

鉄道分野におけるデータモデリング技術

近年、鉄道分野では旅客データや設備データ、運行データなど日々発生する大量のデータを分析し、利用客へのサービス向上や輸送品質向上に活用する研究が進んでいる。データ間には関連、相関、因果があるが、データ間の因果関係を推定することで、より効果的に異常検出や原因推定が可能となる。本稿では因果関係のモデル開発アプローチと、因果関係の推定技術について紹介する。

矢野 浩仁 | Yano Kojin

鈴木 哲司 | Suzuki Tetsushi

岡田 健一郎 | Okada Kenichiro

王 偉 | Wei Wang

高柳 泰介 | Takayanagi Taisuke

1. はじめに

鉄道分野でのデータ利活用は輸送、営業、保守などでさまざまな活動があるが、近年は特に保守分野での利活用のニーズが高まっている。鉄道分野での保守は、主に設備・車両が対象であるが、基本的にあらかじめ決められた期間、走行キロ以内に点検・交換することで行っている。これまで安全・安定輸送を追求する過程で、過去に発生した故障を分析することにより、点検・交換の周期および対象を決定し、安全を確保してきた。一方で、技術の進歩に伴い各設備機器・車両機器の信頼性が向上していることから、過剰な周期で点検・交換している可能性があり、安全を確保したうえでの保守業務の適正化が求められている。

また、鉄道業界の抱える問題として、設備老朽化に伴

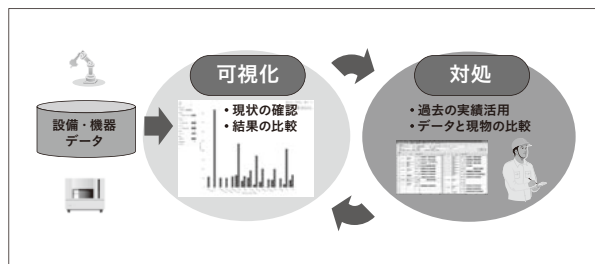
う事故リスクの高まりや、ベテラン社員の大量退職に伴う業務の省力化の観点からも、これまでの保守体系の継続ではなく、状況を踏まえた改善が求められている。近年、IoT（Internet of Things）の進展に伴い、各地上設備や車両の各機器からさまざまなデータが得られるようになってきている。そのため、問題の解決方法としてデータ利活用に基づいた新たな運用が求められるようになってきた。

2. 鉄道分野でのデータ利活用の取り組み

鉄道の設備・車両の保守分野で取得されるデータの多くは、設備・機器の稼働ログとして時系列に記録されたデータである。車両機器の動作データであれば、時刻情報やキロ程、機器の動作回数、動作時間、制御指示内容、制御結果などである。それらで得られるデータはさまざま

図1 | データ活用の用途

データ活用は可視化、対処に分類される。まず可視化を行うことで、対処方法が見えてくる。対処をした結果を集積することで、どのような対処が適切が見えてくる。



主な用途に活用されているが、大きくは可視化、対処に分類することができる。

可視化については、制御指示の内容と回数およびその結果を列車ごともしくはキロ程別に比較することで、特異な動作をしている場合にアラームを出すなどデータから異常を見つけることができる。

対処についても、ある機器の異常動作が一定期間内に5回以上発生した場合に機器交換を行うなど、データを活用することで対処方針を決定するなどの場面で活用されている（図1参照）。

鉄道保守のさらなる効率化のためには、設備や車両の故障予兆検知が必要であるが、これまで安全・安定輸送を追求してきた鉄道においては、故障の発生事例が極めて少ないため、単一の機器の状況のみをモニタリングしてもこれ以上の大きな改善効果を得られない可能性が高い。そこで、データ利活用を推進していくためには、データ間の関係性を突き止め、モデル化することで、通常とは異なる状況を早期に把握することが必要になる。

3. 分析技術の概要と提案アプローチ

こうしたデータ分析においては、統計手法や機械学習が使われることが多い。統計手法としては回帰分析やクラスタリング、因子分析などが挙げられ、機械学習としてはSVM（Support Vector Machine）、DNN（Deep Neural Network）などが挙げられる。

一方でこうした手法で推定したモデルについては、これまでの傾向はつかめているものの、人間にとって解釈することが難しい場合が多い。また実際のデータ分析の際には、データが十分に集まっていない、データ自体にノイズや欠損などが含まれてモデル推定が容易でない場合が多い。

