

기후변화에 따른 사료작물 생산성변화 빅데이터 분석¹⁶⁾

박성민*, 이성원***, 정애란***, 신문선*, 문상호**

*건국대학교 ICT융합공학부 컴퓨터공학과

**건국대학교 동물과학응용화학과

***충북대학교 정보통계학과

e-mail:msshin@kku.ac.kr

Big data analysis for productivity change in feed crops due to climate change

Seong-Min Park*, Seong-Won Lee***, Ae-Ran Jeong***, Moon-Sun Shin*, Moon Sang-ho**

*Dept. of Computer Science, Konkuk University

**Dept. of Animal Science and Application Chemistry, Konkuk University

***Dept. of Information Statistics, Chungbuk National University

요약

본 논문에서는 기후변화에 따른 목초사료작물의 생산성변화와 기후요소와의 상관관계 분석을 위해 기상청데이터를 활용하여 기후데이터DB를 구축하고 목초 사료작물의 과거 생산량 데이터를 수집하여 기후변화에 따른 생산성변화에 대한 빅데이터 분석을 수행하였다. 목초,옥수수,이탈리안라이그라스 3종의 사료작물에 대한 건물생산량이 기후변화와 기후요소에 따라 어떻게 변화되었는지 알아보기 위해서 기초데이터를 수집하여 DB를 구축하고 상관관계분석과 다중선형회귀분석을 수행하였다. 목초사료작물 생산량은 기후변화의 영향에 민감하여 향후 사료작물의 재배적지 변동에 변화가 있을 것으로 예측되어진다. 연구결과는 향후 기후변화에 따른 목초사료작물 생산성 예측에 활용될 수 있으며 목초사료작물의 재배적지변동 예측의 기초자료로 활용될 수 있다.

1. 연구배경

지구온난화와 기후변화는 미래의 생활에 큰 변화를 가져오게 될 것이며 다양한 측면에서 큰 영향을 끼칠 것으로 예측되고 있다. 기후변화 신시나리오인 RCP 8.5에 의하면 우리나라의 평균 기온이 2010년 기준 5.7℃ 상승하게 되며 가뭄이나 폭염 등 극한기후의 출현이 빈번할 것으로 전망되고 있다. 이러한 기후변화 및 기상재해로 인해서 농업생태 환경이 악화되고 이로 인한 사료작물의 생산성 및 품질 저하 등의 문제 발생이 야기될 것이다.

최근 이상기온 등의 기상재해가 빈번하게 발생하고 있으며 폭염, 가뭄, 고온, 건조 등 이상기상으로 인하여 사료작물의 기상재해 피해가 증가될 것으로 전망되고 있다. 조사료의 안정적 확보를 위해 사료작물의 적지 한계 이동성 변화, 생산량 및 재배환경 변화 실태조사와 더불어 영향·취약성 평가 필요하다. 초지와 사료작물의 생산량, 사료가치, 요구량뿐만 아니라, 대규모 초지나 초원의 식생까지 변화시키는 요인으로 작용함. 특히 목초, 사료작물의 생장기 중의 강수량과 온도의 파동은 조사료 생산에 가장 직접적으로 영향을 끼치는 요인

이 되고 있다.

우리나라 조사료 생산기반은 재배되고 있는 사료작물 초종이 지역별, 지형별로 특화되어 있을 뿐만 아니라 동계, 하계작물로 재배되고 있어서 기후변화에 대한 속도와 정도 또한 다르게 나타나고 있다. 따라서 지역별 맞춤형 기후변화 실태조사 및 취약성 평가가 필요하며 이를 위해서는 기상데이터와 사료작물 생산량 데이터의 상관관계 분석과 연관성 세부분석이 필요하다.

따라서 본 논문에서는 기후변화가 목초 사료작물 생산성에 미치는 영향을 분석하기 위해 기후데이터와 목초 사료작물 데이터 수집 및 전처리 후 기초데이터 DB를 구축하고, 빅데이터 분석을 수행한다. 분석결과는 향후 기후변화에 따른 목초 사료작물 생산성 예측에 활용될 수 있으며 목초 사료작물의 재배적지 변동 예측의 기초자료로 활용될 수 있다.

2. 기초데이터 DB 구축

사료작물에 있어서 생산성 실태조사 및 취약성 평가는 기후변화에 따른 초지 및 조사료 생산 분야의 악영향을 최소화하기 위한 사전 예방적이고 선제적인 기후변화 적응정책 수립의 기본 틀을 제공하므로 반드시 추진되어야 할 것이다. 이러

16) 본 연구는 농촌진흥청 신농업기후변화대응체계구축(R&D)사업(과제번호 PJ015079032021)의 지원으로 수행되었음.

