

# マルコフモデル利用による予測結果の検証と2015年農林業 センサデータの有効性

誌名	関東東海北陸農業経営研究
ISSN	21897646
著者	安武, 正史
巻/号	108号
掲載ページ	p. 53-57
発行年月	2018年2月

## マルコフモデル利用による予測結果の検証と 2015年農林業センサデータの有効性

安武 正史  
(中央農業研究センター)

### I 研究の目的

本報告は農林業センサ個票データをパネルデータ化し、マルコフモデルを用いて農業担い手予測の可能性を検討したものである。具体的には2010年農林業センサの数値から2015年の予測を行い、2015年の結果と比較することにより、マルコフモデルの予測の有効性について検討を行った。その上で2015年センサのデータ利用による予測値算出の信頼性を検討した。

2010年の農林業センサ結果での特徴として集落営農急増があり2010年のデータを用いると偏りのある予測値のなる可能性が高い。本論文ではこの状況についても考慮した。

### II データとモデル

農林業センサは日本の全農業経営体を調査対象としている。2015年では約130万の農業経営体のデータが蓄積されている。

特に日本の農林業センサには構造動態統計という他国には見られない特徴ある集計が行われている<sup>注1)</sup>。この統計は、2回分の農林業センサ(2015年農林業センサの場合は2010年世界農林業センサ及び2015年農林業センサ)で農業経営体、農家及び林業経営体に該当した経営体の調査票情報等を接続し、この5年間の動態の変化を集計した統計である。単に大規模経営が増加した、減少した、ではなく規模拡大した経営、縮小した経営、離農した経営が把握可能なデータとなっている。

また、2009年の統計法改正により統計データ

個票の利用が促進されるようになった。改正前に比べて、手続きの簡素化、利用条件の緩和が進んでいる。

#### 1. データに関する留意事項

2010年の構造動態統計には2つの問題点がある。接続不可の問題と集落営農の問題である。構造動態統計を作成するための基データは調査2時点間のデータをパネルデータ化したもので、原則として2時点間の関係は経営体が継続、新設、経営体以外(離農)の3通りである。「接続不可」とはこれらの関係が確認できない場合である。例えば2015年時点で農業経営体として存在しているが、2010年時点の所在が確認できない場合は「接続不可」と分類される。

当然2010年時点の経営体で、離農が確認できる場合と確認できない場合もあり、確認できない場合はこれも「接続不可」になるが2010年農林業センサでは、特に区別せずに集計された(2015年では区別された)。

第1表 接続不可の図示

		単位: 経営体		
		2010年		
		農業経営体	農業経営体以外	接続不可
2005年	農業経営体	1,560,382	448,998	—
	農業経営体以外	12,480		
	接続不可	106,222		
2010年総経営体数		1,679,084		

注)2010年農林業センサ第6巻構造動態統計報告書より作成

以上の点について第1表で整理する。これは2005年から2010年にかけての構造動態統計から集計した結果である。左上の1,560,382経営体が2005年から継続している農業経営体である。その下12,480経営体が2010年に新設<sup>注2)</sup>

された経営体である。その下の 106,222 経営体が接続不可の経営体で、第 5 表で示される接続不可経営体率はこの数値を用いて計算した。この例で計算すると 6.3% (=106,222÷1,679,084×100) という値になる。右列の 448,998 経営体は 2010 年には経営体以外となった数値である。動向予測の計算では離農と見なした。

原則として全農業経営体のパネルデータ化が行われるが、「接続不可」が 2010 年の平均で 6.3%であった。特に、北海道 (23%)、群馬県 (35%)、福岡県 (53%) で接続不可の比率が高くなった。このため、この 3 道県では偏りが大きい可能性がある (後述の第 5 表参照)。

もう 1 点の集落営農について、一部の地域で 2005 年から 2010 年の間に多くの集落営農が新設された。形としては突然 20ha, 50ha といった大規模経営が数多く新設されたことになっている。この傾向は 2005 年から 2010 年にかけての特異な現象とみるべきである。この傾向は、東北地域、北陸地域、北九州地域で多く見られた。ただし農林業センサスでは経営体が「集落営農」であることは明確にできない。

第 2 表 佐賀県における経営耕地規模別経営体数の推移

	単位:経営体		
	2005年	2010年	2015年
1ha未満	14,663	11,129	9,234
1~3ha	14,307	6,515	5,556
3~5ha	2,342	997	976
5~10ha	688	528	576
10~100ha	103	596	656
100ha以上	0	24	22
合計	32,103	19,765	16,998

