

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 14 日現在

機関番号：32682

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25870803

研究課題名(和文) シミュレーションと観測を融合する新しいバッチ型アルゴリズムによる知識発見原理

研究課題名(英文) Principles for knowledge discovery through new batch type algorithm of combination of simulation and observation

研究代表者

中村 和幸 (Nakamura, Kazuyuki)

明治大学・総合数理学部・専任准教授

研究者番号：40462171

交付決定額(研究期間全体)：(直接経費) 1,300,000円

研究成果の概要(和文)：数値シミュレーションと観測データを融合する手法であるデータ同化について、新しいバッチ型のデータ同化アルゴリズムの構築とその有効性の検証を行った。アルゴリズムについては、繰り返し計算を用いた計算手法が有効である可能性を見出した。さらにその有効性について、いくつかの試験用のモデルを通じた数値計算による検証を行った。検証により、有効性が確認されるとともに、数値シミュレーションに含まれる誤差を表現する確率分布を変えることで、データ同化のためのより適切な数理的枠組構築ができることを確認した。

研究成果の概要(英文)：Data assimilation is the concept and algorithm that combine numerical simulation model with observation data. In this research, new batch type algorithm of data assimilation is constructed. The effectiveness of the algorithm is also verified. As a result, possibility of the effectiveness of the repeated calculation algorithm has been found. Numerical experiments to test the effectiveness of the algorithm through several test model is also conducted. From the results, effectiveness of the approach is confirmed. It is also confirmed that we can construct appropriate formulation for data assimilation by changing probability distributions which represent the errors of the numerical simulations.

研究分野：統計科学

キーワード：データ同化 ベイズ推定

1. 研究開始当初の背景

多くの理学・工学の分野において、自然現象の記述のために、数式に基づいた数理モデルが構築され広く用いられている。数理モデルは、対象としている自然現象についての知見や、実験・観察の結果をもとに、妥当と考えられる仮定を置いた上で構築されている。数理モデルは、数理的な解析や計算機シミュレーションにより妥当性検証がなされ、現象の解明や予測に用いられる。数理モデルとそれに基づくシミュレーションによるアプローチは、モデルが妥当な場合には、再現性が高く知識発見や予測に有用なため、自然科学分野や工学分野で広く用いられている。

一方、実現象を表現するのに数理モデルのみでは困難な自然現象も存在する。実現象を正確に記述するためのデータが不足している大規模現象や直接観測が困難な現象が該当する。これらの現象では、現象を記述する数理モデルに含まれるパラメータや初期・境界条件に大きな不確かさを持つことや、数理モデルで表現しきれなかったことに由来する「モデル化誤差」がある。このモデル化誤差は、数理モデルによる演繹的推論から、適切な現象再現や予測を行うのは困難である。さらに、非線形現象においては、初期値の不確かさが系の持つ非線形性により拡大してしまい、予測がさらに困難になるという問題もある。

このような困難を克服するための方法として、データ同化と呼ばれる手法が、気象・海洋学の分野では開発されて発展してきた。工学分野において発展してきた、逆解析と呼ばれる手法も同様の手法と考えられる。データ同化は、学術研究と産業応用で適用が広がっており、気象海洋・地盤解析分野などにおいて、応用上の有効性が確認されている。

データ同化は、大きく分けて逐次型とバッチ型の2種類が存在する。そのいずれにおいても、数理モデルで現象を記述した数値シミュレーションに対して、観測データとの非整合を統計モデルで表現した状態空間モデルによる推論を行うという点で一致している。その一方、従来のバッチ型では同時事後確率最大化を求めているが、逐次型では各時点のフィルタ分布（各時点の周辺分布）の推定を求めている。

バッチ型データ同化手法で広く用いられているのは4次元変分法である。4次元変分法では、数理モデル部分には誤差が入らないという前提で定式化され、同時事後確率最大化をおこなっている。そのため、数理モデル構築時のモデル化誤差を導入することが本質的にできず、正確な誤差解析が不可能である。一方、逐次型データ同化の場合、モデル化誤差は自然に導入・解析が可能であるが、多次元空間中でのモンテカルロ近似の非効率性や、本来空間的になめらかであるべき変数に非現実的な擾乱を加えてしまうという問題が生じる。

