

# 지리적 표시제에 대한 관심이 농산물 가격변화 예측에 미치는 영향 연구 : 사과를 사례로

최효신 · 손소영<sup>†</sup>

연세대학교 공과대학 정보산업공학과

## Influence of Interests in Geographical Indication on the Prediction of Price Change of Agricultural Product : Case of Apples

Hyo Shin Choi · So Young Sohn

Department of Information and Industrial Engineering, Yonsei University

Geographical Indication (GI) has been used with the expectation to influence customer buying behavior. In this research, we empirically investigate if such relationship exists using apple price changes in Korea along with web search traffic reflecting customers' interest in GI. The experimental results indicate that the apple price of the past, apple supply and web search traffic including GI name were significant on the prediction of price change of Chungju while web search traffic of regional name and that of product were significant for Cheongsong apples with GI. In Yeongcheon with no GI, the apple price of the past turns out to be significant only. The results indicated that interests in GI can help the price prediction but the regional name itself can play the same role, if the GI product is well known in association with the region.

**Keywords:** Geographical indication, Web search traffic, Agricultural products, Price change prediction, Logistic regression

### 1. 서 론

지역 특산 농산물 및 가공품의 품질향상과 지역특화산업 육성 및 소비자 보호를 위해 1999년 국내에 지리적 표시제가 도입되었다. 지리적 표시제는 품종특허, 품종보호권, 상표권 등과 같은 지식재산의 하나이며, 2014년 8월 기준으로 농산물 95건, 임산물 49건, 수산물 19건 등 총 163건이 등록되어 있다.

Kim and Seong(2008)의 연구에 의하면 지리적 표시를 하지 않는 농산물 품목보다 지리적 표시를 한 농산물 품목의 선호도 및 지불의사금액이 높은 것으로 나타났으며, 소비자의 알 권리를 충족시켜 시장차별화 효과가 있는 것으로 연구되었다

(Hong et al., 2009). 이에 지리적 표시 품목에 등록된 농산물과 등록되지 않은 농산물간의 가격 형성 차이가 예상된다.

동일한 품질의 국내 작황 농산물 가격 변동과 관련 있는 요인으로 요일, 월별 및 월중(Yoon and Yang, 2004), 도매시장 반입량, 자연재해 발생, 유통장애(Kim and Cho, 2010) 등을 들 수 있다. 농산물은 비탄력적 공급·수요, 기상변화 등에 따라 생산량의 변화가 심하고 유사 작목 간 생산 대체 관계가 존재하기 때문에 가격이 불안정하다. 또한, 지난 2004년 칠레를 시작으로 잇따라 미국·유럽연합(EU)·호주·캐나다 등과 같은 농업 선진국과 자유무역협정(FTA)을 체결하며 시장 개방을 추진해왔는데, 이로 인해 값싼 수입 농산물의 수입이 크게 늘

본 연구는 '국토교통부 국토공간정보연구사업 국토공간정보의 빅데이터 관리, 분석 및 서비스 플랫폼 기술개발(14NSIP-B081011-01)과제'의 연구비 지원에 의해 연구되었음.

<sup>†</sup> 연락저자 : 손소영 교수, 120-749 서울특별시 서대문구 연세로 50 연세대학교 공과대학 정보산업공학과, Tel : 02-2123-4014, Fax : 02-364-7807, E-mail : sohns@yonsei.ac.kr

2014년 11월 17일 접수; 2015년 2월 2일 수정본 접수; 2015년 3월 20일 게재 확정.

어나고 있다. 이와 같이 공급부문에 예측할 수 없는 변수가 많고 국내의 여건 변화 등으로 인해 농민의 의지에 관계없이 가격변동이 일어나, 농가소득의 변동에 영향을 주는 것은 물론, 유통·가공업자의 계획 차질, 소비자의 가계에 대한 불안을 초래하여 모든 경제주체의 관심 대상이 되고 있다(Lee *et al.*, 2006). 이에 농산물의 가격 변동을 지속적으로 관찰하고 미래 농산물가격 변동을 예측할 필요가 있다.

농수산물 가격 예측을 위해 시계열 모형(Kim, 2005; Luo *et al.*, 2013)과 신경망 모형(Kohzadi *et al.*, 1996; Zou *et al.*, 2007; Guo *et al.*, 2011; Lou *et al.*, 2011; Hemegeetha and Nasira, 2012, 2013; Zu and Liu, 2013)을 적용한 많은 연구가 수행되었다. Zang and Xie(2011)는 자체 가격 데이터와 함께 날씨, 계절, 그 외 관련 가격 등 총 10개의 변수데이터에 Levenberg-Marquardt(L-M) 알고리즘을 바탕으로 BP 신경망분석을 수행하였다. Lee *et al.* (2006)은 Autoregressive Integrated Moving Average(ARIMA) 모형, generalized autoregressive conditional heteroskedasticity(GARCH) 모형 등 자체 가격 시계열만으로 구성되는 단변량 모형뿐만 아니라 복수의 가격이나 물량 변수를 함께 고려하는 Vector Autoregression(VAR), 전이함수 모형의 다변량 모형을 함께 사용하여 가격을 예측하였다. Zang and Xie(2011)의 연구를 제외하고, 대부분의 이전 연구들은 농산물 월별 가격 예측에 소비자의 관심을 반영한 변수를 고려하지 않은 한계점이 있다.

따라서, 본 연구에서는 이러한 소비자의 관심을 반영한 변수를 최근 빅데이터 시대에 사람들이 무의식적으로 남긴 많은 정보로 간주하였다. 그 중 웹 검색 트래픽은 잠재고객들의 관심 정도를 나타낼 수 있는 지표로 많이 사용되고 있는데, 자동차 판매량, 실업수당 청구수, 여행 방문자수, 선거 후보자 당선가능성, 제품 판매량 예측에 활용된 바 있다(Ginsberg *et al.*, 2009; Choi and Varian, 2012; Lui *et al.*, 2011; Jun *et al.*, 2013). 이러한 웹 검색 트래픽 정보는 변동이 심한 농산물 가격을 예측하는데 소비자의 관심을 반영하여 유용하게 활용될 것으로 예상된다.

본 연구는 웹 검색 트래픽 정보, 농산물 가격과 물량 정보 그리고 수확기를 예측변수로 고려한 로지스틱 회귀분석 모형을 이용하여 지리적 표시제 등록된 충주·청송사과와 비등록된 영천사과의 주별 가격 변화에 유의한 중요 변수를 비교 분석하고 가격 변화를 예측하는데 지리적 표시제에 대한 소비자의 관심이 유의한 역할을 하는지 살펴보고자 한다. 농산물 가격 변화예측 모형을 통해 농산물 관련 생산자(단체)와 유통·가공업자의 의사결정에 활용할 수 있는 정보를 제공하고, 웹 검색 트래픽 정보를 활용하여 가격변화 예측모형 개선에 기여할 것으로 예상된다.

본 연구의 구성은 다음과 같다. 제 2장에서는 지리적 표시, 농산물 가격변동 요인, 농산물 가격예측 그리고 웹 검색 트래픽에 대한 선행연구를 요약하였고, 제 3장에는 분석에 사용된 데이터 및 방법론 그리고 로지스틱 회귀모형을 통해 지리적 표시제에 등록된 사과와 비등록된 사과의 가격 변화를 예측·분석한 결과를 설명하였다. 마지막으로 제 4장에 결론을 제시하였다.

