

イチゴ収穫シミュレータに関する研究

竹中智哉*・安部貞昭**

*電子・情報担当・**大分県農林水産研究指導センター

Research on the simulator to predict the harvest of strawberries

Tomoya TAKENAKA*・Sadaaki ABE**

*Electronics and Information Technology Section・**Oita Prefectural Agriculture, Forestry and Fisheries Research Center.

要 旨

平成 25 年度から、産業科学技術センターと大分県農林水産研究指導センターは、大分県のイチゴの基幹品種であるさがほのかについて、生産性や品質の向上のために環境情報および生育情報から収穫期および収量を予測する共同研究を実施している。現在、園芸作物の環境要因が収穫に与える影響に関する研究は黎明期にあたり、環境情報をどのように取り扱い、また組合せて制御した場合に、生産性を向上させるかについては十分な知見が得られていない。そこで、本年度は、大分県農林水産研究指導センターにて、これまでに得られた情報から環境要因(温度や CO2 濃度など)や生育要因が収量に与える影響を分析した。また、分析結果をもとに、収量を予測する回帰式を導出して考察し、予測精度および有意性について検証した。

1. はじめに

本県では、イチゴを戦略品目として生産振興を図っており、個別農家の単収向上が求められている。現在、生産現場では単収 5t/10a を目標とした技術指導が行われているが、経営体の持続的発展のためには、6t/10a が必要である。単収 6t/10a を実現するためには、栽培期間を通した適切な栽培管理が重要であり、収穫期や収量を予測できる収穫シミュレータが必要とされている。収穫シミュレータを用いた栽培管理システムができれば、環境情報(温湿度、CO2 濃度、灌水量など)および生育情報(開花日、花数など)から収穫期や収量を予測し、予測結果に基づいて栽培環境を管理(制御)することで、単収向上やマーケットニーズに応じた収穫パターンの実現を容易にできる。

これまで、国内において農作物の収穫期および収量と環境要因(原因変数)の関係については、主要作物である水稲などの穀物類やオランダの先進事例があるトマトの研究が盛んに行われている。一方で、イチゴなどのトマト以外の園芸作物は、研究が黎明期にあたり、収穫予測のための十分な知見が得られていない。園芸作物は、種類や品種によって性質が多岐にわたるため、園芸作物一般の包括的な解析は難しく、個別に解析する必要がある。

そこで、本年度は、大分県農林水産研究指導センター(以後、農研センター)において、本県の基幹品種であるさがほのかについて得られた情報から環境要因(温度や CO2 濃度など)や生育要因が収量に与える影響を分析した。また、分析結果をもとに、収量を予測する回帰式を

導出して考察し、予測精度と有意性を検証した。

2. データセットの作成

分析に用いたデータおよび変数について解説する。分析には、共通の条件(Table 1)で栽培した豊後大野市三重町の農業研究部場内ビニルハウス(間口 6m、長さ 30m)における 2011 年度と 2012 年度の栽培記録を用いた。

Table 1 栽培条件

栽培方法	・大分方式 Y 型高設栽培システム ・株間 20cm(栽植密度 833 株/a) ・換気温度 28℃ ・炭酸ガス無施用	
作業日	2011 年度	2012 年度
採苗,定植	7/4, 9/15	7/3, 9/18
天井ビニル被覆	10/18	10/19
黒マルチ被覆	10/24	10/22
収穫調査	1 区 10 株 2 反復	1 区 8 株 3 反復
	毎週月, 水, 金曜(祝日を除く)	

目的変数である収量(Yield)は、日別(1日当りの)収量であるが、収穫調査日以外の日は収穫していない。そこで、実測された収量における 7 日間の移動平均を日別収量とした。

説明変数としては、積算有効温度(Temp)[℃ day]と積算昼間 CO2 濃度(CO2)[ppm day]、積算昼間日射量

(Solar)[MJ/m² day], 花序果重指数(Flower), 経過日数指数(Day)を検討した。積算有効温度は, 作物の生育に寄与する基準温度を超えた日平均温度を積算したものである。昨年¹⁾の研究結果¹⁾を踏まえ, 成熟期間との相関が最も高い基準温度を 10℃とした値を用いた。積算昼間 CO₂ 濃度および積算昼間日射量は, 日出から日没までの日平均値を積算したものである。

