A general linear model relating an index of proportional entrainment loss to turbidity and Old and Middle River flow

William Smith, 4 October 2018

1 Introduction

This note documents an effort to model an index of proportional entrainment as a function of existing water operations management quantities for adult delta smelt. An *index* of proportional entrainment is used for two reasons. Primarily, there is a mismatch between the timing of abundance estimates and entrainment estimates. Ideally, seasonal entrainment would be expressed as a fraction of abundance at the beginning of the entrainment season, on December 1; however, reliable abundance estimates were not available until January–February. Abundance estimates for November were available but were not considered as accurate as those measured in January–February. Additionally, the simplified model expressed here does not account for competing natural mortality that occurs simultaneously with entrainment mortality.

2 Methods

response: index of proportional entrainment loss (PEL)

$$(1) \qquad PEL_y = \frac{entrainment_y^{estimate}}{abundance_y^{estimate}},$$

Total December through March entrainment of adult delta smelt $entrainment_y^{estimate}$ were separately estimated using a hierarchical model accounting for survival between entrainment and sampling and the sampling efficiency of fish facilities (TN 36). Adult abundances $abundance_y^{estimate}$ were separately estimated from January–February Spring Midwater Trawl samples for cohorts 1993–2000 (corresponding to years 1994–2001) and from Spring Kodiak Trawl samples for cohorts 2001–2015. Abundance estimates were design-based stratified mean catch densities, expanded by strata water volumes and accounting for gear contact selectivity at length (TN 2 and 12).

covariates: December–February mean OMR and mean secchi (measured throughout the Delta during fish surveys). Three water operations management periods were used to categorize cohort years (calendar year-1) into three management regimes pre-CalFed, CalFed, and BiOp years, corresponding to 1993–1998, 1999–2006, and 2007–2015 and periods of unmanaged OMR flow, management to more negative OMR flow, and management to less negative OMR flow.

model: weighted generalized linear model, using beta regression (betareg in R) and a logistic link

$$(2) \qquad PEL_y = \frac{1}{1+e^{-(\beta_0+\beta_1 * OMR_y+\beta_2 * Secchi_y+\beta_3 * OMR_y * Secchi_y+\beta_4 * Regime_y+\varepsilon_y)}},$$

where ε were normally distributed errors with mean 0. For consistency with existing management parameters, only models of OMR, Secchi, and Regime were explored. Model weights were set equal to the inverse of Monte Carlo simulated variance of PEL_y (Eq. 1). Variances of PEL_y were developed by iteratively resampling random $entrainment_y'$ and $abundance_y'$ values from log-normal distributions with mean and associated errors set equal to the values estimated from Spring Midwater and Kodiak Trawl surveys

$$(3) \qquad entrainment'_y \sim \text{Lognormal}\left(\log\left(entrainment_y^{estimate}\right), \sqrt{\left(1+\frac{entrainment_y^{se}}{entrainment_y^{estimate}}\right)}\right) \text{ and}$$

$$(4) \qquad abundance'_y \sim \text{Lognormal}\left(\log\left(abundance_y^{estimate}\right), \sqrt{\left(1+\frac{abundance_y^{se}}{abundance_y^{estimate}}\right)}\right).$$

model selection: all combinations of models using Secchi, OMR, Regime, and a Secchi-OMR interaction were compared, including an intercept-only model, using AIC.

3 Results

Graphical representations of *PEL* versus OMR and turbidity indicate a negative relationship with each (Fig. 1). The three management regime periods were evident in both the time series of *PEL* and OMR, with moderate *PEL* and highly variable OMR during the (cohort year) 1993-1998 pre-CalFed period, higher *PEL* and more negative OMR during the 1999-2007 CalFed period, and lower, less variable *PEL* and less negative, less variable OMR during the 2008-2015 BiOp period.

Model selection, fit, and diagnostics

AIC model selection indicated the best model was the full model including Secchi, OMR, management regime, and secchi-OMR interaction effects (Table 1). Although the full model was selected, the regression coefficient associated with the BiOp period was not significant (P-value = 0.6). Overfitting was a concern, because 23 observations were used to estimate 6 regression parameters, but AIC is considered robust to overfitting because it penalizes model complexity.

