



## 저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

이학석사학위논문

Prophet을 이용한 계층적 시계열 예측

Hierarchical Time Series Forecasting  
with Prophet



충북대학교 대학원

통계학과 정보통계학전공

방태모

2021년 2월

이학석사학위논문

Prophet을 이용한 계층적 시계열 예측

Hierarchical Time Series Forecasting  
with Prophet

지도교수 나 종 화

통계학과 정보통계학전공

방 태 모

이 논문을 이학석사학위 논문으로 제출함.

2021 년 2 월

본 논문을 방태모의 이학석사학위 논문으로 인정함.

심 사 위 원 장      김 정 연 ⑨

심 사 위 원      정 재 환 ⑨

심 사 위 원      나 종 화 ⑨

충 북 대 학 교 대 학 원

2021 년 2 월

# 차 례

Abstract .....	iii
표 차 례 .....	iv
그림 차례 .....	v
I. 서론 .....	1
II. 방법론 .....	3
1. 시계열 예측과 Prophet 모형 .....	3
(1) 지수평활법 .....	3
(2) ARIMA 모형 .....	4
(3) Prophet 모형 .....	4
2. 계층적 시계열 예측 .....	6
(1) 상향식 방법 .....	8
(2) 하향식 방법 .....	9
(3) 미들 아웃 방법 .....	10
(4) 최적 조합 방법 .....	10
III. 실증 분석 .....	13
1. 자료 소개 및 Prophet 모형 설정 과정 .....	13
(1) 자료 소개 .....	13
(2) Prophet 모형 설정 과정 .....	15
2. 모형 평가 및 비교 .....	16
(1) 예측 비교 방법 .....	16
(2) 중·단기 예측 .....	17

(3) 장기예측 .....	20
3. 모형 예측 성능 평가 .....	21
<b>IV. 결론</b> .....	24
참고 문헌 .....	25

# Hierarchical Time Series Forecasting with Prophet\*

*Bang, Tae Mo*

*Department of Statistics*

*Graduate School, Chungbuk National University*

*Cheongju, Korea*

*Supervised by Professor Na, Jong Hwa*

## Abstract

In this paper, we mainly introduce prophet and hierarchical time series forecasting. We consider four hierarchical time series forecasting approaches: bottom-up, top-down, middle-out, and optimal combination. These methods allow the predicted values to satisfy the hierarchical structure and could help to improve forecast performance. We apply the methods to the prophet. Besides, it is applied to the exponential smoothing and the ARIMA model as a comparison point. the purpose of this paper is to show that the hierarchical time series forecasting helps to improve forecast performance and that prophet has superior forecast performance compared to the other two models. We forecast the daily number of passengers using the wide-area railway, which is disaggregated by routes and stations. In conclusion, we observe that the prophet model performs much better than the other two models, and applying hierarchical time series forecasting to each model shows more accurate forecasting performance.

Keywords: prophet, hierarchical time series forecasting, top-down approach, bottom-up approach, middle-out approach, optimal combination approach

---

\* A thesis for the degree of Master in February 2021.

## 표 차 례

<표 2.1> 지수평활법의 분류 체계 .....	3
<표 2.2> 휴일효과 목록의 예 .....	5
<표 3.1> 분석에 사용된 자료의 계층 구조 .....	13
<표 3.2> 모형 설정에 고려한 휴일 효과 .....	15
<표 3.3> 모형별 $MAPE(k)$ .....	18



## 그 립 차 례

<그림 2.1> 2-수준 계층 구조의 트리 다이어그램 .....	6
<그림 3.1> 노드별 시도표 .....	14
<그림 3.2> 중·단기예측의 롤링 창 기법 .....	16
<그림 3.3> 상향식 방법을 적용한 모형별 $MAPE_h(k)$ 의 시도표 .....	19
<그림 3.4> 모형별 $MAPE_h(k)$ 의 상자그림 .....	20
<그림 3.5> 상향식 방법을 적용한 모형별 $MAPE_h(\cdot)$ 의 시도표 .....	21
<그림 3.6> 자료 분할 .....	21
<그림 3.7> 상향식 방법을 적용한 모형별 총 계열의 시도표 .....	22
<그림 3.8> 상향식 방법을 적용한 모형별 관측값과 예측값의 산점도 .....	23

## I. 서론

시계열은 종종 본질적으로 계층적(hierarchical) 또는 그룹화(grouped) 구조로 분해되는 성질을 갖는다. 예를 들어, 우리나라의 월별 관광 수요 자료는 도별로 분해 후, 여행 목적별로 분해될 수 있으며, 그 반대 순서로도 분해할 수 있다. 다른 예로, 지역별로 여러 지점이 존재하는 의류 브랜드의 분기별 판매량은 도별, 시군별, 지점별로 분해될 수 있다. 전자는 그룹화 시계열로 분해에 고유한 순서가 정해져 있지 않으며, 후자는 계층적 구조가 고유하게 존재하는 계층적 시계열에 해당한다. 그룹화 또는 계층적 시계열 자료의 예측에 단일 시계열 예측모형을 이용할 경우, 예측값들은 계층 구조를 만족하지 않는 문제가 발생한다. 예를 들어, 우리나라의 월별 총 관광 수요와 도별 관광 수요를 각각 지수평활법으로 예측한다고 했을 때, 도별 관광 수요의 예측값의 합이 총 관광 수요의 예측값과 일치하지 않는 문제가 발생할 것이다. 이때 지수평활법으로 계산된 예측값들에 계층적 시계열 예측 방법을 적용할 경우, 예측값들은 계층 구조를 만족하게 되며, 예측 성능의 개선에도 도움을 줄 수 있다.

Athanasopoulos 등(2009)은 지수평활법을 이용하여 계층적 시계열 예측을 호주의 분기별 관광 수요자료에 적용하였으며, Lee와 Seong(2017)은 지수평활법과 ARIMA 모형을 이용하여 계층적 시계열 예측을 우리나라 월별 교통사고 건수 자료에 적용하였다. 본 논문에서는 2018년 Facebook에서 제안한 곡선 적합(curve-fitting) 방식의 시계열 예측모형 Prophet을 이용하여 계층적 시계열 예측을 월별 광역철도 수송량 자료에 적용하였으며, 이와 비교 점으로 지수평활법과 ARIMA 모형을 고려하였다. Prophet 모형은 지수평활법, ARIMA 모형과 달리 쉽게 조정이 가능하면서 기본 설정으로도 뛰어난 예측 성능을 보이며, 결측치와 이상점에 로버스트하고 강한 계절성을 가지는 자료에도 잘 작동한다. 또한, 철도 수송량과 같은 자료는 시계열에 큰 충격(shocks)을 줄 수 있는 휴일효과(holidays and events effects)에 민감할 수 있으며, Prophet 모형에서는 시계열 예측에 전문적인 지식이 없는 비전문가도 휴일효과를 쉽게 반영할 수 있다. 정리하면, 본 논문에서는 계층적 시계열 예측 방법의 적용에 Prophet 모형, 지수평활법, ARIMA 모형을 고려하였으며, 이를 통해 예측 성능의 개선을 이루고 Prophet 모형이 지수평활법과 ARIMA 모형보다 우수한 예측 성능을 가짐을 보일 것이다.

본 논문의 구성은 총 4장으로 이루어져 있다. 먼저 2장에서는 지수평활법과 ARIMA 모형을 간략히 소개하고, Prophet 모형을 중점적으로 소개한다. 그리고, 계층적 시계열을 다룰 때 필요한 기본적인 개념과 계층적 시계열 예측 방법을 소개한다.

4가지 계층적 시계열 예측 방법을 이용하였다. 먼저 가장 흔하고 단순한 방법인 상향식 방법(bottom-up approach)과 Athanasopoulos 등(2009)이 제안한 새로운 하향식 방법(top-down approach)을 소개한다. 그리고, 두 방식을 결합한 미들 아웃 방법(middle-out approach)과 Hyndman 등(2007)이 제안한 최적 조합 방법(optimal combination approach)을 소개한다. 3장에서는 실증 분석에 관한 내용을 정리하였다. 먼저, 본 논문에서 이용한 자료와 Prophet 모형의 설정 과정을 자세하게 소개하였으며, 롤링 창(rolling window) 기법을 이용해 *MAPE*를 기반으로 모형 평가 및 비교를 수행한다. 마지막으로 자료를 훈련자료(training data), 시험자료(test data)로 나누어 원자료와 예측값을 기반으로 모형 예측 성능평가를 수행하였다. 4장에서는 연구 결론을 기술하였다.

