

시계열 분해와 앙상블을 활용한 배추 가격 예측 모형의 예측 성과 분석*

한은수**, 윤성욱***, 윤병삼****

Contents

1. 서론	3
2. 연구 방법 및 절차	5
3. 분석 결과	13
4. 요약 및 결론	23

Keywords

배추(Kimchi Cabbage), 가격 예측(Price Forecasting), 시계열 분해(Time Series Decomposition), 예측 조합(Forecast Combination)

Abstract

본 연구는 배추 월별 도매가격의 단기 예측 성과를 개선하기 위해 시계열 분해와 예측 조합을 결합한 분해-앙상블(decomposition-ensemble) 예측 알고리즘을 제안하고, 예측 성과를 평가하였다. 로그 변환한 배추 도매가격 시계열을 STL 분해를 통해 계절 성분과 계절조정 성분으로 구분하고, 계절 성분에는 계절성 단순 예측법(seasonal naive)을, 계절조정 성분에는 ETS 지수평활, ARIMA, ARIMAX, 신경망 자기회귀(neural network autoregression) 모형을 적용한 뒤, 개별 예측치를 단순 평균, 중앙값, 평균제곱오차(MSE) 기반 가중 평균 방식으로 조합하였다. 분석 결과, 분해-앙상블 모형은 원자료 기반 단일 모형 및 분해 기반 개별 모형보다 전반적으로 낮은 예측오차를 보였다. 또한 KREI 농업관측센터와의 비교 및 Encompassing 검정 결과, 제안 알고리즘과 농업관측센터 전망치는 상호 보완적인 예측 정보를 포함하는 것으로 나타났다. 이는 분해-앙상블 예측 알고리즘이 현행 농업관측 전망 체계에서 정량적 보조도로 활용될 수 있음을 시사한다.

* 이 연구는 한국농촌경제연구원에서 수행한 '배추 가격 예측력 개선 연구'를 수정 및 보완하여 논문화한 것임.

** 한국농촌경제연구원 전문연구원.

*** 한국농촌경제연구원 전문연구원.

**** 충북대학교 농업경제학과 교수, 교신저자. e-mail: bsyoon@chungbuk.ac.kr

Analysis of Forecasting Performance of a Kimchi Cabbage Price Forecasting Model Using Time Series Decomposition and Ensemble Methods*

Han Eun-Su**, Yoon Sung-Wuk***, Yoon Byung-Sam****

Keywords

Kimchi Cabbage, Price Forecasting, Time Series Decomposition, Forecast Combination

Abstract

This study proposes a decomposition–ensemble algorithm that combines time series decomposition with forecast combinations to improve the short-term forecasting performance of monthly wholesale kimchi cabbage prices. The log-transformed price series was decomposed into seasonal and seasonally adjusted components using STL decomposition. The seasonal component was forecast using the seasonal naive method, whereas the seasonally adjusted component was forecast using ETS, ARIMA, ARIMAX, and neural network autoregression (NNAR) models. The resulting forecasts were combined using the simple average, median, and MSE-based weighted average. The empirical results suggested that the proposed decomposition–ensemble model achieves lower forecast errors than models estimated directly from the original series and individual decomposition-based models. In addition, a comparison with the KREI Agricultural Outlook Center's price outlook and an encompassing test indicated that the two forecasts contain complementary information. These findings suggest that the proposed algorithm can serve as a supplementary quantitative tool in the Agricultural Outlook Center's forecasting system.

* This study is based on a revised and expanded version of the research, Improving Kimchi Cabbage Price Forecasting Accuracy, conducted by the Korea Rural Economic Institute.

** Research Associate, Center for Agricultural Outlook, Korea Rural Economic Institute.

*** Research Associate, Center for Agricultural Outlook, Korea Rural Economic Institute.

**** Professor, Department of Agricultural Economics, Chungbuk National University, Corresponding author.
e-mail: bsyoon@chungbuk.ac.kr

1. 서론

농산물 가격의 정확한 예측은 농가, 유통·가공업체, 정부 등 다양한 경제주체의 의사결정과 계획 수립에 중요한 정보를 제공한다는 점에서 중요한 연구 과제로 다루어져 왔다. 특히 최근 재배면적 감소, 이상기상 발생 빈도 증가 등으로 농산물 가격 변동성과 예측의 불확실성이 확대되면서 보다 정확하고 안정적인 가격 예측 방법론의 개발 필요성이 높아지고 있다. 한국농촌경제연구원(KREI) 농업관측센터는 이러한 맥락에서 36개 농축산물을 대상으로 재배면적, 생산량, 가격 등의 관측 정보를 제공하고 있다. 농업관측센터는 다양한 현장 조사와 예측 모형을 활용하여 복수의 예측치를 생산하고 있으나, 복수의 후보 예측치를 최종 전망치에 반영하기 위한 객관적인 기준과 절차가 체계적으로 마련되지 못한 한계가 존재하여 이를 보완하기 위한 정량적 예측 방법론의 논의가 요구된다.

농업관측센터의 이러한 한계를 보완하는 대안적 접근 중 하나로 여러 예측 정보를 조합하여 최종 예측치를 도출하는 예측 조합(forecast combination) 방식이 적용될 수 있다. 예측 조합은 단일 예측치에 의존할 때 발생할 수 있는 예측 불확실성을 완화하는 데 유용하다. 전통적으로 예측 조합은 복수의 예측치를 단순히 평균하는 방식에서 개별 예측치의 상대적 예측 성과에 따라 서로 다른 가중치를 부여하는 방법 등으로 확장되어 다양한 실증 연구에서 활용되었다(Bates & Granger, 1969; Clemen, 1989; Wang et al., 2023).

최근에는 복잡한 시계열의 변동 구조를 보다 세분화하여 반영하기 위해 시계열 분해와 예측 조합을 결합한 분해-앙상블(decomposition-ensemble) 기법이 활용되고 있다(Sun et al., 2023). 분해-앙상블 접근은 원자료 시계열을 추세, 계절성, 불규칙 변동 등 하위 성분으로 구분한 뒤 각 성분에 적합한 예측 모형을 적용하고 이후 성분별 예측 결과를 다시 조합하는 방식이다. 이러한 접근은 원자료를 직접 예측하는 경우보다 시계열에 내재된 이질적 변동 구조를 구분하여 반영할 수 있다는 장점이 있다.

한편, 국내 농산물 가격 예측 연구에서 시계열 분해와 예측 조합을 활용한 실증 연구는 제한적이지만, 최근 채소류를 중심으로 관련 연구가 이루어지고 있다. Yin et al.(2020)은 국내 채소류를 대상으로 STL(Seasonal and Trend decomposition using Loess)을 이용하여 원 시계열을 추세 및 계절 성분으로 분해한 후 LSTM(long short-term memory) 및 Attention 기반 LSTM을 결합한 STL-ATTN LSTM 알고리즘을 제안하였다. 분석 결과, 제안 알고리즘은 원자료를 직접 이용한 LSTM 모형보다 상대적으로 높은 예측 정확도를 보였다. 이는 농산물 가격 시계열에 내재된 추세와 계절성을 분리하여 각 성분의 특성에 맞는 예측 모형을 적용하는 것이 예측력 향상에 기여할 수 있음을 보여준다. 한은수 외(2025)는 양파, 마늘, 건고추를 대상으로 시계열 모형 및 인공신경망 알고리즘을 이용하여 개별 가격 예측치를 산출하고, 이를 조합하기 위해 단순 평균, 회귀 분석, 과거 예측 성과, 정보 기준에 기반한 다양한 예측 조합 방식을 적용하였다. 분석 결과, 조합 예측치의 예측오차가 개별 예측 모형보다 낮게 나타나 서로 다른 모형의 예측 정보를 결합하는 전략이 농산물 가격 예측의 불확실성을 완화하는

데 유용할 수 있음을 확인하였다.

이러한 선행연구는 각각 시계열 분해와 예측 조합이라는 서로 다른 측면에서 농산물 가격 예측력을 개선할 수 있는 방법론적 가능성을 제시한다. 본 연구는 이들 접근을 통합하여 Yin et al.(2020)의 시계열 분해 방식을 바탕으로 가격 시계열을 계절 성분과 계절조정 성분으로 구분하고 각 성분에 적합한 예측 모형을 적용한 후 예측치를 재조합한다. 이어 한은수 외(2025)의 예측 조합 접근을 확장하여 분해 기반 모형에서 산출된 복수의 예측치를 단순 평균, 중앙값, 평균제곱오차(mean squared error: MSE) 기반 가중 평균 방식으로 조합한다. 따라서 본 연구는 시계열 분해, 성분별 예측과 재조합, 그리고 모형 간 예측 조합을 하나의 단계적 예측 절차로 구성하였다는 점에서 기존 연구와 차별된다.

국내 농산물 품목 중 배추는 월별 계절성과 봄·여름·가을·겨울의 작형별 출하 구조가 뚜렷한 품목이지만, 이상기상 및 작황 변화 등이 복합적으로 작용하면서 가격 변동성이 크게 나타나는 특징이 있다. 특히 평시에는 일정한 계절 패턴을 보이나, 기상 급변 등의 외부 충격이 발생할 경우 단기간 가격이 급등할 수 있다. 이는 배추 가격 시계열에 계절성, 추세, 불규칙 충격이 동시에 포함되어 있음을 의미한다. 따라서 원자료 가격을 직접 예측하는 단일 모형보다는 시계열을 구조적 성분으로 분해하고, 복수의 예측 모형 결과를 결합하는 분해-앙상블 접근이 배추 가격 예측에 적합한 대안이 될 수 있다.

국내 배추 가격 예측 연구는 초기 수급 구조를 반영한 관측 모형 구축 연구에서 가격 시계열의 통계적 특성과 변동성 분석, 기상 및 생육 등 다양한 외생변수의 활용, 머신러닝 및 딥러닝 기반 예측 모형의 적용으로 연구 범위가 점차 확대되어 왔다(한석호·김병률, 2004; 김배성, 2005; 서교 외, 2008; 박지연·박영구, 2013; 유도일, 2016; 정호중·정원호, 2016; 장수희 외, 2017; 박소연 외, 2022; 이창민 외, 2022; 박종현 외, 2023; 박찬·이경순, 2023; 허다솜·정승권, 2024; 강동곤 외, 2024). 다만, 대부분의 연구가 주로 개별 예측 모형의 예측력 비교나 특정 정보의 추가 활용 가능성 검토에 초점을 두고 있어, 복수의 예측치를 체계적으로 조합하여 예측의 정확성과 안정성을 제고하려는 논의는 상대적으로 제한적이었다.

이러한 배경하에서 본 연구는 배추의 월별 도매가격을 대상으로 시계열 분해와 예측 조합을 결합한 분해-앙상블 예측 알고리즘을 제안하고, 예측 성과를 평가하는 데 목적이 있다. 이를 바탕으로 시계열 분해 없이 원자료에 직접 예측 모형을 적용하는 방식, 분해 후 성분별 예측 모형을 적용하는 방식, 그리고 분해 기반 예측치를 복수의 조합 방법으로 결합하는 방식의 세 가지 접근을 단계적으로 비교함으로써 시계열 분해와 예측 조합이 배추 가격 예측의 정확도 개선에 실질적으로 기여하는지를 검토한다.

아울러 본 연구는 제안 알고리즘의 예측치와 전문가 판단 기반 전망치의 상호 보완성을 실증적으로 검토함으로써 분해-앙상블 전략의 예측 안정성 제고 가능성과 정량 모형·전문가 판단 결합에 관한 방법론적 논의에 기여한다는 점에서 학술적 의의가 있다.

본 연구는 제2절에서 연구 방법 및 절차를 제시하고, 제3절에서는 분석 결과를, 마지막 제4절에서는 요약 및 결론을 제시한다.

2. 연구 방법 및 절차

2.1. 분석 자료

본 연구에 이용된 배추 가격은 서울시농수산식품공사로부터 입수한 가락시장 상품(上品) 기준 월별 도매가격 자료이다. 도매시장 반입량은 가격 자료와 동일하게 가락시장 월별 자료를 이용하였다. 김치 수입량은 한국무역통계진흥원의 월별 자료를 활용하였다. 분석 자료의 표본 기간은 각각 2001년 1월부터 2025년 12월까지이며, 전체 관측치 수는 변수별로 300개이다.

