KJCEM 15. 4. 150~158
JULY 31, 2014

www.kicem.or.kr (pISSN 2005-6095)
http://dx.doi.org/10.6106/KJCEM.2014.15.4.150

하모니 검색 알고리즘을 이용한 포트홀 발생 개수 예측 모형

김도완[1] · 이상염* · 김동호[2]
[1]서울과학기술대학교 건설시스템공학과 · [2]서울과학기술대학교 건설시스템디자인공학과

A Predictive Model for the Number of Potholes Using Basic Harmony Search Algorithm

Kim, Dowan[1] · Lee, Sangyum* · Kim, Dongho[2]
[1]Road Pavement Research Division, Seoul National University Science and Technology
[2]Road Pavement Research Division, Seoul National University Science and Technology

Abstract : A bunch of asphalt roads have been damaged frequently in relation to the rapid climate change. To solve and prevent this type of problems, many nationalities in the world have performed various researches. In this regard, the objective of this study is to develop prediction model as to the number of potholes occurred in seoul. At the same time, we have utilized empirical and statistical approaches in order for us to identify factors which is affecting the actual occurrence. The predictive model was determinded by using BHS (Basic Harmony Search) algorithm. Prediction was based on the weather and traffic data as well as data occurrence data of porthole. To assess the influences which are PAR(Pitch Adjusting Rate) and HMCR(Harmony Memory Considering Rate), we determined suitability by changing the values. In the process of the determining a predictive model, the predictive model composed Training data (2011, 2012 and 2013yrs data). To determine the suitability of the model, we have utilized Testing Set (2009 and 2010 yrs data). The suitability of the basic prediction model has been from RMSE(Root Mean Squared Error), MAE(Mean Absolute Error) and Coefficient of determination.

Keywords : BHS Algorithm, HM, HMCR, MAE, PAR, Pothole, RMSE, Visco-Elasticity

1. 서 론

최근 급격히 변하는 기후영향에 적절하게 대응하기 위하여 다양한 분야에서 연구가 수행되고 있다. 이와 관련한 기후변화에 대한 영향은 기온, 풍향, 습도, 강수량, 적설량 등이 있다. 특히, 우리나라는 사계절의 영향을 뚜렷하게 받고 있기 때문에 기후변화에 매우 민감한 국가라고 할 수 있다. 다시 말해, 급격한 기후변화로 인해 직 · 간접적 피해를 입을 수 있다는 것이다. 예를 들어, 여름철에는 갑작스러운 기온 상승 및 폭염, 2011년 월 최대 총 강수량이 1131mm를 기록했지만, 2012년에는 496.4mm를 기록하는 등 예년에는 경험하지 못했던 급격한 변화의 강수 영향을 받으며, 겨울철에는 최심신적설(最深新積雪)량이 2010년에는 169.6cm로 기록되었던 반면, 2012년에는 그에 반도 미치지 못하는 55.4cm를 기록하였다. 이러한 예측 불가능한 기상현상이 지속되고 있는 상황에서 이에 대한 공학적인 관심이 대두되고 있다.

이와 관련하여, 아스팔트로 포장되어진 도로는 기온, 노면의 수분상태 및 차량하중 등의 조건에 매우 민감하기 때문에 급격한 기후변화에 영향을 많이 받는 구조물이라고 할 수 있다. 실제, 아스팔트는 일반적으로 점탄성체의 동적 저항능력을 표현하는 마스터커브(Master Curve)로 그 물성을 설명할 수 있다. 마스터커브에서 온도에 의해 물성치가 변화하는 양상을 보인다. 또한, 강수현상 또는 재설작업에 의해 노면 공극이나 파손부위가 물로 채워진 경우, 이 노면을 지나는 교통하중에 의해 파손이 발생될 가능성이 높아진다. 이러한 관점에서 볼 때, 하중조건 및 기온, 수분 등의 영향에 매우 민감한 아스팔트는 폭염, 폭설, 폭우 등을 경험하는 우리나라에서 1년 동안 크게 손상될 수 있다는 것을 의미한다.

이러한 도로의 파손은 유지 및 보수의 면에서 경제적인 손실을 발생시키고 2차 사고를 유발할 수 있는 가능성을 갖으며 도로파손의 상태는 도시미관상 보행자에게 불쾌감을

* Corresponding Author: Lee, Sangyum, Ph. D., The Pavement Research Center, Quality Inspection Office, Seoul Metropolitan Government, Republic of Korea
E-mail: yummy10041004@gmail.com
Received May 30, 2014: revised June 12, 2014
accepted June 19, 2014

주는 등 다양한 문제를 야기한다. 이와 관련하여, 본 연구에서는 아스팔트 도로에서 발생할 수 있는 포트홀(Pothole)로 인한 피해를 예방하고 방지하는 차원의 초기 대응으로 포트홀 발생 수와 관련한 예측모델을 개발하였다. 또한, 이 예측 모델에서 결정되어진 비표준화계수(Unstandardized Coefficients)를 표준화계수(Standardized Coefficients)로 변환하여 실제 포트홀에 영향을 주는 인자를 찾는 것에 목적을 두었다.

예측모델에 사용되는 미지수를 결정하기 위하여 하모니 검색 알고리즘(Harmony Search Algorithm)을 사용하였다. 하모니 검색 알고리즘에 적용되어지는 독립변수는 한 달에 기초한 최대기온, 평균기온, 최저기온, 상대습도의 평균, 최저 상대습도, 증발량 평균, 총 강수량, 총 강수일수, 최대 연속강수일수, 최심적설 총량, 최대 적설량의 11가지 기상자료와 교통량 자료를 사용하였다. 독립변수에 대한 자료는 예측모델을 생성하기 위해 2011년, 2012년 그리고 2013년간의 Training Set으로 구성하였으며, 이로 결정되어진 예측모델의 실제 적합성을 평가하기 위해 2009년과 2010년의 Testing Set으로 적용되어졌다. 예측모델 적용 지역은 대도시로 인구가 집중되는 도시집중화 현상을 고려하여, 우리나라 내 서울시 전역으로 결정하였다.