Residuals of the full model indicated errors were normally distributed, with no concerning patterns (Fig. 2), but standardized residuals were larger than expected and regression parameter standard errors were smaller than expected. Both were the consequence of regression weights. Removing all weights resulted in identical model selection and similar parameter estimates, but smaller residuals and higher regression parameter standard errors. Only *PEL* for cohort year 1993 was associated with a large residual and high leverage, as indicated by Cook's distance.

Model application to predict index of proportional entrainment loss

A table of predicted indices of proportional entrainment losses was developed under two turbidity conditions (secchi = mean and mean-sd), a range OMR conditions, and the three management regimes (Fig. 3).

Table 1. Beta regression model results. Covariates were standardized, so parameter estimates may be interpertted as effect sizes. The best model is indicated by lowest AIC (ΔAIC = 0). P-values for all parameters were highly significant (P-value < 0.001).

			Estimate (se)					
Model	AIC	ΔAIC	intercept	Secchi	OMR	CalFed Regime	BiOp Regime	Secchi* OMR
intercept	-193	48	-3.11 (<0.001)	--	--	--	--	--
Secchi	-205	35	-3.25 (<0.001)	-0.55 (<0.001)	--	--	--	--
OMR	-196	44	-3.4 (<0.001)	-0.91 (<0.001)	--	--	--	--
Regime	-210	30	-3.42 (<0.001)	0.87 (0.006)	-0.68 (0.061)	--	--	--
Secchi+OMR	-215	26	-3.56 (<0.001)	-0.61 (<0.001)	-0.94 (<0.001)	--	--	--
Secchi*OMR	-237	3	-4.51 (<0.001)	-1.21 (<0.001)	-2.86 (<0.001)	-1.58 (<0.001)	--	--
Secchi+Regime	-223	18	-3.82 (<0.001)	-0.54 (<0.001)	1.12 (<0.001)	-0.23 (0.543)	--	--
OMR+Regime	-209	32	-3.42 (<0.001)	-0.34 (0.102)	0.68 (0.07)	-0.74 (0.0473)	--	--
Secchi+OMR+Regime	-223	18	-3.74 (<0.001)	-0.54 (<0.001)	-0.4 (0.036)	0.8 (0.0184)	-0.37 (0.331)	--
Secchi*OMR+Regime	-241	0	-4.56 (<0.001)	-1.07 (<0.001)	-2.23 (<0.001)	0.58 (0.0511)	-0.18 (0.6)	-1.27 (<0.001)

Figure 1. Time series of the index of proportional entrainment loss (PEL), mean December–February Old and Middle River flow (OMR), and mean secchi depth in the Delta, and the relationships among the index of PEL, OMR, and secchi depth.

