

分析機器共用実績データを用いた 長期的視点に基づくデータマネジメントの試行

阿部真育¹、江端新吾¹、網塚浩²、上原広充¹、田島さとみ²、吉沢友和²
(¹: 北海道大学 URA ステーション / グローバルファシリティセンター、²: 北海道大学
グローバルファシリティセンター)

1. はじめに

平成18年に閣議決定された第3期科学技術基本計画において、科学技術振興のための基盤強化を目的に先端大型研究設備の整備と共用促進が提言された。その後の第4期並びに第5期科学技術基本計画においても、先端分析機器の共用化は研究基盤強化の柱の一つであり続けている。

施策としての研究基盤強化が成熟に向かいつつある一方、第4期科学技術基本計画以降、大学や研究機関側にはPDCAサイクルの確立といった研究基盤管理システム全体の具体的なマネジメント方法の提言が求められるようになってきている。しかしながら全国的に見て、限られた予算内で機器の維持・購入の優先順位を定量的に求める方法論は、筆者の知るところでは未だ確立されていない。そのような状況の中、北海道大学では先端機器の共用化政策を先進的かつ体系的に進めており、分析機器共用関連のデータ（以下、共用実績データ）の蓄積年数は既に10年を超える。このため、事業の評価や見直し方針の検討を定量的に行える土台が整いつつある。

そこで本研究では、共用実績データマネジメント手法を確立することで、短期的な分析機器メンテナンス方針から長期的な分析機器マネジメント方針までを定量的に整理し、次世代の研究基盤戦略構築への提言を試みる。具体的には時系列分析を主たる分析方法として用いることで、北海道大学が保有している分析機器の利用状況推移を予測も含めて把握し、持続可能な分析機器運用の可能性を示す。

2. ロジックモデル

新行政マネジメント（NPM：New Public Management）理論¹⁾によれば、すべての施策・事業には必ずその活動によってどのような成果が生み出されるのかという論理・道筋の仮説が存在する。また近年はストーリー思考に代表されるように、『現在置かれた状況からどのようなストーリーを立てれば最終的な目標に至るか』を論理的に検討する方法も、様々な分野で着目されている²⁾。

そのような方法論の中で、ロジックモデルはNPM理論を支援する基本的ツールとして定着しており、行財政改革の実践の中で適用されてきた実績を有している。ロジックモデルとは、最終的な成果を設定しそれを実現するために具体的にどのような中間的な成果を必要とし、またそのような成果を得るためにはどのようなデータ収集を行う必要があるのかを体系的に明示するためのツールである（図-1）。欧米諸国ではロジックモデル作成のマニュアルも提案され、幅広く適用されている³⁾。わが国においても、平成13年に『行

政機関が行う政策の評価に関する法律』が施行され、各省庁において、政策評価活動のための基本計画策定にロジックモデルが活用されている例が多く見られる。

ロジックモデルの形式的な特徴としては、1)実際の日常業務から最終的な成果に至るまでの過程を1本もしくは複数の線によって繋げること、2)成果の段階を複数段階に分けて提示すること、が挙げられる。成果の段階とは、図-1に示すような具体的な活動から最終的な成果に至るまでの中間段階にどのようなことが起こりうるかの検討事項に該当する。

ロジックモデルを作成する最大の利点は、事業計画の立案者、実施者、管理者、評価者、利害関係者等の様々な主体が、本格的な政策論争を行い導き出した共通認識を具現化できるところにある。それによりどの立場から見ても見直すべき点が抽出されることとなり、具体的な将来像と現状の課題を常に把握しながら、事業計画の見直しを行うことが可能となる。

