SCIENTIFIC AND STATISTICAL COMMITTEE REPORT ON 2021-2022 CALIFORNIA CURRENT ECOSYSTEM STATUS REPORT AND SCIENCE REVIEW TOPICS

The Scientific and Statistical Committee (SSC) met with representatives of the California Current Integrated Ecosystem Assessment (CCIEA) team, Drs. Toby Garfield (Southwest Fisheries Science Center) and Chris Harvey (Northwest Fisheries Science Center). The SSC's discussion with the CCIEA team encompassed three topics, which are reported upon below in turn: 1) the 2022 California Current Ecosystem Status Report (CCIEA Team Report 1), 2) the report of the August 31 - September 1 2021 SSC Ecosystem Subcommittee (SSCES) meeting (appended to the end of this statement), and 3) ecosystem science review topics proposed for 2022 (CCIEA Team Supplemental Report 2).

Review of the 2022 CCIEA Ecosystem Status Report

The Ecosystem Status Report provides important information on environmental, biological, social, and economic indicators and provides an ecosystem perspective on West Coast fish stocks, fisheries, and coastal communities for the Council process. The SSC commends the CCIEA team's openness and responsiveness to Council and SSC questions and recommendations, and their continuing efforts to improve the Status Report each year. Significant additions to the report this year include an indicator of krill biomass off northern California potentially relevant to Klamath River Fall Chinook, quantitative marine survival projections in association with some salmon indicators, information from acoustic trawl surveys for Coastal Pelagic Species, albacore diet information, analyses of overlap between wind energy areas and non-confidential limited-entry groundfish bottom trawl fishing activity, expansion of the fishery participation network analyses, and a climate change appendix.

Recent changes in distribution and abundance of species and in fisheries have been accompanied by rare combinations of extremes that make it difficult to identify drivers of observed changes, a challenge that is likely to be further intensified by ongoing climate change. Further development of the climate change appendix would be an important step in attempting to address this challenge.

An overarching theme in this year's Ecosystem Status Report is that oceanic indicators largely returned to states similar to generally favorable pre-2013 conditions, aside from a marine heatwave that largely remained far offshore, while conditions in freshwater were characterized by drought, record heat, and reduced snowpack and flows.

The SSC discussed several issues that could affect the interpretation of the indicators in the report including:

1. Natural-area Sacramento River Fall Chinook (SRFC) escapement in 2019 (giving rise to the dominant age-3 age class for 2022 fisheries) is described as "relatively favorable" for natural-area Central Valley Fall Chinook and noted as having "met goal" in Table 3.3.2. However, while natural-area escapement in 2019 was higher than other years under consideration, there is currently no natural-area escapement goal established for SRFC, and

multiple studies indicate that natural production would be maximized at substantially higher escapements than the typical natural-origin fraction of the current combined goal.
2. The boundaries between colors in stoplight charts are based on ranks and are sensitive to the time period used for reference, and subject to change over time.
3. Increased incorporation of predator-prey considerations into salmon indicators could be warranted.
4. A "sawtooth" pattern of strong upwelling followed by relaxation events is apparent in 2021, which is thought to be characteristic of good conditions for productivity. Further quantifying this pattern through the development of indicators that capture the variation in upwelling over relevant timescales and link to productivity could strengthen the biological relevance of upwelling indicators to managed species.

Ecosystem Science Review Report from 2021

The SSC reviewed the SSC Ecosystem Subcommittee (SSCES) report from its meeting held via webinar on August 31 and September 1, 2021 and discussed the report with SSCES Chair Dr. Kristin Marshall (NWFSC). The SSC agrees with the SSCES recommendations that:

1. Future CCIEA reports should identify times when environmental conditions are beyond thresholds associated with poor salmon forecast performance in the past;
2. A size-based krill indicator would be a useful addition to the Klamath River Fall Chinook stoplight table and a biomass indicator would be a helpful addition (these were included in the 2022 report);
3. It would be useful to include port-level network analyses in future CCIEA reports (some analyses were added to the 2022 report); and
4. It would be helpful to see more cross references between the CCIEA report and other Council materials, for example mention of the CCIEA report in salmon reports and vice versa.
5. The SSC also agreed with the SSCES recommendations to develop robust juvenile groundfish abundance indices based on spatially-explicit information on size and abundance from the West Coast Bottom Trawl Survey to inform management between assessments, such as through assessment prioritization scoring or scientific uncertainty buffers. These analyses should focus on species that are well-sampled by the survey and not associated with rocky habitats.

The SSCES reviewed six documents, five of which were journal articles that are cited in the SSCES report and publicly available (but in some cases behind paywalls) and one unpublished addendum to an earlier article produced specifically for the SSCES meeting. The addendum is attached at the end of the SSCES report. In the future, the SSC recommends establishing a process for posting items (or at least links for copyrighted material) provided for SSCES review to the briefing book under an appropriate agenda item to facilitate transparency and public access to materials reviewed by the SSC. A single, easily found repository of all previous SSC Ecosystem Subcommittee reports and a consistent process for distributing the Subcommittee reports would also be useful and increase transparency.

Proposed Ecosystem Science Review Topics for 2022

The CCIEA team has proposed three potential topics for review in September 2022 (Supplemental IEA Team Report 2): 1) strategic review of the salmon indicator portfolio, 2) reference periods for plotting recent means and trends in fishery landings and revenues, and 3) development of the climate change appendix.

The SSC and CCIEA team agreed that point 2 could be addressed without the need for a meeting, with input from relevant advisory subpanels, and then the end product could be reviewed during the March meeting review of the first Ecosystem Status Report to incorporate it. The SSC recommends scheduling a half day for each of the two remaining topics in a September 2022 meeting of the SSCES, noting that both are complicated and ambitious topics that will likely require multiple meetings with additional advisory bodies to fully address. The salmon review could focus on identifying stocks where ecosystem information would be most useful, and how it could be used to better inform management, which could include data-poor stocks and/or stocks whose life histories are not amenable to conventional forecasting techniques. The climate change review could initially focus on technical discussions between the SSCES and the CCIEA team as well as identifying potential processes for involving additional advisory bodies. Given the focus on salmon, the SSC Salmon Subcommittee, Salmon Technical Team, and Salmon Advisory Subpanel should be invited to attend the SSCES meeting in September.

PFMC
03/10/22

SCIENTIFIC AND STATISTICAL COMMITTEE'S
ECOSYSTEM SUBCOMMITTEE REPORT

Pacific Fishery Management Council
Via Webinar

August 31 and September 1, 2021

The Scientific and Statistical Committee's Ecosystem Subcommittee (SSCES) met via webinar August 31 and September 1, 2021 to review new analyses conducted by the NMFS California Current Integrated Ecosystem Assessment (CCIEA) team that may potentially inform future annual Ecosystem Status Reports (hereafter CCIEA report) to the Pacific Council on the state of the California Current Ecosystem. The SSCES reviewed four topics: A) Threshold Relationships Between Environmental Drivers and Performance of Salmon Preseason Abundance Forecasts, B) Krill-based Indicators, C) Year Class Strength and Distribution of Small Groundfish, and D) Port-level Linkages Between Fisheries using Network Analysis. Dr. Kristin Marshall chaired the meeting. Meeting participants are listed in Appendix A.

A. Threshold Relationships Between Environmental Drivers and Performance of Salmon Preseason Abundance Forecasts

Dr. William Satterthwaite (NOAA, Southwest Fisheries Science Center) presented his paper "Ecological thresholds in forecast performance for key United States West Coast Chinook salmon stocks" (Satterthwaite et al. 2020) and an addendum he prepared for the SSCES, "Ecological thresholds in forecast performance for key United States West Coast Chinook salmon stocks – Addendum". This research evaluated whether the performance of Chinook salmon abundance forecasts are related to environmental conditions, focusing on non-linear threshold relationships. Non-linear relationships have potential to disrupt fisheries management and are not incorporated in current forecast models. Satterthwaite et al. (2020) focused on stocks of high priority for US west coast fisheries management and of predicted importance as prey for Southern Resident Killer Whales. The authors tested 2688 stock-driver-time lag combinations and found 65 non-linear relationships. Of these, 60 demonstrated threshold relationships, determined to exist when the 95% confidence interval of the second derivative of the nonlinear function excluded zero. Among indices capable of explaining at least 33% of variance in forecast performance, oceanic environmental indices were much more common than freshwater or local environmental indices. This may be because forecasts already make use of some measure of cohort strength (e.g., jack returns) that takes place after freshwater and ocean-entry conditions have had their immediate effects. There were mechanistic explanations for many of the observed relationships. When many of the relationships were re-examined with updated datasets (see Addendum), in almost all cases where non-linear relationships had been previously selected, they were re-selected. This work could help fisheries managers identify environmental thresholds past which increased precaution may be warranted. For example, as suggested in the Addendum, NPI could be added to the annual CCIEA report as extreme values of it appear to predict poor Sacramento River Fall Chinook forecast performance and it may be relevant for interpreting some Puget Sound abundance forecasts. Dr. Satterthwaite also demonstrated a straw-person method by which fisheries managers could quantify uncertainty in forecasts and increase precaution when a threshold is exceeded, using similar logic to how groundfish are managed.

While this was the first presentation on forecast thresholds to the SSCES, Dr. Satterthwaite addressed many comments the SSC and SSCES made in previous discussions of threshold approaches. The SSCES appreciated that feedback had been incorporated in many instances, and that acknowledgment was made where previous SSC recommendations were applicable but not yet incorporated. For example, a null model randomization procedure approach was used to look at the chance for false positives and Bonferroni corrections were used for p-values, as previously suggested by the SSC and others when large numbers of tests were conducted in screening for relationships. In the Addendum, Dr. Satterthwaite also examined how robust the identified relationships were to new data, as suggested by previous SSC reviews of threshold work. The relationships tended to remain non-linear, and R^2 tended to decrease with small increases in new data. When examining thresholds, one expects little change when “average” new data is added but more substantial changes when more extreme observations are added; this pattern was, in most cases, observed. Non-stationarity in the threshold relationships is not addressed in this work but is worth consideration in further work.

