

위성 및 장기예측자료의 기계학습을 통한 가뭄예측

Drought forecasting based on machine learning of remote sensing and long-range forecast data

이진영

위성 및 장기예측자료의 기계학습을 통한 가뭄예측

Drought forecasting based on machine learning of remote sensing and long-range forecast data

이진영

발간사

2015년은 미국과 유럽을 비롯한 세계 곳곳이 심각한 가뭄에 시달린 한 해로 기록되었습니다. 우리나라도 지역에 따라 수년간 이어지는 극심한 가뭄으로 많은 피해를 입은 국민들이 여전히 괴로움을 겪고 있습니다. 가뭄으로 인한 피해를 줄이려는 다양한 노력이 이어지고 있지만 다가올 가뭄에 지혜롭게 대처할 수 있도록 하는 가뭄예측 연구는 그리 활발하게 이루어지지 않아 아쉬움이 있었습니다.

이에 본 센터에서는 가뭄정보에 대한 사용자들의 수요를 알아보고 위성과 장기예측에서 얻을 수 있는 자료를 이용한 가뭄예측 모델을 개발하는 연구를 수행하여 장기예측의 성능이 향상되는 경우에 가뭄예측이 얼마나 나아질지 향상 가능한 범위를 제시하였습니다. 가뭄피해를 최소화하려면 무엇보다 가뭄예측 정보가 특정지역에 적용될 수 있도록 상세히 제공되어야 할 것입니다. 본 연구에서 제시한 결과가 가뭄예측 후속연구에 기준선을 제시하고 앞으로 가뭄예측 연구가 활발히 이루어져 서로 상호보완적으로 활용되길 기대합니다.

끝으로 연구를 맡아 수행해 주신 본 센터의 이진영 박사의 노고에 감사드립니다. 그리고 바쁘신 가운데 자문과 조언을 해주신 세종대학교 배덕효 교수, 부산대학교 신현석 교수에게도 진심으로 감사의 말씀을 전합니다.

2016년 1월
APEC 기후센터
소장 정진승

ABSTRACT

The impact of droughts can be reduced through sustainable drought management and proactive measures against drought disasters. Accurate and timely provision of drought information is the most essential. This study developed a drought forecasting model to provide high-resolution drought information based on drought indicators that is useful for decision-makers. The purposes of the study are (1) to identify the needs of decision-makers regarding drought information through a two-part survey; (2) to develop a drought forecasting model using remote sensing and long-range forecast data based on machine learning, then assess its performance; and (3) to provide improvable ranges of drought forecasting capability in case of enhanced long-range forecast.

A two-part survey, in the form of a questionnaire, was conducted. Fifty-one government officials from 17 municipal governments were asked to participate in the survey. These government officials' main duties include addressing issues on droughts, heatwaves, disasters, wildfires, and urban greening. The first round of surveys had a return rate of 53% (27 participants), and the second round of surveys had a return rate of 19% (5 participants). Although the return rate is low, each response was considered seriously and applied to the design of the study because the survey was conducted replacing one-on-one interview.

If an institution, that is trusted by the public, were to provide drought information tailored to different sectors, the use of this drought information would be improved. Most respondents preferred to have weekly temporal resolution while only some preferred daily or monthly information. For administrative districts, most preferred

spatial information while some preferred spatially distributed information (with spatial resolution as high as 1×1 km). The need for 1-month to 1-year lead times of drought forecasting was also identified. Judging by these needs, it seems to be necessary to develop a drought forecasting model with a weekly temporal resolution and very fine spatial resolution. This study attempts to start the process of getting to that goal by developing a drought forecasting model with a monthly temporal resolution matching the resolution of the long-range forecast and a $0.5 \times 0.5^\circ$ spatial resolution, which matches the resolution of remote sensing-based data. Drought forecasting models with finer temporal and spatial resolutions will be developed in following studies.

The survey responses also called for drought information for gauges, as well as spatially distributed information based on remote sensing data. Sixty percent of the respondents of the second survey intend to use remote sensing-based drought indicators, even if the accuracy is lower than gauges. Drought risk assessment was performed for general drought impact categories, but no significant results were obtained due to the small number of respondents. Despite this, it is still meaningful that the respondents gave high priority to environmental and social drought impacts, which have long been ignored in comparison to other categories. This calls for a more detailed drought risk assessment for each drought impact category.

Several drought forecasting models were developed that predict drought indices of the 6-month Standardized Precipitation Index (SPI6), the 6-month Standardized Precipitation Evapotranspiration Index (SPEI6), as well as the Normalized Difference Water Index (NDWI). The Multiquadric Spline Interpolation method and three different machine learning models of Decision Tree (DT model), Random Forest (RF model), and Extremely Randomized Trees (ET model) were tested. Interpolation methods have traditionally been used to derive information for ungauged locations using

information from neighboring gauges and their distances. Machine learning models were tested to enhance the provision of initial conditions of droughts based on remote sensing data, as initial conditions are some of the most important factors for drought forecasting.

