

## 다중회귀분석을 이용한 포트홀 개수 예측 모델 개발

# Development of Predictive Model for the Number of Potholes Using Multi Regression Analysis

이상엽\* · 김도완\*\* · 최연우\*\*\*

Lee, Sangyum\* · Kim, Do-Wan\*\* · Choi, Yeon-Woo\*\*\*

### Abstract

Recently, the number of potholes in the principal cities of Korea has been increased because of the rapid climate change and the city population concentration. This study is to predict the number of potholes occurred in the Seoul metropolitan area. The prospective models for the pothole can help preventing the risk factors for the pothole. To finding the effect elements for the pothole occurrence, the empirical approach regarding the road engineering and analytical approach with stochastic analysis were used. The design of experiment regarding the surface response method was used to constitute the predictive model. The primary factors were utilized in the constitutive equation model using the multi regression analysis. The empirically predictive model makes little odds in p-value, coefficient of determination value and root mean squared error compared to the analytically predictive model. But, the analytical model has a lower incidence of error than the empirical model when the MPE(Mean Percentage of Error) analysis related to the maximum and minimum MPE was conducted.

**Key words** : Multi-linear regression, Standardization Coefficient, Surface response analysis, Pothole, Empirical & Analytical approach

### 요 지

최근 우리나라 주요도시의 인구집중현상과 급격한 기후변화로 인해 포트홀의 적출수가 크게 증가하고 있는 추세이다. 이와 관련하여 본 연구에서는 서울시 내에서 발생하는 포트홀의 수를 예측할 수 있는 모델을 개발하고자 한다. 포트홀 개수를 예측하기 위한 모델은 향후 포트홀로부터 발생할 수 있는 위험요소들을 사전에 방지하는 것에 기여할 것으로 예상된다. 포트홀의 발생에 크게 기여하는 요소를 찾기 위해서 도로공학적 지식에 기반한 경험적 접근방법과 확률적 분석에 근거한 분석적 접근법이 사용되었다. 또한 예측모델의 구성을 결정하기 위하여 실험계획법 중 표면반응분석법을 활용하였다. 이러한 방법으로 발견된 영향요소들은 다중 회귀분석에 사용되어졌다. 그 결과, 경험적 접근과 분석적 접근에 의한 예측모델은 P-값과 결정계수에서 큰 차이를 보이지 않았지만, MPE분석에서 경험적 접근에 의한 예측모델이 더 큰 오차발생률을 갖는 것으로 판단하였다.

**핵심용어** : 다중-선형 회귀분석, 표준화 계수, 표면반응분석, 포트홀, 경험적 & 분석적 접근법

## 1. 서 론

도로에 시공되어져 있는 아스팔트-콘크리트 및 시멘트-콘크리트 포장재료는 역학적으로 기후 영향, 교통특성 등의 요인에 민감하게 반응하는 것으로 알려져 있다. 이와 관련하여 급변하는 기후 및 교통특성 등의 영향에 의해, 도로에 균열

또는 파괴로 생성되는 포트홀(Pot hole), 도로의 하부 지반침하에 의한 싱크홀(Sink hole), 바다의 영향에 의해 육지가 수몰되어 나타나는 블루홀(Blue hole) 등이 발생할 수 있다.

이러한 관점에서 볼 때, 시간이 지날수록 대도시를 중심으로 인구가 집중되는 현상과 예측범위를 벗어나는 기후변화가 복합·다발적으로 이루어짐에 따라, 대도시 지역의 도로파손

\*정회원. 서울특별시청 도시안전실 도로관리과(E-mail: yummy10041004@gmail.com)

Member. Road Management Division, Seoul Metropolitan Government, Republic of Korea

\*\*교신저자. 정회원. 서울과학기술대학교 건설시스템공학과 석사과정(Tel: +82-2-970-6959, Fax: +82-2-970-6959, E-mail: marinater@hanmail.net)

Corresponding Author. Member. Master's course, The Road Pavement Research Division, No.43, Seoul National University of Science and Technology

\*\*\*서울특별시청 도시안전실 도로관리과

MS, Official, The Road Management Division, Seoul Metropolitan Government

발생현상에 대한 사회적 문제가 대두되고 있는 실정이다. 실제 세계 각 국에 걸쳐, 도로가 파손되는 현상으로 인하여 인명 및 재산피해 등의 심각한 손해를 입고 있으며 심지어 이러한 현상은 2차 사고를 유발하기도 한다. 예를 들어, 미국에서 작년에 발생한 전례에 없던 폭설, 폭우 및 폭염에 의해 다량의 포트홀과 싱크홀이 발생함에 따라, 올해 겨울 뉴욕에서 신고된 포트홀의 개수만 약 8만 개소 이상이 접수되었으며, 최근 그 피해에 대한 심각성이 점차 커지고 있다. 이와 비슷한 사례로 우리나라에서도 유사한 원인에 의해서 지난 2012년 서울시 내에서 41,313건으로 발생한 도로 파손 수에 비해 2013년에는 약 74,122건으로 급증하여 약 1.8배 증가하는 양상을 보였다. 이때 연 최고기온이 예년에 비해 약 1도정도 상승하였으며, 월 최대 강수량 역시 약 200 mm이상 증가하였고, 월 최대적설량 또한 8 cm이상 증가한 것으로 조사되었다(기상청 기상연보. 2009, 2010, 2011, 2012, 2013). 이러한 피해를 예방하고 줄이기 위해서 세계 각 국에서는 경제·사회·기술적 등의 다양한 면에서 확고한 대책 및 정책을 제시하고 있는 상황이다.

