

A stepwise stochastic simulation approach to estimate life history parameters for data-poor fisheries

Marc O. Nadon and Jerald S. Ault

Abstract: Coastal fisheries are typically characterized by species-rich catch compositions and limited management resources, which typically leads to notably data-poor situations for stock assessment. Some parsimonious stock assessment approaches rely on cost-efficient size composition data, but these also require estimates of life history parameters associated with natural mortality, growth, and maturity. These parameters are unavailable for most exploited stocks. Here, we present a novel approach that uses a local estimate of maximum length and statistical relationships between key life history parameters to build multivariate probability distributions that can be used to parameterize stock assessment models in the absence of species-specific life history data. We tested this approach on three fish species for which empirical length-at-age and maturity data were available (from Hawaii and Guam) and calculated probability distributions of spawning potential ratios (SPR) at different exploitation rates. The life history parameter and SPR probability distributions generated from our data-limited analytical approach compared well with those obtained from bootstrap analyses of the empirical life history data. This work provides a useful new tool that can greatly assist fishery stock assessment scientists and managers in data-poor situations, typical of most of the world's fisheries.

Résumé : Les pêches côtières sont typiquement caractérisées par des compositions de prises à forte richesse spécifique et des ressources de gestion limitées, ce qui mène typiquement à des situations de manque de données marqué pour l'évaluation des stocks. Certaines approches parcimonieuses d'évaluation des stocks reposent sur des données de composition selon la taille peu coûteuses, mais ces approches nécessitent également des estimations de paramètres du cycle biologiques associés à la mortalité naturelle, la croissance et la maturité, des paramètres non disponibles pour la plupart des stocks exploités. Nous présentons une nouvelle approche qui utilise une estimation locale de la longueur maximum et des relations statistiques entre des paramètres du cycle biologique clés pour établir des lois de probabilité multivariées qui peuvent être utilisées pour paramétrer des modèles d'évaluation des stocks en l'absence de données sur le cycle biologique propres à l'espèce. Nous avons testé cette approche sur trois espèces de poissons pour lesquelles des données empiriques sur la longueur selon l'âge et la maturité sont disponibles (d'Hawaii et de Guam) et calculé les lois de probabilité du rapport des potentiels de reproduction (RPR) pour différents taux d'exploitation. Les lois de probabilité des paramètres du cycle biologique et du RPR générées sur la base de notre approche analytique pour des données limitées concordent bien avec celles obtenues d'analyses par autoamorçage de données empiriques sur le cycle biologique. Ces travaux fournissent un nouvel outil pouvant être très utile aux chercheurs et gestionnaires qui travaillent à l'évaluation des stocks de poissons dans des situations de données limitées typiques de la plupart des pêches du monde. [Traduit par la Rédaction]

Introduction

Coastal fisheries provide livelihoods and sustenance for hundreds of millions of people worldwide, particularly in poorer tropical countries (Pauly et al. 2005). These fisheries are often characterized by highly diverse catch compositions, sometimes reaching hundreds of species, and are mostly managed with limited financial and human resources, if managed at all. For most of these fisheries, limited information exists on historical catches, fishing effort, and baseline population abundances (Fenner 2012). These issues make assessing these fisheries highly problematic. Recently, methods have been proposed for data-poor situations that rely mainly on cost-effective size composition data and some basic demographic knowledge of growth, maturity, and longevity (Ault et al. 1998, 2008, 2014; Gedamke and Hoenig 2006; Hordyk et al. 2015). However, even these simple data requirements are often unmet because of a lack of life history (LH) information for

either a particular region or even globally. Specifically, Froese and Binohlan (2000) have estimated that only 1200 out of 7000 (about 17%) of exploited species worldwide have some LH data. Here, we propose a novel, standardized approach to obtain probability distributions of LH parameters for under- and unstudied species by employing a stepwise Monte Carlo simulation approach based on a meta-analysis of published parameters and some relatively well-known relationships between individual parameters.

Beverton and Holt (1959) and Beverton (1963) were the first to identify fundamental linkages between LH parameters in fishes. They observed that (i) natural mortality rate (M) is positively correlated with von Bertalanffy's Brody growth coefficient (K), such that the ratio of M/K is typically close to 1.5 (i.e., whereas longer-lived fishes tend to grow more slowly); (ii) length at 50% maturity (L_{mat}) is positively correlated with asymptotic length (L_{inf}), where the ratio L_{mat}/L_{inf} is generally around 0.66 (i.e., species tend

Received 18 June 2015. Accepted 10 May 2016.

M.O. Nadon. Joint Institute for Marine and Atmospheric Research, 1000 Pope Road, Honolulu, HI 96822, USA; University of Miami, Rosenstiel School of Marine and Atmospheric Science, 4600 Rickenbacker Causeway, Miami, FL 33149, USA; Pacific Island Fisheries Science Center, NOAA Fisheries, 1845 Wasp Boulevard, Honolulu, HI 96818, USA.

J.S. Ault. University of Miami, Rosenstiel School of Marine and Atmospheric Science, 4600 Rickenbacker Causeway, Miami, FL 33149, USA.

Corresponding author: Marc Nadon (email: marc.nadon@noaa.gov).

Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from [RightsLink](https://www.copyright.com).

to mature at a specific fraction of their maximum size); and (iii) von Bertalanffy growth model coefficients L_{inf} and K are negatively correlated following a general power function of the form $L_{\text{inf}} \sim K^{-h}$, where h is typically about 0.5. These ratios, generally referred to as Beverton–Holt (BH) invariants (Charnov 1993), have been shown to be related to the maximization of reproductive output (Jensen 1996, 1997; Charnov 2008). As a consequence, these relationships have likely been conserved through natural selection, as originally theorized by Beverton and Holt (1959). Some recent studies (e.g., Prince et al. 2015) noted that these “invariants” may actually differ significantly between taxonomic groups and thus are better retained within taxa.

One important interest in studying these relationships is their potential use in the estimation of elusive LH parameters such as M by relating these to more tractable ones, such as L_{inf} , K , and perhaps water temperature (see Kenchington 2014 for a review of other empirical M relationships). Recent studies have proposed other relationships related to maximum length and optimum length (Froese and Binohlan 2000, 2003; Jarić and Gačić 2012). These studies were mainly concerned with imputing estimates for individual unavailable parameters. However, it is not clear how these relationships might be used to generate complete multivariate distributions for all key LH parameters (e.g., L_{inf} , K , L_{mat} , M) given that these are all typically correlated to some degree. Standard multivariate distributions (e.g., multivariate normal) can describe the variance–covariance structure of multiple parameters, but preclude the use of more complex relationships between parameters (e.g., power, exponential, polynomial) and non-normal error distributions (e.g., lognormal, gamma).

As a solution, we propose using a stepwise stochastic simulation approach that seeks to preserve the inherent correlation structure between these parameters, an approach analogous to the “sequence of regressions” (or fully conditional specification, FCS) method used for multiple imputation of missing data (Raghunathan et al. 2001; van Buuren 2007; Ellington et al. 2015). To accomplish this, we first reviewed the literature on the LH of six commonly targeted families of coral reef fishes to generate family-specific models linking the four main LH parameters: $L_{\text{inf}} \sim L_{\text{max}}$, $K \sim L_{\text{inf}}$, $M \sim K$, and $L_{\text{mat}} \sim L_{\lambda}$, where L_{max} is a local estimate of maximum length and L_{λ} is the expected length at oldest recorded age (i.e., the length predicted by a specific growth curve for the oldest age; see Table 1 for definitions). With these models, it is theoretically possible to generate the underlying complex multivariate probability distributions for these parameters by taking successive random samples from the estimated regression models, starting with a local estimate of L_{max} . Here, we tested this approach for three well-studied species for which empirical data on lengths, ages, and maturity were available. We used bootstrap analyses to generate probability distributions of LH parameters from biological studies and compared these distributions with those obtained through our stepwise approach. Finally, we derived spawning potential ratios (SPRs) at various fixed exploitation rates using these imputed LH distributions and compared the two resulting data sets.

Methods

The approach presented here only requires a local estimate of L_{max} to generate probability distributions of missing LH parameters. To do so, it is first necessary to model the family-level relationships between these parameters using published estimates from growth and maturity studies.

Table 1. List of life history parameters with definitions.

Parameter	Definition
L_{λ}	Expected length at the oldest recorded age
L_{mat}	Length at which 50% of females reach maturity
L_{max}	Longest length in a growth study or 99th percentile of lengths in a population survey
L_{inf}	Expected length at infinite age
K	Brody growth coefficient of the von Bertalanffy growth curve
a_0	Theoretical age at which length equals zero
a_{λ}	Oldest recorded age (i.e., longevity)
M	Instantaneous natural mortality rate

Note: All lengths are total lengths.

