

A study on the estimation of onion's bulb weight using multi-level model

Junki Kim^a · Seung-cheon Choi^a · Jaehwi Kim^a · Hong-Seok Seo^{a,1}

^aDepartment of Agricultural Outlook, Korea Rural Economic Institute

(Received August 24, 2020; Revised October 5, 2020; Accepted October 11, 2020)

Abstract

Onions show severe volatility in production and price because crop conditions highly depend on the weather. The government has designated onions as a sensitive agricultural product, and prepared various measures to stabilize the supply and demand. First of all, preemptive and reliable information on predicting onion production is essential to implement appropriate and effective measures. This study aims to contribute to improving the accuracy of production forecasting by developing a model to estimate the final weight of onions bulb. For the analysis, multi-level model is used to reflect the hierarchical data characteristics consisting of above-ground growth data in individual units and meteorological data in parcel units. The result shows that as the number of leaf, stem diameter, and plant height in early May increase, the bulb weight increases. The amount of precipitation as well as the number of days beyond a certain temperature inhibiting carbon assimilation have negative effects on bulb weight. However, the daily range of temperature and more precipitation near the harvest season are statistically significant as positive effects. Also, it is confirmed that the fitness and explanatory power of the model is improved by considering the interaction terms between level-1 and level-2 variables.

Keywords: multi-level model, onion bulb weight, interaction effect, above-ground growth factor, meteorological factor

1. 서론

양파는 주로 노지에서 재배되는 작목으로 기상에 따른 작황 변동성이 크고, 수요와 공급이 가격에 비탄력적인 기초 농산물이어서 수급과 가격의 변동폭이 크다. 이에 정부는 수매비축, 산지폐기, 출하정지, 수입비축 등 다양한 수급조절 대책을 마련하는데, 이를 위해서는 정확한 수요량과 생산량 예측이 필수적이다. 농산물의 특성상 수요는 단기에 큰 변화가 나타나지 않아, 생산량을 결정하는 재배면적과 단수의 신뢰도 높은 예측치가 요구된다. 양파 재배면적은 통상 4월 중순에 발표하기 때문에, 정부는 향후 작황 변동 가능성을 감안한 단수 예측치에 근거하여 추가적인 수급 안정대책을 발표함으로써 수급불안을 해소하고자 한다. 본 연구에서는 생산량 비중이 큰 중만생종 양파를 대상으로 5월 초 기상부 생육요인과 5월 초부터 수확기까지 기상요인이 최종 구중에 미치는 효과를 추정함으로써 향후 양파 생산량 예측치의 정확도를 높이는데 기여하고자 한다.

양파 생산량 예측과 관련된 선행연구는 통계청에서 발표하는 시·도 단위의 양파 단수를 예측하는 연구와 기상변수 등 환경요인을 인위적으로 통제하여 실험한 결과를 토대로 분석한 연구로 분류된다. 양파 단수

¹Corresponding author: Models Research Team, Department of Agricultural Outlook, Korea Rural Economic Institute, 601, Bitgaram-ro, Naju-si, Jeollanam-do 58321, Korea. E-mail: hongseokseo@krei.re.kr

예측은 다양한 모형을 이용하여 연구되어 왔다. Choi와 Back (2016)은 양파 주산지 시군의 기상정보를 활용하여 공간가중패널회귀모형으로 양파 단수를 추정하였다. Lim 등 (2016)은 농촌진흥청의 상세기상정보와 흙토람 농경지 정보를 활용하여 다중회귀모형 구조의 단수예측모형을 구축하였다. Lee와 Yang (2017)은 기상자료를 입력변수로 활용한 인공지능망을 이용하여 양파 단수를 예측하고, 패널모형, 적응기대모형과 예측력을 비교하였다. Jung (2019)은 최고기온, 최저기온, 상대습도, 강수량 등의 기상자료와 초장의 변화량 자료를 이용하여 Stepwise, LASSO, PCA, PLS의 4가지 통계기법을 활용한 단수 결정모형 구축을 시도하였다. Lee 등 (2020)은 도 단위 기상자료와 영농투입재 비용을 함께 고려한 패널회귀분석 고정효과모형을 이용하였다. 위 연구들은 종속변수인 단수와 독립변수에 사용된 자료들 간에 공간 단위의 불일치가 존재한다. 종속변수인 양파 단수는 통계청에서 선정한 표본구를 대상으로 집계한 자료이다. 이에 반해 독립변수로 도입된 기상자료는 해당 도 지역에 위치한 기상청 중관기상관측(ASOS) 지점 자료를 토대로 산출한 값이며, 생육 자료, 영농투입재 자료, 토양 자료 등도 통계청의 단수 산출에 사용된 표본과 달라 추정 결과의 신뢰도를 담보하기 어렵다.

해외에서는 1970년대부터 양파의 구 비대와 수량의 결정요인을 탐색한 연구를 활발하게 수행하였다. 양파 구 비대에 영향을 미치는 환경조건으로는 일장, 온도, 광, 시비량 등이 복합적으로 작용하는 것으로 알려져 있다 (Riekels, 1972; De Ruiter, 1986; Henriksin, 1987). 특히 구 비대에는 일장과 온도가 중요한 요인이며, 장일형 양파는 14시간 이상, 단일형 양파는 11시간 이상의 일장조건이 되면 구 비대가 시작되는 것으로 알려져 있다 (Huh 등, 2002). 이러한 한계일장 이상에서는 온도가 높을수록 구 비대가 빠르고 (Brewster, 1990), 고온 조건이 되면 한계일장이 짧아진다고 알려져 있다 (Steer, 1980). 대표적인 국내 연구로 Ahn과 Suh (2013)은 양파 구 비대의 결정 요인 중 하나인 일장을 조절하여 생육에 미치는 영향을 분석하였다. Song 등 (2017)은 생육기 온도 상승이 극조생 양파의 생육과 구 비대에 미치는 영향력을 검증하였고, Lee 등 (2019)은 양파의 고온과 다습조건이 생육, 생리반응 및 수량에 미치는 영향을 규명하였다. 이상의 연구들은 관심 있는 환경요인을 통제하는 실험을 통해 특정한 처리에 따라 변화하는 양상을 확인하는 연구로서 다양한 환경요인을 복합적으로 고려하지 못했다는 한계가 있다. 더군다나 양파는 주로 노지에서 재배되는 작목으로 통제된 실험 환경에서 도출된 결과를 온전히 신뢰하기 어렵다.

