Tukey's Ladder of Power Transformation을 이용한 태풍예측함수 개발

Development of Typhoon Damage Prediction Function using Tukey's Ladder of Power Transformation

김연수* · 김태균** · 이보림***

Kim, Yonsoo*, Kim, Taegyun**, and Lee, Borim***

Abstract

In this study, we use multiple linear regression to develop a function for predicting damage caused by typhoons. The number of explanatory independent variables constituting the damage prediction function is large and varied because typhoon damage is attributed to a combination of factors such as heavy rain, strong wind, and waves. However, existing data of typhoon damage are insufficient for developing a damage prediction function. To resolve this problem, a model for prediction power only was developed, and the leave-one-out-cross-validation (LOOCV)-forward selection method was applied to select the variables. In addition, the power transformation of Tukey's Ladder of Powers was adopted to create linearity in the dependent and independent variables. The typhoon damage prediction function was developed for 22 districts with more than 16 datasets. The transformation factor λ , representing these 22 districts, was shown to vary according to the regional characteristics of each city and district. The best results occurred when $\lambda = 0.2$ and $\lambda = 0.3$ for the normalized root-mean-square error (NRMSE) and coefficient of determination (R^2) standards, respectively.

Key words : Tukey's Ladder of Power, Typhoon Damage, Prediction Function

요 지

본 연구에서는 태풍피해를 예측하기 위한 태풍피해예측함수를 다중선형회귀모형을 이용하여 개발하였다. 태풍 피해는 호우, 강풍, 풍랑으로 인한 피해가 복합적으로 나타나므로, 피해예측함수를 구성하는 설명변수(독립변수)의 수가 많고 다양하다. 그러나 태풍피해예측함수를 개발하기에는 태풍 피해 자료는 작고 충분치 않다. 이러한 문제점을 해결하기 위하여, 예측력만을 위한 모형을 개발하였고, 변수선정방법은 LOOCV-전진선택법을 적용하였으며, 종속변수와 독립변수들의 선형성을 확보하기 위하여 Tukey's Ladder of Power의 멱승변환을 이용하였다. 태풍피해예측함수는 16개 이상의 자료를 가진 전국 22개 시군구를 대상으로 개발되었으며, 시군구별로 지역특성에 따라 다양한 멱승변환계수 λ 를 가지는 것으로 나타났다. 22개 시군구를 대표할 변환계수 λ 는 NRMSE 기준 $\lambda=0.2$, R^2 기준 $\lambda=0.3$ 일 때 가장 좋은 결과를 나타내었다.

핵심용어 : Tukey's Ladder of Power, 태풍피해, 예측함수

1. 서 론

IPCC에서는 기후변화로 홍수, 폭염, 폭한 등의 기상이변

이 더욱 심각해질 것이고, 그에 따라 자연재해의 증가로 인명 및 재산피해가 지속적으로 증가하고, 규모 역시 커질 것으로 경고하였다.

*정회원, (주)LIG시스템 위험관리연구소 선임연구원(E-mail: civil.engineer@hanmail.net)

Member, Researcher, Risk Management Institute, LIG System Co., Ltd.

**교신저자, 정회원, 경남과학기술대학교 조경학과 교수(Tel: +82-55-751-3306, Fax: +82-55-571-3309, E-mail: tgkim@gntech.ac.kr)

Corresponding Author, Member, Professor, Department of Landscape Architecture

***(주)강산 사원(E-mail: hoyoungqr@naver.com)

Researcher, Kang San Co., Ltd.

우리나라도 기후변화의 영향으로 국지성 호우의 증가와 그로 인한 피해가 증가하고 있으며, 지속적으로 수행되고 있는 국토 개발과정에서 풍수해에 의한 재난을 줄이기 위하여 다양한 구조적, 비구조적 대책을 마련하고 있고, 재해영향평가, 사전재해영향성검토 및 도시계획수립지침 등에 따라 재해에 대한 위험성을 평가하고, 재해에 대비하고 있다(Kim et al., 2017).

우리나라 자연재해의 원인별 발생빈도는 호우와 태풍이 대부분을 차지하는 것으로 조사되었고(Ahn et al., 2015), 호우나 태풍 등으로 인한 자연재해를 사전에 방지하거나, 피해를 경감하려는 노력이 실질적 효과가 있는지를 판정하였다(NDMI, 2001). Jang et al. (2016)은 외국사례를 근거로 국내 실정에 맞는 풍수해피해예측시스템 개발 방향을 제시하였다. Jun et al. (2008)은 재해연보의 호우·태풍 피해자료를 이용하여, 지역특성과 발생횟수 및 피해액에 따라 구분하고 지역별 특성에 따른 피해정도를 분석하였다.

풍수해 피해예측을 위한 연구는 손실·손상함수를 구하여 피해를 예측하는 방법(Kim, Kim, Lee, and Chung, 2014; Kim, Kim, Lee, and Park, 2014; Lee et al., 2017)과 재해통계를 이용하여(Choi, Bak, et al., 2016; Choi, Kim, et al., 2016; Song et al., 2016; Choo, Kwon, et al., 2017; Choo, Yun, et al., 2017; Kim et al., 2017; Kwon and Chung, 2017)으로 구분할 수 있다.

