

Prophet 모형을 이용한 정보시스템 접속 건수 예측에 관한 연구

(A study on predicting the number of accesses to information systems using the Prophet model)

정성윤*

(Seong Yun Jeong)

요약

정부의 정보화 예산을 지원받고 있는 건설기술정보시스템도 업무 성과 진단을 받기 위해서는 정확한 성과 목표 설정이 중요하다. 본 연구는 시스템 접속 건수를 업무 성과 지표로 정하고 최근 2년간의 접속 건수를 예측하고자 하였다. 이를 위해 2025년 1월부터 2026년 12월까지 접속 건수를 수집하였고, 수집한 데이터가 비정상적 추세형 시계열 데이터라고 판단하였다. 계량 경제학에서 시계열 데이터의 특성에 대표적으로 사용하는 ARIMA 모형, GARCH 모형 및 Prophet 모형을 선정하였다. 이들 모형을 적용하여 2025년과 2026년 간의 접속 건수의 예측치와 변동성을 구하였다. 예측치를 가지고서 ARIMA 모형과 Prophet 모형을 비교하였다. 접속 건수에 대해서는 Prophet 모형이 ARIMA 모형보다 좀 더 통계적으로 유의미한 것을 확인하였다. 또한, GARCH 모형을 적용하여 예측치의 변동성 특성을 분석하였다. 접속 건수처럼 일정한 수준으로 증감하는 추세형 시계열 데이터는 변동성을 추정할 수 있으나 변동성의 강도와 지속성을 반영하는 데에는 유의하지 않다는 결과를 얻었다.

■ 중심어 : 시계열 분석 모형 ; ARIMA 모형 ; GARCH 모형 ; Prophet 모형 ; 미래 예측

Abstract

Even for the Construction Technology Information System, which receives support from the government's informatization budget, accurate performance target setting is crucial for assessing performance. This study defined the number of system accesses as a performance indicator and aimed to predict the number of accesses over the past two years. To achieve this, accesses were collected from January 2025 to December 2026 and determined to be non-stationary trend-type time series data. The ARIMA, GARCH, and Prophet models, which are representative models used in econometrics to characterize time series data, were selected. These models were applied to obtain forecasts and volatility of accesses between 2025 and 2026. The forecasts were compared between the ARIMA and Prophet models. The Prophet model was found to be more statistically significant than the ARIMA model for accesses. Furthermore, the GARCH model was applied to analyze the volatility characteristics of the forecasts. We found that trend-type time series data, such as the number of connections, which increase and decrease at a constant level, can estimate volatility, but are not significant in reflecting the intensity and persistence of volatility.

■ keywords : Time Series Analysis Model ; ARIMA Model ; GARCH Model ; Prophet Model ; Future Prediction

I. 서론

정부가 2000년대 초반부터 전자정부 구현을 위한 여러 정보화 계획을 발표하면서 정보시스템 개발

이 급속히 많아졌다. 이런 사유로 정보시스템의 유지관리 비용이 급격히 늘어나게 되었다. 정부는 정보시스템을 효과적으로 운영하기 위해 2018년 2월에 「전자정부 성과관리 지침」을 발표하였다[1]. 이 지침에 따라 전자정부 정보화 사업을 통해 개발

* 정회원 한국건설기술연구원 미래스마트건설연구본부 연구위원 정성윤

이 논문은 2025년도 정부(국토교통부)의 출연금으로 한국건설기술연구원이 수행한 "25 건설기술정보 DB 및 서비스 시스템 운영" 과제의 연구임
접수일자 : 2025년 08월 22일

