

# 機械学習による代替モデル生成のための実行基盤

大辻 弘貴<sup>1</sup> 清水 香壺<sup>2</sup> 古屋 篤史<sup>2</sup> 土肥 義康<sup>1</sup> 富田 安基<sup>1</sup> 中島 耕太<sup>1</sup>

**概要:** 機械学習の応用範囲の広がりにより、観測等による実データに留まらず、コンピュータシミュレーションにより生じたデータに対して学習を行い、未知の条件に対しても迅速に結果を得るための代替モデル生成が行われつつある。代替モデルの生成にあたっては、学習対象とするデータの性質や量をコントロールすることが、精度向上やモデル生成に要する時間を短縮する上で重要である。シミュレーションデータを学習に用いる場合の特徴として、条件設定を変更することで、任意の学習データを随時追加可能であることが挙げられる。ところが、既存事例では既に生成されたシミュレーションデータに対して学習を行っており、この利点が活かされていない。シミュレーションと学習を並行して実行することにより、データ生成のためのシミュレーション条件の最適化や、モデルパラメータのチューニングをより早い段階で行うことが可能となり、従来事例よりも高い精度のモデルが短時間で生成されることが期待できる。本研究では、電磁界シミュレーションにおける事例により、36%から44%のシミュレーション実行数で代替モデルを生成できる可能性を示した。本稿においては、シミュレーションと機械学習を並行実行するための基盤および、電磁界シミュレーションの事例を対象とした代替モデル精度特性の評価について述べる。

## 1. はじめに

コンピュータシミュレーションは設計や製造の現場で広く用いられており、生産性の向上や試作コストの削減に貢献している [1]。特に設計の最適化においては、小規模な修正とシミュレーションを繰り返すことで最適値を探るため、試行回数を増やすことがより良い設計につながり、迅速に結果を得ることが求められている。並列処理により大規模シミュレーションを高速化することは一般に行われているが、限られた問題サイズの1回のシミュレーションを短縮することは、スレッドあたりの性能向上が大幅には期待できない状況下では困難さを増している [2]。そこで、シミュレーションの入力条件と結果の組を多数用意し、これらを用いて学習器を訓練することにより、代替モデルを生成する試みが行われている。機械学習の推論処理は短時間で完了することから、元のシミュレーションと比較しても極めて高速に結果を得ることができる。また、モデルの学習のためのシミュレーション実行は、ほぼ独立した並列実行であることから、実行効率が高い。

代替モデルには、短時間で妥当な精度の結果を求められることが期待されているが、精度は与えられるシミュレ

ーションデータの性質や量に依存することから、これらの特性を理解した上で適切に学習データの生成を行う必要がある。一方で、与えられるデータ量や入力条件の分布と、代替モデルの精度の関係性については、予測することが難しい。このため、シミュレーションデータは実データとは異なり、必要に応じて生成することが可能である利点を活かし、学習・精度検証とデータ生成を同時に行うことにより、生成するデータの組み合わせや量を制御することで、より少ないシミュレーション実行で高い精度のモデルを構築することを目指した。このようなワークフローを実現するためには、シミュレーションに与える入力条件の生成や、複数シミュレーションの並列実行、データ形式の変換、進捗状況に応じた学習と精度確認をアプリケーションごとにカスタマイズして行わなければならない。このような処理を連携して行う場合、個別の実装となるため実装コストも高く、再利用性も低い。このため、本研究では以上の処理を行うための実行基盤を実装し、一例として電磁界シミュレーションにおけるインダクタの損失を予測する代替モデルを生成し、その評価を行った。

## 2. 関連研究

大量のシミュレーションデータを利用した機械学習による代替モデルの生成や分析は、気象、天体、化学をはじめとした多くの分野で取り組みが見られる。[3] は、気象シミュレーションのデータに対して、Summit 上で学習処理