## 2. 선행 연구

### 2.1 지리적 표시

지리적 표시란 상품의 원산지를 표시하는 동시에 원산지에 의존하는 제품의 품질 또는 특징에 관한 정보를 전달하는 표시이다. 이 표시는 해당 제품이 전통적인 지명과 더불어 소비자에게 인식되길 바라는 브랜드로서의 역할을 갖고 있다(Korean Intellectual Property Office, 2007). 또한, 지리적 표시는 역사적인 배경과 품질, 농산물에 대한 명성 등과 같은 요소들이 결합됨으로써 시장에서의 차별적 가치로 인식된다.

Kim and Sung(2008)은 선택형 실험(choice experiment)을 이용하여 성주참외를 대상으로 지리적 표시에 대한 소비자 가치를 측정하였다. 표본 소비자에 대한 설문조사 실시 후 다항로짓 모형(multinomial logit model)으로 선호도 유의성을 확인하고, 지리적 표시에 대한 지불의사금액을 계산하였다. 지리적 표시를 한 경우가 하지 않는 경우에 비해 소비자들의 선호도가 높게 나타났으며, 친환경 인증과 상표, 등급, 가격에 비해 매우 높은 가치를 부여하고 있음을 확인할 수 있었다. 이러한 결과는 성주참외의 경우 지리적 표시제에 의한 시장차별화가 달성되고 있다는 것을 반영한다.

Hong *et al.*(2009)는 사과의 지리적 표시제 효과를 알아보기 위해 사과의 가격, 거래단위, 사과 개수, 지리적 표시 등록 여부, 요일, 월, 품종 등과 같은 특성 변수들을 이용하여 회귀분석을 하였고, 사과의 특성과 속성에 대한 잠재가격을 추정하기 위해 헤도닉가격 모델을 사용하였다. 지리적 표시제를 시행함으로써 도매시장에서 사과의 경락가격이 향상되는 것으로 분석되었다. 즉 지리적 표시제 도입이 안정적으로 정착된다면 이로 인한 시장차별화 달성을 예견 하였다.

Kim and Kim(2010)은 컨조인트 분석기법을 통해 소비자의 구매선택행위에 미치는 지리적 표시의 효과를 측정하였다. 상당수의 일반 소비자들은 지리적 표시를 원산지 표시의 일종으로 여타 정부 보증 마크로 오인하여 정부 보증에 대한 신뢰의 차원에서 의미를 두고 구매의사결정에 고려하였고, 지리적 표시에 대해 정보를 제공한 후에는 전반적으로 이전보다 더 높은 가중치를 부여함을 확인할 수 있었다. 그러므로 정부 등의 공공부문이 원하는 정책적 효과를 거두기 위해서는 마크나 라벨 등을 부여하는 노력만큼이나 일반 소비자들을 대상으로 한 홍보에 보다 적극적인 노력을 기울여야 함을 시사한다.

### 2.2 농산물 가격 변동요인

농산물 가격은 연도 간 변동이 심할 뿐 아니라 같은 연도 내에서도 변동이 심하여 농민, 유통·가공업자 그리고 정책자가 의사결정 하는데 어려움이 있다. 하지만 변동요인을 분석하는 선행 연구는 많지 않다.

Yoon and Yang(2004)은 양념채소인 고추, 마늘, 양파(1991년 12월 31일~2003년 12월 31일까지의 일일 경락가격)를 분석

대상으로 하여 요일효과, 월별효과 및 월중효과를 검정하였다. 이 연구는 가격변화율을 종속변수로 하고 각 요일 및 월 더미를 설명변수로 하는 회귀분석을 하였다. 분석결과 요일효과는 양과, 월별효과는 마늘, 월중효과는 특정기간에서 마늘(1996~1999년), 양과(2000~2003년)에서만 나타남을 보였다. 즉, 계절효과가 해당 품목의 계절적 특성에 기인할 수도 있지만 양념 채소시장의 정보 비효율성에 기인할 수도 있음을 지적하였다.

Kang(2007)은 배추, 감자, 양과의 도매시장 경락가격과 반입량간의 상호관계를 VAR(Vector Autoregression) 을 이용하여 분석하였다. 배추의 경우 도매시장 가격이 도매시장 반입량을 주도하지만, 감자나 양과에 대해서는 오히려 반입량이 가격을 주도하는 것으로 나타났다. 저장성이 있는 양과와 감자는 출하자들이 가격형성에 어느 정도 영향을 주고 있으나, 배추의 경우 출하자들이 상대적으로 더 가격순응적임을 의미한다.

Kang(2008)은 GARCH-GJR 모형을 응용하여 서울 가락도 매시장에서 거래되는 신선채소류 중 배추, 무, 풋고추, 오이(취정)의 일일 반입량과 가격간에 유의적인 음의 관계가 있다는 것을 확인하였다. 또한, 오이를 제외하고 배추, 무, 풋고추의 경우, 반입량 변동성이 커질 때 가격 변동성이 작아지며 출하자들이 조직화하여 반입량을 변화시켜 반입량의 변동성에 영향을 줌으로 이는 출하량의 변동성이 가격의 안정화에 기여한다고 볼 수 있다.

Kim and Cho(2010)은 가락동 도매시장의 대과 가격을 대상으로 도매시장 반입량, 산지 생산량, 월중 자연재해 발생과 유통장애 등에 따른 시장 공급지연 그리고 연중 주기성 등의 가격변동 요인을 바탕으로 회귀분석 하였고, 자기상관 문제가 있는 경우 코크렌 오크트 반복추정법(Cochrane-Orcutt Method)를 이용하여 가격변동요인을 분석하였다. 그 결과, 대과의 가격은 월중 반입량보다는 산지 생산량에 영향을 받고, 공급지연 상황이 자주 발생할수록 저장성이 낮은 대과의 시장가격은 크게 증폭될 여지가 있다는 것을 밝혔다.

### 2.3 농산물 가격 예측

농산물 가격 예측에 관한 선행연구는 주로 시계열 분석법과 신경망 분석법을 통해 수행되었다.

Kohzadi *et al.*(1996)은 ARIMA 모형과 피드포워드 신경망(feedforward neural network) 모형을 통해 육우와 밀의 월별 가격 변화를 예측하였다. 가격 데이터는 United States Department of Agriculture(USDA)의 통계고시에서 1950년부터 1990년까지 정보를 이용하였는데, 1970년부터 6, 12, 18, 24개월의 시차를 적용하여 3년 단위로 7번 분석하여 1990년까지 예측하였다. 그 결과, 피드포워드 신경망 모형의 평균제곱오차(MSE)가 ARIMA 모형보다 약 27% 낮은 오차율을 보였다.

Kim(2005)은 선행연구에서 변동성이 높다고 평가한 마늘, 양과, 배추 무를 대상으로 가락시장 도매가격 자료(1987년 1월~2004년 12월까지의 월별자료)를 이용하여 ARIMA 모형, GARCH

모형, 인공신경망 모형(Artificial Neural Networks model, ANN)을 통하여 예측력을 비교하였다. 그 결과 시계열 모형인 ARIMA와 GARCH의 예측능력이 ANN보다 뛰어난 것을 보였으며, ARIMA와 GARCH 두 모형간의 예측력에는 큰 차이가 없음을 보였다.

