

국내 유역에 대한 GCM 정보의 확률론적 불확실성 분석

Uncertainty Analysis of GCM Information in Korea Using Probabilistic Diagnostics

정 창 삼* / 허 준 행** / 배 덕 효***

Jeong, Chang Sam / Heo, Jun Haeng / Bae, Deg Hyo

Abstract

The objective of this study is to examine the usefulness of climate model simulations (GCM) in Korea water resource management. The methods are based on probabilistic measures of the effectiveness of GCM simulations of an indicator variable for discriminating high versus low regional observations of a target variable. The formulation uses the significance probability of the Kolmogorov-Smirnov test for detecting differences between two variables. AMIP-II(Atmospheric Model Intercomparison Project-II) type GCM simulation done by ECMWF(European Centre for Medium-Range Weather Forecasts) was used for indicator variable and observed mean average precipitation(MAP) values on 7 major river basins were used as target variable. Monte Carlo simulation is used to establish the significance of the estimator values. The results show that GCM simulations done by ECMWF are skillful in discriminating the high from the low of the observed MAP for wet season in all seven basins of Korea, but not enough for dry season.

Keywords : GCM simulation, indicator variable, target variable, K-S test

요 지

기상예보모형 중 장기예측에 널리 사용되는 GCM 모의결과를 이용하여 확률론적 불확실성 해석기법의 적용을 통해 유역단위로 관리되는 국내 수자원 운영에 대한 활용 가능성을 분석하였다. 연구된 기법은 GCM 모의값이 관측값의 크고 작음을 얼마나 잘 구분하는지를 확률적으로 분석하는 방법으로 Kolmogorov-Smirnov 검정을 사용한다. GCM 모의값으로는 ECMWF에서 AMIP-II 형태로 모의한 결과로부터 표면강수량을 추출하여 사용하였으며, 관측값은 국내 7개 유역에 대해 면적강수량을 산정하여 사용하였다. 또한, 어느 정도의 구분능력이 적정한가를 판단하기 위한 유의수준(significance threshold)을 결정하기 위해 Monte Carlo 모의를 사용하였다. 이러한 분석을 통해 우리나라의 7개 유역에 대해서는 ECMWF의 GCM 자료가 우기(6월~9월)에 대해 인근 노드점의 자료를 평균하여 사용되어 질 경우 효율적인 것으로 나타났으나, 건기(10월~5월)의 경우 구분 능력이 부족한 것으로 판단된다.

핵심용어 : GCM 모의, 모의값, 관측값, K-S 검정기법, 불확실성 해석

* 연세대학교 토목공학과 박사수료

Ph. D. Candidate, Dept. of Civil Eng. Yonsei Univ. Seoul. 120-749, Korea
(E-mail : chs1229@yonsei.ac.kr)

** 연세대학교 사회환경시스템공학부 교수

Prof. Dept. of Civil Eng. Yonsei Univ. Seoul. 120-749, Korea

*** 세종대학교 토목환경공학과 부교수

Associate Prof. Dept. of Civil and Environmental Eng. Sejong Univ. Seoul. 143-747, Korea

1. 서론

수자원의 적절한 운영을 위해서는 가용한 수자원의 미래 상황에 대해 불확실성이 가미된 정보를 바탕으로 결정을 내려야 하는 경우가 대부분이다. 오늘날 기상예보는 수자원 운영자 및 계획입안자에게 여러 가지 정보를 쏟아내고 있다. 비록 기상예보정보가 많은 지역적인 오차를 가지고 있기는 하지만, 궁극적으로 이에 대한 적절한 이용이 상당한 경제적 편익을 가져다 주리라는 사실은 많은 연구결과에 의해 증명이 되어왔다 (Georgakakos 등, 1998, Georgakakos 등, 2000, Carpenter와 Georgakakos, 2001, Yao와 Georgakakos A., 2001). 이러한 기상예보 정보 가운데 가장 보편적으로 이용되는 수치모형인 GCM(Global Climate Model)으로부터 얻은 정보를 이용하여 실제 수자원 분야 적용과 같은 기상정보의 활용가능성을 연구하다 보면 다음과 같은 문제점에 봉착하게 된다. 첫째, GCM 모의 결과가 모형 자체 규모(scale)를 기반으로 한 예보에 유용한가 하는 문제와 둘째, GCM 모의 결과가 더 작은 유역에 대한 예보를 위해 사용되어질 수 있는가 하는 문제이며, 또한 만일 사용되어질 수 있다면 어떻게 상관관계를 설정할 수 있는가 하는 것을 들 수 있다. 첫 번째 질문의 경우 기존의 많은 연구에서 긍정적인 결과를 보여 왔으나, 두 번째 질문에 대해서는 명확한 분석이 부족한 편이었다. 최근 이러한 연구의 일환으로 Georgakakos(2003)는 GCM 가운데 하나인 ECHAM3 앙상블 모의결과를 이용하여 미국 전지역에 대한 기후모형의 확률론적 불확실성을 앙상블과 공간 규모를 중심으로 모의하여 긍정적인 결과를 도출한 바 있다.

본 연구에서는 이러한 문제에 답하기 위해 국내 GCM 정보의 확률론적 불확실성 해석을 수행하고자 한다. 연구된 기법은 GCM 모의값이 관측값의 크고 작음을 얼마나 구분해 낼 수 있는지를 분석하는 확률론적 방법이다. 이러한 분석을 통해 수문학자나 수자원 운영

자가 기후모형의 다양한 예측결과 가운데, 특정 목적에 어떤 결과를 사용해야하며, 다양한 결과 가운데 어떤 형태로 자료를 적용해야하는가와 각각의 정보에 얼마만큼의 신뢰도를 부여해야하나 같은 문제들에 정량적 기준을 마련할 수 있는 하나의 방안을 제시하고자 한다.

2. 분석 방법론

본 연구는 GCM 모의값과 지상 관측값을 K-S 검정 기법의 유의 확률값 개념을 이용하여 상관관계를 규명하는 것이다. 그림 1은 이러한 방법론을 개념적으로 나타낸 것으로 GCM 모의값과 관측값의 과거 상관관계 규명을 통해 이를 이용한 미래 수자원 상황예측에 GCM정보를 활용하기 위함이다.

2.1 수학적 표현

특정유역에 대해 GCM의 격자화된 모의결과를 분석하기 위해서는 그림 2에서 나타낸 것과 같은 창(window) 개념의 분석이 필요하다. 유역의 중심을 (x_0, y_0) 라하고, GCM 격자 간격을 e_x, e_y 라하고, 창의 크기를 λ 라고 할 경우, 유역 중심에 대한 설정 창은 아래 식 (1)과 같이 표현될 수 있다.

$$\Omega_\lambda = \{n \in \Omega_\lambda : (|x_n - x_0| = d_x) \cap (|y_n - y_0| = d_y)\} \tag{1}$$

여기서 x_0, y_0 : 대상유역 중심점(centroid)의 위·경도이고, Ω_λ : 설정 창 내부의 GCM 격자 집합, e_x, e_y : 격자점의 x축, y축 방향 간격, $d_x = \lambda e_x, d_y = \lambda e_y$ 이다.