2.2. 분석 방법

본 연구는 배추 월별 도매가격의 단기 예측력 개선을 위해 시계열 분해, 개별 예측 모형, 예측 조합을 결합한 분석 절차를 적용하였다. 시계열 분해 방법으로 STL을, 개별 예측 모형으로는 시계열 예측에서 널리 활용되는 ETS 지수평활, ARIMA 및 ARIMAX 모형, 그리고 비선형 변동을 고려할 수 있는 신경망 자기회귀 알고리즘을 활용하였다. 이후 개별 모형에서 산출된 예측치는 단순 평균, 중앙값, MSE 기반 가중 평균 방식으로 조합하였으며, 예측 성과는 MAE, RMSE, sMASE를 이용하여 평가하였다. 마지막으로, Encompassing 검정을 시행하여 농업관측센터 전망치와 본 연구에서 제안한 알고리즘의 예측치가 서로를 포괄하는 관계에 있는지를 검정하였다.

2.2.1. 시계열 분해

시계열 자료는 장기적인 방향을 나타내는 추세, 일정한 주기로 반복되는 계절성, 그리고 추세와 계절성으로 설명되지 않는 잔차 성분으로 구성될 수 있다. STL(Seasonal and Trend decomposition using Loess)은 LOESS(locally estimated scatterplot smoothing) 국소 회귀를 이용하여 시계열을 추세, 계절, 잔차 성분으로 분해하는 방법으로, 시계열에 내재된 구조적 패턴을 구분하고 각 성분의 특성에 적합한 예측 모형을 적용하는 데 유용하다(Yin et al., 2020; Hyndman & Athanasopoulos, 2021).

배추 가격 시계열을 y_t , 추세 성분을 T_t , 계절 성분을 S_t , 잔차 성분을 R_t 라고 하면, 덧셈형(additive) 분해는 $y_t = T_t + S_t + R_t$, 곱셈형(multiplicative) 분해는 $y_t = T_t \times S_t \times R_t$ 로 표현할 수 있다. 덧셈형 분해는 시계열 수준 변화와 관계없이 계절 변동의 절대적 크기가 비교적 일정한 경우에 적합한 반면, 곱셈형 분해는 가격 수준이 높아질수록 변동 폭도 함께 확대되는 경우에 적합하다. 배추 가격은 가격 수준이 높은 시기에 변동 폭도 확대될 수 있으므로 원자료에서는 곱셈적 변동 구조가 나타날 가능성이 있다. 이에 본 연구에서는 가격 수준에 따른 분산 확대를 완화하고, 곱셈적 관계를 덧셈적 관계로 전환하기

위해 배추 가격에 로그 변환을 적용하였다. 즉, 곱셈형 구조에 로그를 취하면 덧셈형 구조인 $\log(y_t) = \log(T_t) + \log(S_t) + \log(R_t)$ 와 같이 변환된다.

STL은 계절 주기와 평활 모수 설정이 가능하여 월별 자료와 같이 계절성이 뚜렷한 시계열에 유연하게 적용할 수 있다. 또한 강건 추정 방식을 적용할 경우 일시적인 가격 급등락이나 이상치가 추세 및 계절 성분 추정에 과도하게 반영되는 것을 완화할 수 있다. 다만 농산물 가격의 큰 변동은 이상기상, 작황 변화 및 수급 불균형 등에 따른 실제 시장정보를 포함할 수 있으므로 강건 추정을 적용할 경우 수급 충격에 의한 가격 변동의 영향이 추세 및 계절 성분 추정에서 과도하게 축소될 가능성이 있다. 이에 본 연구에서는 실제 수급 충격에 의한 가격 정보를 분해 과정에 온전히 반영하기 위해 비강건 STL을 기본 사양으로 적용하였다.

2.2.2. 가격 예측 모형

가. 시계열 모형

본 연구는 배추 월별 도매가격을 예측하기 위한 시계열 모형으로 ETS(Error, Trend, Seasonality) 지수평활 모형, ARIMA 및 ARIMAX를 활용하였다.

ETS 지수평활 모형은 과거 관측값에 지수적으로 감소하는 가중치를 부여하여 예측치를 산출하는 방법이다. ETS 모형은 시계열을 오차(error), 추세(trend), 계절성(seasonality)으로 구분하고, 각 구성요소를 덧셈 또는 곱셈 형태로 조합하여 다양한 모형을 구성한다. 예를 들어 ETS(A,N,A)는 덧셈 오차, 추세 없음, 덧셈 계절성을 갖는 모형을 의미하며, ETS(M,Ad,N)은 곱셈 오차, 덧셈 감쇠 추세,¹⁾ 계절성 없음을 가정한 모형을 의미한다.

ARIMA 모형은 자기회귀(autoregressive: AR) 과정과 이동평균(moving average: MA) 과정을 결합한 대표적인 시계열 모형이다. ARIMA는 일반적으로 ARIMA(p,d,q)로 표기되며, p는 자기회귀 차수, d는 차분 횟수, q는 이동평균 차수를 의미한다. 시계열이 비정상성을 가지면 차분을 통해 정상성을 확보한 후 모형을 추정한다. 월별 가격 자료에서 계절적 자기상관이 존재할 경우에는 비계절 ARIMA 구조에 계절 ARIMA 구조를 추가한 SARIMA 모형이 고려되며, 이는 ARIMA(p,d,q)(P,D,Q)[m]로 표기된다. 여기서 P, D, Q는 각각 계절 자기회귀, 계절 차분, 계절 이동평균 차수이며, m은 계절 주기를 의미한다. 월별 자료의 경우 일반적으로 m=12가 적용된다.

ARIMA 모형에 배추 가격에 영향을 미치는 외생변수를 포함하면 ARIMAX로 확장될 수 있다. ARIMAX는 가격 시계열의 자기상관 구조뿐만 아니라 도매시장 반입량, 수입량 등 수급 관련 정보를 함

¹⁾ 감쇠 추세(damped trend)는 예측기간이 길어질수록 일정 비율로 약화되도록 설정한 형태이다. 이를 통해 장기 예측에서 추세가 지속적으로 확대되어 예측값이 과도하게 증가하거나 감소하는 문제를 완화할 수 있다.

게 반영할 수 있다는 점에서 농산물 가격 예측에 유용하게 활용될 수 있다.

나. 신경망 자기회귀(Neural network autoregression: NNAR) 모형

신경망 자기회귀(NNAR)는 인공신경망 구조를 시계열 예측에 적용한 모형이다. 인공신경망은 입력층(input layer), 은닉층(hidden layer), 출력층(output layer)으로 구성되며, 입력값이 가중치를 통해 은닉층으로 전달되고 활성화 함수(activation function)를 거쳐 비선형적으로 변환된다. 이러한 구조적 특성을 통해 인공신경망은 선형 모형으로 포착하기 어려운 복잡한 비선형 관계를 학습할 수 있다.

NNAR 모형은 시계열의 과거 관측값을 입력변수로 활용하고, 은닉층을 포함한 순방향 신경망(feed-forward neural network)을 통해 미래값을 예측한다. 과거 가격 정보에 기초하여 미래 가격을 예측한다는 점에서 자기회귀 모형의 성격을 가지나, 입력값과 예측값 간의 관계를 선형으로 제한하지 않는다는 점에서 전통적인 자기회귀 모형과 차별화된다. 일반적으로 NNAR 모형은 $NNAR(p, P, k)[m]$ 형태로 표현되며, 여기서 p 는 비계절 시차 수, P 는 계절 시차 수, k 는 은닉층 노드(node) 수, m 은 계절 주기를 의미한다.

2.2.3. 예측 조합 방법

예측 조합은 서로 다른 예측 모형에서 산출된 개별 예측치를 조합하여 최종 예측치를 도출하는 방법이다. 개별 예측 모형은 자료에 내재한 서로 다른 특징을 반영하므로, 특정 모형의 예측 성과가 모든 시점에서 일관되게 가장 우수하다고 보기 어렵다. 따라서 복수의 예측치를 조합하면 개별 모형의 일시적 예측 실패 가능성을 완화하고, 보다 안정적인 예측 결과를 얻을 수 있다.

본 연구는 예측 조합 방법으로 단순 평균, 중앙값, 평균제곱오차(mean squared error: MSE) 기반 가중 평균을 적용하였다. 단순 평균 조합은 개별 예측치에 동일한 가중치를 부여하여 평균값을 최종 예측치로 사용하는 방식이다. 이는 별도의 가중치 추정 과정이 필요하지 않아 적용이 간단하고, 가중치 추정오차에 따른 불확실성을 피할 수 있다는 장점이 있다.

중앙값 조합은 개별 예측치를 크기순으로 정렬한 뒤 중앙값에 위치한 값을 최종 예측치로 사용하는 방식이다. 단순 평균에 비해 극단적인 예측치의 영향을 상대적으로 덜 받기 때문에 일부 모형에서 일시적인 예측 실패가 발생하거나 가격 급등락으로 예측치 간 차이가 확대되는 경우 보다 강건한 조합 방식으로 활용될 수 있다.

MSE 기반 가중 평균은 과거 검증 구간에 대한 예측오차가 작은 모형에 더 큰 가중치를 부여하는 방식이다. k 번째 모형의 과거 검증 구간 평균제곱오차를 MSE_k 라고 하면, 해당 모형의 가중치 w_k 는 식 (1-1)과 같이 정의할 수 있다(Stock & Watson, 1999).

$$w_k = \frac{MSE_k^{-1}}{\sum_{j=1}^N MSE_j^{-1}}, \quad MSE_k = \frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_{k,t})^2 \quad (1-1)$$

여기서 MSE_k 는 과거 검증 구간에서 산출한 k 번째 모형의 평균제곱오차, T 는 검증 구간의 관측치 수, y_t 는 실제값, $\hat{y}_{k,t}$ 는 k 번째 모형의 예측치를 의미한다. 식 (1-1)에 따르면 과거 예측오차가 작을수록 해당 모형의 가중치는 커지고, 예측오차가 클수록 가중치는 작아진다.

최종적으로 MSE 기반 조합 예측치는 식 (1-2)와 같이 개별 예측치에 가중치를 곱한 뒤 합산하여 산출된다.

$$\hat{y}_t^{MSE} = \sum_{k=1}^N w_k \hat{y}_{k,t} \quad (1-2)$$

2.2.4. 예측 정확도 평가 방법

배추 가격 예측 모형의 정확도는 실제값과 예측값의 차이인 예측오차를 기준으로 평가할 수 있다. 본 연구는 예측 성과를 평가하기 위해 평균절대오차(mean absolute error: MAE), 평균제곱근오차(root mean squared error: RMSE), 계절성 평균절대스케일오차(seasonal mean absolute scaled error: sMASE)를 활용하였다. 평가 기간의 관측치 수를 N , t 시점의 실제값과 예측값을 각각 y_t 와 $\hat{y}_t$ 라고 하면, MAE와 RMSE는 각각 식 (2-1), 식 (2-2)와 같이 정의된다.

$$MAE = \frac{1}{N} \sum_{t=1}^N |y_t - \hat{y}_t| \quad (2-1)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (y_t - \hat{y}_t)^2} \quad (2-2)$$

MAE는 예측치가 실제값에서 평균적으로 얼마나 벗어나는지를 원 단위로 나타내는 지표이다. 배추 가격과 같이 동일 단위의 예측 모형을 비교할 경우 해석이 직관적이며, 평균적인 오차 수준을 파악하는데 유용하다. RMSE는 큰 오차에 더 큰 가중을 부여하므로, 가격 급등락 시기와 같이 예측 실패 규모가 크게 나타나는 구간의 성과를 확인하는 데 적합하다.

