

광역규모 예측인자를 이용한 한반도 계절 강수량의 장기 예측

김명기^{1,*} · 김화수¹ · 곽종흠¹ · 소선섭¹ · 서명석¹ · 박정규²

¹공주대학교 대기과학과, 314-701 공주시 신관동 182

²기상청 기후국 기후예측과, 156-720 서울시 동작구 신대방동 460-18

Long-term Forecast of Seasonal Precipitation in Korea using the Large-scale Predictors

Maeng-Ki Kim^{1,*} · Hwa-Su Kim¹ · Chong-Heum Kwak¹
Seon-Sup So¹ · Myoung-Seok Suh¹ · Chung-Kyu Park²

¹Department of Atmospheric Science, Kongju National University, Gongju, 314-701, Korea

²Climate Prediction Division, Climate Bureau, Korea Meteorological Administration,
Seoul 156-729, Korea

Abstract: A super ensemble model was developed for the seasonal prediction of regional precipitation in Korea using the lag correlated large scale predictors, based on the empirical orthogonal function (EOF) analysis and multiple linear regression model. The predictability of this model was also evaluated by cross-validation. Correlation between the predicted and the observed value obtained from the super ensemble model showed 0.73 in spring, 0.61 in summer, 0.69 in autumn and 0.75 in winter. The predictability of categorical forecasting was also evaluated based on the three classes such as above normal, near normal and below normal that are clearly defined in terms of a priori specified by threshold values. Categorical forecasting by the super ensemble model has a hit rate with a range from 0.42 to 0.74 in seasonal precipitation.

Keywords: seasonal prediction, super ensemble model, predictors, long range forecast

요약: 경험적 직교함수(EOF) 분석법과 다중회귀법에 기초하여 지연상관된 광역규모 예측인자로부터 3개월 이전에 계절 강수량을 예측할 수 있는 슈퍼앙상블 모델이 개발되었다. 이 모델의 예측성이 교차검증법에 의해 평가되었다. 관측 값과 예측값사이의 상관계수는 봄철에 0.73, 여름철에 0.61, 가을철에 0.69, 겨울철에 0.75로 나타났다. 이러한 값은 유의수준 $\alpha=0.001$ 에서 유의한 값이다. 슈퍼 앙상블 방법의 범주형 예측성이 3개 범주로 나누어진 사례에 대해서 평가되었다. 3개 범주는 계절 누적강수량의 상위 33.3%를 과우해, 하위 33.3%를 소우해, 그 나머지를 평년해로 구분하였다. 범주형 예측의 적중률은 계절에 따라 42%에서 74%로 나타났다.

주요어: 계절 예보, 슈퍼앙상블 모델, 예측인자, 장기예보

서론

최근 기상청에서 발표하고 있는 단기예보의 정확도는 상당한 수준에 이르렀으나 장기예보(1개월 이상)의 정확도는 높지 않아서 이를 높이기 위한 많은 연구가 필요하다. 강수에 대한 계절 예측은 수문, 산림, 식품, 의류 등을 포함한 사회 경제적 파급 효과가 매

우 크지만 아직 그 예측은 매우 어렵다. 특히 강수량은 여러개의 지역으로 구분될 만큼 지역에 따라 그 변동성이 다르다(허창희와 강인식, 1988; 문영수 외, 1995). 현재 수개월 예측은 통계적인 방법 및 역학모델에 기초한 연구가 동시에 활발하게 진행되고 있다. 최근에 Anderson *et al.*(1999)은 역할 모델의 모수화가 획기적으로 개선되지 않는 한 통계모델이 역학모델과 경쟁할만하거나 더 나은 성능을 가질 수 있다고 지적하였다.

*Corresponding author: mkkim@kongju.ac.kr

국내의 연구로는 예측변수보다 몇 개월 선행하는 광역규모를 예측인자를 사용하여 다중회귀모델을 구성한 후 장기예측을 시도한 연구들이 있었다(강인식 외, 1992; 강인식과 백희정, 1993; 이준이 외, 1999). 강인식 외(1992)는 장마기간 동안의 강수량을 예측하기 위하여 6, 7월보다 2~12개월 이전의 열대 태평양 해수면온도와 500 hPa 지오폠펠셜 고도장을 이용하여 예측인자를 찾고 이로부터 최적의 예측인자를 선정한 후 장기 예측을 시도하였다. 이준이 외(1999)는 기존의 방법을 보완한 단계적 회귀방법(stepwise regression)을 사용하여 한반도 봄철 기온의 예측하였으며, 0.6 정도의 기능평점(skill score)을 얻었다. 특히, 엘니뇨 해와 평년으로 구분하여 구축한 예측 모형에서 기능평점은 0.78로 향상되었다. 회귀분석의 또 다른 형태로 주성분분석(principal component analysis)을 이용하여 광역규모 예측인자로부터 장기 예측을 시도한 연구들도 있다(Shabbar and Barnston, 1994; He and Barnston, 1996; Hwang *et al.*, 2001).

현재 일본은 역학모델을 이용한 1개월 역학예보, 통계모델을 이용한 3개월 이상의 예보가 생산되고 있으며, 미국의 역학모델은 주로 6개월 예보에 중점을 두고 개발중에 있다. 우리 나라의 경우 3개월까지의 기상전망을 생산하고 있다. 따라서 향후 3개월 이상의 기간에 대해서 다양한 예측방법을 연구해야 할 필요가 있다. 앞서의 연구들은 대부분 예측의 대상이 한반도 평균에 대한 것이었다. 한반도는 기후적인 관점에서 볼 때 비록 같은 기후대임에는 틀림이 없으나 지역적으로 상당히 다른 기후적 특성을 나타낸다. 따라서 장기 예측을 할 때 지역에 대한 특성을 예측할 수 있는 방법의 개발이 필요하다. 이 연구에서는

Kim *et al.*(2002)의 방법을 따라 한반도 계절 강수량을 한 계절 이전에 예측하였으며 그 예측성을 조사하였다.

