Assessing goodness of fit for LESV model
(ECASA internal paper)
Napier University

Enrique Portilla, Paul Tett

October 1, 2007

1 Introduction

ECASA WP4's objective 2 cover the development of operational tools, especially models, which capture the functional relationship between environment and aquacultural activities, and which embody the chosen indicators. The WP's deliverable is a"Toolpack" report on the merits of the chosen indicator set including best methodologies for collection, analysis and interpretation, and on the recommended set of models, including criteria for choice of models depending on spatial scale and farm size, and guidance on the use of models to estimate site and water body assimilative capacity and sustainable production, and on the reliability of model predictions.

Here, we present the results of the simplified procedures to assess the *reliability of model predictions* presented to all ECASA modellers in February. Modellers were are asked to use this procedure, so that the project can report a comparable set of model (and indicator) assessments.

This report applies a standard test for the accuracy and precision of all ECASA models. The technique used is based on comparison between simulated data and observed data using linear regression. Models are judged on the basis of (i) the proportion of variance explained by the regression, and (ii) the regression coefficients. These criteria are similar to those described by Mesple et al. (1996).

Following Oreskes et al. 1994, we judge a model as *excellent, good, fair* or *poor*, accepting their argument that this terminology encourages model use and improvement.

This document has three sections. The first part presents the result of the application of procedures agreed by all ECASA modellers, to simulations made with the L-ESV model for 1975 and 2003. The second part is Appendix A with goodness of fit tabulated according to the agreed format. The third part is Appendix B, which gives more details of the *Major*

Table 1: Criteria established to classify the model predictions (X) in different grades when compared with data (Y). Four groups are formed according to the estimates of the slope (β_1) and intercept (β_0) from the regression line of the form $Y = \beta_1 X + \beta_0 + \epsilon$ (being ϵ the error term of the regression model). Within each of the groups the quality of the simulation increases with increasing values of R^2 (the coefficient of determination). The statement $\neq$ corresponds to the result of a t-test being *not significantly different from*, at the level $\alpha = 0.05$

Model category	Slope	Intercept	Interpretation
Excellent	$\beta_1 = 1$	$\beta_0 = 0$	The model is regarded as perfectly simulating, on average, the observations
Good (effect 1)	$\beta_1 = 1$	$\beta_0 \neq 0$	The model over or underestimate, on average, the observations
Good (effect 2)	$\beta_1 \neq 1$	$\beta_0 = 0$	The difference between model predictions and observation is proportional to the values predicted.
Fair	$\beta_1 \neq 1$	$\beta_0 \neq 0$	Effects 1 and 2 combined
Poor	$\beta_1 = 0$		There is no relationship between model and observations

Axix Regression method (Sokal and Rohlf, 1995) used in the comparisons of simulations and observations.

2 Material and Methods

2.1 Introduction

We have confronted simulations with the Loch Ecological State Vector (L-ESV) model with data from Loch Creran in two different years, 1975, prior to fish-farming, and 2003. The L-ESV model predicts daily changes on microplankton (phytoplankton plus pelagic microheterotrophs) in sea lochs. The output is a time series of Ecological State Variables (ESV) in a zone B scale water body.

The L-ESV model has been developed from the single-box, dynamic CSTT model described by Laurent et al. 2006. The version used for the goodness of fit tests was that described by Portilla and Tett (2007) and includes FJORDENV-like physics in 3 layers (Gillibrand and Inall, 2006; Portilla and Tett, 2006).

2.2 Variables used in the test

The variables simulated by the L-ESV model are daily values, in the 3 layers, of: i) Chlorophyll for two micro-plankton compartments (a diatom like microplankton and a flagellate like microplankton); ii) Dissolved Available Inorganic Nitrogen, DAIN (composed of nitrate, nitrite and ammonium), iii)Dissolved Inorganic Phosphorus, DIP, iv) and Dissolved Silicate, DSi; v) Dissolved Oxygen, DO; vi) Suspended particulate Matter, SPM, vii) and Salinity. The model was used to simulate chlorophyll in the fjordic loch Creran in 1975 and 2003, with forcing from daily values of meteorological variables and a climatological boundary condition from outside the sea-loch. In 2003, a salmon farm was adding significant N and P to the loch, in daily amounts calculated by the 'Blackfish' routine in L-ESV from data on monthly amounts of feed. Two runs of the LESV model were performed for 2003, trying to produce the best fit to the data. The first simulation used standard values of the grazing loss rate parameter ($L_0 = 0.1$ d^{-1} at 10°) for both microplanktons. The second simulation used a higher value ($L_0 = 0.1$ d^{-1} at 10°) for the diatom like micro-plankton compartment.

We confronted the model's numerical simulations with observations of total chlorophyll and two dissolved limiting nutrients (DAIN and DIP). The samples were collected from the upper layer of the main basin of Loch Creran in 1975 and 2003. Data for 1975 were taken from the thesis by (Jones, 1979), and data for 2003 are those reported by (Laurent et al., 2006).

