一般化加法モデルを用いた気温過程の 平均・分散構造解析

総合研究大学院大学 複合科学研究科 統計科学専攻 若 浦 雅 嗣

要 旨 気温日間変動データの平均と分散の構造を明らかにすることは、ウェザー・デリバティブの価格評価における関心事であるが、その構造はパラメトリックモデルで表現できるほど単純なものではない。本稿は、日本における気温の平均と分散の長期トレンドと周期トレンドを一般化加法モデルで分析し、わが国の気温変動の時系列的特徴を記述した。とりわけ、分散については、明確な年周期の構造を持ち、地域的な類型化が可能ながことが明らかとなった。

1. 気温変動過程のモデル化のニーズとノンパラメトリック回帰の必要性

本稿は、気温の日間変動の時系列データについて、その平均と分散の長期トレンドならびに周期トレンドを、ノンパラメトリック回帰を用いてモデル化したものである。

この種のモデリングは、ウェザー・デリバティブという金融商品の設計にとって大きな関心事である。ウェザー・デリバティブは、金融派生商品、つまり金融デリバティブ(一般には単に「デリバティブ」と称するが、ウェザー・デリバティブと区別するために、本稿では金融派生商品を「金融デリバティブ」という。)が、株や金利、為替といった金融取引の変動によるリスクをヘッジするのと同様に、気象要素の変動によるリスクをヘッジする商品である。金融デリバティブにおいては、そのリスクの対象となる原資産の時系列の変動をボラティリティ(分散)という形で定式化し価格評価するという手法が用いられてきた。同様にウェザー・デリバティブを価格評価するとすれば、気温等の気象要素について、その時系列の分散を解析する必要がある。

金融デリバティブのオプション¹の評価式としては、Black and Scholes (1973) が有名であるが、これは、株などの原資産が市場で取引され裁定がないことが前提となっている。だが、気温等の気象要素は市場で取引はされない。つまり、気温そのものに価格はない。したがって、株などの市場の無裁定を利用した評価はできない。したがって、気温オプション²の場合には気温変動を予測し、

¹ オプションとは、対象となる原資産(株式や債権)を所定の期日までに所定の価格(行使価格)で買い取るあるいは売り付ける権利で、原資産の価格と行使価格の比較により権利の行使が決定され、これにより原資産の価格変動のリスクを回避することができる。

² 気温オプションとは、ウェザー・デリバティブの契約形態の一つで、主に、気温のインデックス値(日次平均気温値が所定の値を上回った場合や下回った場合の乖離値や日数をカウントするものがある)を一定期間累積した値が、所定の累積値(行使値)を上回った場合や下回った場合に、その累積値と行使値との差について単位インデックスあたりの約定金額を乗じた金額の授受が生じる契約。

ここから当事者の気温変動に基づく期待損益を求めて、それに経験的なリスク・プレミアム³をのせて価格を導出するという手法が用いられることとなる。そこで、価格評価においては、いかに気温の変動を予測するか、あるいはシミュレーションするかということが焦点となり、気温等の気象要素の時系列変動を記述するモデルの構築が課題となる。

これまで、ウェザー・デリバティブにおけるモデル構築のアプローチについては、発祥の地であるアメリカにおける取引が主に気温に関するものであったことから、気温過程を中心に先行研究が行われてきた。

先行研究におけるモデルを紹介すると、まず、Vasicek (1977) の提案した金利の変動過程を表す平均回帰過程モデルを、気温過程 T_n ($n = 1, 2, \dots, N$) に置き換えて離散化した次のモデルが挙げられる。

$$T_{n+1} = T_n + (\Theta_{n+1} - \Theta_n) + a(\Theta_n - T_n) + \varepsilon_n.$$

ここで、 Θ_n ($n = 1, 2, \dots, N$) は n 日の標準(平均)気温であり、 a は平均値に回帰する速さを表すパラメータで、 ε_n は誤差項である。また、次の自己回帰モデルも利用されている。

$$\begin{aligned} Z_n &= T_n - \Theta_n, \\ Z_n &= \sum_{i=1}^p \phi_i Z_{n-i} + \varepsilon_n. \end{aligned}$$

