

サケマス回帰率向上のための ICT を活用したビッグデータ取得と利活用に関する研究 (162301002)

A Study on Obtaining and Utilizing Big Data for Improvement of Salmon Return by ICT

研究代表者

塩谷 浩之 (室蘭工業大学)

Hiroyuki Shioya (Muroran Institute of Technology)

研究分担者

春日井 潔 北海道総合研究機構さけます・内水面水産試験場)

Kiyoshi Kasugai (Hokkaido Research Organization, Fisheries Research Development, Salmon and Freshwater Fisheries Research Institute)

研究期間 平成 28 年度～平成 30 年度

概要

本研究開発では、AI や機械学習を活用する形式にて北海道のふ化放流データ構築を行った。結果として 20 年分のふ化放流データが整備された。ふ化放流に対応する大規模な入力に対応する AI の構成には、深層学習による予測モデルが適切である。しかしながら 20 年では年数としては少ないことから、知識モデルによるデータ改善、およびノイズによるデータ改善を行ったことでデータ学習は改善され、AI の予測能力に対応する汎化誤差を抑えることができた。北海道の全海区の来遊数の予測では、従来法のシブリング法と比較して、予測誤差を半分程度となる来遊予測法を実現した。稚魚の育成過程では様々な過程があり、水産増殖のための本格的なデータ化には ICT の導入が不可欠である。ふ化場の運用効率化につながる情報活用の一環として、捕獲した親魚のふ化場の養魚池に適切に配置するためのアプリ開発を行った。親魚捕獲アプリを根室海区のふ化場で利用を通じて、今後のデータの充実化と利活用につなげた。

1. まえがき

サケマス漁は全国 80% を占める北海道の重要な水産業であるが、最近では不漁が続いており、水揚げ確保は重要な課題である。親魚捕獲から採卵、受精、稚魚育成・放流をふ化放流事業において行っている。稚魚放流数に対する回帰数の向上が課題である。北海道では 10 億の稚魚を放流し、回帰するのは全体の 4% (4000 万尾) である。情報技術を活用し、本申請課題においては、ふ化事業のデータ蓄積と利活用により回帰予測を行う。近年の温暖化等の影響により、日本海サケマス、太平洋秋サケ、オホーツクカラフトマスの回帰率低下が著しい。そのなかで、回帰率向上のためのビッグデータの利活用を進めることを目的として、データ取得と予測技術の研究開発を行った。

2. 研究開発内容及び成果

北海道内の全海区において、放流河川、時期、尾数、体重について利用できるデータは 1996 年から 20 年程となる。一か月を 3 旬に分けた放流旬と放流尾数がデータとなる。補助的に沿岸の海水温も併用する。20 年程度のサンプルで学習を行うには大幅に不足している。説明変数が多く巨大データであっても、統計性を支えるための年数について不足している。よって、繰り返し年数の少なさを補うデータ補完技術が不可欠である。

機械学習においては、データが少ないと予測精度は期待できない。その対応策として、既存事実として確立している知識があるときは、それをデータ化することで学習に活用できる。これを知識モデルデータと呼ぶことにする。根室海区について、放流時期と海水温に関する対応関係を調査した。その結果、放流時期の旬の早期にシフトと来遊数減少の対応が知識モデルとすることにした。稚魚にとって適切な海水温よりも早めの放流と良くない回遊結果を対応させる知識モデルも併用できる。これらを早期放流に関する知識モデルデータとして構成した。放流尾数は決定論的数ではなく、ある種のノイズデータである。放流グループからの稚魚のサンプルにおける体重の平均を放流する稚魚全体の重量で換算したものを放流尾数としている。結果

として、放流尾数は、ガウスノイズの知識モデルとなる。既存のデータを Raw Data、早期放流の知識データを加えたデータを Model data、ガウスモデルとして放流尾数からデータを発生させたものを Gaussian data としてガウスノイズを用いて増分データを用意した。データ学習と予測で用いる深層学習として、多層ニューラルネットワークを用いる (図 1 上)。入力層としては放流尾数と海水温で構成した。訓練データの学習過程から、知識モデルやガウスモデル双方で改善が見られた (図 1 左下)。各データセットにおいて、訓練用のデータによる学習誤差と未学習データに対する予測誤差を示す。導入モデルで予測性能が向上している (図 1 右下)。

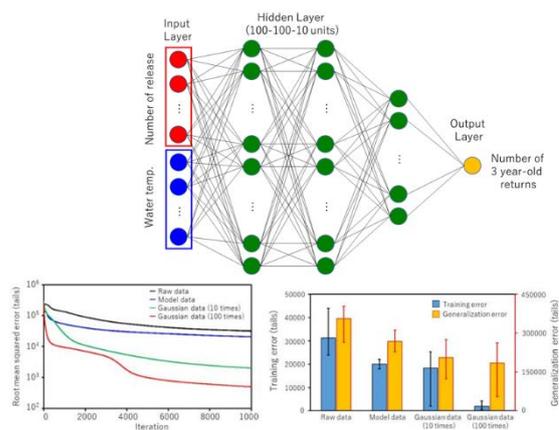


図1: 3年魚来遊数予測のための深層学習、訓練誤差の減少プロセスと予測誤差における、従来法と知識モデル導入との比較

一般的には全年齢が来遊数予測するには、3 年魚を例年数にて定め、前年に来遊した尾数から 4、5、6 年魚来遊数をシブリング法によって予測する。このような従来法では 3 年魚の来遊数は予測の初期値となり、そのズレによる影響は 4 年魚来遊数に影響する。本研究では 3 年魚を機械