이와 관련하여 본 연구에서는 과거 기상자료와 교통량을 독립변수로 판단하여 종속변수인 포트홀의 발생 수를 예측할 수 있는 모델을 개발하고자 한다.

## 2. 경험적 연구 접근법

### 2.1 개요

실제 서울시 내에 포장되어 있는 도로는 크게 아스팔트 포장과 콘크리트 포장의 2가지로 나눌 수 있다. 아스팔트 포장의 경우 서울시 내의 도로 대부분을 차지하고 있으며, 콘크리트 포장은 서울외곽순환도로와 같은 일부 자동차 전용도로 구간에만 포장되어 있다. 이에 대한 파손의 형태는 콘크리트 포장과 아스팔트 포장에서 다르게 발생할 수 있으며 그에 대한 보수공법도 다르게 선택되어 진다.

이 중 본 연구에서는 아스팔트 포장에서 발생할 수 있는 포트홀의 수만을 고려하여 예측 모델을 개발하였다. 포트홀의 경우, 하부지반의 다짐상태 및 배수 상태가 불량하거나 기층의 파손, 택코트 불량, 집중호우, 국부적 결함 및 박리 등의 원인에 의해서 발생할 수 있다.

일반적으로 점탄성체로 분류되어 지는 아스팔트의 경우, 이에 대한 특성을 마스터커브를 이용함으로써 설명 가능하다. 횡축을 점탄성체가 받는 외력에 대한 주파수로, 종축을 점탄성체의 동적 저항력으로 표현하는 점탄성체에 대한 마스터커브는 같은 온도에서 외력의 주기가 길수록, 같은 외력의 주기에서 온도가 높을수록 재료의 동적 저항력이 저하되는 양상을 보인다. 이는 온도가 높을수록, 외력의 주기가 길수록 비교적 아스팔트혼합물의 손상이 더 심하게 발생할 수 있다는 것을 의미한다.

이러한 점탄성체의 역학적 특성을 고려하여 경험적 영향요소들을 판단하였다.

### 2.2 경험적 독립변수의 결정

본 연구에서 사용한 독립변수는 11가지 기상자료(월 평균 기온, 월 최고기온, 월 최저기온, 상대습도, 최저상대습도, 증발량, 월 총 강수량, 최심신적설량, 총 강수일, 최대 연속강수일, 최대 적설량)와 월 총 교통량을 사용하였다.

우리나라는 4계절의 영향을 뚜렷하게 받는 국가이기 때문에, 포트홀 발생에 기여하는 기상현상 중 영향요소를 판단하기 위해 2009년부터 2013년까지 서울시내 도로에서 발생한 월 별 총 포트홀의 수를 다음 Fig. 1과 같이 도식화하였다.

마스터커브의 점탄성체의 기본적인 성질로 볼 때, 아스팔트 도로에 가장 영향을 많이 미칠 것으로 판단되는 것은 온도와 관련된 월 평균, 최고, 최저 기온과 하중주기와 관련된 교통량이다.

Fig. 1을 보면 알 수 있듯이, 서울시 내의 여름(6~9월)과 겨울철(12~2월)에 상대적으로 다른 계절에 비하여 많은 포트홀이 발견된 것을 확인 할 수 있다. 이러한 현상과 기상자료를 관련 지어 판단해 본다면, 여름철에는 아스팔트의 동적 저항치가 줄어 파손이 발생할 수 있다고 판단할 수 있다. 기온자료만 고려한다면 월 최고기온은 실제 도로의 동적 저항력이 최하가 될 수 있다는 것을 내포한다. 따라서 월 최고기온과 평균기온을 활용하도록 한다. 그러나 매년 각 7월의 기온과 포트홀 발생 개수를 비교해 보면 확실한 영향을 파악하기 힘들다.

또한, 이 기간 내에 우리나라에서는 장마가 발생하기 때문에 강수량 역시 영향을 준다고 판단할 수 있다. 이는 강수에 의해 도로 내의 취약부분에 수분이 침투하고 이 지점을 통과하는 주행차량에 의해 파손이 발생된다는 것을 근거로 제시할 수 있다. 다루고자 하는 강수자료와 관련된 자료로는 월 총 강수량, 월 총 강수일수, 최대 연속 강수 일수가 존재하며, 실제 Fig. 1의 7월 데이터에서 볼 수 있듯이 강수량이 많은 해부터 순차적으로 그 포트홀이 다량 적출되었다. 또한 이러한 양상은 월 총 강수일수와도 관련이 있는 것으로 파악되었다.

