

Box-Cox 変換と ARCH(1) による日最大風速の時系列データに含まれる分散不均一性の分析

法政大学 正会員 北川徹哉

1. はじめに 各種の時系列データにおいては、ある時間帯で大きな変動が連続的に発生し、別の時間帯では小幅な変動が持続するといった分散不均一性がみられることがある。本研究においては、日最大風速の時系列データに内在する分散不均一性を Box-Cox 変換¹⁾と ARCH(1)²⁾を併用して調べることを提案する。その前半部分を昨年度の年次大会において報告しており、本稿では主に後半部分について述べる。

2. 分散不均一性の分析方法 まず、昨年度に示した内容を大まかに説明する。 y 年における元旦からの日数を t とし、ある観測地点の日最大風速³⁾を $U_{y,t}$ (> 0) とする。日最大風速の時系列データは非正規過程であることが多く、まず $U_{y,t}$ に Box-Cox 変換を適用 ($\hat{U}_{y,t}$: 式(1)) して正規過程に近づけ、 $\hat{u}_{y,t} = (\hat{U}_{y,t} - \hat{M}_y) / \hat{S}_y$ のように標準化する。ここで、 $\hat{M}_y$ と $\hat{S}_y$ はそれぞれ $\hat{U}_{y,t}$ の標本平均値と標本標準偏差である。さらに、 $\hat{u}_{y,t}$ からドリフト $\hat{m}_{y,t}$ ($\hat{u}_{y,t}$ の 11 次までのフーリエ級数展開) を差し引いたものを偏差時系列 $\hat{\varepsilon}_{y,t}$ とし、 $\hat{\varepsilon}_{y,t}$ に ARCH(1) (式(2)~(5)) を適用する。式(2)の $\hat{\sigma}_{y,t}^2$ が $\hat{\varepsilon}_{y,t}$ の ARCH(1) 分散の時系列であり、 ω_y は一定成分、 α_y は変動成分 ($\hat{\varepsilon}_{y,t-i}^2$ への依存度) を表している。例えば図 1~3 はそれぞれ 1994 年に東京管区気象台 (以降、東京) で観測された $U_{1994,t}$ とその $\hat{\varepsilon}_{1994,t}$ 、そして $\hat{\sigma}_{1994,t}^2$ である。図 3 を見ると $\hat{\sigma}_{1994,t}^2$ は年間にわたって不均一であり、 $\hat{\varepsilon}_{y,t}$ は冬季に分散が大きく、夏季には小さいといった季節変動性が観察される。

以降では上で示した方法をさらに進め、日最大風速 $U_{y,t}$ の分散不均一性を求める。上記の $\hat{U}_{y,t}$ から $\hat{\varepsilon}_{y,t-1}$ を求める過程を逆方向にたどると、 $\hat{U}_{y,t}$ の平均値 $\hat{\mu}_{y,t}$ と分散 $\hat{\varsigma}_{y,t}^2$ はそれぞれ式(6)、(7)のように表すことができる。また、 $U_{y,t}$ と $\hat{U}_{y,t}$ は Box-Cox 変換 (式(1)) を介して互いに変数変換の関係にあることから、両者の確率密度関数をそれぞれ $P(U_{y,t})$ ならびに $p(\hat{U}_{y,t})$ とすると $P(U_{y,t})dU_{y,t} = p(\hat{U}_{y,t})d\hat{U}_{y,t}$ が成り立つ。この関係より、 $P(U_{y,t})$ が式(8)のように導出される。ただ

$$\hat{U}_{y,t} = \begin{cases} \frac{U_{y,t}^{\lambda_y} - 1}{\lambda_y} & (\lambda_y \neq 0) \\ \ln U_{y,t} & (\lambda_y = 0) \end{cases} \quad (1)$$

$$\hat{\sigma}_{y,t}^2 = \omega_y + \alpha_y \hat{\varepsilon}_{y,t-1}^2 \quad (2)$$

