

## 다변량 통계분석을 이용한 남부 내륙지역 태풍피해예측모형 개발

김연수·김태균\*\*

(주)LIG시스템 위험관리연구소

\*경남과학기술대학교 조경학과

## Development of Typhoon Damage Forecasting Function of Southern Inland Area By Multivariate Analysis Technique

Yonsoo Kim·Taegyun Kim\*\*

*Risk Management Research Institute, LIG System Co., Ltd.*

*\*Department of Landscape Architecture, Gyeongnam National Univ. of Science & Tech.*

(Received : 19 August 2019, Revised: 17 October 2019, Accepted: 17 October 2019)

### 요약

본 연구에서는 남부 내륙지역에 속한 시군구별 태풍으로 인한 피해를 예측할 수 있는 태풍피해예측모형을 개발하였다. 내륙 지역의 태풍 피해는 호우, 강풍으로 인한 피해가 복합적으로 발생하므로, 모형을 구성하는 변수가 많고 다양하나, 내륙지역 시군구 단위의 피해사례는 모형을 개발할 만큼 충분하지 않다. 태풍피해 관련 수문기상 자료는 3시간 간격 지속기간별 최대 강우량, 총강우량, 1-5일 선행강우량, 최대풍속 및 제주도 인근 지역에서의 태풍중심기압을 이용하였다. 피해자료의 부족을 고려하기 위해 군집화를 하였으며, 강우 관련 자료의 다중공선성을 제거하기 위하여 주성분분석 등 다변량 통계분석을 이용하여 권역별(경남, 경북, 전남, 전북)로 피해예측모형을 개발하였다. 모형에 의한 태풍피해추정치와 실측치는 최대 2.2배 정도까지 차이가 발생하였는데, 이는 강풍에 의한 피해를 추정하기 어렵고, 전국 69개 ASOS 관측소에 의한 강우자료가 지역적 강우특성을 제대로 반영하지 못하기 때문인 것으로 추정된다.

핵심용어 : 다변량 통계분석, 태풍 피해예측 모형, 재해통계

### Abstract

In this study, the typhoon damage forecasting model was developed for southern inland district. The typhoon damage in the inland district is caused by heavy rain and strong winds, variables are many and varied, but the damage data of the inland district are not enough to develop the model. The hydrological data related to the typhoon damage were hour maximum rainfall amount which is accumulated 3 hour interval, the total rainfall amount, the 1-5 day anticipated rainfall amount, the maximum wind speed and the typhoon center pressure at latitude 33° near the Jeju island. The Multivariate Analysis such as cluster Analysis considering the lack of damage data and principal component analysis removing multi-collinearity of rainfall data are adopted for the damage forecasting model. As a result of applying the developed model, typhoon damage estimated and observed values were up to 2.2 times. this is caused it is difficult to estimate the damage caused by strong winds and it is assumed that the local rainfall characteristics are not considered properly measured by 69 ASOS.

Key words : Multivariate Analysis Technique, Typhoon Damage Forecasting Model, Disaster Statistics

## 1. 서 론

우리나라는 호우, 태풍, 가뭄 등 자연재해로 인한 인명 및 재산상의 피해가 빈발하므로, 재해영향평가, 사전재해영향성검토(자연재해대책법) 및 도시계획수립지침 등을 수립하여 자연재해에 대한 위험성을 평가하고, 자연재해를 줄이기 위한 구조적, 비구조적 대책을 수립하도록 하고 있다.

우리나라 자연재해 발생빈도는 1980년대 41.9%에서 2000년대 19.3%로 줄어들었으나, 피해가 집계된 이후 총 피해액 중 1980년대 11.4%에서 2000년대 60.6%로 증가하여, 발생회수는 감소하였으나 피해가 대형화되었음을 알 수 있다. 또 기상재해 원인으로는 복합재해(5.1%)보다는 대부분이 단일 기상재해(94.9%)로 조사되었다. 피해회수 기준으로 호우 37.8%, 태풍 8.7%였으나, 피해액 기준으로 태풍이 45.6%, 호우가 35.4%를 차지하는 것으로 조사되었다(Ahn et al., 2015). 자연재해로 인한 피해예측연구는 손실·손상함수를 구하여 피해를 예측하는 방법(Kim et al, 2014) (Kim et al., 2014; Lee et al., 2017)과 재해통계를 이용한 방법(Kwon & Chung.,

\* To whom correspondence should be addressed.  
Department of Landscape Architecture, Gyeongnam National Univ. of Science & Tech.  
E-mail: tgkim@gntech.ac.kr

2017)(Kwon and Chung, 2017; Kim et al, 2018; Kim et al, 2017b; Kim et al, 2017a; Kim et al, 2016; Oh and Chung, 2017; Lee and Choi, 2016; Lee et al, 2015; Choi et al, 2017a; Choi et al, 2017b; Choi et al 2016b; Choi et al, 2016a; Choo et al, 2017a)으로 구분할 수 있다.

태풍재해는 호우, 강풍, 저기압으로 인한 해수면상승과 파랑으로 인한 피해가 동시에 발생하는 특징을 가진다. 태풍으로 인한 지자체별 피해특성에 대하여, Jun et al(2018)은 230개 시군구를 내륙, 해안, 산지, 평지 등 지역적 특성에 따라 구분하고, 재해연보에 나타난 발생횟수 및 피해액을 이들 특성에 따라 분석하였고, Lee et al(2016)은 태풍 내습 시 강풍, 호우 등에 따른 지역별 피해액의 지역적 특성과의 관계를 연구하였다.

Park et al(2013)과 Yang and Yi(2013)는 태풍에 동반된 강풍에 의한 피해를 연구하였으며, 태풍에 따른 재해 위험도와 관련하여, Lee and Chang(2009)는 GIS를 이용한 태풍 위험도 추정과 피해대상 자료의 구축에 대한 연구를 수행하였고, Lee and Choi(2016)은 시군구별 사회적, 지형적, 경제적 홍수피해 특성을 반영할 수 있는 홍수피해지표를 개발하였다.

