NOTE: This document was excerpted from a larger report and is not citable.

This section of the *Illex illecebrosus* Research Track Stock Assessment will be presented to the MAFMC SSC by John Manderson, Open Ocean Research

GENERALIZED DEPLETION MODEL (GDM; Manderson & Mercer 2022)

Introduction

Depletion models can use short time steps (days to months) in the analysis of fishery catch and effort to estimate the abundance of the vulnerable portion of a population required to support fishery catches. The approach is used in many parts of the world to assess cephalopod populations because the time scale of the analyses can be made to match the dynamics of short lived, semelparous, and environmentally controlled populations that are typically not well sampled by fishery independent surveys and with respect to age composition (see Arkhipkin *et al.* 2020). Depletion modeling does not provide a full stock assessment of population abundance, productivity and condition relative to fishery reference points. Instead, depletion models develop minimal assessments of species catchability by the fishery, the abundance of the population vulnerable to the fishery (N_o), fishing mortality (F) with respect to the vulnerable population, and natural mortality (M). The results of intra-annual depletion analysis can used in a hierarchical manner to inform assessment models that develop a full suite of fishery reference points (e.g. Roa-Ureta *et al.* 2021)

Depletion models can also be used to estimate fishery escapement (H) relative to sustainable biomass targets by calculating the proportion of the vulnerable stock (N_0) remaining at the end of the fishing season (see Roa-Ureta 2012, Arkhipkin *et al.* 2008, Lin *et al.* 2017, Maynou *et al.* 2021). This is an important application in assessments of squid that have sub-annual life cycles and therefore lack portfolios of age classes that can buffer populations from recruitment failure (Beddington *et al.* 1990). Depletion modeling is considered useful for in-season assessment and management of cephalopod fisheries (Robert 2010). Since the early 1990s, depletion based methods have been discussed in U.S. *Illex illecebrosus* assessments as potentially useful for developing flexible, responsive strategies of in-season fishery management (NEFSC 1992, 1994, 1996, 1999, 2003, 2006).

Classical depletion models combine submodels for the harvested population and the fishery observation process. In the simplest case (Leslie model) the population submodel is:

$$1.\ N_t = N_0 - K_{t-1}$$

where N_t is population size at time t, N_0 is the initial size of the population just before the fishery begins and K_{t-1} is cumulative catch prior to time t.

The observation submodel is:

$$2.\ X_t = qN_t$$

where X_t is observed catch per unit effort and q is catchability of the stock per unit effort in the fishery and N_t is the latent population size at time t.

Substituting the submodel of the harvested population into the observation model yields

$$3.\ X_t = q(N_0 - K_{t-1})$$

or

$$4.\ X_t = qN_0 - qK_{t-1}$$

This linear expression applied to data yields estimates of the catchability coefficient of the fishery (-1*slope=q) and initial population abundance before the fishery begins (intercept/q). In real world applications, depletion models account for losses of individuals resulting from natural morality (M) during the fishing season using Pope's (1972) recursive relation (Chapman 1974). Classical depletion models make the following assumptions:

1) The population is physically and demographically closed to processes of immigration and/or recruitment that increase abundance or emigration that results in losses of animals from the fishing area during the fishing season independently of the fishing process and natural mortality,

2) Losses due to natural mortality are constant over the fishing season,

3) Fishery catch and population abundance are linearly related by the catchability (q) parameter which can therefore be used to estimate vulnerable population abundance (N_0) from the fishery catch,

4) Catchability (q) is constant over the period of fishery removals. A large pool of animals in the population does not have a refuge from the fishery and therefor a q=0,

5) Units of fishing effort are independent and do not compete with each other,

6) Fishing capacity is sufficiently large to deplete the vulnerable portion of the stock at levels required to estimate of output parameters; at a minimum q and N_0, and

7) Assumptions of linear regression are met including random sampling, error free measurement of catch and effort, and independent and normally distributed errors.

Recently, Rago (2021) applied classical Leslie-Davis depletion analysis to weekly U.S. catches of and fishing effort for *Illex illecebrosus* during 19 of the years from 1997 through 2019. Landings per unit effort (LPUE) decreased continuously to produce significantly negative regression slopes as required by classical depletion modeling in only 4 of the 19 years. In 7 years, LPUE increased substantially over fishing seasons, regressions produced positive slopes, negative qs, negative x axis intercepts and thus negative initial population size estimates. Rago (2021) concluded *"the failure of the Leslie-Davis depletion models suggests that migrations into the fishing area, variations in growth, and recruitment overwhelm depletions associated with the fishery" (i*.e. violations of the open population assumption [#1 above] and possibly assumptions of sufficient fishing capacity in the fleet [#6 above]*)*.

Roa-Ureta (2012, 2015, 2020) and others (Paya 2009, 2016, McAllister 2004, Robert 2010) have modified the Leslie-Davis method to develop a Generalized Depletion Modeling approach (GDM) that relaxes the closed population and linear catchability assumptions. GDM relaxes the closed population assumption by modifying the population submodel to account for successive perturbations in catch associated with immigration or recruitment into the fishery, or emigration out of the fishery that reset and restart the depletion process. The method has been successfully applied to a catadromous fish stock (glass eels, *Anguilla japonica*) with transient dynamics on fishing grounds that are completely controlled by immigration and emigration (see Lin *et al.* 2017). GDM also modifies the fishery observation submodel to allow nonlinear relationships between fishery catch and effort and fishery catch and the abundance of the vulnerable population (*assumption #3 above*). Multiple fleets in a fishery can be modeled explicitly and multiple years can be analyzed in a Multi-Annual Generalized Depletion modeling (MAGD) framework (Roa-Ureta 2015, Maynou 2015, Maynou *et al.* 2021), or results of intra-annual GDM can be used as inputs to inter-annual stock recruitment models (e.g. Pella-Tomlinson surplus production model; Roa-Ureta 2020, Roa-Ureta *et al.* 2021). The method, originally developed for the assessment of Patagonian longfin squid (*Doryteuthis gahi*), has been successfully applied to data-limited fisheries for spanish mackerel (Roa-Ureta 2015), glass eels (Lin *et al.* 2017), sand eels (Maynou *et al.* 2021), octopus (Roa-Ureta *et al.* 2021) and to evaluate the degree to which the establishment of artificial reefs led to increases in fish production or increased fish aggregation (Roa-Ureta *et al.* 2019). The method can be implemented using the R package *CatDyn (*Fishery Stock Assessment by Catch Dynamics Models CatDyn version 1.1-1 2018-12-18).

