

論文・解説

# 17 Neural Network を活用したバーチャルセンサー開発について

## Virtual Sensor Development with Neural Network

小林 謙太<sup>\*1</sup> 矢野 佑樹<sup>\*2</sup> 和田 幸史朗<sup>\*3</sup>  
Kenta Kobayashi Yuki Yano Koshiro Wada  
森実 健一<sup>\*4</sup> 柚木 伸夫<sup>\*5</sup>  
Kenichi Morizane Nobuo Yunoki

### 要約

マツダは従来から取り組んできた MBD プロセスに加えて、機械学習を活用することによるクルマの進化に挑戦している。具体的な取り組みとして、エンジンに搭載されているセンサーが抱える課題を解決するために、物理量を Neural Network モデルで予測することで、センサーと置き換える「バーチャルセンサー開発」がある。本稿ではディーゼルエンジンのインテークマニフォールド温度とターボチャージャー回転数を予測するために、Neural Network モデルを開発するプロセスを紹介する。また、多種多様な条件下で使用されるクルマへの適用を想定し、入力パラメーターが学習領域外に存在する場合の対応についても示す。

### Abstract

Mazda is taking up the challenge to evolve vehicle performance by utilizing machine learning in addition to the MBD process that we have continued to work on. What we specifically have done to solve the issues related to the sensors mounted on the engine is “virtual sensor development” that replaces sensors by predicting physical quantities with a Neural Network model. This paper introduces the process of developing the Neural Network model to predict the intake manifold temperature and the turbocharger rotational frequency of a diesel engine. Additionally, this paper shows how to deal with the situation when the input parameter falls outside the learning domain, assuming practical applications to vehicles that will be used under a wide variety of conditions.

**Key words** : Electronics and control, Engine control, New measuring technique, Neural Network

### 1. はじめに

マツダは「2050年カーボンニュートラルへの挑戦」に向けて、各国の電動化政策や規制強化の動向、お客様のニーズや受容性を踏まえたマルチソリューション戦略を実現するために、パートナー企業とともに段階的にパワートレインの電動化を進めていくことを表明している。電動化に対応した上でお客様に「走る歓び」と「優れた環境性能」を提供し続けるためにはエンジンを含めた複雑なパワートレインシステムの構成や制御の最適化を効率的に行うことが必要不可欠であり、マツダではVプロセスの左バンクにおいて流体や熱流れ、化学反応などの物理現象が数式化されたモデルを用い目指す商品性能を達成するための機能を徹底的に追求する Model Based

Development (以下、MBD) を実践してきた<sup>(1)</sup> (Fig. 1)。

他方で、車載 Electronic Control Unit (以下、車載 ECU) の能力向上により自動車開発における機械学習、特に Neural Network (以下、NN) は、自動運転における画像処理、モデルの代替<sup>(2)(3)</sup>、制御モデルの研究<sup>(3)</sup> 等への活用が進められている。NN は非線形の複雑な現象を高い精度で再現することができ、実時間よりも高速に演算可能である<sup>(4)</sup>。これらの NN の特性に着目し、マツダではこれまで実践してきた従来型の MBD による機能開発に加えて、NN を活用したパワートレインの更なる進化に向けた取り組みを進めている。その一例として、エンジン内部の物理量を NN で予測することで、既存のセンサーを削減する取り組みがある。マツダは電動化が進む 2030 年においても大半のクルマへのエンジン搭載

\*1~3 MBD 革新部  
MBD Innovation Dept.

\*5 PT 制御システム開発部  
PT Control System Development Dept.

\*4 パワートレイン技術開発部  
Powertrain Technology Development Dept.