자연재해에 따른 피해액을 수문기상학적 인자를 이용하여 직접 예측하는 연구도 이루어졌는데, Kim et al(2017b), Kim et al(2018)은 태풍의 수문기상학적 인자를 설명변수로 하는 다중선형회귀분석을 이용한 피해액 산정 방법을 개발하였다.

자연재해 중 호우(Kim et al, 2017a; Kim et al, 2016; Choi et al, 2017a; Choi et al, 2017b; Choi et al, 2016a; Choi et al, 2016b)피해예측은 지속기간별 최대강우량, 총강우량, 선행강우량, 시군구의 면적 등을 주성분 분석과 로지스틱회귀모형을 이용하였고, 태풍피해(Kim et al, 2018; Kim et al, 2017b)는 호우와 같은 수문기상인자와 풍속, 태풍중심기압을 이용하여 다중선형회귀모형을 이용하였다. 풍랑(Choo et al, 2017a; Choo et al, 2017b; Choo et al, 2018a; Choo et al, 2018b)은 파랑과 관련된 풍속, 파고, 연안재해영향지수, 노출지수, 민감도지수 등의 군집화를 시행한 후 다중선형회귀분석을 실시하였다. 대설(Kwon & Chung, 2017; Oh and Chung, 2017) 피해는 신적설심, 평균습도, 최저기온 등 수문기상인자와 농촌인구밀도, GRDP 등을 이용하여 다중선형회귀분석 하였다. 강풍피해예측 모형은 전국의 시군구자료를 대상으로 피해액과 최대순간풍속, 피해액과 최대풍속의 단순선형회귀분석을 실시하였으나 결정계수는 낮게 나타났다(Song and Yang, 2016).

자연재해피해예측은 재해원인별 수문기상인자와 다양한 지역적, 사회-경제 원인을 이용하여 모형을 개발하였고, 통계학적 방법으로 주성분분석, 군집화, 비선형 회귀식, 다중회귀모형, 로지스틱회귀 모형 등을 이용하여 규명하였으며, 시군구의 특성에 따라 정확도는 다양하게 나타난다(Kwon and Chung, 2017; Kim et al, 2018; Kim et al, 2017a; Lee et al, 2017; Choi et al, 2017a; Choi et al, 2017b).

재해관련 업무의 최소 단위가 시군구이기 때문에 본 연구에서는 태풍피해 모형의 공간적 범위를 시군구 단위로 하여 개발하고자 하였다. 지속기간별 최대강우량, 선행강우량 등 호우 관련 수문자료와 이들 간의 다중공선성 문제를 해결하기 위하

여 주성분분석(Park et al, 2003; Choi et al, 2017b)을 실시하였다. 본 연구는 태풍피해발생 회수의 부족으로 인한 피해모형 개발이 어려운 시군구에 대하여 군집화를 통해 태풍피해를 예측할 수 있는 피해예측모형을 개발함으로써, 지자체 중심으로 이루어지는 재난 관련 행정(대비, 대응, 조사, 복구)에 활용될 수 있도록 하는데 그 목적이 있다.

## 2. 이론적 배경

### 2.1 계층적 군집화

계층적 군집화(Hierarchical Clustering)란 많은 자료를 속성이 비슷한 몇 개의 그룹(군집)으로 나누어 각 그룹의 특성을 파악하고 공통점을 발굴하여, 이를 효율적으로 활용하고자 하는 통계적 분석 방법이다. 군집분석의 기본원리는 분석하고자 하는 여러 특성들을 유클리드 거리(Euclidean distances) 또는 유클리드 제곱거리를 구하고, 이를 유사성(similarity) 거리로 환산하여 상대적으로 거리가 가까운 개체들을 동질적인 집단으로 군집화하는 것이다.

군집화 방법에는 계층적 군집화, k-Means clustering, Gaussian mixture model, Self-organizing maps 등 다양한 방법이 있으며, 시군구별 태풍피해특성을 계층적 군집화 방법에 의해 군집화하고자 한다. 계층적 군집화에는 최단연결법, 최장연결법, 평균연결법 등이 있으며, 이 중 최단연결법을 적용하고자 한다. 최단연결법은 관찰된 데이터 중 가까운 데이터끼리 묶어가며 덴드로그램(dendrogram)으로 표시되는 클러스터 트리를 생성하여 클러스터의 다단계 계층을 구축하는 방법으로 덴드로그램(dendrogram)을 직접 보고 군집을 나눌 수 있는 장점이 있다.

본 연구에서는 시군구별 태풍피해와 피해발생 당시의 강우량, 선행강우량, 최대풍속, 중심기압 등 상관관계를 분석하고, 상관계수를 이용하여 시군구별 태풍피해 군집화를 시도하였다.

### 2.2 주성분 분석

주성분분석은 서로 상관관계가 높고 독립적이지 않은 다수의 변수를 조합하여, 그 변수들의 정보를 가장 많이 나타낼 수 있는 새로운 변수를 만들어 내기 위한 통계적 기법으로, 고차원의 자료를 저차원의 자료로 축소하거나 복잡한 고차원의 자료를 단순화하여 해석하기 위한 다변량자료 분석기법이다.

주성분분석은 주어진 자료의 통계적 특성을 이용한 방법으로, 자료들의 분산이 최대가 되는 기본벡터를 찾아 기존 자료를 대체할 새로운 변수를 결정하는 과정이다. 주성분을 추출하는 과정은 주어진 자료를 이용하여 변수들의 정보를 최대한 반영할 수 있는 주성분을 먼저 추출한다. 두 번째 주성분은 첫 번째 주성분과 직교하는 벡터로 분산이 두 번째로 크다. 이러한 과정을 변수들의 정보를 모두 반영할 때까지 순차적으로 변수의 수만큼 주성분을 추출하게 된다.