The SSCES discussed several other technical aspects of the work and makes the following suggestions:

- The SSCES agreed that R^2 is a useful metric, but should not be the sole metric to evaluate the utility of models because the performance variable is truncated and thus non-normal, but R^2 is from a normal model
- Multi-variable responses or multidimensional indices. Multi-variable responses are certainly possibilities, but difficult. More localized indicators with clear mechanistic hypotheses would be a good place to start such an investigation. Multi-dimensional indices could be useful, but only if the components are related to forecast performance with the same lag
- Another potential approach to explore is using a logit response; if a forecast breaks down with extreme values in either direction (a u-shaped rather than sigmoidal response) then logit might not capture that
- Results showing strong linear relationships should be investigated for inclusion in forecast models
- Consider exploring additional ways to quantify errors in forecasting because this approach is less likely to capture under-forecasts than over-forecasts and multiple metrics may be needed to fully capture the magnitude of error, proportion of error, and the consequences of errors to management.

The SSCES appreciated this innovative work and supports using the approach in the CCIEA report to characterize conditions when salmon forecasts may perform poorly. In previous reviews of threshold research, the SSC recommended that the CCIEA report include a small set of pressure variables where a threshold is indicated. The report currently includes recent PDO, and it may be useful to add a “now-cast” or a forecast, as well as include the NPI. These indices could aid in categorizing the risk associated with Sacramento River Fall Chinook and certain Puget Sound Chinook salmon forecasts. If an indicator is in a range that is a threshold for any fish stock, the SSCES notes that it is worth mentioning in the CCIEA report. At the same time, the SSCES recognizes that nuance is needed in describing errors in forecasting. An indicator being above a threshold does not imply that a forecast will be wrong, but it does mean that more caution might be warranted if the consequences of forecast error are undesirable and forecast error is more likely due to environmental conditions. In general, it would be helpful to see more cross references between the CCIEA report and other Council materials, for example mention of the CCIEA report in salmon statements and vice versa.

B. Krill-based Indicators

Dr. Eric Bjorkstedt and Ms. Roxanne Robertson (Humboldt State University) presented an overview of data and methods behind krill-based indicators entitled "Size of adult *Euphausia pacifica* along the Trinidad Head Line: an ecosystem indicator for the California Current." The review was suggested by the SSC in March of 2021 to better understand how the mean krill size data presented in the CCIEA report could be interpreted in the absence of relative abundance data, given the nuanced nature of interpreting size data alone with respect to population trends and abundance.

Dr. Bjorkstedt described the indicator as representing density-weighted mean body length of adult *Euphausia pacifica* captured in standard bongo net sampling along the Trinidad Head Line (THL), just north of Cape Mendocino, based on biweekly to bimonthly sampling of five standard stations that run along the continental shelf and slope (35 to 780m depth). The region is characterized as having considerable mesoscale variability in ocean conditions and advection patterns, and a key motivation for the location of this survey line was the hope that the resulting data could help inform regional productivity of Klamath river salmon stocks. The survey began in 2007 and is ongoing, details on survey methods and a great many additional survey results are reported in a publication (Robertson and Bjorkstedt 2020) that was also made available to the subcommittee. Data collected in this survey include hydrographic sampling (temperature, salinity), water sampling (nutrients, chlorophyll) and zooplankton sampling (krill, other zooplankton, ichthyoplankton).

Importantly, in this survey adult krill are identified by maturity factors rather than size thresholds, and the results of their analysis indicate that there would be considerable misclassification of adults and juveniles during warm periods if based on size alone. Adult krill are more abundant over the outer shelf and upper slope, although they are often found inshore, though at lower densities, during the upwelling season (and are often larger on such occasions). There are clear indications of shifts in the size distribution over time, for example, in 2008 krill catches were of generally larger individuals, while in 2015 (during the large marine heatwave) adult krill tended to be considerably smaller. While the authors estimate and have reported biomass indices in the literature, they also noted that numerical abundance (the number of individuals) does not change substantially over time, such that a considerable fraction of the change in total biomass is driven by changes in size. This suggests that changes in adult size represent an integrative index of krill in this region and reflect insights into both available biomass and how it is "packaged and distributed." The authors also note that they have not yet attempted to develop population models, or relate spawning biomass to recruitment, in order to better evaluate the consequences of smaller females to potential spawning output and productivity.

Considerable effort has focused on relating shifts in size distributions from this dataset to environmental conditions. Among the findings are that low frequency shifts in size distributions appear to reflect changes in upper water column ocean conditions, particularly temperature, with convergence towards median adult sizes at warm temperatures. Seasonal increases in average size of mature adults are reduced under warm conditions, and size increases with colder years and with higher chlorophyll levels. Conversely, early furcilia stages are larger during warm years. There is some suggestion that dynamics are preconditioned at some level, with population and size trajectories for spring and summer reflecting observed patterns during the winter. However, there is uncertainty regarding the extent to which shifts in size structure might result from advection rather than local dynamics, as there is evidence for advective drivers of some observations, such as a rapid and steep drop in mean size coincident with the arrival of "warm blob" waters at the coast in late 2014, which happened too fast to reflect localized population dynamics.

For Klamath salmon, it was noted that early ocean survival rates appear to have some general relationship to krill size, such that juvenile salmon rarely have high survivorship when krill are small as adults. The dataset also includes potential assemblage indicators, through the relative abundance of species with warm or cool water affinities.

In discussion, the SSCES asked about the spatial representativeness of the index, and the extent to which this indicator is localized or reflects larger scale trends. The proponents suggest that the index is likely to represent the region between Cape Blanco and Cape Mendocino, and thus could be a useful indicator for Klamath River salmon, but differences in oceanography make it uncertain whether the THL index to be a robust indicator of krill demographic or abundance trends reliably beyond this region. However, Dr. Bjorkstedt noted that earlier investigations found that the THL copepod time series (which is behind several years on data processing) correlated well with the Newport Hydrographic Line (NHL), several hundred km to the North, though with important differences in composition and within-season timing. The SSCES suggested that more comparisons among krill surveys could be helpful to get a sense of the scale of variability in krill across the California Current Ecosystem, and some surveys that occur less frequently over broader spatial areas could also inform this scale.

The subcommittee also discussed the extent to which mean adult size is the most appropriate indicator, or whether the addition of or shift to a biomass based indicator could be more appropriate or informative. The potential benefits of combining or adding biomass to length, or adding assemblage-based indicators was discussed, recognizing that the precise mix of indicators to report would depend on how the indicators would be used or intended to represent. The SSCES suggested greater development of both biomass and size indicators for future CCIEA reports. The potential for "growth products" (e.g., indicators of individual growth rates) was discussed, as were indicators related to shifts in the distribution of mass.

The subcommittee recognized all of these products as helpful indicators of key ecosystem processes in this region but was uncertain regarding just how to integrate the results into informing management in a useful manner. The potential for helping to inform early marine survival indicators for Klamath salmon was discussed, although it was noted that the current assessment model for salmon fisheries is based on sibling regressions, which reflect information obtained after fish have gone through the presumably more variable initial marine survival phase. However, it could be that an indicator could provide an extra year or more of lead time, which could be helpful given that Klamath River Fall Chinook are currently under a rebuilding plan. Additionally, forage indicators also reflect the conditions that 2, 3, and 4 year old fish are facing in the ocean, and thus krill (or krill predators) could still be affecting later maturation and mortality rates in Klamath salmon. Moreover, as river returns are observed with error, modeling approaches (such as state-space models) that forecast based on multiple indicators of cohort strength could be more robust than univariate approaches that ignore uncertainty. Finally, the SSCES suggested that the size-based indicator or other indicators could be useful in the Klamath River stoplight table.

C. Year Class Strength and Distribution of Small Groundfish

Dr. Nick Tolimieri (NOAA, Northwest Fisheries Science Center) presented an analysis of juvenile groundfish habitat and abundance proposed for inclusion in a future CCIEA report. A recent publication was the basis of the presentation (Tolimieri et al., 2021). The motivation for the work is to inform Essential Fish Habitat for juvenile groundfish and identify important nursery areas. This research could also potentially lead to an index of recruitment for some species.

The analysis used lengths and abundance for 13 species from the West Coast Bottom Trawl Survey. The survey ages a subsample of fish. To estimate age for the measured but unaged fish, length was converted to age using a fixed age-length key for each species. For some species, there were not enough individuals in the smallest age-class (age-0 or age-1), age classes were combined (grouped) for the analysis. In discussion, the SSCES suggested that for species with sufficient data, a year-specific age-length key would better account for variability in growth.

Abundance was standardized using the Vector Autoregressive Spatio-temporal (VAST) package, assuming a common intercept across years and spatial variation was explained by spatial and spatiotemporal autocorrelation. The SSCES suggests further investigation of the variance surfaces (in addition to abundance) to better understand how the assumption of a common intercept might be affecting the results. For example, a comparison could be done by fitting a temporal model without the spatial field. The SSCES caution against extrapolating into areas that have particularly high variance. Investigating alternative approaches to VAST (e.g., sdmTMB) may also allow for more flexibility in the fixed spatial field.

The resulting juvenile spatial distributions were qualitatively categorized as: distinct hotspots (dover sole, shortspine thornyhead, splitnose), distinct hotspots that were temporally variable (hake, darkblotched rockfish), large distinct areas of high juvenile abundance (arrowtooth flounder, English sole, sablefish), and limited latitudinal distributions but no obvious hotspots (Pacific grenadier, lingcod, longspine thornyhead, petrale sole). The SSCES agreed that these spatial distributions are a useful starting point for defining juvenile habitat groundfish habitat. Due to multiple distinct patterns, the SSCES recommends continuing to focus on species-specific distributions and cautions against combining species into a single juvenile groundfish distribution map.

Validation of the juvenile abundance indices was explored by comparing against the recruitment deviations from the stock assessment model for sablefish, arrowtooth flounder, lingcod, and hake. Only sablefish appeared to have strong agreement. However, the SSCES noted in discussion that there are many reasons the two indices may not align, including the structure and assumptions of the assessment model. Therefore, it should not be assumed that the assessment recruitment deviations represent a "true" recruitment index.