Lee *et al.*(2006)의 연구는 청과물 14품목의 가격의 계절성을 고려하고 1996년 1월부터 2005년 12월까지 가락시장의 평균 가격을 통해 가격예측을 하였다. ARIMA 모형, GARCH 모형 등 자체 가격 시계열만으로 구성되는 단변량 모형을 추정하였고, 복수의 가격을 포함하거나 물량 변수를 함께 포함하는 단변량 모형으로 VAR 모형을 추정하였다. 또한, 시계열 모형과 회귀모형(가격신축성함수)의 결합을 위한 시도로 공급량을 외생변수로 한 일종의 전이함수 모형을 추정하였다. 품목마다 예측력이 높은 모형은 다르나 예측오차가 15% 이하로 대체로 양호한 모형을 발견할 수 있었다.

Zou *et al.*(2007)은 중국 정저우 곡물 도매시장의 월별 밀 가격 예측을 ANN 모형, ARIMA 모형 그리고 ANN과 ARIMA를 결합한 모형 통해 예측력을 비교하였다. 그 결과, ANN 분석 모형의 R-square가 0.735로 ANN(0.0572)과 결합한 모형(0.477)보다 더 좋은 예측성능을 보였다.

Lou *et al.*(2011, 2013)은 농산물 가격예측과 관련하여 베이징 내의 도매시장의 버섯을 대상으로 2003년부터 2007년까지 월별 가격 데이터를 사용하여 분석하였는데, genetic algorithm neural network 모형이 traditional BP neural network(BPNN) 모형과 Radial basis function network(RBFN) 모형보다 더 예측 성능이 높았다. 또한, 도매시장 양배추의 3일씩 평균 낸 가격과 월별 도매시장 오이 가격을 예측하였는데, 중국 칭타오 남쪽 양배추의 경우 PSO-RBFNN(Particle Swarm Optimization and Radial Basis Function Neural Network) 모형이, 오이의 경우 SARIMA(seasonal auto regressive integrated moving average) 모형이 좋은 모형으로 나타났다.

Zhang and Xie(2011)은 L-M 알고리즘을 사용한 BP neural network 모형을 바탕으로 중국의 광저우 바이윈 구의 월별 양배추가격을 예측하였다. 이 연구는 농산물 가격만을 반영한 다른 연구들과 다르게 기온, 일조량, 강수량 등의 날씨변수, 조립비, 유통비, 판매비 등의 비용변수 그리고 계절변수를 고려하여 예측하였다. 그 결과, 실제값과 예측값의 최대차이가 0.4 위안으로 적게 나타남을 보였다.

Nasira and Hemageetha(2012, 2013)는 2009년 1월부터 2011년 5월까지의 인도 코임바토르 시장의 토마토 가격 데이터를 이용하여 BP neural network 모형을 통해 주별과 월별 토마토 가격을 예측하였다. 그 결과, 주별 가격예측 모형에서 75%, 월별 가격예측 모형에서 90%의 정확도를 보였다. 그 다음해엔 동일한 데이터로 주별 토마토 가격 예측을 위해 BPNN와 RBFN 모형을 이용하여 비교·분석하였는데, 정확도가 각각 77.42%, 85.55%로 RBFN의 예측능력이 BPNN보다 뛰어난 것을 보였다.

이처럼 대부분의 농산물 가격 예측 연구들은 여러 모델들 예측력 향상에 목적을 두고 있다. 하지만, 농산물 가격과 관련

된 요일, 월별, 농산물 물량 등의 요인이 있음에도 불구하고, 해당 농산물 과거 가격 자체만을 변수로 반영하여 사용한 한계점이 있다.

## 2.4 웹 검색 트래픽

최근 빅데이터 분석을 통해 방대한 양의 데이터 속에서 가치를 창출하려는 시도가 여러 분야에서 일어나고 있다(Yoon, 2013; Kim *et al.*, 2014; Lee and Sohn, 2015). 그 중, 구글 트렌드, 네이버 트렌드가 사람들이 무의식적으로 남긴 정보인 웹 검색 트래픽 정보를 표준화한 검색 통계를 실시간으로 공급 하면서 이를 활용한 여러 가지 연구가 진행되었다. 검색 트래픽을 활용한 연구의 장점은 실시간에 가깝게 동향을 분석할 수 있다는 것으로 이는 먼 미래인 장기적 예측보다는 중단기적인 가까운 미래를 예측하는데 강점을 보이고 있다.

Ginsberg *et al.*(2009)는 구글의 검색엔진에서 얻은 데이터를 통해서 독감 유행 수준을 예측하였다. 가공하기 전의 검색 데이터를 이용하여 미국 질병통제센터(Centers for Disease Prevention and Control)에서 발표되는 기존 보고서보다 1주에서 2주 정도 더 빨리 독감 바이러스의 활성을 예측하는 실시간 감시 시스템을 제시하였다. 이 연구가 의미하는 점은 검색 트래픽이 기존 보고서 보다 빠르게 변화나 현상을 감지할 수 있어 독감 확산을 모니터링 하는데 도움이 되고, 평일 수백만 명이 사용하는 방대한 결과를 활용하기 때문에 모집단의 대표성을 가지고 있는 것이다. 그리고 검색 트래픽과 사회적 현상 간에 높은 상관관계가 있어 다양한 다른 현상 예측에 활용할 수 있는 가능성이 높다는 점이다. 이 연구를 시작으로 웹 검색 트래픽을 활용한 다양한 예측가능성이 세상에 알려졌다.

Choi and Varian(2012)는 검색 트래픽을 자동차 판매, 실업수당 청구, 여행 방문자 수, 소비자 신뢰지수와 같은 수요 예측을 위해 이용이 가능하다고 언급했다. 그들은 먼 미래가 아닌 가까운 미래를 예측하는 것이 유용할 때가 있다고 주장하였다. 만일 특정 장소의 “부동산 중개업소”에 관한 검색이 늘어나고 있다면, 이는 가까운 미래에 이 지역에서 주택판매가 늘어날 수 있음을 생각을 할 수 있다는 것이다. 이 연구는 종종 검색 트래픽과 여러 경제지표와의 상관관계가 있으며, 단기간의 경제적 예측에 도움이 것이라고 하였다.

Lui *et al.*(2011)은 검색 트래픽이 다른 잠재적 용도로서 하나의 예가 될 수 있음 제안하였다. 구글 트렌드의 검색 트래픽을 통해 2008년과 2010년 미국 의회선거에서 후보자의 당선 가능성을 연구하였다. 하지만 전통적인 여론조사에 비해 33.3%와 39%라는 낮은 예측력을 보여주었는데, 이는 긍정적인 정보 검색 뿐만 아니라 부정적인 검색도 포함되기 때문이라고 주장하였다.

Vosen and Schmidt(2011)는 자기회귀 모형을 통해 미국에서 소비 동향을 파악할 수 있는 지표로 미시건 대학이 발표하는 소비자 심리지수 MCSI(University of Michigan Consumer Sentiment Index)와 소비자 신뢰지수(Consumer Confidence Index)를

구글 트렌드가 제공하는 검색 트래픽의 결과와 비교하여 연구하였다. 그 결과, 구글이 제공한 검색 트래픽의 예측력이 설문조사 기반의 지수보다 우월함을 보여주었고, 구글 트렌드 데이터는 이미 소비자 지출을 전망할 수 있는 잠재력이 높다는 것을 증명하였다.

