

인공 신경망을 이용한 채소 단수 예측 모형 개발 : 고추를 중심으로*

이춘수** · 양성범***

Development of Yield Forecast Models for Vegetables Using Artificial Neural Networks: the Case of Chilli Pepper

Lee, Choon-Soo · Yang, Sung-Bum

This study suggests the yield forecast model for chilli pepper using artificial neural network. For this, we select the most suitable network models for chilli pepper's yield and compare the predictive power with adaptive expectation model and panel model. The results show that the predictive power of artificial neural network with 5 weather input variables (temperature, precipitation, temperature range, humidity, sunshine amount) is higher than the alternative models. Implications for forecasting of yields are suggested at the end of this study.

Key words : *artificial neural network, yield forecast, panel model, adaptive expectation model, root mean square prediction error*

I. 서 론

가격 불안정성이 큰 노지채소의 수급안정을 위한 선제적 대응 측면에서 생산량 예측에 활용되는 단위면적당 생산량(이하 단수) 예측이 중요하다. 기존의 단수 예측은 주로 기상요인, 기술수준 등을 설명변수로 설정한 계량경제학 모형을 통해 이루어졌다. 그러나 단수의 경우 기상, 토양 등 여러 요인으로부터 복합적인 영향을 받고, 해당 요인들과 단수가 비선형적인 관계를 맺고 있다는 점에서 전통적인 계량경제학 모형보다 인공 신경망(artificial

* 본 연구는 농촌진흥청 공동연구사업(과제번호: PJ012262)의 지원에 의해 이루어진 것임.

** 단국대학교 환경자원경제학과 강사

*** Corresponding author, 단국대학교 환경자원경제학과 조교수(passion@dankook.ac.kr)

neural network)이 예측에 적합할 수 있다.

인공 신경망 모형은 단순하고 확정적인 비선형 모형(simple deterministic nonlinear model)이 규칙성을 전혀 가지지 않은 듯이 보이는 상태(purely random process)를 발생시킬 수 있다는 혼돈이론에 기초한다(Yun et al., 2016). 모형 개발 초기 Hu (1964)가 기상예측에 인공 신경망을 활용한 이후 다층 신경망(multi-layer networks)에 대한 훈련 알고리즘 제약으로 제대로 활용되지 못했다. 그러나 Rumelhart 등(1986)이 다층 신경망 훈련이 가능한 역전파 알고리즘(back propagation algorithm)을 개발한 이후 다양한 영역에서 예측에 활용되고 있다(Lamba and Dhaka, 2014).

최근 해외에서는 인공 신경망을 이용하여 단수를 예측한 연구가 활발히 이루어지고 있으나 대부분 옥수수, 밀 등의 곡물 단수 예측에 집중되어 있다(Panda et al., 2011; Laxmi and Kumar, 2011; Guo and Xue, 2014; Lamba and Dhaka, 2014). 국내에서는 모든의 생산성을 예측한 Lee와 Choe(2009), 벼 수확량을 예측한 Sung과 Lee(2005)의 연구에서 인공 신경망을 이용한 바 있다. 그러나 채소 단수 예측에 인공 신경망을 이용한 사례는 전무하고, 채소 단수 예측에 관한 선행연구는 대부분 계량경제학 모형을 이용하였다(Lee, 1996; Lee et al., 2005; Kim et al., 2010; Han et al., 2011; Park and Park, 2013; Lee and Yang, 2017).

본 연구는 기상변수를 입력변수로 활용한 인공 신경망을 이용하여 채소 단수를 예측하고, 패널모형(panel model) 및 적응 기대 모형(adaptive expectation model)과 예측력을 비교한다는 점에서 선행연구와 차별성을 가진다. 단수 예측 품목은 5대 노지채소(배추, 무, 마늘, 양파, 고추) 중 단수 예측 연구가 미흡한 고추로 선정하였다.

이후의 내용은 다음과 같다. 2장에서는 고추 단수 예측에 활용한 인공 신경망, 패널모형 및 적응 기대 모형에 대해 설명하고, 모형별 예측력 비교 방법을 제시한다. 3장에서 단수 예측에 활용된 자료를 설명하고, 4장에는 모형별 예측력 분석결과를 정리하였다. 결론에는 채소 단수 예측력 향상을 위한 함의를 제시하였다.