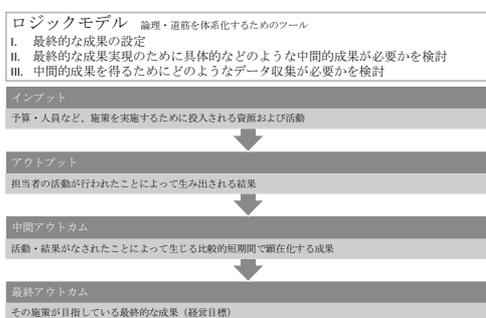


図-1. ロジックモデルの概念

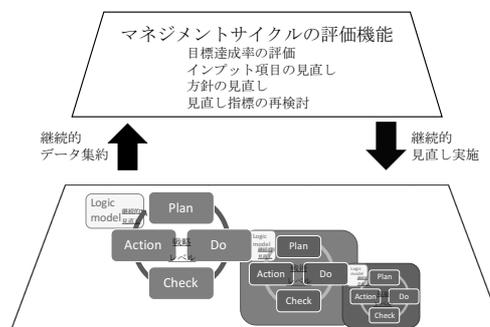


図-2. 階層的 PDCA サイクルとロジックモデル及びマネジメントサイクルの評価機能の位置づけ

3. 継続的事業評価方法としてのロジックモデル

一度立案した維持管理計画を逐次改善していくためには、PDCA サイクルにおける Check と Action を特に実施していくべきである。しかしながら現状では、企画段階のみに対して多くのマンパワーが投入され Check と Action の段階が適切に機能していない状況が非常に多い⁴⁾。

その課題を解決するために、中林ら⁵⁾は高速道路の維持管理項目に対して、ロジックモデルを用いた事後評価手法の提案を行っている。さらに青木ら⁶⁾は維持管理業務によって取得されたデータに対してベンチマーク評価を実施する事により、業務プロセスの要改善箇所を抽出し、改善による維持管理自体のパフォーマンス評価としてロジックモデルを活用する方法論を提案している。

このようにロジックモデルを有効に活用することで運用の成果を継続してモニタリングし、逐次ベンチマーキング評価を行うことで、大学研究基盤全体の効率化並びに先進化に資する業務サイクルが、現場レベル、戦術レベル、戦略レベルのどのレベルにおいても機能することになる(図-2)。

人事的なローテーションが頻繁な組織の場合、システム構築時の方針をノウハウとして効率的に継承することが重要となるため、継続の見直しを行うために客観的な数値指標を設けることが重要である。本研究では、主に『利用人数の記録』の戦略レベルに着目して議論を進める。なお、本来のロジックモデルでは、最終アウトカムは数十年単位の長期目

標と位置付けられることが一般的であるが、本研究においては、既に蓄積されている11年分のデータを用いて時系列分析による将来予測を行い、試験的に評価指標の算出を試みる。

4. 解析方法

ある現象を定式化する分析方法として一般的に、『回帰分析』と『時系列分析』がよく知られている。回帰分析は考えられる要因と未知の要因から定式化を行う。それに対し時系列解析は、ある一定期間（ラグ）前の情報から定式化を行う。共用実績データは、時間経過に伴う状態の変化を表しているため、本研究では時系列分析を用いて分析を行う。

時系列モデルには様々な種類が存在するが、議論の簡便化のため、基本的かつ一般的に用いられている時系列モデルを本研究では考慮する。本研究で考慮する時系列モデルの特徴を表-1にまとめる。なお、ホワイトノイズモデルは現在の事象が過去の事象に全く依存しない現象の場合に用いるモデルであり、時系列モデルによって表現できない事象等をモデルに組み込むために用いられる。蓄積された時系列データを用いて時系列モデルの係数パラメータを推計することで、対象の現象を表現する予測モデルを定式化できる。

モデルの定式化の際には、『データの自己相関』と『データが定常性を満たすかどうか』の確認を行う。データの自己相関は、過去の時点の現象が現時点の現象に対してどの程度影響を与えているかの指標である。一方、データの定常性に関しては一般的に、弱定常性を満たしていれば定常性を満たしていると言われている。現象群を確率変数列と見た時に、①平均が一定、②分散が一定、③自己共分散がラグのみに依存する、といった3つの条件を満たす場合に弱定常性を持つと定義される。