주성분분석은 선형결합으로 Eq. (1)과 같이 나타난다. Eq. (1)에서 성분점수로 이루어진 행렬 PC는 각 기존 자료와 주성분들의 관계를 나타낸다. X는  $(p \times 1)$ 을 가지는 기존 자료의 행렬,  $v$ 는 고유값(eigenvalue)에 대한 고유벡터(eigenvector)이

다. 고유값은 Eq. (2)로 나타나는데, 앞에서 설명한 바와 같이 분산으로부터 산정된다. 또한 주성분들은 독립이므로 Eq. (3)에서 공분산은 '0'의 값을 가지고, 공분산행렬은 고유값  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$  가 된다.

$$PC = V' \cdot X \tag{1}$$

$$Var(PC_i) = v_i' \sum v_i = \lambda_i, \quad i = 1, 2, \dots, p \tag{2}$$

$$Cov(PC_i, PC_k) = v_i' \sum v_k = 0, \quad i \neq k \tag{3}$$

여기서 i번째 주성분은 다음의 3가지 제약조건하에서 분산을 최대화하는 선형결합으로 나타낼 수 있다.

- ① 분산은 선형결합계수로 사용되는 고유벡터 V에 따라 증가될 수 있기 때문에 요소들의 승합이 1이 되도록 제한한다.
- ② 주성분분석에서 두 개의 선형결합관계는 서로 독립이라고 가정하기 때문에 두 주성분의 공분산은 0이 되어야 한다.
- ③ 자료에서 최대의 고유값을 산출하기 위해 특성방정식(characteristic equation)  $|A - \lambda I| = 0$ 가 성립하여야 한다. 여기서 A는 기존자료의 상관계수행렬 또는 표준화된 변수의 분산-공분산 행렬, I는 단위행렬을 의미한다(Choi et al, 2017b).

### 2.3 다중회귀모형

다중회귀모형(multiple regression model, MR)은 반응변수 y에 유의한 영향을 미칠 것으로 기대되는 설명변수들인  $x_1, \dots, x_k$ 을 고려하여, 반응변수가 여러 개의 설명변수들로 설명되는 아래와 같은 구조를 갖는다.

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \varepsilon_i \tag{4}$$

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

태풍피해를 예측하기 위하여 개발된 다중회귀분석모형의 적합도(Goodness-Of-Fit)는 RMSE (Root Mean Square Error)를 정규화한 NRMSE(Normalized RMSE)를 적용하여 평가하였다.

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

$$NRMSE = \frac{SE}{(\max(y_i) - \min(y_i))} \times 100$$

Eq. (5)에서 RMSE는 실제값과 예측값 사이의 평균 차이를 의미하며, NRMSE는 RMSE를 실제값의 범위로 표준화한 값으로 단위는 백분율(%)이다.

## 3. 분석 및 결과

### 3.1 태풍피해자료 수집

본 연구에서는 태풍피해 예측모형을 개발하기 위한 기초자료로 행정안전부에서 발간하는 재해연보를 이용하였다. 1994년~2016년까지 발생한 태풍 가운데 우리나라에 직·간접적인 영향을 미쳐 피해가 발생한 35개 태풍에 대하여 피해액을 조사하였는데, 태풍과 호우태풍을 태풍 피해와 관련된 원인으로 분류하고 있으며, 과거 발생한 태풍목록은 Table 1과 같다.

재해연보에는 기초자치단체인 시군구별로 하천, 도로, 철도 등

Table 1. 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					

14가지 공공시설과 건물, 농지 등 11가지 사유재산 피해액을 복구비 기준으로 나타내고 있으며, 태풍이 영향을 미치지 시작한 날과 종료한 날이 모두 기재되어 있다. 재해연보에 나타난 태풍 피해 자료의 정확성을 확인하기 위하여 국가태풍센터의 태풍자료와 비교하였고, 오류가 있는 태풍은 제외하였다. 제외된 원인은 국가태풍센터에 태풍발생이 없는 기간에 피해가 발생한 사례(2004년 7월 2일부터 7월 17일 피해 등 총 6개 피해사례), 태풍이 소멸한 후에 우리나라에 피해를 일으킨 사례(1997년 OLIWA 등 총 8개 피해사례) 등이 있다. 과거 태풍피해액을 현재시점 화폐가치로 산정하기 위해 2016년 생산자물가지수를 고려하여 환산기준을 반영하였으며, 생산자물가지수는 한국은행 (<http://ecos.bok.or.kr/>)의 자료를 이용하였다.

35개 태풍으로 인한 시군구별 태풍피해 횟수는 Fig. 1과 같으며, 전남, 부산, 울산, 경남, 경북 및 강원도 연안지역과 산청군, 함양군 등의 시군구가 태풍피해를 많이 발생한 것을 알 수 있다. 강원도 및 경북 일부 지방을 제외한 연안지역 시군구의 피해는 세력을 어느 정도 유지한 상태에서 내륙으로 상륙하거나 연안지역을 스쳐지나가면서 발생하는 피해이며, 상륙 후에는 세력의 약화로 피해가 작게 나타난다. 강원도 및 경북 북부지역의 연안지역 시군구 및 경남 산청군 및 함양군은 인근에 위치한 백두대간으로 인한 산악효과로 강우량이 증가하며, 피해도 자주 발생하는 것을 알 수 있다. 시군구별 태풍피해회수는 35개 태풍에 대하여 최대 20회에서부터 서울 등 경기도 지역은 한번도 태풍피해가 발생하지 않은 지역까지 다양한 분포를 가진다.

