

原 著

季節自己回帰和分移動平均モデルによる と畜検査データの時系列分析法

足立泰基^{1)†} 蒔田浩平²⁾

1) 北海道釧路保健所 (〒084-0925 釧路市新野 28 番地)

2) 酪農学園大学獣医学群獣医学類衛生環境学分野 (〒069-8501 江別市文京台緑町 582)

(2014 年 7 月 16 日受付・2014 年 12 月 22 日受理)

要 約

多くの食肉衛生検査所は、衛生的な食肉生産を推進するために生産者にデータ還元を行っているが、と畜検査データの統計処理に関する報告は少ない。時系列分析は、公衆衛生分野を含む医学生物学の幅広い分野で用いられており、廃棄率の顕著な変化から疾病のまん延や薬物治療の有効性を判定するのに利用できる。今回、季節自己回帰和分移動平均 (SARIMA) モデルを用いて廃棄率データを分析するための好適条件を検討した。10 農場からの搬入豚の 4 疾病による廃棄率データについて、SARIMA モデルの当てはまりと対照モデルである指数平滑法の当てはまりを比較した。39 モデルで対照モデルより逸脱度が小さく、27 モデルで有意であった。本法は生産者に対して農場の衛生状態の合理的な判断指標を提供することによって、衛生的な食肉生産を促進するツールとして利用可能である。

——キーワード：と畜検査データ, SARIMA モデル, 時系列分析。

----- 日獣会誌 68, 189~197 (2015)

と畜場に搬入された家畜は、食肉衛生検査機関の獣医師であると畜検査員によって検査され、疾病にかかり、若しくは異常があり食用に供することができないと認められた時等は、と殺禁止・全部廃棄・部分廃棄のいずれかの措置がとられる。北海道では、と畜検査結果は、希望する生産者に送付還元されているが、還元データが、臓器廃棄数 (率) の経時的羅列に留まっていることも少なくない。しかし、農場全体における有病率が一定であったとしても、臓器廃棄率は、搬入動物群の偏りなどによっても変化するため、廃棄率の増減のみから、農場における衛生状態を評価することは難しい。したがって、還元データを科学的に評価するための統計学的処理方法が必要である。

時系列分析は、時間とともに出現するデータの列を分析する方法であり、指数平滑法 [1, 2] や自己回帰移動平均 (ARMA) モデル [3] などのモデルで分析が行われる。

指数平滑法は①式であらわされる。

$$\textcircled{1} \hat{y}_{t+1} = \beta \times y_t + (1 - \beta) \times \hat{y}_t$$

y_t : 時点 t における観測値

\hat{y}_t : 時点 t における予測値

β : 平滑化定数

この方法は 1 つのパラメータ (β) を推定するだけでよいため、簡易性の面からよく用いられている。一方、ARMA モデルは自己回帰 (AR) モデルと移動平均 (MA) モデルの和であり、指数平滑法を含む多様なモデリングが可能である (指数平滑法は、ARMA (1,1) モデルに相当する)。しかし、ARMA モデルは傾向変動や季節変動を含むデータをモデル化することに難がある [4]。自己回帰和分移動平均 (ARIMA) モデル [3] は傾向変動を含む時系列データにも対応でき、季節自己回帰和分移動平均 (Seasonal ARIMA : SARIMA) モデル [3] は、通常の ARIMA モデルと季節階差に関する ARIMA モデルを合わせたものであり、季節変動があるものに用いられる。SARIMA モデルは②式によってあらわされる。

† 連絡責任者：足立泰基 (北海道釧路保健所)