다만 MAE와 RMSE는 원 단위 오차를 제시한다는 장점이 있으나, 계절성이 강한 시계열에서 단순한

계절 반복 예측 대비 상대적 개선 정도를 직접 보여주지는 못한다. 이에 본 연구는 계절성 단순 예측(seasonal naive)을 기준으로 한 sMASE를 추가로 산출하였다. 월별 자료의 계절 주기를 $m=12$, 학습 자료의 관측치 수 T , 평가 기간의 관측치 수를 N 이라고 하면, sMASE는 식 (2-3)과 같이 정의된다. 식 (2-3)에서 분자는 평가 대상 모형의 평균절대오차이며, 분모는 학습자료에서 계절성 단순 예측을 적용했을 때의 평균절대오차를 의미한다. 따라서 sMASE가 1보다 작으면 해당 모형이 계절성 단순 예측보다 우수한 예측 성과를 보인 것으로 해석할 수 있으며, 1보다 크면 기준 모형보다 예측오차가 큰 것으로 해석할 수 있다.²⁾

$$sMASE = \frac{\frac{1}{N} \sum_{t=1}^N |Y_t - \hat{Y}_t|}{\frac{1}{N-1} \sum_{t=m+1}^T |Y_t - Y_{t-m}|} \quad (2-3)$$

2.2.5. 시계열 분해-앙상블 기반 배추 가격 예측 알고리즘

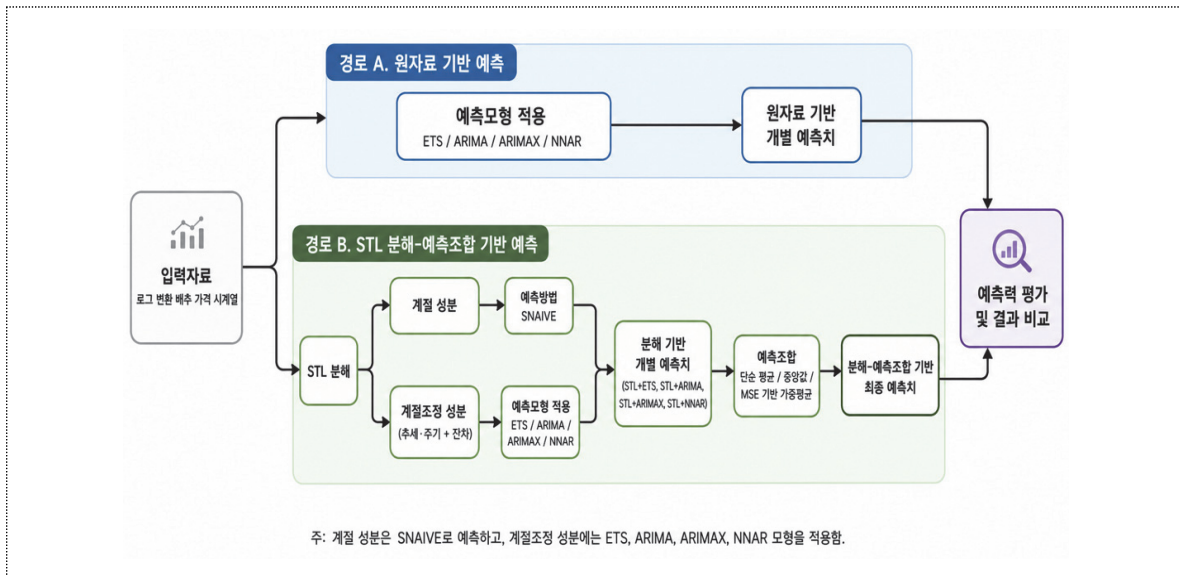
본 연구는 시계열 분해와 예측 조합을 순차적으로 적용하여 배추 가격 예측 알고리즘을 구성하였다. 제안 알고리즘은 로그 변환된 배추 가격 시계열에 STL 분해를 적용하여 계절 성분과 계절조정 성분을 구분한 후, 각 성분의 특성에 적합한 예측 모형을 적용하는 방식으로 이루어진다. 여기서 계절조정 성분은 계절성을 제거한 추세 성분과 잔차 성분의 합을 의미한다.

구체적으로 계절 성분은 전년 동월의 계절 패턴을 활용하는 계절성 단순 예측(SNAIVE)을 이용하여 예측하였다. 계절조정 성분은 ETS 지수평활, ARIMA, ARIMAX, NNAR 모형을 적용하여 복수의 예측치를 산출하였다. 이후 계절 성분 예측치와 계절조정 성분 예측치를 결합하여 STL 분해 기반 개별 예측치를 구성하였다.

마지막으로 분해 기반 개별 예측치는 단순 평균, 중앙값, MSE 기반 가중 평균 방식으로 조합하여 최종 분해-앙상블 기반 예측치를 산출하였다. 또한 시계열을 분해하지 않은 원자료 기반 예측 모형의 결과를 비교 대상에 포함하여, 원자료에 직접 예측 모형을 적용한 경우와 분해 기반 개별 예측치 및 분해-앙상블 기반 예측치의 예측 성과를 함께 비교하였다. <그림 1>은 본 연구에서 제안한 시계열 분해-앙상블 기반 배추 가격 예측 알고리즘의 전체 구조를 제시한 것이다.

²⁾ 한편, 본 연구에서는 일반적으로 많이 적용되는 평균절대비율오차(mean absolute percentage error: MAPE)를 평가 지표에서 제외하였다. MAPE는 실제값을 분모로 사용하기 때문에 가격 수준이 낮은 시기에는 동일한 원 단위 오차라도 과도하게 크게 평가될 수 있으며, 반대로 가격 수준이 높은 시기에는 오차가 상대적으로 작게 평가될 수 있다. 특히 배추 가격과 같이 계절별 가격 수준 차이와 급등락이 반복되는 시계열에서는 이러한 특성이 모형 간 예측 성과 비교를 왜곡할 가능성이 있다.

그림 1. 시계열 분해-양상불 기반 배추 가격 예측 알고리즘 구조



주: 본 그림의 도식화 과정에서 생성형 AI 도구인 ChatGPT를 보조적으로 활용하였으며, 최종 내용을 저자가 검토 수정하였음.

가. 변수 정의 및 구성

본 연구에서 사용한 주요 변수의 정의와 산식은 <표 1>에 제시하였다. 원자료 기반 모형은 로그 가격을 직접 예측 변수로 설정하였고, 분해 기반 모형은 STL 분해를 통해 도출된 계절 성분과 계절조정 성분을 각각 예측 대상으로 설정하였다.

설명변수는 가격 기반 파생 변수와 수급 관련 변수로 구성하였다. 우선, 가격 기반 파생변수로 전년 동월 대비 로그 가격 변화는 동일 계절 조건에서 최근 가격 수준이 전년보다 높은지 또는 낮은지를 반영하기 위한 변수이다. 최근 3개월 평균 가격의 전년 동일 기간 대비 차이는 특정 월의 일시적 가격 변동보다 최근 가격 흐름의 지속성을 반영하기 위해 사용하였다. 동일 작형 평균 대비 가격 수준은 예측 대상 월이 속한 작형의 통상적인 가격 수준과 비교하여 현재 가격이 상대적으로 높은지 또는 낮은지를 파악하기 위한 변수이다. 수급 관련 변수로는 도매시장 반입량과 김치 수입량의 1~3개월 시차 변수를 사용하였다. 반입량은 도매시장에서 실제로 공급되는 물량을 나타내는 단기 공급지표이며, 김치 수입량은 국내 배추 및 김치 수급 여건과 관련된 외부 수급지표로 볼 수 있다.³⁾

본 연구는 설명변수 구성 과정에서 예측 대상 월의 실제 가격이나 사후적으로 확인되는 정보의 사용을 배제하였다. 즉, 예측 대상 월 직전까지 관측 가능한 가격, 반입량, 수입량 자료만을 이용하여 변수를 생성함으로써 정보 누출 가능성을 최소화하고, 예측 모형의 실무 적용 가능성을 높이하고자 하였다. 한편, 기상정보도 유용한 설명변수가 될 수 있으나, 예측 대상 월의 실현 기상자료는 예측시점에 이용할

³⁾ 본 연구에서는 배추 반입량과 김치 수입량의 월별 변동이 작형별 출하 시기와 국내의 수급 상황에 관한 정보를 포함할 수 있다는 점을 고려하여 별도의 계절조정 없이 로그변환한 시차 변수로 활용하였다. 또한 외생변수를 반영한 이후에도 가격의 계절조정 성분에 남아있는 시계열적 자기상관은 ARIMA 오차구조를 통해 반영하였다.

수 없으며 기상예보를 사용할 경우 예측오차가 가격 예측의 불확실성으로 추가될 수 있다는 점을 고려하여 본 연구의 설명변수에는 포함하지 않았다.

표 1. 배추 가격 예측 모형의 주요 변수 정의 및 산식

구분		변수		산식
예측 대상 변수	원자료 기반	P_t	t 월의 배추 도매가격	P_t (t 는 예측 대상 월)
		p_t	로그 변환한 배추 도매가격	$p_t = \log(P_t)$
예측 대상 성분	분해 기반	S_t	STL 분해로 추출한 계절 성분	$p_t = T_t + S_t + R_t$ (T_t, S_t, R_t 는 각각 STL 분해의 추세 성분, 계절 성분, 잔차 성분)
		SA_t	STL 분해 후 계절 성분을 제거한 계절조정 성분	$SA_t = p_t - S_t$
설명변수 (예측변수)	가격 기반 변수	$\Delta^{(12)}p_{t-1}$	예측 대상 월의 전월 기준 전년 동월 대비 로그 가격 변화	$\Delta^{(12)}p_{t-1} = p_{t-1} - p_{t-13}$
		$G_{t-1}^{MA_3}$	예측 대상 월의 전월까지 최근 3개월 평균 로그 가격과 전년 동일 기간 평균 간 차이	$G_{t-1}^{MA_3} = MA_3(p)_{t-1} - MA_3(p)_{t-13}$ ($MA_3(p)_t$ 는 t 월까지의 최근 3개월 로그 가격 이동평균)
		C_{t-1}	예측 대상 월의 전월 로그 가격과 동일 작형 과거 평균 로그 가격 차이	$C_{t-1} = p_{t-1} - \bar{p}_{c,t-1}$ ($\bar{p}_{c,t-1}$ 는 예측 대상 월이 속한 작형 c 와 동일한 작형에 해당하는 과거 월별 로그 가격의 평균)
	반입량	q_{t-k}	배추 반입량의 로그 변환 시차 변수	$q_{t-k} = \log(Q_{t-k}), k = 1, 2, 3$
	수입량	m_{t-k}	김치 수입량의 로그 변환 시차 변수	$m_{t-k} = \log(M_{t-k}), k = 1, 2, 3$

자료: 저자 작성.

나. 학습자료와 평가자료 구간의 구분

본 연구는 배추 가격 예측 알고리즘의 예측 성과를 평가하기 위해 2001년 1월~2025년 12월의 월별 자료($n=300$)를 표본 내 모형 설정 및 추정(in-sample specification and estimation) 구간, 검증(validation) 구간, 표본 외 평가(out-of-sample evaluation) 구간으로 구분하여 활용하였다.

표본 내 모형 설정 및 추정 구간은 2001년 1월~2018년 12월($n=216$)로 설정하였다. 이 구간은 STL 분해 방식과 개별 예측 모형의 기본 구조를 설정하는 데 활용하였다. 즉, 계절 성분과 계절조정 성분을 구분하기 위한 STL 분해 구조를 검토하고, ETS, ARIMA, ARIMAX, NNAR 등 개별 예측 모형의 기본 구조를 결정하였다.

