

東京大学 大学院新領域創成科学研究科  
基盤科学研究系  
先端エネルギー工学専攻

令和元年度

修士論文

データ駆動型アプローチによる水レジストジェットスラス  
タにおける運用時推力および欠損データの推定方法

A Method Estimating Thrust and Missing Data by Data  
Driven Approach When Utilizing a Water Resistojet  
Thruster

2020年1月28日提出  
指導教員 小泉 宏之 准教授

47186061 王 啓航

---

---

# 目次

---

---

第1章	緒言.....	10
1.1	機械学習の衛星利用.....	10
1.1.1	超小型衛星の登場.....	10
1.1.2	衛星領域における機械学習の発展.....	10
1.1.3	データ駆動型の機械学習手法.....	11
1.1.4	データ駆動型機械学習の応用例.....	11
1.2	問題点と研究目的.....	13
1.2.1	問題点.....	13
1.2.2	研究目的.....	14
第2章	データ取得方法.....	15
2.1	対象とする推進システム.....	15
2.1.1	背景.....	15
2.1.2	作動原理.....	15
2.1.3	構造.....	17
2.2	取得したデータの試験内容.....	18
2.3	実験装置.....	19
2.3.1	真空チェンバー.....	19
2.3.2	推力測定装置.....	25
2.3.3	AQUARIUSにおけるセンサー.....	28
2.4	データセット.....	28
第3章	機械学習による解析手法.....	31
3.1	概要.....	31
3.2	機械学習モデル.....	33
3.2.1	主成分分析(Principal Component Analysis PCA).....	33
3.2.2	K-means.....	34
3.2.3	Synthetic Minority Over-sampling Technique (SMOTE).....	35
3.2.4	回帰木と決定木.....	36
3.2.5	ランダムフォレスト.....	38
3.2.6	LightGBM.....	40

3.3	特徴量エンジニアリング	43
3.3.1	ラグ特徴量	43
3.3.2	時間積分値特徴量	44
3.3.3	Rolling Window による特徴量	44
3.3.4	他の特徴量	45
3.3.5	推進システムのモデルから得られる特徴量	45
3.4	機械学習を用いた推力推定	47
3.4.1	K-means でクラス分け	47
3.4.2	LightGBM を用いた推力推定	49
3.4.3	ランダムフォレストを用いた推力推定	53
3.5	機械学習を用いた気化室圧力推定	58
3.5.1	蒸発モード、非蒸発モードの判断	58
3.5.2	気化室圧力の推定	59
第4章	解析結果	61
4.1	推力推定解析の結果	61
4.1.1	K-means のクラスタリングの結果を主成分分析で可視化	61
4.1.2	LightGBM とランダムフォレストを用いた推力推定	67
4.1.3	特徴量の重要度	70
4.2	気化室圧力の推定解析結果	70
4.2.1	蒸発モードと非蒸発モードの判断	70
4.2.2	LightGBM とランダムフォレストを用いた気化室圧力データの推定結果	71
4.2.3	特徴量の重要度	72
第5章	考察	73
5.1	推力推定の考察	73
5.1.1	推力推定において SMOTE の有効性評価	73
5.1.2	推定精度評価	73
5.1.3	特徴量評価	74
5.2	気化室圧力データ推定の考察	76
5.2.1	蒸発モード判断への考察	76
5.2.2	液滴蒸発の物理モデルとの比較する	76
5.2.3	特徴量評価	80
第6章	結論	81

---



---

## 記号表

---

$A$	:	較正係数
$c_m$	:	$m$ の葉ノードにおいて推力の平均値
$C_t$	:	気化室とノズルの間のコンダクタンス
$f$	:	木モデル
$f_N$	:	$N$ 回目のブースティングで作ったモデル
$f_{\text{MSE}}(R)$	:	$R$ ノード内の平均二乗誤差
$FTR$	:	真陽率
$F_{\text{Thrust}}$	:	推力
$I(x \in R_k)$	:	入力 $x$ が $R_k$ 内にある時に 1, そうでない時に 0
$I_{p_{vc}}$	:	気化室圧力の重要度
$J(\mu_k, q_{ik})$	:	類似性評価関数
$L_c$	:	決定木用の評価関数
$L_R$	:	回帰木用の評価関数
$m$	:	水の質量
$\dot{m}$	:	一サイクルの平均流量
$M(\mu_k)$	:	$k$ 番目のクラス
$N_j(R)$	:	$R$ ノードのうち $j$ クラスに属するデータの数
$N(R)$	:	$R$ ノードのデータの数
$p_i$	:	$i$ 番目の主成分分析用の座標変換ベクトル
$p_R$	:	右ノードに属する確率
$p_L$	:	左ノードに属する確率
$P_{vc}$	:	気化室圧力
$P_{sat}$	:	飽和水蒸気圧力
$P_n$	:	ノズル前圧力
$\Delta P$	:	気化室とノズル前の圧力損失
$q_{ik}$	:	$x_i$ が $M(\mu_k)$ に属しているかどうかを表す帰属変数
$R$	:	気体定数
$R_{node}$	:	木のノード
$R_R$	:	分割後右側のノード
$R_L$	:	分割後左側のノード
$Re^*$	:	スロートレイノルズ数
$s(X)$	:	$X$ 特徴量の標準偏差
$s$	:	選定される特徴量
$t$	:	レギュレーションバルブオープン後の経過時間
$T_{tube}$	:	チューブ温度
$T_s$	:	水の表面温度
$T_n$	:	ノズル温度
$T_{VC}$	:	気化室温度
$TPR$	:	偽陽率
$\bar{x}$	:	$X$ 特徴量の平均値
$x_t$	:	スラストスタンドの変位
$x_i$	:	$i$ 番目 $X$ 特徴量の値
$x_{std_i}$	:	$i$ 番目の標準化された $X$ 特徴量

$y_i$	:	$i$ 番目目の真値
$\hat{y}_i$	:	$i$ 番目目の推定値
$y_i^N$	:	$N$ 回目のブースティングで入力データ $\mathbf{x}_i$ に対応する推定するターゲット
$\alpha$	:	学習率
$\gamma$	:	L1 正則化の係数
$\boldsymbol{\mu}_k$	:	$k$ 番目のクラスに属するデータの代表ベクトル
$\mu^*$	:	粘性度
$\sigma$	:	蒸発係数
$\lambda$	:	L2 正則化の係数
$\omega_j$	:	$f$ の各葉ノードの平均値
$\rho_L$	:	水の密度

---



---

## 目次

---

Figure 1-1 Abnormality detection in data driven approach method proposed by Yairi et al. <sup>19</sup>	12
Figure 2-1 Simple model for AQUARIUS concept.	15
Figure 2-2 An example of the operation history of AQUARIUS, the green line shows the pressure inside the vaporization chamber, and the black line shows the thrust.	16
Figure 2-3 AQUARIUS system structure, red letters are abbreviations summarized in Table 2-2 <sup>22</sup> .	18
Figure 2-4 Experimental system for thrust measurement chamber.	21
Figure 2-5 External view of chamber 1 for unit operation test.	21
Figure 2-6 Rotary pump	22
Figure 2-7 Turbo molecular pump	22
Figure 2-8 Cryopump	23
Figure 2-9 Compressor	23
Figure 2-10 Chamber evacuation system diagram for satellite integrated propulsion test.	24
Figure 2-11 System diagram of thrust measurement system.	25
Figure 2-12 Thrust stand appearance	26
Figure 2-13 Example of time history of calibration test.	27
Figure 2-14 The relationship between the displacement and the thrust calculated from the displacement and the mass of the weight is approximated by a straight line, and the slope is used as a calibration coefficient.	27
Figure 3-1 Separate input features and output features.	31
Figure 3-2 Separate train dataset and test dataset.	32
Figure 3-3 Orange triangles are data samples with few classes. The triangle inside the red frame is one of the randomly selected data.	35
Figure 3-4 Find N samples of the same class around each randomly selected data.	35
Figure 3-5 Green small triangles are samples artificially created between two data of the same class.	36
Figure 3-6 Example of visualization diagram of regression tree and decision tree.	37
Figure 3-7 Example of random forest of decision tree.	39
Figure 3-8 Level -Wise, Green nodes will grow in the next step, red nodes will not grow in the next step.	42
Figure 3-9 Leaf -Wise, green nodes will grow in the next step, red nodes will not grow in the next step.	42
Feature 3-10 Rolling Window	44

Figure 3-11 LightGBM learning algorithm without feature engineering and without SMOTE.	51
Figure 3-12 LightGBM learning algorithm with feature engineering and without SMOTE. ....	52
Figure 3-13 LightGBM learning algorithm with feature engineering and SMOTE. ....	53
Figure 3-14 Random forest learning algorithm without feature engineering and without SMOTE. ....	55
Figure 3-15 Random forest learning algorithm with feature engineering and SMOTE. ....	56
Figure 3-16 Random forest learning algorithm with feature engineering and without SMOTE.	57
Figure 4-1 Result of K-means classification class in thrust estimation for DVT1. ....	61
Figure 4-2 Result of K-means classification class in thrust estimation for DVT2. ....	61
Figure 4-3 In DVT1's case, the absolute value taken for each component of the eigenvector of the first principal component, and the top 10 features and the absolute value of the component. ....	62
Figure 4-4 In DVT1's case, the absolute value taken for each component of the eigenvector of the second principal component, and the top 10 features and the absolute value of the component. ....	62
Figure 4-5 In DVT1's case, the absolute value taken for each component of the eigenvector of the third principal component, and the top 10 features and the absolute value of the component. ....	63
Figure 4-6 In DVT2's case, the absolute value taken for each component of the eigenvector of the first principal component, and the top 10 features and the absolute value of the component. ....	63
Figure 4-7 In DVT1's case, the absolute value taken for each component of the eigenvector of the second principal component, and the top 10 features and the absolute value of the component. ....	63
Figure 4-8 In DVT1's case, the absolute value taken for each component of the eigenvector of the third principal component, and the top 10 features and the absolute value of the component. ....	64
Figure 4-9 Visualization of clustering results by K-means for DVT1. The horizontal and vertical axes are the first and second principal components. ....	64
Figure 4-4-10 Visualization of clustering results by K-means for DVT1. The horizontal and vertical axes are the first and third principal components. ....	65
Figure 4-11 Visualization of clustering results by K-means for DVT1. The horizontal and vertical axes are the second and third principal components. ....	65
Figure 4-12 Visualization of clustering results by K-means for DVT2. The horizontal and vertical axes are the first and second principal components. ....	66

Figure 4-13 Visualization of clustering results by K-means for DVT2. The horizontal and vertical axes are the first and third principal components.....	66
Figure 4-14 Visualization of clustering results by K-means for DVT2. The horizontal and vertical axes are the first and third principal components.....	67
Figure 4-15 Thrust error evaluation at each time. The vertical axis shows RMSE. Error bars indicate standard deviation of RMSE. ....	69
Figure 4-16 RMSE of average thrust of each injection cycle and standard deviation of RMSE. ....	69
Figure 4-17 RMSE of average thrust of each injection cycle and standard deviation of RMSE. ....	72
Figure 5-1 Two steps evaporation.....	75
Figure 5-2 RMSE histogram of measured and estimated values of vaporization chamber pressure when using injection data of pattern 1 and DVT1. The curve means the kernel estimation curve.....	77
Figure 5-3 RMSE histogram of measured and estimated values of vaporization chamber pressure when using injection data of pattern 1 and DVT2. The curve means the kernel estimation curve.....	78
Figure 5-4 RMSE histogram of measured and estimated values of vaporization chamber pressure when using injection data of pattern 2 and DVT1. The curve means the kernel estimation curve.....	78
Figure 5-5 RMSE histogram of measured and estimated values of vaporization chamber pressure when using injection data of pattern 2 and DVT2. The curve means the kernel estimation curve.....	79
Figure 5-6 Error comparison between physical model estimates and data-driven approaches...	80



---



---

## 表目次

---

Table 2-1 Explanation of each AQUARIUS parts. ....	17
Table 2-2 Explanation of AQUARIUS ground test experiment.....	19
Table 2-3 Specifications of rotary pump. ....	20
Table 2-4 Specifications of turbo molecular pump. ....	20
Table 2-5 Specifications of cryopump. ....	20
Table 2-6 Specifications of rotary pump used for integration test. ....	24
Table 2-7 Specifications of mechanical booster pump used for integration test. ....	24
Table 2-8 Specifications of oil diffusion pump used for integration test. ....	24
Table 2-9 Specifications of LED displacement meter. ....	25
Table 2-10 Results of weight mass measurement. ....	26
Table 2-11 Specifications of pressure sensors in water tank and bladder. ....	28
Table 2-12 Specifications of vaporization chamber pressure sensors and nozzle pressure sensors.....	28
Table 2-13 Platinum thermometer for tubes, nozzles, vaporization chambers.....	28
Table 2-14 Explanatory table for data set features.....	29
Table 2-15 Range of each feature value in the experiment using DVT1. ....	29
Table 2-16 Range of each feature value in the experiment using DVT1. ....	30
Table 3-1 Hyper-parameter explanation table of random forest. ....	40
Table 3-2 Hyperparameter description table used for LightGbm.....	42
Table 3-3 Description table of lag features. ....	43
Table 3-4 Description table of time integral value features. ....	44
Table 3-5 Description table of features that created by Rolling Window.....	45
Table 3-6 Description table of other features. ....	45
Table 3-7 Features based on physics.....	46
Table 3-8 Number of injections of test data and learning data.....	47
Table 3-9 Input features used for thrust estimation.....	47
Table 3-10 Cases of thrust estimation by Random forest.....	49
Table 3-11 Initial value of LightGBM hyper parameter.....	49
Table 3-12 LightGBM Hyperparameter Grid Search Range and Priority.....	50
Table 3-13 Hyperparameter grid search range for random forest. ....	54
Table 3-14 Features used for mode judgment.....	58
Table 3-15 Definition of True positive $TP$ , False positive $FP$ , False negative $FN$ , True negative $TN$ .....	59

Table 3-16 Cases of vaporization chamber pressure estimation. ....	59
Table 3-17 Features used to estimate vaporization chamber pressure (VPS1). ....	60
Table 4-1 Thrust estimation cases. ....	67
Table 4-2 LightGBM hyperparameter results in DVT1 thrust estimation selected by cross-validation and grid search. ....	68
Table 4-3 LightGBM hyperparameter results in DVT2 thrust estimation selected by cross-validation and grid search. ....	68
Table 4-4 Randomforest hyperparameter results in DVT1 thrust estimation selected by cross-validation and grid search. ....	68
Table 4-5 Randomforest hyperparameter results in DVT2 thrust estimation selected by cross-validation and grid search. ....	68
Table 4-6 Most important feature in thrust estimation. ....	70
Table 4-7 Evaluation result of evaporation mode judgment. ....	70
Table 4-8 LightGBM hyperparameter results in DVT1 vaporization chamber pressure estimation selected by cross-validation and grid search. ....	71
Table 4-9 LightGBM hyperparameter results in DVT2 vaporization chamber pressure estimation selected by cross-validation and grid search. ....	71
Table 4-10 Pressure estimation cases. ....	71
Table 4-11 Most important feature in pressure estimation. ....	72

---

# 第1章 緒言

---

## 1.1 機械学習の衛星利用

### 1.1.1 超小型衛星の登場

第二次世界大戦終結後、米ソはハイテクノロジーの領域において、冷戦中に激しい軍事競争を実施した。特に総合工学である宇宙開発は互いに国の総合力をアピールできるため、冷戦中ハイスピードで発展した。1961年から1972年の間、アメリカは合計6回有人月面着陸に成功した<sup>1</sup>。ソビエト連邦は世界初の人工衛星スプートニク1号<sup>2</sup>、宇宙ステーションサリュート1号<sup>3</sup>、火星着陸マルス3号<sup>4</sup>などの実績を残した。日本、中国や欧州の各国も積極的に宇宙開発の活動に取り組んでいた。このような背景で、通信衛星、地球探査衛星、科学衛星など様々な衛星が次々に登場し、我々の生活も画期的に変わった。例えば、GPS衛星の登場により、交通の利便性が大幅に上がった。気象衛星の登場により、天気予測ができるようになった。このような衛星は大型や中型衛星が多数であり、開発費用は1基あたり数百億円が必要である。このため、衛星業界では民間企業や大学参入の壁が高い。しかし、2000年前後、超小型衛星の登場により、宇宙業界全体がパラダイムシフトを迎えた。超小型衛星は大型衛星と比べると、体積が小さい、開発費用が安い、開発期間が短いなどのメリットがあって、民間企業や大学の注目を集めている。このため、近年、超小型衛星の打ち上げ数も年々増加している<sup>5</sup>。各国において、多くの民間企業が超小型衛星市場に参入しており、今後5年間で2600基の増加が予測されている<sup>6,7,8,9</sup>。

### 1.1.2 衛星領域における機械学習の発展

超小型衛星の登場により、保守的な宇宙業界に、新規技術を積極的に活用する流れができてきた。代表例としてISSに使用されている新型Li電池やイオンエンジンに利用されているGaNトランジスタなどが挙げられる<sup>10,11</sup>。それらに加えて、近年では特に人工知能の技術である機械学習の利用が注目されている。

人工知能の研究は前世紀の1950年代から3回のブームを迎えた。1950年代後半から、コンピューターで特定の問題に対してある程度「探査と推論」が可能になった。このような背景で人工知能技術は1回目のブームを迎えた。しかし、その時の人工知能技術は現実問題の複雑な問題を解くことができないため、冬の時代に入った。1980年代から、人工知能は2回目のブームを迎えた。様々な分野の専門家は「知識」をコンピューターが認識できる形で記述し、人工知能は知識表現できるようになった。しかし、世の中の全ての知識をコンピューターが理解できるように記述することが難しいため、2回目のブームは再び

冬の時代に入った。2000年代から3回目の人工知能ブームを迎えた。ディープラーニングなどの機械学習手法により大量のデータを用いることで、人工知能が特定の知識のパターンを抽出できるようになった。機械学習の技術により、画像認識、自然言語処理、異常検知など特定の領域において人間より上回ったため、工業分野で利用されている<sup>12</sup>。この技術は人工衛星にも活用されている。膨大な画像データの解析が必要な、農業、航行、防災など様々な領域では画像認識技術が応用されている。また、データ駆動型機械学習手法は衛星の異常検知や自立制御への応用が期待されている。

### 1.1.3 データ駆動型の機械学習手法

データ駆動機械学習の方法は大きく分けると教師あり学習、教師なし学習2種類に分類できる。教師あり学習は回帰予測したい或いは分類したいラベル特徴量のデータが既知の場合、他の特徴量からラベル特徴量を予測する機械学習方法である。教師なし学習はデータをクラスタリングする或いは次元圧縮する方法であり、ラベル特徴量が必要ない<sup>13</sup>。

### 1.1.4 データ駆動型機械学習の応用例

衛星は多くのデータを取得するため、データ駆動型機械学習との相性が優れている。衛星の地球観測、衛星の異常検知、衛星制御など様々なところで応用されている。ここで、いくつかの応用例を取り上げる。

#### ➤ 機械学習の地球観測への応用

地球観測衛星は我々の生活と密につながっている。災害の監視、気象観測、農業利用など、衛星で様々なところで応用されている。機械学習は画像認識のような分類問題が得意であり、このため、機械学習と衛星の画像認識との親和性が優れている。

吉原らは、リモートセンシング取得した衛星画像を用いて、教師あり学習である Convolutional Neural Networks (CNN)を用いて災害地の被覆に対する分類を研究した。結果、既存の手法より高精度な分類が実現された<sup>14</sup>。Xiaowei Jia らは、衛星画像を機械学習に用いて、農業のマッピングを正確でタイムリーな農業情報を得た<sup>15</sup>。Yong Hoon Kim らは COCI 衛星データを使って、ランダムフォレスト、Cubist, SVR などの手法を使って、海水の水質を予測した<sup>16</sup>。結果、すべての予測結果が決定係数 0.9 より上回った。木モデルのアンサンブル学習を利用することで、予測用の特徴量の重要度を解明した。重要度の結果はほかの研究先行研究の結果と非常に一致している。

➤ 機械学習の衛星異常検知への応用

衛星の異常検知に関して、従来は単変量の閾値から異常検知していたため誤報が多い。また、閾値の設定には各コンポーネントの深い理解と知識が要求された。一方で、機械学習を使ったデータ駆動型アプローチでは多変量かつ時系列の異常検知により高精度を実現し、加えて閾値の設定に各コンポーネントの知識はほぼ必要ない。

Lachlan Gunn らは、LSTM (Long Short Term Memory) を RNN に応用し、多変量かつ時系列の異常検知を実現できて、単変量の閾値で判断する方法を改善した<sup>17</sup>。Shahroz Tariq らは LSTM と確率 PCA の混合異常検知を提案し、韓国の KOMPSAT-2 に応用し、評価した<sup>18</sup>。矢入らは JAXA の小型衛星 SDS-4 のデータを用いて、クラスタリングと次元削減の機械学習技術による衛星のモニタリング有効性を示した<sup>19</sup>。図 1-1 に示したように、この方法はいくつかのステップに分ける。最初は衛星からテレメトリーデータを地上に送り、データの標準化などの前処理を行う。そして、訓練とテストの二つの段階に分ける。訓練段階では過去の正常データからモデルを訓練する。そして、テスト段階では訓練されたモデルを新しく下したデータに適応する。その後、異常かどうかの情報を出力する。

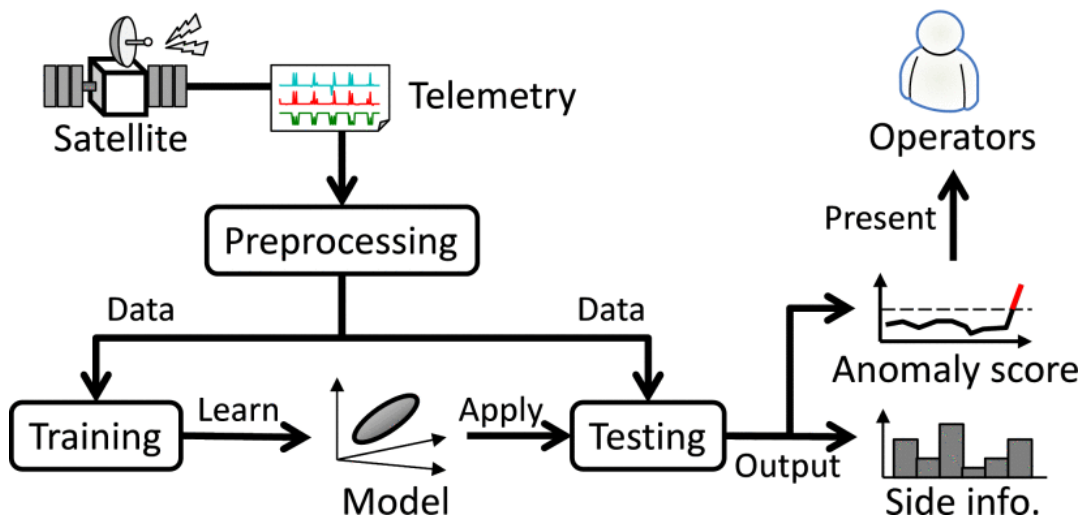


Figure 1-1 Abnormality detection in data driven approach method proposed by Yairi et al.<sup>19</sup>

➤ 機械学習の衛星制御への応用

データ駆動型機械学習は衛星制御へ適用できる。東京工業大学の松永らの研究グループは予め地図の情報を衛星に搭載し、衛星軌道から陸上を撮影し、撮影した画像情報から3軸姿勢を推定する計測技術を実衛星で実証した<sup>20</sup>。他に、ホールスラスタの制御にも適用できる。ホールスラスタは比推力が高いため、注目されている。磁場形状と磁場強さホールスラスタの推進システムの性能に大きく影響する。損耗により、磁場形状と磁場強さは変わっていくため、推力性能が変化する。淵上らはニューラルネットワークを使って、電

流変化を予測することができた<sup>21</sup>。計算コストが高いためこの予測プログラムを衛星に実装することができないが、小型計算機性能向上により、今後ホールスラストの電流変化を予測した上で、変化に合わせた推進システムの制御が期待されている。

## 1.2 問題点と研究目的

### 1.2.1 問題点

本研究では超小型衛星課題の二つの問題点に注目する。一つ目は一部モデル化が不十分な推進システムに対して、物理モデルに基づいた高精度なリアルタイム推力推定が難しいこと。二つ目は限られたスペースの衛星の中で少ない冗長系の実現が難しいこと。これらの問題はデータ駆動型の機械学習で解決できる。本節ではこれらの問題点について詳しく説明する。

超小型衛星が実現できるミッションの幅を広げるために、推進システムを実装することが必要である。姿勢制御を実現するために、リアクションホイールで衛星の角運動量を吸収する方法はよく使われる。しかし、長時間作動の作動ではリアクションホイールの角運動量が飽和する。そうすると、姿勢制御ができなくなり、衛星の運動量が飽和する前にリアクションホイールの角運動量の解放が必要となる。また、深宇宙探査のミッションで軌道維持や軌道遷移が必要となるケースが多いため、並進運動を実現するための推進機を超小型衛星に実装することが必要となる。そして、超小型衛星の生産コストが低い点はフォーメーションフライトの親和性と非常に優れており、フォーメーションフライトを実現するには推進システムが不可欠である。一方、超小型衛星はセカンドペイロードで相乗りのケースが多い。推進剤の漏れによる主衛星への危害を避けるために、超小型衛星の推進システムは厳しい安全審査が要求される。体積が小さく、推進システムを搭載することができないため、小型高性能な推進システムの開発が求められている。このため、様々な推進システムが次々と開発されている。しかし、これらの推進システムでは、推力に関する理論がまだ完全に解明されていないものもある。衛星運用時の軌道遷移において、スラストの推力を知ることがミッションに関わっている。推力を知る手段としては理論計算と観測の二つの手段が考えられる。計算の場合、物理モデルからの演繹的計算が考えられる。この場合に、センサーが取得したデータを物理モデルに代入することで、推力を計算することができる。例えば、イオンスラストはイオン飽和電流が分かると、推力を理論計算できる。ラバールノズルを使うコールドガスジェットはラバールノズルの式によって、圧力のデータから理論計算ができる。しかし、一部の推進システムでは理論がまだ完全に確立されていない。例えば粘性、希薄流体を扱う常温水レジストジェットスラストは非粘性流体を使うため、ラバールノズルの推力理論と乖離し、理論計算の結果と実値と一致する結果を得ることが難しい。推力の観測はいくつかの手段が考えられる。一つ目は、加速度セン

サーのデータから推力を計算すること。しかし、推力が低い推進システムに対して、加速度センサーの分解能制限により、加速度を得ることが難しい。二つ目は、リアクションホイールから推力を算出することある。しかし、リアクションホイールからリアクションコントロールシステム (RCS) の推力を逆算出することができるが、並進運動による角運動量の変化から推力を得ることができない。他の方法としては、ドップラーシフトから衛星のインパルスビットを計算できる。しかし、この方法は推力の積分値を得ることができるが、直接にリアルタイムの推力を得ることができない。このため、一部の既存の理論から離れる推進システムに対して、理論から直接にリアルタイムの推力を高精度で計算することが難しい。これらの推進システムに対して、正確なリアルタイム推力推定の実現が問題になっている。

深宇宙探査、フォーメーションフライトなどのミッションを実行する際に、推進システムが重要な役割を果たしている。推進システムの特徴に合わせて、制御と異常検知用に必要なコンポーネントにセンサーを設置する。宇宙の過酷な環境で、軌道上で衛星の電気回路の異常により、データの欠落、観測センサーの破損、性能低下など発生することがある。対策としては、高い信頼性のセンサーと複数の冗長系を設置することが挙げられる。しかし、超小型衛星は電力とスペースの制限があるため、数多くの冗長系観測センサーを設けることができない。

## 1.2.2 研究目的

水レジストジェットスラスタのような物理モデルがまだ完全に確立されていない推進システムに対して、推力測定などのテストデータを活用することで、データ駆動型アプローチによって、帰納的にリアルタイムの推力と故障したセンサーデータの値を推定することができる。また、リアルタイムでセンサーの値を推定し、冗長系として使える同時に、推定値の信頼区間を検討することで、簡単な異常検知を実現することができる。

本研究は水レジストジェットスラスタの地上推力測定試験データを使用し、データ駆動型アプローチによる推力及び欠損データを推定する手法を提案し、適用する際の予測精度を確認することである。

---

## 第2章 データ取得方法

---

### 2.1 対象とする推進システム

ラバールノズルの推力理論から乖離するスラスタとしては、常温蒸発の気液分離の水レジストジェットシステムがある。その一例として、本研究は東京大学が開発した水を推進剤とするレジストジェットスラスタシステム AQUARIUS の地上試験データを解析データとして使用する<sup>22,23</sup>。本節はAQUARIUSの背景、作動原理、構成について紹介する。

#### 2.1.1 背景

AQUARIUSの開発背景について説明する。水は毒性がないため、安全な推進剤として使える。また、水は常温常圧で液体貯蔵可能なので、高圧タンクが不要となる。このため、水は推進剤として超小型衛星の推進システムとの親和性が優れている。特に電力を用いて水を加熱し、それをノズルから噴射する水レジストジェットは、構造がシンプルで、比推力はコールドガスジェットより高い。このため、これまでに多く水レジストジェット推進システムが研究された。水レジストジェットの作動原理は、水を蒸発させ、水蒸気をノズルから噴射し推力を得ることである。従来の水レジストジェットは水の凝縮に対する対策が不十分であり、確実に気液分離を保障することができない。一方、AQUARIUSは気化室を設けることで、水がノズルから噴射するまでに完全な気液分離を保障できる。

#### 2.1.2 作動原理

AQUARIUSの作動原理について説明する。簡単のために、AQUARIUSの簡易モデルをFigure 2-1にまとめた。Figure 2-1のように、タンク、気化室、ノズルの三つのコンポーネントがある。タンク内のブラダに水を貯蔵している。気化室、ノズルにヒーターを設置しており、水を加熱することで凝縮を防ぐことができる。気化室の隣に通信機を設置しており、通信機の排熱を利用することで気化室を温めるための電力を節約することができる。

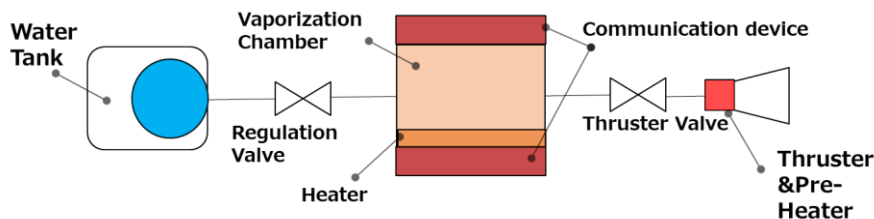


Figure 2-1 Simple model for AQUARIUS concept.



Figure 2-2 はこの推進システムの作動履歴を示している。作動時は次のような三つの段階に分けることができる。

- 1) Thruster Valve 運用時オープン：Thruster Valve をオープンし、ノズルから気化室までの間が真空状態になる。
- 2) Regulation Valve オープン：水タンクと気化室間の Regulation Valve を瞬時オープンし、事前に決めた噴射量の水を気化室に噴射し、供給することができる。本研究で、この「供給」を「供給噴射」と呼ぶ。この後、気化室内の水が蒸発し、水蒸気はノズルから外に噴射される。本論文ではこの時の状態を蒸発モードと呼ぶ。
- 3) 待機：水は潜熱が大きいいため、気化室は水の潜熱により温度が低下する。水が完全に蒸発し切ったら、気化室の温度が既定の温度までに回復することを待つ必要がある。本論文ではこの時の状態を非蒸発モードと呼ぶ。また、AQUARIUS では気化室圧力と気化室温度を監視することで、次の Regulation Valve オープン時間を決める。

最初の1)で気化室を真空引きした後、2) 3) の過程を繰り返すことで、AQUARIUS は脈動式噴射で作動する。Figure 2-2 は履歴の例を示している。

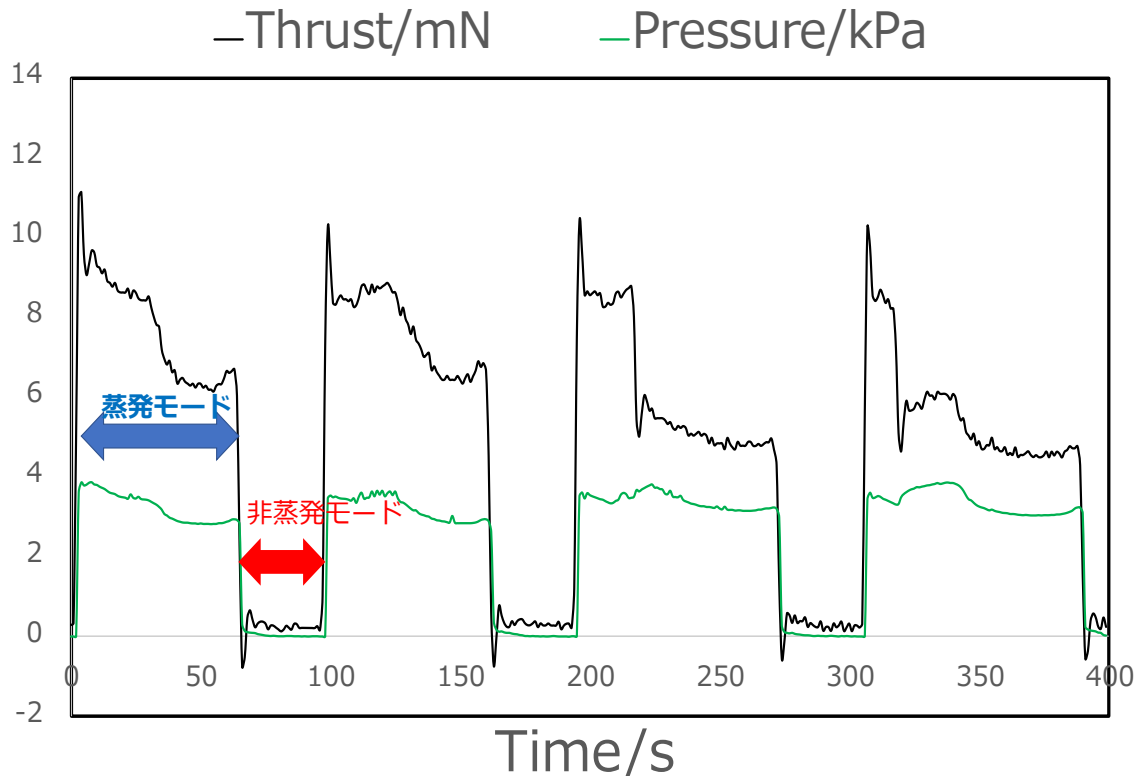


Figure 2-2 An example of the operation history of AQUARIUS, the green line shows the pressure inside the vaporization chamber, and the black line shows the thrust.

## 2.1.3 構造

この推進システムでは多くのセンサーを設置している。水が蒸発し切ったかどうかを判定するには気化室圧力を監視するために二つの圧力センサーが気化室に設置されている。気化室温度は水の飽和水蒸気圧力と関係しているため、気化室の温度を制御する必要がある。AQUARIUSはBangBang制御で決まったレンジの温度をヒーターのON-OFFで制御している。温度を監視するために、二つの温度センサーが気化室に設置されている。ノズルで水の凝縮を防ぐために、ノズルにヒーターと温度計を設置していて、ノズルの温度を監視しながらヒーターのON-OFFをコントロールする。また、これらのセンサーを設置することで、推進システムが異常かどうかを監視することができる。Figure 2-3はAQUARIUSの構造を示している。Table 2-1は各パーツの説明を示している。

Table 2-1 Explanation of each AQUARIUS parts.