## II. 방법론

### 1. 시계열 예측과 Prophet 모형

본 논문에서 계층적 시계열 예측의 적용에 고려한 모형은 지수평활법, ARIMA 모형, Prophet 모형이다. 지수평활법과 ARIMA 모형을 간략하게 소개하고, Prophet 모형을 중점적으로 소개한다.

#### (1) 지수평활법

지수평활법은 이전 시점의 과거 관측값에 지수적으로 감소하는 가중치를 곱하여 시계열 예측을 수행하는 모형이다. 예를 들어, 그중 가장 단순한 방법은 추세와 계절성을 고려하지 않는 단순 지수평활 기법이다. 표 2.1은 Pegels(1969)가 제안하고 Gardner(1985), Hyndman(2002)와 Taylor(2003)에 의해 발전된 지수평활 기법의 기본적인 분류 체계이다.

<표 2.1> 지수평활법의 분류 체계

추세 성분	계절 성분		
	N	A (가법)	M (승법)
N	N, N	N, A	N, M
A (가법)	A, N	A, A	A, M
$A_d$ (가법 감쇠)	$A_d, N$	$A_d, A$	$A_d, M$
M (승법)	M, N	M, A	M, M
$M_d$ (승법 감쇠)	$M_d, N$	$M_d, A$	$M_d, M$

(Athanasopoulos 등, 2009)

표 2.1에서 제시되는 15종류의 모형 각각은 추세 성분과 계절 성분을 가지며, 추세에서는 가법과 승법뿐만 아니라 감쇠 기법까지 고려하였다. 단, N은 해당 성분을 고려하지 않은 경우를 말한다. 예를 들어, (N, N)은 단순 지수 평활 기법이며, (A, A)는 홀트-윈터스 가법 모형이다. ( $A_d, A$ )는 홀트-윈터스 가법 감쇠 모형으로, 추세를 일정하게 고려하지 않고 감쇠 매개변수  $\phi$ 를 이용해 추세를 감소시켜 미래 어떤 시점에는 추세가 상수가 되도록 하는 모형이다.

## (2) ARIMA 모형

ARIMA 모형은 대표적인 통계적 시계열 예측모형으로, 현재값을 과거값과 과거 예측 오차를 통해 설명한다. 본 논문에서는 식 (2.1)과 같은 계절 ARIMA 모형  $ARIMA(p, d, q)(P, D, Q)_s$ 를 고려하였다.

$$\phi_p(B)\Phi_P(B^s)(1-B^s)^D(1-B)^d y_t = \delta + \theta_q(B)\Theta_Q(B^s)\epsilon_t, \quad (2.1)$$

여기서  $\delta$ 는 상수,  $B$ 는 후향 연산자(backshift operator),  $s$ 는 계절 주기,  $\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ ,  $\theta_q(B) = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q)$ ,  $\Phi_P(B^s) = (1 - \Phi_1 B^s - \Phi_2 B^{2s} - \dots - \Phi_P B^{Ps})$ ,  $\Theta_Q(B^s) = (1 + \Theta_1 B^s + \Theta_2 B^{2s} + \dots + \Theta_P B^{Qs})$ 를 나타낸다. 예를 들어, 후향 연산자를 이용하면  $y_{t-b}$ 는  $B^b y_t$ 로 나타낼 수 있다.

## (3) Prophet 모형

Prophet 모형은 2018년 Facebook에서 제안한 시계열 예측모형으로, 시계열의 시간 종속적인 특성을 고려하는 기존의 시계열 모형들과는 달리 곡선 적합으로 모형을 적합한다. 이를 통해 Prophet은 많은 장점을 갖는다. 예를 들면, 결측치와 이상점에 로버스트하고 강한 계절성을 가지는 자료에도 잘 작동한다. Prophet 모형은 식 (2.2)와 같이 시계열을 세 가지 요소로 분해(Harvey와 Peters, 1990)하고 각 요소를 시간의 함수로 가법적으로 모형화한다. 이는 일반화가법모형(Hastie와 Tibshirani, 1987)과 유사하다.

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t, \quad (2.2)$$

여기서  $g(t)$ 는 시계열 값의 비-주기적 변화를 모형화하는 추세 함수,  $s(t)$ 는 주별 또는 연별 계절성과 같은 주기적 변화를 반영하는 계절 함수,  $h(t)$ 는 불규칙하게 발생할 가능성이 있는 휴일효과를 모형화한다. 각 함수를 좀 더 자세하게 설명하면 다음과 같다.

추세 함수  $g(t)$ 는 데이터의 추세를 선형 또는 비선형으로 모형화한다. 기본적으로 Prophet의 추세 모형화는 변화점(changepoints)을 반영하며, 자동으로 선택될 수도 있고 분석자에 의해 명시될 수도 있다.

계절 함수  $s(t)$ 는 푸리에 급수(fourier series)를 통해 임의의 계절효과를 추정한다.

$$s(t) = \sum_{n=1}^N (a_n \cos(\frac{2\pi nt}{P}) + b_n \sin(\frac{2\pi nt}{P})), \quad (2.3)$$

여기서  $P$ 는 주기를 나타낸다. 연도별 자료면  $P = 365.25$ , 주별 자료면  $P = 7$ 이 될 것이다. 계절성의 적합에는  $2N$ 개의 모수 추정이 필요로 되는데,  $N$ 이 커질수록 계절 패턴이 더 빠르게 변화할 수 있도록 허용하며, 이는 과적합(overfitting)을 일으킬 수 있다. 연도별, 주별 계절성의 경우 보통 각각  $N = 10$ ,  $N = 3$ 로 설정하면, 대부분 문제에서 잘 작동한다. 이러한 모수들의 선택은 AIC와 같은 측도를 사용한 모형 선택 과정을 통해 자동으로 이루어진다(Taylor와 Letham, 2018).

마지막으로 휴일효과를 모형화하는  $h(t)$ 는 Prophet 모형이 갖는 큰 장점 중 하나라고 할 수 있다. 휴일효과는 많은 경제 시계열에 큰 충격 효과를 주며, 주기적인 패턴을 따르지 않는 경우가 많아 모형화되기 어렵다. 예를 들어, 미국의 추수감사절은 매년 11월 넷째 주 목요일로 주기적으로 발생하지만, 슈퍼볼과 같은 이벤트는 1월 또는 2월의 일요일에 불규칙적으로 발생하므로 모형화하기 어렵다. 또한, 우리나라를 포함한 여러 나라의 중요한 휴일은 음력을 따르는 문제도 존재하며, 이렇게 불규칙적 패턴으로 나타나는 효과들을 시계열 예측에 반영하는 것은 매우 중요하다. Prophet 모형에서는 이러한 문제를 해결할 수 있도록, 분석자가 과거에 발생한 또는 미래에 발생할 사건들의 목록을 만들어 이를 모형에 반영할 수 있도록 한다. 다음의 표 2.2는 미국의 휴일효과를 작성한 예이다.

<표 2.2> 휴일효과 목록의 예

휴일	국가	연도	날짜
추석	미국	2015	2015-11-26
추석	미국	2016	2016-11-24
추석	미국	2017	2017-11-23
추석	미국	2018	2018-11-22
크리스마스	*	2015	2015-12-25
크리스마스	*	2016	2016-12-25
크리스마스	*	2017	2017-12-25
크리스마스	*	2018	2018-12-25

(Taylor와 Letham, 2018)

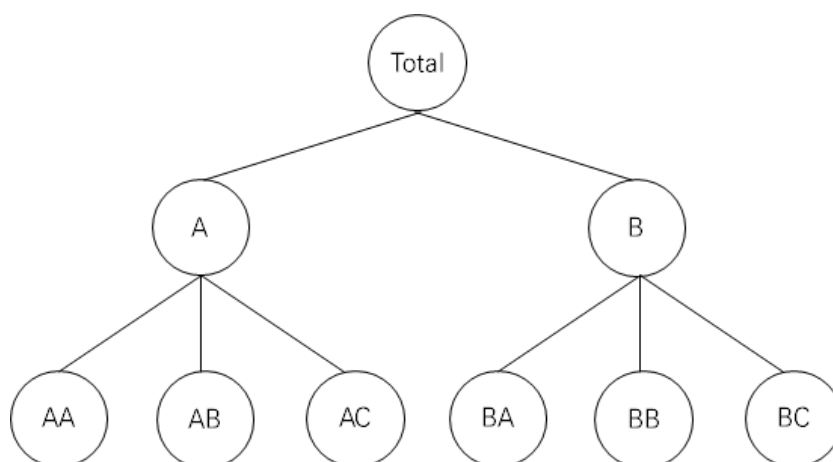
표 2.2와 같이 작성되는 휴일효과의 목록은 분석자가 자유롭게 조정하여 모형 적합

단계에서 설정할 수 있으며, 휴일효과는 독립적이라는 가정하에 반영된다. 휴일효과에서 추가로 고려할만한 중요한 사항은 우리나라의 추석과 같은 특정 휴일의 경우 해당 날짜 전후에 효과를 포함해야 한다는 점이다. 이를 휴일의 창(a window of days)이라고 하며, Prophet 모형에서는 이러한 창을 고려할 수 있도록 하는 추가적인 모수가 존재한다.

