

## REGULARIZATION OF A PARAMETER ESTIMATION PROBLEM USING MONOTONICITY AND CONVEXITY CONSTRAINTS \*

SAM SUBBEY<sup>1</sup>

**Abstract.** In marine science, it is usually assumed that there is a functional relationship between the parental population size and subsequent offsprings. The function is referred to as the Stock Recruitment Function (SRF). Determining the SRF translates to the optimization problem of estimating a set of parameters using past and sparse observation, which are usually of modest accuracy. The problem is challenging because several candidate functions exist in the literature, and the choice of best function is non-trivial, due to data sparsity and uncertainty.

This paper formulates the problem as a constrained optimization task, and uses B-spline basis functions to represent the functional family to which the SRF belongs. Regularized solutions are obtained by requiring that the derived functions are both monotone and convex.

The approach presents two major contributions to the existing computational challenges:

- It avoids the non-trivial problem of choosing the functional form *a priori*.
- Regularization of the problem using constraints ensures that parameter estimates are realistic.

Numerical examples are presented to compare  $\ell_1$  and  $\ell_2$ -norm solutions.

**Résumé.** The paper presents a method that uses *a priori* information to regularize an inverse problem in fisheries science. It demonstrates the efficacy of the methodology in dealing with sparse and uncertain data, and limiting assumptions on the solution space.

### 1. INTRODUCTION

This paper is particularly concerned with the fisheries science problem of determining a functional relationship between the size of offsprings (referred to as recruits) and the size (usually biomass) of their parental population. The functional relationship, referred to as the stock recruitment function (SRF) is a characteristic of the population, irrespective of whether the population is being exploited. To estimate the functional relationship, one is usually given the following:

- a set of  $N \in \mathbb{N}_{>0}$  observations of the parental population size  $\tilde{\mathbf{x}} \equiv \{\tilde{x}_{t_1-\tau}, \tilde{x}_{t_2-\tau}, \dots, \tilde{x}_{t_N-\tau}\}$  and corresponding offsprings  $\tilde{\mathbf{y}} \equiv \{\tilde{y}_{t_1}, \tilde{y}_{t_2}, \dots, \tilde{y}_{t_N}\}$ , where  $0 \leq \tau < t_1 < \dots < t_N$ , and  $\tau$  is usually assumed known, and derivable from the life history of the stock, and
- a set of  $m \in \mathbb{N}_{>0}$ , real-valued candidate functions  $\mathbf{f} \equiv \{f_1, \dots, f_m\}$ , where for each  $f \in \mathbf{f}$ ,  $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} : \tilde{\mathbf{y}} \approx f(\boldsymbol{\theta}, \tilde{\mathbf{x}})$ .  $\boldsymbol{\theta} \in \mathbb{R}^d$  is a set of unknown parameters, and  $d$  depends on the choice of  $f \in \mathbf{f}$ .

The general SRF problem can then be formulated as the unconstrained optimization Problem 1.

---

\* Funding is acknowledged from the following IMR Programs: MarPro (Project # 14412), and Barents Sea (Project # 84126).

<sup>1</sup> Inst. of Marine Research, Pb 1870, N-5817 Bergen, Norway. Hjort Centre for Marine Ecosystem Dynamics, Bergen, Norway.

© EDP Sciences, SMAI 2017

**Problem 1** (The general SRF problem). *Given observations  $\tilde{\mathbf{x}} \in \mathbb{R}^N$  and  $\tilde{\mathbf{y}} \in \mathbb{R}^N$ , determine the function  $f(\boldsymbol{\theta}_p^*, \mathbf{x}) \in C^2$ , where*

$$\boldsymbol{\theta}_p^* = \underset{\boldsymbol{\theta} \in \mathbb{R}^d}{\operatorname{argmin}} \|\tilde{\mathbf{y}} - f(\boldsymbol{\theta}, \tilde{\mathbf{x}})\|_p. \quad (1)$$

In general,  $f$  is expected to be nonlinear since it subsums several stochastic ecological processes [20]. In the literature, choosing an appropriate function form has been identified as perhaps the most challenging in the solution of Problem 1 [38]. This is partly because there are no robust statistical techniques for exploring the fundamental biological processes in the SRF. Furthermore, information theoretic measures such as the Akaike (AIC) and Bayesian (BIC) [5] information criteria may fail in providing unique guidelines to the choice of  $f$ , see e.g. [36]. Simulation studies have also shown that the AIC can lead to a choice different from that which generated the observation data [9]. According to [20]: *the most important and generally most difficult problem in biological assessment of fisheries is the relationship between stock and recruitment*. Problem 1 is thus central to quantifying fish population dynamics.

In the fisheries literature, the most common functional forms are the classical Beverton-Holt [2] and Ricker [31] functions, which are given respectively by (2) and (3). For  $\alpha, \beta \in \mathbb{R}_{>0}$ ,

$$f_1(\alpha, \beta, x) = \frac{\alpha x}{1 + \beta x}, \quad (2)$$

$$f_2(\alpha, \beta, x) = \alpha x e^{-\beta x}. \quad (3)$$

Note that  $f_1$  and  $f_2$  are derived from different ecological assumptions (see [20]). Given the sparsity of the data, it is usually difficult to determine which assumption underlie the observations. In practice, the choice of functional form can be very arbitrary.

Consistent with the notations introduced, we write

$$\boldsymbol{\theta} \equiv \{\alpha, \beta\}, \quad (4)$$

$$\mathbf{f} \equiv \left\{ \frac{\alpha x}{1 + \beta x}, \alpha x e^{-\beta x} \right\}. \quad (5)$$

Using (4) and (5), this paper seeks to address issues related to (i) functional representation of the SRF, and (ii) parameter estimation, concentrating on the most common functions and using artificially generated data as proxy for real observations. The particular drawbacks with current approaches to these problems are discussed shortly, followed by a presentation of the alternative approach, which is the subject matter of this paper.