〒084-0925 釧路市新野 28 番地

☎ 0154-57-9366 FAX 0154-57-9388

E-mail : adachi.yasumoto@pref.hokkaido.lg.jp

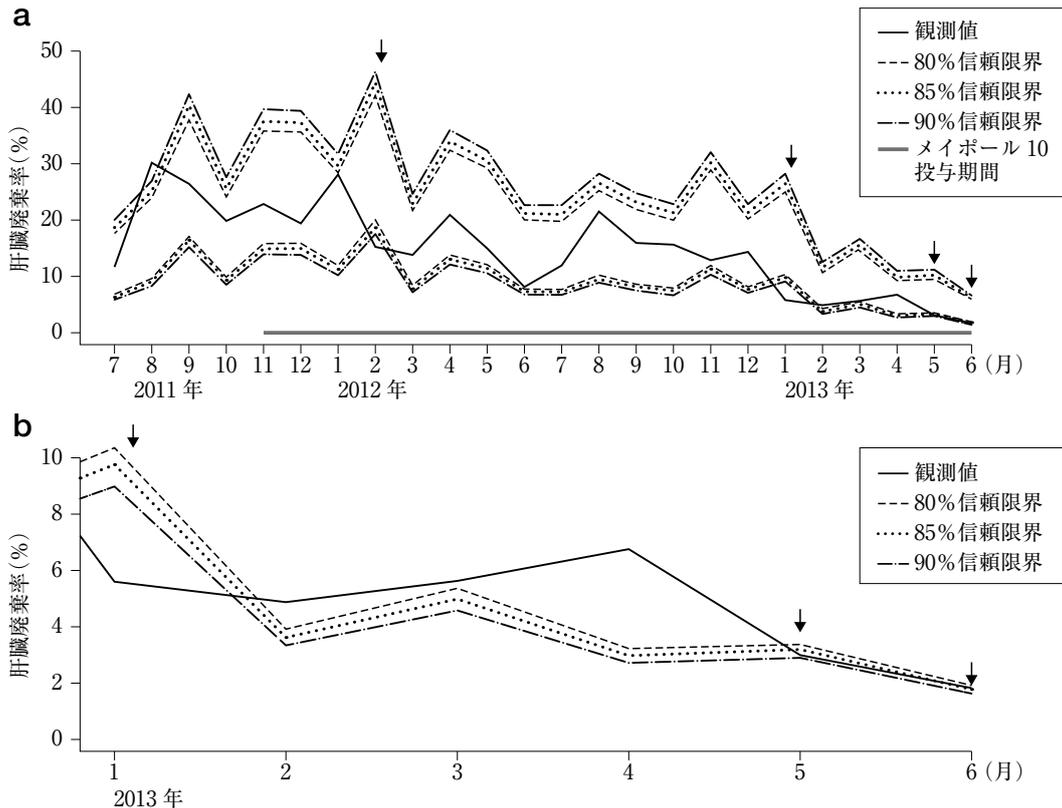


図1 I農場におけるメイポール10投与期間と寄生虫肝炎による廃棄率と信頼限界の推移
 廃棄率がいずれかの信頼限界を下回った時を↓で示す。
 a: 2011年7月～2013年6月の上方及び下方信頼限界
 b: 2013年1月～6月の下方信頼限界

$$\textcircled{2} \phi(\mathbf{B})\Phi(\mathbf{B}^s)\nabla_s^d\nabla^d\mathbf{y}_t = \theta(\mathbf{B})\Theta(\mathbf{B}^s)\mathbf{e}_t, \{\mathbf{e}_t\} \sim N(\mathbf{0}, \sigma^2)$$

②式において、

$$\phi(\mathbf{B}) = (1 - \phi_1\mathbf{B} - \phi_2\mathbf{B}^2 - \dots - \phi_p\mathbf{B}^p)$$

自己回帰の係数

$$\Phi(\mathbf{B}^s) = (1 - \Phi_1\mathbf{B}^s - \Phi_2\mathbf{B}^{2s} - \dots - \Phi_p\mathbf{B}^{sp})$$

季節自己回帰の係数

$$\theta(\mathbf{B}) = (1 + \theta_1\mathbf{B} + \theta_2\mathbf{B}^2 + \dots + \theta_q\mathbf{B}^q)$$

移動平均の係数

$$\Theta(\mathbf{B}^s) = (1 + \Theta_1\mathbf{B}^s + \Theta_2\mathbf{B}^{2s} + \dots + \Theta_q\mathbf{B}^{sq})$$

季節移動平均の係数

\mathbf{B} はラグ演算子 ($\mathbf{B}^n\mathbf{y}_t = \mathbf{y}_{t-n}$)

∇ は連続階差を示す差分演算子 ($\nabla^n\mathbf{y}_t = \mathbf{y}_t - \mathbf{y}_{t-n}$)

∇_s は季節階差を示す差分演算子 ($\nabla_s^n\mathbf{y}_t = \mathbf{y}_t - \mathbf{y}_{t-sn}$)

である。

ARIMAモデルまたはSARIMAモデルは医学・生物学や公衆衛生分野を含む自然科学の広い分野における時系列分析で使用されている [5]。東藻琴食肉衛生検査所では、10種類の臓器または疾病による廃棄率について12カ月前から3カ月前までの毎月の廃棄率期待値の信頼限界をSARIMAモデルによって計算し、生産者に還元している。本研究では、生産者に還元するためのデー

タ処理において疾患の有意な増加や薬物投与などの有効性を判定するためのSARIMAモデルの条件を検討するとともに、モデルの当てはまりのよさを指数平滑法(対照モデル)と比較検討した。

材料及び方法

駆虫剤投与農場の廃棄率データによる信頼水準の選択: 東藻琴食肉衛生検査所管轄と畜場に2003年4月～2014年3月に搬入されたI農場産の6カ月齢肉豚409,067頭の寄生虫性肝炎による廃棄率データを用いた。I農場産肉豚の寄生虫性肝炎による肝臓の廃棄率は2011年8月に30%を超えたが、2011年11月からの駆虫剤(メイポール10: 明治製菓製)の飼料添加後、2013年6月に1.8%まで減少している。この廃棄率下落局面を薬物投与の有効性として検出できるよう、以下に述べる手順で信頼水準を選択した。