3.2 대상지역 및 모형 변수 선정

앞에서 언급한 바와 같이, 시군구별 태풍피해 횟수는 지역적 편차가 심하며, 내륙지역의 시군구는 태풍피해예측모형을 개발하기 어려울 정도로 태풍피해가 작다. 본 연구에서는 피해 횟수가 작은 지역을 군집화하여 자료수를 확장하고, 군집별 태풍피해예측모형을 개발하고자 한다. Fig. 2.에는 연안지역과

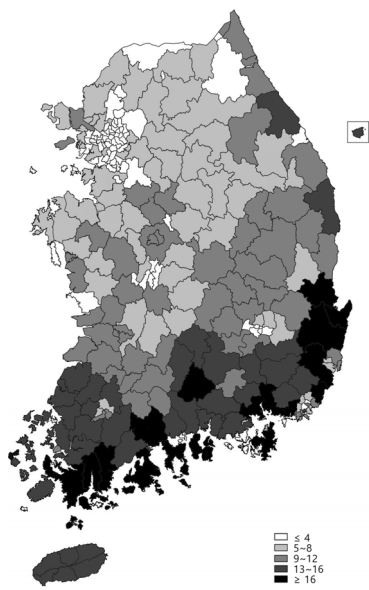
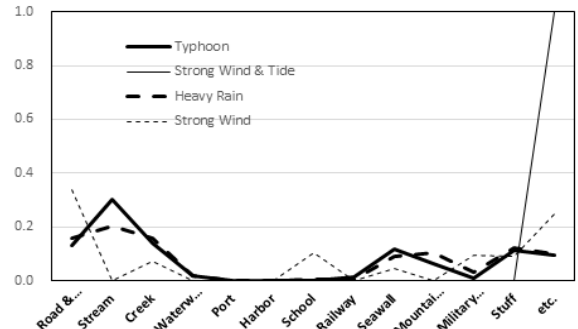
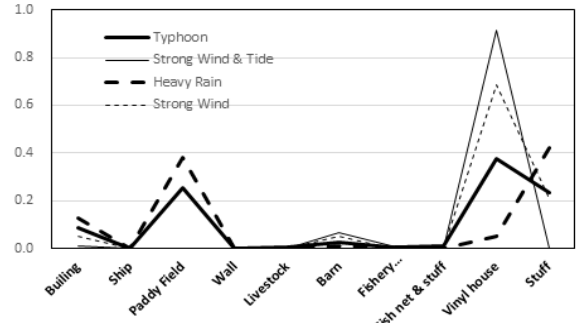


Fig. 1. Number of Damage

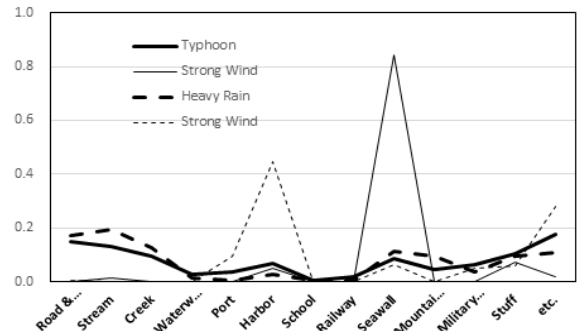
내륙지역의 시설물별 호우피해액과 태풍피해액의 구성비가 나타나 있다. 내륙지역의 경우는 태풍피해와 호우피해가 비슷한 양상을 보이는데 반하여, 연안지역은 바람과 조위상승으로 인한 풍랑의 피해가 상대적으로 커짐을 알 수 있다. 전환동 등 (Jun et al, 2008)은 시군구를 6개 지역으로 구분하여 피해양상이 다름을 보인 바 있다.



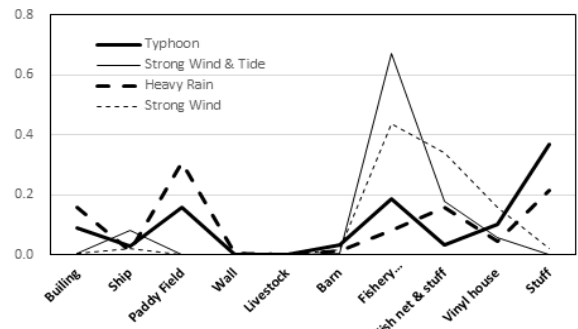
(a) Damage Ratio of Inland Districts Public Structures



(b) Damage Ratio of Inland Districts Private Facilities



(c) Damage Ratio of Coastal District Public Structures



(d) Damage Ratio of Coastal Districts Private Facilities

Fig. 2. Damage Ratio of Structures an Facilities by Disaster Type

본 연구에서는 태풍피해가 비교적 많이 발생하는 낙동강 유역권역인 부산, 대구, 울산, 경남, 경북 지역과 영산/섬진강 유역권인 광주, 전남, 전북 등 남부 지역의 내륙에 위치한 시군구를 대상으로 태풍피해예측모형을 개발하고자 한다.

### 3.3 모형 변수 선정

태풍발생시 태풍예보는 기상청에서 6시간 간격으로 태풍의 진로, 중심기압 위치, 중심기압 세기 등을 발표하며, 1시간 간격 예상강우량을 풍수해피해예측시스템(NDMS)에 제공하며, 제주도 인근을 기점으로 직접적인 피해가 예상될 경우에는 3시간 간격으로 기상특보를 발표한다.

본 연구는 태풍내습시 발생할 수 있는 피해를 예측하는 시스템의 피해모형을 산정하는 것이므로, 태풍예보시 제공되는 이용 가능한 자료를 피해예측모형의 변수로 선정하고자 하며, 풍수해피해예측시스템에 제공되는 1시간 강우량을 이용한 3시간 간격(3, 6, 9, 12, 15, 18, 21, 24시간) 강우지속기간별 최대강우량과 총강우량, 선행강우량, 최대풍속 및 중심기압을 예측모형의 변수로 선정하였다.

