

# Time Series Analysis for Traffic Flow Using Dynamic Linear Model

Kim Hong Geun<sup>†</sup> · Park Chul Young<sup>\*\*</sup> · Shin Chang Sun<sup>\*\*\*</sup> ·  
Cho Yong Yun<sup>\*\*\*</sup> · Park Jang Woo<sup>\*\*\*\*</sup>

## ABSTRACT

It is very challenging to analyze the traffic flow in the city because there are lots of traffic accidents, intersections, and pedestrians etc. Now, even in mid-size cities Bus Information Systems(BIS) have been deployed, which have offered the forecast of arriving times at the stations to passengers. BIS also provides more informations such as the current locations, departure-arrival times of buses. In this paper, we perform the time-series analysis of the traffic flow using the data of the average travel time and the average speed between stations extracted from the BIS. In the mid size cities, the data from BIS will have a important role on prediction and analysis of the traffic flow. We used the Dynamic Linear Model(DLM) for how to make the time series forecasting model to analyze and predict the average speeds at the given locations, which seem to show the representative of traffics in the city. Especially, we analysis travel times for weekdays and weekends separately. We think this study can help forecast the traffic jams, congestion areas and more accurate arrival times of buses.

**Keywords** : BIS, DLM, Traffic

## 동적 선형 모델을 이용한 교통 흐름 시계열 분석

김 홍 근<sup>†</sup> · 박 철 영<sup>\*\*</sup> · 신 창 선<sup>\*\*\*</sup> · 조 용 윤<sup>\*\*\*</sup> · 박 장 우<sup>\*\*\*\*</sup>

## 요 약

도시 내의 교통흐름을 정확히 분석하는 것은 매우 어렵다. 특히, 교통사고나 보행자, 교차로 등은 도시 내의 교통흐름을 분석하는데 있어서 어려움을 가중시킨다. 현재 소도시(예를 들어 전라남도 순천시)들에도 버스 정보시스템(Bus Information System, BIS)이 보급되어 있고 이를 통하여 도착시간 예측 등과 같은 정보를 제공하고 있다. BIS는 버스의 위치, 구간별 이동시간, 출발-도착 시간등을 제공하고 있다. 따라서 본 논문에서는 BIS로부터 정류장 간의 평균 이동 시간, 그리고 이동 거리 등을 시간대 별로 추출하여 도시 내의 교통흐름을 시계열 분석하고자 한다. 소도시의 경우 버스 정보는 도시교통 흐름을 설명하는 중요하고 효과적인 자료이다. 앞서 언급한 신호등 지연, 보행자, 교차로 등은 교통흐름을 분석하고 예측하는데 어려움을 더한다. 본 논문에서는 동적 선형 모델(Dynamic Linear Model, DLM)을 이용하여 주요 구간의 교통흐름을 시계열 분석하고 예측하는 방법을 제시한다. 이때 구간별 통행 속도를 평일과 주말로 나누어 분석한다. DLM을 이용하여 구간별 이동속도의 시계열 분석을 통하여 도시 내의 교통흐름을 파악하여 향후 교통 체증 및 혼잡 구간을 예보하고 버스의 정확한 도착시간을 예측하는데 도움을 줄 수 있으리라 생각한다.

**키워드** : 버스정보시스템, 동적선형모델, 교통흐름

## 1. 서 론

버스정보시스템(Bus Information System, 이하BIS)은 현재 버스의 위치를 전송하여 도착예정시간을 예측한 후 시민들에게 정보를 제공하는 매우 유용한 수단이다[1].

BIS와 지능형교통시스템(Intelligent Transport System; 이하 ITS)은 실시간으로 수많은 데이터를 수집하고 있으며 이를 활용하여 직접적인 교통흐름을 예측하는 것은 중요하다. 특히 현재와 과거의 측정치를 활용하여 교통흐름을 예측한 후 교통 흐름을 제어한다면 더욱 좋을 것이다. 지방 중소규모의 도시에서는 버스전용차선 운영이 없으며 대체로 대도시에 비하여 교통흐름이 원활하기 때문에 BIS를 이용해 전체 교통흐름을 파악하는 것이 가능하리라 생각한다[2].

BIS에서 핵심서비스는 버스도착예정시간을 제공하는 것이다. 버스도착예정시간은 현재 버스의 실시간 위치와 남은

※ 본 연구는 2016년도 순천대학교 학술기반조성비로 연구되었음.  
† 준 회 원 : 순천대학교 전기·전자·정보통신공학과 박사과정  
\*\* 비 회 원 : 순천대학교 전기·전자·정보통신공학과 박사과정  
\*\*\* 정 회 원 : 순천대학교 정보통신공학과 부교수  
\*\*\*\* 정 회 원 : 순천대학교 정보통신공학과 교수  
Manuscript Received : December 13, 2016  
Accepted : January 25, 2017  
\* Corresponding Author : Park Jang Woo(jwpakr@sunchon.ac.kr)

거리를 이용하여 계산된 값으로 각 버스 정류장에 위치한 단말기를 통해 제공하고 있다. 또한 이동평균필터 기반의 알고리즘을 통해 구간 통과시간과 구간거리를 예측모형에 사용하지만 도로의 교통흐름, 신호주기, 이상상황, 데이터 결측 등의 상황은 고려되지 않은 상태로 정보를 제공하고 있다[3].

따라서 본 논문에서는 BIS로부터 구간별 이동 시간, 통행속도 등의 시계열 데이터를 추출하고 모델링 하여 도시 내의 교통 흐름을 분석하여 활용하고자 한다. 본 논문에서 교통 흐름 시계열 데이터 분석에는 동적 선형 모델(Dynamic Linear Models, DLM)을 이용했다. 이 방법은 관측 모델과 상태 모델로 구성되며 관측 데이터의 근본적인 형태를 예측할 수 있다. 또한 이 모델은 관측데이터의 가우시안 분포를 가정하므로 적용이 쉬운 장점을 갖는다. 구간별 통행속도를 통해 평일, 주말 등에 따른 교통흐름을 나누어 분석하고, 이를 통하여 버스도착 예측 시간의 정확도를 보정하는데 이용하고자 한다.

