深層混合処理地盤のFEM解析結果を用いた機械学習によるサロゲートモデルの構築

片山　遥平[1]・笠間　清伸[2]

[1]正会員　五洋建設株式会社　技術研究所（〒329-2746 栃木県那須塩原市四区町1534-1）
E-mail: youhei.katayama@mail.penta-ocean.co.jp (Corresponding Author)
[2]正会員　九州大学大学院教授　工学研究院社会基盤部門（〒819-0395 福岡県福岡市西区元岡744）
E-mail: kasama@civil.kyushu-u.ac.jp

深層混合処理工法などの地盤改良工法を港湾構造物の基礎地盤や背後地盤に適用する場合，本体工の要求性能に応じてFEM解析によるレベル2地震動に対する変形量の照査が必要となる．レベル1信頼性設計法では改良地盤が有する空間的ばらつきを直接的に考慮できないため，改良体強度に係数を乗じて大きく低減することで安全性を確保している．一方，レベル3信頼性設計法ではこの空間的ばらつきを直接的に考慮することができ，構造物の安定性と経済性の両面から合理的な設計が可能となる．しかし本設計法ではモンテカルロ・シミュレーションが必須で，FEM解析に適用する場合は解析コストが非常に大きくなる．本研究では解析コストを低減することを目的として，機械学習を用いて地震後の岸壁の残留水平変位を推定するサロゲートモデルの構築について検討した．

Key Words: *level 3 reliability-based design, spatial variability, machine learning, surrogate model*

1. 研究の背景

港湾の施設の技術上の基準・同解説は，2007年の改定で仕様規定から性能規定へと移行し，信頼性設計法が照査手法として使用可能となった[1]．この経緯は北詰らの報告が詳しい[2]．実務では部分係数法とも呼ばれるレベル1信頼性設計法[3]が広く使用されている．これは，従来は安全率という唯一の係数で安全性の余裕を確保していたものを，部分係数の導入により，材料や荷重が本来持っている物理的な不確実性や統計的推定誤差を考慮することができるというものである．

深層混合処理工法をはじめとする地盤改良工法を港湾構造物の基礎地盤や背後地盤に適用する場合，本体工の要求性能に応じてFEM解析によるレベル2地震動[4]に対する変形量の照査が必要となる．また，各種の地盤改良工で改良された地盤は，その物性に大きな空間的ばらつきを持つことが知られている[5]．これは，原地盤が元来有する土質の不均質性と，施工過程で生じる不均一性に起因する．前述のレベル1信頼性設計法ではこのような空間的ばらつきを直接的に考慮することはできないため，通常は一様な地盤として解析を実施する．この際，改良体強度の特性値に係数を掛けて低減することで安全性を確保している．対して信頼性設計法の中で最高位に属するレベル3信頼性設計法[6]は，モンテカルロ・シミュレーション（以下，MCS）等の確率論的方法に基づいて破壊確率を直接的に算定可能な手法である．空間的ばらつきを直接的に考慮することができ，これにより前述の強度低減が不要となり，構造物の安定性と経済性の両面から合理的な設計が可能となる．レベル3信頼設計法について，あまり多くはないが，実務適用された事例も存在する[7,8]．

ここで，FEM解析にレベル3信頼性設計法を適用する場合，MCSが必須となることから，解析コストは非常に大きなものとなる．筆者らが実施した検討では，改良体要素の強度の空間的ばらつきを変えながら，1ケースにつき100回のFEM解析を35ケース，すなわち合計3,500回実施し，合計で約7,000時間を要した[9]．学術研究が目的であれば，このような長期の検討期間を確保することも可能だが，実務の場合はその限りではない．レベル3信頼性設計法の実務適用があまり進まない要因の一つとして，このように解析コストが非常に大きいということが考えられる．

ここで，サロゲートモデルという手法が存在する．これは解析コストの大きい物理シミュレーションの入力情報・出力結果を学習し，より高速に推定可能なサロゲートモデル（代理モデル）を構築するというものである．

サロゲートモデルを構築する方法として，応答曲面法，モード分解および機械学習などがある．大竹らは，少数の FEM 解析から応答曲面を構築し，説明変数の不確実性情報に基づいて MCS を実施して破壊確率を求める信頼性設計スキームを提案している[10]．住岡らは，特異値分解により FEM 解析の膨大な時空間データから重要なモードのみを抽出するモード分解を用いて代替モデルを構築し，精度を維持しながら港湾構造物の二次元地震応答解析を効率化する手法を検討している[11]．これらの研究はレベル 3 信頼性設計法に基づいているが，地盤の要素一つ一つに異なる物性値を入力した FEM 解析を行っているわけではない．笠間らは，浸透固化処理地盤の強度の空間的ばらつきを考慮した FEM 解析で MCS を実施し，これを機械学習で代替する検討を行っている[12]．

機械学習には様々な手法が存在するが，本研究では教師あり学習[例えば 13]について言及する．教師あり学習の中にも様々な手法が存在し，どの手法においても大量の教師データが必要となるが，前述のとおり 3,500 回分の解析結果が既に存在する[9]．そこで本研究では，大量の解析結果を有効活用できる機械学習によってサロゲートモデルを構築することとした．これにより空間的ばらつきを考慮した FEM 解析を代替し，解析コスト低減を目的に検討を実施した．

2. 使用した FEM 解析データについて

(1) 断面，地盤パラメータ，入力地震動

前述の検討[9]では，深層混合処理工法により改良された地盤を対象に，FEM 解析プログラムの FLIP[14]を用いて地震応答解析を行っている．使用した断面を**図-1**，地盤パラメータを**表-1**，入力地震動を**図-2** に示す．これらは港湾・空港における深層混合処理工法技術マニュアル[15]（以下，CDM マニュアル）の設計事例に示されているデータを一部変更したものである．この断面の改良体要素は 33 行，36 列の合計 1,188 個となっている．