注) 2005年、2010年、2015年農林業センサス調査表情報に基づき独自集計

これを集落営農が急増した佐賀県の数値で示すと第 2 表のようになる。100ha 以上の経営体が 2005 年の 0 から 2010 年に 24 となっており、10~100ha 層は 103 から 596 へ 6 倍近くに急増している。一方で 1~3ha 層は 14,307 から 6,515 へ半減している。3~5ha 層も同様である。これが 2010 年から 2015 年にかけては各階層とも

微減傾向が示されている。

すなわち 2005 年から 2010 年にかけて、中小規模層の経営体が集落営農として大規模組織経営に集約され中小規模層の経営体は離農したのではなく集落営農の構成員となり農業経営体の定義から外れたため統計数値は表 2 のような数値として集約された。

## 2. マルコフモデルについて

マルコフモデルは以下の式で表される。

$$\begin{aligned}\pi(n) &= \pi(n-1)P \\ &= \pi(0)P^n\end{aligned}$$

$\pi(n)$  は  $n$  期における状態確率分布、 $P$  は推移確率行列を表す。

マルコフモデルは予測モデルとして利用するに際して問題点が 2 つあげられる。1 つは因果関係を反映できないという点である。動きを確率で示すモデルのため明確な因果関係の表現が困難となる。もう 1 点は 2 時点間の結果を引き延ばすため、2 時点間の結果以前の動向は反映できないという点である。これは回帰分析のように推計値の有意水準を示すことができないということでもある。

前者の点については一概に欠点とは言えない。規模拡大・縮小、離農は様々な要因が複雑に影響しての結果であり、地域全体の予測については回帰分析のように細かい要因を抽出しての予測値より実際に近い予測値が得られる可能性もある。また、後者については、一般的な統計手法が利用できないため過去のデータを基にした予測結果と実際の数値を単純に比較し、乖離率で評価することとした。

なお 1980 年に神谷は農林業センサスデータをマルコフモデルによって 50 年以上先までの予測値を示している<sup>(注3)</sup>。

## III. 予測結果の検証

### 1. 検証方法と結果

2010 年のデータを用いて予測した 2015 年の

予測値と 2015 年に公表された実際値を都府県で比較した。ここで北海道を除外したのは、経営耕地規模別の経営体数の分布が都府県と全く異なり、経営耕地面積を指標とした予測に際して経営耕地規模階級が都府県と異なるため一律の比較が困難と考えたためである。

経営体数全体での乖離率は概ね 5%以内（約半分の 24 都府県では 2%以内）であった（第 3 表）。乖離が大きかったのは、2011 年の東日本大震災で原発事故のあった福島県（16%）と接続不可の多かった群馬県（31%）、福岡県（56%）、集落営農の成立が多かった佐賀県（30%）である。

震災前の 2010 年のデータで予測すれば、実際値より多くの経営体数が存在するという予測値となる。群馬県と福岡県はデータの不備からの結果である。佐賀県は、データの上からは大規模層の新設と小規模層の離農比率が高い数値となっているため、予測値の大規模層は倍増、小規模層は半減となった。特に福岡県は佐賀県同様、集落営農の成立も多かったため二重に誤差が大きくなったとみられる。このほか、東京都、神奈川県、山梨県、大阪府では 10ha 以上層の乖離率が大きくなっているが、大規模経営体数がわずかであり率としてみると大きな乖離率となった（東京都で公表値 7 に対して予測値 3）。

すなわち乖離率が大きくなった要因は明確であり、これらの要因がない都道府県については近似した予測値が得られると評価できる。

## 2. マルコフモデルの有効性

ここで、データの不備からくる群馬県についてマルコフモデルの推移確立行列（Ⅱ-2 で示した式の P の部分）を農業構造が類似しているとみられる隣の茨城県の推移確立行列で計算すると乖離の少ない予測値が得られた（第 4 表）。群馬県のデータで予測した値はかなり小さな値となった。群馬県のデータに偏りがあり不適切な推移確立行列（実際より離農比率が高い数値）が算出されたとみられる。

