

## 科学研究費助成事業 研究成果報告書

平成 29 年 5 月 22 日現在

機関番号：15301

研究種目：基盤研究(C)（一般）

研究期間：2014～2016

課題番号：26330041

研究課題名（和文）大規模時空間情報の階層構造分析のためのプラットフォーム構築とその応用

研究課題名（英文）Platform construction for hierarchical structure analysis of large scale spatio-temporal data and its application

研究代表者

栗原 考次（Kurihara, Koji）

岡山大学・環境生命科学研究所・教授

研究者番号：20170087

交付決定額（研究期間全体）：（直接経費） 3,600,000円

研究成果の概要（和文）：本研究では、時間の推移とともに多次元空間の領域で多変量情報が観測されるマルコフ確率場において、階層構造をもつ同位相領域の分類及びホットスポット検出を種々の分野に統一的な枠組みでプラットフォームを構築した。さらに、大規模情報も含めた多変量時空間情報に対する真の集積性や複雑性に関する新手法を確立した。また、ホットスポット検出及び地域構造分析のためのアルゴリズム作成及びソフトウェア化を各分野に統一的な枠組みで行なった。

研究成果の概要（英文）：In this study, we have constructed a platform with a unified framework for hierarchical classification of in-phase and hotspot detection in a Markov random field where multivariate information is observed in spatiotemporal space. In addition, we established a new method on true integration and complexity for multivariate spatiotemporal information including large-scale information. We also created algorithms and software for hotspot detection and regional structure analysis in a unified framework for each field.

研究分野：総合領域

キーワード：時空間情報 エンセロン解析 ホットスポット

## 1. 研究開始当初の背景

近年、各分野において種々の情報は、時間・空間的位置情報をともなった大規模データとして得られることが多いが、データの可視化やデータに含まれる特性を簡潔に要約するとともに内存するメカニズムを的確に表現するモデルの構築が重要である。

海外共同研究者の Myers 博士は、2次元空間データの位相的構造を可視化するエシェロン解析法(1997)を提案した。研究代表者の栗原は、この考え方を応用し、3次元空間上の位置情報を持つ空間データの2次元空間上へ可視化(Kurihara, 2004)するためのアルゴリズムを提案した。近年、栗原や共同研究者を中心に、時空間データのエシェロン階層構造分類及びホットスポット検出に関する理論及び環境、生態、医学、公衆衛生などの分野への応用に関する最先端の研究が推進され、多変量空間データへの理論と応用(Kurihara et al., 2006)、エシェロン構造と尤度に基づくホットスポット検出(Myers et al., 2006)、3次元時空間データのホットスポット検出と環境汚染データの影響評価(Ishioka et al., 2010)、エシェロン階層構造を利用したゲノムデータのLDブロック同定(Tomita et al., 2007)、エシェロン階層構造を利用した森林の分類(Oda et al., 2012)、自殺データのホットスポット検出(Kubota et al., 2013)、多次元時空間データの分類(Kurihara et al., 2013)等の研究成果が公表されている。こうした状況下、本研究では、種々の分野に統一した枠組みで、時空間情報のエシェロン階層構造分類のためのプラットフォームを構築し、大規模データも含めた真の集積性や複雑性に対する理論と応用の研究を行う。

## 2. 研究の目的

本研究では、時間の推移にとともに多次元空間の領域で多変量情報が観測されるマルコフ確率場において、空間的相互作用を明確にしたモデルの基に階層構造をもつ同位相領域

の分類及びホットスポット検出を種々の分野に統一した枠組みでプラットフォームを構築するとともに、マルコフ確率場における大規模情報も含めた多変量時空間情報に対する真の集積性や複雑性に関する新手法を確立することを目的としている。さらに、リモートセンシングを含む画像解析学、疾病率や自殺データなどを対象とした疫学、放射能汚染率などの環境学、地価や商圈などの経済学等で見られる大規模時空間情報に対して、各分野に統一した枠組みでホットスポット検出及び地域構造分析のためのアルゴリズム作成及びソフトウェア化を行い、その応用分野の発展に寄与する。具体的には、以下のような問題を取り扱う。

- (1) 位置情報の多次元化とホットスポットの検出
- (2) 地域データに対するピークを中心とした圏による分類
- (3) 時空間情報を利用した大規模放射能データの解析
- (4) 河川の生態構造分析
- (5) 小規模格子データに対する直接計算

## 3. 研究の方法

時間・空間的位置情報をともなったデータに対して、データに内存するメカニズムを的確に表現するモデルやホットスポットを検出するためのプラットフォーム構築のために、研究の目的で記述した5つの課題に以下のような研究の方法で取り組む。