Classification of drought categories and regression of the values of drought indicators were performed. The input variables of machine learning models predicting SPI6 or SPEI6 include the: 6-month accumulated precipitation (PRCP), 6-month accumulated potential evapotranspiration (PET), Normalized Difference Vegetation Index (NDVI), Normalized Difference Water Index (NDWI), Daytime Land Surface Temperature (LST_DAY), Nighttime Land Surface Temperature (LST_NIGHT), Multivariate ENSO Index (MEI), Arctic Oscillation Index (AOI), and month (MONTH). The target variables are SPI6 or SPEI6. The models predicting NDWI use the periods matching the lead time for PRCP and PET instead of using 6 months.

Sensitivity analyses were performed to determine model parameters of the three machine learning models. In order to avoid overfitting, Leave-One-Year-Out cross-validation was applied by excluding one year repeatedly from 2003-2015 input data for training and using the data for the test. The performance measures used are: the accuracy of classifying drought categories such as Extreme Drought (ED), Severe Drought (SD), and Moderate Drought (MD); and the Mean Absolute Error (MAE) for regression. They were evaluated for 61 Automated Synoptic Observing System (ASOS) gauge locations.

The performance of long-range forecast is out of the scope of this study, although it is the most important factor of drought forecasting. Instead, the performance of long-range forecast data was evaluated against the use of climatological data (baseline), for the purposes of filling the future period of the lead time. A perfect forecast was simulated and improvable ranges were provided. The interpolation method was compared to machine learning models for drought forecasting for

ungauged areas.

When the Long-Range Forecast Method (F method) was compared to the Climatology Method (C method) for gauges, the C method outperformed for both classification and regression. Classification accuracy rapidly decreased and regression MAE increased with longer lead times.

The Climatology-Interpolation Method (C-I method), the Long-Range Forecast-Interpolation Method (F-I method), the Climatology-Machine Learning Method (C-ML method), and the Long-Range Forecast-Machine Learning Method (F-ML method) were compared for drought forecasting in ungauged areas. Machine learning-based methods performed better than interpolation methods for both classification and regression, and the methods using climatology data outperformed the methods using long-range forecast. In most cases, the C-ML method performed the best in drought forecasting.

A perfect forecast was simulated and the Perfect Forecast-Interpolation (PF-I method) and the Perfect Forecast-Machine Learning (PF-ML method) were compared to the F-I and F-ML methods, respectively. Classification accuracy of the PF-I method was in the range of 0.34-0.38 for SPI6 and SPEI6, indicating that the classification accuracy will be still below 0.4 despite a perfect forecast. The PF-ML method produced higher classification accuracy values ranging from 0.39-0.56. The ET model performed the best for SPI6 and SPEI6 forecast. Regression MAE of PF-I method was in the range of 0.47-0.53, and MAE for the methods using machine learning models produced an even lower range of 0.35-0.42. The RF and ET models performed very well. The performance of the F-ML method is expected to surpass the C-ML method as the skill of long-range forecast improves.

The Vegetation-Climatology-Machine Learning Method (V-C-ML method) and the Vegetation-Long-Range Forecast-Machine Learning (V-F-ML method) were compared for NDWI forecasting. The classification accuracy of the V-C-ML method

is 0.52-0.53 for 1 to 3-month lead times making it similar to the PF-ML method's (ET model) performance for SPEI6 forecast. The V-F-ML method (ET model) performed the best for the 4 to 6-month lead time with the classification accuracy of 0.51-0.53. The increase of MAE with lead time was not steep. The performance either decreased or was only slightly improved with perfect forecast simulations.

The model that outperformed overall, was the one that was based on climatological data and utilized the machine learning method. Although the contribution of long-range forecast for drought forecasting was not large, the use of machine learning modeling based on remote sensing data contributed to the enhancement of drought forecasting skill. Drought forecasting based on long-range forecast is expected to outperform the forecasting based on climatological data, as the skill of the long-range forecast improves.

In conclusion, it is recommended to forecast SPI6 or SPEI6 values based on machine learning using climatological data to provide spatially distributed drought information with a spatial resolution of $0.05 \times 0.05^\circ$, as used in this study. The classification accuracy will be in the range of 0.47-0.52 with 1-month lead time, and will decrease to 0.21-0.35 with 6-month lead time for ungauged areas. The regression MAE will be in the range of 0.41-0.47 with 1-month lead time, and in the range of 0.56-0.59 with 6-month lead time for ungauged areas. The long-range forecast will be more useful with the improvement of their forecasting skill, the classification accuracy with perfect forecast will reach 0.50-0.56, and the regression MAE will reach to 0.35-0.40 for ungauged areas.

