the same data used in the original analysis from the source identified by the authors. In some cases, this data was available directly, and in others, annual means of data needed for analysis were derived from raw data sets using the same or similar methods originally described by the authors, including applying data transformations (indicated using superscripts in Table 1, with more detail available in Supporting Information Table S1). In some cases, we were not able to reproduce the data using the methods described by the authors and either used a modified method to obtain reasonably similar data or, when this was not possible, excluded the relationship from our analysis. Population data for these relationships were derived from the California Department of Fish and Wildlife's fall midwater trawl (FMWT), summer totnet survey (TNS), beach seine surveys, and salvage surveys; the Bay Study otter trawl (Bay OT) and midwater trawl (Bay MWT) surveys; and the U.S. Fish and Wildlife Service's Chinook salmon trawls. Environmental data was derived primarily from Hutton et al. (2015) and calculations therein for X2, from the California Department of Water Resources DayFlow data portal for flows, from environmental data collected alongside population data as part of the FMWT survey, and in some cases directly from the study authors (Table 1, Supporting Information).

Analysis

We conducted a quantitative reanalysis of selected environment–recruitment relationships. Source data for the analyses was transformed (as per original analyses; see Table 1), subset to the relevant timeframes, and standardized (by subtracting the mean and dividing by the standard deviation) to allow estimated relationships to be comparable across species and environmental variables. All relationships selected for reanalysis were based on linear correlations and so scaled data was analyzed using ordinary least squares regression with an abundance measure as the response variable and one or more environmental factors as the predictor variable. None of the ultimately selected relationships included interactions among variables.

Several relationships selected for reanalysis (e.g., Kimmerer et al. 2009) included a step-change or data-splitting before and after the introduction of the invasive Asian clam (*Corbula amurensis*) into the Bay Delta in 1986. Introduction of this clam is thought to be responsible for a sudden and substantial decline of zooplankton and, subsequently, of fish in the Bay Delta (Kimmerer et al. 1994). Where published analyses found statistical evidence for a split or step change in 1987 (using a dummy

variable), we chose to reanalyze the relationship using only data after this cut-off both to simplify the analyses and because prior data could be considered irrelevant for interpreting how relationships have changed in the context of additional data.

For each relationship, we quantified the strength (R^2 ; coefficient of determination) and magnitude (slope; in standardized units) of the relationships based on the original and updated time series and then compared them to quantify the extent to which the relationships still held when confronted with new data. We also quantified prediction error using unscaled original and extended time series for each relationship to characterize how well the observed relationships would be expected to predict future (out of sample) observations and to provide an indication of how useful the relationship may be from a decision making and management perspective. Prediction error was estimated as the normalized root mean squared prediction error (CV_n):

$$CV_n = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{\hat{y}_i - y_i}{1 - h_i} \right)^2}}{\frac{1}{n} \sum_{i=1}^n y_i} \quad (1)$$

where h_i is the diagonal element of the operator matrix that produces the least squares fit (i.e., hat matrix). This measure of prediction error can be interpreted as the percent error in future predictions relative to the average observed abundance for a given relationship. For example, a prediction error of 100% would mean that the relationship allows us to predict future abundances to within $\pm 50\%$ of the mean predicted abundance.

We compared the raw data, parameters, and fit of relationships using the original time series to those reported in the original publications to ensure that our approach successfully replicated the previously published relationships before proceeding to retesting. Instances where we were not able to reproduce past relationships were not considered further (detailed in Table 1).

Our retests of environment–recruitment relationships using data that has accumulated since a relationship was first established has the potential to be biased by an imbalance in the number of pre- to post-retest observations. In instances where there are many (e.g., 30) years of pre-retest observations and only a few (e.g., 10) post-retest observations it is possible that the pre-retest observations obscure what is otherwise a weakening or different relationship in the post-retest observations. To quantify the extent to which this was the case with the relationships we retested, we carried out a secondary balanced analysis where we randomly subsampled the pre-retest data so that there was an equal number of pre- and post-retest data points, and then, as above with the full time series, quantified the strength and magnitude of the relationship based on the subsetted data. We repeated this exercise 1000 times for each relationship, and then compared the median strength and magnitude of the relationship based on the original (subsetted) and updated time series.

Table 1. Summary of previously published environment–recruitment relationships retested with additional years of data.

Source	Species	Response variable	Predictor variable(s)	Time series (original and/extended)
Bennett et al. (2005)	Delta smelt	TNS index ^a	Log total salvage ^a	1979–2002/2016
Brandes et al. (2006)	Fall Chinook salmon	Chinook salmon fry m ⁻³ at Chipps Island (Jan–mar)	Mean daily Feb flow at Freeport (cfs)	1985–2005/2011
	Fall Chinook salmon	Chinook salmon fry m ⁻³ at Sacramento (Apr–Jun)	Mean daily Feb flow at Freeport (cfs) ^c	1985–2005/2011
	Fall Chinook salmon	Chinook salmon fry m ⁻³ at Chipps Island (Apr–Jun)	Mean daily Apr to Jun flow at Rio Vista (cfs)	1978–2005/2015
Kimmerer et al. (2009)	American shad	FMWT abundance index ^a	Spring X2 (km)	1988–2007/2014
	Bay goby	Bay study OT abundance index ^a	Spring X2 (km)	1988–2007/2014
	Bay shrimp	Bay study OT abundance index ^a	Spring X2 (km)	1988–2007/2012
	Longfin smelt	Abundance index (three models: FMWT; bay study MWT; OT) ^a	Spring X2 (km)	1988–2007/2014
	Sacramento splittail	Abundance index (FMWT) ^a	Spring X2 (km)	1988–2007/2014
	Starry flounder	Abundance index (bay study OT) ^a	Spring X2 (km) (1 yr lag)	1988–2007/2013
Sommer et al. (1997)	Sacramento splittail	Abundance index (FMWT) ^b	Delta outflow (cfs) ^a	1967–1995/2014
	Sacramento splittail	Abundance index (FMWT) ^b	Days Yolo bypass inundated	1967–1995/2014
Sommer et al. (2007)	Longfin smelt	Abundance index (FMWT) ^a	Flows Jan to June (cfs) ^a	1988–2006/2014
Stevens (1977)	Striped bass	Abundance index (TNS 38 mm index) ^a	Mean daily Delta outflow (Jun–Jul) (cfs) ^a	1959–1970/2016
	Striped bass	Abundance index (TNS Suisun Bay) ^a	Mean daily Delta outflow (Jun–Jul) (cfs) ^a	1959–1976/2011
Feyrer et al. (2007)	Delta smelt	Abundance index (TNS)	Mean annual conductance (μS cm ⁻¹ ; FMWT)	1987–2004/2016
			Fall stock abundance index (FMWT)	
	Delta smelt	Abundance index (TNS)	Mean annual Secchi depth (m) (FMWT)	1987–2004/2016
			Fall stock abundance index (FMWT)	
	Delta smelt	Abundance index (TNS)	Mean annual conductance (μS cm ⁻¹) (FMWT)	1987–2004/2016
			Mean annual Secchi depth (m) (FMWT)	
			Fall stock abundance index (FMWT)	
Miller et al. (2012)	Delta smelt	Fall-to-summer survival ^c (July TNS index/previous year's FMWT index)	Previous fall abundance (FMWT) ^d	1972–2006/2014
			Previous–previous fall abundance (FMWT) ^d	
			Average <i>E–P</i> density, Apr–Jun (#/m ³) ^d	
			Proportional entrainment	
			Average <i>E–P</i> density, Jan–Mar (#/m ³) ^d	

(Continues)

Table 1. Continued

Source	Species	Response variable	Predictor variable(s)	Time series (original and/extended)
	Delta smelt	Summer-to-fall survival ^c (FMWT index/July TNS abundance in the same year)	July STN abundance ^e Average <i>E-P</i> density, Sep–Dec (#/m ³)	1972–2006/2014
	Delta smelt	Fall-to-fall survival ^c (FMWT/previous year's FMWT index)	Previous FMWT fall abundance ^d Previous–previous FMWT fall abundance ^d Average <i>E-P</i> density, Apr–Jun (#/m ³) ^d Average <i>E-P</i> density, Sep–Dec (#/m ³) ^d	1972–2006/2014

Superscripts indicate that this variable has been transformed by one of the following methods, (a) \log_{10} , (b) $\log_{10} + 1$, or (c) \ln , (d) divided by 1000 or (e) divided by 10,000, in accordance with the original analysis. Abbreviations as follows: TNS = summer tow net survey, FMWT = fall midwater trawl survey, MWT = Bay Study midwater trawl survey, OT = Bay Study otter trawl survey, *E-P* = *Eurytemora* sp. + *Pseudodiaptomus* sp., two dominant copepods of the upper San Francisco Estuary and major prey items for delta smelt.

All analyses were carried out in the R statistical software suite (R Core Team 2017), and we provide the source code, and data, for our analyses in the Supporting Information.