## 2. 계층적 시계열 예측

계층적 시계열이란, 계층 구조가 고유하게 존재하는 시계열을 말한다. 예를 들어, 광역철도의 수송량은 노선별로 구분되며, 이는 다시 역별로 분해된다. 다음의 그림 2.1은 2-수준 계층 구조를 가지는 계층적 시계열을 도식화하여 표현한 것이다.

<그림 2.1> 2-수준 계층 구조의 트리 다이어그램



계층적 시계열의 각 계열은 노드라고 표현되기도 한다. 그림 2.1의 최상단에 있는 완전히 합산된 계열을 총 계열(total series)이라 하며, 해당 층을 최상위 수준 또는 수준 0이라 한다. 그리고, 이를 분해한 노드  $A, B$ 가 존재하는 층을 수준 1이라 한다. 계층적 시계열은 아래로 내려가면서 이러한 분해가 반복되는 구조를 취하며, 완전히 분해된 계열들로 구성되는 최하위 수준(bottom level)을 통상적으로 수준  $K$ 로 나타낸다. 수준  $K$ 로 구성되는 계층적 시계열을  $K$ -수준 계층 구조라고 표현하며, 이러한 이유로 그림 2.1과 같은 구조를 2-수준 계층 구조라고 한다.

다음으로 계층적 시계열의 표기법을 소개한다. 먼저,  $Y_{X,t}$ 는 노드  $X$ 의  $t$  번째 관측치( $t = 1, \dots, n$ )를 나타낸다. 예를 들어, 그림 2.1의 노드  $A$ 의  $t$  번째 관측치는  $Y_{A,t}$

라 표현할 수 있다. 보통 개별 노드를 나타내기 위해 영어 문자열을 사용하는데, 총 계열은 예외로  $Y_t$  라 표현한다. 또한, 수준별 노드 개수는  $m_i$  로 나타내며, 이는 수준  $i$  의 노드 개수를 의미한다. 예를 들면, 그림 2.1의 수준 1에는 2개의 노드가 존재하므로  $m_1 = 2$  라 할 수 있다. 그리고,  $m$  을 노드의 총 개수라 하면  $m = m_0 + m_1 + \dots + m_K$  로 정의할 수 있다.

본 논문에서 소개할 계층적 시계열 예측 방법들의 이해를 위해 앞서 배운 표현들을 벡터로 확장하면,  $t$  시점 모든 노드의 관측치를 요소로 갖는 벡터  $Y_t$  는 다음과 같이 쓸 수 있다.

$$Y_t = [Y_t, Y'_{1,t}, \dots, Y'_{K,t}]', \quad (2.3)$$

여기서  $Y_t$  는 총 계열의  $t$  시점 관측치이며,  $Y_{1,t}$  는 수준 1의 모든 노드의  $t$  시점 관측치를 요소로 갖는 벡터,  $Y_{K,t}$  는 최하위 수준의 모든 노드의  $t$  시점 관측치를 요소로 갖는 벡터이다.  $Y_t$  는 합산 행렬(summing matrix)  $S$  를 이용하면 식 (2.5)와 같이 최하위 수준 노드들의 합으로 표현할 수 있다.

$$Y_t = SY_{K,t}, \quad (2.5)$$

여기서  $S$  는  $m \times m_k$  합산 행렬이다. 식 (2.5)를 이용하면 그림 2.1의 계층 구조는 다음과 같이 쓸 수 있다.

$$\begin{bmatrix} Y_t \\ Y_{A,t} \\ Y_{B,t} \\ Y_{AA,t} \\ Y_{AB,t} \\ Y_{AC,t} \\ Y_{BA,t} \\ Y_{BB,t} \\ Y_{BC,t} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Y_{AA,t} \\ Y_{AB,t} \\ Y_{AC,t} \end{bmatrix}.$$

즉, 합산 행렬  $S$  는 계층적 시계열이 갖는 계층 구조에 대한 정보를 담고 있다고 할 수 있다. 계층적 시계열에서 결국 우리가 관심이 있는 것은 계층 구조를 반영한 예측



이므로, 예측값을 표기하는 방식에 대해서도 알아보자.  $\hat{Y}_{X,n}(h)$ 는  $n$  개의 관측치를 기반으로 얻은 노드  $X$ 의  $h$  시점 이후 예측값을 의미한다. 이를 벡터를 사용하여 수준  $i$ 로 확장한 것이  $\hat{Y}_{i,n}(h)$ 이며, 모든 노드로 확장할 경우  $\hat{Y}_n(h)$ 가 된다. 계층적 시계열 분석에서는  $\hat{Y}$ 을 기저 예측값(base forecasts)이라 표현한다. 이는 모든 계층적 시계열 예측 방법의 초기에 필요로 되는 값으로, 단일 시계열 예측모형을 이용한 각 노드의 독립적인 예측을 통해 구해진다. 즉, 기저 예측값은 계층 구조가 반영되지 않은 값이라 할 수 있다. 이러한 기저 예측값에 계층적 시계열 예측 방법을 적용한 것을 최종 수정된 예측값(the final revised forecasts)이라 하며,  $\tilde{Y}$ 로 표기한다. 즉, 최종 수정된 예측값은 계층 구조를 만족하며, 예측 성능의 개선에도 도움을 줄 수 있다.  $\tilde{Y}$ 의 표기법은 앞서 배운 기저 예측값의 표기법을 그대로 따른다. 예를 들어,  $n$  개의 관측치를 기반으로 얻은  $h$  시점 이후 모든 노드의 최종 수정된 예측값을 요소로 갖는 벡터는 기호로  $\tilde{Y}_n(h)$ 로 표기한다.

계층적 시계열 예측 방법을 통해 구해지는  $\tilde{Y}_n(h)$ 은 다음과 같이 표현할 수 있다.

$$\tilde{Y}_n(h) = SP\hat{Y}_n(h), \quad (2.6)$$

여기서  $S$ 는  $m \times m_k$  합산 행렬,  $P$ 는  $m_k \times m$  행렬이며,  $\hat{Y}_n(h)$ 는  $n$  개의 관측치를 기반으로 얻은  $h$  시점 이후 모든 노드의 기저 예측값을 요소로 갖는 벡터이다.  $P$ 는 어떤 계층적 시계열 예측 방법을 사용하느냐에 따라 달라진다. 식 (2.6)을 통해 우리는 최종 수정된 예측값이 기저 예측값의 선형 결합으로 표현될 수 있음을 알 수 있다. 이제 본 논문에 사용한 4가지 계층적 시계열 예측 방법을 소개한다.

### (1) 상향식 방법

상향식 방법은 가장 단순하고 흔히 적용되는 방식 중 하나로, 최하위 수준 기저 예측값들의 합산을 통해 상위 수준의 최종 수정된 예측값들을 생성한다. 상향식 방법에서  $P$ 는 다음과 같이 표현된다.

$$P = [0_{m_k \times (m - m_k)} | I_{m_k}], \quad (2.7)$$

여기서  $P$ 의 역할은 최하위 수준 노드들의 기저 예측값을 추출하는 것이며, 합산 행렬  $S$ 를 이용해 상위 수준 노드들의 최종 수정된 예측값을 생성한다. 즉, 상향식 방법

에서 최하위 수준 노드들의 최종 수정된 예측값은 기저 예측값과 같다. 상향식 방법의 가장 큰 이점은 시계열이 최대로 분해된 지점으로부터 모형화함으로써, 합산으로 인해 발생할 수 있는 정보 손실을 방지할 수 있다는 것이다. 이러한 이유로, 상향식 방법은 각 노드가 가지는 개개의 특성을 잘 잡아낼 수 있다. 반대로, 최하위 수준을 구성하는 노드들에 잡음(noise)이 많은 경우 오히려 좋지 않은 예측 결과를 초래할 수 있다.

