



Abstract—The von Bertalanffy growth function is the model most widely applied to describe growth in fish populations. Parameters describing this function usually are estimated from observed lengths at different ages by using maximum likelihood and by assuming Gaussian distributed errors. In harvested populations, observed length at age usually involves a high level of skewness and extreme values because of the size-selective sampling process. Some approaches, based on the maximum-likelihood method for making inferences, have been developed to resolve such issues. We propose a Bayesian framework for estimating growth parameters for nonlinear regression models—a framework that is based on the family of log-skew- t distributions and which provides an approach that is flexible enough for modeling the presence of asymmetries and heavy tails. This framework based on a method in which 1) the error accounts for both skewness and heavy-tailed distributions of a log-skew- t model, and 2) the observed length at each age has a heteroscedastic error distribution. The proposed method was applied and compared with the methods of previous models by using observed length-at-age data for the southern blue whiting (*Micromesistius australis*), an important fish species harvested in the southeast Pacific. Comparisons indicated that the proposed model is the best for describing data on southern blue whiting.

Manuscript submitted 6 January 2016.
Manuscript accepted 30 September 2016.
Fish. Bull. 115(1):13–26 (2017).
Online publication date: 27 October 2016.
doi: 10.7755/FB.115.1.2

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.

Flexible Bayesian analysis of the von Bertalanffy growth function with the use of a log-skew- t distribution

Freddy Omar López Quintero¹

Javier E. Contreras-Reyes (contact author)^{1,2}

Rodrigo Wiff³

Reinaldo B. Arellano-Valle⁴

Email address for contact author: javier.contreras@ifop.cl

¹ Departamento de Matemática
Universidad Técnica Federico Santa María
Avenida España 1680
Valparaíso 2390123, Chile

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

³ Center of Applied Ecology and Sustainability
Pontificia Universidad Católica de Chile
Avenida Libertador Bernardo O'Higgins 328
Santiago 8331150, Chile

⁴ Departamento de Estadística
Pontificia Universidad Católica de Chile
Avenida Vicuña Mackenna 4860
Santiago 7820436, Chile

Growth is one of the most important measurable life-history traits in individual organisms because it is fundamental in creating an understanding of both population and ecosystem functions. Several models have been proposed to describe animal growth. The most widely applied model, however, is the von Bertalanffy growth function (VBGF; von Bertalanffy, 1938). This model has been used to describe length at age for a wide range of species across several taxa, such as mammals (English et al., 2012), birds (Tjørve and Tjørve, 2010), and reptiles (Lehman and Woodward, 2008), although it is most extensively applied to fish species (Pardo et al., 2013). The VBGF is based on principles underpinning the physiology of growth (von Bertalanffy, 1938; Wiff and Roa-Ureta, 2008), gives an adequate description of growth with the use of only 3 parameters, and states that the rate of growth of an individual is determined by the difference between the buildup of body mass and loss due to energy expenditures for maintenance.

In harvested fish populations, the usual data available to estimate these parameters are cross-sectional, and a single length and age measurement is taken from each sampled individual. The VBGF describes the expected growth rate for the population, on the basis of length-at-age data composed of individuals with variable growth rates.

Maximum-likelihood techniques derived from Gaussian and log-Gaussian errors normally are used to estimate VBGF parameters (Millar, 2002; Siegfried and Sansó, 2006). Yet, in fish populations, this assumption often fails because these distributions are typically skewed, present heavy tails or have extreme values. Skewed distributions usually result from the size-selective sampling process (Montenegro and Branco, 2016). In addition, in harvested fish populations, an accumulative effect of fishing exploitation exists for size at age. Growth rates vary among individuals (Sainsbury, 1980), and fishing selectivity removes faster growing individuals from each particular

age class. Hence, the bias in sampling for length at age that favors fast-growing individuals of each age class (Taylor et al., 2005). Therefore, the assumption of Gaussianity to estimate parameters of the VBGF is not adequate (Contreras-Reyes and Arellano-Valle, 2013; Montenegro and Branco, 2016). Moreover, the assumption of Gaussianity implies that length may take negative values and, therefore, is nonsensical (Xiao, 1994; Millar, 2002).

Different approaches have been proposed to overcome this drawback and fitting the VBGF. They can be separated roughly into 2 categories. In the first one, models, such as the one in Taylor et al. (2005), provide a mechanistic approach to dealing with skewed length-at-age data, with a combined process of growth, selectivity, and mortality when fitting the VBGF. The second category is a more empirical approach in which skewed and heavy-tailed length-at-age data are modeled by using the maximum-likelihood method and assuming a non-Gaussian distribution (Contreras-Reyes and Arellano-Valle, 2013) and by using Bayesian analysis (Millar, 2002; Siegfried and Sansó, 2006). Millar (2002) proposed a Bayesian framework to estimate parameters of the VBGF, using a multiplicative error model with log-normal distribution. Contreras-Reyes and Arellano-Valle (2013) calculated the maximum-likelihood estimates for the VBGF with the family of skew-distributions (Azzalini and Capitanio, 2003), a flexible class that extends the known Student distribution (e.g., Geweke, 1993). Such models can incorporate asymmetric and heavy-tailed errors, with presence of heteroscedasticity (Montenegro and Branco, 2016). Contreras-Reyes et al. (2014) reanalyzed the skew-approach to incorporate a log-skew- t distribution under multiplicative error distribution.

In this study, we examined our proposed Bayesian method for estimating the VBGF parameters on the basis of a log-skew- t distribution. This new framework merges the benefits provided by Bayesian analysis (Siegfried and Sansó, 2006) and the log-skew- t distribution (Contreras-Reyes et al., 2014) for estimating parameters of the VBGF for harvested fish populations. Additionally, our approach allows for heteroscedasticity in errors, modeled as power and exponential functions (Contreras-Reyes and Arellano-Valle, 2013). This Bayesian framework is applied to data of length-at-age composition of southern blue whiting (*Micromesistius australis*), an important species fished in the southeast Pacific.

Materials and methods

Log-skew- t von Bertalanffy growth model

We let $L(x_i)$ be the expected value of the length related to an i th individual at age x_i , $L_\infty > 0$, $K > 0$, $t_0 < \min\{x_1, \dots, x_n\}$, and n is the sample size. The VBGF defines growth in length as

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

Equation 1 represents the simplest formulation of the VBGF, described by 3 parameters:

- where L_∞ = the asymptotic length (in length units, e.g., centimeters);
- K = the growth rate coefficient expressed per unit of time; and
- t_0 = the theoretical age (usually in years) when the length is zero.

Parameters of the VBGF usually are estimated from observed length-at-age pairs, such as (x_i, y_i) , $i = 1, \dots, n$, where y_i is the i th observed length at age x_i . Equation 1 was described in terms of multiplicative structure (Millar, 2002; Siegfried and Sansó, 2006; Contreras-Reyes et al., 2014) for random errors:

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

where ε_i = non-negative random errors, usually assumed to be independent, identically distributed errors with a mean of 1. Given this assumption, the VBGF in Equation 2 corresponds to the nonlinear regression with multiplicative random errors. We easily recovered the additive structure of the original model in Equation 2 by applying log scale in the following way:

$$y'_i = L'(x_i) + \varepsilon'_i, \text{ with } \log y'_i, L'(x_i) = \log L_i = L'_i, \text{ and} \\ \varepsilon'_i = \log \varepsilon_i, i = 1, \dots, n, \quad (3)$$

in which ε'_i were assumed to be independent, identically distributed, random errors with zero mean.

Contreras-Reyes et al. (2014) assumed a log-skew- t distribution (Azzalini et al., 2003) for the multiplicative and heteroscedastic random errors. Specifically, they assumed that the multiplicative errors ε_i , $i = 1, \dots, n$, were independent random variables following a log-skew- t distribution with parameters $\mu_i \in \mathbb{R}$ (location), $\sigma_i^2 > 0$ (scale and dispersion), $\lambda_i \in \mathbb{R}$ (skewness and shape), and $\nu > 0$ (degrees of freedom), a distribution that is denoted by

$$\varepsilon'_i \sim LST(\mu_i, \sigma_i^2, \lambda, \nu), i = 1, \dots, n. \quad (4)$$

This approach is equivalent to considering that the transformed errors ε_i , $i = 1, \dots, n$, are independent and have a skew distribution (Branco and Dey, 2001; Azzalini and Capitanio, 2003) denoted by

$$\varepsilon'_i \sim ST(\mu_i, \sigma_i^2, \lambda, \nu), i = 1, \dots, n. \quad (5)$$

In turn, this notation indicates that the transformed response variables (lengths) are derived from

$$y'_i \sim ST(\mu_i + L'_i, \sigma_i^2, \lambda, \nu), i = 1, \dots, n. \quad (6)$$

namely, that the density of y'_i is given by

$$f(y'_i | x_i, \mu_i, \sigma_i^2, \lambda, \nu) = \frac{2}{\sigma_i} t(z_i; \nu) T\left(\lambda z_i \sqrt{\frac{\nu+1}{\nu+z_i^2}}; \nu+1\right), \quad (7)$$

$$y'_i \in \mathbb{R},$$

where

$$z_i = (y'_i - \mu_i - L'_i) / \sigma_i \text{ is a standardized version of } y'_i,$$

$$t(z; \nu) = \frac{\Gamma\left[\frac{\nu+1}{2}\right]}{\Gamma(\nu/2)(\pi\nu)^{\frac{1}{2}}}\left(1 + \frac{z^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad (8)$$

