



장려상

고랭지배추의 무게와 생육환경의 가변적 시차 관계분석

이윤숙

고랭지배추의 무게와 생육환경의 가변적 시차 관계분석

이윤숙*

Key words

고랭지배추(highland chinese cabbage), 임의의 시차변수(arbitrary-time lag), 동적시간 왜곡(dynamic time warping), 가변 시차 인과관계(Variable-lag Granger causality)

ABSTRACT

The study analyzes dynamic time effects between weight of highland chinese cabbage and growth environmental variables and suggests an alternative approach to modelling in order to improve the existing production models. The weight of agricultural products is affected by arbitrary-time lags of growth and weather conditions in a growth period. However, general regression models used in most studies about cabbage production modelling assume fixed and constant time lags. The study applies for dynamic time warping algorithm to find arbitrary-time lags between weight and growth environmental variables. Based on the findings, we compare with well-known Granger causality and variable-lag Granger causality. The growth environmental variables are collected from the public data system operated by Rural Development Administration (wds.agdc.m.kr). The daily data start from May 1, 2019 to September 31, 2019 and May 1, 2020 to September 31, 2020. We find the existence of arbitrary-time lags between weight and growth and weather variables. Based on the existence of arbitray-time lags, we find variable-lag Granger causality analysis perform better than Granger causality.

차례

- | | |
|---------------|------------|
| 1. 서론 | 4. 분석 결과 |
| 2. 선행연구 검토 | 5. 요약 및 결론 |
| 3. 분석 방법 및 자료 | |

* 강원대학교 농업자원경제학과 조교수

1. 서론

농업에서도 정보통신기술(Information and Communication Technology, ICT)과 사물인터넷(Internet of Things, IoT)을 융합하여 생산에서부터 소비까지 데이터를 기반으로 한 관리 및 예측 시스템을 구현을 위한 다양한 연구가 진행되고 있다. 정보시스템을 기반으로 축적된 데이터는 농업의 생산성 향상, 기후변화 및 병충해 등의 외부환경에 따른 작황 변화에 대응 및 균등한 품질의 농산물 안정적 공급 등에 활용되어 농업발전에 이바지할 수 있다. 농업에서 데이터의 역할은 첨단기술과의 융합으로 더욱 그 중요성이 강조되고 있다. 최근 시설뿐 아니라 노지 농산물의 생산성과 직접적으로 연관된 작물의 생육 및 환경요인들의 정보도 주기와 형식에 따라 축적되고 있다. 이렇게 축적되고 있는 데이터의 활용이 새로운 농업의 발전을 위해 필요한 시점이다.

본 연구에서는 다양한 실측조사 농산물 중에서 배추를 선택하였다. 배추는 김치의 주요 재료로 연중 소비되기 때문에 안정적 생산과 공급이 매우 중요한 채소이다(이진형 등 2017). 저온성 채소인 배추는 서늘한 기후를 좋아하기 때문에 기온이 높아지는 여름철에는 고랭지 이외의 지역에서는 재배가 어렵다(이승호 등 2014; 김기덕 등 2015). 고랭지배추는 여름철 김치의 주원료로 중요한 위치를 차지하고 있지만, 고온에 약한 특성이 있어 30℃ 이상의 기온과 가뭄이 2주일 이상 지속되면 생육이 떨어지고, 결구¹⁾불량, 석회결핍증, 무름병 등의 질병 발생률도 높아진다.²⁾ 지구온난화로 인한 빈번한 기후변화 문제는 고랭지배추의 불안정한 작황을 야기하고 농가의 소득감소로 이어진다. 농가의 소득감소는 다시 국내 고랭지배추의 재배면적과 생산량 감소에 영향을 미친다. 게다가 저렴한 중국산 김치의 수입 증가도 고랭지배추 생산 감소에 영향을 미치고 있다.

고랭지배추의 불안정한 생산에 대응하고 예측 가능한 생산모델 구축을 통해 영농계획과 출하시기조절 등의 의사결정에 도움을 줄 수 있는 적절한 생산모델이 필요하다. 여러 연구에서

1) 결구란 배춧잎이 여러 겹으로 겹쳐져 둥글게 속이 드는 상태를 말함(출처: <https://www.korea.kr/news/pressReleaseView.do?newsId=155971079>).

2) 출처: 농사로 고랭지배추 주요 재배기술(<http://www.nongsaro.go.kr/portal/ps/psv/psvr/psvre/curationDtl.mo?menuId=PS03352&srchCurationNo=1313>)

생산모형을 제안하고 있지만, 대부분은 기상요인이 배추의 단수에 미치는 영향을 회귀분석을 통해 추정하고 예측하였다(이진희 등 2015; 이승호 등 2018). 이러한 연구는 농산물의 생육 과정에서 영향을 미치는 다양한 요인들 사이에는 시간 차이가 존재하고 이러한 시차의 영향은 매우 가변적이라는 사실은 간과하고 있다. 나명환 등(2017)은 토마토의 최적생육환경을 찾기 위해 생장과 환경요인 사이의 시간 차이는 고려하고 있지만, 고정된 시차를 가정하여 분석하고 있어 빈번한 기후변화에 대응할 수 있는 생산모델로는 한계가 있다. 본 연구에서는 농산물 생육과정에서 발생하는 생육환경요인들의 역동적 시간 차이를 반영할 수 있는 분석 방법을 제안하고자 한다.

2. 선행연구 검토

배추의 생산 및 김치 수요의 변화를 반영하고 예측할 수 있는 수급모형 구축과 연관된 다양한 연구가 진행되어 왔다. 김연중 등(2006)은 기존의 배추 수급모형을 보완하기 위해 계절별 배추의 상호 연관성을 반영하였다. 또한, 계절에 따라 변화는 김치 수입량을 국내 공급량에 영향을 미치는 내생변수로 간주하여 생산부문의 모델을 구체화하고 김치 소비변화에 대응할 수 있는 수급모형을 구축하였다. 그러나 그들의 모형은 국내 배추 생산에 가장 큰 영향을 미치는 기후변화를 직접적으로 반영하지 못하는 한계가 있었다. 박지연 등(2013)은 이러한 단점을 보완하기 위해 생육단계별 기상요인을 연계한 개선된 모델을 제안하였다. 이승호 등(2018)은 기상요인이 지역별 고랭지배추 단위면적당 생산량에 미치는 영향을 생육단계별로 추정하였고, 이를 바탕으로 미래 기후변화(온실가스 배출 변화) 시나리오를 바탕으로 단위면적당 생산량 변화율을 전망하였다. 이러한 연구는 생육단계별 나타나는 차이는 고려하였지만, 생육 및 기상요인이 배추의 생장에 영향을 미치는 가변적인 시간 차이를 고려하지 않았다는 점에서 국내 생산량을 예측하는 데 한계가 있다. 나명환 등(2017)은 토마토 생장에 영향을 주는 요인들은 시차를 갖는 변수로 전환하여 시계열 분석을 수행하였다. 이 연구에서는 시차를 고려하여 분석을 시도하였지만 각 변수에 적용된 지연변수는 1주 전 평균온도, 7주 전 일사량,

5주 전 물흡수량, 7주 전 공급 pH농도 등 고정된 시차를 가정하고 있어 가변적인 시차의 영향을 분석하기에는 한계가 있다.

