확률론적 클러스터링 기법을 이용한 한반도 태풍경로 및 종관기후학적 분석

Assessment of Typhoon Trajectories and Synoptic Pattern Based on Probabilistic Cluster Analysis for the Typhoons Affecting the Korean Peninsula

김 태 정* / 권 현 한** / 김 기 영***

Kim, Tae-Jeong / Kwon, Hyun-Han / Kim, Ki-Young

Abstract

Lately, more frequent typhoons cause extensive flood and wind damage throughout the summer season. In this respect, this study aims to develop a probabilistic clustering model that uses both typhoon genesis location and trajectories. The proposed model was applied to the 197 typhoon events that made landfall in the Korean peninsula from 1951 to 2012. We evaluate the performance of the proposed clustering model through a simulation study based on synthetic typhoon trajectories. The seven distinguished clusters for typhoons affecting Korean peninsula were identified. It was found that most of typhoon genesis originated from a remote position ($10^{\circ} \sim 20^{\circ}\text{N}$, $120^{\circ} \sim 150^{\circ}\text{E}$) near the Equator. Cluster, type B can be regarded as a major track due to the fact that its frequency is approximately about 25.4% out of 197 events and its direct association with strong positive rainfall anomalies.

Keywords : typhoon track, clustering, synoptic pattern

요 지

최근 빈번하게 발생하는 태풍사상은 극심한 홍수 및 바람 재해를 유발 시키고 있다. 이러한 점에서 본 연구에서는 1951년부터 2012년까지 한반도에 내습한 총 197개의 태풍사상을 대상으로 태풍의 발생위치 및 태풍의 궤적을 기준으로 태풍을 범주화 할 수 있는 확률론적 클러스터링 기법을 개발하였다. 모의실험을 통하여 개발된 모형의 적합성을 확인할 수 있었으며, 태풍 경로에 적용이 가능한 방안으로 평가되었다. 1951년부터 2012년까지 한반도 내습한 197개의 태풍사상을 대상으로 확률론적 클러스터링 기법을 적용한 결과 한반도를 내습한 태풍사상은 총 7개의 클러스터로 분류되었으며, 대부분 위도 $10^{\circ} \sim 20^{\circ}\text{N}$, 경도 $120^{\circ} \sim 150^{\circ}\text{E}$ 해수면에서 발생하여 한반도를 향하여 진행하는 것으로 나타났다. 클러스터 B의 경우 약 25.4%의 발생빈도를 가지며, 전선의 방향도 한반도를 직접 향하고 있어 상대적으로 한반도에 영향이 가장 큰 클러스터로 분석되었으며 한반도 전체에 걸쳐서 강한 양(positive)의 강수량 Anomaly를 갖는 것을 확인할 수 있었다.

핵심용어 : 태풍경로, 클러스터링, 수문인자

* 전북대학교 토목공학과 석사과정, 방재연구센터 (e-mail: kim.t.j@jbnu.ac.kr)

Master Course, Department of Civil Engineering, Chonbuk National University, Jeonju, Korea

** 교신저자, 전북대학교 토목공학과 부교수, 방재연구센터 (e-mail: hkwon@jbnu.ac.kr)

Corresponding Author, Assistant Professor, Dept. of Civil Engineering, Chonbuk National University, Jeonju, Korea

*** 한국수자원공사 K-water 연구원 기반시설연구소 기반연구 2팀장 (e-mail: kky@kwat.or.kr)

General Manager, Infrastructure Research Center, K-water Institute

1. 서 론

최근 지구온난화로 인한 전지구적 기상변동성의 증가는 우리나라를 비롯한 동아시아 지역의 다양한 기상재해를 발생시키는 원인으로 작용하고 있다. 이러한 기상재해의 증가는 기후변화와 맞물려 더욱 증가할 것으로 전망되고 있으며, 한반도의 경우 극한강우의 발생빈도 및 양적 증가경향이 더욱 두드러지게 나타나고 있다. 특히, 지구온난화로 인해 동아시아 지역의 태풍의 발생빈도 및 강도가 과거에 비해 지속적으로 증가하는 경향성이 있으며, 문헌에서나 언급되었던 슈퍼태풍 ‘하이옌’이 발생하여 2013년 필리핀 지역을 포함하여 동아시아 지역의 막대한 사회·경제적 피해를 발생시켰다. 우리나라도 결코 슈퍼태풍의 안전지대가 될 수 없으며 지속적인 태풍 모니터링, 수문기상학적 연계 분석 등을 통한 효과적인 분석체계를 갖출 필요가 있다.

태풍은 적도 부근이 극지방에 비하여 태양열을 더 많이 받기 때문에 생기는 열적 불균형을 없애기 위해 저위도 지방의 따뜻한 공기가 바다로부터 수증기를 공급받으면서 강한 바람과 많은 강우를 동반하며 고위도로 북상하는 기상현상을 의미하는 것으로 발생 장소에 따라 명칭을 각각 달리하고 있다. 북태평양 서부에서 발생하는 것을 태풍(Typhoon), 북대서양·카리브해·멕시코만 및 북태평양 동부에서 발생하는 것을 허리케인(Hurricane), 인도양·아리비아해 그리고 벵골만에서 발생하는 것을 사이클론(Cyclone)이라고 부른다. 세계기상기구(World Meteorological Organization, WMO)에서는 태풍 중심부근의 최대풍속(Maximum Sustained Wind Speed, MSWS)에 따라 TD (Tropical Depression, $MSWS < 17 \text{ m/s}$), TS (Tropical Storm, $17 \text{ m/s} < MSWS < 25 \text{ m/s}$), TY (Typhoon, $MSWS \geq 33 \text{ m/s}$)로 분류하고 있으며, TS 등급 이상부터 태풍의 발생 번호와 이름이 부여된다.

