



ROC evaluation for MLP ANN drought forecasting model

Jeong, Min-Su^a · Kim, Jong-Suk^b · Jang, Ho-Won^c · Lee, Joo-Heon^{c*}

^aDisaster Management Research Center

^bDept. of Civil Engineering, University of Seoul

^cDept. of Civil Engineering, Joongbu University

Paper number: 16-048

Received: 8 June 2016; Revised: 17 August 2016 / 7 September 2016; Accepted: 7 September 2016

Abstract

In this study, the Standard Precipitation Index(SPI), meteorological drought index, was used to evaluate the temporal and spatial assessment of drought forecasting results for all cross Korea. For the drought forecasting, the Multi Layer Perceptron-Artificial Neural Network (MLP-ANN) was selected and the drought forecasting was performed according to different forecasting lead time for SPI (3) and SPI (6). The precipitation data observed in 59 gaging stations of Korea Meteorological Administration (KMA) from 1976~2015. For the performance evaluation of the drought forecasting, the binary classification confusion matrix, such as evaluating the status of drought occurrence based on threshold, was constituted. Then Receiver Operating Characteristics (ROC) score and F score according to conditional probability are computed. As a result of ROC analysis on forecasting performance, drought forecasting performance, of applying the MLP-ANN model, shows satisfactory forecasting results. Consequently, two-month and five-month leading forecasts were possible for SPI (3) and SPI (6), respectively.

Keywords: Meteorological Drought, Standardized Precipitation Index (SPI), MLP ANN, ROC, F score

MLP ANN 가뭄 예측 모형에 대한 ROC 평가

정민수^a · 김종석^b · 장호원^c · 이주현^{c*}

^a(사)방재관리연구소, ^b서울시립대학교 토목공학과, ^c중부대학교 토목공학과

요 지

본 연구에서는 기상학적 가뭄지수인 표준강수지수(Standardized Precipitation Index, SPI)를 이용하여 우리나라 전역에 대한 가뭄예측의 시공간적인 평가를 수행하였다. 또한 다층 퍼셉트론 인공신경망(Multi Layer Perceptron-Artificial Neural Network, MLP-ANN) 예측 기법을 이용하여 SPI(3), (6)에 대한 선행예보시간별 가뭄 예측을 실시하였다. 입력 자료는 기상청 산하의 59개 관측소에서 관측된 기상자료를 활용하였고, 관측자료 기간은 1976~2015년이다. 예측 모형의 성능평가는 기준점(Threshold)에 따른 가뭄 발생유무와 같은 이진분류 혼동행렬을 구성하여 Receiver Operating Characteristics (ROC) score와 조건부 확률에 따른 F score를 산정하여 예측 성능평가를 수행하였다. 예측성능에 대한 ROC 분석결과 다층 퍼셉트론 인공신경망(MLP-ANN) 모형을 적용한 가뭄예측성능이 매우 우수한 것으로 나타났으며, SPI (3)은 2개월, SPI (6)는 5개월 정도의 선행예측이 충분히 가능한 것으로 나타났다.

핵심용어: 기상학적 가뭄, 표준강수지수(SPI), MLP ANN, ROC, F score

1. 서 론

IPCC 제 5차 기후변화평가 보고서(IPCC, 2014)에 따르면 인위적인 복사강제력 증가로 인한 기후시스템의 변동에 따라 여러 지역에 가뭄, 홍수 등 다양한 극한 기후현상 및 도시지역

*Corresponding Author. Tel: +82-31-8075-1615
E-mail: leejh@joongbu.ac.kr (J.-H. Lee)

물부족이 발생되고, 동아시아 지역도 몬순 변동성과 관련한 극한 기후현상에 따른 취약성을 보일 수 있는 것으로 전망하였다. 최근 우리나라의 경우에도 가뭄이 연속적으로 발생되고 있으며, 특히 2015년에 소양강 댐의 수위가 1978년 이후 두 번째로 최저수위를 기록하는 등 극심한 가뭄 현상이 발생되고 있어 이에 대한 정부부처간의 다각적인 대처가 시급하게 요구되고 있는 실정이다.

가뭄발생의 피해를 최소화하기 위한 노력의 일환으로 통계적인 기법을 활용한 기상학적 가뭄의 발생특성을 분석하고 이를 장래 예측에 활용하는 다양한 연구가 진행되어 왔다 (Kwon et al., 2009; Bae et al., 2012; Lee et al., 2013). Kwon et al. (2009)은 전지구적인 기후변화 시나리오에 기반을 두고 미래의 기상 변화를 전망하여 중장기적인 시공간적 분포특성에 대한 가뭄발생을 제시하였고 Bae et al. (2012)은 기상정보와 구축된 지형자료를 이용한 실시간 양상불 가뭄전망정보 연구를 수행한 바 있다. 또한 Lee et al. (2013)은 수문시계열이 갖는 비정상성을 고려하여 다층 퍼셉트론의 네트워크 구조에 따른 인공신경망 모형을 제시한 바 있다.

시계열 수문자료에 대한 통계적 또는 물리적 예측기법을 적용한 결과를 평가하는데 있어 예측 결과의 정확도(Accuracy)를 판단하는 것이 기본적인 모델의 성능 평가 방법이라 할 수 있다. 반면에 이원화된 위험군을 갖는 경우 대한 예측기법 성능 평가방법으로 제시되는 것이 Receiver Operating Characteristics (ROC) 성능 평가 기법이다(Egan, 1975). ROC 기법은 예측성능을 기초로 계급 구간을 선택하고 Hit Rate (HR)와 False Alarm Rate (FAR) 사이의 적정성을 묘사하기 위한 것으로 시그널 탐지 이론에서 오랜 연구가 이루어져 왔다(Egan, 1975; Swets et al., 2000). 기준점(Threshold)에 따른 이원화된 위험군에 대한 평가는 가뭄과 같은 수문기상분야와 관련된 위험 평가기법으로의 확장성을 가지며, Kim and Lee (2011)는 가뭄에 대한 평가를 수행하고 이를 실제 가뭄과 비교평가를 위해 ROC 분석을 수행한 바 있다. Yoo et al. (2013)은 일단위 Standardized Precipitation Index (SPI)를 산정하고 ROC 모형을 활용하여 예측결과에 대한 적중률을 산정하였고 Son et al. (2015)은 대기와 해양모델을 결합한 GloSea5 모델과 VIC 지면모델을 결합하여 가뭄지수를 산정하고 ROC 기법을 적용한 바 있다.

