

■ 論 文 ■

일반국도 상시 교통량자료를 이용한 교통량 결측자료 추정

Missing Data Imputation Using Permanent Traffic Counts on National Highways

하 정 아

(한국건설기술연구원 연구원)

박 재 화

(중앙대학교 컴퓨터공학부 교수)

김 성 현

(한국건설기술연구원 선임연구원)

목 차

- I. 서론
- II. 이론적 고찰
 - 1. 교통량 조사의 개요
 - 2. 분석방법
 - 3. 기존 연구 고찰
- III. 교통량 결측 자료 추정 방법
- IV. 모형검증 및 분석
 - 1. 모형검증 및 예측
 - 2. 분석결과
- V. 결론 및 향후 연구과제

Key Words : 상시조사, 자기회귀분석, 시계열분석, 교통량결측, 교통량보정
 Permanent Traffic Counts, Autoregression Analysis, Time Series Analysis, Missing Traffic Data, Traffic Data Imputation

요 약

일반국도의 상시조사 자료는 교통량 조사 장비를 통해 수집되며, 수집된 자료가 누락되거나 불량일 경우 통계자료의 객관성을 유지하기 위해서 보정을 해야 한다.

교통량 결측 자료의 보정을 통계적인 방법으로 접근하여 신뢰성을 높이고자 본 연구에서는 보정 대상 시간과 동일 시간의 자료를 적용할 수 있는 자기회귀분석과 보정 대상 지점과 동일 지점의 자료를 적용할 수 있는 계절 시계열 분석을 이용하여 보정하는 방안을 제시하였다.

계절 시계열 분석을 적용하여 결측 자료를 보정한 결과, 결측 기간이 길어질수록 오차가 커지는 것으로 분석되었다. 이것은 단기예측의 경우 실제자료를 이용하여 예측값을 제시하지만, 장기예측의 경우 예측된 자료를 이용하여 예측값을 제시하기 때문에 신뢰성이 떨어지기 때문이라 판단된다.

자기회귀분석을 적용하여 결측 자료를 보정한 결과, 시계열분석에 비해서 오차가 적은 것으로 분석되었다. 이것은 교통량 자료는 과거 패턴보다 현재 시점의 영향을 더 많이 받는 것이기 때문이라 판단된다. 하지만 자기회귀분석은 인근에 패턴이 유사한 지점이 있어야 가능하며, 인근에 유사한 지점이 있더라도 그 지점의 자료가 불량일 경우 보정이 불가능하다는 단점이 있다. 이러한 경우에는 과거자료를 이용해서 보정할 수밖에 없으며, 단기 결측의 경우에는 시계열분석을 이용할 수 있다

Up to now permanent traffic volumes have been counted by Automatic Vehicle Classification (AVC) on National Highways. When counted data have missing items or errors, the data must be revised to stay statistically reliable.

This study was carried out to estimate correct data based on autoregression and seasonal AutoRegressive Integrated Moving Average (ARIMA). As a result of verification through seasonal ARIMA, the longer the missed period is, the greater the error. Autoregression results in better verification results than seasonal ARIMA. Traffic data is affected by the present state more than past patterns.

However, autoregression can be applied only to the cases where data include similar neighborhood patterns and even in this case, the data cannot be corrected when data are missing due to low quality or errors. Therefore, these data should be corrected using past patterns and seasonal ARIMA when the missing data occurs in short periods.

I. 서론

도로 교통량 조사는 도로를 이용하는 각종 통행 차량의 통과 대수를 종류별, 방향별 및 시간대별로 관측하는 조사이다. 도로를 통과하는 단위 시간당의 교통량은 도로 시설물의 효율 척도로서 사용되며, 다른 지점과의 상대적 비교를 통하여 각 도로 구간의 역할을 추정 평가할 수 있는 지표로 사용된다. 또한 도로의 계획, 설계, 운영 등에 폭넓게 이용되며 교통 계획 수립을 위한 기초 자료로 활용된다. 따라서 도로 및 교통과 관련된 다양한 분야의 연구에 있어 그 활용 빈도가 높은 중요한 자료라 할 수 있다[1].

교통량 조사 장비를 이용해서 조사를 실시할 경우 도로공사, 천재지변, 교통사고 또는 장비의 고장으로 인하여 일정기간 자료가 누락되는 경우가 발생할 수 있다. 장비를 통해 수집된 자료 중에서 누락이 발생하였거나 불량 자료로 판단되는 경우에는 통계자료의 객관성을 유지하기 위해서 보정을 해야 한다.

현재 교통량 자료 관리에 있어서 결측 자료는 크게 단기 결측 자료의 보정과 장기 결측 자료의 보정으로 나누어 보정하고 있다. 일반적으로 단기 결측 자료의 보정에는 산술평균으로 보정하는 방법, 총량을 시간대별 비율로 분할하여 보정하는 방법, 상행과 하행 교통량의 비율을 이용하여 보정하는 방법 또는 시간대 교통량과 차종별 교통량의 차이를 이용하는 방법이 적용되며, 장기 결측 자료의 보정에는 일반적으로 증감율법이 적용된다.

하지만 이러한 방법은 적용이 용이하다는 장점은 있으나, 통계적으로 신뢰성은 부족하다고 할 수 있다. 이를 위해 본 연구에서는 보정 대상 시간과 동일 시간의 자료를 적용하는 방법과 보정 대상 지점과 동일 지점의 자료를 적용하는 방법을 제시하고자 한다. 이는 교통량의 특성상 동일 지점에서는 시간적인 추세에서 일정한 시간적 패턴을 형성함으로써 그 지점의 특성을 살릴 수 있다는 것과, 동일 시간에 결측 교통량이 발생한 지점과 교통특성이 유사한 지점의 교통량 패턴 역시 유사할 것이라는 것에 착안한 것이다. 이러한 점에 근거하여 교통량 결측 자료보정에 시계열분석과 자기회귀분석을 적용하고자 한다.

교통량 패턴을 분석하고 예측값을 제시하기 위해, 본 연구에서는 교통량 조사 장비를 통해 수집되는 일반국도 상시지점의 시간대별 자료를 이용하였다. 연평균 일교통량(Annual Average Daily Traffic: AADT)은 차종

분류조사나 투자개선계획 및 도로구조물 설계시에 대단히 중요한 자료이며 시간대별 교통량은 도로의 기하구조 설계에 이용되는 중요한 자료이다[11]. 그러므로 일 교통량의 정확한 값을 추정하는 것만큼 시간대별 교통량의 정확한 값을 추정하는 것도 중요하다고 할 수 있다.

본 연구에서는 통계적인 방법으로 교통량 결측 자료에 대한 보정방법을 제시하여 단순한 자료 보정방법에서 벗어나 보다 더 신뢰성이 높은 자료 보정 방법론을 제시하고자 한다. 이러한 방법은 교통량 자료의 보정에 있어서 좀 더 객관적인 방법을 제시하며, 추정값의 오차를 줄여 교통량 자료의 정확성을 높일 것으로 기대된다.

