

機械学習を用いたトカマクのプラズマパラメータの 予測

岡下, 啓人
九州大学大学院総合理工学府総合理工学専攻プラズマ・量子理工学メジャー

<https://hdl.handle.net/2324/7430256>

出版情報：九州大学, 2025, 修士, 修士
バージョン：
権利関係：



令和7年度
九州大学大学院 総合理工学府
総合理工学専攻
プラズマ・量子理工学メジャー

修士論文

論文名

機械学習を用いたトカマクの
プラズマパラメータの予測

氏名

岡下啓人

指導教員名

長谷川真

目次

目次	1
第 1 章 序論.....	1
1.1. 核融合の可能性	1
1.2. 球状トカマク装置 QUEST	3
1.3. 研究目的	6
第 2 章 トカマクプラズマの力学的平衡	7
2.1. プラズマの力学的平衡	7
2.2. Grad-Shafranov 方程式	7
2.3. 平衡計算	9
第 3 章 ディープラーニング.....	16
3.1. ディープニューラルネットワークの仕組み	16
3.2. ハイパーパラメータ	17
3.3. Optuna を使ったハイパーパラメータチューニング	19
3.4. モデルの学習	20
3.4.1. 学習の流れ	20
3.4.2. 探索条件.....	21
第 4 章 学習モデルの作成	23
4.1. モデルの作成	23
4.1.1. モデルの入出力パラメータ	23
4.1.2. 教師データ	26
4.1.3. データの前処理.....	31
4.2. モデル作成の条件.....	31
4.3. 誤差の評価方法	32
第 5 章 予測結果	34
5.1. 各モデルのハイパーパラメータ	34
5.2. 各モデルの比較	34
5.2.1. 中央値でみる誤差	43

第6章	まとめ	48
6.1.	まとめ	48
6.2.	課題	48
参考文献		50
謝辞		52

第1章

序論

1.1.核融合の可能性

全世界規模での気候変動、エネルギー需要の増大が進行する現代で持続可能かつ、大規模な電源の確保は大きな課題である。核融合発電は軽い原子同士が高温・高圧の下で衝突・融合した際に発生する莫大なエネルギーを利用して発電する仕組みである。これは気象に依存しないことや、発電の安定性、資源の豊富さなどの点から、現在、普及しつつある再生可能エネルギーと比較しても多くの利点を持っている。しかし、技術面などの課題から未だ実用化には至っておらずグローバルに研究が行われている。以下に核融合発電の特徴を示す。

第一に、核融合発電は既存の原子力発電が利用する核分裂反応に比べて、非常に安全性が高い。実現性が最も高いとされる D-T 反応で使用される放射性物質（主に三重水素）の半減期は 12.3 年[1]とされており、これは核分裂反応に用いられるプルトニウムやセシウムと比較して圧倒的に短く、廃棄物管理の課題が軽減される。また、核融合反応を持続的に起こしエネルギーを取り出すためには、発電装置内のプラズマを 1 億 2000 万度、密度を 10^{20}m^{-3} 、閉じ込め時間 1 秒以上という条件[2]を満たす必要がある。これらの条件は核融合発電の実現の難しさの要因の一つであると同時にその安全性を示すものでもある。核分裂反応では、反応時の熱により分裂が進み、冷却システムのシステムが喪失した際にも反応が止まらず、メルトダウンや外部への放射性物質の放出などのリスクが伴う。一方で、核融合反応の場合、事故などで先の条件が崩れた瞬間にプラズマが消失し反応が終了する。即ち、災害などの不測の事態で融合炉に損壊が発生した場合であっても、外部から特別な操作をすることなく反応が止まり、制御不能な状態に陥る可能性が低いという点で核融合発電は安全性が高いといえる。

第1章 序論

第二に、燃料供給の点においても利点がある。核融合の燃料である重水素は海水中に豊富に存在し、三重水素は自然界にはほとんど存在しないものの、中性子とリチウムを反応させることで生成することができる。リチウムは地球上に広く存在し、中性子は核融合反応によって生じるものを利用するため、半永久的に燃料を供給することが可能である。これら2つの資源でエネルギーを取り出すため、核融合発電の実現は化石燃料の枯渇などの燃料供給の課題における一つの解決策となりうる。

また、核融合発電では二酸化炭素の排出をしないため地球温暖化抑制に効果があり、気候変動への対策としても有用な技術であるとされている。

しかし、核融合発電の実用化に多数の技術的な課題が残っている。核融合反応を起こすためには原子核同士を十分な距離に近づける必要があり、そのためには、電子核やクーロン障壁を突破するほどのエネルギーを与えプラズマ化させなければならない。さらに、核融合炉内の真空容器にガスを注入し、プラズマ化した原子を高速で衝突させることで核融合反応を起こすためには、拡散・散逸しようとするプラズマを核融合反応が起こる条件を満たすよう閉じ込めておく必要がある。磁場閉じ込め方式の融合炉は強力な磁場を用いてプラズマを空間的に閉じ込め、長時間核融合反応を維持する方式であり、定常運転・連続発電との親和性が高い。その中でもトカマク型は全世界でも最も研究が盛んにおこなわれている型式であり、フランスで建設が進められている ITER もこの方式を採用している。

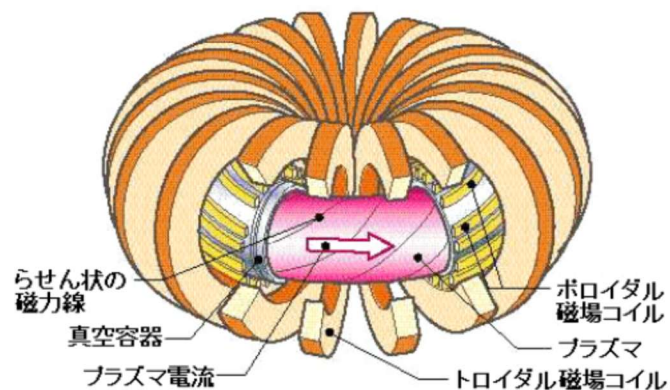


図 1-1 トカマク型核融合装置

[出典] 日本原子力研究所那珂研究室(編): 核融合炉をめざして-核融合研究の進展と拡がり- 平成12年度(2000年11月)p.9

1.2.球状トカマク装置 QUEST

この節では、本研究において学習モデルの対象としている九州大学の球状トカマク装置 QUEST について記述する。QUEST は核融合炉の実現を目指し、国内最大の球状トカマク装置として2008年に建設された。球状トカマクは通常のトカマクに比べて高いプラズマ圧力を弱い磁場で維持する、高 β 運転が可能であるという利点がある。高い β 値では、同じ磁場強度でもより高いプラズマ圧力、もしくは同じプラズマ圧力でもより低い磁場強度でプラズマを維持することができるため炉の小型化・低コスト化に繋がる。

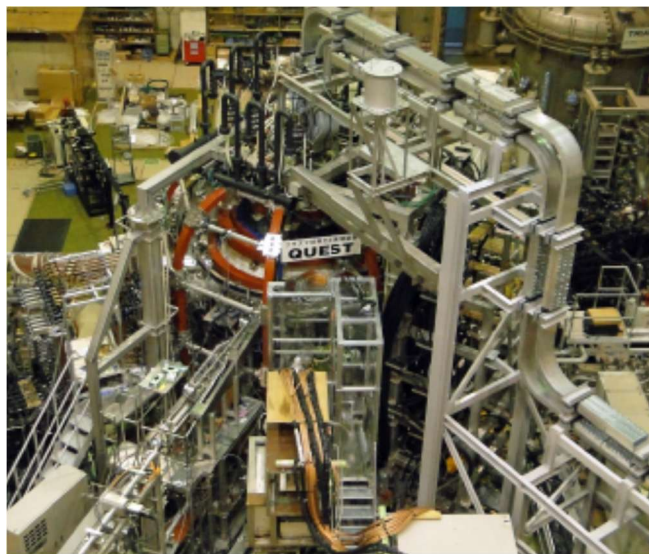


図 1-2 QUEST 実験装置(大半径 0.68m、小半径 0.40m、アスペクト比 1.70、トロイダル磁場 0.25T)

第1章 序論

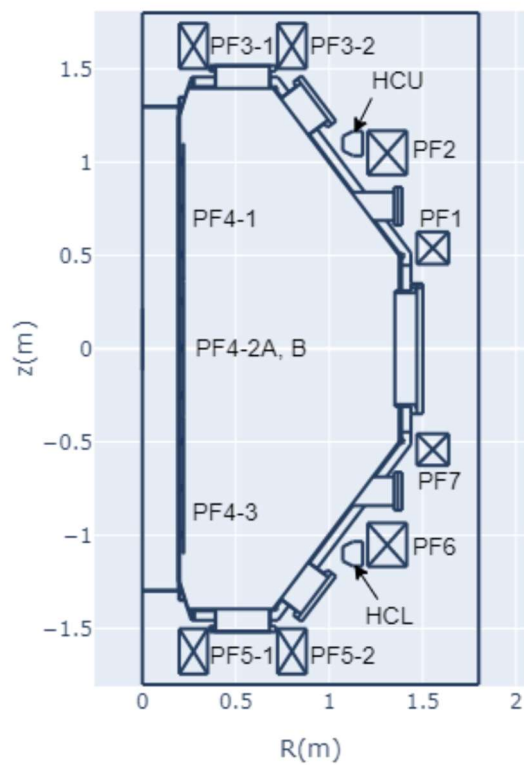


図 1-3 QUEST ポロイダル断面図と使用するコイルの名称

第1章 序論

表 1-1QUEST で使用されるコイルの詳細

名称	巻数	R 位置(mm)	Z 位置	幅	高さ	抵抗(mΩ)	インダクタンス(mH)
PF1	12	1555.2	+540	164.4	165.6	1.02	0.886
	8	1527.8	+540			0.67	0.407
	4	1500.4	+540			0.32	0.106
PF2	72	1311.6	+1100	203.2	228.6	27.47	23.2
	36	1260.8	+1100			13.21	6.05
	18	1235.4	+1100			6.47	1.56
PF3-1	41	273.0	+1615	117.4	194.7	3.34	0.88
PF3-2	41	800.0	+1615	117.4	191.5	9.7	4.3
PF4-1	72	157.4	+740	22.8	590.4	11.8	0.662
PF4-2A,B	72	163.2	0	34.4	791.9	6.3	0.542
PF4-2AB	144	163.2	0	34.4	791.9	12.6	2.17
HCU	16	1125.8	+1150	203.2	228.6	5.27	1.199
	6	1100.4	+1150			1.92	0.187
CC	1	1850	±492				
	2	1850	±508				
	4	1850	±540				
トロイダル 磁場コイル	16						

名称	Z 位置(mm)	上下位置対称の 的コイル
PF4-3	-740	PF4-1
PF5-1	-1615	PF3-1
PF5-2	-1615	PF3-2
PF6	-1100	PF2
PF7	-540	PF1

1.3.研究目的

磁場閉じ込めによる配位の定常的な維持、及びプラズマの制御にはプラズマの位置形状・圧力や β 値、安全係数などのプラズマパラメータのリアルタイム把握が必要になる。現在、QUESTでは磁気データ、画像データなどの計測データを用いて磁気面の再構成を行っているが、その計算は誤差を小さくするために繰り返しの計算を必要とし時間を要するため、実時間でパラメータの把握をすることは難しく制御に用いることはできない。そのため機械学習により、計測データから位置形状やその他のプラズマパラメータをリアルタイムで予測する学習モデルを作成することで上述の計算過程を省略し高速でプラズマパラメータの同定を試みる。また学習モデルの作成にあたり、教師データも併せて作成する。教師データには平衡計算コードで作成した平衡プラズマのデータを用いる。

本研究では、QUESTでの制御に利用することを目的として、数パターンで教師データを作成し、それぞれに学習モデルを構築、比較検討を行うことで最適なモデルを模索していく。

第2章

トカマクプラズマの力学的平衡

2.1. プラズマの力学的平衡

トカマクにおける力学的平衡状態は、高温で高い圧力を持つプラズマとプラズマに作用する磁場からの力が釣り合っている状態を指し、トカマク内でプラズマが安定して存在するための条件となる。この二つの力が釣り合っていないければプラズマの維持は困難であり、長時間運転は不可能である。

まず、運動方程式において平衡状態を考える。ローレンツ力とプラズマの圧力勾配が等しいので

$$\nabla p = \mathbf{j} \times \mathbf{B} \quad (2.1)$$

∇p は圧力勾配、 \mathbf{j} はポロイダル電流密度、 \mathbf{B} は磁束密度とする。

式(2.1)より、

$$\mathbf{B} \cdot \nabla p = 0$$

$$\mathbf{j} \cdot \nabla p = 0$$

が成り立つ。これらは、プラズマの等圧面がそれぞれ磁気面、電流密度に沿うことを示している。[3]

2.2. Grad-Shafranov 方程式

この節では、教師データの作成に使用した平衡計算コードで用いられている Grad-Shafranov 方程式について記述する。

用いる座標系は円筒座標系 (r, θ, z) とし、直交座標系と以下のような関係を持つ。

$$x = r \cos \theta$$

$$y = r \sin \theta$$

$$z = z$$

第2章 トカマクプラズマの力学的平衡

この時、磁束密度は、

$$\begin{aligned} B_r &= -\frac{1}{2\pi r} \frac{\partial \phi}{\partial z} \\ B_\theta &= \frac{\mu_0 I}{2\pi r} \\ B_z &= \frac{1}{2\pi r} \frac{\partial \phi}{\partial r} \end{aligned} \quad (2.2)$$

また電流密度は、

$$\mu_0 \mathbf{j} = \nabla \times \mathbf{B} = -\frac{\partial B_\theta}{\partial z} \mathbf{e}_r + \left(\frac{\partial B_r}{\partial z} - \frac{\partial B_z}{\partial r} \right) \mathbf{e}_\theta + \frac{1}{r} \frac{\partial (r B_\theta)}{\partial r} \mathbf{e}_z$$