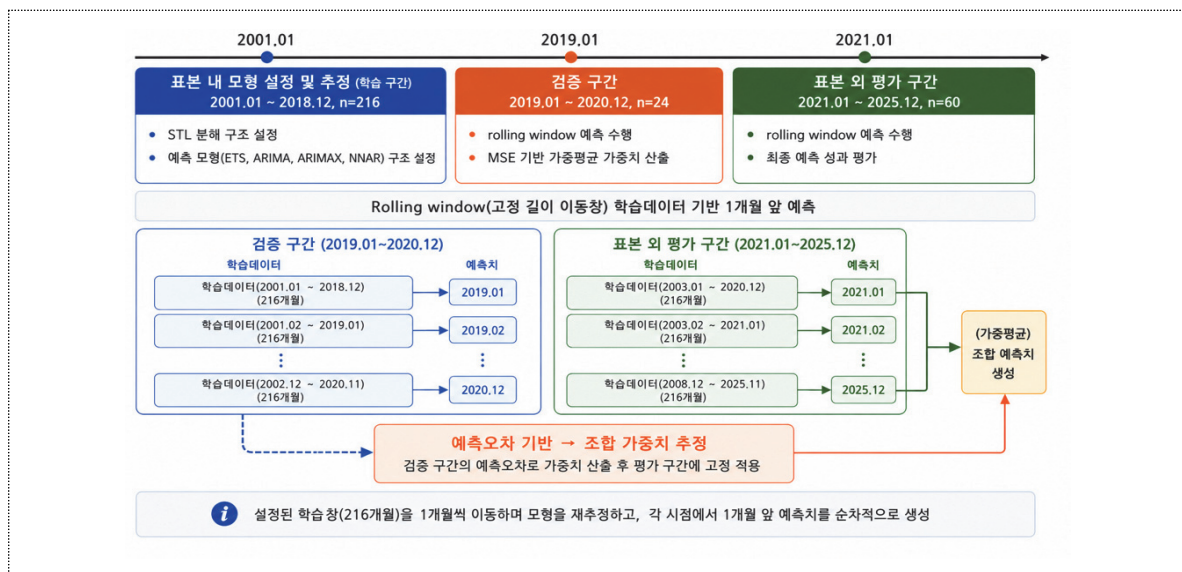
검증 구간은 2019년 1월~2020년 12월($n=24$)로 설정하였다. 이 구간은 표본 내 구간에서 설정한 모형의 구조를 유지한 상태에서 예측 성과를 점검하고, MSE 기반 가중 평균 조합에 필요한 모형별 가중치를 산출하는 데 활용하였다. 검증 구간의 예측치는 고정 길이 이동창(fixed-length rolling

window) 방식으로 산출하였으며, 각 시점에서 최근 216개월 자료를 이용하여 모형을 재추정한 후 1개월 앞 예측치(1-step-ahead-forecast)를 도출하였다. 예를 들어 2019년 1월 예측에는 2001년 1월~2018년 12월의 216개월 자료를 사용하고, 2019년 2월 예측에는 1개월 이동하여 2001년 2월~2019년 1월의 자료를 활용한다. 이와 같은 방식으로 2020년 12월까지 총 24개의 1개월 앞 예측치를 산출한다.⁴⁾ 본 연구는 검증 구간에서 산출한 MSE 기반 조합 가중치를 표본 외 평가 구간 전체에 고정 적용하였다. 이는 평가 구간의 실제값이 가중치 산정에 반영되는 것을 방지하고, 모든 예측 시점에 동일한 결합 기준을 적용하여 예측 성과를 일관되게 비교하기 위한 것이다. 최근 예측오차에 따라 가중치를 매월 갱신하는 동적 조합 방식도 적용할 수 있으나, 가중치 산정기간과 갱신 기준에 따라 결과가 달라질 수 있으므로 본 연구에서는 분석 대상에서 제외하고 향후 연구과제로 남겼다.

표본 외 평가 구간은 2021년 1월~2025년 12월(n=60)로 설정하였다. 이 구간은 앞선 구간에서 확정한 분해 방식, 개별 예측 모형 구조, 예측 조합 가중치를 바탕으로 최종 예측 성과를 평가하는 데 활용하였다. 표본 외 평가 구간에서도 검증 구간과 동일하게 고정 길이 이동창 방식을 적용하여 매월 1개월 앞 예측치를 산출하고, 실제값과 예측값 간의 예측오차를 바탕으로 모형별 예측 성과를 평가하였다.

〈그림 2〉는 본 연구에서 설정한 학습자료와 평가자료 구간의 구분을 제시한 것이다.

그림 2. 학습자료와 평가자료 구간의 구분



주: 본 그림의 도식화 과정에서 생성형 AI 도구인 ChatGPT를 보조적으로 활용하였으며, 최종 내용을 저자가 검토 수정하였음.

⁴⁾ 검증 구간 설정에 따른 영향을 확인하기 위해 기본 rolling window 구조를 유지한 상태에서 검증 구간 종료 시점을 2019년 12월, 2020년 6월, 2020년 12월로 달리하여 MSE 기반 조합 가중치를 산출하였다. 분석 결과, 모형별 가중치와 평가 구간 예측오차의 변화 폭은 제한적으로 나타나, 본 연구의 주요 결과가 검증 구간 설정에 크게 의존하지 않음을 시사한다. 학습 구간 길이와 검증 구간 길이를 동시에 변화시키는 경우 모형 재추정 구간과 STL 분해 결과가 함께 달라지므로, 본 연구에서는 고정된 rolling window 구조하에서 제한적 민감도 분석을 수행하였다.

2.2.6. Encompassing 검정

본 연구는 농업관측센터 전망치와 제안 알고리즘 예측치가 서로 다른 예측 정보를 포함하는지, 그리고 조합을 통해 예측력 개선이 가능한지를 검토하기 위해 Harvey et al.(1998)의 Encompassing 검정을 수행하였다. 본 연구에서는 농업관측센터의 전망치를 기준 예측치로, 제안 알고리즘 예측치를 대안적 예측치로 설정하였다. 시점 t 에서 농업관측센터 전망치의 예측오차를 e_{1t} , 제안 알고리즘의 예측오차를 e_{2t} 라고 하면, Encompassing 검정식은 식 (3-1)과 같이 표현할 수 있다.

$$e_{1t} = \alpha + \lambda(e_{1t} - e_{2t}) + \epsilon_t \quad (3-1)$$

여기서 $H_0 : \lambda = 0$ 이라는 귀무가설이 기각되지 않으면, 농업관측센터 전망치가 제안 알고리즘 예측치에 포함된 예측 정보를 포괄하는 것으로 해석할 수 있다. 반대로 귀무가설이 기각되면($\lambda \neq 0$), 제안 알고리즘 예측치가 농업관측센터 전망치에 포함되지 않은 추가적인 정보를 제공할 가능성이 있으며, 이는 두 예측치를 결합하여 활용할 수 있는 근거로 해석될 수 있다. Encompassing 검정 결과 두 예측치의 조합 가능성이 확인되는 경우, 농업관측센터 전망치와 제안 알고리즘 예측치는 식 (3-2)와 같은 선형 결합 형태로 활용될 수 있다.

$$\hat{f}_{c,t} = (1 - \lambda)f_{1t} + \lambda f_{2t} \quad (3-2)$$

여기서 $f_{c,t}$ 는 조합 예측치, f_{1t} 는 농업관측센터 전망치, f_{2t} 는 본 연구의 제안 알고리즘 예측치를 의미한다. $(1 - \lambda)$ 는 농업관측센터 전망치에 부여되는 가중치, λ 는 제안 알고리즘 예측치에 부여되는 가중치이다.

3. 분석 결과

3.1. 기초 통계량

〈표 2〉는 본 연구의 분석에 활용된 배추 월별 도매가격의 기초 통계량을 제시한 것이다. 원자료 기준 배추 도매가격의 평균은 6,658원/10kg, 최솟값은 1,073원, 최댓값은 24,874원, 표준편차는 4,028원으로 나타났다. 또한 왜도는 1.32, 첨도는 5.54로 분포가 우측으로 치우쳐 있으며, 정규분포보다 뾰족

한 형태를 보였으나, 로그 변환 후에는 왜도가 -0.24로 나타나 분포의 비대칭성이 완화되었고, 첨도는 2.54로 낮아져 극단값의 영향도 감소한 것으로 나타났다. 이러한 결과는 배추 도매가격에 로그 변환을 적용하는 것이 가격 시계열의 분포 왜곡과 극단값의 영향을 완화하는 데 유효함을 시사한다. 또한 로그 변환을 통해 STL 분해가 전제하는 덧셈형 구조에 보다 부합하는 형태로 자료를 변환하였다.

표 2. 배추 도매가격(원/10kg)에 대한 기초 통계량 (n=300)

구분	평균	중앙값	최대값	최소값	표준편차	왜도	첨도
원자료(전체)	6,658.0	5,755.0	24,874.3	1,073.0	4,028.2	1.32	5.54
겨울배추	6,297.9	5,849.0	14,914.5	1,073.0	3,581.2	0.47	2.43
봄배추	5,040.7	4,546.5	10,349.6	1,718.0	2,568.4	0.57	2.16
여름배추	8,437.6	6,910.9	24,874.3	1,396.0	4,886.3	1.22	4.27
가을배추	5,436.4	4,998.0	11,352.4	1,461.0	2,711.0	0.48	2.09
로그 변환	8.62	8.66	10.12	6.98	0.62	-0.24	2.54

주: 작형은 겨울배추 1~4월, 봄배추 5~6월, 여름배추 7~10월, 가을배추 11~12월로 구분하였음.

자료: 서울시농수산식품공사, 가락시장 유통 정보(<https://www.garak.co.kr/>).

3.2. STL 시계열 분해 결과

본 연구는 배추 도매가격 시계열의 계절성과 비계절적 변동을 구분하기 위해 STL 분해를 적용하였다. STL 분해의 주요 파라미터는 표본 내 모형 설정 및 추정 구간(2001년 1월~2018년 12월) 자료를 기준으로 결정하였다. 계절 주기는 월별 자료의 특성에 따라 12개월로 설정하였으며, 계절 성분은 고정된 월별 패턴이 반복되는 방식(s.window=periodic)으로 추정하였다. 이는 작형별 출하 시기에 따라 형성되는 배추 가격의 월별 계절 패턴이 비교적 안정적으로 반복된다는 점을 고려한 설정이다.

다만 고정형 계절성(s.window=periodic) 설정은 계절 성분을 예측기간 전체에 단일하게 고정한다는 의미가 아니다. 본 연구는 최초 모형 설정 및 추정 구간에서 결정한 STL 파라미터를 고정 길이 이동창 전반에 동일하게 적용하되, 각 예측 시점마다 해당 시점의 216개월 학습자료에 STL을 재적용하여 추세·계절·잔차 성분을 새롭게 산출하였다. 즉, 예측 시점별로 새롭게 추가되는 자료의 특성은 분해 결과에 반영되나, 각 이동창 내에서는 월별 계절 패턴이 일정하게 반복된다고 가정한다.

추세 평활화 구간(t.window)과 지역통과 필터 평활화 구간(l.window)은 월별 자료의 계절 주기($m=12$)와 고정 계절성 설정에 따라 R의 stl() 함수 기본값 산출 방식을 적용하여 각각 19와 13으로 결정하였다.⁵⁾ 내부 반복(inner)과 외부 반복(outer)은 각각 2와 0으로 설정하여 강건 반복을 적용하지

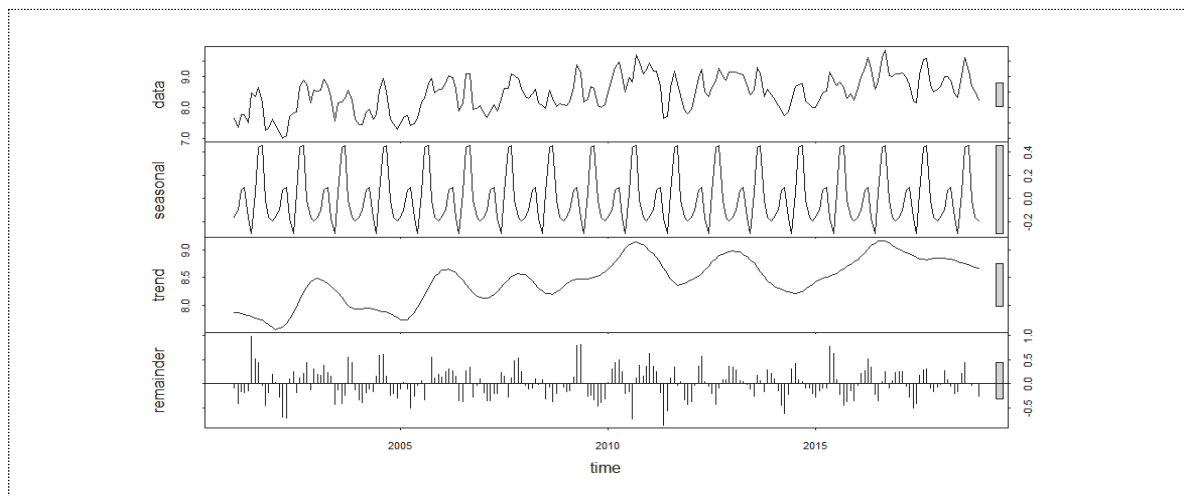
5) R의 stl() 함수에서 l.window는 계절 주기보다 작지 않은 가장 작은 홀수로 설정되며, 계절 주기가 12개월인 월별 자료에서는 l.window=13이 적용된다. t.window는 계절 주기 m 과 계절 평활창 W_s 를 이용하여 $1.5m/(1-1.5/W_s)$ 를 계산한 후, 해당 값보다 작지 않은 가장 작은 홀수로 결정된다. s.window=periodic으로 설정하면 계절 성분이 시간에 따라 변하지 않는 고정 패턴으로 추정되며, 이는 계절 평활창 W_s 가 무한대 ($W_s \rightarrow \infty$)에 해당하는 경우로 볼 수 있다. ($W_s \rightarrow \infty$)이면 $1.5/W_s \rightarrow 0$ 이 되어 분모가 1에 수렴하므로 $1.5 \times m = 1.5 \times 12 = 18$ 이 되어 18보다 작지 않은 가장 작은 홀수인 t.window=19가 산출된다.

