

機械学習を利用した内湾における 浮魚類群集の空間分布量の推定

濱田 孝治¹・吉田 司²・岡村 寛³・原 武史²・鈴木 輝明⁴

¹正会員 株式会社シャトー海洋調査 環境調査部 (〒111-0053 東京都台東区浅草5-23-9)

E-mail:hamada-chateau@to.email.ne.jp

²非会員 全国水産技術者協会 (〒107-0052 東京都港区赤坂1-9-13)

E-mail: yosida.tu20@gmail.com

³非会員 水産研究・教育機構 中央水産研究所 資源研究センター

(〒236-8648 神奈川県横浜市金沢区福浦2-12-4)

E-mail:okamura@fra.affrc.go.jp

⁴非会員 名城大学大学院 総合学術研究科 (〒468-8502 名古屋市天白区塩釜口1-501)

E-mail: suzuteru@meijo-u.ac.jp

浮魚類群集(イワシ類)は内湾における主要な漁獲対象種であり, 自然的, 人為的な環境変動に伴う漁獲量の変動や漁場形成の変化は大きな関心事である. 本研究では機械学習の一つであるGradient Boostingを用いて内湾の浮魚群集の分布量を推定する統計モデルを構築し, 伊勢・三河湾のカタクチイワシ漁獲量の時空間的分布の推定を行った. モデルは標本漁船による実測結果をよく再現した.

Key Words : Anchovy, *Engraulis japonicus*, Machine learning, XGBoost, Fish catch estimation

1. 背景と目的

カタクチイワシ *Engraulis japonicus* は北太平洋・中部太平洋の西部に広く分布し¹⁾, 平均水深15 m, 閉鎖型内湾である伊勢・三河湾 (図-1) を索餌, 成長, 成熟, 産卵の場として利用している²⁾³⁾. 2016年の三重県・愛知県のカタクチイワシ漁獲量(シラスを除く)は40,858 t, 生産額は1,714百万円であった⁴⁾. これは, 2県の海面漁業の漁獲量の16.5%, 生産額の3.8%, また, 全国のカタクチイワシ漁獲量および生産額のそれぞれ23.9%, 13.2%となり, 内湾環境の変化に伴う漁獲量の変動や漁場形成の変化は大きな関心事となっている.

魚類等の資源量, 漁獲量やその分布の推定の方法としては, 魚類動態モデルによる方法⁵⁾, 統計モデルによる方法⁶⁾, AI(人工知能)による方法⁷⁾などがある. 魚類動態モデルによる方法では, 室内実験の結果などにに基づき環境条件に対する生物の選好性をモデル化し, その行動をシミュレートする. しかし, 選好性に関する情報は必ずしも豊富にあるわけではない. 一方, 統計モデルによる方法では, 実測データを用いてGLM(一般化線形モデル), GAM(一般化加法モデル)などのモデルを最適化する. AIによる方法は統計モデルによる方法において「統計モデル」の部分に「機械学習」や「ニューラルネット」な

どで置き換えたものにあたり, より柔軟に実測データに適応させることができるため, 近年の計算機能力の向上に伴い注目を浴びている.

以上をふまえ本研究では, 伊勢・三河湾のカタクチイワシを対象として, 機械学習の一方法であるGradient Boostingを利用した統計ソフトウェアR⁸⁾のパッケージXGBoost⁹⁾¹⁰⁾によって, 標本漁船により得られたその時空間的分布量の再現を試みた.

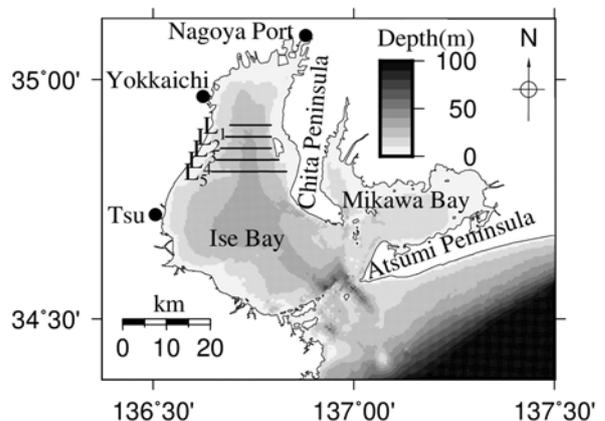


図-1 伊勢・三河湾. L₁~L₅は計量魚探測線を示す.

