

# 비례위험모형을 이용한 상수관로의 상대적 파손위험을 분석

## Analysis of Relative Breakage Hazard Rate of Water Mains Using the Proportional Hazards Model

박수완\*, 김정욱\*\*, 임광채\*\*\*, 이형석\*\*\*\*

Suwan Park\*, Jung Wook Kim\*\*, Gwangchae Im\*\*\*, Hyeong Seok Lee\*\*\*\*

### 요 지

본 연구에서는 상수도 배수관로의 내·외부적 특성에 따라 개별관로를 정의하는 방법을 연구대상 지역의 배수관로 파손 데이터베이스에 적용하여 비례위험모형을 구축하였다. 연구에 사용된 자료는 연구대상지역의 배수관로의 제원 및 파손시기를 포함하는 관로 파손데이터베이스, 관로매설지역의 급수인구 및 수압범위에 관한 자료를 포함하는 GRID 데이터베이스와 관로매설지역의 토지개발 정도에 관한 자료를 포함한다. 이러한 자료를 이용하여 관로를 순차적 파손경험에 따라 7개의 생존시간군(STG I ~ VII)으로 구분하고 각 생존시간군에 대한 비례위험모형(Model I ~ VII)을 구축하였다. 이러한 모형을 이용하여 관로의 파손횟수가 증가하는 동안 파손에 영향을 미치는 인자의 변화와 그 효과를 파악하였으며, 또한 추정된 공변수의 위험비율을 분석함으로써 관로의 제원 혹은 매설환경, 급수인구 등에 따른 위험률의 상대적인 변화를 분석하였다. 또한 비례위험모형의 구축과정에서 관로의 파손에 영향을 미치는 공변수의 비례성 가정을 검토하여 시간종속형 공변수를 모형화하였으며, 모형의 이탈잔차(deviance residual)를 분석하여 모형의 적합성을 검토하였다.

본 연구에서 구축된 비례위험모형에 대해 Shoenfeld 잔차를 이용한 스코어 잔차의 변화(score process)를 검토한 결과, Model I 과 Model II 에 대해서는 공변수의 시간종속 효과가 발견되었다. Model I에 대해서는 관로재질과 급수인구의 영향이 시간에 따라 변하며 Model II에서는 급수인구의 영향만이 시간에 따라 변하는 것으로 나타났다. 한편 Model III ~ Model VII 들에 대해서는 공변수의 영향이 시간에 따라 변하지 않는 것으로 나타났다. 각 생존시간군에 대해 관로재질, 토지개발정도, 관로길이 및 급수인구의 변화가 관로의 상대적 누수위험률에 미치는 영향을 상대위험률의 95% 신뢰구간을 고려하여 정량적으로 산정하였고, 시간 종속형 공변수로 모형화된 공변수는 시간에 따른 공변수 영향의 변화를 분석하였다.

순차적 파손사건에 대한 비례위험모형의 구축 결과 생존시간군(STG) I의 기저위험률은 매설 후 대략 450 개월까지는 파손 위험률이 '0'에 가까우나 그 이후로 급격히 증가하다가 매설 후 약 700개월에 이르러서는 약간 감소하고 약 850개월 이후에는 다시 급격히 증가한다. STG II의 기저위험률은 첫 번째 파손 후 약 300 개월이 되면 위험률이 급격히 증가하는 것으로 나타났다. STG III ~ STG VII의 기저위험률은 이차함수의 형태를 띄며, 특히 STG V, STG VI 및 STG VII의 기저위험률은 욕조형 곡선(bathtub curve)의 형태를 가진다. 각 생존시간군의 기저생존함수의 생존확률 '0.5'에 해당하는 기저중간생존시간에 대한 분석으로부터 파손횟수가 많아질수록 순차적 파손사건 사이의 경과시간은 감소하는 것으로 나타났다. 이러한 기저생존시간에 대한 경향은 관로의 파손횟수가 많아질수록 관로의 일반적인 내구성은 감소하기 때문인 것으로 분석된다.