유역내의 몇몇 지상관측지점의 관측자료를 이용하여 면적 평균된 기상관측값을 산정한 후 이 값을 유역중심 (x_0, y_0) 에 대한 대표값으로 가정하여 창의 개념의

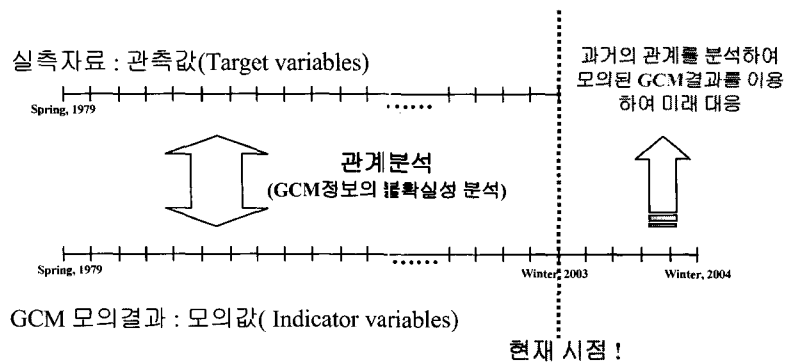


그림 1. GCM 정보의 불확실성 분석

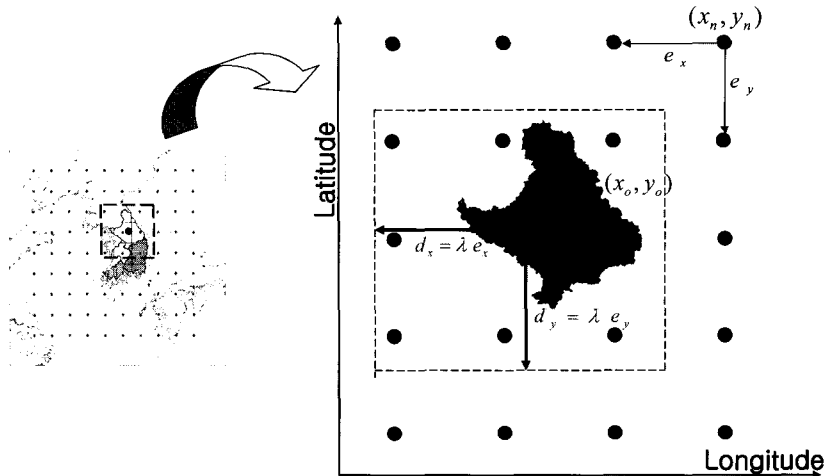


그림 2. GCM 분석을 위한 대상구역 창의 정의

통해 주변 격자점의 GCM 모의값들과 관계를 설정할 경우, 관측값과 모의값의 공간적 상관성이 설정될 수 있다. 이 경우 대상 구역의 각종 관측값들은 참값으로 가정하며 모든 분석의 기준이 되는 관측값이 되고, 이에 상응하는 설정 창 내부 GCM 격자점의 각종 모의자료들은 모의값이 된다. 여기서 사용되는 관측값들은 면적평균강수량, 면적평균온도, 구역 유출량 등의 각종 수문관측자료가 될 수 있으며, 모의값은 GCM 모의에서 추출 가능한 값들로 관측값과 상관성을 지니는 표면강수(surface precipitation), 표면온도(surface temperature), 700hPa기압에 해당하는 고도 등의 기상변수가 해당된다. 관측값과 모의값은 반드시 같을 필요는 없으나, 상관성을 지니는 것을 상호 분석해야 한다. 예를 들어 면적평균강수량-표면강수, 면적평균온도-표면온도, 구역 유출량-표면강수 등이 이에 해당하며, 본 연구에서는 면적평균강수량-표면강수의 예를 적용하였다.

구역면적평균값으로 구성된 관측값과 GCM 노드점 단위의 모의값을 바탕으로 그림 2에서와 같이 설정된 창 안에 있는 노드점들의 모의값들을 이용하여 임의의 기준(2등분, 혹은 3등분)으로 모의값들을 구분하고, 이러한 구분에 상응하도록 연도별로 관측값을 나눈 후 이 결과를 이용하여 아래와 같이 두 개의 누가분포함수(Cumulative Density Function, CDF)를 구성한다. 예를 들어 2등분의 경우, 모의값들의 중앙값(median)을 기준으로 이상과 미만으로 구분하거나, 3등분의 경우 각 영역을 33.3%로 나누어 구분한 후 이 기준을 가지고 관측값을 구분한 후 각각의 자료에 대한 CDF를 작성하는 것이다. 설정 창 내에 9개의 노드점이 있다면 모두 9개의 대해 각각의 CDF가 작성되게 된다.

$$F_L(u_i^{(j)}/n) = \{j/N_L; u_i \in \Omega_u: (P_n^m = q_l)/(n)\} \quad (2)$$

$$F_U(u_i^{(j)}/n) = \{j/N_U; u_i \in \Omega_u: (q_u = P_n^m)/(n)\} \quad (3)$$

여기서, $F_L(u_i^{(j)}/n)$: 모의값에 의해 중앙값 미만 혹은 하위 33.3%로 구분된 관측값의 CDF, $F_U(u_i^{(j)}/n)$: 모의값에 의해 중앙값 이상 혹은 상위 33.3%로 구분된 관측값의 CDF, q_l, q_u : 표면강수량과 같은 GCM 모의값들 중 구분이 되는 하한, 상한 값이고, N_L, N_U : q_l, q_u 에 의해 구분되는 관측값의 총수, P_n^m : m기간에 해당하는 GCM 모의값(n격자), $u_i^{(j)}$ 는 j번째 크기의 관측값, n은 각 격자이다.

이렇게 모의값에 의해 구분된 두 관측값들의 통계적 특성이 서로 다르다면(두 CDF가 다르다면), 모의값이 관측값을 구분하는데 매우 효과적이라는 사실을 알 수 있으며, 이와 반대로 통계적 특성이 같다면 결국 구분 조건, 즉 GCM정보가 관측값을 구분하는데 아무런 효과를 가지고 있지 않다는 뜻이 된다.

그렇다면 어떠한 방법을 통해 구성된 두 자료들(CDF)의 통계적 특성(분포형)이 같고 다름을 구분할 것인가 하는 문제를 해결해야 한다. 이 문제에서 우리가 얻어야 하는 결과는 각각이 어떤 분포를 가지고 있어 통계적으로 다르냐 하는 것이 아니라 단지 각각이 어느 정도로 다른가를 정량적으로 알고 싶은 것이다. 이를 위해 적합도 검정기법 가운데 하나인 K-S 검정기법을 도입하였다. K-S 검정기법은 각 대상자료에 대한 CDF를 이용하여 두 자료군이 동일한 분포를 가지는가

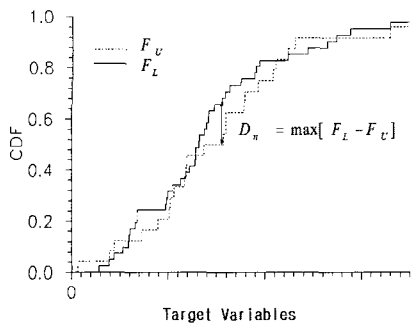


그림 3. 두 CDF의 최대차 (D_n,k)

아닌가를 판별해 주는 검정기법이다. K-S 검정기법의 특징은 각각의 자료군에 대한 정보(각각의 자료의 수)나 사용자가 설정하는 값(유의수준)만을 가지고 통계적인 검정을 할 수 있다는 점이다.

