



AALBORG UNIVERSITY
DENMARK

Aalborg Universitet

Inaccuracy in Traffic Forecasts: Part 1

Flyvbjerg, Bent; Holm, Mette K. Skamris; Buhl, Søren L.

Published in:
Expressways and Motorways

Publication date:
2007

Document Version
Publisher's PDF, also known as Version of record

[Link to publication from Aalborg University](#)

Citation for published version (APA):
Flyvbjerg, B., Holm, M. K. S., & Buhl, S. L. (2007). Inaccuracy in Traffic Forecasts: Part 1. *Expressways and Motorways*, 50(10), 56-65.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

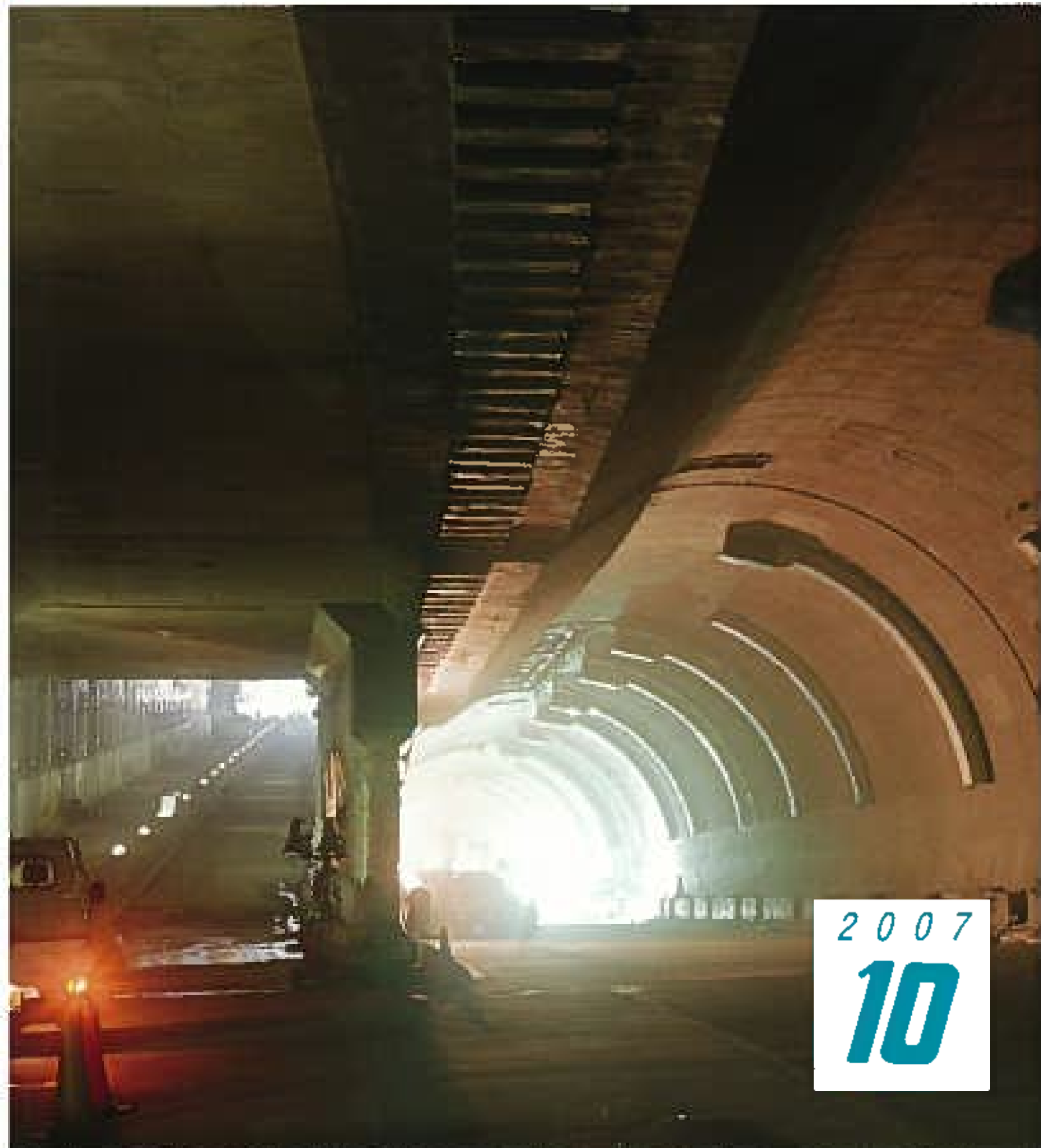
Take down policy

If you believe that this document breaches copyright please contact us at vbn@aub.aau.dk providing details, and we will remove access to the work immediately and investigate your claim.

高速道路と自動車

VOL. 50

NO. 10



2007

10

多くの交通インフラ・プロジェクトの需要予測を統計的に分析し、プランナーが妥当で信頼できるリスク評価をはじめのようなデータとアプローチを示すとともに、誤差の要因や認知のバイアスの存在を踏まえ、説明責任の改善と参照クラス予測という解決策を論ずる。

Bent Flyvbjerg, Mette K. Skamris Holm and Søren L. Buhl
 "Inaccuracy in Traffic Forecasts", *Transport Reviews*, Vol. 26, No. 1,
 1-24, January 2006.