게재확정일 : 2025년 09월 03일

교신저자 : 정성윤 e-mail : syjeong@kict.re.kr

한 정보시스템은 매년 업무 성과를 진단받고 있다. 진단 결과에 따라 정보시스템을 계속 유지, 재개발, 기능 고도화 또는 폐기 등 유지관리 방향이 결정된다. 이 결정에 따라 정보시스템의 예산을 편성한다. 따라서 업무 성과 목표를 얼마나 정확하게 설정하느냐에 따라 시스템의 유지관리 방향과 예산 규모에 영향을 받는다. 정부의 정보화 예산을 지원받고 있는 건설기술정보시스템[2]도 이 지침에 따라 매년 업무 성과를 진단받고 있다. 본 연구는 2025년과 2026년의 성과 목표로 접속 건수를 얼마나 정할지를 결정하고자 하였다. 이를 위해 업무 성과의 목표치를 정량적으로 제시하기 위해 시스템의 접속 건수를 대국민 서비스의 지표로 정하였다. 2015년 1월부터 2024년 12월까지 월 단위로 시스템 접속 건수를 수집하였다. 수집한 데이터를 볼 때 시계열 데이터 유형 중 비정상적(non-stationary) 추세형에 가까운 특징이 있다고 판단하였다. 계량 경제학에서 비정상적 추세형 시계열 예측에 대표적으로 사용하는 ARIMA 모형과 GARCH 모형, 비교적 최근에 발표한 Prophet 모형을 예측 모형으로 선정하였다. 이들 모형에 관한 이론과 선행 연구 사례를 살펴보았다. 다음으로, 수집한 접속 건수에 관한 기초 통계량을 분석하였다. 파이썬의 ARIMA, GARCH 및 Prophet 라이브러리를 적용하여 2025년 1월부터 2026년 12월까지의 월 단위 예측치와 변동성을 추정하였다. 예측치의 기술 통계량의 특징을 분석하여 Prophet 모형이 ARIMA 모형보다 통계적 유의성이 우수하다는 결과를 얻었다. 또한, GARCH 모형을 통해 접속 건수의 예측치가 일정한 수준의 패턴으로 증감하는 것을 확인하였다. 일정한 수준으로 증감하는 과거 접속 건수의 변동성은 현재의 변동성에 영향을 주지만 변동성의 강도와 지속성은 통계적으로 유의하지 않다는 결과를 얻었다. 따라서 비정상적 추세형 시계열 데이터의 특징을 갖는 접속 건수의 예측에는 Prophet 모형이 적합하다고 판단하였다.

II. 이론 및 선행 연구 사례 고찰

1. 비정상 시계열 모형의 이론

앞으로 일어날 현상을 예측하기 위해서는 보통 과거의 데이터를 활용한다. 시계열 예측 모형은 시간의 흐름에 따라 순차적으로 기록된 데이터에 내재된 특성을 통계적으로 계산하여 미래의 예측치를 추정한다. 이때 시계열 데이터의 구성 요소에 따라 추세형, 계절형, 순환형 및 불규칙형으로 구분한다. 또한, 시계열 데이터의 평균과 분산, 공분산이 일정하냐에 따라 정상 또는 비정상 여부를 판단한다. 금융, 재화, 공급/수요, 에너지, 거시경제 등의 시계열 데이터는 대부분 비정상 특성을 갖는다. 시계열 예측 모형을 많이 다루는 계량 경제학에서는 비정상 시계열 데이터를 표본으로 하여 미래의 예측치와 변동성을 추정하는데 ARIMA(Autoregressive Integrated Moving Average) 모형, GARCH(Generalized Autoregressive Conditional Heteroskedasticity) 모형, VAR(Vector Autoregression) 모형을 대표적으로 사용한다. 이 중 VAR 모형은 2개 이상의 비정상 시계열 변수 간의 상호 관계를 분석하여 변수 간에 미치는 영향 정도를 확률적으로 추정한다. 본 연구에서는 접근 건수라는 단일 변수를 사용한다. 따라서 VAR 모형은 본 연구에는 적합하지 않다고 판단하였다. 한편, Prophet 모형은 짧은 기간의 예측치 추정에 적합하도록 고안되었다. 본 연구는 추세형 비정상 시계열 데이터를 이용하여 짧은 기간의 예측치와 변동성을 구하는데 ARIMA 모형, GARCH 모형 및 Prophet 모형이 적합하다고 판단하였다. 이들 모형에 관한 이론과 선행 연구 사례를 살펴보았다.

(1) ARIMA 모형

ARIMA 모형은 1970년 George Box와 Gwilym Jenkins가 고안한 Box-Jenkins 모형[3]을 확장한 시계열 예측 도구를 말한다. 이 모형은 식 (1)과

같이 자기회귀(Autoregressive, AR(p)), 차분(Integrated, I(d)), 이동평균(Moving Average, MA(q))의 구성 요소를 조합하여 시간의 흐름에 따라 평균과 분산이 변하는 데이터를 분석하여 미래를 예측한다.

ARIMA(x , order, include, mean) (식 1)

여기서, x 는 일변량 시계열 데이터를, order는 입력될 차수를 order= $\alpha(p, d, q)$ 의 형태를 의미한다. 이때 p 는 AR의 차수를, q 는 MA의 차수를 말한다. include, mean은 모형에 절편 항을 포함할 것인지를 결정하는 데 사용한다.

ARIMA 모형은 짧은 기간에 시계열 데이터의 선형 추세, 예측 및 계절적 특성을 포착하여 정확하게 예측할 수 있어 계량 경제학에서 가장 대표적으로 사용한다. 하지만 장기간의 예측에서는 오차가 크고 변동성 군집과 불균등 분산에는 취약하다는 단점이 있다.

(2) GARCH 모형

1982년에 Engle이 발표한 ARCH 모형[4]을 1986년에 Bollerslev가 시계열 데이터에 존재하는 오차의 실제 크기를 일반화하여 새롭게 확장한 GARCH 모형[5]을 만들었다. GARCH 모형은 식 (2)와 같이 분산 방정식을 사용하여 과거의 오차항 제곱과 과거의 조건부 분산을 이용한 이분산성(heteroskedasticity)과 시계열 데이터의 변동성 군집(volatility clustering) 현상을 모형화하여 예측한다. 이 모형은 시간 경과에 따른 변동성 변화를 정확하게 추정할 수 있어 계량 경제학에서 시계열 예측에 많이 사용한다.