II. 연구방법

1. 인공 신경망을 이용한 단수 예측 모형

인공 신경망은 인간의 두뇌가 사물을 인지, 추론하는 원리를 컴퓨터에 적용한 모형이다. 인공 신경망 중 가장 널리 사용되는 모형이 오류 역전파 알고리즘(error back-propagation algorithm)을 이용한 다층 퍼셉트론(multi-layer perceptron)이다.

오류 역전파 알고리즘을 이용한 다층 퍼셉트론은 크게 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)으로 구분된다. 입력층에 입력된 정보 $x_{1,t}$ 에서 $x_{p,t}$ 는 식 (1)의

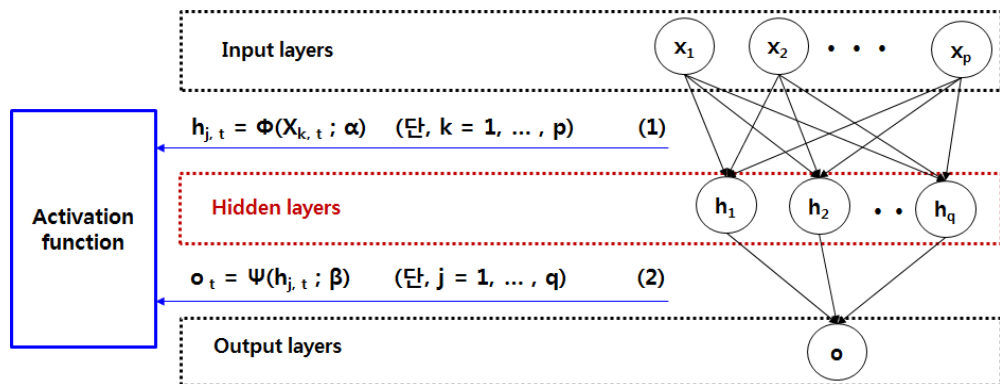
활성화함수 $\Phi(\cdot)$ 에 의해 가공되어 $h_{j,t}$ 의 형태로 은닉층에 전달된다. 은닉층에 전달된 정보 $h_{1,t}$ 에서 $h_{q,t}$ 는 식 (2)와 같이 또 다른 활성화함수 $\Psi(\cdot)$ 에 의해 가공되어 인공 신경망에 의한 예측 값인 o_t 을 도출한다. 식에서 α 와 β 는 전통적인 계량경제학 모형의 모수 추정치에 해당하는 가중치이고, 입력층과 은닉층의 정보는 각각이 하나의 노드(node)를 형성한다.

$$(1) \quad h_{j,t} = \Phi(X_{i,t} : \alpha) \quad (\text{단, } i = 1, \dots, p)$$

$$(2) \quad o_t = \Psi(h_{j,t} : \beta) \quad (\text{단, } j = 1, \dots, q)$$

출력층에서는 도출된 예측 값 o_t 와 실제 값을 비교하여 오차(오류)를 계산한다. 오류 역전파 알고리즘은 출력층에서 계산된 오차를 은닉층과 입력층으로 역전파하여 가중치를 조정한다. 이러한 과정이 끝나면 1회 학습이 완료되고, 학습과정은 사전에 설정한 학습 횟수를 달성하거나 오류가 임계치 이하로 떨어질 때까지 반복된다.

Fig. 1은 오류 역전파 알고리즘을 이용한 다층 퍼셉트론 형태의 인공 신경망의 구조를 나타낸 그림이다. 인공 신경망은 입력 정보를 가공하는 은닉층의 존재로 모형의 해석이 불가하고, 예측력 향상에 많은 시행착오가 따른다는 문제점이 있다. 그러나 전통적인 계량경제학 모형과 달리 자유도(degree of freedom)에 제약을 받지 않고, 다중공선성 문제(multi-collinearity problem)가 발생하지 않아 활용변수의 폭이 넓다. 이에 가용한 정보를 충분히 활용하기 때문에 높은 예측력을 기대할 수 있다(Yun et al., 2016).



Source: Yun et al. (2016)

Note: α and β denote weights.

Fig. 9. Structure of artificial neural network model.