パーツ名	略語	説明
Water Tank		水タンク，押しガスとブラダを内部に入れている。
Water Bladder		水ブラダ，内部に水を入れている。
Filter		バルブに不純物の混入を防ぐために使うフィルターである。
$\Delta V$ Thruster valve	DTV	DVT と VC の間のバルブ。
Delta-V Thruster	DVT	並進運動用のスラスタ，合計二つある。それぞれ DVT1, 2 とする。
Gas Drain Valve	GDV	タンク内のガスを抜きたり充填したりする時に使うバルブである。
Nozzle P sensor	NPS	ノズル前の圧力を測るために使うセンサーである。DVT1,2 のセンサーをそれぞれ DPS1,DPS2 とする。
RCS Thruster	RCT	リアクションコントロールシステム用のスラスタ，合計四つある。それぞれ RCT1, 2, 3, 4 とする。
RCS Thruster Valve	RTV	RCS のノズルと気化室の間のバルブ，使うノズルを選ぶ時にそれに対応するバルブをオープンする。
Regulation Valve	RV	タンクと気化室の間のバルブである。
Tank Pressure sensor	TPS	タンクの圧力センサー，タンク内の押しガスの圧力を測るセンサーである。
Vaporization Chamber	VC	気化室，水の気液分離用。

VC Drain Valve	VDV	気化室内のガスを排気するために使うバルブである.
VC P sensor	VPS	気化室の圧力を測るために使う圧力センサーである.
Water Grain Valve	WDV	タンク内の水を充填する際に使うバルブである.
Water Pressure sensor	WPS	ブラダ内の水の圧力を把握するために使うセンサーである.

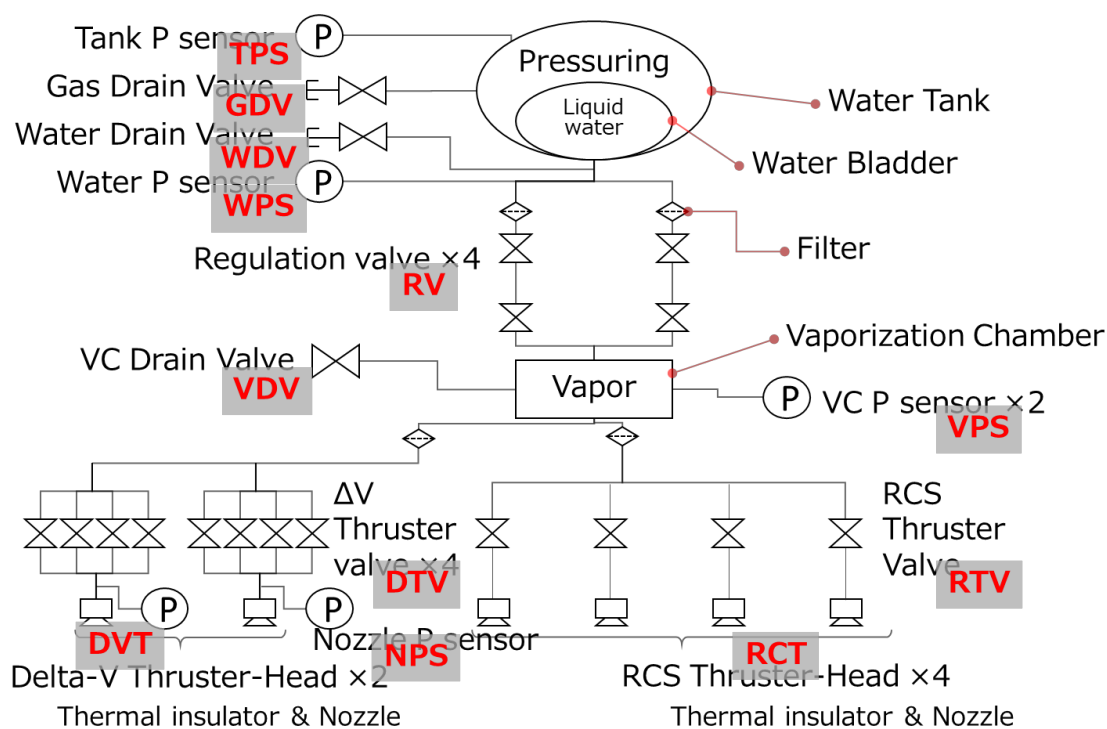


Figure 2-3 AQUARIUS system structure, red letters are abbreviations summarized in Table 2-2 <sup>22</sup>.

## 2.2 取得したデータの試験内容

本研究は推力測定の試験として、以下の試験で得られたデータを使用する.

Table 2-2 Explanation of AQUARIUS ground test experiment.

試験名	対象	説明
振動試験前の 単体作動	推進システ ム単体	推進システム単体が環境試験である振動試験を受ける前 に実施した推進システムの性能確認試験。目的は推進シ ステムが目標の性能に達しているかどうかを確認するこ と。
振動試験後の 単体作動	推進システ ム単体	推進システム単体が振動試験後に受けた推進システムの 性能確認試験。目的は振動試験後、推進システムの性能 変化があるかどうかを確認すること。
温度サイクル 試験後の単体 作動	推進システ ム単体	推進システム単体が温度サイクル試験後に実施した性 能確認試験。目的は温度サイクル試験後推進システ ムの性能変化があるかどうかを確認すること。
パラメータ振 り試験	推進システ ム単体	ノミナル作動試験の条件を決めるために実施した性能 確認試験。システム側の要求を満たす作動条件を決め ること。
ノミナル作動 試験	推進システ ム単体	パラメータ振り試験結果の考察から決めたノミナル作 動条件下での性能確認試験。目的はノミナル作動条件 を見つけること。
衛星統合後の 推力測定試験	衛星統合	推進システムが EQUULEUS と統合後の作動試験。目 的は衛星統合後、推進システムは決めたノミナル作動 条件で推力変更があるかどうかを確認すること。

## 2.3 実験装置

### 2.3.1 真空チェンバー

本研究は宇宙環境を模擬するために、二つの真空チェンバー環境で、AQUARIUS の推進システム単体作動の推力測定試験と EQUULEUS という 6U 衛星の衛星統合推力測定試験を行った。

推進システムの単体作動の推力測定試験では直径 1.2m、長さ 1.7m のステンレス製チェンバーを用いた。以下チェンバー 1 と呼ぶ。Figure 2-4 はチェンバー 1 のシステム図と推力測定の系統図を示している。真空引き用に、ロータリーポンプ、ターボ分子ポンプ、クライオポンプのポンプは三つを使った。Table 2-3 から Table 2-5 にそれぞれのポンプの仕様を示している。

Table 2-3 Specifications of rotary pump.

メーカー	アルバック社
Model	VD401
排気速度	670 L/min
到達圧力	0.67 Pa
モーター形式	防滴保護型 3 相交流モーター
モーター消費電力	1.5 kW, 200V
吸気口	VG40 相当

Table 2-4 Specifications of turbo molecular pump.

メーカー	大阪真空
Model	TG900MVAB
排気速度	900 L/s (N <sub>2</sub> )
到達圧力	1×10 <sup>-6</sup> Pa
最大ガス流量	3000 sccm (N <sub>2</sub> )
吸気口	VG150

Table 2-5 Specifications of cryopump.

メーカー	ジェネシス
本体 Model	ICP250L
ヘリウムコンプレッサ Model	ヘリウムコンプレッサ 2.1
排気速度	2500 L/s (N <sub>2</sub> )
到達圧力	1.33×10 <sup>-7</sup> Pa
標準冷却時間	120 min
標準再生時間	150 min

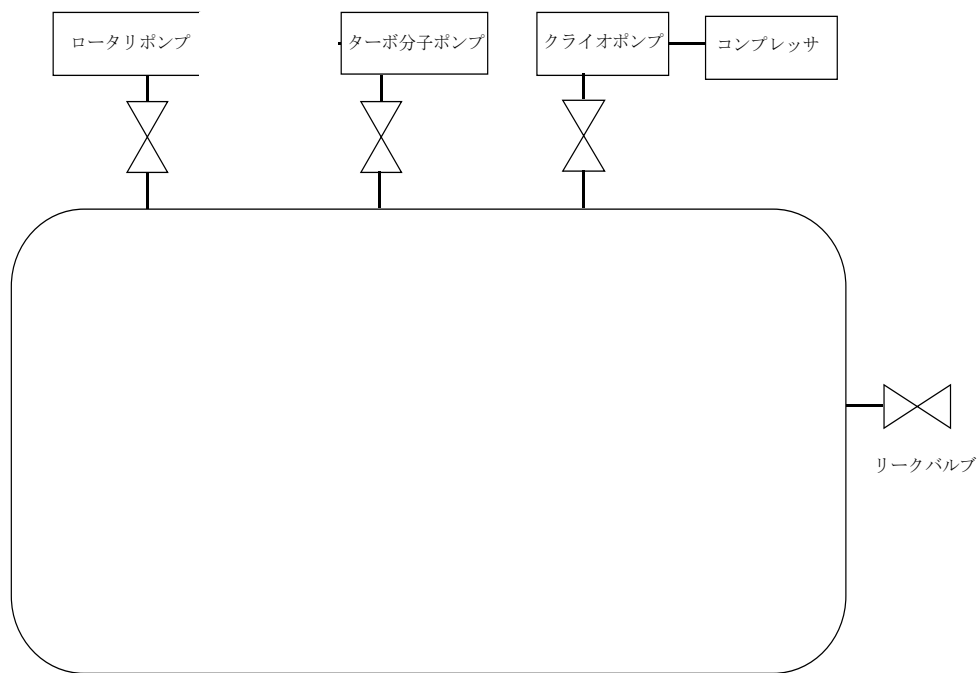


Figure 2-4 Experimental system for thrust measurement chamber.

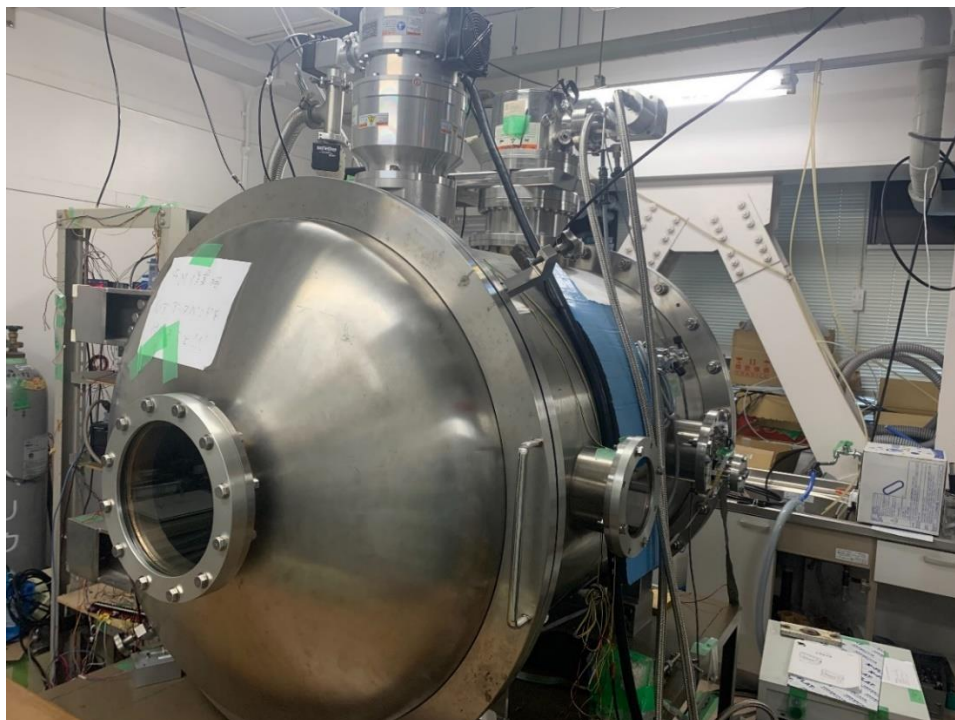


Figure 2-5 External view of chamber 1 for unit operation test.



Figure 2-6 Rotary pump



Figure 2-7 Turbo molecular pump



Figure 2-8 Cryopump



Figure 2-9 Compressor

EQUULEUS と統合した 6U の衛星の推進システムの推力測定試験は直径 2m, 長さ 3m のステンレス製チェンバーチェンバーを使用した. 以下チェンバー2 と呼ぶ. 真空引き用に二つのロータリーポンプ, メカニカルブースターポンプ, オイルディフュージョンポン



プを使っている。チェンバー2の系統図は Figure2-10 に示している。Table2-6 から Table 2-9 はそれぞれのポンプの仕様を示している。

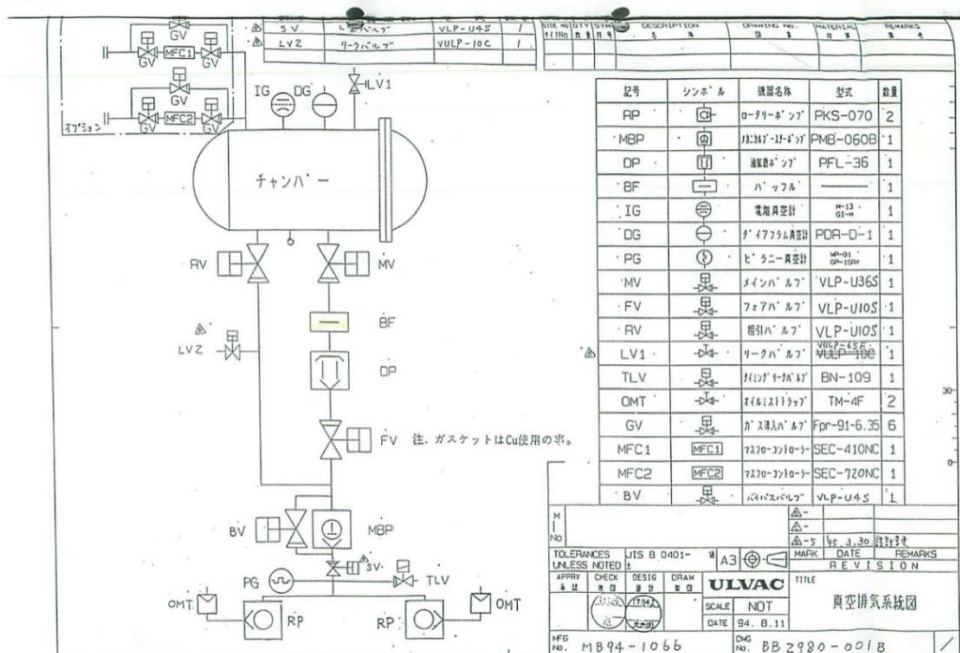


Figure 2-10 Chamber evacuation system diagram for satellite integrated propulsion test.

Table 2-6 Specifications of rotary pump used for integration test.

メーカー	ULVAC
Model	PKS-070
排気速度	116.7 L/s
到達圧力	2.7 Pa

Table 2-7 Specifications of mechanical booster pump used for integration test.

メーカー	ULVAC
Model	PMB-060B
排気速度	1721.7 L/s
到達圧力	$4 \times 10^{-1}$ Pa

Table 2-8 Specifications of oil diffusion pump used for integration test.

メーカー	ULVAC
Model	PFL-36
排気速度	34000 L/s
到達圧力	$3 \times 10^{-4}$ Pa

### 2.3.2 推力測定装置

浅川ら<sup>24</sup>によって設計した重力振り子式スラストスタンドの測定装置を用いて推力を測定した。測定系の系統図は Figure 2-11 に示している。以下測定方法について説明する。

まず、推進システムをスラストスタンドに載せる。振り子を中心位置に確保するために、重心を合わせる必要がある。カウンターウェイトの位置と重さを調節することで、重心を合わせる。スラストスタンドが力を受ける時にスタンドが変位する。この重量振り子式スラストスタンドは力を変位に変換することができる。変位と力の関係は Figure 2-14 に示した推力較正系から較正することで解明できる。推力較正系において、スラストスタンドと接続されている糸に質量が既知の重りを載せる。スラストスタンドの変位を測定するために、Table 2-9 仕様の LED 変位計を使用する。LED 変位計からスタンドが重りから受けた力に対応する変位を測ることができる。実験中、推進剤の質量が減るため、重心の位置が変わる。このため、実験の前後で較正試験を行う。また、較正係数の変化は質量変化と比例するとして、補整をかける。Figure 2-13 は較正試験の時間履歴の一例、Figure 2-14 は校正の解析で得られた変位と力の関係を示している。

Table 2-9 Specifications of LED displacement meter.

メーカー	オムロン
Model	Z4D-F04A
検出範囲	$4 \pm 1.25\text{mm}$
分解能	$5\mu\text{m}$ 以下
アナログ出力	$1-5\text{V}(1.6\text{V}/\text{mm} \pm 10\%)$
電源電圧	$\text{DC}12\text{V}-24\text{V} \pm 10\%$

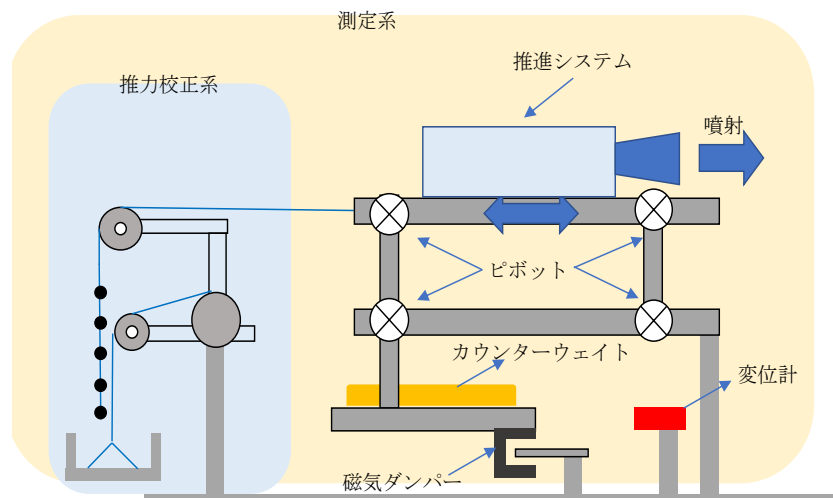


Figure 2-11 System diagram of thrust measurement system.

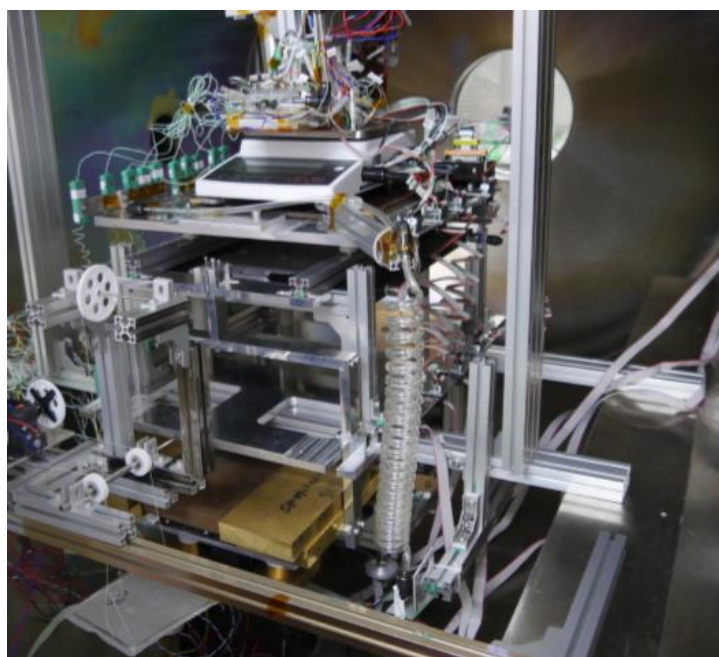


Figure 2-12 Thrust stand appearance

Table 2-10 Results of weight mass measurement.

テーブルに乗せた重りの数	質量/mg
1 個	212.6
2 個	471.1
3 個	671.363
4 個	935.633
5 個	1166.722
6 個	1398.97

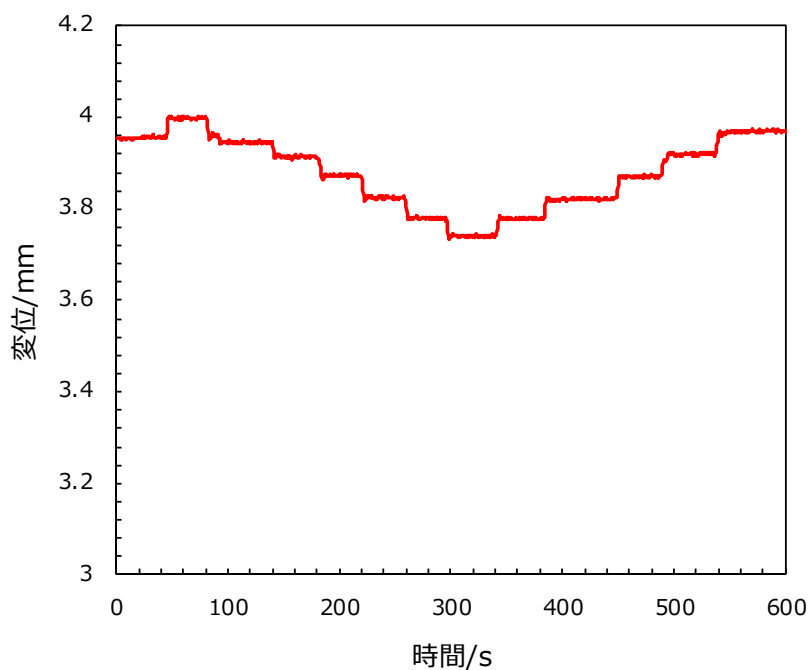


Figure 2-13 Example of time history of calibration test.

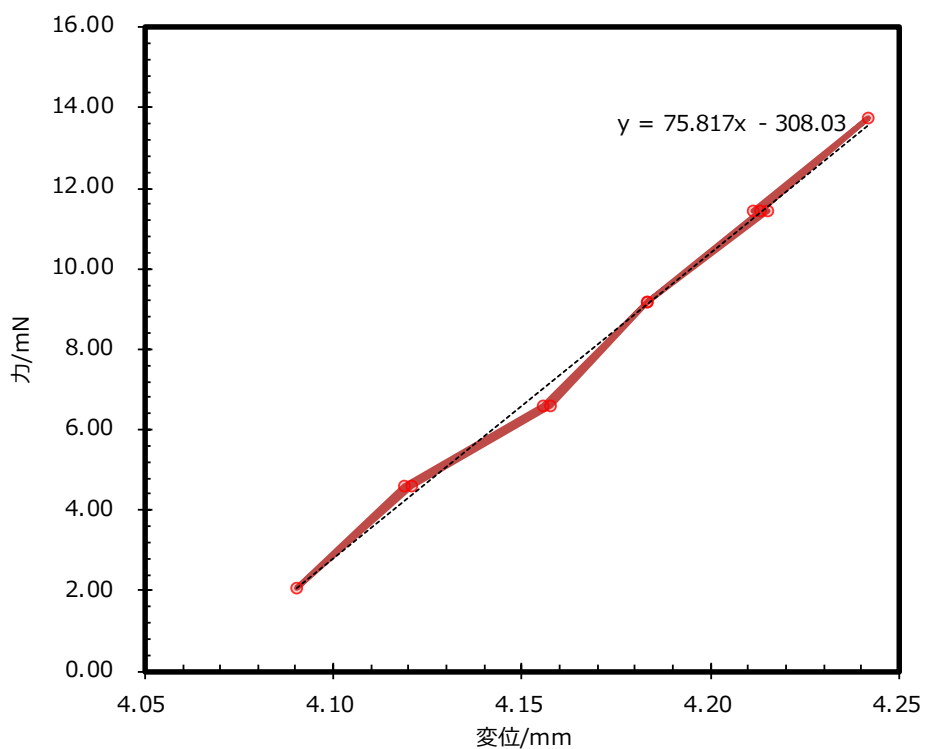


Figure 2-14 The relationship between the displacement and the thrust calculated from the displacement and the mass of the weight is approximated by a straight line, and the slope is used as a calibration coefficient.

推力が $F_T$ ，変位が $x_t$ ，変位と力の関係式の傾きを較正係数 $A$ とすると，

$$F_T = Ax_t \quad (1)$$

として計算することができる．推力の誤差は主に変位計の標準偏差と較正係数の標準偏差から由来する．このため，推力誤差 $\Delta F_T$ が次のように表すことができる．

$$\Delta F_T = \frac{\partial F_T}{\partial A} \Delta A + \frac{\partial F_T}{\partial x_t} \Delta x_t = x_t \Delta A + A \Delta x_t \quad (2)$$

ここで，推力測定試験の較正結果の一例を使用し，誤差を評価する．静止している変位計の標準偏差は0.0015，較正係数の標準偏差は0.003．最大変位は0.13mm．較正係数は76mN/mm．これらを用いて，最大誤差を計算すると，0.114mNとして得られる．

### 2.3.3 AQUARIUS におけるセンサー

ここで AQUARIUS における各圧力センサーと温度センサーの仕様書をまとめる．

Table 2-11 Specifications of pressure sensors in water tank and bladder.

出力範囲	0-15/psi
対応できる温度範囲	-20-85°C

Table 2-12 Specifications of vaporization chamber pressure sensors and nozzle pressure sensors.

対応できる温度範囲	-20-85°C
出力範囲	-10-10/kPa
誤差範囲	±1%FFS
破裂圧力	250kPa

Table 2-13 Platinum thermometer for tubes, nozzles, vaporization chambers.

出力範囲	-50-500°C
誤差範囲	±0.05%FFS
応答時間	0.1s

## 2.4 データセット

本研究で用いたデータは上記の実験から取得されたデータである．機械学習ではそれぞれのデータの名前を特徴量と呼ぶ．本研究は並進運動用の DVT1 と DVT2 の作動データを使

用する. Table 2-14 は使用する特徴量を示している. また Table 2-15 と Table 2-16 はこれらのデータのレンジを示している.

Table 2-14 Explanatory table for data set features.

特徴量	説明
DTS1/°C	DVT1 の温度
DTS2/°C	DVT2 の温度
DVT1_PH/W	DVT1 のヒーターへの印加電力
DVT2_PH/W	DVT2 のヒーターへの印加電力
DPS1/kPa	DTV1 のノズル前圧力
DPS2/kPa	DTV2 のノズル前圧力
Thrust/mN	スラストスタンドを用いて, 変位計のデータを推力に変化した後の値.
Time/s	一サイクルの RV オープン開始からの経過時間
VH/W	気化室のヒーターへの印加電力
VPS1/kPa	気化室圧力その 1
VTS1/°C	気化室温度その 1
VTS2/°C	気化室温度その 2
WPS/kPa	ブラダ内部の圧力

Table 2-15 Range of each feature value in the experiment using DVT1.

特徴量	最小値	最大値
DTS1/°C	54.88	144.97
DVT1_PH/W	0	1.66
DPS1/kPa	-0.052	2.29
Thrust/mN	-2.65	18.10
Time/s	0	107
VH/W	0	19.33
VPS1/kPa	-0.20	5.99
VTS1/°C	28.01	42.91
VTS2/°C	25.57	41.84
WPS/kPa	23.50	69.14

Table 2-16 Range of each feature value in the experiment using DVT1.

特徴量	最小値	最大値
DTS2/°C	44.14	79.10
DVT2_PH /W	0	20.50
DPS2/kPa	-0.095	1.98
Thrust/mN	-2.65	14.17
Time/s	0	257
VH/W	0	20.50
VPS1/kPa	-0.35	5.16
VTS1/°C	27.55	39.69
VTS2/°C	25.41	38.15
WPS/kPa	25.84	79.83

---

## 第3章 機械学習による解析手法

---

### 3.1 概要

本研究では地上試験で得られた AQUARIUS のデータを使用し、並進運動用のスラスターの推力を推定する。また、気化室圧力センサーとノズル前圧力センサーが破損しているのを仮定し、気化室圧力を推定する。本節では機械学習の解析手法の概要について説明する。

▶ データを前処理する

まず、Figure 3-1 のように、入力特徴量と出力特徴量を分ける必要がある。入力特徴量は出力特徴量を推定するために用いる特徴量である。出力特徴量は目標設定によってそれぞれ異なる。詳細は 3.4 と 3.5 で説明する。

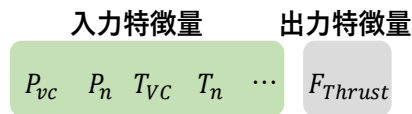


Figure 3-1 Separate input features and output features.

次に、データの特徴に合わせて、既存のデータから新しい特徴量を生み出す。この方法は、特徴量エンジニアリングと呼ばれている。本研究は時系列のデータを扱うため、ラグ特徴量、累積特徴量、Rolling Window による特徴量などの特徴量を作る。また、この推進システムの物理モデルを合わせて、新しい特徴量を作る。データの不均一性を確認することと不均一性を取るために、特徴量の分布を考慮した上でオーバーサンプリングあるいはダウンサンプリングする方法が考えられる。本研究では K-means 法で入力データをクラスタリングする。また、推力推定において、K-means の可視化するために、主成分分析で得られた第一、第二、第三主成分を使い、K-means の結果を可視化する。

更に、Figure 3-2 のように、ランダムにテストデータと学習データを分ける。学習データに対して、本研究は SMOTE というアルゴリズムを使って、オーバーサンプリングする。実際作った機械学習モデルはどれくらいの精度なのかを評価する必要がある。全学習データで評価すると、モデルが過学習に落ちているかどうかを知ることができないため、テストデータと学習データを分けることで、テストデータの精度から誤差評価する。



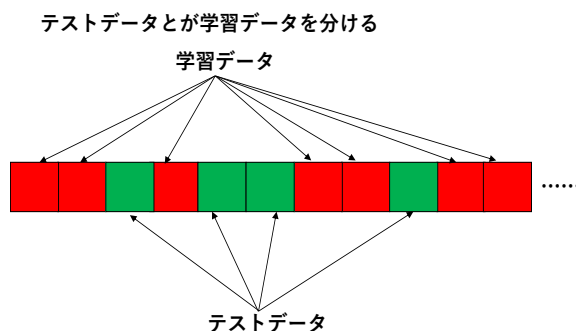


Figure 3-2 Separate train dataset and test dataset.

➤ モデルの選定

機械学習モデルの信頼性、精度と解釈性が本研究の機械学習モデル選定基準である。本研究は機械学習モデルの訓練結果を考察するために、特徴量の重要度を知ることができる決定木と回帰木系の複数のモデルを組み合わせる手法であるアンサンブル学習手法を使用する。決定木と回帰木は結果の解釈がしやすいことに加え、質的データと量的データが混在している状態でも解析が可能という特徴がある。さらに、木のアンサンブル学習手法のランダムフォレストと LightGBM を使うことで、データセットの外れ値を抑える同時に、各特徴量の重要度を出すことができる。このため、本研究は LightGBM とランダムフォレストとの機械学習モデルを採用する。

➤ モデルを訓練する

事前に機械学習モデルの条件を決める時に使用するパラメータをハイパパラメータと呼ぶ。前処理で分けて選られた学習データを使うことで、モデルのハイパパラメータを調整しながら、モデルを訓練する。ハイパパラメータを調整する方法としては、本研究はグリッドサーチと交差検証を使用し、適切なハイパパラメータを選定する。グリッドサーチは、事前にハイパパラメータの組み合わせを設定し、適切なハイパパラメータを選定する方法である。適切さの指標としては、交差検証を用いる。データを数セットに分け、1セットをテストデータとして、残りを学習データとする場合に、機械学習モデルを学習させる。決定係数を計算する。このように、すべてのセットに対して、このような計算を行い、決定係数の平均値を求める。決定係数が一番高い時に使用したハイパパラメータが適切なハイパパラメータとして使用する。

➤ モデルを評価する

テストデータセットでモデルの推定精度を評価する。評価方法は決定係数、平均二乗誤差などがあるが、本研究の推力推定ではテストデータの各サイクルの平均推力の二乗平均平方根誤差 (RMSE) とテストデータのすべての推力の RMSE を評価する。RMSE を用い

た理由は、RMSE の次元は推定値と同じ次元である、絶対精度として使用できる。圧力の推定では、テストデータの圧力推定値と測定値の RMSE を計算し、評価する。(3)式は RMSE を表している。

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3)$$

ここで、 $y_i$  は実値、 $\hat{y}_i$  は推定値を意味している。

## 3.2 機械学習モデル

### 3.2.1 主成分分析 (Principal Component Analysis PCA)

主成分分析 (PCA) はデータセットの特徴量の数、つまりデータセットの次元を圧縮する方法である<sup>25</sup>。多次元のデータを主成分分析することで、主成分でデータセットを表すことができる。それぞれの特徴量のレンジが異なるため、まず、主成分分析を実行する特徴量を同一に評価する。これらの特徴量に対して、データセットを標準化する。例えば  $X$  特徴量を標準化する場合に、標準化された  $X$  特徴量の  $i$  番目  $x_{std_i}$  は

$$x_{std_i} = \frac{x_i - \bar{x}}{s(x)} \quad (4)$$

として表すことができる。ここで、 $\bar{x}$  は  $X$  特徴量の平均値、 $s(X)$  は

$$s(X) = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (5)$$

であり、 $X$  特徴量の標本標準偏差である。

標準化することで、標準化された  $X$  特徴量は平均が 0、標準偏差が 1 になる。  $n$  次元のデータセットを主成分分析する際に、次のような列ベクトルを考える。

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \quad (6)$$

ここで、 $X_{i, i \in \{1, 2, 3, \dots, n\}}$  は、これらの特徴量を意味している。そして、次のような共分散行列を作る。

$$\Sigma = \begin{bmatrix} \text{Var}[X_1] & \cdots & \text{Cov}[X_1, X_n] \\ \vdots & \ddots & \vdots \\ \text{Cov}[X_1, X_n] & \cdots & \text{Var}[X_n] \end{bmatrix} \quad (7)$$

ここで、 $\text{Var}[X], \text{Cov}[X, Y]$  はそれぞれ、 $X$  の分散、 $X, Y$  の共分散を意味している。数式は次のように表すことができる。

$$\text{Var}[X] = s(x)^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (8)$$

$$\text{Cov}[X, Y] = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (9)$$

共分散行列 $\Sigma$ の(i,j)成分は次のように表すことができる.

$$\Sigma_{i,j} = \begin{cases} \text{Var}[X_i] & (\text{if } i = j) \\ \text{Cov}[X_i, X_j] & (\text{if } i \neq j) \end{cases} \quad (10)$$

そして, 対称行列である共分散行列に対して, 固有値と正規化された固有ベクトルを求める. また, 固有値の大きさを順番に並び替え, 次のように対角化することができる.

$$\begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} = U^{-1} \Sigma U \quad (11)$$

ここで,  $U = [\mathbf{p}_1, \mathbf{p}_2 \dots \mathbf{p}_n]$ ,  $\mathbf{p}_1, \mathbf{p}_2 \dots \mathbf{p}_n$ はそれぞれ $\lambda_1, \lambda_2 \dots \lambda_n$ の正規化された固有ベクトルである.  $U$ によって, 元々の特徴量を座標変換するで, 主成分に変換することができる.

本研究でPCAはscikit-learnという機械学習ライブラリを使用する.

### 3.2.2 K-means

K-means法はデータの類似性を尺度に,  $K$ 個のクラスにデータを分類する教師なしの機械学習手法である<sup>26</sup>.  $N$ 組のデータを例として説明すると,  $k$ 番のクラスに属するデータの代表ベクトルを $\boldsymbol{\mu}_k$ として,  $i$ 番目のデータ $\mathbf{x}_i$ が $k$ 番目のクラス $M(\boldsymbol{\mu}_k)$ に属しているかどうかを表す帰属変数は

$$q_{ik} = \begin{cases} 1 & (\text{if } \mathbf{x}_i \in M(\boldsymbol{\mu}_k)) \\ 0 & (\text{other}) \end{cases} \quad (12)$$

として定義する. 帰属変数を使うことで, 類似性の評価関数を次のように定義できる.

$$J(\boldsymbol{\mu}_k, q_{ik}) = \sum_{i=1}^N \sum_{k=1}^K q_{ik} \|\mathbf{x}_i - \boldsymbol{\mu}_k\|^2 \quad (13)$$

評価関数を最小にすることでクラスを分類できる. この場合, 代表ベクトルは

$$\boldsymbol{\mu}_k = \frac{\sum_{i=1}^N q_{ik} \mathbf{x}_i}{\sum_{i=1}^N q_{ik}} \quad (14)$$

である.