연구 방법

한반도 관측자료는 12개 관측소의 계절 누적 강수량이며 1954년부터 1998년까지 45년간의 자료가 이용되었다(Table 1). 광역규모 예측인자는 해면기압(SLP: Sea Level Pressure), 850 hPa 지오폠펠셜고도와 SST(Sea Surface Temperature) 자료가 사용되었다. 이 자료는 미국대기과학연구소(NCEP: National Center for Atmospheric Research) 재분석 자료이며, 해상도는 위도 2.5도와 경도 2.5도이다. 여기서 예측변수는 계절 누적 강수량이고, 예측인자는 월별 지오폠펠셜고도, 해수면온도, 해면기압이다.

이 연구에서 사용한 장기 예측모델 구축 방법은 Kim *et al.*(2002)의 방법과 동일하다. 이를 간단히 요약하면 다음과 같다. 우선 계절 누적 강수량에 경험적 직교함수 분석 EOF(Empirical Orthogonal Function)을 적용하여 주요 변동 모드를 추출한다. 이때 추출된 EOF의 시간 계수는 한반도 강수량의 주요 시간 변동을 설명하는 일종의 대리자 역할을 한다. 따라서 이 연구에서는 한반도 강수량을 사용하지 않고 이 시간계수(TC)를 사용하였다. 그 이유는 EOF의 특성상 이 시간계수와 고유벡터를 선형 결합시키면 한반도 강수량의 공간 분포를 알 수 있기 때문이다(EOF 복원). 강수의 경우에는 처음 2개의 모드가 전체 변동의 약 63%(여름철)에서 84%(겨울철)를 설명하기 때문에 이 연구에서는 장기 예측 모델의 구

Table 1. List of the surface observation stations.

Station No.	Station	Latitude (N)	Longitude (E)	Height (meter)
159	Busan	25° 06'	129° 02'	69.2
165	Mokpo	34° 49'	126° 23'	37.9
108	Seoul	37° 34'	126° 58'	85.5
143	Daegu	35° 53'	128° 37'	57.6
105	Gangneung	37° 45'	128° 54'	25.9
146	Jeonju	35° 49'	127° 09'	53.5
156	Gwangju	35° 10'	126° 54'	70.5
168	Yeosu	34° 44'	127° 45'	66.1
152	Ulsan	35° 33'	129° 19'	34.7
112	Incheon	37° 28'	126° 38'	68.9
135	Chupungnyeong	36° 13'	128° 00'	242.5
138	Pohang	36° 02'	129° 23'	1.9

측에 처음 두 개의 모드를 사용하였다.

모델 구축의 첫 단계는 시간계수와 예측인자사이의 지연상관 관계를 조사하는 것이다. 시간 지연은 3개월 이상을 사용하였다. 즉, 3개월 이상의 시간 지연을 갖으면서 가장 상관성이 높은 상위 3개의 예측인자를 선정한다. 다음으로 가장 높은 상관성을 갖는 예측인자로부터 단순회귀식에 의해 예측한 값과 관측으로 얻은 시간계수사이의 차이를 구한다. 이 차이를 오차 시간계수(ΔTC)라 할 수 있다. 여기서 n 는 각 모드에 해당한다. 이 오차를 보정하기 위하여 TC 와 같은 절차를 거쳐 새로운 예측인자를 추출한다. 이 과정으로부터 첫 번째 모드와 연관된 두 개의 예측인자를 얻게 된다. 다음으로 두 번째 모드에 대해서도 같은 방법을 적용시키면 두 개의 예측인자를 더 얻게 된다. 따라서 4개의 서로 다른 예측인자로부터 두 모드에 대한 시간계수를 추정할 수 있게 되고, 이를 각각 EOF 복원시키면 각 모드와 연관된 강수의 공간 분포를 얻게되며, 이 두 공간분포를 합하면 각 관측소에 대한 예측된 공간 분포를 얻게된다. 이러한 과정을 모두 포함시켜 LR 모델이라 명명하기로 한다.

같은 방법으로 두 번째, 세 번째 상관성이 높은 예측인자를 이용하여 한반도 강수량을 예측할 수 있다. 이러한 방법으로 각 관측소에 대해 3개의 예측값을 얻게 되는데, 예측의 안정성과 예측성을 높이기 위하여 이 3개의 예측값에 서로 다른 가중치를 주어 최적의 예측값을 얻게 된다. 이를 Kim *et al.*(2002)에서는 수퍼 앙상블(super ensemble) 계절 예측이라 명명하였다. 또한 이 연구에서는 예측인자를 선정할 때 지연상관계수의 통계적 유의성 여부를 판단하여 유의하면 예측인자로 선정하였다. 상관계수에 대한 통계적 유의성은 양측 t 검정을 이용하였다(김우철 등, 1999).