LH parameter data sets

We obtained LH parameters from a synthesis of the literature for six commonly exploited families of coral reef fishes: surgeonfishes (Acanthuridae), jacks (Carangidae), emperors (Lethrinidae), snappers (Lutjanidae), goatfishes (Mullidae), and parrotfishes (Scaridae) (see online Supplementary Material S1¹ for sources). Unfortunately, we could not assemble a data set composed of raw length, age, and maturity data given that such information is rarely published and generally difficult to obtain for the large number of species that we targeted. These data could have been useful in a hierarchical meta-analysis modeling context (Helsler et al. 2007; Thorson et al. 2014, 2015). Instead, we obtained fitted parameter estimates for growth (L_{inf} , K , a_0) and maturity (L_{mat}), in addition to longevity (a_{λ}) and L_{max} (see Table 1 for definitions). In all of these studies, length dependent on age data were fitted using the von Bertalanffy growth equation:

$$(1) \quad L_a = L_{\text{inf}}[1 - e^{-K(a-a_0)}]$$

where L_a is the expected length at age a . Although we could have potentially taken L_{max} values from other local data sources (e.g., visual surveys, catch records), we had limited access to such data sets. We also avoided using reported “world-record” lengths as L_{max} , since these can represent anomalous individuals and are not reported for all species. If multiple studies existed for a single species, we kept the parameters from the most in-depth and recent studies, with the exception of longevity, for which we kept the greatest value found in any study. If an individual study provided separate parameters for different localities, we averaged these parameters (if the study had not already provided the averaged values). Most growth studies either did not attempt to separate sexes (67%) or did not identify differences in growth between sexes (28%). Only 5% of studies reported some differences in growth between sexes. For these, growth parameters fitted separately for males and females were averaged together. On the other hand, maturity studies were generally sex-specific; for species with different length-at-maturity between sexes, the female L_{mat} was retained (since spawning potential ratio is related to female spawning biomass; see details further down). Parameters L_{inf} , L_{mat} , and L_{max} reported in standard or fork lengths were converted to total length (TL, mm) using published conversion factors. Time-based parameters (K , a_0 , a_{λ}) expressed in units other than years were converted to that time unit. Finally, natural mortality rates M were estimated from longevity by applying the procedure of Alagaraja (1984), similar to Hoenig (1983) and Hewitt and Hoenig (2005), assuming that 5% of a cohort survives to the observed maximum age (a_{λ}):

¹Supplementary data are available with the article through the journal Web site at <http://nrcresearchpress.com/doi/suppl/10.1139/cjfas-2015-0303>.

Table 2. Summary of life history parameter data sets by family (L_{inf} , K , and a_λ show the ranges of these parameters).

	Acanthuridae (surgeonfishes)	Carangidae (jacks)	Lethrinidae (emperors)	Lutjanidae (snappers)	Mullidae (goatfishes)	Scaridae (parrotfishes)	All families
Species count	81	148	38	110	82	100	559
Growth studies	27	40	22	39	12	27	167
Maturity studies	11	20	14	22	11	19	97
SST < 20 °C	0	10	0	0	3 ^a	2	15
Excluded outliers	1	0	1	1	1	2	6
L_{inf} (mm)	142–619	255–2170	180–712	216–1264	187–547	116–818	116–2170
K (year ⁻¹)	0.2–1.2	0.1–0.4	0.1–0.9	0.1–1.0	0.2–0.8	0.1–1.8	0.1–1.8
a_λ (years)	20–53	5–23	7–36	8–55	3–11 ^b	5–33	3–55
L_{max}/L_{inf}	1.1 (0.1)	0.9 (0.2)	1.2 (0.1)	1.1 (0.1)	0.9 (0.1)	1.2 (0.1)	1.1 (0.2)
M/K	0.3 (0.1)	1.6 (0.7)	0.5 (0.3)	0.8 (0.6)	1.6 (0.6)	0.5 (0.3)	0.8 (0.7)
L_{mat}/L_λ	0.79 (0.08)	0.63 (0.08)	0.78 (0.07)	0.67 (0.12)	0.63 (0.11)	0.69 (0.07)	0.70 (0.11)
L_{mat}/L_{inf}	0.79 (0.08)	0.55 (0.11)	0.78 (0.07)	0.64 (0.13)	0.55 (0.09)	0.68 (0.08)	0.66 (0.13)

Note: See Table 1 for parameter definitions. All lengths are total lengths. SST < 20 °C refers to studies where annual sea surface temperature was lower than 20 °C on average.

^aKept for growth and L_{mat} , but not longevity estimate.

^bIncludes cold-water species.

$$(2) \quad M = \frac{-\ln(0.05)}{a_\lambda}$$

We used the 5% cohort survivorship value based on the analyses of Nadon et al. (2015), which showed that this is an appropriate survivorship value for coral reef fishes. We did not have independent estimates of M per se and had to rely on this longevity-based approach. Although there are other data-poor methods for estimating natural mortality, involving other parameters (e.g., K , L_{inf} , L_{mat} , water temperature), two recent comprehensive reviews on the subject clearly suggest that longevity-only methods perform the best (Kenchington 2014; Then et al. 2015). These studies also raised two important points concerning this approach: (i) the importance of selecting an appropriate survivorship value; and (ii) the potential difficulty in obtaining a representative longevity value in heavily exploited stocks. Regarding the first point, although we selected a survivorship value of 0.05, for the reason explained above, the M estimates obtained using our data-poor approach can easily be converted to M estimates based on a different survivorship value by simply converting M to longevity (by inverting eq. 2) and recalculating M from longevity using a different survivorship value. Regarding the second point, we tried to alleviate this concern as much as possible by selecting the oldest recorded age, regardless of location, as our measure of longevity. It was, unfortunately, impossible to only select longevity estimates from unexploited stocks given that these are, of course, extremely rare.

We rejected studies where growth parameters came from length–frequency or mark–recapture analyses (i.e., we only kept studies based on hard structure ageing, which are generally considered more reliable), a_0 was very negative (< -2 years), the authors expressed concerns about their results, or parameter estimates were derived from empirical relationships. We also rejected LH studies from regions with annual mean sea surface temperatures below 20 °C because of the well-known effects of water temperature on growth and longevity (Pauly 1980; Jobling 1994; Choat and Robertson 2002). We decided not to add temperature as a variable in our regression models, since the families targeted in this current study were mostly tropical, resulting in only a few studies that were discarded (15 out of 167, mostly jacks). However, we did retain growth and maturity information from two goatfish studies from colder water (<20 °C) regions (i.e., Mediterranean Sea) because of a low sample size for this family, but we did not retain their longevity values. It is also important to note that we did not attempt to control for intraspecific regional differences in LH parameters; thus, we analyzed all parameters globally. This was due mainly to the limited number of growth and

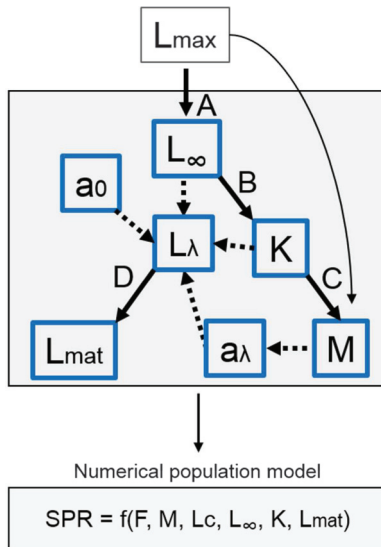
maturity studies, where most species have one set of LH parameters from a single location. Intraspecific regional variability in LH parameters are thus simply added to the overall error variability in our models, in addition to interspecific variability (where most of the variability in LH parameters is likely to be found).

After completing the literature review described above, we found growth information for 167 of the 560 species that are currently listed for the six families included in our analyses (Table 2). Of those 167, we found size-at-maturity information for 97 species. A total of 15 species were discarded because their growth and longevity information came from regions where mean sea surface temperature was below 20 °C. Furthermore, six species were discarded in some of the models for being clear outliers. The L_{mat} values for *Lethrinus laticaudis* (Ayvazian et al. 2004) and *Lutjanus campechanus* (White and Palmer 2004) corresponded to very low L_{mat}/L_{inf} ratios (0.46 and 0.37, respectively), which were far lower than for other species in their family. These ratios were also far lower than the predicted ratio based on the theoretical maximization of reproductive output (~0.66; Jensen 1996). We removed the L_{inf} and K values for the goatfish *Mulloidichthys vanicolensis*, since they were clear outliers; the ageing for this species was done using unvalidated annual otoliths rings, which may be a reason for this discrepancy (Cole 2009). Similarly, we discarded growth and longevity parameters for the surgeonfish *Acanthurus triostegus*, since Longenecker et al. (2008) used (presumed) daily growth rings that resulted in a lifespan estimate much lower than anything previously recorded for surgeonfishes (3 versus 20 years for the next surgeonfish species with the shortest reported lifespan). We removed the small parrotfish *Sparisoma atomarium* from our meta-analysis, since this is an extremely small species (L_{inf} = 110 mm) with very fast growth (K = 1.8) and short lifespan (3 years) that could not be properly modeled with other parrotfishes and was a clear outlier. Finally, we removed a very low (and unlikely) longevity estimate for the parrotfish *Calotomus carolinus* from a study in Guam (Taylor and Choat 2014). Other parrotfish longevity estimates in this study were low compared with values reported elsewhere. For example, the parrotfish *Scarus rubroviolaceus* has a maximum lifespan estimate of 22 years around Oahu, Hawaii (Howard 2008), and 20 years in the Seychelles (Grandcourt 2002), while the estimate for the same species around Guam was only 6 years (Taylor and Choat 2014).

LH parameter models

A preliminary look at the distribution of LH parameters and Beverton–Holt invariants (Charnov 1993) revealed important differences in the distributions and ranges of certain parameters between families (Table 2), which is consistent with observations

Fig. 1. Diagram presenting one iteration of the stepwise simulation chain used to obtain a single spawning potential ratio (SPR) estimate. Solid arrows represent steps that are derived using the four modeled relationships (represented by a letter). Dashed arrows represent deterministic steps where certain values are used to calculate others.



from previous studies (Beverton and Holt 1959; Choat and Robertson 2002; Prince et al. 2015). We therefore conducted all analyses at the family level.