$z \in \mathbb{R}$ is the symmetric Student- t density with ν degrees of freedom, and $T(z; \nu)$ represents the respective cumulative distribution function. In other words, we assumed that the original response y_i followed a log-skew- t distribution (Marchenko and Genton, 2010), which is denoted by

$$y_i \sim LST(\mu_i + L_i, \sigma_i^2, \lambda, \nu), i = 1, \dots, n. \quad (9)$$

We assumed that $\nu > 1$ and considered the first moment of the skew- t distribution (Branco and Dey, 2001); therefore, the extra parameter μ_i had to be chosen as

$$\mu_i = \sqrt{\frac{\nu}{\pi}} \frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)} \frac{\lambda \sigma_i}{\sqrt{1+\lambda^2}}, \quad (10)$$

so that the transformed errors ε'_i have a zero mean. This condition ensured that $E(y'_i) = L'_i$ and allowed us to identify the constant in the additive version of the regression model.

Heteroscedasticity was introduced by means of the dispersion parameters σ_i^2 and modeled by using a non-negative function $m(\rho; x_i)$ depending on age x_i and a heteroscedastic parameter $\rho \in \mathbb{R}$ as $\sigma_i^2 = \sigma^2 m(\rho; x_i)$, where $\sigma^2 > 0$. When $\rho = 0$, homoscedasticity is recovered as $\sigma_i^2 = \sigma^2 m(0; x_i) = \sigma^2$. In our study, we considered 2 specific functions for modeling heteroscedasticity: the exponential function $m(\rho; x_i) = e^{\rho x_i}$ and the power function $m(\rho; x_i) = x_i^{2\rho}$. In both functions, if $m(0; x_i) = 1$, it corresponds to the homoscedastic case.

Asymmetry and heavy tails produced by extreme values of length-at-age data were controlled by the parameters of shape (λ) and degrees of freedom (ν).

Extension to a Bayesian framework

We advanced a Bayesian analysis for the log-skew- t VBGF described in the previous section. Therefore, we first noted from the independence assumption and Equation 7 that the likelihood function of the unknown parameter vector $\theta = (\beta^\top, \sigma^2, \rho, \lambda, \nu)^\top$ is

$$f(\mathbf{y}' | \mathbf{x}, \theta) = \prod_{i=1}^n \left(\frac{2}{\sigma_i}\right) t(z_i; \nu) T\left(\lambda z_i \sqrt{\frac{\nu+1}{\nu+z_i^2}}; \nu+1\right), \quad (11)$$

where $\beta = (L_\infty, K, T_0)^\top$ are VBGF parameters, $\mathbf{y}' = (y'_1, \dots, y'_n)^\top$, $\mathbf{x} = (x_1, \dots, x_n)^\top$, and z_i is as it was previously defined. To complete our Bayesian model specification, we needed to elicit a prior distribution for the unknown parameter vector, say $\pi(\theta)$. Therefore, the Bayesian inference on θ (or function of θ) was based on the posterior distribution $\pi(\theta | \mathbf{x}, \mathbf{y}') \propto f(\mathbf{y}' | \mathbf{x}, \theta) \pi(\theta)$. This posterior distribution does not have a closed form (Cancho et al., 2011), but an estimation could still be calculated by using a Markov chain Monte Carlo (MCMC) algorithm (Chib and Greenberg, 1995; Cowles and Carlin, 1996; Robert and Casella, 2004).

Given the available methods, we chose to implement a hand-tailored component-wise Metropolis–Hasting

transition scheme; in other words, we selected an approach in which θ is divided into individual pieces that are easily updated sequentially with a random walk algorithm. Our selection was based on simplicity and the need to control all steps in the sampling. Other options included the use of variants of BUGS language (Lunn et al., 2012), the AD ModelBuilder (Fournier et al., 2012), or Stan software (Gelman et al., 2015). Note that our selected approach is different from that followed by Siegfried and Sansó (2006), who employed an algorithm that included both Gibbs, as well as Metropolis–Hasting steps. An advantage of using this MCMC scheme is that the procedure is subdivided into several univariate steps. All proposal distributions were tuned to achieve acceptance rates of 25–45% (Robert and Casella, 2004). Specifically, we considered the different components of θ as independent (Siegfried and Sansó, 2006); in other words, $\pi(\theta)$ becomes the product of the marginal prior distributions of $(\beta^\top, \sigma^2, \rho, \lambda, \nu)$. It follows that only these marginal prior distributions must be elicited to complete our Bayesian model.

Prior distributions for the VBGF parameters β were chosen as follows. Given that L_∞ is strictly positive, we assumed a left truncated normal distribution with large variance (e.g., 100) as the prior distribution for this parameter. Other possible and natural choices were log-normal, gamma, or even distributions with support in a reasonable and restricted interval. Because parameters K and $-t_0$ are both positive, we used gamma as prior distributions for them (Xiao, 1994; Siegfried and Sansó, 2006). Contreras-Reyes et al. (2014) reported estimates of around 0.16 for southern blue whiting, a value that incidentally conforms to the value obtained by Siegfried and Sansó (2006) for blue shark (*Prionace glauca*). We used this information to specify that the mean of the gamma prior distribution of K was around 0.15. For $-t_0$, Contreras-Reyes et al. (2014) obtained $-\hat{t}_0 = 2.5$, which indicates a prior distribution for this parameter.

For the scale parameter σ^2 , we considered the classical inverse gamma prior distribution suitable for this type of parameter (Zhang et al., 2009). The heteroscedastic parameter ρ usually takes positive or negative values. To give the full power of estimation to the data, we chose a noninformative prior, $\pi(\rho) \propto 1$. For the shape parameter λ , we set a normal prior distribution with a zero mean and large variance. The parameter defining the degrees of freedom, ν , should be strictly larger than 2 to ensure the existence of variance in the log-skew- t model; therefore, we considered an exponential prior distribution with mean equaling 2 (Geweke, 1993; Cancho et al., 2011) and truncated at the interval $(2, \infty)$. These prior specifications are summarized in Table 1.

Comparisons and selection of models

For sake of comparison, we considered 2 additional models with constant variance function derived from the log-normal distribution. The first one (hereafter, re-

Table 1

Elicited prior specifications for log-normal and log-skew- t models used to examine a Bayesian analysis of the von Bertalanffy growth function. $TN_{(0,\infty)}(0,100)$ represents the $N(0,100)$ -density truncated at $(0,\infty)$, and $TE_{(2,\infty)}(0.5)$ denotes the exponential density truncated at the $(2,\infty)$ interval. The parameters are the asymptotic length (L_∞), growth rate coefficient (K), theoretical age in years when the length is zero ($-t_0$), heteroscedasticity (ρ), inverted dispersion (σ^{-2}), skewness (λ), and degrees of freedom (ν).

Parameter	Log-normal (type I)	Log-normal (type II)	Log-skew- t
L_∞ (cm)	$\pi(L_\infty) \propto 1$	$TN_{(0,\infty)}(0,100)$	$TN_{(0,\infty)}(0,100)$
K (y^{-1})	$Gamma(15,100)$	$Gamma(15,100)$	$Gamma(15,100)$
$-t_0$ (y)	$Gamma(10,4)$	$Gamma(10,4)$	$Gamma(10,4)$
ρ	–	–	$\pi(\rho) \propto 1$
σ^{-2}	$Gamma(0.1,0.1)$	$Gamma(0.1,0.1)$	$Gamma(0.1,0.1)$
λ	–	–	$N(15,100)$
ν	–	–	$TE_{(2,\infty)}(0.5)$

ferred to as *type-I model*) was similar to that developed by Siegfried and Sansó (2006), and the second one included a modification of the prior distribution of L_∞ so that it was the same as that proposed in the log-skew- t model. All these prior specifications are summarized in Table 1. The following models were considered:

- Log-normal (type I) with constant variance function;
- Log-normal (type II) with constant variance function;
- Log-skew- t with constant variance function;
- Log-skew- t with exponential variance function;
- Log-skew- t with power variance function.

Selecting the “best” model is an important aspect in statistical analysis. In the rest of this section, we describe how we implemented the deviance information criterion (DIC) and the widely applicable information criterion (WAIC) for model selection.