分析に用いた目的変数および説明変数は, 日単位の時系列データである。昨年¹⁾の研究結果¹⁾のとおり, 目的変数である収量は花序と関係があるため, 時間経過が収量と関係していると考えられる。そこで, 時間経過による趨勢的傾向を表す指数として, 花序果重指数および経過日数指数を算出した。

花序果重指数は, 農研センターイチゴチームが算出した。さがほのかの果房内における花の着生は一定の法則があり, 最初に咲いた花から次の花序へ進むときの開花間隔は, 時期により一定の傾向がある。そして, 開花間隔日数と花序が進むに従い, 果重が小さくなる。この法則を花序果重指数として表した。

経過日数指数は, 頂果の収穫開始日からの経過日数(以後, 経過日数)と収量の関係を表した指数である。経過日数ごとに 2 カ年の平均値を算出し, 収量の 2 カ年全体平均値で割ることで指数を求めた。Fig.1 に頂果房における収量と経過日数のグラフを示す。

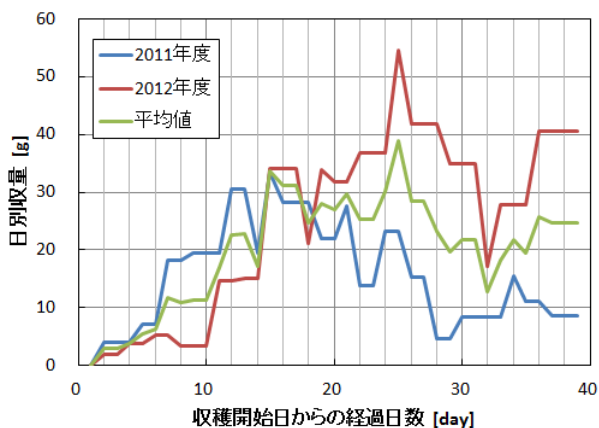


Fig. 1 頂果房における収量と経過日数

2 カ年を比較すると, 20 日までの増加傾向や 24 日ごろのピークとそれ以降の下降傾向など, 趨勢的傾向に類似点が見られる。これより, 経過日数指数を説明変数として検討することに意義があると考えた。

3. 収量と説明変数の相関分析

重回帰分析を行う前に, 頂果房収穫期におけるデータの特性を把握するため, 説明変数の基本統計量や収量と説明変数との相関係数および散布図を確認した。相関係

数と散布図を, Table 2 と Fig.2 に示す。昨年¹⁾の研究結果¹⁾から成熟日数は, 30 日から 60 日であるため, 説明変数の積算期間は 30 日とした。

Table 2 相関係数

	Yield	Temp	CO ₂	Z.CO ₂	Solar	Flower	Day
Yield	1.00						
Temp	-0.41	1.00					
CO ₂	-0.27	0.85	1.00				
Z.CO ₂	0.55	0.32	0.28	1.0			
Solar	-0.68	0.46	0.44	-0.44	1.00		
Flower	0.46	-0.60	-0.13	-0.05	-0.29	1.00	
Day	0.68	-0.45	-0.08	0.02	-0.36	0.68	1.00

Table 2 から, 標準化された積算昼間 CO₂ 濃度(以後, Z.CO₂)と積算昼間日射量, 経過日数指数の収量との単相関係数が 0.5 を超えており, 相関があることがわかる。Z.CO₂ は, 式 (1) を用いて, 年度毎に積算昼間 CO₂ 濃度のデータを平均値 0, 不偏標準偏差 1 となるように変換した値である。

$$Z_i = \frac{x_i - \bar{x}}{\hat{\sigma}} \quad (1)$$

$$\text{ただし } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \hat{\sigma} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$$

ここで, Z_i は積算昼間 CO₂ 濃度の i 番目データの Z.CO₂, x_i は積算昼間 CO₂ 濃度の i 番目データ, \bar{x} は平均値, n はデータ数, $\hat{\sigma}$ は不偏標準偏差を示す。Z.CO₂ を算出した理由については後述する。

また, 説明変数間の相関は 0.9 を超えていないので, 多重共線性(以後, マルチコ)が起きる可能性は低いことがわかる。マルチコとは, 説明変数間に直線関係があるなど, 説明変数間の関係が強い状態をいう。マルチコが起きると, 算出された偏回帰係数が不安定になったり, 偽相関によって本来は目的変数に直接関係のない説明変数でも回帰式に取り込まれてしまったり, 結果の信頼性が落ちてしまう。

