

岡村 寛

国立研究開発法人水産総合研究センター中央水産研究所  
資源管理研究センター資源管理グループ グループ長

海洋生態学と機械学習法の融合によるデータ不足下の生態系評価手法の開発

## § 1. 研究実施体制

(1) 「生態系モデル」グループ (水産総合研究センター中央水産研究所)

- ① 研究代表者: 岡村 寛 (中央水産研究所資源管理研究センター グループ長)
- ② 研究項目
  - ・生態系モデルの開発
  - ・生態系データ解析

(2) 「生態系指標」グループ (情報・システム研究機構統計数理研究所)

- ① 主たる共同研究者: 江口 真透 (統計数理研究所数理・推論研究系学習推論グループ教授)
- ② 研究項目
  - ・生態系指標の開発
  - ・極端に情報が少ない中で生態系を評価する手法の開発

## § 2. 研究実施の概要

生態系データは、様々な不確実性の影響を被っている。それ故に、そのような不確実性を無視したとき、生態系の評価を誤る可能性がある。漁業活動は海洋生態系の中で大きな役割を果たしているため、生態系について知るための重要なデータソースとなる。しかし、漁業データには欠測やバイアスの混入など、生態系の中の資源の状態を知るのを困難にする様々な要因がある。一方、漁業データや調査データを統合的に解析して、海中の生物資源に対する評価がなされ、個体数の状態などの情報が得られている生物種も多い。そのような生物種に対しては、資源が健全な状

態にあるかどうかを知ることが可能である。

資源状態の情報をクラスラベルと考えてやる時、これは機械学習などで使用される用語で「教師付き」データとなる。そこで、資源状態が分かっている生物種のデータを教師付きデータとみなして、漁獲量やその生物の特徴(最大体長、栄養段階など)だけを使用し、資源の良し悪しを判定する予測モデルを構築してやることを考える。このようにして、資源状態の情報がない生物資源のようなデータ不足下にある生態系についても、その評価を可能にしようというのがこの研究のひとつの目的である。

生物資源評価のグローバルデータとして、RAM Legacy Data というデータセットがある。これを使用して、限られたデータから生態系中の資源の状態を予測する予測モデルを構築する。それによって、世界の資源の漁獲量統計 (FAO が公表している) から、世界中の海洋生物の資源状態を予測することが可能となる。しかし、RAM Legacy Data には問題があり、それは資源状態の良し悪しに関するサンプルサイズがかなり不釣り合いで、悪い状態のデータが不足していることである。これは、通常の予測モデルを用いた場合、悪い状態の予測を難しくする。しかし、生物保全の立場からは、悪い状態の予測がより重要なのである。

そこで、我々は、サンプルサイズの不釣り合いの影響を考慮した非対称ロジスティック回帰モデルと呼ばれる新しい評価手法を開発した。この方法を使用した結果、全世界の生物資源の崩壊確率を推定すると、従来 (のロジスティックモデルによる結果) とは異なる推定値が得られることが分かった (図 1)。この結果は、生態学のトップジャーナルのひとつである *Methods in Ecology and Evolution* 誌で公表された (Komori et al. 2016)。

また、我が国周辺の生物資源の放射性物質による汚染リスクを評価する統計モデルを開発し、リスク評価を行った。この成果については、*Proceedings of the National Academy of Sciences* にて出版された (Okamura et al. 2016)。

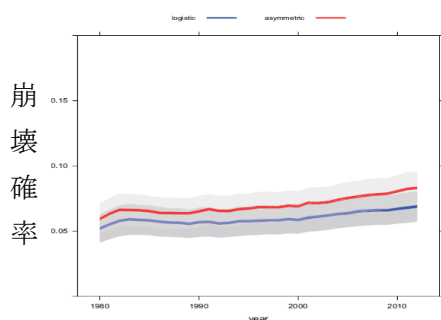


図 1. 世界の生物資源の崩壊確率 (赤: 新モデル, 青: 従来モデル)