ここで、 ϕ_i はパラメータである。なお、平均回帰過程モデルを拡張したモデルとしては、Dischel (1999) と Torró et al. (2001) がある。

特に、Torró et al. (2001) では、STI と呼ばれるスペインの地域人口で加重平均した気温インデックスの過程を対象とし、その特徴として、明確な季節周期性、高い自己相関、平均回帰性を挙げ、また、条件付分散変動の構造が示唆される点から、それらを加味した次の一般化モデルを提示している。

$$\begin{aligned} \Delta T_n &= [a_0 + a_1 T_n + a_2 \cos(2a_3 \pi n / 365 + a_4)] \Delta n + \varphi_n T_n^\gamma \Delta W_n, \\ \Delta W_n &= \varepsilon_{n+\Delta n} \sqrt{\Delta n}, \quad \varphi_{n+\Delta n}^2 = b_0 + b_1 \varepsilon_n^2 + b_2 \varphi_n^2. \end{aligned}$$

ここで、 Δn は時間間隔を表し、 $\Delta T_n = T_{n+\Delta n} - T_n$ は気温の変化を表す。また、 W_n は標準ブラウン運動を離散化したもので、 φ_n はその拡散係数である。一方、 ε_n は気温の平均値に対する残差で、標準正規分布に従うものと仮定している。上記モデルはパラメータ a_0, a_1, a_2, a_3, a_4 により平均の周期トレンドを、 b_0, b_1, b_2 で条件付分散変動の構造を表現し、さらにパラメータ γ の設定により、Vasicek (1977), Cox et al. (1985), Brennan and Schwartz (1982) 等の平均回帰過程モデルを表現できるように一般化されている。

Torró et al. (2001) は、上記モデルの各パラメータの設定を行い、それぞれ性能の比較を行ったところ、モデルのパフォーマンスとしては、季節トレンドに平均回帰する構造と、条件付分散変動の構造を加えることによって適合は単純な平均回帰モデルより有意に改善され、当てはまりと予測性能が向上することを確認した。Torró et al. (2001) のモデルは、このように気温過程の平均に関する季節周期性と分散に関する構造を取り入れた点で興味あるモデルと言える。

³ リスクを引き受ける際に、リスクに応じて上乗せされる利益。

これに対して, Cao and Wei (2000) は, 気温過程における分散の周期性について, そのモデリングの必要性を示唆し, 自己回帰モデルを拡張し気温過程を次のようにモデル化した.

$$\begin{aligned} Z_{y,t} &= T_{y,t} - \Theta_{y,t}, \\ Z_{y,t} &= \sum_{i=1}^p \phi_i Z_{y,t-i} + \sigma_{y,t} \xi_{y,t}, \\ \sigma_{y,t} &= \sigma_0 - \sigma_1 |\sin(\pi t / 365 + d)|, \quad \xi_{y,t} \sim \text{i.i.d. } N(0, 1). \end{aligned}$$

なお, 添え字の y は年次, t は日次数列 ($t = 1, 2, \dots, 365$) を表す. そして, $\xi_{y,t}$ は標準正規分布に従う確率変数で, $\sigma_{y,t}$ はその分散係数として季節的な周期構造 $\sin(\pi t / 365 + d)$ を持つ形で表現されている.

一方, 我が国の気温過程の変動にこの種の単純な周期性を想定するのは考えにくい. 先行研究では, 主として欧米地域の気象データ解析をもとに論じられてきたが, 気温等の気象データは地域的, 気候的な影響が大きく, 日本のように四季が明確で海洋の影響が大きい地域と, アメリカやヨーロッパなどの大陸的な気候特性を持つ地域ではその構造は異なるものと考えられる.