最初はランダムに $\boldsymbol{\mu}_k$ を振って,  $q_{ik}$ を計算する. その後, 上記の手順で新たに $\boldsymbol{\mu}_k, q_{ik}$ を計算する. 状態変化が収束するまでに計算を繰り返す. この後, 次のようにクラスの領域を決めることができる.

$$R_k = \{\mathbf{x} \mid \|\mathbf{x} - \boldsymbol{\mu}_k\|^2 < \|\mathbf{x} - \boldsymbol{\mu}_i\|^2 \text{ for all } i \neq k\} \quad (15)$$

本研究でのK-meansはscikit-learnという機械学習ライブラリを使用する.

### 3.2.3 Synthetic Minority Over-sampling Technique (SMOTE)

データセットに不均一性が存在する場合に、推定結果はしばしば多くの観測クラスに偏る。オーバーサンプリング或いはダウンサンプリングすることでデータセット不均一性を取り除く方法である。ChawlaらはオーバーサンプリングのSMOTEアルゴリズムを提案した<sup>27</sup>。SMOTEはK-NNというアルゴリズムを基に人工的に新たなサンプルを生成する方法である。新しいサンプルを生成する手順は次の通り：

- 1) Figure3-3のように、クラスが少ないn個のサンプルをランダムに選択する。

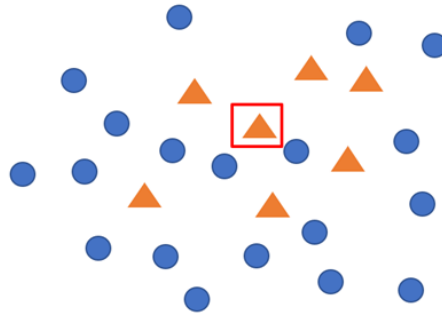


Figure 3-3 Orange triangles are data samples with few classes. The triangle inside the red frame is one of the randomly selected data.

- 2) ランダムに選択したデータと同じクラスの周囲N個のサンプルを見つける。

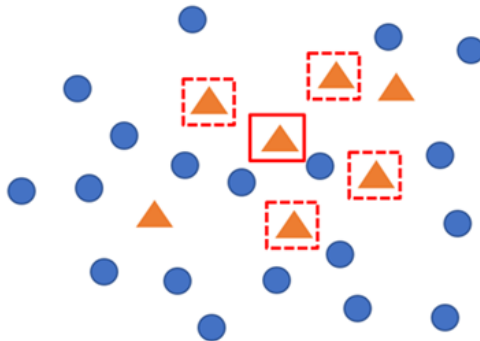


Figure 3-4 Find N samples of the same class around each randomly selected data.

3) 同じクラスの周囲データの間ランダムな位置で新しいサンプルを作る。

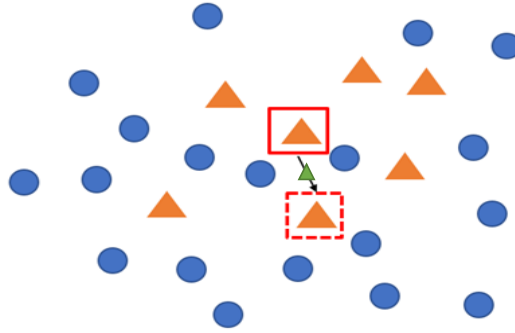


Figure 3-5 Green small triangles are samples artificially created between two data of the same class..

本研究での SMOTE は imbalanced-learn API を使用する。

### 3.2.4 回帰木と決定木

#### ➤ 回帰木<sup>28</sup>

回帰木と決定木はそれぞれ入力特徴量から、出力を計算する回帰モデルと分類モデルである。Figure 3-6 のように、 $R_0$  から  $R_n$  までの番号が付けられたノードとノード間を結 Yes, No とラベル付けされたリンクが示される。一番上にある  $R_0$  のノードは木の始まりという意味で根ノードと呼ばれる。四角で示されたノードは木の終端ノードであり、葉ノードとも呼ばれる。それ以外のノードは内部ノードと呼ばれる。例えば、データセットを  $(\mathbf{x}_i, \mathbf{y}_i)$  とすると、簡単のために、本研究で用いた特徴量の一部を使用し、入力の  $\mathbf{x}_i$  と出力の  $\mathbf{y}_i$  をそれぞれ次のように定義する。

$$\mathbf{x}_i = (P_{vc}(i), P_n(i), T_{VC}(i), T_{tube}(i), T_n(i)) \quad (16)$$

$$\mathbf{y}_i = F_{Thrust}(i) \quad (17)$$

$P_{vc}(i), P_n(i), T_{VC}(i), T_{tube}(i), T_n(i), F_{Thrust}(i)$  はそれぞれ  $i$  番目のデータの気化室圧力、ノズル前圧力、気化室温度、チューブ温度、ノズル温度を表す。回帰木モデルはデータセットを Figure 3-6 のように分割する。すべてのデータ  $\mathbf{x}_i$  に対して、特徴量の中に気化室圧力は閾値 0.5kPa より大きければ、一つのデータセットを二つのデータセットに分け、ノード  $R_1, R_2$  を作る。また、同じ操作を繰り返し行い、回帰木を作る。

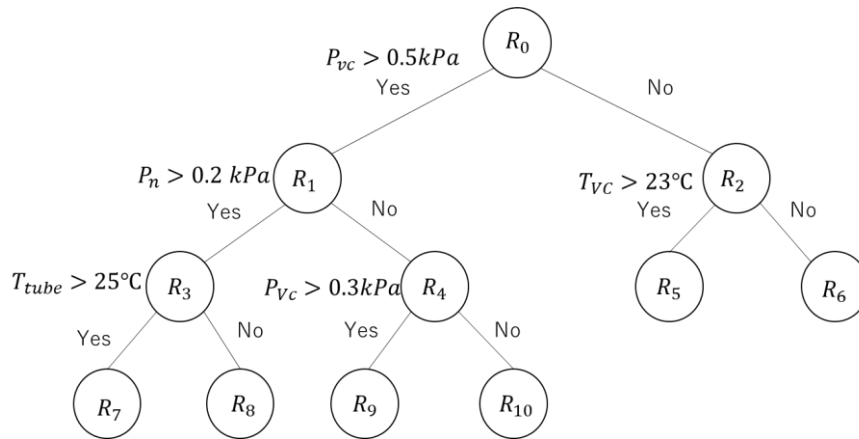


Figure 3-6 Example of visualization diagram of regression tree and decision tree.

最適の分割基準は分岐後それぞれの二つのグループにおけるモデルで選んだ出力データの平均値と実際の出力平均二乗誤差(MSE)が最小になるように選定する. この平均二乗誤差はこのモデルの損失関数という.

$R_{node}$  ノードの MSE は次のように定義する.

$$L_R(R_{node}) = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \quad (17)$$

$j$  は  $\mathbf{x}_i$  の  $j$  番目のパラメータとして, 最適分割パラメータ  $j$  や閾値  $s$  の選定の最適化は(18)式に従う.

$$\min_{j,s} \left[ \min_{c_1} \sum_{\mathbf{x}_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{\mathbf{x}_i \in R_2(j,s)} (y_i - c_2)^2 \right] \quad (18)$$

ここで  $c_1, c_2$  は  $R_1, R_2$  における  $y_i$  の平均値を表している.  $R_1, R_2$  のケース式で説明すると,  $y_i$  はその領域におけるそれぞれの時間における推力データを表している. 例えば  $R_m$  のノードにおいて推力平均値は(19)に表すことができる.

$$c_m = \text{ave}(y_i | \mathbf{x}_i \in R_m) \quad (19)$$

式(3)の最適化は  $R_1, R_2$  における推力の平均二乗誤差 (MSE) が最小になるように選定する. 上記の過程を繰り返すことで, 回帰木を生成する.

$$y = f(\mathbf{x}) = \sum_{k=1}^m I(\mathbf{x} \in R_k) \times c_k \quad (20)$$

$I(\mathbf{x} \in R_k)$  は入力  $\mathbf{x}$  が  $R_k$  内にある時に 1, そうでない時に 0.

このモデルを実行する際に, どこまで分岐を実行するかを決める必要があり, 回帰木のモデルにおいて, 木の深さや葉ノードにあるサンプル数がハイパパラメータである.

単一の回帰木の場合, 最初の特徴量の選定, ハイパパラメータの選定による, 予測が悪くなる可能性もあるため, ロバスト性が悪い.

### ➤ 決定木<sup>28</sup>

決定木は入力特徴量から, 入力データが属しているグループを計算する分類モデルである. 決定木の基本構造は回帰木と同じで, 違いとして, 分岐の基準は MSE の代わりに, 交差エントロピー, ジニ係数, 誤り率を用いるケースが多い. 本研究は蒸発時かどうかを

判断する時に、ジニ係数を用いた。この関数は決定木の損失関数とする。式(21)はジニ係数を示している。

$$L_c(R) = \sum_{i=1}^k P(C_i|R_{node}) (1 - P(C_i|R)) \quad (21)$$

ノード $R_{node}$ は $i$ 番目のクラスに属す確率を $P(C_i|R_{node})$ とする。ジニ係数はすべてのクラスの分散の和を表している。つまり、分類の不純度を表している。ラベル付きのデータセットを $\{(x_i, R_i)\}$ とする。クラス $j$ に属するデータ数を $N_j$ とすると、事前確率は

$$P(C_j) = \frac{N_j}{N} \quad (22)$$

で表すことができる。  $R_{node}$ ノードのデータ数を $N(R)$ とすると、この $N(R_{node})$ のうち、クラス $j$ に属するデータ数を $N_j(R_{node})$ とする。そうすると、クラス $j$ のデータがノード $R_{node}$ に属する確率は

$$P(C_j, R_{node}) = P(C_j)P(R_{node}|C_j) = \frac{N_j(R_{node})}{N(R_{node})} \quad (23)$$

と表すことができる。

ノード $R_{node}$ から子ノードを分割際に、左側のノードを $R_L$ 、右側のノードを $R_R$ とすると、 $\mathbf{x}$ がどちらのノードに属する確率は

$$p_L = \frac{p(R_L)}{p(R_{node})} \quad (24)$$

$$p_R = \frac{p(R_R)}{p(R_{node})} \quad (25)$$

に従って決まる。

ノード $R_{node}$ で分割する際に、この不純度が一番減らせる $s$ を選ぶことで、分類を実現する。この時に、ジニ係数の変化は次の式のように表すことができる。

$$\Delta L(s, R_{node}) = L(RR_{node}) - (p_L L(R_L) + p_R L(R_R)) \quad (26)$$

$$\min_{j,s} \left[ \min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right] \quad (27)$$

### 3.2.5 ランダムフォレスト

ランダムフォレストはバグgingのアンサンブル学習である<sup>29</sup>。このモデルはランダム特徴量を選定し、学習データからブートストラップサンプルを生成する。また、この操作を複数回行い、選定されたデータからいくつかの回帰木或いは決定木を生成する。ランダムに特徴量を選定するかつ複数の木を生成するため、ランダムフォレストと呼ばれている。図3-7に示したように、回帰木のランダムフォレストの出力は各回帰木の出力の平均値を取る。決定木のランダムフォレストの場合に、入力が多数決で属しているクラスを決める。ランダムの特徴量を選定したため、比較的ロバスト性が優れていて、過学習になる可能性は低い。

ランダムフォレストは特徴量の重要度を出すことができる。本研究で重要度を出したのは回帰木のランダムフォレストであるため、回帰木のランダムフォレストの重要性について説明する。3.1.3 で回帰木構成の説明を示したように、ノードを分岐するたびに、最適分割パラメータの選定を行っている。その特徴量の選定により、分割されたノードにおける平均二乗誤差(MSE)の減少はその特徴量の重要度に関連していると考えられる。

$R_m$ ノードにおける推力データ MSE を $f_{MSE}(R_m)$ として、 $R_0$ から $R_1$ と $R_2$ に分割する例を考えると、

$$\Delta f_{MSE}(R_0) = f_{MSE}(R_0) - (f_{MSE}(R_2) + f_{MSE}(R_1)) \tag{28}$$

すべての分割において、式 (12) のように MSE を計算できる。それらの和を $\Sigma f_{MSE}$ とする。

例えば図 3-2 のように、気化室圧力 $P_{vc}$ がノード $R_0, R_4$ の分岐に使われている。この際に、気化室圧力の重要度 $I_{P_{vc}}$ を次のように定義する。

$$I_{P_{vc}} = \frac{\Delta f_{MSE}(R_0) + \Delta f_{MSE}(R_4)}{\Sigma f_{MSE}} \tag{29}$$

他の特徴量の重要度も同じように計算できる。

回帰木のランダムフォレストの場合は、それぞれの回帰木におけるこのような計算を行い、ジニ係数を評価基準として、重要度を評価する。

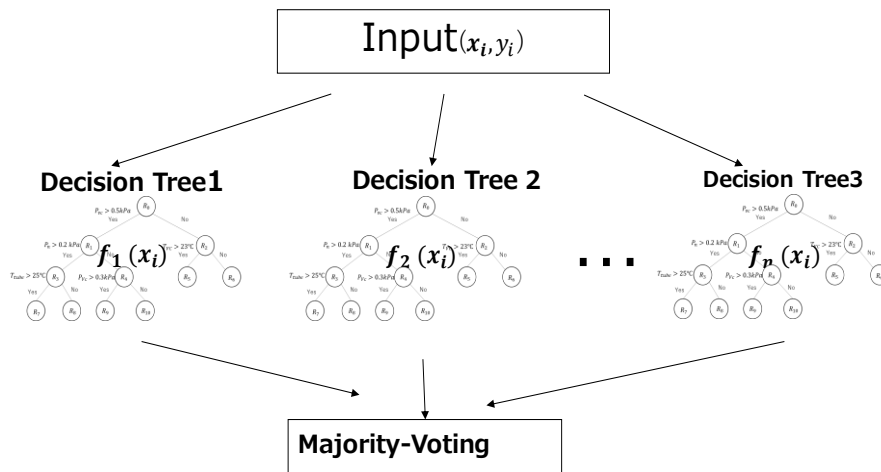


Figure 3-7 Example of random forest of decision tree.

本研究でのランダムフォレストは scikit-learn という機械学習ライブラリを使用する。Table 3-1 は本研究で使ったランダムフォレストのハイパパラメータを示している。



Table 3-1 Hyper-parameter explanation table of random forest.

ハイパパラメータ	説明
n_estimators	ランダムフォレストに使われる木の数
min_samples_leaf	各ノードに必要な最小サンプル数
max_depth	各木の最大の深さ, このハイパパラメータを設定しないと, 葉ノードのデータが min_samples_leaf になるまで成長する.

### 3.2.6 LightGBM

LightGBM はマイクロソフトが 2017 年にオープンソース化した決定木と回帰木のアンサンブル学習アルゴリズムの一つであり, 勾配ブースティングのフレームワークの機械学習手法である<sup>30,31</sup>.

数式で説明すると, 全部  $n$  組の訓練データに対して, 最初に入力データ  $\mathbf{x}_i$  に対応する推定するターゲット出力  $y_i$  の平均値を求める.

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i \quad (30)$$

次に,  $y_i$  と  $\bar{y}$  の差  $y_i^1$  を求める.

$$y_i^1 = y_i - \bar{y} \quad (31)$$

そして,  $y_i^1$  を推定するターゲットとして, モデルを学習させる. モデル  $f_1$  が一つ目の木モデルとして, すべての入力  $\mathbf{x}_i$  から, 次のターゲット

$$y_i^2 = y_i^1 - \alpha f_1(\mathbf{x}_i) \quad (32)$$

を推定する. ここで,  $\alpha$  は学習率である.

更に,  $y_i^2$  を二つ目のモデルの推定するターゲットとして, 二つ目木モデルを学習させる. これを繰り返すことで, 複数のモデルを作成する. モデル誤差が収束する時に,  $N$  個の木が生成されるとすると,  $N$  番目の推定結果は

$$y_i^{N+1} = y_i^N - \alpha f_N(\mathbf{x}_i) \quad (33)$$

として計算することができる.

最後に, これらの結果を全部足し合わせることで, 推定値の最終出力

$$\hat{y}_i = \sum_{j=1}^{N+1} y_i^j \quad (34)$$

とする.

LightGBM は勾配ブースティングの決定木と回帰木の扱い方に工夫した. 本研究に用いた LightGBM の従来の勾配ブースティングからの改良点は以下四つである.

#### 1) 損失関数

過学習を抑えるために L1, L2 正則項の  $\gamma T, \lambda \sum_{j=1}^T \omega_j$  を導入した。LightGBM の損失関数は次のように定義する。

$$L_L(\phi) = \sum_{i=1}^n (\hat{y}_i - y_i)^2 + \sum_{k=1}^N \Omega(f_k) \quad (35)$$

$$\Omega(f) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j \quad (36)$$

ここで、 $T$  はモデル  $f$  の木のノードの数、 $\omega_j$  は  $f$  の各葉ノードの平均値である。 $f_k$  は  $k$  番目の木である。

## 2) ヒストグラムアルゴリズム

従来の決定木と回帰木は厳密な枝分かれポイントを探す代わりに、訓練データの特徴量をヒストグラム化することで、最適な分割ポイントを探す。こうすることで、計算コストを大幅軽減できた同時に、推定精度を確保することもできる。

## 3) Exclusive Feature Bundling (EFB)

同時にゼロが多い疎になっている特徴量が存在する場合、特徴量を統一することで、最適な分割点を探す計算コストを大幅軽減できる。EFB を実行する問題は NP-Hard 問題であり、マイクロソフトはこれを着色問題に変換し、貪欲法を用いてこの問題を解決した。LightGBM はヒストグラムアルゴリズムを利用するため、EFB を使用することで計算コストをさらに軽減することができる。

## 4) Leaf wise

一般的に木の成長方法はレベルから成長する Level-Wise 方法と木の葉から成長していく Leaf-Wise の方法がある。従来の Gradient Boosting は Level-Wise を使うが、LightGBM は複雑な木を作ることができる Leaf-Wise を使う。図 3-8 に示されたように、Level-Wise の場合、ノードからデータを分割した後、分割の左右のノードに対して、さらに分割していく。Level-Wise は予測精度が悪い一方、過学習を抑えることができる。図 3-9 に示されたように、Leaf-Wise の場合は一方のノードから成長して行き、成長し終わったら、片方を成長させる。Level-Wise より予測精度が優れている。

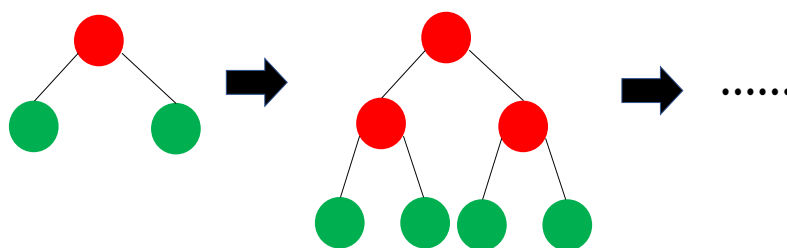


Figure 3-8 Level -Wise, Green nodes will grow in the next step, red nodes will not grow in the next step.

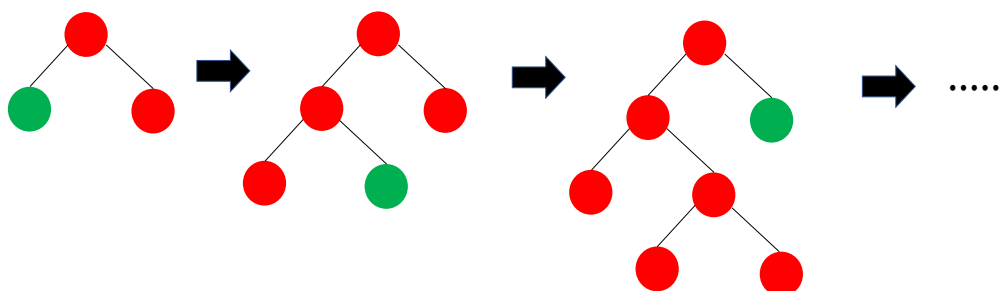


Figure 3-9 Leaf -Wise, green nodes will grow in the next step, red nodes will not grow in the next step.

本研究ではマイクロソフトが開発した LightGBM のライブラリを利用する。LightGBM は訓練時間が短い、メモリー効率が高い、予測精度が高いメリットがある。欠点としては、ランダムフォレストのようなバグギングのアンサンブル学習手法より過学習となるか可能性が高い。

LightGBM に使われるハイパパラメータを表 3-2 にまとめた。

Table 3-2 Hyperparameter description table used for LightGbM.

ハイパパラメータ	説明
num_leaves	一つの木にある最大の葉の数
max_depth	一つの木の最大深さ
learning_rate	学習率
feature_fraction	毎回木を生成する際にランダム選ぶ特徴量の割合
reg_alpha	過学習を抑えるために使う損失関数の L1 正則化の係数
reg_lambda	過学習を抑えるために使う L2 正則化の係数

### 3.3 特徴量エンジニアリング

本研究は時系列のデータを使っているため、ラグ特徴量、Rolling Window による特徴量を使った。また、推進システムのモデルから無次元数などの特徴量を作った。これから詳細について説明する。

#### 3.3.1 ラグ特徴量

過去のデータが現在を推定したい特徴量に影響がある場合に、過去のデータを新しい特徴量として使い、推定することが有効である。本研究は過去の3sの各所の圧力、温度データを新しい特徴量として使った。また、ある時点から数秒後のデータも現時点に推定したい特徴量と関係がある場合、未来のデータから現時点のデータを推定することもできる。ただし、この特徴量を使う場合、リアルより数秒の遅延で推定することになる。本研究でこの特徴量を用いた。説明を Table 3-3 にまとめた。

Table 3-3 Description table of lag features.

ラグ特徴量	説明
DPS1_shift(n)/kPa	n 秒前後 DPS1 センサーが記録したノズル前圧力データ、ここで n は-1,-2,-3,1,2,3 である。
DPS2_shift(n)/kPa	n 秒前後 DPS2 センサーが記録するノズル前圧力データ、ここで n は-1,-2,-3,1,2,3 である。
DTS1_shift(n)/°C	n 秒前後 DTS1 センサーが記録した気化室温度データ、ここで n は-1,-2,-3,1,2,3 である。
DTS2_shift(n)/°C	n 秒前後 DTS2 センサーが記録した気化室温度データ、ここで n は-1,-2,-3,1,2,3 である。
VPS1_shift(n)/kPa	n 秒前後 VPS1 センサーが記録した気化室圧力データ、ここで n は-1,-2,-3,1,2,3 である。
VTS1_shift(n)/°C	n 秒前後 VTS1 センサーが記録する気化室温度データ、ここで n は-1,-2,-3,1,2,3 である。
VTS2_shift(n)/°C	n 秒前後 VTS2 センサーが記録する気化室温度データ、ここで n は-1,-2,-3,1,2,3 である。

## 3.3.2 時間積分値特徴量

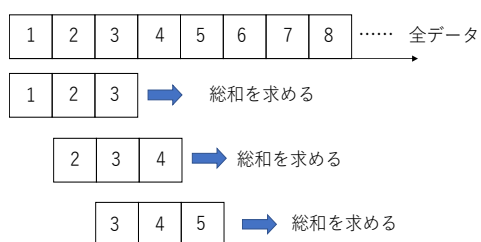
過去のデータの時間累積値が現在推定したい特徴量に影響がある場合に、過去のデータの累積値を使って推定することが有効である。本研究は噴射開始時からの気化室の圧力データ、ノズル前圧力、気化室印加電力、ノズル印加電力の時間積分値の特徴量を作った。

Table 3-4 Description table of time integral value features.

時間積分特徴量	説明
VPS1_integral/kPas	気化室圧力の時間積分値
DPS1_integral/kPas	DPS1 が記録したノズル前圧力の時間積分値
DPS2_integral/kPas	DPS2 が記録したノズル前圧力の時間積分値
VH_integral/J	気化室ヒーターの印加電力の時間積分値
DVT1_PH_integral/J	DVT1 ノズルヒーターの印加電力の時間積分値
DVT2_PH_integral/J	DVT2 ノズルヒーターの印加電力の時間積分値

## 3.3.3 Rolling Window による特徴量

Feature 3-13 のように推定したい時刻のターゲット特徴量の過去の入力特徴量からの移動平均或いは総和を計算し、新たな特徴量を作る方法。この方法は時系列データを扱う推定において有効である<sup>33</sup>。本研究は3秒の間のデータの総和という特徴量を作った。作った特徴量を Table 3-5 にまとめた。



Feature 3-10 Rolling Window

Table 3-5 Description table of features that created by Rolling Window.

## 3.3.4 他の特徴量

蒸発中の情報は推力と一番関連していると考えられるため、蒸発モード中の気化室平均温度、平均圧力、そしてノズル平均温度、平均ノズル前圧力の特徴量を作った。また、ノズルへの印加電力、気化室への印加電力の総和の特徴量を作った。説明を Table 3-6 にまとめた。

Table 3-6 Description table of other features.

Rolling Window 特徴量	説明
DVT1_PH_sum/J	蒸発中 DVT1 への印加電力の総和
DVT2_PH_sum/J	蒸発中 DVT2 への印加電力の総和
VH_sum/J	蒸発中気化室への印加電力の総和
VTS1_mean/°C	VTS1 が記録した蒸発中気化室温度の平均値
VTS2_mean/°C	VTS2 が記録した蒸発中気化室温度の平均値
VPS1_mean/kPa	VPS1 が記録した蒸発中の気化室圧力の平均値
DPS1_mean/kPa	DPS1 が記録した DVT1 のノズル前圧力の平均値
DPS2_mean/kPa	DPS2 が記録した DVT2 のノズル前圧力の平均値

## 3.3.5 推進システムのモデルから得られる特徴量

平均流量、水蒸気の粘性度、スロートレイノルズ数、気化室内の水蒸気圧の特徴量を作った。これからそれぞれの特徴量について説明する。

## ➤ 噴射量

AQUARIUS を天秤に載せて、噴射量を測定する実験した。AQUARIUS の噴射量制御を制御するには、RV のオープン時間をコントロールしている。ブラダ内の圧力は WPS センサーの測定値  $P_{wps}$  として、オープン時間は  $t_{open}$  とすると、噴射量  $m$  は下式のように表すことができる。

$$m = aP_{wps}^n t_{open} \quad (37)$$

ここで、 $a = 0.0311$ ,  $n = 0.696$  であり、AQUARIUS を天秤に乗せた実験からの実験値からのフィッティング結果である。分散は  $5.0447 \times 10^{-4}$  である。

## ➤ 平均流量

実験データから水タンクの圧力と RV オープン時間から噴射量の質量を計算することができる。ここで質量を  $m/kg$  とする。そして、1 サイクルのデータ中で気化室圧力が 0.2kPa 以上の時間を統計し、蒸発時間  $t_v/s$  とする。

1 サイクルの平均流量を以下のように定義する。

$$\dot{m} = \frac{m}{t_v} \quad (38)$$

➤ 気化室内に残った水の質量

タンクから気化室へ噴射した水の質量と平均流量から気化室内に残った水の質量を計算できる.

$$m_l = m - \dot{m}t \quad (39)$$

➤ 粘性度

サザーランドの式を用いて, 粘性度を計算した.

$$\mu^* = \mu_0 \left( \frac{T_n}{T_0} \right)^{\frac{3}{2}} \left( \frac{T_0 - S}{T_n - S} \right) \quad (40)$$

ここで,  $T_0=288.15\text{K}$ ,  $\mu_0=9.69.6 \times 10^{-6}\text{Pa}\cdot\text{s}$ ,  $S = 620.15\text{K}$ .

➤ スロートレイノルズ数

$$Re^* = \frac{4\dot{m}}{\pi D_t \mu^*} \quad (41)$$

ここで $D_t$ はスロートの開口径である.

飽和水蒸気圧:

$$P_{sat} = 6.1112 \exp\left(\frac{17.6T_{vc}}{T_{vc} + 243.5}\right) \quad (42)$$

本研究は $T_{vc}$ に気化室の温度である VTS2 の温度を代入した.

➤ 圧力損失

気化室とノズルの間の圧力損失,  $P_{vc}$ は気化室圧力,  $P_n$ はノズル前の圧力を表している.

$$\Delta P = P_{vc} - P_n \quad (42)$$

上記の物理に基づいて計算した特徴量を Table 3-7 にまとめる.

Table 3-7 Features based on physics.

モデルから得られる特徴量	説明
mas/g	気化室に送った水の噴射量
mass flow rate/(g/s)	蒸発中の平均流量
mass_left/g	気化室内に残った水の残量
Viscosity/Pas	粘性度
Re	スロートレイノルズ数
Psat/kPa	気化室内の水の飽和水蒸気圧力
P_Delta/kPa	気化室とノズル前の圧力損失

### 3.4 機械学習を用いた推力推定

#### 3.4.1 K-means でクラス分け

本研究では、DVT1, DVT2 ノズルに対して、それぞれ推力推定を行った。まず、Table3-8のようにランダムに噴射回数のIDを選択した。

Table 3-8 Number of injections of test data and learning data.

ノズル名	テストデータの噴射回数	学習データの噴射回数
DVT1	81	323
DVT2	89	358

そして、3.3で説明したように、特徴量エンジニアリングすることで新たな特徴量を作った。データセットに不均一性が存在する可能性があるため、不均一性を取り除くために、K-meansというアルゴリズムを用いて、入力用の特徴量を5クラスに分類した。

また、分類された結果を主成分分析特徴量の次元を落とし、可視化した。

Table 3-9はDVT1とDVT2の推力推定に使われた特徴量を示している。

Table 3-9 Input features used for thrust estimation.

DVT1の推力推定に使用する入力特徴量	DVT2の推力推定に使用する入力特徴量
Time/s	Time/s
DTS1/°C	DTS2/°C
VTS2/°C	VTS2/°C
VTS1/°C	VTS1/°C
WPS/kPa	WPS/kPa
VPS1/kPa	VPS1/kPa
DPS1/kPa	DPS2/kPa
VH/W	VH/W
DVT1_PH/W	DVT2_PH/W
mass/g	mass/g
DPS1_shift(n)/kPa	DPS2_shift(n)/kPa
DTS1_shift(n)/kPa	DTS2_shift(n)/kPa
VPS1_shift(n)/kPa	VPS1_shift(n)/kPa
VTS1_shift(n)/°C	VTS1_shift(n)/°C



VTS2_shift(n)/°C	VTS2_shift(n)/°C
DTS1_rolling/°C	DTS2_rolling/°C
VTS2_rolling/°C	VTS2_rolling/°C
VTS1_rolling/°C	VTS1_rolling/°C
VPS1_rolling/kPa	VPS1_rolling/kPa
VH_rolling/J	VH_rolling/J
DVT1_PH_rolling/J	DVT2_PH_rolling/J
DPS1_rolling/kPas	DPS2_rolling/kPas
DVT1_PH_sum/J	DVT2_PH_sum/J
VH_sum/J	VH_sum/J
VTS1_mean/°C	VTS1_mean/°C
VTS2_mean/°C	VTS2_mean/°C
DTS1_mean/°C	DTS2_mean/°C
DPS1_mean/kPa	DPS2_mean/kPa
VPS1_mean/kPa	VPS1_mean/kPa
P_Delta/kPa	P_Delta/kPa
mass flow rate/(g/s)	mass flow rate/(g/s)
viscosity/Pas	viscosity/Pas
Re	Re
Psat/kPa	Psat/kPa
mass_left/g	mass_left/g
VPS1_integral/kPas	VPS1_integral/kPas
DPS1_integral/kPas	DPS2_integral/kPas
VH_integral/J	VH_integral/J
DVT1_PH_integral/J	DVT1_PH_integral/J

一部2つの変位計から推力を測定する実験があり、それらの実験において2つの測定値のRMSEは約0.32mNである。LightGBMとランダムフォレストの推定結果と0.32mNと比較し、妥当性を検証する。また、特徴量エンジニアリングとオーバーサンプリングの手法の有効性を検証するために、Table3-10のように3つのケースでそれぞれLightGBMとランダムフォレストを用いて、推力を推定し、精度を比較する。

Table 3-10 Cases of thrust estimation by Random forest.

	特徴量エンジニアリング	オーバーサンプリング
CASE1	なし	なし
CASE2	あり	なし
CASE3	あり	あり

### 3.4.2 LightGBM を用いた推力推定

上記で説明した K-means で分類後、学習データのそれぞれの 5 組に対して、SMOTE でオーバーサンプリングした。次に、モデルを学習させる。また、過学習を防ぐために、木を生成するたびに、テストデータを用いて、テストデータの RMSE を計算する。学習データとテストデータの RMSE 両方が収束する際に、学習を止める。

全データセットに対して、ハイパパラメータのチューニングはグリッドサーチと交差検証を用いる。LightGBM のハイパパラメータが多いため、初期設定から今回は優先度順に 1,2,3 の順にハイパパラメータをグリッドサーチと交差検証で選定した。交差検証では 10 分割で、決定係数を基準とする。優先度が 1,2,3 の順番で選定した後、ハイパパラメータの初期値を更新する。初期値を Table 3-11 にまとめた。

Table 3-11 Initial value of LightGBM hyper parameter.

ハイパパラメータ	初期値
num_leaves	40
max_depth	6
learning_rate	0.1
feature_fraction	0.85
reg_alpha	0
reg_lambda	0

グリッドサーチに使うハイパパラメータのサーチ範囲と各ハイパパラメータの優先度を Table 3-12 にまとめた。

Table 3-12 LightGBM Hyperparameter Grid Search Range and Priority.

ハイパパラメータ名	サーチ範囲	優先度
reg_alpha	0.5,0.6,0.7,0.8,0.9,1	1
reg_lambda	0,1,2,3,4,5,6	1
num_leaves	5,6,7,8,9,10	2
max_depth	30,40,50,60,70	2
learning_rate	0.05,0.1,0.2,0.3	3
feature_fraction	0.7,0.8,0.85,0.9	3

LightGBM は計算が早いため、400 回ランダムに学習データとテストデータを分け、それぞれ学習させる。評価は次の 2 パターンで行う。

- 1) 推定精度の汎化誤差を評価するために、400 回ランダムにデータを選定し、モデルを学習させる。そして、400 回のテスト結果の RMSE と RMSE の標準偏差を求め、RMSE の信頼区間を評価する。
- 2) テストデータに対して、一サイクルの推定値の平均推力を求める。また、各サイクルの噴射の平均推力の RMSE と、RMSE の標準偏差を求め、RMSE の信頼区間を評価する。

それぞれのアルゴリズムを Figure 3-11 から Figure 3-13 にまとめた。

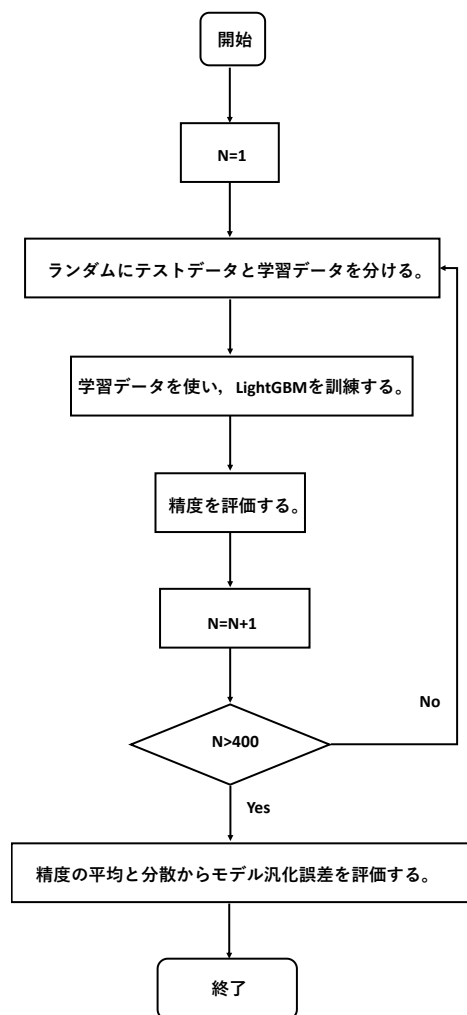


Figure 3-11 LightGBM learning algorithm without feature engineering and without SMOTE.

