生息場適性指数モデリングのための ステークホルダー向け視覚分析環境*

尾上 洋介**, 坂本 尚久***, 小山田 耕二***

Visual Analysis Environment of Habitat Suitability Index Modeling for Stakeholders

Yosuke ONOUE, Naohisa SAKAMOTO and Koji KOYAMADA

1. はじめに

気候変動に対応可能な漁場探索技術及び水産資源変動予測手法の開発が活発に推進されている。気候変動は、前例のない漁場変化を引き起こしており、中長期的な資源変動予測と短期的なピンポイント漁場予測の両方が、持続的な漁業及び食料安全保障のために必要とされている。このような背景のもと、気候変動適応研究推進プログラム (RECCA)¹⁾では、青森県でのアカイカ漁を対象に、実際の漁獲実績データと3次元時系列多変量海洋シミュレーションデータを用いて漁場予測モデルの構築を行っている²⁾。本稿では、アカイカのピンポイント漁場予測プロジェクトにおける、予測モデル構築のための視覚分析環境に関する研究を紹介する。

本研究では、生息場適性指数 (HSI: Habitat Suitability Index) モデルを漁場予測モデルとして用いている。HSI は野生生物の生息環境を定量的に評価するための指標としてアメリカ合衆国内務省で開発された³⁾。HSI モデリングでは、複数の環境要因を用いて対象生物種の生息場適性を説明するモデルを構築する。海洋生物に対する環境要因としては、水温や塩分濃度、流速、海面高度などがある。従来のアカイカの HSI モデリングでは、アカイカの生態系の理解に焦点が当てられていたため、時空間共に粗い解像度のデータが用いられていた^{4,5)}。また、そこで使用されているデータはリモートセンシングで得られた観測データであるため、将来時刻の漁場予測という目的での利用はできない。

本プロジェクトでは、海洋シミュレーションデータを環境要因として用いることで、未来の環境要因の予測とデータの高精細化が達成されピンポイント漁場予測が可能となった。一方で、データの高精細化によって、モデル

ルに組み込む環境要因の候補数が増大し、モデル構築における環境要因選択の作業負担が増大してきた。

また、HSI モデルは従来、専門家が統計分析ソフトウェアなどを用いて構築する場合が多く、非専門家が HSI モデリングに参画するのは困難である。しかしながら、HSI モデルの精度を高め、漁場予測の精度を高めるためには、海洋物理学者、海洋生物学者、統計学者といった専門家だけでなく、実際に漁業に携わる漁業者も含めた様々なステークホルダーの知見を取り入れる必要がある。しかし、非専門家ステークホルダーの典型である漁業者の中にはマウスとキーボードを使った従来の PC 操作に不慣れな人も少なくなく、システムのユーザビリティに配慮しなければ効果的な参画は望めない。

本研究の主な目的は、(1)多数の環境要因からモデル構築を行う作業の効率化と(2)漁業者などの多様なステークホルダーのモデリングへの参画の実現の二点である。本研究では、キーボードとマウスによる PC 操作に不慣れなステークホルダーにも HSI モデリング作業への参画が可能ないようにユーザビリティを改善した視覚分析環境を提案する。提案システムはタブレット端末向けに最適化され、マウスやキーボードを必要とせず、タッチによる直感的な操作が可能ないように設計された。

本稿の残りの構成は次のとおりである。第2節では、本研究で対象とする HSI モデリングプロセスの詳細を述べる。第3節では、本研究で開発した HSI モデリングの視覚分析環境の設計と実装について述べる。第4節では、本研究で開発した視覚分析環境のユーザー評価について述べる。そして第5節では、まとめと今後の課題について述べる。

2. 生息場適性指数 (HSI) モデリング

HSI モデルは複数の環境要因から生息場適性を決定する関数であると考えることができる。生息場適性は、単位努力量当たり漁獲量 (CPUE: Catch Per Unit Effort) などを用いて表される。 n 個の環境要因において、海上の地点 j の環境要因値は $(x_{j1}, \dots, x_{jn})$ と記述できる。

* 原稿受付 2014 年 7 月 15 日

** 非会員 京都大学大学院 工学研究科 電気工学専攻
(〒606-8501 京都市左京区吉田二本松町,
E-mail: onoue@viz.media.kyoto-u.ac.jp)

*** 正会員 京都大学 国際高等教育院

Fig. 1 An image of SI model.