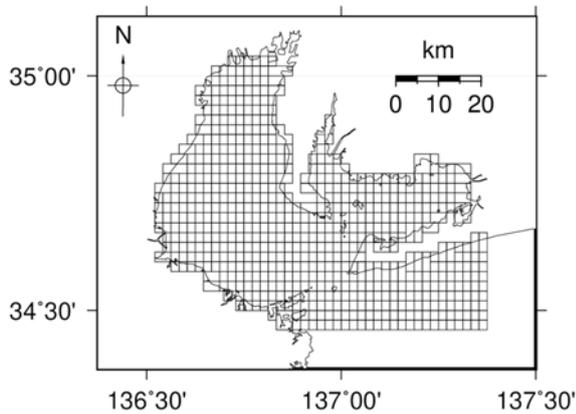


図-2 標本船メッシュ

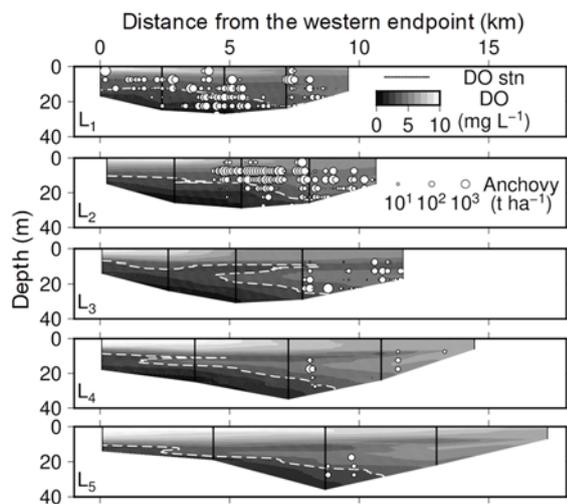


図-3 湾奥の DO およびカタクチイワシの量の断面分布。
横軸は各測点西端からの距離 (km, 図-1 参照),
縦軸は水深 (m), 点線は $DO = 3 \text{ mg L}^{-1}$ を表す。
魚種判定は計量魚探調査と同一測線上で行った
パッチ網の試験操業結果によった。

2. データと方法

(1) 特徴量と出力

本研究では伊勢・三河湾を対象海域とし、メッシュ幅 $1.25'$ として対象海域を167メッシュに分割し(図-2)、特徴量(説明変数)をメッシュごとの水温 (T), 塩分 (S), 溶存酸素濃度 (DO), クロロフィルa濃度 (Chla), 動物プランクトン量 (Zoo) とした。

魚類は変温動物であり、水温 (T) は魚の生理に支配的な影響を及ぼすだけでなく、体温の変化を最小限にするための行動が、漁場形成や回遊経路などの分布に大きな役割を果たす¹¹⁾。カタクチイワシ成魚の塩分 (S) に対する選好性は明らかでないが、たとえば Oozeki et al. (2007)¹²⁾ はカタクチイワシが主に塩分23~34.5 (psu) の海域を産卵場としていることを述べており、塩分はカタクチイワシ成魚の分布に影響を及ぼしていることが考えられる。溶

存酸素 (DO) は魚類の生存に必須であり、浮魚は貧酸素水塊を忌避する。図-3に、2015年9月12日に測定した測線L₁~L₅(図-1)における、計量魚探 (SIMRAD社製 EK60: 周波数 70 kHz) によるカタクチイワシの量の断面分布および同日のDO断面分布を示す。カタクチイワシが貧酸素水塊 (3 mg L^{-1} 以下) を避けて分布していることがわかり、DOを説明変数とすることは妥当であるといえる。植物プランクトン(本研究ではクロロフィルa濃度 (Chla) を植物プランクトンの指標とする) と動物プランクトン (Zoo) は、カタクチイワシの主要な餌である^{13),14)}。

