

自然界の時系列予測問題と時系列の特徴に関する考察

Consideration Concerning Feature of Time Series Prediction Problem and Time Series of the Natural World

渡邊 達男

Tatuo WATANABE

1. 序論

自然界のさまざまな現象は複雑で、その多くは予想するのが大変困難である。もし自然現象を予測できれば、災害や、それに伴う人的被害も回避する可能性があり、切望されるところである。しかし実際には自然界はとても複雑で、特に長期の予測は困難を極める。

最近の非線形科学は少しずつであるが、困難な自然界の予測を行おうと試みている^{1,2,3)}。様々な方法で予測を行おうとしているが、自然界をトータルで予測するのは難しい。しかし、モデルを簡単にして、1変数の関数として、さらに短期の予測するのはそれほど難しくもない事が最近知られるようになってきた。

自然界の予測では、古くはLorenzが気象現象を予測する為に、3変数の予測関数系、いわゆるLorenz系を提案している³⁾。この系は非線形の3変数の微分方程式であり、残念ながら気象現象を予測するのには不足していたが、後のカオス系の研究の火種になり、多くの提案をした。

時系列予測に関してはその後、幾つかの予測器が考えられている^{1,2)}。その幾つかは以前の筆者の論文で考察した^{4,5,6)}。Lorenzの類推法では、数ステップ先までは予測できるが長期予測は非常に困難であると思われた。最近よく話題になっている動径基底関数ネットワーク法^{1,2)} (RBFN法, Radial Based Function Network) では、大変少ない過去のデータで、比較的良い短期予測が可能であったが、複雑な自然現象、例えば台風の上陸数等は予測が難しかった。

この論文では、自然界の幾つかの時系列を例に挙げ、その予測を試みる。その際、RBFN法を用いる。用いる理由は、少ない情報量で比較

的良い予測ができるからである。自然界の過去のデータは観測の結果得られるものであり、必ずしも多く得られるとも限らない。Lorenz類推法では沢山のデータが必要で、自然界の予測には難しい側面がある。その点RBFN法は少ないデータで予測可能である。予測に際してさらにデータの性質を予測のためのパラメータに反映する。用いるのはデータのスペクトルである。これらのデータ特性が予測にどのように関係してくるかを簡単に見て行く事にする。またフラクタル次元がどのように時系列に関わるかの試みも行う。

2章はRBFN法の説明に、3章はフラクタル次元の概説に、4章は計算による結果を、5章は考察、6章はまとめに当てられる。

2. 動径基底関数ネットワーク法^{1,2)}

動径基底関数ネットワーク法 (RBFN法)は、変形された2層ニューラルネットワークの一種で、動径基底関数という局在化された基底関数を重ね合わせることで、任意関数補完の方法として考えられて来た。過去の論文の繰返しになるが、以下に簡単に説明をする。

もともと関数補間の方法であったが、M.Casdagli⁷⁾はRBFN法をカオス時系列予測問題に応用した。

今、時系列 $x(0), x(1), \dots, x(t)$ を要素を持ったベクトル $v(t) = \{x(0), x(1), \dots, x(t)\}$ に対して、次の式を考える。

$$x(t+1) = \sum_i^M \lambda_i \Phi(|v(t) - c_i|) \quad (1)$$

ここで、 λ_i は重み、 c_i はセンターと呼ばれる点、 Φ は基底関数である。

基底関数 Φ には幾つかのものが使われる。よく使われるものに、ガウシアン関数があり、

$$\Phi(r) = \exp(-r^2/b) \quad (2)$$

で与えられる。

シンプレートスプライン関数、

$$\Phi(r) = r^{2k+1} \log r \quad (3)$$

ベル型関数、

$$\Phi(r) = \frac{1}{1 + \cosh(br)} \quad (4)$$

等もある。今回はガウシアン関数 (2) を用いて計算した。

RBFN法を用いて時系列を予測する場合には、あらかじめ既に知られている時系列(データと呼ぶ)を用いて重み λ_i を計算する。 λ_i は2層ニューラルネットの結合係数に相当するが、学習を用いて決定するのではなくて、一般には式 (1) に知られているデータ時系列を代入し、連立方程式を解くと λ_i が求められる。

今回の報告では、過去のデータ時系列を用いて、 λ_i を決定する方法を用いた。

センター c_i は任意である。センターはベクトルで、一般的には時系列 $v(t)$ と同じ次元でなくてはならない。センター c_i に関しては筆者の論文⁶⁾において、過去の時系列データを用いることで良い予測ができることがわかっている。今回は過去の時系列データをセンターに採用することにする。