The SSCES was asked to provide guidance on additional species that could be investigated with this approach and offers the following suggestions:

- Choose species that are well-sampled by the survey. Flatfish are likely good candidates

- Consider using survey selectivity estimated in the assessment models to guide size cut-offs. Assessments typically do not use length at age at very small sizes because they are not well sampled by the trawl
- Avoid applying this method to species that are rock-associated, particularly with the VAST approach. These likely include widow rockfish, darkblotched rockfish, shortbelly rockfish, and possibly chilipepper rockfish.
- Prioritize species that are important to fisheries

The SSCES discussed with the CCIEA team how to include this analysis into future CCIEA reports. The SSCES suggests the analysts consider developing indices representing temporal and/or spatial stability. This would condense the distribution maps into annual anomalies in hotpots or area and distribution of juvenile habitat, for example. The SSCES suggests that a future application of this work could be to use robust juvenile abundance indices to inform management between assessments, such as through assessment prioritization scoring, scientific uncertainty buffers, or other approaches.

D. Port-level Linkages Between Fisheries using Network Analysis

Dr. Jameal Samhouri (NOAA, Northwest Fisheries Science Center) provided an overview of the network analysis approach that has been developed to describe West Coast port groups. An initial set of network diagrams was included in the 2021 CCIEA report. The methods have since been revised and additional work was done in response to feedback following a presentation to the SSCES in January 2021 and the SSC in March. In addition to a PowerPoint provided at the SSCES meeting, Dr. Samhouri provided the SSCES with two publications (Fuller et al. 2017; Fisher et al. 2021) that use similar methods.

Dr. Samhouri presented a number of different networks that were responsive to suggestions made by the SSCES in January including:

(1) vessel-level networks with scaling of nodes based on the median proportion of revenue a fishery contributes to vessels in that fishery, alternative minimum revenue thresholds for determining which vessels to include, and different methods of determining edge weights based on the amount and evenness of revenue, or the number of vessels, associated with each fishery pair;

(2) aggregate port-level and state-level networks with fisheries node inclusion determined by a minimum proportion of port or state revenue and node scaling based on relative total revenue; and

(3) time series of vessel-level network diagrams for two ports showing how networks have changed between 2004 and 2019.

Dr. Samhouri discussed work published in Fisher et al. 2021, illustrating how network characteristics of edge density, centrality and modularity influence the response of participants in a network to a shock. The example focused on HAB-related crab closures in California and suggests that fishers in denser networks are more likely to move to other fisheries while those in less dense networks are more likely to cease fishing. The analysis also shows that for centralized networks impacts vary depending on the centrality of the fishery subject to a shock.

The SSCES appreciates the responsiveness of the analysts to its comments and suggestions and finds the new analyses and network diagrams useful. The networks provide a visual description of the fisheries/species groups of importance to particular port groups and the degree to which they are connected by cross-participation and movement of fishers between them. Fisheries are defined by the same species groupings used in the diversification indices in the annual CCIEA report (rather than by métier as was done in the earlier work by Fuller et al. 2017). The network diagrams complement the diversification indices by providing information about the characteristics of fishery diversification strategies and how they vary across ports.

The network analysis has the potential to contribute to our understanding of how shocks to fisheries may impact particular communities (defined by port group) and potentially reverberate across fisheries. This may be apparent to some degree from simply viewing the network diagram, but quantitative network metrics may provide additional insight into overall stability of networks, and potentially resilience or vulnerability of fishers in a port to shocks to fisheries. These metrics include edge density, centrality, and modularity. Of these, edge density appears to have the clearest relationship to resilience. Networks with high edge density suggest that fishers have greater ability to move effort between fisheries and thus substitute for lost revenue from a fishery that is closed or has a poor year. The effects of centrality and modularity of networks appears to be very context dependent. For example, if the central fishery is closed in a network with high centrality, the impact would be great while it would be small if a non-central fishery was closed. Networks with high modularity would have increased impacts within a module but less outside it. More analysis will be needed to get a better general sense of how and when centrality and modularity mediate impacts of fishery shocks and affect the resilience of fishing communities.

There was some confusion about the scaling of the nodes in the network diagrams that was clarified after the meeting. The scaling of nodes for the vessel level network is based on the median percent of individual vessels' revenue that the fishery contributed to the fishers that participated in it. The node is large if the fishery provides a large proportion of individuals' revenue to at least half of the fishers involved in that fishery. Even a fishery that contributes a relatively small share of revenue at the port level might be shown as a large node. For example, in the 2019 crab year (Nov 2019-Oct 2020) tuna in Astoria only contributed about 2% of total revenue as compared to 15% for non-DTS groundfish but it had a node similar in size to crab which contributed 33% while non-DTS groundfish had a small node. This approach to node scaling has the advantage of showing relative importance of each fishery to those who fish in it, but it does not necessarily reflect the overall importance of the fishery to the port. If this approach to scaling nodes is used, it needs to be clearly explained, or it may lead to confusion. It would be useful to provide some supplementary information about port level revenue such as a pie chart showing the proportion each fishery contributes to port revenue. In contrast, for the aggregate port or state networks, both fishery inclusion and node scaling are based on the proportion of revenue the fishery contributes to the port or state's total revenue. This approach highlights fisheries that contribute a large proportion of total revenue yet it may exclude fisheries that are very important to a subset of fishers. Both approaches have strengths and weaknesses and the SSCES sees merit in both. Whichever approach is used, the methods used for fishery inclusion, scaling of nodes, and defining edge weights should be clearly explained.

The SSCES is supportive of including port-level network analyses in future CCIEA reports. The following observations and comments arose in discussion and may be helpful to the analysts in preparing future network analyses:

- It should be made clear in any publications and presentations that the analysis reflects revenue by "crab years" (Nov-Oct) as opposed to calendar years for all fisheries.
- Node size and edge weights are comparable within ports but not across ports. While Dr. Samhouri noted that this could be changed to allow comparison across ports, it could be problematic to do so given large differences in absolute revenue and fleet sizes for different fisheries in different ports.
- In contrast to edge weights based only on the number of vessels in fishery pairs, revenue connectivity edges have edge weights that are higher when revenue is higher but also more evenly distributed between the nodes. This may provide more insight into what will happen when a shock happens to one or the other (e.g., more impacts are likely if revenue is more evenly distributed than if one node dominates). While more complex than edges based on the number of vessels, this may be more useful for understanding impacts of shocks. The analysis of network metrics (modularity and centrality) has been based on the revenue connectivity definition and may be less applicable when edges are based on vessel numbers only.
- For the aggregate port level diagrams Dr. Samhouri showed on slide 19, the scaling of nodes was based on the ratio of port revenue for that fishery relative to the revenue from the fishery with the highest revenue for that port. It was discovered after the meeting that there was an error in the diagram for Fort Bragg caused by one tuna fish ticket that had a misplaced decimal point. Tuna should not have had a large node in that diagram and other fisheries should have been included.
- For aggregate level networks, the 10% of total port revenue cut-off results in very few fisheries for some ports. An alternative might be a cut-off based on absolute revenue (e.g., over $100K) or a smaller percent of revenue. Supplementary diagrams at the end of the PowerPoint showed aggregate networks including fisheries that includes at least 5% of revenue which substantially increased the numbers of fisheries included. This lower cut-off might be preferable for aggregate networks.
- For Washington fish tickets reported Port may mean different things for groundfish, salmon and shellfish and this should be checked.
- It was suggested that it would be worth considering the vulnerability of the species themselves and tying that to the vulnerability of the networks (e.g., in a network with mostly species impacted by upwelling will be more vulnerable than one that has species that are not impacted by upwelling.
- It was suggested that it might be useful to go back before 2004 for time series analysis and to combine groups of years and look at changes over longer time periods or networks.
- Most of the SSCES members that commented found vessel level analysis more useful than the aggregate port-level analysis. The aggregate networks did not provide substantial information that could not be provided with a simple bar chart of share of revenue by fishery for each port. However, the SSCES assumed at the time that node scaling for vessel level networks already reflected the relative proportion of port revenue, which it did not.
- It was suggested that a network analysis could provide insight on community impacts when developing a groundfish rebuilding plan that largely affects a portion of the fishery. Doing

so might require different exclusion criteria to focus the network on the groundfish fishery similar to the approach used by Fisher et al. (2021) for crab.

References

Fisher, Mary C., Stephanie K. Moore, Sunny L. Jardine, James R. Watson, and Jameal F. Samhouri. Climate shock effects and mediation in fisheries. *Proceedings of the National Academy of Sciences* 118, no. 2 (2021). https://doi.org/10.1073/pnas.2014379117

Fuller, E.C., Samhouri, J.F., Stoll, J.S., Levin, S.A. and Watson, J.R., 2017. Characterizing fisheries connectivity in marine social–ecological systems. *ICES Journal of Marine Science*, 74(8), pp.2087-2096. https://doi.org/10.1093/icesjms/fsx128

Robertson, R.R. and Bjorkstedt, E.P., 2020. Climate-driven variability in *Euphausia pacifica* size distributions off northern California. Progress in Oceanography, 188, p.102412. https://doi.org/10.1016/j.pocean.2020.102412

Satterthwaite, W.H., Andrews, K.S., Burke, B.J., Gosselin, J.L., Greene, C.M., Harvey, C.J., Munsch, S.H., O'Farrell, M.R., Samhouri, J.F. and Sobocinski, K.L., 2020. Ecological thresholds in forecast performance for key United States West Coast Chinook salmon stocks. ICES Journal of Marine Science, 77(4), pp.1503-1515. https://doi.org/10.1093/icesjms/fsz189

Tolimieri, N., Wallace, J. and Haltuch, M., 2020. Spatio-temporal patterns in juvenile habitat for 13 groundfishes in the California Current Ecosystem. PloS one, 15(8), p.e0237996. https://doi.org/10.1371/journal.pone.0237996

Appendix A. Meeting Participants
SSC Ecosystem Subcommittee Members Present

Dr. Kristin Marshall (Subcommittee Chair), National Marine Fisheries Service Northwest Fisheries Science Center, Seattle, WA
Dr. John Field, National Marine Fisheries Service Southwest Fisheries Science Center, Santa Cruz, CA
Dr. Marisol Garcia-Reyes, Farallon Institute, Petaluma, CA
Dr. Michael Harte, Oregon State University, Corvallis, OR
Dr. Dan Holland, National Marine Fisheries Service Northwest Fisheries Science Center, Seattle, WA
Dr. Galen Johnson, SSC Chair, Northwest Indian Fisheries Commission, Olympia, WA
Dr. André Punt, University of Washington, Seattle, WA
Dr. William Satterthwaite, National Marine Fisheries Service Southwest Fisheries Science Center, Santa Cruz, CA
Dr. Ole Shelton, National Marine Fisheries Service Northwest Fisheries Science Center, Seattle, WA
Dr. Cameron Speir, National Marine Fisheries Service Southwest Fisheries Science Center, Santa Cruz, CA

