

双方向型計算様式に基づいた神経回路モデルによる時系列予測： Mackey-Glass カオス系列を用いた獲得信号変換の評価

和久屋 寛^{**} ・ 信太 克規^{**}

Time Series Prediction with a Neural Network Model Based on Bi-directional Computation Style: A Study on Acquired Signal Transformation with Mackey-Glass Chaotic Time Series

By

Hiroshi WAKUYA and Katsunori SHIDA

Abstract : Most studies on time series prediction have been used *uni-directional* computation flow to predict future signals from present signals. On the contrary, a *bi-directional* computing architecture, consisting of two mutually connected subnetworks for future and past prediction, is adopted in this article and applied to several prediction tasks. Since the coupling effects between the two subsystems promote the model's ability, bi-directionalization of the computing architecture makes it possible to improve its performance. Furthermore, in order to investigate the acquired signal transformation, one of the famous chaotic time series proposed by Mackey and Glass is adopted. As a result of computer simulations, it has been found experimentally that the future-past information integration gives the bi-directional model an advantage over the uni-directional one.

Key words : bi-directional computation style, bi-directional neural network model, time series prediction, Mackey-Glass chaotic time series, signal transformation

1 まえがき

時系列予測課題とは、過去から現在に至る一連の情報から未来の値を予測する課題である。これまで予測精度の向上を目指して多くの研究者が取り組んでいるが、その結果、線形回帰モデルなどの線形情報処理よりも神経回路網に代表される非線形情報処理が有効であるとの報告が為されている。これは、時系列データがしばしば有する非線形性に負うところが大きい⁽¹⁾。

ところで、従来の時系列予測に関する研究は、現在までの情報から未来の値を予測するという順時間方向の信号変換が主流であった。神経回路網を用いた研究

についても、ネットワーク構造と学習アルゴリズムの改良に主眼をおいたものが大部分であった⁽²⁾⁻⁽⁵⁾。しかし、近年、これとは異なる発想から逆時間方向の信号変換系を別途用意し、順時間方向の信号変換系との相互作用でモデル全体の時系列予測の精度向上を図る手法（以下、双方向型計算様式と呼ぶ）が提案されている⁽⁶⁾。この手法は、これまでに太陽黒点数の年次変化の予測⁽⁶⁾⁽⁷⁾、非対称な多次元井戸型ポテンシャル内に存在する小物体の運動軌跡の予測⁽⁸⁾⁽⁹⁾などの課題に適用されて良好な成績を得ているが、その処理メカニズムに関する言及はほとんど為されていない。そこで本論文では、先行研究とは異なるタイプの時系列データであるカオス系列を採用し⁽¹⁰⁾、その概周期性に着目した獲得信号変換の評価を試みる。

以下本論文では、**2.**で双方向型計算様式に基づいた時系列予測の概要について紹介し、**3.**で Mackey-Glass

平成14年5月1日受理

* 大学院工学系研究科生体機能システム制御工学専攻

** 電気電子工学科

©佐賀大学理工学部

系列を用いた時系列予測の計算機シミュレーションについて述べる。そして4.で双方向型モデルの時系列処理メカニズムについて考察を行い、最後に5.で結論を述べる。

2 双方向型計算様式に基づいた時系列予測

2.1 基本方針

ここで述べる双方向型計算様式では、順逆変換に相当する2つの信号処理系を想定し、相互結合を介して両者の情報統合を図ることで、単独の場合よりも信号処理能力が改善されることを期待している。これまで、時系列の発生と認識を同時に取り扱うモデルとして提案された双方向型神経回路モデル⁽¹¹⁾では、時系列発生を担う処理系と時系列認識を担う処理系が双方向型の信号変換を扱うことで、その時系列学習能力が著しく改善されるとの報告が為されている。例えば、日常生活において単語を覚えるときなど、我々は声に出してその単語を読み、その結果を耳で聞いて確認することが多いが、このように発声系や聴覚系などの感覚運動系を巧みに組み合わせて利用することと類似している。また、このアイデアを拡張した運動制御系と感覚受容系の統合型神経回路モデル⁽¹²⁾でも、同様の学習能力の改善が認められている。

そこで本論文では、双方向型神経回路モデルに時系列予測課題を適用することを考える。まず、2つの信号処理系に次のような信号変換を割り当てる。

- 現在までの時系列情報から、“未来”のある1点の値を予測する（未来予測課題）。
- 現在までの時系列情報から、“過去”のある1点の値を予測する（過去予測課題）。

以下では、割り当てる信号変換の内容に基づいて、それぞれ未来予測系あるいは過去予測系と呼ぶ。このとき、未来予測系で学習する信号変換は従来の時系列予測課題と同一内容であるのに対して、過去予測系で学習する信号変換は本研究で新たに導入したものである。

2.2 双方向型神経回路モデル

図1に双方向型神経回路モデルの構造を示す⁽¹¹⁾。この中で、丸印は内部結合を持たない単一のニューロン層、矢印は隣接するニューロン層間の結合荷重を表す。また、上半分の信号処理系が未来予測系、下半分の信号処理系が過去予測系に相当し、それぞれ4層からなる階層型ネットワークで構成されている。ここで、未来予測系の各層の出力信号をそれぞれ $y_i^{[0]}$, $y_i^{[1]}$, $y_i^{[2]}$, $y_i^{[3]}$ ($i = 1, 2, \dots, n_l$) と表し、状態遷移則を以下のよ

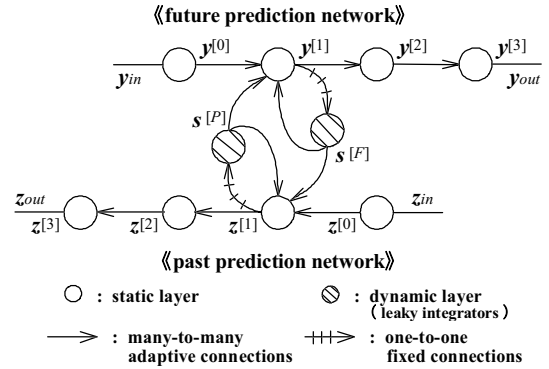


Fig.1 A bi-directional neural network model.