2003年4月を時点 $t=1$ として毎月の廃棄率 p_t =廃棄数/と畜頭数を $t=1\sim 132$ について計算し、おのこの p_t を以下の③式

$$\textcircled{3} \text{logit}(p_t) = \ln\left(\frac{p_t}{1-p_t}\right)$$

によって対数ロジット変換した。変換後のデータを用い

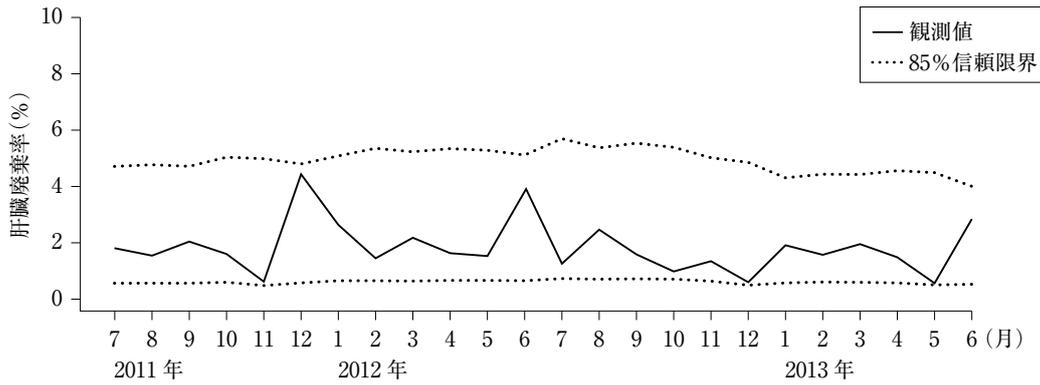


図2 図1aと同時期のD農場における寄生虫性肝炎による廃棄率と信頼限界の推移

て、②式における $P+Q+p+q$ の最大次数が5となるすべての組み合わせについて SARIMA モデルの最尤推定を行った。SARIMA モデルの階差の次数 D 及び d を決定するのに必要な単位根の有無の判定を Augmented Dickey-Fuller (ADF) 検定によって行った。われわれは、パラメータを推定したすべてのモデルの中で最小の赤池情報量規準 (AIC) を持つモデルを選択した。なお、最尤推定を行う際に、まず 2003 年 4 月から 2004 年 3 月までのデータを用いて 2004 年 4 月の期待値計算のためのモデリングを行い、2004 年 4 月のデータを追加してから 2004 年 5 月の期待値計算のためのモデリングを行う一歩先予測 (one step prediction) を 120 回連続することによって 2004 年 4 月から 2014 年 3 月のモデリングを行った。推定されたモデルから計算された標準偏差と期待値から信頼水準を 80%, 85% 及び 90% とした場合についてそれぞれ上方及び下方信頼限界を計算した。計算された信頼限界を、④式によってロジスティック逆変換し、グラフに表示した (図 1a, b)。用いた統計処理ソフトウェアは、R (Ver.3.0.2) [6] であり、<http://cran.r-project.org/> から Windows32/64bit 版をダウンロードし、コンピュータにインストールして用いた。また、SARIMA モデルの最尤推定に auto.arima 関数 [7] を用いた。

$$\textcircled{4} f(x) = \frac{e^x}{1+e^x}$$

われわれは、I農場の廃棄率が下落した 2011 年 8 月～2013 年 6 月において、観測された廃棄率が最大頻度で下方信頼限界を下回る信頼水準のうち、最大のものを選択した。

I農場における廃棄率下落局面と同時期に寄生虫性肝炎による肝臓廃棄率の顕著な増減がみられなかった D 農場についても I農場のデータを用いて行った処理と同一の処理を行い、グラフに表示した (図 2)。ただし信頼水準については I農場の検討で選択されたもの (85%) を用いた。

分析システムの構築と還元：毎日のと畜検査の結果

は、と畜場でタブレット端末に入力後、オンラインで北海道庁のサーバーに送信され、蓄積されている。われわれは、サーバーよりダウンロードされた各生産者の月次と畜検査データ (搬入動物数、疾病ごとの臓器廃棄数) をパーソナルコンピュータに取り込んで時系列分析し、結果をグラフに表示して印刷するという処理を自動的に行うシステムを構築した。このシステムにおける処理手順を以下に説明する。

毎月分の廃棄率 ($p_t = \text{廃棄数} / \text{と畜頭数}$) を計算し、おのおのの p_t を③式によって対数ロジット変換した後、信頼水準の選択時と同じ条件で SARIMA モデルの最尤推定を行った。ただし、一歩先予測により直近 12 カ月分の期待値計算用のモデリングを行い、さらに最終月のデータを加えて推定したモデルを今後 3 カ月間の期待値計算用モデルとした。推定されたモデルから計算された標準偏差と期待値から上方及び下方信頼限界を計算し、④式によって % 値に変換した後、折れ線グラフに表示した。なお、時系列分析を行う廃棄率データの項目を、臓器 4 項目 (肺、大腸、小腸及び肝臓) と Swine enzootic pneumonia (SEP) 様肺炎、肺胸膜炎、心膜炎、肝包膜炎、寄生虫性肝炎及び肝変性による臓器廃棄率とした。

各調査農場における廃棄率モデルの妥当性及び対照モデルとの比較：2003 年 4 月から 2014 年 3 月までの 11 年間に東藻琴食肉衛生検査所管轄のと畜場に搬入された 10 農場 (A～J 農場) 産の 6 カ月齢肉豚 1,193,837 頭の 4 疾病 (SEP 様肺炎、肺胸膜炎、心膜炎及び寄生虫性肝炎) の月次集計データを用いた。

毎月の廃棄率 ($p_t = \text{廃棄数} / \text{と畜頭数}$) を計算し、おのおのの p_t を③式によって対数ロジット変換した後、11 年分のデータすべてを用いて信頼水準 (85%) の選択時と同じ条件で SARIMA モデルの最尤推定と①式に示した指数平滑法の最尤推定を行った。推定した SARIMA モデルのパラメータを、SARIMA (p, d, q) (P, D, Q) S+C の形式で記載することとした (表 1)。ここで、p: 自己回帰項の次数 d: 階差の次数 q: 移動平

SARIMA モデルによると畜検査データの時系列分析法

表 推定された各モデルにおけるパラメータ数と有意差が認められたパラメータ数, 残差の無作為性が認められた最大時間差, 尤度比検定 (逸脱度) 及び絶対平均比率誤差 (MAPE)

