1   # Durability of environment-recruitment relationships in aquatic

2   # ecosystems: insights from long-term monitoring in a highly

3   # modified estuary and implications for management

4   Natascia Tamburello[1†*](ntamburello@essa.com), Brendan M. Connors[2†](Brendan.Connors@dfo-

5   mpo.gc.ca), David Fullerton[3] (dfullerton@mwdh2o.com), and Corey C. Phillis[3]

6   (CPhillis@mwdh2o.com)

7   [1] ESSA Technologies Ltd., Vancouver, BC, V6H 3H4, Canada

8   [2] Fisheries and Oceans Canada, 9860 W. Saanich Rd, Sidney, BC V8L 5T5, Canada

9   [3] Metropolitan Water District of Southern California, 1121 L Street, Suite 900, Sacramento, CA

10   95814, USA

11   [†] These authors contributed equally to this work.

12   * Corresponding author (ntamburello@essa.com)

13   **Running Head:** Durability of environment-recruitment correlations

14   **Keywords:** environment-recruitment, long-term monitoring, regime shift, correlation, fish, step

15   change, resource management

16

17   # Abstract

18   The environment can strongly influence the survival of aquatic organisms and their resulting

19   dynamics. Our understanding of these relationships, typically based on correlations, underpins

20   many contemporary resource management decisions and conservation actions. However, such

21   relationships can break down over time as ecosystems evolve. Even when durable, they may not

22   be very useful for management if they exhibit high variability, context dependency, or non-

23   stationarity. Here, we systematically review the literature to identify trends across environment-

24 recruitment relationships for aquatic taxa from California's San Francisco Bay and Sacramento-

25 San Joaquin Delta Estuary. This is one of the most heavily modified aquatic ecosystems in North

26 America, and home to numerous species of concern whose relationships with the environment

27 inform regulatory actions and constraints. We retested 23 of these relationships spanning 9

28 species using data that have accumulated in the years since they were first published (9-40

29 additional years) to determine whether they persisted. Most relationships were robust (i.e., same

30 or stronger in magnitude) to the addition of new data, but the ability to predict how a species will

31 respond to environmental change did not generally improve with more data. Instead, prediction

32 error generally increased over time and in some cases very quickly, suggesting a rapid regime

33 shift. Our results suggest that more data alone will not necessarily improve the ability of these

34 relationships to inform decision making. We conclude by synthesizing emerging insights from

35 the literature on best practices for the analysis, use, and refinement of environment-recruitment

36 relationships to inform decision making in dynamic ecosystems.

37

Introduction

38

39 The environment can have a profound and complex influence on aquatic organisms and their

40 population dynamics (e.g., Szuwalski et al. 2015). Understanding when and how the

41 environment influences the survival, abundance, and recruitment of fishes and other aquatic

42 organisms has long fascinated and perplexed fish and fisheries scientists and managers (Hjort

43 1914; Cushing 1995; Jacobson and MacCall 1995). Knowing how the environment influences

44 the survival and dynamics of fishes is of general ecological interest because of the light it can

45 shed on, for example, the relative influence of bottom-up and top-down control in ecosystems.

46 Quantifying how the environment influences recruitment can in theory help inform fisheries and

47 improve management. Given the now pervasive influence of humans over the world's aquatic

2

48   ecosystems (e.g., Halpern et al. 2015), understanding when and how the environment -- and

49   human influences on it -- affect the dynamics and abundances of aquatic organisms is critical to

50   many decisions in natural resource management.

51   Relationships between the environment and recruitment, defined here as any relationship

52   between the number of individuals in a population (or their survival rate) and their environment

53   (e.g., river flow, pesticide concentration, water temperature), may break down over time. In a

54   now classic review of fisheries literature, Ransom Myers (1998) found that only 22 of 77

55   environment-recruitment relationships still held after being re-examined with new data. The

56   relationships that were most likely to stand the test of time were those with temperature at the

57   limit of a species' range, where the influence of physical tolerance thresholds outweighs that of

58   more complex ecological interactions. Even when such correlations are reliable, they may not be

59   very useful for informing management if the described relationship is characterized by high

60   variability, context dependency, or non-stationarity as is often the case with recruitment data.

61   Nonetheless, relationships between the environment (which we define broadly as both natural

62   and those aspects under human control) and fish recruitment are central to contemporary

63   resource management decision-making and the conservation of aquatic species. Considering this,

64   and the potential for environment-recruitment relationships to break down over time, there is a

65   pressing need for guidance on best practices for the analysis, use, and refinement of

66   environment-recruitment relationships to inform decision making in natural resource

67   management.

68   California's San Francisco Bay and Sacramento-San Joaquin Delta Estuary (hereafter "Bay

69   Delta") (Figure 1) is an ideal system in which to examine the durability and usefulness (i.e.,

70   predictive power) of environment-recruitment relationships and their implications for decision

71   making. The Bay Delta system has been continuously monitored in a systematic manner for long

72   periods of time, it is heavily altered and managed, and it is home to numerous endangered

73   species and associated regulatory actions and constraints including some which are based upon

74   environment-recruitment relationships (reviewed in Kimmerer 2004). Numerous relationships

75   have been described for taxa in the Bay Delta, but very few have been revisited to test whether

76   these correlations are reliable in the face of new data.

77   Here, we systematically review the literature to identify environment-recruitment relationships

78   for Bay Delta taxa. We then reanalyze the relationships where new data are available to quantify

79   the extent to which the relationships still hold when confronted with new data. Finally, we

80   synthesize emerging insights from the literature on best practices for the analysis, use, and

81   refinement of environment-recruitment relationships to inform decision making in natural

82   resource management.

83   # Methods

84   *Study area*

85   The Bay Delta is made up of a large interior delta formed by the Sacramento and San Joaquin

86   Rivers feeding into a series of basins separated by narrow, deep tidal channels, which flow into a

87   seaward region and ultimately into the San Francisco and San Pablo Bays which are connected to

88   the Pacific Ocean (Figure 1A). The Bay Delta is one of the most heavily modified estuaries in

89   the United States, and is strongly influenced by state and federal water project operations. Two

90   pumping facilities export delta inflows to meet metropolitan and agricultural water needs. Water

91   exports can affect fish directly through entrainment into the pumping facilities, and indirectly

92   though the influence of reduced flows on a wide range of abiotic and biotic variables (Figure

93   1B). Established quantitative relationships between flow characteristics and fish abundance,

94   survival, and migration underpin regulatory decisions regarding levels of allowable exports of

4

95   river flows and minimum outflow from the Bay Delta that maintain fish production and habitat

96   within acceptable bounds (CDFW 2016a,b).

97

98   *Literature review*

99   We carried out a literature search of peer-reviewed publications, grey literature, and government

100   agency reports to identify published examples of relationships between environmental variables

101   and population parameters of aquatic Bay Delta species. Our initial search was carried out using

102   Google Scholar and various combinations of search terms based on variables expected to

103   influence organism abundance as described in comprehensive reviews of Bay Delta ecology

104   (e.g., Kimmerer 2004; CDWR and USBR 2016). These search terms included "fish",

105   "invertebrate", "abundance", "survival", "entrainment", "migration", "environment-

106   recruitment", "environmental variable", "X2", "flow", "conductivity", "turbidity", and "prey

107   density". The resulting set of publications and reports was then supplemented by consulting

108   regulatory documents to identify additional relationships that underpin contemporary

109   management decisions.

110   We created a catalogue of all identified publications to document their various characteristics,

111   including publication year, focal species, the number of relationships examined, and whether the

112   publication is cited in regulatory documents as informing management decisions (included here

113   as Supplementary Online Material). Because most of the government agency and grey literature

114   reports were review documents that reproduced results from primary literature, we chose to focus

115   further investigations only on the peer-reviewed literature. For each peer-reviewed study in our

116   catalogue, we extracted each individually reported relationship into a second catalogue and

117   documented characteristics including the focal species, predictor and response variables, type of

118   analysis, timeframe, primary data source, and reported model outputs including intercept and

119   slope parameters, R^2, p-value, AIC, and others. This catalogue of relationships was used both to

120   carry out a qualitative analysis of trends in the study of such relationships in the Bay Delta, and

121   to select a subset of environment-recruitment relationships suitable to retesting to quantify the

122   extent to which the relationships still hold when confronted with new data.

123   *Relationship selection criteria*

124   We developed three tiers of criteria used to screen the full set of published environment-

125   recruitment relationships to identify those suitable for reanalysis.

126   Tier 1 criteria excluded relationships for which reanalysis would be impractical due to (1)

127   inability to obtain new data for reanalysis because of reliance on either one-time experiments

128   (e.g., paired releases of tagged fish), data collection programs that have since ended, or second-

129   order variables derived via complex integration of many other environmental variables; (2) use

130   of analytical methods that would be impractical to replicate for a review study of this scale (e.g.,

131   whole ecosystem simulation models) or which make it difficult to compare strength and

132   statistical support across relationships (e.g., non-linear correlations such as GAMs, rank

133   analyses, etc.); or (3) because the published relationship is so recent there would be few

134   additional data points available (i.e., we excluded relationships if the number of years of new

135   data were less than the number in the original time series or less than 10 years, whichever was

136   less). In practice, this resulted in retention of linear correlations based on data from long-term

137   monitoring programs with publicly available data sets, which formed a large majority (~75%) of

138   all the published relationships in our dataset examined prior to the application of screening

139   criteria.

