

東京における日照と雲量と降水に関する 日毎の自己回帰性の検証 —東京オリンピックに向けて—

瀧川 綾乃¹ Lim Jing Zhi² 鈴木 寿³

概要：気象庁公開の北緯 35 度 41.5 分，東経 139 度 45.0 分標高 25.5[m]（東京）における 2001 年 6 月 1 日から 2001 年 9 月 1 日までの日合計全天日射量（日照），日平均雲量（雲量），降水量の日合計（降水）の各時系列データについて自己回帰性を検証した。観測日 i は開始日を 1 とする序数，また $x(i)(i = 1, \dots, 93)$ は観測日 i における日照を表すとき，次数 n の自己回帰モデル： $i = 1, \dots, 93$ に対し $x(i) = a(1)x(i-1) + \dots + a(n)x(i-n) + e(i)$ （便宜上 $x(-n) = \dots = x(-1) = 0$ とおく）を考え， $n \in 2, \dots, 93-1$ 毎に推定誤差 $E(n)$ が最小値をとるように係数 $a(1), \dots, a(n)$ を選んだうえで，推定誤差がそれ以上小さくならない次数 n を求めたところ， $n = 27$ が得られ，同様に雲量に関し次数 25，降水に関しては次数 16 が得られた。

1. はじめに

気象予報の分野において気象の自己回帰性を利用し，モデル化する試みがおこなわれている。天候を自己回帰モデルで表現するには，モデルの次数を決定する必要があり，赤池の情報量基準 [1]，リスク関数 [2] に基づいて決定する方法や，二分したデータの正準相関を用いる方法 [3] が提案されている。本論文では，誤差の自乗平均の非負平方根（推定誤差）を基準として用いて，東京における日照，雲量，降水の時系列データに当てはめ自己回帰次数を求めた。そして，求めた次数が妥当であることを統計的仮説検定を用いて検証した。

2. 自己回帰モデル

定常確率過程 $\{x(t); t = \dots, -1, 0, 1, 2, 3, \dots\}$ の時刻 $t-1$ までの観測値にのみ基づき次の時刻 t における値を予測することを考える。

$$x(t) = \sum_{i=1}^m a_i x(t-i) + e(t) \quad (1)$$

式 (1) を満たすとき，互いに全く独立な動きをする確率変数の系列が確率変数 $x(t)$ の系列に影響することを表す。この過程を自己回帰過程（autoregressive process），単に AR

過程と呼び， m は AR 過程の次数， a_i は i 次の自己回帰係数と呼ぶ。また，誤差項 $e(t)$ は平均 0，分散 σ^2 の正規分布に従うホワイトノイズである。 $e(t)$ の仮定より， m 次元確率変数 $x_m = (x(1), \dots, x(m-1))^T$ の確率分布は，

$$V_m = \begin{bmatrix} R_{xx}(0) & R_{xx}(1) & \dots & R_{xx}(m-1) \\ R_{xx}(1) & R_{xx}(0) & \dots & R_{xx}(m-2) \\ \vdots & \vdots & \ddots & \vdots \\ R_{xx}(m-2) & \dots & R_{xx}(0) & R_{xx}(1) \\ R_{xx}(m-1) & R_{xx}(m-2) & R_{xx}(1) & R_{xx}(0) \end{bmatrix} \quad (2)$$

を自己共分散行列として持つ m 次元正規分布であり，その確率密度関数は

$$h(x_m | \sigma^2, a_m) = \frac{\exp\{-\frac{1}{2}x_m^T V_m^{-1} x_m\}}{\sqrt{(2\pi)^m (\det V_m)}} \quad (3)$$

で与えられる [4]。

2.1 推定誤差

式 (1) の AR モデルに対し $n \in \{2, \dots, 93-1\}$ 毎に $e(\cdot)$ の自乗平均の非負平方根

$$E(n) = \sqrt{\frac{e(n+1)^2 + \dots + e(93)^2}{93-n}} \quad (4)$$

を定義する。これを次数 n における推定誤差と呼び，作成した AR モデルから求めた予測値と実測値の誤差を残差と呼ぶ。

¹ 中央大学大学院理工学研究科情報工学専攻
² 中央大学大学院理工学研究科電気・情報系専攻
³ 中央大学理工学部情報工学科

3. 手法

東京における 2001 年 06 月 01 日から 2001 年 09 月 01 日までの日照、雲量、降水の各時系列データ [?] *1 に対し、推定誤差 $E(n)$ が最小となる自己回帰係数を設定する。作成した AR モデルに対し、統計的仮説検定を用いて次数が妥当であると確認する。

3.1 データ形式

利用したデータは以下の形式である。

- 雲量の最小値は 0, 最大値は 10. 露場において観測者が目視によって計測した。
- 日照の最小値は 0, 最大は 24. 測風塔または屋上において、回転式日射計または日照計を用いて計測した。
- 降水量の最小値は 0. 露場において転倒ます型雨量計, または感雨器を用いて計測した。