HSI モデルは、それぞれの環境要因の環境要因適正指数 (SI; Suitability Index) モデルから構成される。SI モデルは一つの環境要因と生息場適性の関係を表した関数である (Fig. 1)。HSI 及び SI の値を地図上にプロットしたものをそれぞれ HSI マップ、SI マップと呼ぶ。HSI 及び SI の値は 0 から 1 の範囲を取るように正規化される。ある地点の値が 0 であることは生息場として全く適さないこと、1 であることは生息場として非常に適していることを表している。

HSI モデルには、SI モデルの構成方法や SI モデルから HSI モデルを構成する方法によって様々なバリエーションが存在する。Tian らは、海面水温、水深 35 メートルの水温、水深 317 メートルの水温、海面塩分濃度、海面高度の 5 つの環境要因を用いた HSI モデルを提案した⁴⁾。Tian らは、SI モデルの構成に平滑化スプラインを採用した。さらに、SI モデルの被説明変数として CPUE と努力量を比較し、努力量を採用した。そして、5 つの SI モデルの幾何平均を HSI モデルとした。Chen らは、海面水温と海面高度の 2 つの環境要因を用いて複数のモデル構成法の比較を行った⁵⁾。Chen らは、SI モデルの被説明変数として CPUE、努力量それぞれの対数を用いて、二次式への回帰を行った。Chen らは、HSI モデルの構成においても幾何平均と算術平均の比較を行った。本研究では、SI モデルに CPUE を用いた平滑化スプライン、HSI モデルの構成に幾何平均を用いる方法を対象とする。

i 番目の環境要因に対する SI モデルは以下の式で表される。

$$SI_i(x_{ji}) = \frac{f_i(x_{ji}) - \min f_i}{\max f_i - \min f_i} \quad (1)$$

$f_i(x_{ji})$ は、漁獲高と環境要因の関係を表した回帰曲線である。 $\min f_i$ と $\max f_i$ は、それぞれ環境要因の定義域における $f_i(x_{ji})$ の最小値、最大値である。 $f_i(x_{ji})$ の推定には、3 次平滑化スプラインによるノンパラメトリック回帰を用いる。モデル作成者は、海洋生物学及び海洋物理学の知見に基づいて、曲線の形状や SI マップを確認しつつ適切な平滑化係数を選択する。

Fig. 2 は、海面塩分濃度についての SI モデルの例である。グラフの横軸が海面塩分濃度、縦軸が CPUE 値

Fig. 2 SI model of sea surface salinity. X-axis and y-axis mean salinity and logarithm of CPUE, respectively. Points are fish catch data and the red line is a regression curve.

Fig. 3 SI map of sea surface salinity.

の対数を表している。プロットされている点が漁獲 (CPUE) データ、赤線が回帰曲線を表している。Fig. 3 は、このモデルに対する SI マップの例で、範囲は北緯 34 度から 46 度、東経 180 度から 200 度である。赤い領域は SI 値が高く、青い領域は SI 値が低いことを表している。

幾何平均による HSI モデルは以下の式で表される。

$$HSI(x_{j1}, \dots, x_{jm}) = \sqrt[n]{\prod_{i=1}^n SI_i(x_{ji})} \quad (2)$$

幾何平均モデルにおいて、ある海域の HSI 値が高いために、全ての環境要因の SI 値が高い必要がある。

従来のアカイカの HSI モデリングでは、環境要因としてリモートセンシングで得られた観測データが用いられていた。そのため、深さ方向の詳細なデータを得るには限界があり、HSI モデリングに使用できる環境要因数にも限りがあった。近年の海洋シミュレーションの大規模化や解像度向上、精度向上に伴ってそれらのデータが HSI モデリングの環境要因として利用されるようになった。このような海洋シミュレーション結果は従来とは異なり、深さ方向の詳細なデータを含んでいる。海洋現象は、水平方向の影響と比較すると、深さ方向の影響は非常に軽微であり、異なる深さのスライスを異なる環境要因として利用することができる。その結果、利用可能な環境要因数が急激に増大し、HSI モデルに使用する環境要因の選択という問題が生じてきた (Fig. 4)。HSI モ

Fig. 4 SI selection.