<sup>1</sup> 株式会社富士通研究所  
コンピュータシステム研究所  
先端コンピュータシステムプロジェクト

<sup>2</sup> 富士通株式会社  
次世代テクニカルコンピューティング開発本部  
アプリケーション開発統括部

を行い、極端現象の検出を行ったものである。並列学習により半精度換算で1 EFLOPSを超える学習処理を行ったことが注目を集めたが、学習対象のデータは既に生成されたものであり、学習状況に応じた動的な変更などは行われていない。また、最大規模における並列学習では、時間あたりの処理量はスケールする一方で、学習精度は悪化しており、このような仕組みは設計プロセスに用いる上では課題がある。[4]は、N体シミュレーションによって得られたダークマターの分布から宇宙論パラメータを推測する機械学習モデルを生成した例である。こちらも同様に、学習過程の精度などの情報をデータ生成段階で利用する仕組みにはなっていない。[5]は、本稿と同じく電磁界シミュレーションを対象とした代替モデル生成であるが、学習データの事前生成を前提としている点では他の関連研究と同様であり、また本稿とは異なり、静的な定常状態の予測のみを目的としたモデルになっている。

本研究は、シミュレーションデータに基づく代替モデルの生成において、データ生成が自在に可能であることを活用しつつ、汎用性の高い実行方法を実現することを目的としている。

### 3. 代替モデル生成実行基盤

#### 3.1 構成

機械学習によりシミュレーションの代替モデルを生成するためには、シミュレーションプログラムを幅広い多数のパラメータで実行し、それらの結果を用いて学習器を訓練する必要がある。図1に、本研究における処理の流れと、図2にシステムのモジュール間のやり取りを示す。最初に、代替モデルの生成部がシミュレーションプログラムの入力条件を生成する。変更すべきパラメータの種類や変動幅はシミュレーション対象により異なるが、形状や入力（外力や電流など）といった値であり、代替モデルにおいて使われうる値の範囲を設定し、これらの値をシミュレーションプログラムの入力条件ファイルとして出力する。

シミュレーションデータの生成にあたっては、複数の条件のシミュレーションを並列動作させれば良く、基本的に相互の依存関係は存在しないため、極めて高い効率で並列動作させることが可能である。シミュレーション実行制御部は、利用可能な計算ノードにおいて、実行状況を監視しつつ最大数のシミュレーションを並列実行する。

シミュレーション結果はアプリケーション固有のフォーマットやHDF5[6]などの構造化フォーマットで出力されるが、そのままでは学習フレームワークの入力には適さない場合があるため、事前あるいはオンザフライにデータ変換を行う。変換されたシミュレーション結果を、学習用データと検証用データに分離した上で、前者のデータを用いて学習器の訓練を行い、代替モデルを生成する。この際、学習器に対する入力はシミュレーション条件として与えたパ

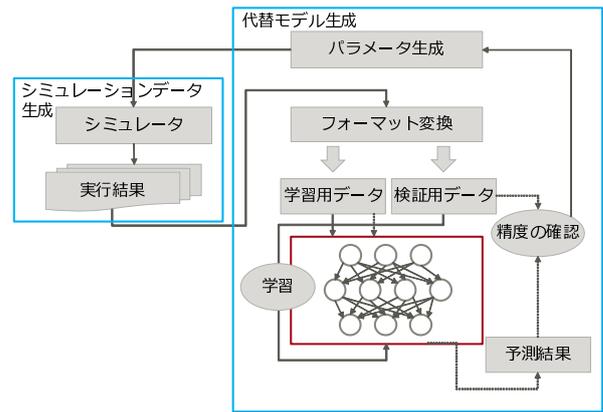


図1 システム全体の構成図を表している。代替モデル生成部が、シミュレーション用のパラメータを生成し、その結果をもとに学習を行う。

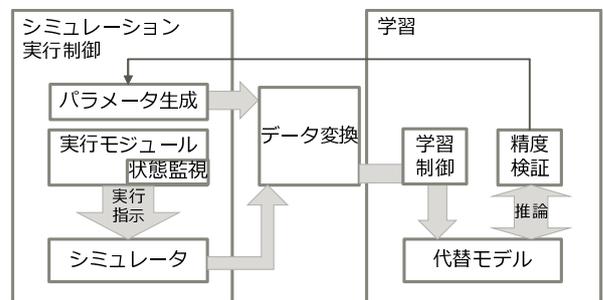


図2 システムのモジュール構成とインタラクション

ラメータであり、出力はシミュレーション実行結果である。

検証用データで生成されたモデルの精度を確認するが、精度検証部とパラメータ生成部の間でやり取りをすることにより、必要に応じてシミュレーションパラメータを追加することも可能である。加えて、すべてのシミュレーションデータが揃う前に、学習パラメータ（ハイパーパラメータ）の最適化を並行して行うことも可能である。