(2) 漁獲量および水質データ

漁獲量の実測値は2014年~2016年、7~9月のパッチ網を用いた標本船 (12ヶ統) の操業記録より得た。パッチ網は大目の翼網と細目の袋網からなる二艘曳きの網で、漁獲対象魚群を翼網によって徐々に袋網口へ誘導し入網させることを意図している¹⁵⁾。パッチ網一曳網あたりの漁獲量に対応するメッシュに配分し、各月の一週間分を積算して漁獲量実測値とした。

生息環境 (T, S, DO, Chla, Zoo) には伊勢湾シミュレーター¹⁶⁾の現況再現された計算値を用いた。伊勢湾シミュレーターは、非静水圧の3次元流動モデルと微生物を考慮した低次生態系モデルから構成される数値モデルである¹⁷⁾。本研究では、2014年から2016年の3年間の計算結果を使用し、各月の一週間分を平均し、また、カタクチイワシの主な遊泳層を考慮し水深5~15 mで鉛直的に平均した値を用いた。

(3) 漁獲量の推定法

魚群はパッチ状に存在し、魚群と遭遇しない場合漁獲量は0となるため、0を多く含んだデータとなる。そこで、水産資源学の分野で用いられるdelta model の考え方を適用し、データを分布量の有無 (Binary; B) と正の分布量 (Positive; P) に分離した¹⁸⁾。そして、実測値が正となる確率を求める学習器 (B) と、実測値が正の場合の分布量の期待値を求める学習器 (P) を構築し、推定分布量は両者の積とした。

本研究では、機械学習の一方法であるGradient Boosting (勾配ブースティング) を用いた。Gradient Boosting を利用したソフトウェアのうち、特に性能が高いことで知られるRのパッケージのXGBoost (eXtreme Gradient Boosting) を使用した。XGBoost は決定木からなる弱い学習器を複数重ね合わせて一つの強い学習器を作る。XGBoost はモデルの能力を評価する目的関数 (objective function) が最小となるよう逐次的に決定木を構築し追加していく方法である。目的関数は推定値と実測値の差をあらわす損失項 (loss term) と木構造の複雑さを表す正則化項 (regularization term) の和となっており、モデルが複雑になりすぎ過学

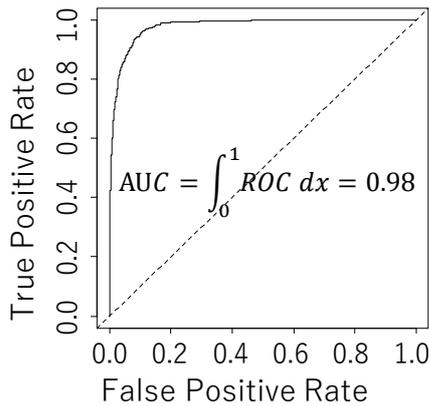


図-4 学習器(B)のROC曲線とAUC

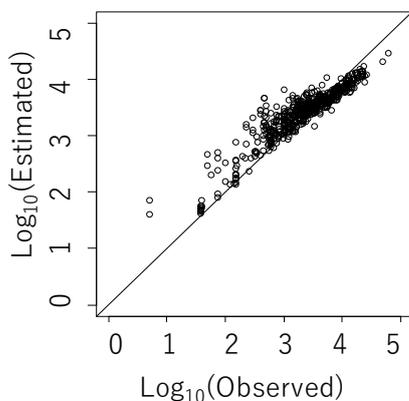


図-5 学習器(P)の漁獲量実測値と推定値の関係

表-1 クロスバリデーションの結果

run	AUC	run	AUC
1	1.00	6	1.00
2	0.96	7	0.91
3	0.92	8	0.95
4	0.94	9	0.93
5	0.93	10	0.94
average=0.95			

習(後述)が起こることを防ぐようになっている。本研究では、カテゴリデータを扱う学習器(B)ではロジスティック回帰¹⁹⁾、連続データを扱う学習器(P)ではTweedie回帰²⁰⁾を適用した。

(4) 推定能力の確認法

学習器(B)、すなわち漁獲の有無の推定能力はROC解析²¹⁾によって確認した。ROC曲線(Receiver operating Characteristics)はカットオフ値を媒介変数としたとき x =感度(True positive rate), y =特異度(False positive rate)が描く曲線