そのため、人間にとって解釈が容易なモデルとして、データ間の統計的な関係性だけでなく、データ間の因果関係を階層的にモデル化するアプローチも用意している。図2は、従来方式のモデル化例と、因果関係のモデル化例を示している。従来方式では、モデルの入力、出力が決まったあとに、その関係性を表すモデルを推定しているが、推定されたモデルはブラックボックスであることが多く、関係性を理解することが容易でない。

一方で因果関係のモデル化では、原因と結果を段階的に表していくため、人間にとって理解しやすいものとなっている。同図の例では、出力した因果関係を見ていくと、部材の損傷確率は、検査員の補修回数が多いと減り、少ないと増えていく。さらに検査員の補修回数は、

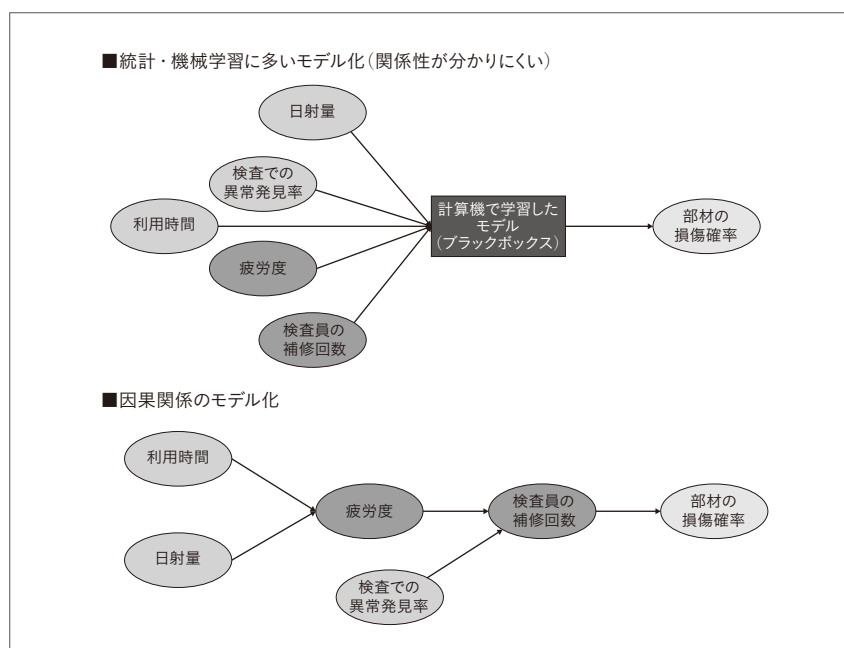


図2 | 従来方式とモデル化の比較

従来方式でのモデル化例と、因果関係のモデル化例を示す。従来方式ではデータの関係性を読み解くのが難しい。

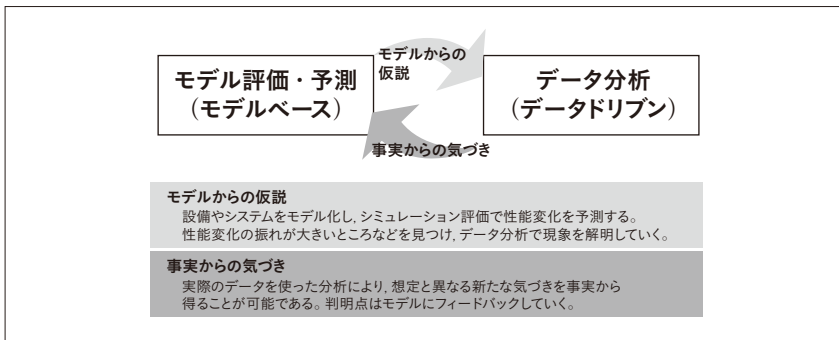


図3| データ分析のアプローチ

データからのモデル化だけでなく、設計知識などのノウハウのモデル組み込みも並行して進め、モデル推定を早く、より正確にしていく。

検査での異常発見率と、そのときの疲労度によって決定される。その疲労度は、部材の利用時間と日射量が関係していることが分かる。このように段階的に表すことで、どこを改善すれば効果的かという推測も容易となる。