6

140   Tier 2 criteria screened out relationships with little or no statistical support in the initial analysis,

141   given that the objective of our work was to evaluate the extent to which established relationships

142   hold over time. For this purpose, we defined little or no statistical support as $p > 0.1$, or $\Delta AIC >$

143   2 from the top model or, where neither is reported in older publications, the absence of a fitted

144   regression line and equation on a scatter plot of the relationship.

145   Lastly, Tier 3 criteria were used to reduce redundancy in the remaining set of relationships by

146   selecting only a single relationship for each unique species – variable pair for reanalysis. When

147   more than one relationship existed for the same species-variable pair relying on the same source

148   data (e.g., delta smelt ~ X2), the following sub-criteria were used to select a single relationship

149   that was: (1) most recently updated among the set to avoid reproducing past retests; (2), the

150   longest time-series among the set for greater statistical power; (3) based on indices of abundance

151   rather than extrapolated abundance estimates to reduce propagation of error, and (4) where all

152   relationships in the set are statistically significant, the relationship with the strongest support

153   based on variation explained (e.g., via R^2) or other rationale provided by the authors. We

154   considered similar relationships using different source data (e.g., delta smelt $_{(Fall\ Midwater\ Trawl)}$ ~

155   X2 and delta smelt $_{(Bay\ Study\ Midwater\ Trawl)}$ ~ X2) distinct and therefore were retained due to known

156   differences in methodology, target life stage, and conclusions that can be drawn from alternative

157   source surveys.

158   The full catalogue of original and screened publications is available in the Supplementary Online

159   Materials.

160   *Data sources*

161   For each relationship retained for reanalysis, we sought to obtain the same data used in the

162   original analysis from the source identified by the authors. In some cases, this data was available

163   directly, and in others, annual means of data needed for analysis were derived from raw data sets

164 using the same or similar methods originally described by the authors, including applying data

165 transformations. In some cases, we were not able to reproduce the data using the methods

166 described by the authors and either used a modified method or excluded the relationship from

167 our analysis (Table 1). Population data for these relationships were derived from the California

168 Department of Fish and Wildlife's fall midwater trawl (FMWT), summer townet survey (TNS),

169 beach seine surveys, and salvage surveys; the Bay Study otter trawl (Bay OT) and midwater

170 trawl (Bay MWT) surveys; and the U.S. Fish and Wildlife Service's Chinook salmon trawls.

171 Environmental data was derived primarily from Hutton et al., (2015) and calculations therein for

172 X2, from the California Department of Water Resources DayFlow data portal for flows, from

173 environmental data collected alongside population data as part of the FMWT survey, and in

174 some cases directly from the study authors (Table 1, Supplementary Online Materials).

175 *Analysis*

176 We conducted a quantitative reanalysis of selected environment-recruitment relationships.

177 Source data for the analyses was transformed (as per original analyses; see Table 1), subset to the

178 relevant timeframes, and standardized (by subtracting the mean and dividing by the standard

179 deviation) to allow estimated relationships to be comparable across species and environmental

180 variables. All relationships selected for reanalysis were based on linear correlations and so scaled

181 data was analyzed using ordinary least squares regression with an abundance measure as the

182 response variable and one or more environmental factors as the predictor variable, with none of

183 the ultimately selected relationships including interactions among variables

184 Several relationships selected for reanalysis (e.g., Kimmerer et al. 2009) included a step-change

185 or data-splitting before and after the introduction of invasive Asian clam (*Corbula amurensis*)

186 into the Bay Delta in 1986. Introduction of this clam is thought to be responsible for a sudden

187 and substantial decline of zooplankton and, subsequently, of fish in the Bay Delta (Kimmerer et

188   al. 1994). Where published analyses found statistical evidence for a split or step change in 1987

189   (using a dummy variable), we chose to reanalyze the relationship using only data after this cut-

190   off both to simplify the analyses and because prior data could be considered irrelevant for

191   interpreting how relationships have changed in the context of additional data.

192   For each relationship, we quantified the strength (R^2; correlation coefficient) and magnitude

193   (slope; in standardized units) of the relationship based on the scaled original and updated time

194   series and then compared them to quantify the extent to which the relationships still held when

195   confronted with new data. We also quantified prediction error using unscaled original and

196   extended time series for each relationship to characterize how well the observed relationships

197   would be expected to predict future (out of sample) observations and to provide an indication of

198   how useful the relationship may be from a decision making and management perspective.

199   Prediction error was estimated as the normalized root mean squared prediction error (CV_n):

200     $$CV_n = \left. \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{y}_i - y_i}{1 - h_i}\right)^2} \middle/ \frac{1}{n}\sum_{i=1}^{n} y_i \right. \qquad \text{(eqn. 1)}$$

201   where h_i is the diagonal element of the operator matrix that produces the least squares fit (i.e.,

202   hat matrix). This measure of prediction error can be interpreted as the percent error in future

203   predictions relative to the average observed abundance for a given relationship. For example, a

204   prediction error of 100% would mean that the relationship allows us to predict future abundances

205   to within +/- 50% of the mean predicted abundance.

206   We compared the raw data, parameters, and fit of relationships using the original time series to

207   those reported in the original publications to ensure that our approach successfully replicated the

9

208   previously published relationships before proceeding to retesting. Instances where we were not

209   able to reproduce past relationships were not further considered (detailed in Table 1).

210   Our retests of environment-recruitment relationships using data that has accumulated since a

211   relationship was first established has the potential to be biased by an imbalance in the number of

212   pre- to post-retest observations. In instances where there are many (e.g., 30) years of pre-retest

213   observations and only a few (e.g., 10) post-retest observations it is possible that the pre-retest

214   observations obscure what is otherwise a weakening or different relationship in the post-retest

215   observations. To quantify the extent to which this was the case with the relationships we retested,

216   we carried out a secondary analysis where we randomly sub-sampled the pre-retest data so that

217   there was an equal number of pre- and post-retest data points, and then, as above with the full

218   time series, quantified the strength and magnitude of the relationship based on the subsetted data.

219   We repeated this exercise 1000 times for each relationship, and then compared the median

220   strength and magnitude of the relationship based on the original (subsetted) and updated time

221   series.

222   All analyses were carried out in the R statistical software suite (R Core Team 2017), and we

223   provide the source code, and data, for our analyses in the Supplementary Online Materials.

Results

Literature review

226   Our literature search identified 98 publications describing environment-recruitment relationships

227   in the Bay Delta ecosystem. Of these, 40 were reviews citing primary literature or offered only

228   general observations on raw data without conducting any analysis, and 3 were not available

229   online and so not examined further. The remaining 55 peer-reviewed publications were retained

230   for further analysis. This subset included papers published between 1977 and 2017, with a mean

231   publication date of 2002. These studies examined an average of 10 relationships per study, and a

232   minority of papers examined a large number (>100) of competing models describing the same

233   relationship. Each study examined between 1 and 17 focal species (mean ~ 3).

234   These 55 peer-reviewed studies described 420 individual relationships which overwhelmingly

235   focused on examining the influence of environmental variables on population abundance as

236   opposed to other biological characteristics (Figure 2A). This is likely a result of the fact that

237   roughly 70% of all relationships relied on publicly available long-term abundance survey data

238   that has been collected in the Bay Delta for decades (e.g., California Department of Fish and

239   Wildlife fall midwater trawl (FMWT), summer townet survey (TNS), and San Francisco Bay

240   Study otter trawl (OT) and midwater trawl (MWT) surveys), whereas other types of population

241   variables are generally not routinely collected for most species and were obtained using one-off

242   experiments or surveys to inform 18% of these relationships. The distribution of relationships

243   was also strongly biased towards species that are either currently or historically listed as

244   Threatened or Endangered (Figure 2B). There was a more even distribution of environmental

245   variables examined across studies, but the most frequently examined variables were X2 (the

246   distance in kilometers from the Golden Gate up the axis of the estuary to where the tidally-

247   averaged near-bottom salinity is 2 ppt or 2×10^{-9} mg/L; Jassby et al. 1995), or other measures of

248   salinity, and flow (Figure 2C). Most relationships examined only a single environmental variable

249   (maximum of 7) through simple linear correlation (60% of all relationships). However, the

250   number of variables has generally increased in more recent studies as researchers continue to

251   adopt more complex analytical methods including multivariate models, generalized additive

252   models, and whole-ecosystem simulations. While about 42% of all relationships have been

253   identified as underpinning regulatory decision-making, only 14% of all relationships have been

11

254   retested to quantify their durability (i.e., extent to which their magnitude of effect holds up in the

255   face of new data).

256   Across the published relationships there was broad variation within and among environmental

257   variables in the amount of variation in the abundance that was explained. For example, some

258   environmental variables like flow and salinity tended to explain over half the variation, while

259   others such as temperature and water export rates (measured in cfs - cubic feet per second - or

260   m^3/s, where 35.31 cfs = 1 m^3/sec) explained much less variation in abundance (Figure 3).