선행연구의 한계를 극복하고자 본 연구에서는 한국농촌경제연구원(Korea Rural Economic Institute; KREI)에서 실시한 생육 실측조사 자료와 국립농업과학원에서 추정된 필지의 기상자료를 토대로 양파의 최종 구중을 추정하는 다층모형 구축을 시도하였다. 초장, 엽초경, 엽수 등 지상부에서의 양분 흡수량과 성장은 수확기 40일 전 전후로 감소하면서 구의 비대가 빠르게 진행되기 때문에 (Lee, 2015) 이 시점을 최종 구중의 결정요인 탐색과 수량을 추정하는 시점으로 삼았다. 자연적으로 노출된 외부 환경에서 수집한 생육 변수를 사용하였으며, 최종 구중을 비롯한 생육변수 자료와 기상변수 자료의 공간적 단위를 일치시켰다는 점에서 신뢰도가 높다고 판단된다. 또한 기존 농산물 수량 추정 연구에서는 t -검정, 분산분석이나 회귀분석과 같이 단층구조의 분석기법을 주로 사용하였으나, 본 연구에서는 위계적 구조를 가진 자료의 분석기법을 활용하였다는 점에서 방법론적 의의가 있다. 본 논문은 분석에 사용된 생육자료와 기상자료의 위계적 구조를 제 2장에서 기술하고, 제 3장에서는 위계적 구조를 갖는 자료를 분석하는 경우 적합한 통계기법인 다층모형에 대해 설명한다. 제 4장에서는 분석결과를 제시하며, 제5장에서는 결론 및 향후과제를 제안한다.

2. 분석자료

2.1. 생육 실측조사 자료

2020년 KREI에서 실시한 양파의 생육 실측조사는 양파를 5년 이상 재배한 농가 중 106개 농가를 대상으

Table 2.1. Elementary statistics of growth data by lot

Time	Growth factor	Mean	Std	Min	Max
5월 초 (7회차 조사)	초장(cm)	81.99	10.89	48.14	99.36
	엽수(매)	8.86	1.02	6.20	11.13
	엽초장(cm)	16.67	3.07	7.59	22.05
	엽초경(mm)	21.94	3.28	16.03	29.54
수확기 (최종 조사)	생구무게(g)	319.59	75.56	186.29	577.50

Table 2.2. Agricultural climate detailing method based on ASOS

Weather factor		Detailed description	Resolution(m)
기온	최고	관측소별 최고기온 관측값을 IDW 방법으로 공간 내삽한 후 동적 해발 고도 보정, 일사효과를 적용	30
	최저	관측소별 최저기온 관측값을 IDW 방법으로 공간 내삽한 후 동적 해발 고도 보정, 도시 열섬효과, 온난대 보정, 냉기류 보정을 적용	
	평균	최고기온과 최저기온의 평균	
강수량	PRISM 모형을 이용		270
일조시간	시간별 음영기복도를 이용하여 추정		
일사량	관측소별 일적산일사량 관측값을 IDW 방법으로 공간 내삽한 후 수평 면일사량의 지형 보정치인 상대일사량지수 분포도를 이용하여 보정		

IDW = inverse distance weighting; PRISM = parameter-elevation regression on independent slopes model.

로 진행하였다. 조사 시기는 2020년 2월 15일부터 6월 11일까지 총 11회를 실시하였으며, 조사는 15일 주기로 하되 수확 임박 시에는 10일 간격으로 실시하였다. 조사 대상 농가의 대표 필지 1곳을 3개 조사구역으로 구분한 후 각 조사구역에서 5개체씩 지상부와 지하부의 생육자료를 수집하였다.

표본교체, 병해발생, 측정오류, 측정불가 등을 제외한 모든 조사 회차에서 조사가 완료된 필지로 한정하여 본 연구의 신뢰도를 높이고자 하였다. 국내 양파 생산량의 80% 수준을 차지하는 중만생종 양파를 분석 대상으로 삼았으며, 측정 수치의 정확도가 의심되는 개체는 제외하였다. 본 연구에서는 총 106개 필지 중 41개 필지의 564개 개체를 대상으로 구비대 시점인 5월 초(7회차 조사)의 지상부 생육자료와 최종 수확기의 생구 무게 자료를 사용하였다. 필지 단위 생육자료의 기초통계량은 Table 2.1과 같다. 수확기 생구 무게의 평균은 319.59g, 표준편차는 75.56g으로 5월 초 지상부 생육에 비해 넓게 분포하는 것을 볼 수 있다.

2.2. 소기후 모형의 기상 자료

현재 기상청 기상데이터포털 중관기상관측(ASOS)자료는 관측소가 설치된 시군 단위의 기상 정보를 얻을 수 있다. 하지만 생육 실측조사 대상 필지의 기상 정보는 얻을 수 없어 구중에 영향을 미치는 기상의 한계 효과를 정확히 추정할 수 없다. 본 연구에서는 국립농업과학원이 소기후 모형을 통해 도출한 필지 단위의 일별 기상자료 추정치를 사용하였다. 소기후 모형은 우리나라 전역의 필지 단위까지 공간적으로 정밀한 농업기상/기후 자료를 추정할 수 있는 모형이다. 일별 기상청 중관기상자료를 기반으로 각 요소별 상세화 기술은 Table 2.2와 같다. 이를 적용하여 30~270m의 격자 해상도의 전국 기본 기후자료를 계산한 다음 생육 실측조사의 필지 단위로 일별 평균기온, 최고기온, 최저기온, 강수량, 일조시간, 일사량의 6가지 기상 정보를 추정한다 (Kim 등, 2019).

6가지 기본 기상자료 외에 여러 선행연구를 참고하여 양파 생육에 영향을 미치는 기상 파생변수를 추가

Table 2.3. Display weather derived variable

Derived variable	Define	Formula
증발산량	증발량과 증산량의 총합, 즉 지면에서 대기로부터 증발한 수증기량과 식물의 잎 표면에서 대기로부터 증산된 수분의 양을 합한 것 (FAO Penman-Monteith 증발산 모형)	$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} U_2 (e_s - e_a)}{\Delta + \gamma(1+0.34U_2)}$ $R_n = \text{순복사량 (MJm-2day-1)}, G = \text{토양열속밀도 (MJm-2day-1)}$ $\gamma = \text{건습계상수 (kPa°C-1)}, T = 2m \text{ 높이의 일평균기온 (°C)}$ $U_2 = 2m \text{ 높이의 평균 풍속 (ms-1)}, e_s = \text{포화수증기압 (kPa)}$ $e_a = \text{실제 수증기압 (kPa)}, \Delta = \text{기온에 따른 포화수증기압 기울기}$
생육도일(GDD)	평균기온이 4°C 이상인 날들의 (평균기온-4) 값의 합	$\sum_i \max(0, X_i - 4), X_i = \text{평균기온}$
최적기온일수	최고기온이 15°C ~ 25°C 인 날들의 수	$\sum_i I(15 \leq X_i^{\max} \leq 25), X_i^{\max} = \text{최고기온}$
일교차	최고기온-최저기온의 평균	$E[X_i^{\max} - X_i^{\min}], X_i^{\max} = \text{최고기온}, X_i^{\min} = \text{최저기온}$
고온일수	최고기온이 30°C 이상인 날들의 수	$\sum_i I(X_i^{\max} \geq 30), X_i^{\max} = \text{최고기온}$
강우일수	일강수량이 5mm 이상인 날들의 수	$\sum_i I(Y_i \geq 5), Y_i = \text{일강수량}$
강우피해지수	전날 강수량과의 합이 50mm 이상이 되는 날의 일 강수량	$\sum_i Y_i^{dm}, Y_i^{dm} = \text{전날 강수량과의 합이 50mm 이상이 되는 날의 일 강수량}$
탄소동화저해일수	최고기온이 25°C 이상인 날들의 수	$\sum_i I(X_i^{\max} \geq 25), X_i^{\max} = \text{최고기온}$
수확전강수량	수확일 전 이틀간의 누적 강수량	$X_{d-1} + X_{d-2}, X_d = \text{수확당일 강수량}$

