

第IV部門

時空間モデルを用いたマレーシアにおけるヘイズ災害のシミュレーション

京大大学院	学生会員	○尾形 誠一郎
京大大学院	学生会員	松田 明広
京大大学院	正会員	松島 格也
京大大学院	フェロー	小林 潔司

1. はじめに

東南アジア地域におけるヘイズ災害は多方面に重大な影響を及ぼす問題として捉えられている。本研究では、最初にヘイズの構成物質に対し、1997年前期(1月－6月)において平常時におけるモデルを構築する。それをもとに1997年後期(7月－12月)にかけてインドネシアで発生した大規模な森林火災に対して、構築された平常時のモデルを用いシミュレーションを行う。近年の研究で時系列データの多くが長期記憶性を持つと報告されている。そのため、シミュレーションには短期記憶性と長期記憶性を持つ二つのモデルを用いる。

2. Seasonal Space-Time ARIMA Model

STARIMAモデルは異時点・異地点間において観測されたデータの依存関係を記述する。このモデルは、一変数時系列のARIMAモデルの空間への拡張であり、時点 $t = 1, 2, \dots, T$ 及び観測地点 $i = 1, 2, \dots, N$ において、空間と時間の双方でラグを取った観測値と誤差の重み付け線形式として表現される。この表現の基本的メカニズムは各地点の近傍の階層性と空間重み付け行列による影響である。またSTARIMAモデルを拡張し、 s 期はなれた変数間にSTAR, STMA過程をあてはめたモデルがSeasonal STARIMAモデルである。

$\mathbf{z}_t$ を t 期における N ヶ所の観測地点の観測値で構成される $N \times 1$ の行列とする。このとき、STARIMAモデルは、

$\Phi_{P,\Lambda}(B^S)\phi_{p,\lambda}(B)\nabla_S^D\nabla^d(\mathbf{x}_t - \mu) = \Theta_{Q,M}(B^S)\theta_{q,m}(B)\epsilon_t$ と表現される。なお、

$$\begin{aligned}\Phi_{P,\Lambda}(B^S) &= I - \sum_{k=1}^P \sum_{l=0}^{\Lambda_k} \Phi_{kl} W_l B^{kS} \\ \phi_{p,\lambda}(B) &= I - \sum_{k=1}^p \sum_{l=0}^{\lambda_k} \phi_{kl} W_l B^k \\ \Theta_{Q,M}(B^S) &= I - \sum_{k=1}^Q \sum_{l=0}^{M_k} \Theta_{kl} W_l B^{kS} \\ \theta_{q,m}(B) &= I - \sum_{k=1}^q \sum_{l=0}^{m_k} \theta_{kl} W_l B^k\end{aligned}$$

である。

この時、 $\Phi_{k,l}, \phi_{k,l}$: 時間ラグ k 及び空間ラグ l における季節・非季節の自己回帰パラメータ、 $\Theta_{k,l}, \theta_{k,l}$: 時間ラグ k 及び空間ラグ l における季節・非季節の移動平均パラメータ、 P, p : 季節・非季節の自己回帰オーダー、 Q, q : 季

節・非季節の移動平均オーダー、 Λ_k, λ_k : k 番目の自己回帰項における季節・非季節の空間オーダー、 M_k, m_k : k 番目の移動平均項における季節・非季節の空間オーダー、 W_l : l 番目の空間オーダーにおける $N \times N$ の空間重み付け行列、 D, d : それぞれ季節・非季節で差分をとる回数、 ∇_S^D, ∇^d : 季節・非季節の差分オペレータであり、季節ラグ S に対して $\nabla_S^D = (I - B^S)^D, \nabla^d = (I - B)^d$ と定義される。ここで B は、 $N \times N$ のラグオペレータであり $B^k \mathbf{x}_t = \mathbf{x}_{t-k}$ である。また ϵ_t は時点 t における正規分布を取るランダム誤差項である。

Seasonal STARIMAモデルは単一時系列に対するARMAモデルの空間への拡張であり、モデル構築プロセスはARMAのモデル構築に用いられるBox-Jenkins法による。Box-Jenkins法によるモデル構築は大きく分けてモデルの同定、パラメータの推計、推計されたモデルの検定の三つのプロセスからなる。この三つのプロセスを繰り返し、診断検定をパスするモデルが求めれば、そのSeasonal STARIMAモデルを用いて予測を行うことができる。