We built statistical regression models for the empirical relationships between pairs of LH parameters for each of the six families (Bolker 2008; Fox 2008). These models were built using TL reported in millimetres. We fitted maximum-likelihood models to the parameter pairs with the closest relationships: $L_{\text{inf}} \sim L_{\text{max}}$, $L_{\text{mat}} \sim L_{\text{inf}}$, $K \sim L_{\text{inf}}$, and $M \sim K$ (Beverton and Holt 1959; Jensen 1996; Froese and Binohlan 2000, 2003). We also tested the use of a two-variable model to predict M , which included the L_{max} variable (in addition to K , similarly to the model described in Pauly 1980). Additionally, we tested L_{λ} (expected length at the oldest measured age) as a replacement for L_{inf} as a predictor of L_{mat} (Fig. 1). For all four models, we tested three deterministic functions (linear, power, and exponential) and two error probability distributions (normal and lognormal). We selected the best model based on an in-depth review of scatterplots and other diagnostics (to select an adequate linear or nonlinear function) and residual error distributions (to select an appropriate probability distribution). Based on published empirical and theoretical results, we did expect certain relationships to be linear ($L_{\text{inf}} \sim L_{\text{max}}$ and $L_{\text{mat}} \sim L_{\text{inf}}$ or L_{λ}) and others to be curvilinear ($K \sim L_{\text{inf}}$ and $M \sim K$), and these relationships were typically explored first. However, we did not assume a priori any model structure (i.e., functional form or error distribution); we simply let the data and diagnostics determine the most appropriate model. For curvilinear relationships, we log-transformed the response variable (K or M) and fitted either an exponential or power function with a lognormal error distribution. The a_0 parameter was originally described using a triangular probability distribution independently from the other parameters, as we did not observe any significant correlation between this parameter and L_{inf} or K (as opposed to Helsen et al. 2007). However, after running some preliminary analyses, we found no measurable impact of a_0 variability on L_{λ} and L_{mat} , and we simply fixed this parameter to -0.3 (the overall average for the species in our data set).

Once these four LH parameter models were established for each family, we used the iterative approach described in Fig. 1 to build

multivariate probability distributions of the four principal LH parameters (L_{inf} , K , M , and L_{mat} ; referred to as the “data-poor” approach in the rest of the document). The first step in this process was to draw a random L_{max} and use the $L_{\text{inf}} \sim L_{\text{max}}$ model (step A in Fig. 1) to draw a random L_{inf} . From this random L_{inf} , a K value was drawn from the $K \sim L_{\text{inf}}$ model (step B), which was then used to draw a random M from the $M \sim K + L_{\text{max}}$ model (step C). Finally, we tested two models to draw a random L_{mat} : the $L_{\text{mat}} \sim L_{\text{inf}}$ and $L_{\text{mat}} \sim L_{\lambda}$ models (step D). For the latter model, we had to first calculate L_{λ} using the von Bertalanffy equation from the L_{inf} , K , a_0 , and longevity parameters drawn during the same iteration (longevity being derived back from M using eq. 2). We tested both model fits (using the coefficient of determination r^2) and ultimately decided on the $L_{\text{mat}} \sim L_{\lambda}$ model, which is the reason why it is presented as step D in Fig. 1 (see Results for more details). It is important to note that the error distributions do not include the uncertainty in the regression coefficients (Fox 2008). To do so, we used the regression coefficient variance–covariance matrix to draw random sets of coefficients from a multivariate normal distribution and then calculated expected values at each iteration. This entire process (Fig. 1) was repeated for 5000 iterations to build multivariate distributions describing the variance and covariance of all four LH parameters. At each iteration, we also calculated SPRs (see next section).

Testing the data-poor approach

We evaluated the precision and accuracy of the data-poor approach by comparing the probability distributions of LH parameters obtained in this way versus the probability distributions originating from well-conducted growth and maturity studies (referred to as the “data-rich situation” in the rest of the text). We selected three growth and maturity studies from Hawaii and Guam for which we had access to empirical length, age, and maturity data: *Scarus rubroviolaceus* (Howard 2008), *Naso unicornis* (Eble et al. 2009), and *Lethrinus harak* (Taylor and McIlwain 2010). We first used the approach described above to obtain data-poor probability distributions of LH parameters for those three species. The L_{max} values used for this approach came from an extensive underwater visual census data set collected by the National Oceanic and Atmospheric Administration in Hawaii and Guam (Williams et al. 2011). We used the 99th percentile of lengths in this data set to obtain L_{max} to reduce the risk of using spurious length measurements from divers. Furthermore, since there was uncertainty associated with L_{max} estimates derived from our survey data set, we used a bootstrapping approach to estimate the L_{max} probability distributions from which random values could be sampled.

To generate probability distributions for the data-rich situation, we bootstrapped the empirical length, age, and maturity observations for the three selected studies, since these did not report the error associated with their fitted parameter estimates. Fish collected for growth and maturity studies are typically selected to be representative of an entire species’ size range (i.e., once researchers have sufficient average-sized specimens, they shift their focus to capturing small and large individuals until they have a representative size range). To reflect this, we stratified the bootstrapping of the raw data set for each species in three size categories (small — bottom 30% of length range, medium, and large — top 30% of length range). We ran the bootstrapping procedure for 5000 iterations to obtain probability distributions of the LH parameters presented in these studies.

We compared the probability distributions of all four LH parameters obtained in the way described above with the ones obtained using our novel data-poor approach. To assess precision, we compared the widths of the data-poor distributions (using standard deviations, SD) with those of the data-rich distributions (SD ratio = $\text{SD}_{\text{data-poor}}/\text{SD}_{\text{data-rich}}$). To assess accuracy, we calculated the standardized distance between the medians of the data-poor and

Table 3. Model functional form, error probability distribution, and variance–covariance matrix of regression coefficients for models relating various life history parameters.

Family	(A) $L_{\text{inf}} \sim L_{\text{max}}$	(B) $K \sim L_{\text{inf}}$	(C) $M \sim K + L_{\text{max}}$	(D) $L_{\text{mat}} \sim L_{\lambda}$
Acanthuridae	[L,N] -39.4, 1.02, (18.1) 137, 1.04e-3, -0.345	[E,LN] -0.278, -1.80e-3, (0.420) 3.63e-2, 2.42e-7, -8.49e-5	[I,N] 9.77e-2, (2.28e-2) 1.99e-5	[L,LN] -12.6, 0.828, (0.102) 194, 3.05e-3, -0.693
Carangidae, small (<900 mm)	[L,LN] 28.0, 1.00, (0.130) 1.65e+3, 8.12e-3, -3.38	[E,LN] -1.10, -4.60e-4, (0.359) 1.77e-2, 2.21e-8, -1.72e-5	[P,LN] 2.27, 3.02e-2, -0.515, (0.307) 0.764, 2.20e-2, 1.94e-2, -2.14e-3, -0.118, 5.21e-3	[L,N] 33.6, 0.573, (44.0) 542, 8.60e-4, -0.612
Carangidae, large (>900 mm)	[L,LN] -27.9, 1.20, (0.242) 8.62e-9, 1.15e-2, 9.94e-6	— —	— —	— —
Lethrinidae	[L,N] -1.44, 0.872, (17.1) 162, 7.53e-4, -0.334	[P,LN] 4.66, -0.930, (0.439) 3.13, 8.87e-2, -0.526	[P,LN] 0.541, 0.538, -0.284, (0.297) 2.12, 2.49e-2, 6.24e-2, -9.77e-2, -0.362, 1.94e-2	[L,N] 36.0, 0.672, (21.3) 396, 2.86e-3, -1.02
Lutjanidae, small (<500 mm)	[L,N] -5.08, 0.912, (21.3) 952, 7.72e-3, -2.65	[P,LN] 5.23, -1.06, (0.397) 0.760, 1.90e-2, -0.120	[P,LN] 2.04, -8.64e-3, -0.631, (0.393) 1.72, 2.51e-2, 5.32e-2, -0.139, -0.300, 2.73e-2	[L,LN] 63.9, 0.519, (0.148) 423, 2.16e-3, -0.853
Lutjanidae, large (>500 mm)	[L,N] -81.8, 1.06, (85.9) 6711, 1.08e-2, -8.30	— —	— —	— —
Mullidae	[L,N] -9.34, 1.16, (16.7) 519, 7.94e-3, -1.97	[I,N] 0.467, (0.173) 2.71e-3	[I,LN] 0.674, (0.348) 6.90e-3	[L,N] 49.3, 0.415, (22.1) 345, 4.36e-3, -1.14
Scaridae	[L,N] 0.191, 0.849, (23.2) 379, 2.09e-3, -0.848	[E,LN] 0.681, -3.54e-3, (0.331) 2.68e-2, 1.93e-7, -6.59e-5	[P,LN] -0.769, 0.646, -3.95e-2, (0.242) 3.47, 3.70e-2, 0.105, -0.258, -0.602, 4.64e-2	[L,N] -7.97, 0.714, (22.0) 183, 1.15e-3, -0.425

Note: For each family, the first line presents model functional forms and error distributions in brackets. Functional forms are I = intercept-only ($Y = b_0$), L = linear ($Y = b_0 + b_1 \cdot X$), E = exponential ($\log Y = b_0 + b_1 \cdot X$), P = power ($\log Y = b_0 + b_1 \cdot \log X$). Error distribution types are N = normal and LN = lognormal. The first line also shows regression coefficients followed by the standard deviation parameter for each distribution in parentheses. The second line presents the variance–covariance matrix of the regression parameters for one-variable models (b_0 variance, b_1 variance, b_0 – b_1 covariance) and two-variable models (b_0 variance, b_1 variance, b_2 variance, b_0 – b_1 covariance, b_0 – b_2 covariance, b_1 – b_2 covariance).

data-rich distributions where relative error (RE) = $(\text{median}_{\text{data-poor}} - \text{median}_{\text{data-rich}}) / \text{median}_{\text{data-rich}} \times 100$.

