

# リカレントニューラルネットワークによる材料構成則モデリングの試み Hysteretic Modeling for Structural Analysis by Using Recurrent Neural Network

鈴木 琢也\*<sup>1</sup>  
Takuya Suzuki

松岡 康友\*<sup>1</sup>  
Yasutomo Matsuoka

\*<sup>1</sup> 竹中工務店  
Takenaka Corp.

In nonlinear structural analysis for seismic design, it is very important to use an accurate hysteretic model. However, there are some models for which the human development and application of a hysteretic model are too complex. For the solution to these problems, we think that Recurrent Neural Network (RNN) system is available. RNN is one of the Neural Network (NN). RNN has a characteristic to be able to change by a system depending on past hysteresis of input and output. In this paper, for modeling of very complex non-linear hysteresis, applicability of RNN modeling to normal bi-linear model is confirmed. First, we identify RNN model by training data consisted of bi-linear history data of displacement (input) and force (output). Then, it is confirmed that identified RNN model can become the system which is equal to theoretical bi-linear model.

## 1. はじめに

建築構造解析分野においてその解析精度を確保するためには、部材の非線形履歴挙動を高精度に追跡できるモデルの構築が重要である。しかし、構造部材の挙動は、人が分析しモデルを構築するには複雑すぎるため、一般的には比較的単純なモデルに置換されることが多く、その分精度が犠牲にされている。このような課題に対する解決策として、著者らは、リカレントニューラルネットワーク(以降、RNN)を用いた非線形履歴挙動のモデリングが有効であると考えている。RNN に対して多くの教師データを与えることで、複雑な挙動の特性を人が理解しなくても高精度な履歴挙動をモデリングすることができれば、非常に有用である。

そこで、本報においては、RNN による非線形モデリング手法の確立に向けて、その嚆矢として、基本的な力学モデルであるバイリニアモデルに対する、RNN モデルの適用性を確認する。具体的には、バイリニアの理論モデルへの入力(変位)と出力(力)の時系列データを教師データとして同定された RNN モデルが、理論モデルと同じ入出力システムとなりうるかどうかを確認する。

## 2. 対象とする履歴則

RNN でモデル化対象とする履歴モデルは鋼材のモデリングなどで最も一般的に用いられている移動硬化型のバイリニアモデル (Fig.1) とする。このモデルはスケルトンの初期剛性  $K$ 、降伏荷重  $F_y$ 、降伏後の二次剛性  $K'$  の 3 種の定数パラメータでその特性が決定されるモデルである。今回モデル化対象としたバイリニアモデルの諸元を Table1 に示す。

Fig.2 には、この移動硬化型バイリニアモデルを人がモデル化する場合のコーディングの例を示す。この例では、あるステップにおける現在の変位 ( $\delta_{now}$ ) が入力された場合の出力である反力 ( $F_{now}$ ) を、3 つの直線(前ステップのプロットを通る傾き  $K$  の直線、正側負側の降伏後のスケルトン直線)のいずれかの式を用いて計算している。この例での反力の計算には、スケルトンを決定づける 3 種の定数パラメータのほか、一つ前のステップにおける変位  $\delta_{old}$  と反力  $F_{old}$  が必要になる。これらの値は毎ステ

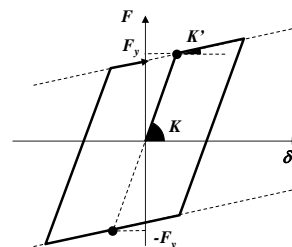
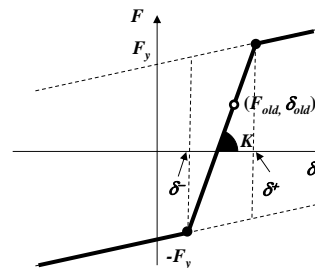


Fig.1 Bi-Linear Model (kinematic hardening)

