

# 市民参加による魚類画像データベースの構築と 学術研究への応用、そして課題

瀬能 宏 (神奈川県立生命の星・地球博物館)

北海道から沖縄まで、日本の沿岸にはダイビングポイントが点在し、毎日のようにどこかで誰かがダイビングを楽しんでいる。フィリピンやインドネシア、オセアニアなど、海外にでかけるダイバーも多い。水中カメラの普及により、ダイバーの多くは生物の撮影を趣味としている。魚類は主要な被写体であり、これまでにダイバーの手元に蓄積された魚類の画像は膨大な数になっているはずである。画像には撮影地、撮影水深、撮影年月日といった属性が付随するため、被写体を同定できれば画像を標本と同じように一次資料として機能させることができる。このような背景と観点から、神奈川県立生命の星・地球博物館では、1994年から魚類の画像収集とデータベース化を進めてきた。当初は画像提供者のほとんどがダイバーであったが、最近では釣り人やアクアリストからの提供も増えつつあり、登録された画像は2009年10月現在で80,645件に達している(図1)。

このようにして集積された画像は、例えば研究者も知らなかった未知の魚類の発見(図2)に役立つことがあるが、中でも著しい有用性を示した分野が生物地理学的研究である。ある地点における魚類相の解明には、熟練した研究者による長期間の調査が必要であり、技術的、法的、道義的な困難を伴うことも多いが、画像データベースを活用すればごくわずかな期間で数百種から千種近い網羅的目録を作成することが可能となった。実際、本州や四国の太平洋沿岸、琉球列島や伊豆・小笠原諸島の複数地点の魚類目録を完成させ、それらを相互に比較することで日本の沿岸魚類相に黒潮がどのような役割を演じているのかといったことまで明らかにすることができた(図3)。

しかしながら、このデータベースにも大きな問題がある。画像の収集は無作為であり、提供画像の選択は提

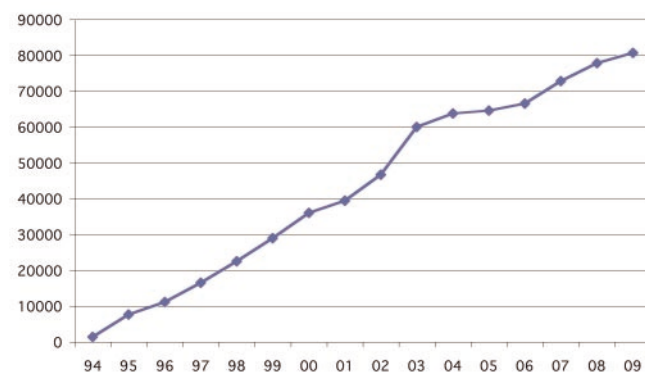


図1 画像の年度別登録実績(累計)

供者の判断に委ねられている。1種につき1点の画像提供で事が足りるため、ある地点での出現種を網羅する画像は集まりやすい。一方、出現頻度や出現期間など、ある魚種の動態を定量的あるいは経時的にとらえるような画像は集まりにくい。なぜなら、既に登録されている種の画像は必要ないだろうという提供者側の心理が働くからである。今後、例えば地球温暖化による魚類相の変化を捉えるといったテーマ性のある収集方針を打ち出し、その成果を情報発信することができれば、さらなる発展につながるかも知れない。