More specifically, GDMs describe the true fishery catch in numbers of individuals (C_t) as a function of observed fishery effort (E_t) and the size of the vulnerable portion of the population (N_t) such that

$$5.\ C_t = f(E_t, N_t) = f_E(E_t)f_N(N_t) = kE^{\alpha}_tN^{\beta}_te^{-M/2}$$

where t is the time step, C, E and N are as defined above. Effort (E) is assumed to be observed without error, while population size, N, is unobserved and latent. M is natural mortality per time step and is assumed to be constant over the time steps. Parameters associated with catchability are k, a constant scaler (similar to q where CPUE=C/E=qN); α, an effort response parameter; and β, an abundance response parameter. The effort response (α) modulates the output of catch so that it can be saturable $\alpha < 1$ (fishing gear catches proportionally less with additional effort), proportional $\alpha \sim 1$ (catch is proportional to effort), or synergistic $\alpha > 1$ (additional effort yields a disproportionate increase in catch). The abundance response (β) reflects the degree to which fishers perceive true population abundance where $\beta<1$ indicates hyperstability and a stable catch rate when population abundance declines, while $\beta >1$ indicates hyperdepletion and a catch rate that declines faster than population abundance. Spatial aspects of the fishing process are implicit in the approach.

GDM uses Pope's (1972) recursive equation to make abundance manifest by including N_0 and M. It can also include in-season perturbations of catch abundance $\{P_j\}$ (immigration, recruitment or emigration) that reset and start the depletion process within the fishing season. A multi-fleet (*f*) GDM with abundance perturbations associated with immigration and emigration is specified as follows:

$$6.\ C_t = \sum_f C_{t,f}$$

$$= \sum k_f E^{\alpha f}_{f,t}\, m(N_0 e^{-Mt} - m[\sum^{t-1}_{i=1} C_{f,i} e^{-M(t-i-1)}] + \sum^{j=P}_{j=1} I_j P_{j,f} e^{-M(t-\tau_{j,f})} - \sum^{j=P}_{j=1} J_j P_{j,f} e^{-M(t-v_{j,f})})^{\beta f}\, e^{-M/2},\ v_j > \tau_j$$

where $_f$ indexes fleet, $_j$ indexes abundance perturbations, P is the total number of perturbations. N_0 and M per time step are as described above. $m = \exp(-M/2)$ and is an adjustment that makes all catch occur instantaneously during the middle of the time step. The term in the square brackets in equation 6 accounts for losses due to fishery catch. The summation following losses to the fishery accounts for perturbations associated with immigration events (I) at time steps τ detectable by each fleet. Emigration events (J) have time steps of v for each fleet and are accounted for in the last summation. Note that natural mortality M of emigrants as well as immigrants is accounted for. If no perturbations of emigration are specified are specified and J=0, the GDM resolves to a model with perturbations associated with only with immigration. If no in-season abundance perturbations are specified at all, GDM resolves to a pure depletion model with a closed population assumption, but the possibility of a nonlinear catchability assumption (e.g. equation 5). Expected catch in numbers (C_t) at each time step is assumed to be a random variable with a known distribution. The discrete time equation is formulated so that survivorship is calculated for the middle of each time step using exponential terms.

The nonlinear regression can be solved using maximum likelihood statistical inference in CatDyn for a variety of candidate distributions for the time series of observed catch ($C_{t,f}$;Poisson, negative binomial, as well as normal, lognormal, gamma; Roa-Ureta 2020). The package optimx (Nash and Varadhan 2011) is called for the optimization. The parameters M and N0 are common to all fleets in a fishery. Fleet specific parameters include observation (~ fishery catchability) parameters kf, αf, and βf as well as perturbations to abundance Pf and their timings (τf,vf). It is assumed that fleet effects that are additive and provide complementary information about the fish population. Competition between fleets is not accounted for. Fishing mortality per time step is also estimated from abundance, natural mortality and catch using a numerical resolution of the Baranov equation (Roa-Ureta 2020). In a complex GDM with multiple fleets and perturbations of abundance due to immigration or emigration the number of free parameters is 2(M and N0) + 3*f(kf, αf, βf) + 2*P*f (the magnitude and timing of abundance perturbations that can be fleet specific). It is important to note that the method requires at least ~ 3 times more data than the number of parameters, for parameters to be estimated accurately and with sufficient certainty

(Roa-Ureta 2012, 2015, 2020). A multi-fleet GDM with abundance perturbations as specified above requires data with a short time step and high temporal resolution.

To develop GDM based hypothesis related to immigration, recruitment, or emigration during the fishing season Roa-Ureta (2012) developed a fleet specific catch spike statistic (S_t) using fishery dependent data that accounts for spikes in catch that are independent of by variations in fishing effort (E_t). Lin *et al.* (2017) developed a complementary parametric catch spike and used fishery independent information about environmental drivers of glass eel migration to develop alternative start values for the timing of abundance perturbations. Calculation of the parametric and nonparametric spike statistics are described in the methods section. Spikes in catch can result when animals migrate onto or off of fishing grounds or when fishing fleets shift to more productive fishing grounds or when forced to poorer grounds by fishery regulation (Roa-Ureta 2015). It is therefore important to inspect the spatial dynamics of fishing fleets to understand whether catch perturbations are the result of movements of animals or of fishing fleets.