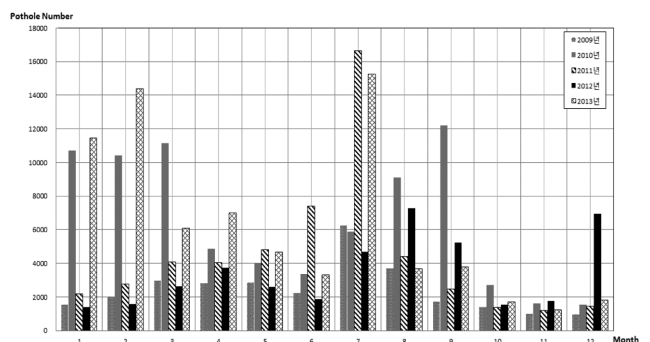


Fig. 1. The Number of Potholes Per a Month.

그러나 최대 연속 강수일수는 포트홀 발생 수와 비교한 결과 많은 영향을 미치지 못하는 것으로 판단하였다.

적설량의 경우를 보면 2010년과 2013년 1,2월에 두드러지는 포트홀 발생 수가 적출되었는데 이때 최심신적설 총량 및 최대 적설량이 2009년, 2011년, 2012년에 비해 상당한 차이를 보인다. 이를 근거로, 적설 또한 포트홀의 발생에 상당한 영향을 미친다고 판단할 수 있다. 더욱이, 이는 겨울철에 도로가 결빙되어 파괴되는 경우, 차량의 스노우 체인에 의한 영향 및 제설제(i.e. 염화칼슘( $CaCl_2$ ) 등)의 사용으로 인한 영향의 이유를 근거로 둘 수 있다. 적설량만을 월 단위로 고려할 때, 다양한 적설에 대한 피해를 방지하기 위해 제설제를 포설하기 때문에 적설의 영향을 생각해 볼 때, 최심신적설의 총량이 실제 도로에 쌓인 양의 월 총량이기 때문에 최대 적설량 보다 영향을 많이 미칠 것으로 판단하였다.

기상자료 중 상대습도자료를 확인해 본 결과, 1,2월에 비해 11, 12월에 상대습도가 더 높은 반면, 포트홀 발생 수는 더 적은 양상을 보이는 등 그 관련성이 매우 떨어지는 것으로 판단하였다. 또한 상대습도(월 평균, 최소 상대습도), 증발량의 경우 강수량 및 기온과 밀접한 연관이 있다고 할 수 있다. 실제 강수량이 많은 7, 8월의 경우 상대습도가 높은 반면 증발량 또한 감소한다. 따라서 기온과 강수량, 적설량을 포트홀 발생 개수에 영향을 미치는 독립변수로 결정하였기 때문에 상대습도와 증발량이 직접적인 영향을 미치지 않는 다고 판단하였다.

경험적, 수치자료에 대한 안목적인 방법을 통해 포트홀 발생에 영향을 미칠 것이라고 판단된 독립변수는 평균기온, 월 총 강수량, 최심신적설 총량, 월 총 강수일, 교통량의 5가지로 선정하였다.

### 3. 분석적 연구 접근법

#### 3.1 개요

포트홀 발생에 영향을 미치는 독립변수를 찾기 위해 각 독립변수와 종속변수에 대한 상관관계 분석과 다중-선형 회귀 분석을 실시하였다.

독립변수는 2장에서 필터링하기 전에 고려하였던 11가지의 기상자료와 교통량 자료를 그대로 사용하였으며, 종속변수는 서울시 전역에서 발생한 매 달의 총 포트홀 발생수로 선정하였다.

#### 3.2 다중-선형 회귀분석

서로 복합인 관계에서 어떠한 독립변수들이 영향을 미치는 지 파악하기 위하여 다중-선형 회귀분석을 실시하였다. 회귀 분석을 실시하는 과정에서 독립변수들은 모두 선형의 관계를 갖는다고 가정하여 다중-선형 회귀분석을 실시하였다.

다중-선형 회귀방정식은 각 데이터들의 선형적인 합의 관계

Table 1. Independent Variable information

Variable	Information	Variable	Information
$x_1$	Aver. Temperature	$x_7$	Precipitation
$x_2$	Max. Temperature	$x_8$	Max New Snowfall
$x_3$	Min. Temperature	$x_9$	Precipitation Days
$x_4$	Aver. Relative Humidity	$x_{10}$	Max Continuous Precipitation Days
$x_5$	Min. Relative Humidity	$x_{11}$	Max Snowfall
$x_6$	Aver. Evaporation	$x_{12}$	Traffic Volume

가 결국 종속변수와 관련이 있다고 설명할 수 있으며, 다음 식 (1)과 같은 다중-선형 회귀방정식에 기초하게 된다.

$$PN = \beta_0 + \beta_1 x_1 + \dots + \beta_{12} x_{12} \quad (1)$$

여기서,  $PN$ 는 각 달에 서울시에서 발생한 포트홀의 수를 의미한다.

Table 1은 각 독립변수에 대한 설명을 나타낸다.

다중회귀분석에서 기초되는 것이 최소 잔차 자승법(Method of Least residual Squares)으로 종속변수인 포트홀 발생수와 다중 회귀식의 차인 잔차를 이용하는 방법이다.

