

**科学研究費助成事業 研究成果報告書**

平成 27 年 6 月 30 日現在

機関番号：82657

研究種目：基盤研究(A) (一般)

研究期間：2010～2014

課題番号：22240030

研究課題名(和文) 科学的政策決定のための統計数理基盤整備とその有効性実証

研究課題名(英文) The Infrastructure Development of Statistical Analysis for Evidence-based Policy Making, and Verifying Validity

研究代表者

北川 源四郎 (KITAGAWA, Genshiro)

大学共同利用機関法人情報・システム研究機構(新領域融合研究センター及びライフサイ・新領域融合研究センター)・センター長

研究者番号：20000218

交付決定額(研究期間全体)：(直接経費) 37,700,000円

研究成果の概要(和文)：本研究では、科学的情報収集に基づく社会価値選択、価値を決定する構造モデル導出、価値のモデル上での最適化、最適化された価値の社会への還元からなる情報循環設計を科学的政策決定の統計数理科学的枠組みと位置づけ、政策の科学的決定に資する統計数理体系構築を目的とした。本研究を通じて、公的ミクロ情報分析統計基盤の確立、情報循環加速ツールの開発、時空間可視化ツールの開発を達成し、同成果を自殺予防対策研究、観光政策研究、産業環境政策研究に応用し、それぞれの政策立案に資する新たな知見を得るとともに、データに基づく政策を提言した。

研究成果の概要(英文)：The purpose of this study is to construct frameworks of statistical analysis for risk and evidence based policy making, by establishing the information spiral-up composed of the following steps (1) social value selection based on technical surveillance, (2) construction of the structured model for decision making to attain the social value, (3) the value optimization with respect to the structured model, (4) implementation of the optimized social value. In this study, we had achieved: (1) establishment of statistical analysis-base for micro-data for official statistics, (2) development of the tool for accelerating the information spiral-up (3) development of the tool for visualization of spatio-temporal data. We applied them to the research for suicide prevention policy, tourism policy, and industrial environmental policy, and, proposed various evidence-based policies by obtained new findings on the policy fields.

研究分野：統計科学

キーワード：データサイエンス

### 1. 研究開始当初の背景

統計数理研究所では、2005 年以来 NOE (Network Of Excellence) を 44 研究組織と形成し、価値選択、モデル化、最適化、価値還元というフェイズからなる「情報循環設計科学」の大枠を提唱し、国際標準化活動も展開してきた。この情報学的枠組みは、NOE 活動の中で、災害、医療、ビジネスなど様々なリスク領域に適用可能なことを当該分野専門家と確認してきたところである。更に、「情報循環設計科学」は、2008 年度に横断型基幹科学技術研究団体連合が経済産業省に提出した「分野横断型科学技術アカデミックロードマップ」にも採録された。その際、これがリスク科学のみならず、サービス科学や経営の高度化など、多くの社会課題を解決するエンジンの役割を果たすものと位置づけられた。なお、情報循環設計科学構想の端緒となったのは、日本学術会議 18 期吉田民人副会長の下で「学術の体系」常置委員会が提示した「新しい学術の体系」である。吉田は、社会のための科学、あるべきものを探求する「設計科学」を従来の「あるものを探究」する「認識科学」と対峙させた。これ以来、わが国でも「社会の価値」選択に依存した伝統的工学以外の「社会技術」、Regulatory Science など人間・社会分野の設計科学構築が加速した。「認識科学」の主要な方法論であった実証的仮説検証に代わる方法論として吉田が提唱したのが、「情報循環」というベンチマーキング的試行錯誤サイクルを伴う設計科学の方法論であった。しかし、木村英紀を中心に「新しい学術の体系」の再解釈を行った日本学術会議小委員会では、情報循環論に代表される「設計科学の文法」基盤の脆弱性が課題とされた。これに応じて構想を開始した統計数理科学が「情報循環設計科学」である。

