

深層学習を応用した 2 溝付き真円軸受の軸受最高温度に関する簡易設計用モデル式
Model Expression for Simple Design of Maximum Bearing Temperature of Cylindrical Journal
Bearing with Two Axial Oil Grooves Applying Deep Learning

九工大・情工（正）*畠中 清史 九工大（院）内田 溪太郎

Kiyoshi Hatakenaka*, Keitaro Uchida*

*Kyushu Institute of Technology

1. はじめに

近年、大型化、高速化が進む産業用の回転機械を支えるジャーナル軸受の潤滑面には、軟質金属のホワイトメタルが広く使われている。ホワイトメタルは低融点材料であるため高速回転時に流体潤滑油膜に接する軸受の最高温度が高くなると最小油膜厚さ位置の付近で軸受面が損傷する危険度が高まる。高速回転機械用のジャーナル軸受を設計する際には、運転中の軸受最高温度が許容値（上限値）を下回ることを確認する必要がある。

著者の一人は、軸受最高温度の簡便な予測式を提案している¹⁾。これは、熱流体潤滑(Thermohydrodynamic lubrication, 以下 THL) モデルにより予測した軸受性能（以下、THL 性能）をもとに構築したデータベース（以下、THL データベース）値から導出した近似多項式（以下、モデル多項式）である。モデル多項式の予測精度は実用上許容可能な範囲に収まる。軸受最高温度を簡便かつ迅速に算出できるので、この式を軸受設計に応用すると、軸受最高温度に関わる制約を満たす設計解を短時間で見出せることになる。これにより軸受の THL 設計に要する時間の大幅な短縮が期待できる。

モデル多項式¹⁾は、1つの式で表されるものの、その対象が反負荷側（鉛直上方）に給油溝をもつ真円軸受の軸受最高温度の場合、70 もの項からなる式になり、使い勝手がやや悪い。導出の際には、重回帰分析を応用している。これに対し、THL データベース値の近似には深層学習（Deep learning, 以下 DL）の応用も可能である。この場合の回帰式（以下、DL モデル式）は、単純な形の関数を多重の入れ子構造にした合成関数として表現される。THL データベース値の予測は複数の式を経て行うことになる。

本報では、軸受最高温度を簡便に予測する DL モデル式を導出すること、また、これによる予測値と検証用データとの比較を通してその評価を行うことを目的とする。なお、対象は前報²⁾で THL データベースの構築を終えた 2 溝付き真円軸受（Fig. 1）とする。

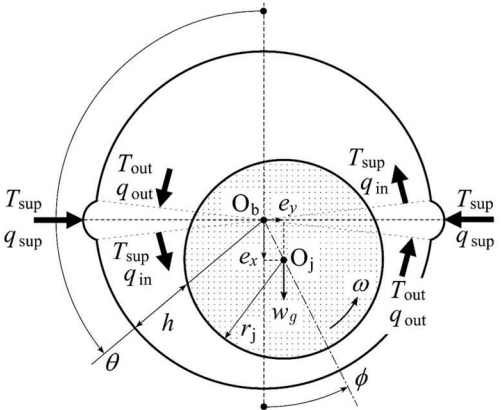


Fig. 1 Cylindrical journal bearing with two-axial oil grooves

2. DL モデル式

2 溝付き真円軸受に関する THL データベース²⁾は、THL 性能（軸受最高温度、最小油膜厚さ、最大油膜圧力、損失、側方漏れ流量などの無次元量）をもとに構築されている。これらの THL 性能は、簡易 THL モデルの 1 つである ALE-ISOADI THL モデル³⁾に 4 つの無次元軸

Table 1 Values of S , Λ , Pe and Kh for THL database	
S	0.5, 0.75, 1, 1.25, 1.5, 2, 2.5, 3, 4, 5, 6, 8, 10, 12, 16, 20
Λ	0.25, 0.5, 0.75, 1.0, 1.25, 1.5, 1.75, 2.0
Pe	5, 10, 20, 40, 80, 160
Kh	0.005, 0.01, 0.02, 0.04, 0.08, 0.16

受設計変数（ゾンマーフェルト数 S 、軸受幅径比 Λ 、ペクレ数 Pe 、コンサリ数 Kh ）を入力値として、Table 1 に示すこれらの組合せ N_d （ $=16 \times 8 \times 6 \times 6 = 4608$ ）ケースに対して算出されている。

本章では、これら 4 つの軸受設計変数（無次元量）を説明変数、軸受最高温度（無次元量）を目的変数とし、Fig. 2 に示す DL モデルを用いて説明変数から目的変数の予測値を算出するための DL モデル式を導出する。

2.1 DL モデルの構成

DL モデルは、THL データベース、前処理層、隠れ層、出力層、後処理層からなる多層パーセプトロンにより構成する。前処理層には、説明変数と同じ 4 つのノードを配置する。前処理層と出力層の間には L_h 層の隠れ層を配置し、各々の隠れ層には I_h 個のニューロンをノードとして配置する。出力層と後処理層は各々を 1 つのノードとみなす。隣接する層のノード同士は全結合させる。

THL データベースから後処理層に向けて計算を進める順伝播の過程では、THL データベースが上流側、後処理層が下流側になる。前処理層の各々のノードは、THL データベース値を入力値とし、これを前処理関数により変換した値を出力値とする。隠れ層の各々のノードは、上流側の層に配置されたすべてのノードの出力値を入力値とし、その重