3시간 간격 강우지속기간별 최대강우량을 선정한 이유는, 태풍이 직접적인 영향을 미칠 때는 3시간 간격으로 기상특보를 발생하므로, 3시간 간격으로 예상강우량이 변하기 때문이고, 선행강우량은 NRCS의 초과우량산정방법에서 5일 선행강우량을 기준으로 선행토양수분함수를 결정하므로 1-5일 선행강우량을 선택하였다. 중심기압은 태풍예보에서 제공하는 태풍진로에 따른 위치, 중심기압을 이용하여 제주도 남단 마라도 인근지역인 위도 33° 지점을 통과할 때의 예상 중심기압을 변수로 채택하였다.

35개 태풍을 이용한 피해예측모형 개발을 위한 기왕의 수문기상 인자중 시강우량은 69개 ASOS 관측소의 시우량, 최대풍속 및 위도 33° 지점을 통과할 때의 중심기압은 기상청 (<http://www.kma.go.kr>)으로부터 수집하였다. 시우량자료는 2005년부터 AWS 관측자료가 있으나, 피해자료가 수집된 기간이 1994년부터이므로, 자료의 일관성유지를 위하여 AWS자료를 이용하지 않고 ASOS자료를 이용하였다.

## 4. 태풍피해예측모형 개발 및 분석

### 4.1 태풍피해예측모형의 개발

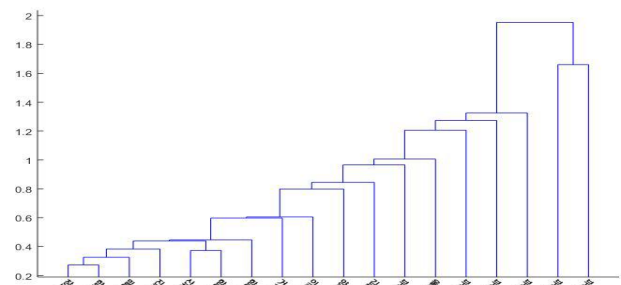
Fig. 2.에서 알 수 있는 바와 같이, 태풍피해형태는 연안지역과 내륙지역이 파랑과 바람의 강도의 차이로 다른 형태를 보이므로, 연안지역을 제외한 시도별로 내륙지역만을 대상으로 한다. 또 4개 광역시에 속한 기초 자치단체(군구)의 경우는 면적이 작고, 태풍피해도 작으므로, 경남권(부산, 울산, 경남), 경북권(대구, 경북), 전남권(광주, 전남), 전북권(전북)으로 구분하여 모형을 개발하고자 하며, 모형 개발 절차는 다음과 같다.

- ① 권역별 시군구별 태풍피해와 강우관련 변수, 최대풍속, 중심기압의 상관계수 산정
- ② 상관계수를 이용한 계층적 군집화

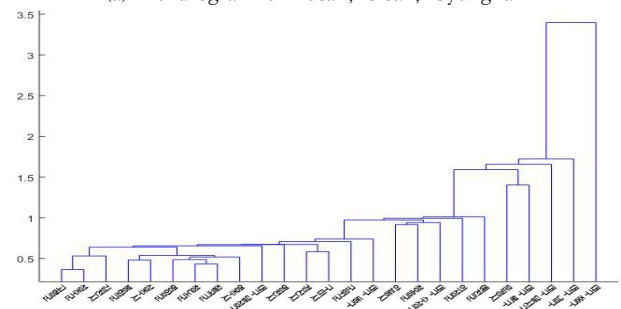
- ③ 군집별 주성분분석 및 주성분을 이용한 태풍피해예측모형 개발
- ④ 태풍피해예측모형 적용 및 특성 분석

### 4.2 권역별 군집화

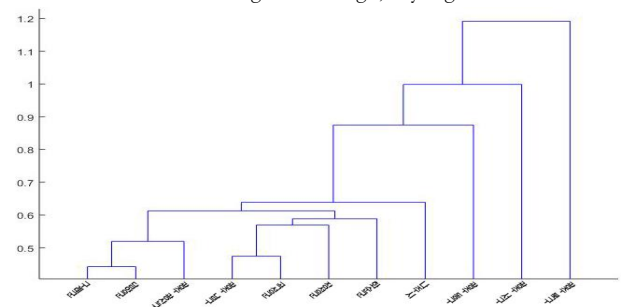
권역별 군집화를 위한 변수로 태풍피해액과 3시간 간격 강우지속기간별 최대강우량(8개), 총강우량, 1-5일 선행강우량(5개), 최대풍속, 중심기압 등 총 16개 태풍관련 수문기상인자의 상관계수를 산정하였고, 이를 근거로 4개 권역에 대하여 계층적 군집화 방법에 의해 군집화를 시행하였으며, 그 결과가 Fig. 3.에 나타나 있다.



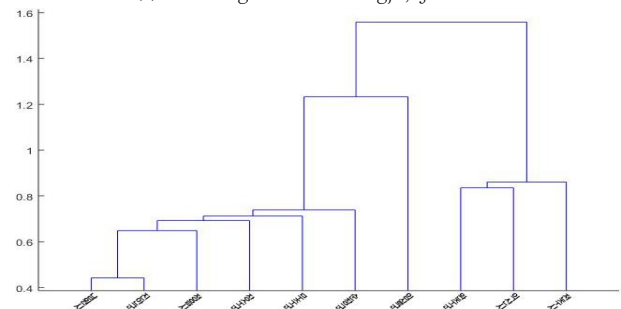
(a) Dendrogram of Busan, Ulsan, Gyeongnam



