

■ 論 文 ■

## 철도건널목 사고요인 분석에 관한 연구

A Study on Crash Causations for Railroad-Highway Crossings

**오 주 택**

(교통개발연구원 책임연구원)

**신 성 훈**

(교통개발연구원 연구원)

**성 낙 문**

(교통개발연구원 책임연구원)

**박 동 주**

(공주대학교 건설환경공학부 부교수)

**최 은 수**

(철도기술연구원 선임연구원)

### 목 차

- |   |  |
|---|--|
| <p>I. 서론</p> <ol style="list-style-type: none"> <li>1. 연구의 배경 및 목적</li> <li>2. 연구의 범위 및 절차</li> </ol> <p>II. 기존 문헌고찰</p> <p>III. 국내 철도건널목 사고현황</p> <ol style="list-style-type: none"> <li>1. 국내 철도건널목 사고발생현황</li> <li>2. 국내 철도건널목 사고의 사고발생 원인분석</li> </ol> <p>IV. 모델 개발을 위한 변수자료의 수집</p> <ol style="list-style-type: none"> <li>1. 사고자료 수집</li> </ol> | <ol style="list-style-type: none"> <li>2. 교통량 및 주변환경자료 수집</li> </ol> <p>V. 모델개발 및 검증방법</p> <ol style="list-style-type: none"> <li>1. 포아송 회귀 모델</li> <li>2. 음이항 회귀 모델</li> <li>3. 모델검증</li> </ol> <p>VI. 철도건널목 사고빈도 예측모델</p> <ol style="list-style-type: none"> <li>1. 철도건널목 사고빈도 예측모델 개발</li> <li>2. 결과분석</li> </ol> <p>VII. 결과 및 향후 연구과제<br/>참고문헌</p> |
|---|--|

Key Words : 철도건널목, 안전, 사고모델, 포아송 회귀분석, 과분산

### 요 약

철도건널목사고는 일반 도로사고보다는 상대적으로 사고의 발생도는 적으나, 사고의 심각도면에서는 대형사고로 연결될 수 있다는 점에서 철도건널목에서의 안전성 확보는 결코 소홀히 다루어질 수 없다. 본 연구에서는 건널목 사고모델을 통해 건널목 사고에 영향을 미치는 요인들을 분석하고 이를 정량화하여 교차로 안전성 향상에 기여하고자 한다. 본 연구에서는 건널목사고 분석에 있어서 사고데이터 특성을 고려하여 비선형 회귀분석 중 적정 모델식을 적용하였다.

철도건널목 관련 변수들을 이용하여 분석한 결과, 분산값이 0에 가까운 값을 나타내어 포아송 회귀분석이 적합한 것으로 나타났다. 또한 본 연구에서는 주 모델과 후보모델 통해 건널목 사고에 영향을 미치는 7개의 주요설명변수들을 규명했는데, 그 변수들은 차량 교통량과 철도교통량, 상업지역, 제어거리, 경보시간차, 건널목유형, 과속방지턱으로 분석되었다.

Railroad crossing crashes are fewer than road crashes, but with regard to crash severity, they can be serious injury crashes. There should be, therefore, enormous efforts to increase the safety of railroad crossings. The objective of this paper is to identify and understand factors associated with railroad crossing crashes. Statistical models are used to examine the relationships between crossing accidents and geometric elements of crossings.

The results show the Poisson model is the most appropriate method for the crossing accidents, because overdispersion was not observed. This study identifies seven significant factors associated with railroad crossing crashes through the main and variant models. With regard to explanatory factors on crossing safety, the total traffic volume, daily train volume, presence of commercial area around crossings, distance of train detector from crossings, time duration between the activation of warning signals and gates, crossing types, and speed hump were found to affect the safety of railroad crossings.

# 1. 서론

## 1. 연구의 배경 및 목적

철도건널목사고는 일반 도로사고보다는 상대적으로 사고의 발생빈도는 적으나, 대형사고로 연결될 수 있다는 사고의 심각도면에서 철도건널목에서의 안전성 확보는 결코 소홀히 다루어질 수 없다. 그러나 지난 수십년간 이루어진 도로의 안전성 향상을 위한 노력들에 비해 철도건널목에서는 체계적인 안전성 분석이 상당히 미비한 실정이다.

외국의 선진국에 비해 국내의 경우, 철도건널목에서의 안전성 분석은 신뢰할 수 있는 통계적인 방법론에 의존하기보다는 안전전문가에 의한 주관적인 방법에 의해 안전성 평가 이루어지는 경우가 종종 발생하였다. 사고분석에 있어서 때로는 사고의 원인을 분석하는 전문가가 사고의 근본적인 원인을 이해하고 있다는 측면에서 안전 분야에서 중요한 역할을 하지만, 보다 체계적인 사고원인의 분석을 위해서는 정량화된 원인과 결과의 관계를 규명하는 연구가 필요하다. 특히 국내 철도건널목의 경우, 사고의 원인을 운전자의 과실에 두는 경향이 많은 반면, 시설운영 및 관리의 문제점을 분석하고 개선방향을 도출해내는 연구가 도로 등 다른 안전 관련 영역들에 비해 상대적으로 부족한 실정이다.

본 연구에서는 철도건널목사고의 원인을 분석하기 위해 지난 5년간 국내에서 일어난 철도건널목사고, 철도건널목을 구성하고 있는 건널목제어시스템, 그리고 건널목기하구조 등에 대한 자료를 수집하여, 철도건널목에서 발생하는 차량 대 열차사고의 원인과 결과를 체계적으로 분석하고자 한다. 이를 위해 철도건널목사고가 산발적으로 일어나는 점을 고려하여 이들 사고를 설명할 수 있는 포아송 회귀분석을 우선 시도하고, 건널목사고의 분포가 포아송 분포를 따르지 않을 경우 포아송 회귀분석 이외의 다른 사고분석방법들도 시도한다.

