

マルチエージェント社会シミュレーションにおける データ同化の可能性評価の枠組み

川口 英俊[†] 野田 五十樹[‡]
東京工業大学^{†‡} 産業技術総合研究所^{‡†}

1 はじめに

本稿では、マルチエージェント社会シミュレーション (Multi-agent Social Simulation 以下、MASS) のデータ同化の可能性評価の枠組みを提案し、実験によりその効果を検証する。データ同化とは、観測データへシミュレーションの出力値を合わせ未知のパラメータ (初期条件など) を推定する手法およびその作業である。データ同化は、予測や分析などのシミュレーションの現実問題への応用を考えた時、重要な役割を担う。

ビッグデータという言葉の流行にもある通り、観測データを得やすくなった。そのため、莫大な観測データを使ったデータ同化への応用も期待できる。これまで、各々独自の方法での観測データを元にした MASS の再現性を高めようという研究は存在するものの [1, 2]、MASS へのデータ同化の理論的研究は少ない。

しかし、このままビッグデータを MASS へのデータ同化へ応用できるかは疑問である。そもそもデータ同化は洗練化されたモデルを前提とし、その上で尚未知なパラメータを推定する手法である。MASS のモデルは発展途上であり、未解明な部分が多くその前提を満たしているとは言い難い。そのため、予めデータ同化の可能性を評価することは重要な要素である。

2 提案手法

提案手法の構成を図 1 に示す。現実の対象事象をデータ同化する MASS に置き換えて、仮想的にデータ同化を行いその可能性を確認する。提案手法は以下の 2 フェーズに分けられる。

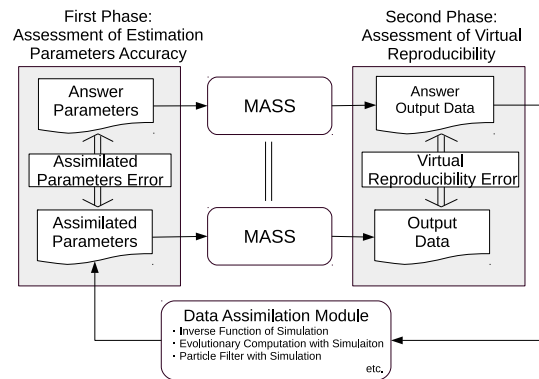


図 1: 提案手法

第 1 フェーズ:パラメータ推定精度評価

入力パラメータの推定精度を確認し、データ同化条件を評価する。物理分野のデータ同化における双子実験に相当する。解答データとなる入力パラメータで実行し、その出力データを擬似観測データとする。その観測データを元にデータ同化を行いパラメータを推定し、解答データと比較することによってそのデータ同化の条件の評価を行う。

第 2 フェーズ:仮想再現性評価

パラメータ推定精度評価で算出した入力パラメータにより再度 MASS を実行し、その入力と出力の対応関係をとる。そしてその出力データと擬似観測データを比較し誤差を算出する。MASS のモデルはその複雑さから、推定した入力パラメータが観測データを再現出来ているか不明な場合が多い。そのため、再度 MASS を実行し、入力パラメータがどの程度解答データの出力側を再現できているかを確認する。

3 実験

人流シミュレーションのデータ同化を行うという想定で提案手法の適用実験を行う。本実験でのデータ同

Possibility Assessment Framework of Data Assimilation for Multi-agent Social Simulation
[†] Hidetoshi Kawaguchi, TITECH, AIST
<kawaguchi.h.ac@m.titech.ac.jp>
[‡] Itsuki Noda, AIST, TITECH <i.noda@aist.go.jp>