次に, 収量と説明変数の関係を散布図で確認すると, 積算有効温度は, 2011 年度が凸型で 2012 年度が負の相関となっている。(Fig.2 (a)) 相関関係を明らかにするには, データ数をさらに増やして分析する必要がある。

積算昼間 CO₂ 濃度は, 2011 年度と 2012 年度で値が大きく異なり, 年度毎に正の相関となっている。(Fig.2 (b-1)) 積算期間における日別昼間 CO₂ 濃度の平均値と不偏標準偏差は, 2011 年度が 438ppm, 33.40ppm, 2012 年度が 381ppm, 29.75ppm であった。今回の栽培条件では, 炭酸ガス無施用であり, CO₂ 濃度の分布に大きな違いは起きにくい。年度間で差を生じたのは CO₂ センサの校正不

備等が考えられるため、本分析では標準化を行い、Z.CO2 を用いることとした。

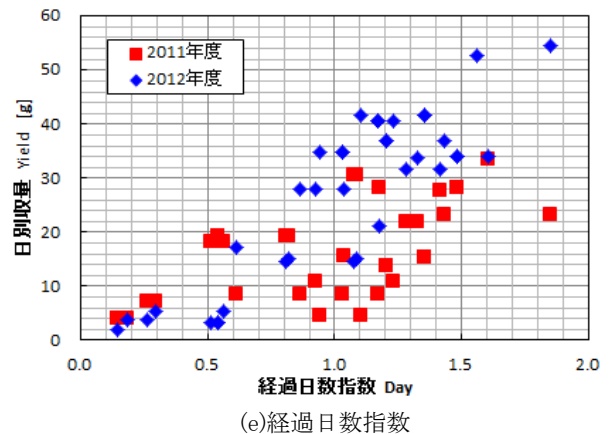
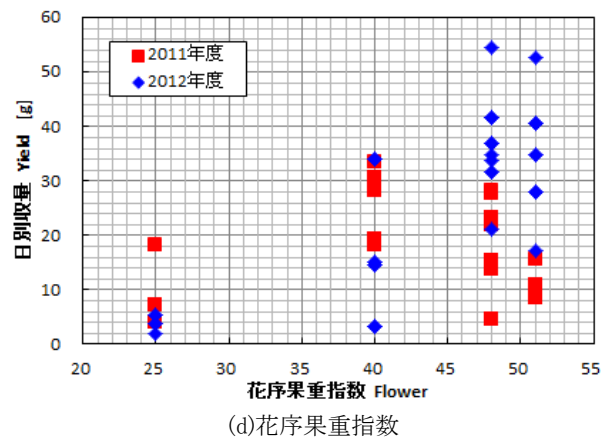
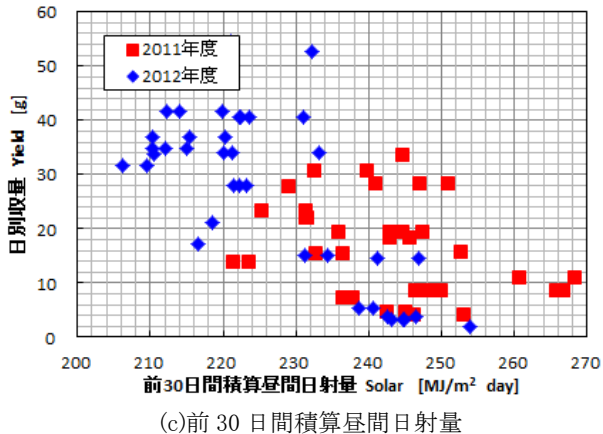
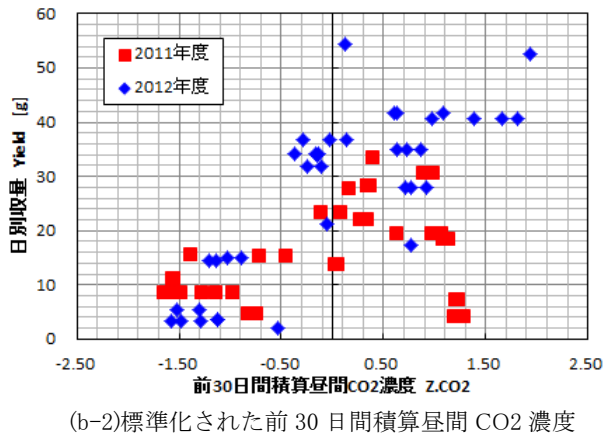
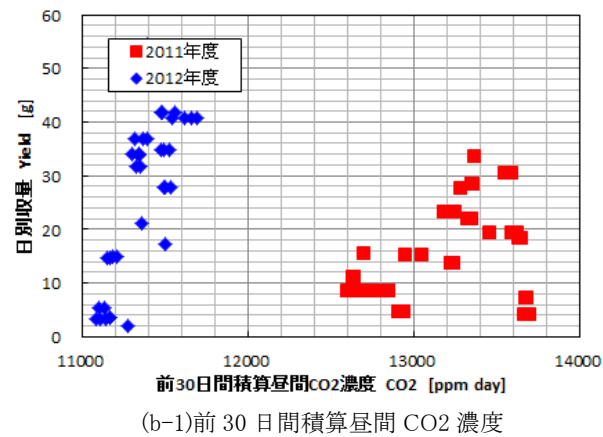
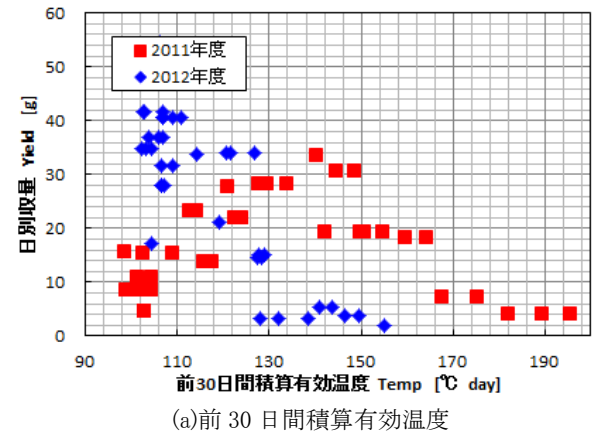