적도지방에서 발달한 열대성저기압인 태풍은 강풍과 호우를 유발시킴으로써 자연재해를 불러일으킨다. 최근 통계를 보면 기상재해의 피해원인은 호우(37%), 호우·태풍(22%), 태풍(15%)의 순으로 나타났으며, 기상재해의 경제적 피해규모 또한 호우, 태풍의 순으로 나타나는 등 대부분의 기상재해가 여름 및 가을철에 발생하는 태풍에 기인한다(Lee, 2001). 소방방재청에 의하면 태풍에 의한 가장 큰 피해액은 2002년 태풍 ‘루사’로 5조 1,480억 원의 피해가 발생하였으며, 2003년 태풍 ‘매미’의 피해액도 4조 2,220억 원에 이르렀다. 또한, 2007년 9월에도 제주도와 고흥지역을 강타한 태풍 ‘나리’가 13명의 인명피해와 천억여 원의 재산피해를 내는 등 우리나라에는 매년 태풍에 의한

많은 인명피해와 재산피해가 발생하고 있다. 이렇게 태풍으로 인한 재해는 꾸준히 증가하고 있으며 태풍에 의한 최대강우량도 과거에 비해 증가하였다(Oh et al., 2008).

국내의 태풍관련 연구는 주로 태풍에 의해 발생하는 강우사상의 특성 등에 주안점을 주고 수행되어왔다. Lee et al. (1992)은 1960년부터 1989년까지 30년간 동아시아 중위도에 접근하면서 한반도에 영향을 미친 태풍에 대하여 중심 기압과 지상 최대풍속(maximum wind velocity)을 통계적으로 분석하고 태풍경로 특성에 따른 태풍을 분류하였다. Back et al. (1999)은 1945년부터 1999년까지의 53개년의 자료를 이용하여 태풍에 의한 한반도의 기후학적 특성을 분석하였다. 또한 최대풍속이 50 m/s 를 상회하는 태풍은 8월말에서 9월초에 주로 발생하는 것으로 분석되었으며, 또한 1973년 이후에 발생하고 있는 태풍의 강도와 태평양 해수면 온도와와의 연관성을 분석하였다. Yoo and Jung (2000)은 2000년 한반도에 영향을 미친 태풍 ‘사오마이’와 ‘쁘라삐룬’의 피해상황과 과거 유사한 경로를 가지는 태풍들과 비교하였으며, 이외에도 Park et al. (2006)은 한반도에 영향을 미치는 태풍을 정의하고, 기후변화에 따른 태풍의 발생 빈도 및 발생지역의 변화 양상을 평가하였다. 또한 한반도에 영향을 미친 태풍을 진로에 따라 유형을 구분하고 각 유형별로 태풍의 강도와 한반도에 영향을 미치는 정도를 파악하여 태풍의 진로와 강도에 따른 태풍 특성을 분석하였다. Oh et al. (2007)은 우리나라에 발생하는 강우사상을 태풍과 집중호우로 구분하여 호우 원인별로 지속시간별 연 최대 강우량을 구축한 후, 최종적으로 태풍과 집중호우에 따른 강우의 기본적인 특성과 더불어 확률 강우량을 산정하여 비교 검토하였다. 최근 연구로서 태풍의 경로에 따른 집중호우의 특성 분석이 필요성을 인지하고 비매개변수적 모의발생기법과 지역가중다항식을 이용하여 태풍의 확률강우량을 산정하였다(Oh et al., 2009).

국외의 경우에는 태풍으로 인한 해일과 연안침식 등을 평가하는 Empirical Simulation Technique 기법이 개발되었으며(Scheffner et al., 1996) 기후변화로 인해 태풍과 동반되는 강우량이 증가할 가능성이 매우 큰 것으로 나타났다(Knutson et al., 2010; Emanuel et al., 2008). Dimego and Bosart (1982)는 1972년 Agens 사례를 통해 태풍이 중위도에 접근하면서 기상학적변동성으로 인해 태풍 특성이 변화하는 양상으로 분석하고 고찰하였다.

앞서 살펴본바와 같이 국내의 태풍관련 연구는 태풍의 발생빈도의 경향성분석 및 태풍과 일반호우를 분리하고 확률강우량을 비교하는 연구 등이 주로 이루어지고 있다. 그러나 태풍의 발생위치, 경로, 이동시간 등을 정량적으로

해석함과 동시에 이들 태풍의 주요 특성과 연계한 수문 기상학 및 종관기후학적 연구는 상대적으로 미흡한 실정이다. 이러한 점에서 본 연구에서는 태풍의 특성을 효과적으로 고려할 수 있는 확률론적 클러스터링 기법을 기반으로 한반도에 영향을 미친 태풍의 대표경로를 추적하였다. 도출된 태풍의 대표경로를 기준으로 태풍의 발생위치를 분리한 후 각 경로에 따른 수문기상학적, 종관기후학적 분석을 실시하였다.

본 논문의 구성은 다음과 같다. 1장에서는 논문의 배경 및 목적에 대해서 언급하였으며, 2장에서는 대상자료 및 확률론적 클러스터링 기법에 대하여 기술하였다. 3장에서는 한반도에 내습한 태풍의 경로를 클러스터링 한 결과를 요약하여 나타내었으며 마지막으로 결론 및 토의를 4장에 수록하였다.

2. 대상자료 및 분석방법

2.1 대상자료

태풍은 열대 해상에서 발생하는 전선을 갖지 않는 대류권내 저기압성 순환을 열대저기압으로 총칭하는 것으로 북태평양 서부에서 발생하는 열대저기압 중에서 중심 부근의 최대풍속이 17 m/s 이상의 강한 폭풍우를 동반하고 있는 것을 말하는데, 폭풍우는 반드시 태풍에만 동반되는 것이 아니고 온대 저기압에서도 발생하는 경우가 많다. 태풍백서(기상청, 2011)에서 한반도 영향 태풍은 32°~40°N, 120°~135°E 의 영역을 통과하는 태풍으로 정의하고 있으며 본 연구에서도 동일한 정의를 사용 하였다.

