ニューラルネットワークを用いた大気汚染濃度の短期予測[†]

近 藤 正*

Short-Term Prediction of Air Pollution Concentration by a Neural Network

Tadashi KONDO*

A short-term prediction of air pollution concentration by a neural network is described. The neural network used in this paper can identify a nonlinear system whose structure is very large and complex. By using the time series data of the SO₂ concentration in Tokushima, Japan, we intend to find a suitable model for predicting SO₂ concentration a few hours in advance. Eight different prediction models obtained by the neural network are compared to find a suitable neural network structure and suitable input variables in the prediction model. The predicted results obtained by the neural network are compared with the results obtained by linear and nonlinear statistical models, such as a linear regression model, a linear auto-regressive model and a nonlinear GMDH model. It is shown that the neural network developed in this paper gives a better prediction result as compared with the statistical models.

Key Words : neural network, short-term prediction, air pollution concentration, nonlinear system identification, GMDH

1. ま え が き

近年、脳の超並列学習情報処理過程を工学的な手法を用いて研究するニューロコンピューティング(神経回路網計算学)の研究が活発に行われ実際問題への応用も行われている^{1),2)}。ニューラルネットワークを実際問題に適用する場合、複雑な構造をした非線形システムを対象にすると、ニューラルネットワークの中に含まれる必要とされるニューロン(神経細胞)の個数が非常に多くなり、その結果ネットワークの構造が複雑になる。このような実際問題への応用を目的とした複雑な構造をしたニューラルネットワークを構成する場合、まず始めに問題となることは、ニューラルネットワークの構造をどのように決定すれば汎化能力が大きく予測精度の良いニューラルネットワークを得ることができるのかということである。階層構造をしたニューラルネットワークを対象にす

る場合、階層構造を何層にすれば良いのか、また、各層におけるニューロンの個数をいくつにすれば良いのかといったことが問題となり、このようなことが適切に決定されなければ良い予測精度を得ることができない。ニューラルネットワークに含まれるニューロンの個数を多くし、ニューラルネットワークの規模を大きくすると、学習に用いたデータに対する精度は向上するが、ニューロンの個数を必要以上に多くして、ニューラルネットワークの構造が過度に複雑化してくると、未学習データに対する予測精度は逆に悪くなり汎化能力は小さくなる³⁾。ニューラルネットワークの実際問題への応用においては、適切な規模のネットワーク構造を見つけ出して、未学習データに対する予測精度を向上させ、大きな汎化能力を得ることが、応用上非常に重要な問題となる。特に、対象分野においてすでに使用されている予測手法がある場合、それらの手法と比較してニューラルネットワークの予測精度が向上して汎化能力が優れていなければ、その分野へのニューラルネットワークの応用上の有効性をみいだすことは難しい。

本研究では、複雑な構造をした非線形システムの同定を目的としたニューラルネットワーク⁴⁾を用いて、このニューラルネットワークの大気汚染濃度の短期予測問題

[†] 第34回システム制御情報学会研究発表講演会で発表(1990・5)

* 徳島大学医療技術短期大学部 徳島市蔵本町 3-18-15

* School of Medical Sciences, The University of Tokushima, Tokushima
(Received April 24, 1992)

(Revised December 14, 1992)

への有効性について明らかにする。大気汚染濃度の短期予測では、従来から観測データを統計的に解析することにより作成する統計モデルがよく用いられており、線形構造をした自己回帰モデルや重回帰モデル、さらに非線形構造をした GMDH モデルなどを用いた予測モデルが研究されている^{4),5)}。ここでは、複雑な構造をした非線形システムの同定を目的としたニューラルネットワークを用いて、大気汚染濃度の短期予測モデルを構成する。そして、従来から使用されている統計モデルによって得られた短期予測結果と比較することにより、本研究で用いたニューラルネットワークでは予測精度が大幅に改善されることを示して、このニューラルネットワークの大気汚染濃度の短期予測問題への有効性を示す。

2. ニューラルネットワークを用いた大気汚染濃度の短期予測モデル

ここでは、ニューラルネットワークを用いて同定する大気汚染濃度の短期予測モデルについて述べる。大気汚染濃度の短期予測モデルの入出力変数としては、OX, NO, NO₂, CO, SO₂ などの濃度と風向、風速などが考えられるが、ここでは、SO₂ 濃度に注目して、SO₂ 濃度の 1~3 時間先の短期予測をニューラルネットワークを用いて行う。以下に、大気汚染濃度の短期予測モデルの構造について具体的に検討する。時刻 t の SO₂ 濃度を変数 $x(t)$ により表わす。SO₂ 濃度の 3 時間先の値は $x(t+3/t)$ により表わす。