Table 2.4. Elementary statistics of weather data by lot

Weather factor	Mean	Std	Min	Max
평균기온(°C)	18.09	0.43	17.31	19.28
최고기온(°C)	23.99	1.01	21.91	25.50
최저기온(°C)	12.19	0.86	9.94	13.62
일조시간(시간)	6.15	0.21	5.68	6.54
증발산량	3.70	0.26	3.29	4.48
일교차(°C)	11.80	1.67	8.55	14.74
수확전강수량(mm)	4.93	8.50	0.00	31.84
강우피해지수(mm)	20.51	26.48	0.00	105.72
일사량(MJ/m ²)	665.87	34.05	573.28	707.17
강수량(mm)	110.43	34.00	72.19	188.48
GDD(°C)	596.92	14.06	571.15	636.22
최적기온일수(일)	4.15	1.46	2.00	8.00
고온일수(일)	1.07	1.01	0.00	3.00
탄소동화저해일수(일)	12.68	5.22	2.00	21.00
강우일수(일)	5.56	1.47	3.00	8.00

GDD = growing degree days.

생성하여 분석에 활용하였다(Table 2.3). 파생변수를 포함한 모든 기상자료는 구 생장이 빠르게 증가하는 시기인 5월 초부터 최종 수확기인 6월 초까지 약 40일 동안의 기상이 생육에 어떠한 영향을 미치는지 확인하기 위하여 기상변수의 특성에 따라 누적 또는 평균값으로 변환하였다. 예를 들어 일강수량은 분석기간의 누적강수량으로 산출하였다. 이를 통해 Table 2.4는 필지 단위 기상자료의 기초통계량을 보여준다. 본 연구에서 사용된 생육자료 및 기상자료와 생육 실측조사에 관한 상세한 내용은 농업관측통계정보시스템(OASIS)에서 확인할 수 있다(<http://oasis.krei.re.kr/index.do>).

2.3. 위계적 자료

생육자료는 양파 개체 단위로 수집되었으나, 최종 구중에 영향을 미치는 기상자료는 필지 단위로 수집되

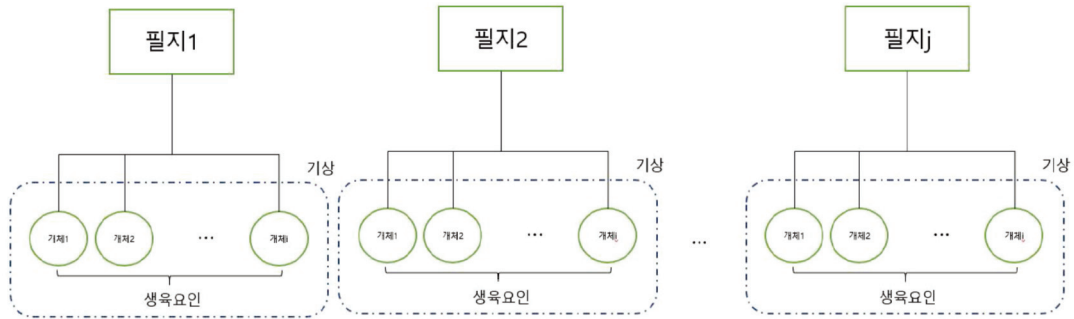


Figure 2.1. Diagram of hierarchical multilevel data structure.

어 있어 본 연구의 분석자료는 Figure 2.1과 같은 위계적 다층자료(hierarchical multi-level data) 구조를 갖는다. 동일한 필지에 있는 양파 개체들은 온도, 강수량, 일사량 등 수많은 기상요인을 공유하므로, 이들의 최종 구중은 서로 영향을 미친다. 즉 동일한 필지 내 개체들의 최종 구중은 독립적이지 않고 어느 정도 상관성을 갖는 반면, 다른 필지에 있는 양파 구중과는 독립적이다. 필지가 다르면 구중에 영향을 미치는 기상여건이 상이하하여 다른 필지에 있는 양파와는 상관성을 갖지 않는다. 관찰단위인 양파 구중은 동일한 필지 내에서는 상호 의존성을 갖지만, 필지가 다르면 독립성을 갖는다. 이는 위계적 다층자료의 대표적인 특성으로 생육요인의 개체 수준과 기상요인의 필지 수준을 모두 분석 단위로 받아들일 수 있는 모형 설정이 필요하다.

3. 모형구조 및 설명

3.1. 다층모형(multi-level model)

다중회귀분석이나 분산분석과 같은 전통적인 통계분석이 가정하는 자료의 구조는 단층구조로 하나의 분석 단위를 고려한다. 하지만 본 연구에서 사용되는 자료는 1-수준인 개체 단위의 생육자료와 2-수준인 필지 단위의 기상자료로 위계적 구조를 갖는다. 동일한 필지 내 양파 개체들은 동일한 기상여건에 있고, 개체 간 상호작용이 존재한다. 전통적인 회귀분석모형으로 개체 단위 자료를 분석하면 독립성의 가정을 위반하므로 오차분산이 과소 추정되고 1종오류(type I error)가 증가하여 가설검정의 타당성이 위협을 받는다 (Burstein, 1980; Cooley 등, 1981; Cronbach, 1976; Cronbach와 webb, 1975). 반면 개체별 생육변수의 값을 필지별 평균값으로 집계하여 2-수준 단위에서 분석하면 분석 단위의 독립성 가정을 충족하므로 오차분산 추정량의 정밀도에는 문제가 없다. 그러나 여러 가지 필연적인 한계를 갖게 되는데 첫째, 필지 단위에서 분석을 하면 표본크기가 불과 41개인 소규모 자료로 전환되어 모든 추정량의 정밀도가 낮아지고 분석 결과의 신뢰도가 낮아진다. 둘째, 작은 표본으로 인하여 모형에 포함하는 독립변수의 수가 제한되고, 많은 독립변수를 포함하면 모형의 결정계수(R^2)의 거표효과가 나타난다. 셋째, 필지 단위 통계 모형이므로 개별 개체의 정보를 반영하는 생육정보 변수의 사용 가능성을 원천적으로 차단하는 모형이다. 넷째, 필지 수준에서 관찰된 변수만 분석에 사용하므로 모든 해석이 필지의 특성들 사이의 관계로 해석되어야 하는데, 양파 개체 수준으로 해석하는 생태학적 오류(ecological fallacy)를 범할 수 있다 (Robinson, 1950). 왜냐하면 필지 수준 모형에서 추정된 회귀계수는 집계화의 효과(aggregation effect)로 인해 개체 수준 모형으로 추정된 값보다 크게 되며, 이 결과는 필지 수준에서의 정보로 해석을 제한해야 한다. 본 연구에서는 개체가 필

지에 속하는 위계적 구조의 자료 분석에 적절한 통계모형인 다층모형을 사용하였다. 수준별로 모형을 분리하여 설명하면, 1-수준 모형은 각 필지별 양과 개체 수준의 회귀분석모형이다.