GARCH(p, q) (식 2)

여기서, p 는 GARCH (p)항의 차수로써 과거 시점의 조건부 분산이 현재 시점의 조건부 분산에 미치는 영향 정도를 말한다. q 는 ARCH 항(q)의 차수로써 과거 시점에서 오차항 제곱이 현재 시점의 조건부 분산에 미치는 영향 정도를 의미한다.

GARCH 모형은 과거 시계열 데이터의 변동성이 현재의 변동성에 영향을 주는 자기회귀 특성

을 통해 변동성 군집 현상을 추정한다. 이 군집 현상을 토대로 시계열 데이터의 변동성을 예측한다. 변동성 군집 현상으로 인해 변동성의 크고 작음에 따라 변동성 추이 변화가 지속되는 경향이 있다. 따라서 변동성의 폭이 클수록 변동성을 정확하게 추정할 수 있다. 시계열 데이터의 특성에 따라 변동성의 비대칭 효과를 설명하기 위해 EGARCH, TGARCH, GJR-GARCH, GARCH-M 등 ARCH를 확장한 모형을 사용하기도 한다.

(3) Prophet 모형

미국 페이스북 회사는 2018년에 비즈니스 예측에 특화된 Prophet 모형을 발표하였다. 이 모형은 추세형, 계절형, 불규칙형의 시계열 데이터의 구성 요소를 가지고서 시계열을 예측하는 분석 도구이다. 이 모형은 장기보다는 짧은 기간에 추세의 변동성을 예측하는 데 적합하다는 장점을 가지고 있다. Prophet 모형은 응답 변수(시계열 데이터)와 예측 변수(시간 관련 특징) 간의 비선형 관계를 허용함으로써 선형 모형을 확장하는 개념을 기반으로 두었다. 이 모형은 식(3)과 같이 추세형, 계절형, 불규칙형의 구성 요소를 각각 분리하여 계산한 후에 이를 합산하여 미래의 예측치를 추정한다.

$y(t) = g(t) + s(t) + h(t) + \epsilon_t$ (식 3)

여기서, $g(t)$ 는 지수 함수를 적용한 추세형 모형을, $s(t)$ 는 푸리에 급수를 적용한 계절형 함수를, $h(t)$ 는 비주기적인 휴일이나 이벤트와 같이 특정 시점에서 발생하는 불규칙형 함수를 말한다. ϵ_t : 정규분포를 가지면서 불규칙 변동(error)에서 나타내는 오차항을 의미한다.

Prophet 모형은 시계열 데이터에 존재하는 이상치로 인한 변동성의 급격한 변화를 탐지하여 유연하고 신뢰할 수 있는 예측치를 계산한다. 이런 이유로 ARIMA 모형과 같이 추세형과 계절형 시계열 데이터를 예측하는 데 많이 사용한다. 페이스북 회사는 파이썬에서 Prophet 모형을 활용할 수 있도록 Prophet 라이브러리를 공개하고 있다[6].

2. 선행 연구 고찰

본 연구에서는 구글의 학술 검색 사이트에서 ARIMA 모형, GARCH 모형, Prophet 모형과 정보시스템의 접속 또는 이용 건수를 검색어로 사용하여 연구 사례를 찾았다. 검색 결과, 국내외에서 발표한 정보시스템 접속 건수 연구 사례를 찾지 못했다. 표 1은 ARIMA 모형, GARCH 모형, Prophet 모형을 결합하여 예측한 유사 연구 사례를 정리한 것이다.