Jun *et al.*(2013)은 ARIMA 모형을 통해 자동차 프리우스의 판매량 예측을 하였다. 제품 브랜드 검색 트래픽과 그 속성에 대한 검색 트래픽이 기존의 특허나 뉴스와 같은 서지분석 변수보다 판매량을 설명하는데 상당한 영향을 미치고 있음을 밝혔다.

## 3. 실증 분석

### 3.1 데이터

이번 장에서는 농산물 지리적 표시가 가격변화 예측에 미치는 영향을 분석하기 위해 사용된 데이터와 방법론을 소개하였다.

서울시농수산물공사가 운영하는 서울가락시장은 한국에서 가장 큰 공영도매시장으로서 전국에서 생산되는 농수산물이 유통되며, 공휴일을 제외한 산지별 일별 농산물 가격과 물량을 홈페이지(<http://www.garak.co.kr>)에 게시하고 있다. 산지에서 서울가락시장으로 반입되는 거래는 산지에 따라 매일 이루어지지 않기 때문에 본 연구는 거래빈도가 높은 산지 중, 지리적 표시에 등록된 제 23호 충주사과와 제 34호 청송사과 그리고 지리적 표시에 등록되어있지 않은 영천사과를 선택했으며 이 3곳의 가격 변화 예측을 위해 사과품종 중 생산량이 가장 많은 후지 15kg 상자의 최고가격과 총 물량을 일별 데이터로 제공받아 사용했다. 충주지역에 지리적 표시제를 등록 받은 등급은 ‘특품’이며 충주의 경우에는 ‘특품’과 ‘상품’으로 산지별 등급 가격을 사용해야하나, 이에 대한 자료가 없으므로 산지별 최고가격이 산지별 ‘특품’과 ‘상품’의 가격과 같다고 가정하였다. 변수들의 크기를 맞추기 위해 최고가격은 원/kg 단위에서 1,000을 나눈 천원/kg 단위를 적용하였고, 총 물량은 kg단위에서 15를 나누어 총 사과 박스 단위로 환산 후 다시 100을 나눈 값을 사용하였다.

웹 검색 트래픽 데이터는 구글과 네이버 두 곳에서 제공하는데, 우리나라의 경우 검색엔진 중 네이버의 점유율이 80% 이상으로 구글 점유율 5% 미만인 것에 비해 월등히 높아 네이버의 웹 검색 트래픽인 네이버트렌드(<http://trend.naver.com>)의 데이터를 이용하였다. 구글 트렌드 데이터는 모든 검색 통계 결과에서 환경적 영향을 없애기 위해 총량으로 나누는 표준화 과정을 거친다. 하지만, 네이버의 경우 검색기간 내의 전체 검색량을 표준화시키지 않고 각각의 검색어에 대한 최소값(0)과 최대값(100)으로 환산된 결과값을 제공한다(Jun *et al.*, 2013). 본 연구에 사용된 웹 검색 트래픽은 사과가 1년에 한 번 수확되어 다음 해 수확기까지 물량이 소모되는 특성에 따라 검색기간을 연 단위로 한 후 주별 자료를 이용하였다. 충주·청송·영천 세 지역 모두 주요 사과 생산지이므로 네이버 검색

단어 ‘지역명’, ‘지역명+사과’와 사과 검색시 연관성 높은 단어 ‘사과’, ‘사과가격’, ‘사과 직거래’의 웹 검색 트래픽을 사용했다. 이 웹 검색 트래픽 데이터는 1주 전부터 4주 전까지의 시차를 반영한 파생변수를 생성하여 사용했다. 수확기(10월 말부터 11월 말까지)와 비수확기에 따른 더미변수를 추가했다.

단기 가격 분석을 위해 일별 데이터를 주 단위로 평균낸 주별 가격을 사용했다. 모형추정에 사용된 모든 자료는 2007년 1월 2일부터 2013년 12월 31일까지의 자료를 사용하였으며, re-

sampling 방법인 Jackknife 방법을 전체 데이터에 적용하여 이를 예측에 반영했다. 그러나, <Figure 1>과 같이 사과 물량거래가 없는 시점부터 사과가 수확하기 시작하는 10월 말 직전의 시점까지는 사과가격이 없으므로 각 모형에서 제거하여 사용했다. <Figure 2>는 산지에서 도매시장으로 반입되는 물량 그래프로 수확기 이후에 사과(후지) 물량이 급격히 증가하다가 3월 이후부터 점점 줄어드는 것을 확인할 수 있다. 웹 검색 트래픽은 아래 <Figure 3>과 같은 추세를 보인다.

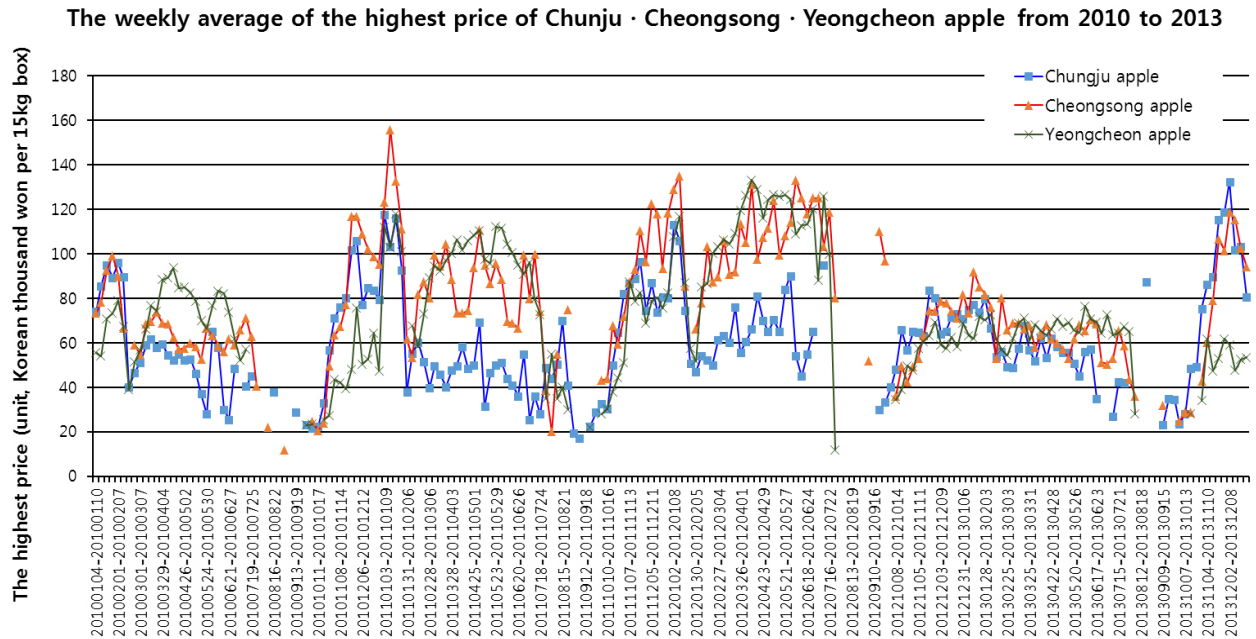


Figure 1. The weekly average of the highest price of Chungju · Cheongsong · Yeongcheon apple from 2010 to 2013 : (unit : korean thousand won per 15kg box)