While GDM relaxes the closed population assumption (# 1); the linear catchability assumption (#3), and allows greater flexibility in the choice of error distributions, several assumptions are made about the migration process when the method is applied to fisheries for fully transient stocks on fishing grounds (e.g. Lin *et al.* 2017).

These include:

8) immigrants mix well with and have the same catchability as the stock on the fishing ground, and

9) abundance added by an immigration pulse at time τ_j is removed from fishing ground when it emigrates at time v_j.

It is important to note that like other catch-only methods, GDM assumes inferences about stock characteristics can be made from fishery catch (Ovando *et al.* 2021). GDMs continue to rely on assumptions #2, #4, #5, #6 listed above. Assumption #4, which posits that the stock does not have refuge from the fishery, is important to consider when making inferences about the *Illex illecebrosus* population in the Northwest Atlantic using GDM. The directed U.S. fishery is a seasonal fishery that operates from May through October on the outer edge of mid Atlantic Bight continental shelf, from Oregon Inlet, North Carolina to Welkers Canyon off southeast Massachusetts. *I. illecebrosus* is believed to produce approximately 4 overlapping cohorts throughout the year and we believe the directed U.S. fishery harvests 1-2 of these cohorts. In the western Atlantic, *Illex illecebrosus* ranges from the Florida Straits northeast to Labrador, the Flemish cap, Baffin Island and Southern Greenland (Trites 1983, Dawe and Beck 1985, Jereb and Roper 2010), so only a small part of the species range is vulnerable to the fishery (Manderson *et al.* 2022). Market forces, technical aspects of processing squid and fisheries regulations also strongly influence fishing effort and catch in US Illex fishery (see Mercer *et al.* 2022).

If, as Rago (2020) suggests, the principle cause of the failure of classical depletion analysis when applied to the U.S. *Illex illecebrosus* fishery is the violation of the closed population assumption (assumption #1 above*),* the GDM approach may more successfully applied to *Illex* to provide information useful for assessing the risk of overfishing for the stock. In this work the application of GDM is evaluated using the R library CatDyn (Fishery Stock Assessment by Catch Dynamics Models v1-1.1) to five years of landings, effort, and individual squid weight data collected in the U.S. *Illex illecebrosus* fishery, aggregated to a weekly time step. Intra-annual models were developed for 2 years (2012, 2016) in which fishery performance was evaluated as poor and three years (2012, 2018, 2019) in which performance was evaluated as good, based on statistical and qualitative industry based assessment of fishery performance (Mercer *et al.* 2022). The steps needed to develop a "best" GDM with an open population assumption for a single year are described and the precision of parameter estimates are presented, and quantities of interest to fishery assessment are calculated. The precision and accuracy of parameter estimates and quantities produced by "best" GDMs are developed for the 5 years. Finally, recommendations are provided for the next steps required to determine whether GDM can be operationalized in the assessment, including in-season assessment, of the risk of overfishing in the U.S. *Illex illecebrosus* fishery.

Fishery Data

Landings and effort data

To assessed the utility of generalized depletion modeling to produce information useful for assessment of the risk of overfishing in the Illex fishery we analyzed 5 years of weekly landings data including years of poor (2013, 2016) and good fishery performance (2017, 2018, 2019) as described in (Mercer *et al.* 2022).

The cumulative landing biomass of squid in kilograms was calculated for ISO 8601 standardized weeks (Monday-Sunday). Several metrics of fishing effort were developed from vessel trip reports (VTR) including number of days fishing (DF), number days absent (DA), and numbers of unique fishing permits responsible for weekly landings. Days absent (DA) is the difference between date-time landed and date-time sailed and includes time steaming between ports and fishing grounds and between fishing grounds. Days fishing (DF) eliminates steaming times and was used here as the preferred effort metric. Effort metrics could only be developed for landings for which there was a 1-to-1 match in dealer reports and VTR records. Fishing effort was curtailed by fishing regulations and seasons closed early in 2017 (09-17; week 37), 2018 (08-14; week 33) and 2019 (08-21; week 34).

We partitioned weekly landings and effort amongst two fleets; vessels that process and freeze squid at sea ("freezer trawlers") and vessels that store squid in recirculating seawater systems (RSW) or on ice and deliver them to shore side plants for processing and freezing onshore (hereafter called "wet boats"). We developed two fleet GDM models for the following reasons. a) Freezer trawlers that process squid at sea can search larger areas for longer times (as long as 14 days) but have longer prey handling times than wet vessels that catch squid and must transport them quickly to nearby shoreside processing plants before the squid spoil (usually within 72 hours; see Mercer et al 2020 for more details). As a result, we expected the scaling, effort and abundance response parameters (k, α, β) as well as timings of detection and magnitudes of pulses of ingressing or egressing squid catch to be different in the two fleets. b) An examination of historical fishery data indicates the ratio of fishing effort in the two fleets fluctuates with fishery performance. In the years landings were low freezer trawlers dominated the fishery and there were sometimes much fewer wet boats. During years of "good" fishery performance wet boats are dominant while many of the freezer trawlers continue to operate in the fishery (Table 5.13). Since 2016, fishery performance has been "exceptional" and some freezer trawlers have been converted to wet boats.

Fishery weigh-out data

Depletion analyses require landed biomass to be converted to number of individuals captured in the fishery for the analysis. We used weights of squid in grams collected for the purposes of inventory and marketing by the two primary processor/dealers in the fishery: Lunds Fisheries, Cape May, New Jersey and Seafreeze Ltd. Davisville, Rhode Island. Whole body weight samples (g) are routinely collected from every trip because the product is marketed by body weight size category. Seafreeze Ltd. primarily inventories squid caught on fishing grounds both north and south of the Hudson Shelf valley that have been frozen at sea on freezer trawlers.
Squid from 69 to 100% of trips are thawed and measured in a given year and sample sizes of freezer trawler body weights are typically very high.

