

서울시 비구조적 침수 저감대책 수립을 위한 복합 포아송 회귀모형 개발*

김학열** · 강일석*** · 허태영****

Estimation of a Compound Poisson Linear Model for Non-Structural Measures to Mitigate Flood-Damaged Areas in Seoul Metropolitan Area*

Hag-Yeol Kim**·Il Suk Kang***·Tae Young Heo****

요약 : 본 연구는 분포적 특수성으로 인하여 모형적합이 어려운 침수면적에 대하여, 이론적 근거를 바탕으로 복합 포아송(Compound Poisson) 회귀모형을 추정하여, 그 모형의 적정성을 분석하고, 침수면적의 감소와 관련 있는 유의미한 변수들에 대하여 정책적 시사점을 도출하고자 한다. 모형 추정을 위해 서울시에 사상 유례없는 침수피해를 일으켰던 지난 2010년 9월과 2011년 7월의 침수자료와 '강우 → 차단 · 침투 · 저류 → 유출'과 연관된 여러 인자들을 도입하였다. 모형 추정 결과, 복합 포아송 회귀모형의 적절성이 통계적으로 인정되었으며, 강우강도 및 토지이용/토지피복에 관련된 변수들이 유의미한 것으로 분석되었다. 독립변수의 단위증가에 따른 침수면적의 한계증가량은 최대강우강도, 시가화율, 상업 업무면적, 주상 혼합면적 변수에서 상대적으로 높게 나타났다. 특히 2011년의 모형 추정 결과에는 토지이용과 연관된 여러 변수의 유의미성이 두드러지게 나타나, 토지 이용강도 및 토지 피복상태의 통제, 개선, 조정과 관리가 침수피해를 저감시키는 데 주요한 역할을 할 수 있음이 확인되었다.

주제어 : 침수 피해면적, 복합 포아송, Tweedie 분포, 수해방지 대책

ABSTRACT : Based on theoretical backgrounds, this study estimates a compound poisson linear regression model for flood damage areas which seldom fit a statistical model due to their unconventional distribution pattern. Then, the appropriateness of the estimated model is analyzed and the implications for its significant variables are addressed. For the model estimation, following two datasets are introduced: (1) the damage area data unprecedentedly occurred on September, 2010 as well as July, 2011 at Seoul Metropolitan Area, (2) various influential variables related to the process of rainfall, interception, infiltration, detention, and runoff. The results show that this compound poisson model is assessed to be statistically appropriate for the flood-damaged area data. The variables related to rainfall intensity and land cover turn out to be statistically significant. Marginal effect for flood area by the increase of an independent variable is revealed as relatively high in independent variables

* 본 연구는 국토해양부 첨단도시개발사업의 연구비지원(11첨단도시G09)에 의해 수행되었습니다.

** 서경대학교 도시공학과 부교수(Associate Professor, Department of Urban Planning and Engineering)

*** 충북대학교 통계학과 석사과정(Master's Student, Department of Information Statistics, Chungbuk National University)

**** 충북대학교 통계학과 부교수(Associate Professor, Department of Information Statistics, Chungbuk National University), 교신저자(E-mail: theo@chungbuk.ac.kr, Tel: 043-261-3741)

of maximum rainfall intensity, urbanized area ratio, commercial area, and residential-commercial mixed area. Especially, the estimated model for year 2011 shows that the coefficients of the land use-related variables are highly significant. This result implies that the control, improvement, and management of land use and land cover can play important roles in mitigation of flood damage.

Key Words : flood-damaged area, compound poisson model, Tweedie distribution, flood mitigation policy

I. 서론

2010년 9월 20일 시간당 100mm가 넘는 기록적인 폭우가 쏟아지며 화곡동, 신월동 등의 상습침수지역들이 피해를 입었을 뿐만 아니라, 서울의 중심이라고 할 수 있는 세종로까지도 침수되었다. 더욱이 2011년 7월 27일에는 기록적인 폭우를 쏟아내면서 그 전날부터 내린 비의 양이 300mm가 넘어, 주택, 도로를 침수시키고, 강남 여러 곳의 물난리와 처참한 우면산 산사태의 원인이 되었다.

서울은 오랜 세월동안 수해에 대비하여 하수시설, 배수시설 등의 구조적인 대책을 추진해오면서 국내에서는 침수로부터 상대적으로 안전한 도시로 알려져 있다. 그러나 기후변화에 따른 유례없는 집중호우와 도시화로 인한 침투·저류 기능의 상실로 말미암아, 더 이상 기존의 구조적인 대책만으로 이러한 문제를 해결할 수 없음이 상기의 2010년, 2011년의 수해로 드러나게 되었다.

이에 기존의 구조적인 대책과 병행하여 도시계획 등 비구조적인 대책을 활용한 저감대책에 대한 관심이 높아지고 있다. 더구나 도시계획 및 설계에 관한 대책은 시민의 실생활과 재산권에 직접적으로 영향을 미치기 때문에 그 대책의 적용가능성, 영향범위 등이 사전에 분석될 필요가 있다.

그러나 이러한 관심도와 필요성에 비하여 비구조적인 대책의 효과성에 대한 실증적인 분석은 상

당히 부족한 게 사실이다. 그간의 침수관련 연구는 복잡한 수리수문학적인 분석과 SWMM 같은 소프트웨어의 시뮬레이션에 집중됨으로써, 선행연구의 결과에 기초하여 비구조적인 대책을 마련하기가 쉽지 않은 상황이다.

