

水産資源の持続性を統計手法で中立的に評価



▲江口真透教授

世界の水産資源が枯渇する!? 「2048年問題」の衝撃

世界の海で現在のような乱獲や汚染が続けば、人間が食べるシーフードは、2048年までに消滅する——。2006年、カナダの海洋生物学者ウォームらの研究グループが、アメリカの科学雑誌「サイエンス」にそんな論文を発表した。このニュースは、魚食文化の根づく日本はもちろん、世界に衝撃を与えた。

その後、この研究のデータ解析の仕方には問題があるとして、反論する意見や論文が出され、現在ではウォームのシナリオがそのまま現実となる可能性は低いという見方が主流になって

いる。とはいえ、この論文が契機となり、人間活動が海洋の生物多様性に与える影響に関する研究が各国で盛んになったことは確かだ。

2012年度に、科学技術振興機構（JST）の戦略的創造研究推進事業「CREST」に採択された「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」もその一つ。水産総合研究センター中央水産研究所の岡村寛グループ長が代表を務める研究チームに、統計数理研究所からは統計的機械学習研究

センターの江口真透教授、数理・推論研究系の小森理特任助教（現在、福井大学特命講師）が参加している。

水産資源の将来予測については、環境保護か資源維持かの立場の違いによって、異なる考察が出回っている。江口は「水産資源に関する知識ではなく、データに基づく統計手法によって中立的に推論する必要がある」と話す。ここに、統計の専門家が参画する意義がある。

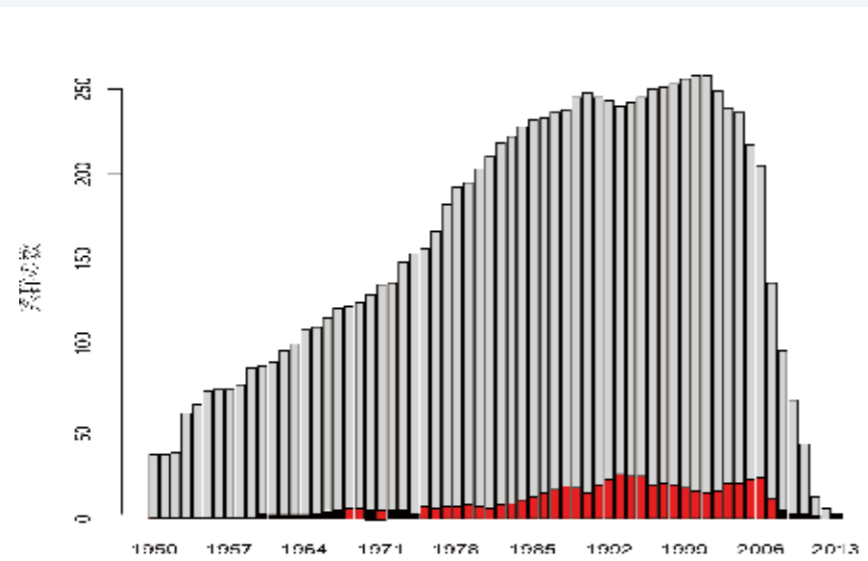


図1：枯渇した系群（赤）と枯渇していない系群（灰色）の数の年推移。



▲小森理 福井大学特命講師

「健全」か「枯渇」か。2値回帰分析に非対称ロジスティックモデルを考案

江口らはこのプロジェクトで、漁獲量と栄養段階、生息地域、最大体長などのデータから2値化したバイオマスデータを予測する統計的モデルを構築。機械学習の手法を用いて、魚などの系群ごとに生態的に「健全」か「枯渇」かを予測する「教師あり学習」を提案した。

ところが、当初からこの手法の確立に成功したわけではなかった。「最先端の機械学習を使えば解決できると思いましたが、あまり良い結果は出ませんでした。もっと根源的な問題があるとわ

かったのです」と江口は打ち明ける。それが、ラベルごとのサンプルの持つ不確定性の相違だ。健全な系群の分布と枯渇しそうな系群の分布とでは、前者が圧倒的に多く、後者はごくわずかなので、データ数に大きな開きがある。つまり、非対称度が高い。一般に非対称度が高いほど予測は難しくなる。

そこで、江口らが考案したのが、「2値回帰分析のための非対称ロジスティックモデル」だ。2値が非対称であっても、全世界または地域ごと、魚の種類ごとなどで、枯渇の懸念があるかないかを明らかにできる。「検証の結果、従来のモデルによる解析よりも有

効なパターン認識が得られることがわかりました」と江口。適切なモデリングを行うことによって漁獲量と栄養段階、生息地域、最大体長など他の変数をうまく組み合わせれば、資源の健全さを予測できる。これによって、「ここまでなら獲っても影響がない」という量が算出できるわけだ。

このテーマでまとめた論文は、2015年9月にイギリスの生態学会誌「メソッド・イン・エコロジー・アンド・エボリューション」に掲載された。江口は「予測の精度が上がり、ウォームの論文への反論を統計的な立場で再評価することができました」と、このプロジェクトにおける成果を説明する。