交通量予測の精度 (上)

B. フ リ ュ ビ ア^{*}
 M. K. スカムリス ホルム^{*}
 S. L. ブール^{*}

要 約

本稿は、交通インフラ・プロジェクトにおける交通量予測に関する、最初の統計的に有意な研究からの結論を示すものである。使用した標本は、この種の研究では最大であり、14カ国210プロジェクト580億米ドル相当額を扱っている。研究は、予測者が全般的に交通インフラ・プロジェクトの需要推定に不十分な仕事しかしていないことを非常に高い統計的有意性ととも示している。その帰結は実質的に財務的経済的リスクを悪化させる面を持つ。考察された30年間に、予測精度がより高くなることはなかった。精度の高い需要予測に到達する専門技術と特殊技術が時間の経過につれて改善されたとしても、しばしば予測者が主張するように、これはデータには表れない。鉄道プロジェクト10のうち9つには、旅客数の予測が過大に推計されており、過大推計の平均は106%である。鉄道プロジェクトの72%には、予測は2/3以上過大に推計されている。道路プロジェクトの50%には、予測交通量と実際交通量との差が±20%を上回り、道路プロジェクトの25%には、その差が±40%を上回る。

道路に対する予測は鉄道より精度が高く、バランスがとれており、過大予測と過小予測の頻度にそれほど差はない。しかし、鉄道と道路の両プロジェクトとも、需要予測が許容範囲内で正しくないのが実質的なリスクが存在する。予測精度の低さの原因は鉄道と道路のプロジェクトで異なり、政治的原因は道路よりも鉄道に大きな役割を果たしている。解決法は、さらなる説明責任と参照クラス予測である。大きな標準偏差と結びついた、きわめて精度の低い交通量予測は大規模な財務的経済的リスクに転嫁される。しかし、社会的経済的厚生への損害へのそのようなリスクは概して、プランナーや意思決定者によって無視されるか、軽視されている。本論文が示すデータ、接近法で、プランナーが妥当かつ信頼に足るリスク評価をしはじめるかもしれない。

序

交通インフラに費やされている莫大な金額にもかかわらず、驚くことに、関係する費用と便益、リスクについて体系的知識はあまり存在しない。交通インフラ・プロジェクトが予測通り成果をあげているかどうかという核心の自明な疑問に対し、文献は統計的に妥当な答に欠けている。プロジェクトが期待を下回る

* オールボー大学開発・計画学部

と、これは不運な事情による単発的な例としてしばしば説明される。つまり、それは概して成果が期待を下回る交通インフラ・プロジェクトに一般的なパターンの特別な表現とは見なされないのである。この研究領域における知識不足ゆえ、期待はずれの成果が例外か規則かを適切に反駁ないし確証することはこれまで不可能だった。

先行する3論文で、本稿の著者は費用、費用関連リスクに関してプロジェクト成果の疑問に詳細に答えた。プロジェクトは費用に関して予測通りの成果をあげておらず、10のうちほとんど9つのプロジェクトが顕著な費用増大という犠牲を払っていることが明らかにされた。著者はまた、そのような期待はずれな成果の原因と解決法も研究した (Flyvbjerg *et al.*, 2002, 2003b, 2004; さらに Flyvbjerg *et al.*, 2003a)。本稿は、投資の便益面に焦点をあて、プロジェクトが必要と取入リスクに関して、予想通りの成果をあげているかどうかの疑問に答える。それは、多数の交通インフラ・プロジェクトに対して、需要に関して予測された成果を実際の成果と比較している。費用リスク、便益リスク、複合リスクについての知識は、プロジェクトを展開したり建設すべきかを決定するとき、プランナーと意思決定者にとってきわめて重要である。交通インフラ・プロジェクトについて、関係する費用と便益はしばしば数億ドルにのぼり、リスクはそれに応じて高くなる。

Pickrell (1990) と Richmond (1998) に指摘されたように、プロジェクトの財務的妥当性の評価は、交通需要予測の精度に強く依存する。そのような予測はまた、交通インフラ・プロジェクトの社会経済的および環境上の評価の基礎でもある。交通量、交通の空間分布、交通のモード間分布を扱う交通部門の需要予測の精度について得られた経験によれば、費用予測と同様の方法で、モデル作成におけるあらゆる科学的前進にもかかわらず、需要予測が交通インフラ・プロジェクトの評価における不確実性とリスクの主要な源泉だという証拠が存在する。

交通量予測は交通インフラ・プロジェクト構築に常套的に用いられている。ここで、予測の精度は希少な資金の効果的配分にとり、著しく重要性をもつ問題である。例えば、20億USドルのバンコクのスカイト

レインがかなり過大な規模を示したのは、旅客数予測が実際の交通量の2.5倍以上だったからだ。結果として、駅のホームは現行システムの短編成の列車には長すぎる、それらへのニーズがないので多くの列車と車両が列車車庫に留められ遊休している、ターミナルが大きすぎる等があげられている。プロジェクト会社は財務的に困難な状況に陥ってしまい、たとえ都市鉄道がバンコクのような混雑し大気の汚染された都市にとって良い考えだったとしても、ムダな輸送力への過剰な投資は決して最適な資源使用方法ではなく、特に投下資本不足の発展途上国においてはそうでない。逆に、英国会計検査庁 (UK National Audit Office) の研究は、交通量予測が過小だったため、過小規模となった多くの道路プロジェクトを識別した。これもまた数百万ポンドの非効率をもたらしたが、それは、既存の十分利用されている道路に容量を加えることはその容量をあらかじめ建設しておくよりもずっと多額の費用を要するからである (National Audit Office, 1988)。こうした理由で、交通量予測の精度は重要なのである。

それにもかかわらず、厳密な精度研究はほとんどない。そのような研究は存在するところでは、特徴的に少数調査なのである。すなわち、単一事例の研究であるか、体系的な統計分析と認めるには、小さすぎたり規則的でなさすぎるプロジェクトの標本ばかり扱っているのである (Webber, 1976, Brooks and Trevelyan, 1979, Mackinder and Evans, 1981, Fullerton and Openshaw, 1985, National Audit Office, 1983, 1992, Fouracre *et al.*, 1990, Kain, 1990, Pickrell, 1990, Walmsley and Pickett, 1992, World Bank, 1994, Richmond, 1998)。他の点での価値にもかかわらず、これらや他の研究をもってして、交通インフラ・プロジェクトに対して交通量予測はどれくらい精度が高いかという問題に統計的に満足のいく答を与えることはこれまで不可能であった。

本研究の目的は、交通インフラ・プロジェクトについて、精度の問題に統計的に妥当な答を与えるに十分な大きさの標本を確立することによって、この状況を変えることであった。この知的目的に加えて、交通需要に関してプロジェクトの現実的で妥当なリスク評価を実行するツールをプランナーに与えるという実践