For both the data-poor and data-rich situations, we calculated SPR, a measure of stock sustainability status, at various fishing mortality rates ranging from close to zero to $4 \times M$, where M was derived from the oldest measured age for all three species (12 years for *L. harak*, Taylor and McIlwain 2010; 54 years for *N. unicornis*, A. Andrews, personal communication; 22 years for *S. rubroviolaceus*, Howard 2008). We used an age-structured numerical population model to make these calculations. In this model, numerical abundance at age a was estimated through the use of an exponential mortality function. Length-at-age was derived using the von Bertalanffy growth equation. This numerical model was used to obtain spawning stock biomass (SSB) by summing over individuals in the population between the age of sexual maturity (a_m ; age where 50% of individuals are mature, with knife-edge assumption) and 1.5 times the oldest recorded age (a_{λ}):

$$(3) \quad \text{SSB} = \sum_{a_m}^{1.5a_{\lambda}} \bar{N}_a \bar{W}_a$$

where \bar{N}_a is the mean abundance at age a , and \bar{W}_a is the mean weight of individuals at age a . The model was run using weekly time steps. Mean abundance at age was modeled using the following equation:

$$(4) \quad \bar{N}_{a+\Delta a} = \bar{N}_a e^{-(S_a F + M)\Delta a}$$

where F and M are fishing and natural mortality, respectively. S_a , proportional selectivity at age a , was treated as “knife-edged” selectivity and was derived from a specific length based on the size composition of the catch and the translation of length to age via the von Bertalanffy growth relationship (similarly to Nadon et al. 2015). We used data from the Hawaii commercial trip report and Marine Recreational Information Program for *S. rubroviolaceus* and *N. unicornis* and from the Guam creel survey data for *L. harak* to look for a clear jump in selectivity in the smaller size bins (i.e., a disproportionate increase in abundance from the previous size bin) to select an appropriate size at first capture. These size-at-first-capture values were 260 mm TL for *S. rubroviolaceus* and *N. unicornis*

and 160 mm TL for *L. harak*. We examined the sensitivity of our model comparisons to different selectivity values and did not find important differences. Weight (W) dependent on length (L) relationship parameters (α , β , where $W = \alpha \cdot L^{\beta}$), necessary for SPR calculations, were obtained from the literature (e.g., Kulbicki et al. 2005). These values are easily calculated and are typically available from local studies. We obtained these values from the same studies that provided growth and maturity information for our three test species. Finally, a stock’s theoretical maximum reproductive biomass occurs when there is no fishing (i.e., $F = 0$). SPR at various fishing mortality rates F was computed as the ratio of the SSB at a given F relative to that of an unexploited stock:

$$(5) \quad \text{SPR}_F = \frac{\text{SSB}_F}{\text{SSB}_{F=0}}$$

We compared the SPR distributions between the data-poor and data-rich scenarios with the same approach used for the LH parameter distribution comparisons.

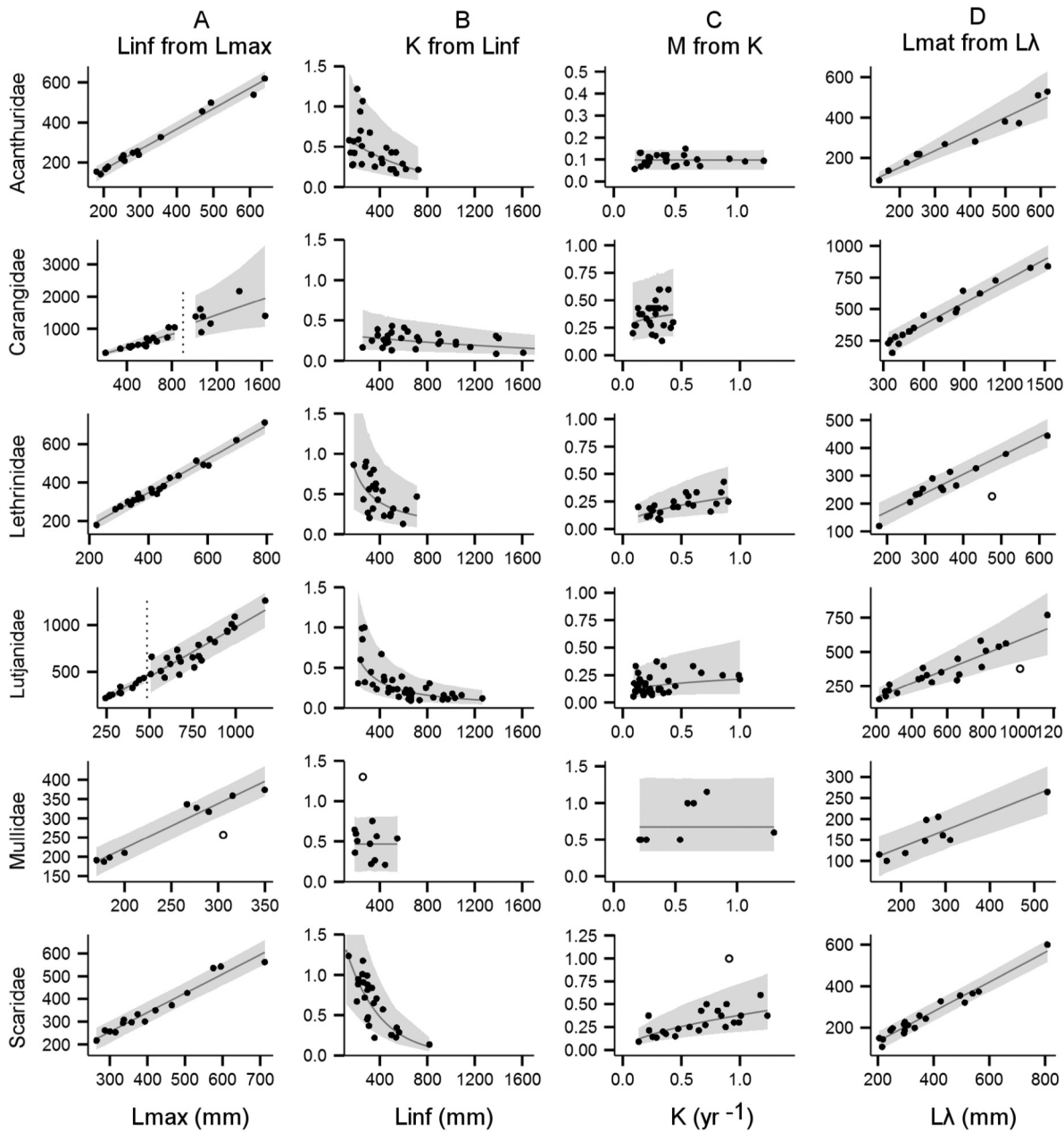
Finally, owing to concerns with biased L_{max} values in heavily fished stocks, we tested the effect of using L_{max} values 10%, 20%, and 30% smaller than the “true” value on LH parameters and SPR estimates.

Results

LH parameter models

Maximum size in a representative local sample (L_{max}) was generally a good predictor of L_{inf} with a reasonably small standard error of about 20 mm for most families (Table 3). We used a normal error distribution for most families, except for jacks where we used a lognormal distribution to take into account increasing L_{inf} variability with L_{max} (i.e., the variance of the lognormal distribution is proportional to the square of its mean). We also had to use a size break point in our models for jacks and snappers, as there was an important increase in variability in L_{inf} beyond certain L_{max} values that could not be taken into account by the properties of the lognormal distribution (900 mm TL for jacks and 500 mm TL for snappers; Fig. 2). The growth parameter K followed a decreasing, curvilinear trend with increasing L_{inf} , which was described using either an exponential or power function, depending on the family (Fig. 2). Variability in K generally decreased with

Fig. 2. Modeled statistical relationships between four life history parameter pairs (columns) for six fish families (rows). Gray areas are 95% confidence intervals; open circles represent removed outliers (see text for justifications). Model C also includes the L_{\max} variable for most families, but only the one-variable $M \sim K$ models are presented here. Vertical dotted lines represent model break points between size categories.



increasing L_{inf} which was properly described by a lognormal error distribution (Table 3). Goatfishes were an exception with no clear trend between L_{inf} and K or increases in variability with L_{inf} . K values for this family were simply described by their mean (0.47) with a normal error distribution (SD = 0.17; Table 3). Natural mortality (M) generally followed a weak increasing curvilinear trend with higher K values and a weak negative curvilinear trend with L_{max} (i.e., larger species in general grew slower and lived longer; Fig. 2). This was true for jacks, snappers, emperors, and parrotfishes (Fig. 2; Table 3). However, there was no clear relationship between M and K or L_{max} for either surgeonfishes or goatfishes (i.e., species in these families had a similar range of longevity regardless of size). Surgeonfishes were all fairly long-lived, with lifespans between 20 years (*Acanthurus chirurgus*, a Caribbean species) and 54 years (*N. unicornis*, a widespread Indo-Pacific species).