Table1 Parameters of Bi-Linear Model

Initial Stiffness $K$ (kN/m)	Yield Force $F_y$ (kN)	Second Stiffness $K'$ (kN/m)
100.0	20.0	5.0



$$F_{now} = \begin{cases} K'(\delta_{now} - F_y/K) + F_y & \delta_{now} > \delta^+ \text{ のとき} \\ K'(\delta_{now} - \delta_{old}) + F_{old} & \delta^- < \delta_{now} < \delta^+ \text{ のとき} \\ K'(\delta_{now} + F_y/K) - F_y & \delta_{now} < \delta^- \text{ のとき} \end{cases}$$

where,  $\delta^- = (K\delta_{old} - F_{old}) / (K - K') - F_y / K$   
 $\delta^+ = (K\delta_{old} - F_{old}) / (K - K') + F_y / K$

Fig.2 Cording Example of Kinematic Hardening Bi-Linear Model

連絡先: 鈴木琢也, 竹中工務店, 千葉県印西市大塚 1-5-1, 0476-47-1700, 0476-47-6460, suzuki.takuya@takenaka.co.jp

ップ変化し、それに応じて同じ入力に対しても出力が変わり、荷重変位関係が Fig.1 に示すようなループを描くようになる。

このように、変数パラメータを適切に選定することで、過去の履歴すべてを記憶することが不要になっている。換言すれば、この変数パラメータは、人の手によって抽出された履歴ループの特徴量である。このような履歴ループの特徴量とその利用法を、入力と出力の教師データのみから RNN モデルが再現できるか、できる場合には、どのような NN によって再現できているのかを確認することが本研究の目的である。

### 3. RNN モデルの構築

#### 3.1 構成

システム同定に用いる RNN モデルの全体構成について説明する。RNN の全体像を Fig.3 に示す。図中の中括弧内は渡されるベクトルの長さ(成分の数)を示している。図に示すように、今回用いた RNN は、階層型とし、入力層、出力層および 1 層の中間層(隠れ層)からなる。各層のニューロンの数はそれぞれ 1,1,400 である。

システム全体としては、1 つの変位入力に対して、1 つの力を出力するシステムとなる。一般的な NN との相違点は、中間層からの出力が出力層のみではなく、次のステップにおいて中間層への入力となる点である。そのため、中間層の入力は、システムに入力された 1 つの値に加えて、ひとつ前のステップでの中間層からの出力である 400 個の数値で構成される 401 の成分を持つ列ベクトルとなる。これにより、過去の出力に応じてシステムは時々刻々と変化し、同じ入力値に対しても常に同じ出力がなされるわけではなく、異なる値を出力できる余地を残している。

入力層(L=1)および出力層(L=3)は、一般的な NN と同様 1 入力 1 出力のニューロン 1 つとし、それぞれ活性化関数はいない。そのため、各層への下層からの入力を  $u$ 、上層への出力を  $z$  で表記した場合、各層での出力( $z$ )は入力( $u$ )と同じ値となる。

$$z_{L=1}^i = u_{L=1}^i \quad (1)$$

$$z_{L=3}^i = u_{L=3}^i \quad (2)$$

中間層のニューロンは LSTM(Long Short Term Memory) ブロックで構成される。LSTM とは RNN において長期にわたる記憶を実現できるようにするための手法の一つである[Hochreiter 1997]。LSTM ブロック 1 つの機能を Fig.4 に示す。図に示すように 1 つの LSTM ブロックは通常の NN のニューロンの 1 入力 1 出力ではなく、入力値を 3 つ加えた 4 入力 1 出力となる。図中、 $u$ 、 $z$  は通常のニューロンも有する入出力であり、 $u_{IG}$ 、 $u_{OG}$ 、 $u_{FG}$  が追加された入力値である。 $u_{IG}$ 、 $u_{OG}$ 、 $u_{FG}$  はそれぞれ入力ゲート(InputGate)、出力ゲート(OutputGate)、忘却ゲート(FogetGate)に入力され、入出力および記憶をコントロールするために利用される。図中●はメモリセルを表しこの値が 1 時刻を隔てて、メモリセル自身に帰還することで記憶を実現する。