본 연구에서는 한국 기상청(KMA)과 일본 기상청(JMA)에서 제공하고 있는 태풍자료를 사용하였으며, 일본 기상청 산하의 지역특별기상센터(Regional Specialized meteorological Centers-Tokyo Typhoon Center)에서 제공하는 자료는 태풍의 이름, 위·경도 위치, 최소 중심기압, 중심 최대풍속 등의 변수가 6시간 간격으로 이루어져 있다. 우리나라의 근대적 기상관측이 개시된 이래 최근까지 109년 동안(1904~2012년) 우리나라에 영향을 미친 크고 작은 태풍의 수는 총 335개이다. 우리나라는 한해에 약 3.1개 정도의 태풍이 우리나라에 영향을 미치며 태풍내습의 최대월은 8월, 7월, 9월 순으로 7월과 8월 두 달 동안에 내습한

태풍의 수는 전체의 66%를 차지하고 있다. 그러나 최근에 지구온난화로 인한 기상변동성 증가로 인하여 5월, 6월 및 10월에도 태풍이 내습하는 사례가 발생하고 있다.

우리나라는 매년 태풍이 내습하여 극심한 집중호우와 강풍을 유발시킴으로써 막대한 사회·경제적 피해를 발생시키고 있다. 한반도에 영향을 미치는 태풍은 남해안을 통과하는 경로의 빈도가 가장 많으며 최근에는 서해안을 통과하는 태풍 발생으로 인해 과거와 다른 형태의 피해가 초래되고 있다. 특히, 최근 홍수사상의 추세를 살펴보면 과거에 비해 그 특성이 변화하고 있을 뿐만 아니라 불확실성이 증가하고 있는 추세이며, 지역적인 강우특성의 차이 또한 두드러지게 나타나고 있다.

Figs. 1 and 2는 1951년부터 2012년까지 우리나라에 영향을 미친 197개 태풍사상의 발생횟수를 적도지역의 태풍 발생 개수와 비교하여 도시한 결과로 우리나라에 영향을 끼친 태풍의 개수와 적도지역에서 발생한 태풍의 개수 사이의 상관관계수는 0.122로 유의한 상관성은 없는 것으로 판단되었다. 따라서 적도지역에서 발생한 태풍 중에서 우리나라에 영향을 미친 태풍만을 분류하여 분석을 실시하였다. 적도지역에서 발생한 태풍 중 1951년에서 2012년까지 우리나라에 영향을 끼친 태풍은 총 197개로 연평균 3.21개이며, 연도별 최대값은 7개로 1959년에 가장 많은 태풍이 우리나라에 영향을 미친 것을 확인하였다.

Fig. 1. No. of Typhoons Affecting the Korea Peninsular and Total No. of Typhoons at Equator during 1951~2012 Years

Table 1. Effected Typhoon in Korea during 1904~2012

Month	1	2	3	4	5	6	7	8	9	10	11	12	Total	Mean
Count	-	-	-	-	2	19	97	125	84	8	-	-	335	3.1

Fig. 2. Monthly Frequency of No. of Typhoons during 1951~2012 Years

2.2 분석방법

2.2.1 클러스터링 알고리즘

클러스터링(clustering method)이란 유사한 특징을 보이는 것들을 집단화하는 작업으로 여러 공학 분야에서 사용되고 있으며, 주어진 자료에서 유사한 속성, 형태 등의 기준을 통하여 다차원의 자료를 소규모의 그룹으로 범주화시키는데 활용된다. Jeong and Bae (2004)는 월 태풍입량을 예측하기 위하여 퍼지 클러스터링 방법 중 하나인 차감 클러스터링 기반의 태풍 유입량 예측기법을 도입하였다. Kim et al. (2011)은 퍼지 *c*-means 클러스터링 기법(FCM)을 1965년부터 2006년까지 북서태평양에서 발생한 855개의 태풍사상에 대하여 7개의 클러스터를 사용하여 분류하였으며, 각 클러스터는 태풍의 수명, 이동거리, 강도, 계절적 변화와 태풍으로 인한 강우량의 분포를 평가하였다. Kim et al. (2013)은 조건부 활성화함수 기반의 퍼지 *c*-means 클러스터링 기법과 신경회로망 패턴분류기를 활용하여 여름철 초단기 강수예측 패턴모델에 적용하였다. Lee and Kwon (2011)은 *k*-means 방법의 속성정보로서 관측 지점별 순환통계량(circular statistics)과 극치강수량을 활용하여 전국을 다섯 개의 군집으로 분류하였으며 최종적으로 지역빈도해석을 위한 군집분석에 활용하였다. Kim and Lee (2008)는 *k*-means 알고리즘을 이용한 저수량(low flow)에 대한 수문학적 동질유역을 분류하고, Bayesian 다중회귀분석을 이용한 저수량 지역 빈도분석을 수행하였다.

클러스터링 방법에는 대표적으로 *k*-Means Clustering, 퍼지 *k*-Means Clustering, Mountain Clustering, Subtractive Clustering 방법 등이 있다. 태풍경로의 분류를 위해서 *k*-means분석은 여러 연구에 적용된 사례가 있다. Elsner and Liu (2003)는 북서태평양의 태풍경로에 적용

하였으며 Elsner (2003)는 북대서양의 허리케인 경로를 범주화시키기 위해서 적용되었다. 태풍경로를 범주화시키기 위한 속성정보로서 위의 두 연구는 허리케인의 최대강도의 위치정보와 허리케인이 종료되는 지점의 위치정보를 활용하였으며 3개의 대표경로를 추출하였다. Blender et al. (1997) 또한 *k*-means 방법을 기초로 6시간마다 기록된 허리케인의 위도와 경도 정보를 활용하여 대표경로를 범주화하였다. 앞서 언급되었듯이, *k*-means 방법은 각각의 태풍 경로를 *K*개의 그룹으로 범주화 하는데 가장 일반적으로 적용되는 방법으로서 그룹들 사이의 총 분산(total variance)을 최소화함으로 최적의 범주화 개수를 찾는 과정이다. 그러나 태풍마다 지속시간의 상이함으로 인해 서로 다른 길이를 가지며 이런 경우 *k*-means방법을 적용하는데 한계가 있다. 이러한 점에서 Harr and Elsberry (1995)는 다른 태풍 경로의 특성과 연계된 공간적 패턴을 평가하기 위해서 퍼지 범주화 방법과 경험적 직교함수(empirical orthogonal functions, EOF)를 활용하였다.