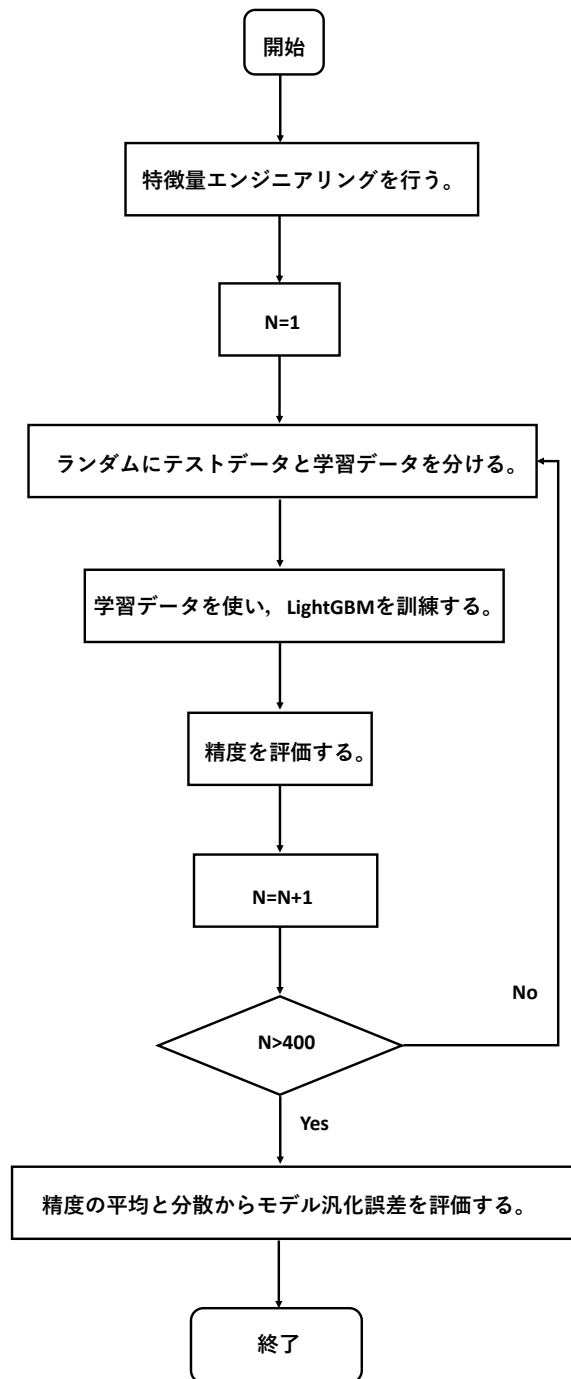


Figure 3-12 LightGBM learning algorithm with feature engineering and without SMOTE.

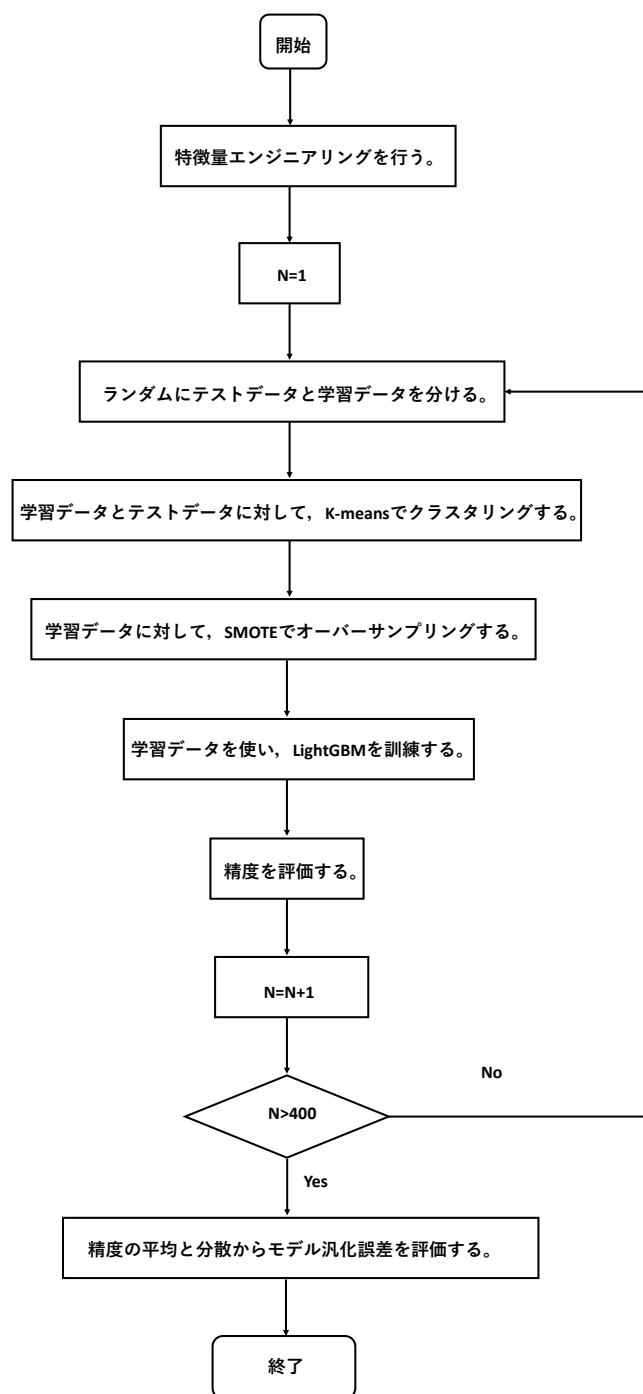


Figure 3-13 LightGBM learning algorithm with feature engineering and SMOTE.

### 3.4.3 ランダムフォレストを用いた推力推定

グリッドサーチと交差検証でハイパパラメータをチューニングする。Table 3-13 にランダムフォレストでサーチするハイパパラメータの範囲を示している。

Table 3-13 Hyperparameter grid search range for random forest.

ハイパパラメータ名	サーチ範囲
min_samples_leaf	5,10,20,30
max_depth	5,10,50,100

ランダムフォレストの計算コストは LightGBM と比べると高いため、本研究では汎化誤差の評価は 30 回ランダムにデータを選ぶ。誤差評価は次の二つのパターンを行う。

- 1) 30 回のテスト結果の RMSE と、RMSE の標準偏差を求める。
- 2) テストデータに対して、一サイクルの推定値の平均推力を求める。また、各サイクルの噴射の平均推力の RMSE と、RMSE の標準偏差を求める。

それぞれのアルゴリズムを Figure 3-14 から Figure 3-17 にまとめた。

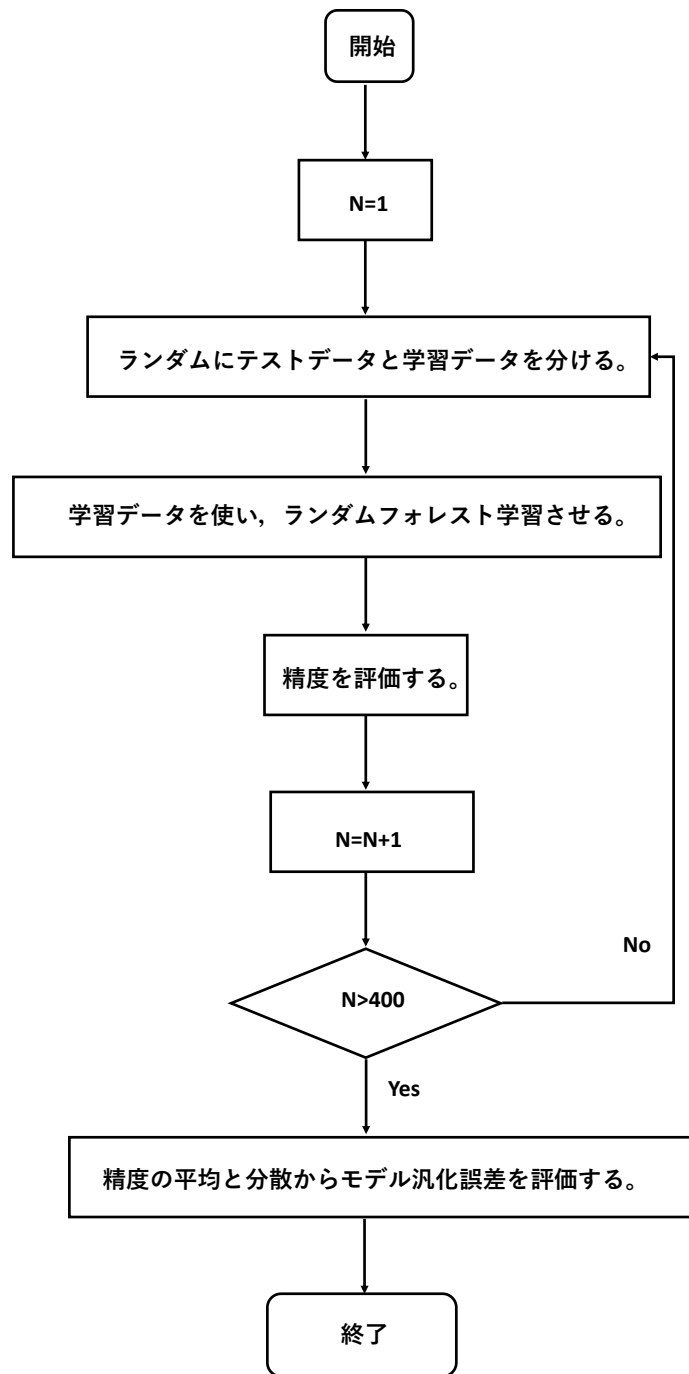


Figure 3-14 Random forest learning algorithm without feature engineering and without SMOTE.



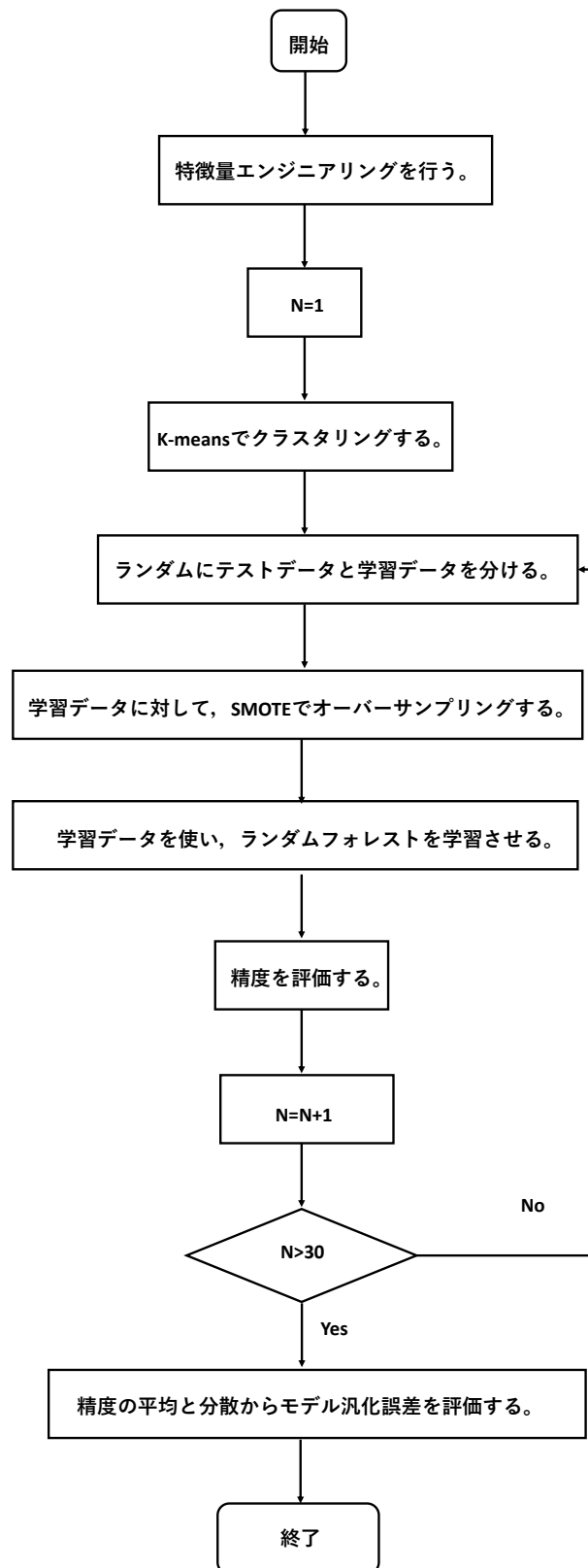


Figure 3-15 Random forest learning algorithm with feature engineering and SMOTE.

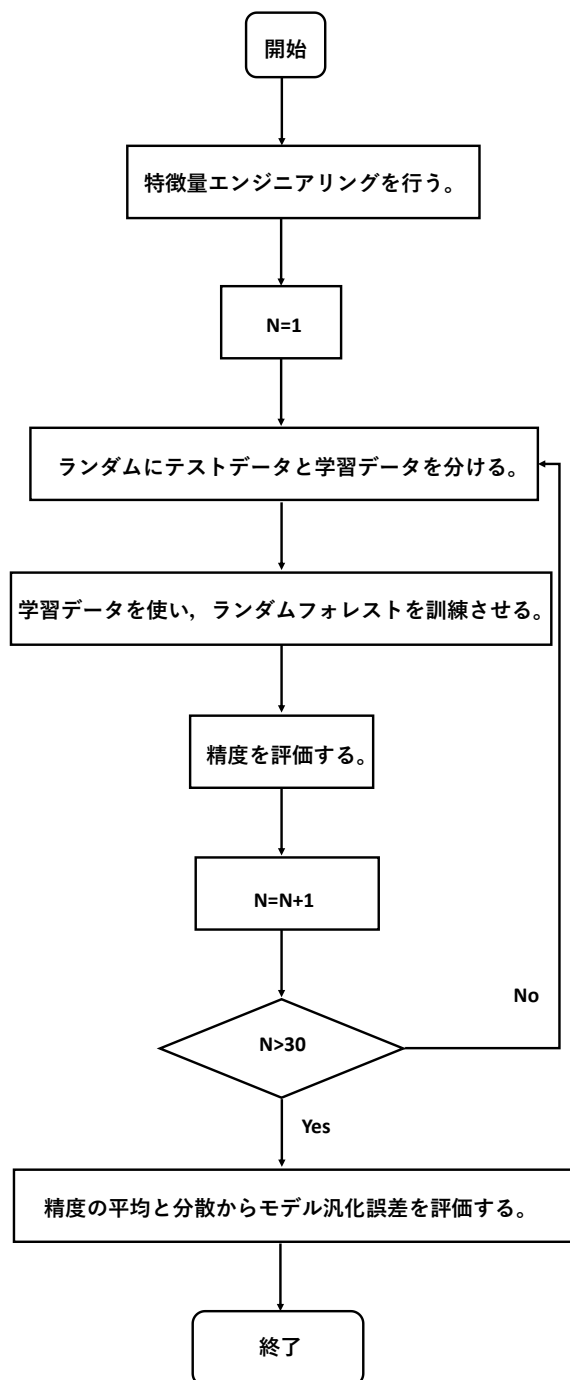


Figure 3-16 Random forest learning algorithm with feature engineering and without SMOTE.

### 3.5 機械学習を用いた気化室圧力推定

#### 3.5.1 蒸発モード，非蒸発モードの判断

推進システム作動時に，気化室は蒸発モードの状態なのか，非蒸発モードの状態なのかは気化室圧力変動と直結している．圧力計の最大誤差は 0.2kPa であるため，本研究では気化室圧力が 0.2kPa 以下の時に，非蒸発モード，0.2kPa 以上の時に蒸発モードとする．気化室圧力を推定する際に，圧力情報が不明を仮定するため，本研究では提案するのは，まず蒸発モードかどうかを決定木のランダムフォレストで判断をする．そして，蒸発モードのみに機械学習を用いて圧力を推定する．非蒸発モードの場合圧力を 0kPa とする．モード判断の機械学習モデルでは Table3-14 のような特徴量を選定する．

Table 3-14 Features used for mode judgment.

DVT1 を使用する場合に蒸発モードかどうかを推定する場合用いた特徴量	DVT2 を使用する場合に蒸発モードかどうかを推定する場合用いた特徴量
Time/s	Time/s
DTS1/°C	DTS2/°C
VTS2/°C	VTS2/°C
VTS1/°C	VTS1/°C
WPS/kPa	WPS/kPa
VH/W	VH/W
DVT1_PH/W	DVT2_PH/W
mass/g	mass/g
DTS1_shift(n)/kPa	DTS2_shift(n)/kPa
VTS1_shift(n)/°C	VTS1_shift(n)/°C
VTS2_shift(n)/°C	VTS2_shift(n)/°C
DTS1_rolling/°C	DTS2_rolling/°C
VTS2_rolling/°C	VTS2_rolling/°C
VTS1_rolling/°C	VTS1_rolling/°C
VH_rolling/J	VH_rolling/J
DVT1_PH_rolling/J	DVT2_PH_rolling/J
viscosity/Pas	viscosity/Pas
Re	Re
VH_integral/J	VH_integral/J
DVT1_PH_integral/J	DVT1_PH_integral/J

データセットの 80%を学習データとしてランダムに選び、残りの 20%をテストデータとして、モデルの精度を検証する。偽陽率 $FPR$ と真陽率 $TPR$ を用いる。それぞれの定義は以下となる。

$$FPR = \frac{FP}{FP + TN} \quad (43)$$

$$TPR = \frac{TP}{TP + FN} \quad (44)$$

ここで、 $TP$ 、 $FP$ 、 $FN$ 、 $TN$ を Table 3-16 にまとめる。 $TP$ は真陽性、推定結果が蒸発モード、現実では蒸発モードのデータの数を意味している。 $FP$ と偽陽性、推定結果が蒸発モード、現実では非蒸発モードのデータの数を意味している。 $FN$ とは、推定結果が非蒸発モード、現実では蒸発モードのデータの数を意味している。 $TN$ とは、推定結果が蒸発モード、現実では非蒸発モードのデータの数を意味している。

Table 3-15 Definition of True positive  $TP$ , False positive  $FP$ , False negative  $FN$ , True negative  $TN$

		現実	
		蒸発モード	非蒸発モード
推定	蒸発モード	$TP$	$FP$
	非蒸発モード	$FN$	$TN$

ハイパパラメータの選定は 3.4 の推力推定と同じサーチ範囲と初期値のグリッドサーチと交差検証を用いる。

### 3.5.2 気化室圧力の推定

気化室圧力を推定する。LightGBM とランダムフォレストを用いて、蒸発モード下の気化室圧力（VPS1）のデータを推定する。特徴量エンジニアリングの効果と SMOTE によるオーバーサンプリングの効果を検証するために、Table 3-16 のように、2つのモデル共に以下3つの推定を行った。アルゴリズムは 3.4 と同じである。

Table 3-16 Cases of vaporization chamber pressure estimation.

	特徴量エンジニアリング	オーバーサンプリング
CASE1	なし	なし
CASE2	あり	なし
CASE3	あり	あり

ハイパパラメータの選定は 3.4 の推力推定と同じサーチ範囲と初期値のグリッドサーチと交差検証を用いる。また、気化室圧力、ノズル前圧力が未知であることを仮定するため、特徴量エンジニアリング後の特徴量としては Table 3-17 のように入力用の特徴量を使用する。

Table 3-17 Features used to estimate vaporization chamber pressure (VPS1).

DVT1 を使用する場合 VPS1 センサーデータを推定する場合用いた特徴量	DVT2 を使用する場合 VPS1 センサーデータを推定する場合用いた特徴量
Time/s	Time/s
DTS1/°C	DTS2/°C
VTS2/°C	VTS2/°C
VTS1/°C	VTS1/°C
WPS/kPa	WPS/kPa
VH/W	VH/W
DVT1_PH/W	DVT2_PH/W
mass/g	mass/g
DTS1_shift(n)/kPa	DTS2_shift(n)/kPa
VTS1_shift(n)/°C	VTS1_shift(n)/°C
VTS2_shift(n)/°C	VTS2_shift(n)/°C
DTS1_rolling/°C	DTS2_rolling/°C
VTS2_rolling/°C	VTS2_rolling/°C
VTS1_rolling/°C	VTS1_rolling/°C
VH_rolling/J	VH_rolling/J
DVT1_PH_rolling/J	DVT2_PH_rolling/J
DVT1_PH_sum/J	DVT2_PH_sum/J
VH_sum/J	VH_sum/J
viscosity/Pas	viscosity/Pas
mass_left/g	mass_left/g
VH_integral/J	VH_integral/J
DVT1_PH_integral/J	DVT1_PH_integral/J

---

## 第4章 解析結果

---

### 4.1 推力推定解析の結果

#### 4.1.1 K-means のクラスタリングの結果を主成分分析で可視化

Figure 4-1 は DVT1 に対して特徴量のエンジニアリングされた特徴量を K-means で分類した後、各クラスに属すデータの数の結果を示している。

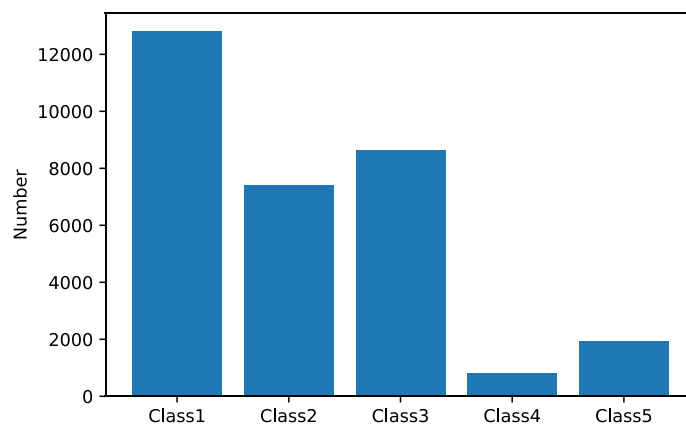


Figure 4-1 Result of K-means classification class in thrust estimation for DVT1.

Figure 4-2 は DVT2 に対して特徴量のエンジニアリングされた特徴量を K-means で分類した後、各クラスに属すデータの数の結果を示している。

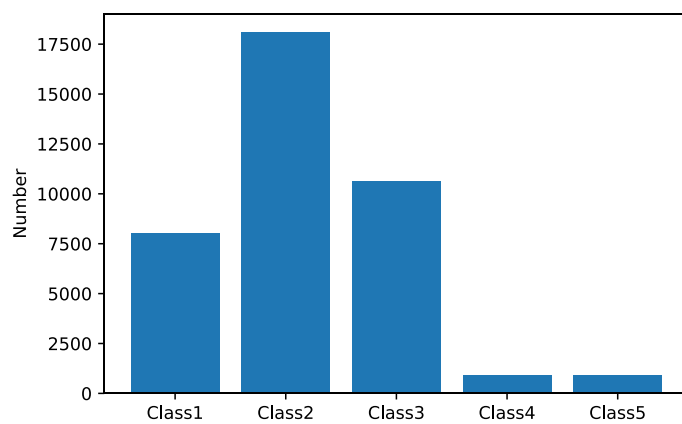


Figure 4-2 Result of K-means classification class in thrust estimation for DVT2.

Figure 4-1 と Figure 4-2 から分かれたクラスには不均一性が存在している。また、主成分分析から多次元データを圧縮することで、クラス分けを可視化することができる。DVT1 と DVT2 に対して、第一、二、三主成分の固有ベクトルの成分に絶対値取って、上位の 10 個特徴量の結果を Figure 4-3 から Figure 4-9 にまとめた。

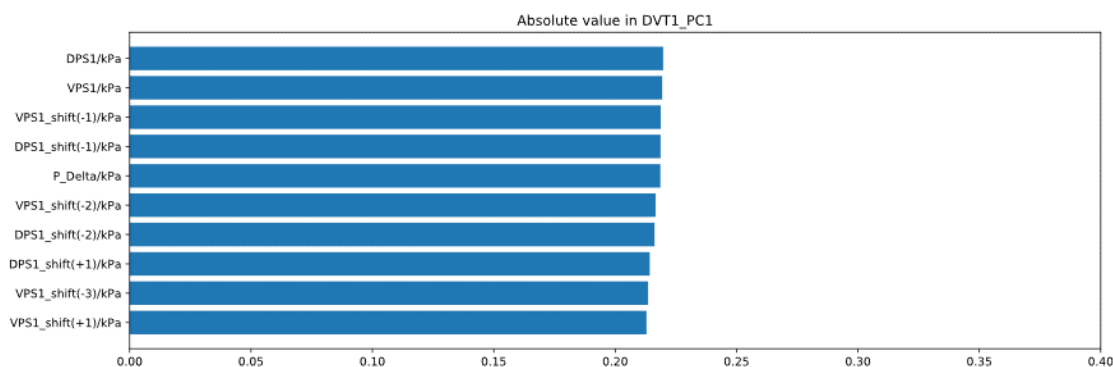


Figure 4-3 In DVT1's case, the absolute value taken for each component of the eigenvector of the first principal component, and the top 10 features and the absolute value of the component.

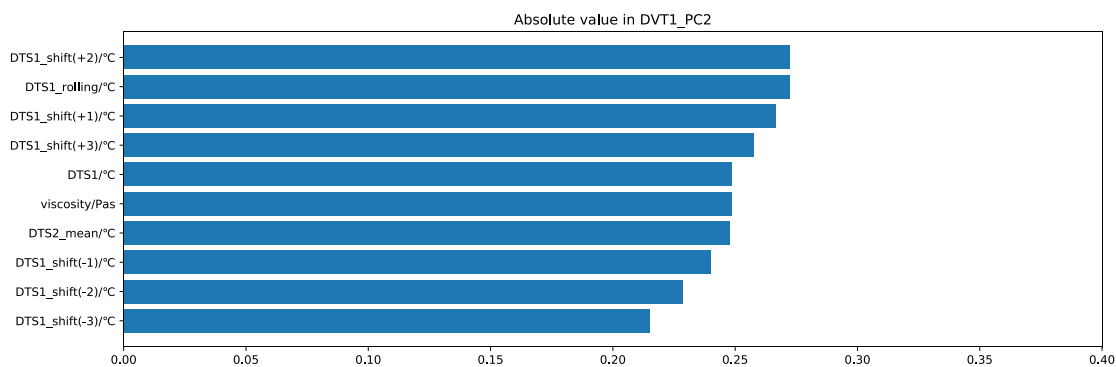


Figure 4-4 In DVT1's case, the absolute value taken for each component of the eigenvector of the second principal component, and the top 10 features and the absolute value of the component.

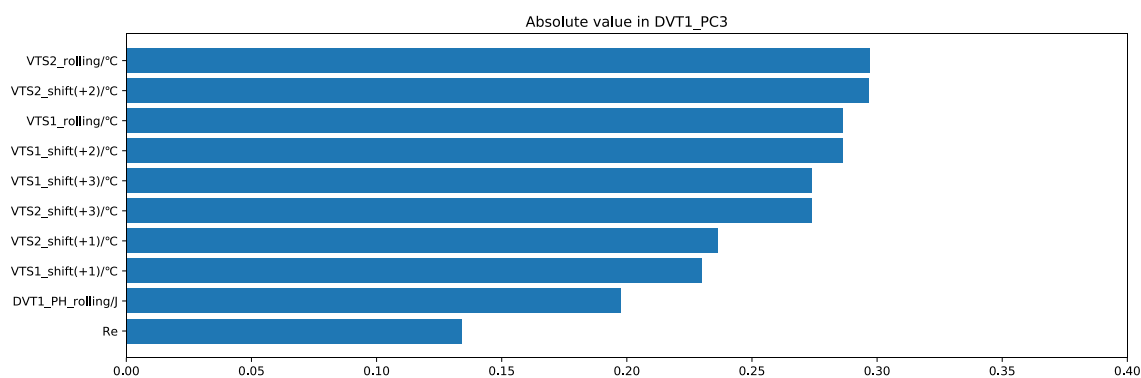


Figure 4-5 In DVT1's case, the absolute value taken for each component of the eigenvector of the third principal component, and the top 10 features and the absolute value of the component.

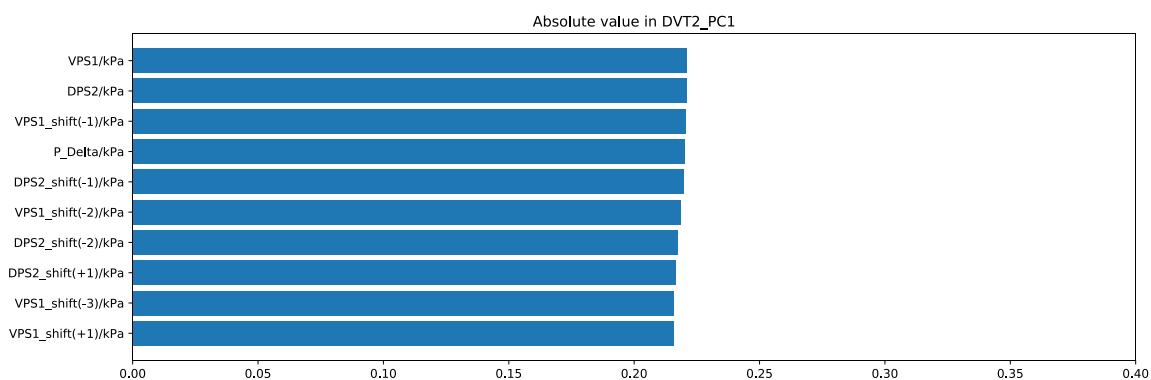


Figure 4-6 In DVT2's case, the absolute value taken for each component of the eigenvector of the first principal component, and the top 10 features and the absolute value of the component.

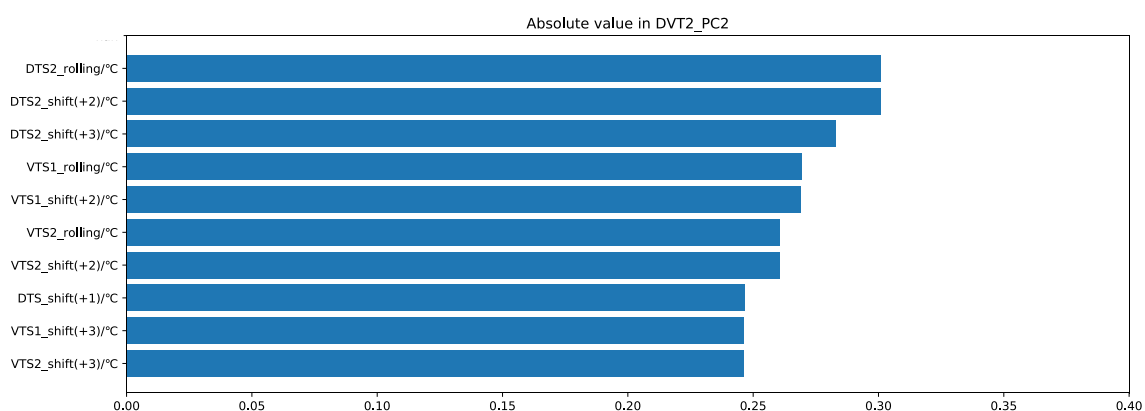


Figure 4-7 In DVT1's case, the absolute value taken for each component of the eigenvector of the second principal component, and the top 10 features and the absolute value of the component.



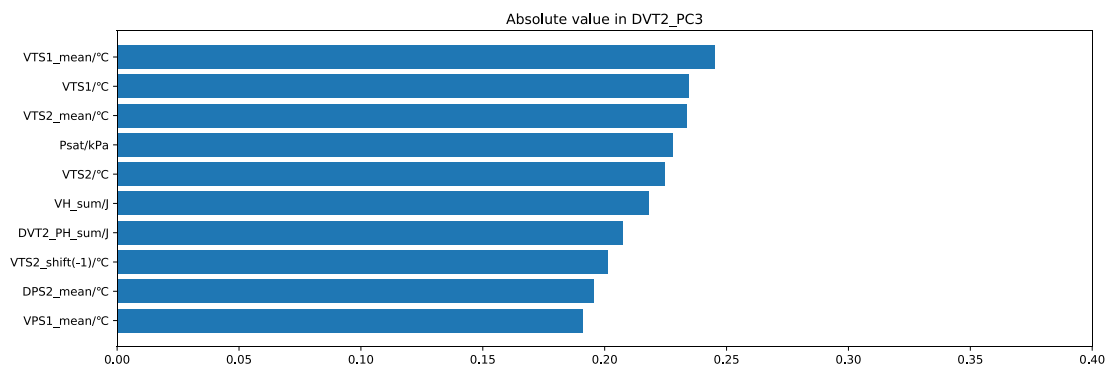


Figure 4-8 In DVT1's case, the absolute value taken for each component of the eigenvector of the third principal component, and the top 10 features and the absolute value of the component.

Figure 4-3 から Figure 4-9 の結果から、第一主成分は主に気化室圧力とノズル圧力の情報からなっている。第二主成分は主にノズル温度からなっている。粘性も入っているが、粘性はノズル温度から計算された値である。第三主成分は主に、気化室温度からなっている。以下、Figure 4-9 から Figure 4-14 は DVT1, DVT2 に対して主成分分析で得られた主成分を横軸と縦軸にした場合のグラフを示している。

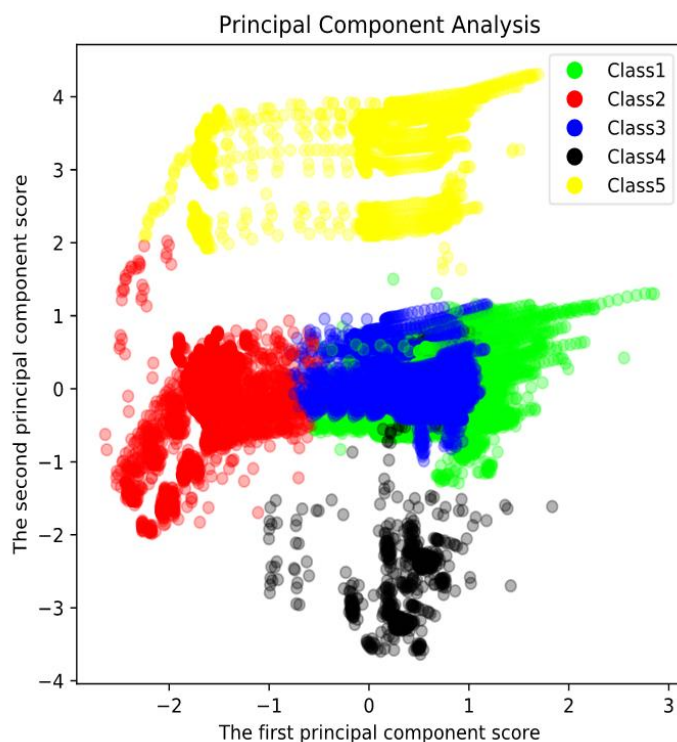


Figure 4-9 Visualization of clustering results by K-means for DVT1. The horizontal and vertical axes are the first and second principal components.