이 논문에서 사용한 BIS 데이터는 지방중소규모 도시(전라남도 순천시)의 2015년도 데이터이다. 이 도시의 버스 노선은 대부분 시내권 위주로 집중되어 있으며, 평균 운행 거리는 짧고 평균배차간격이 크다. 2015년 한해 운행 중인 버스의 1대당 평균 운행 거리는 29.59 km이고 평균 운행 속도는 32.41 km/h이다. 순천시는 인근지역 도시(광양, 여수)와 인접하여 광역노선을 운행하고 있으며 이들 도시 중 주요 주거지 특성을 갖는다.

## 2. 관련 연구

본 논문에서는 BIS를 통해 각 구간에 따른 통행속도를 시간대별로 구분하여 특성을 분석하고, 통행속도 예측을 통해 교통흐름 분석 및 버스도착예정시간의 정확도를 개선할 수 있는 방안을 제시하고자 한다. 이에 따라 기존에 BIS에서 예측을 위해 사용되고 있는 이동평균필터 및 본 연구에서 모델 구축을 위해 사용한 동적선형모델에 대해 알아보고, 모델 검증을 위한 방법에 대해 선행 연구를 수행했다.

### 2.1 Moving Average Filter

MAF는 시간에 따라 변하는 데이터를 확인할 때, 일정한 크기의Queue와 Buffer를 활용하여 Buffer의 크기 내에서만 최근의 데이터를 계속 반영하면서 평균을 적용하는 필터이다.

$$y_t = \frac{1}{M} \sum_{i=0}^{M-1} x_{t+i} \quad (1)$$

$x_{t+i}$ 는 통행속도에 대한 관측값이고,  $y_t$ 는 통행속도에 대한 예측값이며,  $M$ 은 평균을 낼 입력의 개수(큐의 크기)를 나타내고,  $t$ 는 실시간으로 들어오는 데이터의 순번을 나타내며,  $i$ 는 데이터의 수를 나타낸다.

위의 식을 반복하기 위한 재귀식은 다음과 같다.

$$\bar{y}_t = \bar{y}_{t-1} + \frac{1}{M} \sum_{i=0}^{M-1} x_{t+i} \quad (2)$$

재귀식의 최종값  $\bar{y}_t$ 에서  $\bar{y}_{t-1}$ 는 직전단계의 예측값을 나타내고, 가장 최근 데이터, 가장 오래된 데이터, 데이터의 수를 이용하여 계산하는 형태로 구성된다.

### 2.2 Dynamic Linear Model

상태 공간 모델은 동향 및 계절성 성분의 결과를 자연적 해석하기 위해 비정상 단변량 시계열을 모델링을 하기 위해 사용된다. R 언어 및 통계 컴퓨팅을 위한 환경에서 최대우도 추정(Maximum Likelihood Estimation; 이하 MLE)을 통해 매개변수를 추정하였다. 상태 공간모델에서 요하는 주 목표는 눈에 띄지 않는 통행속도에 대한 상태와 관측 순서의 일부를 기반으로 하는 미래의 측정치에 대한 예측을 추론하는 것이다. 통행속도에 대한 상태의 조건부 분산 계산을 위해 사용 가능한 정보, 측정치, 완전한 예측 정보가 주어진다. 동적 상태 공간모델에서, DLM은 상태 벡터의 현재 추론을 업데이트 하기 위한 수식을 제공한다. 일반적인 상태식에 의해 통제되는 이산시간 제어 프로세스의  $x_t$  상태를 추정하는 DLM의 일반적인 해결방법은 측정(상태) 방정식과 시스템 방정식으로 나뉘며 다음과 같이 표현할 수 있다[3-5].

$$x_t = G_t x_{t-1} + w_t, \quad w_t \sim N_p(0, W_t) \quad (3)$$

$$y_t = F_t x_t + v_t, \quad v_t \sim N_m(0, V_t) \quad (4)$$

$x_t$ 는  $t$ 시간의 상태벡터이고,  $y_t$ 는  $t$ 시간에 측정된 값이다.  $G_t$ 는 전이행렬이고,  $F_t$ 는  $t$ 시간에서의 관측(또는 설계) 행렬이다.  $w_t$ 는 상태에 대한 시스템 에러이고,  $v_t$ 는  $t$ 시간에서의 측정 에러이다.  $W_t$ 는 시스템 개선을 위한 공분산 행렬이고,  $V_t$ 는 관측 잡음 공분산 행렬이다.

DLM의 추정 과정은 시스템 모델로부터 이전 상태 추정값으로 현재의 상태 값을 예측(Prediction)하는 과정과 측정값을 사용하여 예측된 값을 교정(Correction)하는 두 단계로 이루어진다. 예측단계는 재귀 상태 예측 및 측정 업데이트 알고리즘으로 구성될 수 있다.

$$\begin{aligned} x_0 &\sim N_p(m_0, C_0) \\ x_{t-1}|y_{1:t-1} &\sim N(m_{t-1}, C_{t-1}) \end{aligned} \quad (5)$$

상태 추정(선행적):

$$\bar{x}_{t|t-1} = E(x_t|y_{1:t-1}) = G_t \bar{x}_{t-1|t-1} \quad (6)$$

에러 공분산 추정(선형적):

$$P_{t|t-1} = \text{Var}(x_t|y_{1:t-1}) = G_t P_{t-1|t-1} G_t' + Q_t \quad (7)$$

예측 상태 추정(선형적)은 현재 시간 단계 정보를 포함하지 않는다. 측정이 수행된 후, 상태 추정을 구체화하기 위해 예측을 비교하게 된다. 측정 업데이트식은 다음과 같다.

측정치 개선

$$\bar{y}_t = y_t - F_t' \bar{x}_{t|t-1} \quad (8)$$

공분산 개선

$$S_t = F_t P_{t|t-1} F_t' + R_t \quad (9)$$

칼만이득 계산

$$K_t = P_{t|t-1} F_t' S_t^{-1} \quad (10)$$

상태 추정(후행적):

$$\bar{x}_{t|t} = \bar{x}_{t|t-1} + K_t \bar{y}_t \quad (11)$$

에러 공분산 추정(후행적):

$$R_{t|t} = P_{t|t-1} - K_t F_t P_{t|t-1} \quad (12)$$

DLM은 상태 벡터의 관측을 위한 예측 값을 분산/공분산 행렬과 함께 one-step-ahead로 계산되어진다. DLM의 one-step-ahead 예측은 한 주기 후 좋은 예측 값을 생성할 수 있다.