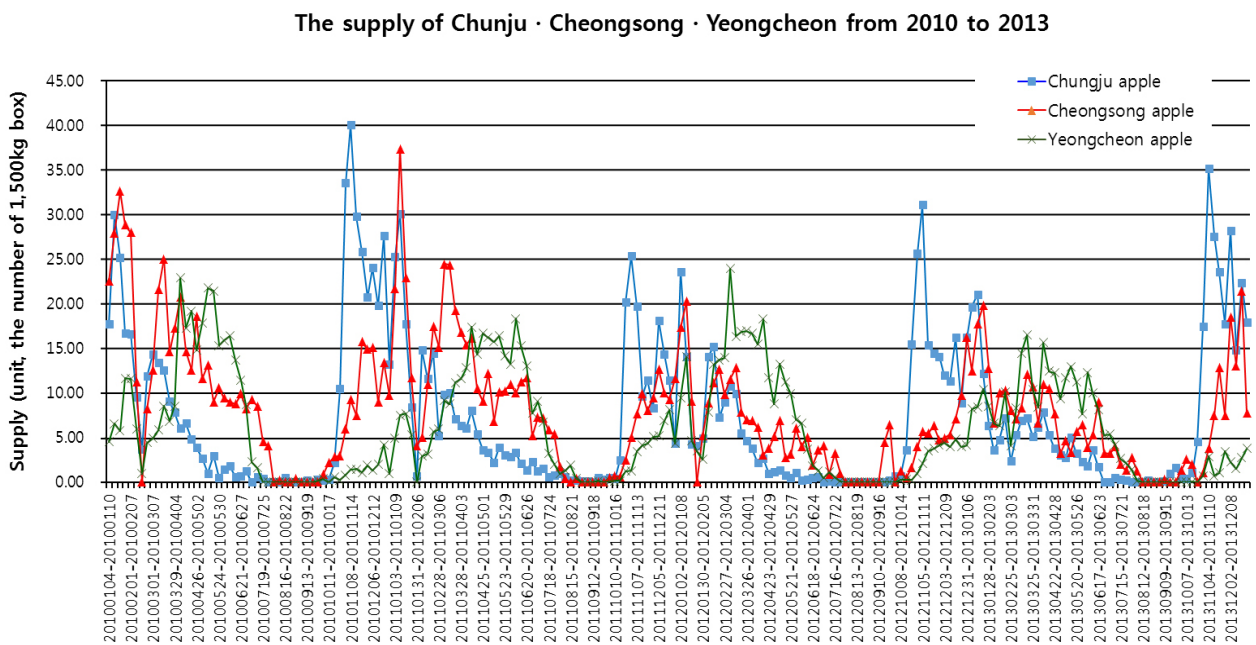


Figure 2. The supply of Chungju · Cheongsong · Yeongcheon apple from 2010 to 2013 : (unit : the number of 1,500kg box)

The normalized web search frequency of keyword from 2010 to 2013

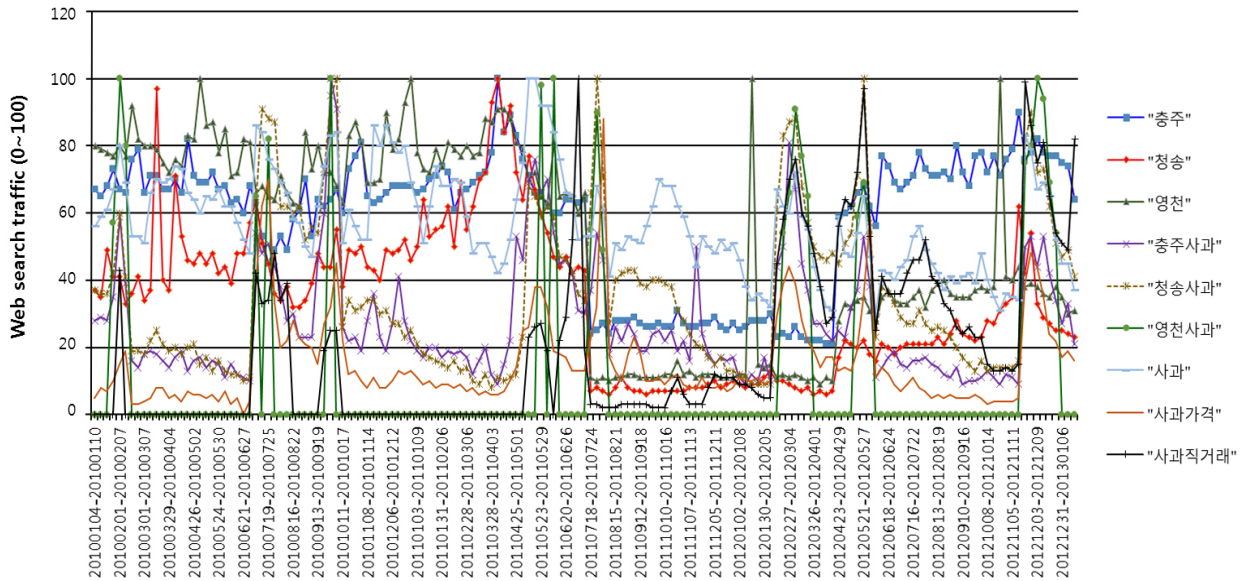


Figure 3. The normalized web search frequency of keyword from 2010 to 2013

앞에서 기술한 바와 같이 본 연구에서 지리적 표시제 등록 <Table 1>과 같으며, 독립변수의 Time lag(n)은 4로 설정하여 유무에 따른 농산물 가격변화 예측을 위해 사용된 변수는 분석하였다.

Table 1. Definition of variables

Model	Name	Variable	Data type	Variable type
Common variable	AT <sub>(t-n)</sub>	(t-n)시점의 웹 검색어 '사과'의 트래픽	Interval	Independent variable
	APT <sub>(t-n)</sub>	(t-n)시점의 웹 검색어 '사과가격'의 트래픽	Interval	Independent variable
	DMT <sub>(t-n)</sub>	(t-n)시점의 웹 검색어 '사과직거래'의 트래픽	Interval	Independent variable
	HAR	사과 수확기 유무(0 = 비수확기, 1 = 수확기)	Binary	Independent variable
Price variation model of Chungju apple	CJP <sub>(t)</sub>	t시점의 충주사과 최고가격0 = (t)시점 가격이 (t-1)시점 가격에 비해 감소한 경우, 1 = (t)시점 가격이 (t-1)시점 가격에 비해 증가한 경우	Binary	Dependent variable
	CJP <sub>(t-n)</sub>	(t-n)시점의 충주사과 최고가격	Interval	Independent variable
	CJS <sub>(t-n)</sub>	(t-n)시점에 충주사과 물량	Interval	Independent variable
	CJT <sub>(t-n)</sub>	(t-n)시점의 웹 검색어 '충주'의 트래픽	Interval	Independent variable
	CJAT <sub>(t-n)</sub>	(t-n)시점의 웹 검색어 '충주사과의 트래픽	Interval	Independent variable
Price variation model of Cheongsong apple	CSP <sub>(t)</sub>	t시점의 청송사과 최고가격0 = (t)시점 가격이 (t-1)시점 가격에 비해 감소한 경우, 1 = (t)시점 가격이 (t-1)시점 가격에 비해 증가한 경우	Binary	Dependent variable
	CSP <sub>(t-n)</sub>	(t-n)시점의 청송사과 최고가격	Interval	Independent variable
	CSS <sub>(t-n)</sub>	(t-n)시점의 청송사과 물량	Interval	Independent variable
	CST <sub>(t-n)</sub>	(t-n)시점의 웹 검색어 '청송'의 트래픽	Interval	Independent variable
	CAST <sub>(t-n)</sub>	(t-n)시점의 웹 검색어 '청송사과의 트래픽	Interval	Independent variable
Price variation model of Yeongcheon apple	YCP <sub>(t)</sub>	t시점의 영천사과 최고가격0 = (t)시점 가격이 (t-1)시점 가격에 비해 감소한 경우, 1 = (t)시점 가격이 (t-1)시점 가격에 비해 증가한 경우	Binary	Dependent variable
	YCP <sub>(t-n)</sub>	(t-n)시점의 영천사과 최고가격	Interval	Independent variable
	YCS <sub>(t-n)</sub>	(t-n)시점의 영천사과 물량	Interval	Independent variable
	YCT <sub>(t-n)</sub>	(t-n)시점의 웹 검색어 '영천'의 트래픽	Interval	Independent variable
	YCAT <sub>(t-n)</sub>	(t-n)시점의 웹 검색어 '영천사과의 트래픽	Interval	Independent variable