인공 신경망을 이용하여 단수를 예측하기 위해서는 활성화 함수의 형태, 알고리즘, 각 층의 노드 수, 은닉층의 수, 데이터 표준화 방식 등을 결정해야 한다. 본 연구는 활성화 함수

로 로지스틱 함수(logistic function), 알고리즘으로 가중치의 역추적을 허용하는 ‘탄력적 역전과 방식(resilient-backpropagation with weight backtracking)’의 알고리즘을 적용한다.1)

입력층과 출력층의 노드 수는 각각 투입변수와 출력변수의 수이다. 본 연구는 주요 기상변수가 고추의 생육에 미치는 영향을 정리한 RDA(2013)을 참고하여 월평균 기온, 일교차, 누적강수량, 상대습도 및 합계일조시간을 투입변수로 활용하였다. 고추의 경우 5월에서 6월 사이 정식을 한 후 10월경 수확하므로 5월에서 10월까지의 월별 기상변수를 투입변수로 설정하였다. 계량경제학 모형을 활용한 선행연구에서 주로 기온과 강수량을 설명변수로 설정하였는데, 투입변수에 따른 예측력 차이를 비교하기 위해 기온과 강수량만을 투입변수로 이용한 모형과 5개 기상변수(기온, 강수량, 일교차, 상대습도, 일조시간)를 모두 이용한 모형을 구분하여 예측하였다.

은닉층과 은닉층별 노드 수는 시나리오별 예측력을 비교하여 결정해야 한다(Yun et al., 2016). 본 연구는 분석의 단순화를 위해 2개 은닉층 시나리오(1개, 2개)와 노드 수 시나리오(5개, 10개) 등 총 4개 시나리오를 설정하고, 예측력이 가장 우수한 시나리오의 은닉층과 노드 수를 단수 예측 모형에 적용하였다.

인공 신경망 모형을 효율적으로 이용하기 위해서는 데이터 표준화 작업이 필요하다. 본 연구는 Yun 등(2016)에서 상대적으로 효율적이라고 평가한 식 (3)의 표준화 방식을 단수 예측에 적용하였다. 식에서 $\max(x)$ 와 $\min(x)$ 는 각각 x 의 최대값과 최소값이고, 표준화된 값 z 는 0과 1 사이의 값을 가진다.

$$(3) \quad z = \frac{x - \min(x)}{\max(x) - \min(x)}$$

2. 패널모형을 이용한 단수 예측 모형

단수 예측을 위한 패널모형으로 식 (4)의 일원 오차 성분 모형(one-way error component model)을 이용하였다. 설명변수는 인공 신경망의 투입변수로 활용된 5개 기상변수이고, 종속변수는 단수이다. 식에서 $y_{i,t}$ 는 t 시점 i 지역의 단수, $X_{k,i,t}$ 는 t 시점 i 지역의 k 번째 기상변수, u_i 는 관찰되지 않는 시간 불변 오차항(unobserved time-invariant error term), $e_{i,t}$ 는 순수 오차항(idiosyncratic error term)이다.

$$(4) \quad y_{i,t} = \alpha + \sum_{k=1}^K \beta_k X_{k,i,t} + u_i + e_{i,t}$$

1) 탄력적 역전과 알고리즘의 내용에 대한 설명은 Gunther and Fritsch (2010)을 참고할 것.

일원 오차 성분 패널모형은 관찰되지 않는 시간 불변 오차항 u_i 을 확률변수(random variable)로 간주하는 확률효과 모형(random effect model)과 추정해야 할 모수로 간주하는 고정효과 모형(fixed effect model)으로 구분된다. 추정 모형은 논리적 추론이나 통계적 방법인 하우스만 검정(Hausman test)을 이용하여 선택할 수 있다. 만약 하우스만 검정에서 $COV(X_{k,i,t}, u_i) = 0$ 인 경우 고정효과와 확률효과 모형의 추정량이 모두 일치추정량이므로 상대적으로 더욱 효율적인 확률효과 모형이 적합하고, $COV(X_{k,i,t}, u_i) \neq 0$ 인 경우 확률효과 모형의 추정량이 일치추정량이 되지 못하여 고정효과 모형이 적합하다(Min and Choi, 2010). 단수 예측 모형의 설명변수인 기상변수의 경우 각 지역의 지형 등 모형에 반영되지 못한 시간 불변 변수의 영향을 받기 때문에 본 연구는 패널 고정효과 모형을 추정하였다.

패널 고정효과 모형은 within 추정량을 통해 도출할 수 있는데, 오차항 u_i 을 고정효과로 가정할 경우 지역별 더미변수를 두어 상수항을 서로 다르게 추정하면 패널 within 추정량과 동일한 결과를 얻을 수 있다(Min and Choi, 2010). 이에 본 연구는 지역별 더미를 포함한 패널 고정효과 모형을 추정하였다.²⁾

패널모형도 인공 신경망 모형과 유사하게 설명변수에 따라 모형의 예측력이 달라질 수 있다. 이에 본 연구는 선행연구에서 널리 활용되는 기상과 강수량 변수를 설명변수로 활용한 모형과 5개 기상변수를 모두 설명변수로 활용한 모형의 설명력을 비교한다.³⁾ 각 기상변수는 인공 신경망과 동일하게 5월에서 10월까지의 월별 변수를 설명변수로 적용하였다.