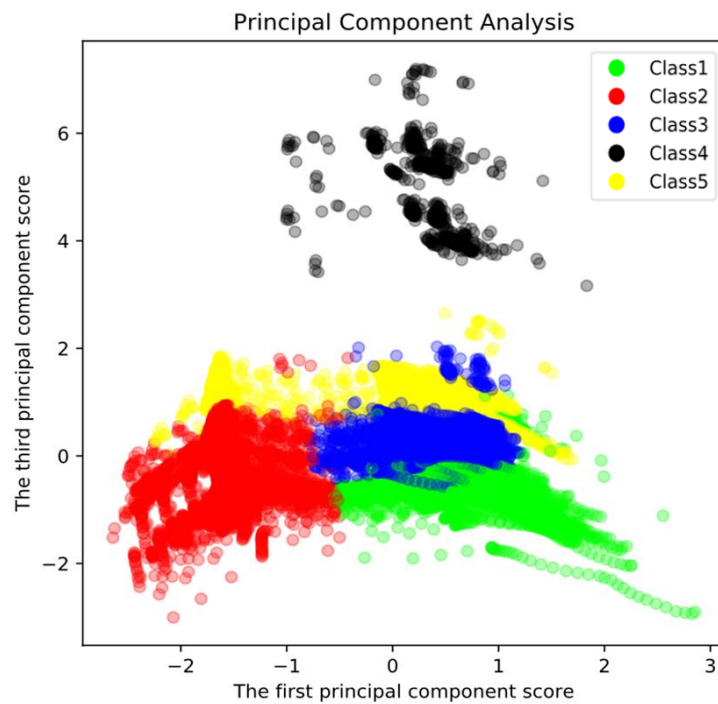


Figure 4-4-10 Visualization of clustering results by K-means for DVT1. The horizontal and vertical axes are the first and third principal components.

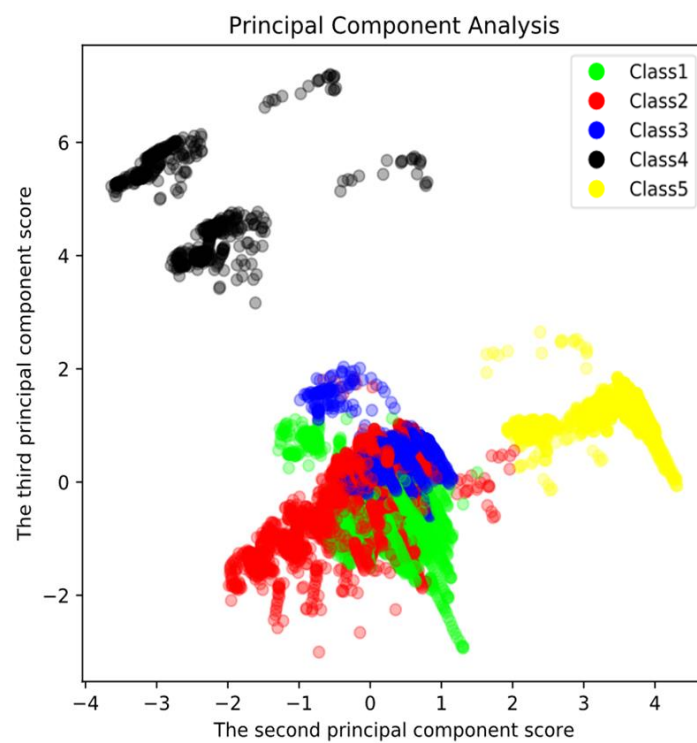


Figure 4-11 Visualization of clustering results by K-means for DVT1. The horizontal and vertical axes are the second and third principal components.

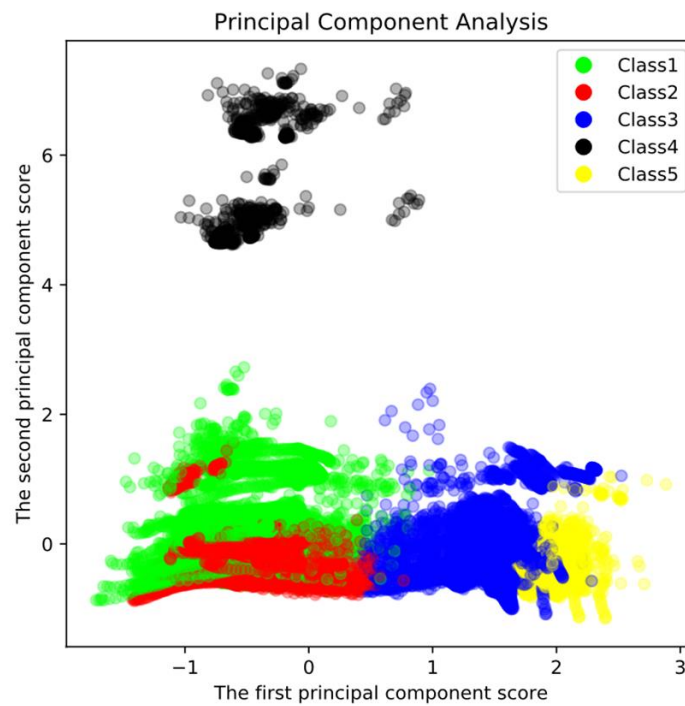


Figure 4-12 Visualization of clustering results by K-means for DVT2. The horizontal and vertical axes are the first and second principal components.

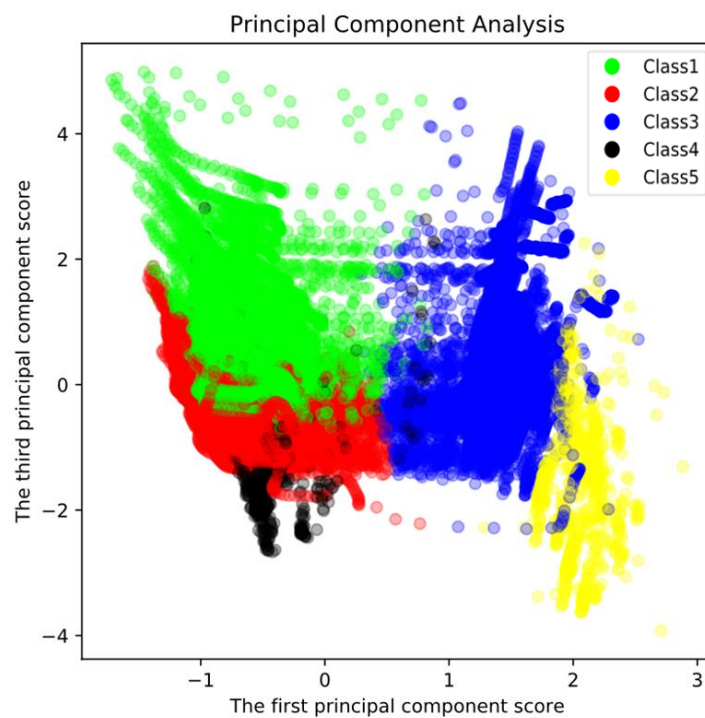


Figure 4-13 Visualization of clustering results by K-means for DVT2. The horizontal and vertical axes are the first and third principal components.

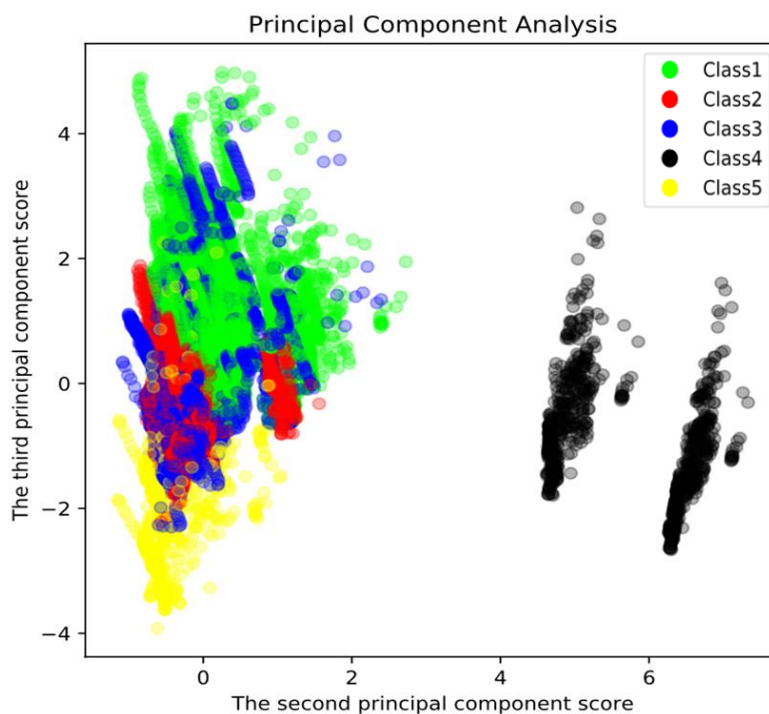


Figure 4-14 Visualization of clustering results by K-means for DVT2. The horizontal and vertical axes are the first and third principal components.

#### 4.1.2 LightGBM とランダムフォレストを用いた推力推定

簡単のために、3章で説明した3つのケースに対して、DVT1 と DVT2 ノズルを使用する場合に、ノズル名の後ろに Table 4-1 のような略称を付けた。例えば DVT1 の CASE3 において LightGBM を使用する場合に DVT1\_ SMOTE\_FE\_LGBM とする。

Table 4-1 Thrust estimation cases.

	LightGBM	ランダムフォレスト
CASE1	LGBM	RFT
CASE2	FE_LGBM	FE_RFT
CASE3	SMOTE_FE_LGBM	SMOTE_FE_RFT

また、本研究では交差検証とグリッドサーチを用いて、LightGBM とランダムフォレストのハイパラメータを選定した。Table 4-2, Table 4-3 は交差検証で選ら得たハイパラメータの結果を示している。

Table 4-2 LightGBM hyperparameter results in DVT1 thrust estimation selected by cross-validation and grid search.

ハイパパラメータ名	値
reg_alpha	3
reg_lambda	1
num_leaves	50
max_depth	10
learning_rate	0.05
feature_fraction	0.8

Table 4-3 LightGBM hyperparameter results in DVT2 thrust estimation selected by cross-validation and grid search.

ハイパパラメータ名	値
reg_alpha	2
reg_lambda	2
num_leaves	70
max_depth	10
learning_rate	0.05
feature_fraction	0.85

Table 4-2 と Table 4-3 から、それぞれ近いハイパパラメータが選定されていたことが分かった。

ランダムフォレストのチューニング結果を Table 4-7 と Table 4-8 にまとめた。

Table 4-4 Randomforest hyperparameter results in DVT1 thrust estimation selected by cross-validation and grid search.

ハイパパラメータ名	値
min_samples_leaf	5
max_depth	30

Table 4-5 Randomforest hyperparameter results in DVT2 thrust estimation selected by cross-validation and grid search.

ハイパパラメータ名	値
min_samples_leaf	5
max_depth	30

評価 1, 各時間におけるテストデータの推力と推定値推力の RMSE と RMSE の標準偏差で評価する. 結果を Figure 4-15 にまとめた.

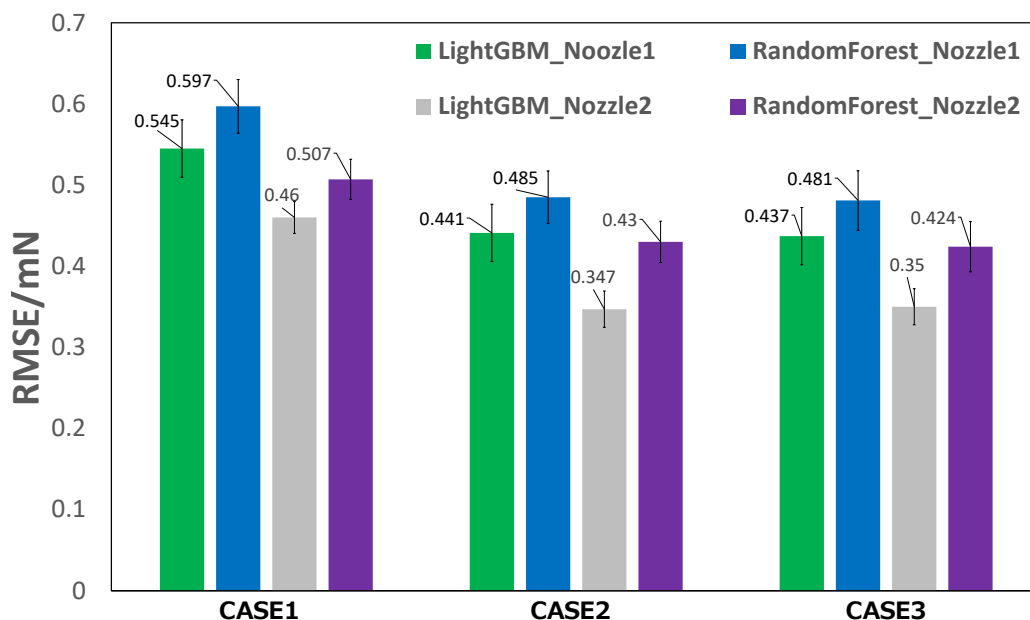


Figure 4-15 Thrust error evaluation at each time. The vertical axis shows RMSE. Error bars indicate standard deviation of RMSE.

評価 2. 各噴射サイクルの平均推力の RMSE と, RMSE の標準偏差で評価する. DVT1, DVT2 に対するの推定精度の評価 2 の結果を Figure 4-16 にまとめた

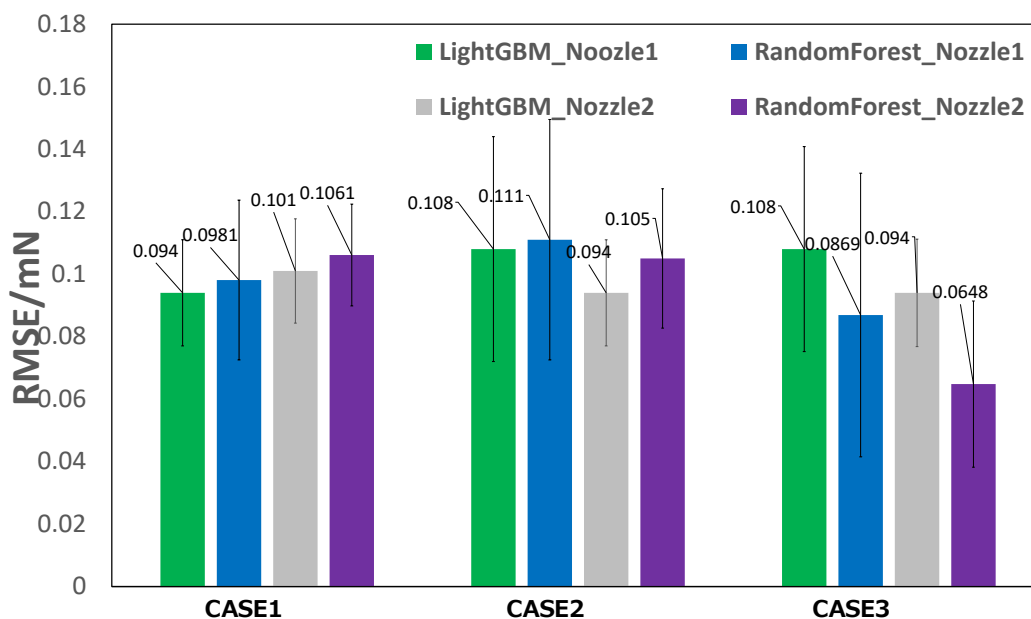


Figure 4-16 RMSE of average thrust of each injection cycle and standard deviation of RMSE.

評価1の結果から、RMSEは他の変位計から測定した推力のRMSEの0.32mNより高いが、その倍以内に抑えることができる。また、平均推力の誤差は約0.1mNであり、各サイクルにおいて、誤差の正と負がバランスよく分布していると考えられる。

#### 4.1.3 特徴量の重要度

それぞれのモデルにおいて、特徴量の重要度を計算した。最も重要度が高い特徴量をTable 4-6にまとめた。

Table 4-6 Most important feature in thrust estimation.

	DVT1	DVT2
LGBM	Time/s	Time/s
FE_LGBM	WPS/kPa	DPS2_rolling /kPa
SMOTE_FE_LGBM	WPS/kPa	TPS/kPa
RFT	VPS1/kPa	DPS2/kPa
FE_RFT	P_Delta/kPa	VPS1_rolling/kPa
SMOTE_FE_RFT	DPS1/kPa	VPS1_rolling/kPa

Table 4-6 から時間、気化室圧力、ノズル前圧力、水タンクの圧力などの特徴量の重要度が高いことが分かった。定性的にラーバルノズルの理論式と一致している。

## 4.2 気化室圧力の推定解析結果

### 4.2.1 蒸発モードと非蒸発モードの判断

グリッドサーチと交差検証によるチューニング結果、DVT1使用の場合、max\_depth: 20, min\_samples\_leaf: 50が最も適切である。DVT2使用の場合、max\_depth:10,min\_samples\_leaf: 50が最も適切である。Table 4-7はFPRとTPRでの評価結果を示している。

Table 4-7 Evaluation result of evaporation mode judgment.

評価指標	DVT1 使用	DVT2 使用
FPR	7.3%	3.3%
TPR	98.5%	99%

DVT2使用のデータ数が多い、パラメータ振り試験の実行が原因で、DVT2使用の推定精度はDVT1使用の推定精度より優れている。また、誤差が存在している原因は蒸発モードの開始時と終了時の判定が難しいからである。

## 4.2.2 LightGBM とランダムフォレストを用いた気化室圧力データの推定結果

交差検証とグリッドサーチで得られた適切なハイパパラメータを次の表にまとめた。

Table 4-8 LightGBM hyperparameter results in DVT1 vaporization chamber pressure estimation selected by cross-validation and grid search.

ハイパパラメータ名	値
reg_alpha	3
reg_lambda	0.9
num_leaves	50
max_depth	7
learning_rate	0.05
feature_fraction	0.8

Table 4-9 LightGBM hyperparameter results in DVT2 vaporization chamber pressure estimation selected by cross-validation and grid search.

ハイパパラメータ名	値
reg_alpha	2
reg_lambda	3
num_leaves	30
max_depth	9
learning_rate	0.05
feature_fraction	0.85

簡単のために、以下のように略称を付けた。

Table 4-10 Pressure estimation cases.

	LightGBM	ランダムフォレスト
CASE1	LGBM_P	RFT_P
CASE2	FE_LGBM_P	FE_RFT_P
CASE3	SMOTE_FE_LGBM_P	SMOTE_FE_RFT_P

推定結果の評価は VPS1 で測定した気化室圧力と推定値の RMSE を用いる。Figure 4-17 はそれぞれの結果を示している。



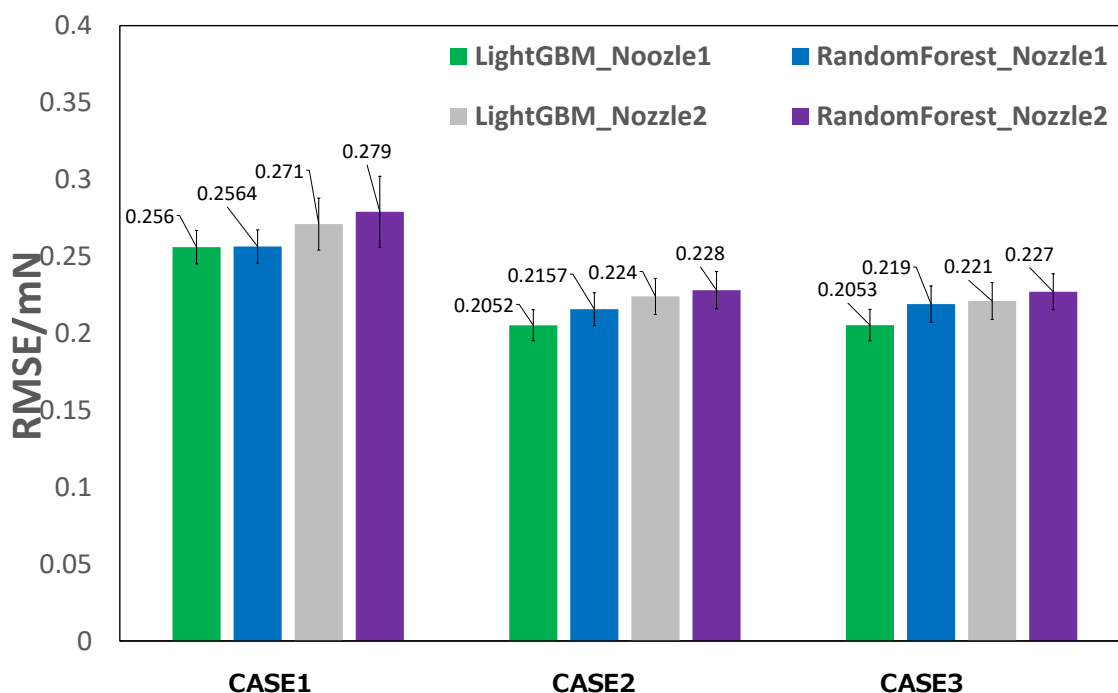


Figure 4-17 RMSE of average thrust of each injection cycle and standard deviation of RMSE.

### 4.2.3 特徴量の重要度

それぞれのモデルにおいて、特徴量の重要度を計算した。最も重要度が高い特徴量を Table 4-11 にまとめた。

Table 4-11 Most important feature in pressure estimation.

	DVT1	DVT2
LGBM	Time/s	Time/s
FE_LGBM	mass_left/g	mass_left/g
SMOTE_FE_LGBM	mass_left/g	mass_left/g
RFT	Time/s	Time/s
FE_RFT	Time/s	Time/s
SMOTE_FE_RFT	VTS2_shift(-3)/°C	mass_left/g

Table 4-11 から時間，気化室に残った水の質量，気化室温度の重要度が高い。気化室に残った水の質量は流速と時間から計算された特徴量であるため，時間の重要度が高いことと一致している。

---

## 第5章 考察

---

### 5.1 推力推定の考察

#### 5.1.1 推力推定において SMOTE の有効性評価

SMOTE の有効性については、LightGBM とランダムフォレストの評価 1 の推定結果から、SMOTE の実行による推定精度変化がない。そして、LightGBM の評価 2 から、SMOTE による推定精度の変化がない。しかし、ランダムフォレストの評価 2 から、平均推力の RMSE は 0.04 程高くなった。これは、ランダムフォレストの推力推定結果は、測定値の推力との正と負のばらつきは SMOTE を実行しない場合より均一になっていると考えられる。全体的に SMOTE を実行することで、推定精度を改善できなかった。原因としては二つ考えられる。一つ目は K-means の分類は適切ではなかった。二つ目は SMOTE によ

これらの結果から見ると、第一主成分は主に気化室圧力とノズル圧力の情報からなっている。第二主成分は主にノズル温度からなっている。第三主成分は主に、気化室温度からなっている。DVT1、DVT2 の第一、第二、第三主成分は定性的に一致している。また、K-means の可視化結果から、分割を実現できたと判断できる。以上を踏まえ、分類には問題がないと考えられる。このため、新たに SMOTE で作ったデータが適正でないため、SMOTE を使用することで、不均一性を取り除けなかった可能性はある。

#### 5.1.2 推定精度評価

評価 1 の推定結果から、以下 3 点が言えると考えられる。

- 特徴量エンジニアリングすることで RMSE は 0.1 以上向上した。
- 本研究で提案した手法で得られた推定値は、他の変位計から得られた推力との RMSE より高いが、その倍以内に抑えたことができた。更なる精度を求める場合、SMOTE の妥当性などを検討する必要がある。
- DVT1 に対しての推力推定精度は DVT2 より低い。
  - 原因としては二つ考えられる。
    - ◇ 一つ目は、モデルを学習させる時に、DVT2 の方がいい噴射回数が多い。このため、より高い精度で推定できたと考えられる。
    - ◇ 二つ目はパラメータ振り試験で全部 DVT2 を使用したため、サンプリングのレンジとしては DVT1 より広い。

評価 2 の推定結果から以下 2 点が言えると考えられる。

- サイクル平均から見ると、3者はほぼ同じぐらいの性能で推定を実現できた。
- 評価1で特徴量エンジニアリングされることで大幅改善できた一方、評価2でほぼ変わらない結果が得られた。一噴射サイクルで見ると、推定誤差を打ち消すことができると考えられる。

### 5.1.3 特徴量評価

ランダムフォレストの重要度から、LightGBMと異なる重要度が得られた。最大の原因はランダムフォレストにおける重要度評価方法はLightGBMの評価方法にある。ランダムフォレストは特徴量が分割基準として選ばれた時に、RMSEはどれ程減少するかを対象として、RMSE減少量の割合を相対重要度として計算される一方、LightGBMは特徴量が選ばれた回数を重要度として計算される。ここでいくつが特徴量の重要度について考察する。

#### ➤ WPS

LightGBMを使用する場合、ブラダ内の圧力であるWPSの重要度が一番高い。WPSの値はブラダ内に残った水の量と関係しており、物理的に考えると、RVオープン時間はWPSの値によって制御される場合、ブラダ内に残った水の量と推力の関係はない。しかし、実際の環境試験を挟んだ推力測定試験ではRVオープン時間を一定としている。このような推力測定試験において、WPSの値は気化室への噴射量と関連する。重要度から見ると、気化室への噴射量は上位1/3にあり、ランダムフォレストで推力を推定する際にも1%以上の重要度を占めているため、推力と関連する可能性はある。もう一つの原因として、ブラダ内の水の量からどの推力推定試験であるかと関連している。試験ごとに熱環境などが異なるため、これが原因で大きな重要度を占めている可能性がある。

#### ➤ DPS1, DPS2

DPS1, DPS2はそれぞれDVT1, DVT2のノズル前圧力である。ノズル前圧力によって作られ特徴量の重要度は上位にある。AQUARIUSのような低レイノルズ数の性流体を扱う推進システムの場合、推力は理論値と乖離する一方、一般的なラバール理論のノズル前の圧力が大きければ大きいほど推力が大きくなるという性質は変わらない。このため、DPS1とDPS2関連の特徴量の重要度が高いことは理論と一致している。

#### ➤ VPS1

各モデルにおいて、気化室圧力であるVPS1 或いはVPS1によって特徴量エンジニアリングされた特徴量の重要度上位にある。VPS1はDPS1とDPS2の上流の圧力である。このVPS1の圧力値変化によってDPS1とDPS2の値が変化する。チューブによる圧力損失があるため、上流のVPS1圧力は下流のDPS1, DPS2の圧力より高い、圧力変化も激しい。回帰木のような機械学習では、変動が大きな特徴量の方が情報価値高い、このためVPS1に関連する特徴量は大きな重要度を持っている。

➤ VTS1 と VTS2

VTS2 の重要度は VTS1 より低いという結果が得られた。AQUARIUS 作動時の温度履歴を見ると、VTS2 の温度変化は VTS1 の温度変化より激しいことがわかる。VTS2 は VTS1 より細かく気化室内部の温度変化を観測できるため、大きな情報量を持っている。このように冗長系の重要度を解析することで、冗長系の位置設定において役に立つと考えられる。

➤ Time

各モデルで RV オープンからの経過時間大きな重要度を占めている。原因としては三つ考えられる。一つ目は蒸発モードの開始と終了の間に、スタンドの変位の減衰を学習したことである。本研究は推力測定値に対し、このような減衰に対して処理していない。開始の時に、推定値の値と測定値を一致させるために、機械学習モデルはこの部分も学習する。このため、時間が分岐基準として選ばれた。二つ目は、蒸発モードと非蒸発モードが存在することである。蒸発モードにおいて、推力は発生する。非蒸発モードではほとんど推力は出ない。経過時間で蒸発モードかどうかをある程度評価できるため、重要度が高い。三つ目は蒸発モード中に二段階の蒸発が存在することである。Figure 5-1 のように、蒸発モードの途中で推力が若干落ちた。これは二つの蒸発部が存在すると考えられる。一つの蒸発部で蒸発し切ったらモード変化がある。気化室内に残った水の量と RV オープンからの経過時間などの特徴量から識別できると考えられる。今後物理モデルを改善する際に、このような現象も視野に入れる必要があると考えられる。

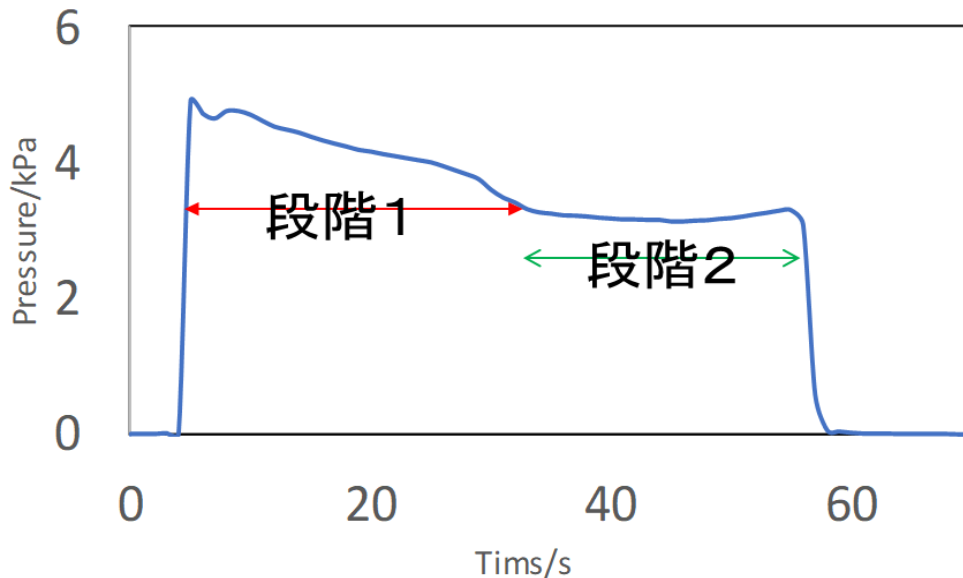


Figure 5-1 Two steps evaporation.

## 5.2 気化室圧力データ推定の考察

### 5.2.1 蒸発モード判断への考察

RV オープンからの経過時間とノズル、気化室ヒーターへの印加電力の重要度が高いという結果が得られた。原因はヒーターからの入熱は蒸発時間に大きく影響すると考えられる。また、真陽率が約 99% であることから、推定を実現できたと考えられる。確認したところ、蒸発開始と蒸発終了の 0.2kPa 前後のデータの分類が間違えた。

### 5.2.2 液滴蒸発の物理モデルとの比較する

物理モデルから演繹的に気化室圧力を算出することができる。ここで西井が提案した蒸発モデルを使う<sup>33</sup>。まず、流量のコンダクタンスを次のように定義する。

$$C_t = \frac{\int_0^{T_e} P_{vc} dt}{m} \quad (45)$$

流量は次のように計算することができる。

$$\dot{m} = P_{vc} C_t \quad (46)$$

そして、Bolton の式を用いて、飽和水蒸気圧力は

$$P_{sat} = 6.1112 \exp\left(\frac{17.67T_s}{T_s + 243.5}\right) \quad (47)$$

である。 $T_s$  は水の表面温度。推力推定において VTS2 の重要度が VTS1 高いため、ここで、VTS2 の値を表面温度とする。飽和水蒸気圧力から水の蒸発速度は次のように計算できる。

$$\dot{m}_{ev} = S_L \sigma (P_{sat} - P_{vc}) \sqrt{\frac{1}{2\pi R(273.15 - T_s)}} \quad (48)$$

ここで、 $\sigma$  は蒸発係数である。文献値によると、0.03-0.15 程度ばらついている<sup>34</sup>。本研究では 0.03, 0.06, 0.09, 0.12, 0.15 の 5 パターンを試した。

液滴は半球の形として仮定すると、液滴の表面積  $S_L$  は

$$S_L(t) = 2 \left( \frac{3m(t)\sqrt{\pi}}{4\rho_L} \right)^{\frac{2}{3}} \quad (49)$$

として計算できる。蒸発速度と流量が一致すると仮定すると、気化室圧力は次のように計算することができる。

$$P_{vc} = \frac{P_{sat}}{1 - \frac{C_t}{S_L \sigma} \sqrt{2\pi R(273.15 - T_s)}} \quad (50)$$

気化室が既知である場合に、 $C_t$ の計算はできるが、一方気化室圧力が未知の場合 $C_t$ を計算することができない。本研究は以下二つのパターンの計算を行った。

パターン1. ランダムに30回、DVT1, DVT2に対してそれぞれ噴射データ80%の噴射IDを選び、これらの噴射データの $C_t$ 平均値を計算する。残りの20%噴射に対して、 $C_t$ 平均値を使い、気化室圧力を算出する。評価基準はRMSEを用いる。

パターン2. ランダムに30回、DVT1, DVT2に対してそれぞれ噴射データ80%の噴射IDを選び、それぞれの噴射において気化室圧力が既知である条件で $C_t$ を計算し、気化室圧力を算出する。評価基準はRMSEを用いる。

計算結果を Figure 5-2 から Figure 5-5 にまとめた。

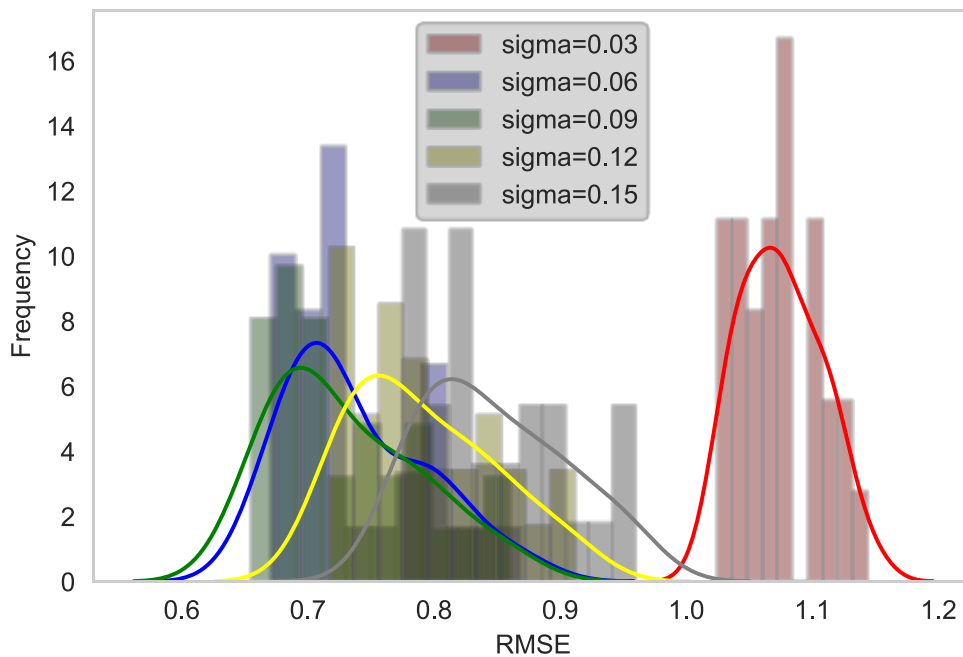


Figure 5-2 RMSE histogram of measured and estimated values of vaporization chamber pressure when using injection data of pattern 1 and DVT1. The curve means the kernel estimation curve.