The models predicting NDWI were less sensitive to the performance of the long-range forecast, compared to the models predicting SPI6 or SPEI6, and showed higher classification accuracy as well as lower regression MAE. NDWI can be useful as an indicator related to environmental drought impacts. It is recommended to: (1) improve the performance of the long-range forecast in order to provide drought

information to decision-makers using drought indicators familiar to them; (2) use models predicting NDWI; and (3) produce and disseminate drought information related to environmental drought impacts using NDWI. The classification accuracy of the V-C-ML method ranges between 0.51-0.53 and the regression MAE ranges between 0.45-0.50. Although this study used NDWI as a drought indicator, remote sensing-based drought indicators combining several variables, such as the Scaled Drought Condition Index (SDCI), can also be used.

The performance measure results derived from the models tested in this study can be used as baselines for future drought forecasting studies in the study area. To reduce uncertainty of drought forecasting, the developed model can be used with the existing drought information system as an ensemble.

국문 초록

가뭄재난에 대한 지속적인 관리 및 선제적인 대응을 통한 가뭄피해 최소화를 위해 무엇보다 정확하고 신속한 가뭄정보의 제공이 이루어져야 한다는 점을 인식하고 이를 위해 의사결정자의 목적에 맞는 가뭄지표를 이용한 지역에 쉽게 적용할 수 있는 고해상도 가뭄정보 제공 모델을 개발하였다. 연구의 목적은 (1) 설문을 통해 의사결정자들의 가뭄정보에 대한 수요를 파악하고, (2) 위성자료 및 장기예측 자료를 이용한 기계학습 모델링 기반 가뭄예측 모델을 구축하고 성능을 평가하며, (3) 장기예측의 성능이 향상되는 경우 가뭄예측이 얼마나 향상될 지 그 향상 가능한 범위를 제시하는 것이다.

의사결정자의 가뭄정보 수요 파악 및 활용의사 도출을 위해 17개 광역지방자치단체의 가뭄 환경피해와 관련있는 가뭄, 폭염, 재난, 재해, 산불, 도시숲 관련 담당자 51명을 대상으로 설문조사를 수행하였다. 제1차 설문조사에는 27명이 응답하였으나(회수율 53%) 1차 응답자를 대상으로 한 제2차 설문조사에서는 5명의 응답만을 얻을 수 있었다(회수율 19%). 표본수는 크지 않으나 실제 담당 공무원을 대상으로 하는 인터뷰 대신에 수행된 설문조사이므로, 개개의 응답을 의미 있게 받아들여 연구의 설계에 참고하였다.

응답자들의 업무에 직접적으로 활용될 수 있는 구체적인 가뭄정보가 공신력 있는 기관에 의해 제공된다면 활용도를 높일 수 있을 것으로 보인다. 가뭄정보의 시간해상도는 일별, 월별 정보가 필요하다고 응답한 경우도 있었으나 대체적으로 주별 가뭄정보를 원하였다. 공간해상도는 시군구와 같은 행정구역 단위의 정보에 대한 요구가 가장 많았으며 공간적으로 연속된 고해상도 격자(1×1 km 이상)에 대한 선호도 있었다. 가뭄예측 정보의 선행 시간은 1개월에서 1년까지 다양한

요구가 있었다. 이에 따르면 주별 시간해상도 및 상당히 높은 공간해상도의 가뭄 정보 생산을 위한 모델 구축이 뒤따라야 하나, 본 연구에서는 우선 장기예측 자료의 월별 시간해상도와 위성기반 식생 자료의 $0.05 \times 0.05^\circ$ 공간해상도에 맞추어 가뭄예측 모델을 개발하였다. 상세한 시간해상도 및 공간해상도 모델은 후속연구에서 개발할 것이다.

미계측지역의 가뭄정보 공간유형에 있어서는 지점 관측 자료를 이용한 가뭄정보와 위성 자료를 이용한 가뭄정보에 대한 수요가 모두 있었으며, 가뭄지표의 선정에 있어 2차 설문 응답자의 60%가 위성 기반 가뭄지표의 활용 의사를 밝혔다. 정확도에 있어 위성 기반 가뭄지표의 정확도가 지점의 관측 기반 정보보다 낮은 경우에도 위성 기반 가뭄지표를 사용하겠다고 답한 응답자도 있었다. 일반적인 가뭄피해유형에 대한 가뭄리스크 평가가 수행되었는데, 응답자가 적어 유의한 결론을 이끌어내지는 못하였다. 그러나 응답자들이 가뭄의 발생빈도와 피해강도를 평가함에 있어 그간 간과되어 온 가뭄의 환경적 피해와 사회적 피해를 다른 피해 유형과 마찬가지로 높게 평가한 점은 작은 표본 수에도 불구하고 주목할 만 하다. 본 결과를 기반으로 추후 각 피해 유형의 상세한 리스크 평가가 수행되어야 할 것이다.