的目的もある。既存研究はほぼすべて交通量予測には過大推計される強い傾向があると結論づけている (Mackinder and Evans, 1981, p. 25; National Audit Office, 1985, appx 5.16; World Bank, 1986; Fouracre *et al.*, 1990, p. 1, p. 10; Pickrell, 1990, p. x; Walmsley and Pickett, 1992, p. 2; Thompson, 1993, pp. 3-4)。本稿は以下で、この結論が既存研究で用いられた小規模標本からの結果であることを示すだろう。つまり、その傾向はプロジェクト母集団にはあてはまらない。プロジェクトの標本が10-20の要素によってより代表的なものに拡張されると、そこには、1つの別な関式が見出される。すなわち道路プロジェクトには、予測の問題はただ単に交通量の過大評価というものでない一方、鉄道には、過大推計がまさに重要な問題なのである。

交通量予測の精度の計測

交通量予測は交通インフラ・プロジェクトの構築と規模を正当化するため常套的に用いられている。そのような予測の精度を評価するため、予測交通量を実際交通量と比較する必要がある。ここでは一般の慣例に従い、交通量予測の精度は予測交通量に対する百分比における交通量の実際量マイナス予測量として定義される。実際交通量は供用1年日（あるいは開業年）に計測される。予測交通量はプロジェクト建設決定時点で評価された供用1年日（あるいは開業年）の交通量推計である。このように予測は問題のプロジェクト建設を決定した意思決定者が基礎をおく、利用可能な推計である。もし建設決定時に利用可能な推計がなかったならば、もっとも時期的に近い利用可能な推計が用いられたが、概して後の推計からは、精度の低さについての我々の計測には控えめなバイアスがかかる結果に帰した。精度のこの定義によれば、完全な正確さは0で示される。例えば-40%の精度は、実際交通量が予測交通量を40%下回ることを示すのに対し、+40%の精度は、実際交通量が予測交通量を40%上回ることを意味するだろう。

比較の基礎としての1年目

プランナーと推進者はときに交通量予測における精度のこの計測方法に異論を吐えることがある

(Flyvbjerg *et al.*, 2003a)。概して時の経過につれてより精度を増す予測値を用いて、プランニングと実行の別々な段階で様々な予測がなされていると彼らは述べている。それゆえ、建設決定時点での予測は最終的なものからかけ離れる。したがって、そのような初期の推計値は非常に精度が低いと予想されるだけなので、この推計値を交通量予測の精度評価の基礎として使用するのには不公平だといったような反論が続く。しかし、意思決定に、それゆえ意思決定者が利用可能な情報の精度に焦点がおかれるとき、それはまさしく建設決定時点での予測交通量であるから、ここではこの方法が支持される。さもなければ、建設か否かの情報が伝えられたと評価することは不可能であろう。建設決定後になされる予測はこの決定と無関連の定義による。精度の低い予測の理由が何であれ、立法者と市民、あるいは民間で資金調達されるプロジェクトの場合の民間投資家には予測交通量と収入の不確実性を知る権利がある。さもなければ、透明性と説明責任が損なわれている。初期の交通量推計の精度の低さが単に不完全情報の問題、遠い将来を予測することに特有の困難さだったならば、プロジェクトの推進者と予測者がしばしば11にするように、精度の低さをランダムであるかランダムに近いものと期待することも観察される。しかし、以下で示されるように、精度の低さは顕著できわめて興味深いバイアスを有している。

プランナーと推進者は予測の精度を測る基礎として供用1年日（あるいは開業年）の交通量を用いることにもときに反対することがある。ここで研究されたプロジェクトの1つ、ドーバー海峡トンネルの所有者であるユーロトンネル管理者は、いくつかの我々の業績へのコメントで次のように述べている。「プロジェクトの開業初年の交通量収入に基づき成功か失敗かを判断するが誤りを導く（ユーロトンネルから著者への手紙, 1999)。」初期問題をプロジェクトが経験し、海底トンネルにとって非常に重大な問題だったとするなら、これは当初、交通量に負の効果を与えるかもしれないが、この議論に従えば、それは一時的なものにすぎず、この基礎に基づき予測の精度を測定することは判断を誤らせることになる。初期問題が終わると、正常な運用が確実になり交通量は増大するので、これが精度を測定する基準であるべきだというように、議論

は続く。さらに、移動者が事実上新たな交通施設を見つけ利用するまでと、それに応じて彼らの移動行動を変えるまでには時間がかかる。惰性が1つの要因なのである。供用1年間に予測交通量を下回るプロジェクトは、数年後には予測量にうまく追いつくかもしれないので、その基礎に基づき精度を測定する方が適切であろう。もし供用1年目が比較の基礎として用いられるならば、不十分な成果のプロジェクトが多すぎるということを確認する結果になるだろう、などといったほどの反対がこの基礎を用いることにはある。

一見、この議論には説得力があるように思われ、原則上(実践上とは異なり)、精度計測の基礎として、供用1年目でなく別の期間を用いることを妨げるものは何もない。例えば、供用5年目を用いることが決定されるかもしれないが、その理由はそのときまでには初期問題は解消されると期待されるかもしれないことにある。一方、例えば土地利用における重要な外的変化はこの時点でも完全には発展していないだろう。しかし、より最近の研究では、精度計測の基礎として供用1年目を支持する以下の理由が存在する。

第1に、供用開始後1年以上にわたる予測交通量と実際交通量のデータが存在するようなプロジェクトにとって、供用1年目に交通量が予測量を下回ったプロジェクトは、後年も交通量が予測量を下回る傾向にあることがわかっている。それゆえ、精度計測の基礎として供用1年目を用いることは、もし比較の基礎として別の期間が用いられたとすれば識別されなかったプロジェクトを不十分な成果と誤って識別してしまうことはあまりないようである。このタイプのプロジェクトでは、実際交通量が急速に予測交通量に追いついたりしないのは明らかであり、決して追いつかないこともある。Pickrell (1990) によって分析された、都市鉄道プロジェクト10のうち7つに関する追跡研究は、時間の経過につれて、大きな利用客の獲得はないことを示した。実際、バルチモア (Baltimore)、バッファロー (Buffalo)、そしてピッツバーグ (Pittsburgh) では、利用客は時間の経過につれて減少したのだった (Richmond, 1998)。一般開放後5年以上のドーバー海峡トンネルでは、列車ユーロスターの旅客数は、開業年に予測された数値の45%にしかならなかった。鉄道貨物輸送は予測された数値の40%であっ