II. 이론적 고찰

1. 교통량 조사의 개요

도로 교통량은 도로 시설물의 효율 척도로서 사용되며, 다른 지점과의 상대적인 비교를 통하여 각 도로 구간의 역할을 추정 평가할 수 있는 기준으로 사용된다. 일반국도의 교통량 조사는 크게 수시조사와 상시조사로 구분된다. 수시조사는 기본 교통량 자료가 필요하다고 판단되는 모든 구간에 대하여 광범위하게 실시하는 조사로서, 전체적인 도로 이용 상황을 파악하기 위한 조사이다. 상시조사는 특정 지점에 고정식 조사 장비를 설치하여, 1년 이상의 장기간에 걸쳐 그 특정지점을 통과하는 차량수를 측정하고 기록하는 조사이다. 수시조사는 이동식 조사 장비를 이용하여 지점별로 연 1회~5회 실시하며, 상시조사는 고정식 조사장비를 이용하여 1년 365일 교통량을 모두 측정한다. 모든 조사지점에 고정식 조사장비를 설치하여 1년 365일 동안 교통량을 조사하는 것이 이상적이나, 예산 등의 문제로 주요 지점에만 고정식 조사장비로 교통량을 조사하고, 나머지 지점에 대해서는 이동식 조사 장비를 이용하여 수시조사를 시행하고 있다. 수시조사는 매년 같은 날짜나 같은 요일의 교통량을 조사하는 것이 아니라 365일 중 주말과 공휴일, 휴가철을 제외한 평일 교통량을 연 1회~5회 걸쳐 조사하는 것이므로, 시간대별 패턴이나 일 교통량의 패턴을 찾는 일이란 불가능하다. 따라서 본 연구에서는 수시조사 교통량에 대한 내용은 분석 대상에서 제외하고, 상시 조사 지점의 교통량을 활용하여 결측 자료를 보정하는 방법에 관한 연구를 수행한다.

2. 분석방법

1) 계절 시계열 모형

시계열분석은 시간의 흐름에 따라 변하는 현상을 관측함으로써 얻어지는 자료를 분석하며 이를 설명할 수 있는 모형을 설정하고 미래를 예측하는 분석을 말한다. 여러 통계분석에서는 일변량 계열에 내재하는 관측값이 통계적으로 독립이라고 가정한다. 하지만 일변량 ARIMA 분석에서는 '자료계열에서 시간의 흐름에 따라 얻어진 관측값은 통계적으로 종속적이다'라고 가정한다. 즉, '계열 내 관측값 사이에 존재하는 상관관계'라는 통계적 개념을 이용한다. 일변량 ARIMA 분석에서는 시점 t 에 해당하는 관측값(y_t)과 t 의 앞 시점들의 관측값(\dots, y_{t-1}) 사이의 상관관계를 알고자 한다(8).

일변량 ARIMA 모형은 특히 단기예측에 적합하다. 왜냐하면 ARIMA 모형은 먼 과거보다는 최근 시점에 가까운 과거 관측값에 더 많은 비중을 주기 때문이다. 최근 시점에 가까운 과거에 비중을 더 준다는 것은 ARIMA 모형에 근거하여 얻어진 장기 예측값이 단기 예측값에 비해 신뢰성이 적다는 의미이다.

2) 자기회귀모형

회귀분석(Regression Analysis)이란 어떤 현상이 변수들의 인과관계에 의하여 나타날 때 그 관계를 수학적으로 설명하기 위하여 사용되는 통계적 방법을 말한다. 회귀분석에서 사용되는 변수는 종속변수와 독립변수가 있다. 종속변수(dependent variable)는 서로 관계를 가지고 있는 변수들 중에서 다른 변수에 의해 영향을 받는 변수를 말하며, 독립변수(independant variable)는 종속변수에 영향을 주는 변수를 말한다. 회귀분석은 이러한 종속변수와 독립변수들 간의 함수관계를 규명하기 위하여 어떤 수학적 모형을 설정하고, 주어진 자료로부터 모형을 추정하거나 검증하는 방법이다.

회귀분석에서 가장 간단한 모형은 단순선형회귀분석(simple linear regression analysis)이며 다음과 같이 표현된다.

$$Z_t = \beta_0 + \beta_1 x_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2) \text{ 이고}$$

서로 독립, $t = 1, 2, \dots, n$

여기서 ϵ_t 는 t 번째 측정치의 오차(error)이고, β_0 와

β_1 은 미지의 회귀계수(regression coefficient)로서, 이것은 최소제곱법(method of least squares)로 추정한다. 오차항에 대해서는 위의 식에서 보는 바와 같이 정규성(normality), 불편성(unbiasedness), 등분산성(equal variance), 독립성(independence)을 일반적으로 가정하고 있다.

만약 주어진 자료들이 시계열적 특성을 띠고 있어 오차항들 사이에 자기상관이 존재한다면 회귀분석에서 가정하는 독립성을 위배하게 되므로 추정된 회귀식의 수정이 필요하다. 오차항들 사이에 자기상관이 존재하는 것은 Durbin-Watson 통계량(d)를 이용하여 검정할 수 있다. d통계량을 이용하여 자기회귀 여부를 검정결과 자기상관이 존재한다면 주어진 자료의 오차항은 시계열모형을 따르는 확률변수로 생각할 수 있고, 이 확률변수를 ϵ_t (t :시차)라고 한다면 위의 회귀식을

$$Z_t = \beta_0 + \beta_1 X_t + \epsilon_t$$

여기서, $\epsilon_t = \rho \epsilon_{t-1} + \delta_t, \delta_t \sim N(0, \sigma_\delta^2)$,

$$Cov(\delta_t, \delta_{t+i}) = 0$$

으로 바꾸어 생각할 수 있다. 이와 같은 모형을 자기회귀모형이라 한다.

3) EM 알고리즘

EM 알고리즘(Expectation-Maximization Algorithm)은 E(=Expectation) 단계와 M(=Maximization) 단계로 구성된다. 총자료 X 를 관측부분 X_{obs} 와 결측부분 X_{mis} 로 구분하여 표기할 때, E단계에서는 관측 X_{obs} 에 조건화하여 결측 X_{mis} 에 대한 추정값을 구하여 X_{mis} 를 대체한다. (파라메타 θ 의 잠정 추정값 θ_0 을 사용하여) 그리고 M단계에서는 X_{obs} 와 X_{mis} 를 모두 써서 θ 에 대한 우도(likelihood), 즉 $L(\theta X_{obs}, X_{mis})$ 를 최대화한다. 여기서 구한 θ 값을 θ_0 으로 놓고 다시 E단계로 돌아간다. 이와 같이 E 단계와 M단계를 반복함으로써 X_{obs} 에서의 파라메타 θ 에 대한 가능도 $L(\theta X_{obs})$ 를 최대화 하는 θ 값을 찾는다(4).

3. 기존 연구 고찰

교통량 자료의 불량 자료를 보정하기 이전에, 전년도

교통량을 이용하여 도로유형을 파악하고, 도로유형에 따라 지점이 어떤 특성을 가질 것이라는 것을 먼저 판단하고 있다면 보다 더 합리적이고 객관적인 예측값을 제시할 수 있다. 또한 전년도 교통량이 없더라도, 지점의 주변 환경이나 정상자료의 교통량으로도 도로유형을 판단하여 예측을 할 수 있다.

Williams(1997) 외 2명은 도시부 도로의 단기간 교통량의 예측을 위해 계절 시계열 모형을 적용하였다(14). 시계열 모형은 어떤 특정 기간의 순환모형으로 특성화되며, 시계열모형의 적용을 위해서 버지니아 북부의 두 개 지점을 대상으로 하였다. 교통량 자료는 15분 자료이며 계절주기는 96(1일 기준)으로 분석하였다. 분석결과 계절 ARIMA모형이 인근 지점의 교통량을 이용하는 방법과 신경망 분석, 과거 교통량의 평균을 이용하는 방법보다 성능이 우수하다고 분석되었다.

