

はじめに

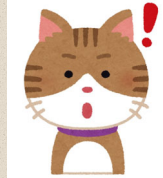
この資料の狙い

種分布モデル (SDM: Species Distribution Model) に
馴染みのない方に基本的な概念・考え方を伝える

詳細な注意点はOpenSDM等で実際に触りながら

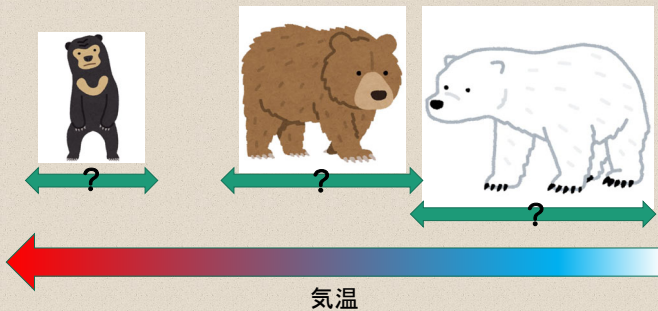
対象: 非専門家

(なので、分からない点などは適宜コメント 頂ければ)



種分布モデル (SDM: Species Distribution Model) とは?

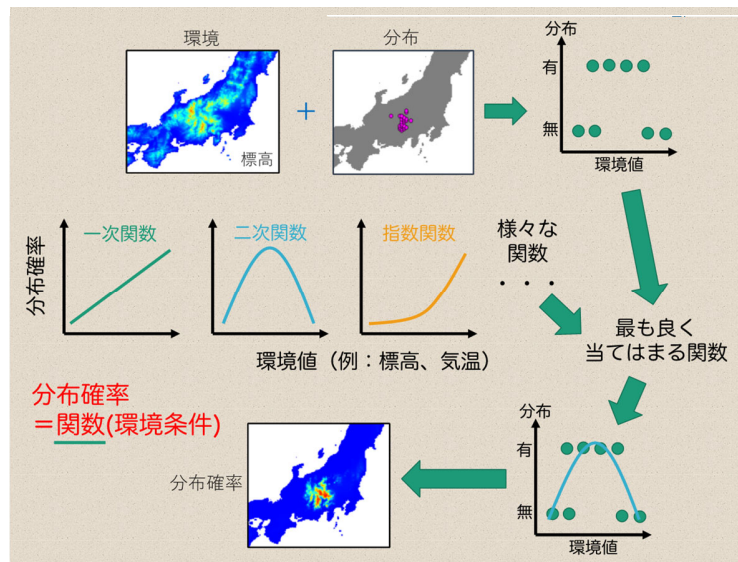
→ 生物種が生息可能な環境の幅 (ニッチ) を推定するツール



- ✓ 種の分布と環境条件 (気温など) との対応関係をモデル化
- ✓ どこからどこまで分布できるのか、定量的に
- ✓ 分布確率で表現

種分布モデルとは、SDM (Species Distribution Model) や Species Niche model、Envelope Model などと呼ばれる統計的なモデルで、生物種が生息可能な環境の幅 (ニッチ: Niche) を推定するツール (道具) です。私たちは大雑把に、北極に生育しているシロクマはすごく寒いところに、温帯に生息しているヒグマは少し寒いところに、熱帯に生息しているマレーグマは暖かいところに分布していることはイメージできます

が、具体的に何°Cから何°Cまでの気温帯にそれぞれの種は分布している (できる) のでしょうか? これを、分布確率を使って定量的に表現してくれる (例: 気温が〇〇°Cの場所の分布確率はXX%だ) のが種分布モデルです。



えておいてください。

3

この分布確率は、実際の生物の分布の有無と、環境条件との間の関係を、適当な関数（函数）（統計モデル）に当てはめることによって算出されます。環境条件と分布の関係は色々なものがあり得ますが、得られたデータや種の性質、モデル利用時の目的によって使い分けます。この点については後ほど（スライド9）もう少し詳しく述べますが、ここでは「色んな統計モデルを使って生物の分布と環境条件との関係を表現する」ことだけ抑

4

種分布モデルを使って何ができるのかというと、まだ調査していない現在もしくは将来の場所で、環境条件から見て分布できそうな場所がどこにあるのかが分かるようになります。分布できそうな場所の中でも、分布確率が高い確実に分布していると考えられる場所と、分布確率が低くて脆弱と考えられる場所とを分けて捉えることもできます。これらの分布確率を描いたマップを使えば、どこを観測・モニタリングすれば温暖化に伴う生物の分布変化を実測できるか（環境条件の変化に敏感な脆弱な場所であれば、温暖化による微細な変化を検出しやすい）や、どこの分布域を保護すれば希少種を絶滅から守ることができるのか（分布確率の高い確実な場所であれば、希少種が実際に生息している可能性が高く効果的に保護しやすい）が分かります。一方で生態系や農業に悪影響を与える生物においては、外来種の駆除や野生動物の食害対策などの課題においても活用できます。

SDMによる分布予測

- ✓ 天気予報的？
- ✓ 確率＝当たったり外れたり（確実ではない）
- ✓ 確認しながら使うべき
- ✓ 利点：備えることができる

うまく使えば

- ✓ 生物の保護・保全
 - ✓ 気候変動対策
 - ✓ 生物資源の管理
 - ✓ 外来種対策
 - ✓ etc.
- をより効果的に



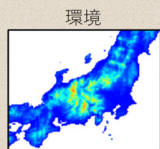
5

ただし、種分布モデルはあくまで確率に基づく予測であるため、天気予報の様に当たったり外れたりがあるものであることは予め想定しておいてください。確実なものではないので、確認しながら使うべきものです。少し使いにくそうではありますが、この不確実性を認識しつつ使えば、予想される問題に予め備えることができます。そしてうまく使えば、希少種の保護・保全や、将来的な気候変動に対する生物多様性の保持、漁獲資源・山