Fig. 2 収量に対する散布図

Fig.2 (b-2)に Z.CO2 の散布図を示す．概ね正の相関となっている．

積算昼間日射量は負の相関となり，経過日数指数は正の相関となっている．(Fig.2 (c), (e)) 花序果重指数は，バラツキが大きく相関関係が弱い．(Fig.2 (d))

4. 収量予測式の導出と考察

重回帰分析を行い，頂果房における収量の予測式を導出した．頂果房における予測技術が確立すれば，その他の果房も同様の手法で予測できると考えられる．統計計算には「エクセル統計 2012」((株)社会情報サービス社製)とフリーソフト「R」を用いた．重回帰分析の結果を Table3 に示す．変数選択には，変数選択基準を AIC として変数増減法を用いた．変数増減法では，Z.CO2 および経過日数指数，積算有効温度の他に花序果重指数も選択された．しかし，花序果重指数は偏回帰係数と単相関係数が異符号となり，マルチョコが疑われたため，除外して再び変数選択を行い，Table3 の回帰式を求めた．Table3 の説明変数は偏回帰係数と単相関係数が同符号であり，VIF の値は 10 より小さいため，マルチョコはないと判断できる．偏回帰係数は各説明変数の重みを表し，標準偏回帰係数は説明変数同士の単位を合せているため，値の大