まず始めに、ニューラルネットワークの出力変数について検討する。ニューラルネットワークの出力変数としては、3 時間先の SO₂ 濃度の値 $x(t+3/t)$ を直接用いることが考えられる。しかしながら、変数 $x(t+3/t)$ をニューラルネットワークの出力変数として用いると、変数 $x(t+3/t)$ の変動幅が非常に大きいためニューラルネットワークの学習回数をかなり多く必要とする。さらに、変数 $x(t+3/t)$ をニューラルネットワークの出力変数に直接用いた場合、つぎのような問題が生じる。学習に用いるデータの個数に制限があり学習が十分に行えない場合や、学習に用いた測定データの汚染濃度の値よりもかなり高い汚染濃度の予測を行う場合、また、逆にかなり低い汚染濃度の予測を行う場合などでは、まだ学習が十分に行われていないためにニューラルネットワークの出力値が初期値の近傍にあり真値から大きく異なって予測精度が極端に悪くなることがある。実際問題を取り扱う場合では、このようないろいろな理由により、学習が十分に行えない場合が生じるが、このようなときに予測モデルの精度が大幅に悪くなると信頼性がなくなり同定されたモデルは実際問題に適用できなくなる。しかし

ながら、ニューラルネットワークの出力変数として 3 時間先の SO₂ 濃度の値 $x(t+3/t)$ を直接用いるのではなく、現在時刻の SO₂ 濃度からの 3 時間先の SO₂ 濃度の変化量 $\Delta x(t+3/t)$ を用いて、この値から 3 時間先の SO₂ 濃度を予測するようにすると、前述のような問題がある程度克服できる。ここで、ニューラルネットワークの出力として SO₂ 濃度の 3 時間先の変化量 $\Delta x(t+3/t)$ を用いる場合には、 $\Delta x(t+3/t)$ の初期値がゼロに近い値を取るように、ニューラルネットワークの結合荷重の初期値を設定する。ニューラルネットワークの出力として SO₂ 濃度の 3 時間先の変化量 $\Delta x(t+3/t)$ を用いると、まず、 $x(t+3/t)$ と比較して変数の変動幅が非常に小さいために、ニューラルネットワークの学習回数を大幅に減少させることができる。さらに、学習に用いるデータの個数に制限があり学習を十分に行うことができずに、ニューラルネットワークの出力 $\Delta x(t+3/t)$ の値が初期値の近傍にあり真値から大きく異なるような場合でも、 $\Delta x(t+3/t)$ が初期値のゼロに近い値をとることは、3 時間先の SO₂ 濃度が現在の SO₂ 濃度からほとんど変化しないことと同じ現象となるため、極端な予測精度の悪化を避けることができる。このために、ニューラルネットワークの出力変数としては、 $x(t+3/t)$ を用いて直接に 3 時間先の SO₂ 濃度を予測するよりも、 $\Delta x(t+3/t)$ を用いて予測するほうがより望ましいと考える。このような理由により、本研究では、ニューラルネットワークの出力変数として $\Delta x(t+3/t)$ を用いる。

つぎに、ニューラルネットワークの入力変数について検討する。入力変数としては、現在時刻から τ 時間過去までの SO₂ 濃度 $x(t), x(t-1), \dots, x(t-\tau)$ と時刻 t を用いる場合と、現在時刻から τ 時間過去までの SO₂ 濃度の 1 時間ごとの変化量 $\Delta x(t), \Delta x(t-1), \dots, \Delta x(t-\tau)$ と時刻 t を用いる場合などが考えられる。時刻 t は SO₂ 濃度に 24 時間の周期性があるために入力変数として考慮する。そのほかの入力変数としては、SO₂ 濃度以外の大気汚染物質の濃度や、風向、風速などが考えられるが、ニューラルネットワークでは、入力変数を増加させるとネットワークの規模(ニューロンの個数など)が急激に大きくなり、それによって小型計算機などでは学習のための計算が困難になる。本研究では、小型計算機(処理速度: 15 MIPS, 主メモリ(ユーザーズエリア): 8 M バイト, MPU: RISC 型)を用いて処理可能な範囲内でニューラルネットワークの学習を行う。このために、SO₂ 濃度の値と時刻 t のみを入力変数として用い、つぎに示す予測モデルをニューラルネットワークを用いて同定する。

