NRC Research Press

ARTICLE

Lessons learned from data-limited evaluations of data-rich reef fish species in the Gulf of Mexico: implications for providing fisheries management advice for data-poor stocks[1]

Skyler R. Sagarese, William J. Harford, John F. Walter, Meaghan D. Bryan, J. Jeffery Isely, Matthew W. Smith, Daniel R. Goethel, Adyan B. Rios, Shannon L. Cass-Calay, Clay E. Porch, Thomas R. Carruthers, and Nancie J. Cummings

Abstract: Specifying annual catch limits for artisanal fisheries, low economic value stocks, or bycatch species is problematic due to data limitations. Many empirical management procedures (MPs) have been developed that provide catch advice based on achieving a stable catch or a historical target (i.e., instead of maximum sustainable yield). However, a thorough comparison of derived yield streams between empirical MPs and stock assessment models has not been explored. We first evaluate trade-offs in conservation and yield metrics for data-limited approaches through management strategy evaluation (MSE) of seven data-rich reef fish species in the Gulf of Mexico. We then apply data-limited approaches for each species and compare how catch advice differs from current age-based assessment models. MSEs identified empirical MPs (e.g., using relative abundance) as a compromise between data requirements and the ability to consistently achieve management objectives (e.g., prevent overfishing). Catch advice differed greatly among data-limited approaches and current assessments, likely due to data inputs and assumptions. Adaptive MPs become clearly viable options that can achieve management objectives while incorporating auxiliary data beyond catch-only approaches.

Résumé : L'établissement de limites de prises annuelles pour les pêches artisanales, les stocks de faible valeur économique ou les espèces de prises accessoires pose problème en raison des données limitées sur lesquelles il repose. De nombreuses procédures de gestion (PG) empiriques ont été élaborées qui fournissent des avis sur les prises basés sur l'atteinte de prises stables ou d'une cible historique (c.-à-d. plutôt que le rendement équilibré maximal). Une comparaison exhaustive des différents rendements obtenus de PG empiriques et de modèles d'évaluation de stock n'a toutefois pas été effectuée. Nous évaluons d'abord les compromis touchant à la conservation et aux paramètres du rendement pour des approches pour données limitées par l'évaluation des stratégies de gestion (ESG) de sept espèces de poissons récifaux dans le golfe du Mexique pour lesquelles les données sont abondantes. Nous appliquons ensuite, pour chaque espèce, des approches pour des situations de données limitées et comparons les avis sur les prises en découlant à ceux issus de modèles d'évaluation actuels basés sur l'âge. Les ESG établissent que les PG empiriques (p.ex. qui utilisent l'abondance relative) constituent un compromis entre les exigences en matière de données et la capacité d'atteindre régulièrement les objectifs de gestion (p.ex. la prévention de la surpêche). Les avis sur les prises diffèrent considérablement entre les approches à données limitées et les évaluations actuelles, vraisemblablement en raison des données entrées et des hypothèses sous-jacentes. Les PG adaptatives deviennent des options manifestement viables qui peuvent permettre l'atteinte des objectifs de gestion tout en incorporant des données auxiliaires autres que celles qui entrent dans les approches basées uniquement sur les prises. [Traduit par la Rédaction]

Introduction

Nearly 80% of global catch comes from fisheries stocks characterized as lacking formal stock assessments, with less than 1% of species assessed using quantitative approaches due to costs and intensive data requirements (Costello et al. 2012). Unassessed species are often deemed "data-poor", meaning available data are insufficient to support the use of traditional quantitative methods of assessing the status of the stock relative to the levels that will produce optimum yield and its associated biological reference points (Geromont and Butterworth 2015*a*). Conventional fisheries stock assessments estimate these quantities by use of statistical models that integrate information on catch, relative abundance,

---

Received 30 October 2017. Accepted 24 June 2018.

S.R. Sagarese, J.F. Walter, M.D. Bryan,* J.J. Isely, M.W. Smith, D.R. Goethel,† A.B. Rios, S.L. Cass-Calay, C.E. Porch, and N.J. Cummings. Southeast Fisheries Science Center, Sustainable Fisheries Division, 75 Virginia Beach Drive, Miami, FL 33149, USA.
W.J. Harford. Southeast Fisheries Science Center, Sustainable Fisheries Division, 75 Virginia Beach Drive, Miami, FL 33149, USA; Cooperative Institute for Marine and Atmospheric Studies, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, FL 33149, USA.
T.R. Carruthers. Institute for the Oceans and Fisheries, The University of British Columbia, 2202 West Mall, Vancouver, BC V6T 1Z4, Canada.
Corresponding author: Skyler R. Sagarese (email: skyler.sagarese@noaa.gov).
*Present address: Alaska Fisheries Science Center, 7600 Sand Point Way N.E., Seattle, WA 98115, USA.

†Daniel R. Goethel currently serves as an Associate Editor; peer review and editorial decisions regarding this manuscript were handled by Melissa Karp.

[1] This article is being published as part of the special issue "Under pressure: addressing fisheries challenges with Management Strategy Evaluation" arising from two related theme sessions sponsored by the American Institute of Fishery Research Biologists at the 147th Annual Meeting of the American Fisheries Society, Tampa, Florida, USA, August 2017.

Can. J. Fish. Aquat. Sci. **76**: 1624–1639 (2019) dx.doi.org/10.1139/cjfas-2017-0482

size or age composition, and biology. However, these models demand more data and analytical support than are available for many stocks. Alternative approaches that cope with data limitations are therefore needed to support determination of catch limits. Data-limited approaches can serve as interim solutions while data collection improves, at which time full stock assessments can be implemented. However, in certain instances, data-limited approaches may be required as long-term solutions where data-poor circumstances are likely to persist (Harford and Carruthers 2017).

Numerous data-limited methods (DLMs) and management procedures (MPs, also known as harvest strategies; Carruthers et al. 2016) have been developed over the last few decades, which provide catch advice utilizing data ranging from catch-only to catch plus auxiliary data (e.g., catch per unit effort, CPUE). Two basic types of data-limited approaches exist: (*i*) empirical MPs that aim to achieve a stable catch or a historical target but not necessarily maximize yield (i.e., not seeking maximum sustainable yield (MSY)); and (*ii*) DLMs that attempt to achieve MSY (i.e., MSY-seeking). Empirical MPs use data streams of recent catch, CPUE, or mean length (Geromont and Butterworth 2015*a*). In the absence of an estimate of MSY and B_{MSY} (the biomass at MSY), and therefore lacking the ability to move the stock towards B_{MSY}, these approaches can be used to avoid further stock declines. As additional information becomes available (e.g., vital rates or age composition data and estimates of age-based selectivity), DLMs (e.g., yield-per-recruit analyses; Beverton and Holt 1957) or age-structured assessment models (e.g., Stock Synthesis (SS); Methot and Wetzel 2013) can be applied to determine optimal reference points such as MSY or MSY proxies.

The Magnuson–Stevens Reauthorization Act in the US requires federal fishery management plans to prescribe annual catch limits designed to prevent overfishing while achieving the optimum long-term yield for each stock (MSFCMA 2006). As a result, there is a need to pursue data-limited approaches for stocks not previously assessed. In the United States, nearly 60% of stocks are considered data-poor (Newman et al. 2015), but the percentage varies by region and is typically higher (>75%) for biodiverse areas in the southeastern US (Berkson and Thorson 2015; Newman et al. 2015). The high percentage of data-poor stocks in the southeastern US and US Caribbean persists despite the high economic impact of its commercial and recreational fisheries (NOAA 2013). Lacking a quantitative assessment due to data limitations, annual catch limits for many data-poor stocks have been based on catch-only techniques. While data collection improvements for data-limited stocks are being made (e.g., Bryan et al. 2016), it is likely that the need for data-limited approaches will continue into the future in areas such as the US Caribbean.

Despite DLMs and MPs being widely applied, there has been relatively limited comparison of resulting catch advice with that from model-derived assessment-management frameworks. Depletion-based stock reduction analysis (DBSRA) was relatively effective for estimating sustainable yields when compared with data-rich stock assessments (Dick and MacCall 2011). Performance of simple MPs was comparable to data-rich assessments, even for stocks that exhibited retrospective patterns (Geromont and Butterworth 2015*b*). Understanding how catch advice may differ among approaches is particularly important as data collection programs mature and stocks begin to move out of the data-poor realm and become data-moderate. Here we examine the performance of a suite of DLMs and empirical MPs for seven reef fish species reflecting varying life histories and fishery characteristics in the Gulf of Mexico. The objectives of this study are to (*i*) determine through management strategy evaluation (MSE) which approaches are able to adhere to management objectives (e.g., preventing overfishing) and at what economic trade-off (e.g., yield reductions); (*ii*) identify commonalities in DLM or MP performance across species groups; and (*iii*) compare data-limited catch advice with forecasted catch advice from existing data-rich stock assessment models to address whether similar catch advice could have been achieved with less data, with fewer data sources, or with computationally less-intensive methods. The results of this study provide one of the first demonstrations of how yield streams compare between data-limited approaches and age-structured population models and indicate which DLMs and MPs appear most robust in respect to Gulf of Mexico reef fish species.

Materials and methods

Modeling approach

The Data-Limited Methods Toolkit (DLMtool, version 3.2.2; http://www.datalimitedtoolkit.org/; Newman et al. 2014; Carruthers and Hordyk 2016) package in R (R Core Team 2016) was developed to broaden the accessibility of various DLMs and MPs (refer to online Supplemental material, Table S1[2]) and facilitate the evaluation of their efficacy using MSE (Punt et al. 2016). A step-by-step approach has been recommended by the DLMtool developers for implementation: (*i*) determination of feasible methods based on data availability; (*ii*) simulation testing of feasible methods (through MSE) to eliminate methods that exhibit pathological behavior (e.g., chronic overfishing) and identify viable methods; and (*iii*) application of viable methods to produce management advice (SEDAR 2016*b*; Hordyk et al. 2017).

Species and available data

This study focuses on seven economically important species that have accepted stock assessments (using SS) for providing management advice. These species span four families exhibiting different life histories or fishery characteristics: greater amberjack (*Seriola dumerili*), gag grouper (*Mycteroperca microlepis*), red grouper (*Epinephelus morio*), yellowedge grouper (*Hyporthodus flavolimbatus*), red snapper (*Lutjanus campechanus*), vermilion snapper (*Rhomboplites aurorubens*), and gray triggerfish (*Balistes capriscus*).

