— 総 説 —

鳥類モニタリングデータの使い方と集め方を考える：カウントデータの誤差とつきあう技術

藤田 剛 *

東京大学大学院農学生命科学研究科 〒113-8657 東京都文京区弥生 1-1-1

(2010年11月15日受付；2010年12月19日受理)

キーワード：カウントデータ，観察誤差，増減判定，プロセス誤差，分布推定.

日本鳥学会誌

*Japanese Journal of
Ornithology*

© The Ornithological Society
of Japan 2011

Go Fujita*. 2011. How to manage uncertainties of count data from avian monitoring surveys: a brief review. *Jpn. J. Ornithol.* 60: 3–11.

Abstract. Among various types of monitoring data on avian populations, count data are among the most common. Such bird count data play an important role in conservation activities of target bird species and other organisms living in the same habitats. Count data on birds, however, often contain serious uncertainties, such as environmental covariates affecting real bird densities and detection rates of individual birds. Here, for professional and amateur Japanese ornithologists, I have made a brief review of ways to manage these uncertainties within the count data, i.e. a numerical simulation using a modified Ricker equation, or using hierarchical models. These approaches provide a reliable way of estimating the process and measurement errors of count data separately, and of inferring population parameters that can be used for evaluating population trends. In addition, using those methods, we may control for uncertainties generated by environmental covariates, including the effects of differences in observer experience. Because we have a series of long term, nationwide count-data of birds in Japan, it will be useful to apply these models to evaluating population trends and/or spatial patterns at large scales, throughout Japan, or even East Asia.

Key words: Count data, Measurement error, Process error, Evaluation of population trend, Inference of spatial pattern.

Laboratory of Biodiversity Science, School of Agriculture and Life Science, University of Tokyo. Yayoi 1-1-1, Bunkyo-ku, Tokyo 113-8657, Japan.

はじめに

鳥類モニタリングの重要な役割のひとつに、監視対象とする種の増減を判定することと、それらの鳥類にとって重要な生息地タイプや地域を明らかにすることが挙げられる (Gregory *et al.* 2005). これらの役割を主に担うのは、国全域など広域に

わたるモニタリングで得られる個体数のカウントデータである。欧米では、これらのデータにもとづき、保全上重要な生息地や保全優先種の選定が行われている (Gregory & Strien 2010 の総説)。

しかし、このカウントデータには誤差という重要な問題が存在する。たとえば、大雨や濃霧など調査日の天候によって確認できる鳥の個体数は大きく変化する。その影響を無視して、わずか数年のデータで個体数の増減を判定すると、たまたま調査日が晴天だった年のあとに悪天候の年が続いた場合、実際に個体数は変化していないにもかかわらず、減少していると判断してしまうことがある。同様に、経験豊かな調査員の後で、まだ調査に慣

* E-mail: go@es.a.u-tokyo.ac.jp

doi:10.3838/jjo.60.3

JOI:JST.JSTAGE/jjo/60.3

この論文は、2010年1月9日に東京大学で開催されたシンポジウム「鳥類モニタリングデータの集め方と使い方を考える」(主催：日本鳥学会企画委員会)の講演内容に基づくものである。

れておらず鳥の発見率の低い調査員が調査をした結果をそのまま使うと、やはり誤って減少傾向にあるという判断をくだす可能性が高まる。さらに、このような天候や調査員の特性が場所によってちがうことで生じるデータ精度の地域差は、たぶん普通に起っているだろう。

これらの誤差問題は、とくに全数調査が困難な生息地に生息する種、たとえば森林性小鳥類や農地のシギ・チドリ類などのモニタリングにとって、重要な問題となる。日本では広域モニタリングの解析が遅れている（天野 2009）。その要因のひとつとして、筆者はこのような誤差が、研究者や保全関係者にとって非常にやっかいなものとしてとらえられてきたことがあると考えている。

しかし近年、この状況は改善されつつある。ベイズ推定などの統計手法を生態学に応用する例が増え、これらの複雑な誤差をあつかえる手法もより身近なものになってきた。加えて、これら新しい統計モデルの計算やその有効性を数値実験で調べることが簡単にできるソフトウェアも普及し、鳥や生態学の調査にかかわる者にも、新しい手法を実践できる状況が整ってきた。しかし、日本語でこの手法を紹介した例はまだ限られている（天野 2010）。