국내 문헌 중 교통량 예측 자료의 대체 기법으로 계절 시계열 모형을 이용한 사례가 있었다(4). 이것은 기존에 사용되던 전후일 평균방법과 전후주 평균, 회귀모형, EM 알고리즘과 시계열을 비교하여 분석하였고, 시계열 모형이 오차가 가장 적은 것으로 분석하였다. 하지만 이 분석은 일반국도 3호선 상에 위치한 단일지점에 대한 교통량에 대해 교통량이 가장 안정적이라고 판단되는 수요 일만의 교통량을 이용하여 공휴일이나 휴가철 등의 교통량에 대해서는 설명할 수 없다는 것에 한계가 있다.

sharma(2003)의 연구에서는 교통량의 결측데이터를 추정하기 위해 ARIMA 모형과 유전자 회귀모형, 신경망 모형을 이용하여 캐나다 앨버타주의 6개 상시교통량 조사지점의 데이터를 활용하였다(15). 이모형은 AADT와 DHV를 추정하기 위해 적용되었으며, 유전자 회귀모형이 결측 교통량 추정에서 가장 효과적인 것으로 제시하였으며, ARIMA 모형은 신경망 모형보다 우수한 것으로 판단되었다. 모형의 추정에 대한 정확성은 도로의 종류 및 기능 분류에 의해 영향을 받으며, 교통량 패턴이 안정적이거나 도로의 기능 분류가 보다 체계적일 때 AADT와 DHV의 추정이 더욱 정확해 지는 것으로 나타났다.

Yiannis Kamarianakis(2003) 외 1명은 그리스의 아테네에 있는 도시부의 실시간 자료를 이용하여 단기간의 교통류를 예측하기 위한 방법으로 일변량(univariate) 분석과 다변량(multivariate) 분석으로 분류하여 분석하였다(13). 일변량 분석은 각각의 검지기에서 독립된 모형으로 제시되며, 다변량분석은 전체 대상구간의 검지기에서 하나의 모형이 제시된다. 일변량 분석 모형은 비교적 적용하기가 쉽고 다른 시간대의 특정 교통특성을 적용할 수 있

으며, 다변량 분석은 인근에 위치하거나 교통류의 패턴이 유사한 지역의 자료를 결합함으로써 정확성을 기대할 수 있다. 자료의 정확성 비교를 위해서 일변량 분석 방법으로는 과거 자료를 평균한 값과, ARIMA를 이용한 모형을 제시하였으며, 다변량 분석 모형으로는 VARMA(Vector Autoregressive Moving Average)와 STARIMA(Space Time Autoregressive Integrated Moving Average)를 이용한 모형을 제시하였다. 4개의 모형으로 실제 검지기 자료를 이용하여 상대속도를 예측하였으며, 비교분석 결과 과거 자료를 평균한 값이 가장 나쁜 예측력을 보였다. 이는 단기간의 자료로는 교통류를 예측하는 것이 어렵기 때문인 것으로 판단된다. 또한 특정 지점에 돌발상황이 발생하였을 때의 교통류 속도에 대한 예측에서는 일변량 모형에서는 예측이 어려운 반면 다변량 모형에서는 인근 지점과 다른 지점의 상황을 반영할 수 있도록 모형화되기 때문에 예측이 가능하다고 분석되었다.

실시간 교통량 자료의 결측값 보정방법으로 과거 5주 교통량의 평균을 이용하는 방법과 회귀분석, 자기회귀분석 방법을 비교분석한 예가 있다(3). 이 분석에서 이용한 자료는 ITS(Intelligent Transport System) 영상검지기로부터 얻은 자료이며, 5분 단위로 수집된 자료이다. 실시간으로 수집이 되는 1분, 5분 교통량 자료는 변동이 매우 심하여 그 추이를 예측하기에 오차가 아주 크다고 알려져 있다. 5분 교통량 자료로 검지기들 간에 상관분석을 하여 유사검지기를 찾은 결과, 상관성이 높은 검지기는 적은 것으로 나타났다. 따라서 유사 검지기를 선택하기 위하여 특정시점이 아니라 하루의 평균적인 교통 변화 패턴이 유사한 인근 검지기를 주중, 토요일, 일요일 별로 선택하여 이 유사 검지기의 정상 자료를 이용한 보정을 제안하였다. 유사 검지기의 정상자료를 이용한 보정으로 자기회귀분석을 제시하였으며, 독립변수는 유사 검지기의 직전 5주의 주중, 토요일 5분 과거 자료를 이용하였다. 분석결과 유사 검지기와 결측 검지기의 상관계수가 높을 경우에는 자기회귀분석이 효과적이며, 그렇지 않을 경우 직전 5주의 과거자료를 이용하여 주중, 토, 일요일 별로 평균으로 보정하는 것이 좋다는 결론을 얻었다. 이 연구는 실시간 ITS 영상검지기 자료를 이용한 것으로 online자료에 해당된다. 본 연구에서 분석하고자 하는 자료는 일반국도 상시조사 장비에서 얻은 자료로 offline 자료에 해당되므로 ITS 실시간 자료의 분석결과와 비교하였을 때 다른 결과를 가져오리라 판단된다.

본 연구에서는 기존 연구방법에서 좋은 예측력을 나타낸

시계열 모형과, 실시간 ITS 영상검지기의 결측자료를 보정할 때 이용된 자기회귀분석을 비교분석하여 일반국도 상시조사 자료의 결측자료를 보정하는 모형을 제시하고자 하였다.

III. 교통량 결측 자료 추정 방법

일반국도 상시 교통량 자료는 장비를 통해서 수집되며, 실시간 자료가 1시간 단위로 데이터베이스에 저장된다. 시간대별로 조사된 교통량은 일 교통량으로 환산되며, 일교통량으로 AADT가 산출된다. 시간대별 교통량과 AADT는 도로의 계획 및 설계 등에 폭넓게 이용되므로, 정확한 값의 산출이 요구된다.

교통량 자료를 보정하기 이전에 수집된 자료가 정상인지 불량인지를 먼저 판단해야 한다. 자료가 불량이라고 판단되는 지점에 대해서는 보정을 해야하며, 예측값에 대해서 오차가 클 것이라 판단되는 지점에 대해서는 보정이 불가능하다고 판단을 내려야 한다. 본 연구에서는 자료의 보정을 위해 계절 시계열 모형과 자기회귀모형을 이용하여 비교·분석하였다. 보정대상 지점은 일반국도 상시조사 지점 중 2004년과 2005년에 교통량자료가 모두 있고, 자기회귀보정이 가능하다고 판단되는 지점을 선택했으며 총 245지점이다.

교통량 자료를 보정하는 방법은 크게 두 가지로 나눌 수 있다. 첫 번째 방법으로는 결측이 발생한 지점과 시간대별 교통량 패턴이 비슷한 검지기가 있을 경우, 그 검지기의 자료를 참고하여 보정하는 방법이다. 이 방법은 보정시점과 동일한 시점의 교통량을 참고하여 보정하는 방법이라는 점에 있어서 그 당시의 교통 상황들을 고려할 수 있다는 장점이 있다. 하지만 이 방법은 보정 대상지점의 교통특성을 인근 유사검지기의 자료로 추정해야 한다는 단점을 가지고 있다. 두 번째 방법은 과거 교통량의 패턴을 이용해서 보정하는 방법이다. 이 방법은 보정하고자 하는 지점의 교통량을 참고하기 때문에 보정 대상 지점의 특성은 살릴 수 있지만, 보정 시점의 교통 상황들을 반영할 수 없다는 단점이 있다.