である。また、AUC(Area Under the Curve)はROC曲線を特異度0から1の範囲で積分した値である。ROC曲線が(0,1)を通るとき、適当なカットオフ値に対して学習器は漁獲量の有無を100%正確に判別する。AUCが1に近いほど、学習器の推定能力は高いとされる。学習器(P)、すなわち漁獲があった時の漁獲量の予測能力は実測値と推定値の散布図によって確認した。

機械学習ではしばしば過学習が問題となる。過学習とは既知のデータ(トレーニングデータ)に対しては適合できるが、未知のデータ(テストデータ)に対して適合できない状態を指す。XGBoostの過学習に対する強さを確認するため、全データを80%のトレーニングデータと20%のテストデータにランダムに分割して学習器(B)のクロスバリデーションを行った。

(5) 漁獲量への環境要素の寄与度

生息環境が漁獲に与える影響について考察するため、各特徴量の寄与度(gain)⁹⁾を求めた。寄与度とは、木構造へノードを追加したときの目的関数の減少量を特徴量毎に足し合わせて得られる値を正規化したものである。

3. 結果と考察

(1) ROC解析によるモデルの性能の検討

図-4に、学習器(B)のROC曲線を示す。AUCは0.98であり、学習器(B)の推定能力は高いと判断された。図-5に、正の分布量の実測値と学習器(P)の推定値の関係を示す。分布量の低い側で過大評価、高い側で過小評価になる傾向はあるものの、全体の傾向はよく再現されている。表-1に、全データの80%をトレーニングデータ、20%をテストデータとして10回のクロスバリデーションを行った結果を示す。10回ともAUCは0.91以上でその平均値は0.95であった。このことから、XGBoostが未知のデータに対しても高い推定能力を示すことが確認された。

(2) 時間変動および空間分布の再現性の検討

図-6に、分布量の水平分布の実測値を示す。カタクチイワシの大まかな分布傾向は、7月に分布量が最も多い海域は知多半島西岸沿いと伊勢湾奥の四日市沖で、三河湾口と知多半島東岸がそれに続いていた。分布の多い海域は、8月に湾央、9月には湾奥となった。図-7にXGBoostによる推定値(B×P値)を示す。漁獲量実測値の水平分布の季節変化はXGBoostを用いた推定でもおおむね再現されていることがわかる。表-2に、実測値と推定値の一週間集計漁獲量を示す。2014年は他の年に比べ漁獲量が多く、8月に最大であった。2015年は7月の漁獲量が最も多く、8月に一旦減少したのち9月に回復した。