최대우도추정(MLE; Maximum Likelihood Estimate)를 사용하여 정규분포를 가정하고, 분포의 모수를 추정하는데 활용된다. 어떤 모수  $\theta$ 로 결정되는 확률변수들의 모임  $D_\theta = (x_1, x_2, \dots, x_n)$ 이 있고,  $D_\theta$ 의 확률밀도함수나 확률질량 함수  $f$ 가이고, 그 확률변수들에서 각각 값  $x_1, x_2, \dots, x_n$ 을 얻었을 경우, 가능도  $L(\theta)$ 는 다음과 같다.

$$L(\theta) = f_\theta(x_1, x_2, \dots, x_n) \quad (13)$$

여기서 가능도를 최대로 만드는  $\theta$ 는 다음과 같다.

$$\bar{\theta} = \underset{\theta}{\text{argmax}} L(\theta) \quad (14)$$

이 때,  $x_1, x_2, \dots, x_n$ 이 모두독립적이고 같은 확률분포를 가지고 있다면,  $L$ 은 다음과 같이 표현이 가능하다.

$$L(\theta) = \prod_i f_\theta(x_i) \quad (15)$$

로그함수는 단조 증가하므로,  $L$ 에 로그를 씌운 값의 최대값은 원래값  $\bar{\theta}$ 와 같고, 이 경우 계산이 비교적 간단해 지고 다음과 같다.

$$L^*(\theta) = \log L(\theta) = \sum_i \log f_\theta(x_i) \quad (16)$$

또한, 모델 검증을 위해 RMSE 및 MAPE를 통해 검증한다. RMSE (Root Mean Square Error)는 모델이 예측한 값과 관측 값의 차이에 대한 측도를 나타내는 것으로, 잔차들을 하나의 측도로 종합하여 정밀도를 판단하는데 활용한다.

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (17)$$

MAPE (Mean Absolute Percent Error)는 RMSE와 마찬가지로 차이에 대한 측도를 백분율로 표현하여 정밀도를 판단하는데 활용한다.

$$\frac{1}{n} \sum \left| \frac{(y_t - \bar{y}_t)}{y_t} \right| \times 100 \quad (18)$$

### 3. 데이터 수집 및 분석

본 연구에서는 순천시BIS를 통해 수집된 '15년도에 데이터를 활용하였고, 구간별 특성 분석을 위해 교통 혼잡도가 높은 구간을 선정하였다. 본 논문에 사용된 자료는 순천시 버스정보시스템의 도착예측정보 계산 및 제공을 위해 수집된 독실<->선거관리위원회와 국민은행<->금당고(조례초) 구간의 2015년 1월 1일~12월 31일 간의 자료를 활용하였다. 수집된 버스데이터는 정류장과 정류장 사이의 데이터로 정차시간을 제외한 속도 데이터이다. Table 1은 분석 구간에 대한 세부사항을 나타낸다.

분석구간은 주거지역과 시내권 사이의 구간으로 출퇴근 시간대 및 주간 시간대의 교통혼잡도가 매우 높은 구간이다. 순천시의 '15년도 전구간 버스의 평균 통행속도는 32.4 1 km/h로 A 구간 16.2 km/h, B 구간 16.3 km/h, C 구간 22.33755 km/h, D 구간 12.30512 km/h로 모두 낮은 통행속도를 나타낸다.

Fig. 2는 2015년도에 시내권에서 주거지역으로 이동하는 A 구간의 버스 통행속도를 시간대에 따른 평일의 그래프이다.

Table 1. Detail Information of the Bus Route

	Length of Route(m)	Average Speed(km/h)	Lane Number (One way)
Route A (Duksil->National Election Commission)	454.6	16.2	2
Route B (National-Election Commission->Duksil)	452.6	16.3	2
Route C (Kukmin Bank ->KumdangHigh-School)	464.2	22.33755	4
Route D (Jorye-Elementary-School->Kukmin Bank)	461.3	12.30512	4

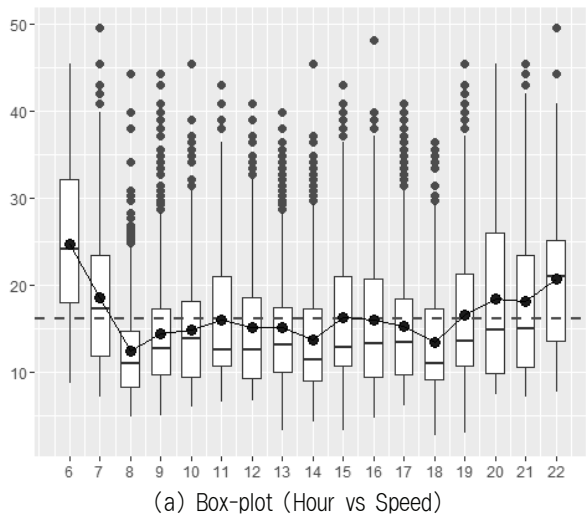


(a) Route A and Route B

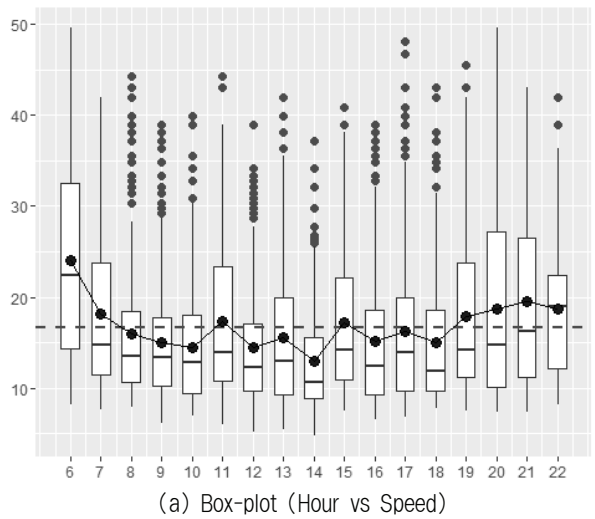


(b) Route C and Route D

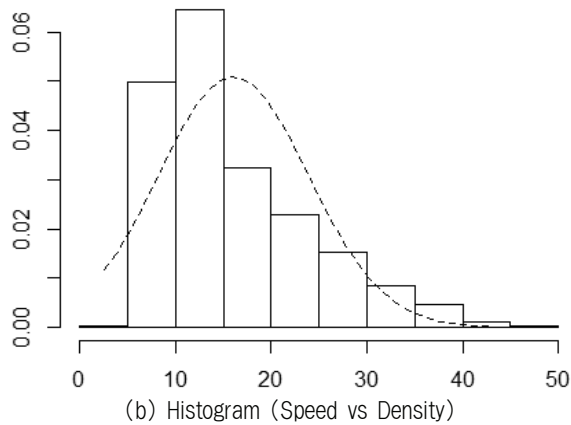
Fig. 1. Detail Information of the Bus Route  
(출처: 순천시 교통관제센터, <http://its.sc.go.kr>)