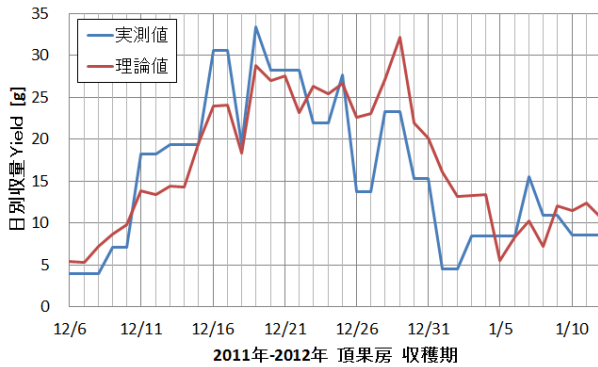
Table 3 重回帰分析の結果

$$\hat{Y} = 9.2101x_{Z.CO2} + 15.3965x_{Day} - 0.1017x_{Temp} + 22.0034$$

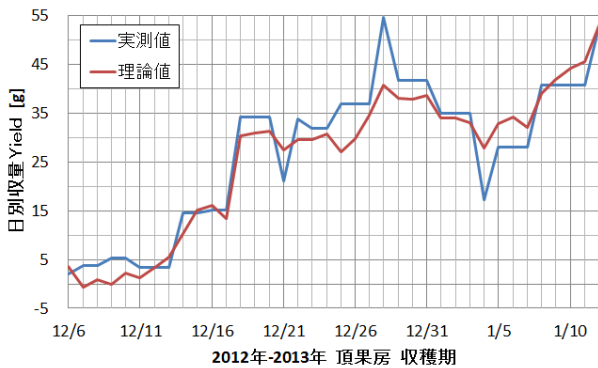
説明変数	偏回帰係数	標準誤差	標準偏回帰係数	F 値	t 値	P 値	単相関係数	VIF
積算昼間 CO2 濃度 Z.CO2	9.2101	0.6190	0.6750	221.4223	14.8803	1.162E-23	0.5541	1.1555
経過日数指数 Day	15.3965	1.5391	0.4807	100.0742	10.0037	2.908E-15	0.6799	1.2964
積算有効温度 Temp	-0.1017	0.0125	-0.4136	66.6376	-8.1632	7.603E-12	-0.4133	1.4414
定数項	22.0034	3.0826		50.9485	7.1378	6.153E-10		
重相関係数 R=0.9337, 決定係数 R ² =0.8718, 自由度調整済み寄与率 修正 R ² =0.8664								
分散分析 (包括的検定)	F=163.1786, P=4.9264E-32							
ダービン・ワトソン比	0.9163			テコ比 平均値			0.053	

きを比較でき、説明変数の目的変数への影響度を比較する際に用いる。Table3 より、Z.CO2>経過日数指数>積算有効温度の順に収量に対する影響度が高いことがわかる。一方で今回の分析では、日射量は収量に対して影響度が高いとはいえない結果となった。昨年の研究結果¹⁾においても、CO2 は果重に影響を与える分析結果が出ており、本回帰式はそれを裏付けた。また、時間経過による趨勢的傾向を表す指数として、本回帰式では経過日数指数がよいという結果になった。本指数では、収量に対する花序の影響も考慮されたと考えられる。この指数は、データの収集数を増やすことで、より相関が高くなる可能性がある。

次に、回帰式の精度について考察する。頂果房の収量



(a)2011 年度



(b)2012 年度

Fig. 3 頂果房 収量の時系列グラフ

における実測値と回帰式による理論値との比較を Fig.3 に示す。理論値は実測値と近い値となり、大まかな傾向をつかめている。そして、回帰式の精度を判断する指標である自由度調整済み寄与率は、Table3 のとおり 0.8664 と 1 に近く、実測値にあてはまりがよいことを示している。

最後に、重回帰分析の結果に問題がないかを検証した。まずは、回帰係数および回帰式の有意性を確認するために、説明変数に対して検定を、回帰式に対して分散分析を行った。結果、P 値は Z.CO2 が 1.162E-23、経過日数指数が 2.908E-15、積算有効温度が 7.603E-12、回帰式が 4.9264E-32 となった。これらの P 値は、有意水準 0.001 よりも小さく、回帰係数および回帰式の有意性が示された。この検定および分散分析は、誤差に対して以下の仮定を置いている。

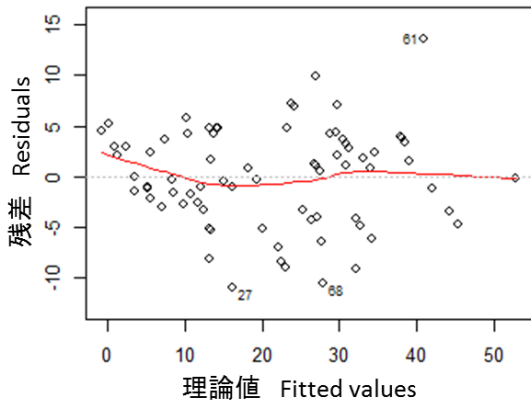
- (1) 誤差項は互いに独立 (独立性)
- (2) 誤差項の期待値 (平均) は 0 (不偏性)
- (3) 分散 σ^2 は説明変数に依存せず一定 (等分散性)
- (4) 誤差項は正規分布にしたがう (正規性)