より、

$$\begin{aligned} j_r &= -\frac{1}{\mu_0} \frac{\partial B_\theta}{\partial z} \\ j_\theta &= \frac{1}{\mu_0} \left(\frac{\partial B_r}{\partial z} - \frac{\partial B_z}{\partial r} \right) \\ j_z &= \frac{1}{\mu_0} \frac{1}{r} \frac{\partial (r B_\theta)}{\partial r} \end{aligned} \quad (2.3)$$

ϕ : ポロイダルフラックス

I : ポロイダル電流

μ_0 : 真空の透磁率

と表わされる。ここで、 z 軸を中心に半径 r を貫くフラックスをポロイダルフラックス、電流をポロイダル電流とする。

式(2.1)の r 座標成分は

$$j_\theta B_z - j_z B_\theta = \frac{\partial p}{\partial r} \quad (2.4)$$

ここで、

$$\begin{aligned} j_z B_\theta &= \frac{1}{\mu_0} \frac{1}{r} \frac{\partial (r B_\theta)}{\partial r} B_\theta \\ &= \frac{1}{\mu_0} \frac{1}{r^2} \frac{\partial (r B_\theta)}{\partial r} (r B_\theta) \\ &= \frac{1}{2\mu_0} \frac{1}{r^2} \frac{\partial (r B_\theta)^2}{\partial r} \end{aligned}$$

第2章 トカマクプラズマの力学的平衡

$$= \frac{\mu_0}{8\pi^2} \frac{1}{r^2} \frac{\partial I^2}{\partial r} \quad (2.5)$$

これと式(2.2)の z 座標成分を式(2.4)に代入して

$$j_\theta \frac{1}{2\pi r} \frac{\partial \phi}{\partial r} - \frac{\mu_0}{8\pi^2} \frac{1}{r^2} \frac{\partial I^2}{\partial r} = \frac{\partial p}{\partial r} \quad (2.6)$$

ここで

$$\frac{\partial p}{\partial r} = \frac{dp}{d\phi} \frac{\partial \phi}{\partial r}$$

$$\frac{\partial I^2}{\partial r} = \frac{dI^2}{d\phi} \frac{\partial \phi}{\partial r}$$

であるので、これらを式(2.6)に代入して整理すると、

$$j_\theta = 2\pi r \frac{dp}{d\phi} + \frac{\mu_0}{4\pi r} \frac{dI^2}{d\phi} \quad (2.7)$$

となり、これは z 座標成分からも導出することができる。また、 θ 座標成分は両辺ともに 0 となる。[4]

式(2.7)を Grad-Shafranov 方程式といい、軸対称なトカマクプラズマにおける磁気流体平衡を記述する式である。

本研究ではこの Grad-Shafranov の式を用いた平衡計算コードを利用して、学習モデルの教師データを作成する。

2.3. 平衡計算

教師データの作成には、前節で示した Grad-Shafranov の式を用いた平衡計算コードを使用している。

また、この節では Grad-Shafranov の式をトカマクに適用することを前提とするため以下の通り、一部の文字を変更する。

方位角成分	θ	\rightarrow	φ
中心からの距離	r	\rightarrow	R
プラズマ圧力	p	\rightarrow	P

第2章 トカマクプラズマの力学的平衡

ポロイダルフラックス	ϕ	\rightarrow	ψ
------------	--------	---------------	--------

よって Grad-Shafranov の式は以下のように記述する。

$$j_\phi = 2\pi R \frac{dP}{d\psi} + \frac{\mu_0}{4\pi R} \frac{dI^2}{d\psi} \quad (2.7)'$$

以下に平衡計算の手順を示す。

1. プラズマ電流密度 $j_\phi(R, z)$ をパラボリック分布として仮定。
2. コイル電流とプラズマ電流による磁束の総和 $\psi(R, z)$ を計算する。
3. 最外殻磁気面を算出し計算領域を決定する。計算領域の一例を図 2-1 に示す。
4. 計算領域をメッシュで分割し、各メッシュにおけるプラズマ電流密度 j_ϕ を $j_{\phi 0}(i, j)$ とおく。((i, j) はメッシュ番号。)
5. $\frac{dP}{d\psi}$, $\frac{dI^2}{d\psi}$ を、 $x = 0$ で定数、 $x = 1$ で 0 となるようにそれぞれ以下のように置く。

$$\frac{dP}{d\psi}(x) = \sum_{n=0}^{n_p} a_n x^n - x^{n_p+1} \sum_{n=0}^{n_p} a_n \quad (2.8)$$

$$\frac{dI^2}{d\psi}(x) = \sum_{n=0}^{n_p} b_n x^n - x^{n_p+1} \sum_{n=0}^{n_p} b_n \quad (2.9)$$

x は標準化したフラックスとして以下のように定義する。

$$x = \frac{(\psi - \psi_M)}{(\psi_B - \psi_M)} \quad (2.10)$$

$$\therefore x = \begin{cases} 0, & \psi = \psi_M \\ 1, & \psi = \psi_B \end{cases}$$

ここで ψ_M は磁気軸位置での磁束、 ψ_B は最外殻磁気面での磁束を表す。

第2章 トカマクプラズマの力学的平衡

6. 式(2.8)、式(2.9)および式(2.7)'により、 $j_{\varphi 1}(i, j)$ を a_n 、 b_n の線形結合として表現し、 $j_{\varphi 0}(i, j)$ との誤差が最小となる係数 a_n 、 b_n を最小二乗法によって探索する。

$$error = \frac{1}{2} \sum_{(i,j)} \left(j_{\varphi 1}(i, j) - j_{\varphi 0}(i, j) \right)^2 \quad (2.11)$$

7. 誤差が十分に小さければ計算を終了し、そうでなければ最小二乗法によって得られた係数で算出される j_{φ} を使用して手順2より再計算する。

図 2-2 に計算手順のフローチャートを示す。

以下に、手順6で示した最小二乗法による係数の算出の方法を示す。

j_{0i} をはじめに仮定した電流密度、 j_{1i} を式(2.8)、(2.9)、(2.7)'を用いて算出した電流密度とする。 a_j 、 b_j をまとめて c_j を係数として以下のようにまとめられる。

$$j_{1i} = \sum_j f_{ij} c_j \quad (2.12)$$

ここで、添え字 i は計算領域におけるメッシュ番号、添え字 j は式(2.8)、(2.9)における次数に相当する。また、式(2.11)は

$$\begin{aligned} E &= \frac{1}{2} \sum_i (j_{1i} - j_{0i})^2 \\ &= \frac{1}{2} \sum_i \left(\sum_j f_{ij} c_j - j_{0i} \right)^2 \end{aligned} \quad (2.13)$$

と表わされる。この E が最小となるときの c_j を求めることが目的となる。

E を c_k で微分すると

$$\begin{aligned} \frac{\partial E}{\partial c_k} &= \sum_i \left(\sum_j f_{ij} c_j - j_{0i} \right) f_{ik} \\ &= \sum_{i,j} f_{ik} f_{ij} c_j - \sum_i f_{ik} j_{0i} \\ &= 0 \end{aligned}$$

となる。また、行列式で示すと、

第2章 トカマクプラズマの力学的平衡

$$F^T F c = F^T j_0$$

この c が求めたい係数となる。[5]

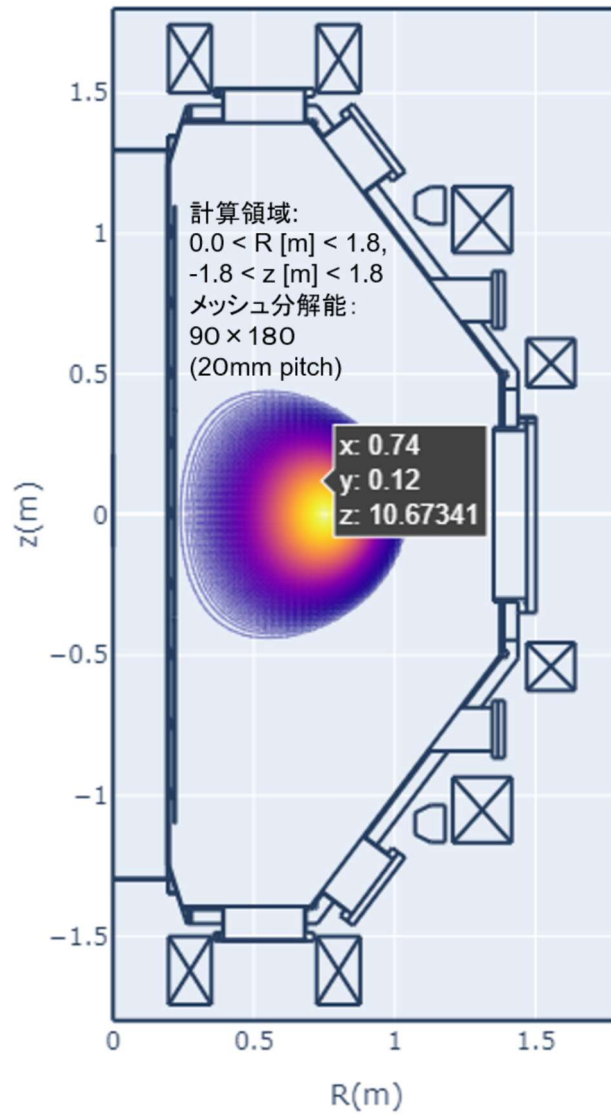


図 2-1 計算領域を示すポロイダル断面図

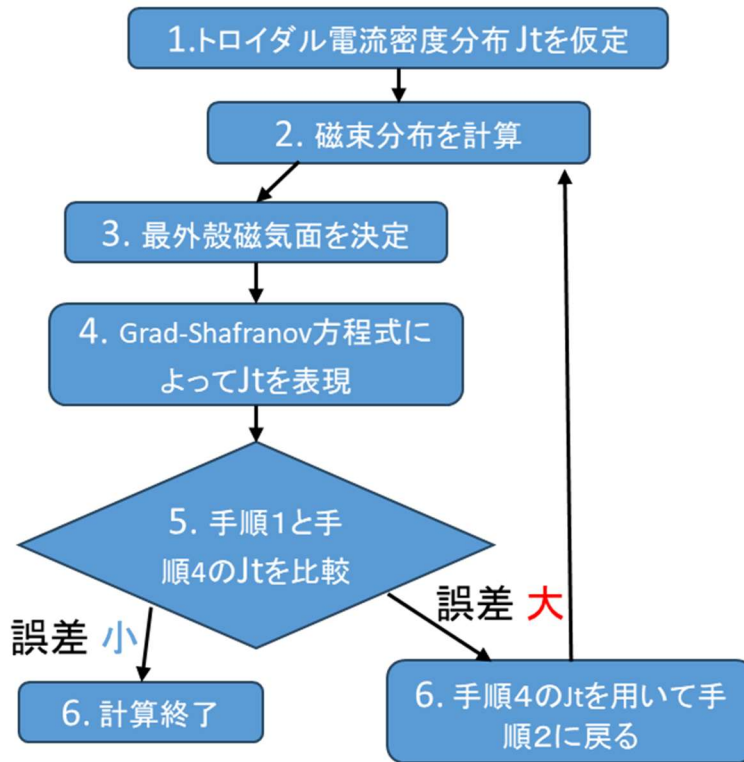


図 2-2 平衡計算のフローチャート

式(2.8)は微分形であるため、実際に圧力分布を求めるには積分する必要がある。式(2.8)は以下のように振る舞う。

$$\frac{dP}{d\psi}(x) = \begin{cases} a_0, & x = 0(\text{axis}) \\ 0, & x = 1(\text{boundary}) \end{cases} \quad (2.14)$$

また、式(2.8)の a_n の係数は

$$x^n - x^{n_{p+1}} \quad (2.15)$$

で与えられる。これを $x = 1$ において 0 となるように積分すると、

$$\frac{x^{n+1} - 1}{n+1} - \frac{x^{n_{p+2}} - 1}{p+2} \quad (2.16)$$

となる。また、式(2.10)より、

$$d\psi = (\psi_B - \psi_M) dx \quad (2.17)$$

(2.16)と式(2.17)から

$$P(\psi) = \int d\psi \frac{dP}{d\psi}(x)$$

第2章 トカマクプラズマの力学的平衡

$$\begin{aligned}
 &= (\psi_B - \psi_M) \int dx \left(\sum_{n=0}^{n_p} a_n x^n - x^{n_p+1} \sum_{n=0}^{n_p} a_n \right) \\
 &= (\psi_B - \psi_M) \sum_{n=0}^{n_p} a_n \left(\frac{x^{n+1} - 1}{n+1} - \frac{x^{n_p+2} - 1}{n_p+2} \right) \tag{2.18}
 \end{aligned}$$

特に、 $n_p = 0$ の場合について書くと、

$$\frac{dP}{d\psi}(x) = a_0(1 - x)$$

$$P(\psi) = a_0 \left((x - 1) - \frac{x^2 - 1}{2} \right)$$

[6]