농산물 가격은 수요, 공급, 시장, 정책, 기후 등 매우 다양한 요인에 영향을 받아 변화하므로 정확한 가격을 예측하는 것은 매우 어렵다. 그리하여 본 연구는 분석대상인 사과 가격의 증감 여부를 예측하는 모형을 개발하고자 로지스틱 모형을 사용했다. 이 때 독립변수를 선택하기 위한 방법으로 전진선택법(forward selection)과 후진선택법(backward selection)은 변수가 선택되거나 제거되었을 때 그 변수에 관해서는 다시 고려되지 않는 단점이 있어 단계적선택법(stepwise selection)을 이용했다. 로지스틱 회귀분석은 선형회귀분석과 같이 독립변수들과 종속변수들간의 인과관계를 분석하는 통계기법이며, 특히 종속변수가 이분형 범주만을 가질 때 사용된다. 로지스틱 회귀분석은 종속변수와 관련하여 의미있는 변수들을 추출하는데 활용된다(Kim et al., 2010; Park et al., 2014; Oh et al., 2014).

3.2 분석 결과

분석 결과 <Table 2>을 보면, 충주사과 가격 변화여부 모형에서는 (t-1)시점의 충주 사과가격, (t-2)시점의 충주 물량, (t-1)시점의 웹검색어 ‘충주사과’의 트래픽, 수확기 더미변수 그리고 청송사과모형에서는 (t-1)시점의 청송 산지가격, (t-4)시점의 청송 산지가격, (t-3)시점의 청송 물량, (t-1)시점의 웹검색어 ‘청송’과 ‘사과’의 트래픽 변수와 유의한 관계가 있는 것으로 나타났다. 마지막으로 영천사과모형에서는 (t-1)시점과 (t-2)시점의 영천 산지가격 변수가 사과 가격변화에 유의한 관련이 있는 것으로 나타났다.

세 사과모형 모두 과거 가격이 현재 가격변화에 유의한 예측 변수임을 알 수 있다. 공통적으로 (t-1)시점의 사과가격이 현재 가격과 음(-)의 관계가 있어, 이는 한 주전 가격이 증가하면 그 다음주 가격은 감소하거나 한 주 전 가격이 감소하면 그 다음주 가격은 증가함을 나타낸다. 또한 청송사과모형의 (t-4)시점 청송사과가격과 영천사과모형의 (t-2)시점 영천사과가격이 증가하면 각 모형의 (t)시점의 사과가격이 증가하는데 관계가 있음을 보였다. 영천의 (t-1)시점의 사과가격과 (t-2)시점의 사과가격 또한 충주사과와 같은 현상을 보인다.

물량은 충주사과모형과 청송사과모형에서 유의하나, 영천사과모형에서는 유의하지 않은 것으로 나타났다. 충주와 청송산지에서 가락도매시장으로 반입되는 물량이 영천에 비해 많기에 물량 변화에 따라 가격에 관련이 있는 것으로 분석된다.

웹 검색 트래픽 변수의 경우, 지리적 표시제에 등록된 충주와 청송 사과에서 지역명이 포함된 변수들이 산지 사과가격 변화에 관련이 있으나, 검색어 단어 형태가 다르게 나타났다. 충주사과모형에서 ‘지역명+사과’ 형태인 (t-1)시점의 ‘충주사과’ 트래픽이 양(+)의 관계가 있어 충주사과 가격변화가 지리적 표시제와 관계가 있음을 보인다. 청송사과모형에서는 ‘지역명’ 형태인 (t-1)시점의 ‘청송’ 트래픽이 양(+)의 관계 그리고 지리적표시 등록 품목인 (t-1)시점의 ‘사과’ 트래픽이 음(-)의 관계를 가지고 있다. 하지만, ‘청송사과’ 트래픽 변수와는

Table 2. Logistic regression model for weekly apple price increase

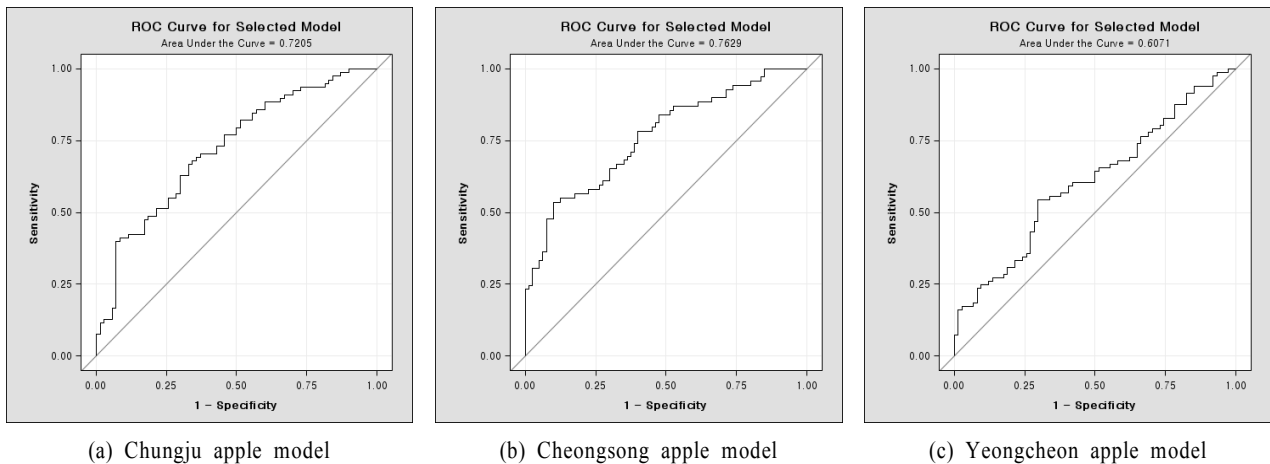
Model	Variable	Odds Ratio	Coefficient	P-value
Price variation model of Chungju apple	Intercept		3.0829	0.0007 <sup>***</sup>
	CJP <sub>(t-1)</sub>	0.957	-0.0437	0.0007 <sup>***</sup>
	CJS <sub>(t-2)</sub>	1.07	0.0681	0.0296 <sup>**</sup>
	CJAT <sub>(t-1)</sub>	1.019	0.019	0.0947 <sup>*</sup>
	HAR	0.211	-1.5551	0.0125 <sup>**</sup>
Price variation model of Cheongsong apple	CSP <sub>(t-1)</sub>	0.955	-0.0458	<.0001 <sup>***</sup>
	CSP <sub>(t-4)</sub>	1.027	0.0263	0.0122 <sup>**</sup>
	CSS <sub>(t-3)</sub>	0.851	-0.1619	0.0002 <sup>***</sup>
	CST <sub>(t-1)</sub>	0.984	-0.0164	0.0986 <sup>*</sup>
	AT <sub>(t-1)</sub>	1.065	0.0631	0.0003 <sup>***</sup>
Price variation model of Yeongcheon apple	YCP <sub>(t-1)</sub>	0.967	-0.0338	0.0114 <sup>**</sup>
	YCP <sub>(t-2)</sub>	1.022	0.0216	0.0937 <sup>*</sup>

<sup>\*</sup>, <sup>\*\*</sup>, <sup>\*\*\*</sup> indicate statistical significance at the significance level  $\alpha = 10\%$ ,  $5\%$ ,  $1\%$  respectively.

