



**Abstract**—Natural mortality ( $M$ ) is one of the most important life history attributes of functioning fish populations. The most common methods to estimate  $M$  in fish populations provide point estimates which are usually constant across sizes and ages. In this article, we propose a framework for incorporating uncertainty into the length-based estimator of mortality that is based on von Bertalanffy growth function (VBGF) parameters determined with Bayesian analysis and asymmetric error distributions. Two methods to incorporate uncertainty in  $M$  estimates are evaluated. First, we use Markov chains of the estimated VBGF parameters directly when computing  $M$  and second, we simulate the posterior distribution of VBGF parameters with the copula method. These 2 approaches were applied and compared by using the extensive database available on age and growth for southern blue whiting (*Micromesistius australis*) harvested in the southeast Pacific. The copula approach provides advantages over Markov chains and requires far less computational time, while conserving the underlying dependence structure in the posterior distribution of the VBGF parameters. The incorporation of uncertainty into length-based estimates of mortality provides a promising way for modeling fish population dynamics.

Manuscript submitted 30 August 2016.  
Manuscript accepted 11 April 2017.  
Fish. Bull. 115:355–364 (2017).  
Online publication date: 18 May 2017.  
doi: 10.7755/FB.115.3.6

The views and opinions expressed or implied in this article are those of the author (or authors) and do not necessarily reflect the position of the National Marine Fisheries Service, NOAA.

## Incorporating uncertainty into a length-based estimator of natural mortality in fish populations

**Freddy O. López Quintero**<sup>1</sup>

**Javier E. Contreras-Reyes (contact author)**<sup>2,3</sup>

**Rodrigo Wiff**<sup>4</sup>

Email address for contact author: [javier.contreras@ifop.cl](mailto:javier.contreras@ifop.cl)

<sup>1</sup> Departamento de Matemática  
Instituto Venezolano de Investigaciones Científicas  
Carretera Panamericana, Km 11, Altos de Pipe  
Estado Miranda 1020-A, Venezuela

<sup>2</sup> División de Investigación Pesquera  
Instituto de Fomento Pesquero  
Avenida Almirante Blanco Encalada 839  
Valparaíso 2361827, Chile

<sup>3</sup> Instituto de Estadística  
Universidad de Valparaíso  
Avenida Gran Bretaña 1111  
Valparaíso 2360102, Chile

<sup>4</sup> Center of Applied Ecology and Sustainability  
Pontificia Universidad Católica de Chile  
Avenida Libertador Bernardo O'Higgins 328  
Santiago 8331150, Chile

Natural mortality ( $M$ ) rate is one of the most important parameters shaping the population dynamics of fish populations (Siegfried and Sansó<sup>1</sup>; Brodziak et al., 2011). It is defined as the death rate of fish due to causes other than fishing, such as predation, senescence, cannibalism, starvation, and other natural factors. Despite the key importance of  $M$  in fish and fisheries modeling, this parameter is extraordinarily difficult to estimate accurately. Methods for determining  $M$  in fish populations generally entail one of 2 approaches: 1) direct methods which estimate  $M$  from observations on survival, with methods derived from tagging or telemetry experiments, 2) indirect methods that estimate mortality from other, more easily obtained parameters, often from life history traits, such as age and growth, and maturity. Direct methods provide the most precise estimates of  $M$ , but those approaches

are data intensive and usually cost prohibitive and therefore preclude their application to a large number of fish stocks. Indirect methods are therefore commonly applied because they are easy, fast, and cheap to implement for most harvested fish populations.

Indirect estimators are based on correlations of  $M$  in well-studied stocks with other life history attributes, such as individual growth, longevity and maturity. The underlying assumption in all indirect methods is that the relationship between  $M$  and other life history parameters is the same for thoroughly studied stocks and lesser-studied ones where this method is usually applied. For most indirect methods, an estimate of  $M$  is computed, which is invariant across age and size classes, although this parameter is dependent on age and size (Gislason et al., 2010). The relationship between  $M$  and age and size is usually defined as a negative exponential function in which early life stages include much greater mortality than later ones, especially after reaching sexual maturity. Nevertheless,  $M$  estimators, whether they

<sup>1</sup> Siegfried, K. I., and B. Sansó. 2009. A review for estimating natural mortality in fish populations. Southeast Data, Assessment, and Review SEDAR 19-RD29, 31 p. [Available from [website](http://www.seafoodwatch.org).]

are size dependent or not, provide only point estimates of mortality. Uncertainty in  $M$  estimates from indirect methods have not been investigated in detail and have usually been based on ad-hoc approaches (Cubillos et al., 1999; Quiroz et al., 2010; Wiff et al., 2011). These methods incorporate uncertainty in  $M$  by taking growth parameters from the literature and their associated uncertainty (e.g. standard deviation, covariance, confidence interval), and then by drawing empirical distributions of these parameters to propagate uncertainty in  $M$  estimates, usually assuming Gaussian error.

A promising method for estimating  $M$ -at-length based on life history theory, has recently been proposed by Gislason et al. (2010) and Charnov et al. (2013), and depends entirely on von Bertalanffy growth function (VBGF) parameters. For these researchers, appropriate uncertainty in growth estimates become a key issue in assessing uncertainty in  $M$  estimates. Most ad-hoc approaches to incorporate statistical variability in  $M$  are based on a 2-step model that keeps estimations of growth parameters and  $M$  separated. Therefore, in this article, we consider the inclusion of 2 different sources of uncertainty in the modeling, one over the growth parameters and another from the proposed structural  $M$ -at-length mortality estimator.

Uncertainty in  $M$  estimates derived from indirect methods come from 2 main sources. First, uncertainty depends on the variability among species or stocks for which the empirical relationship has been proposed. This source of uncertainty is usually referred to as “method error” because it represents how “accurate” the empirical model is (Quiroz et al., 2012). The second source of uncertainty is related to the error within the species-specific parameters that feeds into the indirect method (e.g., growth parameters). This source of uncertainty is called “trait error” because it represents the uncertainty in life history parameters of the stock or population for which  $M$  is being estimated.

The trait-error of estimated growth parameters can be converted into  $M$  estimates as the iterative parameter updates of a Bayesian estimation with a non-Gaussian distribution. In particular, Gaussian error distribution for estimating growth parameters can resolve some even nonsensical shortcomings, for instance negative length values. Distributions based on age-length fishing selectivity are usually skewed as a result of a size-selective sampling process (Contreras-Reyes et al., 2014; Montenegro and Branco, 2016). In addition, in harvested fish populations, an accumulative effect of fishing on size-at-age exists. Growth rates vary among individuals (Sainsbury, 1980) and fishing selectivity removes faster growing individuals from each particular age class. Moreover, in some studies, in which VBGF parameters are estimated, the assumption of Gaussian error distribution sometimes lacks adequacy, especially with the presence of outliers, which could lead to questionable estimates (see Contreras-Reyes and Arellano-Valle, 2013; Contreras-Reyes et al., 2014, and references therein).