た。結果はそれぞれ破産に近い状態となっている。英国のハンバー (Humber) 橋では、一般開放後16年、実際交通量は依然予測量の約半分だけだった。デンマークでは、ニューリトルベルト (New Little Belt) 橋の実際交通量が予測交通量に追いつくまでに20年以上を要し、数年間、予測量と実際量の差は小さくなるどころかむしろ大きくなった。そのような知見は、長期的に予測誤差は相殺し合う傾向にあるという、予測における慣例的な知恵は誤りだとするミルゼジョウスキ (Mierzejewski) の観察 (1995, pp. 31-32) と適合している。誤差はしばしば、より後の年に計測されるとき、供用1年目に対する計測時と比較して、後の年に対する計測時に不正確さが増すという帰結と互いに補強し合うことがある。この論理に従えば、精度計測の基礎として供用1年目を用いることは、交通量予測の全体的な精度の低さを過小評価する傾向をもつことになる。

第2に、ここで焦点があてられる大規模交通インフラ・プロジェクトの種類に対しては、供用開始1カ月の間、観光交通が相当あるかもしれず、それらの多くは、できるだけ効率的にA地点からB地点に客貨を運ぶよう設計された平凡な輸送機械であることに加えて、建築的かつ技術的な驚異である。観光交通は、例えば新設の橋梁や鉄道新路線のような、問題の新交通施設を見たり試してみたい人々の欲求に基づいたプロジェクトにより惹きつけられる交通である。事例を示せば、スウェーデン-デンマーク間のオアスン (Øresund) 橋では、供用後最初の1カ月間の道路交通量は後年同月の交通量より19%多かった。2つの月の差は主に観光交通量に帰因するものであり、プロジェクト会社が予想したよりやや少なかった (Trafikministeriet, Finansministeriet and Sund & Bælt Holding, Ltd, 2002, appx 4:2)。観光交通は、少なくとも衆目を十分に集めるプロジェクトでは、上述の初期問題から交通需要量へのありうる負の影響を相殺するのに役立つかもしれない。プロジェクトの開業当初の間の交通量に対するそのような相殺する影響力の存在は、なぜ供用1年目が交通量予測の精度計測にとってかなり正確な基礎となる傾向があるかを説明する助けとなる。

第3に、経験的事実として、予測者とプランナーは

概して供用1年目を予測を行う原則的な基礎として用いることが観察されるかもしれない。所与のプロジェクトにとって、一般的にこれが意思決定者に与えられる主要予測であり、それは意思決定者が建設するか否かの決定で保有する情報の重要部分を形成している。そのような決定に適切な情報が与えられたかどうかを評価したければ、評価を要するのはこの予測の精度であり、それゆえ、開業1年目の実際交通量をその年の予測交通量と比較する必要がある。

第4に、實際上、例えば開業5年目の交通量予測が存在するプロジェクトはわずかしは見出されず、かつ、実際交通量がこの年に計測されたので精度の低さはこの年に体系的に計測されたのかもしれない。より多くのプロジェクトが後年よりも供用1年目に対する予測交通量と実際交通量についての情報をもっていることが見出されるが、その理由は予測者と予測精度の評価者の双方にとり、作業の基礎として供用1年目を用いることは共通の慣行らしいことにある (Touracre *et al.*, 1990; Pickrell, 1990; National Audit Office, 1992; Walmsley and Pickett, 1992; World Bank, 1994)。

最終的に、多くのプランナーや推進者によって主張されるように、もし新たに開放された交通インフラ・プロジェクトが交通量増大前に規則的な適応期間を持つならば、これは交通需要のモデル化において統合されうるし、そうされるべきである。このようにして、適応は予測に反映されることになり、それらの外にはない。

データの利用可能性

交通量予測の精度を算定させてくれるデータは残念ながら比較的少ない。公的部門のプロジェクトについては、データが簡単にもたらされることはあまりない。そして、データ提供の意図があるところでさえ、プロジェクトは予測交通量と実際交通量の比較を困難か不可能にする方法で展開しているかもしれない。例えば、所与のプロジェクトに対する予測交通量は、開業年に対し推計されたかもしれないが、よくある遅延のせいで、実際の開業データは予測されたデータよりも数年遅れになるので、実際の開業年に対しなされた交通量予測は存在しない。より一般的な言葉では、どのように、そして何を基礎に予測交通量と実際交通量

が推計されるかの方法論的相違が比較をしばしば困難か不可能にする。結局、大規模プロジェクトでは、予測がなされてから建設決定まで、着工まで、プロジェクト完成まで、供用開始まで、そして実際交通量の最終的算定ができるまでに時間の経過は5年、10年あるいはそれ以上になるかもしれない。そのような長期にわたり、予測の基礎となる仮定が実際交通量の測定方法の基礎となる仮定と比較されるとき、時代遅れで同じ基準で測定されないか、実際交通量を予測交通量と比較する当初計画は断念されるか、単に忘れられるかもしれない。

民間部門のプロジェクトについて、交通量は概してプロジェクトの事業主への収入を生じさせる。予算計上と会計は商業的であり、それゆえ、交通量予測と交通量計測の傾向が、予測交通量と実際交通量の比較研究にとり、公的部門のプロジェクトの場合よりも体系的で資するものとなる。しかし、これは研究者には概してあまり役立たない。というのは、民間プロジェクトの交通量データは競合者から守るために機密扱いされていることがよくあるからだ。そして公的プロジェクトと民間プロジェクトの双方に対して、予測交通量と実際交通量の比較を可能とするデータがプロジェクト事業主と管理者によって提出されることはないかもしれない。その理由は、予測交通量と実際交通量の差異の大きさと方向性が、例えば実際交通量が予測交通量よりも相当少ない場合、もし公になれば、プロジェクトを衆目上悪く見せる類だからかもしれないことにある。