의사결정자의 목적에 맞는 가뭄지표를 이용한 고해상도 가뭄예측 모델 개발은 보간법과 기계학습 모델링을 통한 가뭄지수, 정규수분지수(NDWI)의 예측을 통해 수행하였다. 보간법은 주변의 지점 정보와 그 지점까지의 거리를 이용하여 정보를 도출하는 방법으로 전통적으로 많이 사용되어 왔으며, 기계학습은 가뭄예측의 경우 초기조건이 중요한 역할을 한다는 점을 인지하여 초기조건에 대한 정보를 보강하고 위성자료를 이용하여 공간적으로 분포된 정보를 활용하기 위해 적용되었다. 보간법으로는 Multiquadric Spline 방법을 이용하였으며 기계학습은 Decision Tree(DT모델), Random Forest(RF모델), Extremely Randomized Trees(ET모델)의 세 가지 모델을 이용하였다.

가뭄지표의 분류와 회귀를 각각 수행하였다. 6개월 표준강수지수(SPI6), 6개

월 표준강수증발산지수(SPEI6)를 목표변수로 하는 기계학습의 경우 입력변수는 가뭄지수를 예측하는 경우 위성기반의 6개월 누적 강수량(PRCP), 6개월 누적 잠재증발산량(PET), 정규식생지수(NDVI), NDWI, 주간지표면온도(LST_DAY), 야간지표면온도(LST_NIGHT), 다변수ENSO지수(MEI), 북극진동지수(AOI), 그리고 월(MONTH)을 이용하였다. 가뭄의 환경피해와 관련된 정보를 제공할 수 있는 NDWI 예측의 경우 입력 변수 중 강수량과 잠재증발산량의 누적 기간은 6개월로 하지 않고 예측 선행시간과 같은 기간으로 하였다.

기계학습 모델의 매개변수 선정을 위한 민감도 분석을 수행하였으며 예측모델의 과적합(overfitting)을 피하기 위해 2003-2015년의 입력자료 중 일 년씩 제외시킨 값으로 모델을 훈련시켜 그 일 년의 값을 예측하는 Leave-One-Year-Out 교차검증을 실시하였다. 가뭄예측 성능은 우리나라 61개 종관기상관측소(ASOS) 위치에 대하여, 가장 극심한 세 가뭄 카테고리에 대한 분류의 정확성과 회귀 오차(Mean Absolute Error, MAE)의 크기를 통해 평가하였다.

가뭄예측에서 기후모델의 예측 성능이 중요한 역할을 하게 되는데 이의 향상에 대한 연구는 본 연구의 범위에 포함되지 않았다. 여기에서는 장기예측 자료의 예측성을 평가하고자 누적 강수량 및 잠재증발산량의 기간 중 미래 기간에 대해 기후값을 이용하는 경우를 기준선으로 선정하고 장기예측 자료를 이용하는 경우의 성능을 평가하였으며 완벽한 예측을 가정하여 향상시킬 수 있는 성능의 범위를 제시하였다. 미계측지역의 가뭄정보 제공을 위하여 지점에서의 가뭄정보를 보간법을 이용하여 제공하는 경우와 기계학습을 통해 도출하는 방법을 비교하였다.

지점에 대하여 장기예측 자료를 이용한 경우(Long-Range Forecast Method, F방법)를 기준선인 기후값을 이용한 경우(Climatology Method, C방법)와 비교하였을 때 분류 및 회귀 모두 C방법이 우수하였다. 예측의 선행시간이 커질수록 분류 정확성이 급격히 감소하였으며 회귀 오차도 크게 증가하였다.

미계측지역에서의 가뭄예측을 위해 지점 결과를 공간보간한 방법(Climatology-Interpolation Method, C-I방법; Long-Range Forecast-Interpolation Method,

F-I방법)과 기계학습 모델링을 통해 예측한 경우(Climatology-Machine Learning Method, C-ML방법; Long-Range Forecast-Machine Learning Method, F-ML방법)를 비교하였다. 보간법과 기계학습 모델링을 비교하면 분류와 회귀의 경우 모두 기계학습 모델링을 이용한 방법이 보간법보다 성능이 우수했으며 장기 예측 자료를 이용한 경우와 기후값을 이용한 경우를 비교하면 대부분의 경우 기후값을 이용한 경우가 우수했다. 따라서 대부분의 경우 C-ML방법이 가장 우수한 가뭄예측 성능을 나타냈다.