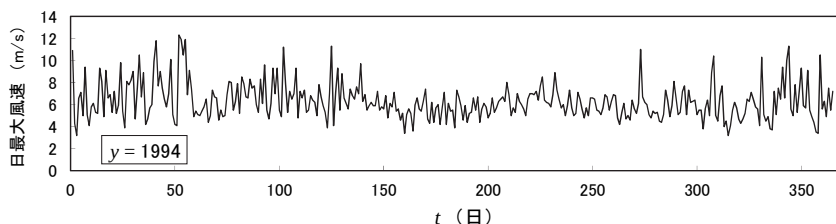
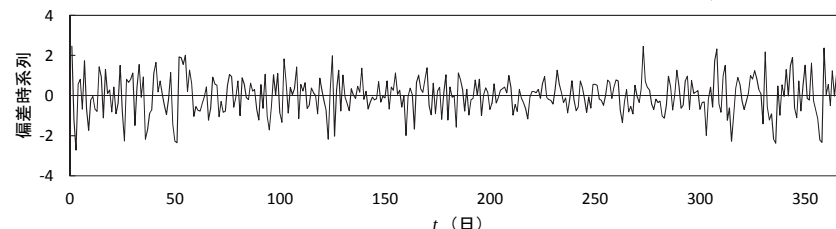
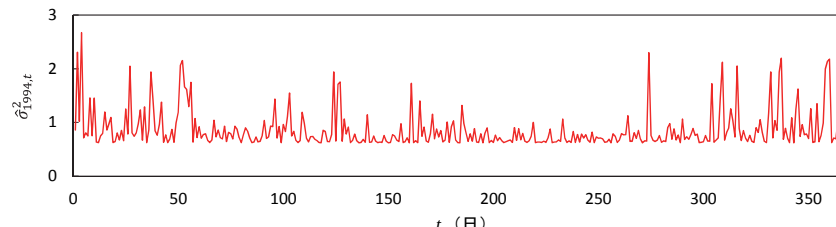
$$\hat{\varepsilon}_{y,t} = \hat{\sigma}_{y,t} v_{y,t} \quad (3)$$

$$v_{y,t} \sim N(0, 1) \quad (4)$$

$$\omega_y > 0, \quad 0 \leq \alpha_y < 1 \quad (5)$$

$$\hat{\mu}_{y,t} = \hat{M}_y + \hat{S}_y \hat{m}_{y,t} \quad (6)$$

$$\hat{\varsigma}_{y,t}^2 = (\hat{S}_y \hat{\sigma}_{y,t})^2 \quad (7)$$

図 1 1994 年・東京における日最大風速の時系列 $U_{1994,t}$ 図 2 1994 年・東京における日最大風速の偏差時系列 $\hat{\varepsilon}_{1994,t}$ 図 3 図 2 の ARCH(1) 分散 $\hat{\sigma}_{1994,t}^2$ の時系列

$$P(U_{y,t}) = \begin{cases} \frac{1}{\sqrt{2\pi\hat{\varsigma}_{y,t}^2}} \exp \left\{ - \left(\frac{U_{y,t}^{\lambda_y} - 1}{\lambda_y} - \hat{\mu}_{y,t} \right)^2 / 2\hat{\varsigma}_{y,t}^2 \right\} U_{y,t}^{\lambda_y - 1} & (\lambda_y \neq 0) \\ \frac{1}{\sqrt{2\pi\hat{\varsigma}_{y,t}^2}} \exp \left\{ - (\ln U_{y,t} - \hat{\mu}_{y,t})^2 / 2\hat{\varsigma}_{y,t}^2 \right\} U_{y,t}^{-1} & (\lambda_y = 0) \end{cases} \quad (8)$$

キーワード 時系列データ, 日最大風速, 分散不均一性, ARCH, Box-Cox 変換

連絡先 〒102-8160 東京都千代田区富士見 2-17-1 法政大学人間環境学部 TEL 03-3264-4921

し、 $p(\hat{U}_{y,t})$ が完全に正規分布(平均値 $\hat{\mu}_{y,t}$ 、分散 $\hat{\sigma}_{y,t}^2$)となっていることが前提である。式(8)の1次モーメントおよび2次モーメントを計算すれば、 $U_{y,t}$ の各 t における平均値と分散が求まる。

3. 分散不均一性の分析結果 前章において示した手法を各観測地点(116箇所)・各年(地点による)における日最大風速に適用し、平均値と分散の時系列を求めた。例として、図4に1994年・東京における日最大風速(図1)の分散時系列を赤色実線で示す。分散の t に応じた変化が示されており、特に冬季において日最大風速そのものの標本分散(約2.74; 黒色実線)を上回る頻度が高く、雨季から秋にかけて下回る日が続く。図4の分散から標準偏差を求め、平均値±標準偏差の時系列を図5に赤色実線で示す。ここで、平均値の時系列を赤色破線で、日最大風速の時系列(図1)を黒色実線で示している。本手法によって推定された平均値±標準偏差はおおむね日最大風速の変動と連動して推移している。しかしながら、例えば $t=273$ 日の台風26号にともなう強風に着目すると、本手法によって推定された平均値+標準偏差は小さく評価されている。これに対して図6の同年・甲府の場合、 $t=273$ 日の台風26号

による日最大風速の変動はその前後の変動とさほど変わらず、本手法による平均値+標準偏差はこれをおおむね評価できている。一方、図7の洲本の場合、台風26号による日最大風速の変動は非常に大きく、本手法による平均値+標準偏差は過小評価となっている。以上のように、本手法は孤立した大きなスパイク状の変動については過小評価となる。最後に、本検討対象の全ての観測地点・年の分散時系列の標本平均値と日最大風速そのものの標本分散との比較を図8に示す。本手法による分散時系列の標本平均値は日最大風速の標本分散よりも全体的に小さく、両者の比はおよそ0.77:1であった。

参考文献 1) Box, G. E. P. and Cox, D. R. : An analysis of transformations, *Journal of Royal Statistical Society, Series B*, Vol.26, pp. 211-252, 1964. 2) Engle, R. F. : Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, Vol. 50, No. 4, pp. 987-1007, 1982. 3) 気象庁ホームページ : <http://www.jma.go.jp/jma/index.html>, 2022.