**핵심용어 : 비례위험모형, 배수관로, 상대위험률, 상수도, 파손**

\* 정회원 · 부산대학교 사회환경시스템공학부 조교수 · E-mail : [swanpark@pusan.ac.kr](mailto:swanpark@pusan.ac.kr)

\*\* 비회원 · 성동조선해양(주) 설계연구소 센터 설계연구부 · E-mail : [wjddnrq@naver.com](mailto:wjddnrq@naver.com)

\*\*\* 비회원 · 부산대학교 사회환경시스템공학부 석사과정 · E-mail : [toigch@yahoo.co.kr](mailto:toigch@yahoo.co.kr)

\*\*\*\* 비회원 · 한국수자원공사 밀양댐관리단 · E-mail : [dlgudtjr97@hanmail.net](mailto:dlgudtjr97@hanmail.net)

## 1. 서론

본 연구에서는 상수도 배수관로의 순차적 파손사건에 대한 비례위험모형을 이용하여 연구대상 지역 배수관로의 파손특성을 분석하였다. 연구대상지역의 배수관로의 파손자료와 관로의 파손에 영향을 미치는 자료는 배수관로에 대한 관로의 제원 및 파손시기를 포함하는 관로 파손데이터베이스, 관로매설지역의 급수인구 및 수압범위에 관한 자료를 포함하는 GRID 데이터베이스와 관로매설지역의 토지개발 정도에 관한 자료를 포함한다.

본 연구에서 제시된 비례위험모형은 연구대상 지역의 자료를 바탕으로 박수완 등(2007)의 비례위험모형의 구축을 위한 개별관로 정의 방법을 이용하여 관로의 내·외부적 특성에 따른 개별관로를 정의하는 것으로부터 시작한다. 관로를 특성에 따라 개별관로로 정의한 후 Park(2004)이 제안한 배수관로의 순차적 파손횟수에 따른 비례위험모형 구축 방법에 따라 7 개의 비례위험모형(Model I ~ Model VII)을 구축한 후 각 모형의 생존함수를 추정한다.

개별관로에 대해 공변량을 부여하기 위해서 토지개발정도(DL)는 도시지역(urban land) 및 비도시지역(non-urban land)으로 구분하였고, 관로재질 및 연결방식은 Spun-Rigid 주철관(SR), Spun-Flex 주철관(SF) 및 Pit 주철관(Pit-CI)으로 구분하였다. 또한 GRID에 따른 관내 수압을 공변수로 고려하기 위하여 관내 수압은 PT1, PT2, PT3 및 PT4로 구분하였다. 관로의 길이(L) 및 급수인구(C) 공변수의 값은 공변량의 분포가 정규분포를 따라야 한다는 선형회귀분석의 가정을 만족시키고 통계적으로 유의한 모형의 구축을 위하여 관로의 길이 및 급수인구 공변량들의 자연로그 값을 구한 다음 평균이하의 값과 평균 미만의 값으로 구분하여 각 개별관로에 부여하였다. 따라서 모델링에 고려된 공변수는 토지개발정도(DL), 길이(L) 및 GRID 내의 급수인구(C), 관로재질 및 연결방식(SR, SF)과 수압유형(PT1, PT2, PT3)이며 모든 공변수는 이분변수(binary variable)로 정의되었다. 단 관로재질 및 연결방식은 Model I 과 II 에서는 연속형 변수 TYPE으로 취급하였으며, 또한 Model I에서 길이(L) 및 급수인구(C)에 관한 공변수도 연속형 변수로 채택하였다.

## 2. 비례위험모형의 구축

각 생존시간군의 최종적인 비례위험모형을 구축하기 위하여, 먼저 모든 공변수에 대한 단일분석(univariate analysis)을 통하여 각 변수가 생존율과 어떠한 연관이 있는지를 살펴보고, 생존율에 영향을 준다고 판단되는 변수를 일차적으로 선정하였다. 또한  $-2\log PL$  및 Akaike Information Criterion(AIC)을 계산하여 공변수의 가능한 모든 조합 및 공변수 간의 교호작용을 검토하여 각 생존시간군에 대한 공변수와 그 회귀계수값을 추정하였다. 또한 각 생존시간군에 대해 채택된 공변수들의 비례성 검토를 실시하여 비례성 가정을 위반할 경우에는 그 공변수를 시간종속형변수로 추가적으로 채택하였다. 모형에 채택될 공변수와 그 회귀계수값을 결정한 후 기저생존함수 및 기저위험함수를 추정하면 각 생존시간군에 대한 비례위험모형의 구축이 완료되며, 그 결과는 Table 1과 같다. 모형에 포함될 공변수의 선택과 최대우도 추정법을 이용한 회귀계수의 추정은 통계분석용 소프트웨어인 SAS System을 이용하였다.