Conversely, goatfishes were all short-lived (longest lifespan of 6 years or 11 years if colder water species are included). Lastly, L_{λ} (expected size at maximum age) was an equal or better predictor of L_{mat} for all families compared with L_{inf} , especially for jacks and goatfishes (Table 4). The r^2 value for the $L_{\text{mat}} \sim L_{\lambda}$ regression was 0.91 versus 0.83 for the $L_{\text{mat}} \sim L_{\text{inf}}$ regression for all families combined. We therefore selected L_{λ} as our predictor of L_{mat} instead of the typical L_{inf} . A simple linear equation with a normal error distribution fitted the $L_{\text{mat}} \sim L_{\lambda}$ data for all families, except surgeonfishes and snappers, where a lognormal error distribution was used because of the increasing variance with larger L_{mat} values (Table 3; Fig. 2). The variance-covariance matrices for all model parameters are available online in the supplementary material (Supplementary Material S2¹), and an R code example of our approach is also available online (Supplementary Material S3¹).

Fig. 3. Probability distributions of life history parameters from the data-rich (red areas; solid lines) and data-poor (blue areas; dashed lines) situations for the three selected fish species. [Colour online.]

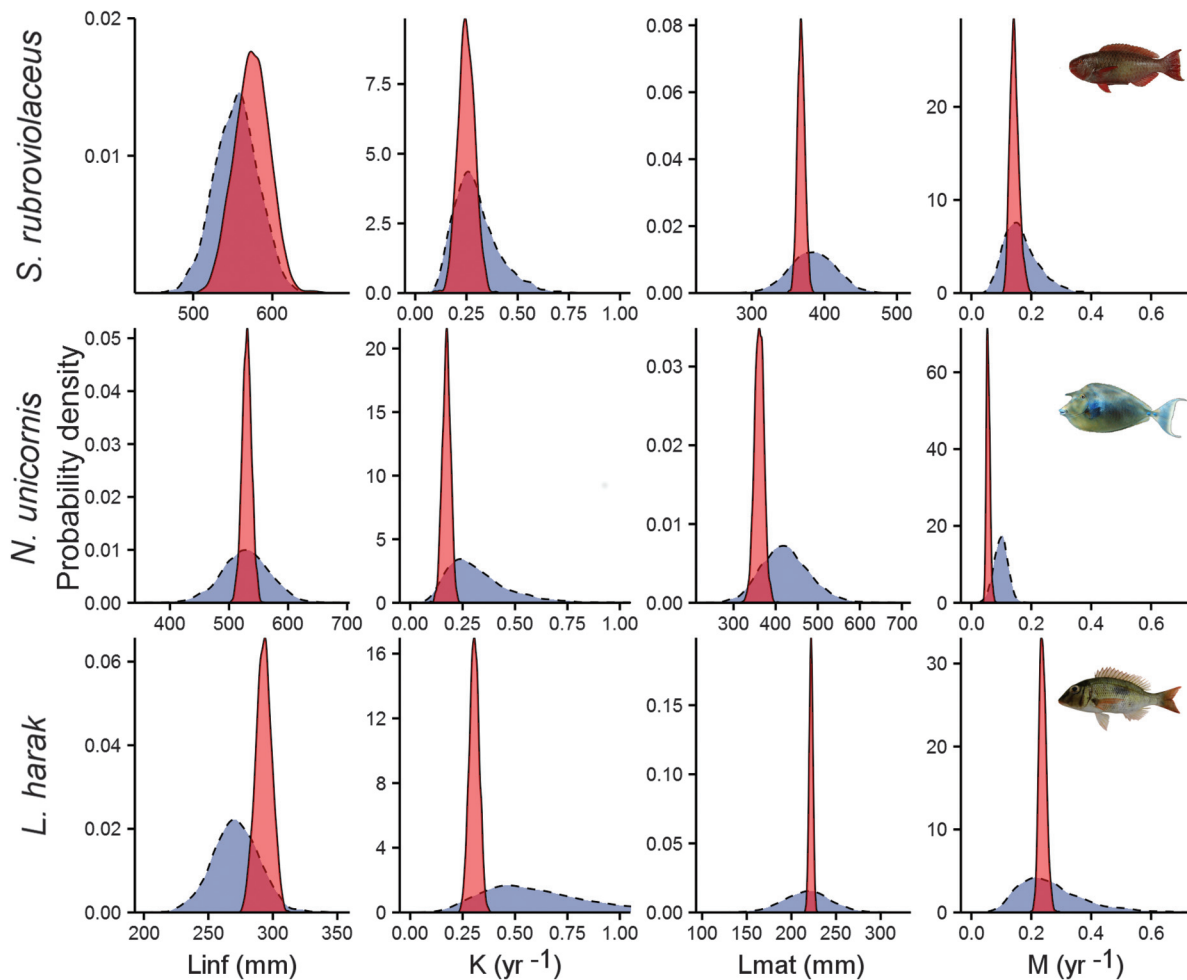


Table 4. Family and overall coefficient of determination (r^2) values for the linear models relating length-at-maturity (L_{mat}) dependent on either expected length at infinite age (L_{inf}) or expected length at oldest recorded age (L_{λ}).

Family	$L_{\text{mat}} \sim L_{\text{inf}}$	$L_{\text{mat}} \sim L_{\lambda}$
Acanthuridae (surgeonfishes)	0.95	0.95
Carangidae (jacks)	0.85	0.95
Lethrinidae (emperors)	0.93	0.93
Lutjanidae (snappers)	0.79	0.77
Mullidae (goatfishes)	0.77	0.84
Scaridae (parrotfishes)	0.94	0.96
All families	0.83	0.91

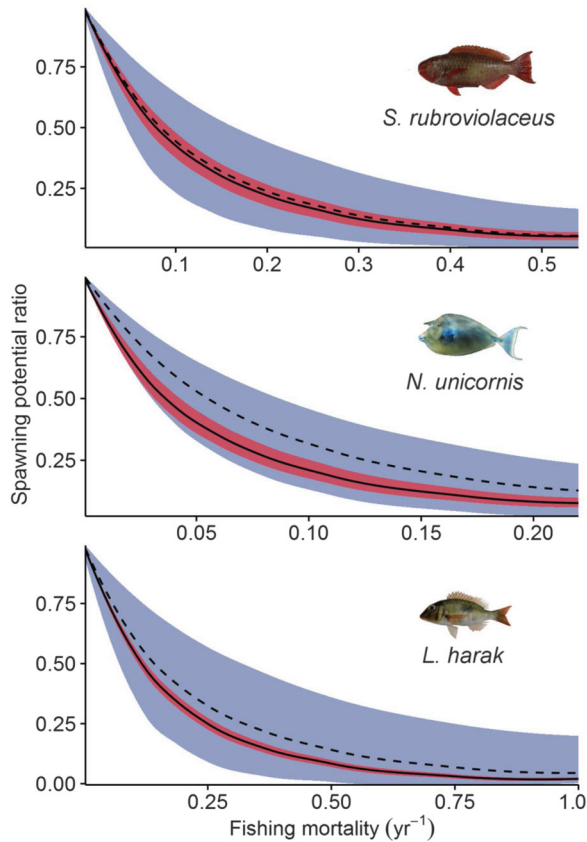
Test of the data-poor approach

For all three test species, we compared LH parameter and SPR probability density distributions obtained through the data-poor approach with those from the data-rich situation. Figures 3 and 4 present these distributions for these species, and Table 5 presents the descriptive statistics associated with these distributions. The LH parameter estimates were all more variable in the data-poor situation compared with the data-rich situation (on average six times larger SD; Table 5). This was especially true for the growth parameter K (eight times larger SD, on average, especially for *L. harak*) and L_{mat} parameter (eight times larger SD, on average). The data-poor approach provided estimates of the L_{inf} and M pa-

rameters that were generally close to the data-rich estimates in terms of variability and accuracy, although the M estimates for *L. harak* were fairly variable (Fig. 3; Table 5). In general, the data-poor parameter estimates for the parrotfish *S. rubroviolaceus* were the most accurate and precise, with probability distributions closer to those originating from the data-rich situation for all parameters (mean SD ratio ~ 3 and mean RE $\sim 8\%$). The estimates for the surgeonfish *N. unicornis* were slightly less precise and accurate than those for *S. rubroviolaceus* when compared with the data-rich situation (mean SD ratio ~ 5 and mean RE $\sim 40\%$; Table 5), especially for the growth parameter K . Both the L_{mat} estimates and M estimates were slightly higher than the data-rich estimates. Finally, the LH comparisons from the data-poor approach fared worst compared with the data-rich situation for the emperor *L. harak* (mean SD ratio ~ 11 and mean RE $\sim 25\%$), mainly due to the K parameter (Table 5). It is important to note that the SD from data-rich situations were smaller for this species, likely owing to the high sample sizes used in the growth-maturity study ($n = 409$), and this makes the data-poor results for *L. harak* appear worse than for the other two species.

The medians of the data-poor SPR distributions for both the emperor *L. harak* and the parrotfish *S. rubroviolaceus* were fairly close to those of data-rich distributions (RE $\sim 15\%$). The median value of the data-poor SPR distributions for the surgeonfish *N. unicornis* were less accurate, being larger than those for data-rich distributions, especially at $F = M$ (median SPR of 0.51 versus 0.38, respectively; RE $\sim 34\%$; Fig. 4; Table 5). The variability of the

Fig. 4. Spawning potential ratio (SPR) at various fishing mortality rates (F) using life history parameters from the data-rich (red areas; solid lines) and data-poor (blue areas; dashed lines) situations. Fishing mortality rates ranged from zero to four times natural mortality (M). Lines represent median values, and shaded areas represent 95% confidence intervals. [Colour online.]