本解析[9]の対象構造物は重力式護岸（ケーソン式）で，レベル 2 地震動作用後のケーソン天端部の残留水平変位

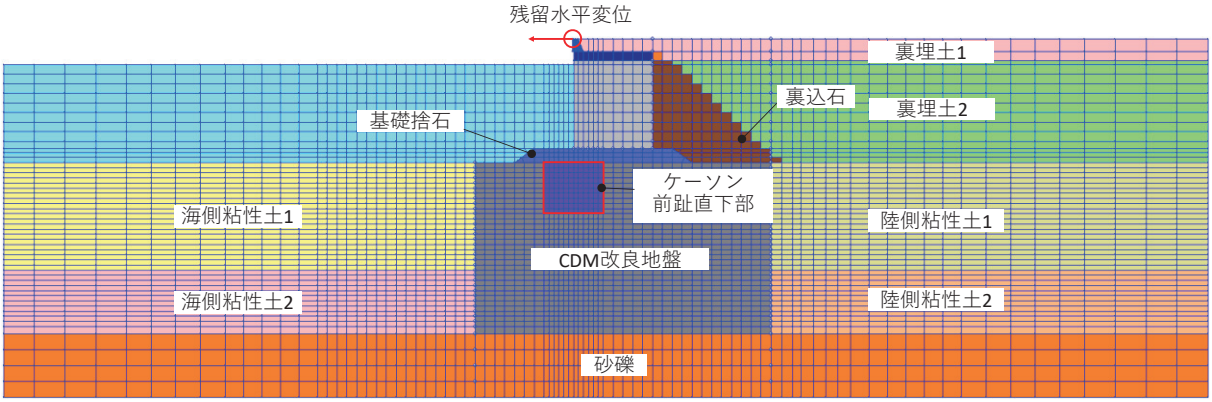


図-1　検討に使用した FLIP 断面[15]を一部変更

表-1　地盤パラメータ[15]を一部変更

土層区分	ρ	ν	σ'_{ma}	G_{ma}	K_{ma}	c	ϕ'_f	n	h_{max}
	t/m^3	-	kN/m^2	kN/m^2	kN/m^2	kN/m^2	°	-	-
海側粘性土1	1.48	0.33	28.2	8,714	22,725	25.63	30.0	0.55	0.20
海側粘性土2	1.73	0.33	82.7	18,700	48,767	55.00	30.0	0.55	0.20
陸側粘性土1	1.48	0.33	193.3	32,868	85,714	96.67	30.0	0.55	0.20
陸側粘性土2	1.73	0.33	247.8	42,123	109,845	123.89	30.0	0.55	0.20
裏埋土1	1.84	0.33	98.0	62,345	162,586	0.00	39.0	0.45	0.24
裏埋土2	2.04	0.33	98.0	62,345	162,586	0.00	39.0	0.45	0.24
砂礫	2.04	0.33	98.0	175,100	456,500	0.00	43.4	0.45	0.24
裏込石	2.04	0.33	98.0	180,000	469,412	20.00	35.0	0.45	0.24
基礎捨石	2.04	0.33	98.0	180,000	469,412	20.00	35.0	0.45	0.24
CDM改良地盤	1.48	0.33	0.0	式(2a)	式(2b)	式(2c)	0.0	0.45	0.24

土層区分	液状化特性値算定に用いる定数			液状化特性値					
	N_{65}	σ'_v	F_c	ϕ_p	s_1	w_1	p_1	p_2	c_1
	-	kN/m^2	%	°	-	-	-	-	-
裏埋土2	6	65	14	28	0.005	5.436	0.500	0.980	2.249

を目的変数としている．これは，緊急物資を積載した船舶を係留するため，地震後でも法線の変形を一定の水準以下に抑える必要があるためである．

(2)　空間的ばらつきの表現方法

空間的ばらつきを表現するため，対数正規分布に基づいて改良体の一軸圧縮強さ（以下，q_u）にばらつきを与えた．q_uの空間的ばらつきの一例を**図-3**に示す．ばらつきの生成に必要となるq_uの平均値，変動係数および自己相関距離の組合せを**表-2**に示す．

平均値について，現場平均強度を 3,340 kN/m^{2}としたとき，変動係数を考慮して低減した数値が 2,238 kN/m^2，さらに係数を乗じて低減した特性値が 1,790 kN/m^{2}である．1,000 kN/m^{2}と 500 kN/m^{2}は切りの良い数値を設定した．

変動係数について，実務で一般に使用される値として CDM マニュアルを参考に 0.33，現場強度のばらつきが大きい場合として事前混合処理工法技術マニュアル [16]を参考に 0.80 を設定した．

自己相関距離について，鉛直方向は Honjo の研究 [17]を参考に 0.4 m，水平方向は Navin et al.の研究 [18]を参考に 3.0 m と設定した．両者とも式(1)に示す指数関数型の自己相関関数により求められている．

$$\rho(\Delta x) = exp\left(-\frac{\Delta x}{\theta}\right) \tag{1}$$

ここに，Δx：2 点間の距離，θ：自己相関距離である．

実際の解析では，FLIP においてq_uと関連する基準初期せん断弾性係数G_{ma}，基準初期体積弾性係数K_{ma}および粘着力cを変動させている．q_uとこれらのパラメータの変換式を式(2a)～(2c)に示す [19]．なお，これらの変換式の変換誤差については考慮していない．

$$G_{ma} = \frac{E}{2(1+\nu)} = \frac{500q_u}{2(1+0.33)} \tag{2a}$$

$$K_{ma} = \frac{2}{3}\frac{1+\nu}{1-2\nu}G_{ma} \tag{2b}$$

$$c = \frac{1}{2}q_u \tag{2c}$$

ここに，E：ヤング係数(kN/m^2)，ν：ポアソン比である．

またq_uをばらつかせる際に，深層混合処理地盤としては考えにくい強度（例えば 1 kN/m^{2}や 40,000 kN/m^{2}など）を取ることがある．このような値が出現しないよう，**表-2**に示すように下限値と上限値を設定し，この条件を満たさない断面は棄却した．

(3)　解析ケース

表-3, 4, 5 に解析ケースを示す．**表-3** の解析ケース 0 は

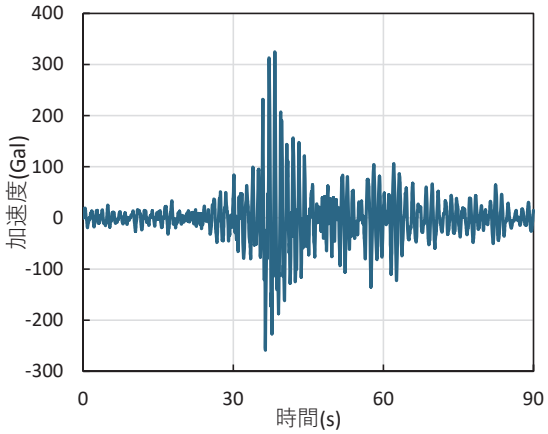


図-2　入力地震動 [15]を一部変更

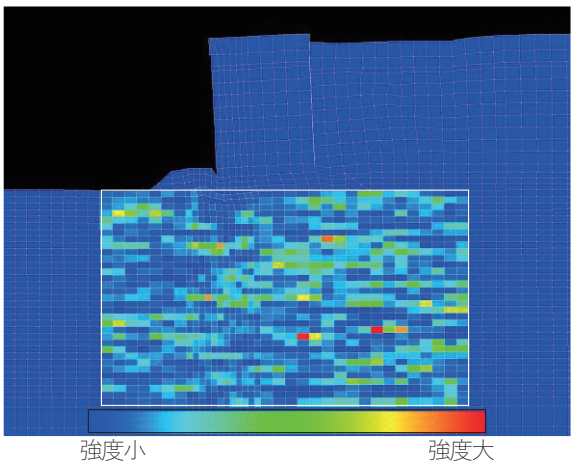


図-3　空間的ばらつきの例

表-2　ばらつき生成に必要なパラメータ

改良体強度			変動係数	自己相関距離	
下限値	平均値	上限値		鉛直	水平
kN/m^2	kN/m^2	kN/m^2	-	m	m
50	500	3,000	0.33	0.4	3.0
50	1,000	6,000			
50	1,790	10,700			
50	2,238	13,400	0.80		
50	3,340	19,900			

表-3　解析ケース 0

ケース	強度
	kN/m^2
0_1	500
0_2	1,000
0_3	1,790
0_4	2,238
0_5	3,340

空間的ばらつきを考慮しない通常の解析，**表-4** の解析ケース 1 は空間的ばらつきを考慮した解析を指す．**表-5** の解析ケース 2 は，解析ケース 1 の解析結果から目的変数への影響が大きい改良範囲をクリティカル部として特定したうえで，クリティカル部とそれ以外の範囲に異なる施工仕様を適用した解析を指す．改良体要素のq_uの空間分布以外の解析条件は 3 ケースとも同じである．

(4)　得られた結果

本検討 [9]で得られた結果の概要は下記のとおりである．
①空間的ばらつきを直接的に考慮した解析（解析ケース 1）を基準とした場合，これと従来の均一地盤における解析（解析ケース 0）の結果が等しくなる強度低減率は，20.0～33.0 %の間にあるという結果であった．実務で一般に使用される強度低減率は 33.0 %であり，これはおおむね適正な数値であることを確認した．また空間的ばらつきを直接的に考慮した解析を行うことで，若干ながら経済的な設計を実現できる可能性があることを確認した．
②解析ケース 1 より，ケーソン前趾直下部をケーソン天端の残留水平変位への影響が大きいクリティカル部と特定した．
③解析ケース 2 より，クリティカル部に大きな平均強度と小さな変動係数，それ以外の範囲に比較的小さな平均強度（セメント添加量減）と大きな変動係数（施工時間減）を設定することで，経済性と安全性を両立する合理的な設計が可能となることを確認した．

