* Current state of ecological forecasting: mini-lit review. **Papers selected which: 1) use models, 2) quantify uncertainty to make a probabilistic forecast, and 3) run the model outside a given training period. (it is hard to find studies that make *probabilistic* forecasts…maybe this can also be an emphasis of the current state and what is missing, rather than a need to exclude papers that don’t include probabilities/quantified uncertainties)** 
  1. Estes et al 2013: empirical does better than mechanistic at predicting observed productivity and suitability of crops in south Africa
     1. Paper aimed at comparing empirical and mechanistic approaches
     2. Uses four models (2 mechanistic, 2 empirical)
     3. Uncertainty: sort of, false positive and true positive estimates?
     4. Tested outside the training period, yes
     5. Use: basic (uses an applied topic but is not directly applied through this study)
  2. Lindegren 2013: process-based models to forecast Baltic cod
     1. Models: yes, mechanistic
     2. Uncertainty: uses CI, accounts for climate and process uncertainty in projections
     3. Tested outside training period: yes, with multiple scenarios, which include recent past as validation
     4. Use: basic
  3. ~~Fenocci et al 2019: use process-based (GLM) to hindcast +validate phytos in large lake~~
     1. ~~Model: yes, GLM-AED~~
     2. ~~Uncertainty: discuss, but do not quantify parameter uncertainty~~
     3. ~~Tested outside training period:~~
     4. ~~Table 1 includes more studies~~
     5. ~~Use: basic~~
  4. Woodbury et al. 1998 Assessing potential climate change effects on loblolly pine growth: A probabilistic regional modeling approach
     1. Model: empirical, probabilistic growth model
     2. Uncertainty:
     3. Tested outside training period:
     4. Use:
  5. Thuiller et al 2004, Effects of restricting environmental range of data to project current and future species distributions: uses GAM for projections of species distributions
     1. Model: empirical
     2. Uncertainty: yes, probabilistic forecasts
     3. Tested outside training period: yes, calibrated on 70% of data, and projected on 30% of data
     4. Use: basic
  6. ~~Araujo, Thuiller, and Pearson 2006, Climate warming and the decline of amphibians and reptiles in Europe: uses multiple empirical model approaches to develop species distributions (70% of data for calibration, 30% for validation) and then develops projections using climate scenarios~~
     1. ~~Model: multiple empirical models (GLinearM, GAM, regression trees, ANN)~~
     2. ~~Uncertainty: discuss model uncertainty which is their reasoning for using so many models to produce so many projections, but do not produce probabilistic forecasts~~
     3. ~~Tested outside training period: yes, 70% calibration, 30% validation, then projections~~
     4. ~~Use: basic~~
  7. White and Nemani 2004: mechanistic model framework to forecast soil ecology
     1. Model: mechanistic, TOPS
     2. Uncertainty: yes, They identify meteorological forecast error and the amount of error required to induce statistically different outputs
     3. Tested outside training period: validated from 1982-1997, then ran the model on different input conditions over the same time period (I think?) to see how much they varied from the control (no change in inputs)
     4. Use: basic
  8. Araujo et al 2005, Reducing uncertainty in projections of extinction risk from climate change: uses multiple empirical model approaches to model species ranges and compares to observed and makes projections
     1. Model: empirical models
     2. Uncertainty: sort of? Uncertainty is looked at in terms of the variability over the different projections and is ‘reduced’ through grouping of the various model outputs
     3. Tested outside training period: yes, 70% calibration and 30% validation; also looks at multiple time periods
     4. Uses ‘ensemble forecasting’ deciding the best forecast based on the likelihood of a number of forecasts
     5. Use: basic
  9. Martinez-Meyer et al 2004, Ecological niches as stable distributional constraints on mammal species , with implications for Pleistocene extinctions and climate change projections for biodiversity: empirical models for current species niches and validation on past data (Pleistocene distributions)
     1. Model: empirical models, GARP
     2. ~~Uncertainty: no?~~
     3. Tested outside training period: yes calibrated using current data and validated during Pleistocene times
     4. Use: basic
  10. Thomas et al 2018, The predictability of a lake phytoplankton community , over time-scales of hours to years:
      1. Model: machine learning, random forests
      2. Uncertainty: sort of? They assess model performance with different parameters and input variables, so the uncertainty between these models is quantified? “Our approach quantifies the decline in predictability with increasing time lag, identifies the predictors that contribute to predictive power and points towards realistic trade-offs and parameterisations through the examination of partial effects.”
