

기후변화가 북한의 농업생산에 미치는 영향분석 : 쌀 농업을 중심으로

서울대학교 북한·해외농업연구소

(임창식*, 문용식**, 류영렬***, 임정빈****)

* 서울대학교 농경제사회학부 박사과정

** 서울대학교 농경제사회학부 석사과정

*** 서울대학교 조경·지역시스템공학부 교수

**** 서울대학교 농경제사회학부 교수 및 농업생명과학연구원 겸임연구원(사업책임자)

차 례

1. 서론	3
1.1. 연구의 배경 및 목적	3
1.2. 선행연구 검토	3
1.3. 북한의 기후변화	5
1.4. 논문의 구성	7
2. 기후변화가 북한 쌀 생산에 미치는 영향과 남한과의 비교분석	9
2.1. 분석자료	9
2.1.1. 기상데이터	9
2.1.2. 농업생산량	11
2.2. 분석 모형	14
2.3. 분석 결과	18
2.3.1. 기후변화가 북한의 쌀 생산에 미친 영향분석	18
2.3.3. 기후변화가 남한의 쌀 생산에 미친 영향분석	19
2.3.3. 기후변화가 북한과 남한의 쌀 생산에 미친 영향 비교분석	22
3. 북한의 쌀 생산성과 식생지수 간 상관관계 분석 및 기후변화 영향평가	24
3.1. 분석자료	24
3.1.1. 위성 자료를 활용한 식생지수 구축	24
3.1.2. 북한 논 지역의 식생지수	28
3.2. 분석결과	30
3.2.1. 식생지수와 북한의 쌀 생산량과의 상관관계 분석	30
3.2.2. 식생지수를 활용한 기후변화가 북한의 쌀 생산량에 미치는 효과 분석	31
4. 위성 및 기상자료를 활용한 북한의 쌀 생산량 예측	33
4.1. 분석자료	33
4.2. 분석방법	36
4.3. 분석결과	38
5. 요약 및 시사점	42
참고문헌	44

1. 서론

1.1. 연구의 배경 및 목적

기후변화는 전세계에서 일어나고 있는 현상으로, 시간이 지날수록 가속화되는 양상을 보이고 있다. IPCC(Intergovernmental Panel on Climate change)(2023)는 전 세계 평균기온 변화 비교의 기준선으로 산업화 이전 시기인 1850~1900년 기간의 평균을 제시하였다¹⁾. 이를 기준으로 비교하였을 때, 2011~2020년의 지표면 온도 평년값²⁾은 산업화 이전에 비해 약 1.1℃ 상승하였다. 또한 WMO(2024)에 따르면, 지구 표면 온도는 2023년에 기준선 대비 약 1.45℃ 상승하였다. 이러한 상승 폭은 파리협정에서 제시된 산업화 이전 대비 온도 변화의 상한선인 1.5℃에 근접한 수치로, 기후변화가 빠르게 진행되고 있음을 보여준다. 북한 또한 이러한 기후변화의 영향으로 인해 평균기온이 상승하고, 극한 기후 및 자연재해 발생 빈도가 높아지는 문제가 발생하고 있다.

하지만 이러한 기후변화에 대한 북한의 적응은 미흡한 실정이다. EU 산하의 재난 위기 관리 지식센터에서 제공하는 기후변화 위기(INFORM Climate Change Risk Index³⁾) 평가 결과, 북한은 2022년 기준 191개 국가 중 147등 수준으로 평가되어, 기후변화에 취약한 것을 확인할 수 있다. 또한 2021년 북한이 UN에 제출한 자발적 국가검토보고서(2021)에 따르면, 북한은 기후변화로 인해 농업생산이 감소하고 농업 기반 시설이 파괴되는 등 식량 안보 차원의 문제를 겪고 있다.

또한, 북한은 지속적인 식량부족 문제를 겪고 있으며, 통계청(2024)에 의하면 북한의 식량 작물 생산은 지속적으로 감소하고 있다. 쌀, 옥수수, 맥류, 두류, 서류, 잡곡을 포함한 북한의 식량 생산량은 2012~2016년 기간 평균 약 472만 톤 수준이었지만, 2019~2023년 기간에는 평균 약 461만 톤으로 감소하였다. 또한 FAO(2024)에서 제공하는 통계에 따르면 영양실조에 걸린 북한 인구수는 지속적으로 증가하여 2017~2019년 기간 동안 평균 약 1,160만 명에 도달하였다. 해당 수치는 같은 기간 북한 추계인구 누적 평균인 약

1) 1850~1900년 기간은 전 지구적 기후 관측 자료를 확보할 수 있는 가장 이른 시기로, IPCC 5차 보고서부터 지속적으로 활용되고 있는 기준이다 (IPCC, <https://www.ipcc.ch/sr15/faq/>).

2) 기후 평년값은 '0'으로 끝나는 연도를 기준으로 최근 10년 혹은 30년간의 기후요소의 누적 평균값으로, 장기적인 기후변화 추세 파악에 활용된다.

3) 기후변화로 인한 위험을 평가하고 예측하기 위해 개발된 지표로 특정 국가나 지역의 기후 변화 취약성과 위험성 파악에 활용된다.

2,500만 명의 절반 수준으로, 북한이 심각한 식량난을 겪고 있음을 보여준다. 이처럼 지속적인 식량부족 문제를 겪고있지만 기후변화로 인하여 이러한 피해가 더욱 커질 것으로 예상되고 있는 상황에서 기후변화가 북한 농업생산량에 미치는 영향을 분석하는 연구는 국제협력 차원에서 중요한 정보를 제공할 수 있을 것이다.

한편 북한은 현재 농업생산량 관련 통계를 공표하지 않아, 작황 파악 및 관련 분석에 어려움이 있다. 이러한 어려움을 극복하기 위해 주로 활용되고 있는 자료는 FAO(Food and Agriculture Organization of the united nations), WFP(World Food Programme), USDA(United States Department of Agriculture)와 농촌진흥청에서 발표하는 북한 농업생산 관련 통계자료와 MODIS(Moderate Resolution Imaging Spectroradiometer), LANDSAT(Land Satellite) 등의 위성 자료가 활용되고 있다. 다만, 농업생산 통계자료는 생산량과 단수를 제공하기 때문에 분석 시 직관적인 해석이 가능하다는 장점은 있지만, 공표 기관에 따라 추정치가 달라 연구 활용 시 주의가 필요하다. 위성자료의 경우 직접 관측된 정보를 활용할 수 있는 정확도가 높은 자료이지만 농업생산량을 직접적으로 확인할 수 있는 지표가 아니기에 추가적인 분석이 필요하다. 따라서 본 연구에서는 농업생산 통계자료와 인공위성 자료를 식생지수를 상호보완적으로 활용하여, 기후변화가 북한 식량 생산량에 미치는 영향을 다각적으로 분석하고자 한다. 이를 통해 북한 농업생산 관련 데이터 부족에 대한 분석적인 한계를 극복하고, 향후 통일대비 남북 농업협력과 식량안보 강화를 위한 전략 수립에 활용될 수 있는 시사점을 제공하고자 한다.

1.2. 선행연구 검토

북한의 기후변화 및 농업생산과 관련된 다수의 연구가 진행되어 왔다. 명수정·권현한(2010)은 북한 전역의 강수 특성을 분석하고 IPCC의 기후변화 시나리오를 기반으로 강수량 및 강수일수의 변화를 예측하였다. Won-Ho Nam et al.(2015)은 기후변화로 인한 북한 내 기온과 강수량의 시공간적 패턴 변화를 연구하였으며, Sang-Hyun Lee et al.(2017)은 북한의 기후변화가 가뭄 발생에 미치는 영향을 분석하였다. 또한, 오삼언·박소영(2022)은 기후변화와 산림황폐화에 대한 북한의 대응 방안을 내부적 전략과 대외적 전략으로 나누어 분석하였다. Yongho Choi and SungJu Cho(2024)는 기후 관측 지점을 기반으로 공간적 가중 평균을 산출하여 극한 기온이 북한의 쌀 및 옥수수 생산에 미치는 영향을 분석하였다.

이와 더불어 위성 자료를 활용한 연구도 활발히 이루어지고 있다. 홍석영 외(2012)는 German Aerospace Center에서 운용 중인 RapidEye 위성 영상을 활용하여 북한의 지역별 논 면적을 추정하였다. 김관호 외(2014)는 RapidEye 위성영상 기반 농업 토지 피복 분류 결과와 KOMPSAT-2(Korean Multi-Purpose Satellite-2) 위성영상 기반 농업생산 기반 추출 결과를 결합하여 북한 신천군의 농업생산 기반 실태를 분석하였다. 홍석영 외(2015)는 MODIS 위성 자료에서 산출된 NDVI와 강수량 자료를 활용하여 북한지역의 쌀 단수를 예측하였다. Yeom et al.(2018)은 COMS(Communication, Ocean, and Meteorological Satellite) 위성 자료를 기반으로 북한 쌀 생산성을 분석하였으며, Seungtaek Jeong et al.(2022)은 LSTM(Long Short-Term Memory) 및 CNN(Convolutional Neural Network) 모형에 MODIS 위성 자료를 적용하여 픽셀 단위로 북한 쌀 단수를 예측하였다.

본 연구와 가장 비슷한 연구는 최용호 외(2021, 2024), 김다울 외(2022)의 연구에서는 기후변화가 북한의 농업생산에 미치는 영향을 분석하였다. 최용호 외(2021, 2024)는 기후변화를 고려한 북한 농업 실태 파악을 위해 종속변수로 FAO 지역별 쌀 단수를 활용하였고, 김다울 외(2022)는 북한 작황에 이상기후가 미치는 영향을 분석하기 위해 북한 기상 관측 지점의 9월 NDVI를 종속변수로 활용하여 북한 작황 관련 연구를 수행한 바 있다. 극한 기후지표인 가뭄과 폭우가 북한의 농업생산성에 부정적인 영향을 미치는 것으로 분

석하였다.

이와 같이 많은 선행연구들에서 북한의 기후변화와 농업생산을 다양한 관점에서 탐구하고 있으며, 데이터 확보가 어려운 북한지역의 작황 파악과 예측에 기여하고 있다. 본 연구에서는 더 나아가 위성 자료와 기후 데이터를 모두 활용하여 북한 농업생산에 기후변화가 미치는 영향을 분석하고 이를 남한과의 비교를 통해 북한의 기후변화에 대한 상대적 취약성을 분석한다는 차별성이 있다. 또한, 북한의 농업생산 통계자료와 식생지수와의 상관관계를 밝히고, 공식적으로 발표되고 있지 않은 북한의 지역별 농업생산량을 예측한다는 점에서 기존 연구들과 차별성을 가진다고 할 수 있다.

1.3. 북한의 기후변화

북한의 기후는 사계절이 뚜렷한 온대성 기후이다. 따라서 여름 날씨는 흐리고 무더운 편이고, 겨울 날씨는 춥고 맑은 것으로 알려져 있다. 또한, 지형적 특성과 계절풍의 영향으로 인해 동해안이 서해안보다 온난하고, 해안 지역이 내륙지방보다 온난한 것으로 알려져 있다(통일부, 2024).

이러한 북한은 기후변화로 인해 기온은 상승하고 강수량은 감소하고 있다. 1991~2020년(신평년) 기준 북한의 연평균 기온 평년값은 8.9℃, 최고 기온 평년값은 14.5℃, 최저 기온 평년값은 4.0℃인 것으로 파악되었다. 이는 북한의 1981~2010년(구평년) 기준 평년값과 비교했을 때, 연평균 기온과 최고 기온은 0.4℃, 최저 기온은 0.3℃ 상승한 것이다. 반면 강수량 평년값의 경우, 신평년의 값은 912mm로 구평년값에 비해 7.7 mm 감소한 것으로 나타났다(기상청, 2021). 이러한 변화는 기후변화로 인한 북한의 기온과 강수량의 변화가 남한보다 큰 것을 의미한다(표 1).

표 1. 남북한 기후 평년값 변화 비교

구분	남한			북한		
	구평년	신평년	차이	구평년	신평년	차이
평균기온(℃)	12.5	12.8	+0.3	8.5	8.9	+0.4
최고기온(℃)	18.1	18.3	+0.2	14.1	14.5	+0.4
최저기온(℃)	7.7	8.0	+0.3	3.7	4.0	+0.3
강수량(mm)	1,307.7	1,306.3	-1.4	919.7	912.0	-7.7

자료: 기상청(2021)

기후변화는 대기 순환과 해양 순환을 변화시켜 기상 패턴의 변화를 유발하여 북한의 극한기후 및 자연재해 발생 빈도를 높이고 있다. 명수정 외(2022)에 따르면 북한에서 발생한 평균 폭염일수는 구평년 10.1일에서 신평년 11.8일로 1.2일 증가하였다. 또한, 강택구 외(2021)에 따르면 북한 내 자연재해⁴⁾ 발생 횟수는 1991~2000년 기간 평균 5.5회에서 2011~2020년 기간 평균 11.2회로 약 2배 증가하였다.

⁴⁾ 태풍, 홍수, 호우, 강풍, 해일, 대설, 가뭄, 지진, 황사, 한파, 한파, 붕괴 및 낙석을 포함한다.

1.4. 논문의 구성

본 연구의 2장에서는 쌀을 중심으로 기후변화가 북한 식량 생산에 미치는 영향을 남한과 비교 분석한다. 남한과 북한 모두 식량 작물 중 쌀 생산 비중이 가장 높기 때문에, 식량 작물로서의 대표성을 가지는 쌀을 중심으로 비교분석을 진행하도록 한다. 본 연구에서 수행하는 기후변화 영향에 대한 정량적 비교를 통해 남북 농업협력 추진 및 관련 정책 마련에 기여할 수 있을 것으로 기대한다.

3장에서는 위성 자료를 활용하여 공식적인 자료가 부족한 북한 농업 생산량 정보를 보완하고 이를 분석에 활용하고자 한다. 우선 위성 자료를 기반으로 계산되는 다양한 식생지수 중 북한 쌀 생산량과 상관관계가 높은 식생지수를 선별한다. 이후, 선별된 식생지수를 북한 쌀 작황에 대한 대리변수로 활용하여 기후변화가 쌀 생산량에 미치는 영향을 분석한다. 본 분석은 원격 탐사를 통한 농업 분석에 가장 널리 활용되고 있는 식생지수인 NDVI를 다양한 식생지수와 비교하여 실제 성능을 객관적으로 평가한다는 점에서 연구적 의의가 있을 것이다. 또한, 이 과정에서 식생지수의 다양한 통계값을 포함하는 구조적 데이터 지표(Metric)를 활용했다는 점에서 기존 연구와의 차별성을 가진다고 할 수 있다.