み付き和を活性化関数により変換した値を出力値とする．出力層は，最終段の隠れ層に配置されたすべてのノードの出力値を入力値とし，その重み付き和を目的関数の予測値として出力する．後処理層は，目的関数の予測値と THL データベース値をもとに，損失関数や誤差の算出を行う．

これに対し，後処理層から前処理層に向けて計算を進める逆伝播の過程では，後処理層を上流側，前処理層を下流側とし，誤差逆伝播法⁴⁾を適用して，すべての重みとバイアスの勾配を上流側から下流側に向けて順に算出していく．

DL では，THL データベースをもとに順伝播と逆伝播の過程を経て求めた重みとバイアスの勾配を用いてこれらの値を更新するという学習を繰り返し，目的変数の予測値による回帰の適合度を良好にしていく．本報では，学習の良し悪しが THL データベース値の選定順に左右されないように，バッチ処理を採用し，THL データベース値の一部（場合によってはすべて）からなる訓練データのすべてに対する重みとバイアスの勾配を算出してから，重みとバイアスの値を一括して更新することにした．

2.2 DL モデルにおける学習

順伝播の過程において，前処理層では，4 つの説明変数 $x_1^{(0)} (=S)$ ， $x_2^{(0)} (=A)$ ， $x_3^{(0)} (=Pe)$ ， $x_4^{(0)} (=Kh)$ を前処理関数 φ に入力して出力値 $z_i^{(0)}$ ($I_{pre}=4$ として $i=1 \sim I_{pre}$) に変換する．ただし，第 n 組目 ($n=1 \sim N_t$) の THL データベース値を用いる場合， $z_{in}^{(0)} = \varphi(x_{in}^{(0)})$ ($i=1 \sim I_{pre}$) である．第 l 段 ($l=1 \sim L_h$) の隠れ層では，Fig. 3 に示す計算グラフ⁴⁾における右向きの直線（エッジ）の上に記したように，前処理層の出力値 $z_j^{(l-1)}$ ($j=1 \sim I_h$) もしくは上流側の隠れ層の出力値 $z_j^{(l-1)}$ ($j=1 \sim I_h$) と重み $w_{ij}^{(l)}$ ，バイアス $b_i^{(l)}$ から重み付き和 $x_i^{(l)}$ ($i=1 \sim I_h$) を算出し，これを活性化関数 ψ に入力して出力値 $z_i^{(l)}$ ($i=1 \sim I_h$) に変換する．この計算は，第 n 組目 ($n=1 \sim N_t$) の THL データベース値から出発した場合，

$$x_{in}^{(l)} = \sum_{j=1}^{I_h} w_{ij}^{(l)} z_{jn}^{(l-1)} + b_i^{(l)} \quad (1.a)$$

$$z_{in}^{(l)} = \psi(x_{in}^{(l)}) \quad (1.b)$$

により表される．出力層では，最終段 ($l=L_h$) の隠れ層の出力値 $z_j^{(L_h)}$ ($j=1 \sim I_h$) と重み $w_{1j}^{(L_h+1)}$ ($j=1 \sim I_h$)，バイアス $b_1^{(L_h+1)}$ から重み付き和 $z_1^{(L_h+1)}$ を算出し，これを出力値とする．この計算は，第 n 組目 ($n=1 \sim N_t$) の THL データベース値から出発した場合，

$$z_{1n}^{(L_h+1)} = \sum_{j=1}^{I_h} w_{1j}^{(L_h+1)} z_{jn}^{(L_h)} + b_1^{(L_h+1)} \quad (2)$$

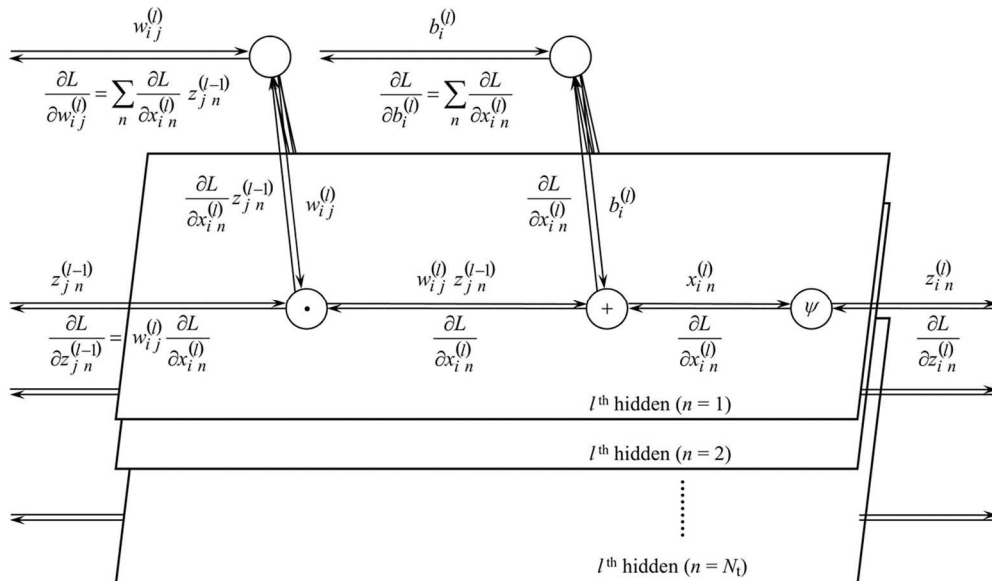


Fig. 3 Computational graph concerning forward and backward propagations in the hidden layer