생육 및 환경의 특성을 중점적으로 반영한 생산모델 구축에 관한 연구도 다양하게 수행되고 있다. 김진희 등(2015)은 배추 성장에 온도의 중요성을 강조하였다. 그들은 온도 구간에 따라 배추의 생육을 정량화하였고, 이를 바탕으로 정식 후 경과기온만으로 수확시기를 추정하는 성장모형을 구축하였다. 김기덕 등(2015) 고랭지배추 생산량 예측모델을 구축하기 위해 비파괴 측정 생육량 변수에 생육도일을 포함한 구중³⁾추정 모델을 도출하였다. 이 구중추정식을 바탕으로 엽생장량, 실측 생장량의 비, 토양수분에 따른 생육속도와 생장단계 및 기간에 따른 상대생장률을 적용하여 고랭지배추의 총생산량을 예측할 수 있는 기반을 구축하였다. 이진형 등(2017)은 김기덕 등(2015)이 제안한 모델을 활용하여 봄배추와 가을배추의 정식시기에 따라 실시간으로 측정되는 생육지표 값과 재배기간 중의 기상요소를 기반으로 배추의 생산량을 추정하였다. 농산물의 생육 및 환경 특성을 반영한 생육모델 구축과 연관된 연구들은 실험용 데이터를 사용하였거나 데이터에 대한 접근이 제한적이어서 일반적인 수급예측 모델으로 활용되기 어려운 점이 있다.

선행연구에서 제안하고 있는 생산량 예측모델은 농산물의 생육단계에서 다양한 생육 및 환경요인들이 임의의 시점으로부터 영향을 받을 수 있다는 점을 간과하였거나, 접근성이 제한적인 데이터를 사용하여 모델을 일반화하기 어려운 점이 있었다. 본 연구에서는 공공데이터를 사용하여, 생육 및 환경요인 사이에서 발생하는 가변적 시차의 영향을 반영할 수 있는 관계를 분석을 수행하였다.

3) 외엽과 외엽이 붙었던 줄기 부위를 제외한 무게임(출처: 농촌진흥청 농사로).

3. 분석 방법 및 자료

3.1. 분석 방법

농산물은 파종 때부터 수확 전까지 임의의 시점(arbitrary-time lags)에서 나타나는 생육 변수와 기상요인으로부터 영향을 받는다. 배추의 생육과정은 정식(아주심기) 이후부터 결구가 발생하기 직전까지의 기간을 생육기, 결구 발생 이후부터 수확 전까지를 결구기라고 한다(김진희 등 2015). 배추가 결구의 형태를 보이기까지는 정식 이후 대략 30일 정도가 소요되고, 결구 형성을 경계로 생육 초기에는 20~22°C에서 생육이 촉진되고, 결구기는 7~8월로 15~16°C에서 생육 촉진된다(안재훈 등 2008; 김진희 등 2015; 농촌진흥청 2007). 배추의 상품성에 핵심적인 영향을 미치는 결구의 형성은 초기생육환경과 밀접한 연관이 있다. 이러한 배추의 성장과 환경요인의 동적인 관계를 설명하기 위해서는 기존의 시계열 분석에서 활용되고 있는 시차 간 반응 경로의 일대일 매칭이 아니라 가변적인 요소를 고려해야 한다.

본 연구에서는 고랭지배추의 실무게(g)와 가변적인 시차의 존재를 확인하기 위해 동적시간 왜곡(Dynamic time warping) 알고리즘이 반영된 유사도 추정(Similarity measure)을 수행하였고, 이 결과를 바탕으로 생육 및 환경변수 사이의 동적인 시차 관계를 분석하기 위해 가변 시차 인과관계(Variable-lag Granger causality) 방법을 활용하였다. 유사도 추정을 통해 배추의 무게와 설명변수 사이에는 가변적 시차에 의한 동적인 관계가 존재함을 시각적으로 제시하였고, 유사도 추정 결과를 바탕으로 전통적인 그랜저 인과관계(Granger causality) 검증과 가변적 시차 인과관계 검정의 비교를 통해서 고랭지배추의 무게와 생육 및 환경변수의 동적인 시차 관계를 분석하였다.

3.1.1. 유사도⁴⁾ 추정(Similarity measure)

시계열 데이터는 일정한 시간 간격으로 배치된 순서를 가지는 데이터를 의미한다. 농산물

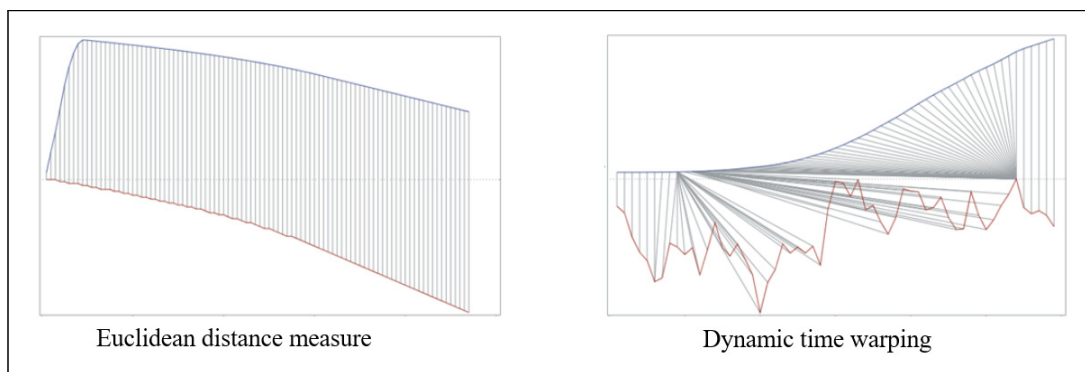
4) 두 데이터 사이의 비슷한 정도를 나타내는 척도로 유사도(similarity)는 n 차 상의 공간에서 두 벡터 사이의 거리를 의미함(출처: Wikipedia).

실측조사 자료는 1주 혹은 1일 간격으로 측정되고 있고, 기상변수는 실시간으로 측정되어 데이터가 조사되는 시간 간격이 다르다. 이러한 실측조사 자료를 통해 두 변수 관계를 분석하기 위해서는 시간 단위를 맞추는 과정을 거치는 것이 일반적이다. 그러나 데이터 변환과정 데이터의 특성이 변형될 수 있다(Batista et al. 2001). 그래서 이러한 생육환경 정보 사이의 가변적인 시차의 영향을 반영하기 위해 동적시간왜곡(Dynamic Time Warping, DTW) 알고리즘을 적용하여 유사도(거리)를 추정해 볼 수 있다.

DTW는 두 시계열 데이터의 길이 변환을 비교하는 알고리즘이다(Coelho 2012). DTW 알고리즘은 음성 인식 분야에서 말하는 속도가 다른 피실험자의 음성과 표준화된 음성의 유사도 측정을 목적으로 개발되었다(이환철 등 2020). 그러나 최근에는 시차 사이에 가변적인 특성이 있는 시계열 데이터를 분석하기 위해 다양한 분야에서 활용하고 있다. 이석준 등(2011)은 선물시장에서 최적의 매매 타이밍을 도출하여 거래 시 적합한 전략을 찾기 위해 DTW 알고리즘을 반영한 트레이딩시뮬레이션 모델을 제안하였다. Franses et al.(2020)와 Wang et al.(2012)은 시차 간의 가변성이 높은 금융시장 정보를 분석하기 위하여 DTW 알고리즘을 시계열 모델에 적용하여 분석하였다. Csillik et al.(2019)는 기후변화에 따른 농산물 지도를 구축하기 위해 경작지 위성 이미지 자료와 기상요인 변수의 관계를 분석하기 위해 DTW 알고리즘을 적용하였다.