- (1) 位置情報の多次元化とホットスポットの検出

3次元空間では、規則格子データに対して、6、26近傍系における同位相分類を行うとともに、具体的な3次元データに対してホットスポット検出を行う。

4次元空間上の超ボクセルデータは、可視化ですら困難であり、具体的に観測すること自体無理がある。本研究では、(1) 3次元空間上のボクセルデータがいくつ

かの時点で観測された4次元時空間データの8近傍系での分類、(2)4次元空間上の超ボクセルデータの8、80近傍系での分類、を行う。また、順序カテゴリーデータから生成される4次元の分割表に空間構造を取り入れる新手法について検討する。

(2) 地域データに対するピークを中心とした圏による分類

本研究では、都道府県別などの地域データに対して、エシェロン解析で得られた階層構造に基づき、ピークの都道府県及びピークに隣接する都道府県により構成される圏を客観的かつ一意に検出する方法を新たに提唱する。

(3) 時空間情報を利用した大規模放射能データの解析

本研究では、横断的に得られる大規模放射能クロスセクショナルデータに対して、各時間の隣接情報を定義し、エシェロン情報に基づき時空間に渡る階層構造を求め、ホットスポットの時間的な遷移を調べる。この改良により、ホットスポットの時間的な拡大、縮小、移動、合併などの時間的な遷移の検出が可能となる。

(4) 河川の生態構造分析

本研究では、海洋や河川における魚類の居住状況データに対して、エシェロン階層構造に基づき同位相を持つパッチ領域を求め、海洋や河川の生態構造分析を行う。

(5) 小規模格子データに対する直接計算

小規模格子データに対しては、安易に計算機支援によりモンテカルロ法などに依ることなく、統計量の理論的な確率分布を求めることが望まれる。本研究では、北海道大学の湊教授らとの共同研究において、ZDD法を利用することにより、空間スキャン統計量の計算に必要となる隣接ブロックの数え上げの直接計算を行う。

#### 4. 研究成果

(1) 位置情報の多次元化とホットスポットの検出

4次元空間データとして、(1)3次元空間上のボクセルデータがいくつかの時点で観測された4次元時空間データ、(2)4次元空間上の超ボクセルデータ、の2種類の空間データに対するホットスポット検出への拡張について研究を行った。(1)では、環境汚染物質データに対して、8近傍系での同位相分類及びホットスポット検出を行った。(2)では、4次元空間において、8、80近傍系を定義し、人工データでの同位相分類及びホットスポット検出を行った。また、順序カテゴリーデータから生成される4次元の分割表に空間構造を取り入れ、セルの分類を行なうとともに、有意に連関の源になっているセル群の検出を行う解析法について検討した。

(2) 地域データに対するピークを中心とした圏による分類

地域データの例として、都道府県別人口密度データに対して、本研究で開発したピークを中心とした圏による分類を適用し、関東圏、関西圏、中部圏、九州圏、中四国圏、東北北海道圏、北陸圏に分けられることを示した。

(3) 時空間情報を利用した大規模放射能データの解析

日本原子力研究開発機構(JAEA)が公開している福島第一原子力発電所近隣の帰還困難区域7村の2012年4月から2015年3月までの空間線量率データに対して、時空間的観点から空間線量率が有意に高い領域(ホットスポット)の検出を行った。ホットスポットの時間的遷移の検出が可能でエシェロンスキャン法を利用することにより、時間と共にホットスポットの領域が縮小していることを解明した。

#### (4) 河川の生態構造分析

河川におけるサーモン居住に関する生態調査には、レーザーを用いる方法と簡略版のソナーを用いる2種類の測定法がある。2種類の測定法の比較するため、各々の測定法に対して生息場適性指数を用いてエシロン階層構造を求め、パッチを検出することにより、簡略版のソナーの特性について検討した。

#### (5) 小規模格子データに対する直接計算

本研究では、北海道大学の湊教授らとの共同研究において、ZDD法を用いることにより全国の都道府県のホットスポット検出の際必要となる都道府県の組合せからなる隣接ブロックを網羅的に数え上げに初めて成功し、組合せ隣接ブロックの総数が1098億5712万5702通りであること公表している。

これらの研究成果は、台北(台湾)で開催された第3回IMS-APRM学会での招待講演、京都で開催されたKyoto International Conference on Modern Statistics in the 21st Century、シンガポールで開催された第9回国際計算機統計学会での招待講演、ポロニヤで開催された国際分類学会(IFCS2015)等の国際会議、2016年にオビエドで開催された欧州地区国際計算機統計学会(COMPSTAT2016)等の国際会議及び統計関連学会連合などの国内学会で講演するとともに、国内外の査読論文として公表した。

#### 5. 主な発表論文等

[雑誌論文](計9件)

Ishioka, F. and Kurihara, K., Detection of space-time clusters for radiation data using spatial interpolation and scan statistics, *Proceedings of COMPSTAT2016 (22nd International Conference on Computational Statistics)*, 査読有, 2016, 85-97, ISBN: 978-90-73592-36-0.