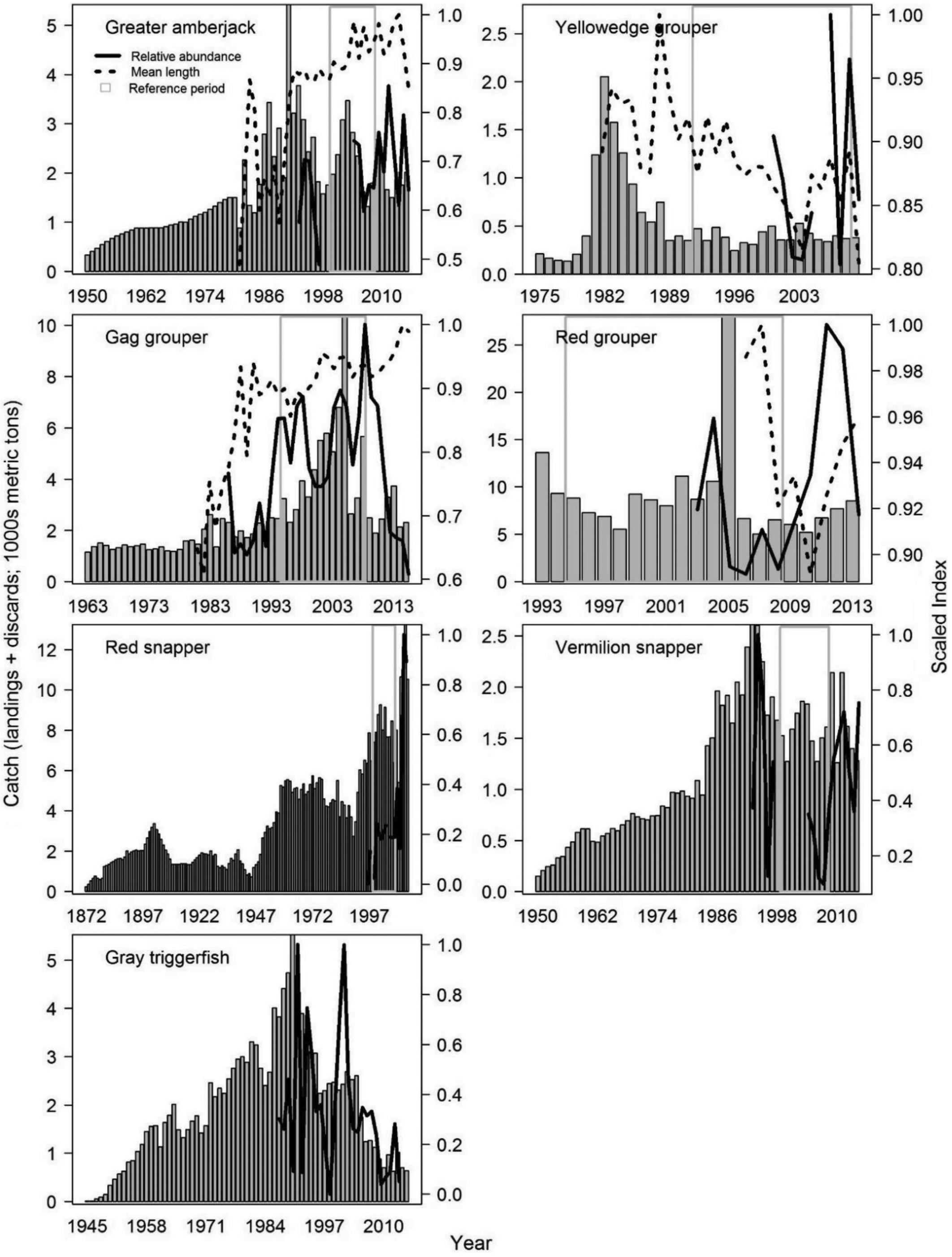
Data inputs for the MSE and for developing catch advice were extracted from the most recent peer-reviewed Southeast Data Assessment and Review (SEDAR) assessment documents for each species (sedarweb.org). For the MSE, a representative fishing fleet was required for each stock to simulate historical fishing effort and mimic the exploitation history (Carruthers et al. 2016). The fishing fleet responsible for the highest proportion of total removals (landings and dead discards) was deemed most representative. In addition, an estimate of current depletion was obtained from each stock assessment model. Specifics on sources and data inputs required for the MSE are provided in Tables S2–S3[2].

For developing catch advice using available data, the index of relative abundance selected for analysis was based on considerations of source (i.e., fishery-independent chosen over fishery-dependent sources), extent of spatial coverage with respect to the spatial distribution of the stock, and sampling effort or catch of the target species in multispecies fisheries. Where available, annual mean length was extracted for each fishing fleet or fishery-independent data source excluding recruitment data sources. Time series of catch and relative abundance were available for all species examined (Fig. 1), whereas mean length was not reported for red snapper, vermilion snapper, or gray triggerfish within their respective assessments. Details on sources and data inputs for determining catch advice are provided in Tables S4–S5[2].

---

[2] Supplementary data are available with the article through the journal Web site at http://nrcresearchpress.com/doi/suppl/10.1139/cjfas-2017-0482.

Fig. 1. Available data, including total removals (bars), an index of abundance (considered most representative), and (or) index of mean length (derived from the predominant fishery) for reef fish species in the Gulf of Mexico. The predominant fishery and representative index of abundance are provided in Tables 3 and S5[2], respectively.



Data-limited and data-rich methods

Various DLMs and MPs as available in DLMtool were examined (Tables 1, S1[2]) and are discussed briefly. Simple catch-only methods (e.g., constant catch; Geromont and Butterworth 2015*a*) yield catch advice based on the mean catch derived from a specified time period (e.g., recent or reference). Empirical MPs aim to maintain a stable catch or historical target and do not provide information on MSY (Geromont and Butterworth 2015*a*). For example, index-based and length-based MPs assume changes in relative abundance or mean length are direct and indirect indicators of stock abundance, respectively (Geromont and Butterworth 2015*a*; Carruthers et al. 2016). Depletion-corrected average catch (DCAC) adjusts historical catches using assumptions about life history characteristics (MacCall 2009) and an expert-informed estimate of depletion to provide a sustainable catch.

DLMs produce reference points such as MSY or indicators of stock status such as B_{MSY} or spawner potential ratio proxies (Brooks et al. 2008). These approaches require data such as relative abundance, current stock depletion, or catch-at-age (Table 2). DLMs tested included delay-difference (DD; Deriso 1980; Schnute

Table 1. Description of methods applied and model outputs.

Type	Method	Description	Output	Reference
Catch-only MPs	CC1	Recent mean catch (last 5 years)	Catch limit	Geromont and Butterworth 2014
	CC_Ref	Reference mean catch (reference period)	Catch limit	
Index-based MPs	Islope (5 years or 10 years)	CPUE slope (adjust catch advice based on slope in CPUE for last 5 or 10 years)	Catch limit	
	Itarget	CPUE target (adjust catch advice to achieve a target CPUE, where target = 1.5 × mean CPUE during reference period)	Catch limit	
Length-based MPs	LstepCC	Stepwise constant catch using mean length (catch adjusted based on ratio of recent to reference mean length)	Catch limit	
	Ltarget	Length target (adjust catch advice to achieve a target mean length, where target = 1.05 × mean length during reference period)	Catch limit	
Depletion-based DLMs	DCAC	Depletion-corrected average catch	Sustainable yield	MacCall 2009; Carruthers et al. 2014
	DBSRA	Depletion-based stock reduction analysis	Carrying capacity, B_{MSY}, F_{MSY}, MSY	Dick and MacCall 2011; Carruthers et al. 2014
Data-moderate	DD	Delay-difference	Depletion, MSY	C. Walters (in Carruthers et al. 2014)
Age-based DLMs	BK_CC	Beddington–Kirkwood life history approach, uses a catch curve to estimate current abundance from catches and recent F	F_{MSY} proxy, catch limit	Beddington and Kirkwood 2005
	YPR_CC	Yield-per-recruit analysis, catch curve as above	F_{MSY} proxy, catch limit	M. Bryan (in Carruthers and Hordyk 2016)
	Fdem_CC	Demographic F_{MSY} approach, catch curve as above	F_{MSY}, catch limit	McAllister et al. 2001
	Fratio_CC	Fixed F_{MSY}/M ratio method, catch curve as above	F_{MSY}, catch limit	Gulland 1971; Walters and Martell 2002; Martell and Froese 2013
Integrated analysis	SS	Stock Synthesis statistical age-structured population model	MSY, B_{MSY}, $\mathrm{SPR}_{\mathrm{MSY}}$, F_{MSY}, SSB, and SPR target reference points	Methot and Wetzel 2013; Methot 2012

Note: Reference point outputs include maximum sustainable yield (MSY), spawning biomass that produces MSY (B_{MSY}), exploitation rate corresponding to MSY (F_{MSY}), spawning stock biomass (SSB), and spawner potential ratio (SPR). MP, management procedure; DLM, data-limited method; CPUE, catch per unit effort. Equations and assumptions are provided in Table S1[2].

Table 2. Data requirements (×) for data-limited methods and management procedures (MPs) applied.

	Catch-only		Empirical MPs				Depletion-based		Data-moderate	Age-based			
Data inputs	CC1	CC_Ref	Islope	Itarget	LstepCC	Ltarget	DCAC	DBSRA	DD	Fratio_CC	BK_CC	YPR_CC	Fdem_CC
Natural mortality (Mort) fixed							×	×	×	×	×	×	×
von Bertalanffy growth fixed									×		×	×	×
Maximum age (MaxAge)												×	×
Steepness (steep) fixed													×
Length at 50% maturity (L50)								×	×				×
Weight–length parameter *a* (wla)									×			×	×
Weight–length parameter *b* (wlb)									×			×	×
Total removals (Cat)	×	×	×	×	×	×		×	×	×	×	×	×
Average catch (AvC)							×						
Length at first capture (LFC)											×		
Length at full selection (LFS)												×	
Catch-at-age (CAA)										×	×	×	×
Mean length time series (ML)					×	×							
Ratio of F_{MSY} to Mort (F_{MSY}/M)							×	×		×			
Ratio of B_{MSY} to virgin biomass (B_{MSY}/B_0)							×	×					
Index of abundance (Ind)			×	×					×				
Depletion over specified time period *t* (Dt)							×						
Stock depletion in terminal year (Dep)								×					

1985; Carruthers et al. 2014); DBSRA, which adjusts the catch history with knowledge about life history and expert-informed estimates of depletion (Dick and MacCall 2011); and age-based (i.e., catch curve) DLMs that estimate current abundance based on catches and a recent exploitation rate (F) estimate from age composition (Tables 1–2).

Each of the seven species has an existing tailored-made, peer-reviewed SS assessment model, which was accepted as the best

Table 3. Summary of assessment model structure for reef fish species in the Gulf of Mexico.

Model dimensions	Red snapper	Vermilion snapper	Red grouper	Gag grouper	Yellowedge grouper	Greater amberjack	Gray triggerfish
Source	SEDAR31 update	SEDAR45	SEDAR42	SEDAR33 update	SEDAR22	SEDAR33 update	SEDAR43
Start year (unfished)	1872 (Yes)	1950 (Yes)	1993 (No)	1963 (No)	1975 (Yes)	1950 (No)	1945 (Yes)
Terminal year	2014	2014	2013	2015	2009	2015	2015
Life history							
Natural mortality fixed	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Growth fixed	Yes	Yes	Yes	No	No	Yes	Yes
Steepness fixed	Yes	No	Yes	Yes	No	Yes	No
Fishery							
Largest annual catch (metric tons)	13 233	2 611	28 215	10 425	2 053	5 422	5 528
No. of fleets	14	4	5	6	4	4	5
Predominant fleet (landings plus discards; percentage over last 10 years)	Commercial handline west (24%)	Commercial handline east (48%)	Commercial longline (44%)	Recreational (MRFSS private) (56%)	Commercial longline east (66%)	Recreational (MRFSS) (68%)	Recreational east (85%)
Composition							
Annual length comp. observations	0	20	55	207	209	195	0
Annual age comp. observations	467	55	84	124	1 656	74	147
Abundance							
No. of fishery-independent indices	10	3	3	3	2	2	3
No. of fishery-dependent indices	8	8	4	5	2	4	6
Depletion level in terminal year	0.15	0.31	0.36	0.39	0.31	0.09	0.19
Projections							
Reference point	SPR26%	SPR30%	SPR30%	SPR30%	SPR30%	SPR30%	SPR30%

Note: MRFSS, Marine Recreational Fisheries Statistics Survey; SPR, spawner potential ratio.

available science through the National Oceanic and Atmospheric Administration (NOAA), Center for Independent Experts peer review program process (http://www.st.nmfs.noaa.gov/science-quality-assurance/). SS (Methot 2012) is a biological and statistical framework used in more than 60 fishery stock assessments worldwide (Methot and Wetzel 2013) spanning data-poor (e.g., Cope 2013) to data-rich applications (Methot and Wetzel 2013). The SS framework consists of four submodels: (1) a population submodel; (2) an observational submodel that calculates predicted values for various observed data sources; (3) a statistical submodel that quantifies the goodness of fit by calibrating predictions against observations; and (4) a forecast submodel that projects future catch and abundance based on a user-specified target reference point (Methot and Wetzel 2013). Each assessment differs in the number of data sources, fisheries modeled, time-series length, and species-specific biology (Table 3).

