Natural and anthropogenic effects on the early life stages of European anchovy in one of its essential fish habitats, the Guadalquivir estuary

Gustavo F. de Carvalho-Souza, Enrique González-Ortegón, Francisco Baldó, César Vilas, Pilar Drake, Marcos Llope

*Corresponding author: marcos.llope@ieo.es, marcos.llope@ibv.uio.no

Marine Ecology Progress Series https://doi.org/10.3354/meps12562

Supplement

Additional information on datasets and variables as well model outputs and diagnostics.

Supplementary Methods

Regression assumptions: Independence

Some models (eg. water mass I) showed autocorrelation in the residuals. To assess the importance of violating this regression assumption we created five alternative models or sub-models (Figure S1). Each of the models accounted for this problem in a different way: (i) alternative model 1 included 'month' as a temporal factor, (ii) sub-model 2 included the response variable (anchovy abundance) as an explanatory term lagged by one month, (iii) in sub-model 3 the trigonometric functions $\sin(2\pi t/12)$ and $\cos(2\pi t/12)$ were added as covariates, where t is the monthly index (Stenseth et al. 2006), (iv) model 4 used 'month' again, but this time entered with cubic spline basis, and finally (v) model 5 formulation incorporated an autoregressive term in a mixed model framework (GAMM) (Wood 2006, Hefley et al. 2017). All sub-models showed non-correlated residuals (Figure S2) and therefore corrected for the lack of independence. Some of the sub-models (models 1 and 3, for example) did not retained some of the covariates (freshwater discharges and *M. slabberi*) while models 2, 4 and 5 kept practically the same structure. All things considered we can conclude that autocorrelation, when present, does not affect the main results.

Model validation

Overall, model predictions matched well the observed anchovy densities although some extreme values were not efficiently captured (Fig. S11a). However, it is worth noticing that the models captured well the detrimental effects that the events of high and persistent turbidity (*sensu* González-Ortegón et al. 2010) and high freshwater inputs had on the abundance of the early stages of this fish. Predictions for *M. slabberi* also reproduced the observed dynamics fairly well (Fig. S12a). The seasonal cycle was well resolved, particularly for the summer months when their abundances are at their highest (Figs. S11b & S12b).

Figure S1. Sub-models 1-5 partial plots showing the partial effect of each predictor on the response variable. The R-squared (R^2), generalized cross validation (GCV), deviance explained (Dev. Expl.) and the number of observations (*n*) for the model are presented to the right of the plots.

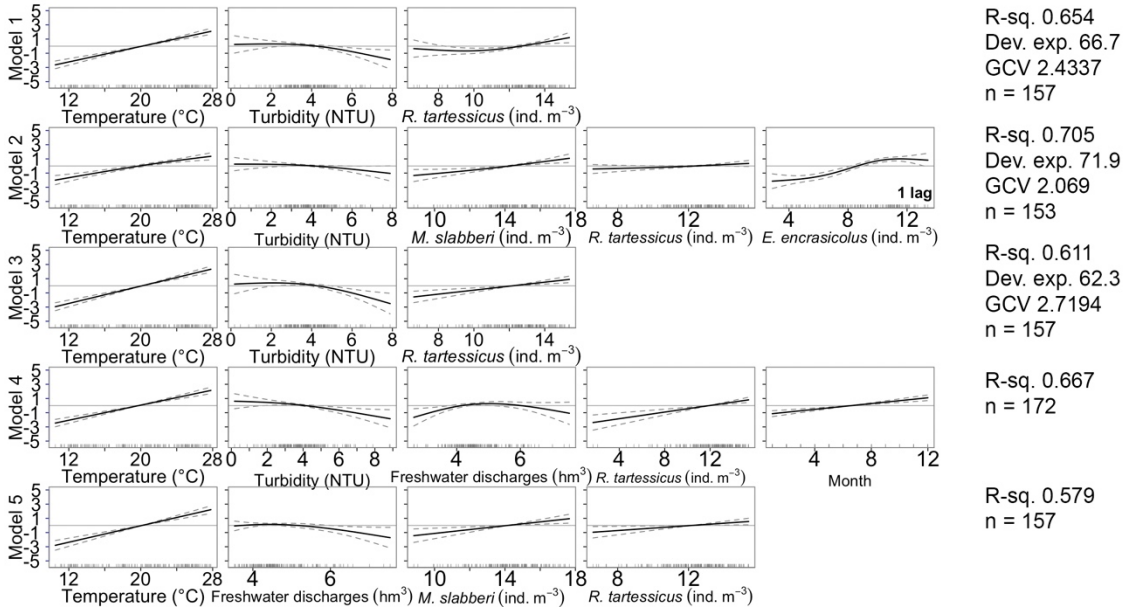
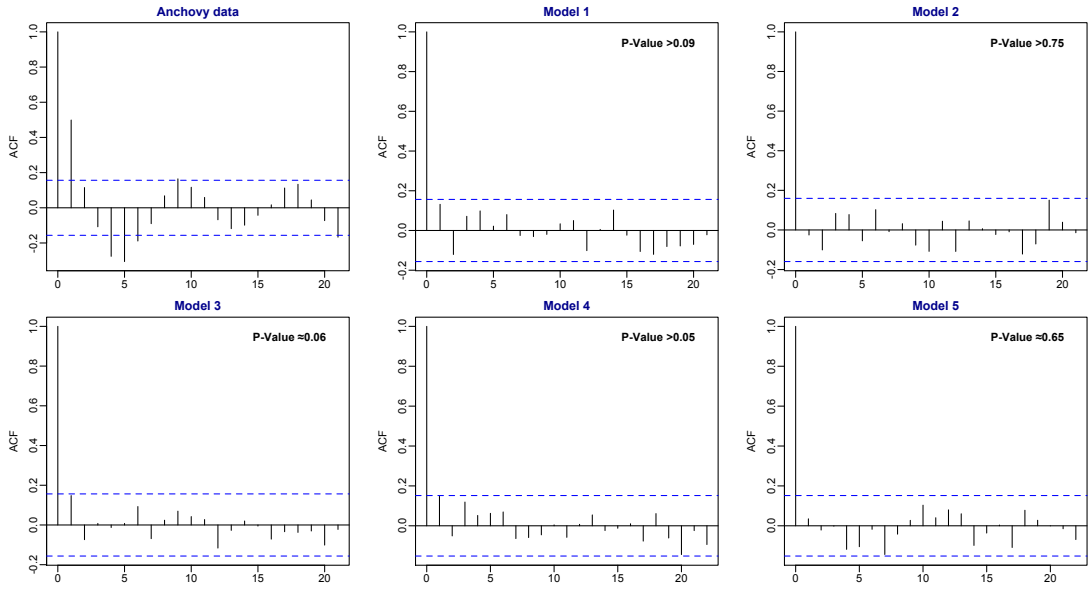


Figure S2. Residual autocorrelation plots corresponding to the five alternative model formulations (Fig. S1a). The first plot shows the structure of the response variable (anchovy density).