○予測モデル(1) ($K=2$)

$$\Delta x(t+3/t) = f_1(\Delta x(t), \Delta x(t-1), \Delta x(t-2),$$

$$\Delta x(t-3), t) \quad (1)$$

○予測モデル(2) ($K=20$)

$$\Delta x(t+3/t) = f_2(\Delta x(t), \Delta x(t-1), \Delta x(t-2), \Delta x(t-3), t) \quad (2)$$

○予測モデル(3) ($K=2$)

$$\Delta x(t+3/t) = f_3(x(t), x(t-1), x(t-2), x(t-3), t) \quad (3)$$

○予測モデル(4) ($K=20$)

$$\Delta x(t+3/t) = f_4(x(t), x(t-1), x(t-2), x(t-3), t) \quad (4)$$

ここで、 K はニューラルネットワークの内部において連続データを離散データに量子化するとき用いるパラメータを示しており、このパラメータを量子化係数と呼ぶ。本研究で用いる複雑な構造をした非線形システムを同定するニューラルネットワーク⁶⁾では、量子化係数(K)の値を決めることにより、ニューラルネットワークの構造を比較的容易に決定することができる。 K の具体的な内容については次章のニューラルネットワークの概略を参照されたい。

つぎに、時刻 t が予測モデルの精度にどの程度影響を及ぼすのかを調べるために、入力変数に時刻 t を含まない場合の予測モデルを同様にニューラルネットワークを用いて同定する。

○予測モデル(1)' ($K=2$)

$$\Delta x(t+3/t) = f_1'(\Delta x(t), \Delta x(t-1), \Delta x(t-2), \Delta x(t-3)) \quad (5)$$

○予測モデル(2)' ($K=20$)

$$\Delta x(t+3/t) = f_2'(\Delta x(t), \Delta x(t-1), \Delta x(t-2), \Delta x(t-3)) \quad (6)$$

○予測モデル(3)' ($K=2$)

$$\Delta x(t+3/t) = f_3'(x(t), x(t-1), x(t-2), x(t-3)) \quad (7)$$

○予測モデル(4)' ($K=20$)

$$\Delta x(t+3/t) = f_4'(x(t), x(t-1), x(t-2), x(t-3)) \quad (8)$$

以上の考察により、本研究では大気汚染の短期予測モデルとして、(1)式から(8)式までの8種類の予測モデルをニューラルネットワークを用いて同定し、おのおのの予測モデルの精度を比較することにより、8種類の予測モデルの中からより望ましい短期予測モデルの構造を明らかにする。

3. 複雑な構造をした非線形システムの同定を目的としたニューラルネットワーク⁶⁾

階層型ニューラルネットワークの学習方式に関する理論的研究はバックプロパゲーション法を中心にかなり詳

しく研究されているが^{1),2)}、実際問題を取り扱えるような大規模ニューラルネットワークの構造同定に関する研究は、まだ、これから活発に行われるものと思われる。本研究では、複雑な構造をした非線形システムを同定できるニューラルネットワーク⁶⁾を用いて大気汚染の短期予測を行う。このニューラルネットワークは非線形システムから得られる入出力データを用いて学習を繰り返すことにより、対象となる非線形システムの特徴をネットワークの内部に記憶保存する機能をもっている。このような複雑な構造をした非線形システムの同定を目的とするニューラルネットワークを構成する場合、対象となる非線形システムの構造の複雑さを十分に考慮して、入出力データに対して十分な汎化機能を備えたネットワーク構造を構成することが重要である。つまり、過度にニューラルネットワークの構造を複雑化しても未学習データに対して良い予測精度を得ることができないために、ニューラルネットワークの階層構造を何層にすればよいのか、さらに、各層におけるニューロンの適切な個数をどのように見つけ出すのかといったニューラルネットワークの構造決定に関する問題が良い予測精度を得るために重要な問題となる。本研究で用いるニューラルネットワークは、小脳皮質内の情報処理機構を数学モデル化したCMAC(Cerebellar Model Arithmetic Computer)⁷⁾の考え方をを用いて、これを多層構造化することにより、階層型ニューラルネットワークの構造を比較的容易に決定している。このニューラルネットワークのアルゴリズムの詳細については、すでに文献6)に報告されているが、このニューラルネットワークを実際問題へ応用する場合には、ニューラルネットワークの構造決定方法やニューラルネットワークに含まれる設計パラメータの設定方法などが、実用上十分な予測精度を得るために非常に重要な問題になるために、以下では、このような点を中心にニューラルネットワークの概略について簡単に述べる。