## 1.1. Limitations to Current Estimation Methodology

### *Choice of functional representation*

In most practical cases, the choice of function ( $f_1$  or  $f_2$ ) is arbitrary since it is based on visual inspection of the (sparse) data. Furthermore, the use of analytical (parametric) functions with limited flexibility constrains the solution space of the problem. There is no guarantee therefore, that the derived function is an optimal representation of the observation data. A viable alternative is to represent the SRF by a non-parametric function. This approach has been reported by [4], and [6], who used the R-SCAM package [28, 29] to fit a non-parametric function to the observation data. Reparametrization using a flexible function (e.g., B-spline basis functions) has the attraction that the design points (spline knots) can be chosen independent of the observation. If too high number of basis functions are used, oscillatory (non-monotonic) and biologically implausible solutions result. In the R-SCAM package, regularization of the problem is addressed by adding a term in the objective function with a smoothing parameter, which penalizes divergence from smoothness. The smoothing parameter, which can be determined by e.g. generalized cross-validation (GCV), is usually data dependent [35]. More generally, the choice of penalizing term and how it is determined, will dictate the type of solution obtained.

For instance, methods such as GCV and unbiased risk (UBR) have been developed under the assumption that the data is from independent observations [36]. When the independent observation assumption is violated, the results obtained are underestimates of the optimal smoothing parameter.

#### Parameter estimation

In general,  $\ell_p$  estimators involve minimization of the  $p$ -Minkowsky metric (see [11, 12])  $d^p(\boldsymbol{\delta})$  for  $p \geq 1$ , defined by (6), where  $\boldsymbol{\delta} \in \mathbb{R}^n$  is a vector representing the difference between two points in  $\mathbb{R}^n$ .

$$d^p(\boldsymbol{\delta}) = \left[ \sum_{i=1}^n |\delta_i|^p \right]^{\frac{1}{p}}. \quad (6)$$

Partly because of mathematical convenience and computational ease, and also because  $\ell_2$ -norm approaches (and generalizations) have well known Gaussian properties, this norm has enjoyed unparalleled popularity in statistics and fisheries science. It is therefore the standard solution norm for Problem 1. Unfortunately the  $\ell_2$ -norm poorly performs poorly when the data are derived from a heavy-tailed distribution or sparse, see e.g. [1]. In particular, a single outlier or extreme value can have remarkable influence on the solution. Hence the  $\ell_2$ -norm solution is not guaranteed to be robust, given the nature of the observation data.

## 2. A NEW COMPUTATIONAL FRAMEWORK

The goal of the new computational approach is threefold – (i) to circumvent the problem of making an *a priori* choice of the SRF, (ii) avoid the use of smoothing parameters for regularization of the problem, and (iii) reduce bias in estimates of parameters, caused by the existence of few extreme values in the observation data.

The approach is to first examine the common characteristics (e.g., continuity, monotonicity, convexity, etc.) of the function set defined by  $\mathbf{f}$ . In the next step, a non-parametric function is used to represent the members of  $\mathbf{f}$ , where the common characteristics are imposed as constraints. The non-parametric function is then fitted to the observation data, and the optimization problem solved to obtain the  $\ell_1$ - and  $\ell_2$ -norm solutions. The  $\ell_1$ -norm is known to be robust (see e.g. [12]) to the influence of data that deviate largely from the mean. Hence it is particularly suited for this problem, given the nature (sparsity and uncertainty) of the data.

Rather than the traditional approach of choosing between functions, the aim is to develop a generic functional solution, and then determine which ecological assumption (Beverton-Holt or Ricker) best fits the solution obtained. The generic solution must therefore have characteristics of both recruitment functions. We introduce a variable transformation and show that the transformed functions are continuous, monotone and convex. Recall that  $x > 0$  (only positive parental biomass) and  $f(\boldsymbol{\theta}, x) \geq 0$  (only positive biomass of offsprings). Now, define  $u(\boldsymbol{\theta}, x) \in \mathbb{R}$  by (7).

$$x \cdot u(\boldsymbol{\theta}, x) = f(\boldsymbol{\theta}, x). \quad (7)$$

It is straightforward to prove that  $u(\boldsymbol{\theta}, x)$  satisfies (8).

$$u(\boldsymbol{\theta}, x) \geq 0, \quad \frac{\partial u(\boldsymbol{\theta}, x)}{\partial x} \leq 0, \quad \frac{\partial^2 u(\boldsymbol{\theta}, x)}{\partial x^2} \geq 0. \quad (8)$$

I introduce the following definition (see [14]):

**Definition 1.** (*Monotonicity and Convexity*): Let  $g(x)$  be a continuous function on an arbitrary interval  $I$  and at least twice differentiable on its interior. Then  $g(x)$  is decreasing (monotone) if  $g'(x) \leq 0$  and convex iff  $g''(x) \geq 0$ ,  $\forall x$  in the interior of  $I$ ,

The function  $u(\boldsymbol{\theta}, x)$  is thus positive, monotone and convex with respect to  $x$ . Table 1 summarizes components of  $\boldsymbol{\theta}$  for  $f_1$  and  $f_2$  defined in (2) and (3). Let  $u_1$  and  $u_2$  define the transformations corresponding to  $f_1$  and  $f_2$ , respectively. Table 1 summarizes revised expressions for the parameters  $\alpha$  and  $\beta$ .