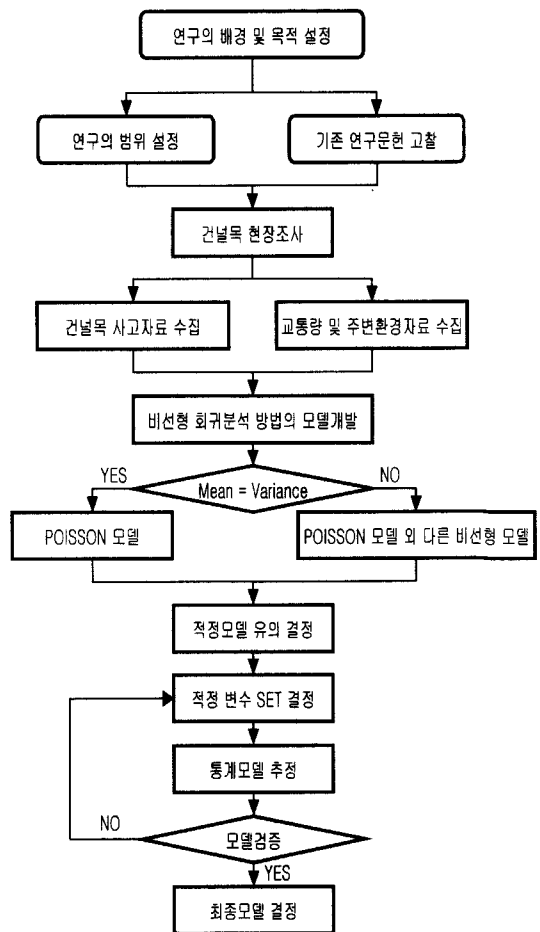
본 연구의 결과를 통해 그동안 운전자의 과실에 비중을 둔 기존의 철도건널목 안전성 분석에 대한 관점을 운전자의 과실을 유발시키는 설명요소들에 대한 체계적이고 정량화된 분석으로 전환시킴으로써 보다 안전한 철도건널목 건설에 이바지 하고자 한다. 또한 본 연구에서는 사고예측모델 개발을 통해 철도건널목에서 발생할 수 있는 사고빈도를 예측함으로써 철도건널목의 안전성을 정량적으로 평가할 수 있는 기초 자료를 제공하

고자 한다.

## 2. 연구의 범위 및 절차

본 연구에서는 최근 5년간(1998년~2002년) 철도건널목의 각종 안전설비 및 시설이 변하지 않은 건널목에서 수집된 자료만으로 연구범위를 한정하였다. 이들 건널목은 수도권을 기준으로 2~3시간 거리에 위치한 건널목 중 162개를 선정하여 조사하였다.

연구방법에 있어서는 stratified random sampling 방법을 적용하여 철도건널목에서 발생할 수 있는 사고와 철도건널목의 차량 및 열차교통량, 주변 환경과의 관계를 포아송 회귀분석과 기타 비선형 회귀분석을 이용하여 통계분석을 한 후 최종모델을 결정하였다. <그림 1>은 본 연구의 수행 흐름도를 도시화한 것이다.



<그림 1> 연구수행절차

## II. 기존 문헌고찰

철도 건널목 사고빈도 예측모델과 관련된 연구는 건널목 안전관련 연구가 활발히 이루어지고 있는 미국과 유럽 등지에서는 비교적 많은 연구가 수행되고 있다. 반면, 국내에서는 강승규(1995) "철도건널목의 사고예측모형 개발에 관한 연구"가 이루어졌으며, 철도청(2003)은 "철도건널목 위험요인 분석 및 개량방안 연구"를 통하여 건널목 구조의 위험요인과 교통량 등 교통환경 위험요인 분석을 하였으나 국내의 실정에 맞는 사고빈도 예측모델을 개발하여 분석을 한 것이 아니라 기존 국외의 연구를 바탕으로 분석이 이루어지는 등 전반적으로 건널목 사고빈도 예측과 관련된 연구는 매우 미흡한 실정이다.

국외의 연구로는 버지니아주 교통연구협회에서 연구 평가한 적이 있으며 철도건널목 위험요소에 대해 예상할 수 있는 Peabody Dimmick 모델과 New Hampshire Index 모델, NCHRP 50 Method, US DOT 모델 등이 있다. 국·내외에서 개발된 철도건널목 사고빈도 예측모델들의 일반적인 특징 및 식은 다음과 같다.

먼저 강승규(1995)의 모델은 1종 건널목에서는 안내원의 유무가 사고확률을 결정하는 가장 큰 요인으로 조사되었고, 2종 건널목에서는 열차속도, 시거, 종단구배차, 교차각도, 철도와 도로교통량 및 조명설비 등이 사고확률을 결정하는 요소로 분석되었다.

$$V_{1n(1)} = k_0 + k_f F + k_t T + k_s S + k_a A + k_g G + k_v V + k_n N$$

$$V_{1n(2)} = k_0 + k_v V + k_a A + k_s S + k_t T + k_g G + k_l L$$

여기서,  $V_{1n(1)}$  : 1종 건널목의 사고확률 모형값

$V_{1n(2)}$  : 2종 건널목의 사고확률 모형값

- $k_0$  : 상수
- $k_f$  : 안내원유무에 대한 계수
- F : 안내원유무(1=유, 0=무)
- $k_v$  : 철도와 도로교통량의 계수
- V : 철도와 도로교통량 합외 로그값
- $k_a$  : 선로와 도로의 교차각도의 계수
- A : 직각교차에 대한 상대적인 교차각도
- $k_n$  : 선로수의 계수

- N : 선로수
- $k_s$  : 시거에 대한 계수
- S : 시거(m)
- $k_t$  : 통과열차의 속도계수
- T : 통과열차의 최고속도 등급(1~6)
- $k_g$  : 선로와 접근도로의 구배차에 대한계수

Peabody Dimmick 모델은 철도건널목에서 5년간 발생한 사고이력자료를 기반으로 작성되었으며, 5년간의 사고발생 건수를 예측하는데 적용하였다. 안전설비계수는 해당 건널목에 설치된 시설의 특성에 따라 산출하였다.

$$I = 1.28 \frac{V^{0.170} \times T^{0.151}}{P} + K$$

- 여기서, I : 5년간 예측사고건수(건)
- V : 연평균 일일교통량, AADT(대/일)
- T : 평균 일일철도교통량(회/일)
- P : 교통안전설비계수
- K : 추가보정계수

New Hampshire Index는 안전설비 계수를 여러 가지로 변경하여 적용하는 것이 특징이다. 이 식은 현재 주마다 열차의 속도, 선로수, 차량속도, 표지판 거리, 교차 각, 인접 교차로 수, 인구, 학교버스 수, 승객 수 등 각기 다른 특성을 살려 사용하고 있다.

$$HI = V \times T \times P_f$$

- 여기서, HI : 위험지수
- V : 연평균 일일교통량, AADT(대/일)
- T : 평균 일일철도교통량(회/일)
- $P_f$  : 안전설비계수

NCHRP 50 Method는 철도건널목을 횡단하는 도로교통량과 열차교통량 및 안전설비계수를 사용하여 사고빈도를 예측하며, NCHRP 50 Method 보고서에 제시된 위험지수는 복합식으로 표현될 수도 있고 표와 그래프를 사용하여 계수를 가져오는 간이식으로 사용할 수 있다.

$$EAF = A \times B \times T$$

여기서,  $EAF$  : 사고발생빈도

$A$  : 도로교통량 계수

$B$  : 안전설비 계수

$T$  : 평균 일일철도교통량(회/일)

