

인공신경망 모형을 이용하여 토양 화학성으로 벼 수확량 예측

성제훈 이동훈

Rice Yield Prediction Based on the Soil Chemical Properties Using Neural Network Model

J. H. Sung D. H. Lee

Abstract

Precision agriculture attempts to improve cropping efficiency by variable application of crop treatments such as fertilizers and pesticides, within field on a point-by-point basis. Therefore, a more complete understanding of the relationships between yield and soil properties is of critical importance in precision agriculture. In this study, the functional relationships between measured soil properties and rice yield were investigated. A supervised back-propagation neural network model was employed to relate soil chemical properties and rice yields on a point-by point basis, within individual site-years. As a results, a positive correlation was found between practical yields and predicted yields in 1999, 2000, 2001, and 2002 are 0.916, 0.879, 0.800 and 0.789, respectively. The results showed that significant overfitting for yields with only the soil chemical properties occurred so that more of environmental factors, such as climatological data, variety, cultivation method etc., would be required to predict the yield more accurately.

Keywords : Soil chemical properties, Yield prediction, Neural network

1. 서론

농지는 포장별 또는 포장내 위치별로 토양의 이화학적 성질, 비옥도, 균평 정도, 물의 흐름, 병해충 저항성 등이 달라 농산물은 질과 양이 다를 수 있다. 그러나 지금까지의 관행 농업기술은 포장 각 위치별 이러한 변수의 차이를 고려하지 못하고 있다. 포장 전체를 하나로 보고 획일적으로 일정한 양의 관개수나 화학제를 공급하여 환경오염을 일으킬 수도 있다. 이러한 포장내 변이를 고려하여 농자재를 투입하는 것이 정밀농업의 목적 중 하나이다. 정밀농업의 기본 개념은 포장내 위치별 작물 생산능력과 작물 생육상태에 차이가 있음을 인정하고, 각 위치에 대하여 적당한 시기에 적절한 처리를 수행함으로써 농자재의 투입을 줄이고 농자재의 과다투입으로 인한 환경오염을 줄이면서 농지를 적정하게 활용하는 농법이

다(Sudduth 등, 1996).

정밀농업을 위해서는 포장 각 위치의 농산물 생산 잠재력에 대한 정보를 취득해야 한다. 취득한 생산 잠재력 정보 중 토양 화학성을 분석해서 포장내 위치별 시비처방전을 만들게 된다. 시비처방은 양분의 균형공급과 양분축적경감에 의한 환경오염을 방지하기 위한 기준으로서 작물별 표준시비량과 토양검정에 의한 시비기준량을 포함하고 있다. 표준시비량은 질소, 인산, 칼륨으로 구분하여 제시한다. 일반적으로 시비처방전에서 질소 시비량은 유기물과 규산 함량에 따라 결정되고, 인산 시비량은 인산 함량에 따라 결정되며, 칼륨량은 양이온치환능력과 칼륨 포화도에 의해 결정된다.

포장이 가지고 있는 작물생산능력을 이용하여 수확량을 예측하고자하는 연구는 여러 차례 수행된 적이 있다. Sudduth 등(1996)은 다양한 선형 분석 방법으로 데이터를 해석해서

The article was submitted for publication in October 2005, reviewed in December 2005, and approved for publication by the editorial board of KSAM in December 2005. The authors are Je Hoon Sung, KSAM member, Research Engineer, National Institute of Agricultural Engineering, RDA, Suwon, 441-707, Korea, Dong Hun Lee, Assistant Researcher, National Institute of Agricultural Engineering, RDA, Suwon, 441-707, Korea. The corresponding author is D. H. Lee, Assistant Researcher, National Institute of Agricultural Engineering, RDA, Suwon, 441-707, Korea; Fax : +82-31-290-1900; E-mail : <jhsung@rda.go.kr>

토양내 화학성, 수확량, 기상 등이 서로 연관되어 있음을 보고했고, 선형 분석만으로도 수확량을 예측할 수 있음을 보였다.

신경망이나 퍼지 같은 비선형 해석기술을 이용해서 수확량을 예측한 연구도 다수 있다. Liu 등(1999, 2001)은 역전파 알고리즘을 사용해서 수년 동안의 토양, 기상, 경영요인을 이용하여 옥수수 수확량을 예측했다. Sudduth 등(1996)과 Drummond 등(2003)은 역전파 알고리즘을 이용한 신경망을 포함한 다양한 분석기법을 이용하여 선형 분석법보다 더 정밀하게 수확량을 예측하는 방법을 제시했다.