## (2) 하향식 방법

하향식 방법도 단순하고 흔히 적용되는 방식이며, 총 계열의 기저 예측값을 일정 비율로 최하위 수준의 각 노드에 분배하고 합산 행렬  $\mathbf{S}$ 를 이용해 상위 수준으로 합산하여 최종 수정된 예측값을 생성한다. 즉, 하향식 방법에서는 상향식 방법과 반대로 총 계열의 최종 수정된 예측값은 기저 예측값과 같다. 하향식 방법에서  $\mathbf{P}$ 는 다음과 같이 주어진다.

$$\mathbf{P} = [\mathbf{p} | \mathbf{0}_{m_k \times (m-1)}], \quad (2.8)$$

여기서  $\mathbf{P}$ 의 역할은 총 계열의 기저 예측값을 일정 비율로 최하위 수준의 노드로 분배하는 것이다.  $\mathbf{p} = [p_1, p_2, \dots, p_{m_k}]'$ 는 총 계열의 기저 예측값의 분배 비율의 집합에 해당하며,  $\mathbf{p}$ 를 계산하는 방식에 따라서 하향식 방법은 크게 3가지로 나뉜다. Gross와 Sohl(1990)이 제안한 과거 비율의 평균(the average historical proportions), 과거 평균의 비율(proportions of the historical averages)을 이용하는 것이 있으며, 본 논문에서는 Athanasopoulos 등(2009)이 제안하고 훌륭한 성능을 보여준 예측 비율(forecasted proportions)을 이용한 하향식 방법을 사용하였다.

예측 비율을 이용한 하향식 방법에서  $\mathbf{p}$ 의 원소들이 계산되는 과정을 간략히 설명한다. 설명의 편의를 위해 최상위 수준과 최하위 수준만 존재하는 1-수준 계층 구조를 갖는 계층적 시계열 자료라 가정하자. 예측 비율을 이용한 하향식 방법은 기본적으로 모든 노드의 기저 예측값이 필요하다. 이를 초기 예측값이라 한다. 그리고 최하위 수준 노드들의 초기 예측값의 합 대비 최하위 수준의 각 노드의 초기 예측값이 갖는 비율을 계산한다. 최하위 수준 노드별로 계산된 각 비율들을 우리는 예측 비율이라 부르며, 총 계열의 초기 예측값을 최하위 수준의 노드들로 분배하기 위해 사용한다. 이러한 과정의 일반화를 위해  $K$ -수준 계층 구조를 고려하면, 상술한 과정이 최상위 수준부터 최하위 수준까지 반복된다. 즉,  $K$ -수준 계층 구조에서  $\mathbf{p}$ 의 원소가 계산되는 방식은 다음과 같이 표현할 수 있다.

$$p_j = \prod_{l=0}^{K-1} \frac{\hat{Y}_{j,n}^{(l)}(h)}{\hat{S}_{j,n}^{(l+1)}(h)} \text{ for } j = 1, 2, \dots, m_k, \quad (2.9)$$

여기서  $(l)$ 은 노드  $j$ 의 상위  $l$ 수준,  $\hat{S}_{j,n}^{(l+1)}(h)$ 는 노드  $j$ 의 상위  $(l+1)$ 수준과 연결된 하위 수준 모든 노드의 기저 예측값들의 합,  $\hat{Y}_{j,n}^{(l)}(h)$ 은 노드  $j$ 의 상위  $(l)$ 수준 노드의 기저 예측값을 의미한다. 예를 들어 그림 2.1의 총 계열에 대한 기저 예측값은 다음과 같이 쓸 수 있다.

$$\hat{S}_{AA,n}^{(2)}(h) = \hat{S}_{\text{Total},n} = \hat{Y}_{A,n}(h) + \hat{Y}_{B,n}(h), \quad (2.10)$$

Gross와 Sohl(1990)이 제안한 과거 비율에 근거한 두 방식은 예측 비율에 근거한 방식과 달리 총 계열의 예측만이 필요로 되므로 단순하고 빠르다는 장점이 있었으나, 비율이 시간에 따라 변할 수도 있다는 점을 고려하지 못하여 하위 수준으로 갈수록 예측 성능이 떨어지는 경향이 있다. 이러한 문제는 최상위 수준의 기저 예측값을 분배하는 비율로 과거 데이터가 아닌 예측 비율을 이용함으로써 개선될 수 있었다(Athanapoulos 등, 2009). 그러나, 하향식 방법과 이전에 소개한 상향식 방법이 갖는 공통적인 단점은 기저 예측값이 비편향성을 만족해도 최종 수정된 예측값은 비편향성을 만족하지 못한다는 것이다(Hyndman 등, 2011).

### (3) 미들 아웃 방법

미들 아웃 방법은 상향식 방법과 예측 비율을 이용한 하향식 방법을 결합한 것이다. 먼저, 중간 수준(middle level)을 고르고 해당 수준의 모든 노드에 대해 기저 예측값을 생성한다. 그리고, 중간 수준 위로는 상향식 방법을 통해 예측을 수행하고, 아래로는 예측 비율을 이용한 하향식 방법을 사용한다. 즉, 미들 아웃 방법의 경우 중간 수준의 최종 수정된 예측값은 기저 예측값과 같다.

### (4) 최적 조합 방법

최적 조합 방법은 Hyndman 등(2007)이 제안한 계층적 시계열 예측 방법으로, 계층 구조에 관한 정보를 가진 합산 행렬을 예측 변수로 최하위 수준의 기저 예측값들을 회귀적으로 추정하고, 상위 수준의 기저 예측값들을 생성하는 방법이다. 즉, 이 방법의 기본적인 아이디어는 모든 노드의  $h$ 시점 이후 기저 예측값을 선형 회귀 방정식의

로 표현하는 것에 있다. 최적 조합 방법에서  $\hat{Y}_n(h)$ 는 식 (2.11)과 같이 표현할 수 있다.

$$\hat{Y}_n(h) = S\beta_n(h) + \epsilon_h, \quad (2.11)$$

여기서  $\beta_n(h) = E[\hat{Y}_{K,n}(h) | Y_1, \dots, Y_n]$ 로 최하위 수준  $K$ 의 모든 노드의 알려지지 않은 조건부 기댓값이며,  $\epsilon_h$ 는 선형 회귀에 대한 오차로 평균으로 0을 갖고 공분산 행렬로  $\text{Var}[\epsilon_h] = \sum_h$ 를 갖는다. 즉, 우리가  $\sum_h$ 를 알고있다면 다음의 식 (2.12)와 같이 일반화최소제곱 추정을 통해  $\beta_n(h)$ 의 최소분산 불편추정량을 얻을 수 있다.

$$\hat{\beta}_n(h) = (S' \sum_h^+ S)^{-1} S' \sum_h^+ \hat{Y}_n(h), \quad (2.12)$$

여기서  $\sum_h^+$ 는  $\sum_h$ 의 무어-펜로즈 일반화 역행렬(the Moore-Penrose generalized inverse)이다. 그러나, 일반적으로 실제 문제에서  $\sum_h$ 에 대한 정보가 존재하기는 어렵다. 이러한 추정의 어려움을 Hyndman 등(2011)은 식 (2.13)과 같은 계층 구조 모든 시계열의  $h$ 시점 이후 기저 예측이 만들어 내는 오차가 근사적으로 계층적 구조를 만족한다는 가정을 통해 해결했다.

$$\epsilon_h \approx S \epsilon_{K,h}, \quad (2.13)$$

여기서,  $\epsilon_{K,h}$ 는 선형 회귀 방정식 식 (2.11)에서 최하위 수준 노드들의 기저 예측값이 만들어 내는 오차이다. 이 가정하에서는  $\sum_h$ 에 의존하는  $\beta_n(h)$ 의 일반화최소제곱 추정은 다음의 식과 같이 매우 간단해진다.

$$\hat{\beta}_n(h) = (S' S)^{-1} S' \hat{Y}_n(h). \quad (2.14)$$

식 (2.14)은  $\beta_n(h)$ 의 최량선형불편추정량(best linear unbiased estimator; BLUE)에 해당한다. 해당 결과는 오차가 근사적으로 계층적 구조를 만족해야 한다는 가정 때문에 항상 정확하진 않으나, 적절한 기저 예측값들의 집합에 대해서는 근사적으로 정확한 값을 추정할 것이다. 최적 조합 방법의 최종 수정된 예측값은 다음의 식과 같이 표현되며, 이는 식 (2.11)에 식 (2.14)을 대입하면 얻어진다.

$$\tilde{\mathbf{Y}}_n(h) = \mathbf{S}\hat{\boldsymbol{\beta}}_n(h) = \mathbf{S}(\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'\hat{\mathbf{Y}}_n(h) = \mathbf{SP}\hat{\mathbf{Y}}_n(h), \quad (2.15)$$

여기서  $\mathbf{P} = (\mathbf{S}'\mathbf{S})^{-1}\mathbf{S}'$  이다. 최적 조합 방법에서  $\mathbf{SP}$ 는 선형 회귀의 해트 행렬(hat matrix)과 같은 역할을 하며, 이때  $\mathbf{SP}$ 를 기저 예측에 곱해지는 가중치 개념으로 보기도 한다(Hyndman 등, 2011).

