

Working Paper

2011-WP-16

서울시 도시고속도로 사고취약구간 선정에 관한 연구

강승모

2011

서울시 도시고속도로 사고취약구간 선정에 관한 연구

2011

■ 연구진 ■

연구책임 강승모 • 고려대학교 건축사회환경공학부 교수

이 보고서의 내용은 연구진의 견해로서
서울특별시의 정책과는 다를 수도 있습니다.

목 차

I. 연구의 개요	1
제1절 연구의 배경 및 목적	1
제2절 주요 연구내용	2
II. 국내외 사고취약구간 선정 방법론 검토	3
제1절 국내 사고 많은 곳 선정 기준	3
제2절 Safety Performance Function	5
제3절 Empirical Bayes 법	9
제4절 Continuous Risk Profile 법	14
제5절 Geographically Weighted Regression 법	18
제6절 Kernel Density Estimation 법	22
III. 고속도로 사고취약구간 선정 방법론 적용 사례	31
제1절 사례 연구 개요	31
제2절 대상 구간 데이터 구축	31
제3절 GWR 법과 KDE 법을 이용한 사고 취약구간 선정	34
1. GWR 법의 적용	34
2. KDE 법의 적용	36
3. GWR과 KDE의 비교	38
IV. 결론	41
참고문헌	43

표 목 차

〈표 1-1〉 2007년 OECD 회원국 도로교통사고 비교	2
〈표 2-1〉 국내 교통사고 잦은 곳 선정기준	4
〈표 2-2〉 Kwak et al.(2010)에 사용된 변수	7
〈표 2-3〉 국내 고속도로 대상 SPF 연구	8
〈표 2-4〉 전역적 통계량과 국지적 통계량의 차이	18
〈표 2-5〉 커널밀도를 적용한 기존 연구 사례	28
〈표 3-1〉 분석에 사용된 변수	33
〈표 3-2〉 전역회귀분석과 지리적 가중회귀분석의 결과 비교	35
〈표 3-3〉 GWR의 그룹별 결과 비교	35
〈표 3-4〉 사상자별 가중빈도 상위 10% 구간(KDE_W)	38
〈표 3-5〉 사고취약구간 선정 결과 비교	39
〈표 3-6〉 GWR과 KDE의 비교	40

그림목차

〈그림 2-1〉 국내 교통사고 잦은 지점 선정 과정	4
〈그림 2-2〉 Empirical Bayes 법 계산식	10
〈그림 2-3〉 Sliding Window 법	12
〈그림 2-4〉 Peak Search 법	13
〈그림 2-5〉 CRP 다이어그램 예시	15
〈그림 2-6〉 CRP 다이어그램을 이용한 사고취약구간 선정	15
〈그림 2-7〉 서울방향 264km~273km 구간 CRP	17
〈그림 2-8〉 Simpson's paradox 예시	19
〈그림 2-9〉 사고예측모형을 이용한 사고취약구간 판별 개념	21
〈그림 2-10〉 커널함수의 예	23
〈그림 2-11〉 커널밀도의 개념	23
〈그림 3-1〉 고속도로 데이터 지오코딩 자료	32
〈그림 3-2〉 전체 노선구간 사고밀도지도	36
〈그림 3-3〉 국지적 범위의 사고밀도지도	37

I

연구의 개요

제1절 연구의 배경 및 목적

- 세계 각국은 사고가 자주 발생하는 지점 혹은 구간을 선별하여 적절한 안전개선 조치를 취함으로써 사고를 줄여나가고자 하는 노력을 지속적으로 기울이고 있으며, 이에 따라 이러한 지점이나 구간을 보다 합리적으로 선정하기 위한 연구의 중요성 역시 계속적으로 강조되고 있음.
- 사고취약구간 선정문제는 비용 효과적인(cost-effective) 측면에서 제한된 도로 안전 개선사업 예산을 최대한 효과적으로 집행할 수 있도록, 사업 시행의 우선순위를 부여하는 중요한 의사 결정 과정임.
- 우리나라의 교통안전 수준은 OECD 회원국 중 하위권을 벗어나지 못하고 있는 실정임.
 - 2009년 기준 도로교통사고 사망률(자동차 1만 대당 2.8명)이 OECD 평균(32개국 평균 1.3명)보다 2배 이상 높으며, 특히 스위스, 일본과 영국 (이상 0.70명)에 비해서는 4배 더 높은 수치를 기록하고 있음.
 - 뿐만 아니라 교통사고로 인해 발생하는 사회적 손실 비용이 2010년 기준 연간 12조 9,599억원 규모로 같은 해 국내 총생산 1,173조원의 약 1.1%에 해당하는 것으로 조사된 바 있음.
 - 서울시의 교통사고 발생건수도 2010년 기준 41,662건으로 2006년 38,234건에 비해 9% 증가하였고 사망자 또한 2010년 기준 연간 429명에 달하고 있음.
- 특히 고속도로의 경우 우리나라 전체 발생건수 기준 사고발생비율은 3.28%이나 사망자 비율은 전체의 7.06%로 사고발생비율의 약 2.15배에 이르고 있으며, 이는 차

량이 고속으로 주행하는 특성상 고속도로 구간에서 발생하는 사고가 훨씬 심각한 결과를 초래한다는 것을 의미함.

〈표 1-1〉 2009년 OECD 회원국 도로교통사고 비교

구 분	한국	일본	영국	미국	스위스
교통사고 사망자수(명)	5,838	5,772	2,337	33,808	349
자동차 1만대당 사망자수(명)	2.8	0.70	0.70	-	0.70
인구 10만명당 사망자수(명)	12.0	4.5	3.8	11.0	4.5
주행거리 10억km당 사망자수	20.1	8.1	5.2	8.0	5.6

- 서울시는 내부순환로, 강변북로, 올림픽대로, 북부간선도로, 동부간선도로, 서부간선도로, 경부고속도로 등 총연장 175.6km에 달하는 다양한 구간의 도시고속도로를 관리하고 있으며, 노들길, 언주로, 남부순환로 등 자동차 전용도로를 포함할 경우 그 연장이 202.8km에 달하고 있음.
- 이러한 도시고속도로에서의 사고율 및 치사율을 줄이기 위해서는 장기적인 계획을 통한 집중적인 예산투입이 필수적이며, 이러한 예산 투입 방법과 시기를 결정하는데 있어 합리적이고 과학적인 사고취약구간 선정 과정의 정립이 절실한 실정임.
- 따라서, 이 연구는 지금까지 소개된 사고취약구간 선정 방법론 및 적용사례를 분석함으로써, 서울시 도시고속도로 실정에 맞는 합리적인 방안을 도출하는 것을 목적으로 함.

제2절 주요 연구내용

- 이 연구는 우선 현재 우리나라에서 적용되고 있는 사고 많은 곳 선정기준을 분석하고, 선진국에서 적용되고 있거나 최근에 연구된 사고 취약구간 선정 방법론을 소개함.
- 이렇게 제시된 사고 취약구간 선정 방법론들이 구체적으로 어떻게 적용될 수 있는지, 국내 고속도로에 적용한 사례를 분석함.
- 마지막으로 이러한 방법론 및 적용사례를 통해 서울시 도시고속도로 실정에 맞는 사고취약구간 선정 방안을 제시함.

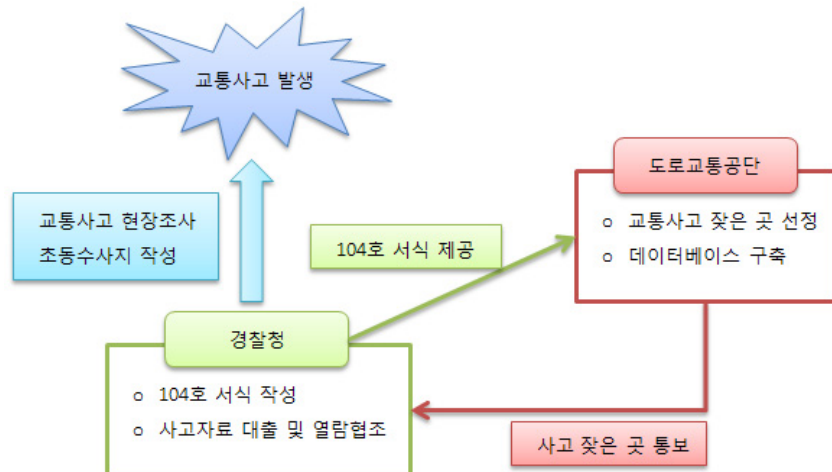
II

국내외 사고취약구간 선정 방법론 검토

제1절 국내 사고 많은 곳 선정 기준

- 우리나라는 도로교통공단에서 경찰청의 협조를 얻어 매년 교통사고 잦은 지점을 선정하여 안전개선사업을 시행하고 있는데, [교통사고 자료조사 및 분석 → 사업대상지점 선정 → 현장조사 및 사고요인 분석 → 개선방안 수립 → 기본개선계획 보고서 작성 → 개선공사 시행 → 개선효과분석 평가]의 단계로 추진됨.
- 첫 단계인 자료조사 및 분석을 위해 교통사고 전수조사를 실시하는데, 이는 조사시행 전년도에 전국에서 발생한 모든 교통사고를 대상으로 함.
 - 이 단계에서 교통사고 잦은 곳으로 선정되면, ‘도로시설의 환경적 측면’에서 문제가 있어, ‘개선사업 수행 시 사고감소 효과가 클 것으로 기대되는 지점’을 대상으로, 우선순위 결정방법에 따라 사업대상지점을 선정하여 개선공사를 시행함.
 - 여기서의 우선순위 결정은 각 시·도 도로관리부서와 협의를 통한 개선설계 요구지점 및 교통안전 시범도로 지정구간을 제1순위로, 사망사고 많은 지점 및 다수의 민원제기 지역을 제2순위로, 교통사고(인적+물적) 다발지점 및 사고위험이 예상되는 개선 필요지점을 제3순위로 하고 있음(김숙희 외, 2005).
- 이러한 과정들이 대부분 조사원들의 수작업을 통해 진행되므로 많은 시간과 노력, 그리고 비용이 소요됨.
- 현재 교통사고 잦은 지점의 선정에 이용되는 방법은 다음 <표 2-1>의 기준들에 근거한 사고건수법으로, 지난 1년간 발생한 교통사고건수를 바탕으로 도로 유형별로 기준 건수 이상의 사고가 발생한 구간을 개선대상지점으로 분류하고 있음.

○이 방법은 사고 및 도로 구간에 대한 상세한 분석을 통해 추진된다기보다 단순히 교통사고 발생 건수만을 기준으로 후보지를 선정한 후 실무자들의 경험적 이해를 통해 안전도를 평가하므로 우선순위 선정 및 예산집행에서 비효율성을 초래할 수 있다는 단점이 있음.



〈그림 2-1〉 국내 교통사고 잦은 지점 선정 과정(박신형, 2010)

〈표 2-1〉 국내 교통사고 잦은 곳 선정기준

구분		4차 사업 (‘08년 이후)	3차, 4차 사업 (‘08년 이전)	1차, 2차 사업
지역	특별시	5건 이상	7건 이상	10건 이상
	광역시	5건 이상	7건 이상	7건 이상
	일반시	3건 이상	5건 이상	5건 이상
	기타	3건 이상	3건 이상	3건 이상
도로형태	교차로 및 횡단보도		차량 정지선에서 후방 30m 이내	
	기타	시가지	반경 100m 이내	
	단일로	기타, 고속도로	반경 200m 이내	
	대상사고		인피사고	인피+물피사고

○<표 2-1>에서 1차에서 4차까지 진행되어 온 교통사고 잦은 곳 개선사업에 적용된 기준을 보면, 전반적으로 사고건수 기준이 낮아지고 있고, 기타 단일로 중 시가지 도로를 제외한 고속도로 등의 구간조사 범위도 넓어진 것을 볼 수 있음.

–또한 한동안 물피사고까지 선정 기준에 포함하여 고려하였으나 2008년 이후부터는 다시 제외하였음.

○이러한 기준의 변화는 다음과 같이 볼 수 있음.

—첫째로, 그간 꾸준히 지속되어 온 교통사고를 줄이기 위한 다각도의 노력이 반영된 결과로 전반적인 사고 발생건수가 줄어들었기 때문에, 검토에 필요한 일정 수준 이상의 대상구간을 선정하기 위해 사고건수 및 구간의 조사범위 기준을 완화한 것이라고 볼 수 있음.

—하지만 통계학적인 측면에서 볼 때, 전체적인 사고 자료의 수가 감소함에 따라 사고 잦은 구간의 선정을 위한 사고건수 기준이 낮아지면, 지난 한 해 동안 일시적으로 사고가 증가한 구간이 우선적으로 선정되는 것과 같은, 사고의 우연성이 구간 선정에 큰 영향을 미치는 문제가 발생하게 됨.

—둘째로, 단순물피사고나 경미한 부상사고는 경찰에 신고하지 않고, 보험이나 당사자 간 합의에 따라 처리하는 경우가 많아 정확한 집계가 어렵기 때문에 대상사고에서 물피사고를 다시 제외하였음.

—즉, 심각도가 낮은 사고들의 잘 보고되지 않는(under-reported) 경향으로 인해 구간별로 오차가 커질 수 있고, 이는 조사의 신뢰성에 영향을 미칠 수 있게 됨.

—쉬운 예로, 교통량이나 기하구조 등 제반조건이 비슷한 두 개의 지방부 고속도로 구간 A와 B가 있을 때, A에서는 2건의 인피사고와 함께 보고되지 않은 4건의 물피사고가 있었고, B에서는 인피사고가 없었으나 실제 발생한 5건의 물피사고가 모두 보고되었다면, 조사과정에서 1차적으로 A 대신 B가 사고 잦은 곳으로 선정되는 문제가 생길 수 있음.

○이러한 해석을 고려하였을 때, 사고건수를 기준으로 사고 잦은 지점을 판별하는 기존의 방식보다 더 과학적이고 합리적인 방법론을 도입할 필요가 있음.