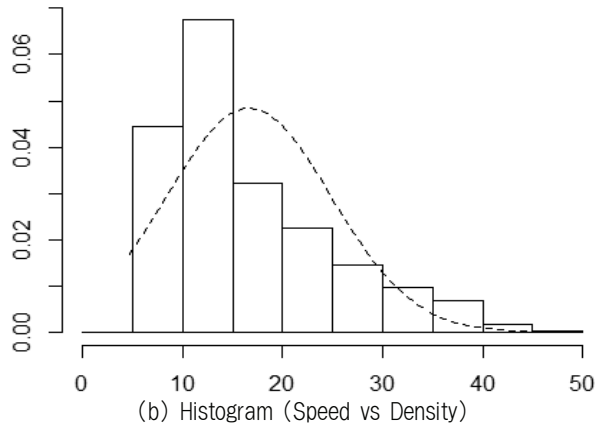
(a) Box-plot (Hour vs Speed)



(a) Box-plot (Hour vs Speed)



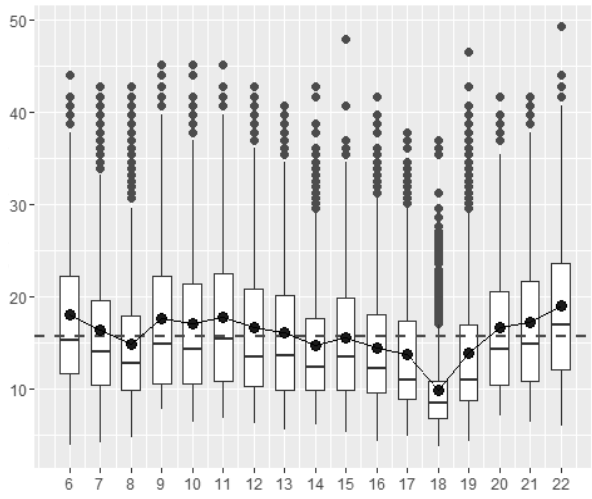
(b) Histogram (Speed vs Density)



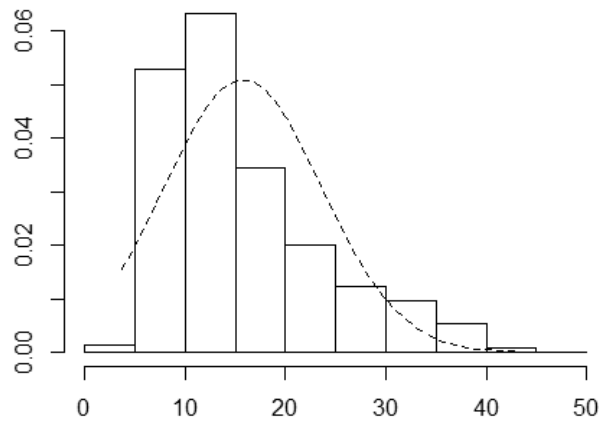
(b) Histogram (Speed vs Density)

Fig. 2. Weekly Data of Route A Based on Speed vs Hour (Year: 2015) for (a) and (b)

Fig. 3. Weekend Data of Route A Based on Speed vs Hour (Year: 2015) for (a) and (b)

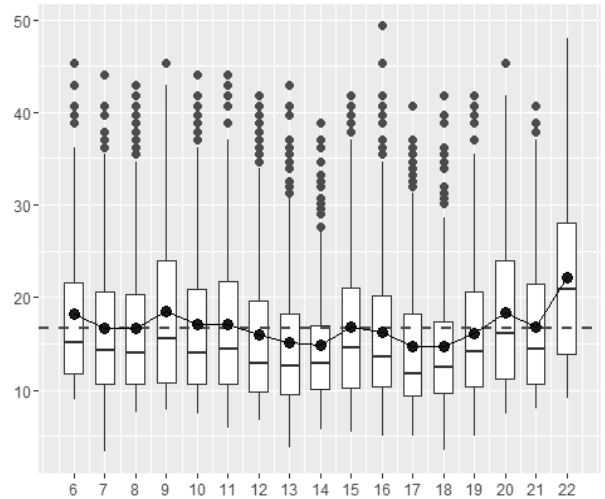


(a) Box-plot (Hour vs Speed)

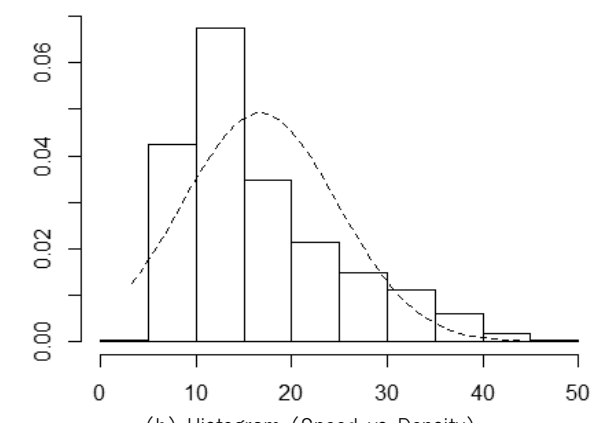


(b) Histogram (Speed vs Density)

Fig. 4. Weekly Data of Route B Based on Speed vs Hour (Year: 2015) for (a) and (b)



(a) Box-plot (Hour vs Speed)



(b) Histogram (Speed vs Density)

Fig. 5. Weekend data of Route B Based on Speed vs Hour (Year: 2015) for (a) and (b)

A 구간은 8시와 18시 경의 출퇴근 시간대 통행속도가 가장 낮게 나타나며 평균 속도 보다 낮은 시간대가 많다. 히스토그램은 평균속도보다 낮은 영역에서 분포가 형성되어 있다.