#### 3.2 学習計画

ここでは、RNN モデルの重みとバイアスを調整し、Fig.1 に示したノーマルバイリニアの理論モデルと同等のシステムとなるようにするための学習計画について述べる。なお、学習には計算プログラム Chainer を用いる。

学習に用いる教師データ(訓練データ)は、Fig.2 に示す人が作成した理論モデルに対する入力(変位)と、その入力に対応する出力(力)の履歴データとする(Fig.5)。数多くの「理論モデ

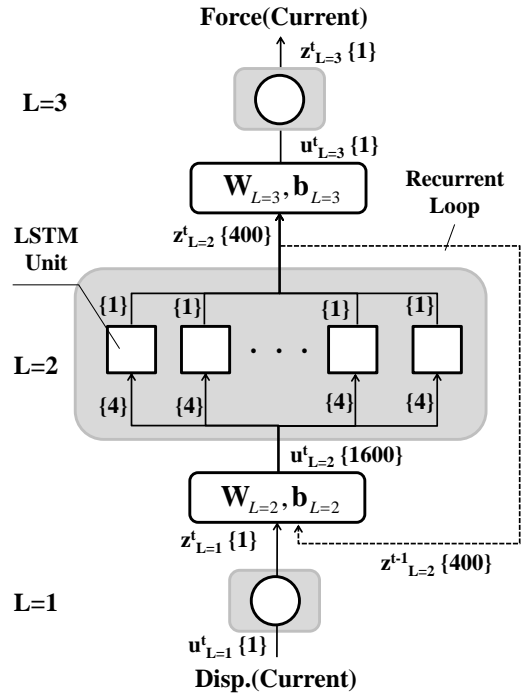


Fig.3 RNN architecture

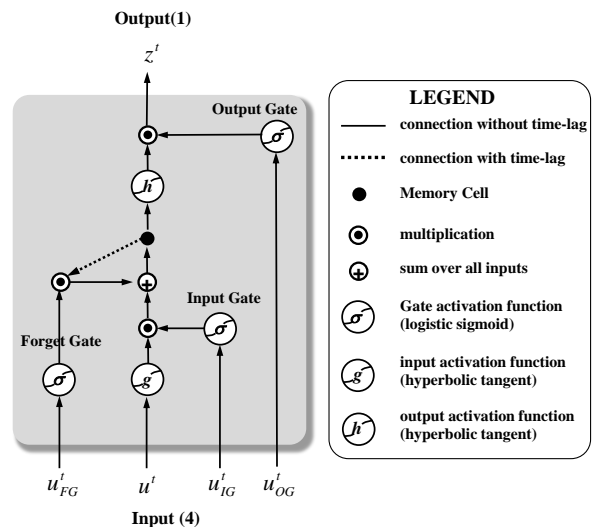


Fig.4 LSTM block

ルの入出力関係)を学習させ、理論モデルと同様のシステムの構築を図る。

教師データとする変位入力データは、同定される RNN モデルが建物応答を概ね現実的かつ網羅的に模擬できるものとなるように、観測地震動を 1 質点系モデルに入力した場合の変位応答履歴とする。観測地震動は、防災科学研究所が公開している K-net, KiK-net 観測網の観測記録とし、観測日が 1996 年～2015 年、最大加速度が 100gal～600gal の範囲内の合計 8069 波を選定した。この地震動を周期 0.1, 0.2, 0.5, 1.0, 2.0, 5.0, 10.0 秒の 7 つの 1 自由度モデルに入力して、合計  $8069 \times 7 = 56483$  波の変位波形を作成する。変位波形はその最大値と最小値が 1 から -1 の範囲(塑性率 5 以内)に収まるよう係数倍し、

さらにこの変位波形を Fig.2 に示す理論モデルに入力して、力の時刻歴波形を算出し (Fig.6 参照), 入力変位波形と併せて教師データとする。

RNN の同定は、与えられた教師データ ( $x_n, d_n$ ) とシステムからの出力 ( $y_n$ ) から計算される誤差関数 ( $E(\mathbf{w})$ ) を最小にすることで行われる。ここでは、学習に用いる誤差関数は、一般的な MSE (誤差二乗平均) を用いる。

$$E(\mathbf{w}) = \frac{1}{N} \sum_{n=1}^N \|d_n - y(x_n; \mathbf{w})\|^2 \quad (3)$$

