**Brief Description of the Ranking of Hybrid Species Distribution Models by Species**

Dr. Betsy Barber

INRAE

EABX - Écosystèmes aquatiques et changements globaux

50, avenue de Verdun, Gazinet

F-33612 Cestas Cedex

[inrae.fr](https://inrae.fr/)

1. **Objective**

This document provides a ranking of calibrated hybrid species distribution models developed for 11 European diadromous fish species. Models were ranked according to the average of three indices. These indices provide a measure of the quality of the calibrated models developed for each species, which are then used in the models to estimate forward predictions. **It is important to keep in mind that the calibrated models do not currently include a measure of anthropogenic mortality, meaning that predictions of future distributions are made based only on environmental conditions and dispersal dynamics. Future predictions do not take into account population losses due to fisheries mortality, poor dam passage, habitat degradation, or other anthropogenic effects.**

1. **Three indices for ranking calibrated species models**

Three separate indices were used to rank the hybrid species distribution models, each focusing on a different aspect of the modeling. First, the true skill statistics (TSS, Allouche et al. 2006) was calculated for each species (Table 1). This index measures the accuracy of each calibrated model by comparing the model accuracy to how well the model performs by chance. This index can range from -1 to +1, with a value above 0 indicating that the model is performing better than random and a value of +1 indicating perfect accuracy. TSS is calculated as follows:

where a = true presence, b = false presence, c = false absence, and d = true absence.

In the context of our hybrid models, TSS measures how well the calibrated model for each species matches the actual presence and absence data based on the environmental predictor variables. Briefly, we used boosted regression trees to relate both the physical characteristics of the basins as well as environmental predictor variables (averaged from 1901-1911 and including sea surface temperature, continental precipitation, ocean salinity, and mixed layer depth) to the recorded presence and absence of each species within the same time period (R, package gbm). This analysis produced an estimate of habitat suitability for each river basin that ranges between 0 (not suitable) and 1 (perfectly suitable), which was converted back to presence and absence using the threshold value that produces the most accurate results (R, evaluate function in package dismo). The estimated presence/absence data was then compared to the actual data to determine how accurately the calibrated model performs. The calculated value for TSS was than ranked according to the scale established in Landis and Koch 1977 (Table 2).

**Table 1**. Ranking of species models according to three indices. Average ranking is average of TSS Rank, Survey Index, and Modeler Ranking. Smaller values indicates a “better” ranking across all three indices.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Species** | **TSS Index** | **TSS Rank** | **Survey Index** | **Modeler Ranking** | **Average Ranking** |
| Salmon | 0,950 | 1 | 0,723 | 1 | 0,908 |
| Allis Shad | 0,776 | 2 | 0,946 | 2 | 1,649 |
| Twaite Shad | 0,760 | 2 | 0,973 | 2 | 1,658 |
| Flounder | 0,878 | 1 | 1,226 | 3 | 1,742 |
| Sea Trout | 0,697 | 2 | 0,668 | 3 | 1,889 |
| Smelt | 0,885 | 1 | 1,693 | 3 | 1,898 |
| Sea Lamprey | 0,876 | 1 | 0,733 | 4 | 1,911 |
| Mullet | 0,847 | 1 | 2,185 | 3 | 2,062 |
| Eels | 0,919 | 1 | 0,810 | 5 | 2,270 |
| River Lamprey | 0,874 | 1 | 2,134 | 4 | 2,378 |
| Sturgeon | 0,722 | 2 | 3,277 | 3 | 2,759 |

**Table 2**. TSS value, ranking, and interpretation according to Landis and Koch 1977.

|  |  |  |
| --- | --- | --- |
| **TSS value** | **Ranking** | **Interpretation** |
| < 0 | 6 | no agreement |
| 0 to 0.2 | 5 | slight |
| 0.2 to 0.4 | 4 | fair |
| 0.4 to 0.6 | 3 | moderate |
| 0.6 to 0.8 | 2 | substantial |
| 0.8 to 1 | 1 | perfect |

Second, an index was calculated to incorporate the level of uncertainty in the estimate of the model parameters for each species. Eight model parameters were estimating using a survey of European diadromous species experts. For each parameter, a weighted mean and standard deviation was calculated based on group answers. Each participant indicated a level of confidence in their answer for each model parameter and each species, which was used as a weight in the final estimates. To calculate the survey index used for ranking the species in Table 1, the weighted estimate of uncertainty (upper SD – lower SD) was divided by the weighted mean for each parameter estimated in the survey. This value was then divided by the number of participants who provided a value for each parameter and species. To calculate one value per species, the index was summed across all parameters. To interpret this index, a larger number indicates a higher level of uncertainty for that species across all survey questions.