CCIEA Team Members Present

Dr. Eric Bjorkstedt, National Marine Fisheries Service Southwest Fisheries Science Center, La Jolla, CA
Dr. Toby Garfield, National Marine Fisheries Service Southwest Fisheries Science Center, La Jolla, CA
Dr. Chris Harvey, National Marine Fisheries Service Northwest Fisheries Science Center, Seattle, WA
Ms. Roxanne Robertson, Humboldt State University, Arcata, CA
Dr. Jameal Samhouri, National Marine Fisheries Service Northwest Fisheries Science Center, Seattle, WA
Dr. Jarrod Santora, National Marine Fisheries Service Southwest Fisheries Science Center, La Jolla, CA
Dr. Nick Tolimieri, National Marine Fisheries Service Northwest Fisheries Science Center, Seattle, WA

Others Present

Mr. Kelly Andrews, National Marine Fisheries Service Northwest Fisheries Science Center, Seattle, WA
Ms. Marlene Bellman, Northwest Indian Fisheries Commission, Olympia, WA
Ms. Anna Bolm, Oregon State University, Corvallis, OR
Mr. Alan Byrne, Idaho Department of Fish and Game, Boise, ID
Mr. Jon Carey, National Marine Fisheries Service West Coast Region, Lacey, WA
Ms. Susan Chambers, West Coast Seafood Processors Association, Charleston, OR
Dr. Kit Dahl, Pacific Fishery Management Council, Portland, OR

Mr. John DeVore, Pacific Fishery Management Council, Portland, OR
Dr. Jeff Dorman, Farallon Institute, Petaluma, CA
Dr. Michael Drexler, Ocean Conservancy, St. Petersburg, FL
Ms. Robin Ehlke, Pacific Fishery Management Council, Portland, OR
Ms. Jennifer Fisher, National Marine Fisheries Service Northwest Fisheries Science Center, Newport, OR
Mr. Craig Foster, Oregon Department of Fish and Wildlife, Clackamas, OR
Dr. Tommy Garrison, Columbia River Inter-Tribal Fish Commission, Portland, OR
Ms. Grace Ghrist, California Department of Fish and Wildlife, Santa Rosa, CA
Dr. Owen Hamel, National Marine Fisheries Service Northwest Fisheries Science Center, Seattle, WA
Ms. Ashton Harp, Northwest Indian Fisheries Commission, Forks, WA
Mr. Brian Hoffman, Hoh Indian Tribe, Port Angeles, WA
Dr. Diego Holmgren, Tulalip Tribe, Everett, WA
Ms. Helen Killeen, University of California, Davis, CA
Ms. Gway Kirchner, The Nature Conservancy, Newport, OR
Mr. Hap Leon, Makah Tribe, Neah Bay, WA
Dr. Laura Lilly, Scripps Institution of Oceanography, La Jolla, CA
Mr. Pete McHugh, California Department of Fish and Wildlife, Santa Rosa, CA
Dr. Oliver Miler, Northwest Indian Fisheries Commission, Olympia, WA
Ms. Rebecca Miller, National Marine Fisheries Service Southwest Fisheries Science Center, Santa Cruz, CA
Dr. Tommy Moore, Northwest Indian Fisheries Commission, Olympia, WA
Ms. Kandice Morgenstern, California Department of Fish and Wildlife, Santa Rosa, CA
Mr. Corey Niles, Washington Department of Fish and Wildlife, Olympia, WA
Dr. Michael O'Farrell, National Marine Fisheries Service Southwest Fisheries Science Center, Santa Cruz, CA
Dr. Kiva Oken, National Marine Fisheries Service Northwest Fisheries Science Center, Seattle, WA
Ms. Tiffany Petersen, Makah Tribe, Neah Bay, WA
Ms. Corey Ridings, Ocean Conservancy, Santa Cruz, CA
Ms. Michele Robinson, Oceanbeat Consulting, Olympia, WA
Mr. Gordon Rose, Northwest Indian Fisheries Commission, Olympia, WA
Ms. Mindy Rowse, National Marine Fisheries Service Northwest Fisheries Science Center, Seattle, WA
Dr. Jason Schaffler, Muckleshoot Indian Tribe, Auburn, WA
Dr. Casey Schmidt, Suquamish Tribe, Bainbridge Island, WA
Ms. Kate Self, Columbia River Inter-Tribal Fish Commission, Portland, OR
Mr. Jeremiah Shrovnal, Washington Department of Fish and Wildlife, Olympia, WA
Dr. Julie Thayer, Farallon Institute, Petaluma, CA
Dr. Theresa Tsou, Washington Department of Fish and Wildlife, Olympia, WA
Mr. Kyle Van de Graaf, Washington Department of Fish and Wildlife, Olympia, WA
Ms. Lynn Langford Walton, All Gear Group, Centralia, WA
Mr. Verner Wilson, Friends of the Earth, Seattle, WA

Ecological thresholds in forecast performance for key United States West Coast Chinook salmon stocks – Addendum

Will Satterthwaite (NMFS/SWFSC) – will.satterthwaite@noaa.gov

This document, prepared to assist review of nonlinear relationships between environmental indicators and Chinook salmon forecast performance by the PFMC's SSC Ecosystem Subcommittee (SSCES), extends Satterthwaite et al. (2020) by:

1) Discussing expectations for model responses to additional data under different scenarios for how environmental conditions relate to forecast performance,
2) Re-evaluating select environmental driver – forecast performance relationships identified in Satterthwaite et al. (2020) upon the addition of more recently available data, and
3) Presenting a more detailed stepwise analysis of the highlighted relationships for Sacramento River Fall Chinook (SRFC, an indicator stock of high relevance to PFMC-managed ocean fisheries and one requiring ABC specification from the SSC) and discussing potential further uses of this approach or elements thereof for this stock.

I. Expected effects of adding new data

As acknowledged in Satterthwaite et al. (2020), a large number of environmental driver – location – lag – stock combinations were investigated, creating the risk of identifying spurious relationships. Satterthwaite et al. (2020) attempted to address this, at least in part, through randomization tests as well as considering the biological plausibility of the highlighted relationships. An additional test, suggested in the SSCES' earlier review of environmental driver - biological response threshold analysis (SSCES 2021), involves evaluation of model performance as new data are added.

The expected consequences of data addition depend on the nature of the indicator-response relationship (in this case, the response is forecast performance rather than a biological process *per se*). A strong, nonlinear relationship with high explanatory power over the full range of indicator values should be able to predict responses based on indicator values in novel years, as well as retaining support via the model selection procedure for nonlinear models, retaining R^2 at similar levels as more data are added, and creating new point estimates for threshold locations generally within the intervals calculated previously[1]. However, another type of "threshold" relationship, where extreme (and thus potentially rarely encountered) values of indicators elicit a response but "typical" values do not, may yield a more complicated pattern. It is possible that in a threshold scenario, the indicator will have little predictive power (especially on top of factors already informing the forecast) when environmental conditions are close to "normal". As a result, while model selection should continue to favor a nonlinear model as new data are added, the model may have poor predictive power for new data from "typical" years and adding data from "typical" years may decrease R^2 while additional data from "extreme" years should increase R^2.

These considerations should be kept in mind when evaluating these and other threshold relationships. When it comes to forecasting, if an indicator has high predictive power for forecast

[1] Perhaps counterintuitively, the interval describing the possible location for the threshold might not shrink as data are added, and might even widen, since with increased data we might be confident that the second derivative is not zero over a broader range.

performance across the full range of indicator values, it should be considered for direct incorporation into the forecast. Conversely, if an indicator only has predictive power at environmental extremes, there may be little utility in incorporating it into routine forecasting, but it may provide qualitative guidance on when environmental conditions are such that forecast performance may be poorer than usual, and managers might be advised to consider the consequences of increased uncertainty in forecasts under those conditions.

II. Re-evaluating select environmental indicator – forecast performance relationships from Satterthwaite et al. 2020 based on the latest available data

I first compiled updated data on forecast performance. Satterthwaite et al. (2020) used data on forecast performance from 1995-2016 for Puget Sound summer-fall Chinook stocks and 1995-2017 for SRFC (no strong nonlinear relationships were found for Klamath River Fall Chinook, so that stock is not considered further in this update[2]). Data for 2017-2019 (Puget Sound) and 2018-2020 (SRFC) were extracted from the 2021 Preseason Report 1[3] (PFMC 2021). Although some values for earlier years reported in the most recent version of Preseason Report 1 (PFMC 2021) differed from the values reported earlier and used by Satterthwaite et al. (2020), I judged these differences to be minor and did not update the old data in Satterthwaite et al. (2020).

Next, I updated the timeseries for select environmental indicators, using the same source documents/websites as the CCIEA report (CCIEA 2021a) when possible. Due to the workload associated with updating some indicators that are not routinely reported, I only updated a subset of environmental indicators. I updated environmental indicators if they were included in the nonlinear relationships documented for SRFC, if they were included in any relationships with $R^2 \geq 0.60$ for Puget Sound stocks, and for the single strongest relationship with a freshwater indicator.

In updating the indicator data sets (see Appendix A1 for two updates to sources), I discovered that due to a miscommunication among coauthors, the seasonal indexing of different indicators was not always consistent – i.e., "winter" corresponded to December-February for some indicators and January-March for others. Thus, seasonal indicators in this document are subscripted by a range of months. In addition, the selected freshwater indicator used in Satterthwaite et al. (2020) was a composite indicator for the Salish Sea and Washington coast ecoregion. Since the writing of that paper, a Puget Sound specific freshwater indicator has been developed (Stuart Munsch, NWFSC, pers. comm.), and its performance was compared to the original coarser indicator.

Having updated the datasets for forecast performance and these selected indicators, I refit each relationship from Satterthwaite et al. (2020)'s Table 2 for which new data were obtained – thus I re-evaluated all relationships for SRFC, all relationships for Puget Sound stocks with R2≥0.6 and select other relationships that involved indicators that had been updated based on their explanatory power for other stocks.