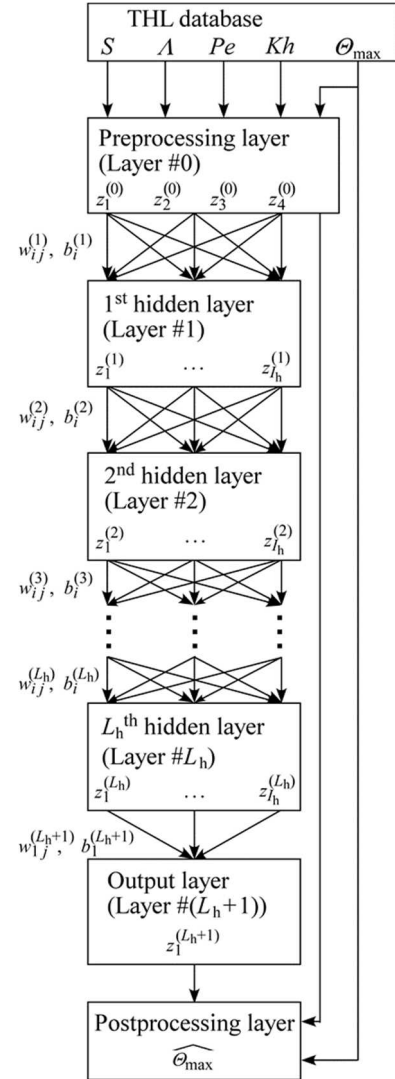


Fig. 2 Calculation model for predicting maximum bearing temperature from four bearing design variables – S , A , Pe and Kh – by applying deep learning

により表される．後処理層では，軸受最高温度（無次元量）の THL データベース値 θ_{\max} を前処理関数 ϕ により変換した $\phi(\theta_{\max})$ と出力層の出力値 $z_1^{(L_h+1)}$ から損失関数 L を

$$L = \frac{1}{2} \sum_{n=1}^{N_t} \left[z_1^{(L_h+1)} - \phi(\theta_{\max n}) \right]^2 \quad (3)$$

により算出するとともに， $z_1^{(L_h+1)}$ を前処理関数の逆関数 ϕ^{-1} により変換した $\phi^{-1}(z_1^{(L_h+1)})$ を目的変数の予測値 $\widehat{\theta_{\max}}$ として，これと THL データベース値 θ_{\max} から算出した相対誤差の絶対値の最大値（以下，最大相対誤差） $\delta_{\theta_{\max \max}}$ を

$$\delta_{\theta_{\max \max}} = \max_n \left| \frac{\widehat{\theta_{\max n}} - \theta_{\max n}}{\theta_{\max n}} \right| \quad (4)$$

により求める．学習は，最大相対誤差 $\delta_{\theta_{\max \max}}$ が判定基準値 $\varepsilon_{\theta_{\max}}$ を下回る，もしくは学習回数が規定値 $M_{\theta_{\max}}$ に達した時点で終了する．

学習を継続する場合は，逆伝播の過程に進む．後処理層では損失関数 L の勾配 $\partial L / \partial L = 1$ を出力する．出力層では，後処理層の出力 1 を入力値とし，これと $z_1^{(L_h+1)}$ に関する損失関数 L の偏導関数 $\partial L / \partial z_1^{(L_h+1)} = z_1^{(L_h+1)} - \phi(\theta_{\max n})$ とを乗じ，これを勾配として出力する．第 l 段 ($l = L_h \sim 1$) の隠れ層では，Fig. 3 に示す計算グラフ 4)における左向きの直線（エッジ）の下に記したように，出力層の出力値 $\partial L / \partial z_1^{(L_h+1)}$ もしくは上流側の隠れ層の出力値 $\partial L / \partial x_i^{(l)}$ ($i = 1 \sim I_h$) を入力値とし，これを活性化関数の導関数 $\partial \psi / \partial x$ に入力して $\partial L / \partial x_i^{(l)}$ ($i = 1 \sim I_h$) に変換した後，これに重み $w_{ij}^{(l)}$ を乗じた $w_{ij}^{(l)} (\partial L / \partial x_i^{(l)})$ を勾配として出力する．重みとバイアスの勾配は，それぞれ，

$$\frac{\partial L}{\partial w_{ij}^{(l)}} = \sum_n \frac{\partial L}{\partial x_i^{(l)}} z_{jn}^{(l-1)} \quad (i = 1 \sim I_h, j = 1 \sim I) \quad (5.a)$$

$$\frac{\partial L}{\partial b_i^{(l)}} = \sum_n \frac{\partial L}{\partial x_i^{(l)}} \quad (i = 1 \sim I) \quad (5.b)$$

により算出する．

本報では，重みとバイアスの勾配を用いて重み $w_{ij}^{(l)}$ とバイアス $b_i^{(l)}$ の更新を行う際に，Adam⁵⁾を適用する．すなわち， $i = 1 \sim I_h$ ， $j = 1 \sim I$ ， $l = 1 \sim L_h$ のそれぞれに対して，0 から 1 の範囲にあるランダムな値を与えた初期値 $w_{ij}^{(l)(0)}$ ， $b_i^{(l)(0)}$ から出発し，第 m 回目の更新を

$$w_{ij}^{(l)(m)} = w_{ij}^{(l)(m-1)} - \eta \frac{\sqrt{1 - \beta_2^m}}{1 - \beta_1^m} \frac{m_{w_{ij}^{(l)}}^{(m)}}{\sqrt{v_{w_{ij}^{(l)}}^{(m)}}} \quad (6.a)$$

$$b_i^{(l)(m)} = b_i^{(l)(m-1)} - \eta \frac{\sqrt{1 - \beta_2^m}}{1 - \beta_1^m} \frac{m_{b_i^{(l)}}^{(m)}}{\sqrt{v_{b_i^{(l)}}^{(m)}}} \quad (6.b)$$