      3. Tested outside training period: OOB (out of bag) prediction, so the forests are built with only a subset of the data and can therefore be ‘validated’ using the data not included
      4. “as time lag increased, including environmental predictors led to larger improvements in predictability” over including just the AR term
      5. Use: basic
  11. Perretti et al 2013, Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data: kinda confused, they used ‘mechanistic control models’ to produce a simulated time series?
      1. Model: use several empirical model approaches, but still confused about their mention of using mechanistic models to produce time series? I guess this is just the training/validation data and it isn’t real but that doesn’t matter?
      2. Uncertainty: YES process noise and observation error! This paper does an awesome job of talking and quantifying uncertainty!!
      3. Tested outside training period: yes, 50-yr training period and 50-yr validation period
      4. **Take home:** the empirical models did better than mechanistic
      5. **Use:** basic
  12. ~~Liu et al 2006 Overview of NOAA Coral Reef Watch Program’s Near-Real- Time Satellite Global Coral Bleaching Monitoring Activities~~
      1. ~~Model: empirical, based on historical data~~
      2. ~~Uncertainty: yes~~
      3. ~~Tested outside of training period: yes, updated regularly~~
      4. ~~Use: applied~~
      5. Scratched this one because they’re not actually making predictions yet
  13. Stow et al 2003, Comparison of Estuarine Water Quality Models for Total Maximum Daily Load Development in Neuse River Estuary
      1. Model: 2 process-based models and one probabilistic Bayesian model (empirical)
      2. Uncertainty: yes
      3. Tested outside training period: yes, trained with data pre-2000 and then validated on 2000 estuarine chlorophyll data
      4. Summary: used both process-based and empirical models to predict riverine chl at various spatial scales, but they report that none do well. Their model assessment metrics are a little hard to compare because they don’t include all of the usual metrics (r2) but their RMSE values actually look pretty good to me
      5. Use: applied
  14. Brown et al 2013, Ecological forecasting in Chesapeake Bay: Using a mechanistic–empirical modeling approach
      1. Model: mechanistic and empirical
      2. Uncertainty: yes! They have a whole distribution of probabilities for species distributions over a geographic range, not as much discussion on uncertainty partitioning but the empirical forecasts are probabilistic
      3. Tested outside training period: yes, because the model is tested for accuracy each time a prediction is made and the future becomes present
      4. Summary: uses a mechanistic model to make predictions of multiple variation (water temps, chla, zoop, etc.) and then feeds that output into an empirical model to predict species relative abundance for some habs and other nuisance species
      5. Use: applied
  15. Dean et al 2004, Forecasting landscape-level carbon sequestration using gridded, spatially adjusted tree growth
      1. Model: process-based, CAR4D
      2. Uncertainty: forecasts are probabilistic
      3. Tested outside training period:
      4. Use: basic
  16. Gonzalez-Benecke et al. 2017, Pinus taeda forest growth predictions in the 21st century vary with site mean annual temperature and site quality
      1. Model: process-based, forest simulation model 3-PG
      2. Uncertainty: climate model uncertainty only (driver uncertainty?)
      3. Tested outside training period: use multiple scenarios
      4. Use: basic
  17. Hazen et al 2017, WhaleWatch : a dynamic management tool for predicting blue whale density in the California Current
      + Model: multiple empirical models
      + Uncertainty: yes? They mention uncertainty
      + Tested outside training period: yes, models developed and then 8-day forecasts made which were then validated as time passed
      + Use: applied
  + Seeing lots more empirical studies, not sure if 15 is enough?

Payne et al 2017, Lessons from the First Generation of Marine Ecological Forecast Products

* + 1. More of a review of marine ‘forecasting’
    2. Table 1 includes other forecast studies and what type of model they used