For each indicator-response relationship, Table 1 compares the selected model form and R^2 estimated from the dataset used in Satterthwaite et al. 2020 to the updated dataset. In all cases

[2] Though for completeness, I verified that no promising new relationships with the previously explored indicators emerged when adding new data for KRFC.

[3] SRFC forecast performance is reported in figure but not tabular form, so I obtained numeric values from Michael O'Farrell (SWFSC). Robin Ehlke provided an Excel version of the Puget Sound tables, which in some cases reports to a higher precision than the rounded values (nearest 100 fish) in the published document.

but one, a GAM was still selected as the best model form. The one exception was the relationship between South Puget Sound natural forecast performance and 1-day maximum river flow the previous year in the Salish Sea / Washington Coast ecoregion, which appeared best described by a linear model with autocorrelation (LMAC) in the updated dataset. Replacing this flow indicator with one specific to Puget Sound resulted in a GAM being selected as the best model, with a slight increase in R^2. R^2 often declined modestly with additional data, although it increased slightly in four cases (highlighted in bold) and decreased markedly in one case (highlighted in red, South Puget Sound hatchery versus autumn [October-December] sea level height off Alaska the previous year). The reason for this decrease in R^2 seems apparent based on the 2018 and 2019 datapoints in Figure 1 and raises concern that this apparent relationship is likely to be spurious or at least transient. The location of the apparent threshold near mean environmental conditions (i.e., near z=0) is also unexpected, although the modeled response does not deviate much from the mean response until well beyond the threshold.

Table 2 compares point estimates and intervals estimated for potential threshold locations between the previously-used and extended datasets. In all cases where a new threshold location was estimable, it was within the previous interval (but note that for the relationship between South Puget Sound natural and NPGO in December-February, a GAM was still the best-supported model but no threshold value could be identified). In all cases where estimates were made, point estimates were similar between datasets. Intervals were generally comparable as well, with the exception of a narrower interval for the South Puget Sound Natural – lagged spring SST off the WA coast relationship in the updated dataset.

The consistently high R^2 for the relationship between Tulalip hatchery forecast performance and NPI in December-February three years ago suggests there may be merit in considering NPI as a factor in that stock's forecast, especially if a reason for the large forecast error (not predicted by the fitted GAM) in 2013 can be identified and addressed in the future. It may also be useful to explore more proximate ecological processes for which NPI might be a proxy. However, given the shallow slope of the relationship over much of the range of the indicator, it may be more useful as a "threshold" warning of potential poor performance in extreme conditions (Figure 2, see also Satterthwaite et al. 2020 Figure 3b).

The modest decreases in R^2 for many other relationships cautions against premature adoption of these indicators into the respective forecasts, although small decreases in R^2 are not unexpected if these indicators simply warn of the potential for poor forecast performance in extreme environmental conditions. On the other hand, decreasing R^2 over time is consistent with non-stationary relationships (Litzow et al. 2020) or other cases where relationships between stock productivity and environmental indicators tend to break down over time (Myers 1998, Winship et al. 2015). More data from "extreme" years would be needed to better test potential "threshold" relationships that are mainly apparent during extreme conditions.

Table 1. Selected model form and adjusted R^2 for original (Satterthwaite et al. 2020) versus updated datasets for select indicator-forecast performance relationships. Rows in italics use new indicators compared to those used in Satterthwaite et al. 2020. Increases in R^2 are highlighted with bold text, and a large decrease in R^2 is highlighted with red text. NA = not fit to original dataset.

Sacramento River Fall Chinook Model Performance				
	1995-2017		1995-2020	
	Best Model	adj. R^2	Best Model	adj. R^2
SRFC / PDO.mar-may.lag0	GAM	0.44	GAM	**0.45**
SRFC / NPI.dec-feb.lag0	GAM	0.41	GAM	0.35
Puget Sound summer-fall Chinook Model Performance				
	1993-2016		1993-2019	
$R^2 \geq 0.6$	Best Model	adj. R^2	Best Model	adj. R^2
South Puget Sound natural / satSST.apr-jun.WA.lag1	GAM	0.75	GAM	0.57
South Puget Sound natural / NPGO.dec-feb.lag0	GAM	0.66	GAM	0.42
South Puget Sound natural / NPI.sep-nov.lag3	GAM	0.64	GAM	0.54
Tulalip hatchery / NPI.dec-feb.lag3	GAM	0.69	GAM	**0.70**
Hood Canal combined / SLH.jan-mar.AK.lag1	GAM	0.67	GAM	0.43
Stillaguamish natural / SLH.oct-dec.AK.lag1	GAM	0.62	GAM	0.55
Stillaguamish natural / SLH.apr-jun.AK.lag1	GAM	0.60	GAM	0.41
strongest freshwater relationship				
South Puget Sound natural / RivFlow.1daymax.SW.lag2	GAM	0.49	LMAC	0.48
South Puget Sound natural / RivFlow.1daymax.PS.lag2	*NA*	*NA*	*GAM*	*0.54*
$0.33 \leq R^2 < 0.6$, already extracted index for above				
South Puget Sound natural / satSST.apr-jun.WA.lag2	GAM	0.45	GAM	0.38
South Puget sound hatchery / SLH.oct-dec.AK.lag1	GAM	0.43	GAM	0.25
Hood Canal combined / SLH.oct-dec.AK.lag1	GAM	0.58	GAM	0.47
Hood Canal combined / SLH.apr-jun.AK.lag1	GAM	0.54	GAM	0.44
Snohomish hatchery / NPGO.dec-feb.lag0	GAM	0.49	GAM	**0.50**
Strait of Juan de Fuca / SLH.apr-jun.AK.lag1	GAM	0.47	GAM	**0.49**

Figure 1. Relationship between forecast performance (proportional error relative to mean absolute proportional error, with positive values indicating overforecasting) for South Puget Sound hatchery and autumn (October-December) sea level height off Alaska the previous year. Dotted line is fitted relationship, overlain grey thick line is potential location of threshold, red arrow is point estimate of threshold location. Environmental indices are z-scored.

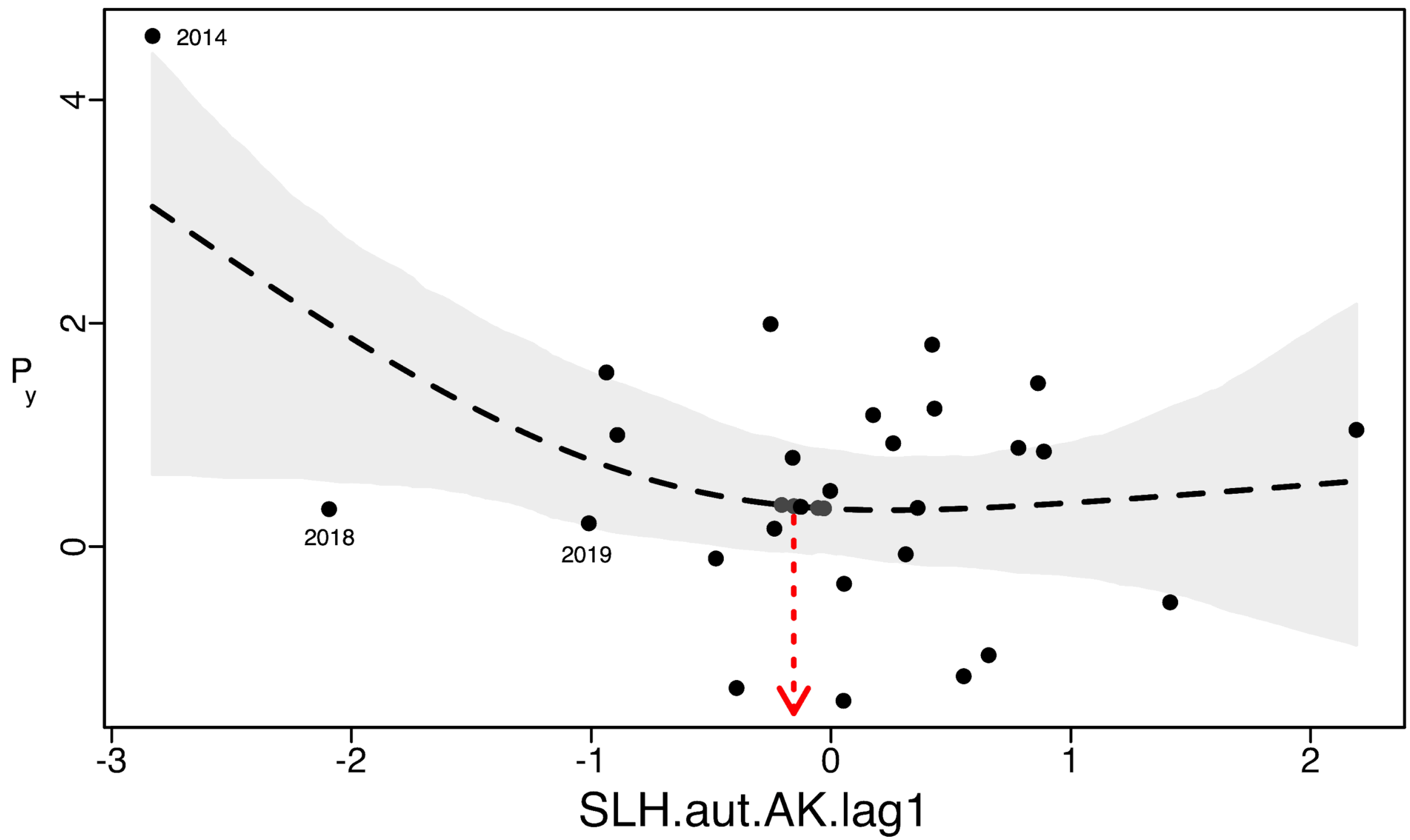
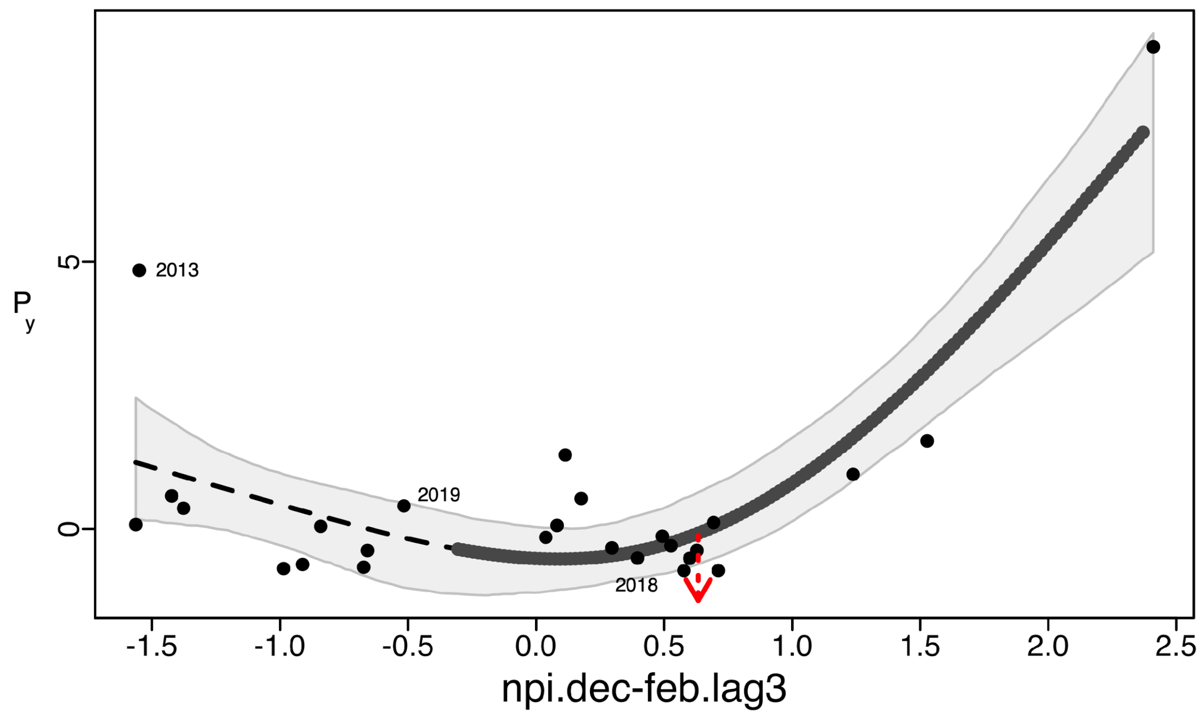