3. Seasonal Space-Time ARFIMA Model

ARFIMAモデルのコンセプトは、整数差分パラメータ d を一般化し、実数差分パラメータ d を導入したものである。ラグオペレータ B と実数差分オペレータは、

$$\nabla^d = (1 - B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k$$

のようになる。ARFIMA(0, d , 0)過程において $-1/2 < d < 1/2$ を満たせば、系列 $\{x_t\}$ は定常でありかつ以下のように反転可能である。

$$\begin{aligned}x_t &= \psi(B)\epsilon_t = \sum_{k=0}^{\infty} \psi_k \epsilon_{t-k} = \sum_{k=0}^{\infty} \frac{\Gamma(k+d)}{\Gamma(d)\Gamma(k+1)} \epsilon_{t-k} \\ \pi(B)x_t &= \sum_{k=0}^{\infty} \pi_k x_{t-k} = \sum_{k=0}^{\infty} \frac{\Gamma(k-d)}{\Gamma(-d)\Gamma(k+1)} x_{t-k} = \epsilon_t\end{aligned}$$

ここで $\Gamma(\cdot)$ はガンマ関数である。Seasonal fractional STARIMAモデルは、 $\psi_j = \sum_{i=0}^j \psi_0(d, i) \psi_s(d_s, j-i)$ を用いることによって以下のようになる。

$$\begin{aligned}(\mathbf{x}_t - \mu) &= \frac{\Theta_{Q,M}(B^S)\theta_{q,m}(B)}{\Phi_{P,\Lambda}(B^S)\phi_{p,\lambda}(B)} \sum_{j=0}^{\infty} \sum_{i=0}^j \psi_0(d, i) \psi_s(d_s, j-i) \epsilon_{t-j}\end{aligned}$$

ここで、 $\psi_0(d, j) = \Gamma(j+d)/\Gamma(d)\Gamma(j+1)$ 、また $\psi_s(d, j) = \psi_0(d, j/s)$ ($j = 0, s, 2s, \dots$) ; $= 0$ ($j \neq 0, s, 2s, \dots$) である。差分パラメータとSTARMAパラメータは、 $\mathbf{x}_t = \epsilon_t = 0$

表1. 推計されたパラメータ

Seasonal STARFIMA model

Parameter	d	d_s	ϕ_{10}	ϕ_{20}	ϕ_{11}	ϕ_{21}	ϕ_{21}
Estimation	0.25415	0.19703	0.24124	0.04368	0.19957	0.04089	-0.07189
F-Value	-	-	1971.13	136.252	1348.68	33.3433	196.353
					CSS	σ^2	AIC
					-287239	394.86	574495

Seasonal STARIMA model

Parameter	ϕ_{10}	ϕ_{20}	ϕ_{30}	ϕ_{11}	ϕ_{10}	ϕ_{20}	ϕ_{20}	ϕ_{40}
Estimation	0.47247	0.10433	0.05302	0.16901	0.08560	0.02927	0.00695	0.01177
F-Value	15295.0	430.669	177.366	986.059	422.324	38.5506	2.69651	7.13110
Parameter	ϕ_{30}	ϕ_{60}	ϕ_{70}	ϕ_{10}	Likelihood	σ^2	AIC	
Estimation	0.01464	0.00552	0.01106	0.9474	-287926	403.28	575889	
F-Value	27.3599	2.32654	16.3007	3939.89	-	-	-	

($t \leq 0$)のもとで条件付最小二乗法(CSS)によって推計される。最もよく系列を表現するSTARFIMAモデルは、CSSとパラメータ数から求められる赤池情報量基準(AIC)を最小化するものとして定義される。

4. ケーススタディ

本研究では、半島マレーシア西海岸の15地点で観測された微小粒子状物質(PM10)の時間毎平均データを用いる。PM10は1997年ヘイズ災害においてもやを引き起こした主要物質であり、同時にマレーシアで用いられている大気汚染指数(API)の主要なsub-indexでもある。1997年を前期と後期に分ける。前期を平常な状態として定義し、この期間(1月-6月)の観測結果を用いてモデルを構築する。そして、そのモデルを用い大規模ヘイズ災害が発生した期間を含む後期(7月-12月)の予測を行いモデルの当てはまりを見る。