The sampling was done throughout the year at different locations. The standard error associated with each sampling day is related to the variability in the samples across the different locations sampled. This error could be used to weight the least squares algorithm – i.e., performing what is called a Weighted Regression (Sokal and Rohlf, 1995). This is one method for compensating for lack of homogeneity in the sample variability. Instead, we used logarithmic transformation of the chlorophyll concentration to correct for heterogeneity, because the variability in concentration was found by Tett and Wallis (1978) to be proportional to the concentration.

For the two years, and three state variables, observations were plotted against corresponding simulated values for layer 1 in the model. Linear regressions were fitted using the well-known Ordinary Least Squares (OLS) method. In addition, the coefficients of the regression line were estimated by using a more robust technique, Major Axis Regression (MAR), which allows for error variability in x-axis as well as y-axis variables. We provide a brief description of its computation in Appendix B. A comparison between the OLS and MAR will change the perception of the performance of the model for certain state variables, variables and conditions.

3 Results

We assessed the LESV model in three cases by using the (OLS) and the Major Axis Regression (MAR) methodologies (table 2). Figure 1 compares simulations with observations of Chlorohyll, DAIN and DIP in the upper layer of Loch Creran for 1975. Figure 2) compares simulations forced with data from 2003 against Chlorophyll, DAIN and DIP oberved in the upper layer during 2003 In the third case, we modified one coefficient of the model, forced the model with 2003 data, and compared the resulting simulation with observations of Chlorophyll, DAIN and DIP during that year (figure 3). The goodnes–of–fit classifications from the OLS regression and the MAR were similar in all cases apart from that of Chlorophyll in the second simulation for 2003. Nevertheless, the MAR procedure tended to suggest better performance by the model than that implied by the OLS method (table 3).

The assessment of the model for 1975 produces reasonably good results for the three state variables considered (figure 1) obtaining a large r^2 ($r^2 > 0.7$). The test for the slope and intercept for the OLS shows that only the intercepts for DAIN and DIP cannot reject the hypothesis of $\beta_0 = 0$. Nevertheless, the model will pass the test if we asses the goodness of fit with the MAR for the 3 state variables, i.e. 0 and 1 are within the confidence intervals of the values estimated for the slope and the intercept respectively.

The model did not simulate 2003 very well (figure 2). The goodness of fit obtained for 1975 was not obtained either for the standard loss rates of when the loss rate for diatomy microplankton was increased in order to force total chlorophyll to decrease to the same level as the observed values (figure 3).

For the first simulation in 2003, the chlorophyll values predicted by the model were higher than expected (figure 2a). This was highlighted by the low regression line found (table 2). For the nutrient variables, the fit was consistent and both intercept and slope estimate were satisfactory.

Finally for the second simulation in 2003, with higher loss rate, the chlorophyll concentrations were simulated at the same level as the observed values, and much lower than in the first simulation for 2003 (figure 3a). Nevertheless, the comparative plot between observed and simulated values (figure 3c) suggest that we have to take some care when assessing the model. Although the OLS assessment consider the model to be poor for chlorophyll, the MAR assessment consider the model to perform better (table 3. The MAR intercept and slope did not differ significantly from the expected values of 0 and 1. However, the ranges of values for both parameters were quite large with the zero value included in that range. This result arises from the large variation recorded in both the observations and the simulations.

The second simulation in 2003 did not perform as well as the first for nutrients in general (low value of r^2 for both nutrients) and for DAIN in particular. DAIN predictions were systematically higher than the observed values (figure 3c-d). Although a similar situation can be observed for DIP (figure 3e-f), we cannot in this case reject the hypothesis of the intercept being equal to zero. What is remarkable, though, is that the slope for both nutrients was equal to the ideal value of 1 (table 2).

4 Discussion

4.1 Methods

We tried two regression methods for comparing observations with simulations, and found that the assessment of goodness–of–fit differs between the method. The use of the MAR is a robust choice for assessing the goodness–of–fit of models because models are sometimes subject to error in their predictions. The sources of model errors can be widespread, and include error associated with the parameters or error in the forcing variables. The presence of those errors, when estimating the regression line, leads to wrong estimates of the slope and the intercept when using the OLS. Therefore, the use of the MAR yields a better model classification to the scientist than that of the OLS method.

A different matter is the error associated with the field data used to validate the models. Linear regression is very sensitive to spatial or temporal variability in the field data. These variability can produce type II errors when testing hypotheis. One might decrease the rate of *false negatives* by increasing sampling frequency, for instance. Also the further transformation of the data, or the use of Weighted Regression or (see (Sokal and Rohlf, 1995)) might be investigated further. We used Major axis regression (see (Sokal and Rohlf, 1995) and section B in this document), transformed the data for normality when necessary, but ignore the weights for the data when it was available.

4.2 Goodness of fit in 1975 and 2003

Simulation with the L-ESV model proved to be an *excellent* fit to the seasonal variations of chlorophyll and nutrients in Loch Creran for 1975. For 2003 case, the model did not perform so well, and the fit is classified as *good*. Overall we classify the model as *good*, based on the criteria in Table 1 and the results in Table 2 These results raise a question about why the model, which simulated 1975 well, not be able to fit 2003 data so well? As Table 1 shows, we could fit either the nutrient data, or the chlorophyll data, but not both together. We have therefore to consider (a) what might have changed in loch Creran between 1975 and 2003 in addition to nutrient enrichment, and (b) what should be added to L-ESV to be able to simulate this change. These are matters for further study.