최소 잔차 자승법은 실제 이 오차가 최소가 되도록 하기 위해, 함수를 각 미지수에 대하여 미분을 하고 이 미분 값이 0인 값을 역으로 찾아가는 방법이다. 본 연구에서 사용한 상관관계 계수는 표준편차에 대한 공분산을 의미하는 피어슨의 적률상관계수, 결정계수(Coefficient of Determination)와 조정 후 결정계수(Adjusted R-Square)이며, 결정계수는 상관관계수( $R$ )의 제곱으로 나타낼 수 있다. 또한, 데이터의 총 수( $n$ )가 독립변수의 수( $k$ )보다 비교적 많기 때문에, 영향력이 비교적 낮은 독립변수를 추가적으로 포함하여 상관관계( $R^2$ )를 증가시키는 것을 방지하기 위해 독립변수의 수와 표본의 수로써 조정되어진 조정 후 결정계수를 사용하였다. 이는 다음 식 (2)과 같이 표현 가능하다.

$$R_{adj}^2 = 1 - \frac{(1-R^2)(n-1)}{n-k-1} \quad (2)$$

각 독립변수간의 선형적 판단으로 인해 예측되어진 포트홀의 수와 측정되어진 포트홀의 값을 다음 Fig. 2과 같이 비교하였다.

만약, 도출된 회귀방정식이 정확성을 설명하지 못한다면 본 연구에서 실시한 회귀분석의 신뢰할 수 없기 때문에, 도출되어진 다중-선형회귀방정식을 평가하기 위하여 상관계수, 결정

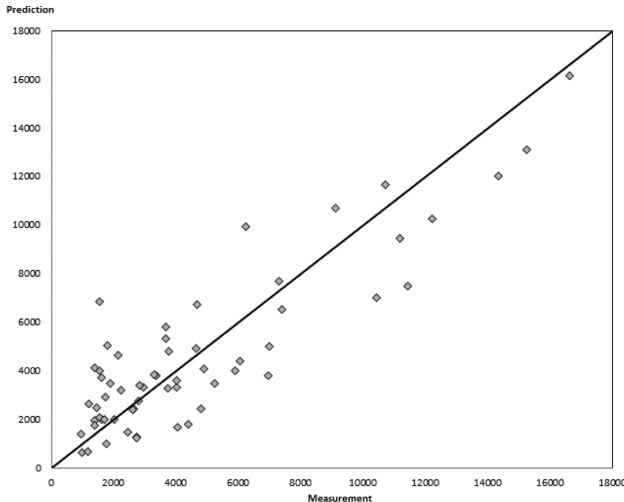


Fig. 2. Multi-Linear Regression Results.

계수, 수정 후 결정계수 및 오차를 계산하였다.

오차는 RMSE를 사용하였으며, 이러한 오차는 다음 식 (3)와 같이 표현가능 하다.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (M_n - P_n)^2} \quad (3)$$

여기서,  $M$ 과  $P$ 는 각각 측정된 포트홀의 수, 예측된 포트홀의 수이며,  $N$ 은 총 관측자료의 수 이다.

이러한 방법으로 도출되어진 서울시 전체 포트홀 발생수에 대한 다중-선형회귀 방정식 중 각 미지수( $\beta$ )값은 비표준화 계수(Unstandardized Coefficient, UC)를 나타내며, 이 비표준화 계수는 표준화계수(Standardized Coefficient, SC)를 구하는데 사용된다. 본 연구에서 표준화계수를 사용하는 이유는 각 독립변수마다 표준편차가 상이하고 단위마저 다르기 때문에, 이들을 단순히 전체적인 영향력에 비교하는 것은 수학적으로 상당한 오류를 갖기 때문이다. 따라서 독립변수들에 대해 표준화하여 비교, 분석할 필요가 있다.

이러한 분석 외에도 각 설명변수간의 상관관계를 확인하기 위해서 다중 공선성 분석(Multi Collinearity Analysis)을 실시하였다. 다중 공선성 분석은 설명변수간의 상관관계 분석에 의존하며, 일반 상관관계 분석은 두 독립변수간의 상관관계만을 설명하기 때문에, 각 독립변수와 나머지 모든 독립변수의 관계를 설명하기 위해서는 공차한계(Tolerance)와 분산팽창요인(Variance Inflation Factor, VIF)을 확인하면 된다. 공차한계는 판독된 공선성을 점검하기 위해 사용되는 지표이다. 또한, 분산팽창요인은 공차한계의 역수로 표시되며, 이는 다시 말해 다중공선성의 심각성을 진단할 수 있는 지표가 된다.

Fig. 3은 SPSS를 이용한 다중-선형 회귀분석에 대한 결과를 나타낸다.