In our analysis we used 2 methods to incorporate

uncertainties in the  $M$ -at-length estimates. Both depend on information that can be drawn from a Bayesian estimation derived from VBGF parameters. The first method takes advantage of the Markov chains and, after convergence, those chains are directly incorporated into the  $M$  calculation. In particular, we used the Bayesian results generated by López Quintero et al. (2017) as a baseline. For the second method, we propose taking the dependent structure, which is concentrated in the posterior distribution of parameters, and using this structure for simulating a distribution with the same dependent features based on the copula method (Nelsen, 2006). This last approach has the advantage that, even without the precision in the joint distribution of parameters, we can obtain samples that preserve the dependence between the observed variables by only approximating its marginal distributions, at a much lower computational cost. We apply the length-based  $M$  estimators, using a data set corresponding to 24,942 individuals of southern blue whiting (*Micromesistius australis*) collected from Chilean continental waters over the period 1997–2010 (J. Contreras-Reyes, unpubl. data). Their total lengths and ages were recorded and assigned by studying their otoliths. Further information about this data set can be found in Contreras-Reyes et al. (2014) and references therein.

## Material and methods

### Growth model and $M$ -at-length

Let  $L(x)$  be the theoretical expected value of the length related to an individual at age  $x$ . The (specialized) VBGF function is defined as

$$L(x) = L_{\infty}(1 - e^{-K(x-t_0)}). \quad (1)$$

This equation represents the simplest formulation of the VBGF (Essington et al., 2001), which is described by 3 parameters:  $L_{\infty}$  represents asymptotic length (in length units e.g. centimeters);  $K$  represents the growth rate coefficient usually expressed in inverse time units; and  $t_0$  is the theoretical age (usually in years) at length zero. Parameters of the VBGF are estimated from observed length-at-age pairs such as  $(x, y)$ , where  $y$  is the length at age  $x$ .

Equation 1 can also be modeled in terms of a multiplicative structure for random errors

$$y_i = L(x_i)\varepsilon_i, \quad (2)$$

where  $y_i$  = the length at age  $x_i$  of the  $i$ th sampled subject,  $i = 1, \dots, n$ ;  $L_{\infty} > 0$ ,  $K > 0$ ,  $t < \min\{x_1, \dots, x_n\}$ ; and  $\varepsilon_i$  are independent (not necessarily identical) non-negative random errors (Contreras-Reyes et al., 2014).

With this assumption, the VBGF in Equation 2 corresponds to the multiplicative nonlinear regression with logarithmic random errors. The additive structure of the original model in Equation 2 is easily recovered by applying logarithmic properties, such as  $y'_i = L'(x_i) + \varepsilon'_i$ ,

with  $y'_i = \log y_i$ ,  $L'(x_i) = \log L(x_i)$ , and  $\varepsilon'_i = \log \varepsilon_i$  which are independent random errors. The estimated parameters can be used to compute a  $M$ -at-length estimator.

Gislason et al. (2010) proposed an indirect method for estimating  $M$ -at-length, using previous estimations of  $M$ , habitat temperature and the VBGF parameters. Their proposed  $M$ -at-length estimator is only a function of the VBGF parameters and habitat temperature. Charnov et al. (2013) re-analyzed the data set assembled by Gislason et al. (2010) and concluded that coefficients of growth parameters of the previous  $M$ -at-length estimator did not differ from the general life history theory described in Charnov (1993). The analysis in Charnov et al. (2013) also provides a theoretical basis for estimating indirectly  $M$ -at-length from life history theory. In this reductionist approach, the  $M$ -at-length in fish populations is based entirely on the VBGF parameters and can be estimated by the expression

$$M(L) = K \left( \frac{L_\infty}{L} \right)^{3/2}, \quad (3)$$

where  $M$  = the natural mortality rate at length  $L$ ; and the other parameters are estimated from the VBGF.

#### Natural mortality model

Method error was incorporated in this model by following a log-normal regression model of the form

$$\log M(L) = \beta_0 + \beta_1 \log \frac{L_\infty}{L} + \beta_2 \log K + \log \eta, \quad (4)$$

where  $\log \eta$  = an independent additive Gaussian error,  $\log \eta \sim N(0, \sigma_\eta^2)$ .

Parameter  $\sigma_\eta^2$  corresponds to the uncertainty given by the mortality expression (Eq. 3), which came from the Gislason et al. (2010) database and comprises 168 more or less independent estimates for marine and brackish water fish. In order to have an approximation of  $\sigma_\eta^2$ , as well as  $\beta_0$ ,  $\beta_1$ , and  $\beta_2$ , we estimate the log-normal regression model with normal prior distributions with mean 0 and large variance for  $\beta_i$ ,  $N(0, \sigma_\eta^2)$ ,  $i = 0, 1, 2$ , and an inverse gamma distribution,  $\text{Inv-}\Gamma(\alpha, \beta)$ , for with small values for the scale and rate hyperparameters,  $(\alpha, \beta)$  on a Bayesian framework.

In the Results section, the estimated parameters are presented with this approach for the same data set used by Charnov et al. (2013).

#### Life history parameters

**Bayesian approach** Assuming the multiplicative structure of Equation 3, Contreras-Reyes et al. (2014) considered a skew- $t$  distribution for the errors  $\varepsilon'_i$ , denoted by  $\varepsilon'_i = \log \varepsilon_i \sim ST(\mu_i, \sigma_i^2, \lambda, \nu)$ ,  $i = 1, \dots, n$ . More specifically, we assumed that the original multiplicative errors  $\varepsilon_i$ ,  $i = 1, \dots, n$ , are independently distributed random variables following a log-skew- $t$  distribution (Azalini et al., 2002; Marchenko and Genton, 2010) with parameters of location, dispersion, shape and degrees of freedom given by  $\mu_i$ ,  $\sigma_i^2$ ,  $\lambda$ , and  $\nu > 2$ , respective-

ly. Then, the log-transformed lengths are  $y'_i = \log y_i \sim ST(\mu_i + L_i, \sigma_i^2, \lambda, \nu)$ ,  $i = 1, \dots, n$ , with  $L'_i = \log L(x_i)$ . Thus, the density of  $y'_i = \log y_i$  is

$$f(y'_i | x_i, \mu_i, \sigma_i^2, \lambda, \nu) = \frac{2}{\sigma_i} t(z_i; \nu) T \left( \lambda z_i \sqrt{\frac{\nu+1}{\nu+z_i^2}}; \nu+1 \right), y'_i \in R, \quad (5)$$

where  $z_i = (y'_i - \mu_i - L'_i) / \sigma_i$  is a standardized version of  $y'_i$ , and  $t(z; \nu)$  and  $T(z; \nu)$  represent the usual symmetric Student's  $t$  density and cumulative distribution function, respectively. In this case, we assume that the original lengths follow a log-skew- $t$  distribution, denoted by  $y_i \sim LST(\mu_i + L_i, \sigma_i^2, \lambda, \nu)$ ,  $i = 1, \dots, n$ . The power function  $m(\rho; x_i) = x_i^{2\rho}$  introduces heteroscedasticity into the dispersion parameter  $\sigma_i^2 = \sigma^2 m(\rho; x_i)$  (Contreras-Reyes and Arellano-Valle, 2013).