4장에서는 2장과 3장에서 활용한 기상변수와 위성 자료를 활용하여 북한의 농업생산량을 예측한다. LASSO(Least Absolute Shrinkage and Selection Order) 회귀모형과 RIDGE 회귀분석, 그리고 ELASTIC-NET 회귀분석을 통해 북한의 지역별 농업생산량을 예측한다. 본 예측은 북한의 공표 여부와 관계없이 확보할 수 있는 기후 데이터와 위성 자료를 활용하여 북한 작황을 예측할 수 있다는 점에서 연구적 의의가 있을 것이다.

5장에서는 앞서 진행한 세 가지 분석 결과에 대한 요약을 제시하고, 남북 농업 협력에 대한 본 연구의 시사점을 제시한다.

2. 기후변화가 북한 쌀 생산에 미치는 영향과 남한과의 비교분석

본 장에서는 남한과 북한의 농업생산에 기후변화가 미치는 영향을 비교 분석한다. 구체적으로, 남한과 북한의 쌀 단수가 기후 변인에 어떠한 영향을 받는지 비교하였다. 이를 통해 북한 농업생산의 기후 취약성을 남한과 비교하여 상대적으로 평가하였고, 이는 추후 정책 마련의 기틀로 활용할 수 있을 것이다.

2.1. 분석자료

2.1.1. 기상데이터

남한과 북한의 기상 데이터는 세계기상기구(World Meteorological Organization)에서 관측한 기상 데이터와 Global drought Monitor에서 제공하는 표준 강수 증발산 지수인 SPEI(Standardized Precipitation-Evaporation Index)를 활용하였다. WMO 기상 데이터는 관측 지점을 기준으로 측정된 기온, 시정, 바람, 구름, 기압, 강수, 습도 항목에 대한 정보를 분, 시간(정시), 일, 월, 연 단위로 제공하고 있으며, 기상청 홈페이지⁵⁾를 통해 내려받을 수 있다. 본 연구에서는 분석 기간인 2002년부터 2020년까지 연속적으로 일별 및 월별 기상 데이터를 제공하는 남한의 67개 지점과 북한의 27개 지점을 선별하여 연구에 활용하였다. 다만 WMO 기상 데이터는 남한과 북한 모두 시도별로 제공되지 않기 때문에, 본 연구에서는 티센망(Thiessen network) 분석을 활용하여 관측지점별 데이터를 가중합하여 시도별 데이터로 변환하였다. 성재훈 외(2019)에 따르면 티센망 분석은 수문학(Hydrology)에서 특정 지역을 인접한 관측소들의 가중 평균으로 표현하는 방법이다. 본 연구에서 각 관측 지점은 주변 관측 지점과의 인접도를 기준으로 <그림 1>과 같이 분할되어 관측 지점마다 티센 폴리곤(polygon)을 형성한다.

⁵⁾ <https://data.kma.go.kr/data/grnd/selectNkRltmList.do>

그림 1. 관측 지점별 가중치 산출을 위한 남북한 티센 폴리곤 형성



자료: WMO 기상 데이터를 활용하여 저자 작성.

특정 시도의 면적 대비 각 관측 지점의 티센 폴리곤이 중첩되는 영역 비율을, 특정 시도에 대한 각 관측 지점의 가중치로 활용한다. 중첩되는 부분이 없는 경우 가중치는 0으로 계산되며, 각 시도에 대한 관측 지점 가중치 합은 1이 되도록 한다. 이와 같이 산출된 가중치를 활용하여 각 시도의 평균기온, 누적 강수량, 기온 변이계수, 강수강도, 폭염일수, 호우일수를 산출하여 연구에 활용하였다.

SPEI는 위도 경도 1°의 공간 해상도(약 100km)와 한 달의 시간 해상도를 가지고 있는 가뭄지수로, 강수량과 증발산량을 활용하여 산출한다. SPEI는 Global drought monitor를 통해 48개월 누적치까지 제공되고 있다. 본 연구에서는 남한과 북한의 각 시도에 중첩되는 영역에서 2002년부터 2020년까지 값을 추출하여 연구에 활용하였다. SPEI 값에 대해 기상청(2024)에서 제공하는 단계별 해석과 SPEI 시각화 예시는 <그림 2>와 같다.

그림 2. SPEI의 단계별 해석 및 시각화 예시

가뭄단계		지수범위	SPEI 시각화 예시
습함		$1.00 \leq \text{SPEI}$	
정상		$-0.99 \leq \text{SPEI} \leq 0.99$	
약한 가뭄(관심)		$-1.49 \leq \text{SPEI} \leq -1.00$	
보통 가뭄(주의)		$-1.99 \leq \text{SPEI} \leq -1.50$	
심한 가뭄(경계)		$-2.00 \leq \text{SPEI}$	
극심한 가뭄(심각)		-2.0 이하 20일 이상 지속	

자료: 기상청(2024)과 Global drought monitor(2024)을 활용하여 저자 작성.

2.1.2. 농업생산량

가. 북한의 쌀 생산량

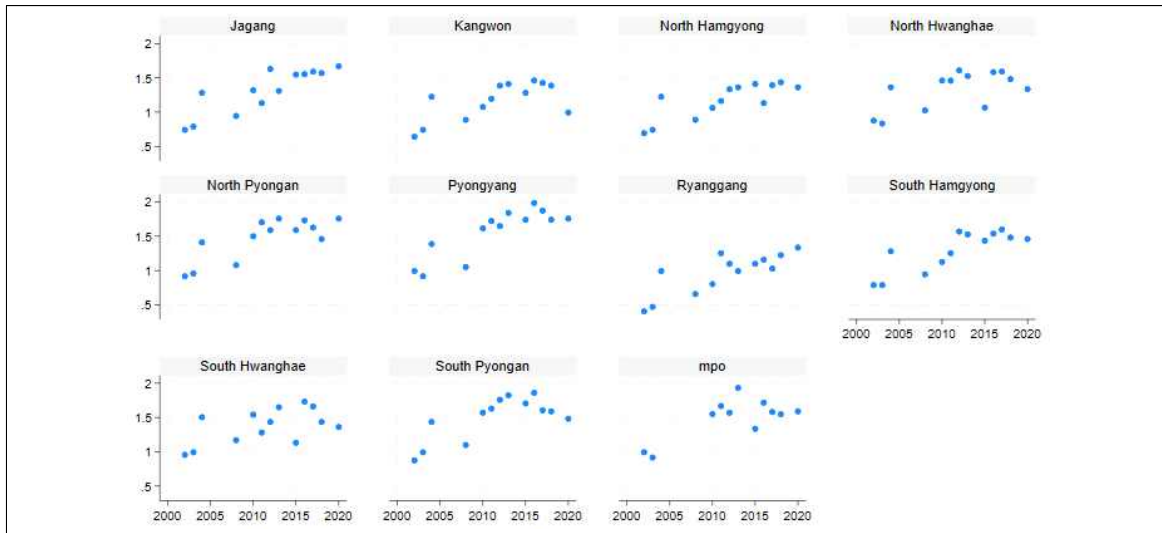
북한의 농업생산 통계는 공식적으로 발표되고 있지 않지만, FAO, WFP, USDA에서 지역별 농업생산통계를 간헐적(2002, 2003, 2004, 2008, 2010, 2011, 2012, 2013, 2015, 2016, 2017, 2018, 2020)으로 발표하고 있어 이를 통합하여 분석에 활용하였다. 북한의 쌀 생산과 관련된 기초통계는 <표 2>와 같다. 재배면적과 생산량이 가장 큰 지역은 황해남도로 북한의 쌀 생산 중심지인 것으로 나타났다. 단수의 경우 평양시가 5.01ton/ha로 가장 높았으며, 가장 단수가 낮은 양강도(2.72) 대비 85% 높은 수치를 보여 지역별 쌀 생산성의 차이가 큰 것으로 나타났다. 또한, 북한의 연도별 쌀 생산성<그림 3>은 2002년부터 2020년까지 지속해서 증가하는 추세를 보이고 있다.

표 2. 북한의 지역별 쌀 생산량 기초통계

구분	재배면적(ha)		생산량(ton)		단수(ton/ha)	
	평균	표준편차	평균	표준편차	평균	표준편차
자강도	6.6	0.5	25.1	6.2	3.89	1.13
강원도	27.5	8.5	87.0	25.1	3.30	0.82
함경북도	25.7	1.0	85.2	21.3	3.31	0.77
황해북도	48.3	14.8	187.2	75.8	3.88	0.97
평안북도	101.3	1.8	455.8	115.5	4.50	1.18
평양시	18.3	6.7	82.7	16.3	5.01	1.53
양강도	1.6	0.4	4.3	1.0	2.72	0.72
함경남도	58.0	3.1	217.6	54.6	3.77	0.99
황해남도	140.3	9.3	568.9	134.1	4.06	0.99
평안남도	84.4	16.0	380.0	83.6	4.65	1.29
남포시	22.6	4.4	107.5	42.7	4.61	1.25

자료: FAO, WFP, USDA 각 연도

그림 3. 북한의 지역별·연도별 쌀 생산성(단위: log(ton/ha))



자료: FAO, WFP, USDA 각 연도

나. 남한의 쌀 생산량

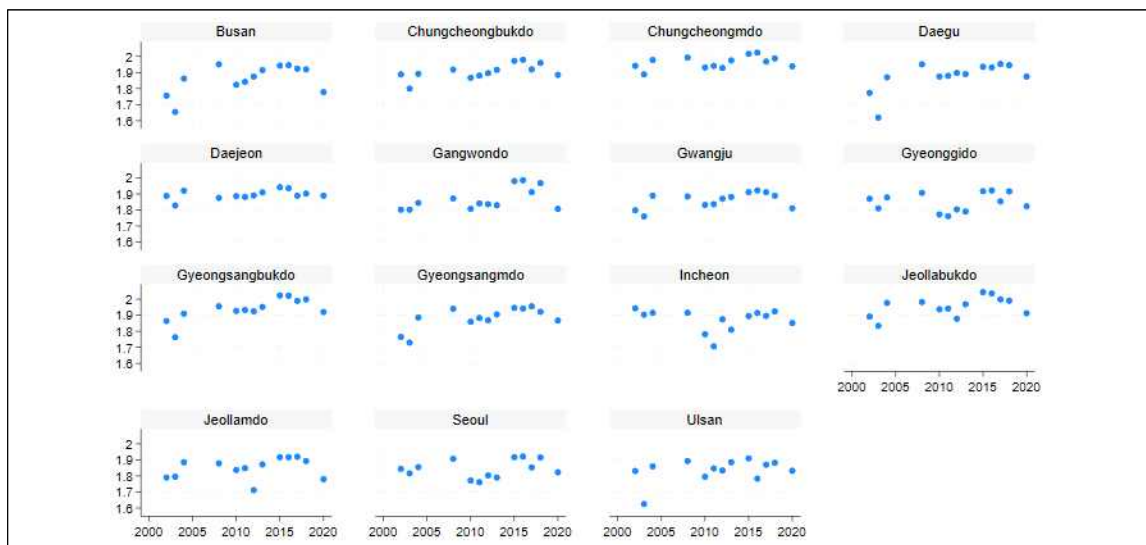
남한의 경우 연속성이 있는 자료구축을 위해 세종과 제주를 제외한 15개의 시도를 기준으로 데이터를 통계청의 작물생산조사 자료를 통해 수집하였다. 재배면적과 생산량이 많은 지역은 전라남도, 충청남도, 전라북도 순으로 나타났다. 단수는 충청남도가 7.12ton/ha로 가장 높은 것으로 나타났으며 북한의 평양시와 비교하면 약 42% 높은 쌀 생산성을 보인다. 남한의 연도별 쌀 생산성은 북한과 달리 일정 수준에서 안정적으로 유지되고 있는 것으로 나타났다.

표 3. 남한의 쌀 생산 기초통계

구분	재배면적(ha)		생산량(ton)		단수(ton/ha)	
	평균	표준편차	평균	표준편차	평균	표준편차
부산	34.4	10.1	219.0	54.7	6.45	0.55
충청북도	442.2	87.1	2,961.4	505.8	6.73	0.32
충청남도	1,514.7	142.2	10,748.4	896.1	7.12	0.27
대구	35.9	10.3	231.2	50.8	6.56	0.55
대전	16.1	5.5	107.0	35.2	6.65	0.19
강원	360.9	62.5	2,326.1	326.3	6.49	0.46
광주	61.2	12.5	392.1	69.6	6.44	0.31
경기도	926.8	144.5	5,883.0	941.3	6.36	0.37
경상북도	1,136.3	141.8	7,855.9	671.9	6.95	0.47
경상남도	803.6	119.2	5,257.8	603.2	6.58	0.44
인천	126.3	24.7	824.9	187.8	6.51	0.42
전라북도	1,302.5	129.8	9,174.9	799.6	7.06	0.43
전라남도	1,749.3	148.7	11,126.3	981.5	6.37	0.40
서울	3.1	1.8	19.8	11.6	6.33	0.36
울산	59.1	15.2	368.7	88.0	6.28	0.43

자료: 통계청, 작물생산조사, 각 연도

그림 4. 남한의 지역별·연도별 쌀 생산성(단위: log(ton/ha))



자료: 통계청, 작물생산조사, 각 연도

2.2. 분석 모형

본 연구에서 활용할 수 있는 북한의 지역별 쌀 생산성 자료는 11개의 지역, 13개의 연도로 분석에 활용할 수 있는 관측치 수는 143⁶⁾개이다. 패널데이터로 구축이 가능하긴 하지만 적은 패널(지역), 짧은 기간(시계열)으로 심층적인 분석에 활용되기에는 한계가 있다. 관측치와 관련된 경험법칙(Rule of Thumb)을 고려하면 모형의 타당성을 확보하고자 할 때 최대 7~8개 이하의 설명변수로 구성된 모형이 적합할 것으로 판단된다.⁷⁾ 이에 본 연구에서는 기상정보 및 극한기후요소를 대표할 수 있는 정보를 우선적으로 고려하고자 하였다. 농업생산에 영향을 미치는 가장 주된 요소인 기온과 강수량을 기준으로 평균적인 기후지표, 기후 변동성 지표, 극한기후 지표로 3가지로 구분하였다.