本総説では、鳥類モニタリングデータとしてもっとも広く集められている個体数カウントデータに注目し、モニタリングデータの応用として使われる可能性の高い、1) 個体数増減の判定と、2) 増加率や個体数の場所間比較の2つを取り上げ、これらの2種類の応用を実践する際に関係する誤差にはどのようなものがあり、それらの誤差をコントロールする技術を提案した研究を紹介する。

カウントデータの誤差

カウントデータ誤差の問題を具体的に示すため、調査を担当した観察者が3人いる状況を想定したカウントデータを、コンピュータ上で実験的に生成した（図1上）。このデータ発生は、統計用フリーソフトウェア R（R Development Core Team 2010）をもちい、「増減判定に関わる誤差の推定」の節で説明する Hovestadt & Nowicki (2008) の方法を単純化して行った。この計算で用いた真の個体数（1年目の個体数：100）は20年間を通して増加しているのだが（増加率：1.02；図1下）、観察者の鳥発見率にちがいがあすることで、1年目から10年目まで個体数が増加した後、20年目まで減少しているように見える。ここでは、真の個体

図1. 観察者が3回交代した状況を想定し、実験的に発生したカウントデータ（上図）。計算に用いた真の個体数は20年を通して増加しているが（増加率1.02；下図）、観察者の影響を無視すると、個体数は10年目まで増加したあと、20年まで減少しているように見える。発見率は、例えば大雨や濃霧などで見通せる範囲が変化することを想定し、年によってばらつくものとした（正規分布、平均値：0；標準偏差：0.1）。また、真の個体数も、気候変動などを想定しばらつかせた（ポアソン分布）。Fig. 1. Artificially generated count data simulating observer effects (top). Different color dots represent data from different observers with different detection rates (observer A: 0.5, B: 0.7, C: 0.3). Although the true number of birds I used in the model follows an increasing pattern (bottom; increasing rate: 1.02), the count data apparently shows a unimodal pattern without considerations of observer effects. In this simulation, I incorporated random effects to detection rates (Gaussian distribution: mean = 0; σ = 0.1) and to true population size (Poisson distribution) to simulate, for instance, weather conditions affecting detection rates, and true population size in each year.

数変化を既知のものにするため実験的に発生したデータをもちいたが、実際にも観察者による発見率のちがいが個体数増減推定に大きな影響を与えることを示した例は少なくない（たとえば、Sauer *et al.* 1994, Link & Sauer 1998）。つまり、カウントデータから増加率を正確に知るためには、たとえば観察者の特性（この例では発見率のちがい）などで生じる誤差の大きさも、推定する必要があるのだ。

このようなカウントデータ解析にあたって注意

すべき誤差は、大きくプロセス誤差と観察誤差の2つに分けられる。プロセス誤差とは、現実の自然界に存在する確率のプロセスで生じるカウントデータのばらつきである。たとえば、偶然気温が高かったり大雨が降ったりするなど、天候の影響で鳥の分布が変化して生じるばらつきは、この誤差に含まれる。また、ある年、たまたま死亡数が多かった結果個体数が少なくなるといった人口学的なゆらぎで生じるカウント数のばらつきも、この誤差に含まれる。

一方、観察誤差とはデータ収集の過程で生じるカウントデータのばらつきである。たとえば、図1で挙げたような、調査員間の発見率の差によって起るカウントのばらつきや、調査した日の天候のちがいで発見率が変化して生じるカウント数のばらつきは、この観察誤差に含まれる。カウントデータから個体数の増減を判定する場合、先に述べたプロセス誤差と観察誤差のそれぞれが個体数の増減判定にどう影響するのかを把握しないと、実際には減少している種を増加していると判断してしまうなどの過ちを犯す危険がある。

保全や生態学の研究には、どの生息地、あるいはどの地域で対象種の個体数が多いのかといった、いわゆる分布推定を目的とするものが多い。それらの結果から、対象種の保全上優先すべき生息地や地域を明らかにすることができる。しかし、このカウントデータの場所間比較にも、困難な問題があることが指摘されている (Hovestadt & Nowicki 2008)。たとえば、生息地や地域の環境要因が個体数と同時に鳥の発見率にも影響するといった問題がある。この関係を無視すると、真の個体数が多い生息地を保全上優先すべきところを、単に発見率の高い場所を優先して保全してしまうなどの過ちを犯す可能性がある。場所によって調査員の発見能力にちがいがあると、この影響はさらに複雑なものになる。