이 방법은  $\hat{\mathbf{Y}}_n(h)$ 의 추정에 회귀 방정식을 이용하므로 각 수준의 계열 간의 상관과 상호작용을 고려할 수 있다. 또한, 앞서 소개한 계층적 시계열 방법들과 달리 기저 예측값이 불편성을 만족할 경우, 최적 조합 방법을 통해 구해지는 최종 수정된 예측값도 불편성을 만족하게 된다는 장점이 있다.

### Ⅲ. 실증 분석

이번 장에서는 Prophet 모형, 지수평활법, ARIMA 모형에 2장에서 설명했던 4가지 계층적 시계열 예측 방법을 적용하고, 롤링 창 기법을 이용해 모형 평가 및 비교를 중·단기예측, 장기예측 관점에서 수행한다.

#### 1. 자료 소개 및 Prophet 모형 설정 과정

##### (1) 자료 소개

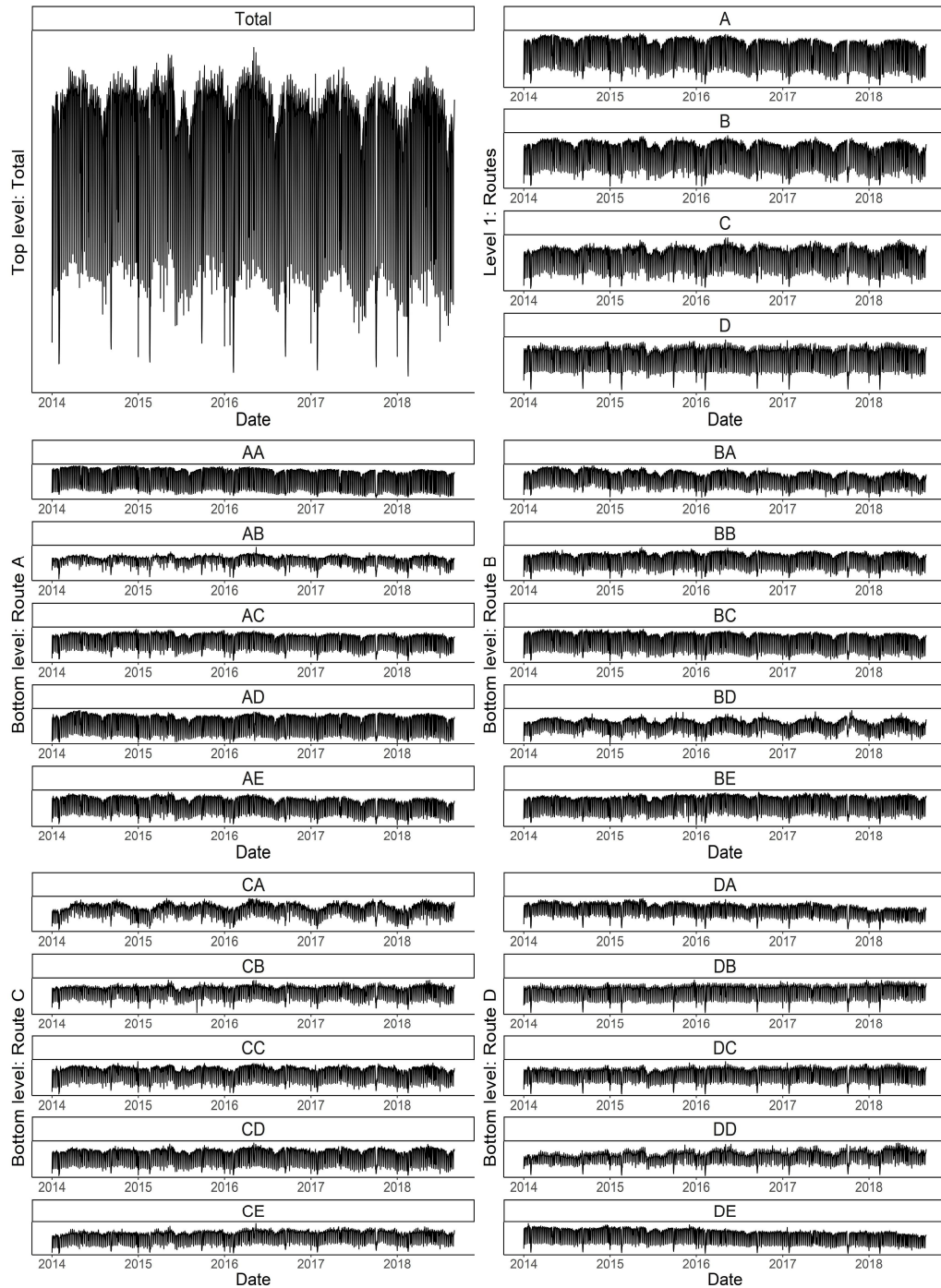
본 논문에서는 2014년 1월 1일부터 2018년 8월 31일까지 기록된 노선별, 역별로 구분된 광역철도 일별 수송량 자료를 이용하였다. 실증 분석에서 광역철도의 모든 노선을 이용하진 않았으며, 비교적 역의 수가 많은 4개의 노선을 택하고 그중 이상점이 존재하지 않는 안정적인 패턴을 띠는 역을 위주로 5개씩 선택했다. 다음의 표 3.1은 실증분석에 사용된 자료의 계층 구조를 나타낸 것이다.

<표 3.1> 분석에 사용된 자료의 계층 구조

수준	노드의 수	수준 별 노드의 수
최상위 수준: 총 계열	1	1
수준 1: 노선	4	4
최하위 수준: 역	5	20

그리고, 해당 자료의 2015년 4월 9일에 특별한 날이 아님에도 불구하고 지나치게 많은 수송량을 보여 이상점으로 판단하고, 제거 후 각 역의 평균을 대치하여 분석을 진행하였다. 평균 대치에는 R의 {imputeTS} 패키지(Moritz와 Bartz-Beielstein, 2017)를 이용하였다. 그림 3.1은 이상점 대치 작업을 완료한 모든 노드의 시도표에 해당한다. 여기서 A는 경부선, B는 경원선, C는 중앙선, D는 분당선에 해당한다.

<그림 3.1> 노드별 시도표



## (2) Prophet 모형 설정 과정

본 논문에서 이용한 Prophet 모형의 설정 과정에 대해 자세히 설명한다. 추세는 선형으로 모형화하고, 변화점은 자동으로 선택되도록 하였다. 또한, 계절효과는 승법적으로 설정했다. 표 3.2는 모형 적합에 고려한 휴일효과를 간략히 나타낸 것이다.

<표 3.2> 모형 설정에 고려한 휴일 효과

	날짜	Lower window	Upper window
구정	2014-01-31	-2	2
	2015-02-19	-2	2
	⋮	⋮	⋮
	2019-02-05	-2	2
추석	2014-09-08	-2	2
	2015-09-27	-2	2
	⋮	⋮	⋮
	2019-09-13	-2	2
Overday	2014-02-02	0	0
	2014-09-07	0	0
	⋮	⋮	⋮
	2019-09-15	0	0
크리스마스	2014-12-25	-1	0
	2015-12-25	-1	0
	⋮	⋮	⋮
	2019-12-25	-1	0
공휴일	2014-05-05	0	0
	2014-05-06	0	0
	⋮	⋮	⋮
	2019-10-09	0	0

2014년부터 2019년까지의 법정 공휴일을 고려하였다. 2018년 8월 31일까지의 자료임에도 불구하고, 2019년까지의 법정 공휴일을 고려한 이유는 예측에도 휴일 효과를 반영하기 위함이다. 설과 추석 당일의 전후 2일, 크리스마스의 경우 이브를 고려하여 1일 전을 휴일의 창으로 설정하여 해당 기간 내 일차에 따라 영향이 다르도록 설정하였다. 또한, 설과 추식에 일요일이 겹치는 경우(overday) 날짜를 따로 지정하여 예측 값이 지나치게 줄어드는 효과를 보정하였다. 그리고, Prophet 모형과 비교를 위해 고려한 지수평활법과 ARIMA 모형의 경우 AIC를 기준으로 최적의 모형을 선택했다.