수문기상자료 및 지역특성 인자를 이용하여 피해액을 산정하는 연구는 주로 호우 또는 태풍으로 인한 하천의 범람과 침수시 예상되는 재산상의 피해를 산정하는 것으로, Yang et al. (2008)은 홍수범람에 의한 홍수피해예측 시스템 구축에 대한 연구를 수행하였다. Kim, Kim, Lee, and Chung (2014)은 도시지역의 침수심별 피해 추정함수를 건물피해를 중심으로 개발하였으며, Lee et al. (2017)은 학교건물에 대한 침수심별 손상함수 개발하였다. Lee et al. (2011)은 침수지역에 대한 DEM 자료와 다차원법을 이용한 피해액 산정 방법을 이용한 홍수피해액 산정 방법을 모의하였다.

한편 자연재해로 인한 피해액과 재해 발생 당시의 수문기상학적 인자를 이용한 피해액 산정과 관련하여서는 호우, 태풍, 대설, 강풍, 풍랑 등 다양한 재해에 대하여 연구되고 있다. Kim et al. (2017)은 자료가 부족한 지역의 태풍 피해액을 예측하기 위하여 LOOCV 방법을 이용한 태풍피해함수를 개발하였고, Lee et al. (2016)은 시군구별 태풍내습시 극한기후현상의 강도와 재산피해규모를 예측하였으나 상관성은 떨어지는 것으로 조사되었다. Lee and Chang (2009)은 GIS를 이용한 태풍 위험도 추정과 피해대상 자료의 구축에 대하여, Park et al. (2013)와 Yang and Yi (2013)는 태풍에 동반된 강풍에 의한 피해에 대하여 피해액을 산정하기 위한 모형을 개발하였다.

호우로 인한 피해액 산정은 수문기상학적 인자와 피해 지역의 특성과의 관계(Choi and Seo, 2013, Shim and Kim, 2012)를 이용한 하거나, 수문기상인자와 하천특성을 관계를 이용한 피해액 산정 방법(Kim, Lee, et al., 2016), 수문기상인자

와 피해액 관계를 직접 유도(Choi, Bak, et al., 2016; Choi, Kim, et al., 2016; Kim, Choi, et al., 2016) 하였다. 태풍, 호우 이외에도 강풍에 의한 피해(Song et al., 2016), 대설에 의한 피해(Kwon and Chung, 2017; Kwon et al., 2017; Oh and Chung, 2017), 풍랑에 의한 피해(Ko et al., 2006; Choo, Kwon et al., 2017; Choo, Yun, et al., 2017) 등 다른 자연재해 유형에 대한 연구도 이루어졌다.

수문기상학적 인자, 사회경제적 인자, 유역특성 인자와 피해액 관계는 추정하는 통계학적 방법으로는 비선형 단순 회귀분석을 이용하는 경우가 대부분을 차지하며(Kim, Kim, Lee, and Chung, 2014; Song et al., 2016; Choo, Kwon, et al., 2017; Choo, Yun, et al., 2017; Kim et al., 2017; Kwon and Chung, 2017; Kwon et al., 2017; Lee et al., 2017; Oh and Chung, 2017), Choi and Seo (2013)는 패널모형을, Shim and Kim (2012)는 robust 회귀분석, Choi, Bak et al. (2016) 과 Choi, Kim, et al. (2016)은 시군구별 호우피해함수를 혼합분포를 이용한 로지스틱 회귀분석을 이용하였다.

본 연구에서는 태풍피해예측함수를 다중선형회귀분석을 이용하여 개발하였다. 종속변수와 독립변수의 관계가 비선형이면, 선형성을 확보하기 위하여 종속변수 또는 독립변수에 log 값을 취하여 회귀분석을 시행한다. 선형성 확보를 위하여 Tukey's Ladder of Powers Transformation을 이용하여 종속변수와 독립변수의 선형성을 확보하고, 이에 근거하여 선형다중회귀분석을 실시하여 태풍피해예측함수를 개발하였다. 선형다중회귀식은 결정계수(R^2)가 최대가 되는 식을 개발하는 것이 보통이나, NRMSE (Normalized Root Mean Square Error)를 최소로 하는 식을 개발하였으며, 참고로 결정계수(R^2)를 기준으로 한 함수식도 개발하여, 2가지 기준에 의한 차이도 살펴보았다.

2. 통계적 방법론

2.1 Tukey's Ladder of Powers

Tukey (1977)는 멱함수 변환(power transformation)을 이용하여 변수들 간의 관계의 규칙성을 찾는 방법을 Eq. (1)과 같이 설명하고 있다.

$$y = b_0 + b_1x^\lambda \text{ or } y^\lambda = b_0 + b_1x \quad (1)$$

여기서 λ 는 x , y 의 관계를 가능한 한 가장 직선에 가깝게 만들 수 있는 변환계수이다. 만약 x^λ , y^λ 와 같은 멱함수 변환을 이용하여 Eq. (1)과 같은 변수들 간의 선형관계를 유도할 있다면, 멱함수 변환을 이용하여 변수들을 변환하고, 통계학적인 분석을 수행할 수 있을 것이다. 이때 λ 에 대한 어떠한 통계학적 제약도 없으며, 대표적인 λ 에 대한 변환 형태는 Table 1과 같다.