Table 1. Estimated Regression Coefficients and Hazard Ratios

Model	Covariate	Parameter	<i>p</i> -value (Pr < Ch-square)	Hazard ratio	95% Hazard ratio confidence limits	
					Lower limit	Upper limit
I	<i>TYPE</i>	1.34242	<.0001	3.828	3.443	4.257
	<i>DL</i>	0.54472	<.0001	1.724	1.543	1.926
	<i>L</i>	0.000497	<.0001	1.000	1.000	1.001
	<i>C</i>	0.01968	<.0001	1.020	1.019	1.021
	<i>DL · L</i>	0.000259	0.0093	1.000	1.000	1.000
	<i>TYPE · C</i>	-0.01284	<.0001	0.987	0.987	0.988
	<i>TYPE · time</i>	-0.02302	<.0001	0.977	0.977	0.978
	<i>C · time</i>	-0.00003	<.0001	1.000	1.000	1.000
II	<i>TYPE</i>	0.17846	<.0001	1.195	1.117	1.279
	<i>DL</i>	0.34444	0.0183	1.411	1.060	1.879
	<i>L</i>	1.70421	<.0001	5.497	4.609	6.556
	<i>C</i>	1.25702	<.0001	3.515	2.309	5.351
	<i>DL · C</i>	-0.531	0.0153	0.588	0.382	0.903
	<i>C · time</i>	-1.6262	<.0001	0.197	0.177	0.218
III	<i>SR</i>	0.25651	0.3334	1.292	0.769	2.173
	<i>SF</i>	-0.06004	0.8321	0.942	0.541	1.640
	<i>L</i>	0.85493	<.0001	2.351	1.934	2.859
	<i>C</i>	-0.16995	0.0679	0.844	0.703	1.013
IV	<i>L</i>	0.72912	<.0001	2.073	1.612	2.666
V	<i>SR</i>	-0.87835	0.0606	0.415	0.166	1.040
	<i>SF</i>	-1.01847	0.0408	0.361	0.136	0.958
	<i>DL</i>	0.64732	0.0572	1.910	0.980	3.723
	<i>L</i>	0.32466	0.0457	1.384	1.006	1.902
VI	<i>L</i>	0.44147	0.0340	1.555	1.034	2.339
VII	<i>L</i>	0.66856	0.0097	1.951	1.176	3.239

각 생존시간군에 대해 추정된 비례위험모형의 적합도를 검증하기 위하여 각 모형의 이탈잔차 (deviance residual)를 산정하였다. 이탈잔차는 기록된 파손시간과 모형을 이용하여 추정되는 파손시간의 기댓값 사이의 관계를 나타내는 것으로서, 모형으로 추정되는 개별관로의 이탈잔차의 절대값이 클수록 기록된 파손시간이 모형으로 추정되는 파손시간의 기댓값과 차이가 많다는 것을 나타낸다. 따라서 이탈잔차는 모형의 이상치(outlier)를 확인하기 위하여 사용하며, 이탈잔차의 절대값이 3을 초과할 경우를 모형의 outlier로 정의한다(Allison, 1995). 이탈잔차는 Model I에서만 35개만 발견되며 다른 Model에서는 이탈잔차가 발견되지 않았다. 따라서 각 생존시간군에 대해 구축된 비례위험모형은 기록된 자료에 대체적으로 적합한 것으로 판단된다.

### 3. 결론

Table 1은 분석대상 관로들의 순차적 파손사건에 대한 공변수의 회귀계수 및 상대위험비율 등을 나타내고 있다. 이러한 결과를 분석함으로써 관로의 파손횟수가 증가하는 동안 파손에 영향을 미치는 인자의 변화와 그 효과를 파악할 수 있으며, 또한 추정된 공변수의 위험비율을 분석함으로써 관로의 제원 혹은 매설환경 급수인구 등에 따른 위험률의 상대적인 변화를 파악할 수 있다.

Table 1에 나타난 바와 같이 파손횟수가 증가할수록 모형에 포함되는 공변수의 개수가 대체적으로 감소함을 알 수 있다. 파손횟수가 6회 7회가 되면 관로의 파손은 재질, 토지개발정도 및 급수 인구로는 관로의 파손을 설명할 수 없으며, 관로의 길이만이 관로의 파손확률에 영향을 미치는 것으로 분석된다.