제2절 Safety Performance Function(SPF)

○미국 연방도로국(Federal Highway Administration)이 개발하여 많은 주(state)에서 공식 기법으로 활용하고 있는 Safety Performance Function(SPF)은 특정 도로구간에

- 대한 교통량, 도로 기하구조 특징 등을 독립변수로 하여 해당 도로 구간의 단위시간당 예측되는 사고 빈도수를 추정하는 모형으로 정의될 수 있음(Zhong 외, 2009).
- 이는 도로의 계획과 설계 과정에서 안전도를 고려하는 데 있어 중요한 역할을 할 수 있음.
 - 도로를 신설하는 경우에는 보다 안전한 도로계획 및 선형설계에 영향을 미칠 수 있으며, 기존 도로를 개선하는 경우에는 예산의 효율적 집행을 위해 안전도 개선 사업의 우선순위를 산정하는 척도로 이용될 수 있음.
 - 특정 도로에서 예측되는 사고빈도수를 추정하는 SPF에 대한 연구는 국내외에 걸쳐 상당히 활발히 진행되고 있음.
 - 강정규, 이성관(2002)은 고속도로를 직선부와 곡선부로 구분하여 도로의 선형요소(길이, 곡선반경, 종단경사, 편경사) 등을 독립변수로 하는 선형회귀식을 개발하였음.
 - 강민욱 외(2002)는 고속도로의 평면선형상 곡선부에 대한 사고의 분포가 음이항 분포를 따른다는 것을 규명하였으며, 교통량에 관계된 도로의 노출도(exposure) 변수와 곡선반경, 단위거리당 편경사 등의 기하구조 관련 변수를 이용하여 음이항 회귀모형을 개발하였음.
 - 강승림, 박창호(2003)는 곡선반경과 곡선길이의 조합별 위험도, 곡선반경, 인접한 직선길이, 곡률차의 조합별 위험도, 곡선반경과 종단경사의 조합별 위험도 등 여러 가지 기하요소를 조합하여 다중선형회귀모형을 구축하였음.
 - 박효신(2007)은 사고의 분포를 음이항분포로 가정하여 고속도로 트럼펫 인터체인지 연결로 전체, 연결로 형식별(직결, 준직결, 루프연결로)로 각각에 대한 음이항 회귀모형을 개발하였으며, 이때 사용된 변수로는 exposure 변수, 중차량 비율, 곡선장, 곡률, 유출입 여부 등이 있음.
 - 한상진 외(2008)는 사고예측을 위해 Hauer (2004)가 제시한 방법론을 2004년 한국의 전국 고속도로 자료를 이용하여 적용하였으며, 이때 이용된 변수는 AADT와 종단구배임.
 - Zhong 외(2009)는 exposure 변수와 선형관련 변수 외에 환경관련 변수들로 통행방향 변수, 지방부와 도시부 구분 변수, 인터체인지 여부 변수 및 트럭 혼입률,

트럭과 승용차 간 속도 차이 등 비교적 다양한 변수들을 고려하였으며, 여타 연구와 달리 시간당 교통량을 분석단위로 설정하였음.

- 한국 고속도로에서의 SPF를 개발하는 연구를 수행한 Kwak et al.(2010)은 교통량 반영 변수인 노출도와 기하구조 및 도로환경변수를 이용하였으며, 분석 결과 유의미한 변수로 선정된 노출도, 평면선형, 진입부, 차로수, 톨게이트, 트럭혼입률 등을 반영하여 다음의 SPF를 도출하였음.

$$Y_i = EXPO \cdot \exp(-4.285395 + 0.0000293 \cdot Hor_{radius} + 0.3109837 \cdot IC_{in} - 0.2348637 \cdot Lane + 0.7697941 \cdot TG + 2.191668 \cdot Truck\%)$$

<표 2-2> Kwak et al.(2010)에 사용된 변수

구분	변수	설명
Exposure	EXPO	노출도(ADT)
Alignment	Hor_radius	곡선반경
Environment	IC_in	진입부
	Lane	차로수
	TG	Toll Gate
Traffic	Truck%	트럭 혼입률

- 국내에서 고속도로를 대상으로 한 SPF에 관한 연구는 <표 2-3>과 같이 정리될 수 있음.
- 이들 연구에서는 교통사고의 발생분포가 대체로 정규분포나 포아송 혹은 음이항분포를 따를 것으로 가정하고 모형식을 추정하고 있으며, 모수 추정을 위해서는 최소자승법(Least Square Estimation; LSE) 혹은 최우추정법(Maximum Likelihood Estimation; MLE)을 사용하고 있음.
- 또한 교통사고를 설명하기 위한 변수로는 주로 교통량에 관계된 exposure 변수와 곡선반경, 편경사 등의 기하요소가 이용되었음.
- SPF는 도로의 기하구조와 교통조건이 사고에 영향을 미치는 정도를 분석하여 도로 구간별 예상사고변수를 추정함으로써 도로위험도를 평가하고 도로안전개선사업을 위한 구간을 선정하는 기법으로 활용되고 있음.

〈표 2-3〉 국내 고속도로 대상 SPF 연구(박신행, 2010)

연구자	공간적 영향 반영	적용 모형(식)
강정규, 이성관(2002)	평면 선형에 따라 각각 본선구간과 유출입부로 구간 분할	다중선형회귀모형
강민욱 외(2002)	곡선부에 대한 구간 분할법 적용	음이향 회귀모형
강승림, 박창호(2003)	기하요소의 조합	다중선형회귀모형
박효신 외(2007)	트럼펫 인터체인지 연결로 전체, 연결로 형식별(직결, 준직결, 루프연결로)로 구분	음이향 회귀모형
한상진 외(2008)	교통량(AADT), 종단구배, 곡선반경을 변수로 반영	음이향 회귀모형
Kwak et al.(2010)	기하구조와 교통량, 도로시설물 등을 변수로 반영	음이향 회귀모형

○ 하지만 도로 및 교통환경의 특성을 변수로 반영하여 유사한 조건을 가진 구간들은 비슷한 결과를 얻을 수 있으나, 사고의 발생을 독립적으로 간주하고, 구간 내 사고 확률을 동일하게 부여하며, 인접한 구간 또는 사고와의 상호작용을 반영하지 않는다고 가정하는 문제가 있음.

○ Chung 외(2009)는 러한 가정의 문제점을 다음과 같이 세 가지로 지적하였음.

— 첫째, SPF는 고속도로에서의 사고가 서로 상관되지 않고(uncorrelated), 공간적으로 독립적이라고 가정하는데, Chung 외(2009)는 고속도로의 상습정체구간(recurrent bottleneck)과 그 주변 구간의 사고지점들을 도시한 결과, 사고들이 시간적, 공간적으로 관련되어 있는 경우가 관찰되었으므로, 개별 사고가 독립적으로 발생한다는 가정은 현실에 위배된다고 주장하였음.

— 이러한 공간적 의존성을 줄이고 첫 번째 가정을 만족시키는 방법은 도로 구간의 길이를 늘이는 것을 생각할 수 있지만, 도로 구간의 길이를 늘리면 공간적으로 이질적인 기하구조가 변하는 구역까지 포함시킬 수 있음. 또한 국지적으로 높은 사고 빈도 또는 사고율을 나타내는 지점이 전체 구간길이에 대해 평균화(average out)되어 버림으로써 사고 잦은 곳으로 판별되지 못하는 문제가 발생하기도 함.

- 둘째, SPF는 동일 구간 내 각 지점에서의 사고 확률이 모두 동일하다고 가정하는데, 현실적으로 같은 구간 내에서도 지점별로 사고가 발생할 확률이 달라질 수 있음.
- 마지막으로, 잦은 사고를 유발하는 요인은 그 구간 내에 존재한다는 가정인데, 실제로 사고는 연속된 도로상에서 인접한 구간의 영향을 받아 발생하는 경우가 많을 뿐만 아니라, 반대로 해당 구간에서 잦은 사고를 유발하는 요인이 인접구간까지 영향을 미칠 수 있기 때문에 이는 부적절한 가정이라 볼 수 있음.

제3절 Empirical Bayes(EB) 법

- EB 법은 고속도로상의 특정 지점에 대하여 장기간의 기대 사고건수를 예측하는 방법을 말함.
- EB 법은 미연방도로국(FHWA)에서 개발한 SafetyAnalyst 소프트웨어 등에서 네트워크 스크리닝 방법의 하나로 쓰이고 있으며, 최근 제정된 교통안전편람(Highway Safety Manual; HSM)에서도 추천되고 있음.
- Bayesian 이론은 사전분포(Prior distribution)의 불확실성을 사후분포(Posterior distribution)를 통하여 보정할 수 있다는 논리를 기반으로 하고 있음. 이는 다시 말하면 위험도를 가진 어떤 특정 지점은 과거의 사고기록뿐만 아니라 그와 유사한 지점들의 집단인 Reference Sites에도 속해 있으며 이러한 Reference Sites는 개체의 위험도에 관한 정보를 제공한다는 뜻임(정성봉 외, 2009).
- EB 법은 어떤 특정 지역의 교통사고 현황을 그 대상 지점과 Reference Sites의 교통사고 기록과 결합하여 교통사고의 불확실성으로 인한 문제를 극복할 수 있고 확률적 해석이 가능하다는 측면 때문에 전통적인 사고예측방법과 많은 점이 다름.
- Hauer(1992)는 EB 법을 교통사고 취약구간 선정에 다음과 같은 방법으로 적용하였음.
 - 첫 번째 단계에서는 특정 종류의 도로구간들(참조 모집단)에 다변량 모형을 적합(fitting)시키고, 이 정보를 이용하여 참조 모집단에서 기대사고건수의 분포를 얻음.

- 이 분포에서 기대사고건수가 어떤 값을 초과하면($m > m^*$) 이 도로구간은 정상적이지 않다고 판단하도록 하는 임계값인 m^* 를 선택함.
- 다음으로 도로구간의 사고 이력을 이용하여 해당 도로구간의 기대사고건수에 대한 확률밀도함수를 구함.
- 마지막으로 도로구간의 기대사고건수가 정밀한 조사를 통해 참조 모집단에서 선택된 m^* 를 초과할 확률을 결정하고, 이 확률이 크다면 해당 도로구간은 더 상세한 조사가 요구된다고 볼 수 있음.
- EB 법에서 예측되는 기대사고건수는 SPF와 그 지점에서 관측되는 사고건수의 가중평균으로 <그림 2-2>와 같이 계산됨.
- 이때, 가중치(w)는 Reference Sites의 교통사고건수와 교통사고건수의 분산값을 이용하여 결정할 수 있으나, 현실적으로 Reference sites의 사고건수와 분산값을 예측하기 어렵기 때문에 Hauer(1992)는 교통사고예측모델을 이용하여 이를 해결하는 방안을 제시하였음.

Estimate of Expected Collisions for a site (E) = $w \times (\text{SPF prediction}) + (1 - w) \times (\text{Observed collision frequency})$

where:
 $0 \leq w \leq 1$

Estimate of Expected Excess Collisions for a site = (Estimate of Expected Collisions for a site) - (SPF Prediction)

<그림 2-2> Empirical Bayes 법 계산식(Srinivasan 외, 2011)

$$w = \frac{1}{k(P + 1/k)},$$

여기서, k= overdispersion parameter, P=SPF Prediction

- EB 법은 단순사고건수에도 적용될 수 있으며, 사상자 혹은 물피금액에 가중치를 둔 가중 사고건수에도 적용될 수 있음.
- EB 법은 Regression-to-Mean(평균으로의 회귀)로 인한 오차의 제거에 매우 효과적이라고 평가되고 있음.
- Regression-to-Mean은 통계학적으로 어떤 지점에서 교통사고가 일시적으로 크게 증가하거나 감소하더라도 궁극적으로는 그 지점에 대한 장기적인 평균값에 수렴

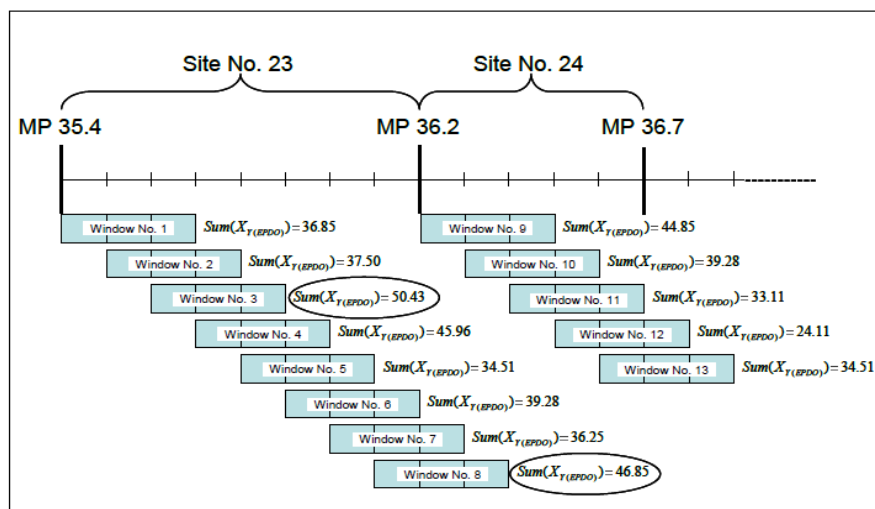
한다는 것을 의미함.

- 따라서 이를 고려하지 않고 짧은 기간의 사고 자료를 바탕으로 사고 잦은 곳을 선정하게 되면, 통계적으로 사고가 유의하게 발생하지 않는 곳임에도 일시적으로 사고가 많이 발생하여 선정되는 경우가 있고, 반대로 통계적으로 사고가 발생할 경향이 높은 곳임에도 일시적으로 사고가 적게 발생하여 선정되지 못하는 경우가 생길 수 있음.

○EB 법을 국내사례에 적용한 연구는 다음과 같음.