このように計算することで最終的に図 2-3、図 2-4 のような圧力分布や電流分布を得ることができる。

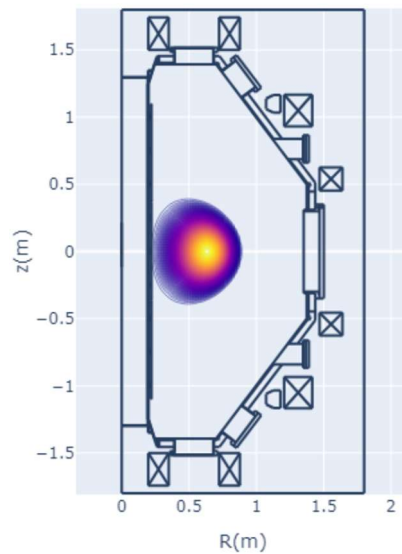


図 2-3 平衡計算によって得られるプラズマ圧力分布の例

第2章 トカマクプラズマの力学的平衡

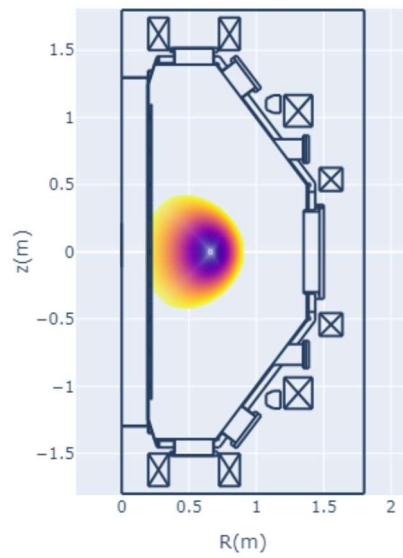


図 2-4 平衡計算によって得られるプラズマ電流密度の例

第3章 ディープラーニング

3.1. ディープニューラルネットワークの仕組み

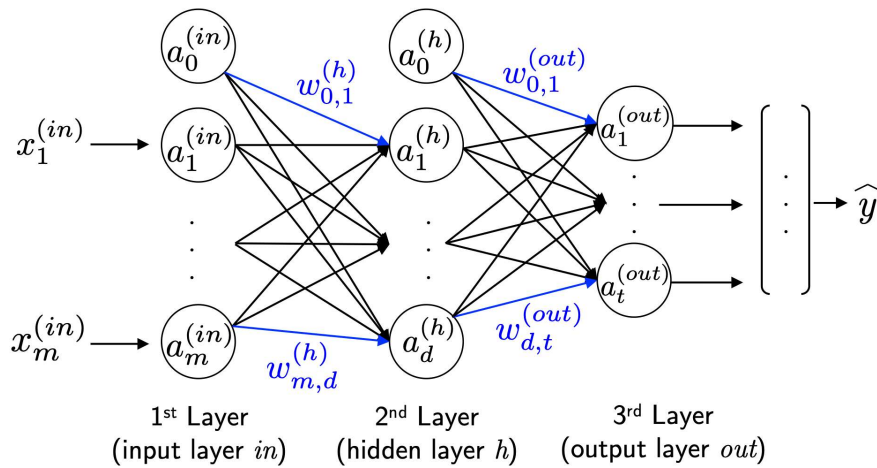


図 3-1 DNN の概念図

本研究では、プラズマパラメータの予測にディープニューラルネットワーク (Deep Neural Network, DNN) を用いている。DNN とは人工ニューラルネットワークの一種で一般的に入力層・隠れ層・出力層のうち、複数の隠れ層から構成される深い構造を持つモデルである。多層の隠れ層により、複雑なパターンの出力を予測することができるため、画像認識、音声認識や自然言語処理などの多くの技術にも利用されている。

機械学習では大きく分けて、教師あり学習と教師なし学習がある。本研究では、前述の平衡計算を利用したコードにより、様々な平衡プラズマのデータを作成しそれを教師データとした教師あり学習を行う。

学習では、各層で前の層から受けとったデータに「重み」を掛け合わせて次の層に値を渡していき、最終的な出力層で「教師データの真の値」「隠れ層から受け取った予測値」をもとにコスト関数(損失関数) $L(\theta)$ を定義する。これを最小化するように、各層での「重み」を更新していくことでモデルとしての精度を高めていく。基本的な重みの更新式は以下のようになる。

$$\theta_{t+1} = \theta_t - \eta \frac{\partial L}{\partial \theta_t} \quad (3.1)$$

η を学習率といい、一度の重みの更新での変化量を決めるパラメータとなる。このように損失関数の勾配をもとに重みを更新していくことで誤差を最小化していく方法を勾配降下法という。実際にはデータ1つずつに対して重みの更新を行っているのは、データ数が多いほどモデルの収束に時間がかかってしまうため、いくつかのデータをまとめて学習し重みの更新を行う。これをミニバッチ勾配降下法という。また、重みの更新方法には、これ以外にも多くの種類があり、アルゴリズムによって更新式は異なる。

3.2. ハイパーパラメータ

機械学習のモデルには、訓練時に設定するパラメータがあり、これをハイパーパラメータという。ハイパーパラメータには、前述の学習率や、隠れ層の数、ノード数などがあり、これらを適切に設定しなければモデルの精度はよくなる。様々なハイパーパラメータのうち、本研究では以下のハイパーパラメータを探索した。

- ・ 隠れ層の数
モデルの入力層と出力層の間にある層の数。モデルの深さを決めるパラメータで、増やすことで複雑なパターンを学習することができるが、過学習のリスクが増加する。
- ・ ノード数
モデルの各層における計算ユニットの数。隠れ層の数と同じく複雑なパターンの学習が可能になるが、過学習のリスクや計算コストの増加が懸念される。
- ・ 学習率
一度の重みの更新での変化量。大きすぎると、モデルが収束せず発散や振動してしまうが、小さすぎると学習が進まなかったり、極小値にとらわれてしまったりするリスクがある。
- ・ 最適化方法

第3章 ディープラーニング

重みの更新のアルゴリズム。学習率を学習の度合いに応じて変更するものや、直前の更新の大きさの影響を考慮するものなどがある。以下に本研究で試行した最適化方法の特徴と重みの更新のアルゴリズムについて記述する。

◇ SGD(Stochastic Gradient Descent)

最も基本的な勾配降下法。勾配そのものだけを用いて重みを更新する。汎化性能がいいことが多いが、学習率の調整が難しいなどデメリットも多い。

重みの更新式：

$$\theta_{t+1} = \theta_t - \eta g_t \quad (3.2)$$

θ : 重み

η : 学習率

$g_t = \nabla L(\theta)$: 勾配

◇ RMSprop(Root Mean Square Propagation)

勾配の2乗の移動平均をもとに、学習率を変動させながら重みを更新する。勾配が大きい場合（微量の更新で損失関数が急激に変動する点）では、学習率が小さくなり慎重に更新を進める。一方で学習率が小さくなりすぎて学習が進まないなどのデメリットも抱えている。

重みの更新式：

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{v_t + \epsilon}} g_t \quad (3.3)$$

$$v_t = \rho v_{t-1} + (1 - \rho) g_t^2$$

ρ : 減衰率

ϵ : 0 で割ることを防止する値

◇ Adam(Adaptive Moment Estimation)

前述の RMSprop と過去の勾配を考慮して学習率を決定する

「momentum」を同時に利用する方法。収束が早く安定した学習ができる。

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t \quad (3.4)$$

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t}, \quad \hat{v}_t = \frac{v_t}{1 - \beta_2^t}$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2$$

m_t : 指数移動平均

v_t : 勾配の2乗の移動平均(RMSpropのものと同じ)

[7]

3.3. Optuna を使ったハイパーパラメータチューニング

適切なハイパーパラメータの組合せを探索することをハイパーパラメータチューニングという。本研究では、python ライブラリの一つである「Optuna」[8]を用いてハイパーパラメータチューニングを行う。

探索の方法には、大きく分けて「グリッドサーチ」「ランダムサーチ」「ベイズ最適化」などがある。「グリッドサーチ」は、すべての組み合わせを網羅していく探索方法である。全組み合わせを検証するため、計算効率が悪く最適解を見つけるために非常に時間がかかってしまう。「ランダムサーチ」は探索範囲の中からランダムに組合せを試していく方法である。グリッドサーチよりは計算コストは低く、試行回数を無限に増やしていけば最適解を得られる確率は高い。しかし、実務の上では無限に試行することは不可能なことに加えて、過去の試行の情報を使用しないため効率が良いとは言えない。「ベイズ最適化」は前述の2つの方法とは異なり、過去の試行のデータを利用して次の探索点を決定する。そのため、有限の試行回数において最適解に到達する可能性が高い。「Optuna」ではデフォルトの探索方法がベイズ最適化になっている。また、ベイズ最適化にもいくつか種類が存在しており、Optuna では TPE (Tree-structured Parzen Estimator) を採用している。これは、良い試行の分布を推定し次の試行の点(組合せ)を決める方法である。この方法は計算コストが比較的軽く、限られた試行数で良い解に到達しやすいというメリットがある。しかし、探索初期の良い点周辺での探索に集中しやすく、探索が局所化するリスク

第3章 ディープラーニング

も抱えている。そのため本研究では、探索の途中で定期的にランダムな試行点をつくるように改良を加え探索の局所化の可能性を低くした方法を使用した。

表 3-1 探索方法の違い

探索方法	特徴	長所	短所
グリッドサーチ	全ての組合せを網羅的に試行する	単純な仕組みで実装が容易	計算コストが非常に高い
ランダムサーチ	無作為に選んだ組合せを試行する	グリッドサーチより効率的	過去の結果を参照しないため無駄が出る
ベイズ推定	過去の試行のデータを用いて試行点を選ぶ	限られた計算量で最適解を得られる確率が高い	初期条件などにより、探索の局所化などが起こる可能性がある

3.4. モデルの学習

3.4.1. 学習の流れ

学習モデルの具体的な訓練方法を示す。実際のモデルの訓練では、3.3節で述べたハイパーパラメータチューニングに加えて、教師データの分け方を変えてモデルの汎化性を高める交差検証なども合わせることで精度の良いモデルの完成を試みる。

以下は訓練の流れを示す。

1. 教師データを訓練データと検証データに分割する。
2. 訓練データをランダムにシャッフルし、バッチ数（一度に学習するデータ数）毎に分割する。
3. バッチから損失（誤差）を計算する。
4. 損失から勾配を計算し、重みを更新する。
5. 3～4を訓練データが尽きるまで繰り返す。
6. 1で分割した検証用データを使ってモデルを評価する。（損失を計算する）

第3章 ディープラーニング

7. 2～6を設定した epoch 数繰り返す。(epoch=訓練回数)

※実際には、設定した epoch 数を繰り返す前にモデルの精度の向上が止まる(誤差の減少が止まる)タイミングで訓練を終了させる。(early stopping)

上記が基本的な訓練の手順である。しかし、手順1の訓練データと検証データの分割パターンによってモデルの精度に偏りが出ることがある。これを防ぐために交差検証を行う。交差検証とは、手順1の分割パターンを数パターン試し、その平均の性能を調べるものである。これにより、より汎化性の高い(未知のデータに対して強い)モデルかどうかを判断することができる。これと3.3節で記述したハイパーパラメータチューニングも行った。まとめると、1つのハイパーパラメータの組合せに対して、教師データ、訓練データの分割パターンを5パターン訓練し、そのすべての分割パターンでの誤差の値の平均値が最も低かったハイパーパラメータの組合せを選択する。

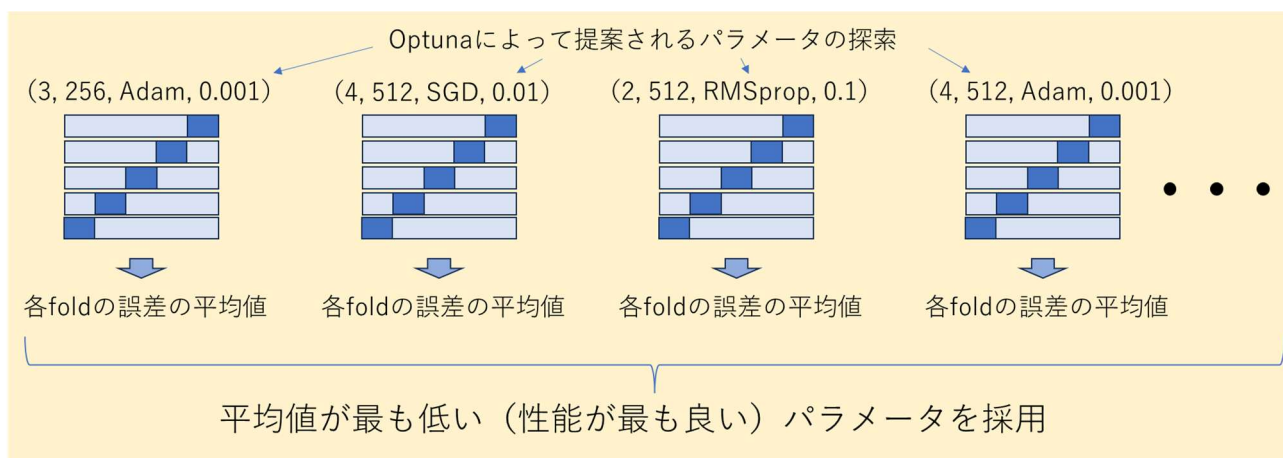


図 3-2 交差検証とハイパーパラメータの組合せ

3.4.2. 探索条件

Optuna では探索するハイパーパラメータの種類と探索範囲を指定することができる。ノード数、隠れ層は1以上の整数、学習率は0～1の間、最適化方法は上述した3種の中から選択する。ハイパーパラメータの組合せはほぼ無限

第3章 ディープラーニング

に存在するため、試行回数を事前に設定することは難しい。そこで本研究では、試行回数ではなく計算時間に対して24時間という制限を設けてハイパーパラメータ探索を行った。