본 연구에서는 검정의 통과 여부에 대한 단순한 판단보다는 통과 여부를 통계적으로 표현한 후 이를 이용한 진단변수(estimator)를 구성하여 모의값의 관측값 구분능력을 정량화 하는데 목적이 있다. 이를 위해 Press 등(1992)이 제시한 식 (4)의 검정 결과를 확률로 표현하는 방법을 도입하였다. 다른 많은 검정 방법이 있지만, 본 연구에서 K-S 검정기법을 사용한 가장 큰 이유가 바로 결과를 식 (4)와 같이 확률로 표현해 줄 수 있다는 것으로, 검정 결과 서로 다른 통계적 특성을 가질 확률이 높을 경우 이는 모의값이 관측값을 구분해내는데 효과적이라는 사실을 알 수 있다. 이와 반대로 동일한 통계적 특성을 가질 확률이 높을 경우 모의값이 관측값을 구분해내지 못하는 것으로 상관관계가 없음을 의미한다. K-S 검정기법은 아래의 그림 3과 같이 구성된 두 CDF의 차이가 제시된 기준에 적합한가 아닌가를 계산하여 동일한 통계적 분포를 가지는지를 평가하는 기법이다. 그림에서와 같이 구성된 두 경험적 CDF의 최대 차이를 D_n라하고 식 (5)와 같이 표현할 수 있다.

$$P_{KS}(D > D_n) = 2 \sum_{j=1}^{\infty} (-1)^{j-1} e^{-2j^2 A^2}; \quad (4)$$

$$A = [\sqrt{N_e} + 0.12 + 0.11 / \sqrt{N_e}] D$$

여기서 $N_e = \frac{N_L N_U}{N_L + N_U}$

$$D_n = \max [|F_L(u_i^{(j)} / n) - F_U(u_i^{(j)} / n) |] \quad (5)$$

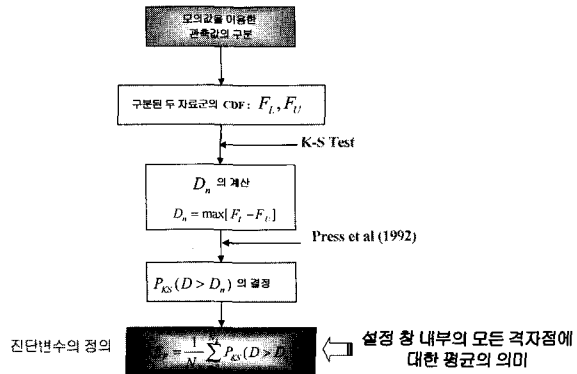


그림 4. 구성된 모형의 흐름도

식(4)는 식(5)에서 표현된 D_n값을 이용하여 두 자료군의 통계적 특성의 차이를 K-S검증기법을 이용하여 확률적으로 표현한 것이다(Press 등, 1992). 일반적으로 위의 식 (4)는 N_e ≥ 4인 경우에 유효한 것으로 알려져 있으며, 작은 P_{KS}값은 더 큰 D_n값을 의미하므로 결국 두 CDF가 다를 확률이 높음을 나타내어 모의값이 관측값을 구분해낼 가능성이 높다는 것을 의미한다. 이 P_{KS}값을 그림 2에서 설정된 모의값과 관측값의 공간적 상관관계와 앙상블로 표현되는 다양한 GCM 모의값들의 상관관계를 설정하기 위해 아래와 같은 진단변수를 정의하였다. 정의된 진단변수인 E_p는 설정 창 내부의 모든 격자에 대한 평균을 의미하며, 식 (6)과 같이 표현된다.

$$E_p = \frac{1}{N_L} \sum_{n=1}^{N_L} P_{KS}(D > D_n) \quad (6)$$

2.2 분석절차

2.1절에서 제시한 수학적 표현을 바탕으로 실제 적용을 위해 먼저, 특정 기간(예를 들어 계절 혹은 월)에 대해 연도별로 장기간 생성된 GCM 모의결과를 크기 순으로 나열한 후 중앙값을 중심으로 양분한다. 이러한 구분을 바탕으로 그림 5에서와 같이 해당하는 연도의 과거의 관측자료를 중앙값 미만(below the median)과 중앙값 이상(over the median)으로 구분한다. 만일 이러한 조건이 유효할 경우 구분된 관측자료는 서로 다른 통계적 특성을 나타낼 것이고, 반대로 조건이 무효할 경우 과거자료를 무작위로 추출하게 되는 것이므로 서로 동일한 통계적 특성, 즉 모든 구간의 자료 그대로의 통계적 특성을 지니게 될 것이다. 결국 주어진 조건(GCM 정보)이 유효한가 아닌가 하는 문제를 이러한

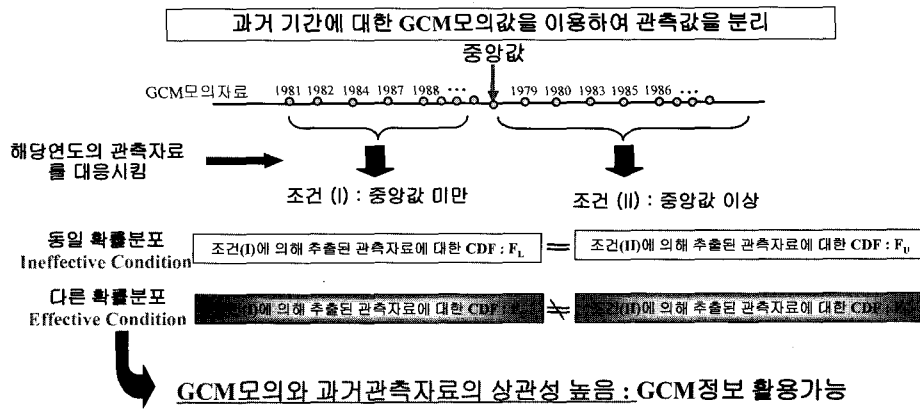


그림 5. 모의값을 이용한 관측값의 분리

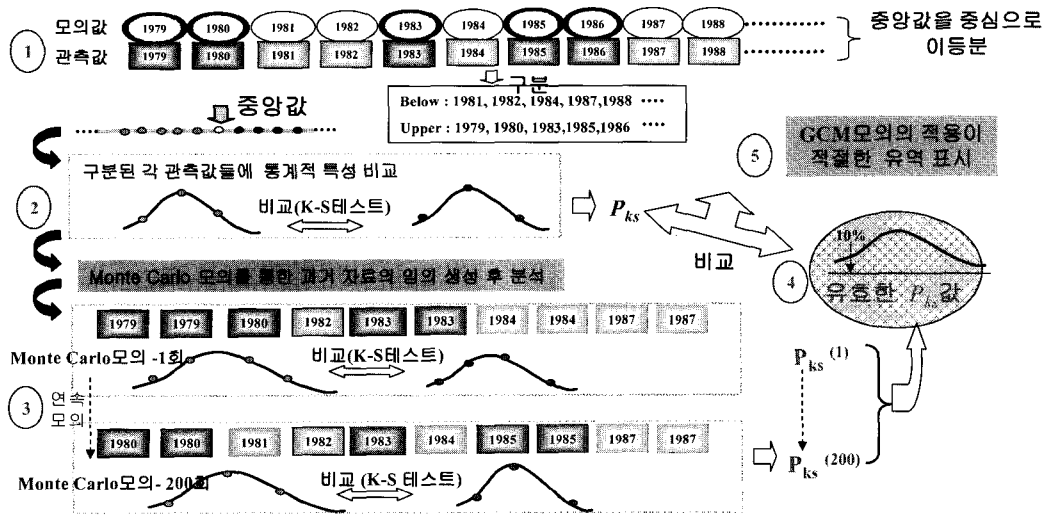


그림 6. Monte Carlo 모의를 통한 적정 유의수준의 결정

비교를 통해 판단하는 것이 본 연구의 기본적인 개념이라 할 수 있다.