A Bayesian analysis for the log-skew- $t$  VBGF is proposed. If independence is assumed, the likelihood function  $f(y' | x, \theta)$  of the unknown parameter vector  $\theta = (L_\infty, K, t_0, \sigma^2, \rho, \lambda, \nu)^\top = (\beta^\top, \sigma^2, \rho, \lambda, \nu)^\top$  is considered with Equation 5. The Bayesian model specification involves a prior distribution for each parameter of  $\theta$  to be inserted into Equation 3 in the case of  $L_\infty$  and  $K$ , respectively and inference on  $\theta$  rests on posterior distribution  $\pi(\theta | x, y') \propto f(y' | x, \theta) \pi(\theta)$ .

In the specific case of southern blue whiting (López Quintero et al., 2017), the prior distributions were the following:  $L_\infty \sim TN_{(0, \infty)}(0, 100)$ ;  $K \sim \Gamma(15, 100)$ ;  $-t_0 \sim \Gamma(14, 4)$ ;  $\pi(\rho) \propto 1$  (a non-informative prior density);  $\sigma^2 \sim \Gamma(0.1, 0.1)$ ;  $\lambda \sim N(0, 100)$ ; and  $\nu \sim TE_{[2, \infty)}(0.5)$ , where  $TN_{(0, \infty)}(0, \sigma^2)$  denotes the truncated normal density at interval  $(0, \infty)$  with zero mean and variance  $\sigma^2$ ,  $TE_{[c, \infty)}(\lambda)$  denotes the truncated exponential density at interval  $[c, \infty)$  with parameter  $\lambda$ , and  $\Gamma(\alpha, \beta)$  represents the gamma distribution whose shape and rate parameter are  $\alpha$  and  $\beta$ , respectively.

**Copula approach** The copula method is used when considering 2 random variables represented by their joint (cumulative) distribution function,  $H(u, v) = H(F(x), G(y))$ , where  $u$  and  $v$  denote the values of  $F(x)$ , and  $G(y)$ , respectively. Our main objective is to simulate random variables from  $H(u, v)$ .

Most of the sampling methods are based on the assumption that  $F(x)$  and  $G(y)$  are known but the joint distribution  $H(F(x), G(y))$  is not. A naive framework would be that in which samples are drawn independently from each marginal distribution, yet this approach is would be misleading because underlying dependent structures in  $H$  will be lost. Alternatively, a copula  $C$  between  $F$  and  $G$  can be used as a function such that  $C(F(x), G(x)) = H(u, v)$ . This function can be recognized as a distribution function on  $[0, 1]^d$  (for a dimension  $d$ ) whose arguments are the marginal distribution of the joint distribution. Thus, as in any statistical distribution,  $C$  depends on some parameters, say  $\theta$ .

A copula can be simulated by using different approaches depending on the family they belong to, e.g., Gaussian, Archimedean, Extreme-value, and others.

Each of these families have their own set of parameters (dos Santos Silva and Lopes, 2008). Owing to an important result from copula theory, we are able to simulate from  $H$ , even if unknown, only requiring the marginal distributions,  $F$  and  $G$ , and the copula  $C_\theta$  (see Sklar’s Theorem in Nelsen, 2006).

In order to imitate the joint distribution  $H$ , the copula  $C_\theta$  and the marginal distributions  $F$  and  $G$  need to be estimated. Doing so could create various situations, namely: if  $F$  and  $G$  belong to known families, we could use a parametric framework to obtain  $\hat{C}_\theta$ , by either estimating all parameters associated with both marginal and copula inside a single routine in a procedure known as *full maximum-likelihood*, or by estimating the parameters in 2 stages. The first stage includes estimating parameters of marginal distributions and afterwards the parameters in copula. This approach is commonly known as *inference functions for margins* (Cherubini et al., 2004). There are also intermediate procedures where the marginal distributions are estimated nonparametrically. Results are then plugged into the copula parameters obtained by the maximum likelihood estimation. This latter approach is called “canonical maximum likelihood” or “maximum pseudo-likelihood” and is based on the empirical estimation of the distribution functions  $F$  and  $G$ . This strategy is applied when the families of marginal distributions are unknown (McNeil et al., 2005). In all these cases, to estimate the copula  $C$  we need approximations for  $F$  and  $G$ , denoted by  $\hat{F}$  and  $\hat{G}$ , respectively. These quantities are known as “pseudo-samples” or “pseudo-observations” and provide valuable information for the type of dependency between the original variables (see Fig. 2 and McNeil et al., 2005).

We obtained the empirical posterior distributions for  $K$  and  $L_\infty$ , where the exact marginal and joint family distributions are unknown. We consider the Gaussian copula with parameter  $\rho$ ,  $|\rho| < 1$ , which is the usual correlation coefficient that measures the linear association between 2 variables.

The Gaussian copula has a nonclosed form,

$$C_\rho(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} e^{\frac{w^2 - 2\rho wz + z^2}{2(1-\rho^2)}} dw dz, \quad (6)$$

and to estimate it, the likelihood function,

$$\ell(\rho) = \sum_{i=1}^n \log [c_\rho \{F(\hat{x}_i), G(\hat{y}_i)\}], \quad (7)$$

needs to be maximized, where  $c_\rho(a, b) = \frac{\partial^2}{\partial u \partial v} C_\rho(u, v)$  and  $\Phi(\cdot)$  represents the cumulative normal distribution. We are following the maximum pseudolikelihood approach where the marginal distributions  $F(x)$  and  $G(y)$  are estimated nonparametrically as follows:

$$F(\hat{x}_i) = \frac{1}{n+1} \sum_{i=1}^n I\{K \leq x_i\} \text{ and} \quad (8)$$

$$G(\hat{y}_i) = \frac{1}{n+1} \sum_{i=1}^n I\{L_\infty \leq y_i\}. \quad (9)$$

Using the copula estimate parameters, we draw samples from  $C_\theta = C_\rho$ , as suggested in Nelsen (2006).