により行う．式(6)による更新時に速度の役割を果たす変数 $m_{w_{ij}^{(l)}}^{(m)}$ ， $m_{b_i^{(l)}}^{(m)}$ の初期値は $m_{w_{ij}^{(l)}}^{(0)} = m_{b_i^{(l)}}^{(0)} = 0$ ，勾配の 2 乗の累積に相当する変数 $v_{w_{ij}^{(l)}}^{(m)}$ ， $v_{b_i^{(l)}}^{(m)}$ の初期値は $v_{w_{ij}^{(l)}}^{(0)} = v_{b_i^{(l)}}^{(0)} = 0$ とし，減衰率 β_1 ， β_2 および学習率 η にはそれぞれ 0.9，0.999，0.001 を与えた．

2.3 ハイパーパラメータの調査範囲

上記 DL モデルにおいて，前処理関数 ϕ と φ ，訓練データ数 N_t ，活性化関数 ψ ，隠れ層の数 L_h ，各隠れ層に配置するノードの数 I_h はハイパーパラメータである．ただし，これらを設定する際の明確な指針は存在しないので，まずは，これらの調査範囲について検討する．

なお，前処理関数 ϕ については，文献¹⁾において，恒等関数と 10 を底とする対数関数（以下，対数関数）のそれぞれを使用した場合の回帰の適合度が調査され，後者の場合に回帰の適合度が高くなることが示されている．Table 1 に示す 4 つの説明変数 S ， A ， Pe ， Kh は，文献¹⁾と同様，各々が範囲が異なっている．このため，本報でも文献¹⁾に倣い，前処理関数 ϕ として対数関数を用い，前処理層の出力値を $z_i^{(0)} = \log_{10} x_i^{(0)}$ ($i = 1 \sim I_{\text{pre}}$) により算出することにした．

軸受最高温度 θ_{\max} の前処理関数 ϕ についても，文献¹⁾において，恒等関数と対数関数のそれぞれで回帰の適合度が調査され，後者の場合に回帰の適合度が高くなることが示されている．しかし，本報では多項式による回帰から合成関数による回帰への変更を行ったので，各々を採用した場合の適合度について調べることにした．

次は，訓練データ数 N_t について検討する．THL データベースは，軸受最高温度 θ_{\max} が小さな値となるケースを複数含んでいる．一般には，その値が小さい場合は THL モデルの解析対象にはならない．本報では，THL データベースの中から N_t 組の訓練データを任意に選定するのではなく，軸受最高温度 θ_{\max} がその下限 θ_{ll} を上回るケースのすべてを訓練データとして採用し，DL モデルの学習を進めることにする．訓練データ数 N_t は， $\theta_{\text{ll}} = 0.0$ の場合は $N_t = 4\,608$ ($= N_d$)， $\theta_{\text{ll}} = 0.1$ の場合は $N_t = 4459$ ， $\theta_{\text{ll}} = 0.2$ の場合は $N_t = 3532$ となる．

次に，活性化関数 ψ について検討する．DL では，隠れ層を多層化する利点を生かすために，ReLU 関数やシグモイド関数など，さまざまな非線形関数が活性化関数として提案されている⁶⁾．本報では，ReLU 関数とその派生（合わせ

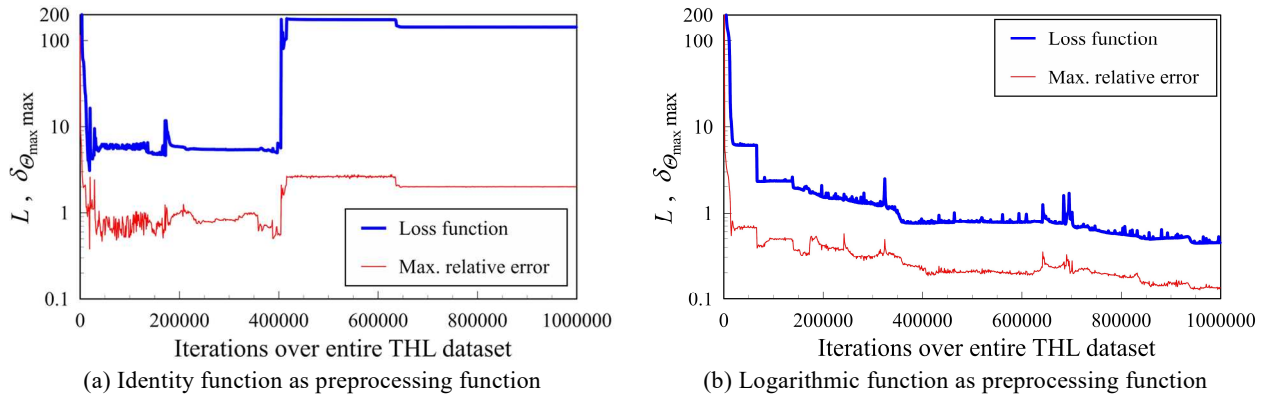


Fig. 4 An example of variation of loss function and maximum absolute value of relative error with iterations of learning over entire THL dataset