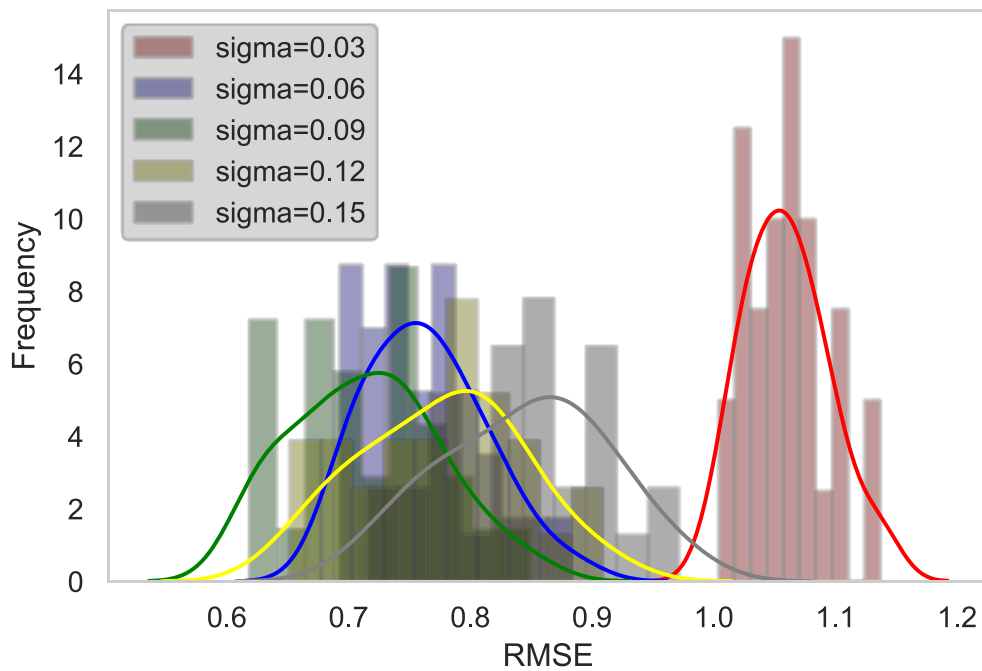


Figure 5-3 RMSE histogram of measured and estimated values of vaporization chamber pressure when using injection data of pattern 1 and DVT2. The curve means the kernel estimation curve.

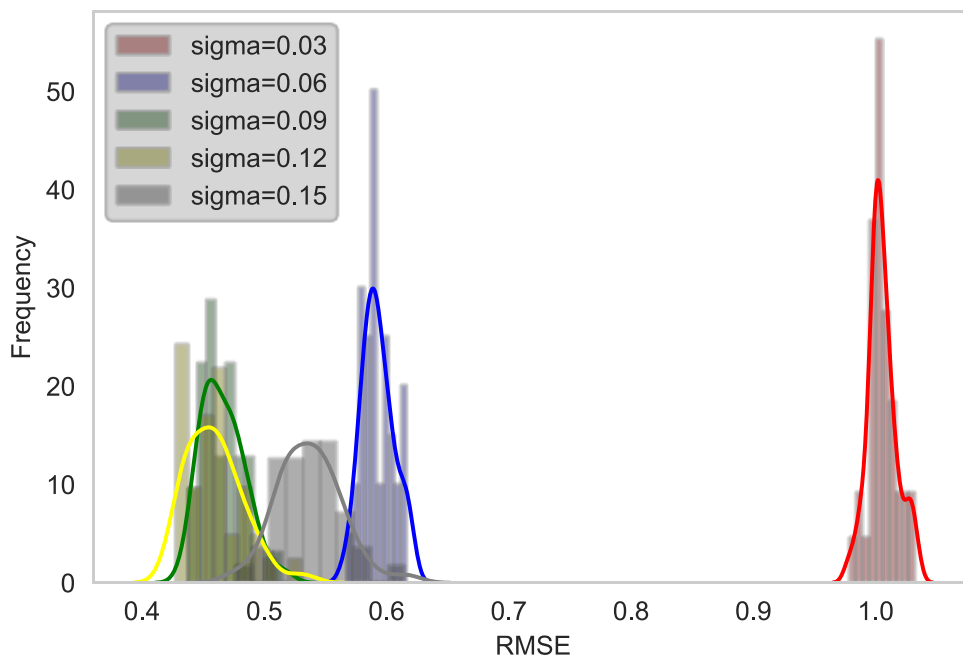


Figure 5-4 RMSE histogram of measured and estimated values of vaporization chamber pressure when using injection data of pattern 2 and DVT1. The curve means the kernel estimation curve.

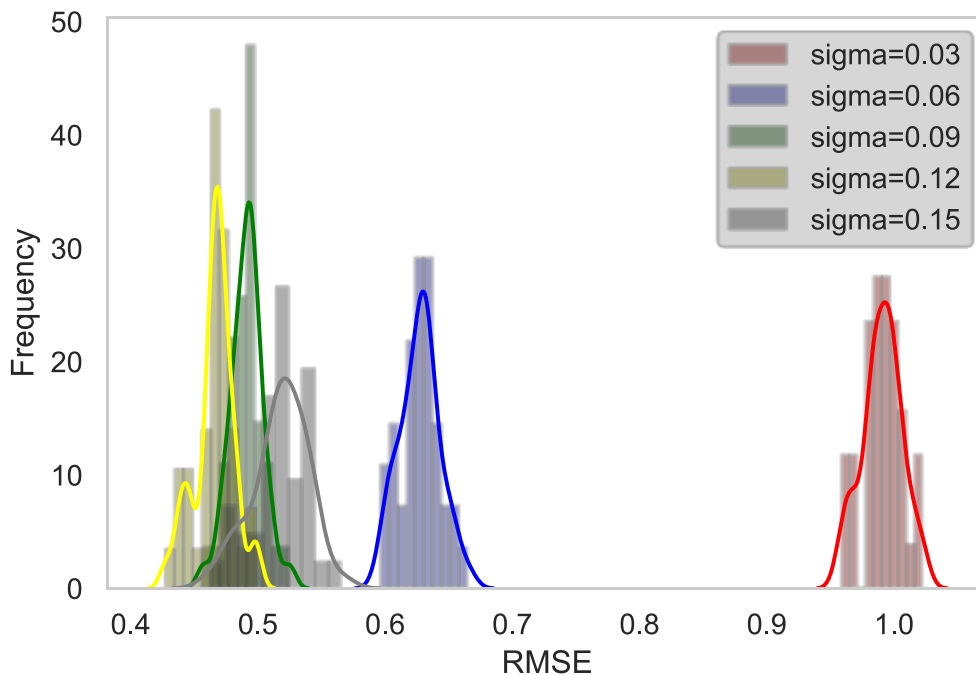


Figure 5-5 RMSE histogram of measured and estimated values of vaporization chamber pressure when using injection data of pattern 2 and DVT2. The curve means the kernel estimation curve.

計算結果から、 $\sigma$  が 0.09 と 0.12 の場合、RMSE が小さい。また、パターン 1 とパターン 2 の推定精度の RMSE は 0.2 程違いがある。このため、 $C_t$  の変化により、推定精度が大きく変わることが分かった。また、 $\sigma$  を変えることで推定精度が大きく変わった。推定値を測定値が異なる原因はいくつが考えられる。一つ目は、このモデルでは二つの蒸発部を考えていなかった。二つ目は、いくつかの仮定がモデルに入っているため、推定に誤差がある。例えば、VTS2 の温度は水の方面温度と同じ仮定、水滴の形は半球である仮定、一サイクルの噴射で  $C_t$  は変わらないなどの仮定がある。

ランダムフォレストと LightGBM は圧力情報一切使わなくても RMSE が 0.22 程度実現できた。また、VPS2 を実装した場合の試験において、推定値の代わりに、VPS2 の値を使い RMSE を計算すると、RMSE は 0.084 である。機械学習による推定値は VPS2 ほど冗長系として使えないが、予備としては使えると考えられる。そして、推定値から、推定値の RMSE を信頼区間として設けることで、実際作動時に、気化室或いは圧力に異常があるかどうかを検知できると考えられる。

物理理論からの計算値と実際の値の RMSE、本研究で提案した推定値の中で一番精度が優れているモデルと比較するグラフを Figure 5-6 にまとめた。



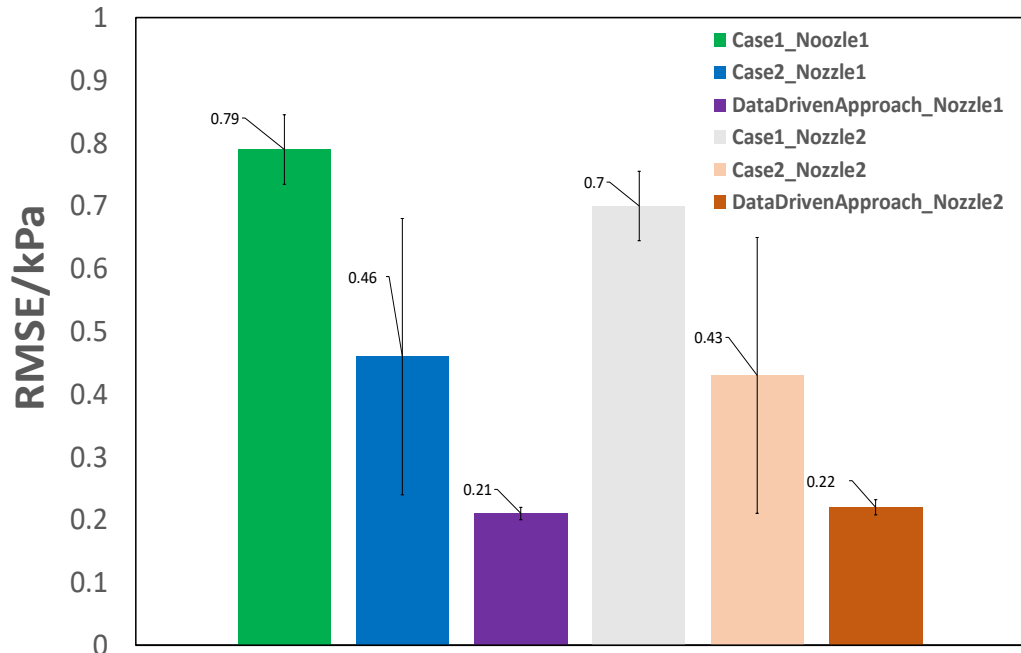


Figure 5-6 Error comparison between physical model estimates and data-driven approaches.

また、一部の試験では2つの圧力センサーを使用した、それらの実験データにおいて、センサー間の RMSE は約 0.32。データ駆動型アプローチによる推定値の誤差はそれ以内に抑えることができる。

### 5.2.3 特徴量評価

それぞれの推定モデルから得られた特徴量の重要度から、Time, mass\_left と VTS2 の重要度が高いことがわかった。これは推力推定において、Time の重要度が高い結果と一致する。気化室二つの蒸発部の存在するため、蒸発の経過時間によって、ある程度どの蒸発部であるかを判断できると考えられる。気化室内部の残った水の量が重要な理由は残った水の量は液滴の表面積とつながるため、式 (48) から、気化室圧力は液滴の表面積と関係している。このため、mass\_left は大きな重要度を占めていると考えられる。また、VTS1, VTS2 の温度情報は液滴の表面温度と関係するため、これらに関連する特徴量の重要度が高い。特に、VTS2 の変動が激しいため、大きな情報量を持っている。このため、VTS2 は VTS1 より高い重要度を占めている。

---

## 第6章 結論

---

AQUARIUS のような希薄粘性流体を扱う水レジストジェットスラスタでは、理論計算からの推力推定が難しい。また、超小型衛星にはスペースが限られているため、多くの冗長系を設置することが難しい。本研究ではデータ駆動型アプローチによる推定方法を提案した。推力の推定において、LightGBM とランダムフォレストを用いることで、推力を推定した。物理現象と時系列特有な性質を考慮することで、推定精度を大幅改善できた。

冗長系の課題に対して、本研究では、推進システムの一部の圧力計が使えないのを仮定した上で、別のセンサーデータから圧力データを推定した。また、物理モデルからの計算方法と比較した。

本研究の結論は以下の通りである。

- 本研究で提案したデータ駆動型アプローチによる推力推定方法は、RMSE0.4mN 程度、水レジストジェットスラスタの推力を推定できる。
- 推進システムの特徴に合わせて特徴量エンジニアリングすることで、推力推定精度を RMSE0.1mN 程度向上できる。
- データ駆動型アプローチにより、一部のセンサーデータを用いて、欠損データを推定できる。
- 選定した木系の機械学習モデルから得られた特徴量の重要度を物理現象と比較し、定性的に一致していると考えられる。
- 提案した圧力推定手法は物理モデルから推定した圧力より精度が優れている。

---

## 参考文献

---

1. Charles Murray, Catherine Bly Cox. *Apollo*. South Mountain Books, 2004
2. Dickson, Paul, *Sputnik: The Shock of the Century*, Walker & Company, 2007.
3. “Salyut 1”. NASA - NSSDC.2019年12月31日閲覧.
4. “旧ソ連の火星探査機「マルス3号」とみられる物体を発見”. AstroArts. (2013年4月16日)2019年12月31日閲覧.
5. Poghosyan, A. & Golkar, A. ‘CubeSat evolution: Analyzing CubeSat capabilities for conducting science missions’. *Progress in Aerospace Sciences* 88(2017), pp. 59–83.
6. “人工的に「流れ星」作り出す 超小型衛星打ち上げ”.NHK NEWS WEB(2019年12月6日). 2019年12月31日閲覧.
7. “欧州発、進化する宇宙ベンチャー”.そらこと.(2019年2月13日)2019年12月31日閲覧.
8. “中国の宇宙産業が活況、民間スタートアップが次々に参入”(2019年9月17日)2019年12月31日閲覧.
9. SpaceWorks',2018 Nano/Microsatellite Market Forecast C Williams, B Doncaster, J Shulman, 2018.
10. “GS ユアサの国際宇宙ステーション用リチウムイオン電池の第2回打ち上げが決定”.GSYUASA,(2018年9月6日),2020年1月6日閲覧.
11. Y Kobayashi *et al.* 'X-band, 15-W-class, highly efficient deep-space GaN SSPA for PROCYON mission', *IEEE Transactions on Aerospace and Electronic Systems*,52,3 (2016), pp. 1340-1351.
12. 土方 嘉徳「第3次人工知能ブームのすべて」『エコノフォーラム21：学生と教職員のインターコ ミュニケーション誌』25号, 2019年.
13. 平井 有三「初めてのパターン認識」, 2012.
14. 吉原 篤「深層学習を用いた被災地衛星画像の被覆分類」『神戸大学都市安全研究センター研究報告』第22号, 2018年.
15. Xiaowei Jia *et al.* ‘Automated Monitoring Cropland Using Remote Sensing Data: Challenges and Opportunities for Machine Learning’. Cornell University, 2019.
16. Yong Hoon Kim *et al.*, ‘Machine learning approaches to coastal water quality monitoring using GOCI satellite data’, *GIScience & Remote Sensing*, 51, 2014.
17. Lachlan Gunn *et al.* ‘Anomaly Detection in Satellite Communications Systems using LSTM Networks’,2018 Military Communications and Information Systems Conference, MilCIS 2018 – Proceedings, 2018.
18. ShahrozTariq *et al.*, ‘DetectingAnomaliesinSpaceusingMultivariateConvolutional LSTMwithMixturesofProbabilisticPCA’, *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 2123-2133.

19. Takehisa Yairi.*et al.* ‘A Data-Driven Health Monitoring Method for Satellite Housekeeping Data Based on Probabilistic Clustering and Dimensionality Reduction’,*IEEE Transactions on Aerospace and Electronic Systems*,53, (2019),pp. 1384-1401.
20. 菊谷侑平,「深層学習を用いた革新的地球センサ・スタートラッカ「DLAS」の開発」第18回宇宙科学シンポジウム(2018).
21. 淵上「ホールスラスタによけるニューラルネットワークを用いた放電電流予測」『日本航空宇宙学会論文集』66,5,143-145(2018).
22. ASAKAWA, J. *et al.* ‘Fundamental Ground Experiment of a Water Resistojet Propulsion System: AQUARIUS Installed on a 6U CubeSat: EQUULEUS’. *TRANSACTIONS OF THE JAPAN SOCIETY FOR AERONAUTICAL AND SPACE SCIENCES, AEROSPACE TECHNOLOGY JAPAN* 16(2018),pp. 427–431.
23. ASAKAWA, J. ‘Water Micro-Propulsion System Using Room-temperature Evaporation : Thrust Performance and its Application to Spacecraft’. 2019 Doctor Thesis
24. ASAKAWA, J. ‘Water Micro-Propulsion System Using Room-temperature Evaporation : Thrust Performance and its Application to Spacecraft’. 2019 Doctor Thesis
25. SVANTE WOLD. *et al.* ‘Principal Component Analysis’. *Chemometrics and Intelligent Laboratory Systems*. 2,1,3, pp.37-52.
26. J. MacQueen, ‘Some Methods for classification and Analysis of Multivariate Observations’,*Proc. of the Fifth Berkeley Symposium on Math., Stat. and Prob.*, 1(1967),pp.281-296..
27. Nitesh V. Chawla. *et al.* ‘Synthetic Minority Over-sampling Technique’, *journal of Artificial Intelligence Research*,16(2002), pp.321-357.
28. J.R. QUINLAN. ‘Classification and regression trees’, *Machine Learning* , 1,1(1986), pp.81-106.
29. Leo Breiman. ‘Random Forests’, *Machine learning*, 45,1(2001),pp.5-32,.
30. Guolin Ke. *et al.* ‘LightGBM: A Highly Efficient Gradient Boosting Decision Tree’, *NIPS*, (2017) pp.3149–3157.
31. Microsoft. LightGBM Documentation, 2018. <https://lightgbm.readthedocs.io/en/latest/Experiments.html>.
32. Jui-Sheng Chou. *et al.* ‘Time series analytics using sliding window metaheuristic optimization-based machine learning system for identifying building energy consumption patterns’,*Applied Energy*,177,1(2016),pp.751-770.
33. Nishii, K. ‘Vaporizing Mechanisms and Micro-nozzle flows of the Water Resistojet Thruster for CubeSats’. 2018 Master Thesis
34. Persad, A. H., and Ward, C. A., “Expressions for the Evaporation and Condensation Coefficients in the HertzKnudsen Relation,” *Chemical Reviews*, 116, 2016, pp. 7727–7767.

## 謝辞

私は小さい頃から宇宙機や飛行機が好きで、いつか自分が宇宙開発に参加したいと思っています。留学生はあまり宇宙開発系のプロジェクトに参加させないと聞いたことがありますが、研究室在籍二年間、水推進システムAQUARIUSの開発に参加させていただきました。掛け替えのない経験でした。最初に研究室に入った時に、何もわからなかったし、みなさんに迷惑をかける一方でした。それにも関わらず、丁寧に色々教えていただきました。2012年から来日した以来、一番成長を感じたのは小泉研の二年間でした。

多大なるご指導・ご助言を賜りました小泉宏之准教授に深く感謝しております。毎週のMTGで推進に関する専門的な知識、電気回路や数値計算の知識など、常に新しい知識を教えいただけます。さらに、質問の方法、問題分析の方法と質問への回答方法なども身に付けました。このようなディスカッションミーティング場を提供していただき、ありがとうございます。

MTGを通して、私は多思考習慣を学び、自分自身の多くの欠点にも気づきました。これらの指導は私の成長を促進し、これらの習慣はこれから社会人になる私に非常に役に立ちます。

小紫教授と川嶋助教に深く感謝します。輪講発表で貴重なアドバイスを頂き、本論文を仕上げることができました。

水班の皆さんに感謝します。いつもAQUARIUSのプロジェクトや研究でアドバイスをしていただく浅川さんに感謝します。浅川さんのアドバイスがないと、私はAQUARIUSの組み立て手順書などの一つ一つのタスクを完成することができません。西井さんと服部さんに感謝します。彼らのリードのお陰で、プロジェクトが順調に進むことができ、組み上げと一つ一つ試験をクリアできました。西井さんは物理現象の分析、数値計算の方法、実験設備使い方など様々なことを丁寧に説明していただきました。西井さんは常に熱心で、研究だけでなく、生活の悩み相談、ワークバランスの調整方法など色々なアドバイスを私にしていました。服部さんは、いつも真剣で、プロジェクトが忙しい時期に、いつも最後まで研究室に残っていました。また、毎日の朝会で、AQUARIUSのメンバーの皆さんに細かく一つ一つのタスクを説明していただきました。服部さんから研究やプロジェクトのアドバイスだけでなく、就活についても色々アドバイスをいただきました。深く感謝します。特任研究員であった齋藤さんにも非常にお世話になりました。同期の菊池、秋山、室原に感謝します。菊池は前PIとして、服部さんが卒業後私たちをリードし、様々な試験をクリアできました。彼の口癖は「いつでもいるんで」です。個人の時間を犠牲し、プロジェクトのために、頑張った彼に感謝します。最初の1年で同じ水ハイブリット班の秋山に感謝します。秋山は粘り強い人で、彼女から研究への熱意を深く感じました。分光温度測定試験など色々協力していただき、ありがとうございます。室原に感謝します。彼は今まで私が知っている日本人の中で一番面白い、ツッコミが強い人です。彼のお陰で研究室はより楽しくなりました。また、彼のアドバイスのお陰でこの論文を仕上げられました。感謝します。後輩の井澤くん、万波くんにも感謝します。彼の協力でデータ収集が順調

に進むことができ、データベースを構築できました。そして、留学しに行った西くんに感謝します。無重力実験系の修正で色々協力していただき、ありがとうございます。

班が異なりますが、先輩の中川さん、柳沼さん、関根さん、そして吉竹さんからもたくさん研究のアドバイスを頂き、感謝します。小紫研の先輩の田中さん、松井さんから分光器の使い方など教えていただき、感謝します。特別研究員の陳さんから、実際中国での推進システム開発などの話を伺うことができ、大変勉強になりました。また、王さん、田畑さんからも色々なアドバイスを頂き、ありがとうございました。

同期の龍野、安宅、内藤からMTGと輪講発表でいつも質問やコメントをしていただき、ありがとうございました。そして、小紫研の同期の竹内、山田、神田からもたくさんアドバイスを頂き、ありがとうございました。

最後に、両親に感謝したいです。いつも私がやりたいことを支えてくれて、ありがとうございます。18歳から日本に来て、両親の支えがないと、充実した日本留学生活ができません。当然本論文を完成することができません。深く感謝します。

## 関連研究発表

2018 宇宙輸送シンポジウム 発表

- 題目：6UCubeSat：EQUULEUS 推進システム AQUARIUS の推力特性分析
- 著者：王，西井，服部，浅川，齋藤，菊池，秋山，安宅、室原，小泉，船瀬，小紫
- 概要：2020 年に打ち上げ予定の 6UCubeSat：EQUULEUS には，推進システムである AQUARIUS が搭載される．AQUARIUS は気化した水をノズルから噴射するレジストジェットスラストである．AQUARIUS の推力性能を正しく取得するためには，各試験パラメータと推力の関係を正確に把握する必要がある．しかしながら，数時間にわたる長時間作動に加え，定常ではなく脈動作動を採用する AQUARIUS の作動試験データは，非常に膨大である．そこで，本研究では AQUARIUS の開発段階で取得された膨大な試験データをもとにデータ分析を行った．本講演では，分析結果のうち，試験パラメータである各所圧力，温度，および投入電力が推力に与える影響についてまとめ，報告する

---

# Appendix

---

## Appendix

本章は3つの部分から構成されている。

A.1. 主成分分析の一部の解析結果

A.2. 推力推定の結果

A.3. 気化室圧力推定の結果

### A. 1. 主成分分析の一部の解析結果

主成分分析において、共分散行列の第1, 2, 3主成分に対する固有ベクトルを表にまとめた。Table A.1はDVT1に対する第1, 2, 3主成分に対する固有ベクトル、Table A.2はDVT1, DVT2に対する第1, 2, 3主成分に対する固有ベクトルを表している。

Table A.1 Eigenvector results of principal component analysis for DVT1.

特徴量	PC1	PC2	PC3
Time/s	-0.20391	0.03676	-0.05302
DTS1/°C	-0.01434	0.248471	0.112383
VTS2/°C	0.020152	-0.03477	-0.05956
VTS1/°C	-0.00201	0.051349	-0.05372
WPS/kPa	0.021441	-0.1418	-0.08192
VPS1/kPa	0.21934	0.001439	0.012535
DPS1/kPa	0.219745	0.006678	0.015359
VH/W	0.052801	-0.07867	-0.07984
DVT1_PH/W	0.021116	-0.17084	-0.11704
mass/g	0.019332	-0.14977	-0.07774
DPS1_shift(+1)/kPa	0.214185	0.033044	-0.04508
DPS1_shift(+2)/kPa	0.20485	0.048867	-0.08439
DPS1_shift(+3)/kPa	0.191278	0.051322	-0.09349
DTS1_shift(+1)/°C	-0.01523	0.26679	-0.00359
DTS1_shift(+2)/°C	-0.01902	0.272252	-0.07407
DTS1_shift(+3)/°C	-0.02641	0.257468	-0.08538
VPS1_shift(+1)/kPa	0.212893	0.028254	-0.04901
VPS1_shift(+2)/kPa	0.202694	0.04445	-0.08936
VPS1_shift(+3)/kPa	0.188246	0.046927	-0.09866



VTS2_shift(+1)/°C	0.009789	0.080244	-0.23636
VTS2_shift(+2)/°C	0.002372	0.113346	-0.29672
VTS2_shift(+3)/°C	-0.00855	0.106758	-0.27375
VTS1_shift(+1)/°C	-0.00433	0.117629	-0.22988
VTS1_shift(+2)/°C	-0.01061	0.138465	-0.28629
VTS1_shift(+3)/°C	-0.00855	0.106758	-0.27375
DPS1_shift(-1)/kPa	0.218695	-0.00251	0.036684
DPS1_shift(-2)/kPa	0.216185	-0.00436	0.042153
DPS1_shift(-3)/kPa	0.212843	-0.00545	0.045845
DTS1_shift(-1)/°C	0.013672	0.240095	0.121528
DTS1_shift(-2)/°C	0.034064	0.228798	0.125318
DTS1_shift(-3)/°C	0.049013	0.215105	0.124132
VPS1_shift(-1)/kPa	0.218759	-0.00665	0.033158
VPS1_shift(-2)/kPa	0.216613	-0.00786	0.038378
VPS1_shift(-3)/kPa	0.213551	-0.0086	0.042011
VTS2_shift(-1)/°C	0.057082	0.023848	0.017404
VTS2_shift(-2)/°C	0.075137	0.03731	0.039731
VTS2_shift(-3)/°C	0.086286	0.038766	0.047264
VTS1_shift(-1)/°C	0.050842	0.063201	0.024973
VTS1_shift(-2)/°C	0.071997	0.066511	0.044373
VTS1_shift(-3)/°C	0.086286	0.038766	0.047264
DTS1_rolling/°C	-0.01908	0.272205	-0.07421
VTS2_rolling/°C	-0.0014	0.112637	-0.29699
VTS1_rolling/°C	-0.01181	0.138194	-0.2863
VPS1_rolling/kPa	0.208713	0.043371	-0.08352
VH_rolling/J	0.045249	-0.05278	-0.1308
DVT1_PH_rolling/J	0.016101	-0.12246	-0.19771
DPS1_rolling/kPas	0.210579	0.047598	-0.07848
DVT1_PH_sum/J	0.009348	-0.1593	-0.09539
VH_sum/J	0.012339	-0.10434	-0.05674
VTS1_mean/°C	0.018613	0.071051	-0.07293
VTS2_mean/°C	0.025173	-0.01186	-0.12672
DTS2_mean/°C	-0.01601	0.247717	0.11052
DPS1_mean/°C	0.060701	-0.08431	-0.10448
VPS1_mean/°C	0.060738	-0.10964	-0.10483

P_Delta/kPa	-0.21859	0.001465	-0.01094
mass flow rate/(g/s)	0.048849	-0.15637	-0.12457
viscosity/Pa·s	-0.01435	0.248428	0.112378
Re	0.047395	-0.18775	-0.134
Psat/kPa	0.020625	-0.02982	-0.05732
mass_left/g	0.186298	-0.05065	0.070704
VPS1_integral/kPa·s	-0.18107	0.045238	-0.06834
DPS1_integral/kPa·s	-0.17853	0.05314	-0.06821
VH_integral/J	-0.15011	-0.0377	-0.10143
DVT1_PH_integral/J	-0.13901	-0.08082	-0.11908

Table A.1 Eigenvector results of principal component analysis for DVT2.

特徴量	PC1	PC2	PC3
Time/s	0.185722	-0.05513	-0.04763
DTS2/°C	-0.04497	-0.11924	-0.10623
VTS2/°C	0.046324	0.037813	0.224442
VTS1/°C	0.04911	0.035152	0.234243
WPS/kPa	-0.04537	-0.08389	-0.16543
VPS1/kPa	-0.2211	0.016072	-0.00181
DPS2/kPa	-0.22084	0.020364	0.004211
VH/W	-0.04601	-0.05512	-0.13532
DVT2_PH/W	-0.0103	-0.0621	-0.1285
mass/g	-0.01695	0.007858	-0.01574
DPS2_shift(+1)/kPa	-0.21662	-0.03573	0.014398
DPS2_shift(+2)/kPa	-0.20974	-0.07354	0.014686
DPS2_shift(+3)/kPa	-0.19961	-0.08329	0.005172
DTS2_shift(+1)/°C	-0.0251	-0.24676	-0.0316
DTS2_shift(+2)/°C	-0.01262	-0.30089	-0.01335
DTS2_shift(+3)/°C	-0.00012	-0.28308	-0.02387
VPS1_shift(+1)/kPa	-0.21572	-0.04206	0.008812
VPS1_shift(+2)/kPa	-0.20769	-0.08139	0.009226
VPS1_shift(+3)/kPa	-0.19641	-0.09133	-0.00055
VTS2_shift(+1)/°C	0.027717	-0.18496	0.175198
VTS2_shift(+2)/°C	0.026543	-0.26064	0.149161
VTS2_shift(+3)/°C	0.031509	-0.24638	0.111341

VTS1_shift(+1)/°C	0.0246	-0.20324	0.143542
VTS1_shift(+2)/°C	0.025898	-0.26924	0.119935
VTS1_shift(+3)/°C	0.031509	-0.24638	0.111341
DPS2_shift(-1)/kPa	-0.21972	0.035708	0.001546
DPS2_shift(-2)/kPa	-0.21736	0.039356	0.002567
DPS2_shift(-3)/kPa	-0.21419	0.041919	0.003994
DTS2_shift(-1)/°C	-0.06926	-0.07319	-0.00676
DTS2_shift(-2)/°C	-0.08444	-0.05186	0.022632
DTS2_shift(-3)/°C	-0.09432	-0.039	0.026937
VPS1_shift(-1)/kPa	-0.22062	0.030693	-0.00259
VPS1_shift(-2)/kPa	-0.21874	0.034028	-0.00051
VPS1_shift(-3)/kPa	-0.21596	0.036487	0.00138
VTS2_shift(-1)/°C	-0.01701	0.032602	0.201182
VTS2_shift(-2)/°C	-0.04226	0.033498	0.184392
VTS2_shift(-3)/°C	-0.05688	0.034044	0.161347
VTS1_shift(-1)/°C	-0.02954	0.02541	0.1729
VTS1_shift(-2)/°C	-0.05377	0.027065	0.158134
VTS1_shift(-3)/°C	-0.05688	0.034044	0.161347
DTS2_rolling/°C	-0.01251	-0.30094	-0.01344
VTS2_rolling/°C	0.030626	-0.26069	0.147745
VTS1_rolling/°C	0.027468	-0.26929	0.11898
VPS1_rolling/kPa	-0.21213	-0.07733	0.01199
VH_rolling/J	-0.0364	-0.10176	-0.12217
DVT2_PH_rolling/J	-0.00481	-0.14197	-0.10623
DPS2_rolling/kPas	-0.21389	-0.0693	0.01742
DVT2_PH_sum/J	-0.03138	-0.06527	-0.20722
VH_sum/J	-0.02347	-0.022	-0.21814
VTS1_mean/°C	0.014933	0.005504	0.245298
VTS2_mean/°C	0.03079	-0.00973	0.233769
DTS2_mean/°C	-0.04373	-0.11934	-0.10745
DPS2_mean/kPa	-0.07068	-0.06088	-0.1953
VPS1_mean/kPa	-0.07137	-0.06009	-0.19092
P_Delta/kPa	0.220145	-0.01341	0.005416
mass flow rate/(g/s)	0.007057	0.02201	0.14149
viscosity/Pas	-0.04495	-0.11923	-0.10613

Re	0.010468	0.031322	0.148447
Psat/kPa	0.045904	0.037871	0.227733
mass_left/g	-0.17321	0.066115	0.035758
VPS1_integral/kPas	0.146582	-0.07468	-0.12146
DPS2_integral/kPas	0.140881	-0.07748	-0.12644
VH_integral/J	0.155433	-0.06293	-0.13188
DVT2_PH_integral/J	0.132583	-0.08205	-0.10943

## A. 2. 推力推定の結果

### A. 2. 1. 1 推力推定結果の RMSE と、RMSE の標準偏差

評価 1 に対して、Table A.2.1 と Table A.2.2 はそれぞれのケースの RMSE と、RMSE の標準偏差を示している。ヒストグラム図は Figure A.2.1 から Figure A.2.4 にまとめた。

Table A.2.1 Results of RMSE and standard deviation of RMSE for thrust in each time, trained by randomly dividing test data and training data 400 times when LightGBM is used.