TABLE 1. Summary of derived functional expressions

$j$	$u_j(\boldsymbol{\theta}, x)$	$u'_j(\boldsymbol{\theta}, x) = \frac{\partial u_j(\boldsymbol{\theta}, x)}{\partial x}$	$\alpha = \boldsymbol{\theta}(1)$	$\beta = \boldsymbol{\theta}(2)$
1	$\frac{\alpha}{1+\beta x}$	$-\frac{\beta}{\alpha} u^2(\boldsymbol{\theta}, x)$	$[1 + x\beta] u(\boldsymbol{\theta}, x)$	$-\frac{u'(\boldsymbol{\theta}, x)}{xu'(\boldsymbol{\theta}, x)+u(\boldsymbol{\theta}, x)}$
2	$e^{\alpha-\beta x}$	$-\beta u(\boldsymbol{\theta}, x)$	$\ln u(\boldsymbol{\theta}, x) + x\beta$	$-\frac{u'(\boldsymbol{\theta}, x)}{u(\boldsymbol{\theta}, x)}$

**Remark 2.1.** *Characteristics of  $\alpha$  and  $\beta$*

Recall that for the classical Ricker and Beverton-Holt function, the parameters  $\alpha$  and  $\beta$  are strictly positive constants. In the approach adopted in this paper  $\alpha$  and  $\beta$  are no more constants but dynamic parameters, which are functions of instantaneous values of  $x$ . The approach does not guarantee that  $\alpha(x), \beta(x) > 0, \forall x \geq 0$ .

**2.1. An Approximation Problem**

The approach is to select  $n$  basis functions  $v_1, \dots, v_n$  and approximate  $u$  by  $\tilde{u}$  such that

$$\tilde{u}(\mathbf{c}, x) \approx \sum_{r=1}^n c_r v_r(x), \tag{9}$$

where  $c_r \in \mathbb{R}$  are coefficients to be determined, and introduce a *merit number* to define the goodness of fit associated with a feasible vector  $\mathbf{c}$ . Define the vector  $\boldsymbol{\delta} \in \mathbb{R}^N$  by setting

$$\delta_i = U_i - \sum_{r=1}^n c_r v_r(x_i), \quad i = 1, \dots, N, \tag{10}$$

$$U_i = f(\boldsymbol{\theta}, x_i)/x_i. \tag{11}$$

Thus  $\delta_i \in \boldsymbol{\delta}$  is the difference between the measured value  $U_i$  and the predicted value given by the sum in (10). Next, define an ordering among the  $N$ -dimensional vectors  $\boldsymbol{\delta}$  by selecting a convex, nonnegative function  $d$  which is defined on  $\mathbb{R}^N$  and such that

$$d(\boldsymbol{\delta}) \geq 0, \quad d(\boldsymbol{\delta}) = 0 \Rightarrow \boldsymbol{\delta} = \mathbf{0}. \tag{12}$$

Hence  $d(\boldsymbol{\delta})$  is a norm on  $\mathbb{R}^N$  and a merit number associated with  $\boldsymbol{\delta}$ , which in turn is defined by  $\mathbf{c}$  via (10). We arrive at the approximation task (where the strict inequalities are relaxed):

**Problem 2** (Constrained Optimization Problem). *We seek a solution to the problem:*

$$\boldsymbol{\delta}^* = \underset{\boldsymbol{\delta} \in \mathbb{R}^N}{\operatorname{argmin}} d(\boldsymbol{\delta})$$

subject to

$$\sum_{r=1}^n c_r v_r(x) \geq 0, \quad \sum_{r=1}^n c_r v'_r(x) \leq 0, \quad \sum_{r=1}^n c_r v''_r(x) \geq 0, \tag{13}$$

where

$$d(\boldsymbol{\delta}) = \sum_{i=1}^N |\delta_i|^p, \quad \delta_i = U_i - \tilde{u}(\mathbf{c}, x_i), \quad \tilde{u}(\mathbf{c}, x_i) = \sum_{r=1}^n c_r v_r(x_i),$$

The constraints in (13) are consistent with (8). The aim is to choose the basis functions  $v_r$  such that the constraints may be represented as a finite set of linear conditions. Then Problem 2 takes the form of a linear program (LP) having finitely many variables and constraints. This may be done using finite elements and polynomial basis functions. This paper used normalized B-splines, as choices of  $v_r$ .

## 2.2. Using B-Spline Basis Functions

By restricting the solution to quadratic splines  $v(x)$ , a sufficient and necessary condition for monotonicity [34] is an ordering of the spline coefficients according to

$$1 \geq c_1 \geq c_2 \geq c_3 \cdots \geq c_m \geq 0, \quad m \leq n. \quad (14)$$

Equation (14) could be written in a matrix form as

$$G\mathbf{c} \geq \mathbf{s}, \quad (15)$$

$$\begin{bmatrix} -1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & \cdots & 0 & 0 \\ \vdots & & & & & & \\ 0 & 0 & 0 & \cdots & 0 & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_{m-1} \\ c_m \end{bmatrix} \geq \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \quad (16)$$

where  $G \in \mathbb{R}^{(m+1) \times m}$ , and  $\mathbf{s}, \mathbf{c} \in \mathbb{R}^{m \times 1}$ . Convexity condition requires that

$$v_r''(x_i) \geq 0, \quad i = 0, 1, \dots, N.$$

However, since  $v_r(x)$  is piecewise quadratic,  $v_r''(x)$  is independent of  $x$ , and depends only on the coefficient of the quadratic term in  $x$ . The convexity constraint then translates into a matrix inequality equation given by

$$H\mathbf{c} \geq \mathbf{w}, \quad (17)$$

where  $H \in \mathbb{R}^{(m-2) \times m}$ , and  $\mathbf{w} \in \mathbb{R}^{(m-2) \times 1}$  is a zero vector. The elements of  $H$  are defined by

$$H(l, j) = \begin{cases} 2a_{l,j}, & j = l(1)l + 2 \\ 0, & \text{Otherwise.} \end{cases}$$

where  $l$  is an index of the inter-nodal interval for which the coefficient of the quadratic term,  $a_{l,j}$ , applies to spline number  $j$ .