Supplementary References

Hefley, TJ, Broms, KM, Brost, BM, Buderman, FE, Kay, SL, Scharf, HR, Tipton, JR, Williams, PJ Hooten, MB (2017) The basis function approach for modeling autocorrelation in ecological data. Ecology 98:632–646

Stenseth NC, Llope M, Anadón R, Ciannelli L, Chan K-S, Hjermann DØ, Bagøien E, Ottersen, G (2006) Seasonal plankton dynamics along a cross-shelf gradient. Proc R Soc Lond B 273:2831–2838

Supplementary Tables

Table S1. List of predictors used in the models: acronyms, definition and source. Long-term Program = Guadalquivir Monitoring program; AEMET = Agencia Española de Metereologia (http://www.aemet.es/es/eltiempo/observacion/ultimosdatos); CHG = Confederación Hidrográfica del Guadalquivir; EBD/CSIC = Estación Biológica de Doñana; NOAA = National Oceanic & Atmospheric Administration.
Monthly Mean North Atlantic Oscillation (NAO) index were obtained from the Climate Prediction Center (NOAA) (http://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml).

Acronym	**Definition**	**Data source**
Temp	Temperature (˚C)	Long-term Program
Sal	Salinity	Long-term Program
Tur	Turbidity (NTU)	Long-term Program
Vol4, Vol7, Vol15, Vol30	Freshwater discharges (hm^3) at 4, 7, 15 and 30 days prior to sampling day	CHG
P4, P7, P15, P30	Precipitation at 4, 7, 15 and 30 days prior to sampling day	EBD/CSIC
Un4, Un7, Un15, Un30	Average velocity (km) of levanters (4, 7, 15 and 30 prior to sampling day)	AEMET
Up4, Up7, Up15, Up30	Average velocity (km) of westerlies (4, 7, 15 and 30 prior to sampling day)	AEMET
D7PW25, D15PW25, D30PW25, D7PW30, D15PW30, D30PW30	Days per month (7, 15 and 30 prior to sampling day) with winds >25 and 30 km	AEMET
NAO	North Atlantic Oscillation Index	NOAA
AncA	Anchovy abundance (ind. m^{-3})	Long-term Program
Mesop	Mesopodopsys slabberi abundance (ind. m^{-3})	Long-term Program
Neom	*Neomysis integer* abundance (ind. m^{-3})	Long-term Program
Rhop	*Rhopalophthalmus tartessicus* abundance (ind. m^{-3})	Long-term Program

Table S2. Variance-inflation factor (VIF) of the significant covariates in each selected model. Acronyms for the predictors are indicated in Table S1 and for the models are described as follows: I – Water mass I models (Bonanza station at ebb tide); II – water mass II models (Bonanza station at flood tide); III – water mass III models (Tarfía station at ebb tide); IV – water mass IV models (Tarfía station at flood tide).

Predictor	***Engraulis encrasicolus***				***Mesopodopsis slabberi***				***Neomysis integer***				***Rhopalophthalmus tartessicus***			
	I	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**	**I**	**II**	**III**	**IV**
Temp	1.36	1.25	1.61	-	1.37	1.75	1.51	1.52	-	-	1.14	1.13	1.12	-	-	-
Sal	-	-	1.66	-	1.86	1.77	1.54	1.27	-	-	1.30	1.12	1.08	1.04	1.41	1.41
Tur	1.04	-	1.23	1.03	1.03	1.10	1.89	-	1.01	1.00	-	-	-	1.01	-	-
Vol4	-	-	-	-	1.75	1.56	1.94	-	-	-	1.17	-	-	-	-	-
Vol7	-	1.06	1.36	-	-	-	-	-	-	-	-	-	-	-	-	-
Vol30	1.03	-	-	-	-	-	-	1.22	1.06	1.01	-	-	-	-	-	-
Un4	-	-	-	1.04	-	-	-	-	-	-	-	-	-	-	-	-
Un7	-	-	-	-	-	-	1.14	-	-	-	-	-	-	-	-	-
Un15	-	-	-	-	-	-	-	1.39	-	-	-	-	-	-	-	-
Up30	-	-	-	-	-	1.34	-	-	-	-	-	-	-	-	-	-
AncA	1.45	1.34	1.23	1.65	-	-	-	-	-	-	-	-	-	-	-	-
AncA1	-	-	-	-	1.19	-	-	-	-	-	-	1.04	-	-	-	-
Mesop	1.21	1.31	1.07	1.64	1.06	1.14	1.08	1.10	-	-	-	-	-	-	-	-
Neom	-	-	-	-	-	-	-	-	1.04	1.00	1.06	1.06	-	-	-	-
Rhop	1.13	1.33	-	-	-	-	-	-	-	-	-	-	1.04	1.05	1.04	1.04

Table S3. Model summary statistics per species and water mass (I-IV). The generalized cross validation (GCV), R-squared (R^2), deviance explained (Dev. Expl.) and number of observations (*n*) are indicated. Model number correspondences (I-IV) are given in Table S2.

Variable Response	GAMs	GCV	R2(adj.)	Dev. Expl. (%)	n
Engraulis encrasicolus	I	2.8704	0.61	63	172
	II	2.5361	0.61	62.9	162
	III	3.7709	0.7	68.3	140
	IV	4.3227	0.46	50.3	103
Mesopodopsis slabberi	I	2.4691	0.49	52.8	171
	II	2.8172	0.70	71.4	190
	III	4.6138	0.68	70.5	165
	IV	6.1983	0.6	62.1	173
Neomysis integer	I	5.5684	0.2	22.9	100
	II	7.0734	0.18	20.2	140
	III	7.1005	0.09	10.8	177
	IV	6.8478	0.11	14.7	111
Rhopalophthalmus tartessicus	I	5.0627	0.095	11	172
	II	5.9878	0.307	32.4	158
	III	7.2323	0.253	27.1	43
	IV	4.4111	0.14	18.1	27

Table S4. Model results for each predictor. Estimated degrees of freedom (edf), significance value (*p-value*), absolute and partial deviance (relative to total deviance) partitioned per predictor (%) are provided. Model number correspondences (I-IV) are given in Table S2.