うに定義する。ただし、 n_l ($l = 1, 2, 3, 4$) は第 l 層のニューロン数である。

$$y_i^{[0]} = [y_{in}]_i \quad (1)$$

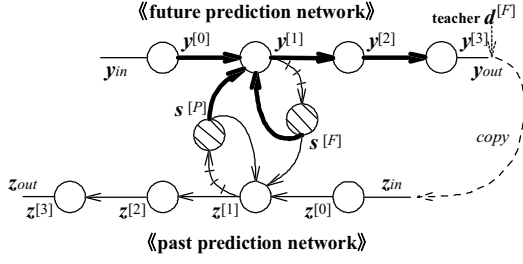
$$\begin{aligned} y_i^{[1]} &= f_1(u^{[y1]}) \\ &= f_1\left(\sum_j w_{ij}^{[1]} y_j^{[0]} + \sum_j w_{ij}^{[F]} s_j^{[F]} + \sum_j w_{ij}^{[P]} s_j^{[P]}\right) \end{aligned} \quad (2)$$

$$\tau \frac{ds_j^{[F]}}{dt} + s_j^{[F]} = y_j^{[1]} \quad (3)$$

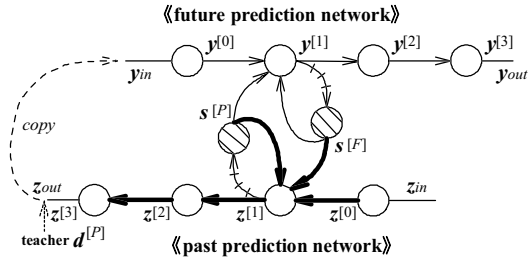
$$y_i^{[2]} = f_2(u_i^{[y2]}) = f_2\left(\sum_j w_{ij}^{[2]} y_j^{[1]}\right) \quad (4)$$

$$[y_{out}]_i = y_i^{[3]} = f_3(u^{[y3]}) = f_3\left(\sum_j w_{ij}^{[3]} y_j^{[2]}\right) \quad (5)$$

ここで、 y_{in} と y_{out} はそれぞれ未来予測系の入力信号と出力信号を表し、 τ は1次遅れ特性を有する動的なニューロンの時定数である。この動的なニューロンは、過去の時系列情報を“内部状態”という形で保持する役割を担っており、その結果を未来予測系と過去予測系の両方にフィードバックすることで、所望の信号変換の学習に利用している。換言すると、式(2)からも明らかなように、 $s^{[F]}$ と $s^{[P]}$ という未来予測系と過去予測系の内部状態を基に両者の情報統合を図ることで、未来予測系の信号変換の学習に利用している。なお、第1層は単なるバッファ、第2層と第3層は通常のシグモイド型の非線形変換細胞層、第4層は $[0, 1.0]$ という範囲の出力値制限を取り除くための線形変換細胞層である。



(a) training phase for the future prediction system
— future prediction mode —



(b) training phase for the past prediction system
— past prediction mode —

Fig.2 A training procedure of the bi-directional model. All connections depicted by thick arrows are modified in each training mode.

$$f_1(x) = f_2(x) = \frac{1}{1 + \exp(-x)} \quad (6)$$

$$f_3(x) = x \quad (7)$$

過去予測系についても、各ニューロン層の出力信号を $z_i^{[0]}$, $z_i^{[1]}$, $z_i^{[2]}$, $z_i^{[3]}$ ($i = 1, 2, \dots, n_l$) とおいて、同様の状態遷移則を定義する。

$$z_i^{[0]} = [z_{in}]_i \quad (8)$$

$$\begin{aligned} z_i^{[1]} &= f_1(u^{[z1]}) \\ &= f_1\left(\sum_j v_{ij}^{[1]} z_j^{[0]} + \sum_j v_{ij}^{[P]} s_j^{[P]} + \sum_j v_{ij}^{[F]} s_j^{[F]}\right) \end{aligned} \quad (9)$$

$$\tau \frac{ds_j^{[P]}}{dt} + s_j^{[P]} = z_j^{[1]} \quad (10)$$

$$z_i^{[2]} = f_2(u_i^{[z2]}) = f_2\left(\sum_j v_{ij}^{[2]} z_j^{[1]}\right) \quad (11)$$

$$[z_{out}]_i = z_i^{[3]} = f_3(u^{[z3]}) = f_3\left(\sum_j v_{ij}^{[3]} z_j^{[2]}\right) \quad (12)$$

次に、双方向型神経回路モデルの学習方法について述べる。上述のように、双方向型モデルは大別して未来予測系と過去予測系という2つの信号処理系から構成されている。そこで、各々の信号処理系に限定した局所的な学習フェーズ — i) 未来予測系の学習フェーズ, ii) 過去予測系の学習フェーズ — を仮定し、これを繰り返すことでモデル全体の学習を実現する (図2)。まず、未来予測系の誤差 e_f を

$$e_f = \sum_t \sum_i \left\{ [y_{out}(t)]_i - d_i^{[F]}(t) \right\}^2 \quad (13)$$

と定義する。ここで、 $d^{[F]}$ は未来予測系の教師信号、 t は教師信号を課す時刻である。次に、この誤差 e_f が減少するように結合荷重 $W^{[3]}$, $W^{[2]}$, $W^{[1]}$, $W^{[F]}$, $W^{[P]}$ を誤差逆伝搬学習 (real-time recurrent learning: RTRL)⁽¹³⁾によって更新する。

$$\Delta w_{\mu\xi}^{[3]} = -\eta_f \frac{\partial e_f}{\partial u_{\mu}^{[y3]}} y_{\xi}^{[2]} \quad (14)$$

$$\Delta w_{\mu\xi}^{[2]} = -\eta_f \frac{\partial e_f}{\partial u_{\mu}^{[y2]}} y_{\xi}^{[1]} \quad (15)$$

$$\begin{aligned} \Delta w_{\mu\xi}^{[1]} &= -\eta_f \left\{ \frac{\partial e_f}{\partial u_{\mu}^{[y1]}} y_{\xi}^{[0]} + \sum_i \frac{\partial e_f}{\partial u_i^{[y1]}} w_{i\mu}^{[F]} \right. \\ &\quad \left. \times \left(1 + \tau \frac{d}{dt}\right)^{-1} [f_1'(u_{\mu}^{[y1]}) y_{\xi}^{[0]}] \right\} \end{aligned} \quad (16)$$

$$\begin{aligned} \Delta w_{\mu\xi}^{[F]} &= -\eta_f \left\{ \frac{\partial e_f}{\partial u_{\mu}^{[y1]}} s_{\xi}^{[F]} + \sum_i \frac{\partial e_f}{\partial u_i^{[y1]}} w_{i\mu}^{[F]} \right. \\ &\quad \left. \times \left(1 + \tau \frac{d}{dt}\right)^{-1} [f_1'(u_{\mu}^{[y1]}) s_{\xi}^{[F]}] \right\} \end{aligned} \quad (17)$$

$$\Delta w_{\mu\xi}^{[P]} = -\eta_f \frac{\partial e_f}{\partial u_{\mu}^{[y1]}} s_{\xi}^{[P]} \quad (18)$$