SPR probability density distributions for both the data-poor and data-rich approaches varied with fishing mortality, with lower variability at low ($0.1M$) and high ($2M$) fishing mortality rates (Table 5). The probability density distributions at moderate fishing mortality rates ($F = M$) were the most variable (Fig. 4). The SDs from the data-poor approach were, on average, seven times as wide as those from the data-rich approach, which is a similar mean ratio than for the LH distributions (six times wider on average). The data-poor and data-rich SPR distributions for the emperor *L. harak* were generally less precise than those from the other two species, similarly to the LH parameter comparisons (Fig. 4; Table 5).

Finally, the effects of negatively biased L_{\max} values were strongest on K and M estimates and led to these parameters being positively biased (Table 6). The obvious exception was the surgeonfish *N. unicornis*, where M was modeled as a simple mean, independent of K and L_{\max} (Table 6). For L_{\inf} and L_{mat} , the biases were negative and directly proportional to the L_{\max} bias (i.e., a 20% bias in L_{\max} lead to a 20% bias in these parameters). This pattern was expected given the linear relationship between these parameters. It is interesting to note that the simulated negative bias in L_{\max} consistently led to a positive bias in SPR. The positive bias in K and negative bias in L_{mat} will push SPR values lower, as both age-at-first-capture and age-at-maturity are reached more quickly (i.e., a greater portion of the mature population is exposed to fishing). However, the negative bias in L_{\inf} and positive bias in M will result in higher SPR values, as age-at-first-capture is reached more slowly and the relative contribution of F to total mortality is re-

duced. From the results shown in Table 6, it seems that biased L_{\inf} and M parameters have a greater influence on SPR than biased K and L_{mat} parameters, thus leading to an overall positive bias in SPR. This is also evident from the smaller impact of biased L_{\max} values on SPR for *N. unicornis* due to M being independently estimated for surgeonfishes (Table 6).

Discussion

This study presents a novel way of obtaining probability distributions for key LH parameters (L_{\inf} , K , L_{mat} , and M) in data-poor situations where the only available information for a stock is taxonomic group (i.e., family) and a local estimate of maximum size. It is well-known that these parameters are correlated to varying degrees (Beverton and Holt 1959; Beverton 1992; Charnov 1993) and therefore that their probability distributions should not be described independently (e.g., species with large L_{\inf} typically have large L_{mat} , but small K and M values). The relatively intuitive stepwise approach presented here allows researchers to build relatively complex multivariate probability distributions for these parameters that preserve their correlative structure and associated benefits (Pulkkinen et al. 2011). These distributions can be used to parameterize population models or provide prior information for Bayesian analyses. Although our study was focused on six families found in tropical coastal areas, this approach is quite flexible and can be easily extended to other families in other geographical regions (e.g., groupers, Pacific rockfishes, etc.).

Our stepwise approach is analogous to the FCS method (Raghunathan et al. 2001; van Buuren 2007) used for multiple imputation. Multiple imputation seeks to generate missing values in an incomplete data set by calculating the posterior distribution of the variables and then drawing samples at random from this distribution (Rubin 1987). This can be accomplished through joint modeling under the multivariate normal distribution (or other distributions), which is theoretically elegant, but inflexible to important data features such as nonlinearity, censoring, and non-normal distributions (van Buuren 2007). FCS removes this important limitation by imputing missing data variable-by-variable using a specific model for each variable. This allows complex models to be built for each variable (Raghunathan et al. 2001). In our stepwise approach, we do not regress individual variables against all others, but rather focus on the key relationships between variables (i.e., LH invariants). One reason for doing so was that it is unlikely that modeling a variable such as L_{mat} using variables beyond L_{λ} would improve our predictions given that L_{λ} explained up to 96% of the variability in L_{mat} . Another difference was that our approach did not seek to fill all missing values in a data set to run further analyses on this completed data set, but instead attempted to build missing parameter probability distributions for an individual set of observations (i.e., a species' LH parameters).

The first step in our stepwise approach was to build statistical models describing four key relationships connecting LH parameters with one another. These models allowed for an intercept similarly to Froese and Binohlan (2000) and so are not directly comparable to the BH invariants, which are typically calculated as ratios (e.g., M/K versus $M = b_0 + b_1 \cdot K$). However, we did calculate two BH invariants (M/K and L_{mat}/L_{\inf}), as well as the L_{\max}/L_{\inf} ratio, as a first-step comparison of the families in our study, as well as to compare these values with those from previous studies (Table 2). First, it was clear that these ratios varied by taxonomic group, in this case family, as reported in previous studies (Beverton 1992; Prince et al. 2015). The families fell into two broad categories characterized by determinate or indeterminate growth. Individuals from families with mostly determinate growth do not grow as adults and will, on average, reach lengths close to L_{\inf} before dying. Individuals from families with mostly indeterminate growth will continue growing as adults, usually at a decreasing rate, and will not, on average, reach lengths near L_{\inf} . Jacks and

Table 5. Descriptive statistics for the probability distributions of life history parameters and SPR values for the data-rich and data-poor situations.

Parameter	<i>S. rubroviolaceus</i>		<i>N. unicornis</i>		<i>L. harak</i>	
	Data-rich	Data-poor	Data-rich	Data-poor	Data-rich	Data-poor
L_{max} (mm)	—	653 (12)	—	558 (35)	—	312 (3)
L_{inf} (mm)						
Median	574	555	530	528	293	271
SD	22	28	8	40	6	18
SD ratio	—	1.3	—	5	—	3.0
RE (%)	—	-3	—	1	—	-8
K (year ⁻¹)						
Median	0.25	0.28	0.17	0.29	0.31	0.58
SD	0.04	0.11	0.02	0.14	0.02	0.30
SD ratio	—	2.8	—	7	—	15
RE (%)	—	12	—	70	—	87
L_{mat} (mm)						
Median	368	384	361	420	222	218
SD	5	31	10	56	2	26
SD ratio	—	6.2	—	5.6	—	13
RE (%)	—	4	—	16	—	-2
M (year ⁻¹)						
Median	0.14	0.16	0.06	0.10	0.24	0.25
SD	0.02	0.06	0.01	0.02	0.01	0.11
SD ratio	—	3.0	—	2.0	—	11
RE (%)	—	14	—	67	—	4
SPR $F = 0.1M$						
Median	0.92	0.90	0.89	0.93	0.95	0.89
SD	0.01	0.03	0.01	0.02	0.002	0.04
SD ratio	—	3.0	—	2.0	—	20
RE (%)	—	-2	—	4	—	-6
SPR $F = M$						
Median	0.35	0.38	0.38	0.51	0.29	0.36
SD	0.03	0.11	0.03	0.08	0.01	0.13
SD ratio	—	3.7	—	3.0	—	13
RE (%)	—	9	—	34	—	24
SPR $F = 2M$						
Median	0.16	0.18	0.19	0.29	0.11	0.16
SD	0.02	0.08	0.02	0.08	0.01	0.10
SD ratio	—	4.0	—	4.0	—	10
RE (%)	—	13	—	53	—	36

Note: The median L_{max} used to run the models are presented with SD in parentheses. SD ratio is the ratio of the data-poor to data-rich standard deviations, and RE (relative error) is the standardized distance between the data-poor and data-rich medians.

Table 6. Life history parameter estimates from the data-poor approach at various levels of negative bias in L_{max} (percentage reduction in L_{max} compared with “true” value).

Species	Bias in L_{max} (%)	L_{max} (mm)	L_{inf} (mm)	K (year ⁻¹)	M (year ⁻¹)	L_{mat} (mm)	SPR
<i>S. rubroviolaceus</i>	0	653	555	0.28	0.16	384	0.38
	-10	588	499	0.36	0.19	347	0.42
	-20	522	443	0.44	0.22	307	0.49
	-30	457	388	0.53	0.25	268	0.56
<i>N. unicornis</i>	0	558	528	0.29	0.10	420	0.51
	-10	502	472	0.36	0.10	376	0.52
	-20	446	415	0.38	0.10	329	0.54
	-30	391	358	0.42	0.10	284	0.58
<i>L. harak</i>	0	312	271	0.58	0.25	218	0.36
	-10	281	244	0.71	0.30	198	0.37
	-20	250	218	0.77	0.33	179	0.43
	-30	218	190	0.88	0.35	158	0.53

Note: F was set to M for SPR calculations in this table.

goatfishes had relatively high M/K ratios (1.6 versus <1 for other families), which corresponded to an indeterminate growth curve (i.e., high M coupled with slow growth; Prince et al. 2015). This type of growth curve typically leads to maximum observed sizes (L_{max}) being smaller than L_{inf} , as indicated in the mean L_{max}/L_{inf}

ratios for these families, which is <1 (Table 2). The other four families followed the more typical determinate growth patterns, with lower mean M/K ratios (0.3 to 0.8) and L_{max}/L_{inf} ratios > 1.