않는 표준 STL 분해 방식을 사용하였다.⁶⁾

〈그림 3〉은 로그 변환한 배추 가격 시계열에 STL 분해를 적용한 결과이다. 분해 결과, 원 시계열(data)에서는 시기별 가격 수준의 변화와 반복적인 계절 패턴이 함께 확인된다. 계절 성분(seasonal)은 월별로 유사한 형태가 반복되는 패턴을 보이며, 이는 배추 가격에 작형 및 출하 시기와 관련된 계절성이 내재해 있음을 의미한다. 추세 성분(trend)은 표본 기간 동안 완만한 상승과 하락을 반복하여, 계절 변동과는 독립적으로 중장기적인 가격 수준 변화가 존재함을 보여준다. 잔차 성분(remainder)은 전반적으로 0을 중심으로 분포하였으나, 일부 시점에서는 비교적 큰 양(+) 또는 음(-)의 변동이 확인되었다. 이는 추세 및 계절 성분으로 설명되지 않는 일시적 가격 충격과 불규칙 변동이 존재함을 시사하며, 기상 여건, 작황 변화 등 단기 수급 충격 요인과 관련될 가능성이 있다.

이러한 분해 결과를 바탕으로 본 연구는 계절 성분에는 계절성 단순 기법(SNAIVE)을 적용하고, 계절 조정 성분에는 ETS, ARIMA, ARIMAX, NNAR 모형을 적용하여 분해 기반 예측치를 산출하였다.

그림 3. 배추 가격 시계열의 STL 분해 결과



주: 로그 변환한 월별 배추 도매가격 시계열에 STL 분해를 적용한 결과임. 계절 주기는 12개월, 계절 성분은 고정된 월별 패턴이 반복되는 방식(s.window=periodic)으로 설정하였으며, 추세 평활화 구간(t.window)은 19, 저역통과 필터(l.window)는 13, 내부 반복(inner)은 2, 외부 반복(outer)은 0을 적용함.

자료: 저자 작성.

3.3. 배추 가격 예측 모형 추정 결과

〈표 3〉은 배추 가격 예측에 적용한 개별 모형의 구조와 주요 설정을 제시한 것이다. 각 모형의 기본 구조는 표본 내 모형 설정 및 추정 구간(2001년 1월~2018년 12월) 자료를 기준으로 선정하였다. 이

⁶⁾ 본 연구의 자료와 평가 기간에서는 강건 반복의 적용에 따른 일관된 예측력 개선이 확인되지 않아 비강건 STL을 기본 사양으로 유지하였다.

후 rolling window 예측 과정에서는 예측 시점별로 새롭게 추가되는 자료를 반영하여 모형의 계수, 상태값 및 신경망 가중치를 재추정하되, <표 3>에서 제시한 ETS 구성 형태, ARIMA 차수, NNAR 입력 시차 및 은닉 노드 수 등 기본 구조는 전 예측 시점에 걸쳐 동일하게 유지하였다.

원자료 기반 모형은 로그 변환한 배추 가격 시계열에 예측 모형을 직접 적용하는 방식으로 구성하였으며, ETS, ARIMAX, 외생변수를 포함하는 신경망 자기회귀 모형(NNARX)을 포함한다. ETS(A,N,A) 모형은 덧셈 오차와 덧셈 계절성을 반영하여 별도의 추세항은 포함하지 않는 구조이다. ARIMAX(1,1,2)(2,0,0)[12]는 비계절 차분을 적용한 계절 ARIMA 구조에 배추 반입량과 김치 수입량의 시차 변수를 추가한 모형이다.⁷⁾ NNARX(3,1,2)[12]는 비계절 시차 3개와 계절 시차 1개를 입력변수로 활용하고, 은닉층 노드 수를 2개로 설정하였으며, 가격 기반 파생변수를 추가 입력변수로 포함하였다.

분해 기반 모형의 경우 STL 분해를 통해 구분된 계절 성분은 계절성 단순 기법(SNAIVE)를 이용하여 예측하였고, <표 3>의 모형 추정 결과는 계절조정 성분에 적용한 결과이다. STL+ETS(A,N,N)는 계절성이 제거된 계절조정 성분에 별도의 추세 및 계절항을 포함하지 않는 단순 지수 평활 구조를 적용한 것이다. STL+ARIMA 및 STL+ARIMAX는 계절조정 성분에 내재한 자기상관 구조를 반영하기 위해 설정하였으며, STL+NNAR 및 STL+NNARX는 비선형 자기회귀 패턴을 고려하기 위해 설정하였다.

표 3. 배추 가격 예측 모형의 구조 및 주요 설정

구분	모형	구조(structure) 및 주요 설정(specifications)	
원자료 기반	ETS 지수평활	① ETS(A,N,A)	• 덧셈 오차(A), 추세 없음(N), 덧셈 계절성(A)
	ARIMA	② ARIMAX(1,1,2)(2,0,0)[12]	• 계절성 ARIMA(p,d,q)(P,D,Q)[m] • 외생변수: 반입량·김치 수입량 시차 변수
	인공신경망	③ NNARX(3,1,2)[12]	• 계절성 NNAR(p,P,k)[m] • 비 계절 시차(p): AR(p)로 최적 시차 결정(AIC 기준) • 계절 시차(P): 기본값 1 • 은닉층의 노드 수(size): 2 • 외생변수: 가격 기반 특성 변수
분해 기반	ETS 지수평활	④ STL-ETS(A,N,N)	• 덧셈 오차(A), 추세·계절성 없음(N)
	ARIMA	⑤ STL+ARIMA(1,1,2)(2,0,1)[12]	• 계절성 ARIMA(p,d,q)(P,D,Q)[m]
		⑥ STL+ARIMAX(0,1,5)(2,0,0)[12]	• 계절성 ARIMA(p,d,q)(P,D,Q)[m] • 외생변수: 반입량·김치 수입량 시차 변수
	인공신경망	⑦ STL+NNAR(3,1,2)[12]	• 계절성 NNAR(p,P,k)[m]
		⑧ STL+NNARX(2,1,2)[12]	• 계절성 NNAR(p,P,k)[m] • 외생변수: 가격 기반 특성 변수

자료: 저자 작성.

7) KPSS 검정 결과, 로그 가격과 STL 계절 조정 성분의 검정 통계량은 각각 1.830과 1.885로 1% 유의수준에서 정상성 귀무가설이 기각되었다. 1차 차분 후에는 각각 0.029, 0.023으로 유의수준 10% 기준에서 귀무가설이 기각되지 않아 차분 이후 정상성이 확보되었음을 확인하였다. 이에 비계절 차분 차수를 $d=1$ 로 설정하였다. 계절 차분 필요성 검토 결과 로그 가격과 계절 조정 성분 모두 계절 차분 차수 $D=0$ 으로 나타나 추가적인 계절 차분은 적용하지 않았다.

3.4. 예측 조합을 위한 가중치 산출 결과

〈표 4〉는 분해 기반 개별 예측 모형의 예측치를 조합하기 위해 산출한 조합 방식별 가중치 및 중앙값 선택 비율을 제시한 것이다. 단순 평균 조합은 5개의 개별 예측 모형에 동일한 가중치를 부여하는 방식으로, 각 모형에 0.200의 가중치가 적용된다. MSE 기반 조합 가중치는 검증 구간(2019년 1월~2020년 12월)의 예측오차를 이용하여 산출하였으며, 평균제곱오차가 작은 모형일수록 더 큰 가중치가 부여되도록 설정하였다. 산출된 가중치는 표본 외 평가 구간(2021년 1월~2025년 12월)의 조합 예측치 산출에 고정 적용하였다. MSE 기반 가중치 산출 결과, STL+ARIMAX 모형의 가중치가 0.212로 가장 높게 나타났으나, 전체적으로 모형 간 가중치 차이는 크지 않아 조합 대상 모형들의 예측 성과가 비교적 유사한 것으로 나타났다. 한편, 중앙값 조합은 사전적으로 고정된 가중치를 부여하는 방식이 아니므로, 〈표 4〉에는 표본 외 평가 구간에서 각 모형의 예측치가 중앙값으로 선택된 비율을 제시하였다. 중앙값 선택 비율은 STL+ARIMA 모형이 0.292로 가장 높게 나타났다.

표 4. 분해 기반 개별 예측 모형의 조합 방식별 가중치 및 선택 비율

모형		구조 및 주요 설정	단순 평균 가중치	중앙값 선택 비율	MSE 기반 가중치
분해 기반	ETS 지수평활	④ STL+ETS(A,N,N)	0.200	0.125	0.209
	ARIMA	⑤ STL+ARIMA(1,1,2)(2,0,1)[12]	0.200	0.292	0.165
		⑥ STL+ARIMAX(0,1,5)(2,0,0)[12]	0.200	0.208	0.212
		⑦ STL+NNAR(3,1,2)[12]	0.200	0.167	0.206
	인공신경망	⑧ STL+NNARX(2,1,2)[12]	0.200	0.208	0.208

자료: 저자 작성.

3.5. 가격 예측 정확도 평가 결과

〈표 5〉는 표본 외 평가 구간인 2021년 1월~2025년 12월을 대상으로 원자료 기반 모형, STL 분해 기반 개별 모형, 본 연구에서 제안한 분해-앙상블 기반 모형의 예측 정확도를 비교한 결과이다. 예측 정확도는 RMSE, MAE, sMASE를 기준으로 평가하였으며, 각 지표는 작을수록 예측오차가 작음을 의미한다. 예측 모형은 로그 변환 가격을 기준으로 추정하였으나, 예측 정확도 평가는 원 단위 가격 기준으로 수행하였다.⁸⁾

분석 결과, 본 연구에서 제안한 분해-앙상블 기반 모형은 전반적으로 낮은 예측오차를 보였다. RMSE 기준으로 분해-앙상블 모형은 2,457.1~2,472.2로 원자료 및 분해 기반 개별 예측 모형보다 우수한 예측 성과를 보였다. MAE와 sMASE 기준에서는 STL+NNARX를 제외한 대부분의 개별 예측 모

8) 로그 예측치 $\hat{p}_t$ 를 원 단위 가격으로 환산할 때 단순 역변환 $\exp(\hat{p}_t)$ 을 적용하면 과소 추정될 가능성이 있다. 이에 본 연구에서는 로그 예측오차의 분산 $\hat{\sigma}^2$ 을 반영한 보정 역변환을 적용하였으며, 원 단위 예측치는 $\hat{p}_t = \exp(\hat{p}_t + \hat{\sigma}^2/2)$ 로 산출하였다.

형보다 분해-양상불 모형의 예측오차가 낮게 나타났다. 한편, 모든 모형에 대해 sMASE가 1보다 낮게 나타나 계절성 단순 예측 대비 우수한 예측 성과를 보였다.

개별 모형 간 비교에서는 STL 분해 기반 모형이 원자료 기반 모형보다 대체로 낮은 예측오차를 보였다. 특히 ARIMA 계열과 인공신경망 계열에서는 STL 분해를 적용한 모형의 RMSE가 원자료 기반 모형보다 낮게 나타났다. 원자료 기반 ARIMAX의 RMSE는 2,882.4였으나, STL+ARIMAX는 2,730.8로 나타났고, 원자료 기반 NNARX의 RMSE는 2,989.7이었으나, STL+NNARX는 2,504.7로 낮게 나타났다. 다만 ETS는 원자료 기반 모형의 RMSE가 분해 기반보다 다소 낮아, STL 분해가 예측 성과 개선에 기여할 수 있으나, 모형 구조와 자료 특성에 따라 달라질 수 있음을 시사한다.