이와 같이 대부분의 국내 ROC 평가기법은 가뭄 평가 지수를 산정하고 실제 가뭄 발생을 조사하여 ROC 비교평가를 수행하여 왔다. 반면에 Spackman (1989)은 기계학습(Machine Learning)과 관련하여 ROC curve를 이용한 알고리즘의 비교 및 평가를 제시하였으며 단순 계급 정확도에 관한 성능평가를

위한 부분에서 ROC 그래프의 사용성은 증가되어 왔다 (Provost and Fawcett, 1998; Provost et al., 1998). 한편으로 ROC 성능 평가기법은 기본적으로 혼동행렬(Confusion Matrix)을 구성하게 되는데 위험 발생과 비위험 발생의 불균형이 행렬의 구성에 영향관계를 가짐에 따라 기존의 ROC 성능 평가방법과 함께 F score 등의 다각적인 성능 검토 및 비교평가의 필요성을 갖는다.

본 연구에서는 중장기 가뭄예측을 위해서 Multi Layer Perceptron - Artificial Neural Network (MLP-ANN) 모형을 적용하였으며, 월단위의 SPI를 적용하여 가뭄예측을 수행하였다. 또한 예측결과의 정확성을 확인하고 통계적인 기법을 적용한 가뭄의 선행예측이 몇 개월까지 가능한지 확인하기 위하여 SPI의 보통가뭄에 대한 기준점(Threshold)에 따른 예측치 및 관측치 각각의 사건-사건발생유무의 확인 및 혼동행렬을 구성하고 이를 이용한 ROC score 및 F score를 산정하여 예측성능 평가를 수행하였으며 보간기법을 이용하여 우리나라 전 영역에 대한 공간적 분석을 실시하였다.

2. 입력자료 구축 및 전망 시스템 구축

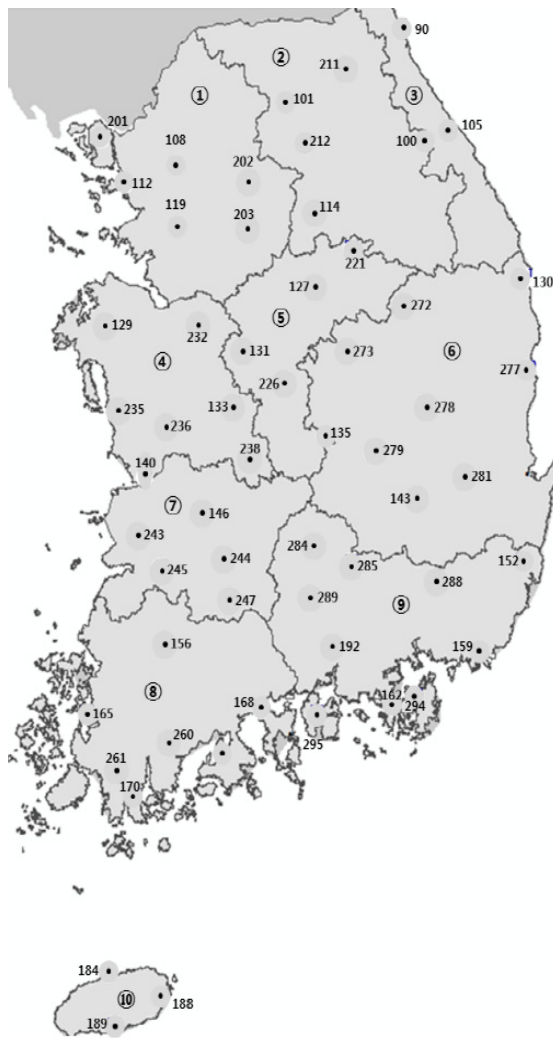
2.1 입력자료 구축

본 연구는 기상학적 가뭄지수인 SPI를 산정하고 SPI를 활용하여 가뭄재해에 미리 대처하기 위한 방안으로서 단기 또는 중장기 가뭄 예측 및 성능 평가를 수행하고자 한다. 입력자료 구축을 위해 기상청 산하 59개 강우 관측지점을 선정하였고 대상기간은 1976~2015년까지를 선정하였으며 Fig. 1과 같이 행정구역별로 구분하여 나타내었다.

2.2 표준강수지수(SPI)

표준강수지수(SPI)는 여러 가뭄평가지표 중 하나로 수분 공급원이 될 수 있는 강수의 과함 또는 부족 상태에 대한 정량적 해석이다(McKee et al., 1995). 따라서 표준강수지수는 인간 활동보다는 기상조건 상황에 따른 대기상태의 표현으로 국부적인 강우 발생의 가뭄상태를 표현하는 것이라 할 수 있다. McKee et al.(1995)은 시간단위별로 강수를 산정하여 유역의 건조/습윤 상태를 정량적으로 평가할 수 있는 기준을 Table 1과 같이 마련하였다.