菜・お花畑などの生物資源の管理や、外来種対策などで効果的に利活用できるものです。

SDMはどうやって作られるか？

→ 必要なものは以下の3つ



環境

環境データ
(どんな場所？)



分布

対象種の分布データ
(いつどこで見られたか？)

最も良く
当てはまる関数

両者をつなぐ統計モデル
(どんな関係性？)

6

そんな種分布モデルはどうやって作られるのか、ここからはその構成要素である3つ（種分布データ、環境データ、統計モデル）に関してそれぞれもう少し詳しく説明いたします。

まず生物種の分布データですが、このデータは図鑑などに記載されている分布域の記述を参考にしながら、なるべく対象種の分布域全域をカバーできるようにデータを集める必要があります。分布全域からデータを集めるのは、例えばその種にとって寒い場所の分布データが欠落すると、モデル上でどこまで寒い場所へ分布できるかを過小評価してしまうためです。全域をカバーする必要があるために、自力で全域調査することはかなり難しい

生物の分布情報を収録したデータベースのうち、比較的幅広い分類群を含んでいて公開されているデータベースとして、3つほどここでご紹介いたします。

地球規模生物多様性情報機構（GBIF: Global Biodiversity Information Facility: <https://www.gbif.org/>）は、全球規模で生物の情報をまとめて収録したデータ基盤です。ここには主に博物館に収集された標本のデータなどが収録されています。ホームページ上でも個別の生物種の分布データを地図上で参照できますし、アカウントを作成すれば分布の緯度

分布データ（ポイントデータ）

- なるべく種の分布域（図鑑参照）全域で
- 自力での全域収集はほぼ不可能
- 先人のデータベースを使用

環境省
自然環境保全基礎調査

GBIF: Global Biodiversity Information Facility

- 全球データベース
- 博物館の標本データなど
- Rを使った抽出・解析も
(<https://uriibo.hatenablog.com/entry/2015/12/25/125732>)

iNaturalist

- 全球データベース
- 市民の観察報告（スマホ）
- Rを使った抽出・解析も
(https://qiita.com/ocean_f/items/aa0cdd4f00fd73079699)

Fagus crenata Blume
(<https://www.gbif.org/ja/species/2882300>)

観察記録
(https://www.inaturalist.org/observations?place_id=any&subview=map&taxon_id=362505)

と言えますので、ここは先人の培ってくれたデータベースを活用してください。

サイエンスミュージアムネット

- 国内データベース
- 博物館の標本データなど
(<https://science-net.kahaku.go.jp/>)

いきものログ

- 国内データベース
- 市民の観察報告（スマホ）
(<https://ikilog.biodic.go.jp/>)

在のみデータ（標本、観察）

在・不在データ（調査）

モデル構築時
不在データの推定・出力

調査区の不在も
専門家が調査・報告

経度を含む詳細なデータをホームページからダウンロードも可能です。また統計プログラミング言語である R を用いて直接データを抽出・解析するためのパッケージ（参考：

<https://uriibo.hatenablog.com/entry/2015/12/25/125732>）も開発・提供されているので、R に慣れている方はそちらを使うとかなり使い勝手が良いと思います。GBIF は全球的なスケールで収集しているデータセットですが、国内の博物館などで収録されている標本情報のデータベースであるサイエンスミュージアムネット（S-Net: <https://science-net.kahaku.go.jp/>）でも、同様にデータのダウンロードが可能になっています。

iNaturalist (<https://www.inaturalist.org/>) は、同じく全球規模で生物の情報を収録したデータ基盤ですが、こちらは専門家だけでなく一般市民によって観察されたデータも含む、いわゆる市民データのデータベースとなっています。近年のスマートフォンの普及に伴い、スマートフォンのカメラと GPS 情報を用いて、機械学習による自動的な種同定システムが普及したこともあり、近年データが増えて

きています。それでも GBIF に比べるとまだまだデータ数は少ないですが、市民の方と一緒にモニタリングを展開して、普及啓発しながら生物の変化を観測する際などには、有用性の高いプラットフォームと言えるでしょう。GBIF と同じく HP 上で分布マップの確認やダウンロードが可能ですし、R のパッケージを用いたデータのやり取りも可能（参照：https://qiita.com/ocean_f/items/aa0cdd4f00fd73079699）です。iNaturalist のような市民データにおいても、環境省が国内を対象に展開するいきものログ（<https://ikilog.biodic.go.jp/>）があり、データのダウンロードが可能になっています。

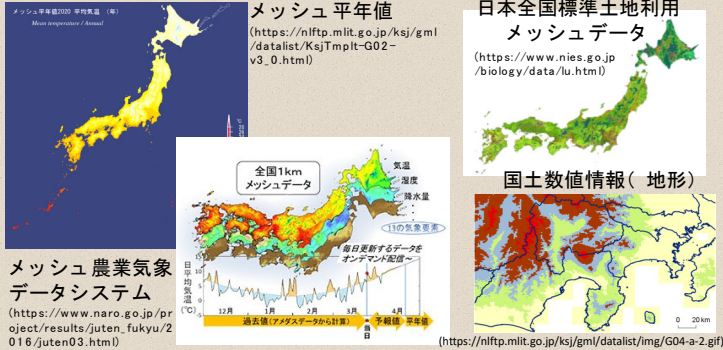
環境省の自然環境保全基礎調査（https://www.biodic.go.jp/kiso/fnd_list_h.html）は、我が国における多種多様な生物種・生態系における様々な調査結果を収録しています。いろいろな生物種を対象にしており、調査ごとに個別のリンクで説明されているので少しホームページ内を探す必要はありますが、生物種や個体レベルまで記載した詳細なデータが公開されているものもありますので、詳しく解析したい場合は有用性が高いと考えられます。一方で希少種の場合は盗掘などの危険を考慮して公開制限をかけているので、閲覧・ダウンロードなどはできないはずです。