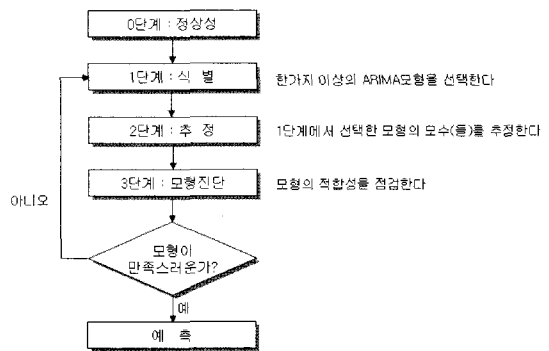
1. 계절 시계열 모형을 이용한 자료보정

일반국도 상시조사 장비로부터 얻어진 자료는 1시간

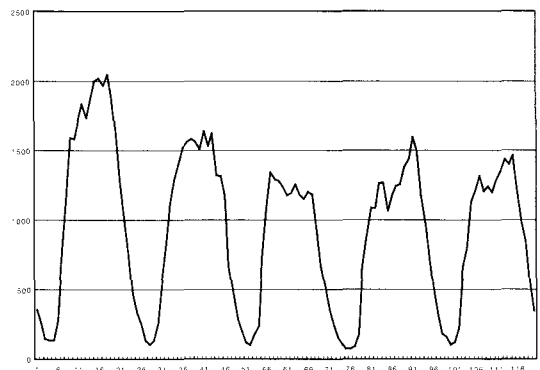
단위로 조사되며, 시간의 흐름에 따라 변하는 현상을 관측하므로 시계열자료라고 볼 수 있다. 시계열분석은 일반적으로 Box-Jenkins가 제안한 3단계 절차를 이용하며, 그 단계는 <그림 1>과 같다.

시계열자료를 분석하기에 앞서 먼저 순차도표를 그려 보고 자료의 흐름을 판단해야 한다. <그림 2>은 일반국도 44호선 강원도 홍천군 홍천읍 갈마곡리에 위치한 상시지점의 시간대별 교통량의 일부이다¹⁾. 이 지점은 도시부로 분석되었으며, 2003년 AADT는 21,343대이고 변동계수(Coefficient of Variance)는 28.1%로 나타났다.

<그림 2>를 보면 알 수 있듯이 시간대별 교통량은 24시간을 주기로 비슷한 패턴을 갖는다. 교통량과 같은 시계열 자료는 일정한 시간 간격을 두고 동일한 현상이 반복되는 경향을 가지고 있다. 이와 같이 반복적인 현상이 계속 일어나는 시간간격을 계절주기라고 하며, 이러한 특징을 지닌 시계열을 계절 시계열(seasonal time-series)이라고 한



<그림 1> ARIMA 모형 설정을 위해 필요한 절차



<그림 2> 지점의 시간대별 교통량 예

1) 요인분석을 활용한 일반국도 유형분류(한국도로학회 논문집, 2005. 09)에서 일반국도 상시조사 지점의 교통량으로 분석한 결과 3가지 그룹으로 분류되었으며, 교통지표로 보았을 때 3개의 그룹은 관광부, 지방부, 도시부로 판단되었다.

다. 관측된 시계열에 대하여 계절통합혼합모형을 적합시켜서 얻은 잔차들이 ARIMA 모형을 따를 때, 이 모형을 계절통합모형에 승법으로 결합한 모형을 일반화 승법 계절통합혼합모형이라고 하며 이를 식으로 나타내면 다음과 같다.

관측된 시계열 Z_t 에 대하여 계절통합혼합모형을 적합시키면 $\phi(B)(1-B)^d Z_t = \theta(B)e_t$

이고, 여기서 잔차 e_t 의 모형이 ARIMA(p,d,q)이므로 모형 식은

$$\phi(B)(1-B)^d e_t = \theta(B)a_t, \quad a_t \text{는 백색잡음}^2)$$

이다. 두 모형을 승법으로 결합하면

$$\phi(B)\Phi(B)(1-B)^d(1-B)^D Z_t = \theta(B)\Theta(B)a_t$$

2004년 상시 교통량 자료로 승법 계절통합혼합모형식에 적합시킨 결과, 전체 지점 모두 계절차분을 해야 정상성을 만족시킬 수 있었다.

지점별로 시간대별 교통량의 추정식을 구하고, 추정식으로부터 산출된 추정값과 참값의 오차를 MAPE로 비교한 결과는 <표 1>과 같다.

<표 1> 시계열분석 추정오차의 범주별 빈도수

MAPE(%) 범위	해당 지점수
0~10	65
10~20	157
20~30	20
30~40	2
40 이상	1
계	245

2. 자기회귀모형을 이용한 자료보정

계절 시계열 모형처럼 지점의 과거 교통량을 이용하여 보정할 경우 해당지점의 특성은 살릴 수 있지만, 보정시점의 주변 환경이나 날씨 등의 영향을 전혀 감안할 수 없다는 점에 있어서 오차를 가져올 수 있다. 자기회귀모형은 이러한 단점을 보완하여 동일시점에서 인근에 교통패턴이 유사한 지점의 교통량을 참고하여 보정하는 방법이다. 자기회귀모형을 이용하여 보정하는 절차는 <그림 3>과 같다.

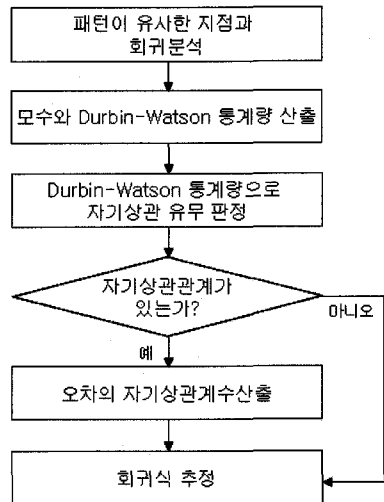
자기회귀모형을 이용하여 보정하기 위해서는 우선 패턴이 유사한 지점을 찾아야 한다. 패턴이 유사한 지점을 찾는 방법으로는 회귀분석의 선행 작업으로 자주 이용되는 상관분석을 이용했다. 동일노선 내에 보정대상 지점과 상관계수가 높은 지점이 많을 경우 인근지점을 선택하였으며, 인근 지점이 모두 선택되면 인근지점 중 상관계수가 더 높은 지점을 우선적으로 선택하였다.

2004년 일반국도 상시조사 지점 353개 중 유사한 패턴을 가지는 지점을 분석한 결과, 312개 지점이 인근에 패턴이 유사한 지점이 존재하는 것으로 나타났다.

2004년 자료로 회귀분석 결과, Durbin-Watson 통계량의 분포는 <그림 4>와 같다.

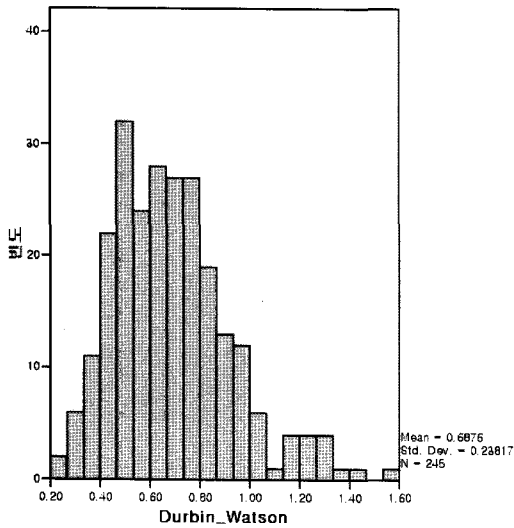
Durbin-Watson의 통계량(d)의 최대값은 1.55였으며, 전 지점에서 d값이 0.2에서 1.6 사이에 존재함을 알 수 있었다. 즉, 시간대별로 조사된 교통량자료의 회귀분석 결과 전 지점에서 오차항들의 자기상관이 존재한다고 분석되었다. 오차항들의 자기상관은 흔히 시계열 자료에서 나타나는 현상이며, 이 경우 오차항의 추정식도 회귀분석 모형에 포함시켜야 한다.