を想定しているが、エンジンには物理量を計測するために多数のセンサーが搭載されており、コスト高、レイアウトへの制約、重量増加等の幾つかの課題を抱えている。計測された物理量はエンジン制御に反映されるため正確な数値の把握が重要であるが、仮に他の計測された物理量から対象の物理量を予測することでセンサーを削減できれば前述の課題の解決策となり得る。またクルマはさまざまな環境条件下で使用されるため、NN に対する入力パラメーターが事前の学習領域外に存在する状況も考慮すべきである。そこで本稿では、ディーゼルエンジンに搭載されているインテークマニフォールド（以下、インマニ）の温度とターボチャージャー回転数（以下、TC 回転数）を題材として、NN による物理量の予測手法と、入力パラメーターが学習領域外に存在する場合の対応を紹介する。

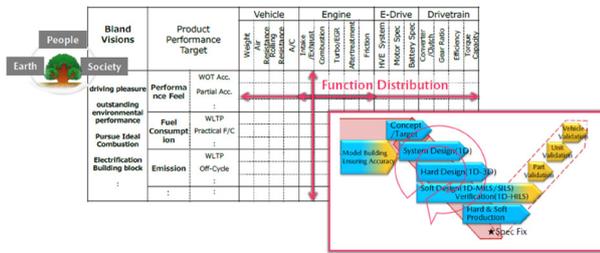


Fig. 1 MBD Process in Mazda

## 2. NN によるセンサー値予測

### 2.1 研究概要

本研究では 1.8L のディーゼルエンジンである SKYACTIV-D 1.8 を題材とした。SKYACTIV-D 1.8 のエンジンシステム図を Fig. 2 に示す。SKYACTIV-D 1.8 の吸排気システムには低圧、高圧の 2 系統で構成される排気ガス再循環装置（以下、EGR）や水冷式インタークーラー、シングル可変ジオメトリー（VG）ターボチャージャーなどが採用されており、インマニ温度や TC 回転数はこれらの部品間で発生するエネルギー収支の結果として決定される。

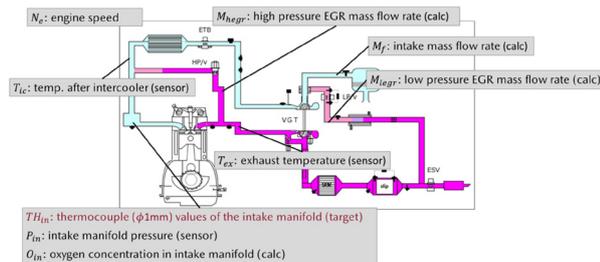


Fig. 2 System Diagram of SKYACTIV-D 1.8

次に NN による物理量予測の工程を Fig. 3 に示す。学習データの収集、入力パラメーターの選定、NN の作成

や学習を実施し、NN の精度や寄与度を分析するプロセスを繰り返すことで精度の高い NN を構築し、最後に開発車両を用いたテストを行うため Simulink モデルに統合した。以降に各工程の詳細について述べる。

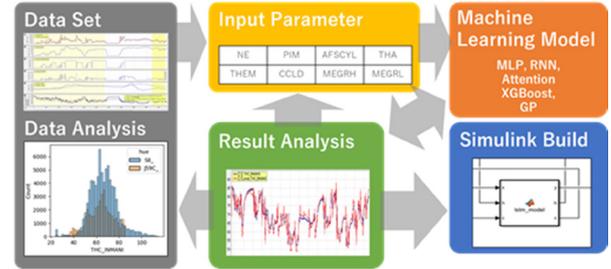


Fig. 3 Procedure of Virtual Sensor Development

### 2.2 学習データ収集

機械学習において重要である学習データの収集には、車載 ECU 内の計測・計算値を取得することが可能な開発車両を用いた。開発車両でマツダ本社内やテストコースを実際に走行し、市街地走行や高速走行を模擬した計 23000 秒分の学習データを収集した。またインマニ温度に関しては市販されている車両に搭載されるサーミスタ型センサーではなく、計測精度や応答性の観点から  $\phi 1\text{mm}$  のシース型熱電対を採用し、こちらも学習データとして活用した。