The distinction between these two type of growth curves was also apparent for the reproductive load values as usually defined (L_{mat}/L_{inf}), with low ratios of ~0.55 for the jacks and goatfishes (i.e., indeterminate growth). These estimates were well below the theoretical value of 0.66 that maximizes reproductive output (Jensen 1996). However, we believe these low reproductive load ratios for families with indeterminate growth curves are actually an artifact of L_{inf} not being representative of the expected size at the oldest age (L_{λ}). In the case of highly indeterminate growth, L_{inf} plays more of a fitting parameter role, rather than having any biological meaning. For example, L_{inf} for the giant trevally (*Caranx ignobilis*) equaled 1840 mm, while L_{λ} was only 1260 mm (as compared with the world-record size of ~1800 mm TL). As a consequence, we found that for families with indeterminate growth, L_{λ} was a better predictor of L_{mat} , and the L_{mat}/L_{λ} ratio was closer to 0.66 (i.e., 0.63). This observation likely explains the previously reported negative relationship between L_{mat}/L_{inf} and M/K ratios (see Prince et al. 2015); as M/K increases there is concomitant decrease in the biological significance of L_{inf} as a representation of the expected length at oldest age (L_{λ}), and thus, the L_{mat}/L_{inf} ratio decreases accordingly. Beverton (1992) showed a similar relation-

ship between $L_{\text{mat}}/L_{\text{inf}}$ and Ka_{max} that is likely related to this artifact as well. Support for this conjecture was apparent in our data set when using L_{inf} but disappeared when we replaced L_{inf} with L_{λ} . For this reason, we selected L_{λ} to infer L_{mat} in our model, and further, we suggest replacing the $L_{\text{mat}}/L_{\text{inf}}$ LH invariant with $L_{\text{mat}}/L_{\lambda}$, when possible.

Although we looked at the BH-LH ratios as a preliminary source of information, our primary focus was to go beyond these relatively simple ratios and explore more flexible statistical models, an approach similar to Froese and Binohlan (2000). In general, these relationships followed similar patterns to those previously published, with linear relationships between $L_{\text{inf}} \sim L_{\text{max}}$ and $L_{\text{mat}} \sim L_{\text{inf}}$ (or L_{λ}) and curvilinear relationships between $K \sim L_{\text{inf}}$ and $M \sim K$, with a few notable exceptions. For example, there was no clear $K \sim L_{\text{inf}}$ and $M \sim K$ relationships for goatfishes, although this may be simply related to the small sample size and narrow L_{inf} range. The $K \sim L_{\text{inf}}$ relationship for jacks was linear, but this is likely due to the lack of smaller species in our data set. The clearest departure from the expected relationship was the flat trend between $M \sim K$ (and L_{max}) for surgeonfishes, despite a wide range of K and L_{max} values. Choat and Robertson (2002) also reported a lack of significant relationship between K or L_{max} and longevity for surgeonfishes (note: M was directly derived from longevity in our study). It is not clear why surgeonfishes seem to so clearly violate the M/K LH invariant. However, a lack of pattern between M and K has also been observed for other fishes, such as walleye (*Sander vitreus*) (Beverton 1987) and brown trout (*Salmo trutta*) (Vøllestad et al. 1993). Despite the lack of a predictor variable for M , the range of longevity (and associated M) values found in these families was consistent and fairly limited; surgeonfishes had the longest maximum age range (20–54 years) and goatfishes had a low longevity range (3–6 years). For these families, taxonomic group was the sole predictor of longevity (and thus M). Consequently, biases in L_{max} had no impact on M estimates for surgeonfishes and goatfishes, as shown in Table 5 for *N. unicornis*.

For the three selected test species, the precision of the data-poor stepwise approach compared reasonably well with the data-rich situation, especially for *S. rubroviolaceus* and *N. unicornis* (i.e., SDs about three times larger on average). However, the data-poor to data-rich SD ratios were significantly larger for *L. harak*, due in part to the much larger sample size used in the *L. harak* study ($n = 409$ versus $n = 180$ for the two other species), which resulted in narrower distributions for the data-rich situation for this species. Furthermore, the emperor (Lethrinidae) family LH relationships were generally less precise than for other species, contributing to wider data-poor parameter distributions and higher mean SD ratios (~11 versus 4 for the other species). In terms of accuracy, the LH parameter and SPR medians were, on average, about 22% off the “true” median values of the data-rich situation, although this varied between species. *Scarus rubroviolaceus* median values were off the true median by about 8% on average compared with ~35% for *N. unicornis*. The median M for *N. unicornis* was 67% larger than the data-rich median, which likely explains the positive bias in SPR median values as well (a higher M leads to increased resilience to high fishing mortality rates and higher SPR values). The larger M estimate is explained by the fact that the growth study used for this species reports the oldest recorded age for any surgeonfish (54 versus 45 years for the next oldest species). The natural mortality of *N. unicornis* is thus slightly overestimated by our simple model for surgeonfishes, which only uses an average M value for this family (since K and L_{max} were independent from M for this family). In general, LH and SPR parameter distributions were surprisingly accurate, with data-poor median values only about 20% off their “true” values, on average.

The stepwise stochastic simulation approach presented here provides relatively accurate and precise LH parameter estimates that allows for first-step stock assessments in data-poor situations where perhaps only size structure information is available. The

probability distributions generated by our approach can also be used as prior information in a Bayesian statistical analysis framework. Although this study presented the models necessary to run this approach for six families, other important families could easily be added through extended meta-analyses (e.g., groupers, rockfishes, etc.). It is important to note that this study presented a way of obtaining SPR when fishing mortality rates have been quantified, which is generally not the case for data-poor situations. However, it is possible to estimate fishing mortality rates through the use of length-based methods (Ehrhardt and Ault 1992; Ault et al. 2005; Gedamke and Hoenig 2006). We plan on extending and integrating these concepts in future studies.

Acknowledgements

This study was a collaborative effort between the NOAA Pacific Island Fisheries Science Center (PIFSC), Fisheries Research and Monitoring Division (Integrated Reef Ecosystem Evaluation Framework iREEFS, Contract No. WE-133F-12-SE-2099) and the University of Miami (NOAA Fisheries Coral Reef Conservation Program Grant No. NA17RJ1226 through the NOAA Southeast Fisheries Science Center). Funding for reef fish surveys was provided by the NOAA PIFSC Coral Reef Ecosystem Program as part of the Pacific Reef Assessment and Monitoring Program. This manuscript was greatly improved by the comments of S.G. Smith, E.A. Babcock, N.M. Ehrhardt, J.A. Bohnsack, G.T. DiNardo, J. O'Malley, F. Carvalho, A. Punt, and several anonymous reviewers.

References

- Alagaraja, K. 1984. Simple methods for estimation of parameters for assessing exploited fish stocks. *Ind. J. Fish.* 31(2): 177–208.
- Ault, J.S., Bohnsack, J.A., and Meester, G.A. 1998. A retrospective (1979–1996) multispecies assessment of coral reef fish stocks in the Florida Keys. *Fish. Bull.* 96(3): 395–414.
- Ault, J.S., Smith, S.G., and Bohnsack, J.A. 2005. Evaluation of average length as an estimator of exploitation status for the Florida coral-reef fish community. *ICES J. Mar. Sci.* 62(3): 417–423. doi:10.1016/j.icesjms.2004.12.001.
- Ault, J.S., Smith, S.G., Luo, J., Monaco, M.E., and Appeldoorn, R.S. 2008. Length-based assessment of sustainability benchmarks for coral reef fishes in Puerto Rico. *Environ. Conserv.* 35(3): 221–231. doi:10.1017/S0376892908005043.
- Ault, J.S., Smith, S.G., Browder, J.A., Nuttle, W., Franklin, E.C., Luo, J., DiNardo, G.T., and Bohnsack, J.A. 2014. Indicators for assessing the ecological dynamics and sustainability of southern Florida's coral reef and coastal fisheries. *Ecol. Indic.* 44: 164–172. doi:10.1016/j.ecolind.2014.04.013.
- Ayvazian, S., Chatfield, B., Gaughan, D., Keay, I., and Nowarda, G. 2004. The age, growth, reproductive biology and stock assessment of grass emperor, *Lethrinus laticaudis* in Shark Bay, Western Australia. Western Australian Marine Research Laboratories, North Beach, Western Australia.
- Beverton, R.J.H. 1963. Maturation, growth and mortality of clupeid and engraulid stocks in relation to fishing. *Rapp. P.-v. Réun. Cons. Perm. Int. Explor. Mer.* 154: 44–67.
- Beverton, R.J.H. 1987. Longevity in fish: some ecological and evolutionary considerations. In *Evolution of longevity in animals*. Edited by A.D. Woodhead and K.H. Thompson. Springer, USA. pp. 161–185.
- Beverton, R.J.H. 1992. Patterns of reproductive strategy parameters in some marine teleost fishes. *J. Fish Biol.* 41: 137–160. doi:10.1111/j.1095-8649.1992.tb03875.x.
- Beverton, R.J.H., and Holt, S.J. 1959. A review of the lifespans and mortality rates of fish in nature, and their relation to growth and other physiological characteristics. In *Ciba Foundation Symposium — The Lifespan of Animals (Colloquia on Ageing)*. Edited by G.E.W. Wolstenholme and B.A. O'Connor. John Wiley & Sons, Ltd. pp. 142–180.
- Bolker, B. 2008. *Ecological models and data* in R. Princeton University Press, Princeton, New Jersey.
- Charnov, E.L. 1993. Life history invariants. Oxford University Press, Oxford.
- Charnov, E.L. 2008. Fish growth: Bertalanffy k is proportional to reproductive effort. *Environ. Biol. Fishes.* 83(2): 185–187. doi:10.1007/s10641-007-9315-5.
- Choat, J.H., and Robertson, D.R. 2002. Age-based studies on coral reef fishes. In *Coral reef fishes: dynamics and diversity in a complex ecosystem*. Academic Press, San Diego, Calif. pp. 57–80.
- Cole, K.S. 2009. Size-dependent and age-based female fecundity and reproductive output for three Hawaiian goatfish (Family Mullidae) species, *Mulloidichthys flavolineatus* (yellowstripe goatfish), *M. vanicolensis* (yellowfin goatfish), and *Parupeneus porphyreus* (whitesaddle goatfish). Report to the Division of Aquatic Resources Dingell-Johnson Sport Fish Restoration.
- Eble, J.A., Langston, R., and Bowen, B.W. 2009. Growth and reproduction of Hawaiian Kala, *Naso unicornis*. Division of Aquatic Resources report, Honolulu.
- Ehrhardt, N.M., and Ault, J.S. 1992. Analysis of two length-based mortality mod-