Generally landings from fishing grounds south of the Hudson Shelf Valley (~39.5 N) are delivered to Lunds Fisheries which primarily accepts unculled squid for shoreside processing and freezing from wet boats. Weight sampling of the wet boat fleet is less comprehensive. The weight data were provided with the date of measurement but without attribution to fishing ground or the type of vessel responsible for the landing. These data were used in the evaluation of the GDM. The operationalization of the method will require squid weight data that are more representative the fleets, fishing grounds and the time step of analysis.

Greater than 10,000 weights were available for each year and a median of 1690 measurements were available during most weeks of the fishing seasons (5% and 95% quantiles =100 and 4573 wk^{-1}). To develop mean weights of squid (grams) for each week, local polynomial regression (loess) with R defaults (span=0.75; fitting by weighted least squares) was applied to describe the relationship between the weight of squid in grams and day of the year within each fishing year for the fishery as a whole. The loess regressions were then used to predict mean weights (and standard errors) of squid for the day falling in the middle of each week for which landings were reported. This approach has been applied in other studies developing GDMs for data limited fisheries (e.g. Roa-Ureta 2015, Mayou 2021).

GDM Modeling

Strategy for intra-annual generalized depletion model development using CatDyn

The GDM was applied to catch data for the two fishing fleets independently for each fishing year. It is possible to develop Multi-Annual Generalized Depletion models (MAGD; see Maynou *et al.* 2021) but they assume levels of strengths of stock-recruitment relationships and inter-annual autocorrelation in natural mortality not justified for *Illex illecebrosus*. It is also possible to use the results of intra-annual generalized depletion models to inform assessment models in a hierarchical manner and more formally develop biological reference points (Roa-

Ureta 2020, Roa-Ureta *et al.* 2021). This type of hierarchical analysis is beyond the scope of this work. Here, the potential utility of the GDM for the assessment the Illex fishery is evaluated by developing intra-annual GDMs using data with a weekly time step for recent years classified as good (2017-2019) and poor fishing years (2013, 2016).

Intra-annual GDMs are developed in three steps: 1) the development and selection of model variants for hypothesis about movements of squid onto and off of fishing grounds during the fishing season beginning with a pure depletion model (Null Model) that assumes the fishery is closed, 2) the selection of the "best" hypotheses about in-season movements of squid from the suite of hypotheses developed for the specific fishing season, and finally 3) the use of the "best" model variant reflecting the "best" hypotheses to develop parameter estimates and derived quantities useful for stock assessment in a given year.

Step 1. Development and selection of GDM variants for closed and open fishery hypotheses

Generalized depletion models representing closed and open fisheries were developed for each fishing season using 12 different likelihoods assumptions. Eight model variants applied the same likelihood and distribution assumption to freezer trawler and wet boat fleets [normal (n), adjusted profile normal (apn), lognormal (ln), adjusted profile lognormal (apln), poisson (p), negative binomial (nb), and gamma (g)]. Four additional models were developed for each model hypothesis that applied different likelihoods to the two fleets (freezer trawler: wet boat fleet, apn:apln; apln:apn; n:ln; ln:n). Each of these 12 "likelihood" model variants were fit using four different numerical optimization methods ("spg", "CG", "Nelder-Mead", "BFGS"). Thus, 48 possible model variants were produced for each model hypothesis in each season.

Reasonable initial starting values for parameters are required for fitting GDM model variants to data using maximum likelihood through calls from CatDyn to the optimx function in the R package optimx. Initial values for M, N_0 as well as the fleet specific parameters k, α, β are required for fitting pure depletion models (H0). Model variants with an open population assumption also require starting values for the timing and magnitudes of in-season pulses of squid. Starting parameters were refined before model fitting based on the visual inspection plots generated with the *CatDynExp* function that allows visually exploration of initial parameter values. We selected starting values that minimized the difference between observed and predicted catch, and temporal trends in deviance residuals. The ranges for starting parameter values are reported in the following text table.

Ranges for initial starting values for fitting model variants that were refined using the CatDyn exploratory tools and plots of observed and predicted catch, and temporal trends in deviance residuals.		
Initial Starting Parameter	Value or equation	Logic
M (weekly)	0.01-0.15	Literature values from Roa-Ureta & Arkhipkin 2007; Arkhipkin *et al.* 2021b; Hendrickson 2004, Hendrickson & Hart, 2006; Hoenig 2005 as described in text below. Adjusted on the basis of weekly trends in observed vs predicted catch. Lower threshold relaxed in null models or those missing pulses evident in catch perturbation analysis.
$N_{0,a}$ *(M unaccounted for)*	Total Annual Catch * 1/0.25 pure depletion models Total Annual Catch * 1/0.33 when pulses evident	Industry estimates of industry net efficiency ~0.25 (Mercer *et al.* 2022; Manderson *et al.* 2022) Reduced when vulnerable population open and in-season pulses are probable
$N_{0,b}$ *(M accounted for)*	Total Annual Catch * 1/0.25 + (Total Catch * 1/0.25) * 1-exp(-M*number of fishing weeks)/2 As above but with M accounted for when pulses evident	Account for net efficiency as well as losses due to natural mortality over ½ of fishing season
Fleet specific start values		
k_f	1/total days fished by fleet /n	n ~100-350. Often different for the fleets
α_f Effort Response	0.8-1	Assume close to linearity
β_f Abundance Response	1-1.3	Assume close to linearity
P_f (Sign, timing)	Determined on the basis of catch perturbation analysis described in text below	Assumes that inferences about in-season immigration/emigration can be developed soley on the basis of rom perturbation analysis of fishery dependent data.
$P1_f$ magnitude	¼ of N_0 * proportion of cumulative catch caught by fleet	Assumes that size of incoming pulses is smaller than the size of the vulnerable fraction before the fishery began. This assumption is probably not valid.
multiplicative dispersion parameter	0.25*sd(log(total catch))^2	When the likelihood is lognormal, negative binomial, gamma
additive dispersion parameter	0.25*sd(total catch)^2	When the likelihood is the normal distribution

Probabilistic model selection criterion (AIC, BIC, MDL) can only be used to compare models that assume the same likelihood and numerical optimization method. As a result, GDM variants of model hypotheses were selected using a set of criteria associated with statistical properties and biological realism. Statistical criteria included the selection of converged models that produced numerical gradients for parameter estimates < |1| (Thorson *et al.* 2015, Roa-Ureta *et al.* 2021). Remaining models were compared with respect to their ability to produce asymptotic standard errors, %CVs for parameter estimate, and multi-collinearity among parameter estimates.