우리나라에서는 Hong 등(1976)이 비옥도 구성인자를 설명변수로하고 수확량을 목적변수로 하여 다중회귀분석을 실시해서 결정계수가 0.43임을 보고했다. 동시에 비옥도 인자간의 상호작용을 고려하여 직선회귀모델을 이차방정식으로 변형시켜 생산력을 추정한 결과 결정계수가 0.59임을 보고했다. 또한 쌀의 생산력을 지배하는 인자는 비옥도뿐만 아니라 환경인자의 영향도 있으므로 비옥도군, 수분공급의 양호정도, 토양배수의 양호정도, 기후도, 표토 점착성, 토양 강도, 토성 및 하층토의 구조발달과정 등을 설명변수로 하면 결정계수를 0.9까지 올릴 수 있다고 보고했다. Kim 등(2001)은 벼 수량의 기상반응을 종합적으로 검토하여 벼 수량예측모델을 구축하고자 기상에 대한 수량반응에 대해 최대경계선분석을 실시하여 수량예측모형을 설정하였고 결정계수가 0.61 이상임을 보고했다. 그러나 토양 화학성을 이용해서 수확량을 예측한 보고는 없다.

본 연구는 시비처방전을 만드는 데 필요한 토양 화학성 6항목과 작물 수확량간의 관계를 분석하고, 인공신경망을 이용하여 시비처방전을 만드는 데 필요한 포장내 위치별 토양 화학성을 이용해서 그해 위치별 수확량 예측성능을 평가하는데 그 목적이 있다.

2. 재료 및 방법

가. 공시 포장 및 시료

시험포장은 100 m×30 m 직사각형 형태의 논이다. 관계수는 그림 1과 같이 오른쪽 아래에서 왼쪽위로 흐른다. 1999년부터 2002년까지 매년 4월 중순에 일품벼를 파종하여 5월 중순에 이앙하였으며, 이앙 직전에 토양조사 결과를 이용하여 환경보전형 시비기준(농협중앙회, 2000)에 따라 기비했다. 질소, 인산, 칼륨 비료는 각각 요소, 용과린, 염화기리를 사용하였다. 질소의 경우 기비, 분얼비를 각각 70%, 30%로 나누어 시비하였으며, 인산은 전량 기비로 하였다. 칼륨은 기비와 분얼비를 각각 70%, 30%로 나누어 시비하였다. 인산과 칼륨

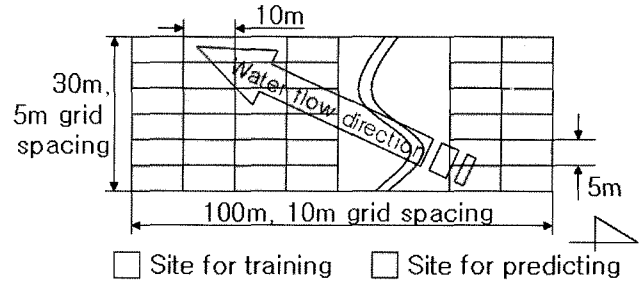


Fig. 1 The experimental paddy field for soil sampling, yield monitoring, and site for neural network training and predicting.

시비추천량은 최저 3 kg/10a였다.

포장내 위치별 토양 화학성과 수확량 측정을 위해 그림 1과 같이 5 m×10 m 단위로 60개의 구획으로 구분하였다. 각 격자의 중심부분에서 토양 시료를 채취하였다. 토양 시료는 전년도 수확후에 각 구획별 중앙부분 표토에서 5 cm~15 cm 깊이로 3점을 채취하여 혼합한 후에 변량시비 추천량 계산을 위한 토양 화학성 분석을 실시했다. 토양 분석은 농촌진흥청 농업과학기술원 토양식물체 분석법(농업과학기술원, 2000)에 따라 pH와 EC는 초자전극법, 유기물은 Walkley-Black법, 치환성양이온은 1M-NH₄OAc(pH 7.0) 용액으로 침출하여 유도결합 플라즈마 분광 광도계로 측정하였으며, 질소는 Kjeldahl법, 인산은 몰리브덴청법, 칼륨은 원자흡광법, 규산은 1M CH₃COONa 침출법으로 각각 분석하였다. 수확량은 3조식 자탈형 콤팩인(model KHD515, Gukje, Korea)으로 측정된 후 수분 14%로 보정했다.

