Estimating a smooth common transfer function with a panel of time series – inflow of larvae cod as an example

Elizabeth Hansen,[1] **Kung-Sik Chan**[1,*] **and Nils Chr. Stenseth**[2]

[1]Department of Statistics & Actuarial Science,

University of Iowa, Iowa City, IA 52245, U.S.A.

[2] Centre for Ecological and Evolutionary Synthesis (CEES),

Department of Biology, University of Oslo,

P.O. Box 1050 Blindern, N-0316 Oslo, NORWAY

April 2004

SUMMARY. We consider the problem of estimating a smooth common transfer function shared by a panel of short time series that are contemporaneously correlated. We propose an estimation scheme using a likelihood approach that penalizes the roughness of the common transfer function. For a known smoothness parameter, the estimation can be done via an iterative procedure. The method of cross-validation can be used to determine the smoothness parameter. We illustrate the proposed method with a biological example of indirectly estimating the spawning date distribution of North Sea cod. Some simulation results are reported on the empirical performance of the proposed method.

Key words: cod-spawning-date distribution, (generalized) cross-validation, seemingly unrelated regression, multimodality

*email: kung-sik-chan@uiowa.edu

1. Introduction

Consider the following stochastic regression model describing how the response depends on the aggregate effects of a covariate:

$$Y_t = \alpha^T W_t + \sum_{j=m_1}^{m_2} \psi_j X_{t,j} + e_t, t = 1, 2, \cdots, T, \qquad (1)$$

where Y_t are the responses, W_t and $X_{t,j}$ are vector-valued and scaler-valued covariates, α and ψs are parameters and $\{e_t\}$ is a sequence of independent and identically distributed random variables of zero mean and finite variance; the superscript T denotes the transpose. The sampling scheme is as follows. Both Y and W are measured over regular, basic sampling intervals. Each basic sampling interval is further sub-divided into, say M, equal intervals over each of which X is measured. In the biological application to be discussed in section 4, Y and W were measured annually whereas X was measured daily. The measurement of X in the jth sub-interval of the tth basic sampling interval is denoted by $X_{t,j}$. The summation limits m_1 and m_2 are assumed to be known integers. The model defined by (1) is also known as a transfer-function model (Box et al., 1994) or distributed-lag model (Almon, 1965).

The main interest is to estimate ψ_j as a function of j. Often, the functional form of ψ is unknown. Empirical parametric models such as the rational transfer function model and the Almon polynomial lag model are popular methods for estimating ψ, but they are less useful with complex functional forms. For example, in our biological application, ψ may be a multimodal function, in which case both the rational transfer function model and the Almon polynomial lag model require many parameters for providing an adequate description of ψ. Shiller (1973) introduces a nonparametric approach for estimating a smooth ψ function by postulating a smoothness prior on the second difference of ψ, but otherwise putting no constraints on ψ. See, also, Kitagawa and Gersch (1996). (Shiller also discussed briefly the use of higher differences, but we shall not pursue this point here.) Specifically, it is assumed that

$$(1-B)^2\psi_j = \eta_j, \tag{2}$$

where B is the backshift operator defined by $B\psi_j = \psi_{j-1}$, and η_j are iid normal with zero mean and variance $\sigma_\eta^2 > 0$. That is, a hierarchical model is employed. Shiller discussed both the use of a fully Bayesian analysis with non-informative priors as well as a sort of empirical Bayes approach where σ_η^2 is specified by some rule of thumb. Note that the approach introduced by Shiller is similar to spline smoothing; see Wahba (1990), Wood (2000) and Gu (2002). The nonparametric approach of Shiller can cope with complex functional form of ψ.

However, even the nonparametric approach fails if the number of data cases is small compared to $m_2 - m_1 + 1$, the number of lags of X appearing in the model. This problem may be circumvented if there exist a panel of time series that share the same transfer function so that information can be pooled across series for estimating ψ. Here, we consider this situation so that the sth series is generated by the model:

$$Y_{t,s} = \alpha^T W_{t,s} + \sum_{j=m_1}^{m_2} \psi_j X_{t,j} + e_{t,s}, t = 1, 2, \cdots, T, \tag{3}$$

where we note that the same X's enter into the equation for each component series, but W and e may vary across series. For the panel data, the errors are often contemporaneously correlated although they may be serially independent. Here, we "extend" Shiller's approach to a multivariate stochastic regression model with contemporaneously correlated errors that subsumes the common transfer function model defined by (3). However, our approach differ from Shiller's approach in that we use a penalized likelihood approach (Green and Silverman, 1994).

