**Tracking and forecasting community responses to climate perturbations in the California Current Ecosystem**  
Mary E. Hunsicker1,\*,†, Eric J. Ward2\*, Michael A. Litzow3, Sean C. Anderson4, Chris J. Harvey2, John C. Field5, Jin Gao6, Michael G. Jacox7, Sharon Melin8, Andrew Thompson9

**Supporting Information**

S1 Appendix: Standardization of time series from spatially resolved datasets.  
 Datasets collected through the Rockfish Recruitment and Ecosystem Assessment Survey (RREAS; pelagic juvenile fish and invertebrates survey) and the California Cooperative Oceanic Fisheries Investigations (CalCOFI, ichthyoplankton survey) include spatial attributes and were standardized using Generalized Additive Models (GAM) to create a univariate time series for each species included in our analysis.

*RREAS Survey*

Because many species in the RREAS survey were absent from a large number of observed trawls, we modeled species occurrence and abundance separately, using a delta GAM approach with two sub-models (Hastie 1990, Guissan 2002). In the first sub-model, species occurrence (presence-absence) was modelled using a binomial GAM with a logit link. In the second model (‘positive model’), species abundance (count) conditional on the catch of at least one individual was modelled using a Poisson GAM with a log link. The formulations of the two sub-models were analogous and as follows:

where *P*, the probability of species occurrence*,* or *C*, an estimate of species abundance when the speciesis present, is a function of year (*yr,* as factor), latitude (φ) and longitude (λ), and Julian day (*jday*). A two-dimensional smoothing function is denoted by *s* and indicates an error term.

Using the fitted GAMs for each species, we generated predictions of overall abundance of individual species. First, we created spatial occurrence and abundance distribution profiles by creating a grid of all combinations of model explanatory variables: geographic coordinates (latitude and longitude), Julian day, and year. We restricted the range of geographic coordinates to 20 values within the upper 0.8 and lower 0.2 quantiles of the data set, and bounded the range of Julian days to the upper 0.9 and lower 0.1 quantiles. This was done to avoid edge effects that can result in unrealistic predictions. All unique sampling years were included in the grid. Next, the probability of occurrence and estimates of abundance were predicted for each combination of the explanatory variables, and then the predictions from the two sub-models were multiplied to determine the overall abundance of individual species. Lastly, we calculated the mean standardized abundance of each species in each year from the prediction grid to generate the univariate time series of species abundance used in our study analyses.

*CalCOFI survey*

The CalCOFI time series of ichthyoplankton densities were standardized using a Tweedie GAM (power parameter fixed at 1.25) (Tweedie 1984, Dunn ad Smyth 2002). The model formulation was as follows:

where species density (*D*) is a function of year (*yr*), season (spring, summer), and latitude () and longitude (. Thetwo-dimensional smoothing function and error term are indicated by s and , respectively.

Similar to above, we used the fitted GAMs for each species to generate predictions of species densities. We first created a spatial abundance distribution profile by creating a grid of different combinations of the model covariates. Here we restricted the range of geographic coordinates (latitude and longitude) to those from sampling stations that were sampled 20 years. We included all sampling years in the grid and limited the season to spring only. Species densities were then predicted for each combination of the model covariates, and the univariate time series of species abundance was generated by calculating the mean standardized density of each species in each year from the prediction grid.

The delta GAMs and Tweedie GAMs were run using the ‘mgcv’ package (v1.8-34; Wood 2011, 2017) in R.

*References*

Dunn PK, Smyth GK. Series evaluation of Tweedie exponential dispersion model densities. Statistics and Computing 2005; 15: 267-280.

Guisan A, Edwards TC, Hastie T. Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecol. Modell. 2002; 157: 89–100.

Hastie T, Tibshirani R. Generalized additive models. London: Chapman and Hall; 1990.

Tweedie MCK. An index which distinguishes between some important exponential families. Statistics: Applications and New Directions. Calcutta: Indian Statistical Institute. Proceedings of the Indian Statistical Institute Golden Jubilee International Conference (Eds. J. K. Ghosh and J. Roy) 1984; pp. 579-604.

Wood SN. Generalized Additive Models: An Introduction with R (2nd edition). Chapman and Hall/CRC; 2017.

Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. Journal of the Royal Statistical Society (B) 2011; 73:3 -36.

S1 Table: A list of climate and biology time series included in the analyses and the associated units of measurements, data transformations (if applicable), years of data, and data sources.



1UCSC ROMS model, 2NOAA Marine Mammal Lab, 3Point Blue / USFW, 4CalCOFI, 5NOAA Rockfish Recruitment & Ecosystem Assessment Survey

S2 Table. Summary information for the Bayesian DFA biology-covariate one-trend candidate models and the top two biology only models (years 1981-2017). Only one trends models are shown because they outperformed all two and three trend models. The table indicates whether process error was estimated, the number of model trends, estimated log predicted densities (ELPD), standard error of ELPD, and the environmental covariate included in model. All models had an AR(1) coefficient on the trend and Student-t deviations. BEUTI = Biologically Effective Upwelling Transport Index; CUTI = Coastal Upwelling Transport Index; ILD = Isothermal Layer Depth; SST = Sea Surface Temperature; BV = Brunt-Väisälä frequency; SSH = Sea Surface Height.