Table 1. Tukey's Ladder of Powers Transformation

λ	-2	-1	-1/2	0	1/2	1	2
y	$\frac{1}{x^2}$	$\frac{1}{x}$	$\frac{1}{\sqrt{x}}$	$\log x$	$\sqrt{x}$	x	x^2

만약 x 값이 음(-)수이면, 대수적으로 불완전한 경우가 발생할 수 있어 주의하여 변환하여야 한다. 따라서 대부분의 경우 $x > 0$ 인 경우에 멱함수 변환하는 것이 바람직하다. 또한 λ 이 (-)값을 가지면, 멱함수 변환된 값은 분수의 값을 가지게 되고, 만약 x_i 가 증가하는 경향의 값을 가지면, $1/x_i^\lambda$ 는 감소하는 경향을 나타내게 되어 변수들 간의 특성이 변하게 된다. 이와 같은 문제를 해결하기 위하여 Tukey's Ladder of Powers를 Eq. (2)와 같이 다시 정의하고, 대표적인 변환계수 λ 에 대한 변환 형태는 Table 2와 같다.

$$y = \begin{cases} x^\lambda & \text{if } x > 0 \\ \log x & \text{if } x = 0 \\ -(x^\lambda) & \text{if } x < 0 \end{cases} \quad (2)$$

Table 2. Modified Tukey's Ladder of Powers Transformation

λ	-2	-1	-1/2	0	1/2	1	2
y	$\frac{-1}{x^2}$	$\frac{-1}{x}$	$\frac{-1}{\sqrt{x}}$	$\log x$	$\sqrt{x}$	x	x^2

최적의 선형관계를 유도할 수 있는 멱함수 변환의 변환계수 λ 값은 x^λ 와 y 의 상관계수(r , Correlation Coefficient)를 산정하고 상관계수가 가장 큰 λ 를 선정하면 된다.

2.2 다중선형회귀모형

다중선형회귀모형(multiple linear regression model)은 종속 변수(y)에 유의한 영향을 미치는 다수의 독립변수 $x_1, \dots, x_k$ 을 이용하여, 종속변수를 나타내는 것으로 Eq. (3)과 같이 나타낼 수 있다.

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i \quad (3)$$

$$= \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i \quad \text{for } i = 1, \dots, n$$

여기서 총 n 개 자료 중 i 는 i 번째 자료이고, β_0 는 절편(intercept), $\beta_1, \dots, \beta_k$ 는 회귀계수(regression coefficient)라고 하며, β_j 는 j 번째 설명변수 x_j 의 계수이고, ε 는 오차(Error)로써, 반응변수 y 가 설명변수 $x_1, \dots, x_k$ 에 의해 설명되지 않는 나머지 부분을 나타낸다. 회귀모형에서 오차항에 대한 가정으로는 정규분포(normal distribution)를 따르면서 평균이 0이고 동일한 분산을 가지고, 동시에 서로 다른 오차들이 서로 통계적으로 독립임이 가정된다. 회귀모형 추정치는 잔차제곱합(Eq. (4))을 최소화하는 값이다.

$$\sum_{i=1}^n (y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ij})^2 \quad (4)$$

추정한 다중회귀분석모형의 적합도 평가 방법은 결정계수(R^2)를 산정하여 평가하는데, 본 연구에서는 Eq. (5)에 나타나 있는 NRMSE를 이용하여, NRMSE가 최소가 되는 식을 개발하였다. 이는 태풍피해예측을 위하여 개발된 식이므로 예측력을 적합도 평가기준으로 설정한 것이다.

$$SE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

$$NRMSE = \frac{SE}{(\max(y_i) - \min(y_i))}$$

2.3 LOOCV 변수선택법

Kim et al. (2017)은 독립변수가 많고, 자료수가 부족한 경우에 LOOCV (Leave-One-Out Cross Validation)법을 이용하여 태풍피해예측함수를 추정하였다. Kim et al. (2017)은 태풍피해예측함수를 구성하기 위한 설명변수(독립변수)의 가능성이 있는 변수로 호우, 바람, 풍랑을 고려할 수 있는 총 32개의 변수를 선정하고, 이중 어떤 설명변수가 피해액(종속변수)을 가장 설명하는지 결정하기 위하여 LOOCV-NRMSE방법을 이용하였다. LOOCV 방법은 회귀분석을 위한 자료가 부족한 경우에 채택할 수 있는 방법이며, 가장 좋은 예측력을 보이는 설명변수들의 조합을 결정하는 기준을 NRMSE로 하였다. LOOCV-NRMSE 방법에서 설명변수의 결정방법은 모든 설명변수에 대한 모형을 구성하고, 각 모형의 예측력을 NRMSE에 의하여 평가하는 방법은 소위 “Curse of Dimension” 문제를 야기하므로, Kim et al. (2017)은 상관계수가 높은 변수부터 1개씩 추가하여 모형을 구성하고, 추가한 변수에 의하여 NRMSE가 개선되지 않을 때까지 변수를 추가하는 방법을 선택하였다. 이와 같은 과정은 변수 선정 방법 중 전진선택법과 동일한 과정이므로, 이 방법을 LOOCV-전진선택법으로 명명하였다.

본 연구에서 채택한 LOOCV-전진선택법은 오직 태풍피해예측함수 추정 과정에서 예측력을 극대화할 것으로 기대되는 변수선택을 달성하고자 하는데 목적이 있다.