Deviance information criterion The DIC proposed by Spiegelhalter et al. (2002) is based on the posterior mean of the deviance, and it can be approximated by the MCMC algorithm as follows:

$$\widehat{DIC} = 2 \sum_{i=1}^n \left(\log f(y'_i | \mathbf{x}_i, \bar{\theta}) - \frac{4}{B} \sum_{s=1}^B \log f(y'_i | \mathbf{x}_i, \theta_s) \right), \quad (12)$$

where $\bar{\theta} = \frac{1}{B} \sum_{s=1}^B \theta_s$ is the mean of a sample $\theta_1, \dots, \theta_B$ obtained from the posterior distribution $\pi(\theta | \delta)$. The DIC is related to the effective number of parameters:

$$\hat{p}_{DIC} = 2 \left(\log f(\mathbf{y}' | \mathbf{x}, \bar{\theta}) - \frac{1}{B} \sum_{s=1}^B \log f(\mathbf{y}' | \mathbf{x}, \theta_s) \right). \quad (13)$$

The widely applicable information criterion The WAIC (e.g., Gelman et al., 2014) is based on the computed log-pointwise-posterior-predictive density, complemented by a correction for the effective number of parameters to adjust for overfitting:

$$\widehat{WAIC} = \sum_{i=1}^n \log \left(\frac{1}{B} \sum_{s=1}^B \log f(y'_i | \mathbf{x}_i, \theta_s) \right) - \hat{p}_{WAIC}. \quad (14)$$

Also, the WAIC is related to the effective number of parameters:

$$\hat{p}_{WAIC} = 2 \sum_{i=1}^n \left(\log \left(\frac{1}{B} \sum_{s=1}^B \log f(y'_i | \mathbf{x}_i, \theta_s) \right) - \frac{1}{B} \sum_{s=1}^B \log f(y'_i | \mathbf{x}_i, \theta_s) \right). \quad (15)$$

Compared with DIC, WAIC has the property of averaging over the posterior density by using each iterated θ_s , instead of being replaced by the mean $\bar{\theta}$. In addition, \hat{p}_{WAIC} is more numerically stable than \hat{p}_{DIC} because it averages separately for each observation y'_i (Gelman et al., 2014).

Influential analysis

The statistical stability of the proposed models exposed to perturbations of the data were analyzed by using influential analysis. We considered the Kullback–Leibler (KL) divergence measure (Kullback and Leibler, 1951) to quantify the effect on the inferences produced by excluding one observation or a group of observations from the full data set. The KL-divergence had been considered previously in Bayesian influential analysis for elliptical and skew-elliptical models (Arellano-Valle et al., 2000; Vidal et al., 2006).

We let $P = \pi(\theta | S)$ and $P_{-i} = \pi(\theta | S_{-i})$ be the posterior distribution of θ obtained from the full data $S = (\mathbf{x}, \mathbf{y}')$ and the data without the i th observation $S_{-i} = (\mathbf{x}_{-i}, \mathbf{y}'_{-i})$, respectively. The KL-divergence between P and P_{-i} was given by

$$K(P, P_{-i}) = \int \pi(\theta | S) \log \left\{ \frac{\pi(\theta | S)}{\pi(\theta | S_{-i})} \right\} d\theta. \quad (16)$$

To identify influential observations, Peng and Dey (1995) showed that if $p_i \gg 1/2$, where

$$p_i = \frac{1}{2} (1 + \sqrt{1 - e^{-2K(P, P_{-i})}}), \quad \frac{1}{2} \leq p_i \leq 1, \quad (17)$$

then the i th observation is considered influential. Additionally, because the integral in Equation 16 cannot be written in closed form, it still can be approximated by sampling from the posterior distribution of θ via the MCMC algorithm. In fact, if $\theta_1, \dots, \theta_B$ is a sample of size B from $\pi(\theta|S)$, then the MCMC estimator of $K(P, P_{-i})$ is computed as

$$\hat{K}(P, P_{-i}) = \log \left\{ \frac{1}{B} \sum_{s=1}^B \frac{1}{f(y'_i | \mathbf{x}_s, \theta_s)} \right\} + \frac{1}{B} \sum_{s=1}^B \log f(y'_i | \mathbf{x}_s, \theta_s), \quad (18)$$

with $f(y'_i | \mathbf{x}_s, \theta_s)$ given by Equation 7 and $\theta_s = (\beta_s, \sigma_s^2, \rho_s, \lambda_s, v_s)^\top$ (e.g., Cancho et al., 2011). It should be noted that we computed the KL-divergence between P and P_{-i} using the i th marginal sample density $f(y'_i | \mathbf{x}_s, \theta_s)$, but we did so without considering the posteriors $\pi(\theta|S)$ and $\pi(\theta|S_{-i})$.

In addition, given the new sample with removed observations, we quantified the change produced for each new estimate with respect to the full sample. In several cases, the estimates were notably different for these samples (Contreras-Reyes and Arellano-Valle, 2013). We used restricted data S_{-j} in which a set of J observations was removed. Then, we computed the percentage of relative change (RC) of estimates by following Contreras-Reyes and Arellano-Valle (2013). The RC was defined by

$$RC(\hat{\theta}_k, \hat{\theta}_{k,j}) = 100 \left| 1 - \frac{\hat{\theta}_{k,j}}{\hat{\theta}_k} \right|, \quad (19)$$

where $\hat{\theta}_{k,j}$ and $\hat{\theta}_k$ are the posterior median estimates of k th component of θ obtained from the posterior distributions $\pi(\hat{\theta}|S_{-j})$ and $\pi(\hat{\theta}|S)$, respectively. Therefore, we computed the change (in percentage) of each parameter of the VBGF.

Application

Data We evaluated the performance of the proposed model, using the available data for southern blue whiting. This dataset was based on 24,942 individuals collected from a region spanning latitudes from 46°S to 6°S over the period 1997–2010 by the Instituto de Fomento Pesquero (Contreras-Reyes, unpubl. data). Random samples of fish were collected by onboard scientific observers during each catch haul of southern blue whiting were caught. All these fish were measured to the nearest centimeter, and both otoliths of each fish were extracted onboard. Otoliths were then taken to the laboratory, where age was determined by reading annual growth increments in the sagittal otoliths. The southern blue whiting is assumed to recruit once a year; therefore, age is treated as a discrete variable with a 1-year interval. Otolith age assignment involved killing sampled fish; therefore, each data point represents 1 individual fish. Fish in the catch had observed ages between 1 and 25 years and a size range of 20–75 cm in total length (Cés-

pedes et al., 2013). Contreras-Reyes et al. (2014) reported extreme values in young and old age classes and reported asymmetry caused by fishing selectivity. Both these issues justify the use of heavy-tailed and skewed distributions in VBGF errors.

MCMC sampling For inference, 4 chains were selected from each applied Bayesian model. The length of the chains necessary to reach convergence differed depending on the treated model: it was around 20,000 iterations for all log-skew- t models and around 100,000 iterations for the log-normal homoscedastic model. We considered a burn-in period to be 10,000 iterations for the first model and 20,000 for the log-normal homoscedastic model. In addition, all models conformed with the traditional diagnostic convergence tests, such as Geweke and Heidelberger-Welch, when tests were applied to individual chains (Cowles and Carlin, 1996; Carlin and Louis, 2000). However, results of the Raftery–Lewis test, also applied to individual chains, indicated that we should take the largest thinning of chains (Link and Eaton, 2011) because values were highly correlated. In addition, visual examination throughout trace and autocorrelation plots (not shown) indicated that convergence was reached for all parameters in all models. This situation was transferred to the effective sample size, which can be interpreted as the number of independent samples necessary to yield the same precision as the (serially dependent) MCMC samples. Effective sample size is especially important in resampling and should not be confused with the degree of over dispersion usually found in length-at-age compositions.

For all models, the parameters with higher and lower values of effective sample size were v and K , respectively. Finally, \hat{R} Gelman's indexes (Gelman and Rubin, 1992) were all near 1, indicating that the specific parameter had good convergence after the burn-in period was eliminated. This test was applied to 4 chains for each parameter and each model.

Software Statistical methods used in this article were implemented in the software R, vers. 3.1.0 or higher (R Core Team, 2014). The MCMC was developed in C++ embedded in the R package ReppArmadillo, vers. 0.4.300.0 or higher (Eddelbuettel and Sanderson, 2014). Diagnostic analysis was conducted with the coda package, vers. 0.16-1 or higher, in R (Plummer et al., 2006). Von Bertalanffy growth curves were estimated in each realization by simulating estimated parameters several times (e.g., 10,000) with the models proposed previously in the *Comparisons and selection of models* section. Such simulated observations are called *fake data*, according to Gelman and Hill (2007, Ch. 16). Afterward, the 95% highest posterior density (HPD) was computed across ages. The generation of the log-skew- t values was conducted by using fake data in the R package sn, vers. 0.4-11 (Azzalini, 2008).

Table 2

Values from models fitted to simulated length-at-age data after log-skew- t distribution with constant variance, with asymptotic length (L_∞) = 59.72, growth rate coefficient (K) = 0.16, theoretical age in years when the length is zero ($-t_0$) = 2.5, heteroscedasticity (ρ) = -0.5, and dispersion (σ^2) = 0.05 considered by ranging the parameters of skewness (λ) and degrees of freedom (ν) in the set $\{-3, -1, 0\} \times \{3, 12, 100\}$.