図 1 は, 地上気象観測データ (気象庁提供) の大手町における日次平均気温データ (1969~1998 年/30 年分, ただし, 閏年の 2 月 29 日は除いた.)⁴ をプロットしたものである. 大きな年周期と日々の微小な変動によって構成されるデータであることは一見して理解される. 図 2 は 1969 年から 1998 年までの 30 年間のデータについて, 1 月 1 日から 12 月 31 日の日別の標本分散値をプロットしたものであるが, 日毎に分散値が異なっていることがわかる. また, 図 3 は 1998 年の大手町のデータに単純な三角関数を当てはめたものであるが, 系統的乖離があることがわかる. 気温自体は, 気象という自然現象の営みの結果なので, 気温データは自然のサイクルに起因する周期的な構造を持っていると考えられる. しかし, その周期性は意外と複雑な構造であるため, ノンパラメトリック回帰を用いて柔軟に構造を抽出するのに適している.

図 1. 東京・大手町日次平均気温データ
(1969~1998 年/閏年の 2 月 29 日を除く)

⁴ 本稿におけるデータについては, すべて気象庁提供の地上気象観測データの日次平均気温値 (1969~1998 年/30 年分) を用いた. なお, 年周期を 365 日とするため, 閏年の 2 月 29 日は除いたが, データのスケール及び解析の目的から影響はないものとする.

図 2. 日別標準分散

図 3. 1998 年のデータとサインカーブ

2. 方 法

ノンパラメトリックな手法による季節調整の手法としては Cleveland et al. (1990) があるが、本稿では分散の構造を表現するために、若干アドホックではあるが次の 2 段階推定を採用した。まず、第 1 段階は平均トレンド処理として平均の長期変動と周期変動(季節変動)を柔軟に表現するためにノンパラメトリック回帰モデルによる解析を試みた。そして、第 2 段階として、第 1 段階の回帰残差の平方に対して、ガンマ変量に関する一般化加法モデルを当てはめ、分散の長期トレンドと季節性を抽出した。

第 1 段階では、気温過程の平均の変動を表現するために平滑化スプライン関数を用いて、日次気温データ T_n ($n = 1, 2, \dots, N$) に対して、次の加法モデルをあてはめる。

$$\begin{aligned} T_n &= \mu_n + \varepsilon_n, \\ \mu_n &= s_\mu^{(1)}(n) + s_\mu^{(2)}(\sin(2\pi\psi(n)/365 - \pi/2)) + s_\mu^{(3)}(\cos(2\pi\psi(n)/365 - \pi/2)), \\ \psi(n) &= \begin{cases} 365 & , n \nmid 365 \\ n - 365 \cdot [n/365] & , n \mid 365 \end{cases} \quad [\cdot] : \text{整数部分}, | : \text{整除される}, \nmid : \text{整除されない}. \end{aligned} \quad (1)$$

ここで、 ε_n は日次の平均値 μ_n に対する残差である(ここでは、 ε_n の分散が n に依存しないことを仮定して加法モデルを作成するが、 ε_n の分散がいくらか n に応じて変化するとしても得られる加法モデルは有効であろう)。また、 $s_\mu^{(1)}(\cdot)$, $s_\mu^{(2)}(\cdot)$, $s_\mu^{(3)}(\cdot)$ はスプライン関数を表し、第 1 項は 1 から N の整数列を当てはめることで長期トレンドを表現し、第 2 項、第 3 項の和を用いてノンパラメトリックに年周期成分を表現する。

なお、スプライン関数としては平滑化スプラインを用いた。その導出にあたっては、 $X_{n1} = n$, $X_{n2} = \sin(2\pi\psi(n)/365 - \pi/2)$, $X_{n3} = \cos(2\pi\psi(n)/365 - \pi/2)$ として、次の罰金付残差平方和 (penalized residual sum of squares) を最小にする。

$$PRSS = \sum_{n=1}^N \left(T_n - \sum_{i=1}^3 s_\mu^{(i)}(X_{ni}) \right)^2 + \sum_{i=1}^3 \lambda_i \int_{-\infty}^{\infty} \left(\frac{\partial^2 s_\mu^{(i)}(x_i)}{\partial x_i^2} \right)^2 dx_i.$$

ここで、 λ_i ($i = 1, 2, 3$) は平滑化パラメータである。実際には、 $s_\mu^{(i)}(\cdot)$ ($i = 1, 2, 3$) を $(n+2)$ 個の