Na, M., Moon, S. and Kurihara, K., Linear

Discriminant Analysis Technique Based on Tuning Parameters for Enhancing Discriminant Power, *Proceedings of 2015 International Workshop for JSCS 30th Anniversary in Okinawa*, 査読有, 2015, 11-12.

Hayashi, K. and Kurihara, K., Detecting a change point using statistical sensitivity analysis based on the influence function. *2014 Joint 7th International Conference on and Advanced Intelligent Systems (ISIS)*, *IEEE*, 査読有, 2014, 506-511, DOI:10.1109/SCIS-ISIS.2014.7044767.

Kim, S., Hayashi, K. and Kurihara, K., Geostatistical Data Analysis with Outlier Detection, *Journal of the Korean Data Analysis Society*, 査読有, 2014, 16(5), 2285-2297.

石岡文生、林邦好、植田琢也、水藤寛、栗原考次, 胸部大動脈ステント治療に関する統計的アプローチ, *応用数理*, 査読有, 2014, 24(2), 2-10.

Hayashi, K., Ishioka, F., Raman, B., Sze, Y., Suito, H., Ueda, T. and Kurihara, K., Assessment of the Relationship Between Native Thoracic Curvature and Endoleak Formation after TEVAR Based on Linear Discriminant Analysis, *German-Japanese Interchange of Data Analysis Results*, (Edited by W. Gaul et al.), Springer, 査読有, 2014, 179-192, DOI 10.1007/978-3-319-01264-3\_16.

林邦好、石岡文生、植田琢也、水藤寛、栗原考次, 胸部大動脈瘤ステントグラフト治療に関わる予後予測, *計算機統計学*, 査読有, 2014, 26(2), 59-77.

Hayashi, K., Ishioka, F., Raman, B., Sze, D., Suito, H., Ueda, T. and Kurihara, K., Statistical Assessment for Risk Prediction of Endoleak Formation after TEVAR Based on

Linear Discriminant Analysis, *Analysis and Modeling of Complex Data in Behavioural and Social Sciences*, (Edited by D. Vicari, A. Okada, G. Ragozini, C. Weihs ), Springer, 査読有 , 2014, 143-151.

Kim, S., Hayashi, K. and Kurihara, K., The optimal number of lags in variogram estimation in spatial data analysis, *COMPSTAT 2014 Proceedings in Computational Statistics*, 査読有 , 2014, 507-514.

[学会発表](計29件)

石岡文生、栗原考次、ZDDによるグラフ列挙技法を利用したホットスポット検出と空間スキャン法の性質評価、第37回大規模データ科学に関する研究会、2017.2.3、北海道大学。

梶西将司、石岡文生、栗原考次、空間疫学における集積性の検出Shinyへの実装、日本計算機統計学会第30回シンポジウム、2016.11.25、沼津市。

石岡文生、栗原考次、空間スキャン統計量による集積性検出の新たなアプローチ、第33回大規模データ科学に関する研究会、2016.9.23、北海道大学。

石岡文生、栗原考次、空間データに対するホットスポット検出手法の性質評価について、2016年度統計関連学会連合大会、2016.9.5、金沢大学。

小田牧子、Saija Koljonen、石岡文生、Petteri Alho、水藤寛、Timo Huttula、栗原考次、ハビタットの空間的分割とその比較、2016年度統計関連学会連合大会、2016.9.5、金沢大学。

小田牧子、久保田康裕、楠本聞太郎、正木隆、石岡文生、栗原考次、パッチベースでの森林の比較について、第127回日本森林学会大会、2016.3.29、藤沢市。

石岡文生、栗原考次、放射線量モニタリングポストデータに基づいた空間補完と

その時空間クラスターについて、第32回大規模データ科学に関する研究会、2016.3.4、北海道大学。

栗原考次、石岡文生、放射線量モニタリングポストデータの位相的階層構造を利用した時空間集積性の検出、科学研究費・基盤研究(A)科研費シンポジウム「空間データと災害の統計モデル」、2015.12.24、同志社大学。

Ishioka, F. and Kurihara, K., Detection of cluster for radiation monitoring data based on scan statistic, The 9th Conference of the Asian Regional Section of the IASC, (Invited Paper), 2015.12.17, Singapore.