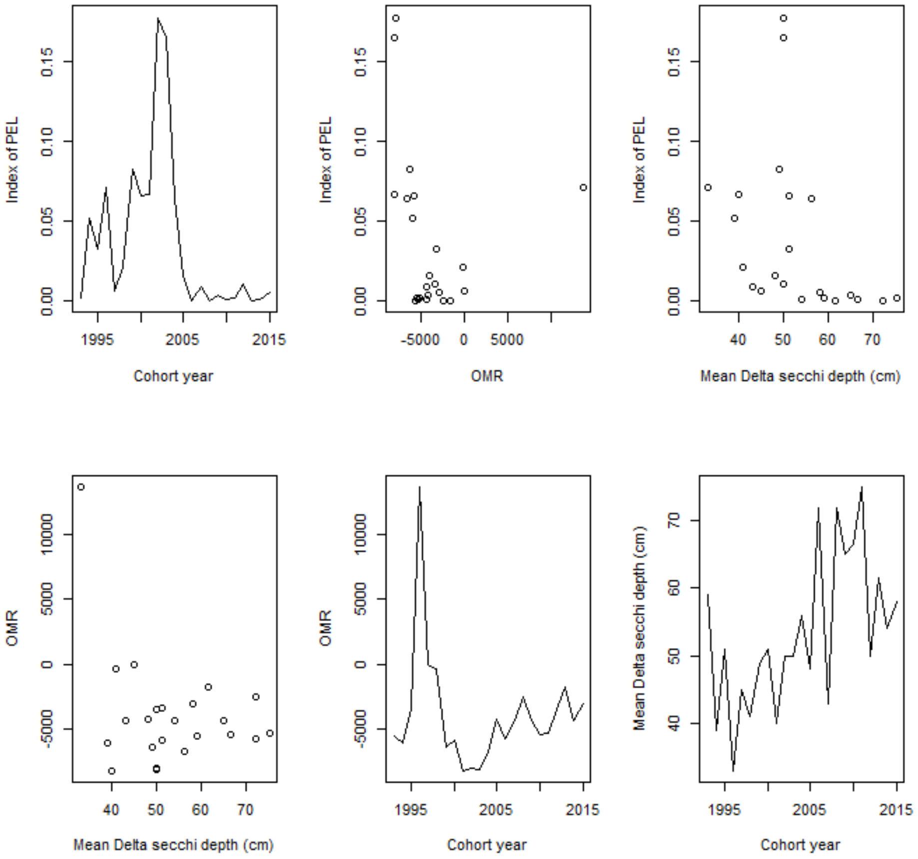

Figure 2. Residual plots for the model expit$(PEL_y) = \beta_0 + \beta_1 * Secchi_y + \beta_2 * OMR_y + \beta_3 * Secchi_y * OMR_y + \beta_4 * Regime_y$.

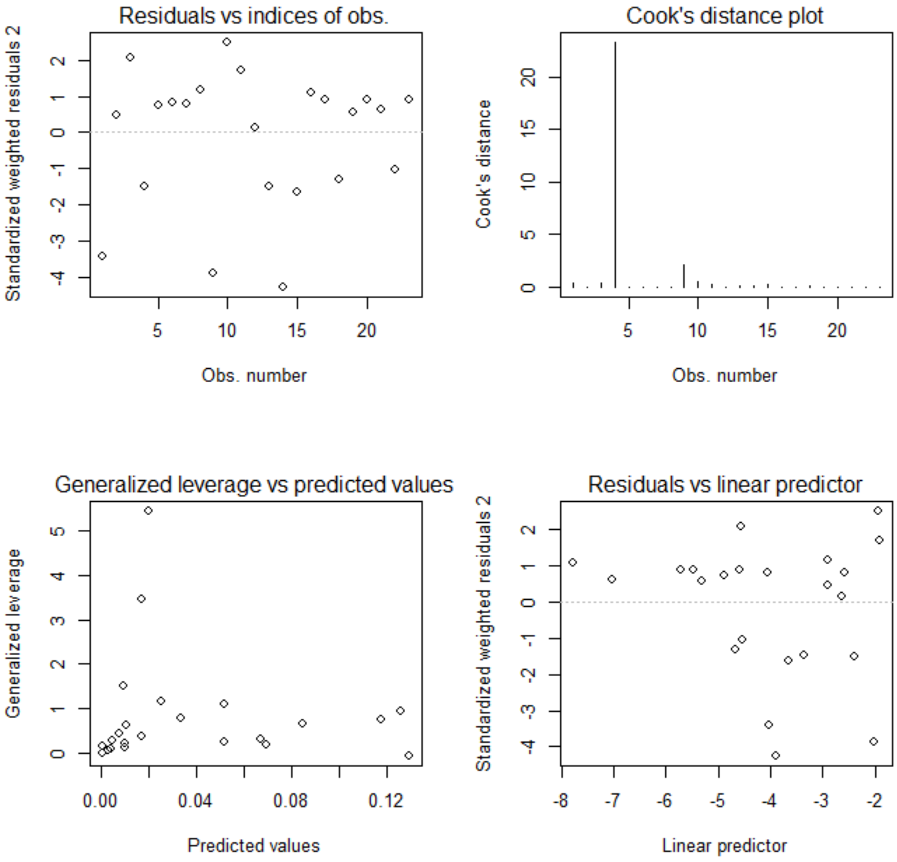

Figure 3. Model predictions from the best PEL model identified using AIC under muddy and average secchi depth conditions. Black lines indicate mean predictions, and red lines indicate 95% confidence intervals of the mean.