

降水量予測に適したニューラルネットワーク構造

小泉 耕*・平沢 正信**

要旨

降水量予測に適したニューラルネットワークを構築するために、ネットワーク構造として4層フィードフォワード構造と3層ハイブリッド構造、また、出力層の遷移関数として線型関数とシグモイド関数を用いた4種類のネットワークを作り、降水量予測の問題に適用してスキルスコアを比較した。その結果、3層ハイブリッド構造とシグモイド出力の組み合わせがもっとも良好な性能を示した。

このことにより、ある問題に最適なニューラルネットワーク構造というものは、取り扱うデータの性質によって変わることが示唆される。

1. はじめに

ニューラルネットワークは、今日では帰納的な統計モデルの1つとして多くの分野で用いられるようになり、気象の予測に関わる分野でも、次々と成果が報告されている(最近の報告として Marzban and Stumpf, 1996, 1998; Baik and Hwang, 1998; Koizumi, 1999 など)。

これらの成果を見ると、ニューラルネットワークが、従来この分野で用いられてきた線型モデル(主として重回帰モデル)よりすぐれたものであることは、自明であるかのように思われる。しかしながら、実際に様々の問題にニューラルネットワークを適用してみると、必ずしも重回帰モデルより良い成績が自動的に得られるわけではない、ということに気がつく。

ニューラルネットワークを特定の問題に応用しようとする場合、ネットワーク構造や、出力層における遷移関数の選択については任意性があり、対象とする問題に最も適した構造を発見することは容易でない。そのため、時には不適切な構造を選択したことが原因で所期の性能が得られない、といったことも起こる可能性がある。

本稿では、降水量予測という問題に対して、ネットワーク構造および出力層の遷移関数の異なる4種類のニューラルネットワークを適用して予測精度を比較し、降水量予測に適したネットワーク構造について考察した。

2. 問題設定

予測対象は、関東地方を覆う120個のメッシュ(1メッシュの大きさは約20 km 四方)におけるメッシュ平均降水量の3時間積算値(またはその階級)である(第1図)。

この領域をカバーする数値モデルの格子点値および実況観測値を予測資料として用い、3時間ごとに最新の実況値を用いて12時間先までの予測を行うという問題を設定した。

3. データ

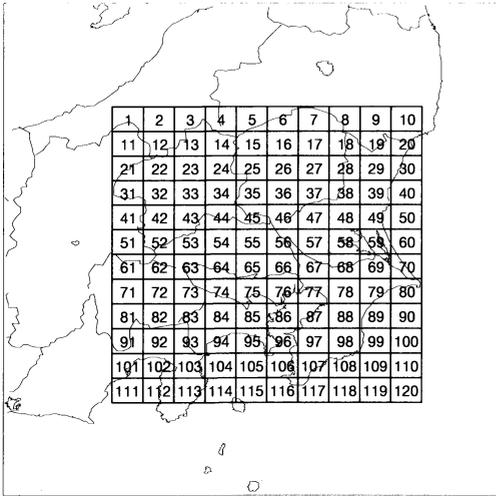
データは1993年1月から1995年12月の3年分を用いた(ただし、1994年については1、2月の資料が欠けている)。このうち1993、1994年の2年間のデータでニューラルネットワークの学習および重回帰式の作成を行い、1995年のデータで精度を検証した。