3.　機械学習手法について

本研究では，第2章に示したFEM解析結果を用いて，機械学習によるサロゲートモデルの構築について検討する．

(1)　手法の選定

連続値の回帰を行う機械学習手法は説明変数の取り扱いにより様々な手法が存在するが，ここでは説明変数をテーブルデータもしくは画像とする場合について考える．前者で用いられるアルゴリズムには，ランダムフォレスト，ディープニューラルネットワークおよび勾配ブースティング決定木などがある[例えば 20]．後者には，畳み込みニューラルネットワーク（以下，CNN）[21]および Vision Transformer[22]などがある．

表-4, 5 に示した全 35 ケースにおいて，ケース間の解析条件の差は改良体要素のq_uの空間分布のみであるため，これを唯一の説明変数とする．目的変数はケーソン天端部の残留水平変位で，連続値を推定する回帰タスクとなる．

表-4　解析ケース 1

ケース	平均強度	変動係数
	kN/m^2	-
1_1	500	0.33
1_2	500	0.80
1_3	1,000	0.33
1_4	1,000	0.80
1_5	1,790	0.33
1_6	1,790	0.80
1_7	2,238	0.33
1_8	2,238	0.80
1_9	3,340	0.33
1_10	3,340	0.80

表-5　解析ケース 2

ケース	ケーソン前趾直下		その他	
	平均強度	変動係数	平均強度	変動係数
	kN/m^2	-	kN/m^2	-
2_1	500	0.33	500	0.80
2_2	500	0.80	500	0.33
2_3	500	0.33	1,000	0.33
2_4	500	0.33	1,790	0.33
2_5	500	0.33	2,238	0.33
2_6	500	0.33	3,340	0.33
2_7	1,000	0.33	500	0.33
2_8	1,000	0.80	500	0.80
2_9	1,000	0.33	500	0.80
2_10	1,790	0.33	500	0.80
2_11	1,790	0.33	500	0.33
2_12	1,790	0.33	1,000	0.80
2_13	1,790	0.80	500	0.80
2_14	1,790	0.80	1,000	0.80
2_15	1,790	0.33	1,000	0.33
2_16	2,238	0.33	500	0.80
2_17	2,238	0.33	1,000	0.80
2_18	2,238	0.80	500	0.80
2_19	2,238	0.80	1,000	0.80
2_20	2,238	0.33	1,000	0.33
2_21	3,340	0.33	500	0.80
2_22	3,340	0.33	1,000	0.80
2_23	3,340	0.80	500	0.80
2_24	3,340	0.80	1,000	0.80
2_25	3,340	0.33	1,000	0.33

説明変数をテーブルデータとする場合，全 1,188 要素のq_uを 1 行に，すなわち 1,188 次元のデータを作成することになる．ここで第 2 章 4 節で示したように，特定の改良範囲が残留水平変位に大きく影響することがわかっている．この方法では改良体の各要素を独立な変数として扱うため，位置関係（空間的構造）を特徴量に含めることができない．

一方で説明変数を画像とする場合，CNN では，畳み込みを行う際に近傍の値を参照するため，局所的なパターンや連続的な傾向などの空間的構造を特徴量に含めることができる．原理は異なるが，Vision Transformer も同様の特性を持つ．このような理由から，説明変数を画像とする手法を採用した．本研究では，参考となる書籍や文献が多く，Python での実装のハードルが比較的低い CNN を採用した．ライブラリは Tensorflow[23]をバックエンドとした Keras[24]を使用した．

(2) CNN の構造設計

CNN で回帰を実施するにあたり，ネットワークの構造を設計する必要がある．CNN の構造[例えば 25]は大きく 3 つの層に分けられ，それぞれ入力層，中間層および出力層である．中間層はさらに大きく 3 つに分けられ，それぞれ畳み込み層，プーリング層および全結合層である．これらのパーツを適切に組み合わせて，自身のタスクに適したネットワーク構造を設計する必要がある．そのためには検討対象（ここではケーソン式護岸）に関する専門知識に加えて，プログラミングや機械学習に関する知識も不可欠である．筆者も含め，いち土木技術者には非常にハードルが高い作業と言える．

ここで，転移学習[26]およびファインチューニング[27]と呼ばれる手法が存在する．これらは，既存のネットワーク構造（以下，CNN アーキテクチャ）と学習済みの重みを自身のタスクに流用する手法である．重みとは入力値の重要性を数値化したものである．転移学習では学習済みの重みをそのまま全て使用する．ファインチューニングでは一部の層の重みの固定（これを凍結という）を解除（これを解凍という）し，自身のデータセットを用いて重みを再学習する．これらの手法を用いることで，一から CNN の構造を設計する必要がなくなる．さらに，データセットが少なく十分な学習ができない場合でも，学習済みの重みを流用することで，精度よく推定できるモデルを構築できる可能性がある．ただし，詳細は第 5 章で述べるが，両手法とも自身のタスクに応じて出力層を変更する必要がある．

(3) CNN アーキテクチャ

今回の検討で使用した 7 つの CNN アーキテクチャ[28),29),30),31)]を**表-6** に示す．いずれも画像認識モデルとして著名なものである．これらは Keras に含まれており，インポートするだけで簡単に実装することができる．学習には ImageNet[32]という 100 万枚を超える様々な画像を含むデータセットが使用されている．具体的には動物や建築物など，一般的によく見られる物体などを収めた画像である．これらの CNN アーキテクチャは後述の本研究の入力データとは大きくかけ離れたデータで学習されているが，大量のデータセットから「学習の仕方」を習得しており，これを活用することで高精度に推定できることが期待される．これら 7 つの CNN アーキテクチャで転移学習およびファインチューニングを検討する．

4. 説明変数の作成について

本章では，第 2 章に示した全 35 ケース，全 3,500 メッシュ（断面データを指す）の FEM 解析結果を用いた説明変数の作成について検討する．

(1) 入力データのサイズ

表-6 に各 CNN アーキテクチャのデフォルトの入力サイズを示している．EfficientNetV2 の 3 種は他の 4 つとサイズが異なるが，本研究では224×224 ピクセルに統一した．そのため，**図-4** に示すようにq_uの空間的ばらつきを 224×224 ピクセルの配列データとして表現した．これにより，EfficientNetV2 の 3 種は性能が若干低下する可能性があるが，本研究では検証していない．**図-1** に示した断面の改良体要素は 33 行，36 列であり，1 要素 1 ピクセルとすると余白が大きすぎるため，1 要素を 4×4 ピクセル

表-6　CNN アーキテクチャ

名称	ブロック数	デフォルトの入力サイズ
VGG16	5	224×224×3
VGG19	5	224×224×3
ResNet50	5	224×224×3
MobileNetV2	16	224×224×3
EfficientNetV2S	6	384×384×3
EfficientNetV2M	7	480×480×3
EfficientNetV2L	7	480×480×3