## 2.3. The $\ell_2$ -norm Problem in Standard Form

We use the matrices  $G$ , and  $H$ , and the vectors  $\mathbf{s}$  and  $\mathbf{w}$  to express Problem 2 in a standard form.

**Problem 3** ( $\ell_2$ -norm Optimization Problem in Standard Form). *We seek a solution to the problem:*

$$\boldsymbol{\delta}^* = \underset{\boldsymbol{\delta} \in \mathbb{R}^N}{\operatorname{argmin}} d(\boldsymbol{\delta})$$

subject to

$$\begin{bmatrix} G \\ H \end{bmatrix} \mathbf{c} \geq \begin{bmatrix} \mathbf{s} \\ \mathbf{w} \end{bmatrix} \quad (18)$$

where the elements of  $G$ ,  $H$ ,  $\mathbf{s}$ ,  $\mathbf{w}$  and  $d(\boldsymbol{\delta})$  are defined as previously.

For  $p = 2$ , Problem 2 can be solved with constrained optimization algorithms (see e.g. [37]).

### 2.4. The $\ell_1$ -norm Optimization Problem

In practice, the quality of the measured values is unknown. Furthermore, some observations could deviate considerably from the average. Hence a robust measure of the deviation between the measured values  $U_i$  and the calculated ones,  $\tilde{u}(\mathbf{c}, x_i)$ , is required. This is achieved by using the  $\ell_1$ -norm, i.e. for  $p = 1$ .

Let (see [16])

$$|\delta_i| = \delta_i^+ + \delta_i^-, \quad i = 1, \dots, N,$$

where  $\delta_i^+, \delta_i^- \geq 0$  and  $\delta_i = \delta_i^+ - \delta_i^-$ . Problem 2 is then equivalent to the following task:

**Problem 4** (General  $\ell_1$  approximation with side constraints). *Solve:*

$$\operatorname{argmin}_{\delta^+, \delta^- \in \mathbb{R}^N} \sum_{i=1}^N (\delta_i^+ + \delta_i^-),$$

subject to

$$A\mathbf{c} + \boldsymbol{\delta}^+ - \boldsymbol{\delta}^- = \mathbf{U}, \quad \begin{bmatrix} G \\ H \end{bmatrix} \mathbf{c} \geq \begin{bmatrix} \mathbf{s} \\ \mathbf{w} \end{bmatrix}, \quad \mathbf{c} \geq 0, \quad \boldsymbol{\delta}^+ \geq 0, \quad \boldsymbol{\delta}^- \geq 0.$$

Here  $A \in \mathbb{R}^{N \times n}$  contains the coefficients of the approximation problem involving the basis function, and all other matrices and vectors are defined previously.

The formulation of the  $\ell_1$ -norm optimization problem is such that it can be solved with the same algorithm types as for the  $\ell_2$ -norm (see [16]).

## DESIGN OF NUMERICAL EXPERIMENTS

The numerical experiments are designed to test the ability of the framework to reconstruct the true (but unknown) function  $f$ , based on corrupted data  $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ , and to compare  $\ell_1$  to  $\ell_2$  solutions.

In the first step, synthetic data of recruitment ( $f$ ) was generated at 20 uniformly distributed values of  $x \in [0.125 \ 1.00]$ . Note that for real cases, this simply implies rescaling of the  $x$ -axis by the maximum observed  $x$  value. Proxy observation data,  $\tilde{u}$ , was generated by adding random noise to values of  $f$ , according to:

$$\begin{aligned} \tilde{x} &= |x + \mathcal{N}(0, \sigma)|, \\ \tilde{y} &= f(\boldsymbol{\theta}, \tilde{x}), \end{aligned}$$

where  $\mathcal{N}(0, \sigma)$  represents normally distributed pseudo-random numbers with zero mean and standard deviation  $\sigma$ . The sought function as approximated using quadratic splines with six uniformly spaced knots  $t = [0.0 \ 0.25 \ 0.50 \ 0.75 \ 1.0]$ . We derive the functional relationship between  $f$  and  $x$  based on the data  $(\tilde{\mathbf{x}}, \tilde{\mathbf{y}})$ , for different error levels, represented by  $\sigma$ . The  $p$ -norm errors (see (6)),  $d^p(\boldsymbol{\delta})$  ( $p = 1, 2$ ) are reported for each solution. Several numerical simulations have been conducted with good results, especially for low  $\sigma$  values. The results presented here are for high values of uncertainty in  $f$ , induced by the uncertainty in  $x$ .

## RESULTS AND DISCUSSION

Figure 1 shows example plots of the original,  $f(x)$ , rescaled,  $u(x)$ , and derivative,  $u'(x)$ , for the Beverton-Holt and Ricker functions, while Fig. 2 shows example spline basis functions used in this manuscript.

Figure 3 shows results for the Ricker function with  $\sigma = 0.25$  (Fig. 3a) and 0.5 (Fig. 3b) and Table 2 summarizes the performance statistics. The results in Fig. 3a show identical performance for  $\ell_1$  and  $\ell_2$  (see Table 2), especially for  $d^2$ . This is not surprising, given the low variability in the data nearer to the function maximum and the tail. In contrast, Fig. 3b shows large observation data variability, notably at the tail. Large data variability, especially at the maximum and tail, presents a special estimation (and inference) challenge to current methodologies. For this particular case, observe that the  $\ell_2$  solution is highly influenced by the extreme values at the tail whereas the  $\ell_1$  appears to be unaffected (see also the performance statistics in Table 2).