3.1 ニューラルネットワークの概略

本研究で用いたニューラルネットワークは4層から構成されており、以下に各層の入出力関係を示す。

〈第1層〉

$$u_j^1 = x_j \quad (j=1, 2, \dots, p) \quad (9)$$

ただし、 $x_j (j=1, 2, \dots, p)$ はシステムの入力変数、 u_j^1 は第1層における j 番目のニューロンの出力値を示す。

〈第2層〉

$$z_h^2 = \sum_i w_{ih}^{12} \cdot u_i^1 + \sum_j \sum_j w_{jh}^{12} \cdot u_j^1 \cdot u_j^1 - w_{0h}^{12} \cdot \theta_0^1 \quad (10)$$

$$u_h^2 = 1 / (1 + \exp(-\alpha(|z_h^2| - K/2))) \quad (11)$$

Fig. 1 Structure of the neural network

ただし、 $\theta_0^1=1$ 、 K は量子化係数、 w_{ik}^{12} 、 w_{ik}^{13} および w_{ik}^{14} は第1層と第2層のニューロン間の結合荷重を示す。 α は量子化に用いる釣鐘状関数の形状を決めるパラメータを示す。

〈第3層〉

$$z_j^3 = \sum_i w_{ij}^{23} \cdot u_i^2 - w_{0j}^{23} \cdot \theta_0^2 \quad (12)$$

$$u_j^3 = 1 / (1 + \exp(-z_j^3)) \quad (13)$$

ただし、 $\theta_0^2=1$ である。

〈第4層〉

$$\phi_j = \sum_i w_{ij}^{34} \cdot u_i^3 \quad (j=1, 2, \dots, q) \quad (14)$$

第2層では入力変数の非線形結合とその量子化を行う。第3層における結合荷重 w_{ij}^{23} 、 w_{0j}^{23} と第4層における結合荷重 w_{ij}^{34} はバックプロパゲーション法を用いて学習させる。第2層と第3層におけるニューロン数、さらに第2層における結合荷重 w_{ik}^{12} 、 w_{ik}^{13} 、 w_{ik}^{14} の値および第3層における結合荷重 w_{ij}^{23} 、 w_{0j}^{23} の初期値はCMACの考え方を用いて決定する。第2層における結合荷重 w_{ik}^{12} 、 w_{ik}^{13} 、 w_{ik}^{14} の値は、ニューラルネットワークの構造決定の段階でCMACの考え方を用いて決めた値をそのまま用いて学習は行わない。また、CMACでは、第4層の結合荷重の学習のみを行っているが、ここでは、第4層と、さらに第3層の結合荷重もバックプロパゲーション法を用いて学習させることにより、CMACの考え方を用いて決めた初期値からより望ましい結合荷重の値へ修正する。

Fig. 1 にニューラルネットワークの構造を示す。

3.2 ニューラルネットワークの構造決定方法

まず、ここで使用するニューラルネットワークとCMACとの対応について述べる。ニューラルネットワークの第2層のニューロンは、CMACの苔状線維に信号を与える起始細胞に対応させ、第3層のニューロンはCMACの顆粒細胞に対応させる。さらに、第4層のニューロンはCMACのブルキンエ細胞に対応させる。このように対応付けを行うことにより、以下に示すようにニューラルネットワークの構造を決定できる。

(第2層における構造決定)

ニューラルネットワークの第2層における情報処理は、CMACにおいて入力ベクトルから苔状線維への情報処理に対応する。つまり、(10)式と(11)式を用いて、入力変数とその非線形結合の量子化を行い、第3層の顆粒細胞へ信号を伝達する。このため、第2層のニューロンの個数は、第3層の顆粒細胞へ信号を伝達する苔状線維の起始細胞の個数に一致する。つまり、(11)式の量子化係数 K の値を決めることにより、第2層のニューロンの個数と結合荷重の値を決定することができる。