2. 점탄성 아스팔트 도로

기존의 연구(Y. A. Jimoh(2012))에서는 기후영향들을 배제하고 교통하중에 대한 영향을 지정기간 내의 차종과 교통량으로 세분화 하여 포트홀의 발생 수를 예측하였다. 예측모델의 적용은 Linear, Natural, Logarithm, Quadratic, Exponential의 회귀분석(Regression Analysis)에 기초하였다.

그러나 본 연구에서는 아스팔트 자체의 성질과 급격한 기후변화를 고려하여 기상자료를 추가하되, 교통하중에 의한 영향은 단지 서울시에서 활동하는 모든 차량의 수만 고려하였다.

점탄성체인 아스팔트 혼합물은 온도와 하중의 조건에 매우 민감한 특성을 가지고 있다. 이와 관련하여, 실제 온도가 높은 경우, 서울시 내에서 발생한 포트홀의 수가 크게 증가하는 것을 Fig. 1에서 같이 확인 할 수 있다. Fig. 1은 월 별 수분 관련 포트홀의 수를 나타낸다. 그러나 점탄성체로 분류되는 도로는 다른 현상에도 민감하기 때문에 이러한 현상을 오로지 온도에만 관련되어있다고 단정 지을 수는 없다.

또 다른 점탄성체의 민감한 요소라고 할 수 있는 것이, 하중조건이다. 같은 크기의 하중에서 그 주기에 따라 동적 물리특성이 다르다. 다시 말해, 하중의 크기가 같더라도, 버스전용차로와 같이 하중 주기가 긴 경우, 그 동적인 저항력이 저하되며, 고속도로와 같은 하중 주기가 짧은 경우 그 동적 저항력이 상대적으로 높다.

위 두 가지의 하중 및 기온에 따른 민감 요소는 아스팔트에 관련된 주요 특성이며 이를 시각적으로 표현하기 위해 마스터커브를 활용하여 점탄성체의 물성을 나타낸다.

실제 서울시에 설계 및 포장되어 있는 아스팔트 도로는 각 지역의 특성에 맞추어져 있다. 따라서 그 지역적 특성에 따라 다른 물리적 특성을 갖는데, 투수성포장 및 저소음포장의 경우 일반 아스팔트 도로와 비교해 공극을 갖는 포장체로 다소 외력의 저항성에서 낮은 단점을 보이며, 수분에 상대적으로 민감한 반응을 보인다.

이러한 기후와 하중조건의 영향 중 실제 포트홀 발생에 영향을 주는 요인을 찾기 위해 표준화계수를 산정하여 비교하였다.

3. 데이터 수집

기상관련 자료는 기상청에서 제공하고 있는 기상월보에 의해 자료를 수집하였으며, 도로의 파손은 아스팔트에서 발생할 수 있는 다양한 파손형태 중 포트홀만 고려하였다.

3.1 독립변수 자료

기상자료는 기상청에서 제공하는 서울시 기상월보 및 일보의 자료에 기초하였다. 교통량의 자료는 서울시 유입·출지점의 합산으로 한 달 동안의 서울시 내 총 유동차량을 계산하였다. 다음 Table 1은 총 12개 독립변수에 대한 각 독립변수 별 설명을 나타낸다.

Table 1. Definition of Independent Variables(IV)

IV.	Information	IV.	Information
x_1	Aver. Temp	x_7	Total Rainfall
x_2	Max. Temp	x_8	Total Max. Snowfall
x_3	Min. Temp	x_9	Total Rainy Day
x_4	Aver. Relative Humidity	x_{10}	Max. Continuos Rainy Day
x_5	Min. Relative Humidity	x_{11}	Max. Snowfall
x_6	Aver. Evaporation	x_{12}	Traffic Volume

예측하는 과정에서 각 독립변수의 평균과 분산의 차이가 크고, 기온의 경우 다른 독립변수와 달리 음의 값을 갖을 수 있으며, 여름철에 적설자료의 부재와 같은 강수와 적설자료에서 존재하지 않는 경우의 문제들을 고려하여 독립변수와 종속변수를 -1에서 1사이의 값을 갖도록 하기 위해 정규화작업을 실시하였다. 이러한 정규화 작업은 다음 식 (1)과 같은 방법으로 수행되어 진다.

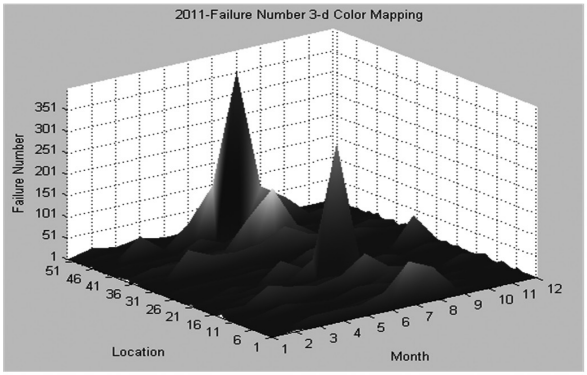
$$X_{n,i} = \frac{2x_{n,i} - x_{\max,i} - x_{\min,i}}{x_{\max,i} - x_{\min,i}} \quad (1)$$

$X_{n,i}$: 정규화 값
n : 관측자료의 순번
i : 독립변수의 순번
$x_{n,i}$: i에 속하는 독립변수의 정규화 작업 전의 값
$x_{\max,i}$: $x_{n,i}$ 중 최대값
$x_{\min,i}$: $x_{n,i}$ 중 최소값