上記の方法はスタンダードRBFN法と呼ばれているが、式 (1) に線形項を加えたアファインプラスRBFN(APBFN)もある。この方法はノイズ成分を線形項に吸収できるので、ノイズを含む時系列に有効である¹⁾。

$$x(t+1) = A v(t) + B + \sum_{i=1}^M \lambda_i \Phi(|v(t) - c_i|) \quad (5)$$

3. フラクタルとフラクタル次元

フラクタルとは自己相似性を持つ図形のことである。フラクタルを定量的に表す量はフラク

タル次元である。フラクタル次元及びフラクタル次元を求める方法は幾つか有る。その中で、時系列データに用いられる方法の一つに相関積分を用いた相関次元がある¹⁾。

相関積分とは、時系列を要素とした次元 m のベクトル $v(t)$ に対して、

$$C^m(r) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^N I(r - |v(i) - v(j)|) \quad (6)$$

を求める。ここで $I(t)$ は単位ステップ関数であり、

$$I(t) = \begin{cases} 1 & (t \geq 0) \\ 0 & (t < 0) \end{cases} \quad (7)$$

である。

もし、 $v(t)$ が自己相似性を持っていると、

$$C^m(r) \propto r^{\nu(m)} \quad (8)$$

の関係があり、対数を取ると、

$$\log C^m(r) \propto \nu(m) \log r \quad (9)$$

である。ここで $\nu(m)$ が相関次元である。 $\log r$ と $\log C^m(r)$ とのグラフの傾きが $\nu(m)$ を与える。

実際の自然界の時系列では N を大きくとり、 m を増加させたとき、 $\nu(m)$ がある値で飽和する。その値が相関次元である。この方法をGP法と言う。

なお、RBFN法で用いる次元 M とフラクタル次元を求める際のベクトルの次元 m とを大文字、小文字で分けて表記した。

4. 時系列データの特徴を考慮した予測結果

次に過去の時系列データの特徴を用いて、その特性を反映することで、予測がどう変化するかを見る事にする。特徴として、過去の時系列データのスペクトル値を用いた。また時系列のフラクタル次元に関しても試みた。

4. 1 スペクトル

過去の時系列データのスペクトルは何らかの予測に役立たないかと考えるのは普通のことである。ここでは、時系列データのスペクトル値を計算し、スペクトル値が最も高い周波数をベクトルの次元に取る事により、どの程度予測が

良くなるかを調べた。

図1は、栃木県小山市の月ごとの降水量のスペクトル値のグラフである⁹⁾。

図では横軸が周期（月）、縦軸がその周期に対するスペクトル値を表している。図を見ると、予想通り周期12（月）の部分でピークが現われる。周期が12以外に6、7付近にピークが現われる。

図1の結果を元に、RBFN法でベクトルの次元M=12で計算した結果を以下に示す。

図2は先の論文で発表したものと同一ものである。M=12であるとよく合う。しかし細かい部分のフラクテーションは追従してない。

M=6,7で行うと、予測精度は悪い。従って、このようなある程度周期的な変動を持つ時系列で、あるスペクトル値が高い場合には、その値をベクトルの次元にとると予測が良くなる事が言える。

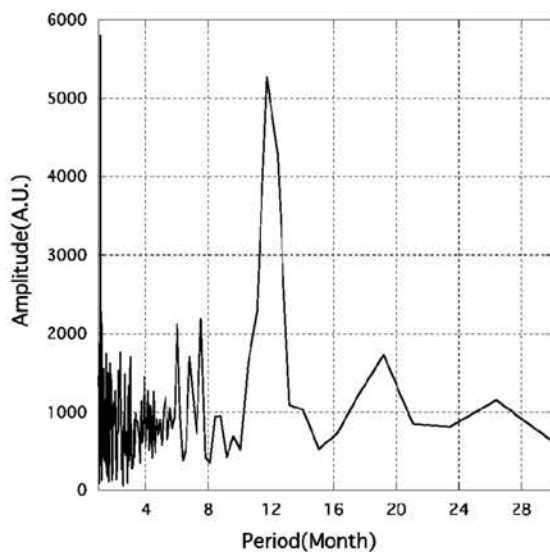


図1 小山市の月降水量のスペクトル
(1990/1-2007/8)