나. 신경망 학습

역전파 신경망은 다층 퍼셉트론 구조를 가지고, 지도학습 규칙을 따르도록 C 프로그램(ver. 6.0, Microsoft)을 이용하여 만들었다. 학습에 사용한 신경망 구조는 그림 2와 같다. 입력노드는 시비처방전을 만드는 데 필요한 토양화학성 6항목이고 출력노드는 수확량이다. 30 m×100 m 시험포장을 5 m×10 m 기본 격자 60개로 구분한 후, 그림 1에서처럼 5 m×10 m

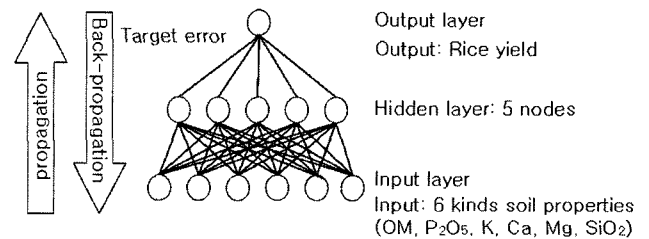


Fig. 2 Neural network structure consisting of input layer with six soil properties and output layer with rice yield.

기본 격자를 하나씩 건너면서 토양 화학성을 학습에 사용했고(30개 패턴), 그 사이 격자에서 측정된 값(30개 패턴)을 학습에서 예측한 값과 비교했다. 학습 형태는 학습 데이터 이외의 시료를 예측에 포함하지 않도록 일괄학습(batch training) 형태를 사용했다. 연결 가중치의 변화율을 결정하는 학습률은 0.1, 오차진동을 적게 하여 수렴속도를 높이는 모멘텀은 0.9로 하였다. 학습횟수는 100,000이다. 기술통계와 상관 분석은 SAS(ver. 8.01, SAS Institute Inc., NC)를 이용하였으며, 지도 작성은 Surfer(ver. 8.03. Golden software)를 이용했다.

최종 예러는 상대평균오차(MSE, relative mean square error)를 썼으며 사용식은 식 (1)과 같다.

$$\frac{1}{N} \sum_{k=1}^N (E_k - t_k)^2 \quad (1)$$

where

N = count of output layer

E_k = predicted value of k node

t_k = reference of k node

3. 결과 및 고찰

가. 기술통계

수확량 예측을 위한 신경망 학습과 검증에 사용한 포장내 위치별 토양 화학성 및 수확량의 기술통계를 표 1에 나타냈다. SiO₂는 변이계수가 17.9%~29.6%로 다른 화학성에 비해 포장내 위치별로 변이가 큰 것으로 나타났다. 상대적으로 유기물은 변이계수가 4.3%~10.3%로 다른 화학성에 비해 포

Table 1 Descriptive statistics of soil chemical properties and rice yield for training and predicting of test field