3. 적응 기대 모형을 이용한 단수 예측 모형

적응 기대 모형은 사람들의 예상이 과거 경험에 의존한다는 적응 기대가설(adaptive expectation hypothesis)을 바탕으로 과거 단수를 예측에 활용하는 모형이다. 생산량 예측 시 과거 일정 기간 동안의 평균 단수를 예측치로 적용하는 방식이 적응 기대 모형을 이용한 예측이다. 본 연구는 직전연도와 최근 5개년 동안의 평균 단수를 예측 단수로 활용한 적응 기대 모형을 설정하였다(식 (5), 식 (6)). 식에서 Y_{t+1} 은 t+1 시점의 단수, Ω_t 는 t 시점에 활용 가능한 정보 집합, Y_t 에서 Y_{t-k} 는 t시점에서 t-k 시점까지의 단수를 의미한다.

$$(5) E(Y_{t+1}|\Omega_t) = Y_t$$

$$(6) E(Y_{t+1}|\Omega_t) = \frac{1}{K+1} \sum_{k=0}^K Y_{t-k}$$

2) 패널 within 추정은 Min과 Choi (2010, pp. 124-133)을 참고할 것.

3) 전통적 계량경제학 모형에서는 설명변수 설정과 관련하여 누락변수 편의(omitted variable bias)나 불필요한 변수로 인한 효율성 저하 문제(redundant variable problem)가 발생할 수 있다. 그러나 본 연구는 인공 신경망의 예측력을 평가하는데 목적이 있으므로 이러한 계량경제학적 이슈는 고려하지 않았다.

4. 단수 예측 모형별 예측력 비교 방법

모형의 예측력은 롤오버(roll-over) 방식으로 t 시점까지의 자료를 이용하여 단수를 추정하고, 단수 추정에 사용하지 않은 t+1 시점의 지역별 기상변수를 이용하여 t+1 시점의 단수를 예측하는 방법(out-of-sample validation)을 이용하여 분석하였다. 일반적인 롤오버에서는 모형 추정에 이용되는 기간을 고정하여 새로운 표본이 모형 추정 기간의 표본(in-sample)에 포함될 경우 추정 시작 시점을 조정한다. 그러나 본 연구는 Yun 등(2016)에서 적용한 바와 같이 새로운 정보 추가로 예측력이 향상될 수 있는 인공 신경망의 특성을 고려하여 단수 모형 추정 시작 시점을 고정하였다. 예를 들어 1996년에서 2013년까지의 자료로 모형을 추정하여 2014년의 단수를 예측하고, 2015년의 단수는 1996년에서 2014년까지의 자료를 이용하여 추정하였다.

예측 대상은 2014년, 2015년, 2016년의 지역별 단수이다. 단수 예측 추정에 사용된 자료의 기간은 각각 기간 1(2014년 예측) 1996~2013년, 기간 2(2015년 예측) 1996~2014년, 기간 3(2015년 예측) 1996~2015년이다(Fig. 2).

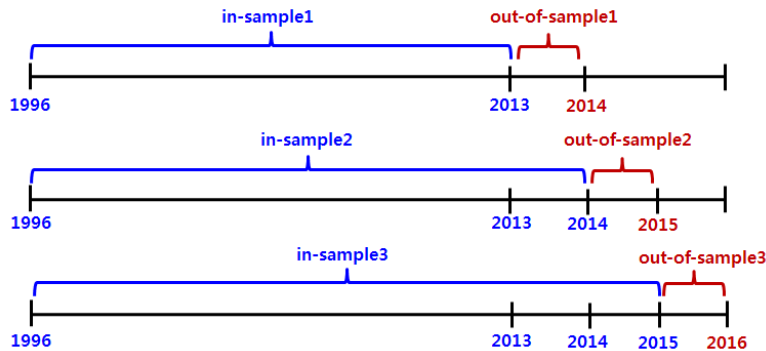


Fig. 10. Roll-over by forecasting period.