따라서 본 연구에서는 도시계획과 연관된 다양한 인자를 도입하여 서울시 침수면적에 적합한 통계모형을 구축하고자 한다. 특히, 실증분석을 위해 2010년과 2011년에 발생한 서울의 대규모 침수사태에 영향을 미쳤으리라 예상되는 토지 이용특성 인자를 중심으로 회귀모형을 추정하고자 한다. 또한 그 결과를 통해 침수 저감을 위한 도시계획적 수단을 도출하고, 그 효과를 분석하고자 한다. 본 연구는 다음과 같은 과정으로 진행된다. 첫째, 통계 모형을 중심으로 침수와 관련된 선행연구를 검토한다. 둘째, 침수면적에 가장 적합한 회귀모형을 도출한다. 셋째, 연도별 침수면적, 토지이용 및 토지피복 인자, 지형인자, 입지인자, 강우인자 등을 이용하여 통계 모형을 추정한다. 마지막으로, 모형 추정 결과를 이용하여 정책적 시사점을 제시한다.

II. 침수면적 통계 모형 정립

1. 강우에 대한 선행연구

실제 자료를 통한 복합 포아송 모형의 활용성은 보험분야와 기상분야에서 많은 연구자들에 의해 연구되었다. 특히 기상분야의 경우 비가 오는 경우와 오지 않는 경우에 대한 이산형 변수와 비가 올 경우 비가 내린 양에 대한 연속형 변수인 강우량에 대해 모두 모형 구축이 가능한 복합 포아송 모형을 많이 활용하고 있다. 강우량은 비가 오지 않는 경우 0의 값을 가지며, 비가 오는 경우에는 연속형 자료의 형태를 가지는 특성을 갖고 있어 강우에 대한 모형을 구축하기 위해서는 0을 포함하고 있으면서 연속형 자료를 가지는 복합적인 형태의 모형이 필요하게 되었으며, 이러한 특성을 가지는 강우자료에 대한 복합 포아송 모형에 대한 연구는 현재도 진행중에 있다(Revfeim, 1984; Thompson, 1984, Dunn, 2004; Garavaglia et al., 2011). 따라서 강우량과 비슷한 형태의 자료 특성을 가지고 있는 침수면적에 대해 복합 포아송 모형을 적용하여 그 활용성을 분석하는 것은 학술적·현실적 의미가 높다고 판단된다.

2. 토지피복과 유출 및 침수면적의 관계

토지이용·피복과 유출에 관한 연구는 총우량과 유효우량에 관한 전통적인 수문학이론에 근거하여 현재까지 많은 연구가 진행되어 오고 있다. 즉 직접유출량(유효우량)은 유역의 잠재 보유수량에 의해 결정되며, 이 잠재 보유수량은 토지이용 및 토지식생피복 상태와 관련이 높은 유출곡선지수(CN)에 따라 산출가능하기 때문에(윤용남, 2003), 토지이용·피복 상태와 유출과의 관계가

이 분야의 주요한 관심분야이다.

토지피복, 표면재료 등과 우수유출의 관계를 분석한 연구는 토지피복 등의 변화와 함께, 실제 혹은 인위적 강우강도를 다양하게 적용하면서 실제 유출량을 측정하거나 컴퓨터시뮬레이션을 통하여 그 상관관계를 분석한 연구들이 대부분을 차지하고 있다(임장혁 외, 2007; 김성준 외, 2005; 최지연 외, 2012, 박민지 외, 2005). 이에 관한 거의 모든 연구들은 도시적 토지이용과 인공 포장면적이 증가될수록, 강우강도가 증가될수록 우수유출이 증가하는 것으로 결론을 내리고 있다.

한편 토지이용·피복과 침수면적과의 연관성을 분석한 연구로는, 강우-유출-침수 과정에서 발생하는 복잡한 메커니즘과 이와 관련된 다양한 영향인자들을 동시에 고려해야 하는 이유로 인하여 다른 연구주제보다는 그 연구사례가 극히 제한적이다. 단지 불투수면적 변화의 다양한 시나리오를 구성하여 침수지역을 모사한 연구사례가 몇몇 존재하나(박창열·신상영, 2014), 실제 침수면적을 이용하여 토지이용과의 관련성을 모형화한 연구 사례는 서울시정개발연구원(2011)의 연구 외 다른 국내 사례를 찾기가 힘들다. 특히 서울시정개발연구원(2011)의 연구는 서울시 2010년 9월, 2011년 7월의 침수면적자료에 대해 강우특성, 지형, 토지피복, 토지용도 등의 자료를 연관시켜 OLS 모형, 토빗(Tobit) 모형, 로지스틱 회귀모형 등을 추정함으로써, 토지이용과 침수면적의 관계를 설명하였다. 비록 이 연구에서 개발한 분석모형의 유의미성은 통계적으로 인정되었으나, 모형의 적합도가 낮게 나타나 설명모델로 활용하는 데 한계가 있음이 나타났다.