前期のデータによって構築されたseasonal STARIMAモデルは $(3_{100}, 0, 0) \times (7_{0...0}, 1, 1_0)_{24}$ のオーダーを持ち、式で表すと以下である。

$$(I - \sum_{k=1}^3 \phi_{k0} W_0 B^k - \phi_{11} W_1 B)(I - \sum_{k=1}^7 \Phi_{k0} W_0 B^{24k}) \times (I - B^{24})(\mathbf{x}_t - \mu) = (I - \Theta_{10} W_0 B^{24}) \varepsilon_t$$

seasonal STARFIMAモデルは、 $(2_{11}, d, 0) \times (1_0, d_s, 0)_{24}$ のオーダーを持ち、式で表すと以下である。

$$(I - \sum_{k=1}^2 \phi_{k0} W_0 B^k - \sum_{k=1}^2 \phi_{k1} W_1 B^k)(I - \Phi_{10} W_0 B^{24}) \times (I - B)^d (I - B^{24})^{d_s} (\mathbf{x}_t - \mu) = \varepsilon_t$$

各パラメータの推計値と統計量は表1の通りである。結果を見ると、どちらのモデルにおいても空間項が含まれる。STARFIMAモデルは短期的な変動と長期的な変動の両方を反映する。そのためSTARMA過程のパラメータ数を減らすことができていく。長期的な変動は、二つの実数差分オペレータ $(I - B)^d$ と $(I - B^{24})^{d_s}$ により表現される。

両方のモデルを用いて24時間予測を行った結果を図1に示す。また、その結果から得られる24時間平均二乗残差とPM10観測値から得られる分散、24時間平均について前期のデータより求められた標準偏差 σ を用

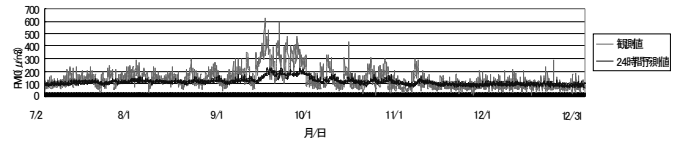
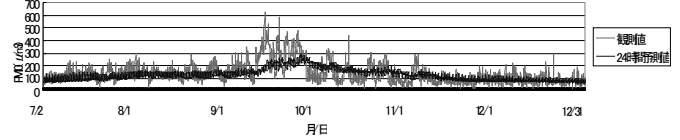
 STARFIMA $(2_{11}, d, 0)(1_0, d_s, 0)_{24}$

 STARIMA $(3_{100}, 0, 0)(7_0, 1, 1_0)_{24}$


図-1: PM10予測値と観測値

1997年後期 地点CA012(Kuala Lumpur)

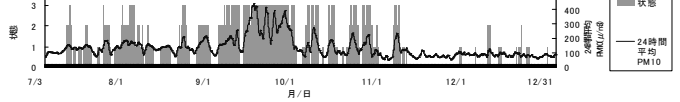
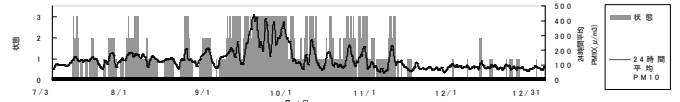
 STARFIMA $(2_{11}, d, 0)(1_0, d_s, 0)_{24}$

 STARIMA $(3_{100}, 0, 0)(7_0, 1, 1_0)_{24}$


図-2: 状態(0-3)とPM10観測値24時間平均

1997年後期 地点CA012(Kuala Lumpur)

いて 2σ 区間を求める。後期において、その三つ全てが区間外の場合を状態3、二つを状態2、一つを状態1、全てが区間内の場合を状態0として図-2に示す。状態が0ということは、観測値と予測誤差の両方がヘイズの発生していない前期と同じ水準にあることを示す。図-1のようにヘイズ災害発生時においてはSTARIMAモデルのほうが当てはまりがよい。しかし特に長時間予測においては、ヘイズが発生した後図-2のようにSTARFIMAモデルを用いたほうが予測値との二乗残差を含む状態が早く0に戻る。このことは長期記憶の影響で過去の観測値や誤差の影響を反映しているからだと思われる。そのためヘイズが収まり大気汚染の状態が平常に戻った後、長期記憶モデルを用いて長期予測を行った場合、短期記憶モデルを用いた場合に比べ早く精度の良い予測値が得られるようになる。

5. 結論

本研究では、時間と空間両方に依存し季節性を持つ長期記憶モデルを用い、半島マレーシアにおけるヘイズ災害のシミュレーションを行い結果を示した。また、短期記憶モデルを用い同様にシミュレーションを行い、両モデルの比較を行った。