(第3層における構造決定)

ニューラルネットワークの第3層における情報処理は、CMACの苔状線維から顆粒細胞への情報処理に対応する。つまり、(12)式により第2層のニューロンの出力を複数個組み合わせて、第3層のニューロンの入力とする。そして、(13)式により入力信号の強さがある閾値を越えた場合、出力信号を発生させる。第3層のニューロンの個数は、顆粒細胞の個数に一致する。つまり、第2層のニューロンの出力の組合せ方法を決めることにより、第3層のニューロンの個数と結合荷重の初期値が決まる。ここで、第2層のニューロンの出力の組合せ方法としてはいろいろ考えられるが、CMACにおいて相異なる2組の入力ベクトルを組み合わせるように、ニューラルネットワークの第2層のニューロンの出力を組み合わせると汎化能力が大きくなることがわかっている。

(第4層における構造決定)

ニューラルネットワークの第4層における情報処理は、CMACの顆粒細胞からブルキンエ細胞への情報処理に対応する。つまり、(14)式により第3層のニューロンの出力を複数個組み合わせて重み付き加算した後に出力する。第4層のニューロンの個数は、出力変数の個数に一致する。

以上の手続きにより、複雑な構造をした非線形システムを同定できるニューラルネットワークの構造を決定できる。

4. ニューラルネットワークによる短期予測

複雑な構造をした非線形システムを同定するニューラルネットワークを用いて、 SO_2 濃度の3時間先の短期予測を行う。また、 SO_2 濃度の1時間先の短期予測結果も示す。さらに、ほかの予測モデルとの比較を行うために、従来から短期予測に用いられてきた線形統計モデルおよび改良形GMDHによる非線形統計モデルの短期予測結果を示す。線形統計モデルおよび改良形GMDHによる非線形統計モデルの短期予測結果は、文献5)に報告されているものを引用し、測定データも同じものを使用する。

4.1 ニューラルネットワークによる短期予測結果

測定データは、1975 年 5 月から 6 月に徳島県で測定された大気汚染データを使用する。ニューラルネットワークを用いた学習は 5 月 27 日から 5 月 31 日までの 5 日間の測定データを用いて行い、6 月 1 日から 6 月 15 日までの 15 日間の大気汚染濃度を予測する。本研究で使用するニューラルネットワークでは、量子化係数(K)を変化させることによりネットワーク構造が大きく変化して予測精度が異なってくる。ここでは、量子化係数の値として最も小さい($K=2$)をとる場合と、かなり大きな値($K=20$)をとる場合の二つのケースを考える。さらに、ニューラルネットワークの入力変数の違いにより予測精度がどのように変化するかを調べるために、前述の(1)式から(8)式の予測モデル(1)から予測モデル(4)'を、ニューラルネットワークを用いて同定しその予測結果を検討する。

(1) 量子化係数(K)と入力変数(x , Δx)の影響についての検討

まず始めに、(1)式から(4)式までの予測モデル(1)から予測モデル(4)によって得られた予測結果を調べることにより、入力変数に SO_2 濃度の 1 時間ごとの測定値(x)を用いた場合と、 SO_2 濃度の 1 時間ごとの変化量(Δx)を用いた場合の予測精度の違いについて検討する。さらに、量子化係数(K)を変化させることにより、ニューラルネットワークの構造が予測精度にどの程度影響を及ぼすかについて検討する。Fig. 2 に、予測モデル(1)から予測モデル(4)によって得られた予測精度を示す。ただし、 j 日目の予測精度は

$$J = 1/24 \sum_{k=1}^{24} \{x_j(k) - \hat{x}_j(k/k-3)\}^2 \quad (15)$$

によって評価する。 $\hat{x}_j(k/k-3)$ は 3 時間先の予測値を示

Fig. 2 The prediction error at three hours in advance for the prediction model (1)~(4)