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{1ij} + \beta_{2j}X_{2ij} + \cdots + \beta_{pj}X_{pij} + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma^2), i = 1, 2, \dots, n_j, j = 1, 2, \dots, J. \quad (3.1)$$

위 식에서 Y_{ij} 는 j 필지에 속한 i 개체의 최종 구중이며, X_{ij} 는 각 개체별 수확기 40일 전 지상부 생육변수(엽수, 초장, 엽초경 등)이다. β_{pj} 는 j 번 필지에서 p 번째 생육변수(X_{pij})가 종속변수 Y 에 대한 효과를 반영하는 모수이다. 이때 계수 β_{pj} 는 농가에 따라 값이 다를 수 있는데, 예를 들어 초장이 최종 구중에 영향을 미치는 영향은 필지에 따라 다를 수 있다. 오차항 ϵ_{ij} 는 1-수준의 잔차이며 정규분포를 갖고 각 필지의 잔차 분산은 등분산성을 갖는다.

β_{pj} 는 각 필지별로 다르므로, 이 값은 필지의 특성에 따라 달라지게 되며 β_{pj} 의 필지 간 차이를 설명하는 모형이 2-수준 모형이다.

$$\beta_{pj} = \gamma_{p0} + \gamma_{p1}W_{1j} + \gamma_{p2}W_{2j} + \cdots + \gamma_{pQ}W_{Qj} + v_{pj}, \quad p = 0, 1, 2, \dots, P. \quad (3.2)$$

2-수준 모형은 종속변수가 $(P + 1)$ 개인 다변량 모형이며(multi-variate model), 2-수준에서 잔차는 분산-공분산행렬을 갖는다. 즉 $[v_{0j}, v_{1j}, v_{2j}, \dots, v_{pj}]^T \sim N(0, T_j)$ 이며, T_j 는 각 필지별로 $(P + 1) \times (P + 1)$ 차원의 분산-공분산 행렬이다.

식 (3.2)를 식 (3.1)에 대입하면 2-수준 다층모형의 일반식을 도출할 수 있다.

$$\begin{aligned} Y_{ij} &= \gamma_{00} & (1) \\ &+ \sum_{q=1}^{Q_0} \gamma_{0q}W_{qj} & (2) \\ &+ \sum_{p=1}^P \gamma_{p0}X_{pj} & (3) \\ &+ \sum_{p=1}^P \sum_{q=1}^{Q_p} \gamma_{pq}W_{pq}X_{pij} & (4) \\ &+ v_{0j} + \sum_{p=1}^P v_{pj}X_{pij} + \epsilon_{ij}. & (5) \end{aligned} \quad (3.3)$$

위의 식 (3.3) 중 (1)은 절편계수, (2)는 2-수준 변수인 기상요인의 고정효과, (3)은 1-수준 변수인 지상부 생육요인의 고정효과, (4)는 1-수준 변수와 2-수준 변수간의 상호작용 고정효과, (5) 랜덤효과를 나타낸다. 여기서 랜덤효과는 필지별 생육요인 효과를 기상요인 변수로 설명한 이후에 남은 1-수준과 2-수준 잔차의 합을 의미한다.

다층모형 분석에서는 중심화 변환을 사용하여 모형 절편에 구체적인 의미를 부여하고 특정 설명변수가 종속변수에 미치는 효과를 직관적이고 명료하게 해석하고자 한다. 이러한 중심화 변환은 다중공선성 문제를 보완해주는 이점이 있어 모형 추정에 필수적이며 이는 통계지수인 분산팽창지수(variance inflation factor)를 통하여 확인한다. 중심화 변환 방법은 다양한데 연구자의 분석 목적에 따라 변환 방법을 선택한다. 연구자들은 보통 2-수준 변수는 '전체평균 중심화'로 1-수준 변수는 '집단평균 중심화' 또는 '전체평균 중심화'

를 선호한다. 1-수준 변수를 전체평균으로 중심화하면 설명변수를 필지에 상관없이 모두 동일한 조건으로 통계한 후 분석한다는 장점이 있다. 하지만, 본 연구에서는 식 (3.4)와 같이 중심화 변환 방법 중 필지별 회귀선의 신뢰도가 가장 높고 1-수준 오차도 가장 작은 집단평균 중심화로 1-수준 변수를 변환하였다 (Kang, 2016).

$$\begin{aligned} \text{1-수준모형중심화변환: } X_{ij} &\rightarrow (X_{ij} - \bar{X}_{.j}), \\ \text{2-수준모형중심화변환: } W_j &\rightarrow (W_j - \bar{W}). \end{aligned} \tag{3.4}$$

다층모형 분석 결과로 얻을 수 있는 유용한 정보로 집단 내 상관계수와 오차감소비율이 있다. 집단 내 상관계수(intraclass correlation coefficient; ICC)는 양파의 최종 구중에 영향을 미치는 1-수준 변수(지상부 생육)와 2-수준 변수(기상)의 영향력은 어느 정도인지 확인하는 값이다. 식 (3.5)와 같이 모형 내 전체 분산에서 2-수준 단위인 필지 간 분산이 차지하는 비율로 산출된다.

$$ICC = \frac{\text{Var}(\beta_{0j})}{\text{Var}(Y_{ij})} = \frac{\tau}{\sigma^2 + \tau} = \text{Corr}(Y_{ij}, Y_{i'j}) = \frac{\text{Cov}(Y_{ij}, Y_{i'j})}{\sqrt{\text{Var}(Y_{ij})\text{Var}(Y_{i'j})}}, \quad (i \neq j). \tag{3.5}$$

또한, ICC는 가장 우측 식과 같이 동일 필지 내에 있는 양파 개체 간의 상관계수로 정리된다. 따라서 ICC는 두 가지 해석이 가능한데, 전체 관찰분산에서 필지 간 분산이 차지하는 비율과 동일 필지 내 개체들 간의 상관계수를 의미한다.

오차감소비율(proportional reduction in error; PRE)은 다층모형에서 1-수준 변수, 2-수준 변수, 교호작용 등을 추가함으로써 분산의 변화량을 확인하여 모형의 개선 정도를 확인하는 수치이다. PRE값을 통해 독립 변수의 효과가 어떤 수준에서 발생하며, 각 수준별로 어느 정도의 효과 크기를 갖는지 분석이 가능하다. 산출방식은 식 (3.6)과 같이 기존 모형 대비 새로운 모형의 분산이 얼마나 감소하였는지 비율로 계산한다. 개선 전 모형의 1-수준 분산은 $\sigma_{previous}^2$ 개선 후 모형의 1-수준 분산은 σ_{next}^2 이고 마찬가지로 τ 는 서로 다른 2-수준의 분산을 의미한다. $PRE_{level-1}$ 는 1-수준 분산에서 $PRE_{level-2}$ 은 2-수준 분산에서 그리고 PRE_{total} 는 전체 분산에서 감소한 비율을 의미한다.