Results

Literature review

Our literature search identified 98 publications describing environment–recruitment relationships in the Bay Delta. Of these, 40 were reviews citing primary literature or offered only general observations on raw data without conducting any analysis, and three were not available online and so not examined further. The remaining 55 peer-reviewed publications were retained for further analysis. This subset included papers published between 1977 and 2017, with a mean publication date of 2002. These studies examined an average of 10 relationships per study, and a minority of papers examined a large number (> 100) of competing models describing the same relationship. Each study examined between 1 and 17 focal species (mean ~ 3).

These 55 peer-reviewed studies described 420 individual relationships which overwhelmingly focused on examining the influence of environmental variables on population abundance as opposed to other biological characteristics (Fig. 2A). This is likely a result of the fact that roughly 70% of all relationships relied on publicly available long-term abundance survey data that has been collected in the Bay Delta for decades (e.g., California Department of Fish and Wildlife FMWT, TNS, and San Francisco Bay Study OT and MWT surveys), whereas other types of population variables are generally not routinely collected for most species and were obtained via one-off experiments or surveys (18% of the relationships). The distribution of relationships was also strongly biased toward species that are

either currently or historically listed as threatened or endangered (Fig. 2B). There was a more even distribution of environmental variables examined across studies, but the most frequently examined variables were X2, or other measures of salinity, and flow (Fig. 2C). Most relationships examined only a single environmental variable (maximum of seven) through simple linear regression (60% of all relationships). However, the number of variables has generally increased in more recent studies as researchers continue to adopt more complex analytical methods including multivariate models, generalized additive models, and whole-ecosystem simulations. While about 42% of all relationships have been identified as underpinning regulatory decision-making, only 14% of all relationships have been retested to quantify their durability (i.e., extent to which their magnitude of effect holds up in the face of new data).

Across the published relationships there was broad variation within and among environmental variables in the amount of variation in the abundance that was explained. For example, some environmental variables like flow and salinity tended to explain over half the observed variation in abundance, while others such as temperature and water export rates (measured in cfs—cubic feet per second—or $\text{m}^3 \text{s}^{-1}$, where $35.31 \text{ cfs} = 1 \text{ m}^3 \text{ s}^{-1}$) explained much less variation in abundance (Fig. 3).

Reanalysis of Bay Delta environmental–recruitment relationships

Applying our screening criteria to the full set of relationships resulted in 31 relationships suitable for retesting (Table 1). The publications originally reporting these relationships varied widely in the details provided on

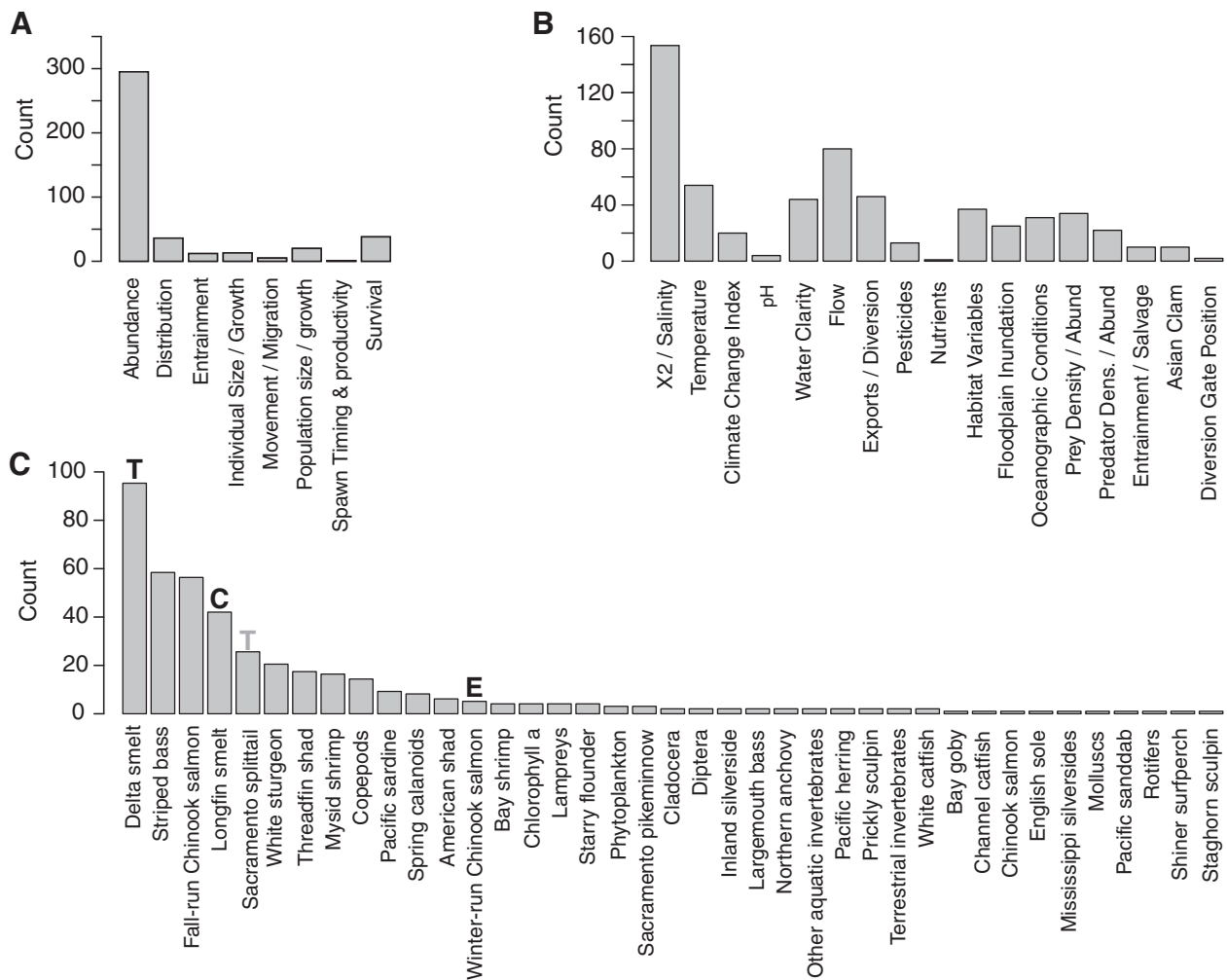


Fig. 2. Frequency distributions of the type of response variable (A), environmental variables (B), and focal species (C), represented in the 420 individual environment–recruitment relationships for the Bay Delta region examined in this study, showing the relative emphasis of this body of work. Superscripts in panel C indicate the species current (black) or historical (gray) conservation status as endangered (E), threatened (T), or candidate (C) as reported by the U.S. fish and wildlife service environmental conservation online system.

methodology, and it was often difficult to determine how the environmental variables were derived or the source of the data used to derive them. As a result, we were unable to complete reanalyses for 8 of these 31 relationships, because of either an inability to recreate the input variables using the methodology described by the authors or because of errors or missing information in the original publications. This left us with 23 relationships (17 univariate and 6 multivariate) from 8 publications which were fully reanalyzed using the most recent available data (between 9 and 40 additional years, median 9 yr). In recognition of the bias introduced by retesting only previously significant relationships, we also attempted to replicate the full model selection process in the two publications using multivariate analyses to determine whether the top models remained the same in light of new data. However, we were unable to reproduce all the variables for the more extensive model selection example (Miller

et al. 2012), and obtained nearly identical model selection results for the more limited model selection example (Feyrer et al. 2007). The general dearth of model selection approaches in this system makes it difficult to draw conclusions about the behavior of previously discarded models in the face of new data.

When updated data were used to retest previously published relationships, the direction and statistical significance of the relationships remained the same (Fig. 4; Table 1). Of the univariate relationships, 9 of 17 became stronger (i.e., either more negative or positive depending on the original relationship), 3 of 17 became weaker and 5 of 17 remained nearly identical. These general patterns remained the same for the multivariate relationships (Table 1) and when the dataset used to retest relationships was balanced to achieve an equal number of original and updated data points (Fig. 4; Table 1).

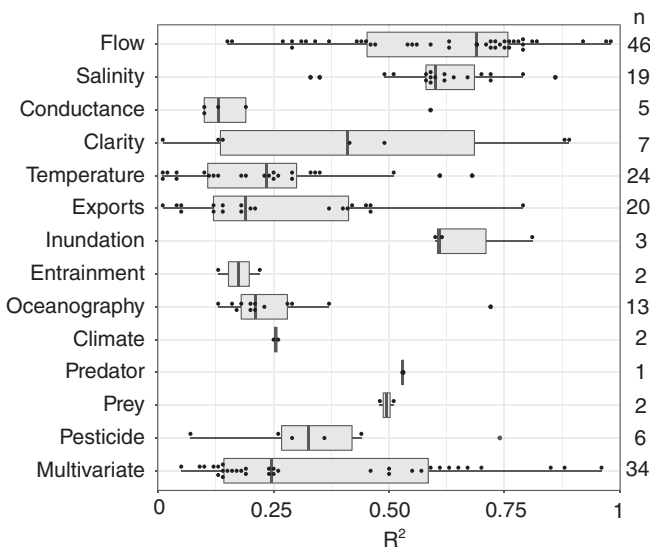


Fig. 3. Boxplots and underlying estimates of the distribution of the strength (R^2) of published relationships for each broad type of environmental variable identified in the literature review (184 papers reported this statistic). Some variables have stronger relationships (e.g., flow and salinity) with the abundance of species than others that are more variable (e.g., temperature and volume of exports) and may therefore depend more on the species and context involved. The “multivariate” environmental variable encompasses relationships that include two or more of the other environmental variables listed here. Black dots outside the range of the box and whiskers represent true outliers (i.e., beyond 1.5 times the interquartile range), while the jittered black dots represent the distribution of remaining data points.