Figures 3c-d show plots of values for  $\alpha$  and  $\beta$  as functions of  $x$ . Observe that in the approximate interval  $0.4 \leq x \leq 0.6$  (Fig. 3a) and  $0.5 \leq x \leq 0.8$  (Fig. 3b),  $\alpha(x) \leq 0$ . Thus for these particular intervals, it is impossible to reconcile the data with any of the biological assumptions underpinning either the Ricker or Beverton-Holt models. Observe (contrast Fig. 4c.-d.) to Fig. 3c.-d.) that for these particular examples,  $\alpha, \beta > 0, \forall x \in [0, 1]$ . Another way to view the parameters  $\alpha(x)$  and  $\beta(x)$  is that they serve to partition the  $f(x)-x$  plane, representing different recruitment regimes.

For the Beverton-Holt model, cases for  $\sigma < 0.5$  are not of particular interest since simulation results show that the true recruitment functions are easily recovered with high precision. The results obtained with higher uncertainties (in this case,  $\sigma \geq 0.5$ ) are however variable and Fig. 4 shows example simulation results for  $\sigma = 0.5$ . Table 3 summarizes the performance statistics for the Beverton-Holt function.

This paper has presented a simple approach to fitting analytical models of the Ricker and Beverton-Holt types to observation data. In contrast to current methods, the computational approach adopted allows the data to dictate its own best functional representation. Though the emphasis has been on the classical Beverton-Holt and Ricker recruitment models, the computational framework is a generic solution to fitting analytic functions to recruitment data.

A significant contribution from this paper is the treatment of model parameters  $\theta$ , as dynamics, rather than static, as is current practice. There are several arguments in the literature, in support of the approach in this paper. Fish populations can be considered as dynamic systems subject to diverse perturbations including natural mortality, predation and harvesting. Recruitment to a fish stock is often characterized by periodic very large values and very occasionally exceptionally low values [20]. In addition, the productivity of a system may shift either up or down resulting in a sustained change in the underlying stock and recruit relationship [32]. This is sometimes referred to as a regime shift [27]. It is therefore inconceivable that the stock parameters  $\alpha$  and  $\beta$  will remain invariant even when the recruitment  $f$  changes (non-linearly or otherwise) with the spawning stock biomass  $x$ . Since these parameters have management relevance for many stocks, the approach presented permits inference on whether the recruitment dynamics remains invariant over the whole range of observed spawning stock biomasses. Further it is natural to expect that a precautionary management (see e.g. [15]) will seek to dampen the influence of extreme value observations (or outliers) on management decisions. However, outlier identification and removal is not always possible in most practical cases. This is either because of the low number of observations (typical for stock recruitment data) or that in some cases, it is impossible or impractical to separate outliers from observations. For stock recruitment data, there is also a third possibility. Outliers may represent different states of the recruitment dynamics. The computational framework presented in this paper accommodates cases representing all three possible scenarios of extreme observations, and also lends interpretation to these scenarios through analysis of the parameters  $\alpha$  and  $\beta$ , which characterize the recruitment dynamics. This is particularly simple in the computational framework presented since  $\alpha$  and  $\beta$  are expressed as analytical functions of stock size.

A sequel manuscript (in preparation) will seek to apply the computational framework in this paper to stock recruitment data across several fish stocks and ecosystems. This sequel manuscript will attempt to give more ecological interpretation to the dynamic parameters,  $\alpha(x)$  and  $\beta(x)$ .

TABLE 2. Performance statistics for the Ricker function

$\sigma = 0.25$			$\sigma = 0.50$		
Solution norm	$d^1$	$d^2$	Solution norm	$d^1$	$d^2$
$\ell_1$	3.0866	0.5540	$\ell_1$	2.2224	0.4226
$\ell_2$	2.6708	0.5733	$\ell_2$	5.4578	1.1960

TABLE 3. Performance statistics for the Beverton-Holt function

$\sigma = 0.50$ , Fig. 4a			$\sigma = 0.50$ , Fig. 4b		
Solution norm	$d^1$	$d^2$	Solution norm	$d^1$	$d^2$
$\ell_1$	2.0223	0.3782	$\ell_1$	6.3351	1.0908
$\ell_2$	2.7072	0.6233	$\ell_2$	8.2869	1.5407