평균 기후지표에는 벼 재배기간 동안의 평균 기온, 누적 강수량이 포함되며, 기후 변동성 지표로는 기온의 변이계수, 강수강도 변수를 활용하였다. 극한기후 지표로는 WMO와 IPCC에서 제안하고 있는 폭염일수, 폭우일수, 가뭄지수(SPEI6)를 활용한다. 호우일수와 강수강도 변수간의 상관계수는 남북한 모두 0.75 이상이기에 동시에 두 변수를 활용하기에는 다소 무리가 있을 것으로 판단된다.

표 4. 분석에 활용된 변수들의 정의 및 단위

변수명	정의	단위
단수	단위 면적당 연간 쌀 생산량	ton/ha
평균 기온	벼 재배기간 동안의 일별 기온의 평균	℃
누적 강수량	벼 재배기간의 일 강수량을 누적하여 더한 값	cm
기온 변이계수	일별 기온의 표준편차를 일별 기온의 평균으로 나눈 값	-
강수강도	벼 재배기간의 누적 강수량을 강수일수(일 강수량이 1mm 이상인 날)로 나눈 값	cm/일
폭염일수	벼 재배기간 중 일 최고 기온이 33℃ 이상인 날의 수	일
호우일수	벼 재배기간 중 일 강수량이 80mm 이상인 날의 수	일
가뭄 여부	9월 기준 SPEI6의 값이 -1 이하이면 1, 아니면 0	더미변수

⁶⁾ 남포시의 경우 11개의 연도에 대해서만 사용이 가능하다.

⁷⁾ 이러한 자료의 부족으로 생기는 문제를 해결하기 위한 대안으로 본 연구의 3장에서는 인공위성 자료를 활용하여 쌀의 생육주기별로 구분하여 파종기, 생장기, 수확기로 구분하고 추가적인 분석을 진행하였다.

본 연구에서 활용한 모형은 공통추세가 고려된 지역 고정효과모형을 활용한 패널 분석이며 이는 수식 (1)과 같이 표현할 수 있다.

$$y_{it} = \alpha + X_{it}\beta + \mu_i + \gamma t + \varepsilon_{it} \quad (1)$$

y 는 로그 변환된 쌀의 단수, i 는 시도 단위의 지역, t 는 연도, α 는 절편, X 는 기후 지표, μ_i 는 지역 고정효과, γt 는 공통추세, ε 는 오차항을 의미한다. 지역 고정효과를 통제하는 것은 북한의 지역 간의 차이가 아니라 지역 내에서 시간에 따른 기후변화가 농업생산량에 어떠한 영향을 미치는지를 보고자 하기 때문이다.

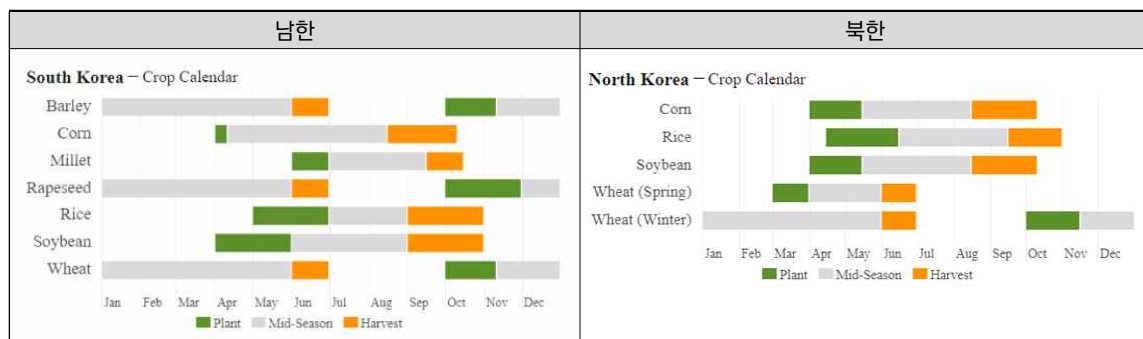
한편, 본 연구와 가장 유사한 연구(김다울 외, 2022; Choi et al., 2024)에서는 기후와 북한의 농업생산성 관계를 분석한 패널모형에서 각각 연도 고정효과와 개별 추세모형을 활용하였다. 각 연구의 방법과 목적에 따라 다양한 방법론을 활용할 수 있지만 본 연구에서는 연도 고정효과나 개별추세를 통제하기보다는 공통추세를 모형에 포함하여 시간에 따른 공통의 변화를 통제하고자 한다. 이는 북한의 모든 지역의 농업생산성이 시간에 따라 동일한 추세로 연간 γ 만큼 변화한다고 해석할 수 있다. 이러한 선형추세가 공통의 흐름에서 통제해야 될 것으로 판단되며 이로부터 벗어나는 변동은 중요한 정보를 제공하는 고유한 변동으로 간주된다. 예를 들어 특정 연도의 기후 현상으로 인해 북한의 모든 지역에서 기후지표(설명변수)가 동일하게 한 단위 증가하였다면, 이러한 변화는 쌀 생산성(종속변수)의 변화를 설명하는데 사용되지만, 연도별 고정효과를 사용하게 되면 모든 지역에서 동일하게 변화한 기후요인은 종속변수의 변화를 설명하는데 사용되지 않고 연도 고정효과로 설명되기 때문이다. 그리고 개별 추세모형은 패널(지역)별로 각자의 다른 계수를 갖도록 추정하는 것인데 FAO, WFP, USDA에서 제공하는 데이터는 11개의 지역에 대한 13개 연도인 자료로 짧은 시계열로 인해 이를 통한 개별 추세항을 따로 추정하였을 경우 타당성이 확보되기 어려울 것으로 판단하기 때문에 모든 지역에서 공통의 추세를 갖는 모형을 설정하였다.

또한, 표준오차를 계산에는 클러스터 표준오차를 활용하고자 한다. 농업생산량 자료는 동일한 지역에서 다년간의 데이터가 수집되는 구조로 동일한 지역 내에서의 상관성을 고려하지 않으면 효율적인 추정량을 얻지 못하게 된다. 이러한 동일한 지역 내의 상관성을 고려하기 위하여 일반적으로 클러스터 표준오차가 사용된다. 하지만 50개 미만의 클러스

터(지역)을 가질 때 이러한 클러스터 표준오차 추정치의 불안정성이 존재하게 되는데, 본 연구에서 사용되는 남북한의 패널을 모두 합하여도 30개 이하이다. 클러스터의 개수가 적을 때 이를 교정할 수 있는 방법으로는 Cameron, Gelbach, and Miller(2008)에서 제안된 와일드부트스트랩이나 Bell and McCaffrey(2002)이 제안하는 소표본 자유도 조정을 통해 이러한 문제를 해결할 수 있다. 본 연구에서는 분석에 유연하게 적용될 수 있는 소표본 자유도 조정을 통해 표준오차를 추정하여 클러스터의 수가 적을 때 표준오차의 왜곡문제를 해결하고자 하였다.

또한, 벼 재배기간은 위도가 다른 남한과 북한을 고려하기 위하여 USDA IPAD에서 제공하고 있는 Cropcalendar에 의해 벼 작물의 생산주기를 구분하였다. 북한의 경우 4월 중순(20일)부터 5월 초(10일)까지를 파종기, 5월 초순(11일)부터 9월 중순(20일)까지를 생장기, 9월 중순부터 10월까지를 수확기로 구분하였다. 남한의 경우 5~6월을 파종기 7월~8월을 생장기, 9~10월을 수확기로 정의하고 분석을 진행하였다.

그림 5. 한반도의 작물 생산주기



자료: USDA

본 연구에서 활용한 기초통계는 다음과 같다. 남한의 평균 쌀 단수는 6.593ton/ha로 북한의 쌀 단수(3.965ton/ha)보다 약 1.7배 높은 것을 알 수 있다. 또한, 남한의 평균기온과 강수량 모두 북한보다 높은 것으로 나타났다. 기온의 변동성을 나타내는 기온 변이계수는 북한이 높은 것으로 나타났으나, 강수강도, 폭염일수, 호우일수, 가뭄여부는 모두 남한이 북한보다 높은 것으로 나타났다.

표 5. 분석에 활용된 변수들의 기초통계

구분	남한				북한			
	평균	표준편차	최솟값	최댓값	평균	표준편차	최솟값	최댓값
단수	6.593	0.472	5.05	7.72	3.965	1.232	1.5	7.272
ln(단수)	1.883	0.731	1.619	2.044	1.326	0.332	0.405	1.984
평균 기온	21.151	0.960	18.276	23.541	17.927	2.102	11.911	20.719
누적 강수량	106.643	30.641	42.085	183.837	87.426	31.341	23.710	209.966
기온 변이계수	0.221	0.026	0.158	0.300	0.303	0.064	0.215	0.532
강수강도	1.928	0.498	0.852	3.356	1.516	0.489	0.527	3.181
폭염일수	12.159	10.991	0	51	2.839	4.944	0	25
호우일수	2.05	1.531	0	9	1.336	1.472	0	7
가뭄 여부	0.236	0.426	0	1	0.133	0.341	0	1
공통추세	10.462	5.653	1	19	10.462	5.658	1	19
관측치 수	195개				144개			
벼 재배기간	5월 ~ 10월				4월 20일 ~ 10월			

2.3. 분석 결과

2.3.1. 기후변화가 북한의 쌀 생산에 미친 영향분석

기후요소가 북한의 쌀 생산성에 미치는 영향을 분석한 결과는 <표 6>와 같다. <표 6>의 분석 모형 중 모형 (1)은 평균 기후지표만을 포함한 모형, 모형 (2)는 평균 기후지표와 기후 변동성지표를 포함한 모형, 모형 (3)은 평균 기후지표와 극한 기후지표를 포함한 모형, 모형 (4)와 모형 (5)는 평균 기후지표와 변동성 지표, 극한 기후지표 모두를 포함하는 모형이다. 모형 (4)와 (5)의 강수량과 호우일수의 경우 두 변수의 상관계수가 큰 것으로 나타나 다중공선성 문제를 피하기 위하여 각각의 모델에서 구분하여 통제하고 분석을 진행하였다.

평균 기후지표만을 고려한 모형 (1)에서는 평균 기온만이 쌀 단수에 긍정적인 영향을 미치는 것으로 나타났다. 변동성 지표까지 포함한 모형(2)에서는 평균 기온과 누적 강수량은 쌀 생산성에 긍정적인 영향을 미치지만, 강수량은 부정적인 영향을 미치는 것으로 나타났다.

평균 기후지표와 극한 기후지표를 포함하는 모형 (3)에서는 평균 기온은 쌀 생산성에 긍정적인 영향을 미치지만, 폭염일수는 부정적인 영향을 미치는 것으로 나타났다. 극한기후 지표 중 호우일수와 가뭄더미 변수의 경우 약한 유의수준에서 쌀 생산성에 부정적인 영향을 미치는 것으로 나타났다.

평균 기후지표, 변동성 지표, 극한 기후지표를 모두 포함한 모형 (4)에서는 평균 기온은 쌀 단수에 긍정적인 영향을 미치는 것으로 나타났으며 누적 강수량은 약한 유의수준에서 긍정적인 영향을 미치는 것으로 나타났다. 또한, 극한 기후지표인 강수량과 가뭄더미는 쌀 생산성에 부정적인 영향을 미치는 것으로 나타났다. 모형(5)에서는 평균 기온은 양의 영향을 미치는 것으로 나타났으며, 가뭄더미는 부정적인 영향을 미치는 것으로 나타났다. 폭염일수와 호우일수는 약한 유의수준에서 쌀 단수에 부정적인 영향을 미치는 것으로 나타났다.

누적 강수량 변수의 경우 쌀 생산성에 양의 영향을 미치는 것이 확인은 되나 단독으로 활용되거나 극한 기후지표가 포함될 때 유의성이 낮아지는 결과를 보이는데, 이는 강수량

자체보다는 강수량의 변동성이 쌀 생산성에 큰 영향을 미치고 강수량과의 영향이 서로 상쇄되는 것으로 판단된다.

모형 (4)와 모형 (5) 를 기준으로 기후요소들이 쌀 생산성에 미치는 효과의 크기를 살펴보면 모형 (4)에서는 다른 조건이 일정할 때 평균 기온이 1℃ 높아지면 쌀의 단수는 17.7% 높아지는 것으로 나타났으며, 강수강도(강수일 당 누적 강수량)가 10mm 높은 경우 쌀의 단수는 27.3% 감소하는 것으로 나타났다. 또한, 가뭄인 경우 쌀의 단수가 10.1% 감소하는 것으로 나타났다. 모형 (5)의 경우 평균 기온이 1℃ 높아지면 쌀의 단수는 15.4% 증가, 가뭄인 경우 9.7% 감소하는 것으로 나타났다.

모형의 설명력의 경우 호우일수보다 강수강도 변수를 포함한 모형 (4)의 R²가 0.664로 가장 크게 나타났다.

표 6. 기후변화가 북한의 쌀 생산에 미치는 영향(종속변수: log(쌀의 단수))

독립변수 \ 모형	(1)	(2)	(3)	(4)	(5)
평균 기온	0.137*** (0.032)	0.140*** (0.041)	0.133*** (0.032)	0.163*** (0.045)	0.143** (0.051)
누적 강수량	0.001 (0.001)	0.004** (0.001)	0.001 (0.001)	0.003* (0.001)	0.001 (0.001)
기온 변동성		-0.427 (0.463)		-0.077 (0.492)	0.216 (0.570)
강수강도		-0.330*** (0.098)		-0.319*** (0.096)	
폭염일수			-0.004** (0.002)	-0.004 (0.002)	-0.004* (0.002)
호우일수			-0.038* (0.018)		-0.037* (0.018)
가뭄더미			-0.100* (0.046)	-0.107** (0.043)	-0.102** (0.044)
공통추세	0.026*** (0.004)	0.029*** (0.003)	0.029*** (0.004)	0.030*** (0.003)	0.029*** (0.004)
상수항	-1.486** (0.521)	-1.189 (0.763)	-1.420** (0.527)	-1.627* (0.857)	-1.651 (0.995)
지역 고정효과	YES	YES	YES	YES	YES
Within R ²	0.618	0.651	0.645	0.664	0.645

주1: *, **, ***는 10%, 5%, 1% 유의수준 하에서 회귀계수가 0이라는 귀무가설이 기각됨을 의미함.