We now outline the organization of the rest of the paper. In section 2, we elaborate on the framework of a multivariate stochastic regression model that subsumes the common transfer function model. A cross-validation approach is outlined for estimating the smoothness parameter. Some large-sample properties of the estimator are derived. In section 3, a simulation study is reported where the simulation model is motivated by the real application. In the simulation study we investigate the empirical power of our approach for detecting multimodality in the ψ function. In section 4 we apply the proposed method to a biological data set motivating the study. In particular, we estimate the probability density function of the egg spawning date of North Sea cod indirectly based on data on sea current, spawning biomass and counts of half-year old cod in eight fjords in Southern Norway. We briefly conclude in section 5.

2. A multivariate stochastic regression model

Consider the following general regression model with multivariate response and covariate.

$$\mathbf{Y}_t = X_t\beta + \mathbf{e}_t;\ t = 1, \ldots, T \tag{4}$$

where the dimension of $\mathbf{Y}_t$ is $n \times 1$, X_t is $n \times k$, and the coefficient vector β is $k \times 1$. The $\mathbf{e}_t$'s are independent and identically distributed as normal with mean zero and variance-covariance matrix Ω, and $\mathbf{e}_t$ is independent of X_t. (Here, we restrict the errors to be normally distributed for convenience; extension to non-normality is straight-forward but it will complicate the iterative estimation procedure below.) This model is rather general and includes the common transfer-function model discussed at the end of the preceding section. Consider the case that the dimension of β is high compared to the sample size resulting in multicollinearity. The multicollinearity problem can be mitigated by exploiting some known "smoothness" property of β. Suppose that the roughness of β can be quantified by the Euclidean norm of $\eta = A\beta$ where A is a known $m \times k$ matrix. (In the case of the common transfer function model defined by (3), $A\beta$ is the vector of second differences of ψ.) We can now construct a penalized log-likelihood where a quadratic penalty for $A\beta$ is used:

$$\ell(\beta) = -\frac{T}{2}\log|\Omega| - \frac{1}{2}\sum_{t=1}^{T}(\mathbf{Y}_t - X_t\beta)^T\Omega^{-1}(\mathbf{Y}_t - X_t\beta) - \frac{1}{2}\frac{\beta^T A^T A\beta}{\sigma_\eta^2}, \quad (5)$$

where the coefficient $\sigma_\eta^2 > 0$ quantifies the trade-off between badness of fit and roughness of the parameter; σ_η^2 will be determined by the method of cross-validation. Here, the penalized log-likelihood has the Bayesian interpretation that the components of η have joint prior independent and identical normal distribution of zero mean and variance σ_η^2.

Notice that if $\Omega = [\Omega_{ij}]$ and the smoothness parameter $\tau^2 = \frac{\Omega_{11}}{\sigma_\eta^2}$ are known, β can be estimated by minimizing the following expression.

$$\begin{aligned} &\Omega_{11}\left\{\sum_{t=1}^{T}(\mathbf{Y}_t - X_t\beta)^T\Omega^{-1}(\mathbf{Y}_t - X_t\beta) + \frac{\beta^T A^T A\beta}{\sigma_\eta^2}\right\} \\ = &\sum_{t=1}^{T}\left\{\left(\frac{\Omega}{\Omega_{11}}\right)^{-\frac{1}{2}}\mathbf{Y}_t - \left(\frac{\Omega}{\Omega_{11}}\right)^{-\frac{1}{2}}X_t\beta\right\}^T\left\{\left(\frac{\Omega}{\Omega_{11}}\right)^{-\frac{1}{2}}\mathbf{Y}_t - \left(\frac{\Omega}{\Omega_{11}}\right)^{-\frac{1}{2}}X_t\beta\right\} \\ &+\tau^2\beta^T A^T A\beta. \end{aligned} \quad (6)$$