これらの処理により、シミュレーションの代替モデルの生成と精度特性の分析が完了する。

#### 3.2 並行実行

図3に、代替モデル生成実行基盤をクラスタ上で動作させる場合の構成を示す。クラスタは、GPUを搭載するノードから構成され、共有ストレージを有する。将来的には学習専用アクセラレータの導入などにより、均一でないリソース管理が必要になる可能性があるが、現時点ではGPUにおいてモデルの学習を行うため、GPUを利用するシミュレーションの実行と資源を共有することが可能である。

シミュレーション入力条件や実行結果、変換後の学習用データはストレージを介して共有されるため、任意ノードを用いて実行可能である。

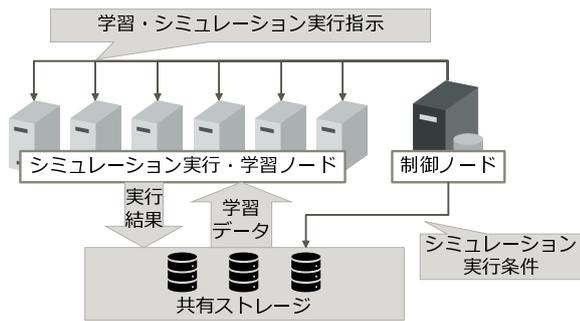


図 3 ノードごとの処理を表している。GPU を有するノードにおいてシミュレーションの並列実行や代替モデルの学習を行う。シミュレーションの入力条件や実行結果は、共有ストレージを介してアクセスされる。

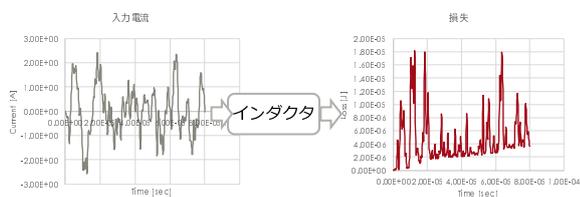


図 4 題材とした電磁界シミュレーションの入出力を示した図である。ある形状のインダクタに対して任意の電流波形を入力することで、エネルギー損失を求めることができる。

## 4. 電磁界シミュレーションの代替モデル生成

### 4.1 シミュレータ概要

これまでに示した枠組みの効果を明らかにするための題材として、電磁界シミュレーションを対象とした代替モデルの生成を行った。対象とするアプリケーションは、A- $\Phi$ 法を用いた有限要素解析を行う富士通製の電磁界シミュレータ [7] である。解析対象のメッシュデータと外部磁場、電流を入力条件として、時系列で全要素の状態や特性値が得られる。ここでは、電流波形を入力とし、シミュレータの出力のうちインダクタにおけるエネルギー損失の推移を教師データとして用いる。

### 4.2 入出力データ形式

本実験においては、解析対象の形状は一定とし、様々な入力電流に対する損失を予測する代替モデルを生成した。

モデルを学習させるにあたり、様々な入力電流に対するシミュレーション結果を用意する必要がある。図 5 は入出力データの流れを表したものである。本実験において用いているシミュレータでは、入力電流は正弦波であれば周波数と振幅、直流オフセット値で入力を受け取り、任意波形の場合は任意点数の 1 次元配列として電流値を与えることが出来る。本実験においては、学習器に対する入力を容易にするため、時間方向のステップ数を 640 に固定した上で、これらを 1 次元配列として入力として用いている。教