보통가뭄의 경우 -1.0과 -1.5 사이 범위로 정의하고 있으며 -1.5는 보통가뭄과 심한가뭄의 구분 기준임에 따라 -1.5를 가뭄 위험에 대한 기준으로 분석에 적용하였다. 대상기간은 2000년 1월에서 2015년 12월까지이며, 이 기간 내에 가뭄 발



- ① **Seoul·Incheon·Gyeonggi**
 108 Seoul
 112 Incheon
 119 Suwon
 201 Ganghwa
 202 Yangpyeong
 203 Inheon
- ② **Gangwon (Yeongseo)**
 101 Chuncheon
 114 Wonju
 211 Inje
 212 Hongcheon
- ③ **Gangwon (Yeongdong)**
 90 Sokcho
 100 Daegwallyeong
 105 Gangneung
- ④ **Daejeon·Chungnam**
 129 Seosan
 133 Daejeon
 232 Cheonan
 235 Boryeong
 236 Buyeo
 238 Geumsan
- ⑤ **Chungbuk**
 127 Chungju
 131 Cheongju
 135 Chupungnyeong
 221 Jecheon
 226 Boeun
- ⑥ **Daegu·Gyeongbuk**
 130 Uljin
 138 Plhang
 143 Daegu
 272 Yeongju
 273 Mungyeong
 277 Yeongdeok
 278 Uiseong
 279 Gumi
 281 Yeongcheon
- ⑦ **Jeonbuk**
 140 Gunsan
 146 Jeonju
 243 Buan
 244 Imsil
 245 Jeongeup
 247 Namwon
- ⑧ **Gwangju·Jeonnam**
 156 Gwangju
 165 Mokpo
 168 Yeosu
 170 Wando
 260 Jangheung
 261 Haenam
 262 Goheung
- ⑨ **Busan·Ulsan·Gyeongnam**
 152 Ulsan
 159 Busan
 162 Tongyeong
 192 Jinju
 284 Geochang
 285 Hapcheon
 288 Miryang
 289 Sancheong
 294 Geoje
 295 Namhae
- ⑩ **Jeju**
 184 Jeju
 188 Seongsan
 189 Seogwiop

Fig. 1. Administrative district classification according to the rainfall stations of Korea meteorological administration (KMA)

Table 1. SPI Drought categories (McKee et al., 1995)

Range	Condition
More than 2.00	Extremely Wet
1.50~1.99	Very Wet
1.00~1.49	Moderately Wet
-0.99~0.99	Near Normal
-1.49~-1.00	Moderately Dry
-1.99~-1.50	Severely Dry
Less than -2.00	Extremely Dry

생 해는 Yoo et al. (2013)은 2000, 2001, 2008, 2009, 2012년, Son et al. (2015)은 2001, 2008, 2009, 2012, 2013년을 가뭄 피해 연도로 제시한 바 있다. 본 연구에서는 2014년과 2015년에도 연이은 가뭄피해를 고려해 2000, 2001, 2008, 2009,

2012, 2013, 2014, 2015년 등 8년의 기간을 가뭄 발생해로 선정하여 비교 검토를 수행하였다.

2.3 다층 퍼셉트론 인공신경망(MLP ANN)

인공신경망은 뇌신경 생리학에서 시작된 것으로 시냅스 결합으로 네트워크를 형성하고 학습을 통해 전반적인 시냅스 결합세기를 변화시키면서 문제를 해결해가는 모델을 의미한다(Kang and Lee, 2011). 인공신경망은 목적에 따라 다양하게 적용될 수 있으며 복잡한 자료구조를 갖는 시계열 목적에 따라 자료의 장래 예측문제 해결을 위한 비선형 모델로 분류될 수 있다. 본 연구에서는 기상학적인 비선형 SPI를 모델의 입력자료로 구축하기 위해 지수를 0과 1 사이에 값을 갖는 데이터로 변환하여 사용하였으며(Lee et al., 2013) 정규화는 다음 Eq. (1)과 같다.

$$X_o = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

여기서, X_o 는 표준화를 통해 정규화된 값이며 X 는 관측된 SPI 값으로 X_{min} 과 X_{max} 는 각각 SPI의 최솟값 및 최댓값을 나타낸다. 다층 퍼셉트론과 달리 다층 퍼셉트론은 여러 개의 층으로 구성하여 모델의 한계점과 선형분리 문제에 대한 단점을 보완한 방법으로 입력 노드를 제외한 각 처리노드에 비선형 활성화함수를 가지며 역전파(Backpropagation) 알고리즘이라 불리는 지도학습기법을 이용하여 비선형 시계열 예측모델로 활용성을 갖는다. 각 처리노드의 활성화함수에는 여러 활성화함수가 있으며 본 연구에서는 여러 활성화함수 중에서 Eq. (2)와 같은 시그모이드(Sigmoid) 함수를 적용하였다.

$$Sigmoid(a) = \frac{1}{1 + e^{-a}} \tag{2}$$

Fig. 2는 다층 퍼셉트론 인공지능망을 이용한 가뭄 예측 절차를 나타내며 크게 데이터 구축, 학습 및 매개변수 조정, 검증 및 예측으로 이루어져 있다(Lee et al., 2013).

본 연구에서는 2000~2015년까지 16년간 월별 192개 시계열 자료에 대해 one-step-forward 방식으로 예측을 하였으며 예측된 결과와 관측치에 대한 비교 검토를 통해 예측 모델의 예측성능 평가를 수행하였다. 성능평가에는 Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Correlation Coefficient (COR)을 이용하였다.

2.4 이원화된 위험군 분류기법 평가 및 ROC

Murphy (1993)는 기존의 상관분석 등을 통한 예측 성능평가와 달리 예측치 및 관측치가 속해 있는 포함군의 비교를 통해 예측검증을 제시한 바 있다. Bae et al. (2013)은 가뭄심도 평가와 함께 결과의 불확실성 평가를 통한 의사결정을 제시한 바 있으며 이러한 이원화된 위험군 분류기법을 통한 성능평가 방법이 ROC 분석이다(Kim and Lee, 2011). ROC 분석은 적중률(Hit Rate, HR)과 비적중률(False Alarm Rate, FAR)로 나타내어 이를 확률적 근거를 통해 가장 높은 확률을 찾아내는 방법이다(Heckerling, 2002). ROC 분석은 Table 2와 같이 혼동행렬 구성이 선행되어야 한다.