미계측지역에 대해 완벽한 예측을 가정한 경우(Perfect Forecast-Interpolation, PF-I방법; Perfect Forecast-Machine Learning, PF-ML방법)를 F-I방법 및 F-ML방법과 비교하였다. 분류 정확도는 보간법을 이용한 PF-I방법의 경우 SPI6와 SPEI6에 대하여 대략 0.34-0.38 정도의 값을 가져 장기에측의 예측 성능이 완벽해 진다고 하더라도 미계측지역에 대한 분류 정확도는 최대가 0.4 이하에 머무르는 것으로 나타났다. 기계학습 모델링을 이용한 방법은 0.39-0.56에 이르는 분류 정확도를 보였는데, SPI6와 SPEI6 예측의 경우 모두 ET모델이 가장 우수한 성능을 보였다. 회귀의 경우 완벽한 예측을 가정한 PF-I방법의 경우 오차가 0.47-0.53 수준이었는데, 기계학습을 이용한 경우 이보다 더 낮은 0.35-0.42의 오차 수준을 보였으며 특히 RF모델과 ET모델의 성능이 매우 우수했다. 결과적으로 장기에측의 예측 성능이 향상되는 경우 F-ML방법의 성능이 PF-ML에 가까워짐에 따라 가뭄예측의 성능이 C-ML방법을 능가하게 될 것으로 예상된다.

NDWI 예측 모델의 경우 미래 기간에 대해 기후값을 사용하는 경우(Vegetation-Climatology-Machine Learning Method, V-C-ML방법)과 장기에측 자료를 사용하는 경우(Vegetation-Long-Range Forecast-Machine Learning, V-F-ML방법)를 비교하였다. 분류의 경우 선행시간 1-3개월에서 V-C-ML방법의 분류 정확도가 0.52-0.53에 이르러 매우 높았는데, 이는 가뭄지수 예측의 경우 완벽한 예측을 가정한 경우 가장 높은 성능을 보인 SPEI6 예측을 위한 PF-ML방법(ET모델)의 성능과 유사한 값이다. 선행시간 4-6개월에는 특히 장기에측 자료를 이용

한 V-F-ML방법(ET모델)의 분류 정확도가 0.51-0.53으로 가장 높게 나타났다. 회귀의 경우도 선행시간에 대한 오차증가의 폭이 크지 않았다. 완벽한 예측을 가정한 경우 많은 경우 오히려 예측 성능이 약간 감소하거나 향상되는 범위가 미미하였다.

종합적으로 보면 기후값을 이용하여 기계학습 모델링을 적용시킨 경우가 대부분의 경우 가장 가뭄지수 예측 성능이 우수하게 나타났다. 장기예측 자료가 가뭄 예측에 기여하는 바는 크지 않았으나 위성 자료를 이용한 기계학습 모델링을 통해 가뭄예측의 성능 향상에 기여할 수 있었으며 앞으로 장기예측의 성능이 향상됨에 따라 가뭄예측의 성능도 향상되어 기후값을 이용한 기준값보다 우수한 결과를 나타낼 것으로 기대된다.

결론적으로 본 연구에서 사용된 $0.05 \times 0.05^\circ$ 해상도에서 SPI6, SPEI6 가뭄지수 예측을 수행하는 경우 장기예측 자료를 사용하지 않고 위성자료의 기후값을 이용하여 기계학습 모델링을 수행, 공간적으로 연속적으로 분포된 가뭄정보를 생산하는 것이 가장 적합하다는 결론을 얻을 수 있다. 미세측지역에서 C-ML방법을 이용하여 기대할 수 있는 분류 정확도는 1개월 선행시간의 경우 0.47-0.52 수준이며 선행시간이 길어질수록 감소하여 선행시간 6개월의 경우에는 0.21-0.35 정도의 값을 가지게 된다. 회귀 오차 MAE의 경우 1개월 선행시간에서 0.41-0.47 수준을 얻을 수 있으며 선행시간이 길어질수록 증가하여 6개월 선행시간에서 0.56-0.59 정도를 기대할 수 있다. 장기예측의 성능이 향상되면 장기예측을 이용하는 경우가 기후값을 이용하는 경우보다 예측 성능이 더 커질 수 있는데, 이 경우 장기예측 자료와 위성자료를 이용하여 기계학습 모델링을 수행하는 경우에 완벽한 예측을 가정한 가뭄예측 성능의 상한은 분류 정확도가 0.50-0.56, 회귀 오차 MAE는 0.35-0.40 수준이다.