MSE of data-limited methods

MSE is the process of using simulation testing to examine the performance of candidate management strategies under uncertainty (Butterworth et al. 2010; Punt et al. 2016). In an MSE, the robustness of a management system is tested by simulating the system dynamics via an underlying operating model and testing the ability of a given management strategy to achieve management goals utilizing "observed" data (i.e., biological sampling) that informs a scientific analysis (e.g., stock assessment), the outputs of which resulting management measures are based on (Sainsbury et al. 2000; Kell et al. 2007; Punt et al. 2016). Species-specific MSEs were conducted within the DLMtool using applicable data-limited approaches based on data availability. For comparison with a best-case scenario, the true simulated F_{MSY} was multiplied by abundance to derive the true catch advice (reference F_{MSY}). A brief overview of MSE is below, but specifics can be found in Harford et al. (2016*b*), Carruthers et al. (2014), and Carruthers et al. (2016) and associated supplementary material. In a traditional sense, our applications are not complete MSEs because they generally require multiple years of analyses and feedback. However, our framework did incorporate feedback from stakeholders on objectives and performance metrics obtained during the Gulf of Mexico data-limited assessment (discussed below).

Operating model

An operating model was developed to mimic the life history, stock dynamics, and fleet characteristics (e.g., effort, selectivity) for each species. These operating models are assumed to comprise reasonable ranges of parameters for each species; realistically represent the biological components of the system to be managed; and portray realistic fisher behavior in response to management actions (Kell et al. 2007; Maunder 2014). Specific data requirements along with justifications are documented in Tables S2–S3[2]. Notably, current depletion in the simulation was defined as the ratio of current spawning biomass to unfished spawning biomass (Carruthers et al. 2014) and was estimated within SS to be below 0.4 (i.e., overexploited) for the majority of stocks (Table 3).

Simulated stock dynamics

Stock dynamics in the DLMtool MSE framework are age-structured for ages 1 (recruitment age) to a maximum age at which less than 1% of the cohort survives. Natural mortality is age- and sex-invariant, spawning stock biomass consists of both sexes combined, and maturity is a logistic function of age. The DLMtool simulates a single fishery operating on a fish stock. Vulnerability-at-age is calculated through a user-specified double-normal distribution, which offers the flexibility to simulate a variety of patterns including dome-shaped or asymptotic vulnerability. Data inputs are simulated through an observation submodel that introduces imperfect information into the specification of harvest decisions. In proceeding from "true" simulated values to values used

Table 4. Levels of bias (units as coefficients of variation) defining the accuracy and precision of "observed" data or expert knowledge used to implement the various data-limited approaches for each species-specific MSE.

MSE attribute	Value	Source
Observation Error		
Recruitment	0.1–0.3	
Catch	0.1–0.5	Carruthers et al. 2014
Current F	0.15–1.2	Range of current F (Table S3[2])
Absolute biomass	0.2–0.5	
Relative abundance index	0.29–2.28*	Table S3[2]
Depletion (Dep)	0.05–0.83*	Table S3[2]
Lognormal variability in length-at-age	0.10–0.67*	Table S3[2]
Parameter bias		
Catch, L50, Mort, vbLinf, vbK, vbt0, LFC, LFS	0.05	
B_{MSY}/B_0, steepness, F_{MSY}	0.1	
Intrinsic rate of increase, Dep, current F	0.2	
F_{MSY}/M	0.25	
Unfished biomass	0.5	
Absolute biomass	0.33–3.0	
Annual length or age observations	100–200	Desired range
Effective sample size	10–20	

Note: No source indicates default values. Parameters are as defined in Table 2, with ranges representing the lower and upper bounds of a uniform random variable.

*Species-specific values are provided in Table S3[2].

by DLMs and MPs, error is introduced to reflect user-specified levels of imperfect knowledge (Table 4). For the purposes of this analysis, imperfect knowledge was introduced in the form of imprecision, which refers to random interannual variation in observable quantities around respective "true" simulated values. Bias, referring to inaccuracy in a given quantity that occurs for the duration of a simulation, was considered minimal in the present study to reduce its influence on the results.

In DLMtool version 3.2.2, inputs for von Bertalanffy growth parameters were drawn from user-specified uniform distributions of von Bertalanffy growth parameters. The existing MSE framework was modified to allow correlated uniform distributions to be simulated using copulas, which are functions that join together multiple univariate distributions to form multivariate probability distributions (Nelsen 2005). Correlation coefficients between von Bertalanffy growth parameters were obtained from a literature review of primarily temperate species, due to a paucity of information available for tropical species (SEDAR 2016*b*). Verification testing revealed no errors in modified R code and the achievement of specified levels of correlation between variables (Harford et al. 2016*b*).

Temporal trends in the simulation are divided into a historical period (pre-assessment, no catch limits) and a simulation period (where data-limited approaches are used to set catch limits). The historical period simulates the development of the fishery prior to the implementation of the candidate DLM or MP. The population is initiated in an unfished equilibrium condition and then subjected to a series of annual F rates that are proportional to a user-specified time series of fishing effort, but rescaled so as to achieve a user-specified level of stock depletion at the end of the historical period. The simulation period captures the population response to the data-limited approach and was set at 40 years. Periodic "assessments" are implemented every 3 years, and the corresponding catch advice is imposed each year until the next assessment.

Comparison of performance

Each stochastically generated simulated run is retained, making it possible to subject each data-limited approach to exactly the same sets of simulated conditions and ensuring performance is measured against equivalent sequences of ecological and observational events (Carruthers et al. 2014). Within the MSE, MSY reference points were calculated for each simulation by projecting the operating model forward for 100 years and numerically optimizing ("optimize" function; R Core Team 2016) for the fishing effort that provided the maximum yield (Carruthers et al. 2014). Inherently, this approach assumes that future recruitment is deterministically related to the stock–recruit relationship and that no changes in catchability or fishery targeting have occurred (Carruthers et al. 2014).

MSE was used to identify data-limited approaches that adhere to management objectives and to evaluate trade-offs among a variety of conservation- and exploitation-based performance metrics. Management objectives were based on the US Magnuson–Stevens Fishery Conservation and Management Act, National Standard 1 Guidelines mandate to prevent overfishing and an overfished stock status (NMFS 2009). Three conservation metrics were calculated over the last 10 years of the simulation period, including (*i*) the probability of not overfishing (i.e., the fraction of simulation years where $F < F_{MSY}$, averaged across simulations); (*ii*) the probability of not being overfished (i.e., the fraction of simulation years where $B/B_{MSY} > 0.5$, averaged across simulations); and (*iii*) the probability of the stock not collapsing (i.e., the fraction of simulation years where $B/B_{MSY} > 0.2$, averaged across simulations). A threshold of 50% was specified for the probabilities of not overfishing and of not being overfished in concordance with National Standard 1 Guidelines (NMFS 2009) and previous work in the Gulf of Mexico (SEDAR 2016*b*) and US Caribbean (SEDAR 2016*a*). In addition, three metrics relating to fishery yields were considered, including (*i*) short-term and (*ii*) long-term yields, each defined as the fraction of simulations achieving over 50% F_{MSY} yield over the first and final 5 years of the simulation period, respectively; and (*iii*) the variability in yield, defined as the fraction of simulations achieving less than 15% average annual variability in yield during the entire simulation period (Carruthers et al. 2016). A threshold of 50% was chosen for the probability of interannual variability in yield remaining within 15% of the previous year's yield to capture desirable fluctuations in year-to-year catches (SEDAR 2016*a*, 2016*b*). In reality, stakeholders and managers would prioritize the long-term and short-term yield metrics (e.g., desire higher yield in long-term, short-term, or a compromise between both?). One thousand simulations were deemed appropriate after identifying convergence of performance metrics for each data-limited approach (Carruthers and Hordyk 2016).

Reference period selection for developing data-limited catch advice

Reference mean catch (CC_Ref), CPUE target (Itarget), length target (Ltarget), and LstepCC require a reference period (see Table S1[2]), which specifies the catch series and the target value to be achieved (if required). Ideally, the reference period should be chosen to reflect a stable stock size when the fishery was at or near a sustainable equilibrium (Berkson et al. 2011). Peer-reviewed reference periods have not been identified for all of the data-rich species analyzed in this study, because stock assessments were already developed and data-limited approaches were not previously applied. However, reference periods have been specified for data-limited conspecifics (e.g., lesser amberjack, *Seriola fasciata*) or groupings of species (e.g., shallow-water grouper) by the Gulf of Mexico Fishery Management Council Scientific and Statistical Committee based on expert evaluation of the best scientific information available (e.g., landings period with no trend, minimal influence from management regulations, and perceived stock status; GMFMC 2011). For each species, a simple linear regression on the catch data was used to explore stability during the reference period specified for similar species or for Gulf of Mexico Fishery Management Council-defined reference periods (vermilion snapper). Since no reference period was available for gray triggerfish, we explore multiple periods to demonstrate how selection of reference periods can influence CC_Ref catch advice.

Comparison of catch advice from data-rich and data-limited assessments

For each data-limited approach, a probability density function of catch advice was derived using 10 000 random draws from parameter distributions defined by the input mean and coefficient of variation (Table S5[2]). The median of the probability density function was used for the purpose of comparison (Carruthers et al. 2016). Data-limited catch advice was produced for the year following the terminal year of data and was held constant between assessments. To enable comparisons of catch advice from DLMs and MPs with SS-derived catch advice, all applicable data-limited approaches were used to produce catch advice; however, in reality, only those DLMs and MPs satisfying the performance criteria in the MSE would be considered candidate approaches for providing catch advice.

From SS, catch advice equivalent to the overfishing limit was determined by the prescribed optimal target reference point and current stock status and was extracted from SS projections (via the forecast submodel) 3 years after the terminal assessment year (Table 3). Because quotas are set for a number of years in advance (to account for time lags caused by data collation and assessment implementation), assumed catches are fixed at predetermined quota levels for the first 2 years of the projection in SS. The distribution of the catch recommendation from SS was assumed normal and was obtained using a maximum likelihood approach. Thus, the third year of the projection represents the first year of catch advice; therefore, the forecasted catch (extracted as a point estimate with standard deviation) was used for comparison. Although the years being compared are not identical (e.g., terminal year of data = 2015, data-limited catch advice = 2016, SS catch advice = 2018), the approach to developing catch advice is similar (i.e., produce catch advice for next possible year).

To quantitatively compare DLM and MP catch advice with SS catch advice (i.e., data-rich projection from current stock assessment model; $\mathrm{OFL}_{\mathrm{assessment}}$), the relative difference (RD) was calculated for each species with the following equation:

$$(1)\qquad \mathrm{RD}\ (\%) = \frac{(\mathrm{DLM} - \mathrm{OFL}_{\mathrm{assessment}})}{\mathrm{OFL}_{\mathrm{assessment}}} \times 100$$

Positive RD values indicate higher data-limited catch advice compared with SS catch advice. The SS model reflects the best available science for each species (Cope et al. 2015).