year		OM g kg ⁻¹	Av. P ₂ O ₅ mg kg ⁻¹	Exch. cations			Av. SiO ₂ mg kg ⁻¹	yield kg/10a
				K	Ca	Mg		
				cmol _c kg ⁻¹				
1999 (training)	avg	2.36	122.3	0.68	4.79	1.00	93.70	594.96
	std	0.125	14.441	0.144	0.700	0.164	24.181	78.581
	CV (%)	5.3	11.8	6.5	14.6	16.4	25.8	13.2
1999 (predicting)	avg	2.32	117.22	0.68	4.82	1.01	96.27	588.23
	std	0.168	14.60	0.039	0.641	0.161	24.489	68.096
	CV (%)	7.3	12.5	5.8	13.3	16.0	25.4	11.6
2000 (training)	avg	2.31	79.98	0.63	4.44	1.00	107.08	471.33
	std	0.143	10.990	0.052	0.385	0.131	31.672	73.141
	CV (%)	6.2	13.7	8.3	8.7	13.2	29.6	15.5
2000 (predicting)	avg	2.29	80.28	0.62	4.50	1.02	108.08	460.15
	std	0.111	9.246	0.044	0.428	0.107	30.206	95.175
	CV (%)	4.9	11.5	7.1	9.5	10.6	27.9	20.7
2001 (training)	avg	2.43	91.87	0.75	5.46	1.16	80.73	570.72
	std	0.104	15.305	0.367	0.844	0.146	22.478	80.567
	CV (%)	4.3	16.7	48.6	15.5	12.6	27.8	14.1
2001 (predicting)	avg	2.44	92.31	0.69	5.42	1.16	78.28	558.22
	std	0.119	15.817	0.663	0.824	0.162	22.774	84.829
	CV (%)	4.9	17.1	9.2	15.2	14.0	29.1	15.2
2002 (training)	avg	2.67	71.93	0.68	5.06	1.36	78.45	703.01
	std	0.234	11.752	0.074	0.644	0.144	16.787	43.004
	CV (%)	8.8	16.3	11.0	12.8	10.6	21.4	6.1
2002 (predicting)	avg	2.77	72.66	0.66	4.832	1.31	80.65	690.03
	std	0.284	10.527	0.077	0.564	0.140	14.443	54.087
	CV (%)	10.3	14.5	11.7	11.7	11.0	17.9	7.8

OM=organic matter, Av.=available, Exch.=exchangeable

장내 위치별로 변이가 크지 않은 것으로 나타났다. P₂O₅, K, Ca, Mg와 수확량의 변이계수는 각각 11.5%~17.1%, 4.9%~11.7%, 8.6%~15.5%, 10.6%~16.4%, 6.1%~20.7%였다.

나. 상관계수

수확량 예측을 위한 신경망 학습과 검증에 사용한 포장내 위치별 토양 화학성 및 수확량의 상관관계를 분석했다(표 2). 1999년의 경우 Ca와 Mg간의 피어슨 상관계수는 0.973이었으며, SiO₂와는 Ca와 Mg가 0.7 이상의 피어슨 상관계수를 가지고 있었다. 그러나 수확량과 의미 있는 상관을 가진 토양 화학성은 없었다. 2000년의 경우 Ca와 Mg간의 피어슨 상관계수는 0.732이었으며, P₂O₅가 수확량과 0.5 이상의 부의 상관을 가지고 있었다. 2001년의 경우 Ca와 Mg간의 피어슨 상관계수는 0.880이었으며, SiO₂와는 Ca와 Mg가 0.88과 0.69 이상의 피어슨 상관계수를 가지고 있었다. 1999년과 마찬가지로 수확량과 의미 있는 상관을 가진 토양 화학성은 없었다.

2002년도 Ca와 Mg간의 피어슨 상관계수는 0.982이었으며, SiO₂와는 Ca와 Mg가 0.78 이상의 피어슨 상관계수를 가지고 있었다. 그러나 수확량과 의미 있는 상관을 가진 토양 화학성은 없었다. 따라서 토양 화학성 한두 개만 가지고 수확량을 예측할 수 없음을 알 수 있다.

수확량 예측을 위한 신경망 학습과 검증에 사용한 포장내 위치별 토양 화학성의 연도별 상관관계를 분석했다. 표 3에서 보는 바와 같이 유기물과 K는 연도별로 의미 있는 상관이 없는 것으로 나타났다. 수확량도 연도별로 의미 있는 상관이 없는 것으로 나타났다. P₂O₅는 연도별로 의미 있는 상관은 있으나 피어슨 상관계수가 0.5 이하로 나타났다. 그러나 Mg와 Ca는 연도별로 의미 있는 상관이 높아 연도별로 값이 크게 변하지 않음을 알 수 있다.

따라서 토양 화학성 중 Mg와 Ca는 연도별로 크게 변하지 않으나 다른 화학성은 해마다 변하는 것을 알 수 있고, 수확량도 매년 변함을 알 수 있다.