B-スプライン基底の線形結合で表し、後退あてはめ(backfitting)アルゴリズムにより $s_{\mu}^{(i)}(\cdot)$ の全てを求める。その際、 $s_{\mu}^{(i)}(\cdot)$ に対応する平滑化行列を $\mathbf{S}_{\mu}^{(i)}$ ($i = 1, 2, 3$) とすると、 $\mathbf{S}_{\mu}^{(i)}$ はそれぞれ λ_i に依存し、そのトレースにモデルの定数項の自由度を考慮し $\text{tr}(\mathbf{S}_{\mu}^{(i)}) - 1$ としたものは等価自由度(equivalent degrees of freedom)と呼ばれ平滑度を間接的に定める。本稿では、等価自由度をそれぞれ 4 としたが、その時適切と思われる滑らかさを持つ回帰関数が得られた。また、解析にあたっては、S-PLUS の gam オブジェクトを用いたが、gam ではデータ数が多い場合、データの予測変数の部分を全て節点にするのではなく、一部を減らして推定する方法が選択される。

なお、モデル (1) の右辺第 2 項と第 3 項の $\sin(\cdot)$ 関数と $\cos(\cdot)$ 関数について、本来、ここで定義された関数 $\psi(\cdot)$ などを用いずに、それぞれ $\sin(2\pi n/365 - \pi/2)$, $\cos(2\pi n/365 - \pi/2)$ とすべきであるが、推定において、プログラムの計算精度によると思われる誤差が生じるケースもあったことから関数 $\psi(\cdot)$ を用いて調整した。これは、計算機の実数精度や有効桁数の設定等によって、 $\sin(2\pi n/365 - \pi/2)$, $\cos(2\pi n/365 - \pi/2)$ の値が厳密に同一にならないときスプライン関数の回帰係数の計算が数値的に不安定になり、回帰関数が正しく推定されないことがあるためである。これは、gam に特有の現象ではなく平滑化スプラインに関するものと考えられ、loess では同様の現象は起こらない。

続いて、第 2 段階では (1) 式の当てはめ残差 $e_n = T_n - \hat{\mu}_n$ の 2 乗 e_n^2 に対して次の対数リンク関数とガンマ分布を想定した一般化加法モデル(Hastie and Tibshirani (1990))を当てはめる。

$$\begin{aligned}\sigma_n^2 &= E[e_n^2], \\ \log \sigma_n^2 &= s_{\sigma}^{(1)}(n) + s_{\sigma}^{(2)}(\sin(2\pi\psi(n)/365 - \pi/2)) + s_{\sigma}^{(3)}(\cos(2\pi\psi(n)/365 - \pi/2)), \\ \psi(n) &= \begin{cases} 365 & , n \nmid 365 \\ n - 365 \cdot [n/365] & , n \nmid 365 \end{cases} \quad [] : \text{整数部分}, | : \text{整除される}, \nmid : \text{整除されない}.\end{aligned}\tag{2}$$

モデル (1) 同様、 $s_{\sigma}^{(1)}(\cdot)$, $s_{\sigma}^{(2)}(\cdot)$, $s_{\sigma}^{(3)}(\cdot)$ は平滑化スプラインを用いて、第 1 項は 1 から N の整数列を当てはめることで長期トレンドを表現し、第 2 項と第 3 項の和で年周期成分を表現する。また、等価自由度も同様にそれぞれ 4 とした。ただし、モデル (2) においてはガンマ分布を想定しているため、罰金付残差平方和を最小にする替わりに罰金付尤度(penalized likelihood)を最大とすることになり、想定分布とリンク関数に応じた作業反応変量(working response)と重みが局所スコア(local-scoring)アルゴリズムにより求められる。