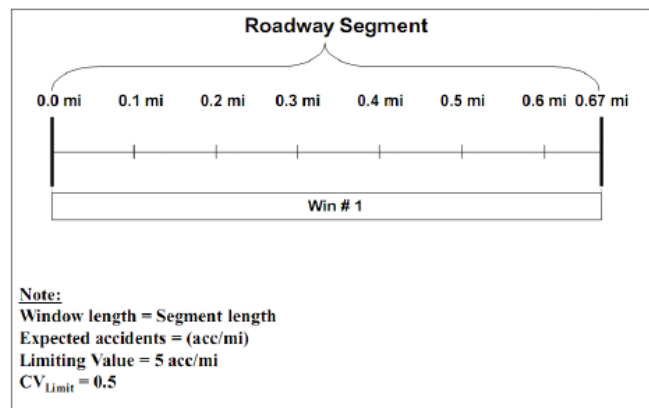
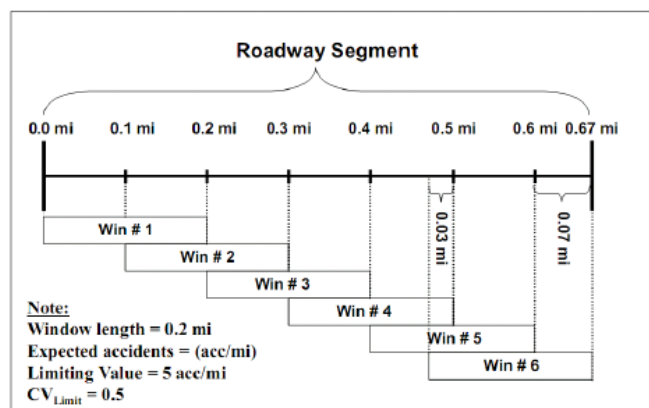
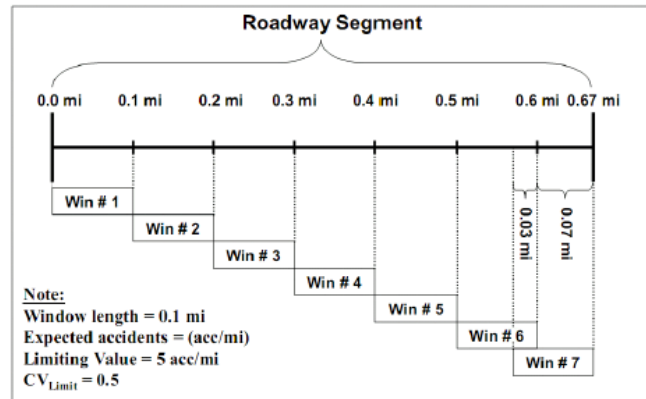
- 박민호 외(2006)는 EB 법을 이용하여 중앙분리대 설치에 따른 사고전환효과 분석을 실시하였으며, 분석 결과, 중앙분리대 설치로 인해 정면충돌 사고는 감소하였으나 곡선부에서의 후미추돌사고와 측면사고는 증가하는 것으로 나타났음.
- 또한, 사고심각도 측면에서 심각도가 높은 사망사고는 감소하지만 심각도가 다소 낮은 부상사고는 증가하는 것으로 분석되었음.
- 박규영 외(2006)는 도로안전시설물의 효과도를 비교하였는데 EB 법을 이용하여 시설물별 효과도 산정모형을 구축하였고 승산비(Odds Ratio)를 이용하여 상대효과도 산정기법을 제시하였음.
- 분석 결과, 중앙분리대와 가드레일은 상대효과가 크지 않게 나타났으나, 갈매기표지나 미끄럼방지포장 등은 구간사업에서 상대효과도가 매우 크게 나타나 향후 시설물 설치 시 구간측면에서 고려가 필요함을 확인하였음.
- 성낙문(2003)은 도로의 안전도를 평가하는 데 있어 교통사고에 내재되어 있는 특성(교통사고의 불확실성, 복잡한 교통사고 요인, 각 지점에 내재되어 있는 고유의 사고위험성)을 반영하는 방안으로 교통사고예측 모델을 이용하여 도로의 안전도를 평가하는 방법을 제시하였음.
- 정성봉 외(2009)는 EB 법을 이용하여 사고 잦은 곳 개선사업 우선순위 판정기법을 개발하였는데, 총 사고건수 추정 시 기존의 사고 예측방법인 비선형 회귀모형과 EB 법의 결과 값 모두 예측력이 높은 것으로 나타났지만 지점별 사고건수 예측력을 함께 고려할 경우엔 EB 법이 우수하다고 발표하였음.
- 강현건 외(2009)는 경상북도 내에서 발생한 4년간의 교통사고 자료를 대상으로 EB 법을 이용하여 예상 사고건수를 예측하였는데, 심각도는 대물피해환산법을 적

- 용하였고, 군집분석을 통해 유사한 지역을 선정하여 SPF를 도출하여 적용하였음.
- EB 법을 고속도로에 적용하는 데 있어, 그 구간을 정의할 때 분기점이나 IC 구간은 그 영향권으로 정의할 수 있으나, 일반 도로 구간에서는 주로 Sliding Window 법 혹은 Peak Search 법을 사용함.
 - Sliding Window 법은 고정된 길이의 Window가 정해진 길이만큼 이동하면서 그 Window 내에서 사고건수가 계산되는 방법임.
 - 각각의 도로 구간에서의 사고밀도는 그 구간 내에서 가장 큰 사고건수를 보이는 Window에 의해 정의됨.
 - 그러나, 이 기법에서 사용하는 고정길이 Window는 사고 위험도가 그 윈도우 내에서는 불변이라고 가정하고, 윈도우 내에서 사고위험지역과 그렇지 않은 지역이 공존할 경우, 국지적으로 높은 위험을 갖는 지점이 구간 내 사고위험이 낮은 지역으로 인해 위험도가 하향평준화(average out) 될 수 있음.
 - 또한 사고 위험을 높이는 모든 요인이 그 윈도우 내에 존재한다고 가정하므로 교통 혼잡으로 인한 주변 하류부의 2차 사고 발생을 함께 고려할 수 없다는 단점도 있음.



〈그림 2-3〉 Sliding Window 법(Srinivasan 외, 2011)

- Peak Search 법은 점진적으로 길이가 증가하는 Window로 그 구간을 나누며, Window는 구간의 첫 시작지점에서 출발하여 끝나는 지점에 닿을 때까지 길이가 증가함.



〈그림 2-4〉 Peak Search 법(Srinivasan 외, 2011)

- 예를 들면, 0.5km 길이의 구간에서는 증가 단위가 0.1km라고 가정할 때, 0.1, 0.2, 0.3, 0.4, 0.5km의 Window가 생성됨.
- 그 중 가장 큰 사고건수를 가지는 Window가 그 구간의 사고밀도를 정의하며, 통계적 유의성을 검증하게 됨.

- 유의성 검증은 CV(Coefficient of Variation)를 이용하며, 이는 추정값의 표준편차를 그 추정값으로 나눈 수치와 같음.
- 구간 내의 Window의 CV 값이 미리 정해진 경계 CV 값보다 낮으면, 그 구간은 단위거리당 추정 사고건수에 의해 순위가 매겨짐.
- 구간 내의 Window의 CV 값이 미리 정해진 경계 CV 값보다 높으면, Window 길이가 증가하게 되고, 위의 과정을 반복함.
- Peak Search 법의 장점은 너무 큰 Window 길이로 인해 무시될 수 있는 특정 지점의 문제들이 고려될 수 있으며, 이러한 추정값들이 자료의 임의성(randomness)이 아닌 통계적인 검증을 통과하여 신뢰성을 가질 수 있다는 점임.

제4절 Continuous Risk Profile(CRP) 법

- CRP 법은 현재 미국 캘리포니아 교통국(Caltrans)에서 도로 위험도 평가기법으로 사용하고 있는 Sliding Moving Window기법을 보완하고 나아가 연방정부에서 채택하고 있는 SPF법의 한계를 개선하기 위해 새로이 고안된 기법임(Chung 외, 2009).
- CRP 법은 먼저 도로를 따라 누적 사고횟수를 계산하고, 사용자에게 의해 정의된 참조 위험도와 비교하여 초과위험도를 추정하며, 좁은 범위에서 나타나는 변동 혹은 잡음(noise)을 제거한 뒤, 전체 도로구간에 걸쳐 연속적으로 초과위험도를 도출함.
- Moving Window의 크기를 $2L$, 1회 이동거리를 l 이라고 할 때, 구간 $[d-l/2, d+l/2]$ 내의 사고건수를 $A(d)$, 윈도우 $[d-L, d+L]$ 내의 평균사고건수를 $M(d)$ 라고 하면 수식은 다음과 같음.

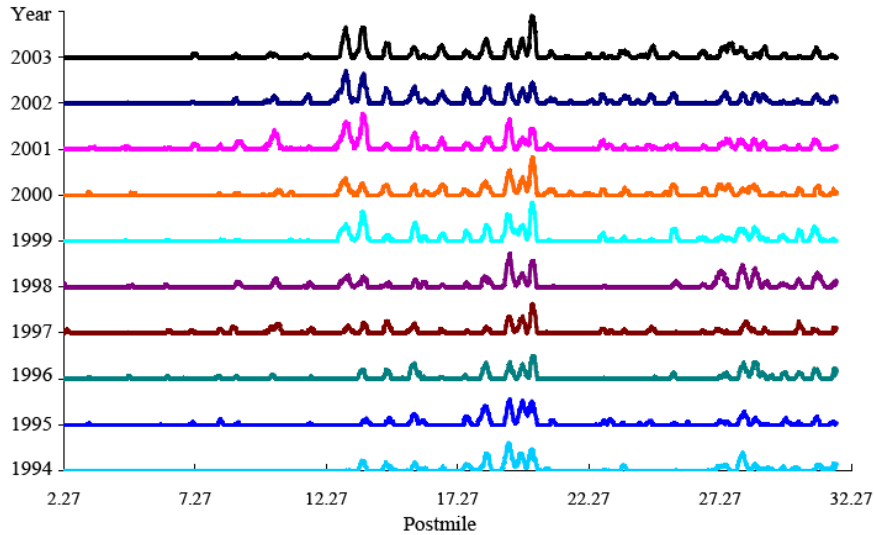
$$M(d) = \frac{\sum_{i=-\min(L/l, (d-d_0)/l)}^{\min(L/l, (d_{end}-d)/l)} A(d+i \times l)}{\min(L/l, (d_{end}-d)/l) + \min(L/l, (d-d_0)/l) + 1}$$

여기서,

$$d = d_0 + k \times l, \quad k = 1, 2, \dots, \frac{d_{end} - d_0}{l},$$

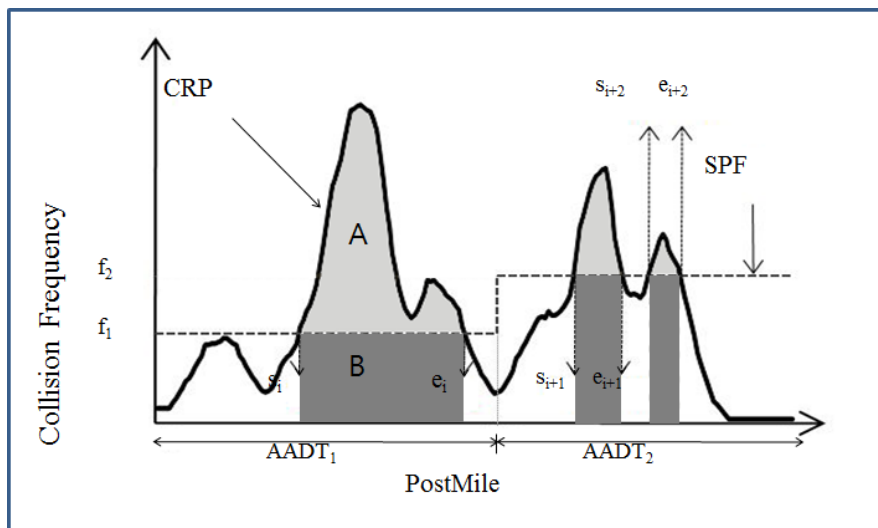
d_0 는 구간의 시작이정, d_{end} 는 구간의 끝이정, $d_0 < d_{end}$

- 이러한 단계에 따라 고속도로 연장에 걸쳐 연도별 CRP 다이어그램을 도시한 예는 <그림 2-5>와 같음.



<그림 2-5> CRP 다이어그램 예시(Chung 외, 2007)

- 이러한 방법에 따라, CRP 다이어그램이 구해지면, SPF 등 Reference Value 등과 비교하여, 그보다 통계적으로 큰 값을 가지는 부분을 사고취약구간으로 선정할 수 있음.



<그림 2-6> CRP 다이어그램을 이용한 사고취약구간 선정

- <그림 2-6>에서 A 구간이 사고취약구간으로 선정될 수 있으며, 보이는 바와 같이 이때의 사고취약구간은 CRP가 연속적인 함수이므로 해당 구간의 사고분포에 따라 Window 길이와 무관하게 나타나게 됨.

○ Chung 외(2007)는 캘리포니아의 샌프란시스코만 주변지역(Bay area)의 고속도로 413마일 구간에 CRP를 적용하였고, 그 결과 Sliding Moving Window 기법을 적용하였을 때보다 false positive 오류가 크게 줄어드는 결과를 얻었음.

○ 국내 고속도로에서 CRP 법을 적용한 연구들은 다음과 같음.

- 유준석(2008)은 window의 크기(L)를 곡선반경에 따라 차등 적용하여 국내 고속도로 4개 구간에 CRP 법을 적용하였으며, 그 결과 현행 국내 사고 잦은 구간 선정방법에서 선정된 구간들은 CRP 법을 통해서도 역시 사고위험도가 높은 구간으로 나타났음. 이는 CRP 법의 개념상, Window 내에 포함된 사고건수를 구간의 사고밀도로 구하기 때문에 당연한 결과로 해석할 수 있음.

