

# 군집기법을 이용한 연강수량 예보개선에 관한 연구

## A Study on the Improvement of Quantitative Precipitation Forecast using a Clustering Method

김광섭\*, 조소현\*\*

Gwangseob Kim, So Hyun Jo

### 요 지

연 및 계절강수량의 정확한 예보는 수자원관리에 매우 중요하다. 예보 정확도를 높이기 위한 다양한 연구가 계속 진행되어 왔다. 그럼에도 불구하고 강수량자료가 가지는 매우 큰 불확실성 때문에 예보의 정확도 향상은 계속되는 숙제로 우리에게 남아 있다. 이를 개선하기 위하여 본 연구에서는 군집화 기법을 이용한 연 및 계절 강수량 예측개선에 대한 연구 결과를 제시하였다. 이를 위하여 연강수량, 계절강수량 및 월강수량의 예측을 위하여 전구에서 일어나는 각종 기후 인자들과의 상관성 분석은 대단히 중요하다. 전 세계적으로 어느 특정 지역에서의 선행 기후인자 변화 양상이 우리나라의 강수량에 높은 상관성을 가지며 영향을 미친다면 예측을 위한 매우 유용한 정보라 하겠으나 국내 강수량과 기후 지수 사이의 선형 상관성은 매우 낮을 뿐만 아니라 지체상관성도 특정 지체에서 매우 큰 상관성을 보이는 인자를 찾기 어려움을 알 수 있다. 이를 극복하기 위하여 본 연구에서는 k-mean clustering을 이용하여 우리나라 주변의 기후조건을 분류하고 기후조건에 따른 강수량의 변화를 분석하였다. 남중국해역(105°E~135°E, 0°N~35°N), 우리나라 연안 해역(110°E~150°E, 20°N~40°N), 인도양 해역(75°E~105°E, 0°N~25°N) 및 아라비아 해역(45°E~75°E, 0°N~30°N) 평균 해수면 온도 변화에 따라 8개 군집으로 분류한 분석결과로 분석결과 2008년도는 그룹 5에 해당하며 그룹 5의 기후 상태는 근해와 남중국해역의 평균 해수면 온도가 평년보다 낮고 인도양 해역과 아라비아 해역의 평균 해수면 온도는 평년값과 비슷한 상태를 나타낸다. 그룹 5에 해당하는 기후조건에서 차년의 강수평균은 평년값 보다 적음을 보였다. 이러한 특성은 전체 유역에 걸쳐 동일하게 나타났다. 이에 대한 계절적 평균 분포는 군집 5에 대한 차년도 강수의 평균 계절분포는 전체적으로 평년값보다 낮게 나타났다. 이에 근거하여 올해 연 평균 강수량은 평년값보다 적을 것이며 전체 계절에 대하여도 평년값보다 적은 강수량이 올 것으로 판단된다. 이는 기상청의 2009년 봄철 기후전망과 유사한 예측 결과를 보여준다.

**핵심용어 : 강수량, clustering, 기후지수, 강수예측**

### 1. 서 론

IPCC 2007 보고서에 따르면 아시아 전역은 전구 평균을 상회하는 기온상승이 예상되며 특히 동아시아지역은 여름철 heat wave와 hot spell의 기간이 길어지고 강도와 빈도는 증가하는 경향을 보이며 여름철 강수량과 호우일수가 증가하는 것으로 나타났다. 이는 우리나라의 과거 관측자료를 분석한 기후변화 경향성과 일치한다. 특히 아시아지역의 기후변화를 이해하는데 있어 몬순 특성은 가장 중요한 요소의 하나로 몬순의 영향을 받는 지역의 자료를 수집하고 국내 기후변화와의 상관 분석을 수행하였다. 이를 위하여 남중국해역(105°E~135°E, 0°N~35°N), 우리나라 연안 해역(110°

\* 정회원-경북대학교 건축토목공학부 교수E-mail : [kimsgs@knu.ac.kr](mailto:kimsgs@knu.ac.kr)

\*\* 경북대학교 건축토목공학부 석사과정E-mail : [juicylucy7340@hotmail](mailto:juicylucy7340@hotmail)

E~150°E, 20°N~40°N), 인도양 해역 (75°E~105°E, 0°N~25°N) 및 아라비아 해역 (45°E~75°E, 0°N~30°N) 평균 해수면 온도 자료를 구축하고 우리나라 전체 평균 및 5대강 유역 평균 강수량자료와의 상관관계 및 변동 특성을 분석하였다. 국내 강수 변동 특성상 특정지역의 기상자료와 높은 선형적인 상관성을 보장하는 자료를 찾을 수 없고 최근 우리나라는 풍수년의 강수형태를 보이고 있으나 장기자료의 변동특성을 볼 때 심각한 갈수년의 가능성이 있으며 특정지역의 강수메커니즘이 풍수년이나 갈수년일 때 상태를 지속적으로 유지하려는 특성을 고려할 때 기후변화의 영향평가의 불확실성을 최소화하여 신뢰성 있는 수자원 부존량의 변동성을 정량적으로 전망하기 위하여 전술한 자료와 군집화 기법을 사용하여 상태 분류에 따른 강수 예측성 형상가능성을 제시하고자 하였다.