ハイパーパラメータ	探索範囲
隠れ層	1~10
ノード数	32~512
学習率 η	0.001~0.1
最適化方法	SGD, RMSprop, Adam

第4章

学習モデルの作成

4.1. モデルの作成

本研究では複数パターンの教師データと、入出力のパラメータの組合せを試行し、それぞれのモデルでの比較を行った。

4.1.1. モデルの入出力パラメータ

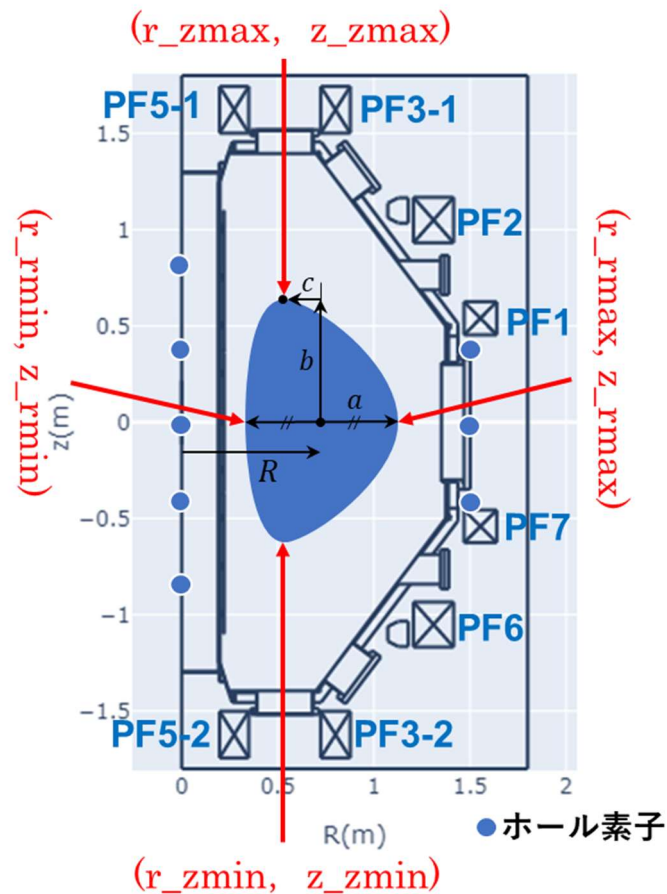


図 4-1 入出力パラメータを示したポロイダル断面図

図 4-1 に球状トカマク装置 QUEST のポロイダル断面と各 PF コイル、ホール素子の位置と最外殻磁気面の頂点座標を示す。

第4章 学習モデルの作成

また、入力・出力に使用するパラメータを以下の表に示す。

- ・入力パラメータ

inputs	
PF17	PF1, PF7コイルの電流値
PF26	PF2, PF6コイルの電流値
PF4	PF4コイルの電流値
PF35-1	PF3-1, PF5-1コイルの電流値
PF35-2	PF3-2, PF5-2コイルの電流値
h1~h8	ホール素子位置での磁場
diamag	反磁性ループの値
p1~p10	z=0軸上での圧力値

- ・出力パラメータ

outputs	
I_p	プラズマ電流[A]
axis-r, axis-z	磁気軸位置[m]
major-radius	大半径(R)[m]
minor-radius	小半径(a)[m]
elongation	楕円度(b/a)
triangularity	三角度(c/a)
r-rmax, z-zmax, r-rmin, z-zmin	最外殻磁気面の各頂点[m]
ρ_v, ρ_s	プラズマ平均圧力[Pa]
pbeta, tbeta, nbeta	β 値
q-center, q-edge	安全係数
li	内部インダクタンス[H]
stored_energy	内部エネルギー[J]

- ・ ρ_v : 体積平均、 ρ_s : 断面積平均
- ・ β 値 [9]

第4章 学習モデルの作成

プラズマの圧力と磁場強度の比。高いほど、同じ磁場で高温、高圧のプラズマを閉じ込めている。

poloidal beta(pbeta)

$$\beta_p = \frac{8\pi \langle p \rangle_s S}{\mu_0 I^2}$$

toroidal beta(tbeta)

$$\beta_t = \frac{\langle p \rangle_v}{B_{t0}^2 / 2\mu_0}$$

normalized beta(nbeta)

$$\beta_t[\%] = \beta_N \frac{I_p}{a\beta_t} [MA/m.T]$$

$\langle \rangle_s$: 面積平均

$\langle \rangle_v$: 体積平均

B_{t0}^2 : 磁気軸での真空トロイダル磁場

B_t : トロイダル磁場

- 安全係数 q [10]

磁力線のねじれ。磁力線がトロイダル方向に1周する間にポロイダル方向に周回する数の逆数。出力値には最外殻磁気面での値を q_edge、磁気軸での値を q_center としている。

$$q = \frac{d\phi}{d\psi}$$

ϕ : トロイダルフラックス, ψ : ポロイダルフラックス

- 内部インダクタンス [11]

電流密度の分布を表す指標となる値。0.5 付近でフラットな分布、大きいほど中心ピークな分布となる。

$$l_i = \frac{4\pi S \langle B_\theta^2 \rangle_s}{\mu_0^2 I^2}$$

S: プラズマ断面積

第4章 学習モデルの作成

B_θ : プラズマ由来のポロイダル磁場

- ・ 内部エネルギー [10]

プラズマに蓄えられたエネルギーの総量。

$$W = 2 \times \frac{3}{2} \int p dv = 3 \langle p \rangle_v V$$

$\langle \rangle_v$: 体積平均

V : プラズマ体積

4.1.2. 教師データ

複数作成した教師データの主な違いは、2.3節の平衡計算で利用した Grad-Shafranov の式の $\frac{dP}{d\psi}$ 、 $\frac{dI^2}{d\psi}$ を近似した多項式（基底関数）である式(2.8)、式(2.9)の違いにある。

Grad-Shafranov の式

$$j_\varphi = 2\pi R \frac{dP}{d\psi} + \frac{\mu_0}{4\pi R} \frac{dI^2}{d\psi} \quad (2.7)'$$

$$\frac{dP}{d\psi}(x) = \sum_{n=0}^{n_p} a_n x^n - x^{n_p+1} \sum_{n=0}^{n_p} a_n \quad (2.8)$$

$$\frac{dI^2}{d\psi}(x) = \sum_{n=0}^{n_p} b_n x^n - x^{n_p+1} \sum_{n=0}^{n_p} b_n \quad (2.9)$$

以下、作成したデータの特徴を記述する。

- ・ データセット 1 (n0-equilibrium)

基底関数の次数 $n_p = 0$ である単純なものを使用した。

$n_p = 0$ の時、式(2.8)は

第4章 学習モデルの作成

$$\begin{aligned}\frac{dP}{d\psi}(x) &= \sum_{n=0}^0 a_n x^n - x^{n_p+1} \sum_{n=0}^0 a_n \\ &= a_0 x^0 - x^{0+1} a_0 \\ &= a_0(1-x)\end{aligned}$$

また、積分すると

$$P(\psi) = a_0 \left((x-1) - \frac{x^2-1}{2} \right)$$

となる。 $\frac{dI^2}{d\psi}$ についても同様に計算される。この教師データでは、上記のように基底関数が単純であるため、表現できる $\frac{dP}{d\psi}$ 、 $\frac{dI^2}{d\psi}$ が限られており計算できる平衡プラズマのパターンも限定的である。

- データセット2 (n2-equilibrium)

教師データ1では表現できなかったプラズマを学習させるため基底関数を複雑化したデータセット。

基底関数の次数 $n_p = 2$ とすることで表現できるプラズマのパターンを増やした。

$n_p = 2$ の時、式(2.8)は

$$\begin{aligned}\frac{dP}{d\psi}(x) &= \sum_{n=0}^2 a_n x^n - x^{n_p+1} \sum_{n=0}^2 a_n \\ &= a_0 x^0 + a_1 x^1 + a_2 x^2 - x^{2+1}(a_0 + a_1 + a_2) \\ &= (1-x^3)a_0 + (x-x^3)a_1 + (x^2-x^3)a_2\end{aligned}$$

となる。この教師データでは、前述の教師データ1に比べて、基底関数が複雑化しているため、より多くのパターンの平衡プラズマを計算することができる。

- データセット3 (one-step-equilibrium)

第4章 学習モデルの作成

教師データ1、教師データ2はすべて上下位置が対象なプラズマのみを計算していた。しかし、実際に制御に使うことを想定する場合、上下位置が非対称なプラズマを扱う必要がある。そこで、計算が終了したプラズマをランダムに上下方向に少しずらした。このデータセットの目的は、上下位置が非対称なプラズマが登録されたデータセットを教師データに用いることで、上下位置が非対称なプラズマの予測も可能であることを検証することにある。

なお、このデータセットに用いた平衡計算コードにおける基底関数の次数は $n_p = 0$ である。

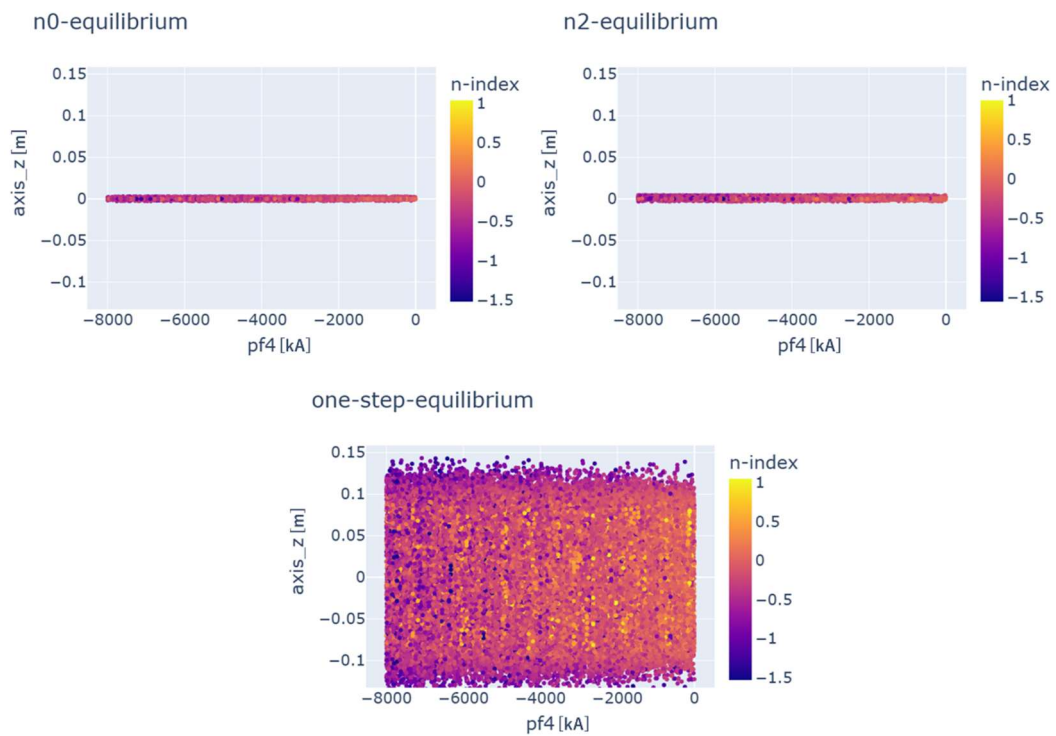


図 4-2 各データセットの磁気軸位置の分布

図 4-2 は縦軸に磁気軸の z 位置、横軸に PF4 コイルの電流値を取った散布図である。カラーは n -index を表している。 n -index は、プラズマの上下位置の不安定性を示す値であり、正の値であるときに上下位置安定となる。n0-equilibrium と n2-equilibrium の散布図を比較すると、ほとんど分布に差がないことがわかる。この2つは基底関数の次数が異なり、プラズ

第4章 学習モデルの作成

マ内部の電流分布の表現の幅に大きな差がある。しかし、QUESTのPFコイルが上下対称に存在しており、対称位置にあるPFコイルには等しい電流値が流れることと、プラズマの初期位置として与える初期条件で $z=0$ で固定であることからプラズマは上下方向にほとんど移動しない。それにより、この2つの磁気軸の上下位置の分布は $z=0$ 付近に存在している。一方、one-step-equilibriumは、平衡計算収束後、上下方向に微量ずらし再度2.3節の平衡計算ループを1回行ったデータを登録している。そのため、このデータセットでは約 $-0.15\sim 0.15$ の間に分布しており、これが他2つのデータセットと最も異なる点である。

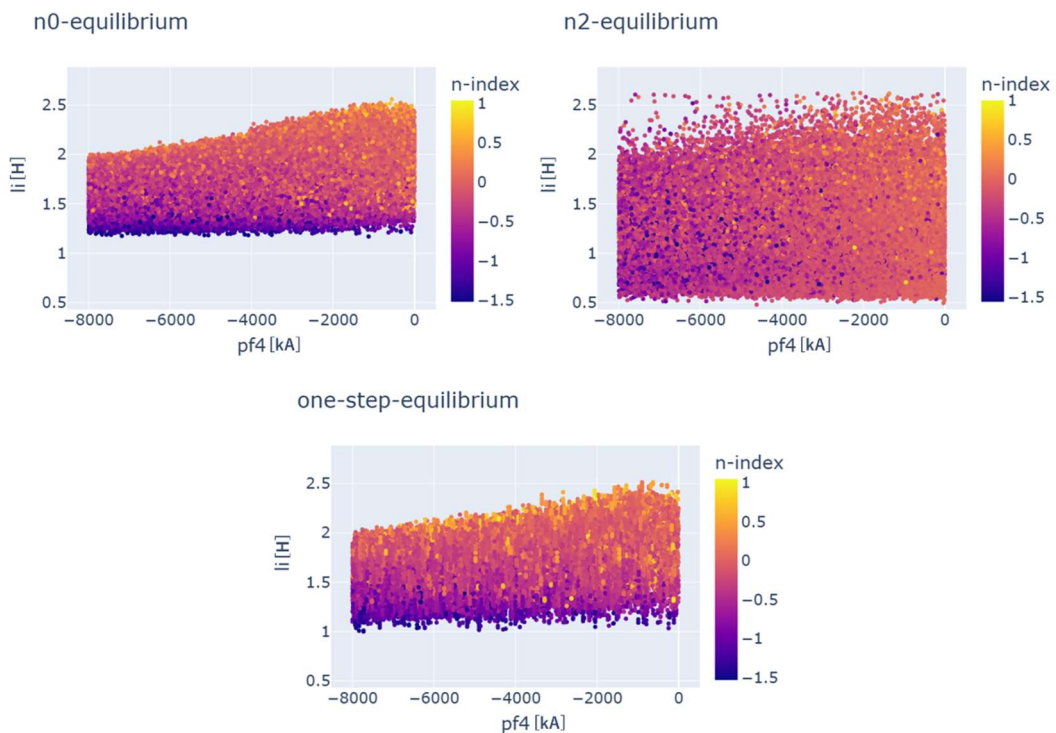


図 4-3 各データセットの内部インダクタンスの分布

図 4-3 は縦軸に内部インダクタンス、横軸に PF4 コイルの電流値を取ったものである。n0-equilibrium と n2-equilibrium を比較すると分布の範囲が大きく異なることが分かる。内部インダクタンスは電流分布と大きく関連のあるパラメータである。基底関数の次数を増やし、プラズマの電流分布のパターンを増加させたことで、内部インダクタンスの分布も大きく広が