Table 2 The Pearson correlation coefficients between the soil chemical properties and rice yield

index		P ₂ O ₅	K	Ca	Mg	SiO ₂	Yield
1999	OM	0.004	0.070	-0.086	-0.107	-0.074	-0.282*
	P ₂ O ₅	1.000	0.128	-0.564***	-0.549***	-0.502***	-0.112
	K		1.000	-0.078	-0.098	-0.064	0.058
	Ca			1.000	0.973***	0.741***	0.148
	Mg				1.000	0.703***	0.193
	SiO ₂					1.000	0.002
2000	OM	-0.077	0.051	0.405**	0.302*	-0.065	0.045
	P ₂ O ₅	1.000	-0.178	-0.221	-0.236	0.232	-0.523***
	K		1.000	-0.217	-0.109	0.039	0.002
	Ca			1.000	0.732***	0.232	0.172
	Mg				1.000	-0.032	0.296*
	SiO ₂					1.000	-0.318
2001	OM	0.049	-0.033	0.227	-0.052	0.315*	0.097
	P ₂ O ₅	1.000	0.021	-0.754***	-0.698***	-0.655***	-0.135
	K		1.000	-0.005	0.004	-0.048	-0.096
	Ca			1.000	0.881***	0.882***	0.081
	Mg				1.000	0.690***	-0.068
	SiO ₂					1.000	0.103
2002	OM	0.281*	-0.110	-0.214	-0.248	-0.137	-0.052
	P ₂ O ₅	1.000	0.097	-0.409**	-0.438**	-0.298*	0.071
	K		1.000	0.223	0.231	0.212	0.081
	Ca			1.000	0.983***	0.794***	0.104
	Mg				1.000	0.788***	0.114
	SiO ₂					1.000	0.019

*, **, *** significant at 0.05, 0.01 and 0.001 probabilities, respectively

Table 3 The Pearson correlation coefficients of soil chemical properties and rice yields in three years

index	year	year		
		2000	2001	2002
OM	1999	0.097	-0.029	-0.197
	2000		0.308*	0.063
	2001			0.095
P ₂ O ₅	1999	0.077	-0.048	0.112
	2000		0.459**	0.295*
	2001			0.448**
K	1999	-0.047	-0.064	-0.048
	2000		0.038	-0.217
	2001			0.31
Ca	1999	0.361**	0.339**	0.563***
	2000		0.896***	0.810***
	2001			0.809***
Mg	1999	0.488***	0.474***	0.592***
	2000		0.737***	0.749***
	2001			0.828***
SiO ₂	1999	0.071	0.241	0.365**
	2000		0.178	-0.180
	2001			0.687***
Yield	1999	0.151	0.154	0.295*
	2000		-0.036	-0.068
	2001			-0.087

*, **, *** significant at 0.05, 0.01 and 0.001 probabilities, respectively

다. 신경망 학습

포장내 위치별 토양 화학성으로 수확량을 예측하기 위해 역전파 신경망 알고리즘을 학습시켰다. 학습 입력은 매년 시비처방전을 만드는데 필요한 토양 화학성 6항목, 30위치이고 출력은 위치별 그해 수확량이다. 식 (1)을 써서 구한 목표값과 학습에 의한 모의값 사이의 최종 오차는 표 4와 같다.

라. 수확량 예측

역전파 신경망 알고리즘으로 5 m×10 m 기본 격자를 하나씩 건너면서 토양 화학성을 학습에 사용했고, 학습결과를 이

Table 4 Training iteration and final error of neural network

year	training iteration	final error
1999	100,000	0.075
2000	100,000	0.117
2001	100,000	0.221
2002	100,000	0.214

training rate 0.1, momentum 0.9

용해서 학습에 사용한 위치의 수확량을 예측했다. 신경망으로 예측한 수확량과 그 위치의 실제 수확량간의 상관관계를 분석한 결과 피어슨 상관계수가 1999년에는 0.916, 2000년에는 0.879, 2001년에는 0.800, 2002년에는 0.789로 나타났다.

학습된 역전파 신경망의 적합성을 검증하기 위해 학습에 사용하지 않은 위치(학습에 사용한 격자를 하나씩 건너뛰면서 배치)의 수확량을 이용했다. 학습위치와 예측위치가 달라 결과를 직접 비교하기는 곤란하지만, 그림 3에서처럼 학습에 사용한 수확량, 동일 위치를 예측한 수확량, 바로 옆 위치를 예측한 수확량이 서로 비슷함을 알 수 있다.