In some instances, the addition of more years of data resulted in more variation in the relationship being explained (i.e., higher R^2 ; Fig. 5A,B). However, on average, considering more recent data did not increase the estimated strength of the relationships, and in some cases reduced it (Fig. 5A). For example, relationships between striped bass abundance and flow saw substantial declines in R^2 with the addition of more years of data, despite the fact the magnitude of the relationship remained nearly identical (Fig. 4). This occurred because the overall abundance of striped bass declined between the original time series and the updated one, which has been attributed to introduction of the Asian clam in 1987, but abundance still increased with increases in flow during both periods (Supporting Information Fig. S1).

In most cases, considering more years of data did not improve the predictive power of the environment–recruitment relationships. Instead, counterintuitively, the prediction error of each relationship typically increased with the addition of more years of data (Fig. 5C,D). The median percent increase in prediction error across the relationships retested was 30%. As with the variation explained, these changes were most pronounced for relationships between striped bass abundance and flow. Increases in prediction error could be attributed to an increase in overall variance because of the addition of more data points, particularly if the new data are drawn from a previously undersampled part of the overall distribution. However, the observations that most

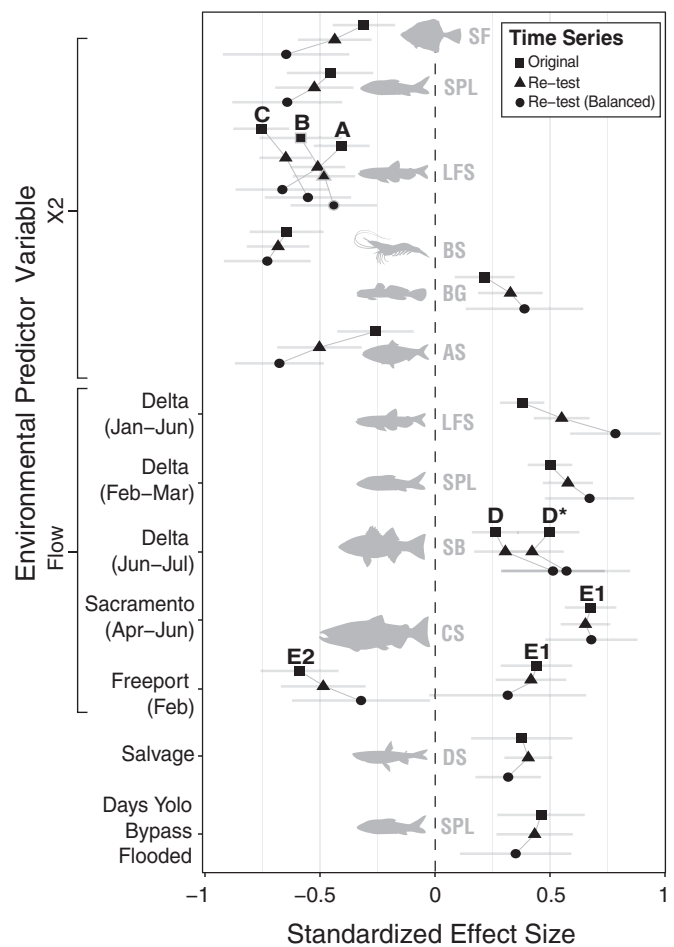


Fig. 4. Magnitude of environment–recruitment relationships for 17 univariate analyses spanning nine species based on original (■), extended (▼), and extended but balanced (●) time series. “Balanced” refers to estimates based on analyses where we randomly subsampled (100 times) the pre-retest data so that there was an equal number of pre- and post-retest data points. Where there is more than one relationship per species–variable combination, letters indicate the source of data as being from the FMWT (A), bay study MWT (B), bay study OT (C), TNS (D for the overall Delta, and D* for Suisun Bay only), or Chinook salmon trawls (E1 at Chipps Island, E2 at Sacramento). The standardized effect is the slope of the relationship between abundance and the environmental variable under consideration in standard deviation units. For example, based on the updated time series, a one standard deviation unit increase in flow at Freeport in February is expected to result in a 0.4 standard deviation unit increase in juvenile fall-run Chinook salmon index abundance in the Chipps Island trawl. Error bars are ± 1 SE.

new data points are spread across the full range of values in the original data set (Supporting Information Fig. S1), that the error around coefficient estimates does not differ markedly between extended and balanced models (Fig. 4), and that prediction error sometimes increases drastically within the span of a few years (Fig. 6) do not appear to support this hypothesis.

Discussion

Our review of the literature identified 98 publications examining 420 individual environment–recruitment relationships in

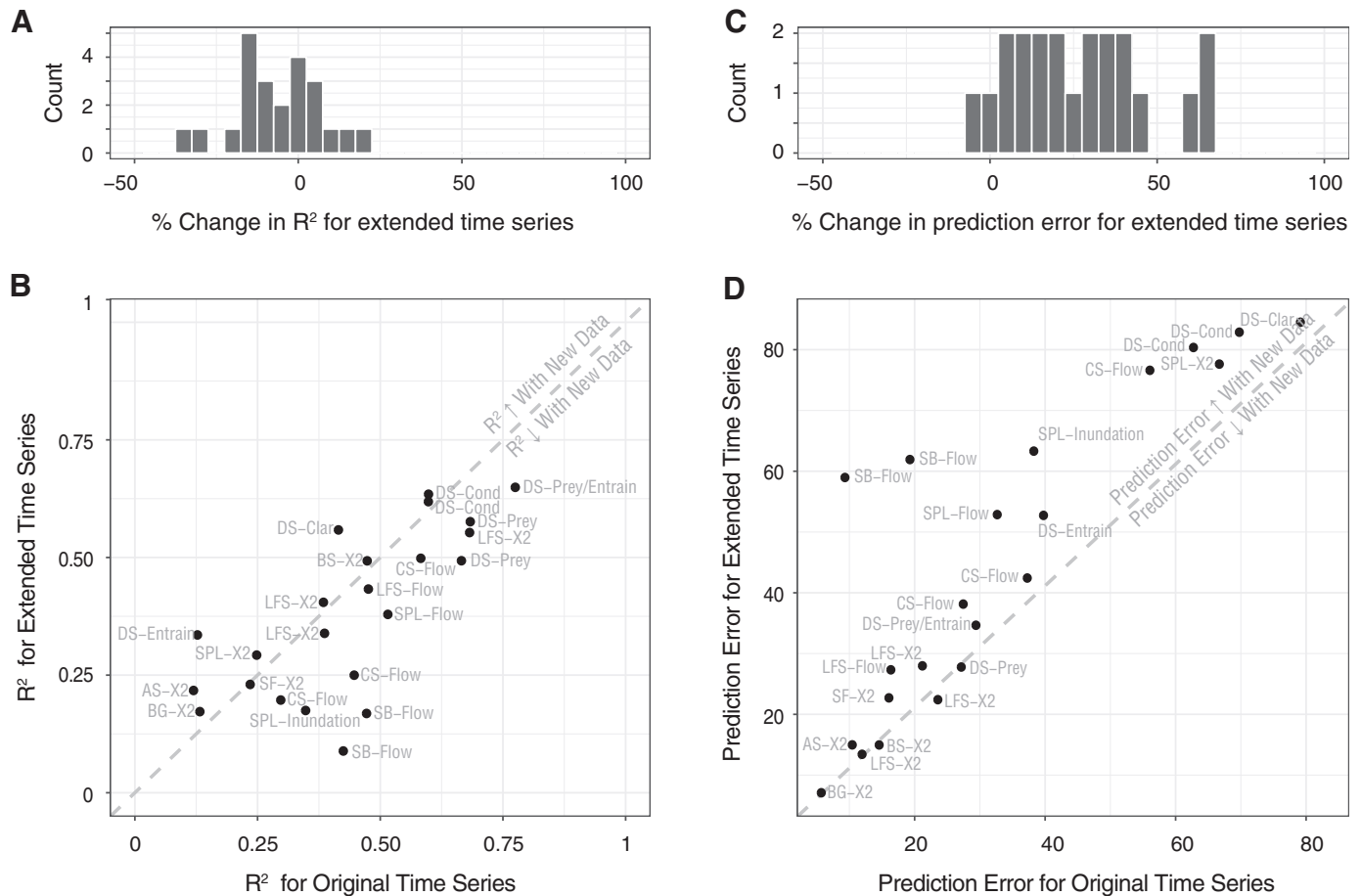


Fig. 5. Percent change as well as absolute variation explained (R^2 ; **A**, **B**) and predictive power (prediction error; **C**, **D**) of Bay Delta environment–recruitment relationships when additional years are added to the relationships. Estimates of percent change in prediction error for the striped bass–Flow relationship are not included in (**C**) and (**D**) because of their very large values ($\sim 600\%$; see Table 1). Prediction error (see Eq. 1) is the percent error in future predictions relative to the average observed abundance for a given relationship. For example, a prediction error of 50% would mean that the relationship is expected to predict future abundances within $\pm 25\%$ of the mean predicted abundance. The first two letters of each label are species codes as follows: AS = American shad, BG = bay goby, BS = bay shrimp, DS = Delta smelt, LFS = longfin smelt, CS = fall-run Chinook salmon, LFS = longfin smelt, SPL = Sacramento splittail, SF = starry flounder, SB = striped bass.

