Dynamics of Two Anadromous Species in a Dam Intersected River: Analysis of Two 100-Year Datasets

Ulisses M Azeiteiro, Mário J Pereira, Amadeu MVM Soares, Heitor O. Braga, Fernando Morgado, Magda C. Sousa, João M. Dias and Carlos Antunes

Supplementary Material

Correlation between environmental variables

This analysis showed only one significant correlation between average sea surface temperature and average air temperature (Table S1).

Table S1. Correlation coefficients between the possible pairs of environmental variables.

CORRELATION COEFFICIENT	SST	air TºC	NAO	Precipitation
SST				
air TºC	0.57			
NAO	-0.01	-0.09		
Precipitation	0.04	-0.11	0.09	
P-VALUE	**SST**	**air TºC**	**NAO**	**Precipitation**
SST				
air TºC	0.0000			
NAO	0.8805	0.3803		
Precipitation	0.7140	0.2593	0.3380	

Time series analysis of fish captures

Analysis of the behaviour of fish capture data as a time series was identified using Differencing Methods. This resulted in a decomposition of the time series between a trend and random error (as no seasonal effects were recorded; Figures S1 and S2).

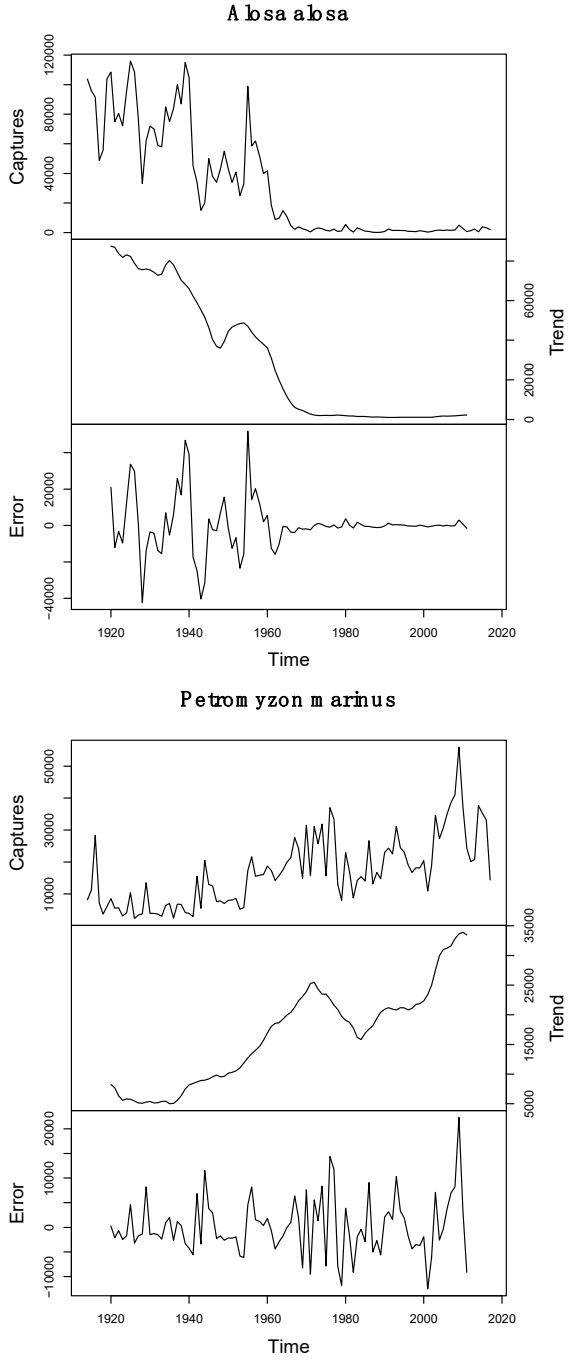

Figure S1. Decomposition of the time series data of A) *Alosa alosa* and B) *Petromyzon marinus*. The top panel shows original captures, the middle panel shows the resulting trend, and the bottom panel shows the random error.

Multiple linear regression of the fish capture data

Linear models were fitted for *A. alosa* and *P. marinus* captures using precipitation, air temperature, NAO, and upwelling index using only the years for which all variables had data (intersect data). Model results are presented in Tables S2 and S3.

Table S2. Model results of the Linear Regression for *Alosa alosa* captures with environmental variables as regressors (intersect data).

Coefficients	Estimate	Standard Error	t-value	P-value

Alosa alosa captures (units)	3040.884	4419.357	0.688	0.498
Monthly Precipitation	6.950	6.631	1.048	0.305
Monthly Air Temperature	-170.248	296.022	-0.575	0.571
North Atlantic Oscillation	-240.427	452.217	-0.532	0.600
Upwelling Index	2.355	1.956	1.204	0.240
Residual Standard Error: 1269 on 24 Degrees of Freedom (DF) ǀ Multiple R^2: 0.1031 ǀ Adjusted R^2: -0.04643 ǀ F-statistic: 0.6894 on 4 and 24 DF, P-value: 0.6064				

Table S3. Model results of the Linear Regression for *Petromyzon marinus* captures with environmental variables as regressors (intersect data).