予測対象の降水量の真値は、分解能5 kmのレーダー・アメダス解析雨量の一時間値を第1図のメッシュごとに平均した後、3時間分積算して求めたもの

* 気象庁予報部数値予報課. ** 気象研究所予報研究部。

—1999年7月7日受領—

—2001年9月14日受理—



第1図 予測対象領域。番号の付いているメッシュそれぞれについて領域平均の3時間降水量（またはその階級）を予測する。各メッシュの大きさは約20 km 四方。

第1表 ニューラルネットワークへの入力要素と、それぞれの要素において用いられた主成分の数。ASM はアジア域、JSM は日本域のスペクトル数値予報モデル。

種別	要素名	主成分の数
ASM (格子点数144)	Qベクトルの発散	19
	相当層厚の勾配の絶対値	23
JSM (格子点数100)	相当層厚による相対湿度	4
	SSI	2
	500 hPa の温度移流	14
	700 hPa の ω	14
	850 hPa の風の東西成分	6
	850 hPa の風の南北成分	6
	900 hPa の水蒸気フラックスの発散	21
降水量の3時間積算値	14	
実況観測値 (アメダスは110の格子点に内挿して使用)	アメダスの風の東西成分	35
	アメダスの風の南北成分	33
	アメダスの風の発散	38
	レーダー・アメダス解析3h雨量	21

である。

ニューラルネットワークへは、第1表に掲げた要素を入力することとした。ただし、格子点値をそのまま用いるのではなく、要素ごとに主成分分析を行ない、各格子点の変動の80%を説明するのに必要なだけの主成分を選んで、その主成分スコアを入力値とした。こ

れは入力変数相互の共線性を除去し、学習後のネットワークの解析を容易にするため、および、入力要素の数を減らしてネットワークサイズおよび計算量を軽減するための操作である。

4. ニューラルネットワークの諸元

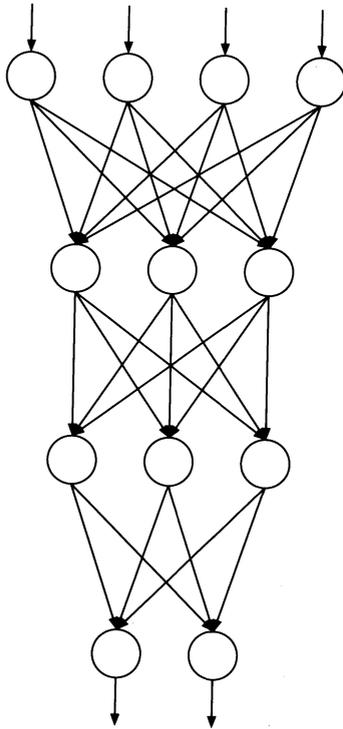
4.1 学習方式

ニューラルネットワークの学習アルゴリズムについては、広く用いられているバックプロパゲーションではなく、Koizumi(1999)と同様に、学習効率の良い共役勾配方向の直線探索を採用した。共役勾配方向の計算には Fletcher and Reeves (1964) の方式を用いている。バックプロパゲーション法とは異なり、共役勾配法においては「逐次学習」を行うことはできない。つまり、教師データ全体について共役勾配方向を求め、その方向にパラメータを変化させることで誤差を減少させる、という操作を繰り返すことになる。繰り返し回数は最大200回としたが、極小点の探索ができなくなった場合にはそこで止めることとした。

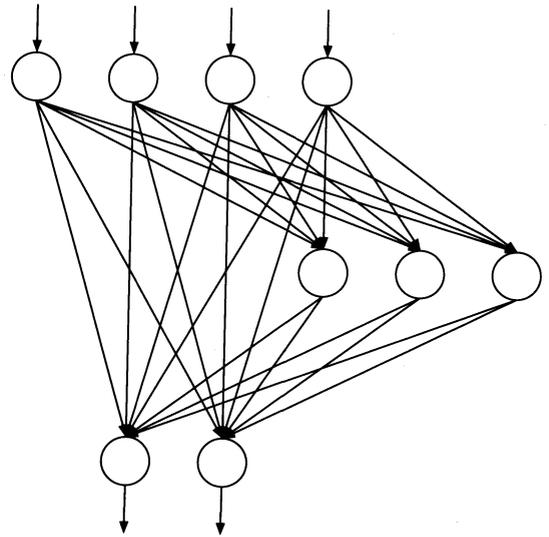
加えて、過適応を防ぐ措置として Weight Decay アルゴリズムを併用した。ここで用いた Weight Decay アルゴリズムは、極小化すべき目標関数として、誤差評価関数にネットワークパラメータ（各層間の線形結合係数）の二乗和を付加したものを設定するもので、誤差を小さくする過程でパラメータが大きくなりすぎないように抑制する働きを持つ。