本研究班を政策決定の科学化を目的に、この時期に立ち上げたのは、2008 年の新統計法全面施行により政府収集情報が研究利用可能となり、本研究の有効性を事実に基づいて検証する情報環境がわが国アカデミアに確立したからである。海外では、2004 年に OECD が、「Statistics, Knowledge and Policy: Key Indicators to Inform Decision Making」に関する第 1 回国際フォーラムを実施し、政策決定に対する統計情報利用の意義を強調した。これを受けて、英国下院科学技術委員会は 2006 年に「Risk and Evidence Based Policy Making」に関する 69 項目の勧告を行い、パイロット研究の有効性などを既に認めている。一方、この種の研究を支えるのは、経済学関係者に閉じており、その要素技術たる情報収集、モデリング、意思決定に必要な Risk 最小化数理、及び、これら要素技術の有機的連携などに関する統計数理科学研究並びに情報システムによる支援研究は、重要視されていない。本研究班は、政策決定自体への経済・社会的提案

ではなくて、科学的政策決定に必要な情報循環設計を、基本数理、要素統計数理技術、前提となるデータ基盤、それらを有機的に連携させるための情報システム設計を通じて行うものである。

### 2. 研究の目的

本研究は、科学的情報収集に基づく社会価値選択、価値を決定する構造モデル導出、価値のモデル上での最適化、最適化された価値の社会への還元からなる情報循環設計を科学的政策決定の統計数理科学的枠組みと位置づけ、政策の科学的決定に資する統計数理体系構築を目的とする。このため、別途開発を進めている国土・地域の経済、社会、健康、自然環境など多様な情報に関する多変量時空間系列データベースの完備化、並びに、当該データベースに投入される予定の多様な社会情報に関する構造モデル化を地域・時間変動が捕捉可能なベイズ型モデルで実現し、諸変数から政策関連変数への総合的影響を定量化し、政策評価を行う方法を開発するとともに、一連の方法論の有効性を政府・自治体と共同して自殺・事故予防政策、地域文化・環境継承と経済成長との均衡政策立案に適用して検証する。

### 3. 研究の方法

#### (1) 科学的政策決定に必要な統計数理方法論研究

わが国の科学的政策決定に必要と考える以下の 2 つの統計数理的方法論について的高度化を行い、別途活動する政策決定高度化事例開発の中で適用し、その有効性を実証する。

#### 人間・社会・経済を俯瞰する情報環境開発支援研究

国立情報学研究所と別途共同で開発する国土時空間メッシュデータベースに投入される予定の経済、社会、公衆衛生、自然環境に関する不完備情報に関して時空間統計モデルのフィルタリング技法に基づく仮想完備化技術を研究し、上記データベースに適用する。

#### 人間・社会・経済を俯瞰する構造化モデルの開発と政策による社会制御性の定量化研究

上記データベースに投入予定の多様な情報間に予想される因果構造とその時間変動並びに地域間変動を把握するベイズ型非線形潜在構造モデルを開発し、上記データベースに適用し、わが国の自然環境、人間健康、地域活力、経済活力に関わる諸変数が相互に与えている影響を定量的に把握すると共に、それらの変数の短期予測を可能とする。また、これらアウトカム変数が中央政府ないしは自治体による政策投入でどのように影響を受けてきたかを Impulse 応答ないしは Step 応答を推定することで明らかにする。これに

より立案した政策を経済・社会全般への影響を基に総合評価することが可能となる。

## (2) 科学的政策決定の現状分析に基づく課題抽出と解決

わが国ないしは各国の科学的政策決定のケースを収集し、方法論視点から分析して、情報収集、モデル化、最適意思決定、コミュニケーションの4分野について高度化すべきニーズをそれぞれ1-2件抽出し、それに対して統計数理研究所内ないしはNOEの適切な人材と連携して解決策を研究する。

## (3) 科学的政策決定への統計数理的方法論の適用と効果実証

当研究班が連携可能な次の2つの政策研究について、当研究班が開発する方法論の有効性に関する実証を行う。

### 自殺・事故予防政策の効果予測研究

地域別自殺・事故率情報に影響を与える自然環境、健康、地域活力、経済活力情報などをベイズ型非線形構造モデルによって明らかにし、特定の変数に対する政策目標が将来自殺率や事故率に与える影響を定量化すると共に、必要な政策目標を明らかにし、それを行わない場合の機会損失を評価する。

### 地域文化並びに環境の継承と経済成長をバランスさせる政策の立案研究

京都地区を対象にして、工業化が進展している地域、文化、伝統を継承している地域、自然環境を保存している地域を抽出し、人間・社会・経済・環境変数の相互関係にどのような違いがあるかを構造モデル同定から明らかにし、伝統文化に根付いた社会ないしは自然環境を次世代に継承するという政策目標と、経済成長を両立させる総合的政策立案方法を明らかにする。