US DOT 모델은 두 가지의 독립적인 계산을 하여 사고를 예측하며, 건널목의 특성을 고려하여 초기 예측식을 제공한다. 각 건널목에서 일정 기간동안 발생한 실제사고발생 이력을 사용하는 것이 특징이며, 향후 년간 발생하는 사고가 계산에 적용한 기간에 걸쳐 평균 사고발생 이력비율과 같을 것이라는 가정하였다.

$$A = \frac{T_0}{T_0 + T} \times a + \frac{T}{T_0 + T} \times \frac{N}{T}$$

여기서,  $A$  : 최종 예측사고건수(건/년)

$a$  : 초기 예측사고건수(건/년)

$N/T$  : 연간 사고이력 건수,

$N$ 은  $T$ 년간 관측된 사고건수

$T_0$  : 가중계수  $T_0 = \frac{1.0}{0.05 + a}$

### III. 국내 철도건널목 사고현황

#### 1. 국내 철도건널목 사고발생현황

최근 5년간(1998년~2002년) 발생한 철도건널목 사고를 보면 <표 1>과 같이 1998년 116건 발생되었던 것이 2002년에는 57건으로 매년 감소하는 추세이다. 철도건널목 사고는 일반 교통사고의 비율에 비해 사고건수는 상대적으로 적지만 사고 심각도를 고려할 때 대형사고를 유발할 수 있는 확률은 더 높다고 할 수 있다.

건널목 종별 사고 발생 건수를 비교해 보면 최근 5년간의 연평균사고율을 볼 때 전체 80.4건 중 1종 건널목 75건, 2종 건널목 3.2건, 3종 건널목 2.2건의 사고가 발생하였다. 건널목 사고 사상자수는 사망 78명, 중상 88명, 경상 98명, 피해 없음이 243명인데, 피해 없음의 사고자는 대부분이 열차 접근 직전에 탈출 및 도주 건수이다. 대부분의 사고가 1종 건널목에서 발생하였는데, 이는 1종 건널목의 차량통행량이 많은 것에 기인된다고 볼 수 있다.

(표 1) 최근 5년간 국내 철도건널목 사고발생현황

| 구분      |    | 연도    |       |       |       |       |    |
|---------|----|-------|-------|-------|-------|-------|----|
|         |    | '98   | '99   | '00   | '01   | '02   |    |
| 건널목 수   | 1종 | 1,477 | 1,719 | 1,711 | 1,651 | 1,611 |    |
|         | 2종 | 290   | 25    | 19    | 21    | 32    |    |
|         | 3종 | 120   | 100   | 106   | 104   | 101   |    |
| 총 건널목수  |    | 1,887 | 1,844 | 1,836 | 1,776 | 1,744 |    |
| 사고발생건수  |    | 116   | 95    | 75    | 59    | 57    |    |
| 종별구분    | 1종 | 104   | 88    | 72    | 55    | 56    |    |
|         | 2종 | 9     | 2     | 1     | 3     | 1     |    |
|         | 3종 | 3     | 5     | 2     | 1     | -     |    |
| 사상자 발생  | 사망 | 1종    | 20    | 21    | 8     | 12    | 13 |
|         |    | 2종    | 2     | -     | 1     | -     | -  |
|         |    | 3종    | 1     | -     | -     | -     | -  |
|         | 계  |       | 23    | 21    | 9     | 12    | 13 |
|         | 중상 | 1종    | 24    | 18    | 14    | 10    | 10 |
|         |    | 2종    | 8     | -     | 1     | -     | -  |
|         |    | 3종    | 2     | -     | 1     | -     | -  |
|         | 계  |       | 34    | 18    | 16    | 10    | 10 |
|         | 경상 | 1종    | 33    | 14    | 22    | 5     | 6  |
|         |    | 2종    | 2     | 5     | 2     | 3     | -  |
|         |    | 3종    | 2     | 2     | -     | 2     | -  |
|         | 계  |       | 37    | 21    | 24    | 10    | 6  |
| 피해 없음   | 1종 | 53    | 59    | 40    | 32    | 46    |    |
|         | 2종 | 7     | -     | -     | 1     | 1     |    |
|         | 3종 | 1     | 2     | 1     | -     | -     |    |
| 계       |    | 61    | 61    | 41    | 33    | 47    |    |
| 총 사상자 수 |    | 155   | 121   | 90    | 65    | 76    |    |

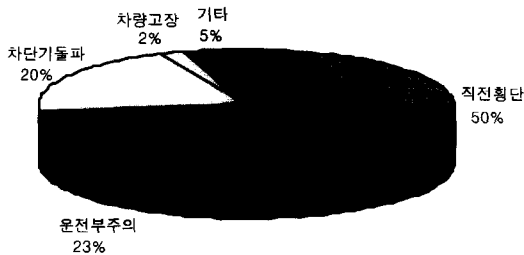
#### 2. 국내 철도건널목 사고의 사고발생 원인분석

최근 5년간의 건널목 사고발생 원인을 보면 <표 2>, <그림 2>와 같다. 건널목 차단직전 횡단과 차단기돌파로 인한 사고 등의 범규위반 관련 사고가 전체사고의 69.6%로 가장 높은 비중을 차지하고 있으며, 운전부주의로 인한 사고 23.4%, 차량고장으로 인한 사고 2.0% 순으로 발생하였다.

특히 건널목 차단직전 자동차가 횡단하는 경우 주로 열차 접근에 의한 경보직후 차단기가 하강시까지의 시간(5~8초)에 건널목을 통과하려다가 진행방향의 차단기 하강으로 인하여 건널목에 갇혀서 열차와 충돌하는 사고가 대부분 이었다. 또한 운전부주의로 인한 사고의 원인으로는 건널목을 통과하는 자동차 운전자의 운전미숙으로 건널목 통과 중 차량시동정지, 보판 이탈 등의 사고가 대부분이었다.