Model	Parameter	(λ, ν)								
		(-3,3)	(-3,12)	(-3,100)	(-1,3)	(-1,12)	(-1,100)	(0,3)	(0,12)	(0,100)
Log-normal (type I)										
Constant	L_∞	56.789	57.568	57.458	57.519	57.790	57.923	59.717	59.665	59.653
	K	0.158	0.157	0.160	0.155	0.161	0.163	0.161	0.164	0.164
	$-t_0$	2.603	2.541	2.530	2.652	2.491	2.427	2.430	2.381	2.354
	σ^2	0.004	0.002	0.001	0.006	0.002	0.002	0.011	0.003	0.003
Log-normal (type II)										
Constant	L_∞	56.797	57.578	57.456	57.595	57.810	57.938	59.739	59.635	59.640
	K	0.157	0.157	0.160	0.154	0.161	0.163	0.160	0.164	0.164
	$-t_0$	2.606	2.545	2.529	2.684	2.502	2.435	2.445	2.377	2.351
	σ^2	0.004	0.002	0.001	0.006	0.002	0.002	0.011	0.003	0.003
Log-skew- t										
Constant	L_∞	56.570	57.179	57.346	57.395	57.965	57.859	59.809	59.614	59.654
	K	0.160	0.161	0.161	0.160	0.159	0.163	0.153	0.163	0.163
	$-t_0$	2.521	2.453	2.501	2.515	2.518	2.460	2.644	2.403	2.373
	σ^2	0.003	0.003	0.003	0.003	0.004	0.003	0.003	0.004	0.004
	λ	-3.842	-3.627	-3.266	-1.375	-1.683	-1.718	-0.472	-0.885	-1.224
	ν	3.626	16.640	18.471	3.640	11.739	17.457	3.130	13.739	16.752
Log-skew- t										
Exponential	L_∞	56.886	57.651	57.520	57.645	57.801	58.059	59.993	59.624	59.700
	K	0.158	0.156	0.159	0.155	0.161	0.160	0.153	0.163	0.162
	$-t_0$	2.490	2.510	2.491	2.590	2.473	2.490	2.640	2.403	2.372
	ρ	-0.125	-0.121	-0.126	-0.107	-0.117	-0.144	-0.135	-0.069	-0.049
	σ^2	0.005	0.005	0.004	0.006	0.006	0.006	0.005	0.005	0.005
	λ	-3.517	-3.528	-2.917	-1.196	-1.713	-1.458	-0.249	-0.545	-1.126
	ν	3.217	11.846	15.961	3.296	10.977	15.346	2.804	12.430	15.929
	ν	3.217	11.846	15.961	3.296	10.977	15.346	2.804	12.430	15.929
Log-skew- t										
Power	L_∞	56.929	57.645	57.495	57.691	57.829	58.049	59.951	59.610	59.705
	K	0.155	0.155	0.159	0.153	0.160	0.160	0.153	0.163	0.162
	$-t_0$	2.570	2.553	2.512	2.635	2.503	2.487	2.647	2.403	2.382
	ρ	-0.023	-0.021	-0.020	-0.021	-0.023	-0.025	-0.022	-0.007	-0.003
	σ^2	0.004	0.004	0.003	0.005	0.005	0.004	0.004	0.004	0.004
	λ	-3.702	-3.670	-3.075	-1.291	-1.778	-1.585	-0.342	-0.863	-1.198
	ν	3.341	13.131	17.094	3.476	11.855	16.640	2.940	13.372	16.574
	ν	3.341	13.131	17.094	3.476	11.855	16.640	2.940	13.372	16.574

Results

Simulations

To assess the effect of error distribution in the VBGF parameters, length-at-age data were simulated from a log-skew- t distribution with a constant variance and the estimated growth parameters of Contreras-Reyes et al. (2014) (see the *Comparisons and selection of models* section). Different cases were evaluated by considering a range for λ and ν in the set of $\{-3, -1, 0\} \times \{3, 12, 100\}$ (Table 2). This procedure permits assessment of the closeness of estimates in absence and presence of skewness and heavy-tailed simulated data. Each simulation considered 30,000 realizations. For $\lambda = -3$, estimates of L_∞ indicated the largest differences with real values,

and the smaller differences were reported in K and $-t_0$. The largest differences of ν for estimated and simulated data were produced for $\nu = 100$ (which approximates the log-skew-normal distribution) and for $\lambda = 0$ (which approximates the log-normal distribution). The log-skew- t model presented estimates similar to the initial parameters $\lambda = -1$ and $\nu = 12$.

Modeling data from southern blue whiting

In the case of the parameters of error distribution, the σ^2 posterior estimates are small values because of the application of the log-transformation, and the ν posterior estimates are smaller than 15, indicating the presence of extreme values (Table 3). In these models, the shape parameters λ are close to -1 , indicating a non-

Table 3

Estimates from fitted log-normal and log-skew- t models, with standard deviations (SDs) and 95% highest posterior density (HPD) intervals. The parameters are the asymptotic length (L_∞), growth rate coefficient (K), theoretical age in years when the length is zero ($-t_0$), dispersion (σ^2), heteroscedasticity (r), skewness (λ), and degrees of freedom (ν).

Model	Parameter	Estimate	SD	95% HPD
Log-normal (type I)				
Constant	L_∞	59.249	0.091	(59.064, 59.439)
	K	0.167	0.001	(0.165, 0.170)
	$-t_0$	2.323	0.035	(2.250, 2.396)
	σ^2	0.004	3×10^{-5}	(0.0042, 0.0044)
Log-normal (type II)				
Constant	L_∞	59.249	0.090	(59.060, 59.425)
	K	0.167	0.001	(0.165, 0.170)
	$-t_0$	2.323	0.034	(2.248, 2.391)
	σ^2	0.004	0.000	(0.0042, 0.0044)
Log-skew- t				
Constant	L_∞	59.212	0.086	(59.055, 59.386)
	K	0.166	0.001	(0.164, 0.169)
	$-t_0$	2.382	0.034	(2.322, 2.454)
	σ^2	0.005	1×10^{-4}	(0.0047, 0.0053)
	λ	-1.012	0.051	(-1.105, -0.916)
	ν	11.020	0.643	(9.853, 12.210)
Log-skew- t				
Exponential	L_∞	59.666	0.085	(59.527, 59.815)
	K	0.161	0.001	(0.159, 0.163)
	$-t_0$	2.488	0.036	(2.428, 2.549)
	ρ	-0.039	0.002	(-0.043, -0.035)
	σ^2	0.008	2×10^{-4}	(0.007, 0.008)
	λ	-1.080	0.050	(-1.181, -0.977)
	ν	13.351	0.884	(11.757, 15.144)
	ρ	-0.039	0.002	(-0.043, -0.035)
Log-skew- t				
Power	L_∞	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)
	ρ	-0.180	0.009	(-0.197, -0.161)
	σ^2	0.011	0.001	(0.010, 0.013)
	λ	-1.096	0.053	(-1.200, -0.997)
	ν	14.322	1.047	(12.457, 16.586)

symmetric length-at-age distribution. In addition, log-normal models yielded very similar posterior estimates for VBGF and σ^2 . Interestingly, standard errors for all parameters were very precise, in a similar way to that of previous studies of frequentist inference (Contreras-Reyes and Arellano-Valle, 2013; Contreras-Reyes et al., 2014). This level of precision probably is a result of a strong underlying structure of the data.

Using the DIC and WAIC criteria of Equations 12 and 14, respectively, we found the log-skew- t model with power heteroscedastic function to be the best model (Table 4). As Table 4 indicates, the log-normal model is the least useful among the selected models.

The fitted curve of the power-variance log-skew- t model to the observed length-at-age data is presented in Figure 1A. The model is adequate for younger ages (1–8 years), but for older ages (>15 years) the observed length tends to converge to $\hat{L}_\infty = 59.52$ (Table 3). The

log-skew- t model provides more precise 95% HPD intervals for older ages (>13 years; Fig. 1B) and less precise for young ages (0–5 years) in comparison with the log-normal model. Intervals of 95% HPD of log-skew- t model fit indicate that the observations were affected by the negative heteroscedastic parameter $\hat{\rho}$ (Fig. 1C). In addition, constant variance was assumed for the log-normal model and, therefore, the model underestimated the real variance in the age at length containing extreme values. The posterior densities of VBGF and variance parameters corresponding with the homoscedastic log-normal and power-variance log-skew- t models are compared in Figure 2. The asymmetry and dispersion of the posterior densities of VBGF were similar for the different error distributions. However, for the variance parameter, the posterior density was leptokurtic when log-normal error distribution was used.

Considering the boxplots of residuals by age from

Table 4

Summary of chain diagnostics for the fitted models: effective sample size (ESS), credibility \hat{R} of Gelman index, Geweketest (G), Heidelberger–Welch test (HW), and Raftery–Lewis test (RL). In addition, the deviance information criterion (DIC) and widely applicable information criterion (WAIC) values for each model are reported, with their estimated number of parameters, p_{DIC} and p_{WAIC} , respectively. The parameters are the asymptotic length (L_∞), growth rate coefficient (K), theoretical age in years when the length is zero ($-t_0$), dispersion (σ^2), heteroscedasticity (ρ), skewness (λ), and degrees of freedom (ν).