Fig. 3은 A 구간의 주말에 대한 데이터로 평일 데이터와는 달리 출퇴근 시간대의 패턴은 보이지 않는다. 11시와 15시 이외의 오전 오후 시간에는 평균 이하의 속도를 유지하는 것을 확인할 수 있다. A 구간은 시내권 방향으로 이동하는 구간으로 주말과 평일 모두 평균속도보다 낮은 통행속도를 보인다.

Fig. 4는 2015년도에 주거지역에서 시내권으로 이동하는 B 구간의 버스 통행속도를 시간대별로 평일의 데이터를 표현한 그래프이다. B 구간은 18시 경의 퇴근 시간대 통행속도가 매우 낮게 나타난다.

Fig. 5는 B 구간의 주말에 대한 데이터로 퇴근 패턴은 나타나지 않으며, 낮은 통행속도를 나타내고 있다.

## 4. 데이터 모델링 및 검증

### 4.1 Data Modelling

현재 순천시 BIS는 도착예정 시간 예측 서비스를 제공하는데 MAF 방법을 사용한다. 본 연구에서는 기존의 서비스를 개선하기 위해 단변량 시계열을 모델링하는데 주로 사용되는 동적 예측 모델인 DLM 모델을 적용하였다. 각 구간별 데이터의 정밀 분석을 위해 평일과 주말 데이터로 구분하였고, 단위 구간별 DLM 모델의 구현을 위해 MLE를 활용하여 추정(상태) 및 시스템 오차를 구한 결과는 Table 2와 같다.

Fig. 6(a)는 A 구간의 평일 데이터를 활용하여 관측 데이터(회색선)에 대해 MAF(청색선)와 DLM(적색선) 모델을 적용하여 비교한 결과이다. Fig. 6(b)는 DLM을 통해 예측된 값을 활용한 잔차의 정규성 검증을 위해 분위수-분위수 플

Table 2. System and Observation(State) Error Using MLE

	Weekend data of Route A	Weekly data of RouteA	Weekend data of Route B	Weekly data of RouteB
DV	2153.138	2371.12	2149.175	2295.924
DW	1076.569	1185.56	1074.588	1147.962

롯(Quantile-Quantile Plot; 이하 Q-Q Plot)을 나타냈고, Fig. 6(c)는 Histogram을 나타냈다. One-step-ahead 예측 데이터를 표현한 결과로 MAF보다 DLM 모델의 속도 추정 결과가 관측데이터의 특성을 잘 반영하고 있다, Figs. 6(b)와 6(c)을 통해 보면 45도 기울기의 직선상에 측정값과 예측값에 대한 분포가 양끝에서 이상치를 보이고 있으나, 대체적으로 정규분포의 형태를 따르고 있다.

Fig. 7(a)는 A 구간의 주말 패턴에 대해 모델을 적용한 결과를 비교한 그래프이다. Figs. 7(b)와 7(c)는 1.8 이상에서 벌어지기는 하지만 정규분포의 형태를 따르고 있다.

Fig. 8(a)는 B 구간의 평일 데이터에 대해 모델을 적용한 결과를 비교한 그래프이고, Figs. 8(b)와 8(c)는 Fig. 6과 Fig. 7 보다 정규분포의 형태를 잘 따르고 있는 것을 확인할 수 있다.

Tab 3. RMSE and MAPE Calculation

	Weekly data of Route A		Weekend data of Route A	
	MAF	DLM	MAF	DLM
RMSE	7.362723	4.316624	7.6961	4.584328
MAPE (%)	41.51491	23.53085	40.9204	23.71282
	Weekly data of Route B		Weekend data of Route B	
	MAF	DLM	MAF	DLM
RMSE	7.360572	4.409726	7.666814	4.661685
MAPE (%)	41.37651	24.07824	41.19096	24.59674

Fig. 9(a)는 B 구간의 평일 데이터에 대해 모델을 적용한 결과를 비교한 그래프이고, Figs. 9(b)와 9(c)는 분포의 양 끝에서 정규분포보다 높게 형성되나, 대체적으로 정규분포의 형태를 따르고 있다.

4.2 Model Validation

본 논문에서는 BIS 관측데이터를 통해 DLM을 구축하였고, 기존에 사용되는 MAF 방법과 비교 분석을 수행하였다.

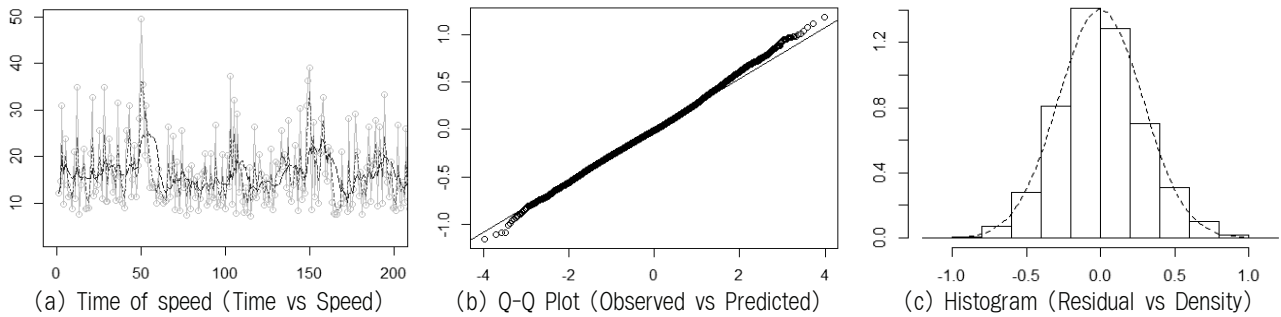


Fig. 6. Weekly Data of Route A based on Time Series of Speed (Year: 2015) for (a) and (b)