この仮定が満たされているかを検証した。

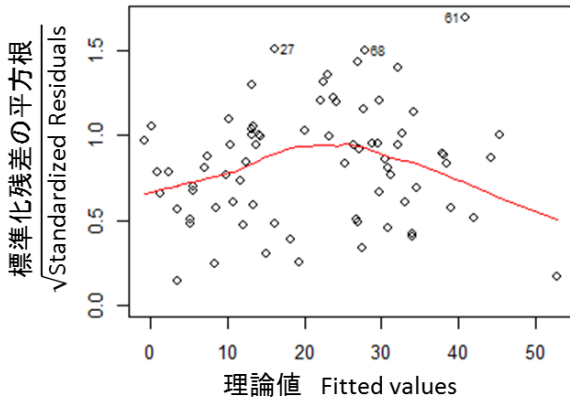
(1) の独立性は、ダービン・ワトソン比で検証できる。ダービン・ワトソン比は、その前のデータの誤差の影響を受けている (系列相関がある) かを調べる統計量である。本データのように、時系列データは系列相関があることが多い。検証の結果、ダービン・ワトソン比は、0.9163 で 2 よりも小さいため、正の系列相関があり、誤差は独立性を満たしているとはいえない。本データは時系列データであり、仕方がない結果ではあるが、独立性を満たしていないことを注意して、検定と分散分析の結果を取り扱う必要がある。

(2), (3) の不偏性と等分散性は、理論値と残差 (実測値と理論値の差) の散布図から判断する。この散布図を Fig.4 に示す。標準化残差の平方根とは、標準化した残差の絶対値の平方根を表す。Fig.4 (a) より、平均 0 を中心として残差がばらばらで、プロットに大きな偏りは見られないため、不偏性と等分散性は問題ないと考えられる。ただし、Fig.4 (b) のとおり、標準化残差の平方根

と理論値の関係は、やや凸型の曲線となる。したがって、説明変数として 2 乗項や 3 乗項を追加するとよい可能性がある。



(a)残差と理論値



(b)標準化残差の平方根と理論値

Fig. 4 残差と理論値(フィット値)のプロット

(4) の正規性は、正規確率プロット (QQ プロット) から判断する。この正規確率プロットを Fig.5 に示す。点がほぼ直線上に並んでおり、残差の分布は正規分布に近く、正規性は問題ないと考えられる。

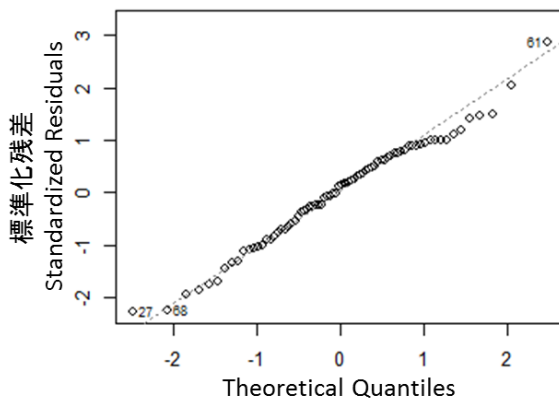


Fig. 5 残差の正規確率プロット

次に、外れ値がないかを確認した。外れ値判断の基準

としては、下記が用いられる。

- (1) テコ比：平均値×2.5 以上
- (2) 標準化残差：絶対値 3.00 以上
- (3) クックの距離：0.5 以上

この基準をもとに、Table3 のテコ比の平均値と Fig.6 の残差と影響力プロットで確認したところ、基準を超えるデータはなく、外れ値はないと判断した。ただし、61 番目(2012 年 12 月 28 日(金)), 68 番目(2013 年 1 月 4 日(金))のデータが、他のデータよりも離れている。

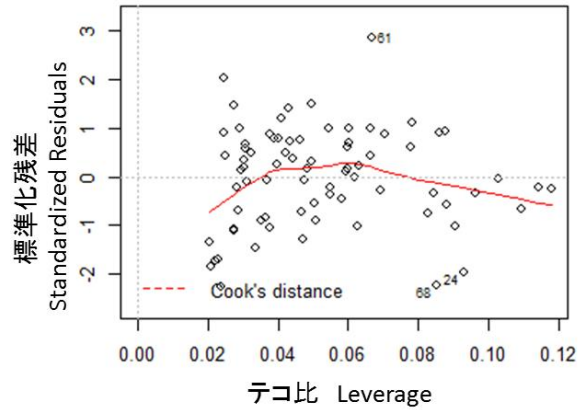


Fig. 6 残差と影響力プロット

12 月 28 日は年末年始休暇前の最後の収穫日であり、休暇期間はしばらく収穫できないことから、9 部着色になる前の果実も収穫した影響ではないかと考えられる。そのため、休暇明けの 1 月 4 日の収量が減っているのではないかと推測できる。