ただし、 η_f は学習速度係数、 $(1 + \tau \frac{d}{dt})^{-1}$ は1次遅れを表す演算子である。ここで、双方向型モデル内のフィードバック結合を循環する誤差は1周分のみ[†]を考慮し、簡単のため2周目以降は省略する。過去予測系についても、誤差 e_p を

$$e_p = \sum_t \sum_i \left\{ [z_{out}(t)]_i - d_i^{[P]}(t) \right\}^2 \quad (19)$$

[†]図2(a)の $y^{[1]} \rightarrow s^{[F]} \rightarrow y^{[1]}$ と図2(b)の $z^{[1]} \rightarrow s^{[P]} \rightarrow z^{[1]}$ という経路を通るものが相当する。

と定義して同様の学習を実施する。以下では、これら2種類の局所的な学習の際の動作モードを、それぞれ未来予測モードあるいは過去予測モードと呼ぶ。

なお学習期間中には、2.1で述べた発話時の音声フィードバックに相当する経路を用意し、未来予測系と過去予測系の間で協調的な信号変換が形成されやすい状況を整える。具体的には、未来予測モードでは未来予測系の出力 y_{out} を過去予測系の入力 z_{in} にフィードバックし (図2(a)の破線状の矢印)、過去予測モードでは過去予測系の出力 z_{out} を未来予測系の入力 y_{in} にフィードバックする (図2(b)の破線状の矢印) という操作を行う。

3 時系列予測に関する計算機シミュレーション

時系列予測を扱った研究では、1.で述べたように種々の構造を有する神経回路モデルや学習アルゴリズムの提案が大部分であり、その性能評価についても実に様々な時系列データが利用されている。それらの中には太陽黒点データ(2)(5)-(7)やサンタフェ時系列予測コンテスト[†]で提供されたデータ(1)(3)-(5)(8)(9)、Mackey-Glass系列(10)(14)などのように、頻繁に用いられるものも存在する。本論文では、非周期ではあるが概周期的な特徴をもつカオス系列に着目し、その中から Mackey-Glass 系列を取り上げる。

3.1 訓練用時系列データ

Michael C. Mackey と Leon Glass という2人の研究者の名前を冠した時系列データは、1977年に発表した彼らの論文(15)に端を発している。ヒトの健康状態を表す観測量 (例えば、血中 CO_2 濃度、白血球数など) は、ある種の病気にかかった場合に安定状態 (定常点や周期運動) から逸脱するという不安定現象を示すことがある。この論文の本来の目的は、そのような挙動を記述する数式モデルの提案にあったが、そのカオス状の振る舞いから時系列予測にしばしば利用されている。

まず、時間 t の関数である観測量を $X(t)$ とおいて、一般的な1階微分方程式

$$\frac{dX(t)}{dt} = \lambda - \gamma X(t) \quad (20)$$

を考える。もし、時間 t を十分大きくすれば、

$$\lim_{t \rightarrow \infty} X(t) = \frac{\lambda}{\gamma} \quad (21)$$

[†]1992年7月14日から17日までの4日間にわたって、米国ニューメキシコ州サンタフェで開催されたコンテスト。詳細は文献(1)を参照。

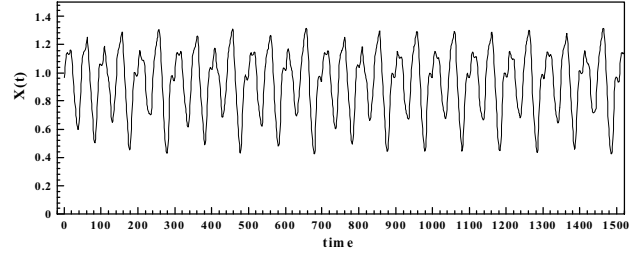


Fig.3 A part of the Mackey-Glass chaotic time series $X(t)$.

となって安定点に収束することは明らかである。しかし、通常は λ と γ のいずれも一定値ではないため、ここでは単純化して次の2つの場合を仮定する。

(1) λ は $X(t - t_0)$ の関数

(2) γ は $X(t - t_0)$ の関数

ただし、 t_0 は遅延時間である。まず (1) のとき、遅延時間 t_0 を大きくすると周期的な振動現象を生じるようになる。また、(2) についても

(2a) γ は $X(t - t_0)$ に対して単調減少の関数

(2b) γ は $X(t - t_0)$ に対して単峰性の関数

を考えると、いずれも遅延時間 t_0 を大きくすると周期的な振動現象を生じるようになるが、(2b) については、遅延時間 t_0 をさらに大きくすると非周期的な振動現象、すなわちカオス状態になるというものである。この時系列データの一部を図3に示す。

3.2 学習課題と学習方法

本論文では、Mackey-Glass 系列の最初の数周期分に相当する550点を用いて以下のように学習する。

- **未来予測系**： 入力信号として $X(t)$ を $y_{in}(t')$ に印加し、教師信号として $X(t+a)$ を $y_{out}(t')$ に与える。
- **過去予測系**： 入力信号として $X(t+a)$ を $z_{in}(t')$ に印加し、教師信号として $X(t)$ を $z_{out}(t')$ に与える。

ここで、 t は時系列データに含まれる数値の通し番号 (index number)、 t' は計算機シミュレーションの時刻のステップ数 (time step) とする。また、 $X(t)$ は第 t 番目の数値データ、 a は予測点までの時間幅である。本論文で取り扱う時系列データは単一の情報のため、入出力ニューロンもそれぞれ1個用意し、1つずつ順番に数値データを印加しながら対応する予測値を順