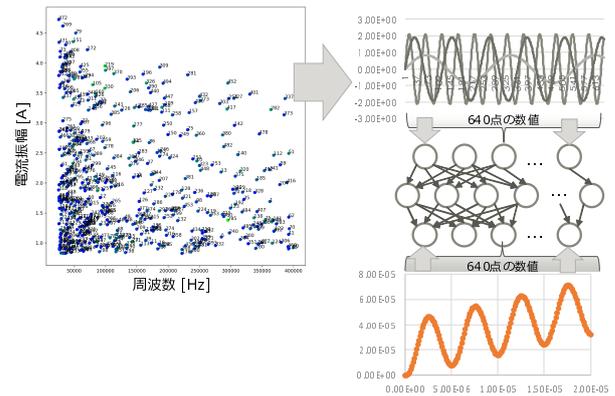


図 5 パラメータで表された波形を学習器に与える流れを表した図である。ここでは一定範囲の電流振幅と周波数の中からランダムに値を選択し、それらに対応する正弦波波形を生成している。

師データは、シミュレータの実行結果から得られたデータを入力する際と同様、1次元の時系列データとして学習器に与えている。

## 5. 評価

### 5.1 評価環境

本実験に用いた評価環境を表 1 に示す。これらのノードを図 3 に示した構成で用いて実験を行った。

表 1 評価環境

CPU	Intel(R) Xeon(R) CPU E5-2695 v4 2.10GHz 18C36T x2
RAM	DDR4-2400 128 GB
OS	CentOS 7.2
GPU	NVIDIA Tesla P100 x2
使用ノード数	9
学習フレームワーク	Keras 2.1.5 (Tensorflow 1.4.1)
ノード間接続	InfiniBand EDR x4 (100 Gbps)
共有ファイルシステム	FEFS (Fujitsu Exascale File System) (Lustre 2.1.0)

### 5.2 評価条件

#### 5.2.1 学習に用いたデータ

本稿に示す代替モデル生成に関して、4章に示す電磁界シミュレータの事例により評価を行った。代替モデルの学習にあたっては、表 2 に示すデータを生成した。三角波は、図 6 に例示する波形であり、周波数・振幅・デューティ比・直流オフセットの 4 パラメータを持つ。それぞれのパラメータについて乱数を用いて、複数の入力波形を生成した。ランダム波形は、高周波成分を抑制するため、入力ウィンドウ (640 ステップ) に対して、4~40 周期の正弦波波形を用意し、それらに対してランダムな増幅率をかけたものを足し合わせることで生成した。波形の具体例は既に図 4 に示したとおりである。

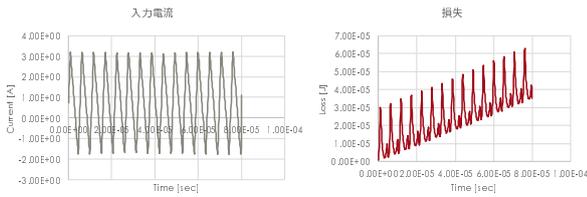


図 6 三角波の入力電流の例と損失の推移を示す。

表 2 生成したデータおよび処理時間

三角波	10,000 パターン
ランダム波形	13,000 パターン
1 回のシミュレーションに要する時間	1371 秒 (ランダム波形 7 サンプル平均)

### 5.2.2 シミュレーション実行

学習データを生成するため、多数のシミュレーションを並列実行した。独立したシミュレーションを並列実行しているため、実行結果の入出力に伴う I/O を除き、スケーラビリティに制約はない。

### 5.2.3 学習条件

学習時の各種パラメータを表 3 に示す。モデルは全結合を用いており、中間層の活性化関数は ReLU [8] の一種である Leaky ReLU [9] を使い、出力層では Linear を用いている。また、エポック数は過学習を防ぐため、サンプル数に応じて 140~600 としている。全サンプル中 2% のデータを検証に使用している。その他、最適化アルゴリズムは AdaDelta を用いた。