第 3 表 予測値と公表値の乖離

	単位:%		
	1~5ha	10ha以上	経営体数合計
青森県	4.1	-2.6	3.0
岩手県	-2.4	-5.5	-2.8
宮城県	0.4	-2.6	1.0
秋田県	-5.2	9.5	-1.9
山形県	-10.2	-1.7	-4.0
福島県	22.8	7.5	16.2
茨城県	8.5	-1.7	2.3
栃木県	4.2	4.2	1.4
群馬県	-30.0	-16.5	-31.2
埼玉県	3.9	4.3	-0.1
千葉県	9.3	-5.2	4.6
東京都	9.3	-57.1	6.8
神奈川県	4.8	-40.0	-0.3
新潟県	0.1	-1.0	-1.9
富山県	-7.8	5.4	-10.2
石川県	-0.1	10.4	-1.7
福井県	-7.8	-8.5	-9.4
山梨県	4.7	-40.4	1.1
長野県	-5.5	-0.8	-5.0
岐阜県	3.4	4.3	-2.9
静岡県	6.1	-4.9	0.5
愛知県	5.8	7.1	2.2
三重県	8.4	1.0	2.4
滋賀県	6.3	-11.9	-1.7
京都府	3.1	-5.1	0.2
大阪府	12.1	-75.0	-1.9
兵庫県	3.3	-18.3	0.4
奈良県	7.3	-10.0	-1.1
和歌山県	4.3	-33.3	-0.5
鳥取県	2.1	-5.3	-0.1
島根県	-1.7	-2.0	-0.2
岡山県	3.6	-10.5	0.5
広島県	1.0	-2.6	-3.6
山口県	3.5	-7.7	-0.3
徳島県	8.8	-4.2	1.5
香川県	-1.9	6.8	-5.8
愛媛県	5.3	-15.9	0.7
高知県	7.9	-34.2	0.4
福岡県	-60.8	-46.9	-56.4
佐賀県	-46.1	31.9	-29.9
長崎県	3.4	16.5	-0.7
熊本県	-5.2	4.9	-3.8
大分県	2.0	-5.4	-2.3
宮崎県	7.3	-12.6	0.6
鹿児島県	3.2	1.4	-2.9
沖縄県	-7.8	-5.9	-14.9
都府県	-0.1	-1.6	-2.8

注) 乖離率 = (予測値 - 公表値) ÷ 公表値 × 100

また、集落営農の影響が大きかった佐賀県については、集落営農の影響が少ない北九州の3県（長崎県、熊本県、大分県）の推移確立行を用いて予測した結果はやはり乖離率が小さくなった。10年の予測でも誤差が単純に5年の乖離率の2倍程度と考えると一部で問題点はあるもののマルコフモデルの予測モデルとしての有効性は高いとみることができる。

第4表 群馬県における経営耕地面積規模別経営体数の推移と予測

公表値	単位:経営体、%				
	計算方法A		計算方法B		乖離率
	予測値	乖離率	予測値	乖離率	
1ha未満	14,572	9,622	-34.0	13,364	-8.3
1～3ha	8,604	5,956	-30.8	9,706	12.8
3～5ha	1,325	967	-27.0	1,555	17.4
5～10ha	1,226	1,030	-16.0	1,097	-10.5
10ha以上	508	424	-16.5	547	7.7
合計	26,235	17,999	-31.4	26,269	0.1

注) 計算方法Aは群馬県のデータで予測、Bは茨城県のデータで予測

#### IV. 2015年農林業センサスデータによる予測の有効性

##### 1. 「接続不可」経営体数

2010年と2015年農林業センサスについての接続不可割合を第5表に示す。この表から示されるように2015年の接続不可割合は2010年に比べて各都道府県とも低く最も高い沖縄県で18%、全国平均では2.3%である。2010年に比べて接続不可割合が3分の1程度になっており、問題点の少ないデータと考える。

##### 2. 急増した集落営農の動向

2010年段階では集落営農が急増し、この傾向を基に予測を行うと大規模経営体の増加と中小規模経営体の減少を過大推計する恐れが出た。逆に、2015年段階で急増した集落営農が解散等で急減すると大規模経営体の増加、中小規模経営体の減少を過小推計する恐れがある。そこで、2010年新設の大規模経営の動向について第6表に示す。この表で、2010年時点経営体数の欄には、2005年から2010年にかけて新設された