한 생산성 실태조사 및 취약성평가를 위하여 선행되어야 할 것은 기후변화에 따른 목초 및 사료작물의 생산성 변화 DB구축 및 생산성변화 분석을 수행하는 일이다.

향후 고온, 가뭄, 건조 등 극한 기후변화와 이상기상으로 인하여 사료작물의 기상재해 피해에 대비하고, 기후변화에 따른 초지 및 사료작물 생산 분야의 악영향을 최소화하기 위한 사전 예방적이고 선제적인 기후변화 적응정책을 마련하기 위해 과거 생산성 실태조사와 기후데이터를 수집하여 각 지역별 연도별 건물생산량과 기상데이터를 기초데이터로 생성하였다. 기상데이터는 건물생산량 데이터의 연도별 지역별 기상데이터를 기상청 공공데이터를 통해 확보하였으며, [그림 1]은 구축된 기초데이터 DB이다. 작물은 이탈리아 라이그라스, 호밀, 청보리, 목초, 옥수수 데이터로 구성되었으며 건물수량은 kg/ha 단위이다. 기후데이터는 실험년도, 지역을 기반으로 기상청 데이터를 활용하여 생성하였으며, 연평균기온, 최고기온, 최저기온, 폭염일수, 최대폭염지속일수, 강수량을 활용하였다. 향후 추가되는 기후, 지질 등의 다양한 조건을 추가하며 지속적으로 연구를 수행할 계획이다.

연도	지역	건물수량(kg/ha)	연평균기온	최고기온	최저기온	고온일수	연고온일수(일)	폭염일수(일)	연평균강수량(mm)	최대폭염지속일수(일)	강수량
2006	1	10668	11.5	-15.6	33.3	6	3	1043.3	67.6	20719	46.15
2006	2	18966	16.2	-1.8	34	7	2	1527.4	84.1	17047	38.32
2006	3	8600	11.2	-17.5	34.8	8	3	1360.1	66.6	19721	44.27
2006	4	6193	16.1	-11.1	36.4	20	10	1195.2	69.9	1933.3	61.74
2007	1	16038	12.5	-13.6	34.9	3	2	1562.4	69.2	1934.3	43.83
2007	2	20672	16.5	-8.9	34.4	6	2	2138.9	83.2	1671.6	33.87
2007	3	14666	11.5	-13.8	34.3	1	1	1286.3	67.3	1854.5	41.68
2007	4	14543	14	-7.2	34.8	16	5	1471.7	69.9	1895	40.58
2008	1	17900	10.3	-13.4	36.2	7	3	879.1	69.8	2107.4	47.22
2008	2	11361	11.7	-15.2	35.2	8	6	1518.2	68.1	2127.1	47.88
2008	3	20717	16	-6.5	34.6	4	2	1268.8	63.9	1607	27.94
2008	4	13840	12	-11.1	33.5	2	1	908.6	74.3	2095.5	46.96
2009	1	14620	13.9	-10.8	35.3	19	5	1260	65.2	1993.2	44.67
2009	2	11198	10.3	-18.2	33.3	1	1	998.9	68.6	2151.8	48.32
2009	3	5007	11.7	-15	34.1	4	3	1527.8	66.9	2097.1	46.89
2009	4	15469	16	-2.5	33.6	6	2	1304.8	62.3	1764.7	39.07
2009	5	8543	12.3	-13.9	32.5	0	0	1074.3	73.4	2174.3	48.83
2009	6	52700	13.6	-12.9	34.4	9	2	1161.9	63.9	2071.4	47.76
2010	1	8884	12.2	-18.5	34.9	14	5	1378.3	70.4	1970.9	44.26
2010	2	5041	10.9	-19.5	31	1	1	1939.5	67.7	1865.9	43.44
2010	3	16231	15.6	-2.3	35.8	13	5	1584.9	66.5	1681.8	38.07
2010	4	7460	11.7	-13.2	31.8	0	0	2141.8	73.9	1984.3	44.56
2010	5	10291	13.5	-10.9	35.7	26	7	1462.3	64.1	1827.1	40.06
2011	1	9057	12	-18.3	34.3	6	4	1845.8	67.3	2045.5	46
2011	2	2020	16.7	-9.1	35.2	6	3	2311.2	63.9	2111.8	47.61
2011	3	16161	13.6	-2.8	34.7	8	3	1438.6	68.7	1720	38.66
2011	4	8208	11.6	-16.7	33	1	1	1704.4	73.1	2007	40.07
2011	5	5724	13.1	-13.4	36.2	16	6	1621.8	66.6	1954.6	43.91
2011	6	13826	12.3	-19.2	33.6	14	6	1095.6	70.8	2468.8	53.47
2011	7	11628	11.4	-20.7	34.3	5	2	1470.9	67.3	2288.2	53.84
2011	8	13437	16.5	-1.8	36.3	23	5	891.1	72.1	2395.2	44.73
2011	9	15820	11.8	-16.6	33.4	2	1	1016.7	69.9	2082.7	53.51
2011	10	8191	13.8	-15.3	37.8	39	18	1264.7	67.3	1911.5	50.01
2011	11	9138	10.7	-14.8	35.1	7	8	84	61.1	2148.4	48.27