$$\begin{aligned} PRE_{level-1} &= \frac{\sigma_{previous}^2 - \sigma_{next}^2}{\sigma_{previous}^2}, \\ PRE_{level-2} &= \frac{(\tau_{00} + \tau_{11} + \tau_{10})_{previous} - (\tau_{00} + \tau_{11} + \tau_{10})_{next}}{(\tau_{00} + \tau_{11} + \tau_{10})_{previous}}, \\ PRE_{total} &= \frac{\sigma_{previous}^2 + (\tau_{00} + \tau_{11} + \tau_{10})_{previous} - \sigma_{next}^2 + (\tau_{00} + \tau_{11} + \tau_{10})_{next}}{\sigma_{previous}^2 + (\tau_{00} + \tau_{11} + \tau_{10})_{previous}}, \\ \text{where, } \epsilon_{ij} &\sim N(0, \sigma^2), \quad \begin{pmatrix} v_{0j} \\ v_{1j} \end{pmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix} \right). \end{aligned} \tag{3.6}$$

다층모형 고도화에 따른 모형의 적합도 개선을 통계적 기법으로 확인하기 위해 우도비율검정(likelihood ratio test; LRT)을 사용하였다. 비교 모형과 기존 모형의 최대 우도함수 값을 각각 추정하여 그 차이를 반영하는 편차량을 이용하여 적합도를 평가한다. 우도비율검정의 가설과 검정통계량은 식 (3.7), 식 (3.8)과 같으며, 로그 변환하여 정리한 식 (3.9)와 같이 χ^2 분포의 검정통계량을 가지는 LRT를 통하여 모형의 적합도 개선이 통계적으로 유의한지 검정한다.

$$H_0: \theta = \theta_0; \quad H_1: \theta = \theta_1, \tag{3.7}$$

Table 3.1. Definition for variables by the level

Factor		Variable						
Growth factor (level-1)	x1	엽수	x2	엽초경	x3	엽초장	x4	초장
	w1	강수량	w2	강우일수	w3	강우피해지수	w4	고온일수
Weather factor (level-2)	w5	수확전강수량	w6	일교차	w7	누적 일사량	w8	일조시간
	w9	증발산량	w10	최고기온	w11	최저기온	w12	최적기온일수
	w13	탄소동화저해일수						

$$\Lambda(x_1, x_2, \dots, x_n) = \frac{L(\theta_0|x_1, x_2, \dots, x_n)}{L(\theta_1|x_1, x_2, \dots, x_n)}, \quad (3.8)$$

$$\lambda_{LR} = -2\ln(\Lambda(x_1, x_2, \dots, x_n)) = -2(\ln(L(\theta_0|x_1, x_2, \dots, x_n)) - \ln(L(\theta_1|x_1, x_2, \dots, x_n))). \quad (3.9)$$

1-수준과 2-수준, 교호작용의 설명변수 선택에 따라 다양한 모형이 존재하는데 (Akaike information criterion; AIC)와 (Bayesian information criterion; BIC) 값을 기준으로 적합성을 비교하여 최적 모형을 선택하였다. 산출 방법은 식 (3.10)과 같으며, $L(\theta) = L(\theta|x_1, x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n|\theta)$, K 는 모형에서 추정할 모수의 개수이며 n 은 표본 크기이다.

$$\begin{aligned} \text{AIC} &= -2\log(L(\theta)) + 2K, \\ \text{BIC} &= -2\log(L(\theta)) + K\log(n). \end{aligned} \quad (3.10)$$

위계적 자료의 분산 성분을 추정하기 위해 최대우도 추정법(maximum likelihood; ML)을 사용하였다. ML은 표본의 가능도를 최대화하고 표본의 관측값과 모형의 예측값 간 차이를 최소화시키는 방식으로 모수를 찾는 추정방법이다. 따라서 위계적 자료와 같이 자료의 불균형 문제에 대해 자유로우며, 표본의 크기가 클수록 모수 값에 수렴하는 특성인 일치성과 추정량 가운데서도 분산이 작은 특성인 효율성을 갖는다. 또한 상위 수준인 필지 표본수가 41개로 적지 않기 때문에 랜덤효과의 과소 추정 위험도가 낮다 (Finch 등, 2014; Snijders와 Bosker, 2011).

3.2. 모형 선택 과정

개체 단위 지상부 생육자료와 필지 단위 기상자료를 활용한 최적의 다층모형을 선택하기 위하여, 다음과 같은 과정을 통해 모형을 선택하였다. 모형 선택은 AIC, BIC, LRT의 모형적합도를 기준으로 판단하였으며, 가용할 수 있는 변수들의 모든 조합 중 AIC 값을 기준으로 가장 낮은 5개 모형을 나타낸다. Table 3.1은 본 연구에서 고려한 생육요인과 기상요인 변수이다.

(Step1) 1-수준 요인인 지상부 생육 변수만을 활용하여 최종 구증을 추정하는 적정모형을 선택하며, 상수항과 개체 변수에 랜덤효과 적용 여부도 같이 고려하였다. Table 3.2에서 1-수준 변수만을 고려한 분석결과 엽초장을 제외한 엽수, 엽초경, 초장 변수가 투입된 G1 모형을 선택했다.

1-수준 모형에서 선택된 모형 G1을 기준으로 랜덤효과의 적용 여부를 위한 분석하였는데, 엽초경을 랜덤효과로 설정한 R1의 적합도가 가장 높았다. 1단계에서 선택된 R1 모형은 식 (3.11)과 같다.

$$\begin{aligned} Y_{ij} &= \beta_{0j} + \beta_{1j}(X_{1ij} - \bar{X}_{1\cdot j}) + \beta_{2j}(X_{2ij} - \bar{X}_{2\cdot j}) + \beta_{4j}(X_{4ij} - \bar{X}_{4\cdot j}) + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma^2), \\ \beta_{0j} &= \gamma_{00} + v_{0j}, \\ \beta_{2j} &= \gamma_{20} + v_{2j}, \end{aligned} \quad (3.11)$$

Table 3.2. Level-1 model selection process

	Level-1 model selection process	AIC	BIC	Model selection
G1	엽수, 엽초경, 초장	6647.2	6673.2	
G2	엽수, 엽초경, 엽초장, 초장	6647.4	6677.7	
G3	엽수, 엽초경, 엽초장	6660.4	6686.4	G1
G4	엽초경, 초장	6663.0	6684.7	
G5	엽초경, 엽초장, 초장	6663.2	6689.2	

Table 3.3. Random effect model selection process

	Level-1 + Random effect model	AIC	BIC	Model selection
R1	G1, 엽초경	6622.2	6656.9	
R2	G1, 엽수, 초장	6641.5	6689.2	
R3	G1, 엽수	6637.6	6672.3	R1
R4	G1, 초장	6644.0	6678.7	
R5	G1, 엽초경, 초장	6624.4	6672.1	

Table 3.4. Level-2 model selection process

	Level-1 + Level-2 model	AIC	BIC	Model selection
GW1	R1, 강수량, 수확전강수량, 탄소동화저해일수, 고온일수, 일교차	6615.8	6672.1	
GW2	R1, 강수량, 수확전강수량, 탄소동화저해일수, 고온일수, 최저기온	6616.3	6672.7	
GW3	R1, 강수량, 수확전강수량, 탄소동화저해일수	6616.5	6664.2	GW1
GW4	R1, 강수량, 수확전강수량, 탄소동화저해일수, 고온일수, 일교차, 일조시간	6616.7	6674.6	
GW5	R1, 강수량, 수확전강수량, 탄소동화저해일수, 고온일수, 일교차, 강우일수	6616.7	6675.0	