따라서 배추 가격 예측에서 원자료 기반 단일 모형보다 STL 분해 기반과 분해-양상불 기반 모형의 예측 성과가 대체로 양호하였다. 이러한 결과는 배추 가격과 같이 계절성, 추세 변화 및 일시적 충격이 복합적으로 나타나는 농산물 가격 시계열의 경우, 시계열 분해와 예측 조합을 결합한 분해-양상불 전략이 예측오차를 완화하는 데 효과적일 수 있음을 시사한다.⁹⁾

표 5. 배추 가격 예측 정확도 평가 결과

모형			RMSE	MAE	sMASE
원자료 기반	ETS 지수평활	① ETS(A,N,A)	2,606.2	2,017.2	0.680
	ARIMA	② ARIMAX(1,1,2)(2,0,0)[12]	2,882.4	2,187.7	0.737
	인공신경망	③ NNARX(3,1,2)[12]	2,989.7	1,985.4	0.669
분해 기반	ETS 지수평활	④ STL-ETS(A,N,N)	2,635.8	2,010.5	0.678
	ARIMA	⑤ STL+ARIMA(1,1,2)(2,0,1)[12]	2,548.6	1,991.7	0.671
		⑥ STL+ARIMAX(0,1,5)(2,0,0)[12]	2,730.8	2,051.5	0.691
	인공신경망	⑦ STL+NNAR(3,1,2)[12]	2,673.6	1,957.2	0.660
		⑧ STL+NNARX(2,1,2)[12]	2,504.7	1,807.4	0.609
본 연구 제안 알고리즘	분해-양상불	평균값 조합	2,457.1	1,858.7	0.626
		중앙값 조합	2,472.2	1,858.9	0.627
		예측성과(MSE) 기반 조합	2,458.9	1,856.0	0.626

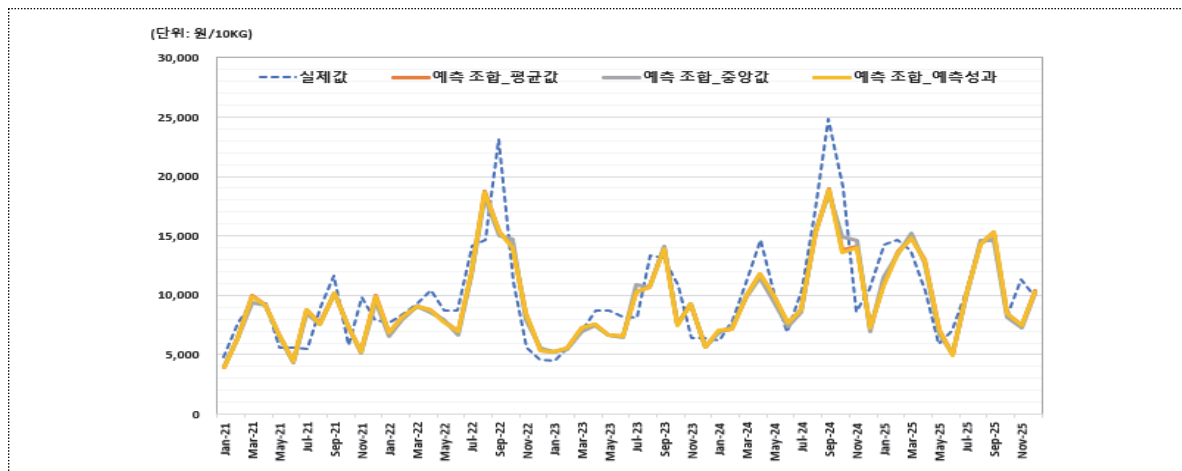
자료: 저자 작성.

〈그림 4〉는 본 연구에서 제안한 분해-양상불 기반 알고리즘의 배추 가격 예측치와 실제값을 비교한 결과이다. 제안 알고리즘은 실제 배추 가격의 전반적인 상승·하락 흐름과 계절적 변동 양상을 비교적 잘 반영하는 것으로 나타났다. 다만 가격이 급격하게 상승하는 일부 시점에서는 예측오차가 확대되었으며, 이는 기상 여건, 작황 변화 등 단기 수급 충격에 따른 영향으로 해석할 수 있다. 따라서 본 연구의

⁹⁾ 본 연구는 예측오차 평가 결과에서 확인된 모형 간 예측력 차이가 통계적으로 유의한지를 검토하기 위해 Modified Diebold and Mariano(MDM) 검정을 추가로 수행하였다. MDM 검정은 Diebold & Mariano(1995)의 예측력 차이 검정을 소표본 편의 문제에 대응하여 수정한 방법으로, 구체적인 내용은 Harvey et al.(1997)에 제시되어 있다. 구체적인 검정 통계량과 p-value는 부록 〈부표 2〉에 제시하였다. MDM 검정 결과, 본 연구의 분해-양상불 모형은 일부 개별 예측 모형에 비해 예측 손실이 낮게 나타났으며, 그 차이가 통계적으로 유의하게 나타났다. 다만 유의성은 비교 모형과 손실함수(제곱오차, 절대오차)에 따라 다르게 나타나, 본 연구의 제안 알고리즘이 모든 개별 모형에 대해 일관되게 통계적으로 우월하다고 해석하기보다는, 예측오차 지표와 MDM 검정 결과를 함께 고려하여 상대적 예측 성과를 판단하는 것이 적절하다.

제안 알고리즘은 배추 가격의 전반적인 변동 흐름과 계절적 패턴을 파악하는 정량적 예측 도구로 활용될 수 있으나, 단기 수급 충격에 따른 급격한 가격 변동을 보완적으로 점검하기 위해서는 현장 정보 기반의 예측 전략이 병행될 필요가 있다.

그림 4. 본 연구 제안 알고리즘의 예측치와 실제값 비교



자료: 저자 작성.

3.6. 본 연구 제안 알고리즘의 실무 적용 가능성 검토

본 연구의 제안 알고리즘은 예측 대상 월 직전까지 관측된 과거 가격 및 관련 정량 자료에 기반하여 사전에 설정된 분석 절차에 따라 해당 월의 평균 도매가격을 1개월 앞 예측하는 방식이다. 반면 KREI 농업관측센터 전망치는 예측 대상 월의 실제 평균가격이 확정되기 전에 전망 시점까지 수집된 산지 작황, 출하 동향, 시장 상황 및 전문가 판단을 종합적으로 반영하여 산출된다.¹⁰⁾ 따라서 두 전망치는 동일한 대상 월의 월평균 배추 도매가격을 예측하고 동일한 실제값을 기준으로 평가된다는 점에서 비교 가능하지만 활용 정보와 산출 방식에는 차이가 있다. 본 연구는 이러한 차이를 전제로 두 전망치의 상대적 예측 성과와 정보의 상호 보완 가능성을 검토하였다.

배추는 단기 수급 변화와 가격 급등락이 빈번하게 나타날 수 있는 품목이므로, 현장 정보를 반영하는 농업관측센터 전망치는 단기 수급 변동 요인을 직접적으로 포착하는 데 강점이 있다. 반면 본 연구의 제안 알고리즘은 동일한 절차에 따라 반복적으로 예측치를 산출하는 구조이므로, 전망 과정에서 체계적이고 일관된 정량적 참고 정보를 제공할 수 있다는 점에서 보완적 의의가 있다. 이에 본 연구는 농업관측센터 전망치와 제안 알고리즘의 예측 정확도를 비교하고, Encompassing 검정을 통해 두 예측치가 상호 보완적으로 활용될 수 있는지를 검토한다.

〈표 6〉은 본 연구 제안 알고리즘과 농업관측센터 전망치의 예측 정확도를 비교한 결과이다. 농업관

10) 구체적인 농업관측센터의 배추 가격 전망치 산출 구조는 부록 3에 제시하였다.

측센터의 배추 가격 전망치는 매년 3~12월에 공식적으로 제공되므로, 비교는 2021~2025년 중 전망치가 존재하는 공통 평가 표본을 기준으로 수행하였다. 분석 결과, 농업관측센터 전망치가 제안 알고리즘보다 전반적으로 낮은 예측오차를 보였다. RMSE 기준으로 농업관측센터 전망치는 2,362.6인 반면, 제안 알고리즘은 조합 방식에 따라 2,623.2~2,654.0 수준으로 나타났다. MAE와 sMASE 기준에서도 농업관측센터 전망치가 더 낮은 예측오차를 보여, 공통 평가 표본에서는 농업관측센터 전망치의 예측 성과가 제안 알고리즘보다 전반적으로 우수한 것으로 나타났다.

표 6. 본 연구 제안 알고리즘과 농업관측센터 전망치의 예측 정확도 비교

구분			RMSE	MAE	sMASE
본 연구 제안 알고리즘	분해-양상별	평균값 조합	2,623.2	2,025.6	0.683
		중앙값 조합	2,654.0	2,032.6	0.685
		예측성과(MSE) 기반 조합	2,623.2	2,020.4	0.681
KREI 농업관측센터 전망치			2,362.6	1,642.9	0.554

자료: 저자 작성.

〈표 7〉은 제안 알고리즘과 농업관측센터 전망치의 예측 정확도를 배추 작형별로 비교한 결과이다. 본 연구는 농업관측센터의 작형 구분을 준용하여 겨울배추는 1~4월, 봄배추는 5~6월, 여름배추는 7~10월, 가을배추는 11~12월로 설정하였다.

작형별 분석 결과, 겨울 작형에서는 제안 알고리즘의 RMSE가 농업관측센터 전망치보다 낮게 나타났다. 여름 작형에서는 두 전망치 간 예측오차 차이가 크지 않았다. 반면 봄과 가을 작형에서는 농업관측센터 전망치가 제안 알고리즘보다 낮은 예측오차를 보였으며, 특히 가을 작형에서 두 전망치 간 격차가 상대적으로 크게 나타났다. 가을 작형은 김장 수요 집중, 정부 수급안정대책 시행, 출하 조절 등 정책·시장 요인이 단기간에 집중적으로 작용하는 시기로, 이러한 요인들은 과거 자료와 정량 수급 변수만으로는 포착하기 어렵다. 따라서 가을 작형에서 나타난 농업관측센터 전망치의 예측 우위는 현장 조사 및 전문가 판단을 반영하는 농업관측센터 전망체계가 정성적 시장 정보를 반영하는 데 구조적 강점을 지니는 시사한다. 반면 겨울 작형에서 제안 알고리즘이 상대적으로 양호한 성과를 보인 것은 저장 배추 출하 구조와 연동된 겨울철 가격이 비교적 안정적인 계절 패턴을 형성하여 정량모형으로 포착 가능한 변동 구조를 갖기 때문으로 해석된다.

따라서 공통 표본 전체에서는 농업관측센터 전망치의 예측 성과가 상대적으로 우수하나, 작형 및 시기에 따라 두 전망치의 상대적 예측 우위가 달라질 수 있음을 보여준다. 이는 두 전망치가 단순한 대체관계에 있다기보다 활용 정보와 산출 방식의 차이로 인해 서로 다른 가격 변동 국면에서 각각의 강점이 나타나는 상호 보완적 관계에 있음을 시사한다.

