

スケトウダラ太平洋系群ピア・レビュー報告書

庄野 宏（広島工業大学・工学部）2022/1/5(水)

1. はじめに

スケトウダラ太平洋系群では、年別年齢別漁獲尾数データを使用し、沖合底びき網漁業の年齢別標準化 CPUE に基づくチューニング VPA（近年の若齢魚の漁獲係数 F に制約を加えたリッジ VPA）を利用して資源評価を実施し、資源評価結果に基づく広範なシナリオによる長期間の将来予測を実施している。

資源評価や将来予測に関する手法の選択は概ね妥当と考えられ、推定精度は高いと思われる。その上で、私が事前に提出した質問票（コメントを含む）に対する回答、2021年12月8～9日に水産資源研究所（横浜市）にて開催されたピア・レビュー委員会での資源評価担当者によるプレゼンテーションおよび関連する質疑応答などに基づき、今後の改良点も含めた検討事項について、項目毎に分けて簡潔に記述したい。

2. 資源評価に用いられたデータおよび生物学的パラメータについて

今回使用したスケトウダラの成長曲線や成熟率について、知見が少し古いように感じられるため、最近のデータ収集に基づき適当な時期に改良を試みていただきたい。

自然死亡係数 M については一般に推定精度が低いため、M の値を変化させた場合の感度分析に基づきチューニング VPA による推定された年別年齢別資源尾数のトレンドおよび絶対値の差異をチェックすることに注力してみても良いのではないだろうか。

本筋とは関係ないが、年別年齢別漁獲尾数データのコピー&ペーストの誤りの説明があった。このような集計作業については、機械とは異なり人間のミスが起こりうるという前提で考えていただきたい。具体的には、ログ（記録）が残るプログラミング言語などを利用するか、MS-EXCEL を使用する場合でも、マクロの記録などの機能を利用して後から見直すことにより、多少なりともミスを減らせるのではないだろうか。

3. 資源評価に用いられた前提条件および手法の妥当性について

(1). 資源量指数（沖底 CPUE の標準化）について

Delta 型 2 段階モデルの 2nd-step で船 ID を説明要因として組み込んでいる。船 ID は漁獲量や操業位置、操業時期など多くの説明変数の要素を含むため取り扱いには十分注意し、VIF などに代表される多重共線性のチェックを欠かさずに行うべきである。

通常は有意になることが多い Year*Area の交互作用を欠損の関係で含められなかったこと、固定効果で組み込む場合はエリア分けを変えて（i.e. 各々のサブエリアを大きく区切り）欠損をなくす方法もあるが、将来的には変量効果も検討すべきである。

漁法を説明変数として CPUE 標準化に組み入れているが、可能ならば漁法別に標準化を行い、両者の推定された CPUE 年トレンドを比較してみたら良いのではないか。

説明変数の取捨選択に BIC を使用している。AIC との比較については 4.(1)節で後述するが、AIC 系統と BIC 系統の情報量規準から得られる推定 CPUE 年トレンドの両方をチューニング VPA の資源量指数として利用し、結果を比較することは重要である。

変数選択のための ANOVA Table において Type II SS を利用しているが、Type II はカテゴリーカル説明変数の各水準に関するデータ数を考慮した平方和である。CPUE 標準化のコンセプトと合うのは、各水準のウェイトを等しくした Type III SS ではないだろうか。結果的に両者の違いがみられないこともあるが、可能ならば検討されたい。

説明変数の取捨選択 (i.e. 推定された説明変数の組合せ) と CPUE 年トレンドの間には規則性がなく、CPUE 年トレンドのわずかな違いが、VPA によって推定される年別年齢別資源尾数の大きなギャップの原因となることも多いため、注意が必要である。

(2). チューニング VPA について

プラスグループの F とプラスグループを除く最高齢の F の比率である F-ratio ($\alpha = F_{10+}/F_9$) について、計算期間を通じて一定の値 (1.0) で固定しているが、あまりにも強い制約条件ではないか。全期間を通じて一定として推定した場合は 0.4275 と低い値になり産卵親魚重量 (SSB) のトレンドも合わせて非現実的とのこと、毎年推定するのはパラメータが多すぎて不可能とのこと、主観は入ってしまうが、卓越年級群を考慮して F-ratio が一定と考えられる数年毎のグループ分けをして推定したら良いのではないか。もしくは F-ratio に関する一次の自己相関モデル (AR(1)) の適用を推奨したい。