주2: 소표본 자유도가 조정된 클러스터 표준오차를 활용함.

2.3.3. 기후변화가 남한의 쌀 생산에 미친 영향분석

기후변화가 남한의 농업생산에 미치는 영향을 분석한 연구는 매우 활발히 진행되어왔다. 남한의 경우 1970년대부터 시군별 농업생산자료가 조사되어 왔기 때문에 기후변화의

영향을 분석함에 있어 평균적인 생산성과 생산의 변동성, 공간적·시간적 해상도를 고려한 분석, 공간적 의존성을 반영한 연구, 기후변화 시나리오에 따른 생산량 예측 등 심도 있는 연구가 가능하다(권오상&김창길, 2008; 조현경 외, 2013; 조현경·권오상, 2014; 김용준, 2019). 그러나 이러한 분석은 본 연구의 범위를 벗어나므로, 본 절에서는 기후변화가 남한과 북한의 농업 생산성에 미치는 영향을 비교분석하기 위해 북한의 자료와의 통일성을 갖춘 데이터를 구축하고 이를 계량경제 모형을 통해 비교분석을 진행하였다.

기후 지표가 남한의 쌀 생산성에 미치는 영향은 북한의 결과와 상이한 것으로 나타났다. 극한 기후지표를 포함하지 않은 모형(1)과 모형(2)를 제외한 모든 모형에서, 평균 기온은 쌀 생산성에 통계적으로 유의한 영향을 미치지 않는 것으로 분석되었다. 반면, 폭염일수는 모든 모형에서 쌀 생산성에 긍정적인 영향을 미치는 것으로 나타났다. 이는 남한의 쌀 생산이 현재까지는 온난화로 인한 피해보다는 높은 온도로 인해 긍정적인 영향을 받고 있음을 시사한다. 이러한 결과는 농촌진흥청에서 수행한 서명철 외(2020)의 연구와 일치하는데, 선행연구에 따르면 1996년 이전에 개발된 벼 품종은 높은 온도에서 수확량이 감소하는 경향을 보였으나, 이후 개발·보급된 품종들은 온난화에 적응하여 수확량이 안정적으로 유지되고 있는 것으로 판단하고 있다.

모든 모형에서 누적 강수량은 쌀 생산성에 부정적인 영향을 미치는 것으로 나타났다. 남한의 경우 관개 인프라가 잘 구축되어 있어 높은 수리답율을 보이고 있어, 이러한 결과는 벼 생산에서 농업용수 부족보다는 강수로 인한 양분유실(권오상·김창길, 2008), 병해충 발생(김용준, 2019) 등의 문제가 더 큰 것으로 판단된다.

남한의 기후지표가 쌀 생산성에 미치는 영향의 크기를 살펴보면 모든 변수를 통제한 모형 (4), 모형 (5)에서는 벼 재배기간동안의 누적 강수량이 10mm 상승하면 쌀 단수는 0.1% 감소하며, 일 평균 기온이 33℃가 넘는 폭염일수가 하루 늘어날수록 쌀 단수는 0.3% 증가하는 것으로 나타났다. 한편, 기온의 변이계수는 약한 유의수준 하에서 쌀 생산성에 부정적인 영향을 미치는 것으로 나타났다.

표 7. 기후변화가 남한의 쌀 생산에 미치는 영향(종속변수: log(쌀의 단위))

독립변수 \ 모형	(1)	(2)	(3)	(4)	(5)
평균 기온	0.013* (0.006)	0.012* (0.006)	-0.007 (0.007)	-0.011 (0.009)	-0.010 (0.008)
누적 강수량	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
기온 변동성		0.196** (0.090)		-0.332* (0.174)	-0.346* (0.172)
강수강도		0.002 (0.011)		-0.017 (0.013)	
폭염일수			0.002*** (0.000)	0.003*** (0.001)	0.003*** (0.001)
호우일수			0.004 (0.003)		0.004 (0.003)
가뭄더미			-0.003 (0.009)	-0.006 (0.008)	-0.008 (0.009)
공통추세	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)
상수항	1.724*** (0.142)	1.698*** (0.134)	2.125*** (0.156)	2.283*** (0.206)	2.277*** (0.197)
지역 고정효과	YES	YES	YES	YES	YES
Within R ²	0.573	0.577	0.629	0.634	0.636

주1: *, **, ***는 10%, 5%, 1% 유의수준 하에서 회귀계수가 0이라는 귀무가설이 기각됨을 의미함.

주2: 소표본 자유도가 조정된 클러스터 표준오차를 활용함.

2.3.3. 기후변화가 북한과 남한의 쌀 생산에 미친 영향 비교분석

기후지표가 북한과 남한의 농업생산에 미치는 이질적 영향을 분석하기 위하여 북한 더미변수와 기후변화의 상호작용항을 모형에 포함시켜 분석을 진행하였다. 한 모형안에 변수와의 상호작용항을 포함하여 두 그룹간의 비교를 하는 것을 일반적으로 많이 활용되고 있으며, 이러한 방법은 개별 모형을 통해 추정된 계수 값을 비교하는 것보다 통계적 가설검정까지 동시에 수행할 수 있다는 장점이 있다.

〈표 8〉처럼 북한 더미변수와 각 변수들 간의 상호작용항의 계수의 통계적 유의성 검정을 통해 기후변화가 남한과 북한의 쌀 생산성에 미치는 영향의 차이를 확인할 수 있다. 모든 조건들을 통제한 모형 (4)와 모형 (5)를 통해 기후요소가 남한과 북한의 쌀 단수에 미치는 영향의 크기를 비교해보면, 다른 모든 조건들이 동일할 때 북한의 쌀 생산성에 평균 기온과 누적 강수량은 남한보다 더 큰 긍정적인 영향을 미치는 것으로 나타났다. 극한기상 지표인 폭염일수, 호우일수, 가뭄더미 변수도 남한보다 북한에 더 큰 부정적인 영향을 미치는 것으로 나타났다.

표 8. 기후변화가 남한과 북한의 쌀 생산성에 미치는 영향 비교(종속변수: log(쌀의 단위))

독립변수 \ 모형	(1)	(2)	(3)	(4)	(5)
평균 기온	0.013* (0.006)	0.012* (0.006)	-0.007 (0.007)	-0.011 (0.009)	-0.010 (0.008)
누적 강수량	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
기온 변동성		0.196** (0.090)		-0.332* (0.174)	-0.346* (0.172)
강수강도		0.002 (0.011)		-0.017 (0.013)	
폭염일수			0.002*** (0.000)	0.003*** (0.001)	0.003*** (0.001)
호우일수			0.004 (0.003)		0.004 (0.003)
가뭄더미			-0.003 (0.009)	-0.006 (0.008)	-0.008 (0.009)
공통추세	0.002*** (0.001)	0.002*** (0.001)	0.002** (0.001)	0.002** (0.001)	0.002** (0.001)
평균 기온*북한더미	0.124*** (0.033)	0.127*** (0.041)	0.140*** (0.033)	0.174*** (0.046)	0.153*** (0.051)
누적 강수량*북한더미	0.002*** (0.001)	0.005*** (0.001)	0.003** (0.001)	0.004** (0.002)	0.003** (0.001)
기온 변동성*북한더미		-0.623 (0.472)		0.255 (0.522)	0.562 (0.596)
강수강도*북한더미		-0.333*** (0.098)		-0.302*** (0.097)	
폭염일수*북한더미			-0.006*** (0.002)	-0.007** (0.003)	-0.007*** (0.002)
호우일수*북한더미			-0.042** (0.018)		-0.042** (0.019)
가뭄더미*북한더미			-0.097** (0.047)	-0.102** (0.043)	-0.094** (0.045)
공통추세*북한더미	0.024*** (0.004)	0.027*** (0.003)	0.028*** (0.004)	0.028*** (0.003)	0.027*** (0.004)
상수항	0.377 (0.234)	0.487 (0.330)	0.637** (0.239)	0.642 (0.379)	0.629 (0.433)
지역 고정효과	YES	YES	YES	YES	YES
Within R ²	0.615	0.647	0.644	0.662	0.645

주1: *, **, ***는 10%, 5%, 1% 유의수준 하에서 회귀계수가 0이라는 귀무가설이 기각됨을 의미함.

주2: 표준오차는 소표본 자유도 조정을 통하여 계산함.

3. 북한의 쌀 생산성과 식생지수 간 상관관계 분석 및 기후변화 영향평가

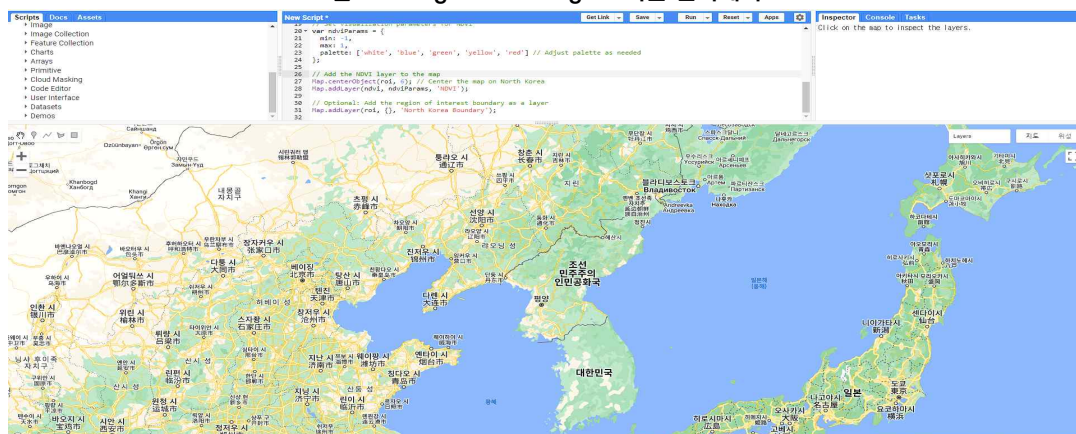
본 장에서는 북한의 쌀 생산성과 가장 상관관계가 높은 식생지수를 선별한 후, 해당 식생지수에 기후변화가 미치는 영향을 평가한다. 본 장에서는 식생지수인 NDVI, NIRvP, OSAVI에 대한 구조적 데이터 지표를 적용하여 다양한 형태의 식생지수와 북한 쌀 생산성 간 상관관계를 분석한다. 이를 통해 식생지수의 설명력과 기후변화 영향평가의 정확도를 제고할 수 있을 것이다.

3.1. 분석자료

3.1.1. 위성 자료를 활용한 식생지수 구축

본 연구에서는 Google Earth Engine을 활용하여 식생지수 데이터를 구축하였다. Google earth engine은 google에서 개발한 클라우드 기반 플랫폼으로, 다양한 위성 자료 및 GIS(Geographic Information System) 데이터를 시각화 및 전처리할 수 있는 기능을 가지고 있다. Google earth engine은 <그림 6>과 같은 플랫폼 내 개발 환경, 또는 API를 활용한 코딩으로 전처리가 일정 부분 완성된 형태의 영상을 제공받을 수 있어 데이터 구축이 편리하다는 장점이 있다. 또한 위성 자료의 필요한 부분만을 선택적으로 내려받을 수 있어, 하드웨어 용량 확보에 대한 부담이 적다는 장점이 있다.

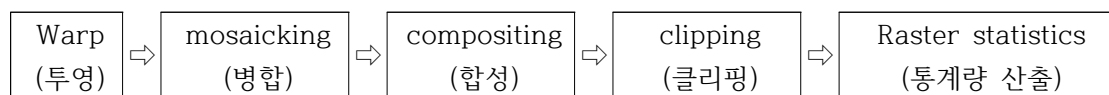
그림 6. Google earth engine 기본 인터페이스



자료: Google earth engine

〈그림 7〉은 위성 자료 다운로드 이후의 식생지수 구축 절차를 나타낸다. 먼저 내려받은 위성 자료를 좌표 변환 기능인 ‘투영’을 통해 원하는 좌표계로 통일한다. 그 후, ‘병합’을 통해 여러 위성 자료를 연결하여 하나의 연속된 이미지로 통합한다. ‘합성’ 단계에서는 동일한 위치의 영상을 겹쳐 특정 기간 또는 특정 조건의 데이터를 구축한다. 본 연구에서는 각 밴드값의 월평균 값을 산출한 것이 ‘합성’에 해당한다. 이후 ‘클리핑’ 단계에서는 위성 자료를 원하는 공간 경계에 따라 잘라내어 관심 지역의 데이터만 추출한다. 본 연구에서는 북한 논 지역을 공간 경계로 설정하여 해당하는 지역만 활용하였다. 일련의 과정을 거친 후, 원하는 공간 구획에 따른 ‘통계량 산출’로 분석에 적합한 형태의 데이터를 구축할 수 있다. 본 연구에서는 Google earth engine을 활용하여 ‘투영’에서 ‘클리핑’까지의 과정을 자동화하였다.

그림 7. 식생지수 데이터 구축 절차



자료: 저자 작성.

앞서 제시한 구축 과정에서 활용된 북한 논 지역 공간 경계는 환경부 대분류 피복지도와 NESEA-Rice10 데이터를 합성하여 구축하였다. 환경부 대분류 피복 지도는 〈표 9〉와 같이 1998년부터 10년 주기로 남한과 북한 전역을 7개의 토지 이용 항목(시가화·건조, 농업, 산림, 초지, 습지, 나지, 수역)으로 분류하고 제공하는 공간정보 DB이다.