学習係数は、初期値を 1.0 とし、10000 回の重みの更新 (1epoch) が済むごとに学習係数を 1.03 で割って、徐々に学習係数を小さくする方式とする。そのため、150 Epoch の学習が済む頃には、学習係数が 0.00014 程度まで低くなる。なお、この学習係数は別途実施した 1000 回以上の数値実験を経て試行錯誤の結果辿り着いた経験上の最適値である。

誤差関数の勾配は RNN で一般的に用いられる BPTT 法 [Werbos 1990] を用い、RNN を時間方向に展開し、順伝播型に書き換えて誤差逆伝播計算を行う。

求めた誤差関数の勾配によるシステムの更新は、全訓練サンプルに対して計算される誤差関数から求める方法と、訓練サンプルの一部に対して計算される誤差関数から求める方法 (所謂 mini batch) があるが、本検討においては後者の確率的勾配降下法 (SGD) [Robbins 1951] を用いる。

なお、本検討において、誤差は 25 時刻分の入出力データから求める。これは、誤差の逆伝播は、それ以前の全ての時刻までさかのぼることはせずに、25 時刻前までで打ち切ることに相当する。また、mini batch は 50 に設定する。即ち 1 回のシステム更新で用いる訓練サンプル群は  $25 \times 50 = 1250$  Step の教師データである。学習係数同様、切り時刻や mini batch のハイパーパラメータの設定値についても、数値実験を経て経験的に得た最適値である。試行錯誤の方法によって学習効果の高い設定値を採用した。

なお、SGD においては、選択したサンプル群の特徴による過剰な学習 (重み変更) により重み値がゼロや無限になってしまうのを防ぐために、勾配に上限値を設けることがあり、本検討においては勾配ノルムの上限値を 0.5 とし、学習があまり早く進まないようにしている。具体的には、勾配ノルムが 0.5 を超える場合には、ノルムが 0.5 になるように勾配ベクトルを係数倍する。

併せて、学習時にネットワークの自由度を強制的に小さくし、過適合を防ぐために、中間層 ( $L=2$ ) から出力層 ( $L=3$ ) において、50% のドロップアウトを適用する [Srivastava 2014]。この手法は、多層ネットワークのニューロンを確率的に選別して学習する方法である。具体的には中間層のニューロンの 50% をランダムに選出し、それ以外を存在しないものとして学習することになる。なお、学習終了後同定されたモデルにおいては、中間層の出力は一律に 50% を乗じる。

#### 4. 同定されたモデルの精度

同定された RNN モデルの精度は、設計で用いられる地震動を 1 質点系モデルに入力した時の変位応答履歴を同定された RNN モデルに入力した場合の出力を、Fig.5 の理論モデルの出力と比較することによって行う。確認誤差計算に用いる変位波形は、Fig.7 に示す ElcentroNS 波の地震動に対する、 $T=1.0$ (s) の 1 質点系モデルの応答変位波形 (最大変位を 1m に正規化したもの) とする。

Fig.8 には RNN モデルの学習曲線を示す。縦軸は二乗平均誤差を、横軸は学習データ数 (サンプル群の数) を表している。

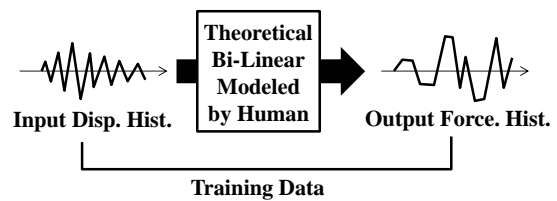


Fig.5 Training Data made by Theoretical Bi-linear Model

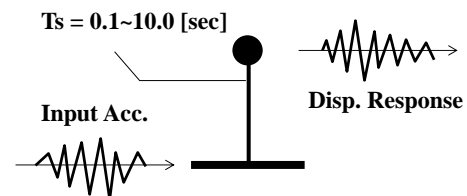


Fig.6 Input Disp. History Data from Single Node Dynamic

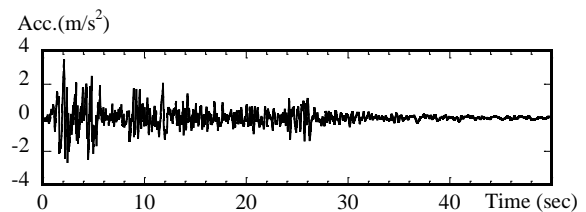


Fig.7 Time History of Input Acceleration (ElcentroNS)