표 1. 선행 연구 사례

예측 모형	주요 연구 내용
ARIMA와 GARCH 결합모형	- ARIMA 모형과 GARCH 모형을 사용하여 주식 가격 예측(Soumith 외[7]) - 일별 S&P500 지수 예측을 위한 순환신경망 기반 비선형 ARMA-GARCH 모형을 제안(정용진[8]) - 계절형 ARIMA-GARCH 모형을 활용하여 일일 발전용 가스 공급량 예측(김시영 외[9]) - ARMA-GARCH 모형을 활용하여 중국 금 선물 시장 가격변동 분석 및 예측(이문화 외[10]) - ARMA 모형과 GARCH 모형을 활용하여 양파, 마늘, 마른 고추, 배추, 무 등 농산물의 가격변동 분석(이문석[11])
ARIMA와 Prophet 결합모형	- Prophet 모형과 ARIMA 모형을 활용하여 코로나바이러스 질병 예측과 Prophet 모형이 ARIMA 모형보다 우수하다는 것을 입증(Satrio 외[12]) - 계절형 ARIMA 모형과 Prophet 모형을 활용하여 대기 오염 수준 예측(Samal 외[13]) - 주가 예측을 위한 ARIMA, Prophet, LSTM 모형을 비교하여 Prophet 모형이 정확한 가치 예측보다는 추세 분석에 최적화되었고 시장 방향 및 계절적 패턴 파악에 장점이 있다는 것을 보여줌(Somisett 외[14]) - 전력 발전 시스템의 전력량 예측을 통한 Prophet 모형과 SARIMA 모형을 비교하여 Prophet 모형이 SARIMA 모형보다 정확도가 높다는 것을 입증(Atasever 외[15]) - Prophet 모형을 이용한 마늘 가격의 장기 예측 및 ARIMA 모형과 LSTM 모형과의 차이 분석(오승원 외[16])
Prophet과 GARCH 결합모형	- 주간 현물 원유 원가 예측을 위해 ARMA-GARCH-MLP(Multi-Layer Perceptron) 결합모형을 활용하여 얻은 결과를 Prophet 모형에 적용하여 Prophet 모형의 정확도를 평가(Sari-Hassoun 외[17]) - 주가 수익률 예측을 위한 GJR-GARCH 모형과 Prophet 모형을 비교하여 일일 시계열 데이터 따라 GJR-GARCH 모형과 Prophet 모형의 정확도가 다르다는 것을 보여줌(Tasi'u 외[18])

표 1에서 보듯이 주식, 공급/수요, 전력수요, 농산물, 전염병, 거시경제 등 여러 분야에서 미래의 변동성 예측을 위해 전통적인 시계열 모형인

ARIMA 모형과 GARCH 모형 및 비교적 최근에 발표된 Prophet 모형을 활용하였다. 특히, 표 1의 사례처럼 각각의 모형을 개별적으로 적용하기보다는 변동성 군집 및 이분산성(Heteroskedasticity)에 취약한 ARIMA 모형과 시계열 데이터의 평균과 선형 추세를 포착하면서 예측하는데 한계가 있는 GARCH 모형을 상호 보완하기 위해 두 모형을 결합(hybrid)하여 미래의 변동성을 예측하였다. 즉, 개별적인 모형을 활용하기보다는 결합한 모형을 활용할 때 예측의 정확도를 높일 수 있었다. 다만, 국내에서는 Prophet 모형과 GARCH 모형을 결합하여 미래의 변동성을 예측한 연구 사례는 찾지 못하였다. 또한, 대부분의 연구는 새로운 시계열 예측 이론과 모형은 제시하기보다는 시계열 데이터의 특성에 맞는 모형을 서로 결합하거나 기존의 시계열 모형을 확장한 후에 예측의 정확도나 성능을 검증하는 사례가 많았다. 딥러닝을 응용하는 사례가 많아지면 LSTM을 접목하여 시계열 데이터를 학습하여 예측하는 연구가 늘고 있다. 한편, Satrio 외[12]과 Atasever 외[15]의 연구 사례에서 짧은 기간에서의 예측은 ARIMA 모형보다는 Prophet 모형이 우수하다고 하였다. 다만, Tasi'u 외[17]가 말했듯이 시계열 데이터의 특성에 따라 미래 예측의 정확도가 달라질 수 있다. 이러한 현상은 과거의 시계열 데이터가 갖는 변동성의 증감 폭과 변동성이 지속적으로 미치는 기간에 기인한 것으로 보였다.

III. 기초 통계량 및 결과 해석

1. 기초 통계량의 특성

본 연구는 검색 건수를 대상으로 한 성과 목표를 예측하기 위해 2015년 1월부터 2024년 12월까지 120개의 매월 시스템의 접속 건수를 수집하였다. 표 2와 그림 1은 접속 건수의 통계적 특성과 시간 흐름에 따른 접속 건수의 변화에 대한 추세선을 나타낸 것이다.

표 2. 접속 건수의 기초 통계량

구분	통계치	구분	통계치
평균	338,883.8	표준편차	213,070.2
표준오차	19,450.6	분산	45,398,918,086.3
중앙값	339,329.5	첨도	-0.8
최솟값	51,267.0	왜도	0.5
최댓값	906,377.0	범위	855,110.0
관측수	120.0	신뢰수준(95.0%)	38,514.1