このアルゴリズムにおいては、パラメータの二乗和の項に適当な係数を乗ずることで、学習によって得られる非線形関数の滑らかさをコントロールすることができる。係数の値をいくつか変えて調べたところ、係数が小さすぎると関数の非線形性が強くなり、過適応が起こって検証データでの精度が落ちるが、ある程度以上係数が大きければ検証データにおける精度はそれほど変化しないことがわかった。そこで、以下では4層フィードフォワード型ネットワークについては50、3層ハイブリッド型ネットワークについては10という値に係数を固定して用いている。これらの値は経験的に設定したものであるが、4層フィードフォワード型ネットワークの方が非線形性をより表現しやすいことを考えると、こちらの方に大きな係数が設定されることは妥当である。

なお、極端に大きな Weight Decay 係数で学習を行うと、完全に滑らかな関数（つまり、定数）になってしまうので、注意が必要である。



第2図 4層フィードフォワード型ニューラルネットワークの模式図。データは上から下へ流れるものとして、円で表わされたものは、ニューロンで、ここでは上流から受け取った値を特定の遷移関数で変形して下流に渡している。



第3図 3層ハイブリッド型ニューラルネットワークの模式図。

4.2 ネットワーク構造の選択

一般にニューラルネットワークの応用においては3層ないし4層のフィードフォワード型ネットワーク(第2図)が用いられることが多い。

本稿では、4層フィードフォワード型(2つの中間層のニューロン数は280と130とした)の他に、3層ハイブリッド型という構造を採用し比較を行った。3層ハイブリッド型は、第3図にあるとおり、一般的な3層フィードフォワード構造に、入力層と出力層を直結するリンクを加えたものである(ここでこの中間層ニューロン数は250としている)。調節可能なパラメタの数はそれぞれ123090と123730となり、ほぼ同程度である。

一般にニューラルネットワークの各層では、上流側の層からの出力値の線形結合を適当な遷移関数によって加工して出力する。一般的には、入力層の遷移関数としては線形関数、中間層の遷移関数としては、シグ

モイド関数($1/(1+\exp(-x))$)が用いられる。シグモイド関数は、 x が0より大きいと急速に1に近づき、 x が0より小さいと急速に0に近づくという性質を持っているので、おおまかに見れば、 x の値に応じて0か1かを与える関数と見なすことができる。第1中間層においては、上式の x は入力変数の線形結合になるから、第1中間層の各ニューロンでは線形結合の係数に応じて設定される分割面によって、入力変数空間を2分割していることになる。このように、通常のフィードフォワード構造においては、中間層ニューロンごとに分割面が設定されていて、それらによって入力変数空間が細かく分割された後に、それぞれの部分空間ごとに線形関数の当てはめが行われていると考えることができる。一方3層ハイブリッド構造は、入力変数空間を分割して記述する通常のニューラルネットワーク構造に加えて、入力変数空間全体に1つの線型関数を当て嵌める項を付加したものと見える。

4.3 出力層の遷移関数の選択

出力層の遷移関数は、必ずしもシグモイド関数である必要はない。降水量予測に利用できる出力層の遷移関数のうち、代表的なものとしては次の3種類がある。

- (a) 線型関数
- (b) 指数関数
- (c) シグモイド関数

このうち(a)は中間層からの出力値の線形結合をそのままニューロンの出力値として扱うものであり、(b)