RV	GAM	factors	edf	F	P-value	% Absolute Deviance	% Partition Deviance
Engraulis encrasicolus	I	Temperature	1	77.71	<0.001	36	57
		Turbidity	1.209	4.70	0.03	8	13
		Fresh. discharges	2.238	3.57	0.01	2	3
		M. slabberi	1	4.19	0.04	11	17
		R. tartessicus	1	7.99	0.005	6	10
	II	Temperature	2.286	13.33	<0.001	24	38
		Fresh. discharges	2.725	2.94	0.02	7	12
		M. slabberi	1	22.47	<0.001	22	34
		R. tartessicus	1	4.59	0.03	10	16
	III	Temperature	1.120	11.14	<0.001	22	32
		Salinity	2.605	3.18	0.04	10	15
		Turbidity	1	4.38	0.03	4	6
		Fresh. discharges	1.737	5.10	0.005	3	5
		M. slabberi	2.519	9.93	<0.001	29	43
	IV	Turbidity	2.901	5.88	0.01	9	17
		M. slabberi	1.731	25.38	<0.001	33	65
		Levanters	2.801	2.96	0.02	9	18
Mesopodopsis slabberi	I	Temperature	2.18	20.42	<0.001	22	43
		Salinity	2.53	3.64	0.03	10	19
		Turbidity	2.04	4.30	0.01	7	13
		Fresh. discharges	2.48	3.82	0.04	12	23
		E. encrasicolus	1	6.15	0.01	1	2
	II	Temperature	1.66	9.42	<0.001	18	26
		Salinity	1.94	2.39	0.01	11	15
		Turbidity	2.69	14.97	<0.001	21	30
		Fresh. discharges	2.64	4.39	0.003	16	22
		Westerlies	1	8.48	0.004	6	7
	III	Temperature	1	84.61	<0.001	33	46
		Salinity	2.53	5.08	0.002	15	22
		Turbidity	2.03	5.88	0.002	8	11
		Fresh. discharges	2.50	4.74	0.004	10	14
		Levanters	1	8.93	0.003	5	7
	IV	Temperature	1	96.43	<0.001	34	54
		Salinity	1.56	4.69	0.009	11	18
		Fresh. discharges	2.73	6.20	<0.001	8	13
		Levanters	1	4.90	0.02	9	15
Neomysis	I	Turbidity	2.06	3.83	0.01	13	57

		Fresh. discharges	1	7.86	0.006	10	43
	II	Turbidity	1.85	5.03	0.05	5	24
		Fresh. discharges	1	12.98	0.004	15	76
	III	Temperature	1	18.69	<0.001	7	62
		Salinity	1	12.00	<0.001	3	28
		Fresh. discharges	1	5.41	0.02	1	9
	IV	Temperature	1.36	6.94	0.008	5	35
		Salinity	1.41	3.14	0.03	4	26
		E. encrasicolus	1	10.86	0.001	6	39
Rhopalophthalmus tartessicus	I	Temperature	1	9.48	0.002	6	53
		Salinity	2.26	2.97	0.04	5	47
	II	Salinity	2.31	14.83	<0.001	22	55
		Turbidity	1.95	10.84	<0.001	18	45
	III	Salinity	1	15.21	<0.001	27	100
	IV	Salinity	1	5.52	0.02	18	100

Table S5. Reference points were calculated based on the effects that turbidity and freshwater discharges had on the abundance of anchovy and *M. slabberi* (Fig. 3b,c,g,l,m,o and Fig. 4c,d,h,i,m,n,r, respectively). The coloured scaled and correspondence between raw vs. logged values is given as follows (approximate values):

	turbidity	log(turbidity)	freshwater	log(freshwater)
	NTU	NTU	hm^3	hm^3
red	>500	>6.2	>1000	>6.9
orange	>300 & <=500	>5.7 & <=6.2	>500 & <=1000	>6.2 & <=6.9
yellow	>150 & <=300	>5.0 & <=5.7	>150 & <=500	6.2> & <=5.0
green	>0 & <=150	>0 & <=5.0	>0 & <=150	>0 & <=5.0

Supplementary Figures

Figure S3. Correlation Matrix of covariates at water mass I. The colour scale indicates Pearson correlation values.