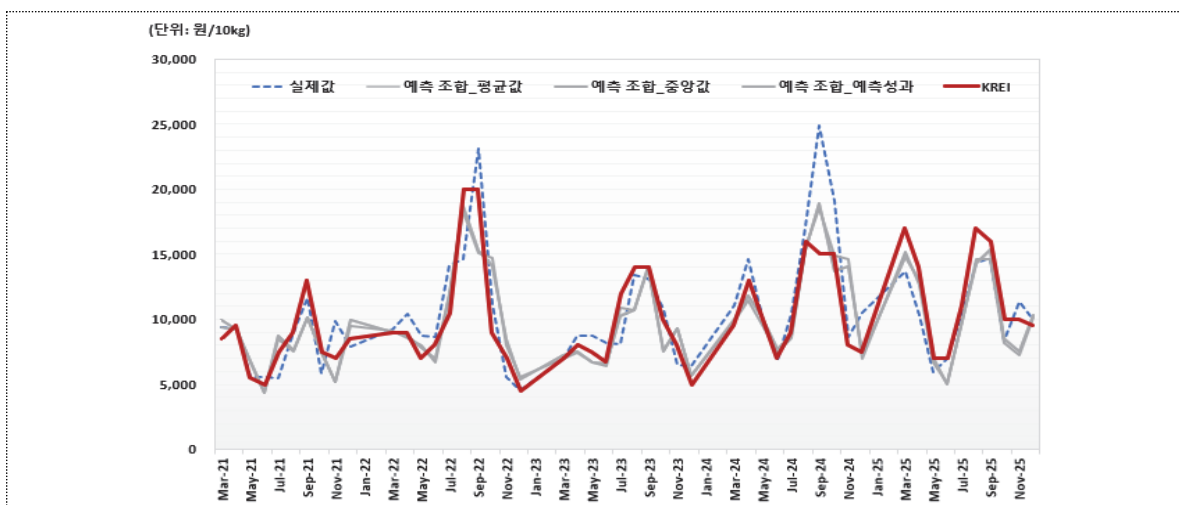
표 7. 본 연구 제안 알고리즘과 농업관측센터 전망치의 작형별 예측 정확도 비교

구분				RMSE	MAE	sMASE
본 연구 제안 알고리즘	분해- 앙상블	평균값 조합	겨울(1~4월)	1,422.6	1,095.3	0.369
			봄(5~6월)	1,390.6	1,272.0	0.429
			여름(7~10월)	3,191.7	2,497.3	0.842
			가을(11~12월)	3,141.6	2,694.8	0.908
		중앙값 조합	겨울(1~4월)	1,394.6	1,077.8	0.363
			봄(5~6월)	1,401.7	1,273.6	0.429
			여름(7~10월)	3,187.9	2,505.0	0.844
			가을(11~12월)	3,238.7	2,714.3	0.915
		예측성과(MSE) 기반 조합	겨울(1~4월)	1,429.0	1,096.4	0.370
			봄(5~6월)	1,398.9	1,280.1	0.431
			여름(7~10월)	3,196.0	2,491.1	0.840
			가을(11~12월)	3,131.8	2,680.6	0.904
KREI 농업관측센터 전망치			겨울(1~4월)	1,813.7	1,362.0	0.459
			봄(5~6월)	923.8	714.3	0.241
			여름(7~10월)	3,251.6	2,406.4	0.811
			가을(11~12월)	1,619.1	1,325.5	0.447

자료: 저자 작성.

〈그림 5〉는 본 연구 제안 알고리즘과 농업관측센터 전망치를 실제값과 비교한 결과를 제시한 것이다. 농업관측센터 전망치는 일부 가격 급등 시점에서 실제 가격 상승을 비교적 잘 반영하는 것으로 나타났다. 다만 모든 시점에서 농업관측센터 전망치가 실제값에 더 근접한 것은 아니며, 일부 시점에서는 제안 알고리즘의 예측치가 실제 가격 수준과 유사하게 나타나는 경우도 확인되었다. 이는 두 전망치가 서로 다른 정보 기반에 의해 산출되기 때문에 가격 변동을 포착하는 방식에도 차이가 있음을 보여준다. 따라서 예측 성과의 단순 비교와 함께, 두 전망치가 서로 다른 정보를 제공하는지를 검토할 필요가 있다.

그림 5. 본 연구 제안 알고리즘 및 농업관측센터 전망치와 실제값 비교



자료: 저자 작성.

〈표 8〉은 2021~2025년 공통 평가 기간을 대상으로 농업관측센터 전망치와 제안 알고리즘 간 Encompassing 검정 결과를 제시한 것이다. 분석 결과, λ 의 추정치는 0.347로 나타났으며 5% 유의수준에서 통계적으로 유의하였다. 이는 농업관측센터 전망치가 제안 알고리즘의 예측 정보를 완전히 포괄하기 어려우며, 제안 알고리즘이 농업관측센터 전망치에 포함되지 않은 추가적인 정량 정보를 제공할 가능성이 있음을 시사한다. 조합 가중치 측면에서 제안 알고리즘에 부여되는 가중치(λ)는 0.347, 농업관측센터 전망치에 부여되는 가중치($1 - \lambda$)는 0.653으로 해석할 수 있다.

표 8. 농업관측센터 전망치와 제안 알고리즘 간 Encompassing 검정 결과

구분	상수항	λ
회귀계수	371.9	0.347**
	(0.261)	(0.047)

주: () 안의 값은 p-value이고, **, *는 각각 5%, 10% 유의수준에서 통계적으로 유의함을 의미함.
자료: 저자 작성.

Encompassing 검정 결과는 두 전망치가 단순한 대체 관계가 아니라 상호 보완적 관계에 있음을 시사한다. 추정된 가중치(농업관측센터 0.653, 제안 알고리즘 0.347)는 농업관측센터 전망치를 중심으로 하되 제안 알고리즘의 정량 정보를 부가적으로 반영하는 구조로 해석되며, 이를 향후 두 전망치를 조합하여 활용하는 방안의 통계적 근거로 이해할 수 있다.

이상의 결과를 종합하면, 제안 알고리즘은 기존 배추 가격 전망체계를 대체하는 수단이 아니라 전망 과정에서 활용 가능한 정량적 보조도로 이해하는 것이 적절하다. 농업관측센터 전망치는 현장 조사 및 전문가 판단을 반영한다는 점에서 전반적으로 우수한 예측 성과를 보이며, 제안 알고리즘은 사전에 설정된 분석 절차에 따라 일관된 예측치를 산출할 수 있다는 점에서 보완적 가치를 지닌다. 향후 실제 전망 업무에서 두 전망치를 함께 축적하고 예측오차와 조합 가중치를 주기적으로 점검함으로써 전망의 정확도와 운영 안정성을 제고할 수 있을 것으로 기대된다.

한편, 본 연구의 분석 결과에서 확인된 바와 같이, 제안 알고리즘은 가격 급등락 시기에 예측오차가 확대되는 구조적 한계를 지닌다. 이는 단일 예측값만을 제공하는 점 예측 방식으로는 가격 급변 가능성을 사전에 파악하기 어렵다는 점을 시사한다. 따라서 예측값의 불확실성 범위를 함께 제시하는 확률예측 모형을 보완적으로 도입한다면, 가격 수준 전망뿐만 아니라 수급 위험에 대한 조기경보 기능도 강화할 수 있을 것으로 판단된다.

4. 요약 및 결론

농산물 수급 예측 정보는 생산 농가, 유통·가공업체, 정부 등 다양한 경제주체의 의사결정과 계획 수립에 활용되므로 예측 정확도 제고를 위한 실증 연구가 지속적으로 이루어지고 있다. 특히 배추는 월별 계절성과 작형별 출하 구조가 뚜렷하면서도 이상기상과 작황 변화 등으로 가격 변동성이 크게 나타나, 계절성, 추세, 불규칙 충격이 복합적으로 내재된 시계열이다. 이로 인해 원자료 가격에 단일 예측 모형을 직접 적용하는 방식만으로는 안정적인 예측 성과를 확보하기 어렵다. 이에 본 연구는 배추 월별 도매 가격의 단기 예측력을 개선하기 위해 시계열 분해와 예측 조합을 결합한 분해-앙상블 예측 알고리즘을 제안하고 예측 성과를 평가하였다.

분석에는 2001년 1월부터 2025년 12월까지의 월별 도매가격($n=300$)을 활용하였으며, 가격 분포의 비대칭성과 분산 확대를 완화하기 위해 로그 변환을 적용하였다. 제안 알고리즘은 STL 분해를 통해 시계열을 계절 성분과 계절조정 성분으로 구분한 뒤, 계절 성분에는 계절성 단순 예측(SNAIVE)을, 계절조정 성분에는 ETS 지수평활, ARIMA, ARIMAX, NNAR 모형을 적용하여 복수의 예측치를 산출하고, 이를 단순 평균, 중앙값, MSE 기반 가중 평균 방식으로 조합하였다.

분석 결과는 다음과 같다. 첫째, 표본 외 평가 구간에서 분해-앙상블 모형의 RMSE는 2,457.1~2,472.2로 원자료 기반 모형 및 분해 기반 개별 모형보다 대체로 낮은 예측오차를 보였으며, 모든 모형의 sMASE는 1보다 작아 계절성 단순 예측 대비 우수한 성과를 나타냈다. ARIMA 및 신경망 자기회귀계열에서는 STL 분해를 적용한 모형이 원자료 기반 모형보다 낮은 예측오차를 보였으나, ETS의 경우 원자료 기반 모형이 분해 기반 모형보다 낮게 나타났다. 이는 시계열 분해의 예측 성과 개선 효과가 모형 구조와 자료 특성에 따라 달라질 수 있음을 시사한다.

둘째, 제안 알고리즘과 KREI 농업관측센터 전망치를 비교한 결과, 공통 평가 표본 전체에서는 농업관측센터 전망치의 예측 성과가 상대적으로 우수하였으나, 작형별로는 겨울 작형에서 제안 알고리즘이 더 낮은 예측오차를 보이는 등 시기와 작형에 따라 상대적 우위가 달라지는 것으로 나타났다. 또한 Encompassing 검정에서 제안 알고리즘의 추정 가중치(λ)가 0.347로 5% 유의수준에서 통계적으로 유의하게 나타나, 제안 알고리즘이 농업관측센터 전망치에 포함되지 않은 추가적인 정량 정보를 제공할 가능성이 확인되었다. 이는 두 전망치가 단순한 대체 관계가 아니라 상호 보완적 관계에 있음을 시사한다.

본 연구는 분해-앙상블 전략을 국내 배추 가격 예측에 적용하여 단일 모형 대비 예측 안정성을 제고할 수 있음을 실증하였다는 점에서 의의가 있다. 나아가 Encompassing 검정을 통해 제안 알고리즘이 농업관측센터 전망치에 포함되지 않은 정량 정보를 제공할 수 있음을 확인함으로써 제안 알고리즘이 기존 전망 체계의 대체 수단이 아니라 전망 과정에서 객관적 근거를 보완하는 보조도구로 기능할 수 있음을 제시하였다.

최근 농산물 가격 예측 분야에서는 인공지능(AI) 기반 기법이 빠르게 도입되고 있으나, 정량 기법만

으로는 단기 수급 충격이나 정책 및 현장 요인을 즉각적으로 반영하는 데 한계가 있다. 따라서 향후에는 본 연구의 분해-양상불 알고리즘과 같은 정량 예측과 농업관측센터의 현장 조사 및 전문가 판단을 결합하는 인간-AI 협업(human-AI collaboration) 기반 전망체계를 단계적으로 구축함으로써 정량모형의 예측 일관성과 전문가 판단의 현장 대응력을 상호 보완하는 방향으로 발전시켜 나갈 필요가 있다.



이 저작물은 크리에이티브 커먼즈 저작자표시-비영리-변경금지 4.0 국제 라이선스에 따라 이용하실 수 있습니다.
This work is licensed under CC BY-NC-ND 4.0.