またデータ分析を進めていく場合に、データのみからモデルを推定する(データドリブン)手法だけではなく、これまでの設計知識や業務知識を基本的なモデルとして入れ込んで分析する(モデルベース)手法の両方を組み合わせるアプローチを提案している。これにより従来よりも早く、正確なモデル化が期待できる(図3参照)。

このアプローチを効果的に進めるうえでは、データ分析で得られたモデルが人間にとって解釈が容易であることが求められる。因果関係のモデル化は、このアプローチを効果的に回すことが可能になる。具体的にはデータ分析によりデータの因果関係を抽出し、それを分析者が確認することで、モデルの妥当性の判断や、ノウハウをモデル化し因果関係に組み込むことができる。さらに、構築されたモデルに対し、再度データで検証を続けていくことで、より詳細で精度の高いモデル構築が実現できる。

4. データモデリングを支援する因果推定技術

4.1

因果推定技術の概要

上述の因果関係のモデル化を効率よく行うため、データからの因果推定技術の開発を行っている。その概要を図4に示す。

入力情報は、機器を観測するセンサーデータとそのときの検査履歴から構成される。これら入力情報の中から保守の修繕・交換基準となる目的変数を1つ定め、その因果関係の階層構造を自動的に推定する。これにより図2で示したデータ分析において、対象とするデータの

関係性を早くつかむことができ、事実からの気づきを促すことができる。

因果関係は物理的、確率的にさまざまな関係性が考えられるが、本技術開発では、主に土木、機械系の設備の分析を対象とし、物理的な関係性をデータから明らかにすることに主眼を置いている。そのため、ここでは因果関係は多項式で表されるものとし、多項式で構成される説明変数は、そのまま使う場合と、微分・積分を通して使う場合があるものとし、それらの多数の組み合わせの中から有意性の高い因果関係の構造を抽出するようにしている。

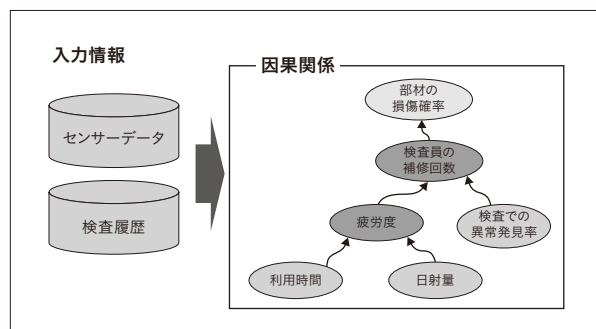
4.2

鉄道保守への応用

得られた因果関係を使った業務イメージについて、鉄道の空調保守を例に説明する。図5はデータからの因果推定を行っている様子を示す。同図左側はデータ選択部分であり、右側は選択されたデータからの因果関係モデルの自動推定結果を示している。データにはノイズ・欠損が含まれている場合や、共線性と呼ばれる同様の動きをするデータが含まれている場合もあるため、出力された因果関係の構造は、保守の担当者でも説明が付かないこともある。こうした場合には、因果の自動推定の途中処理も記憶しておき、因果として採用されなかった第二、