は中間層からの出力値の線形結合を指数関数に通してから出力するものである。(a)の場合は負の値が出力される可能性があるため、降水量予測の場合は、必ず0以上の値を出力する、という意味では(b)の方が好都合といえる。(c)については、関数の値域が(0, 1)であることから、出力値を降水量そのものとして扱うことはできないが、出力値が降水量階級の出現確率を表わしていると見なして利用することはできる。本稿では、次節に述べる階級区分に従った降水量階級の出現確率を与えるものとして扱った。なお、この場合には、階級の数に応じて出力すべき値の数も増えることになるので、これを防ぐために確率値の「折り込み」を行って、全体の計算量が増えないようにした(付録参照)。

調査の段階では、出力層の遷移関数として上記の3種類全てを試したが、本稿では線形関数とシグモイド関数の結果についてのみ述べる。指数関数を用いたものについては、出力値の頻度分布が若干異なる他は、線形関数の場合と大差なかった。

5. 検証結果

5.1 検証方法

降水量は、全事例の半数以上で0となる特殊な頻度分布を示す(第4図の黒のヒストグラム)。このため、2乗平均平方根誤差のように、ガウス分布する資料の評価に適した指標を用いても、解釈することが難しい。

ここでは、利用者の実感をできるだけ反映した評価方法ということを考えて、降水量を7つの階級に区分し、7×7の分割表を作ってそこからHeidkeのスキルスコアを計算することとした。降水量を7階級に区分するためのしきい値は、0.5, 1, 2, 5, 10, 15 mmとした。

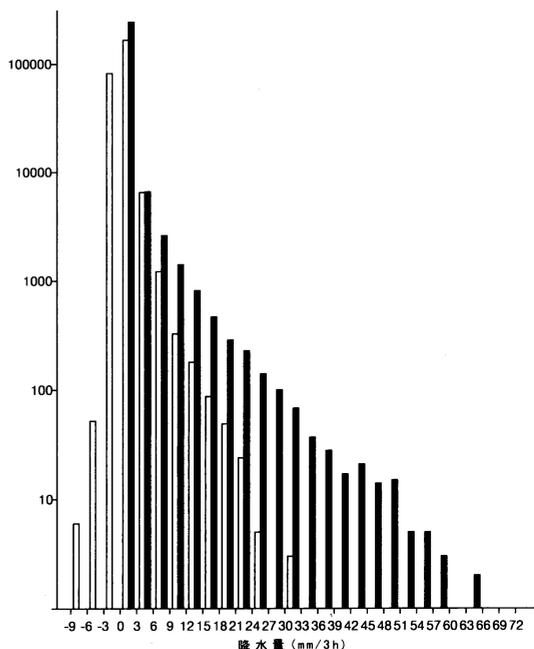
Heidkeのスキルスコア S は次の式で計算される(このスコアの性質についての詳しい議論は菊地原(1988)を参照)。

$$S = (T - C) / (N - C) \quad (1)$$

ただし、 N はデータの総数、 T は的中数、 C は気候学的に期待される的中数(デタラメに予測を行った場合のまぐれ当たりの数、という解釈も可能)で、以下のように計算される。

$$C = (\sum F_i O_i) / N \quad (2)$$

ここでは、 i は1から7までの値をとり、 F_i は階級



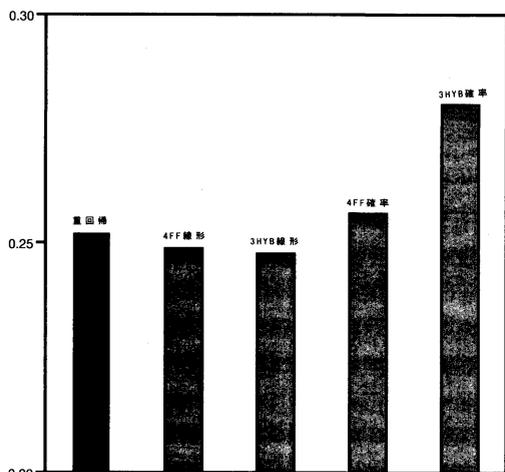
第4図 1995年1～12月の、第1図の各メッシュにおける3時間降水量の頻度分布。縦軸は対数目盛。黒いヒストグラムは実際の降水量のもので、白いヒストグラムは重回帰式の予測値のものである。

