Update onCross Validation Papers

Laurence Kell, Rishi Sharms

June 21, 2018

Short and Long Term Prediction Skill

- Use cross validation to evaluate overfitting and hindcasting to evaluate prediction skill for emprical and model based Management Procedures and Operating Models conditioned using a stock assessment paradigm.

- When conducting Management Strategy Evaluation the stock assessment in the Management Procedure is intended to provide advice in the short term, while the Operating Model is designed to characterise uncertainty in the long term. In the former case there is requirement for good short term prediction skill and in the later case, when conditioning an Operating Model using a stock assessment paradigm, good long term prediction skill.

- Of particular concern when conditioining an Operating Model is that the past may not represent the future. To date two ways have been used to deal with this: either expand the level of uncertainty moving forwards compared to the past, or to invoke Exceptional Circumstances and revise the Management Procedure if future data turn out to be outside the range considered in the trials upon which the Management Procedure had been selected. In the former if quotas are set based on risk then this will result in a reduction in yield while in the later there is a danger of invoking Exceptional Circumstances when the Management Procedure produces an unpalatable result and also negating one of the reasons for conducting Management Strategy Evaluation which is to reduce the frequency of stock assessments.

- Poor prediction skill can result from overfitting where the model is able to describe the past but has poor ability in forecasting. Therefore a more rigorous way is to perform cross validation to test the prediction skill of a model using a set of data not used when fitting. There is often insufficient data, however, in stock assessment datasets to allow some of it to be kept back for testing. A more sophisticated way to create test datasets is, like the jackknife, to leave out one (or more) observation at a time. Cross validation then allows prediction residuals to be calculated, i.e. the difference between fitted and predicted values where the later is calculated from the out-of-sample predictions. A comparison of the variance of the model and prediction residuals can be used to identify over fitting.

- Prediction skill in the future can be evaluated by use of a hindcast where a model is fitted to the first part of a time series and then projected over

the period omitted in the original fit. Prediction skill is then evaluated by comparing the projections $(\hat{Y}_t)$ with observations (Y_t) at time t for a given horizon (h) using the Mean Absolute Scaled Error

$$MASE = \frac{\sum_{t=1}^{T}|Y_t - \hat{Y}_t|}{\frac{T}{T-h}\sum_{t=h}^{T}|Y_t - Y_{t-h}|}$$

- Compare model based and emprirical procedures for short-term forcasts for use in Management Procedures

  Model based Use a biomass dynamic assessment model as the stock assessment

  Emprirical A variety of approaches have been used for short term forecasting, methods include Support Vector Regression (SVR), an autoregressive integrated moving average (ARIMA) model and a multivariate adaptive regression spline (MARS).

To Do

Rishi

- Set up SS grid for North Atlantic swordfish Operating Model.
- Implement and run hindcast and cross-validation for Operating Model.

Laurie

- Set up empirical and model based Management Procedures
- Perform a cross test using Operating Model grid, i.e. run OEM without feedback and estimate population status for 1,2,...n years

Laurie & Rishi

- Summarise prediction skill of Operating Model, i.e. how do hindcasts compare with observations
- Summarise prediction skill Management Procedures, i.e. how does estimated stock status from Management Proceedure compare to that from Operating Model