農場	SEP 様肺炎				肺胸膜炎				心膜炎				寄生虫性肝炎			
	モデル	有意パラメータ数	全パラメータ数		モデル	有意パラメータ数	全パラメータ数		モデル	有意パラメータ数	全パラメータ数		モデル	有意パラメータ数	全パラメータ数	
1. パラメータ数	A	SARIMA(2,1,1) (2,0,0)12	1	5	SARIMA(1,0,2) (2,0,0)12+C	3	6		SARIMA(0,1,3) (1,0,1)12	2	5		SARIMA(1,0,1) (1,0,2)12+C	3	6	
	B	SARIMA(1,1,1)	1	2	SARIMA(2,1,1) (2,0,0)12+C	2	6		SARIMA(0,1,1)	1	1		SARIMA(0,0,1) (2,0,0)12	1	3	
	C	SARIMA(4,1,0)	2	4	SARIMA(0,1,1) (2,0,1)12	3	4		SARIMA(1,0,1) (1,0,1)12+C	3	5		SARIMA(1,0,1) +C	3	3	
	D	SARIMA(3,1,0) (2,0,0)12	3	5	SARIMA(0,1,1) (2,0,2)12	1	5		SARIMA(1,1,1) (2,0,1)12	3	5		SARIMA(0,1,2) (1,0,0)12	1	3	
	E	SARIMA(0,1,1)	1	1	SARIMA(1,1,1)	2	2		SARIMA(3,0,0) (2,0,0)12+C	3	6		SARIMA(1,0,0) +C	2	2	
	F	SARIMA(1,0,1) (2,0,0)12+C	4	5	SARIMA(0,1,1) (1,0,1)12	1	3		SARIMA(4,1,1)	5	5		SARIMA(2,0,1) +C	3	4	
	G	SARIMA(1,0,1) (2,0,0)12+C	4	5	SARIMA(0,0,1) (0,1,1)12	2	2		SARIMA(1,1,1)	2	2		SARIMA(4,0,0) (1,1,0)12	3	5	
	H	SARIMA(1,1,1)	2	2	SARIMA(0,0,2) (0,0,1)12	3	3		SARIMA(0,0,2) (0,0,2)12	3	4		SARIMA(0,1,4)	3	4	
	I	SARIMA(1,0,0) (2,0,2)12+C	4	6	SARIMA(2,0,0) (2,0,1)12+C	4	6		SARIMA(2,1,3)	6	6		SARIMA(3,1,0) (1,1,1)12	3	5	
	J	SARIMA(3,1,0) (2,0,0)12	4	5	SARIMA(2,1,0)	2	2		SARIMA(4,1,0)	2	4		SARIMA(1,0,1) (2,0,0)12+C	3	5	
合計		26	40		23	39			30	43			25	40		
2. 残差の無作為性 (残差の無作為性が認められた最大時間差 (月))	A	10 ≤			10 ≤				10 ≤				10 ≤			
	B	10 ≤			10 ≤				10 ≤				10 ≤			
	C	10 ≤			10 ≤				10 ≤				10 ≤			
	D	10 ≤			10 ≤				10 ≤				10 ≤			
	E	10 ≤			10 ≤				10 ≤				10 ≤			
	F	10 ≤			10 ≤				10 ≤				10 ≤			
	G	4			10 ≤				10 ≤				10 ≤			
	H	10 ≤			10 ≤				10 ≤				10 ≤			
	I	10 ≤			10 ≤				10 ≤				10 ≤			
	J	10 ≤			10 ≤				10 ≤				10 ≤			
3. 尤度比検定 ^{※1}	A	261.6*	272.4	186.7*	196.3	199.1	204.9	267.8	275.9							
	B	224.7	225.7	74.8	83.9	154.8	155.9	232.8	235.5							
	C	112.8*	120.7	43.3***	81.4	121.9*	132.5	271.5	274.3							
	D	226.1	225.1	36.4*	44.8	160.1***	203.6	272.5	275.7							
	E	149.3	152.6	93.8*	98.2	131	138.3	288.7***	310							
	F	140.6***	163.7	26.2	29.8	81.2***	100.1	269.2***	285							
	G	120.0***	145.9	-46.7***	-35.4	-4.0**	5.2	171.4	178.8							
	H	57.3*	63.7	-51.2***	-39.3	-72.9**	-60.6	184.3***	202.1							
	I	30.5***	54.8	-173.3***	-131.7	-64.0**	-50.8	155.8***	200.5							
	J	253.1**	263.6	160.0**	170.7	194.0*	200.2	340.3**	349.8							
4. 絶対平均比率誤差 (MAPE) ^{※2}	A	24.52	25.64	20.37	20.75	22.26	22.31	13.67	14.05							
	B	17.39	17.44	12.10*	12.75	10.04	10.11	12.37	12.49							
	C	11.78	12	9.89***	12.47	9.45	9.59	12.07	12.27							
	D	14.36	14.49	13.01*	13.59	9.41***	12.73	12.3	12.56							
	E	17.42	17.52	15.94	16.07	10.96	10.75	13.05*	14.83							
	F	12.99*	14.16	11.94	12.17	8.13	8.61	11.77**	12.72							
	G	13.35	14.5	11.57	11.86	7.04	6.96	8.46	8.69							
	H	12.78**	13.67	10.71*	11.43	5.77	5.83	10.92	11.49							
	I	9.58	10.44	5.93***	7.76	5.71	5.95	14.02**	19.02							
	J	21.85	22.18	19.39	19.96	13.2	13.55	19.07	19.35							

※1: 数値は④式~⑦式によって求められた逸脱度である。逸脱度の差を用いて行った尤度比検定の結果をアスタリスクで示し、両モデルのうち逸脱度の小さい数値を太字で表示した。

※2: MAPE の値を指数平滑法と SARIMA モデルで比較し、小さい数値を太字で表示した。また t 検定の結果をアスタリスクによって示した。

*: $P < 0.05$ **: $P < 0.01$ ***: $P < 0.001$

均項の次数 P：季節自己回帰項の次数 D：季節階差の次数 Q：季節移動平均項の次数 s：季節変動の期間 C：定数項を意味している。

推定された p, q, P, Q, C の各項をそれぞれの標準偏差で除した値を t 値とし、パラメータの有意性検定として有意水準 0.05 の t 検定を行った。

さらに、残差の無作為性検定 (Ljung-Box 検定 [8]) として⑤式で計算される Q 値が自由度 h の χ^2 分布の 5% 分位点を上回る場合に、帰無仮説を棄却した。具体的には R の tdiag 関数のデフォルト条件 ($h \leq 10$) で表示される p 値のチャートが 0.05 を下回っている場合に残差の無作為性なしと判定し、その時の h-1 を残差の無作為性が認められた最大の時間差 (月) とした。

$$\textcircled{5} Q = n(n+2) \sum_{j=1}^h \frac{\hat{p}_j^2}{n-j}$$

n：標本数

h：時間差

\hat{p}_j ：時間差 j における自己相関係数

以下に示す⑥式によって、各モデルの尤度 L を計算し、⑦式によって逸脱度 D を計算した。

$$\textcircled{6} L = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\varepsilon_t^2}{2\sigma^2}}$$