모형별 단수 예측력은 측정 단위의 영향을 제거하여 예측오차를 비율로 표시한 RMSPE (root mean square prediction error)로 평가하였다(식 (7)). 식에서 T와 I는 각각 예측 대상 시점과 지역의 수, $Y_{i,t}^s$ 는 t 시점 i 지역의 예측 단수, $Y_{i,t}$ 는 실제 단수를 의미한다.

$$(7) \quad \text{RMSPE} = \sqrt{\frac{1}{T \times I} \sum_{t=1}^T \sum_{i=1}^I \left(\frac{Y_{i,t}^s - Y_{i,t}}{Y_{i,t}} \right)^2}$$

인공 신경망의 예측력은 패널모형, 적응 기대 모형과 비교한다. Yun 등(2016)은 인공 신경망에 포함되는 정보의 증가가 예측력을 증가시키는지를 평가한 바 있다. 본 연구에서도 인공 신경망과 패널모형의 예측력을 평가할 때 기온과 강수량 정보만을 반영한 경우와 5개

기상변수를 모두 반영한 경우의 예측력을 비교한다.

Ⅲ. 자 료

본 연구는 1996년에서 2016년까지의 8개 주산지 광역시도(경기, 강원, 충북, 충남, 전북, 전남, 경북, 경남)의 연도별 자료를 이용하여 단수 예측 모형을 추정하였다.⁴⁾ 단수는 통계청에서 제공하는 주산지 광역시도의 단수(kg/10 a)를 이용하였다. 주산지 광역시도의 재배 면적과 생산량은 최근 5개년(2012~2016) 기준으로 각각 전체 대비 96.8%와 97.1%로 높다 (Table 1).

Table 1. Acreage and Production of chilli pepper (1996~2016)

	Acreage (1,000ha)			Production (1,000ton)			Proportion of chief producing district	
	Chief producing district	Others district	Total	Chief producing district	Other district	Total	Acreage	Production
1996	89.3	1.5	90.8	215.5	3.0	218.5	98.4%	98.6%
2000	72.5	1.9	74.5	189.4	4.3	193.8	97.4%	97.8%
2005	59.6	1.7	61.3	157.4	4.0	161.4	97.2%	97.5%
2010	43.4	1.1	44.6	93.1	2.3	95.4	97.5%	97.6%
2011	41.4	1.2	42.6	75.3	1.9	77.1	97.2%	97.6%
2012	44.2	1.3	45.5	101.1	3.1	104.1	97.2%	97.0%
2013	43.8	1.5	45.4	114.4	3.4	117.8	96.6%	97.1%
2014	34.9	1.2	36.1	82.3	2.7	85.1	96.7%	96.8%
2015	33.3	1.2	34.5	95.0	2.7	97.7	96.6%	97.2%
2016	31.2	1.0	32.2	83.2	2.3	85.5	97.0%	97.3%
Mean (2012-2016)	37.5	1.2	38.7	95.2	2.9	98.0	96.8%	97.1%

Source : National Statistical Office

Note : Chief producing districts of chilli pepper are 8 provinces (Gangwon, Gyeonggi, Gyeongbuk, Gyeongnam, Jeonbuk, Jeonnam, Chungbuk, Chungnam).

4) 단수 예측을 위한 인공 신경망의 투입변수와 패널모형의 설명변수인 기상변수는 동일 광역시도 내에서도 지역별 편차가 발생하기 때문에 최소한 주산지 시군 단위의 분석이 적합하다. 그러나 고추 주산지 조사가 2016년 이후 중지되어 향후 자료 이용이 불가하고, 2015년 이전의 경우에도 조사 자료가 누락된 주산지 시군이 많아 광역시도 기준의 자료를 이용하여 분석하였다.

주산지 광역시도의 연도별 단수는 Table 2에 정리하였다. 지역별 재배면적 비중을 가중치로 사용한 연도별 평균 단수는 1996년 241 kg/10 a에서 2016년 266 kg/10 a로 증가 추세를 보인다. 가중 평균과 표준편차를 이용하여 산출한 변이계수는 1996년 0.146에서 2016년 0.080으로 감소하는 추세를 보이는데 이는 지역별 단수의 편차가 줄어든다는 의미이다. 최근 5개년(2012~2016)을 기준으로 경북 지역의 단수가 274 kg/10 a로 가장 많고, 그 다음이 충남, 전남, 전북 등의 순이며, 단수가 가장 작은 지역은 경기(219 kg/10 a)이다.