Sensitivity of data-limited catch advice to data inputs

Sensitivity analyses explored the impact of selected data inputs on catch advice. Data inputs explored included available indices of relative abundance (index-based MPs) and mean length (length-based MPs) and depletion estimates and related catch series (DCAC).

Results

Management strategy evaluations of data-limited methods

Method viability

Methods meeting the performance criteria (>50%) relating to overfishing, being overfished, and interannual variability were identified for all species except gray triggerfish (Fig. 2). Empirical MPs (excluding Ltarget) consistently met the performance criteria when feasible (Fig. 2). DCAC met the performance criteria solely for grouper species (note lack of convergence for greater amberjack), whereas DD met the performance criteria for greater amberjack and gag grouper (Fig. 2). Ltarget met the performance criteria solely for gag and red groupers (Fig. 2). Recent mean catch and age-based methods generally did not meet either conservation or yield performance criteria, whereas DBSRA consistently led to highly variable yields. Although reference F_{MSY} resulted in desirable performance metrics related to not overfishing and not being overfished, high interannual variability in yields was noted, possibly due to large variability in stock recruit deviations.

Trade-offs in conservation and yield

Index-based MPs often displayed intermediate metrics in terms of lower conservation metrics (meeting thresholds), but moderate probabilities of getting high yields in the short term, long term, or both (Fig. 2). LstepCC frequently resulted in relatively high conservation metrics at the expense of lower probabilities of getting high long-term yield (Fig. 2). Both Islope configurations resulted in nearly identical performance in the MSE. Trade-offs in long-term and short-term yield metrics were common for most species and approaches considered, particularly for empirical MPs (Fig. 2).

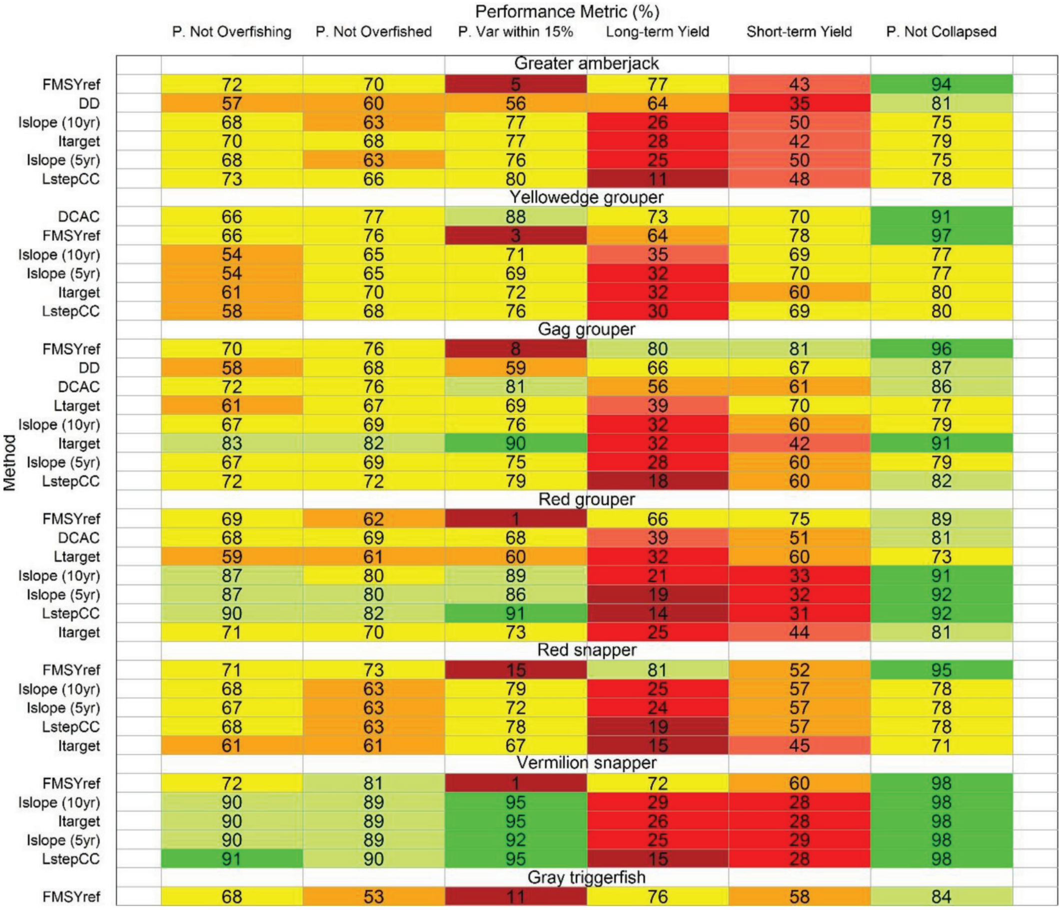
Reference period selection for developing data-limited catch advice

Reference periods exhibiting stable catches with no trend (i.e., $p > 0.05$) were identified for all species (Table 5). For gray triggerfish, catches were stable only during the most recent period (Table 5). The importance of the reference period can be ascertained by examining the mean catch of gray triggerfish during each potential reference period, which forms the basis of catch advice for many empirical MPs. For example, implementing a recent reference period for gray triggerfish resulted in a lower mean catch (792 metric tons, t) in contrast with an older reference period (e.g., 1992–2008, 2409 t; Fig. 3).

Data-rich versus data-limited comparison of catch advice

Catch advice for data-limited approaches that satisfied the performance criteria in the MSE was highly variable and uncertain, with standard deviations (SDs) greatest for DD and smallest for DCAC where viable (Table 6). Nearly all DLM and MP applications exhibited SD exceeding that of SS (Table 6). The degree of agreement between data-limited catch advice and SS was inconsistent across species and methods, although at least one DLM (DBSRA, DCAC) or MP (Islope, LstepCC) recommended catch advice within 25% of SS for nearly all species (Fig. 4). Higher data-limited catch advice was prevalent for CC_Ref, the majority of empirical MPs, and DD, as well as greater amberjack, gag grouper, and red snapper (Fig. 4). DD recommended catch advice substantially higher (two- to fourfold) than the largest catch on record (Table 3) and with the greatest variability (Table 6).

Fig. 2. Method (Table 2) performance, restricted to data-limited methods (DLMs) and management procedures (MPs) meeting performance criteria, in species-specific management strategy evaluations for reef fish species in the Gulf of Mexico. A gradation colour scheme from red (low probability; poor) to green (high probability; good) is used to highlight differences within metrics for each species. [Colour online.]



Method	Performance Metric (%) P. Not Overfishing	P. Not Overfished	P. Var within 15%	Long-term Yield	Short-term Yield	P. Not Collapsed
Greater amberjack						
FMSYref	72	70	5	77	43	94
DD	57	60	56	64	35	81
Islope (10yr)	68	63	77	26	50	75
Itarget	70	68	77	28	42	79
Islope (5yr)	68	63	76	25	50	75
LstepCC	73	66	80	11	48	78
Yellowedge grouper						
DCAC	66	77	88	73	70	91
FMSYref	66	76	3	64	78	97
Islope (10yr)	54	65	71	35	69	77
Islope (5yr)	54	65	69	32	70	77
Itarget	61	70	72	32	60	80
LstepCC	58	68	76	30	69	80
Gag grouper						
FMSYref	70	76	8	80	81	96
DD	58	68	59	66	67	87
DCAC	72	76	81	56	61	86
Ltarget	61	67	69	39	70	77
Islope (10yr)	67	69	76	32	60	79
Itarget	83	82	90	32	42	91
Islope (5yr)	67	69	75	28	60	79
LstepCC	72	72	79	18	60	82
Red grouper						
FMSYref	69	62	1	66	75	89
DCAC	68	69	68	39	51	81
Ltarget	59	61	60	32	60	73
Islope (10yr)	87	80	89	21	33	91
Islope (5yr)	87	80	86	19	32	92
LstepCC	90	82	91	14	31	92
Itarget	71	70	73	25	44	81
Red snapper						
FMSYref	71	73	15	81	52	95
Islope (10yr)	68	63	79	25	57	78
Islope (5yr)	67	63	72	24	57	78
LstepCC	68	63	78	19	57	78
Itarget	61	61	67	15	45	71
Vermilion snapper						
FMSYref	72	81	1	72	60	98
Islope (10yr)	90	89	95	29	28	98
Itarget	90	89	95	26	28	98
Islope (5yr)	90	89	92	25	29	98
LstepCC	91	90	95	15	28	98
Gray triggerfish						
FMSYref	68	53	11	76	58	84

Table 5. Reference years and trend analyses for each reef fish species assessed.

Species	Similar species	Reference years	Significant trend? (R^2, p)
Greater amberjack	Amberjacks	2000–2008	No (0.27, 0.15)
Red grouper	Shallow-water grouper	1995–2008	No (0.03, 0.54)
Gag grouper	Shallow-water grouper	1995–2008	No (0.25, 0.07)
Yellowedge grouper	Deep-water grouper	1992–2008	No (0.00, 0.92)
Red snapper	Midwater snapper	1999–2008	No (0.09, 0.40)
Vermilion snapper	None	1999–2008	No (0.00, 0.98)
Gray triggerfish	None	1992–2008	Yes (0.74, 0.00)
		1995–2008	Yes (0.57, 0.00)
		1999–2008	Yes (0.67, 0.00)
		2000–2008	Yes (0.70, 0.01)
		2006–2015	Yes (0.60, 0.01)
		2011–2015	No (0.24, 0.40)

Note: Years are from similar species in the Gulf of Mexico Reef Fish Fishery Management Plan where necessary (reference period exists for vermilion snapper; GMFMC 2011). R^2 = coefficient of determination from linear regression; p = probability value.

Sensitivity of data-limited catch advice to data inputs

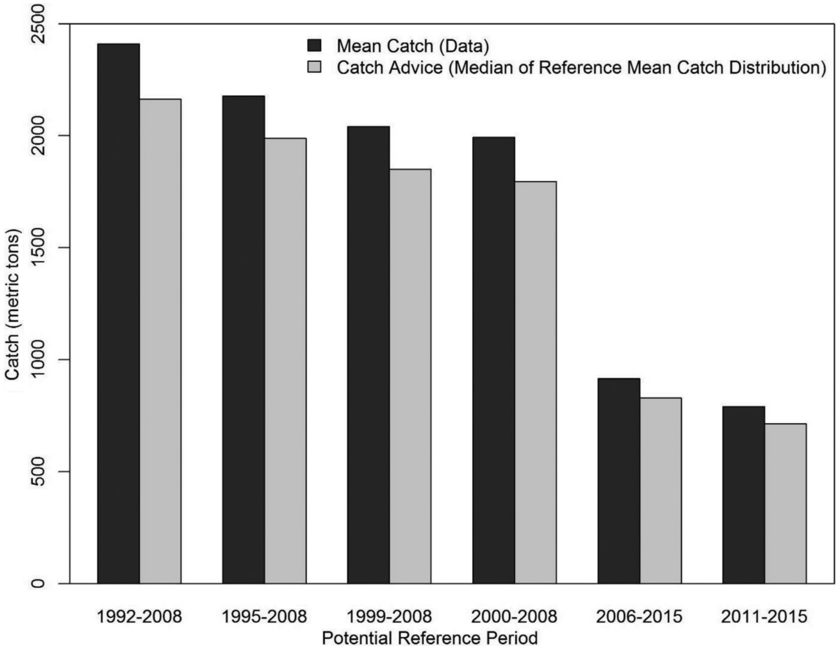
Relative abundance

Available indices of relative abundance ranged from four for yellowedge grouper to 18 for red snapper (Table 3) and often showed conflicting trends in recent abundance patterns (Fig. S1[2]). In many cases, the median Islope catch advice was similar across available indices, although catch advice was slightly higher and less variable when 10 years was used to capture recent trends in contrast with 5 years (Fig. 5). Distributions of catch advice were more variable across indices for Itarget, owing to the specification of target index values, which may differ across indices (Fig. 5). For example, the larger distribution of catch advice for vermilion snapper and the video survey is due to high contrast in relative abundance between the reference (low, 1999–2008) and recent (high, 2010–2014) periods.