261

262   *Reanalysis of bay delta environmental-recruitment relationships*

263   Applying our screening criteria to the full set of relationships resulted in 31 relationships suitable

264   for retesting (Table 1). The publications originally reporting these relationships varied widely in

265   the details provided on methodology, and it was often difficult to determine how the

266   environmental variables were derived or what the source of the data was that was used to derive

267   them. We were unable to complete re-analyses for 8 of these 31 relationships due to an inability

268   to re-create the input variables using the methodology described by the authors or because of

269   errors or missing information in the original publications, leaving 23 relationships (17 univariate

270   and 6 multivariate) from 8 publications which were fully re-analyzed using the most recent

271   available data (between 9 and 40 additional years, median 9 years).

272   When updated data were used to retest previously published relationships, the direction and

273   statistical significance of the relationships remained the same (Figure 4; Table 1). Of the

274   univariate relationships, 9 of 17 became stronger (i.e., either more negative or positive depending

275   on the original relationship), 3 of 17 became weaker and 5 of 17 remained nearly identical. These

276   general patterns remained the same for the multivariate relationships (Table 1) and when the

277   dataset used to retest relationships was balanced to achieve an equal number of original and

278   updated data points (Figure 4; Table 1).

279   In some instances, the addition of more years of data resulted in more variation in the

280   relationship being explained (i.e., higher R^2; Figure 5A, B). However, on average considering

281   more recent data did not increase the estimated strength of the relationships, and in some cases

282   reduced it (Figure 5A). For example, relationships between striped bass abundance and flow saw

283   substantial declines in R^2 with the addition of more years of data, despite the fact the magnitude

284   of the relationship remained nearly identical (Figure 4). This occurred because the overall

285   abundance of striped bass declined between the original time series and the updated one, which

286   has been attributed to introduction of the Asian clam in 1987, but abundance still increased with

287   increases in flow during both periods (Figure S1).

288   In most cases, considering more years of data did not improve the predictive power of the

289   environment-recruitment relationships. Instead, counterintuitively, the prediction error of each

290   relationship typically increased with the addition of more years of data (Figure 5C, D). The

291   median percent increase in prediction error across the relationships retested was 30%. As with

292   the variation explained, these changes were most pronounced for relationships between striped

293   bass abundance and flow.

Discussion

295   Our review of the literature identified 98 publications examining 420 individual environment-

296   recruitment relationships in the Bay Delta. These relationships overwhelmingly focused on the

297   influence of environmental variables on population abundance as opposed to other biological

298   characteristics and were biased towards species that are either currently or historically listed as

299   Threatened or Endangered. About half of these relationships are used in regulatory decision-

13

300   making, but only one in five relationships have been retested to quantify the extent to which they

301   stand the test of time. We retested 23 relationships using data that have accumulated in the years

302   since they were first published. In contrast to Myers (1998), who found that the proportion of

303   relationships that held up when retested was low, we found that when new data were used to

304   retest previously published Bay Delta relationships, the direction and statistical significance of

305   the relationships remained the same, though the amount of variation explained by the

306   relationships and our ability to predict how a species will respond to change in their environment

307   did not generally improve with more data. Instead, in most cases, prediction error actually

308   increased when extending the time series, suggesting that accumulating more data will not

309   necessarily improve the ability of these relationships to inform decision making (insofar as

310   predictive power is useful for decision making). Perhaps this should not come as a surprise given

311   the original relationships examined here were typically identified based on their ability to explain

312   observed data (e.g., R^2 from linear regression) as opposed their ability to predict future

313   observations using, for example, approaches like data splitting and cross-validation (Power 1993;

314   Harrel 2015). However, a large reduction in predictive power when relationships are re-tested

315   with more data may be diagnostic of an established relationship that is breaking down (i.e., as is

316   the case with striped bass and potentially Sacramento splittail in Figure 5), and might prompt

317   action to search for unmeasured drivers of this change.

318   Our review of published environment-recruitment relationships in the Bay-Delta also highlights

319   some methodological shortcomings of studies in this discipline. First, our review has made clear

320   the great value of long-term data collection programs that follow standardized and consistent

321   protocols to detect and validate long-term trends in biological variables, and has shown that a

322   large share of studies in this space leverage these datasets. However, the availability and

323   accessibility of such survey data may also reduce the likelihood that researchers in this region

324   will embark on independent data collection to study other species and biological variables that

14

325   are not the focus of existing surveys. Second, despite the accessibility of long-term survey data

326   and the simplicity of analysis by correlation, we were still unable to reproduce originally

327   published parameters for many relationships otherwise meeting our criteria for re-testing. In

328   some cases, this was due to errors in the original work, unreported assumptions about data

329   transformations that only became clear after contacting the authors, or the prior use of

330   interpolated data that has since been corrected at the source by survey operators. These

331   challenges highlight the importance of reproducibility in research in general, and into

332   environment-recruitment relationships in particular, echoing a growing call for greater

333   reproducibility both in ecology and across other scientific disciplines (Cassey and Blackburn

334   2006; Nosek et al. 2015; Borregaard and Hart 2016). Trends towards the use of coding for

335   ecological analysis, and for the publication of that code alongside manuscripts as we do in our

336   Supplementary Online Materials for this study, will play a significant role in improving the

337   reproducibility of ecological research going forward (Mislan et al. 2016).

338   Our conclusions should be considered in light of potential biases in both our study selection

339   criteria and in the type of data these studies draw upon. As with all studies based on literature

340   review, our results are subject to publication bias (Cooper et al. 2009) relating to our decision to

341   focus on peer-reviewed studies, a propensity towards publication of significant relationships in

342   peer-reviewed journals, and a disproportionate number of publications on particular variables

343   (e.g., X2 and flow) coming from a few very active authors in this field. In addition, recent work

344   has shown that the long-term survey data used to create many of these relationships may itself be

345   inherently biased by unquantified changes in detection probability. Detection probability, or

346   catchability, may vary considerably over time with (1) overall abundance (i.e., it is more difficult

347   to catch a rarer species or size class) (Mahardja et al. 2017), (2) with changing environmental

348   conditions (e.g., catchability may decrease with increasing water clarity as fish are better able to

349   see and avoid survey gear in clearer water) (Latour 2016), and (3) with differences in gear type

15

350   across surveys (J. T. Peterson and M. Barajas, unpubl.). When not accounted for, these changes

351   in detectability may be incorrectly interpreted as real changes in abundance or occupancy.

352   Notwithstanding these potential biases, we believe that the breadth of species, environmental

353   variables, and survey types covered in our analysis still allows us to draw some more general

354   conclusions about the utility of environment-recruitment models, and to synthesize insights from

355   the literature on best practices for the analysis, use, and refinement of environment-recruitment

356   relationships to inform decision making in natural resource management.

357   *Correlation, causation, and strength of evidence*

358   Correlations underpin most natural resource management decisions where one must predict how

359   the environment responds to human action. These correlations are based on experience, and

360   accumulated observations of the relationship between a species (e.g., abundance or survival) and

361   its environment (e.g., flow, temperature, prey abundance, etc.). Such historic relationships are

362   usually assumed to be causative, but we are often reminded of the adage "correlation does not

363   equal causation". Ideally, manipulative experiments can be used to determine whether a specific

364   human action causes a response in an ecosystem component. Such learning by manipulation

365   embraces the three key elements of experimental design: controls, randomization of treatments,

366   and replication. Controls (i.e., monitoring systems that are not subject to a specific management

367   intervention) provide critical contrast needed to disentangle the effects of the management

368   actions (treatment) from other system change. Randomization ensures that the choice of which

369   systems receive the management intervention is based on chance, thereby reducing the potential

370   for confounding and bias due to factors not accounted for in the experimental design. Lastly,

371   replication, whereby treatment and control groups are replicated over space and/or time, allows

372   for more precise estimates of the effect of the management intervention.

16

373   Much has been learned in natural resource management through manipulative experiments. For

374   example, hatcheries manipulate the timing and size of fish released to determine which

375   combination result in optimal survival (e.g., Irvine et al. 2013) and hydroelectric facility

376   operators manipulate timing and magnitude of flow releases to determine which flows are most

377   likely to improve fish survival (e.g., Bradford et al. 2011). While manipulative experiments are

378   the gold standard approach to establishing causation, scope for manipulative experiments

379   decreases at increasing scales, and so we are left interpreting correlative relationships in order to

380   manage some of the largest human-influenced ecosystems in the world.

381   When opportunities for learning by manipulation are limited or impossible, the weight of

382   evidence for a hypothesized causal correlation can be assessed based on the strength,

383   consistency, specificity and plausibility of the mechanism underlying the relationship (e.g., Hill

384   1965; Hilborn 2016). The stronger the association between two variables and implied magnitude

385   of effect (e.g., a small [2%] versus large [50%] change in abundance), the more likely the

386   relationship is causal. When the same relationship is found repeatedly across space (e.g., in other

387   systems or species) and over time (e.g., with addition of more years of data), the stronger the

388   evidence that it is indeed directly causal or related to the same causal driving variable (e.g.,

389   relationships to flows ultimately driven by precipitation) and not due to another unmeasured

390   confounding factor. How likely are alternative explanations (hypotheses) for a given

391   relationships? When few or no alternative explanations exist, causality is more likely than under

392   circumstances where alternative explanations abound. Lastly, the plausibility of the mechanism

393   underlying the association between two variables can shed light on the likelihood it is causal;

394   when there is clear evidence of a mechanism that could be responsible for the relationship (e.g.,

395   from laboratory experiments) then there is stronger evidence that the relationship is indeed

396   causal relative to instances where support for (or nature of) the underlying mechanisms is

397   unknown.