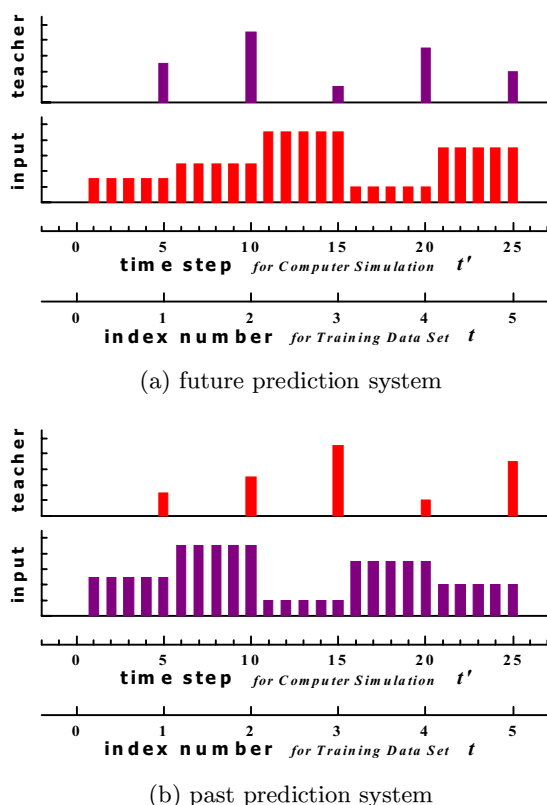


Fig.4 An example of input data and teacher data. — single-index prediction task —

次取り出すという構成にする。ただし、動的なニューロンが急激な出力変化を処理できないために、入力信号の場合は個々の数値データは5ステップ間継続して印加し、教師信号の場合は第5ステップのみに課すという手段を講じる (図4参照)。このとき、 $\Delta t' = 5\Delta t$ という関係が成立し、上述の“点 (index number)”と“ステップ (time step)”は同義語となるので、以下ではおもに後者のステップという表現のみを用いる。

その他のパラメータは以下のように設定する。まず、各隠れ層のニューロン数はすべて9個とおく。動的なニューロンの時定数はシミュレーションの時間刻みの10倍 ($\tau = 10t'$) とするが、この値を小さくすると過去の履歴情報を保持できず、逆に大きくすると新しい情報を取り込むことができないために成績が悪化する傾向にある。結合荷重の初期値は $[-1.0, 1.0]$ の範囲で一様に分布する乱数で決定し、10種類の異なるものを用意する (試行No.1~10)。学習は初期段階で過度の結合荷重の更新が為されないように、学習速度係数を $\eta_f = \eta_p = 0.001$ という小さな値に定め、誤差 e_f, e_p がともに1.0未満になるまで一括処理 (batch training) を繰り返した。また20,000回学習した後でも誤差が収

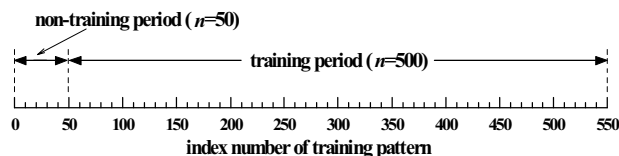


Fig.5 Schematic diagram of a training pattern.

束しない場合は、“学習不能”としてそれ以降の学習を打ち切る。

なお、双方向型モデルへの入力パターンの呈示は3つの部分に分けて行う (図5参照)。最初は長さ500ステップの無入力区間であり、モデルを状態空間の定常点に収束させることで初期状態に依存した先験情報を排除している。次が250ステップ (50点) の非学習区間であり、訓練用データをモデルに与えるが学習は行わない区間を用意することで、入力信号を呈示した直後の不連続性を吸収している。そして最後が長さ2500ステップ (500点) にわたる学習区間であり、実際に訓練用データをモデルに与えて学習を行っている。

3.3 双方向型神経回路モデルの時系列予測能力

3.3.1 双方向型モデルの学習能力

時系列予測課題を取り扱う場合、予測点までの時間幅という観点から分類すると、1時刻先を予測する課題 (single-index prediction : SIP) と数時刻先を予測する課題 (multi-index prediction : MIP) に大別できる。また後者については、直接数時刻先を予測するもの (direct MIP) と1時刻先を予測する手続きを繰り返すもの (iterated SIP) がある。一般に、予測時間幅が大きくなるほど精度よい予測を実現するための難度は増大するが、本論文では双方向型モデルの時系列処理能力を検討するという目的から敢えて direct MIP を採用する。

表1は、予測点を1時刻先から5時刻先まで変化させたときの計算機シミュレーションの結果である。この表では、双方向型計算様式の性能を従来の計算様式 (以下、一方向型計算様式と呼ぶ) と比較するため、双方向型モデルを構成する未来予測系と過去予測系を分割し、それぞれ単独で学習させた場合の結果も併記している。また、神経回路網においては隠れニューロン数が時系列処理能力と密接な関わりがあるため、その総数が双方向型モデルと同じになる未来予測系も用意した。これらの結果から以下のことがわかる。

- (1) 過去予測系のみの場合、いずれの予測時間幅でもすべての試行で学習が完了している。これは予測すべき値が、数時刻前の入力信号として事前に

Table 1 Results of computer simulations with the Mackey-Glass chaotic time series.

Prediction index width	Bi-directional model			Uni-direct. model		Uni-direct. model		Uni-direct. model	
	future pred. system	past pred. system	(rate)	future pred. system	(rate)	future pred. system (II)	(rate)	past pred. system	(rate)
1	0.884	0.474	(¹⁰ / ₁₀)	0.999	(¹⁰ / ₁₀)	0.999	(¹⁰ / ₁₀)	0.992	(¹⁰ / ₁₀)
2	0.989	0.620	(¹⁰ / ₁₀)	1.000	(¹⁰ / ₁₀)	1.036	(⁹ / ₁₀)	0.997	(¹⁰ / ₁₀)
3	1.040	0.563	(⁶ / ₁₀)	1.103	(¹ / ₁₀)	1.185	(³ / ₁₀)	1.000	(¹⁰ / ₁₀)
4	1.997	0.700	(⁰ / ₁₀)	2.062	(⁰ / ₁₀)	1.992	(⁰ / ₁₀)	1.000	(¹⁰ / ₁₀)
5	2.929	0.774	(⁰ / ₁₀)	2.969	(⁰ / ₁₀)	2.880	(⁰ / ₁₀)	1.000	(¹⁰ / ₁₀)

Note : “Future prediction system (II)” in the seventh column indicates the future prediction system with 18 neurons in each hidden layer.