References

Gillibrand, P. A., Inall, M. E., March 2006. Improving assimilative capacity modelling for Scottish coastal waters: I. a model of physical exchange in scottish sea lochs. Tech. Rep. 167, Scottish Association for Marine Science, Dunstaffnage Marine Laboratory, Oban. 2

Jones, K. J., 1979. Studies on nutrient levels and phytoplankton growth in a Scottish sea loch. Phd, University of Strathclyde. 3

Laurent, C., Tett, P., Fernandes, T., Gilpin, L., Jones, K. J., 2006. A dynamic CSTT model for the effects of added nutrients in Loch Creran, a shallow fjord. Journal of Marine Systems. 61 (3-4), 149–164. 2, 3

Mesple, F., Troussellier, M., Casellas, C., Legendre, P., 1996. Evaluation of simple statistical criteria to qualify a simulation. Ecological modelling. 88 (1), 9–18. 1, 21

Oreskes, N., Shrader-Frechette, K., Belitz, K., 1994. Verification, validation, and confirmation of numerical models in the earth sciences. Science 263, 641–646. 1

Portilla, E., Tett, P., 2006. Report on the dynamic CSTT (dCSTT) model: Sarf012 milestone 4. Tech. rep., School of Life sciecnes, Napier University, Edinburgh. 2

Portilla, E., Tett, P., 2007. Report on the Loch Ecological State Vector (LESV) model: Sarf012 milestone 5. Tech. rep., School of Life sciecnes, Napier University, Edinburgh. 2

Sokal, R., Rohlf, F., 1995. Biometry: The Principles and Practice of Statistics in Biological Research., 3rd Edition. Freeman and Co., New York. 2, 3, 5

Tett, P., Wallis, A., 1978. The general annual cycle of chlorophyll standing crop in Loch Creran. Journal of Ecology 66, 227–239. 3

Warton, D., Wright, I., Falster, D., Westoby, M., 2006. Bivariate line-fitting methods for allometry. Biological Reviews 81 (2), 259–291. 22