## 4. 研究成果

### (1) 科学的政策決定に必要な統計的基盤研究

#### (1-1) 公的ミクロ情報分析基盤の確立

わが国のエビデンスに基づく政策研究を支援するデータベースとその分析拠点を整備するため、(独)統計センターと情報・システム研究機構の間で平成22年7月に連携協定を締結し、「匿名化データ提供拠点」として、総務省4調査に関する匿名化データを統計数理研究所に受け入れる仕組みを確立した。並行して、政府メッシュデータなど入手し、当該分析拠点に投入した。

その後、本分析拠点を政府統計の個票データ等をセキュアな環境で利用するための設備や監視体制を整備し、「オンサイト利用拠点」としての認可を統計センターから受け、全国の研究者がデータに基づく政策研究を可能とする場を構築した。

オンサイト拠点の研究者利用についての政府方針決定の遅れを受けて、公的統計個票

データのみならず、当該拠点に厚生労働省が提供するレセプト・ナショナル・データベース(NDB)の個票を平成23年度上半期に申請し、医薬品の相互作用などの検出に関する安全性研究を推進しようとしたが、探索的研究をオンサイト拠点で行うことは許可されなかった。平成24年度に厚生労働省が開発したNDBのサンプリングデータセットがオンサイト拠点に提供され、島根大学医学部のオンサイト拠点利用を通じて、併用禁忌の医薬品の投与実態などがある程度解明された。オンサイト拠点は、平成26年3月の閣議決定で今後の公的統計ミクロデータ提供のモデルの一つと位置づけられたが、実際に各大学が同様な監視付環境を整備することは難しく、より簡便な方法を学官で開発する必要がある。

この他にも、当研究班は、国立精神神経医療研究センターが総合自殺対策に資することを目的として人口動態統計個票データ等から作成した「自殺統計」開発を支援し、これらのデータ基盤を研究者に提供することで、時空間集積性など自治体の自殺対策に資する資料を提供した。このデータを用いた具体的政策分析については、(3)にのべる。

#### (1-2) 情報循環加速ツールの開発

当研究班は、政策活用可能なデータ収集の基本的研究基盤として、2つに注目した。一つは、(3)で詳しく述べるが、サイバースペース上の情報を自動収集し、それと公的統計情報を用いてバイアスを補正したリアルタイム統計である。特にリアルタイム観光統計の作成とその成果については、(3)を参照されたい。もう一つの方法が、データから得られた分析結果をフィードバックすることによってデータ収集を加速するという枠組みである。これを産業環境政策に適用するために、やはり(3)に示す、「el-Platform」というツールを開発した。

#### (1-3) 時空間可視化ツールの開発

当研究班は、イベントが地域空間ならびに時間軸上どこで多く起きているかといった時空間集積性をEchelon解析に基づき可視化するためのツールを開発した。これは自殺の発生頻度やその年次変化を可視化し、自治体の対策に活かすツールとして活用された。この成果は、国際学会企画セッションなどでも紹介された。

## (2) 科学的政策決定の現状分析に基づく課題抽出と解決

金融政策領域に関する理論研究の成果を整理し、非ガウス型時系列モデリングの深化や信用リスク評価モデルの開発を進めるとともに、ソプリンリスクや流動性リスクに関わる、政府、地方政府、中央銀行の金融政策、商品先物市場に関わる制度設計、政府労働政策への応用について取りまとめをすること