Table 3.5. Interaction model selection result

	Level-1 + Level-2 + Interaction model	AIC	BIC	Model selection
GW1I	GW1, 엽수*탄소동화저해일수, 초장*고온일수, 초장*수확전강수량	6607.4	6676.8	
GW12	GW1, 엽수*강수량, 초장*고온일수, 초장*수확전강수량	6607.9	6677.3	
GW13	GW1, 엽수*탄소동화저해일수, 초장*고온일수, 초장*수확전강수량, 엽초경*고온일수	6608.0	6681.7	GW1I
GW14	GW1, 엽수*고온일수, 초장*고온일수, 초장*수확전강수량	6608.2	6677.6	
GW15	GW1, 엽수*일교차, 초장*고온일수, 초장*수확전강수량	6608.2	6677.6	

(Step2) 최적의 1-수준 변수와 랜덤효과로 선택한 R1 모형을 기준으로 필지별 기상자료를 추가한 2-수준 모형을 고려하여 기상변수 중 모형에 가장 적합한 변수를 탐색하였다. 최적의 모형은 강수량, 수확전강수량, 탄소동화저해일수, 고온일수, 일교차 변수가 투입된 GW1으로 식 (3.12)와 같다.

$$\begin{aligned}
 Y_{ij} &= \beta_{0j} + \beta_{1j}(X_{1ij} - \bar{X}_{1\cdot j}) + \beta_{2j}(X_{2ij} - \bar{X}_{2\cdot j}) + \beta_{4j}(X_{4ij} - \bar{X}_{4\cdot j}) + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma^2), \\
 \beta_{0j} &= \gamma_{00} + \gamma_{01}(W_{1j} - \bar{W}_{1\cdot}) + \gamma_{04}(W_{4j} - \bar{W}_{4\cdot}) + \gamma_{05}(W_{5j} - \bar{W}_{5\cdot}) + \gamma_{06}(W_{6j} - \bar{W}_{6\cdot}) \\
 &\quad + \gamma_{013}(W_{13j} - \bar{W}_{13\cdot}) + u_{0j}, \\
 \beta_{2j} &= \gamma_{20} + u_{2j},
 \end{aligned}
 \tag{3.12}$$

(Step3) 2-수준을 적용한 GW1 모형을 기준으로 변수 간 교호작용의 유의성 등을 확인하였다. 분석결과, 교

Table 4.1. Results of final multi-level model analysis

	Fixed effect	Estimate	Std. error	t-value	Pr(> t)
Level-1	절편(γ_{00})	312.79	8.85	35.338	0.000***
	엽수(γ_{10})	14.79	3.20	4.619	0.000***
	엽초경(γ_{20})	6.58	1.70	3.882	0.000***
	초장(γ_{40})	1.89	0.48	3.953	0.000***
Level-2	강수량(γ_{01})	-2.12	0.72	-2.947	0.006**
	고온일수(γ_{02})	-39.54	19.00	-2.082	0.044*
	수확전강수량(γ_{03})	4.55	1.78	2.554	0.015*
	일교차(γ_{04})	22.97	12.78	1.796	0.081.
	탄소동화저해일수(γ_{05})	-12.26	4.12	-2.973	0.005**
Interaction terms	엽수*강수량(γ_{11})	0.26	0.10	2.537	0.011*
	초장*고온일수(γ_{44})	-1.20	0.60	-2.011	0.045*
	초장*수확전강수량(γ_{45})	-0.17	0.08	-2.254	0.025*
	Random effect		Variance		Variance
Level-1	오차항(σ^2)		5533.3		51.48
Level-2	랜덤절편(τ_{00})		2650.2		7.09
	랜덤기울기(τ_{11})		50.2		74.39

***p < 0.001, ** p < 0.01, *p < 0.05, .p < .0.1

호작용항인 엽수*탄소동화저해일수, 초장*고온일수, 초장*수확전강수량이 투입된 GWII 모형을 최종 모형으로 선택하였고 식 (3.13)과 같다.

$$\begin{aligned}
 Y_{ij} &= \beta_{0j} + \beta_{1j}(X_{1ij} - \bar{X}_{1.j}) + \beta_{2j}(X_{2ij} - \bar{X}_{2.j}) + \beta_{4j}(X_{4ij} - \bar{X}_{4.j}) + \epsilon_{ij}, \quad \epsilon_{ij} \sim N(0, \sigma^2), \\
 \beta_{0j} &= \gamma_{00} + \gamma_{01}(W_{1j} - \bar{W}_{1.}) + \gamma_{04}(W_{4j} - \bar{W}_{4.}) + \gamma_{05}(W_{5j} - \bar{W}_{5.}) + \gamma_{06}(W_{6j} - \bar{W}_{6.}) \\
 &\quad + \gamma_{013}(W_{13j} - \bar{W}_{13.}) + u_{0j}, \\
 \beta_{1j} &= \gamma_{10} + \gamma_{11}(W_{1j} - \bar{W}_{1.}), \\
 \beta_{2j} &= \gamma_{20} + u_{2j}, \\
 \beta_{4j} &= \gamma_{40} + \gamma_{44}(W_{4j} - \bar{W}_{4.}) + \gamma_{45}(W_{5j} - \bar{W}_{5.}).
 \end{aligned} \tag{3.13}$$

4. 분석결과

모형 선택과정을 통해 최종 선정한 다층모형의 분석결과는 Table 4.1에 제시되어 있다. 우선 고정효과 모수 추정치를 살펴보면 전체 양파 구중의 평균값을 의미하는 절편계수(γ_{00}) 값은 312.8g으로 추정되었다.

지상부 생육변수 중에서 1-수준 변수로 선정된 엽수, 엽초경, 초장은 모두 양의 효과로 유의수준 0.1%에서 통계적으로 유의하였다. 엽수(γ_{10}), 엽초경(γ_{20}), 초장(γ_{40})의 분석단위가 1단위 증가하면 수확기 구중은 각각 14.8g, 6.6g, 1.9g 증가하는 것으로 추정되었다. 구가 비대하는 시기에 지상부에서 탄소동화작용을 통해 생성한 탄수화물과 보유하고 있는 양분을 지하부인 구로 전달하게 되는데, 구 비대 시점인 5월 초에 엽수, 엽초경과 엽초장의 생장 상황이 수확기 구중에 미치는 영향이 큰 것으로 해석된다. Lee (2015)는 엽수가 많은 상황에서 구 비대가 진행되면 잎집의 개수가 많아져 생구가 커진다고 하였다. 기상변수인 2-수준 변수의 고정효과는 선행연구와 유사하게 추정되었다. 우선 강수량(γ_{01})의 한계효과는 -2.1로 1% 수준에서 통계적으로 유의하였다. 양파 구가 굽어지는 시기에 잦은 강수로 토양 수분이 많아지면 잎에 질소 성분이 많아져 구의 비대가 늦춰지고, 토양이 과습되면 뿌리가 빠르게 썩고 병이 발생할 가능성이 높다(Lee, 2015).