販売金額1位稲かつ経営耕地面積20ha以上の経営体数を示している。この定義が新設の集落営農と一致するとは限らないが、一応の目安と

第5表 構造動態統計の接続不可割合

都道府県	単位:経営体、%			
	2010年		2015年	
	接続不可経営体数	接続不可割合	接続不可経営体数	接続不可割合
北海道	10,776	23.1	1,854	4.6
青森県	1,777	4.0	984	2.7
岩手県	1,892	3.3	869	1.8
宮城県	1,382	2.7	662	1.7
秋田県	925	1.9	867	2.2
山形県	864	2.1	495	1.5
福島県	1,706	2.4	730	1.4
茨城県	1,568	2.2	613	1.1
栃木県	954	2.0	404	1.0
群馬県	11,254	34.6	388	1.5
埼玉県	1,331	2.9	466	1.2
千葉県	1,355	2.4	1,641	3.6
東京都	587	7.9	217	3.6
神奈川県	554	3.5	285	2.1
新潟県	1,673	2.5	668	1.2
富山県	845	3.7	264	1.5
石川県	691	3.9	358	2.6
福井県	943	4.7	313	2.0
山梨県	943	4.4	281	1.6
長野県	3,317	5.2	994	1.8
岐阜県	2,051	5.5	414	1.4
静岡県	1,197	3.0	751	2.3
愛知県	2,222	4.9	782	2.2
三重県	1,343	4.0	405	1.5
滋賀県	971	3.8	317	1.6
京都府	1,202	5.5	566	3.1
大阪府	599	5.6	232	2.5
兵庫県	3,300	5.7	873	1.8
奈良県	799	5.2	201	1.5
和歌山県	920	3.8	348	1.6
鳥取県	926	4.2	438	2.4
島根県	1,077	4.3	373	1.9
岡山県	2,190	4.9	687	1.9
広島県	1,586	4.4	718	2.4
山口県	1,153	4.2	470	2.2
徳島県	968	4.4	227	1.2
香川県	1,073	4.2	257	1.2
愛媛県	1,771	5.3	697	2.6
高知県	1,040	5.5	383	2.4
福岡県	22,888	53.1	604	1.7
佐賀県	615	3.1	281	1.7
長崎県	1,003	3.9	393	1.8
熊本県	1,791	3.7	842	2.0
大分県	1,509	4.9	901	3.5
宮崎県	1,777	5.6	1,098	4.2
鹿児島県	3,403	7.2	2,615	6.7
沖縄県	1,511	9.6	2,759	18.4
全国	106,222	6.3	31,985	2.3

注1) 2010年、2015年農林業センサス第6巻構造動態統計報告書より作成

注2) 接続不可割合 = 接続不可経営体数 ÷ 総経営体数 × 100

第6表 2010年新設された販売金額1位稲かつ  
経営耕地面積20ha以上の経営の動向

都道府県	単位:経営体、%		
	2010年時 点経営体 数	2015年組 織経営継 続	残存率
北海道	24	20	83.3
青森県	33	21	63.6
岩手県	127	102	80.3
宮城県	166	146	88.0
秋田県	250	197	78.8
山形県	169	127	75.1
福島県	18	9	50.0
茨城県	27	24	88.9
栃木県	52	41	78.8
群馬県	12	10	83.3
埼玉県	4	2	50.0
千葉県	8	4	50.0
東京都			
神奈川県			
新潟県	132	122	92.4
富山県	106	98	92.5
石川県	56	50	89.3
福井県	58	55	94.8
山梨県			
長野県	42	30	71.4
岐阜県	34	26	76.5
静岡県	1	1	100.0
愛知県	1		0.0
三重県	8	7	87.5
滋賀県	16	15	93.8
京都府	3	2	66.7
大阪府			
兵庫県	5	4	80.0
奈良県			
和歌山県			
鳥取県	10	7	70.0
島根県	26	23	88.5
岡山県	2	2	100.0
広島県	38	37	97.4
山口県	30	27	90.0
徳島県	1	1	100.0
香川県	26	18	69.2
愛媛県	3	1	33.3
高知県			
福岡県	76	72	94.7
佐賀県	347	309	89.0
長崎県	4	4	100.0
熊本県	98	74	75.5
大分県	16	13	81.3
宮崎県			
鹿児島県			
沖縄県			
全国	2,029	1,701	83.8

注) 2010年、2015年農林業センサス調査表情報  
を基に独自集計

考えられる。実際、新設の集落営農が多かった東北、北陸、北九州で高い数値が示されている。2015年組織経営継続の欄には、これら新設された経営で2015年も存続している経営体数が示されている。多くの経営体は2015年も継続していることがわかる(全国平均で84%)。ここから考えると集落営農の増減による影響は小さいと考えることができる。

### 3. 予測値の有効性

予測モデルとしてのマルコフモデルの有効性に加えて、2015年農林業センサスのデータは2010年と比べても問題点が少ない。また2010年急増した集落営農の多くは2015年段階でもほぼそのまま継続している形で統計が取られている。従って2015年農林業センサスによる予測はかなりの確かな予測値が得られると期待できる。

注1) 神谷(1962)は引用文献「新しい農業分析」でマルコフモデルの有用性について紹介。例としては農業部門の労働力予測を示す。このあと1965年農林業センサスから構造動態統計(当初の名称は「農家調査抽出集計報告書」)が開始された。

注2) 「新設」といっても自給的農家から農業経営体へ移行したケースが多い。

注3) 神谷(1980)は引用文献『日本農業の進路をさぐる』において1975年農林業センサスのデータを基に農家数の予測でマルコフモデルの終局値(50年以上になる)までを算出している。

### 【引用文献】

神谷慶治(1962)「第7章 職業移動マトリックス—マルコフモデルについて—」、神谷慶治・沢村東平監修『新しい農業分析』、東京大学出版会  
神谷慶治(1980)『日本農業の進路をさぐる』、筑波書房