따라서 본 연구에서는 침수면적자료의 고유한 특성을 파악하여 이에 적합한 모형을 도출하고 여

기에 토지이용·피복에 관한 변수를 적극 도입하여 모형을 추정함으로써 비구조적인 대책에 시사점을 제공할 수 있는 설명력 높은 통계 모형을 제시하고자 한다.

3. 침수면적에 적합한 통계 모형

1) Tweedie 분포

변수들 간의 관계를 알아보는 통계적 방법으로 일반선형모형(generalized linear model, GLM)을 비롯하여 여러 모형이 널리 사용되고 있다. 예를 들어, 기후 분야에서 비가 내리는 사건의 횟수를 종속변수로 하는 경우에는 연결함수(link function)를 로그(log)로 취하는 포아송 선형회귀 모형을 기반으로 한다. 하지만 사건의 횟수가 아닌 각 사건에 따른 어떠한 양의 실수값을 종속변수로 하는 경우에는 포아송 회귀모형을 이용하기에는 적절하지가 않다. 이처럼, 각 분야별 변수가 가지는 특별한 성질로 인해 모형의 적합도가 현저하게 낮아지거나 모형을 적합 시키지 못하는 경우를 종종 겪는다(Smyth, 1989).

종속변수가 양의 값을 가지는 실수이고 셀 수 있는 사건이 발생함에 따라 종속변수가 얻어지는 경우에 복합 포아송 분포(compound poisson distribution)를 따르는데, 위의 경우를 예를 들면, 침수면적은 호우로 인한 침수가 발생한 경우에만 얻을 수 있는 변수값이며 항상 0 이상의 실수값을 가지기 때문에 복합 포아송 분포를 이용한 모형을 사용하는 것이 적절할 것이다. 실제로 순보험료, 강우량, 교통 사고에 따른 피해액 등을 종속변수로 가지는 경우에 복합 포아송 분포를 이용한 모형을 실제 응용연구에서 많이 사용하고 있다

(Jørgensen and Souza, 1994; Smyth and Jørgensen, 2002).

먼저 복합 포아송모형을 포함하고 있는 지수 산포모형(exponential dispersion model)에 대해 간략한 설명을 하면 다음과 같다. 지수 산포 모형은 Tweedie(1984)에 처음으로 제안되어 Tweedie 모형으로 불리기도 한다.

Jørgensen(1987)에 의해 연구된 지수 산포를 이용한 모형은 여러 가지 통계적 분석에 많이 쓰이는 모형으로 일반적인 지수 산포 분포의 확률밀도함수는 식 (1)과 같다.

$$f(y|\theta, \phi) = a(y, \phi) \exp\left(\frac{y\theta - \kappa(\theta)}{\phi}\right) \quad (1)$$

여기서 a 는 상수함수이며 κ 는 누율생성함수(cumulant generating function)이며, θ, ϕ 는 모수이다. 특히, ϕ 는 0 이상의 값을 가지는 산포 모수(dispersion parameter)이다(Jørgensen, 1987).

Tweedie 분포는 누율생성함수의 고유한 성질을 이용하여 기댓값과 분산인 $E(y) = \kappa'(\theta)$ 와 $Var(y) = \phi \kappa''(\theta)$ 를 쉽게 구할 수 있으며, θ 와 μ 는 일대일 함수임을 통하여 $\kappa''(\theta)$ 를 μ 의 함수 $V(\mu)$ 로 표현할 수 있고 이를 멱분산함수(power variance function)라고 명명한다(Kaas, 2005; Barndorff-Nielsen 1978).

멱분산함수의 일반적인 형태는 $V(\mu) = \mu^p$ 로 나타내며, 이때 분산함수의 p 는 지시모수(index parameter)이다. 지수 산포 모형에서 멱분산함수는 중요한 성질을 가지는데, 멱분산함수의 형태만 결정된다면 지수 산포 모형의 확률분포를 알 수 있다(Jørgensen, 1997).

Tweedie 분포는 지시모수의 형태에 따라 분포

가 달라지는 특성을 가지고 있는 장점을 가지고 있다. 지시모수가 $0 < p < 1$ 인 경우에는 확률분포가 존재하지 않으며, $p = 0$ 인 경우에는 정규분포를 따르고, $p = 1$ 일 때는 포아송 분포, $p = 2$ 일 때는 감마분포를 따른다고 알려져 있다. 특히, $1 < p < 2$ 일 때는 복합 포아송 분포를 따르며, 복합 포아송 분포는 0에서 양의 확률을(확률질량함수) 가지는 이산형 변수와 0 이상의 양의 실수 값에서는 확률밀도함수를 가지는 연속형 변수가 결합된 분포로 알려져 있다.

그러나 지시모수 p 를 추정하는 데 있어서 복합 포아송 분포의 확률밀도함수의 특성상 계산의 어려움으로 인해 다양한 모수 추정 방법들이 제안되었다. Cox and Reid(1987)은 프로파일가능도(profile likelihood)방법을 제안하여 p 를 추정하였으며, Nelder and Pregibon(1987)은 확장된 준 우도(extended quasi-likelihood) 방법을 기반으로 한 방법을 제안하였다. 최근 연구에서는 컴퓨팅의 발전을 통해 보다 정확한 복합 포아송 분포를 구할 수 있다(Dunn and Smyth, 2005, 2008).