増減判定に関わる誤差の推定

上述のように個体数増減率の推定は、カウントデータ活用のもっとも重要な役割のひとつである。たとえばEUでは、6年間にわたる年減少率が1%あるかどうか「保全が望まれる種」選定の基準として使われている (European Commission 2005)。また、英国では、カウントデータから推定した個体数指数 (Landres *et al.* 1988, 天野 2009, 2010, Gregory & Strien 2010 を参照) にもとづき、過去40年間でもっとも減少した種の多かった農地が、

保全対象として重視されるなど、保全優先度の高い生息地選定の基準としても役立っている (Butler *et al.* 2007, Gregory & Strien 2010)。

個体数増減評価のために、内的自然増加率 r や環境収容力 K 、さらに環境悪化による環境収容力の低下率などの個体群パラメータの推定が行われることが多い。これらのパラメータ推定のため、理論生態学者は多くのモデルを構築している (たとえば, Sibly *et al.* 2005)。しかし、残念ながら、現実個体群ではこれらの理論モデルの前提を満たすものは少ない。その第1の理由として、現実の個体群には人口学的なゆらぎや環境変動による攪乱が存在することが挙げられる。理論モデルの多くは決定論的なもので、これらの攪乱によって生じるプロセス誤差を組み込んではいない (Hovestadt & Nowicki 2008)。第2の理由としては、個体群サイズ自体を測定することは非常に困難であるため、増減推定にはカウントデータが多く使われることが挙げられる。このため、カウントデータには、個体群の一部をサンプリングすることで生じる確率的なばらつきと観察誤差が含まれることになる。

また、解析に使われる統計モデルでは、プロセス誤差と観察誤差とを分けて推定されないことが多い。プロセス誤差と観察誤差を混同すると、実際には減っているのに増加していると判断したり、減少率を過小評価したりする可能性がある (Holmes 2001, McNamara & Harding 2004)。したがって、観察誤差をプロセス誤差と分け、その影響を推定することが望ましい。

Hovestadt & Nowicki (2008) は、プロセス誤差と観察誤差が増加率推定にどう影響するのかを、数値シミュレーションで調べている。その際、解析にもちいる年数の影響も加えることで信頼できる増減評価に必要な年数も明らかにした。どれくらいの期間にわたるカウントデータがあれば、どの程度の精度で増減評価ができるのか、という問題もモニタリング解析では重要な問題である。

Hovestadt & Nowicki (2008) は、まず修正リッカー式 (たとえば, Meir & Fagan 2000 を参照) をもちい、さまざまな程度のプロセス誤差と観察誤差を設定して実験的に個体数データを生成した。この修正リッカー式は、 t 年の個体数 N_t とその翌年 $t+1$ 年の個体数 N_{t+1} の関係を示したモデルである。

$$N_{t+1} = N_t \cdot e^{r(1 - \frac{N_t}{K(1-\alpha)}) + \varepsilon_p}$$

ここで e は自然対数の底 (≈ 2.718)、 r は内的自然

増加率, K は環境収容力, α は人為開発など環境劣化による環境収容力の減少を示している. 右辺の r に $(1-N_t/K(1-\alpha)^t)$ を乗じることによって, 増加率に密度依存性をもたせている. つまり, ある年の個体数 N_t がその年の環境収容力 $K(1-\alpha)^t$ に近づくにしたがって $(1-N_t/K(1-\alpha)^t)$ は 0 に近づく. そして, ε_p がプロセス誤差を示しており, 平均値 0, 標準偏差 σ_p の正規乱数で与えられる.

$$\varepsilon_p \approx N(0, \sigma_p)$$

したがって, このモデルで生成される個体数の増加率は, 密度効果だけでなくランダムに生じる ε_p の影響が加わった形で決まる.

つぎに, この生成データをもとに, 以下の式をもちいて観測データを生成する.

$$N_{t,m} = N_t \cdot e^{\varepsilon_m}$$

$N_{t,m}$ は t 年の個体数測定値を指し, 上式で生成した t 年の個体数 N_t に e^{ε_m} を乗じた値である. そして, ε_m が観察誤差で, 以下のように平均値 0, 標準偏差 σ_m の正規乱数で与えられる.

$$\varepsilon_m \approx N(0, \sigma_m)$$

最後に, そのカウントデータにもとづき, 次の統計モデルで増加率とその誤差を推定した. この統計モデルの推定には, R の非線形最小二乗法 `nls()` を使い, 以下の式に回帰した.

$$Y_t = \ln\left(\frac{N_{t+1,m}}{N_{t,m}}\right) = r_f \left(1 - \frac{N_{t,m}}{(K - \Delta_f)(1 - \alpha_f)^t}\right) + \varepsilon_T$$

r_f と K が内的自然増加率と環境収容力, Δ_f が真の K からの回帰残差, ε_T が全誤差である. さまざまな調査期間の, さまざまなプロセス誤差と観察誤差の値で, 真の増加率と, 統計モデルによる観察誤差の関係を調べた (各条件の繰り返し 5,000 回, σ_p : 0.1–0.5, σ_m : 0–0.75, 期間: 10–40 年, $r = 1$, $K = 1,000$, α : 0–0.75).