これらのデータベースは、扱っている分布データの種類として「在のみデータ」と「在・不在データ」の2種類に分けられます。1つ目は「在のみデータ」と呼ばれるもので、「いつどこにどの生物が分布していた」というデータだけを集めたものになります。博物館の標本データや、市民データなどにおける観察データはこれに含まれ、前述の GBIF、S-Net、iNaturalist、いきものログのデータベースはこの在のみデータを収録しています。この在のみデータはデータ収集が比較的簡便であるためデータ数も多いのですが、種分布モデルの構築時にはどこに不在データがあると想定すべきか考えて、それに関するパラメータを提供する必要があります。この不在データの推定が大きな不確実性を含んでいるため、在のみデータはモデル構築時に課題があると言えます。

もう1つの分布データの種類は、「在・不在データ」と呼ばれるもので、「いつどこにどの生物が分布していた or 分布していなかった」というデータを集めたものになります。先述のデータベースの中だと、自然環境保全基礎調査における植生調査データベース（http://gis.biodic.go.jp/webgis/files/veg_survey_db_H12-R3.pdf、環境省生物多様性センターより提供）などがこの在・不在にあたります。このデータは、調査を行う方形区（枠、区画）や道などのルート、観察時間を区切った地点での観察調査を行い、その場所や時間内で出現したすべての種を記載する調査によって得られます。すべての種を記載しているため、そこに記載されていない種は分布が見られなかった「不在データ」として扱うこともできるため、これをもとに解析対象の種ごとに在・不在データを収集できます。この在・不在データでは不在データも観察データとして入力できるため、種分布モデル構築時の不確実性が低いというメリットがありますが、一方でどのように、どの程度、どこで調査をするのかを適切に判断する必要があり、専門家による調査やサポートが必要になる点に課題があります。

環境データ（ラスターデータ：地図）

- 気候値（気温、降水量、最大積雪深、日射量、etc.）
 - 土地利用（森林、農地、都市、etc.）
 - 地形（傾斜、斜面方位、尾根谷、凸凹、集水面積、etc.）
 - 地質土壌（石灰岩、蛇紋岩、褐色森林土、黒ボク土、etc.）
- 対象種の特性や調査の空間スケールで使い分け



8

次に環境データですが、気候値、土地利用、地形、地質土壌など、様々なものが候補として挙げられます。対象種の分布に影響していそうなものは全て集めたいところではあるのですが、全国スケールの種分布モデルを構築する場合であれば、ある程度主要な変数である気温、降水量の気候値と、地形、土地利用のデータがあればひとまず良いのではないかと思います。崖や砂浜、石灰岩や蛇紋岩などの特殊な立地や地質などに特異的に出現

する種であれば、植生図や地質などを含めて解析したいところですが、全国という広域スケールで空間解像度を粗く（1km メッシュ程度）解析するにはそれらの変数はそこまで強く影響してこないことが多いです。

気温や降水量などの気象値は、気象庁によって全国の気象観測地点でデータを観測しています。しかし何処どここの観測地点の値がわかるだけでは生物の分布データがある場所での値はわからないため、これらの気象観測地点のデータを使って気象モデルを構築し、観測所のないメッシュの値を推定して穴埋めをして、日本の陸域全域での値を推定したメッシュデータをここでは主に使います。これを使えば、気象観測所がない場所の生物分布データに対応した気候値が参照できます。

また気候値として使用する際、ある程度長い時間間隔（20 年～30 年程度）における平均値（平年値）を使う必要があります。ある年のデータだけ使うと、その年がたまたま暑い年だったり寒い年だったりした際に、生物の分布を正確に把握することが難しくなるため、こうした平年値を使います。

気象庁が提供しているメッシュ平年値（https://nftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-G02-v3_0.html）は、30 年間の平均値として算出した値を、全国の 1km メッシュあたりで全ての陸域を対象に算出したデータとなっています。年値・月値で気温、降水量、最深積雪深、全天日射量などを揃えています。スタンダードな気候値であるのですが、全国のデータを一括でダウンロードすることは難しく、80×80km のメッシュごとにダウンロードする必要があります。また 2022 年度に作成された第 3 版は 1991～2020 の 30 年間の値が収録されていますが、1981～2010（第 2 版：<https://nftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-G02.html>）、1971～2000（第 1 版）などもあり、生物の分布データの観測時期に合わせて使い分けるなどもできます。