참고문헌

- 강동곤, 장영민, 이주석, 이성수. (2024). 장단기 기억 신경망을 사용한 다변수 데이터 농산물 가격 예측 모델. *전기전자학회논문지*, 28(3), 451-457. <http://doi.org/10.7471/ikeee.2024.28.3.451>
- 김배성. (2005). 채소가격 예측을 위한 응용기법별 예측력 비교. *농업경제연구*, 46(4), 89-113.
- 박소연, 김동언, 김민수. (2022). 생육시기 기상을 활용한 배추 가격 예측. *한국자료분석학회지*, 24(5), 1667-1679. <http://doi.org/10.37727/jkdas.2022.24.5.1667>
- 박종현, 임영우, 임도현, 최윤성, 안현철. (2023). 중장기 농산물 가격 예측을 위한 다단계 시계열 예측 모델. *한국컴퓨터정보학회논문지*, 28(2), 201-207. <http://doi.org/10.9708/jksoci.2023.28.02.201>
- 박지연, 박영구. (2013). *배추 무 예측 모형 고도화 방안*. 한국농촌경제연구원 기타연구보고.
- 박찬, 이경순. (2023). 농산물 가격의 특이치 탐지 및 처리를 통한 딥러닝 기반 가격 예측. *한국디지털콘텐츠학회논문지*, 24(8), 1899-1906.
- 서교, 이정재, 김태곤. (2008). 신뢰성 해석기법을 이용한 배추 가격폭등 예측 모형의 개발. *한국농공학회논문집*, 50(3), 91-104.
- 서울시농수산식품공사. 가락시장 유통정보. <<https://www.garak.co.kr/>>. 검색일: 2026. 2. 23.
- 유도일. (2016). 기상변인을 고려한 채소가격 예측 모형 개발. *농업경제연구*, 57(1), 1-24.
- 이창민, 송성광, 정성욱. (2022). 앙상블 Voting 기법을 활용한 배추 가격 예측에 관한 연구. *융합정보논문지*, 12(3), 1-10. <http://doi.org/10.22156/CS4SMB.2022.12.03.001>
- 장수희, 전희주, 조인호, 김동환. (2017). 비정형 농업기상자료를 활용한 배추 도매가격 예측 모형 연구. *한국데이터정보과학회지*, 28(3), 617-624. <http://doi.org/10.7465/jkdi.2017.28.3.617>
- 정호중, 정원호. (2016). 배추수급모형 추정을 통한 가격변동 시나리오 분석. *농업경영·정책연구*, 43(4), 897-920.
- 한국무역통계진흥원. 한국무역통계 정보포털 TRASS. <<https://www.bandtrass.or.kr>>. 검색일: 2026. 2. 23.
- 한석호, 김병률. (2004). 시장개방하의 배추 수급 모형과 전망. *농촌경제*, 27(3), 35-54.
- 한은수, 김선웅, 윤병삼. (2025). 예측 조합을 활용한 양념채소 가격 예측 및 예측성과 분석. *농촌경제*, 48(3), 1-31. <http://doi.org/10.36464/jrd.2025.48.3.001>
- 허다삼, 정승권. (2024). 고랭지배추의 생육 시기별 주산지 기상정보를 활용한 출하시기 도매가격 예측 모형. *한국정보기술학회논문지*, 22(9), 11-22. <http://doi.org/10.14801/jkiit.2024.22.9.11>
- Bates, J. M. & C. W. J. Granger. (1969). The Combination of Forecasts. *Journal of the Operational Research Society*, 20(4), 451-468. <https://doi.org/10.1057/jors.1969.103>
- Clemen, R. T. (1989). Combining Forecasts: A Review and Annotated Bibliography. *International Journal of Forecasting*, 5(4), 559-583. [https://doi.org/10.1016/0169-2070\(89\)90012-5](https://doi.org/10.1016/0169-2070(89)90012-5)
- Diebold, F. X. & R. S. Mariano. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253-263. <https://doi.org/10.1080/07350015.1995.10524599>
- Harvey, D. I., S. J. Leybourne & P. Newbold. (1997). Testing the Equality of Prediction Mean Squared Errors. *International Journal of Forecasting*, 13(2), 281-291. [https://doi.org/10.1016/s0169-2070\(96\)00719-4](https://doi.org/10.1016/s0169-2070(96)00719-4)
- Harvey, D. I., S. J. Leybourne & P. Newbold. (1998). Tests for Forecast Encompassing. *Journal of Business & Economic Statistics*, 16(2), 254-259. <https://doi.org/10.1080/07350015.1998.10524759>
- Hyndman, R. J. & G. Athanasopoulos. (2021). *Forecasting: Principles and Practice* (3rd ed.). OTexts.
- Stock, J. H. & M. W. Watson. (1999). A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series. In R. F. Engle & H. White (Eds.), *Cointegration, Causality, and Forecasting: A Festschrift in Honour of Clive W. J. Granger*, 1-44. Oxford University Press. <https://doi.org/10.1093/oso/9780198296836.003.0001>
- Sun, F., X. Meng, Y. Zhang, Y. Wang, H. Jiang & P. Liu. (2023). Agricultural Product Price Forecasting

- Methods: A Review. *Agriculture*, 13(9), 1671. <https://doi.org/10.3390/agriculture13091671>
- Wang, X., R. J. Hyndman, F. Li, & Y. Kang. (2023). Forecast Combinations: An Over 50-Year Review. *International Journal of Forecasting*, 39(4), 1518-1547. <https://doi.org/10.1016/j.ijforecast.2022.11.005>
- Yin, H. L., D. Jin, Y. H. Gu, C. J. Park, S. K. Han & S. J. Yoo. (2020). STL-ATTLSTM: Vegetable Price Forecasting Using STL and Attention Mechanism-Based LSTM. *Agriculture*, 10(12), 612. <https://doi.org/10.3390/agriculture10120612>

부록 1: 개별 예측 모형의 구조 및 선정 근거

부표 1. 개별 예측 모형의 최종 구조 및 선정 근거

모형 계열	최종 구조	선정 근거
ETS	원자료: ETS(A,N,A) STL 계절조정 성분: ETS(A,N,N)	- 각 자료에 적용 가능한 오차, 추세 및 계절성 구조를 비교하고 AIC가 가장 작은 모형을 선정
ARIMA· ARIMAX	원자료: ARIMAX(1,1,2)(2,0,0)[12] STL 계절조정 성분: ARIMA(1,1,2)(2,0,1)[12], ARIMAX(0,1,5)(2,0,0)[12]	- KPSS 검정과 OCSB 검정을 토대로 각각 (d=1), (D=0)으로 설정하고, (p,q,P,Q)는 AIC 최소 기준으로 선정 - ARIMAX의 외생변수 시차는 1~3개월의 연속 시차 중 검증 구간 RMSE가 가장 낮은 구성을 선택
NNAR· NNARX	원자료: NNARX(3,1,2)[12] STL 계절조정 성분: NNAR(3,1,2)[12], NNARX(2,1,2)[12]	- 비계절 입력 시차는 선형 AR 모형의 AIC를 기준으로 선정하고 월별 계절성을 반영하여 계절 시차를 설정 - 은닉노드 수는 2개로 고정하고 20회 반복학습을 실시

자료: 저자 작성.

부록 2: 예측력 차이 검정(Modified Diebold and Mariano: MDM)

〈부표 2〉는 본 연구 제안 알고리즘과 개별 비교 모형 간 예측력 차이의 통계적 유의성을 검토하기 위해 수행한 MDM 검정 결과이다. MDM 검정은 두 모형 간 예측 손실의 차이가 통계적으로 유의한지를 평가한다. 본 연구에서는 제안 알고리즘(예측 조합_예측성과(MSE) 기반 조합)을 기준 모형으로, 원자료 기반 및 분해 기반 개별 모형을 비교 대상 모형으로 설정하였다. 손실함수로는 제곱 오차와 절대 오차를 각각 적용하였으며, MDM 통계량이 음(-)의 값을 보이는 경우 기준 모형인 제안 알고리즘의 손실값이 비교 모형보다 작음을 의미한다.

부표 2. 예측력 차이의 통계적 유의성 검정: MDM 검정 결과

기준 모형: 본 연구 제안 알고리즘(“예측 조합_예측성과(MSE)”)			제곱 오차	절대 오차
비교 대상 모형				
원자료 기반	ETS 지수평활	① ETS(A,N,A)	-0.661	-1.186
			(0.256)	(0.120)
	ARIMA	② ARIMAX(1,1,2)(2,0,0)[12]	-1.689	-1.709
			(0.048)**	(0.046)**
인공신경망	③ NNARX(3,1,2)[12]	-1.482	-0.658	
		(0.072)*	(0.257)	
분해 기반	ETS 지수평활	④ STL-ETS(A,N,N)	-0.679	-1.042
			(0.250)	(0.151)
	ARIMA	⑤ STL+ARIMA(1,1,2)(2,0,1)[12]	-0.882	-1.605
			(0.191)	(0.057)*
		⑥ STL+ARIMAX(0,1,5)(2,0,0)[12]	-1.792	-2.071
			(0.039)**	(0.021)**
인공신경망	⑦ STL+NNAR(3,1,2)[12]	-1.406	-0.931	
		(0.083)*	(0.178)	
	⑧ STL+NNARX(2,1,2)[12]	-0.283	0.443	
		(0.389)	(0.670)	

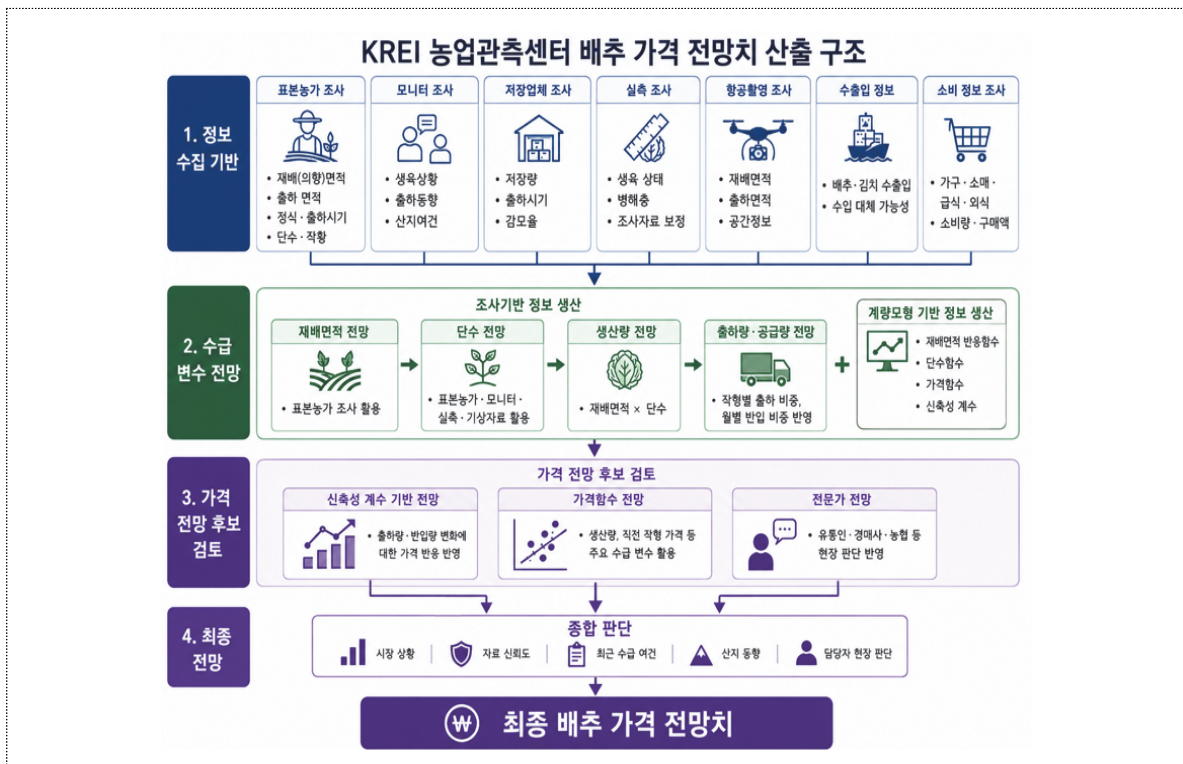
주: **, *는 각각 5%, 10% 유의수준을 의미함. () 안의 값은 p-value임. MDM 통계량의 음(-)의 값은 기준 모형인 본 연구의 제안 알고리즘의 손실값이 비교 모형보다 작음을 의미함.

자료: 저자 작성.

부록 3: KREI 농업관측센터의 배추 가격 전망치 산출 구조

KREI 농업관측센터의 배추 가격 전망치는 조사자료, 계량모형 분석 결과, 현장 정보 및 전문가 판단을 종합하여 산출한다. 조사자료는 표본농가·모니터 조사, 항공촬영, 소비 정보 등을 통해 재배면적, 단수, 생산량 및 출하 여건을 파악하는 데 활용되며, 계량모형은 수급 변수 간의 관계를 정량적으로 분석하여 이를 보완한다. 수급 변수에 대한 전망은 우선, 재배면적과 단수 전망치를 결합하여 생산량을 산출한 후 작형별 출하 비중, 월별 출하 패턴, 저장량 및 도매시장 반입 동향 등을 반영하여 월별 출하량과 공급량을 전망한다. 가격 전망은 출하량·반입량 변동에 가격 신축성 계수를 적용하는 방식과 생산량, 반입량, 직전 작형 가격 등을 활용한 가격함수 방식으로 산출한다. 또한 산지유통인, 경매사, 농협 등에서 입수한 현장 정보를 토대로 정량화하기 어려운 시장 상황을 반영한 전문가 전망을 병행한다. 최종 전망치는 복수의 가격 전망 후보를 바탕으로 농업관측센터 품목 담당자가 자료의 신뢰도와 최근 수급 여건을 종합적으로 검토하여 결정한다.

부도 1. 농업관측센터 배추 가격 전망치 산출 구조



주: 본 그림의 도식화 과정에서 생성형 AI 도구인 ChatGPT를 보조적으로 활용하였으며, 최종 내용을 저자가 검토 수정하였음.