第4章 学習モデルの作成

った。電流分布がフラットのとき、内部インダクタンスは0.5になる。n0-equilibriumの散布図では0.5近傍にはデータ点が存在していないことから、このデータセットには電流分布がフラットなプラズマが存在していないことがわかる。一方、n2-equilibriumの散布図は0.5付近から2.5付近まで広く分布している。このことから、n2-equilibriumの方がより多くのパターンのプラズマを表現している。また、one-step-equilibriumの基底関数はn0-equilibriumと同じであるため内部インダクタンスの分布も類似していると考えられる。

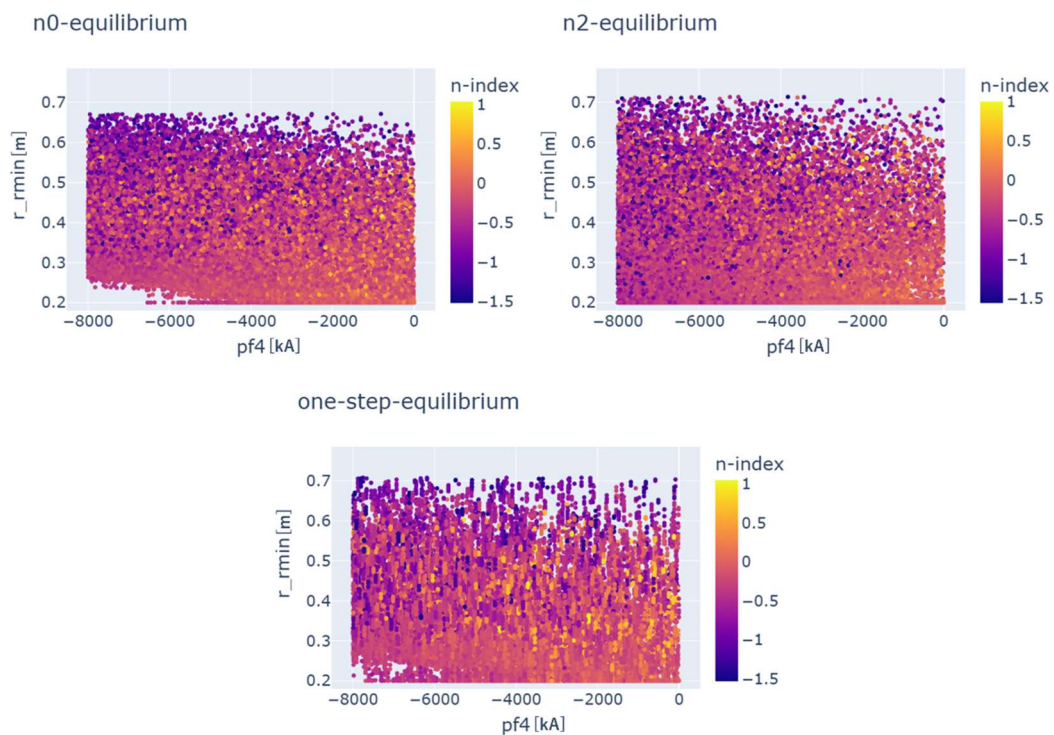


図 4-4 各データセットの最外殻磁気面の頂点座標の分布

図 4-4 は、縦軸にプラズマの r_{rmin} 、横軸には PF4 コイルをとった散布図である。この散布図では先に示した内部インダクタンス、磁気軸位置を縦軸に取った散布図と異なり 3 つのデータセット間で分布の範囲に大きな差異はない。これにより、この 3 つのデータセットでは、基底関数の違いによりプラズマの内部インダクタンスなどのパラメータには表現の幅の差

第4章 学習モデルの作成

異が大きくあるが、位置形状についてはどれも同等に表現できていると考えられる。

4.1.3. データの前処理

教師データは学習モデルに入力する前に前処理を行う必要がある。前処理は2段階で行った。

1. 3シグマ法

データセットを

$$\mu \pm 3\sigma$$

μ : 平均値

σ : 標準偏差

の間で抽出し、外れ値を除外する。

2. 標準化

$$x'_i = \frac{x_i - \mu}{\sigma}$$

上式に従って、各データを標準化する。標準化とは、データの平均を0、標準偏差を1とする操作である。学習に用いるデータは様々なプラズマパラメータを含んでいる。それらのパラメータはそれぞれ異なるスケールであるため、学習モデルの訓練時にスケールの大きな特徴量に影響されて正常に学習が完了しない可能性がある。標準化の操作をすることでこういったパラメータのスケールに左右されずに訓練を行うことができる。[11]

4.2. モデル作成の条件

4.1節で示した入力・出力パラメータと教師データを組み合わせて5つの学習モデルを作成した。

表 4-1 各モデルの概要

	教師データ	入力値	出力値
model1	n0-equilibrium	A	全て
model2	n0-equilibrium	A + diamag, p1~p10	全て
model3	n2-equilibrium	A	全て
model4	n2-equilibrium	A + diamag, p1~p10	全て
model5	one-step-equilibrium	A	全て
A	(PF17), (PF26), (PF4), (PF35-1), (PF35-2) (h1), (h2), (h3), (h4), (h5), (h6), (h7), (h8)		

4.3. 誤差の評価方法

本研究では予測したパラメータの中に、物理的な意味の異なる出力が混在しているため、誤差の評価方法をパラメータの種類に応じて使い分けた。

- MAPE (Mean Absolute Percentage Error)

$$MAPE = \frac{1}{n} \left(\sum \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) \times 100$$

n : データ数

y_i : 真値

\hat{y}_i : 予測値

平均絶対パーセント誤差。スケールに対して相対的に誤差の評価が必要な圧力や電流値などの大きさを示す値に対して用いた。

- MAE (Mean Absolute Error)

$$MAE = \frac{1}{N} \sum |y_i - \hat{y}_i|$$

第4章 学習モデルの作成

絶対平均誤差。最外殻磁気面の頂点座標などのように座標を表すパラメータに対して使用した。相対的な評価ではなく、単位が存在する絶対的な評価を行う。[12]

MAPEで評価したパラメータ	
	I_p
	major-radius
	minor-radius
	elongation
	triangularity
	ρ_v, ρ_s
	pbeta, tbeta, nbeta
	q-center, q-edge
	li
	stored_energy

MAEで評価したパラメータ
axis-r, axis-z
r-rmax, z-zmax, r-rmin, z-rmin, r-zmin, z-zmin

図 4-5 パラメータの種類による誤差評価方法の違い

第5章

予測結果

5.1. 各モデルのハイパーパラメータ

入力値、出力値、使用するデータセットが異なると、最適なハイパーパラメータの組合せも変化する。本研究ではそれぞれの入出力パラメータに適したハイパーパラメータを用いて作成したモデルの精度を比較する。以下にチューニングによって選択された各モデルのハイパーパラメータを示す。

表 5-1 ハイパーパラメータチューニング結果

	隠れ層	ノード数	学習率	最適化方法
model1	4	333	0.0903	SGD
model2	4	509	0.0968	SGD
model3	3	80	0.0010	Adam
model4	3	266	0.0560	SGD
model5	6	402	0.0807	SGD

5.2. 各モデルの比較

- モデル1 vs モデル2

この比較では、同一のデータセットに対して入力値の違いが及ぼす影響を確認する。

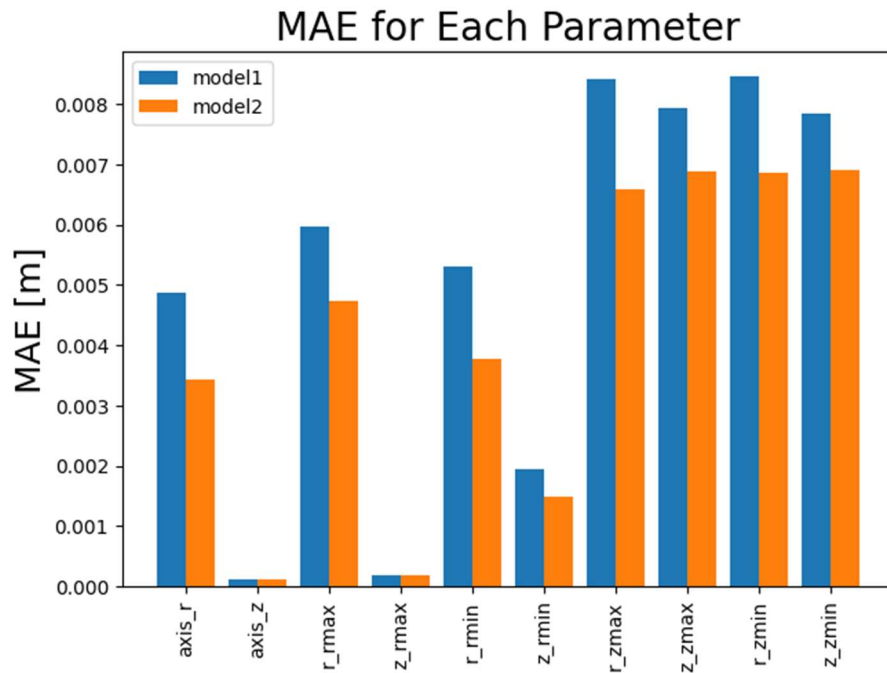


図 5-1 モデル 1 とモデル 2 の MAE 比較グラフ

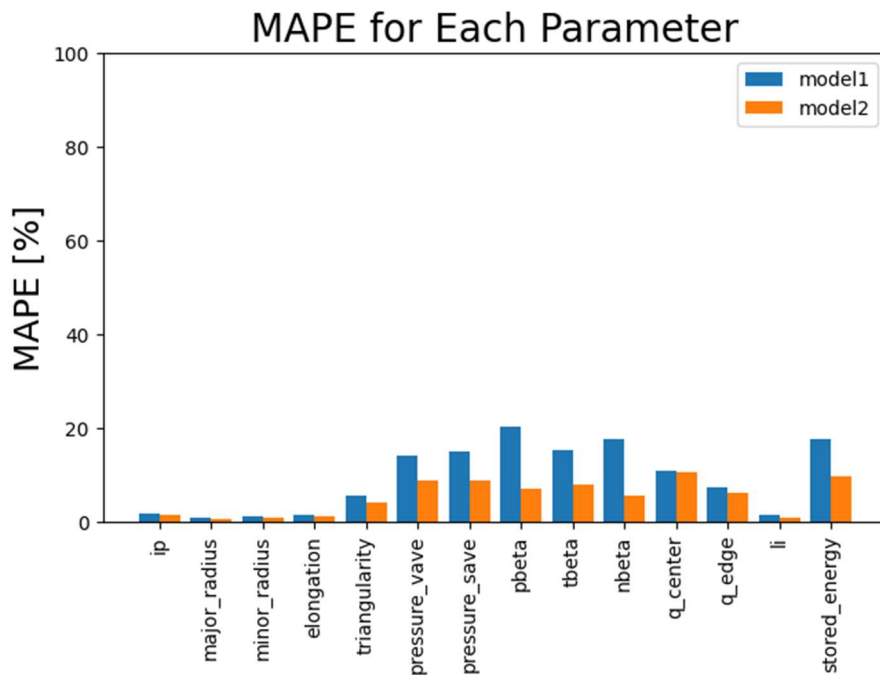


図 5-2 モデル 1 とモデル 2 の MAPE 比較グラフ

モデル 2 では、 $z=0$ 軸上の計 10 点の圧力値と、反磁性の値が入力値として追加されている。この 2 つの値は、主にプラズマ内部のパラメータと相

第5章 予測結果

関が強いため、内部インダクタンスや内部エネルギー、 β 値などの予測精度が上がることを期待される。MAE のグラフは、縦軸は単位を持っている。そのことを踏まえてモデル1とモデル2を比較すると数 mm 程度ではあるが精度が向上していることが分かる。これにより増加した入力値が位置座標には関係が薄いものであっても予測精度の向上が認められた。一方 MAPE のグラフでは、内部エネルギーや β 値などと相関の高い入力値を加えたことでそれらのパラメータの予測精度が向上していることがわかる。ただし、モデル1でも誤差は大きくても20%程度と、ある程度の予測は行えている。これらのパラメータはモデル1の入力値からは予測が難しいものであるが、ある程度の予測ができている理由として教師データの基底関数が単純で内部パラメータの分布範囲が狭いことが考えられる。