the Bay Delta. These relationships overwhelmingly focused on the influence of environmental variables on population abundance as opposed to other biological characteristics and were biased toward species that are either currently or historically listed as threatened or endangered. About half of these relationships are used in regulatory decision-making, but only one in five relationships has been retested to quantify the extent to which they stand the test of time. In contrast to Myers (1998), who found that the proportion of relationships that held up when retested was low, we found that when new data were used to retest previously published Bay Delta relationships, the direction and statistical significance of the relationships remained the same, though the amount of variation explained by the relationships and our ability to predict how a species will respond to change in their environment did not generally improve with more data. Instead, in most cases, prediction error actually increased when extending the time series, suggesting

that accumulating more data will not necessarily improve the ability of these relationships to inform decision making (insofar as predictive power is useful for decision making). Perhaps this should not come as a surprise given the original relationships examined here were typically identified based on their ability to explain observed data (e.g., R^2 from linear regression) as opposed to their ability to predict future observations using, for example, approaches like data splitting and cross-validation (Power 1993; Harrel 2015). However, a large reduction in predictive power when relationships are retested with more data may be diagnostic of an established relationship that is breaking down (i.e., as is the case with striped bass and potentially Sacramento splittail in Fig. 5), possibly because of changes in data collection methods, misspecification of the original model, or an environmental shift that causes another factor to become limiting. Such losses in predictive power might prompt action to search for unmeasured drivers of this change.

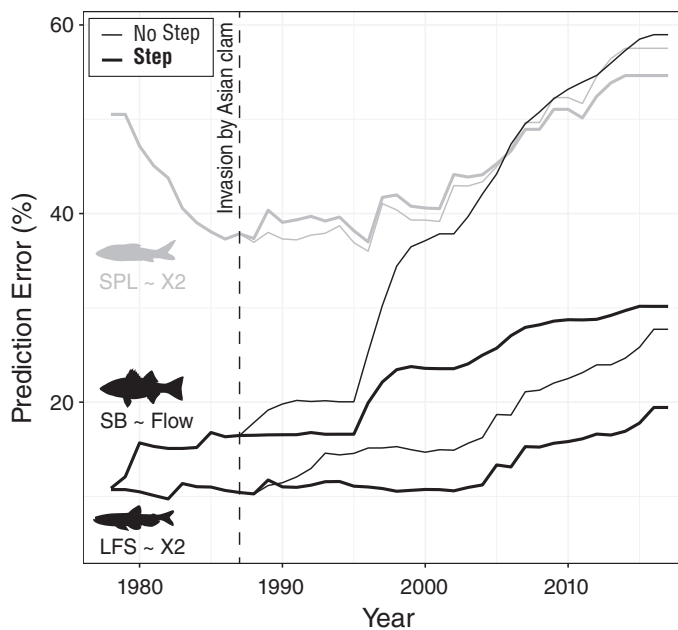


Fig. 6. Prediction error (Eq. 1) relative to the logged index values used in the models over time for three illustrative Bay Delta environment–recruitment relationships modeled without (thin lines) and with (heavy lines) a step change to account for the regime shift associated with the introduction of the Asian clam circa 1987. Prediction error in each year is based on all years of data up to that point in time. Some species (black: SB = striped bass, LFS = longfin smelt) show strong responses to the regime shift and benefit from reduced prediction error with the inclusion of a step change in the models, while others (gray: SPL = Sacramento splittail) do not.

The finding that the environment–recruitment relationships retested here are generally durable is surprising in light of the previous work by Myers (1998). The durability of the Bay Delta relationships might reflect the quality of the underlying data, consistent with Myers' finding that relationships were more durable where they used reliable data, such as research surveys and reconstructed population estimates. Rather than relying on fisheries-dependent data that can obscure environment–recruitment patterns through changes in fishing effort and capture efficiency during the exploitation phases of a fishery (Hilborn and Walters 1992), studies in the Bay Delta benefit from multiple long-term monitoring surveys. Further, whereas the relationships reviewed by Myers came from marine, freshwater, and estuarine ecosystems, the Bay Delta relationships focus entirely on species that reside in an estuarine habitat for part or all their life. It is perhaps unsurprising then that freshwater flow and salinity are the most frequently inspected response variables in the Bay Delta (Fig. 2B), and inasmuch as these relationships are durable, may reflect species' adaptations to the determinants of estuarine habitat.

However, closer inspection of the individual relationships suggests their apparent durability may be misleading. Fourteen

of the 23 relationships we retested either split (Kimmerer et al. 2009; Sommer et al. 1997, 2007) or excluded altogether data prior to 1987 under the assumption that the invasion of the Asian clam fundamentally altered the estuary food web (Kimmerer et al. 1994). In some cases the slope of the environment–recruitment relationship held, but the intercept changed after 1987, as in the case of longfin smelt (Kimmerer et al. 2009), implying a shift in the abundance index. In the case of delta smelt abundance, only the postclam invasion period produced significant relationships with water quality predictors (Feyrer et al. 2007). Constantly refining these relationships by adding step changes or excluding portions of the time series to account for changes in the ecosystem gives the appearance that the relationships are durable, but their practical value to inform policy is limited when the present or future state of the ecosystem is unknown. For example, Stevens (1977) reported a strong striped bass response to delta outflow into the 1970s, but visual inspection of the data suggests the relationship has broken down in the years since (Supporting Information Fig. S1A,B). Similarly, the slope of the relationship between flow at Freeport and Chinook salmon catches in Sacramento and Chipps Island trawls are driven by flow conditions only observed in the original time series (Supporting Information Fig. S1O,P); it is unclear if this relationship still holds. Indeed, despite the apparent durability of the retested Bay Delta relationships, their generally poor—and declining—predictive power (Fig. 5, Supporting Information Table S1) should be acknowledged before any of the relationships are used to inform policy.

Our review of published environment–recruitment relationships in the Bay-Delta also highlights some methodological shortcomings of studies in this region. First, our review illustrates the value of long-term data collection programs that follow standardized and consistent protocols to detect and validate long-term trends in biological variables. A large share of studies in this space leverage these datasets. However, the availability and accessibility of such survey data may also reduce the likelihood that researchers in this region embark on independent data collection to study other species and biological variables that are not the focus of existing surveys. Second, despite the accessibility of long-term survey data and the simplicity of correlative analyses, we were still unable to reproduce originally published variables for many relationships otherwise meeting our criteria for retesting. In some cases, this was because of errors in the original work, unreported assumptions about data transformations that only became clear after contacting the authors, or the prior use of interpolated data that has since been corrected at the source by survey operators. These challenges highlight the importance of reproducibility in research in general, and into environment–recruitment relationships in particular, echoing a growing call for greater reproducibility both in ecology and across other scientific disciplines (Cassey and Blackburn 2006; Nosek et al. 2015; Borregaard and Hart 2016). Trends toward the use of open source programming

languages for ecological analysis, and for the publication of the underlying code alongside manuscripts as we do in our Supporting Information, should help play a role in improving the reproducibility of ecological research going forward (Mislan et al. 2016).

Our conclusions should be considered in light of potential biases in both our study selection criteria and in the type of data these studies draw upon. As with all studies based on literature review, our results are subject to publication bias (Cooper et al. 2009) relating to our decision to focus on peer-reviewed studies, a propensity toward publication of significant relationships in peer-reviewed journals, and a disproportionate number of publications on particular variables (e.g., X2 and flow) coming from a few very active authors in this field. In addition, recent work has shown that the long-term survey data used to create many of these relationships may itself be inherently biased by unquantified changes in detection probability. Detection probability, or catchability, may vary considerably over time with (1) overall abundance (i.e., it is more difficult to catch a rarer species or size class) (Mahardja et al. 2017); (2) with changing environmental conditions (e.g., catchability may decrease with increasing water clarity as fish are better able to see and avoid survey gear in clearer water) (Latour 2016); and (3) with differences in gear type across surveys (Peterson and Barajas, in press). When not accounted for, these changes in detectability may be incorrectly interpreted as real changes in abundance or occupancy. Notwithstanding these potential biases, we believe that the breadth of species, environmental variables, and survey types covered in our analysis allows us to draw general conclusions about the utility of environment–recruitment models, and to synthesize insights from the literature on best practices for the analysis, use, and refinement of environment–recruitment relationships to inform decision making in natural resource management.