그렇다면 어느 정도의 P_{KS} 값이 조건을 유효하게 하는 낮은 값이라 판단할 수 있는가 하는 문제가 발생한다. 이러한 문제를 해결하기 위해 Monte Carlo 모의 기법을 사용하였다. 방법은 주어진 자료에 대한 초기모의(actual run)를 통해 산정된 P_{KS} 값과 Monte Carlo 모의(기존의 관측 시계열 자료를 바탕으로 200개의 자료군을 모의 발생 후 각각에 대한 P_{KS} 값을 산정하고 산정된 값들을 이용하여 경험적 PDF를 구성한 후 하위 10%(임의 지정 가능)에 해당하는 값을 통해 얻어진 P_{KS} 값의 크기를 비교하여 초기모의의 결과가 더 작게 나타나는 경우를 유효한 지역으로 결정하게 되며, 이 값을 유의수준이라 정의한다. Monte Carlo 모의를 적용

하는 이유는 과거자료의 임의발생을 통해 미래의 발생 상황을 임의로 모의하는 것으로, 무작위 선정(random sampling)을 통해 생성된 자료가 임의의 사상이 될 수 있다는 가정 하에 확장된 자료를 이용하여 모형의 유효성을 평가하기 위함이다. 아래의 그림 6은 Monte Carlo 모의를 적용하여 적절한 유의수준을 산정하는 일반적인 방법을 나타내고 있다. 아래의 그림 6을 통해 위에서 언급한 내용을 과정별로 정리하면 다음과 같다.

- ① GCM 모의결과를 중앙값을 중심으로 2등분한다.
- ② 초기모의를 통해 초기 P_{KS} 값과 진단변수(I)를 산정한다.
- ③ Monte Carlo 모의를 통해 200개의 자료군을 발생시킨다(모의된 자료는 각 지점별 관측자료와 동일한 분포형과 평균값을 가지도록 원 자료를 무작위 추출함).

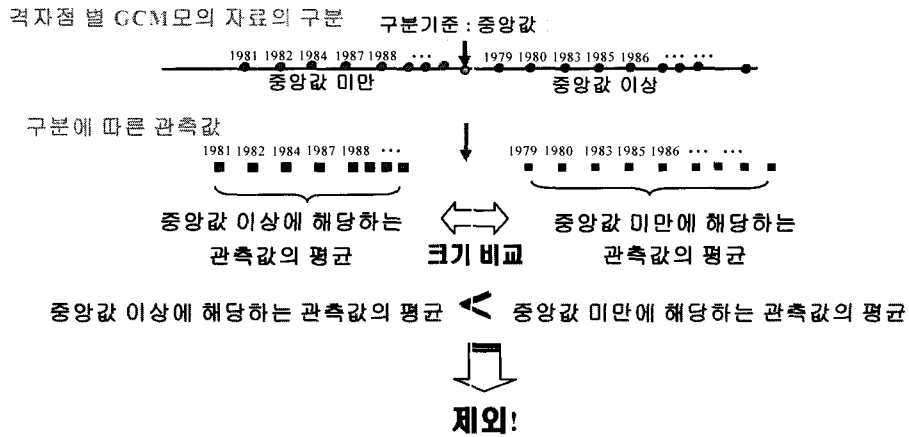


그림 7. 평균비교분석

④ 200개의 자료군에 대한 P_{KS} 값과 진단변수(I)을 산정한 후 이를 경험적 PDF를 이용하여 10%유의수준 값에 해당하는 진단변수(I)값을 산정한다.

⑤ 초기모의에서 산정된 진단변수값과 Monte Carlo 모의에서 구한 적정유의수준값에 해당하는 진단변수(I) 값을 비교하여 초기모의에서 구한 진단변수값이 더 작게 나타내는 유역에 대해 GCM 모의결과를 이용하기에 적합하다고 판단하고 이를 적절하게 유역도로 표현한다.

⑥ 구분된 자료의 평균을 계산하여 자료가 역전되지 않도록 제한 조건을 주어 계산에서 제외하는 평균비교 분석(Correct association)을 실시한다 (아래 그림 7 참조).

⑦ 평균비교분석을 통과하는 자료를 이용하여 ⑤의 과정을 반복한다.

3. 모형의 적용

본 연구를 위해서는 GCM의 과거 모의자료(모의값)와 GCM의 활용성을 검증할 과거 관측자료가 요구된다. 사용된 관측값은 유역면적평균강수량으로 관측값 작성을 위한 유역 구분은 수문단위지도를 기초로 하여 우리나라를 7개 유역(5대강유역, 안성천-삼교천 유역, 만경

강-동진강 유역)으로 구분하여 분석하였다. 또한 모형의 적용성에 대한 전반적인 판단을 위해 7개 유역을 크게 2개의 유역으로 통합 구분하여 모의값과 관측값의 상관성을 비교·분석하였다.

3.1 GCM 자료

모의값으로 선정된 자료는 ECMWF에서 AMIP-II 형태로 모의한 $2. \times 2.$ 의 고해상도 GCM의 1979~1996의 월별 지표면 강수량 자료의 일 평균값이다. 그림 8(a)와 같이 한반도를 중심으로 한 30개의 격자 자료를 추출하여 사용하였다. 그림 8(b)는 본 연구의 분석 대상이 되는 유역전반과 GCM 격자를 동일한 그림에 표현한 것이다.

3.2 관측 강수량 자료

관측자료로는 모의값과 동일하게 강우량을 선정하였으며, 동일한 자료 기간에 대해 관측값을 가지는 35개 관측지점의 자료를 이용하여 면적평균강수량을 산정하여 사용하였다. 그림 9(a)는 수문단위지도의 티센망을 중심으로 35개 기상청 지점의 위치를 나타내고 있다. 관측값은 그림 9(b)와 같이 우리나라내의 5대강 유역과 안성천-삼교천 유역, 만경강-동진강 유역의 7가지 유역

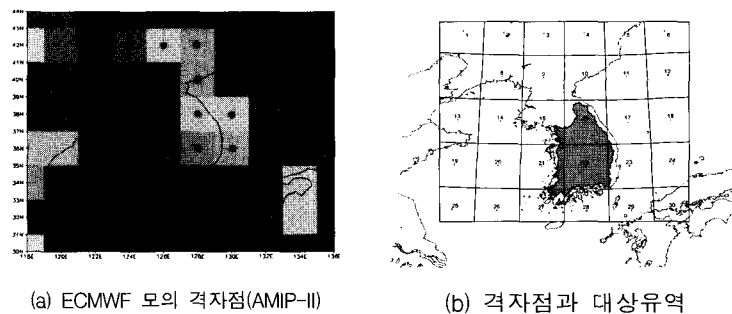


그림 8. ECMWF의 GCM 격자점과 대상유역

으로 구분하여 각 유역에 대한 면적평균강수량을 산정하여 사용했으며 모의값인 GCM 격자점과의 분석을 위한 창 설정의 중심이 되는 유역중심점은 무게중심의 개념을 이용하여 CAD상에서 작업하여 설정하였으며 동일한 그림에 나타내었다. 또한 모의값과 관측값의 자료 분석을 통해 모형의 형태를 결정하기 위해 그림 9(c)와 같이 유역을 크게 두 개로 구분하여 자료를 비교 분석하였다. 북쪽으로 구분된 유역은 한강, 금강, 안성천-삼교천 유역을 합친 것이며, 남쪽으로 구분된 유역은 낙동강, 섬진강, 영산강, 만경강-동진강 유역을 합친 것이다. 관측값 산정을 위해 선정된 기상청 지점의 명칭과 유역 면적 평균값 산정을 위해 결정된 티센계수는 표 1과 표 2에 나타내었다.