Table 4.2. Comparison of multi-level models by level-1, level-2, and interaction

	M1	M2	M3	M4	M5	
Random effect	랜덤절편(τ_{00})	3830.0	4320.7	2807.0	2649.8	2650.2
	랜덤기울기(τ_{11})		61.3	203.9	60.5	50.2
	공분산(τ_{10})		0.35	0.42	0.30	0.27
	오차항(σ^2)	9109.0	5637.0	6066.5	5645.3	5533.3
Fixed effect (Level-1)	절편	316.68***	317.53***	298.44***	312.81***	312.79***
	엽수(매)		14.20***		14.21***	14.79***
	엽초경(mm)		7.09***		7.17***	6.58***
	초장(cm)		1.71***		1.69***	1.89***
Fixed effect (Level-2)	강수량(mm)			-2.23**	-2.20**	-2.12**
	고온일수(일)			-37.63	-37.97	-39.54*
	수확전강수량(mm)			4.77**	4.73*	4.55*
	일교차(°C)			22.76	22.92	22.97
			-12.36**	-12.36**	-12.26**	
Fixed effect (Interaction)	엽수*강수량					0.26*
	초장*고온일수					-1.20*
	초장*수확전강수량					-0.17*
Explanatory power	ICC	29.60	42.36	29.89	31.26	31.80
	PRE(Based on M2)				0.169	0.195
	PRE(Based on M1)		0.212	0.281	0.345	0.362
	AIC	6826.3	6622.2	6678.4	6615.8	6607.9
	BIC	6839.3	6656.9	6721.8	6672.1	6677.3

***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; ·: $p < 0.1$

고온일수(γ_{02})와 탄소동화저해일수(γ_{05})의 고정효과 추정치는 각각 -39.5, -12.3으로 최고기온이 특정 기온을 넘는 일수가 많아질수록 최종 생구 무게는 통계적으로 유의하게 감소하는 것으로 분석되었다. 양파는 고온에서 비교적 약한 작물로서 30°C 이상 온도에서는 잎이 말라버리며 25°C 이상 온도에서는 잎이 탄소동화 능력을 잃어버려 구 비대에 부정적인 영향을 미치는 것으로 알려져 있다 (Lee, 2015). 수확전강수량(γ_{03})이 많을수록 일교차(γ_{04})가 커질수록 수확기 생구 무게는 통계적으로 유의하게 증가하는 것으로 추정되었다. 수확 임박한 시기에 내린 강수로 인해 양파 구의 수분 흡수량이 많아져 최종 구중이 증가한 것으로 판단된다. 또한, 일교차가 크면 마늘과 같이 낮에 저장한 양분의 소모가 적어 상대적으로 많은 양분을 저장할 수 있을 것으로 판단되어 구 성장에 긍정적인 것이라고 보인다. 교호작용은 2-수준 변수에 따라 1-수준 변수의 영향을 변화시키는 효과를 의미한다. 분석결과 엽수와 초장이 생구 무게에 미치는 영향은 강수량, 고온일수, 수확전강수량에 의해 조절되는데 각각 5% 수준에서 통계적으로 유의하게 나타났다. 강수량이 1단위 증가하면 엽수가 최종 구중에 미치는 효과(γ_{11})는 0.26만큼 증가하는 것으로 추정되었다. 즉 강수량은 최종 구중에 -2.1g의 부정적인 영향을 미치지만 엽수와 교호작용으로 인해 생구 증가에 긍정적인 영향을 미치는 것으로 분석되었다. 반면 고온일수와 수확전강수량이 1단위 증가하면 초장의 효과(γ_{44} , γ_{45})는 각각 1.20, 0.17만큼 감소하는 것으로 추정되었다.

다층모형의 단계별 명세화에 따른 구중 결정 모형의 분석결과를 Table 4.2와 Table 4.3에 정리하였으며, 모형 적합도 및 설명력의 개선을 확인할 수 있다. 설명변수를 전혀 고려하지 않은 기초모형(M1)의 ICC는 29.6%로 전체 분산에서 개체 간 분산이 70.4%로 상당히 크게 나타났다. 즉 최종 구중의 예측력을 높이려면 양파 개체별로 갖고 있는 특성인 수확기 40일 전 생육 변인을 고려해야 한다는 것을 알 수 있다. 생육

Table 4.3. Result of likelihood ratio test and random effect test about M5

	Chisq	Df	Pr(> Chisq)
LR test(Based on M1)	244.41	13	0.000***
LR test(Based on M2)	30.29	8	0.000***
LR test(Based on M3)	82.52	6	0.000***
LR test(Based on M4)	13.89	3	0.003**
Random effect test(업초경)	21.53	2	0.000***

***: $p < 0.001$; **: $p < 0.01$; *: $p < 0.05$; ·: $p < 0.1$

변인을 고려한 모형(M2)의 ICC는 42.4%로 개체 간 분산 비중이 큰 폭으로 낮아졌으며, 전체 오차분산은 M1 대비 21.2% 감소하였다. 2-수준 단위인 기상 변인을 고려한 모형(M3)은 M1 대비 오차분산은 28.1% 감소하였으며 우도비율 검정 결과 0.1% 유의수준에서 통계적으로 개선된 것으로 나타났다. 이는 1-수준인 생육 변인뿐만 아니라 2-수준인 기상 변인도 분석단위로 함께 고려하는 다층모형 활용의 당위성을 보여준다. 생육 변인과 기상 변인을 함께 고려한 M4 모형으로 추정된 결과, 전체 오차분산은 더 크게 감소(M1 대비 34.5%, M2 대비 16.9%)하였으며, 우도비 검정을 통해 M2와 M3보다 통계적으로 적합도가 개선된 것을 확인하였다. 1-수준과 2-수준 간 교호작용도 고려한 최종 모형(M5)은 기존 모형의 전체 오차분산보다 작아졌으며, M4 대비 통계적으로 5% 유의수준에서 개선된 것으로 확인되었다.

5. 결론 및 향후과제

양파의 최종 구증을 추정하기 위해 구 비대 초기의 지상부 생육자료와 수확기까지의 추정 기상자료를 활용한 다층모형을 분석모형으로 도입하였다. 분석 결과, 5월 초 업수, 업초경, 초장의 생육이 좋을수록 생구무게는 증가하는 것으로 나타났다. 기상요인에서는 강수량, 고온일수, 탄소동화저해일수가 생구 무게에 음의 효과가 나타났으며, 일교차와 수확전강수량은 양의 효과로 통계적으로 유의하였다. 또한 생육요인과 기상요인의 상호작용을 고려함으로써 독립변수에 대한 조절효과를 부여해 모형의 적합도 향상을 확인하였다. 선행연구에서 주로 사용했던 단층 구조의 전통적인 통계모형이 아닌 위계적 자료의 특성을 반영한 다층모형을 제안하였다는 점에서 농산물 수량 추정 연구 분야에 기여하는 바가 크다고 판단된다. 또한 수확기 40일 전 지상부 생육 정보를 토대로 최종 구증과 생산량을 예측할 수 있는 방안을 모색하였다는 점에서 의의가 있다고 판단된다. 하지만 구증 및 생산량 예측에 적극적으로 활용하려면 본 연구에서 구축한 모형을 기초로 개선 작업을 지속해야 한다. 1년차 자료만으로 분석에 활용되었기에 모형의 신뢰도가 다소 미흡하다. 향후 축적되는 생육자료와 기상자료를 토대로 모형의 정확도 및 신뢰도를 높여야 한다. 또한 최종모형의 랜덤효과 모수 추정치를 살펴보면, 필지별 교정평균의 차이를 기상요인으로 설명한 이후에도 잔차분산(τ_{00} 과 τ_{11})은 통계적으로 유의하게 존재하였다. 이는 기상요인 외에 토양 정보와 시비량 정보 등 제3의 변수들로 필지 간 차이를 설명할 여지가 남아 있다는 의미이다. 향후 실측조사에서는 위와 같은 다양한 생육환경 요인을 수집하여 보다 개선된 최종 구증 예측모형으로 활용되길 기대한다.