与えられているためと考えられる。

- (2) 1~2時刻先を予測する未来予測課題では、いずれのモデルでも大部分の試行で学習が完了している。
- (3) 4~5時刻先を予測する未来予測課題では、いずれのモデルでも学習が全くできていない。
- (4) 3時刻先を予測する未来予測課題では、双方向型モデルでは半数以上が学習できている (⁶/₁₀) のに対して、未来予測系のみではほとんど学習できていない (¹/₁₀, ³/₁₀)。

要するに、本来の目的である未来予測課題のみに着目すると、予測時間幅が3の場合に限って、双方向型モデルと未来予測系のみとの間に成績の優劣が認められる。

そこで上記5つの課題について、その入出力関係をグラフに描いたものが図6である。ここで横軸は入力信号 $X(t)$ 、縦軸は所望の出力信号 $X(t+a)$ ($a = 1, 2, 3, 4, 5$) を表す。一般に、カオス系列は概周期的な特徴をもつため、その入出力関係の時間変化を表す軌道はある特定の形状を示す。図6によると、Mackey-Glass 系列では右肩上がりの楕円形に近いものになっている。このとき、図6(a)(b)のような細長い場合はほぼ入力のみ依存するため学習が比較的簡単であるが、図6(d)(e)のように徐々に幅を持つようになると系の構造が複雑となり、過去の履歴情報が必要となるなど学習が難しくなることがわかる。よって境界付近に相当する図6(c)のときに、双方向型モデルと未来予測系のみとの間で時系列学習能力の優劣が顕在化すると考えられる。以下では、これらの点を勘案して、予測時間幅が3のときに限定した議論を行う。

ところで表1によると、未来予測課題を20,000回学習したときの誤差 e_f の平均値は、双方向型モデルが

1.040であるのに対して、未来予測系のみでは1.103、隠れニューロンが18個の未来予測系では1.185となっている。すなわち、提案手法の方が成績がよい。これを定量的に評価するため、両者の比で定義した予測精度改善指数 (index for improvement quality: IIQ) という指標を導入する。これを今回の計算機シミュレーションの結果に適用すると、

$$IIQ_{f1} = \frac{[e_f]_{\text{bi-direct.model}}}{[e_f]_{\text{future pred.sys.}}} = \frac{1.040}{1.103} \approx 0.943 < 1.0 \quad (22)$$

となり、従来の一方型よりも双方向型の計算様式の方が有効であることがわかる。また、各隠れ層のニューロンを同数とした場合も、

$$IIQ_{f2} = \frac{[e_f]_{\text{bi-direct.model}}}{[e_f]_{\text{future pred.sys.(II)}}} = \frac{1.040}{1.185} \approx 0.878 < 1.0 \quad (23)$$

となり、やはり双方向型計算様式の方が有効であることが確認できる。

3.3.2 双方向型モデルの応答

双方向型モデルの働きを解明するためには、時系列信号処理を行う際の振る舞いを検討することが重要である。まず最初に、3時刻先を予測する課題の学習結果を図7に示す。書式は図6と同様で、図7(a)は訓練用データ ($51 \leq t \leq 550$)、図7(b)(c)は未来予測系と過去予測系で学習によって獲得された信号変換 ($\mathbf{y}_{in} \rightarrow \mathbf{y}_{out}$, $\mathbf{z}_{in} \rightarrow \mathbf{z}_{out}$) の一例である。理想的には、図7(b)(c)が図7(a)の形状に一致すればよい[†]のであるが、実際に