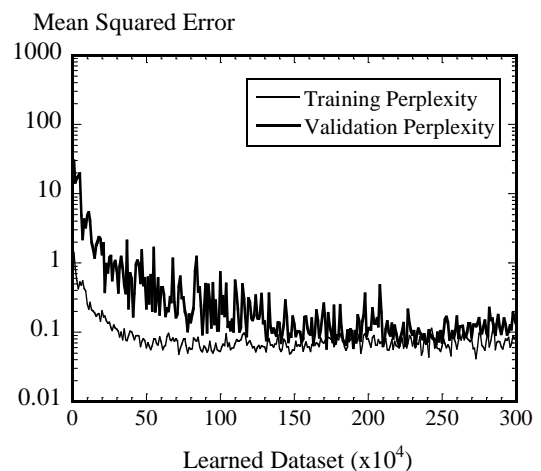


Fig.8 Learning Curve

細実線が訓練誤差、太実線は確認誤差である。図より、学習回数が増大に伴い誤差は減少し、 $300 \times 10^4$  set 程度で概ね収束しており過適合も生じていない。また、訓練誤差と確認誤差を比較すると、訓練誤差の方が小さくなっており、一般的な訓練誤差と確認誤差の関係と同様の結果となった。

つづいて、学習数  $N_L=50 \times 10^4$ ,  $270 \times 10^4$  set のタイミングでテストを行った際の RNN モデルの出力を、理論モデルの出力と比較する。Fig.9 は出力された力の時刻歴波形である。

まず、学習が比較的浅く、確認誤差が収束しきっていない  $N_L=50 \times 10^4$  の場合を比較すると、主要な傾向は捉えられているが主要動に含まれる 1000step 前後、および振動が収まりつつある 3500step 以降で理論モデルと RNN モデルに差異が現れている。

一方、更に学習を進め、確認誤差も概ね収斂したと考えられる  $N_L=270 \times 10^4$  の場合においては、かなりの精度で理論モデルの出力を再現できていることがわかる。

Fig.10 には荷重変位関係を示す。時刻歴波形と同様学習が進むと理論モデルと RNN モデルの差異は小さくなっており、学習の浅い  $N_L=50 \times 10^4$  の場合でもバイリニアの履歴を概ね良好に捉えられていることがわかる。なお、いずれのモデルにおいても、両者の差異は折れ点周りで生じる傾向がある。理論モデルの鋭角な折れ線を RNN モデルでは表現しきれず、やや曲率をもって降伏状態に移行している。これは学習が進んだ  $N_L=270 \times 10^4$  のモデルでも同様である。このように RNN モデルで鋭角な折れ線が表現できないのは、RNN モデルの中間層の活性化関数がシグモイド関数や双曲線正接関数であり、折れ点の鋭角部の変化を表現しづらいためと考えられる。中間層の一部で活性化関数を ReLU 等のランプ関数[Glorot 2011]に置き換えることによって対応できる可能性があるといえる。一方、現状のままであっても、バイリニアモデルの鋭角部の変化は概ね捉えられており、活性化関数にランプ関数を用いなくても十分な学習を行うことで、鋭角部についても一定程度の精度は確保できるといえる。