한편 관내수압은 수압유형에 따른 생존경험분석 및 상호작용 검토 등에서 통계적으로 유의하지 않은 것으로 나타났으며, Collett(2003)의 공변수 선택 방법에 의해서도 유의하지 않아 모든 모형에서 사용되지 않았다. 이러한 결과는 관내 수압의 절대적인 크기가 파손에 영향을 미치기보다는 수압의 급격한 변화가 관로의 파손에 영향을 줄 것으로 예상되나 본 연구에서는 이러한 자료가 사용되지 못했기 때문에 사료된다. 각 모형에서 최종적으로 선택된 공변수와 상호작용 변수 및 시간종속변수의 위험비율에 대한 분석은 다음과 같다.

1. Model I에서의 TYPE(Pit-CI=0, SR=1, SF=2)은 매설 초기 위험비율은  $Pit-CI=1 < SR=3.828 < SF=7.656$  으로 Pit-CI의 위험률이 가장 낮으나, TYPE의 시간종속 효과로 인하여 매설 후 시간이 경과함에 따라 SR, SF의 위험률이 Pit-CI에 비해 상대적으로 감소하는 것으로 나타났다. 또한 TYPE의 시간종속 효과로 인하여 매설 후 대략 5년을 기점으로 위험률은  $SF < SR < Pit-CI$ 의 순서로 변화할 것으로 추정된다.
2. Model II에서의 관종별(TYPE) 위험비율은  $SF < Pit-CI < SR$ 의 순서로 나타난다.
3. Model I과 Model II의 급수인구의 시간종속 효과로 인한 위험비율을 보면 매설초기에는 급수인구가 많은 부분의 관로의 위험률이 높은 것으로 나타나고 있으나, 시간이 지남에 따라 급수인구에 의한 위험률은 감소하는 것으로 분석된다. 이는 급수인구가 많은 지역의 관로는 그렇지 않은 지역보다 수압이 높고, 수압의 증감이 클 것으로 예상된다. 이러한 원인들에 의해 매설초기 관로의 파손위험률이 비교적 높은 것으로 판단되며, 시간이 지남에 따라 관로의 조인트 부위 등이 안정화됨에 따라 위험비율이 감소하는 것으로 보인다.
4. 토지개발정도를 나타내는 공변수 DL의 위험비율의 신뢰구간을 종합적으로 분석하면 도시지역의 파손 위험률이 비도시지역에 비해 높은 것으로 판단된다.
5. 관로의 길이를 나타내는 공변수 L의 위험비율의 평균값은 약 2.00으로 나타나고 있으며, 전체적으로 파손횟수가 증가함에 따라 L의 위험비율은 대체적으로 감소하는 경향을 보인다. 이러한 경향은 파손횟수가 많지 않을 때에는 관로의 길이가 긴 만큼 파손이 일어날 물리적인 범위가 크므로 파손 위험률이 높으나 파손횟수가 많아질수록 파손위험률에 대한 길이의 상대적인 영향이 감소하는 것으로 사료된다.
6. GRID내의 급수인구를 나타내는 공변수 C는 Model I과 Model II에서 급수인구가 많을수록 파손위험률이 큰 것으로 나타났으며, Model III에 대해서는 위험비율의 신뢰구간이 1을 포함하고 있으므로 이에 대한 위험률의 차이가 있다고 단정할 수 없는 것으로 보인다. 급수인구와 관로 파손간의 상관관계는 인구집중에 수반되는 교통량의 증가가 관로 파손위험을 증가시키는 것으로 추론할 수 있으나 이에 대해서는 추가적인 연구가 필요할 것으로 사료된다.

## 참 고 문 헌

- 박수완, 김정옥, 전환돈(2007). 상수도 배수관로의 특성에 따른 개별관로 정의 방법을 이용한 파손사건 사이의 비례위험모델링, 수질보전, 한국물환경학회, 제23권 제1호, pp. 87-96.
- Allison, P. D.(1995). *Survival Analysis using SAS: A Practical Guide*, Cary, U.S.A., SAS Institute, pp. 173-174.

- Collet, D.(2003). *Modelling Survival Data in Medical Research*, Chapman & Hall/CRC, Boca Raton, U.S.A., pp. 80-87.
- Park, S.(2004). Identifying the Hazard Characteristics of Pipes in Water Distribution Systems by using the Proportional Hazards Model: 1. Theory, *KSCE Journal of Civil Engineering*, Vol. 8, No. 6, pp. 663-668.