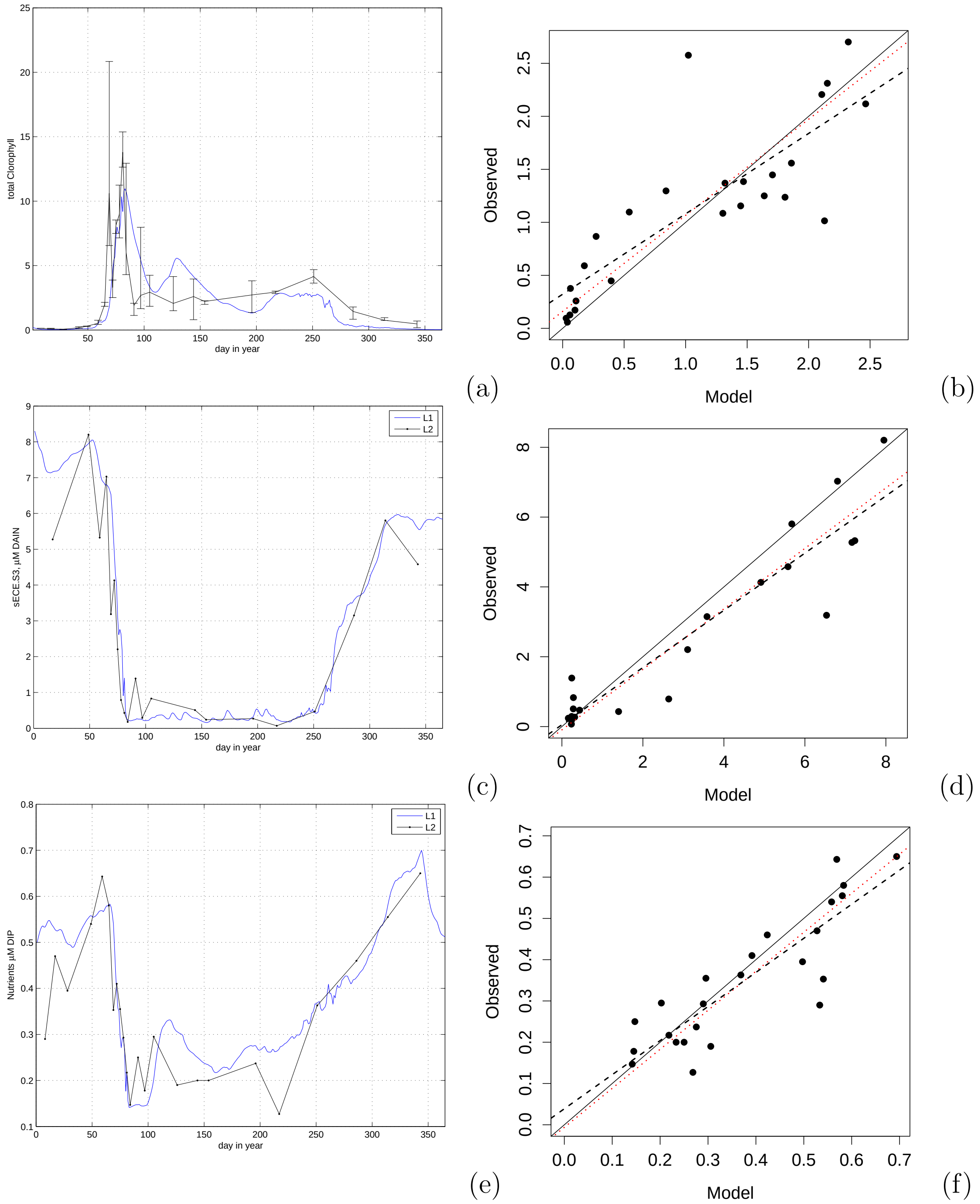

Figure 1: Test of the LESV model for 1975. The left column contains the time series of observed values (black) and simulated values (blue) for total chlorophyll (a) DAIN (c) and DIP (e). The error bars in (a) are quantiles 5 and 95, so they contain 90% of the data. The right column contains plots of simulated values on the x-axis versus observed median values on the y-axis for total chlorophyll (b) DAIN (d) and DIP (f). The chlorophyll data are given as $\ln(\mathrm{mg\ m^{-3}} + 1)$. The continuous line is the one to one relationship expected in the case of a perfect fit. The broken black line is the relationship found by OLS regression and the dotted red line is the relationship found by MAR.

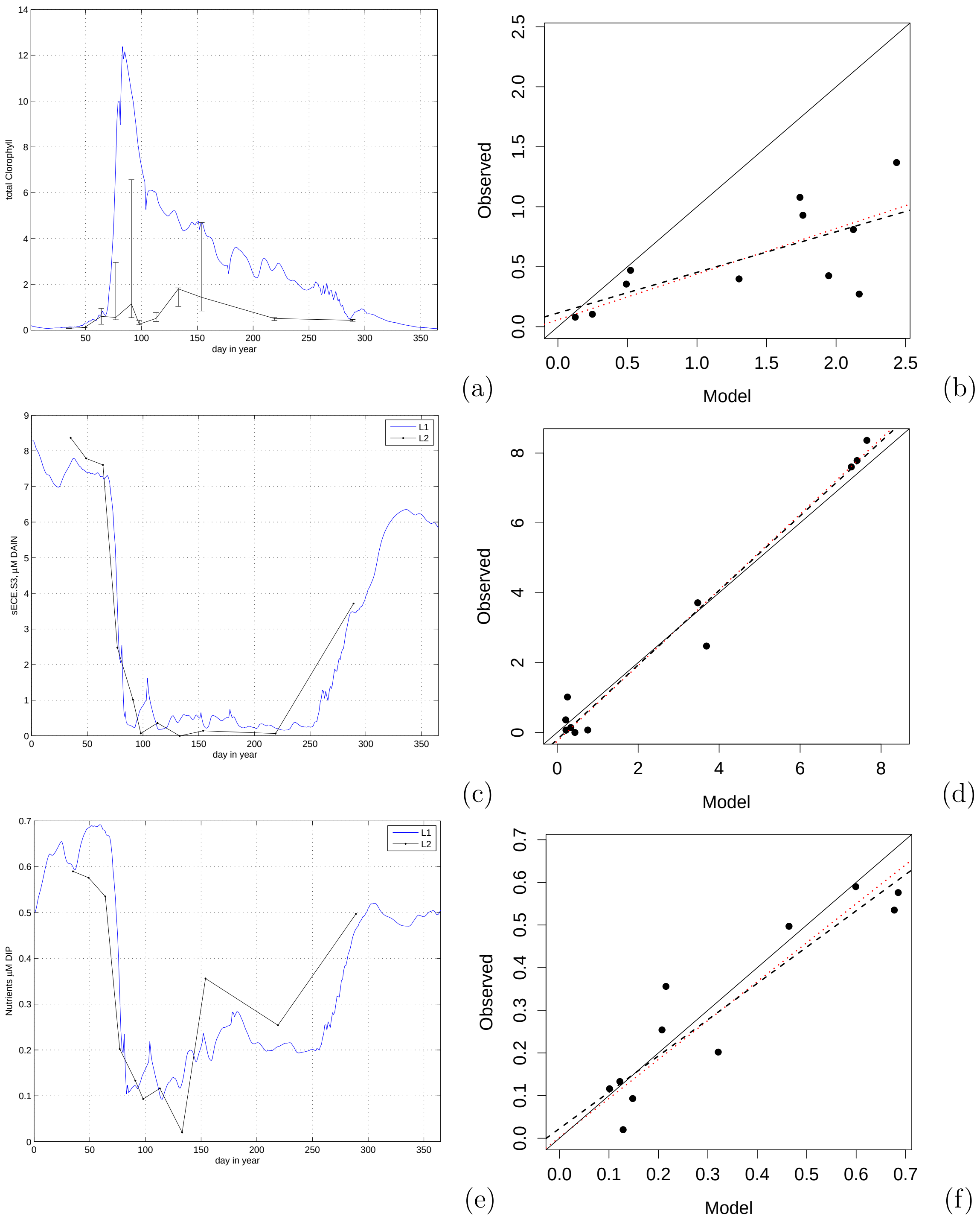

Figure 2: Same as figure 1 but for year 2003. Loss rate L_0 equal to the standard value of $0.1\ \mathrm{d}^{-1}$ at 10°C for both microplanktons.

