

## 科学研究費助成事業（科学研究費補助金）研究成果報告書

平成25年 3月31日現在

機関番号：62603

研究種目：基盤研究（C）

研究期間：2008～2012

課題番号：20500266

研究課題名（和文） 環境データ解析のためのベイズ的方法の開発とその応用

研究課題名（英文） Development of Bayesian methods for analyzing environmental data and their application

研究代表者

柏木 宣久（KASHIWAGI NOBUHISA）

統計数理研究所・モデリング研究系・教授

研究者番号：50150032

研究成果の概要（和文）：環境問題における統計的問題、特に発生源解析、環境化学物質測定の原因分析、生態リスク評価、降雨リモートセンシング、東京湾水質の長期変動等の問題に現れる識別問題を解決するためのベイズ的方法を開発した。同時に、開発した方法を実際の環境問題に適用し、方法の実用性を検証すると共に、環境問題の解決に尽力した。

研究成果の概要（英文）：I have developed Bayesian methods for solving statistical problems in an environmental issue, especially identification problems appearing in source receptor modeling, factorial experiment of the measurement of environmental chemicals, assessment of ecological risk, radar remote sensing of rainfall, and trend analysis of water quality in Tokyo Bay. The methods developed in this research have been applied to actual environmental problems to verify practicality of the methods and to contribute to solve those environmental problems.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	1,000,000	300,000	1,300,000
2009年度	600,000	180,000	780,000
2010年度	600,000	180,000	780,000
2011年度	600,000	180,000	780,000
2012年度	700,000	210,000	910,000
総計	3,500,000	1,050,000	4,550,000

研究分野：統計数理

科研費の分科・細目：情報学・統計科学

キーワード：環境統計、識別問題、ベイズ、発生源解析、要因分析、生態リスク評価、降雨リモートセンシング、東京湾水質

### 1. 研究開始当初の背景

近年のベイズ的方法の進展は著しく、進展以前の観念論に過ぎないとの批判を覆し、実用的方法として広く一般に受け入れられるようになってきた。特に従来は解決困難とされてきた識別問題、非線形問題、大規模問題等の解決に威力を発揮している。例えば、申請者の提案だけでも、識別問題については、井戸水位の時空間変化に基づく地下帯水構

造の逆推定（Honjo and Kashiwagi, 1999）や環境汚染に対する未知発生源寄与率の推定（Kashiwagi, 2004）、非線形問題については、ポアソン系列の平滑化（Kashiwagi and Yanagimoto, 1994）や線形トレンドにおける多数の構造的変化の同時検知（Kashiwagi, 1996）、大規模問題については、平面平滑化（Kashiwagi, 1993）や時空間季節変動調整（Kashiwagi, et al., 2003）等がある。これ

ら方法の開発の契機になったのが状態空間、マルコフ連鎖モンテカルロといった計算法の普及に伴う計算可能性の拡大である。これにより現実には限られたモデリングが可能になった。そしてまた、問題解決におけるモデリングの重要性が一層認識されるようになった。

モデリングを行うには、述べるまでもなく、現象に対する理解が不可欠である。そのため、申請者の場合、方法の開発に際し該当分野研究者との共同研究を実施している。その一環として、環境分野研究者との交流があり、統計の立場から環境問題全般を観察する機会を得た。その際に気付いたのが、環境問題では識別不能な推論を強いられる場面が極めて多いという点である。本来的に識別不能な場合もあるが、目立つのはデータの未整備に起因する識別不能である。環境問題は注目を集めてから未だ日が浅く、次々と新たな問題が浮上してくるため、データの整備に手がまわらないのが実状である。それでも、問題解決に向け迅速に決断を下さなければならない。こうした状況に対処するには、識別不能な推論を実現する方法を開発するのが最善と考えられた。

## 2. 研究の目的

本研究では、環境問題に現れる識別問題を解決するためのベイズ的方法を開発する。同時に、開発した方法を実際の環境問題に適用し、方法の実用性を検証すると共に、環境問題解決に貢献する。以下では、研究課題毎に説明する。