The introduction of Ω_{11} in the definition of τ^2 is to make the latter interpretable as the signal-to-noise ratio; alternatively, we could use $\mathrm{tr}(\Omega)/n$, i.e. the average variance, instead of Ω_{11}, where for any square matrix A, $\text{tr}(A)$ is the sum of diagonal elements of A. Assume for the moment that the smoothness parameter τ is known. Define $L = (\Omega/\Omega_{11})^{-\frac{1}{2}}$. Then the preceding penalized sum of least squares can be minimized by an iterative process. The iterative procedure bears resemblance to the method of seemingly unrelated regression technique, see Zellner (1962) and Hamilton (1994). We first find $\hat{\beta}^{(0)}$ by regressing $\mathbf{Y}_t$ on X_t. For $i = 0, 1, 2, \ldots$, compute $\hat{\Omega}^{(i)} = \frac{1}{T}\sum_{t=1}^{T}(\mathbf{Y}_t - X_t^T\hat{\beta}^{(i)})(\mathbf{Y}_t - X_t^T\hat{\beta}^{(i)})^T$ and $\hat{L}^{(i)}$ obtained via the Cholesky decomposition of $\hat{\Omega}^{(i)}$. Define

$$\mathbf{Y}^{(i)} = \begin{pmatrix} \hat{L}^{(i)}\mathbf{Y}_1 \\ \hat{L}^{(i)}\mathbf{Y}_2 \\ \vdots \\ \hat{L}^{(i)}\mathbf{Y}_T \\ \mathbf{0} \end{pmatrix}, X^{(i)} = \begin{pmatrix} \hat{L}^{(i)}X_1 \\ \hat{L}^{(i)}X_2 \\ \vdots \\ \hat{L}^{(i)}X_T \\ \tau A \end{pmatrix}.$$

Next, we update $\hat{\beta}^{(i+1)}$ by regressing $\mathbf{Y}^{(i)}$ on $X^{(i)}$. The iterative procedure can be stopped by using some stopping criteria, e.g., when the relative change in the L^1-norm of the $\beta^{(i)}$ or the objective function defined in (5) is smaller than some prespecified tolerance level. At the end of the iteration and letting $\hat{\Omega}$ be the estimator of Ω, it can be readily checked that

$$\hat{\beta} \;=\; (A^TA/\hat{\sigma}_\eta^2 + \sum_t X_t^T\hat{\Omega}^{-1}X_t)^{-1}\sum_t X_t^T\hat{\Omega}^{-1}Y_t, \tag{7}$$

where $\hat{\sigma}_\eta^2 = \tau^2/\hat{\Omega}_{11}$. Hence, the asymptotic covariance matrix of $\hat{\beta}$ is approximately given by $(A^TA/\hat{\sigma}_\eta^2 + \sum_t X_t^T\hat{\Omega}^{-1}X_t)^{-1}(\sum_t X_t^T\hat{\Omega}^{-1}X_t)(A^TA/\hat{\sigma}_\eta^2 + \sum_t X_t^T\hat{\Omega}^{-1}X_t)^{-1}$.

The smoothness parameter τ can be determined by minimizing the cross-validation (CV) or the generalized cross-validation (GCV); see Wahba (1990) and Green and Silverman (1994). In order to visualize the calculation of

CV or GCV, let $\mathbf{Y}$ be the vector obtained by stacking up the Y_t's, $\mathbf{X}$ the corresponding design matrix and $\hat{\mathbf{Y}}$ the fitted values so that

$$\hat{\mathbf{Y}} = \mathbf{X}\hat{\beta} = H\mathbf{Y}, \tag{8}$$

where the matrix $H = [h_{ij}]$ is the hat matrix equal to $\mathbf{X}W$ where W is implicitly defined by (7) in the vector form $\hat{\beta} = W\mathbf{Y}$. The values of CV $=$ $CV(\tau)$ and GCV $= GCV(\tau)$ can be found as follows.

$$CV(\tau) = \sum_{i=1}^{n} \frac{(Y_i - \hat{Y}_i)^2}{(1 - h_{ii})^2}, \tag{9}$$

$$GCV(\tau) = \sum_{i=1}^{n} \frac{(Y_i - \hat{Y}_i)^2}{(1 - \frac{1}{n}\sum_{j=1}^{n} h_{jj})^2}. \tag{10}$$