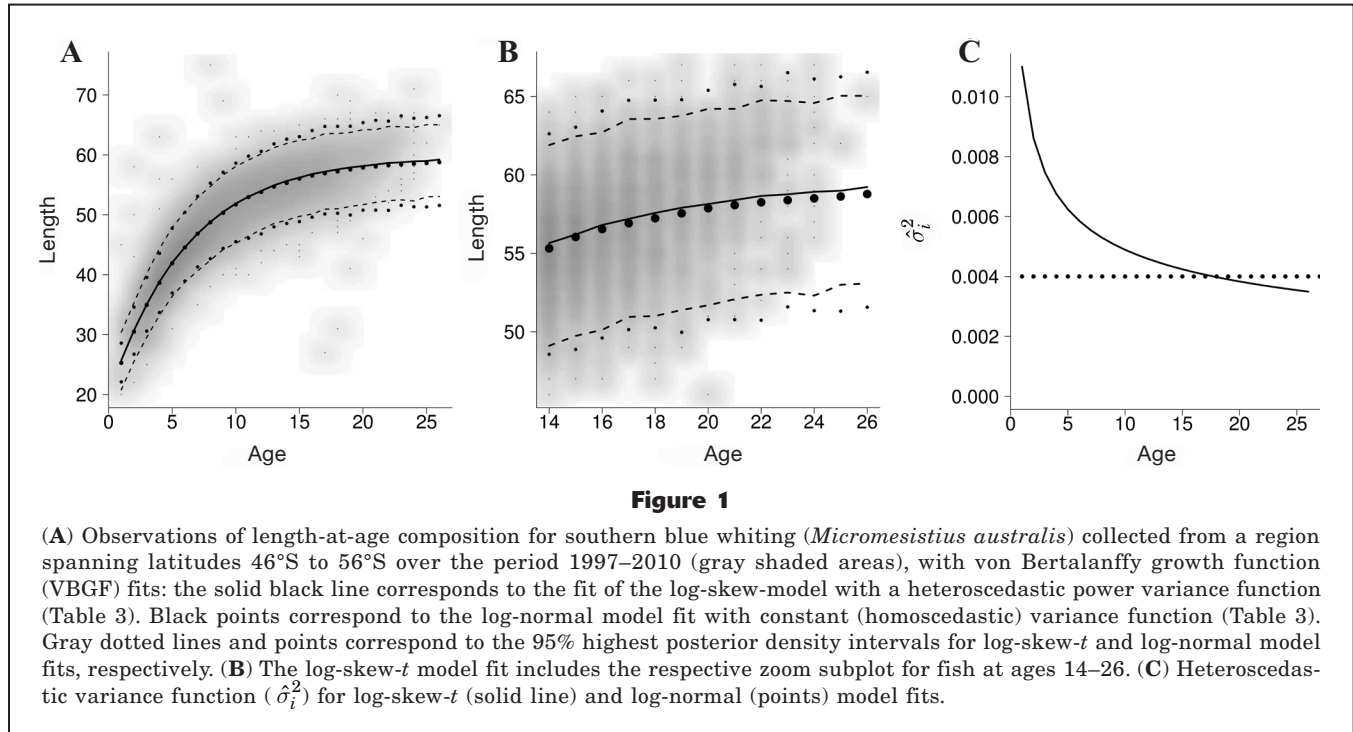
Model	Parameter	ESS	\hat{R}	G	HW	RL	DIC	p_{DIC}	WAIC	p_{WAIC}
Log-normal (type I)										
Constant	L_∞	1134.673	1.002	1.543	0.817	39.4	-65275.66	4.338	-65275.66	7.079
	K	760.156	1.003	-1.241	0.850	82.6				
	$-t_0$	934.834	1.003	1.518	0.799	63.0				
	σ^2	52477.971	1.000	1.111	0.572	4.6				
Log-normal (type II)										
Constant	L_∞	749.129	1.005	-1.891	0.226	32.4	-65275.659	4.338	-66623.850	8.181
	K	463.168	1.008	1.499	0.050	50.3				
	$-t_0$	569.197	1.008	-1.817	0.456	35.3				
	σ^2	29883.554	1.000	1.510	0.093	3.4				
Log-skew- t										
Constant	L_∞	211.801	1.022	1.665	0.143	14.9	-66231.757	2.918	-66227.642	6.133
	K	81.302	1.037	-1.118	0.500	30.0				
	$-t_0$	91.705	1.041	1.230	0.393	60.5				
	σ^2	656.456	1.002	1.662	0.654	6.7				
	λ	596.318	1.004	-1.598	0.638	13.3				
	ν	1927.670	1.005	-1.232	0.408	5.1				
Log-skew- t										
Exponential	L_∞	204.714	1.054	-1.174	0.436	8.2	-66559.569	6.420	-66556.535	9.048
	K	112.927	1.077	1.232	0.471	43.7				
	$-t_0$	162.885	1.078	-1.696	0.264	16.9				
	ρ	561.808	1.013	-1.335	0.504	8.5				
	σ^2	179.131	1.029	1.015	0.673	116.0				
	λ	1404.850	1.010	1.490	0.247	8.7				
	ν	856.653	1.007	-0.771	0.159	10.9				
Log-skew- t										
Power	L_∞	200.944	1.008	-1.293	0.087	7.7	-66625.510	6.482	-66623.850	8.181
	K	99.489	1.052	0.979	0.086	16.0				
	$-t_0$	128.952	1.053	-1.223	0.052	16.9				
	ρ	706.436	1.008	-0.580	0.096	8.7				
	σ^2	189.195	1.029	0.631	0.615	23.6				
	λ	1656.937	1.004	1.700	0.355	5.6				
	ν	289.711	1.016	-0.481	0.495	24.5				

the fitted log-skew- t model (Fig. 3), we can observe that residuals indicate a flat pattern and that their mean is concentrated around zero. We noted also a decreasing variance in older fish, produced in part by the negative value of the estimated heteroscedasticity ($\hat{\rho} = -0.18$). Furthermore, extreme values for younger and older fish (<6 and >15 years) were detected by the estimated degree of freedom ($\hat{\nu} = 14.32$; Table 3).

Influential analysis

Peng–Dey’s criterion (Eq. 17, $p_i = 0.5$) is suitable for certain nonlinear regression models with normal errors and many observations are considered influential

(Fig. 4). As expected, we found that the log-normal model has more influential observations than the power log-skew- t model for each probability. Therefore, we selected, in Table 5, only the probabilities 0.70, 0.60, 0.55, and 0.51 for those influential observations in log-normal and power log-skew- t models. For the selected model, when $p_i = 0.51$, the largest number of restricted observations was recorded and the RC of the error distribution parameters was raised. When the number of influential observations increased (in terms of the p_i) and were removed, the degree of freedom parameter also increased. Because several of these observations are extreme values (Contreras-Reyes et al., 2014), the error distribution shifts from log-skew- t to log-skew- t



normal. Among the VBGF parameters, $-t_0$ showed the largest RC variation, although RC values of the VBGF parameters were small given the absence of influential observations.

The relationship between the paucity of observations for young (1–5 years) and old (16–24 years) age classes and heteroscedastic variance can be interpreted from Table 5 for the cases when $p_i = 0.55$ and 0.51 as follows. For $p_i = 0.55$, 67.5% and 17.5% of young and old individuals, respectively, were obtained for the sample from 40 influential observations. For $p_i = 0.51$, 35.4% and 23.0% of young and old individuals were obtained for the sample, respectively, from 3106 influential observations. When the estimates for $\hat{\sigma}^2$, \hat{v} , and $\hat{\rho}$, from the *Log-skew- t von Bertalanffy growth model* section, were considered, the heteroscedastic variance σ_i^2 decreased mainly when young and old individuals (extreme values) were excluded from the sample.