て、以下、ReLU 系関数) およびシグモイド関数、双曲線正接関数、逆正接関数 (3 つ合わせて、以下、シグモイド系関数) を活性化関数 ψ の調査対象とする。

最後は、隠れ層の数 L_h 、各隠れ層に配置するノードの数 I_h について検討する。これらが大きい場合、回帰の適合度が高い合成関数が DL モデルとして導出できたとしても、多数の重みとバイアスを含む式になるので、簡易とは言い難い予測式になる。このため、本報では、まず L_h と I_h を 2, 4, 6 のいずれかに絞り込んでから、その前後の値に対する調査を行うことで本 DL モデルの L_h と I_h を決定することにする。

2.4 ハイパーパラメータの決定

2.4.1 軸受最高温度の前処理関数

まずは、軸受最高温度 θ_{\max} の前処理関数 ϕ の影響について調べる。前節で記したように、ランダムに与えた重みとバイアスの初期値から学習を開始するため、計算結果はこれら初期値に大きく依存する。これを踏まえ、5 回の試行を行った。活性化関数 ψ として ReLU 関数を採用し、軸受最高温度の下限を $\theta_{\min} = 0.0$ とした場合の学習回数 m に対する損失関数 L および最大相対誤差 $\delta_{\theta_{\max} \max}$ の変動の一例を Fig. 4 に示す。図中、太線と細線はそれぞれ損失関数 L と最大相対誤差 $\delta_{\theta_{\max} \max}$ を表す。図の(a)は恒等関数を、(b)は対数関数をそれぞれ前処理関数 ϕ に採用した場合の結果である。学習回数 m が少ないうちは、いずれの前処理関数も損失関数 L は低下する傾向を示す。しかし、学習回数 m がある程度多くなると、恒等関数では過学習となり損失関数 L は大幅な増加に転じた後、一定値に落ち着く。一方、対数関数では、わずかながらではあるが、損失関数 L の低下が進む。いずれの前処理関数でも最大相対誤差 $\delta_{\theta_{\max} \max}$ は損失関数 L に連動して変化している。5 回行った各試行で学習回数 m が規定値 $M_{\theta_{\max}} = 1000000$ に達するまでの間に達成した最大相対誤差 $\delta_{\theta_{\max} \max}$ の最小値を Table 2 に示す。最大相対誤差 $\delta_{\theta_{\max} \max}$ の最小値は、恒等関数の場合で 0.351、対数関数の場合で 0.128 であり、回帰の適合度は、多項式回帰 ⁹⁾ の場合と同様、DL モデルでも対数関数の方が高くなりうる事が分かった。これを受け、本 DL モデルでは、前処理関数 ϕ として対数関数を採用することに決めた。

2.4.2 訓練データ数

次は、訓練データ数 N_t の影響について調べる。ただし、前節で記したように、訓練データ数 N_t そのものではなく、軸受最高温度の下限 θ_{\min} を変更する。 θ_{\min} の値を 0.0, 0.1, 0.2 のいずれかに設定して 5 回の試行を行った。最大相対誤差 $\delta_{\theta_{\max} \max}$ を Table 3 に示す。 $\theta_{\min} = 0.0$ に対する結果は Table 2 に示した対数関数の場合と同じである。最大相対誤差 $\delta_{\theta_{\max} \max}$ の最小値は、 $\theta_{\min} = 0.0$ の場合で 0.128、 $\theta_{\min} = 0.1$ の場合で 0.139、 $\theta_{\min} = 0.2$ の場合で 0.125 であり、軸受最高温度の下限 θ_{\min} を 0.2 に設定することで回帰の適合度が高くなりうる事が分かった。このため、本 DL モデルでは、 $\theta_{\min} = 0.2$ として DL モデルの学習を進めることにする。

2.4.3 活性化関数

次は、活性化関数 ψ の影響について検討する。本報では、9 つの活性化関数 ψ (ReLU 系関数 6 種とシグモイド系関数 3 種) に対して 5 回ずつの試行を行った。ReLU 系関数とシグモイド系関数を採用した場合の最大相対誤差 $\delta_{\theta_{\max} \max}$ をそれぞれ Table 4 と Table 5 に示す。Table 4 の ReLU 関数に対する結果は Table 2 に示した対数関数の場合と同じであ

Table 2 Dependence of maximum absolute value of relative error on preprocessing function

Type of pre-processing function ϕ	L_h	I_h	$\delta_{\theta_{\max} \max}$
Identity function	2	2	0.448 ~ 1.017
		4	0.398 ~ 0.853
		6	0.351 ~ 3.926
Logarithmic function	4	2	0.432 ~ 4.425
		4	0.268 ~ 3.096
		6	0.128 ~ 5.826

Table 3 Dependence of maximum absolute value of relative error on lower limit of maximum bearing temperature

θ_{\min}	L_h	I_h	$\delta_{\theta_{\max} \max}$
0.0	4	6	0.128 ~ 5.826
0.1	6	6	0.139 ~ 1.978
0.2	6	6	0.125 ~ 2.598