3. 태풍피해예측함수 개발

3.1 대상지역 선정 및 자료 수집

태풍피해예측함수를 개발하기 위한 기초자료로 행정안전부에서 발간하는 재해연보를 이용하였다. 자료기간은 1994년부터 2016년까지 조사하였으며, 우리나라에 직·간접적인 영향을 미쳐 피해가 발생한 총 35개 태풍에 대하여 피해액을 조사하였으며, Table 3에 나타나 있다.

Table 3. Typhoon List (1994-2016)

No.	Typhoon	Affecting start	affecting end	Pressure at 33°	No.	Typhoon	Affecting start	affecting end	Pressure at 33°
1	BRENDAN	1994-07-31	1994-08-02	984hPa	19	SHANSHAN	2006-09-16	2006-09-18	970hPa
2	DOUG	1994-08-09	1994-08-12	978hPa	20	MAN-YI	2007-07-13	2007-07-16	977hPa
3	ELLIE	1994-08-14	1994-08-16	972hPa	21	NARI	2007-09-13	2007-09-18	963hPa
4	SETH	1994-10-10	1994-10-12	972hPa	22	KALMAEGI	2008-07-19	2008-07-21	1007hPa
5	FAYE	1995-07-23	1995-07-24	962hPa	23	DIANMU	2010-08-09	2010-08-12	983hPa
6	TINA	1997-08-07	1997-08-09	982hPa	24	KOMPASU	2010-09-01	2010-09-03	949hPa
7	OLIWA	1997-09-15	1997-09-17	972hPa	25	MALOU	2010-09-05	2010-09-07	989hPa
8	YANNI	1998-09-29	1998-10-01	979hPa	26	MUIFA	2011-08-06	2011-08-10	981hPa
9	ANN	1999-09-17	1999-09-24	993hPa	27	KAHNUN	2012-07-17	2012-07-19	984hPa
10	PRAPIROON	2000-08-30	2000-09-01	966hPa	28	BOLAVEN	2012-08-25	2012-08-30	970hPa
11	SAOMAI	2000-09-12	2000-09-16	972hPa	29	SANBA	2012-09-14	2012-09-17	954hPa
12	RAMMASUN	2002-07-05	2002-07-06	986hPa	30	DANAS	2013.10.07	2013.10.10	970hPa
13	RUSA	2002-08-30	2002-09-01	972hPa	31	NEOGURI	2014.07.08	2014.07.10	996hPa
14	SOUDELOR	2003-06-18	2003-06-19	973hPa	32	NAKRI	2014.08.01	2014.08.04	974hPa
15	MAEMI	2003-09-11	2003-09-13	940hPa	33	HALONE	2014.08.08	2014.08.10	974hPa
16	MEGI	2004-08-17	2004-08-19	976hPa	34	GONI	2015.08.23	2015.08.27	949hPa
17	NABI	2005-09-06	2005-09-08	962hPa	35	CHABA	2016-10-03	2016-10-06	967hpa
18	WUKONG	2006-08-17	2006-08-21	995hPa					

태풍피해예측함수는 재해연보에서 피해를 나타내는 최소 행정단위인 시군구별로 개발하고자 한다. 전국 229개 시군구중 최대 태풍피해를 입은 회수는 총 20회이며, 서울, 경기도 지역은 태풍피해가 한 번도 없는 시군구도 있다. 태풍피해예측함수를 개발하기 위하여 시군구별로 16회 이상 피해를 입은 22개 시군구를 대상으로 선정하였다.

재해연보에서는 자연재해로 인한 피해액을 공공시설물과 사유재산으로 구분하여 제공하고 있기에, 본 연구에서는 공공시설물과 사유재산 피해를 합한 총 피해액을 추정하기 위한 태풍피해예측모형을 개발하고자 한다. 총 피해액의 현재 가치 환산은 한국은행에서 발간하는 2010년 생산자물가지수를 이용하여 환산하였다.

각 태풍 영향시 수문기상인자는 재해연보에 나타난 영향 시작시간부터 영향 종료시간까지 시간우량은 국가수자원관리종합정보시스템(www.wamis.go.kr), 최대풍속은 기상청

(www.kma.go.kr), 위도 33°지점(제주 마라도 인근 지역)을 통과할 때의 중심기압은 국가태풍센터(typ.kma.go.kr)로부터 수집하였다.

태풍피해예측함수를 개발하기 위하여서는 태풍예보에서 제공되는 자료를 변수로 이용하여야 하며, Table 4는 2016년 태풍 ‘차바’ 내습시 10월 4일 21시 시점에서의 태풍예보를 보여준다. Table 4에서 태풍위치, 중심기압을 알 수 있으므로 위도 33° 지점에서의 중심기압을 구할 수 있고, 최대풍속도 파악할 수 있다. 시군구별 피해액을 추정을 위한 시간우량은 기상청에서 운영하는 유인관측소(ASOS)의 자료와 Thiessen 가중평균법을 이용하여 면적강우량을 산정하였다. ASOS에 비하여 AWS (Automatic Weather System)가 조밀하게 분포하고 있으나, AWS가 지금과 같이 조밀하게 설치되기 전에 발생한 태풍피해도 포함하기 위하여 유인관측소 자료를 이용하였다.