- 또한 곡선반경별로 L의 길이를 차등 적용한 경우가 일률적으로 적용하였을 경우보다 위험도가 높은 지점이 더욱 뚜렷하게 표현되어 더 효과적인 것으로 분석되었음.

- 김용일 외(2010)는 1994년부터 2002년까지의 경부선 사고 자료를 이용하여 경부선 전 구간에 대한 CRP를 연도별로 구현하고, 구현된 CRP와 1994년부터 1999년까지 도로개선사업이 시행된 구간 목록을 바탕으로 사고 추이분석과 사전사후분석을 실시하여 사고 잦은 곳 선정 및 개선사업의 효과분석에 대한 CRP의 활용가능성을 제시하였음.

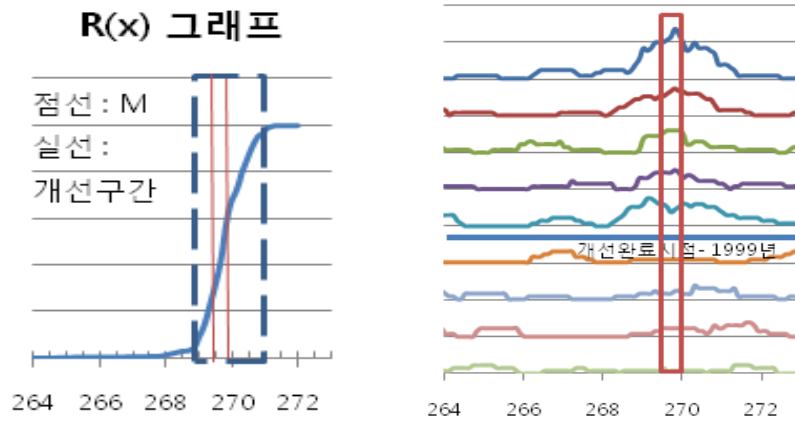
- 추이분석을 통해서, 시간의 경과에 따라 어떤 지역에서 CRP가 연도별로 계속 증가하는 경우에는 사고 잦은 곳으로 선정될 수 있는 기준을 만족시키기 이전에라도 위험도의 증가를 확인할 수 있기 때문에 예방적 조치를 취할 수 있는 장점이 있다는 점을 파악하였음.

- 또한, 개선사업이 시행된 구간에서는 사고위험도가 감소하는 현상과 개선구간의 인접구간으로 위험도가 전이되는 현상이 발견되었는데, 이러한 경우는 개선사업 전후의 효과분석 시 활용이 가능할 것으로 판단됨.

- 특히 개선사업 시행연도의 전후 연도에 대해 다음과 같이 CRP 값의 변화율을 구하는 함수(Chung et al., 2010)를 이용해 급격한 변화가 이루어지는 구간을 판별할 수 있었음.

$$R(x) = \frac{\int_a^x \{C_{i+1}(x) - C_{i-1}(x)\}^2 dx}{\int_a^b \{C_{i+1}(x) - C_{i-1}(x)\}^2 dx}$$

여기서, $C_i(x)$ 는 i 년도에 지점 x 에서의 CRP 값임.



〈그림 2-7〉 서울방향 264km~273km 구간 CRP(김용일 외, 2010)

- CRP 법은 교통량이나 점유율 데이터를 요구하지 않고 단지 사고데이터와 사고지점의 위치정보만 있으면 되기 때문에 적용이 아주 간편하며, 위험도분석을 위한 단위구간이라고 할 수 있는 window의 길이를 사용자가 정할 수 있다는 장점이 있으며, 시간대별로 반복적으로 나타나는 고속도로 상에서의 위험도의 변화를 보여줄 수 있음.
- 하지만 비교적 최근에 제안된 방법론이기 때문에 아직까지는 Sliding Moving Window 기법을 보완하는 수준에서 연구되었으므로 이론적으로 개선할 부분이 많이 남아 있음.
- 특히 사고의 심각도를 반영하지 않고 모든 사고를 등가로 간주하였으므로 사고건수법과 크게 다르지 않고, Moving Window와 연속된 그래프로 도시되는 프로파일을 통해 고정된 구간을 기준으로 하는 방법들의 한계¹⁾를 어느 정도 해소할 수는 있지만, 공간적 상호작용을 반영하기에는 부족하다는 단점도 있음.

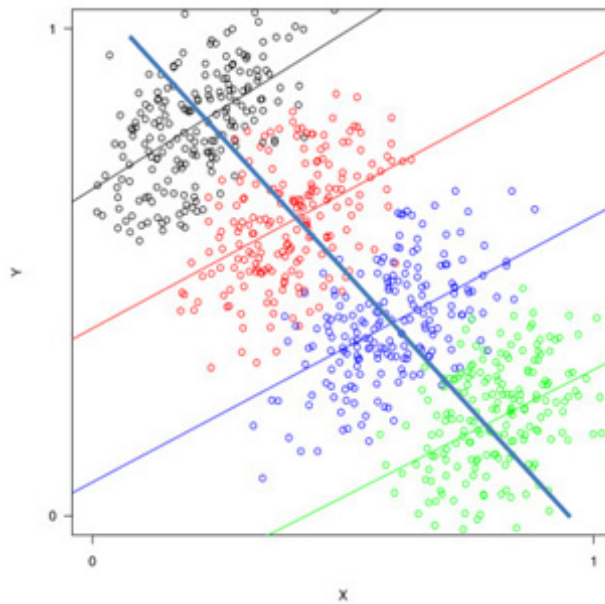
1) 동일 구간 내에서는 모든 지점에서 사고 확률이 동일하며, 잦은 사고를 유발하는 요인은 그 구간 내에서만 존재한다는 가정

제5절 Geographically Weighted Regression(GWR) 법

- GWR은 Fortheringham, Brunson, Martin의 공동연구로 개발된 공간 통계분석 기법으로 공간상에서 서로 다른 위치에 존재하는 데이터들에 각각 다른 회귀모델을 적용하여 공간적 비정상성(non-stationarity)을 탐색하는 방법임(김익희, 2005).
- 사고예측에 활용된 기존의 다양한 통계적 모형들은 연구 지역 전체에 걸쳐 평균적이며 전역적인 효과를 전제로 하는 것으로, 전체 연구대상범위에 대해서 변수별로 하나의 계수(coefficient)만을 제시하므로 지역적 특성에 관계없이 그 변수의 영향은 연구지역 전체에서 동일한 것으로 해석함.
- 반면, GWR은 지역별로 각각 다른 계수를 추정함으로써 변수들의 국지적인 영향력을 파악할 수 있는 결과를 제시하여, 기존의 회귀분석에서 고려할 수 없는 인접지역과의 공간적 상호작용을 반영할 수 있음.
- 일반적인 회귀분석에서는 데이터들 사이의 거리가 모델에 반영되지 않지만, GWR은 회귀지점에서 멀어질수록 낮아지는 차별적인 가중치를 데이터들에 적용한 후 분석을 실시하므로, 데이터들 사이의 거리가 고려됨으로써 Tolber의 지리학 제1법칙을 반영할 수 있음.
- 한편 전역적 평균에 관한 정보를 제공하는 것은 현실을 왜곡할 수 있는데, Simpson's paradox가 대표적인 경우임.
- <그림 2-8>에서 국지적 범위에서는 각각의 그룹에서 양의 관계를 이루고 있으나, 데이터를 통합하여 전역적 범위에서 분석했을 경우에는 음의 관계를 확인할 수 있으며, 이렇듯 전역적 분석은 현실을 반대로 해석할 수도 있는 위험성을 갖고 있음.

〈표 2-4〉 전역적 통계량과 국지적 통계량의 차이(Fotheringham 외, 2002)

전역적 통계 방법	국지적 통계 방법
전체 연구지역의 자료 요약	전역적 통계량의 국지적 분해
한 가지 결과만 도출	서로 다른 결과들이 도출
지도화할 수 없음	지도화할 수 있음
GIS에 부적합	GIS에 적합
비공간적 혹은 공간적으로 제약	공간적
공간적 동질성을 강조	공간적 이질성을 강조
규칙성 혹은 법칙 탐색	예외 혹은 지역적 특징 탐색



〈그림 2-8〉 Simpson's paradox 예시(박신행, 2010)

- GWR은 전통적인 회귀식을 사용함으로써 친숙한 방법을 제시하면서 동시에 회귀 구조에 지역적 공간 관계들을 반영할 수 있음(Fotheringham 외, 2002).
- 즉 GWR은 대부분의 사고예측모형과 마찬가지로 회귀모델을 사용한다는 점에서 기존 모형들의 장점들이 적용된다고 할 수 있으며, 나아가 기존 모형들로는 파악하기 어려운 공간적으로 이질적인 패턴을 확인할 수 있다는 장점을 가지고 있음.
- 또한 GWR 분석 결과를 토대로 회귀계수들을 GIS 맵을 통해 시각화함으로써 데이터들의 본질과 정확성에 대해 통찰할 수 있고, 공간상에 존재하는 변수들의 관계와 변동에 대한 상세한 이해를 할 수 있으며, 공간적 비정상성을 무시하는 전통적 접근 방식의 문제점을 확인할 수 있음(Fotheringham et al., 1996).
- GWR 법은 이러한 장점 덕택에 사회학이나 지리학 분야의 연구에 활발하게 적용되고 있으며, 교통 분야에 적용된 사례로는 연평균일일교통량(Annual Average Daily Traffic; AADT)을 추정하기 위해 GWR을 활용한 Zhao and Park(2004)의 연구가 있음.
 - 이 연구에서는 교통량 측정지점이 있는 도로구간의 차로수, 지역별 경제활동이 교통량에 미치는 영향, 측정지점으로부터 일정 반경 이내의 인구 및 고용자 수, 그리고 측정지점이 위치한 도로가 고속도로로 직접 연결되는지 여부 등 총 5개의 변수를 이용하여 OLR(Ordinary Linear Regression)과 GWR로 AADT를 추정하였음.

– 그 결과 GWR 모형이 독립변수들의 설명변수(AADT)에 대한 지역별 영향을 밝히는데 보다 정확하고 유용한 결과를 도출할 수 있다는 결론을 얻었음.

- GWR을 교통사고의 발생에 적용한 것은 박신행(2010)의 연구가 대표적임.
- GWR의 이론적인 면을 살펴보면, 일반적인 전역적 회귀 모형식은 다음과 같이 나타남(박신행, 2010).

$$y_i = \beta_0 + \sum_{k=1}^p \beta_k x_{ik} + \epsilon_i$$

- GWR은 이러한 전통적인 회귀식을 다음과 같은 좌표가 부여된 지리적인 모델로 확장한 것임.

$$y_i = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) x_{ik} + \epsilon_i$$

여기서,

y_i : 지점 i 에서 종속변수($i = 1, 2, \dots, n$. n 은 관측 수),

x_{ik} : 지점 i 에서 k 번째 독립변수,

β_k : 모형에서 추정된 k 번째 파라미터,

ϵ_i : 지점 i 에서의 오차항,

p : 파라미터의 개수,

(u_i, v_i) : i 번째 회귀중심점의 좌표,

$\beta_k(u_i, v_i)$: 회귀 분석의 중심점에서 k 번째 독립 변수의 회귀계수

- GWR에서는 일반적인 선형회귀모형과 달리, 독립변수별로 모든 회귀분석 중심점들에 대해 회귀계수들이 추정됨.

– 먼저, 선형회귀계수의 추정치는 다음 공식에 의해서 계산됨.

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

– 반면, GWR에서 회귀계수 β 의 추정치는 다음과 같이 계산됨.

$$\hat{\beta}(u_i, v_i) = (X^T W(u_i, v_i) X)^{-1} X^T W(u_i, v_i) y$$

– 여기서, $\hat{\beta}$ 은 추정된 파라미터의 벡터이며, X 는 독립변수 값을 포함하는 행렬이고, y 는 종속변수 값의 벡터임.

– $W(u_i, v_i)$ 는 행렬의 대각요소들을 제외한 값들이 0인 $n \times n$ 행렬로, 대각선에 존재하는 값들은 회귀중심점 i 에 대해 주변 n 개의 관측치에 부여하는 지리적 가중치를 의미함.

$$W(i) = \begin{bmatrix} w_1(i) & 0 & \dots & 0 \\ 0 & w_2(i) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & w_n(i) \end{bmatrix}$$

—가중치 행렬 $W(i)$ 에서 $w_k(i)$ 는 지점 i 에 대한 계산에서 data point k 에 부여되는 가중치를 의미하며, 거리가 멀어짐에 따라 공간적인 영향이 감소하는 현상을 반영하기 위해 Gaussian 함수를 적용할 수 있음.

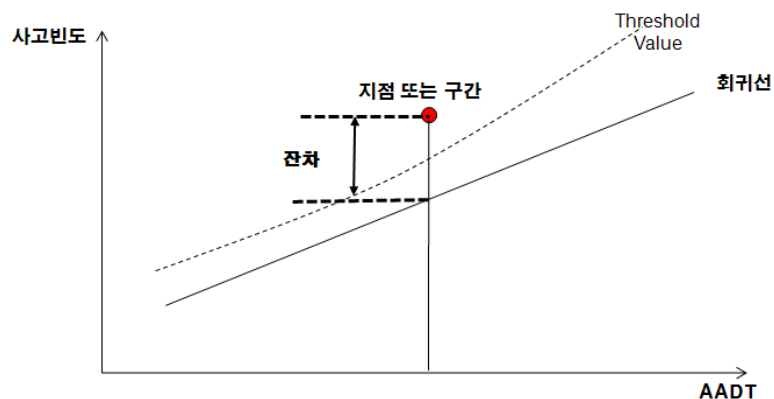
$$W_k(i) = \frac{1}{\sqrt{2\pi}} e^{-0.5(d_k(i)/h)^2}$$

—여기서 $w_k(i)$ 는 데이터 세트에서 지점 i 에 대해 k 번째 관측치의 상대적인 지리적 가중치이고, $d_k(i)$ 는 k 번째 관측치와 지점 i 사이의 거리를 의미하며, h 는 가중치가 적용되는 범위를 나타냄.