기계학습 모델링을 이용한 NDWI의 예측은 가뭄지수 예측과는 달리 장기예측 자료의 예측성에 크게 민감하지 않으면서 모든 선행시간에 대해 비교적 높은 분류 정확도와 작은 회귀 오차를 보였다. 가뭄지수 예측의 경우 선행시간이 길어짐에

따라 예측 성능이 감소하는 단점을 보완할 수 있으며, 식생의 수분 상태를 나타내는 지수로서 가뭄의 환경피해와 연관지어 가뭄지표로서 활용할 수 있다. 따라서 장기예측 자료의 예측성을 개선시켜 의사결정자에게 익숙한 여러 가뭄지표의 예측 성능을 함께 향상시키는 노력을 지속함과 동시에 NDWI와 같이 가뭄의 환경피해를 나타내는 가뭄지표의 예측을 통한 가뭄정보를 생산하고 제공하여야 한다. V-C-ML방법의 경우 기대할 수 있는 분류 정확도는 0.51-0.53 정도이며 회귀 오차는 정규화된 MAE를 보았을 때 0.45-0.50 정도가 된다. 본 연구에서는 정규수분지수인 NDWI를 그 자체로 가뭄지표로 삼아 예측하였으나 이러한 위성 기반의 지수를 다른 변수와 융합하여 가뭄지수를 구성하는 SDCI(Scaled Drought Condition Index) 등 여러 가지 가뭄의 환경피해와 관련된 가뭄지표를 활용할 수 있다.

본 연구에서 기후값 및 단순한 방법론을 이용한 장기예측 자료를 통해 가뭄지수 예측의 성능을 도출하였는데, 이를 우리나라에서 수행하는 가뭄지수 예측 연구에 기준선으로서 활용할 수 있을 것이다. 또한 의사결정에 있어 불확실성에 효과적으로 대응하고자 본 연구를 통해 개발된 지역 규모의 활용을 위한 고해상도 가뭄예측 방법을 기상청의 가뭄정보시스템 등과 상호보완적으로 결합하여 활용하길 기대한다.

목 차

1. 서론	1
1.1 연구의 배경 및 필요성	1
1.2 연구의 목적 및 범위	4
1.2.1 연구의 목적	4
1.2.2 연구의 범위	5
1.3 연구 방법 및 자료	7
1.3.1 연구 방법	7
1.3.2 자료	15
1.4 선행연구와의 차별성	20
2. 가뭄의 정의 및 분류	23
2.1 가뭄의 정의와 가뭄 지표	23
2.1.1 가뭄의 정의	23
2.1.2 가뭄 지표	24
2.2 가뭄 발생의 원인	29
2.2.1 국외 가뭄 발생의 원인	29
2.2.2 국내 가뭄 발생의 원인	30

2.2.3	국내 과거 가뭄 발생	34
2.3	가뭄의 영향과 분류	39
2.3.1	정규수분지수를 이용한 가뭄 영향 분석	39
2.3.2	가뭄 피해 유형 분류	43
3.	가뭄 정보의 활용	45
3.1	가뭄 정보의 활용 현황	46
3.1.1	과거 가뭄 피해	46
3.1.2	가뭄 정보의 활용	50
3.2	가뭄 리스크 평가	52
3.3	가뭄 정보 수요	55
3.3.1	가뭄감시 정보의 수요	55
3.3.2	가뭄예측 정보의 수요	55
3.3.3	미계측지역의 가뭄 정보 수요	61
3.3.4	가뭄 정보의 정확성 수용	62
3.4	소결	62
4.	가뭄예측 모델의 개발	64
4.1	지점에 대한 가뭄지수 예측	65
4.1.1	가뭄지수의 선정	65
4.1.2	성능 평가 기준	65
4.2	미계측지역에 대한 가뭄지수 예측	67
4.2.1	공간보간법	67
4.2.2	기계학습 모델링	68

4.3 정규수분지수 예측	76
4.3.1 기계학습 모델링	76
5. 모델 간 비교 결과	81
5.1 지점에 대한 가뭄지수 예측 결과	81
5.1.1 분류 정확도	81
5.1.2 회귀 성능	81
5.2 미계측지역에 대한 가뭄지수 예측 결과	84
5.2.1 분류 정확도	84
5.2.2 회귀 성능	84
5.2.3 항상 가능한 범위	87
5.2.4 변수의 상대적 중요도	90
5.3 미계측지역에 대한 정규수분지수 예측 결과	91
5.3.1 분류 정확도	92
5.3.2 회귀 성능	94
5.3.3 항상 가능한 범위	96
5.3.4 변수의 상대적 중요도	98
6. 토론 및 결론	100
6.1 가뭄 정보 수요 및 활용 의사	100
6.2 기후값과 장기예측	101
6.3 보간법과 기계학습 모델링	102
6.4 가뭄지수와 정규수분지수	102
6.5 결론 및 제언	102