表 3 モデルおよび学習パラメータ、処理時間

モデル	全結合 4 層 入力/出力ノード数 640
活性化関数	中間層: Leaky ReLU 出力層: Linear
エポック数	140~600
バッチサイズ	128
損失関数	mse (平均二乗誤差)
最適化	AdaDelta
1 回の推論に要する時間	平均 0.69 ms
学習時間	155.6 sec (11000 サンプル時)

## 5.3 学習状況

10000 サンプルのランダム波形を前節に示す条件で学習させた場合の学習状況を図 7 に示す。training loss と validation loss はともに低下しており、過学習の傾向がないことも分かる。

## 5.4 精度と特性評価

### 5.4.1 評価の概要と目的

代替モデルによる推論精度は、1 章でも述べたように、与えるデータに左右される。代替モデルの精度を向上させる上では多様なデータで学習させることが望ましいが、す

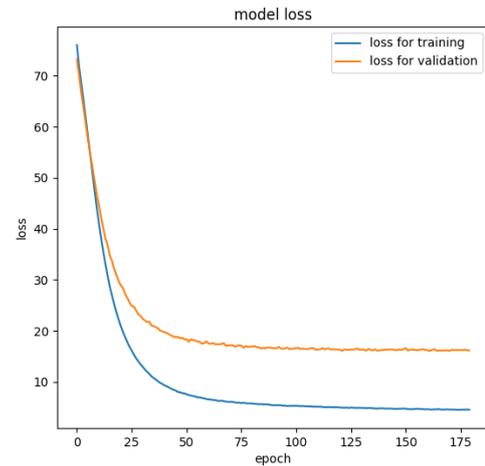


図 7 ランダム波の入力電流に関して学習を行った際の、学習 loss と validation loss のグラフである。横軸はエポック数、縦軸は loss (mse) である。

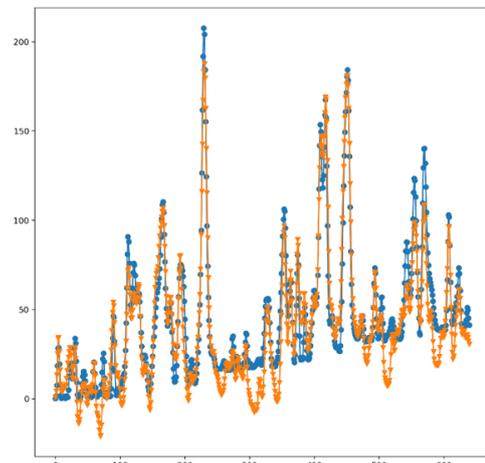


図 8 図 7 のケースにおける、代替モデルによる損失の推測値とシミュレーションによる値 (教師) の比較例である。横軸は時間軸 (640 タイムステップ)、縦軸は損失であり、オレンジ (▲) は代替モデルによる推論値、青 (●) はシミュレーションによる値である。推論対象は、テストデータであり、学習用データには含まれていないものである。

べての条件を網羅することはシミュレーションに要する演算コスト上難しい。汎用性の高い学習データで代替モデルを訓練することにより、用途特化に必要な学習データを生成するためのシミュレーション回数を減らすことが可能であるならば、比較的少ない演算コストで代替モデルの精度を向上させることが可能になる。そこで、一旦汎用性の高い学習データを用意した上で、代替モデルの利用状況に応じて用途別の特化学習データを追加して、代替モデルを再訓練させるシナリオのもと、精度特性の評価を行った。

### 5.4.2 学習データ

前述の評価を行うため、表 2 に示すデータに関して、三角波の入力電流に関する推論を用途特化の一例として実験を行った。代替モデルに与えるデータの種類と数を、図 9 に示す。図中、上段のケースでは三角波の入力電流波形

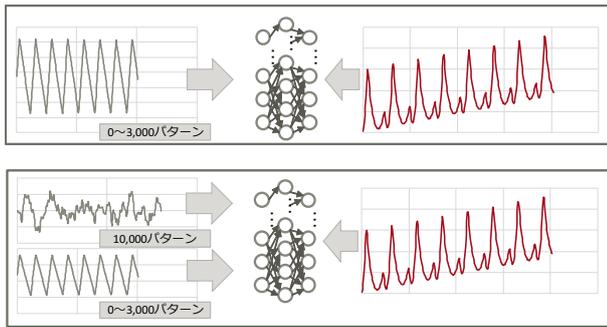


図 9 ランダム波の入力電流に関して学習を行った際の、学習 loss と validation loss のグラフである。横軸はエポック数、縦軸は loss (mse) である。