결과 및 토의

계절 강수량의 시공간 변동

한반도 봄철과 가을철 강수량은 10년당 약 9.2 mm, 5.6 mm 정도 감소하였으며, 여름철과 겨울철에는 10년당 약 21.9 mm, 0.9 mm 증가하였다. 그러나 어느 계절의 경향도 유의한 수준은 아니었다. 따라서 이 연구에서는 장기 예측모델을 구축하기 위하여 사전에 선형 경향을 제거하지 않았다. 그러나 기온의 경우는 선형 경향성이 분명하기 때문에 분석

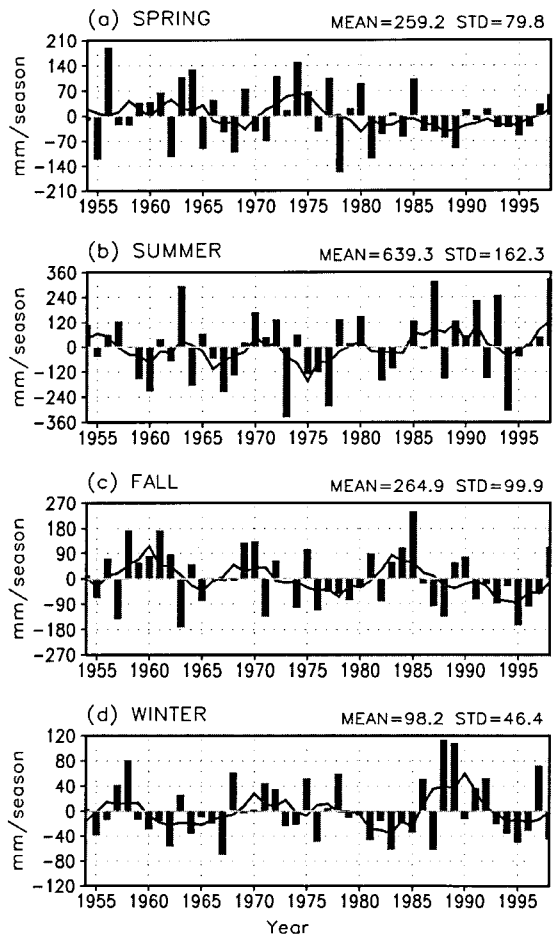


Fig. 1. Time series of the averaged seasonal-mean (bar) and 5-year running mean (solid line) precipitation of 12 stations in Korea. (a) Spring, (b) Summer, (c) Fall, (d) Winter.

에 앞서 보통 선형경향을 제거한다(Kim *et al.*, 2002). Fig. 1의 상단에 누적 강수량의 평균(MEAN)과 표준편차(STD)가 제시된 것처럼, 계절 누적 강수량의 경년변동은 여름철에 가장 크며, 겨울철에 가장 적다. 그러나 평균값에 대비한 표준편차(변동계수)는 그 반대이다. 봄철은 과거보다 최근 20여년 동안에 변동성이 상당히 줄어든 것을 알 수 있으며, 대체로 평균보다 적은 강수가 빈번하게 나타났다. 다른 계절의 경우는 경년변동이 두드러지게 나타나고 있는데 비해서 겨울철의 경우는 약 15년 정도의 주기가 경년 변동보다 더 두드러지게 나타난다.

경험적 직교함수 분석은 보통 어떤 변수의 시공간 변동 특성을 몇 개의 주요 모드로 설명하기 위하여 이용된다. 보통 기온의 경우 첫 번째 모드가 전체 변