그림 3의 위 왼쪽은 1999년도 토양 화학성을 이용해서 그해 수확량을 예측한 수확량 지도다. 왼쪽은 신경망을 학습하는 데 사용한 데이터를 이용해서 그린 수확량지도고, 중앙은 학습된 신경망을 이용해서 그 위치의 수확량을 예측한 것이며, 맨 오른쪽은 학습된 신경망을 이용해서 5 m×10 m 떨어진 곳의 수확량을 예측한 결과다. 그림 3의 위 오른쪽, 아래 왼쪽, 아래 오른쪽은 각각 2000년, 2001년, 2002년 수확량 지도이다.

작물 생육과 수확량은 토양 화학성 외에도 작물의 품종, 재배법, 기상 등 여러 요인에 따라 변하므로, 토양 화학성 외에 기상, 재배법, 품종 등의 환경 인자를 추가해서 수확량을 예측한다면 더 정밀한 예측이 가능할 것으로 판단된다. 앞으로의 연구를 통해서 토양 화학성으로 수확량 예측 뿐만아니라 수확물 속에 있는 아밀로스, 단백질 등의 영양분 함량도 예측할 수 있을 것이다.

4. 요약 및 결론

시비처방전을 만드는데 필요한 토양 화학성 6항목과 작물 수확량간의 관계를 분석한 후, 인공신경망을 이용하여 시비처방전을 만드는데 필요한 포장내 위치별 토양 화학성을 이용해서 그해 위치별 수확량 예측성능을 평가했다.

기술통계를 분석한 결과 SiO₂는 변이계수가 17.9~29.6%로 다른 화학성에 비해 포장내 위치별로 변이가 컸으며, 상대적으로 유기물은 변이계수가 4.3~10.3%로 다른 화학성에 비해 포장내 위치별로 변이가 크지 않았다.

포장내 위치별 토양 화학성 및 수확량의 상관관계를 분석한 결과, 전체적으로 Ca, Mg, SiO₂ 사이에는 0.7 이상의 상관을 가지고 있었다. 그러나 수확량과 의미 있는 상관을 가진 토양 화학성은 없었다. 따라서 토양 화학성 한두 개만 가지고 수확량을 예측할 수 없음을 알 수 있다.