Correlation analysis

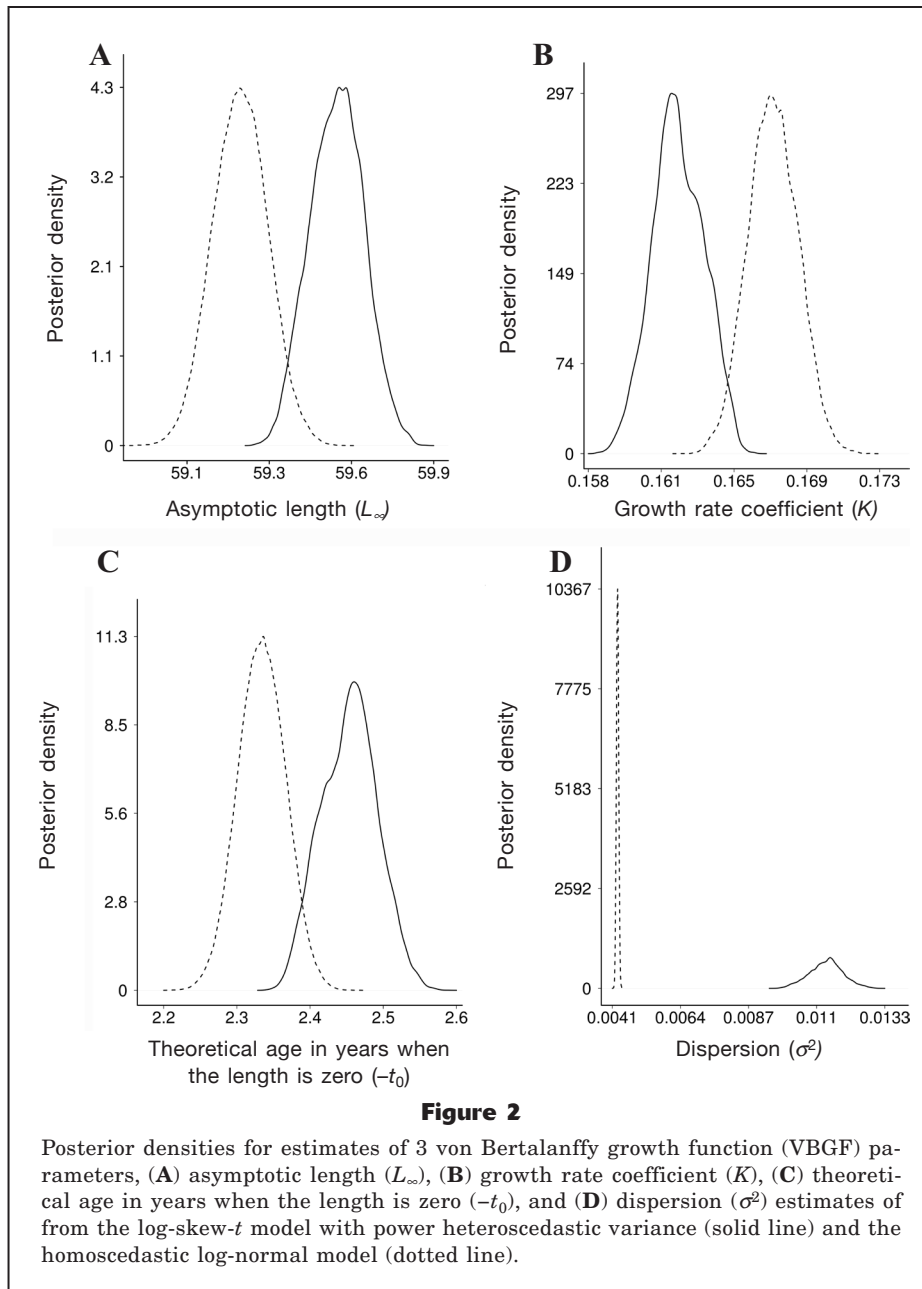
An important aspect in fisheries research related to VBGF analysis is the correlation between parameters (Pilling et al., 2002; Siegfried and Sansó, 2006; Shelton and Mangel, 2012). High correlation among the 3VBGF parameters is common in fish populations (Ratkowsky, 1986; Pardo et al., 2013). Correlation between parameters was analyzed by using the scatter plots in Figure 5. The highest correlation was found between K and $-t_0$ (-0.94), followed by the correlation between L_∞ and K (-0.89) (Xiao, 1994; Pilling et al., 2002; Siegfried and Sansó, 2006) and by the correlation between L_∞ and

$-t_0$ (0.71) (Ratkowsky, 1986; Pilling et al., 2002). Ratkowsky (1986) found that correlations between VBGF parameters may depend on the parameters that are used. Other choices of parameters should produce a low correlation between the VBGF parameters. However, in our model, the solution for K is affected by values of L_∞ and $-t_0$ under the classical VBGF parametrization.

The relationship between estimates of L_∞ and K are similar to the ones found by Siegfried and Sansó (2006), but, in contrast to their results, we found a large correlation between $(K, -t_0)$ and $(L_\infty, -t_0)$. This finding could have occurred for different reasons, such as the species studied and the specific Bayesian method employed. However, the use of maximum-likelihood estimation (not shown) also verified high correlations between those parameters. The scatter plots did not show a clear correlation for error distribution parameters, except for the relationship between σ^2 and ρ given by the heteroscedastic power function, where the correlation was -0.82 .

Discussion

In this study, we embedded previous log-skew- t distribution analyses in a Bayesian framework. This approach, namely using log-skew- t distribution, has several advantages over previous frequentist inference. First, in a Bayesian framework, prior knowledge of the model parameters can be included in the modeling process in terms of a prior distribution, and our inferences were based on the posterior distribution, therefore, allowing



for probability statements for our statistical conclusions. Second, the degrees of freedom parameter ν is directly estimated from the posterior density, whereas, in the frequentist approach used by Contreras-Reyes and Arellano-Valle (2013) and Contreras-Reyes et al. (2014), they are obtained manually by using profiles of the log-likelihood function. Thirdly, an additional parameter λ is also considered by the log-skew- t approach, allowing us to model different degrees of skewness in data—something that the traditional log-normal model does not make possible. Finally, boundary restrictions on each prior density can be incorporated in Bayesian analysis, and avoids deriving nonsensical parameters of the VBGF (Gasbarra et al., 2007).

Contreras-Reyes et al. (2014) computed the VBGF for both sexes in southern blue whiting using the maximum-likelihood method and a heteroscedastic log-skew- t model. Estimates reported in Contreras-Reyes et al. (2014) are similar to those reported here for southern blue whiting, except for the heteroscedastic parameter ρ . In this study, $\hat{\rho}$ was significantly higher than the one reported in Contreras-Reyes et al. (2014) because a prior distribution was specified for the heteroscedastic parameter. This specification allows us to model the decreasing variance of lengths better across ages, given the paucity of observations in young (1–5 years) and older (16–24 years) fish (Fig. 1C). An adequate modeling of variance, especially in young ages,

Table 5

Summary of the log-skew- t model fitted with power heteroscedastic variance function for the full and restricted data and its respective percentage of relative change (RC) for the probabilities (p_i), of 0.70, 0.60, 0.55 and 0.51. The parameters are the asymptotic length (L_∞), growth rate coefficient (K), theoretical age in years when the length is zero ($-t_0$), heteroscedasticity (ρ), dispersion (σ^2), skewness (λ), degrees of freedom (ν), sample size (n), and number of influential observations (Influ. observ.).

Parameters	Restricted data (S_{-1})							
	$p_i = 0.70$	RC (%)	$p_i = 0.60$	RC (%)	$p_i = 0.55$	RC (%)	$p_i = 0.51$	RC (%)
L_∞ (cm)	59.559	0.024	59.550	0.039	59.445	0.215	59.588	0.025
K (y^{-1})	0.163	0.617	0.163	0.617	0.164	1.235	0.161	0.617
$-t_0$ (y)	2.446	0.326	2.455	0.041	2.391	2.567	2.501	1.915
ρ	-0.179	0.556	-0.180	0.000	-0.168	6.667	-0.141	21.667
σ^2	0.012	9.091	0.013	18.182	0.012	9.091	0.008	27.273
λ	-1.181	7.755	-1.231	12.318	-1.305	19.069	-1.801	64.325
ν	19.198	34.046	22.304	55.732	30.223	111.025	80.796	464.139
n	24936		24932		24902		21836	
Influ. observ.		6		10		40		3106

will improve the estimation of t_0 . In addition, considering the influential analysis, the estimated parameters from the restricted data indicate significant differences with those obtained with the full data set, particularly for the degree of freedom parameter of the error distribution. However, numerous subjects were not evaluated in this article—topics such as other sources of uncertainty or data-related problems that can lead to bias in an estimation (Ortiz and Palmer¹). Particularly, we did not address the direct influence of the prior specification on the final estimates (Fig. 5C), a topic that will be of interest for future research.

Siegfried and Sansó (2006) and Hamel (2015) considered log-normal distributions to be appropriate for the asymmetry observed in the length-at-age data in harvested fish populations. However, in such data, we can usually find different degrees of skewness and heavy-tailed and extreme values in which log-normal distribution does not give a good description of observations. The log-normal model may underestimate the real variance contained in the data (Slatkin, 2013). In such cases, log-skew- t models (such as the one proposed here), could yield a fair description of the observed length-at-age data, as was the case for the southern blue whiting, in which the log-skew- t model turned out to be the best among all competing models. In addition, the proposed model gives great flexibility in modelling heteroscedasticity by adding a function dependent on the scale of σ^2 and a heteroscedastic parameter ρ . The assumption of asymmetry and heavy tails and the log-transformed nature of the log-skew- t model reduces the standard errors of the estimated parameters of the VBGF (Contreras-Reyes et al., 2014).

¹ Ortiz, M., and C. Palmer. 2008. Review and estimates of von Bertalanffy growth curves for king mackerel Atlantic and Gulf of Mexico stock units. SEDAR16-DW-12, 20 p. [Available at [website](#).]