Mean length

Length composition from fishery-dependent or fishery-independent data sources was reported in the stock assessment for four species

Fig. 3. Demonstration of the sensitivity of catch advice for gray triggerfish to the selected reference period on mean catch.



(Fig. S2[2]). Trends in mean length were highly variable over time, with breaks in the indices indicative of either low sample sizes (<10 individual observations in year) or lack of samples. Overall, the median catch advice for LstepCC was generally unchanged across available sources of mean length (Fig. 6). Median catch advice was more variable for Ltarget (Fig. 6), because target mean length during the reference period differed between data sources. For example, the larger distribution of catch advice for red grouper using the Marine Recreational Fisheries Statistics Survey index is due to high contrast in mean length between the reference (smaller, 1995–2008) and recent (larger, 2009–2013) periods.

Depletion estimate

DCAC catch advice was consistently higher but more variable when assuming current stock depletion of 0.8 (Fig. 7). For the 0.1 and 0.8 depletion scenarios tested, coefficients of variation ranged from 0.09 to 0.44 and from 0.29 to 2.85, respectively. Interestingly, the species with the longest catch history did not show the smallest difference (red snapper, 30% difference). Instead, differences in catch advice were smallest for greater amberjack (14%) and largest for yellowedge grouper (89%).

Depletion time series

The influence of the length of the time series on DCAC catch advice was inconsistent across species. The catch advice for both the full and recent half of the time series was generally within 25%, with the exception of yellowedge grouper (52%). The full time series provided higher catch advice for gray triggerfish, red grouper, and yellowedge grouper, all species where recent catches were noticeably lower than earlier in the time series (Fig. 7).

Discussion

The re-examination of data-rich assessment-management frameworks using data-limited approaches revealed common patterns and highlighted potential challenges in developing catch advice for data-poor stocks. Application of empirical MPs based on relative abundance or mean length showed considerable promise by consistently satisfying conservation performance metrics (e.g., probability of not overfishing) while displaying intermediate yields based on simulation, with the exception of gray triggerfish (discussed below). These approaches require far fewer data inputs and offer interim and practical solutions to setting catch advice while fishery improvement programs are being developed. However, application of data-limited approaches can be time-consuming when conducting MSE and simultaneously considering various approaches. As observed in the demonstration of reference periods impacting mean catch for gray triggerfish, empirical MP catch advice is largely dependent upon data inputs and assumptions. Ideally, data-limited approaches should only be implemented in the short term where no other approach is feasible, because of high variability in catch advice when compared with SS. However, in areas such as the US Caribbean, application of these approaches may be longstanding due to severe data limitations (Sagarese et al. 2018).

Simple empirical MPs such as Islope and LstepCC showed little sensitivity to the actual index of abundance and mean length used, respectively. In many instances, catch advice from these MPs was within 25% relative difference of SS catch advice, albeit these similarities may have occurred by chance since these approaches are not attempting to optimize yield. Greater variability in catch advice was observed for empirical MPs requiring targets (Itarget and Ltarget) where reference target levels differed across data sources. The stability of catch advice improved for Islope when a longer index time series was considered (e.g., 10 versus 5 years). The flexibility of empirical MPs could enable more frequent assessments and updates compared with data-rich assessments, and, therefore, future collection of abundance and size data by means of high-precision surveys are encouraged for data-limited species. In a data-rich context, these MPs or other empirical MPs could be applied between benchmark SS assessments to adjust SS-derived catch advice.

Empirical MPs were shown to lead to stock rebuilding for overexploited herring (*Clupea* spp.), Atlantic bluefin tuna (*Thunnus thynnus*), and rockfishes (*Sebastes* spp.) life histories (Carruthers et al. 2016). Index-based MPs also showed better performance in terms of risk compared with depletion-based methods for the South African panga (*Pterogymnus lanniarius*) stock (Geromont and Butterworth 2016). Although empirical MPs have gained momentum in recent years (Wayte 2009; Geromont and Butterworth 2015*a*; Carruthers et al. 2016), a number of limitations exist. Previous investigators have cautioned against interpreting changes

Table 6. Catch advice (metric tons) derived from data-limited approaches (Table 2) and Stock Synthesis (SS) for reef fish species in the Gulf of Mexico.

	Percentile			
Method	25th	50th	75th	Standard deviation
Greater amberjack				
DD	10 827	20 268	37 257	28 347
LstepCC	1 497	2 035	2 742	1 015
Islope (5 years)	1 141	1 533	2 064	776
Itarget	1 008	1 352	1 833	661
Islope (10 years)	1 429	1 777	2 214	608
SS	—	770	—	151
Gag grouper				
DD	15 019	28 595	54 446	42 973
LstepCC	3 091	4 143	5 544	1 994
Ltarget	3 044	4 043	5 443	1 958
Islope (5 years)	1 786	2 394	3 175	1 161
Islope (10 years)	2 173	2 690	3 303	884
Itarget	503	674	903	330
DCAC	1 899	2 075	2 224	254
SS	—	1 973	—	301
Red snapper				
Itarget	12 735	17 092	23 155	8 596
Islope (5 years)	9 176	12 442	16 814	6 477
Islope (10 years)	7 777	9 714	12 157	3 441
SS	—	4 571	—	344
Gray triggerfish				
SS	—	1 107	—	266
Yellowedge grouper				
LstepCC	266	356	477	171
Islope (5 years)	243	327	437	157
Islope (10 years)	312	384	473	127
Itarget	188	251	336	118
DCAC	149	215	289	97
SS	—	375	—	3
Red grouper				
Ltarget	10 185	14 113	19 616	7 990
Itarget	7 127	9 913	13 723	5 469
LstepCC	6 314	8 761	12 228	4 926
Islope (5 years)	4 472	6 224	8 667	3 554
Islope (10 years)	6 935	8 786	11 188	3 372
DCAC	5 353	5 943	6 489	840
SS	—	7 127	—	646
Vermilion snapper				
Itarget	1 864	2 500	3 389	1 276
Islope (5 years)	1 023	1 390	1 879	690
Islope (10 years)	1 251	1 546	1 928	537
SS	—	1 902	—	72

Note: Results are only shown for methods meeting performance criteria.

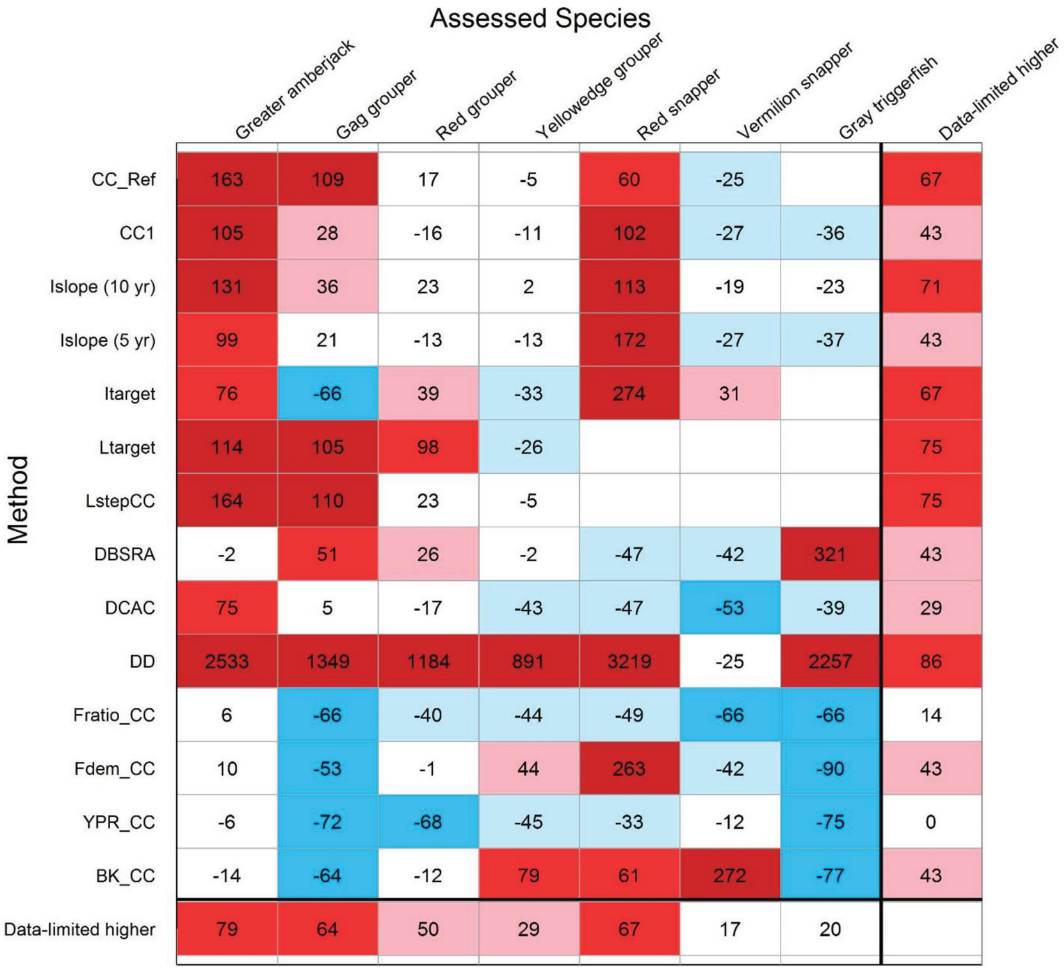
in mean length as a response in the stock abundance, because of the strict assumptions (e.g., constant recruitment; Table S1[2]), as well as concerns that noise in mean length will be interpreted as a signal in the data (Geromont and Butterworth 2016). Although length-based MPs may be influenced by recruitment variability, in our study LstepCC still led to reasonable management outcomes. Other limitations acknowledged are simplifying assumptions such as a single fishing fleet and the representativeness of the operating model. For example, the apparent disconnect between MSE performance of Islope (i.e., high probability of not overfishing) and application using actual data (i.e., Islope catch advice often exceeds SS advice) could indicate a mismatch between actual dynamics in the operating model and reality, although for some species this may result from trends in actual versus simulated data (e.g., recent catches exceed the ACL for greater amberjack).