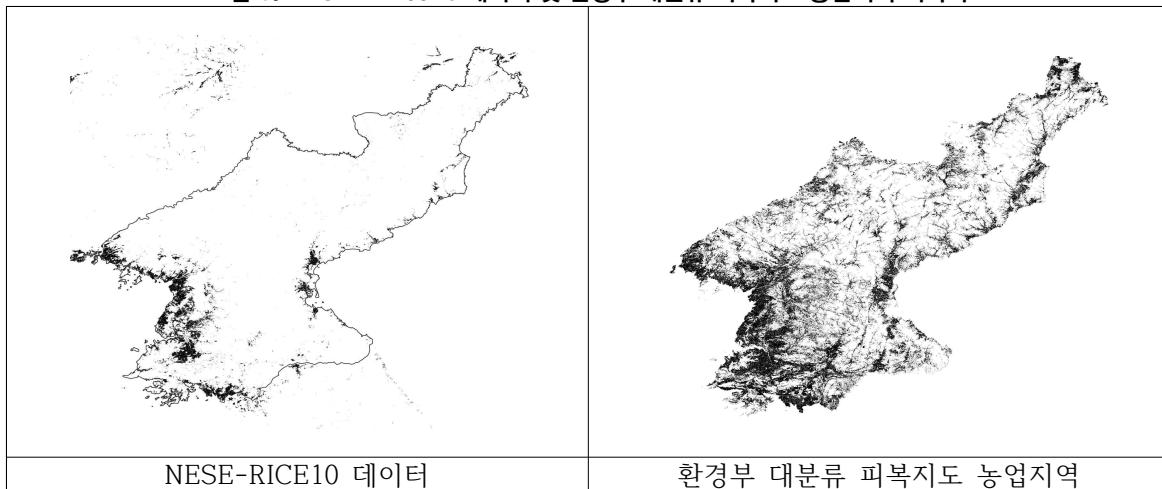
표 9. 환경부 대분류 피복지도 제작 현황

제작 연도	영상 연도	제작 활용 위성 자료	자료형식 및 물량
1998년 2000년	1987 ~ 1989년	Landsat TM(30m)	- 포맷 : GeoTiff, PDF - 해상도 : 30M급 - 물량 : 축척 1/50,000 488도엽 - 분류항목 : 7개 항목
2000년	1997 ~ 1999년	Landsat TM(30m)	- 포맷 : GeoTiff, PDF - 해상도 : 30M급 - 물량 : 축척 1/50,000 488도엽 - 분류항목 : 7개 항목
2010년	2008 ~ 2010년	Landsat 7(30m)	- 포맷 : GeoTiff, PDF - 해상도 : 30M급 - 물량 : 축척 1/50,000 488도엽 - 분류항목 : 7개 항목
2019년	2018 ~ 2019년	- Landsat 8 (30M) - 아리랑 2호 (1M) - 아리랑 3A호(1M)	- 포맷 : GeoTiff, PDF - 해상도 : 30M급 - 물량 : 축척 1/50,000 488도엽 - 분류항목 : 7개 항목

자료: 환경공간정보서비스 홈페이지를 참고하여 저자 작성(<https://egis.me.go.kr/intro/land.do>).

NESEA-Rice10 데이터는 2017년에서 2019년까지 MODIS 및 SENTINEL-1 위성 자료를 활용하여 논 지역을 추정한 공간 해상도 10m급 위성자료이다(Jichong. H. et al., 2021). NESEA- Rice10 데이터와 환경부 대분류 피복 지도를 시각화하면 다음 <그림 8>과 같다.

그림 8. NESEA-Rice10 데이터 및 환경부 대분류 피복지도 농업지역 시각화



자료: NESEA-Rice10 데이터 및 환경부 대분류 피복지도를 활용하여 저자 작성.

NESEA-Rice10의 북한 논 지역 추정 결과는 남한, 일본 등 다른 국가의 추정 결과에 비해 검증 수단이 제한적이기 때문에, 추가적인 검증 과정이 필요하다. 이를 위해, 환경부 대분류 피복 지도의 농업 지역 데이터와의 합성을 진행하여 북한 논 지역 공간 경계에 대한 정확성을 검증하였다.

본 연구에서 분석자료로 활용된 식생지수는 NDVI(Normalized Difference Vegetation Index), NIRvP(Near-infrared Reflectance of Vegetation multiplied by Photosynthetically active radiation), 그리고 OSAVI(Optimized Soil-Adjusted Vegetation Index)이다. 식생지수 구축을 위해 미항공우주국에서 제공하는 MODIS 위성 자료를 활용하였다. 구체적으로, NDVI 밴드값은 MODIS/061/MOD13Q1, NIR 및 RED 밴드값은 MODIS/061/MOD09A1, PAR 밴드값은 MODIS/061/MCD18C2 위성 자료를 활용하였다. 각 위성 자료에 대한 기본적인 제원은 <표 10>과 같다. 서로 다른 위성 자료의 시간 해상도에 대한 통일성을 확보하기 위해, 각 위성 자료에서 제공하는 밴드값의 월평균 값을 산출하여 연구에 활용하였다.

표 10. 분석 활용 위성 자료 기본 제원

위성 자료	시간 해상도 ⁸⁾ (Temporal resolution)	공간 해상도 (Spatial resolution)	용도
-------	---	--------------------------------	----

MODIS/061/MOD13Q1	16 day composite	250m	NDVI 추출
MODIS/061/MOD09A	8 day composite	500m	OSAVI 산출
MODIS/061/MCD18C2	daily	5,600m	NIRvP 산출

자료: USGS 홈페이지를 참고하여 저자 작성.

NDVI는 위성 영상을 활용하여 식생의 밀도와 건강 상태를 정량적으로 나타낼 수 있는 식생지수이다(Rouse et al., 1974). NDVI는 식생의 밀도가 높고 건강할수록 적색 파장(Red)에 대한 흡수량은 증가하고, 근적외선(Near-Infrared, NIR)에 대한 흡수량은 감소하는 원리를 활용한다. 일반적으로 작황 분석을 위해 가장 많이 활용되는 식생지수이기 때문에, 본 연구에서도 NDVI를 활용하였다. NDVI의 산출식은 다음과 같다.

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (2)$$

NIRvP는 앞서 언급한 NDVI에 근적외선 반사율과 태양 복사량(Photosynthetically Active Radiation)을 곱해 산출하는 개선된 형태의 식생지수이다(Dechant. B. et al., 2022). NIRvP는 식생의 광합성을 정량화할 수 있는데, 이를 통한 북한지역 작황 분석 정확도 개선 가능성을 확인하기 위해 본 연구에서 활용하였다. NIRvP의 산출식은 다음과 같다.

$$NIRvP = NDVI * NIR * PAR \quad (3)$$

OSAVI는 식생지수 산출 시 토양 반사율의 영향을 고려하기 위해 개발된 지표이다. 기존 NDVI에서 활용하는 적색 파장과 근적외선 이외에 토양 보정을 위한 상수(0.16)를 추가한 형태를 가지고 있다. Rachel R. et al.(2018)에 따르면 OSAVI는 토양 반사의 영향이 큰 건조 지역에 대한 식생 파악의 정확도를 개선할 수 있다. OSAVI는 관수 시설이 미흡한 북한 농경지에 대한 분석 정확도 개선을 도모하기 위해 본 연구에 활용되었다. OSAVI의 산출식은 다음과 같다.

$$OSAVI = \frac{NIR - RED}{NIR + RED + 0.16} \quad (4)$$

8) composite는 정해진 기간에 확보된 데이터를 합성하여 하나의 최종 결과물을 산출하는 것을 의미함.

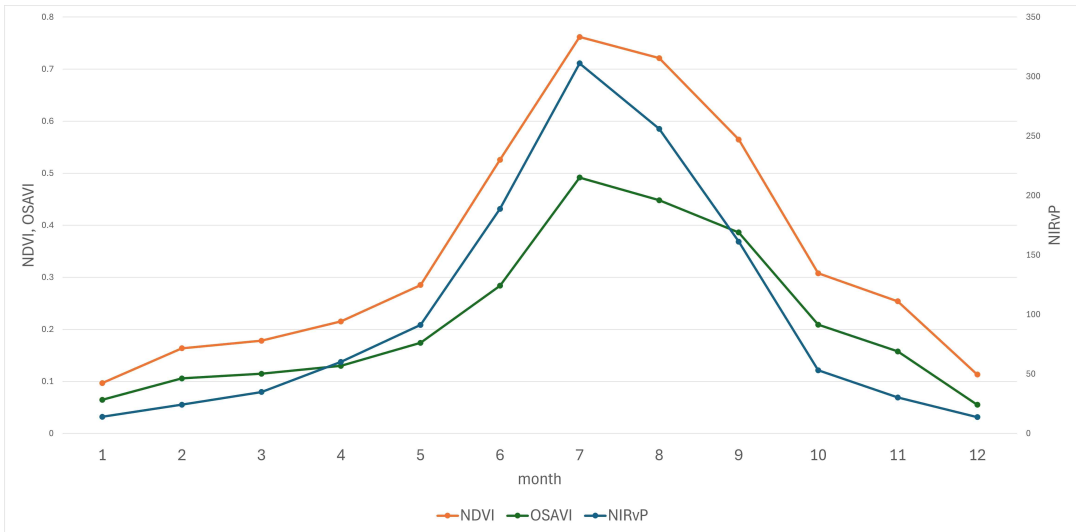
3.1.2. 북한 논 지역의 식생지수

앞 절에서 제시한 세 가지 식생지수 NDVI, NIRvP 그리고 OSAVI의 기초 통계량은 <표 11>과 같고, 2023년 월 평균값 추세는 다음 <그림 8>와 같다. <그림 8>에서 세 식생지수 모두 파종이 시작되는 4월부터 7월까지 지속적으로 상승하다가 이후 감소하는 패턴을 보인다. FAO에서 제시하는 북한 쌀 수확 시기인 9월보다 이른 7월부터 식생지수 값이 감소하는 현상에 대해서는 다양한 해석이 가능하다. 가령 벼가 등숙기에 접어들어 잎과 줄기의 엽록소가 감소하면 NDVI의 수치가 감소할 수 있다. 혹은 자연재해나 극한기후로 인해 작황이 나빠진 것이 원인이 될 수 있다.

표 11. NDVI, NIRvP, OSAVI의 기초 통계량

변수명	평균	표준편차	최솟값	최댓값
NDVI	0.334	0.214	0.008	0.802
NIRvP	92.423	86.378	1.344	307.182
OSAVI	0.208	0.139	0.006	0.569

그림 8. NDVI, NIRvP, OSAVI 월 평균값 추세(2023년)



자료: MODIS/061/MOD13Q1, MODIS/061/MOD09A, MODIS/061/MCD18C2 위성 자료를 활용하여 저자 작성,

이처럼 작황 관련 상황을 명확히 규명할 수 없는 상태에서 농업생산 분석의 정확도를 높이기 위해서는 식생지수의 성능을 다각도에서 검토하는 과정이 필요하다. Michael Wilton(2021)은 식생지수에 다양한 메트릭을 적용하여 그 설명력을 제고한 바 있다. 본

연구 또한 이러한 접근 방식을 차용하여 식생지수의 성능 평가 및 향상을 위해 다양한 형태의 메트릭을 적용하여 비교 분석하였다. 구체적으로, 본 연구에서는 각 식생지수의 평균값(Tian et al., 2016), 누적값 (Reed et al., 1996), 차분 값(Jinglian Tian et al., 2024) 그리고 비율 값(Michael Wilton, 2021)을 활용하였다.

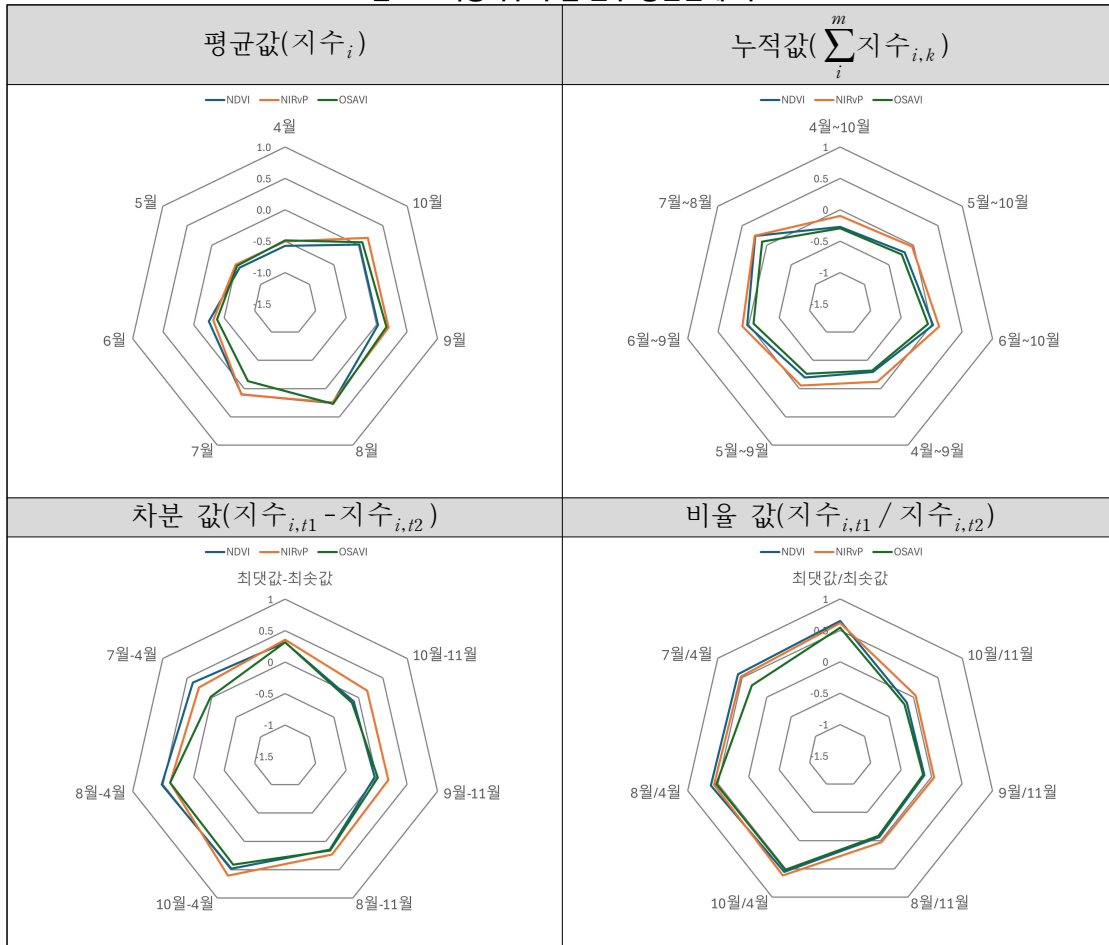
식생지수의 평균값은 특정시점에서의 생육환경을 가장 잘 나타낼 수 있는 형태의 메트릭이다. 누적값의 경우, 평균값보다 장기적인 생육 상태 반영이 가능한 형태의 메트릭이다. 차분 값과 비율 값은 모두 특정 시점의 차이를 나타내며 생육상태의 변화정도를 나타내고, 차분 값은 절대적인 차이를 반영하고, 비율 값은 상대적 차이를 반영한다는 점에서 차이가 있다. 본 분석에서는 파종이 시작되지 않은 1, 2, 3월과 수확이 완전히 종료된 시점인 12월은 분석 대상에서 제외하였다.