17

398   While strength, consistency, specificity and plausibility can help guide the degree of support for

399   a given relationship, they should not come at the cost of maintaining multiple working

400   hypotheses, and evaluating the evidence for each simultaneously when using correlations to

401   guide decision making (Hilborn and Mangel 1997; Plowright et al. 2008). An illustrative

402   example of the simultaneous consideration of multiple working hypotheses is the development

403   and application of a state-space multi-stage life cycle model to investigate for drivers of

404   population decline in delta smelt by Maunder and Deriso (2011).

405   *Adaptive management in the face of ecosystem change*

406   Even when relationships are truly causative, using past relationships to guide future management

407   decisions can fail to have the intended effect when the system in which they occur changes over

408   time (i.e., exhibits non-stationary). Such non-stationarity in aquatic systems can arise from both

409   slow-moving environmental change or rapid regime shifts and "tipping points" between

410   alternative stable states (Scheffer et al. 2001, 2009). There is widespread evidence of regime

411   shifts in aquatic ecosystems arising from both natural (e.g., climate) and human (e.g., pollution,

412   species introductions) caused factors (Carpenter 2003; Hunsicker et al. 2016) and the state of the

413   ecosystem can have a strong influence on the outcomes of management actions. For example,

414   large releases of hatchery salmon reduce the survival of endangered wild salmon but only during

415   periods of poor ocean conditions (Levin et al. 2001), and translocation of wild juvenile salmon

416   past hydropower installations carries greater benefits for their ocean survival in cooler but not

417   warmer oceanic regimes (Gosselin et al. 2017). These examples highlight the fact that the

418   benefits of management interventions (e.g., reducing hatchery releases to minimize impacts to

419   wild fish, translocation of fish past barriers) are contingent upon the ocean regime the system is

420   experiencing in any given year. In the Bay Delta, the introduction and rapid expansion of the

421   invasive Asian clam in the late 1980s is believed to have caused a major increase in grazing

18

422 pressure on phytoplankton, leading to a persistent decline in overall pelagic food resources

423 (Carlton et al. 1990; Nichols et al. 1990; Baxter et al. 2010). The so-called "step change" towards

424 this new stable state has had varying influences on different species within the ecosystem. For

425 example, the regime shift due to Asian clam has led to a change in the overall magnitude

426 (intercept) but not the rate of change (slope) of existing abundance-flow relationships for striped

427 bass (Kimmerer et al. 2009, and reproduced in this study), but has been suggested to have driven

428 a new abundance-flow relationship for threatened Delta smelt that brought with them significant

429 implications for the way flows in this system are managed (CDFW 2016a,b).

430 Given the ubiquity of regime shifts and non-stationarities in aquatic systems (Möllmann and

431 Diekmann 2012), including in the Bay Delta (Kimmerer 2002; Kimmerer et al. 2009; Thomson

432 et al. 2010), how should one evaluate the evidence for environment-recruitment relationships and

433 use them to inform decision making when system change is suspected? In some cases, change in

434 a system will be so pronounced that there is little question about when it occurred and so the

435 nature of a relationship post regime change should be the one that is used to inform future

436 management actions. In other instances, the timing of abrupt change, and indeed whether it has

437 occurred at all, will be uncertain for several years and so debated (at times vociferously) until

438 clear evidence accumulates that a regime shift has occurred. In such cases, quantifying the

439 statistical support for changes in system state using, for example, change-point analyses

440 (Thomson et al. 2010) or dummy variables in linear regression (e.g., Kimmerer 2002; Kimmerer

441 et al. 2009), is one way to quantitatively evaluate the evidence for system change. In addition,

442 and perhaps more importantly, one should quantitatively evaluate the decision-making

443 consequences of incorrectly assuming a regime shift has or has not occurred so as to be able to

444 understand and clearly communicate the costs (e.g., biological, economic and social) of getting it

445 wrong. Lastly, it has also been found that when non-stationarity is present (or suspected to be

19

446   present), using recent observations to predict the consequences of alternative management

447   actions can improve management outcomes (e.g., Ianelli et al. 2012).

448   In systems that have undergone dramatic change (e.g., tipping points), failure to regularly re-

449   evaluate the durability and predictability of environment-relationships risks making management

450   decisions based on information with increasingly large margins of error, with the potential for

451   negative ecological, social and economic consequences. To illustrate this point we estimated

452   prediction error for a few Bay delta environment-recruitment relationships spanning the period

453   before and after the Asian clam invasion in 1987 for species that have (striped bass and longfin

454   smelt; Kimmerer et al. 2009), and have not (Sacramento splittail; Kimmerer et al. 2009),

455   responded to the invasion (Figure 6). For those species that declined in abundance coincident

456   with the invasion, failure to account for this change results in relationships with increasing

457   prediction error as time goes by after the invasion, relative to relationships that account for the

458   change by including a step change in 1987 (striped bass and longfin smelt; Figure 6). In contrast,

459   for Sacramento splittail, which did not appear to respond to the invasion, there is no benefit to

460   including a step change in the relationship. Interestingly, these analyses suggest that the

461   environment-recruitment relationships for striped bass and longfin smelt have experienced

462   subsequent regime shifts (~1995 for striped bass and ~ 2005 for longfin smelt) that have further

463   eroded their predictive power. This subsequent shift may be explained by observed changes in

464   distribution likely to affect catchability in long-term surveys. Prior studies have suggested that a

465   reduction on pelagic food resources due to overgrazing by Asian clam appears to have driven

466   shifts in the distribution of young fish in the 1980s and 1990s away from the primary sampling

467   regions of long-term surveys and towards areas characterized by fewer clams and better foraging

468   prospects. This manifested as a lateral shift from deeper channel habitat preferentially sampled

469   by annual surveys towards shallower slough habitat for striped bass (Sommer et al. 2011), and as

470   a longitudinal shift from upstream habitat towards more saline downstream habitat for longfin

20

471   smelt (Baxter et al. 2008; Sommer et al. 2011). Thus, these two shifts in prediction error align

472   and illustrate with two stages rapid environmental change driven by a trophic cascade, the first

473   likely related to a species introduction, and the second likely related to two different behavioural

474   responses to adapt to the consequences of this introduction.

475   Our findings suggest that when environment-recruitment relationships underpin decision-making

476   they should be re-evaluated on a regular basis as part of a broader adaptive management

477   approach to ensure that they remain robust in the face of new data and continue to provide an

478   accurate representation of a continually evolving ecosystem. Such an adaptive approach to

479   evaluating environment-recruitment relationships is aligned with broader calls for increasing the

480   implementation of more proactive adaptive management in Bay Delta ecosystems to address

481   accelerating environmental change (Delta Independent Science Board 2015, 2016; Zandvoort et

482   al. 2017).

483   *Environment-recruitment relationships in decision making*

484   The widespread use, and at times misuse, of environment-recruitment relationships to inform

485   decision making has produced several general insights into best practices for incorporating such

486   relationships into natural resource management.

487   First, uncertainty should be both quantified and propagated in any analysis that seeks to predict

488   the consequences of alternative management actions and identify those actions most likely to

489   achieve desired objectives. This uncertainty comes in at least four distinct forms (e.g., Peterman

490   2004): (1) natural variation in both physical and biological processes, (2) uncertainty due to

491   imperfect assessment arising from measurement error, (3) structural uncertainty due to

492   incomplete understanding of how a system functions, and lastly (4) outcome uncertainty or

493   implementation error in how well a given management target (e.g., increase flow by 20%) is

21

494   achieved by a management action (e.g., releasing water from a reservoir). While uncertainty

495   resulting from points 2-4 can in theory be reduced by improved measurements, greater

496   understanding of system function, and better management control, all of which can be

497   accomplished to some extent with the collection of more years of data, uncertainty arising from

498   natural variability is irreducible. We found that prediction error was not reduced for Bay Delta

499   environment-recruitment relationships with the accumulation of more years of data, and in fact

500   increased in many cases. This finding suggests that natural variation in the physical or biological

501   processes of the system is an important and ongoing source of uncertainty that will limit the

502   extent to which improvements in predictions of how taxa from the Bay Delta will respond to

503   changing environmental conditions and human action can be achieved.