デルに組み込む環境要因数の増大は、過学習を引き起こしモデルの頑健性を損なう。特に、本研究で対象とする HSI モデルはピンポイント漁場予測を目的としており、未来の未知のデータに対して適切な予測ができる必要があるため頑健性を保ったままモデル構築をする必要がある。

実際に、本研究で扱うデータが含む環境要因の候補数は、深さレベル 30 の 3 次元データが 5 個と 2 次元データが 3 個の合計 153 個に上る。先行研究で用いられた環境要因数が Tian らのモデル⁴⁾で 5 個、Chen らのモデル⁵⁾で 2 個であることからすると大幅な増大であり、10 個程度まで環境要因数を絞り込むには多大な労力が必要である。環境要因数の増大は、精度の高い HSI モデル構築に有用であるが、統計学、海洋物理学、海洋生物学の全ての観点から齟齬のないモデルを構築するための負担も大幅に増大した。そのため、相関係数のような統計指標だけではなく、SI 曲線の形状や、SI マップ、HSI マップを見ながら、モデルの判断を総合的かつ効率よく行う必要が生じてきた。また、以上のような HSI モデリング作業は、専門家が統計分析ソフトウェアなどを利用して行うことが一般的であり非専門家が行うことは容易ではなかった。そのため、漁業者などの多様な非専門家ステークホルダーの暗黙知の集約も困難であった。

3. HSI モデリングの視覚分析環境

3.1 視覚分析環境の要求要件

提案システムの開発にあたってはじめに、HSI モデリングのワークフローを分析することで、提案システムに対する要求を明確化した。HSI モデリングのワークフローは、初期設定と SI モデル構築、HSI モデル構築の 3 工程に大きく分かれる (Fig. 5)。

初期設定では、対象海域、漁獲情報、予測対象日などの入力を行う。対象海域や漁獲情報を変更する場合、それまでに構築したモデルは無効になるのでこれらはシステム起動時に入力を行い変更はしないものとする。予測対象日は、モデルを無効にすることなく変更できるため、モデリング途中での変更を可能にする。予測対象日を変更しつつ HSI マップ、SI マップを確認することはモデルの頑健性を確認する上で必要である。

SI モデル構築は、(1)環境要因の探索、(2)回帰曲線の調整、(3)妥当性検証の 3 ステップから構成され、十分な数の SI モデルを構築するまでこれらのステップを繰り返す。

Fig. 5 Procedure of HIS modeling.

返す。(1)環境要因の探索では、環境要因の種類（水温、塩分濃度、流速など）とスライスする深さの決定を行う。ここでは、環境要因値が地図上でどのように分布しているか確認することや、選択された環境要因と CPUE の相関を深さ毎に確認することが必要になる。(2)回帰曲線の調整では、環境要因と CPUE の回帰曲線がオーバーフィッティングなどの不適切な状態にならないように、回帰曲線として使用する平滑化スプライン曲線の平滑化係数の調整を行う。(3)妥当性検証では、回帰曲線の当てはまり具合や、SI マップの状態を確認しつつ構築した SI モデルが HSI モデルの部品として適切か検証する。予測対象日として過去の日付を使用することで、構築したモデルと実際にあった漁獲の情報を照らし合わせることができ、妥当性検証においては重要な作業である。以上のワークフローから、SI モデル構築作業には、環境要因値及び SI 値の地図上へのマッピングや、環境要因値と CPUE の変数空間上での関係の表示、回帰曲線の表示、深さ毎の回帰曲線の相関係数の表示が必要であることがわかる。また、ユーザーからシステムへの操作は、変数の種類の選択、スライスする深さの選択、平滑化係数の選択、構築した SI モデルの保存の 4 つである。

HSI モデル構築では、構築した SI モデルの中から最終的に HSI モデルに組み込むものを選択する。このときに必要なのは、構築した SI モデルの回帰曲線の形状及び SI マップを HSI マップと同時に再度確認できることである。ここでのユーザーの操作は、回帰曲線と SI マップを表示する SI モデルを選択することと、選択した SI モデルの HSI モデルへの追加及び削除である。

また、非機能要件として、マウスとキーボードによる操作に不慣れた漁業者にも扱いやすいユーザーインターフェースの実現が挙げられる。

3.2 視覚分析環境の実装

前節の要求要件に基づいて、提案システムの設計を行った。第一に、マウスとキーボード操作に不慣れた

が一覧されている。チェックボックスを用いて HSI モデルへの使用・不使用を選択することができる。また、ボタンによって SI マップ、散布図への表示をさせることができる。SI モデル構築画面と同様にユーザーが操作する部分は画面右側に配置している。