본 연구는 DTW 알고리즘이 반영된 유사도(거리) 추정방법을 사용하여 고령지배추의 생육 및 환경변수와 무게 사이에 시차 간의 가변적인 관계가 존재하는 것을 분석하고자 한다. 이석준 등(2011)은 유클리디안 거리 측정(Euclidean distance measure)과 DTW 알고리즘을 비교하여 두 시계열의 반응거리를 분석에 하였다. 그 결과 DTW 알고리즘을 적용한 거리 측정 방법이 유클리디안 거리 측정 방법보다 가변적인 시차 관계를 더 적절하게 분석한다는 것을 보여주었다. <그림 1>은 가변적 시차의 특성을 갖는 실제 시계열 데이터에 유클리디안 거리 측정 방법과 DTW 알고리즘을 적용한 분석한 사례를 보여주고 있다. 두 변수 사이에 가변적 시차의 존재함에도 불구하고 일정한 패턴과 순서에 의존하는 일부 시계열 모델을 사용한 추정값은 잘못된 결과를 도출할 수 있다(Giorgino et al. 2009).

그림 1. 시계열 데이터 시차 관계 비교방법



주: 1) DTW 그래프는 2019년 5월 1일부터 6월 30일 동안의 안반데기 고랭지배추 엽수와 무게 실측 데이터를 활용함.
2) SAS 프로그램을 사용함.

3.1.2. 가변 시차 인과관계(Variable-lag Granger causality)

그랜저 인과관계(Granger causality) 분석은 변수 사이 및 변수의 과거와의 원인과 결과 관계를 분석할 때 일반적으로 사용되고 있는 기본적인 시계열 분석 방법이다(Granger 1969). 그러나 흔히 사용되고 있는 그랜저 인과관계 분석은 일정하게 고정된 시차(fixed-lag)를 기반으로 변수 간의 인과관계를 분석한다(McCrorie et al. 2006). 그러나 이러한 그랜저 인과관계 분석은 농산물의 생육과정 및 자연현상의 가변적인 시간차 관계를 분석하기에 한계가 있다. 본 연구에서는 DTW 알고리즘을 기반으로 임의의 시간 차이를 반영한 가변 시차 인과관계(Variable-lag Granger causality) 분석을 활용하였다. 본 연구에서는 이러한 분석 접근법 제시를 통해 배추 생산량 예측 모델을 고안하기 위해서는 가변적 시차가 반드시 고려되어야 함을 강조하였다.

전형적인 그랜저 인과관계와 가변적 시차 인과관계 검정의 차이를 설명하기 위해 $X(t)$ 는 농산물의 생육 및 환경변수로, $Y(t)$ 는 무게로 가정하였다. 먼저 일반적으로 사용되는 고정 시차 인과관계 회귀식을 살펴보았다.

$$(1) \gamma_Y(t) = Y(t) - \sum_{i=1}^{\delta_{\max}} a_i Y(t-i)$$

$$(2) \gamma_{YX}(t) = Y(t) - \sum_{i=1}^{\delta_{\max}} (a_i Y(t-i) + b_i X(t-i))$$

단, $\delta_{\max}$ 는 최대 시차, r_Y 는 Y의 회귀식의 잔차, γ_{YX} 는 X와 Y의 회귀식의 잔차, a_i 와 b_i 는 i 번째 시차에서 잔차를 최소화하는 상수를 의미한다. (1)과 (2) 식으로부터 γ_{YX} 의 분산값이 r_Y 의 분산값 보다 작으면 X와 Y는 그랜저 인과관계가 존재한다고 표현할 수 있다.

X(t)와 Y(t) 사이의 가변적 시차가 인과관계를 분석하기 위해 아래와 같은 회귀식을 적용하였다.

$$(3) \gamma_{YX}^*(t) = Y(t) - \sum_{i=1}^{\delta_{\max}} (a_i Y(t-i) + b_i X(t-i) + c_i X^*(t-i)),$$

$$X^*(t-i) = X(t-i+1 - \Delta_{t-i+1})$$

단, $\Delta_t > 0$.

(1), (2)와 (3) 식으로부터 $\gamma_{YX}^*(t)$ 의 분산값이 r_Y 와 γ_{YX} 각각의 분산값 보다 작으면 X와 Y는 가변적 시차가 인과관계가 존재한다고 표현할 수 있다.

3.2. 분석 자료

본 연구에서는 고랭지배추의 생육과정에 따른 동적 관계 분석하기 위하여 농촌진흥청(Rural Development Administration, RDA)에서 운영하는 배추, 마늘 주산지 생육정보 시스템(wds.agdcn.kr)이 제공하는 배추의 생육 및 기상정보를 활용하였다.

한국농촌경제연구원 농업관측통계시스템(Outlook & Agricultural Statistics Information System, OASIS)에서 제공하는 작물의 실측정보는 지역 및 품목별로 생육정보가 세분화되어 있어 횡단자료는 충분히 제공되고 있었다. 그러나 시간 이동에 따른 생육과정을 분석할 수 있는 시계열자료는 부족하였다. 고랭지배추의 경우 농가에 따라 2주 혹은 3주 정도의 데이터를

제공하고 있어 생육과정에서 시간에 따른 동적 변화를 분석할 수 있을 만큼의 시계열 정보가 충분하지 않았다. 농촌진흥청의 배추, 마늘 주산지 생육정보 시스템에서는 제공하는 생육정보는 엽수(매)와 실무게(g/포기당) 제공하고 있어, 횡단자료 측면에서는 OASIS 자료보다는 부족한 면이 있다. 그러나 생육과정 정보가 1일 기준으로 연속적으로 제공되고 있어 고랭지 배추의 동적인 생육변화를 추정하기에 더 적절하였다.

본 연구에서는 강원도 영월군 매봉산과 강릉시 안반데기 지역의 고랭지배추의 생육자료와 기상자료를 사용하였고, 정식부터 수확기까지를 고려하여 2019년 5월 1일부터 9월 30일, 2020년 5월 1일부터 9월 30일까지 일별 자료를 사용하였다. 고랭지배추의 실무게(g/포기당)는 생산량을 대체하여 사용하였고, 무게에 영향을 미치는 생육변수로는 엽수(매), 기상변수로는 최저기온($^{\circ}\text{C}$), 최고기온($^{\circ}\text{C}$), 강수량(mm), 일사량(MJ/m^2), 일조시간(h)을 사용하였다. 김기덕 등(2015)과 이진형 등(2017)은 온도를 배추의 생산에 중요한 영향을 미치는 기상요인으로 간주하였고, 안재훈 등(2014)과 이인혜 등(2018)은 평균기온, 최저기온, 최고기온, 최대연속건조일수, 최대연속강우일수, 호우일수 등을 고랭지배추의 생장에 중요한 영향을 미치는 기상요인으로 사용하였다. 본 연구에서는 데이터의 접근성을 고려하여 선행연구에서 사용한 유사한 생육 및 환경변수를 선정하였다.