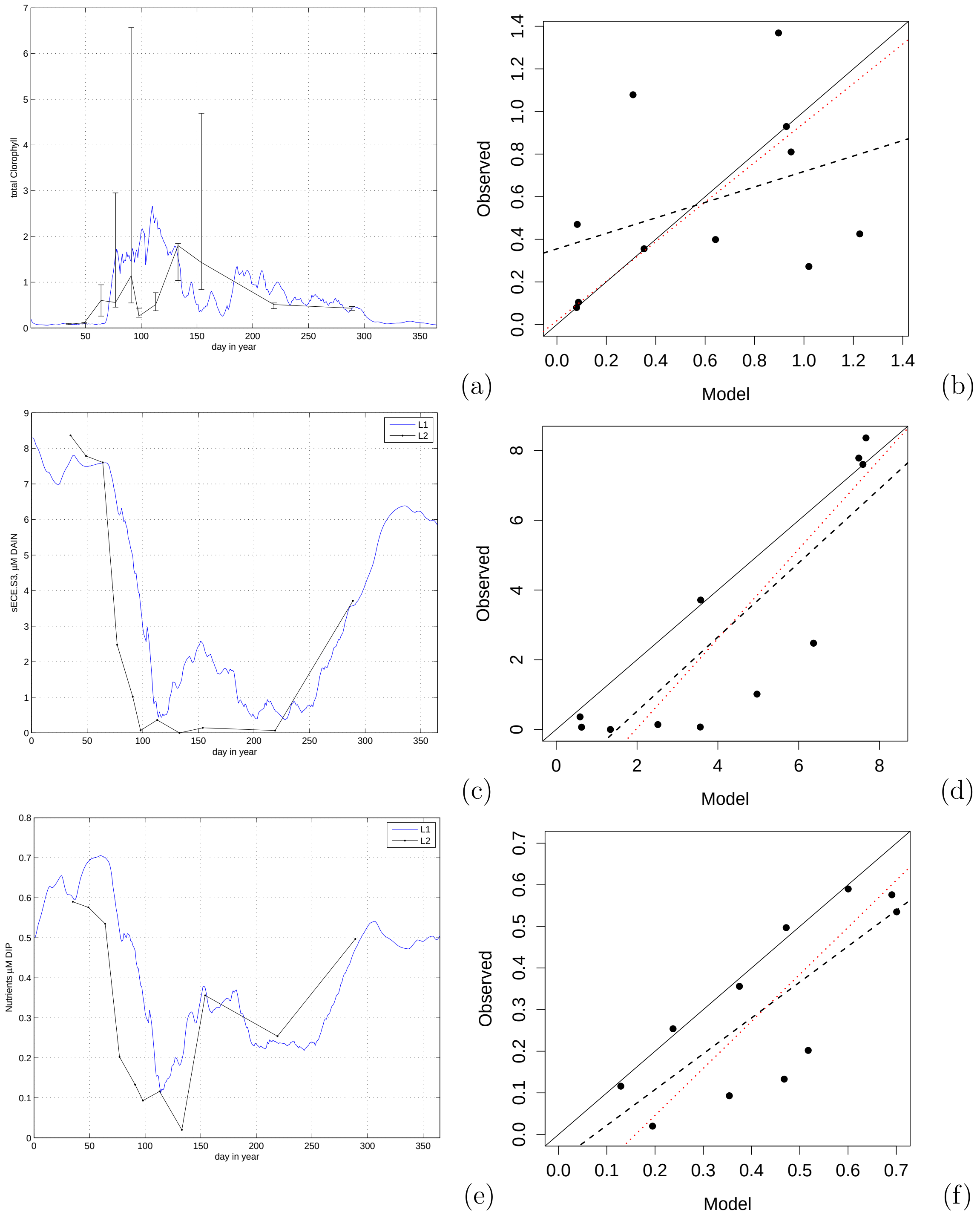

Figure 3: Same as figure 2 but with loss rate (L_0) equal to 0.7 d^{-1} at 10°C for microplankton life form 1

Table 2: Statistical results from the assessment of goodness of fit of LESV simulations against 1975 and 2003 data. The 2003bis case used modified loss rate parameters. The results shown are the slope (β_1) and intercept (β_0) from the OLS and MAR line-fits. The lower and upper values are 95% confidence limits to the estimates. We look for the confidence limits of the intercept to include zero, and the confidence limits of the slope to include 1, in order to say a model fits the data well. Parameter estimates that do not do this, are shown in bold.