혼동행렬의 열은 위험요소에 대한 관측사건의 발생 유무이고 행은 예측사건 발생유무로 실제 위험요소 발생 및 예측 성공인 True Positive (TP), 실제 위험 미발생에 대한 예측 실패인 False Negative (FN), 위험요소 발생에 대한 예측 실패인 False Negative (FP), 위험 미발생 및 예측 성공을 의미하는 True Negative (TN)으로 분류된다.

분류된 4가지 요소는 위험의 이원화된 관점으로 비용민감 학습(Cost-Sensitive Learning)에 따른 트레이드오프 관계로

Table 2. Confusion matrix of ROC

Threshold : -1.5		Actual	
		Positive	Negative
Predicted	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	False Negative (FN)

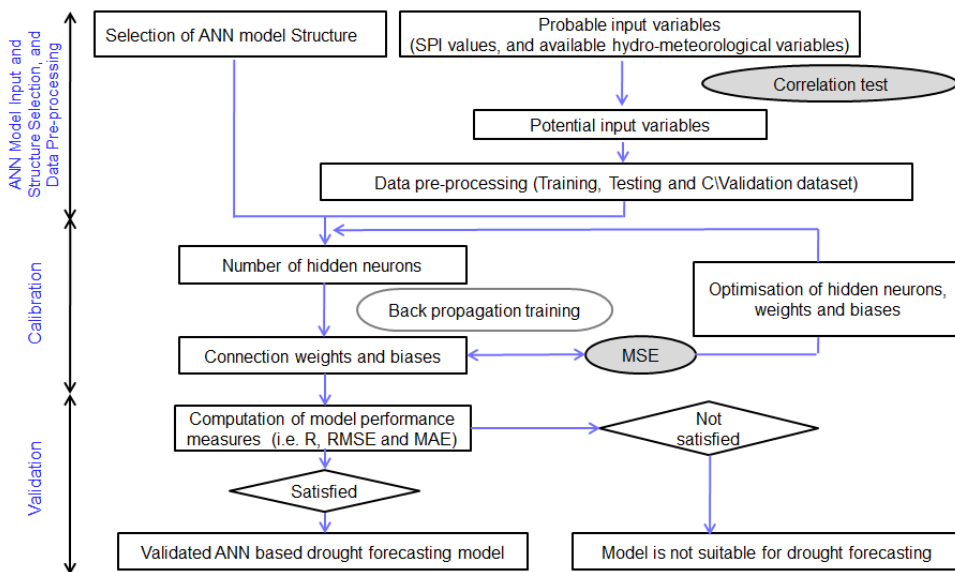
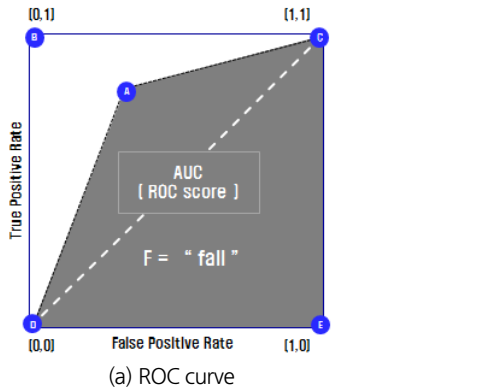


Fig. 2. Flowchart of ANN based drought forecasting process (Lee et al., 2013)



(a) ROC curve

ROC score	Evaluation
0.9 < AUC ≤ 1.0	Excellent
0.8 < AUC ≤ 0.9	Good
0.7 < AUC ≤ 0.8	Fair
0.6 < AUC ≤ 0.7	Poor
AUC ≤ 0.6	Fail

(b) ROC score

Fig. 3. ROC curve and ROC score categories

볼 수 있으며 X축에 가까울수록 적은 비용발생, Y축에 가까울수록 적은 편익을 의미한다. 또한 분류된 4개의 요소 중에서 TN은 위험 및 위험대비 비용이 없는 건전한 상태를 의미하고 TP는 위험발생에 대한 편익의 최소화로 이해 될 수 있다. Fig. 3은 ROC curve와 ROC score의 비용과 편익 사이의 트레이드 오프 관계를 표현하며, ROC curve의 적중률과 비적중률은 Eqs. (3)~(4)와 같다.

$$HR = \frac{TP}{Total\ Positive} \tag{3}$$

$$FAR = \frac{FP}{Total\ Negative} \tag{4}$$

2.5 혼동행렬을 이용한 기타 성능평가 방법

혼동행렬은 ROC 기법과 달리 Accuracy, F score 등과 같은 확률적 성능평가 방법으로 제시될 수 있다. Accuracy는 행렬의 각 성분을 예측 성공과 실패의 두 개의 범주로 재분류하여 성능평가를 하는 것으로 ROC score와 유사한 결과를 도출하며 Eq. (5)와 같다.