○교통사고 예측모형을 기반으로 위험구간을 판정하는 기존의 도로 안전도 평가방법은 구간의 특성을 나타내는 변수들을 독립변수로 채택하고, 모형을 통해 해당 구간의 사고 예측치인 기대 사고건수를 산출한 다음, 이를 관측치와 비교하며, 이때, 관측된 사고건수가 예측 사고건수보다 유의미하게 높은 구간을 개선대상구간으로 선정하게 됨.

○<그림 2-9>에서 지점 또는 구간의 관측 사고빈도와 회귀선의 Y축 값인 예측치의 차가 잔차이며, 어떤 구간에서 실제 발생한 사고건수가 해당 구간에 대해 예상되는 사고건수와 얼마나 차이가 나는지를 의미함.

○이 값을 전체 구간의 잔차들에 대해 표준화한 것이 표준잔차(standard residual)이며, 이를 통해 신뢰구간을 벗어나는 자료를 쉽게 찾을 수 있음.



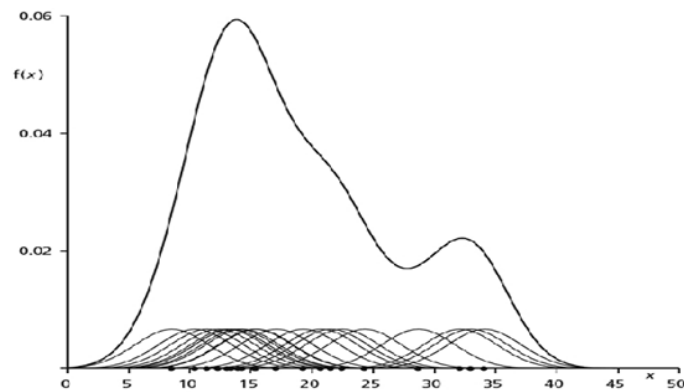
<그림 2-9> 사고예측모형을 이용한 사고취약구간 판별 개념(박신형, 2010)

- 이 그림에서는 임계값(threshold value)이 유의수준을 의미하는데, 예를 들어, 잔차의 크기가 전체 잔차들 중에서 상위 1%에 해당하는 구간들은 구간의 환경적 특성에 비추어 유의미하게 높은 사고가 발생하는 구간, 즉 사고취약구간이라 할 수 있음.
- 여기서 사고취약구간의 판별 기준으로는 캘리포니아 교통국(California Department of Transportation, 2002)에서는 0.5%, Kononov 외(2003)에서는 표준편차의 1.5배를, 성낙문(2003)의 연구에서는 5%를 쓰고 있음.
- 이러한 기준의 설정은 투입 가능한 예산의 규모에 따라 달라질 수 있는데, 예산이 충분히 확보되었을 경우에는 기준을 완화하여 보다 많은 구간을 선정하고, 예산이 부족할 경우에는 기준을 강화하여 보다 유의미한 구간들을 선정할 수 있음.
- 하지만 이 기준을 완화할 때는 시급한 개선을 요하지 않는 구간이 추가 선정될 수 있으므로 false positive 오류 발생가능성이 커지고, 기준을 강화할 때에는 시급한 개선이 필요한 구간이 제외될 수 있으므로 false negative 오류 발생가능성이 커질 우려가 있기 때문에 연구자의 적절한 판단이 필요함.

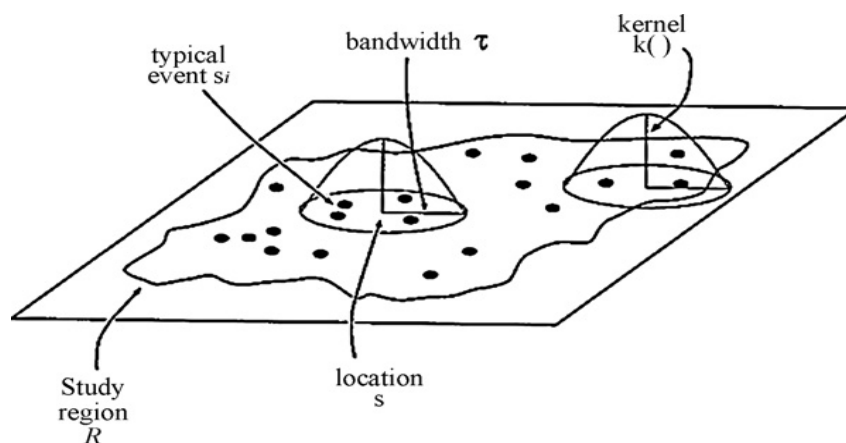
제6절 Kernel Density Estimation(KDE) 법

- 교통사고밀도를 이용한 공간분석은 공간적으로 어느 지역에 사고가 집중되어 있는지 식별하기 쉽도록 시각적으로 효과적인 정보를 제공할 수 있다는 점에서 유용한 기법임.
- 일반적으로 사고 지도는 두 건 이상의 사고가 발생한 지점의 사고 집중도를 정확히 반영하지 못하는데, 그 이유는 한 지점에서 여러 건의 사고를 표시하면, 개별 사고가 하나로 중첩되어 표시되므로 사고건수가 구분되어 보이지 않기 때문임 (Pulugurtha 외, 2007).
- 따라서, 사고의 집중도를 지도에 표현할 때, 지점표시 대신에 적절한 커널함수를 이용하면 사고밀도의 분포를 알기 쉽게 표현할 수 있음.
- <그림 2-10>와 <그림 2-11>에서와 같이 커널밀도는 point나 line feature들로부터 커널함수를 사용하여 단위지역당 밀도를 계산하고 부드럽게 처리된 형태로 표현되며,

- 개념적으로, 커널의 값은 사고 발생지점에서 가장 크고, 그 지점으로부터의 거리가 멀어짐에 따라 값이 줄어들다가 탐색반경의 경계에 이르면 0이 됨(Silverman, 1986).
- 이 탐색반경을 대역폭(bandwidth)이라고 부르는데, 이 값은 커널밀도 분석에서 이벤트의 영향범위를 의미하기 때문에 가장 중요한 기준이라고 볼 수 있음(Silverman, 1986; Fotheringham et al., 2002).
 - 사고취약구간 분석에서 bandwidth는 사고가 발생할 가능성이 있는 수평적 범위를 의미하며, 어떤 지점에서 사고가 발생하였다는 것은 그 지점의 도로구조 및 교통 환경으로 인한 공간적인 효과가 작용한 결과라고 할 수 있음. 따라서 이 지점을 중심으로 한 일정 범위 내의 지역은 이러한 공간적 영향에 의해 사고가 발생할 가능성이 존재한다고 볼 수 있음.



〈그림 2-10〉 커널함수의 예



〈그림 2-11〉 커널밀도의 개념(Bailey and Gatrell, 1995)

- bandwidth는 그 크기를 어떻게 결정하느냐에 따라 추정량의 편의(bias)와 분산(variance) 사이에 trade-off 관계가 존재하는데, bandwidth 값이 너무 작으면 국지적 추정량의 분산이 커지는 반면, bandwidth 값이 너무 크면 편의가 커지게 됨(박신행, 2010).
- Silverman(1986)은 bandwidth의 결정에 대해, 커널추정의 목적이 통계적 모형과 가설을 제시하기 위한 데이터 탐색에 있다면 연구자의 주관적 판단에 따르도록 제안하고 있으며, Anderson(2009) 역시 bandwidth 결정 과정은 분석의 목적에 따라 어느 정도 주관적인 판단에 따르는 것이라 기술하고 있음.
- 이론적으로, 통계학적인 측면에서는 AIC(the corrected Akaike Information Criterion) 값을 최소화하거나(Akaike, 1974), Cross Validation(CV)을 최소화하는 방법이 있지만(Fotheringham et al., 2002), 현실적으로는 분석의 목적에 맞게 bandwidth를 설정하는 것이 더욱 바람직함.
- 어떤 주어진 도로에 대한 전반적인 위험도를 파악하기 위해 커널밀도를 분석할 경우에는 우선 상대적으로 큰 값의 bandwidth를 적용함으로써 전역적 추정량(global estimator)을 살펴보는 것이 좋은 출발점이 될 수 있음(Flahaut et al., 2003).
 - 그 후에 위험도가 높은 것으로 확인된 구간에 대해 확대된 축척 레벨에서 보다 작은 값의 bandwidth를 적용함으로써, 상세한 위험도의 변동을 탐색하여 정확한 위험구간의 위치를 확인할 수 있음.
- Xie & Yan(2008)은 위험구간 판별에 적합한 bandwidth를 찾기 위해 국지적 범위와 전역적 범위에서 20m, 100m, 250m, 500m, 1000m, 2000m의 bandwidth를 적용하여 bandwidth의 크기에 따른 영향을 분석하였음.
 - 그 결과 작은 값의 bandwidth(20m, 100m, 250m)를 적용한 경우에는 사고의 국지적 영향이나 위험구간의 판별에 적합한 패턴이 나타났고, bandwidth가 커질수록 국지적인 위험구간이 주변 지역의 위험구간과 결합하여 더 큰 군집이 형성되는 것을 발견하였음.
 - 보다 넓은 지역을 관찰하는 경우에는 더 큰 값의 bandwidth(500m, 1000m, 2000m)를 사용하는 것이 효과적임을 파악하였음.
- 커널밀도의 형태를 결정하는 또 다른 파라미터인 커널은 커널함수의 값을 의미함.

- 즉 사고 발생 지점에서 가장 큰 값을 가지고, 거리가 멀어짐에 따라 그 값이 감소하며, bandwidth의 경계에서 0의 값을 가지는 사고의 확률이나 사고위험도를 뜻한다고 할 수 있음.
- 대부분의 커널밀도를 활용한 관련 연구들에서는 모든 사고를 1이라는 등가로 간주하여 커널밀도를 추정하였기 때문에 사고 건수법에 의한 사고 잦은 곳 선정 방식과 사실상 차이가 없었으며, 단지 결과를 시각적으로 제시하는 것 이상의 의미는 갖지 못했음.
- 개념적으로, 커널함수는 공간을 사용자가 결정한 크기의 셀(cell)로 정의하고, 각 cell은 함수를 통해 계산되는 값을 갖게 되는데, 전술한 바와 같이 사고가 발생한 지점에서부터 사용자가 미리 설정한 탐색반경인 bandwidth에 이르기까지 그 값이 차등적으로 계산됨.
- 또한 여러 사고지점으로부터 영향을 받는 위치에 있는 cell은 각각의 사고지점들로부터 얻게 되는 cell 값들을 중첩한 값을 갖게 되는데, 예를 들어 Q라는 cell이 a, b, c 세 지점에서 발생한 사고의 탐색반경 안에 위치하고 있다면, 각각의 지점으로부터 계산된 Q의 cell 값이 $Q(a)$, $Q(b)$, $Q(c)$ 라고 할 때 Q의 최종 cell 값은 $Q_{total} = Q(a) + Q(b) + Q(c)$ 로 계산됨.
- 바꾸어 말하면, Q라는 지점이 갖는 특성이 사고 a, b, c의 발생에 공히 영향을 미친다고도 해석할 수 있으므로, cell 값이 큰 지점은 사고가 잦을 가능성이 높은 지점이라고 판단할 수 있음(박신행, 2010).
- cell 값을 결정하는 커널함수의 종류는 여러 가지이나, 일반적으로 사용되는 커널함수는 다음과 같음.

$$\begin{aligned}
 \text{Bi-square} & : \text{Kern}(u) = 1 - u^2, |u| \leq 1 & , 0 \text{ otherwise,} \\
 \text{Quadratic} & : \text{Kern}(u) = 0.75(1 - u^2), |u| \leq 1 & , 0 \text{ otherwise, and} \\
 \text{Gaussian} & : \text{Kern}(u) = (2\pi)^{-1/2} \exp(-u^2/2)
 \end{aligned}$$

- 사고 지점 t_i 에 인접한 cell t에서의 커널값을 $Q_i(t)$ 라고 하면,

$$Q_i(t) = \text{Kern}\left(\frac{t_i - t}{h}\right)$$

로 나타낼 수 있는데, 여기서 $t_i - t$ 는 사고 발생지점(t_i)과 인근지점(t) 간의 집중도를 나타내고, 이 값은 bandwidth인 h 의 값에 따라 크기가 조절됨.

- 만약 cell t가 n개의 사고지점으로부터 영향을 받는다고 하면, cell t의 값은 다음 수식을 통해 구할 수 있음.

$$Q(t) = \sum_{i=1}^n Q_i(t)$$

- 박신형(2010)은 이러한 개념을 확장하여 심각도를 반영한 커널함수(Severity-weighted Kernel Function)를 고안하였음.

- 기본 개념은 심각도 지수를 커널함수에 곱의 형태로 반영하여 cell마다 사고의 심각도에 따라 수직적인 강도를 반영하는 것으로, 심각도 지수는 다음 식으로 나타낼 수 있음.

$$SI(i) = W_F(i) \cdot C_F + W_I(i) \cdot C_I + PDO$$

여기서,

$SI(i)$: i지점에서의 심각도 지수,

$WF(i)$: i지점에서 발생한 사고가 사망사고일 때 부여되는 가중치,

$WI(i)$: i지점에서 발생한 사고가 부상사고일 때 부여되는 가중치,

CF : 사망자 수 또는 사망사고 발생건수,

CI : 부상자 수 또는 부상사고 발생건수,

PDO (Property Damage Only) : 대물사고에 대해서는 1의 값을 부여

- 커널함수로 Gaussian 함수를 사용한다면, 심각도가 반영된 새로운 커널함수는 다음과 같이 기술될 수 있음.

$$K = (2\pi)^{-1/2} \exp(-(t_i - t)^2 / 2h^2) \cdot SI$$

- 커널밀도를 적용하여 사고취약구간의 판별을 시도한 연구는 다음과 같이 정리될 수 있음.