⑥式において、 $\varepsilon_t = \text{logit } p_t - \text{logit } \hat{p}_t$

p_t ：t 時点における実際の廃棄率

\hat{p}_t ：t 時点におけるモデルによる廃棄率期待値

σ ：モデルと観測値の標準偏差

n：時点数

である。

$$\textcircled{7} D = -2 \log L$$

尤度比検定として、SARIMA モデルと指数平滑法によるモデルの逸脱度 D について、モデルと対照モデルのパラメータ数の差を自由度とし、 $P < 0.05$ を有意水準とする χ^2 検定を行った。

⑧式により、絶対平均比率誤差 (MAPE) を計算し、農場及び疾病ごとに指数平滑法と SARIMA モデルで数値の大小比較を行い、さらに t 検定 (両側検定) を行った。

$$\textcircled{8} MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\text{logit } p_t - \text{logit } \hat{p}_t}{\text{logit } p_t} \right|$$

成 績

駆虫剤投与農場の廃棄率データによる信頼水準の選択：I 農場における寄生虫性肝炎による肝臓廃棄率の経時変化 (実線) と期待値の上方及び下方信頼限界 (80%, 85%, 90%) 並びに医薬品使用報告書に基づくメイポール 10 の投与豚搬入期間を図 1a, b に示した。肝臓廃棄率は、2011 年 8 月に上方信頼限界 (80~90%) を超え

て増加した。2011 年 11 月から 2013 年 6 月までの間に廃棄率は、30.2~1.8% に下落し 80% 及び 85% 信頼限界を 2012 年 2 月、2013 年 1, 5, 6 月の 4 回下回ったが、90% 信頼限界の場合は、2012 年 2 月及び 2013 年 1 月の 2 回下回ったのみであった。以上の結果より、最大頻度で下方信頼限界を下回る信頼水準のうち、最大の信頼水準として 85% を選択した。選択した信頼水準 85% を用いて D 農場における寄生虫性肝炎による肝臓廃棄率データを I 農場と同様に処理した結果を図 2 に示す。廃棄率 (実線) は上方及び下方信頼限界 (点線) の間にあり、両信頼限界はおおむね横ばいであった。

分析システムの構築と還元：システムからの出力例として 2014 年 8 月末日における D 農場の解析結果を示した (図 2)。図 2 に例示した処理結果には、10 の臓器または疾病による廃棄率について 12 カ月前から 3 カ月先までの毎月の廃棄率期待値の信頼限界を表示しており、東藻琴食肉衛生検査所では 2014 年 4 月より希望する生産者にはこの様式の処理結果を郵送している。

各調査農場における廃棄率モデルの妥当性及び対照モデルとの比較：推定された各モデルにおけるパラメータ数と有意差が認められたパラメータ数、残差の無作為性が認められた最大時間差、尤度比検定 (逸脱度) 及び絶対平均比率誤差 (MAPE) を表 1 に示した。全パラメータ数 162 個のうち、有意なパラメータは、104 個 (64.2%) であり、疾病別ではそれぞれ、SEP 様肺炎では 26/40 (65.0%)、肺胸膜炎では 23/39 (59.0%)、心膜炎では 30/43 (69.8%)、寄生虫性肝炎では 25/40 (62.5%) であった。

残差の無作為性検定 (Ljung-Box 検定) における時間差 h には明確な基準はないが、10 以上を選択することが多く、本研究で使用した tdiag 関数の通常条件では時間差 10 までの p 値が表示される。本研究では、残差の無作為性が認められる最大時間差数が 10 未満であったのは、G 農場 SEP 様肺炎のモデルのみであった (表 2. 残差の無作為性の欄)。

表 3. 尤度比検定の欄に示した逸脱度の差は、同じ農場の同じ疾病について指数平滑法の逸脱度から SARIMA モデルの逸脱度を差し引いたものである。SARIMA モデルの逸脱度が大きいのは D 農場における SEP 様肺炎の場合のみで、他の 39 農場・疾患の比較では、SARIMA モデルの逸脱度が小さく、そのうち 27 の組み合わせにおいて有意差が認められた。

38 モデルで SARIMA モデルの MAPE が小さく、40 モデル中 11 モデルで有意差 ($P < 0.05$) が認められた。尤度比及び MAPE のいずれの検定においても、指数平滑法モデルの方が有意に小さいと判断されたものはなかった。