本研究における使用データ

上述のデータの希少性に関する問題にも関わらず、データ収集と改良の1年後、著者たちは予測交通量と実際交通量に対し比較可能なデータをもつ210の交通インフラ・プロジェクトの標本を開発した。標本は実際原価(2003年価格)で約580億米ドル相当額のプロジェクト資産構成からなる。プロジェクトの種類は都市鉄道、高速鉄道、在来鉄道、橋梁、トンネル、幹線道路、高速道路である。プロジェクトは5大陸14カ国に位置し、先進国と途上国双方を含む。プロジェクトは1969年から1998年の30年間に完成された。プロジェクトの規模は建設費2,200万米ドルから100

億米ドルの範囲にわたり（2003年価格）、最小規模のプロジェクトは概してより大規模な道路スキームの道路延伸、最大規模のプロジェクトは鉄道リンクと固定リンク（トンネルと橋梁）である。知られる限り、これは、この種のプロジェクトのために確立された、予測交通量と実際交通量の比較可能データを持つプロジェクトの最大の標本である。

プロジェクトはデータの利用可能性に基づく標本に選ばれた。予測交通量と実際交通量に関する比較可能なデータが入手できたと知られているプロジェクトはすべて標本に包含されるよう考慮された。これは485プロジェクトあった。不明瞭ないし不十分なデータの質のため、全部で275のプロジェクトが棄却された。より明確には、棄却された275プロジェクトのうち、124が棄却された理由は、精度の推定に決められた方法（前述参照）と異なり比較できない方法で不正確さが評価されてしまったことにある。151のプロジェクトが棄却された理由は、これらのプロジェクトに対する予測の精度が、実際交通量に決められたような実際データを用いる代わりに、調整されたデータの基礎に基づいて推計されたことにある。妥当で信頼に足るデータが利用可能なプロジェクトはすべて標本に含まれた。これは、著者がデータを収集したプロジェクト、他の研究で他の研究者がデータを収集したプロジェクトのどちらも網羅している。

著者自身のデータ収集はヨーロッパの大規模プロジェクトに集中したが、その理由は比較研究を可能とするにはこの種のプロジェクトにとって存在するデータが少なすぎるからである。主要データが交通量予測の精度に関してデンマーク、フランス、ドイツ、スウェーデン、英国における31プロジェクトに対し収集され、実際交通量と推計交通量の双方にとり信頼に足るデータをもつヨーロッパの大規模プロジェクトの数は数倍に増えたので、統計的手法が適用されるこの種のプロジェクトにとり最初の比較研究を可能にさせた。他プロジェクトは以下の研究から標本に含まれた：Webber (1976)、Hall (1980)、National Audit Office (1985, 1988)、Fouracre *et al.* (1990)、Pickrell (1990)、Walmsley and Pickett (1992)、Skamris (1994) and Vejdirektoratet (1995)。統計的な検定から、我々自身の調査を通じて収集されたデータと他の

研究者により実施された研究から収集されたデータとの間に違いがないことが示された。

どの標本に関しても、重大な問題は標本が母集団を代表しているかどうか、ここでは、標本に含まれたプロジェクトが交通インフラ・プロジェクトの母集団を代表しているかどうかである。標本の抽出基準がデータの利用可能性、妥当性、信頼性だったので、この問題は利用可能で妥当な信頼に足るデータを持つプロジェクトが代表的であるかに言い換えられる。これは4つの理由でおそらくあてはまらない。第1に、そのようなデータがプロジェクト管理によってプロジェクトを監視するために用いられるとき、成果の評価を可能にするデータの存在自体がまさに成果の改善に寄与するかもしれないと主張されてきた (World Bank, 1994, p. 17)。そのようなプロジェクトは平均よりましな、すなわち非代表的な成果をもつだろう。第2に、交通量予測に関して特に悪い実績をもつプロジェクトの管理者と推進者は交通量データを利用可能にさせないことに関心があると推測されたかもしれず、これはそのとき標本のそのようなプロジェクトの代表を下回る結果になっただろう。逆に、交通量予測に対して良好な実績を持つプロジェクトの管理者と推進者は、これを公開することに関心があるかもしれず、これらのプロジェクトの代表を上回る結果になる。第3に、管理者は交通量データを利用可能にした場合でさえ、プロジェクトを都合よく位置づけるデータを示す選択をしてきたかもしれぬ。しばしば、所与の時点で所与のプロジェクトから選択されるいくつかの予測交通量と、それに対する実際交通量をいくつか編集したものが存在する。よくある事実のように、研究者が調査質問票の手段によってデータを収集するならば、最も適合する予測交通量と実際交通量の組み合わせ、おそらくそのプロジェクトをよく見せる組み合わせを選ぶ誘惑が管理者にはあるかもしれぬ。ヨーロッパの大国のある経験豊富な研究者は、その国に対する我々の研究へのフィードバックを提示し、データ収集に関してコメントしている。

[研究]の多くが質問票への[国鉄]の回答に基づいている。これがシステムティック・バイアス（系統的偏向）を生み出しているらしい。

[国鉄]が これらの問題に関して、あなた方

に真実を伝えているとは信じられない。あなた方がよく知っているように、これらの問題における「真実」の概念は特にろい。[国鉄]が予測に対して最も適合する数値をとる誘惑、この誘惑は大きいと違わず、私は彼らがそれに抵抗できるとは思わない。あなた方にとって[より良好なデータを得るために]必要となるのは、理想的には([国鉄]から)ではなく)交通省の公文書からの、決定に使用された予測の原資料でしょう。

他の研究がデータのそのような「調理」の存在を立証してきた(Wachs, 1990)。残念ながら、現実には、予測の原資料に近づくことは困難か不可能なことがよくあると知られている。これが、我々に加えて他の研究者が、ときに調査質問票による改善の方法に依拠しなければならない理由である。これがまた、データにバイアスがかけられそうな理由でもある。最後に、異なる副次標本の代表性の差も非代表的なデータ、例えば鉄道、道路間の差をもたらすかもしれない。この点は以下で振り返ることにしよう。