## 2. 군집화 기법을 이용한 강수 예측

우리나라 강수 자료의 유역 변화 특성을 분석하기 위하여 지상 관측망이 전국적으로 구축되어 관측된 1973년 이후자료에 대하여 분석하였다. 먼저 강수량 anomaly에 대한 기본 통계량 분석 결과는 유역별로 다소 차이가 있으나 한강 유역은 다우월(본 연구에서 월 평년값 보다 100mm이상 강수가 많이 오는 월로 정의) 빈도가 소우월(월 평년값 보다 100mm이상 강수가 적게 오는 월로 정의) 빈도의 약 1.5배에 해당하며(432개월 동안 32:22) 낙동강과 금강유역은 다우월과 소우월비가 비슷하고(29:28, 28:27) 섬진강과 영산강 유역은 소우월이 다소 많다(각각 36:38, 31:34). 우리나라 평균 월 강수량이 경우 다우월이 소우월이 약 두 배에 해당한다(30:17). 뿐만 아니라 계절별 강수량 변화 분석 결과, 봄철(3,4,5월: MAM)과 여름철(6,7,8월: JJA) 표준편차가 매우 커서, 봄철과 여름철의 강수량에 대한 장기 예측이 매우 어려움을 나타낸다. 이러한 연강수량, 계절강수량 및 월강수량의 예측을 위하여 전구에서 일어나는 각종 기후 인자들과의 상관성 분석은 대단히 중요하다. 이를 위하여 Table 3.5.1에서 제시한 바와 같이 34개의 기후인자와 우리나라 강수량 변동 상관분석을 실시하였다. 전 세계적으로 어느 특정 지역에서의 선행 기후인자 변화 양상이 우리나라의 강수량에 높은 상관성을 가지며 영향을 미친다면 예측을 위한 매우 유용한 정보라 하겠으나 전술한 바와 같이 선형 상관성은 매우 낮을 뿐만 아니라 지체상관성도 특정 지체에서 매우 큰 상관성을 보이는 인자를 찾기 어려움을 알 수 있다.

MacQueen(1967)에게서 제안된 K-mean clustering을 이용하여 우리나라 주변의 기후조건을 분류하고 기후조건에 따른 강수량의 변화를 분석하였다. Table 1는 남중국해역(105°E~135°E, 0°N~35°N), 우리나라 연안 해역 (110°E~150°E, 20°N~40°N), 인도양 해역 (75°E~105°E, 0°N~25°N) 및 아라비아 해역 (45°E~75°E, 0°N~30°N) 평균 해수면 온도 변화에 따라 8개 군집으로 분류한 분석 결과로 분석결과 2008년도는 그룹 5에 해당하며 Fig 1에서 제시한 바와 같이 그룹 5의 기후 상태는 근해와 남중국해역의 평균 해수면 온도가 평년보다 낮고 인도양 해역과 아라비아 해역의 평균 해수면 온도는 평년값과 비슷한 상태를 나타낸다. 그룹 5에 해당하는 기후조건에서 차년의 강수평균은 평년값 보다 적음을 보였다. Table 1에서 제시한 바와 같이 그룹 5는 이러한 특성은 전체 유역에 걸쳐 동일하게 나타났다. 이에 대한 계절적 평균 분포는 Fig 2에서 제시한 바와 같이 군집 5에 대한 차년도 강수의 평균 계절분포는 전체적으로 평년값보다 낮게 나타났다. 이에 근거하여 올해 연 평균 강수량은 평년값보다 적을 것이며 전체 계절에 대하여도 평년값보다 적은 강수량이 올 것으로 판단된다. 이는 기상청의 2009년 봄철 기후전망과 유사한 예측 결과를 보여준다.

Table 1. Annual Precipitation of Each Group in This Year and Next Year.

	그룹	G1	G2	G3	G4	G5	G6	G7	G8
	그룹해당총년수	3	2	1	3	4	7	10	6
전체 평균	당해년 연강수량	1527.6	1243.5	1340.6	1420.2	1214.3	1195.9	1405.3	1358.3
	차년도 연강수량	1530.5	1175.4	1111.3	1507.0	1082.6	1305.6	1454.5	1255.5
한강유역	당해년 연강수량	1512.1	1172.4	1288.0	1303.2	1241.0	1229.0	1405.4	1385.2
	차년도 연강수량	1434.8	1158.6	1181.1	1371.4	1093.2	1407.5	1420.3	1297.0
낙동강유역	당해년 연강수량	1566.5	1196.8	1392.4	1365.8	1101.8	1164.9	1374.5	1315.0
	차년도 연강수량	1595.4	1186.8	1007.0	1395.0	1049.6	1214.2	1428.4	1239.2
금강유역	당해년 연강수량	1336.0	1158.3	1296.4	1370.7	1141.4	1131.8	1340.1	1291.5
	차년도 연강수량	1386.7	1114.6	999.4	1515.8	1004.6	1273.7	1373.5	1132.2
섬진강유역	당해년 연강수량	1634.3	1423.6	1416.5	1670.3	1338.0	1207.7	1486.0	1418.3
	차년도 연강수량	1680.0	1204.5	1189.3	1759.2	1127.1	1299.7	1577.0	1323.0
영산강유역	당해년 연강수량	1448.3	1404.5	1247.9	1564.7	1312.0	1153.2	1378.3	1291.8
	차년도 연강수량	1522.8	1111.4	1054.7	1689.4	1003.4	1176.8	1509.6	1258.8