References

- Ahn, S. R. and Suh, J. K. (2013). Effect of Different Day Length on the Production of Onion, *Protected horticulture and Plant Factory*, 22, 43-47.
- Burstein, L. (1980). Chapter 4: The analysis of multilevel data in educational research and evaluation, *Review of Research in Education*, 8, 158-233.

- Brewster, J. L. (1990). The influence of cultural and environmental factors on the time of maturity of bulb onion crops, *Acta Horticulturae*, **267**, 289–296.
- Choi, S. and Baek, J. (2016). Onion yield estimation using spatial panel regression model, *The Korea Journal of Applied Statistics*, **29**, 873–885.
- Cooley, W. W., Bond, L., and Mao, B. B. (1981). Analyzing multilevel data. In R. A. Berk (Ed.), *Educational Evaluation Methodology*, Baltimore, MD: Johns Hopkins University Press.
- Cronbach, L. J. (1976). Research on classrooms and schools: Formulation of questions, design, and analysis, *Occasional paper of the Stanford Evaluation Consortium*.
- Cronbach, L. J. and Webb, N. (1975). Between-class and within-class effects in a reported aptitude * treatment interaction: Reanalysis of a study by G. L. Anderson, *Journal of Educational Psychology*, **67**, 717–724.
- De Ruiter, J. M. (1986). The effects of temperature and photoperiod on onion bulb growth and development. *Proceedings Agronomy Society of NZ*, **16**, 93–100.
- Finch, W. H., Bolin, J. H., and Kelley, K. (2014). *Multilevel Modeling Using R*, CRC Press, FL.
- Henriksin, K. (1987). Effect of N- and P-fertilization on yield and harvest time in bulb onions (*Allium cepa* L.), *Acta Horticulturae*, **208**, 207–215.
- Huh, E. J., Cho, K. S., Kwon, Y. S., and Woo, J. G. (2002). Effects of temperature and photoperiod on bulbing and maturity of spring sown onions in highland, *The Journal of Horticultural Science and Biotechnology*, **43**, 587–590.
- Jung, I. (2019). Comparison of the prediction model of the production unit and the area of the onion-comparison of stepwise, LASSO, PCA and PLS methods (Master Thesis), Chonnam National University.
- Kang, S. (2016). Multilevel Model, *Hakjisa*, 290–296.
- Kim, D. J., Kim, S. O., Kim, J. H. and Yun, E. J. (2019). Establishment of Geospatial Schemes Based on Topo-Climatology for Farm-Specific Agrometeorological Information, *Korean Journal of Agricultural and Forest Meteorology*, **21**, 146–157.
- Lee, C. and Yang, S. B. (2017). Development of yield forecast models for onion and garlic using artificial neural network, *Horticultural Science and Technology*, **35**, 128–129.
- Lee, H. J., Lee, S. G., Kim, S. K., An, S., Lee, J. H., Lee, H. S., Kim, C. W., Kwon, Y. S., and Han, J. W. (2019). Effects of High-Temperature and Soil Moisture Conditions on the Physiological Response of Onion, *Horticultural Science and Technology*, **37**, 571–578.
- Lee, J. (2015). Onion: From one year old to organic and no-till cultivation method, *deulnyeok*.
- Lee, S. I., Son, S., and Lee, H. (2020). An analysis of changes onion yields in Korea using panel regression analysis and Bayesian network model, *Journal of Rural Development*, **43**, 1–28.
- Lim, C. H., Kim, G. S., Lee, E. J., Heo, S., Kim, T., Kim, Y., and Lee, W. K. (2016). Development on crop yield forecasting model for major vegetable crops using meteorological information of main production area, *Journal of Climate Change Research*, **7**, 199–203.
- Riekels, J. W. (1972). The influence of nitrogen on the growth and maturity of onions grown on organic soil, *Journal of the American Society for Horticultural Science*, **97**, 37–40.
- Robinson, W. S. (2009). Ecological Correlations and the Behavior of Individuals, *International Journal of Epidemiology*, **38**, 337–341.
- Snijders, T. B. S. and Bosker, R. J. (2011). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, Sage Publishers, London.
- Song, E. Y., Moon, K. H., Wi, S. H., Kim, C. H., Lim, C. K., Oh, S. J., and Son, I. C. (2017). Impact of Elevated Temperature in Growing Season on Growth and Bulb Development of Extremely Early-Maturing Onion, *Korean Journal of Agricultural and Forest Meteorology*, **19**, 223–231.
- Steer, B. T. (1980). The bulbing response to day-length and temperature of some Australasian cultivars of onion (*Allium cepa* L.), *Australian Journal of Agricultural Research*, **31**, 511–518.

다층모형을 활용한 양파 구중 추정 연구

김준기^a · 최성천^a · 김재휘^a · 서홍석^{a,1}

^a한국농촌경제연구원 농업관측본부 모형연구팀

(2020년 8월 24일 접수, 2020년 10월 5일 수정, 2020년 10월 11일 채택)

요약

양파는 기상여건에 따른 작황의 변동성이 커 생산량 및 가격 변화가 크다. 정부는 양파를 수급 민감 품목으로 지정하여 다양한 수급 안정대책을 마련하고 시행하는데 이를 위해서는 선제적이고 신뢰도 높은 양파 생산량 예측 정보가 필수적이다. 본 연구에서는 양파의 5월 초 지상부 생육정보와 5월 초부터 수확기까지의 기상정보를 이용하여 최종 생구 무게에 미치는 영향을 추정함으로써 생산량 예측의 정확도 개선에 기여하고자 한다. 위계적 특성을 갖고 있는 자료를 통해 개체별 생육요인인 1-수준 자료와 필지별 기상요인인 2-수준 자료, 그리고 두 수준 간 상호작용을 고려한 다층모형을 도입하여 분석하였다. 분석 결과, 5월 초에 엽수, 엽초경, 초장의 생육이 좋을수록 최종 생구 무게는 증가하는 것으로 추정되었다. 5월 초부터 수확기까지의 기상요인에서는 강수량, 고온일수, 탄소동화저해일수가 생구 무게에 음의 효과가 나타났으며, 일교차와 수확전강수량은 양의 효과로 통계적으로 유의하였다. 또한 1-수준과 2-수준의 교호작용항을 고려하여 모형의 적합도와 설명력을 향상시켰다.

주요어: 다층모형, 양파 생구무게, 상호작용, 지상부 생육요인, 기상요인

¹ 교신저자: (58321) 전라남도 나주시 빛가람로 601, 한국농촌경제연구원, 농업관측본부, 모형연구팀.
E-mail: hongsookseo@krei.re.kr