・ モデル1 vs モデル3

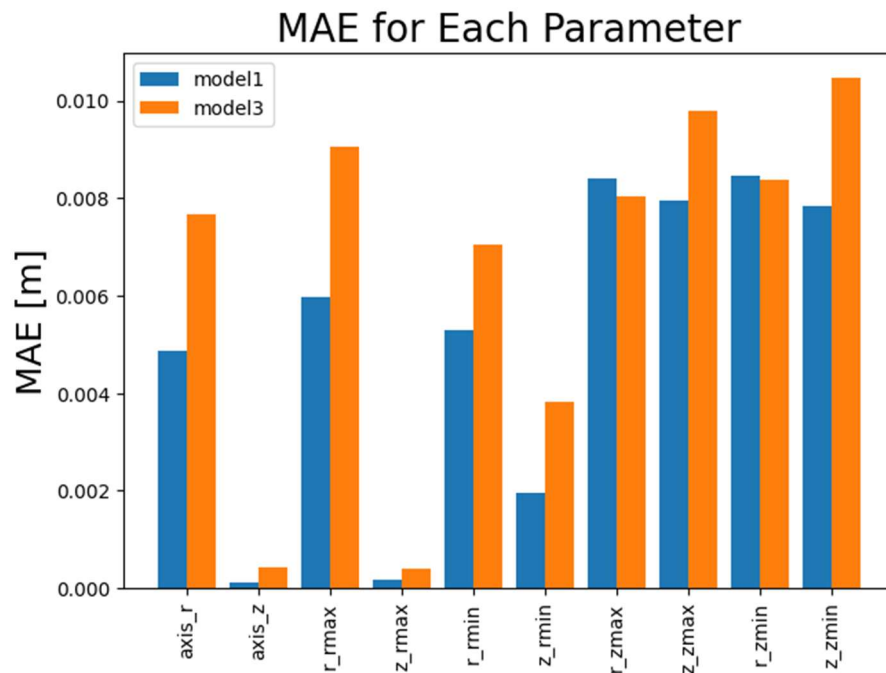


図 5-3 モデル1とモデル3の MAE 比較図

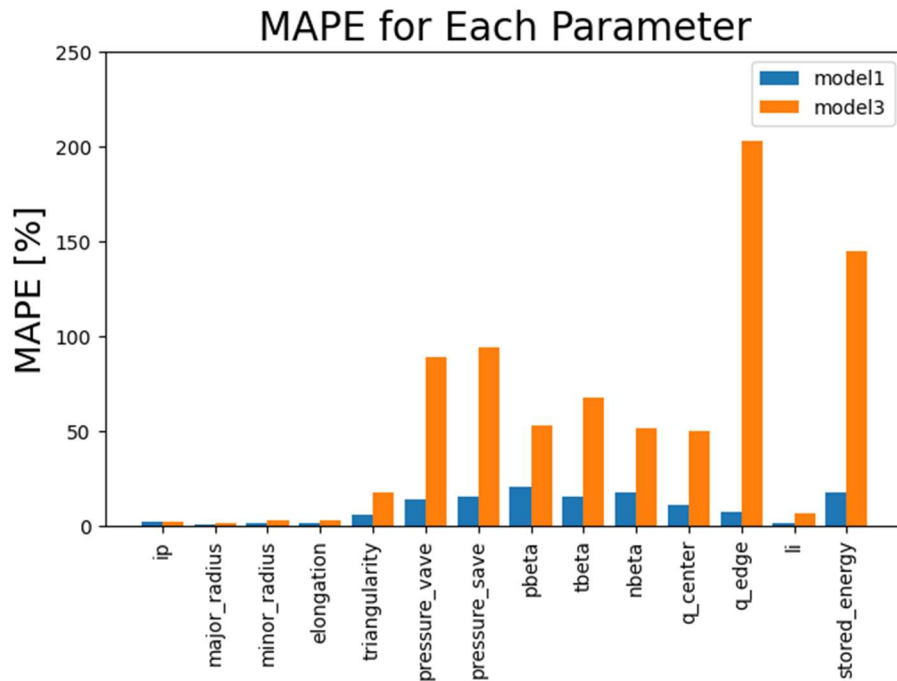


図 5-4 モデル1とモデル3の MAPE 比較図

モデル1とモデル3は入力値が等しく、教師データが異なったモデルである。モデル3に使用した教師データは、基底関数が複雑化しているため、より多くのパターンの平衡プラズマがデータセットに含まれている。MAEで評価した位置形状のパラメータについては、モデル3の方が精度が悪いがその差は数ミリ程度であり、実用上問題はないと考えられる。一方で、主に MAPE で評価しているパラメータは基底関数の違いの影響を大きく受けており、予測誤差も非常に大きくなってしまっている。特に誤差の大きな pressure_vave, q_edge, stored_energy について、予想値と真値の比較図を図 5-5 に示す。図は縦軸に予測値、横軸に真値を取ったもので図中の赤い点線は $y=x$ を示すものであり、これに沿っているほど予測精度が高いと考えられる。モデル3の q_edge のグラフは全体的に膨らんでしまっており、予測精度が低いことがこの図からもわかる。一方で stored_energy のグラフは MAPE のグラフの結果とは異なり、 $y=x$ の直線に沿っていることから予測誤差は大きくないように見える。原因として、MAPE の計算時に数点の非常に大きなパーセント誤差が平均値を大きく引っ張り上げてしまっていることが考えられる。ただし、モデル3のグラフにおいて大きな

第5章 予測結果

値付近では予測値は系統的に低くなっていることがわかる。これは、その他の出力パラメータ同士の相関によって引き起こされたものだと推察される。

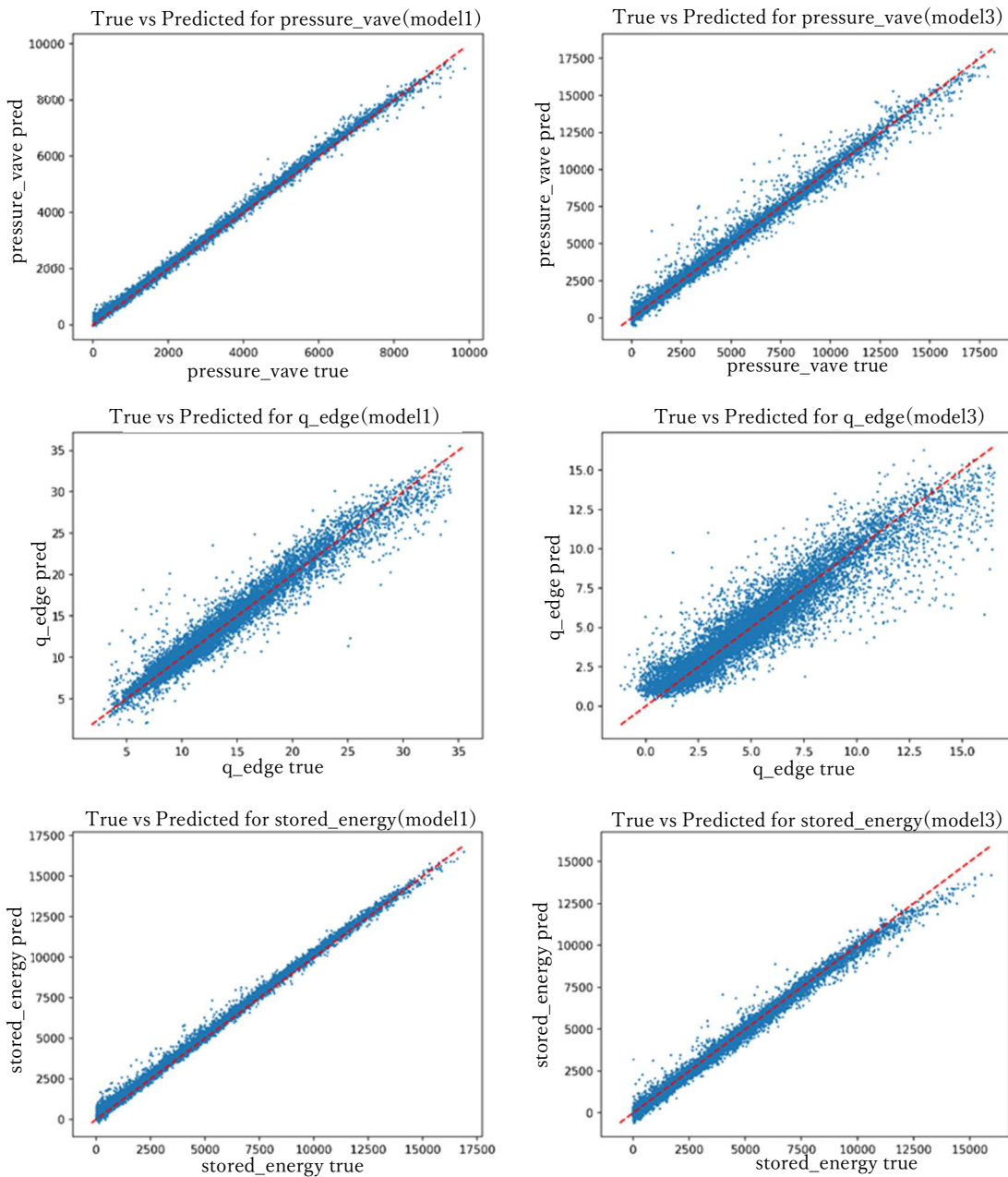


図 5-5 予測値 vs 真値プロット (モデル 1 とモデル 3)

- モデル 3 vs モデル 4

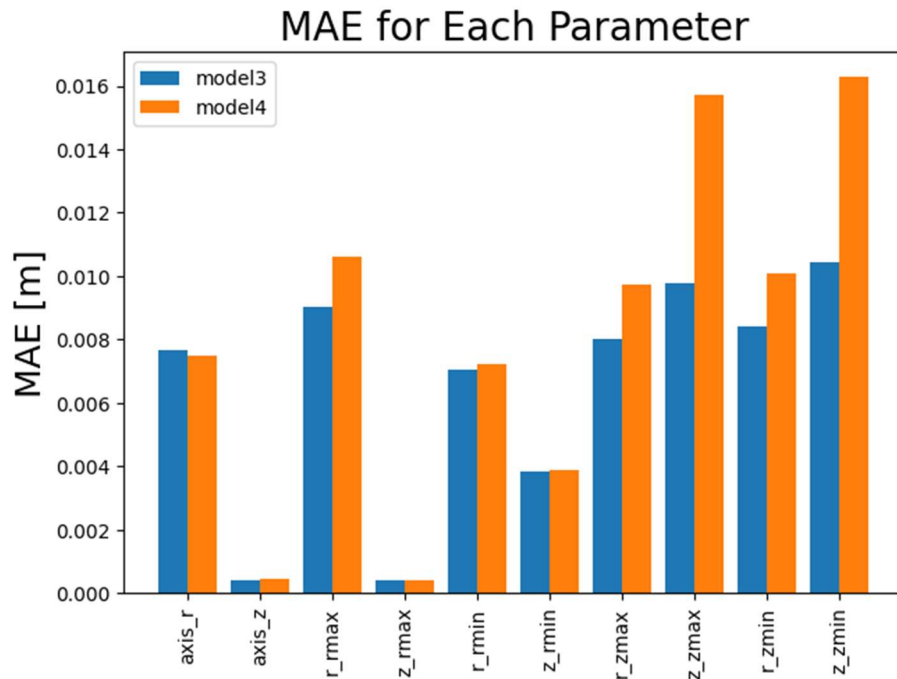


図 5-6 モデル3とモデル4の MAE 比較図

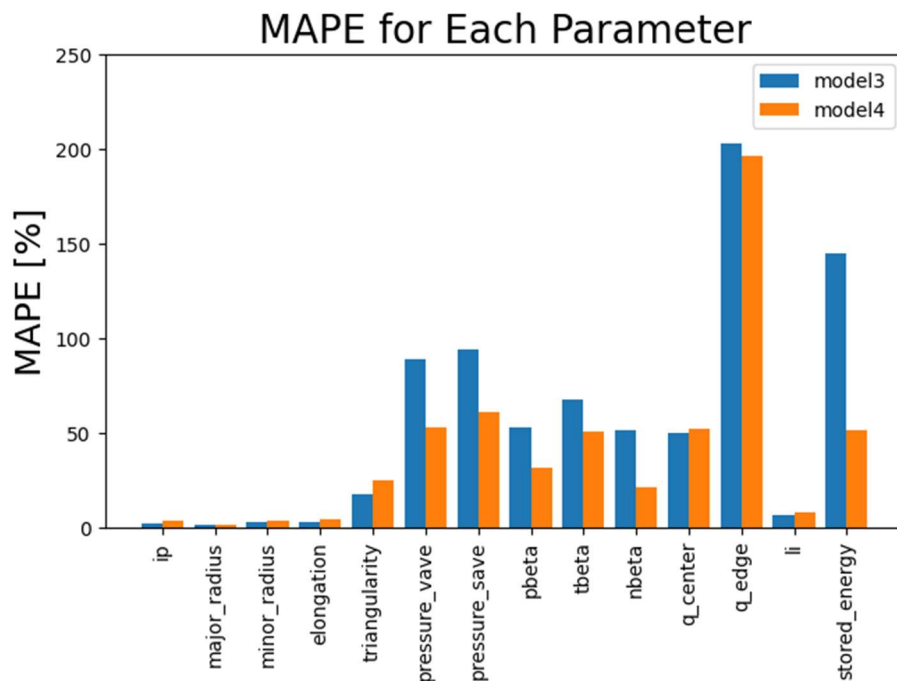


図 5-7 モデル3とモデル4の MAPE 比較図

モデル3とモデル4は同じ教師データを用いており、入力値が異なっている。MAPEのグラフを見ると特に内部エネルギーや平均圧力の予測精度

第5章 予測結果

が大きく向上していることが分かる。これは、入力値に内部のパラメータと相関のある反磁性の値や、圧力の値を加えたからであると推察される。しかし、同様に内部のパラメータである内部インダクタンスの精度は低下、安全係数の値の MAPE はモデル4でも約200%近くの誤差を出してしまっており、予測精度が高いとは言えない。また、位置形状の値も予測精度が低下しており、微差であるため大きな問題ではないが精度の向上という目的には反する。

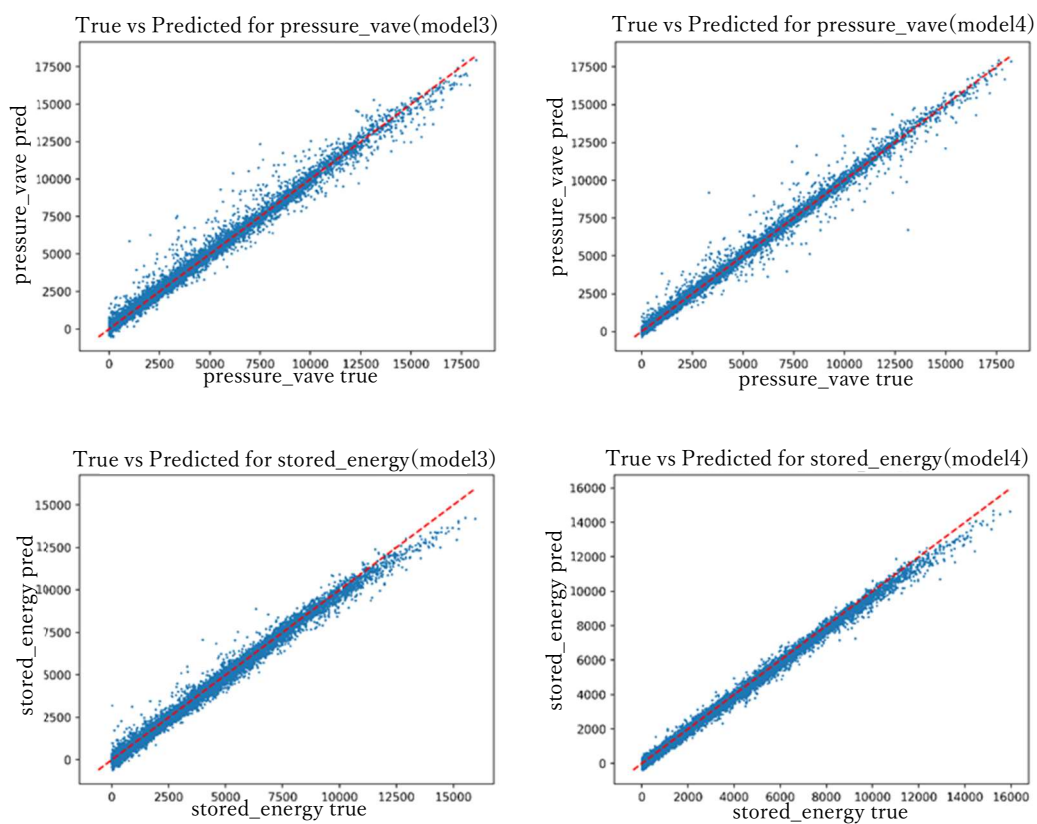


図 5-8 予測値 vs 真値 (モデル3 とモデル4)