그림 10은 유역면적평균값으로 환산된 관측값과 모의값의 월 평균값과 월 표준편차를 북측과 남측으로 구

분하여 비교하였다. 그림에서 실선은 월평균을 의미하고 점선은 월 표준편차를 의미한다. 유역별로 차이는 있지만 어느 정도 유사한 경향을 나타내고 있으며, 특히 남측 하천의 경우 유사한 경향을 나타내고 있다. 또한 자료에 대한 검정을 위해 모의값과 관측값을 중심으로 건기, 즉 10월~5월에 해당하는 8달과 우기, 즉 6월~9월의 4달로 구분하여 각 해당연도에 대한 자료를 평균하여 산정한 후 1979~1996년의 18년 자료에 대해 표준화를 실시하여 각각에 대해 초과확률함수를 구하였다. 그림 11은 북측과 남측 하천의 경우에 대한 초과확률함수를 나타내고 있으며, 자료가 유사한 경향을 나타내고 있어 모형을 적용할 경우 타당한 결과를 도출할 가능성이 있음을 제시하고 있다. 특히 북측 하천에서 표준화된 자료를 이용할 경우 보다 유사한 경향을 나타내고 있음을 알 수 있다.

표 1. 유역의 구분과 사용된 티센계수 (7개 유역)

한강권역	서울	양평	원주	이천	인제	제천	춘천	충주	홍천	문경	영주	보은	청주	합계				
	0.0934	0.0621	0.0658	0.0573	0.2314	0.1331	0.1207	0.067	0.0814	0.0098	0.0563	0.017	0.0047	1				
낙동강권역	남원	임실	거창	구미	대구	문경	밀양	부산	산청	영주	영천	의성	진주	추풍령	합천	보은	합계	
	0.0082	0.0001	0.0662	0.0664	0.0823	0.0705	0.1082	0.0234	0.0531	0.1479	0.0743	0.1286	0.061	0.0358	0.0706	0.0034	1	
금강권역	이천	충주	임실	거창	추풍령	금산	대전	보은	부여	청주	합계							
	0.0105	0.0183	0.0358	0.0343	0.0926	0.1713	0.1264	0.0868	0.2199	0.2041	1							
섬진강권역	남원	순천	임실	장흥	광주	여수	거창	산청	진주	군산	합계							
	0.2858	0.2707	0.2309	0.0613	0.0293	0.0164	0.002	0.0525	0.0388	0.0123	1							
영산강권역	남원	순천	장흥	광주	목포	군산	합계											
	0.0083	0.0307	0.1096	0.6124	0.2163	0.0227	1											
안성,삼교	서울	이천	부여	청주	합계													
	0.2928	0.2646	0.1987	0.2439	1													
만경,동진강	임실	군산	금산	부여	합계													
	0.2369	0.5289	0.1357	0.0985	1													

표 2. 유역의 구분과 사용된 티센계수 (2개 유역 구분)

북측 유역	서울	양평	원주	이천	인제	제천	춘천	충주	홍천	임실	거창	문경	영주	추풍령	군산	대전	보은	부여	청주	합계					
	0.0936	0.0399	0.0423	0.0697	0.1487	0.0856	0.0775	0.0475	0.0523	0.0087	0.0083	0.0063	0.0362	0.0225	0.0416	0.0307	0.032	0.0761	0.0805	1					
남측 유역	남원	순천	임실	장흥	광주	목포	여수	거창	구미	대구	문경	밀양	부산	산청	영주	영천	의성	진주	추풍령	합천	군산	금산	부여	청주	합계
	0.0463	0.0409	0.053	0.0194	0.0646	0.0214	0.0023	0.0449	0.0447	0.0555	0.0475	0.0729	0.0158	0.0431	0.0998	0.0501	0.0867	0.0466	0.0242	0.0476	0.05	0.0118	0.0023	0.0086	1

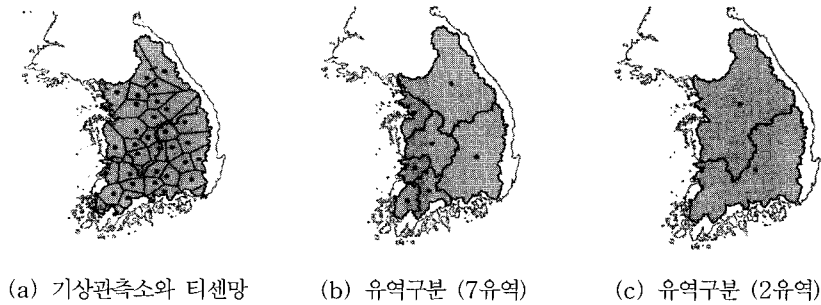
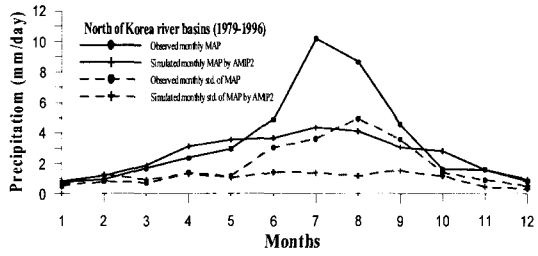
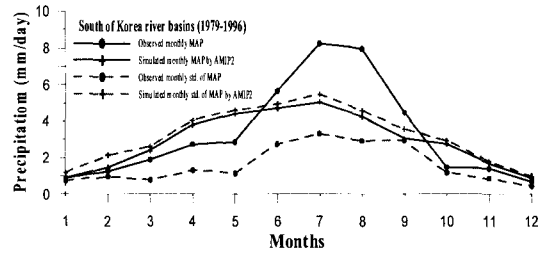


그림 9. 분석을 위한 유역구분

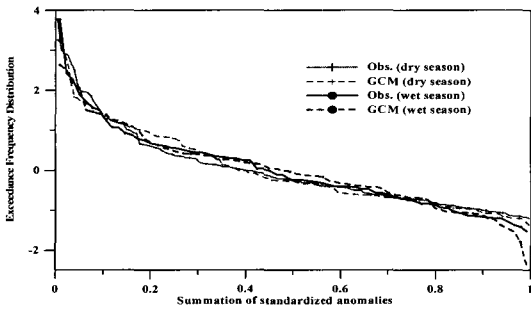


(a) 북측 하천 (한강, 금강, 안성천-삼교천)

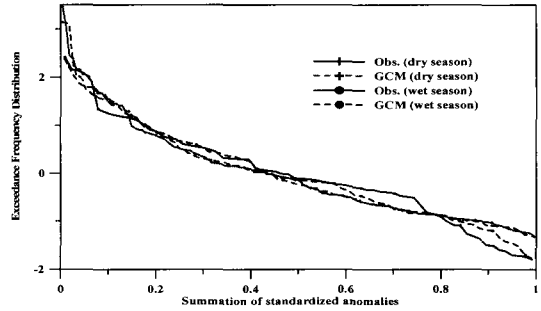


(b) 남측 하천 (낙동강, 영산강, 섬진강, 동진-만경강)