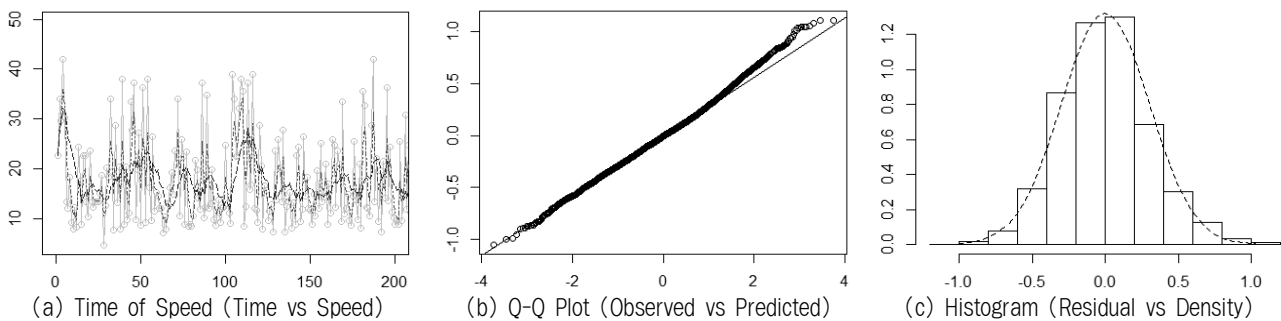


Fig. 7. Weekend Data of Route A based on Time Series of Speed (Year: 2015) for (a) and (b)

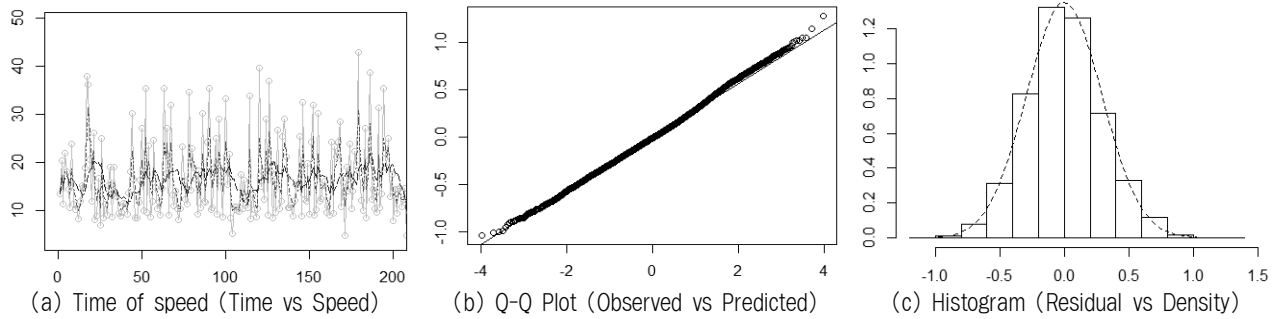


Fig. 8. Weekly Data of Route B based on Time Series of Speed (Year: 2015) for (a) & (b) & (c)

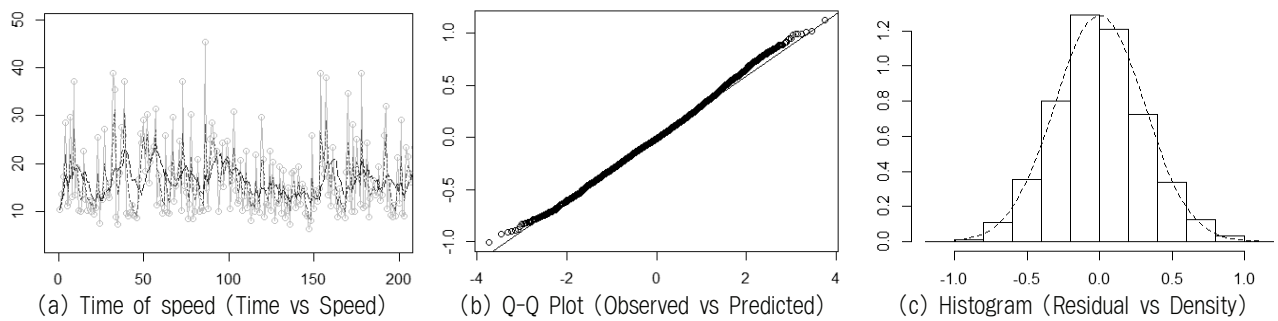


Fig. 9. Weekend Data of Route B based on Time Series of Speed (Year: 2015) for (a) & (b) & (c)

동적 선형 모델을 이용한 결과의 검증에 위해 RMSE와 MAPE를 계산한 결과를 Table 3에 나타냈다.

관측값과 예측값 사이에서 MAF 모델에 비해 DLM의 결과가 보다 관측값의 특성을 정밀하게 반영하고 있다. 본 논문에서는 A, B 구간에 대해 DLM을 적용하였다. 주중과 주말 데이터의 경우 DLM을 이용한 방법이 정밀한 추정치를 보이며 관측데이터와의 오차가 작다. 현재 시스템에 이용되는 MAF에 비해 데이터의 특성을 잘 반영하고 있다.

본 논문에서 적용된 동적 선형 모델을 C와 D 구간에 대해서도 동일하게 적용했지만, D구간에 대해서는 경험적 분포와 이론적 분포가 근접하게 이루어지고 있어 정규화가 더 필요한 상태이며, 이는 5장에서 다루기로 한다.

### 5. 결론 및 향후 연구

본 연구에서는 BIS의 각 단위 구간별 통행속도를 활용하여 DLM 기반의 예측 모델을 구현하였다. DLM의 매개변수들은 MLE를 통해 추정하였고, 이를 활용하여 기존의 MAF 방법과 비교분석을 수행하였다. 생성된 DLM은 정규분포를 가정하여 A, B 구간의 통행속도에 대해 모델을 적용한 결과 일부 구간을 제외하고는 본 모델을 적용할 수 있을 것으로 예상된다.

본 연구에서 제시한 모델은 차로수, 신호주기, 출/퇴근 시간의 영향도, 버스전용차선 미운영 등의 순천시와의 교통여건이

유사한 중소규모 도시에서 실시간 통행속도 예측하여 도착예정 시간을 계산하는데 MAF 방법보다 유용한 모델이 될 것이다.

본 논문에서 언급하지 않은 C 구간과 D 구간은 9시~17시까지의 통행량에 대한 이상점(Outlier)의 분포가 크고, 2분 위 4분위선의 폭이 좁게 퍼져 있는 것을 확인하였다. 이 두 개의 구간에 대해 평일과 주말에 따른 Box-Plot을 그려 확인하였고, DLM 및 MAF를 적용하여 예측해본 결과로 A, B 구간과 유사한 양상을 보이나 이상점(Outlier)의 분포가 많은 D 구간의 경우 일반화 되지 않는다.