まず, 環境収容力の劣化 α を含まないモデルでは, ほぼすべての条件で増加率や他の変数が推定可能だったが, α を組み込んだ場合に推定不可能な場合が生じ, 観察誤差が大きく ($\sigma_m > 0.5$), 調査期間が短い (<15) 条件で推定できなかった.

観察誤差の方が, プロセス誤差より増加率 r 推定への影響が大きかった. つまり, 観察誤差が小さいときは増加率を過大評価し, それが大きくなると過小評価する傾向が見られた. プロセス誤差は, その大小に関わらず増加率をわずかに過小評価する傾向があった. また, 環境収容力の劣化 α の大小は, 増加率推定に影響せず, 環境収容力推定への影響も比較的小さかった. しかし, これに

観察誤差が大きいという条件が加わると増加率を過小評価する傾向があった. 一方, 調査年の長短は増加率推定に大きく影響し, 20 年より短い調査年データでは一貫して増加率を過大評価する傾向があった.

そして, 保全上重要な意味をもつ環境収容力の劣化 α の推定には, 観察誤差の大きさと調査年の長さが強く影響し, 観察誤差が大きく, かつ, 調査年が 15 年より短い場合に, 過小評価される傾向が見られた (図 2).

これらの結果から, Hovestadt & Nowicki (2008) は次のような結論を導いている. 1) 調査年が長い場合には環境収容力 K と生息地悪化による収容力低下の程度 α は比較的精度高く推定できる. 2) ただし, 内的自然増加率の推定には, 観測誤差が強く影響する. 3) また, 解析にもちいたデータ年数が短い場合, 増加率を過大評価するだけでなく, 生息地悪化の影響を過小評価する危険が高まる. 先に挙げたように, EU は 6 年間で年 1% の速度で減少している種を見つけることを要求している. しかし, Hovestadt & Nowicki (2008) は, この結果から過去 6 年のみのデータから年 1% の減少を評価することは困難であると警告している.

増減率や個体数の地域間比較に関わる誤差の推定

個体群の増減率や個体数が生息地や地域によってどうちがうかを比較することで, 保全優先地域や生息地をしぼることができる. 特定の生息地などで増加率や生息密度が高くなっている生態学的要因を解明できれば, 減少をくいとめる方策や具体的な生息地管理の提言が可能になる.

これらの地域間比較のため, 生息地や地域間のカウントデータ比較が頻繁に行われる. これらは比較が目的であるため, カウントデータを相対的な個体数指数と見なし, そのまま使われることが多い. しかし, 個体数に影響する環境要因が, 発見率にも影響することは少なくないと考えられる. たとえば, 森林内の低木密度が高い場所は, 餌昆虫が多いなどの理由で鳥が高密度で営巣するが, 視界が遮られるため個体発見率が下がるというような場合が考えられる. このような, 個体数には正, 発見率には負の方向に影響を与える要因が存在するとき, カウントデータにはその両方向の影響が打ち消し合う形で含まれることになる. このカウントデータをそのまま使うと, この発見率への影響を無視することになり, 発見率のちがいによって生じたカウント結果のちがいを, 真の個体

図2. プロセス誤差 σ_p と観測誤差 σ_m 、解析に用いたデータの年数 T_m 、環境収容力劣化の度合い α 、増加率 r が、 α の推定におよぼす影響。観測誤差は右へ行くほど大きくなり、プロセス誤差は上へ行くほど大きくなる。等高線上の数値が内的自然増加率 r 。等高線は、収容力の変化が検出される ($P < 0.05$) 変数値の組み合わせを結んでいる。この等高線の右上が、収容力の変化を検出できる領域。例えば、下段右端のグラフ (観測誤差 $\sigma_m = 0.5$ 、プロセス誤差 $\sigma_p = 0.3$) では、 r が 1 の場合、収容力の変化を検出できる領域は右上 (収容力低下の度合いが大きく、年数の長い場合) に限られることが分かる。この領域は、観測誤差と測定誤差が小さい場合 (上段左端) 広がる。また、解析に用いるデータの年数が短くなると、 α が大きい場合にも検出可能な領域が狭いことが分かる。Hovestadt & Nowiki (2008) の図5を許可を得て転載。

Fig. 2. The effect of process error, σ_p , measurement error, σ_m , the length of the time series analyzed, T_m , and the magnitude of the trend, α , on the probability of finding significant evidence for the existence of a temporal decline in habitat capacity. σ_m increases from left to right, σ_p from the top to the bottom row. Contour lines demark the probability that a single NLR (at a given parameter combination) will return a significant ($\alpha \neq 0$, $P < 0.05$) estimate for a temporal trend (Fig. 5 from Hovestadt & Nowiki 2008 included with permission).