Models (usually ≤ 4) that produced standard errors for most estimates, relatively small standard error/ parameter ratios, and relatively small parameter inter-correlations were compared with respect to biological realism criteria, placing greatest weight on natural mortality, M wk^{-1} estimates. Unrealistic values of M were identified based on the literature (Roa-Ureta & Arkhipkin 2007, Arkhipkin *et al.* 2021b, Hendrickson 2004, Hendrickson & Hart 2006, Hoenig 2005). Weekly M estimates generally exceed 0.01 wk^{-1} for Ommastrepid squid. It has recently been shown with GDM that natural mortality for *Illex argentinus* (Argentinian shortfin squid) and *Doryteuthis gahi* (Patagonian longfin squid) in the Falklands island is ~ 0.092 wk^{-1} (Roa- Ureta & Arkhipkin 2007, Arkhipkin *et al.* 2021a). Hendrickson (2004) reported a maximum age of 215 days for squid captured on US fishing grounds. For animals with this max age, the Hoenig (2005) age based equation predicts M wk^{-1} ~ 0.134 (SE=0.017). Therefore, models producing M wk^{-1} estimates between 0.01 and 0.15 were considered plausible. A relatively conservative approach was adopted that did not include the higher M wk^{-1} values estimated for mature squid by Hendrickson and Hart (2006). M wk^{-1} estimates > 0.1 were rarely estimated in the more than 750 GDM model variants examined in this work. M wk^{-1} estimates <0.01 were often estimated in pure depletion model variants and variants that did not include in-season pulses evident in catch perturbation analysis. The lower bound to M wk^{-1} was relaxed under these circumstances.

Generalized depletion modeling is an exercise in multi-model inference; the selection and evaluation of hypotheses about the timing and magnitudes in-season pulses of animals onto and off of fishing grounds. Typically these hypotheses are developed from catch perturbation analysis of fishery dependent data sources because fishery independent data describing stock movements are unavailable. This is indeed the case with *Illex illecebrosus* in U.S. waters. The perturbation analyses conducted in this work summarized 4 lines of evidence:

1) Development of pure depletion models (H0) that assumed the fraction of population vulnerable to the fishery was closed to immigration or emigration during the season; identified weeks when observed catch was higher or lower than predicted catch and tallied these residuals (e.g. Figure 5.17).

2) Use of fleet specific catch nonparametric and parametric spike statistics to identify weeks when catches in the fleets were disproportionally high or low when compared to fishing effort (e.g. Figure 5.18). The nonparametric spike statistic $S_{l,t}$, internally generated in CatDyn, was developed by Roa-Ureta (2012) to identify spikes in observed catch (X_t) unexplained by variations in fishing effort (E_t). Such that

$$7.\quad S_{l,t} = 10(X_{l,t}/\max(X_{l,t}) - E_{l,t}/\max(E_{l,t}))$$

We also computed the complementary parametric residual catch spike statistic developed by Lin *et al.* (2017) that uses plots of residuals from the regression

$$8.\quad \log (X_t) = A + \alpha \log (E_t)$$

where the intercept,

$$9.\ A = \beta \log (N_t) - M/2$$

3) We examined plots of gridded weight frequencies of squid landed in the fishery to identify weeks of the season when weight classes of squid may have entered or exited the fishery (e.g. Figure 5.19)

4) Finally, to determine if pulses of squid evident in statistics above reflected the in-season movements of squid into or out of the fishery or movements of the fleet to different fishing grounds, we examined spatially explicit plots of the relative magnitudes of catches of squid reported by NOAA observers, and self-reported by fishers on vessel trip reports VTR and the NOAA study fleet program (e.g. Figure 5.20).

These lines of evidence were entered in a perturbation summary table (e.g. Table 5.14) and synthesized into a set of alternative hypotheses about the timings of in-season movements of squid onto or off of the fishing ground. Hypotheses were only developed for timings when two more indicators recorded in the perturbation summary table coincided in time.

Model variants were developed and selected for each open population hypothesis using methods described above. The "best" model variant (distribution assumption, numerical method) for each hypothesis was selected using the statistical and biological realism criteria described above.

Step 2) Selection of the "best" hypotheses about in-season movements of squid

We used statistical and biological realism criteria described above as well as the protocol of Lin *et al.* (2017) who used the Akaike information criterion (AIC) to objectively select the "best" hypothesis for a fishing season. This required selecting likelihood and numerical optimization algorithms resulting in convergence for all model hypotheses including the pure depletion model (H_0). In most cases several combinations of likelihood and numerical method could be found, but this was a compromise since one or more of the model hypothesis failed to meet the criteria described in the previous section. We selected the hypothesis with the lowest AIC, from a consensus of the likelihood and numerical method combination that could be used to compare the hypotheses. Once the "best" hypothesis was selected, the "best" model variant meeting statistical and biological realism criteria in step 1 representing "best" hypothesis was used in step 3, to develop final parameter estimates and derived quantities useful for stock assessment.

Step 3) Develop parameter estimates and derived quantities useful for stock assessment using "best" model variant reflecting the selected hypothesis

The best model variant (step 1) reflecting the selected hypothesis (step 2) was used to produce final estimates of fishing and population parameters, including N_0, M wk^{-1} , fishing mortality (F wk^{-1}), exploitation rate, the magnitude and timing of in season pulses of squid into the fishery and escapement biomass at the end a fishing season. All of these quantities are computed within CatDyn software or can be developed from computed quantities.

