

水稲作況試験データによる重回帰モデル作成システム

佐々木俊彦・日塔 明広・中野 修*

(宮城県農業センター・*宮城県農業普及課)

Production System of Multiple Regression Model by Paddy-rice Crop Situation

Toshihiko SASAKI, Akihiro NIITOU and Osamu NAKANO

(Miyagi Prefectural Agricultural Research Center・*Agricultural Extension)

Service Section of Miyagi Prefectural Government Office

1 はじめに

重回帰式は、従来から作物の生育モデル作成に活用されてきたが、様々な目的変数に対して予測式を作成する場合、気象・栽培データの集計に多くの時間と労力を要した。そのため、データベース化された気象・作況試験データをもとに、主要品種の生育ステージや生育量の重回帰式を作成するシステムを開発したので、その内容を報告する。

2 システムの特徴

システムは、①個々の目的変数に対してあらかじめ設定されている説明変数群(30~60程度)から説明変数を任意に組合せて重回帰式を作成するプログラム、②あらかじめ決められた数の説明変数を生産的に組合せて重回帰式を作成するプログラム、③作成された重回帰式をもとに生育予測を行うプログラムから構成されている。

重回帰式で予測する対象項目は、最高分げつ期・幼穂形成期・出穂期の生育ステージと成熟期を含めた各ステージにおける主な生育量(草丈、茎数、葉数、止葉数、出穂期の稈長・ m^2 当穂数、成熟期の m^2 当玄米重・玄米千粒重・登熟歩合等)である。

Cross-validation 法(以下CR)を利用して予測誤差を計算し、作成された重回帰式の予測精度が確認できる。

気象データと作況試験の調査データは、データファイル

として独立に管理され、新しい年次のデータがファイルに追加される度に、年毎に新しい重回帰式が作成できる。

3 プログラムの構造

(1) 重回帰式の任意作成プログラム

本プログラムの操作は、最初に目的変数と説明変数のデータを検索するために、各種のパラメータを設定する。

説明変数は、対象とする目的変数に応じて、作況試験データ(草丈、茎数、葉数等とその2乗値)及び名取圃場の気象データ(任意期間の最高・最低・平均気温、日較差、日照時間とその2乗値)から構成されている。

対象品種は宮城県農業センターが行っている作況試験の中から主に10年以上のデータの蓄積のある品種を対象とする。気象と作況データの対象期間は概ね10~15年の範囲で指定する。次に作況と気象のデータファイルから指定されたデータを検索して、各目的変数に対して用意されている30~60程度の説明変数群から、3~5個の説明変数を選択して重回帰分析を行い、精度の高い重回帰式の各種係数(決定係数、定数、回帰係数、目的・説明変数名等)をファイルへ保存する。

(2) 重回帰式の自動作成プログラム

図1は重回帰式の自動作成プログラムの概要図である。目的変数と説明変数のデータを検索するところまでは(1)とほぼ同様である。それ以後の処理は、各目的変数に対して

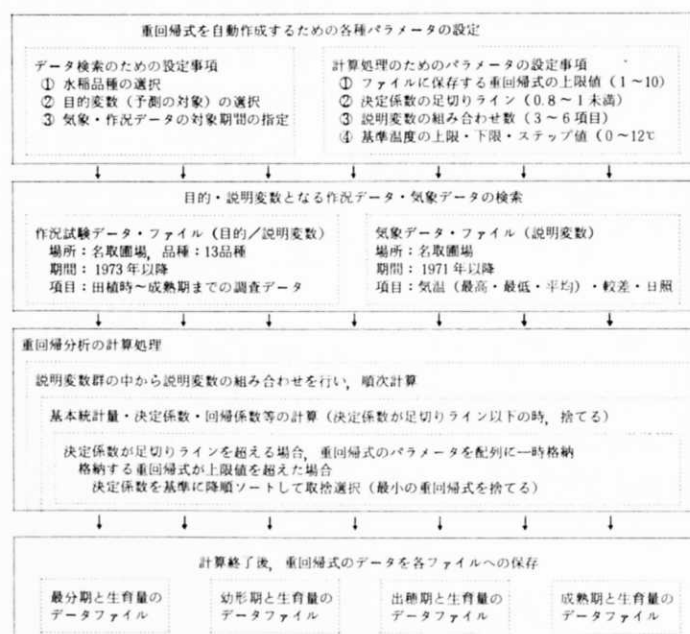


図1 重回帰モデルの自動作成手順の概要図

用意されている説明変数群から、3～5個の説明変数の組合せを作り、順次重回帰分析を行う。更に計算された重回帰式の中から決定係数の高い順に上位数個を選抜して、それぞれの重回帰式の各種係数等をファイルへ保存する。

(3) 生育予測プログラム

重回帰式の各種係数等を保存したファイルから任意の式を選択し、説明変数に生育量や気象のデータを当てはめて予測を行う。この場合、1～3項目の説明変数に刻みを付した11ランクのデータを当てはめ、予測値を1～3次の表形式で出力できる。