Correlation, causation, and strength of evidence

Correlations underpin most natural resource management decisions, including those in the Bay-Delta. Such correlations are usually assumed to be causative, but we are often reminded of the adage “correlation does not equal causation.” Ideally, manipulative experiments can be used to determine whether a specific human action causes a response in an ecosystem component. Such learning by manipulation embraces the three key elements of experimental design: controls, randomization of treatments, and replication. Much has been learned in natural resource management through manipulative experiments. For example, hatcheries manipulate the timing and size of fish released to determine which combination result in optimal survival (e.g., Irvine et al. 2013) and hydroelectric facility operators manipulate timing and magnitude of flow releases to determine which flows are most likely to improve fish survival

(e.g., Bradford et al. 2011). In the Bay Delta, the Vernalis Adaptive Management Plan evaluated how juvenile salmon migration survival through the Delta responded to experimentally manipulated San Joaquin flow, exports, and physical barriers (San Joaquin River Group Authority 2013). While manipulative experiments are the gold standard approach to establishing causation, scope for manipulative experiments decreases at increasing scales, and so we are left interpreting correlative relationships in order to manage some of the largest human-influenced ecosystems in the world.

When opportunities for learning by manipulation are limited or impossible, the weight of evidence for a hypothesized causal correlation should be assessed based on the strength, consistency, specificity and plausibility of the mechanism underlying the relationship (e.g., Hill 1965; Hilborn 2016). However, while strength, consistency, specificity and plausibility can help guide the degree of support for a given relationship, they should not come at the cost of maintaining multiple working hypotheses, and evaluating the evidence for each simultaneously when using correlations to guide decision making (Hilborn and Mangel 1997; Plowright et al. 2008). An illustrative example of the simultaneous consideration of multiple working hypotheses is the development and application of a state-space multistage life cycle model to investigate for drivers of population decline in delta smelt by Maunder and Deriso (2011).

Adaptive management in the face of ecosystem change

Even when relationships are truly causative, using past relationships to guide future management decisions can fail to have the intended effect when the system in which they occur changes over time (i.e., exhibits nonstationarity). Such nonstationarity in aquatic systems can arise from both slow-moving environmental change or rapid regime shifts and “tipping points” between alternative stable states (Scheffer et al. 2001, 2009). There is widespread evidence of regime shifts in aquatic ecosystems arising from both natural (e.g., climate) and human (e.g., pollution and species introductions) caused factors (Carpenter 2003; Hunsicker et al. 2016) and the state of the ecosystem can have a strong influence on the outcomes of management actions. For example, large releases of hatchery salmon reduce the survival of endangered wild salmon but only during periods of poor ocean conditions (Levin et al. 2001), and translocation of wild juvenile salmon past hydropower installations carries greater benefits for their ocean survival in cooler but not warmer oceanic regimes (Gosselin et al. 2017). These examples highlight the fact that the benefits of management interventions (e.g., reducing hatchery releases to minimize impacts to wild fish, translocation of fish past barriers) are contingent upon the ocean regime the system is experiencing in any given year. In the Bay Delta, the introduction and rapid expansion of the invasive Asian clam in the late 1980s is believed to have caused a major increase in grazing

pressure on phytoplankton, leading to a persistent decline in overall pelagic food resources (Carlton et al. 1990; Nichols et al. 1990; Baxter et al. 2010). The so-called “step change” toward this new stable state has had varying influences on different species within the ecosystem. For example, the regime shift because of Asian clam has led to a change in the overall magnitude (intercept) but not the rate of change (slope) of existing abundance-flow relationships for striped bass (Kimmerer et al. 2009, and reproduced in this study), but has been suggested to have driven a new abundance-flow relationship for threatened delta smelt that brought with them significant implications for the way flows in this system are managed (CDFW 2016). Managers may also need to anticipate future changes in ecosystem state. For example, climate change may erode or replace existing relationships as warming water temperatures exceed the range of historical observations in this region and begin to approach species’ thermal limits (Cloern et al. 2011). Temperature was a commonly explored environmental variable in the studies we reviewed (Fig. 2), but its limited explanatory power excluded it from the retested relationships (Fig. 3; Table 1). However, during California’s most recent drought, water temperatures regularly approached or exceeded the thermal tolerance of Bay Delta species, including longfin smelt (Jeffries et al. 2016). Indeed, a limiting factors approach (e.g., Hamilton and Murphy 2018) may reveal established flow–recruitment relationships are superseded in certain years by species’ physiological response to increased water temperature.

Given the ubiquity of regime shifts and nonstationarities in aquatic systems (Möllmann and Diekmann 2012), including in the Bay Delta (Kimmerer 2002; Kimmerer et al. 2009; Thomson et al. 2010), how should one evaluate the evidence for environment–recruitment relationships and use them to inform decision making when system change is suspected? In some cases, change in a system will be so pronounced that there is little question about when it occurred and so the nature of a relationship post regime change should be the one that is used to inform future management actions. In other instances, the timing of abrupt change, and indeed whether it has occurred at all, will be uncertain for several years and so debated (at times vociferously) until clear evidence accumulates that a regime shift has occurred. To date, statistical support for changes in the state of the Bay Delta system have used retrospective approaches, for example, change-point analyses (Thomson et al. 2010) or applying dummy variables in linear regression (e.g., Kimmerer 2002; Kimmerer et al. 2009), to quantitatively evaluate the evidence for system change. However, dynamic linear models may be better suited for forecasting population responses to changes in the system state because they treat regression parameters as time-varying, thereby explicitly acknowledging the evolution of the functional relationship between the environmental variable and recruitment (e.g., Scheuerell and Williams 2005). In addition, and perhaps more importantly, one should quantitatively

evaluate the decision-making consequences of incorrectly assuming a regime shift has or has not occurred so as to be able to understand and clearly communicate the costs (e.g., biological, economic, and social) of getting it wrong. Lastly, it has also been found that when nonstationarity is present (or suspected to be present), using recent observations to predict the consequences of alternative management actions can improve management outcomes (e.g., Iannelli et al. 2012).

In systems that have undergone dramatic change (e.g., tipping points), failure to regularly re-evaluate the durability and predictability of environment–recruitment relationships risks making management decisions based on information with increasingly large margins of error, with the potential for negative ecological, social and economic consequences. To illustrate this point, we estimated prediction error for a few Bay Delta environment–recruitment relationships spanning the period before and after the Asian clam invasion in 1987 for species that have (striped bass and longfin smelt; Kimmerer et al. 2009), and have not (Sacramento splittail; Kimmerer et al. 2009), responded to the invasion (Fig. 6). For those species that declined in abundance coincident with the invasion, failure to account for this change results in relationships with increasing prediction error as time goes by after the invasion, relative to relationships that account for the change by including a step change in 1987 (striped bass and longfin smelt; Fig. 6). In contrast, for Sacramento splittail, which did not appear to respond to the invasion, there is no benefit to including a step change in the relationship. Interestingly, these analyses suggest that the environment–recruitment relationships for striped bass and longfin smelt have experienced subsequent regime shifts (~ 1995 for striped bass and ~ 2005 for longfin smelt) that have further eroded their predictive power. This subsequent shift may be explained by observed changes in distribution likely to affect catchability in long-term surveys. Prior studies have suggested that a reduction on pelagic food resources because of overgrazing by Asian clam appears to have driven shifts in the distribution of young fish in the 1980s and 1990s away from the primary sampling regions of long-term surveys and toward areas characterized by fewer clams and better foraging prospects. This manifested as a lateral shift from deeper channel habitat preferentially sampled by annual surveys toward shallower slough habitat for striped bass (Sommer et al. 2011), and as a longitudinal shift from upstream habitat toward more saline downstream habitat for longfin smelt (Baxter et al. 2008; Sommer et al. 2011). Thus, these two shifts in prediction error align with and illustrate two stages of rapid environmental change driven by a trophic cascade, the first likely related to a species introduction, and the second likely related to two different behavioral responses to adapt to the consequences of this introduction.

Our findings suggest that when environment–recruitment relationships underpin decision-making they should be re-evaluated on a regular basis as part of a broader adaptive management approach to ensure that they remain robust in

the face of new data and continue to provide an accurate representation of a continually evolving ecosystem. Such an adaptive approach to evaluating environment–recruitment relationships is aligned with broader calls for increasing the implementation of more proactive adaptive management in Bay Delta ecosystems to address accelerating environmental change (Delta Independent Science Board 2015, 2016; Zandvoort et al. 2018).

Environment–recruitment relationships in decision making

The widespread use, and at times misuse, of environment–recruitment relationships to inform decision making has produced several general insights into best practices for incorporating such relationships into natural resource management.