学習でふ化放流データから推測している。サケの水揚げには豊漁と不漁があると考えると、例えばロジスティックによる回帰を用いても、原理的に対応ができないことがわかる。以上から、シプリング法への根本的改善、あるいは他手法を組み合わせるなどの改善を行う必要がある。本研究では AIC (An Information Criteria) を用いて推定値によって豊漁か不漁の判定を構成した。AIC とは情報量規準の一種であり、データから計算される負の対数尤度と情報量が最小となるパラメータ数から得られる。データ群における豊漁不漁のクラス分けに K-means 法を用いることで、提案法を AIC クラスタリング手法と呼ぶことにする。全海区について、提案法による来遊数に関する検討を行った。図 2 に観測値 (黒)、従来のシプリング法 (青)、シプリング法に AIC クラスタリングで改善した手法 (赤) による結果を示す。改善度合は様々であるが、5 海区の合計では、2016 年 2017 年では双方で従来法と比較して、AIC クラスタリング手法によって誤差を半減させており、来遊予測精度の改善成果を示している。

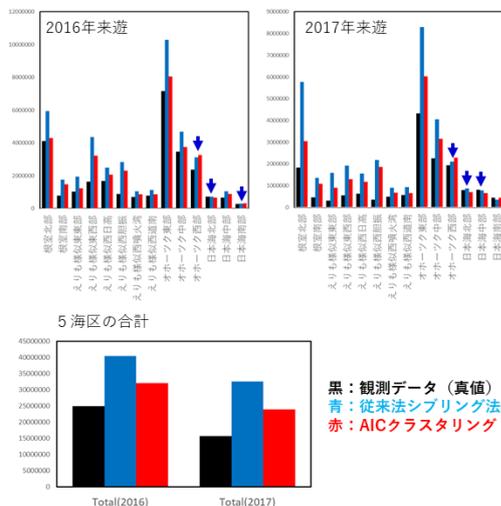


図2：全海区におけるサケの来遊予測の比較、従来法 (シプリング法) と AIC クラスタリング法と実際の来遊数

一般論としてデータの充実化は、来遊予測の将来的な向上につながる。ふ化場においては、親魚捕獲と管理は、ふ化放流の出発点となる重要業務である。それへの ICT 支援の試験的開発研究として、根室海区のふ化場で聞き取り調査を行い、親魚捕獲とデータを集約するアプリ開発を行った。

本研究開発で作成した親魚捕獲のデータ管理システムを図 3 に示す。捕獲数、蓄養する尾数、採卵他で使用した尾数について、ふ化場から管内さけ・ます増殖事業協会の事務所に報告する。捕獲場 (河川) とふ化場での両方の作業で生じるデータの総合管理を行えるアプリを作成し、実業務で試験使用を実施した。

サケのふ化放流業務に ICT をより積極的に導入することが、現場とデータ解析を行う機関との間で、データ収集への業務負担などの垣根を低くする方向に働くので、今後の積極的な ICT 提供を続ける必要がある。

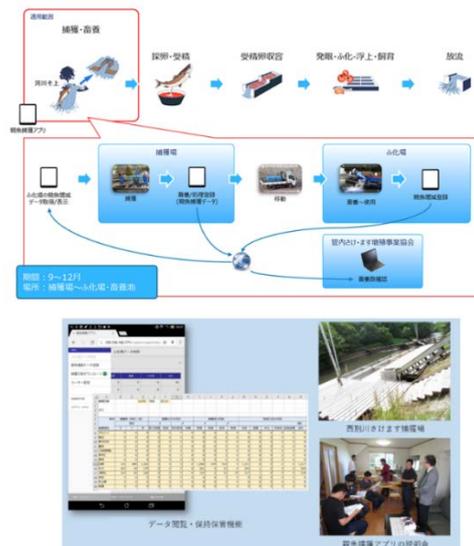


図3：親魚捕獲のためのICT開発、捕獲プロセスとデータ蓄積 (西別川における試作アプリ使用)

3. 今後の研究開発成果の展開及び波及効果創出への取組

本研究開発によって改善された予測手法を構築した。次の段階として、実用に繋げる研究開発を行う必要がある。そのためには、公表のための来遊予測計算システム開発を進める。それに加えて、来遊数データの信頼性を向上させる取り組みが必要となる。来遊数データの調査の規模拡大のための研究開発が不可欠である。サケの調査では、捕獲場、漁港、市場等でウロコを採取する。現状の方式で調査数を増やすには莫大なコストがかかり、調査結果を得るまで期間がかかる。そこで現場におけるデータ化と迅速な情報共有のための総合化した情報機器のシステム開発を進めることで、データの充実化を進め、来遊数予測の向上につなげる。

4. むすび

本研究開発では、3 年魚とふ化放流データに AI 機械学習を利活用した手法を組合せ、加えて、情報理論的なモデル手法を援用することで、来遊予測を改善できた。これまでの水産科学で培った予測方法と、情報科学の理論モデルの導入により、水産学と情報学の新しい共同研究を進めることができた。

【誌上発表リスト】

- [1] 小林 賢哉、佐藤 寛己、塩谷 浩之、春日井 潔、“サケのふ化放流データに関する学習モデル”、第 80 回情報処理学会 全国大会、1k-06 (平成 30 年 3 月 13 日)
- [2] Y. Maehara, H. Shioya, K. Kasugai, Y. Miyakoshi and M. Fujiwara, “Using hatchery-release data of eastern Hokkaido to predict returns of chum salmon,” 10th Soft Computing and Intelligent Systems and 19th International Symposium on Advanced intelligent Systems in conjunction with Intelligent Systems Workshop 2018, pp.1068-1073 (2018 年 12 月 7 日)