降水や日照のデータに関して、観測器の不調や日照現象が発生せず、欠損した場合には前後の値から線形補完する。データの観測中に観測環境の変化はない。

3.2 推定誤差が最小であることの検証方法

作成した n 次の AR モデルの残差に対し、 $n+1$ 次の AR モデルの残差との有意差の有無を仮説検定を用いて検定する。仮説検定は以下の手順でおこなう。

- (1) AR モデルの次数 n を決定する。
- (2) 作成した AR モデルの残差を求める。
- (3) $n+1$ 次の AR モデルを作成し、同様に残差を求める。
- (4) 仮説検定に用いる二つの変数のうち (2) を変数 1, (3) を変数 2 として t 検定をおこなう。
- (5) n を増やし、(1) から (4) を繰り返す。

帰無仮説を「二群の平均は等しい」、対立仮説を「二群の平均は等しくない」とし、検定をおこなう。有意水準 α は 0.05 とする。

4. 結果

日照 ($n = 27$), 雲量 ($n = 25$), 降水 ($n = 16$) の場合の検定結果をそれぞれ表 4, 表 5, 表 ?? に示す。日照に関して $n = 27$ のとき、検定結果より $n+1$ 次の AR モデルの残差に対し有意差が認められない。よって、残差が最小になる次数は 27 次であるといえる。同様に雲量に関して次数 25, 降水に関して次数 16 が得られた。以上の意味で、東京では日照に関し 27 日, 同様に雲量 25 日, 降水 16 日の程度の自己回帰性があるといえる。

5. おわりに

本論文では、推定誤差を用いて東京における日照、降水、

表 1 統計値: 日照

項目	変数 1	変数 2
平均	2.05	2.06
分散	2.65	2.53
観測数	66	65

表 2 統計値: 雲量

項目	変数 1	変数 2
平均	1.74	1.74
分散	1.36	1.42
観測数	68	67

表 3 統計値: 降水

項目	変数 1	変数 2
平均	4.244	4.29
分散	65.8	66.2
観測数	77	76

表 4 検定結果: 日照

項目	値
仮説平均との差異	0.0217
自由度	129
t	-0.106
P(T<=t) 両側	0.915
t 境界値 両側	1.98

表 5 検定結果: 雲量

項目	値
仮説平均との差異	0.00807
自由度	133
t	$-2.29 * 10^5$
P(T<=t) 両側	1.00
t 境界値 両側	1.98

表 6 検定結果: 降水

項目	値
仮説平均との差異	0.051
自由度	151
t	-0.0776
P(T<=t) 両側	0.938
t 境界値 両側	1.98

雲量の各項目に対して自己回帰性を検証した。なお、自己回帰性が観測される物理現象の機序を気象学的に解明することは、別途の課題である。

参考文献

- [1] H.Akaike, "Information Theory and an Extension of the Maximum Likelihood Principle," *E.Parzen, K.Tanabe, G.Kitagawa(eds) Selected Papers of Hirotugu Akaike. Springer Series in Statistics (Perspectives in Statistics)*. Springer, New York, NY.
- [2] P.Gates, H.Tong, "On Markov Chain Modeling to Some Weather Data," *Journal of Applied Meteorology*, Vol.15, Nov. 1976.
- [3] H.Tong, "Autoregressive Model Fitting with noisy data by Akaike's information criterion," *IEEE Transactions on Information Theory*, pp. 476-480, Jul. 1975.
- [4] 尾崎統 北川源四郎, "時系列解析の方法," 統計科学選書, 朝倉書店, 1998.

*1 気象庁 | 過去のデータ・ダウンロード: <https://www.data.jma.go.jp/gmd/risk/obsdl/>

正誤表

下記の箇所に誤りがございました。お詫びして訂正いたします。

訂正箇所	誤	正
1 ページ 2 行目	日合計全天日射量	日照時間
2 ページ 2 行目	各時系列データ[*] ^{*1}	各時系列データ ^{*1}
2 ページ 3.1 節 5 行目	回転式日射計	回転式日照計
2 ページ 3.2 節 10, 11 行目	「二群の平均は等しい」 「二群の平均は等しくない」	“二群の平均は等しい” “二群の平均は等しくない”
2 ページ 4 章 2 行目	表??	表 6
2 ページ 表 5	$-2.29 \cdot 10^5$	-2.29×10^5
2 ページ 表 5	P (T<=t) 両側	P(T<=t) 両側
2 ページ 表 6	P (T<=t) 両側	P(T<=t) 両側
2 ページ 脚注	*1 気象庁 過去のデータ・ダウンロード: https://www.data.jma.go.jp/risk/obsdl/	*1 気象庁 過去のデータ・ダウンロード: https://www.data.jma.go.jp/risk/obsdl/ (2019/11/1 ダウンロード)