Table 4. Forecasting Information for Typhoon CHABA (2016)

status	date	Location		Pressure (hPa)	Max. Wind Vel. (m/s)	storm radius (km)	intensity	size	direction	moving velocity (km/h)
		Lat. (N)	Long. (E)							
Current	10/04/21:00	31.0	125.9	945	45	300	Very Strong	Middle	N	30
Forecasting	10/05/03:00	32.9	126.1	955	40	280	Strong	Small	N	35
	10/05/09:00	34.6	127.8	965	37	250	Strong	Small	NE	41

3.2 태풍피해예측함수 개발

3.2.1 태풍피해예측함수 개발

태풍피해예측함수를 결정하기 위한 독립변수로는 강우 지속기간 1~24 시간별 최대강우량, 총강우량, 1~5일 선행강우량, 최대풍속, 위도 33°지점을 통과할 때의 중심기압 등 총 32개 값을 변수로 선택하였고, 종속변수로는 태풍으로 인한 시군구별 총 피해액을 선택하였으며, 개발 절차는 Fig. 1과 같다.

Fig. 1의 태풍피해예측함수 개발절차를 살펴보면, 22개 시군구에 대하여 변환계수 $\lambda=0.0$ 으로 변환한 종속변수(피해액)과 32개 독립변수중 가장 낮은 NRMSE를 가지는 첫 번째 독립변수를 구한다. 다음으로 첫 번째 독립변수와 나머지 31개 독립변수를 차례로 대입하여 가장 낮은 NRMSE를 가지는 두 번째 독립변수를 구한다. 이때 두 번째 NRMSE가 첫 번째 NRMSE보다 크면 작업을 중단하고, 해당 변환계수에 대한 예측함수는 1차함수가 된다. 이러한 작업을 NRMSE가 더 이상 개선되지 않을 때까지 계속한다. 이때 변환계수 λ 는 0.1씩 증가하여 1이 될 때까지 계속하여 가장 낮은 NRMSE를 가지는 λ 와 그때의 함수를 해당 시군구의 태풍피해예측함수로 결정한다.

Fig. 1. Flowchart

Table 5. Optimum λ According to Criteria R^2 and NRMSE

District	Criteria by NRMSE				Criteria by R^2				data
	λ	order	NRMSE	R^2	λ	order	NRMSE	R^2	
Gangjin gun	0.7	9	3.5	0.996	0.7	9	3.5	0.996	18
Naju city	0.5	4	4.8	0.985	0.5	4	4.8	0.985	16
Pohang city	0.6	8	6.3	0.971	1	9	8.8	0.988	18
Gangseo gu	0.2	9	7.3	0.975	0.2	9	7.3	0.975	20
Haenam gun	0.9	6	8.7	0.941	1	6	10.6	0.956	19
Sacheon city	0.1	6	9.5	0.963	0.1	6	9.5	0.963	16
Tongyeong city	0	9	10.9	0.953	0	9	10.9	0.953	19
Jangheung gun	1	5	11.1	0.927	0.7	6	12.9	0.928	20
Geoje city	0.2	7	11.2	0.927	0.2	7	11.2	0.927	19
Namhae gun	0.3	7	11.7	0.937	0.4	7	11.8	0.947	17
Sancheong gun	0	5	12.3	0.928	0	5	12.3	0.928	17
Gyeongju city	0.3	6	12.5	0.913	0.3	6	12.5	0.913	17
Changwon city	0.1	5	12.7	0.843	0.2	5	14.9	0.869	19
Gimhae city	0.2	5	13.6	0.855	0.3	6	13.9	0.888	16
Hapcheon gun	0	3	13.6	0.840	0.4	7	20.1	0.907	16
Goheung gun	0.1	6	15.0	0.817	0.7	9	18.9	0.877	19
Yeosu city	0.2	4	15.2	0.833	0.2	4	15.2	0.833	20
Wando gun	1	5	15.5	0.770	1	5	15.5	0.770	20
Gijang gun	0.2	4	16.7	0.749	0.3	5	17.8	0.773	19
Boseong gun	0	5	17.4	0.863	0	5	17.4	0.863	16
Ulju gun	0.2	3	17.9	0.743	0.2	3	17.9	0.743	17
Yeongam gun	1	4	18.3	0.875	1	4	18.3	0.875	16

Fig. 2. Result of Prediction Function of the Typhoon Damage

Fig. 1의 절차에 따라 LOOCV-전진선택법을 이용하여 22개 시군구의 태풍피해함수를 산정하였으며, 그 결과가 Table 5에 나타나 있다. Table 5에 나타나 있는 22개 시군구별 태풍피해함수를 살펴보면, RMSE의 표준값인 NRMSE를 기준으로 개발된 태풍피해함수와 R^2 를 기준으로 개발된 태풍피해함수가 일치하는 경우와 포항시, 해남군 등의 경우와 같이 일치하지 않기도 한다. 그러나 일치하지 않는 경우라도 Tukey's Ladder of Powers Transformation의 변환계수 λ 값이 거의 비슷한 것을 알 수 있다. Table 5에서 NRMSE와 R^2 를 기준으로 λ 가 다른 경우는 색을 구분하여 표시하였다. 또 태풍피해예측함수 개발의 결과로 구할 수 있는 예측피해액과 실측치를 Fig. 2에 나타내었다.