In order to estimate τ using this technique, we will start the optimization at a value that "equate" the information from the data with that from the smoothness "prior", i.e. the initial value $\tilde{\tau}$ is set to be

$$\tilde{\tau} \doteq \frac{\mathrm{tr}(\sum_t X_t^T X_t)}{\mathrm{tr}(A^T A)}. \tag{11}$$

In other words, the initial value corresponds to a probably over-smoothed model with the amount of smoothing having as much weight as the data information. Since the value of τ is nonnegative, the optimization can be more easily done by applying Newton's method on $\log(\tau)$ with a starting value of $\log(\tilde{\tau})$ to find a minimum on either the CV or GCV functions. Once an optimum value of τ is found, it can be used in the procedure shown above.

3. Simulation

We investigate the empirical performance of the proposed method by simulations. For the cod example in section 4, the ψ function appears to be bimodal. This motivates us to study two matters: the number of significant modes detected and the error involved in the estimates. We consider six cases of the following model motivated by the cod example:

$$\mathbf{Y}_t = \kappa + \zeta b_t + \psi^T \mathbf{c}_t b_t + \mathbf{e}_t; t = 1, \ldots, T. \tag{12}$$

The vector $\mathbf{Y}_t$ is of the dimension $F \times 1$, and the dimension of the vectors ψ and $\mathbf{c}_t$ is $D \times 1$ for both. The error vector $\mathbf{e}_t$ has a multivariate normal distribution with mean zero and variance-covariance matrix $\sigma_e^2 I$. In the estimation, we impose the smoothness constraints

$$\psi_i - 2\psi_{i+1} + \psi_{i+2} = \eta_{i+2}; \;\; i = 1, \ldots, D-2, \tag{13}$$

as well as end constraints:

$$\begin{aligned} \psi_1 &= \eta_1 \\ -2\psi_1 + \psi_2 &= \eta_2 \\ \psi_{D-1} - 2\psi_D &= \eta_{D+1} \\ \psi_D &= \eta_{D+2}, \end{aligned}$$

to ensure that the ψ function estimates are smooth across the boundaries beyond which they are zero.

For each of the two examples below, $T = 50$, $F = 3$, $D = 61$, $\kappa = [1, 0, -1]^T$, and $\zeta = 1$. The value of b_t is determined by taking a random number from a normal distribution with mean zero and standard deviation one, while the values of $\mathbf{c}_t$ are determine by drawing sixty-one random numbers from a normal distribution with zero mean and unit variance. The error standard deviation was either 0.05 or 0.1, and the ψ_j equal the probability density function at j of an equal mixture of two normal distributions, namely $N(30-\Delta, 9)$ and $N(30+\Delta, 4), \Delta = 10, 5, 0$. Hence ψ has two modes that are separated by either 20 units, 10 units, or 0 units (thus making one mode). Each case was simulated 1,000 times. The plots of the three different sets of ψs used can be found in figure 1.

Recall the simulation was used to study two matters: the number of significant modes detected and the error involved in the estimates. Results of both of these can be found in table 1. In regards to the first matter, the simulation was able to detect unimodality well in the case where the modes were not separated. In the cases where there was separation between the modes, bimodality was detected about 64% of the time when the modes were clearly separated and/or the error variability was low. (Note that we count the number of modes only for the ψ_j's that are significantly different from zero, hence there could be no mode in the curve if none of the ψ_j's are significant.) As far as the second matter, the mean absolute deviation and mean deviation were small as compared to the maximum value of the ψ's being estimated. Their standard deviations were small as well and depended proportionately on the error variance.

[Figure 1 about here.]

[Table 1 about here.]