[†]ここでは、比較が簡単のように図7(c)の横軸と縦軸を入れ替えているので、図7(b)と(c)の軌道が類似していればよい。

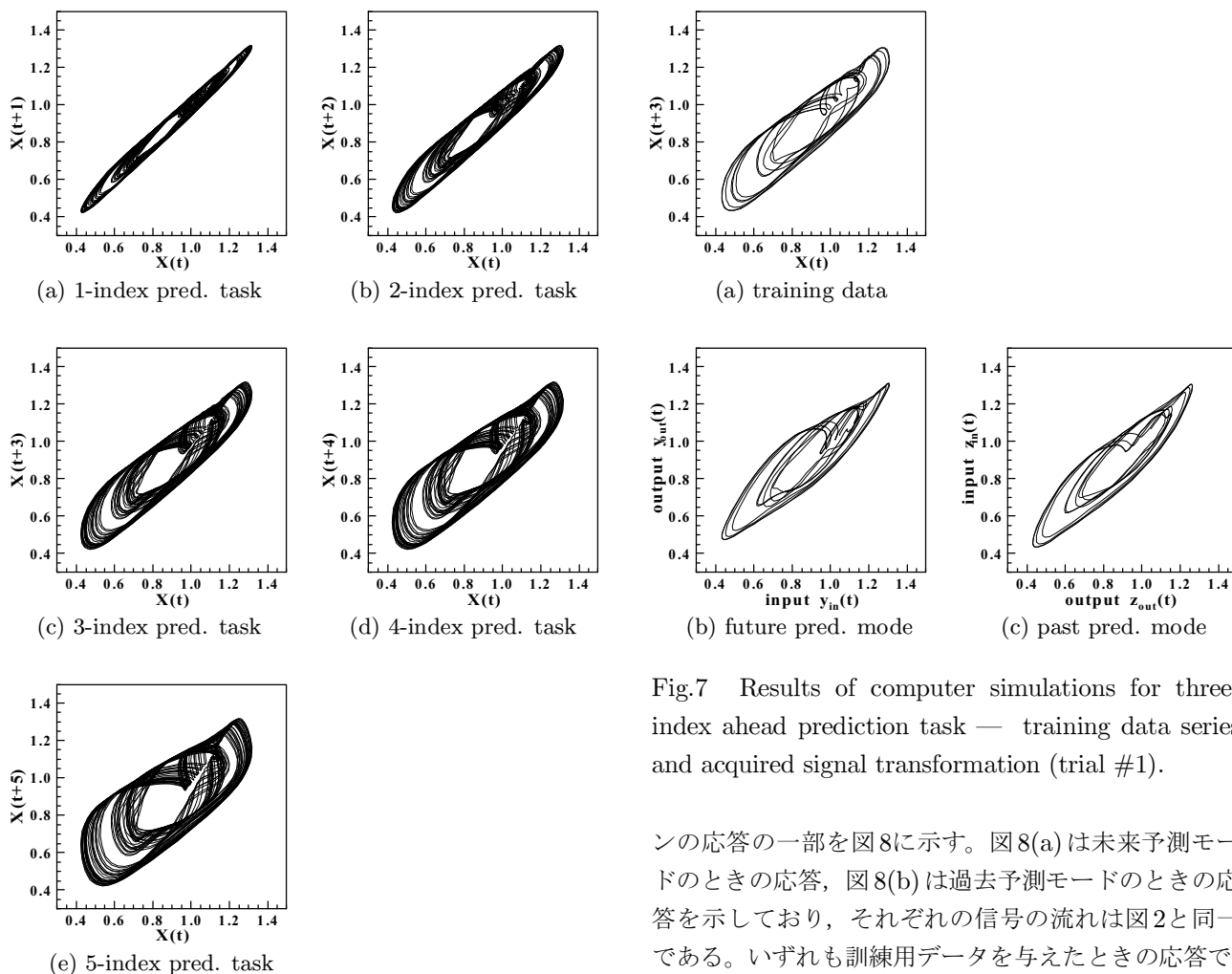


Fig.6 Input-output relationship of the future prediction tasks (prediction index width: 1-5).

はいくつか異なる点が認められる。例えば、双方向型モデルでは右上と左下の角の部分はやや尖っているが、ここは訓練用データの特徴を完全には獲得できていないことを意味しており、誤差の原因となっている。しかし、楕円軌道の概形や2周してほぼ元に戻ることで、その途中で軌道が交差する点の位置、右上の部分で軌道が一部捻じれていることなどの特徴は、学習によってほぼ獲得されている。よって、双方向型モデルは未来予測モードと過去予測モードで所望の順逆変換をほぼ求めることができたと言ってもよいであろう。未来予測系や過去予測系を単独で学習させた場合も、入出力関係の時間変化を描いた軌道は双方向型モデルと類似していたが、そこで認められる差異が学習能力の優劣として現れていることになる。

次に、このときの双方向型モデルの入出力ニューロ

Fig.7 Results of computer simulations for three-index ahead prediction task — training data series and acquired signal transformation (trial #1).

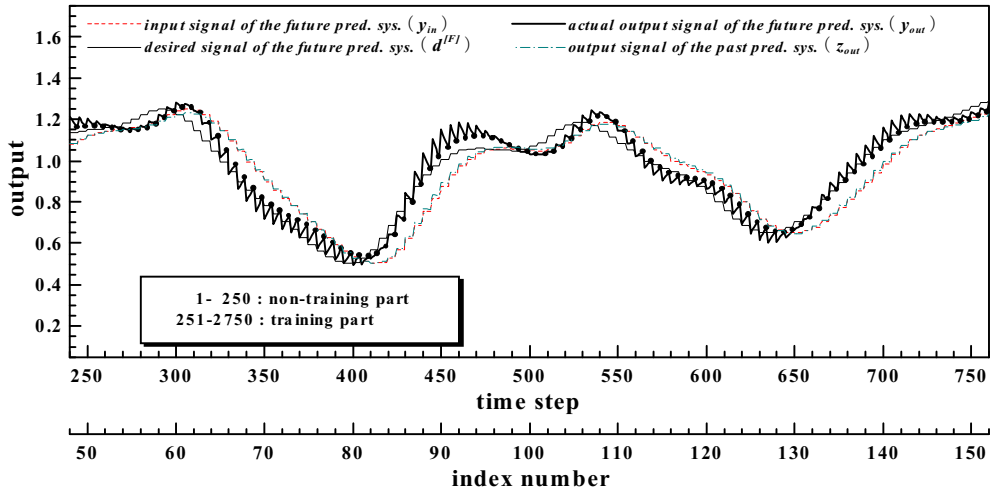
ンの応答の一部を図8に示す。図8(a)は未来予測モードのときの応答、図8(b)は過去予測モードのときの応答を示しており、それぞれの信号の流れは図2と同一である。いずれも訓練用データを与えたときの応答で、図中の黒丸の部分に教師信号を課した点を表している。なお3.2で述べたように、横軸の“index number”は時系列データの数値の通し番号 t に、“time step”は計算機シミュレーションの時間刻み t' にそれぞれ対応している。以下簡単のために、未来予測系で行われる信号変換、式(1)~(5)をまとめて

$$\mathbf{y}_{out} = \Phi(\mathbf{y}_{in}, \mathbf{s}^{[F]}, \mathbf{s}^{[P]}) \quad (24)$$

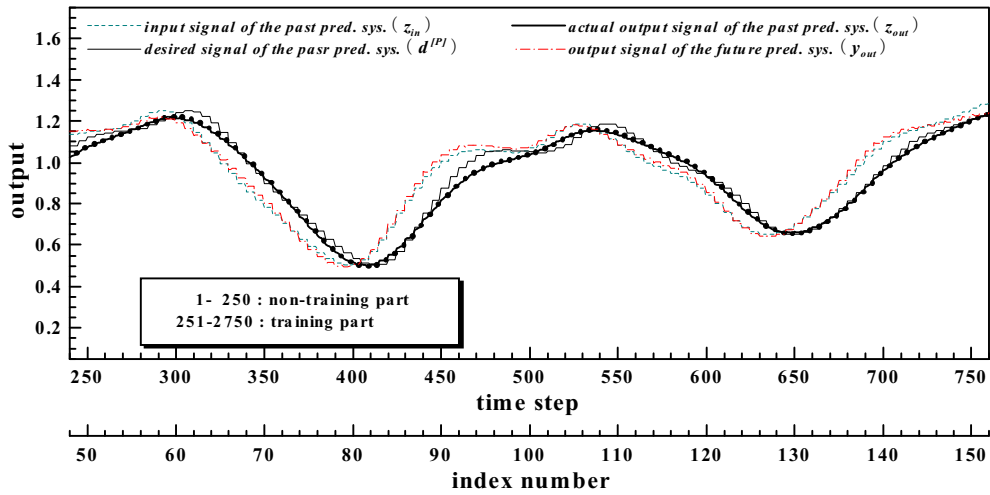
とおき、同様に過去予測系で行われる信号変換、式(8)~(12)をまとめて

$$\mathbf{z}_{out} = \Psi(\mathbf{z}_{in}, \mathbf{s}^{[F]}, \mathbf{s}^{[P]}) \quad (25)$$