First, uncertainty should be both quantified and propagated in any analysis that seeks to predict the consequences of alternative management actions and identify those actions most likely to achieve desired objectives. This uncertainty comes in at least four distinct forms (e.g., Peterman 2004): (1) natural variation in both physical and biological processes; (2) uncertainty because of imperfect assessment arising from measurement error; (3) structural uncertainty because of incomplete understanding of how a system functions leading to model misspecification; and lastly (4) outcome uncertainty or implementation error in how well a given management target (e.g., increase flow by 20%) is achieved by a management action (e.g., releasing water from a reservoir). While uncertainty resulting from points 2 to 4 can in theory be reduced by improved measurements, greater understanding of system function, and better management control, all of which can be accomplished to some extent with the collection of more years of data, uncertainty arising from natural variability is irreducible. We found that prediction error was not reduced for Bay Delta environment–recruitment relationships with the accumulation of more years of data, and in fact increased in many cases. This finding suggests that natural variation in the physical or biological processes of the system is an important and ongoing source of uncertainty. Improving predictions of how Bay Delta taxa respond to changing environmental conditions and human action will require identifying predictors of currently unaccounted for variation. In the interim, modeling natural variability as a stochastic process should be favored over retrospective change-point and dummy variable evaluations because the former accounts for the irreducible uncertainty and propagates it into better-calibrated predictions than standard linear regression (Clark 2005).

Second, there is increasing recognition of the value of developing and using life cycle models to evaluate the predicted consequences of alternative management actions on species of concern in the face of this uncertainty (e.g., Good et al. 2007; Ruckelshaus et al. 2002; Zeug et al. 2012). In contrast to single life stage, habitat type, or environmental relationships, life-cycle models simultaneously consider extrinsic

(environment, management action) and intrinsic (density dependence) factors influencing multiple life stages. Life cycle models can either be mechanistic where survival between life stages is based on specific mechanisms (Scheuerell et al. 2006) or statistical where life stage specific survival is not defined by specific mechanistic relationships (Nobriga and Rosenfield 2016). The use of life cycle models allows for more realistic and comprehensive evaluation of the predicted outcomes of alternative management actions than considering single life stage, habitat type, or environmental relationships, because it considers environmental effects across linked stages in a life cycle while also accounting for population processes (e.g., growth, movement, mortality, and reproduction).

Even when uncertainty is successfully incorporated into modeling approaches, the broader question remains—how can we account for and propagate uncertainty into the broader management of a complex system with many conflicting management objectives when our understanding of that system is not, and will never be, complete? In a review of management approaches in the Bay-Delta ecosystem for the National Research Council (NRC), the Committee on Sustainable Water and Environmental Management proposes that agencies should adopt management approaches that assume “universal nonstationarity,” or the idea that all aspects of the environment will constantly be changing.” (NRC 2012). Such approaches may prove challenging for many traditional decision-making pathways, which are often constrained by static or slow-moving policy frameworks (Aladjem 2013; Delta Independent Science Board. 2016). However, a number of approaches with shared characteristics have been developed to help facilitate formalized decision-making in complex systems characterized by their uncertainty and are among the approaches recommended by the NRC review. Among these are decision analysis (Peterman and Anderson 1999), decision scaling (NRC 2012), risk assessment (Peters and Marmorek 2001), and management strategy evaluation (Punt et al. 2016).

Conclusions

Moving forward, there is a growing recognition of the importance of maintaining multiple working hypotheses when quantifying the support for correlations in environmental management (Hilborn 2016), that quantitative assessment of policies that consider these relationships should be done using approaches that allow for realistic incorporation and propagation of multiple sources of uncertainty (Peterman 2004), and that, ultimately, managers in the Bay-Delta and elsewhere should identify policies that are robust to a range of alternative hypotheses (National Research Council 2012; Schindler and Hilborn 2015).

Despite advances in the tools available to improve our assessment of environment–recruitment relationships and their consideration in decision making, we should remain humble in

our zeal to either accept them as fact or discount them entirely because they are “just correlations.” As Hill emphasized in his 1965 Presidential Address on correlation and causation to the Royal Society of Medicine (Hill 1965): “All scientific work is incomplete—whether it be observational or experimental... [and] is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a particular time.”

REFERENCES

- Aladjem, D. 2013. When all you have is a hammer: Scientific uncertainty and Bay-Delta policy, 1995 to present. *Nat. Res. Environ.* **27**: 36–39. Available from: <https://www.jstor.org/stable/24426044>
- Baxter, R., and others. 2008. Pelagic organism decline progress report: 2007 synthesis of results. Interagency Ecological Program for the San Francisco Estuary. doi:10.1016/j.cveq.2008.10.006
- Baxter, R., and others. 2010. Interagency Ecological Program 2010 Pelagic Organism Decline work plan and synthesis of results. Interagency Ecological Program for the San Francisco Estuary. doi:10.4061/2010/712370
- Bennett, W. A. 2005. Critical assessment of the delta smelt population in the San Francisco estuary, California. *San Francisco Estuar. Watershed Sci.* **3**: 1–71. doi:10.15447/sfews.2005v3iss2art1
- Borregaard, M. K., and E. M. Hart. 2016. Towards a more reproducible ecology. *Ecography* **39**: 349–353. doi:10.1111/ecog.02493
- Bradford, M. J., P. S. Higgins, J. Korman, and J. Snee. 2011. Test of an environmental flow release in a British Columbia river: Does more water mean more fish? *Freshw. Biol.* **56**: 2119–2134. doi:10.1111/j.1365-2427.2011.02633.x
- Brandes, P., Burmester R., and Apeegle J. 2006. Estimating relative abundance and survival of juvenile Chinook Salmon in the Sacramento—San Joaquin Estuary. Interagency Ecological Program for the San Francisco Estuary, Newsletter. **19**: 41–46. Available from: https://water.ca.gov/LegacyFiles/iep/newsletters/2006/IEPNewsletter_spring2006final5.pdf
- California Department of Fish and Wildlife (CDFW). 2016. CDFW rationale for summer delta flow augmentation for improving delta smelt survival (CDFW 8 July 2016). Available from https://www.nrdc.org/sites/default/files/media-uploads/cdfw_outflow.pdf.pdf [accessed 2017 September 9]
- California Department of Water Resources and U.S. Bureau of Reclamation. 2016. Final Environmental Impact Report/Environmental Impact Statement for the Bay Delta Conservation Plan/California WaterFix—Volume I. Final EIR/EIS for the BDCP/California WaterFix. December (DOE/EIS-0515) (ICF 00139.14). Prepared by ICF International, Sacramento, CA.
- Carlton, J. T., J. K. Thompson, L. E. Schemel, and F. H. Nichols. 1990. Remarkable invasion of San Francisco Bay (California, USA) by the Asian clam *Potamocorbula amurensis*. I. Introduction and dispersal. *Mar. Ecol. Prog. Ser.* **66**: 81–94. doi:10.3354/meps066081
- Carpenter, S. R. 2003. Regime shifts in lake ecosystems: Patterns and variation. In Volume 15 in outline of excellence in ecology series. Ecology Institute, Oldendorf/Luhe. doi:10.2190/GPN8-DF42-DT0M-9BV4
- Cassey, P., and T. M. Blackburn. 2006. Reproducibility and repeatability in ecology. *Bioscience* **56**: 958–959. doi:10.1641/0006-3568(2006)56[958:RARIE]2.0.CO;2
- Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. *Ecol. Lett.* **8**: 2–14. doi:10.1111/j.1461-0248.2004.00702.x
- Cloern, J. E., N. Knowles, L. R. Brown, D. Cayan, M. D. Dettinger, T. L. Morgan, D. H. Schoellhamer, M. T. Stacey, M. van der Wegen, R. W. Wagner, and A. D. Jassby. 2011. Projected evolution of California’s San Francisco Bay-Delta-River system in a century of climate change. *PLoS ONE* **6**: e24465. doi:10.1371/journal.pone.0024465
- Cooper, H., L. V. Hedges, and J. C. Valentine [eds.]. 2009. The handbook of research synthesis and meta-analysis. Russell Sage Foundation.
- Cushing, D. H. 1995. The long-term relationship between zooplankton and fish IV. Spatial/temporal variability and prediction. *ICES J. Mar. Sci.* **52**: 611–626. doi:10.1016/1054-3139(95)80076-X
- Delta Independent Science Board. 2015. Flows and fishes in the Sacramento–San Joaquin Delta, research needs in support of adaptive management, Sacramento, CA. 37 p. Available from: <http://deltacouncil.ca.gov/docs/delta-isb-s-final-report-flows-and-fishes-sacramento-san-joaquin-delta-research-needs-support> [accessed 2017 September 9], doi:10.1039/c4sc90043j
- Delta Independent Science Board. 2016. Improving Adaptive Management in the Sacramento–San Joaquin Delta, Sacramento, CA. 58 p. Available from: <http://deltacouncil.ca.gov/docs/final-delta-isb-adaptive-management-review-report> [accessed 2017 September 9]
- Feyrer, F., M. L. Nobriga, and T. R. Sommer. 2007. Multidecadal trends for three declining fish species: Habitat patterns and mechanisms in the San Francisco estuary, California, USA. *Can. J. Fish. Aquat. Sci.* **64**: 723–734. doi:10.1139/f07-048
- Good, T. P., T. J. Beechie, P. McElhany, M. M. McClure, and M. H. Ruckelshaus. 2007. Recovery planning for endangered species act-listed Pacific salmon: Using science to inform goals and strategies. *Fisheries* **32**: 426–440. doi:10.1577/1548-8446(2007)32[426:RPFESL]2.0.CO;2
- Gosselin, J. L., R. W. Zabel, J. J. Anderson, J. R. Faulkner, A. M. Baptista, and B. P. Sandford. 2017. Conservation planning for freshwater–marine carryover effects on Chinook salmon survival. *Ecol. Evol.* **1–14**: 319–332. doi:10.1002/ece3.3663
- Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., Lowndes, J.S., Rockwood, R.C., and others. 2015. Spatial and temporal changes in cumulative human impacts