포장내 위치별 토양 화학성의 연도별 상관관계를 분석한 결과, 토양 화학성 중 Mg와 Ca는 연도별로 크게 변하지 않음

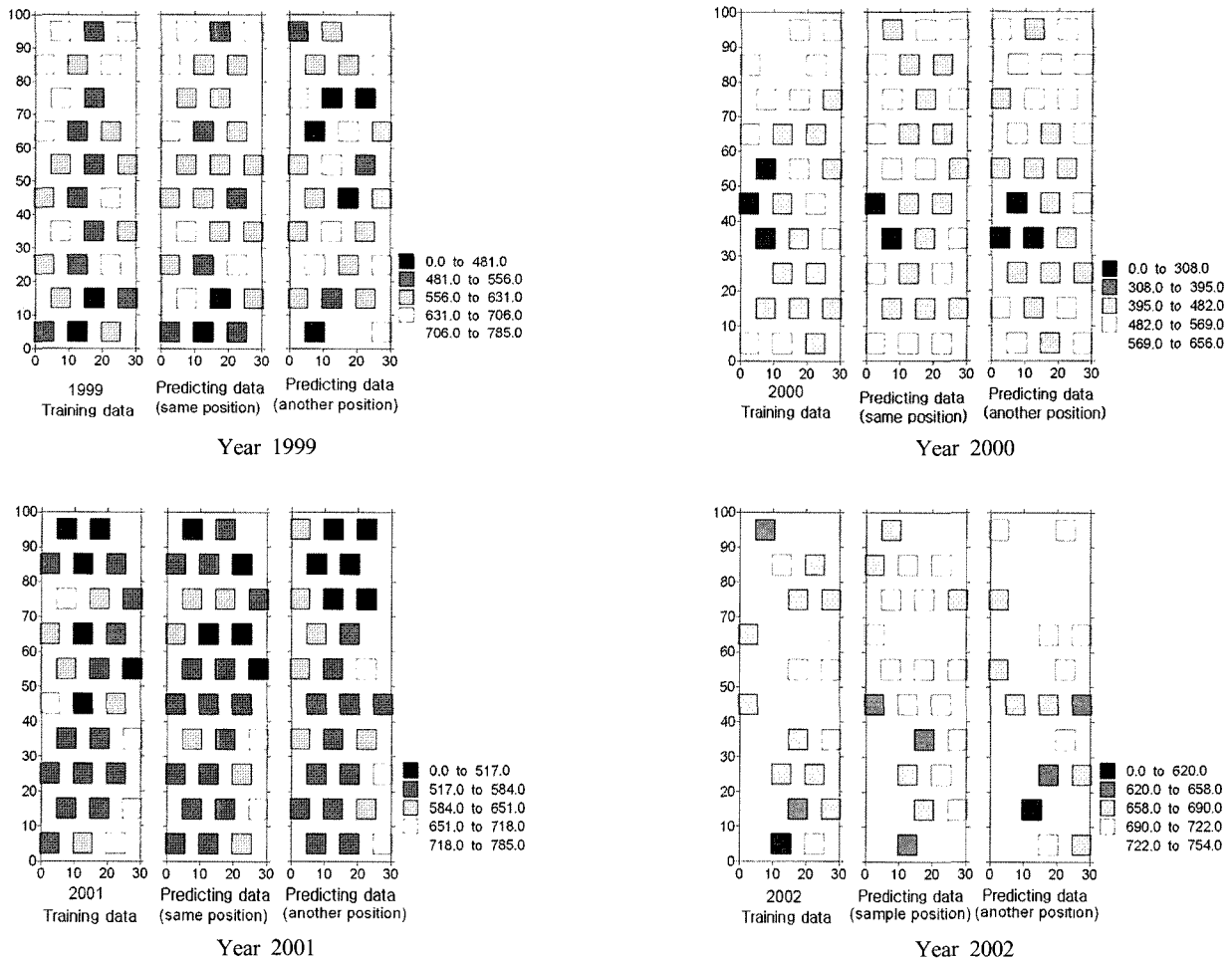


Fig. 3 Yield maps generated by soil properties in 1999 to 2002

나 다른 화학성은 해마다 변하는 것을 알 수 있고, 수확량도 매년 변함을 알 수 있다.

역전과 신경망 알고리즘으로 5 m×10 m 기본 격자를 하나씩 건너면서 토양 화학성을 학습에 사용했고, 학습결과를 이용해서 학습에 사용한 위치의 수확량을 예측했다. 신경망으로 예측한 수확량과 그 위치의 실제 수확량간의 상관관계를 분석한 결과 피어슨 상관계수가 1999년에는 0.916, 2000년에는 0.879, 2001년에는 0.800, 2002년에는 0.789로 나타났다. 이는 인공지능망을 이용하여 토양 화학성 6항목으로 그해의 수확량을 예측할 수 있음을 의미한다.

참고 문헌

1. 농업과학기술원. 2000. 토양 식물체 분석법.
2. 농협중앙회. 2000. 흙 살리기와 시비기술.
3. Drummond, S. T., K. A. Sudduth, A. Joshi, S. J. Birrell and N. R. Kitchen. 2003. Statistical and neural methods for site-specific yield prediction. *Trans. ASAE* 46(1):5-14.
4. Hong, K. C., D. W. Maong, K. Kyuma, H. Furukawa and Y. S. Suh. 1976. A method of capability evaluation for Korean paddy soils. *J. Korean Soc. Soil Sci. fert.* 12(1):15-23.
5. Kim, C. K., B. W. Lee and W. S. Han. 2001. Upper Boundary Line Analysis of Rice Yield Response to Meteorological Condition for Yield Prediction I. Boundary Line Analysis and Construction of Yield Prediction Model. *Korean Journal of Crop Science* 46(3):241-247. (In Korean)
6. Liu, J and C. E. Goering. 1999. Neural network for setting target corn yields. *ASAE paper 99-3040*, Totonto, Ontario, Canada, July 18-21.
7. Liu, J., C. E. Goering and L. Tian. 2001. A neural network for setting target yields. *Trans. ASAE* 44(3):705-713.
8. Sudduth, K. A., S. T. Drummond, S. J. Birrell and N. R. Kitchen. 1996. Analysis of spatial factors influencing crop yield. In *Proc. 3rd Int. Conf. On precision Agriculture*, Robert, O. C. et al. (ed.) pp. 129-140.