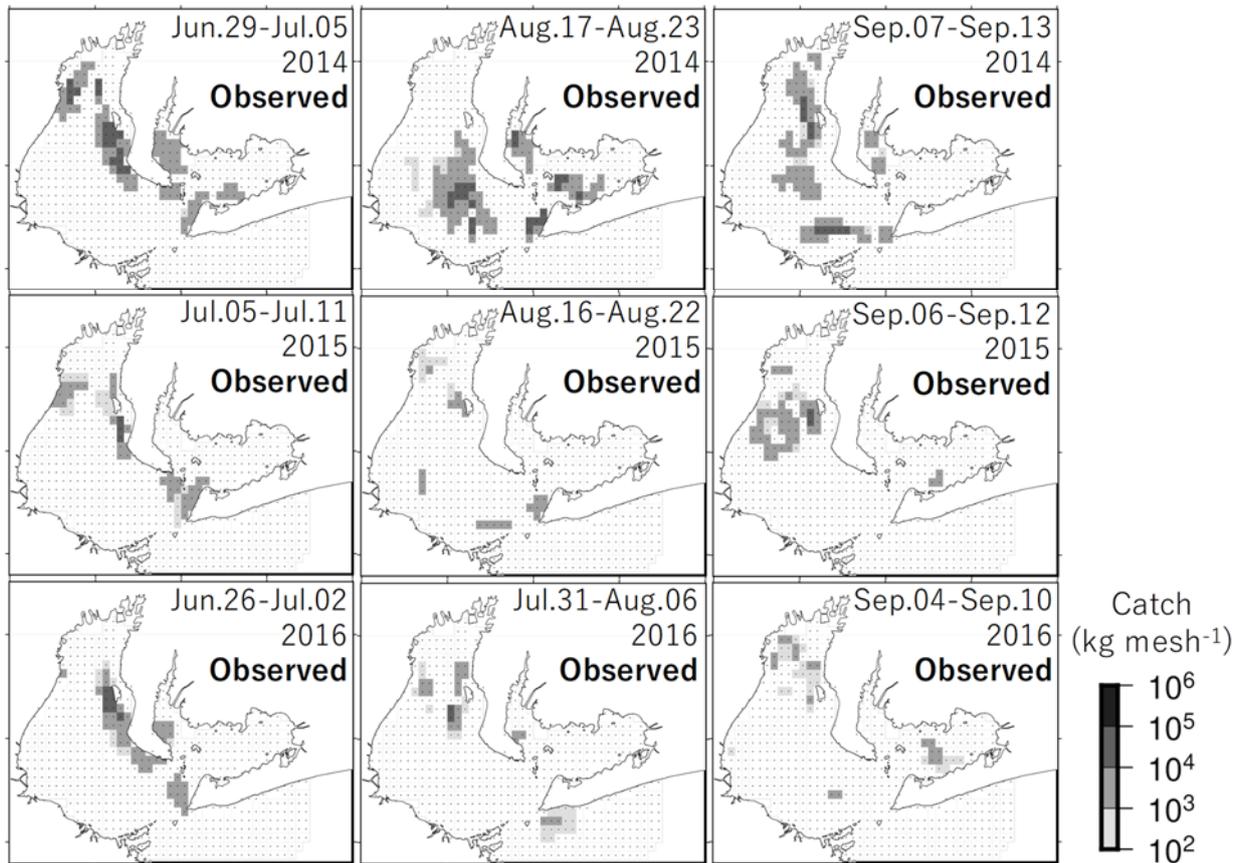


図-6 漁獲量水平分布(実測値, kg mesh⁻¹)