Fig. 3의 결과를 이용하여 도출된 다중-선형회귀방정식을 통해 예측한 서울시 전역의 포트홀 발생 수 결과는 Fig 2와

모형	비표준화 계수		표준화 계수		t	유의확률	공선성 통계량	
	B	표준오차	베타	베타			공차	VIF
1 (상수)	8960.100	7649.532			1.171	.247		
일평균기온평균	126.999	245.417	.356	.517	.607	.010	.010	97.025
일최고기온평균	-718.465	436.625	-1.997	-1.645	.107	.003	.003	301.560
일최저기온평균	660.014	478.475	1.844	1.379	.174	.003	.003	365.579
일상대습도평균	-35.317	114.961	-.070	-.307	.760	.095	.095	10.570
일최소상대습도평균	-177.007	101.971	-.444	-1.736	.089	.075	.075	13.398
증발량평균	299.450	337.338	.125	.888	.379	.245	.245	4.088
월총강수량	15.899	2.446	.906	6.501	.000	.252	.252	3.971
최심적설	14.866	37.378	.045	.398	.693	.374	.374	2.672
월총강수일	6.986	127.869	.009	.055	.957	.170	.170	5.893
최대연속강수일수	65.936	225.907	.036	.292	.772	.325	.325	3.078
최대적설량	479.071	108.099	.612	4.432	.000	.256	.256	3.900
교통량	3.321E-5	.000	.233	2.742	.009	.677	.677	1.477

Fig. 3. Multi-Linear Regression Analysis Results.

같으며, 피어슨의 적률상관계수, 결정계수, 조정후 결정계수는 각각 0.88, 0.77, 0.71로 도출되었으며 RMSE는 1812 개소이다.

### 3.3 분석적 독립변수의 결정

P-Value는 ‘독립변수가 종속변수에 영향을 미친다’는 귀무가설을 기각하는 정도를 나타내기 때문에 적을수록 이에 대한 영향도가 높다고 판단할 수 있다. 표준화계수는 종속변수에 대한 각 독립변수가 미치는 영향력을 나타내는 지표이기 때문에 이 값이 클수록 상대적으로 독립변수가 종속변수에 영향을 많이 미친다고 할 수 있다. 또한 공선성을 평가하는 지표인 분산팽창요인은 공차한계의 역수이며, 공차한계 값이 적을수록 나머지 독립변수가 설명하는 정도가 크다는 것을 의미하며, 분산팽창요인의 값이 클수록 나머지 독립변수가 설명하는 정도가 크기 때문에 공선성의 문제가 있을 것으로 판단할 수 있다.

이러한 이유를 근거로 볼 때, 가장 중점적으로 확인해야 할 것이 표준화계수이며, 여기서, 자료의 수치에 관계없이 경험적인 면으로 판단해 볼 때 교통량은 도로파손에 크게 기여하는 인자라는 것을 알 수 있다. 또한, 단적인 방향으로써 귀무가설을 기각을 약 10%로 판단했을 때 독립변수 자료에서 나머지 독립변수를 제어하도록 한다. 다중 공선성면에서 보면, 기온과 상대습도를 제외한 나머지 독립변수들은 공선성을 갖지 않는 개별적 독립변수로 분류된다. 따라서 분류된 기온에 대한 독립변수 중 최대기온의 데이터만 활용하도록 한다.

결국, 다중-선형 회귀분석에 의해 필터링된 독립변수는 그 영향력의 크기 순서로, ‘최대기온, 강수량, 최대 적설량, 최대 상대습도, 교통량’이다.

## 4. 실험계획법

본 장에서는 예측결과에 영향을 미치는 인자들에 대한 최적의 조건식을 찾기 위해서 실험계획법 중 반응표면분석(Response Surface Analysis)을 실시하였다(Lee *et al.*, 2012).

반응 표면분석은 종속변수에 영향을 주는 최적의 독립변수를 발견하여 이에 대한 최적의 조건을 찾고자 하는 경우 사용

되는 분석방법이다. 2장과 3장의 방법은 엄밀히 따지면, 실험 계획법에 속하는 방법이지만 본 연구에서는 이를 미리 수행한 후, 반응표면분석에서 목적방정식을 결정토록 하였다.

경험적인 선택방법을 통해서 평균기온, 월 총 강수량, 최신 신적설 총량, 월 총 강수일, 교통량의 독립변수가 선택되었으며, 상관관계 및 회귀분석적 방법을 통해서 최대기온, 강수량, 최대 적설량, 최저상대습도, 교통량의 독립변수를 선택하였다.

이러한 독립변수에 의한 예측결과는 다음 식 (4)와 같이 표현되어질 수 있다.

$$y = f(x_1, \dots, x_5) + \varepsilon \quad (4)$$

여기서,  $y$ 는 결정되어질 포트홀의 수를 의미하고,  $f$ 함수는 독립변수가 구성하는 회귀방정식이고  $\varepsilon$ 은 통계적 오차값이다.

결국, 독립변수에 의한 회귀방정식과 통계적 오차의 합으로 정확한 예측을 기대할 수 있다는 결론을 내릴 수 있다. 그러나, 독립변수 및 독립변수에 의한 회귀방정식이 서로 다른 통계적 분포를 가지고 있다. 만약 이러한 입력변수들이 평균이 0인 표준정규분포를 갖는다면, 이에 대한 평균적 함수는 다음 식 (5)과 같이 정리될 수 있다.

$$E(y) = E[f(x_1, \dots, x_5)] + E(\varepsilon) = f(x_1, \dots, x_5) \quad (5)$$

위 식 (5)의  $E(y)$ 는 반응함수(Response Function)를 의미한다.

위에서 언급한 바와 같이 이러한 조건을 만족시키기 위해 최소값이 -1, 최대값을 1로 정규화하는 작업을 수행하였다.