利用可能なデータからは、不実表記問題の大きさを正確に実証的に査定させてはくれない。しかし、前述した理由のため、大抵バイアスがかかった標本らしいこと、そのバイアスは控えめであることが結論される。言い換えれば、標本から推計された交通量予測の精度はプロジェクト母集団の交通量予測の精度よりも高くなるだろう。これは、以下の統計分析からの帰結を解釈するとき、留意されるべきである。標本はどうしても完全ではない。依然、それがこの研究領域における現行の最高技術水準で付与された、得られる最善の標本なのである。

統計分析上、標本における実際交通量と予測交通量との百分比上の差異は、特に断らない限り、正規に分布すると考えられる。残りの作図(非表示)は、正規分布が完全には充足されず、鉄道データは2つのはずれ値をもち、道路への分布は上向きの大きな尾をもち幾分不均衡であることを示している。後者では、対数変換が正規性を改善するが、これは注目に値するとは考えられなかった。その理由は、部分的には検定が正規性からのずれに対しかなり頑健だからであり、部分的にはそれが解釈を複雑にするからである。

分析の一部として実行される標本の細分化はそれ自

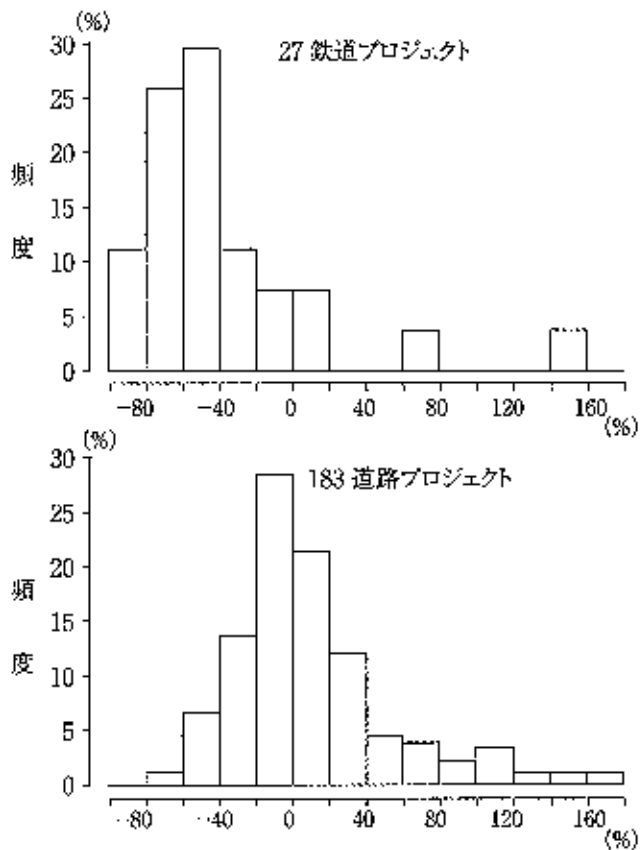
体の方法論的問題を必然的に伴っている。それゆえ、副次標本の別な組み合わせにおける観察結果の表記は検討されるデータに対し相当不均衡である。表記がより単調だったならば、分析は著しく改善されただろう。部分的かつ全面的な混乱が生じる、すなわち、もし2つ以上の効果の組み合わせが意義をもつならば、いずれか一方ないし双方が差異の原因であるかを決定することが困難なことがある。相互作用に対しては、全ての組み合わせが表示される訳ではないことがよくあるか、表記がきわめて稀なことがありうる。実際、副次標本に関する有益な結論はこれらの理由のためわずかしか見出されないの、我々のデータの解釈がこれらの制約に適用されてきた。もし良好なデータが収集されたならば、より明確な結論が導かれることがありえた。

使用された統計モデルは線形正規モデルであり、適切なF検定とt検定を伴う分散分析と回帰分析がなされた。方法に関して仮説の検定は正規性からのズレに頑健であると知られている。各検定に対して、 p 値が報告された。

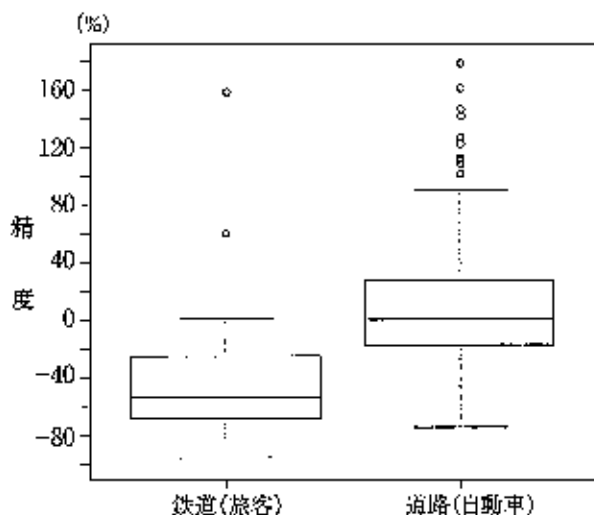
鉄道と道路、予測精度が高いのは？

図-1と図-2は、プロジェクトを鉄道と道路に分けた標本で210プロジェクトの交通量予測精度の分布を示している。既述のとおり、精度は実際交通量マイナス予測交通量を予測交通量の百分比として測定されている。それゆえ、完全な正確さはゼロで示される。負の数値は実際交通量が予測交通量より少ないこと、正の数値は実際交通量が予測交通量よりも多いことを示している。図-1と図-2の最も注目する特性は鉄道と道路のプロジェクト間の顕著な差である。鉄道旅客数の予測は、道路交通量の予測よりずっと精度が低く、偏って(水増しされて)いる。

検定が示しているのは、統計分析に含まれた27鉄道プロジェクトのうち、ドイツの2つのプロジェクトは統計的はずれ値と考えられるべきだということである。これらは、図-1の鉄道ヒストグラムでは右側に位置する2つの柱状部と、図-2に示された鉄道箱線図では上方の2つの点によって表される2プロジェクトである。2つの統計的はずれ値を伴うものと伴わないものの検定は、都市鉄道のみが合理的表示をもつと