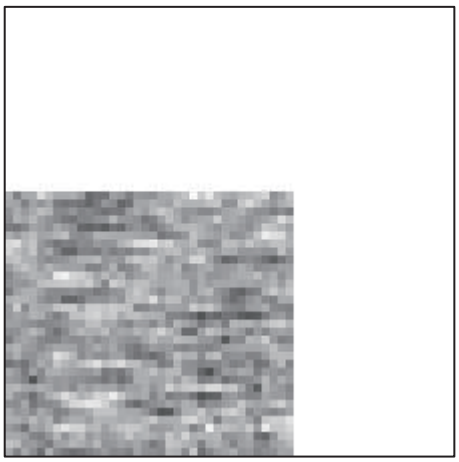


図-4　CNN に入力する配列データのイメージ

に拡張した．なお今後の展望として，構造形式，改良工法，改良範囲および入力地震動などが異なる FEM 解析データを組み合わせ，多入力一出力の汎用的な推定モデルを構築することを目指している．そのため改良範囲の違いを吸収できるよう，余白を残すデータ作成方法を採用した．

(2)　色による q_u の表現

表-4, 5 に示した全 35 ケースにおいて，q_u の最小値および最大値はそれぞれ 50 kN/m^2，19,900 kN/m^2 である．この数値のばらつきを色で表現した．コンピュータ上では色は RGB のピクセル値で表現され，それぞれ赤色，緑色および青色の強さを表す．0-255 の 256 段階で表現され，全てが 0 の場合は黒，全てが 255 の場合は白となる．ばらつきを考慮するパラメータが 2 つ，または 3 つの場合は，RGB のチャンネルにそれぞれのパラメータをあてはめることが考えられる．今回は q_u のみであるが，転移学習およびファインチューニングを採用する都合上，3 つのチャンネルを用意する必要がある．よって RGB のそれぞれのチャンネルに同じ q_u を入力した．そのため **図-4** のようにグレースケールで表現されることになる．

q_u の幅は 19,850 kN/m^2 で，これを 256 で割ると 77.5 kN/m^2 となる．つまり q_u を線形的に色に変換した場合，同じピクセル値の中に q_u が 77.5 kN/m^2 異なる要素が存在することになる．10,000 kN/m^2 と 10,077.5 kN/m^2 ならともかく，50 kN/m^2 と 127.5 kN/m^2 の差は解析結果に大きく影響すると考えられる．そのため q_u の対数をとり，q_u が小さい時はピクセル値の変化が小さく，q_u が大きい時は変化が大きくなるようにした．50 kN/m^2 の要素のピクセル値を 254，19,900 kN/m^2 を 0 とし，余白部は 255 とした．q_u とピクセル値の変換式を式(3)に示す．

$$\text{ピクセル値} = -42.429 \times \ln q_u + 419.984 \qquad (3)$$

この場合，ピクセル値 253 と 254 の強度差は約 1.2 kN/m^2 であり，低強度の範囲を感度よく表現できる．反対にピクセル値 0 と 1 の強度差は約 463.5 kN/m^2 と非常に大きいが，解析には大して影響しないと考えた．

5.　解析条件について

(1)　重みの再学習を開始するブロック

前述のとおり，ファインチューニングでは CNN アーキテクチャの一部の層を解凍して重みを再学習する必要がある．どの層から先を再学習するべきかの知見が無いため，CNN アーキテクチャごとに総当たりで検証した．例として**図-5** に VGG16 の構造イメージを示す[33]を一部変更．入力層や出力層などを除いて，大きく 5 つのブロックで構成されている．なお VGG16 のタスクは 1,000 通りのラベルへの分類であり，単一の連続値の回帰タスクへ流用するためには，出力層の変更が必要となる．具体的には，元の全結合層および Softmax 出力層を除去したうえで，GlobalAveragePooling2D[34]により出力をベクトル化し，Dropout[35]により過学習を抑制し，最後に単一出力の全結合層を付け加えている．

ブロック単位で重みを再学習し，転移学習を含めて合

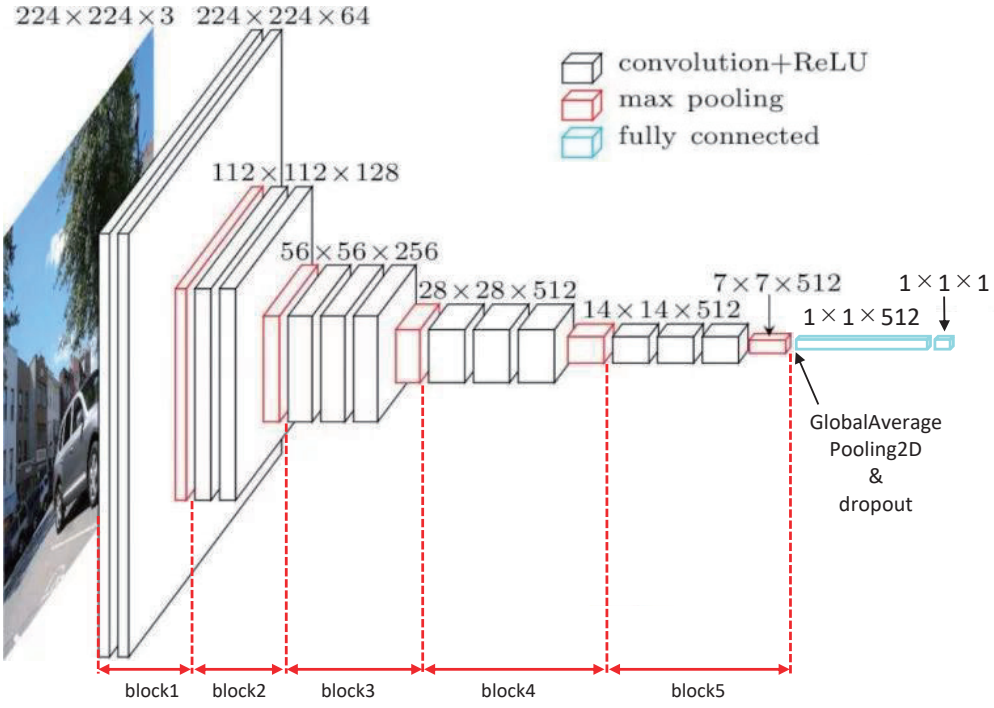


図-5　VGG16 の構造のイメージ[33]を一部変更

計 6 パターンで検証した．他の 6 つの CNN アーキテクチャのブロック数は**表-6** に記載したとおりである．

(2) ハイパーパラメータ

機械学習を実施するにあたって，人間が事前に設定しなければならないパラメータをハイパーパラメータ（以下，HP）と言う．機械学習のアルゴリズムによって程度は異なるが，精度よく推定を行うためには HP の最適化が必須である．CNN の HP は多く存在し，その中には中間層の数などネットワーク構造に関するものも含まれる．ただし本研究では転移学習およびファインチューニングを実施するため，構造に関する検討は必要ない．よって**表-7** に示す 4 つのみを最適化することとした．なお過学習を防止するため early stopping [例えば25)]という手法を採用し，HP の一つである patience によって学習が途中で打ち切られる仕組みとした．最大の epoch 数（学習回数）は 100 と設定した．HP の探索は総当たり法のグリッドサーチで実施した．なお，7 つの CNN アーキテクチャそれぞれについて最適なブロックを求めること，さらに後述のデータの分割を合わせると解析実行数が非常に多くなるため，ここでの HP の探索範囲は最小限とした．最適な CNN アーキテクチャおよびブロックが決まった段階で HP の探索範囲を広げて最適化することにした．

ここで，**表-8** に本研究で使用した PC のスペックを示す．CNN は画像を扱う手法であるため，GPU を搭載した PC を使用することが望ましい．CPU のみでも CNN を実行することは可能だが，解析に要する時間が非常に長くなる．GPU のスペックを決める項目の一つに GPU メモリがある．その他の項目の性能が低くても解析に要する時間が増えるだけだが，GPU メモリが不足すると解析そのものが実行できなくなる場合があることに留意されたい．実際に，本研究で使用した CNN アーキテクチャの中で構造が最も大きい EfficientNetV2L において，ファインチューニングで batch size が 16 の解析は実行できなかった．