図は pressure_vave、stored_energy についての予測値と真値のグラフである。特に stored_energy のグラフを比較すると僅かではあるが、モデル4の方が $y=x$ 直線に沿っていることがわかる。

- モデル1 vs モデル5

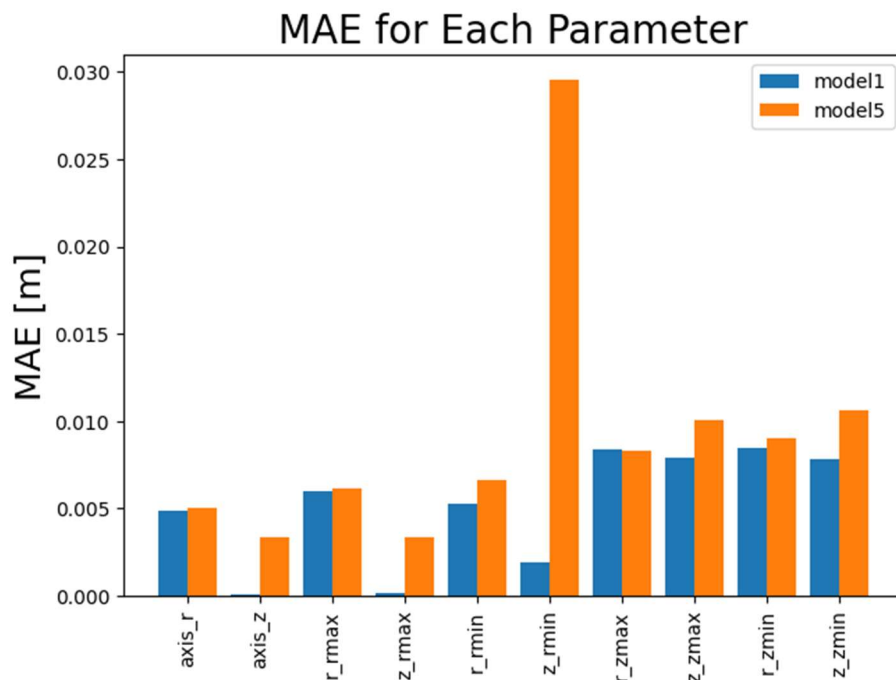


図 5-9 モデル 1 とモデル 5 の MAE 比較図

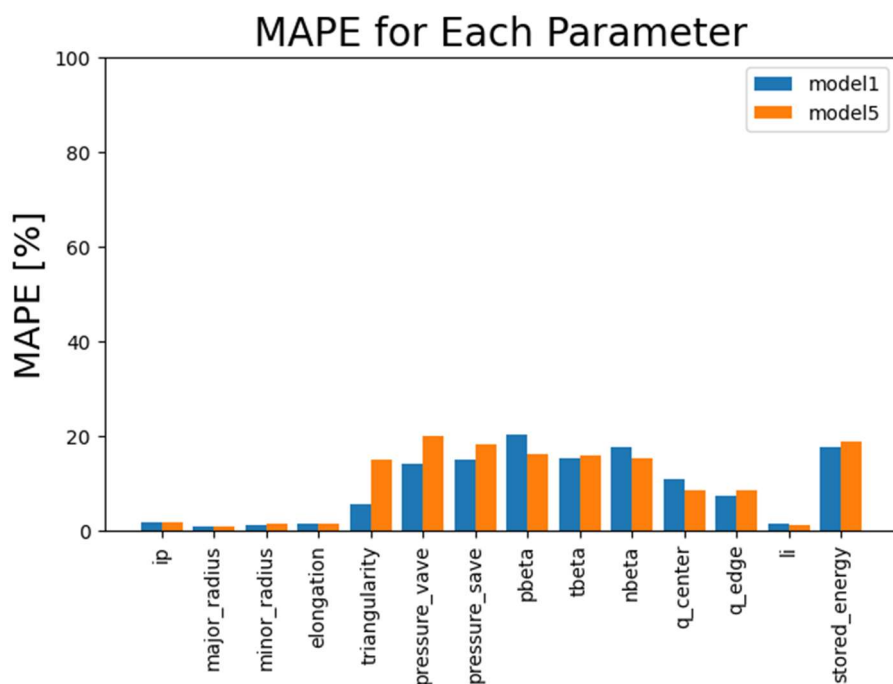


図 5-10 モデル 1 とモデル 5 の MAPE 比較図

モデル 1 とモデル 5 の入力値は同じである。モデル 5 に使用した教師データには上下位置非対称なプラズマが存在している。モデル 1 のデータセ

第5章 予測結果

ットでは、プラズマは上下に動かない。MAEのグラフを見ると磁気軸位置である $axis_z$ 、最外殻磁気面の z_rmin , z_rmax はモデル1ではそもそも動かないことから誤差もほとんどない。一方でモデル5の教師データでは上下位置に変動があり、予測誤差は大きくても3cmに収まっておりこれは、データセットの分布範囲と比較すると大きな誤差ではない。MAPEのグラフでは両モデル間の誤差にほとんど差がないが、これは基底関数が同一のものであるためだと考えられる。結果、上下位置がずれてしまったプラズマであっても位置の予測は十分可能であることが分かった。

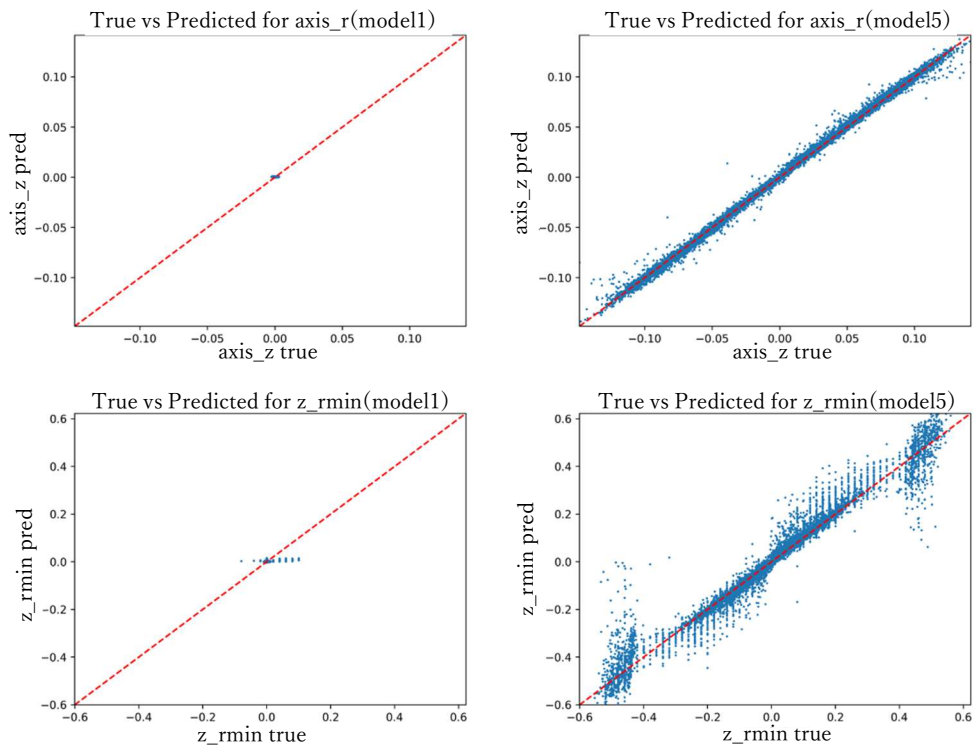


図 5-11 予測値 vs 真値 (モデル1とモデル5)

モデル1における $axis_z$ および、 z_rmin は教師データそのものの値域が狭いため、上図では $y=x$ 直線に沿ってなくてもほとんど誤差がない。一方でモデル5では、上下位置を移動させているため、値域もほかのモデルと比較して広がっている。 $axis_z$ の図を見ると、 $y=x$ 直線に沿っていることから予測の精度もいいことが分かる。しかし、 z_rmin の図では誤差に系統的な法則があるように見られる。教師データの作成時において z_rmin の計

算過程で系統的なズレを生んでしまっている可能性などが原因と考えられる。

5.2.1. 中央値でみる誤差

モデル3、モデル4のMAPEのグラフ(図5-7)を見ると、パラメータによって値が200%近くまで高くなってしまっていることが分かる。しかし予測値-真値のグラフ(図5-8)では大きくずれている様子は確認できない。そこで誤差の平均値ではなく中央値(データを大きさ順に並べたときに真ん中に位置する値)で比較すると以下のようなになる。

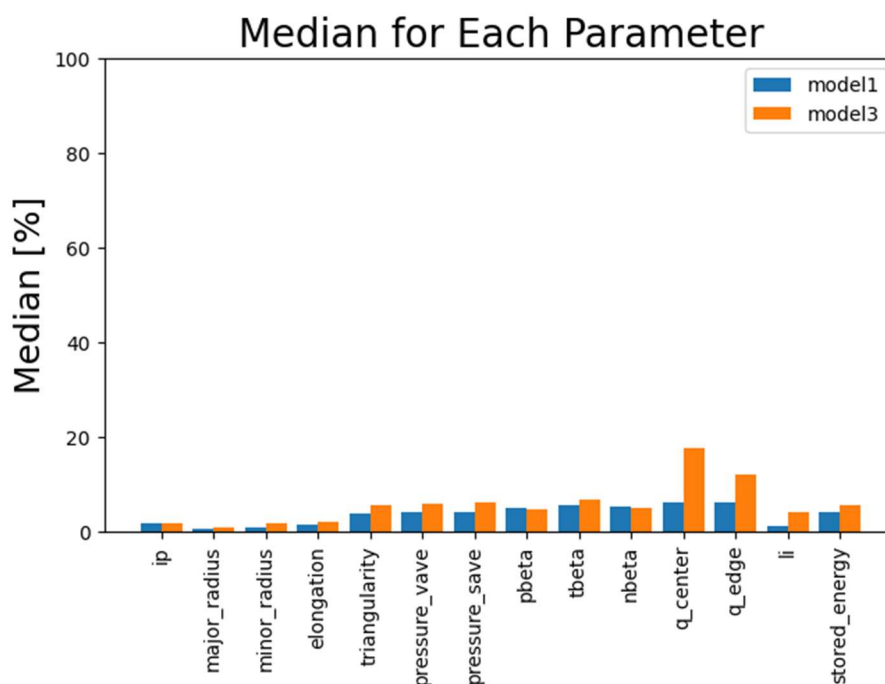


図 5-12 モデル1とモデル3の中央値比較

モデル1とモデル3の比較をしてみると平均値比較に比べて、モデル3の誤差は大きくはないが、基本的にはモデル3の方が予測精度は低く、やはり教師データの基底関数の複雑化が起因しているとみられる。