4. Inflow of larvae cod as an example

Recent genetic analysis by Knutsen et al. (2004) suggested that the young (half-year old) cod sampled in some fjords in the Skagerrak, Norway, resembled adult cod in the North Sea in year 2001 but less so in year 2000. It was, furthermore, found that in 2001 when the sampled young cod of Skagerrak were genetically similar to the adult cod of North Sea, there was higher than average inflow of sea current from the North Sea to the Skagerrak, but not so in 2000 when the resemblance switched to the local adult cod. Thus, Knutsen et al. (2004) suggested the hypothesis that the North Sea cod stock might have contributed to the local cod population in the Skagerrak via transportation of cod eggs by sea current from North Sea into the Skagerrak. Stenseth et al. (2004) tested this hypothesis using a long-term monitoring beach seine data on the annual counts of young cod, the (annual) spawning biomass of North Sea cod and daily inflow of sea current from North Sea to Skagerrak. It is believed that the cod spawn, or breed, in the months of March and April, but it is not known specifically when the majority of the spawning took place. Stenseth et al. (2004) computed the average daily inflow (from North Sea to the Skagerrak) over several windows of 2-week period between March and April, and tested the transportation hypothesis using a regression model with a covariate that is the product of average sea influx times log spawning biomass, a proxy for the transportable amount of cod eggs, the coefficient of which is non-zero under the transportation hypothesis and zero otherwise. Stenseth et al. (2004) found that the transportation hypothesis is consistent with the data, with stronger, significant result when the mean inflow is computed over the second half of March. Clearly, which two-week period over which the mean inflow is computed is critical as the test can be made more powerful by aligning the period with the main period when the cod spawned.

The developed method allows us to study the problem from a different perspective. Instead of searching for an optimal window for averaging the daily inflow, we consider the distribution of the cod spawning date. Let S be the date, counted from the beginning of March of each year, when a randomly selected adult cod spawn. Let ψ_j be the probability that $S = j$. The daily contribution of North Sea cod to the Skagerrak is postulated to additively contribute, on the logarithmic scale, to the young cod counts by an amount proportional to $\psi_j c_{t,j} b_t$ where b_t is the log spawning biomass in the tth year and $c_{t,j}$ be the mean inflow on the jth day of the tth year; for simplicity of notation, the proportional constant is absorbed into ψ_j so that they need not sum to 1. In other words, the total annual North-Sea-cod contribution equaled $\sum_{j=1}^{61} \psi_j c_{t,j} b_t$, under the transportation hypothesis.

Let $n^0_{t,s}$ be the logarithm of the number of young cod caught in fjord s in year t. We confine the analysis to eight fjords in the Southern Norway, over the period from 1971 to 1997 over which we have complete data. These eight fjords are reported in the earlier analysis by Stenseth et al. (2004) to admit significant transportation effects. In fact, $n^0_{t,s}$ are part of a longer residual series from a stochastic regression model using a longer database that has adjusted for the intra-specific and the inter-specific effects, as well as the environmental effects on the local cod in the Skagerrak; see Chan et al. (2003a), Chan et al. (2003b) and Stenseth et al. (2004). We now state the model.

$$n^0_{t,s} = \kappa_s + \zeta b_t + \sum_{j=1}^{61} \psi_j c_{t,j} b_t + e_{t,s}; \; t = 1, \ldots, 26; \; s = 1, \ldots, 8. \qquad (14)$$

The κ_s's can be interpreted as the fjord-specific effect on the cod population, and they may be expected to be close to zero because $n^0_{t,s}$ are part of a long residual series. The term ζb_t can be interpreted as the contributions of the North Sea adult cod by directly swimming to the Skagerrak and spawning there.

Our prior knowledge of the amount of spawning that occurs outside of the months of April and March (see, e.g., Knutsen et al., 2004) allows us to impose the following end constraints:

$$\begin{aligned} \psi_1 &= \eta_1 \\ -2\psi_1 + \psi_2 &= \eta_2 \\ \psi_{D-1} - 2\psi_D &= \eta_{D+1} \\ \psi_D &= \eta_{D+2}, \end{aligned} \tag{15}$$