以上の検証結果から、時系列データであるため独立性は満たしていないが、その他の検証項目は問題なく、重回帰分析の結果は有意であると考えられる。

5. まとめ

本年度は、農研センターにおいて、さがほのかについて得られた情報から環境要因や生育要因が収量に与える影響を分析した。また、分析結果をもとに、頂果房の収量を予測する回帰式を導出して考察し、予測精度および有意性についても検証した。得られた回帰式は、説明変数として、積算昼間 CO₂ 濃度と経過日数指数、積算有効温度が選択され、積算昼間 CO₂ 濃度の影響度が最も高かった。そして、自由度調整済み寄与率が 0.8664 であり、予測精度が高いことがわかった。今回扱ったデータは時系列データであり系列相関があることから、独立性は満たしていないが、その他の検証項目は問題なく、重回帰分析の結果は有意であると結論付けた。

昨年からの研究成果を併せてまとめると、収穫期の予測には、積算有効温度と開花日の情報を用いる。具体的

には、基準温度を 10℃とした積算有効温度が、開花日から 190℃ day となる日が収穫開始の予想日となる。そして収量の予測には、本研究と同様の手順で導出する回帰式を用いる。頂果房を除く、その他の果房についてはまだ分析できていないが、頂果房と同様の分析で対応できると考えられる。この予測技術をもとにすれば、収穫シミュレータを構築することができる。

ニーズにあった収穫シミュレータを構築するためには、収量予測のさらなる高精度化が必要である。高精度化の方法としては、以下の方法が考えられる。

- (1) 説明変数の追加
- (2) 標本サイズ(データ数)の追加
- (3) 他の回帰分析手法の検討

(1) については、農研センターの研究で収量と関係があることが明らかになっている液肥の情報などが考えられる。液肥の情報とは、液肥の種類や灌水量、高設ベッドから排出される液肥(排液)の量と EC である。また、既出の説明変数も積算日数や統計期間を最適化したり、平均値とは別の最大値や最小値、レンジ(最大値と最小値の差)といった統計量から最適なものを選択したりする必要がある。(2)は、統計解析において一般的な手法であり、Fig.2 (a)の積算有効温度のような傾向をつかめていない説明変数の傾向を明らかにすることができる。(3)は、非線形解析等の適用が考えられる。一般的に、生育データ間には非線形な関係になることも多く、Fig.4 (b)の分析結果から説明変数として 2 乗項や 3 乗項を追加した方がよい可能性があることも明らかになっている。

今後は、大分県内の農家やデータロガーを製造するメーカーとの共同研究を検討しており、国の競争的資金の獲得に向けて話を進めている。その他にも、大分県内の 5 戸のイチゴ農家(玖珠町 1 戸、国東市 1 戸、大分市 3 戸)に、当センターにて昨年度に試作した排液量センサ¹⁾・²⁾を接続したデータロガー((株)四国総合研究所製ハッピー・マインダー)が設置され、データ収集が開始されている。(Fig.7)今後の解析では、これらのデータの活用も期待できる。



Fig. 7 県内農家へのセンサ設置(玖珠町 ひぐま農園)

謝 辞

本研究へのデータ提供をして頂き、多大なるご支援を頂いた大分県農林水産研究指導センター 佐藤 如氏に心より御礼申し上げます。また、昨年度の排液量センサの試作に加え、県内農家への排液量センサの設置作業にご協力して頂き、多大なるご支援を頂いた機械・金属担当の水江 宏氏に心より御礼申し上げます。そして、県内農家への排液量センサの設置作業および設置後のメンテナンスを担当して頂いている大分県農林水産研究指導センター 椎原 誠一氏に心より御礼申し上げます。

参考文献

- (1) H25 年度経常研究 研究報告書「イチゴ収穫シミュレータに関する研究」(2014)
- (2) H25 年度ハウス内環境制御技術普及促進事業 研究報告書「施設園芸用液肥の排液量計測(転倒ます式)に関する研究」(2014)