3.2. 분석결과

3.2.1. 식생지수와 북한의 쌀 생산량과의 상관관계 분석

NDVI, NIRvP, OSAVI의 메트릭 비교분석 결과는 다음 <그림 10>과 같다. 평균값은 7월부터 유의미한 상관관계가 나타나기 시작했으며, 평균적으로 ‘8월’ 평균값(NDVI: 0.3, NIRvP: 0.2, OSAVI: 0.3)이 쌀 단수와 가장 상관관계가 높은 것으로 나타났다. 이는 8월이 수확을 앞두고 논벼의 생장이 가장 뚜렷하게 식별되는 시기이기 때문으로 보인다. 누적값 비교분석에서는 ‘7월~8월’ 누적값(NDVI: 0.23, NIRvP: 0.24, OSAVI: 0.09)이 평균적으로 쌀 단수와 가장 높은 상관관계를 가진 것으로 나타났다. 이 또한 논벼 생장의 식별 정도에 따른 결과인 것으로 사료된다. 차분 값 비교분석의 경우, ‘10월-4월’ 값(NDVI: 0.48, NIRvP: 0.60, OSAVI: 0.41)에서 평균적으로 가장 높은 상관관계가 식별되었는데, 이는 파종기와 수확기의 식생지수 차이가 쌀 단수와 상관관계가 있음을 시사한다. 전체 비교 결과 중 가장 높은 상관관계는 ‘최댓값/최솟값’ 값(NDVI: 0.65, NIRvP: 0.61, OSAVI: 0.54)인 것으로 확인되었다. 이는 연중 식생지수 최댓값과 최솟값 간의 비율이, 그 차분 값보다 쌀 단수와의 상관성을 더 명확히 설명할 수 있음을 보여준다.

그림 10. 식생지수와 쌀 단수 상관관계 비교



자료: 저자 작성.

3.2.2. 식생지수를 활용한 기후변화가 북한의 쌀 생산량에 미치는 효과 분석

본 절에서는 북한의 쌀 생산량과 상관관계를 보이는 식생지수를 종속변수로 설정하고, 2장에서 수행한 동일한 분석을 적용하였다. 분석 결과, 기온과 관련해서는 2장 <표 6>의 결과와 상이하였으나, 강수량 변수와 관련된 결과는 유사한 경향을 보였다. 구체적으로 평균 기온은 쌀 생산성에 유의한 영향을 미치지 않는 것으로 나타났고, 반대로 기온의 변동성이 높을수록 쌀 생산성이 증가하는 경향이 나타났다. 누적 강수량 변수는 쌀 생산에 긍정적인 영향을 미치는 것으로 나타났으며, 강수강도가 강하거나 호우일수가 많아지면 쌀 생산성에 부정적인 영향을 미치는 것으로 나타났다.

이와 같은 결과는 식생지수를 종속변수로 설정함으로써 북한의 제한적인 농업생산 관측치를 확장하려는 본 연구의 논리적 타당성이 충분히 확보되지 못했을 가능성을 시사한다.

이를 보완하기 위해 북한지역을 서부 저지대, 동부 해안지대, 북부 고원지대 등으로 농업 기후대를 구분하고 쌀 생육주기를 반영하거나, 기상 관측 지점의 일정 반경 이내의 논 지역에서 산출된 식생지수만 활용하는 방안을 고려할 수 있을 것이다. 추후 연구에서는 이를 보완하여 북한의 쌀 단수를 종속변수로 한 결과와 식생지수를 종속변수로 한 분석 간의 일관적인 결과를 도출하고, 이를 확장하여 북한의 농업을 대상으로 시간적 해상도를 향상시키거나 생산의 변동성을 반영하는 등 다양한 방법론을 적용함으로써 심층적인 분석 결과를 도출할 수 있을 것으로 기대한다.

표 12. 기후변화가 북한의 쌀 생산에 미치는 영향(종속변수: log(식생지수))

독립변수 \ 모형	(1)	(2)
평균 기온	0.008 (0.009)	0.003 (0.009)
누적 강수량	0.001*** (0.003)	0.001* (0.001)
기온 변동성	0.540** (0.184)	0.586*** (0.199)
강수강도	-0.007*** (0.001)	
폭염일수	0.000 (0.001)	-0.004 (0.001)
호우일수		-0.012** (0.005)
가뭄더미	-0.009 (0.006)	0.005 (0.007)
공통추세	0.005*** (0.002)	0.006*** (0.002)
상수항	0.932*** (0.176)	0.947*** (0.183)
지역 고정효과	YES	YES
Within R ²	0.192	0.163

주1: *, **, ***는 10%, 5%, 1% 유의수준 하에서 회귀계수가 0이라는 귀무가설이 기각됨을 의미함.

주2: 소표본 자유도가 조정된 클러스터 표준오차를 활용함.

4. 위성 및 기상자료를 활용한 북한의 쌀 생산량 예측

본 장에서는 다양한 기후변수와 앞선 분석에서 산출된 다양한 식생지수들을 활용하여 북한 쌀 생산량을 예측한다. 성격이 다른 세 가지 정규화 회귀분석 기법을 활용하여 북한 쌀 생산량 예측 정확도를 개선할 수 있을 것이다.

4.1. 분석자료

북한 쌀 생산량을 예측하는 모형은 다음과 같다.

$$\begin{aligned}
 Prod_{i,t} = & \beta_0 + \sum_{k=4}^{10} \beta_1 temp_{i,k,t} + \sum_{k=4}^{10} \beta_2 temp^2_{i,k,t} + \sum_{k=4}^{10} \beta_{3,k} temp-max_{i,k,t} + \sum_{k=4}^{10} \beta_{4,k} temp-min_{i,k,t} \\
 & + \sum_{k=4}^{10} \beta_{5,k} prec_{i,k,t} + \sum_{k=4}^{10} \beta_{6,k} prec^2_{i,k,t} + \sum_{k=4}^{10} \beta_{7,k} prec-max_{i,k,t} + \sum_{k=4}^{10} \beta_{8,k} extreme-hot_{i,k,t} \\
 & + \sum_{k=4}^{10} \beta_{9,k} extreme-rain_{i,k,t} + \sum_{k=4}^{10} \beta_{10,k} frost_{i,k,t} + \sum_{k=4}^{10} \beta_{11,k} m8-VIS_{i,k,t} + \sum_{k=4}^{10} \beta_{12,k} ratio-VIS_{i,k,t} \\
 & + \sum_{k=4}^{10} \beta_{13,k} diff-8-VIS_{i,k,t} + \sum_{k=4}^{10} \beta_{14,k} cumm-VIS_{i,k,t} + \sum_{k=4}^{10} \beta_{15,k} SPEI-1_{i,k,t} + \sum_{k=4}^{10} \beta_{16,k} SPEI-3_{i,k,t} \\
 & + \sum_{k=4}^{10} \beta_{17,k} SPEI-6_{i,k,t} + \sum_{i=1}^{11} \gamma_i + \delta Time_t + \delta Time_t^2 + \epsilon_t
 \end{aligned}$$

첨자 i, k, t는 각각 지역, 월, 연도를 나타낸다. $temp$ 는 월평균기온, $temp^2$ 는 월평균의 제곱, $temp_{max}$ 는 월 최고 기온, $temp_{min}$ 은 월 최저 기온, $prec$ 은 월 누적 강수량, $prec^2$ 는 월 누적 강수량의 제곱, $prec-max$ 는 월 최대 강수량, $extreme-hot$ 은 폭염일수, $extreme-rain$ 은 폭우 일수, $frost$ 는 서리일수를 나타낸다. VIS 는 식생지수 벡터로 NDVI, NIRvP, OSAVI를 포함한다. $m8-VIS$, $ratio-VIS$, $diff-VIS$, $cumm-VIS$ 은 각각 평균, 비율, 차분, 누적 메트릭에서 가장 상관관계가 높았던 조합을 나타낸다. $SPEI-1$ 는 월별 가뭄지수, $SPEI-3$ 은 3개월 누적 가뭄지수, $SPEI-6$ 은 6개월 누적 가뭄지수를 나타낸다. γ 는 지역 고정 효과를 나타내며, $Time$ 은 추세항, $Time^2$ 는 추세항의 제곱을 의미한다.

본 연구에 활용된 기후 변수는 모두 4월부터 10월까지의 월별 수치를 활용하였다⁹⁾. 또한 평균기온과 월 누적 강수량의 제곱 항을 변수로 추가한 것은, 기온과 강수가 북한 쌀 생

9) FAO의 crop calender을 기반으로 변수의 시계열 활용 구간을 결정하였다.

산에 미치는 비선형적 영향을 반영하기 위함이다.

상기 서술한 기상변수에 대한 기초 통계량은 , <표 13>와 같다.

표 13. 예측에 활용된 변수들의 기초 통계량

변수명	시기	평균	표준편차	최솟값	최댓값
평균 기온	4월	9.235	2.525	0.824	13.201
	5월	15.307	2.182	9.616	18.874
	6월	19.804	2.240	13.942	23.058
	7월	22.718	2.107	14.691	26.229
	8월	23.035	2.272	16.126	27.256
	9월	18.39022	2.392493	11.57013	22.77426
	10월	10.948	2.781	2.334	15.411
월 최고 기온	4월	23.632	3.562	13.824	30.669
	5월	28.357	2.149	19.566	32.769
	6월	30.817	2.114	23.177	34.673
	7월	31.894	2.193	25.447	36.904
	8월	32.375	2.158	26.285	38.166
	9월	28.475	1.984	22.154	32.679
	10월	24.114	2.194	16.826	29.661
월 최저 기온	4월	-2.310	3.475	-15.881	3.297
	5월	3.841	3.050	-3.519	9.916
	6월	7.741	5.047	0.000	15.751
	7월	12.080	6.586	0.000	21.010
	8월	10.511	6.356	0.000	20.952
	9월	6.538	4.200	-5.085	14.730
	10월	-1.827	3.591	-12.979	5.552
월 누적 강수량	4월	50.061	26.074	9.170	191.264
	5월	79.390	50.265	2.740	276.640
	6월	91.497	52.808	9.021	369.993
	7월	295.705	163.455	28.080	933.743
	8월	249.087	166.451	30.634	902.043
	9월	100.101	77.710	6.668	482.802
	10월	48.858	42.054	2.045	227.088

폭염일수	4월	0	0	0	0
	5월	0	0	0	0
	6월	0.142	0.423	0	3
	7월	1.431	2.981	0	14
	8월	1.579	3.113	0	17
	9월	0	0	0	0
	10월	0	0	0	0
호우일수	4월	0.005	0.075	0	1
	5월	0.039	0.195	0	1
	6월	0.068	0.294	0	2
	7월	0.517	0.868	0	5
	8월	0.460	0.893	0	5
	9월	0.102	0.339	0	2
	10월	0.022	0.149	0	1
SPEI_1	4월	-0.022	0.759	-1.981	1.258
	5월	0.190	1.157	-2.483	2.294
	6월	-0.060	1.122	-2.179	3.280
	7월	0.054	1.064	-2.451	2.761
	8월	0.103	0.914	-1.912	2.408
	9월	-0.131	1.094	-2.359	2.461
	10월	0.032	1.146	-1.953	3.486
NDVI 최댓값/최솟값		3.744	0.470	2.557	4.913
NIRvP 최댓값/최솟값		5.233	0.746	3.376	6.956
OSAVI 최댓값/최솟값		4.246	0.564	2.599	5.399
NDVI 7, 8월 누적값		1.437	0.056	1.308	1.580
NIRvP 7, 8월 누적값		512.031	35.625	431.350	609.028
OSAVI 7, 8월 누적값		0.891	0.049	0.762	0.993
NDVI 최댓값-최솟값		0.538	0.036	0.434	0.617
NIRvP 최댓값-최솟값		214.799	18.125	169.597	257.497
OSAVI 최댓값-최솟값		0.376	0.039	0.239	0.461
8월 NDVI 평균값		0.728	0.029	0.604	0.799
8월 NIRvP 평균값		260.766	20.288	196.620	307.182
8월 OSAVI 평균값		0.495	0.039	0.363	0.569

4.2. 분석방법

회귀분석을 통한 계량모형의 목적은 모형의 설명과 예측, 이 두 가지로 구분할 수 있다. 본 연구의 2장과 3장에서는 기후변수와 농업 생산성의 관계를 설명하고자 하였다면 본 절에서는 예측(설명)변수를 활용하여 북한의 농업생산량을 예측하고자 한다. 예측에서 관측치의 개수가 설명변수의 개수보다 월등히 많지 않다면 추정값에 변동이 존재하고 과적합으로 인하여, 학습에 사용되지 않은 데이터를 예측할 때 성능이 좋지 않게 된다. 또한, 관측치의 수가 설명변수의 수보다 작다면 최소제곱 추정치가 유일하게 결정되지 않는다.