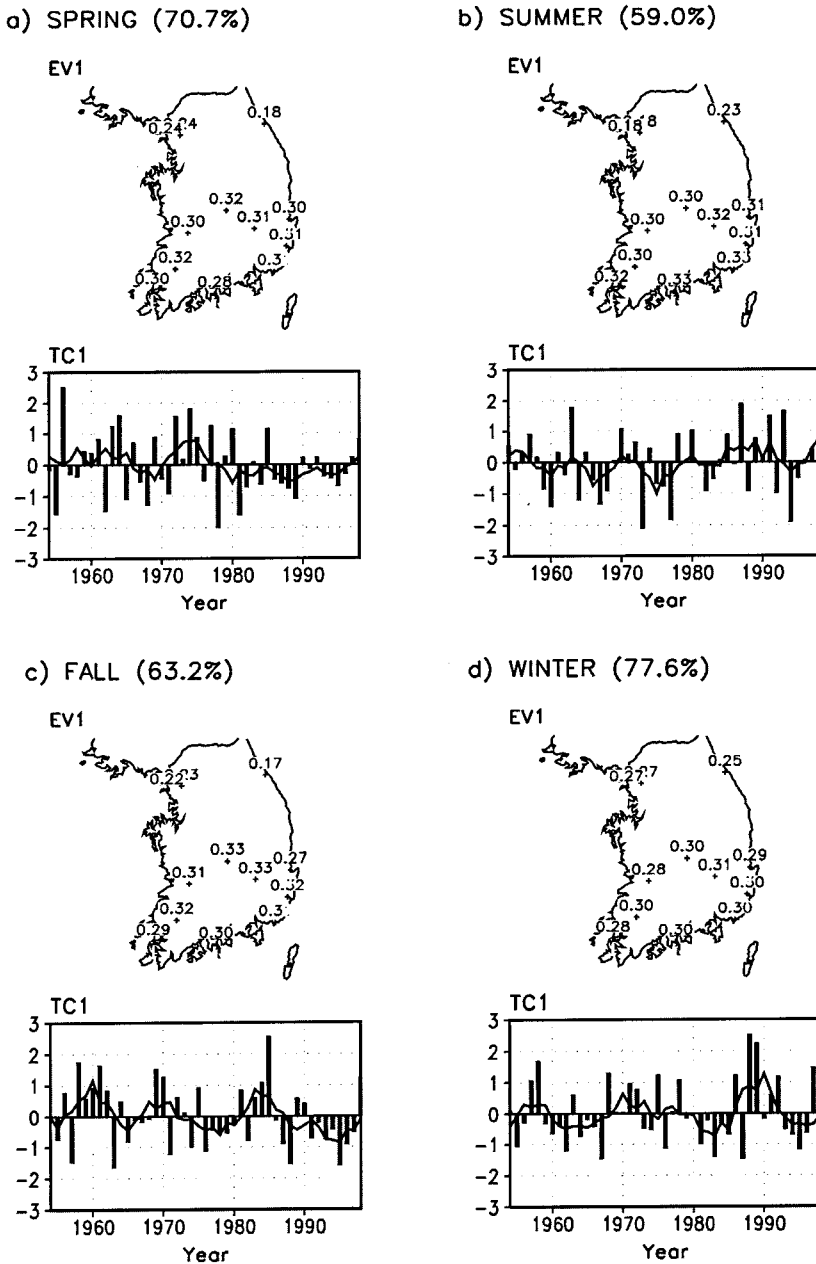


Fig. 2. Eigen vector (upper) and associated time coefficient (lower) of the first EOF mode for seasonal precipitation of (a) spring, (b) summer, (c) fall and (d) winter. The year to year and 5-year running mean time series are indicated by bar and solid line, respectively.

동의 80~90% 정도를 설명하지만(Kim *et al.*, 2002), 강수의 경우는 첫 번째 모드가 전체의 약 60~70% 정도를 설명하고 두 번째 모드가 약 10% 내외를 설명한다(Fig. 2). 첫 번째 모드에서 보여주는 시간계수의 시계열은 한반도 평균 시계열과 거의 같은 변동

을 보여준다(Fig. 2와 Fig. 1). 이는 첫 번째 모드가 통상 한반도의 평균적인 변동을 대변하는 모드라는 것을 의미한다. 또한 고유벡터가 모든 관측소에서 같은 부호를 나타내고 있는 것으로 보아서 적어도 첫 번째 모드만을 놓고 해석해 본다면 한반도는 같은

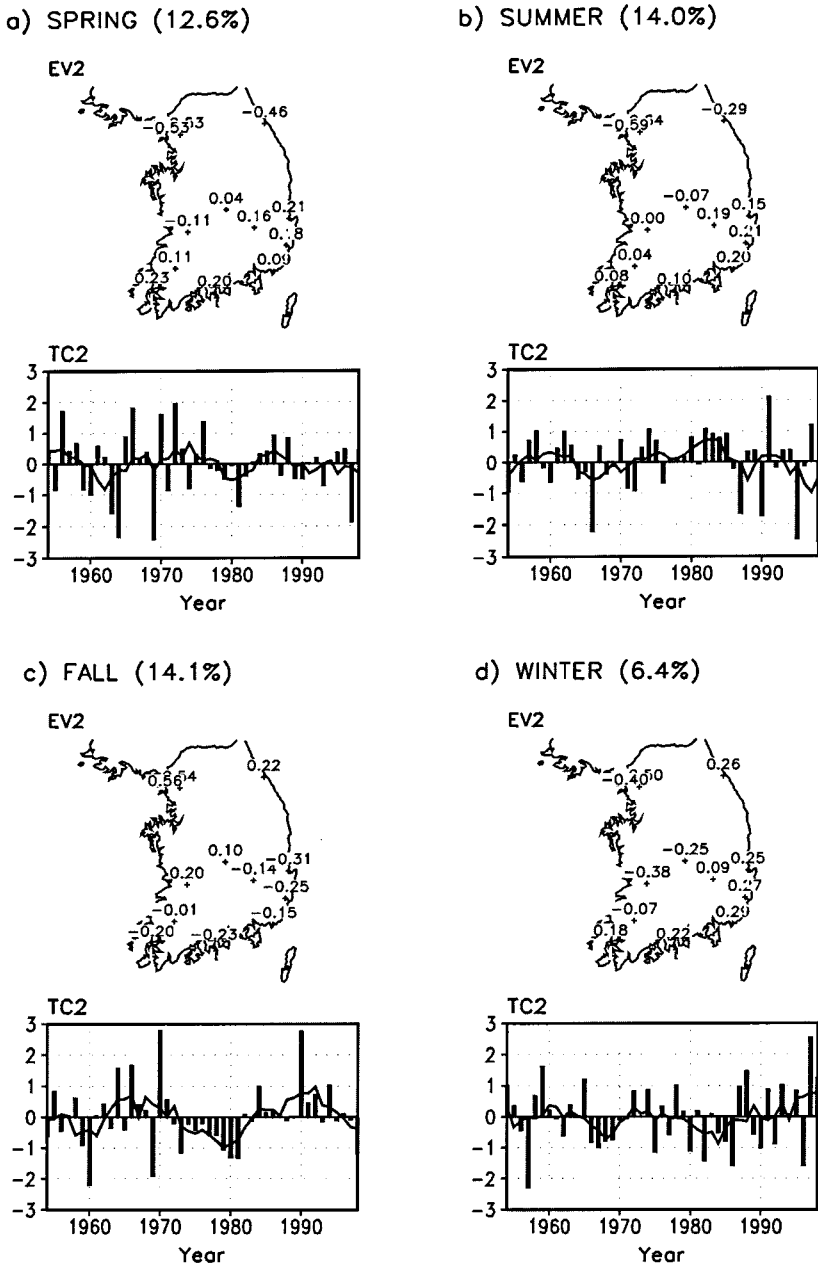


Fig. 3. Eigen vector (upper) and associated time coefficient (lower) of the second EOF mode for seasonal precipitation of (a) spring, (b) summer, (c) fall and (d) winter. The year to year and 5-year running mean time series are indicated by bar and solid line, respectively.