(3). リッジ VPA について

リッジ VPA (Okamura et al, 2017) について、リッジ回帰 (Hoerl et al, 1970, 2 編の論文) を適用した VPA と考えていたところ、リッジ回帰とは大きく異なるコンセプトに基づく計算が行われており、非常に驚いた。誤解を招かないようにするためにも、リッジ VPA のネーミングを再考しても良いのかもしれない。

スケトウダラ太平洋系群の資源評価で実施していたように、近年の若齢魚の F に関するペナルティを与えたリッジ VPA と通常の VPA で資源量指数である CPUE のフィッティングを比較してみることを、本節の最初に提案したい。

リッジ VPA とリッジ回帰の違いは、私の理解では大別して次の 3 点になる。

- 1) リッジ回帰では最小二乗項が未知パラメータの線形結合になっているのに対し、リッジ VPA ではパラメータ F の非線形関数になっていること
(リッジ VPA では CPUE でフィッティングを行っているため)
- 2) ペナルティ項の比率を表すパラメータ λ について、リッジ回帰では目的関数に含めて推定しているのに対し、リッジ VPA では外部情報に基づき固定していること
(この点について、私(庄野)と提案者の岡村さんの間に見解の相違が認められた)
- 3) リッジ回帰では全ての未知パラメータをペナルティ項に含めているのに対して、

リッジ VPA では直近の若齢魚の漁獲係数 F のみをペナルティ項に含めていること特に一番の違いはリッジ VPA における 1) の非線形性であり、それゆえに 2) の λ の推定を難しくしている。

1) について、最小二乗項が未知パラメータの線形結合の場合と非線形関数の場合では、挙動が全く異なり、リッジ回帰のみならず代表的な L_p 正則化である LASSO

(Tibshirani, 1996) や Elastic Net (Zou and Hastie, 2005) でも線形関数のみを取り扱っている。岡村さんは 2018 年に発表した論文で VPA の構造に基づいた F (漁獲係数) を未知パラメータとした非線形関数でシミュレーションを行い性能が良いことを確認したとのこと、これらの結果を資源評価レポートに参考資料として記載することは有用と思われるが、私は非線形関数の取り扱いには慎重に検討すべきという立場である。

私の認識として、 λ は目的関数に含めて同時に推定するか、 λ が閉区間 $[0, 1]$ の間にあることを利用して、0.01 刻みなどのグリッドサーチに基づく非線形数値的最適化を利用した逐次推定を行うのが基本である、難しい場合には **n-fold cross-validation** などを利用した情報量規準による選択を行うべきと考えている。リッジ VPA では、Mohn (1999) によるレトロスペクティブ・バイアスという方法を利用して λ の値を定めており、私は外部情報を用いて固定していると捉えている一方、岡村さんは VPA と類似の構造に基づいているため **cross-validation** の一種と考えており、見解の相違がある。

リッジ VPA では非線形性ゆえに λ の他の未知パラメータとの同時推定が困難で、さらに各々の未知パラメータから平均を引いて標準偏差で割るなどの基準化に代表されるスケールが難しいため λ を複数設定する必要があり、グリッドサーチに基づく逐次推定を複雑にしている。

この点について、拡張カルマンフィルター等非線形関数を線形化する方法の適用も一案であるが、 F の値が 1 を超えることはめったに起こらないことを考慮し、 $[0, 1]$ 区間に **mapping** させるスケールを基準化の代わりに利用したら良いのではないかと。

また、リッジ VPA の最小二乗項を形式的に線形化する方法、すなわち CPUE に代えて F でフィッティングする VPA を感度分析の 1 つとして試しても良いのではないだろうか (**parametrization** などの工夫が必要なため簡単な作業ではないかもしれない)。

なお、ペナルティ項における未知パラメータの選択について、直近の若齢魚の漁獲係数 F のみをデフォルトとするにせよ、**robustness test** の一環として全てのパラメータをペナルティ項に含めたチューニング VPA を実施することも合わせて推奨したい。

オリジナル論文の引用の是非について、一般に広く参照されている教科書で代用できるという考え方に私は反対であり、その理由として、以下に 2 つ挙げておきたい。

- ・教科書とオリジナル論文の記述が異なる場合もあり、読者の **misleading** に繋がる可能性があること
- ・教科書は引用文献が多く、読者がオリジナル論文に辿り着けない可能性があること