### 2.3 入力パラメーター選定

NN の入力パラメーターは機械学習の予測精度に大きく影響する。そこでインマニ温度に関係するパラメーターを機能展開図から選択する工学的経験からの手法、インマニ温度を物理式で導く際のパラメーターを選択する数値流体力学からの手法、共分散等のデータ分析からインマニ温度に相関のあるパラメーターを選択する統計的な手法の三手法を組み合わせることで最も高い予測精度となる入力パラメーターを選定した。予測対象を現在のインマニ温度の熱電対値として、入力パラメーターはセンサー値と ECU 内の計算値からエンジン回転数、インマニ圧力、吸気流量、インタークーラー後ガス温度、タービン前排気温度、インマニ内酸素濃度、高圧 EGR 流量、低圧 EGR 流量を設定した。TC 回転数についても同様の手法により入力パラメーターの選定を実施した。

### 2.4 検証環境と NN モデル

机上検証は Python により実施しており、NN の作成と学習には代表的な機械学習のフレームワークと NN ライブラリである Tensorflow と Keras を使用した。また、開発車両の走行中に 0.1 秒間隔で ECU 適合ツールと車両間の通信をしてリアルタイム予測を行える環境も Simulink 上で構築した。

インマニ温度や TC 回転数を求めるためには過去から

現在までのエネルギー収支を計算する必要があり、NNはRecurrent Neural Network (以下、RNN)の採用が一般的である。ただし、RNNは長期の時系列データを処理する際の勾配消失という問題を抱えているため、今回は記憶セルの導入によって長期の過去情報を記憶可能なLong Short Term Memory (以下、LSTM)を用いた (Fig. 4)。今回のLSTMはユニット数が10で活性化関数をtanhとsigmoidとしたLSTM層が一層のモデルを使用している。また、出力層は活性化関数を恒等写像とした全結合層である。学習の設定は誤差関数として平均二乗誤差を、最適化アルゴリズムとしてAdam<sup>(5)</sup>を採用した。

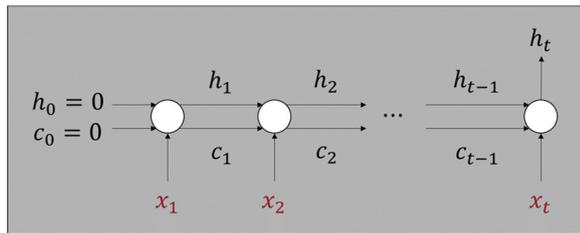


Fig. 4 Structure of LSTM

2.5 結果

学習領域内の入力データによってインマニ温度とTC回転数を予測した結果を Fig. 5, 6 に示す。ここでNeはエンジン回転数、TH<sub>in</sub>はインマニ温度、TCはTC回転数である。実測値は実線で、予測値は破線で表している。Fig. 5は145km/hまで加速する開発車両の走行データを基に机上予測した結果であり、実測と一致していることが確認できる。また、Fig. 6は開発車両の走行中にリアルタイム予測を145km/hまで複数回加速する走行において実施した結果である。開発車両によるリアルタイム予測においても遅延なく演算できていることに加え、予測精度の許容誤差であるインマニ温度±5℃以内、TC回転数±10000rpm以内を全予測期間のうち97%の領域で達成している。