表-7　最適化するハイパーパラメータ

名称	範囲	概要
patience	10	学習打ち切りの判断基準
dropout	0.0, 0.2	不活性化するニューロンの割合
learning rate	10^{-3}, 10^{-4}	学習率，学習の進む速度を制御
batch size	8, 16	一度に処理するデータ量

表-8　使用した PC のスペック

OS	Windows11
CPU	Intel(R) Core(TM) i5-14400F　2.50 GHz
メモリ	32GB
GPU	NVIDIA GeForce RTX 4060 Ti
GPUメモリ	16GB

(3) データの分割

HP を最適化するにあたって，手持ちのデータを学習用と評価用に分ける必要がある．流れとしては，①HP の組合せを決定，②学習データと評価データに分割，③学習データで重みを学習，④学習した重みを用いて評価データを推定，⑤得られた精度で HP の組合せの良し悪しを評価，という形になる．①～⑤の工程を 1 回繰り返すだけでは，たとえ良好な精度が得られたとしても，データの組合せが偶然良かっただけという可能性を否定できない．そこでクロスバリデーションと呼ばれる手法を採用した．この手法では，まずデータを n 個のグループ（これを fold という）に分け，このうちの一つを評価用とする．評価用のデータを入れ替えながら①～⑤の工程を繰り返し，n 回の検証結果の平均値により HP の組合せを評価するというものである．さらに，クロスバリデーションには関わらないテストデータをあらかじめ分離しておくことで，完全に未知のデータに対する推定精度を評価することができる．なお，FEM 解析を全く実施していないケースを推定することを模擬して，ここではケース単位でデータを分割した．本研究におけるクロスバリデーションのイメージを**図-6** に示す．

全 35 ケースを 7 ケースずつの 5fold に分割し，うち 1 つをテストデータ，残りの 4 つを学習・評価データとした．テストデータはできるだけ偏らないよう，平均強度と変動係数の設定が異なるケースを人為的に選定した（**表-4, 5** の赤字のケース）．**図-7** に各ケースの最大変位－最小変位のヒストグラムを示す．なお**表-4** のケース 1_2，**表-5** のケース 2_2 はクリティカル部に小さな平均強度(500 kN/m^2)と大きな変動係数(0.80)を設定したケースで，他と比較して極端に大きな変位が生じている特異ケースである．これらのケースの推定は難しいと考え，テストデータには含めなかった．学習・評価データの分割はランダムに実施した．

6. 最適な学習条件の決定について

(1) CNN アーキテクチャとブロックの決定

転移学習を含めて，7 つの CNN アーキテクチャで合計 58 ブロックで解析を実施した．その結果を**表-9** に示す．平均値とは，HP の組合せごとに 4fold 分の RMSE（二乗平均平方根誤差）の平均値を計算し，さらに HP の組合せ 8 通りの平均値を取ったものである．最小値は同様に，HP の組合せ 8 通りの最小値である．VGG16・block3 および VGG19・block3 が評価データの RMSE の最小値が最も小さい結果であった．このうち，評価データの RMSE の

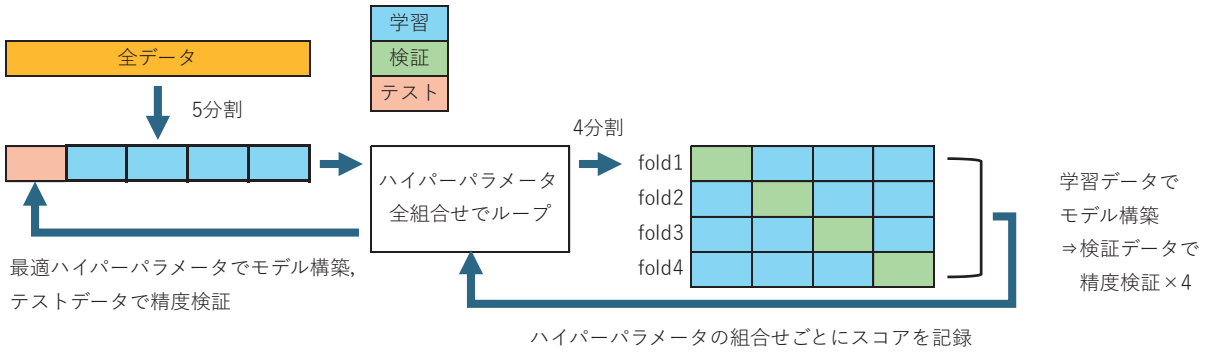


図-6　クロスバリデーションのイメージ

平均値がより小さく，安定して良好な精度が得られるVGG16・block3に決定した．

(2)　ハイパーパラメータの再探索

VGG16・block3において，**表-10**に示すHPで再度解析を実施した．**表-9**を見ると，学習データよりも評価データのRMSEの方が小さい場合がある．これは学習回数の不足が原因と考え，patienceの範囲を大きい側に広げることとした．これにより過学習のリスクが高まるため，dropoutも範囲を広げている．

解析結果を確認すると，learning rateが10^{-3}のときのRMSEの平均は0.071 m，10^{-4}のときは0.021 mであった．よってlearning rateは小さい方が高精度と判断した．残りのパラメータとRMSEの関係を示した図を**図-8**に示す．batch sizeはプロットの色で表現した．これより，patienceは15以上，dropoutは0.0もしくは0.2，batch sizeは8以下が望ましいと判断した．再度HPを設定して解析した結果，dropout: 0.0，learning rate: 10^{-5}，batch size: 8をHPの最適な組合せと決定した．なお，これらのHPを用いて推定モデルを構築する際はearly stoppingを実施しないため，patienceの設定は不要である．その代わり，4fold分の学習打ち切り時の学習回数の平均値44をepoch数として設定した．

7.　推定結果の評価

ここまでの検討により，CNNアーキテクチャ，再学習を開始するブロックおよびHPの最適な組合せが決定した．学習・評価データを結合したものにこれらの学習条件を適用し，推定モデルを構築した．本章ではこの推定モデルを用いてテストデータ，すなわちFEM解析を実施していないケースの残留水平変位を推定し，その精度を確認する．

(1)　個々の解析結果の推定精度

7ケース分のテストデータの推定結果を**図-9**，**表-11**に示す．**表-11**には，RMSEに加えてNRMSE（正規化したRMSE）を示している．これはRMSEをケースごとの残留水平変位の最大値－最小値で除した値である．

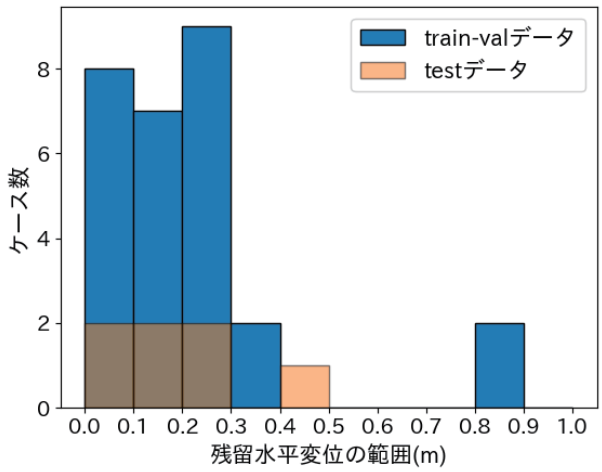


図-7　残留水平変位の範囲のヒストグラム

図-9を見ると，ケース1_3およびケース2_25を除いてFEM解析値と推定値が等しいことを意味する黒の実線付近におおよそのプロットが位置しており，全体の傾向は良く捉えられていることがわかる．**表-11**を見ると，RMSEは0.01～0.04 m程度で，解析結果がおおよそ1.0 m前後であることを考えると，誤差は小さいように見える．ここで**図-7**を見ると，ほとんどのテストケースの残留水平変位の範囲は0.30 m以下となっている．残留水平変位の範囲を考慮した指標であるNRMSEは全てのテストケースで5 %を超えており，推定精度は低い結果となった．これより，全体のおおよその傾向を捉えることはできても，個々の解析結果を精度よく推定するのは難しいことが分かった．