관계가 없는 것으로 나타나 지리적 표시제에 대한 관심여부에 상관없이 지역명 과 사과에 대한 관심만으로도 청송사과 가격변화 예측이 가능한 것으로 보여진다. 농촌진흥청의 사과구매패턴 조사결과에 따르면, 지역 사과브랜드 중 청송사과 구매빈도 및 구매액이 월등하게 높은 것으로 조사되었다. 이렇듯 청송사과는 소비자에게 인지도 및 영향력을 가지고 있으므로 청송 대표 특산물인 ‘사과’와 지역명인 ‘청송’에 대한 웹 검색 트래픽이 청송사과가격과 유의한 관계를 보이는 것으로 해석된다. 영천사과가격 변화 모형의 경우에는 웹 검색 트래픽과 영천사과 가격변화와의 유의한 관계가 없는 것으로 나타나 지리적 표시제에 등록되지 않은 것뿐만 아니라 인지도 및 선호도가 높지 않은 것을 반영 한 것으로 보여진다. 참고로 검색어 ‘사과’, ‘영천’, ‘영천사과’의 트래픽 변수들간의 상관관계가 0.5 미만으로 낮으므로, 이들간의 높은 상관관계에 의해 가격변화에 유의한 관계가 없는 것으로 나온 것이 아님을 확인할 수 있다.

나머지 ‘사과가격’과 ‘사과직거래’의 웹 검색 트래픽은 세 지역 모두 사과가격변화와 유의한 관계가 없는 것으로 나타났다. 그리고 유의한 수확기 더미 변수는 충주사과 가격변화에서만 관계가 있는 것으로 나타났는데, 이는 수확기 직후 충주사과가격이 급증하여 최고가격을 기록하고 점점 감소하는 현상을 나타내는 것으로 보인다.

추정된 모형의 적합도를 알아보기 위해 ROC curve 검증 결과를 살펴보았다. ROC curve는 아래의 <Figure 4>과 같다. 충주사과모형의 ROC curve 아래의 면적(area under the roc curve, AUC)은 0.7205, 청송사과모형의 AUC는 0.7629 그리고 영천사과모형의 AUC는 0.6071로 나타났다. 이 때 AUC가 1에 가까울수록 예측모형의 성능이 좋음을 나타내며, 충주사과모형과 청송사과모형은 우수한 예측 정확도를 보이나, 영천사과모형의 경우 비교적 낮은 값을 보인다.



(a) Chungju apple model

(b) Cheongsong apple model

(c) Yeongcheon apple model

Figure 4. ROC curve of apple model

더불어 지리적표시제 관련 웹 검색 트래픽이 가격변동예측에 주요한 영향을 미치는지 알아보고자, 앞서분석한 충주사과 모형에서 지리적표시제 관련 웹 검색 트래픽을 제거하여 로지스틱 회귀분석을 하였다. 그 결과, AUC가 0.7205에서 0.7040으로 0.0165 감소하였다. 이는 지리적표시제 관련 웹 검색 트래픽 변수 없이 사과가격과 물량 변수만으로도 비교적 가격변화 예측모형을 설명할 수 있다는 것을 나타낸다. 즉, 지리적표시제에 대한 사람들의 관심이 사과가격변화에 미치는 영향이 미비하다는 것을 알 수 있다.

#### 4. 결론

지리적 표시제는 품종특허, 품종보호권, 상표권 등과 같은 지적재산의 하나이며, 지리적 특성을 가진 농수산물·농수산가공품의 품질 향상과 지역특화산업 및 소비자 보호를 위한 지리적 표시 등록제도이다. 2002년 보성녹차가 제1호로 선정된 이후 지리적 표시제에 대해 많은 연구가 수행되었다. 하지만 지리적 표시제 농산물 가격 변화에 대한 연구는 없었다.

본 연구에서는 사과의 산지별 가격 자료를 사용하여 지리적 표시제에 등록된 농산물과 비등록된 농산물 가격 변화에 관련 있는 요소를 비교·분석하고 예측하기 위하여 로지스틱 회귀분석을 이용하였다. 분석대상으로는 지리적 표시 농산물 제 23호 충주사과와 제 34호 청송사과 그리고 지리적 표시에 등록되어있지 않은 영천사과 총 3가지를 선택하였고, 사과가격 변화 예측 분석에 산지별 사과(후지) 최고가격과 총물량, 웹 검색 트래픽, 수확기와 비수확기에 따른 더미가 사용되었다.

로지스틱 회귀분석 모형에 따른 연구 결론은 다음과 같다.

첫째, 기존 농산물 가격예측 선행연구에서 농산물 가격에 영향을 주는 변수로 사용된 과거 자신의 농산물 가격과 산지에서 도매시장으로 반입되는 물량이 지리적 표시제 등록된 농

산물 가격 변화와 관계가 있음을 확인하였다.

둘째, 지리적 표시제 등록된 품목의 가격이 소비자의 관심을 나타내는 웹 검색 트래픽과 관계가 있는 것으로 나타났다. 하지만, 지리적 표시제에 등록된 품목이라고 하여 해당 품목 가격변화와 지리적 표시제에 관련된 웹 검색 트래픽과 관련이 있는 것이 아님을 확인할 수 있었다. 본 연구에서 충주와 청송 사과가격 변화예측 모형이 충주와 청송 두 곳 모두 지리적 표시제에 함께 등록되어있지만, 충주의 경우는 지리적 표시제 관련된 ‘충주사과’의 웹 검색 트래픽과 관계가 있는 반면 청송의 경우에는 ‘청송’, ‘사과’의 웹 검색어 트래픽이 유의한 것으로 나타나 지리적 표시제 명칭 없이 지역명과 특산물이 청송사과 가격변화와 관계가 있음을 알 수 있다.

셋째, 지리적 표시제 등록된 품목의 가격변화 예측을 위해 웹 검색 트래픽을 사용할 경우, 사용하지 않은 경우보다 예측력이 미비하게 향상되었다. 이는 아직 지리적표시제에 대한 소비자 인지도가 낮기 때문으로 보여지며, 2013년 농림수산식품교육문화원이 실시한 지리적표시제 실태조사에서도 지리적표시제의 소비자 인지도는 35.4%에 불과한 것으로 나타나 이를 뒷받침한다. 그러나, 향후 지리적표시제의 홍보 및 마케팅을 강화하여 인지도를 높인다면 가격 변동이 크고 예측하기 어려운 농산물 가격 예측 연구에 새로운 변수로 활용될 수 있을 것으로 기대된다.