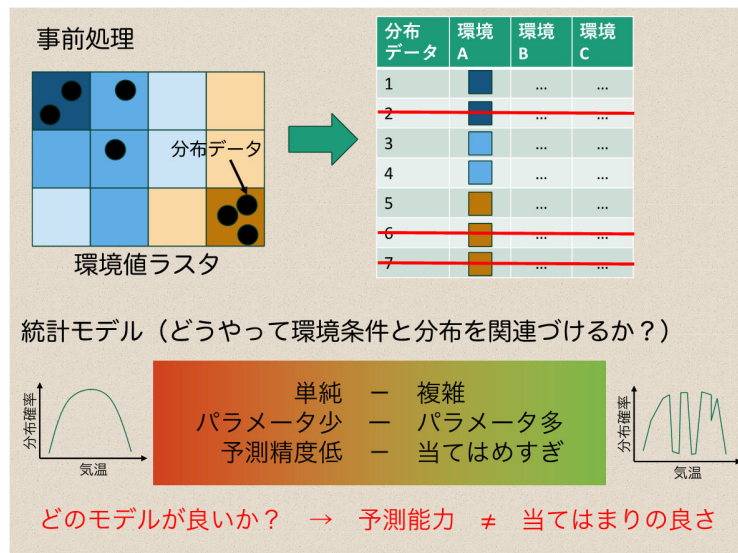
農業・食品産業技術総合研究機構（農研機構）が提供するメッシュ農業気象データ（<https://amu.rd.naro.go.jp/>）は、全国の日別気象データを、1km メッシュ単位にオンデマンドで提供してくれるデータベースです。14 種類の気象要素（気温、降水量、積雪深、積雪水量、全天日射量、長波放射量など）を 1980 年 1 月 1 日から現在まで、のみならず現在から 1 年後の 12 月 31 日までの未来のデータまで提供しています。かなり自由度が高いのですが、それをユーザーがオンデマンドで

操作するために、アカウント登録がまず必要で、さらにプログラミング言語である Python 言語をユーザーが書いてダウンロードする必要があるなど、ちょっとハードルが高い気候値と言えるかもしれません。また長期平均に当たる平年値を用いる場合は、自分で計算する必要があります。

土地利用のデータは様々なものがありますが、気候値の空間解像度である 1km メッシュあたりで集計していて、自然生態系における分類が多めになっている日本全国標準土地利用メッシュデータ (<https://www.nies.go.jp/biology/data/lu.html>) が、生物の分布予測の際にはある程度使い勝手が良いかと思います。このデータベースでは植生図や国土数値情報などのデータを用いて比較的多くの土地被覆分類を持っていて、他の多くの土地利用図では単に「森林」としている分類に対しても、「自然林」「二次林」「人工林」という形で細かく分類しており、生物種の分布予測に有用な情報となります。

地形のデータも衛星や航空機によるリモートセンシングを使って得られたデータをまとめたデータベースがいくつかありますが、全国スケールで標高・傾斜角度・傾斜方位などについてデータをまとめた国土数値情報の標高・傾斜度 3 次 (1km) ～ 5 次 (250m) メッシュデータ

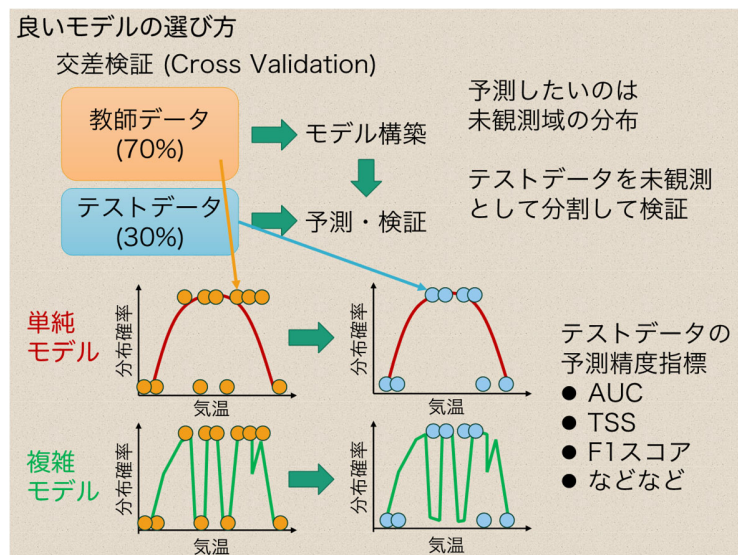
(<https://nlftp.mlit.go.jp/ksj/gml/datalist/KsjTmplt-G04-a.html>) などは基礎的な地形データになるかと思います。より細かい解像度の地形データが必要な場合は、5m や 10m の解像度で地点あたりの標高値を収録した数値標高データ (<https://fgd.gsi.go.jp/download/menu.php>) からデータをダウンロードします。標高や傾斜だけでなく、尾根や谷などの地形を抽出したい際には、さらにこれらダウンロードしたデータを GIS ソフトウェア (ArcGIS、QGIS など) を使って解析し、データを出力する必要があります。



9

次に統計モデルです。まずモデル作成の前に、得られた緯度経度のポイント形式の分布データと、地図的なラスタ形式（テレビの様に地上を緯度経度に沿った細かい区画に分けて環境値を入力した形式）の環境データを対応づけたデータシート（どの分布データの環境値はこの値と対応づけたもの）を作成します。この際、あまりにも分布データ間が近すぎる（データの空間密度が高い）部分がある場合、そのデータのモデルが過度に引っ張られてしまうことを避けるために、重複のある分布データを除外します。具体的には、多くの場合気候値のデータが1kmメッシュで提供されていることもあるため、同じ1kmメッシュ内に複数存在する分布データに関しては、1つに絞るなどの操作をしたりします。