4 重回帰式の作成例

表1は、上記プログラムで作成された重回帰式の一例で

表1 システムで作成された重回帰式の例 (ササニシキ・稚苗)

目的変数	決定調査	重回帰式の	CRの予測	1.	2.	3.	4.	5.
	係数	時期	標準偏差	標準偏差	変数名	変数名	変数名	変数名
					変数名	変数名	変数名	変数名
					／回帰係数	／回帰係数	／回帰係数	／回帰係数
①最分期	1.796	田植期	1.796	3.562	776.67	草丈 0.2717	最低5/10 ～6/30 -131.242	最高5/20 ～6/30 15.0179
②幼形期	0.895	田植期	0.895	1.477	68.852	草丈 7.274	葉数 -6.492	較差6/1 ～7/10 -1.5633
③止葉期	0.146	最分期	0.146	0.206	5.943	茎数 0.0014	葉数(*) 6.400	較差 10～25(*) 0.024
④出穂期	1.159	最分期	1.159	2.259	202.30	草丈 -0.1375	最高10～25 34.963	平均10～30 -50.92
⑤m ² 当穂数	8.494	最分期	8.494	12.932	388.24	草丈 4.6516	日照 2.8819	較差 10～25(*) 3.675
⑥登熟歩合	3.216	出穂期	3.216	4.503	224.13	m ² 当穂数 -0.228	平均 1穂数 -1.005	日照1～30 2.731
⑦玄米千粒重	0.276	出穂期	0.276	0.442	65.665	m ² 当穂数 -0.2195	m ² 当穂数(*) 2.116	日照 1～20 1.173
⑧m ² 当玄米重	13.788	出穂期	13.788	21.245	489.57	m ² 当穂数 -0.1663	最高1～30 -89.442	平均1～35 94.640

注. 調査期は、重回帰式の中で、生育量の説明件数を調査した時期である。データの期間は1976～1990年である。最高・最低・平均は～気温、較差は気温の日較差、*は自乗値、/は百分の一、//は一万分の一である。自由度補正済みの決定係数である。

表2 Cross-Validation法による予測誤差の例

(表1の⑧の重回帰式)

目的指数: m²当玄米重, 重相関係数: 0.9772

決定数: 0.9549 決定係数(自由度調整済み): 0.9368

No.	年次	調査値	予測値	予測誤差	予測値(CR)	予測誤差(CR)
1	1976	449 kg	431.78 kg	17.22	424.85	24.15
2	1977	476	490.96	-14.96	514.26	-38.26
3	1978	551	537.95	13.05	530.16	20.88
4	1979	558	556.88	1.12	556.62	1.38
5	1980	449	443.53	5.47	439.80	9.20
6	1981	430	439.47	-9.47	466.42	-36.42
7	1982	418	426.31	-8.31	428.07	-10.07
8	1983	436	430.69	5.31	426.48	9.52
9	1984	530	539.79	-9.79	541.66	-11.66
10	1985	559	584.53	-25.53	600.67	-41.67
11	1986	478	466.29	11.71	464.05	13.95
12	1987	484	487.27	-3.27	487.77	-3.77
13	1988	361	365.80	-4.80	371.44	-10.44
14	1989	451	457.01	-6.01	458.15	-7.15
15	1990	598	569.75	28.25	560.86	37.14
予測誤差の標準偏差				13.79		21.25

このことから本システムでは、重回帰分析にCRを併用していることによって、現実には即した精度の検証が可能で

ある。表2は、m²当玄米重を目的変数として作成された重回帰式(表1の⑧式)について、CRを併用しなかった場合と併用した場合の各年次別の予測値と予測誤差の計算結果であり、前者の予測誤差の標準偏差は13.79であるが、後者の計算結果は21.25であった。

これは、一般の重回帰分析は予測式を作成したデータの中の当該年次のデータも含まれるため、常に予測誤差を過小に評価しており、一方CRを併用した場合は、当該年次のデータを除いた他の年次のデータによって作成された重回帰式から、当該年次の予測値及び予測誤差を計算しているため誤差が大きくなるためである。

ある。

5 おわりに

このシステムを使用することによって、水稻の生育モデルとして高い精度の重回帰式が容易に作成されるが、CRによる予測誤差の計算から、自由度調整済みの決定係数で概ね0.9以上の精度が必要であると考えられる。

作成された重回帰式は、宮城県農業センターの圃場で観測された気象データを使用しているため、この重回帰式を他の地域へ適用する場合、詳細な検討が必要である。

アメダスの気象データやメッシュ気候値を使用した場合の予測精度の検証は、今後の課題である。

参考文献

- 1) 竹澤邦夫, 田村良文, 小野祐幸, 1989, 作物の発育ステージのノンパラメトリック推定の有効性, 農業気象45(3): 151-154.