회귀분석과 자기회귀분석으로 교통량 자료를 추정한 결과, 대부분의 지점에서 자기회귀분석으로 추정된 추정값의 오차가 회귀분석으로 추정된 추정값의 오차보다 더 적게 나타났다.



<그림 3> 자기회귀모형을 이용한 보정절차

2) 백색잡음(White noise)이란 ARIMA 프로세서 내에 있는 e_t 항들 모두가 평균이 0이며 분산이 상수인 정규분포를 하는 확률변수이며 서로 독립인 변수들을 말한다.



〈그림 4〉 Durbin-Watson 통계량 분포도

지점별로 자기회귀분석을 이용하여 시간대별 교통량의 추정식을 구하고, 추정식으로부터 산출된 추정값의 오차를 MAPE로 비교한 결과는 〈표 2〉와 같다.

〈표 2〉 자기회귀분석 추정오차의 범주별 빈도수

MAPE(%) 범위	해당 지점수
0~10	60
10~20	132
20~30	37
30~40	11
40 이상	5
계	245

MAPE 비교결과 오차의 범주로 봤을 때, 시계열분석의 추정값과 회귀분석의 추정값의 오차는 크게 차이나지 않았으며, 지점별로 MAPE를 비교해 보았을 때, 자기회귀분석보다 시계열분석이 MAPE가 적은 지점들이 더 많은 것으로 분석되었다.

IV. 모형검증 및 분석

1. 모형검증

3장에서는 2004년의 지점별·시간대별 교통량 자료를 이용하여 식을 추정하고, 추정 값을 구하여 오차를 비교·분석하였다. 본 절에서는 분석 대상 자료의 2005년 10월~12월 교통량을 결측이라고 가정하고, 3장에서 제

시한 방법으로 추정 값을 구하여 오차를 검증하였다.

1) 계절 시계열 모형

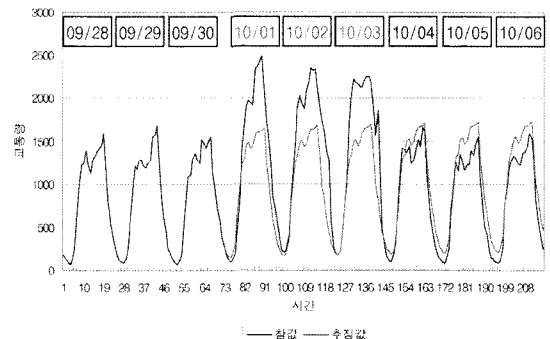
3장에서 추정한 2004년 시간대별 교통량의 추정식과는 별개로 본 절에서는 2005년도 1월부터 9월달 자료로 시계열분석을 하고, 분석결과 얻어진 시계열 식으로 10월~12월의 교통량을 추정하였다.

〈그림 5〉는 일반국도 44호선 강원도 홍천군 홍천읍 갈마곡리에 위치한 상시지점의 2005년 교통량의 실제값과 계절 시계열 모형으로 추정한 추정값을 비교한 그림이다.

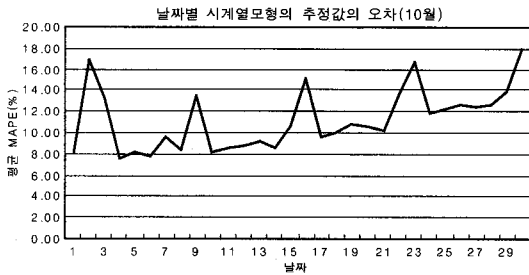
〈그림 5〉에서 짙은색 실선으로 표현된 부분은 참값이며, 옅은색 실선으로 표현된 부분은 시계열 모형으로 추정된 추정값이다. 이 지점의 2005년 1월부터 9월까지 시간대별 교통량으로 추정된 시계열 모델은 ARIMA(1,1,0)(1,1,1)₂₄로 분석되었다. ARIMA(1,1,0)(1,1,1)₂₄의 일반적인 추정식은 다음과 같다.

$$(1 - \phi_1 B)(1 - B)(1 - \phi_1 B^{24})(1 - B^{24})Z_t = (1 - \theta_1 B^{24})a_t$$

ARIMA(1,1,0)(1,1,1)₂₄의 후진연산자의 최고차수는 50이며, 이것은 보정시점의 시간보다 최고 50시간 전의 교통량을 참고하여 추정값을 산출한다고 해석할 수 있다. 따라서 10월 1일의 1시 교통량은 9월 28일 23시 교통량에서 10월 1일 0시 교통량 사이의 시간대별 교통량을 이용하여 10월 1일의 1시 교통량을 구한다. 만약 10월 4일의 7시 교통량을 ARIMA 모형으로 추정한다면, 10월 2일 5시 교통량에서 10월 4일 6시 교통량을 이용하여 추정한다. 이 경우 10월 2일 5시 교통량은 참값이 아닌 ARIMA 모형으로 추정한 교통량이므로, 10월 4일 교통량을 추정할 때에는 참값이 아닌 추정값으로



〈그림 5〉 시계열 모형의 참값과 추정값 비교의 예 (시간대별 교통량)



〈그림 6〉 날짜별 시계열 모형의 예측값의 오차(MAPE)(10월)

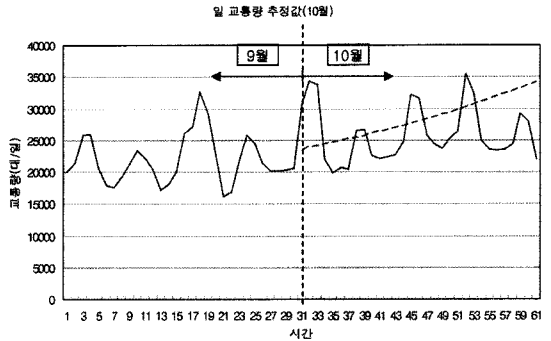
또다른 추정값을 제시하게 되어 신뢰성을 잃게 된다. 〈그림 5〉에서 나타나듯이 추정값은 과거 교통량을 토대로 거의 일정한 수준을 유지함을 알 수 있다. 2005년 10월 1일은 토요일, 10월 2일은 일요일, 10월 3일은 개천절로 공휴일이어서 10월 1일부터 3일까지는 연휴이므로 평소의 교통량과 교통패턴이 다를 것이라 예상된다. 3) 실제로 〈그림 5〉를 보면 추정된 교통량은 실제 교통량과 차이가 있음을 알 수 있다.

〈그림 6〉은 보정 대상지점을 대상으로 실제 교통량과 추정 교통량 간의 평균 MAPE를 그래프로 나타낸 것이다.

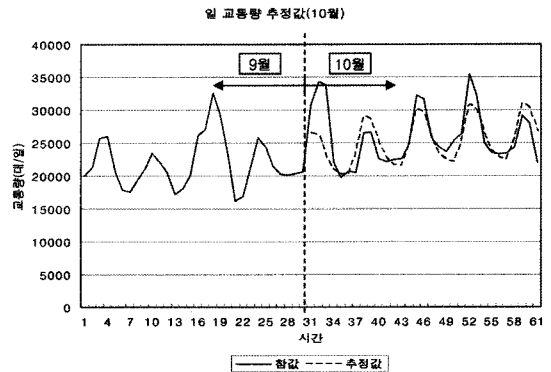
〈그림 6〉을 보면 알 수 있듯이 시계열 모형으로 추정된 추정값이 공휴일이나 주말의 경우 오차가 큰 것으로 나타났으며, 예측하고자 하는 날짜가 길어질수록 오차가 계속 커지고 있음을 알 수 있다. 〈그림 7〉은 국도 44호선 지점의 일 교통량의 참값과 시간대별 교통량으로 추정된 일 교통량의 추정값을 그래프로 그린 것이다.