〈표 1〉은 연구에 사용된 지역별 고랭지배추의 생육 및 환경변수의 기초 통계량을 보여주고 있다. 매봉산 고랭지배추 포기당 실무게는 2019년에 비해 2020년 평균 무게가 더 무거웠다. 평균 엽수도 2020년이 더 많았다. 평균 최저기온은 2019년이 더 높았고, 평균 최고기온은 2020년이 더 낮았다. 강수량은 2020년이 더 많았고, 일사량과 일조시간은 2020년보다 더 낮았다. 안반데기의 경우 포기당 실무게는 2019년 평균 무게가 2020년보다 무거웠다. 평균 엽수도 2019년이 더 많았다. 평균 최저기온과 평균 최고기온의 차이는 뚜렷하지 않았다. 평균 강수량은 2020년이 더 많았다. 평균 일사량과 평균 일조시간의 차이는 뚜렷하지 않았다.

표 1. 고령지배추 생육 및 환경변수의 기초 통계량

영월군 매봉산(2019)						
변수명	관측치 수	단위	평균	표준편차	최솟값	최대값
실무게	153	g/포기당	1758.4	901.7	0.9	2432.6
엽수	153	매	45.4	14.5	3.5	56.6
최저기온	153	℃	12.3	4.0	0.0	21.3
최고기온	153	℃	20.0	3.6	9.9	26.8
강수량	153	mm	6.9	17.4	0.0	114.5
일사량	153	MJ/m ²	17.5	8.3	5.5	29.7
일조시간	153	h	5.6	2.9	1.1	9.7
영월군 매봉산(2020)						
실무게	153	g/포기당	2151.9	1160.9	1.1	3024.2
엽수	153	매	52.4	18.1	3.8	65.9
최저기온	153	℃	13.2	3.7	1.1	22.0
최고기온	153	℃	18.6	3.8	7.2	26.4
강수량	153	mm	10.3	20.4	0.0	126.9
일사량	153	MJ/m ²	16.2	8.9	5.4	29.8
일조시간	153	h	4.9	3.2	1.1	9.6
강릉시 안반데기(2019)						
실무게	153	g/포기당	2151.8	1160.8	1.1	3024.1
엽수	153	매	52.4	18.1	3.8	65.9
최저기온	153	℃	13.2	3.7	1.1	22.0
최고기온	153	℃	18.6	3.8	7.2	26.4
강수량	153	mm	10.3	20.4	0.0	126.9
일사량	153	MJ/m ²	16.2	8.9	5.4	29.8
일조시간	153	h	4.9	3.2	1.1	9.6
강릉시 안반데기(2020)						
실무게	153	g/포기당	1719.9	897.8	1.1	2385.6
엽수	153	매	45.6	14.7	3.8	57.1
최저기온	153	℃	13.2	4.0	0.6	20.9
최고기온	153	℃	18.4	4.0	8.5	27.0
강수량	153	mm	11.8	26.2	0.0	150.6
일사량	153	MJ/m ²	16.1	8.9	5.2	30.0
일조시간	153	h	5.0	3.3	1.1	10.1

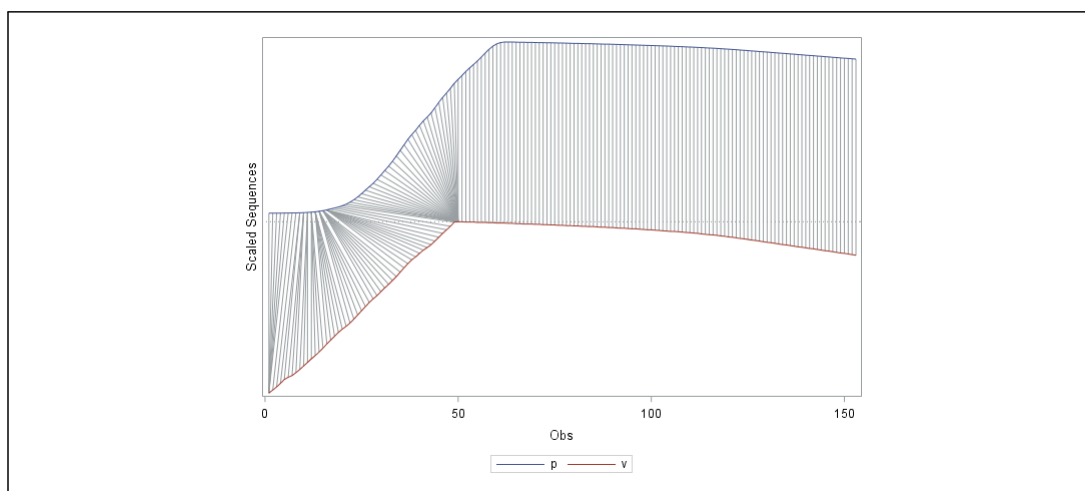
자료: 배추, 마늘 주산지 생육정보시스템(wds.agdcm.kr)

4. 분석 결과

4.1. 유사도 추정 결과

본 연구에서는 고랭지배추 무게(g/포기당)에 영향을 미치는 각각의 설명변수와 짝을 이루어 DTW 유사도 추정이 수행되었다. <그림 2>는 2019년 5월 1일부터 9월 30일 강원 영월군 매봉산 고랭지배추의 무게와 생육환경변수를 분석한 결과이다. <그림 2>는 매봉산 고랭지배추의 무게와 엽수의 DTW 유사도 추정 결과를 보여주고 있다. 생육초기(관측점 0~60일) 동안에 무게와 엽수의 시차 사이에서 나타나는 역동적인 대응 관계를 나타나고 있었다. 60일 이후로 엽수와 무게의 시차 간격은 일대일로 매칭이 되는 것을 볼 수 있었다.

그림 2. DTW를 적용한 엽수와 무게와의 관계(영월군 매봉산)



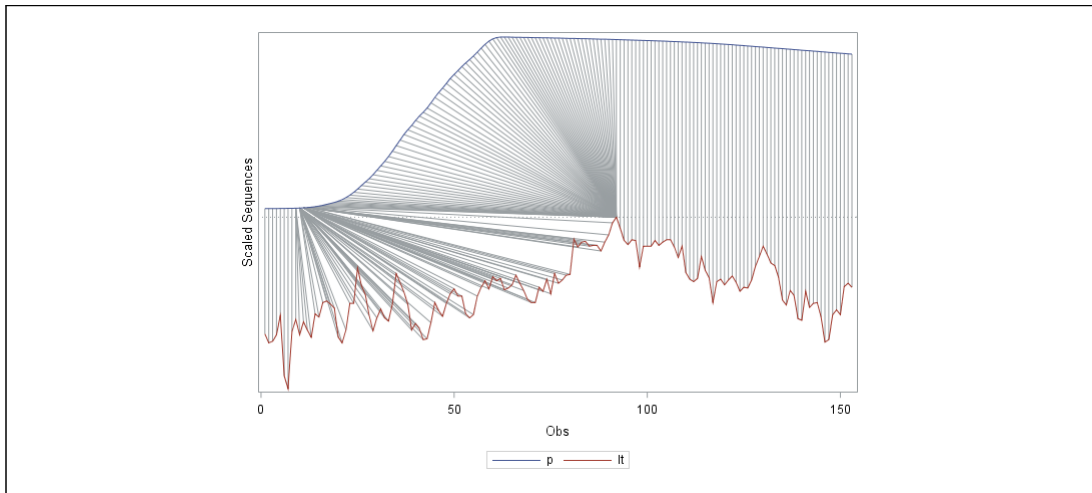
주: 1) p는 무게, v는 엽수를 의미함.

2) 태백시 매봉산 지역에서 2019년 5월 1일부터 9월 30일까지 수집된 데이터임.