The results section proceeds as follows. We describe the development of a "best" general depletion model with an open population assumption for a single the fishing year of 2016. We selected 2016 because the final model was relatively well behaved with respect to the statistical and biological realism criterion described above. Models were constructed in the manner described in all years. We then discuss parameter estimates and their uncertainties from the "best" GDMs developed for all 5 years before presenting quantities that could be of use in assessments. We discuss the potential role of model time step and sample size in determining parameter uncertainties and suggest next steps that may lead to use of generalized depletion models in the operational assessment, including in-season assessment if necessary, of the risk of overfishing in the US *Illex illecebrosus* fishery.

GDM Model Results

Development of GDM with open population assumption for the 2016 fishing year

Statistical and qualitative industry based assessments of fishery performance indicated that the 2016 fishing year was relatively poor (Table 5.13). The fishery landed 7,004,000 kg of squid during the season. Overall fishing effort was relatively low and 12 vessels were responsible for the landings (J. Didden personal communication). Only 5 of vessels landed more 226 mt. While the freezer trawler fleet accounted for 68% of total landings and 56% of total landings, no effort was recorded for freezer trawlers in 6 of the 19 weeks of the season. This probably occurred because the duration of most freezer trawler trips was longer than 7 days. Nevertheless since there were 19 weeks in the season, 38 datum were available to estimate the population parameters N_0 and M, while 19 datum were available to estimate fleet specific parameters (k, α, β) and the timings (τ, or v) and magnitudes of perturbations (P_j) associated with in-season ingress or egress of squid.

Seven of 48 pure depletion model variants converged, had numerical gradients less than 1, and 2 or fewer incalculable standard errors (Table 5.15a). Natural mortality estimates (M wk 10^{-3} – 10^{-7}) were much lower than the lower bound indicating that squid may have moved onto the fishing grounds during the fishing season. The pure depletion model variant with the smallest CVs (Table 5.15b, others not shown) and parameter correlations (Figure 5.21) assumed a normal distribution for freezer trawler catches, a lognormal distribution for "wet boats" and was optimized with the CG algorithm.

In the catch perturbation summary table (Table 5.14) we recorded residuals from the time series plots of observed vs predicted catch for the freezer trawler and wet boat fleets (Figure 5.17) developed from the "best" pure depletion model. We also tallied high positive and low negative values for nonparametric and parametric catch spike statistics (Figure 5.18), weeks of appearance for weight classes of squid (Figure 5.19), along with the weekly positions where the majority of catches were relative to the Hudson Canyon (Figure 5.21) in the catch perturbation summary table used to develop hypothesis about in-season ingress or egress of squid for the 2016 season.

Using the multiple lines of evidence in the 2016 catch perturbation summary table (Table 5.14) we framed 3 hypothesis for the development and evaluation of GDM model variants with an open population assumptions. Hypothesis H1, posited ingress of squid into the fishery detected by the wet boat fleet landing in week 33 and in the freezer trawler fleet in week 34. Hypothesis H1 was supported by weight frequencies, residuals of the pure depletion model, and spike statistics in the freezer trawler fleet, and spike statistics in wet boat fleet. Hypothesis H2 added a second pulse of ingress detectible in landings in in weeks 37 (freezer trawlers) and 38 (wet boats). H2 was supported by the parametric and non-parametric catch spike statistics.
Hypothesis H3 added early season egress in both fleets to Hypothesis H2a that was a post hoc modified the H2 inference. H3 posited that the second pulse of ingress was only detectible in the wet boat fleet. Some model variants with input perturbation parameters reflecting H1, H2 and H2b converged when fit to the data using maximum likelihood. Model variants for H3 did not converge. H3 had 9 fleet specific parameters for the wet boat fleet with a parameter to data ratio of 0.67 (9/19; 7 parameters in the freezer trawler fleet) and a total of 18 parameters and parameter to data ratio of 0.47 (18/38).

Model variants best reflecting H1 and H2 and H2b were selected on the basis of statistical and biological realism criteria described in the methods and demonstrated for the pure depletion model H0 above. For the sake of brevity only parameter estimates of “best” model variants reflecting the 4 (H0-H2b) hypotheses are presented here (Table 5.15b-d). The H1 model with a single ingress of squid into the fishery produced an M falling within the bounds of biological realism criteria and %CVs for parameter estimates that were less than 57% except for the catch-ability scaler for wet boats (k CV=168%) and the magnitude of the relatively of ingresses into the freezer trawler fleet (CV >5000%, 4c) which was 0.24% of the magnitude of the pulse detected in the wet boat fleet. Parameters associated with the fishing process (k, α, β) for the H1 model variant appear unrealistic. Developing H2 by adding the second in-season pulse of squid ingress to both fleets increased CVs for nearly all the parameters including M which fell below the lower bound unless the error is considered (Table 5.15d). Model variants for H2b with the first ingress event detectible in both fleets and the second detectible only in the wet boat fleet were developed in response to statistical properties of the best H2 variant and further inspection of the perturbation summery (Table 5.14). This scenario was supported by more lines of evidence in catch perturbation analysis, including weight frequencies (Table 5.14). M Fell within the bounds and parameter CVs for the best H2b variant were somewhat smaller than for the H2 variant (Table 5.15e). However fewer standard errors were produced in the H2b variant. Parameter estimate correlations were also higher for the best H2b variant than the H2 variant.

The H1 model variant (apln apln BFGS) appears to best explain the weekly fishery dependent data available for the 2016 season compared with the other hypothesis variants based on most of the statistical and biological realism criterion. This is supported by the AIC comparison of hypotheses H0-H2b made by holding constant distribution assumptions and numerical operationalization method constant for all the hypothesis developed for the 2016 fishing season (Table 5.16).