그림 10. 관측값과 모의값의 월 평균값과 월 표준편차 비교



(a) 북측 하천 (한강, 금강, 안성천-삼교천)



(b) 남측하천(낙동강, 영산강, 섬진강, 동진-만경강)

그림 11. 표준화된 관측값과 모의값의 초과확률함수

4. 결과 분석

본 연구에서는 국내 하천을 7개 유역으로 나누고, 각각의 유역에 대해 총 35개 지점 관측값들을 이용한 유역 면적평균강수량을 관측값으로 설정하였으며, 모의값으로는 ECMWF에서 AMIP type-2 형태로 모의한 2. × 2. 의 고해상도 GCM 모의 결과를 이용하여 한반도 주변 30개 격자점의 표면강수량 자료를 추출하여 선정하였다. 추출된 자료의 기간은 1979년~1996년의 18년 자료로 분석은 월간 자료를 이용하여 건기(10월~5월)와 우기(6월~9월)로 크게 구분하여 분석을 실시하였으며, 자료는 중앙값 미만과 중앙값 이상의 2등분법을 사용하였다. 또한 본 연구에서는 모의값에 의해 구분된 각 관측값의 평균을 계산하여 중앙값 미만의 평균이 중앙값 이상의 평균보다 크거나, 그 반대인 경우를 제외하는 평균비교분석을 동시에 실시하였다.

표 3은 각 격자점에 대한 모의값의 중앙값을 기준으로 한 구분의 결과를 나타내고 있고, 이러한 구분조건을 이용한 진단변수의 초기모의결과는 표 4~7과 같다. 이중 표 4~5는 평균비교분석을 고려하지 않은 경우의 결과를 나타내고 있으며, 표 6~7은 평균비교분석을 고

려한 결과를 나타내고 있다. 또한 표 8~9는 Monte Carlo 모의를 통해 관측값을 200회 모의 발생시켜 각 창 크기에 대해 건기와 우기로 구분하여 진단변수를 구한 후 이 값 가운데, 상위 10% 및 20% 유의수준에 해당하는 값들을 표현하고 있다. 표 8은 평균비교분석을 고려하지 않은 것이고, 표 9는 고려한 것이다. 본 연구에서 적용한 Monte Carlo 모의는 과거 기간에 발생한 사상이 임의로 해당 기간만큼 무작위로 추출되는 방법을 사용하였다. 그림 12와 13은 표 4~7의 결과를 표 8~9에서 제시한 유의수준과 비교하여 적절하다고 판단되는 유역을 경우별로 그림으로 나타낸 것이다. 그림에서 볼 수 있듯이 건기의 경우 평균비교분석을 적용하지 않은 경우 설정 창의 크기에 따라, 일부 유역에 대해 GCM 모의결과가 유효한 것으로 나타났지만, 평균비교분석을 고려한 경우 그림 13의 (a)와 (b) 같이 GCM 모의결과가 유효한 유역은 하나도 존재하지 않았다. 우기의 경우 평균비교분석의 적용 유무에 관계없이 대부분의 유역에서 유효함을 볼 수 있었다. 이를 다시 표현하면 ECMWF의 GCM 모의결과는 우기에 매우 유효함을 알 수 있으며, 유역단위의 수자원 운영에 적절히 사용되어질 수 있을 것으로 기대된다고 할 수 있다.

표 3. 각 격자점의 모의값을 이용한 관측값의 구분

	건기		우기	
	중앙값 미만 해당 연도	중앙값 이상 해당 연도	중앙값 미만 해당 연도	중앙값 이상 해당 연도
격자점 1	79,81,83,84,87,92,94,95,96	80,82,85,86,88,89,90,91,93	80,82,84,85,87,89,93,94,96	79,81,83,86,88,90,91,92,95
격자점 2	79,80,81,83,87,88,92,94,95	82,84,85,86,89,90,91,93,96	80,82,85,87,88,89,93,94,96	79,81,83,84,86,90,91,92,95
격자점 3	79,80,81,83,87,88,92,94,96	82,84,85,86,89,90,91,93,95	80,82,85,87,88,89,93,94,96	79,81,83,84,86,90,91,92,95
격자점 4	79,80,81,83,87,88,92,94,95	82,84,85,86,89,90,91,93,96	81,84,85,87,89,90,93,94,96	79,80,82,83,86,88,91,92,95
격자점 5	79,80,81,83,87,88,92,94,95	82,84,85,86,89,90,91,93,96	81,83,84,85,87,89,90,93,96	79,80,82,86,88,91,92,94,95
격자점 6	81,83,86,87,88,92,94,95,96	79,80,82,84,85,89,90,91,93	79,81,83,84,86,87,90,93,96	80,82,85,88,89,91,92,94,95
격자점 7	79,80,81,83,86,87,92,94,95	82,84,85,88,89,90,91,93,96	82,84,85,89,92,93,94,95,96	79,80,81,83,86,87,88,90,91
격자점 8	79,80,81,83,86,87,92,94,95	82,84,85,88,89,90,91,93,96	79,82,85,89,92,93,94,95,96	80,81,83,84,86,87,88,90,91
격자점 9	79,80,83,86,87,88,92,94,95	81,82,84,85,89,90,91,93,96	79,82,84,85,89,91,93,94,96	82,81,83,86,87,88,90,92,95
격자점 10	80,82,83,86,87,91,93,94,95	79,81,84,85,88,89,90,92,96	79,81,83,84,86,88,92,93,95	80,82,85,87,89,90,91,94,96
격자점 11	79,80,82,86,87,91,93,94,95	81,83,84,85,88,89,90,92,96	79,81,84,86,92,93,95,93,96	80,82,83,85,87,89,90,91,94
격자점 12	79,80,82,86,87,91,93,94,95	81,83,84,85,88,89,90,92,96	81,84,86,87,88,92,93,95,96	79,80,82,83,85,89,90,91,94
격자점 13	79,80,82,86,87,89,91,93,94	81,83,84,85,88,90,92,95,96	83,84,86,88,89,92,93,95,96	79,80,81,82,85,87,90,91,94
격자점 14	79,80,82,86,87,91,93,94,95	81,83,84,85,88,89,90,92,96	82,83,84,86,88,89,92,95,96	79,80,81,85,87,90,91,93,94
격자점 15	80,82,85,86,87,91,93,94,95	79,81,83,84,88,89,90,92,96	80,82,83,85,86,89,92,95,96	79,81,84,87,88,90,91,93,94
격자점 16	79,80,82,83,85,86,91,93,94	81,84,87,88,89,90,92,95,96	81,83,84,88,91,92,93,95,96	79,80,82,85,86,87,89,90,94
격자점 17	79,80,82,85,86,87,91,93,94	81,83,84,88,89,90,92,95,96	81,84,88,91,92,93,94,95,96	79,80,82,83,85,86,87,89,90
격자점 18	79,80,82,85,86,87,91,93,94	81,83,84,88,89,90,92,95,96	81,83,84,88,90,92,93,95,96	79,80,82,85,86,87,89,91,94
격자점 19	79,81,82,85,86,90,92,93,94	80,83,84,87,88,89,91,95,96	80,82,83,85,87,91,92,94,96	79,81,84,86,88,89,90,93,95
격자점 20	79,81,82,85,86,90,92,93,94	80,83,84,87,88,89,91,95,96	80,83,85,87,91,92,94,95,96	79,81,82,84,86,88,89,90,93
격자점 21	79,81,83,85,86,90,92,93,94	80,82,84,87,88,89,91,95,96	80,83,86,87,91,92,94,95,96	79,81,82,84,85,88,89,90,93
격자점 22	79,81,85,86,88,90,92,93,94	80,82,83,84,87,89,91,95,96	83,85,87,88,91,92,94,95,96	79,80,81,82,84,86,89,90,93
격자점 23	79,81,84,85,86,90,92,93,94	80,82,83,87,88,89,91,95,96	81,82,85,87,88,91,94,95,96	79,80,83,84,86,89,90,92,93
격자점 24	79,81,84,85,86,90,92,93,94	80,82,83,87,88,89,91,95,96	79,81,82,84,85,88,91,94,95	80,83,86,87,89,90,92,93,96
격자점 25	79,81,82,84,85,90,92,93,94	80,83,86,87,88,89,91,95,96	80,83,87,90,91,92,94,95,96	79,81,82,84,85,86,88,89,93
격자점 26	79,81,82,84,85,86,90,92,93	80,83,87,88,89,91,94,95,96	80,83,87,90,91,92,94,95,96	79,81,82,84,85,86,88,89,93
격자점 27	79,81,83,84,85,86,90,92,93	80,82,87,88,89,91,94,95,96	80,83,87,89,91,92,94,95,96	79,81,82,84,85,86,88,90,93
격자점 28	80,81,82,84,85,89,91,92,93	79,83,86,87,88,90,94,95,96	79,81,82,84,86,90,91,93,95	80,83,85,87,88,89,92,94,96
격자점 29	80,81,82,84,85,89,91,92,93	79,83,86,87,88,90,94,95,96	79,82,84,90,91,93,94,95,96	80,81,83,85,86,87,88,89,92
격자점 30	80,82,84,85,87,89,91,92,93	79,81,83,86,88,90,94,95,96	79,82,85,86,90,91,94,95,96	80,81,83,84,87,88,89,92,93