Table 2. Estimated threshold locations and intervals of potential threshold locations for original (Satterthwaite et al. 2020) versus updated datasets for select indicator-forecast performance relationships. Rows in italics use new indicators compared to those used in Satterthwaite et al. 2020. NA = not applicable (not fit to original dataset), NT = no threshold estimated (either because linear model selected, or GAM was selected but no threshold identified based on second derivative).

Sacramento River Fall Chinook						
	1995-2017			1995-2020		
	Threshold	(interval)		Threshold	(interval)	
SRFC / PDO.mar-may.lag0	-0.65	-1.79	0.25	-0.54	-1.83	0.28
SRFC / NPI.dec-feb.lag0	0.64	0.15	0.64	0.58	0.14	0.58
Puget Sound summer-fall Chinook						
	1993-2016			1993-2019		
$R^2 \geq 0.6$	Threshold	(interval)		Threshold	(interval)	
So. Puget Sound natural / satSST.apr-jun.WA.lag1	-0.14	-1.76	0.53	-0.17	-0.58	0.38
South Puget Sound natural / NPGO.dec-feb.lag0	-0.57	-1.67	0.11	NT	NT	NT
South Puget Sound natural / NPI.sep-nov.lag3	0.64	-0.30	2.08	0.64	-0.29	2.07
Tulalip hatchery / NPI.dec-feb.lag3	0.60	-0.28	2.35	0.63	-0.27	2.41
Hood Canal combined / SLH.jan-mar.AK.lag1	-0.81	-2.03	0.01	-0.68	-1.77	0.06
Stillaguamish natural / SLH.oct-dec.AK.lag1*	-1.10	-3.16	-0.11	-0.94	-2.78	-0.08
Stillaguamish natural / SLH.apr-jun.AK.lag1	-1.01	-2.31	-0.22	-0.93	-2.17	-0.19
strongest freshwater relationship						
South Puget Sound natural / RivFlow.1daymax.SW	0.76	-0.05	2.05	NT	NT	NT
South Puget Sound natural / RivFlow.1daymax.PS	*NA*	*NA*	*NA*	*0.95*	*0.09*	*2.34*
$0.33 \leq R^2 < 0.6$, already extracted index for above						
So. Puget Sound natural / satSST.apr-jun.WA.lag2	-0.49	-2.24	0.17	-0.50	-2.29	0.14
South Puget sound hatchery / SLH.oct-dec.AK.lag1	-0.06	-0.41	0.11	-0.16	-0.16	0.02
Hood Canal combined / SLH.oct-dec.AK.lag1	-1.10	-3.16	-0.14	-0.94	-2.78	-0.13
Hood Canal combined / SLH.apr-jun.AK.lag1	-0.78	-2.30	0.32	-0.65	-2.16	0.32
Snohomish hatchery / NPGO.dec-feb.lag0	0.48	0.06	0.48	0.52	-0.01	0.52
Strait of Juan de Fuca / SLH.apr-jun.AK.lag1	0.55	-0.24	0.62	0.60	-0.21	1.04

*This relationship has two separate threshold locations identified. I reported the lower one as it seems to have the stronger response.

Figure 2. Relationship between forecast performance (proportional error relative to mean absolute proportional error, with positive values indicating overforecasting) for Tulalip hatchery forecast performance and NPI in December-February three years ago. Dotted line is fitted relationship, overlain grey thick line is potential location of threshold, red arrow is point estimate of threshold location. Environmental indices are z-scored.



III. Detailed analysis of indicators of SRFC forecast performance

I selected SRFC for a more detailed analysis of the potential indicator-forecast performance relationships identified because it is a major contributor to Council-area ocean fisheries and because I know a consistent forecast methodology was used throughout[4] the dataset. Because this analysis was restricted to a single stock, to make the response variable more directly interpretable I looked at proportional error $\left(\frac{forecast-observed}{observed}\right)$. I dropped the denominator from equation 1 of Satterthwaite et al. 2020 since it served to put forecasts for different stocks, varying in their overall performance, on a similar scale; but this is not needed when only looking at performance for a single stock. Thus, for the plots that follow (but not the plots in Satterthwaite et al. 2020), a forecast performance score of 0 corresponds to an accurate forecast (this is also true for Satterthwaite et al. 2020), while scores of -0.5, +0.5, and +1.5 correspond to a forecast 50% lower than the postseason estimate, 50% higher than the postseason estimate, and 150% higher than the postseason estimate, respectively. Note that proportional error is bounded below at -1 (provided forecasts cannot be negative) but there is no hard upper bound.

Considering 10 years as a minimum sufficient dataset, I repeated the model fitting and model selection process for SRFC datasets corresponding to 1995-2004 and subsequent datasets with one more year added each iteration.

For relationships with PDO in March-May the year of return, for datasets prior to 2008 the best-supported model was linear and R^2 was very low (Table 3). Adding data from 2008 (the first year of very low spring PDO and severe overforecasting) still led to a linear model being selected, with very low R^2, but for that year a GAM offered a substantially higher R^2 of 0.36. For 2009 (another year of low PDO and overforecasting), a GAM was selected and R^2 increased to 0.67. GAMs were selected every subsequent year except 2015, which corresponded to a poorly predicted year (Figure 3), followed by slightly lower R^2 and wider intervals on the threshold location for every year thereafter.

For relationships with NPI in December-February the year of return, linear models were selected, with low R^2 values, for 2004-2008. In 2009, a linear model was still selected, but R^2 increased to 0.35. Starting in 2012, a GAM was selected, with R^2 initially increasing to 0.56 and then slowly decreasing, with a notable drop in 2015 (Table 4, Figure 4). The drop in R^2 from the 1995-2018 value reported in Satterthwaite et al. 2020 to the 1995-2020 value reported here appears attributable to underforecasting in 2019 despite moderately high NPI (Figure 4).

[4] In fact, a new forecasting methodology was adopted for SRFC in 2014, but I was able to work with Michael O'Farrell (SWFSC) to obtain retrospective estimates of what the forecasts would have been if using the current approach for earlier years, based on data available at the time.

Table 3. Selected model, adjusted R^2, and point estimates and intervals of threshold locations for the relationship between SRFC forecast performance and PDO in March-May the year of return, for datasets of varying lengths.

Dataset	Best Model	adjusted R^2	Threshold Location	(interval)	
1995-2004	Linear	0.22	-	-	-
1995-2005	Linear	0.30	-	-	-
1995-2006	Linear	0.29	-	-	-
1995-2007	Linear	0.03 (GAM similar)	-	-	-
1995-2008	Linear	-0.06 (GAM 0.36)	-	-	-
1995-2009	GAM	0.67	-0.52	-0.80	-0.02
1995-2010	GAM	0.68	-0.47	-0.73	0.08
1995-2011	GAM	0.67	-0.11	-0.69	0.20
1995-2012	GAM	0.52	-0.49	-0.69	0.00
1995-2013	GAM	0.53	-0.66	-0.77	0.05
1995-2014	GAM	0.54	-0.66	-0.81	-0.06
1995-2015	Linear	0.00 (GAM 0.51)	-	-	-
1995-2016	GAM	0.47	-0.54	-1.83	0.34
1995-2017	GAM	0.44	-0.64	-1.83	0.30
1995-2018	GAM	0.44	-0.52	-1.83	0.30
1995-2019	GAM	0.45	-0.52	-1.83	0.30
1995-2020	GAM	0.45	-0.54	-1.83	0.30

Figure 3. Relationship between forecast performance (proportional error, with positive values indicating overforecasting) for SRFC forecast performance and PDO in March-May the year of return. Dotted line is fitted relationship, overlain grey thick line is potential location of threshold, red arrow is point estimate of threshold location. Environmental indices are z-scored. (Note the grey thick line at the very far right suggesting a potential additional threshold at high PDO values.)