이러한 정규작업을 마친 독립변수들에 의한 방정식은 1차 방정식과 2차 방정식으로 표현되어질 수 있다(고차방정식은 고려하지 않음). 그러나 1차 방정식은 그 결과에 대한 유연성(자유도)이 낮기 때문에 본 연구에서는 다음 식 (6)과 같은 2차 모델 방정식을 선정하였다.

$$\begin{aligned} E = & \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_{11} x_1^2 + \beta_{22} x_2^2 \\ & + \beta_{33} x_3^2 + \beta_{44} x_4^2 + \beta_{55} x_5^2 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{14} x_1 x_4 + \beta_{15} x_1 x_5 \\ & + \beta_{23} x_2 x_3 + \beta_{24} x_2 x_4 + \beta_{25} x_2 x_5 + \beta_{34} x_3 x_4 + \beta_{35} x_3 x_5 + \beta_{45} x_4 x_5 \end{aligned} \quad (6)$$

여기서,  $E$ ,  $\beta$ 는 각각 예측되어진 포트홀의 정규화 작업에 대한 수치와 회귀계수를 의미한다.

## 5. 예측 모델

### 5.1 개요

본 장에서 실시하는 예측모델의 구성은 지역적 특성을 반영하기 위하여, 서울특별시 전체를 총 6개의 관할 사업소 별로

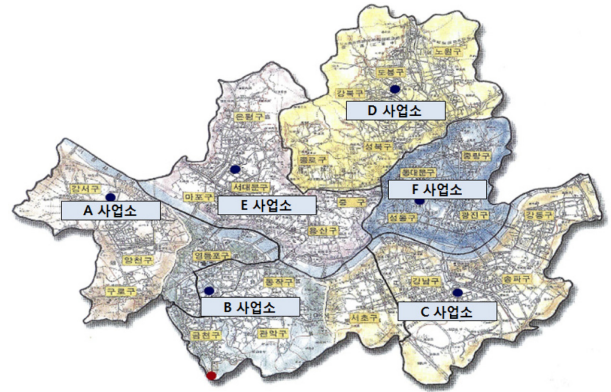


Fig. 4. Sections for Road Management.

나누었으며, 서울시 내의 자동차전용도로를 관리하는 서울시 시설공단(이하 PC, Public Corporation)을 따로 분류하여 총 7개의 서울시 내 지역적, 도로적 특성을 구분하였다. 이에 대한 지역적 분류는 Fig. 4와 같다.

### 5.2 다중 회귀분석

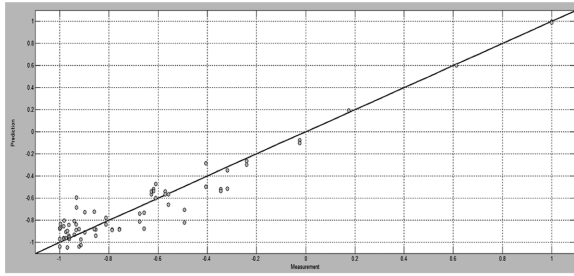
다중 회귀 방정식을 구성하는 변수는 최소 오차 자승법에 의하여 도출되어지며, 이에 대한 기본 이론은 3장에서 설명하였다.

다만, 4장에서 언급된 정규화 작업을 거친 자료를 본 예측 모델을 위해 사용하였으며, 예측모델 적용함수는 식 (6)과 같다. 총 7구간의 예측모델을 형성하였으며, 경험적으로 추려진 5개의 독립변수와 분석적으로 걸러진 5개의 독립변수로 고려하여 결론적으로 14개의 다중 회귀 예측모델을 구성하였다. 다중 회귀분석에 의해 도출되어진 회귀계수에 대한 결과는 Fig. 5와 Fig. 6과 같다.

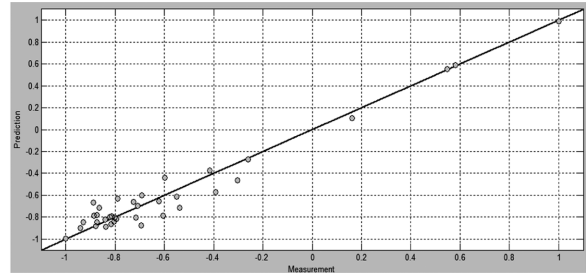
본 연구에서 개발되어진 다중회귀예측모델에 대한 적합성을 판단하기 위한 방법 중, 추세에 대한 적합성을 판단하기 위해 피어슨의 적률상관계수에 의한 결정계수를 사용하였으며, 오차에 대한 적합성을 판단하기 위하여 RMS 오차를 사용하였다. 각 지역 및 각기 다른 방법론적 영향요소를 적용한 다중 회귀 예측모델의 추세, 오차, P-value 결과는 다음 Table 2와 같다.