504   Second, there is increasing recognition of the value of developing and using life cycle models to

505   evaluate the predicted consequences of alternative management actions on species of concern in

506   the face of this uncertainty (e.g., Good et al. 2007; Ruckelshaus et al. 2002; Zeug et al. 2012). In

507   contrast to single life stage, habitat type, or environmental relationships, life-cycle models

508   simultaneously consider extrinsic (environment, management action) and intrinsic (density

509   dependence) factors influencing multiple life stages. Life cycle models can either be mechanistic

510   where survival between life stages is based on specific mechanisms (Scheurell et al. 2006) or

511   statistical where life stage specific survival is not defined by specific mechanistic relationships

512   (Nobriga and Rosenfield 2016). The use of life cycle models allows for more realistic and

513   comprehensive evaluation of the predicted outcomes of alternative management actions than

514   considering single life stage, habitat type, or environmental relationships, because it considers

515   environmental effects across linked stages in a life cycle while also accounting for population

516   processes (e.g., growth, movement, mortality and reproduction).

517   Even when uncertainty is successfully incorporated into modelling approaches, the broader

518   question remains – how can we account for and propagate uncertainty into the broader

519   management of a complex system with many conflicting management objectives when our

520   understanding of that system is not, and will never be, complete? In a review of management

521   approaches in the Bay-Delta ecosystem for the National Research Council, the Committee on

522   Sustainable Water and Environmental Management proposes that agencies should adopt

523   management approaches that assume " 'universal nonstationarity,' or the idea that all aspects of

524   the environment will constantly be changing." (NRC 2012). Such approaches may prove

525   challenging for many traditional decision-making pathways, which are often constrained by

526   static or slow-moving policy frameworks (Aladjem 2013; Delta Independent Science Board.

527   2016). However, a number of approaches with shared characteristics have been developed to

528   help facilitate formalized decision-making in complex systems characterized by their uncertainty

529   and are among the approaches recommended by the NRC review. Among these are decision

530   analysis or decision scaling, risk assessment, and management strategy evaluation.

531   Decision analysis, a form of risk assessment, is a systematic approach to incorporating

532   uncertainties in the managed system into decision making using models to project outcomes of

533   alternative actions (Peterman and Anderson 1999). These outcomes (or performance measures)

534   act as criteria for comparing and ranking the actions against specific objectives (e.g., ecological,

535   social and economic). Decision analysis can be coupled with stochastic simulations (e.g., Peters

536   and Marmorek 2001) to identify management "actions" that explicitly consider multiple sources

537   of uncertainty, where management actions refer to a combination of sampling design for data

538   collection, approaches to analyzing the data, and decision rules for how to manage a specific

539   aspect of the system. However, such risk assessments and decision analysis can overestimate risk

540   by failing to account for management reactions to the information provided by future data. An

541   alternative version of the approach known as "decision scaling" begins with a bottom-up analytic

23

542   framework that first identified the most critical management uncertainties that influence

543   decision-making, and uses this information to tailor top-down environmental monitoring and

544   modelling to provide the most credible information at appropriate temporal and spatial scales to

545   project the outcomes of alternative actions (e.g., different flow management regimes) under a

546   range of possible futures (e.g., wet and dry years) (NRC 2012). Explicitly considering

547   management feedback is also a central feature of what has been variously referred to as closed-

548   loop simulations (Walters 1986), Management Strategy Evaluations (Sainsbury et al. 2000), and

549   Management Procedure Evaluations (Butterworth and Punt 1999), which are in effect risk

550   assessment methods and formal decision analyses which consider management feedback. While

551   most widely applied in relation to fisheries and cetacean conservation and management,

552   Management Strategy Evaluations have also been used to inform decision making focused on

553   multi-species and ecosystem objectives (reviewed by Punt et al. 2014). All of these approaches

554   lend themselves to robust evaluation of the value of different sources of information, including

555   environment-recruitment relationships and the uncertainty inherent therein, into a broader

556   assessment of the best way forward for natural resource management decisions in complex

557   systems.

558   *Conclusions*

559   Moving forward, there is a growing recognition of the importance of maintaining multiple

560   working hypotheses when quantifying the support for correlations in environmental management

561   (Hilborn 2016), that quantitative assessment of policies that consider these relationships should

562   be done using approaches that allow for realistic incorporation and propagation of multiple

563   sources of uncertainty (Peterman 2004), and that, ultimately, managers in the Bay-Delta and

564   elsewhere should identify policies that are robust to a range of alternative hypotheses (NRC

565   2012; Schindler and Hilborn 2015).

24

566   Despite advances in the tools available to improve our assessment of environment-recruitment

567   relationship and their consideration in decision making, we should remain humble in our zeal to

568   either accept them as fact or discount them entirely because they are "just correlations". As Hill

569   emphasized in his 1965 Presidential Address on correlation and causation to the Royal Society of

570   Medicine (Hill 1965): "All scientific work is incomplete – whether it be observational or

571   experimental… [and] is liable to be upset or modified by advancing knowledge. That does not

572   confer upon us a freedom to ignore the knowledge we already have, or to postpone the action

573   that it appears to demand at a particular time."

574

Acknowledgements

585

References

587   Aladjem, D. 2013. When All You Have Is a Hammer: Scientific Uncertainty and Bay-Delta

588   Policy, 1995 to Present. Nat. Res. & Env. 27(4):36-39.

589   Baxter, R., Breuer, R., Brown, L., Chotkowski, M., Feyrer, F., Gingras, M., Herbold, B.,

590   Mueller-Solger, A. and others. 2008. Pelagic organism decline progress report: 2007

591   synthesis of results. Interagency Ecological Program for the San Francisco Estuary.

592   Baxter, R., Breuer, R., Brown, L., Conrad, L., Feyrer, F., Fong, S., Gehrts, K., Grimaldo, L., and

593   others. 2010. Interagency Ecological Program 2010 Pelagic Organism Decline work plan

594   and synthesis of results. Interagency Ecological Program for the San Francisco Estuary.

595   Bennett, W.A. 2005. Critical assessment of the delta smelt population in the San Francisco

596   Estuary, California. San Francisco Estuar. Watershed Sci. 3(2): 1-71.

597   Borregaard, M.K. and Hart, E.M. 2016. Towards a more reproducible ecology. Ecography.

598   39(4): 349-353.

599   Bradford, M.J., Higgins, P.S., Korman, J. and Sneep, J. 2011. Test of an environmental flow

600   release in a British Columbia river: does more water mean more fish? Freshwater

601   Biol. 56(10): 2119-2134.

602   Brandes, P., Burmester R., and Apeegle J. 2006. Estimating Relative Abundance and Survival of

603   Juvenile Chinook Salmon in the Sacramento - San Joaquin Estuary. Interagency Ecological

604   Program for the San Francisco Estuary. 7 p

605   Butterworth, D.S. and A.E. Punt. 1999. Experiences in the evaluation and implementation of

606   management procedures. ICES J. Mar. Sci. 56:985-998.

607   California Department of Fish and Wildlife (CDFW). 2016a. CDFW Rationale for Summer

608   Delta Flow Augmentation for Improving Delta Smelt Survival (CDFW 7/8/16). Retrieved

609   from: https://www.nrdc.org/sites/default/files/media-uploads/cdfw_outflow.pdf.pdf

610   Retrieved on: 09/09/2017

611   California Department of Water Resources and U.S. Bureau of Reclamation (CDFW). 2016b.

612   Final Environmental Impact Report/Environmental Impact Statement for the Bay Delta

613   Conservation Plan/California WaterFix—Volume I. Final EIR/EIS for the BDCP/California

614   WaterFix. December. (DOE/EIS-0515.) (ICF 00139.14.) Prepared by ICF International,

615   Sacramento, CA.

616   Carlton, J.T., Thompson, J.K., Schemel, L.E. and Nichols, F.H., 1990. Remarkable invasion of

617   San Francisco Bay (California, USA) by the Asian clam *Potamocorbula amurensis*. I.

618   Introduction and dispersal. Mar. Ecol. Prog. Ser. pp.81-94.

619   Carpenter, S. R. 2003. Regime shifts in lake ecosystems: patterns and variation. Volume 15 in

620   Outline of Excellence in Ecology Series. Ecology Institute, Oldendorf/Luhe, Germany.

621   Cassey, P. and Blackburn, T.M. 2006. Reproducibility and repeatability in ecology. BioScience.

622   56(12): 958-959.

623   Cooper, H., Hedges, L.V. and Valentine, J.C. eds. 2009. The handbook of research synthesis and

624   meta-analysis. Russell Sage Foundation.

625   Cushing, D.H. 1995. The long-term relationship between zooplankton and fish IV.

626   Spatial/Temporal Variability and Prediction. ICES J. Mar. Sci. 52(3-4): 611-626.

627   Delta Independent Science Board. 2015. Flows and Fishes in the Sacramento-San Joaquin Delta,

628   Research Needs in Support of Adaptive Management, Sacramento, CA. 37 pp. Available

629   online from: http://deltacouncil.ca.gov/docs/delta-isb-s-final-report-flows-and-fishes-

630   sacramento-san-joaquindelta-research-needs-support Retrieved on: 09/09/2017

27

631   Delta Independent Science Board. 2016. Improving Adaptive Management in the Sacramento-

632   San Joaquin Delta, Sacramento, CA. 58 pp. Retrieved from:

633   http://deltacouncil.ca.gov/docs/final-delta-isb-adaptive-management-review-report

634   Retrieved on: 09/09/2017

635   Feyrer, F., Nobriga, M.L. and Sommer, T.R. 2007. Multidecadal trends for three declining fish

636   species: habitat patterns and mechanisms in the San Francisco Estuary, California,

637   USA. Can. J. Fish. Aquat. Sci. 64(4): 723-734.

638   Good, T. P., Beechie, T. J., McElhany, P., McClure, M. M., & Ruckelshaus, M. H. 2007.

639   Recovery planning for Endangered Species Act-listed Pacific salmon: using science to

640   inform goals and strategies. Fisheries, 32, 426–440.