방향의 변동을 한다. 다시 말하면 평년보다 비가 적게 오는 해에는 모든 지역에서 비가 적게 온다. 다만 고유벡터의 부호는 같지만 크기가 다른 것으로 보아 편차의 크기는 지역적으로 차이가 있다. 두 번째 모드의 경우는 계절에 따라 전체변동의 약 6~14% 정

도를 설명한다(Fig. 3). 고유벡터의 크기가 첫 번째 모드와 달리 지역에 따라 부호가 크기가 다르다. 이는 이 모드가 한반도 내의 국지적인 강수량 변동을 설명하는 모드라는 것을 의미한다. 특히 겨울철의 고유벡터에서 동서 패턴의 양상을 나타낸다. 특히 태백

Table 2. The lagged correlation between the predictors and the surrogate of seasonal temperature in Korea. The surrogates are the time coefficients (TC) and the error time coefficients (ΔTC). Rank indicates three cases with the highest correlation.

Season	Rank	Time coefficient	Predictors	Location (lat, lon)	Correlation coefficient
Spring	1st	TC_1	SST of September	57N,172W	0.55
		ΔTC_1	SLP of September	15N,127.5E	-0.52
		TC_2	SST of December	33N,132W	0.54
		ΔTC_2	SST of May	41N,150E	0.58
	2nd	TC_1	SLP of March	50S,32.5W	-0.54
		ΔTC_1	850mb GPH of July	27.5N,112.5W	0.50
		TC_2	SLP of May	60N,75E	0.48
		ΔTC_2	SST of June	17N,114W	0.54
	3th	TC_1	850mb GPH of September	37.5N,10E	0.54
		ΔTC_1	SST of June	19N,148E	0.59
		TC_2	850mb GPH of September	60N,135W	-0.47
		ΔTC_2	850mb GPH of July	50N,135E	0.54
Summer	1st	TC_1	SST of October	11N,138E	-0.54
		ΔTC_1	SST of March	35S,26E	-0.55
		TC_2	SLP of July	70N,80E	-0.56
		ΔTC_2	SST of August	35S,116E	0.61
	2nd	TC_1	SLP of July	12.5S,75W	-0.48
		ΔTC_1	SST of October	45N,16W	-0.54
		TC_2	SST of February	45S,104W	0.53
		ΔTC_2	SLP of July	32.5N,122.5W	0.54
	3th	TC_1	850mb GPH of October	15S,100W	-0.47
		ΔTC_1	SST of November	43N,22W	-0.53
		TC_2	850mb GPH of July	70N,75E	-0.52
		ΔTC_2	SST of August	35S,116E	0.57

산맥에서 소백산맥으로 이어지는 경계선을 중심으로 동서 또는 남북간에 부호가 다르다. 반면 여름철과 가을철의 경우는 남북패턴이 두드러지게 나타난다 (Fig. 3b, c).

예측인자의 선정과 시간계수의 예측성

앞서 Fig. 2와 3에서 제시한 고유벡터의 두 모드는 한반도의 강수량 변동의 대부분 설명하는 모드이다. EOF의 특성상 시간계수를 예측할 수 있다면 예측한 시간계수와 고유벡터의 선형결합에 의해 관측소별 강수량의 예측 값을 얻을 수 있다. 시간 계수를 예측하기 위하여 시간계수와 예측인자간의 지연(lag) 상관관계를 조사하였다. 특히 이 연구의 관심이 한 계절 앞서서 계절 강수량을 예측하는 것이므로 시간지연의 월수를 3개월 이상으로 하였다. Table 2와 3은 각 계절 누적 강수량에 대한 EOF 시간계수와 예측인자사이의 상관관계를 나타낸 것이다. Rank에 표시되어 있는 1st, 2nd, 3th 는 지연상관계수가 가장 큰 상위 3개의 예측인자를 나타낸다. ΔTC 와 TC 에 있는 첨자는 1(2)인

경우 첫(두) 번째 모드의 시간계수를 의미한다. 봄철 강수량에 대한 TC_1 은 전년 9월의 북태평양(57°N, 172°W) SST와 약 0.55의 상관관계를 보였으며, TC_2 는 전년 12월 북태평양 동부 해역(33°N, 132°W)의 해면 기압과 -0.54의 상관관계를 보였다. 반면 여름철 강수량에 대한 TC_1 은 전년 10월에 서태평양(11°N, 138°E)의 해수면온도 변동과 -0.54의 상관관계를 보였으며, TC_2 는 전년 7월의 북극해역(70°N, 80°E)의 해면기압과 -0.56의 상관관계를 보였다. 가을철 강수량에서 TC_1 은 같은해 2월의 남 대서양(52.5°S, 35°W) 해역의 해면 기압 변동과 약 0.62의 높은 상관관계를 보였다. 반면 두 번째 모드인 TC_2 는 같은해 5월의 북극해 주변(75°N, 120°E)의 850 hPa GPH 변동과 약 0.55의 상관관계를 나타냈다. 겨울철 강수량에 대한 TC_1 은 6월 북대서양 동안의 850 hPa GPH 변동과 약 0.58의 상관관계를 나타냈다. 유의수준 $\alpha=0.001(0.1\%)$ 에 대한 상관계수는 0.47이다. Table 2와 3에 제시한 인자들은 모두 이 조건을 만족하는 인자들이다. 한반도 강수량과 연관된 예측인자는 태평양, 북대서양, 북극, 남극 등

Table 3. The lagged correlation between the predictors and the surrogate of seasonal temperature in Korea. The surrogates are the time coefficients (*TC*) and the error time coefficients (ΔTC). Rank indicates three cases with the highest correlation.