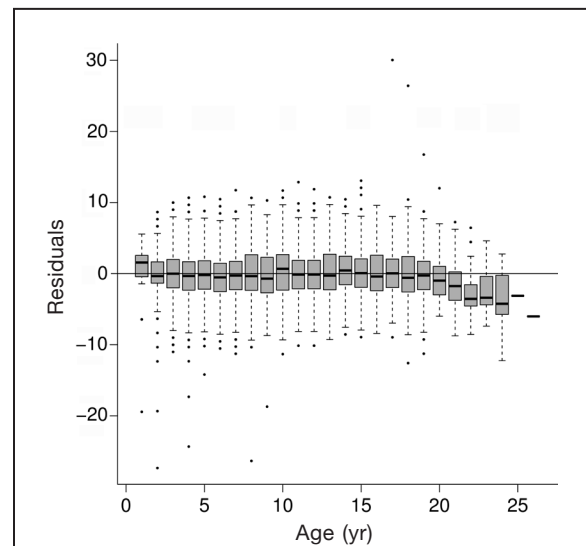
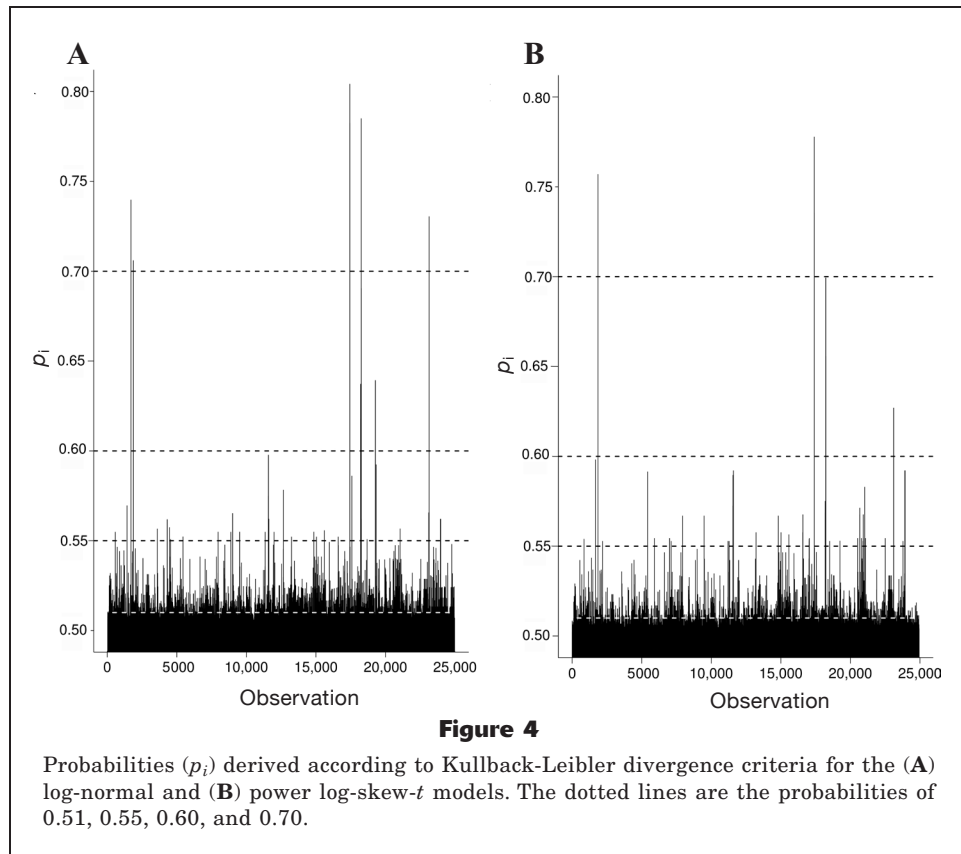


Figure 3

Boxplots of residuals versus ages of southern blue whiting (*Micromesistius australis*) from the log-skew- t model with power heteroscedastic function. The dark black lines correspond to the observed median, the gray shaded boxes represent the observed interval from the 25% residual quartile to the 75% residual quartile, the error bars are the observed interval from minimum to maximum residual value, and the dots are atypical residual values.

The Bayesian analysis that we developed and describe in this article provides a flexible framework that allows biologically meaningful estimates of the VBGF. This method also takes into account the uncertainty and kurtosis produced by extreme values common in



length-at-age data (Quiroz et al., 2010; Montenegro and Branco, 2016). This framework is particularly relevant when dealing with harvested fish populations for which length-at-age data collected from fishing operations usually contain missing observations and indicate bias toward fast-growing individuals of each age class. A correct specification of the VBGF is critical because growth is an important aspect of contemporary stock assessment models (Zhu et al., 2016) upon which biomass estimates and conservation measures are based.

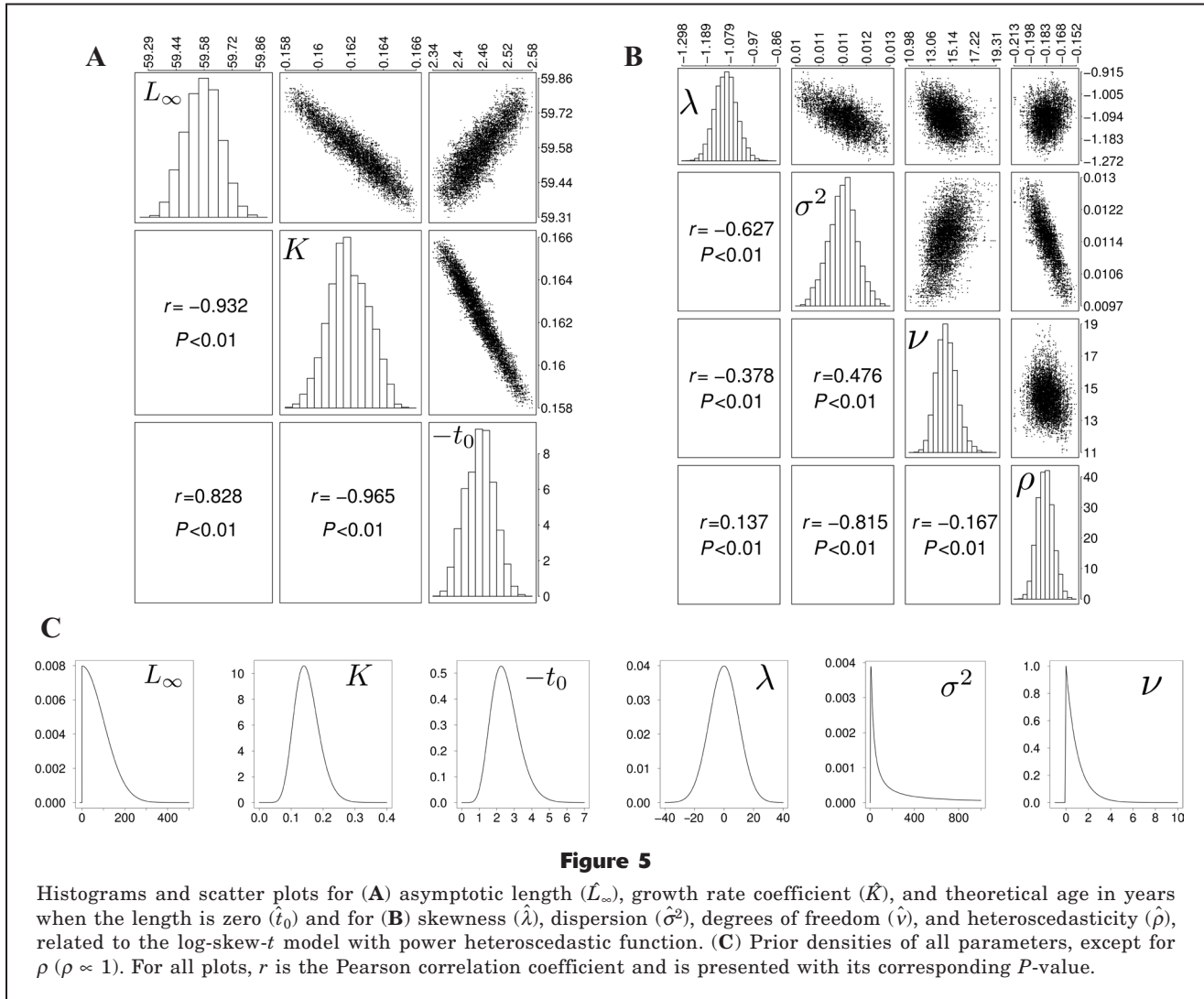
Acknowledgments

The authors are grateful to the Instituto de Fomento Pesquero (Valparaíso, Chile) for providing access to the data used in this work. The research of F. López Quintero was supported partially by a doctoral grant from Pontificia Universidad Católica de Valparaíso (Valparaíso, Chile). The research of J. Contreras-Reyes was supported partially by Comisión Nacional de Investigación Científico y Tecnológico (CONICYT) doctoral scholarship 2016 number 21160618 (Res. Ex. 4128/2016). R. Wiff was funded by Fondo Nacional de Desarrollo Científico y Tecnológico (FONDECYT) post-doctoral project number 3130425 and by the Center of Applied Ecology and Sustainability (CAPES) project CONICYT FB 0002. R. Arellano-Valle was funded by

FONDECYT (Chile) grants 1120121 and 1150325. We are sincerely grateful to 3 anonymous reviewers for their comments and suggestions that greatly improved an early version of this manuscript.