図2 フィリピンやインドネシアで撮影されたハゼ科魚類。ほとんどが未記載種(いわゆる新種)と考えられる。

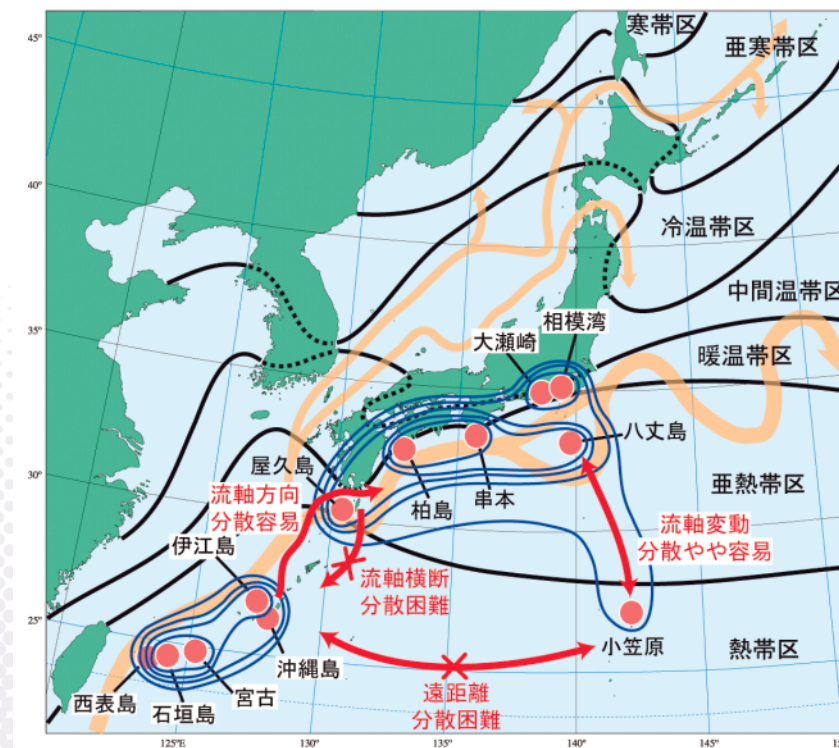


図3 日本近海の生物地理区分(黒線:西村、1992)と、調査地点(赤丸)および沿岸魚類相の類似関係(青線)。薄いオレンジの線は黒潮の流路を示す。黒潮はこれまで南方系の魚類を運んでくることだけが強調されてきたが、この図から琉球列島とそれ以外の地域の魚類相を分断するバリア(障壁)としての機能があることがわかる。



# 日本産シダ植物の系統的多様性 — 分布情報と分子情報の融合 —

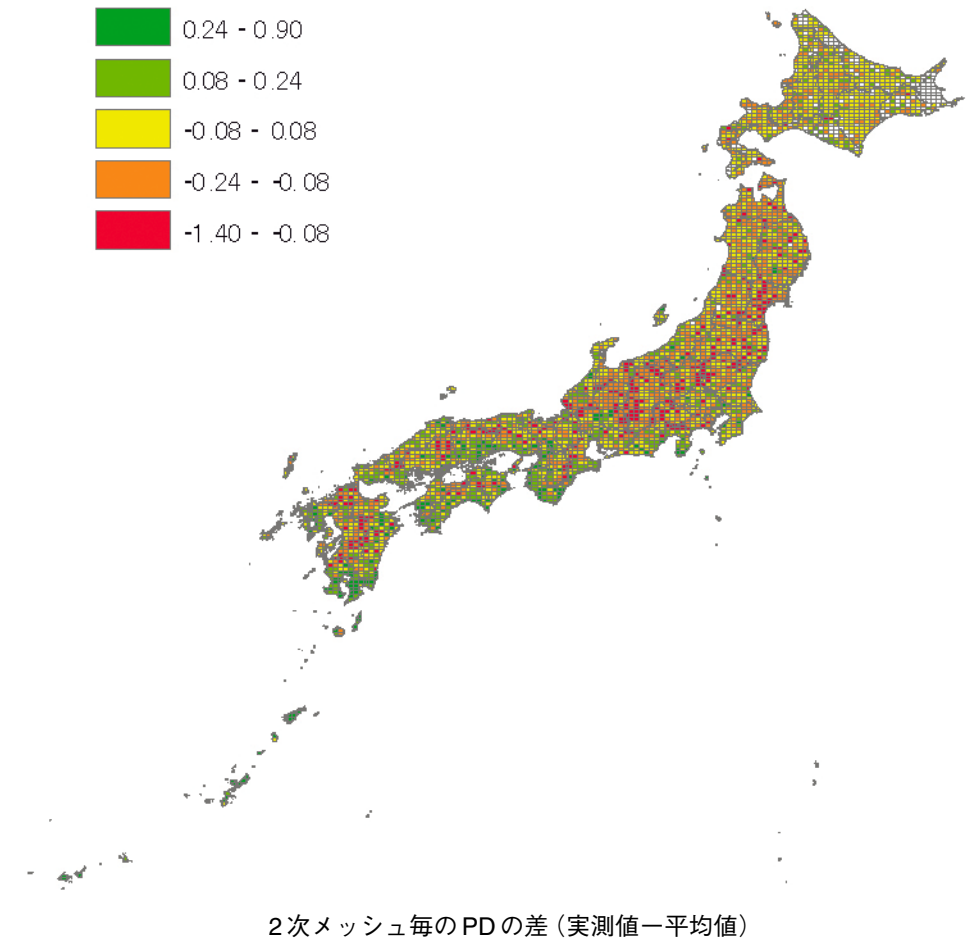
海老原 淳 (国立科学博物館)

生物多様性ホットスポット特定にあたって、一定面積あたりの種数がしばしば指標とされる。しかし、同属10種が分布するのと10科10種が分布するのでは、その意味合いは大きく異なるといえる。ある地域に分布する生物の多様性を評価するためにはそれらが辿った歴史の考慮が重要であり、歴史性の指標として遺伝情報も採り入れる必要がある。