図4| 因果関係の推定アルゴリズムの概要

入力情報の中から目的変数を1つ指定すると、因果関係の階層構造を自動的に推定する。



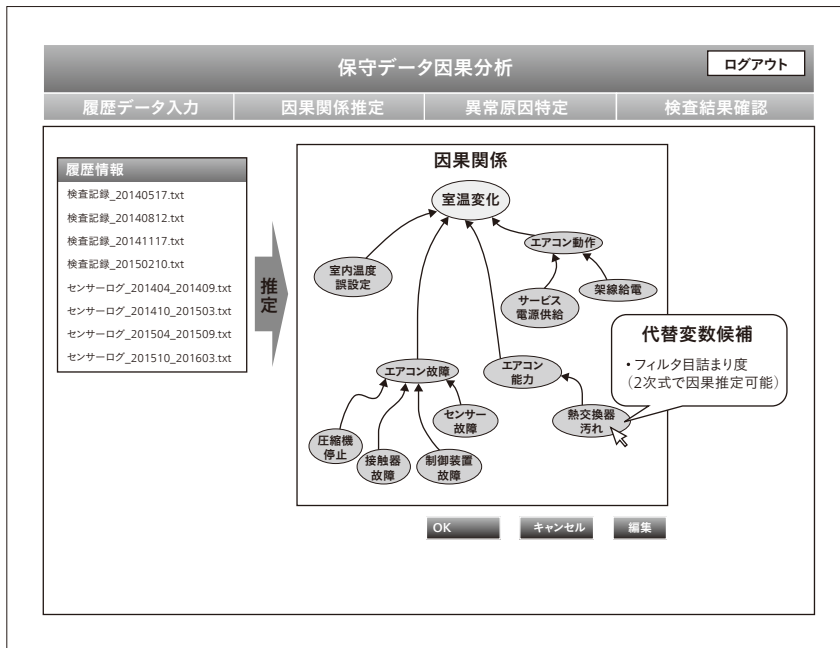


図5 保守での因果関係学習の例

履歴データからの学習の様子を示す。保守の場合、機器のセンサーデータと検査データを基に、機器の不具合の因果関係を推定する。

第三候補を示すことで、説明できる因果関係を構築することが期待できる。同図では一度選択された変数にカーソルを当てることで、採用されなかった第二候補の変数名や、因果関係を表示している。分析者はノウハウに従い変数や因果式を変えていくことで、より精度のよいモデルに直していくことが可能となる。

次に構築した因果関係モデルを使った検査業務の様子を説明する。図6は因果関係モデルを使い、機器の異常時を見つけた様子を示す。

同図の中央のグラフが空調の温度を示しており、実線が実際に観測された空調の温度、点線が因果関係モデルで予測した空調の温度を示している。鉄道車両の室内温

度は、乗車率や外気温の影響を受けるため、室温変化のみでは空調の異常を検出するのは難しい。しかし空調内の圧縮機、熱交換器、電源などの機器の動作の状態から、因果関係モデルによる温度変化の推測が可能である。こうした室温の推測値と、実際の室温を比較することで、温度変化のみでは一見して分かりにくい空調の異常を捉えることが可能となる。

最後に異常箇所推定業務の様子を説明する。図7は因果関係モデルを使った異常箇所の推定の様子を示す。同図右側に因果関係が示されており、これまでの分析で温度変化が異常と判断したときの状態を示している。このとき、因果関係を逆にたどっていくことで、異常箇所の