提案システムは HTML, CSS, JavaScript といった Web 標準技術を用いた Web アプリケーションとして実装され、Web ブラウザのみで利用することができる。Web 標準で実装することで、システムの動作環境をタブレットだけに限定することなく、高解像度ディスプレイを備えたデスクトップ PC などにも対応した。また、システムが Web アプリケーションとして提供されることで、インストール作業が不要となり、各ステークホルダーへシステムを配布・更新する手間が削減された。Web ブラウザが搭載している WebGL や SVG (Scalable Vector Graphics) といったグラフィック描画 API を利用することで、Web アプリケーションでも高速なグラフィック描画を行うことができる。環境要因、SI 及び HSI のマップビューは WebGL で描画されている。WebGL は、組み込み向けの OpenGL サブセットである OpenGL ES 2.0 と同等の能力を持っている⁶⁾。WebGL の命令は JavaScript から呼び出されるが、効率的な GPU アクセラレーションによってネイティブアプリケーションと同等の性能を発揮することができる。一方で、関連グラフと散布図は JavaScript ベースの可視化ツールキットである D3.js⁷⁾を用いて SVG への出力を行っている。D3.js は平面的なチャートの描画を効果的に行うことができる。

環境要因データの取得にデータ転送のアーキテクチャとプロトコルである OPeNDAP⁸⁾を採用することで、タブレット端末上でのストレージ消費量の削減を図った。本研究で予測に使用する環境要因データの総容量は約 2 テラバイトであり、全てのユーザーに全てのデータを配布するにはコストがかかる。特に本システムが対象とするようなタブレット端末は内蔵ストレージ容量や外部ストレージの接続に制限がある場合も少なくはないため、この課題を解決する必要があった。実際には、一度の HSI モデリングで使用されるデータは全データの一部であるため、OPeNDAP を用いてオンデマンドなデータ取得を行うことでデータ配布の問題を解決した。OPeNDAP によって、必要なデータをその都度ネットワーク越しに OPeNDAP サーバーから取得するため、通信量の増大が懸念される。しかし、単精度浮動小数点数を 100×100 個含んだ二次元格子のデータ量は約 16KB である。予測に使用する一つの環境要因のサイズはこの程度であり、100 個の環境要因をモデルとして検討した場合でも総通信量は約 1.6 メガバイトであり、通信量の問題は生じない。さらに、OPeNDAP の採用によって、アプリケーションとデータの境界が明確になったため、異なる環境要因などのデータセットの切り替え

が容易となった。

4. ユーザー評価

本システムの主要なユーザーの一人である HSI モデル作成者を評価者として、インタビューによるシステムの有効性の評価を行った。評価者からは好意的なコメントが多数得られた。得られたコメントの一部を以下に抜粋する。

- ・各変数の分布と SI, HSI の分布の比較が容易になった。
- ・色別に漁獲量の差がわかるので、HSI との対応関係が見やすくなった。
- ・表示している深度が赤点表示になっているのはわかりやすくいい。表示は層の番号が書かれているが、深度(m)もあるとさらにいい。計算している関連の性格上、マイナスの領域はもっと狭くていいのではないか。プラスの領域の幅をもっと広げてくれると見やすいと思う。赤点の脇に相関係数の値が表示できるとさらにいい。
- ・漁獲点のプロットとスプラインカーブが同時に見られるのは非常にわかりやすく、またスプラインの平滑化係数の調整が図上のカーブに反映されるので、適切な自由度を決定するのに非常に役に立つ。平滑化係数は深度によって異なる値を使う事になると思うので深度ごとに設定できるようになるとさらにいい。
- ・SI モデル及び HSI マップ作成工程の全てを対話的に行う事ができるので、モデル構築に必要な試行を整理しながら行う事が可能であるとともに、やり直しや反復作業も容易に行う事ができる。
- ・変数及び鉛直層毎のスプラインカーブの決定やそれを統合した HSI 値の算出、さらにその結果の予測値・解析値へのマッピングの全てを瞬時に行う事ができるので、間違いの少ない HSI モデルをこれまでより短い時間で構築することが可能である。
- ・現在開発中の HSI モデルでは非常に多くの変数の組み合わせを扱う必要があり、しかもその作業を統計学的視点、海洋物理学的視点、海洋生物学的視点の 3 つの立場から齟齬のないモデルを構築する必要がある。これらの必要条件を満たすために必要な多くの試行錯誤を効率的に行う事ができる本システムは、単に統計的に高い精度が得られるモデルを構築するだけでなく、対象となる生物の好適生息域についての新知見を得るための情報抽出が可能であるという観点から非常に有効であると考えます。
- ・現状の作業を著しく効率化できるという観点からは大変満足である。今後は、実利用試験を通じて更なる利便性の向上を図るとともに、機能拡張を行う事ができればさらに頼もしいシステムとなると思う。