ができた。その一部については、雑誌「統計」2013年6月号に「時系列解析と季節調整の新しい流れ」として特集することができた。

### (3) 科学的政策決定への統計数理的方法論の適用と効果実証

自殺予防政策研究においては、国立精神神経医療研究センターと連携し、国勢調査、人口動態統計を目的外利用申請して受領したデータをリンケージして自殺率に関する地域別時空間データを構築し、その時空間分析ならびにモデル当てはめを実施し、自殺の時空間ホットスポットやクールスポットを抽出し、総合自殺対策に資するデータであることを実証した。その後、本DBは研究者向けに公開しており、年度毎にアップデートを実施している。本研究班でも、本DBに自治体居住地域の全家屋の土地平均傾斜度情報に基づいて市町村居城地域平均傾斜度を作成・投入し、これと自殺率との高度に有意な相関を実証し、高台急傾斜地域で孤立する住民の自殺リスクを明らかにした。また、時空間集積性分析技法ないしはポアソン回帰モデル当てはめに基づく自殺の地域差、自殺要因分析など「総合自殺対策」に関わる分析を行い、国際学会のセッションを企画した。ビッグデータ利用の一例としての自殺に関するtweetデータの分析も実施した。当該研究発表に際しては、法務省矯正局少年鑑別所再入所リスク評価担当者から、方法論的類似性も含め大きな関心が寄せられた。

京都地区における統計に支援された政策決定支援研究においては、観光政策・産業環境政策を重点領域と定め、京都市の協力のもと、研究を開始した。

観光政策研究においては、京都のホテル稼働率のリアルタイム推定、観光ルートの最適化研究などを進めた。その成果の一環としてWebサイトから収集したホテル空室データや観光情報を用いて、日次で宿泊施設の稼働率を推定し、また、宿泊施設であるホテルの人気の要因を分析した。その結果、曜日により稼働率が規則的に変動するなど曜日効果があることがわかった。また、この稼働率の推定方法が、実際の稼働率とどの程度異なるという点については、京都市が発表している京都市内の主要14ホテルの実績稼働率と本研究から得られた推定方法との結果を平成23年11月から平成24年10月の期間で月次で比較し、本研究の推定方法は概ね良好な推定精度であることもわかっており、観光政策の立案に資する新たな知見を得ることが出来た。

産業環境政策については、京都市地場産業からの直接大規模データ収集交渉が最終的に不調に終わったため、インドネシアのボゴール市を対象とするプロジェクトを新たに立ち上げた。環境負荷(Life Cycle Assessment: LCA)データの自動収集と政策反映のために、LCAやコンピュータ操作の知

識が十分でなくても使いやすいシステムを指向するWEBベースの簡易LCAシステム「el-Platform」の開発を支援し、基礎データとして、(一社)産業環境管理協会が管理するLCA基礎データ(産総研が開発)を一部利用可能とした。実際の実証実験については、「el-Platform」と連携するスマートデバイス向けアプリケーションのプロトタイプを開発する一方、現地でのアンケート調査や新潟市のリサイクル事業との比較検証を実施した。その結果、回収される廃食油量や再生エネルギーの60%しか公用車両稼働に利用されていないこと等、両者には共通点が多いことが明らかとなった。また、食用油の温室効果ガス排出量としてのライフサイクルインパクトを、実際の製造所データや報告されているデータを使用して推定した結果、廃食油の排水・土壌の投棄、あるいは完全消費と比較した再生化の場合の環境的有利性は、経済的あるいは技術的な制約により、必ずしも再生活動を助長するものではないことも明らかとなった。

### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計95件)

1. Peng, H., Kitagawa, G., Tamura, Y., Xi, Y., Qin, Y. and Chen, X.: "A modeling approach to financial time series based on market microstructure model with jump" *Applied Soft Computing* 66. 40-51 (2015), 査読有  
DOI: 10.1007/s10463-014-0446-0
2. Oka, M., Kubota, T., Tsubaki, H. and Yamauchi, K.: "Analysis of impact of geographic characteristics on suicide rate and visualization of result with Geographic Information System" *Psychiatry and Clinical Neurosciences* 69. 309-385 (2014), 査読有  
DOI: 10.1111/pcn.12254
3. 岡檀, 久保田貴文, 椿広計, 山内慶太: "日本の自殺率上昇期における地域格差に関する考察 - 1973~2002年全国市区町村自殺統計を用いて - " *厚生*の指標 61(8). 8-13 (2014), 査読有
4. Hasuike, T., Katagiri, H., Tsubaki, H. and Tsuda, H.: "Route planning problem with groups of sightseeing sites classified by tourist's sensitivity under Time-Expanded Network" *Proceedings of 2014 IEEE International Conference on Systems, Man and Cybernetics (SMC2014)*. 188-193 (2014), 査読有
5. Koide, N., Okuhara, K., Ichifuji, Y. and Sonehara, N.: "Analysis of Location and Reservation Trend from