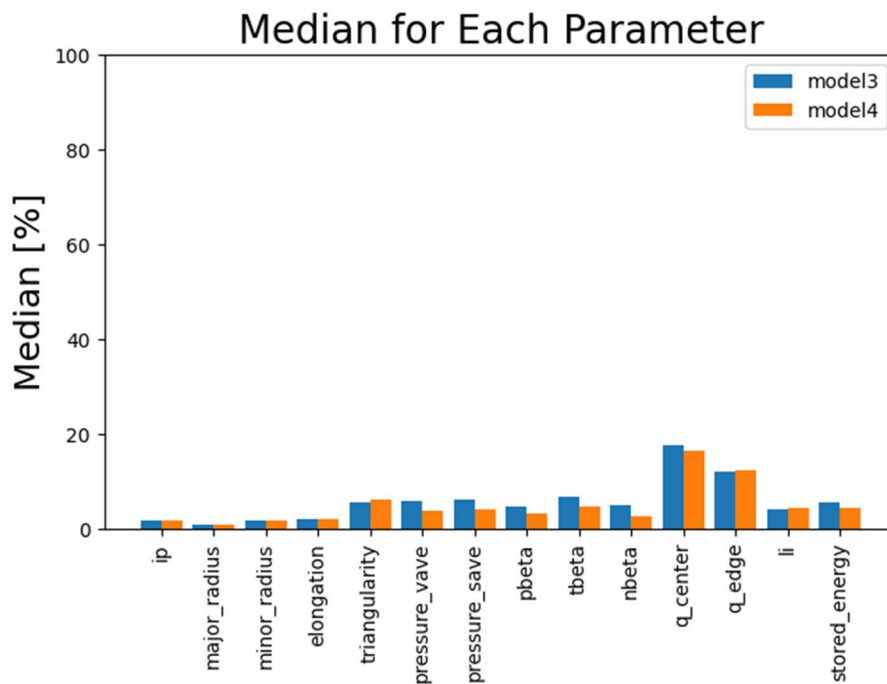


図 5-13 モデル3とモデル4の中央値比較

モデル3とモデル4の比較も中央値で行うと僅かではあるものの、モデル4が精度は高まっていることがわかり、入力値の差により精度に差が出ることを改めて確認できた。

平均値と中央値で結果が大きく異なる原因を考えるため、以下に model3 における「stored_energy」の誤差が大きくな順にデータを並べ替えた図を示す。

第5章 予測結果

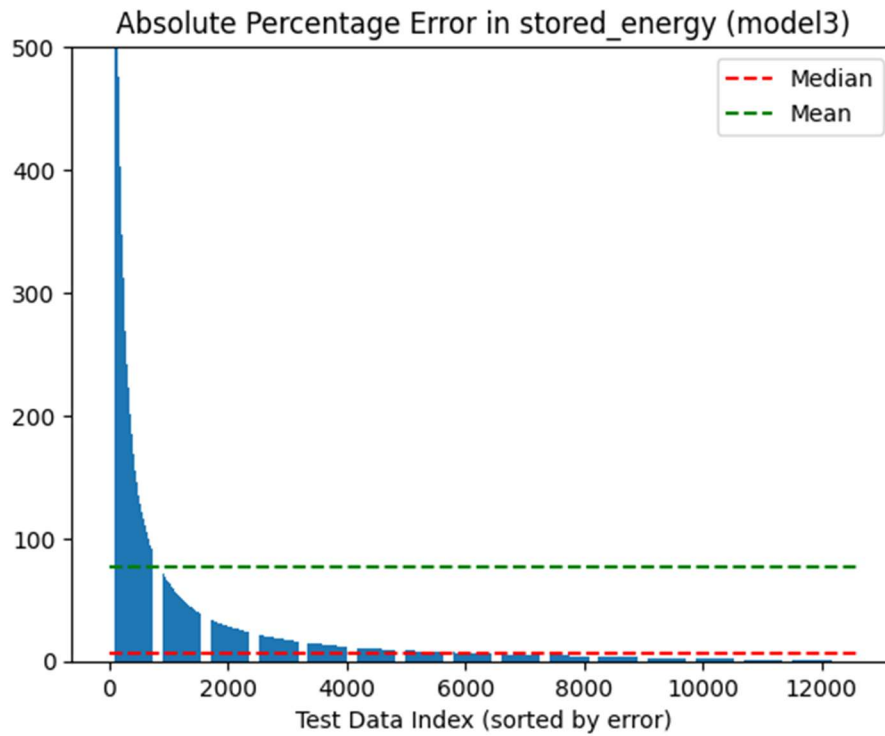


図 5-14 model3 における stored_energy のパーセント誤差 (降順)

図は視認性を考慮し、縦軸の最大値を 500%としてある。データのおよそ 1000 個が誤差 100%をはるかに超えている。

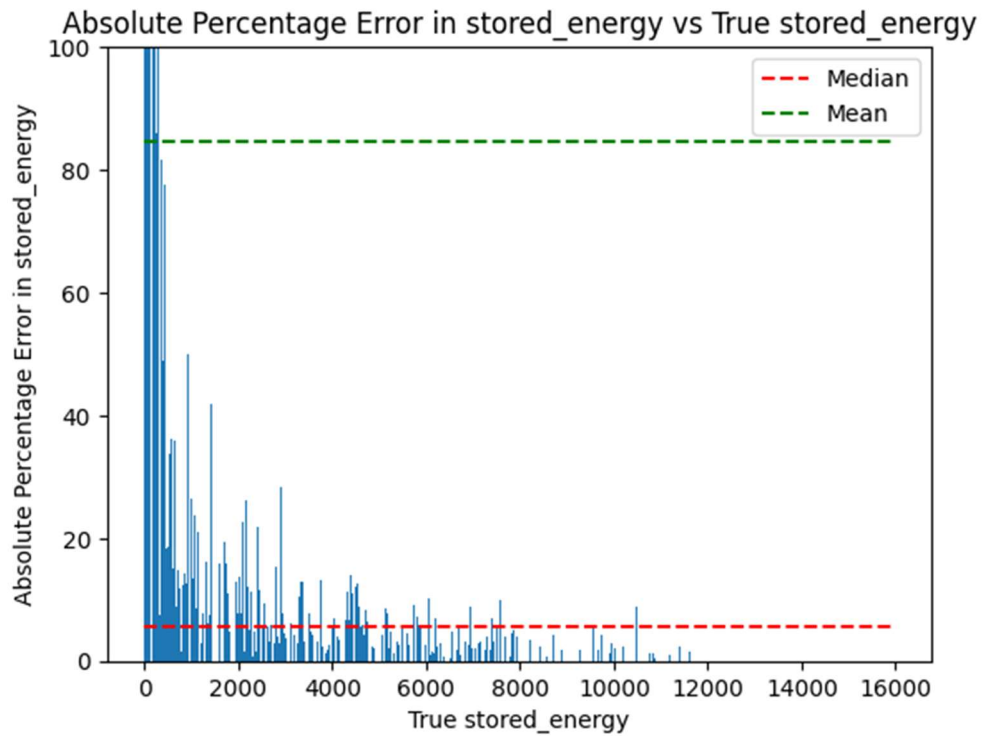


図 5-15 stored_energy のパーセント誤差と真値 (model3)

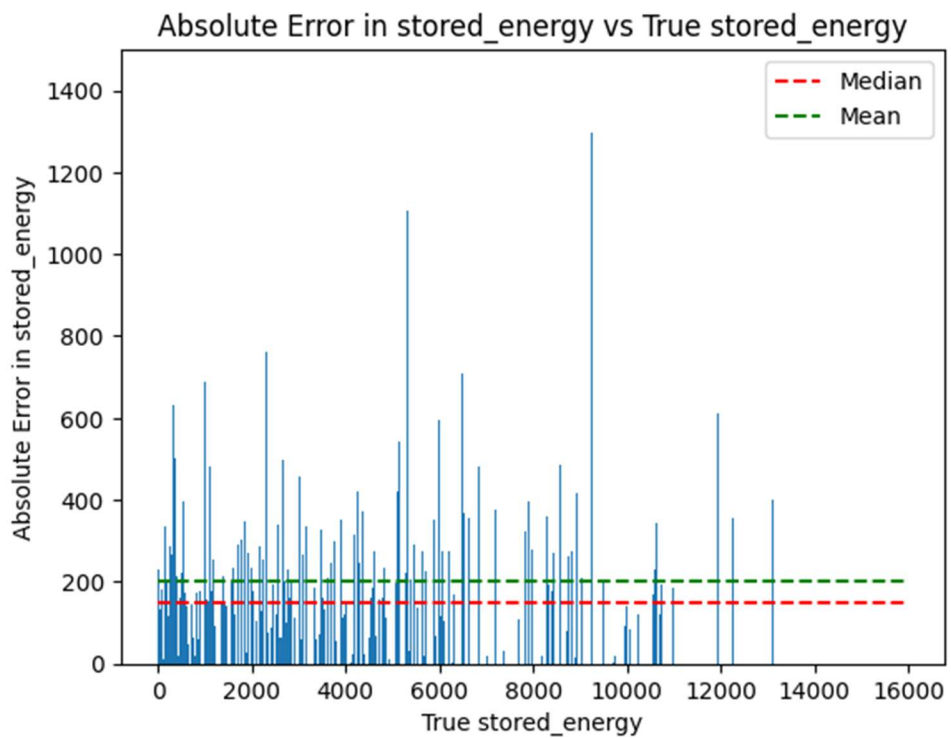


図 5-16 stored_energy の絶対誤差と真値 (model3)

第5章 予測結果

図 5-15 から、パーセント誤差は真値が小さいほど大きく、特に 0 付近で大きく跳ね上がっていることが確認できる。一方で、図 5-16 より、絶対誤差は真値が大きい時でも 800~1200 の値を取っている。絶対誤差の平均値は約 200 であり、真値の値域が 0.034~15957.9であることを考慮するとパーセント誤差でみるほど大きな誤差ではないといえる。

これらの結果を踏まえると、MAPE による評価では元のデータの値が非常に小さい場合パーセント誤差を用いて比較した際に、値が跳ね上がってしまい正確な評価が難しい。ただし、絶対誤差で評価した場合でも真値が 0 付近のデータに対して 100 以上の誤差を出してしまっていることは無視できない。これらの原因として使用したデータセットに本研究で使用した前処理では省けなかった外れ値のようなデータが混在している可能性が考えられる。

第6章

まとめ

6.1. まとめ

本研究では九州大学の球状トカマク装置 QUEST を対象としたプラズマ制御のためのリアルタイムプラズマパラメータ把握用学習モデルの作成とそのため
の教師データの作成を行った。

教師データについて、現時点で使用できるデータセットでは、基底関数の次
数の高いほど多くのパターンの平衡プラズマを表現できることが確認された。
このデータセットを用いた学習モデルでは、基底関数が単純な教師データを用
いて訓練した学習モデルに比べると、精度が落ちてしまったものの、より現実
に即したモデルであると考えられる。

学習モデルの入力値と出力値の関係性について、基本的に入力値の種類が多
い方が精度が高いことが分かった。特に出力値との相関が強いパラメータを入
力値に加えることで対応する出力値の精度が向上することを確認した。

上下非対称なプラズマに対する予測では、上下位置が移動したものであっても磁気軸位置、及び最外殻磁気面の頂点座標などは予測することが可能であることを確認した。

6.2. 課題

今後、実際に制御に用いることを想定した場合、本研究で作成した学習モデルでは以下のような課題があるといえる。

- ・ 誤差の許容範囲

本研究では、パラメータの特性に合わせた誤差の評価を行ったが、この誤差が実際に制御に用いることができるレベルであるかどうかを確かめる

第6章 まとめ

必要がある。そのため実際に制御に組み込みリアルタイムでの把握が可能かどうかの検証を行うことが望ましい。

- 教師データの妥当性

基底関数を複雑化することでより多くのパターンの平衡プラズマをデータセットとして登録し訓練することはできたが、このデータセットに含まれるすべてのプラズマが教師データとして相応しいかはわからない。大きな誤差を生んだデータ点を特定しパラメータの確認を行うこと、さらに別の基底関数などを用いてデータセットに登録するプラズマのパターンを増やしたモデルを訓練することが望ましい。また、上下位置非対称なプラズマのデータについて、本研究では上下位置がずれたものでも予測可能かという立場での教師データの利用であったが、実際に上下にずれたプラズマの制御を視野に入れる場合、データセットに登録されるプラズマが正しいプラズマであるか検証が必要であると考えられる。

- 入出力パラメータの前処理

入力パラメータ同士、出力パラメータ同士での相関関係により、予測値が系統的にずれてしまうことが考えられるため、主成分分析などを用いてそれぞれのパラメータを適切に処理して使用方法を検討する必要がある。

- 制御に要する時間の問題

本研究において作成した学習モデルを制御に組み込む場合。実験装置に取り付けられている計測機器からデータを受け取り、前処理を行った後に予測、出力、その後、後処理でデータを可視化したのちに制御プログラムを起動するという手順を踏む必要がある。これらが許容時間内に完了するか否かの検証も必要である。

参考文献

- [1]環境省, ”トリチウムの環境中での存在量”, 2025
<https://www.env.go.jp/chemi/rhm/current/02-05-18.html>
- [2]長壁正樹, “核融合(フュージョンエネルギー)の概観と課題”.
<https://www.nifs.ac.jp/about/sn/ffs2024-02-01.pdf>.
- [3] 渡邊智彦, プラズマ物理学 講義メモ, 2014, https://www.p.phys.nagoya-u.ac.jp/%7Etomo/PlasmaPhysics2014/PlasmaPhysics_Memo13.pdf
- [4] 長谷川真, Derivation of the grad-shafranov equation, 2025,
https://gitlab.com/hasemac/tokamak_equilibrium/-/blob/main/doc/grad_shafranov_eq.md
- [5] 長谷川真, procedure of equilibrium calculation, 2025,
https://gitlab.com/hasemac/tokamak_equilibrium/-/blob/main/doc/equilibrium_en.md
- [6] 長谷川真, Polynomial basis function, 2025,
https://gitlab.com/hasemac/tokamak_equilibrium/-/blob/main/doc/basis_polynomial.md
- [7] 吉崎響, 機械学習の Adam や RMSProp などの最適化手法まとめ : 効率的なモデル構築の秘訣, 2025
<https://zero2one.jp/learningblog/machine-learning-optimization-methods/?srsltid=AfmBOorLTgfSbmlrAM1SMav58XVYHFHYnHfQptWOV6rf3sfVKL2cD9G0>
- [8] OPTUNA, <https://optuna.readthedocs.io/en/stable/index.html>
- [9]長谷川真, Definition of beta, 2025,
https://gitlab.com/hasemac/tokamak_equilibrium/-/blob/main/doc/def_of_beta.md
- [10]長谷川真, Definition of parameters, 2025

参考文献

https://gitlab.com/hasemac/tokamak_equilibrium/-/blob/main/doc/def_of_params.md

[11]長谷川真, Definition of inductances, 2025

https://gitlab.com/hasemac/tokamak_equilibrium/-/blob/main/doc/def_of_inductance.md

[12] V. M. Sebastian Raschka, *Python 機械学習プログラミング達人データサイエンティストによる理論と実践*(株式会社インプレス, 2022).

[13]シュリンプクラブ, “RMSE、MAE、MAPE、決定係数：評価指標の違いと選び方”, <https://shrimp-crab.com/rmse-mae-mape-r2/3180/>

謝辞

本研究を行うにあたり、0 から機械学習を学ぶ私にお忙しい中、親身になってご指導していただいた長谷川真准教授に深く感謝を申し上げます。

同じく長谷川研究室である南知樹様、田中雄登様には日ごろからご協力いただき誠に感謝いたします。

また、井戸研究室の井戸毅教授、木下稔基助教、末継寅英様、学会や発表の際など、ご助言をしていただきありがとうございました。

最後に QUEST 関係者の皆様に強く感謝を申し上げます。