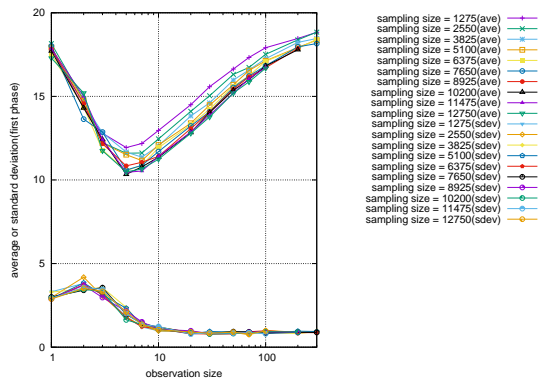


図 2: パラメータ推定精度評価の結果

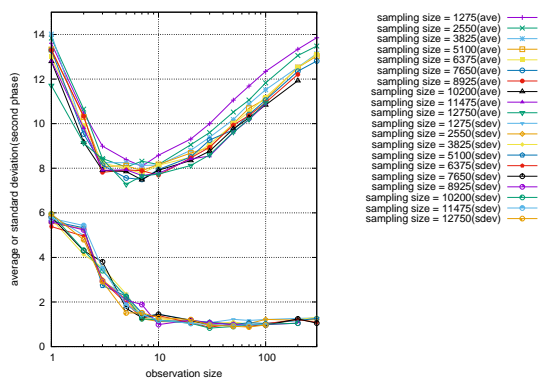


図 3: 仮想再現性評価の結果

化は、評価関数が複数ある場合を想定し、その解の不確か性をゲーム理論におけるパレート解として扱う。

観測サイズ (=データ同化の評価関数の数) がデータ同化へ与える影響について確認していく。観測サイズ毎に 100 通りの評価関数をランダムに選択し、それぞれの入出力の精度の平均、標準偏差を確認する。

シミュレーション設定 本実験の対象シミュレーションは、2 入力 600 出力のエージェントベースの人流シミュレーションとする。入力エージェント数とエージェントの経路選択の際の気まぐれ度の 2 つであり、50 段階と 51 段階設定できる。出力は地図上の特定の 2 箇所それぞれのタイムステップ毎の存在するエージェント数である。1 箇所につき 300 タイムステップ分出力する。実装には山下 [3] の CrowdWalk を使用した。

データ同化の方法 パレート解を許容したランダムサーチをデータ同化手法として用いる。本実験では、入力の全組み合わせ × 乱数シード 5 つで実行した入出力セットをあらかじめ用意しておき、

その中でランダムサーチを行う。サンプリング数は 1275 の倍数 10 段階で行う。解答データの 1 つの出力値とそれに対応するシミュレーション出力値のユークリッド距離を 1 つの評価関数とする。

3.1 実験結果

第 1 フェーズでは観測サイズ毎に推定されたすべての入力パラメータの推定誤差平均値と標準偏差を算出し、図 2 に結果を示す。推定誤差は入力 2 つのユークリッド距離とした。縦軸は誤差の平均と標準偏差であり、横軸は観測地点の数を対数軸で表している。

第 1 フェーズで推定されたすべての入力パラメータでの出力結果と疑似観測データの誤差の平均と標準偏差を第 2 フェーズの結果として 3 に示す。150 タイムステップ目の 2 箇所での出力のユークリッド距離を第 2 フェーズの誤差とした。

観測サイズを増やすと精度は上がるが、ある段階で悪化する。標準偏差から観測点数が少なくても、適切に設定すれば良い精度を得られることが分かった。

4 おわりに

本稿では、データ同化の可能性を評価する枠組みを提案し、実験による評価を行った。パレート解を許容したデータ同化の効果に対する観測サイズの限界があるという結果になった。

謝辞 本研究は JST CREST の助成を受けたものである。

参考文献

- [1] 竹内内篤 *et al.*, “歩行者・自転車マイクロシミュレーションの現況再現性の検証に関する一考察,” 土木計画学研究・講演集, 41, 367, Jun. 2010.
- [2] 上野秀樹 *et al.*, “交通現象を高精度で再現できるマイクロ交通流シミュレータ”, 東芝レビュー, 64, 4, pp.23–26, Apr. 2009.
- [3] 山下倫央 *et al.*, “一次元歩行者モデルを用いた高速避難シミュレータの開発とその応用”, 情報処理学会論文, 53, pp.1732–1744, July. 2012.