(2)　96 %信頼性上限値の推定

ここで，NRMSEではなく96 %信頼性上限値を用いて評価することを考える．96 %信頼性上限値（以下，96 %値）とは，ケースごとの100メッシュのうち，5番目に大きい残留水平変位を指す．つまりこの値が設計許容値よりも小さければ，96 %以上の確率で要求性能を満足す

表-9　最適なCNNアーキテクチャとブロック

モデル	ブロック	学習データ		評価データ	
		平均値	最小値	平均値	最小値
		m	m	m	m
VGG16	転移学習	0.085	0.079	0.079	0.076
	1	0.040	0.012	0.038	0.015
	2	0.044	0.009	0.043	0.015
	3	0.049	0.010	0.044	**0.014**
	4	0.060	0.012	0.055	0.015
	5	0.020	0.013	0.023	0.017
VGG19	転移学習	0.081	0.072	0.076	0.069
	1	0.045	0.014	0.043	0.016
	2	0.056	0.013	0.051	0.019
	3	0.054	0.011	0.050	0.014
	4	0.056	0.016	0.052	0.024
	5	0.037	0.013	0.037	0.019
MobileNetV2	転移学習	0.073	0.060	0.073	0.062
	1	0.090	0.028	0.088	0.025
	2	0.078	0.020	0.076	0.020
	3	0.162	0.022	0.157	0.024
	4	0.084	0.022	0.084	0.026
	5	0.045	0.018	0.047	0.019
	6	0.080	0.023	0.084	0.025
	7	0.070	0.026	0.073	0.029
	8	0.060	0.025	0.065	0.030
	9	0.063	0.018	0.068	0.022
	10	0.051	0.021	0.053	0.024
	11	0.055	0.027	0.055	0.032
	12	0.058	0.026	0.059	0.031
	13	0.052	0.024	0.057	0.027
	14	0.055	0.036	0.054	0.039
	15	0.051	0.025	0.051	0.031
	16	0.073	0.041	0.075	0.046

モデル	ブロック	学習データ		評価データ	
		平均値	最小値	平均値	最小値
		m	m	m	m
ResNET50	転移学習	0.088	0.080	0.077	0.075
	1	0.036	0.026	0.040	0.028
	2	19.407	0.175	19.416	0.190
	3	0.044	0.014	0.046	0.023
	4	0.038	0.021	0.038	0.026
	5	0.048	0.038	0.046	0.036
EfficientNetV2S	転移学習	0.076	0.072	0.073	0.071
	1	0.028	0.014	0.030	0.020
	2	0.022	0.018	0.023	0.018
	3	0.023	0.017	0.022	0.018
	4	0.023	0.018	0.023	0.020
	5	0.027	0.023	0.027	0.024
	6	0.050	0.042	0.052	0.043
EfficientNetV2M	転移学習	0.103	0.102	0.095	0.093
	1	0.028	0.019	0.030	0.022
	2	0.025	0.019	0.025	0.019
	3	0.026	0.019	0.027	0.022
	4	0.030	0.017	0.029	0.017
	5	0.030	0.023	0.032	0.026
	6	0.057	0.042	0.055	0.043
	7	0.077	0.069	0.071	0.064
EfficientNetV2L	転移学習	0.097	0.092	0.087	0.081
	1	0.023	0.019	0.024	0.022
	2	0.029	0.022	0.026	0.021
	3	0.031	0.020	0.025	0.021
	4	0.027	0.020	0.026	0.022
	5	0.030	0.021	0.028	0.018
	6	0.050	0.038	0.044	0.035
	7	0.069	0.065	0.057	0.053

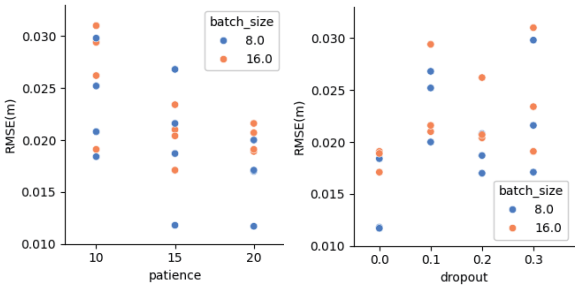
るということになる．MCS では個々の解析結果よりも，このような統計値が重要な評価指標となる．なお既往の検討[9]では，6 番目に大きい残留水平変位を 95 %信頼性上限値として FEM 解析結果の評価に使用していたが，後述のデータ数に応じた標準偏差の検討のため，本研究では 96 %値を使用する．

ここで 96 %値ではなく，FEM 解析結果の平均値および標準偏差を推定することも考えられる．FEM 解析結果が正規分布もしくは対数正規分布に従う場合，平均値と標準偏差から分布形状を仮定し，これより 96 %値を推定できるためである．ただし，シャピロ-ウィルク検定[36]を実施したところ，全 35 ケース中 15 ケースのみが少なくともどちらかの確率分布に従う（有意水準 5 %）という結果であった．そのため本研究では，96 %値を直接推定することとした．

FEM 解析から求めた本来の 96 %値（以下，96 %真値）と，機械学習の推定結果より求めた 96 %値（以下，96 %推定値）を比較した結果を**図-10**，**表-12** に示す．パーセント誤差とは，96%真値と 96%推定値の絶対誤差を

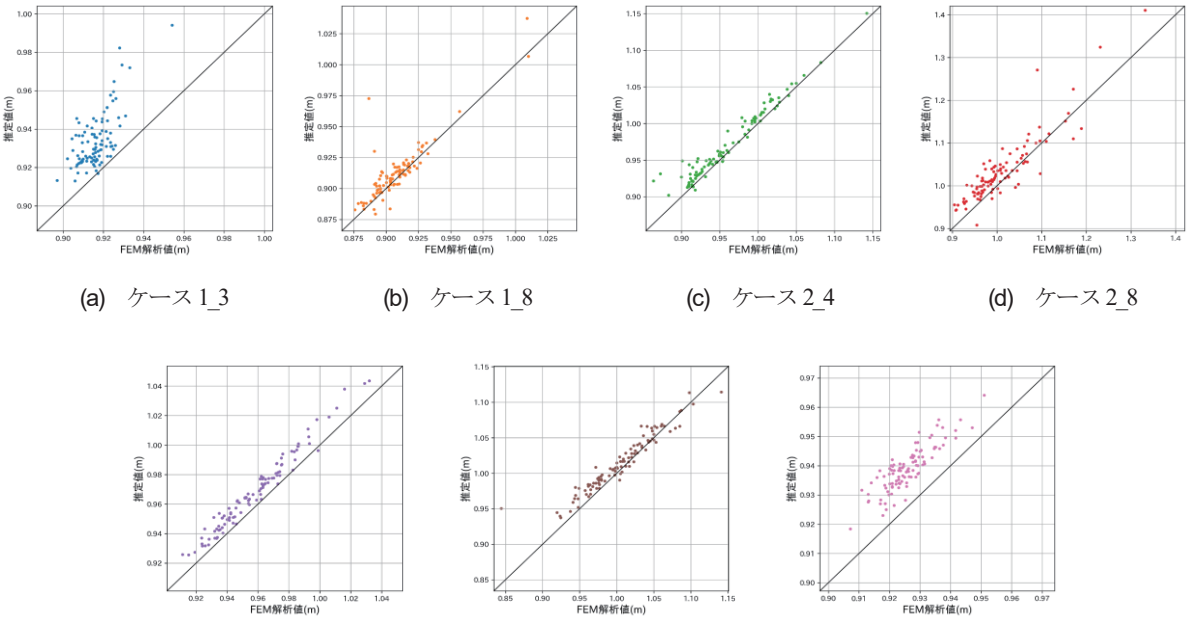
表-10　ハイパーパラメータの再探索範囲

名称	範囲
patience	10, 15, 20
dropout	0.0, 0.1, 0.2, 0.3
learning rate	10^{-3}, 10^{-4}
batch size	8, 16