Fig. 10과 Fig. 11은 C구간의 평일과 주말에 대한 데이터를 나타낸 것으로 C 구간은 잔차 정규성 및 정규분포를 잘 따르는 편이다. 그러나 Figs. 12와 13에 해당하는 D 구간은 Q-Q Plot을 확인한 결과 잔차 정규성이 매우 안좋은 결과를 보이고 있다.

전반적으로 순천시의 모든 구간에서는 본 논문에서 제시한 동적선형모델을 이용한 예측 방법이 잔차의 정규성 및 정규분포의 형태를 보이고 있다. 그러나 일부 구간에서는 수집주기의 표본추출 문제로 이상점이 많고, 평균속도의 편차가 크지 않고 평균속도보다 낮게 형성되어 있어 본 논문에서 구현된 DLM 모델을 적용하기 어렵다. 향후연구로는 일부 구간에 대한 추가적인 모델링이 필요하며, 동적 선형 모델의 정밀도 보정을 위한 상태 분리방법으로 은닉 마르코프 모델(Hidden Markov Model)을 적용하여 혼합 분포에 대응하는 모델링 방법이 필요하다.

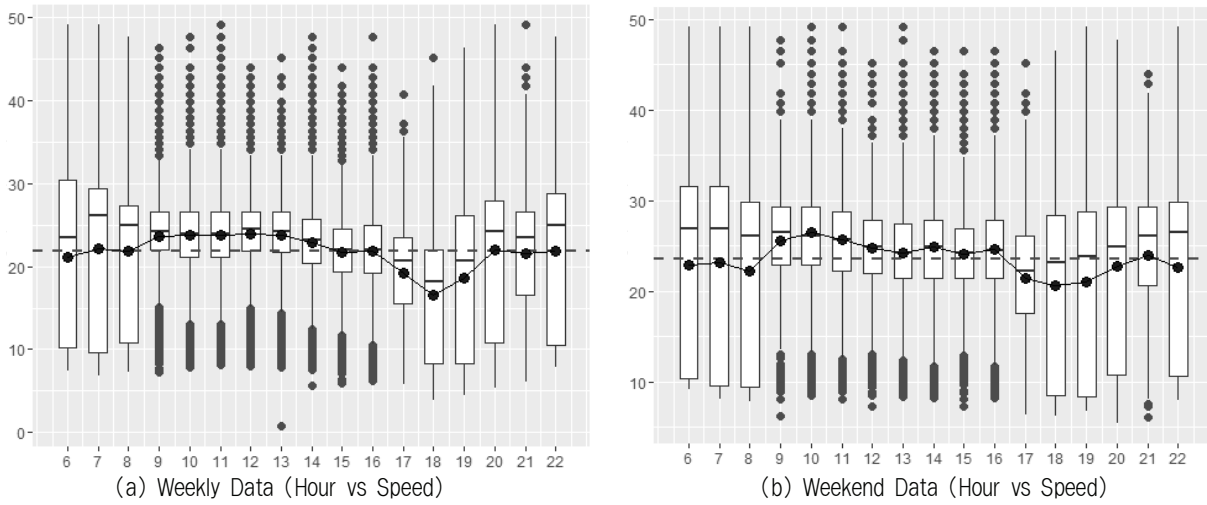


Fig. 10. Route C Based on Speed vs Hour (Year: 2015) for (a) and (b)

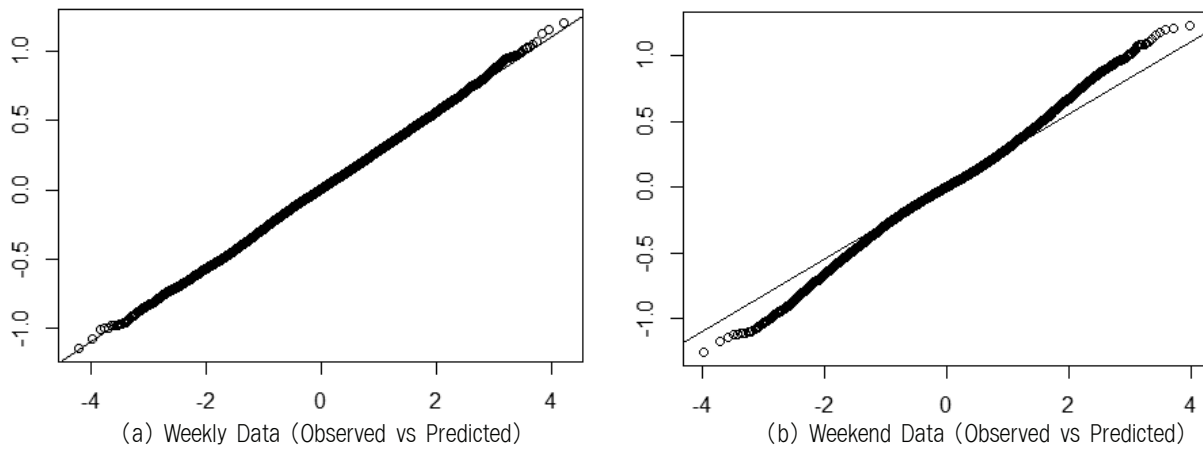


Fig. 11. Q-Q Plot of Route C (Year: 2015) for (a) and (b)