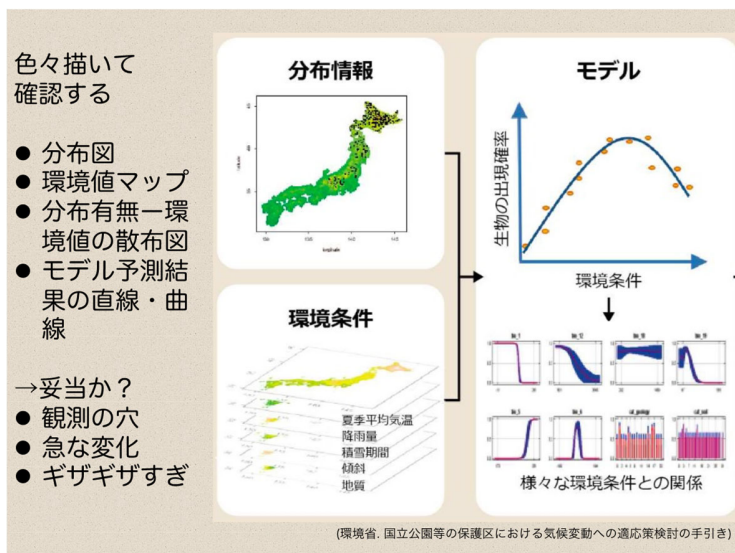
こうして得られたデータシートを使って、統計モデルの構築を行います。統計モデルには様々なものがありますが、ここではそれらのモデルを、線形回帰モデルなどの少ないパラメータ（変数）で表現した比較的単純なモデルと、MaxEnt や RandomForest などの機械学習と呼ばれるような多くのパラメータで表現した複雑なモデルにざっくりと分けて、それぞれについて簡単に説明させていただきます。単純なモデルは、横軸に環境値、縦軸に分布の有無（分布確率）をとった際に、直線関係や一山型など比較的単純な関数の形で両者の関係を表現するモデルです。一方で複雑なモデルは、より複雑な関係であっても表現できる自由度の高いモデルといえます。単純なモデルではデータとの適合度（予測精度）が低くなる一方で、複雑なモデルではデータとの適合度が高くなる傾向があります。しかし、良いモデルというのは当てはまりが良いモデルを指しているわけではありません。種分布予測モデルの場合、良いモデルとは観測できていないところのデータを予測できる能力（予測能力）が高いモデルを指します。この予測能力を検証するために、交差検証と呼ばれる操作でデータシートの中から一部の教師データ（70%など）だけを使ってモデルの構築を行い、残りのテストデータを未観測データとしてモデルからの予測値と観測値を比較して予測能力の評価を行います。こうした予測能力の評価を行うと、複雑なモデルの方が予測能力が低いことも多くあり、こうした場合は複雑なモデルでは与えられた70%の教師データにピッタリ当てはめすぎていて他の場所の予測には使えないと判断されます。テスト



モデルと、MaxEnt や RandomForest などの機械学習と呼ばれるような多くのパラメータで表現した複雑なモデルにざっくりと分けて、それぞれについて簡単に説明させていただきます。単純なモデルは、横軸に環境値、縦軸に分布の有無（分布確率）をとった際に、直線関係や一山型など比較的単純な関数の形で両者の関係を表現するモデルです。一方で複雑なモデルは、より複雑な関係であっても表現できる自由度の高いモデルといえます。単純なモデルではデータとの適合度（予測精度）が低くなる一方で、複雑なモデルではデータとの適合度が高くなる傾向があります。しかし、良いモデルというのは当てはまりが良いモデルを指しているわけではありません。種分布予測モデルの場合、良いモデルとは観測できていないところのデータを予測できる能力（予測能力）が高いモデルを指します。この予測能力を検証するために、交差検証と呼ばれる操作でデータシートの中から一部の教師データ（70%など）だけを使ってモデルの構築を行い、残りのテストデータを未観測データとしてモデルからの予測値と観測値を比較して予測能力の評価を行います。こうした予測能力の評価を行うと、複雑なモデルの方が予測能力が低いことも多くあり、こうした場合は複雑なモデルでは与えられた70%の教師データにピッタリ当てはめすぎていて他の場所の予測には使えないと判断されます。テスト

トデータでの予測値と観測値の適合度（予測精度）を評価する指標も、ROC（Receiver Operating Characteristic）解析の結果出される AUC（Area Under the Curve）という指標や、TSS（True Skill Statistic）指標、再現率（実際に分布ありのものの中で正しく分布ありと予測できた割合）と適合率（分布ありと予測したものの中で実際に分布ありだった割合）の調和平均をとった F1 スコアなど、さまざまなものがありますが、詳しい説明は他の資料を参照いただければと思います。

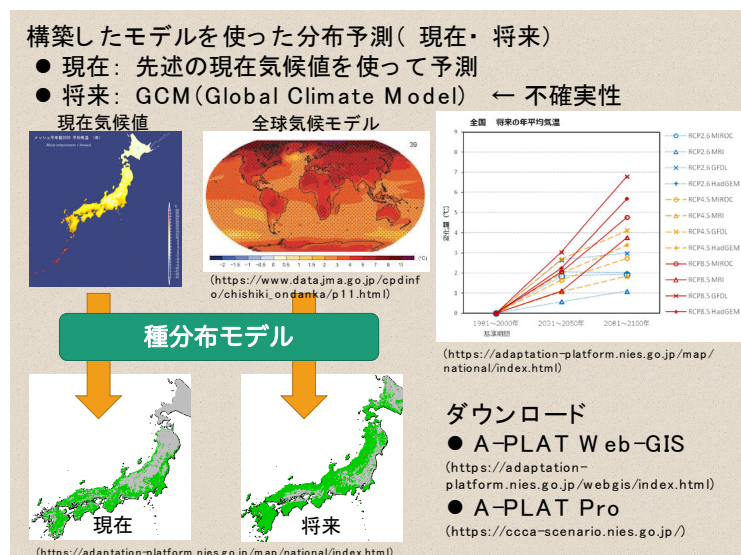
様々なモデルが提唱されたり使われたりしており、それらを全て学習することはかなりハードルが高いですが、少ない分布データ数からでも比較的的確なモデルが構築でき、適度にモデルの複雑さも調整され予測能力が比較的高い機械学習モデルの一つである MaxEnt というモデルが現在世界的によく使われています。近年このモデルを使って個別ユーザー独自の種分布モデルを、プログラムを書かせずにマウスのクリックベースで構築をサポートするプラットフォーム（例：オープン SDM: https://adaptation-platform.nies.go.jp/map/open_sdm.html）なども提供されています。こうしたものから触っていくのが良いかと思います。