試験名	RMSE
DVT1_SMOTE_FE_LGBM	0.437±0.0397
DVT1_FE_LGBM	0.441±0.0405
DVT1_LGBM	0.545±0.0306
DVT2_SMOTE_FE_LGBM	0.350±0.0223
DVT2_FE_LGBM	0.347±0.0224
DVT2_LGBM	0.460±0.0196

Table A.2.2 Results of RMSE and standard deviation of RMSE for thrust in each time, trained by randomly dividing test data and training data 30 times when Random forest is used,

試験名	RMSE
DVT1_SMOTE_FE_RF	0.481±0.0367
DVT1_FE_RF	0.485±0.0323
DVT1_RF	0.597±0.0330
DVT2_SMOTE_FE_RF	0.424±0.0309
DVT2_FE_RF	0.430±0.0255
DVT2_RF	0.507±0.0248

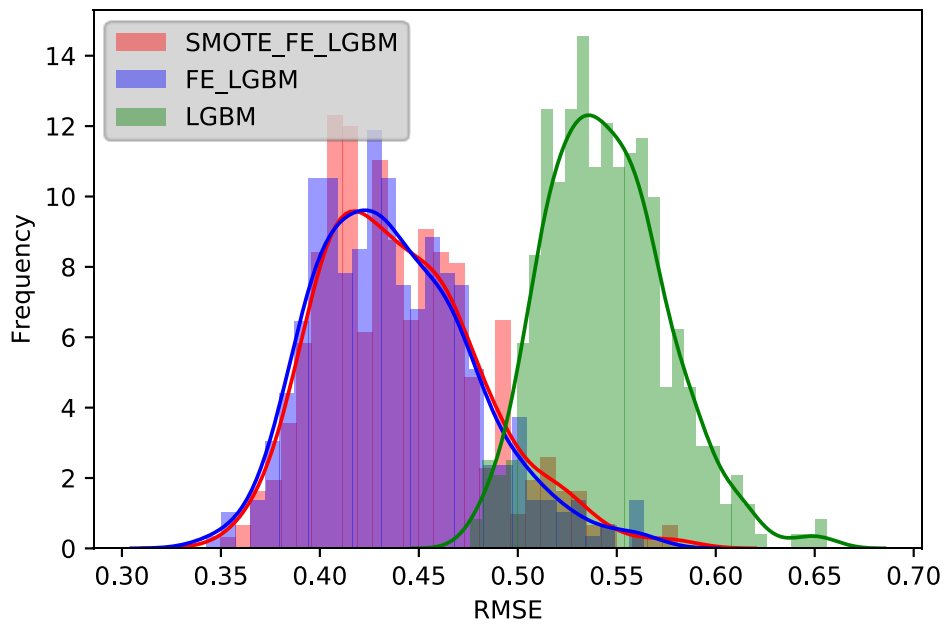


Figure A.2.1 The histogram diagram of the result of evaluation 1 of the LightGBM model for DVT1, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

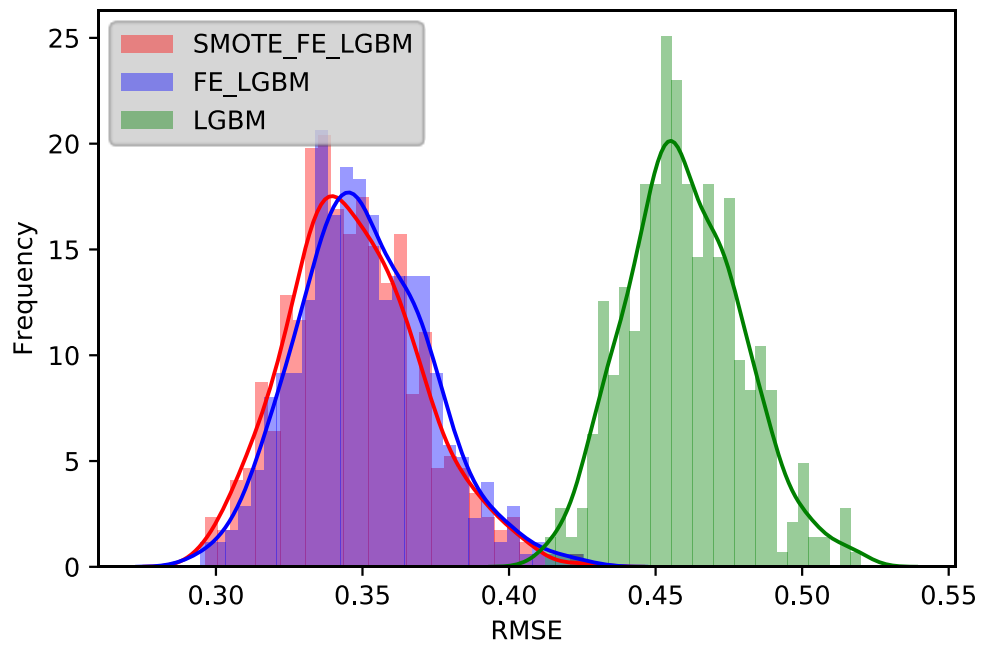


Figure A.2.2 The histogram diagram of the result of evaluation 1 of the LightGBM model for DVT2, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

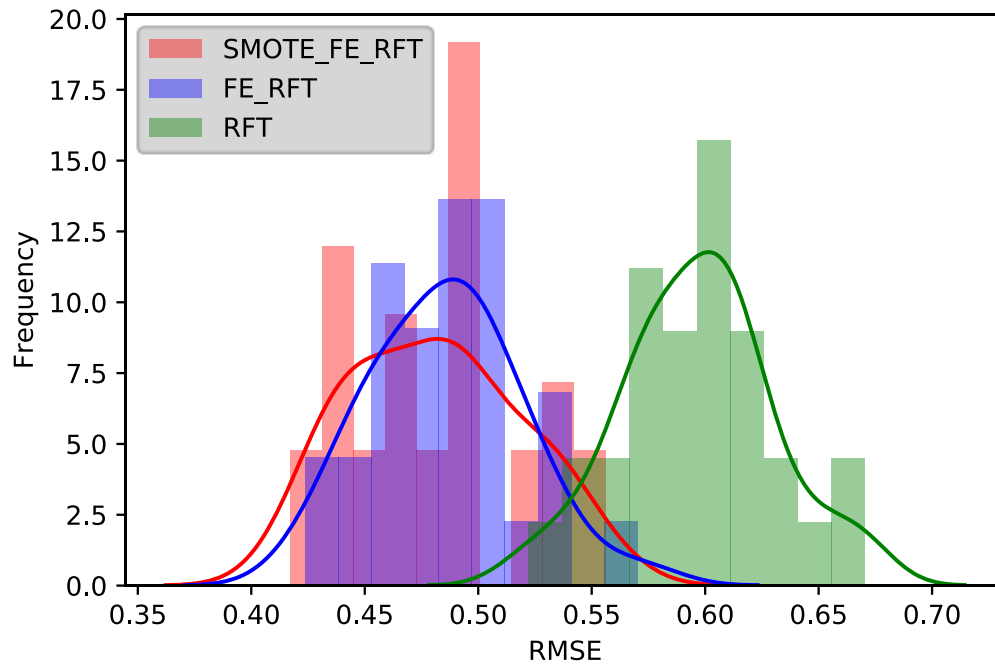


Figure A.2.3 The histogram diagram of the result of evaluation 1 of the Radom forest model for DVT1, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

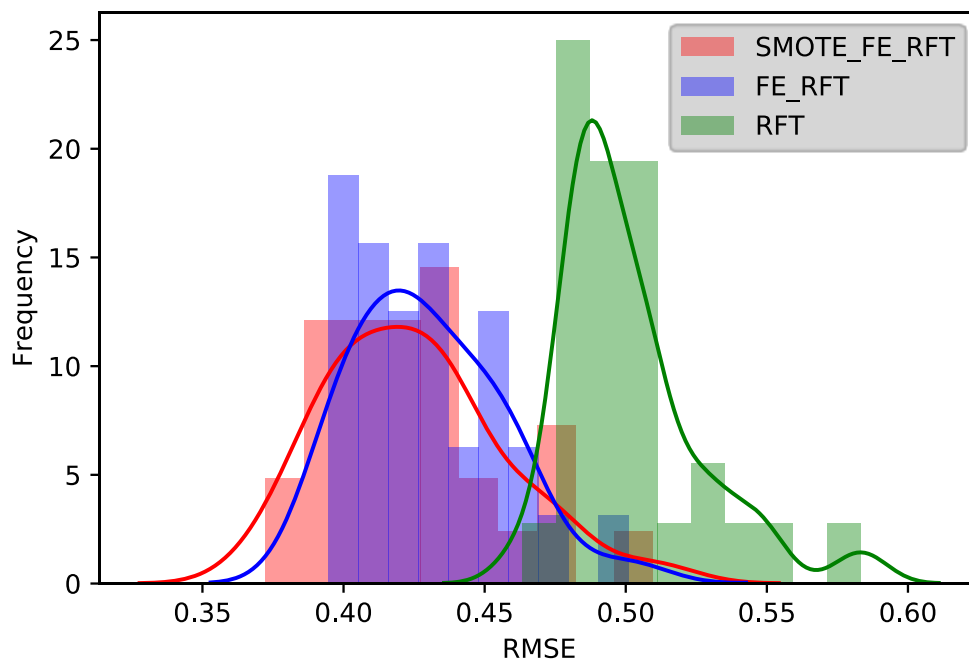


Figure A.2.4 The histogram diagram of the result of evaluation 1 of the Radom forest model for DVT2, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

評価2に対して、Table A.2.3 と Table A.2.4 はそれぞれのケースの RMSE と、RMSE の標準偏差を示している。ヒストグラム図は Figure A.2.5 から Figure A.2.8 にまとめた。

Table A.2.3 Results of RMSE and standard deviation of RMSE for thrust in each injection cycle, trained by randomly dividing test data and training data 400 times when LightGBM is used.

試験名	RMSE
DVT1_SMOTE_FE_LGBM	0.108±0.0328
DVT1_FE_LGBM	0.108±0.036
DVT1_LGBM	0.094±0.017
DVT2_SMOTE_FE_LGBM	0.094±0.0172
DVT2_FE_LGBM	0.094±0.0170
DVT2_LGBM	0.101±0.0167

Table A.2.4 Results of RMSE and standard deviation of RMSE for thrust in each injection cycle, trained by randomly dividing test data and training data 30 times when Random forest is used.

試験名	RMSE
DVT1 の SMOTE_FE_RF	$0.0869 \pm 0.0454$
DVT1 の FE_RF	$0.1110 \pm 0.0385$
DVT1_RF	$0.0981 \pm 0.0256$
DVT1 の SMOTE_FE_RF	$0.0648 \pm 0.0266$
DVT1 の FE_RF	$0.1050 \pm 0.0223$
DVT1_RF	$0.1061 \pm 0.0163$

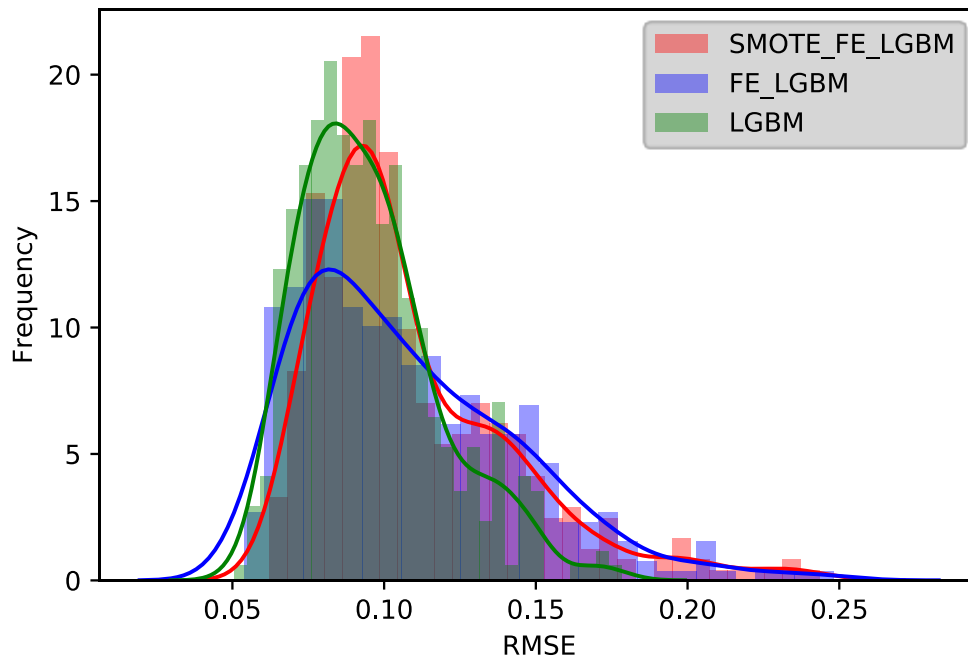


Figure A.2.5 The histogram diagram of the result of evaluation 2 of the LightGBM model for DVT1, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.



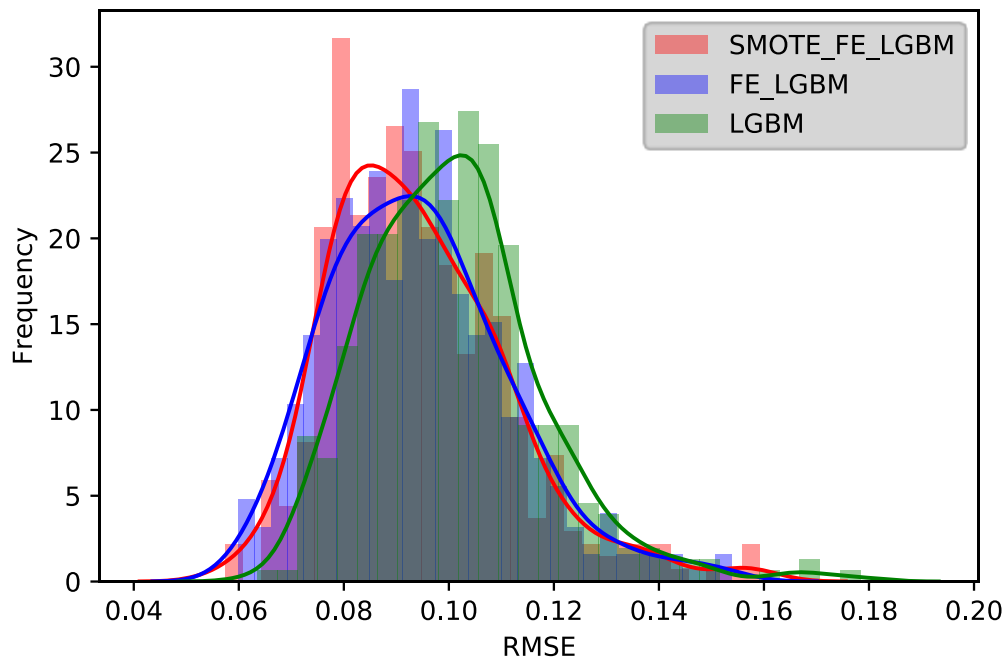


Figure A.2.6 The histogram diagram of the result of evaluation 2 of the LightGBM model for DVT2, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

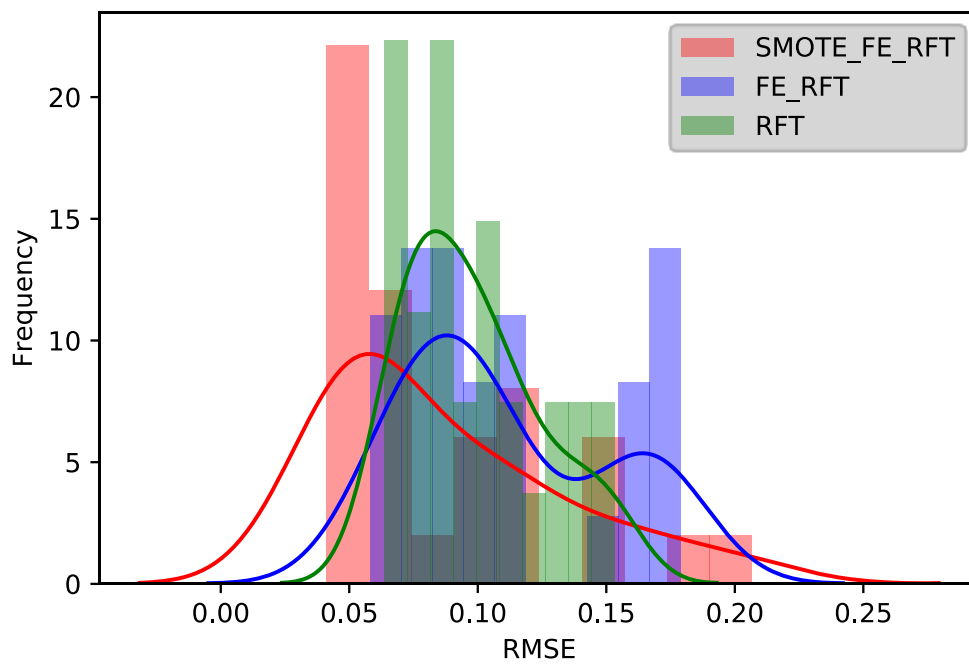


Figure A.2.7 The histogram diagram of the result of evaluation 2 of the Radom forest model for DVT1, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

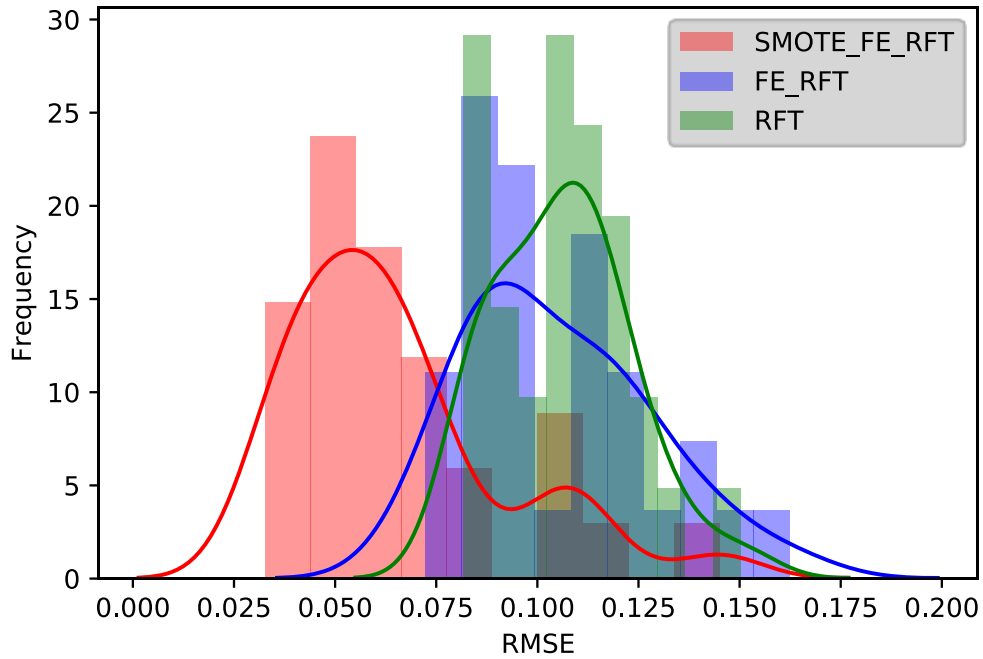


Figure A.2.8 The histogram diagram of the result of evaluation 2 of the Radom forest model for DVT2, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

### A. 2. 1. 2 推力推定重要度

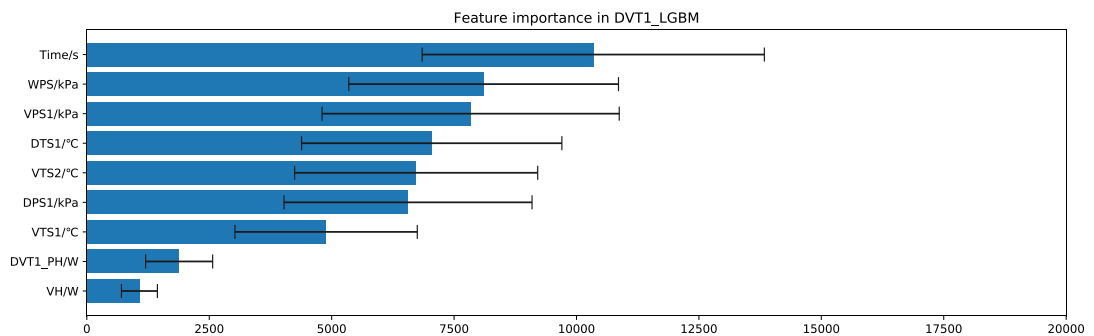


Figure A.2.9 Feature importance in DVT1\_LGBM thrust estimation

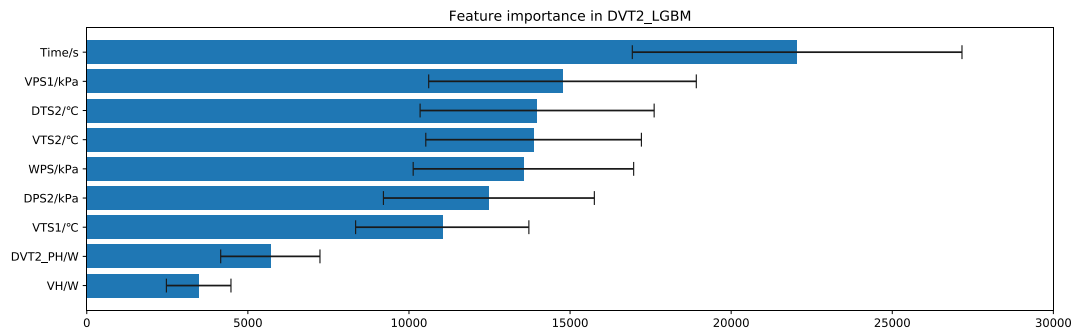


Figure A.2.10 Feature importance in DVT2\_LGBM thrust estimation

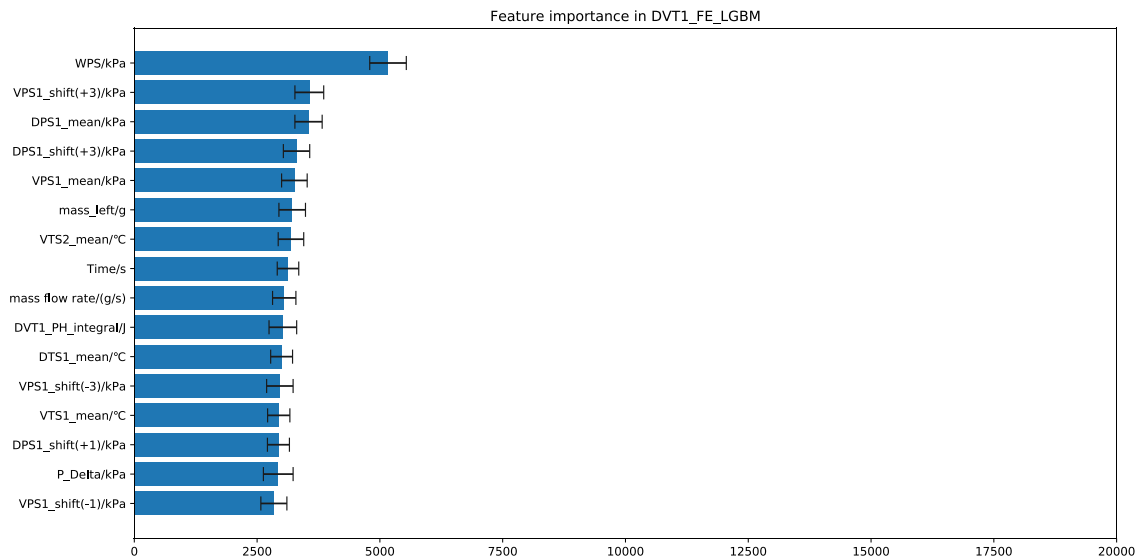


Figure A.2.11 Feature importance in DVT1\_FE\_LGBM thrust estimation

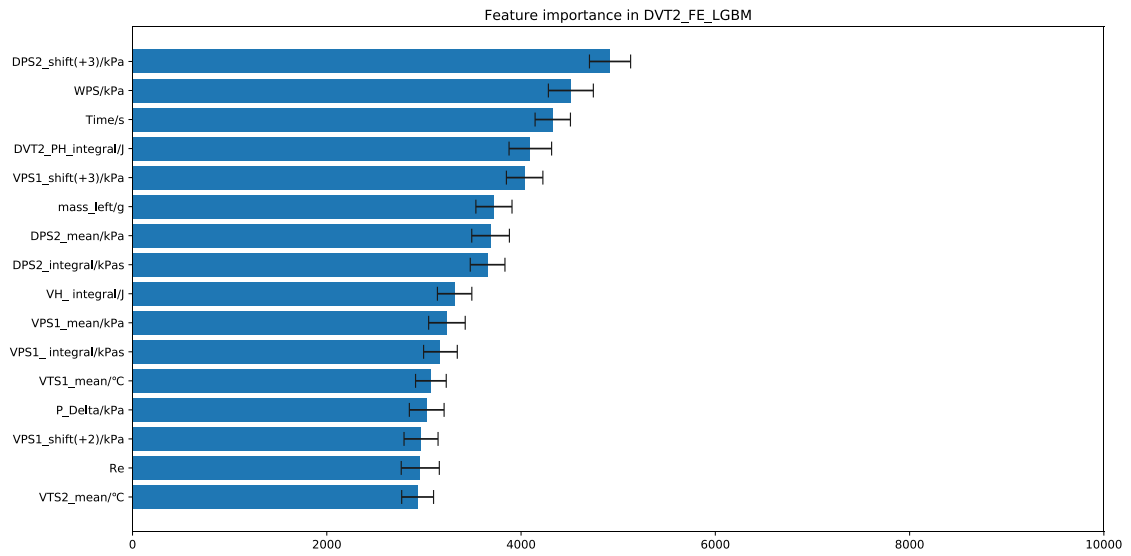


Figure A.2.12 Feature importance in DVT2\_FE\_LGBM thrust estimation

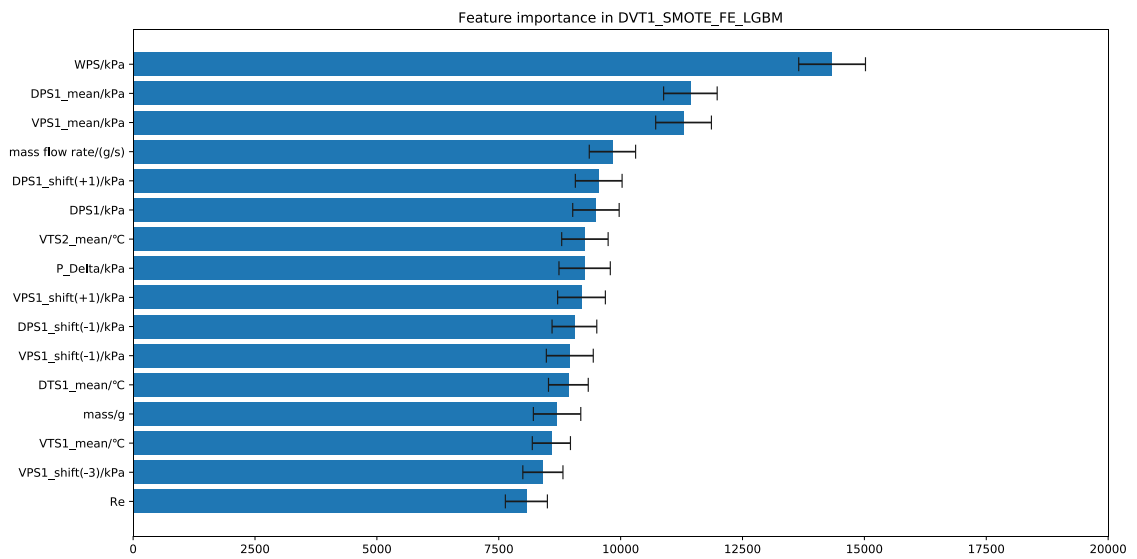


Figure A.2.13 Feature importance in DVT1\_SMOTE\_FE\_LGBM thrust estimation

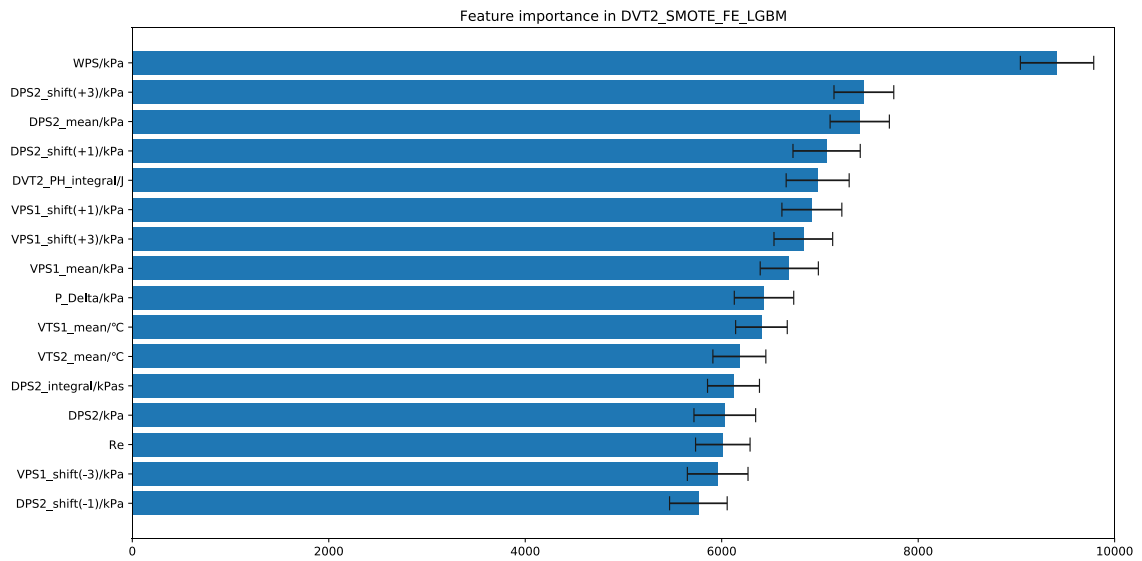


Figure A.2.14 Feature importance in DVT2\_SMOTE\_FE\_LGBM thrust estimation

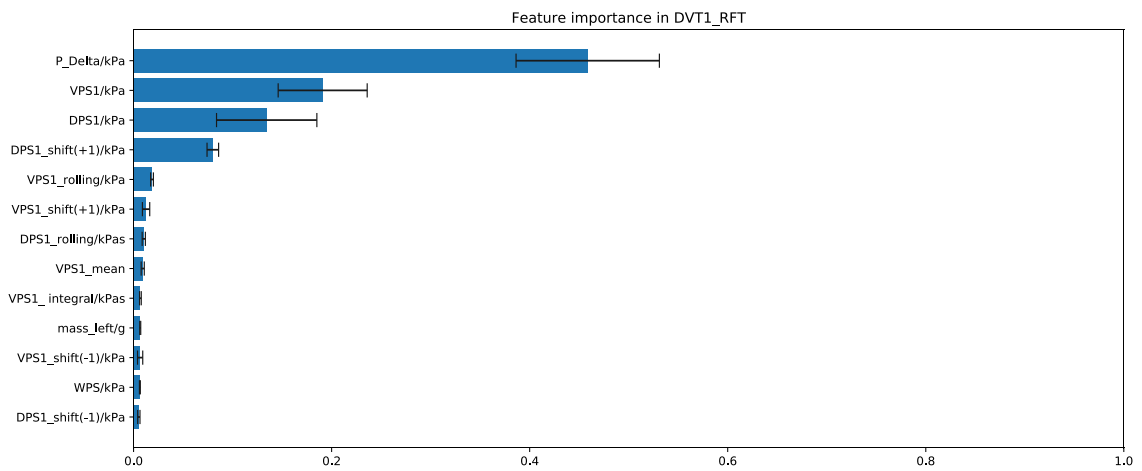


Figure A.2.15 Feature importance in DVT1\_RFT thrust estimation

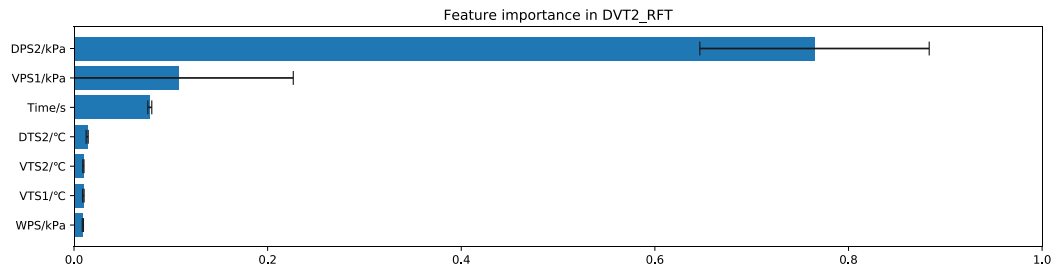


Figure A.2.16 Feature importance in DVT2\_RFT thrust estimation

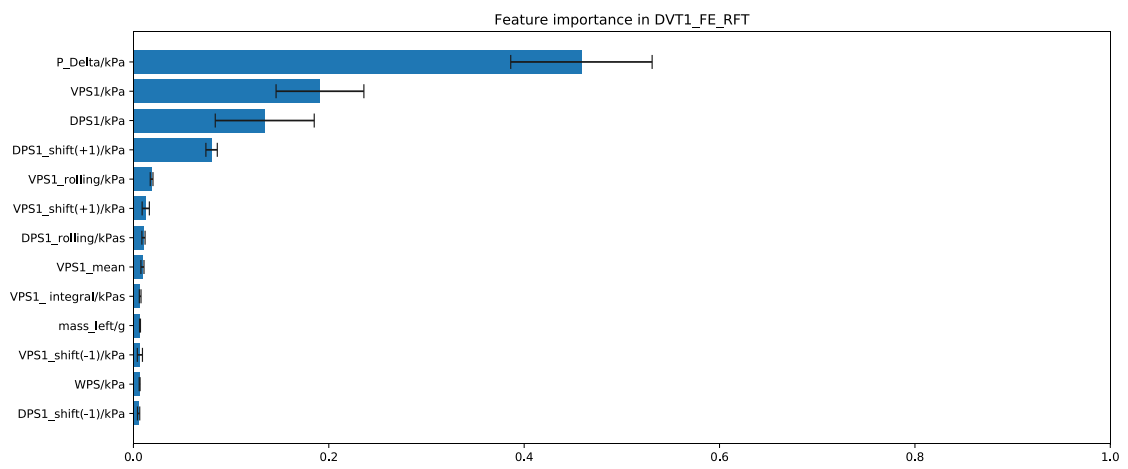


Figure A.2.17 Feature importance in DVT1\_FE\_RFT thrust estimation

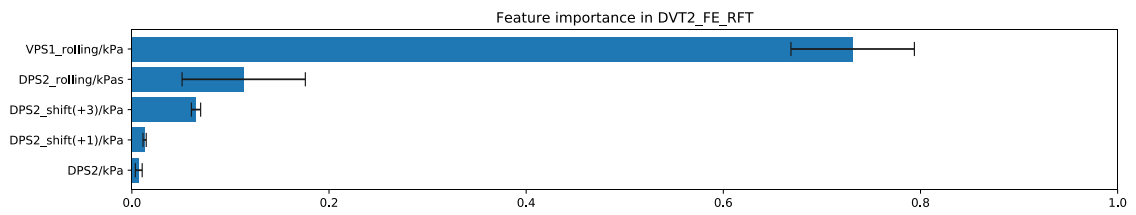


Figure A.2.18 Feature importance in DVT2\_FE\_RFT thrust estimation

Figure A.2.19 Feature importance in DVT1\_SMOTE\_FE\_RFT thrust estimation

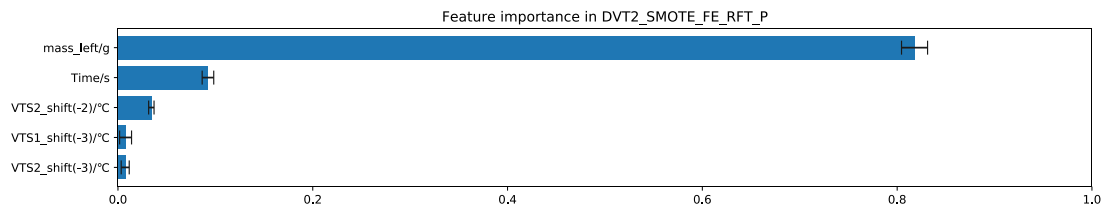
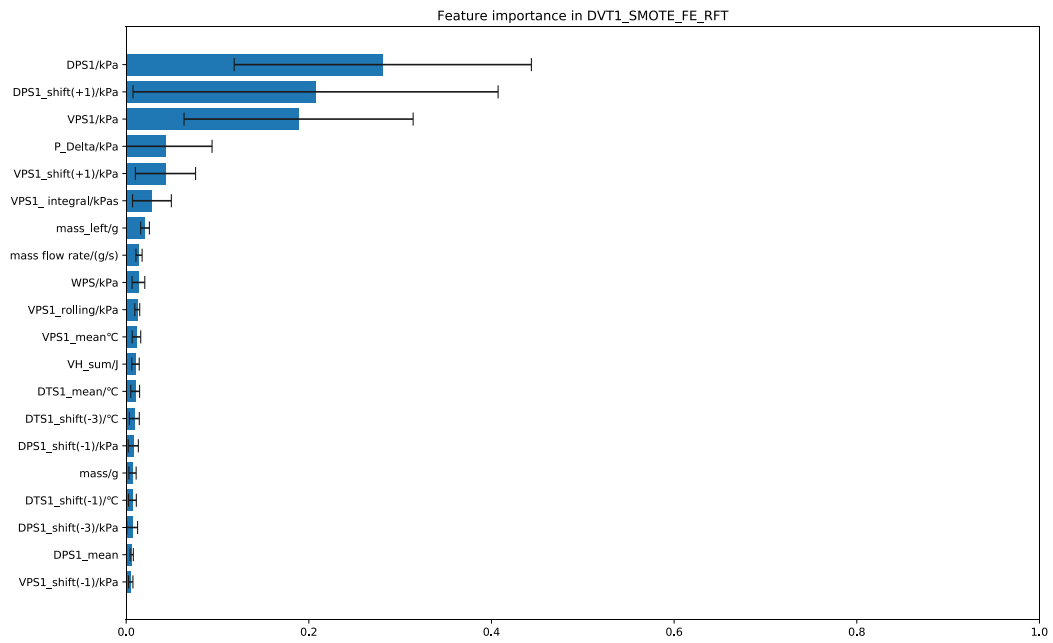


Figure A.2.19 Feature importance in DVT2\_SMOTE\_FE\_RFT thrust estimation

### A. 3.0 圧力推定

まず、蒸発モードか否かの判定においての重要度を Figure A.3.1 と Figure A.3.2 にまとめた。

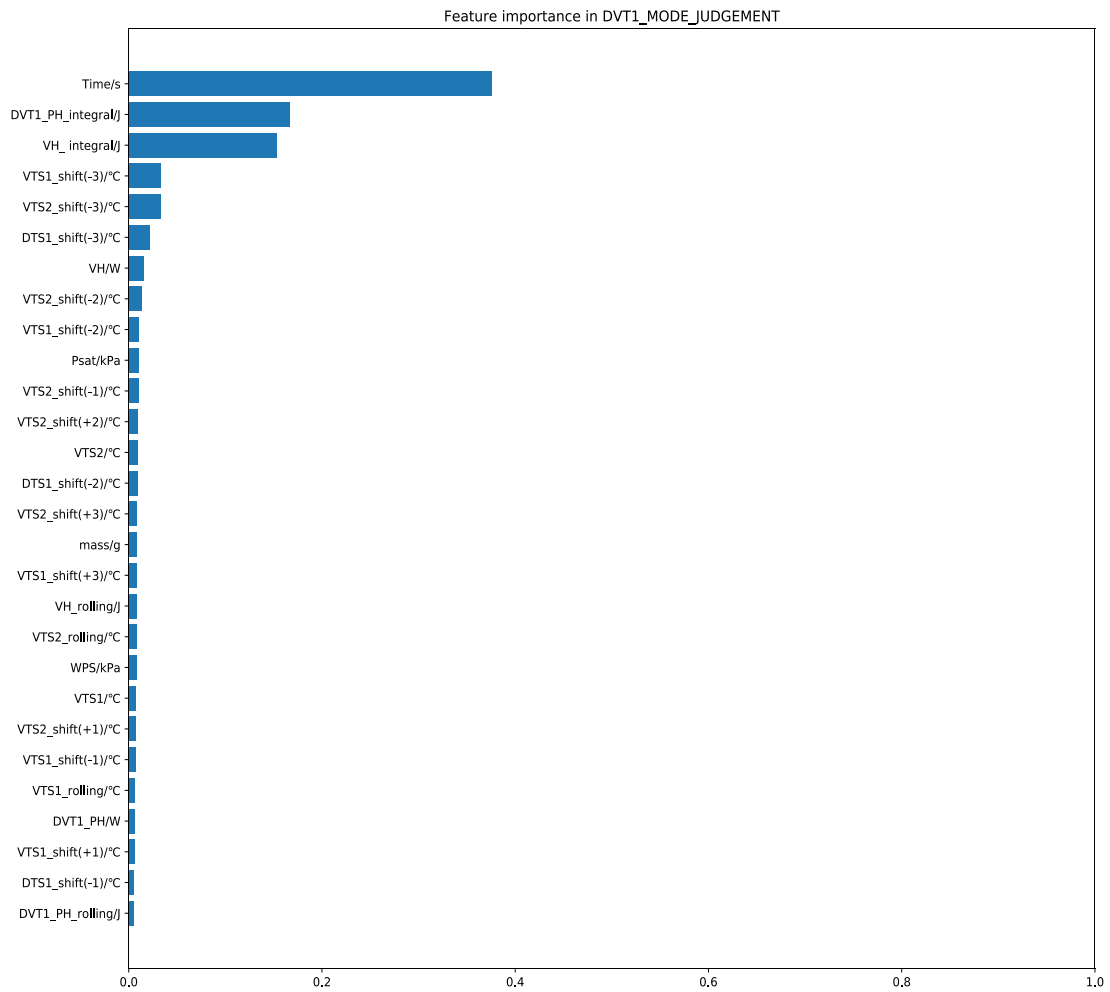


Figure A.3.1 Feature importance in evaporation mode judgement for DVT1.