일 교통량이 요일별로 패턴을 가지고 있음에도 불구하고, 시간대별 교통량으로 추정한 일 교통량은 패턴을 가지지 않고 꾸준히 증가하는 추세를 보이고 있다. 시간대별 교통량을 이용하여 시간대별 교통량을 추정하고, 추정된 시간대별 교통량을 이용하여 일 교통량을 추정할 경우, 단기 예측값의 경우에도 공휴일이나 주말의 교통량은 예측오차가 작지 않음을 알 수 있었다. 이는 교통량의 시간대별 패턴만 고려하고 요일별 교통량의 패턴은 고려하지 않았기 때문이라고 판단되며, 요일별 교통량 패턴을 감안하여 분석을 하기 위해서는 시간대별 교통량으로 시계열분석을 하는 것보다 일 교통량을 이용하여 시계열분석을 해야 한다고 판단된다.

〈그림 8〉은 위 지점의 일 교통량을 이용하여 시계열 분



〈그림 7〉 시계열 모형의 참값과 추정값 비교의 예 I (일교통량)



〈그림 8〉 시계열 모형의 참값과 추정값 비교의 예 II (일교통량)

석을 실시한 결과이다. 실선으로 표현된 자료는 참값이며, 점선으로 표현된 부분은 시계열 분석을 이용한 추정값이다.

이 지점의 2005년 1월부터 9월까지 시간대별 교통량으로 추정된 시계열 모형은 ARIMA(1,0,0)(1,1,1)₇로 분석되었다. ARIMA(1,0,0)(1,1,1)₇의 일반적인 추정식은 다음과 같다.

$$(1 - \phi_1 B)(1 - \phi_1 B^7)(1 - B^7)Z_t = (1 - \theta_1 B^7)a_t$$

ARIMA(1,0,0)(1,1,1)₇의 후진연산자의 최고차수는 15이며, 이것은 보정시점의 시간보다 최고 15일 전의 교통량을 참고하여 추정값을 산출한다고 해석할 수 있다. 따라서 10월 1일 교통량은 9월 16일의 교통량부터 9월 30일의 교통량을 이용하여 10월 1일의 교통량을

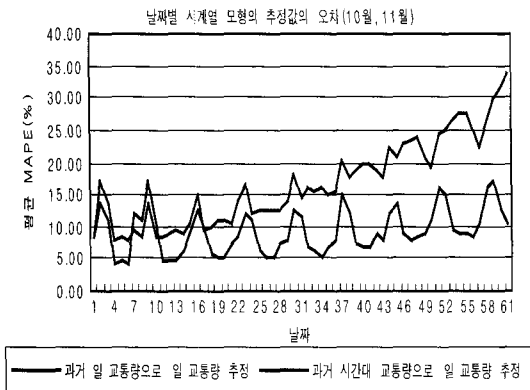
3) 통행시간 추정과 통계자료 보정을 위한 결측 데이터 보정 모형 제작(한국건설기술연구원 2001. 11)(3)에서는 요일별 연평균 하루 교통량 변화패턴의 군집분석 결과 주중요일, 토요일, 일요일, 공휴일로 유의한 차이가 있다고 분석되었다. 또한 공휴일 데이터는 일요일 교통량과도 유의한 차이를 보인다고 판단되지만, 넓은 의미로는 일요일의 교통량 패턴 변화에 준하여 분석해도 된다는 결론을 얻었다.

구한다. 시간대별 교통량 추정과 마찬가지로 10월 16일의 교통량을 추정하기 위해서는 10월 1일의 교통량부터 10월 15일의 교통량을 이용하므로, 역시 추정값으로 또 다시 추정을 하게 되어 신뢰성이 떨어진다.

〈그림 9〉는 시간대별 교통량으로, 일 교통량을 추정한 값의 오차(MAPE)와 일 교통량으로 추정값을 구한 값의 오차(MAPE)를 비교한 것이다. 일 교통량으로 추정된 것도 시간대별 교통량으로 추정된 값과 마찬가지로 시간이 지날수록 오차가 커짐을 알 수 있다. 하지만 시간대별 교통량으로 일 교통량을 추정한 값보다는 오차가 훨씬 줄어들었음을 알 수 있다.

일변량 ARIMA 모형은 먼 과거보다는 최근 시점에 가까운 과거 관측값에 더 많은 비중을 두기 때문에 장기 예측보다 단기 예측에 더 적합하다. 단기 예측과 장기 예측에 대한 정의는 데이터의 수준과 성격에 따라 달라질 수 있고, 결측기간이 짧아 참값으로 시계열 자료를 추정할 경우 단기결측이라 할 수 있고, 결측기간이 길어 추정된 값을 이용하여 시계열 자료의 예측값을 제시하는 경우 장기결측이라고 볼 수 있다.

최근 시점에 가까운 과거에 비중을 더 준다는 것은 ARIMA 모형에 근거하여 얻어진 장기 예측값이 단기 예측값에 비해 신뢰성이 적다는 의미이다. 만약 결측 기간이 단기간이라면 시간대별 교통량을 이용하여 예측을 해도 무방하지만, 결측 기간이 길어지면 시간대별 교통량을 이용하여 예측하는 것은 신뢰성이 떨어질 수 있다. 결측이나 불량인 장기간일 경우에는 시간대별 교통량을 이용하여 예측하는 것보다 일 교통량을 이용하여 예측하는 것이 신뢰도를 높이는 한 방법일 수 있다. 하지만 결측이나 불량인 기간이 너무 길어진다면, 일 교통량을 이용하여 예측하는 것도 예측된 자료로 예측을 해야 하므로 오차를 수반할 수밖에 없다.



〈그림 9〉 일 교통량의 예측값의 오차비교(MAPE) (10월, 11월)

2) 자기회귀모형

자기회귀모형으로 분석하기 위해서는 먼저 상관분석을 통해서 교통량 패턴이 유사한 검지기를 찾아야 한다. 2005년도 자료를 예측하기 위해서 2004년 자료를 이용하여 교통량 패턴이 유사한 검지기를 찾아야 하며, 2005년 상시조사 지점 중 총 245개의 검지기가 유사검지기가 있는 것으로 분석되었다.

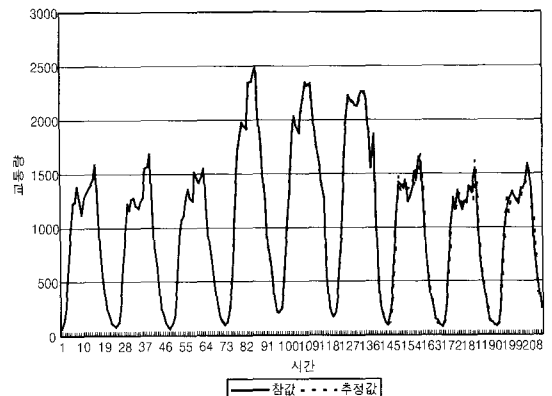
〈그림 10〉은 일부 자료를 결측이라 가정하고 자기회귀분석으로 추정값을 구하여 참값과 비교한 그래프이다.

그래프로 살펴보면 시계열 자료와는 달리 시간대별 교통량의 추정값이 참값과 크게 차이가 나지 않는 것으로 판단된다.