641   Gosselin, J. L., Zabel, R. W., Anderson, J. J., Faulkner, J. R., Baptista, A. M., & Sandford, B. P.

642   2017. Conservation planning for freshwater–marine carryover effects on Chinook salmon

643   survival. Ecol. Evol. 1–14. doi:10.1002/ece3.3663

644   Halpern, B.S., Frazier, M., Potapenko, J., Casey, K.S., Koenig, K., Longo, C., Lowndes, J.S.,

645   Rockwood, R.C., and others. 2015. Spatial and temporal changes in cumulative human

646   impacts on the world's ocean. Nat. Commun. 6.

647   Harrell Jr., F.E. 2015. Regression modeling strategies: with applications to linear models,

648   logistic and ordinal regression, and survival analysis. Springer.

649   Hjort, J. 1914. Fluctuations in the great fisheries of northern Europe viewed in the light of

650   biological research. Rapp. P.-V. Réun. Cons. Int. Explor. Mer. 20: 1–228.

651   Hilborn, R. 2016. Correlation and causation in fisheries and watershed

652   management. *Fisheries*, *41*(1), pp.18-25.

653   Hilborn, R. and M. Mangel. 1997. The ecological detective: confronting models with data.

654   Princeton University Press.

655   Hill, A. B. 1965. The environment and disease: association or causation? Proc. Roy. Soc. Med.

656   58:295–300.

657   Hunsicker, M.E., Kappel, C.V., Selkoe, K.A., Halpern, B.S., Scarborough, C., Mease, L. and

658   Amrhein, A. 2016. Characterizing driver–response relationships in marine pelagic

659   ecosystems for improved ocean management. Ecol. Appl. 26(3): 651-663.

660   Hutton, P.H., Rath, J.S., Chen, L., Ungs, M.J., Roy, S.B. 2015. Nine Decades of Salinity

661   Observations in the San Francisco Bay and Delta: Modeling and Trend Evaluations. J.

662   Water. Res. Plan. Man., 142(3):04015069.

663   Ianelli, J.N., Honkalehto, T., Barbeaux, S., Kotwicki, S., Aydin, K. and Williamson, N. 2012.

664   Assessment of the Walleye Pollock Stock in the Eastern Bering Sea. Alaska Fisheries

665   Science Center, National Marine Fisheries Service, Seattle, WA.

666   Irvine, J.R., O'Neill, M., Godbout, L. and Schnute, J. 2013. Effects of smolt release timing and

667   size on the survival of hatchery-origin coho salmon in the Strait of Georgia. Prog. Oceanog.

668   *115*, pp.111-118.

669   Jacobson, L.D., and MacCall, A.D. 1995. Stock-recruitment models for Pacific sardine

670   (*Sardinops sagax*). Can. J. Fish. Aquat. Sci. 52(3), pp.566-577.

671   Jassby AD, Kimmerer WJ, Monismith SG, Armor C, Cloern JE, Powell TM, Schubel JR, and

672   Vendlinski TJ. 1995. Isohaline position as a habitat indicator for estuarine populations. Ecol.

673   Appl. 5(1):272-89.

674   Kimmerer, W.J. 2002. Effects of freshwater flow on abundance of estuarine organisms: physical

675   effects or trophic linkages? Mar. Ecol. Prog. Ser. 243: 39-55.

676   Kimmerer, W.J. 2004. Open water processes of the San Francisco Estuary: from physical forcing

677   to biological responses. San Francisco Estuar. Watershed Sci. 2(1): 1-140.

678   Kimmerer, W.J., Gross, E.S., and MacWilliams, M.L., 2009. Is the response of estuarine nekton

679   to freshwater flow in the San Francisco Estuary explained by variation in habitat

680   volume? Estuar. Coast. 32(2): 375-389.

681   Latour, R.J. 2016. Explaining patterns of pelagic fish abundance in the Sacramento-San Joaquin

682   Delta. Estuar. Coast. 39(1): 233-247.

683   Levin, P.S., Zabel, R.W., and Williams, J.G. 2001. The road to extinction is paved with good

684   intentions: negative association of fish hatcheries with threatened salmon. P. Roy. Soc.

685   Lond. B. Bio. 268(1472): 1153-1158.

686   Mahardja, B., Young, M.J., Schreier, B. and Sommer, T. 2017. Understanding imperfect

687   detection in a San Francisco Estuary long-term larval and juvenile fish monitoring

688   programme. Fisheries Manage. Ecol. 24(6): 488-503.

689   Maunder, M.N. and Deriso, R.B. 2011. A state–space multistage life cycle model to evaluate

690   population impacts in the presence of density dependence: illustrated with application to

691   delta smelt (*Hyposmesus transpacificus*). Can. J. Fish. Aquat. Sci. 68(7): 1285-1306.

692   Mislan, K.A.S., Heer, J.M. and White, E.P. 2016. Elevating the status of code in ecology. Trends

693   Ecol. Evol. 31(1): 4-7.

694   Miller, W.J., Manly, B.F., Murphy, D.D., Fullerton, D. and Ramey, R.R. 2012. An investigation

695   of factors affecting the decline of delta smelt (Hypomesus transpacificus) in the

696   Sacramento-San Joaquin Estuary. Rev. Fish. Sci. 20(1): 1-19.

697   Möllmann C, Diekmann R. 2012. Marine ecosystem regime shifts induced by climate and

698   overfishing: a review for the Northern Hemisphere. Adv. Ecol. Res. 47, 303–347.

699   Myers, R. A. 1998. When do environment–recruitment correlations work? Rev. Fish Biol.

700   Fisher. 8: 285–305.

701   National Research Council (NRC). 2012. Sustainable water and environmental management in

702   the California Bay-Delta. National Academies Press.

703   Nichols, F.H., Thompson, J.K. and Schemel, L.E. 1990. Remarkable invasion of San Francisco

704   Bay (California, USA) by the Asian clam *Potamocorbula amurensis*. II. Displacement of a

705   former community. Mar. Ecol. Prog. Ser. 66: 95-101.

706   Nobriga, M.L. and Rosenfield, J.A. 2016. Population dynamics of an estuarine forage fish:

707   disaggregating forces driving long-term decline of Longfin Smelt in California's San

708   Francisco Estuary. T. Am Fish. Soc. 145(1): 44-58.

709   Nosek, B.A., Alter, G., Banks, G.C., Borsboom, D., Bowman, S.D., Breckler, S.J., Buck, S.,

710   Chambers, C.D., and others. 2015. Promoting an open research culture. Science. 348(6242):

711   1422-1425.

712   Peterman, R.M. and Anderson, J.L. 1999. Decision analysis: a method for taking uncertainties

713   into account in risk-based decision making. Hum. Ecol. Risk. Asst. 5(2): 231-244.

714   Peterman, R.M. 2004. Possible solutions to some challenges facing fisheries scientists and

715   managers. ICES J. Mar. Sci. 61:1331-1343.

716   Peters, C.N. and Marmorek, D.R. 2001. Application of decision analysis to evaluate recovery

717   actions for threatened Snake River spring and summer Chinook salmon (*Oncorhynchus*

718   *tshawytscha*). Can. J. Fish. Aquat. Sci. 58(12): 2431-2446.

719   Plowright, R. K., S. H. Sokolow, M. E. Gorman, P. Daszak, and J. E. Foley. 2008. Causal

720   inference in disease ecology: investigating ecological drivers of disease emergence. Front.

721   Ecol. Envir. 6:420–429.

722   Power, M. 1993. The predictive validation of ecological and environmental models. Ecol.

723   Model. 68(1-2): 33-50.

724   Punt, A.E., Butterworth, D.S., Moor, C.L., De Oliveira, J.A. and Haddon, M. 2016. Management

725   strategy evaluation: best practices. Fish Fish. 17(2): 303-334.

726   R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for

727   Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

728   Ruckelshaus, M. H., Levin, P., Johnson, J. B., & Kareiva, P. 2002. The Pacific salmon wars:

729   what science brings to the challenge of recovering species. Annu. Rev. Ecol. Syst. 33: 665–

730   706.

731   Sainsbury K.J., A.E. Punt, A.D.M. Smith. 2000. Design of operational management strategies

732   for achieving fishery ecosystem objectives. ICES J. Mar. Sci. 57: 731–741.

733   Scheuerell, M.D., Hilborn, R., Ruckelshaus, M.H., Bartz, K.K., Lagueux, K.M., Haas, A.D. and

734   Rawson, K. 2006. The Shiraz model: a tool for incorporating anthropogenic effects and

735   fish–habitat relationships in conservation planning. Can. J. Fish. Aquat. Sci. *63*(7): 1596-

736   1607.