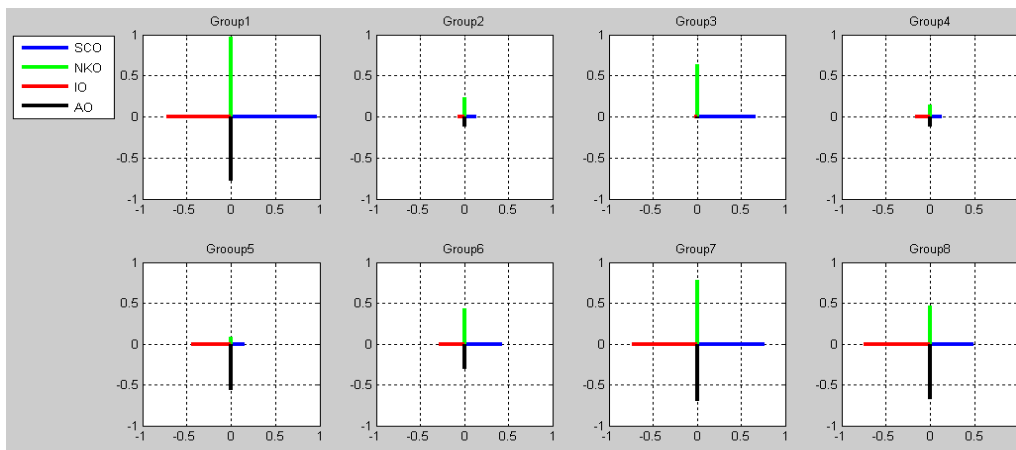


Fig 1. Grouping using K-mean Clustering (SST in South China Sea, Neighborhood Sea of Korea, Indian Ocean and Arabian Sea).

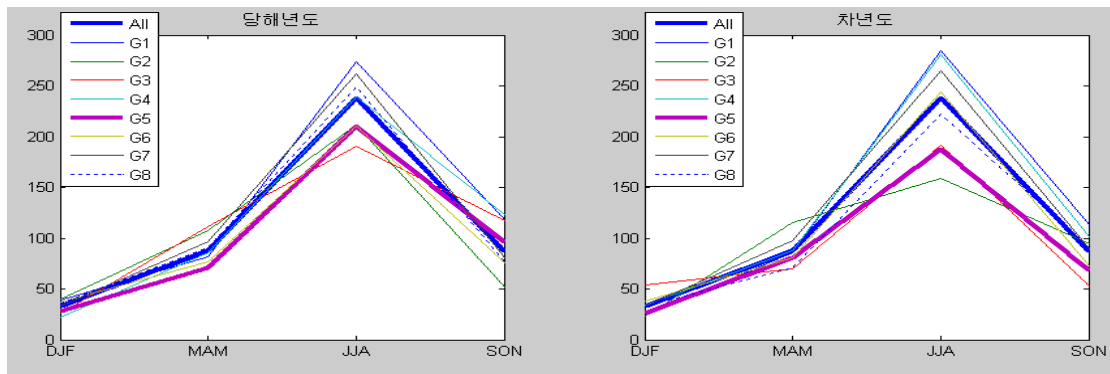


Fig 2. Seasonal Distribution of Annual Precipitation This Year and Next Year.

### 3. 결 론

동아시아 지역과 국내 강수량 상관분석 결과 계절변화를 제거한 자료에서 보여주는 장주기 특성은 전체 자료변화를 제시하는 부분이 미미하였으며 각종 기후 지수와 국내 강수량의 선형 상관성이 낮다. 이러한 여러 기후인자들 사이의 복잡한 상관성을 고려할 기법으로 군집화 기법 적용하여 동아시아 지역 기후패턴의 군집분석 결과를 이용한 국내 강수량 상관성은 당해 연도 및 차년도 강수량이 감소하는 집단으로 2008년이 분류되어 올해(2009년) 연 평균 강수량은 평년값보다 적을 것이며 전체 계절에 대하여도 평년값보다 적은 강수량이 올 것으로 판단된다. 이는 기상청의 2009년 봄철 기후전망과 유사한 예측 결과를 보여 주었다.

### 감 사 의 글

본 연구는 21세기 프론티어 연구개발사업의 수자원의 지속적 확보개발 사업단의 연구비 지원(1-9-3)에 의해 수행되었다.

### 참 고 문 헌

1. MacQueen, J. B. (1967). Some Methods for classification and Analysis of Multivariate Observations, in Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. 1: 281-297, University of California Press
2. Stefano Serafin, Alessio Berto, Dino Zardi, (2005). Application of cluster analysis technique to the verification of quantitative precipitation forecasts, 28th International Conference on Alpine Meteorology (ICAM), Zadar, Croatia