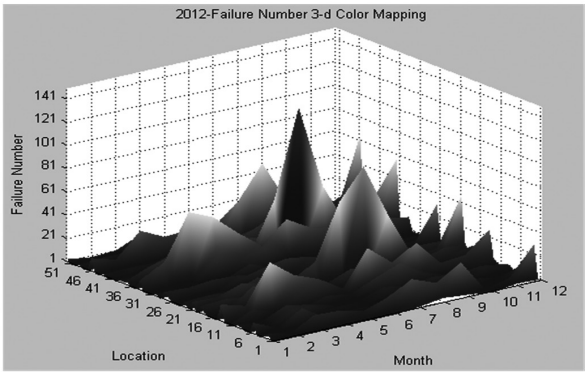
여기서, $X_{n,i}$는 정규화 작업을 거친 정규화값을 의미하며, n과 i는 각각 관측자료와 독립변수의 순번이므로, 총 자료의 수인 36개(총 36개월)은 n에 속하는 순번이 되며, 독립변수의 수인 12까지(총 12개 독립변수)는 i에 속하는 순번이 된다. $X_{n,i}$, $X_{max,i}$ 및 $X_{min,i}$는 i에 속하는 독립변수의 정규화 작업 전 값, 최대값 그리고 최소값을 의미한다. 이러한 정규화 작업은 포트홀의 수에도 같은 방법으로 적용되었다.

3.2 도로 파손 개수

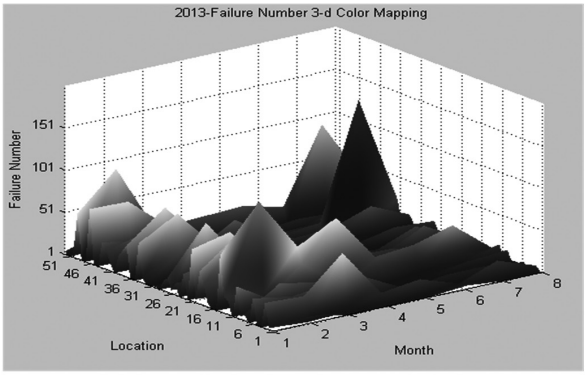
서울시 내에 도로포장의 파손의 수는 포트홀의 수에 초점을 두었으며, 이를 지역적으로 파악하기 위하여 국소지역에 대하여 총 52가지 지역으로 구분함에 따라 다음 Fig. 1.과 같이 3D 칼라맵핑(Color-Mapping)을 시도하였다. 다음 Table 2는 국소 52지역에 대한 설명을 한 표이다.



(a) 2011



(b) 2012



(c) 2013

Fig. 1. Color Mapping for the number of Potholes in Regions

Table 2. Number for Regions regarding roads

No.	Inf.	No.	Inf.	No.	Inf.
1	Gangbyeonnyeok-ro	19	Majang-ro	37	Wangsan-ro
2	Gyeomjae-ro	20	Majo-ro	38	Wangsimni-ro
3	Gosanja-ro	21	Mangu-ro	39	Yongmasan-ro
4	Government Complex-Seoul-ro	22	Myeonmokdong-ro	40	Eungbong-dong-ro
5	Government another Complex-Seoul-ro	23	Myeonmok-ro	41	Imun-ro
6	Gwangnaru-ro	24	Myeonmokcheon-ro	42	Jayang-ro
7	Gucheonmyeon-ro	25	Muhak-ro	43	Jeonnong-ro
8	Gunja-ro	26	Bongujae-ro	44	Jegi-ro
9	Nangye-ro	27	Bonghwasan-ro	45	Jungnangcheon-ro
10	Neungdong-ro	28	Sagajeong-ro	46	Cheonho-daero
11	Neungsan-gil	29	Salgoji-gil	47	Cheonggyecheon-ro
12	Dapsimni-ro	30	Sangbong-ro	48	Cheongsa-naero
13	Dokseodang-ro	31	Seoulsiripdae-ro	49	Hancheon-ro
14	Dongil-ro	32	Seongsu1-ga1-dong-gil	50	Haengdang-ro
15	Dongbu Expressway	33	Sinnae-ro	51	Hwangsan-ro
16	Donge-ro	34	Achasan-ro	52	Hoegi-ro
17	Dongho-ro	35	Achasan-ro 1-gil		
18	Ttukseom-ro	36	Olympic-daero-gil		

위 칼라맵핑을 시도한 것은 월 별 기상자료의 영향을 쉽게 파악하고 해석하기 위함이다.

4. Harmony Search Algorithm(HSA)

하모니 검색 알고리즘(Harmony Search Algorithm, HSA)은 데이터들의 조화로 인하여 그 가중치를 결정하는 최적화 기법 중 하나로써, 이 알고리즘은 즉흥연주를 모방함으로써 전역적 최적화(Global Optimization)적 해를 도출한다는 특징을 가지고 있다는 것 때문에 여러 분야에서 이용되고 있는 알고리즘이다.

4.1 즉흥연주와 하모니 검색 알고리즘

하모니 검색 알고리즘의 원리는 즉흥연주(Improvisation)에 있다. 조화를 이루려고 하는 데이터를

즉흥연주에 비교한다면, 참여연주자가 된다. 참여 연주자들이 기본적인 악보가 없는 상태에서 즉흥연주를 실시한다면, 연주 초기보다 후기에 비교적 더 조화를 잘 갖춘 연주를 할 것이라고 판단할 수 있다. 이는 각 연주자 들이 서로의 음향을 파악하면서 최상의 조화를 이루기 때문이다.

이러한 면에서 판단해 본다면, 데이터는 각 참여 연주자의 개인을 나타내기 때문에, 이를 이용하여 독립변수(Dependent Variable)에 상응하는 결정변수(Decision Variable)를 결정할 수 있다.