자료: 배추, 마늘 주산지 생육정보시스템(wds.agdcm.kr)

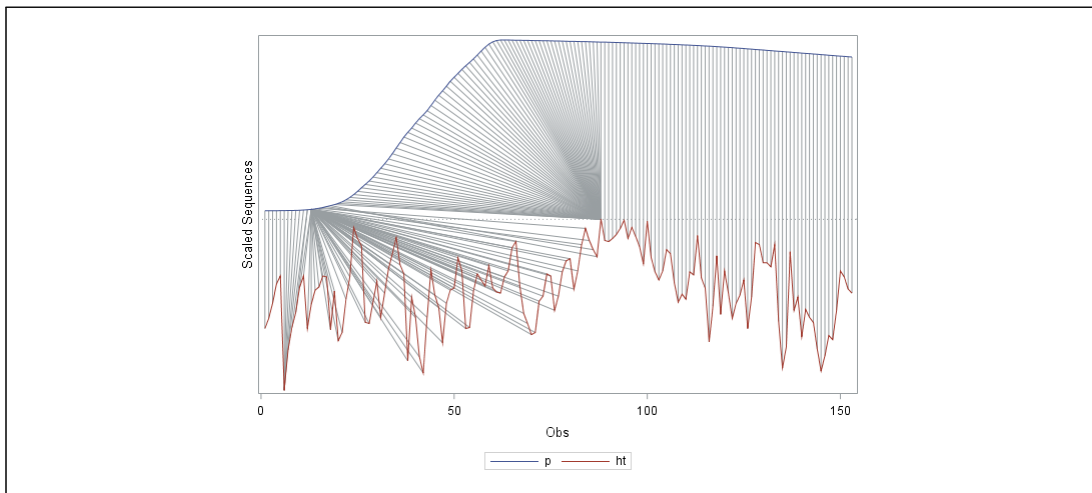
<그림 3>과 <그림 4>는 최저기온, 최고기온과 무게의 사이의 시차 간격이 가변적으로 변화하는 것을 보여주고 있다. 최저기온의 경우 시차의 가변적 특성은 생육초기 뿐 아니라 결구기에도 나타나고 있다.

그림 3. DTW를 적용한 최저기온과 무게와의 관계(영월군 매봉산)



주: p는 무게, lt는 최저기온을 의미함.

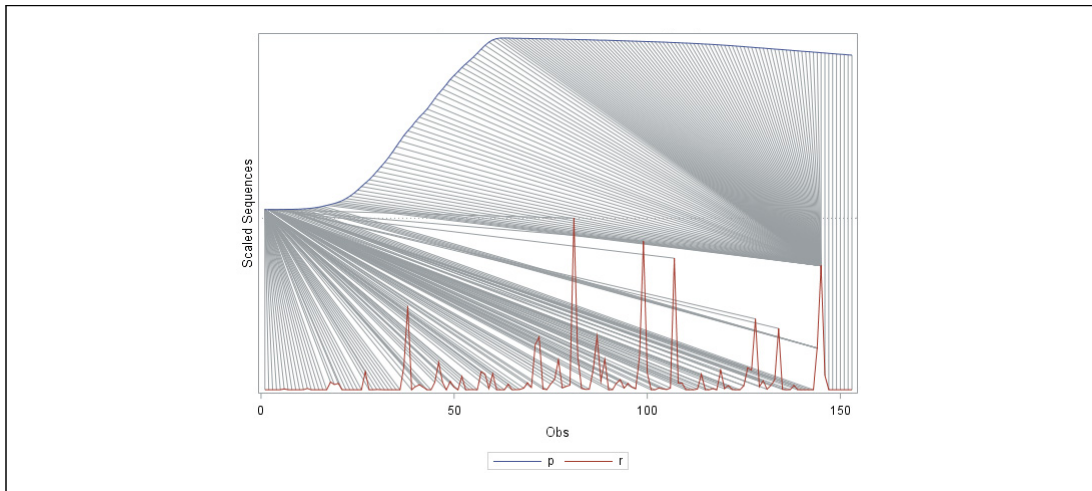
그림 4. DTW를 적용한 최고기온과 무게와의 관계(영월군 매봉산)



주: p는 무게, ht는 최고기온을 의미함.

〈그림 5〉는 고랭지배추의 무게와 강수량 사이에도 가변적 시차의 관계가 나타나고 있음을 보여주고 있다. 강수량의 경우 생육초기, 결구기, 수확 전까지 고랭지배추의 무게에 미치는 시차의 영향은 가변적임을 볼 수 있다.

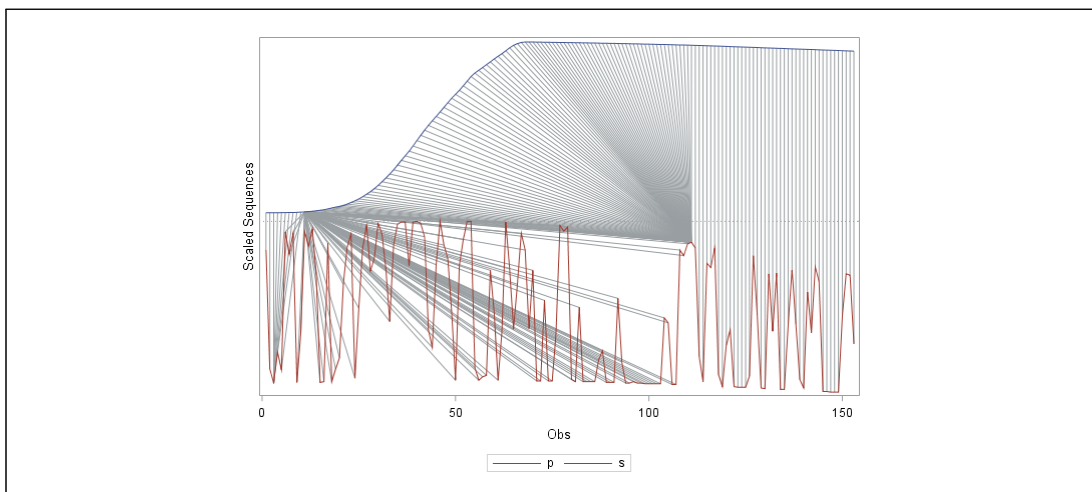
그림 5. DTW를 적용한 강수량과 무게의 관계(영월군 매봉산)



주: p는 무게, r는 강수량을 의미함.

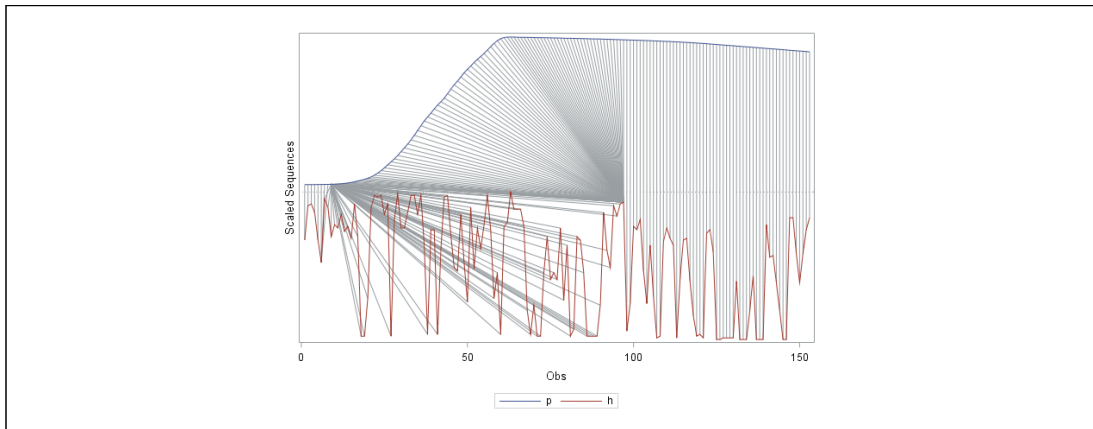
〈그림 6〉과 〈그림 7〉은 일조량, 일조시간과 무게의 가변적 시차 관계를 보여주고 있다. 각각의 기상변수가 고랭지배추 무게에 미치는 시차의 변동성이 다른 기상변수보다 큰 것을 볼 수 있다.