- Cyber Physical Space" ICIC Express Letters 8. 785-791 (2014), 査読有
6. 北川源四郎: "季節調整法から状態空間モデリングへ(特集 時系列解析と季節調整の新しい流れ)" 統計 64-6. 2-7 (2013), 査読無
  7. 畠田誠, 石岡文生, 久保田貴文, 藤田利治: "日本人自殺者数とその増減による空間集積性の評価" 厚生指標 60 (6). 1-8 (2013), 査読有
  8. Kubota, T., Tomita, M., Ishioka, F. and Fujita, T.: "Cluster analysis of age-grouped suicide data in Japan with spatio-temporal structure" COMPSTAT2012 Proceedings in Computational Statistics. 449-460 (2012), 査読有
- 〔学会発表〕(計 202 件)
1. 津田博史, 安藤雅和: "京都・大阪・兵庫のホテルの稼働率の推定とホテルプランの人気要因分析" 第1回観光イノベーション・シンポジウム. (20150306). 京都
  2. 榎広計: "多様なステークホルダーが存在する社会のリスクに基づく意思決定について" 第100回行動計量シンポジウム. (20150125). 同志社大学 室町キャンパス
  3. 榎広計, 北川源四郎: "データ中心政策科学 我が国の取組と情報・システム研究機構プロジェクト" 2014 年度統計関連学会連合大会. (20140914). 東京大学 本郷キャンパス
  4. 津田博史: "データ中心観光政策" 2014 年度統計関連学会連合大会. (20140914). 東京大学 本郷キャンパス
  5. 久保田貴文, 榎広計: "地域ごとの原因・動機別自殺統計に基づく自殺予防総合対策の為に自殺リスクに関する研究" 2014 年度統計関連学会連合大会. (20140914). 東京大学 本郷キャンパス
  6. 小出哲彰, 奥原浩之: "スマートデバイスを用いた環境負荷削減に対する貢献の可視化" 第13回情報科学技術フォーラム. (20140903-05). 筑波大学 つくばキャンパス
  7. Fujita, H., Iijima, W., Koide, N., Satria, D., Santoso, A.D., Prayitno, J., Tsuabaki, H., Kitagawa, G.: "Mobile Application Development for Environmental Informatics & Feedback on Cooking Oil Use and Disposal in Indonesia" 4th International Conference on Technology, Informatics, Management, Engineering and Environment 2014. (20140819). Bandung, Indonesia
  8. Fujita, H., Iijima, W., Nakamo, K., Prayitno, J., Tsubaki, H. and Kitagawa, G.: "A Comparative Study of Waste Cooking Oil Recycling Programs in Bogor and Niigata Cities and GHG Emission Reduction by Recycling" 3rd International Conference on Environmental Science and Development. (20140805). Singapore
  9. 津田博史: "サイバーフィジカル融合社会におけるWebデータの金融分野への応用可能性(会長講演)" 日本金融・証券計量・工学学会 2014 夏季大会. (20140801). 成城大学
  10. 津田博史, 多田舞衣, 山本俊樹, 一藤裕, 曾根原登, 榎広計: "Web データを用いた京都市内のホテルの稼働率の推定" 日本オペレーションズ・リサーチ学会 2014 年春季研究発表会. (20140306). 大阪大学 豊中キャンパス
  11. Nishiyama, Y.: "A NEW APPROACH TO GOODNESS OF FIT FOR ERGODIC MARKOV PROCESSES" Swiss-Kyoto Symposium. (20131121). Zurich, Swiss
  12. Kawasaki, Y. and Ueki, M.: "Multiple choice from competing regression models under multicollinearity based on standardized update" Joint Statistical Meeting 2013. (20130804). Montreal, Canada
  13. 久保田貴文, 榎広計: "自殺に関連するtweetデータの解析" 日本行動計量学会第41回大会. (20130906). 東邦大学 習志野キャンパス
  14. 北川源四郎: "季節調整法から状態空間モデリングへ" 北川源四郎氏大内賞受賞記念シンポジウム「社会科学分野における時系列解析:これまでの発展と今後の展望」. (20130212). 国際文化会館
  15. 佐藤整尚: "状態空間モデルによる金融・経済データ解析" 北川源四郎氏大内賞受賞記念シンポジウム「社会科学分野における時系列解析:これまでの発展と今後の展望」. (20130212). 国際文化会館
  16. 北川源四郎: "ビッグデータ時代の科学的方法論とデータ中心科学リサーチコミュニティ" 応用統計学シンポジウム I 「ビッグデータイニシアティブ:データ科学の高度化と活用の実際」(招待講演). (20130201). 日本大学経済学部
- 〔図書〕(計 9 件)
1. Ohtsu, K., Peng, H., Kitagawa, G.: "Time Series Modeling for Analysis and Control, Advanced Autopilot and Monitoring Systems" Springer. 126 (2015)
  2. 土屋隆裕: "データサイエンス・オンライン講座 社会人のためのデータサイエンス入門 オフィシャルスタディノート" (一財)日本統計協会. 130 (2015)