▮ 참고 문헌	105
▮ 부록 1. 설문조사지	112
▮ 부록 2. 설문조사지	121

1. 서론

1.1 연구의 배경 및 필요성

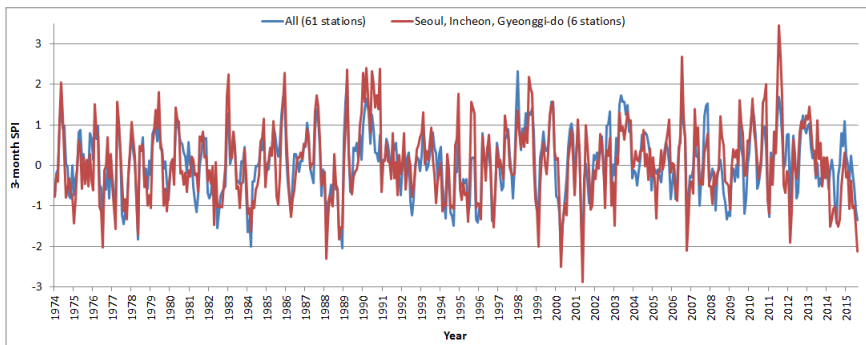
가뭄에 대응하는 것은 인류의 지난한 과제이다. 인간 생활을 물이용과 별개로 생각할 수 없는 만큼 가뭄으로 인한 피해는 언제나 존재하였다. 지금도 그렇지만 특히 과학기술이 현재와 같이 발달하지 못하였던 과거에는 가뭄으로 인해 크나큰 경제적 피해와 많은 사상자가 발생하였다. 재난역학연구소(Center for Research on the Epidemiology of Disasters, CRED)의 국제 재난 데이터베이스(EM-DAT, 2014)에 따르면 1928년 중국에서 발생한 가뭄으로 3백만 명이, 1943년 방글라데시의 가뭄으로 190만 명이 사망하였으며 2002년 말라위의 가뭄으로 5백 명이 사망하였다. 경제적인 피해는 1994년 중국에 발생한 가뭄으로 인한 피해가 137억 달러 이상, 2012년 미국의 가뭄으로 인한 피해는 2백억 달러, 최근 2014년 브라질의 가뭄으로 인한 피해는 43억 달러로 추산된다(EM-DAT, 2014).

우리나라에도 가뭄 피해가 수없이 발생하였는데, 특히 1990년 이전에 대해서는 1939년, 1968년, 1978년, 1982년에 발생한 가뭄을 4대 가뭄으로 들고 있다(심기오, 2009). 1939년 가뭄의 경우 해양성 열대기단인 북태평양 고기압의 지나친 발달에 의해 발생하였으며, 1982년 가뭄은 초기에는 해양성 한대기단의 지나친 발달에 의해서, 후기에는 북태평양 고기압의 지나친 발달에 의해서 발생한 것으로 추정된다(농림부, 2000). 1990년 이후에는 1994-1995년에 전국적으로 매우 심각한 가뭄이 발생하였고, 2000년대에 발생한 가뭄 중에는 2001년 봄 중국에서 이동한 이동성 고기압의 영향을 받아 중부지방을 중심으로 발생하였던 가뭄이 매우 심하였던 것으로 기록되었다(심기오, 2009). 2008-2009년에는 평년보다 적은 강수량으로 가뭄이 발생하였으며 특히 강원도 태백시에 생활용수 부족으로 인한 피해가 있었다(심기오, 2009). 2012년에는 봄철 전국 평균 평년의 30%에 미치지 못하는 강수량으로 전국적으로 심한 가뭄이 발생하여 농업용수 및 생활용수를 공급하는 등 가뭄 극복을 위한 노력이 있었다(김연주 외, 2013).

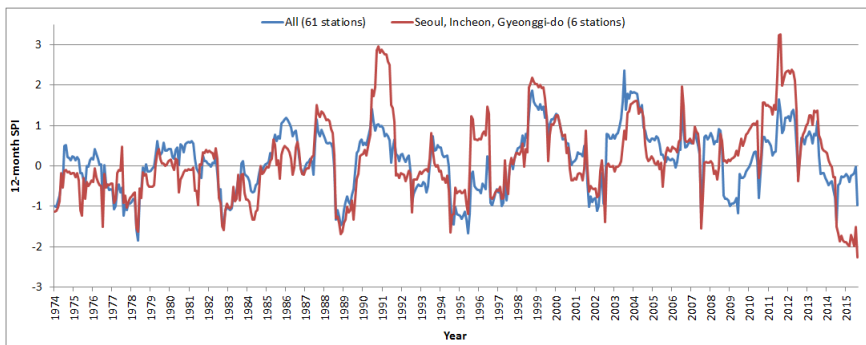
2015년에도 극심한 가뭄이 지속되고 있는데, 1974년에서 2015년 8월까지의 가뭄 상태를 누적 강수량의 정규분포곡선에서의 위치를 통해 가뭄의 심도를 나타내주는