上記定義が存在するため、時系列データが定常性を満たす場合と満たさない場合が存在し、これらは基本的にARMAモデルとARIMAモデルを使い分けることで対応できる。定常性を満たす場合は、ARMAモデルにて分析を行う。ARMAモデルは、現在の事象が過去の事象に影響を受けることを表すAR部分と、現在の事象が過去の事象の分散に影響を受けることを表すMA部分によって成り立つ。時系列データをARMAモデルに当てはめ、最尤法にて各項のパラメータを推計する。時系列データが定常性を満たさない場合においても、一時点前の値との差分を取るなどの処理を行うことで、定常性を満たすようになる場合も存在する。こうしたデータのことを『単位根を持つ』という。時系列データ分析を行うモデルを選択する際には、データが定常性を満たすか、単位根を持つかについて単位根検定を行うことで確認する。単位根を持つ場合は、ARIMAモデルを用いて分析する。ARIMAモデルは単位根を持つデータに対しデータの差分を取ることで、データを定常過程としARMAモデルを当てはめるものであり、最尤法にて各項のパラメータを推計する。

上記の処理過程において、現象を説明する時系列モデルの候補が挙げられることとなり、その中から最適なモデルを選定する作業が必要となる。本研究では、モデル選択の際に一般的に用いられるAIC（Akaike Information Criteria：赤池情報量基準）を用いることとする。AICは統計モデルの良さを評価するための相対的指標であり、『モデルの複雑さと、データとの適合度とのバランスを取る』ために用いられる。AICが最も小さいものを適切なモデルとなる。

表-1. 本研究にて考慮する時系列モデルの概要

モデル	数式	概要説明
ホワイトノイズモデル	$E(e_t) = 0, V(e_t) = \sigma^2 = 1, Cov(e_t, e_{t-s}) = 0$ ここで $e_t \sim$ [任意の確率分布] を, s はラグを示している. 以下, ホワイトノイズは $u(t)$ と表記する.	現在の事象が過去の事象に全く依存しない現象の場合に用いるモデル.
自己回帰モデル (AR モデル)	$y(t) = \mu' + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_{t-p} y(t-p) + u(t)$ ここで μ' と ϕ は係数パラメータである.	現在の事象が前後の時点の事象から影響を受ける場合に用いるモデル.
移動平均モデル (MA モデル)	$y(t) = \mu + u(t) - \theta_1 u(t-1) - \theta_2 u(t-2) - \dots + \theta_q u(t-q)$ ここで μ と θ は係数パラメータである.	現在の事象が過去の事象の分散に影響を受ける場合に用いるモデル. (データを一つずつずらして平均を計算する移動平均とは異なった概念)
自己回帰移動平均モデル (ARMA モデル)	$y(t) = \mu' + \phi_1 y(t-1) + \phi_2 y(t-2) + \dots + \phi_{t-p} y(t-p) + u(t) - \theta_1 u(t-1) - \theta_2 u(t-2) - \dots + \theta_q u(t-q)$ AR の次数が p, MA の次数が q なので ARMA(p,q) と表記する.	AR モデルと MA モデルの両方の対象となる事象を含む現象を説明する場合に用いるモデル.
自己回帰和分移動平均モデル (ARIMA モデル)	$(1 - \sum_{i=1}^p \phi_i L^i)(1-L)^d y(t) = (1 + \sum_{i=1}^q \theta_i L^i) u(t)$ ここで d は階差の値を示す. L はラグオペレータを示しており, $\Delta x_t = (1-L)x_t - x_{t-1}$ のように用いる. AR の次数が p, MA の次数が q, 階差が d なので ARIMA(p,d,q) と表記する.	ある一定時点間の差のデータが ARMA モデルで説明可能な現象の場合に用いるモデル.

5. 結果と考察

共用実績データの 1 つである利用人数に関する時系列データの一例を図-3 に、利用人数に関する共用実績データを用いて予測を行った結果を図-4 に示す。なお、図-4 に示した予測モデルの例では、ARIMA(1, 0, 0) という記載が図の上部にあることより、1 次の AR モデルが最適なモデルとして選定されていることが分かる。

図-3 の共用実績時系列データの例として『フローサイトメーター』を挙げた。蓄積されたデータを図化するだけでも、短期的・中期的・短期的の傾向を定性的に確認することが可能となっていることが分かる。

図-4 は、図-3 に示した学内と学外の利用人数の合計値の時系列データから 1 年、3 年、5 年先の予測を行った結果である。予測結果の赤線とオレンジ色、及び黄色はそれぞれ、予測値と 75% 信頼区間及び 95% 信頼区間を示している。一般的に時系列予測は、サンプルサイズの 10%~20% の期間しか予測値に対する信頼が担保されないとされているが、3 年予測、5 年予測においても信頼区間の幅は 1 年予測の場合と大きく変わらず、長期予測値も戦略検討に利用できる可能性は大きいと言える。