복합 포아송 모형을 적용하기 위해 본 연구의 관심변수인 침수면적에 대해 모형식을 적용하면 다음과 같다. 주어진 기간에 대해서 확률변수 N 을 비가 온 사건의 횟수로 정의한다. 또한 X_i 를 i 번째 비가 내렸을 때의 침수면적으로 정의하며, 이러한 확률변수들의 가정은 식 (2)와 같다.

$$N \sim \text{Pois}(\lambda), X_i \stackrel{iid}{\sim} \Gamma(\alpha, \gamma), N \perp X_i \quad (2)$$

즉, 확률변수 N 은 평균이 λ 인 포아송 분포를 따르며, 확률변수 X_i 는 감마분포를 따르고, N, X_1, X_2, \dots, X_N 은 서로 독립임을 가정한다. 이제

확률변수 Z 를 주어진 기간에 대한 누적 침수면적으로 정의하면 식 (3)과 같다.

$$Z = \sum_{i=1}^N X_i \quad (3)$$

이때 확률변수 Z 를 평균이 λ 인 복합 포아송 분포를 따른다고 한다. 즉, 누적 침수면적 Z 는 X_i 들의 임의의 누적합(random sum)으로 정의되며, 이때 $N = 0$ 인 경우에는 Z 의 값도 0으로 정의한다. 식 (4)의 관계식을 통해 누적 침수면적 Z 의 평균 $\mu = \lambda\alpha\gamma$, 분산 $\phi\mu^p = \lambda\alpha(\alpha+1)\gamma^2$ 을 얻을 수 있다.

$$\begin{aligned} \lambda &= \frac{\mu^{2-p}}{\phi(2-p)} \\ \alpha &= \frac{2-p}{p-1} \\ \gamma &= \phi(p-1)\mu^{p-1} \end{aligned} \quad (4)$$

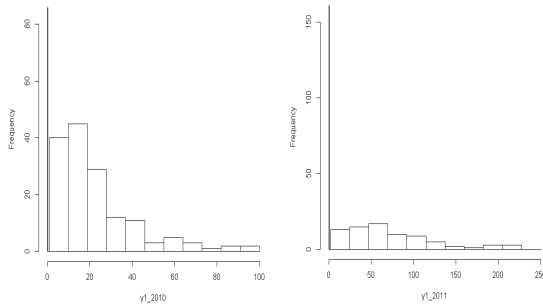
복합 포아송 분포를 따르는 확률변수 Z 는 Tweedie 분포에서 언급한 바와 같이 이산형과 연속형이 혼합된 분포이므로 확률밀도함수를 식 (5)와 같이 표현한다.

$$\begin{aligned} f(y, n|\lambda, \alpha, \gamma) &= f(y|n, \alpha, \gamma)f(n|\lambda) \\ &= \begin{cases} e^{-\lambda}, & y=0 \\ \frac{e^{-y/\gamma}y^{\alpha-1}}{\Gamma(n\alpha)\gamma^{n\alpha}} \frac{\lambda^n e^{-\lambda}}{n!}, & y>0 \end{cases} \end{aligned} \quad (5)$$

2) 침수면적의 복합 포아송 분포 적용

앞절에서 침수면적과 같이 어떠한 사건이 발생하지 않을 시의 0의 값을 갖고, 사건이 발생함에 따라 얻어지는 양의 실수값에 대한 확률변수를 기존의 다른 통계적 모형들을 이용하여 분석하기에는 적절하지 않음을 언급하였다. 따라서 본 연구

에서 활용될 복합 포아송 모형의 타당성을 확인하기 위해 종속변수의 히스토그램을 통해 침수면적의 분포 적용의 적정성 및 적용성을 확인하였다(김명준·김영화, 2009).



〈그림 1〉 2010년, 2011년 침수면적의 히스토그램

〈그림 1〉은 본 연구에서 종속변수로 사용될 2010년, 2011년 발생한 침수면적을 히스토그램으로 나타낸 것이다. 특정 배수분구를 제외하고는 상당수의 배수분구에서 침수지역이 발생되었는데, 서울시 전체 239개의 배수분구 중 2010년에는 86개(약 36%), 2011년에는 160개(약 67%)의 배수분구에서 침수가 발생된 것으로 나타났다. 우측으로 긴 꼬리를 갖는 분포의 형태를 갖는 이러한 분포는 일반적으로 감마분포 또는 로그정규분포를 적용하기도 한다. 침수면적에 관한 변수는 감마분포와 같은 연속형 변수를 적용하여 분석할 수 있지만, 침수가 발생되지 않았을 경우의 0으로 처리되는 값을 설명하기에는 적절하지 않다. 하지만 Tweedie 분포에서 언급한대로 지수모수 p 가 $p \in (1, 2)$ 일 때 얻을 수 있는 복합 포아송 분포를 적용하면 침수면적이 0인 경우의 확률을 부여할 수 있는 이산형 변수의 특징과, 침수면적의 값이 0이 아닌 양의 실수 값을 가질 때의 연속형 변수의 특징을 모두 설명할 수 있다.

III. 분석자료 및 도입변수

본 분석에서 사용되는 데이터는 최근 서울시에서 침수피해가 발생되었던 2010년 9월과 2011년 7월의 강우사상과 이와 연관된 토지이용 특성을 중심으로 하여, 서울의 239개 배수분구를 공간단위로 구축되었다. 비록 이 공간단위가 서울시 행정동 단위 혹은 센서스 집계구 단위보다 면적이 크나, 빗물유출 및 배수에 관한 자연조건과 배수관리체계를 반영한 최소의 공간단위이므로, 그 면적의 크기에 불구하고 최선의 대안으로 사용되었다.