[그림 1] 목초 사료작물 기초데이터 DB

3. 목초 사료작물 빅데이터 분석

구축된 기후요소데이터와 목초 사료작물 생산량 데이터에 대하여 Python 과 SPSS를 활용하여 상관분석과 다중선형회귀 분석을 수행하였다. 6 변수들간의 상관관계분석 결과는 다음 [그림 2]와 같다.

	IRG 건물수량	호밀 건물수량	청보리 건물수량	목초 건물수량	옥수수 건물수량
건물수량	1.000000	1.000000	1.000000	1.000000	1.000000
연평균기온	0.239715	-0.164086	0.207601	0.408202	0.184841
최저기온	0.447601	-0.121840	0.240502	0.398227	0.180419
최고기온	0.000602	-0.358295	-0.008498	0.366174	0.294991
고온일수	-0.089369	-0.242360	0.001771	0.212400	0.210931
연속고온지속일수	-0.163525	-0.232085	-0.015547	0.175447	0.120075
강수량	0.053881	0.011546	-0.119161	-0.236651	-0.255129

[그림 2] 상관 관계 분석 결과

모든 조건(x축)을 변수로 넣어 건물수량(y축)과 함께 선형의 관계를 갖고 있는지를 분석하기 위해 진행 하였다.

계수는 절대값이 0.3이상일 경우 어느 정도 의미가 있다고 볼 수 있다. 동계작물인 IRG(이탈리안라이그라스)의 경우 최저기온과 상관관계가 높은 것으로 나타났으며 하계작물인 옥수수의 경우 최고기온 고온일수에 영향이 있는 것으로 분석결과 확인되었다. 그러나 토양 등 작물의 특성상 여러 조건이 복합적으로 영향을 미치는 등의 변수가 많이 존재하므로 기후요소만으로 생산량이 결정된다고 할 수는 없다.

다음으로 수행된 분석은 건물수량을 종속변수로 기후요소들을 독립변수로 하여 다중회귀분석을 적용하였으며 분석결과는 다음 [그림 3]과 같다. 다중회귀분석은 IRG(이탈리안라이그라스), 목초, 옥수수에 대하여 분석을 진행하였다. 모든 독립변수로 단계선택 방법으로 다중회귀분석을 실시하였으며 R계급값이 순회귀분석의 결과보다 높은 값을 보이고 있어 추정된 선형회귀모형의 유의확률 또한 유의하다는 결과를 보여 주고 있다.

모형	Pearson 상관	LN건물수량	연평균기온			강수량
			평균	최저기온	최고기온	
1	.554**	1	.554**	.552**	-.032	.006
2	.580**	.134	.000	.134	.714	.947
N		134	134	134	134	134
연평균기온	.554**	.000	1	.896**	.240**	.042
최저기온	.552**	.000	.896**	1	.016	.094
최고기온	-.032	.000	.240**	.016	1	-.056
강수량	.006	.042	.094	-.056	-.056	1
N	134	134	134	134	134	134

** 상관관계가 0.01 수준에서 유의합니다(양측).

모형	R	R 제곱	수정된 R 제곱	추정값의 표준오차
1	.554*	.307	.302	.425868
2	.580*	.336	.326	.418514

a. 예측자: (상수), 연평균기온
b. 예측자: (상수), 연평균기온, 최고기온

모형	제곱합	자유도	평균제곱	F	유의확률	
1	회귀 잔차 전체	10.626 132 34.566	1 133	10.626 .181	58.587	.000*
2	회귀 잔차 전체	11.620 22.945 34.566	2 131 133	5.810 .175	33.172	.000*

모형		비표준화 계수		표준화 계수		t	유의확률
		B	표준오차	베타	베타		
1	(상수)	7.098	.275			25.772	.000
	연평균기온	.159	.021	.554		7.654	.000
2	(상수)	9.138	.898			10.179	.000
	연평균기온	.171	.021	.596		8.133	.000
	최고기온	-.063	.026	-.175		-2.383	.019

[그림 3] IRG 다중회귀분석결과

최종적으로 이탈리안라이그라스는 다중회귀분석의 결과가 신뢰도가 있게 나왔으며 추정된 회귀모형은 '건물수량 =

9.138+0.171*연평균기온-0.063*최고기온'으로 나왔다.