수집한 접속 건수의 표준편차, 분산, 첨도, 왜도를 가지고서 분포도의 퍼짐 정도와 형태를 파악하였다. 표준오차의 값이 평균의 5.75% 수준이므로 표본 평균의 값이 서로 비슷하였고 모집단의 평균에 매우 가깝다고 볼 수 있다. 신뢰수준(95.0%)의 값이 38,514.1이므로 신뢰구간으로 [300,369.7, 377,397.9]로 추정할 수 있다. 평균치는 신뢰구간에 존재하였고 중앙값은 신뢰구간에 인접하였다. 한편, 왜도가 양(+)의 값을 가졌으므로 오른쪽으로 약간 치우친 형태의 분포도를 가졌다. 첨도는 음(-)의 값을 가졌으므로 분포도의 꼬리가 얇은 형태를 가졌다. 왜도와 첨도의 값이 0에 가까우므로 정규분포의 행태를 가졌다. 다음으로, 그림 1에서 보면 2019년부터 접속 건수가 지속적으로 증가하다가 2023년 9월부터 감소하였다. 특히, 특정 시점에 변동의 증감 폭이 상당히 컸다. 이는 DB 확충 여부, 회원가입 기능 개설과 폐기 등 외부적인 요인에 의해 나타난 현상으로 보인다. 한편, 그림 1에서 보듯이 계절형, 순환형, 불규칙정보다는 시간의 흐름에 따라 접근 건수가 바뀌는 추세(Trend)형의 특성을 가졌다. 이러한 추세형 특성은 계절성이나 이벤트성과 같은 특정 시기마다 특정한 패턴을 보이기 보다는 기준시점에 따라 평균과 분산이 바뀌는 비정상적인 특성을 갖는다고 볼 수 있다.

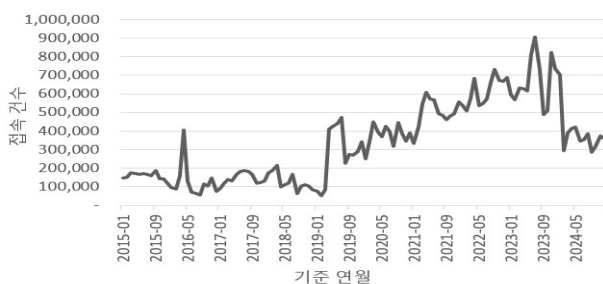


그림 1. 접속 건수의 변화 추이

2. 예측 결과 해석

본 연구는 예측 결과를 성과 목표로 바로 적용하기 위해 2025년과 2026년이라는 비교적 짧은 기간에 접속 건수가 어떻게 변화하는지를 보고자 하였다. 이를 위해 2.0.4 버전의 ARIMA 라이브러리, 1.1.4 버전의 Prophet 라이브러리를 적용한 파이썬을 사용하여 예측치를 계산하였다. 표 3은 ARIMA 모형과 Prophet 모형을 통해 얻은 예측치를 나타낸 것이다. 표 3을 보면 ARIMA 모형과 Prophet 모형의 예측 건수와 변동율의 차이가 크지 않았다.

표 3. ARIMA 모형과 Prophet 모형을 이용한 예측 건수 비교

기준연월	ARIMA 모형		Prophet 모형		차이
	예측건수	변동율(%)	예측건수	변동율(%)	
2025-01	524,668		575,912		51,245
2025-02	650,996	24.08	648,914	12.68	-2,082
2025-03	730,611	12.23	703,347	8.39	-27,264
2025-04	668,462	-8.51	665,783	-5.34	-2,679
2025-05	709,188	6.09	689,713	3.59	-19,475
2025-06	730,074	2.95	702,934	1.92	-27,140
2025-07	637,853	-12.63	632,905	-9.96	-4,948
2025-08	591,190	-7.32	623,430	-1.50	32,240
2025-09	582,082	-1.54	615,382	-1.29	33,300
2025-10	681,678	17.11	674,185	9.56	-7,493
2025-11	676,457	-0.77	675,905	0.26	-552
2025-12	691,675	2.25	671,885	-0.59	-19,790
2026-01	568,447	-17.82	628,223	-6.50	59,776
2026-02	710,798	25.04	705,917	12.37	-4,881
2026-03	795,625	11.93	758,874	7.50	-36,751
2026-04	727,102	-8.61	722,990	-4.73	-4,112
2026-05	772,073	6.18	748,980	3.59	-23,093
2026-06	794,369	2.89	762,749	1.84	-31,620
2026-07	691,314	-12.97	687,390	-9.88	-3,924
2026-08	640,001	-7.42	676,036	-1.65	36,035
2026-09	631,855	-1.27	670,181	-0.87	38,327
2026-10	741,774	17.40	734,319	9.57	-7,455
2026-11	734,705	-0.9%	736,686	0.32	1,981
2026-12	747,220	1.70	728,517	-1.11	-18,703
평균	684,592	0	685,048	0	456

표 4는 표 3을 기초데이터로 하여 얻은 기술 통계를 나타낸 것이다.

