

境界値過剰データに対する時空間モデリング: 海洋資源調査データへの適用

菅澤 翔之助¹, 中川 智之², H. Solvang³, S. Subbey³, S. Alrabeei³
¹東京大学 空間情報科学研究センター, ²東京理科大学 理工学部,
³Marine Mammals Research Group, Institute of Marine Research
 連絡先: <sugasawa@csis.u-tokyo.ac.jp>

(1) **動機:** 比率データを解析する方法としては二項分布に基づいた回帰分析を行うのが一般的である。しかし、海洋資源の時空間データでは、0 や 1 といった比率の境界値をとる割合が非常に多い傾向にあり、二項分布のモデルが必ずしも適合しない。このような境界値過剰の性質を無視した解析は、誤った分析結果を導いてしまう危険性があるため、境界値過剰の性質を考慮した時空間モデリングの方法論は必要不可欠である。

(2) **方法:** このような境界値過剰の性質に対応するために、二項分布と境界上に値をとる 1 点分布の混合分布を考える。今 $t = 1, \dots, T$ を時間のインデックスとし、 $i = 1, \dots, N_i$ を各時点におけるサンプルのインデックスとする。 y_{it} を応答変数、 n_{it} をトライアルの数、 x_{it} を説明変数のベクトル、 s_{it} をデータが得られた位置とする。また、 $z_{it} \in \{-1, 0, 1\}$ を潜在変数として、

$$P(y_{it} = 0 | z_{it} = -1) = 1,$$

$$P(y_{it} = n_{it} | z_{it} = 1) = 1,$$

$P(y_{it} = y | z_{it} = 0) = \text{Bin}(y; p_{it})$, $y = 0, 1, \dots, n_{it}$ とモデル化する。ただし、 $\text{Bin}(\cdot; p_{it})$ は成功確率 p_{it} の二項分布の確率関数である。さらに、 z_{it} および p_{it} を以下のようにモデル化する。

$$p_{it} = \frac{1}{1 + \exp(-\psi_{it})}, \quad \psi_{it} = x_{it}^t \beta + u(s_{it}) + v_t$$

$$z_{it} = \begin{cases} -1 & (\text{if } g_{it} \leq -1) \\ 0 & (\text{if } -1 < g_{it} < 1) \\ 1 & (\text{if } g_{it} \geq 1) \end{cases}$$

$$g_{it} = x_{it}^t \gamma + \xi(s_{it}) + \eta_t + e_{it}$$

ここで、 $e_{it} \sim N(0, \delta^2)$ であり、 $u(s_{it})$ と $\xi(s_{it})$ は時間に依存しない空間効果を表し、 v_t と η_t は時間効果を表す。空間効果は Gaussian predictive process、時間効果は random-walk を用いて推定を行う。モデルのパラメータに事前分布を導入し、ベイズ法による推定を考える。そのために効率的なマルコフ連鎖モンテカルロ法のアルゴリズムを開発した。

このモデルをノルウェー王立海洋学研究所が毎年行うバーレンツ海域における海洋資源調査のデータに適用した。本稿では、2014 から 2019 年の各年における、400 箇所を観測ステーションで観測されたシシャモのうち、体長 14 センチ以上の産卵が可能な成魚の個体数 (y_{it}) を対象とした。また、各

年において調査が実施された日付とサンプル地点での様々な深さの海水温度の情報が得られている。本解析の目的は、大きなサイズのシシャモの割合を海面温度や調査を行った位置情報によって説明することである。一方、このデータの特徴として、単純な比率 (y_{it}/n_{it}) のスケールにデータを変換したときに 0 と 1 の境界値が比較的多いことがわかった。そのため、境界値過剰を考慮した今回の提案モデルを使う必要がある。

(3) **結果:** 提案モデルを用いてサイズの大きいシシャモの生息割合を推定した。その空間分布が図 1 である。この結果から生息割合は年々多少の変化がある一方、空間的には北東地域における生息割合が高いことが確認できる。回帰係数の結果から、海水温度が生息割合に有意な影響を与えていることもわかった。さらに、将来の海水温度の上昇や下降のシナリオに対する生息確率の将来予測も行った。これらの結果の詳細については当日報告する。

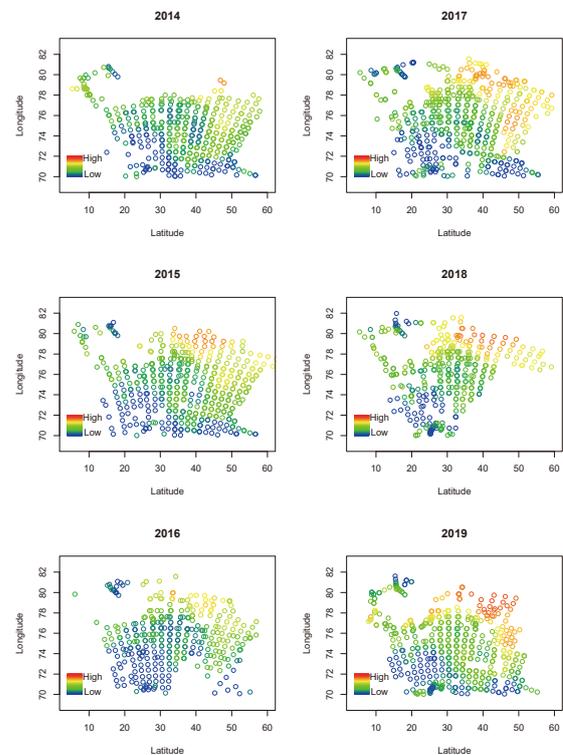


図 1: 提案モデルによって推定されたサイズの大きいシシャモの生息割合の空間分布