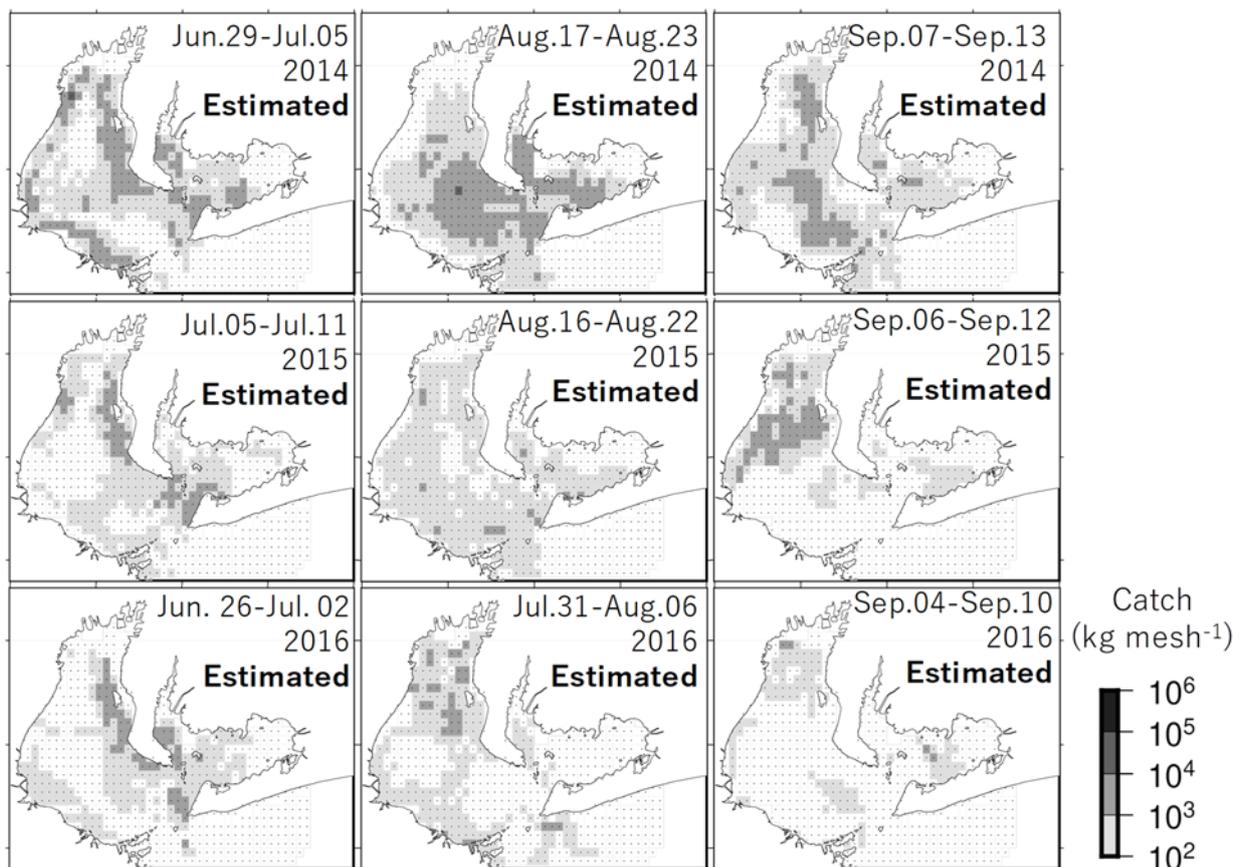


図-7 漁獲量水平分布(推定値, kg mesh⁻¹)

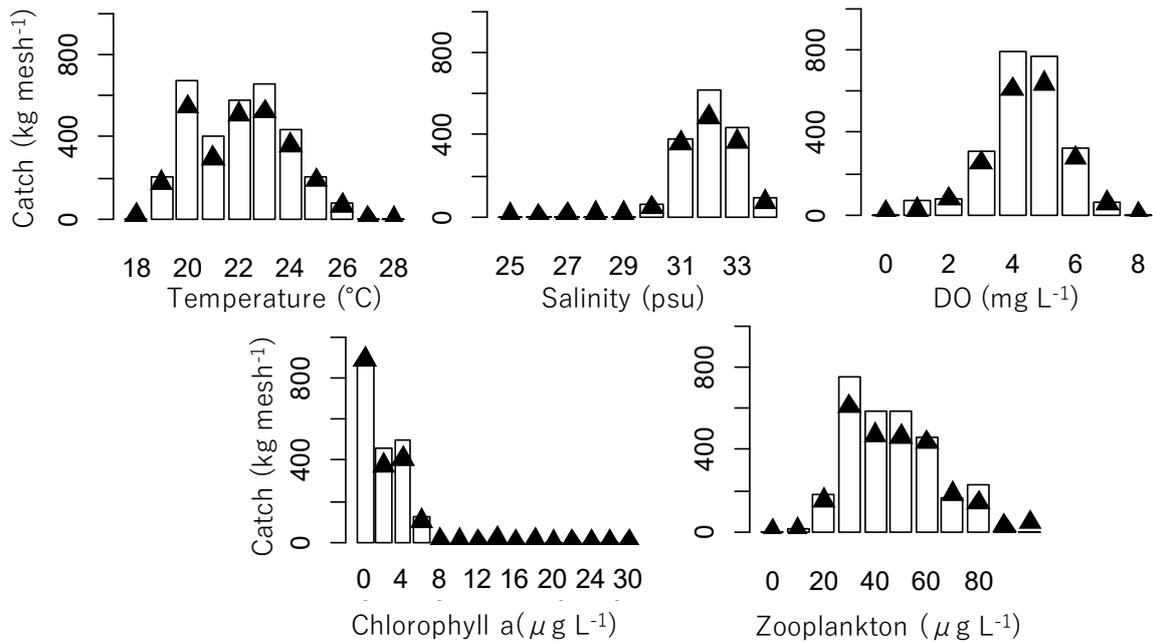


図-8 各特微量と漁獲量の関係、(各特微量の階級ごとにメッシュあたりの平均漁獲量を示した。□は実測値、▲は推定値)