The “best” H1 model variant had a single squid ingress event detectable by the freezer trawler fleet in week 35 and in wet boat fleet in week 33 (Figure 5.23). Deviance residuals did not show a temporal trend for either fleet. However plots of predicted vs observed catch indicate that the wet boat fleet was fit better than the freezer trawler fleet that did not land trips in 6 of the 19 weeks of the fishing season.

Quantities derived from the best H1 GDM variant in 2016 and useful for assessing the risk of overfishing are shown in Figure 5.24. The weekly natural mortality estimate was 0.026 (SE=0.015). The median weekly F for the fishery was 7.75E-5 (2.5 and 97.5 quantiles: 1.00E-09, 4.22E-04) with a cumulative F over the season of ~0.0023 (Table 5.20, Figure 5.24a). The median observed exploitation rate (F/Z) for the fleet was estimated to be 8.96E-07 (7.46E-08, 4.87E-06; Table 5.20, Figure 5.24b). Cumulative F/M was ~0.005 well below the value of 0.667 proposed by Patterson 1992 as biological reference point appropriate for small pelagic fish (Table 5.20, Figure 5.24c). The GDM model predicted escapement biomass at the end of the fishing season to be 3,889,492 metric tons (Table 5.0, Figure 5.24d). This value fell at the 96th percentile of values developed using the Rago (2021) Indirect Estimation Methods approach to developing plausible bounds using spring and fall NEFSC trawl surveys on the continental shelf (Figure 5.25).

Caution is warranted in interpreting some of these quantities given the uncertainties for some parameter estimates, specifically those related to the fishing process and the small in-season pulse of squid in 2016. However, the small Fs and large escapement value are not out of the realm of possibility given the small size of the fleet and its fishing effort.

Review of modeling results for final models for all years 2013, 2016-2019 fishing years

Hypothesis selection

Models that included in-season pulses of squid into the fishery and an open population assumption generally better explained Illex fishery catches than pure depletion models with a closed population assumption (Table 5.16). In 2 years (2016, 2017), models with a single pulse of squid ingress detected in landings by both fishing fleets had the lowest AICs. In 2019, the “best” model had 1 in-season pulse into both fleets followed by a second positive pulse detectible in the wet boat fleet. In 2018 a pure depletion model with a closed population assumption had the lowest AIC; 2018 had the shortest fishing season (12 weeks), the smallest sample size, and AICs for the three 2018 model hypotheses developed were very similar. There was not strong evidence for an in-season pulse of ingress in the 2018 perturbation summary (not shown). Furthermore, the pure depletion model variant that passed statistical criteria produced reasonable estimates of M wk^{-1} (0.045-0.093). In the seasons for which open population assumption GDMs better explained catches, pure depletion model variants with a closed assumption produced mortality estimates orders of magnitude lower than the plausible bounds (0.01-0.15).

There was no evidence for significant in-season pulses of squid emigration from the fishery in the week fishery landings data analyzed in this work. There were no model variants meeting statistical criteria that produced M wk^{-1} estimates greater than 0.1. High values for M wk^{-1}

would be expected for model variants that did account for substantial pulses of in-season emigration. However, weight frequency plots suggested that in some years larger squid leave the fishery by emigration or mortality (e.g. Figure 5.19). Fishers also observe squid moving off of continental shelf fishing grounds response to weather and oceanographic events (Mercer *et al.* 2022). In addition *Illex* are known to move off the continental shelf in a sex specific manner during the fall (O'Dor & Dawe 2013).

We believe the hypotheses about the number in-season pulses of ingress were limited by the weekly time step since ratio of the number of parameters to be estimated to the amount of data was high and increased for open population models. Pure depletion GDMs with 2 fleets (P0P0) that assumed the fishery "closed" during the season estimated eight parameters (N0, M wk^{-1} as well as k, α, β for each of the two fleets; Table 5 & 10). The timings and magnitudes of each in-seasons pulse of squid into or out of the fishery are detected in the landings in a fleet specific manner. Four additional parameters are estimated when a single in season pulse is detected in landings in both fleets for a total of 12 parameters. Sixteen parameters are estimated in models with two perturbations in each fleet. The longest season (2016) produced only 19 data points per fleet. Even in that year the best GDM with a single pulse in each fleet produced a parameter to data ratio of 0.316. Generalized depletion modeling requires a parameter to data ratio less than 0.333 (Roa-Ureta 2012, 2015, 2020). Sample sizes in the weekly Illex fishery data fell as low as 24 in 2018 when the pure depletion model best explained the data.

Models best representing supported hypotheses

GDM variants best representing the hypotheses and chosen on the basis of numerical quality, statistical quality, and biological realism criteria described in step 2 were fit with using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) numerical optimization algorithm. The best variants assumed a variety of data distributions, but with the same distributions applied to both fleets (Table 5.17). The normal distribution best explained the data for both fleets in two years (2017, 2019) while adjusted profile lognormal, negative binomial, and gamma models were also represented.

Parameters describing the fishing process (k, α, β) were extremely variable and difficult to estimate with GDM using data with a weekly time step (Table 5.17). Asymptotic standard errors could not be estimated for all fishing parameters except for 2016 that had the largest sample size and 2018 when the pure depletion model best explained the available data.

Estimates for the scaling parameter k (~q) varied by 13 orders of magnitude for freezer trawlers and 11 orders of magnitude for the wet boat fleet (Table 5.17). The standard errors for k could not be computed for freezer trawlers in 2017 and 2019 and for wet boats in 2013 and 2019. The pure depletion model for the 2018 season produced a k that had particularly low precision (CV=4524%).

Estimates for the effort response parameter were variable for both fleets; saturable ($\alpha < 1$, gear catches proportionally less with additional effort) in some years and synergistic ($\alpha > 1$, additional effort yields a disproportionate increase in catch) in others (Table 5.17). While patterns were not consistent among the fleets, asymptotic standard errors were produced in all cases and CVs for α estimates were < 45% for freezer trawlers and < 64% for wet boats.