Fig. 2에는 부산 강서구, 경북 포항시, 전남 여수시 및 울산 울주군의 조사된 태풍피해와 태풍피해예측함수에 의한 예측치 비교가 나타나 있으며, 22개 시군구의 피해액의 유형과 NRMSE의 크기를 고려하여 4개 시군구만을 도시하였다. 강서구와 포항시는 NRMSE가 6.3%, 7.3%로 우수한 예측력을 보이는 시군구이며, 강서구는 태풍 ‘매미’에 의한 피해가 다른 태풍에 비하여 상대적으로 대단히 크게 나타난 유형이고, 포항시는 태풍 ‘매미’를 포함한 다양한 크기의 피해를 입은 유형이다. 여수시와 울주군은 15.2%, 17.9%로 예측력이 나쁜 시군구로 큰 피해에 대하여서는 작게 추정되고, 작은 피해에 대하여서는 비교적 크게 추정되는 것으로 나타났다.

3.2.2 멱함수 변환계수(λ)에 따른 회귀모형 적합성 변동 특성

회귀분석식의 적합도를 평가하는 기준인 R^2 와 NRMSE 관계는 반비례관계에 있으며, Table 5에서 전반적인 추세는 이러한 경향을 나타내는 것을 알 수 있으나, 항상 이와 같은 형태를 가지는 것은 아니다.

예를 들면, Table 5에서 함천군의 경우 NRMSE가 최소가 되는 λ 를 결정하여 태풍피해함수를 개발한 결과, $\lambda=0$ 일 때 NRMSE는 13.6%로 최소로 선택되었고 $R^2=0.817$ 이다. R^2 가 최대가 되는 λ 는 0.4이며 $R^2=0.907$ 이다. 이때 NRMSE는 18.9%로 나타나 $\lambda=0$ 일 때의 13.6%에 비하여 크게 나타났다.

따라서 Tukey's Ladder of Powers를 이용하여 시군구별 태풍피해예측함수를 개발함에 있어 λ 의 변화에 따른 NRMSE와 R^2 의 변화를 살펴볼 필요가 있으며, λ 의 변화에 따른 NRMSE와 R^2 의 대표적인 변화 형태가 Fig. 3에 나타나 있다.

Fig. 3에는 시군구별 멱함수 변환계수 λ 값 변화에 따른 NRMSE 값과 R^2 변화에 대한 3가지 유형이 나타나 있으며, 22개 대상 시군구의 λ 값 변화에 따른 NRMSE와 R^2 변화가 이들 3가지 형태중 하나를 가지는 것으로 나타났다.

첫 번째 형태는 Fig. 3(a)의 경남 거제시와 같은 형태를 가지는 것으로 작은 λ 값에서 가장 작은 NRSME를 나타내다

Fig. 3. NRMSE and R^2 by λ

Table 6. Comparison of NRMSE and R^2 by λ

Criteria	Statistics	λ											
		0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	
NRMSE (%)	Mean	17.4	16.0	15.6	15.9	17.3	18.4	19.1	19.5	19.7	20.0	20.2	
	Std. Dev.	0.045	0.046	0.048	0.048	0.051	0.059	0.064	0.072	0.069	0.069	0.067	
R^2	Mean	0.72	0.775	0.788	0.792	0.765	0.657	0.605	0.529	0.492	0.476	0.419	
	Std. Dev.	0.178	0.169	0.161	0.157	0.168	0.277	0.311	0.363	0.38	0.377	0.400	

가 λ 가 커질수록 NRMSE가 커지고, 이때 다중회귀분석의 R^2 은 특정 값에서 급격히 작아지는 형태를 가진다. 즉 거제시의 경우 $\lambda=0.2$ 에서 가장 작은 NRMSE를 가지고 이때 R^2 도 최대가 되며, λ 가 증가할수록 NRMSE가 증가하나 $\lambda=0.8$ 이상에서 R^2 은 대단히 나쁜 결과를 나타낸다.

두 번째 형태는 Fig. 2(b)와 같이, 멱함수 변환계수 λ 값 변화에 NRMSE가 크게 영향을 받지 않는 형태로 R^2 변화도 비슷한 형태를 보인다. 전남 장흥군이 이러한 특성을 보이며, 특이하게도 $\lambda=1$, 즉 어떤 변환도 하지 않는 값으로 태풍피해 함수를 개발하는 것이 가장 좋은 결과를 나타내었다. Fig. 3(c)의 세 번째 형태는 $\lambda=0$ 일 때 NRMSE가 최소값을 나타내고 λ 값이 커짐에 따라 NRMSE도 커지며, R^2 도 $\lambda=0$ 일 때 가장 큰 값을 가지고 λ 값이 증가할수록 작아지는 형태를 가진다.

앞에서 살펴본 바와 같이, 멱함수의 변환계수 λ 값의 변화와 NRMSE 또는 R^2 의 관계는 시군구별 태풍피해자료 특성에 따라 달라지므로 시군구별로 다른 λ 를 가진다. 태풍피해가 증가하여 전국 229개 시군구별 피해예측함수를 개발할 경우, 태풍피해액과 수문기상학적 인자 사이의 대표적인 변환계수를 산정하는 것이 바람직하며, 22개 시군구에 대한 변환계수 λ 별 NRMSE와 R^2 에 대한 평균과 표준편차를 산정하였으며, 그 결과가 Table 6에 나타나 있다.