본 연구에서는 태풍의 경로를 효과적으로 묘사하기 위해서 유한혼합(finite mixture)모형을 적용하였으며 다른 길이를 가지는 태풍의 경로를 범주화할 수 있는 확률론적(probabilistic) Curve-Aligned Clustering (PCAC) 모형을 도입하였다. 이러한 점에서 PCAC기법은 기존 *k*-means 방법의 단점을 보완할 수 있는 방안으로 적용이 가능하다. 특히, PCAC기법은 상대적으로 적은 태풍의 경로정보로 인한 단점을 보완할 수 있을 뿐만 아니라 경로의 모양과 위치정보를 동시에 고려할 수 있는 장점이 있다.

PCAC 기법은 적용은 크게 4단계로 이루어진다. 첫째, 클러스터의 개수를 선택한다. 둘째, 태풍의 경로를 묘사하기 위한 회귀다항식의 형태를 결정한다. 셋째, 결정된 회귀다항식의 차수(order)를 결정한다. 마지막 단계로 클러스터링이 이루어진다. 따라서 태풍 경로를 효과적으로 분류하기 위해서 최적의 클러스터의 수, 회귀모델의 형태, 다항식의 차수 등이 정량적인 기준으로 결정되어야 한다(Gaffney, 2004). PCAC기법의 세부적인 내용은 다음과 같다.

우선 관측된 태풍 자료로부터 *i*번째 태풍의 위도와 경도를 나타내는 $n_i \times 2$ 형태로 구성된 Z_i 경로 자료를 구축하여 위도와 경로에 대하여 *p*차의 다항회귀 분석을 실시한다. 아래의 Eq. (1)은 *i*번째의 *k*번째 클러스터에 의한 태풍 경로의 다항회귀식이다.

$$z_i = T_i \beta_k + \epsilon_i, \quad \epsilon_i \sim N(0, \Sigma_k) \quad (1)$$

위 식에서 T_i 는 시간 t_i 와 관련된 $n_i \times (p+1)$ 반데몬드

(vandermonde) 회귀행렬을 나타내며, 시간 t_i 에 상응하는 $(p+1)$ 열로 정의된다. 따라서 m 번째 열에서 시간 t_i 의 요소들은 m ($0 \leq m \leq p$)의 거듭제곱형태를 가지게 된다. β_k 는 $(p+1) \times 2$ 의 구조를 가지는 클러스터 k 에 대한 회귀계수로 이중에서 첫 열은 경도에 대한 회귀계수이고 두 번째 열은 위도에 대한 회귀계수를 나타낸다. 또한 ϵ_i 는 $n_i \times 2$ 의 평균이 0이고 2×2 의 공분산 행렬 Σ_k 를 가지는 다변량 정규분포의 잡음(noise)을 나타낸다. 여기서 공분산 행렬 Σ_k 는 관측된 태풍의 위도와 경도의 잡음 분산(noise variance)을 나타내는 σ_{k1}^2 과 σ_{k2}^2 의 대각행렬을 가지며, 편의상 교차 공분산행렬은 항상 0으로 가정한다. 클러스터 k 에 속하는 i 번째 태풍사상의 조건부함수(conditional density function)는 아래의 Eq. (2)와 같다.

$$p(Z_i|t_i, \theta_k) = f(Z_i|T_i\beta_k, \Sigma_k) \\ = (2\pi)^{-n/2} |\Sigma_k|^{-n/2} \exp\left\{-\frac{1}{2}tr[(Z_i - T_i\beta_k)\Sigma_k^{-1}(Z_i - T_i\beta_k)']\right\} \quad (2)$$

여기서, $\theta_k = \{\beta_k, \Sigma_k\}$ 이다.

위 식은 k 개의 클러스터를 가지는 혼합회귀(regression mixture)모델의 형태로 아래의 Eq. (3)과 같이 나타낼 수 있다.

$$p(Z_i|t_i, \phi) = \sum_k^K \alpha_k p_k(Z_i|t_i, \theta_k) = \sum_k^K \alpha_k f_k(Z_i|T_i\beta_k, \Sigma_k) \quad (3)$$

여기서, α_k 는 클러스터 k 에 속하는 태풍경로에 대한 확률을 나타내며, ϕ 는 혼합분포의 매개변수집합 $\{\beta_k, \Sigma_k, \alpha_k; 1 \leq k \leq K\}$ 을 나타낸다.

n 개의 관측시간에 따른 태풍경로로 이루어진 $Z = \{z_1, \dots, z_n\}$ 와 $T = \{t_1, \dots, t_n\}$ 에 대한 확률밀도함수(probability density function, PDF)를 사용하여 우도(likelihood)를 아래의 Eq. (4)와 같이 나타낼 수 있다.

$$p(Z|T, \phi) = \prod_i^n \sum_k^K \alpha_k f_k(z_i|T_i\beta_k, \Sigma_k) \quad (4)$$

2.2.2 Expectation-Maximization 알고리즘

클러스터링을 수행함에 있어 주어진 자료로부터 우도를 값을 최대화하는 매개변수를 추정하는 방법으로 본 연구에서는 Expectation-Maximization (EM) 알고리즘을 활용하였다. EM 알고리즘은 자료들이 최적의 클러스터의 정보를 도출하기 위해서 초기값을 가정하여 얻은 해를 일차적으로 추정한다. 이렇게 추정된 값을 초기값으로 다시 해를 재추정하는 과정을 반복하면서 해가 수렴할 때까지 변가아 수행하는 방법이다.