표 4. 예측된 접속 건수에 관한 기술 통계량 비교

구분	ARIMA 예측	Prophet 예측	의미
평균	684,592	685,048	유사한 수준으로 예측
표준오차	14,486	9,889	Prophet 예측이 더 정확함
표준편차	70,966	48,445	Prophet 예측이 더 안정적임
첨도	-0.23	-0.34	꼬리가 얇은 정규분포 형태
왜도	-0.49	-0.30	왼쪽으로 치우친 정규분포 형태
신뢰수준(95.0%)	29,966	20,456	ARIMA 예측이 더 정밀함

Prophet 모형에서의 표준오차, 표준편차, 첨도, 신뢰수준은 ARIMA 모형보다 작았다. 반대로, 평균과 왜도의 차이 폭은 ARIMA 모형보다 크

다. ARIMA 모형과 Prophet 모형은 모두 과거의 접속 건수의 분포도는 오른쪽으로 약간 치우친 형태를 가졌으나 예측 결과에서는 왼쪽으로 치우친 형태의 분포도를 가졌다. 이는 예측 건수가 정규분포의 행태에 가깝다고 볼 수 있다. 표준편차는 실제값이 평균과의 차이를 나타내는 표준오차를 가지고서 접속 건수의 변동성을 추정하였다. 추정한 변동성은 Prophet 모형이 ARIMA 모형보다 작았다. 표본 평균의 정확도를 나타내는 표준오차를 볼 때, Prophet 모형이 ARIMA 모형보다 더 신뢰성이 높았다. 따라서 예측한 접속 건수에 관한 기술 통계량을 종합적으로 비교했을 때, 두 모형 모두 예측 평균값은 유사하지만, 예측의 정밀도와 신뢰성 측면에서 Prophet 모형이 ARIMA 모형보다 더 뛰어난 것으로 볼 수 있다. 다음으로, 그림 2는 GARCH 모형을 적용하여 얻는 변동성 결과를 나타낸 것이다.

GARCH Model Summary:

Constant Mean - GARCH Model Results				
Dep. Variable:	returns	R-squared:	0.000	
Mean Model:	Constant Mean	Adj. R-squared:	0.000	
Vol Model:	GARCH	Log-Likelihood:	-580.285	
Distribution:	Normal	AIC:	1168.57	
Method:	Maximum Likelihood	BIC:	1179.69	
		No. Observations:	119	
Date:	Thu, Aug 14 2025	Df Residuals:	118	
Time:	14:32:36	Df Model:	1	

Mean Model				
	coef	std err	t	P> t 95.0% Conf. Int.
mu	2.5731	2.591	0.993	0.321 [-2.505, 7.651]

Volatility Model				
	coef	std err	t	P> t 95.0% Conf. Int.
omega	464.5635	134.048	3.466	5.290e-04 [2.018e+02, 7.273e+02]
alpha[1]	1.0000	0.541	1.847	6.472e-02 [-6.107e-02, 2.061]
beta[1]	0.0000	3.925e-03	0.000	1.000 [-7.692e-03, 7.692e-03]

그림 2. 변동성 결과

그림 2에서 결정계수인 R^2 와 조정된 R^2 는 모두 0의 값을 가졌다. 이는 접속 건수라는 하나의 변수만을 사용하였기 때문에 변동성(분산)을 설명하지 못한 결과라고 볼 수 있다. 그림 2에서의 변동성 모형은 접속 건수의 변화에 따른 변동성(분산)을 어떻게 모형화했는지를 보여주는 상수항이다. 변동성 모형에서 모든 계수(coefficient)는 양의 값을 가졌다. 이는 과거의 변동성이 현재

변동성에 양의 값으로 영향을 미친다고 해석할 수 있다. omega는 과거의 데이터가 갖는 변동성 증감의 폭과 무관하게 존재하는 장기 평균 변동성을 나타내는 척도이다. omega의 계수 값이 클수록 변동성도 크다고 해석한다. 따라서 추정한 계수 값이 양수이고 크지만, 계수 값에 비해 표준오차의 값은 작으므로 추정한 계수 값을 신뢰할 수 있다고 볼 수 있다. 또한, t-통계량이 2보다 크므로 5% 수준이고, $P>|t|$ 의 값이 0.05보다 작으므로 해당 계수는 통계적으로 유의미함을 갖는다고 해석할 수 있다. 신뢰구간이 0을 포함하지 않는다면 신뢰구간이 양의 범위 내에 존재하므로 통계적으로 유의미하다고 판단할 수 있다. 따라서 omega의 결과를 볼 때 추정한 변동성은 통계적으로 유의미함을 갖는다고 해석할 수 있다. 다음으로, alpha[1]는 이전 기간의 예측 오차가 현재 기간의 변동성에 얼마나 영향을 미치는지를 나타낸다. $P>|t|$ 의 값이 $6.472e-2$ 이므로 0에 가깝다. 따라서 통계적 유의성이 있다고 볼 수 있지만 신뢰구간이 음수 값($-6.107e-02$)에서 양수 값(2.061) 사이로 0을 포함할 가능성이 있어 유의미함을 갖지 않는다. 따라서 alpha[1]의 결과를 볼 때 해당 변동성은 통계적으로 유의미하지 않다고 해석할 수 있다. 즉, 과거 변화의 강도가 미래의 변동성에 크게 반응하지 않는다. 끝으로, beta[1]는 과거 기간의 변동성이 예측 기간의 변동성에 얼마나 지속적으로 영향을 미치는지를 나타내는 지표이다. beta[1]의 계수와 t-통계량의 값이 0이고, $P>|t|$ 의 값이 1을 가지므로 beta[1]는 통계적으로 유의하지 않다. 따라서 변동성이 지속하게 영향을 미치지 않는다고 해석할 수 있다. GARCH 모형을 통해 얻은 결과를 종합적으로 볼 때 접근 건수는 과거의 변화량은 현재의 변동성에 영향을 미치지만, 영향력 정도 즉, 강도와 지속성은 통계적으로 유의미하지 않다는 결론을 얻었다. 그림 3은 2025년에서 2026년까지의 접속 건수의 예측 변화량을 나타낸 것이다. 이 그림에서 보면