where $D = 61$. These end constraints merely incorporate the prior assumption that ψ is zero beyond March and April, and maintain the constraint of small roughness across the boundaries. We estimated the model using the method proposed in section 2. Figure 2 plots the ψ function where the central curvy line is the estimated curve and the other two lines enclose the individual 95% confidence limits. The value of τ estimated by the method shown in section 2 is 2,178, with the initial value being 37,037, as determined by the procedure outlined at the end of section 2. Clearly, there is a significant spike in the spawning that begins on March 15th and ends March 27th. To check the robustness of the ψ estimates against the window width, we have also re-done the analysis with the ψs specified as zero outside the period from March 15th to April 8th. It was then found that the estimates are almost unchanged and hence not reported. In conclusion, there is clear evidence that sea current transported the North Sea cod eggs to the Skagerrak, mainly over the second half of March. Furthermore, the data suggest that the North Sea adult cod did not swim to the Skagerrak to spawn there. These conclusions are consistent with the findings of Stenseth et al. (2004) that is obtained by assuming constant ψ over 2-week periods. However, our new method allows for much more refined conclusions of great importance to the field of marine ecology.

[Figure 2 about here.]

[Table 2 about here.]

5. Conclusion

We have demonstrated the potential usefulness of the method of penalized likelihood for estimating a smooth common transfer function with a panel of short, contemporaneously correlated errors. As illustrated with our marine example, the new method provides refined conclusion within the field of marine ecology, particularly with reference to how different populations of a marine fish species are interlinked through larvae inflow. As such, our results are of direct relevance for studies on the ecological effects of climate change (see, e.g., Stenseth et al., 2002).

There are a few interesting future research problems. First, it is of interest to work out the case of non-normal errors in greater details. Second, the common transfer function assumption is a strong one. A more flexible approach is to incorporate random-effects in the transfer function model.

Acknowledgements

We are grateful to Chong Gu for valuable suggestions. The empirical work referred to above was done in collaboration with colleagues at the Institute of Marine Research, Flødevigen IMR, Lysekil and Tjärnö Marine Biological Laboratory; that work was funded by the Norwegian Research Council, the Swedish Council for the Environment and the Nordic Council of Ministers supported this work. Work reported in this paper was supported financially by the Norwegian Research Council, the University of Oslo and the University of Iowa BSI program.

References

Almon, S. (1965). The distributed lag between capital appropriations and expenditures. *Econometrica* **33**, 178 – 196.

Box, G., Jenkins, G. and Reinsel, G. (1994). *Time Series Analysis, Forcasting and Control (3rd edition).* Pearson Education, San Fransisco, California, USA.

Chan, K., Stenseth, N., Kittilsen, M., Gjøsæter, J., Lekve, K., Smith, T., Tveite, S. and Danielssen, D. (2003a). Assessing the effectiveness of releasing cod larvae for stock improvement: using long-term beach-seine monitoring data to resolve a century-long controversy. *Ecological Applications* **13**, 3 – 22.

Chan, K., Stenseth, N., Lekve, K. and Gjøsæter, J. (2003b). Modeling pulse disturbance impact on cod population dynamics; the 1988 algae bloom of Skagerrak, Norway. *Ecological Monographs* **73**, 151 – 171.

Green, P. and Silverman, B. (1994). *Nonparametric Regression and Generalized Linear Models.* Chapman and Hall, London, England.

Gu, C. (2002). *Smoothing Spline ANOVA Models.* Springer-Verlag, New York, U.S.A.

Hamilton, J. D. (1994). *Time Series Analysis.* Princeton University Press, Princeton, New Jersey, USA.

Kitagawa, G. and Gersch, W. (1996). *Smoothness Priors Analysis of Time Series.* Springer-Verlag, New York, USA.

Knutsen, H., André, C., Jorde, P., Skogen, M., Thuróczy, E. and Stenseth, N. (2004). Transport of north sea cod larva into the Skagerrak coastal populations. *Proceedings of the Royal Society of London, B* .

Shiller, R. (1973). A distributed lag estimator derived from smoothness priors. *Econometrica* **41**, 775 – 788.

Stenseth, N. C., Atle, M., Geir, O., Hurrell, J. W., Chan, K.-S. and Lima, M. (2002). Ecological effects of climate fluctuations. *Science* **297**, 1292–1296.

Stenseth, N., Jorde, P., Chan, K., H., K., André, C., Skogen, M. D. and Lekve, K. (2004). Ecological and genetic impact of larval drift. *Biology Letter* (under review).

Wahba, G. (1990). Spline models for observational data. *Philadelphia: SIAM.*

Wood, S. (2000). Modelling and smoothing parameter estimation with multiple quadratic penalties. *Journal of Royal Statistical Society* **B62**, 413–428.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association* **57**, 348–368.