즉흥연주를 실시하는 경우 조화라 하면, 어떠한 연주자가 자신의 연주음이 전체 연주에 조화를 잘 이루고 있다면 자신의 연주에 대해 만족하고 연주를 지속하는 반면, 자신의 연주음이 전체적 연주에 조화를 이루지 못한다고 판단한다면, 현재 자신의 연주 방식을 포기하고 바로 새로운 조화방법을 탐색하게 된다. 이러한 방식으로 즉흥연주의 마지막까지 연주를 계속하게 된다.

4.2 하모니 검색 알고리즘의 절차

하모니검색알고리즘의 절차를 이해하기 위해서는 조화기억장치(Harmony Memory, HM), 조화기억고려율(Harmony Memory Considering Rate, HMCR), 피치조정률(Pitch Adjusting Rate, PAR), 조화기억공간(Harmony Memory Size, HMS) 등을 잘 고려해야 한다. HM은 고려하는 가중의 데이터를 지속적으로 반복시켜나가며 조화기억장치에 저장하는 가상적 영역에 속한다. 이를 수행하는 과정에서 HMS를 초기에 지정해 줌에 따라 가상영역공간에 얼마나 많은 수행에 의해 화음을 저장할지 결정해야한다. 다음 식 (2)는 HM과 HMS의 관계를 나타낸다.

$$HM=\begin{bmatrix} x_1^1 & x_2^1 & \cdots & x_n^1 \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{HMS} & x_2^{HMS} & \cdots & x_n^{HMS} \end{bmatrix} \begin{matrix} \Rightarrow & f(x^1) \\ \Rightarrow & f(x^2) \\ & \vdots \\ \Rightarrow & f(x^{HMS}) \end{matrix} \quad (2)$$

x : 각각의 결정변수
$f(x)$: 목적함수
n : 하모니 메모리 사이즈

여기서 x의 값은 각각의 결정변수를 나타내며, f(x)는 목적함수(Objective Function)을 나타낸다. n은 ND(Number of Decision variable)의 수이다.

결국 조화기억공간에는 초기 계획에 있어서 목적함수에서 필요로 하는 가중치의 개수에 목적함수에 대한 열이 하나 더 생김으로써 총 ND+1개의 열로 구성된다.

이를 반복적으로 수행하는 과정에서 무작위로 선택된 수가 기존의 조화기억장치에 있는 데이터와 조화를 이룰지에 대하여 선택하는 확률인 HMCR에 근거하게 된다. 그리고 이에 대하여 기억회상(Memory Consideration)에서 변수에 대한 피치조정(다시 어느 정도로 기존 화음과 피치를 조율할지 정하는)을 PAR의 결정으로 수행한다. 이 결과의 만족정도에 따라 반복수행의 여부를 결정하며, 기존의 해와 새로운 조화에 의한 해를 비교하여 두 값 중 가장 합당한 것을 다시 반영하는 방식으로 반복된다. 아래 단계는 하모니 검색 알고리즘에 대한 흐름순서이다.

Step 1. DV, 목적함수, HMS, HMCR, PAR, ND, 변수 및 반복횟수 한계 선정

가장초기 단계로 목적함수를 결정하게 되면, 이 함수에 대한 ND와 DV는 자연스럽게 결정되어진다고 할 수 있다. DV는 사용가능한 데이터를 각 행렬로 정리를 하여 조화를 이룰 수 있는 준비를 이 단계에서 미리 해 두어야 한다. HMS, HMCR 및 PAR은 사용자의 편의에 따라 선택할 수 있다. 그러나 사실 이는 임의의 수를 적용하는 것이 아니라 적정한 범위가 기존 연구에 의해 제안되어 있으며, 이에 대한 설명은 다음 단계에서 상세히 설명하도록 한다. 변수 및 반복횟수의 한계 제한은 실제 무작위로 선택되는 변수의 경우, 무한대의 수 중 아무것이나 반영하기에는 상당히 큰 오차를 갖기 때문에 각 변수에 대한 임의의 한계 값을 지정해 주어야 한다. 이는 기술적 또는 경험적인 기술인에 의해 수치를 부여해야 한다.

Step 2. 새로운 조화의 생성 단계

하모니검색알고리즘은 무작위로 어떠한 수를 지정하여 그 수의 오차여부 내에서 확률적으로 변수를 조절하는 알고리즘이기 때문에 HMCR 및 PAR의 영향이 매우 중요하다. 그러나 HMCR 및 PAR은 HM내에 저장되어 있는 이전의 값으로부터 다양하게 선택하는 범위와 채택에 있어서 한 변수의 피치조절을 결정하는 것을 의미하기 때문에 초기 무작위 변수의 한계가 매우 중요한 요소라 할 수 있다. 이러한 HMCR과 PAR은 하모니 검색 알고리즘단계에서 다음과 같이 표현될 수 있다.

$$x_i' \Leftarrow \begin{cases} x_i' \in x_i^1, x_i^2 \cdots x_i^{HMS} \rightarrow HMCR \\ x_i' \in X_i \qquad\qquad\quad \rightarrow (1-HMCR) \end{cases} \quad (3)$$

$$x_i' \Leftarrow \begin{cases} Apply\ the\ adjusting?\ yes \rightarrow PAR \\ Apply\ the\ adjusting?\ no \rightarrow (1-PAR) \end{cases} \quad (4)$$

여기서 x_i는 새로운 하모니 벡터의 각각의 요소(i번)으로 표현한 것이며 X_i은 각각의 변수의 가능한 범위를 나타낸 것이다. 예를 들어, HMCR을 70%로 설정한다면, 하모니검색 알고리즘은 HM내에 이전에 저장되어있는 변수에 대해 70%의 확률 내에서 변수를 결정할 것이며, 반대 상

황에서는 30%(100%−HMCR)의 전체적 범위 내에서 변수를 결정할 것이다. PAR은 새로운 백터해에 변화를 어떻게 주고자 하는 값을 의미하기 때문에 만약 20%라면, 20%×HMCR확률내의 이웃한 변수값으로 조정을 실시한다.