Table 2. Yield of chilli pepper in chief producing provinces (1996~2016)

	Gyeong-gi	Gang-won	Chung-buk	Chung-nam	Jeon-buk	Jeon-nam	Gyeong-buk	Gyeong-nam	mean	SD	CV
1996	168	175	274	203	275	258	249	218	241	35	0.146
2000	223	234	306	254	265	261	270	182	261	28	0.109
2005	214	255	255	221	264	281	292	274	264	26	0.098
2010	168	212	212	209	198	238	228	192	214	20	0.091
2011	116	126	106	124	203	225	224	216	182	50	0.274
2012	199	206	213	205	224	229	260	226	229	21	0.093
2013	192	195	250	256	239	277	306	266	261	37	0.140
2014	225	242	226	247	234	242	243	194	236	13	0.055
2015	257	252	263	316	310	283	289	279	285	20	0.069
2016	223	224	264	298	284	264	271	269	266	21	0.080
Mean (2012~2016)	219	224	243	264	258	259	274	247	255	17	0.068

Source : National Statistical Office

Note 1 : SD and CV denote standard deviation and coefficient of variation respectively.

Note 2 : Yearly mean, standard deviation are weighted values: weight = province's acreage / total acreage.

기상변수는 기상청 기상자료개방포털에서 제공하는 관측소별 자료를 이용하였다. 주산지 광역시도 내 시군별로 재배면적이 다르므로 이를 반영한 가중지표를 사용할 필요가 있으나, 자료 제약으로 각 광역시도 내 관측소별 가중치를 동일하게 적용한 산술평균을 이용하였다.⁵⁾ 기상변수의 월별 기초통계량은 Table 3과 같다.

5) 시군별 재배면적 확보가 가능할 경우 GIS(공간정보시스템) 등을 활용한 가중평균 값을 이용할 경우 재배지역의 기상정보가 더욱 정확히 반영될 수 있으리라 기대된다. 이는 향후 과제로 남겨둔다.

Table 3. Basic statistics of weather variables for chilli pepper yield forecast (1996~2016)

		May	June	July	Aug.	Sept.	Oct.
Daily average temperature (°C)	Mean	17.3	21.4	24.5	25.0	20.5	14.2
	Max	19.0	23.5	26.7	27.5	23.1	17.6
	Min	15.1	18.9	21.2	22.1	17.0	10.6
	SD	0.8	0.8	1.1	1.2	1.2	1.5
	CV	0.046	0.038	0.045	0.046	0.058	0.106
Monthly cumulated precipitation (mm)	Mean	103.2	159.2	315.2	277.1	160.5	58.0
	Max	289.9	397.8	1,001.6	817.4	523.0	212.6
	Min	11.6	29.9	78.1	39.3	11.4	1.4
	SD	53.6	87.0	152.4	157.5	119.0	40.7
	CV	0.519	0.546	0.483	0.569	0.742	0.702
Daily temperature range (°C)	Mean	12.1	9.9	7.7	8.3	9.9	12.0
	Max	14.9	12.4	9.8	10.8	13.4	15.3
	Min	8.6	6.2	5.3	5.8	6.3	7.3
	SD	1.3	1.3	0.9	1.1	1.4	1.4
	CV	0.104	0.128	0.122	0.130	0.142	0.114
Daily average humidity (%)	Mean	65.0	72.1	79.6	78.5	75.8	70.8
	Max	77.4	84.6	88.6	86.3	84.7	80.1
	Min	53.5	61.3	70.8	67.3	63.6	60.4
	SD	4.4	4.4	3.5	3.4	4.0	3.9
	CV	0.067	0.060	0.044	0.043	0.053	0.055
Daily average sunshine amount (h)	Mean	7.4	6.1	4.7	5.4	5.7	6.3
	Max	10.4	9.5	9.6	10.4	10.1	8.2
	Min	5.1	2.9	1.9	2.9	2.5	3.5
	SD	1.1	1.2	1.2	1.5	1.2	1.0
	CV	0.146	0.195	0.254	0.267	0.208	0.152

Source : Open portal for weather variables, Korea Meteorological Administration.

Note : SD and CV denote standard deviation and coefficient of variation respectively.