### (1) 発生源解析

環境管理にとって汚染発生源の同定は本質的な問題である。発生源を同定できなければ如何なる発生源対策も立てようがないからである。発生源を同定するためレセプターモデルと呼ばれる様々な数理的方法が提案されている。典型的には米国環境保護局の標準的方法として採用された Positive Matrix Factorization (Paatero and Tapper, 1993) を挙げられる。しかし、この方法の解に一意性は無いとの指摘があり (Park et al., 2002)、標準的方法が存在しない状況になった。そうした状況を打開するため、観測データが存在しない発生源の組成と寄与率さえ推定できるベイズ的方法を開発し (Kashiwagi, 2004)、指摘された識別問題を解決する糸口を与えた。ただし、開発した方法はある程度の観測の繰り返しを必要とする。環境化学物質の測定には多大なコストが掛かるため、必要とする観測の繰り返しは少ない方が好ましい。本研究ではそうした要求に応えるための方法を開発する。

### (2) 環境化学物質測定における要因分析

実験や測定の要因分析では実験計画法を

用いるのが常道である。ところが、環境化学物質の測定の場合、要因数が極めて多く、測定に多大なコストが掛かるため、実験計画法を限定的にしか利用できないでいる。そのため、偶発的に設定された測定条件とその条件下での測定結果が比較的大量に集まる調査の結果を利用し、要因効果を推定しようと試みられている (環境省, 2006)。しかし、要因効果の分離推定が原理的に困難な不完全データしか集まらないため、試みはことごとく失敗に終わっている。そこで、本研究では、要因の絞り込みを目的に、漫然と集められたデータからでも要因効果の有無を評価できるベイズ的方法を開発する。

## 3. 研究の方法

当初予定していた発生源解析と環境化学物質測定における要因分析の他に、当該分野の研究者から研究協力依頼があった生態リスク評価、降雨リモートセンシング、および東京湾水質の長期変動を研究課題に含めた。研究課題毎に説明する。

### (1) 発生源解析

情報が存在しない発生源の組成と寄与率を推定できるようにしたが、発生源情報は存在した方が推定精度は高くなる。東京都環境科学研究所からダイオキシン類による土壤汚染に関する相談があり、検討した結果、塩素製造に用いられた食塩電解過程の影響が考えられた。ただし、食塩電解過程から生成されるダイオキシン類に関するデータは存在していなかった。そこで、データの採取を実施した。同じく東京都環境科学研究所から PCB 製品による土壤汚染と大気汚染に関する相談があり、データを解析した結果、何れの場合も未確認発生源の寄与率が高く、原因究明が必要になった。検討した結果、発生源である製品原体から環境に到達する間に組成変化が生じている可能性が考えられた。そこで、PCB の環境動態について検討した。

発生源情報を拡充する一方で、モデル改良に備え、情報収集に努めた。宮城県保健環境センターが採取した多種環境試料中の PCB、新潟県保健環境科学研究所が採取した佐渡真野湾底質コアサンプル中のダイオキシン類、神戸市に依頼された神戸港ダイオキシン類汚染、および富山市に依頼された運河底質ダイオキシン類汚染等に関するデータの解析を実施した。

### (2) 環境化学物質測定における要因分析

環境省は、環境測定分析の信頼性確保と精度向上を目的に、環境測定分析統一精度管理調査を実施している。この調査では、均一に調製した環境試料を全国の分析機関に配布し、指定した物質の測定値と分析条件を回収している。今回、この調査により回収された底質中フタル酸ジエチルヘキシルのデータ

を解析の対象にした。データの大きさは86であるが、分析条件は46あり、水準数3以上の条件も含んでいた。そのため、通常の方法では、要因効果は識別困難と言える。そこで、要因効果は分散に現れる傾向があるという予備解析の結果に基づき、データを標準データと非標準データに分類するモデルを考えた。そして、各分析条件の水準を標準条件と非標準条件に整理し、非標準条件の分類に対する効果を事後確率で評価するベイズ的方法を開発した。また、分散を分析条件でモデル化したベイズ的方法も開発し、結果の解釈を補強できるようにした。

### (3) 生態リスク評価

環境中の化学物質を管理するためにリスク評価は不可欠である。生態リスク評価では、ハザード比のような恣意的な指標ではなく、環境濃度や種の感受性の確率変動を明示的に扱える Expected Potentially Affected Fraction (EPAF)を用いるのが合理的である。ただし、環境濃度について、検出下限未満の測定値が多数を占め、データの大きさが時点や地点に依存する、といった問題があり、種の感受性について、毒性値が測定されている種数が藻類、甲殻類、魚類といった分類毎に異なり、毒性データの大きさも極端に小さい、といった問題があり、環境濃度分布や種の感受性分布を合理的に推定できず、EPAFを利用し難い状況が続いていた。状況を改善するため、ベイズ的方法を開発した。