(b) Dendrogram of aegu, Gyeongbuk



(c) Dendrogram of Kwangju, Jeonnam



(d) Dendrogram of Jeonbuk

Fig. 3. Cluster Analysis

Table 2. Districts of Cluster & Province

Province	cluster	District
Gyungnam	1	Geochang-gun, Gimhae-si, Miryang-si, Yangsan-si, Uiryeong-gun, Jinju-si, Changnyeong-gun, Haman-gun, Hamyang-gun, Hapcheon-gun, Geumjeong-gu(Busan), Dongnae-gu(Busan), Busanjin-gu(Busan), Buk-gu(Busan), Jung-gu(Ulsan)
	2	Sasang-gu(Busan), Jung-gu(Busan)
Gyungbuk	1	Gyeongsan-si, Gumi-si, Gunwi-gun, Gimcheon-si, Bonghwa-gun, Sangju-si, Seongju-gun, Andong-si, Yeongyang-gun, Yeongju-si, Yeongcheon-si, Yecheon-gun, Uiseong-gun, Cheongdo-gun, Cheongsong-gun, Chilgok-gun, Dalseong-gun(Daegu), Dong-gu(Daegu), Suseong-gu(Daegu)
	2	Mungyeong-si, Buk-gu(Daegu), Dalseo-gu(Daegu), Nam-gu(Daegu), Jung-gu(Daegu)
Jeonnam		Gwangju Metropolitan City, Gokseong-gun, Gurye-gu, Naju-si, Damyang-gun, Jangseong-gun, Hwasun-gun
Jeonbuk		Namwon-si, Muju-gun, Sunchang-gun, Wanju-gun, Iksan-si, Imsil-gun, Jangsu-gun, Jeonju-si, Jeongeup-si, Jinan-gu

Fig. 3.에서 경남권은 2개, 경북권은 2개의 군집과 군집화에 포함되지 않는 개별 시군구로 대구 달서구, 남구, 전남, 전북은 1개의 군집으로 구분하였고, 권역별 군집은 Table 2에 나타나 있다.

4.3 주성분분석에 의한 태풍피해예측모형 개발 및 분석

Kim et al(2017b)와 Kim et al(2018)는 부산시 강서구를 대상으로 태풍피해예측모형을 개발함에 있어, 강우관련 변수로 1시간-24시간 강우지속기간별 강우량, 총강우량, 1-5일 선행강우량, 최대풍속, 중심기압 등 총 30개의 변수를 이용하였다. 이때 다중공선성 문제가 있음에도 불구하고, 태풍피해자료의 부족과 예측력을 극대화하기 위하여 다중공선성을 고려하지 않고 태풍피해예측모형을 개발하였다.

그러나 지속기간별 최대량 산정과정에서 강우지속기간이 연속시간이어야 하므로, 다중공선성 문제는 당연히 발생할 수 밖에 없다. 3시간간격 강우지속기간별 최대강우량(8개), 총강우량 및 1-5일 선행강우량(5개), 최대풍속 및 중심기압 등 총 16개의 변수가 생산되므로, 대단히 높은 차원의 변수를 가지는 모형이 될 수 있다. 따라서 차원을 줄이기 위하여 16개 변수들을 이용한 주성분분석을 실시하여 태풍피해예측모형을 개발하였다.

주성분분석은 Table 2에 나타나 있는 권역별 군집별 태풍피해자료를 이용하여 실시하였다. 주성분의 선택은 분산비율법에 의하여 누가분산비가 80%를 넘는 성분까지 취한다는 경험적인 기준(Choo et al, 2018b)이 있으나, 본 연구에서 몇 개

군집은 첫 번째 주성분의 공분산이 80%를 넘는 경우가 많아 90%를 넘는 주성분까지 모형개발에 이용하였다.

권역별로 군집별로 채택된 시군구를 대상으로 주성분을 선정하고, 다중회귀분석을 이용하여 태풍피해예측모형을 식 (6) 과 같이 개발하였으며, 그 결과가 Table 3과 Fig. 4.에 나타나 있다.

$$Y = \exp(y), y = \beta_0 + \beta_1 PC_1 + \dots + \beta_n PC_n \tag{6}$$

여기서 Y는 태풍피해액,  $y = \log(Y)$ ,  $\beta_i$ 는 다중회귀분석 계수,  $PC_i$ 는 주성분이다.

Table 3을 살펴보면 태풍피해예측모형에서 공분산이 90%을 넘는 주성분이 최대 3개까지 나타났다. 모형에서 경남 군집 2, 경북 군집 2를 제외한 나머지 군집들의 경우 태풍피해액과 산출된 주성분의 관계가 있는 것을 알 수 있으며, Fig. 4.에 군집별 관측피해액과 태풍피해예측모형에 의한 추정피해액의 관계가 나타나 있다.

경남 군집 2와 경북 군집 2에 속한 시군구는 태풍피해가 작아 일정한 경향을 나타내기에는 자료수 부족한 것으로 판단되므로, 관측치와 추정치의 반비례는 태풍피해 자료가 충분히 수집되면 변경될 것이다.

자연재해에 의한 피해를 예측하는 것은 어려운 일이며, 이와 관련된 연구에서도 호우를 제외하고는 만족할 만한 결과를 얻지 못하였다(Kim et al, 2016; Choo et al, 2018a; Choo et al, 2018b; Choo et al, 2017b)). 본 연구의 결정계수가 낮은

Table 3. Typhoon Damage Forecasting Function & NRMSE

Province	Cluster	$\beta_0$	$\beta_1$	$\beta_2$	$\beta_3$	$R_2$	NRMSE
Gyungnam	1	16.197	0.010	0.001	-	0.292	221.4%
	2	19.739	-0.006	-0.005	-	0.115	97.8%
Gyungbuk	1	16.052	0.010	0.011	-0.002	0.317	111.3%
	2	19.885	-0.002	-	-	0.635	249.6%
Jeonnam		15.672	0.006	0.011	-	0.325	132.7%
Jeonbuk		16.075	0.005	0.013	-	0.179	114.2%