그림 6. DTW를 적용한 일조량과 무게의 관계(영월군 매봉산)



주: p는 무게, s는 일조량을 의미함.

그림 7. DTW를 적용한 일조시간과 무게의 관계(영월군 매봉산)

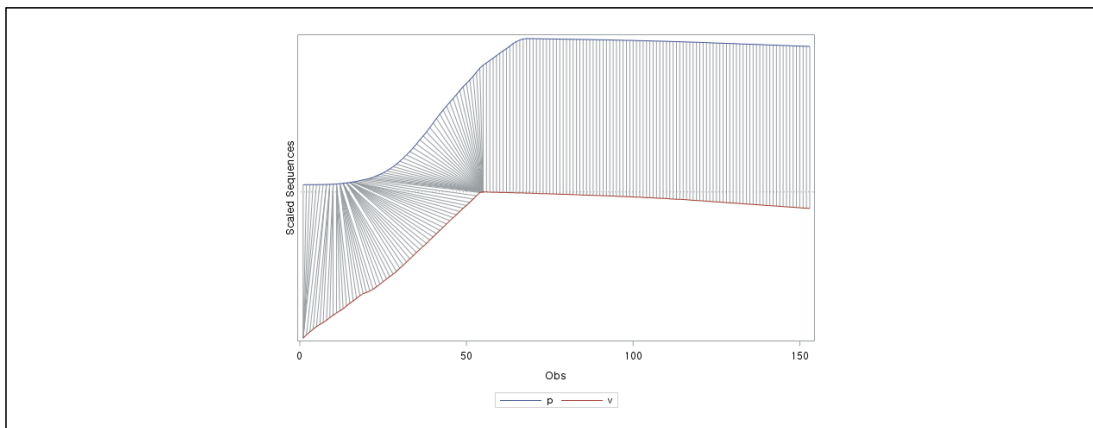


주: p는 무게, h는 일조시간을 의미함.

2020년 5월 1일부터 9월 30일 강원 영월군 매봉산 고랭지배추 자료에서도 무게와 생육 환경변수 사이에 가변적 시차의 특성이 존재하였다. 그래프는 2019년 유사도 결과와 유사하여 생략하였다.

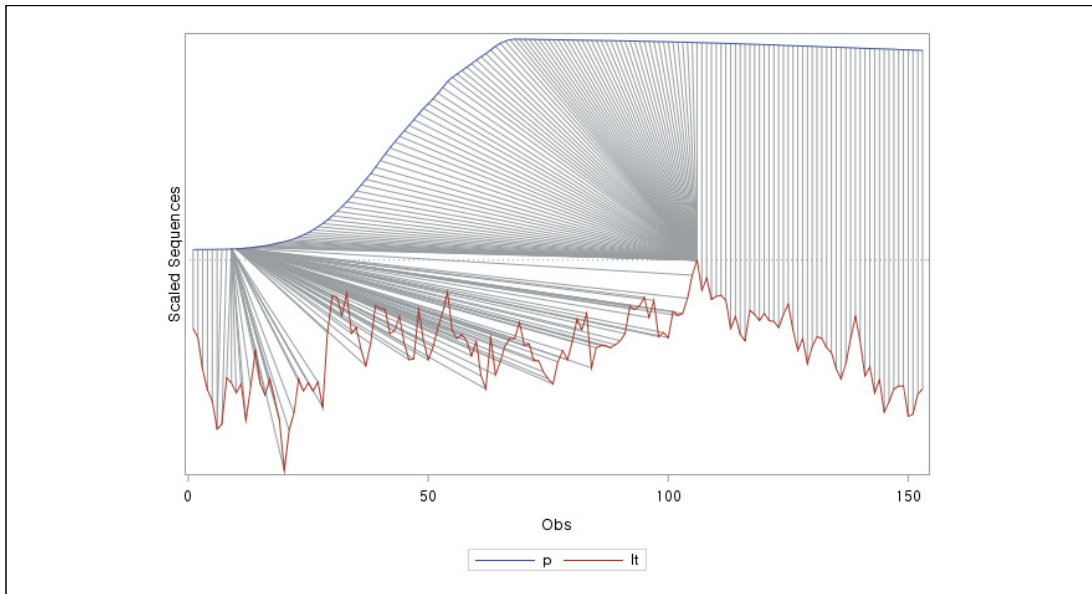
〈그림 8〉은 강릉시 안반데기 지역의 고랭지배추 자료도 사용하여 유사성 추정한 결과를 나타내고 있다. 안반데기 고랭지배추 역시 생육 및 환경변수와 무게 사이에 가변적 시차의 특성이 존재하였다.

그림 8. DTW를 적용한 엽수와 무게의 관계(강릉시 안반데기)



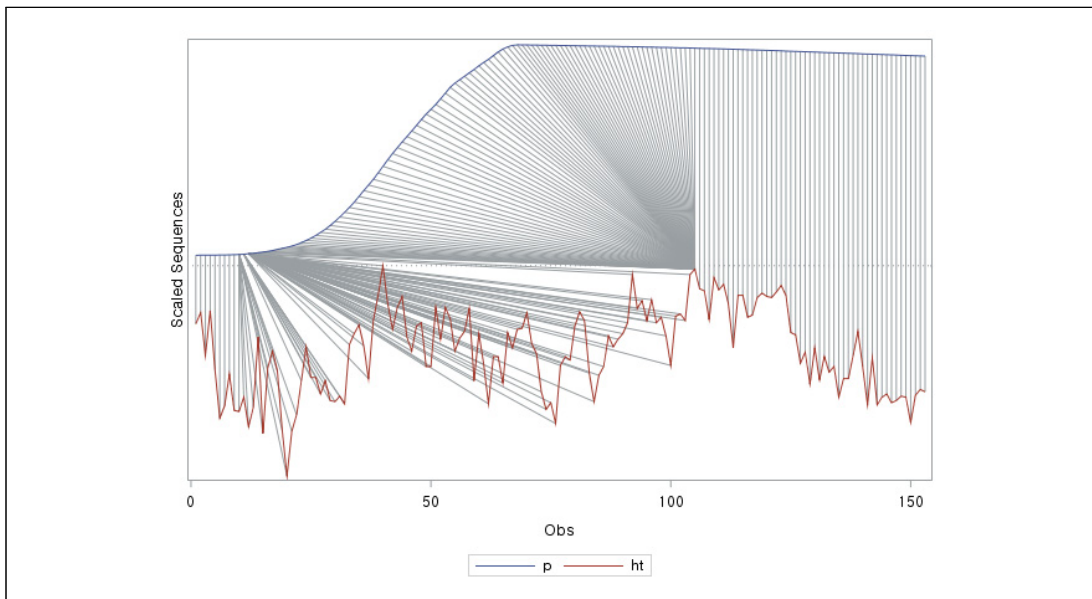
주: p는 무게, v는 엽수를 의미함.