종속변수로 활용될 침수면적 자료는 2010년과 2011년의 서울시 침수흔적도로부터 배수분구별 침수면적을 추출한 자료이다. 2010년 9월 21일과 2011년 7월 27일에 최대강우강도로 인하여 침수가 최대로 발생되었으므로, 침수흔적도는 이때 발생한 침수지역을 나타낸다. 이들 자료는 서울시 하천관리과에서 보유하고 있는 자료로 해당 시기에 적합한 자료만을 추출하여 사용하였다.

독립변수로는 각 배수분구에 대하여 강우사상에 관한 변수, 지형 및 입지에 관한 변수, 토지이용 및 토지피복에 관한 변수, 배수시설 및 기타 변수들로 구성된다. 이 변수들은 ‘강우 → 차단·침투·저류 → 유출’에 이어지는 일련의 과정과 연관된 변수들이고, 특히 토지이용과 토지피복에 상대적인 주안점을 두고 변수가 선정 추출되었다. 〈표 1〉에서 제시된 바와 같이, 지형 및 입지에 관련된 변수로는 경사도가 2% 이하 지역의 면적, 지형습윤지수(TWI: topographic wetness index)의 값이 10 이상인 면적, 하천계획홍수위 이하의 면적, 하천계획홍수위에 2m를 더한 표고보다 낮은 지역의 면적, 하상고보다 낮은 지역의 면적, 하상고에 2m를 더한 표고보다 낮은 지역의 면적 등이

〈표 1〉 모형도입 변수 및 예상 결과

구분	변수명	변수설명(단위)	출처	예상부호
침수면적	Y_1	2010년 침수면적(m^2)	2010 서울시침수흔적도	종속변수
	Y_2	2011년 침수면적(m^2)	2011 서울시침수흔적도	종속변수
강우사상	X_1	최대강우강도(mm/h)	서울시AWS 자료	+
	X_2	최대일강우량(mm)	서울시AWS 자료	+
지형 및 입지	X_3	경사도 2%이하 면적(m^2)	2008 서울시수치지형도	+
	X_4	지형습윤지수(TWI) 10이상 면적(m^2)	2010 서울시도시생태현황도	+
	X_5	하천계획홍수위 이하 면적(m^2)	서울시 하천정비기본계획	+
	X_6	하천계획홍수위+2m 이하 면적(m^2)	서울시 하천정비기본계획	+
	X_7	하상고 이하 면적(m^2)	서울시 하천정비기본계획	+
	X_8	하상고+2m 이하 면적(m^2)	서울시 하천정비기본계획	+
토지이용 및 토지피복	X_9	불투수율이 90%이상인 면적비율(%)	2010 서울시도시생태현황도	+
	X_{10}	단독주택 면적(m^2)	2010 서울시도시생태현황도	-
	X_{11}	공동주택 면적(m^2)	2010 서울시도시생태현황도	+
	X_{12}	주상혼합 면적(m^2)	2010 서울시도시생태현황도	+
	X_{13}	상업업무 면적(m^2)	2010 서울시도시생태현황도	+
	X_{14}	교통시설지 면적(m^2)	2010 서울시도시생태현황도	+
	X_{15}	녹지 및 오픈스페이스 면적(m^2)	2010 서울시도시생태현황도	-
	X_{16}	하천호소 면적(m^2)	2010 서울시도시생태현황도	+
	X_{17}	시가화율, 주거, 상업, 공업용도 면적비율(%)	2010 토지이용현황도	+
	X_{18}	건폐면적률(%)	2008 서울시수치지형도	+
	X_{19}	평균 CN	2010 서울시도시생태현황도	+
배수시설	X_{20}	토지이용다양도지수	2010 서울시도시생태현황도	-
	X_{21}	시가화면적당 하수관거연장(m/ m^2)	2009 서울시 하수도정비기본계획	-
	X_{22}	하수관거연장 대비 통수능부족관거연장(%)	2009 서울시 하수도정비기본계획	+

선정되었다. 여기서 TWI는 상류지역의 유출기여 면적을 해당지점의 경사로 나눈 값으로, 강우유출 현상과 지형특성의 관계를 보여주는 척도이며, 유역의 특정지점에서 상류로부터 내려오는 흐름의 특성과 하류로 통수시킬 수 있는 유하능력을 비교하여 포화되는 정도 또는 침수가능성을 나타내는 지수이다. 일반적으로 TWI가 10을 초과하면 국지적인 우수의 집중이 발생하기 쉬운 지역으로 평가된다. 이 변수들은 모두 양의 상관관계가 나타나리라 예상되었다.