- on the world's ocean. *Nat. Commun.* **6**, 7615. doi:[10.1038/ncomms8615](https://doi.org/10.1038/ncomms8615)
- Hamilton, S. A., and D. D. Murphy. 2018. Analysis of limiting factors across the life cycle of Delta smelt (*Hypomesus transpacificus*). *Environ. Manag.* **62**(2): 365–382. doi:[10.1007/s00267-018-1014-9](https://doi.org/10.1007/s00267-018-1014-9)
- Harrell, F. E., Jr. 2015. Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis. Springer. doi:[10.4081/or.2015.5691](https://doi.org/10.4081/or.2015.5691)
- Hilborn, R. 2016. Correlation and causation in fisheries and watershed management. *Fisheries* **41**: 18–25. doi:[10.1080/03632415.2016.1119600](https://doi.org/10.1080/03632415.2016.1119600)
- Hilborn, R., and M. Mangel. 1997. The ecological detective: Confronting models with data. Princeton Univ. Press. doi:[10.1259/bjr.70.829.9059293](https://doi.org/10.1259/bjr.70.829.9059293)
- Hilborn, R. and C. J. Walters (Eds.). 1992. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. Springer. 570 p. doi: [10.1007/978-1-4615-3598-0](https://doi.org/10.1007/978-1-4615-3598-0)
- Hill, A. B. 1965. The environment and disease: Association or causation? *Proc. R. Soc. Med.* **58**: 295–300. doi:[10.1177/003591576505800503](https://doi.org/10.1177/003591576505800503)
- Hjort, J. 1914. Fluctuations in the great fisheries of northern Europe viewed in the light of biological research. *Rapp. P.-V. Réun.* **20**: 1–228. Available from: <http://hdl.handle.net/11250/109177>
- Hunsicker, M. E., C. V. Kappel, K. A. Selkoe, B. S. Halpern, C. Scarborough, L. Mease, and A. Amrhein. 2016. Characterizing driver–response relationships in marine pelagic ecosystems for improved ocean management. *Ecol. Appl.* **26**: 651–663. doi:[10.1890/14-2200](https://doi.org/10.1890/14-2200)
- Hutton, P. H., J. S. Rath, L. Chen, M. J. Unga, and S. B. Roy. 2015. Nine decades of salinity observations in the San Francisco Bay and Delta: Modeling and trend evaluations. *J. Water Res. Plan. Manag.* **142**: 04015069. doi:[10.1061/\(ASCE\)WR.1943-5452.0000617](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000617)
- Ianelli, J. N., T. Honkalehto, S. Barbeaux, S. Kotwicki, K. Aydin, and N. Williamson. 2012. Assessment of the walleye Pollock stock in the eastern Bering Sea. Alaska Fisheries Science Center, National Marine Fisheries Service. Available from: <https://www.afsc.noaa.gov/REFM/Docs/2012/EBSpollock.pdf>
- Irvine, J. R., M. O'Neill, L. Godbout, and J. Schnute. 2013. Effects of smolt release timing and size on the survival of hatchery-origin coho salmon in the strait of Georgia. *Prog. Oceanogr.* **115**: 111–118. doi:[10.1016/j.pocean.2013.05.014](https://doi.org/10.1016/j.pocean.2013.05.014)
- Jacobson, L. D., and A. D. MacCall. 1995. Stock-recruitment models for Pacific sardine (*Sardinops sagax*). *Can. J. Fish. Aquat. Sci.* **52**: 566–577. doi:[10.1139/f95-057](https://doi.org/10.1139/f95-057)
- Jassby, A. D., W. J. Kimmerer, S. G. Monismith, C. Armor, J. E. Cloern, T. M. Powell, J. R. Schubel, and T. J. Vendlinski. 1995. Isohaline position as a habitat indicator for estuarine populations. *Ecol. Appl.* **5**: 272–289. doi:[10.2307/1942069](https://doi.org/10.2307/1942069)
- Jeffries, K. M., R. E. Connon, B. E. Davis, L. M. Komoroske, M. T. Britton, T. Sommer, A. E. Todgham, and N. A. Fangue. 2016. Effects of high temperatures on threatened estuarine fishes during periods of extreme drought. *J. Exp. Biol.* **219**: 1705–1716. doi:[10.1242/jeb.134528](https://doi.org/10.1242/jeb.134528)
- Kimmerer, W. J., E. Gartside and J. J. Orsi. 1994. Predation by an introduced clam as the likely cause of substantial declines in zooplankton of San Francisco Bay. *Mar. Ecol. Prog. Ser.* **113**: 81–93. doi: [10.3354/meps113081](https://doi.org/10.3354/meps113081)
- Kimmerer, W. J. 2002. Effects of freshwater flow on abundance of estuarine organisms: Physical effects or trophic linkages? *Mar. Ecol. Prog. Ser.* **243**: 39–55. doi:[10.3354/meps243039](https://doi.org/10.3354/meps243039)
- Kimmerer, W. J. 2004. Open water processes of the San Francisco estuary: From physical forcing to biological responses. *San Francisco Estuar. Watershed Sci.* **2**: 1–140. Available from: <http://escholarship.org/uc/item/9bp499mv>
- Kimmerer, W. J., E. S. Gross, and M. L. MacWilliams. 2009. Is the response of estuarine nekton to freshwater flow in the San Francisco estuary explained by variation in habitat volume? *Estuar. Coasts* **32**: 375–389. doi:[10.1007/s12237-008-9124-x](https://doi.org/10.1007/s12237-008-9124-x)
- Latour, R. J. 2016. Explaining patterns of pelagic fish abundance in the Sacramento–san Joaquin Delta. *Estuar. Coasts* **39**: 233–247. doi:[10.1007/s12237-015-9968-9](https://doi.org/10.1007/s12237-015-9968-9)
- Levin, P. S., R. W. Zabel, and J. G. Williams. 2001. The road to extinction is paved with good intentions: Negative association of fish hatcheries with threatened salmon. *Proc. R. Soc. Lond. B Biol.* **268**: 1153–1158. doi:[10.1098/rspb.2001.1634](https://doi.org/10.1098/rspb.2001.1634)
- Mahardja, B., M. J. Young, B. Schreier, and T. Sommer. 2017. Understanding imperfect detection in a San Francisco estuary long-term larval and juvenile fish monitoring programme. *Fish. Manag. Ecol.* **24**: 488–503. doi:[10.1111/fme.12257](https://doi.org/10.1111/fme.12257)
- Maunder, M. N., and R. B. Deriso. 2011. A state–space multi-stage life cycle model to evaluate population impacts in the presence of density dependence: Illustrated with application to delta smelt (*Hypomesus transpacificus*). *Can. J. Fish. Aquat. Sci.* **68**: 1285–1306. doi:[10.1139/f2011-071](https://doi.org/10.1139/f2011-071)
- Miller, W. J., B. F. Manly, D. D. Murphy, D. Fullerton, and R. R. Ramey. 2012. An investigation of factors affecting the decline of delta smelt (*Hypomesus transpacificus*) in the Sacramento–San Joaquin estuary. *Rev. Fish. Sci.* **20**: 1–19. doi:[10.1080/10641262.2011.634930](https://doi.org/10.1080/10641262.2011.634930)
- Mislan, K. A. S., J. M. Heer, and E. P. White. 2016. Elevating the status of code in ecology. *Trends Ecol. Evol.* **31**: 4–7. doi:[10.1016/j.tree.2015.11.006](https://doi.org/10.1016/j.tree.2015.11.006)
- Möllmann, C., and R. Diekmann. 2012. Marine ecosystem regime shifts induced by climate and overfishing: A review for the northern hemisphere. *Adv. Ecol. Res.* **47**: 303–347.
- Myers, R. A. 1998. When do environment–recruitment correlations work? *Rev. Fish Biol. Fish.* **8**: 285–305. doi:[10.1023/A:1008828730759](https://doi.org/10.1023/A:1008828730759)
- National Research Council (NRC). 2012. Sustainable water and environmental management in the California Bay-Delta. The National Academies Press. doi:[10.17226/13394](https://doi.org/10.17226/13394)