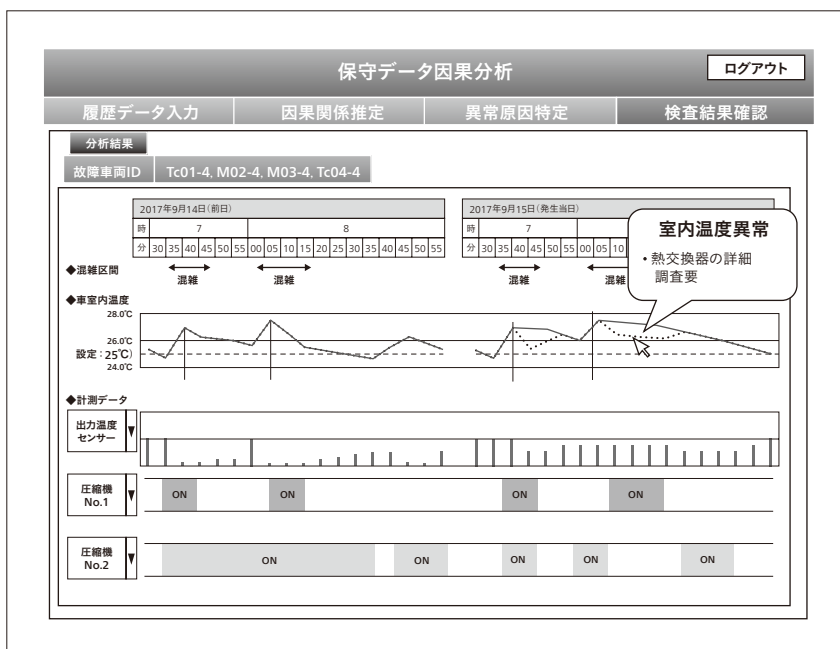


図6 保守での因果関係を用いた異常検知の例

因果関係モデルを使い、異常時を見つけた様子を示す。因果関係によるモデル推定の室温と、実際の温度との乖 (かい) 離を見つけ、それがどこから来るのかを推定する。

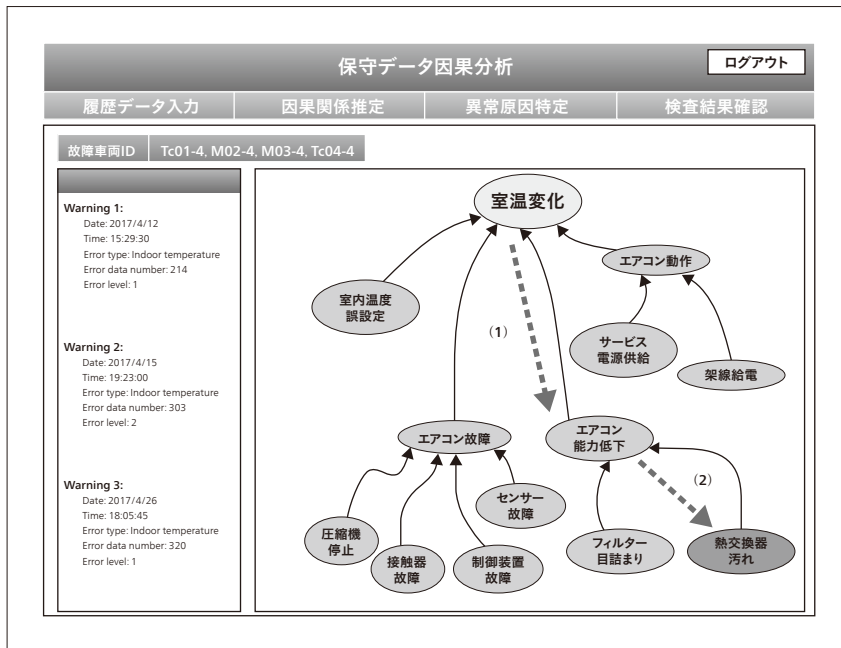


図7 | 保守での因果関係による原因解明の例

因果関係モデルを使い、異常箇所を推定する様子を示す。モデルと実測が違う変数をたどって原因を突き止めていく。

推定が容易になると考えられる。具体的には温度変化の目的変数に因果関係を与えている複数の説明変数に対し、因果関係モデルで推定した説明変数の値と、実際の説明変数の値がずれている説明変数を特定し [同図 (1)]、さらにその説明変数に因果関係のある説明変数の中で、通常と違う値を示している説明変数を特定していく [同図 (2)]。このように異常箇所の推定を支援していくことで、保守業務の工数削減や保守品質の維持が期待できる。

5. おわりに

本稿では、鉄道分野におけるデータ利活用について、特に保守分野についての動向を説明した。従来の単一の機器の状況による判断から、データをより有効に活用していくためには、データの関係性に着目したモデル化が重要であると考えている。これに対し因果関係のモデル化技術の開発を進めており、その技術の適用イメージを紹介した。

今後、実データでの本技術の実証確認事例を積み重ねつつ、鉄道分野へのソリューション提供を行っていく。

参考文献

- 1) 久保俊一：ICTの活用による鉄道メンテナンス技術の革新，第29回鉄道総研講演会（2016.11）
- 2) 森本寛之，外：Hitachi Rail Innovation デジタル技術を活用した鉄道サービスの未来像，日立評論，100，2，182～187（2018.3）
- 3) 長岡晴子，外：ヒトと経営の視点からの顧客価値可視化手法の開発，日立評論，97，11，665～669（2015.11）

執筆者紹介



矢野 浩仁

日立製作所 研究開発グループ
システムイノベーションセンタ 社会システム研究部 所属
現在、交通分野における情報システムおよびシミュレーション研究に従事
博士（工学）
電気学会会員，社会情報学会会員，情報処理学会会員，日本オペレーションズ・リサーチ学会会員



鈴木 哲司

日立製作所 社会ビジネスユニット 社会システム事業部
交通情報システム本部 交通第四システム部 所属
現在、鉄道車両の保守・メンテナンス管理システムの開発に従事



岡田 健一郎

日立製作所 社会ビジネスユニット 社会システム事業部
交通情報システム本部 交通第四システム部 所属
現在、鉄道関連システム全般の開発に従事
プロジェクトマネジメント学会会員



王 偉

日立製作所 研究開発グループ
システムイノベーションセンタ 社会システム研究部 所属
現在、鉄道保守分野におけるデータ利活用技術の研究に従事
博士（理学）



高柳 泰介

日立製作所 研究開発グループ
エネルギーイノベーションセンタ 電磁応用システム研究部 所属
現在、社会インフラのモデリング・シミュレーション研究に従事