737   Scheffer, M., S. Carpenter, J. A. Foley, C. Folke, and B. Walker. 2001. Catastrophic regime

738   shifts in ecosystems. Nature. 413:591–596.

739   Scheffer, M., J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H.

740   van Nes, M. and others. 2009. Early-warning signals for critical transitions. Nature. 461:

741   53–59.

742   Stevens, D.E. 1977. Striped bass (*Morone saxatilis*) year class strength in relation to river flow

743   in the Sacramento–San Joaquin Estuary, California. T. Am. Fish. Soc. 106: 34–42.

744   Sommer, T., R. Baxter, and B. Herbold. 1997. Resilience of splittail in the Sacramento–San

745   Joaquin Estuary. T. Am. Fish. Soc. 126: 961–976.

746   Sommer, T., Mejia, F., Hieb, K., Baxter, R., Loboschefsky, E. and Loge, F., 2011. Long-term

747   shifts in the lateral distribution of age-0 striped bass in the San Francisco Estuary. T. Am.

748   Fish. Soc. 140(6): 1451-1459.

749   Szuwalski, C. S., K. A. Vert-pre, A. E. Punt, T. A. Branch, and R. Hilborn. 2015. Examining

750   common assumptions about recruitment: a meta-analysis of recruitment dynamics for

751   worldwide marine fisheries. Fish and Fish. 16(4): 633-648.

752   Thomson, J.R., Kimmerer, W.J., Brown, L.R., Newman, K.B., Nally, R.M., Bennett, W.A.,

753   Feyrer, F. and Fleishman, E. 2010. Bayesian change point analysis of abundance trends for

754   pelagic fishes in the upper San Francisco Estuary. Ecol. Appl. 20(5): 1431-1448.

755   Walters, C.J. 1986. Adaptive Management of Renewable Resources. MacMillan Press. 374 pp.

756   Zandvoort, M., van der Vlist, M.J. and van den Brink, A. 2017. Handling uncertainty through

757   adaptiveness in planning approaches: comparing adaptive delta management and the water

758   diplomacy framework. J. Environ. Pol. Plan. doi: 10.1080/1523908X.2017.1347035

759   Zeug, S.C., Bergman, P.S., Cavallo, B.J. and Jones, K.S. 2012. Application of a life cycle

760   simulation model to evaluate impacts of water management and conservation actions on an

761   endangered population of Chinook salmon. *Envir. Model. Assess.* 17(5): 455-467.

762

Tables

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	Log total salvage	1979 - 2002 / 2016
Brandes et al. 2006	Fall Chinook salmon	Chinook salmon fry / m^3 at Chipps Island (Jan–Mar)	Mean daily February flow at Freeport (cfs)	1985 - 2005 / 2011
	Fall Chinook salmon	Chinook salmon fry / m^3 at Sacramento (Apr–Jun)	Mean daily February flow at Freeport (cfs)	1985 - 2005 / 2011
	Fall Chinook salmon	Chinook salmon fry / m^3 at Chipps Island (Apr–Jun)	Mean daily April to June flow at Rio Vista (cfs)	1978 - 2005 / 2015
Kimmerer et al. 2009	American shad	FMWT Abundance Index	Spring X2 (km)	1988 - 2007 / 2014
	Bay goby	Bay Study OT Abundance Index	Spring X2 (km)	1988 - 2007 / 2014
	Bay shrimp	Bay Study OT Abundance Index	Spring X2 (km)	1988 - 2007 / 2012
	Longfin smelt	Abundance Index (3 models: FMWT; Bay Study MWT,OT)	Spring X2 (km)	1988 - 2007 / 2014
	Sacramento splittail	Abundance Index (FWMT)	Spring X2 (km)	1988 - 2007 / 2014
	Starry flounder	Abundance Index (Bay Study OT)	Spring X2 (km) (1 year lag)	1988 - 2007 / 2013
Sommer et al. 1997	Sacramento splittail	Abundance Index (FWMT)	Delta outflow (cfs)	1967 - 1995 / 2014
	Sacramento splittail	Abundance Index (FWMT)	Days Yolo Bypass inundated	1967 - 1995 / 2014
Sommer et al. 2007	Longfin smelt	Abundance Index (FWMT)	Flows Jan to June (cfs)	1988 - 2006 / 2014
Stevens 1977a	Striped bass	Abundance Index (TNS 38 mm Index)	Mean daily Delta outflow (June-July) (cfs)	1959 - 1970 / 2016
	Striped bass	Abundance Index (TNS Suisun Bay)	Mean daily Delta outflow (June-July) (cfs)	1959 - 1976 / 2011
Feyrer et al. 2007	Delta smelt	Abundance Index (TNS)	Mean Annual Conductance (µS/cm) (FMWT) Fall Stock Abundance Index (FMWT)	1987 - 2004 / 2016
	Delta smelt	Abundance Index (TNS)	Mean Annual Secchi Depth (m) (FMWT) Fall Stock Abundance Index (FMWT)	1987 - 2004 / 2016
	Delta smelt	Abundance Index (TNS)	Mean Annual Conductance (µS/cm) (FMWT) Mean Annual Secchi Depth (m) (FMWT) Fall Stock Abundance Index (FMWT)	1987 - 2004 / 2016
Miller et al. 2012	Delta smelt	Fall-to-Summer Survival (July TNS Index / previous year's FMWT Index)	Previous Fall Abundance (FWMT) Previous-Previous Fall Abundance (FWMT) Average E-P density, Apr–Jun (#/m^3) Proportional Entrainment Average E-P density, Jan–Mar (#/m^3)	1972 - 2006 / 2014
	Delta smelt	Summer-to-Fall Survival (FMWT Index / July TNS abundance in the same year)	July STN Abundance /1,000 Average E-P density, Sep –Dec (#/m^3)	1972 - 2006 / 2014
	Delta smelt	Fall-to-Fall Survival (FMWT / previous year's FMWT index)"	Previous FMWT Fall Abundance Previous-Previous FMWT Fall Abundance Average E-P density, Apr–Jun (#/m^3) Average E-P density, Sep –Dec (#/m^3)	1972 - 2006 / 2014

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.
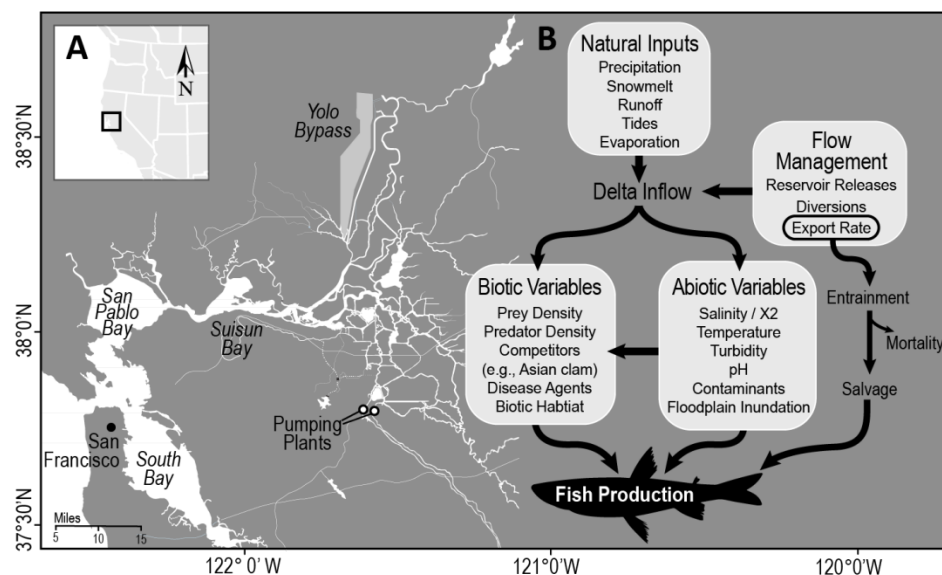
Figures