토지이용 및 토지피복에 관련된 변수로는 각종

토지 이용·피복 변수, 불투수율, 평균 CN값, 토지이용 다양도 지수 등이 선정되었고, 이들은 강우유출 특성과 높은 상관관계가 있을 것으로 예상되었다. 여기서 토지이용 다양도는 각 배수분구에서 용도별 면적비율의 제곱합을 1에서 차감한 심슨지수(Simpson Index)를 이용하여 산출하였다. 토지용도별 면적이 고루 분포하면 지수 값이 높고, 특정용도가 우세하면 그 반대로 나타난다. 이 변수들 중 단독주택면적, 녹지 및 오픈스페이스 면적, 토지이용 다양도지수가 음의 상관관계가 있고, 나머지 변수들은 양의 상관관계가 있을 것으

로 예상되었다. 또한 토지이용관련 각 변수의 계수는 건폐율, 용적률과 같은 토지이용강도와 그에 따른 포장과 같은 토지피복특성을 반영하여 모형 추정 결과가 나타날 것으로 예상되었다.

IV. 분석 및 결과

앞서 추출된 데이터를 복합 포아송 분포를 이용한 모형에 적합시켜 2개년에 대한 각각의 침수면적에 어떠한 변수가 유의하게 영향을 미치는지 알아보고자 한다. 이에 우선 <표 1>에서 제시된 22개의 모든 변수를 모델에 도입하여 1차적으로 모형을 추정하고, 그다음 95% 유의수준에서 유의한 변수만으로 2차 추정을 한 후, 그 결과를 제시한다.

분석에 앞서, 두 종속변수 Y_1 , Y_2 의 단위 관계상 제곱근 변환을 취한다. 이는 복합 포아송 모형에서 ML(maximum likelihood) 추정방법을 기반으로 모수를 추정하므로 변수의 값이 큰 경우 모수추정을 위한 계산 수렴과정에서 종종 수렴되지 않는 문제점을 보완하기 위해 단위를 줄이는 적절한 변수변환이 필요하기 때문이다. 아울러, 복합 포아송 모형을 적용하기 위해 필요한 조건인 지시모수 $p \in (1, 2)$ 를 만족해야 하므로 모형의 결과 해석에 앞서 p 의 추정값 확인이 반드시 선행되어야 한다.

본 연구에서는 복합 포아송 모형의 적용을 위해 통계패키지 R을 사용하여 총 침수면적에 영향을 주는 영향변수들의 계수 추정값과 복합 포아송 활용 여부를 판단하기 위한 모수 p 와 ϕ 에 대한 추정값을 얻을 수 있었다. 종속변수 Y_1 (2010년 침수면적)에 대한 모형에서의 모수 추정값은 $\hat{p} = 1.2786$, $\hat{\phi} = 7.0095$ 이며, Y_2 (2011년 침수면적)에 대한 모수 추정값은 각각 $\hat{p} = 1.3521$,

$\hat{\phi} = 24.776$ 로 추정되어 침수면적에 대한 영향요인을 확인하기 위한 복합 포아송 모형의 사용은 적합하다고 평가되었다.

2010년 침수면적을 종속변수로 하여 모든 변수들을 복합 포아송 모형에 적합하여 최종적으로 얻어진 유의한 변수에 대한 결과는 <표 2>에 제시되어 있다.

<표 2> 2010년 침수면적 모형 추정 결과

변수	추정값	표준오차	P-값
절편	0.3048	0.6472	0.6381
X_1	0.0446	0.0053	<0.0001
X_{12}	0.0061	0.0009	<0.0001
X_{19}	-0.0179	0.0073	0.0146

최종 결과에 대한 유의한 변수의 회귀계수로 살펴보면, 2010년의 침수면적은 최대강우강도(X_1)가 1단위 증가 시 평균적으로 4.46%($e^{0.0446}$: 1.0446배) 증가하며, 주상혼합면적(X_{12})의 경우 면적이 1단위 증가 시 평균적으로 0.61%($e^{0.0061}$: 1.0061배) 증가하는 것으로 분석되었다. 한편, 최대 강우강도는 초기 예상과 같이 다른 변수들보다 침수면적 증가에 높은 영향력을 갖고 있는 것으로 나타났다.

2011년의 침수면적에 관한 모형 추정은 2010년과 마찬가지로 초기 추정 이후에 최종적으로 유의하게 영향을 미치는 변수들에 대한 재추정 작업을 통하여 최종결과를 얻었다. <표 3>은 그 결과를 나타낸 것이며, 이에 대한 결과 해석은 2010년의 침수면적의 해석과 유사하다.

회귀계수를 살펴보면, 환경사면적(X_3), 주상혼합면적(X_{12}), 상업업무면적(X_{13}), 하천호소면적(X_{16}), 시가화율(X_{17}) 등이 유의한 변수로 나타났다. 각 변수의 증가에 따른 침수면적의 한계

증가는 각각 0.0001%, 0.736%, 1.06%, 0.614%, 3.56%로 나타났다. 즉 시가화율, 상업업무면적의 단위 증가에 대한 침수영향력이 타 변수에 비해 상대적으로 큰 것으로 나타나 이러한 변수의 통제가 침수저감 정책으로 유용하게 활용될 수 있음을 보여주었다. 또한 주상혼합면적의 영향력은 2010년과 유사한 것으로 나타나 그 변수의 관리의 효과가 안정적일 수 있음이 확인되었다.

그러나 TWI(X_4)과 불투수율(X_9)은 당초 예상과는 달리 부(-)의 부호가 나타났다. 특히 도시유출의 증가에 있어서 중요한 인자로 알려져 있는 불투수율의 계수가 정(+)의 부호로 나타나지 않은 것은 여러 가지 의문을 남긴다. 이러한 결과는 불투수율이 높은 지역에서 유출이 발생하였으나 경사나 배수시설 등의 여러 영향으로 실제 침수는 다른 지역에서 나타났기 때문에 발생한 결과로 추측된다.