系統的多様性 (PD: Phylogenetic Diversity -- Faith 1992) は分子系統樹を利用して、地域ごとの多様性を分子系統樹の枝長の和で評価する手法である。PDを算出するためには、対象分類群全種について分布情報と分子情報 (特定の遺伝子領域の塩基配列) の両者が完璧に揃っていることが条件となるため、解析例は未だ限られている。日本産シダ植物 (680分類群 = 亜種・変種を含み品種・雑種を除く) では既に2次メッシュ単位の分布情報 (17万2,000件) が網羅されており、分子同定用の葉緑体 *rbcL* 遺伝子の塩基配列も95%近くの種で解析が終了している (海老原、未発表) ことから、解析のモデルとして好適であると考えられる。

PD値は存在種数 (OTU) が増えるのに対応して増加する傾向が見られた。そこで、メッシュ毎に実測値と平均値 (同数の塩基配列のランダムサンプリング時のPD値、1,000回試行の平均) の差を求めることによって、PDが著しい外れ値を示す地域の洗い出しを試みた。種数に対して著しく枝長が長い地域は遺存的系統が集中する地域、逆に、種数に対して著しく枝長が短くなる地域は種分化が盛んに起きている可能性がある地域であると解釈することが可能である。

著しく長い枝長を示す地域は南方の亜熱帯性気候に分布する傾向が見られた。この結果は亜熱帯性気候でのみ生育が可能な系統 (科) の存在と関係があると考えられる。一方、著しく短い枝長を示す地域は本州・四国・九州の内陸部、および北九州に多く分布していた。このうちの一部は分布情報の不足によるアーティファクトの可能性はあるが、九州での分布は無配生殖種の分布密度と関係する傾向が見られた。



## 鳥の渡りを解明する 鳥類標識データベースの概要と今後の展望

山崎 剛史 (財団法人山階鳥類研究所)

本講演では、山階鳥類研究所が実施している鳥類標識調査事業の概要について紹介する。

鳥類標識調査は、野鳥を生きのまま捕獲し、番号が刻印された金属製の足環等で個体識別した後、放鳥することを繰り返す研究手法である(図1、2)。後日、標識した個体が再捕獲されたり、遺体が発見されたりした場合には、個体識別番号をもとに以前の捕獲時の情報と照合が行われる。これによって渡りのルートや寿命等、鳥類についての基礎的なデータを得ることができる。

金属足環を用いる研究手法は、1890年にデンマークの鳥類学者がホシムクドリの個体識別に用いたことを皮切りにして全世界に広まった。日本では、1924年に農商務省が調査を開始したが、その後、戦争の激化等の理由で数度の中断を経験した。1961年以降は、林野庁・米軍・環境庁・環境省等の要請に基づいて山階鳥類研究所がわが国の標識調査事業を一手に引き受け、今日にまで至っている。

現在、山階鳥類研究所は、大正時代に農商務省が行った調査の結果を含む、400万件超の標識調査データを管理しており、近年は毎年およそ17万件的ペースでデータが増加している。これらのデータは、日本全国をカバーする約4,000の地点で集められた(図3)。また、データの収集には、質を担保するため、野外での鳥類の捕獲や分類・同定に関するトレーニングを積んだ“バラタクソノミスト”があたっているが、これも当研究所のデータベースの特色の一つである。山階鳥類研究所は、収集したデータの管理を行うことはもちろん、これらの調査員の育成・指導も進めている。

標識調査における野鳥の捕獲地点のデータは『生物分布情報』とみなすことができるが、そのような情報を解析する技術は、近年、飛躍的な発展を続けている。近い将来、400万件超の標識調査データと、こうした新技術が結びつくことによって、わが国の鳥類相の過去およそ85年間にわたる変遷が解明され、今後の鳥たちのゆくすえを予測することが可能になるに違いない。