3. 石黒真木夫, 椿広計, 宮本道子, 柳本武美, 弥永真生, 岡本基: "法廷のための統計リテラシー - 合理的討論の基盤として -" 近代科学社. 205 (2014)
4. 刈屋武昭, 前川功一, 矢島美寛, 福地純一郎, 川崎能典[編著]: "経済時系列分析ハンドブック" 朝倉書店. 772 (2012)
5. Kitagawa, G.: "Introduction to Time Series Modeling" CRC Press. 296 (2010)

## 6. 研究組織

### (1) 研究代表者

北川 源四郎 (KITAGAWA, Genshiro)  
大学共同利用機関法人情報・システム研究機構・新領域融合研究センター・センター長  
研究者番号: 20000218

### (2) 研究分担者

椿 広計 (TSUBAKI, Hiroe)  
統計数理研究所・データ科学研究系・教授  
研究者番号: 30155436

藤田 利治 (FUJITA, Toshiharu)  
統計数理研究所・データ科学研究系・教授  
研究者番号: 30175575  
(平成 22 年度)

津田 博史 (TSUDA, Hiroshi)  
同志社大学・理工学部・教授  
研究者番号: 90450163

西山 慶彦 (NISHIYAMA, Yoshihiko)  
京都大学・経済研究所・教授  
研究者番号: 30283378

川崎 能典 (KAWASAKI, Yoshinori)  
統計数理研究所・モデリング研究系・准教授  
研究者番号: 70249910

佐藤 整尚 (SATO, Seisho)  
東京大学・経済学研究科(研究院)・准教授  
研究者番号: 60280525

土屋 隆裕 (TSUCHIYA, Takahiro)  
統計数理研究所・データ科学研究系・准教授  
研究者番号: 00270413

久保田 貴文 (KUBOTA, Takafumi)  
多摩大学・経営情報学部・准教授  
研究者番号: 30379705  
(平成 23 ~ 26 年度)

藤田 晴啓 (FUJITA, Haruhiro)

新潟国際情報大学・情報文化学部・教授  
研究者番号: 40366513  
(平成 26 年度)

奥原 浩之 (OKUHARA, Koji)  
大坂大学・情報科学研究科・准教授  
研究者番号: 40284616  
(平成 26 年度)

### (3) 連携研究者

村上 政勝 (MURAKAMI, Masakatsu)  
同志社大学・文化情報学部・教授  
研究者番号: 00000216

片桐 英樹 (KATAGIRI, Hideki)  
広島大学・工学研究院・准教授  
研究者番号: 40325147

宮本 道子 (MIYAMOTO, Michiko)  
秋田県立大学・システム科学技術部・教授  
研究者番号: 30469598

曾根原 登 (SONEHARA, Noboru)  
国立情報学研究所・情報社会関連研究系・教授  
研究者番号: 30390595

富田 誠 (TOMITA, Makoto)  
東京医科歯科大学・医学部附属病院・准教授  
研究者番号: 20399025

笛田 薫 (FUEDA, Kaoru)  
岡山大学・環境生命科学研究科・准教授  
研究者番号: 50253399

蓮池 隆 (HASUIKE, Takashi)  
大阪大学・情報科学研究科・助教  
研究者番号: 50557949

宮原 孝夫 (MIYAHARA, Takao)  
名古屋市立大学・経済学研究科・名誉教授  
研究者番号: 20106256

安藤 雅和 (ANDO, Masakazu)  
千葉工業大学・社会システム科学部・准教授  
研究者番号: 00462169