Table 6에서 NRMSE를 기준으로 $\lambda=0.2$ 일 때 22개 시군구의 평균 NRMSE가 15.6%로 가장 좋은 결과를 나타냈고, R^2 를 기준으로 피해함수를 개발하면 $\lambda=0.3$ 일 때 0.792로 가장 좋은 결과를 나타내는 것을 알 수 있다. NRMSE 기준으

로 선정된 $\lambda=0.2$ 의 표준편차 역시 0.048로 비교적 작은 값을 가지며, R^2 기준의 $\lambda=0.3$ 일 때 0.157로 가장 작은 값을 나타내어, 22개 각각의 시군구별로 예측값과 실제 피해액과의 편차가 작다는 것을 나타낸다.

4. 결론

행정안전부, 국토교통부 등 자연 재해에 대응하기 위한 시스템을 가진 정부부처에서 자연재해를 예측하여 대비하는 것은 피해를 줄이고, 재해 발생시 적절한 대응과 사후 조치를 취할 수 있는 근거를 제공하는 것이다. 대부분의 자연재해에 대한 예측은 기상관측기술의 발달로 상당히 높은 수준까지 가능하며, 예측된 재해에 대한 대비도 가능한 수준에 이르고 있다.

본 연구에서는 진로, 강도, 바람의 세기, 강우량 등을 예측할 수 있는 태풍에 대하여, 피해 여부를 판단하고, 피해액을 예측할 수 있는 태풍피해예측함수를 시군구별로 개발하였다.

1994년부터 2016년까지 피해를 야기한 35개 태풍을 대상으로 16회 이상 피해를 입은 22개 시군구를 선정하여 태풍피해예측함수를 개발하였다.

독립변수로는 1~24시간 강우지속기간별 최대 강우량, 총강우량, 1~5일 선행강우량, 최대풍속, 위도 33°통과시의 중심기압을, 피해액을 종속변수로 다중선형회귀분석을 실시하였다. 선형회귀분석모형의 선형성을 향상하기 위하여 Tukey's Ladder of Powers Transformation을 이용하였고, 변환

계수 $\lambda=0.0\sim 1.0$ 까지 0.1 간격으로 역승변환하여 회귀분석을 시행하였고, 회귀분석식의 적합도 판정기준으로 NRMSE와 R^2 를 적용하였다.

22개 시군구별 변환계수 λ 는 피해액과 수문기상인자들의 관계가 지역적 특성에 따라 결정되는 것임을 알 수 있다. 다중선형회귀모형의 적합성 검증에 NRMSE와 R^2 , 두가지 기준이 적용하였으며, 22개 시군구중 9개 시군구의 변환계수 λ 가 다르게 나타났으며, 태풍피해예측함수를 개발하는 것이므로 관측치와 예측치의 오차가 작은 NRMSE 기준의 변환계수 λ 를 선택하는 것이 바람직하다.

비선형 자료의 회귀모형 적용을 위하여, 변수에 log를 취하여 선형회귀분석을 시행하는 경우가 많으나, 태풍피해와 수문기상인자들의 관계는 역함수 변환계수(λ)를 0.2(NRMSE 기준) 또는 0.3(R^2 기준)으로 하는 것이 모형의 적합도가 가장 좋게 나타났다.

태풍피해는 수문기상학적인자 이외에도 재난을 대비하는 사회경제적 인자의 영향을 받을 것으로 추정되므로, 다양한 사회경제적 인자를 발굴하여 모형에 포함할 필요가 있을 것으로 판단된다.

감사의 글

본 연구는 정부(행정안전부)의 재원으로 재난안전기술개발사업단의 지원을 받아 수행된 연구임[MOIS-재난-2015-05].