る。ReLU 系関数では ELU, Swish, Mish, tanhExp の 4 種が、シグモイド系関数は 3 種いずれもが 0.1 (モデル多項式¹⁾による誤差の目安) を下回る最大相対誤差 $\delta_{\theta_{\max \max}}$ となりうる。DL モデル式の煩雑化を避けるためには、Mish, tanhExp, シグモイド, atan のように、少ない隠れ層でその目安を達成できる関数が望ましい。ただし、本報のように前処理関数 ϕ として対数関数を用いると、 Kh に対する前処理層の出力値 $z_4^{(0)} = \log_{10} x_4^{(0)}$ は負の値であり、また、1 を下回る S , A に対する出力値 $z_1^{(0)} = \log_{10} x_1^{(0)}$, $z_2^{(0)} = \log_{10} x_2^{(0)}$ も負の値になる。このため、重みとバイアスの初期値次第では、活性化関数の入力値が負となるので、ReLU 系関数を採用した場合は、その出力値のほとんどが 0 となり、逆伝播の過程で重みとバイアスの勾配が極端に小さくなるために学習が進みにくくなる。本報では、THL データベース値に加え、検証用データへの適合度も高い DL モデル式の導出を目標としている。最大相対誤差 $\delta_{\theta_{\max \max}}$ の目安 0.1 を下回る値が頻出する活性化関数を採用すれば、少ない試行回数で目標を達成できる可能性が高まる。以上を踏まえ、本報の活性化関数はシグモイド系関数の atan がもっとも相応しいとの判断を下した。

2.4.4 隠れ層の数とノード数

Table 5 によると、活性化関数 atan の場合、 $L_h = 4$, $I_h = 6$ のときに最大相対誤差 $\delta_{\theta_{\max \max}}$ の最小が 7% を下回っている。試行 5 回の平均は 9.9% であった。これは L_h と I_h が偶数の場合の結果である。これらの前後の奇数の場合についても調べ、最大相対誤差 $\delta_{\theta_{\max \max}}$ の最小がさらに低減することが確認できれば、設計向きの簡素な (重みとバイアスの合計数が少ない) DL モデル式を最終的に得ることができるともかもしれない。そこで次に、 $L_h = 3 \sim 5$, $I_h = 5 \sim 7$ の範囲で最大相対誤差 $\delta_{\theta_{\max \max}}$ を調べた。その結果を Table 6 に示す。最大相対誤差 $\delta_{\theta_{\max \max}}$ は $L_h = 5$, $I_h = 7$ のときに最小になっている。しかし、この場合、重みとバイアスの合計数は 267 となり、 $\delta_{\theta_{\max \max}}$ の最小が同程度である $L_h = 4$, $I_h = 6$ の場合の合計数 163 に比べて 100 以上も多くなる。以上を踏まえ、隠れ層の数を $L_h = 4$ 、各隠れ層あたりのノード数を $I_h = 6$ として DL モデル式を導出することにした。

2.5 DL モデル式の導出

前節に記したハイパーパラメータのもとで DL モデル式を導出した後、Table 7 に示す 12 ケースとの比較を行い検証用データへの適合度を調べる。これら 12 ケースは、文献³⁾で ALE-ISOADI THL モデルの検証時に用いたケースである。なお、学習回数 m の規定値は $M_{\theta_{\max}} = 1\,000\,000$ 、最大相対誤差 $\delta_{\theta_{\max \max}}$ の判定基準値 $\varepsilon_{\theta_{\max}}$ は 0.09 ($\delta_{\theta_{\max \max}}$ の目安 0.1 に対して 10% の裕度をもたせた値) とした。50 回の試行を行った結果、15 回で $\delta_{\theta_{\max \max}}$ が $\varepsilon_{\theta_{\max}}$ を下回った。しかし、そのいずれでも検証用データの最大相対誤差 $\delta_{\theta_{\max \max}}$ が 0.09 を下回ることにはなかった。

これを踏まえ、 $M_{\theta_{\max}}$ に至るまでの間で、最大相対誤差 $\delta_{\theta_{\max \max}}$ が判定基準値 $\varepsilon_{\theta_{\max}}$ を下回った場合も学習を継続させ、学習回数 m と最大相対誤差 $\delta_{\theta_{\max \max}}$ との関係を訓練データと検証用データのそれぞれで調べることにした。その結果を Fig. 5 に示す。図中、太線と細線はそれぞれ訓練データと検証用データの最大相対誤差 $\delta_{\theta_{\max \max}}$ であり、破

Table 4 Dependence of maximum absolute value of relative error on activation function of ReLU family

Type of activation function ϕ	L_h	I_h	$\delta_{\theta_{\max \max}}$
ReLU	4	6	0.128 ~ 5.826
leaky-ReLU	4	6	0.102 ~ 0.206
ELU	6	6	0.098 9 ~ 0.888
Swish	6	6	0.086 4 ~ 0.205
Mish	4	6	0.084 8 ~ 0.143
tanhExp	4	6	0.075 4 ~ 0.173

Table 5 Dependence of maximum absolute value of relative error on activation function of sigmoid family

Type of activation function ϕ	L_h	I_h	$\delta_{\theta_{\max \max}}$
sigmoid	2	6	0.083 2 ~ 0.198
tanh	6	6	0.074 5 ~ 2.568
atan	4	6	0.068 7 ~ 0.119

Table 6 Dependence of maximum absolute value of relative error on number of hidden layers and number of nodes per hidden layer

L_h	I_h	$\delta_{\theta_{\max \max}}$
3	7	0.080 5 ~ 0.126
4	6	0.068 7 ~ 0.119
5	7	0.065 0 ~ 0.096 9

Table 7 Twelve cases of combination of S , L , Pe and Kh used for verification of simplified THL model in Ref³⁾