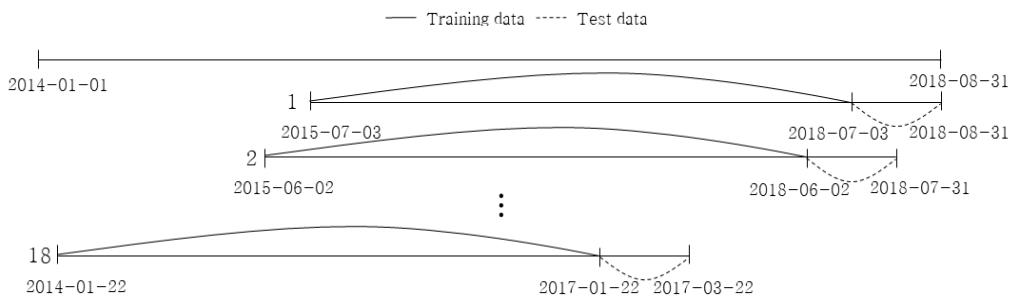


## 2. 모형 평가 및 비교

### (1) 예측 비교 방법

중·단기예측과 장기 예측 두 관점으로 나누어서 비교하였으며, 시계열 자료가 갖는 시간 종속적인 성질을 반영하기 위해 롤링 창 기법을 사용하였다. 먼저, 중·단기예측에서는 최근 관측치로부터 3년 자료를 이용해 60일 후까지 예측을 수행하였고, 표본 간 간격은 한 달로 설정하여 총 18개의 데이터 조각을 적합과 예측에 사용했다. 장기 예측의 경우도 이와 똑같은 설정으로 예측 시점만 180일 후까지 늘려주었다. 결론적으로, 장기예측에서는 총 14개의 데이터 조각이 적합과 예측에 사용되었다. 그림 3.2는 중·단기예측의 롤링 창 기법을 그림으로 표현한 것이다.

<그림 3.2> 중·단기예측의 롤링 창 기법



예측 성능 평가 측도로는  $MAPE$ 를 사용하였다. 롤링 창 기법을 통해 각 예측 시점마다 중·단기예측에서는 18개의 예측값, 장기예측에서는 14개의 예측값이 생성되므로, 우리는 각 시점에서  $MAPE$ 를 얻을 수 있다. 즉, 길이가  $n(t = 1, \dots, n)$ 인 시계열에 대해 중·단기예측의 롤링창 기법으로  $MAPE$ 를 계산한다고 했을 때, 각 노드의 예측 시점  $h$ 에 대한  $MAPE$ 는 식 (3.1)과 같이 정의할 수 있다.

$$MAPE_{i,h} = \frac{1}{18} \sum_{j=1}^{18} \left| \frac{Y_{i,n+h}^{(j)} - \tilde{Y}_{i,n}^{(j)}(h)}{Y_{i,n+h}^{(j)}} \right|, \quad i = 1, \dots, 25, \quad h = 1, \dots, 60, \quad (3.1)$$

여기서  $i$ 는 각 노드,  $j$ 는 데이터 조각,  $h$ 는 예측 시점을 의미한다. 장기예측의 경우  $j = 1, \dots, 14$ ,  $h = 1, \dots, 180$ 일 것이다. 본 논문에서는  $MAPE_{i,h}$ 를 기반으로 다양한 방면으로 모형의 예측 성능을 비교하였다. 앞으로 쓸 식들은 중·단기예측의 롤링창 기

법을 기준으로 한다. 다음의 식 (3.2)는 예측 시점에 따른 성능을 수준별로 비교하기 위해 고려하였다.

$$MAPE_h(k) = \frac{1}{m_k} \sum_{i(m_k)} MAPE_{i(m_k), h}, \quad k = 0, 1, 2, \quad (3.2)$$

여기서  $k$ 는 수준,  $i(m_k)$ 는 각 수준에 속하는 노드,  $m_k$ 는 해당 수준에 속하는 노드의 개수를 의미한다. 또한, 다음의 식 (3.3)은 모형의 전반적인 수준별 성능을 보기 위해  $MAPE_h(k)$ 를 예측 시점에 대해 평균을 취한 식이다.

$$MAPE(k) = \frac{1}{60} \sum_{h=1}^{60} MAPE_h(k). \quad (3.3)$$

마지막으로, 다음과 같이 모든 노드에 대한 전반적인 성능을 비교하기 위해  $MAPE(k)$ 를 수준  $k$ 에 대해 평균을 취하여 비교했다.

$$MAPE(\cdot) = \frac{1}{3} \sum_{k=0}^2 MAPE(k). \quad (3.4)$$

앞서 설명한 식들을 기반으로 모형의 중·단기예측 성능을 다양한 방면으로 평가 및 비교하고, 장기예측에서는 모든 노드에 대한 전반적인 성능을 기반으로 모형 간 비교를 수행한다.

## (2) 중·단기예측

먼저, 모형들의 전반적인 수준별 성능을 비교한다. 다음의 표 3.3은 각각의 계층적 시계열 방법을 적용한 모형별  $MAPE(k)$ 를 나타낸 것이다. 단, 여기서 기저 예측은 계층적 시계열 방법을 적용하지 않은 경우를 의미하며, 하향식은 예측 비율을 이용한 하향식 방법을 말한다.

<표 3.3> 모형별  $MAPE(k)$

	$MAPE(k)$		
	지수평활법	ARIMA	Prophet
<i>최상위 수준: 총 계열</i>			
기저 예측	8.68	9.21	3.43
상향식	8.57	9.17	3.40
하향식	8.68	9.21	3.43
미들 아웃	8.68	9.19	3.41
최적 조합	8.59	9.17	3.40
<i>수준 1: 노선</i>			
기저 예측	9.19	9.58	3.71
상향식	9.04	9.54	3.70
하향식	9.20	9.59	3.72
미들 아웃	9.19	9.58	3.71
최적 조합	9.06	9.55	3.70
<i>최하위 수준: 역</i>			
기저 예측	10.23	10.68	4.63
상향식	10.23	10.68	4.63
하향식	10.32	10.73	4.65
미들 아웃	10.32	10.72	4.64
최적 조합	10.24	10.68	4.63
<i>전체 평균</i>			
기저 예측	9.37	9.82	3.92
상향식	9.28	9.80	3.91
하향식	9.40	9.84	3.93
미들 아웃	9.40	9.83	3.92
최적 조합	9.30	9.80	3.91

표 3.3에서는 각 계층적 시계열 예측 방법을 적용한 모형별 예측 성능을 수준별로 확인할 수 있다. 단, 표의 가장 아래 영역인 전체 평균은 모든 노드에 대한 전반적인 예측 성능을 보기 위해,  $MAPE(\cdot)$ 을 나타낸 값이라 할 수 있다. 여기서 강조할만한 점은 어떤 계층적 시계열 예측 방법을 적용 혹은 적용하지 않더라도 Prophet 모형이 지수평활법과 ARIMA 모형보다 훨씬 더 우수한 성능을 보여준다는 것이다. 그리고, 모든 노드에 대해 전반적인 예측 성능이 가장 우수한 모형은 Prophet 모형에 최적 조합 방법이나 상향식 방법을 적용한 경우였다.

다음으로는, 상향식 방법을 적용한 각 모형의 예측 시점에 따른 수준별 예측 성능을 비교한다. 그림 3.3은  $MAPE_h(k)$ 를 모형별로 나타낸 것이다.