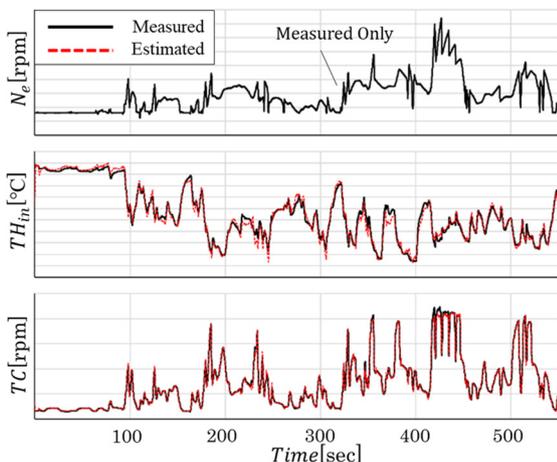


Fig. 5 Prediction for Training Data by LSTM

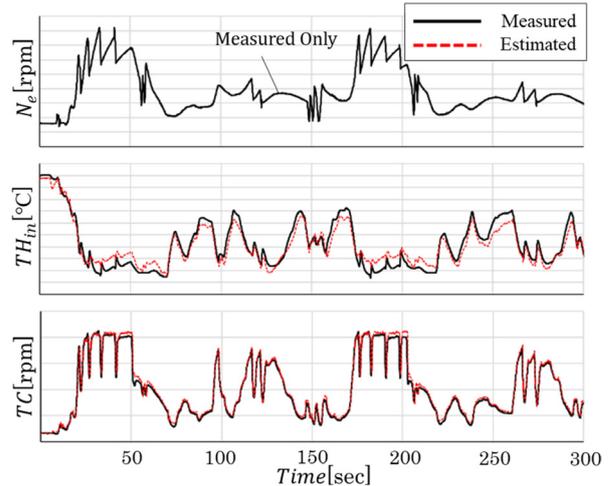


Fig. 6 Real-Time Prediction by LSTM

次に学習領域外における予測精度を検証するために、学習データに高車速走行を加えずに、高車速走行を含む走行パターンを予測した机上検証結果を Fig. 7 に示す。実測の傾向はとらえているものの高車速走行時は絶対誤差が大きくなっている。一般論として、NNは学習領域外において予測精度が低下する傾向が見られる。しかし、実際の市場では多種多様な走行パターンや環境条件が存在し、全ての学習データを事前に収集することは困難である。そこで次項では、学習領域外における予測精度の信頼性を保証するための対応について紹介する。

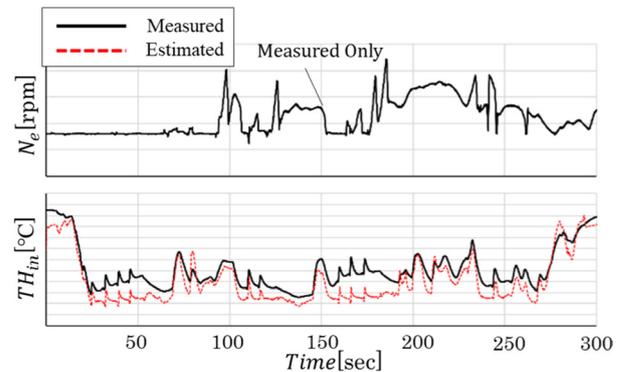


Fig. 7 Prediction for Outside Learning Domain by LSTM

3. バーチャルセンサーの外挿対策

3.1 外挿対策の必要性

NNが学習領域外において精度が悪化する問題を Fig. 8 で説明する。青点はNNの学習領域内のデータ、赤点は学習領域外のデータを表している。実線で示すように、学習領域内のデータに対する予測(内挿)はよく合うが、学習領域外のデータに対する予測(外挿)は信頼性が失われる。これはNNが学習領域内のデータのパターンを学習し、その高い表現力で近似する一方で、学習領域外のデータのパターンは学習時に把握できないためである。

NNの学習時に市場で想定される極冷間や高地といった環境条件下で収集したデータを使用して、可能な限り広い範囲のデータを学習領域内に含めるという方法はあるが、データ収集に莫大な工数とコストがかかる上、全ての条件下で内挿とするのは困難である。そこでNNの外挿を補間するために、以下の手順をとる。

- 1) 高精度予測が可能な NN モデルの学習
- 2) NN の寄与度分析による入力パラメータ絞り込み
- 3) 絞り込み入力パラメータによる重回帰予測
- 4) 学習領域外での重回帰予測への切り替え