と記述する。ここで $\mathbf{s}^{[F]}$, $\mathbf{s}^{[P]}$ は未来予測系と過去予測系が共有しているそれぞれの内部状態である。まず、図8(a)の未来予測モードを注視すると、過去予測系の出力 \mathbf{z}_{out} と未来予測系の入力 \mathbf{y}_{in} の応答が非常に似ていることがわかる ($\mathbf{z}_{out} \approx \mathbf{y}_{in}$)。図2(a)にも示すように、本来、未来予測系の学習フェーズでは、所望の出力 $\mathbf{d}^{[F]}$ が得られるように結合荷重 $W^{[3]}$, $W^{[2]}$, $W^{[1]}$, $W^{[F]}$, $W^{[P]}$ を更新して変換関数 Φ を求めている



(a) responses in the future prediction system



(b) responses in the past prediction system

Fig.8 Responses of the input and the output neurons in the bi-directional neural network model (trial #1) after 4,948-epoch training.

るだけで、過去予測系の出力信号 z_{out} を未来予測系の入力信号 y_{in} に近づける操作は全く行っていない。それにもかかわらず、実際は非常によく似た応答を示している。これは、未来予測系の信号変換 Φ と過去予測系の信号変換 Ψ が順逆変換の関係にあり、図7(b)(c)でも確認できるように実際にこれを満足する信号変換が形成されていることと関連している。すなわち、①未来予測系に y_{in} を与えると信号変換 Φ を施されて y_{out} が得られ、②これを直ちに過去予測系に帰還して ($z_{in} = y_{out}$)、③逆変換 Ψ を行って z_{out} が求まっている。よって、最終的には過去予測系の出力 z_{out} と未来予測系の入力 y_{in} の応答が類似すると考えられる。

図8(b)の過去予測モードについても、これと逆の現象 ($y_{out} \approx z_{in}$) が確認できるが、そのメカニズムも同様であり、入力 z_{in} に順変換 Ψ を施してその結果に逆変換 Φ を行うというものである。以上のことから、未来予測モードと過去予測モードという動作モードの違いに関係なく、双方向型モデルは時系列予測の際に順逆変換の関係にある2つの信号変換 Φ , Ψ を並行して同時に実行していることがわかる。

ところで、双方向型計算様式の最大の特徴は、付加的な課題を実行する過去予測系の存在である。よって、そこで学習される逆時間方向の信号変換 Ψ は学習能力の優劣に関連があると考えられる。すなわち、

- (i) もし過去予測系の出力 z_{out} が未来予測系の入力 y_{in} に似ていれば、未来予測系と過去予測系で確かに順逆変換の関係にある信号変換を実行している。
- (ii) (i) を満たす場合、未来予測系と過去予測系の間で相互結合を介した協調的な信号変換を実現できる可能性がある。
- (iii) (ii) で利用できる内部状態 $s^{[F]}$, $s^{[P]}$ は未来予測系あるいは過去予測系の学習フェーズ (図2参照) で独立に形成することができる。

というものである。本来の課題である未来予測を行う場合を考えると、 $y_{in} \rightarrow y_{out} \rightarrow z_{in} \rightarrow z_{out}$ という経路が双方向型モデルの学習過程で重要となることは間違いない。また学習する課題は異なるものの、上記 (iii) で言及したように、途中の経路で独立に形成した非線形な信号変換を利用することが重要であることも示唆されている⁽¹⁶⁾。以上のことから、未来-過去情報の統合機構は双方向型計算様式を用いた場合の予測精度向上に重要な役割を担っていると考えられるが、詳細については未だ不明な点も多く、さらなる検討が必要である。

3.3.3 双方向型モデルの予測精度

学習終了後の汎化能力の評価については、式(13)(19)で表される誤差 e を時系列データの分散 σ^2 と長さ T で規格化した平均相対変動値 (average relative variance: ARV) という指標を採用する⁽²⁾。

$$ARV = \frac{e}{\sigma^2 T} \quad (26)$$

この ARV を用いて評価した結果をまとめたものが図9である。ここで、評価用データの長さは、図5に示す訓練用データと同一の550点 (非学習区間50点、学習区間500点) とし、Mackey-Glass 系列のデータの中から開始点を50点ずつ遅らせて作成した。なお図中の横軸は、この評価用データを切り出したときの開始点を表している。ここで最左端の点は訓練用データに相当し、550以降の点はすべて各モデルにとって未学習の新規データに属する。いずれの評価用データに対しても ARV は平均的にほぼ一定の値を示しているため、訓練用データと同等の予測精度が常に期待でき、優れた汎化能力を有していることがわかる。また個々のモデルの成績に着目すると、双方向型計算様式 (●) の予測精度が一方向型計算様式 (▲, △) と比較して改善していることも確認できる。

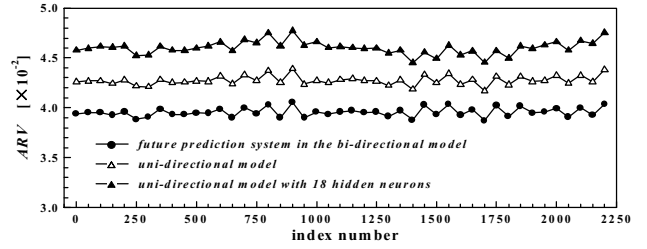


Fig.9 Prediction quality with the trained networks (trial #1) — future prediction mode.