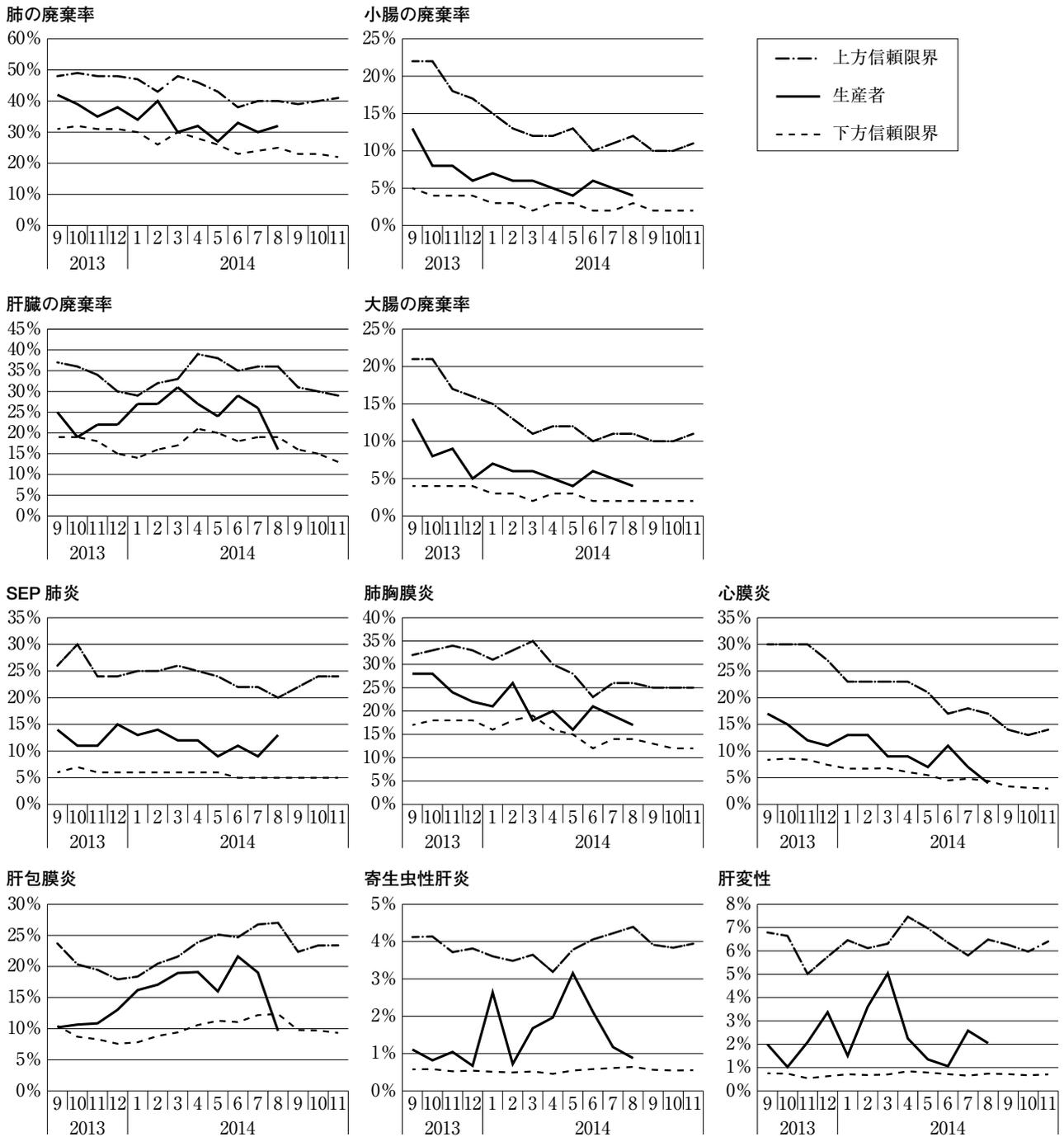


図3 分析システムの出力例 (D農場)

考 察

本研究では、と畜検査データを SARIMA モデルによって分析するための条件検討を行い、推定したモデルの統計学的評価を行った。公衆衛生分野の研究では、感染症患者数の解析に ARIMA モデルまたは SARIMA モデルが用いられている。そのような研究の例として、カンピロバクター食中毒 [9]、重症急性呼吸器症候群 [10]、腎症候性出血熱 [11]、コレラ [12]、オンコセルカ症 [13]、インフルエンザ [14]、マラリア [15]

やデング熱 [16] の研究があげられる。また救急外来患者数の解析に用いられることもある [17, 18]。これらの中には、バイオテロリズムの発生 [17] や、未知の食中毒感染源 [9] の検知を目的とするものもある。時系列分析が、時間とともに出現するデータの分析方法であることからと畜検査データの解析に利用できることは明らかであるが、と畜検査データの時系列分析例は少ない。報告されているのは、主として経時変化をトレンドや周期性などの要素に分解することを目的としたもので [19, 20]、ARIMA モデルを用いてと畜検査データの時

系列モデリングを行った研究として Neumann ら [21] の報告のみが認められる。ただし、ARIMA モデルを利用し、期待値と廃棄率を比較することによって合理的評価が可能であることに言及しているが、評価そのものを行っていない。

と畜検査における廃棄率の経時的羅列（プロット）から把握される数値変動の中から衛生上の問題として捉えるべき変動とそうでない変動を識別する拠り所は分析者の直感によるところが多い。しかし、直感による判断結果は分析者により異なるものであり、同じ結果が得られないことが多い。本分析法では、観測された廃棄率と SARIMA モデルを利用して求められた期待値の信頼限界との大小比較に基づいて上記の識別を行うため、数値の大小比較は分析者に依存せず、直感に依拠するより科学的な方法である。

本分析法では、信頼水準を適宜選択することにより、感度や特異度を調整することができる。信頼水準として 95% が選択される統計解析の例は多いが、気象庁の台風予報円 [22] のように 70% が選択される場合もあり、95% に限定されない。信頼水準を大きくしすぎると両信頼限界の間の領域が非常に大きくなることにより、感度が低下する。逆に信頼水準を小さくしすぎると、特異度が低くなり、些細な廃棄率の変動をも有意な変動として検出するようになる。本研究では駆虫剤投与による肝臓廃棄率の下落過程における検出頻度をもとに 85% を選択したが、この値に限定するものでなく、その他の事例も参照し、感度と特異度はトレードオフの関係にあることを勘案して両者のバランスを調整することが望ましい。具体的には、農場における真の有病率が変動している蓋然性が高い時に観測された廃棄率変動を検出しうる最大の信頼水準を選択することになる。農場における真の有病率を調査することによって把握した有病率の変動を用いてバランスを調整することが一層望ましい。

I 農場の廃棄率データによって選択した信頼水準 85% は、駆虫剤投与時の廃棄率が最大頻度で下方信頼限界を下回る信頼水準のうち、最大のものであり、感度と特異度のバランスが簡便に調整されたものである。2011 年 8 月に見られるように上方信頼限界を超えて廃棄率が増加している場合を廃棄率の著明な増加と判定し、2012 年 2 月、2013 年 1, 5, 6 月にみられるように廃棄率が下方信頼限界を下回る場合を廃棄率の著明な減少と判定するのに用いることができる。以上のように信頼限界を超えるまたは下回ることを判断指標にすれば、著明な廃棄率の増加や減少があったことを生産者は識別することができ、農場における疾病のまん延や衛生対策の効果の指標とすることができる。