Case No.	S	A	Pe	Kh	Remarks
1	3.36	0.8	41.0	0.0781	Brito et al. ⁷⁾
2	16.8			0.0565	
3	2.43		0.0141		
4	0.607			0.0538	
5	1.84	0.55	42.7		0.116
6	3.97			0.0870	
7	7.15		0.0233		
8	2.98			0.00582	
9	0.396	0.5	234		0.0233
10	0.0989			58.6	
11	0.791		234		
12	5.22				

線は判定基準値 $\varepsilon_{\theta_{\max}}$ ($=0.09$) を表す。訓練データ、検証用データともに最大相対誤差 $\delta_{\theta_{\max \max}}$ が増減を繰り返す様相を見て取れる。この試行回の場合、訓練データと検証用データの最大相対誤差 $\delta_{\theta_{\max \max}}$ はそれぞれ所々で判定基準値 $\varepsilon_{\theta_{\max}}$ を下回っているが、双方の $\delta_{\theta_{\max \max}}$ が同時に $\varepsilon_{\theta_{\max}}$ を下回ることではない。この結果は、訓練データの $\delta_{\theta_{\max \max}}$ が $\varepsilon_{\theta_{\max}}$ を下回った時点で学習を終了させる方針のもとでは検証用データへの適合度の高い DL モデル式が得られる見込みが非常に低いことを示唆している。

そこで、汎化性能が高い DL モデル式を学習終了の時点で確実に得るために、訓練データによる学習と並行して検証用データへの適合度調査を行い、両データの $\delta_{\theta_{\max \max}}$ が同時に $\varepsilon_{\theta_{\max}}$ を下回るまで学習を継続することにした。この方針のもと、 $M_{\theta_{\max}}=1\,000\,000$ 、 $\varepsilon_{\theta_{\max}}=0.09$ として 50 回の試行を行った結果、19 回で訓練データの最大相対誤差 $\delta_{\theta_{\max \max}}$ が判定基準値 $\varepsilon_{\theta_{\max}}$ を下回り、そのうち 8 回で訓練データと検証用データの $\delta_{\theta_{\max \max}}$ が同時に $\varepsilon_{\theta_{\max}}$ を下回った。本報では、これら 8 回の中で検証用データの $\delta_{\theta_{\max \max}}$ が最小値 ($=0.0488$) となった重みとバイアスの組合せが DL モデル式にもっとも相応しいとの判断を下した。

2.6 DL モデル式による予測手順

前節までで DL モデル式のハイパーパラメータおよび重みとバイアスのすべてが確定した。これらのもとで 4 つの説明変数 S 、 A 、 Pe 、 Kh から軸受最高温度の予測値 $\widehat{\theta_{\max}}$ を算出するまでの手順を整理すると、次の通りとなる。[1] 説明変数 S 、 A 、 Pe 、 Kh の値を前処理層に入力する。[2] 対数関数によりこれらの変換を行った $z_1^{(0)}=\log_{10} S$ 、 $z_2^{(0)}=\log_{10} A$ 、 $z_3^{(0)}=\log_{10} Pe$ 、 $z_4^{(0)}=\log_{10} Kh$ を前処理層の出力とする。[3] 第 l 段 ($l=1\sim 4$) の隠れ層における i 番目 ($i=1\sim 6$) のノード出力を

$$z_i^{(l)} = \tan^{-1} \left(\sum_{j=1}^I w_{ij}^{(l)} z_j^{(l-1)} + b_i^{(l)} \right) \quad (7)$$

により算出する。[4] 出力層の出力を

$$z_1^{(5)} = \sum_{j=1}^6 w_{1j}^{(5)} z_j^{(4)} + b_1^{(5)} \quad (8)$$

により計算する。[5] 後処理層では、10 を底とする指数関数による変換を行い、軸受最高温度の予測値 $\widehat{\theta_{\max}}$ を

$$\widehat{\theta_{\max}} = 10^{z_1^{(5)}} \quad (9)$$

により求める。

3. DL モデル式の評価

本 DL モデル式は、訓練データと検証用データの最大相対誤差が同時に判定基準値を下回った時点で学習を終了している。データ数 $N_t=3\,532$ の訓練データに対する最大相対誤差は $0.089\,9$ (<0.09) であった。全 12 ケースの検証用データに対する予測値 $\widehat{\theta_{\max}}$ および THL データベース値 θ_{\max} との相対誤差は Table 8 に示す通りとなった。最大相対誤差がケース 9 の 0.0488 (<0.09) であることから、検証用データへの適合度は訓練データに対するよりも高い。特に、Table 1 に示した軸受設計変数の範囲を逸脱している 4 つのケース (ケース 9～ケース 12) でも他の 8 ケースと遜色ない適合度になっている。軸受最高温度の THL データベース値は軸受設計変数の変化に対して急変しないことが確認されている²⁾。また、その THL 予測値は、軸受設計変数がデータベースの範囲から外れても、変化の傾向を大幅に変えることはない。検証用データへの適合度の高さは、このような性質がある目的変数の外挿による予測にも本 DL モデル式が有効であることを示している。これは、本 DL モデル式は合成関数による回帰式であり、文献¹⁾のような多項式による回帰で生じがちなデータ範囲外での大きな変動が起こりにくい性質をもつため