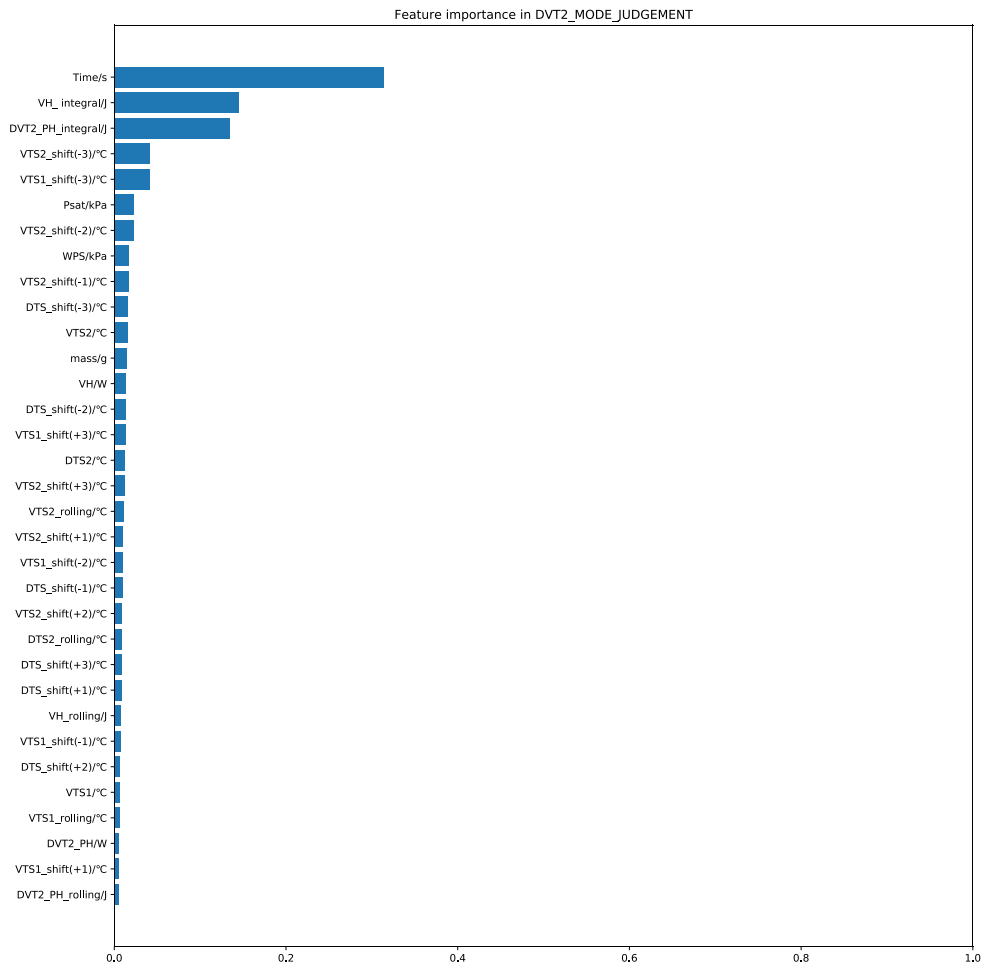


Figure A.3.2 Feature importance in evaporation mode judgement for DVT2.

Table A.3.1 Estimation accuracy when estimating vaporization chamber pressure with LightGBM model.

試験名	RMSE
DVT1_SMOTE_FE_LGBM_P	0.2053 ± 0.0103
DVT1_FE_LGBM_P	0.2052 ± 0.0101
DVT1_LGBM_P	0.2560 ± 0.0109
DVT2_SMOTE_FE_LGBM_P	0.221 ± 0.1196
DVT2_FE_LGBM_P	0.224 ± 0.0117
DVT2_LGBM_P	0.271 ± 0.0168

¥

Table A.3.2 Estimation accuracy when estimating vaporization chamber pressure with Random forest model.

試験名	RMSE
DVT1_SMOTE_FE_RFT_P	$0.2190 \pm 0.0118$
DVT1_FE_RFT_P	$0.2157 \pm 0.0107$
DVT1_RFT_P	$0.2564 \pm 0.0108$
DVT2_SMOTE_FE_RFT_P	$0.227 \pm 0.0117$
DVT2_FE_RFT_P	$0.228 \pm 0.0121$
DVT2_RFT_P	$0.279 \pm 0.0231$

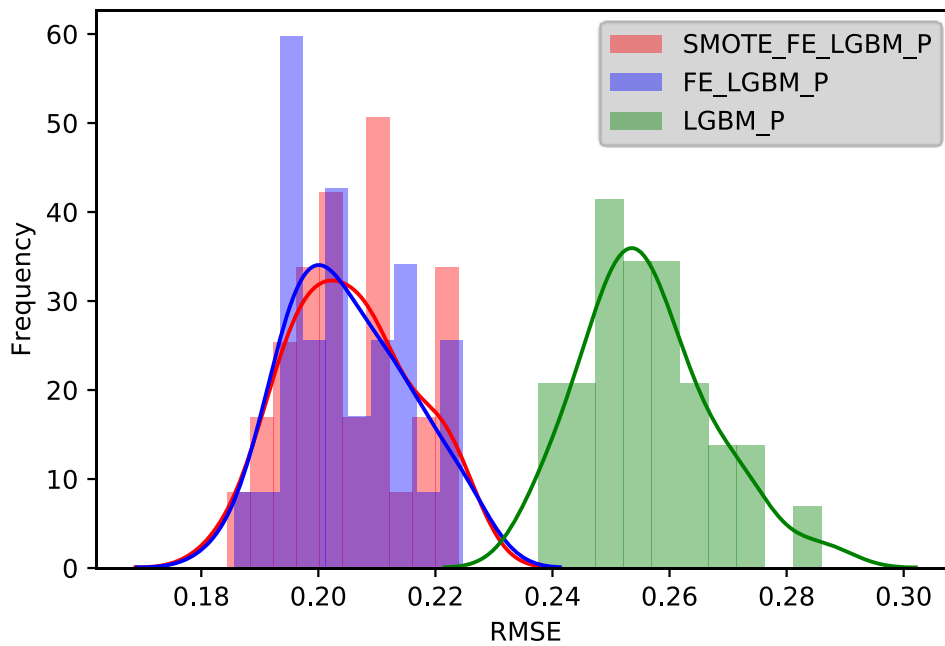


Figure A.3.3 The histogram diagram of the result of the LightGBM model for DVT1 vaporization chamber pressure estimation, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

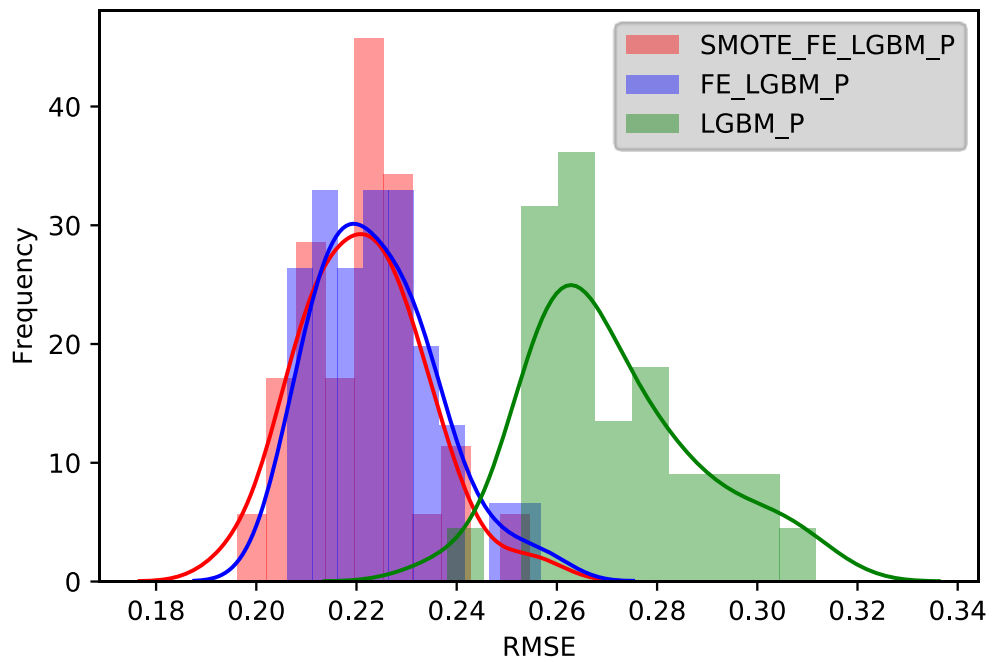


Figure A.3.4 The histogram diagram of the result of evaluation 1 of the LightGBM model for DVT2 vaporization chamber pressure estimation, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

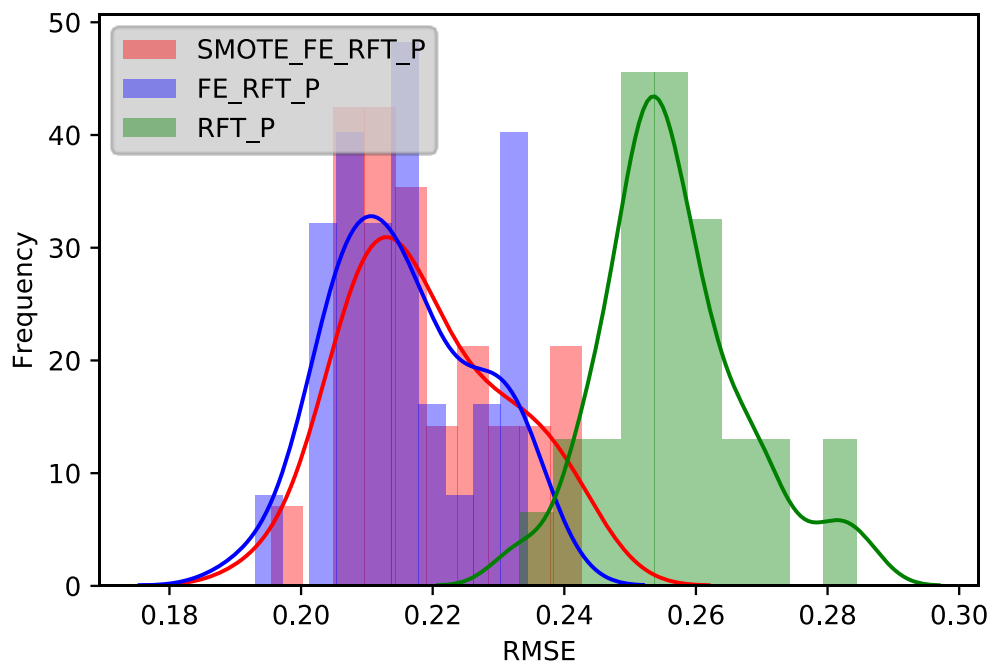


Figure A.3.5 The histogram diagram of the result of evaluation 1 of the Random forest model for DVT2 vaporization chamber pressure estimation, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

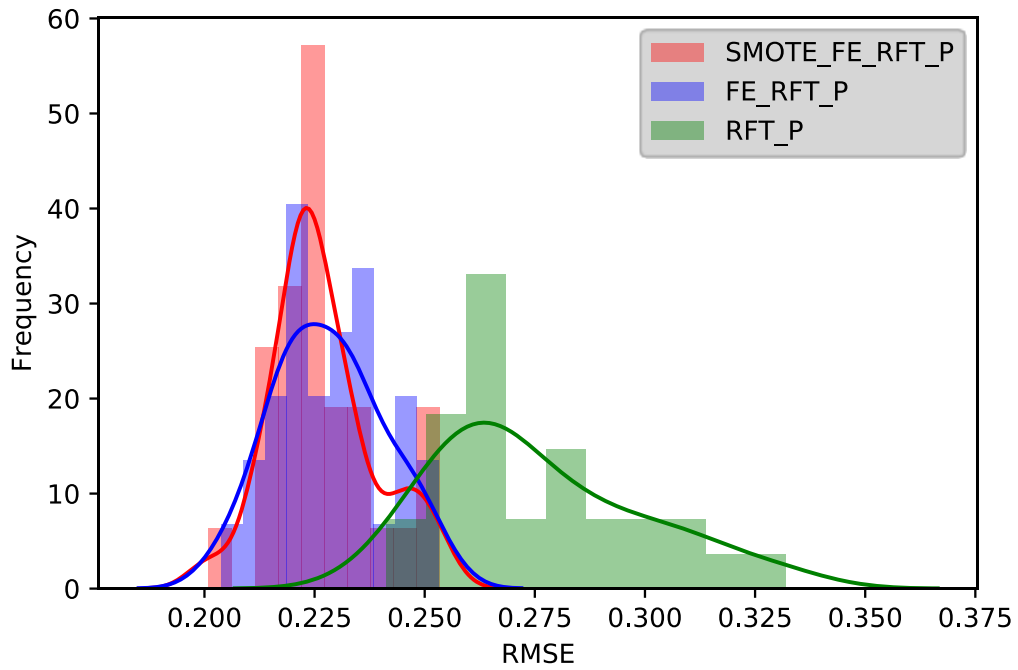


Figure A.3.6 The histogram diagram of the Random forest model for DVT2 vaporization chamber pressure estimation, the horizontal axis RMSE, the vertical axis frequency, and the curve mean the kernel estimation curve.

各モデルで得られた特徴量の重要度を Figure A.3.7 から Figure A.3.20 にまとめた。

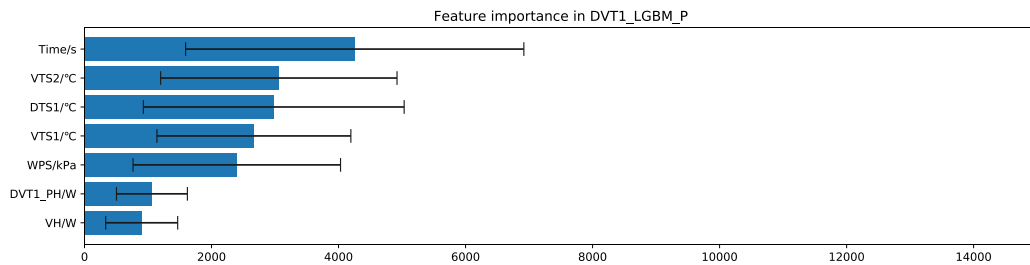


Figure A.3.7 Feature importance in LGBM\_P DVT1 vaporization chamber pressure estimation

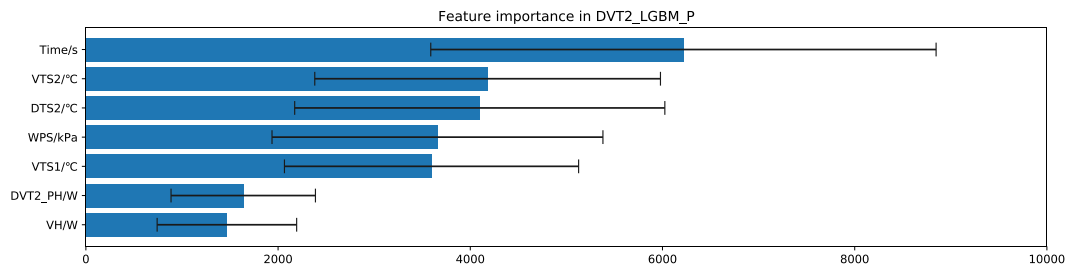


Figure A.3.8 Feature importance in LGBM\_P DVT2 vaporization chamber pressure estimation

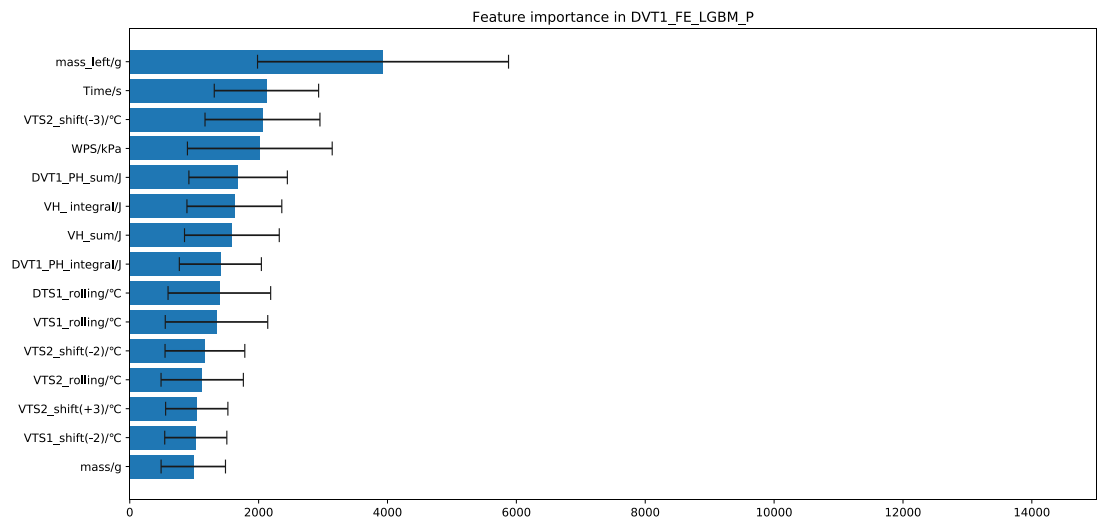


Figure A.3.9 Feature importance in FE\_LGBM\_P DVT1 vaporization chamber pressure estimation.

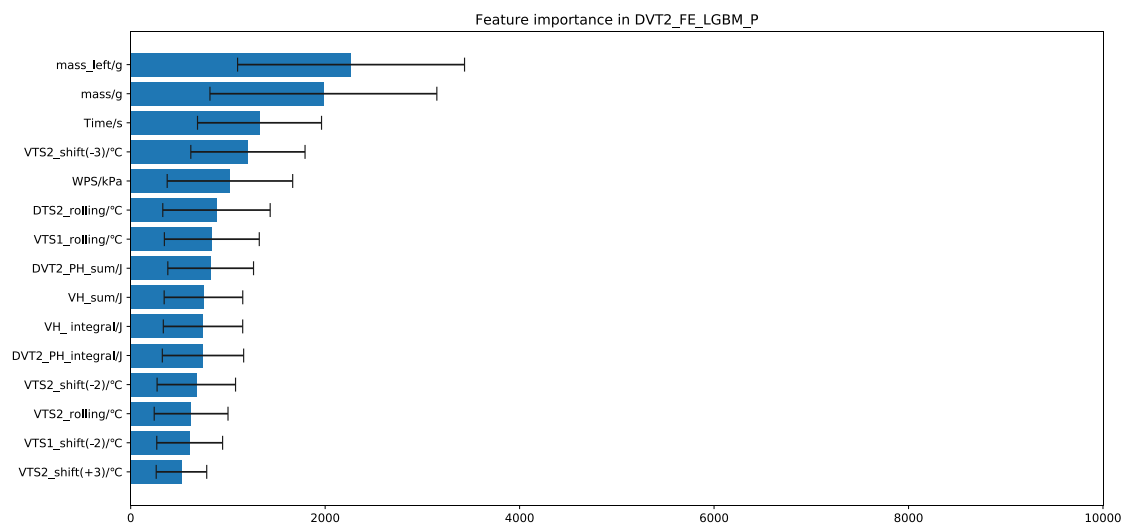


Figure A.3.10 Feature importance in FE\_LGBM\_P DVT2 vaporization chamber pressure estimation.

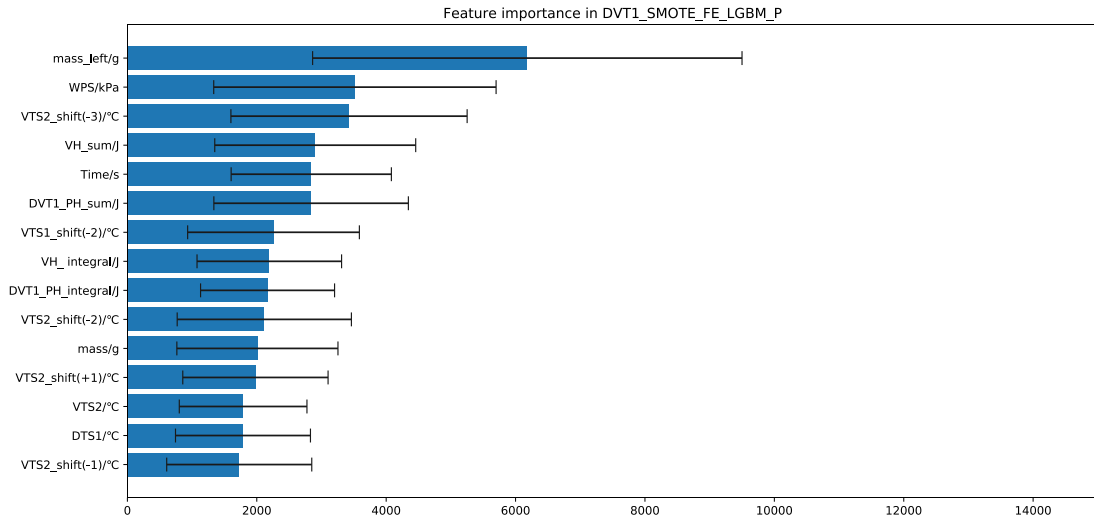


Figure A.3.11 Feature importance in DVT1\_SMOTE\_FE\_LGBM\_P vaporization chamber pressure estimation.

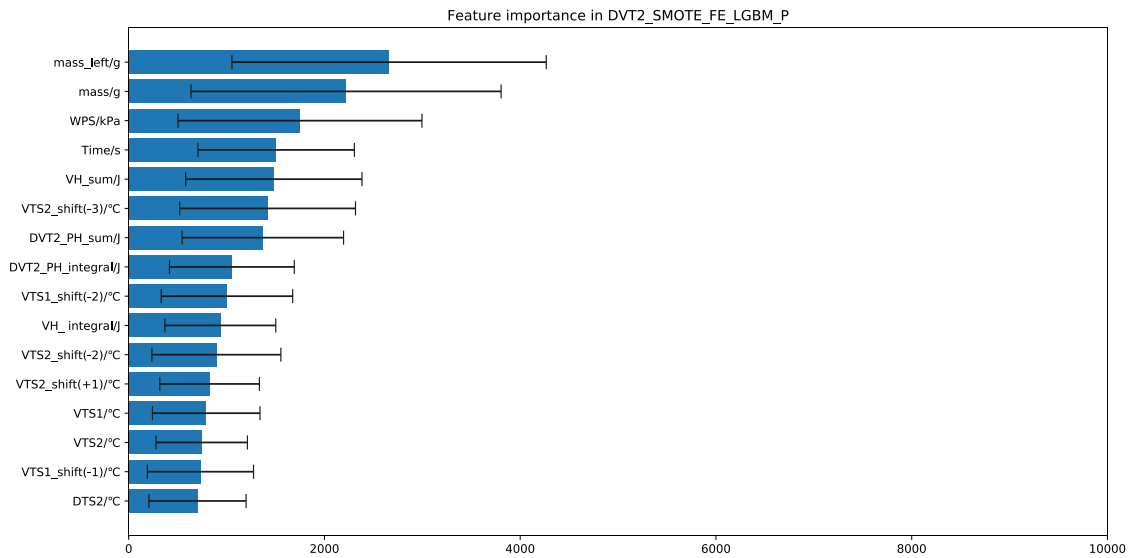


Figure A.3.12 Feature importance in DVT2\_SMOTE\_FE\_LGBM\_P vaporization chamber pressure estimation.

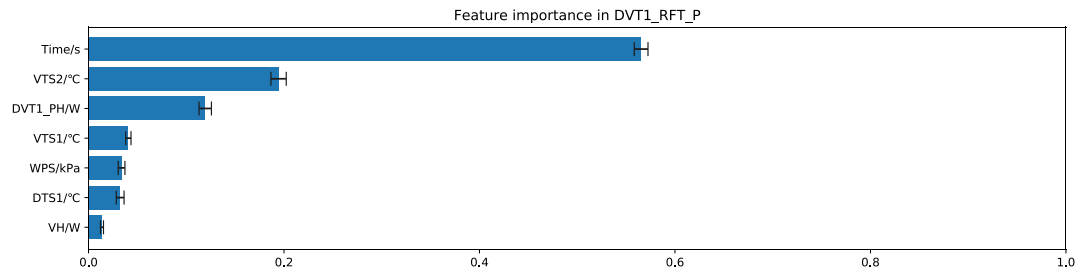


Figure A.3.13 Feature importance in DVT1\_RFT\_P vaporization chamber pressure estimation.

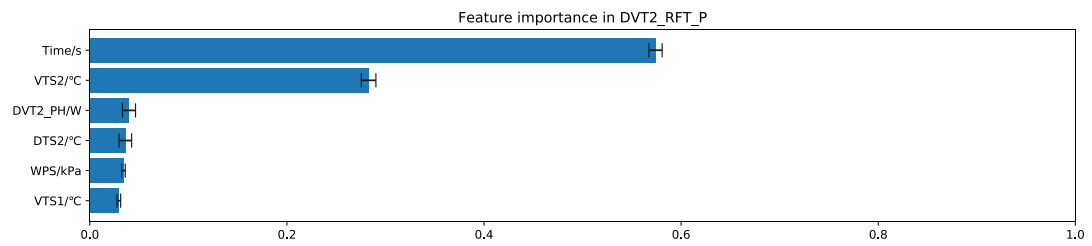


Figure A.3.14 Feature importance in DVT2\_RFT\_P vaporization chamber pressure estimation.

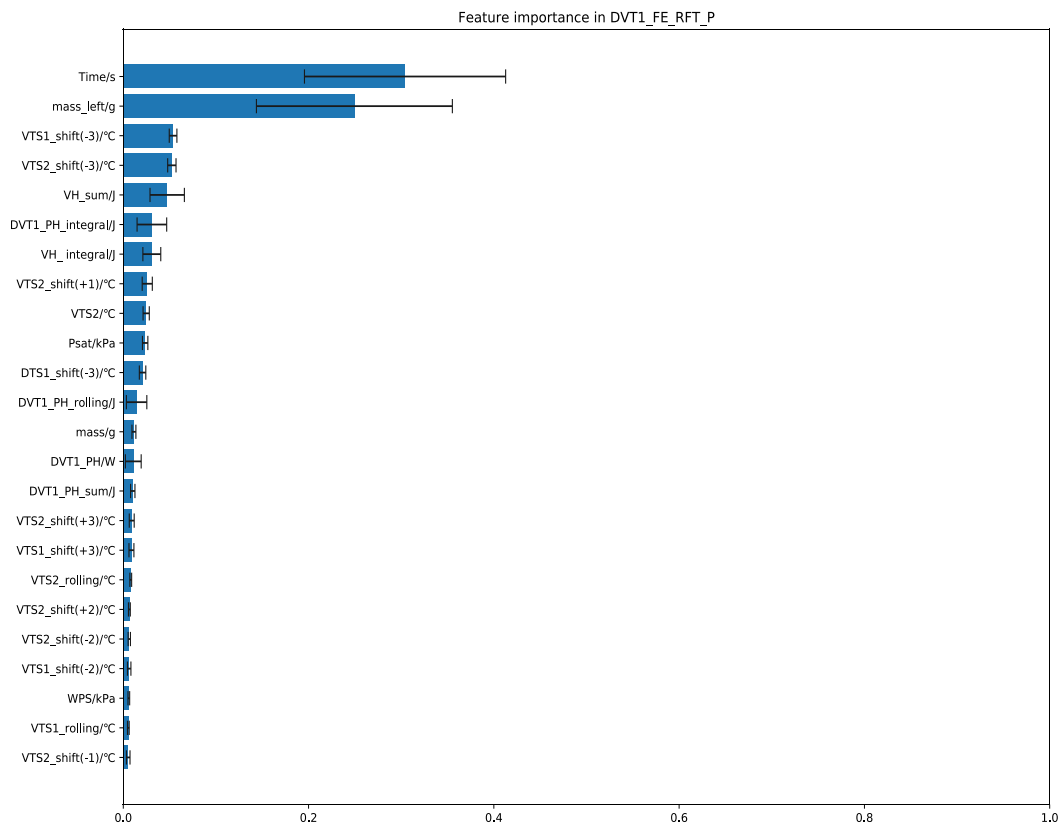


Figure A.3.15 Feature importance in DVT1\_FE\_RFT\_P vaporization chamber pressure estimation.

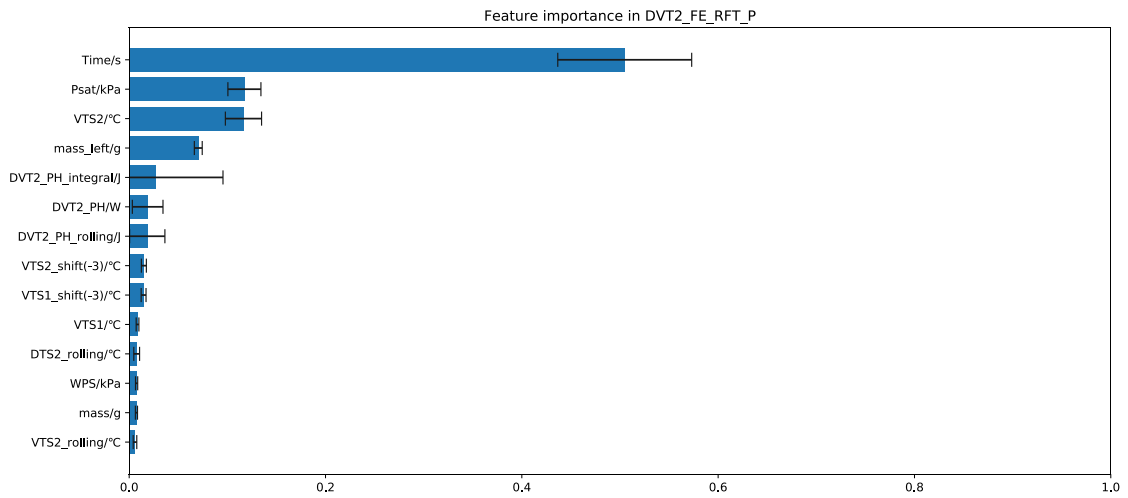


Figure A.3.16 Feature importance in DVT2\_FE\_RFT\_P vaporization chamber pressure estimation.

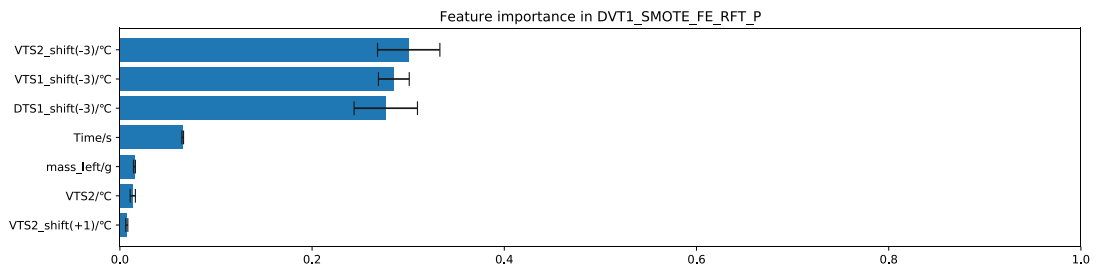


Figure A.3.17 Feature importance in DVT1\_SMOTE\_FE\_RFT\_P vaporization chamber pressure estimation.

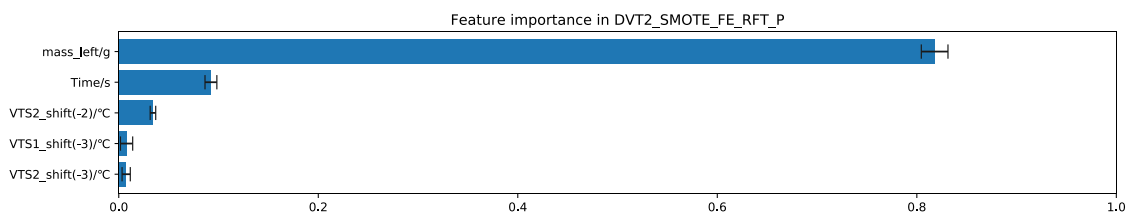


Figure A.3.20 Feature importance in DVT2\_SMOTE\_FE\_RFT\_P vaporization chamber pressure estimation.