3. 解析結果

3.1. 東京・大手町の日次平均気温データの解析

前述の東京・大手町における日次平均気温データ(図 1/ $n = 1, \dots, 10950$)に対し加法モデル (1) を当てはめた結果は図 4~6 のようになる。図 4 は (1) 式の第 1 項 $s_{\mu}^{(1)}(n)$ の偏回帰関数の推定値(実線)であり、気温過程の平均の長期トレンドを表している。プロット自体は、偏残差プロットとして表現されているため偏残差の Y 平均が 0 となるような中心化が行われており、横軸は n である。なお、破線は偏回帰関数の 95% 近似信頼区間を表している。図 4 からは、過去 30 年で気温が上昇する傾向にあることが見てとれる。

一方、図 5 は関数 $s_{\mu}^{(2)}(\sin(2\pi\psi(n)/365 - \pi/2))$ 、図 6 は関数 $s_{\mu}^{(3)}(\cos(2\pi\psi(n)/365 - \pi/2))$ の偏回帰関数推定値をそれぞれプロットしたもの(破線は 95% の信頼区間)である。図 4 同様、偏残

図 4. $s_{\mu}^{(1)}(n)$ の偏残差プロット (破線は 95% 信頼区間)

図 5. $s_{\mu}^{(2)}(\sin(2\pi\psi(n)/365 - \pi/2))$ の偏残差プロット (破線は 95% 信頼区間)

図 6. $s_{\mu}^{(3)}(\cos(2\pi\psi(n)/365 - \pi/2))$ の偏残差プロット (破線は 95% 信頼区間)

差プロットとして表現されているため偏残差の平均が 0 となるような中心化が行われており、横軸はそれぞれ $\sin(2\pi\psi(n)/365 - \pi/2)$ 及び $\cos(2\pi\psi(n)/365 - \pi/2)$ 値である。また、図 7 は、 $s_{\mu}^{(2)}(\sin(2\pi\psi(n)/365 - \pi/2)) + s_{\mu}^{(3)}(\cos(2\pi\psi(n)/365 - \pi/2))$ の推定値(365 日分)をプロットしたものである。図 7 から常識的ではあるが複雑な年周期パターンがわかる。年周期のパターンとしては、2 月に最低気温のピーク、8 月に最高気温のピークがあり、春にあたる部分の膨らみから、その時期に比較的急に気温が上昇する一方、夏から秋にかけては直線的に気温が下降していく様子が見てとれる。なお、30 年間の μ_n の予測値プロット、すなわち気温過程の平均構造推定値のプロットを図 8 に示す。

次に、原データ(図 1)から平均構造の推定値(図 8)を差し引いた残差(図 9)を 2 乗したデータに対して、一般化加法モデル (2) を適用して解析を行った。

図 10 は、(2) 式の第 1 項 $s_{\sigma}^{(1)}(n)$ の偏回帰関数プロットである(破線は 95%の信頼区間、横軸は n)。スケールは小さいが長期的な変動の構造があることがわかる。一方、図 11 は関数 $s_{\sigma}^{(2)}(\sin(2\pi\psi(n)/365 - \pi/2))$ 、図 12 は関数 $s_{\sigma}^{(3)}(\cos(2\pi\psi(n)/365 - \pi/2))$ の偏残差による推定値をそ

図 7. $s_{\mu}^{(2)}(\sin(2\pi\psi(n)/365 - \pi/2)) + s_{\mu}^{(3)}(\cos(2\pi\psi(n)/365 - \pi/2))$ の偏残差プロット

図 8. 平均構造の推定値

図 9. 残差プロット

図 10. $s_{\sigma}^{(1)}$ の偏残差プロット(破線は 95% 信頼区間)

図 11. $s_{\sigma}^{(2)}(\sin(2\pi\psi(n)/365 - \pi/2))$ の偏残差プロット(破線は 95% 信頼区間)

図 12. $s_{\sigma}^{(3)}(\cos(2\pi\psi(n)/365 - \pi/2))$ の偏残差プロット (破線は 95% 信頼区間)

図 13. $s_{\sigma}^{(2)}(\sin(2\pi\psi(n)/365 - \pi/2)) + s_{\sigma}^{(3)}(\cos(2\pi\psi(n)/365 - \pi/2))$ の偏残差プロット