す。図より、予測モデル(1)と予測モデル(2)を比較すると、予測モデル(2)のほうが予測精度が大幅に良くなっている。同様に、予測モデル(3)と予測モデル(4)を比較すると、予測モデル(4)のほうが予測精度が大幅に良くなっている。これらのことから、ニューラルネットワークの構造を決める量子化係数(K)の値としてはより大きな値($K=20$)を用いたほうが予測精度が大幅に良くなることがわかる。つぎに、予測モデル(1)と予測モデル(3)を比較すると、予測モデル(3)のほうが予測モデル(1)よりも良くなっている。同様に、予測モデル(2)と予測モデル(4)を比較すると、最大予測誤差の観点からは予測モデル(2)のほうが良いが、全体的な平均予測精度の観点から見ると予測モデル(4)のほうが予測モデル(2)よりも良くなっている。これらのことから、入力変数としては、 SO_2 濃度の値(x)を使用するほうが、 SO_2 濃度の 1 時間ごとの変化量(Δx)を使用するよりも全体的に予測精度が良くなることがわかる。さらに、予測モデル(1)から予測モデル(4)の四つの予測モデルを比較すると、予測モデル(4)の予測精度が全体的に最も良くなっている。Table 1 に、予測モデル(1)から予測モデル(4)の四つの予測モデルの同定に使われたニューロンの個数を示す。予測モデル(4)は、予測モデル(1)や予測モデル(3)よりもネットワーク内部のニューロンの個数が少ないにもかかわらず、良い予測精度が得られている。このことから、ニューロンの個数を多くしても必ずしも良い予測精度が得られるとは限らないことがわかる。つまり、対象となるシステムの複雑さを考慮して適切な規模のネットワーク構造を構成することが、良い予測精度を得るために重要であると思われる。さらに、Fig. 2 からわかるように、四つの予測モデルの精度はかなり異なっており、このことからニューラルネットワークを用いて予測モデルを同定する場合、良い予測精度を得るためには入力変数の構成や量子化係数(K)の値などをよく検討して適切なものを選択することが必要である。Fig. 3 に予測モデル(4)によって得られた予測値を示す。Fig. 3 より、ニューラルネットワークにより 3 時間先の予測値として全体的に良好な結果が得られていることがわかる。しかしながら、たとえば、6 月 8 日の 10 時から 16 時の間に発生しているような大きな変動は予測モデル(4)により

Table 1 Number of neurons in the prediction models

Layer	Model 1	Model 2	Model 3	Model 4
1	6	6	6	6
2	151	201	161	301
3	4500	800	5120	1800
4	1	1	1	1

Fig. 3 The predicted values at three hours in advance by the prediction model (4)

Fig. 4 The predicted values at one hour in advance by the neural network

予測が十分にできなかった。これは、予測モデル(4)の入力変数として SO_2 濃度と時刻 t のみしか用いていないため、このような大きな変動を起こす要因を入力変数として用いることができれば、このような変動も予測できるようになるものと思われる。また、参考のために予測モデル(4)と同じ入力変数を用いた場合の1時間先の SO_2 濃度の短期予測結果を Fig. 4 に示す。Fig. 4 の結果から、1時間先の短期予測値として全体的に良好な結果が得られていることがわかる。

(2) 入力変数(時刻 t)の有効性についての検討

つぎに、入力変数(時刻 t)の有効性について検討する。Fig. 5 に、予測モデル(1)'から予測モデル(4)'によ

て得られた予測精度を示す。ここでは、前述の Fig. 2 に示した、予測モデル(1)から予測モデル(4)によって得られた予測精度との比較を行う。予測モデル(1)'から予測モデル(4)'と、予測モデル(1)から予測モデル(4)の予測精度をそれぞれ比較すると、入力変数(時刻 t)を予測モデルから取り除くと予測精度が大幅に悪くなることがわかる。特に、予測モデル(1)'と予測モデル(2)'の予測精度が著しく悪くなった。これに対し、予測モデル(4)'と予測モデル(4)の予測精度を比較すると、予測モデル(4)'は予測モデル(4)よりも予測精度が少し悪くなった程度で、予測モデル(4)'の予測精度は15日間の全期間を通してそれほど悪くなっていない。さらに、予

測モデル(1)'から予測モデル(4)'の四つの予測モデルの中では、予測モデル(4)'の予測精度が最も良くなり、変動も少なく安定している。これらのことから、予測モデル(4)や予測モデル(4)'のモデル作成において用いた、量子化係数(K)の値や入力変数(x)の選択がほかの予測モデルと比較して適切であったことがわかる。以上の結果をまとめると、大気汚染濃度の短期予測では、測定データの中に 24 時間の周期性が含まれているために、時刻 t の影響が非常に大きく作用しており、このために、モデリングを行う前に、測定データから 24 時間の周期性を除去するなどの前処理を行わないときには、時