以上の各手法の問題点と、それを踏まえた使い分けの指針について、地盤工学の問題に制約して検討し、一定の結論を得ていた。その検討過程で、モデル化誤差を考慮にいれ事後周辺確率分布を連続分布で近似可能なバッチ型データ同化手法が得ることができれば、従来のアルゴリズムが持つ問題点を全て解決できるという可能性を見出していた。

2. 研究の目的

(1) モデル化誤差を考慮にいれたバッチ型データ同化手法の構築と妥当性検証を通じて、「バッチ型データ同化における事後周辺確率分布の近似推定」という新しい視点の導入とその有効性を明らかにすること、さらにその帰結として、導入した視点と手法が、実問題に応用可能な新知見発見やリスク検知のための計算原理となり得ることを明らかにすることを目的とした。

(2) 前項を実現するために、既存のバッチ型データ同化手法を参考にした新たなデータ同化手法の構築可能性、特に、モデル化誤差を考慮したバッチ型データ同化アルゴリズム構築可能性の検討を具体的な目的の一つとし、さらにその適用性を試験モデルならびに実問題に適用することも目的とした。

3. 研究の方法

(1) アルゴリズム構築に関しては、モデル化誤差を考慮したバッチ型データ同化アルゴリズムの構築を行うこととした。特に、ベイズ推論アルゴリズムの一種と考えることができる、弱拘束4次元変分法ならびにIntegrated Nested Laplace Approximation (INLA) をアルゴリズム構築の起点とし、シミュレーションモデルに対するアルゴリズムの拡張による構築を目指すこととした。

(2) 構築アルゴリズムの適用性と有効性の検証については、試験用数理モデルへの適用による妥当性と有効適用範囲の検討、ならびに実際のモデルへの適用による検証を行うこととした。特に、これまでに他の研究における検証に利用したLorenz 63モデルやLorenz 96モデルによる試験用データ同化ルーチンや、実問題プログラムを援用し、妥当性検証を行うこととした。

(3) 上記に加えて、システムノイズの分布形による結果の変化も検証することとした。これにより、システム誤差構造のモデリングが同化結果に与える影響を評価することが可能となるためである。

4. 研究成果

(1) アルゴリズム構築については、当初計画していたベイズ推論アルゴリズムについては、実際の問題への適用を想定すると、周辺分布を計算する上での非線形性に起因する

困難性の問題。特に、観測モデルにおいて非線形性がある場合の取り扱いに困難性があることが分かった。その一方、手法の整理を行っている中で、従来開発していた粒子フィルタ・平滑化を使用しながら周辺分布を繰り返し再計算する手法の可能性を見出したため、その適用性について検討した。その結果、本研究において目的としている、モデル化誤差を考慮にいれた事後周辺確率分布を求めるアルゴリズムと同様の文脈で整理できることが確認できた。さらに、適切なシステムノイズの分布を導入した上で、平滑化分布の構成も有効である可能性を見出した。そこで、試験用モデルにおいてシステムノイズの分布形も変えた計算結果の検討を行うこととした。

(2) 試験用のモデルとしては、最終的には、レベルシフト型モデル、Lorenz 63 モデルならびに Lorenz 96 モデルの3種類を採用するに至り、これらにおいて、手法の妥当性と有効性の検証を行った。

レベルシフト型モデルについては、誤差分布の分布形をコーシー分布とし、その誤差レベルを変化させることで、誤差構造にどのような影響を与えるかの検討を行った。その際には、バッチ型アルゴリズムに加え、逐次型アルゴリズムによる推定結果も加えることとした。数値実験による検討の結果、コーシー分布におけるパラメータ変化によって、誤差構造を連続的に変化可能であること、特に、パラメータのレベルによって、正規分布を用いた場合の構造に近い構造となることが確認された。本結果、特に誤差構造を変化させたことによる推定結果への逐次型アルゴリズムにおける影響については、応用数理に計算結果を掲載した。

Lorenz 63 モデルと Lorenz 96 モデルについては以下の通りである。まず Lorenz 63 モデルにおいては、再計算型のアルゴリズムの有効性を確認するとともに、複数の状態に同時に誤差を入れた時の影響は小さいことを確認した。Lorenz 96 モデルについては、空間構造があることに由来して、モデル化誤差の空間構造が推定誤差や予測誤差に影響を与える傾向があることを確認した。このことから、モデル化誤差の空間構造をうまく定式化することで、データ同化アルゴリズムを通じたリスク検知に対する可能性を見出した。