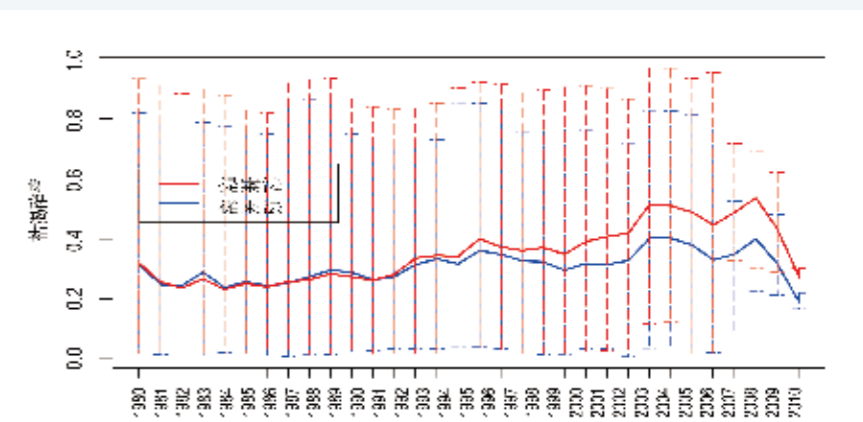


図2：枯渇した系群に対する提案法（赤）と従来法（青）による枯渇率の比較。

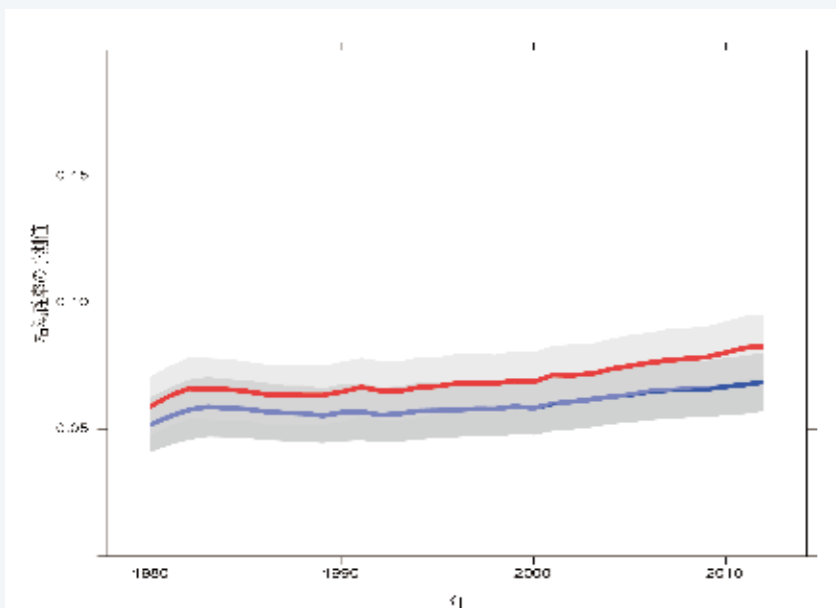


図3：提案法（赤）と従来法（青）による世界的な枯渇確率の推移の予測。



図4：水産資源評価の国際シンポジウムのポスター。

メイヤー博士の遺産として 開示された生態系データ

水産資源の将来予測が研究者によって大きく異なる理由の一つは、データの少なさに起因する。陸上とは違い、海の中で生物を観測するのは容易ではない。このため、入手できるデータには限りがある。「水産資源はミッシングデータのかたまりと言っていいでしょう。観測できるのは、全体の何100万分の1程度なのですから」と江口は話す。

最も大規模かつ広域をカバーしているのは、国連食糧農業機関（FAO）の保有する漁獲量データだ。「2048年に水産資源が枯渇する」としたウォームの予測も、FAOが1950年から2003年までに集めたデータに基づく予測だった。だが、漁獲量だけで生態系の状態を把握しようとするのは無

理がある。なぜなら、漁獲量は市場の需要に大きく影響され、サンプルに偏りが生じるからだ。予測精度を高めるためには、漁獲量以外のバイオマスデータが必要だった。

そこで、江口らが着目したのが、「RAM 記念資源データベース」だ。精力的な研究の半ばで急逝した海洋生態学者ランソン・A・メイヤー博士の功績を記念し、学術研究のためにWeb上で公開されているものだ。ここには、世界的な範囲で200を超える系群のバイオマスを含む年次データがある。「2値回帰分析のための非対称ロジスティックモデル」には、このデータを使用した。モデル化の過程では、メイヤー博士と同じく海洋生態学者であるヒルボーン博士が来日した際に、直接会ってディスカッションをしたという。

水産資源のデータは、長いこと地域

限定のクローズドなものだった。各国の経済戦略や外交に絡む重要機密として扱われていたからだ。それがここ10年ほどのトレンドでは開示されつつあるという。「今回のモデル化も、グローバルデータを入手できたからこそ実現できたのです」と江口は言う。