i を予報した数、 O_i は階級 i が観測された数を表わす。

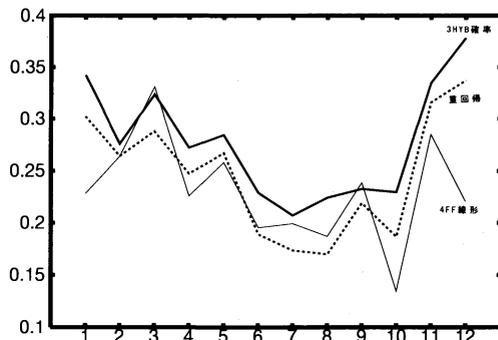
ニューラルネットワークの精度を従来の手法と比較するために、同じ入力データを用いて、第1図のメッシュごとに線形重回帰式を作成した。ただし、前処理としてステップワイズ法による変数選択を行って入力変数の数を11～26個にしぼりこんでから重回帰式を作った。

入力変数の多くがガウス分布に近い頻度分布を示すので、重回帰式による予測値の分布もガウス分布に近いものとなる(ガウス分布に従う資料の線形結合もまたガウス分布に従う)。これに対して、降水量は第4図の黒のヒストグラムで示したような頻度分布を示すため、重回帰式の出力値は値の小さい側に偏る傾向がある(第4図の白のヒストグラムは重回帰式の出力値の頻度分布を表わす)。これを補正するために、重回帰式の出力値を降水量階級に変換する場合のしきい値を以下の手順によって調節すると、検証データでもスコアの値が改善される。

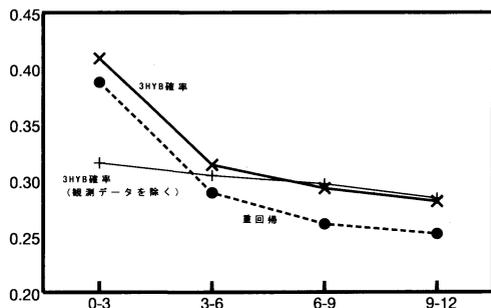
(a) 従属資料における重回帰式の出力値を昇順に



第5図 観測データが与えられてから9~12時間後の予測についての検証期間のスキルスコア。「4FF」は4層フィードフォワード構造、「3HYB」は3層ハイブリッド構造を表わす。「線形」は出力層の遷移関数が線型関数であることを、「確率」はシグモイド関数によって各降水階級の確率が出力されることを表わす。



第6図 観測データが与えられてから9~12時間後の予測についての検証期間の月別スキルスコア。予測法の種別の表示は第5図と同じ。



第7図 観測データが与えられてからの時間によるスキルスコアの変化。実線は3層ハイブリッド・シグモイド出力のニューラルネットワークのもの、破線は重回帰式のものである。なお、細実線は上のニューラルネットワークに、入力として数値モデルGPVのみを与えた場合のスコア。

並べる

(b) 6つのしきい値 (0.5, 1, 2, 5, 10, 15 mm) それぞれについて以下の操作を行う

(b-1) 出力値のしきい値を i 番目の出力値と $i+1$ 番目の出力値の中間に設定し、 2×2 の分割表を作って Critical Success Index (CSI) を計算する。これを従属資料中の全データについて行い、CSI が最大となるしきい値を求める