한편 최대강우강도(X_1)에 따른 2011년의 침수면적의 한계증가량은 평균적으로 4.17% 정도인 것으로 분석되었는데, 이는 2010년의 결과와 상당히 유사한 결과로써, 서울과 같은 대도시지역에서의 강우강도는 침수면적의 증감에 있어서 매우 중요한 영향인자임을 보여주는 결과이다.

〈표 3〉 2011년 침수면적 모형 추정 결과

변수	추정값	표준오차	P-값
절편	-1.63E+00	8.76E-01	0.0643
X_1	4.71E-02	8.70E-03	<0.0001
X_3	1.14E-06	5.11E-07	0.0272
X_4	-1.61E-02	5.94E-03	0.0071
X_9	-3.51E-02	1.15E-02	0.0025
X_{12}	7.36E-03	2.08E-03	0.0005
X_{13}	1.06E-02	3.81E-03	0.0059
X_{16}	6.14E-03	2.47E-03	0.0138
X_{17}	3.56E-02	9.90E-03	0.0004

2010년 및 2011년의 침수면적에 둘 다 유의하게 영향을 주는 변수들을 살펴보면, 최대강우강도, 주상혼합면적의 두 변수가 유의한 것으로 확인된다. 또한 두 변수 모두 추정된 계수의 부호나 그 크기가 상당히 안정적인 것으로 나타났다. 이는 강우강도가 높을수록, 주상혼합면적이 상대적으로 넓을수록 침수면적이 증가한다는 것을 의미한다. 이에 서울시의 토지이용관리나 도시개발에 있어서 주상혼합지의 침투 및 저류기능의 보완·확대를 위한 다양한 정책마련이 필요할 것으로 판단된다. 즉, 주상혼합지의 개발에 대한 도시계획심의·결정에 있어서 기존의 침투·저류시설의 법적인 내용을 강화하거나, 현재 서울시에서 운용중인 ‘빗물관리시설 설치’ 및 ‘자연지반보존’에 관한 친환경인센티브 항목의 계수를 적극적으로 상향하여 빗물유출이 저감되고 지역의 물순환과 환경성이 증진될 수 있도록 유도할 필요가 있다.

V. 결론

본 연구는 분포적 특수성으로 인하여 모형 적합이 어려운 침수면적에 대하여, 이론적 근거를 바탕으로 복합 포아송 회귀모형을 추정하여, 그 모형의 적정성을 분석하고, 침수면적의 감소와 관련있는 유의미한 변수들에 대하여 정책적 시사점을 도출하고자 하였다.

‘강우 → 차단·침투·저류 → 유출’에 이어지는 일련의 과정과 연관된 서울시 배수분구 내 특성인자들을 도입하여, 2010년, 2011년의 각각의 침수면적에 대하여 복합 포아송 모형의 적절성을 확인하기 위해 지시모수 p 추정값의 타당성을 확인한 후, 복합 포아송 회귀모형을 통해 각 변수의 회귀계수를 추정하였다. 그 결과, 2010년 모형에

서는 최대강우강도, 주상혼합면적이 유의미한 변수로 나타났고, 2011년 모형에서는 환경사면적, 주상혼합면적, 상업업무면적, 시가화율 등이 유의미한 변수로 추정되었다. 특히 2011년에는 토지이용과 연관된 여러 변수의 유의미성이 두드러지게 나타나, 토지이용강도 및 토지피복상태가 침수피해를 저감시키는 데 주요한 역할을 할 수 있음이 확인되었다. 특히 주상혼합지는 두 개 연도 모두에서 유의미한 변수로 분석되어 이에 대한 토지이용관리정책과 침투·저류시설 강화 및 자연지반보존·확대에 대한 세부적인 정책마련과 도시계획심의기준이 필요할 것으로 분석되었다. 따라서 침수저감을 위한 수방대책으로는, 침투기능을 담당하는 녹지지역을 보전하고, 상업용도/주상복합용도 지역의 이용강도와 토지피복을 침투저류능이 향상되도록 유도하는 도시계획적 대책이 필요할 것으로 판단된다. 비록 서울과 같은 고밀도의 높은 지가의 지역에서는 이러한 도시계획적 대책을 적용하기가 쉽지 않을지라도, 이러한 대책의 점진적이고 지속적인 추진은 구조적인 침수방지 대책을 보완하고 기후변화에 대응하는 슬기롭고 친환경적인 도시 관리 대책임을 주지할 필요가 있다.