実データがあればこそ 困難な推論を乗り越えられる

「アイデアは理論だけ考えても出てきません。現実の課題を解いていくなから、理論的なヒントを得ることが重要。その意味で、今回はいい経験でした」と江口は振り返る。

このプロジェクトが始動する前、統計研チームは医療分野で抗がん剤の効果に関する評価を手掛けていた。遺伝子発現パターンから、患者の特性ごとに有効な投与量や頻度をモデ

ル化するものだ。だが、モデルを考案してみるものの再現性がなく、難航していた。

そんなとき、このプロジェクトへの参加を打診してきたのが、古くからの研究仲間である岡村教授だった。「原点に戻ってみようと思い、引き受けました」と江口は明かす。抗がん剤の効かないも、水産資源の枯渇するしないも、統計的に見ている「ふるまい」は同じ。今回の成果を横展開できる手応えを感じているという。プロジェクトのほうでも、今後は機械学習の方法も採用してFAOデータの予測問題をを進める予定だ。

大学院生の時代から江口に指導を受け、共に研究を続けている小森は「このモデルは実際のデータを見て、理論を考えるという繰り返りで到達した考え方。不確実性の多い生物の世界に対応したもので、金融や機械故障など応用できる範囲は広いでしょう。この先も、実社会に役立つ手法を考えたい」と展望を語る。

ブームとも言えるビッグデータの波は、今回のプロジェクトがテーマとする海洋生態学の分野にも押し寄せている。それにより、調査時の気象や海域の塩分濃度、深度などのデータがリンケージされるようになった。生物にセンサを取り付けてデータを取得するバイオロギングなど、新しい調査方式も開発されつつある。江口は「生態学のなかでも統計学の役割はますます重要になるでしょう。やるべきことは、まだたくさんあります」と先を見つめている。

（広報室）

RAMデータとFAOデータの違い

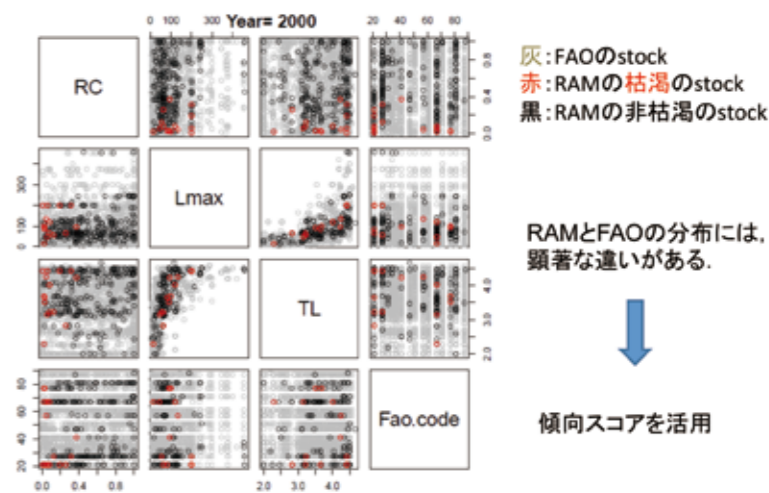
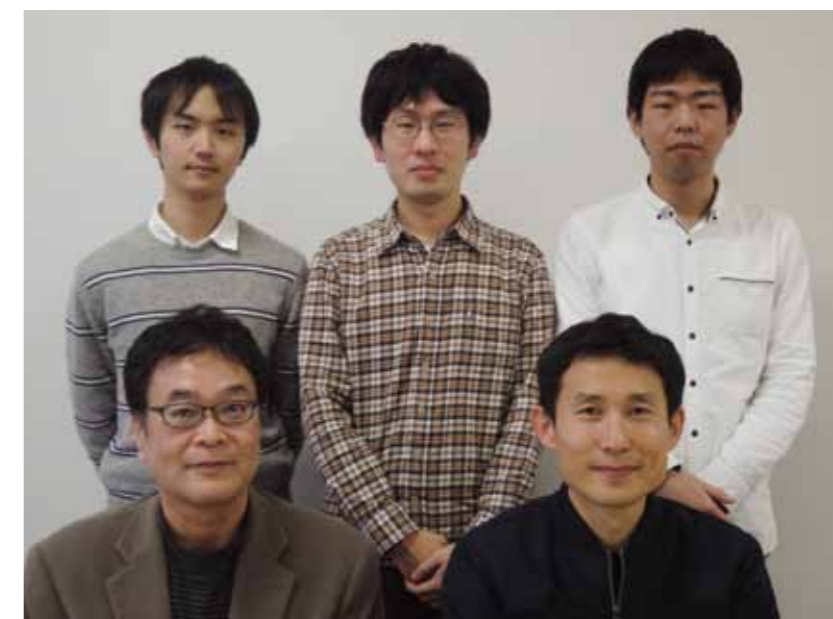


図5：今後の問題としてFAOとRAMのデータの食い違いを注目して選択バイアスの補正にチャレンジしている。



▲プロジェクトチームのメンバー。後列右から東京理科大学大学院博士課程2年の三枝祐輔さん、同修士1年の丸山智久さん、同修士1年の岡田昌之さん。「研究室では理論が中心なので、実データの解析は新鮮な体験」（三枝）、「実データを扱うのは初めてで、勉強になる」（丸山）、「研究室の研究よりも難しい」（岡田）と話す。