(a)　patience　　(b)　dropout

図-8　ハイパーパラメータと RMSE の関係



(a)　ケース 1_3　(b)　ケース 1_8　(c)　ケース 2_4　(d)　ケース 2_8

(e)　ケース 2_12　(f)　ケース 2_18　(g)　ケース 2_25

図-9　テストデータの推定結果

96 %真値で除したものである．**図-9** において，個々の FEM 解析値に対して推定値の正負のバランスが比較的良いケース 1_8，2_4，2_8 および 2_18 はパーセント誤差が 1 %を下回っている．これは個々の推定誤差が大きくても，正負の誤差が上手く打ち消しあうことで，96 %値のような統計値の推定誤差は小さくなったものと考えられる．一方で FEM 解析値よりも推定値の方が明らかに大きくなっている残りの 3 ケースでは，このような効果を受けることができずに誤差が大きくなったと考えられる．ただし 4 %程度の誤差が出ることを見越して，96 %推定値に例えば 1.04 の係数を乗じることで安全性を担保できる．このような形で実務適用は可能だと考える．

(3)　損失関数の変更による正負バランスの改善

本研究では学習時の損失関数に MSE（平均二乗誤差）を採用している．MSE では外れ値の影響を強く受けることから，大きな外れ値に影響されて推定値が全体的に大きく，もしくは小さくなることがある．ここで MAE（平均絶対誤差）は外れ値に寛容であり，MSE と比較して正負の誤差のバランスが良くなることが期待される．推定モデルを構築する際の損失関数を MAE に変更するだけでは大きな変化は確認できなかったが，解析条件の最適化の段階から MAE を用いることで，正負の誤差のバランスを改善できる可能性はある．

(4)　生成メッシュ数と 96 %値の標準偏差の関係

FEM 解析における各ケースの 100 メッシュの空間分布は乱数により生成している．そのため，クリティカル部

表-11　テストデータの推定結果

ケース	変位の範囲	RMSE	NRMSE
	m	m	%
1_3	0.057	0.021	35.91
1_8	0.134	0.012	8.83
2_4	0.279	0.015	5.51
2_8	0.428	0.041	9.71
2_12	0.121	0.011	9.30
2_18	0.297	0.017	5.90
2_25	0.044	0.014	31.53

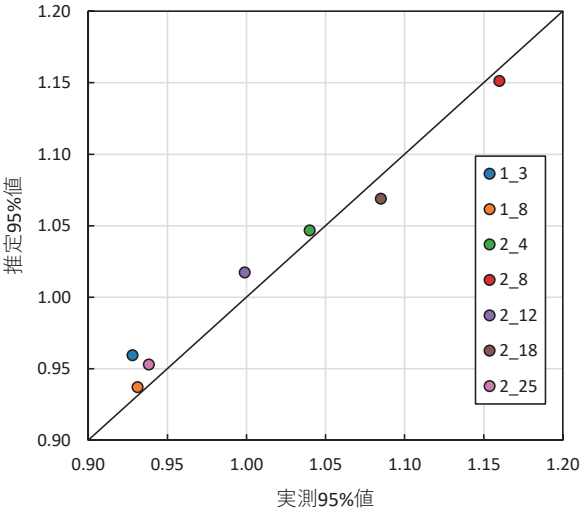


図-10　96 %真値と 96 %推定値の比較

に小さなq_uが集中するメッシュが偶然にも多くなると，96 %推定値が上振れることになる．この場合，生成メッシュ数を増やすことで 96 %推定値のバラつき，すなわち標準偏差を小さくすることができる．

7 つのテストデータにおいて，100 メッシュから 20, 40, 60 および 80 メッシュをランダムに抽出し，メッシュ数×0.05 番目に大きい値を 96 %値とし，96 %真値と 96 %推定値の誤差を算出する．これを 1,000 回繰り返し，誤差の標準偏差を求めた結果を**図-11** に示す．全体的な傾向として，抽出するメッシュ数が多くなるほど標準偏差が小さくなることがわかる．これは，誤差のばらつきは$1/\sqrt{n}$（nはデータ数）に比例して小さくなるためである．メッシュの作成および推定には大して時間はかからないため，例えば 1,000 メッシュを作成することで，96 %推定値の信頼性を高めることができる．

(5)　総括

最後に，本研究の目的である解析コスト低減について言及する．本研究を開始する時点では CNN アーキテクチャ，ファインチューニングおよび HP に関する知見が全く無かったため，総探索という非常に時間のかかる手法で最適化検討を実施した．最適な学習条件を求めるために実施した解析回数は合計 2,128 回で，**表-8** に示したスペックの PC で合計約 220 時間かかった．ただし素人が組んだコードのため，最適化することで低減できる可能性は大いにある．また最適な学習条件を求めてさえしまえば，条件にもよるが，推定モデルの構築 1 回に要する時間は 15 分程度で，これを用いた推定には 1 分もかからない．

今後同様の手法で検討を行う場合は，本研究で得られた知見を活用することで，より短時間で結果を出すことができると考えられる．

8.　おわりに

合理的な設計が可能になる一方で，解析コストが大きく実務適用が難しいレベル 3 信頼性設計法について，機械学習により FEM 解析を代替するサロゲートモデルを構築する検討を行った．その成果を下記に示す．

(1) 改良体の一軸圧縮強さの空間的ばらつきを考慮した FEM 解析の結果を使用し，これを機械学習で代替するサロゲートモデルの構築について検討した．

(2)機械学習手法には CNN を採用し，既存の CNN モデルの構造と学習済みの重みを流用する転移学習およびファインチューニングによって，CNN の構造設計を行った．

(3)改良体要素の一軸圧縮強さの空間分布を説明変数とし，これをピクセル値に対数変換し，224×224 ピクセルの配列データとして表現した．

(4) 総当たり法によって CNN アーキテクチャ，重みの再学習を開始するブロック，およびハイパーパラメータの最適な学習条件を決定した．

(5) 個々の FEM 解析結果を精度よく推定することは難しいが，統計値である 96 %信頼性上限値については，正負の誤差が上手く打ち消しあうことで精度よく推定できることを確認した．正負の誤差のバランスが悪いケースについてはその限りではない．ただし，ある程度の誤差が生じることを見越して，機械学習による推定値に係数を乗じることで安全性を担保できる．

(6) FEM 解析を実施していない未知のケースについても 96 %信頼性上限値を推定できることから，サロゲートモデルの構築によって解析コストを大幅に低減できることを示した．

表-12　96 %真値と 96 %推定値の比較

ケース	実測95 %値	推定95 %値	絶対誤差	パーセント誤差
	m	m	m	%
1_3	0.928	0.965	0.037	3.94
1_8	0.932	0.940	0.007	0.77
2_4	1.044	1.054	0.010	0.98
2_8	1.171	1.170	0.001	0.12
2_12	1.006	1.019	0.013	1.29
2_18	1.085	1.087	0.002	0.19
2_25	0.942	0.953	0.011	1.22