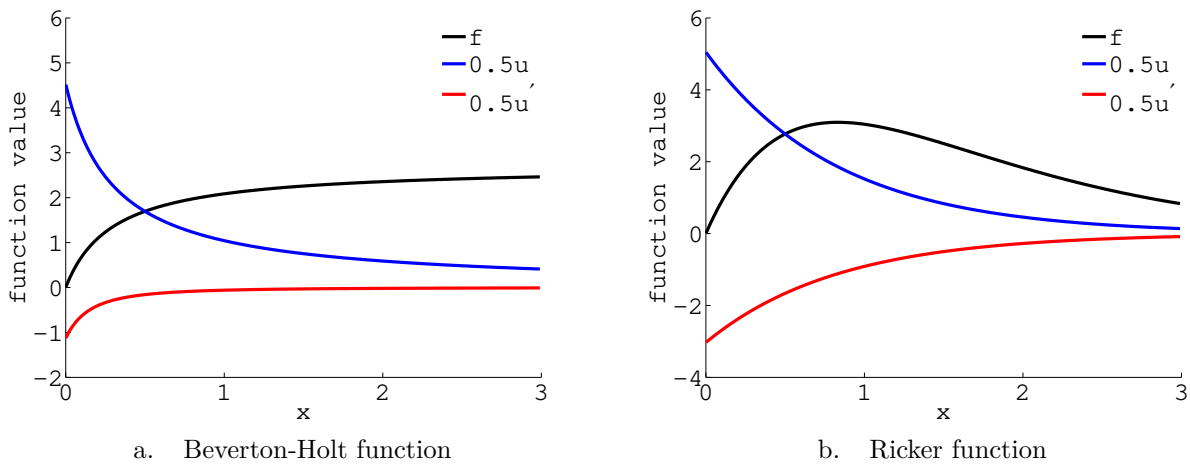


FIGURE 1. Examples plots of the original,  $f(x)$ , rescaled,  $u(x)$ , and derivative,  $u'(x)$ , for the Beverton-Holt ( $\alpha = 2.7$ ,  $\beta = 0.3$ ) and Ricker ( $\alpha = 10.1$ ,  $\beta = 1.2$ ) functions. Note that  $u(x)$  and  $u'(x)$  have been rescaled for the sake of comparison.



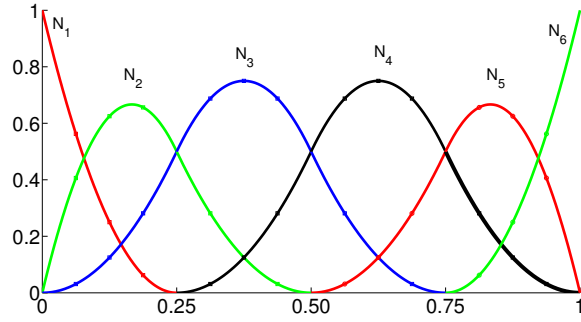


FIGURE 2. Example normalized quadratic spline function with 5 equally spaced knots.

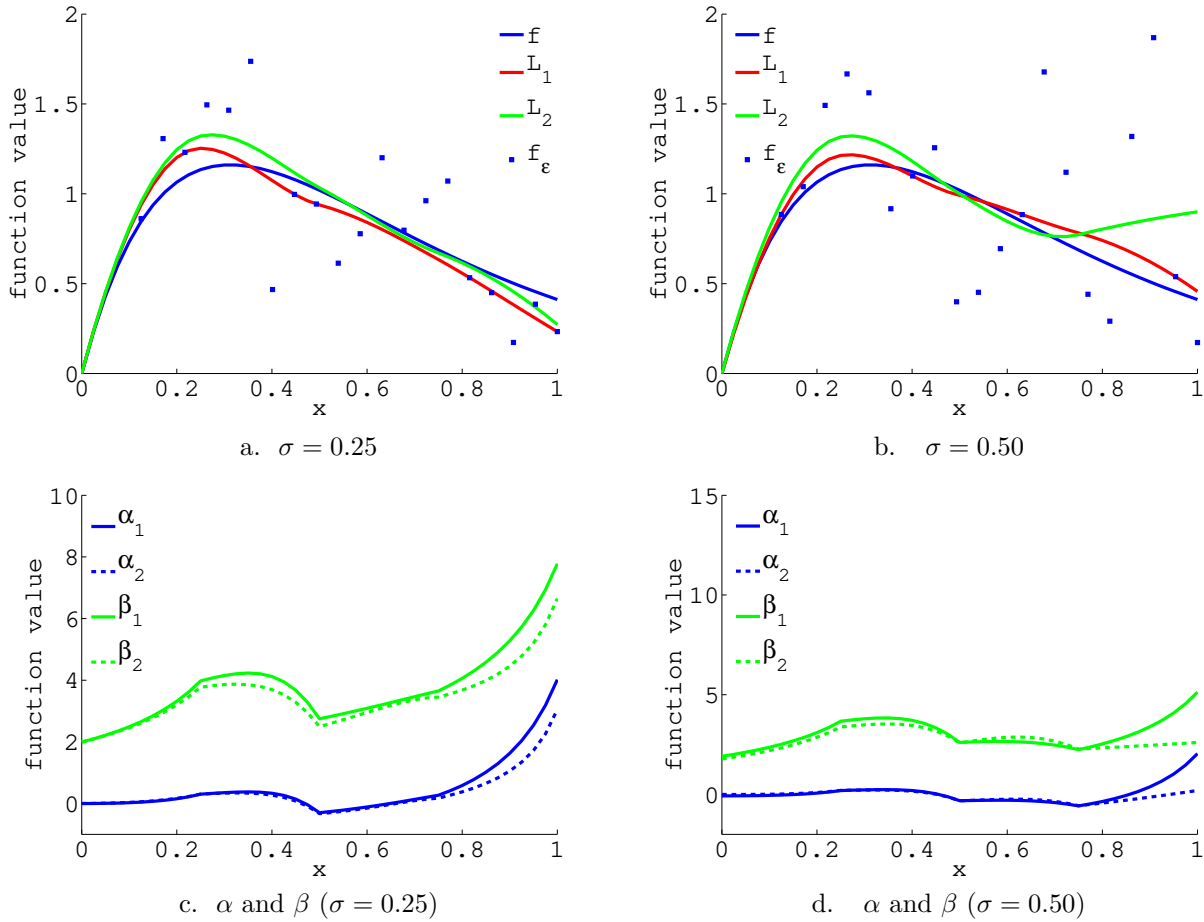


FIGURE 3. Example results for the Ricker function.  $f$ ,  $L_1$ ,  $L_2$ ,  $f_\epsilon$  represent respectively, the true analytical function, the solutions in the  $\ell_1$  and  $\ell_2$  norms, and the *experimental* data. The parameters  $\alpha_i$  and  $\beta_i$  refer to the  $\ell_i$ -solution values for  $\alpha$  and  $\beta$ , ( $i = 1, 2$ ).