Furthermore, to obtain samples on the same scale of  $K$  and  $L_\infty$ , instead of the (0,1) interval, we apply the appropriate quantile functions to the simulated marginal copula. These quantile functions depend on the original marginal distributions.

### M-at-length estimation

To obtain an estimator of Equation 3, we take  $N$  estimated values of parameters  $L_\infty$  and  $K$ , represented as  $L_\infty^{(j)}$  and  $K^{(j)}$ , respectively,  $j = 1, \dots, N$ . All these values are used for each of the values of fish lengths,  $y_i$ ,  $i = 1, \dots, J$ , where the value of  $J$  is one of the distinct values of the age. The estimated values  $K^{(j)}$  and  $L_\infty^{(j)}$  are obtained from either the Bayesian chains values after convergence or from the copula simulation. To complete the estimator, it is then perturbed with the uncertainty  $\eta_i$  (method error) of Equation 4, recovered from the  $M$  modeling,

$$M_{y_i}^{(j)} = K^{(j)} \left( \frac{L_\infty^{(j)}}{y_i} \right)^{3/2} \eta_i, \quad (10)$$

with  $j = 1, \dots, N$ ,  $i = 1, \dots, J$ .

In order to emphasize that the mortality estimator depends on individual  $i$  through the length variable,  $y$ , the subscript  $y_i$  is added to the notation. Using the Equation 10, we draw samples at each length  $i$ . Uncertainty in the VBGF parameters (trait error) is guaranteed by using Markov chains or copula iterations, the incorporation of  $\eta_i$ , whereas  $M$  estimators have their own uncertainty abridged  $\sigma_\eta$  parameter.

Furthermore, to include the dependence structure between age and length in the mortality estimation, we use the predicted value  $\hat{L}(x_i)$  instead of  $y_i$ . Using Equation 1, we obtain the mortality estimator with the following equation:

$$M_{y_i}^{(j)} = K^{(j)} \left( \frac{L_\infty^{(j)}}{\hat{L}(x_i|\beta)} \right)^{3/2} \eta_i, \quad (11)$$

where  $j = 1, \dots, N$ ;

$i = 1, \dots, J$ ; and

$\tilde{\beta} = (\tilde{L}_\infty, \tilde{K}, \tilde{t}_0)^\top$  are point estimates, such as median values from posterior distribution for  $(L_\infty, K, t_0)$  which can be taken from previous studies.

Results from both approaches, with Bayesian chains or copula simulation, are then compared.

All statistical methods used in this study were developed with the software R<sup>2</sup> (vers. 3.1.0 or higher; R Core Team, 2014). Bayesian estimations were carried out with the program JAGS, vers. 3.4 or higher (Plummer, 2003). Copula estimation was conducted with the R package copula, vers. 0.999-12 or higher, developed by Hofert et al. (2015). The length and the

<sup>2</sup> Mention of trade names or commercial companies is for identification purposes only and does not imply endorsement by the National Marine Fisheries Service, NOAA.

**Table 1**

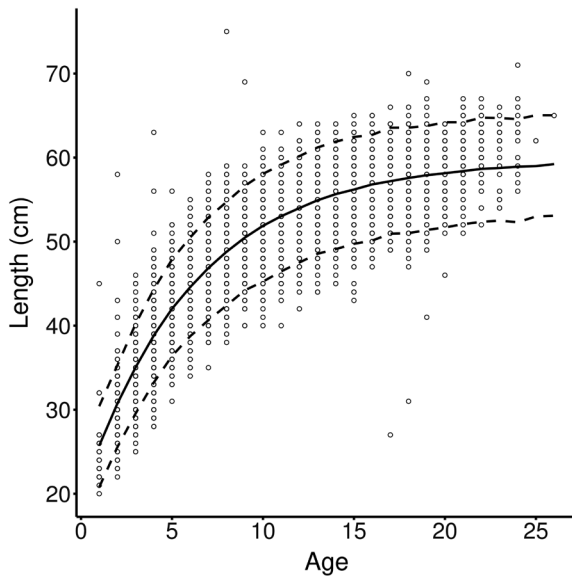
Mean estimates, with standard deviations (SDs) and 95% highest posterior density (HPD) intervals, of Bayesian log-normal model parameters from fitting the regression of natural mortality ( $M$ ) to the von Bertalanffy growth function parameters, the asymptotic length, the growth rate coefficient, and the theoretical age at length zero by using the Equation 4 and data in Charnov et al. (2013).

Parameter	Estimates	SD	95% HPD interval
$\beta_0$	-0.050	0.110	-0.267, 0.163
$\beta_1$	-1.467	0.125	-1.714, -1.226
$\beta_2$	1.007	0.075	0.860, 1.153
$\sigma_\eta$	0.749	0.041	0.666, 0.825

**Table 2**

Mean estimates, with standard deviations (SDs) and 95% highest posterior density (HPD) intervals, of the von Bertalanffy growth function parameters from the power log-skew- $t$  model. The parameters are asymptotic length ( $L_\infty$ ), growth rate coefficient ( $K$ ), negative theoretical age at length zero ( $-t_0$ ), heteroscedasticity ( $\rho$ ), dispersion ( $\sigma^2$ ), skewness ( $\lambda$ ), and degrees of freedom ( $\nu$ ).

Parameter	Estimates	SD	95% HPD interval
$L_\infty$	59.573	0.090	59.386, 59.755
$K$	0.162	0.001	0.159, 0.165
$-t_0$	2.454	0.042	2.367, 2.541
$\rho$	-0.180	0.009	-0.197, -0.161
$\sigma^2$	0.011	0.001	0.010, 0.013
$\lambda$	-1.096	0.053	-1.200, -0.997
$\nu$	14.322	1.047	12.457, 16.586



**Figure 1**

Length-at-age composition for southern blue whiting (*Micromesistius australis*) collected from Chilean continental waters during 1997–2010 (open circles), fitted with the von Bertalanffy growth function: the solid line corresponds to the fit of the log-skew- $t$  model with a heteroscedastic power variance function (Table 2). Dashed lines correspond to the 95% highest posterior density for the fitted log-skew- $t$  model. Fish length was measured as total length in centimeters.