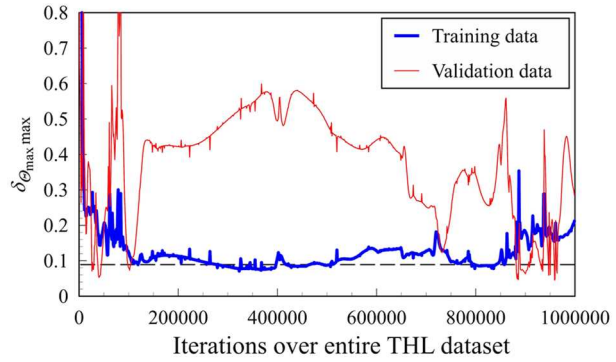


Fig. 5 An example of relationship between iterations of learning over entire THL dataset and minimum of maximum absolute value of relative error

Table 8 Predicted maximum bearing temperature by DL model expression for validation data and its relative error to the THL database value

Case No.	θ_{\max}	$\widehat{\theta_{\max}}$	Relative error
1	1.151	1.138	-0.0108
2	0.890	0.899	0.0092
3	0.995	0.983	-0.0119
4	0.337	0.327	-0.0292
5	1.088	1.082	-0.0051
6	1.538	1.507	-0.0198
7	1.385	1.376	-0.0063
8	1.247	1.237	-0.0080
9	1.622	1.543	-0.0488
10	0.396	0.390	-0.0165
11	1.019	1.057	0.0377
12	0.459	0.439	-0.0447

ある。

次は、データ総数 $N_d = 4\,608$ の THL データベース値のうち訓練データに使用しなかった 1 076 ケースに対する予測精度について調べる。これらは軸受最高温度が下限値 $\Theta_{ll} = 0.2$ を下回るケースであるため、相対誤差の代わりに絶対誤差で評価を行うことにする。その大きさは最大で 0.046 6 であったことから、訓練データではない THL データベース値への適合度も高いことが確認された。

また、本 DL モデルによる軸受最高温度の計算は、Microsoft Excel のような表計算用の応用ソフトウェアを使用すると、簡便にできることが分かった。

以上より、本 DL モデル式の汎化性能は十分に高く、2 溝付き真円軸受の軸受最高温度を短時間で簡便に予測できる。軸受設計の過程で軸受諸元を繰り返し変更しても、軸受最高温度の見積もりに時間を費やす必要はほとんどなくなる。この点で本 DL モデル式は有用性も高いといえる。

4. おわりに

本報では、2 溝付き真円軸受の熱流体潤滑性能に関するデータベースをもとに、深層学習を応用して、この軸受の最高温度を予測する DL モデル式を導出した。このモデル式には 4 つの軸受設計変数（ゾンマーフェルト数 S 、軸受幅径比 A 、ペクレ数 Pe 、コンサリ数 Kh ）を入力値として与える。

導出時に適用した深層学習モデルは、前処理層、隠れ層、出力層、後処理層からなる多層パーセプトロンにより構成され、隣接する層のノード同士は全結合させている。軸受設計変数と軸受最高温度に対する前処理関数には対数関数を、活性化関数には逆正接関数をそれぞれ採用し、4 層の隠れ層の各々に 6 つずつのノードを配置した。訓練データには軸受最高温度が 0.2 以上の THL データベース値を用いた。学習は、訓練データおよび検証用データに対する最大相対誤差が同時に 9% を下回った時点で終了した。

本 DL モデル式は汎化性能が高く、検証用データとの最大相対誤差は 5% 以下 (0.048 8) であり、訓練データとして採用しなかった THL データベース値との絶対誤差の大きさは最大で 0.046 6 であった。

本 DL モデル式は、2 溝付き真円軸受の軸受最高温度を短時間で簡便に予測することができる。以上より、この式は 2 溝付き真円軸受の THL 設計に要する時間を大幅に短縮するための有用なモデル式であると結論づける。

謝辞

本研究は JSPS 科研費 23K03645 の助成を受けて行われた。

文献

- 1) 畠中・本田：簡易熱流体潤滑モデルにもとづく真円軸受の新しい設計法，トライボロジスト，66，7 (2021) 548.
- 2) 畠中・内田：2 溝付き真円軸受の設計への応用を見据えた熱流体潤滑性能データベース，トライボロジスト，69，5 (2024) .
- 3) 畠中・内田：2 溝付き真円軸受の設計への応用を見据えた簡易熱流体潤滑モデル，機論，88，915 (2022) 22-00234.
- 4) 斎藤：ゼロから作る Deep Learning —Python で学ぶディープラーニングの理論と実装— 初版 (5 章)，ティム・オライリー (2016) 123.
- 5) D. P. Kingma & J. L. Ba: Adam: A Method for Stochastic Optimization, The 3rd International Conference on Learning Representations (ICLR), San Diego (2014) DOI:10.48550/arXiv.1412.6980.
- 6) 文献 4 の 3 章，ティム・オライリー (2016) 39.
- 7) F. P. Brito, A. S. Miranda, J. Bouyer & M. Fillon: Experimental Investigation of the Influence of Supply Temperature and Supply Pressure on the Performance of a Two-Axial Groove Hydrodynamic Journal Bearing, ASME J. Tribology, 129, 1 (2007) 98.
- 8) J. Tonnesen & P. K. Hansen: Some Experiments on the Steady State Characteristics of a Cylindrical Fluid-Film Bearing Considering Thermal Effects, ASME J. Tribology, 103, 1 (1981) 107.
- 9) M. K. Fitzgerald & P. B. Neal: Temperature Distributions and Heat Transfer in Journal Bearings, ASME J. Tribology, 114, 1 (1992) 122.