図一 交通インフラ・プロジェクトにおける交通量予測の精度



図二 210 交通インフラ・プロジェクトにおける交通量予測の精度

いう留保条件とともに、鉄道プロジェクトの異なるタイプ間の予測の精度に何ら差を示さない。それゆえ、鉄道プロジェクトは一集合体と考えられる。統計的なはずれ値を除けば、残りの 25 鉄道プロジェクトにつ

いては以下が見出される（2つの統計的なはずれ値を含む結論は括弧で与えられる）。

- ・データは水増しされた鉄道旅客数予測のもつ強力な問題を示している。鉄道プロジェクト 10 のうち 9 以上では、旅客数の予測は過大に推計されており、全ての鉄道プロジェクトの 72% に対して、旅客数の予測は 2/3 を上回って過大に推計されているのである。

[統計的なはずれ値を含む場合、全ての鉄道プロジェクトの 67% に対して、旅客数の予測は 2/3 を上回って過大に推計されている]

- ・鉄道旅客数の予測は、平均 105.6% 過大に推計され（95% 信頼区間（66.0, 169.9））、実際交通量は予測交通量よりも平均 51.4% 低い結果になった（SD=28.1, 95% 信頼区間（-62.9, -39.8））。

[統計的なはずれ値を含む場合、鉄道旅客数の予測は平均 65.2%（95% 信頼区間（23.1, 151.3））過大に推計され、実際交通量は予測交通量よりも平均 39.5% 低いという結果になった（SD=52.4, 95% 信頼区間（-60.2, -18.8））]

- ・鉄道プロジェクトの合計 84% は、実際交通量が予測交通量を 20% より下回り、実際交通量が予測交通量を 20% より上回るものはない。たとえ閾値を倍の 40% に設定しても、全ての鉄道プロジェクトの 72% ではずっと、実際交通量はその限度を下回っている。

[統計的なはずれ値を含む場合、数値はそれぞれ 85% と 74% である]

道路プロジェクトでは、信頼性 95% で、高速道路、橋梁、トンネル（170 の高規格道路、10 の橋梁、3 つのトンネル）における自動車交通量の間で、予測の精度に関して有意な差はない（ $p=0.638$ ）。それゆえ、183 の道路プロジェクトは一集合体と考えられる。検定は以下に示す通り（表一）。

- ・道路プロジェクトの合計 50% は実際交通量と予測交通量の間には $\pm 20\%$ を超える差がある。閾値を倍の $\pm 40\%$ に設定しても、プロジェクトの 25% がこの水準を上回る。

- ・道路自動車交通に対し、予測の水増し対割引きの頻度に関して有意な差はない（ $p=0.822$, 両側二項検定）。プロジェクトの合計 21.3% は -20% を下回る精度をもっている一方、プロジェクトの 28.4% は $+20\%$ を上回る精度をもつ。

表一 交通量予測の精度 (鉄道旅客数と道路交通量)
(%)

	鉄道*	道路
平均的な精度	-51.4 (SD=28.1) [-39.5 (SD=52.4)]	9.5 (SD=44.3)
±20%を上回る低精度プロジェクト	84 [85]	50
±40%を上回る低精度プロジェクト	72 [74]	25
±60%を上回る低精度プロジェクト	40 [41]	13

※ [] 内の数字は、2つの統計的はずれ値を含む。

・道路交通量の予測は平均 8.7% 過小に推計され (信頼区間 95% で, 2.9, -13.7), 予測交通量を平均 9.5% 上回る実際交通量の結果になった (SD=44.3, 95% 信頼区間 (3.0, 15.9))。

それゆえ、道路交通量の予測は許容範囲内で正しくないので実質的なリスクはあるが、そのリスクは鉄道旅客数の予測よりもずっとバランスがとれている。鉄道、道路間の差を検定すれば、非常に高い統計的有意水準で、鉄道旅客数の予測は道路自動車数の予測よりも精度が低く、より水増しされている ($p < 0.001$, ウェルチ 2 標本 t 検定)。しかし、鉄道予測と道路予測に対する標準偏差の間で、有意な差の指標はない。すなわち双方とも高く、両予測タイプに対して不確実性とリスクの要因が大きいことを示している ($p = 0.213$, 両側 F 検定)。鉄道に対する 2 つの統計的はずれ値を除けば、鉄道プロジェクトに対する標準偏差は、依然として高いけれども、道路プロジェクトに対するよりも著しく低い ($p = 0.0105$)。

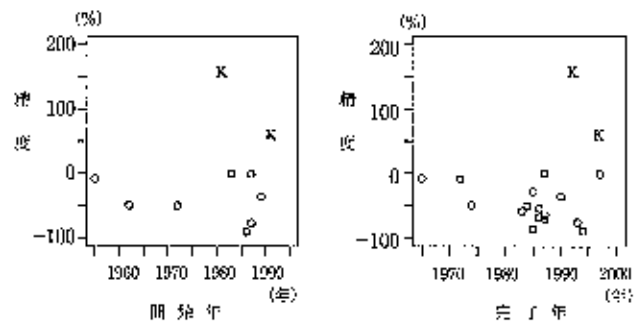
鉄道インフラ開発の意思決定に用いられた交通量推計は高度に、体系的に、注目値するほど誤りを導いていると結論される。鉄道旅客数の予測は一貫して顕著に水増しされている。道路プロジェクトにとり予測を誤らせる問題は鉄道にとってほど深刻でなく、一方的でない。しかし、道路でさえ、プロジェクトの半分にとって、実際交通量と予測交通量の差が ±20% を超えている。この背景にもとづき、プランナーと意思決定者は将来交通量を予言する不確実性をはっきり考慮に入れないなどの交通量予測も割り引いて聞くように