본 연구가 침수면적 분포의 특수성을 고려하고, 강우유출과정에 연관된 다양한 변수들을 도입한 회귀모형을 구축하였으나, 다음과 같은 향후 연구가 진행된다면 보다 적합도가 높은 모형개발이 이루어질 것으로 판단된다. 첫째 침수지역별로 지역 특성이 매우 다양하여 모델적합성이 낮아질 수 있으므로, 개별침수지역들을 유형화하고 그 지역 내·외의 세부인자를 도입하여 분석할 수 있는 정교한 모형 개발이 필요하다. 둘째 강우유출지역과 침수발생지역이 상이함에 따른 문제를 해결할 수

있는 모형을 개발해야 한다. 이에 침수발생 메커니즘을 보다 정교하게 반영할 수 있는 통계모형구조와 도입변수를 연구해야 한다. 셋째, 침수면적의 분포에 적합한 모형을 탐색·개발할 필요가 있다. 2011년 침수면적의 경우, 침수면적이 0값을 가지는 비율이 67%로써 영과잉 복합 포아송 모형(Zero-Inflated Compound Poisson Model)을 고려해 볼 수 있으며, 모수 추정을 위해 베이지안 추정기법을 활용한 복합 포아송 모형의 연구도 필요할 것으로 판단된다(Zhang, 2012).

참고문헌

- 김명준·김영화, 2009, “다양한 모형화를 통한 자동차 보험가격 산출”, 『한국데이터정보과학회지』, 20(3): 515~526.
- 김성준·박근애·전무갑, 2005, “토지이용의 변화가 홍수 유출에 미치는 영향분석”, 『한국수자원학회논문집』, 38(4): 301~311.
- 박민지·권형중·김이형, 2005, “HSPF모형을 이용한 토지피복변화에 따른 유출 변화 분석”, 『한국수자원학회논문집』, 38(6): 495~504.
- 박창열·신상영, 2014, “도시홍수 피해 저감을 위한 불투면 규제 효과 분석”, 『서울도시연구』, 15(1): 85~99.
- 서울시정개발연구원, 2011, 『토지이용특성과 침수피해지역 간의 관계 연구』.
- 윤용남, 2003, 『공업수문학』, 서울: 청문각.
- 임장혁·송재우·박성식·박호상, 2007, “택지개발에 따른 표면재료를 고려한 우수유출저감시설의 침투 특성에 관한 실험연구”, 『한국지반환경공학회』, 8(5): 47~55.
- 최지연·송용선·김이형, 2012, “도시화된 토지이용에서 유출되는 강우유출수의 유출특성분석”, 『한국습지학회지』, 14(2): 159~168.
- Barndorff-Nielsen, O., 1978, *Information and Exponential Families in Statistical Theory*, Chichester: Wiley.
- Cox, D. R. and Reid, N., 1987, “Parameter orthogonality

- and approximate conditional inference”, *Journal of the Royal Statistical Society Series B*, 49: 1~39.
- Dunn, P. K., 2004, “Occurrence and quantity of precipitation can be modelled simultaneously,” *International Journal of Climatology*, 24(10): 1231~1239.
- Dunn, P. K. and Smyth, G. K., 2005, “Series evaluation of Tweedie exponential dispersion models densities”, *Statistics and Computing*, 15(4): 267~280.
- Dunn, P. K. and Smyth, G. K., 2008, “Evaluation of Tweedie exponential dispersion model densities by Fourier inversion”, *Statistics and Computing*, 18: 73~86.
- Garavaglia, F., Lang, M., Paquet, E., Gailhard, J., Garcon, R. and Renard, B., 2011, “Reliability and robustness of rainfall compound distribution model based on weather pattern sub-sampling”, *Hydrology and Earth System Sciences*, 15: 519~532.
- Jørgensen, B., 1987, “Exponential Dispersion Models(with Discussion)”, *Journal of the Royal Statistical Society B*, 49: 127~162.
- Jørgensen, B., 1997, *Theory of Dispersion Models*, Chapman & Hall, London.
- Jørgensen, B. and de Souza, M., 1994, “Fitting Tweedie’s compound Poisson model to insurance claim data”, *Scandinavian Actuarial Journal*, 1: 69~93.
- Kaas, Rob, 2005, “Compound Poisson distribution and GLMs—Tweedie’s distribution.”, *MATHEMATICS DAY*: 3.
- Nelder, J. A. and Pregibon, D., 1987, “An extended quasi-likelihood function”, *Biometrika*, 74: 221~232.
- Ohlsson, Esbjörn and Johansson, Björn, 2010, *Non-life Insurance Pricing with Generalized Linear Models*, Springer.
- Revfeim, K. J., 1984, “An initial model of the relationship between rainfall events and daily rainfalls”, *Journal of Hydrology*, 75: 357~364.
- Smyth, G. K., 1989, “Generalized linear models with varying dispersion”, *Journal of the Royal Statistical Society B*, 51: 47~60.
- Smyth, G. K. and Jørgensen, B., 2002, “Fitting Tweedie’s compound Poisson model to insurance claims data: dispersion modelling”, *ASTIN Bulletin*, 32: 143~157.
- Thompson, C. S., 1984, “Homogeneity analysis of a rainfall series: an application of the use of a realistic rainfall model”, *Journal of Climatology*, 4: 609~619.
- Tweedie, M. C. K., 1984, “An index which distinguishes between some important exponential families”, in *Statistics: Applications and New Directions*, Proceedings of the Indian Statistical Institute Golden Jubilee International Conference: 579~604.
- Zhang, Y., 2012, “Likelihood-based and Bayesian Methods for Tweedie Compound Poisson Linear Mixed Models”, *Statistics and Computing*, 23(6): 743~757.
- 원 고 접 수 일 : 2014년 5월 2일
1차심사완료일 : 2014년 5월 30일
최종원고채택일 : 2014년 6월 9일