일정한 패턴으로 변동성이 증감하는 추이를 보였다. 이러한 변동성 추이 패턴은 계절형이나 이벤트와 같은 불규칙정보보다는 일정한 수준으로 증감하는 추세형 형태를 가졌다. 변동성(분산)의 폭이 큰 시계열 데이터를 예측하는데 유용한 GARCH 모형은 접속 건수 예측에 적합하지 않다고 판단된다. 결론적으로, ARIMA 모형과 GARCH 모형을 적용하기보다는 Prophet 모형만을 가지고서 짧은 기간의 접속 건수를 예측하는 것이 바람직하다고 판단된다.

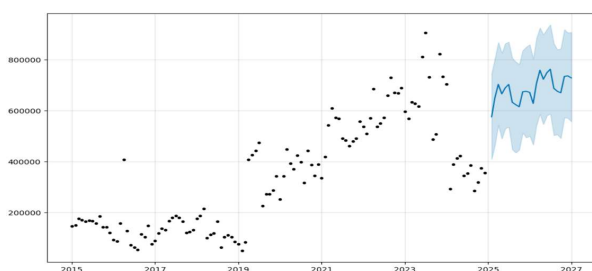


그림 3. 2025년과 2026년 간의 변동성 추세

IV. 결 언

본 연구는 건설기술정보시스템의 업무 성과 목표를 정확하게 설정하기 위해 2025년과 2026년 간의 접속 건수를 예측하고자 하였다. 이를 위해 2015년 1월부터 2024년 12월까지의 접속 건수를 수집하였고, 접속 건수가 갖는 시계열 특성을 분석하였다. 계량 경제학에서 비정상 추세형의 시계열 예측 모형으로 많이 사용하는 ARIMA 모형과 GARCH 모형, Prophet 모형을 적용하여 접속 건수의 예측치와 변동성을 추정하였다. ARIMA 모형과 Prophet 모형의 예측 결과를 비교하였을 때 Prophet 모형이 ARIMA 모형보다 예측치의 통계적 유의미성을 가지고 정확성에서도 우수하였다. 이러한 결과는 Satrioa 외[12]과 Atasever 외[15]에서 입증한 내용과 동일하였다. 하지만 GARCH 모형은 과거의 변동성이 현재의 변동성에 영향을 주지만 변동성의 강도와 지속성은 낮다는 결과를 얻었다. 따라서 짧은 기간의 접속

건수를 예측하는 데에는 ARIMA 모형과 GARCH 모형을 적용하기보다는 Prophet 모형을 활용하는 것이 좀 더 정확하게 예측할 수 있다는 결론을 얻었다. 이처럼 계량 경제학에서 주로 다루는 금융, 재화, 공급/수요, 에너지, 거시경제와 달리 접속 건수를 분석 대상으로 하여 ARIMA 모형, GARCH 모형, Prophet 모형을 적용하여 예측치와 변동성을 추정하고 비교하였다는 점에서 선행 연구와의 차별성이 있다고 사료된다. 끝으로, 본 연구는 접속 건수라는 하나의 변수를 사용하였고, 2년이라는 짧은 기간을 예측하였다. 하지만 대국민 서비스의 성과 목표로 다른 시계열 변수를 추가할 수 있다. 2개 이상의 시계열 변수를 사용할 때 변수 간의 예측치와 변동성 변화에 미치는 영향력은 달라질 수 있다. 따라서 2개 이상의 시계열 변수를 사용하였을 때 가장 적합한 시계열 예측 모형에 관한 연구가 필요하다. 특히, Tasi'u 외[18]에서 언급하였듯이 본 연구에서 사용한 접속 건수는 주가, 에너지, 농산물과 같이 외부요인이나 계절성에 크게 영향을 받지 않는 시계열 특성을 가졌다. 또한, 본 연구는 예측치와 예측치 사이에서만 비교하였다. 좀 더 현실적인 결과를 얻기 위해서는 예측 결과를 실측치와 비교하는 추가적인 연구가 필요하다. 본 연구 결과를 변동성의 강도가 크고 변동성이 지속적으로 미치는 시계열 변수에 적용하는 것은 바람직하지 않다.