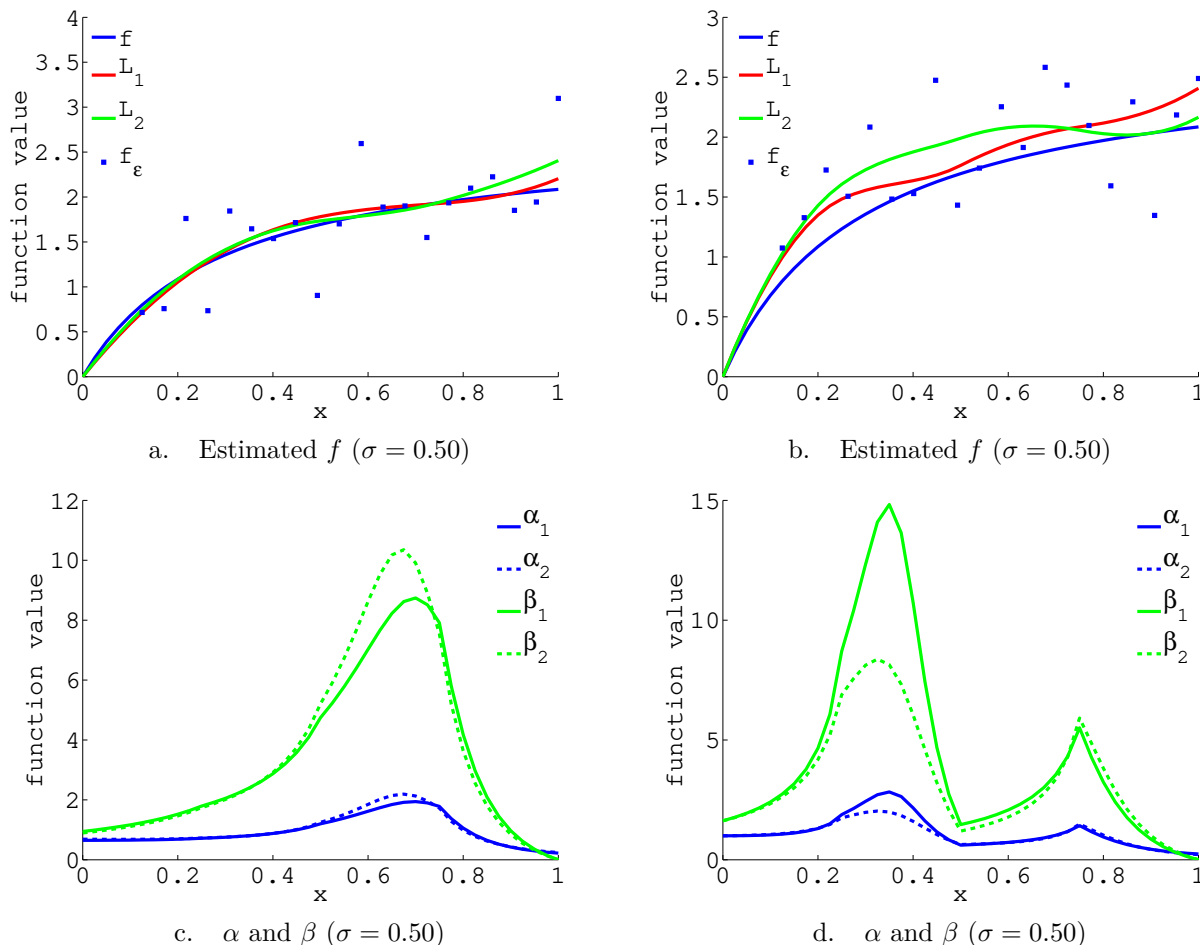


FIGURE 4. Example results for the Beverton-Holt function.  $f, L_1, L_2, f_\epsilon$  represent respectively, the true analytical function, the solutions in the  $\ell_1$  and  $\ell_2$  norms, and the *experimental* data. The parameters  $\alpha_i$  and  $\beta_i$  refer to the  $\ell_i$ -solution values for  $\alpha$  and  $\beta$ , ( $i = 1, 2$ ).

### REFERENCES

[1] Alarcon-Aquino, V. et al. Wavelet-network based on L1-norm minimisation for learning chaotic time series. *Journal of Applied Research and Technology*, 2010, vol. 3, nr. 03.

[2] Beverton, R. J. H. and Holt, S. J. On the Dynamics of Exploited Fish Populations. *Springer Fish and Fisheries* 11, 1993.

[3] Box, G. E. P., et al. *Time series analysis: forecasting and control*. John Wiley & Sons, vol. 734, 2011.

[4] Bravingotn, M. V. and Stockes, T. K. and O'Brien, C. M. Sustainable recruitment: the bottom line, *Marine and Freshwater Research*, CSIRO, 2000, vol. 51, nr. 5, p. 465–475.

[5] Burnham, K. P. and Anderson, D. R. Multimodel inference understanding AIC and BIC in model selection. *Sociological methods & research*, 2004, vol. 33, nr. 2, p. 261–304.

[6] Cadigan, N. G. Fitting a non-parametric stock–recruitment model in R that is useful for deriving MSY reference points and accounting for model uncertainty, 2013, *ICES Journal of Marine Science*, Oxford University Press, vol. 70, nr. 1, p. 56–67.

[7] Charles, A. T. Living with uncertainty in fisheries: analytical methods, management priorities and the Canadian groundfishery experience. *Fisheries Research*, 1998, vol. 37 nr. 1, p. 37–50.