評価者のコメントから、関連グラフの表現などに課題は残るものの、提案システムは HSI モデリングを視覚

的にサポートすることで作業の効率化を実現できていることがわかる。さらに、HSI モデリング作業の効率化は、モデリングにおける試行錯誤を支援し、対象生物種に関する新知見を得る上で有用であることが示唆された。一方で、システムの操作性に関するコメントはあまり多く得られなかった。操作性に関する評価は、ユーザー代表者によるインタビューだけではなく、多数のユーザーによる定量的な調査が必要と考えられる。

5. おわりに

本研究では、HSI モデリングによるアカイカのピンポイント漁場予測を対象に、視覚分析環境の設計・実装を行うことで HSI モデリング作業の効率化及び新知見の発見の促進を行った。提案システムは、タブレット端末向けに最適化された画面レイアウトであり、直感的なタッチ操作によって HSI モデリング作業を行うことができる。また、システムを Web アプリケーションとして実装し、大容量の海洋シミュレーションデータを OPeNDAP で配信することで、システムのインストールの負担などが軽減され、容易にシステムの利用を始めることができるようになった。提案システムを使用することで、今後は多様なステークホルダーの HSI モデリングへの参画が期待される。

本研究では、キーボードとマウスによる操作に不慣れたステークホルダーでもシステムが利用できるようにユーザビリティ設計を行った。しかしながら、ユーザビリティに関する定性的・定量的な評価は不十分であり、フォーマルなシステム評価によってシステムの有効性を明らかにすることは今後の課題である。

異なる魚種や異なる環境要因データセットに対しての提案システムの適用や機能面の拡張も今後の課題として残っている。本研究ではアカイカの HSI モデリングを対象としたが、漁獲データ及び環境要因データをシステムの入力形式に揃えれば他のデータに関しても適用は可能である。様々なデータで実証実験を行うことにより、システムの評価・改善を行っていきたい。機能面では、海洋環境の俯瞰的可視化や渦情報の可視化手法を本システムに取り込むことで、従来の HSI モデリングで考慮

できていなかった点の改善を行う。これらの技術は既に活発に研究がされているものであるが、これらの機能を現行の Web ベースのシステムに統合し、効率よく動作するように実装することは挑戦的な課題である。

3 次元高精細データを用いた HSI モデリングの研究は未だに発展途上である。視覚分析環境の整備によって、HSI モデリングプロセスの改良及び対象魚種の暗黙知の集約が促されることを期待したい。

謝 辞

本研究は、文部科学省「気候変動適応推進プログラム (RECCA)」及び文部科学省「地球環境情報統融合プログラム (DIAS)」の支援により実施された。

参 考 文 献

- 1) 文部科学省, 気候変動適応研究推進プログラム, <http://www.mext-isacc.jp/>.
- 2) 文部科学省, 気候変動に伴う水産資源・海況変動予測技術の革新と実利用化, http://www.mext-isacc.jp/staticpages/index.php/report_awaji_j.
- 3) United States Department of the Interior : Habitat evaluation procedures handbook, (1980).
- 4) Tian, S., Chen, X., Yong Chen, Xu, L., Dai, X. : Evaluating habitat suitability indices derived from CPUE and fishing effort data for *Ommastrephes bartramii* in the northwestern Pacific Ocean, Fisheries Research, Vol.95 No.2-3 (2009) pp. 181-188.
- 5) Chen, X., Tian, S., Liu, B., Chen, Y. : Modeling a habitat suitability index for the eastern fall cohort of *Ommastrephes bartramii* in the central North Pacific Ocean, Chinese Journal of Oceanology and Limnology, Vol.29 No.3 (2011) pp.493-504.
- 6) Marrin, C. : WebGL Specification, Khronos WebGL Working Group, (2011).
- 7) Bostock, M., Ogievetsky, V., Heer, J. : D3 : Data-Driven Documents, IEEE transactions on visualization and computer graphics, Vol.17, No.12 (2011) pp.2301-2309.
- 8) Cornillon, P., Gallagher, J., Sgouros, T. : OPeNDAP : Accessing data in a distributed, heterogeneous environment, Data Science Journal, Vol.2 (2003) pp.164-174.