Figure 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 the Delta Independent Science Board 2015.
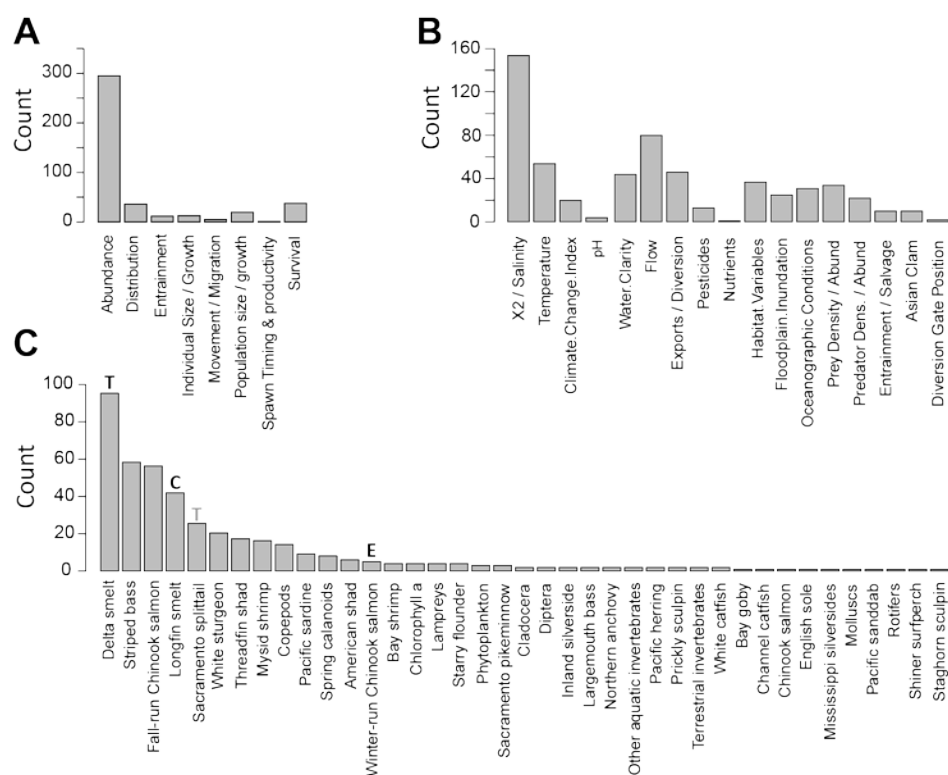
Figure 2: Frequency distributions of the type of response variable (A), environmental variables (B) and focal species (C), represented in the 420 individual environment-recruitment relationships for the Bay Delta region examined in this study, showing the relative emphasis of this body of work. Superscripts in panel C indicate the species current (black) or historical (grey) conservation status as Endangered (E), Threatened (T), or Candidate (C) as reported by the U.S., Fish and Wildlife Service Environmental Conservation Online System.
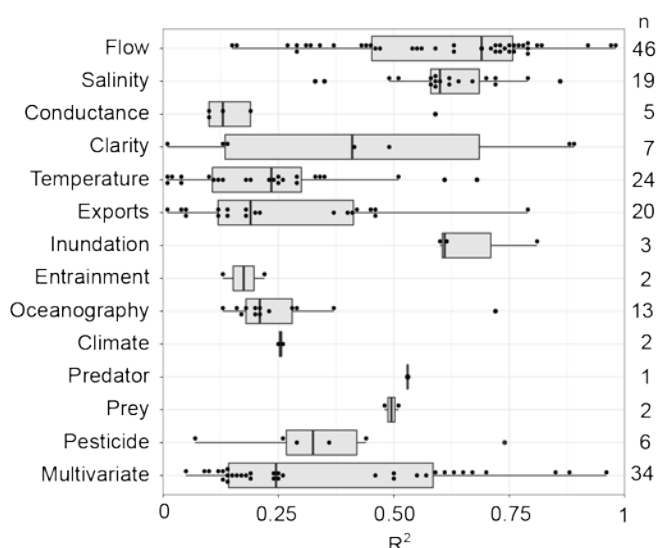
Figure 3: Boxplots and underlying estimates of the distribution of the strength (R^2) of published relationships for each broad type of environmental variable identified in the literature review (184 papers reported this statistic). Some variables have stronger relationships (e.g., flow, salinity) with the abundance of species than others that are more variable (e.g., temperature, 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.
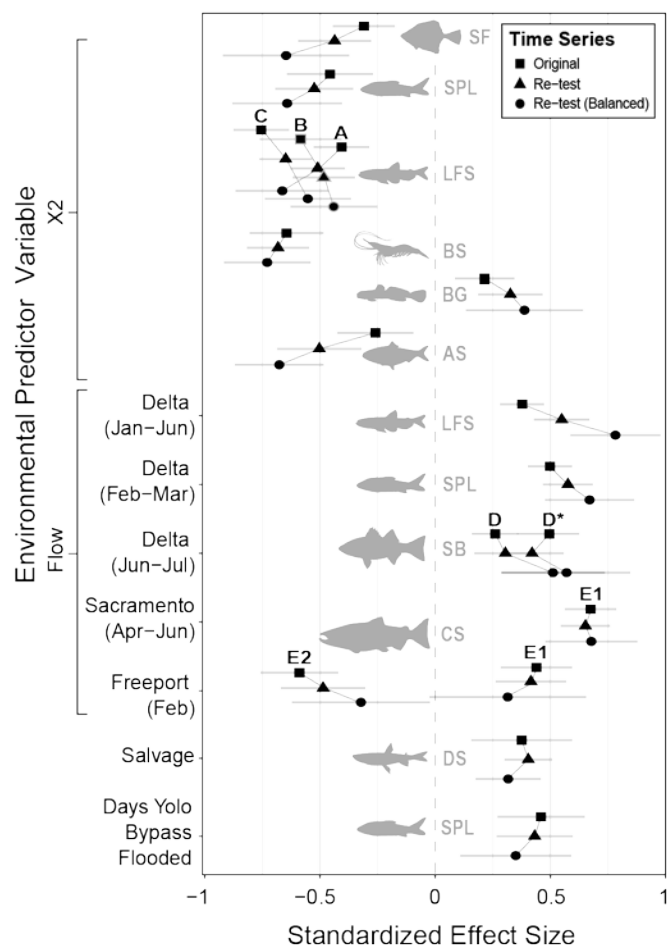
Figure 4: Magnitude of environment-recruitment relationships for 17 univariate analyses spanning nine species based on original (■), extended (▲), and extended but balanced (●) time series. "Balanced" refers to estimates based on analyses where we randomly sub-sampled (1000 times) the pre-retest data 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.
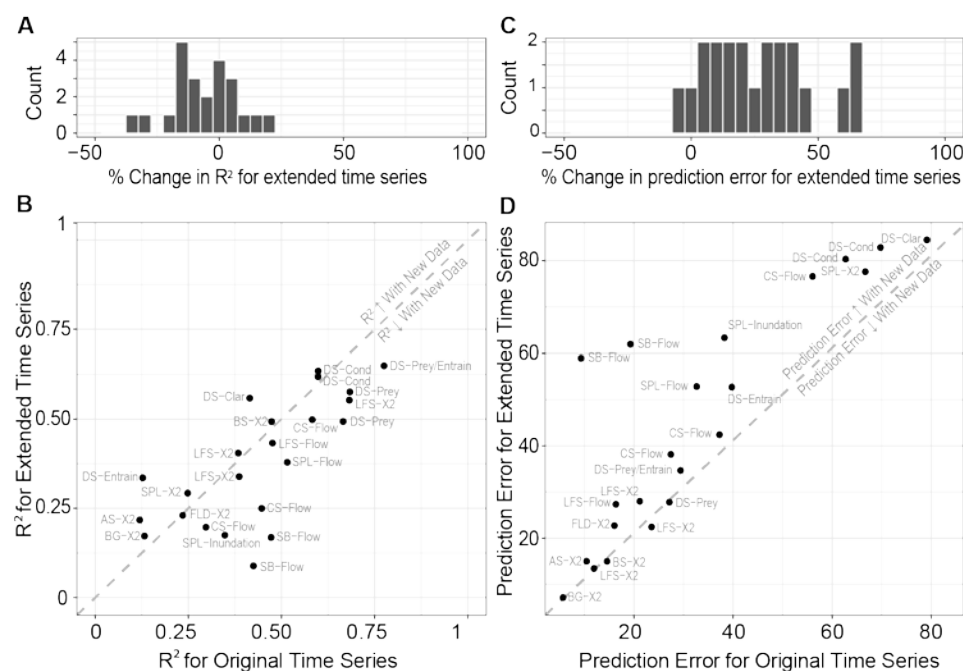
38

Figure 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 due to their very large values (~600%; see Table 1). Prediction error (see equation 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 +/- 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.
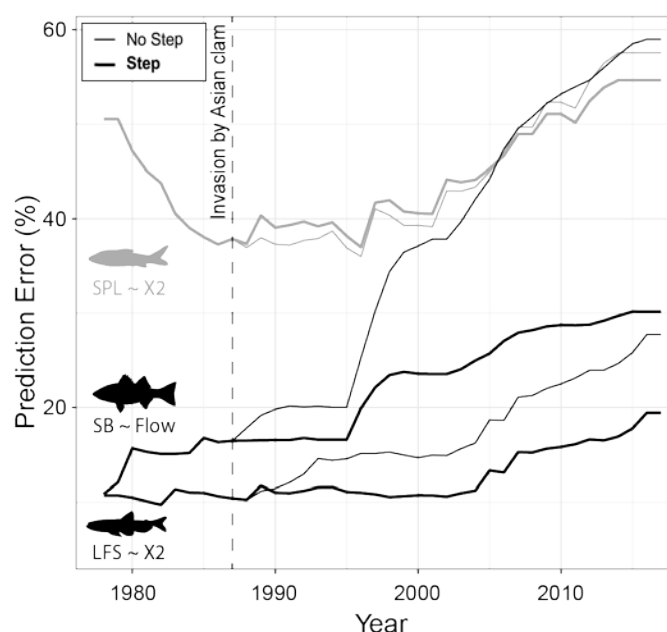
Figure 6: Prediction error (equation 1) relative to the logged index values used in the models over time for three illustrative Bay Delta environment-recruitment relationships modelled without (thin lines) and with (heavy lines) a step change to account for the regime shift associated with the introduction 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 (grey: SPL – Sacramento splittail) do not.