〈표 2〉 국내 철도건널목 사고발생 원인분석

| 구분             | 연도    | '98  | '99 | '00 | '01 | '02 | 합계  | 비율     |
|----------------|-------|------|-----|-----|-----|-----|-----|--------|
|                |       | 직진횡단 | 64  | 56  | 35  | 25  |     |        |
| 사고<br>원인<br>구분 | 운전부주의 | 30   | 17  | 20  | 11  | 16  | 94  | 23.4%  |
|                | 차단기돌파 | 16   | 19  | 17  | 15  | 14  | 81  | 20.1%  |
|                | 차량고장  | 4    | 1   | -   | 1   | 2   | 8   | 2.0%   |
|                | 기타    | 2    | 2   | 3   | 7   | 6   | 20  | 5.0%   |
| 총 사고건수         |       | 116  | 95  | 75  | 59  | 57  | 402 | 100.0% |

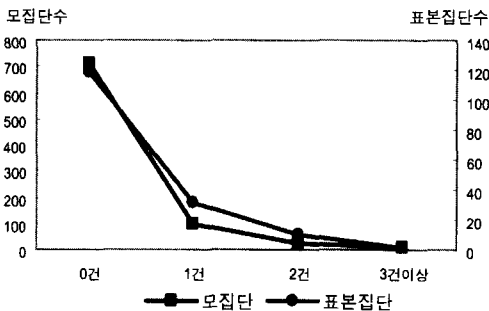


〈그림 2〉 국내 철도건널목 사고발생 원인별 비율

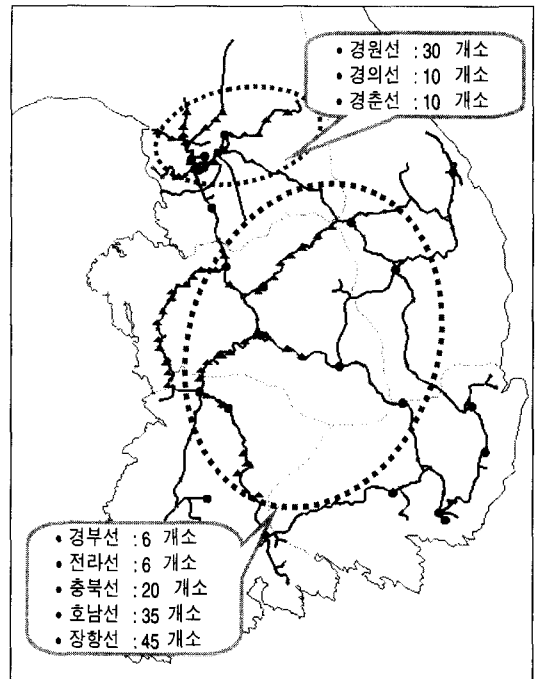
#### IV. 모델 개발을 위한 변수자료의 수집

##### 1. 사고자료 수집

본 연구의 현장조사대상건널목은 전체 1,611개소의 1종 건널목 중 최근 5년간(1998년~2002년) 철도건널목의 각종 안전설비 및 시설이 변하지 않은 건널목 840개를 모집단으로 선정하였다. 이들 건널목 중 수도권을 기준으로 2~3시간 거리에 위치한 건널목 162개를 무작위로 선정하여 조사하였다. 선정된 162개의 건널목의 분포가 고르게 이루어 졌는지 모집단과 선정된 표본집단을 사고건수와 비교하여 살펴보면 〈그림 3〉과 같이 고르게 분포된 것을 알 수 있다. 노선별 현장조사 대상 건널목의 현황은 〈그림 4〉와 같다.



〈그림 3〉 모집단과 표본집단의 분포



〈그림 4〉 조사대상 건널목 분포 현황

##### 2. 교통량 및 주변환경자료 수집

조사대상 건널목의 교통량 및 주변환경자료 수집은 크게 두 가지로 구분된다. 첫번째는 건널목 관리소에서 보관하고 있는 건널목 카드에 기입된 내용의 확인이며, 두번째는 건널목카드에 없는 항목 중 사고요인

분석에 필요하다고 여겨지는 사항들에 대한 조사이다.

차량 및 철도 교통량 자료 및 주변환경에 관한 일반적인 자료는 건널목 카드를 이용하여 수집하였다. 이외에 건널목 카드에 기입되어 있지 않는 항목은 현장에서 직접 조사·수집하였으며, 기존 건널목카드의 기입 내용을 재차 확인 및 실측을 실시하였다. 변수자료 수집 항목 및 통계특성은 〈표 3〉과 같다.

#### V. 모델개발 및 검증방법

철도건널목 사고빈도 예측식을 개발하는데 있어서 사용된 통계적 방법은 포아송 회귀분석과 음이항 회귀 분석을 사용하였으며, 기본적인 통계학적 배경과 일반 모델식은 다음과 같다.

(표 3) 철도건널목 사고모델 개발을 위한 주요변수들의 통계특성

| 변수                 | 빈도수         | 평균     | 최소값   | 최대값     | 중간값    |
|--------------------|-------------|--------|-------|---------|--------|
| 차량교통량(ADT)(차량대수/일) | 162         | 4617   | 10    | 61,199  | 671    |
| 철도교통량(열차대수/일)      | 162         | 70     | 32    | 203     | 57     |
| 선로 형태              | 162         |        |       |         |        |
| 1:단선               | 100(61.73%) |        |       |         |        |
| 2:복선               | 62(38.27%)  |        |       |         |        |
| 도로포장상태             | 162         |        |       |         |        |
| 1:아스콘              | 112(69.14%) |        |       |         |        |
| 2:비포장              | 6( 3.70%)   |        |       |         |        |
| 3:콘크리트             | 44(27.16%)  |        |       |         |        |
| 도로 차선수             | 162         | 1.57   | 1.00  | 6.00    | 1.00   |
| 도로폭(m)             | 162         | 5.91   | 2.00  | 37.5    | 4.40   |
| 건널목 폭(m)           | 162         | 5.93   | 2.50  | 39.70   | 4.15   |
| 주거                 | 162         |        |       |         |        |
| 0:없음               | 109(62.28%) |        |       |         |        |
| 1:있음               | 53(32.72%)  |        |       |         |        |
| 농업(신)              | 162         |        |       |         |        |
| 0:없음               | 37(22.84%)  |        |       |         |        |
| 1:있음               | 125(77.16%) |        |       |         |        |
| 상업                 | 162         |        |       |         |        |
| 0:없음               | 149(91.98%) |        |       |         |        |
| 1:있음               | 13(8.02%)   |        |       |         |        |
| 학교인접거리(m)          | 162         | 3.21   | 0.00  | 200.00  | 0.00   |
| 경보기종류              | 162         |        |       |         |        |
| 1:직립형              | 126(75.93%) |        |       |         |        |
| 2:현수형              | 36(24.07%)  |        |       |         |        |
| 제어거리(m)            | 162         | 827.36 | 0.00  | 1500.00 | 920.00 |
| 경보 시간차(초)          | 162         | 25.46  | 0.00  | 232.00  | 20.00  |
| 과속방지턱              | 162         |        |       |         |        |
| 0:없음               | 134(82.72%) |        |       |         |        |
| 1:있음               | 28(17.28%)  |        |       |         |        |
| 미끄럼방지턱             | 162         |        |       |         |        |
| 0:없음               | 146(90.12%) |        |       |         |        |
| 1:있음               | 16(9.88%)   |        |       |         |        |
| 안내원 근무             | 162         |        |       |         |        |
| 0:없음               | 126(77.78%) |        |       |         |        |
| 1:있음               | 33(22.22%)  |        |       |         |        |
| 건널목 교차각(도)         | 162         | 83.40  | 43.00 | 160.00  | 90.00  |
| 인접교차로              | 162         |        |       |         |        |
| 0:없음               | 88(54.32%)  |        |       |         |        |
| 1:있음               | 74(45.68%)  |        |       |         |        |
| 건널목 유형             | 162         |        |       |         |        |
| 1:Y형               | 29(17.90%)  |        |       |         |        |
| 2:I형               | 43(26.54%)  |        |       |         |        |
| 3:S형               | 23(14.20%)  |        |       |         |        |
| 4:F형               | 12(7.41%)   |        |       |         |        |
| 5:T형               | 55(33.95%)  |        |       |         |        |