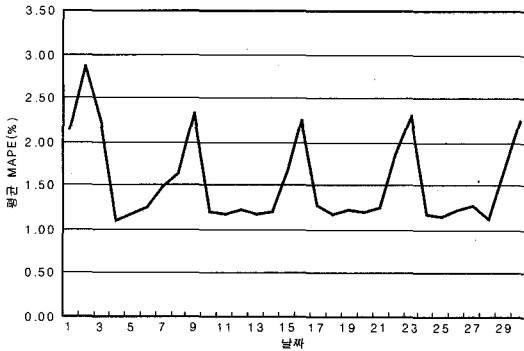
〈그림 11〉은 날짜별로 자기회귀모형으로부터 얻어진 추정값의 MAPE의 평균을 그래프로 나타낸 것이다.

MAPE의 값이 평일에는 1.2% 수준을 유지하다가 공휴일에는 2~3%로 평일과 공휴일에 예측값의 오차의 수준이 다르다는 것을 그래프를 보면 알 수 있다. 하지만 예측기간이 길어질수록 오차가 커지는 시계열분석과는 달리 오차가 안정적임을 알 수 있다. 다시 말하면, 자기회귀모형은 시계열모형과는 달리 장기 결측값의 추정에도 적합하다고 판단된다.

일 교통량을 추정할 때 계절 시계열 모형에서는 과거 시간대별 교통량으로 시계열분석의 추정식을 구하여 일 교통량을 구하는 것보다 과거 일 교통량의 추정식을 구하여 구하는 것이 오차를 줄이는 방안 중 하나라고 분석되었다. 다음은 자기회귀분석으로 일 교통량을 추정할 때 시간대별 교통량과 일 교통량 중 어떤 값을 이용하는 것이 더 좋은가를 두 가지 방법의 오차를 분석하여 알아보고 있다. 두 집단의 속성이 같은지를 알아보는 방법으로 T검



〈그림 10〉 자기회귀모형의 참값과 추정값의 비교의 예



〈그림 11〉 날짜별 자기회귀분석의 추정값의 오차(MAPE)-10월

정이 널리 이용되며, 지점별로 유사 지점의 시간대별 교통량으로 보정지점의 일 교통량을 구하는 방법과 유사지점의 일 교통량으로 보정지점의 일 교통량을 구하는 방법을 비교하고자 유사지점의 시간대별 교통량으로 보정지점의 일 교통량을 구하여 실제 값과 오차를 구한 값과, 유사지점의 일 교통량으로 보정지점의 일 교통량을 구하여 실제 값과 오차를 구한 값이 같은지 다른지를 T검정을 이용하여 비교 분석하였다. 검정에 이용될 가설은 다음과 같다.

- H_0 : 두 방법의 오차는 차이가 없다.
- H_1 : 두 방법의 오차는 차이가 있다.

분석결과 검정통계량 t값은 -3.29로 나타났으며 p값은 0.000으로 귀무가설(H_0)이 기각되었다. T검정 결과 두 방법 간에는 차이가 있는 것으로 분석되었으며, 유사지점의 시간대별 교통량으로 보정 지점의 일 교통량을 추정하였을 때의 MAPE의 평균은 1.94, 유사지점의 일 교통량으로 보정지점의 일 교통량을 추정하였을 때의 MAPE의 평균은 5.23으로 유사지점의 시간대별 교통량으로 보정지점의 일 교통량을 추정하는 것이 오차가 적은 것으로 분석되었다. 즉 자기회귀분석으로 일 교통량을 추정하고자 할 때에는 시계열분석과는 달리 패턴이 유사한 지점의 시간대별 교통량을 이용해서 보정 대상 지점의 일 교통량을 추정하는 것이 더 좋다고 판단된다. 이는 자기회귀모형은 시계열분석과 달리 참값을 토대로 결측값을 추정하며, 독립변수가 세분화될수록 결측 시점의 교통현황을 잘 반영하기 때문이라 판단된다.

2. 분석결과

장비를 통해서 수집되는 교통량 자료의 보정을 위해

3장에서는 시계열 분석과 자기회귀분석으로 분석하였으며, 그 결과 시계열 분석은 단기 결측 자료의 추정에 적합하고 자기회귀분석은 단기결측과 장기결측의 추정 모두 적합하다고 분석되었다.

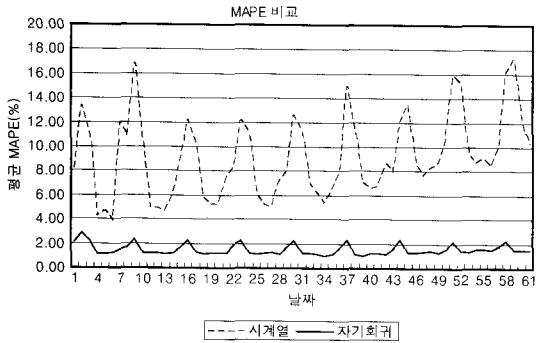
2004년 시간대별 교통량으로 시계열 분석을 하였을 때, 실제 자료를 이용하여 그 다음 시간의 교통량의 추정 값을 구하는 것은 오차가 크지 않은 것으로 분석되었다. 하지만 2005년 시간대별 교통량을 2004년의 방법과 달리, 결측이라고 가정하여 자료가 없는 경우에 추정값을 구하는 것은 오차가 크게 나타났다. 이와 같은 시계열 분석의 내용은 다음과 같이 설명할 수 있다.

실제자료	$y_1, y_2, y_3, \dots, y_{200}$
추정값1	$\hat{y}_{51} = f(y_1, y_2, \dots, y_{50}),$ $\hat{y}_{52} = f(y_2, y_3, \dots, y_{51}), \dots,$ $\hat{y}_{201} = f(y_{151}, y_{152}, \dots, y_{200})$
추정값2	$\hat{y}_{201} = f(y_{151}, y_{152}, \dots, y_{200}),$ $\hat{y}_{202} = f(y_{152}, y_{153}, \dots, y_{201}),$ $\hat{y}_{203} = f(y_{153}, y_{154}, \dots, y_{201}, y_{202}), \dots,$ $\hat{y}_{251} = f(y_{201}, y_{202}, \dots, y_{250})$

위의 식은 계절주기가 50인 경우를 가정하고, 참값에 대하여 추정식을 통해서 참값을 이용하여 추정값을 구한 것과 추정식을 통해서 추정값을 이용하여 추정값을 구한 것을 비교한 것이다. 추정식으로 실제 자료를 이용하여 추정값을 구한 경우 오차가 적은 것으로 분석되었으나, 추정식으로 결측 자료를 예측한 경우 단기간 결측에는 실제 자료를 이용하여 추정하지만 장기간 결측은 추정값을 이용하여 추정을 해야 하므로 오차가 커졌다. 결측 자료가 몇시간 되지 않을 경우에는 시계열분석을 이용해도 오차가 크지 않지만, 장기간 결측이거나 불량일 경우에는 시계열분석은 오차를 수반할 것으로 예상된다.

〈그림 12〉은 시계열 분석과 자기회귀분석의 MAPE를 날짜별로 평균하여 분석한 그림이다.

〈그림 12〉는 2005년 10월 1일부터 2005년 11월 30일까지의 자료가 결측이라고 가정하고, 2005년 1월 1일부터 2005년 9월 30일까지 자료를 이용하여 교통량 추정식을 구하고, 추정식으로 추정값을 구하여 참값과 비교하여 오차를 비교·분석한 것이다. 시계열 분석의 경우 일 교통량을 추정할 때 과거의 일 교통량을 시계열 분석하여 추정하는 것이 결과가 더 좋게 나왔으므로 일 교통량 추정에는 과거의 일 교통량을 이용하였다. 자기회귀분석의 경우 일 교통량을 추정할 때, 패턴이 유사한



〈그림 12〉 시계열 분석과 자기회귀 분석 추정값의 평균 MAPE 분석

지점의 일 교통량으로 추정하는 것보다 시간대별 교통량을 추정하여 일 교통량을 산출하는 것이 오차가 더 적게 나왔으므로 일 교통량 추정에는 패턴이 유사한 지점의 시간대별 교통량을 이용하였다. MAPE 분석결과 시계열 분석보다 자기회귀 분석의 MAPE의 평균이 훨씬 낮음을 알 수 있었다. 또한 시계열 분석의 오차는 결측기간이 길어질수록 커진다는 것도 알 수 있다.