Season	Rank	Time coefficient	Predictors	Location (lat, lon)	Correlation coefficient
Fall	1st	<i>TC</i> ₁	SLP of February	52.5S,35W	0.62
		ΔTC ₁	SLP of June	40S,52.5W	-0.49
		<i>TC</i> ₂	850mb GPH of May	75N,120E	0.55
		ΔTC ₂	SST of December	29N,32W	0.57
	2nd	<i>TC</i> ₁	850mb GPH of February	52.5S,35W	0.61
		ΔTC ₁	SLP of June	40S,52.5W	-0.48
		<i>TC</i> ₂	SLP of May	75N,115E	0.54
		ΔTC ₂	SLP of January	27.5N,120W	-0.55
	3th	<i>TC</i> ₁	SST of April	11S,118E	-0.56
		ΔTC ₁	850mb GPH of May	30N,82.5E	0.55
		<i>TC</i> ₂	SST of September	39N,44W	-0.49
		ΔTC ₂	850mb GPH of April	62.5S,175E	0.57
Winter	1st	<i>TC</i> ₁	850mb GPH of June	42.5N,0	0.58
		ΔTC ₁	850mb GPH of April	45N,45E	-0.55
		<i>TC</i> ₂	SST of December	13N,46E	-0.54
		ΔTC ₂	SST of September	75N,90W	-0.52
	2nd	<i>TC</i> ₁	SLP of June	40N,0	0.57
		ΔTC ₁	SLP of July	25N,62.5W	-0.53
		<i>TC</i> ₂	SLP of January	32.5S,160E	-0.51
		ΔTC ₂	850mb GPH of August	35N,17.5W	0.56
	3th	<i>TC</i> ₁	SST of May	23S,158W	-0.52
		ΔTC ₁	SLP of December	50N,85W	0.60
		<i>TC</i> ₂	850mb GPH of April	47.5N,45E	0.49
		ΔTC ₂	SST of March	45S,78E	-0.54

다양한 지역에 분포하고 있다. 특히, 극 지방의 변동은 최근 부각되고 있는 AO(Arctic Oscillation)이나 AAO (Antarctic Oscillation)에 의해 대변되는 전구 규모의 변동이다. 한반도의 기후변동성도 이러한 극진동(AO), 북대서양 진동(NAO) 등과 상당히 관련되어 있는 것으로 보인다.

앞서 제시한 예측인자를 사용하여 예측인자와 *TC*₁의 예측성을 조사하였다. 예측 표본의 수를 증가시키기 위하여 교차검증법(cross-validation method)을 사용하여 총 45개의 예측 표본을 얻었다(Michaelson, 1987). 첫 번째 모드에 대한 시간계수인 *TC*₁의 경우 관측 값(가로축)과 예측값(세로축)간의 상관계수는 계절에 따라 약 0.64에서 0.68까지의 값을 나타냈다(Fig. 4). 두 번째 모드인 *TC*₂의 경우, 봄철에 약 0.68, 여름철에 0.68, 가을철과 겨울철에 각각 0.62와 0.65로 비교적 높은 상관관계를 보였다.

계절 강수량의 예측성

해당 계절보다 3개월 이전의 예측인자를 이용하여

12개 관측소별 계절 강수량을 예측하였다. 편의상 예측값과 관측값의 상관관계를 한반도 평균에 대해 제시하였다(Table 4). 상관성이 가장 높은 3개의 예측인자 중에서 첫 번째 예측인자를 사용한 *LR*₁ 모델의 예측 강수량은 관측값과 약 0.6 내외의 상관관계를 보였다. 특히일년중 강수량이 가장 적은 겨울철에 0.63의 가장 높은 상관계수를 보였다. 두 번째 예측인자를 사용한 *LR*₂ 모델의 예측 강수량은 *LR*₁ 모델보다 다소 낮은 상관관계를 보였다. 특히, 겨울철 강수량의 예측성은 모든 모델에서 가장 높았으며, 여름철 강수량에서는 가장 낮았다. 반면 슈퍼 앙상블 방법에 의한 예측은 여름철에 0.61에서 겨울철에 0.75의 상관계수를 나타냈다(Fig. 5). 3개의 예측값에 서로 다른 가중치를 주어 예측하는 슈퍼앙상블 방법에 의한 예측은 각각의 *LR*_{*i*} 모델에 의한 예측보다 개선되었다. 슈퍼앙상블 모델은 *LR*_{*i*} 모델보다 약 0.10~0.15 정도 상승된 상관계수를 보였다. 또한 여름철과 가을철 예측보다 봄과 겨울철 예측이 더 많이 개선되었다. 여름철 강수량의 예측성이 다른 계절에 비해

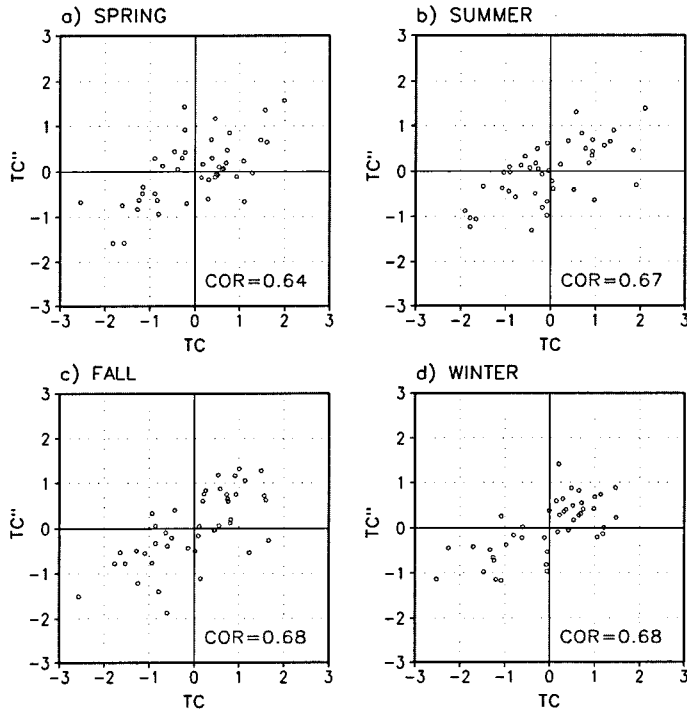


Fig. 4. Scatter diagram of the observed and the predicted of the first EOF time coefficient for (a) spring, (b) summer, (c) fall and (d) winter. The correlation coefficients between two time series are shown in the right lower part of each figure.