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{5}$$

이와 달리 우리나라의 가뭄과 같이 위험요소가 비위험보다 발생이 적은 경우에 따른 불균형을 고려하여 비위험요소를 제외한 성능평가를 고려할 수 있다. 비위험 요소 제거에 따른 위험발생-위험예측의 예측성능 평가를 위해서, 위험을 예측한 경우에 대한 실제 위험의 발생(Precision)과 실제 위험이 발생했을 때 위험의 예측된 경우(Recall)에 대한 조건부 확률로 나타낼 수 있으며 각각 Eqs. (6)~(7)과 같다.

$$Precision = P(TP | Predicted\ Positive) \tag{6}$$

$$Recall = P(TP | Actual\ Positive) \tag{7}$$

불균형 혼동행렬에서 비위험발생-비위험예측을 제외한 Precision과 Recall을 산정하고 두 결과에 대한 조화평균으로 성능평가를 수행하는 것이 F score이며 Eq. (8)과 같다.

$$Fscore = \frac{1}{\frac{1}{Precision} + \frac{1}{Recall}} \tag{8}$$

2.6 결과분석

2.6.1 SPI (3), (6)의 선행예보시간별 ROC Curve 및 Accuracy 공간분포

SPI (3)의 경우는 2000, 2001, 2008, 2009, 2013, 2014, 2015년에 상대적으로 높은 가뭄 발생빈도를 보였으며 SPI (6)의 발생횟수는 SPI (3)보다 다소 적은 발생을 보였으나 SPI (3)과 유사한 분포를 보였다. SPI (3)에 대한 분석 결과, 2007년 이전에는 주로 봄철에서 여름철로 넘어가는 4월과 7, 8월 사이에 가뭄발생 양상을 보였으며, 2008년 이후에는 8, 9월부터 가뭄 발생 양상이 시작되어 이듬해에 초반까지 가뭄의 영향을 주는 양상이 나타났다. SPI (6)은 SPI (3) 보다는 상대적으로 가뭄사상이 적게 발생한 것으로 나타났으나 SPI (3)과 전반적으로 유사한 발생 양상을 보이는 것으로 나타났다.

ROC score 및 Accuracy는 전 예측자료에 대한 모델 성능평가로 기존 RMSE, COR 등과 유사한 성능평가를 얻는 것으로 알려져 있다. SPI (3), (6)의 선행예보시간별 ROC score 및 Accuracy 산정결과는 Fig. 4~5에서 나타내었으며, 행정구역별 Accuracy 산정결과는 Table 3으로 나타내었다.

Fig. 4는 SPI (3)의 ROC score 산정결과 LT-1 (1개월 선행예측)에서는 대부분 (0, 1)에 위치하고 Accuracy 공간분포도 좋은 성능 분포를 보였으며, LT-2 (2개월 선행예측)는 LT-1에 비해 다소 떨어지지만 상당히 좋은 성능 평가로 나타났다. 반면에 LT-3 (3개월 선행예측)에서는 제주도 및 영산강 일부

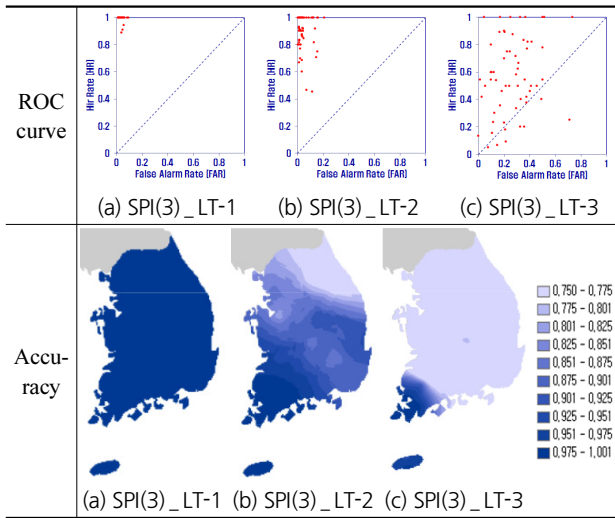


Fig. 4. ROC curve and spatial distribution of SPI (3) based drought forecasting accuracy

지역을 제외하고는 예측 성능이 현저히 떨어지는 것으로 나타났다.

특히, SPI (3), (6)의 ROC curve에서 보이듯이 선형예보시간이 길어짐에 따라 자료의 분포가 왼편으로 편향되어 나타나는 것을 확인할 수 있으며 이는 가뭄이 비가뭄에 비해 발생빈율이 상대적으로 적은 불균형에 대한 결과로 판단된다.

Fig. 5는 SPI (6)에 대한 지체시간별 ROC score 및 Accuracy 공간분포로 SPI (3)과 유사하게 선형예보시간이 길어질수록 예측 성능이 저하되는 것으로 나타났으며, 이러한 변화는 한강유역에서 뚜렷한 것으로 분석되었다. Table 3은 행정구역별 SPI (3)과 (6)의 Accuracy 산정결과 표로 강원(영동)의 경우는 선형예보시간이 늘어남에 따라 타 지역보다 상대적으로