주어진 태풍경로들 중에서 클러스터 k 와 가장 연계되는 태풍의 경로를 추출하기 위하여 EM 알고리즘이 적용되며 Expectation 단계에서는 각 클러스터 k 의 속할 확률(membership probability)을 계산한다.

$$w_{ik} = \frac{\alpha_k f_k(z_i|T_i\beta_k, \Sigma_k)}{\sum_j^K \alpha_j f_j(z_i|T_i\beta_j, \Sigma_j)} \quad (5)$$

i 번째에 해당하는 태풍의 경로에 대한 w_{ik} 는 클러스터 k 에 속하는 것으로 w_{ik} 에 속하는 모든 태풍의 사상은 동일한 우도를 가진다. 또한 w_{ik} 를 원소로 가지는 대각행렬 W_k 는 $N \times N$ 의 구조를 가진다. Maximization 단계에서는 W_k 를 사용하여 혼합(mixture) 매개변수를 계산한다.

$$\hat{\beta}_k = (X' W_k X)^{-1} X' W_k Y \quad (6a)$$

$$\hat{\Sigma}_k = \frac{(Y - X\hat{\beta}_k)' W_k (Y - X\hat{\beta}_k)}{\sum_i^n w_{ik}} \quad (6b)$$

$$\hat{\alpha}_k = \frac{1}{n} \sum_i^n w_{ik} \quad (6c)$$

여기서, Y 는 $N \times 2$ 의 구조인 연결(concatenated) 행렬 $(y_1', \dots, y_n)'$ 이며, X 또한 $N \times (p+1)$ 구조의 연결 행렬로 $(X_1', \dots, X_n)'$ 에 대한 회귀행렬이다. 가중회귀분석을 통하여 W_k 의 원소는 가중치로 표현되어 Y 와 X 에 회귀분석에 적용된다. Expectation 단계와 Maximization 단계는 수렴할 때까지 반복적으로 수행된다.

2.2.3 최적 클러스터의 선택

본 연구에서는 태풍의 경로를 산정하는데 태풍의 중심 위치가 관측된 좌표를 회귀분석기법을 적용하여 태풍의 경로를 산정하였다. 본 연구에서 적용한 PCAC 기법은 혼합회귀다항식(mixture polynomial regression)의 차수를 결정한 뒤 태풍의 경로를 분석하게 된다. 혼합회귀다항식의 최적 차수와 클러스터 수를 산정하기 위해서 우도함수가 최대가 되는 조합을 결정하는 과정이 필요하다. 그러나 우도가 최대화됨에 따라서 모델의 매개변수가 무한히 증가할 수 있다. 이렇게 매개변수가 무한히 증가하는 경우는 모델이 과적합(overfitting)되는 문제점이 있으므로(Kwon et al., 2013), 본 연구에서는 모형 선택 지표로 Akaike Information Criterion (AIC)를 사용하였으며 회귀분석 최적차수와 클러스터의 개수를 결정하였다.

AIC는 선형과 비선형 모형의 식별(model identification)에 주로 사용되는 기준으로 AIC는 모형의 적합도

(model fitting)나 모형 최적화(model optimization)가 균형을 이루는 합리적인 기준이라고 볼 수 있다(Faraway and Chatfield, 1998; Yang and Zou, 2004). AIC는 우도값이 모형내에서의 비교만 가능하게 하다는 단점을 보완한 것으로 모형의 적합(fitting) 정도에 대한 비교지표로 사용되고 있다. AIC의 일반적인 형식은 아래의 Eq. (7)과 같다.

$$AIC = 2k - 2\ln(L) \quad (7)$$

식에서 k 는 모형의 매개변수 개수를 나타내며 L 은 우도값을 의미하는 것으로 모형이 복잡해지고 과적합 될수록 k 값이 증가하므로 AIC값은 증가하게 되므로 모형의 과적합에 대한 벌점(penalty)을 부여한다. 따라서 모형의 매개변수를 최소화하고 우도값을 최대화 하면 AIC는 최소값을 가지게 되는데 AIC가 최소화 될수록 모형의 최적화를 의미한다. 따라서 본 연구에서는 모형의 매개변수의 개수와 우도값으로 산정되는 AIC를 사용하여 모형의 최적 클러스터의 개수를 결정하는 최적화 과정을 수행하였다.

3. 적 용

3.1 모의실험

모의실험을 통하여 PCAC 클러스터링 기법의 적용성을 평가하였다. 우선 1차부터 3차까지의 차수를 가지는 다항식을 임의의 발생한 후 다항식의 계수에 일정한 난수를 가산하여 각 차수별로 50개의 변형된 다항식을 생성하였다. Fig. 3은 임의의 생성된 다항식을 나타낸 것으로 1차 다항식부터 3차 다항식까지 총 150개의 다항식을 나타내고 있다. Fig. 3은 PCAC 클러스터링 기법을 사용하여 클러스터링 한 결과로 기존의 임의의 발생한 다항식을 효과적으로 분류하였다. Fig. 3의 검은색 실선은 초기에 임의의 발생

을 위한 기본다항식을 나타내며 각 차수에 해당하는 다항식이 효과적으로 분류되고 있음을 확인할 수 있었다.

3.2 한반도 태풍경로의 적용

1951년부터 2012년까지 한반도를 내습한 총 197개의 태풍사상의 경로를 대상으로 PCAC 클러스터링 기법을 적용하였다. 회귀분석을 실행함에 있어서 고차다항식 및 Spline회귀방정식의 경우 다수의 매개변수로 인해 과적합이 발생할 개연성이 크다. 이러한 점에서 3차이상의 고차다항식 적용이 가능하지만 상대적으로 단순한 2차의 다항식이 태풍경로의 적합성(goodness of fit)과 해석학적 편의성의 관점에서 가장 적합할 것으로 판단된다(Camargo et al., 2007). 모형의 적용에 앞서 클러스터의 수를 결정하기 위하여 대수우도함수(log-likelihood function)와 AIC를 산정하였으며, 대수우도함수가 클수록 또는 AIC가 작을수록 모형의 적합성이 우수함을 나타낸다. 클러스터 수 K 를 1~10까지 변화하면서 각 클러스터에 해당하는 대수우도함수와 AIC를 계산하여 Fig. 4에 나타내었으며 대수우도함수와 AIC가 수렴하기 시작하는 지점에서 최적의 클러스터 수를 결정하였다. 즉, 클러스터의 수 K 를 4로 증가시키면 대수우도가 상대적으로 큰 증가를 보이며, $K=5$ 에서부터 수렴하는 특성을 보이고 있다. 특히, 클러스터의 개수 $K=7$ 을 중심으로 대수우도와 AIC가 큰 차이를 보이지 않고 빠르게 수렴하고 있으며 본 연구에서는 클러스터의 개수를 7개로 적용하여 태풍경로를 분석하였다.