4. 将来予測に用いられた前提条件および手法の妥当性について

(1). 再生産曲線、管理基準値、KOBÉ-plot などについて

B-H (ベバートン・ホルト), Ricker, H-S (ホッケー・スティック) などの再生産曲線の比較表について、最小絶対値法や最小二乗法という用語が使われている。これは、対数尤度関数の二乗誤差項を絶対誤差項で置き換えた、統計学的に誤った方法と混同される可能性があるため、L1 ノルムと L2 ノルムなど、表記を改めたら良いのではないか。加えて、検算を容易にするため、もし可能ならば標本数や未知パラメータ数、仮定した確率分布の名称および最大対数尤度の値を追記することを推奨したい。

上記の比較表で使用している AICc (AIC の有限修正) について、オリジナルの正しい表記は c-AIC である。CAIC (Consistent AIC; Bozdogan, 1987) や cAIC (conditional AIC; Vaide and Blanchard, 2005) と見間違える可能性もなくはないが、c-AIC は始まりが小文字で - (ハイフン) があるため、識別可能である。記法を改め、参照されることが少ないオリジナル論文 (Sugiura, 1978) も忘れずに引用すべきである、と考えている。

再生産曲線の推定に際しては、AIC (Akaike, 1973) 系統の c-AIC を使用している一方、CPUE 標準化のモデル選択で BIC (Schwarz, 1978) を利用している点について、前者は小標本のため、後者は一貫性と呼ばれる標本数を無限大に近づけた場合に真のモデルを選択する性質を重視した、という説明があった。一貫性については、真のモデルが存在するという前提かつ未知パラメータのネスト構造 (包含関係) が認められる場合にパラメータ数が一番多いフルモデルが真のモデルでないなどの条件が必要なため、無条件に成り立つ訳ではない。そのため、BIC と予測誤差最小のコンセプトに基づく AIC との比較における優劣はケース・バイ・ケースで、必ずしも BIC が優れている訳ではないことに注意が必要である。

(2). 将来予測のシミュレーションについて

漁獲率一定方策や産卵親魚量一定方策と比較して、管理目標が同じ場合に平均漁獲量が少なくなる漁獲量一定方策を取り扱っている点について、KOBÉ-plot などの指標値に基づく近年の資源状態が良いスケトウダラ太平洋系群では影響が少ないのかもしれないが、今後の資源状況によっては他の方策への切り替えが必要ではないか。ただし、多くの漁業者が毎年の漁獲量の大幅な変動を好まない点については理解できる。

5. おわりに

2~4 節で記述したコメントや提案について、重要度や改善に要する時間もまちまちだが、適宜取捨選択して優先順位の高い事項から検討していただけるとありがたい。

スケトウダラ日本海系群やホッケ道北系群についても同様だが、資源評価部分と比べて将来予測の概念や記号の定義など、レポートの内容理解が難しい点が多かったた

め、記号の定義を巻末一覧表にするなど、記述の改良や説明の工夫をお願いしたい。
3(3)の大部分および3(2), 4(1)の一部分の記述は、私が担当したスケトウダラ太平洋系群、日本海系群、ホッケ道北系群の報告書に共通である。ご了承いただきたい。

文献

- Akaike, H. (1973) Information theory and an extension of the maximum likelihood principle. 2nd International Symposium on Information Theory. Petrov, B. N., and Csaki, F. (eds.), Akademiai Kiado, Budapest, p.267-281
- Bozdogan, H. (1987) Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika*, **52**: p.345-370
- Hoerl, A. E. and Kennard, R. (1970) Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*, **12**: p.55-67
- Hoerl, A. E. and Kennard, R. (1970) Ridge regression: Applications to nonorthogonal problems. *Technometrics*, **12**: p.69-82
- Mohn, M. (1999) The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. *ICES Journal of Marine Science*, **56**: p.473-488
- Okamura, H., Yamashita, U. and Ichinokawa, M. (2017) Ridge virtual population analysis to reduce the instability fishing mortalities in the terminal year. *ICES Journal of Marine Science*, **74**: p.2427-2436
- Schwarz, G. (1978) Estimating the dimension of a model. *Annals of Statistics*, **6**: p.461-464
- Sugiura, N. (1978) Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics -Theory and Methods*, **7**: p.13-26
- Tibshirani, R. (1996) Regression shrinkage and selection via the Lasso, *Journal of the Royal Statistical Society B*, **58**: p.267-288
- Vaida, F. and Blanchard, S. (2005) Conditional Akaike Information for Mixed-Effects Models. *Biometrika*, **92**: p.351-370
- Zou, H. and Hastie, T. (2005) Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society B*, **67**: p.301-320