なお、上記の結果を次世代研究基盤戦略に活かすためには、マクロ的な視点による解釈が必要となる。具体的には、継続的に利用が見込まれる機器、近年利用が増加している機器、利用が低下している機器といったトレンドに着目し、今後の新規機器購入やメンテナンス等への投資優先順位を検討するといった戦略や、機器の測定分野別、装置の価格別、設備の設置場所別、学内/学外利用別などによるトレンドの違いを把握することで大学全体としての研究基盤強化に資する検討を行うといった戦略などが挙げられる。本研究では試験的に、機器の共用に供している期間が比較的長い分析装置を対象とし（本研究ではデータ蓄積年数が 11 年の機器を対象とし）、対象の分析装置群を管理する研究所における平均

利用人数の3年予測値トレンドの相関を確認した。この分析により、異なる研究所間（異なる研究分野間）の装置の利用傾向の類似度を比較でき、組織単位の課題対応を行った際に影響が生じる組織を予想することが可能となる。具体的に図-5の例では、A研究所とB研究所の回帰直線は(a-1)のようになり、相関係数は0.00と相関はない。一方、A研究所とC研究所の場合の回帰直線は(b)のようになり、0.49の相関係数を示すように正の相関が見られる。

従って、A研究所とC研究所は、B研究所に比べて同様かつ効果的な機器共用戦略を取ることで、少ない投資で利用人数増加を見込めることが予想される。

6. まとめと今後の課題

本研究において、現場レベルから戦略レベルまでを包括した継続の見直しシステムの一例を、ロジックモデルをメインのツールとして用いることで提案した。広範に及ぶ共用実績データ等をいかにロジックモデルとして体系的に整理できるかは、ロジックモデルの最終目標に関わる組織と密にコミュニケーションを取り続けることが重要と言える。

また、ロジックモデルの評価指標算出のために、時系列分析を用いた共用実績データの活用事例を紹介した。これらの結果は、共用機器に関するデータ蓄積の重要性を提示することが出来たが、以下に挙げる課題も明らかになった。

- 1) 現状蓄積しているデータ項目のみの解析では検討できる内容が限られてしまう
- 2) 全ての機器の補修実績取得は現実的に不可能なため、分析機器の劣化状態の把握は困難である
- 3) 新規導入機器に関してはデータの蓄積が不十分であるため分析対象から除外せざるを得ず、対象選定の段階でバイアスが生じる
- 4) 共用機器と研究業績との紐づけがなされていないため、利用人数が少なくても研究業績を量産しているような装置に対して優遇措置を取る仕組みが組み込めていない

時々刻々と変わる国の施策や研究トレンドを予測しつつ、これらの課題の解決を進めるためには、蓄積すべきデータや用いる統計的手法の改修や変更に関して逐次検討を行う必要がある。これはつまり、見直しのたびに最適な手法の検討が必要となる可能性を意味しているが、ロジックモデルの見直しの際に指標を算出する手法検討がメインの議題になることは避けるべきであり、目標を達成するための手法検討であるという共通認識を関係者間で持つておくことが必要である。