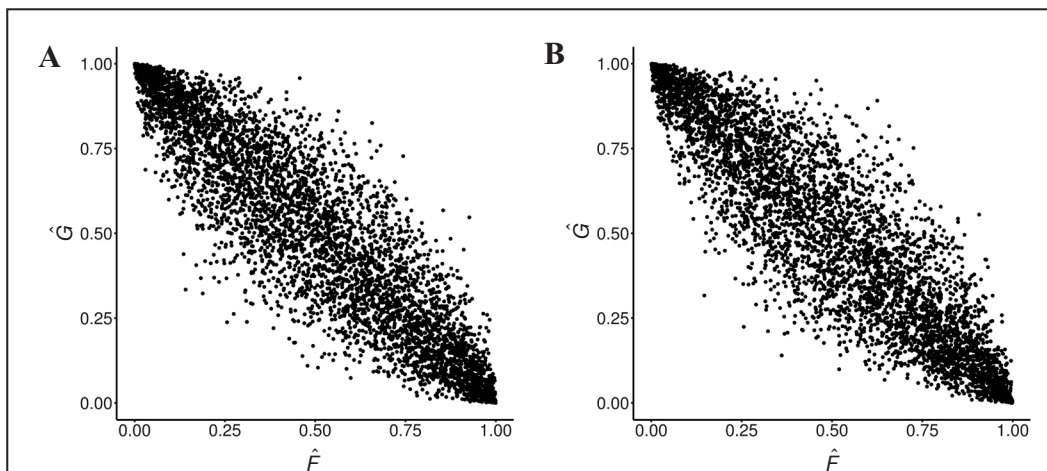
**Results**

Table 1 shows the estimates from the  $M$  model with Bayesian inference. Differences between these results and those reported in Charnov et al. (2013) were caused by the method used to incorporate uncertainty. However, the actual values of the estimated parameters were very similar. The standard deviation,  $\sigma_\eta$ , is then a key parameter underpinning the method error, which is assumed to be a log-normal random variable  $\eta$  in Equations 10 and 11.

Figure 1 shows the application of the power log-skew- $t$  model fitted to the observed length-at-age data (for southern blue whiting) by using the VBGF parameters summarized in Table 2 and estimated with Bayesian inference. This curve was created by simulating 10,000 log-skew- $t$  random values from each age and then by taking the 95% highest posterior density interval. Further details regarding the Bayesian estimation, such as residual diagnostic and sensitivity analysis, can be found in López Quintero et al. (2017). Those authors also showed that the heteroscedastic parameter allows a better model with small variance across observed ages in southern blue whiting. In addition, extreme values for younger and older fish (i.e., <9 and >16 years) were accounted for by the estimated degree of freedom parameter. Moreover, the credibility intervals showed that observations were affected by the negative heteroscedastic parameter,  $\hat{\rho}$ .

The marginal distributions estimated empirically with Equations 8 and 9 and the pseudosamples  $\hat{F}$  and  $\hat{G}$  from copula are shown in Figure 2. We first observed that points concentrate around the diagonal of unit square. The relationship between pseudosamples is linear and negative. Particularly, the Bayesian Markov chain simulation, shown in Figure 2A, which was obtained directly from the estimated VBGF parameters.

burn-in period of the considered chains is 20,000 iterations and 10,000, respectively. Last, a machine running on Linux (kernel 4.6), with an Intel Core i5 processor (Intel Corp., Santa Clara, CA) and 7.7 GB of random-access memory, processed all these computations.



**Figure 2**

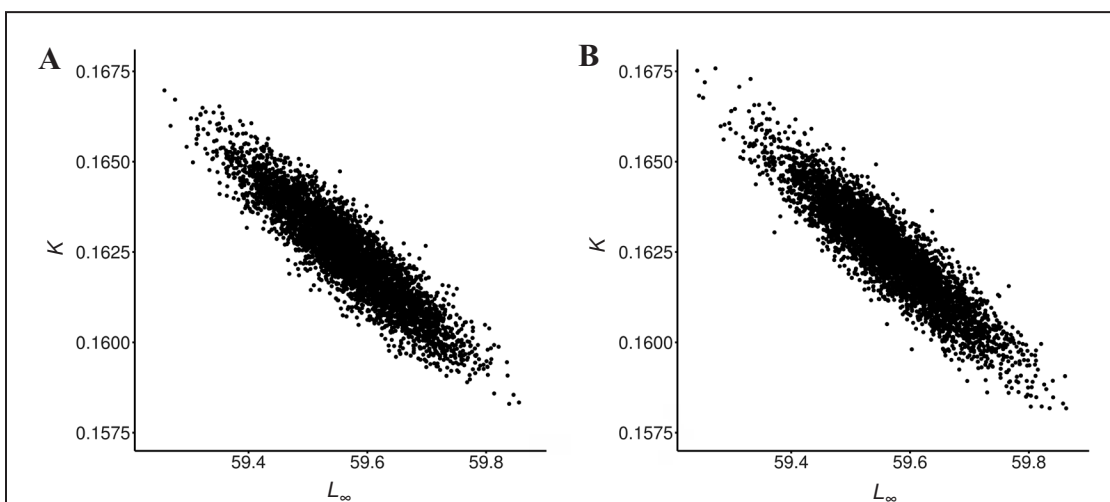
Scatter plots of pseudo-observations, or approximations of distribution functions ( $\hat{F}$  and  $\hat{G}$ ), from simulations (A) where Markov chains of estimated von Bertalanffy growth function (VBGF) parameters were used directly and (B) where copulas were applied to the posterior distribution of VBGF parameters.

In Figure 2B, the relationship between both parameters was recovered by using a Gaussian copula drawing from the empirical posterior distribution. Both plots corresponded with a graphical representation of the intrinsic dependence of the pseudosamples.

Figures 3A and 3B show the estimated parameters from chains and their marginal histograms are displayed in Figures 4A and 4B. The histograms in Figure 4C and 4B were built by assigning a roughly heuristic normal distribution to each marginal, but other alternatives are still possible. These plots ensure the