表-2 各月の1週間分の漁獲量 (kg) 集計

year	month	Observed	Estimated
2014	7	563,120	386,946
	8	671,615	525,369
	9	420,255	269,840
2015	7	200,520	179,532
	8	69,040	121,306
	9	177,765	160,792
2016	7	246,200	202,526
	8	88,400	144,637
	9	50,330	45,748

表-3 各特微量の寄与度

Binary (B)			Positive (P)		
Order	Item	Gain	Order	Item	Gain
1	Zoo	0.28	1	T	0.24
2	Chla	0.21	2	Chla	0.22
3	DO	0.19	3	Zoo	0.22
4	T	0.16	4	DO	0.19
5	S	0.15	5	S	0.12

2016年は7月の漁獲量が最も多く、時間が経つにつれ減少した。こうした変動の傾向は推定値でも再現された。

(3) 環境要素と漁獲量の関係

図-8に、環境要素ごとに実測値と推定値 (B×P値) の分布量の関係を比較した結果を示す。実測値の分布の特徴は推定値でも同様に表れており、定性的ではあるが説明変数と目的変数の関係がモデルによってよく再現されていることがわかる。表-3に、学習器(B),(P)それぞれに対する各特微量の寄与度を示す。寄与度は学習器(B)で動物プランクトン量(Zoo)、クロロフィルa濃度(Chla)、溶存酸素濃度(DO)が、学習器(P)で水温(T)、クロロフィルa濃度(Chla)、動物プランクトン量(Zoo)が高かった。

いずれもクロロフィルa濃度(Chla)と動物プランクトン量(Zoo)の寄与度が高いことから、カタクチイワシが主として餌を求めて伊勢・三河湾内を移動・回遊していること、貧酸素水塊の状況も水温とともにそれに影響を与えていることが示唆された。

4. 結論

近年、環境破壊や乱獲への反省から、持続可能な社会への関心が高まってきている²⁾。持続可能な社会の達成のためには自然的、人為的な環境の変動が資源に与える影響を正確に予測・評価する手法が不可欠である。本研究で用いたXGBoostによる漁獲量推定モデルは環境変動に伴う浮魚群集の空間分布量やその変動の推定に活用可能であり、こうしたニーズにこたえることのできる将来有望な方法である。

謝辞：本研究に使用した諸データの一部は国土交通省中部地方整備局による伊勢湾漁業影響調査の一環として得られたものであり、関係者の方々および調査を実施した、いであ・全国水産技術者協会JV、さらに調査に協力頂いた愛知県漁業協同組合連合会の皆様に深く感謝申し上げます。なお、本研究では作図の一部にグラフィックソフトウェア GMT (Generic Mapping Tools)²³を使用した。