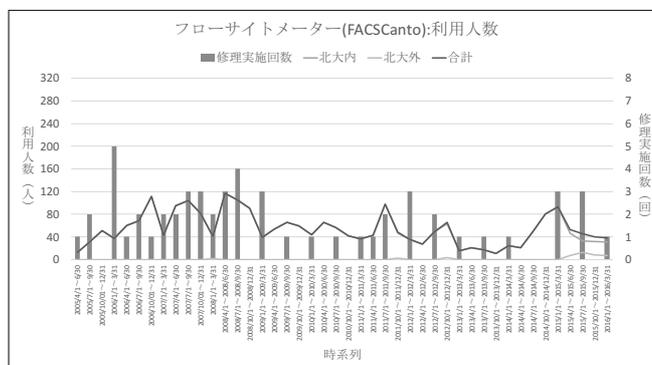


図-3. フローサイトメーターの利用人数時系列データ

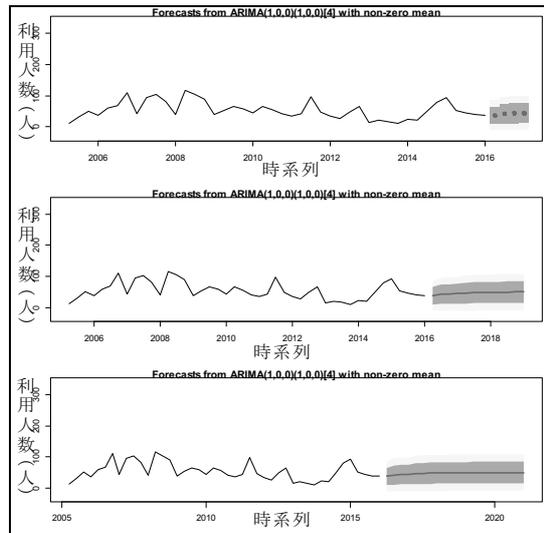


図-4. フローサイトメーターの利用人数の実績 (黒線) と時系列予測結果 (赤線)
 [なお, 上から 1 年, 3 年, 5 年の予測結果であり, オレンジ色, 及び黄色はそれぞれ 75%信頼区間, 及び 95%信頼区間を表す]

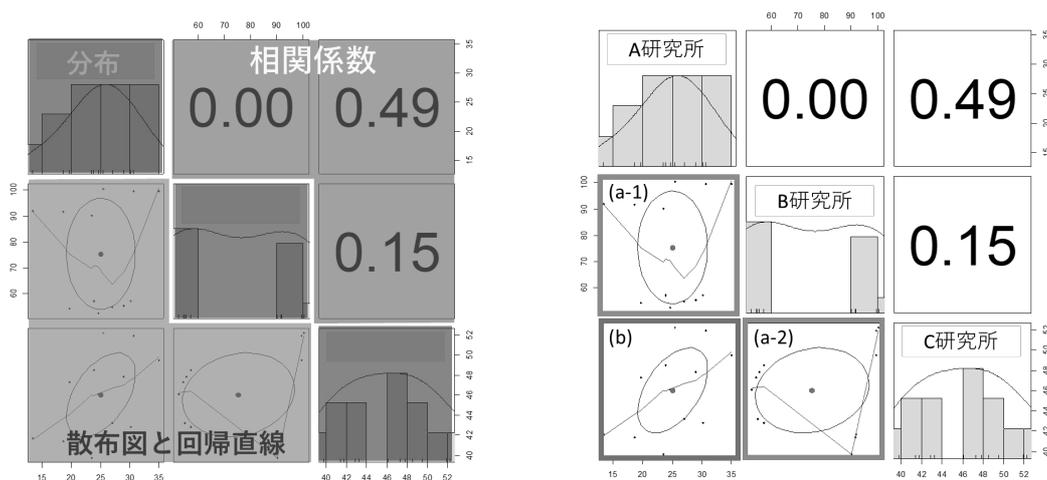


図-5. 利用人数の組織内平均値の 3 年予測トレンドに注目した組織間の予測傾向類似度

【参考文献】

- 1) 大住 荘一郎：ニューパブリック・マネジメント，日本評論社，1999.
- 2) 神田 昌典：ストーリー思考，ダイヤモンド社，2014.
- 3) W.K.Kellogg Foundation: W.K.Kellogg Foundation Evaluation Handbook,1998.
- 4) 松下 哲明, 秀島 栄三：不確実性を考慮した防災事業の予算配分の評価 - 木造住宅の耐震化事業と仮説住宅の備蓄 - , 土木学会論文集 D3, Vol.68, No.5, pp.I_131-I_139, 2012.
- 5) 中林 正司, 西岡 敬治, 小林 潔司：阪神高速道路の維持管理の現状と課題, 土木学会論文集 F, Vol.63, No.4, pp.494-505, 2007.
- 6) 青木 一也, 小田 宏一, 児玉 英二, 貝戸 清之, 小林 潔司：ロジックモデルを用いた舗装長寿命化のベンチマーキング評価, 土木技術者実践論文集, Vol.1, pp.40-52, 2010.