이러한 과정을 거쳐 새로운 조화의 해 백터를 생성한다.

Step 3. 적합성의 판단 및 HM Updating

Step 2.에 의해 결정되어진 새롭게 형성된 변수들을 이용하여 결정된 새로운 HM의 결과가 기존의 목적함수와 비교하여 더 조화를 잘 이룬다고 판단하는 경우 새로운 조화를 HM에 저장을 하게 되고, 그렇지 않은 경우 HM에서 이를 배제 시킨다.

Step 4. 반복적인 작업 수행(Step 2., 3)

이러한 반복 작업은 기준에 만족하지 않는 경우 Step 2.로 돌아가서 새로운 조화를 생성하게 되고 Step 3.의 적합성 판단과정을 거치게 된다.

5. 하모니 검색 알고리즘의 적용

본 연구에서 수행하고자 하는 것은 서울시 전체에서 발생하는 도로의 포트홀 개수를 예측하는 것이다. 그러나 하모니 검색 알고리즘에서 사용되는 PAR이나 HMCR의 영향이 예측모델의 미지수를 결정하는 데 매우 중요한 역할을 하기 때문에 PAR과 HMCR의 변화를 고려하여 예측모델을 결정하도록 하였다.

5.1 기본구성사항

목적함수는 실제 발생한 도로의 파손 개수와 하모니검색을 통해 예측된 도로 파손의 개수의 오차를 최소화 시키는 것에 기반한다. 이러한 오차 최소화 함수는 RMSE(Root Mean Squared Error)에 기초하였으며, 다음 식 (5)와 같이 표현되어 진다.

$$f(x_i) = Min(\mathrm{RMSE}) = \mathrm{Min}\left(\sqrt{\frac{\sum_{n=1}^{i}(\mathrm{FN_n}-\mathrm{PN_n})^2}{\mathrm{N}}}\right) \quad (5)$$

FN_n : 실제 파손 개수
PN_n : 하모니검색을 통한 예측 파손 개수

5.2 포트홀 개수 예측 모형

서울시 내에서 발생하는 도로의 파손은 그 개수 및 특성이 매우 불규칙적이며, 그 변화 또한 매달 매우 큰 폭으로 변화하는 등 규칙성이 거의 존재하지 않는다. 이를 확인하기 위해서 데이터에 대한 표준확률곡선(Normal Probability Curve)를 작성하였으며, 그 결과는 다음 Fig. 2와 같다.

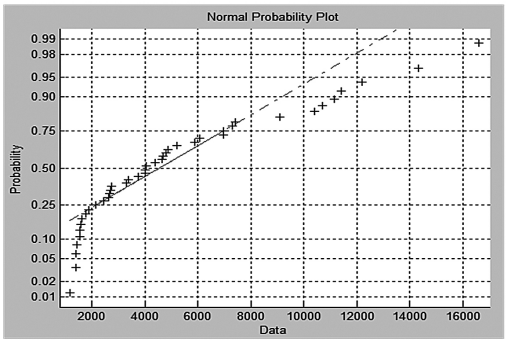


Fig. 2. Normal Probability Curve for Pothole Number

다음 Fig. 2에서와 같이 선형적으로 균형을 이루지 못하고 있으며, 매우 다양한 변화폭을 갖기 때문에 큰 오차가 도출될 것으로 판단하였다. 이 영향을 최소화하기 위하여 식 (1)과 같은 정규화 작업을 실시하였고, PAR과 HMCR은 기존 연구(X.-S. Yang(2009), O. M. Alia, R. Mandava and D. Ramachandram(2009))에서 지정한 경험적 범위(HMCR : 0.7 ~ 0.9, PAR : 0.1 ~ 0.5)에서 0.05간격으로 수행하도록 하였다.

이러한 파손개수의 불규칙적인 고려사항 이외에도 기후변화와 관련된 서울시 전역에 걸친 예측모형을 개발함과 동시에 포트홀에 영향을 미치는 인자를 찾기 위해 예측모델의 구성은 선형방정식에 기초하였으며, 다음 식 (6)과 같이 표현할 수 있다.

$$PN = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8 + \beta_9 X_9 + \beta_{10} X_{10} + \beta_{11} X_{11} + \beta_{12} X_{12} \quad (6)$$

β : 하모니 검색을 통한 비표준화계수
X : 각 독립변수

하모니검색알고리즘의 적용에 있어서 HMCR 및 PAR를 0.9와 0.35로 반영하였다. 2011년, 2012년 그리고 2013년 매달 발생한 파손개수 데이터를 이용해서 예측모형을 구성한 결과, 다음 Fig 3.과 같이 반복작업에 의해 식(5)의 오차를 줄여나갈 수 있었다.

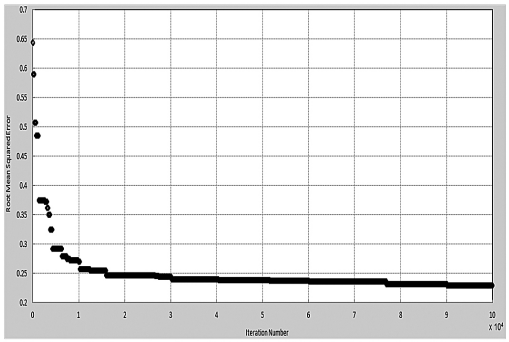


Fig. 3. Reduction of RMSE with Iteration Number

컴퓨터 프로그래밍을 통한 반복작업의 횟수가 100000번 이상인 경우에 오차를 최소화 시키는 것에 한계가 있는 것으로 판단하였다. 하모니 검색 알고리즘에 의해 예측되어진 결과의 적합성을 판단하기 위하여 오차는 식 (4)와 같은 최소화된 RMSE를 활용하였으며, 오차율과 상관계수는 다음 식과 같이 MAE(Mean Absolute Error)와 결정계수(Coefficient of Determination)을 사용하였다.

$$MAE = \frac{100\%}{N}\sum_{n=1}^{36}\left|PN_n - FN_n\right| \tag{7}$$

$$R = \frac{\sum_{n=1}^{36}(PN_n - \overline{PN_n})(FN_n - \overline{FN_n})}{\sqrt{\sum_{n=1}^{36}(PN_n - \overline{PN_n})^2}\sqrt{\sum_{n=1}^{36}(FN_n - \overline{FN_n})^2}} \tag{8}$$

PN: 예측된 포트홀의 수
$\overline{PN}$: 예측된 포트홀의 평균
FN: 관측 포트홀의 수
$\overline{FN}$: 관측 포트홀의 평균

여기서, 피어슨의 적률 상관계수를 제곱하여 결정계수로 활용할 수 있다.