<그림 3.3> 상향식 방법을 적용한 모형별  $MAPE_h(k)$ 의 시도표

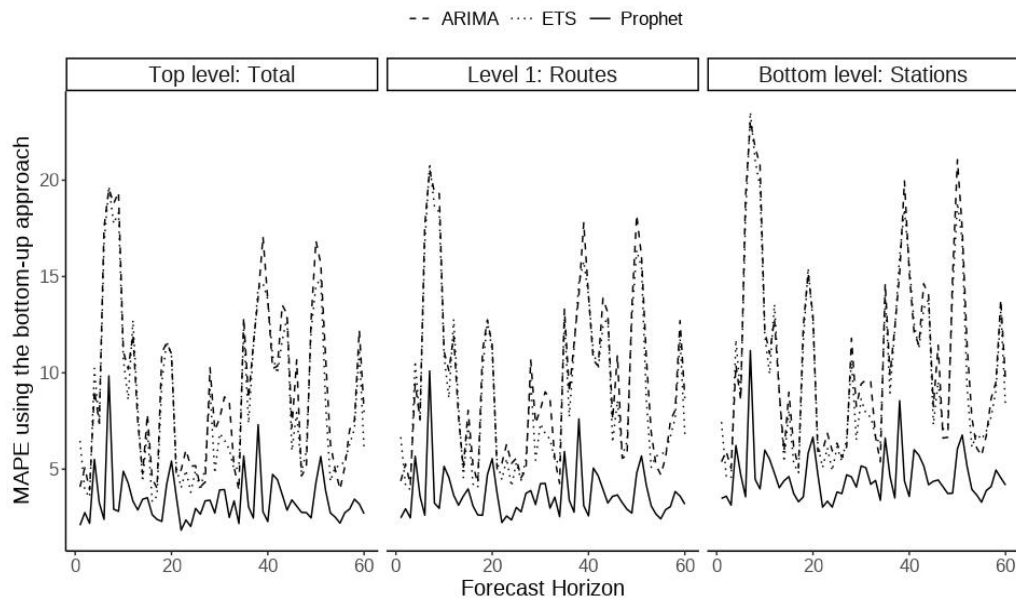


그림 3.3에서도 Prophet 모형의 성능은 잘 드러난다. Prophet 모형은 예측 시점에 관계없이 비교적 안정적인 성능을 보이며,  $MAPE_h(k)$ 의 평균뿐만이 아닌 변동도 매우 작아 보인다. 이는 그림 3.4의 모형별  $MAPE_h(k)$ 의 상자그림을 보면 명확하게 드러난다.

<그림 3.4> 모형별  $MAPE_h(k)$ 의 상자그림

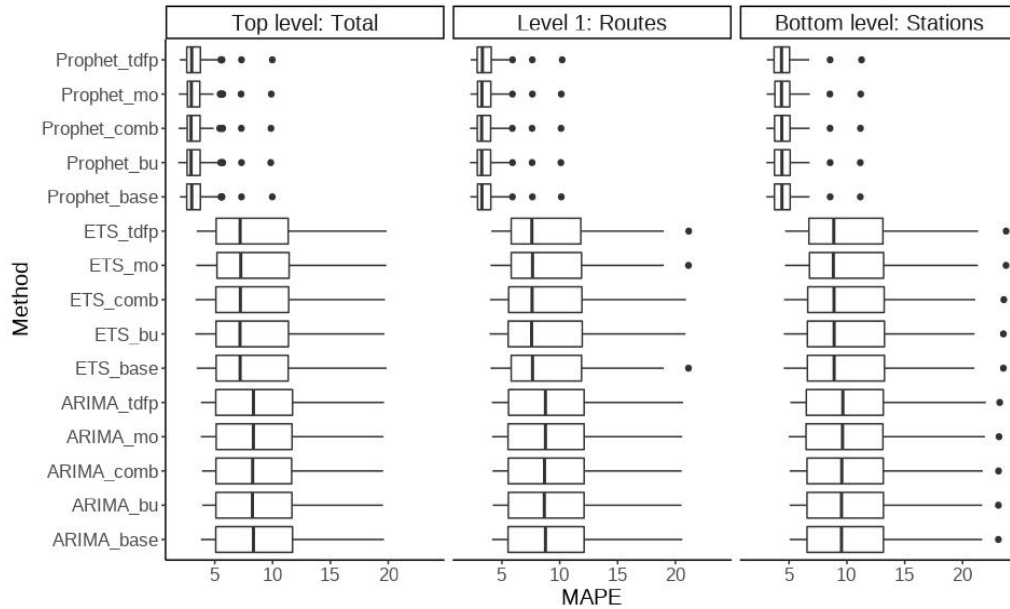
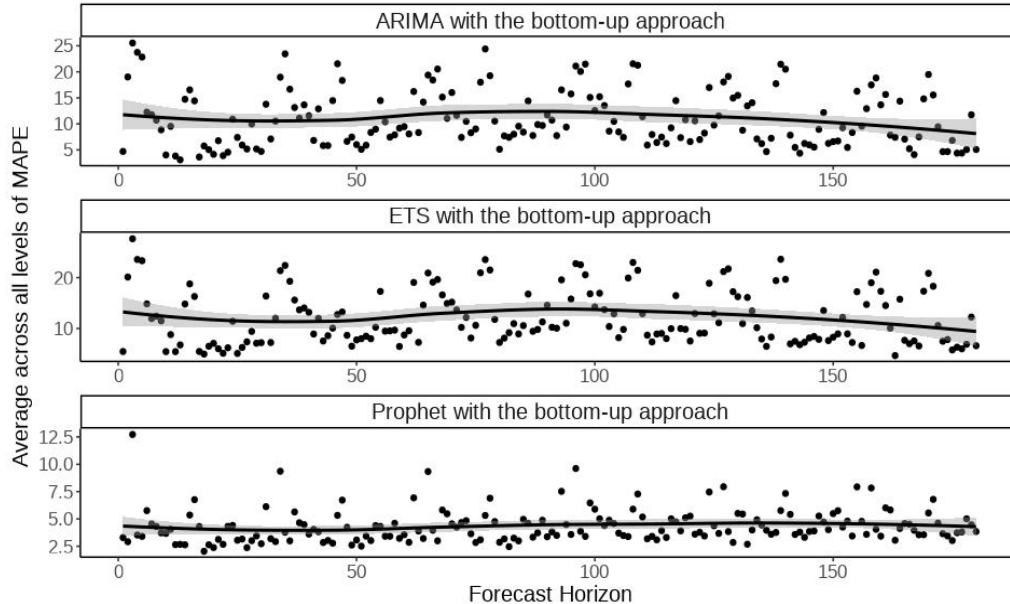


그림 3.4의 tdfp, mo, comb, bu, base는 각각 예측 비율을 이용한 상향식 방법, 미들아웃 방법, 최적 조합 방법, 상향식 방법, 계층적 시계열 예측 방법을 적용하지 않은 경우를 의미한다. Prophet 기반의 모형들은 나머지 모형들과 비교하면  $MAPE_h(k)$ 의 평균이 매우 작으며, 그뿐만 아니라 변동도 매우 작은 모습을 보여주고 있음을 확인할 수 있다.

### (3) 장기예측

예측 시점을 180일 후까지 늘려 모형별 장기예측 성능을 간략히 비교해보았다. 계층적 시계열 예측 방법으로 상향식 방법을 택하였으며, 모든 노드에 대한 전반적인 성능을 그림 3.5를 통해 비교하였다. 그림 3.5의 점들은 모형별  $MAPE_h(\cdot)$  값을 나타내며, 곡선은 점들을 평활하여 구한 것이다.

<그림 3.5> 상향식 방법을 적용한 모형별  $MAPE_h(\cdot)$  의 시도표

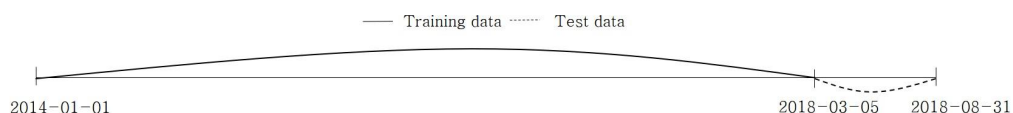


$y$  축의 값을 주의하며 봐야 한다. Prophet 모형은 나머지 두 모형에 비해 훨씬 더 좋은 예측 성능을 보일 뿐만 아니라, 예측 시점에 관계없이 안정적인 성능을 보여주고 있다. 이는 우리가 중·단기예측에서 확인한 결과와 같다. 즉, 장기예측에서도 Prophet 모형이 나머지 두 모형에 비해 우수한 성능을 보인다.