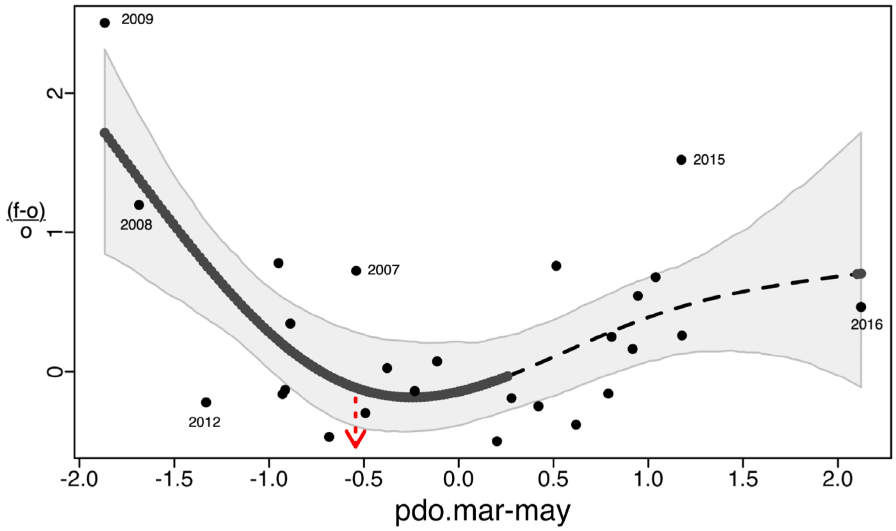
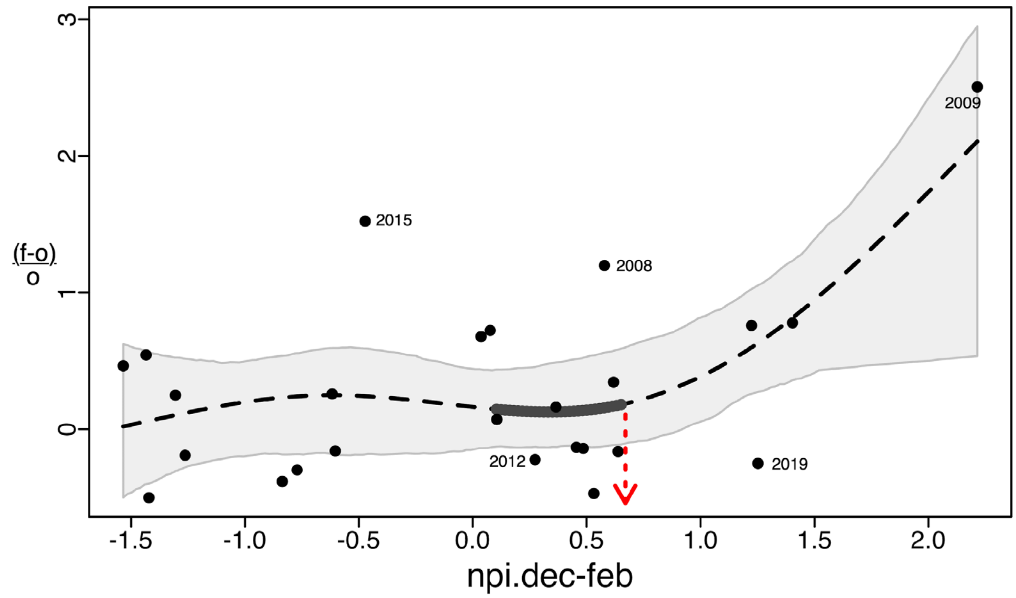


Table 4. Selected model, adjusted R^2, and point estimates and intervals of threshold locations for the relationship between SRFC forecast performance and NPI in December-February the year of return, for datasets of varying lengths.

Dataset	Best Model	adjusted R^2	Threshold Location	(interval)	
1995-2004	Linear	0.06	-	-	-
1995-2005	Linear	-0.08	-	-	-
1995-2006	Linear	-0.07	-	-	-
1995-2007	Linear	-0.09	-	-	-
1995-2008	Linear	-0.04	-	-	-
1995-2009	Linear	0.35	-	-	-
1995-2010	Linear	0.39	-	-	-
1995-2011	Linear	0.40	-	-	-
1995-2012	GAM	0.56	0.34	0.01	0.34
1995-2013	GAM	0.53	0.40	0.05	0.40
1995-2014	GAM	0.53	0.44	0.03	0.44
1995-2015	GAM	0.42	0.49	0.12	0.49
1995-2016	GAM	0.39	0.41	0.03	0.41
1995-2017	GAM	0.41	0.48	0.09	0.48
1995-2018	GAM	0.41	0.59	0.05	0.59
1995-2019	GAM	0.36	0.54	0.16	0.54
1995-2020	GAM	0.35	0.67	0.14	0.69

Figure 4. Relationship between forecast performance (proportional error, with positive values indicating overforecasting) for SRFC forecast performance and NPI in December-February the year of return. Dotted line is fitted relationship, overlain grey thick line is potential location of threshold, red arrow is point estimate of threshold location. Environmental indices are z-scored.



I considered evaluating some formal metric of out-of-sample prediction like the summed log-likelihood of each new year's prediction under the model fitted through data up to the prior year, or plotting how often out-of-sample predictions fell in various quantiles of predictive uncertainties, but since all the model confidence intervals are based on bootstrapping it was not immediately clear how to do this in a computationally efficient way and I ran out of time when faced with conflicting demands.

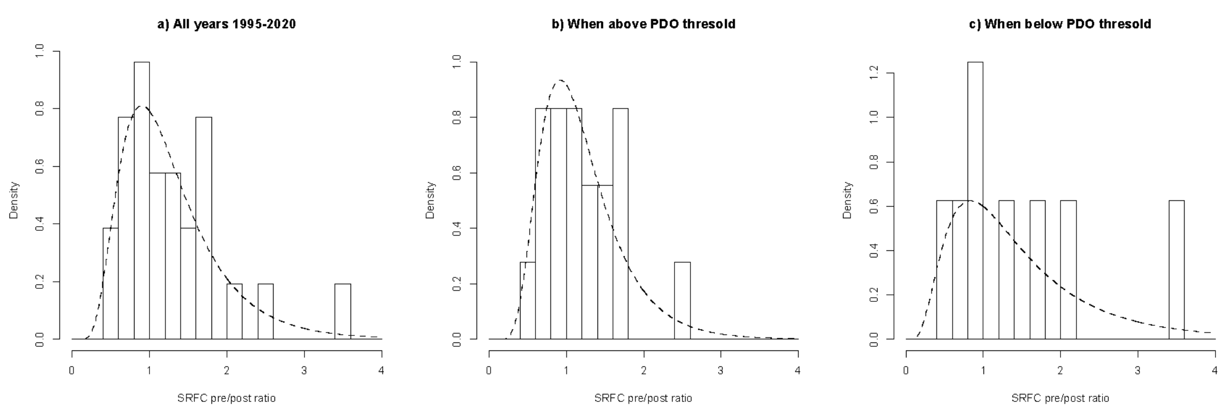
As noted by Satterthwaite et al. (2020), both nonlinear relationships identified for SRFC forecast performance have modest explanatory power when evaluated over the full range of indicator values, and the explanatory power of winter NPI appears to have decreased with the addition of more recent data. As further noted by Satterthwaite et al. (2020), the relationships fitted here are sensitive to a small number of datapoints with high leverage, and randomization tests suggest a substantial probability of spuriously detecting relationships with R^2 at least as high as those found here based on the number of relationships tested. On the other hand, the probability of obtaining spurious results calculated under the null mode is sensitive to distributional assumptions, the identified relationships are at a plausible lag given the nature of the forecast (specifically the information on early cohort strength provided by jack returns) and likely would have been identified as more promising candidates than most others a priori, and poor forecast performance at rarely encountered environmental extremes seems mechanistically plausible.

While increasing the number of indicators tested increases the risk of spurious results, hypothesis-driven inspection of local indicators motivated by mechanistic hypotheses and/or demonstrated relationships to stock productivity could be warranted and might increase explanatory power. The Habitat Committee (HC, CCIEA 2021b) has identified a number of indicators thought to be informative with respect to SRFC productivity, and some of these factors may affect forecast performance, potentially in a nonlinear way. NMFS SWFSC (2021) suggested that thiamine deficiency, potentially linked to local anchovy abundance, could be linked to errors in the SRFC forecast. More localized indicators of ocean conditions than the PDO, for example the Central and/or Northern California MOCI (García-Reyes and Sydeman 2017), might also be informative (Julie Thayer, Farallon Institute, pers. comm.).

Inspection of Figure 3 suggests that the performance of the SRFC forecast is sensitive to environmental conditions and breaks down at low (and possibly high) spring PDO values but does well (or possibly tends to underforecast) when spring PDO is near its mean. Although more work is needed both narrowly to establish the robustness of this relationship and more broadly on how to incorporate forecast error and uncertainty into salmon management, as a thought experiment I considered the implications of a "threshold" value of -0.54[5] for z-scored springtime PDO (assuming it was known or could be reliably forecasted in time to inform preseason management), contrasting the implications of threshold-dependent interpretations of forecasts versus interpretations of forecasts without considering environmental state.

Figure 5a shows the distribution of annual ratios of the SRFC preseason forecast to the postseason abundance estimate for all years 1995-2020, while Figure 5b is restricted to years above the potential PDO threshold and Figure 5c is restricted to years below the threshold. I also fitted lognormal distributions to each set of ratios.

Figure 5. Histograms and fitted lognormal distributions of SRFC forecast performance (preseason : postseason abundance estimate ratio) for all years 1995-2020 combined (a), just years above a z-score of -0.54 in springtime PDO (b), or just years below a z-score of -0.54 in springtime PDO.



Forecast performance (pre:post) for all years combined, with no consideration of environmental state, was best described by a lognormal distribution with median 1.14[6] and log-scale SD 0.49. This implies that, if a median-unbiased forecast was desired and environmental indicators were not considered, the preseason forecast of abundance should be multiplied by 0.88 (the inverse of 1.14). By rough analogy with the P*/sigma approach used in groundfish (PFMC

---

[5] Note that this is on the scale of z-scores for 1995-2014 springtime PDOs rather than a numeric value of -0.54 for the PDO index itself. Any potential thresholds ultimately considered for management use should probably be defined on the index scale rather than z-scores sensitive to the range of years included.