Estimates for the abundance response parameter were also variable, standard errors frequently inestimable, and when they were estimable, they were often quite high (Table 5.17). β estimates for the wet boat fleet were less than <1 in 4 of the years (2016-2019; hyperstability: stable catch rate when the vulnerable fraction of the population changes). However, standard errors could not be produced in 2 years (2013, 2019) and the CV for β was 256% in 2018. β estimates for the freezer trawler fleet were greater than 1 in 3 years (2016-2018, β >1, hyperdepletion: catch rate changes faster than abundance of the vulnerable fraction of the population) less than 1 in 2 years. In 2013 the freezer trawler β estimate was 0.005 and the CV, 1500%.

GDM makes inferences about in-season pulses of squid into or out of fishing areas based on their detectability in landings made by the individual fleets. In a two fleet model half the total sample size is used to estimate the timings and magnitudes of pulses along with the catch-ability parameters. Like catchability parameter estimates, the magnitudes of in-season pulses of squid were very imprecise; probably again a result of the small sample sizes (Table 5.18). The CVs for pulses were greater than 3000% for pulses in the freezer trawler fleet which were more than 400 times smaller than estimates in the wet boat fleet. However, the CVs for pulse magnitudes also exceeded 400% for the wet boat fleet except in one case (2016, CV=22%).

Natural mortality (M wk^{-1}) and the numerical abundance (N_0* 1000) of the vulnerable fraction of the population before the fishing season appeared to have been better estimated with the GDMs (Table 5.19). This is to be expected because the entire datasets (2 fleets * length of the season) are used to estimate population parameters. M wk^{-1} ranged from 0.26 (SE=0.015) to 0.97 (SE=0.012) and CVs were less than 58% in 3 seasons (2016, 2017, 2018). CVs for M wk^{-1} were greater than 100% in 2013 and 2019. It is also important to acknowledge that M wk^{-1} served as a diagnostic criteria for biological realism in our model selection process. However, M wk^{-1} fell into a range between 0.02 and 0.1 in all model variants meeting statistical criteria if in-season pulses evident in the perturbation summaries had been specified. Error estimates for N_0 were produced in all years except 2018. CVs for N_0 were greater than 400% in 2013 and 2019; CVs were less than 75% 2016 and 2017.

Derived quantities of interest for assessing the risk of overfishing and sample sizes

Quantities derived from the final GDM useful for assessing the risk of overfishing (M, F, Exploitation rate, F/M, Escapement biomass) are reported in Table 5.20. Weekly Fs generally fell below 0.027 and observed exploitation rates fell below 3.25E-04 in all years. F/M ratios ranged from 3.89E-08 to 4.08E-01, with seasonal means ranging from 0.002 to 0.115. Finally, in all years except 2013, escapement biomass predictions from the GDMs exceeded 1,000,000 metric tons. These values fell near the upper bounds (>85th percentiles) of escapement estimates develop by Rago (2022) using fishery independent survey estimates and ranges plausible values for survey gear efficiency q and availability v and ratios of F/M.

The uncertainty of the parameter estimates is too large to make a reasonable conclusion about the risk of overfishing for every year analyzed in this work except 2016. Even for 2016, the uncertainty of the estimated fishing process parameters is large. The overall lack of precision in this analysis might be eliminated by moving from a weekly to a daily time step and increasing sample sizes. Moving to a daily time step would reduce the ratio of parameters to data in the models specified here from an average of 38% to %11 (Table 5.21). This could dramatically increase the precision of fleet specific parameters associated with the fishing process and in-season pulses of squid. However, it is also the case that catch data with a daily time step will be noisier due to the influence of nuisance variables (holidays, stormy weather, etc.).

Conclusion

Generalized depletion modeling appears to be a very useful technique providing historical and in-season assessments of the risk of over fishing for stocks that are transient on fishing grounds during the fishing grounds. The technique is used to estimate the abundance of the fraction of the population vulnerable to the fishery (No), natural mortality (M), fishing mortality (F), and fishery escapement (H) in a manner that accounts for in-season migration that confounds the estimation of those parameters. While we believe that some parameters had reasonable precision in some years, overall the GDM analysis indicates that a weekly time step is not sufficient in the U.S. *Illex* fishery to support the precision required for operational stock assessment. This is very much the case during years when the fishing is good and the fishery is closed early upon achieving the quota. We recommend, in order of priority, 4 steps that could allow generalized depletion modeling to be operationalized for the assessment of the U.S. *Illex* fishery in the future:

1) Use analysis of existing landings and weight data along with data simulation to determine the effects of moving to a daily time step on the accuracy and precision of parameter estimates developed using GDM. In the simulations, test the effects of variations in data quality and quantity on parameter estimate accuracy and precision under different and realistic scenarios of in-season ingress and egress of squid onto fishing grounds.

2) Refine existing catch and landings reporting programs to meet requirements determined in #1.

3) Determine the requisite frequency of sampling of squid weights representative of the fishery, its fleets and its fishing grounds required to estimate mean daily weights of squid (and errors) accurately at the time step required for GDM. Co-create with the fishing industry a sampling program based on those findings that is mutually beneficial to fishing businesses as well as fisheries science.

4) Develop a research program investigating the in-season movements and migration of squid through fishing grounds including roles played by oceanographic processes and variations in structure. Develop from this research fishery independent indicators that can supplement fishery catch perturbation analysis used to inform plausible open population hypotheses for GDM.

WG CONCLUSION ON THE GDM

The WG believes GDM approach is promising but requires further research. The GDM results suggest in a qualitative way that F was lower than M (from internal GDM F to M ratios results) and that stock biomass was lightly fished in 2019 (from comparison of the estimated range of annual biomass to the Rago (2021) Mass Balance bounds). The WG concludes that the GDM (as currently configured with weekly fishery landings data) does not provide an adequate quantitative basis for stock status determination using any of the candidate BRPs, including Mass Balance bounds, F to M ratios, or previously published estimates of biological reference points for the stock (i.e., Hendrickson and Hart 2006).