(b-2) CSI が最大となるしきい値を、その階級のしきい値とする

以下では、この方法で調節したしきい値によって重回帰式の出力値を降水量階級に変換してニューラルネットワークと比較した。

5.2 検証結果

第5図は観測データの時刻から9~12時間先の予測についての検証期間におけるスキルスコアを、重回帰式および4種類のニューラルネットワークについて示したものである。標準的なニューラルネットワークである4層フィードフォワード+線型出力の組み合わせは、重回帰式よりもやや低いスコアとなっている。出力層の遷移関数をシグモイド関数にして、降水階級の確率を出力する形にするとスコアは上昇し、特に3層ハイブリッド構造にシグモイド関数を組み合わせたものが最

も高いスコアを示した。

第6図は同じく9~12時間先の予測についてのスコアを、検証期間の月別に示したものである。3層ハイブリッド構造とシグモイド関数を組み合わせたネットワークは、ほぼ全期間にわたって最も高いスコアを示している。

第7図は観測データの時刻からの時間別に見たスキルスコアである。観測データの時刻からの時間が長くなるにつれてスコアは下がっていくが、どの時間においても、3層ハイブリッド構造+シグモイド関数の組み合わせのネットワークは重回帰式より高いスコアを示す。なお、図中の細実線はニューラルネットワークの入力から観測データを除き、数値予報モデルのデー

タのみを与えた場合のスコアである。これを見ると直近の観測データを入力に加えることによって、0～3時間後のスコアは大幅に改善されており、3～6時間後のスコアもわずかながら改善されていることがわかる。ここでは入力資料としてレーダーアメダス解析雨量だけを用いているが、降水短時間予報の予測値を加えれば、3～6時間後のスコアを更に改善することも可能と考えられる。

6. 考察

一般的に利用されているフィードフォワード型ニューラルネットワークを線形出力と組み合わせる降水量予測に適用したところ、重回帰式と同程度以下の性能しか示すことができなかつた。しかし、3層ハイブリッド構造とシグモイド関数を組み合わせたネットワークの場合は、重回帰式よりも高いスコアを示した。このことは、標準的な手順に従ってニューラルネットワークを構築するだけで、自動的に精度の良いものが得られるわけではない、ということを示唆している。

3層ハイブリッド構造とシグモイド関数を組み合わせたネットワークが良好な性能を示した理由については、今のところ次のような仮説を立てている。

降水量は第4図の黒のヒストグラムで示したような頻度分布を示し、降水量の少ない事例の数が圧倒的に多く、降水量の多い事例の数は少ない。重回帰式はこれらの全ての事例について、予測資料との間にただ一つの線形関係を作ろうとするのであるが、その際に最小二乗法が使われていることもあって、得られた結果は数の多い（すなわち、降水量の少ない）事例における関係に比重のかかったものとなっていると考えられる。

一方、ニューラルネットワークは先にも述べた通り、中間層において入力変数空間が分割され、分割された部分空間ごとに関数の当てはめが行われていると見ることができる。ニューラルネットワークが複雑な関数に対して柔軟な当てはめを行うことができるのは、この入力変数空間の分割に依るところが大きい。しかし、降水量予測の場合は、降水量の少ない事例と降水量の多い事例とを分割してしまうと、降水量の多いケースについては、少数の事例だけに基づいた当てはめが行われてしまう可能性がある。

一般に統計モデルにおいては、多数事例から作られたモデルは頑健で一般性が高いのに対して、少数の事例のみから作られたモデルは一般性が低く、新規の入

力に対して適切な出力を与えることが難しい。つまり、重回帰式は全事例について一つの線形関係を与えるため概ね頑健であるのに対して、ニューラルネットワークは入力変数空間の分割のために頑健な部分と脆弱な部分とが混在しているものと考えられる。

3層ハイブリッド構造は、中間層を介する標準的なニューラルネットワーク構造と、入力変数の線形結合を直接出力する線型重回帰式の構造の双方を重ね合わせたものとなっており、そのことがスキルの改善に寄与した可能性がある。