### (4) 降雨リモートセンシング

衛星搭載降雨レーダによる降雨観測について、雨滴粒径分布を同定し降雨量を推定するために必要な統計的方法について検討した。

### (5) 東京湾水質の長期変動

東京湾沿岸の各自治体は、湾内の水質状況を把握するため、水質汚濁防止法に基づく公共用水域水質測定計画に沿って、毎月1回、測定点を分担し合い、多項目の水質測定を実施している。これらの測定で得られたデータを収集し、解析に必要な統計的方法を開発し、開発した方法を用いて東京湾水質の長期変動傾向を推定した。

## 4. 研究成果

### (1) 発生源解析

食塩電解過程については、黒鉛電極を用いると、高濃度のダイオキシンが生成されることが分かった。生成されるダイオキシン類はほとんどがPCDFsであり、反応時間が進むに従い高塩素体の比率が増え、黒鉛電極のバインダーであるピッチの量や存在状態が生成量に大きく影響し、現在使用されているチタン電極では生成されないが、チタン電極にピッチを共存させるとPCDFsが生成されるのも確認した。食塩電解では、ジベンゾフランに高

濃度の塩素が反応して塩素の置換反応が進み、優先的に生成される異性体が存在し、クロマトグラムは特異なパターンを示す。反応時間の経過に伴うダイオキシン類の同族異性体組成の関係を検討した結果、塩素反応パターンに関する知見は発生源解析に利用可能であると判断された。

東京都の土壌汚染については、PCB製品の漏出などにより汚染された土壌を想定したカラム試験を行った。その結果、PCB製品が漏出すると、雨水等の影響により周辺非汚染土壌への二次汚染が起り、二次汚染土では、実際に漏出したPCB製品より低塩素に偏った組成を示すのが確認された。

東京都の大気汚染については、室内にてPCB製品等からのPCB揮散実験を行った。その結果、PCB製品が大気環境へ揮散すると、二次汚染土の場合と同様、PCB製品より低塩素に偏った組成を示すが確認された。また、揮散温度を変化させても同族体組成に大きな変動は認められなかった。一般環境では、気温が高くなると、揮散量は多くなるが、組成変動は極めて小さいと考えられた。

宮城県のデータについては、多重共線性にも対応可能なベイズ型重回帰モデルと未確認発生源にも対応可能なベイズ型半因子モデルを適用し、方法の有用性と問題点について検討した。その結果、従来の方法と比較し、格段に性能が向上しているのを確認できた。問題点としては、想定通り、推定でバイアスが生じるため、ややなまった数値になりやすい点と、観測の繰り返しを確保するのにコストが掛かる点であった。

新潟県のデータについては、測定誤差モデルを適用し、ベイズ的方法の比較対象にした。推定では、サンプルが小さいためもあり、ノイズの影響を強く受けた。やはり、事前情報を付加する必要があるようである。それでも、解析結果から、底質汚染発生源の経年変化を推測出来た。また、コアサンプルは時系列データとしても扱えるのが分かり、モデル開発のヒントになった。

神戸市と富山市の問題については、早急に結論を導く必要があったため、既開発のベイズ型半因子モデルを適用し、解析結果を提供した。ただ、汚染が大規模であったため、何れの場合も、詳細な空間データが採取されており、観測の繰り返しを空間データで代替する方法を開発する際の動機付けになると考えられた。

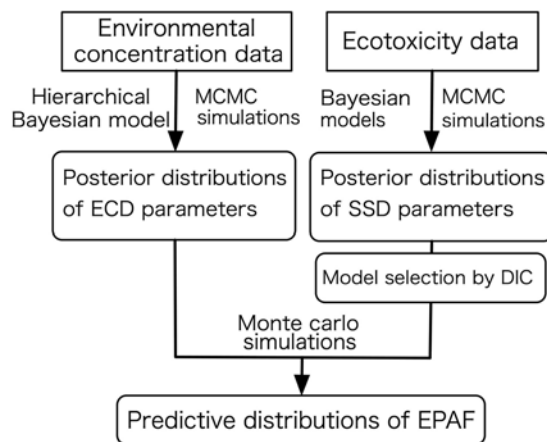
### (2) 環境化学物質測定における要因分析

計算の結果、分類に関しては、効果がある要因数の事後確率と、各要因に効果が有るとする事後確率が求まった。要因数の事後確率のモードは8で、概ねその数に見合う要因が効果が有りとする事後確率により選び出された。興味深いのは、事後確率による結果は、