그림 9. DTW를 적용한 최저기온과 무게의 관계(강릉시 안반데기)



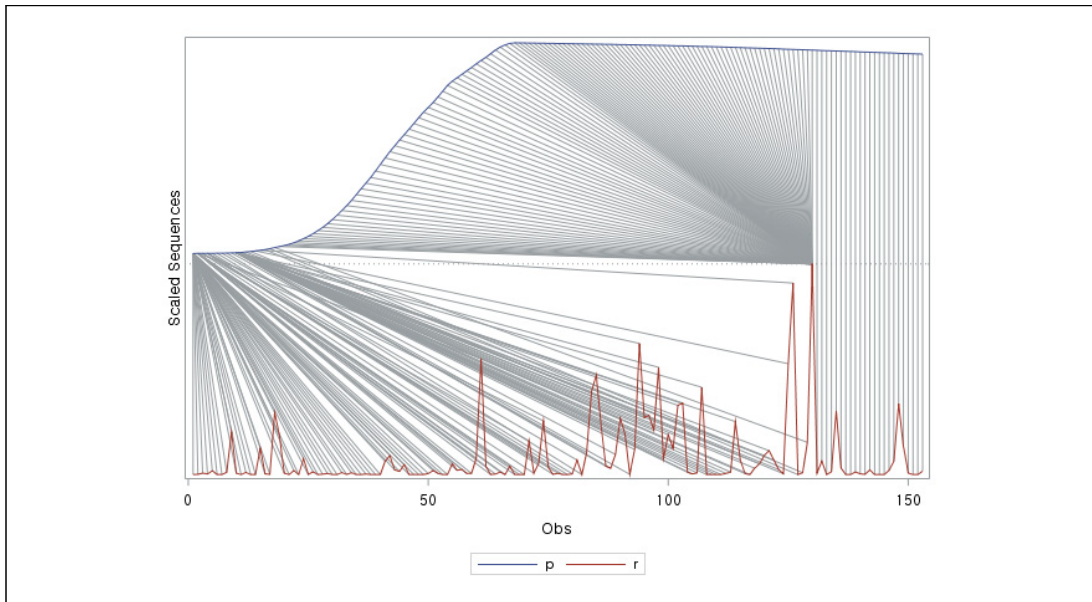
주: p는 무게, lt는 최저기온을 의미함.

그림 10. DTW를 적용한 최고기온과 무게의 관계(강릉시 안반데기)



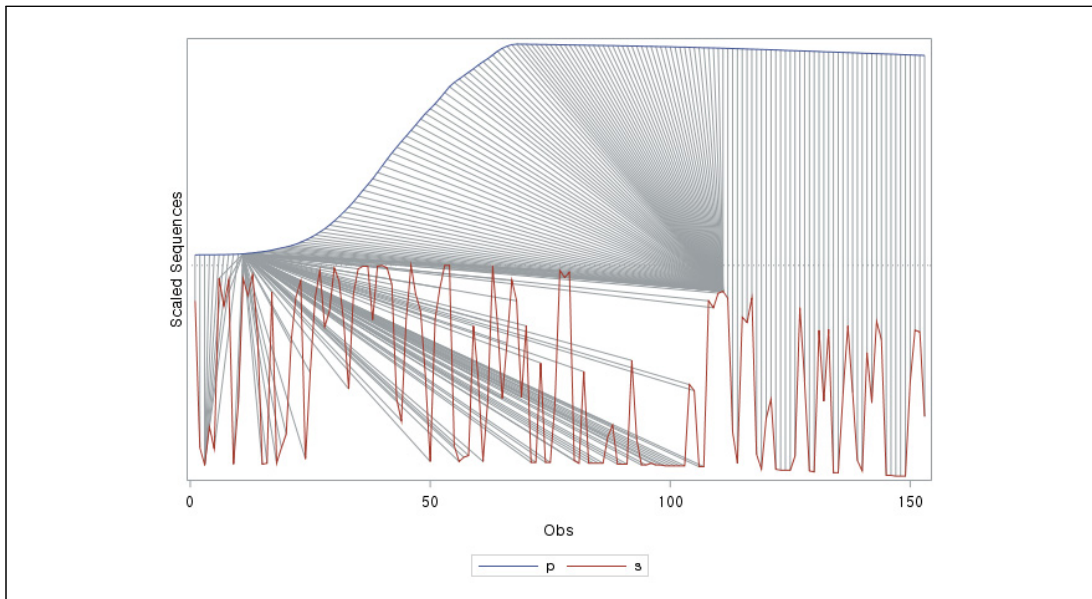
주: p는 무게, ht는 최고기온을 의미함.

그림 11. DTW를 적용한 강수량과 무게의 관계(강릉시 안반데기)



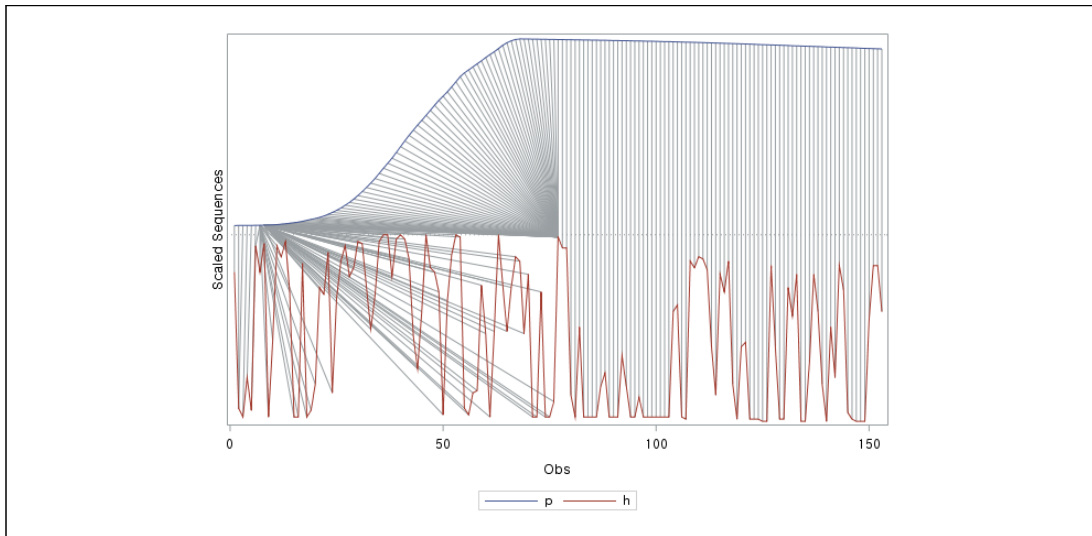
주: p는 무게, r는 강수량을 의미함.

그림 12. DTW를 적용한 일조량과 무게의 관계(강릉시 안반데기)



주: p는 무게, s는 일조량을 의미함.

그림 13. DTW를 적용한 일조시간과 무게의 관계(강릉시 안반데기)



주: p는 무게, h는 일조시간을 의미함.