Fig. 5는 한반도에 영향을 미친 총 197개의 태풍에 대해서 클러스터 A~G까지 총 7개의 클러스터에 해당하는 태풍의 경로 및 발생위치를 나타낸다. 태풍의 발생빈도 및 경로는 적도 부근 해수면 압력특성(sea level pressure, SLP)과 대륙의 지오폠펜설고도(geopotential height)와 같

Fig. 3. Synthetic Trajectories Based on Polynomial Regression and PCAC Clustering Results

Fig. 4. Log-likelihood Function and AIC Value Given No. of Typhoon Clusters

은 기상학적 역학특성에 크게 영향을 받는다. 이러한 점으로 인하여 한반도에 내습하는 대부분의 태풍은 위도 $10^{\circ} \sim 20^{\circ}\text{N}$, 경도 $120^{\circ} \sim 165^{\circ}\text{E}$ 해수면에서 발생하여 한반도로 진행되는 것을 확인하였다. 이렇게 발생한 태풍은 해수면으로부터 수증기를 공급받으면서 북상하며 한반도 남서쪽 해안으로 내습하여 북동쪽 태백산맥을 통과하면서 소멸되는 특성을 나타내고 있다. 붉은색 굵은 선으로 표시된 경로는 각 클러스터에 해당하는 대표경로로서 대부분의 태풍은 위도 30°N 부근에서 선회하여 북서쪽으로 상향 진행되는 것으로 나타났다.

각 클러스터에 속하는 태풍의 수는 Table 2와 같다. 한반도에 내습하는 태풍의 경로 중 가장 많은 특성을 나타낸 클러스터는 클러스터 B인 것을 확인하였다. 클러스터 B의 태풍은 위도 15°N , 경도 135°E 해수면 근처에서 발생하여 다른 클러스터와는 달리 한반도를 향하여 북쪽으로 수직 상향 하면서 기압변화가 이루어지는 위도 30°N 부근에서 북동쪽으로 선회하여 한반도에 영향을 미치는 것으로 확인되었다. 이러한 클러스터 B에 해당하는 대표적 태풍사상으로는 현재까지 인명피해 순위 4위인 ‘베티’(1972년)가 있으며 재산피해액을 기준으로는 ‘올가’(1999년, 3위) ‘재니스’(1995년, 5위) ‘예니’(1998년, 8위) ‘메기’(2004년, 10위)가 있다. 이 중에서 1998년 발생한 ‘예니’는 1998년 9월 30일에 포항에 일 최대강수량 516.4mm 를 발생시켰으며 2004년에 발생한 ‘메기’는 8월 18일과 19일 이틀 동안 한반도에 상륙하여 속초와 강릉시에 각각 244.5mm 와 280.0mm 의 강우량을 기록하였다. 이 밖에 최근에 발생한 태풍으로는 ‘텐무’와 ‘곤파스’(2010년) 그리고 ‘메아리’(2011년) 등이 클러스터 B에 속한다.

Table 2. No. of Typhoon for each Cluster

Cluster Type	Number of Typhoon
A	39
B	50
C	27
D	16
E	17
F	16
G	32

태풍이 발생된 시점의 중관기후학적 패턴을 알아보기 위하여 NCEP/NCAR (National Centers for Environmental Prediction/National Center for Atmospheric Research) 재해석(reanalysis) 자료를 기본적으로 활용하였다. 재해석 자료는 전지구모형(global circulation model, GCM)에 관측값을 입력 자료로 활용하여(forcing) 전지구적인 기상학적 거동을 평가한 자료이다. NOAA에서 제공하는 NCEP 자료 중 x방향 바람벡터(u-wind vector), y방향 바람벡터(v-wind vector) 자료와 850Pa 해수면기압 SLP의 총 3가지 기상인자 자료를 이용하여 각 클러스터 별로 중관기후학적 패턴을 도출하였다.

Fig. 6은 각 클러스터에 해당하는 태풍의 발생일자별 기준으로 작성된 기상합성(climate composite) 분석 결과를 나타낸다. 기상합성분석은 예를들어 태풍 발생일이 2012년 7월 1일과 8월 3일이라면 이날들의 SLP와 바람장벡터를 추출하고 각 격자별로 평균하여 기상장(climate

Fig. 5. Typhoon Trajectories and Derived Mean Tracks Associated with Different Clusters

pattern)을 표시하는 방법으로서 종관기상학적 특성을 파악하는데 유용하다. 본 연구에서 사용된 기상자료는 모두 Anomaly값으로서 1981~2010년의 장기간 평균값(long-term climatology)을 기준으로 계산된 값이다.

기상합성 분석 결과 태풍의 발생위치에는 모든 클러스터에서 강한 저기압이 위치하고 있는 것을 확인할 수 있었으며, 반대로 저기압 상단에는 강한 고기압이 위치하고 있어 전선이 형성되고 있음을 확인할 수 있다. Fig. 6의

Fig. 6. A Map Showing Distribution SLP and Wind Vector of Cluster

(H)에서 확인 할 수 있듯 태풍의 발생지점일 한반도의 동쪽에는 고기압 Anomaly가 위치해 있다. 이러한 한반도 우측에 위치는 고기압 Anomaly로부터 불어오는 북서풍의 영향을 받아 북태평양에서 발생한 태풍을 한반도로 이동시키는 지향류(steering flow)의 역할을 하게 됨으로써

북서태평양에서 발생한 태풍이 해수면에서 다량의 수증기를 포함하여 북상한 후 한반도에 상륙할 수 있게 된다. 가장 빈도수가 높은 클러스터 B의 경우 한반도에 가장 근접하여 태풍이 발생하고 있으며 저기압의 강도 또한 다른 클러스터의 비해 매우 큰 특징을 가지고 있다. 또한 강

Fig. 7. A Map Showing Distribution Precipitation Rate (mm/day) and Mean Track of Cluster B