표 4. 진단변수값 : 건기-평균비교분석을 고려하지 않은 경우

진단변수(Ep)	한강	낙동강	금강	섬진강	영산강	안성-삼교	동진-만경
창 크기 $\lambda = 1.05$	0.6034	0.4853	0.4717	0.8795	0.6916	0.6034	0.9681
창 크기 $\lambda = 2.05$	0.5491	0.6476	0.5818	0.7840	0.4663	0.6590	0.6963

표 5. 진단변수값 : 우기-평균비교분석을 고려하지 않은 경우

진단변수 (Ep)	한강	낙동강	금강	섬진강	영산강	안성-삼교	동진-만경
창 크기 $\lambda = 1.05$	0.3408	0.5160	0.3408	0.3885	0.4717	0.2804	0.1491
창 크기 $\lambda = 2.05$	0.5039	0.4935	0.4241	0.3684	0.3855	0.5007	0.2122

표 6. 진단변수값 : 건기-평균비교분석을 고려한 경우

진단변수 (Ep)	한강	낙동강	금강	섬진강	영산강	안성-삼교	동진-만경
창 크기 $\lambda = 1.05$	1.0000	0.6030	0.6030	0.6030	0.6030	1.0000	0.9575
창 크기 $\lambda = 2.05$	0.7212	0.6030	0.8393	0.6030	0.6030	0.9575	0.9575

표 7. 진단변수값 : 우기-평균비교분석을 고려한 경우

진단변수 (Ep)	한강	낙동강	금강	섬진강	영산강	안성-삼교	동진-만경
창 크기 $\lambda = 1.05$	0.1352	0.4277	0.3408	0.3103	0.4717	0.2804	0.1155
창 크기 $\lambda = 2.05$	0.4319	0.4901	0.4237	0.3491	0.3835	0.5042	0.2087

표 8. Monte Carlo 모의 결과(평균비교분석을 고려하지 않은 경우)

창 크기	기간	진단변수	유의수준	한강	낙동강	금강	섬진강	영산강	안성-삼교천	만경-동진강
$\lambda = 1.05$	건기	Ep	10%	0.5017	0.5196	0.5176	0.5276	0.5305	0.505	0.5279
			20%	0.545	0.549	0.5517	0.5583	0.574	0.5472	0.5637
	우기	Ep	10%	0.4983	0.5225	0.493	0.4977	0.502	0.5006	0.4936
			20%	0.5302	0.5625	0.5314	0.5336	0.5312	0.5332	0.53
$\lambda = 2.05$	건기	Ep	10%	0.5681	0.5576	0.5628	0.5704	0.5562	0.5605	0.5647
			20%	0.5779	0.5703	0.5788	0.5845	0.5851	0.5775	0.5808
	우기	Ep	10%	0.5523	0.5655	0.5568	0.5495	0.5524	0.5575	0.5395
			20%	0.5727	0.5885	0.5759	0.5683	0.5757	0.5769	0.5662

표 9 Monte Carlo 모의 결과 (평균비교분석을 고려한 경우)

창 크기	기간	진단변수	유의수준	한강	낙동강	금강	섬진강	영산강	안성-삼교천	만경-동진강
$\lambda = 1.05$	건기	Ep	10%	0.4499	0.4728	0.4864	0.4892	0.4808	0.4707	0.4719
			20%	0.539	0.5311	0.5366	0.5434	0.5407	0.53641	0.527
	우기	Ep	10%	0.4802	0.5168	0.4508	0.4859	0.4597	0.4723	0.4255
			20%	0.5272	0.5513	0.5073	0.5341	0.5293	0.5253	0.4994
$\lambda = 2.05$	건기	Ep	10%	0.5367	0.5276	0.5586	0.5444	0.5438	0.5391	0.5388
			20%	0.5606	0.5518	0.5748	0.5614	0.5602	0.567	0.5613
	우기	Ep	10%	0.549	0.5545	0.5292	0.5342	0.5291	0.5527	0.5144
			20%	0.5819	0.5854	0.563	0.5655	0.5577	0.5793	0.5456

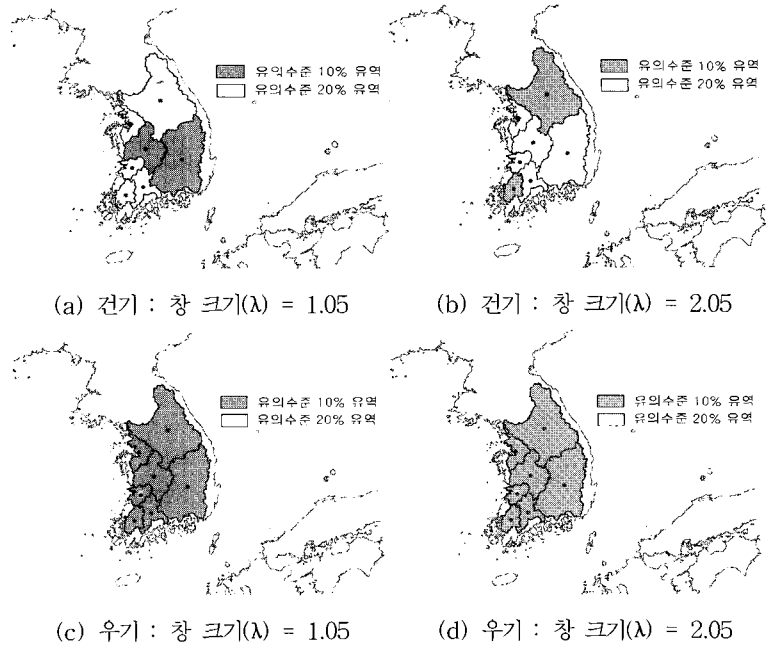


그림 12. GCM정보가 유효한 유역의 도시 (평균비교분석을 고려않은 경우)