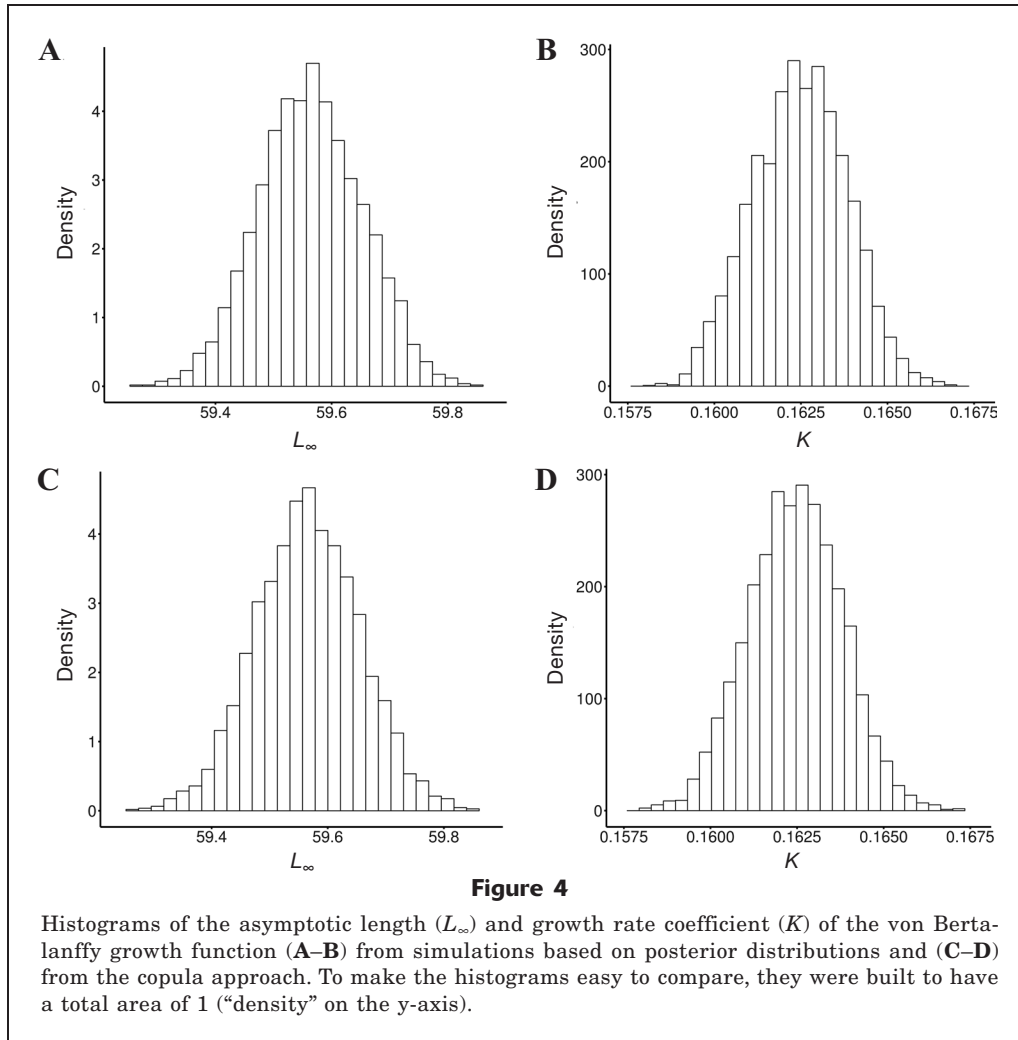
dependence and shape between  $K$  and  $L_\infty$  parameters. We can also include the related quantile functions in the sample generation if the exact distributions for  $K$  and  $L_\infty$  are known. In our study, we particularly considered the Gaussian copula in Figure 3B because of the dependence between pseudosamples generated for both marginal posterior distributions of  $K$  and  $L_\infty$  which looked linear and strongly negative (Fig. 3A; coefficient of correlation  $[r] = -0.912$  (standard error 0.006;  $t$ -value =  $-157.14$ ;  $P$ -value  $< 0.001$ ).

Additionally, both methods showed a strong correla-



**Figure 3**

Scatter plots of asymptotic length ( $L_\infty$ ) and growth rate coefficient ( $K$ ) of the von Bertalanffy growth function showing results (A) from simulations based on posterior distribution and (B) from the copula approach.



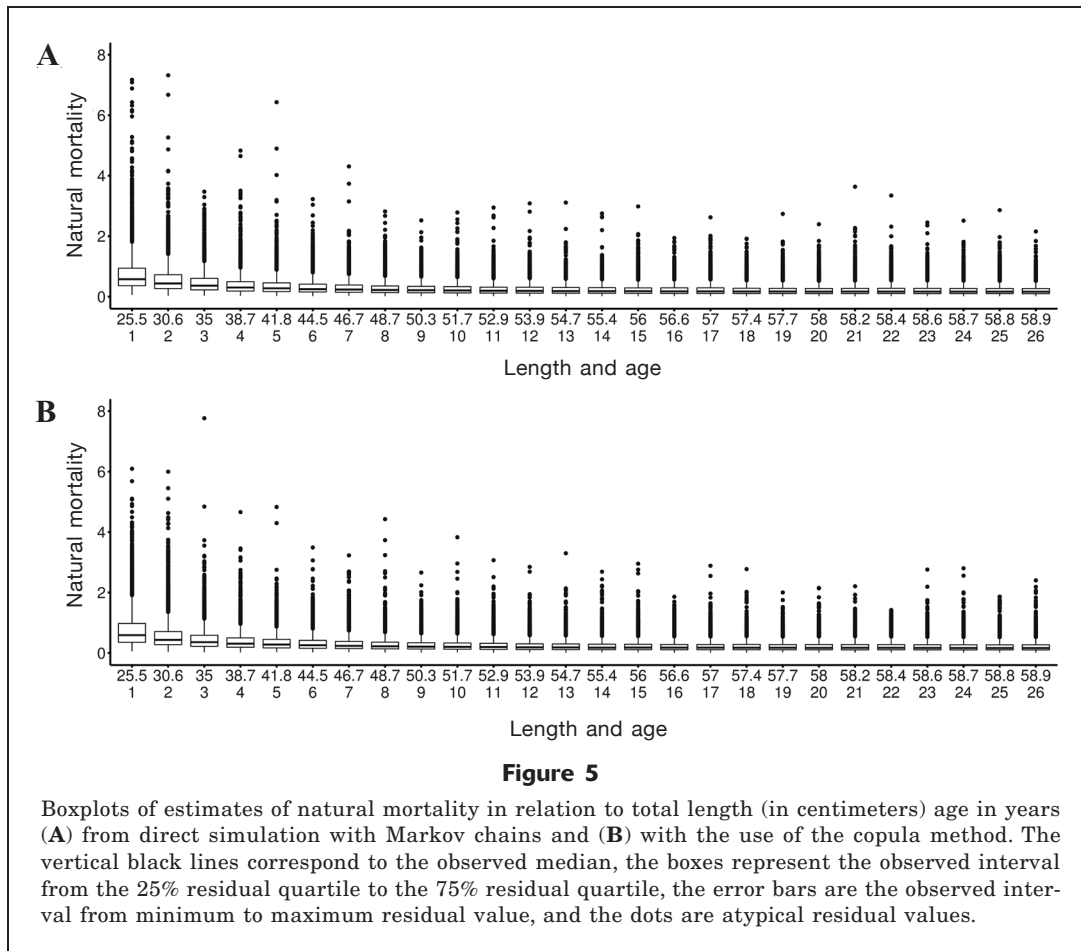
tion between  $K$  and  $L_\infty$ . Note the copula method keeps the correlation structure reported in Siegfried and Sansó (2006) and López Quintero et al. (2017). These simulated values of  $K^{(j)}$  and  $L_\infty^{(j)}$ , from both, Markov chain or from copula, will be used in the expression  $M$ -at-length estimation.