[8] Chen, D. G. Bias and bias correction in fish recruitment prediction. *North American Journal of Fisheries Management*, 2004, vol. 24, nr. 2, p. 724–730.

- [9] De Valpine, P. and Hastings, A. Fitting population models incorporating process noise and observation error. *Ecological Monographs*, 2002, vol. 72, nr. 1, p. 57–76.
- [10] Deriso, R. B. Harvesting strategies and parameter estimation for an age-structured model. *Canadian Journal of Fisheries and Aquatic Sciences*, 1980, vol. 37, nr. 2, p. 268–282.
- [11] Dodge, Y. An introduction to L1-norm based statistical data analysis. *Computational Statistics & Data Analysis*, 1987, vol. 5, nr. 4, p. 239–253.
- [12] Dodge, Y. (Ed.). *Statistical Data Analysis based on the  $L_1$ -norm and Related Methods*. Springer, 2002.
- [13] Draper, N. R., Smith, H. and Pownell, E. *Applied regression analysis*. New York: Wiley, 1966.
- [14] Dudley, R. M. Real analysis and probability. Vol. 74. Cambridge University Press, 2002.
- [15] Garcia, S. M. The precautionary approach to fisheries and its implications for fishery research, technology and management: an updated review, *FAO Fisheries Technical Paper*, 1996, 350.
- [16] Glashoff, K. and Gustafson, S.-Å. *Linear Optimization and Approximation: An Introduction to the Theoretical Analysis and Numerical Treatment of Semi-infinite Programs*. Springer-Verlag, 1983.
- [17] Graybill, F. A. Theory and application of the linear model. Wadsworth and Brooks/Cole. 1976.
- [18] Hammond, T. R. and O'Brien, C. M. An application of the Bayesian approach to stock assessment model uncertainty. *ICES Journal of Marine Science*, 2001, vol. 58, nr. 3, p. 648–656.
- [19] Hilborn, R. Simplified calculation of optimum spawning stock size from Ricker's stock recruitment curve. *Canadian Journal of Fisheries and Aquatic Sciences*, 1985, vol. 42, nr. 11, p. 1833–1834.
- [20] Hilborn, R., et al. Quantitative fisheries stock assessment: choice, dynamics and uncertainty. *Reviews in Fish Biology and Fisheries*, 1992, vol 2, Nr. 2, p. 177–178.
- [21] Jeffrey, D. J., et al. Integration of the signum, piecewise and related functions. In: *Proceedings of the 1997 international symposium on Symbolic and algebraic computation*. ACM, 1997, p. 324–330.
- [22] Liu, W. et al. Extended reconstructed sea surface temperature version 4 (ERSST. v4): Part II. Parametric and structural uncertainty estimations, *Journal of Climate*, 2015, vol. 28, nr.3, p.931–951.
- [23] Ludwig, D. and Walters, C. J. Measurement errors and uncertainty in parameter estimates for stock and recruitment. *Canadian Journal of Fisheries and Aquatic Sciences*, 1981, vol. 38, nr. 6, p. 711–720.
- [24] Ludwig, D. and Walters, C. J. Optimal harvesting with imprecise parameter estimates. *Ecological Modelling*, 1982, 14. Jg., Nr. 3, S. 273-292.
- [25] Patterson, K. R. Evaluating uncertainty in harvest control law catches using Bayesian Markov chain Monte Carlo virtual population analysis with adaptive rejection sampling and including structural uncertainty. *Canadian Journal of Fisheries and Aquatic Sciences*, 1999, vol. 56, nr. 2, p. 208–221.
- [26] Patterson, K. et al. Estimating uncertainty in fish stock assessment and forecasting. *Fish and Fisheries*, 2001, vol. 2, nr. 2, p. 125–157.
- [27] Payne, M. R. et al. Recruitment in a changing environment: the 2000s North Sea herring recruitment failure. *ICES Journal of Marine Science*, 2009, vol. 66, nr. 2, p. 272–277.
- [28] Pya, N. Additive models with shape constraints, PhD dissertation, 2010, University of Bath, UK.
- [29] Pya, N. *Package 'scam'* @ <http://CRAN.R-project.org/>, 2013.
- [30] Quinn II, T. J. and Deriso, R. B. *Quantitative fish dynamics. Biological Resource Management Series. XV*. Oxford University Press: New York, USA, 1999, 542pp.
- [31] Ricker, W. E. *Computation and interpretation of biological statistics of fish populations*. Dept. of Fisheries and Oceans, vol.191, pg. 382. 1987.
- [32] Rothschild, B. J. Fish stocks and recruitment: the past thirty years. *ICES Journal of Marine Science*, 2000, vol. 57, nr. 2, p. 191–201.
- [33] Schnute, J. T. and Richards, L. J. Use and abuse of fishery models. *Canadian Journal of Fisheries and Aquatic Sciences*, 2001, vol. 58, nr. 1, pgs 10–17.
- [34] Schumaker, L. L. Spline functions: Basic Theory. 1981. *John Wiley & Sons*, New York.
- [35] Wahba, G. A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem, *The Annals of Statistics*, JSTOR, 1985, vol. 13, nr. 4, p. 1378–1402.
- [36] Wang, Y. and Liu, Q. Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC) in selection of stock–recruitment relationships. *Fisheries Research*, 2006, vol. 77, nr. 2, p. 220–225.
- [37] Wright, S. and Nocedal, J. Numerical optimization. Springer Science 35 (1999): 67-68.
- [38] Zhou, S. Discriminating alternative stock–recruitment models and evaluating uncertainty in model structure. *Fisheries Research*, 2007, vol. 86, nr. 2, p. 268–279.