### 3. 모형 예측 성능 평가

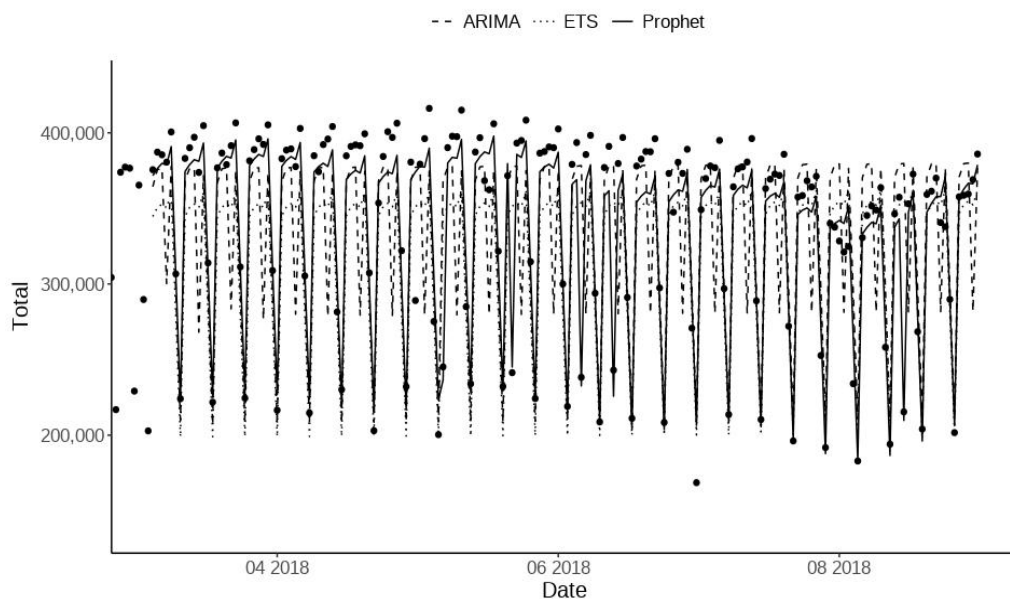
우리는 앞서 Prophet 기반의 모형이 중·단기예측뿐만 아니라, 장기예측 성능에서도 뒤떨어지지 않음을 확인했다. 이번에는 롤링 창 기법을 이용하지 않고, 다음의 그림 3.6과 같이 총 4년 9개월 치 자료의 마지막 6개월을 시험자료로 사용하여 원자료와 예측값을 기반으로 성능을 비교한다.

<그림 3.6> 자료 분할



계층적 시계열 예측 방법으로 상향식 방법을 이용하였으며, 총 계열을 대표적으로 시각화하여 비교하였다. 먼저, 원자료와 예측값의 시도표를 통해 비교해보고자 한다. 다음의 그림 3.7에서 점은 원자료, 선은 상향식 방법을 적용한 각 모형의 예측값을 나타낸다. 모형 간 비교를 용이하게 하기 위해  $x$  축의 날짜를 2017년 3월 5일부터 2018년 8월 31일까지 확대하여 나타냈다.

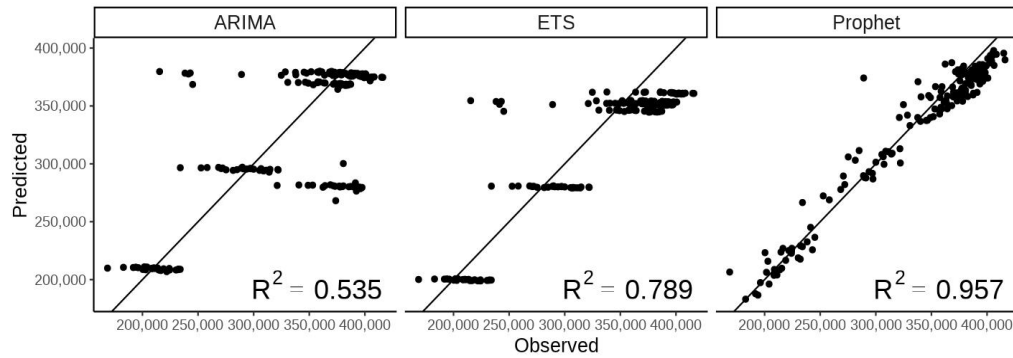
<그림 3.7> 상향식 방법을 적용한 모형별 총 계열의 시도표



ARIMA 모형과 지수평활법의 경우 일정한 구간 내에서 예측값들이 주어지지만, Prophet 모형의 경우 원자료의 패턴을 따라가며 예측을 해내고 있음을 확인할 수 있다.

마지막으로 모형 간 예측 성능을 원자료와 예측값 간의 산점도를 통해 비교하였다. 그림 3.8에서  $x$  축은 관측값이며,  $y$  축은 상향식 방법을 적용한 모형별 예측값, 빨간색 직선은  $y = x$ 을 나타낸다.

<그림 3.8> 상향식 방법을 적용한 모형별 관측값과 예측값의 산점도



ARIMA 모형 및 지수평활법의 예측값은 관측값을 과소 또는 과대 추정하는 경우가 상당히 많이 발생하고 있어, 점들이  $y = x$  직선에서 많이 벗어나고 있다. 반면, Prophet 모형은 점들이  $y = x$  직선상에 잘 놓여있고,  $R^2$ 가 95.7%에 달한다. 즉, Prophet 모형이 나머지 두 모형에 비해 훨씬 더 우수한 예측 성능을 보여주고 있다.



## IV. 결론

본 논문에서는 주로 계층적 시계열 예측과 Prophet 모형에 관한 내용을 다루었으며, Prophet 모형과 비교 점으로 지수평활법과 ARIMA 모형을 고려하였다. 그리고, 계층적 시계열 자료인 광역철도 일별 수송량 자료를 이용해 각각의 계층적 시계열 예측 방법을 적용한 모형별 예측 성능을 비교하는 실증 분석을 수행하였다. 그 결과, Prophet 모형이 나머지 두 모형에 비해 매우 우수한 예측 성능을 보여준을 확인하였고, 각 모형의 예측값에 대한 계층적 시계열 예측 방법의 적용은 한층 더 정확한 예측 성능을 보여주었다.

본 논문에서는 과거 논문들과 달리 계층적 시계열 예측 방법의 적용에 Prophet 모형을 고려하였다. 그러나, R에서 계층적 시계열 예측을 제공하는 {hts} 패키지에서 기본적으로 제공되는 모형은 지수평활법과 ARIMA 모형뿐이다. 그래서, Prophet 모형을 계층적 시계열 방법에 적용하기 위해 {hts} 패키지의 forecast.gts() 함수를 수정하여 사용하였다. 앞으로 계층적 시계열 예측에 다양한 시계열 예측 모형이 실무에서 쉽게 이용되기 위해서는 {hts} 패키지의 추가적인 개발이 필요로 된다.

## 참고 문헌

- [1] 이주은 · 성병찬(2017), “계층적 시계열 분석을 이용한 지역별 교통사고 발생건수 예측,” *한국통계학회*, 30(1), 181-193.
- [2] Athanasopoulos, G., R. A. Ahmed and R. J. Hyndman(2009), “Hierarchical Forecasts for Australian Domestic Tourism,” *International Journal of Forecasting*, 25, 146 - 166.
- [3] Gardner, E.S., Jr.(1985), “Exponential Smoothing: The State of the Art,” *Journal of Forecasting*, 4, 1-28.
- [4] Gross, C. and J. Sohl(1990), “Dissagregation Methods to Expedite Product Line Forecasting,” *Journal of Forecasting*, 9, 233-254.
- [5] Harvey, A.C. and S. Peters(1990), “Estimation Procedures for Structural Time Series Models,” *Journal of Forecasting*, 9, 89-108.
- [6] Hastie, Trevor. and Robert. Tibshirani(1987), “Generalized Additive Models: Some Applications,” *Journal of the American Statistical Association*, 82:398, 371-386.
- [7] Hyndman, R. J., A. B. Koehler, R. D. Snyder and S. Grose(2002), “A State Space Framework for Automatic Forecasting using Exponential Smoothing Methods,” *International Journal of Forecasting*, 18, 439 - 454.
- [8] Hyndman, R. J., R. A. Ahmed and G. Athanasopoulos(2007), *Optimal Combination Forecasts for Hierarchical Time Series*, Working paper 09/07, Department of Econometrics and Business Statistics, Monash University.
- [9] Hyndman, R. J., R. A. Ahmed, G. Athanasopoulos and Shang, H. L.(2011), “Optimal Combination Forecasts for Hierarchical Time Series,” *Computational Statistics and Data Analysis*, 55, 2579 - 2589.

- [10] Moritz, S. and T. Bartz-Beielstein(2017), "ImputeTS: Time Series Missing Value Imputation in R," *The R Journal*, 9(1), 207 - 218.
- [11] Pegels, C. C.(1969), "Exponential Smoothing: Some New Variations," *Management Science*, 12, 311 - 315.
- [12] Taylor, James.(2003), "Exponential Smoothing with a Damped Multiplicative Trend," *International Journal of Forecasting*, 19, 715 - 725.
- [13] Taylor, Sean. and Letham. Benjamin(2018), "Forecasting at Scale," *The American Statistician*, 72:1, 37-45.