References

- Ahn, S.H., Park, K.J., Kim, J.Y., and Kim, B.J. (2015). The characteristics of the frequency and damage for meteorological disasters in Korea. *J. Korean Soc. Hazard Mitig.*, Vol. 15, No. 2, pp. 133-144.
- Choi, C.H., Bak, J., Kim, J., Kim, D., and Kim, H.S. (2016). Development of heavy rain damage prediction function using generalized linear model. *Proceedings of 2016 Annual Conference*, Korean Society of Civil Engineers, pp. 109-110.
- Choi, C.H., Kim, J.S., Han, D.G., Oh, S.H., and Kim, H.S. (2016). Development of heavy rain damage prediction function using mixed distribution. *Korean Society for Disaster Information*, pp. 236-237.
- Choi, Y., and Seo, M.H. (2013). Assessing the impact of the factors of urban characteristics on the damage caused by natural disaster using panel model. *Journal of Korea Planning Association*, Vol. 48, No. 5, pp. 309-318.
- Choo, T.H., Kwak, K.S., Ahn, S.H., Yang, D.U., and Son, J.K. (2017). Development for the function of wind wave damage estimation at the Western coastal zone based on disaster statistics. *Journal of the Korea Association-Industrial cooperation Society*, Vol. 18, No. 2, pp. 14-22.
- Choo, T.H., Yun, G.S., Kwon, Y.B., Park, S.J., and Kim, S.R. (2017). Proposal for wind wave damage cost estimation at the Southern coastal zone based on disaster statistics. *Journal of the Korea Content Association*, Vol. 17, No. 4, pp. 267-274.
- Jang, H., Oh, K., Kim, W., and Sim, O. (2016). A study on establishment of strategy for development of storm and flood damage prediction system. *J. Korean Soc. Hazard Mitig.*, Vol. 16, No. 3, pp. 117-124.
- Jun, H.D., Park, M.J., Kim, G.Y., and Kim, J.W. (2008). Damage analysis of meteorological disasters for each district considering the characteristics of a district. *J. Korean Soc. Hazard Mitig.*, Vol. 8, No. 2, pp. 75-82.
- Kim, J., Choi, C., Kim, D., Bak, J., and Kim, H.S. (2016). Development rain damage prediction functions using artificial neural networks. *Proceedings of 2016 Annual Conference*, Korean Society Civil Engineers, pp. 64-65.
- Kim, J., Kim, T., and Lee, B.R. (2017). An analysis of typhoon damage pattern type and development of typhoon damage forecasting function. *J. Korean Soc. Hazard Mitig.*, Vol. 17, No. 2, pp. 339-347.
- Kim, S.H., Kim, B.S., Lee, C.H., and Chung, J.H. (2014). Development of depth-damage function by investigating flooded area with focusing on building damage. *Journal of Korea Water Resources Association*, Vol. 47, No. 8, pp. 717-728.
- Kim, S.H., Kim, B.S., Lee, C.H., and Park, I.C. (2014). Development of loss damage function for disaster assessment. *Proceedings of 2014 Annual Conference*, Korean Society of Hazard Mitigation, p. 72.
- Kim, S.H., Lee, C.H., Kim, Y.S., and Hwang, S.B. (2016). A study on development of flood damage estimation function for river infra structure with NDMS and WAMIS. *J. Korean Soc. Hazard Mitig.*, Vol. 16, No. 4, pp. 81-87.
- Ko, J.S., Choi, S.C., and Yum, H.M. (2006). Analysis of Ulsan harbor coastal structure by typhoon Nabi. *Magazine of Korean Soc. Hazard Mitig.*, Vol. 6, No. 4, pp. 103-112.
- KMA (Korea Meteorological Administration). Retrieved from <http://www.kma.go.kr>
- Kwon, S.H., and Chung, G. (2017). Estimation of snow

- damages using multiple regression model: The case of Gangwon province. *Journal of the Korean Society of Civil Engineers*, Vol. 37, No. 1, pp. 61-72.
- Kwon, S.H., Lee, J., and Chung, G. (2017). Snow damages estimation using artificial neural network and multiple regression analysis. *J. Korean Soc. Hazard Mitig.*, Vol. 17, No. 2, pp. 315-325.
- Lee, S.S., and Chang, E.M. (2009). Application of GIS to typhoon risk assessment. *The Journal of GIS Association of Korea*, Vol. 17, No. 2, pp. 243-249.
- Lee, C.H., Kim, S.H., Hwang, S.B., and Kim, G.H. (2017). A study on development of flood depth-damage functions focused on school buildings. *J. Korea Water Resour. Assoc.* Vol. 50, No. 8, pp. 513-520.
- Lee, G.H., Park, K.W., Yu, W.S., Jung, K.S., and Jang, C.L. (2011). A study on flood damage estimation using DEM-based flood inundation model and MD-FDA. *J. Korean Soc. Hazard Mitig.*, Vol. 11, No. 5, pp. 327-336.
- Lee, J.S., and Choi, H.I. (2016). Flood vulnerability index estimated by comparing analysis methods of flood damage data. *J. Korean Soc. Hazard Mitig.*, Vol. 16, No. 2, pp. 427-435.
- Lee, S.W., Ahn, S.H., Lim, B.H., and Choi, G.Y. (2016). Relationships between intensity of extreme climate events and magnitude of damages for different typhoon tracks in the Republic of Korea. *Journal of The Korean Association of Regional Geographers*, Vol. 22, No. 2, pp. 450-465.
- NDMI (National Disaster Management Research Institute). (2001). *Study of mitigation efforts for the repetitive flooded area*.
- National Typhoon Center. Retrieved from <http://typ.kma.go.kr>
- Oh, Y., and Chung, G. (2017). Estimation of snow damage and proposal of snow damage threshold based on historical disaster data. *Journal of the Korean Society of Civil Engineers*, Vol. 37, No. 2, pp. 325-331.
- Park, J.K., Jung, W.S., Kim, E.B., and Kim, J.S. (2013). Characteristics of the state and distribution of damage by typhoon. *Proceedings of 2013 Annual Conference*, Korean Society of Hazard Mitigation, p. 40.
- Shim, J.H., and Kim, J.E. (2012). An analysis on the interrelationship between land-use characteristics and damages caused by natural hazards. *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 13, No. 9, pp. 4319-4325.
- Song, C.Y., Lee, H.J., and Lee, C.J. (2016). A simple regression model for predicting the wind damage according to correlation analysis between wind speed and damage: Gyeongsangbuk-do. *Proceeding of 2016 Conference*, Korean Society for Disaster Information, pp. 207-211.
- WAMIS (Water Resources Management Information System). Retrieved from <http://www.wamis.go.kr>
- Yang, D.M., Jang, D.W., and Yoon, J.H. (2008). A study on damage prediction system for natural disaster. *Korea Multimedia Society*, Vol. 12, No. 1, pp. 43-49.
- Yang, W., and Yi, W.H. (2013). Damage pattern analysis of buildings by wind storm and typhoon. *Proceedings of 2013 Annual Conference*, Korean Society of Hazard Mitigation, p. 198.

Received	■ June 11, 2018
Revised	■ June 14, 2018
Accepted	■ August 23, 2018