## 5. まとめ

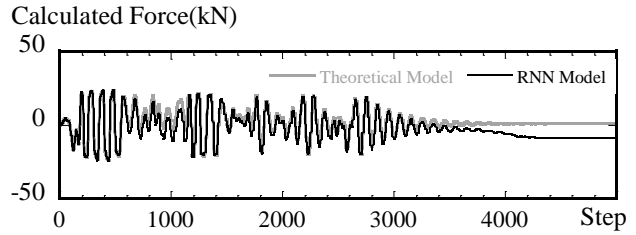
本報においては、LSTM ユニットを用いた 3 層の RNN モデルを構築し、理論モデルの入出力結果を教師データとして基本的な力学モデルであるバイリニアモデルの同定を行った。その結果、同定された RNN モデルは、単純な入力(変位)と出力(力)の履歴から、バイリニアの理論モデルの挙動を良好な精度で同定できることを確認した。ただし、この精度は学習経験のある範囲内でのみ確認されており、適用範囲には十分な注意が必要である。

## 謝辞

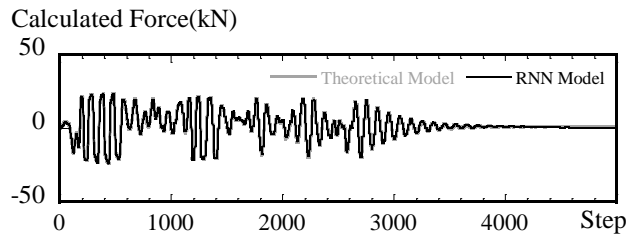
千葉大学大学院工学研究科荒井幸代教授には、人工知能分野の技術全般について多くの貴重な御助言並びに御指導を頂きました。教師データの作成に用いた観測地震動は、防災科学研究所が公開している強震観測網 K-NET, KiK-net の観測記録を利用しました。関係各位に厚く御礼申し上げます。

## 参考文献

- [Glorot 2011] X. Glorot, A. Bordes and Y. Bengio : Deep Sparse Rectifier Neural Networks, Proc. of AISTATS-11, 15, pp.315-323, 2011
- [Hochreiter 1997] S. Hochreiter and J. Schmidhuber : Long short-term memory, Neural Computation, 9(8), pp.1735-1780, 1997
- [Robbins 1951] H. Robbins and S. Monro: A stochastic approximation method, Ann. Math. Statist., 22 (3), pp.400-407, 1951
- [Srivastava 2014] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov: Dropout : A simple way to prevent neural networks from over fitting, Journal of Machine Learning Research, 15, pp.1929-1958, 2014
- [Werbos 1990] P.Werbos: Backpropagation through Time: What it does and how to do it, Proc. of the IEEE, 78(10), pp.1550-1560, 1990

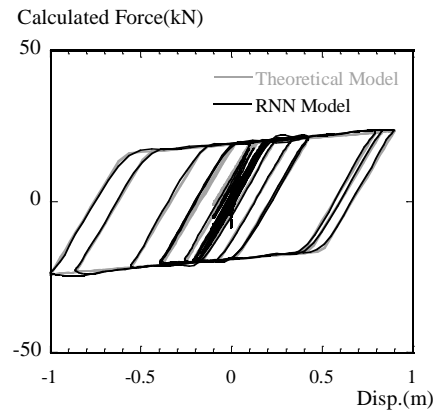


(a)  $N_L=50 \times 10^4$  set

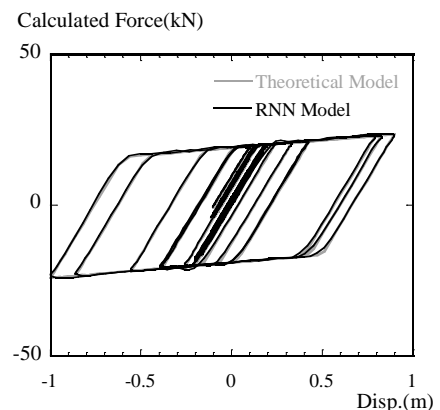


(b)  $N_L= 270 \times 10^4$  set

Fig.9 Comparison of Output Force History between RNN and Theoretical bi-linear



(a)  $N_L=50 \times 10^4$  set



(b)  $N_L= 270 \times 10^4$  set

Fig.10 Comparison of force-disp. relationship between RNN and Theoretical bi-linear