数のちがいと誤って判断する可能性がある。これは、森林に限らず、全数調査が困難な生息地のモニタリングに共通する重要な問題である。

Kéry (2008) は、二項混合モデル (Royle 2004) と呼ばれる階層モデルの一種を使い (階層モデルについては、天野 2010 に導入的な解説がある)、真の個体数への影響と発見率への影響が分離可能であることを示した。まず、ある場所で何度か繰り返して個体数を記録する調査を想定する。その繰り返し数 j 回の調査で得られるカウントデータは、真の個体数を N_i 、発見率を p とする二項分布値の j 回分の積で表されるとする。つまり、

$$g(c_i | N_i, p) = \prod_j \text{Bin}(c_{ij} | N_i, p)$$

左辺 $g(c_i | N_i, p)$ は、調査地点 i のカウントデータ値 c_i が、真の個体数が N_i 、発見率が p という条件下で得られる確率を意味する。右辺の $\text{Bin}(c_{ij} | N_i, p)$ は、二項分布の条件下で調査地点 i の j 回目のカウントデータ c_{ij} が、真の個体数 N_i 、発見率 p という条件下で得られる確率を指し、その左の記号は $\text{Bin}(c_{ij} | N_i, p)$ を j 回乗じることを意味する。

つぎに、真の個体数 N_i と環境要因の関係に注目する。これまでのモデルでは、ひとつのサンプル場所についてひとつの N_i のみが使われることが多かった。しかし、ここではこの N_i をランダム効果とみなし、 θ の大きさでばらつく確率分布と見なすこととする。ランダム効果とは、応答変数のばらつきに影響する効果を指す (Crawley 2005)。個

体数のような自然数の場合、ランダム効果として慣習的にポアソン分布が使われることが多い。ポアソン分布では、平均値 λ に依存してばらつきの大きさも決まるため、この λ だけ推定すれば良いという利点もある。そこで、ある場所に生息する個体数は、平均値 λ のポアソン分布をとることとし、その分布は、以下のように k 種類の環境要素 x との線形の関係にあるとした。

$$\log(\lambda_i) = \alpha + \sum_{k=1}^K \beta_k x_{ik}$$

また、発見率 p (0 から 1 の値をとる) は上で注目した環境要素 x との関係のみで決まるとし、以下のようなロジット式の関係にあるとした。本来、この p には複数の環境要素が影響する可能性が高いが、ここでは計算の都合上、1 種類のみにしていく。

$$\text{logit}(p_{ij}) = \gamma + \delta x_{ij}$$

以上の3つのモデルを、実際のデータにあてはめることで、知りたい変数 ($\alpha, \beta, \gamma, \delta$) を推定することができる。Kery (2008) は、その推定のためにマルコフ連鎖モンテカルロ法 (MCMC 法、動的モンテカルロ法とも呼ばれる) によるベイズ推定をもちいている。ベイズ推定は、推定したい変数にある候補値を与え、手持ちのデータと対比し、その候補値がこれまでに分かっている推定値より「確からしい」場合、ある確率で新しい候補値を採用する作業を繰り返すことで、推定値を真の値に近づけようとする方法である。そして、この確からしさの判断に、条件つき確率の式であるベイズの公式を使う。

この際、推定したい変数の種類が多く、かつ各変数のとる値の範囲も大きい場合、これらの変数候補値の組み合わせが膨大になり、すべての組み合わせの確からしさの計算は不可能になる。この膨大な組み合わせの中から、効率よく真の値に近づけるよう候補値をサンプリングするのがMCMC法である。筆者の知る範囲では、ベイズ統計の入門者向けの数学的な和文解説として伊庭 (2003) が分かりやすく、MCMC法の一種、ギブス・サンプラー Gibbs sampler などの説明も明解である。このベイズ推定を実際に行うには、フリーソフトウェア WinBUGS (<http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml>) が便利である。MacCarthy (2007) には、ベイズ統計の解説以外に、このWinBUGSの説明がある。