Table. 3 Maximum Rainfall Ranking through the Typhoon

Rank	Cluster Type	Typhoon Name	Station Name	Maximum Rainfall (mm)	Observation Date
1	C	RUSA	Gangneung	870.5	2002. 8.31
2	A	AGNES	Jangheung	547.4	1981. 9. 2
3	B	YANNI	Pohang	516.4	1998. 9.30
4	D	GLADYS	Busan	439.0	1991. 8.23
5	G	NARI	Jeju	420.0	2007. 9.16
6	F	MAEMI	Namhae	410.0	2003. 9.12
7	B	BETTY	Haenam	407.5	1972. 8.20
8	C	OLIVE	Samcheok	390.8	1971. 8. 5
9	B	OLGA	Dongducheon	377.5	1999. 8. 1
10	A	JANIS	Boryeong	361.5	1995. 8.25

한 고기압대와 블록을 형성하고 있으며 전선의 방향도 한반도를 직접 향하고 있어 상대적으로 한반도에 영향이 가장 큰 경로로 분석되었다. 클러스터 A의 경우도 한반도의 근접하여 태풍이 발생하고 있으며 제주도 먼바다 및 대만 근처에서 전선이 형성되어 한반도 서쪽으로 크게 선회하는 경로를 가지는 특징이 있으며 주로 강원도 지역에 큰 수해를 발생시킨 태풍사상 ‘루사’ 및 ‘올리브’가 이에 해당한다. 세 번째로 빈도가 높은 클러스터 G의 경우 클러스터 B와 유사한 경로를 가지고 있으나 발생위치에서 판단해보면 클러스터 G에 비해 평균적으로 오른쪽으로 치우쳐 발생하고 있으며 종관기후학적으로 클러스터 B에 비해 저기압강도가

가 상대적으로 약한 점을 Fig. 6에서 확인할 수 있다.

Fig. 7은 한반도를 가장 많이 내습한 클러스터 B에 대하여 태풍 발생일로부터 5일 이후의 강수량의 Anomaly 값의 분포를 나타낸 것으로 태풍의 중심을 기준으로 왼쪽 부근에 강수량이 집중적으로 분포되어 있는 것을 확인할 수 있다. 또한 한반도 통과시에 최대강수량을 기록한 태풍사상을 정리한 결과 클러스터 B에 해당하는 태풍사상이 가장 많은 것으로 나타났다. 이러한 이유는 태풍이 북서쪽으로 진행함에 따라서 반시계방향으로 회전하면 해수면의 수증기를 흡수하여 육지에 도달하는 경우 기압의 변화로 인하여 상대적으로 많은 강수량을 유발시키는 것

과 연관이 있는 것으로 사료된다.

4. 결 론

태풍은 사회·경제적으로 막대한 피해를 주는 기상현상 중 하나로 태풍에 의한 재산피해 및 인명피해가 해마다 발생하고 있다. 태풍에 의한 막대한 피해를 줄이기 위해서는 정확한 태풍의 경로에 대한 분석이 필수적이다. 본 연구에서는 KMA와 JMA에서 제공하고 있는 시간에 따른 태풍경로 자료를 기반으로 확률론적 클러스터링 기법을 적용하여 한반도에 영향을 미친 태풍의 경로의 특성을 분석하였다. 각 클러스터에 해당하는 태풍의 발생위치 및 태풍의 경로를 검증하고 NCEP/NCAR의 재분석 자료를 사용하여 클러스터의 종관기후패턴을 분석하였다. 본 연구를 통하여 도출한 결론은 다음과 같다.

- 1) 본 연구에서는 서로 다른 길이를 가지는 태풍사상의 경로 특성을 고려하여 태풍을 범주화할 수 있는 확률론적 모형을 도입하였다. 본 연구에서 적용한 PCAC 방법은 상대적으로 적은 태풍의 경로정보를 고려할 수 있을 뿐만 아니라 경로의 모양과 위치정보를 동시에 고려할 수 있는 장점을 확인할 수 있었다. 이와 더불어 모의실험을 통하여 모형의 적합성을 확인할 수 있었으며, 태풍 경로에 적용이 가능한 방안으로 평가되었다.
- 2) 한반도에 내습한 태풍의 경우 총 7개의 클러스터로 분류되었다. 클러스터 B의 경우 약 25.4%의 발생빈도를 가지며, 전선의 방향도 한반도를 직접 향하고 있어 상대적으로 한반도에 영향이 가장 큰 경로로 분석되었다. 클러스터 A는 한반도 서쪽으로 크게 선회하는 경로를 가지는 특징이 있으며 주로 강원도 지역에 영향을 많이 주는 태풍 경로로 분석되었다. 클러스터 G의 경우 클러스터 B와 유사한 경로를 가지고 있으나 종관기후학적으로 클러스터 B에 비해 저기압강도가 상대적으로 약한 특징을 가지고 있었다.
- 3) 한반도에 영향을 미치는 대표경로인 클러스터 B에 대하여 강우장의 분포를 도시한 결과 한반도 전체에 걸쳐서 강한 양(positive)의 강우량 Anomaly를 갖는 것을 확인할 수 있었다. 강우량 Anomaly의 값이 양의 값을 가진다는 의미는 구름이 많고 강우가 발생하고 있다는 것으로서 양의 값이 클수록 강한 저기압대가 위치하고 있다는 것을 의미한다. 즉, 태풍 경로를 따라 강한 집중호우 형태의 강우장이 형성되고 있음을 확인할 수 있다.

추후 연구로서 태풍의 경로에 대한 모의 기법을 개발하

고 이에 따른 한반도 주요 지역에 대한 강수의 발생빈도 및 집중호우 가능성을 평가할 수 있는 다변량 Monte Carlo 모의 기법을 개발하고 이를 활용한 태풍 위험도 평가기법을 도출하는 것이 필요할 것으로 사료된다.

감사의 글

본 연구는 국토교통부/국토과학기술진흥원 건설기술 연구사업의 연구비지원(13건설S01)에 의해 수행되었습니다. 저자들은 전북대학교 방재연구센터에 소속되어 연구를 수행하였습니다.