東藻琴食肉衛生検査所では、本研究に係る分析方法に

よる処理結果を生産者宛に送付する以前は、廃棄率を経時的に羅列したグラフを生産者宛に郵送していたが、廃棄率の増減のみから対策を講じるべき廃棄率の変動を識別するのは困難と考える。しかし、信頼限界を超える廃棄率の増減を生産者に示すことにより農場の衛生対策を講じるきっかけを与えたり、投薬などの対策の効果を生産者が確認できることが期待できる。また、今後 3 カ月間の廃棄率を示すことによって、生産者に早めの対応を促すことを期待している。

通常、SARIMA モデルの推定を行う際には、モデルを構成する多項式の各係数がゼロと有意差があるかを確認し、さらにモデルの妥当性の指標の一つである残差の無作為性を確認する必要がある。東藻琴食肉衛生検査所が毎月送付しているおよそ 30 生産者の還元データを処理する過程で生成される約 3,900 個のモデルには、20,000 個程度のパラメータが含まれており、すべてを毎月検定することは困難である。よってこれらの検定を行わず、AIC が最も低いモデルを自動処理で選択するように分析システムを構築した。そこで、本研究では全生産者から一部抽出したデータを用いて推定した 40 のモデルを構成する多項式の各係数の有意性と残差の無作為性を検討した。モデルには最低一つ以上の有意パラメータが存在すれば SARIMA モデルとして機能する。AIC を規準としてパラメータの個数を選択していることから有意差のないパラメータがモデルの当てはまりを悪化させることはないが、有意パラメータの数は多いほうがよい。今回検討した 40 のモデルはすべて有意パラメータを含み、全体の 6 割以上が有意であり、疾病ごとの偏りは認められなかった。また、残差の無作為性が認められなかったのは 40 モデル中 1 モデル (2.5%) のみであった。残差とは、観測データとモデルの差であり、残差が無作為とは残差に時間依存性がないことを意味し、残差が無作為ではないとは、観測データの時間依存性を十分にモデル化できていないことを意味する。本研究の条件では観測データの時間依存性を十分にモデル化できていない場合が 2.5% に留まっていることがわかった。

本研究では、尤度比検定を実施し、さらに MAPE を計算することによって、SARIMA モデルと指数平滑法によるモデルの当てはまりの良さと精度の比較を行った。尤度比検定は、誤差項を含めたモデルの当てはまりの良さを比較するのに用いられるのに対し、MAPE は予測の精度を示している。比較の結果、尤度比検定 (39/40 で良好、27 モデルで有意) と MAPE (38/40 で良好、11 モデルで有意) の両方で SARIMA モデルが良好であり、当てはまりの良さ及び精度の両面で SARIMA モデルが指数平滑法によるモデルより優れていることがわかった。

本研究における分析方法では、自動処理を行うために

AICのみを指標とし、残差の無作為性検定を行わず、有意差のないパラメータを含むモデルをそのまま選択しているが、残差の無作為性が認められなかったのは2.5%に過ぎず、選択されたモデルは指数平滑法などの簡便に推定できるモデルよりも精度が高く当てはまりが良好であることから、われわれは、本方法で選択されたSARIMAモデルをと畜検査データの分析に利用することは妥当であると考え。

これまで、と畜検査データの解析方法に関する研究が少なかったため、廃棄率や廃棄数そのものが生産者に還元されていることが多いものと思われる。本研究はと畜検査データの統計処理方法の例としてSARIMAモデルを示し、おもに直感で判断されていた廃棄率の傾向変化を科学的に判断する指標として利用できることを示した。本指標をより所にして生産現場の衛生向上を行う動機としたり、投薬や消毒などの成果を確認できることにより、高品質で衛生的な食肉生産を支援するツールとして本法による分析結果が生産者に利用されることを期待している。特に投薬や消毒などの行動はコストアップ要因であるために生産者によって躊躇する可能性があるが、効果をより合理的に判断できることにより積極的な対応を支援できる可能性がある。現在、早めの対策を促すために3カ月先までの予測を生産者に還元している。さらなる長期の予測を出力することも可能であるが、生産者側に長期予測値のニーズがあるかどうかいまのところ不明であるため新たなニーズを調査した上で利用法を検討したい。また、他分野における時系列分析法の研究はさらに高度化しており、と畜検査データの分析法についてもSARIMAモデルに留まらず、種々の手法への取り組みを継続していく必要がある。