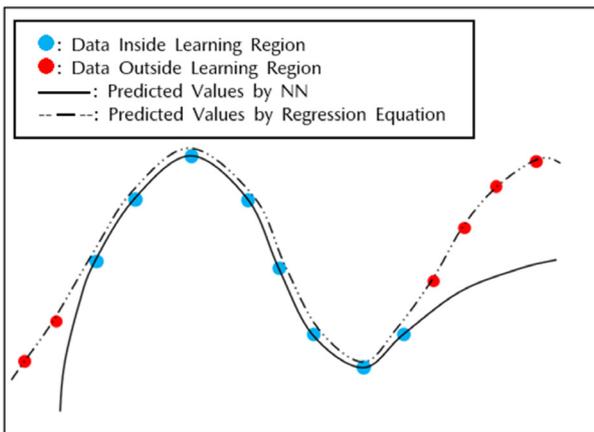


Fig. 8 Schematic Diagram of a Prediction

1) では実車走行で種々の物理センサーから得られたデータ及び2章で構築したNNモデル構造と入力パラメータを活用しNNに学習させる。

2) では予測に対する寄与度の小さい入力パラメータを削除する、特徴量選択を実施する。機械学習における特徴量選択の主な効用には、ノイズの削減、過学習の防止、計算量の削減、モデルの解釈可能性の向上などがある。特徴量選択の手法はさまざまなものがあるが、今回は既に精度よく予測できるNNがあるので、SHAP<sup>(6)</sup>とデータの相関、分布を使って、NNから高精度予測に必要な入力パラメータを絞り込む。ここでSHAPとは機械学習モデルを局所的に線形モデルで近似することで、各入力パラメータの予測への寄与度をデータごとに定量化するツールである。

3) では2)で絞り込んだ入力パラメータを使ってロバストな予測数式を重回帰により立式する。そのような数式は前提をみれば外挿でも精度を確保できる。そのアイデアをFig.8の模式図で説明する。Fig.8ではNNの学習領域内のデータは、sin関数のような数式で近似できる。このとき数式は学習領域内のデータの法則を表現している。この法則が学習領域外でも成り立つ前提がみれば、NN予測の信頼性が失われる学習領域外のデータに対しても、点線のような信頼できる予測結果を出力できる。

4) は予測モデルを実車に搭載する際の方針である。学習領域内での予測精度は基本的にNNが最も優れているため、走行時に学習領域外のデータが入力された場合のみ、予測モデルをNNから3)で構築した重回帰式に切り替え、外挿の信頼性を担保する。

本研究では、高地における性能変化が大きいTC回転数の予測に対して、以上の手法による大気圧に対する外挿対策を適用する。大気圧は高度上昇に対して指数関数的に減少していく法則があり、特に本研究で使用するデータの範囲内では線形に近い形で減少していくため、3)で言及した法則の前提をみている。

### 3.2 TC 回転数の外挿対策

#### (1) 高精度予測が可能な NN モデルの学習

まずNNモデルとして、1層10ユニットのLSTMに、出力層として1ユニットの全結合層を加えたものを用意した。入力パラメータはTC回転数に物理的に関与するであろうものを(1)式のように12個選んだ。

$$\hat{x} = (M_{air}, P_{atm}, P_{in}, P_{ex}, T_{ic}, T_{bic}, T_{ex}, T_w, L_{tc}, Q_{fuel}, M_f, O_{ex}) \quad (1)$$

ここで  $M_{air}$  は新気流量,  $P_{atm}$  は大気圧,  $P_{in}$  はインマニ圧力,  $P_{ex}$  はタービン前圧力,  $T_{ic}$  はインタークーラー後ガス温度,  $T_{bic}$  はインタークーラー前ガス温度,  $T_{ex}$  はタービン前ガス温度,  $T_w$  はエンジン水温,  $L_{tc}$  はVGタービンのベーン開度,  $Q_{fuel}$  は燃料噴射量,  $M_f$  はインマニガス流量,  $O_{ex}$  は排気酸素濃度である。

学習には標高0m, 1000m, 1500m, 2000m, 2500m, 3000mに相当する大気圧下の開発車両走行データを使用した。データのサンプリングレートは0.1秒とし、NN学習時に使用するデータのタイムステップ数は100とした。全学習データ点の合計個数は約13万点だった。Fig.9にモデルの模式図を示す。