刻 t を入力変数として用いることが必要である。さらに、予測モデルの構造としては、予測モデル(4)'に入力変数(時刻 t)を加えた予測モデル(4)の構造が安定した予測精度を示し最も優れていると思われる。

4.2 統計モデルによる予測結果との比較

ここでは、ニューラルネットワークを用いて同定した予測モデルの有効性を検討するために、従来から大気汚染濃度の短期予測に用いられて来た統計モデルによる予測結果との比較を行う。統計モデルとしては、自己回帰モデルと重回帰モデルの二つの線形モデルを、また、非線形統計モデルとして改良形 GMDH モデルを取り上

Fig. 5 The prediction error at three hours in advance for the prediction model (1)'~(4)'

Fig. 6 The prediction error at three hours in advance for the statistical models and the neural network

Fig. 7 The predicted values at three hours in advance by the auto-regressive model⁴⁾

げる。これらの統計モデルを用いた大気汚染濃度の短期予測方法については文献5)を参照されたい。Fig. 6にこれらの統計モデルとニューラルネットワークによって得られた3時間先の予測精度の比較を示す。ニューラルネットワークについては予測モデル(4)の結果を示す。図より、これらの四つのモデルの中では、ニューラルネットワークによって同定された予測モデルが最も予測精度が良いことがわかる。特に、統計モデルを用いて精度良く予測できなかった6月8日、12日、13日、15日における予測精度の改善が著しい。参考までに、Fig. 7に自己回帰モデルによる3時間先の予測値を示し、前述のFig. 3のニューラルネットワークを用いた3時間先の予測値と比較する。ここで、Fig. 7は文献4)から引用した。Fig. 7より、自己回帰モデルでは、3時間先の予測値は現在の予測値がほぼ3時間先にシフトしたような形になっており3時間先の予測がうまく行われていないことがわかる。これに対して、Fig. 3のニューラルネットワークを用いた予測では、このような現象が非常に小さくなっており、3時間先の予測が精度良く行われていることがわかる。

5. む す び

複雑な構造をした非線形システムの同定を目的としたニューラルネットワークを用いて、大気汚染濃度の短期予測を行った。そして、8種類の大気汚染濃度の短期予測モデルを同定し、その結果を比較することにより明らかになったことを以下にまとめる。なお、以下の文中のニューラルネットワークとは、本研究で用いたニューラルネットワークを示す。

- (1) ニューラルネットワークの構造を決定する量子化係数(K)の値は $K=2$ とした場合よりも $K=20$ としたほうがニューラルネットワークを構成するニューロンの個数が少ないにもかかわらず予測精度が大幅に良くなり汎化能力が大きくなることがわかった。このことより、ニューラルネットワークの各層におけるニューロンの個数は、対象システムの入出力関係の複雑さをよく考慮して決定する必要があり、ニューロンの個数をむやみに多くしても大きな汎化能力が得られないことがわかる。
- (2) 予測モデルの入力変数としては、 SO_2 濃度の値(x)をそのまま使用するほうが、 SO_2 濃度の1時間ごとの変化量(Δx)を使用するよりも予測精度が良くなった。このため、ニューラルネットワークの入力変数としては、 SO_2 濃度の値(x)をそのまま使用するほうが、 SO_2 濃度の1時間ごとの変化量(Δx)を使用するよりもより望ましいと思われる。

れる。

- (3) 入力変数(時刻 t)の有効性については、時刻 t を予測モデルの入力変数から取り除くと予測モデルの予測精度が大幅に悪くなることから、時刻 t が入力変数として重要な影響を及ぼしていることがわかる。大気汚染濃度の短期予測では、測定データの中に24時間の周期性が含まれているために、あらかじめ測定データの前処理によりこのような周期性を除去しない場合には、時刻 t を入力変数として用いることが必要である。
- (4) ニューラルネットワークを用いて得られた予測結果と従来から用いられてきた統計モデルによって得られた予測結果との比較により、ニューラルネットワークの大気汚染濃度の短期予測に対する有効性が明らかになった。特に、統計モデルを用いて精度良く予測できなかった日の予測精度の改善が著しかった。また、統計モデルとの予測値の比較により、自己回帰モデルを用いた予測で発生していた予測値のかなり大きな時間遅れなどがニューラルネットワークでは小さくなっており、従来の統計モデルによる予測値よりも良い予測値が得られることが明らかになった。