이와 같은 하모니검색알고리즘의 오차를 줄여나간 결과, 예측된 값을 기존 데이터에 비교하여 다음 Table 3.과 같은 결과를 도출하였다.

Table 3. Results of Predictions regarding HMCR and PAR

HMCR	PAR	RMSE	MAE(%)	R^2
0.7	0.1	0.2407	19.7004	0.8883
0.75	0.1	0.2322	19.1595	0.892
0.8	0.1	0.2301	19.3807	0.8932
0.85	0.1	0.2223	18.2997	0.9013
0.9	0.1	0.2175	18.3266	0.9034
0.7	0.15	0.2458	19.5348	0.8807
0.75	0.15	0.2344	19.1901	0.8885
0.8	0.15	0.2172	17.8591	0.9047
0.85	0.15	0.222	18.5889	0.905
0.9	0.15	0.2157	18.0529	0.906
0.7	0.2	0.247	20.155	0.8836
0.75	0.2	0.2281	19.0214	0.8978
0.8	0.2	0.2204	18.4807	0.9013
0.85	0.2	0.2148	18.2646	0.9081
0.9	0.2	0.2219	18.5331	0.9008
0.7	0.25	0.2446	19.3447	0.8851
0.75	0.25	0.2246	18.2305	0.9048
0.8	0.25	0.2263	18.5429	0.8962
0.85	0.25	0.2191	18.5426	0.9069
0.9	0.25	0.2172	18.4407	0.9042
0.7	0.3	0.2422	20.1086	0.8832
0.75	0.3	0.2365	19.9316	0.8899
0.8	0.3	0.2242	18.9054	0.8972
0.85	0.3	0.2252	18.716	0.9019
0.9	0.3	0.2153	17.6421	0.9054
0.7	0.35	0.2443	19.2301	0.8894
0.75	0.35	0.2318	18.2671	0.8908
0.8	0.35	0.2225	18.6265	0.9019
0.85	0.35	0.2193	18.6435	0.9058
0.9	0.35	0.2204	18.435	0.9026
0.7	0.4	0.2432	18.478	0.8773
0.75	0.4	0.2296	19.2848	0.899
0.8	0.4	0.2241	18.5841	0.9021
0.85	0.4	0.2205	18.3959	0.9031
0.9	0.4	0.2193	18.5661	0.9031
0.7	0.45	0.245	19.2088	0.8775
0.75	0.45	0.2375	19.4751	0.8859
0.8	0.45	0.2273	19.6289	0.8956
0.85	0.45	0.2328	18.363	0.8868
0.9	0.45	0.2193	18.3084	0.9027
0.7	0.5	0.2503	19.5951	0.8726
0.75	0.5	0.2331	18.4237	0.8896
0.8	0.5	0.2265	18.3289	0.8938
0.85	0.5	0.2201	18.1939	0.9048
0.9	0.5	0.2203	18.2886	0.9022

Table 3에서 확인할 수 있듯이, RMSE가 가장 적고, 추세의 일치성이 가장 높은 것은 HMCR과 PAR이 각각 0.85, 0.2인 경우이며, MAE가 가장 낮은 것은 HMCR이 0.9이고 PAR이 0.3인 경우이다. 따라서 위 모형 중 가장 적합하다고 판단한 것은 HMCR과 PAR이 각 각 0.85, 0.2인 경우이다. 이 결과로부터 도출된 비표준화 회귀계수는 다음 Table 4와 같다.

Table 4. Value of Unstandardized Coefficient using Harmony Search Algorithm

β_0	0.997585	β_7	0.990397
β_1	0.093697	β_8	−0.10668
β_2	0.435063	β_9	0.103754
β_3	−0.5316	β_{10}	0.012264
β_4	−0.14429	β_{11}	0.995739
β_5	0.021472	β_{12}	0.136142
β_6	0.145801		

위 비표준화계수를 이용하여 예측한 결과는 정규화 작업에 대한 예측값이기 때문에 이를 실제 포트홀의 수로 변환할 필요가 있다. 정규화된 값을 실제 포트홀의 수로 변환하는 방법은 다음 식과 같다. 이를 변환하는 과정에서 예측된 포트홀의 수는 정수로 산출되어야 하기 때문에 소수점이하는 생략하도록 하였다.

$$PN = \frac{Y_n(PN_{\max} - PN_{\min}) + PN_{\max} + PN_{\min}}{2} \tag{9}$$

Y_n : 예측된 정규화 값
$PN_{\max}$: 실제 관측된 포트홀 수의 최대값
$PN_{\min}$: 실제 관측된 포트홀 수의 최소값

실제 포트홀 수로 변환한 값은 다음 Fig. 4와 같으며, 이에 대한 RMSE오차가 1683개소, 추세에 대한 상관도가 90.83%로 확인되었다.