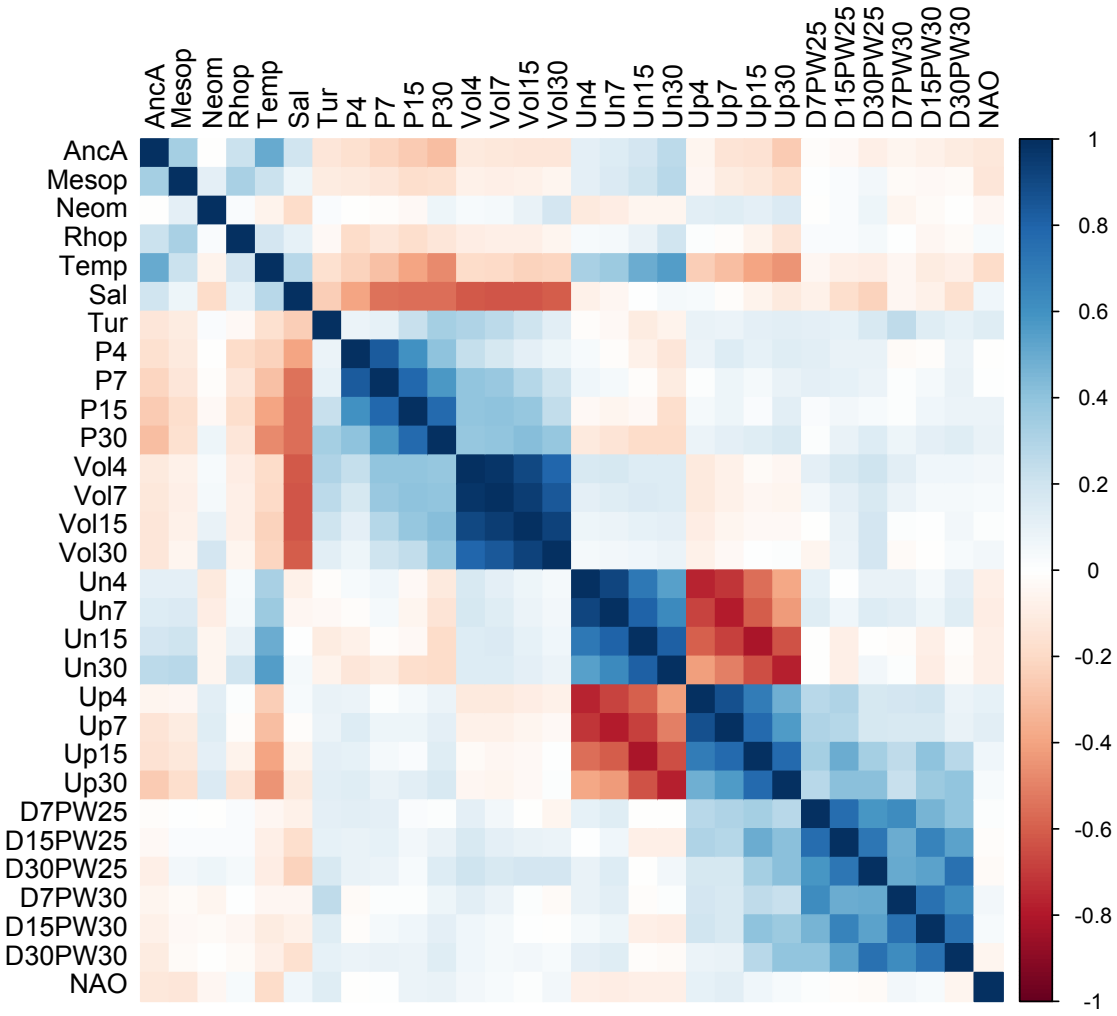


Figure S4. Correlation Matrix of covariates at water mass II. The colour scale indicates Pearson correlation values.

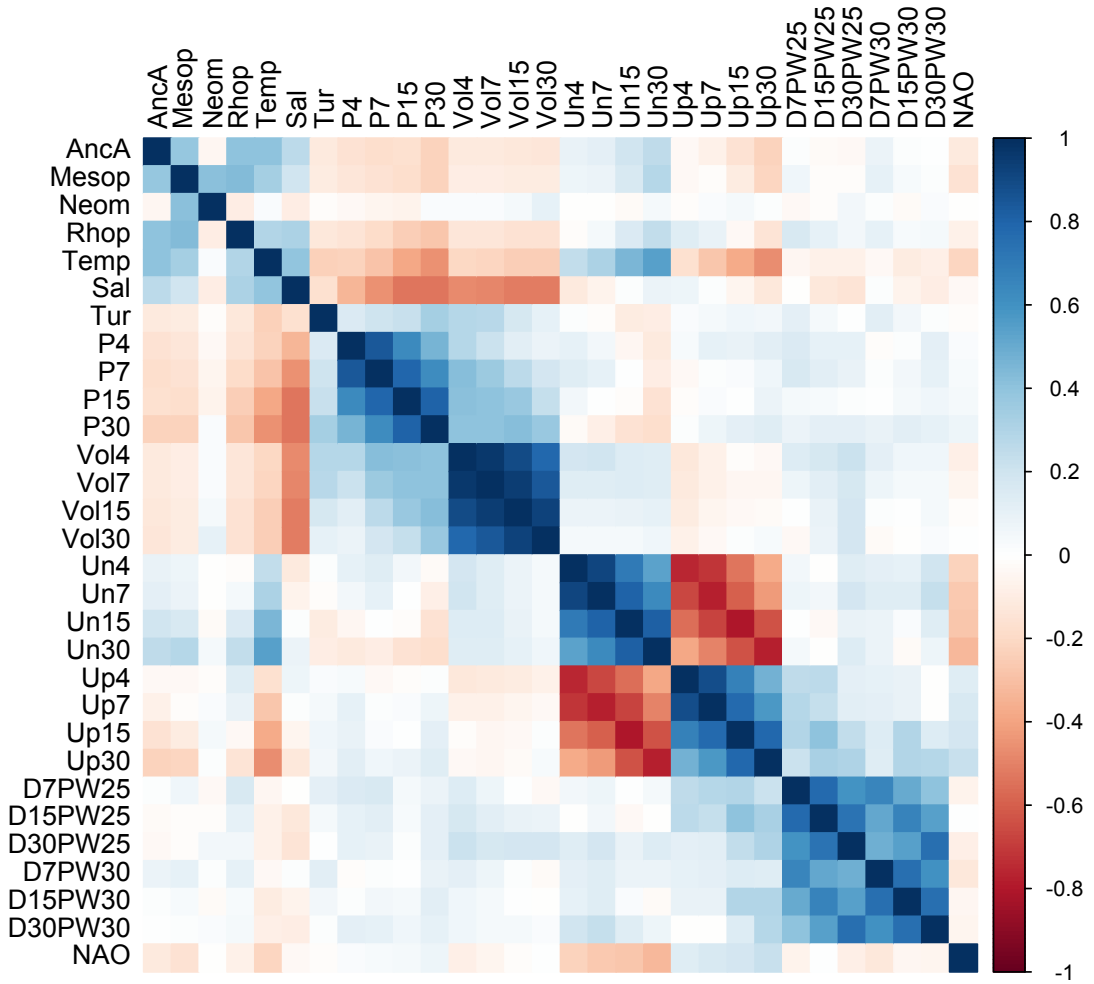


Figure S5 - Correlation Matrix of covariates at water mass III. The color scale indicates Pearson correlation values.