## 6. 결 론

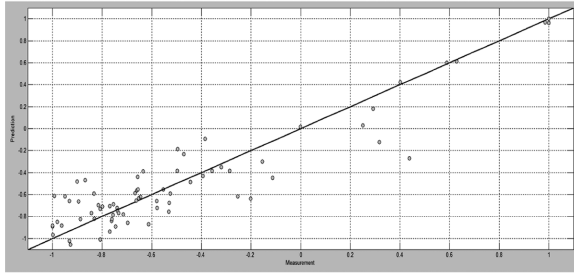
다중회귀분석을 이용하여 두 가지 방법으로 필터링한 독립변수에 대한 예측모델은 모두 포트홀 발생에 영향을 미치는 요소들에 의해 포트홀의 전체적 개수가 영향을 받는다는 귀무가설을 기각하는 P-값에서 유의수준(0.05)에 만족하는 결과를 보였다. 또한 RMSE 오차값과 추세를 의미하는 결정계수에서 큰 차이가 발생하지 않는 것으로 판단하였다. 그러나, 오차의 발생률을 확인한 결과 경험적 방법에 의해 영향요소를 필터링한 회귀예측모델이 최소 2%에서 최대 37%로 발생한



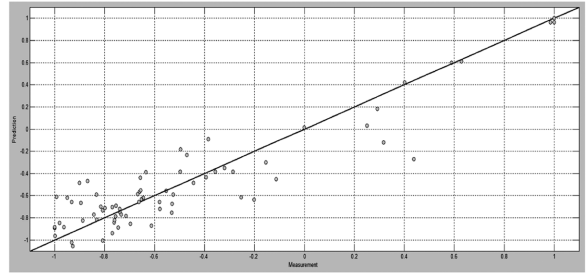
(a) 'PC' Road Management



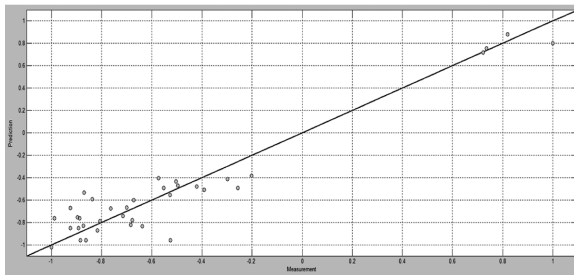
(b) 'A' Road Management



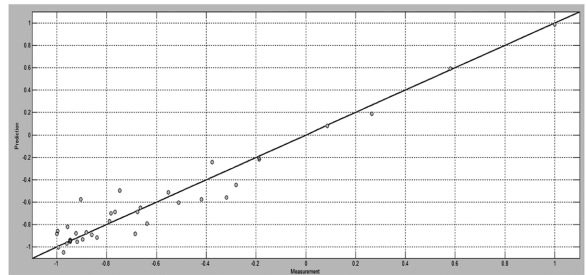
(c) 'B' Road Management



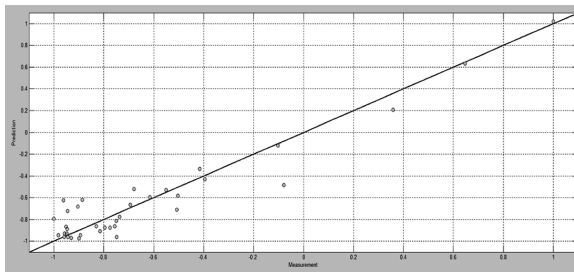
(d) 'C' Road Management



(e) 'D' Road Management

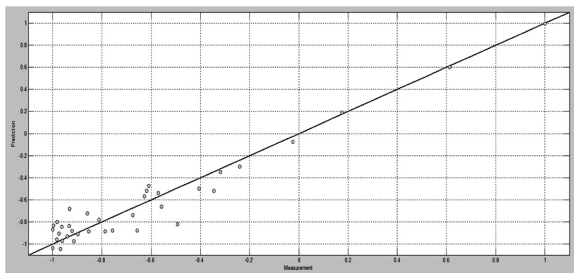


(f) 'E' Road Management

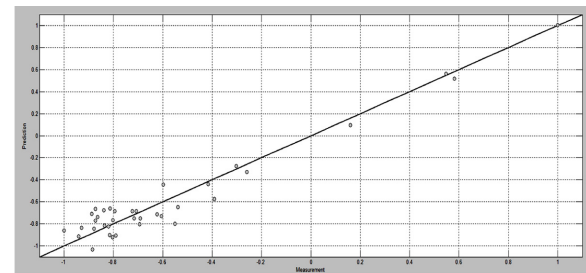


(g) 'F' Road Management

**Fig. 5. Multi Regression Results Using Empirical Approach.**



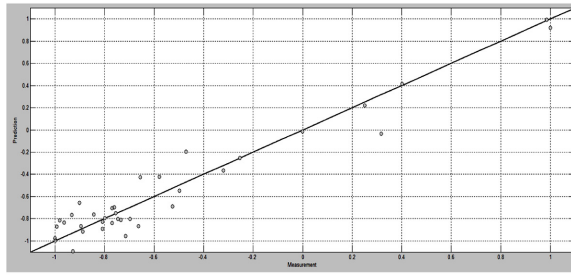
(a) 'PC' Road Management



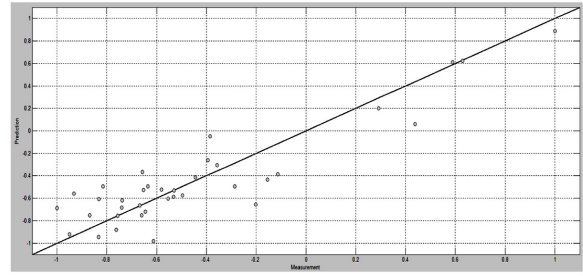
(b) 'A' Road Management

**Fig. 6. Multi Regression Results using Analytical Approach.**

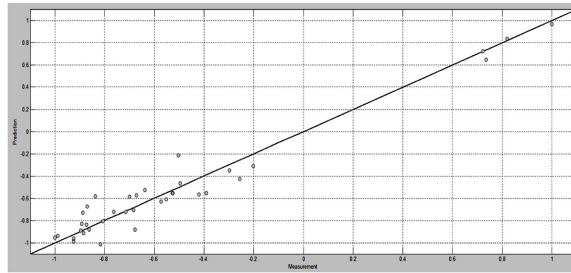




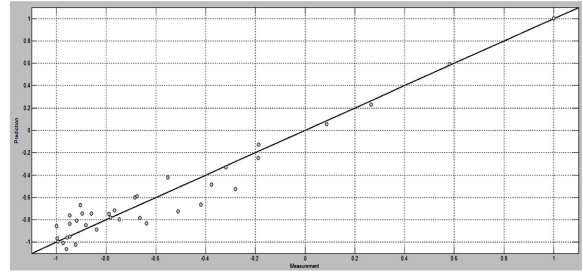
(c) 'B' Road Management



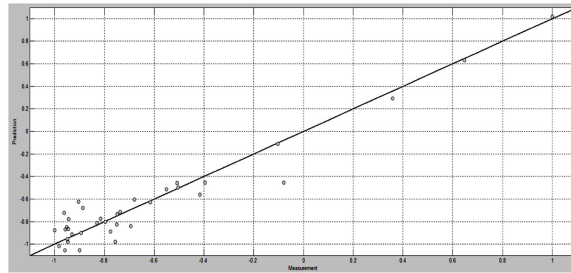
(d) 'C' Road Management



(e) 'D' Road Management



(f) 'E' Road Management



(g) 'F' Road Management

Fig. 6. Continued.

Table 2. Suitability Analysis for Multi Regression Equations

	Empirical Method							Analytical Method						
	PC	A	B	C	D	E	F	PC	A	B	C	D	E	F
RMSE	2426	1842	1588	1074	1534	1913	1066	2432	1874	1516	1050	1463	1917	1045
$R^2$	0.94	0.95	0.89	0.78	0.92	0.94	0.91	0.94	0.94	0.93	0.82	0.95	0.94	0.93
P-value	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00

반면, 분석적 방법에 의해 영향요소를 필터링한 회귀예측모델은 최소 1%에서 최대 31%로 오차발생률이 줄어든 것을 확인하였다.

추가적인 연구로 본 연구에서 개발된 모델의 정확도가 더욱 향상된다면, 향후 도로환경 변화로 발생 예측되는 파손에 대비하여 도로안전과 방재정책을 기획 수립하는데 활용될 수 있을 것이다.

## 감사의 글

이 논문은 2013년도 정부(미래창조과학부)의 재원으로 한국