のみを用いて学習サンプル数を変化させており、下段では 10000 パターンのランダム生成波形データを予めデータセットに追加した状態で、上段と同じく学習データに含まれる三角波波形サンプルの数を変化させている。同等の推論精度に到達するために要する三角波波形（用途特化）のデータ数を比較することにより、汎用性の高い学習データの演算量削減効果を知ることができる。

### 5.4.3 精度評価と考察

本節では、前節に示す実験の結果を示す。図 10 は、用途特化（三角波）の推論精度に関して、汎用性の高い事前データ（ランダム生成波形）によるシミュレーション結果を用いた場合と、用いていない場合における精度比較を行ったグラフである。事前データを用いることにより、同じ特化データの学習サンプル数においても、より高い精度を達成できていることがわかる。この点に着目したものが図 11 であり、一定精度（RMSE4.0 と 3.0）に到達するために要した特化データのサンプル数を比較している。この図から分かるように、事前データを用いることにより、それぞれ 44 %、36 %の特化データサンプル数で同じ精度を達成できている。事前データが代替モデルの多くの用途に共通である限り、その生成に要する演算コストを重視する必要はない一方、目的に応じたシミュレーション実行数が削減できることから、精度の高い代替モデル生成をより少ない演算コストで実現することに貢献するといえる。

特に、本稿が目的とする設計の最適化プロセスに代替モデルを利用する場合、例えば RMSE が 5.0 以下になるモデルを生成するために必要なシミュレーションは 400 回であり、表 1 の環境を用いれば 200~300 並列程度で同時実行が可能であることから、シミュレーション 2 回分の時間で必要な学習データが生成される。一度モデルを生成すれば推論時間は極めて短く済むことから、シミュレーションと学習の時間を含めたとしても、設計のプロセスの高速化が可能である。