References

- Back, H.J., Kwon, W.T., and Oh, J.H. (1999) *Occurred in the western North Pacific climatological characteristics of a strong typhoon*. Korea Water Resources Association annual Meeting, pp. 177-182.
- Blender, R., Fraedrich, K., and Lunkeit, F. (1997). "Identification of cyclone-track regimes in the north atlantic." *Journal of The Royal Meteorological Society*, Vol. 123n No. 539, pp. 727-741.
- Camargo, S.J., Robertson, A.W., Gaffney, S.J., Smyth, P., and Ghil, M. (2007). "Cluster analysis of typhoon track. Part I: General properties" *Journal of Climate*, Vol. 20, No. 20, pp. 3635-3653.
- Dimego, G.J., and Bosart, L.F. (1982). "The transformation of tropical storm agents into an extratropical cyclone. Part II: Moisture, Vorticity and Kinetic Energy Budget." *Monthly Weather Review*, Vol. 110, No. 5, pp. 412-433.
- Elsner, J.B. (2003). "Tracking Hurricanes." *Bulletin of the American Meteorological Society*, Vol. 84, No. 3, pp. 353-356.
- Elsner, J.B., and Liu, K.B. (2003) "Examining the ENSO-typhoon hypothesis." *Journal of Climate Research*, Vol. 25, pp. 43-54.
- Emanuel, K., Sundararajan, E., and Williams, J. (2008). "Hurricanes and global warming: results from downscaling IPCC AR4 simulations." *Journal of American Meteorological Society*, Vol. 89, No. 3, pp. 347-367.
- Faraway, J., and Chatfield, C. (1998). "Time series forecasting with neural networks: a comparative study

- using the airline data.” *Journal of The Royal Statistical Society*, Vol. 47, No.2 pp. 231-250.
- Gaffney, S.J. (2004). *Probabilistic curve-aligned clustering and prediction with regression mixture models*. Ph. D. dissertation, University of California, Irvine, CA 281.
- Harr, P.A., and Elsberry, R.L. (1995). “Large-scale circulation variability over the tropical western north pacific. Part I: Spatial Patterns and Tropical cyclone characteristics.” *Monthly Weather Review*, Vol. 123, No. 5, pp. 1225-1246.
- Jeong, D.-M., and Bae, D.-H. (2004). “Monthly dam inflow forecasts by using weather forecasting information.” *Journal of Korean Water Resources Association*, Vol. 37, No. 6, pp. 449-460.
- Kim, H.-S., Kim, J.-H., Ho, C.-H., and Chu, P.-H. (2011). “Pattern classification of typhoon tracks using the fuzzy c-means clustering method.” *Journal of Climate*, Vol. 24, No. 2, pp.488-508.
- Kim, K.-H., Choi, W.-Y., and Oh, S.-K. (2013). “Design of summer very short-term precipitation forecasting pattern in metropolitan area using optimized RBFNNs.” *Journal of Korean Institute of Intelligent Systems*, Vol. 23, No. 6, pp. 533-538.
- Kim, S.-U., and Lee, K.-S. (2008). “Regional low flow frequency analysis using bayesian multiple regression.” *Journal of Korean Water Resources Association*, Vol. 41, No. 3, pp. 325-340.
- Knutson, T.R., McBride, J.L., Chan, J., Emanuel, K., Holland, G., Landsea, C., Held, I., Kossin, J.P., Srivastava, A.K., and Sugi, M. (2010). “Tropical cyclones and climate change.” *Nature Geoscience*, Vol. 3, pp. 157-163.
- Korea Meteorological Administration. (2011) “Typhoon White Paper.”
- Kwon, H.-H., Kim, T.-J. Hwang, S.-H., and Kim, T.-W. (2013). “Development of daily rainfall simulation model based on homogeneous hidden markov chain.” *Journal of Korean Society of Civil Engineer*, Vol. 33, No. 5, pp. 1861-1870.
- Lee, C.-R. (2001). *A study on meteorological disasters in Korea*. Master’s thesis, University of KyungHee, Seoul, Korea.
- Lee, D.-K., Jang, D.-E., and Wee, T.-K. (1992). “Typhoon approaching Korea, 1960~1989 part I: Statistics and synoptic overview.” *Journal of Atmospheric Sciences*, Vol. 28, No. 2, pp. 133-147.
- Lee, J.-J., and Kwon, H.-H. (2011). “Analysis on spatio-temporal pattern and regionalization of extreme rainfall data.” *Journal of Korean Society of Civil Engineer*, Vol. 31, No. 1B, pp. 13-20.
- Oh, T.-S., and Moon, Y.-I. (2008). “Special quality analysis of extreme rainfall by typhoon.” *Journal of Korean Society of Civil Engineer*, Vol. 28, No. 5B, pp. 459-473.
- Oh, T.-S., Moon, Y.-I., and Ahn, J.-H. (2007). “A study on special quality of hourly precipitation of typhoon happened in Korea.” *Journal of Korean Water Resources Association*, Vol. 40, No. 9, pp. 709-722.
- Oh, T.-S., Moon, Y.-I., Chun, S.-Y., and Kwon, H.-H. (2009). “Evaluation of extreme rainfall based on typhoon using nonparametric monte carlo simulation and locally weighted polynomial regression.” *Journal of Korean Society of Civil Engineer*, Vol. 29, No. 2B, pp. 193-205.
- Park, J.-K., Kim, B.-S., Jung, W.-S., Kim, E.-B., and Lee, D.-G. (2006). “Change in statistical characteristics of typhoon affecting the korean peninsula.” *Journal of Korean Atmosphere Association*, Vol. 16, No. 1, pp. 1-17.
- Scheffner, N., Borgman, L., and Mark, D. (1996). “Empirical simulation technique based storm surge frequency analyses.” *Journal of Waterway, Port, Coastal, Ocean Engineer*, Vol. 122, No. 2, pp. 93-101.
- Yang, Y., and Zou, H. (2004). “Combining time series models for forecasting.” *Journal of Forecasting*, Vol. 20, No. 1, pp. 69-84.
- Yoo, S.-A., and Jeong, J.-S. (2000). “A study on Typhoons affect the peninsula in 2000.” *Korean Meteorological Society Fall Meeting*, Vol. 2000. pp. 302-304.

논문번호: 14-022	접수: 2014.02.24
수정일자: 2014.03.25	심사완료: 2014.03.25