10

種分布モデルが構築できたら、マップやグラフを色々描いてみてどのような分布と環境の関係が描かれているのか確認してみましょう。種分布や環境条件のマップはモデル構築前にも確認する必要がありますが、ここではモデル予測結果である分布確率のマップとも比較しながら確認します。また分布の有無を縦軸に、環境条件の値を横軸にした散布図も、モデル予測結果として示される両者の関係を表す直線・曲線とも比較しながら、対象

種の特徴と照らし合わせて妥当な関係として描けているのかを確認します。この際の確認事項は様々ありますが、分布データ上に観測の穴が無いかや、両者の関係がどこかで急激に変化する不連続な関係が無いか（ある場合なぜその環境値で起きるのか）、分布と環境の関係曲線が極端にギザギザしたりぐにゃぐにゃしたりしていないかなどを見たりします。



11

最後に構築したモデルと気候値を使って、分布予測図を描画します。現在の分布予測図を描く際は、先述した現在気候値を使ってまず分布確率の算出を行い、分布の有無と分布確率の関係を解析してある閾値（境界となる値）を設定し、その閾値よりも分布確率が高い場所を分布可能な場所として描画します。閾値の設定にはいくつか方法がありますが、主要な方法としてはROC（Receiver Operating Characteristic）解析を使って推定

したり、分布有のデータの95%や全ての値を含むように分布確率の閾値を設定したりします。

一方で将来の分布予測には、将来の気候値を持ってきて予測を行います。この将来の気候値は、全球気候モデル（Global Climate Model）などと呼ばれる気候モデルの予測結果を使います。ただ気候モデル自体も予測結果なので、当たったり外れたりの不確実性が存在する点は注意が必要で、この将来気候値における不確実性を表現するために、複数の全球気候モデルの下で将来分布予測図を描画・比較して、予測結果の不確実性を想定範囲に織り込んでおくことが大切です。将来気候値に関してもいくつかのダウンロードサイトがありますが、日本の1kmメッシュという比較的高い解像度までダウンスケール（空間解像度の詳細化）してバイアス補正（観測値を使った誤差の補正）もしており、幾つかのデータセットを提供しているA-PLATのWeb-GIS (<https://adaptation-platform.nies.go.jp/webgis/index.html>) や、アカウントの登録によってユーザーが自動的にA-PLATのデータをダウンロードできるA-PLAT Pro (<https://ccca-scenario.nies.go.jp/>) などが日本国内での使用の際には良いかと思います。

予測結果（マップ）の使い方

- 対照種の管理目標（保護／駆除）
- 4つの状況に分けて使い分け

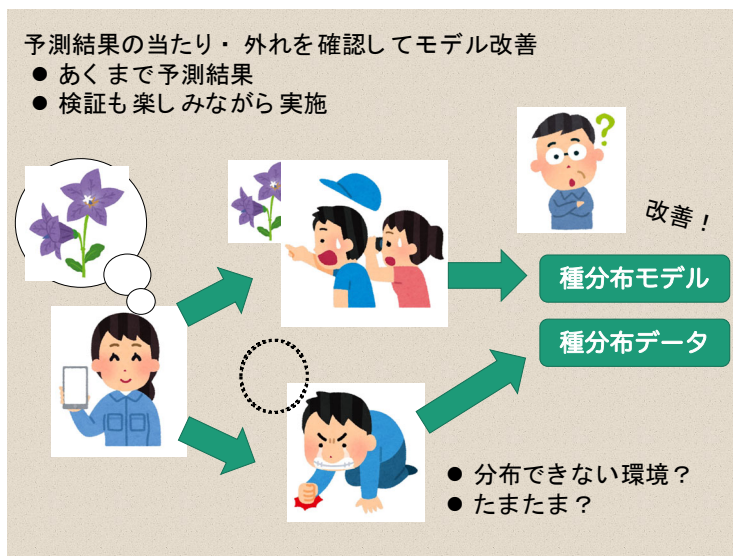
	現在分布なし	現在分布あり
将来分布なし	✓ 基本的に対象外 ✓ (分布辺縁部でモニタリング) D	- 保護（例：希少種） ✓ 気候変動以外の分布阻害要因を除去 ✓ 維持のための管理 - 駆除（例：外来種） ✓ 消失過程のモニタリング ✓ 消失を助長する管理 B
将来分布あり	- 保護 ✓ 積極的な導入・定着のための管理 ✓ 侵入過程のモニタリング - 駆除 ✓ 侵入防止措置 C	- 保護 ✓ 保護区に設定して人為的な撓乱などを除去 ✓ 長期的に保護 - 駆除 ✓ 集中的な駆除 ✓ 予算と効果のバランス検討 A