要因毎に実施した古典的なF検定の結果と異なっていた点である。要因配置は直交からほど遠い状況のため、当然の結果ではある。分散に対する要因効果の推定結果を併せて解釈すると、a)「サロゲート無」は精度悪化の有意な要因、b)「検出下限1」も有意だが、精度悪化の直接的原因ではなく、分析全般の信頼性に関係しており、c)、「抽出回数3」と「内分比率負」も有意と言えるが、例数が少ないため、断定はできず、d)「検量線点数3」と「内分比率1以上」も有力な候補であるが、推定値が安定しないため、有意とまでは言えず、e)「MS装置形式4」は一部の計算で候補として検出されたが、「検出下限1」を含むデータの影響と考えられる、といった結論が導かれた。従来は、こうした解釈さえ困難であった。

### (3) 生態リスク評価

環境濃度に関しては、平均は地点に依存するとし、その平均に無情報的事前分布を仮定した。毒性値に関しては、種により平均と分散が同じ、平均または分散が異なる、および平均と分散が異なるという4つのモデルを想定し、それらの母数に無情報的事前分布を仮定し、最善のモデルを選択するようにした。そして、以下の手続きを用い、EPAFを計算するよう提案した。



提案の方法を東京都河川のモニタリングデータに適用し、主要9物質の生態リスクの定量化及び比較を行った。結果、EPAFの中央推定値で評価した場合、東京都河川においてリスクが最も大きいのはニッケルであり、次いで亜鉛とアンモニアであると示唆された。

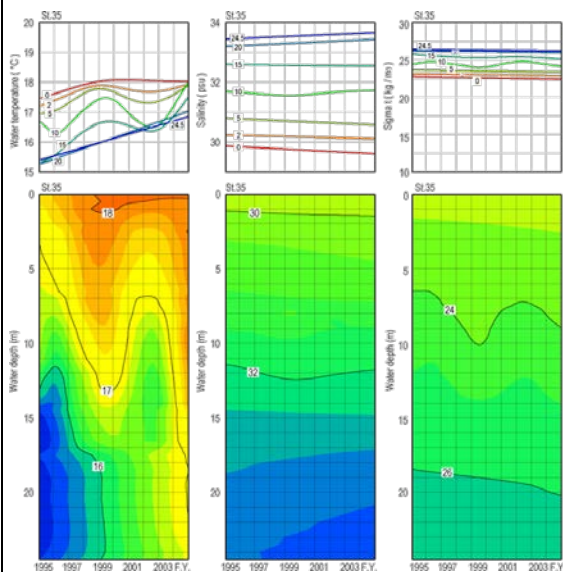
### (4) 降雨リモートセンシング

降雨パラメータ間の統計的關係について知見を得た。また、降雨パラメータの長周期変動と短周期変動を分離推定するための1次元2スケールベイズモデルを開発した。更に、衛星搭載降雨レーダによる降雨観測に関連して雨滴粒径分布の母数の同定について検討し、加えてレーダ反射因子と降雨強度の關係についても検討した。

### (5) 東京湾水質の長期変動

データを収集し、スクリーニングを実施した後、データベースを構築した。また、水質の長期変動を捉えるため、各種季節変動モデルを開発した。多くの地方環境研究所から地球温暖化に伴う沿岸域の水質変化を捉えるため使用したとの要望があったため、ベイズ的方法だけでなく、取り扱いが簡単な重回帰型季節変動モデルも開発した。また、特定の観測点で毎月1回水深50cm毎に水質を測定したデータがあり、鉛直方向に滑らかに変化し、時間方向にトレンドを持ちつつ季節変化していたため、対応するベイズ的方法を開発した。

公共用水域水質測定計画のデータを解析した結果、富栄養化に関しては、対策により過去においてCOD, N, P等は改善したが、最近では横ばい状態であり、肝心の貧酸素水塊は拡大傾向にあるため、新たな対策が必要と考えられた。また、鉛直データを解析した結果、下図のような結果が得られた。左から水温、塩分濃度、密度を表し、縦軸は水深で、横軸は時間である。これらの図から、表層で塩分濃度が低下し、底層で塩分濃度と水温が上昇しており、これらの結果から、表層への淡水流入量が増加し、誘発されて底層への外洋水流入量も増加し、エスチャリー循環が強化されているのが明らかになった。



### 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18件)

①Hayashi, T. and Kashiwagi, N., A Bayesian approach to probabilistic ecological risk assessment: risk comparison of nine toxic substances in Tokyo surface waters, Environmental Science and Pollution

Research, 査読有, 18, 2011, 365-375.