図1 野鳥に金属足環を装着する



図2 使用される金属足環

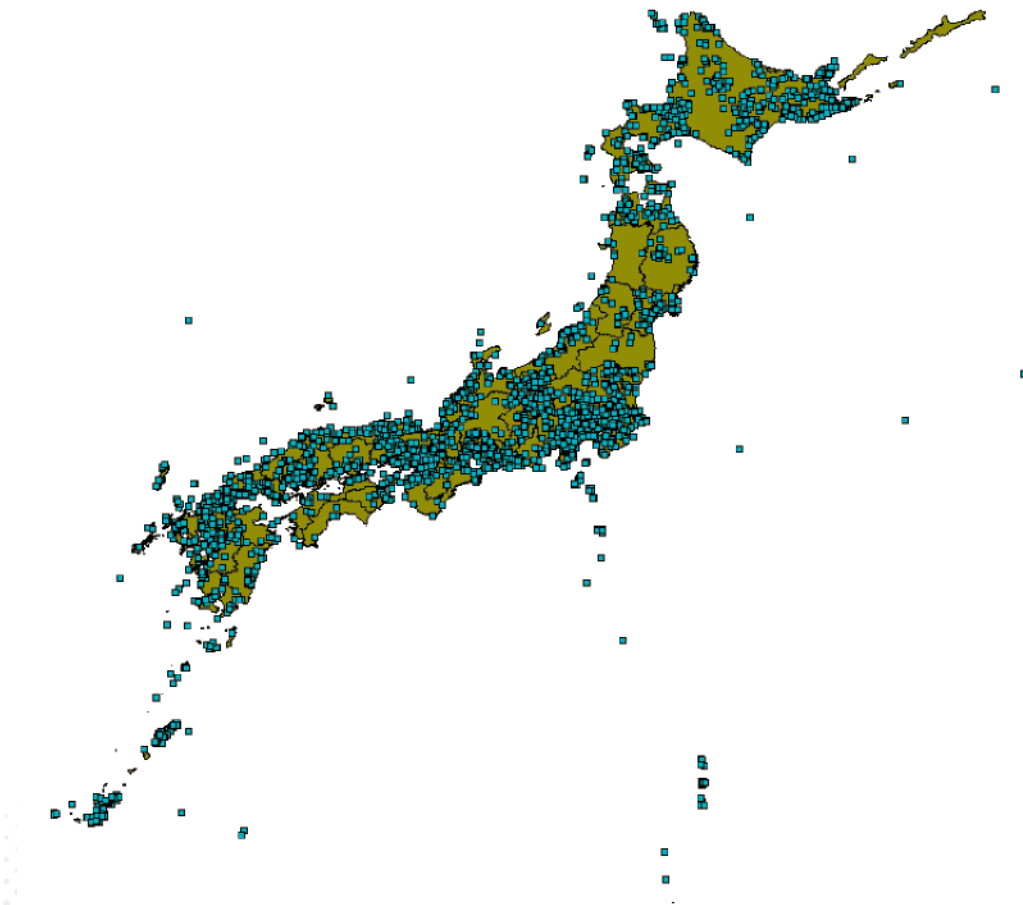


図3 約4,000地点を数える鳥類の捕獲場所



# 生息適地モデルの方法論： GBIF データの活用における課題と展望

三橋 弘宗 (兵庫県立人と自然の博物館)

野生生物の生息適地や種多様性を予測する研究は、ここ10年間で大幅に進歩したことで、自然史分野では最も注目を浴びている分野の一つである。この分野では、生物の分布情報と物理環境要因との関係性を統計学的手法を用いて予測する「生息適地モデル (ecological niche modeling)」を構築し、未調査地のポテンシャル評価や分布を規定する要因の特定、種分化機構の解明、あるいは将来の気候変動との対応を推定することが主な目的となっている(図1)。生息適地の予測は、古くから多数の研究が行われてきたが、近年の飛躍は、GBIFをはじめ生物多様性情報の蓄積と電子情報化が進展したこと、膨大な生物多様性情報や環境情報をパソコンで手軽に扱えるようになったことが大きい。加えて、生息適地モデルから評価地図を作成可能であり、保全計画において効果的に活用されており、社会的な要請が高いことも研究が進展した要因の一つである。

生息適地モデルの方法論は、過去数年間だけでも新たな解析手法が数多く提示されており、方法論の比較がなされている(図2: Guisan et al. 2007)。かつては、重回帰分析やロジスティック回帰分析などの単純な回帰モデルが適用されていたが、計算機性能の向上によって高度な解析方法が適用可能となった。例えば、複雑な環境応答を示す複数の要因に対する解析が可能な「一般化加法モデル (GAM)」、要因を閾値で分岐させて分類を繰り返し予測・判別を行う「樹形回帰」、ニューラルネットワークなどのマシンラーニングの方法を駆使した「GARP」や「MAXENT」、「BRT」といった方法が開発されており、特に「MAXENT」と「BRT」の予測精度は各誌で高く評価されている。統計技法だけでなく、解析するデータの性質や量の問題、環境要因の選定や新たな指標開発、扱う空間スケールと解像度の問題、モデルの適合性の検討、解析目的に応じた手法選定などの課題もある。なかでも、希少種のように情報が少ない場合の解析方法、潜在的には生息ポテンシャルが高いにも関わらず観測が困難なために偽陰性となるケースの扱い、「分布あり」データだけでの予測限界、分布が近傍に密接して空間的自己相関があるケースなど、解析にあたっては様々な注意や工夫が必要となる。