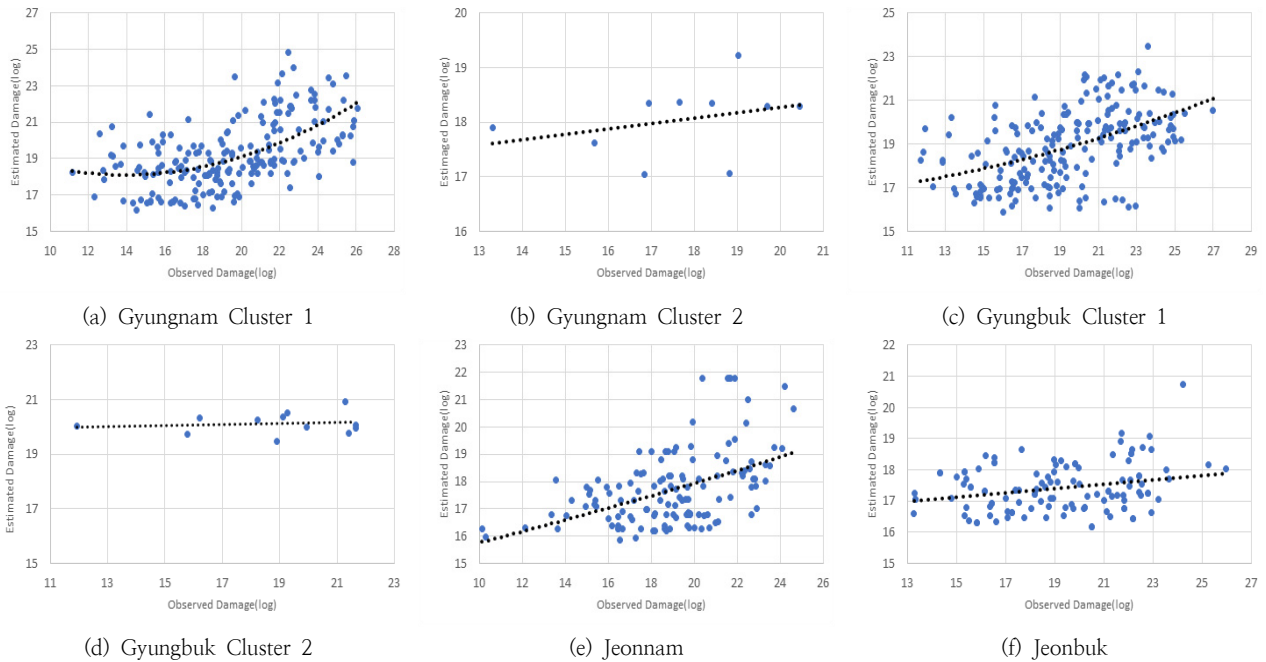


Fig. 4. Comparison between Observed & Estimated Damage

이유로 시군구별 부족한 자료와 강풍으로 인한 피해를 예측하기 어렵기 때문에 판단된다(Song and Lee, 2016). 또 다른 원인중 하나는 강우량 자료에 대한 정확성을 들 수 있다. 전술한 바와 같이 본 연구에서는 2005년부터 집중으로 설치된 AWS 자료를 이용하지 못하고, 자료의 일관성을 유지하기 위하여 69개 ASOS 관측자료를 이용함에 따라 시군구별 강우특성을 정확히 반영하지 못한 결과로 보인다.

## VI. 결 론

본 연구는 태풍내습시 가용한 태풍예보에 포함된 수문기상 자료를 이용하여, 태풍피해를 미리 예측하고 그에 대한 대비, 대응을 할 수 있는 태풍피해예측모형을 개발하였다. 태풍으로 인한 피해는 한반도의 남부지방인 부산, 울산, 대구, 광주, 경상남북도 및 전라남북도에 집중되고, 내륙에 위치한 지역보다는 연안에 위치한 지역이 더 많은 피해를 입은 것으로 나타났다. 태풍으로 인한 시설물 피해는 내륙에 위치한 시군구와 연안에 위치한 시군구에서 서로 다른 양상을 보인다. 내륙지역은 호우로 인한 피해와 비슷한 형태의 시설물 피해를 나타내었고, 연안지역은 강풍 및 파랑으로 인한 피해가 추가적으로 발생하는 것을 알 수 있었다.

본 연구에서는 내륙에 위치한 시군구에 대한 태풍피해예측모형을 개발하였다. 태풍피해예측모형은 태풍피해와 수문기상인자의 상관관계를 이용하여 시도별 군집화를 시도하였고, 군집별로 주성분분석을 이용한 주성분을 산정하고, 태풍피해를 결정변수로 하고 주성분을 설명변수로 하는 다중회귀분석을 실시하여 태풍피해예측모형을 개발하였다.

본 연구의 결론으로는 먼저, 남부 내륙에 위치한 시군구의 태풍피해회수는 연안지역에 위치한 시군구보다 작아 각 시군구별로 피해예측모형을 개발하기 어려웠으나, 군집화를 이용하여 자료를 확충하고 피해예측모형을 개발할 수 있었다. 두 번째로는 수문기상인자의 산정 과정에서 발생할 수 밖에 없는 설명인자들 간의 다중공선성 문제를 주성분분석을 통하여 해결할 수 있었다. 태풍피해예측모형에 의한 추정치와 실측치의 비교결과  $R^2$ 은 0.115~0.635, NRMSE는 97.8%~221.4%로 나타났다.