4 考察

これまでに、双方向型計算様式に基づいた時系列予測について、その有効性を様々な観点から検討してきた。その結果、双方向型モデルを用いた場合に予測精度の改善が認められたが、そのメカニズムについては未だ不明な点が多い。本論文でも、未来予測系と過去予測系の学習フェーズを独立に設定し、そこで獲得された信号変換を利用することの重要性を指摘したが、なお検討の余地が残る。

ところで、時系列予測を行う際に重要なことは、式(24)の変換関数 Φ と式(25)の変換関数 Ψ が何をしているかである。双方向型モデルをオートマトンのように考えると、前者は入力 y_{in} と内部状態 $s^{[F]}$, $s^{[P]}$ から出力 y_{out} を求めることと等価であり、後者は入力 z_{in} と内部状態 $s^{[F]}$, $s^{[P]}$ から出力 z_{out} を求めることと等価である。一般に、双方向型モデルと一方向型モデルにおける未来予測課題の信号処理を比較した場合、 y_{in} は共通であり、 $s^{[F]}$ はともに未来予測系の内部状態なので内容は類似している。したがって、 $s^{[P]}$ としてどのような内部状態を表現できるかが、両者の性能の優劣を決定する。ここで、 $s^{[P]}$ は過去予測系の内部状態であり、式(25)の変換関数 Ψ によって独立に求まるものである。したがって、双方向型計算様式と一方向型計算様式との時系列予測能力の比較を行う場合、 $s^{[F]}$ とは別に $s^{[P]}$ を用意することと、 $s^{[F]}$ の要素数を単に増加させることを比べて、同一の入力信号に対してどちらが多様な内部状態を形成できるかという問題に帰着する。主成分分析を用いた解析によれば前者 (双方向型計算様式) の方が内部状態を表現するために必要な主成分の数が多く、多様性も増していると考えられるが、この件についてはより詳細な検討が必要であろう。

5 むすび

本論文では, 双方向型計算様式に基づいた神経回路モデルの時系列予測課題への適用について検討した。カオス状に変化する時系列を用いた計算機シミュレーションの結果, 双方向型計算様式の方が一方向型計算様式よりも学習能力の優れていることを確認した。また汎化能力も有しており, 予測精度も改善された。このとき, 未来予測系と過去予測系は順逆変換の関係にある2つの信号変換を並行して同時に実行しており, その未来-過去情報の統合作用の重要性について指摘した。

謝辞 本研究は, 第1著者(和久屋)が平成10年度文部省在外研究員として米国ケンタッキー州ルイビル大学滞在中に, Jacek M. Zurada 教授と共同で行ったものである。期間中, Zurada 教授ならびに Computational Intelligence Laboratory の諸氏には, いろいろと貴重な意見をいただいた。記して感謝します。

参考文献

- [1] N.A.Gershenfeld and A.S.Weigend : “The future of time series”, *In Time Series Prediction: Forecasting the Future and Understanding the Past* (eds. A.S.Weigend and N.A.Gershenfeld) , pp.1-70, Addison-Wesley, 1993
- [2] A.S.Weigend, B.A.Huberman and D.E.Rumelhart : “Predicting the future: A connectionist approach”, *International Journal of Neural Systems*, **Vol.1**, pp.193-209, 1990
- [3] E.A.Wan : “Time series prediction by using a connectionist network with internal delay lines”, *In Time Series Prediction: Forecasting the Future and Understanding the Past* (eds. A.S.Weigend and N.A.Gershenfeld) , pp.195-217, Addison-Wesley, 1993
- [4] M.C.Mozer : “Neural net architectures for temporal sequence processing”, *In Time Series Prediction: Forecasting the Future and Understanding the Past* (eds. A.S.Weigend and N.A.Gershenfeld) , pp.243-264, Addison-Wesley, 1993
- [5] T.Cholewo and J.M.Zurada : “Sequential network construction for time series prediction”, *Proceedings of the IEEE International Joint Conference on Neural Networks*, June 9-12, Houston, Texas, USA, pp.2034-2039, 1997
- [6] H.Wakuya and J.M.Zurada : “Bi-directional computing architecture for time series prediction”, *Neural Networks*, **Vol.14**, pp.1307-1321, 2001
- [7] H.Wakuya and J.M.Zurada : “Time series prediction by a neural network model based on the bi-directional computation style”, *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, July 24-27, Como, Italy, **Vol.2**, TE3-1, pp.225-230, 2000
- [8] H.Wakuya, K.Shida and J.M.Zurada : “Bi-directionalization of neural computing architecture for time series prediction: Application to computer generated series ‘Data Set D’”, *Proceeding of the 6th International Conference on Soft Computing*, October 1-4, Iizuka, Japan, C-4-2, pp.660-667, 2000
- [9] 和久屋 寛, 信太克規 : “双方向型計算様式に基づいた神経回路モデルによる時系列予測: 計算機生成時系列データ ‘Data Set D’ を用いた汎化能力の検討”, 電子情報通信学会論文誌 (D-II), **Vol.J85-D-II**, pp.101-111, 2002
- [10] H.Wakuya, K.Shida and J.M.Zurada : “Bi-directionalization of neural computing architecture for time series prediction. II: Application to Mackey-Glass chaotic time series”, *Abstract of the International Conference on Dynamical Aspects in Complex Systems from Cells to Brain* , November 29-December 1, Sendai, Japan, PS58, pp.43-44, 2000
- [11] 和久屋 寛, 二見亮弘, 星宮 望 : “時系列の発生と認識を行う双方向型神経回路モデル”, 電子情報通信学会論文誌 (D-II), **Vol.J77-D-II**, pp.236-243, 1994
- [12] 和久屋 寛, 信太克規 : “双方向型神経回路モデルを用いた運動制御と感覚受容の統合モデル”, 電子情報通信学会論文誌 (D-II), **Vol.J80-D-II**, pp.1929-1938, 1997
- [13] R.J.Williams and D.Zipser : “A learning algorithm for continually running fully recurrent neural networks”, *Neural Computation*, **Vol.1**, pp.270-280, 1989
- [14] A.E.Gaweda and J.M.Zurada : “Fuzzy neural network with relational fuzzy rules”, *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, July 24-27, Como, Italy, **Vol.5**, HA1-1, pp.3-7, 2000
- [15] M.C.Mackey and L.Glass : “Oscillation and chaos in physiological control systems”, *Science*, **Vol.197**, pp.287-289, 1977
- [16] 和久屋 寛, 二見亮弘, 星宮 望 : “双方向型神経回路モデルにおける時系列パターン変換機能の形成過程”, 電子情報通信学会技術研究報告, **NC93-31**, 1993