12

得られた現在と将来の分布予測図をどうやって使うのか、ここではその大まかな考え方・事例をご紹介します。分布予測を行った対象の種に対する管理目標（例：希少種で保護をしたい or 外来種で駆除をしたい）を想定した上で、現在・将来の分布有・無に合わせて4つの場合に分けて考えたのがこちらの表になります。現在も将来も分布がないDは、基本的に想定の対象外なのですが、分布辺縁部においては予備的にモニタリングをしておく

くという考え方もあるかもしれません。逆に現在も将来も分布があるAでは、保護を考える際には

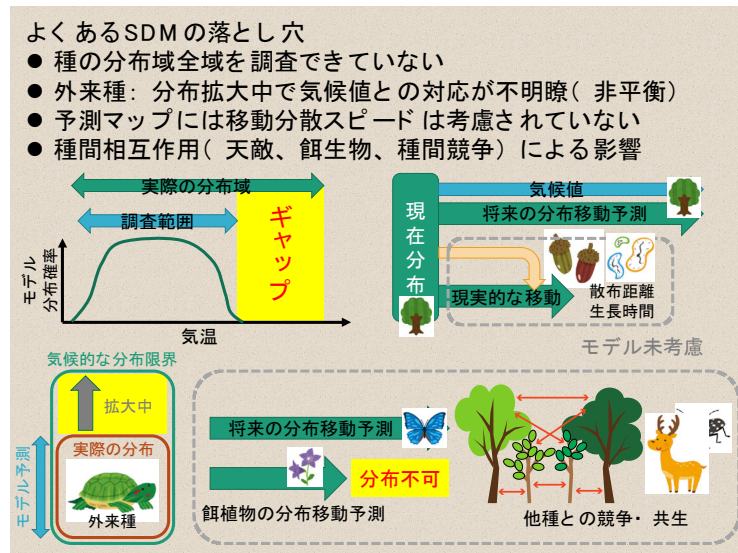
安定的に分布が維持される場所と言えるので積極的に保護区に設定して長期的な保護を図る必要がありますが、駆除を考える際には分布の中心地であるため集中的な駆除を予算と効果のバランスを測りながら考える必要があります。将来分布が消失する B では、保護では将来変化を引き起こす気候変動以外のマイナス要因を除去するなどして地域個体群の維持を図ることを考えるべきですが、駆除では本当に消失していくかモニタリングをするなどが考えられます。さらに将来分布が形成される C では、保護では侵入の妨げとなる他生物などの除外や侵入過程のモニタリングが、駆除では侵入個体の駆除などによる侵入防止措置が考えられます。この表では単純化のために分布の有無で表現しましたが、実際には確率の高低による対応の違いも想定しておく必要があるかと思います。



13

繰り返しになりますが、描画された分布予測図はあくまで予測であるので、当たり外れがあることは理解が必要です。そのためできれば予測された分布を確かめに行って、分布が実際にあった際の驚きやなかった時の悔しさも含めて、楽しみながら当たり外れを確認して、種分布モデルの改善や種分布データのアップデートを行っていけると良いかと思います。なお検証の結果分布が無かった場合には、環境条件としては分布可能だがたまたま

種子が散布されていなかったなどの理由で分布していなかった場合と、環境条件として分布しにくい場所だった場合の2つがある点は注意が必要で、その確認のためにもなるべく多くの検証データが必要です。なかなかここまで計画を立てて実際に行っている事例はないですが、特に将来のモデル予測結果は慎重に受け止めるべきなので、利用の際は検証・確認まで頭に入れていただけると良いかと思います。



14

こうした種分布モデルを構築する上で、陥ってしまいがちな落とし穴をいくつか最後に述べようと思います。

種の分布や気候値の空間的な制限から、対象種の全ての分布域のうち、一部の部分だけ調査して得られたデータを使って種分布モデルを構築した場合は、実際の分布域とモデル上で予測される分布域に大きなギャップが生じてしまいます。これは分布予測域の過小評価につながるため、こうやって作られた種分

布モデルは使ってはいけません。「日本全国のデータを扱うのは骨が折れるので、〇〇県や△△地域のデータだけ使おう」とする場合などは、あり得る落とし穴かと思います。

同様に、現在も分布拡大中と考えられる移入されてからそれほど時間の経っていない外来種に対しても、種分布モデルの構築は注意が必要です。こうした若い外来種では環境条件的に生息できる場所の隅々まで分布が行き渡っていない状況（非平衡状態）と考えられるため、拡大する余地のある部分にモデルと現実的な分布の間のギャップが生じてしまいます。こうした外来種の場合、その外来種が元々生息していた国・地域のデータで種分布モデルを構築して拡大途中の場所（日本）でそのモデルを使って予測した分布域を描いて、どれほど拡大する余地があるかを検討することも可能ですが、外国と日本では周りに住む生物や環境条件も異なるため、全く同じ環境的な分布限界が引かれるわけでもないので注意が必要になります。

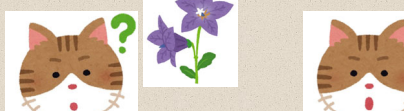
また日本国内で隅々まで分布が行き渡ったと考えられる在来種などにおいても、環境条件と分布との対応関係を解析した種分布モデルでは考慮できていないプロセスも多く含まれるため、特に将来的な分布変化を予測する際には注意が必要です。将来的な温暖化によってより寒い場所へ気候的に分布可能な場所が広がることは予測できますが、実際に分布を広げるためには、例えば樹木の場合であれば広がった領域まで種（たね）を飛ばし、さらにそこである程度の大きさになるまで時間がかかりますが、こうした散布・分散能力や生育にかかる時間は、一般的な種分布モデルでは考慮しきれていません。そして多くの種の場合、特に樹木のように移動せず分布を広げる力が低い種においては、温暖化のスピードに対して種の分布移動のスピードは大きく遅れていることが報告されています。そのため将来の分布予測をする際には、現在の分布域からある程度散布・分散可能な範囲に予測領域を絞って予測結果を使うなどの処理も必要になったりします。