Key issues to consider when examining the performance of data-limited approaches in simulation analysis are the level of depletion assumed to characterize the stock dynamics and the tuning of parameters to achieve yield-risk performance (Geromont and Butterworth 2015*a*). While data are rarely available to infer stock status for data-limited stocks (Carruthers et al. 2014; Cope et al. 2015), the economic performance of data-limited approaches is often dependent on assumed depletion status (Harford et al. 2016*a*). When uncertainty in stock status exists, performance of data-limited approaches can be evaluated across different plausible depletion levels (Carruthers et al. 2016), which would enable managers to examine the trade-offs and risks of specifying catches based on different assumed stock conditions (e.g., risk in probability of not overfishing of assuming a more depleted stock?). In this study, estimates of current stock depletion were available from SS and were at or close to the range assumed by Geromont and Butterworth (2015*a*), suggesting the assumption inherent in the generic "off-the-shelf" configurations of target-based MPs was appropriate to the reef fish stocks that we evaluated (i.e., stocks likely in an overexploited state). Generic off-the-shelf parameterizations were intended to capture "medium productivity" and "severely depleted" (current biomass between 10% and 30% of the pre-exploitation level) stocks in South Africa after tuning to yield-risk performance (Geromont and Butterworth 2015*a*) and may not be applicable to other regions. Any assessment implementing these approaches must evaluate data inputs and fine-tune each MP to ensure assumed parameters reflect the status of the species under evaluation (Geromont and Butterworth 2016). Common practice of borrowing data from similar species (e.g., "Robin Hood" approach; Punt et al. 2011) will not be appropriate if exploitation rates and targeting behavior (e.g., infrequently encountered bycatch species) differ substantially between species (SEDAR 2016*b*).

Tuning an MP entails adjusting the values of its control parameters to achieve particular objectives such as improved yield-risk performance (Geromont and Butterworth 2016). For example, in the off-the-shelf parameterization, the default target scalar value of 1.5 in Itarget implies a target CPUE that is 1.5 times the reference mean CPUE, which assumes that the stock was experiencing overfishing during the reference period (Geromont and Butterworth 2015*a*). Although the need to conserve stocks by reducing fishing effort is sound from a conservation perspective (Worm et al. 2009), the lower realized yields may be unacceptable and can have detrimental impacts on fleet economics (e.g., reduced catches often lead to lower income; Béné 2003). Luckily, these methods offer flexibility in terms of configuring parameters, which specify the target CPUE you want to achieve, the rate of change in the catch advice, and the threshold below which catch advice is greatly reduced (Geromont and Butterworth 2015*a*). Although it has been suggested by Geromont and Butterworth (2016) that applying the precautionary off-the-shelf methods can be effective and may alleviate costs of tuning methods for low-value data-poor stocks, there is considerable concern in the US of setting catch limits too low for nontarget species ("choke" species; Schrope 2010), which can lead to premature closures of other more lucrative fisheries (Baudron and Fernandes 2015). Therefore, it is responsible practice to thoroughly vet assumptions and either exclude methods requiring such assumptions or tune the methods using the best available information on stock status based on expert opinion.

All assessment methods, regardless of complexity, require assumptions and careful consideration of all aspects of the analysis. Even the simplest catch-only approach, CC_Ref, requires the assumption that the reference period exhibits stable biomass with surplus production that is commensurate with the chosen catch level (i.e., minimal influence of management regulations). Thus, if this catch level is continued, it is unlikely to cause overfishing.

Fig. 4. Relative difference (as a percentage) between catch advice derived from data-limited methods (Table 2) and Stock Synthesis for reef fish species in the Gulf of Mexico. The "data-limited higher" column refers to the percentage of times catch advice for the DLM or MP was higher. [Colour online.]



Method	Greater amberjack	Gag grouper	Red grouper	Yellowedge grouper	Red snapper	Vermilion snapper	Gray triggerfish	Data-limited higher
CC_Ref	163	109	17	-5	60	-25		67
CC1	105	28	-16	-11	102	-27	-36	43
Islope (10 yr)	131	36	23	2	113	-19	-23	71
Islope (5 yr)	99	21	-13	-13	172	-27	-37	43
Itarget	76	-66	39	-33	274	31		67
Ltarget	114	105	98	-26				75
LstepCC	164	110	23	-5				75
DBSRA	-2	51	26	-2	-47	-42	321	43
DCAC	75	5	-17	-43	-47	-53	-39	29
DD	2533	1349	1184	891	3219	-25	2257	86
Fratio_CC	6	-66	-40	-44	-49	-66	-66	14
Fdem_CC	10	-53	-1	44	263	-42	-90	43
YPR_CC	-6	-72	-68	-45	-33	-12	-75	0
BK_CC	-14	-64	-12	79	61	272	-77	43
Data-limited higher	79	64	50	29	67	17	20	

Choosing an appropriate reference period corresponding to an underexploited or fully exploited stage in the utilization of the resource is paramount for reliably implementing catch-only methods, and the use of catch-only methods is only supported when more adaptive MPs are unfeasible (Carruthers et al. 2014, 2016; Sagarese et al. 2018). Reference period assumptions are also important for empirical MPs requiring target levels of mean length or relative abundance. In instances where a reference period is not available or supported, Islope could be applied since it is not dependent upon a reference period.

Both DD and DCAC occasionally satisfied the performance criteria in the MSE and exhibited relatively larger probabilities of high yields in the short or long term. However, application of these approaches using actual data can be complicated by a lack of contrast in the data (DD) or the lack of depletion estimates (DCAC) for truly data-poor stocks. For example, DD catch advice developed using actual data was deemed unrealistic for all species because it was extremely variable (Table 6) and was substantially higher than both the largest observed catch (Table 3) and the SS estimate (Table 6). Potential reasons for this discrepancy in performance may be poor contrast in actual index data or an operating model not capturing the actual dynamics (i.e., simulated data are not as poor as in reality).

DCAC performed relatively well for grouper species in the MSE, although additional analyses using actual data showed the sensitivity of this approach to the depletion estimate, an input that is rarely known in reality (Cope et al. 2015; Harford and Carruthers 2017). The time series of catch recommended for assessment was used for DCAC, an assumption that would require additional vetting prior to being used to set catch advice. The implementation of DCAC must be used with caution, as this method is not directly suitable for specifying catches in a stock-rebuilding plan and is a one-time only calculation (MacCall 2009; Carruthers et al. 2014). Data-poor stocks in the Pacific have been assessed using DCAC and DBSRA (Newman et al. 2014, 2015), and these methods have been reviewed favorably by analytical panels (NMFS 2011; Stokes 2011). In our study, simulation analyses revealed poor performance of DBSRA, whereas applications using actual data led to ballpark estimates within SS for a few species, some of which did not possess complete time series since unfished conditions (e.g., red grouper; Table 3).

A goal of this analysis was to evaluate the agreement of catch advice between data-limited approaches and data-rich methods using the same available data, with the SS assessment results treated as benchmarks to compare against because they reflect the best available scientific information. For the application to real data, we simply calculated the next year's data-limited catch advice and did not project forward in time, as typically done using SS. Additional considerations of how short-term data-limited projections compare with SS projections may further uncover trends in DLM and MP behavior. In addition, research exploring how catch advice changes between assessments (e.g., fixed or incremental changes) as well as a reverse retrospective analysis would be interesting to see how data-limited catch advice over time

Fig. 5. Sensitivity of index-based catch advice to the index of abundance (Table S6[2]) for reef fish species in the Gulf of Mexico (fishery-independent shown with shaded bars). Solid, thin horizontal line in each panel identifies SS-derived catch advice, a single line for each index within a panel reflects not applicable (NA), and *y* axes differ among panels.

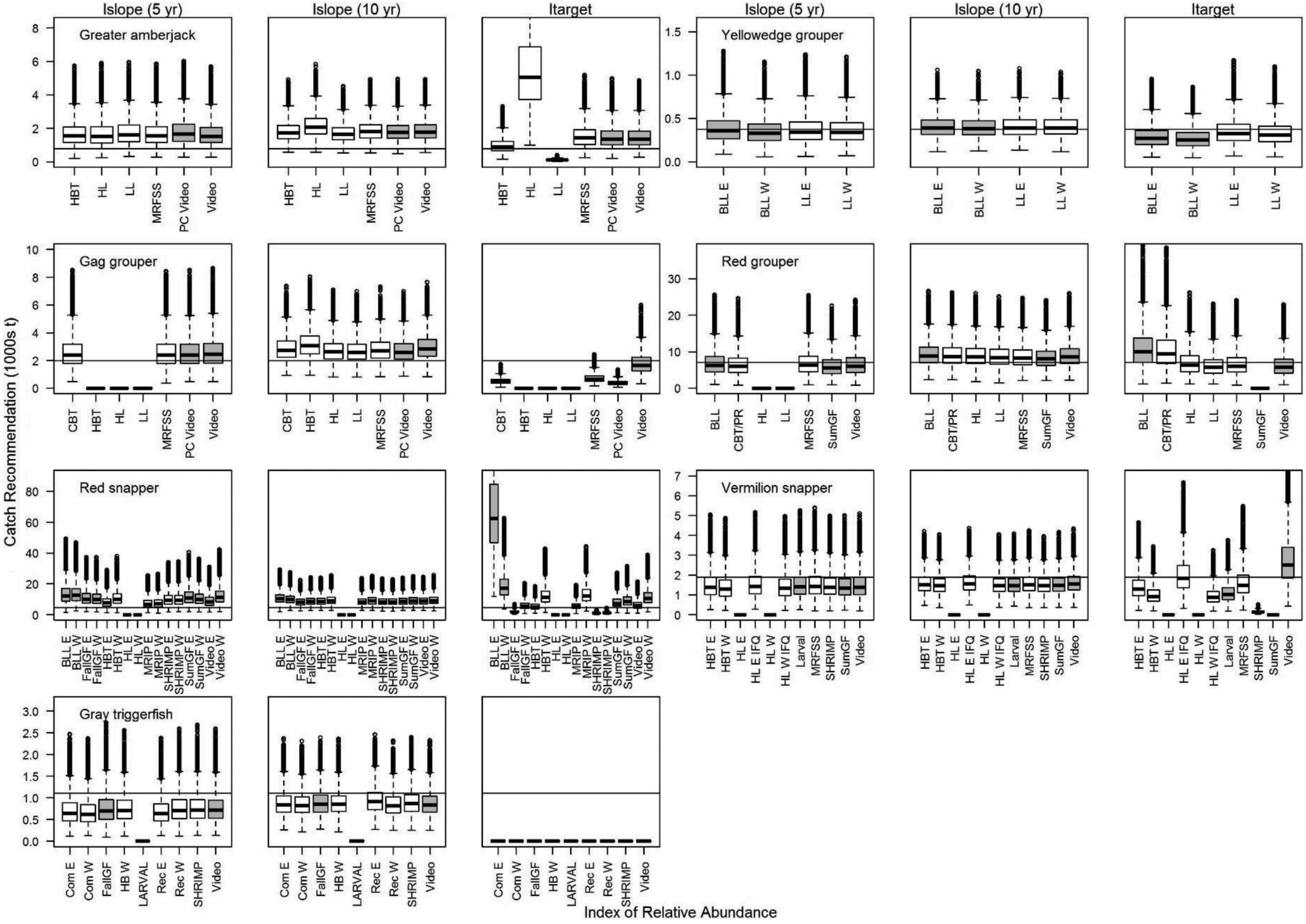
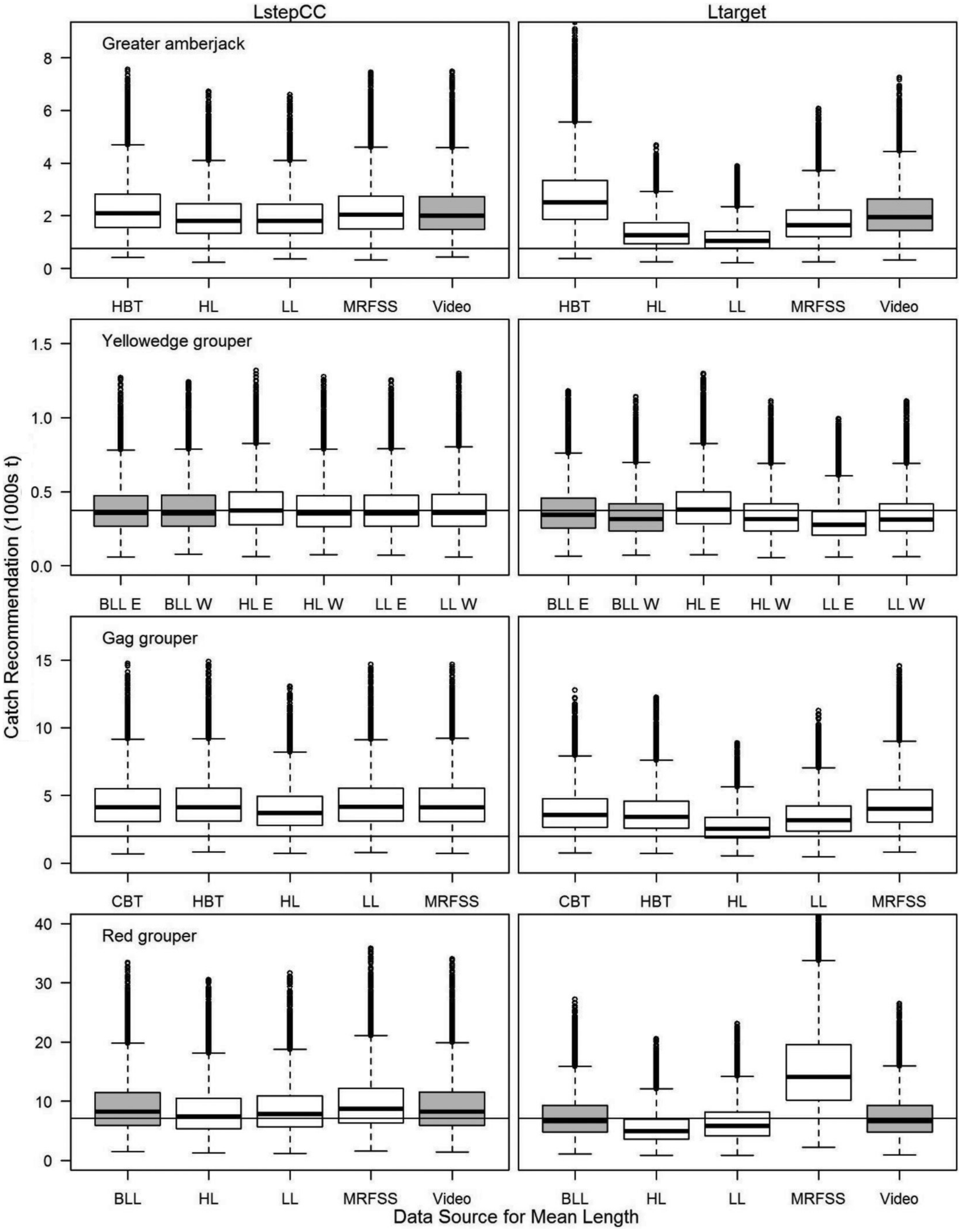