낮은 원인은 변동성이 가장 크기 때문으로 보인다.

슈퍼 양상불 예측의 범주형 예측성이 평가되었다 (Table 5). 이를 위하여 관측 강수량을 3개의 범주로 나누었다. 우선 45년간의 계절별 누적 강수량을 크기 순으로 나열한 후 3등분하였고, 이 중에서 상위 15년은 다우해(Wet), 하위 15년은 소우해(Dry), 중간 15년은 평년(Normal)으로 정의하였다. 만약 예측에 대한 아무런 정보 없이 매년 평년과 같다고 예측한다면 맞을 확률은 33.3%가 된다. Table 5에서 관측값의 각 범주에 대해서 예측값은 3개의 범주 중에서 어느 하나의 범주로 예측을 하게 된다. 강수량에 대한 슈퍼 양상불 방법의 예측성은 모든 계절에 33% 보다는 훨씬 높은 51%에서 74%까지의 적중률(hit rate)을 보였다. 소우해에 대한 적중률은 약 51~56% 정도로 계절에 따른 차이가 적었고, 다우해에 대한 적중률은 42~74%로 나타났다. 특히 여름철에 42%로 낮고, 나머지 계절은 0.7 이상의 높은 적중률을 보였다. 평년과 유사한 해에 대한 적중률은 51~63%로 나타났다. 슈퍼양상불 예측의 또 하나의 두드러진 특징은 극단적으로 틀릴 확률이 상당히 적다는 점이다. 즉, Dry인 경우를 Wet로 예측할 확률은 1~7%이고,

Wet인 경우를 Dry로 예측할 확률은 2~5% 정도에 불과하다.

회귀계수의 안정성을 평가하기 위하여 교차검증에서 얻은 회귀계수의 평균과 표준편차를 조사하였다. 그 결과 전체 회귀계수의 92%에서 표준편차는 평균의 약 10% 이내를 보였다. 이는 통계 모델을 구축하는 기간이 다르더라도 회귀계수는 크게 다르지 않다는 것을 말한다. 회귀모델에서 다중공선성(multicollinearity) 관계가 있으면 회귀계수를 구하기가 어렵고, 회귀계수의 분산이 커지며, 회귀계수의 신뢰구간의 폭이 커져서 구간추정의 의미를 잃게된다(박성현, 1998). 회귀계수가 안정된 값을 갖는다는 것은 다중공선성 관계가 문제가 되지 않는다는 것을 나타낸다.

결론

기존 계절에 대해서 3개월 이전의 광역규모 예측인자를 이용하여 한반도의 계절 강수량을 예측할 수 있는 슈퍼 양상불 모델이 개발되었다. 이 모델은 경험적 직교함수(EOF) 분석법과 다중회귀법에 기초하

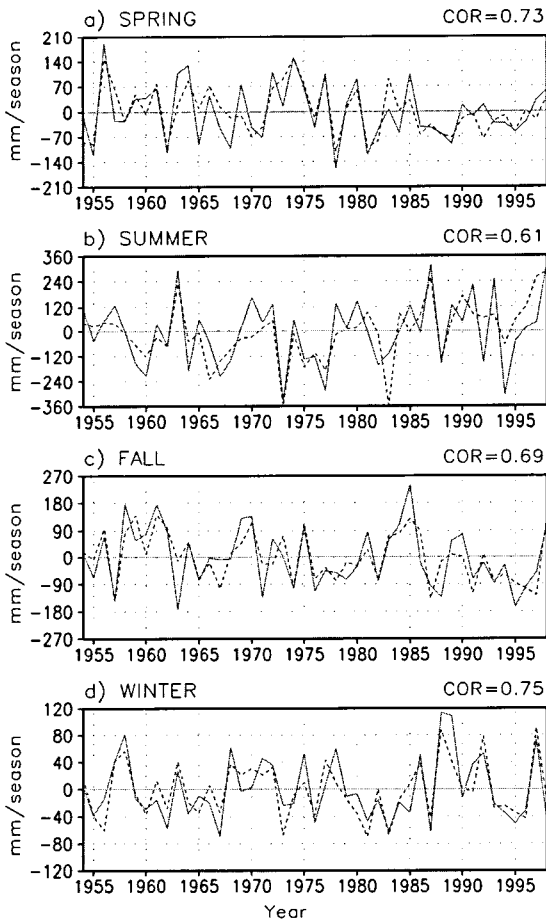


Fig. 5. Time series of the observed (solid line) and the predicted (dashed line) for (a) spring, (b) summer, (c) fall and (d) winter. The prediction was obtained from super ensemble based on the multiple regression analysis.

여 개발되었으며, 모델의 근간은 Kim *et al.*(2002)에 기초하고 있다. 개발된 모델의 예측성을 평가하기 위하여 교차 검증법이 이용되었다.

이 연구에서 개발한 슈퍼앙상블 모델의 특징은 기

Table 4. Correlation skills of LR series and super ensemble model. LR series indicate the predictions of linear regression model from the lagged large-scale predictors (1st, 2nd, and 3th) shown in Table 2.