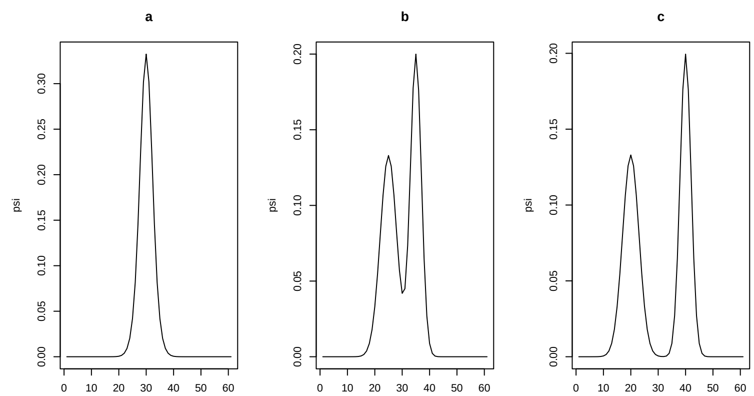

Figure 1. Plots of ψ_j versus j for the simulation model; ψ_j is the probability density function of the an equal mixture of $N(30-\Delta, 9)$ and $N(30+\Delta, 4)$, where $\Delta = 0, 5, 10$, from left to right.

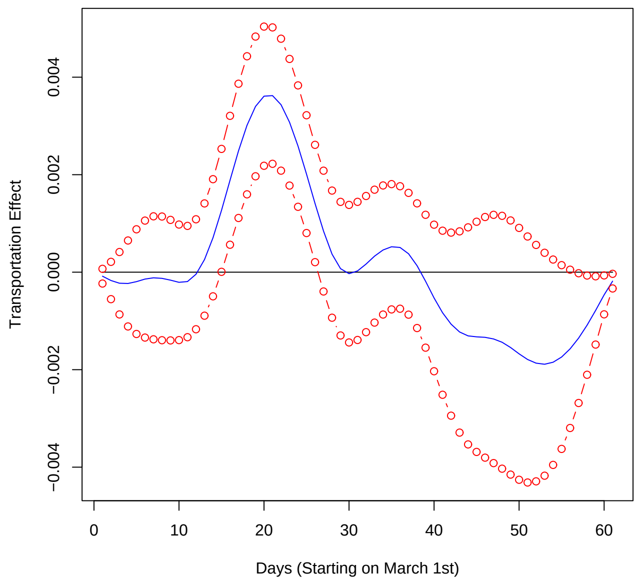

Figure 2. The plot of $\hat{\psi}_j$ versus j for the North Sea cod – the central curvy line. The other two curves envelope the individual 95% confidence limits.

standard deviation = 0.05				standard deviation = 0.1			
	Δ				Δ		
	10	5	0		10	5	0
no. of modes	% of significant modes			no. of modes	% of significant modes		
0	0.5%	10.5%	9.6%	0	0.6%	6.9%	10.5%
1	98.2%	24.0%	19.7%	1	94.0%	46.6%	25.6%
2	1.3%	63.2%	68.9%	2	4.9%	42.6%	60.0%
≥ 3	0.0%	2.3%	1.8%	≥ 3	0.5%	3.9%	3.9%
mean abs. dev.	0.0281	0.0256	0.0252	mean abs. dev.	0.0353	0.0298	0.0307
SD	0.00480	0.00464	0.00546	SD	0.00961	0.01031	0.01275
mean deviation	0.0166	0.0167	0.0163	mean deviation	0.0162	0.0159	0.0159
SD	0.00528	0.00497	0.00491	SD	0.00902	0.00906	0.00970

Table 1

Simulation results for the model defined by (12).

Parameter	Estimate	SE
κ_1	-0.028	0.151
κ_2	-0.042	0.160
κ_3	0.170	0.147
κ_4	-0.049	0.131
κ_5	-0.230	0.165
κ_6	-0.246	0.100
κ_7	-0.122	0.132
κ_8	0.097	0.128
ζ	-0.007	0.008

Table 2

Estimates of model for the Skagerrak cod.