参考文献

- 1) Whitehead, R.J.P., Nelson, G.J. and Wongratana, T.: *Engraulis japonicus*. *FAO Species catalogue*, Vol.7, Clupeoid fishes of the world(suborder CLUPEOIDEI; Part II, Engraulidae), pp. 318-319, 1985.
- 2) 中村元彦・岡田秋芳: 遠州灘西部および伊勢・三河湾におけるカタクチイワシシラスの漁場への来遊様式. 愛知県水産試験場研究報告, 8, pp. 29-36, 2001.
- 3) 吉田司・大竹由華・芝修一・前田隼平・原武史・鈴木輝明: 1978-2003 年における伊勢湾でのカタクチイワシ現存量と動物プランクトン密度, 塩分の関係. 水産海洋研究, 80(3), pp. 207-213, 2016.
- 4) 東海農政局: 第 64 次東海農林水産統計年報 (平成 28~29 年), <http://www.maff.go.jp/tokai/tokei/nenpo/64.html> 参照 2019-4-17.
- 5) 鈴木翔太・多部田茂・中谷祐介・中村義治: 魚類動態モデルを利用したマアナゴ漁業シミュレータの開発. 土木学会論文集 B2(海岸工学), 69(2), pp. I_1201-I_1205, 2013.
- 6) 岡村寛・市野川桃子: 水産資源学における統計モデリング. 統計数理, 64(1), pp. 39-57, 2016.
- 7) 松原雄平・野田英明: ニューラルネットワークによる生態系環境評価システムの開発. 海岸工学論文集, 41, pp. 1136-1140, 1994.
- 8) The R Project for Statistical Computing: <https://www.r-project.org/> 参照 2019-05-01.
- 9) Chen, T., Guestrin, C.: XGBoost: A scalable tree boosting system. *KDD '16 Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785-794, 2016.
- 10) XGBoost R Tutorial: <https://xgboost.readthedocs.io/en/latest/R-package/xgboostPresentation.html> 参照 2019-05-01.
- 11) 土田修二: 沿岸性魚類の温度選好に関する実験的研究. 海洋生物環境研究所研究報告, 4, pp. 11-66, 2002.
- 12) Oozeki, Y., Takasuka, A., Kubota H. and Barange, M.: Characterizing spawning habitats of Japanese sardine (*Sardinops melanostictus*), Japanese anchovy (*Engraulis japonicus*) and Pacific Round Herring (*Etrumeus teres*) in the northwestern pacific. *California Cooperative Oceanic Fisheries Investigations Reports*, 48, pp. 191-203, 2007.
- 13) 山本昌幸・片山知史: 1995 年の瀬戸内海燧灘東部におけるカタクチイワシとマイワシの食性の比較. 水産海洋研究, 76(2), pp. 66-76, 2012.
- 14) 吉田幹英・池内仁・杉野浩二郎・神菌真人: 筑前海沿岸域のカタクチイワシの食性. 福岡県水産技術海洋センター研究報告, 9, pp. 19-24, 1999.
- 15) 野々田得郎: パッチ網によるカタクチイワシ漁とマイワシ漁との間にみられる差異について. 三重大水産研報, 12, pp. 79-89, 1985.
- 16) 永尾謙太郎・田中陽二・中田喜三郎・恩田充: 伊勢湾での詳細な観測データに基づく微生物を含めた物質循環解析. 土木学会論文集 B2(海岸工学), 71(2), pp. I_1411-I_1416, 2015.
- 17) 内藤了二・澤田玲: 伊勢湾再生への取り組み～伊勢湾シミュレータの開発について～. 平成 23 年度国土交通省国土技術研究会プログラム, 2011.
- 18) Lo, N.C.H., Jacobson, L.D. and Squire, J.L.: Indices of relative abundance from fish spotter data based on delta-lognormal models. *Canadian Journal of Fisheries and Aquatic Sciences*, 49, pp. 2515-2526, 1992.
- 19) Friedman, J., Hastie, T. and Tibshirani, R.: Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors). *The Annals of Statistics*, 28(2), pp. 337-407, 2000.
- 20) Smyth, G., Jørgensen, B.: Fitting tweedie's compound Poisson model to insurance claims data: Dispersion Modelling. *ASTIN Bulletin*, 32(1), pp. 143-157, 2005.
- 21) Fawcett, T.: An introduction to ROC analysis. *Pattern Recognition Letters*, 27, pp. 861-874, 2006.
- 22) United Nations: Transforming our world: the 2030 Agenda for Sustainable Development. <https://www.mofa.go.jp/mofaj/gaiko/oda/sdgs/pdf/000101402.pdf> (外務省による仮訳) 参照 2019-05-15.
- 23) Wessel, P., Smith, W.H.F., Scharroo, R., Luis, J., Wobbe, F.: Generic Mapping Tools: improved version released. *EOS*, 94(45), pp. 409-420, 2013.

(2019.3.13 受付)

ESTIMATION OF PELAGIC FISH DISTRIBUTION IN INNER BAY WITH MACHINE LEARNING

Takaharu HAMADA, Tsukasa YOSHIDA, Hiroshi OKAMURA,
Takeshi HARA and Teruaki SUZUKI

Pelagic Fishes are the main target species in the inner bay, and the fluctuation of their catches and the formation of fishing grounds associated with natural and anthropogenic environmental changes are of great concern. In this study, we construct a statistical model to estimate the distribution of floating fishes in the inner bay using Gradient Boosting, which is one of the highly efficient machine learning techniques. Then we estimate catch of anchovy (*Engraulis japonicus*) in Ise Bay and Mikawa Bay. The model reproduced distribution pattern of anchovy well.