②永山恵, 岩淵美香, 小林弘明, 牧秀明, 二宮勝幸, 安藤晴夫, 柏木宣久, 川崎港における海水温の連続調査結果 2, 川崎市公害研究所年報, 査読無, No. 38, 2011, 49-52.

③今井美江, 山崎正夫, 東野和雄, 柏木宣久, 高橋昭宏, PCB の揮発に伴う組成変化について, 東京都環境科学研究所年報, 査読無, 2011, 9-15.

④二宮勝幸, 柏木宣久, 岡敬一, 東京湾西岸域における海水温の上昇と季節のずれ, 横浜市環境科学研究所報, 査読無, No. 35, 2011, 34-40.

⑤古津年章, 柏木宣久, 下舞豊志, 状態空間モデルによるレーダ反射因子対降雨強度 (Z-R) 関係の推定, 電子情報通信学会論文誌 B, 査読有, J93-B, 2010, 1660-1665.

⑥Kozu, T., Masuzawa, K., Shimomai, T. and Kashiwagi, N., Estimation of NO for 2-scale gamma raindrop size distribution model and their statistical properties at several locations in Asia, Journal of Applied Meteorology and Climatology, 査読有, 49, 2010, 676-686.

⑦Hayashi, T. and Kashiwagi, N., A Bayesian method for deriving species-sensitivity distributions: Selecting the best-fit tolerance distributions of taxonomic groups, International Journal of Human and Ecological Risk Assessment, 査読有, 16, 2010, 251-263.

⑧東野和雄, 山本央, 大浦健, 雨谷敬史, 佐々木裕子, 橋本俊次, 柏木宣久, 嶽盛公昭, 高菅卓三, 高橋明宏, 食塩電解過程で生成するダイオキシン類の異性体/同族体組成, 東京都環境科学研究所年報, 査読無, 2010, 1-7.

⑨Kozu, T., Iguchi, T., Shimomai, T. and Kashiwagi, N., Raindrop size distribution modeling from a statistical rain parameter relation and its application to the TRMM precipitation radar rain retrieval algorithm, Journal of Applied Meteorology and Climatology, 査読有, 48, 2009, 716-724.

⑩林岳彦, 柏木宣久, 時間的・空間的な変動性 V および推定の不確実性 U を解析する: 階層ベイズモデルによる不検出値を含む環境中濃度データの解析, 日本リスク研究学会誌, 査読有, 19, 2009, 47-54.

[学会発表] (計 28 件)

①柏木宣久, 清家伸康, 元木裕, 稲生圭哉, 大谷卓, ウリ科野菜における POPs 汚染度診断法の開発Ⅲ: 圃場における POPs の土壌中濃度分布について, 環境化学討論会, 2012. 7. 13, 松山.

②柏木宣久, 安藤晴夫, 東京湾水質の鉛直分布データの特徴, 統計関連学会連合大会, 2011. 9. 6, 福岡.

③ Hayashi, T. and Kashiwagi, N., A Bayesian approach to probabilistic ecological risk assessment using SSDs: risk comparison of nine toxic substances in Tokyo surface waters, SETAC Europe 21st Annual Meeting, 2011. 5. 18, Milan.

④林岳彦, 柏木宣久, ベイズ統計を用いた化学物質の定量的生態リスク比較, 統計関連学会連合大会, 2010. 9. 7, 東京.

⑤柏木宣久, 環境モニタリング結果の統計解析とその活用について, 化学物質環境実態調査環境科学セミナー, 2010. 1. 20, 東京.

⑥古津年章, 柏木宣久, 状態空間法による Z-R 関係の自動推定, 電子情報通信学会ソフトウェア大会, 2009. 11. 27, 名古屋.

## 6. 研究組織

### (1) 研究代表者

柏木 宣久 (KASHIWAGI NOBUHISA)

統計数理研究所・モデリング研究系・教授

研究者番号: 50150032