4.2. 가변 시차 인과관계 결과

DTW 알고리즘이 반영된 유사성 추정 결과를 바탕으로 가변적 시차 인과관계 검정을 수행하였다. <표 2>와 <표 3>은 전형적인 그랜저 인과관계와 가변적 시차 인과관계 비교 결과를 보여주고 있다. False는 두 변수의 인과관계 분석에는 고정된 시차가 반영된 그랜저 인과관계가 더 적절함을 의미하고, True는 두 변수의 인과관계 분석에는 가변적 시차가 반영된 그랜저 인과관계가 더 적절함을 의미한다.

2019년 2020년 매봉산 고랭지배추의 무게와 과거 무게의 인과관계 분석에는 고정 시차 인과관계가 더 적절하였다. 엽수와 무게, 최저기온과 무게, 최고기온과 무게, 강수량과 무게, 일조량과 무게, 일조시간과 무게 사이에서 인과관계 분석에는 가변 시차 인과관계가 더 적절하였다. <표 3>은 2019년 2020년 안반데기 고랭지배추의 그랜저 인과관계와 가변 시차 인과관계 결과를 보여주고 있다. 안반데기 고랭지배추의 무게와 생육환경변수 관계 역시 매봉산 고랭지배추와 유사한 결과를 보여주고 있다.

고정 시차 인과관계와 가변 시차 인과관계 비교 결과는 고랭지배추 무게에 영향을 미치는 생육 및 환경변수의 사이에는 가변적 시차에 의한 인과관계가 존재하였다. 본 연구에서는 간단한 인과관계 분석만을 수행하였지만 보다 정확한 배추의 생산량 추정을 위해서는 시차의 가변성을 고려해야 할 필요성을 보여주고 있다.

표 2. 고정 시차와 가변 시차 인과관계 비교(영월군 매봉산)

		생육 및 기온 변수						
		무게	엽수	최저기온	최고기온	강수량	일조량	일조시간
2019	무게	False	True	True	True	True	True	True
2020	무게	False	True	True	True	True	True	True

주: 1) True는 설명변수와 반응변수 사이에는 variable-lag Granger cause가 적절함 의미함.

2) False는 설명변수와 반응변수 사이에는 fixed-lag Granger cause가 적절함 의미함.

3) R 프로그램을 사용함.

자료: 배추, 마늘 주산지 생육정보시스템(wds.agdcm.kr)

표 3. 고정 시차와 가변 시차 인과관계 비교(강릉시 안반데기)

		생육 및 기온 변수						
		무게	엽수	최저기온	최고기온	강수량	일조량	일조시간
2019	무게	False	True	True	True	True	True	True
2020	무게	False	True	True	True	True	True	True

5. 요약 및 결론

본 연구에서는 고랭지배추 생산성에 직접적인 영향을 미치는 생육 및 기상자료를 활용하여 두 변수 사이에 시차의 영향은 가변적으로 나타남을 분석하였다. 반응변수로는 고랭지배추의 실무게(g/포기당)를 사용하였고, 설명변수로는 엽수(매), 최저기온(°C), 최고기온(°C), 강수량(mm), 일사량(MJ/m²), 일조시간(h)을 사용하였다. 두 변수 사이의 가변적인 시차를 분석하기 위해 DTW 유사도 추정과 가변 시차 인과관계를 적용하였다.

DTW 알고리즘을 기반으로 하는 유사도 추정은 고랭지배추의 무게와 각각의 생육 및 기온

변수 사이에 가변적인 시차가 존재함을 보여주었다. 유사도 추정 결과를 바탕으로 DTW 알고리즘이 적용된 가변 시차 인과관계 검정을 수행하였다. 가변 시차 인과관계 검증에서도 고랭지배추의 무게에 영향을 미치는 생육 및 기온변수 사이에는 가변적인 시차에 의한 인과관계가 존재함을 보여주었다. 본 연구에서는 단순한 분석방법을 활용하여 생육변수들 간의 가변적 시차 관계가 형성되어 있음을 보여주었지만, 이러한 결과는 고정 시차를 가정하는 시계열 모델을 활용한 분석은 잘못된 예측값을 생산할 수 있는 위험이 있음을 시사하고 있다. 기존의 배추 수급모형의 예측력을 높이기 위해서는 배추생육환경을 포함할 수 있고, 생육환경의 가변적인 시간 차이뿐 아니라 가변적 시차의 매칭시점까지도 반영할 수 있는 생산모델의 개선이 필요하다.

참고문헌

- 국립식량과학원. 「고랭지배추 주요 재배기술」.
- 김연중, 김배성, 송성환, 박영구. 2006. 「배추 계절별 수급모형 개발」. 한국농촌경제연구원.
- 김진희, 윤진일. 2015. “온도시간 기반의 배추 생육단계 추정.” 「한국농림기상학회지」 16(4): 333-339.
- 나명환, 박유하, 조완형. 2017. “스마트팜 데이터를 이용한 토마토 최적인자에 관한 연구.” 「한국데이터정보과학회지」 28(6): 1427-1435.
- 농촌진흥청. 2007. 「배추 유기재배 매뉴얼」.
- 박지연, 박영구. 2013. 「배추·무 예측모형 고도화 방안」. 한국농촌경제연구원.
- 안재훈, 김기덕, 이정태. 2014. “고랭지배추의 생장모의.” 「한국농림기상학회지」 16(4): 309-315.
- 이석준, 오경주. 2011. “선물시장의 시스템트레이딩에서 동적시간와핑 알고리즘을 이용한 최적매매 빈도의 탐색 및 거래전략의 개발.” 「한국데이터정보과학회지」 22(2): 255-267.
- 이승호, 허인혜. 2018. “기후가 고랭지배추 생산에 미치는 영향: 강원도를 사례로.” 「대한지리학회지」 53(3): 265-282.
- 이진형, 이희주, 김성겸, 이상규, 이희수, 최장선. 2017. “재배시기, 정식일에 따른 배추의 생육 모델 개발 및 생산량 예측 평가.” 「시설원·예식물공장」 26(4): 235-241.
- Amornbunchornvej, C., E. Zheleva and T. Y. Berger-Wolf. “Variable-lag Granger Causality and Transfer Entropy for Time Series Analysis.” *ACM Trans. Knowl. Discov. Data.* 1(1): 1-30.
- Batista, G. E., D. Gada and V. Puttagunta. 2001. “Distance Measures for Effective Clustering of ARIMA time-series.” Proceedings 2001 IEEE International Conference on Data Mining.
- Csillik, O., M. Belgiu, G. P. Asner and M. Kelly. 2019. “Object-Based Time-Constrained Dynamic Time Warping Classification of Crops Using Sentinel-2.” *Remote Sensing.* 11: 1257-1283.
- Franses, H. P. and T. Wiemann. 2020. “Intertemporal Similarity of Economic Time Series: An Application of Dynamic Time Warping.” *Computational Economics.* 56: 59-75.
- Giorgino, T. 2009. “Computing and Visualizing Dynamic Time Warping Alignment in R: the dtw package.” *Journal of Statistical Software.* 31(7): 1-24.
- Granger, C. W. J. 1969. “Investigating Causal Relations by Econometric Models and Cross-Spectral Methods.” *Econometrica.* 37(3): 424-438.
- McCrorie, J. R. and M. J. Chambers. 2006. “Granger Causality and the Sampling of Economic Processes.” *Journal of Econometrics.* 132: 311-336.
- Wang, G. J., C. Xie, F. Han and B. Sun. 2012. “Similarity Measure and Topology Evolution of Foreign Exchange Markets Using Dynamic Time Warping Method: Evidence from Minimal Spanning Tree.” *Statistical Mechanics and its Applications.* 391(16): 4136-4146.