それぞれプロットしたもの(破線は 95% の信頼区間, 横軸はそれぞれ図 5, 図 6 に同じ)で, また, 図 13 は, $s_{\sigma}^{(2)}(\sin(2\pi\psi(n)/365 - \pi/2)) + s_{\sigma}^{(3)}(\cos(2\pi\psi(n)/365 - \pi/2))$ の偏残差による推定値(365 日分)をプロットしたものである。これにより, 気温の分散に年周期のパターンが顕著にあることがわかる。

年周期のパターンとしては, 春先と梅雨の時期に気温の変動が大きくなり, 比較的冬季の気温の変動は小さいことがわかる。実際, 関東地方では, 春先は三寒四温などと言われ, この点は経験的に言われていることとも合致する。また, 梅雨の時期については, 次第に気温が上昇していく中で, いわゆる梅雨寒の日があったりすることなどとも合致している。Moreno (2000), Cao and Wei (2000), Torró et al. (2001) などでは, アメリカやヨーロッパの気温データ解析において, 分散の値は季節的に冬季に高く夏季に低いという報告がなされているが, 日本の, 特に東京の気温について言えば, 全く様相が異なっているということになる。

なお, 30 年間の推定値 σ_n^2 のプロット, すなわち分散構造の推定値のプロットを図 14 に示す。

図 14. 分散構造の推定値

3.2. 日本各地の日次平均気温データの解析

続いて、東京・大手町以外の地点(気象庁提供の地上気象観測データ 1969～1998 年/30 年分、ただし、閏年の 2 月 29 日は除いた。)についても同様の解析を行ったので、その結果(平均および分散の長期トレンドと年周期成分の偏残差による推定値プロット)を図 15 に示す。それぞれ、(a) 平均長期トレンド(30 年分)、(b) 平均周期成分(365 日分)、(c) 分散長期トレンド(30 年分)、(d) 分散周期成分(365 日分)を表す。ここから、次のような特徴が把握できる。

- (1) すべての地点において、平均気温は長期的に上昇傾向にある。
- (2) 平均の年周期のパターンは地域的な相違はほとんどない。
- (3) 分散の長期的な変動については、各地点毎に異なり、平均のように全国的に上昇傾向にあるといったことはない。
- (4) 分散の年周期のパターンについては、地域的に共通した特徴から次のように分類される。
 - ① [相対的に冬期の分散が大きい地域]このパターンを示すのは主に西日本のエリアで、他に梅雨の時期に一時的に分散が大きくなる特徴があり、W 型のパターンを示す。また、那覇や八丈島といった島部ではほぼ V 型のパターンを示している。
 - ② [相対的に冬期の分散が小さい地域]このパターンを示すのは主に東京を含めた関東地方から長野などのいわば中日本地域にかけてのエリアで、関東地方では、春と梅雨に分散が相対的に大きくなる M 型のパターンを示し、長野や甲府などの中日本地域では梅雨の分散は関東地方ほど小さくなく一山型のパターンを示す。ただ、関東地方にあって銚子などは上記①の W 型のパターンを示しており、八丈島のパターンからすると、海流が何らかの影響を及ぼしている可能性が示唆される。
 - ③ [相対的に冬期・夏期ともに全般的に分散が大きい地域]このパターンを示すのは、主に北日本のエリアで、春・梅雨・秋などの寒候期と暖候期の合間の季節に分散が相対的に小さくなる。

図 15. 主要都市の偏残差プロット

(a) 平均長期トレンド (b) 平均周期分布 (c) 分散長期トレンド (d) 分散周期成分

(6) 新潟

(7) 宇都宮

(8) 前橋

(9) 水戸

(10) 熊谷

図 15. (続き)

(11) 鉾子

(12) 横浜

(13) 三島

(14) 甲府

(15) 長野

図 15. (続き)

(16) 名古屋

(17) 輪島

(18) 大阪

(19) 米子

(20) 広島

図 15. (続き)

(21) 高松

(22) 福岡

(23) 鹿児島

(24) 那覇

(25) 八丈島

図 15. (続き)