S3 Table. Observations, predictions, and prediction errors for single species parameters in 2018. The predictions and error were derived from the top four biology-covariate models shown in Table 2. Species with the lowest predictions are highlighted in bold.



S1 Figure. Climate and biology time series used in the study analyses. They include climate indices (gray), ichthyoplankton abundance indices (orange), seabird reproductive success (teal), juvenile fishes and invertebrates (blue), and California sea lion pup parameters. Time series are scaled by their mean and standard deviation.



S2 Figure. AR(1) coefficient on the southern/central California latent climate trend (top) and support for a heavy-tailed deviations of the latent trend (bottom). Smaller values of the degrees of freedom parameter (e.g., less than 10 or 20) are consistent with heavy-tailed deviations.





S3 Figure. The Student-t deviations degrees of freedom parameter (nu) in the southern/central California biology trend.



S4 Figure. A summary of the effect of the Cumulative Upwelling Transport Index (CUTI) on the individual single species parameter included in the DFA analyses. Cal. = California, Juv. = juvenile fish stage, Larv. = larval fish stage, Juv./adult = juvenile and adult stages combined, all other fish are larval fish. Blue error bars reflect 95% credible intervals.



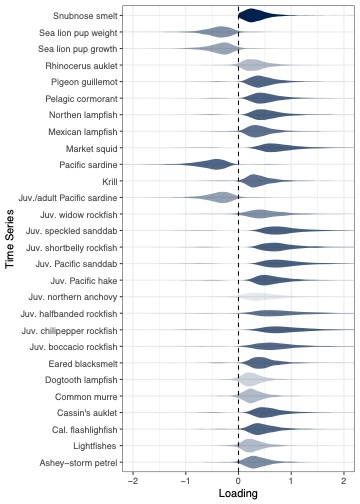
S5 Figure. A summary of the effect of the Isothermal Layer Depth (ILD) on the individual single species parameter included in the DFA analyses. Cal. = California, Juv. = juvenile fish stage, Larv. = larval fish stage, Juv./adult = juvenile and adult stages combined, all other fish are larval fish. Blue error bars reflect 95% credible intervals.



S6 Figure. A summary of the effect of the sea surface temperature on the individual single species parameter included in the DFA analyses. Cal. = California, Juv. = juvenile fish stage, Larv. = larval fish stage, Juv./adult = juvenile and adult stages combined, all other fish are larval fish. Blue error bars reflect 95% credible intervals.



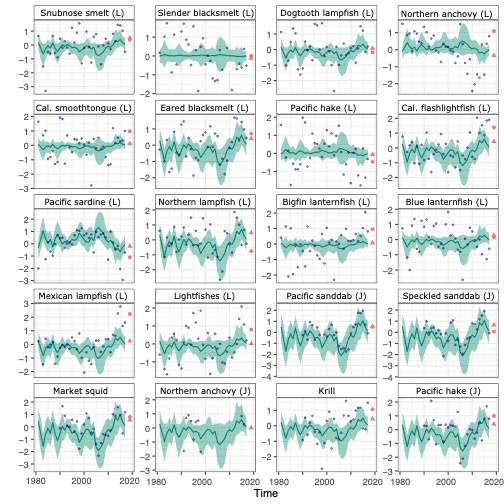
S7 Figure. Community variability in the southern California Current ecosystem (1981-2018). See Fig. 6 for shared trend and below for posterior distributions for loadings on individual time series. Only time series with ≥ 90% of loading distributions above or below zero are shown. Loadings with darker shading indicate time series loading most strongly on the biology trend. See Table S1 for times series details.



S8 Figure. Forecasts and model estimates of the ‘true’ community state in the southern and central California Current in years 2008—2018 (circle, with 95% credible intervals).



S9 Figure. Fitted values for biology-covariate model including BEUTI (nitrate flux) as a covariate (1981-2017). Blue circles = observations; blue line = fitted model, red circles = 2018 observations; red triangles = model predictions of single species parameters in 2018. Fish life stages: L = larval, J = juvenile, A = adult. 2018 observations were not available for juvenile northern anchovy.



S9 Figure cont. Fitted values for SCC biology model including BEUTI (nitrate flux) as a covariate (1981-2017). Blue circles = observations; blue line = fitted model; red circles = 2018 observations; red triangles = model predictions of single species parameters in 2018. Fish life stages: L = larval, J = juvenile, A = adult. 2018 seabird observations were only available for Brandt’s cormorant, Cassin’s auklet, and Common murre.



S10 Figure. Log coefficient of variation (CV) of 2018 predictions of individual species parameters plotted against (a, b) the mean and log CV of loadings related to each species, and (c, d) the mean and log CV of coefficients relating each species to BEUTI (nitrate flux). Model loadings, coefficients and predictions were derived from Model 1 listed in Table 2.

(a)

(c)

(b)

(d)B