- Nichols, F. H., J. K. Thompson, and L. E. Schemel. 1990. Remarkable invasion of San Francisco Bay (California, USA) by the Asian clam *Potamocorbula amurensis*. II. Displacement of a former community. *Mar. Ecol. Prog. Ser.* **66**: 95–101. doi:[10.3354/meps066095](https://doi.org/10.3354/meps066095)
- Nobriga, M. L., and J. A. Rosenfield. 2016. Population dynamics of an estuarine forage fish: Disaggregating forces driving long-term decline of Longfin smelt in California's San Francisco estuary. *Trans. Am. Fish. Soc.* **145**: 44–58. doi:[10.1080/00028487.2015.1100136](https://doi.org/10.1080/00028487.2015.1100136)
- Nosek, B.A., Alter, G., Banks, G.C., Borsboom, D., Bowman, S. D., Breckler, S.J., Buck, S., Chambers, C.D., and others. 2015. Promoting an open research culture. *Science* **348**: 1422–1425, doi:[10.1126/science.aab2374](https://doi.org/10.1126/science.aab2374)
- Peterman, R. M. 2004. Possible solutions to some challenges facing fisheries scientists and managers. *ICES J. Mar. Sci.* **61**: 1331–1343. doi:[10.1016/j.icesjms.2004.08.017](https://doi.org/10.1016/j.icesjms.2004.08.017)
- Peterman, R. M., and J. L. Anderson. 1999. Decision analysis: A method for taking uncertainties into account in risk-based decision making. *Hum. Ecol. Risk Assess.* **5**: 231–244. doi:[10.1080/10807039991289383](https://doi.org/10.1080/10807039991289383)
- Peters, C. N., and D. R. Marmorek. 2001. Application of decision analysis to evaluate recovery actions for threatened Snake River spring and summer Chinook salmon (*Oncorhynchus tshawytscha*). *Can. J. Fish. Aquat. Sci.* **58**: 2431–2446. doi:[10.1139/f01-173](https://doi.org/10.1139/f01-173)
- Peterson, J., and Barajas, M. In Press. An Evaluation of Three Fish Surveys In The Bay-Delta. 1995–2015. San Francisco Estuary and Watershed Science.
- Plowright, R. K., S. H. Sokolow, M. E. Gorman, P. Daszak, and J. E. Foley. 2008. Causal inference in disease ecology: Investigating ecological drivers of disease emergence. *Front. Ecol. Environ.* **6**: 420–429. doi:[10.1890/070086](https://doi.org/10.1890/070086)
- Power, M. 1993. The predictive validation of ecological and environmental models. *Ecol. Model.* **68**: 33–50. doi:[10.1016/0304-3800\(93\)90106-3](https://doi.org/10.1016/0304-3800(93)90106-3)
- Punt, A. E., D. S. Butterworth, C. L. Moor, J. A. De Oliveira, and M. Haddon. 2016. Management strategy evaluation: Best practices. *Fish Fish.* **17**: 303–334. doi:[10.1111/faf.12104](https://doi.org/10.1111/faf.12104)
- R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Available from <https://www.R-project.org/>.
- Ruckelshaus, M. H., P. Levin, J. B. Johnson, and P. Kareiva. 2002. The Pacific salmon wars: What science brings to the challenge of recovering species. *Annu. Rev. Ecol. Syst.* **33**: 665–706. doi:[10.1146/annurev.ecolsys.33.010802.150504](https://doi.org/10.1146/annurev.ecolsys.33.010802.150504)
- San Joaquin River Group Authority. 2013. San Joaquin River Agreement—Vernalis Adaptive Management Plan (VAMP) 2011 annual technical report. 339 p. Available from http://www.sjrg.org/technicalreport/2011/2011_SJRG_AnnualTechnicalReport.pdf, doi:[10.1212/01.CPJ.0000437017.29917.c9](https://doi.org/10.1212/01.CPJ.0000437017.29917.c9)
- Scheffer, M., J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. and others. 2009. Early-warning signals for critical transitions. *Nature* **461**: 53–59, doi:[10.1038/nature08227](https://doi.org/10.1038/nature08227)
- Scheffer, M., S. Carpenter, J. A. Foley, C. Folke, and B. Walker. 2001. Catastrophic regime shifts in ecosystems. *Nature* **413**: 591–596. doi:[10.1038/35098000](https://doi.org/10.1038/35098000)
- Scheuerell, M. D., and J. G. Williams. 2005. Forecasting climate-induced changes in the survival of Snake River spring/summer Chinook salmon (*Oncorhynchus tshawytscha*). *Fish. Oceanogr.* **14**: 448–457. doi:[10.1111/j.1365-2419.2005.00346.x](https://doi.org/10.1111/j.1365-2419.2005.00346.x)
- Scheuerell, M. D., R. Hilborn, M. H. Ruckelshaus, K. K. Bartz, K. M. Lagueux, A. D. Haas, and K. Rawson. 2006. The shiraz model: A tool for incorporating anthropogenic effects and fish-habitat relationships in conservation planning. *Can. J. Fish. Aquat. Sci.* **63**: 1596–1607. doi:[10.1139/f06-056](https://doi.org/10.1139/f06-056)
- Schindler, D. E., and R. Hilborn. 2015. Prediction, precaution, and policy under global change. *Science* **347**: 953–954. doi:[10.1126/science.1261824](https://doi.org/10.1126/science.1261824)
- Sommer, T., F. Mejia, K. Hieb, R. Baxter, E. Loboschewsky, and F. Loge. 2011. Long-term shifts in the lateral distribution of age-0 striped bass in the San Francisco estuary. *Trans. Am. Fish. Soc.* **140**: 1451–1459. doi:[10.1080/00028487.2011.630280](https://doi.org/10.1080/00028487.2011.630280)
- Sommer, T., C. Armor, R. Baxter, R. Breuer, L. Brown, M. Chotkowski, S. Culberson, F. Feyrer, M. Gingras, B. Herbold, and W. Kimmerer. 2007. The collapse of pelagic fishes in the upper San Francisco Estuary: El Colapso de los Peces Pelagicos en La Cabecera Del Estuario San Francisco Fisheries. **32**: 270–277. doi: [10.1577/1548-8446\(2007\)32\[270:TCOPF\]2.0.CO;2](https://doi.org/10.1577/1548-8446(2007)32[270:TCOPF]2.0.CO;2)
- Sommer, T., R. Baxter, and B. Herbold. 1997. Resilience of splittail in the Sacramento–San Joaquin estuary. *Trans. Am. Fish. Soc.* **126**: 961–976. doi:[10.1577/1548-8659\(1997\)126<0961:ROSITS>2.3.CO;2](https://doi.org/10.1577/1548-8659(1997)126<0961:ROSITS>2.3.CO;2)
- Stevens, D. E. 1977. Striped bass (*Morone saxatilis*) year class strength in relation to river flow in the Sacramento–San Joaquin estuary, California. *Trans. Am. Fish. Soc.* **106**: 34–42. doi:[10.1577/1548-8659\(1977\)106<34:SBMSYC>2.0.CO;2](https://doi.org/10.1577/1548-8659(1977)106<34:SBMSYC>2.0.CO;2)
- Szuwalski, C. S., K. A. Vert-pre, A. E. Punt, T. A. Branch, and R. Hilborn. 2015. Examining common assumptions about recruitment: A meta-analysis of recruitment dynamics for worldwide marine fisheries. *Fish Fish.* **16**: 633–648. doi:[10.1111/faf.12083](https://doi.org/10.1111/faf.12083)
- Thomson, J. R., W. J. Kimmerer, L. R. Brown, K. B. Newman, R. M. Nally, W. A. Bennett, F. Feyrer, and E. Fleishman. 2010. Bayesian change point analysis of abundance trends for pelagic fishes in the upper San Francisco estuary. *Ecol. Appl.* **20**: 1431–1448. doi:[10.1890/09-0998.1](https://doi.org/10.1890/09-0998.1)
- Zandvoort, M., M. J. van der Vlist, and A. van den Brink. 2018. Handling uncertainty through adaptiveness in planning approaches: Comparing adaptive delta management and the

water diplomacy framework. *J. Environ. Policy Plan.* **20**: 183–197. doi:[10.1080/1523908X.2017.1347035](https://doi.org/10.1080/1523908X.2017.1347035)

Zeug, S. C., P. S. Bergman, B. J. Cavallo, and K. S. Jones. 2012. Application of a life cycle simulation model to evaluate impacts of water management and conservation actions on an endangered population of Chinook salmon. *Environ. Model. Assess.* **17**: 455–467. doi:[10.1007/s10666-012-9306-6](https://doi.org/10.1007/s10666-012-9306-6)

Acknowledgments

This work was supported through funding from the Metropolitan Water District of Southern California to authors N.T. and B.M.C. The authors would like to thank Ted Sommer, Shawn Acuña, Will Satterthwaite, Alison Collins, and Rebecca Sheehan for their help in identifying additional environment–recruitment relationships; B. J. Miller, Matt Nobriga, Fred Feyrer, Shawn Acuña, and Ted Sommer for their assistance

in recreating the original relationships from their published studies; B. J. Miller, Kathy Hieb of the California Department of Fish and Wildlife San Francisco Bay Study, and Derick Louie of the California Department of Water Resources for their assistance in obtaining raw data needed for retesting relationships; and F. Poulson for his assistance in deriving some of the required abundance indices from raw data.

Conflict of Interest

None declared.

Submitted 20 January 2018

Revised 29 June 2018

Accepted 03 August 2018

Associate editor: Mark Scheuerell