## 1. 포아송 회귀 모델

사고의 발생은 비교적 드문 이산형태의 분포를 나타

내므로 포아송 분포를 사용하는 것이 일반적이다. 건널목에서의 사고건수  $\hat{Y}$ 가 포아송 분포를 따른다는 가정하에  $i$  번째 철도건널목에서  $m$ 개의 변수에 의해 사

고가 발생할 확률의 일반식은 식(1)과 같다.

$$\begin{aligned} \widehat{Y}_i &= \exp(\beta_0 X_0 + \beta_1 X_1 + \beta_2 X_2 + \dots \\ &\quad + \beta_m X_m) \\ &= \exp\left(\sum_{j=0}^m \beta_j X_j\right) \end{aligned} \quad (1)$$

여기서,  $\beta_j$  : 회귀추정계수

포아송 회귀 모델의 경우, 모델계수는 Maximum-likelihood 방법을 통해 계산되었으며 식(2)와 같다.

$$\begin{aligned} L(\beta) &= \prod_{i=1}^n f_i(Y_i) \\ &= \prod_{i=1}^n \frac{[\mu(X_i, \beta)]^{Y_i} \exp[-\mu(X_i, \beta)]}{Y_i!} \end{aligned} \quad (2)$$

만약 모델이 자료와 일치된다면 위의 식은 가능한 최대의 값을 가지게 된다. 이런 상황은 식(3)처럼  $\mu_i$  가  $y_i$  로 대체되면서 나타나게 된다.

$$\frac{\partial \log_e L(\beta)}{\partial \beta} = \sum_i (-\mu_i + y_i) X_i = 0 \quad (3)$$

## 2. 음이항 회귀 모델

음이항 회귀분석은 각 분포의 분산이 같아야한다는 포아송 회귀분석의 제약조건을 완화시키는데 사용될 수 있다. 사고예측모델에 있어서 포아송 회귀분석은 추정된 자료의 평균을 실제자료의 분산이 초과하면 과분산(Overdispersion)으로 나타난다. 과분산은 조사되지 않은 다른 성질의 구간이 다양한 결과를 나타내기 때문이며 음이항 회귀분석은 사고자료가 과분산일 때 사용하면 큰 효과를 나타낼 것으로 기대된다. 음이항 회귀분석은 포아송 회귀분석과 비슷하게  $i$  번째 건널목에서의 사고발생건수는  $q$  개의 변수와 연관성을 가지며 식(4)와 같다.

$$Function(\mu_i) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_q X_{iq} \quad (4)$$

그러나 음이항 회귀분석은 분산을 나타내는 과분산

이 포함된 2차식의 형태이므로 식(5)와 같은 형태를 가지게 된다.

$$P(y_i) = \frac{(y_i + K - 1)!}{y_i! (K - 1)!} \cdot \frac{\mu^{y_i}}{(1 + \mu)^{y_i + K}} \quad (5)$$

여기서,  $K$ 는 과분산 계수이며, 그 분산은 식(6)을 통해 주어진다.

$$Var(y_i) = E[y_i] + KE[y_i]^2 \quad (6)$$

결과적으로 음이항 회귀분석은 조사되지 않은 다른 성분에 의하여 포아송 회귀분석의 분산을 변형시키게 된다. 과분산 계수인  $K$ 가 0에 가까우면 음이항 회귀분석은 포아송 회귀분석으로 바뀌게 되고,  $K$ 가 큰 값을 가지게 되면 자료에 있어서 평균값인  $\mu_i$  에 비해 더 큰 변화성을 가지게 된다.

포아송 회귀분석의 경우 계수  $\beta_j$  는 log-likelihood 인  $\log_e L(\beta)$  를 최대화함으로써 구할 수 있지만 음이항 회귀분석의 경우 추정된  $\widehat{K}$ 에 대하여 추정계수벡터와  $\widehat{Y}$ 가  $\log_e L(\beta, K)$  를 최대화함으로써 구할 수 있게 되며, 관련식은 식(7)과 같다.

$$\begin{aligned} \log_e L(\beta, K) &= \sum_i \left[ \left( \sum_{j=0}^{y_i} \log_e(1 + K_j) \right) \right. \\ &\quad - \log_e(1 + Ky_i) + y_i \log_e \mu_i - (y_i + \frac{1}{K}) \log(1 + K\mu_i) \\ &\quad \left. - \log(y_i!) \right] \end{aligned} \quad (7)$$

본 연구에서는 포아송 회귀분석과 음이항 회귀분석을 하여 과분산 계수인  $K$ 의 값을 확인한 결과 0에 가까운 값이 도출되었으므로 포아송 회귀분석을 최종모델로 선택하였다.

## 3. 모델검증

모델에 대한 적합도 검증(Goodness-of-fit) 방법은 내부검증(Internal validation) 방법과 외부검증(External validation) 방법으로 구분되어 진다.

내부검증방법은 변수선정에 있어서 기존문헌 등을 통해 변수들의 적절성을 평가하고, 모델식의 적용에 있

어서 적절한 통계적 방법을 이용하였는지에 대해 검증하며, 자료수집의 타당성 및 대표성을 살펴보는 과정등을 의미한다. 한편 외부검증방법은 내부검증방법과는 달리 보다 정량화된 방법으로 개발된 모델들을 검증하는 과정으로 Pearson 상관계수, MPB, MAD, 과분산을 통한 검증방법을 의미하며, 이에 대한 설명은 다음과 같다.