Literature

- Arellano-Valle, R. B., M. Galea-Rojas, and P. Iglesias Zuazola. 2000. Bayesian sensitivity analysis in elliptical linear regression models. *J. Stat. Plann. Inference* 86:175–199. [Article](#)
- Azzalini, A. 2008. The R package sn: the skew-normal and skew-t distributions (version 0.4-6). [Available at [website](#).]
- Azzalini, A., and A. Capitanio. 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *J. R. Stat. Soc., B* 65:367–389. [Article](#)
- Azzalini, A., T. del Cappello, and S. Kotz. 2003. Log-skew-normal and log-skew- t distributions as models for family income data. *J. Income Distrib.* 11:12–20.
- Branco, M., and D. K. Dey. 2001. A general class of multivariate skew-elliptical distributions. *J. Multivariate Anal.* 79:93–113. [Article](#)
- Cancho, V. G., D. K. Dey, V. H. Lachos, and M. G. Andrade. 2011. Bayesian nonlinear regression models with scale mixtures of skew-normal distributions: estimation and



- case influence diagnostics. *Comput. Stat. Data Anal.* 55:588–602. [Article](#)
- Carlin, B. P., and T. A. Louis.
2000. Bayes and empirical Bayes methods for data analysis, 2nd ed., 440 p. Chapman & Hall/CRC, Boca Raton, FL.
- Céspedes, R., V. Ojeda, L. Adasme, L. Muñoz, K. Hunt, L. Cid, M. Miranda, and A. Villalón.
2013. Programa de seguimiento de las pesquerías demersales y aguas profundas, 2013: demersales sur austral, 175 p. Instituto de Fomento Pesquero, Valparaíso, Chile.
- Chib, S., and E. Greenberg.
1995. Understanding the Metropolis-Hastings algorithm. *Am. Stat.* 49:327–335. [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. Mariella Canales.
2014. Comparing growth curves with asymmetric heavy-tailed errors: application to southern blue whiting (*Micromesistius australis*). *Fish. Res.* 159:88–94. [Article](#)
- Cowles, M. K., and B. P. Carlin.
1996. Markov chain Monte Carlo convergence diagnostics: a comparative review. *J. Am. Stat. Assoc.* 91:883–904.
- Edelbuettel, D., and C. Sanderson.
2014. RcppArmadillo: accelerating R with high-performance C++ linear algebra. *Comput. Stat. Data Anal.* 71:1054–1063. [Article](#)
- English, S., A. W. Bateman, and T. H. Clutton-Brock.
2012. Lifetime growth in wild meerkats: incorporating life history and environmental factors into a standard growth model. *Oecologia* 169:143–153. [Article](#)
- Fournier, D.A., H. J. Skaug, J. Ancheta, J. Ianelli, A. Magnusson, M. N. Maunder, A. Nielsen, and J. Sibert.
2012. AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim. Methods Softw.* 27:233–249. [Article](#)
- Gasbarra, D., M. Pirinen, M. J. Sillanpää, E. Salmela, and E. Arjas.
2007. Estimating genealogies from unlinked marker data:

- a Bayesian approach. *Theor. Popul. Biol.* 72:305–322. [Article](#)
- Gelman, A., and D. B. Rubin.
1992. Inference from iterative simulation using multiple sequences. *Stat. Sci.* 7:457–511.
- Gelman, A., and J. Hill.
2007. Data analysis using regression and multilevel/hierarchical models, 625 p. Cambridge Univ. Press, Cambridge, UK.
- Gelman, A., J. Hwang, and A. Vehtari.
2014. Understanding predictive information criteria for Bayesian models. *Stat. Comput.* 24:997–1016. [Article](#)
- Gelman, A., D. Lee, and J. Guo.
2015. Stan: a probabilistic programming language for Bayesian inference and optimization. *J. Educ. Behav. Stat.* 40:530–543. [Article](#)
- Geweke, J.
1993. Bayesian treatment of the independent Student-*t* linear model. *J. Appl. Econom.* 8:S19–S40. [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](#)
- Kullback, S., and R. A. Leibler.
1951. On information and sufficiency. *Ann. Math. Stat.* 22:79–86.
- Lehman, T. M., and H. M. Woodward.
2008. Modeling growth rates for sauropod dinosaurs. *Paleobiology* 34:264–281.
- Link, W. A., and M. J. Eaton.
2012. On thinning of chains in MCMC. *Methods Ecol. Evol.* 3:112–115. [Article](#)
- Lunn, D., C. Jackson, N. Best, A. Thomas., and D. Spiegelhalter.
2012. The BUGS book: a practical introduction to Bayesian analysis, 399 p. Chapman & Hall/CRC, Boca Raton, FL.
- Marchenko, Y. V., and M. G. Genton.
2010. Multivariate log-skew-*t* elliptical distributions with applications to precipitation data. *Environmetrics* 21:318–340. [Article](#)
- Millar, R.B.
2002. Reference priors for Bayesian fisheries models. *Can. J. Fish. Aquat. Sci.* 59:1492–1502. [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](#)
- Pardo, S.A., A. B. Cooper, and N. K. Dulvy.
2013. Avoiding fishy growth curves. *Methods Ecol. Evol.* 4:353–360. [Article](#)
- Peng, F., and D. K. Dey.
1995. Bayesian analysis of outlier problems using divergence measures. *Can. J. Stat.* 23:199–213. [Article](#)
- Pilling, G. M., G. P. Kirkwood, and S. G. Walker.
2002. An improved method for estimating individual growth variability in fish, and the correlation between von Bertalanffy growth parameters. *Can. J. Fish. Aquat. Sci.* 59:424–432. [Article](#)
- Plummer, M., N. Best, K. Cowles, and K. Vines.
2006. CODA: convergence diagnosis and output analysis for MCMC. *R News* 6(1):7–11. [Available at [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: language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [Available at [website](#), accessed May 2014.]
- Ratkowsky, D. A.
1986. Statistical properties of alternative parameterizations of the von Bertalanffy growth curve. *Can. J. Fish. Aquat. Sci.* 43:742–747. [Article](#)
- Robert, C. P., and G. Casella.
2004. Monte Carlo statistical methods, 2nd ed., 644 p. Springer, New York.
- Sainsbury, K.J.
1980. Effect of individual variability on the von Bertalanffy growth equation. *Can. J. Fish. Aquat. Sci.* 37:241–247. [Article](#)
- Shelton, A.O., and M. Mangel.
2012. Estimating von Bertalanffy parameters with individual and environmental variations in growth. *J. Biol. Dyn.* 6:3–30. [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](#)
- Slatkin, M.
2013. A method for estimating the effective number of loci affecting a quantitative character. *Theor. Popul. Biol.* 89:44–54. [Article](#)
- Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. vander Linde.
2002. Bayesian measures of model complexity and fit. *J. R. Stat. Soc., B* 64:583–639. [Article](#)
- Taylor, N. G., C. J. Walters, and S. J. D. Martell.
2005. A new likelihood for simultaneously estimating von Bertalanffy growth parameters, gear selectivity, and natural and fishing mortality. *Can. J. Fish. Aquat. Sci.* 62:215–223. [Article](#)
- Tjørve, K. M. C., and E. Tjørve.
2010. Shapes and functions of bird-growth models: how to characterise chick postnatal growth. *Zoology* 113:326–333. [Article](#)
- Vidal, I., P. Iglesias, M. D. Branco, and R. B. Arellano-Valle.
2006. Bayesian sensitivity analysis and model comparison for skew elliptical models. *J. Stat. Plann. Inference* 136:3435–3457. [Article](#)
- von Bertalanffy, L.
1938. A quantitative theory of organic growth (Inquiries on growth laws. II). *Hum. Biol.* 10:181–213.
- Wiff, R., and R. Roa-Ureta.
2008. Predicting the slope of the allometric scaling of consumption rates in fish using the physiology of growth. *Mar. Freshw. Res.* 59:912–921. [Article](#)
- Xiao, Y.
1994. von Bertalanffy growth models with variability in, and correlation between, K and L_{∞} . *Can. J. Fish. Aquat. Sci.* 51:1585–1590. [Article](#)
- Zhang, Z., J. Lessard, and A. Campbell.
2009. Use of Bayesian hierarchical models to estimate northern abalone, *Haliotis kamtschatkana*, growth parameters from tag-recapture data. *Fish. Res.* 95:289–295. [Article](#)
- Zhu, J., M. N. Maunder, A. M. Aires-da-Silva, C., and Y. Chen.
2016. Estimation of growth within Stock Synthesis models: management implications when using length-composition data. *Fish. Res.* 180:87–91. [Article](#)