Coefficients	Estimate	Standard Error	t-value	P-value
Petromyzon marinus captures (units)	8614.60	27646.93	0.312	0.7580
Monthly Precipitation	74.59	41.48	1.798	0.0847
Monthly Air Temperature	181.56	1851.88	0.098	0.9227
North Atlantic Oscillation	1037.57	2829.01	0.367	0.7170
Upwelling Index	17.04	12.24	1.392	0.1767
Residual Standard Error: 7940 on 24 Degrees of Freedom (DF) ǀ Multiple R^2: 0.1971 ǀ Adjusted R^2: 0.06331 ǀ F-statistic: 1.473 on 4 and 24 DF, P-value: 0.2415				

Linear models were also fitted for *A. alosa* and *P. marinus* captures using precipitation, air temperature, NAO, and upwelling index using only the variables for which data were available for the entire time series. Model results are presented in Tables S4 and S5.

Table S4. Model results of the Linear Regression for *Alosa alosa* captures with environmental variables as regressors.

Coefficients	Estimate	Standard Error	t-value	P-value
Alosa alosa captures (units)	104335.60	65870.29	1584	0.116
Monthly Average Precipitation	-62.39	107.66	-0.580	0.564
Monthly Air Temperature	-4643.66	4454.50	-1.042	0.300
North Atlantic Oscillation	-4345.54	5862.32	-0.741	0.460
Residual Standard Error: 37090 on 100 Degrees of Freedom (DF) ǀ Multiple R^2: 0.01784 ǀ Adjusted R^2: -0.01163 ǀ F-statistic: 0.6054 On 3 And 100 DF, P-value: 0.613				

Table S5. Model results of the Linear Regression for *Petromyzon marinus* captures with environmental variables as regressors.

Coefficients	Estimate	Standard Error	t-value	P-value
Petromyzon marinus captures (units)	-3656.54	19486.53	-0.188	0.852
Monthly Average Precipitation	22.73	31.85	0.714	0.477
Monthly Air Temperature	1270.72	1317.78	0.964	0.337
North Atlantic Oscillation	-451.32	1734.26	-0.260	0.795
Residual Standard Error: 10970 on 100 Degrees Of Freedom (DF) ǀ Multiple R^2: 0.01369 ǀ Adjusted R^2: -0.0159 ǀ F-statistic: 0.4625 On 3 And 100 DF, P-value: 0.7091				

Intervention analysis

Preliminary ARIMA models were fitted with the data transformations that presented the best results for each species dataset. Accuracies of these models are presented in Tables S6 and S7.

Table S6. Performance measurements of ARIMA models with the original, log, and Δlog transformed *Alosa alosa* capture data. The model with log transformation performed better.

Transformation	ME	RMSE	MAE	ACF1	AIC	AICc	BIC
log	-0.06710364	0.3378423	0.2432781	-0.08379447	34.75	35.41	39.81
Δlog	-0.06907031	0.3420343	0.2490714	-0.08410868	34.75	35.41	39.81

ME, Mean Error; RMSE, Root Mean Squared Error; MAE, Mean Absolute Error; ACF1, Autocorrelation of errors at lag 1; AIC, Akaike Information Criterion; AICc, Corrected Akaike Information Criterion; BIC, Bayesian Information Criterion

Table S7. Performance measurements of ARIMA models with the original, log, Δlog, and Δlog12 transformed *Petromyzon marinus* capture data. The model with Δlog performed better.

Transformation	ME	RMSE	MAE	ACF1	AIC	AICc	BIC
log	-4.887146e-14	0.5650506	0.4506777	0.221383	73.54	73.86	76.97
Δlog	-0.03174938	0.5689634	0.4524447	0.003259193	73.23	73.56	76.61
Δlog12	0.186168	0.8858761	0.6510311	-0.07279541	78.49	78.97	81.15

ME, Mean Error; RMSE, Root Mean Squared Error; MAE, Mean Absolute Error; ACF1, Autocorrelation of errors at lag 1; AIC, Akaike Information Criterion; AICc, Corrected Akaike Information Criterion; BIC, Bayesian Information Criterion

Although the previous analyses indicated that models with log and Δlog transformation would perform better, ARIMA models were fitted with all data transformations. The model with the data transformation presenting the lowest AIC was chosen as the best to explain the pre-intervention capture data (Tables S8 and S9). For *A. alosa* captures, the best ARIMA model was ARIMA(0,0,2) with non-zero mean and using Δlog transformed data. For *P. marinus* captures, the best model was ARIMA(0,0,0) with non-zero mean and using log transformed data.

Table S8. ARIMA models for pre-intervention (1914-1954) capture data of *Alosa alosa*. Two models were tested, with log and Δlog transformation, as they had previously shown very small differences.