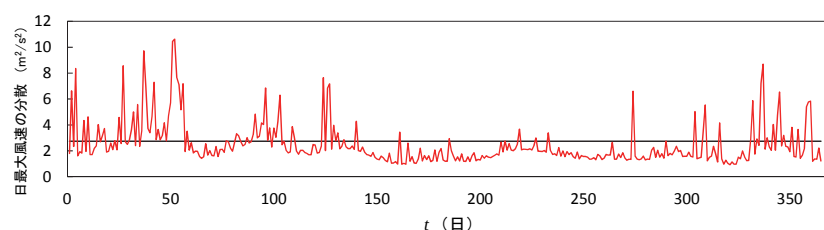


図4 1994年・東京の日最大風速(図1)を対象として本手法によって求めた分散の時系列(赤色実線)と標本分散(黒色実線)との比較

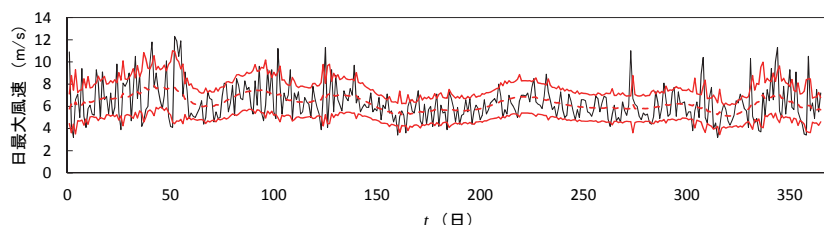


図5 1994年・東京の日最大風速(黒色実線; 図1)を対象として本手法によって推定された平均値の時系列, 平均値±標準偏差の時系列(それぞれ赤色破線, 赤色実線)

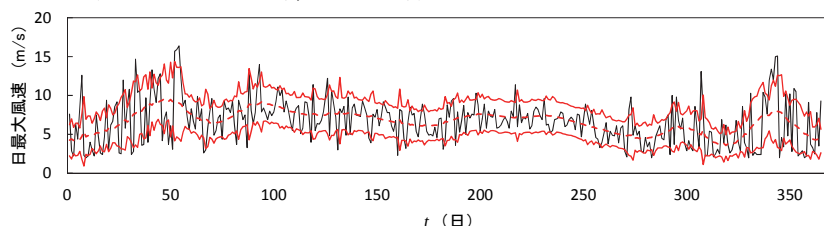


図6 1994年・甲府の日最大風速(黒色実線)を対象として本手法によって推定された平均値の時系列, 平均値±標準偏差の時系列(それぞれ赤色破線, 赤色実線)

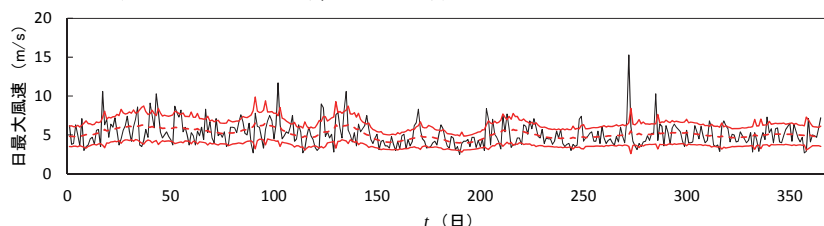


図7 1994年・洲本の日最大風速(黒色実線)を対象として本手法によって推定された平均値の時系列, 平均値±標準偏差の時系列(それぞれ赤色破線, 赤色実線)

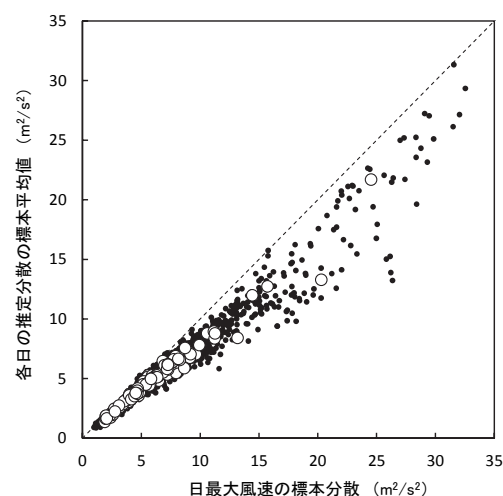


図8 本手法によって推定された各日の分散の標本平均値と日最大風速の標本分散との比較(●:各地点・各年の値, ○:地点ごとの年に関する平均値, 破線:完全相関を表す直線)