기후변수를 활용하여 농업의 생산량을 예측하고자 할 때 영향을 줄 수 있는 기후변수는 수없이 다양하며, 이러한 변수들을 활용하여 만들 수 있는 함수의 개수는 무수히 많게 된다. 이러듯 설명변수의 개수(p)들이 데이터의 개수(n)보다 많은 데이터인 고차원데이터(high-dimensional data)를 분석하기 위해서는 이에 적합한 고차원 회귀분석 방법이 활용되어야 한다. 이러한 방법 중 가장 대표적인 방법은 식 (2)와 같이 표현되는 벌점화 회귀분석(penalized regression analysis)으로 Ridge회귀, Lasso모형, Elastic-net모형 등이 있다.

$$\underset{\beta}{Min} \frac{1}{n} \sum_{i=1}^n (y_i - \sum_{j=1}^p \beta_j x_{ij})^2 + \sum_{j=1}^p J_{\lambda}(\beta_j) \quad (2)$$

$J_{\lambda}(\beta_j)$ 는 벌점함수로 회귀계수에 대한 특정 함수(vector norm)로 정의된다. λ 는 벌점에 대한 축소 정도를 조정하는 튜닝 매개변수다. λ 가 0이라면 벌점이 없는 최소제곱모형과 동일한 추정치를 얻게되고¹⁰⁾, λ 가 ∞ 라면 절편을 제외한 모든 계수들이 0으로 추정된다. λ 가 0과 ∞ 사이라면 본래의 최소제곱 추정량에서 0의 방향으로 계수의 크기를 수축(shrinkage)하게 된다.

Ridge회귀는 Hoerl and Kennard(1970)에 의해 제안되었으며 능형회귀라고도 불린다. 예측의 안정성과 정확성 측면에서 좋은 성능을 갖는 것으로 알려져 있으며 벌점함수가 $J_{\lambda}(\beta_j) = \lambda \beta_j^2$ 로 정의된다.

Lasso(Least Absolute Shrinkage and Selection Operator) 모형은

¹⁰⁾ 하지만 p 가 n 보다 크게 되면 최소제곱추정량이 유일하게 결정되지 않게 된다.

Tibshirani(1996)에 의해 제안되었으며, 벌점함수는 $J_\lambda(\beta_j) = \lambda|\beta_j|$ 로 정의된다. 모형의 이름처럼 계수를 축소(shrinkage)하는 기능을 하면서 추가적으로 변수를 선택(selection)하는 역할까지 하게 된다. Lasso 추정량은 기준점(thresholding value) 이하 일 때는 그 값이 0으로 대체되고, 기준점 이상의 추정량은 0의 방향으로 크기를 수축하게 된다.

Elastic-net 모형은 Zou and Hastie(2005)에 의해 제안되었으며 Ridge회귀모형과 Lasso모형의 벌점함수를 블록결합한 형태의 벌점함수

$$J_\lambda(\beta_j) = \lambda \sum_{j=1}^p [(1-\alpha)|\beta_j| + \alpha \beta_j^2], \alpha \in [0, 1] \text{를 갖도록로 정의된다. 일반적으로 Lasso모}$$

형은 상관관계가 있는 유의미한 변수들은 그 중 하나의 변수만 선택하는 성질을 가지고 있으나, Elastic-net 모형은 상관관계가 있는 모든 변수들을 선택하게 된다.

예측모형의 성능평가를 위해서는 모형 적합에 사용되지 않은 새로운 데이터를 활용하여, 실제값과 예측값의 차이를 나타내는 예측오차의 크기를 측정해야 한다. 이를 위해 학습자료와는 독립적인 평가자료를 통해 예측오차를 계산하며, 일반적으로 식 (3)과 같은 평균제곱오차(mean squared error)를 활용하여 계산한다.

$$\widehat{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{f}(x))^2, \text{ n은 관측치 수} \quad (3)$$

데이터를 모형을 적합하기 위한 학습자료와 성능 평가를 위한 평가자료로 전체자료를 분할해서 평균제곱오차를 추정하게 되지만, 본 연구에서는 관측치 수가 적은 문제로 인해 이를 분할해서 활용하면 학습자료의 크기가 작아져 모형의 적합도가 좋지 않게 되는 문제가 있다. 이러한 점을 보완하기 위하여 자료의 분할과 오차추정을 반복수행하여 안정적인 평균제곱오차를 추정하는 절차인 교차검증(Cross Validation)방법은 식 (4)와 같다. 교차검증 방법 중 전체자료를 K 등분하여 K번 평균제곱오차를 추정하는 K-중첩(fold) 교차검증을 활용하여 예측모형의 성능평가를 하였다. 중첩 횟수(K)의 경우 5-중첩이나 10-중첩을 활용하면 높은 편의와 분산으로 인한 문제 없이 평균제곱오차에 대한 적절한 추정값을 얻을 수 있는 것으로 알려져있어(James et al., 2013), 본 연구에서는 10-중첩 교차검증을 수행하였다.

$$CV_K = \frac{1}{K} \sum_{i=1}^K \widehat{MSE} \quad (4)$$

4.3. 분석결과

본 연구에서 추정한 Lasso, Ridge 그리고 Elastic-net 모형에 대한 검증 적합도 통계량은 다음 <표 x>와 같다. 각 모형의 적합도 판단을 위해 평균제곱오차와 결정계수를 활용하였다.

검증 결과, Lasso 모형의 예측력($R^2 = 0.957$)이 가장 높은 것으로 나타났다. Elastic-net 모형은 벌점 조정 변수인 α 가 1로 완전히 수렴하여 Lasso 모형과 동일한 모형이 도출되었다. 이러한 결과는 본 분석에 활용된 데이터를 통한 예측에는 중요도가 낮은 변수들을 모형에서 제거하는 것이 더 적합할 수 있음을 보여준다. 이는 Lasso모형보다 낮은 Ridge 모형의 결과를 통해서도 확인할 수 있다. Ridge 모형은 학습 단계에서는 다른 두 모형보다 높은 적합도($R^2 = 0.983$)를 보였지만, 검증 단계에서는 더 낮은 예측력($R^2 = 0.951$)을 가진 것으로 나타났다. 이는 중요도가 낮은 변수를 완벽히 제거하지 않는 Ridge 모형의 특성이 과적합을 유발한 것으로 볼 수 있다.

표 14. 예측 모형별 학습 및 검증 적합도 통계량

	표본	MSE	R2	표본 수
Lasso	검증 표본	0.066	0.957	27
	학습 표본	0.046	0.977	114
Ridge	검증 표본	0.076	0.951	27
	학습 표본	0.035	0.983	114
Elastic-net	검증 표본	0.066	0.957	27
	학습 표본	0.046	0.977	114

주: 종속변수인 쌀 생산량은 자연로그 변환하여 사용함.

학습 과정에서 Lasso 모형에 선택된 변수는 다음 <표 15>과 같다.

표15. Lasso 모형에서 선택된 변수 순위

순위	변수명
1	양강도
2	NDVI 최댓값-최솟값
3	황해남도
4	평안북도
5	평안남도
6	자강도
7	OSAVI 최댓값-최솟값
8	함경남도
9	황해북도
10	평양
11	4월 SPEI-6
12	9월 SPEI-6
13	10월 SPEI-1
14	10월 평균 기온
15	5월 호우일수
16	9월 서리일수
17	7월 최저 기온
18	7월 평균 기온
19	8월 호우일수
20	NIRvP 최댓값-최솟값
21	9월 최고 기온
22	NIRvP 최댓값/최솟값
23	5월 최저 기온
24	7월 누적 강수량

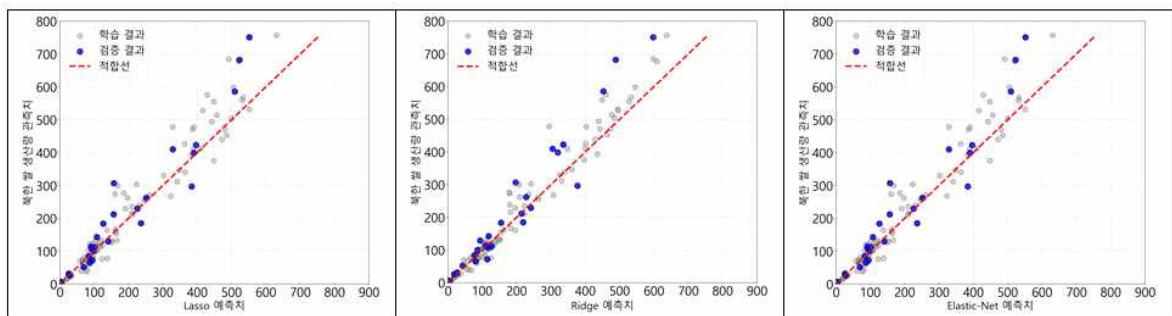
주1: Ridge 모형은 변수를 선택하지 않으므로 생략함.

주2: 순위를 책정함에 있어 상수항은 결과에 반영하지 않음.

상기 순위를 바탕으로 분석한 결과, 지역을 구분하는 변수들이 예측 모형에서 높은 중요도를 가지는 것으로 확인되었다. 이는 북한 작황 예측에서 지역적 특성이 중요한 요인으로 작용한다는 점을 시사한다. 또한, 다양한 식생지수 메트릭이 상대적으로 높은 순위를 차지한 것으로 나타났는데, 이는 메트릭의 활용이 북한 작황 예측 정확도를 높이는 데 중요한 기여를 한다는 점을 의미한다. 아울러, 가뭄지수, 호우일수와 같은 극한기후 관련 변수들이 선택된 점을 고려할 때, 이러한 기후 요인들이 북한 작황에 상당한 영향을 미치는 것으로 분석된다. 이는 극한기후가 작물의 생산성과 생존율에 미치는 영향을 설명하며, 향후 예측 모형 개선 시에도 주요 고려 요인으로 포함될 필요가 있음을 시사한다. 이러한 분석 결과를 바탕으로 구축된 세 가지 예측 모형의 2002년부터 2020년

까지의 예측값을 산포도로 시각화한 결과는 다음 <그림 11>에 제시되어 있다.

그림11. Lasso, Ridge, Elastic-net 예측치 산포도



자료: 저자 작성.

한편, 2020년 이후의 북한의 시도별 작황은 공표되지 않고 있다. 다만, 연도별 쌀 생산량 추정은 농촌진흥청 및 USDA(2024)에 의해 제공되고 있다. 따라서 본 연구에서는 2021년, 2022년, 2023년의 농촌진흥청 및 USDA 자료와 본 연구에서 가장 높은 성능을 보인 Lasso 모형의 예측 결과를 비교하였다. 비교 결과는 다음 <표 16>과 같다.

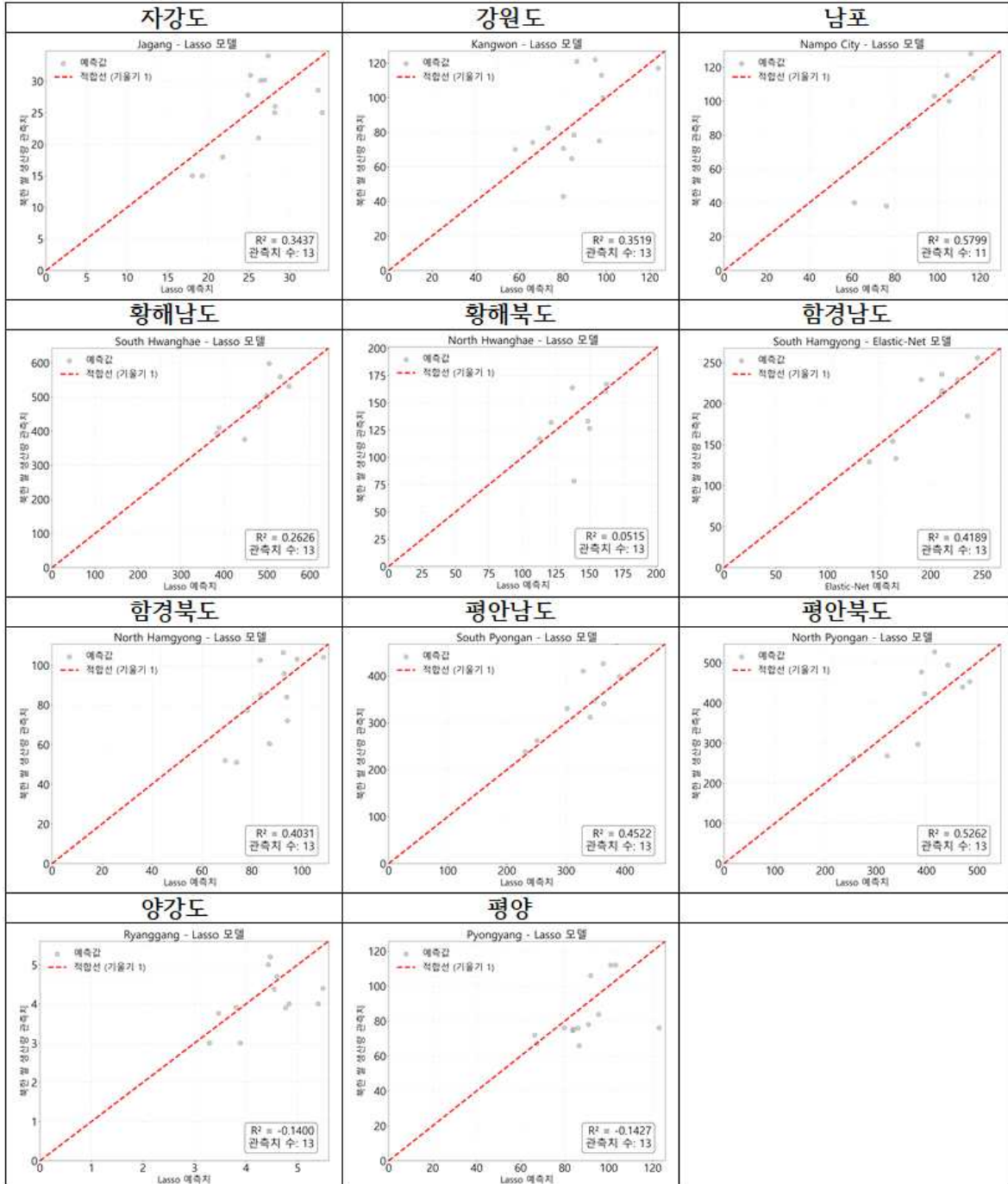
표 16. Lasso 모형 북한 쌀 생산량 예측치 (단위: ton)

	2021	2022	2023
본 연구(Lasso)	1,914,820	2,169,980	2,361,540
농촌진흥청 추정량	2,156,000	2,074,000	2,110,000
USDA 추정량	2,090,000	2,090,000	2,261,380

자료: 농촌진흥청, USDA 자료를 활용하여 저자 작성 참고하여 작성

또한 성능이 가장 높은 것으로 확인된 Lasso 모형의 지역별 예측값 분포도는 다음 <그림 12>와 같다. 지역별 예측치를 비교한 결과, 남포 지역의 예측 결과가 가장 높은 성능을 보이는 것으로 나타났다. 이는 남포 지역의 기상 관측 데이터가 상대적으로 안정적이고 일관된 패턴을 보였기 때문일 가능성이 있다. 반면, 양강도 지역의 예측 성능은 다른 지역에 비해 현저히 떨어지는 것으로 분석되었다. 양강도 지역의 낮은 예측 성능은 해당 지역의 기상 데이터가 상대적으로 부족하거나 변동성이 크기 때문일 가능성이 있다. 이러한 결과는 향후 데이터가 더 긴 기간 동안 축적되고, 추가적인 변수들이 모형에 포함된다면, 모형의 전반적인 성능 향상과 함께 양강도 지역을 포함한 지역별 예측력 또한 더욱 높아질 가능성을 시사한다. 이를 통해 지역적 특성을 보다 정밀하게 반영하는 개선된 예측 모델을 구축할 수 있을 것으로 기대된다.