出力層の遷移関数について、シグモイド関数が良好な結果を与えた理由は明確でない。一つ考えられることとしては、ニューラルネットワークはその柔軟性によって、教師データ内の入力-出力関係を良く再現するが、実はこの入力データはそこまでの情報を内包しておらず、結果として過適応が起こった、というものである。用意した入力データが出力値に対してどの程度の情報を持つか、ということを知るのは難しいが、結果から見ると、ここでは出力値を降水階級へと離散化することで、出力値の情報量が減り、入力データの持つ情報量とのバランスがより適正になった可能性がある。

以上に述べたことはいずれも推測であり、客観的な裏付けはない。今後、なんらかの実験によって、3層ハイブリッド構造やシグモイド関数による確率出力の有効性の理由が解明されれば、予測対象に応じたネットワーク構造・遷移関数の選択がより容易になるものと考えられる。

付録：ニューラルネットワークによる確率値の出力

降水量があるしきい値を越えるか否か、といったような2分割の問題の場合には、ニューラルネットワークの出力層におけるシグモイド関数の出力値を確率と見なすことが許される。ただしそのように扱うときには、学習の過程において極小化すべき誤差関数を、平均二乗誤差ではなく、以下のようなクロスエントロピーに設定しておく必要がある。

$$E = - \sum \{ t \log p + (1-t) \log (1-p) \} \quad (3)$$

ここでの総和は、教師データのサンプルすべてについてのものである。 p はニューラルネットワークが出力した確率値であり、 t は真の降水量が設定されたしきい値を越えていれば1、それ以外は0となる値である。本稿のように降水量を7階級に分ける場合でも、

6つのしきい値それぞれについて2分割を行うと考えればよい。

しかし、降水階級の数が増えると、各々の階級についての確率値を出力しなければならないため、降水量そのものを出力する場合に比べて出力すべき値の数が増えるという難点がある。たとえば、7階級の降水確率を出力しようとするれば、出力すべき値の数は6倍になり、それにともなって中間層と出力層を結ぶパラメタの数も6倍になるため、4層フィードフォワード型の場合、パラメタ数は123090から201690に増え、それに伴って学習に必要な計算量も増大することになる。

一方で経験的には、各階級の確率は完全には独立でない。たとえば、10 mm以上の確率が80%であれば、1 mm以上の確率は80%よりもかなり大きな値であろうと見当がつくし、10 mm以上の確率が上下すれば、1 mm以上の確率も上下すると考えるのが自然である。

そこで本稿においては、降水量がしきい値 θ を越える確率 p を、次のような1つの式で表現することを試みた。

$$p=1/\{1+\exp(-\sum wx+\beta \log \theta)\} \quad (4)$$

ただし、 x は注目している出力層ニューロンと結合している各ニューロンからの出力値、 w はニューロンごとに異なる結合のパラメタであり、総和は結合しているすべてのニューロンについてとるものとする。

出力層ニューロンに与えられる1つの値 $\sum wx$ に対して、各階級の確率値 p は、 θ が大きいほど小さくなるような値として出力されることになる。ここで、 β というパラメタは正の実数で、 θ の違いにともなう確率値の変化の割合を調節している。この定義式そのものは便宜的なものであり、パラメタ β を決定する理論的な根拠はないので、このパラメタも w などの他のパラメタと同様に、学習によって調節することとした。

具体的には次のような手順となる。真の降水量 r が与えられると、それと各階級のしきい値 θ とを比較することで、(3)式の t が1または0に決まり、(3)式から誤差の値 E が求まる。また、 E を各パラメタで微分

した関数から勾配方向が求まるので、それをもとに学習を行うことができるのである。

本稿では降水を7階級に分けたので、6種類のしきい値 θ が存在する。従って、1つの観測値 r に対して、6つの t が求まることになる。これは教師データのサンプル数が6倍になることを意味するが、これによる計算量の増加は、上で述べたパラメタ数の増加に伴う計算量の増加に比べれば問題にならない。この方法により、ネットワークの規模や必要な計算量を、降水量そのものを出力する場合とほぼ同等に抑えることができる。これは異なる階級の確率を1つの関数に「折り込んで」表現しているので、本稿では「折り込み確率」と称する。