[6] Note however that this estimate is accompanied by high uncertainty, with a 95% confidence interval on the median (based on a normal approximation to the 95% confidence interval on the mean of the log ratio) spans from 0.95-1.38. With 26 years' of data and a log-SD of 0.49, the smallest ratio with a 95% confidence interval excluding 1.0 is 1.21, highlighting the challenges of precisely characterizing forecast bias given realistic variability and sample sizes.

2020) and CPS, it might be appropriate to consider using the 0.55 or 0.60 quantile[7] of this distribution rather than the median, implying a multiplier of 0.83 or 0.78.

Forecast performance in years with springtime PDO above the threshold was described by a lognormal distribution with median 1.10 and log-scale SD 0.42. This would imply multipliers of 0.91, 0.86, or 0.82 for median-unbiased forecasts or for levels of risk tolerance analogous to P* of 0.45 or 0.40, respectively. This results in only slightly smaller buffers than the all-years case, due to multiple instances of overforecasting even when above the threshold but does reflect somewhat less uncertainty when avoiding extreme conditions (and excluding high PDO years like 2015 might shrink the buffer further). Conversely, forecast performance when springtime PDO was below the threshold was described by a lognormal distribution with median 1.24 and log-scale SD 0.63. This would imply multipliers of 0.81, 0.75, or 0.69 for median-unbiased forecasts or for levels of risk tolerance analogous to P* of 0.45 or 0.40, respectively; resulting in more substantial buffering in years of extreme environmental conditions.

The proposal to use two different distributions for the forecast buffer, with each distribution corresponding to one side of the threshold, is simpler than an approach which applies a correction factor based on the fitted relationship between forecast error and environmental indicator at the specific value of the indicator for the year in question, and is likely to be more robust than the fitted relationship when making projections for data-sparse regions of parameter space (i.e., we have little confidence in the exact value of the fitted curve when at extreme environmental values). However, this neglects any accounting for how far away from the threshold the current environmental state is, whereas nonlinear relationships often imply substantially worse forecast performance far away from the fitted thresholds but only modest deterioration in performance slightly past the threshold. If the interest is in identifying the "threshold" as corresponding to most rapid response in performance, basing the threshold location on maximizing the first derivative rather than second derivative may be more appropriate (Large et al. 2013). Or the "threshold" might be crossed when confidence intervals on the modeled response beyond the threshold exclude the modeled response under average conditions.

Of course, there is currently no mechanism for applying multipliers to salmon abundance forecasts to account for either bias or uncertainty, nor is it obvious that median-unbiased forecasts are risk neutral[8] when considering the consequences of varying magnitudes and directions of forecast error, especially when factoring mixed-stock constraints on ocean fisheries. It is also unclear whether the correction should be based on estimates of pre-fishing or post-fishing abundance. Nevertheless, the results presented here suggest that simply accepting point estimates is likely not risk-neutral, both because of potential biases and because of potential increases in both bias and uncertainty under certain environmental conditions.

Recommendations

[7] Here, I present preseason:postseason ratios for consistency with the metrics presented in Satterthwaite et al. (2020) and the ratios typically reported in Preseason Report 1. This contrasts with the P*/sigma approach where the distribution being modeled is of true biomass (or more recently, the true OFL) relative to the assessment output, analogous to the postseason:preseason ratio. Thus, using the 0.55 or 0.60 quantile of the pre:postseason ratio is roughly analogous to using a P* of 0.45 or 0.40.

[8] Indeed, the mode of all three distributions is near a pre:post ratio of 1.0. On the other hand, the arithmetic mean ratio (loosely speaking, the expected error) is even further above 1.0 than the median.

• Any of the apparent relationships documented here should be viewed with caution due to the caveats raised in Satterthwaite et al. (2020), along with decreasing R^2 for many relationships in the face of new data. At the same time, we should realize that characterizing responses to rarely encountered environmental conditions is inherently challenging, and so is forecasting under such conditions.

• Continue evaluation of promising relationships between environmental indicators and salmon forecast performance, noting that data from "extreme" conditions is what would be most useful in further testing these relationships. This approach is probably most appropriately applied as a screening tool to identify potential relationships and provide focus for future finer-scale research. Indeed, the improved performance of the Puget Sound-specific freshwater indicator over the regional freshwater indicator in describing South Puget Sound natural forecast performance suggests promise of more finely resolved indicators.

• Consider incorporating winter NPI at a lag of three years in the Tulalip hatchery forecast (and/or attempt to identify ecological mechanisms for which this is a proxy) and look into potential causes of that forecast's poor performance in 2013.

• Consider adding NPI to the CCIEA report. NPI in the wintertime has the advantage of being available sooner than springtime PDO for potentially informing the interpretation of SRFC forecasts, and also may be relevant (and available sufficiently far in advance given the lagged relationships) for interpreting forecasts of South Puget Sound natural and Tulalip hatchery abundances.

• Judiciously apply this approach, including randomization tests and attempts at out-of-sample validation, to potential new indicators tailored for SRFC and KRFC, leveraging indicators being developed by the HC in association with the rebuilding plans as well as other potentially relevant local indicators.

• Consider a precautionary approach to using SRFC forecasts in years of extreme (forecasted) springtime PDO or wintertime NPI. Ideally, this would be done via a formal process for incorporating scientific uncertainty and risk tolerance into ABC specifications, as is done for CPS and groundfish.

• Elements of the approach outlined here may serve as methods for quantifying forecast uncertainty – either overall uncertainty regardless of conditions or indicator-dependent metrics of uncertainty. Although uncertainty buffers typically reduce fishing opportunity compared to reliance on point estimates, an indicator-driven bias correction could potentially increase fishing opportunity in some years.

References cited

California Current Integrated Ecosystem Assessment [CCIEA] Team. 2021a. California Current Integrated Ecosystem Assessment (CCIEA) California Current Ecosystem Status Report, 2021. Report to Pacific Fishery Management Council. Available at: https://www.pcouncil.org/documents/2021/02/i-1-a-iea-team-report-1.pdf/

California Current Integrated Ecosystem Assessment [CCIEA] Team. 2021b. Supplementary Materials to the California Current Integrated Ecosystem Assessment (CCIEA) California Current Ecosystem Status Report, 2021. Report to Pacific Fishery Management Council. Available at: https://www.pcouncil.org/documents/2021/02/i-1-a-iea-team-report-2.pdf/ [particularly Appendix H.5]

García-Reyes, M. and W. J. Sydeman. 2017. California Multivariate Ocean Climate Indicator (MOCI) and marine ecosystem dynamics. Progress in Oceanography 120:352-369. Available at: http://dx.doi.org/10.1016/j.ecolind.2016.08.045

Large, S. I., G. Fay, K.D. Friedland, and J.S. Link. 2013. Defining trends and thresholds in responses of ecological indicators to fishing and environmental pressures. ICES Journal of Marine Science 70:755-767. Available at: https://dx.doi.org/10.1093/icesjms/fst067

Litzow, M.A., M. E. Hunsicker, N.A. Bond, B.J. Burke, C.J. Cunningham, J.L. Gosselin, E.L. Norton, E.J. Ward, and S.G. Zador. 2020. The changing physical and ecological meanings of North Pacific Ocean climate indices. Proceedings of the National Academy of Sciences 117:7665-7671. Available at: https://dx.doi.org/10.1073/pnas.1921266117

Myers, R.A. 1998. When do environment–recruitment correlations work? Reviews in Fish Biology and Fisheries 8:285–305. Available at: https://dx.doi.org/10.1023/A:1008828730759

National Marine Fisheries Service Southwest Fisheries Science Center [NMFS SWFSC]. 2021. NMFS SWFSC Report (March 2021): Thiamine Deficiency in West Coast Salmon. Report to Pacific Fishery Management Council. Available at: https://www.pcouncil.org/documents/2021/02/e-1-attachment-1-thiamine-deficiency-in-west-coast-salmon.pdf/

Pacific Fishery Management Council [PFMC]. 2020. Status of the Pacific Coast Groundfish Fishery. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, Oregon 97220-1384. Available at: https://www.pcouncil.org/documents/2020/09/status-of-the-pacific-coast-groundfish-fishery-stock-assessment-and-fishery-evaluation-september-2020.pdf/#page=292

Pacific Fishery Management Council [PFMC]. 2021. Preseason Report I: Stock Abundance Analysis and Environmental Assessment Part 1 for 2021 Ocean Salmon Fishery Regulations. (Document prepared for the Council and its advisory entities.) Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, Oregon 97220-1384. Available at: https://www.pcouncil.org/documents/2021/03/2021-preseason-report-i.pdf/

Satterthwaite, W.H., K.S. Andrews, B.J. Burke, J.L. Gosselin, C.M. Greene, C.J. Harvey, S.H. Munsch, M.R. O'Farrell, J.F. Samhouri, and K.L. Sobocinski. 2020. Ecological thresholds in forecast performance for key United States West Coast Chinook salmon stocks. ICES Journal of Marine Science 77:1503-1515. Available at: https://doi.org/10.1093/icesjms/fsz189

Scientific and Statistical Committee Ecosystem Subcommittee [SSCES]. 2021. Scientific and Statistical Committee's Ecosystem Subcommittee Report. Report to Pacific Fishery Management Council. Available at: https://www.pcouncil.org/documents/2021/03/i-1-b-supplemental-ssc-report-1.pdf/

Winship, A.J., M.R. O'Farrell, W.H. Satterthwaite, B.K. Wells, and M.S. Mohr. 2015. Expected future performance of salmon abundance forecast models with varying complexity. Canadian Journal of Fisheries and Aquatic Sciences 72:557-569. Available at: https://dx.doi.org/10.1139/cjfas-2014-0247

Appendix A1 updated sources of environmental indicator data

For the most part, environmental indicators were obtained from the sources cited in Satterthwaite et al. (2020). Two of the indicators are reported in CCIEA reports, using different source citations. For those indicators, I obtained new values from the sources cited in CCIEA reports, while verifying that values for previous years were identical.

PDO
ICES paper cited http://research.jisao.washington.edu/pdo/
For extended analysis and consistency with CCIEA report, used http://oceanview.pfeg.noaa.gov/erddap/tabledap/cciea_OC_PDO.csvp?time%2CPDO

NPGO
ICES paper cites http://www.o3d.org/npgo/npgo.php
For extended analysis and consistency with CCIEA report, used http://oceanview.pfeg.noaa.gov/erddap/tabledap/cciea_OC_NPGO.csvp?time%2CNPGO