

ベンチマークデータ公開のための枠組み

内田 敬*

by Takashi UCHIDA*

1. はじめに

“ベンチマークデータ”、“検証手順の標準”をダイナミックな過程・手続きとして捉えて欲しい。実交通現象の量的側面（交通量、交通密度、速度など）をシミュレート（再現）することを目指すとき、ターゲットとなる観測データが第一に必要である。観測データの蓄積、それを用いたシミュレーションモデル検証事例の蓄積の過程において、“自ずから”標準は定まるものであろう（この過程を経ずして標準は設定し得ない）。ここでは、“蓄積”のための枠組み——ベンチマークデータ公開へ向けた仕組み、運用・発展側面における仕組み——について述べる。“標準”の内容（あるいは十分条件）については論じない。

2. 基本理念、組織

ベンチマークデータを巡る主体を、①データ作成者、②モデル開発者、③モデル検証者、④モデル利用者、とする。あるひとつのデータセット、モデルに対しては、①～④それぞれが互いに異なる個人・機関であり得ることを保証することが“標準”の前提である。

ベンチマーク化は、図1に示すように、これら4主体の共働によりなされる。ベンチマーク化に関与する意志を有する個人・機関は、そのときどきでいずれかの主体として関与し、それに伴う図1のような権利・義務を負う。例えば、モデル開発者は、観測データを利用することができる／利用しなければならない。そして、作成したモデルと設定データを第3者が実行可能な形態で提供しなければならない。この様な主体間の連携を、自由参加、透明性、検証可能性の下に実現する機構としてクリアリングハウス（clearinghouse: 手形交換所、情報交換所）が位置する。

3. クリアリングハウス

クリアリングハウスの実体は、a)メタデータのオンラインデータベースと b)履歴参照可能な情報交換システムである。“③モデル検証者”の手順を例にクリ

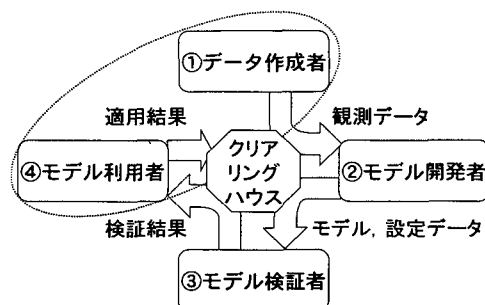


図1 ベンチマークデータ公開の枠組み

アリングハウスの機能を述べる：

- 1)観測データ、モデルそれぞれに関するメタデータオンラインデータベース(a)から検証対象を選定する。
- 2)情報交換システム(b)を通じて、①データ作成者、②モデル開発者に、データ、モデルの提供を要請する。
- 3)検証結果 (i.モデルの再現性、 ii.観測データと評価対象モデルの適合性)をまとめたドキュメントをメタデータ DB に登録し、概要を情報交換システムを通じて公表する。

クリアリングハウスは仲介所に過ぎないが、これを介することで、著作権やデータ改竄の問題を回避し、また、観測データとモデルの任意の組み合わせによる検証を可能とする。データ要求や問い合わせを履歴参照可能な情報システムを介して行うことで、透明性の確保と経験の蓄積が実現できる。この、検証・蓄積過程を経ることで品質保証がなされ、状況ごとにふさわしいベンチマークデータが定まっていくであろう。

4. 課題

- 1) 汎用性のある観測データセットが可能か？
- 2) クリアリングハウスの運用組織、進行役、書記が必要 → 交通シミュレーション実用化促進ワークショップが担える／担うべきか？

参考文献

- 1) Jeff Rothenberg: Metadata to Support Data Quality and Longevity, FIRST IEEE METADATA CONFERENCE, April 16-18, 1996.

* 正会員、博(工)、東北大学大学院工学研究科土木工学専攻
(〒980-8579 仙台市青葉区荒巻字青葉 06)
TEL 022-217-7476, FAX 022-217-7477
E-mail: uchida@civil.tohoku.ac.jp