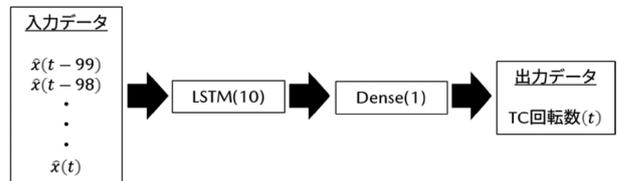


Fig. 9 Schematic Diagram of the Model

Fig.10に標高0mにおけるTC回転数の予測結果を示す。横軸が時間、縦軸がTC回転数で、黒い実線がTC回転数の実測値、赤い点線がNNモデルによるTC回転数の予測値を表す。学習領域内であれば高い予測精度であることが分かる。

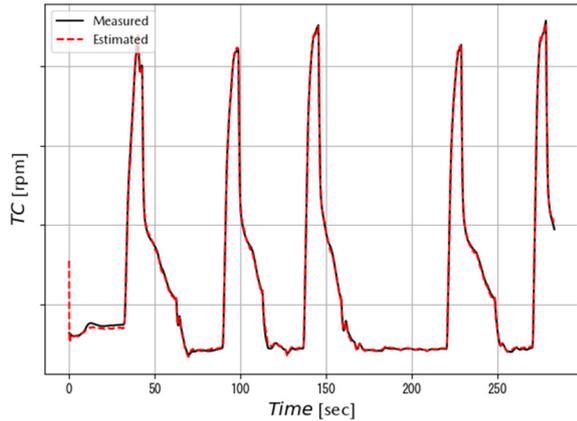


Fig. 10 TC Prediction Results from the NN Model with 12 Inputs

(2) NNの寄与度分析による入力パラメーター絞り込み  
次に重回帰式を立式するために、入力パラメーターの絞り込みを行った。まず学習したNNモデルに対してSHAPによる寄与度分析を適用した。先述したようにSHAPは各入力パラメーターの予測への寄与度をデータごとに定量化するツールであり、今回のような入力パラメーターの絞り込みにも適用可能である。Fig. 11に標高0m相当の大気圧下におけるデータにSHAPを適用した結果を示す。

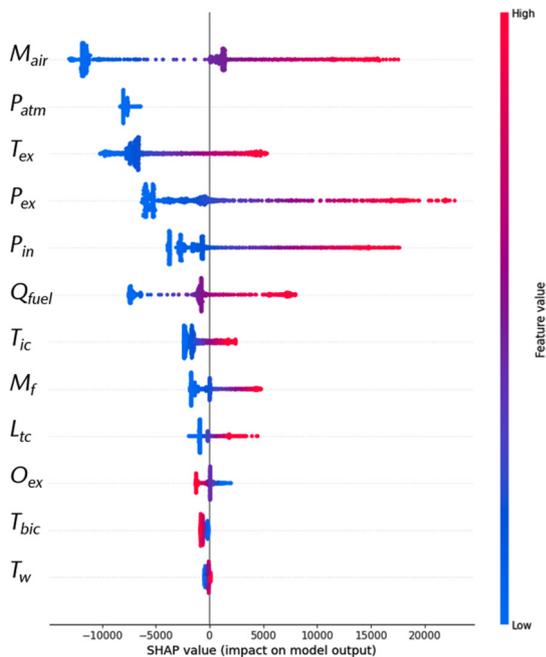


Fig. 11 SHAP Summary Plot under 0m Atmosphere Pressure

図は各タイムステップの各入力（図中の1点が1つのデータに対応する）が、TC回転数予測に与える寄与を表している。例えばタービン前圧力  $P_{ex}$  の行では、赤いplotが図の横軸 SHAP value が1万から2万の間に分布しているのが見てとれる。これはタービン前圧力が大きいと

きに、TC回転数の予測が平均値より1万から2万回転程度上昇することを意味する。実際に今回解析したエンジンでのタービン前圧力とTC回転数の間には強い正の相関があるので、NNは入力パラメーターを正しく考慮して予測を行っていると解釈できる。

ただしSHAPの結果はモデルの入力パラメーター間の相関に強い影響を受ける。これは線形回帰モデルが多重共線性に強い影響を受けるのと同様であり、個々のパラメーターが独立でない場合、SHAPの出力をそのままモデルの予測に対する寄与度と解釈することはできない。したがって実際にはデータの入力パラメーターと出力パラメーターの分布や、パラメーター間の相関関係に注意してSHAPの結果を解釈する必要がある。