また自然環境の中では、それぞれの種は環境条件だけの影響を受けて分布が決まっているわけではなく、実際には競合する他種や餌生物・共生生物などの影響（生物間相互作用）を受けながら分布が形成されています。そのため、温暖化に対する現実的な分布移動スピードがこれら強く影響を受ける他種において予測対象とする生物種と大きく異なる場合には、他種における将来分布予測も構築して

検討する必要があります。そうでなければ、「蝶々で予測された将来的な分布の拡大に対して、餌となる植物の分布の拡大が遅いために実際には移動がある程度制限される」状況や、「樹木の予測分布域に対して、競合する他の樹木種による競争排除や獣害・虫害の拡大によって実際にはある程度分布が制限される」などといった状況をうまく説明・想定できない状況になってしまいます。

SDM：いろんな問題はあるが、有用なツール

- データを使って目に見えないものを可視化
- 怖がらずに触って学びつつ、うまく活用すべき





オープン SDM 種分布モデル 構築補助ツール 使い方・注意点 日本語化中！ →誰でも使える ツールに

15

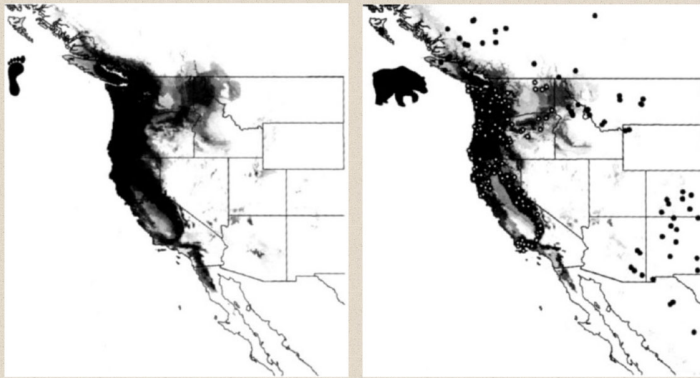
種分布モデルには現状まだ色々な課題もあるのは確かなのですが、データを使って目に見えない分布域というものをうまく可視化・数値化してくれる有用なツールであることは確かです。ある場所である植物を見てもどこまでその種が広がっているかはイメージできませんが、種分布モデルはそこを具体的・定量的に表現してくれます。また将来分布域などはそれこそ全く見えないものですが、種分布モデルはそれを定量的に可視化してくれま

す。こうしたメリットに目を向けつつ、怖がらずに触って学びつつ、うまく利活用する必要があると思います。失敗もちろんあるでしょうが、地域の専門家との繋がりを広げて確認してもらったり、モデル予測を使った適応策を実施する中でモニタリングをしながら、間違っていることが分かった時点で修正・改善をして順応的に管理をすれば、多くの場合対応は可能だと思います。

種分布モデルの構築補助ツールも最近提供されてくるようになってきており、これまでプログラミング言語を駆使して専門家が構築するものだった種分布モデルも、マウスクリックベースで構築できるようにする環境が提供されるようになってきました。このツールのうちの一つを日本語化して使いやすくしたものを国立環境研究所でも提供し始めたところ（オープン SDM: https://adaptation-platform.nies.go.jp/map/open_sdm.html）ですので、こうしたツールも使って種分布モデルに触って楽しんでいただければと思います。

おまけ：面白いSDMの使い方

- ビックフットの分布予測域＝ほぼクロクマの分布予測域



(Lozier et al. 2009. Predicting the Distribution of Sasquatch in Western North America: Anything Goes with Ecological Niche Modelling. J Biogeogr. 36: 1623-1627)

最後に種分布モデル（MaxEnt）を使った面白い研究の事例を一つ紹介して終わりにしようと思います。

アメリカの未確認動物（UMA: Unidentified Mysterious Animal）であるビックフットで種分布モデルを構築したというものです。この論文の面白いところはさらに、北アメリカのクロクマにおいても種分布モデルを構築し、両者の分布予測図を比較することによって、両者の分布域が大きく重複しており、おそら

くビックフットはクロクマの見間違いなのだろうという結論を導き出しているところです。ある意味UMA好きな人々の夢を奪うような論文とも言えてしまいますが、この着眼点と種分布モデルの使い方は個人的に非常に面白くて好きなところです。

種分布モデルもツールの一つと言えるので、使い方次第で楽しみ方は無限大です。ポケモンの種分布モデルだってありますし、「田中さんは本当に田んぼの中に住んでいるのか？」といった疑問へも種分布モデルでアプローチできるでしょう。肩肘を張りすぎずに、是非一度種分布モデルを触ってみて、楽しんでみてください。

図の引用： Lozier et al. 2009. Predicting the Distribution of Sasquatch in Western North America: Anything Goes with Ecological Niche Modelling. J Biogeogr. 36: 1623-1627

筆者：小出大

国立環境研究所 気候変動適応センター

気候変動影響観測研究室 研究員

〒305-8506 茨城県つくば市小野川 16-2

mail: koide.dai@nies.go.jp