IV. 예측력 분석 결과

인공 신경망을 이용하여 고추 단수를 예측한 결과 기온과 강수량을 투입변수로 할 경우 은닉층 1개, 노드 5개 시나리오(S1-1-1), 5개 기상변수를 모두 투입변수로 할 경우 은닉층 2개, 노드 각 10개 시나리오(S2-2-2)의 예측력이 가장 우수하였다. 투입변수에 따라 예측력이 가장 우수한 은닉층과 노드 수가 다르다는 결과는 Yun 등(2016)이 지적한 바와 같이 인공 신경망 활용 시 은닉층과 노드 수 설정이 중요함을 의미한다.

Table 4. Forecast errors of artificial neural network by scenario

Input variables		Temperature, precipitation				Temperature, precipitation, temperature range, humidity, sunshine amount			
		S1-1-1	S1-1-2	S1-2-1	S1-2-2	S2-1-1	S2-1-2	S2-2-1	S2-2-2
Number of nodes by hidden layer	H1	5	10	5	10	5	10	5	10
	H2	-	-	5	10	-	-	5	10
RMSPE		0.2380	0.3088	0.2674	0.2580	0.4073	0.3141	0.2594	0.1384

Note 1 : Critical value = 0.01, hidden layer activation function = logistic, output layer activation function = linear, error function = SSE, algorithm = resilient back- propagation with backtracking, number of repetition for the training process = 1000

Note 2 : "H" in H1 and H2 denotes hidden layer.

Note 3 : "-" denotes no hidden layer.

고추 단수 모형별 예측력 비교 결과 5개 기상변수를 모두 이용한 인공 신경망 모형(M3-2)의 예측오차가 13.84%로 가장 우수하였다(Table 5).⁶⁾ 그 다음으로 우수한 모형은 적응적 기대 모형이고, 패널모형은 3개 모형 중 예측력이 가장 낮게 평가되었다. 이는 적절한 투입변수를 설정할 경우 인공 신경망을 이용함으로써 단수 예측력을 향상시킬 수 있음을 의미한다.

6) 본 연구가 인공 신경망과 패널모형 등 다른 모형 사이의 예측력을 비교하기 위한 연구로 개별 기상변수의 통계적 유의성이나 부호가 큰 의미를 가지지 않아 패널모형의 추정 결과는 생략하였다. 추정 결과를 저자에게 요청할 경우 제공받을 수 있다.

Table 5. Forecast errors of yield forecast models for chilli pepper

Model	Adaptive expectation model		Panel model		Artificial neural network model ¹⁾	
	M1-1	M1-2	M2-1	M2-2	M3-1	M3-2
Input or explanatory variables	Yieldt-1	5-year average yield	Temperature, precipitation	Temperature, precipitation, temperature range, humidity, sunshine amount	Temperature, precipitation	Temperature, precipitation, temperature range, humidity, sunshine amount
RMSPE	0.1637	0.1810	0.1960	0.1947	0.2380	0.1384

Note : Number of hidden layer = 1 (M3-1) and 2 (M3-2), number of nodes by hidden layer = 5 (M3-1) and 10 (M3-2), critical value = 0.01, hidden layer activation function=logistic, output layer activation function = linear, error function = SSE, algorithm = resilient back-propagation with backtracking, number of repetition for the training process = 1000

V. 요약 및 결론

본 연구는 인공 신경망을 이용한 고추 단수 예측 모형을 제시하고, 적응 기대 모형 및 패널모형과 예측력을 비교하였다. 투입정보가 늘어날수록 예측력 향상이 기대되는 인공 신경망의 특성을 고려하여 기온과 강수량만을 투입변수로 설정한 모형과 기온, 강수량, 일교차, 상대습도, 일조시간 등 5개 기상변수를 모두 반영한 모형을 설정하여 예측력을 비교하였다.

모형별 예측력 비교 결과 5개 기상변수를 모두 반영한 인공 신경망의 예측력이 패널모형이나 적응 기대 모형보다 우수하였다. 이는 인공 신경망이 채소 단수 예측력 향상에 기여할 수 있음을 의미한다. 본 연구는 분석의 단순화를 위해 은닉층과 노드 수에 대한 시나리오와 기간을 한정하고, 투입변수도 5개 기상변수로 제한하였으나 향후 다양한 시나리오와 투입변수를 적용할 경우 예측력 향상의 여지가 있다. 최근 기후변화로 인한 기상이변 증가가 노지 채소 단수 변동에 미치는 영향이 큰 상황을 고려할 때 다양한 기상변수를 이용한 (가칭) 기상이변지수 등을 활용할 경우 예측력 향상에 기여할 수 있으리라 기대된다.