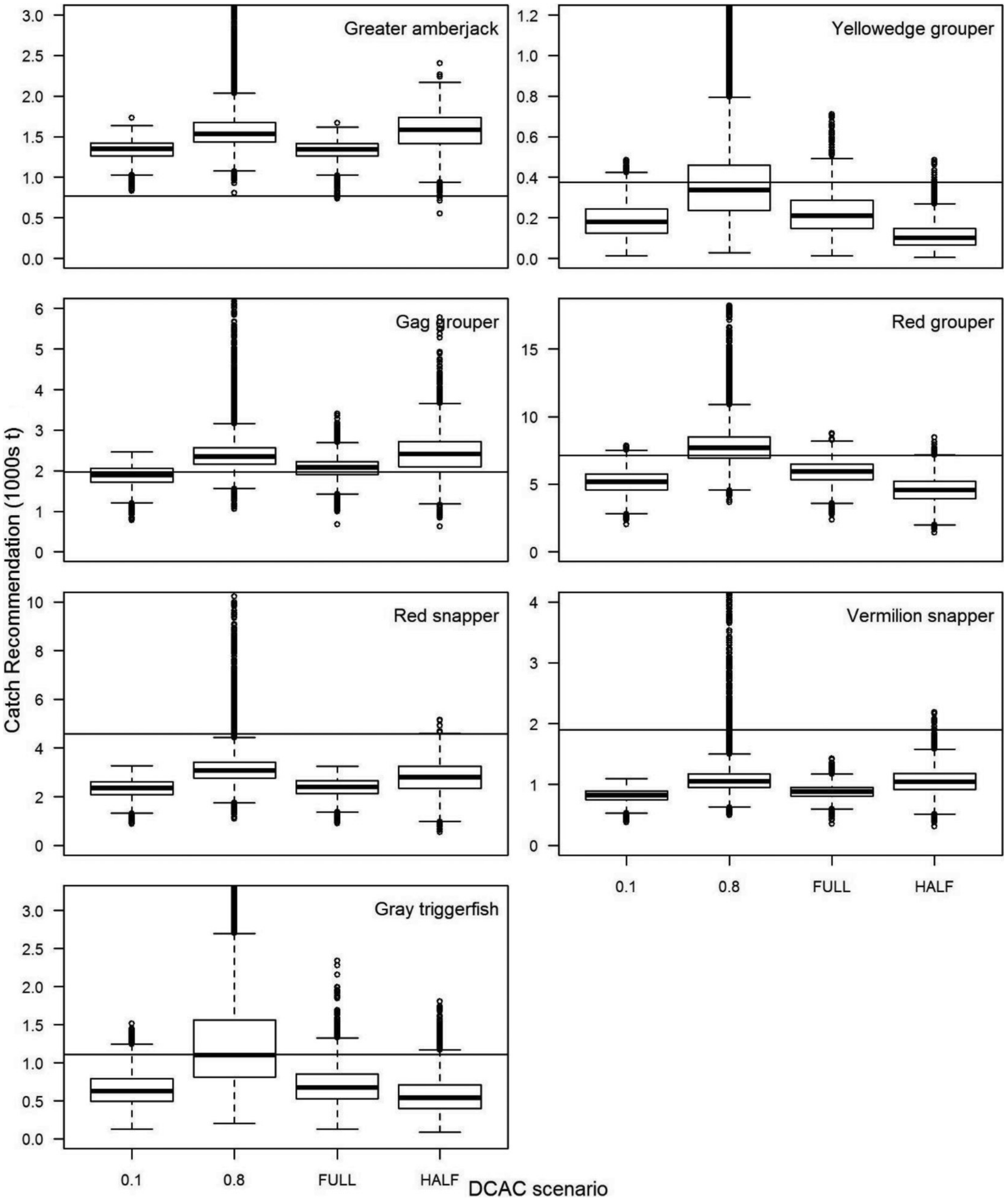
Fig. 6. Sensitivity of length-based catch advice to the index of mean length (fishery-independent shown with shaded bars) for reef fish species in the Gulf of Mexico. See Fig. 5 caption for details.



would compare with data-rich assessment results. Similarity in catch advice does not necessarily imply that **any** of the models are more correct or preferred over another, as all assessment models are intrinsically built upon assumptions. Simulation analysis revealed that multiple plausible assessment models may exist and not just a single "best" assessment model (Deroba et al. 2015). In addition, comparisons between data-rich methods and data-limited approaches are limited due to the difficulty in assessing risks and the inability to quantify bias (Carruthers et al. 2014).

Some results of this analysis were unexpected, such as the absence of any data-limited approach meeting the performance criteria for gray triggerfish. Although the gray triggerfish stock assessment was accepted as the best available science, it was not used to provide management advice because of Scientific and Statistical Committee concerns over higher recommended catch levels. Specific SS assessment concerns included modeled growth and steepness (SEDAR 2015), which were both inputs used in simulating gray triggerfish stock dynamics. Of 21 stocks simulated to date (SEDAR 2016*a*, 2016*b*; this study), this is the only species where no viable methods were identified, suggesting a review of the operating model. Equally as surprising was the performance of age-based methods within the MSE. Additional simulation testing could help determine whether alternative configurations that improve simple catch curve analysis (Carruthers and Hordyk 2016)

Fig. 7. Sensitivity of depletion-corrected average catch (DCAC) catch advice to the estimate of depletion over time (Dt; 0.1 or 0.8) and the length of the time series (HALF or FULL) for reef fish species in the Gulf of Mexico. See Fig. 5 caption for details. Note that scenarios' varying time series length use base depletion levels (Table 3).



could result in better performance. For example, current configurations use the last 2 years of age composition and the terminal year of catch rather than a specified time period of catch.

Both management strategy evaluation and application using actual data identified empirical MPs as promising options for assessing data-poor stocks in the southeastern US and elsewhere. Such approaches are clear improvements over static catch-only methods, which currently serve as the status quo for management of data-limited fisheries in the southeastern US (GMFMC 2011). These adaptive MPs may be the only viable path for fisheries management in areas such as the US Caribbean (Sagarese et al. 2018). Although designated a data-limited exercise, application of data-limited approaches requires considerable input on data availability and quality, sensitivity analyses of all required data inputs (e.g., *M* for DCAC), and defensible decision rules to minimize misuse. First and foremost, emphasis needs to be placed on stakeholder input in terms of management objectives, education on how data-limited approaches are subsequently designed, potential for data collection (e.g., collect length over age data?), and, ultimately, the need to find MSY (i.e., are adaptive approaches acceptable options for management?). Selection of data-limited approaches for providing catch advice requires careful consideration of assumptions (e.g., reference period), tuning, and an understanding of the potential negative consequences of each approach. DLMtool represents a powerful management tool with the capability of increasing throughput of data-limited evaluations, particularly in areas with diverse resources and high reliance on artisanal fisheries (e.g., the US Caribbean Sea).

Acknowledgements

Gratitude is expressed to the data providers and analysts for each SEDAR. This research was initiated as part of an NOAA Stock Assessment Analytical Methods grant initiative and was carried out (in part) under the auspices of the Cooperative Institute for Marine and Atmospheric Studies, a Cooperative Institute of the University of Miami and the NOAA – Southeast Fisheries Science Center, cooperative agreement No. NA17RJ1226.

References

Baudron, A.R., and Fernandes, P.G. 2015. Adverse consequences of stock recovery: European hake, a new "choke" species under a discard ban? Fish Fish. **16**(4): 563–575. doi:10.1111/faf.12079.

Beddington, J., and Kirkwood, G. 2005. The estimation of potential yield and stock status using life-history parameters. Philos. Trans. R. Soc. B Biol. Sci. **360**(1453): 163–170. doi:10.1098/rstb.2004.1582.

Béné, C. 2003. When fishery rhymes with poverty: a first step beyond the old paradigm on poverty in small-scale fisheries. World Dev. **31**(6): 949–975. doi:10.1016/S0305-750X(03)00045-7.

Berkson, J., and Thorson, J.T. 2015. The determination of data-poor catch limits in the United States: is there a better way? ICES J. Mar. Sci. **72**(1): 237–242. doi:10.1093/icesjms/fsu085.

Berkson, J., Barbieri, L., Cadrin, S., Cass-Calay, S., Crone, P., Dorn, M., Friess, C., Kobayashi, D., Miller, T.J., Patrick, W.S., Pautzke, S., Ralston, S., and Trianni, M. 2011. Calculating acceptable biological catch for stocks that have reliable catch data only (only reliable catch stocks – ORCS). NOAA Technical Memorandum NMFS-SEFSC-616.

Beverton, R.J.H., and Holt, S.J. 1957. On the dynamics of exploited fish populations. Springer-Science+Business Media, B.V., London, UK.