연구재단의 지원을 받아 수행된 연구임(No. 2011-0030040).

## References

- Cleopatra Petroutsatou, Sergios Lambropoulos and John-Paris Pantouvakis (2006) Road Tunnel Early Cost Estimates Using Multiple Regression Analysis. *Operational Research. An International Journal*, 6, 311-322.
- Jog, G., Koch, C., Golparvar-Fard, M., and Brilakis, I. (2012) Pot-hole Properties Measurement through Visual 2D Recognition and 3D Reconstruction. *Computing in Civil Engineering* (2012), 553-560.
- Korea Meteorological Administration (2009) 2009 Annual Weather

Report. Korea Meteorological Administration, 58-60.

Korea Meteorological Administration (2010) 2010 Annual Weather Report. Korea Meteorological Administration, 57-59.

Korea Meteorological Administration (2011) 2011 Annual Weather Report. Korea Meteorological Administration, 59-61.

Korea Meteorological Administration (2012) 2012 Annual Weather Report. Korea Meteorological Administration, 58-60.

Korea Meteorological Administration (2013) 2013 Annual Weather Report. Korea Meteorological Administration, 59-61.

Lee, Chang-Joon, Kim, Dowan, Mun, Sungho and Yoo, Pyeong-Jun (2012) Study on a Prediction Model of the Tensile Strain Related to the Fatigue Cracking Performance of Asphalt Concrete Pavements Through Design of Experiments and Harmony Search Algorithm. *Journal of Korean Society of Road Engineering*, Vol. 14, No. 2, pp. 11-17.

Suh, Y., Mun, S., and Yeo, I. (2010) Fatigue Life Prediction of Asphalt Concrete Pavement Using a Harmony Search Algorithm, *KSCE Journal of Civil Engineering*, Vol. 14, No. 5, pp. 725-730.

Vida Mokhtari (2012) Regression Analysis of Factors in the agricultural production cooperative success of Ilam. *Journal of Trends in Advanced Science and Engineering*, Vol. 4, No. 2, pp. 100-109.

---

**Received** ■ April 25, 2014

**Revised** ■ May 2, 2014

**Accepted** ■ June 3, 2014