그림 12. Lasso 모형 지역별 예측치 산포도



자료: 저자 작성.

5. 요약 및 시사점

본 연구에서는 가속화되는 기후변화에 북한 농업생산이 받는 영향을 다각적으로 분석하기 위해 위성 및 기상자료를 활용하여 분석을 진행하였다. 이를 통해 북한의 농업생산에 기후변화가 미치는 영향을 분석하였고 북한의 농업 현황을 간접적으로 살펴볼 수 있는 식생지수를 선정하였다. 추가적으로 북한의 농업생산량을 예측할 수 있는 예측모형을 구축하였다.

첫 번째 분석에서는 기후변화가 북한 농업생산에 미치는 영향을 남한과 비교하였다. 기후 데이터 구축을 위해 남한과 북한의 기상 관측 지점 자료를 활용하였으며, 이 과정에서 티센망 분석을 통해 남북한 각 시도에 대한 기상 관측 지점의 가중치를 산출하여 기후 데이터 구축의 정확성을 제고하였다. 또한, 지리정보의 형태로 제공되는 SPEI를 가공하여 본 연구에 맞게 활용하였다. 패널 고정 모형 분석 결과, 평균기온과 누적 강수량 상승은 남한보다 북한의 농업생산 증가에 더 큰 영향을 미치는 것으로 나타났다. 또한, 극한 기상 발생에 따른 생산량 감소의 폭은 북한이 더 큰 것으로 파악되었다. 이러한 결과는 북한 농업이 남한보다 기후요소에 더 민감하게 반응하며, 특히 극한기후에 대한 적응 능력이 부족한 것으로 해석할 수 있다. 이는 추후 국제협력 시에 기후변화 적응력을 강화하기 위한 체계적인 지원과 대응 전략 마련이 필요함을 시사한다.

두 번째 분석에서는 먼저 위성 자료를 기반으로 구축한 식생지수와 북한 쌀 단수와의 상관관계를 비교 분석하였다. 식생지수 NDVI, NIRvP, OSAVI를 기반으로 구축된 다양한 메트릭과 쌀 단수의 상관관계를 산출하여 가장 상관관계가 높은 메트릭을 선정하고자 하였다. 분석결과, NDVI의 변화 비율이 북한의 쌀 단수와 가장 상관관계가 높은 식생지수로 선별되었다. 이렇게 선정된 식생지수를 기준으로 기후변화의 영향을 살펴보았는데 첫 번째 분석과는 다소 상이한 결과를 나타내었다. 이를 보완하여 추후 연구에서는 기후변화가 북한의 농업생산에 미치는 영향을 심층적으로 분석하고자 한다. 이는 식생지수 메트릭을 활용함으로써 식생지수의 활용 범위를 확장하고자 하였다는 데에 연구적 의의가 있다.

세 번째 분석에서는 앞선 두 분석에서 활용한 다양한 위성 자료와 기후 자료를 활용하여 북한 쌀 생산량 예측 모형을 구축하였다. 고차원 특성을 가진 기후 자료의 특성을 반영하기 위해 벌점화 회귀분석인 Lasso, Ridge, Elastic-net 모형을 활용하였다. 예측의 정확

도를 높이기 위해 평균기온, 강수량, 가뭄지수, 폭염 일수 등 다양한 기후 자료와 더불어, 두 번째 연구에서 쌀 생산성과 가장 상관관계가 높은 것으로 나타난 식생지수 메트릭 조합들을 예측 모형에 적용하였다. 모형의 검증 결과, Lasso 모형이 가장 높은 예측력을 가진 것으로 나타났다. 본 예측은 외부에서 관측할 수 있는 기후 자료 및 위성 자료를 활용하여, 관측이 어려운 북한 쌀 생산량 예측을 시도했다는 점에서 의의가 있다.

종합적으로, 본 연구는 북한 농업생산이 남한보다 기후변화에 더 민감하고, 극한기후에 더욱 취약하다는 점을 실증적으로 입증하였다. 더불어, 위성 자료와 기상 데이터를 활용한 분석 및 예측 모형 구축의 가능성을 제시함으로써, 북한 농업의 특성과 기후변화 간 상호 작용에 대한 이해를 확장하였다. 이러한 연구 결과는 기후변화 대응책 마련에 기여할 수 있는 기반을 제공하며, 남북 농업 협력 정책의 수립과 북한의 기후변화 적응 전략 개발에 중요한 기초 자료로 활용될 수 있다. 나아가, 이 결과는 북한의 식량 안보 문제를 완화하고, 한반도의 지속 가능한 농업 환경 조성을 촉진하는 데 기여할 수 있을 것이다.

참고문헌

- Bell, R. M., & McCaffrey, D. F. (2002). Bias reduction in standard errors for linear regression with multi-stage samples. *Survey Methodology*, 28(2), 169-182.
- Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The review of economics and statistics*, 90(3), 414-427.
- Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: applications to nonorthogonal problems. *Technometrics*, 12(1), 69-82.
- James, G. (2013). *An introduction to statistical learning*.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1), 267-288.
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2), 301-320.
- Choi., S. J. Cho. (2024). Climate Variability and Agricultural Outputs in North Korea: An Econometric Analysis of Rice and Corn Yields, *Journal of Agricultural, Life and Environmental Sciences*, 36(1), 23-26.
- Dechant, R. Youngryel., B. Grayson., K. Philipp., R. Uwe., M. Mirco., Z. Yongguang., T. Giulia., G. Kaiyu. , R. Micol., G. Yves., Z. Yelu., F.Christian., A. B. Joseph. (2022). NIRVP: A robust structural proxy for sun-induced chlorophyll fluorescence and photosynthesis across scales. *Remote Sensing of Environment*, 268, 1-20.
- Global drought monitor,
<https://spei.csic.es/map/maps.html#months=2#month=8#year=2023>
- Hong, Min, B. K., Lee, J. M., Kim, Y., Lee, K. (2012). Estimation of Paddy Field

- Area in North Korea Using RapidEye Images. Korean Journal of Soil Science and Fertilizer. Korean Society of Soil Science and Fertilizer, 45(6), 1194-1202.
- IPCC. (2018). special report global warming of 1.5°C.
- IPCC. (2023). Climate change 2023: Synthesis Report.
- Jeong, J. Ko., J. M. Yeom. (2022). Predicting rice yield at pixel scale through synthetic use of crop and deep learning models with satellite data in South and North Korea, Science of The Total Environment, 802.
- Jichong. H., Zhang. Z., Luo. Y., Cao. J., Zhang. L., Cheng. F., Zhuang. H., Zhang. J., and Tao. F. (2021) NESEA-Rice10: high-resolution annual paddy rice maps for Northeast and Southeast Asia from 2017 to 2019, Earth Syst. Sci. Data, 13, 5969–5986.
- Kim., Sunghack L., Jinyong C. (2014). An Analysis of Agricultural Infrastructure Status of North Korea Using Satellite Imagery, Korean National Committee on Irrigation and Drainage, 21(1), 45-54.
- Lee., S. H. Yoo., J. Choi., S. Bae. (2017). Assessment of the Impact of Climate Change on Drought Characteristics in the Hwanghae Plain, North Korea Using Time Series SPI and SPEI: 1981–2100. Water, 9(8).
- Nam. E. Hong, G. A. Baigorria. (2015). How climate change has affected the spatio-temporal patterns of precipitation and temperature at various time scales in North Korea, International journal of climatology, 36(2), 722-734.
- Rachel, Elliott A. F., Andrea B., Michael L. M. (2018). Suitability of NDVI and OSAVI as estimators of green biomass and coverage in a semi-arid rangeland, Ecological Indicators, 94(1), 16-21.
- Reed., T. R. Loveland., L. L. Tieszen. (1996). An approach for using AVHRR data to monitor U.S. Great Plains Grasslands, Geocarto International, 11, 13-22.
- Rouse, R. H. Haas, A. J. Schell., D. W. Deering. (1974). Monitoring vegetation Systems in the Great Plains with ERTS. Proceedings, 3rd Earth Resources Technology Satellite(ERTS) Symposium, 1, 309-317.
- Tian, J., Y. Tian., W. Wan., C. Yuan., K. Liu., Y. Wang. (2024). Research on the

- Temporal and Spatial Changes and Driving Forces of Rice Fields Based on the NDVI Difference Method. *Agriculture*, 14(7).
- Tian, Martin B., Yi Y. L., Aleixandre V., Torbern T., Abdoul A. D., Kjeld R., Cheikh M., Yunjia W., Rasmus F. (2016). Remote sensing of vegetation dynamics in drylands: Evaluating vegetation optical depth (VOD) using AVHRR NDVI and in situ green biomass data over West African Sahel, *Remote Sensing of Environment*, 177, 265-276.
- USDA. (2024). North Korea MY 2024/25 Seasonal Crop Outlook and Excess Rainfall and Flood-related Analysis
- Wilton. (2021). Crop Yield Estimation Using NDVI: A Comparison of Various NDVI Metrics, University of Manitoba.
- Yang, Bruce K. W., Larry L. T., Bradley C. R. (1998) An Analysis of Relationships among Climate Forcing and Time-Integrated NDVI of Grasslands over the U.S. Northern and Central Great Plains, *Remote Sensing of Environment*, 65(1), 25-37.
- Yeom., Jeong. S., Jeong. G., Ng. C. T., Deo. R. C., Ko. J. (2018). Monitoring paddy productivity in North Korea employing geostationary satellite images integrated with GRAMI-rice model. *Scientific Reports*, 8.
- EU, INFORM Climate Change, <https://drmkc.jrc.ec.europa.eu/inform-index/INFORM-Climate-Change>.
- FAO, <https://www.fao.org/faostat/en/#country/116>
- USDA, <https://ipad.fas.usda.gov/cropexplorer/Definitions/spotveg.htm>
- USGS, <https://lpdaac.usgs.gov/products/mod09a1v061/>
- 권오상, & 김창길. (2008). 기후변화가 쌀 단수변화에 미치는 영향: 비모수적 및 준모수적 분석. *농업경제연구*, 49(5), 45-64.
- 김용준. (2019). 품질효과를 고려한 벼 작기 관리 기후변화 적응의 경제적 효과 분석. *농업경제연구*, 60(4), 1-28.

조현경, 조은빛, 권오상, & 노재선. (2013). 기후변수와 쌀 생산성: 패널 지역자료를 이용한 준모수적 분석. 농업경제연구, 54(3), 71-94.

조현경, & 권오상. (2014). 재배 시기별 기후변수가 논벼의 단위면적당 생산성과 변동성에 미치는 영향 분석. 농업경제연구, 55(3), 115-140.

강택구, 정기웅, 홍윤근, 김일기, & 이준호. (2021). 북한 자연재해 자료 구축과 협력 전략, 사업보고서, 2021(0), 1-154.

기상청, https://hydro.kma.go.kr/frontHelp/menu1200_2.do

김관호, 이성학, & 최진용. (2014). 인공위성영상을 활용한 북한의 농업생산기반 실태분석, 한국관개배수논문집, 21(1), 45-54.

김다울, 김범환, 한하린, & 이대은. (2022). 북한 기후변화 적응을 위한 국제협력방안: 농업과 자연재해를 중심으로. 대외경제정책연구원 전략심층연구 22(1).

명수정, & 권현한. (2010). 북한의 강수특성 분석과 기후변화 전망. 수자원학회:학술대회 논문집, 1443-1447.

명수정, 최영은, & 신상희. (2022). 북한의 기후변화 취약성과 기후변화 대응을 위한 남북 협력(II). 사업보고서, 2022(0), 1-158.

성재훈, 정학균, & 이현정. (2019). 이상기후가 농업부문에 미치는 경제적 영향 분석. 한국 농촌경제연구원 연구보고서 R881.

홍석영, 나상일, 이경도, 김용석, & 백신철. (2015). MODIS NDVI와 강수량 자료를 이용한 북한의 벼 수량 추정 연구, 대한원격탐사학회지, 31(5), 441-448.

홍석영, 민병걸, 이지민, 김이현, & 이경도. (2012). RapidEye 영상을 이용한 북한의 논 면적 산정, 한국토양비료학회지, 45(6), 1194-1202.

자발적 국가검토 보고서, <https://hlpf.un.org/countries/democratic-peoples-republic-of-korea/voluntary-national-review-2021>

통계청, 주요 식량작물 생산량, https://kosis.kr/bukhan/nkStats/nkStatsIdctChart.do?menuId=M_01_02&listNm=%EB%A6%BC%EC%88%98%EC%82%B0%EC%97%85

통일부, <https://nkinfo.unikorea.go.kr/nkp/pge/view.do;jsessionid=xf8MBTWz>

[8HIp8o6rpNM7OGj0cpzsL_CvFnBhkcZY.ins22?menuId=SO321](https://egis.me.go.kr/intro/land.do)
환경공간정보서비스, <https://egis.me.go.kr/intro/land.do>