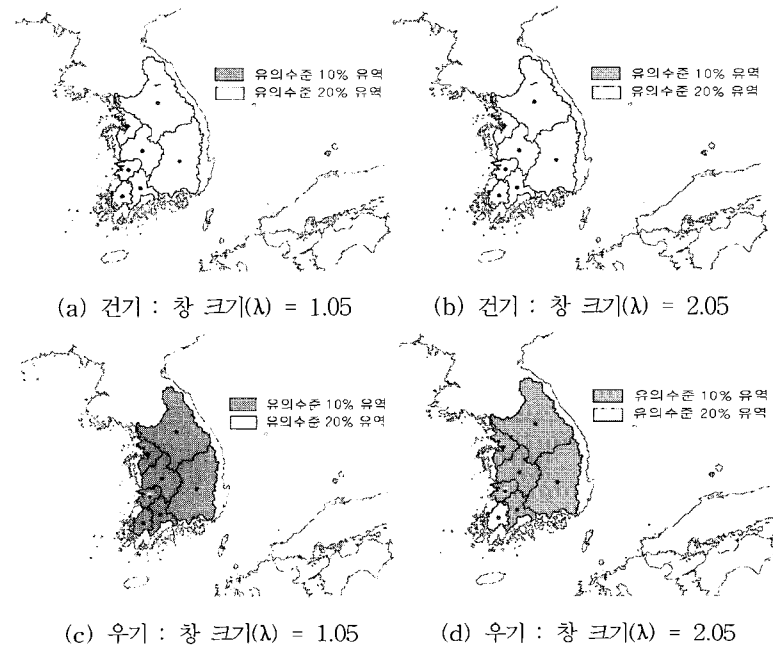


그림 13. GCM정보가 유효한 유역의 도시 (평균비교분석을 고려한 경우)

5. 결 론

본 연구는 GCM 모의결과를 이용하여 관측값을 구분해내는 확률론적 분석을 통해 수자원 활용 가능성을 타진하는 기상정보의 활용가능성에 관한 연구이다. 제안된 방법론을 통해 남한내의 7개 강 유역에 대해 1979년~1996년의 18년 자료를 이용하여 건기 및 우기에 대해 분석을 실시하였으며, 평균비교분석의 적용 유무에 따른 2가지 분석을 각각 실시하였다. 평균비교분석을 고려하지 않은 경우, 창 의 크기를 1.05로 설정하여 창

내부의 모든 격자의 평균을 이용하여 분석을 실시한 경우 건기의 경우 유의 수준 10%에 대해서는 낙동강 유역, 금강 유역의 정보가 적용 가능성이 크게 나타났으며, 창 크기를 2.05로 설정한 경우 유의 수준 10%에 대해서는 한강 유역, 영산강 유역의 정보의 이용가능성이 큰 것으로 나타났다. 우기의 경우 유의수준 10%에 대해서 모든 유역의 정보가 유효한 것으로 기대된다. 평균비교분석을 고려한 경우, 건기의 경우 창 크기에 관계없이 GCM 모의결과를 사용할 수 있는 유역이 없는 것으로 나타났으며, 우기의 경우 창 크기가 1.05인

경우 영산강 유역을 제외한 모든 유역의 정보를 유의 수준 10%내에서 적용 가능성이 높게 나타났으며, 영산강의 경우 유의수준 20%내에서 적용 가능성이 높게 나타났다. 또한 창 크기를 2.05로 설정한 경우 모든 유역의 정보가 유효한 것으로 나타났다. 다시 말해, 우기에는 7대강 유역에 대해 창 크기를 1.05 혹은 2.05로 설정할 경우, 유역내 ECMWF의 GCM 격자점의 평균화된 기상정보는 적용 가능성이 높다고 할 수 있으며, 향후 유역단위의 수자원 운영 등에서의 높은 활용 가능성을 나타낸다. 평균비교분석의 개념을 사용하지 않을 경우 건기에도 과반수 가량의 유역에 대해 GCM 모의 결과를 활용할 수 있을 것으로 나타났지만, 평균비교분석의 개념을 사용할 경우 어떤 유역에도 GCM 모의 결과를 활용할 수 없는 것으로 나타났다. 따라서 건기의 경우 ECMWF의 GCM 모의결과는 사용하기에 적절치 않은 것으로 판단된다. 이러한 결과를 바탕으로 다음과 같은 결론을 내릴 수 있을 것으로 판단된다. 모의값으로 사용된 ECMWF의 GCM 모의결과는 우기에 대해서는 향후 우리나라 7개 유역에 대해 인근한 격자점(window= 1.05, 혹은 2.05)에 대한 정보가 평균되어 사용되어질 경우 실제 상황을 비교적 유사하게 모의하여 유역단위로 운영되는 수자원 관리에 도움을 줄 수 있을 것으로 기대한다.

향후 우리나라와 같이 좁은 지역에 대해 본 방법론의 적용을 통한 보다 다양한 분석을 위해서는 좀더 조밀한 격자를 가진 GCM 모형이나, 기상청에서 운영중인 고해상도 GDAPS 모형을 이용한 분석이나, 기존 모형 결과에 대한 downscaling 기법의 적용 등이 필요하며, 앙상블 모의 결과에 대한 분석도 실시되어야 할 것으로 판단된다. 또한 모형의 분석 단위를 좀더 세분화한 계절에 대한 분석이나, 자료의 확충을 통한 좀더 세밀한 분석도 요구된다고 할 수 있다.

감사의 글

본 연구는 21세기 프론티어연구개발사업인 수자원의 지속적 확보기술개발사업단의 연구비지원(과제번호 1-3-1)에 의해 수행되었습니다.

참고 문헌

- Carpenter, T.M., and Georgakakos, K. P. (2001). "Assessment of Folsom Lake response to historical and potential future climate scenarios, I, Forecasting" *Journal of Hydrology*, Vol. 249, pp. 148~175.
- Georgakakos, K. P. (2003). "Probabilistic Climate-Model Diagnostics for Hydrologic and Water resources Impact Studies." *Journal of Hydro-meteorology*, Vol. 4, pp. 92~105.
- Georgakakos, K. P., Georgakakos, A. P., and Graham, N. E. (1998). "Assessment of benefits of climate forecasts for reservoir management in the GCIIP region." *GEWEX News*, Vol. 8, No. 3, pp. 5~7.
- Georgakakos, K. P., Graham, N. E., and Georgakakos, A. P. (2000). "Can Forecast Accrue Benefits for Reservoir Management? The Folsom Lake Study." *The Climate Report*, Vol. 1, No. 4, pp. 7~10.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical Recipes in FORTRAN, The Art of scientific Computing*. Cambridge University Press, New York, pp. 963.
- Yao, H., and Georgakakos, A. P. (2001). "Assessment of Folsom Lake response to historical and potential future climate scenarios, 2, Reservoir management." *Journal of Hydrology*, Vol. 249, pp. 176~196.

(논문번호:03-88/접수:2003.10.17/심사완료:2004.02.13)