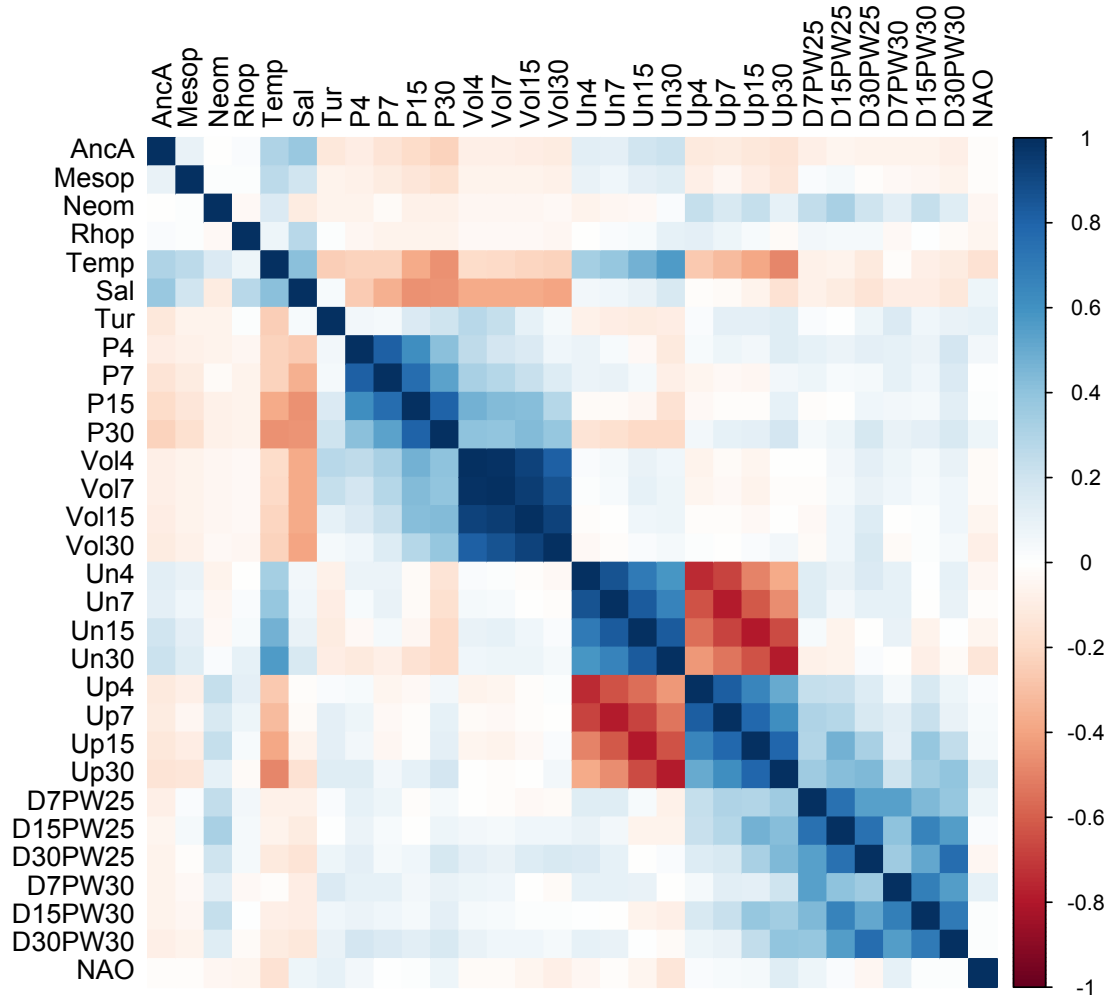


Figure S6 - Correlation Matrix of covariates at water mass IV. The colour scale indicates Pearson correlation values.

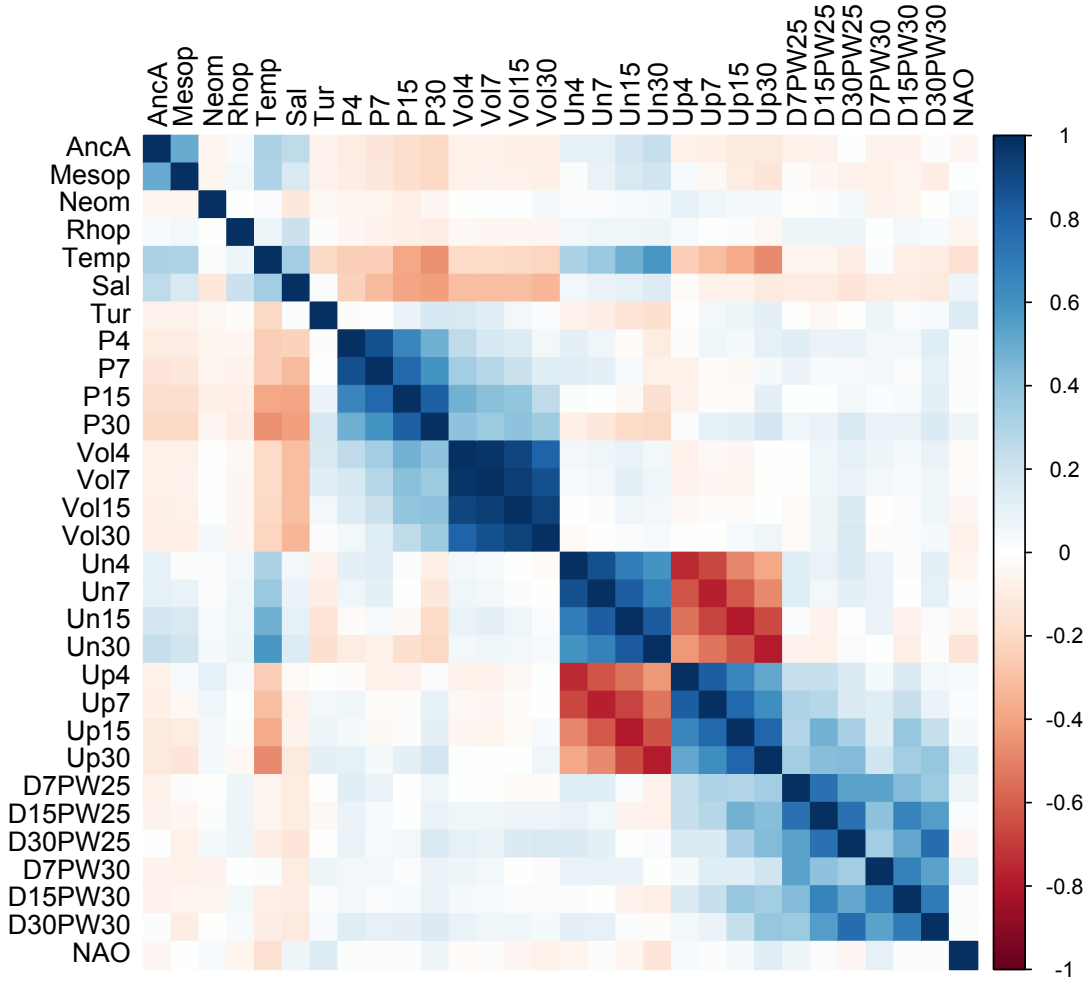


Figure S7. Residuals. The graphs are structured as follows: rows correspond to each of the anchovy individual models. First column shows the autocorrelation within the residuals (correlations are not significant if they lie within the dashed blue lines). Second column assesses homoscedasticity. Third and fourth columns evaluate normality.