例えば Fig. 11 では新気流量  $M_{air}$  の寄与度が大きいですが、 $M_{air}$  はタービン前圧力  $P_{ex}$  と強い相関をもち、かつTC回転数に対するデータ分布の形状は  $P_{ex}$  の方が素性がよいため、 $M_{air}$  は入力パラメーターから削除できる。

以上のようなSHAP、相関、データ分布を合わせた解析により、TC回転数予測に最も寄与するパラメーターを以下の4つに絞り込んだ。

$$\hat{x} = (P_{ex}, P_{atm}, O_{ex}, L_{tc}) \tag{2}$$

実際にこれらのパラメーターを入力として3.2(1)と同様の構成のNNモデルを学習させ、Fig. 10と全く同じ学習領域内のデータのTC回転数を予測した結果をFig. 12に示す。4個のパラメーターによって高い精度で予測できている。

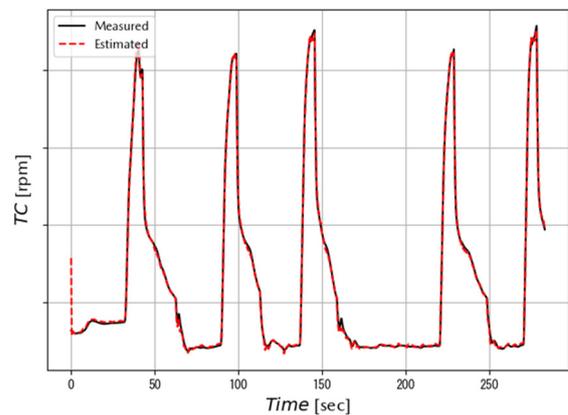


Fig. 12 TC Prediction Results from the NN Model with 4 Inputs

(3) 絞り込み入力パラメーターによる重回帰予測

以下の解析にはPythonの機械学習ライブラリであるscikit-learnを使用した。標高0m~2000mのデータを学習に使い、2500m~3000mのデータを大気圧の外挿検証に用いた。絞り込んだ(2)式の4個のパラメーターを入力とし、TC回転数を予測する重回帰モデルを作成した。具体的にはL2正則化を含むリッジ回帰モデルを採用

した。フィッティングの際にはパラメーターのデータ分布などの考察から、4個のパラメーターを3次(交互作用項含む)まで展開した。結果的に重回帰式は定数項を含めて合計35項となった。式(3)に求めた重回帰式の一部を示す。ただし式(3)の両辺の各項は標準化している。

$$TC \text{ 回転数} = 1.33 P_{ex} - 1.28 P_{ex}^2 * P_{atm} + \dots \quad (3)$$

#### (4) 学習領域外での重回帰予測

式(3)によって学習領域内である標高0mと領域外である標高3000mのTC回転数を予測した結果をFig. 13とFig. 14に示す。Fig. 13の内挿の精度と同様に、Fig. 14で示す外挿の精度も高いことが分かる。

一方数式ではなくNNで、学習領域外の標高3000mでのTC回転数を予測した結果をFig. 15に表す。高回転側でTC回転数の予測と実測の誤差が大きくなっており、学習領域外ではFig. 14で示した(3)式の予測結果の方が高い精度であることが分かる。実際にFig. 14の重回帰式による外挿は、TC回転数±10000rpm以内の要求精度を94%のデータで達成できている。一方Fig. 15のNNによる外挿では、高回転側の精度が低いため、同要求精度を79%のデータでしか達成できていない。以上の結果から、外挿の精度悪化をNN予測から数式予測に切り替えることで補間する、今回採用した手法は有効であるといえる。