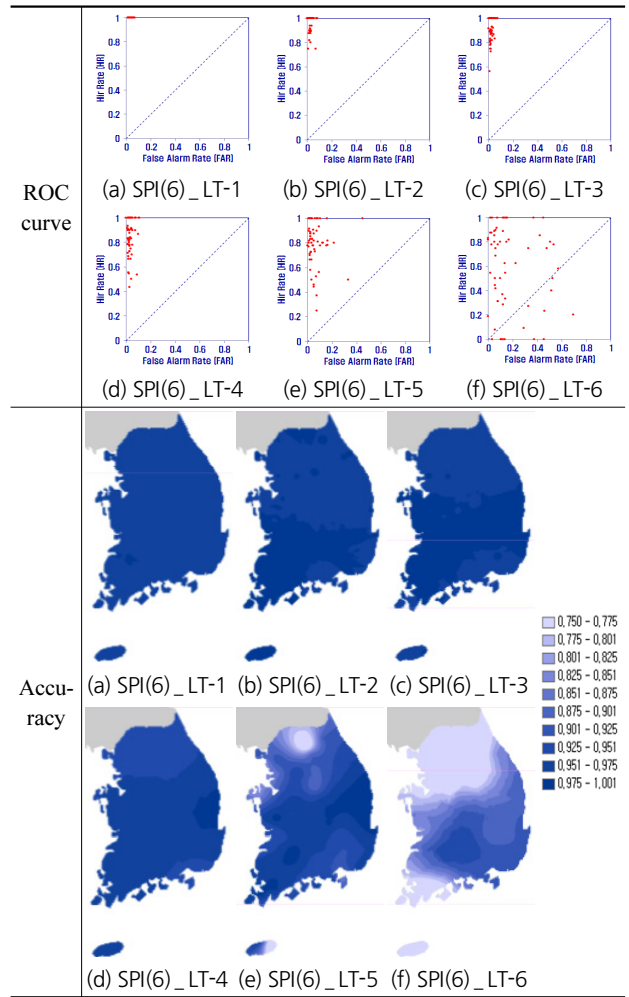


Fig. 5. ROC curve and spatial distribution of SPI (6) based drought forecasting accuracy

Table 3. The calculation results of accuracy according to the administrative district classification

Mean		①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	Total Average
SPI (3)	LT-1	0.96	0.95	0.97	0.96	0.96	0.95	0.97	0.97	0.97	0.97	0.96
	LT-2	0.97	0.92	0.94	0.92	0.93	0.90	0.98	0.96	0.97	0.98	0.95
	LT-3	0.72	0.74	0.84	0.76	0.77	0.69	0.73	0.70	0.65	0.75	0.72
SPI (6)	LT-1	0.97	0.97	0.98	0.95	0.97	0.96	0.97	0.97	0.97	0.97	0.97
	LT-2	0.97	0.96	0.98	0.96	0.97	0.97	0.98	0.97	0.98	0.98	0.98
	LT-3	0.97	0.97	0.96	0.97	0.98	0.98	0.99	0.97	0.97	0.97	0.97
	LT-4	0.94	0.93	0.95	0.96	0.96	0.98	0.96	0.95	0.96	0.97	0.96
	LT-5	0.90	0.84	0.92	0.94	0.94	0.97	0.96	0.94	0.95	0.99	0.93
	LT-6	0.69	0.58	0.88	0.79	0.87	0.89	0.88	0.93	0.70	0.61	0.81

① Seoul·Incheon·Gyeonggi, ② Gangwon(Yeongseo), ③ Gangwon(Yeongdong), ④ Daejeon·Chungnam, ⑤ Chungbuk, ⑥ Daegu·Gyeongbuk, ⑦ Jeonbuk, ⑧ Gwangju·Jeonnam, ⑨ Busan·Ulsan·Gyeongnam, ⑩ Jeju

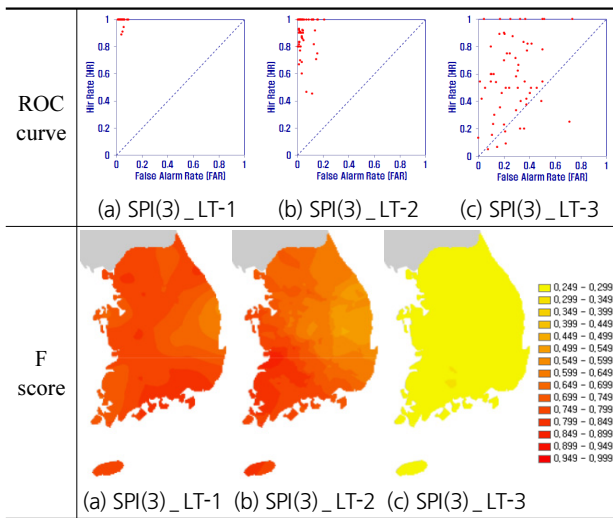


Fig. 6. ROC curve and spatial distribution of F score based on SPI (3) drought forecast

높은 예측 정확도를 보이는 것으로 나타났으며 이는 지형적 효과에 기인한 것으로 사료된다.

2.6.2 SPI (3), (6) 선행예보 기간별 ROC curve 및 F score 공간분포

혼동행렬 구성비 및 그에 따른 ROC curve의 자료분포가 편향성을 보이는 것을 고려할 때 가뭄이 비가뭄 발생보다 현저히 적음을 확인할 수 있으며 이원화된 위험군의 발생유무에 대한 보다 면밀하고 다각적인 성능평가의 제시가 필요하다. 따라서 위험과 비위험요소의 현저한 비대칭적 발생을 고려하여 위험 발생에 대한 예측성능 평가에 초점을 맞춘 F score에 따른 성능평가를 수행하였다. F score는 민감도 및 특이도에

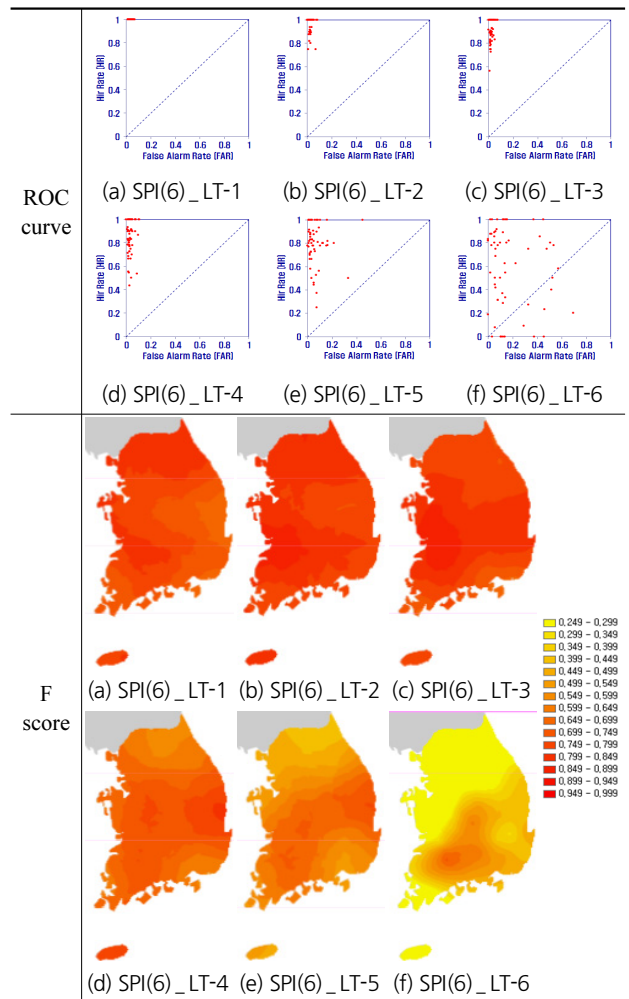


Fig. 7. ROC curve and spatial distribution of F score based on SPI (6) drought forecast