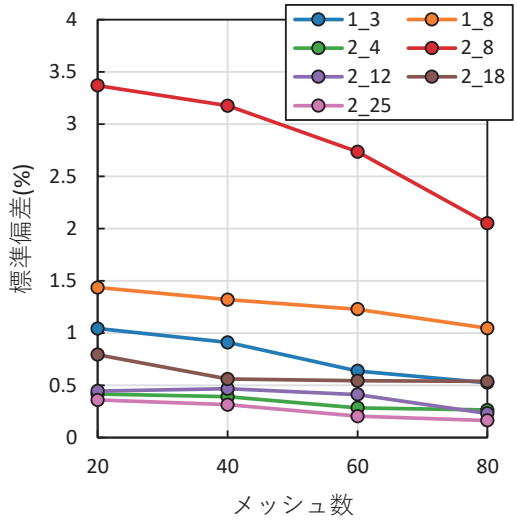


図-11　メッシュ数と 96 %推定値の誤差の標準偏差の関係

参考文献

1)　公益社団法人日本港湾協会：港湾の施設の技術上の基準・同解説，p.34，2007.
2)　北詰昌樹，長尾毅：深層混合処理地盤の性能設計法導入に向けた検討について，港湾空港技術研究所報告，Vol.46，No.1，pp.33-44，2007.
3)　公益社団法人日本港湾協会：港湾の施設の技術上の

基準・同解説，p.1444，2018.
4) 公益社団法人日本港湾協会：港湾の施設の技術上の基準・同解説，p.385，2018.
5) 公益社団法人日本港湾協会：港湾の施設の技術上の基準・同解説，pp.786-787，2018.
6) 公益社団法人日本港湾協会：港湾の施設の技術上の基準・同解説，pp.1453-1454，2018.
7) 笠間清伸，長山達哉，濱口信彦，杉村佳寿，藤井照久，金子智之，善功企：性能規定に基づく浸透固化処理地盤の支持力評価，土木学会論文集 C（地圏工学），Vol.78，No.1，pp.45-49，2022.
8) 笠間清伸，長山達哉，濱口信彦，杉村佳寿，藤井照久，金子智之，善功企：性能規定に基づく浸透固化処理地盤の地震後変形照査，土木学会論文集 C（地圏工学），Vol.78，No.3，pp.235-250，2022.
9) Katayama, Y., Ueno, K., Takahashi, H. and Kasama, K.: Seismic response analysis of ground improved by Deep Mixing considering spatial variability, Japanese Geotechnical Society Special Publication, Vol.11, Issue 7, pp.214-221, 2024.
10) 大竹雄，本城勇介：応答曲面を用いた実用的な地盤構造物の信頼性設計法：液状化地盤上水路の耐震設計への適用，土木学会論文集 C（地圏工学），Vol.68，No.1，pp.68-83，2012.
11) 住岡直樹，宮田正史，福永勇介，大竹雄：特異値分解による線形次元削減と代替モデルに基づく耐震性能照査手法の高度化に向けた基礎的研究～重力式岸壁に対する地震応答解析への適用～，国土技術政策総合研究所資料，No.1100，2020.
12) 笠間清伸，古川全太郎，西田啓介，秋本哲平，堤彩人，片山遥平：機械学習を用いた浸透固化処理地盤の地震時挙動の推定，第 57 回地盤工学研究発表会，DS-9-06，2022.
13) Müller, A. C. and Guido, S.（中田秀基訳）：Python ではじめる機械学習，pp.2-3，オライリー・ジャパン，2017.
14) Iai, S., Matsunaga, Y. and Kameoka, T.: Strain space plasticity model for cyclic mobility, Report of Port and Harbour Research Institute, Vol. 29, No. 4, pp. 27-56, 1990.
15) 一般財団法人沿岸技術研究センター：港湾・空港における深層混合処理工法技術マニュアル（改訂版），pp.251-257，2018.
16) 一般財団法人沿岸技術研究センター：事前混合処理工法技術マニュアル（改訂版），p.233，2019.
17) Honjo, M.: A Probabilistic Approach to Evaluate Shear Strength of Heterogeneous Stabilized Ground by Deep Mixing Method, Soil and Foundations, Vol.22, No.1, 1982.
18) Navin, M. P. and Filz, G. M.: Statistical analysis of strength data from ground improved with DMM columns, Proceedings International Conference on Deep Mixing Best Practice and Recent Advances, 2005.
19) 一般財団法人沿岸技術研究センター：港湾・空港における深層混合処理工法技術マニュアル（改訂版），pp.101, 208，254，2018.
20) Müller, A. C. and Guido, S.（中田秀基訳）：Python ではじめる機械学習，pp.82-90，102-115，オライリー・ジャパン，2017.
21) LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P.: Gradient-Based Learning Applied to Document Recognition, Proceedings of the IEEE, Vol.86, Issue.11, pp.2278-2324, 1998.
22) Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J. and Houlsby, N.: An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, Proceedings of the International Conference on Learning Representations, 2021.
23) Tensorflow：https://www.tensorflow.org/（2025 年 7 月 29 日確認）
24) Keras：https://keras.io/（2025 年 7 月 29 日確認）
25) Deep Learning : https://www.deeplearningbook.org/（2025 年 7 月 29 日確認）
26) Pan, S. J. and Yang, Q.: A Survey on Transfer Learning, IEEE Transactions on Knowledge and Data Engineering, Vol.22, Issue.10, pp.1345-1359, 2009.
27) Yosinski, J., Clune, J., Bengio, Y. and Lipson, H.: How transferable are features in deep neural networks?, Advances in Neural Information Processing Systems 27, pp.3320-3328, 2014.
28) Simonyan, K. and Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, Proceedings of the International Conference on Learning Representations, 2015.
29) He, K., Zhang, X., Ren, S. and Sun, J.: Deep Residual Learning for Image Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.770-778, 2016.
30) Sandler, M., Howard, A., Zhu, M., Zhmoginov, A. and Chen, L. C.: MobileNetV2: Inverted Residuals and Linear Bottlenecks, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp.4510-4520, 2018.
31) Tan, M. and Le, Q. V.: EfficientNetV2: Smaller Models and Faster Training, Proceedings of the International Conference on Machine Learning, 2021.
32) ImageNet: https://www.image-net.org/index.php （2025 年 6 月 17 日確認）
33) VGG16: https://neurohive.io/en/popular-networks/vgg16/（2025 年 6 月 17 日確認）
34) Lin, M., Chen, Q. and Yan, S.: Network In Network, Proceedings of the International Conference on Learning Representations, 2014.
35) Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R.: Dropout: A Simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning Research 15, pp.1929-1958, 2014.
36) Shapiro, S. S. and Wilk, M. B.: An Analysis of Variance Test for Normality (Complete Samples), Biometrika, Vol.52, No.3/4, pp.591-611, 1965.

(Received June 30, 2025)
(Accepted August 31, 2025)

Construction of surrogate model by machine learning using FEM analysis results of ground improved by Deep Mixing

Youhei KATAYAMA and Kiyonobu KASAMA

When ground improvement methods, such as Deep Mixing, are applied to the foundation of port and harbor structures, it is necessary to conduct a seismic response analysis under the level 2 earthquake ground motion. The level 1 reliability-based design cannot directly take into account for spatial variability, so the average unconfined compression strength of in-situ improved ground is multiplied by various coefficients to reduce it to simulate spatial variability. In contrast, the level 3 reliability-based design can directly take into account for this spatial variability, enabling a more rational design from the perspectives of structural stability and economic efficiency. However, this design method requires Monte Carlo simulation, and the analysis cost is very high when applied to FEM. In this study, we investigated the construction of a surrogate model by machine learning to reduce the analysis cost.