ここでは、個体数と環境要因の関係に関わる α と β 、発見率と環境要因の関係に関わる γ と δ をこの方法で推定している。上のモデルで、個体数と発見率への影響が別々に、どの程度の精度で推定できるのかを調べるため、Kery (2008) は、実際のモニタリングデザインにしたがった方法で、コンピュータ上で α から δ の4種類の変数を設定して生成したデータを使った。シミュレーションにもちいた実際のモニタリングデータは、スイス全域を対象に行われているスイス繁殖鳥類調査によるもので、1999年から面積1 km²の調査区画268を対象にそれぞれ繁殖期間中に3回ずつ繰り返しのテリトリーマッピングが行われている。

スイスではさまざまな標高での調査が行われているため、個体数と発見率に影響する環境要因として標高に注目した。この際、個体数は標高が上がるにつれて減少し、発見率は標高が上がると増加する関係を想定した。具体的には、 $\alpha = 2$, $\beta = -0.5$, $\gamma = 1$, $\delta = 2$ という値をもちいている。そして、さまざまな標高に応じて、ポアソン乱数を発生させ、真の個体数と観察値を計算した。このデータ生成にはRをもちいている。

シミュレーションにもちいた α などの係数をもちいると、各コドラートの個体数は標高が高くなるにつれて凹型に減少し、発見率はS字型に増加した。この平均値の変化傾向に、ポアソン誤差による個体数データのばらつきと発見率の影響が重なると、中標高でもっとも個体数の多くなるひと山型になった。この一山型パターンは、標高が高くなるほど減少するという、真の個体数変化パターンとはちがうものである。

にも関わらず、Keryのもちいた混合モデルは、標高が個体数と発見率におよぼす影響を正確に推定した ($\alpha = 2.01 \pm 0.05$ s.d. [真の値: 2], $\beta = -0.51 \pm 0.06$ s.d. [真の値: -0.5], $\gamma = 0.99 \pm 0.10$ s.d. [真の値: 1], $\delta = 2.01 \pm 0.08$ s.d. [真の値: 2])。つまり、混合モデルが、同じ環境要因によって真の個体数と発見率が同時に変化するという複雑な条件でも、発見率の影響を推定した上で真の個体数と環境要因の関係を推定可能であることが示されたのである。Kery (2008) は、この手法が、個体識別や発見個体の正確な位置情報など手間のかかるデータを必要としない、有効で新しい方法であると結論している。

Kery (2008) のモデルでは、発見率は環境要因のみで決まる。しかし、カウントデータでは、図1の例で挙げたように、観察者のちがいが発見率

に影響する可能性も考えられる。とくに調査地点間で観察者がちがっており、かつ、各調査地点で観察者の入れ替わった年にも差があると、その影響も複雑になる。観察者ごとにその影響を推定しようとする、非常に多くの変数推定が必要になり現実的ではない (Link & Sauer 1998)。

この対策のひとつとして、混合モデルの説明変数にランダム効果として観察者の影響を組み込む方法がある (観察者の影響をランダム効果とし、推定変数の数を減らす具体的な方法については、Link & Sauer 1998 に詳しい)。Thogmartin *et al.* (2004) は、この手法によって、北米繁殖鳥類調査で得られた希少種ミズイロアメリカムシクイ *Dendroica cerulea* の 21 年にわたるカウントデータを解析している。対象としたのは、約 1,000 km × 500 km の広大な範囲にわたる乾燥林地帯の鳥類保全地域とその周辺地域を含むデータで、調査ルート数は 140、関わった観察者は 310 人におよぶ。その結果、観察者などの影響を取り除いた上で、湿地林の割合と土壌の潜在的湿り度合いが、ムシクイの個体数決定に重要であることを明らかにしている。また、降水量の少ないところでは、落葉広葉樹林のサイズが大きいと個体数が多く、降水量の多いところでは落葉広葉樹林サイズが小さい地域でも個体数が多いといった、降水量と落葉樹林サイズの交互作用も重要であることが分かった。この解析は、観察者の特性のちがい以外に、0 データが非常に多いことへの対処法や地域によって重要な環境要因がちがって見える効果も組み込むなど、広域にわたるカウントデータの解析にあたって直面する他の問題に対処している点も特徴である。この解析にも、Kery (2008) と同様にベイズ推定がもちいられている。Thogmartin *et al.* (2004) は、このモデルによる予測を分布図にし、鳥類保全地域内には、個体群全体のわずか 15% しか分布していないことを示し、保全地域の新しい設定などの対応が必要であることを提言している。