## 사 사

본 연구는 행정안전부 재난예측및저감연구개발사업의 지원을 받아 수행된 연구임 (MOIS-재난-2015-05)

## References

Ahn, S., Park, K., Kim, J., and Kim, B.(2015), The Characteristics of the Frequency and Damage for Meteorological Disasters in Korea, Journal of Korean Society of Hazard Mitigation, Vol. 15, No. 2, pp.133-144

Choi, C., Bak, J., Kim, J., Kim, D., and Kim, H.S.,(2016a) Development of Heavy Rain Damage Prediction Function Using Generalized Linear Model, 2016 Conference of Korean Society of Civil Engineers, pp.109-110, 201

Choi, C., Kim, J., Han, D., and Oh, S., and Kim, H.S.(2016b), Development of Heavy Rain Damage Prediction Function Using Mixed distribution, 2016 Conference of The Korean Society of Disaster Information, pp.236-237

- Choi, C., Kim, J., Kim, J., Kim, H., Lee, W., and Kim, H.S.(2017a), Development of Heavy Rain Damage Prediction Function Using Statistical Methodology, Journal of Korean Society of Hazard Mitigation, Vol. 17, No. 3, pp. 331-338
- Choi, C., Kim, J., Lee, M., Kim, J., Lee, W., and Kim, H.S.(2017b), Development of Heavy Rain Damage Prediction Function Using Principal Component Analysis and Logistic Regression Model, Journal of Korean Society of Hazard Mitigation, Vol. 17, No. 6, pp.159-166
- Choo, T.H., Cho, H.M., Shim, S.B., and Park, S.J.(2018a), Development of the Wind Wave Damage Estimation Functions based on Annual Disaster Reports : Focused on the Western Coastal Zone, The Journal of the Korea Contents Association. Vol. 18, No. 1, pp. 154-163
- Choo, T.H., Kim, Y.S., Sim, S.B., and Son, J.K.(2018b), Development of the Wind Wave Damage Predicting Functions in southern sea based on Annual Disaster Reports, Journal of the Korea Academia-Industrial cooperation Society, Vol. 19, No. 2, pp. 668 - 675
- Choo, T.H., Kwak, K.S., Ahn, S.H., Yang, D.U. and Son, J.K.(2017a), Development for the function of Wind wave Damage Estimation at the Western Coastal Zone based Disaster Statistics, Journal of the Korea Academia-Industrial cooperation Society, Vol.18, No.2, pp.14-22
- Choo, T.H., Yun, G.S., Kwon, T.B., Park, S.J., and Kim, S.R.(2017b), Proposal for Wind Wave Damage Cost Estimation at the Southern Coastal Zone based on Disaster Statistics, Journal of The Korea Contents Association, Vol.17, No.2, pp. 267-274
- 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, Journal of Korean Society of Hazard Mitigation, Vol. 8, No. 2, pp. 75~82.
- Kim, J., Choi, C., Lee, J., and Kim, H. S.(2017a), Damage Prediction Using Heavy Rain Risk Assessment: (2) Development of Heavy Rain Damage Prediction Function, Journal of Korean Society of Hazard Mitigation, Vol. 17, No. 2, pp.371-379
- Kim, J., Choi, C., Oh, S., Han, D., and Kim, H.S.(2016), Development of Heavy rain damage Prediction functions in Gyeonggi Province. 2016 Conference of The Korean Society of Disaster Information, pp.203-204, 2016
- Kim, J., Kim, T., and Lee, B.(2017b), An Analysis of Typhoon Damage Pattern Type and Development of Typhoon Damage Forecasting Function, Journal of Korean Society of Hazard Mitigation, 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, Y., Kim, T., and L, B.(2018), Development of Typhoon Damage Prediction Function using Tukey's Ladder of Power Transformation, J. Korean Soc. Hazard Mitig., Vol.18, No.6, pp.259-267
- 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
- Lee, B.J., Jung, I.W., and Bae, D.H.(2009), Parameter Regionalization of Semi-Distributed Runoff Model Using Multivariate Statistical Analysis, Journal of Korea Water Resources Association, Vol. 42, No. 2, pp.149-160
- 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. Journal of Korea Water Resources Association, Vol. 50, No. 8, pp. 513-520
- Lee, J.S., and Choi, H.I.(2016), Flood Vulnerability Index Estimated by Comparing Analysis Methods of Flood Damage Data, Journal of Korean Society of Hazard Mitigation, Vol. 16, No. 2, pp.427-435
- Lee, J., Eo, G., Choi, C., Jung, J., and Kim, H.(2015), Development of Rainfall-Flood Damage Estimation Function using Nonlinear Regression Equation, Journal of the Korea Society of Disaster Information, Vol.12, No.1, pp.74 - 88
- Lee, 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, S.W., Ahn, S.H., Lim, B. and Gwangyong Choi, G.(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
- Oh, Y.R., 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, 2013 Conference of Korean Society of Hazard Mitigation, p.40
- Park, S.W., Jun, B.H., and Jang, S.H.(2003) A Study on the Regionalization of Point Rainfall by Multivariate Analysis Technique, Journal of Korea Water Resources Association, Vol. 36, No. 5, pp.879-892
- Song, C.Y., and Lee, J.H.(2016), Development of Categorization System for Efficient Calculation of Damage Cost according to Strong Wind, Journal of the Korean Society of Safety, Vol. 31, No. 2, pp. 127-132



- Song, C.Y., and Yang, B.S.(2016), Gale Disaster Damage nvestigation Process Provement Plan according to Correlation Analysis between Wind Speed and Damage Cost, -Centering on Disaster Year Book-, Journal of the Korean Society of Safety Vol. 31, No. 2, pp. 119-126
- Yang, W.J., and Yi, W.H.(2013), Damage Pattern Analysis of Buildings by Wind Storm and Typhoon, 2013 Conference of Korean Society of Hazard Mitigation, p.198

**〈저자소개〉****Yonsoo Kim**

Risk Management Research Institute, LIG System Co., Ltd. Senior Researcher / [civil.engineer@hanmail.net](mailto:civil.engineer@hanmail.net)

**Taegyun Kim**

Gyeongnam Nattional Univ. of Science & Tech. Professor / [tgkim@gntech.ac.kr](mailto:tgkim@gntech.ac.kr)