この式によれば、ある階級の確率は他の階級の確率と連動して変化することになるが、これが現実の降水現象においても成立しているという保証は無い。しかし、本文中に見る通り、この方法でも良好な成績が得られていることから、実際の降水量の持つ性質から大きくかけ離れたものではないと考えられる。

参考文献

- Baik, J.-J. and H.-S. Hwang, 1998: Tropical cyclone intensity prediction using regression method and neural network, *J. Met. Soc. Japan*, **76**, 711-717.
- Fletcher, R. and C. M. Reeves, 1964: Function minimization by conjugate gradients, *Comput. J.*, **7**, 149-154.
- 菊地原英和, 1988: 気象予測の検証と評価, 気象研究ノート, (161), 189pp.
- Koizumi, K., 1999: An objective method to modify numerical model forecasts with newly given weather data using an artificial neural network, *Wea. and Forecasting*, **14**, 109-118.
- Marzban, C. and G. J. Stumpf, 1996: A neural network for tornado prediction based on Doppler radar-derived attributes, *J. Applied Met.*, **35**, 617-626.
- Marzban, C. and G. J. Stumpf, 1998: A neural network for damaging wind prediction, *Wea. and Forecasting*, **13**, 151-163.

A Neural-network Structure Suitable for Precipitation Amount Forecast

Ko KOIZUMI* and Masanobu HIRASAWA**

* (Corresponding author) Forecast Department, Japan Meteorological Agency 1-3-4 Otemachi
Chiyoda-ku Tokyo, JAPAN

** Meteorological Research Institute.

(Received 7 July 1999 ; Accepted 14 September 2001)

Abstract

Two types of network structure and two types of output-layer transfer function were tested for neural-network precipitation amount forecasts. The neural network of 3-layer hybrid structure (which has links combining input layer and output layer directly) with sigmoid function as output-layer transfer function showed the best skill score among four types of neural networks (including a traditional 4-layer feed-forward network) and a linear multiple regression model.

The result suggests that the neural network which is most skillful for a specified problem cannot be given automatically but the network structure or output-layer transfer function must be chosen properly.

2001年度秋季大会の報告

日本気象学会2001年度秋季大会は、岐阜県民文化ホール未来会館（岐阜市学園町3-42）を会場として2001年10月10日（水）～12日（金）に行われた。参加者数（前納登録者と当日受付者の合計）は624名（一般会員402名、学生会員138名、非会員84名）であった。

2日目午後には、未来会館・長良川ホールにおいて山本・正野論文賞が谷本陽一会員と渡部雅浩会員に、堀内賞が久保田雅久会員と小池 真会員に、日本気象学会奨励賞が出口 一会員、大久保 篤会員および加治屋秋実会員にそれぞれ授与された。授賞式のあと、山本・正野論文賞及び堀内賞の受賞記念講演が行われた。続いて岐阜地方気象台が岐阜県教育委員会の協力を得て取り組んでいる「岐阜県地域気象教育プロジェクト」の概要紹介があった。引き続き、大会シンポジ

ウム「東海豪雨—自然、都市、人間の関わり—」が行われた。

一般講演の発表申込み件数は382件で、その内訳は口頭発表が252件、ポスター発表が130件であった。

会期中およびその前日と翌日には、個別のテーマによる研究会や講演会が5件開かれた。

最後に、今大会事務局として大会準備・運営にご尽力頂いた名古屋地方気象台、岐阜地方気象台、名古屋大学地球水循環研究センター、名古屋大学大学院環境学研究科、愛知教育大学教育学部、岐阜大学工学部、(財)日本気象協会東海支社をはじめとする中部支部の皆様深く感謝の意を表します。

2001年12月 講演企画委員会