Figure 5 shows the simulated  $M$  derived from Equations 10 and 11, by using both the Markov chains and copula approaches. Uncertainty at each age was incorporated by simulating the posterior distributions of each VBGF parameter. The uncertainty for the method error was incorporated through the random variable  $\eta$  (as explained in the *Materials and methods* section). The empirical distributions related to the copula method showed a similar shape to that with the Markov chains method because we have assumed marginals of the Gaussian family. The empirical distributions behavior can change depending on the class of copula used and the available information for marginals (dos Santos Silva and Lopes, 2008). Moreover, it has been assumed that  $M$  follows a slow exponential decay with length (Gedamke and Hoenig, 2006). On

the other hand, recent empirical and theoretical work has shown that  $M$  represents a decreasing exponential function of length in fish populations (Gislason et al., 2010; Charnov et al., 2013). Specifically, for both methods, the percent change in the median values from the first to the sixth year is around 57%. It is important to note that the most of the methods used to estimate  $M$  consider this parameter as constant across ages and lengths within species (e.g. Pauly, 1980; Hewitt and Hoenig, 2005). These indirect methods are based on the assumption that  $M$  remains relatively constant in fish after they reach sexual maturity.

## Discussion

We presented a mortality estimator that incorporates 2 sources of variability or uncertainty. One of them is associated with the VBGF parameters (trait error) and the other is related to the mortality model (model error) which is related to uncertainty in the original method described in Charnov et al. (2013). In relation



to the trait error, we presented 2 ways to incorporate uncertainty in length-based  $M$  estimators, drawing on an empirical distribution of VBGF parameters. First, the applied Bayesian methods simulate the  $M$  estimator directly with the posterior distribution of the VBGF parameters, with the method of López Quintero et al. (2017). This particular approach was preferred over that for traditional distributions (Siegfried and Sansó, 2006; Hamel, 2015), because this type of data usually contains a degree of asymmetry and extreme values. Additionally, an inadequate distribution may underestimate the real variance contained in the data. The model used gives great flexibility in modeling heteroscedasticity by adding a function dependent on the scale  $\sigma^2$  and a heteroscedastic parameter,  $\rho$  (Contreras-Reyes et al., 2014). In addition, a copula method was used to approximate the posterior distribution and calculate the  $M$  estimator. The method proposed in this study provides a way to incorporate uncertainty in the length-based  $M$  estimator proposed by Charnov et al. (2013), while acknowledging both method and traits errors. Furthermore, our scheme can easily be extended to generate values of uncertainty in indirect methods used to estimate mortality, and therefore has the potential to improve actual ad-hoc methods for incorporating

uncertainty, such as in Cubillos et al. (1999); Quiroz et al. (2010) and Wiff et al. (2011). Furthermore, we recommend the copula method instead of the Bayesian Markov chain approach to incorporate uncertainty in the  $M$ -at-length estimates for 2 reasons: 1) the copula method conserves the underlying dependence in the posterior distribution (see Figs. 3 and 4) and 2) it uses less computing time than the Bayesian Markov chain approach. For example, in our case the copula method required 1 s to compute each length class, whereas the same procedure takes at least 20 h with the Bayesian Markov chain approach. These differences in computational time result in part from the Metropolis–Hastings algorithm that is derived from the kernel of likelihood function and discards, by construction, many proposed values of parameters. Other algorithms, such as the Gibbs sampler, can take the advantage of the known conditional distributions, making the sampling process faster.

Additionally, it is important to note that the proposed  $M$ -at-length estimators are not conceptually limited to use a Bayesian estimation. Researchers can simulate the entire  $M$ -at-length structure just by knowing the dependence between parameters  $L_\infty$  and  $K$ . This can be achieved by reviewing specialized litera-



ture about the species under study and incorporating this information in the copula method. In addition, the method proposed is not limited to the use of Charnov et al. (2013) estimator as the underpinning model to relate  $M$  and growth parameters. A method for estimating  $M$  addressing uncertainty and considering environmental factors such temperature (e.g., Hewitt and Hoenig, 2005) can also be used.

As pointed out in the introduction,  $M$  is a key parameter in modeling any animal population but, it is crucial for harvested fish populations. Natural mortality affects these populations concurrently and continuously with fishing mortality to yield the total mortality rate, which determines the decay in the abundance of a population over time and therefore the size of the stock. Estimation of  $M$  within stock assessment models is difficult, and resultant estimates are usually imprecise (Vetter, 1988; Gavaris and Ianelli, 2002). Standard practice is therefore to use a constant value for  $M$  across sizes or ages—a value that is derived from indirect methods when fitting a population model. Most of the current stock assessment models are age or size based, and therefore the incorporation of an age or size constant value for  $M$  is misleading and may introduce a critical source of bias in abundance estimates. Recent stock assessment models recognize the importance of incorporating size-dependent mortality, and thus incorporating uncertainty in size-based models has become highly recommended (e.g., Clark, 1999; Fu and Quinn, 2000; Siegfried and Sansó<sup>1</sup>; Gislason et al., 2010; Lee and Chang<sup>3</sup>). We agree with Quiroz et al. (2010) in the sense that uncertainty, as reported here for size-based  $M$  estimates, can then be integrated into stock assessment models and used for management analysis through methods such as: 1) sensitivity analysis, where an assessment is conducted repeatedly for several values of size-based  $M$  (see McAllister et al., 1994; Patterson, 1999) drawn from the empirical distributions using copula methods; 2) Bayesian framework, by setting the empirical distributions as prior distributions of size-based  $M$ ; and 3) state-space models, where uncertainty in  $M$  is incorporated as one of the random components regulating the stochasticity in the population dynamics (i.e. the model process error; see Millar and Meyer, 2000).

## Acknowledgments

The authors are grateful to the Instituto de Fomento Pesquero for providing access to the data used in this work. J. Contreras-Reyes was supported partially beginning in 2016 by Comisión Nacional de Investigación

Científica y Tecnológica (CONICYT) doctoral scholarship number 21160618. R. Wiff was funded beginning in 2014 by the Centro de Ecología Aplicada y Sustentabilidad (CAPES) project CONICYT FB 0002. We are sincerely grateful to the 3 anonymous reviewers for their comments and suggestions that greatly improved an early version of this manuscript.