목초에 대한 다중회귀분석을 수행하기전 정규화를 진행한 데이터를 기반으로 선형적인 관계를 알아보기위한 상관분석을 먼저 진행하였다. 유의확률이 0.01보다 작은 연평균기온, 최저기온, 강수량이 나왔으며 3가지 변수에 대해 순회귀분석과 모든변수를 단계선택 방법으로 다중회귀 분석을 진행하였다. 목초의 경우 최종적인 다중회귀모형은 '건물수량 = 9.926 +0.124*연평균기온-0.08*최고기온-0.014*최저기온'으로 결과가 나왔으나 추가적인 검토가 필요한 것으로 분석되었다.

LN건물수량	Pearson 상관계수	연평균기온	최저기온	최고기온	강수량
유의확률 (양측)		.282**	.238**	-.006	.165**
N	637	.000	.000	.874	.000
연평균기온			1	.846**	-.142**
유의확률 (양측)	.282**		.000	.000	.000
N	637	637	637	637	637
최저기온				1	.403**
유의확률 (양측)	.238**	.846**			-.098**
N	637	637	637	637	.014
최고기온					1
유의확률 (양측)	-.006	.866**	.403**		-.373**
N	637	.874	.000	.000	.000
강수량					1
유의확률 (양측)	.165**	-.142**	-.098**	-.373**	
N	637	.000	.014	.000	637

** 상관관계가 0.01 수준에서 유의합니다(양측).
* 상관관계가 0.05 수준에서 유의합니다(양측).

모형	R	R 제곱	수정된 R 제곱	추정값의 표준오차
1	.282*	.080	.078	.399941
2	.384*	.147	.145	.385266
3	.401*	.161	.157	.382428
4	.414*	.171	.166	.380367

- a. 예측자: (상수), 연평균기온
- b. 예측자: (상수), 연평균기온, 최고기온
- c. 예측자: (상수), 연평균기온, 최고기온, 강수량
- d. 예측자: (상수), 연평균기온, 최고기온, 강수량, 최저기온

모형	회귀	제곱합	자유도	평균제곱	F	유의확률
1	회귀	8.783	1	8.783	54.909	.000*
	잔차	101.570	635	.160		
	전체	110.353	636			
2	회귀	16.249	2	8.124	54.735	.000*
	잔차	94.104	634	.148		
	전체	110.353	636			
3	회귀	17.776	3	5.925	40.515	.000*
	잔차	92.577	633	.146		
	전체	110.353	636			
4	회귀	18.916	4	4.729	32.686	.000*
	잔차	91.437	632	.145		
	전체	110.353	636			

- a. 종속변수: LN건물수량
- b. 예측자: (상수), 연평균기온
- c. 예측자: (상수), 연평균기온, 최고기온
- d. 예측자: (상수), 연평균기온, 최고기온, 강수량
- e. 예측자: (상수), 연평균기온, 최고기온, 강수량, 최저기온

모형	비표준화 계수 B	표준화 계수 베타	t	유의확률
1 (상수)	8.436	.080	105.179	.000
	.049	.007	.282	.7410
2 (상수)	10.646	.321	33.151	.000
	.090	.009	.514	10.461
	-.079	.011	-.349	-7.092
3 (상수)	10.040	.370	27.147	.000
	.086	.009	.490	9.922
	-.064	.012	-.285	-5.403
	.000	.000	.128	3.232
4 (상수)	9.926	.370	26.819	.000
	.124	.016	.711	7.658
	-.080	.013	-.352	-6.109
	.000	.000	.114	2.863
	-.014	.005	-.210	-2.807

[그림 3] 목초다중회귀분석결과

세 번째로 하계사료작물인 옥수수에 대한 데이터를 가지고 선형관계분석을위한 상관관계분석, 선형관계가 보이는 변수를 대상으로 순회귀분석, 단계적 다중회귀 분석 순으로 분석을 진행하였다. 회귀모형은 '건물수량=8.139+0.052*최고기온+0.00*강수량이다.