Oda, M., Koljonen, S., Ishioka, F., Alho, P., Suito, H., Huttula, T. and Kurihara, K., A novel approach for comparing spatial data obtained by different measurement systems, The 9th Conference of the Asian Regional Section of the IASC, (Invited Paper), 2015.12.17, Singapore.

石岡文生、栗原考次、エシェロン解析を利用した放射線量モニタリングポストデータの時空間クラスターの検討、統計数理研究所共同研究集会「環境・生態データと統計解析」、2015.11.6、統計数理研究所。

石岡文生、富田誠、久保田貴文、栗原考次、市区町村別自殺データの空間集積性の検討、第31回大規模データ科学に関する研究会、2015.8.24、北海道大学。

Ishioka, F. and Kurihara, K., Space-time clustering for radiation monitoring post data based on hierarchical structure, 2015 Conference of the International Federation of Classification Societies, 2015.7.8, Bologna.

Hayashi, K. and Kurihara, K., Analysis of influence scores for detecting a change point, 2015 Conference of the International Federation of Classification Societies,

- 2015.7.7, Bologna.
- 小田牧子、久保田康裕、楠本聞太郎、石岡文生、栗原考次、スキャン統計量を利用した森林帯の評価、日本計算機統計学会第29回大会、2015.5.15、山梨県立図書館。
- 石岡文生、栗原考次、放射線量モニタリングポストデータに対する時空間集積性の検討、日本分類学会第33回大会、2015.3.2、帝京大学。
- 石岡文生、栗原考次、放射線量モニタリングポストデータに対する時空間集積性の検出について、第30回大規模データ科学に関する研究会、2015.2.27、北海道大学。
- 栗原考次、石岡文生、福島第一原子力発電所事故に伴う放射性物質の時空間分布、科学研究費・基盤研究(A)科研費シンポジウム「空間データと災害の統計モデル」、2014.12.21、アクロス福岡。
- 石岡文生、栗原考次、空間集積検出手法に基づいた福島第一原発事故に伴う放射性物質のクラスター検出について、統計数理研究所共同研究集会「環境・生態データと統計解析」、2014.12.5、統計数理研究所。
- Kurihara, K. and Ishioka, F., Spatial Zone Clustering for Lattice Data based on Hierarchical Structure, Kyoto International Conference on Modern Statistics in the 21st Century, 2014.11.17, Kyoto.
- 21 小田牧子、久保田康裕、楠本聞太郎、石岡文生、栗原考次、空間的視点からの森林構造表現と比較、日本計算機統計学会第28回シンポジウム、2014.11.15、沖縄科学技術大学院大学。
- 22 羅明振、栗原考次、チューニングパラメータに基づく判別分析法、日本計算機統計学会第28回シンポジウム、2014.11.15、沖縄科学技術大学院大学。
- 23 小田牧子、Saija Koljonen、Timo Huttula、Petteri Alho、水藤寛、栗原考次、空間的階層構造を利用したサーモンの生息場評価法、2014年度統計関連学会連合大会、2014.9.15、東京大学。
- 24 富田誠、石岡文生、久保田貴文、西山毅、栗原考次、遺伝・精神保健データに対する集積性手法の実践、2014年度統計関連学会連合大会、2014.9.15、東京大学。
- 25 石岡文生、栗原考次、福島第一原子力発電所事故に伴う放射線物質モニタリングポストデータの時空間集積性の検討、2014年度統計関連学会連合大会、2014.9.15、東京大学。
- 26 林邦好、石岡文生、井上幸平、中玉利葉月、植田琢也、水藤寛、栗原考次、樹木モデルに基づく肝臓癌の悪性度に関する判別について、2014年度統計関連学会連合大会、2014.9.15、東京大学。
- 27 石岡文生、栗原考次、大規模空間データに対する集積検出手法の評価について、第27回大規模データ科学に関する研究会、2014.7.22、北海道大学。
- 28 Kurihara, K. and Ishioka, F. Spatial Clustering for Four-dimensional Cubic Data Based on Hierarchical Structure, The 3rd Institute of Mathematical Statistics Asia Pacific Rim Meeting, (Invited paper), 2014.7.3, Taipei, Taiwan.
- 29 金秀証、林邦好、栗原考次、空間隣接情報を考慮した空間データの分析、日本計算機統計学会第28回大会、2014.5.17、中央大学。

〔その他〕

ホームページ等

<http://www.ems.okayama-u.ac.jp/kurihara/>

6. 研究組織

(1) 研究代表者

栗原 考次 (KURIHARA KOJI)

岡山大学・大学院環境生命科学研究科・教授  
研究者番号：20170087