- Flahaut 외(2003)는 사고 잦은 곳을 판별하기 위해 공간적 자기상관 지수(spatial autocorrelation index)를 이용하는 방법과 커널밀도추정법을 각각 적용한 뒤 두 가지 방법을 비교하였음.

- 두 방법 모두 장단점을 갖고 있으나, 여러 값의 파라미터를 각각 적용해본 결과 특정한 파라미터에서 거의 유사한 결과를 얻을 수 있었음.

- Pulugurtha 외(2007)는 보행자 사고 위험지역을 판별하여 이 지역들에 대한 우선순위를 부여하기 위해 보행자 사고의 공간적 패턴을 탐색하는 방법론을 개발하였음.

- 커널밀도 추정을 통해 29개의 위험구간을 선정하고, 이 구간 내에서 여러 가지 위험도 측정 기법을 적용하여 순위를 결정하였음. 먼저, 사고빈도, 심각도를 가중한 사고빈도, 교통량에 기반한 사고율, 인구 대비 사고율, 사고밀도를 구하였고, 이 중에서 보다 의미 있는 것으로 판단한 단위면적당 사고밀도(CD_A)와 교통량 대비 사고율(CR_{VV}), 그리고 인구대비 사고율(CR_{PP})을 결합하여 다음과 같이 순위합과 사고점수를 구하였음.

- 먼저 순위합(Sum-of-the-ranks; SR)은 각각의 기준에 대한 순위를 산술평균하여 구한 것으로, 공식은 아래와 같음.

$$SR = \frac{Rank(CD_A) + Rank(CR_{VV}) + Rank(CR_{PP})}{3}$$

- 사고점수 방법(Crash score; CS)은 기준별로 최대값에 대한 해당 구간 값의 비율을 백분율로 구한 뒤 총합을 계산하는 것으로 수식은 아래와 같음.

$$Score\ CD_A = \frac{CD_A}{maximum\ CD_A} \times 100$$

$$Score\ CD_{VV} = \frac{CD_{VV}}{maximum\ CD_{VV}} \times 100$$

$$Score\ CD_{PP} = \frac{CD_{PP}}{maximum\ CD_{PP}} \times 100$$

$$CS = Score\ CD_A + Score\ CR_{VV} + Score\ CR_{PP}$$

여기서,

CD_A : 단위면적당 사고밀도,

CR_{VV} : 교통량 대비 사고율,

CR_{PP} : 인구 대비 사고율

- 개별적인 방법과 결합방법을 적용하여 순위를 결정한 결과, 결합방법을 이용한 경우가 개별 방법을 이용하였을 때보다 일관성 있는 결과를 얻을 수 있었음.

- 또한 결합방법인 순위합 방법과 사고점수 방법의 결과를 비교하면, 개별 방법들로부터 산출된 점수를 반영한 사고점수 방법이 단순히 개별 방법의 순위의 합으로 순위를 재산정하는 방법보다 사고원인의 규명과 위험지역 순위 결정에 더 강건한(robust) 결과를 제시하는 것으로 분석되었음.

- 하지만 이 방법은 사고건수에 기반하여 커널밀도를 추정하였고, 초기 단계에서

후보지를 결정하는 목적으로만 이용하였으며, 실제 우선순위 결정에 적용된 방법들, 특히 결합방법에 활용된 사고밀도와 사고율은 해당 지역 또는 구간 내에서는 어떤 지점에서도 항상 일정한 값을 갖는다고 가정하기 때문에 공간적인 차이를 반영하지 못하는 한계가 있음.

〈표 2-5〉 커널밀도를 적용한 기존 연구 사례

연구자	심각도 반영 여부	KDE 활용	연구의 특징
Flahaut et al.(2002)	X	사고건수 기반의 사고밀도 분석	· 사고건수를 기반으로 분석
Pulugurtha et al. (2007)	X	상세분석을 위한 1차 후보지 선정	· 사고건수를 기반으로 분석
Erdogan et al. (2008)	X	사고건수 기반의 사고밀도 분석	· 사고건수를 기반으로 분석
Xie & Yan (2008)	X	사고건수 기반의 사고밀도 분석	· 사고건수를 기반으로 분석 · Bandwidth 크기에 따른 다양한 결과 제시
Anderson (2009)	X	사고건수 기반의 사고밀도 분석	· 현실적 활용방안에 중점 · 분석범위와 사고 내용 및 지역적 특성에 따라 hotspot을 cluster와 group으로 범주화하여 적절한 대책 제시
박신형 (2010)	O	사고건수와 심각도 기반의 사고밀도 분석	· KDE에 처음으로 사고 심각도 적용 · GWR과 결합하여 사고취약구간 선정

- Erdogan et al.(2008)은 사고 잦은 곳 판별을 위해 두 가지 방법을 적용하였는데, 하나는 포아송 분포를 이용하여 5% 또는 1%에 해당하는 확률을 갖는 사고건수를 기준으로 하여 그 기준을 초과하는 지역을 선정하는 방법이고, 또 하나는 bandwidth를 500m로 부여한 커널밀도를 활용하는 분석법임.
- 각각의 방법을 적용한 결과, 거의 동일한 지역들이 사고 잦은 곳으로 선정되었는데, 이 연구 또한 공간적 상호작용을 고려하지 않는 기본적인 통계기법인 포아송 분석법과 단순히 사고의 건수만을 반영하는 커널밀도 분석법을 사용하였다는 한계가 있음.
- Anderson(2009)은 KDE를 활용하여 사고건수 기반의 사고밀도 분석을 수행하였는데, 이를 통한 현실적 활용방안에 중점을 두었음.

- 먼저 밀도분석을 통해 사고밀집지역을 hotspot으로 선정하였고, 분석의 공간적 범위와 사고 내용, 그리고 지역적 특성에 따라 hotspot을 cluster로 범주화하고, 비슷한 특성을 가진 cluster들을 다시 group으로 범주화한 후, 각 cluster나 group의 특성에 따라 적절한 대책을 제시하였음.
- 이상의 기존 연구들을 정리하면 <표 2-5>와 같음.



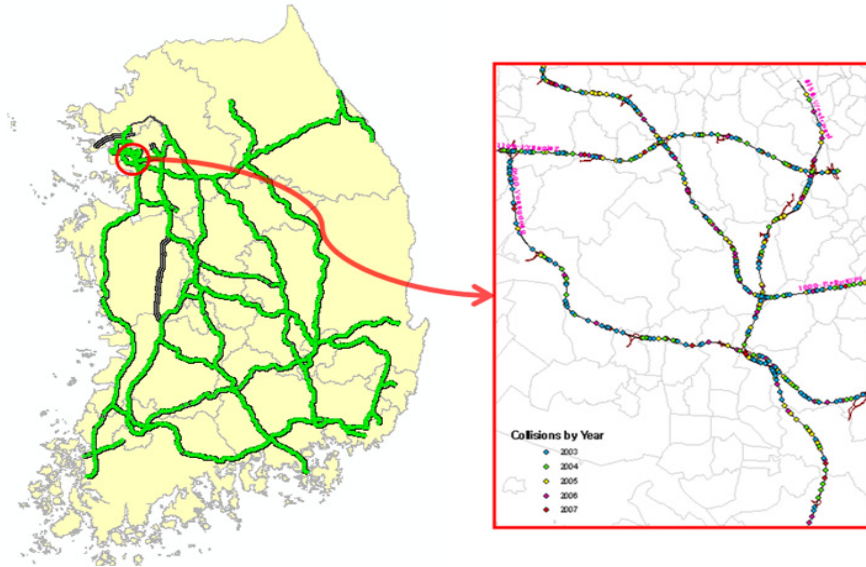
고속도로 사고취약구간 선정 방법론 적용 사례

제1절 사례 연구 개요

- 이 장에서는 고속도로 사고취약구간 선정 방법론 적용사례로, 박신형(2010)의 연구를 소개하고자 함.
- 이 연구는 GWR 법과 KDE 법을 이용하여, 경부고속도로 서울~부산 간 상행선 구간에 대하여 2006년부터 2008년까지 3년간의 데이터를 이용하여 사고취약구간을 선정하였음.
- 이는 최신 방법론 및 데이터를 이용하여, 장거리 구간에 걸쳐 종합적이고 자세하게 사고취약구간을 선정한 예로, 앞 장에서 검토된 최신 방법론을 서울시 도시고속도로 구간에 적용하는 데에 좋은 사례가 될 수 있음.

제2절 대상 구간 데이터 구축

- 분석에 사용된 데이터는 <그림 3-1>에서와 같이 지오코딩된 고속도로 사고 자료 중 2006년부터 2008년까지 경부고속도로 상행선에서 발생한 총 842건의 사고 자료임.



〈그림 3-1〉 고속도로 데이터 지오코딩 자료(박신행, 2010)

- 경부고속도로는 2006년 이후로 고속도로 이정이 변경되지 않았고, 운전자의 이용도가 높은 만큼 사고도 비교적 빈번하게 발생하므로 양적으로 분석을 수행할 수 있을 정도의 데이터 샘플 확보가 용이하였음.
- 이론적으로 교통사고의 연구에 적절한 사고 자료의 기간도 고려되었는데, 현재 국내 고속도로의 사고 잦은 구간은 전년도 1년간 발생한 사고건수를 기준으로 하고 있기 때문에, 일시적인 사고의 증가나 감소로 인한 영향을 크게 받을 수 있음.
 - 기존 방법은 regression to the mean 현상을 적절히 고려하지 못하기 때문에, 지속적으로 높은 위험도를 가진 구간이 선정될 확률이 낮으므로 false positive 및 false negative 오류의 발생 가능성이 커질 수 있음.
 - 2006년부터 2008년까지 3년간의 사고 자료를 이용하는 것은 어느 한 해에만 사고가 많이 발생한 구간에 대해 3년의 기간으로 평균화할 때, 그 영향이 많이 완화될 수 있음.
- 공간적 분석 단위를 정하기 위해 고속도로의 구간별 특징을 나타낼 수 있는 기하구조 요소가 검토되었음.
 - 국내외에서 수행된 많은 선행연구에서는 주로 평면선형, 종단구배, 편구배 등이 사고 발생 및 심각도에 유의한 요소들인 것으로 분석되었는데, 어느 요소가 보다 더 큰 영향을 미치는지에 대해서는 저마다 다른 결과가 제시되었음.

- 여기서는 Kwak 외(2010)와 Park 외(2010)의 연구에서 도출된 결과를 참조하여 공간분석 대상구간의 기준을 선정하였음. 이들 연구에서는 평면선형이 종단구배보다 사고 발생이나 사고 심각도에 더 유의한 영향을 주는 것으로 확인되었으므로, 평면선형을 기준으로 경부고속도로 노선을 524개 구간으로 나누었음.
- 842건의 사고를 구간별로 집계하여 다음과 같이 네 종류의 종속변수와 10개의 독립변수가 생성되었음(<표 3-1>).
- 변수 선정의 기준은 한국도로공사의 사고데이터에서 기록하는 정보들 중 도로의 기하구조와 교통조건에 관계되는 항목임.
- 추가로 과속단속카메라, 휴게소, 터널, 교량 등 도로시설물의 위치정보와 선형에 관한 세부 정보를 수집함.
- 주요변수에 대한 추가적인 설명은 다음과 같음.
- EPDO(Equivalent Property Damage Only)는 한 구간 내의 사고들 중 인명 피해 없이 물질적 피해만 발생한 사고를 1건으로 간주하였을 때, 사망자가 발생한 사고는 12건, 사망자 없이 부상자만 발생한 사고는 3건으로 계산한 사고건수에 대한 가중합을 나타냄.
- Wfreq(Weighted Frequency)는 한 구간 내의 사고건수에 인명피해를 가중하여 계산한 값임. 한 구간에서 사고가 3건이 발생하였는데, 총 사망자수가 3명, 부상자수가 7명이려면, $Wfreq = 12 \times 3 + 3 \times 7 + 3 = 60$ 이 됨.

〈표 3-1〉 분석에 사용된 변수(박신행, 2010)

변수 구분	변수명	설명
종속변수	Freq	해당 구간에서 3년간 발생한 사고건수
	EPDO	각 사고의 피해수준에 따라 사망사고 : 부상사고 : 물피사고를 12 : 3 : 1로 가중한 사고건수
	Wfreq	사고건수에 사망자당 12, 부상자당 3의 가중치를 더하여 계산한 가중 사고건수
독립변수	NumLane	해당 구간 차로 수
	Bridge	구간 내 교량의 개수
	Tunnel	구간 내 터널의 개수
	Camera	구간 내 과속단속카메라의 개수
	Offramp	구간 내 진출부의 개수
	OnRamp	구간 내 진입부의 개수
	Restarea	구간 내 고속도로 휴게소의 개수
	TG	구간 내 톨게이트의 개수
	EXPO	교통량을 기반으로 산출한 노출도

- EXPO(Exposure)는 사고발생 가능성에 대한 노출 정도를 나타내는 변수로, 교통량은 교통사고 발생의 주요 원인 중 하나이며, 일반적으로 교통량이 많은 곳에서 사고빈도도 높게 나타남.
- 따라서 도로구간을 기준으로 봤을 때, 구간의 교통량이 많으면 그만큼 그 구간에서 사고발생요인에 대한 노출이 많아지는 것으로 볼 수 있으므로 노출도를 각 도로 구간의 공간적 특성인 교통조건을 반영하는 변수로 채택되었음.
- 이 변수는 아래 식과 같이 구간별로 연도별 발생 사고건수 및 연평균일일교통량(AADT: Annual Average Daily Traffic)을 가중 평균하여 산출되었음.

$$EXPO = \frac{L \times 365 \times Y \times AADT}{10^6}$$

$$= \frac{L \times 365}{10^6} \times \left(\sum_{i=2006}^{2008} (AADT_i \times \frac{Freq_i}{\sum_{i=2006}^{2008} Freq_i}) \right)$$

여기서,

L: 도로 구간의 길이 (km),

Y: 기간(년)

- 단, 해당구간에서 3년 동안 단 한 건의 사고도 발생하지 않은 경우에는 아래 식과 같이 3년간의 AADT를 평균한 값을 EXPO의 값으로 계산되었음.

$$EXPO = \frac{L \times Y \times 365}{10^6} \times \frac{\sum_{i=2006}^{2008} AADT_i}{3}$$

제3절 GWR 법과 KDE 법을 이용한 사고 취약구간 선정

1. GWR 법의 적용

- 경부고속도로 구간에 GWR 법을 적용한 결과는 <표 3-2>와 같음.
- AIC는 여러 개의 추정된 통계모형의 적합도를 비교하여 보다 우수한 모형을 선택하는 기준인데, 더 작은 값을 가진 모형일수록 더 우수하다고 평가할 수 있음.
 - EPDO가 비슷한 AIC값을 보이고 나머지 변수들에서는 지리적 가중회귀분석을 실시한 모형이 더 우수한 것으로 나타남.