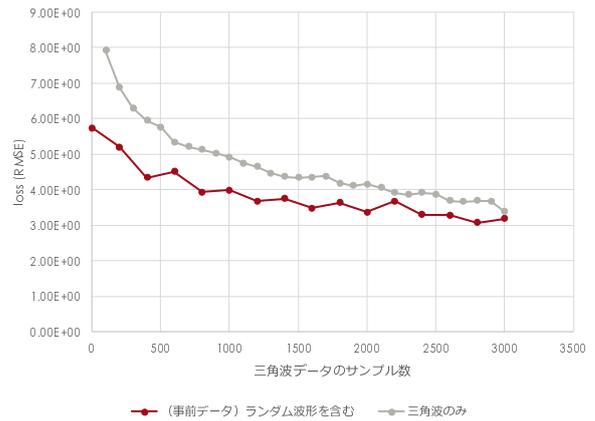


図 10 事前データ（ランダム生成波）の有無による、代替モデル制度の比較。X 軸は特化データ（三角波）のサンプル数で、Y 軸は誤差（RMSE）である。

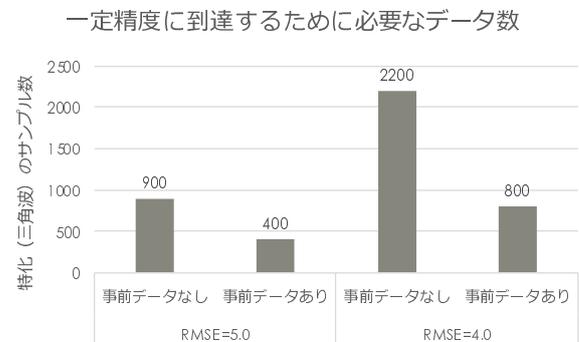


図 11 代替モデルがある一定の精度に到達するまでに必要なデータ数を比較したグラフ

## 6. まとめと今後の課題

### 6.1 まとめ

本稿では、コンピュータシミュレーションを用いた設計プロセスの高速化に資する代替モデル生成に関して、それらの処理を柔軟かつ迅速に行うための処理の枠組みについて示した。また、この枠組みに基づく電磁界シミュレーションの代替モデルの生成を行い、その実行状況と精度特性の評価を行った。評価結果から、代替モデルによって対象アプリケーションの出力結果の一部が推論可能であることと、汎用性のある入力を用いてシミュレーションを実行し、事前データを生成することにより、用途に合わせたモデルを生成する際にも、必要な演算コストを引き下げられることを示した。

### 6.2 今後の課題

本稿で対象としたシミュレーションは出力データ量や種類が限られていたことから、シミュレーション側と学習側のデータのやり取りに課題は生じなかったものの、今後対象をより大規模化した場合、共有ストレージを用いた仕組

みでは性能低下が発生する懸念がある。また、代替モデルがより大規模化した場合、学習処理の並列数を高める必要が生ずる一方、同時に実行されるシミュレーション処理とのリソース配分を動的に行う必要があると考えられる。幅広いシミュレーションに本稿に示した枠組みを適用することで、システム上の課題を明らかにし、解決していくことが求められる。

## 参考文献

- [1] Shephard, M. S., Beall, M. W., O'Bara, R. M. and Webster, B. E.: Toward simulation-based design, *Finite Elements in Analysis and Design*, Vol. 40, No. 12, pp. 1575–1598 (2004).
- [2] Snavely, A., Carrington, L., Wolter, N., Labarta, J., Badia, R. and Purkayastha, A.: A Framework for Performance Modeling and Prediction, *Proceedings of the 2002 ACM/IEEE Conference on Supercomputing, SC '02*, pp. 1–17 (2002).
- [3] Kurth, T., Treichler, S., Romero, J., Mudigonda, M., Luehr, N., Phillips, E., Mahesh, A., Matheson, M., Deslippe, J., Fatica, M., Prabhat and Houston, M.: Exascale Deep Learning for Climate Analytics, *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis, SC '18*, pp. 51:1–51:12 (2018).
- [4] Mathuriya, A., Bard, D., Mendygral, P., Meadows, L., Arneemann, J., Shao, L., He, S., Kärnä, T., Moise, D., Pennycook, S. J., Maschhoff, K., Sewall, J., Kumar, N., Ho, S., Ringenburt, M. F., Prabhat and Lee, V.: CosmoFlow: Using Deep Learning to Learn the Universe at Scale, *Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis, SC '18*, pp. 65:1–65:11 (2018).
- [5] Al-Jarro, A., Tomita, Y., Georgescu, S. and Nakashima, K.: DeepSim-HiPAC: Deep Learning High Performance Approximate Calculation for Interactive Design and Prototyping (2018).
- [6] Folk, M., Cheng, A. and Yates, K.: HDF5: A file format and I/O library for high performance computing applications, *Proceedings of supercomputing*, Vol. 99, pp. 5–33 (1999).
- [7] Shimizu, K., Furuya, A., Uehara, Y., Fujisaki, J., Kawano, H., Tanaka, T., Ataka, T. and Oshima, H.: Loss Simulation by Finite-Element Magnetic Field Analysis Considering Dielectric Effect and Magnetic Hysteresis in EI-Shaped Mn–Zn Ferrite Core, *IEEE Transactions on Magnetics*, Vol. 54, No. 11, pp. 1–5 (online), DOI: 10.1109/TMAG.2018.2839773 (2018).
- [8] Nair, V. and Hinton, G. E.: Rectified Linear Units Improve Restricted Boltzmann Machines, *Proceedings of the 27th International Conference on International Conference on Machine Learning, ICML'10*, pp. 807–814 (2010).
- [9] Maas, A. L., Hannun, A. Y. and Ng, A. Y.: Rectifier nonlinearities improve neural network acoustic models, in *ICML Workshop on Deep Learning for Audio, Speech and Language Processing* (2013).