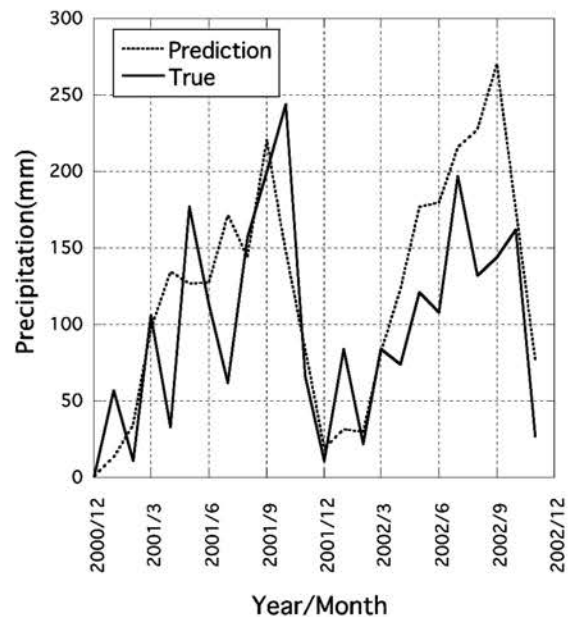


図2 小山市の降水量の予測（再掲載⁶⁾）

しかし、ただスペクトル値が高いからと言って、必ずしも予測が良いとは言えない。

図3は1951年から2006年までの台風の日本への上陸数のスペクトルである¹⁰⁾。横軸が周期（年）、縦軸がスペクトル値である。図を見ると、複雑であるが、周期が1、3、4、6、14年あたりにピークを見る事ができる。そこで、試みとして、RBFN法のベクトル次元M=3,4,6として予測を行い、真値と比較したグラフを図4に示す。

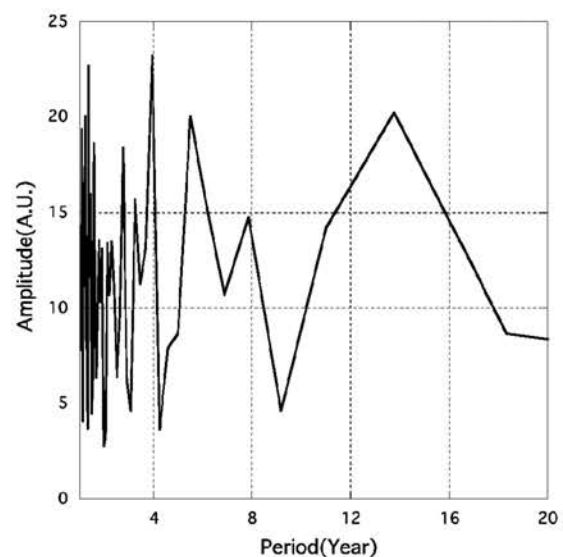


図3 台風の日本への上陸数のスペクトル
(1951-2006)

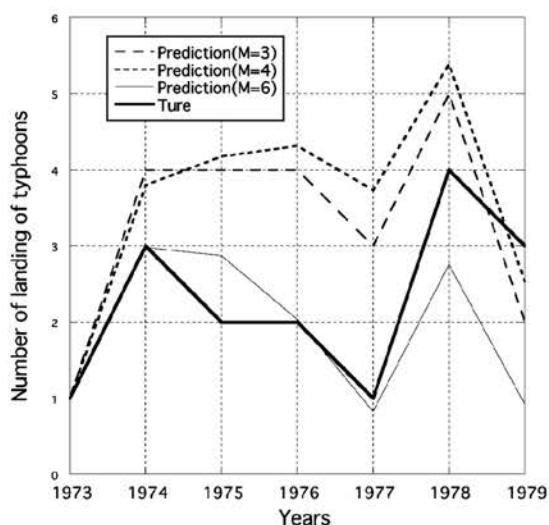


図4 台風の上陸数の予測

図4を見ると、 $M=3,4$ ではあまり予測が当たっていない。本来なら、 $M=4$ でスペクトルがピークなので、何らかの予測の良さを予想したのであるが、そうでもなかった。しかし、 $M=6$ で予測を行うと非常に良く合う。値は4年先まで、傾向は6年先まで当たっている。従って、スペクトルのグラフが予測に何らかの好影響を与えていると考えられる。しかし、スペクトル値だけでは決まっていないうである。

4. 2 フラクタル次元

4. 1の時系列データのフラクタル次元を求めて、何らかの特性が得られないかと試みた。

図5は小山市降水量の時系列データから、 $\log r$ と $\log C^m(r)$ とのグラフを作ったものである。次元 m をパラメータとしてある。これらのグラフの傾きがフラクタル次元 $\nu(m)$ を与える。ただし m を増加したときの傾きの飽和値である。