- els applied to bounded catch length frequencies. *Trans. Am. Fish. Soc.* **121**(1): 115–122. doi:10.1577/1548-8659(1992)121<0115:AOTLMM>2.3.CO;2.
- Ellington, E.H., Bastille-Rousseau, G., Austin, C., Landolt, K.N., Pond, B.A., Rees, E.E., Robar, N., and Murray, D.L. 2015. Using multiple imputation to estimate missing data in meta-regression. *Methods Ecol. Evol.* **6**(2): 153–163. doi:10.1111/2041-210X.12322.
- Fenner, D. 2012. Challenges for managing fisheries on diverse coral reefs. *Diversity*, **4**(4): 105–160. doi:10.3390/d4010105.
- Fox, J. 2008. *Applied regression analysis and generalized linear models*. Sage Publications.
- Froese, R., and Binohlan, C. 2000. Empirical relationships to estimate asymptotic length, length at first maturity and length at maximum yield per recruit in fishes, with a simple method to evaluate length frequency data. *J. Fish Biol.* **56**(4): 758–773. doi:10.1111/j.1095-8649.2000.tb00870.x.
- Froese, R., and Binohlan, C. 2003. Simple methods to obtain preliminary growth estimates for fishes. *J. Appl. Ichthyol.* **19**(6): 376–379. doi:10.1111/j.1439-0426.2003.00490.x.
- Gedamke, T., and Hoenig, J.M. 2006. Estimating mortality from mean length data in nonequilibrium situations, with application to the assessment of goosfish. *Trans. Am. Fish. Soc.* **135**(2): 476–487. doi:10.1577/T05-153.1.
- Grandcourt, E.M. 2002. Demographic characteristics of a selection of exploited reef fish from the Seychelles: preliminary study. *Mar. Freshw. Res.* **53**(2): 123–130. doi:10.1071/MF01123.
- Helser, T.E., Stewart, I.J., and Lai, H.-L. 2007. A Bayesian hierarchical meta-analysis of growth for the genus *Sebastes* in the eastern Pacific Ocean. *Can. J. Fish. Aquat. Sci.* **64**(3): 470–485. doi:10.1139/f07-024.
- Hewitt, D.A., and Hoenig, J.M. 2005. Comparison of two approaches for estimating natural mortality based on longevity. *Fish. Bull.* **103**(2): 433–437.
- Hoenig, J.M. 1983. Empirical use of longevity data to estimate mortality rates. *Fish. Bull.* **82**(1): 898–903.
- Hordyk, A., Ono, K., Valencia, S., Loneragan, N., and Prince, J. 2015. A novel length-based empirical estimation method of spawning potential ratio (SPR), and tests of its performance, for small-scale, data-poor fisheries. *ICES J. Mar. Sci.* **72**(1): 217–231. doi:10.1093/icesjms/fsu004.
- Howard, K.G. 2008. Community structure, life history, and movement patterns of parrotfishes: large protogynous fishery species. Ph.D., University of Hawaii at Manoa, Honolulu.
- Jarić, I., and Gačić, Z. 2012. Relationship between the longevity and the age at maturity in long-lived fish: Rikhter/Efanov's and Hoenig's methods. *Fish. Res.* **129–130**: 61–63. doi:10.1016/j.fishres.2012.06.010.
- Jensen, A.L. 1996. Beverton and Holt life history invariants result from optimal trade-off of reproduction and survival. *Can. J. Fish. Aquat. Sci.* **53**(4): 820–822. doi:10.1139/f95-233.
- Jensen, A.L. 1997. Origin of relation between K and L_{inf} and synthesis of relations among life history parameters. *Can. J. Fish. Aquat. Sci.* **54**(5): 987–989. doi:10.1139/f97-007.
- Jobling, M. 1994. *Fish bioenergetics*. Chapman and Hall, London, UK.
- Kenchington, T.J. 2014. Natural mortality estimators for information-limited fisheries. *Fish. Fish.* **15**(4): 533–562. doi:10.1111/faf.12027.
- Kulbicki, M., Guillemot, N., and Amand, M. 2005. A general approach to length-weight relationships for New Caledonian lagoon fishes. *Cybium*, **29**(3): 235–252.
- Longenecker, K., Langston, R., and Eble, J.A. 2008. Growth, mortality, and reproduction of manini *Acanthurus triostegus sandvicensis*. Contribution No. 2008-006 to the Hawaii Biological Survey, Honolulu, Hawaii.
- Nadon, M.O., Ault, J.S., Williams, I.D., Smith, S.G., and DiNardo, G.T. 2015. Length-based assessment of coral reef fish populations in the Main and Northwestern Hawaiian Islands. *PLoS ONE*, **10**(8): e0133960. doi:10.1371/journal.pone.0133960. PMID:26267473.
- Pauly, D. 1980. On the interrelationships between natural mortality, growth parameters, and mean environmental temperature in 175 fish stocks. *ICES J. Mar. Sci.* **39**(2): 175–192. doi:10.1093/icesjms/39.2.175.
- Pauly, D., Watson, R., and Alder, J. 2005. Global trends in world fisheries: impacts on marine ecosystems and food security. *Philos. Trans. R. Soc. B Biol. Sci.* **360**(1453): 5–12. doi:10.1098/rstb.2004.1574.
- Prince, J., Hordyk, A., Valencia, S.R., Loneragan, N., and Sainsbury, K. 2015. Revisiting the concept of Beverton–Holt life-history invariants with the aim of informing data-poor fisheries assessment. *ICES J. Mar. Sci.* **72**(1): 194–203. doi:10.1093/icesjms/fsu011.
- Pulkkinen, H., Mäntyniemi, S., Kuikka, S., and Levontin, P. 2011. More knowledge with the same amount of data: advantage of accounting for parameter correlations in hierarchical meta-analyses. *Mar. Ecol. Prog. Ser.* **443**: 29–37. doi:10.3354/meps09368.
- Raghunathan, T., Lepkowski, J.M., Van Hoewyk, J., and Solenberger, P. 2001. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Surv. Methodol.* **27**(1): 85–95.
- Rubin, D.B. 1987. *Multiple imputation for nonresponse in surveys*. John Wiley & Sons, Inc., New York.
- Taylor, B.M., and Choat, J.H. 2014. Comparative demography of commercially important parrotfish species from Micronesia. *J. Fish Biol.* **84**(2): 383–402. doi:10.1111/jfb.12294. PMID:24446696.
- Taylor, B., and McIlwain, J. 2010. Beyond abundance and biomass: effects of marine protected areas on the demography of a highly exploited reef fish. *Mar. Ecol. Prog. Ser.* **411**: 243–258. doi:10.3354/meps08672.
- Then, A.Y., Hoenig, J.M., Hall, N.G., and Hewitt, D.A. 2015. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. *ICES J. Mar. Sci.* **72**(1): 82–92. doi:10.1093/icesjms/fsu136.
- Thorson, J.T., Taylor, I.G., Stewart, I.J., and Punt, A.E. 2014. Rigorous meta-analysis of life history correlations by simultaneously analyzing multiple population dynamics models. *Ecol. Appl.* **24**(2): 315–326. doi:10.1890/12-1803.1. PMID:24689143.
- Thorson, J.T., Cope, J.M., Kleisner, K.M., Samhoury, J.F., Shelton, A.O., and Ward, E.J. 2015. Giants' shoulders 15 years later: lessons, challenges and guidelines in fisheries meta-analysis. *Fish. Fish.* **16**(2): 342–361. doi:10.1111/faf.12061.
- Van Buuren, S. 2007. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat. Methods Med. Res.* **16**(3): 219–242. doi:10.1177/0962280206074463. PMID:17621469.
- Vøllestad, L.A., L'Abée-Lund, J.H., and Sæggrov, H. 1993. Dimensionless numbers and life history variation in Brown Trout. *Evol. Ecol.* **7**(2): 207–218. doi:10.1007/BF01239389.
- White, D.B., and Palmer, S.M. 2004. Age, growth, and reproduction of the red snapper, *Lutjanus campechanus*, from the Atlantic waters of the southeastern US. *Bull. Mar. Sci.* **75**(3): 335–360.
- Williams, I.D., Richards, B.L., Sandin, S.A., Baum, J.K., Schroeder, R.E., Nadon, M.O., Zgliczynski, B., Craig, P., McIlwain, J.L., and Brainard, R.E. 2011. Differences in reef fish assemblages between populated and remote reefs spanning multiple archipelagos across the central and western Pacific. *J. Mar. Biol.* **2011**: 1–14. doi:10.1155/2011/826234.