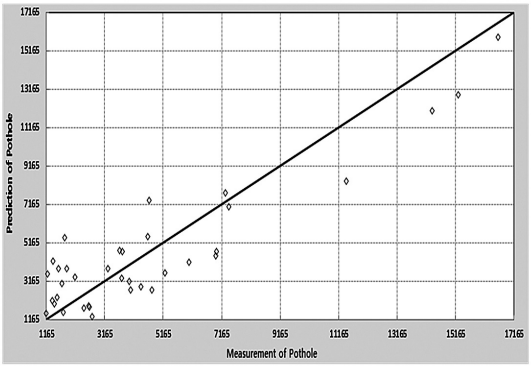


Fig. 4. Results of Predictions using Harmony Search Algorithm(Training Data)

5.3 모델의 과거 자료 적용

위 Fig. 4는 2011년부터 2013년까지 총 36개월의 Training데이터를 이용하여 결정되어진 예측모델의 결과이다. 이 Training데이터를 실제 2009년과 2010년에 적용하여 이 예측모델의 적합성을 검증하고자 하였다. 다음 Fig. 5는 본 연구에서 개발한 예측모델을 적용한 Testing 데이터의 예측결과를 나타낸 것이다.

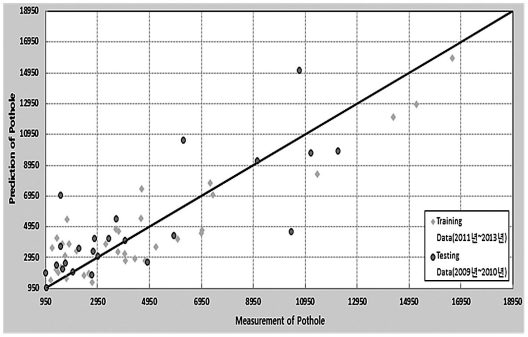


Fig. 5. Results of Predictions using Harmony Search Algorithm(Training and Testing Data)

Training 데이터에 기반한 예측모델을 Testing Set에 적용한 결과 RMSE오차는 1918개소, 추세에 대한 정확성은 85.4%인 것으로 확인되었다.

5.4 영향요소 평가

비표준화회귀계수(Unstandardized Coefficient, UC)는 각 독립변수들의 평균이나 표준편차의 차이로 인해 이를 이용하여 독립변수가 종속변수에 영향을 미치는 정도를 설명하는 것이 불가능하다. 따라서 각 독립변수가 포트홀 발생에 미치는 영향을 판단하기 위해서는 표준화를 시켜줄 필요가 있다. 표준화를 수행하는 과정에서 각 독립변수와 종속변수의 표준편차를 활용하였다. 또한 표준화계수의 값은 각 독립변수가 종속변수에 미치는 상대적 영향력을 나타내지만, 이 영향력의 유의수준이나 공선성을 판단해 주는 역할을 하지 못한다. 따라서 분산팽창요인(Variance Inflation Factor, VIF)을 도출하기 위한 다중공선성분석과 유의확률(P-value)을 계산하기 위한 상관분석을 수행하였다. 단, 표준화 작업은 Testing Set이 예측모델의 구성에 영향을 주지 못하였기 때문에, 예측모델을 구성하는 Training Set에만 적용되어졌다. 다음 Table 5는 각 독립변수에 대한 비표준화계수를 표준화시켜준 결과이다.

Table 5. Results of Statical Analysis for each independent variable

UC	Value of UC	Value of SC	R^2	VIF	P-value
β_1	0.997585	0.11430	0.1623	1.19383	0.344
β_2	0.093697	0.53833	0.1269	1.14532	0.461
β_3	0.435063	−0.62836	0.1846	1.22642	0.281
β_4	−0.5316	−0.14256	0.4928	1.97175	0.002
β_5	−0.14429	0.02160	0.5190	2.07920	0.001
β_6	0.021472	0.13089	0.1303	1.14992	0.448
β_7	0.145801	0.78796	0.6909	3.23556	0.000
β_8	0.990397	−0.04513	0.2522	1.33729	0.138
β_9	−0.10668	0.09868	0.5406	2.17685	0.001
β_{10}	0.103754	0.01078	0.5422	2.18422	0.001
β_{11}	0.012264	0.56710	0.3228	1.47667	0.055
β_{12}	0.995739	0.19986	0.2380	1.31243	0.162

일반적으로, P-value의 신뢰구간의 수준을 95%로 보며, 표준화계수의 값이 -1에서 1사이의 값을 갖고 서로 공선성의 문제가 없는 경우 다항 구성방정식에서 결정계수로 영향력을 판단하기 때문에, 유의수준이 0.005이하면서 결정계수가 상대적으로 높은 평균상대습도와 최소상대습도, 총 강수량, 총 강수일, 최대 연속강수일수가 포트홀 발생에 영향을 미치는 요소로 판단될 수 있다. 그러나 경험적 관점에서 영향요소를 판단할 때 교통이나 온도의 조건 역시 매우 중요한 것으로 판단하여 다음 Fig. 5와 같이 월 별 실제 서울시 전체 포트홀 예측 결과와 측정 결과를 도시화 하였다.

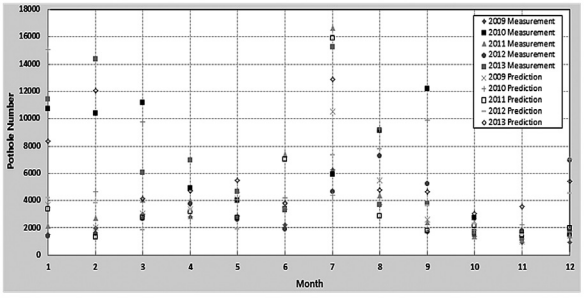


Fig. 6. The Number of Potholes for Prediction and Measurement

6. 결론

본 연구의 목적은 2011년부터 2013년까지 기상자료와 교통량자료의 Training Set을 이용하여 하모니 검색 알고리즘에 기반한 예측모델을 개발하는 것이다. 또한, 예측모델을 개발함과 동시에 실제 다양한 기상자료와 교통량에 대한 독립변수 중 포트홀에 영향을 미치는 인자를 찾고자 하였다. 포트홀 발생에 영향을 미치는 요소를 확인하기 위하여 확률적 방법과 경험적 방법으로 접근하였으며, 본 연구를 수행한 최종적인 결론은 다음과 같다.