Model	ma1	ma2	sma1	mean	AIC	AICc	BIC
log [ARIMA(0,1,2)]	0.2095	-0.7905	-0.1356	n.a.	35.99	37.13	42.75
Δlog [ARIMA(0,0,2) with non-zero mean]	0.0000	-1.0000	-0.0946	-0.0261	34	35.76	42.44

ma1, ma2, sma1, and mean are ARIMA model coefficients; AIC, Akaike Information Criterion; AICc, Corrected Akaike Information Criterion; BIC, Bayesian Information Criterion

Table S9. ARIMA models for pre-intervention (1914-1954) capture data of *Petromyzon marinus*. Three models were tested, with log, Δlog, and Δlog12 transformation, as they had previously shown very small differences.

Model	ma1	sma1	mean	AIC	AICc	BIC
log [ARIMA(0,0,0) with non-zero mean]	n.a.	0.1772	8.7415	73.9	74.55	79.04
Δlog [ARIMA(0,0,1) with non-zero mean]	-0.4133	-0.4133	-0.0133	78.35	79.49	85.1
Δlog12 [ARIMA(0,1,1)]	-0.4261	-0.4261	n.a.	81.5	82.5	85.5

ma1, ma2, sma1, and mean are ARIMA model coefficients; AIC, Akaike Information Criterion; AICc, Corrected Akaike Information Criterion; BIC, Bayesian Information Criterion

It was also investigated whether there were any structural breaks with respect to a constant level as regressor. The structural breaks found for each time series are reported in Figure S4. These breakpoints indicate dates where a specific and abrupt variation in the data was observed.

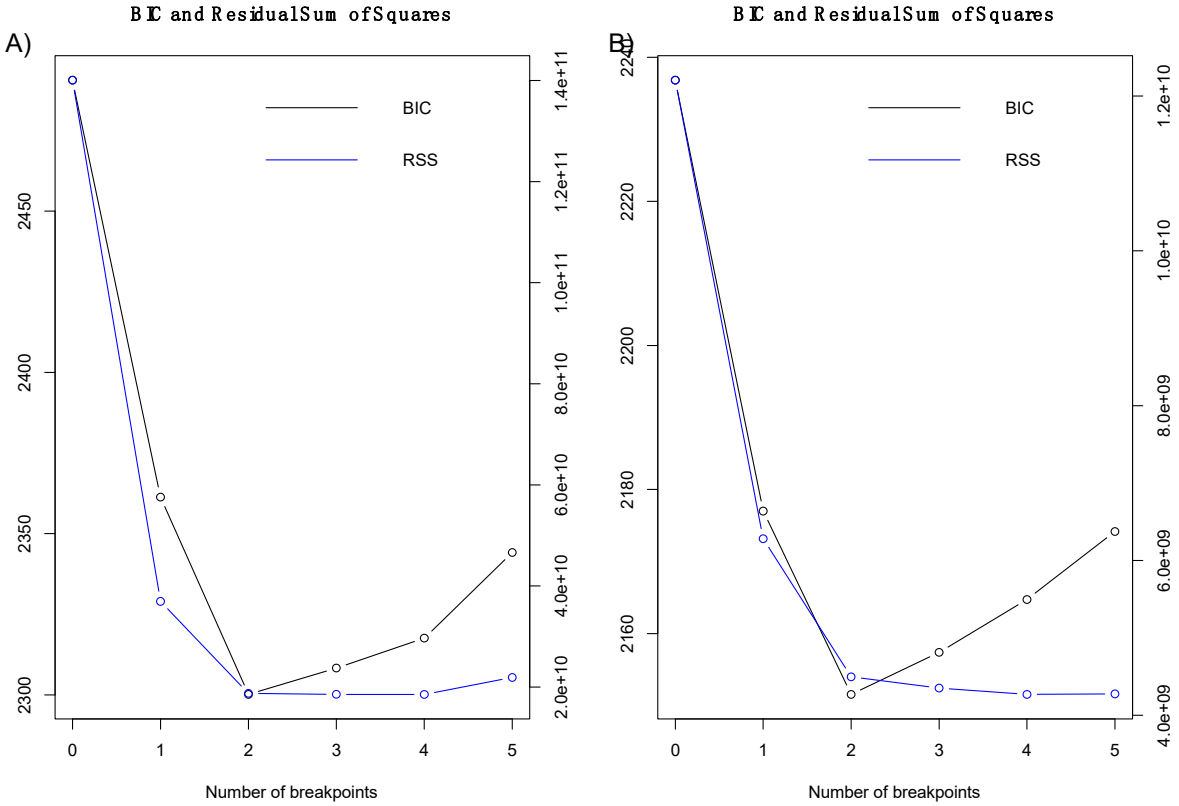

Figure S2. Plot of the Bayesian Inference Coefficient (BIC) value against the number of breakpoints for each species: *Alosa alosa* (A) and *Petromyzon marinus* (B) capture data.