しかし、こうした方法論や生息適地モデルを構築する上での数多くの課題や制約事項については、国内では十分に解説されていない。今回の話題提供では、生息適地モデルを構築する上で課題となっている問題をいくつか取り上げて解析事例を紹介し、今後、GBIF データを活用した解析や評価を進めて行く上での展望を整理する。

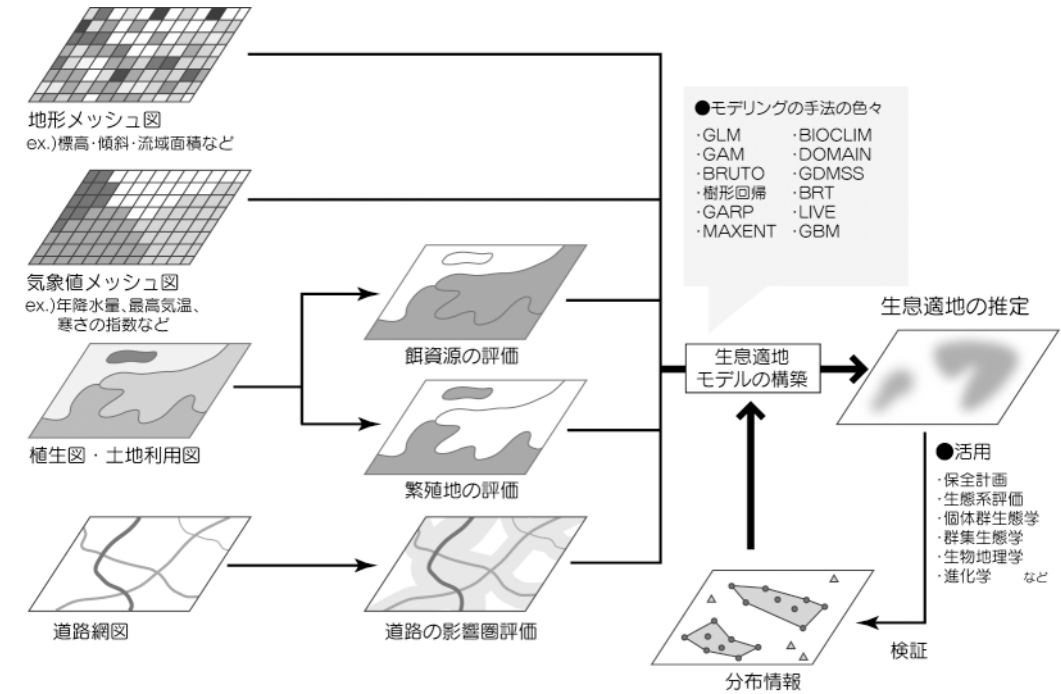


図1 生息適地モデルの構築に関する概略

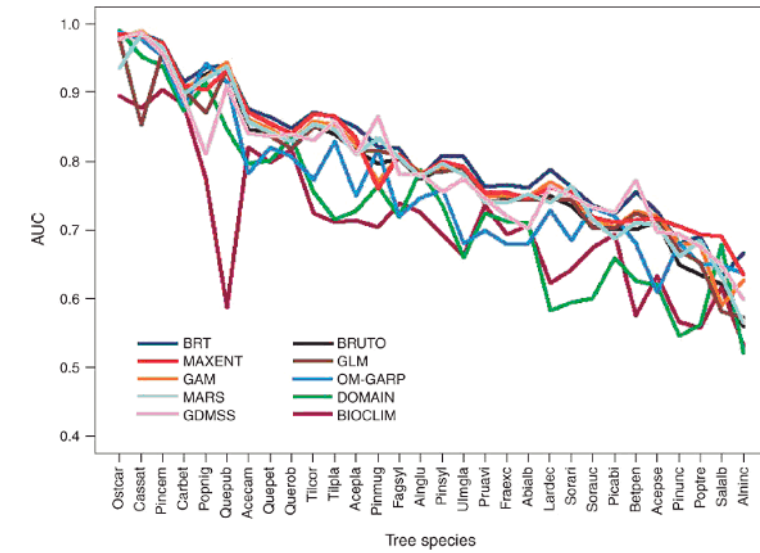


図2 スイス国内の樹木種30種の生息適地モデルを10種類の手法で構築し、解析方法による違いと適合性が比較検証されている (Guisan et al. 2007. Ecol.Monog.77 (4) .615-630)。