Brooks, E.N., Shertzer, K.W., Gedamke, T., and Vaughan, D.S. 2008. Stock assessment of protogynous fish: evaluating measures of spawning biomass used to estimate biological reference points. Fish. Bull. **106**(1): 12–23.

Bryan, D.R., Smith, S.G., Ault, J.S., Feeley, M.W., and Menza, C.W. 2016. Feasibility of a regionwide probability survey for coral reef fish in Puerto Rico and the U.S. Virgin Islands. Mar. Coast. Fish. **8**(1): 135–146. doi:10.1080/19425120.2015.1082520.

Butterworth, D.S., Bentley, N., De Oliveira, J.A., Donovan, G.P., Kell, L.T., Parma, A.M., Punt, A.E., Sainsbury, K.J., Smith, A.D., and Stokes, T.K. 2010. Purported flaws in management strategy evaluation: basic problems or misinterpretations? ICES J. Mar. Sci. **67**: 567–574. doi:10.1093/icesjms/fsq009.

Carruthers, T.R., and Hordyk, A. 2016. Package 'DLMtool', version 3.2.2 [online]. Available from http://cran.r-project.org/web/packages/DLMtool/index.html.

Carruthers, T.R., Punt, A.E., Walters, C.J., MacCall, A., McAllister, M.K., Dick, E.J., and Cope, J. 2014. Evaluating methods for setting catch limits in data-limited fisheries. Fish. Res. **153**: 48–68. doi:10.1016/j.fishres.2013.12.014.

Carruthers, T., Kell, L., Butterworth, D., Maunder, M., Geromont, H., Walters, C., McAllister, M., Hillary, R., Levontin, P., Kitakado, T., and Davies, C. 2016. Performance review of simple management procedures. ICES J. Mar. Sci. **73**(2): 464–482. doi:10.1093/icesjms/fsv212.

Cope, J.M. 2013. Implementing a statistical catch-at-age model (Stock Synthesis) as a tool for deriving overfishing limits in data-limited situations. Fish. Res. **142**: 3–14. doi:10.1016/j.fishres.2012.03.006.

Cope, J.M., Thorson, J.T., Wetzel, C.R., and DeVore, J. 2015. Evaluating a prior on relative stock status using simplified age-structured models. Fish. Res. **171**: 101–109. doi:10.1016/j.fishres.2014.07.018.

Costello, C., Ovando, D., Hilborn, R., Gaines, S.D., Deschenes, O., and Lester, S.E. 2012. Status and solutions for the world's unassessed fisheries. Science, **338**(6106): 517–520. doi:10.1126/science.1223389. PMID:23019613.

Deriso, R.B. 1980. Harvesting strategies and parameter estimation for an age-structured model. Can. J. Fish. Aquat. Sci. **37**(2): 268–282. doi:10.1139/f80-034.

Deroba, J., Butterworth, D.S., Methot, R., De Oliveira, J., Fernandez, C., Nielsen, A., Cadrin, S., Dickey-Collas, M., Legault, C., Ianelli, J.,, et al. 2015. Simulation testing the robustness of stock assessment models to error: some results from the ICES strategic initiative on stock assessment methods. ICES J. Mar. Sci. **72**(1): 19–30. doi:10.1093/icesjms/fst237.

Dick, E., and MacCall, A.D. 2011. Depletion-based stock reduction analysis: a catch-based method for determining sustainable yields for data-poor fish stocks. Fish. Res. **110**(2): 331–341. doi:10.1016/j.fishres.2011.05.007.

Geromont, H., and Butterworth, D. 2015*a*. Generic management procedures for data-poor fisheries: forecasting with few data. ICES J. Mar. Sci. **72**(1): 251–261. doi:10.1093/icesjms/fst232.

Geromont, H.F., and Butterworth, D.S. 2015*b*. Complex assessments or simple management procedures for efficient fisheries management: a comparative study. ICES J. Mar. Sci. **72**: 262–274. doi:10.1093/icesjms/fsu017.

Geromont, H.F., and Butterworth, D.S. 2016. A review of assessment methods and the development of management procedures for data-poor fisheries [online]. United Nations Food and Agriculture Organisation (FAO). Available from https://www.researchgate.net/publication/297195982_FAO_Report_A_review_of_assessment_methods_and_the_development_of_management_procedures_for_data-poor_fisheries_219pp.

GMFMC. 2011. Final generic annual catch limits/accountability measures amendment for the Gulf of Mexico Fishery Management Council's red drum, reef fish, shrimp, coral and coral reefs, fishery management plans. Gulf of Mexico Fishery Management Council, Tampa, Florida.

Gulland, J. 1971. Science and fishery management. J. Cons. Int. Explor. Mer, **33**(3): 471–477. doi:10.1093/icesjms/33.3.471.

Harford, W.J., and Carruthers, T.R. 2017. Interim and long-term performance of static and adaptive management procedures. Fish. Res. **190**: 84–94. doi:10.1016/j.fishres.2017.02.003.

Harford, W., Gedamke, T., Babcock, E., Carcamo, R., McDonald, G., and Wilson, J. 2016*a*. Management strategy evaluation of a multi-indicator adaptive framework for data-limited fisheries management. Bull. Mar. Sci. **92**(4): 423–445. doi:10.5343/bms.2016.1025.

Harford, W.J., Sagarese, S.R., Isely, J.J., and Smith, M.S. 2016*b*. Technical description of operating models in data-limited methods toolkit (DLMtool) [online]. SEDAR49-AW-10, SEDAR, North Charleston, South Carolina. Available from sedarweb.org [accessed February 2018].

Hordyk, A., Newman, D., Carruthers, T., and Suatoni, L. 2017. Applying management strategy evaluation to California Fisheries: case studies and recommendations [online]. Data Limited Methods Toolkit. Available from https://www.datalimitedtoolkit.org/wp-content/uploads/2017/07/Applying-MSE-to-CA-Fisheries-Case-Studies-Recommendations.pdf [accessed December 2017].

Kell, L.T., Mosqueira, I., Grosjean, P., Fromentin, J.-M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M., Poos, J., Scott, F., and Scott, R.D. 2007. FLR: an open-source framework for the evaluation and development of management strategies. ICES J. Mar. Sci. **64**: 640–646. doi:10.1093/icesjms/fsm012.

MacCall, A.D. 2009. Depletion-corrected average catch: a simple formula for estimating sustainable yields in data-poor situations. ICES J. Mar. Sci. **66**(10): 2267–2271. doi:10.1093/icesjms/fsp209.

Martell, S., and Froese, R. 2013. A simple method for estimating MSY from catch and resilience. Fish Fish. **14**(4): 504–514. doi:10.1111/j.1467-2979.2012.00485.x.

Maunder, M. 2014. Management strategy evaluation (MSE) implementation in Stock Synthesis: application to Pacific bluefin tuna. Document SAC-05-10b, Inter-American Tropical Tuna Commission.

McAllister, M.K., Pikitch, E.K., and Babcock, E. 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. Can. J. Fish. Aquat. Sci. **58**(9): 1871–1890. doi:10.1139/f01-114.

Methot, R. 2012. User manual for stock synthesis: model version 3.24f. US Department of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Seattle, Washington.

Methot, R.D., and Wetzel, C.R. 2013. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. Fish. Res. **142**: 86–99. doi:10.1016/j.fishres.2012.10.012.

MSFCMA. 2006. Magnuson–Stevens Fishery Conservation and Management Reauthorization Act of 2006. U.S. Public Law 109–479, 120 Statute 3575.

Nelsen, R.B. 2005. 14: Copulas and quasi-copulas: an introduction to their properties and applications. *In* Logical, algebraic, analytic, and probabilistic aspects of triangular norms. *Edited by* E.P. Klement and R. Mesiar. Elsevier, Amsterdam. pp. 391–413.

Newman, D., Carruthers, T.R., MacCall, A., Porch, C., and Suatoni, L. 2014. Improving the science and management of data-limited fisheries: an evaluation of current methods and recommended approaches. NRDC Report R:14-09-B, Natural Resources Defense Council, New York.

Newman, D., Berkson, J., and Suatoni, L. 2015. Current methods for setting catch limits for data-limited fish stocks in the United States. Fish. Res. **164**: 86–93. doi:10.1016/j.fishres.2014.10.018.

NMFS. 2009. NMFS's National standard guidelines. 50 C.F.R. 600.310 et seq.

NMFS. 2011. Assessment methods for data-poor stocks report of the review panel meeting. National Marine Fisheries Service (NMFS), Southwest Fisheries Science Center (SWFSC), Santa Cruz, California.

NOAA. 2013. Southeast Fisheries Science Center Science Plan 2013–2018. National Oceanic and Atmospheric Administration.

Punt, A.E., Smith, D.C., and Smith, A.D. 2011. Among-stock comparisons for improving stock assessments of data-poor stocks: the "Robin Hood" approach. ICES J. Mar. Sci. **68**(5): 972–981. doi:10.1093/icesjms/fsr039.

Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A., and Haddon, M. 2016. Management strategy evaluation: best practices. Fish Fish. **17**(2): 303–334. doi:10.1111/faf.12104.

R Core Team. 2016. R: a language and environment for statistical computing [online]. Available from https://www.R-project.org/.

Sagarese, S.R., Rios, A.B., Cass-Calay, S.L., Cummings, N.J., Bryan, M.D., Stevens, M.S., Harford, W.J., McCarthy, K.J., and Matter, V.M. 2018. Working towards a framework for stock evaluations in data-limited fisheries. N. Am. J. Fish. Manage. **38**(3): 507–537. doi:10.1002/nafm.10047.

Sainsbury, K.J., Punt, A.E., and Smith, A.D. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES J. Mar. Sci. **57**(3): 731–741. doi:10.1006/jmsc.2000.0737.

Schnute, J. 1985. A general theory for analysis of catch and effort data. Can. J. Fish. Aquat. Sci. **42**(3): 414–429. doi:10.1139/f85-057.

Schrope, M. 2010. Fisheries: What's the catch? Nature, **465**: 540–542. doi:10.1038/465540a. PMID:20520684.

SEDAR. 2015. SEDAR43: Gulf of Mexico Gray Triggerfish Stock Assessment Report [online]. SEDAR, North Charleston, South Carolina. Available from sedarweb.org [accessed February 2018].

SEDAR. 2016*a*. SEDAR46: U.S. Caribbean Data-limited Species Stock Assessment Report [online]. SEDAR, North Charleston, South Carolina. Available from sedarweb.org [accessed February 2018].

SEDAR. 2016*b*. SEDAR49: Gulf of Mexico Data-limited Species Stock Assessment Report [online]. SEDAR, North Charleston, South Carolina. Available from sedarweb.org [accessed February 2018].

Stokes, T.K. 2011. Report on the workshop for data limited stock assessment methods (or review of assessment methods for data-poor stocks). The Center for Independent Reviews, Wellington, New Zealand.

Walters, C., and Martell, S.J. 2002. Stock assessment needs for sustainable fisheries management. Bull. Mar. Sci. **70**(2): 629–638.

Wayte, S. 2009. Evaluation of new harvest strategies for SESSF species. CSIRO Marine and Atmospheric Research, Hobart and Australian Fisheries Management Authority, Canberra.

Worm, B., Hilborn, R., Baum, J.K., Branch, T.A., Collie, J.S., Costello, C., Fogarty, M.J., Fulton, E.A., Hutchings, J.A., Jennings, S.,, et al. 2009. Rebuilding global fisheries. Science, **325**(5940): 578–585. doi:10.1126/science.1173146. PMID:19644114.