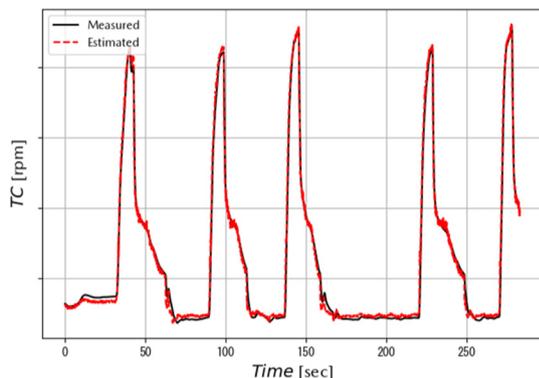


Fig. 13 Interpolate TC Prediction Results from the Multiple Regression Equation

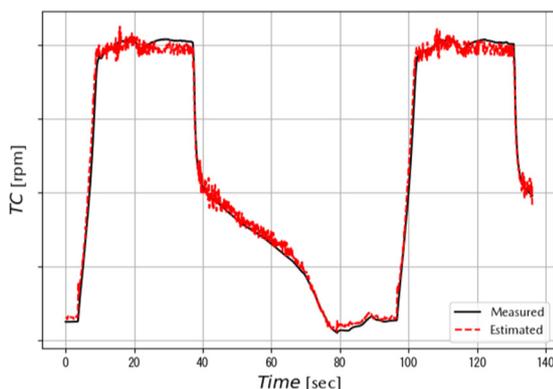


Fig. 14 Extrapolate TC Prediction Results from the Multiple Regression Equation

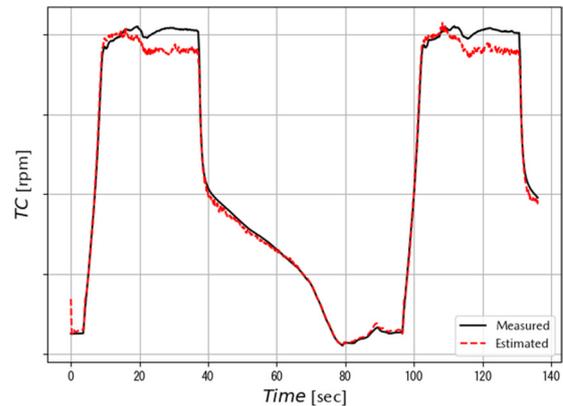


Fig. 15 Extrapolate TC Prediction Results from the NN Model

## 4. おわりに

本研究ではディーゼルエンジンのインマニ温度とTC回転数に対して、NNを活用することによって精度良く、かつリアルタイムに予測できることを示した。また市場におけるさまざまな使用環境下を想定し、入力パラメーターが学習領域外に存在する場合に重回帰式による予測に切り替える手法を提案した。これらの技術は他の物理量に対しても応用可能である。今後はコスト低減、レイアウト自由度の向上、車両重量の削減などのバーチャルセンサー化によって得られる価値を考慮し、対象センサーの選定を実施する予定である。またバーチャルセンサーの実用化に向けて、高精度なCAEなどを用いたより広範囲な学習データの生成技術や、実際のECU上での検証、2章で紹介したインマニ温度予測に対する外挿対策の適用も併せて進めていく。

## 参考文献

- (1) 宮崎ほか：SKYACTIV-D 3.3の開発におけるMBDプロセス革新と適用, [マツダ技報, No.39, pp.43-50 \(2022\)](#)
- (2) 池戸ほか：機械学習によるモデリング手法を活用した後処理状態推定 第2報, 第33回内燃機関シンポジウム予稿集, (2022)
- (3) 森安ほか：機械学習によるディーゼルエンジン吸排気系の実時間MPC設計, 計測自動制御学会論文集, Vol.3, No.3, pp.172-180 (2019)
- (4) 和田ほか：Neural Networkモデルによるセンサー値予測, 自動車技術会春季大会前刷集, No.227 (2023)
- (5) Diederik P. Kingma, Jimmy Ba: Adam: A method for stochastic optimization, In International Conference for Learning Presentations, (2015)
- (6) Scott M. Lundberg, Su-In Lee: A Unified Approach to Interpreting Model Predictions, (2017)

■著者■



小林 謙太



矢野 佑樹



和田 幸史朗



森実 健一



柚木 伸夫