く助言されている。鉄道旅客数の予測にとり、これでは十分でないかもしれない。鉄道プロジェクトと道路プロジェクトのどちらにとっても、交通需要に関するリスク評価と管理がプランニングの統合部分でなければならないことをデータは示している。上記データは、プランナーがそのようなリスク評価と管理を発見できる、実証的な基礎を与えている。

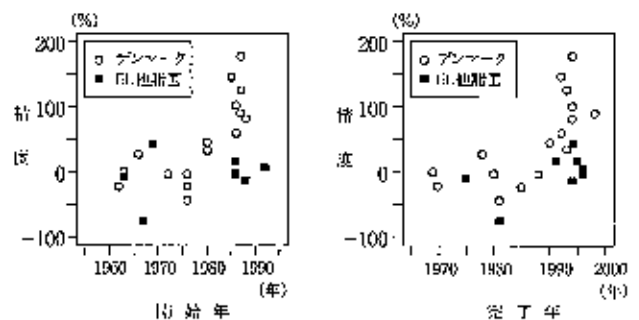
予測は時間の経過につれて精度を増したか？

図一 3 と図一 4 はプロジェクトの建設決定年および完成年かあるいはその一方についての情報と組み合わされた標本において、プロジェクトにとっての予測精度の悪さが時間の経過につれてどう変化しているかを示している。交通量予測が時間の経過につれて精度を増したとする指標はない。道路プロジェクトでは全く逆であり、そこでは予測は期間の終わりほど、精度を悪化させているように見える。統計的な分析がこの印象に確証を与えている。

鉄道プロジェクトでは、予測精度の低さはプロジェクト開始年ともプロジェクト完了年とも独立している。ドイツの 2 つのプロジェクト (図一 3 中の K) が



図一 3 予測精度 (鉄道旅客数)



図一 4 予測精度 (自動車台数)

統計的なはずれ値として扱われるか否かが問題である。鉄道旅客輸送量の予測は時間の経過につれて改善されなかったと結論される。鉄道旅客輸送量は考察された30年間にわたり一貫して過大推計されてきた。

道路プロジェクトでは、期間の終わりほど、交通量が過小に評価されるので、精度はより低くなる。しかし、デンマークと他の国の道路プロジェクトには差がある。デンマークの道路プロジェクトでは、非常に高い統計的有意水準で、精度の悪さは時間経過に伴い変化することが分かった ($p < 0.001$)。1980年以後、デンマークの道路交通量予測は交通量全般の過小推計を伴う誤りに完全に陥った一方、1980年以前のデンマークにも、データのある他の国々にも、これは該当しない。1970年代後半から1980年代後半までの10年間、デンマークの道路交通量予測の精度は悪化し、誤差は3%から55%へ18倍に拡大した(図一五)。

デンマークのプロジェクトでは、建設決定年に対する回帰式は以下である。

$$f = 3.0 + 5.48(Y - 1970)$$

ここで、 f は交通量予測の精度の悪さ(%), Y は建設決定年である。

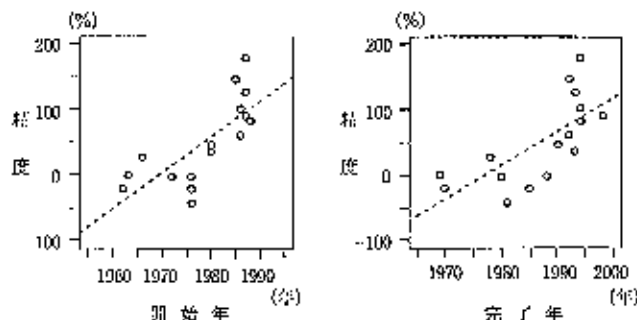
道路交通量予測における精度の悪化の上昇を伴うデンマークの経験は、Ascher (1979, p. 52, pp. 202-203) が「仮説の麻痺」と呼ぶもの、すなわちデータによる妥当性の否定後も仮説の使用が継続されることにより最もよく説明されている。より明確には、交通量予測は概して歴史的データに基づく予測モデルを校正する。1973年と1979年のいわゆるエネルギー危機、それに関連する原油価格上昇プラス実質賃金下落は、一時的にせよ、数十年間ではじめて交通量減少を

伴うという、デンマークの道路交通量に深刻な影響を及ぼした。デンマークの交通量予測担当者は、それらが永続的な傾向を示していたという仮説に対応してモデルを調整し校正した。その仮説が誤りだった。1980年代、二度の石油危機の効果と関連政策手段が徐々に消えていったとき、交通量は再び増加し、1970年代の仮説に基づいてなされた予測を非常に低い精度にした。

交通量予測の精度は、時間につれて改善されなかったと結論される。鉄道旅客数の予測は今日も30年前も同様に不正確であり、つまり水増しされているのである。道路自動車予測は、考察された30年の期間の終わりほど大きな過小推計値をもち、時間の経過に伴いより精度を低くしているようにさえ見える。もし正確な交通量予測に到達する専門的技術と特殊技術が時間につれて改善されたとしても、これはデータには示されていない。これがプランナーに示唆するのは、予測の精度向上に最も有効な方法は、おそらくモデルの改善ではなく、より現実的な仮説、不確実性とリスクの実証的な評価に基づく体系的な利用である。鉄道にとって、特に時間が経過しようともなお高い水増し旅客数予測の執拗な存在は、旅客交通量を過大推計する強い誘因と弱い抑制が、プロジェクト推進者に学ぶべきもの、すなわち過剰推計による旅客数予測が成果をあげると教えてきたかもしれないところで均衡に至ったという憶測を招いている。つまり推計過小な費用との結びつきにおいて、そのような予測は、意思決定者に鉄道プロジェクトの承認と建設を勝ち取るのに役立つ方法で、鉄道プロジェクトについて誤って伝える一助となっている (Flyvbjerg *et al.*, 2003a)。これは、鉄道予測に対する精度の改善が予測の戦略的な不実表記を抑制する強力な説明責任の方法を必要とすることを示唆している。

(以下次号)

(翻訳: 山田 徳彦 白鷗大学経営学部准教授
校閲: 上遠野武司 大東文化大学経済学部教授)



図一五 デンマークのプロジェクトに対する予測精度 (自動車台数)