- R^2 와 Adjusted R^2 의 비교에서도 모든 종속변수들에 대해 전역적 회귀분석을 실시한 결과보다 지리적 가중회귀분석을 적용한 모형의 설명력이 더 우수함을 보여주고 있음.

〈표 3-2〉 전역회귀분석과 지리적 가중회귀분석의 결과 비교(박신행, 2010)

항목	전역적 회귀분석				지리적 가중회귀분석			
	Freq	EPDO	Wfreq	Acc_Cost	Freq	EPDO	Wfreq	Acc_Cost
AIC	2,066	3,390	3,869	19,932	2,040	3,390	3,863	19,888
R ²	0.519	0.477	0.623	0.560	0.586	0.520	0.694	0.633
Adj. R ²	0.510	0.467	0.615	0.551	0.551	0.484	0.668	0.602

- GWR을 통해 국지적으로 도출된 회귀식을 통해 선정된 사고취약구간은 <표 3-3>과 같음.

〈표 3-3〉 GWR의 그룹별 결과 비교(박신행, 2010)

그룹	사고건수	사망자	부상자	사고비용 (천원)	구간 총 연장(m)	단위거리(m)당 사고비용(천원)
GWR_F	104	5	56	4,221,431	17,455	241.847
GWR_W	108	40	139	14,511,146	29,156	497.707
GWR_E	83	31	36	13,764,766	20,410	674.413

- GWR_F 그룹은 사망자 수, 사고비용, 단위거리당 사고비용 모두 가장 적은 값인 반면, GWR_W 그룹은 사고건 수, 사망자 수, 부상자 수, 전체 사고비용 등에서 가장 높은 값을 기록하고 있음.
- 하지만 구간 연장이 GWR_E 그룹보다 50% 가까이 길기 때문에 단위거리당 사고비용은 GWR_E 그룹보다 낮았음.
- GWR_E 그룹은 사고건 수가 가장 적고 부상자 수도 적었으나 단위 비용이 높은 것을 볼 수 있는데, 이것은 구간 내에 포함된 사고들 중에서 전반적으로 사상자가 발생한 사고의 비중이 높은 결과라고 볼 수 있음.
- 다시 말해, 우연적으로 발생하는 경향이 높은 대형사고보다 사상자가 더 발생하지만 평균사상자 수가 적은 사고들의 영향을 많이 받았다고 볼 수 있음.
- 따라서 EPDO를 기준으로 선정된 구간들이 전반적으로 위험도가 높은 구간이며, 단위거리당 사고비용을 감안하면 이들 구간에 대한 안전개선사업 시행 시 경제적

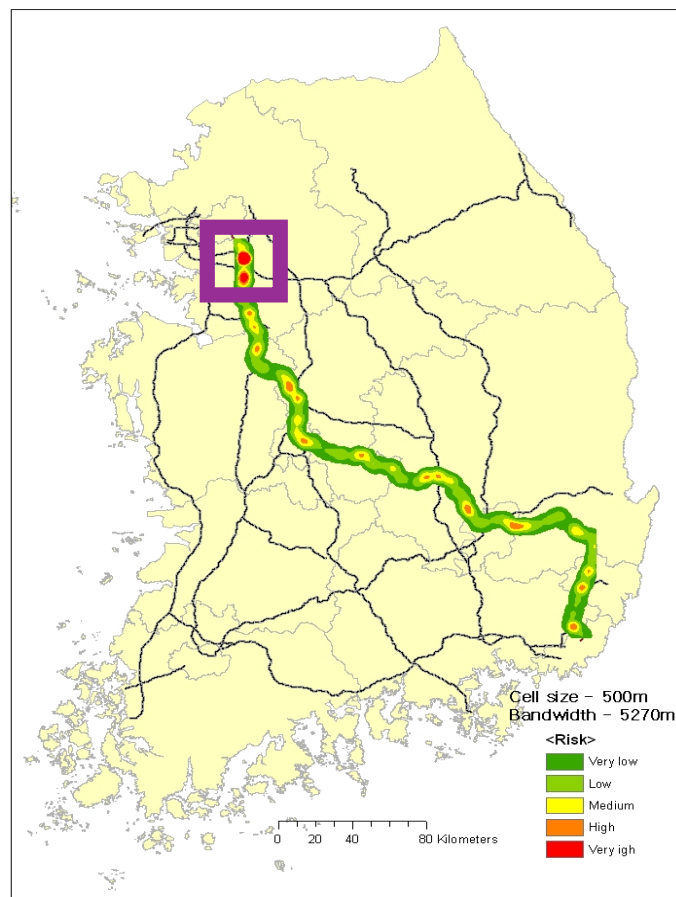
인 효과가 가장 클 것으로 판단됨.

—반면, 심각도를 고려하지 않고 단순히 사고건수로만 선정할 경우에는 도로개선 후에 사고발생 횟수를 줄일 수는 있으나, 교통사고로 인한 사상자 수와 그에 연계하는 사회적 비용이 줄어드는 효과는 상대적으로 적을 것으로 예상됨.

2. KDE 법의 적용

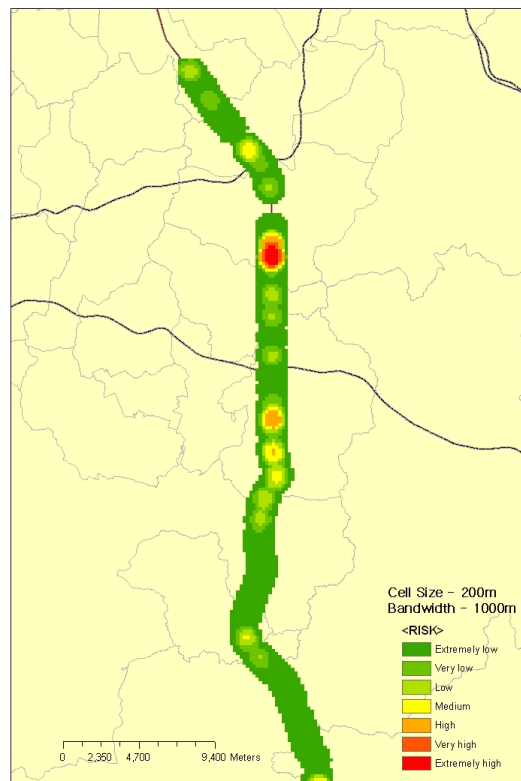
○<그림 3-2>는 심각도를 반영한 커널밀도 추정법을 사용하여 생성된 경부고속도로 구간의 사고밀도 지도를 나타냄.

—bandwidth는 5270m를 사용하였고, cell size는 500m로 부여하였으며, 커널 값을 기준으로 상위 10%에 드는 지역을 사고취약구간으로 표시하였음.



〈그림 3-2〉 전체 노선구간 사고밀도지도(박신형, 2010)

- 다음 단계로 노선 범위에서 판별된 사고취약구간에 대해 지도를 확대하여 상세 분석을 실시하게 되는데, 먼저 <그림 3-2>에서 네모 상자로 표시된 부분을 확대하여 다시 bandwidth를 결정하고 커널밀도를 추정한 결과가 <그림 3-3>에 나타나 있음.
- 이때의 bandwidth는 1000m, cell size는 200m로 설정되었음.



<그림 3-3> 국지적 범위의 사고밀도지도(박신형, 2010)

- Freq 변수와 함께, 심각도를 반영한 변수인 Wfreq와 EPDO를 기반으로 커널밀도 추정을 통해 상위 10%에 해당하는 구간이 사고취약구간으로 선정됨.
- <표 3-4>는 그중 Wfreq를 사용하여 선정된 구간을 나타냄.

〈표 3-4〉 사상자별 가중빈도 상위 10% 구간(KDE_W) (박신행, 2010)

순위	위치정보				사고정보			부문별 구간 합계		
	시작이정	끝 이정	길이(m)	행정 구역	사고	사망	부상	Wfreq	EPDO	사고비용(천원)
1	402.5	403.3	800	성남	13	2	18	91	41	1,069,763
2	391.9	392.4	500	용인	10	0	18	64	12	178,749
3	35.2	36.0	800	울산	5	3	7	62	27	1,323,150
4	367.4	367.6	200	안성	2	0	18	56	4	94,210
5	10.1	10.6	500	양산	3	0	15	48	7	93,819
6	189.5	189.9	400	김천	2	2	7	47	24	885,905
7	299.2	299.6	400	청원	6	2	4	42	21	898,184
8	307.3	307.8	500	청원	3	3	0	39	36	1,341,469
8	119.7	120.2	500	대구	3	1	8	39	16	481,642
10	233.2	233.9	700	영동	4	2	2	34	30	910,410
11	343.2	343.6	400	천안	4	2	2	31	28	916,462

3. GWR과 KDE의 비교

- GWR모형을 통해 선정된 31개 구간과 KDE를 통해 선정된 27개 구간의 위치를 비교하여 공통된 구간을 비교함.
 - KDE를 통해 선정된 KDE_F 그룹에서는 총 12개의 구간 중 10개 구간이 GWR의 31개 구간에 포함되어 있었고, 선정되지 않은 두 구간은 GWR의 유의수준을 5%로 완화했을 때 포함되는 것을 확인할 수 있었음.
 - KDE_W 그룹은 11개 구간 중 10개, KDE_E 그룹은 15개 구간 중 14개 구간이 GWR 선정구간에 포함되어 있었고, KDE_W와 KDE_E 그룹에서 각각 GWR 선정구간에 포함되지 않은 구간은 동일한 구간으로, 역시 5% 유의수준에서는 포함되었음.
 - 전체 31개 GWR 구간 중 KDE에서 공통적으로 선정된 구간 수는 23개이고, KDE를 통해 선정된 27개 구간 중 1% 유의수준에서 GWR 구간에 공통적으로 포함된 구간은 24개였음.
- 이상의 결과를 통해, GWR과 KDE의 결과가 많은 부분 일치하며, 사고건수나 심각도를 기준으로 구간을 선정했을 때 가장 위험도가 높은 것으로 나타나는 구간이 대부분 판별되는 것으로 나타남.

- 포함되는 그룹의 수와 단위거리당 사고비용에 따라 다음 <표 3-5>와 같이 사고취약 구간이 선정되고 우선순위가 부여되었음.

<표 3-5> 사고취약구간 선정 결과 비교(박신행, 2010)

GWR					KDE					포함 그룹 수	우선 순위
시작 이정	끝 이정	길이 (m)	단위사 고비용 (천원)	포함 그룹	시작 이정	끝 이정	길이 (m)	단위사 고비용 (천원)	포함 그룹		
298.7	300.4	1,659	567	E, F, W	299.2	299.6	400	2,245	W7, E14	5	1
389.9	390.7	843	699	E, F, W	389.8	390.4	600	987	F1, E4	5	2
34.9	36.0	1,140	1,151	E, W	35.2	36.0	800	1,654	E8, W3	4	7
188.7	190.1	1,400	652	E, W	189.5	189.9	400	2,215	W6, E11	4	5
307.4	307.9	511	2,625	E, W	307.3	307.8	500	2,683	E3, W8	4	3
343.0	343.7	705	1,300	E, W	343.2	343.6	400	2,291	E7, W11	4	4
391.4	392.4	990	186	W, F	391.9	392.4	500	357	F3, W2	4	8
397.4	407.9	10,432	305	W	402.9	403.3	400	1,500	F2	2	6
					402.5	403.2	700	1,528	W1, E2	3	
268.8	269.7	940	978	E, W	269.1	269.8	700	1,349	E6	3	9
353.9	356.3	2,400	389	E	356.0	356.4	400	1,298	F5, E13	3	10
386.9	388.5	1,590	579	E	386.8	387.3	500	972	E15	2	제외
					388.3	388.8	500	1,813	E8	3	11

* E: EPDO, F: Freq, W: Wfreq, KDE 포함그룹의 숫자는 그룹 내 순위

- GWR은 평면선형에 따라 미리 설정된 구간에 대해 위험도를 평가하므로, 그 길이가 짧게는 수백 미터에서 길게는 10km가 넘을 수 있어, 한 구간 내에 KDE를 통해 선정된 구간이 2개씩 포함되는 경우도 있음.
 - 이는 달리 말하면, 위험도가 높은 구간이더라도 구간이 길기 때문에 실제로 위험하지 않은 지역까지 함께 포함하고 있다고 볼 수 있음.
 - <표 3-5>에서 단위거리당 사고비용이 KDE로 선정한 구간들에서보다 GWR로 선정한 구간들에서 확연히 낮은 것을 보아도 이러한 사실을 확인할 수 있음.
- 반면, KDE에 의해 선정된 구간들은 GWR에서 선정된 구간들보다 짧은 길이의 구

간을 제시하고 있음.