Last, a subjective ranking was performed by the model developers. Models were ranked between 1 (high confidence) and 5 (low confidence). This ranking was primarily based on the performance of the dispersal component of the models given the current parameters and constraints. However, the ranking also included any observations made by the modelers throughout the entire process that may cause lower confidence in the results of a particular species (listed in Table 3).

**Table 3.** Subjective ranking of model performance for each species. Ranking performed by model developers. Models ranked between 1 (high confidence) and 5 (low confidence).

|  |  |  |
| --- | --- | --- |
| **Species** | **Ranking** | **Reasoning** |
| Allis shad | 2 | No obvious problems with dispersal; sensitivity is second lowest, though still “good” (0.82) |
| Eel | 5 | Is a panmictic species and the model equations as developed don't fit this life history strategy well; also raw data for Atlantic Area has presence in all basins, so it is difficult for the model to make absence predictions; Dispersal components of equations are not working correctly |
| Flounder | 3 | Several parameters, including dispersal and proportion of emigrants, are more difficult to estimate for catadromous species; High prevalence of presence in raw data, meaning the model may have a hard time making absence predictions; No obvious problems with dispersal; few experts for parameter estimates |
| Mullet | 3 | Several parameters, including dispersal and proportion of emigrants, are more difficult to estimate for catadromous species; High prevalence of presence in raw data, meaning the model may have a hard time making absence predictions; No obvious problems with dispersal |
| River Lamprey | 4 | Dispersal component of equations is not currently working correctly because of high parameter estimates for proportion of emigrants and emigrant mortality; However, there are methods for improving these estimates. |
| Salmon | 1 | No obvious problems with dispersal; highest predictive accuracy; relatively narrow range of uncertainty in parameter estimates |
| Sea Lamprey | 4 | Dispersal component of equations is not currently working correctly because of high parameter estimates for proportion of emigrants and emigrant mortality; However, there are methods for improving these estimates. |
| Sea Trout | 3 | No obvious problems with dispersal; High prevalence of presence in raw data, meaning the model may have a hard time making absence predictions; specificity is lowest of all models, though still “good” value (0.76) |
| Smelt | 3 | Low prevalence of presence in raw data, meaning the model may have a hard time making presence predictions; no obvious problems with dispersal; predictive accuracy still relatively high; few experts for parameter estimates |
| Sturgeon | 3 | Not as many experts for survey responses; Also lowest sensitivity in calibrated models, though still “good” value (0.78) |
| Twaite shad | 2 | No obvious problems with dispersal; second lowest specificity, though still “good” value (0.82); highest prediction error, though still relatively low (0.12) |

In the final ranking, the average value of these three indices was calculated. Species were then organized in Table 1 according to this Average Ranking, with a smaller value indicating a “better” ranking across the three indices.

1. **Rough estimate of predictive accuracy**

Predictive accuracy was calculated by comparing the predicted presence/absence to the raw data. This estimate only accounted for the effect of environmental predictors on species presence within a basin, and did not take into account the dispersal component of the hybrid models. Also, this estimate of predictive accuracy did not take into account the estimated abundance of a population, only the presence or absence within a river basin.

Annual values of habitat suitability were estimated for each species and river basin from 1950-2019 for three different climate models using the gbm.predict function from the gbm package in R. This method uses the relationship between species presence/absence and environmental predictor variables identified in the calibration of the model to predict future habitat suitability as a value between 0 (not suitable) and 1 (perfectly suitable). For each year, average habitat suitability was calculated for the three climate models in each river basin, and then converted to presence/absence according to the species-specific threshold values as explained in Section 2. Annual values of TSS were calculated by comparing the modeled presence/absence to the raw data for the time period. Annual TSS values were averaged to get the final Predictive TSS value (Table 4) used to rank the species according to the criteria in Table 2.

**Table 4.** Rough estimate of model predictive accuracy, based solely on comparing predicted habitat suitability from 1950-2019 to raw presence/absence data from this time period, and not taking into account anthropogenic effects (i.e. fisheries mortality, poor dam passage, pollution, etc).

|  |  |  |
| --- | --- | --- |
| **Species** | **Predictive TSS** | **Rank** |
| Sea Trout | 0,945 | 1 |
| Eel | 1,000 | 1 |
| Flounder | 0,478 | 3 |
| Mullet | 0,537 | 3 |
| Sturgeon | 0,517 | 3 |
| River Lamprey | 0,560 | 3 |
| Smelt | 0,325 | 4 |
| Allis shad | 0,175 | 5 |
| Twaite shad | 0,161 | 5 |
| Sea Lamprey | 0,162 | 5 |
| Salmon | 0,190 | 5 |

**References:**

Allouche O, Tsoar A, and Kadmon R. 2006. Assessing the accuracy of species distribution models: prevalence, kappa, and the true skill statistic (TSS). Journal of Applied Ecology. 43: 1223-1232.

Landis JR and Koch GG. 1977. The measurement of observer agreement for categorical data. Biometrics. 33: 159-174.