## Literature cited

- Azzalini, A., T. Del Cappello, and S. Kotz.  
2002. Log-skew-normal and log-skew-t distributions as models for family income data. *J. Income Distrib.* 11(3-4):12–20.
- Brodziak, J., J. Ianelli, K. Lorenzen, and R. D. Methot Jr. (eds).  
2011. Estimating natural mortality in stock assessment applications. NOAA Tech. Memo. NMFS-F/SPO-119, 38 p.
- Charnov, E. L.  
1993. Life history invariants: some explorations of symmetry in evolutionary ecology, 184 p. Oxford Univ. Press, Oxford, UK.
- Charnov, E. L., H. Gislason, and J. G. Pope.  
2013. Evolutionary assembly rules for fish life histories. *Fish. Fish.* 14:213–224. [Article](#)
- Cherubini, U., E. Luciano, and W. Vecchiato.  
2004. Copula methods in finance, 310 p. John Wiley & Sons, West Sussex, UK.
- Clark, W. G.  
1999. Effects of an erroneous natural mortality rate on a simple age-structured stock assessment. *Can. J. Fish. Aquat. Sci.* 56:1721–1731. [Article](#)
- Contreras-Reyes, J. E., and R. B. Arellano-Valle.  
2013. Growth estimates of cardinalfish (*Epigonus crassicaudus*) based on scale mixtures of skew-normal distributions. *Fish. Res.* 147:137–144. [Article](#)
- Contreras-Reyes, J. E., R. B. Arellano-Valle, and T. M. Canales.  
2014. Comparing growth curves with asymmetric heavy-tailed errors: application to the southern blue whiting (*Micromesistius australis*). *Fish. Res.* 159:88–94. [Article](#)
- Cubillos, L. A., R. Alarcón, and A. Brante.  
1999. Empirical estimates of natural mortality for the Chilean hake (*Merluccius gayi*): evaluation of precision. *Fish. Res.* 42:147–153. [Article](#)
- dos Santos Silva, R., and H. F. Lopes.  
2008. Copula, marginal distributions and model selection: a Bayesian note. *Stat. Computing* 18:313–320. [Article](#)
- Essington, T. E., J. F. Kitchell, and C. J. Walters.  
2001. The von Bertalanffy growth function, bioenergetics, and the consumption rates of fish. *Can. J. Fish. Aquat. Sci.* 58:2129–2138. [Article](#)
- Fu, C., and T. J. Quinn II.  
2000. Estimability of natural mortality and other population parameters in a length-based model: *Pandalus borealis* in Kachemak Bay, Alaska. *Can. J. Fish. Aquat. Sci.* 57:2420–2432. [Article](#)
- Gavaris, S., and J. N. Ianelli.  
2002. Statistical issues in fisheries' stock assessments. *Scand. J. Stat.* 29:245–267. [Article](#)
- Gedamke, T., and J. M. Hoenig.  
2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosefish. *Trans. Am. Fish. Soc.* 135:476–487. [Article](#)

<sup>3</sup> Lee, H.-H., and Y. J. Chang. 2013. Age-structured natural mortality for Pacific blue marlin based on meta-analysis and an ad hoc model. Working document ISC/13/BILLWG-1/07. Submitted to the International Scientific Committee for Tuna and Tuna-like Species in the North Pacific Ocean. [Available from [website](#).]

- Gislason, H., N. Daan, J. C. Rice, and J. G. Pope.  
2010. Size, growth, temperature and the natural mortality of marine fish. *Fish Fish.* 11:149–158. [Article](#)
- Hamel, O. S.  
2015. A method for calculating a meta-analytical prior for the natural mortality rate using multiple life history correlates. *ICES J. Mar. Sci.* 72:62–69. [Article](#)
- Hewitt, D. A., and J. M. Hoenig.  
2005. Comparison of two approaches for estimating natural mortality based on longevity. *Fish. Bull.* 103:433–437.
- Hofert, M., I. Kojadinovic, M. Maechler, and J. Yan.  
2015. Copula: multivariate dependence with copulas. R package vers. 0.999-13. [Available from [website](#).]
- López Quintero, F. O., J. E. Contreras-Reyes, R. Wiff, and R. B. Arellano-Valle.  
2017. Flexible Bayesian analysis of the von Bertalanffy growth function with the use of a log-skew- $t$  distribution. *Fish. Bull.* 115:13–26. [Article](#)
- Marchenko, Y. V., and M. G. Genton.  
2010. Multivariate log-skew-elliptical distributions with applications to precipitation data. *Environmetrics* 21:318–340. [Article](#)
- McAllister, M. K., E. K. Pikitch, A. E. Punt, and R. Hilborn.  
1994. A Bayesian approach to stock assessment and harvest decisions using the sampling/importance resampling algorithm. *Can. J. Fish. Aquat. Sci.* 51:2673–2687. [Article](#)
- McNeil, A. J., R. Frey, and P. Embrechts.  
2005. Quantitative risk management: concepts, techniques and tools, 538 p. Princeton Univ. Press, Princeton, NJ.
- Millar, R. B., and R. Meyer.  
2000. Bayesian state-space modeling of age-structured data: fitting a model is just the beginning. *Can. J. Fish. Aquat. Sci.* 57:43–50. [Article](#)
- Montenegro, C., and M. Branco.  
2016. Bayesian state-space approach to biomass dynamic models with skewed and heavy-tailed error distributions. *Fish. Res.* 181:48–62. [Article](#)
- Nelsen, R. B.  
2006. An introduction to copulas, 2nd ed., 272 p. Springer-Verlag, New York.
- Patterson, K. R.  
1999. Evaluating uncertainty in harvest control law catches using Bayesian Markov chain Monte Carlo virtual population analysis with adaptive rejection sampling and including structural uncertainty. *Can. J. Fish. Aquat. Sci.* 56:208–221. [Article](#)
- Pauly, D.  
1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. *ICES J. Mar. Sci.* 39:175–192. [Article](#)
- Plummer, M.  
2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. *In* Proceedings of the 3rd international workshop on distributed statistical computing (DSC 2003); Vienna, Austria, 20–22 March (K. Hornik, F. Leisch, and A. Zeileis, eds.), 10 p. [Available from [website](#).]
- Quiroz, J. C., R. Wiff, and B. Caneco.  
2010. Incorporating uncertainty into estimation of natural mortality for two species of Rajidae fished in Chile. *Fish. Res.* 102:297–304. [Article](#)
- R Core Team.  
2014. *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. [Available from [website](#), accessed April 2014.]
- Sainsbury, K. J.  
1980. Effect of individual variability on the von Bertalanffy growth equation. *Can. J. Fish. Aquat. Sci.* 37:241–247. [Article](#)
- Siegfried, K. I., and B. Sansó.  
2006. Two Bayesian methods for estimating parameters of the von Bertalanffy growth equation. *Environ. Biol. Fish.* 77:301–308. [Article](#)
- Vetter, E. F.  
1988. Estimation of natural mortality in fish stocks: a review. *Fish. Bull.* 86:25–43.
- Wiff, R., J. C. Quiroz, V. Ojeda, and M. A. Barrientos.  
2011. Estimation of natural mortality and uncertainty in pink cusk-eel (*Genypterus blacodes* Schneider, 1801) in southern Chile. *Lat. Am. J. Aquat. Res.* 39:316–326.