이를 정리하면, 지리적 표시제 시행한 지역임에도 ‘지리적 표시제명’의 웹 검색 트래픽이 사과가격과 관련이 없는 것으로 나타났다. 그리하여 소비자들이 지리적 표시제에 따른 구매의사결정에 영향을 미칠 수 있도록 지리적 표시제 홍보전략 강화 필요성을 제기한다.

본 연구는 지리적 표시가 적용되는 품종 중 ‘사과’라는 품목에 한해서 분석을 수행한 것으로, 여러 품목에 대해 고려하지 못한 한계를 지니고 있다. 또한, 사과(부사)는 수확기를 지나 저장물량이 점점 감소하여 물량거래가 없는 기간이 나타나 시차를 적용하는데 어려움이 있으므로 이를 조정할 수 있는 연구가 필요하다.



## 참고문헌

- Beck, N.-G. (2012), The Effect on preference cognition and Purchase intention for geographical indication collective mark of agricultural products, *Korean Journal of Customer Satisfaction Management*, **14**(3), 103-127.
- Choi, H. and Varian, H. (2012), Predicting the present with Google Trends, *Economic Record*, **88**(s1), 2-9.
- Ginsberg, J., Mohebbi, M. H., Patel, R. S., Brammer, L., Smolinski, M. S., and Brilliant, L. (2009), Detecting influenza epidemics using search engine query data. *Nature*, **457**(7232), 1012-1014.
- Guo, Q., Luo, C. S., and Wei, Q. F. (2011), Prediction and research on vegetable price based on genetic algorithm and Neural network model. *Asian Agricultural Research*, **3**(5), 148-150.
- Hemageetha, N. and Nasira, G. M. (2013), Radial basis function model for vegetable price prediction, *Informatics and Mobile Engineering (PRIME)*, 2013 International Conference on *Pattern Recognition*, **IEEE**, 424-428.
- Hong, N.-G., Kim, H.-H., and Kim, T.-G. (2009), Effect of Korea Protected Geographical Indication on the Wholesale Price in Garak Market : Focused on Apple, *Korean Journal of Agricultural Management and Policy*, **36**(2), 294-310.
- Jun, S.-P., Kim, Y.-E., and Yoo, H.-S. (2013), A Comparative Study of Consumer's Hype Cycles Using Web Search Traffic of Naver and Google, *Journal of Korea Technology Innovation Society*, **16**(4), 1109-1133.
- Jun, S.-P., Yeom, J., and Son, J.-K. (2014), A study of the method using search traffic to analyze new technology adoption. *Technological Forecasting and Social Change*, **81**, 82-95.
- Kang, T.-H. (2007), An Analysis on the Relationship between Prices and Trading Volumes of Agricultural Products, *Korean Journal of Agricultural Economics*, **48**(4), 45-67.
- Kang, T.-H. (2008), The Influences of Volume Volatilities on Wholesale Price Stabilities for Fresh Vegetables, *Korean Journal of Agricultural Economics*, **49**(1), 21-38.
- Kim, B.-S. (2005), A Comparison of Forecasting Performance of the Application Models for Forecasting of Vegetable Prices, *Korean Journal of Agricultural Economics*, **46**(4), 89-113.
- Kim, B.-S. and Cho, J.-H. (2010), A Review on the Volatility of Green Onion wholesale Price, *Korean Journal of Agricultural Management and Policy*, **37**(4), 747-758.
- Kim, J.-K. and Kim, H.-H. (2010), The Impact of Geographical Indications on Brand Credibility of Fruits, *Korean Journal of Agricultural Economics*, **51**(4), 1-24.
- Kim, S. W., Kim, G. G., and Yoon, B. K. (2014), A Study on a Way to Utilize Big Data Analytics in the Defense Area, *Journal of the Korean Operations Research and Management Science Society*, **39**(2), 1-20.
- Kim, T.-G. and Sung, M.-H. (2008), Measuring the Consumer's Value for Geographical Indication, *Korean Journal of Food Marketing*, **25**(1), 27-41.
- Kim, Y.-S., Yum, B.-J., and Kim, N. (2010), Development of a Recommender System for E-Commerce Sites Using a Dimensionality Reduction Technique, *Journal of the Korean Institute of Industrial Engineers*, **36**(3), 193-202.
- Lee, Y.-S., Choi, B.-O., and Shim, S.-B. (2006), A Time Series Analysis on Prices of Fruits and Vegetables, *Research Report of Korea Rural Economic Institute*, **R537**, 1-104.
- Lee, W.-S. and Sohn, S.-Y. (2015), Topic Model Analysis of Research Trend on Spatial Big Data, *Journal of the Korean Institute of Industrial Engineers*, **41**(1), 64-73.
- Lui, C., Metaxas, P. T., and Mustafaraj, E. (2011), On the predictability of the US elections through search volume activity, In *Proceedings of the IADIS International Conference on e-Society*.
- Luo, C., Wei, Q., Zhou, L., Zhang, J., and Sun, S. (2011), Prediction of vegetable price based on Neural Network and Genetic Algorithm, In *Computer and Computing Technologies in Agriculture IV*, Springer Berlin Heidelberg, 672-681.
- Luo, C. S., Zhou, L. Y., and Wei, Q. F. (2013), Application of SARIMA Model in Cucumber Price Forecast, *Applied Mechanics and Materials*, **373**, 1686-1690.
- Kim, H.-H., Ha, S.-G., Oh, H.-S., Kim, T.-G., Ji, H.-Y., Park, S.-S., Song, W.-G., and Lee, D.-M. (2009), An analysis about agri-food value creation of the geographical indication and effect on rural development, *Report of Ministry for Food, Agriculture, Forestry and Fisheries*, 1-288.
- Oh, Y.-H., Yun, J.-S., and Lee, J.-S. (2014), Using Data Mining Techniques to Predict Win-Loss in Korean Professional Baseball Games, *Journal of the Korean Institute of Industrial Engineers*, **40**(1), 8-17.
- Panel Insight and Canaan Coaching Consulting (2013), Consumer's Awareness and Purchase Actual Condition Research about Fruits and vegetables, 1-42.
- Park, K.-H., Hou, T., and Shin, H.-J. (2011), Oil Price Forecasting Based on Machine Learning Techniques. *Journal of the Korean Institute of Industrial Engineers*, **37**(1), 64-73.
- Vosen, S. and Schmidt, T. (2011), Forecasting private consumption : survey based indicators vs. Google trends, *Journal of Forecasting*, **30**(6), 565-578.
- Xu, Q. and Liu, M. (2013), Simulation and Forecast about Vegetable Prices Based on PSO-RBFNN Model, *Proceedings of 2013 Chinese Intelligent Automation Conference*, Springer Berlin Heidelberg, 255-260.
- Yoon, B.-S. and Yang, S.-R. (2004), A study on the Day-of-the-Week, Monthly and Intra-Month Effects in Spicy Vegetable Prices, *Korean Journal of Agricultural Economics*, **45**(2), 187-210.
- Yoon, M.-Y. (2013), Analysis on Big Data Policy of Major country and Implication, *Science and Technology Policy*, **23**(3), 31-43.
- Zou, H. F., Xia, G. P., Yang, F. T., and Yang, H. (2007), A neural network model based on the multi-stage optimization approach for short-term food price forecasting in China, *Expert Systems with Applications*, **33**(2), 347-356.