〈그림 12〉로 판단하였을 때, 대부분의 지점에서 자기회귀모형을 이용하여 보정을 한 보정값이 계절 시계열모형을 이용하여 보정을 한 보정값보다 오차가 적은 것으로 분석되었다. 하지만 자기회귀모형의 경우 보정하고자 하는 지점의 인근에 패턴이 유사한 상시조사 지점이 있어야 하며, 유사하다고 판단되는 장비는 보정기간 내에는 결측이거나 불량이어서는 안된다. 2004년 상시조사 지점의 시간대별 교통량으로 상관분석을 시행한 결과 353지점 중 312개만이 유사한 조사 지점이 존재하는 것으로 나타났다. 만약 인근에 교통 패턴이 유사한 상시조사 지점이 없거나, 교통 패턴이 유사한 상시조사 지점은 있지만, 그 지점의 자료가 불량하여 교통 패턴이 유사한 지점의 자료를 이용할 수 없다면, 자기회귀분석은 이용할 수 없으며 이 경우 과거 교통량을 이용하여 보정을 할 수밖에 없다. 과거 자료를 이용하여 분석하는 대표적인 방법인 시계열분석은 단기결측에는 적합하지만 장기결측에는 부적합하다고 분석되었다. 그러므로 교통량 자료가 장기적으로 결측 또는 불량 자료가 발생할 경우, 오차를 수반할 수 있으므로 분석에 주의를 기울여야 한다.

V. 결론 및 향후 연구과제

교통량 자료는 도로 건설 및 계획에 있어서 매우 중요

한 자료로 활용된다. 특히 상시조사 자료는 특정 장소의 교통 변동을 시계열적으로 파악함으로써, 연평균 일교통량과 변동계수를 산출한다. 또한 상시조사 지점의 그루핑 결과를 이용하여 수시조사 지점의 정확한 AADT 추정을 할 수 있다.

상시조사 자료가 공사나 장비의 고장으로 인하여 일정 기간 자료가 결측일 경우 보정을 하지 않으면 상시조사 지점의 월별 계절별 특성 뿐만 아니라 시계열적인 변동을 파악하기 힘들게 되어 상시조사의 본래의 목적을 상실하게 된다. 이에 본 연구에서는 상시조사의 결측 및 불량 자료에 대해서 보정하는 방법에 대해 알아보았다. 보정하기에 앞서 보정 대상 지점의 교통특성을 먼저 파악하고 보정하는 것이 정확한 교통량을 추정하는 데 효과적이며, 이를 위해서 도로의 기능분류가 우선적으로 필요하다고 판단된다.

교통량 보정을 위해 본 연구에서는 시계열 분석과 자기회귀분석 방법을 제시하였다. 시계열분석은 보정 대상 지점과 동일 지점의 과거 자료를 이용해서 보정하는 방법이며, 이 방법으로 추정값을 제시한 결과, 단기결측에서의 추정은 가능하나 장기결측에서의 추정은 오차가 크게 나타나서 장기결측에는 부적합하다는 결론을 얻었다. 이 방법은 보정 대상 시점의 주변 상황이나 날씨 등은 고려하지 못하는 방법이며, 과거 자료만을 이용하여 추정한다는 점에 있어서 한계가 있다고 판단된다. 이에 본 연구에서는 이러한 한계점을 극복하기 위해 보정대상 시점과 동일한 시점의 자료를 이용할 수 있는 자기회귀분석을 이용하는 방법을 제시하였다.

자기회귀분석은 인근에 패턴이 유사한 지점의 교통량을 이용하여 보정 대상 지점의 교통량을 예측하는 방법이다. 보정결과 대부분의 지점에서 시계열분석에 비해 오차가 적게 나타났으며, 이것은 교통량 자료는 과거 패턴보다 현재 시점의 영향을 더 많이 받는 것이라 판단된다.

하지만 자기회귀분석은 인근에 패턴이 유사한 지점이 있어야 가능하며, 인근에 패턴이 유사한 지점이 있더라도 그 지점의 자료가 불량일 경우 보정이 불가능하다는 단점이 있다. 이 경우 결측 및 불량인 기간이 짧을 경우 시계열분석을 이용할 수 있으며 장기간일 경우 다른 분석방법을 이용해야 한다고 판단된다.

만약 자기회귀분석이 불가능할 경우에는 과거 교통량을 이용할 수밖에 없다. 하지만 이는 보정시점의 중요한 변수들의 영향을 반영하지 못하므로, 이러한 점을 해결하여 장기 결측의 오차를 줄이는 방법에 대한 추후 연구가 필요하리라 판단된다.

참고문헌

1. 건설교통부(2001), 도로 교통량 조사 지침.
2. 건설교통부(2006), 2005 도로 교통량 통계연보.
3. 한국건설기술연구원(2001), 통행시간 추정과 통계 자료 보정을 위한 결측 데이터 보정 모형 제작.
4. 장진환 외 1인(2005), "교통량 결측 자료 대체기법 연구", 대한토목학회 정기학술대회.
5. 임성한 외 2인(2005), "요인분석을 활용한 일반국도 유형분류", 한국도로학회 논문집.
6. 김동희 외 6인(2001), "통계학 -이론과 응용-", 자유아카데미.
7. 이덕기(1999), "예측모형의 이해", SPSS아카데미.
8. 정동빈 외 1인(2001), "SPSS를 활용한 시계열자료와 단순화분석", SPSS아카데미.
9. 강근석 외 1인(1999), "회귀분석", 교우사.
10. 박우창 외 4인(2003), "데이터마이닝 개념 및 기법", 자유아카데미.
11. 도철웅(1998), "교통공학원론", 청문각.
12. Stathis C. Sharma Birj M, Gulati Samantha N. Rizak(2006), "Statewide Traffic Volume studies and precision of AADT estimates", Journal of Transportation engineering Vol. 122 No. 6, pp.430~439.
13. Yiannis Kamarianakis, Poulicos Prastacos (2003), "Forecasting Traffic Flow Conditions In an Urban Network : Comparison of Multivariate And Univariate Approaches", Transportation Research Record No. 1857 Transportation Network Modeling, pp.74~84.
14. Williams, B M, Durvasula, P K, Brown, D E (1997), "Urban Freeway Traffic Flow Prediction : Application of Seasonal Autoregressive Integrated Moving Average and Exponential Smoothing models" Transportation Research Record No.1644 Traffic Flow Theory, pp.132~141.
15. Satish Sharma, Pawan Lingras, Ming Zhong(2001), "Effect of Missing Value Imputation on Traffic Parameters Estimations from Permanent Traffic Counts", Transportation Research Board 80th Annual meeting, Washington, D.C.

- ♣ 주 작성자 : 하정아
- ♣ 교신저자 : 하정아
- ♣ 논문투고일 : 2006. 10. 18
- ♣ 논문심사일 : 2006. 11. 16 (1차)
2006. 12. 9 (2차)
2006. 12. 17 (2차)
- ♣ 심사판정일 : 2006. 12. 17
- ♣ 반론접수기한 : 2007. 6. 30