次に、高層(850 hpa および 500 hpa の気圧面)のデータ(気象庁提供の 9 時レーウィンゾンデ観測データ 1989~1998 年/10 年分、ただし、閏年の 2 月 29 日を除く)に対しても同様の解析を行った。850 hpa の気圧面は高度約 1500 m 付近で、地表の摩擦や熱の直接の影響がなくなるとされる自由大気の下層であり、一方、500 hpa の気圧面は高度約 5400 m 付近で、気圧と風向・風力の関係など気象の理論的な関係がよく近似される高度とされる。

解析の結果(平均および分散の長期トレンドと年周期成分の偏残差による推定値プロット)を図 16(850 hpa)、図 17(500 hpa)に示す。それぞれ、(a) 平均長期トレンド(10 年分)、(b) 平均周期成分(365 日分)、(c) 分散長期トレンド(10 年分)、(d) 分散周期成分(365 日分)を表す。データ数が 10 年分しかないため、長期トレンドについて地上のデータと比較して明確なことは言えないが、年周期については、平均も分散も地域的なパターンの違いはない。

このことから、分散の年周期のパターンについては、地表の摩擦や熱の直接の影響、すなわち、地形や海洋の影響が大きく影響しているものと考えられる。

3.3. 残差分析

今回の解析は、気温過程の比較的長期の変動をモデル化しているもので数日の短期変動については、特に精緻な分析を行っていない。そこで、大手町のデータ(図 1)について、第 1 段階と第 2 段階から推定された各時点での気温の平均 $\hat{\mu}_n$ と標準偏差 $\hat{\sigma}_n$ を用いて気温系列を

$$\tilde{T}_n = \frac{T_n - \hat{\mu}_n}{\hat{\sigma}_n}$$

と標準化し、この標準化残差時系列データ $\tilde{T}$ の平均構造に ARMA、分散構造に GARCH モデル(Bollerslev (1986))を導入し、次のモデル当てはめを行った。

$$\begin{aligned}\tilde{T}_n &= c + \sum_{i=1}^p \phi_i \tilde{T}_{n-i} + \varepsilon_n - \sum_{i=1}^q \theta_i \varepsilon_{n-i}, \\ \tilde{\sigma}_n^2 &= c^* + \sum_{i=1}^{p^*} \phi_i^* \tilde{\sigma}_{n-i}^2 + \sum_{i=1}^{q^*} \theta_i^* \tilde{\sigma}_{n-i}^2, \quad \tilde{\sigma}_n^2 = E[\varepsilon_n^2] : \text{条件付分散}\end{aligned}$$

なお、 $c, \phi_i, \theta_i, c^*, \phi_i^*, \theta_i^*$ はパラメータである。そして、AIC 規準に基づいて最適なモデル選択を行ったところ、ARMA (2,2) GARCH (2,1) が選ばれた。この標準化残差系列の構造の同定結果から、当然第 1 段階、第 2 段階でのトレンドの推定に用いた平滑化パラメータに強く依存するが、条件付分散変動の構造が示唆され、また、Torró, et al. (2001) らが用いた条件付分散変動構造を取り入れたモデルは、我が国のような複雑な季節構造を持つ気温データにおいては、それなりのトレンド処理を行ってから当てはめるべきものであることは確かと思われる。

4. おわりに

本稿では、気温データの構造について、ノンパラメトリック回帰を用いた解析を行った。その結果、平均気温は長期的には上昇の傾向にあり、またその年周期は単純なサインカーブではないこと、一方、分散についても、長期に変動する構造があり、とりわけ年周期のパターンについては地域的な特徴があり、今回の分析対象となった地点では大きく 3 つの型に分類されることが明

図 16. 高層 850 hpa 面の偏残差プロット
(a) 平均長期トレンド (b) 平均周期分布 (c) 分散長期トレンド (d) 分散周期成分

図 17. 高層 500 hpa 面の偏残差プロット

(a) 平均長期トレンド (b) 平均周期分布 (c) 分散長期トレンド (d) 分散周期成分

らかとなった。これらの特性は、通常のパラメトリック回帰モデルでは把握しにくいものであり、ノンパラメトリック回帰分析の応用上の優位性を示した事例と考える。

実際、ウェザー・デリバティブにおいてリスクの評価を行うためには、分散の構造は無視することはできない。少なくとも、本稿の解析結果からすれば、従来無視されがちな地域的な特性や季節的な特性を取り入れた評価を行うことは必須と考えるし、そのためにも柔軟な気温過程のモデリングは重要である。