1) Pearson 상관계수

Pearson 상관계수  $r$ 은 두 변수  $Y_1$ 과  $Y_2$ 간의 연관성의 정도와 방향을 수량적으로 나타낸 값이다. 상관계수는 -1에서 1사이의 값을 가지며 -1에 가까우면 음의 상관관계, 1에 가까우면 양의 상관관계이며 절대치가 0에 가까우면 상관관계가 거의 없음을 의미한다.

$$r = \frac{\Sigma(Y_{1i} - \widehat{Y}_1)(Y_{2i} - \widehat{Y}_2)}{[\Sigma(Y_{1i} - \widehat{Y}_1)^2 \Sigma(Y_{2i} - \widehat{Y}_2)^2]^{1/2}} \quad (8)$$

$Y_i$  : 모델을 만들기 위하여 종속변수로 사용된 자료  
 $\widehat{Y}_i$  : 모델에 의한 결과값

2) MPB(Mean Prediction Bias)

이 통계수치는 모델을 만들기 위하여 종속변수로 사용된 자료에 대해서 모델에 의한 결과 값이 어느 정도 그리고 어떻게 치우쳐있는지를 판단할 수 있는 기준을 제공해준다. 이 방법에 의한 결과 값이 작을수록 모델의 예측 값은 정확한 것을 의미한다.

$$MPB = \frac{\sum_{i=1}^n (Y_i - \widehat{Y}_i)}{n} \quad (9)$$

$Y_i$  : 모델을 만들기 위하여 종속변수로 사용된 자료  
 $\widehat{Y}_i$  : 모델에 의한 결과값

3) MAD(Mean Absolute Deviation)

MAD는 모델의 예측 값이 평균적으로 얼마나 잘못 예측되었는지를 판단할 수 있는 척도를 제공해준다. 이 방법이 MPB와 다른 점은 각 수치의 음과 양의 차이로 인해 상쇄되지 않는다는 점이다. 결과 값이 0에 가까

울수록 모델이 실제의 관측된 자료에 부합되는 결과를 나타냄을 의미한다.

$$MAD = \frac{\sum_{i=1}^n |Y_i - \widehat{Y}_i|}{n} \quad (10)$$

$Y_i$  : 모델을 만들기 위하여 종속변수로 사용된 자료  
 $\widehat{Y}_i$  : 모델에 의한 결과값

4) 과분산(K)

포아송 회귀모델에서 과분산(Overdispersion)은 모델계수의 분산을 실제보다 적게 예측하게 되는 원인이 된다. 또한 이것은 일부변수들의 중요도를 과장하여 나타내는 결과를 초래하기도 한다. 자유도, n-p 등에 의해 구분되는 모수를 포함하고 있는 모델의 편차는 과분산인지의 여부를 결정할 수 있는 수치를 제공한다. 분산이 포아송 분포보다 더 크거나 작은 경우는 포아송 회귀모델이 적합하지 않음을 나타낸다.

$$DEV(x_0, x_1, x_2, \dots, x_{p-1}) = 2 \left( \sum_{i=1}^n Y_i \text{Log}_e \left( \frac{Y_i}{\mu_i} \right) - \sum_{i=1}^n (Y_i - \mu_i) \right) \quad (11)$$

VI. 철도건널목 사고빈도 예측모델

1. 철도건널목 사고빈도 예측모델 개발

본 연구에서는 철도건널목 사고빈도 예측모델을 개발하기 위하여 통계패키지 LIMDEP 8.0을 이용하여 분석하였다. 사용된 변수는 조사대상 철도 건널목 162개 지점에서 조사된 철도건널목 사고건수, 차량 및 철도 교통량, 주변 환경요소를 이용하여 모델개발에 활용하였다. 이러한 변수들을 이용하여 포아송 회귀모델과 기타 비선형 회귀모델의 적합성을 분석한 결과 과분산계수(K)의 값이 0에 가까운 값을 나타내어 철도건널목 사고빈도 예측모델은 포아송 회귀모델이 적합한 것으로 나타났다.

일반적인 사고빈도 예측모델의 경우 가장 예측성이 높은 주요모델(Best Fit Model)을 기준으로 안전성을 고려하고 있다. 그러나 사고빈도 예측모델의 주요 역할



은 사고빈도 예측을 통한 대상지역의 안전성 평가와 주요설명변수와 사고와의 관계를 규명함으로써 대상지역의 안전성 향상을 유도하는 것이다. 하나의 주 모델(Best Fit Model)만을 가지고 변수들을 설명할 경우 안전성 평가에 주요 설명변수가 생략될 수 있다. 즉, 주요 설명변수를 선택하기 위해선 선택된 모델들의 설명변수들을 종합적으로 비교·평가할 필요가 있다. 따라서 본 연구에서는 철도건널목 사고빈도 예측모델을 철도건널목 사고와 설명변수들간의 상관관계를 가장 잘 설명하는 주 모델과 후보모델로 나누어 개발하였다. 모델의 검증은 변수 선정에 있어서 변수들의 적절성을 평가하고 적절한 통계적 방법을 이용하였는지에 대해 검증하며 자료수집의 타당성 및 대표성을 살펴볼 수 있는 정성적인 내부검증방법과 Pearson 상관계수, MPB, MAD, 과분산 등의 정량적인 검증이 가능한 외부검증 방법을 종합하여 검증하였다.

포아송 회귀모델을 이용하여 개발한 철도건널목 사고빈도 예측모델과 모델을 설명하는 변수들의 통계학적 특성은 <표 4>와 같다. 주 모델의 경우 철도 건널목 사고를 예측하기 위하여 선정된 변수는 차량 교통량과 철도교통량, 상업지역, 제어거리, 경보시간차, 과속방지턱의 6개 변수가 철도 건널목 사고에 영향을 미치는 것으로 나타났다. 이 변수들의 P-Value는 신뢰수준 90%( $\alpha=0.10$ ) 기준에 모두 유의하므로 변수의 선정이 적절하다고 판단할 수 있다.

**2. 결과분석**

포아송 회귀분석을 이용하여 최종 선정된 포아송 회귀 모델의 회귀계수값(Coefficient)은 포아송 확률분포의 우도함수의 모수를 최대화함으로써 얻어진 결과 값이다. 포아송 회귀분석을 이용하여 최종 선정된 주 모