このような混合モデルや、さらに複雑な誤差構造を含む階層モデルの係数推定にベイズ推定をもちいる解析は非常に強力で (天野 2010)、近年盛んに利用され始めている (たとえば、Link & Sauer 2007, Thogmartin & Knutson 2007)。日本でも、Amano *et al.* (2010) は、1975 年から 2008 年までのシギ・チドリ類のカウントデータに基づいた全国レベルの個体数指数推定にこの推定法をもちいている。その結果、シギ・チドリ類データのよう

に年や場所によるばらつきが大きく見えるものでも、個体数指数の推定が可能であること、春に日本を通過する主要 41 種中 12 種、秋に通過する 42 種中 14 種が減少していること、さらに、減少している種の特徴は水田を利用している種や黄海に依存した種であることを明らかにしている。

ま と め

日本には世界に誇るべきモニタリングデータがある。たとえば、環境省の全国ガンカモ一斉調査は 1970 年から、シギ・チドリ類定点調査も 1971 年から形を変えつつも毎年継続されている (環境省生物多様性情報システム <http://www.biodic.go.jp/J-IBIS.html>)。また、すべての鳥を対象とした全国鳥類繁殖分布調査も、1974~78 年に 1 回目の調査が行われている。

しかし、残念ながらそれらの保全への応用は欧米にくらべ遅れていた (天野 2009)。とくに森林性鳥類など全数調査が困難な鳥のカウントデータを、個体数増減や分布推定に応用した例は非常に限られている (限られた例として、Amano *et al.* 2010; カウントデータではないが、分布域増減の解析には、Amano & Yamaura 2007; 本特集の植田 睦之・福井晶子・山浦悠一・山本 裕 2011 全国的生態観測調査「モニタリングサイト 1000」で見えてきた日本の森林性鳥類の分布状況、日鳥学誌 60 などがある)。筆者は、その原因のひとつとして、保全活動やモニタリングデータに関わる研究者が、全国一斉調査などのカウントデータについて、その誤差の複雑さを深刻にとらえすぎていた点があるのではないかと考えている。少なくとも筆者自身はそうであった。

ここでは、混合モデルなどによって、これらの誤差にも対処可能であることを示した。そして、無償でかつ利用者が改善を加え易いソフトウェアの R や WinBUGS の普及によって、これらの手法を簡単に利用できる環境も整い始めている。たとえば、上で紹介した Kery (2008) は、解析にもちいた WinBUGS のコードも公開し、このようなコードに不慣れな場合でも、同様の解析ができるようにしている。日本のモニタリングデータの解析が進み、生物多様性の保全などに活用される準備は整っているのである。本稿が、日本の広域モニタリングデータの解析や今後の調査デザインの改善などに、わずかでも役立てば幸いである。

匿名の査読者には、初稿への有益なコメントをいただいた。本論文は、2010年1月に開催したシンポジウム「モニタリングデータの集め方と使い方を考える」(日本鳥学会企画委員会主催)の発表を基礎に大幅加筆したものである。本シンポジウム実施にあたっては、綿貫 豊氏と嶋田哲郎氏にご協力いただいた。また、天野達也、笠原里恵、神山和夫、新妻靖章氏、植田睦之の各氏にはシンポジウム講演者として、藤田道男氏にはコメンテーターとして全面的にご協力いただいた。これらの方々に感謝申し上げる。

摘 要

さまざまな鳥類モニタリングデータの中で、カウントデータはもっとも良くみられるものであり、対象とする鳥類や同じ生息地の他の生物の保全活動に重要な役割を果たしてきた。しかし、カウントデータには、たとえば同じ環境要因が実際の鳥の個体数に影響すると同時に個体の発見率にも影響するといった誤差が含まれていることが多い。ここでは、不確定要素の影響を推定する技術として、修正リッカー式を用いた数値実験と混合モデルをもちいた解析例を紹介した。これらの方法によって、カウントデータに含まれる観察誤差とプロセス誤差、それぞれの影響の大きさを推定することで個体数増減評価上注意すべき点を明らかにしたり、発見率の場所によるばらつきを考慮した上で増加率や生息密度の高い場所を推定したりすることが可能になる。日本には、30年以上にわたる広域モニタリングによって蓄積された鳥類のカウントデータがある。ここに挙げた方法などをもちいることによって鳥類カウントデータ解析がさらに進むことは、日本の鳥類をはじめとする生物多様性の保全に重要な役割を果たすものと考えられる。