Table 4. The calculation results of F score according to the administrative district classification

Mean		①	②	③	④	⑤	⑥	⑦	⑧	⑨	⑩	Total Average
SPI (3)	LT-1	0.8	0.72	0.82	0.75	0.77	0.66	0.77	0.8	0.76	0.78	0.76
	LT-2	0.79	0.61	0.61	0.6	0.69	0.49	0.85	0.72	0.77	0.82	0.69
	LT-3	0.19	0.21	0.12	0.24	0.17	0.18	0.26	0.21	0.26	0.29	0.21
SPI (6)	LT-1	0.8	0.83	0.86	0.7	0.81	0.69	0.81	0.74	0.7	0.84	0.76
	LT-2	0.83	0.88	0.86	0.73	0.84	0.76	0.87	0.79	0.75	0.9	0.81
	LT-3	0.79	0.8	0.73	0.78	0.86	0.82	0.89	0.74	0.73	0.81	0.79
	LT-4	0.69	0.58	0.6	0.74	0.69	0.78	0.74	0.63	0.66	0.82	0.69
	LT-5	0.53	0.41	0.44	0.62	0.6	0.72	0.68	0.58	0.59	0.46	0.59
	LT-6	0.24	0.13	0.2	0.39	0.17	0.41	0.46	0.53	0.26	0.06	0.33

① Seoul·Incheon·Gyeonggi, ② Gangwon(Yeongseo), ③ Gangwon(Yeongdong), ④ Daejeon·Chungnam, ⑤ Chungbuk, ⑥ Daegu·Gyeongbuk, ⑦ Jeonbuk, ⑧ Gwangju·Jeonnam, ⑨ Busan·Ulsan·Gyeongnam, ⑩ Jeju

대한 합성추정 알고리즘 평가기법으로 조건부 확률을 이용한 Precision과 Recall을 산정하고 이 두 인자에 대한 동등 가중치를 적용할 수 있는 조화평균을 이용하여 F score를 산정한다.

ROC score와 비교검토를 위해 산정된 F score 값 및 보간기법을 적용하여 우리나라 전역에 대한 공간분포를 수행하였으며 결과는 Fig. 6~7과 같다.

Fig. 6은 SPI (3)의 LT-1, LT-2에 대한 행정구역별 F score 산정 결과로 0.6~0.8 사이에 예측 성능 평가가 이루어지는 것으로 나타났으며 부산·울산·경남 지역의 산정 결과가 높게 나타나는 반면에 대구·경북 지역은 다소 낮은 성능 평가를 보였다.

SPI (6)에 대한 지체시간별 F score 산정한 결과는 Fig. 7과 같이 전북 지역에서 다소 높게 나타났다. 행정구역별 F score 산정결과는 Table 4와 같다.

행정구역별 F score 산정결과 SPI (6), LT-2에서 가장 좋은 성능평가가 이루어지는 것으로 나타났다. 또한 행정구역별로는 강원(영서) 지역이 다른 곳에 비해 예측성능이 다소 떨어지는 것으로 분석되었다.

ROC score 평가 기준에 따라 F score는 “Fair” 또는 “Good”에 가까운 평가 결과로 이는 Accuracy에 비해 등급이 1~2 등급 정도 떨어지는 것으로 분석되었다. 따라서 기존 성능평가 및 이원화된 위험 평가 기법에 있어 다소 높은 예측 성능을 나타냈지만 보다 면밀하고 다각적인 예측성능 검토의 경우에는 F score 분석이 예측성 검토를 위한 정보제공에 있어 유용할 것으로 사료된다.

3. 결론

본 연구에서는 기상청 산하 59개 지점에 대한 강우자료를 이용하여 기상학적 가뭄을 정량화하고 통계모델을 통한 단기 예측을 실시하여 우리나라 적용가능성을 평가하였다. 이원화된 위험군에 대한 분류를 통하여 가뭄의 공간적 예측성능 평가를 실시하였으며, 주요 결과를 정리하면 다음과 같다.

- 1) 기상학적 가뭄지수를 통한 우리나라 가뭄을 분석한 결과, 2007년 이전에는 주로 봄철에서 여름철로 전환되는 시기에 가뭄발생 양상을 보였으며, 2008년 이후에는 8, 9월부터 가뭄이 발생하기 시작하여 이듬해 초반까지 영향을 주는 것으로 나타났다. 또한, SPI (6)의 가뭄발생빈도는 SPI (3)보다 다소 적은 발생을 보였으나, 공간적 발생양상은 유사한 것으로 분석되었다.
- 2) SPI에 대한 Accuracy 분석결과, 선행예보시간이 길어질

수록 예측성능이 현저히 떨어지는 경향을 보였으며, SPI (3)의 경우에는 LT-3에서 제주도를 포함한 남부 일부지역을 제외한 유역에서 이러한 변화가 뚜렷하게 분석되었다. SPI (6)에서는 한강과 영산강 일부지역에서 예측성능이 상대적으로 저하되는 특징을 보이는 것으로 나타났다. 그러나 강원(영동)의 경우는 선행예보시간이 늘어남에 따라 타 지역보다 상대적으로 높은 예측 정확도를 보이는 것으로 나타났다.