<표 4> 철도 건널목 사고빈도 예측모델

| 변수              |           | 주 모델   | 후보모델 I | 후보모델 II | 후보모델 III | 후보모델 IV |
|-----------------|-----------|--------|--------|---------|----------|---------|
| 상수              | coeff.    | -5.336 | -5.440 | -4.951  | -5.226   | -3.506  |
|                 | std. err. | 0.918  | 0.909  | 0.856   | 0.957    | 0.635   |
|                 | P-value   | 0.000  | 0.000  | 0.000   | 0.000    | 0.000   |
| 차량교통량<br>(ADT*) | coeff.    | 0.344  | 0.375  | 0.361   | 0.335    | 0.294   |
|                 | std. err. | 0.082  | 0.078  | 0.079   | 0.079    | 0.071   |
|                 | P-value   | 0.000  | 0.000  | 0.000   | 0.000    | 0.000   |
| 철도교통량           | coeff.    | 0.005  | 0.005  | 0.005   | 0.005    | 0.004   |
|                 | std. err. | 0.003  | 0.003  | 0.003   | 0.004    | 0.003   |
|                 | P-value   | 0.109  | 0.099  | 0.144   | 0.139    | 0.247   |
| 상업지역            | coeff.    | 0.967  | 1.005  |         | 0.977    |         |
|                 | std. err. | 0.369  | 0.365  |         | 0.364    |         |
|                 | P-value   | 0.009  | 0.006  |         | 0.007    |         |
| 제어거리 *          | coeff.    | 0.141  | 0.144  | 0.102   | 0.155    |         |
|                 | std. err. | 0.001  | 0.001  | 0.001   | 0.001    |         |
|                 | P-value   | 0.009  | 0.006  | 0.044   | 0.003    |         |
| 경보시간차           | coeff.    | 0.005  |        | 0.005   |          |         |
|                 | std. err. | 0.003  |        | 0.003   |          |         |
|                 | P-value   | 0.141  |        | 0.088   |          |         |
| 건널목유형(T형)       | coeff.    |        |        |         | -0.595   |         |
|                 | std. err. |        |        |         | 0.329    |         |
|                 | P-value   |        |        |         | 0.071    |         |
| 과속방지턱           | coeff.    | -0.853 | -0.966 | -0.872  |          |         |
|                 | std. err. | 0.425  | 0.416  | 0.424   |          |         |
|                 | P-value   | 0.045  | 0.020  | 0.040   |          |         |
| Pearson 상관분석    |           | 0.528  | 0.501  | 0.433   | 0.405    | 0.352   |
| MPB/year        |           | 0.000  | 0.000  | 0.000   | 0.047    | 0.000   |
| MAD/year        |           | 0.076  | 0.077  | 0.080   | 0.069    | 0.424   |

\* ADT\* = Ln(ADT), 제어거리\* = 제어거리/100

텔과 후보모델 변수들의 특성을 분석하여 보면 다음과 같다.

먼저 주모델 변수들의 특성을 분석하여 보면 상업지역과, 차량 교통량의 계수값은 각각 0.967, 0.34로 타 변수에 비해 가장 큰 계수값이 도출되었고, 철도건널목 사고에 가장 큰 영향을 미치는 것으로 나타났다. 이는 상업지역의 경우 차량의 유동이 많이 발생하는 지역이므로 철도 건널목을 이용하는 차량은 증가하여 사고가 발생할 수 있는 확률이 높아질 수 있기 때문이다. 기존의 철도건널목 사고빈도 예측모델인 Peabody Dimmick 모델, New Hampshire Index 모델, NCHRP 50 Method, DOT 모델의 연구 모두 차량의 교통량이 철도건널목 사고와 밀접한 관련이 있는 것으로 나타나 있다. 이러한 결과는 향후 건널목 입체화를 추진할 경우 차량의 교통량이 많이 유발되는 상업지역을 최우선적으로 입체화를 추진하는 것이 사고예방에 효과적이라고 할 수 있다.

제어거리에 대한 계수값은 0.141로서 제어거리가 길수록 철도건널목 사고가 많이 발생한다는 결과가 나타났다. 제어거리가 길어지면 철도건널목이 차단기로 차단된 후 열차가 통과하기까지의 시간이 길어진다. 이것은 건널목을 통과하기 위한 차량의 대기시간이 길어진다는 것을 의미한다. 차량의 대기시간이 길어질 경우 급한 성격을 가진 운전자나 바쁘게 움직여야 하는 운전자 등은 차단기의 정상작동 여부를 의심할 수 있으며, 차단된 차단기를 피하여 건널목을 통과하다 사고가 발생할 수 있다.

경보시간차의 계수값은 0.005로서 경보가 길게 작동하는 경우가 짧게 작동하는 경우보다 사고가 많이 발생한다고 제시하고 있다. 이는 기존의 연구(Berg 등, 1982 ; Hopkins, 1981 ; Mather, 1991)에서도 경보시간차가 사고와 밀접한 관련이 있다고 되어 있다. 이는 경보시간이 길어질 경우 급한 성격을 가진 운전자나 바쁘게 움직여야 하는 운전자 등은 오랜 경보 후에도 열차가 오지 않아 열차 진입경보를 무시하고 건널목을 통과하다가 열차와 충돌하는 경우가 발생할 수 있다. 기존의 건널목 사고 발생유형 중 법규위반으로 인한 사고가 가장 많다는 분석과 일치하는 결과를 제시하고 있다.

과속방지턱의 계수값은 -0.853으로서 과속방지턱이 설치되어 있는 경우가 설치되어 있지 않는 경우보다 사고가 적게 발생한다고 제시하고 있다. 이는 과속방지턱은 차량의 속도를 감속시키는 효과가 있으므로, 차단기

〈표 5〉 사고빈도 예측 모델식

| 구분        | 사고빈도 예측 모델식  |
|-----------|--|
| 주 모델      | $Y_{min} = \exp[-5.336 + 0.344 X_1 + 0.005 X_2 + 0.967 X_3 + 0.141 X_4 + 0.005 X_5 - 0.853 X_7]$ |
| 후보 모델 I   | $Y_1 = \exp[-5.440 + 0.375 X_1 + 0.005 X_2 + 1.005 X_3 + 0.144 X_4 - 0.966 X_7]$                 |
| 후보 모델 II  | $Y_2 = \exp[-4.951 + 0.361 X_1 + 0.005 X_2 + 0.102 X_4 + 0.005 X_5 - 0.872 X_7]$                 |
| 후보 모델 III | $Y_3 = \exp[-5.226 + 0.335 X_1 + 0.005 X_2 + 0.977 X_3 + 0.155 X_4 - 0.595 X_6]$                 |
| 후보 모델 IV  | $Y_4 = \exp[-3.506 + 0.294 X_1 + 0.004 X_2]$   |

참조)  $X_1$ : Ln(평균 일교통량(ADT))  $X_2$ : 철도교통량  $X_3$ : 상업지역  $X_4$ : 제어거리  $X_5$ : 경보시간차  $X_6$ : 건널목유형(T형)  $X_7$ : 과속방지턱

가 동작중일 때 완전히 차단되기 전 건널목을 통과하기 위하여 차량이 가속하는 것을 예방하는 효과에서 기인된다고 판단된다.

이외에 후보모델의 변수를 분석하여 보면 후보모델 I, II, IV의 경우에는 주모델의 변수들과 중복된다. 그러나 후보모델 III에서는 건널목 유형이 변수로 추가되어 철도 건널목 사고와 관련이 높은 것으로 나타났으며, 이중 철도건널목 유형이 T형일 때 사고가 적게 발생하는 것으로 분석되었다. 이는 건널목의 신설 또는 기존 건널목의 개량 시 건널목 유형에 대한 고려가 건널목사고를 줄일 수 있는 중요한 요소 중의 하나임을 나타내고 있다. 본 연구에서 개발된 사고빈도 예측모델 식은 〈표 5〉와 같다.