稿を終えるにあたり、と畜検査データの使用を許可してくださった北海道東藻琴食肉衛生検査所、横山敦志所長に深謝する。

引用文献

- [1] Holt C : Forecasting seasonals and trends by exponentially weighted averages, *Int J Forecast*, 20, 5-13 (2004)
- [2] Winters PR : Forecasting sales by exponentially weighted moving averages, *Management Sci*, 6, 324-342 (1960)
- [3] Box GEP, Jenkins GM : Time series analysis: forecasting and control, rev. ed, 47-412, San Francisco, Holden-Day (1976)
- [4] Brockwell P, Davis R : Introduction of Time Series and Forecasting, 2nd ed, 23-35, Springer, New York (2002)
- [5] Zeger SL, Irizarry R, Peng RD : On time series analysis of public health and biomedical data, *Annu Rev Publ Health*, 27, 57-79 (2006)
- [6] Ihaka R, Gentleman RR : A language for data analysis and graphics, *J Comp. Graph Stat*, 5, 299-314 (1996)
- [7] Hyndman RJ, Khandakar Y : Automatic time series forecasting: The forecast package for R, *J Stat Softw*, 27, 3 (2008), (online), (<http://www.jstatsoft.org/v27/i03>), (accessed 2013-08-03)
- [8] Ljung GM, Box GEP : On a measure of lack of fit in time series models, *Biometrika*, 65, 297-303 (1978)
- [9] Allard R : Use of time-series analysis in infectious disease surveillance, *Bull World Health Organ*, 76, 327-333 (1988)
- [10] Earnest A, Chen MI, Ng D, Sin LY : Using autoregressive integrated moving average (ARIMA) models to predict and monitor the number of beds occupied during a SARS outbreak in a tertiary hospital in Singapore, *BMC Health Serv Res*, 5, 36 (2005), (online), (<http://www.biomedcentral.com/1472-6963/5/36>), (accessed 2013-06-01)
- [11] Liu Q, Liu X, Jiang B, Yang W : Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model, *BMC Infect Dis*, 11, 218 (2011), (online), (<http://www.biomedcentral.com/1471-2334/11/218>), (accessed 2013-06-05)
- [12] Ali M, Kim DR, Yunus M, Emch M : Time series analysis of cholera in Matlab, Bangladesh, during 1988-2001, *J Health Popul Nutr*, 31, 11-19 (2013)
- [13] Lara-Ramírez EE, Rodríguez-Pérez MA, Pérez-Rodríguez MA, Adeleke MA, Orozco-Algarra ME, Arrendondo-Jiménez JI, Guo X : Time series analysis of onchocerciasis data from Mexico: a trend towards elimination, *Plos Neglect Trop D*, 7, e2033 (2013), (online), (<http://www.plosntds.org/article/info%3Adoi%2F10.1371%2Fjournal.pntd.0002033>), (accessed 2014-02-01)
- [14] Soebiyanto RP, Adimi F, Kiang RK : Modeling and predicting seasonal influenza transmission in warm regions using climatological parameters, *PLOS ONE*, 5, e9450 (2010), (online), (<http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0009450>), (accessed 2013-07-01)
- [15] Wangdi K, Singhasivanon P, Silawan T, Lawpoolsri S, White NJ : Kaewkungwal J : Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: a case study in endemic districts of Bhutan, *Malaria J*, 9, 251 (2010), (online), (<http://www.malariajournal.com/content/9/1/251>), (accessed 2013-07-11)
- [16] Bhatnagar S, Lal V, Gupta SD, Gupta OP : Forecasting incidence of dengue in Rajasthan, using time series analyses. *Indian J Publ Health*, 56, 281-285 (2012)
- [17] Reis BY, Mandl KD : Time series modeling for syndromic surveillance, *BMC Med Inform Decis*, 3, 2 (2003), (online), (<http://www.biomedcentral.com/1472-6947/3/2>), (accessed 2013-07-01)
- [18] Sun Y, Heng BH, Seow YT, Seow E : Forecasting daily attendances at an emergency department to aid resource planning, *BMC Emergency Medicine*, 9, 1 (2009), (online), (<http://www.biomedcentral.com/1472-6947/3/2>), (accessed 2013-07-01)

- com/1471-227X/9/1), (accessed 2013-08-01)
- [19] Sanchez-Vazquez MJ, Nielen M, Gunn GJ, Lewis FI : National monitoring of *Ascaris suum* related liver pathologies in English abattoirs: a time-series analysis, 2005-2010, *Vet Parasitol* 184, 83-87 (2012)
- [20] Sanchez-Vazquez MJ, Nielen M, Gunn GJ, Lewis FI : Using seasonal-trend decomposition based on loess (STL) to explore temporal patterns of pneumonic lesions in finishing pigs slaughtered in England, 2005-2011, *Prev Vet Med*, 104, 65-73 (2012)
- [21] Neumann E, Hall W, Stevenson M, Morris R, Ling Min Than J : Descriptive and temporal analysis of post-mortem lesions recorded in slaughtered pigs in New Zealand from 2000 to 2010, *N Z Vet J*, 62, 110-116 (2014)
- [22] 市沢成介 : 台風予報の現状について, *天気*, 22, 281-286 (1975)

Method of Time Series Analysis of Meat Inspection Data Using Seasonal Autoregressive Integrated Moving Average Model

Yasumoto ADACHI^{1)†} and Kohei MAKITA²⁾

- 1) *Kushiro Health Center, Kushiro Subprefectural Bureau, Hokkaido Government, 28 Niino, Kushiro-shi, 084-0925, Japan*
- 2) *Veterinary Epidemiology, Division of Health and Environmental Science, Department of Veterinary Medicine, School of Veterinary Medicine, Rakuno Gakuen University, 582 Midorimachi, Bunkyo-dai, Ebetsu-shi, 069-8501, Japan*

SUMMARY

Although many meat inspection centers provide abattoir condemnation data to each producer, few studies relating to statistical processes on abattoir data have been reported. Time series analysis is widely used in a range of biomedicine, including public health, and can be used to detect outbreaks of disease and to determine the effect of drugs based on significant changes in the condemnation rate. In this study, we carried out investigations to obtain favorable conditions for utilizing the Seasonal Autoregressive Integrated Moving Average (SARIMA) model for the analysis of the condemnation rates. The degree of fit between the condemnation rates and the models estimated using the SARIMA model and those estimated using the exponential smoothing method as a control were compared for 4 diseases of swine from 10 producers. The deviances for 39 out of 40 models were smaller than the control, and statistical significance was shown for 27 out of 40 models. In conclusion, this method can provide an index for determining the sanitation of farms for producers, and can be utilized as a tool to promote hygienic meat production. — Key words : abattoir data, SARIMA model, time series analysis.

† Correspondence to : Yasumoto ADACHI (*Kushiro Health Center, Kushiro Subprefectural Bureau, Hokkaido Government*)
28 Niino, Kushiro-shi, 084-0925, Japan
TEL 0154-57-9366 FAX 0154-57-9388 E-mail : adachi.yasumoto@pref.hokkaido.lg.jp

J. Jpn. Vet. Med. Assoc., 68, 189 ~ 197 (2015)