以上の結果は、8種類の大気汚染濃度の短期予測モデルの比較から得られたものであるが、本研究で用いたニューラルネットワークが大気汚染濃度の短期予測問題への応用に対して十分に有効であることが確かめられた。

本研究で用いたニューラルネットワークでは、内部で入力変数の量子化を行っており、実際問題へ応用して十分な予測精度を得るためには、この入力変数の量子化をかなり小さく行う必要がある。その結果、入力変数が少ないにもかかわらず、ニューラルネットワークは大規模化する傾向にある。このような理由と、さらに、本研究では小型計算機を用いてニューラルネットワークの学習を実行しているために、入力変数の個数を5個程度に制限した。このために、 SO_2 濃度の過去の値や時刻 t などの入力変数のみを用いては予測困難な変動に対しては、十分な予測ができなかった。また、計算時間に関しては、モデル式の構造を決定する段階で過去5日間の全ての測定データを一度に用いて学習をさせる場合には計算時間が非常に多く必要とされるが、モデル式の構造を決定した後にオンラインで使用する場合、つまり1時間に一つの測定値を学習させるようなオンライン学習情報処理においては、学習に要する計算時間は非常に少なくてすむために、小型計算機を用いても実用上全く問題がないものと思われる。

今後の課題としては、

- (i) 高性能大型計算機を用いてニューラルネットワークの学習を実行し、ほかの重要と思われる因子を入力変数に取り込んで解析を行い、モデル式の構造をさらに充実させる。
- (ii) 応用上十分な汎化能力をもち、しかも、少ない個数のニューロンでニューラルネットワークが構成できるようにニューラルネットワークのアルゴリズムを改良する。
- (iii) 本研究で得られた大気汚染濃度の短期予測結果は、複雑な構造をした非線形システムの同定を目的としたニューラルネットワークを用いて得られたものであるために、ほかのニューラルネットワークを用いても同様な短期予測結果が得られるかどうかを調べて、ニューラルネットワークの大気汚染濃度の短期予測問題への有効性をより明らかにする。

などが考えられ、これらのことを実行することにより、今後、さらに大気汚染濃度の短期予測の精度向上と実用化が期待できる。

大気汚染濃度の短期予測の研究を行うにあたり、徳島文理大学 添田喬学長には、以前に測定データをご提供いただくと同時に有益なご討論を賜った。また、大阪大学基礎工学部 田村担之教授には、終始ご厚情溢るるご指導とご鞭撻を頂いた。ここに、両教授に対して深甚の謝意を表す。

参 考 文 献

- 1) 甘利, 後藤(編): 人工ニューラルシステム, bit 臨時増刊号, 共立出版, 21-11(1989)
- 2) 杉江, ほか(編): ニューロコンピュータ, part II, コンピュートロール, 29, コロナ社(1990)
- 3) 喜多一: ニューラルネットワークの汎化能力, システム制御情報学会誌, 36-10, 625/633(1992)
- 4) Y. Sawaragi, T. Soeda, et al: The Predictions of Air Pollution Levels by Nonphysical Models Based on Kalman Filtering Method, ASME Trans., Series G, 98-4, 375/386(1976)
- 5) 田村, 近藤: 改良形 GMDH による大気汚染のモデリングと短期予測, 計測自動制御学会論文集, 15-5, 622/627(1979)
- 6) 近藤正: 複雑な構造をした非線形システムの同定を目的としたニューラルネットワーク, システム制御情報学会論文誌, 4-7, 259/266(1991)
- 7) J. S. Albus(小杉, ほか訳): ロボティクス, 啓学出版(1984)

[著 者 紹 介]

近 藤 正 (正会員)

1979 年, 大阪大学大学院工学研究科博士課程後期課程修了。同年(株)東芝入社。88 年徳島大学医療技術短期大学部助教授, 現在に至る。システム同定, エキスパートシステム, ニューラルネットワーク, 医療情報システムの研究に従事。システム制御情報学会, 日本機械学会, 日本医療情報学会などの会員(工学博士)。