- 3) 가뭄이 비가뭄 발생보다 현저히 적게 발생함에 따라 이원화된 위험군의 발생유무에 대한 평가를 위하여 F score 분석을 실시하였다. SPI (3)에 대한 분석결과, 0.6~0.8 사이에 예측 성능 평가가 이루어지는 것으로 나타났으며, 부산·울산·경남 지역의 산정 결과가 높게 나타나는 반면에 대구·경북 지역은 다소 낮은 성능 평가를 보였다. SPI (6)에 대한 F score 산정 결과, LT-2에서 가장 좋은 성능평가가 이루어졌고, 예측성능이 가장 우수한 지역은 전북 지역으로 나타났다. 강원(영서) 지역은 다른 지역에 비해 다소 낮은 예측성능을 보이는 것으로 분석되었다. 또한, 혼동행렬의 TN(비가뭄 발생 - 비가뭄 예측) 항목을 제외한 예측성능 평가를 위해 Precision과 Recall에 대한 F score를 산정한 결과 “Good” 또는 “Fair”로 1~2 등급 낮게 평가되는 것으로 나타났다.

본 연구는 제한된 관측 자료에 대한 기상학적 가뭄평가와 통계모델을 이용한 단기예측평가를 실시하였지만, 다각적이고 보다 면밀한 검토를 통하여 예측모형의 시공간적 예측성능평가를 실시함으로써 가뭄관리에 대한 선제적 방어전략 수립에 중요한 기초자료를 제공해 줄 수 있을 것으로 기대된다.

감사의 글

본 연구는 기상청 기상기술개발사업(KMIPA 2015-2070) 중 “가뭄정보의 통합 및 공동 활용기술 개발” 과제의 연구비 지원으로 수행되었으며 이에 감사드립니다.

References

- Bae, D. H., Son, K. H., Ahn, J. B., Hong, J. Y., Kim, G. S., Chung, J. S., Jung, U. S., and Kim, J. K. (2012). “Development of real-time drought monitoring and prediction system on Korea & East Asia region.” *Atmosphere Korean Meteorological*

- Society*, Vol. 22, No. 2, pp. 267-277.
- Bae, D. H., Son, K. H., and Kim, H. A. (2013). "Derivation & evaluation of drought threshold level considering hydro-meteorological data on south korea." *J. Korea Water Resources Association*, Vol. 46, No. 3, pp. 287-299.
- Egan, J. P. (1975). "Signal detection theory and ROC analysis, series in cognition and perception." Academic Press, New York.
- Heckerling, P. S. (2002). "Parametric receiver operating characteristic curve analysis using mathematica." *Computer Methods and Programs in Biomedicine*, Vol. 69, No. 1, pp. 65-73.
- IPCC (2014). Fifth Assessment Report (AR5) 「Climate change 2014 : impacts, adaptation, and vulnerability」
- Kang, B. S., and Lee, B. K. (2011). "Application of artificial neural network to improve autitative precipitation forecasts of meso-scale numerical weather prediction." *Journal of Korea Water Resources Association*, Vol. 44, No. 2, pp. 97-107.
- Kim, G. S., and Lee, J. W. (2011). "Evaluation on drought indices using the drought Records." *J. Korea Water Resources Association*, Vol. 44, No. 8, pp. 639-652.
- Kwon, H. H., Moon, J. W., Song, H. S., and Moon, Y. I. (2009). "Climate information and GCMs seasonal forecasts based short-term forecasts for drought." *Conference of Korea Water Resources Association*, pp. 1186-1190.
- Lee, J. H., Kim, J. S., Jang, H. W., and Lee, J. C. (2013) "Drought forecasting using the multi layer perceptron(MLP) artificial neural network model." *J. Korea Water Resources Association*, Vol. 46, No. 12, pp. 1249-1263.
- Mckee, T. B., Doesken, N. J., and Kleist, J. (1995). "Drought monitering with multiple time scales preprints." *9th Conference on Applied Climatology*, 15-20 Janiary, Dallas, TX, pp. 233-236.
- Murphy, A. H. (1993). "What is a good forecast? an essay on the nature of goodness in weather forecasting." *American Meteorological Society*, Vol. 8, pp. 281-293.
- Provost, F., and Fawcett, T. (1998). "Robust classification systems for imprecise environments." In Proceedings of AAAI-98. AAAI Press. To appear. Available: <http://www.croftj.net/~fawcett/papers/aaai98-dist.ps.gz>.
- Provost, F., Fawcett, T., and Kohavi, R. (1998). The case against accuracy estimation for comparing induction algorithms. In: Shavlik, J. (Ed.), Proc. ICML-98. Morgan Kaufmann, San Francisco, CA, pp. 445-453.
- Son, K. H., Bae, D. H., and Cheong, H. S. (2015). "Construction & evaluation of glosea5-based hydrological drought outlook system." *Atmosphere*, Vol. 25, No. 2, pp. 271-281.
- Spackman, K. A. (1989). "Signal detection theory: valuable tools for evaluating inductive learning. In: Proc. Sixth Internat". Workshop on Machine Learning. Morgan Kaufman, San Mateo, CA, pp. 160-163.
- Swets, J. A., Dawes, R. M., and Monahan, J. (2000). "Better decisions through science." *Scientific American* 283, pp. 82-87.
- Yoo, J. Y., Song, H., Kim, T. W., and Ahn, J. H. (2013). "Evaluation of short-term drought using daily standardized precipitation index and ROC analysis." *Journal of The Korean Society of Civil Engineers*, Vol. 33, No. 5, pp. 1851-1860.