図を見ると、 m を変化させると、グラフの傾きが大きく変わっている事がわかる。ただし、 m が大きいくところでは、傾きが飽和しているように思われる。

そこで、図5の傾きのグラフを求めてみた。図6に示す。

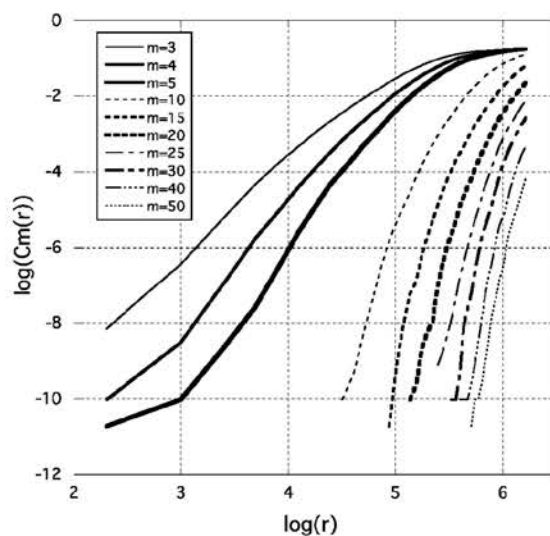
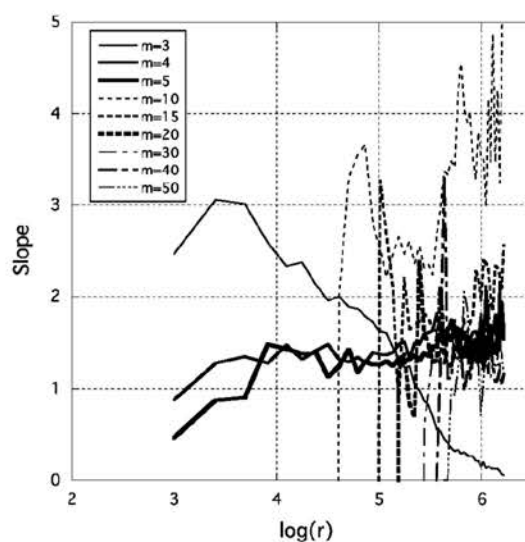
図5 降水量の $\log r$ と $\log C^m(r)$ 

図6 図5の傾き

図6を見ると、グラフの右の方で、だいたい1.5に収束している。従って、単純に判断すると、フラクタル次元は約1.5ということになる。しかし、グラフのどこの傾きを取るかでかなり異なってしまう。

このフラクタル次元をどのように予測問題にどのように扱うかはもう少し考察が必要である。

5. 考察

幾つかの計算を行い、自然現象の予測を行ってきたが、ここで若干の考察を行う。

スペクトル値を用いたRBFN法のベクトル次元の選択と予測はある程度の効果があるように

思われる。しかしスペクトルのピークが高い場合には効果があるが、他のピークと比較してあまり差が無いピークの場合、その周期をベクトル次元にすると必ずしも予測が良くない。この理由はよくわからないが、今後の考察の必要がある。

また、フラクタル次元においては、GP法は多くのデータが必要である。今回のデータは僅か200あまりであった。このデータ数でどの程度までフラクタル次元は求まったかはわからない。また求めたフラクタル次元をどのように扱うかは問題である。

図5において、 $m=40$ での $\log r$ が小さいところでの傾きは約1.2である。これはスペクトルの章で議論してベクトル次元に取ったものであるが、 $\log r$ の小さいところで m と傾きに何かヒントがあるかもしれない。

6. 結論

RBFN法による、自然界の時系列予測問題で、パラメータの選び方を時系列のスペクトルにより決定し、それなりの効果が出た。また時系列のフラクタル次元を計算し、それが何らかの予測に関する指針を与える可能性を考えてきた。スペクトル値をベクトル次元に選ぶと、少し予測が良くなる場合があった。今後はさらに予測を高めて行く方法を考えて行きたい。

参考文献

- 1)合原一幸編:カオス時系列解析の基礎と応用,産業図書(2000).
- 2)北川源四郎:時系列解析入門,岩波書店(2005).
- 3)E.N.Lorenz:Journal of the atmospheric sciences, 20(1963)130.
- 4)渡辺他:小山高専紀要38(2006)101.
- 5)渡辺:小山高専紀要39(2007)107.
- 6)渡辺:小山高専紀要40(2008)91.
- 7)M.Casdagli:Physica D, 35(1989)335.
- 8)合原一幸編:カオスセミナー,海文堂(1994).
- 9)<http://www.jma.go.jp/jma/index.html>
- 10)<http://www.fukuoka-jma.go.jp/fukuoka/chosa/taifu/sekkin.html>

「受理年月日 2008年9月25日」