		1975			2003			2003bis		
	Method	Lowr	Est	Upr	Lowr	Est	Upr	Lowr	Est	Upr
CHL										
β_0	OLS	0.063	**0.322**	0.580	-0.225	0.115	0.454	-0.044	0.355	0.754
	MAR	-0.188	0.159	0.423	-0.408	0.058	0.437	-0.866	0.018	2.059
β_1	OLS	0.571	**0.758**	0.946	0.123	**0.339**	0.554	-0.188	**0.363***	0.914
	MAR	0.665	0.906	1.223	0.100	**0.381**	0.726	-0.492	0.928*	2.489
DAIN										
β_0	OLS	-0.459	0.051	0.562	-0.654	-0.202	0.250	-3.543	-1.603	0.337
	MAR	-0.613	-0.094	0.352	-0.666	-0.247	0.118	-6.062	**-2.534**	-0.588
β_1	OLS	0.699	**0.819**	0.940	0.959	1.067	1.175	0.673	1.064	1.454
	MAR	0.722	0.866	1.034	0.956	1.083	1.228	0.822	1.285	2.123
DIP										
β_0	OLS	-0.032	0.039	0.110	-0.062	0.023	0.108	-0.261	-0.064	0.132
	MAR	-0.110	-0.006	0.075	-0.109	0.002	0.085	-0.654	-0.180	0.046
β_1	OLS	0.652	**0.826**	0.999	0.638	0.851	1.064	0.442	0.861	1.280
	MAR	0.729	0.945	1.220	0.664	0.913	1.245	0.604	1.130	2.229

* Zero is included in the interval, so no relationship between simulations and observations.

Table 3: Model category (Table 1) for LESV from the two regression techniques used here: The Ordinary Least Squares (OLS) and the Major Axis Regression (MAR). The LESV model was tested under three different situations and for three state variables. LESV against 1975 data, LESV against 2003 data and LESV with loss rate parameter modified (2003bis).

	1975		2003		2003bis	
	OLS	MAR	OLS	MAR	OLS	MAR
CHL	Fair	Excellent	Good	Good	Poor	Poor
DAIN	Good	Excellent	Excellent	Excellent	Excellent	Good
DIP	Good	Excellent	Excellent	Excellent	Excellent	Excellent

APPENDIX

A Tables

Table 4:

Model Description			
Model name		L-ESV	
State variable		Chlorophyll-a	
site at which tested		Loch Creran 1975	
n, number of independent observations used in test		25	
Model Performance			
r^2, % of variance	67.58	p, on null hypothesis	$p < 0.001$
$\hat{\beta}_0$, regression intercept	0.322	$se_{\hat{\beta}_0}$	0.151
$t = (\hat{\beta}_0 - 0)/se_{\beta_0}$	2.133	p	0.0438
$\hat{\beta}_1$, regression slope	0.758	$se_{\hat{\beta}_1}$	0.110
$t = (\hat{\beta}_1 - 1)/se_{\beta_1}$	-2.21	p	0.038
Model Conclusion [†]			
Model explaint a signifcant part of variance in observations		YES	
Model reliability Class		3 Fair	

[†] Delete what does not apply

Table 5:

Model Description

Model name	L-ESV
State variable	DAIN
site at which tested	Loch Creran 1975
n, number of independent observations used in test	21

Model Performance

r^2, % of variance	87.89	p, on null hypothesis	$p < 0.001$
$\hat{\beta}_0$, regression intercept	0.051	$se_{\hat{\beta}_0}$	0.295
$t = (\hat{\beta}_0 - 0)/se_{\beta_0}$	0.174	p	0.863
$\hat{\beta}_1$, regression slope	0.819	$se_{\hat{\beta}_1}$	0.070
$t = (\hat{\beta}_1 - 1)/se_{\beta_1}$	-2.587	p	0.0164

Model Conclusion [†]

Model explain a signifcant part of variance in observations	YES
Model reliability Class	2 Good

[†] Delete what does not apply

Table 6:

Model Description			
Model name		L-ESV	
State variable		DIP	
site at which tested		Loch Creran 1975	
n, number of independent observations used in test		24	
Model Performance			
r^2, % of variance	75.21	p, on null hypothesis	$p < 0.001$
$\hat{\beta}_0$, regression intercept	0.039	$se_{\hat{\beta}_0}$	0.041
$t = (\hat{\beta}_0 - 0)/se_{\beta_0}$	0.935	p	0.360
$\hat{\beta}_1$, regression slope	0.836	$se_{\hat{\beta}_1}$	0.101
$t = (\hat{\beta}_1 - 1)/se_{\beta_1}$	-1.722	p	0.098
Model Conclusion †			
Model explain a signifcant part of variance in observations		YES	
Model reliability Class		2 good	

† Delete what does not apply

Table 7:

Model Description			
Model name	L-ESV		
State variable	Chlorophyll-a		
site at which tested	Loch Creran 2003		
n, number of independent observations used in test	11		
Model Performance			
r^2, % of variance	47.97	p, on null hypothesis	$p = 0.018$
$\hat{\beta}_0$, regression intercept	0.114	$se_{\hat{\beta}_0}$	0.018
$t = (\hat{\beta}_0 - 0)/se_{\beta_0}$	0.618	p	0.543
$\hat{\beta}_1$, regression slope	0.339	$se_{\hat{\beta}_1}$	0.117
$t = (\hat{\beta}_1 - 1)/se_{\beta_1}$	-5.627	p	$p < 0.001$
Model Conclusion †			
Model explaint a signifcant part of variance in observations	YES		
Model reliability Class	2 Good		

† Delete what does not apply

Table 8:

Model Description			
Model name		L-ESV	
State variable		DAIN	
site at which tested		Loch Creran 2003	
n, number of independent observations used in test		11	
Model Performance			
r^2, % of variance	97.333	p, on null hypothesis	$p < 0.001$
$\hat{\beta}_0$, regression intercept	-0.202	$se_{\hat{\beta}_0}$	0.246
$t = (\hat{\beta}_0 - 0)/se_{\beta_0}$	-0.820	p	0.420
$\hat{\beta}_1$, regression slope	1.067	$se_{\hat{\beta}_1}$	0.059
$t = (\hat{\beta}_1 - 1)/se_{\beta_1}$	1.139	p	0.266
Model Conclusion [†]			
Model explain a signifcant part of variance in observations		YES	
Model reliability Class		2 Excellent	

[†] Delete what does not apply

Table 9:

Model Description			
Model name		L-ESV	
State variable		DIP	
site at which tested		Loch Creran 2003	
n, number of independent observations used in test		11	
Model Performance			
r^2, % of variance	85.67	p, on null hypothesis	$p < 0.001$
$\hat{\beta}_0$, regression intercept	0.022	$se_{\hat{\beta}_0}$	0.046
$t = (\hat{\beta}_0 - 0)/se_{\beta_0}$	0.490	p	0.628
$\hat{\beta}_1$, regression slope	0.851	$se_{\hat{\beta}_1}$	0.116
$t = (\hat{\beta}_1 - 1)/se_{\beta_1}$	-1.285	p	0.211
Model Conclusion [†]			
Model explain a signifcant part of variance in observations		YES	
Model reliability Class		3 Excellent	

[†] Delete what does not apply

Table 10:

Model Description			
Model name		L-ESVbis	
State variable		Chlorophyll-a	
site at which tested		Loch Creran 2003	
n, number of independent observations used in test		11	
Model Performance			
r^2, % of variance	13.95	p, on null hypothesis	$p = 0.258$
$\hat{\beta}_0$, regression intercept	0.35	$se_{\hat{\beta}_0}$	0.216
$t = (\hat{\beta}_0 - 0)/se_{\beta_0}$	1.63	p	0.116
$\hat{\beta}_1$, regression slope	0.363	$se_{\hat{\beta}_1}$	0.301
$t = (\hat{\beta}_1 - 1)/se_{\beta_1}$	-2.118	p	0.045
Model Conclusion [†]			
Model explaint a signifcant part of variance in observations		YES	
Model reliability Class		4 poor	

[†] Delete what does not apply

Table 11:

<table>
<tr><td colspan="4">Model Description</td></tr>
<tr><td colspan="2">Model name</td><td colspan="2">L-ESVbis</td></tr>
<tr><td colspan="2">State variable</td><td colspan="2">DAIN</td></tr>
<tr><td colspan="2">site at which tested</td><td colspan="2">Loch Creran 2003</td></tr>
<tr><td colspan="2">n, number of independent observations used in test</td><td colspan="2">11</td></tr>
<tr><td colspan="4">Model Performance</td></tr>
<tr><td>r^2, % of variance</td><td>73.51</td><td>p, on null hypothesis</td><td>$p < 0.001$</td></tr>
<tr><td>$\hat{\beta}_0$, regression intercept</td><td>-1.603</td><td>$se_{\hat{\beta}_0}$</td><td>1.058</td></tr>
<tr><td>$t = (\hat{\beta}_0 - 0)/se_{\beta_0}$</td><td>-1.515</td><td>$p$</td><td>0.143</td></tr>
<tr><td>$\hat{\beta}_1$, regression slope</td><td>1.063</td><td>$se_{\hat{\beta}_1}$</td><td>0.213</td></tr>
<tr><td>$t = (\hat{\beta}_1 - 1)/se_{\beta_1}$</td><td>0.298</td><td>$p$</td><td>0.768</td></tr>
<tr><td colspan="4">Model Conclusion [†]</td></tr>
<tr><td colspan="2">Model explain a signifcant part of variance in observations</td><td colspan="2">YES</td></tr>
<tr><td colspan="2">Model reliability Class</td><td colspan="2">3
Excellent</td></tr>
</table>

[†] Delete what does not apply

Table 12:

<table>
<tr><td colspan="4">Model Description</td></tr>
<tr><td colspan="2">Model name</td><td colspan="2">L-ESVbis</td></tr>
<tr><td colspan="2">State variable</td><td colspan="2">DIP</td></tr>
<tr><td colspan="2">site at which tested</td><td colspan="2">Loch Creran 2003</td></tr>
<tr><td colspan="2">n, number of independent observations used in test</td><td colspan="2">11</td></tr>
<tr><td colspan="4">Model Performance</td></tr>
<tr><td>r^2, % of variance</td><td>61.18</td><td>p, on null hypothesis</td><td>0.004</td></tr>
<tr><td>$\hat{\beta}_0$, regression intercept</td><td>-0.064</td><td>$se_{\hat{\beta}_0}$</td><td>0.107</td></tr>
<tr><td>$t = (\hat{\beta}_0 - 0)/se_{\beta_0}$</td><td>-0.600</td><td>$p$</td><td>0.554</td></tr>
<tr><td>$\hat{\beta}_1$, regression slope</td><td>0.861</td><td>$se_{\hat{\beta}_1}$</td><td>0.228</td></tr>
<tr><td>$t = (\hat{\beta}_1 - 1)/se_{\beta_1}$</td><td>-0.609</td><td>$p$</td><td>0.548</td></tr>
<tr><td colspan="4">Model Conclusion †</td></tr>
<tr><td colspan="2">Model explain a signifcant part of variance in observations</td><td colspan="2">YES</td></tr>
<tr><td colspan="2">Model reliability Class</td><td colspan="2">3
Excellent</td></tr>
</table>

† Delete what does not apply

B Major axis regression

When both the response (y–axis) variable and the explanatory (x–axis) variable of the regression model are not strongly controlled by the researcher they can include random variation. In such cases the OLS estimation of regression parameters is biased by the presence of measurement error in the explanatory variable – in this case, simulations by the model – and it may be desirable to fit a Major Axis Regression (MAR) (Mesple et al., 1996). [1]

This Apendix gives methods for fitting a MAR. The regression line obtained by this procedure also corresponds to the first principal component of a PCA. As for the OLS, if the model fits well the data, the slope of the MAR should be 1 and the intercept 0.

The estimated regression line in the MAR is the result of minimising the sum , over all the x and y points, of the squared *Euclidean distance* between the points and the regression line, instead of the *vertical distance* as in the OLS.

For the estimation of the intercept and the slope, we have to extract the two eigen values of the observation and predictions. This is estimated by firstly computing the value of D as follow:

$$D = \sqrt{(s2x + s2y)^2 - 4 \cdot (s_x^2 \cdot s_y^2 - s_{xy}^2)} \tag{1}$$

where x refers to the model predictions and y refers to the field observations; s_x^2 and s_x^2 are the estimated variance of predicted values and observed field values respectively; and s_{xy} is their covariance. Then, D is used to estimate the two eigen values:

$$L_1 = \frac{s_x^2 + s_y^2 + D}{2} \tag{2}$$

$$L_2 = \frac{s_x^2 + s_y^2 - D}{2} \tag{3}$$

and finally the value of L_1 is used to estimate the slope (β_1) of the first major axis is:

$$\beta_1 = \frac{s_{xy}}{L_1 - s_y^2} \tag{4}$$

---

[1] The simulated values may be erroneous for several reasons. These include an incorrectly formulated model, error in parameter values or forcing data, and numerical error arising during numerical integration. We suspect that the main source of simulation errors lies in the boundary condition and related forcing data.

and the intercept is

$$\beta_0 = \bar{y} - \beta_1 \cdot \bar{x} \tag{5}$$

where $\bar{x}$ and $\bar{y}$ are the mean of prediction and observations respectively.

The inference of the slope and the intercept can be worked out by predetermining the upper and lower bounds for their estimates. We can compute the confidence intervals (CI) of of the slope and the intercept at a certain certain confidence level (say α) . Thereafter, we will test, at the level α previously set. , whether a certain value (zero, one or any other) lies inside or outside the range set in (CI) for the slope and intercept estimates The upper and lower limits of the values for the slope can be estimated as follows:

$$\begin{aligned}
H &= \frac{F_{\alpha,1,n-2}}{((L1/L2) + (L2/L1) - 2) \cdot (n-2)} && (6)\\
A &= \sqrt{\frac{H}{1-H}} && (7)\\
\beta_1 low &= \frac{\beta_1 - A}{1 + \beta_1 \cdot A} && (8)\\
\beta_1 up &= \frac{\beta_1 + A}{1 - \beta_1 \cdot A} && (9)\\
& && (10)
\end{aligned}$$

where $F_{\alpha,1,n-2}$ is the quantile for the Fdistribution at level α and degrees of freedom 1 and $(n-2)$, being n the numnber of pairs of observations and $\beta_1 low$ and $\beta_1 up$ are the upper and lower limit of the slope at the α significance level. For the inference of the intercept, we should operate as follows:

$$\begin{aligned}
b_0 low &= \bar{y} - L1 \cdot \bar{x} && (11)\\
b_0 up &= \bar{y} + L2 \cdot \bar{x} && (12)\\
& && (13)
\end{aligned}$$

The estimation of standard errors for the intercept and the slope in the MAR is more complicated. Warton et al. (2006) give details.