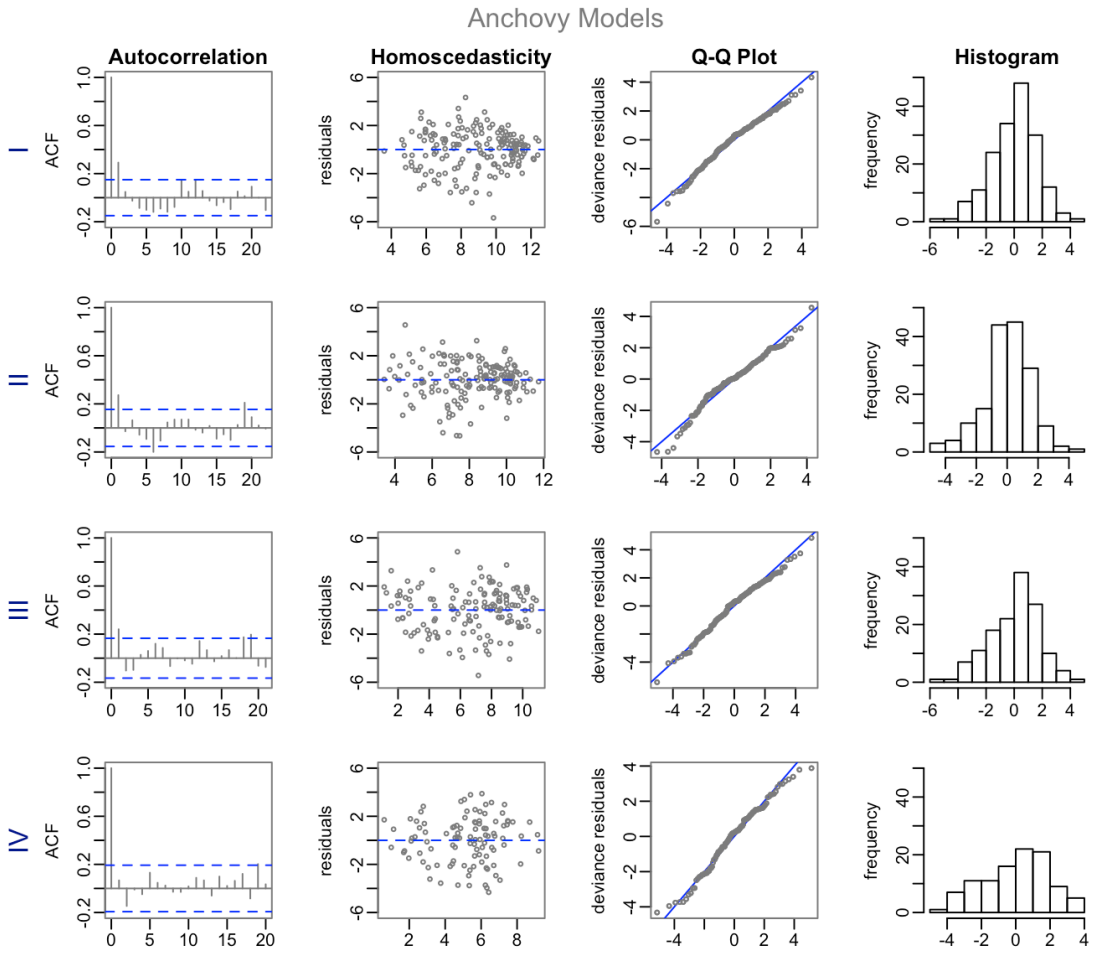


Figure S8. Residuals. The graphs are structured as follows: rows correspond to each of the *M. slabberi* individual models. First column shows the autocorrelation within the residuals (correlations are not significant if they lie within the dashed blue lines). Second column assesses homoscedasticity. Third and fourth columns evaluate normality.