Season	LR1	LR2	LR3	Super Ensemble
Spring	0.58	0.56	0.59	0.73
Summer	0.55	0.51	0.51	0.61
Fall	0.59	0.58	0.59	0.69
Winter	0.63	0.61	0.61	0.75
mean	0.59	0.57	0.58	0.70

존의 연구에 몇 가지 점에서 차이를 보인다. 첫째는 경험적 직교함수 분석에 의해 얻은 시간 계수를 이용하여 예측인자를 찾지 때문에 시간계수를 예측하면 고유벡터와 시간계수의 선형결합에 의해 한반도 지역 강수량을 예측할 수 있다는 점이다. 두 번째는 각 시간계수가 모드별로 추출되므로 각 모드별로 예측인자를 찾을 수 있고, 그로 인해 보다 많은 변동성을 예측할 수 있다는 점이다. 셋째는 상관성이 높은 상위 3개의 예측인자를 이용하여 예측한 후에 각 예측 값에 서로 다른 가중치를 주어 보다 높은 예측성을 확보했다는 점이다. 관측값과 예측값사이의 상관계수는 봄철에 0.73, 여름철에 0.61, 가을철에 0.69, 겨울철에 0.75로 나타났다. 이러한 값은 유의수준 $\alpha=0.001$ 에서 유의한 값이다.

슈퍼 앙상블 방법의 범주형 예측성이 3개 범주에 대해 나누어진 사례에 대해서 평가되었다. 3개 범주는 계절 누적강수량의 상위 33.3%를 다우해, 하위 33.3%를 소우해, 그 나머지를 평년해로 구분하였다. 강수량에 대한 슈퍼 앙상블 방법의 예측성은 모든 계절에 33% 보다는 훨씬 높은 51%에서 74%까지의 적중률(hit rate)을 보였다. 소우해에 대한 적중률은 약 51~56% 정도로 계절에 따른 차이가 적었고, 다우해에 대한 적중률은 42~74%로 나타났다. 특히 여름

Table 5. The hit rate of categorical forecast for the three classes such as above normal (W), near normal (N) and below normal (C). Obs and Fct indicate observation and prediction, respectively.

Season	Obs Fct	D			N			W		
	D	N	W	D	N	W	D	N	W	
Spring	0.56	0.41	0.03	0.21	0.61	0.18	0.05	0.23	0.72	
Summer	0.51	0.42	0.07	0.13	0.63	0.24	0.03	0.55	0.42	
Fall	0.52	0.43	0.04	0.23	0.52	0.24	0.02	0.25	0.73	
Winter	0.53	0.47	0.01	0.23	0.51	0.26	0.03	0.23	0.74	
Mean	0.53	0.43	0.04	0.20	0.57	0.23	0.03	0.32	0.65	

철에 42%로 낮고, 나머지 계절은 70% 이상의 높은 적중률을 보였다. 평년과 유사한 해에 대한 적중률은 51~63%로 나타났다. 슈퍼양상블 예측의 또 하나의 두드러진 특징은 극단적으로 틀릴 확률이 상당히 적다는 점이다. 즉, Dry인 경우를 Wet로 예측할 확률은 1~7%이고, Wet인 경우를 Dry로 예측할 확률은 2~5% 정도에 불과하다.

사 사

이 연구는 농림부에서 주관하는 “농림기술 개발 사업 제4차 기혁연구과제(과제번호: 19990543)”의 지원을 받았습니다.

참고문헌

강인식, 백희정, 1993, 한반도 겨울철 월평균 기온의 장기 예측. 한국기상학회지, 29, 253-262
 강인식, 허창희, 민경덕, 1992, 한반도 여름철 강수량의 장기예측. 한국기상학회지, 28, 283-292
 김우철, 김재주, 박병욱, 박성현, 송문섭, 이상열, 이영조, 전종우, 조신섭, 2000, 현대통계학, 영지문화사, 337 p.
 문영수, 이병근, 권혜영, 1995, 강수량의 장기변동에 의한 한국의 지역구분. 한국지구과학회지, 16, 38-43.
 박성현, 1998, 회귀분석, 민영사, 666 p.

이준이, 강인식, 허창희, 1999, 통계적 방법을 이용한 한반도 봄철 기온의 장기 예측. 한국기상학회지, 35, 372-383.
 허창희, 강인식, 1988, 한국 지역 강수의 변동성에 관한 연구. 한국기상학회지, 24, 38-48.
 Anderson, J., H. van den Dool, Barnston, A., Chen, W., Stern, W., and Ploshay, J., 1999, Present-day capabilities of numerical and statistical models for atmospheric extratropical seasonal simulation and prediction, Bulletin of American Meteorological Society, 80, 1349-1361.
 Hwang, S. N, Schemm, J. K. E., Barnston, A. G., Kwon, W. T., 2001, Long-Lead Seasonal Forecast Skill in Far Eastern Asia Using Canonical Correlation Analysis. Journal of Climate, 14, 3005-3016.
 He Y., and Barnston, A. G., 1996, Long-Lead Forecasts of Seasonal Precipitation in the Tropical Pacific Islands Using CCA. Journal of Climate, 9, 2020-2035.
 Kim, M. K., Kim, H. S., Kwak, C. H., Suh, M. S., Park, C. K., 2002, Seasonal prediction of regional temperature in Korea using the lag-correlated large-scale predictors. Accepted to Journal of the Korean Meteorological Society.
 Michaelson, J., 1987, Cross-validation in statistical climate forecast models. Journal of Climate Application Meteorology, 26, 1589-1600.
 Shabbar, A. and Barnston, A. G., 1996, Skill of seasonal climate forecasts in Canada using canonical correlation analysis. Monthly Weather Review, 124, 2370-2385.

2002년 8월 1일 원고 접수
 2002년 9월 25일 수정원고 접수
 2002년 9월 28일 원고 채택