—결과적으로, 현재 고속도로의 사고 잦은 구간에 근거하여 개선사업을 실시하는 대상 구간들과 비슷한 공간적 범위를 판별해내므로 현장에 적용하기에 적합한 방법이라고 할 수 있음.

—또한 위험도가 높지 않은 구간들은 최대한 배제함으로써 불필요한 예산과 자원을 낭비하게 되는 false positive 오류를 줄여주는 장점이 있음.

○이상의 내용을 정리하여 GWR과 KDE를 비교하면 <표 3-6>과 같음.

<표 3-6> GWR과 KDE의 비교(박신행, 2010)

구분	GWR	KDE
종류	모수적 기법	비모수적 기법
구간선정 방식	기존의 통계적 예측모형에서 사용하는 선정기준 활용 (빈도, 심각도의 관측치와 예측치의 상대적 차)	공간상에서의 위험도 분포를 시각적으로 확인하여 구간 선정(빈도, 심각도의 절대적 크기)
위험도 분석 범위	미리 설정된 구간에 대한 위험도 판정	사고지점을 중심으로 한 위험도 판정
위험구간 범위	설정된 구간의 길이에 따라 위험구간 범위 결정	bandwidth에 따라 위험 구간의 범위 조절 가능
심각도 반영	종속변수 설정 (사상자 가중빈도, EPDO)	심각도 지수 반영 (사상자 가중빈도, EPDO)
공간적 영향 고려	인접 구간에 대해 지리적 가중치 적용	사고발생지점에 인접 지역에 대한 지리적 가중치 적용
오류의 형태	false negative 오류 발생가능성을 줄임	false positive 오류 발생가능성을 줄임

IV

결론

- 교통 안전을 위한 정부의 지속적인 노력으로 인해 도로교통 사고건수 및 인명피해가 지속적으로 감소하고 있는 실정이나, 그 수준은 여전히 높아 OECD 회원국 중 하위권을 벗어나지 못하고 있음.
- 최근 급증하는 시민 복지에 대한 관심은 교통복지 측면에서 교통안전 사업에 대한 필요성이 강조되고 있음. 더불어 한정된 예산의 효율적인 집행에 대한 요구 때문에 교통 안전대책 수립 및 도로 안전 개선계획 실행에서 좀 더 합리적이고 과학적인 방법론의 도입이 중요한 과제가 되었음.
- 따라서, 효율적인 사고 취약구간의 선정은 사고 위험이 높은 지점이나 구간을 논리적으로 선정하는 것에만 국한되는 것이 아니라, 비용 효과적인(cost-effective) 정책 결정을 밑받침하고, 행정적 측면에서 민원을 최소화할 수 있는 방안임.
- 이 연구에서는 이를 위한 기초 정책 연구의 일환으로, 국내외에서 연구되고 있는 사고 취약구간 선정을 위한 방법론을 검토하고, 대표적으로 국내 고속도로에 적용된 사례를 소개하였음.
- 이는 서울시 도시 고속도로를 대상으로 가장 알맞은 방법론을 선정하고, 이를 바탕으로 효율적인 사고 취약구간 선정을 위해 실시할 대규모 연구 사업의 예비 조사 단계라고 할 수 있음.
- 기존 국내에서 사용되고 있는 사고 잦은 곳 선정기준은 다음과 같은 문제점을 가지고 있음.
 - 단순 사고건수 기반으로 인명피해 등 사고의 심각도에 따른 영향을 고려하지 못함.
 - 전년도 사고 기록만을 고려함에 따라 Regression-to-Mean 문제, 즉 우연한 요인으로 인하여 사고확률이 특정 연도에 과다 혹은 과소 추정될 수 있음.

- 구간 단위 길이 설정 및 사고건수 기준에 대한 과학적 설명이 부족함.
- 이와 같은 문제점을 해결하기 위해, 이 연구에서는 SPF를 이용한 방법, EB 법, CRP 법, GWR 법, KDE 법 등 최근 소개되고 있는 방법론을 정리하였음.
- 이 중, SPF는 다른 모든 방법론에서 기준이 되는 사고 예측 함수를 개발하는 것으로 서울시 도시고속도로에 과학적인 방법론을 도입하기 위해서는 반드시 대상 구간에 해당하는 SPF를 개발하는 연구가 선행되어야 함.
- CRP 법과 GWR 법은 최근에 개발된 방법으로 실제 정책 결정과정에서 사용된 사례가 아직까지 많지 않음.
 - CRP 법은 미국 California 이외의 지역에서는 실제 정책 결정과정에서 아직까지 적용된 사례가 없음.
 - GWR 법은 시험적인 연구단계로, 박신행(2010)에서도 나타난 바와 같이, 적용 방법에 따라 선정된 구간이 과다하게 추정되는 문제를 보일 수 있음.
- EB 법은 현재 전 세계적으로 사고취약구간 선정에 가장 광범위하게 사용되는 방법으로, 교통안전편람(HSM)에서도 추천되는 방법론임.
- KDE 법은 사고밀도 지도를 통해 인접 구간에 밀집된 사고의 분포를 시각적으로 명확하게 표시해 줄 수 있는 장점이 있음.
 - 사고건수 기반의 KDE 법은 과거에도 많은 적용 사례가 있었으며, 필요한 경우 박신행(2010)의 연구에 나타난 바와 같이 사고 심각도도 쉽게 반영할 수 있음.
 - cell size를 조정할 경우, KDE 법은 현재 고속도로의 사고 잦은 구간에 근거하여 개선사업을 실시하는 대상 구간들과 비슷한 공간적 범위를 판별해내므로 현장에 좀 더 쉽게 적용할 수 있는 방법임.
- 결론적으로, 이 연구에서는 서울시 도시고속도로 구간의 사고취약구간 선정을 위해 우선적으로 구간별로 SPF를 개발할 것을 제안함.
 - 또한, 향후 이를 바탕으로 EB 법 및 KDE 법을 우선 적용하여, 그 효과를 분석하는 것이 바람직할 것임.
 - 이러한 과학적 방법론의 적용을 위해서는 사고 데이터의 종합적이고 효율적인 수집 및 가공 또한 필수적이므로, 이에 대한 표준화 및 전산화 작업이 이루어져야 할 것임.

참고문헌

- 강민욱 · 도철웅 · 손봉수, 2002, “고속도로 평면선형상 사고빈도분포 추정을 통한 음이
항회귀모형 개발: 기하구조요인을 중심으로”, 『대한교통학회지』, 207, 197~204.
- 강승림 · 박창호, 2003, “고속도로 선형조건과 GIS기반 교통사고 위험도지수 분석: 호
남 · 영동 · 중부고속도로를 중심으로”, 『대한교통학회지』, 211, 21~40.
- 강정규 · 이성관, 2002, “고속도로 선형조건별 교통사고 위험도 평가모형개발: 호남고속
도로를 중심으로”, 『대한교통학회지』, 204, 163~175.
- 강현건 · 강승규 · 장용호, 2009, “Empirical Bayes Method를 이용한 교통사고 예측모형”,
『대한교통학회지』, 275, 135~144.
- 김숙희 · 장정아 · 최기주, 2005, “사고다발지점의 안전성능진단 및 위치별 사고요인분석
수원시를 중심으로”, 『대한교통학회지』, 231, 9~20.
- 김용일 · 박신형 · 김동규 · 고승영, 2010, “Continuous Risk Profile 기법을 이용한 사고찾
은구간 선정: 경부고속도로 사례 분석”, 한국ITS학회 춘계학술대회 발표논문.
- 김익희, 2005, “GWR을 이용한 지가설명함수의 공간적 이질성 Spatial Heterogeneity 탐
색-서울시 서초구 · 강남구 · 송파구를 사례로”, 서울대학교 대학원 석사학위논문.
- 박규영 · 김태희 · 김성욱 · 이수범, 2006, “도로안전개선사업의 공간적 범위에 따른 도로
안전시설물의효과도 비교”, 『대한교통학회지』, 244, 31~42.
- 박민호 · 박규영 · 장일준 · 이수범, 2006, “중앙분리대 설치에 따른 사고전환효과 분석”,
『대한교통학회지』, 242, 113~124.
- 박신형, 2010, “고속도로 사고취약구간 판별을 위한 공간분석기법 연구”, 서울대학교 대
학원 박사학위 논문.
- 박효신 · 손봉수 · 김형진, 2007, “고속도로 인터체인지 연결로에서의 교통사고 예측모형
개발”, 『대한교통학회지』, 253, 123~135.

성낙문, 2003, “교통사고예측모델을 이용한 도로의안전도 평가방법 연구”, 교통개발연구원 연구보고서.

유준석, 2008, “CRP(Continuous Risk Profile) 기법을 활용한 고속도로 사고 잦은 지점 분석방법 연구”, 한양대학교 석사학위논문.

정성봉 · 황보희 · 성낙문 · 이선하, 2009, “EB기법을 이용한 사고 잦은 곳 개선사업 우선순위 판정기법 개발”, 『대한교통학회지』, 273, 81~90.

한상진 · 김근정 · 오순미, 2008, “전통적 사고예측모형의 한계 및 개선방안: Hauer 사고예측모형의 소개 및 적용”, 『한국도로학회논문집』, 101, 19~29.

Akaike, H., 1974, *A New Look at the Statistical Model Identification*, IEEE Transactions on Automatic Control 196, 716~723.

Anderson, T. K., 2009, *Kernel Density Estimation and K-means Clustering to Profile Road Accident Hotspots*, Accident Analysis and Prevention 41, 359~364.

Bailey, T. C. & Gatrell, A. C., 1995, *Interactive Spatial Data Analysis*. John Wiley & Sons, New York.

California Department of Transportation, 2002, *Table C Task Force: Summary Report of Task Force's Findings and Recommendations*.

Chung, K. & Ragland, D. R., 2007, *A Method for Generating a Continuous Risk Profile for Highway Collisions*. Presented at 86th Annual Meeting of the Transportation Research Board, Washington D. C.

Chung, K., Madanat, S., Washington, S., and Jang, K., 2010, *Proactive Detection of High Collision Concentration Locations*. Submitted for The 19th International Symposium on Transportation and Traffic Theory ISTTT.

- Chung, K., Ragland, D. R., Madanat, S. and Oh, S. M., 2009, *The Continuous Risk Profile Approach for the Identification of High Collision Concentration Locations on Congested Highways*, Transportation and Traffic Theory 2009: GoldenJubilee, Papers selected for presentation at ISTTT 18, a peer reviewed series since 1959, William H. K. Lam, S. C. Wong and Hong K. Lo eds., Springer, August 2009, 463~480.
- Erdogan, S., Yilmaz, I., Baybura, T. and Gullu, M., 2008, *Geographical Information Systems Aided Traffic Accident Analysis System Case Study: City of Afyonkarahisar*, Accident Analysis and Prevention 40, 174~181.
- Flahaut, B., Mouchart, M., Martin, E. S. and Thomas, I., 2003, *The Local Spatial Autocorrelation and the Kernel Method for Identifying Black Zones: a Comparative Approach*, Accident Analysis and Prevention 35, 991~1004.
- Fotheringham, A. S., Brunsdon, C. and Charlton, M., 1996, *The Geography of Parameter Space: an Investigation into Spatial Non-stationary*, International Journal of Geographic Information Systems 105, 605~627.
- Fotheringham, A. S., Brunsdon, C. and Charlton, M., 2002, *Geographically Weighted Regression- the analysis of spatially varying relationships*, John Wiley & Sons Ltd, Chichester.
- Hauer, E., 2004, Statistical Road Safety Modeling. Transportation Research Record 1897, TRB, National Research Council, Washington, D. C., 81~87.
- Kononov, J. & Allery, B., 2003, *Level of Service of Safety*, Transportation Research Record 1840, TRB, National Research Council, Washington, D. C., 57~66.

- Kwak, H., Kim, D. K., Park, S. H. and Kho, S. Y., 2010, *Development of a Safety Performance Function for Korean Expressways*. Presented at the 12th World Conference on Transport Research, Lisbon, Portugal.
- Park, S., Park, S. H., Kim, D. K. and K. S. Chon., 2010, *Factors that Influence the Level of Accident Severity in Vehicle Crashes: A Case Study of Accidents on Korean Expressways*, Presented at the 12th World Conference on Transport Research, Lisbon, Portugal.
- Pulugurtha, S. S., Krishnakumar, V. K. and Nambisan, S. S., 2007, *New Methods to Identify and Rank High Pedestrian Crash Zones: An Illustration*, Accident Analysis and Prevention 39, 800~811.
- Silverman, B. W., 1986, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, New York.
- Srinivasan, R., Lyon, C., Persaud, B., Martell, C., Baek, J., 2011, *Methods For Identifying High Collision Concentration Locations Hccl For Potential Safety Improvements - Phase Ii: Evaluation Of Alternative Methods For Identifying HCCL*, California Department of Transportation, CFS Number 2078A DRI Final Report.
- Xie, Z. & Yan, J., 2008, *Kernel Density Estimation of Traffic Accidents in a Network Space*, Computers, Environment and Urban Systems 32, 396~406.
- Zhao, F. & Park, N., 2004, *Using Geographically Weighted Regression Models to Estimate Annual Average Daily Traffic*, Transportation Research Record 1879, TRB, National Research Council, Washington, D. C., 99~107.
- Zhong, L. D., Sun, X. D., He, Y. L., Zhong, X. M. and Chen, Y. S., 2009, *Safety Performance Function for Freeway in China*, Presented at 88th Annual Meeting of the Transportation Research Board, Washington D. C.

Working Paper
2011-WP-16

서울시 도시고속도로 사고취약구간 선정에 관한 연구

발 행 인 이 창 현

발 행 일 2012년 5월 15일

발 행 처 서울시정개발연구원

137-071 서울시 서초구 남부순환로 340길 57

전화 (02)2149-1234 팩스 (02)2149-1025

본 출판물의 판권은 서울시정개발연구원에 속합니다.