(1) BHS(Basic Harmony Search) 알고리즘은 PAR과 HMCR이 각각 0.2, 0.85인 경우 정규화 값의 RMSE오차가 0.2148로 가장 적으며, 0.3, 0.9인 경우 오차발생률인 MAE가 17.64%로 가장 낮은 것을 확인하였다.

(2) 예측모델을 개발함과 동시에 각 독립변수가 포트홀 발생에 영향을 미치는 인자를 확인하기 위해 독립변수에 대한 구성모델은 선형으로 가정하였으며, 이 결과로부터 포트홀 발생에 영향을 미치는 요소를 판단하여 방정식을 재 구성할 경우 예측모델의 적합성이 더 우수할 것으로 판단하였다.

(3) 확률적인 분석면에서는 기상자료 중 상대습도와 강수량이 가장 영향을 많이 미치는 것으로 판단하였다.

(4) 경험적인 면에서는 온도, 강수량, 적설량, 교통량이 포트홀의 발생에 영향력이 클 것으로 판단하였다.

(5) 서울시 내의 도로 중 교통하중의 조건에 따라 파손수가 크게 변하는 것에 반해 본 연구에서는 이를 반영하지 않았다.

이외에도 지역적 특성이나, 도로의 설계수명, 또는 포장도로의 종류 등의 경험적 요소의 중요 부분을 고려하지 않은 서울시 전역에 대한 예측모델을 개발하였다. 따라서 예측모델의 정확성을 높이기 위해 지역적 특성 및 포장 설계수명 등의 상세 정보가 필요한 것으로 판단하였다.

본 연구는 추후 위 결론을 토대로 실험계획법과 포트홀 발생 영향요소의 판단을 이용하여 구성방정식을 보다 합리적으로 결정하도록 하는 연구를 수행할 것이다.

감사의 글

이 논문은 2013년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2011-0030040).

References

Lee, C. , Kim, D. , Mun, S. , and Yoo, P. (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”, Korean Journal of Korea Society of Road Engineers, KSRE, 14(2), pp. 11-17.

Lee, J. , Yu, H. , Kim, Y. , Lee, J. , and Cho, M. (2004). “A Study on the Development of an Automated Pavement Crack Sealer”, Korean Journal of Construction Engineering and Management, KICEM, 5(2), pp. 162-171.

Lee, S. , Mun, S . (2014). “Improving a model for the dynamic modulus of asphalt using the modified harmony search algorithm”, Expert Systems and Applications, SCIE 41(8), pp. 3856-3860.

Mun, S. , and Geem, Z-W. (2009). “Determination of individual sound power levels of noise sources using a harmony search algorithm”, International Journal of Industrial Ergonomics, SCIE, 39, pp. 366-370.

Mun, S. , and Geem, Z-W. (2009). “Determination of viscoelastic and damage properties of hot mix asphalt concrete using a harmony search algorithm”, Mechanics of Materials, SCIE, 41, pp. 339-353.

Mun, S. , and Lee, S. (2011). “Identification of viscoelastic functions for hot-mix asphalt mixtures using a modified harmony search algorithm”. Journal of Computing in Civil Engineering, ASCE, 25, pp. 139-148.

Mun, S. , and Cho, Y. (2012). “Modified harmony search optimization for constrained design problems”, Expert Systems and Applications, SCIE, 39(1), pp. 419-423.

Park, D. , Lee, S. , and Kim, K. (2010). “A Study on the Construction Specification and Quality Assurance Criteria in Clay Paver”, Korean Journal of Construction Engineering and Management, KICEM, 11(6), pp. 111-121.

Suh, Y. , Mun, S. , and Yeo, I. (2010). “Fatigue Life Predictio of Asphalt Concrete Pavement Using a Harmony Search Algorithm”, Journal of Civil Engineering, KSCE, 14(5), pp. 725-730.

Jimoh. Y. A. (2012). “Model and Application for a Pothole in a Flexible Pavement Maintenance Due to Axle Loads”, Espistemics in Science, Engineering and Technology, ESET, 2(1), pp. 1-7.

요약 : 최근 급격한 기후 변화에 관련하여 아스팔트 도로의 손상이 다발적으로 발생하고 있다. 이러한 현상을 해결 및 방지하기 위해서 세계 각국에서 다양한 연구를 수행하고 있다. 이와 관련하여, 본 연구에서는 서울시에서 발생하는 포트홀 수에 대한 예측모형을 개발하였다. 예측모델을 개발함과 동시에 다양한 독립변수 중 실제 포트홀 발생에 영향을 미치는 요소를 파악하기 위해 경험적인 방법과 통계적인 방법을 활용하였다. 예측모형은 BHS (Basic Harmony Search) 알고리즘을 이용하여 결정하였으며, 기후자료와 교통량 자료 및 포트홀 발생 자료를 기반으로 예측이 이루어진다. 하모니 검색 알고리즘을 이용하여 예측모델을 결정하는 과정에서 PAR(Pitch Adjusting Rate)과 HMCR(Harmony Memory Considering Rate)의 영향을 파악하기 위해 이 값을 변화시키며 적합성을 판단하였다. 예측모델은 Training Data(2011년, 2012년 및 2013년 자료)로 인해 구성되며, 결정된 모델의 적합성을 판단하기 위해 Testing Set에 적용하도록 하였다. 기본적인 예측모델의 적합성 판단은 RMSE(Root Mean Squared Error), MAE(Mean Absolute Error), 결정계수(Coefficient of Determination)를 이용하도록 하였다.

키워드 : 하모니검색알고리즘, 하모니메모리, 하모니메모리조정률, 평균절대오차, 피치조정률, 포트홀, RMSE, 점탄성