(3) セミバッチ型の粒子平滑化アルゴリズムに対して、有効粒子数が精度に大きな影響を与えることが確認された。このことを踏まえて、本研究からの派生として、簡易多粒子系データ同化プログラムの構築において有効粒子数を適切に考慮する手続きの追加を行うこととした。実際に他の研究によって得られたモデリングの成果と合わせて、簡易多粒子系のマクロパラメータ推定を行った結

果、推定が適切に行えることが確認できた(雑誌論文)。

5. 主な発表論文等

[雑誌論文](計4件)

K. Nakamura and Y. Kono, "Fast and stable estimation of macroscopic parameters in particle systems by data assimilation," Proceedings of the 47th ISCI International Symposium on Stochastic Systems Theory and Its Applications, 査読有, pp. 132-136, 2016.

A. Kimura, A. Celani, H. Nagao, T. Stasevich and K. Nakamura, "Estimating cellular parameters through optimization procedures: elementary principles and applications," Frontiers in Physiology, 査読有, 03 March 2015 (online journal), doi: 10.3389/fphys.2015.00060, 2015.

中村 和幸, 「統計的時系列解析4: グラフィカルモデルとデータ同化」, 応用数理, 査読無, 24(4), pp. 29-35, 2014.

中村 和幸, 「統計的時系列解析3: 非線形・非ガウス状態空間モデル」, 応用数理, 査読無, 24(3), pp. 33-38, 2014.

[学会発表](計8件)

中村 和幸, 静的・動的システムにおける不確かさの定量化, 第22回信頼性設計技術WS&第35回最適設計研究会, 2015年9月14日, 岡山大学(岡山).

中村 和幸, 粒子法流体解析における誤差統計モデルについて, 2015年度統計関連学会連合大会, 2015年9月9日, 岡山大学(岡山).

中村 和幸, 清水 脩輔, 非圧縮流体の粒子法解析におけるデータ同化と誤差モデリング, 第63回理論応用力学講演会, 2014年9月28日, 東京工業大学(東京).

中村 和幸, 統計数理から見たデータ同化の理論・アルゴリズムとその応用, 気象学におけるビッグデータ同化の数理ワークショップ, 2014年3月19日, 京都大学(京都).

K. Nakamura, Data Assimilation and Bayesian Statistical Analysis: Tools for Understanding Uncertain Natural and Social Phenomena, 2013 Northeast

Asian Symposium, 2013年9月23日, 成都(中国).

中村和幸, データ同化-データを活かしたシミュレーションと予測の方法, 土木学会, 2013年9月4日-6日, 日本大学生産工学部(千葉).

K. Nakamura, Y. Ohya and T. Shiga, Time series analysis and data assimilation for natural and economic hazards, Joint Meeting of the IASC Satellite Conference for the 59th ISI WSC and the 8th Conference of the Asian Regional Section of the IASC, 2013年8月22日-23日, ソウル(韓国).

K. Nakamura, Review on data assimilation methodology, Seventeenth International Conference on Intelligent System Applications to Power Systems, 2013年7月1日-4日, 明治大学(東京).

〔図書〕(計2件)

三村 昌泰(編著), 杉原 厚吉, 青木 健一, 中村 和幸, 高安 秀樹, 砂田 利一, 萩原 一郎, 現象数学の冒険, 全220ページ, 第4章「地球科学の数理」(pp. 95-123)分担執筆, 明治大学出版会, 2015.

鈴木 和幸, 横川 慎二, 長塚 豪己, 山本 涉, 嶋崎 真仁, 河村 敏彦, 貝瀬 徹, 中村 和幸, 横山 真弘 他, 新版 信頼性ハンドブック(日本信頼性学会編), 全915ページ, 「2.7.1 データ同化とは」(pp. 458-462)分担執筆, 日科技連, 2014.

6. 研究組織

(1) 研究代表者

中村 和幸 (NAKAMURA, Kazuyuki)
明治大学・総合数理学部・専任准教授
研究者番号: 40462171