## Ⅷ. 결과 및 향후 연구과제

철도건널목 사고는 도로교차로 사고에 비하여 사고 빈도는 작으나 사고발생시 대형사고로 이어져 막대한 사회적·경제적 손실을 초래한다. 그리고 장래 철도의 복선화, 고속화 및 도로 교통량의 증가로 인해 철도건널목의 사고위험성은 증가될 것이며, 이를 예방하기 위하여 건널목의 입체화가 매년 추진되고 있으나 지형적 문제 및 경제적 여건 등으로 한계성이 있다. 따라서 철도건널목 사고를 줄이기 위하여 인간이 제어가능한 철도건널목의 기하구조와 주변환경요소가 사고와 어떤 관련이 있는가에 대한 심도있는 분석은 매우 중요하며 이러한 분석을 통해 사고를 미연에 방지해야 할 것이다.

본 연구는 162개의 철도건널목의 기하구조와 주변환경요소를 조사하여 그 자료를 근거로 포아송 회귀분석과 음이항 회귀분석을 이용하여 철도건널목 사고예측

에 가장 적합한 회귀분석 모델을 선정하였다. 그 결과 포아송 회귀분석이 철도건널목 사고모델에 가장 적합한 경로 분석되었다. 모델에 적용된 변수로는 평균 일 교통량, 철도교통량, 상업지역, 제어거리, 경보시간차, 과속방지턱이다. 본 연구는 지형적 문제 및 경제적 여건 등으로 건널목의 입체화가 추진될 수 없는 건널목에서 발생하는 사고를 사전에 예방하기 위하여 사고를 예측하고 사고관련요소들을 찾아 사전에 정비 및 보완하는데 그 의의가 있다.

본 연구에서 제시한 철도건널목 사고예측모델을 보완하기 위하여 향후 연구해야할 부분을 살펴보면 다음과 같다.

첫째, 철도건널목 사고의 경우 사고빈도가 적으므로 조사대상 건널목 수가 많을수록 더욱 신뢰할 수 있는 모델을 만들 수 있다. 따라서 본 연구의 표본수보다 더 많은 표본수를 이용한 연구가 지속적으로 이루어져야 할 것이다.

둘째, 철도건널목의 경우 도로교차로와는 달리 차단기라는 물리적인 장치를 이용하여 열차가 건널목 진입시 진행을 통제하고 있다. 그럼에도 사고가 발생한다는 것은 철도건널목의 기하구조와 주변 환경요건 뿐만 아니라 운전자 개개인의 심리와 경험 등이 사고에 밀접한 관련이 있다고 볼 수 있다. 하지만 본 연구에서는 사고자에 대한 개별 연구를 목적으로 한 것이 아니라 다수의 철도건널목을 연구대상지로 선정하여 분석하였기 때문에 철도건널목에서 발생하는 개별사고에 대한 분석은 본 연구의 범위를 벗어난다. 따라서 철도건널목을 이용하는 운전자 개개인의 인적요인을 대상으로 하는 추후 연구가 이루어져야 할 것이다.

셋째, 철도건널목 사고예측을 위해 포아송 회귀모델과 기타 비선형 회귀모델과 같은 확률회귀모델 이외에 다양한 예측이론과의 접목이 필요하다. 한 예로 철도건널목 사고의 사전정보와 사후정보를 동시에 고려한 베이즈이론분석론의 도입은 확률회귀모델의 단점을 보완하는 추후 연구로 시도해 볼만한 좋은 연구과제가 될 수 있을 것이다.

넷째, 원천적으로 건널목 사고를 예방하기 위해서는 건널목의 입체화가 필요하다. 고속철도가 운행되는 경부선과 호남선의 경우 기존 모든 평면 건널목의 입체화를 추진하고 있으며 이외에도 교통량이 많은 도시부의 평면 건널목도 입체화를 추진하고 있다. 하지만 지형적, 경제적 타당성에서 입체화가 불가능한 경우도 있

다. 따라서 본 연구는 입체화가 불가능한 평면 건널목을 대상으로 사고빈도 예측모델을 개발하였지만 향후 입체화 시공에 대한 안전성 연구가 반드시 필요하다.

마지막으로, 정확한 사고분석을 위하여 철도건널목의 기하구조와 주변환경요건에 대한 자료에 대하여 데이터베이스(DB) 구축이 필요하며 합리적이고 체계적인 자료의 정리가 이루어져야 할 것이다.

### 참고문헌

1. 강승규(1995), "철도건널목의 사고예측모형 개발에 관한 연구", 대한교통학회지, 제13권 제2호, 대한교통학회, pp.43~58.
2. 철도청(2003), "철도건널목 위험요인 분석 및 개량방안 연구".
3. 철도청(1998~2002), "철도운전사고 월보".
4. 철도청(1998~2002), "철도사고이력카드".
5. 철도청, "철도건널목카드".
6. Ardeshir Faghri, Michael J. Demetsky(1988), "Reliability and Risk Assessment in the Prediction of Hazards at Rail-Highway Grade Crossings", Transportation Research Record 1160.
7. Berg, W, knoblauch, K, and Hucke, W(1982). Causal factors in Railroad-Highway Grade Crossing Accident, Transportation Research Record 847.
8. F.F. Saccomanno, Congning Ren, Liping Fu(2003), "Collision Prediction models for Highway-Rail Grade Crossings in Canada", TRB 2003 Annual Meeting CD-ROM.
9. Hopkins, J.B(1981). Technological Innovations in Grade Crossing Protection Systems. Report TSC-FRA-71-3, FRA, U.S. Department of Transportation, Cambridge, Mass.
10. Jutaeck Oh, Craig Lyon, Simon Washington, Bhagwant Persaud, Joe Bared(2004), "Validation of the FHWA Crash Models for Rural Intersections: Lessons Learned", Transportation Research Record 1840.
11. Jutaeck Oh, Simon Washington, Keechoo Choi, "Development of accident prediction

- models for rural highway intersections”, TRB 2004 Annual Meeting CD-ROM.
12. Mather, R. A(1991). Seven Years of Illumination at Railroad-Highway Crossings, Transportation Research Record 1316, pp.54~58.
  13. Ross D. Austin, Jodi L. Carson(2002), “An alternative accident prediction model for highway-rail interfaces”, Accident Analysis & Prevention.
  14. Simon P. Washington, Matthew G. Karlaftis, Fred L. Mannering(2003). “Statistical and Econometric Methods for Transportation Data Analysis”, Chapman & Hall/CRC.
  15. William H. Greene, “Limdep version 8.0 Econometric Modeling Guide”, Econometric Software, Inc.

✉ 주 작 성 자 : 오주택

✉ 논문투고일 : 2004. 8. 17

논문심사일 : 2004. 11. 30 (1차)

2005. 2. 2 (2차)

심사판정일 : 2005. 2. 2

✉ 반론접수기한 : 2005. 6. 30