なお、本稿においては、時系列の平均と分散の推定に2段階の推定を行ったが、これを段階的に行うのではなく、ノンパラメトリック回帰を用いて統一的に推定する手法の構築に関しては今後の課題と考える。

謝 辞 本稿は、著者が筑波大学大学院 経営・政策科学研究科 経営システム科学専攻に在籍していた時に行った研究を拡張しまとめたもので、同専攻の椿広計教授には、同専攻の在籍時代から本研究の遂行ならびに本稿をまとめるにあたってきめ細かいご指導を頂いた。また、投稿に際しては、中央農業総合研究センターの竹澤邦夫先生からは大変貴重なご意見・ご指導を頂くとともに、査読者の方々からは適切なコメントを頂いた。ここに記して謝意を表したい。

参 考 文 献

- Brennan, M.J. and Schwartz, E. (1982): An equilibrium model of bond pricing and a test of market efficiency. *Journal of Financial and Quantitative Analysis* **17**, 75–100.
- Black, F. and Scholes, M. (1973): The pricing options and corporate liabilities. *Journal of Political Economy* **3**, 637–654.
- Bollerslev, T. (1986): Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* **31**, 307–327.
- Cao, M. and Wei, J. (2000): Equilibrium Valuation of Weather Derivatives. working paper, University of Toronto (<http://www.rotman.utoronto.ca/~wei/research/hddcdd.pdf>).
- Cleveland, R.B., Cleveland, W.S., McRae, J.E. and Terpenning, I. (1990): STL: A seasonal-trend decomposition procedure based on loess. *Journal of Official Statistics* **6**, 3–73.
- Cox, J.C., Ingersoll, J. and Ross, S. (1985): A theory of the term structure of interest rates. *Econometrica* **53**, 385–407.
- Dischel, B. (1999): The Fledgling Weather Market Takes Off – Part 5: The D1 Stochastic Temperature Model for Valuing Weather Futures and Options. *Applied Derivatives Trading*, April.
- Hastie and Tibshirani (1990): *Generalized Additive Models*. Chapman and Hall, London.
- Moreno, M. (2000): Riding the Temp. Published in FOW, Special Weather Derivatives Supplement (December).
- Torró, H., Meneu, V. and Valor, E. (2001): Single Factor Stochastic Models with Seasonality Applied to Underlying Weather Derivative Variables. working paper in European Financial Management Association.
- Vasicek, O. (1977): An equilibrium characterization of the term structure. *Journal of Financial Economics* **5**, 177–188.

(2003年5月30日受付 10月21日最終修正 11月4日採択)

著者連絡先：〒106-8569 東京都港区南麻布4-6-7
総合研究大学院大学 複合科学研究科 統計科学専攻
若浦 雅嗣

Analysis of Mean and the Variance of Temperature Processes in Japan Using Generalized Additive Models

Masatsugu Wakaura

Department of Statistical Science, The Graduate University for Advanced Studies,
4–6–7, Minami-Azabu, Minato-ku, Tokyo 106–8569, Japan

Abstract

With pricing weather derivatives, it is important to investigate the structures of the mean and variance of the temperature processes. However, complicated features of variation in temperature cannot be modeled in terms of a parametric statistical model. This paper investigates the features of temperature processes in Japan using nonparametric regression. Seasonal-trend decomposition using generalized additive models shows the structure of the mean and the variance clearly. In particular, it is proved that seasonal periodicities in the variance vary among areas due to their different climatic characters.

Key words: nonparametric regression, seasonal-trend decomposition, smoothing splines, weather derivatives.

(Masatsugu Wakaura).

Received May 30, 2003; Received in final form October 21, 2003; Accepted November 4, 2003.