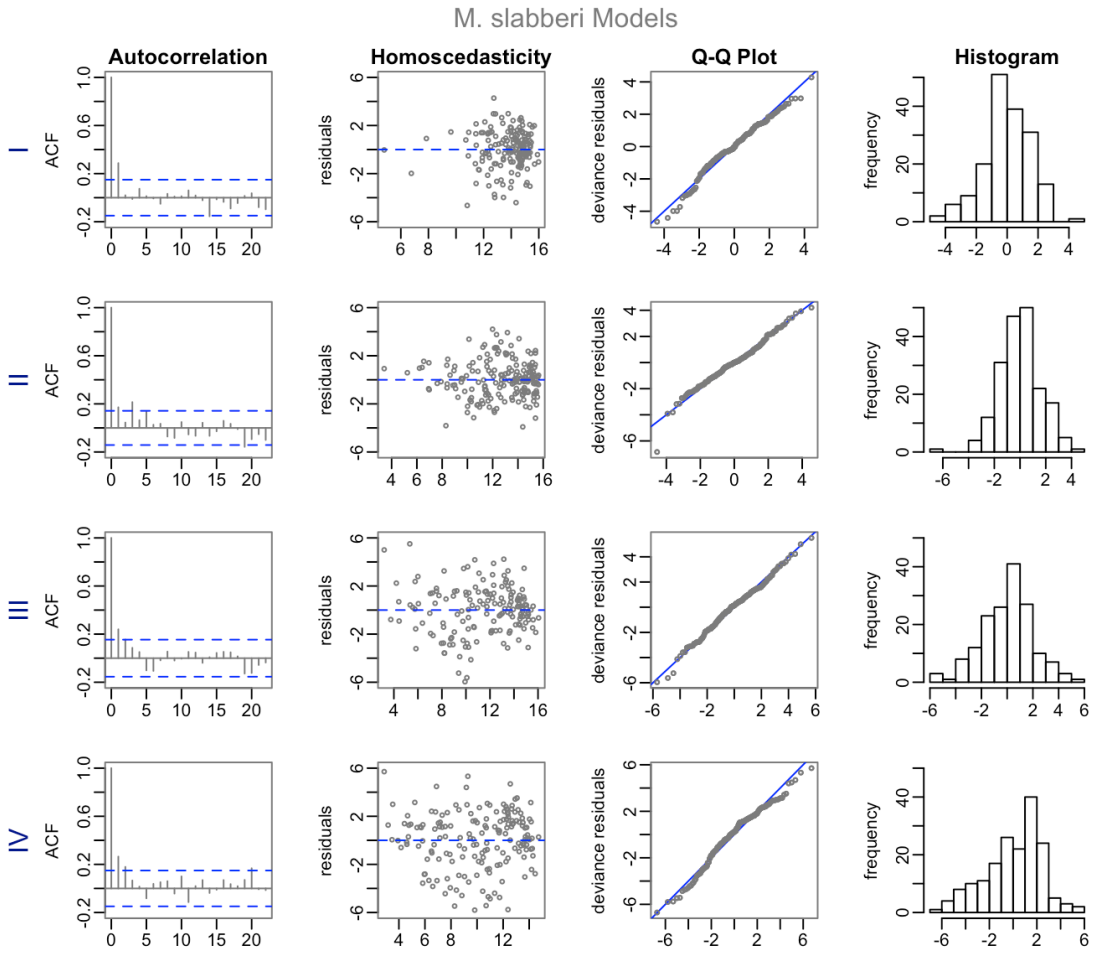


Figure S9. Residuals. The graphs are structured as follows: rows correspond to each of the *N. integer* individual models. First column shows the autocorrelation within the residuals (correlations are not significant if they lie within the dashed blue lines). Second column assesses homoscedasticity. Third and fourth columns evaluate normality.

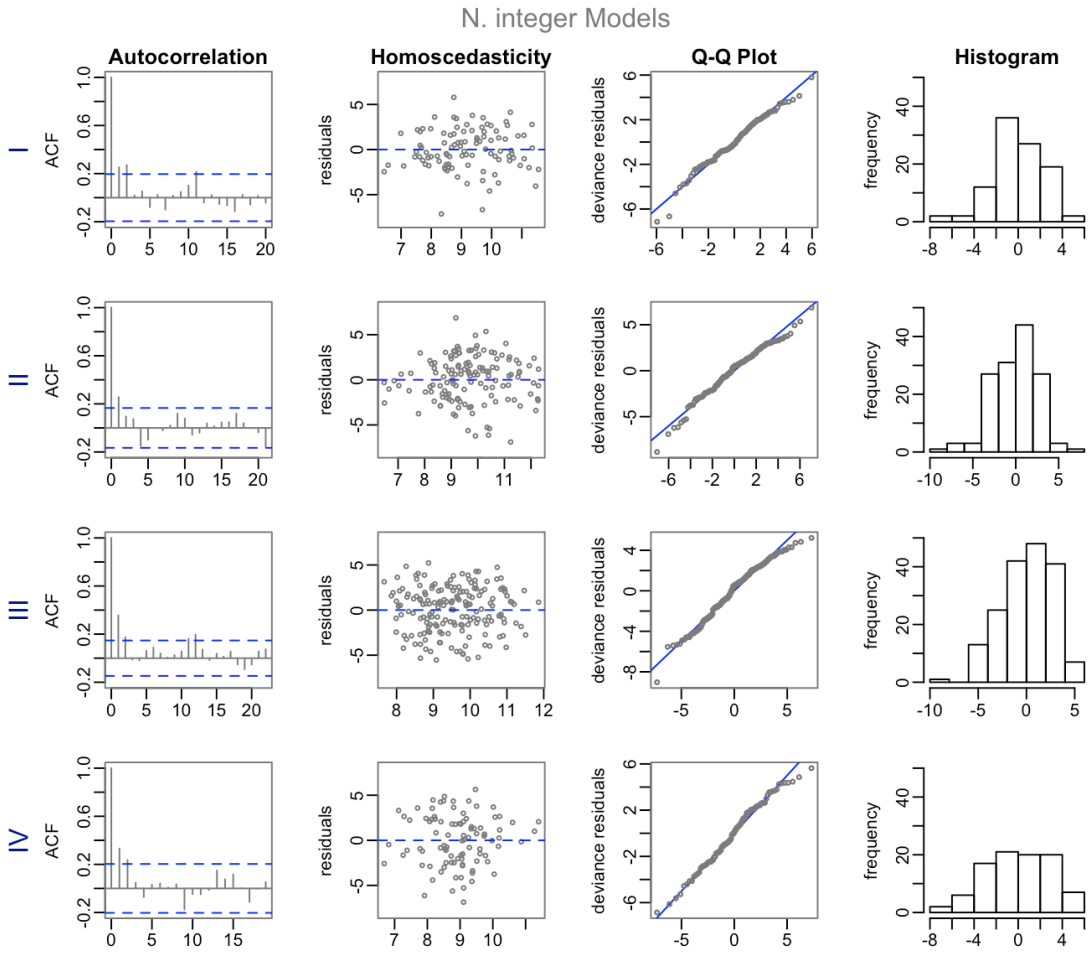


Figure S10. Residuals. The graphs are structured as follows: rows correspond to each of the *R. tartessicus* individual models. First column shows the autocorrelation within the residuals (correlations are not significant if they lie within the dashed blue lines). Second column assesses homoscedasticity. Third and fourth columns evaluate normality.