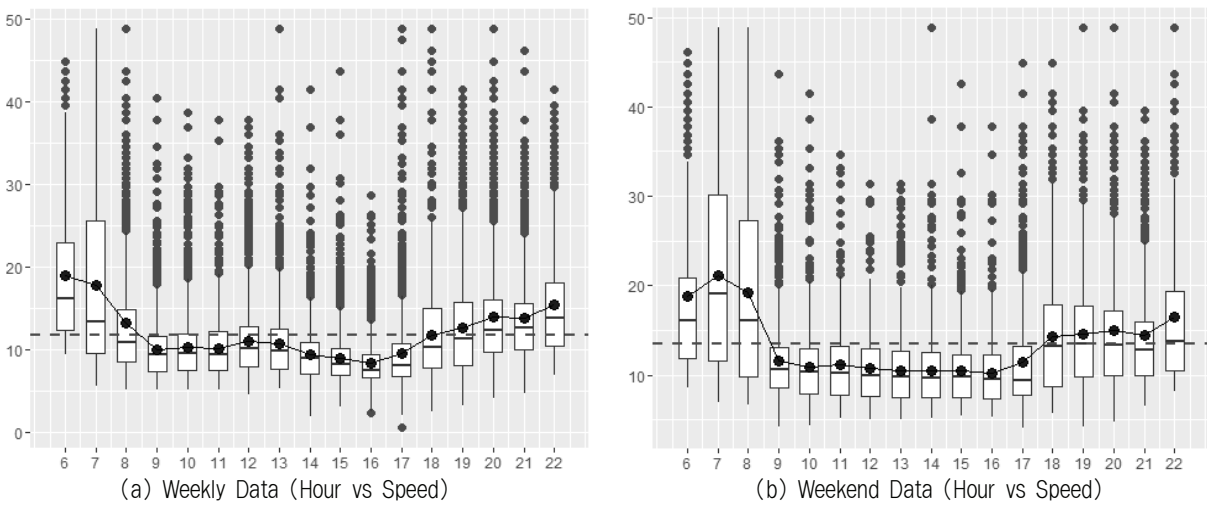


Fig. 12. Route D Based on Speed vs Hour (Year: 2015) for (a) and (b)



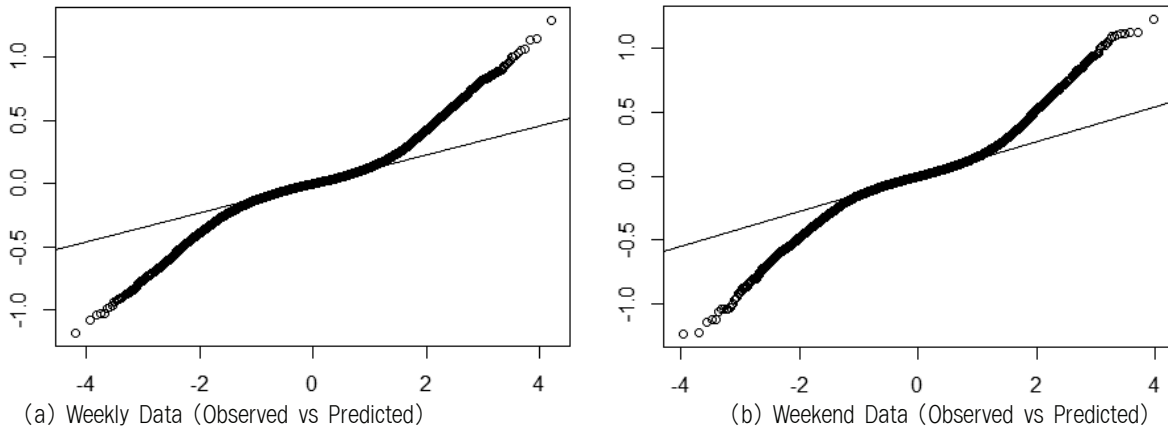


Fig. 13. Q-Q Plot of Route D (Year: 2015) for (a) and (b)

**References**

[1] H. G. Kim, C. Y. Park, D. C. Shin, C. S. Shin, Y. Y. Cho, and J. W. Park, "A Study on Traffic Analysis Using Bus Information System," *The KIPS Transactions on Computer and Communication Systems*, Vol.5. No.9, pp.261-267, 2016.

[2] Y. Y. Lee, "A Study on Estimate to Link Travel Time Using Traveling Data of Bus Information System," *Journal of Korean Society of Transportation*, Vol.30. No.3. pp.241-246, 2010.

[3] B. Portugais and M. Khanal, "State-Space Models With Kalman Filtering for Freeway Traffic Forecasting," *International Journal of Modern Engineering*, Vol.15. No.1, pp.11-14, 2014.

[4] G. Petris and S. Petrone, "State Space Models in R," *Journal of Statistical Software*, Vol.41. No.4. pp.1-25, 2011,

[5] G. Petris, S. Petrone, and P. Campagnoli, "Dynamic Linear Models with R," *Springer Science Business Media*, 2009.

[6] City Bus of Suncheon-si [Internet], [https://ko.wikipedia.org/wiki/%EC%88%9C%EC%B2%9C%EC%8B%9C%EC%9D%98\\_%EC%8B%9C%EB%82%B4%EB%B2%84%EC%8A%A4](https://ko.wikipedia.org/wiki/%EC%88%9C%EC%B2%9C%EC%8B%9C%EC%9D%98_%EC%8B%9C%EB%82%B4%EB%B2%84%EC%8A%A4).



**박철영**

e-mail : naksu21@gmail.com  
 2010년 순천대학교 정보통신공학과(공학사)  
 2012년 순천대학교 정보통신공학과  
 (공학석사)  
 2012년~현 재 순천대학교 전기·전자·  
 정보통신공학과 박사과정

관심분야: 기계학습, 시계열 분석, IoT



**신창선**

e-mail : csshin@sunchon.ac.kr  
 1996년 우석대학교 전산학과(학사)  
 1999년 한양대학교 컴퓨터교육과(석사)  
 2004년 원광대학교 컴퓨터공학과(공학박사)  
 2005년~현 재 순천대학교  
 정보통신공학과 부교수

2016년~현 재 순천대학교 정보전산원 원장

관심분야: 분산컴퓨팅, 실시간 객체모델, 시계열분석



**김홍근**

e-mail : khg\_david@sunchon.ac.kr  
 2011년 순천대학교 정보통신공학과  
 (공학사)  
 2013년 순천대학교 정보통신공학과  
 (공학석사)  
 2013년~현 재 순천대학교 전기·전자·  
 정보통신공학과 박사과정

관심분야: 기계학습, 시계열분석, IoT



**조용운**

e-mail : yycho@sunchon.ac.kr  
 1995년 인천대학교 전산학과(학사)  
 1998년 숭실대학교 컴퓨터학과(공학석사)  
 2006년 숭실대학교 컴퓨터학과(공학박사)  
 2009년~현 재 순천대학교  
 정보통신공학과 부교수

관심분야: 시스템 소프트웨어, 유비쿼터스 컴퓨팅, 기계학습



**박 장 우**

e-mail : jwpark@sunchon.ac.kr

1989년 한양대학교 전자공학과(공학사)

1991년 한양대학교 전자공학과(공학석사)

1993년 한양대학교 전자공학과(공학박사)

1995년~현 재 순천대학교

정보통신공학과 교수

관심분야: SoC, USN, 기계학습, 시계열 분석