본 연구는 자료의 제약으로 기상변수만을 투입변수로 반영하였는데, 단수에 영향을 미치는 토양의 물리적, 화학적 특성이나 병해충 정보 등을 반영할 경우 예측력 향상이 가능하리라 본다. 이를 위해 관련 정보의 구축과 예측 모형 구축을 위한 정량화 작업이 요망된다. 다른 한편으로 본 연구는 예측 시점의 기상변수의 정보를 모두 알고 있다는 가정 하에 예측을 실시하였으나, 실제로 예측을 할 경우에는 투입변수인 기상변수의 예측도 중요하다.

기상변수의 예측 값으로 과거 일정 기간 동안의 평균이나 평년값을 적용할 수 있고, 별도의 예측 모형을 설정할 수 있다. 기상변수의 예측은 본 연구의 범위를 벗어나므로 이는 향후 과제로 남겨둔다.

[Submitted, July. 31, 2017 ; Revised, August. 14, 2017 ; Accepted, August. 21, 2017]

References

1. Gunther, F. and S. Fritsch. 2010. Neuralnet: Training of Neural Networks. *The R Journal*, 2(1): 30-38.
2. Guo, W. W. and H. Xue. 2014. Crop Yield Forecasting Using Artificial Neural Networks: A Comparison between Spatial and Temporal Models, *Mathematical Problems in Engineering*, Volume 2014, Article ID 8578865.
3. Han, S. H., B. H. Lee, M. S. Park, J. H. Seung, H. S. Yang, and S. C. Shin. 2011. A Study on Building Crop Yield Forecasting Model Considering Meteorological Elements. P152. Korea Rural Economic Institute.
4. Hu, M. J. C. 1964. Application of the Adaline System to Weather Forecasting. Master Thesis. Technical Report 6775-1. Stanford Electronic Laboratories. Stanford, CA. June.
5. Kim, B. S., M. S. Park, J. H. Cho, and T. K. Kim. 2010. A Demand and Supply Model of Agricultural and Livestock Products for Midterm Outlook. M103. Korea Rural Economic Institute.
6. Lamba, V. and V. S. Dhaka. 2014. Wheat Yield Prediction Using Artificial Neural Network and Crop Prediction Techniques. *International Journal for Research in Applied Science and Engineering Technology*. 2(4): 330-341.
7. Laxmi, R. R. and A. Kumar. 2011. Weather Based Forecasting Model for Crops Yield Using Neural Network Approach. *Statistics and Application*. 9(1&2): 55-69.
8. Lee, C. S. and S. B. Yang. 2017. Development of Yield Forecast Models for Autumn Chinese Cabbage and Radish Using Crop Growth and Development Information. *Korean Journal of Organic Agriculture*. 25(2): 257-271.
9. Lee, J. W. 1996. A Study on the Production Determinants of Chinese Cabbage and Radish. R346. Korea Rural Economic Institute.
10. Lee, M. S. and Y. C. Choe. 2009. Forecasting Sow's Productivity Using the Machine

- Learning Models. *Journal of Agricultural Extension & Community Development*. 16(4): 939-965.
11. Lee, Y. S., H. K. Jung and S. B. Sim. 2005. A Study on Determinants of Seasonal Supply and Price of Produce in Korea: With Special Emphasis on Weather. R494. Korea Rural Economic Institute.
 12. Min, I. S. and P. S. Choi. 2010. STATA Panel Data Analysis. The Korean Association of STATA.
 13. Panda S. S., D. P. Ames, and S. Panigrahi. 2011. Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques. *Remote Sensing*, 2(3): 673-696.
 14. Park, J. Y. and Y. G. Park. 2013. The Development of Chinese Cabbage and Radish Forecast Models. M125. Korea Rural Economic Institute.
 15. Rumelhart, D. E., G. E. Hinton, and R. J. Williams. 1986. Learning Representations by Backpropagating Errors. *Nature*. 323(6188): 533-536.
 16. Rural Development Administration. 2013. Chilli Pepper. *Agricultural Technology Guide*, 126. Rural Development Administration.
 17. Sung, J. H. and D. H. Lee. 2005. Rice Yield Prediction Based on the Soil Chemical Properties Using Neural Network Model. *Journal of Biosystems Engineering*. 30(6): 360-365.
 18. Yun, S. J., C. S. Lee, and S. R. Yang. 2016. Development of Prices Forecasting Model for International Grains Using Artificial Neural Network. *Korean Journal of Agricultural Economics*, 57(2): 83-101.