REFERENCES

- [1] 행정안전부, “전자정부 성과관리 지침”, 행정안전부 고시 제2018-10호 제정, 2018년 2월 8일
- [2] 건설기술정보시스템, “<https://www.codil.or.kr>”. (accessed Aug., 15, 2025).
- [3] G. Box, G. Jenkins, “*Time Series Analysis: Forecasting and Control*”, Holden-Day, San Francisco, 1970.
- [4] Robert F. Engle, “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation”, *Econometrica*, vol. 50, no. 4, pp. 987-1007, 1982.
- [5] Tim Bollerslev, “Generalized autoregressive conditional

- heteroskedasticity”, *Journal of Econometrics*, vol. 31, np. 3, pp. 307-327, 1986.
- [6] Facebook Prophet, “<https://facebook.github.io/prophet>”. (accessed Aug., 01, 2025).
- [7] D. Soumith, M. Sravani, “A Forecasting Equity Prices Using ARIMA and GARCH Models: A Comparative Study”, *International Journal of Research Publication and Reviews*, vol. 6, no. 6, pp. 7056-7062, 2025.
- [8] 정용진, “순환신경망 기반 비선형 ARMA-GARCH 모형을 이용한 S&P500 지수 예측”, 서울대학교 대학원, 석사논문, 2019년 8월
- [9] 김시영, 황은주, “SARIMA-GARCH 모형을 활용한 일일 발전용 가스 공급량 예측 분석”, *한국데이터정보과학회*, 제25권, 제1호, 41-57쪽, 2025년
- [10] 이몽화, 김석태, “ARMA-GARCH 모형에 의한 중국 금 선물 시장 가격 변동에 대한 분석 및 예측”, *무역학회지*, 제47권, 제3호, 211-232쪽, 2022년
- [11] 이문석, “확장형 ARMA-GARCH 모형을 활용한 농업관측사업의 가격안정화 효과 분석”, 서울대학교 대학원, 석사논문, 2020년 2월
- [12] C.B.A. Satrioa, W. Darmawana, B.U. Nadiaa, N. Hanafiahb, “Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET”, *International Conference on Computer Science and Computational Intelligence*, Vol. 179, pp. 524-532, 2021.
- [13] K.K.R Samal, S.S .Babu, S.K. Das, A. Accharaya, “Time Series based Air Pollution Forecasting using SARIMA and Prophet Model”, *Proceedings of the 2019 International Conference on Information Technology and Computer Communication*, pp. 80-85, 2019.
- [14] S. Somisett, G.P. Manla, “Multi-Model Ensemble Approach for Stock Price Forecasting: An Interactive Dashboard Implementation Using ARIMA, Prophet, and LSTM Models with Advanced Trading Strategy Integration”, *TechRxiv*, 2025.
- [15] S. Atasever, B. Öztürk, G. Bilgiç, “A new approach to short-term wind speed prediction: the prophet mode”, *Energy Sources, Part A: Recovery, Utilization, And Environmental Effects*, vol. 44, no. 4, pp.8826 - 8841, 2022.
- [16] 오승원, 임남희, 이상현, 김민수, “Prophet 모형을 이용한 마늘 가격의 장기 예측 및 트렌드 분석”, *한국자료분석학회 논문지*, 제22권, 제6호, 2325-2336 쪽, 2020년
- [17] S. Sari-Hassoun, N. R Boudjourfa, M. Mekidiche, “Forecasting the Weekly Spot Oil Price Using a Hybrid Model ARMA-GARCH-MLP and Prophet Forecasting Method”, *Economic Computation and Economic Cybernetics Studies and Research*, vpl. 59, no. 1, pp. 156-173, 2025.
- [18] M. Tasi’u, A.A. Umar, U. Abdulaziz, “Does Holidays affect the Returns of Stock Markets Prices? An Application of some GARCH and Prophet Models under Varying Error Distributions”, *International Journal Of Science For Global Sustainability*, vol. 10, no. 4, pp. 60-67, 2024.

 저 자 소 개



정성원(정회원)

1992년 2월 : 한양대학교 전자계산학과 (공학학사)

1994년 2월 : 숭실대학교 컴퓨터공학과 (공학석사)

2018년 2월 : 서울과학기술대학교 산업정보시스템전공 (공학박사)

• 1994년 3월 ~ 현재 : 한국건설기술연구원 미래 스마트건설연구본부 연구위원.

<주관심분야 : 건설정보화, DB, 인공지능, 투자공학>