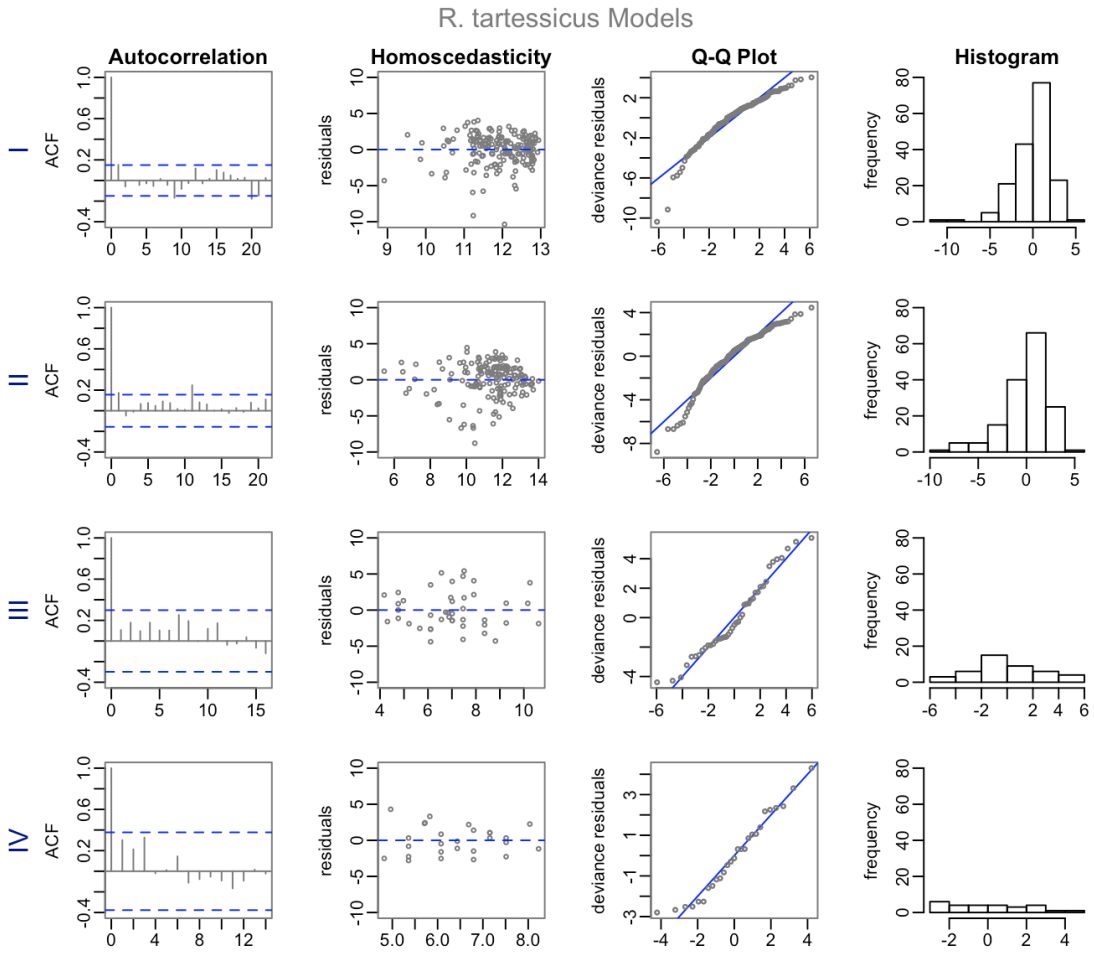


Figure S11. Time series of anchovy abundance: (a) month-to-month and (b) seasonal cycle for the four water masses. Black and red lines represent observations and predictions, respectively. Blue stripes shows the occurrence of high freshwater discharges [>20 and >400 hm^3 (7 and 30 days, respectively) depending on water mass] and grey shading those of high turbidity (>100 and >550 NTU, respectively). Pearson's correlations are presented in the upper left corner. Correlations with a $p < 0.05$ are indicated *, $p < 0.01$**, $p < 0.001$***.

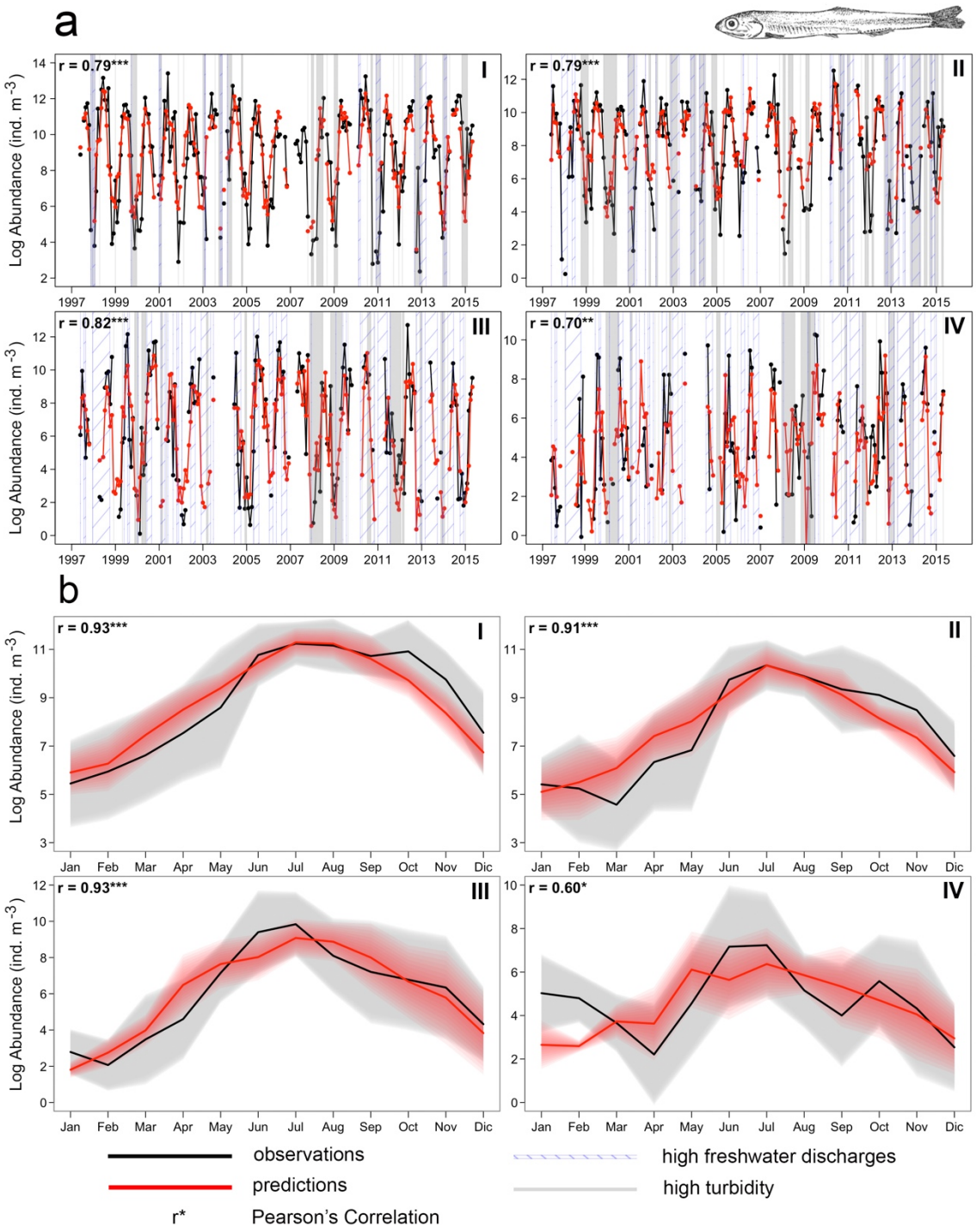


Figure S12. Time series of *M. slabberi* abundance: (a) month-to-month and (b) seasonal cycle for the four water masses. Black and red lines represent observations and predictions, respectively. Blue stripes shows the occurrence of high freshwater discharges [>1.5 and >2.5 hm^3 (4 days, respectively) depending on water mass] and grey shading those of high turbidity (>100 and >250 NTU, respectively). Pearson's correlations are presented in the upper left corner.