LN건물수량	Pearson 상관계수	연평균기온	최저기온	최고기온	강수량
유의확률 (양측)		.182**	.195**	.282**	-.243**
N	437	.000	.000	.000	.000
연평균기온			1	.854**	-.429**
유의확률 (양측)	.182**		.000	.000	.000
N	437	437	437	437	437
최저기온				1	.293**
유의확률 (양측)	.195**	.854**			-.328**
N	437	.000	.000	.000	.000
최고기온					1
유의확률 (양측)	.282**	.458**	.293**		-.224**
N	437	.000	.000	.000	.000
강수량					1
유의확률 (양측)	-.243**	-.429**	-.328**	-.224**	
N	437	.000	.000	.000	437

** 상관관계가 0.01 수준에서 유의합니다(양측).

모형	R	R 제곱	수정된 R 제곱	추정값의 표준오차
1	.282*	.080	.078	.253944
2	.337*	.114	.109	.249513

- a. 예측자: (상수), 최고기온
- b. 예측자: (상수), 최고기온, 강수량

모형	회귀	제곱합	자유도	평균제곱	F	유의확률
1	회귀	2.428	1	2.428	37.651	.000*
	잔차	28.052	435	.064		
	전체	30.480	436			
2	회귀	3.461	2	1.730	27.794	.000*
	잔차	27.020	434	.062		
	전체	30.480	436			

- a. 종속변수: LN건물수량
- b. 예측자: (상수), 최고기온
- c. 예측자: (상수), 최고기온, 강수량

모형		비표준화 계수		표준화 계수		t	유의확률
		B	표준오차	베타	t		
1 (상수)	최고기온	7.617	.355	.282	21.452	.000	
	강수량	.061	.010	.061	6.136	.000	
2 (상수)	최고기온	8.139	.372	.240	21.898	.000	
	강수량	.052	.010	.052	5.174	.000	
		.000	.000	-.189	-4.073	.000	

[그림 4] 옥수수다중회귀분석결과

4. 결론 및 향후 연구

본 논문에서는 사료작물별 생산성변화와 기후요소와의 관계를 분석하고자 기상데이터와 사료작물 생산선량데이터를 수집하여 전처리과정을 수행하고 기초데이터유를 구축하였다. 구축된 사료작물 기초데이터 DB를 대상으로 상관관계분석과 단순선형회귀분석과 다중회귀분석을 수행하여 생산량에 대한 회귀식을 도출하였다. 통계사료작물인 IRG는 연평균기온, 최고기온의 영향을 받으며, 하계사료작물인 옥수수의 경우 최고기온, 강수량의 영향을 많이 받는 것으로 분석되었다. 연구 수행결과 통계사료작물에는 1월, 2월의 평균기온 최저기온 등의 기상데이터에 대한 분석이 필요할 것으로 사려되었으며 하계작물인 옥수수의 경우 7월, 8월 최고기온 평균기온 강수량 및 고온지속일수 등의 기후요소에 대한 추가데이터를 생성하여 분석을 수행할 계획이다. 향후 추가적인 기후요소 데이터를 수집하여 분석결과의 정확도를 높일 수 있도록 연구를 계속 수행할 것이다.

참고문헌

- [1] 김창길, "기후변화에 따른 농업부문 영향분석과 대응전략", Korean Journal of Environmental Agriculture pp. 3-29, 한국환경농학회 춘계워크숍 자료집, 2010년
- [2] 오재호, 민영미, 우수민, "지구온난화에 따른 한반도 및 동아시아 지역의 평균기온 및 이상기온 현상의 변화", 한국

- 제4기학회 추계학술대회, pp.27-33, 2005년.
- [3] 김현애, 김광수, “공간분포모델을 활용한 사료작물 이탈리아 라이그라스의 재배적지 변동예측연구”- 한국작물학회 학술발표대회 논문집, pp. 30-30, 2014년.
- [4] 김현애, 현신우, 김광수, “공간분포모델을 활용한 사료작물 이탈리아 라이그라스 (*Lolium multiflorum* L.) 의 재배적지 변동예측연구”, *Korean Journal of Agricultural and Forest*, 2014년.
- [5] 현신우, 김광수, “국내 옥수수 재배적지 예측을 위한 R기반의 기후적합도 모델 병렬화”, *한국농림기상학회지* 제19권 제3호, pp.164-173, 2017년.