引用文献

- 天野達也 (2009) 水田を利用する鳥類の個体数変化と広域観測. 遺伝 **63**: 51–55.
- 天野達也 (2010) 生物個体数の指標化法. 鷺谷いづみ・宮下 直・西廣 淳・角谷 拓 (編) 保全生態学の技法: 157–178. 東大出版, 東京.
- Amano T & Yamaura Y (2007) Ecological and life-history traits related to range contractions among breeding birds in Japan. *Biol. Conserv.* **137**: 271–282.
- Amano T, Székely T, Koyama K, Amano H & Sutherland WJ (2010) A framework for monitoring the status of populations: an example from wader populations in the East Asian-Australasian flyway. *Biol. Conserv.* **143**: 2238–2247.
- Butler SJ, Vickery JA & Norris K (2007) Farmland biodiversity and the footprint of agriculture. *Science* **315**: 381–384.
- Crawley MJ (2005) *Statistics: an introduction using R*. Wiley, Hoboken. (Crawley MJ (2008) 統計学: Rを用いた入門書. 野間口謙太郎・菊池泰樹 (訳). 共立出版, 東京.)
- European Commission (2005) Assessment, monitoring and reporting of conservation status—Preparing the 2001–2007 report under Article 17 of the Habitats Directive. DG Env. B2/AR D rev3. Brussels, http://forum.europa.eu.int/Public/irc/env/monnat/library?l=/reporting_framework
- Gregory RD & van Strien A (2010) Wild bird indicators: using composite population trends of birds as measures of environmental health. *Ornithol. Sci.* **9**: 3–22.
- Gregory RD, van Strien A, Vorisek P, Meyling AWG, Noble DG, Foppen RPB & Gibbons DW (2005) Developing indicators for European birds. *Phil. Trans. R. Soc. Lond. B* **360**: 269–288.
- Holmes EE (2001) Estimating risks in declining populations with poor data. *Proc. Natl. Acad. Sci. USA* **98**: 5072–5077.
- Hovestadt T & Nowicki P (2008) Process and measurement errors of population size: their mutual effects on precision and bias of estimates for demographic parameters. *Biodivers. Conserv.* **17**: 3417–3429.
- 伊庭幸人 (2003) ベイズ統計と統計物理. 岩波書店, 東京.
- Kery M (2008) Estimating abundance from bird counts: binomial mixture models uncover complex covariate relationships. *Auk* **125**: 336–345.
- Landres PH, Verner J & Thomas JW (1988) Ecological uses of vertebrate indicator species—a critique. *Conserv. Biol.* **2**: 316–328.
- Link WA & Sauer JR (1998) Estimating population change from count data: application to the North American breeding bird survey. *Ecol. Appl.* **8**: 258–268.
- Link WA & Sauer JR (2007) Seasonal components of avian population change: Joint analysis of two large-scale monitoring programs. *Ecology* **88**: 49–55.
- MacCarthy MA (2007) *Bayesian method for ecology*. Cambridge University Press, Cambridge. (MacCarthy MA (2009) 生態学のためのベイズ法. 野間口眞太郎 (訳). 共立出版, 東京.)
- McNamara JM & Harding KC (2004) Measurement error and estimates of population extinction risk. *Ecol. Lett.* **7**: 16–20.
- Meir E & Fagan WF (2000) Will observation error and biases ruin the use of simple extinction models? *Conserv. Biol.* **14**: 148–154.
- R Development Core Team (2010) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Royle JA (2004) N-mixture models for estimating population size from spatially replicated counts. *Biometrics* **60**: 18–115.
- Sauer JR, Peterjohn BG & Link WA (1994) Observer differences in the North American breeding bird survey. *Auk* **111**: 50–62.

- Sibly RM, Barker D, Denham MC, Hone J & Pagel M (2005) On the regulation of populations of mammals, birds, fish, and insects. *Science* **309**: 607–610.
- Thogmartin WE & Knutson MG (2007) Scaling local species-habitat relations to the larger landscape with a hierarchical spatial count model. *Landscape Ecol.* **22**: 61–75.
- Thogmartin WE, Sauer JR & Knutson MG (2004) A hierarchical spatial model of avian abundance with application to cerulean warblers. *Ecol. Appl.* **14**: 1766–1799.
- 植田陸之・福井晶子・山浦悠一・山本 裕 (2011) 全国的な生態観測調査「モニタリングサイト 1000」で見えてきた日本の森林性鳥類の分布状況. *日鳥学誌* **60**: 19–34.