2 | 위성 및 장기예측자료의 기계학습을 통한 가뭄예측

표준강수지수(Standardized Precipitation Index, SPI)로 살펴보았다(Figure 1-1, Table 1-1). 단기가뭄의 심도를 나타내는 3개월 표준강수지수(SPI3) 값을 보면 전국 61개 종관기상관측시스템(Automated Synoptic Observing System, ASOS) 지점 평균의 경우 Moderate Drought 상태를 보이고 있으나, 서울, 인천, 경기의 6개 관측 지점 평균의 경우 Extreme Drought 상태를 나타내고 있음을 알 수 있다(Figure 1-1a). 장기가뭄의 심도를 나타내는 12개월 표준강수지수(SPI12)를 보면 전국 평균의 경우 2014년 7월경의 Moderate Drought에서 회복되던 중 급격히 심화되어 Near Normal 에서 다시 Moderate Drought 으로 넘어가는 상태에 있는 반면, 서울, 인천, 경기의 6개 관측 지점 평균의 경우 2013년 중반부터 가뭄 상태가 꾸준히 악화되어 가장 심각한 단계인 Extreme Drought에 이르러 그림에 나타낸 1974년 이후 가장 심한 가뭄이 지속되고 있음을 알 수 있다(Figure 1-1b).



(a)



(b)

Figure 1-1 Drought conditions based on (a) 3-month Standardized Precipitation Index (SPI3) and (b) 12-month Standardized Precipitation Index (SPI12).

Table 1-1 Drought classes of SPI and SPEI.

Classification	Index Value (Unitless)
Extremely wet(EW)	≥ 2.00
Very wet(VW)	1.50 to 1.99
Moderately wet(MW)	1.00 to 1.49
Near Normal(NN)	0.99 to -0.99
Moderate drought(MD)	-1.00 to -1.49
Severe drought(SD)	-1.50 to -1.99
Extreme drought(ED)	≤ -2.00

우리나라에는 가뭄에 대한 개별 법령은 존재하지 않지만 22개의 법령이 본문에 가뭄에 대한 항목을 포함하고 있는데, 이 중 법률은 「재난 및 안전관리 기본법」, 「자연재해대책법」, 「수도법」, 「농어업재해대책법」, 「내수면어업법」이 있다. 가뭄은 「재난 및 안전관리 기본법」 제3조에 자연재난으로 명시되어 있으며 「농어업재해대책법」은 농업재해의 일종으로 가뭄을 포함시키고 있다. 또한 「자연재해대책법」에 제29조부터 제33조까지 가뭄의 예방 및 대비를 위한 중앙 정부 부처 및 지방자치단체의 책무와 협조를 위한 법적 근거가 마련되어 있다. 중앙재난안전대책본부의 「2015년 가뭄대비 관계부처 합동지침」에 따르면 가뭄 발생 시 국민안전처가 총괄을 맡고 기상청에서 기상정보를 제공하며 농업분야, 상수도분야, 용수공급분야에서 각각 농림축산식품부, 환경부, 국토교통부가 대책을 수립하고 집행하도록 되어 있다. 이로써 가뭄의 발생에 대한 예방, 대비, 대응 및 복구에 이르는 가뭄재난관리의 전 과정을 이루게 되는데, 실제로는 가뭄 발생 후에 가뭄 피해 평가에서부터 대응, 복구에 이르는 일반적인 재해관리 또는 위기관리의 과정에 좀 더 집중하게 된다. 그러나 선제적인 대응을 통한 가뭄 피해 최소화를 위해서는 예방과 대비가 무엇보다도 큰 중요성을 가지며, 이를 위해서는 보다 정확한 가뭄 정보의 제공이 우선적으로 이루어져야 한다.

지금까지 가뭄의 발생을 진단하기 위해 여러 부처에서 가뭄감시 및 조기 경보 시스템을 구축하여 운영하여 왔다. 기상청은 가뭄정보센터를 설립하여 여러 가뭄지수를 제공해 왔으며 소방방재청은 국가 가뭄재해정보시스템 개발 사업을 수행하였고 한국수자원공사에서는 가뭄모니터링시스템을 개발하였으며 한국농어촌공사는 농업 가뭄평가시스템을 운영하고 있다(배덕효, 손경환, 2012). 그러나 이러한 시스템에 대

한 접근은 제한적이어서, 예를 들어 Figure 1-1에서 살펴본 표준강수지수를 통한 가뭄 심도에 대한 시계열 등의 정보 등은 쉽게 얻을 수 없다. 또한 대부분 가뭄감시를 위한 시스템이어서 가뭄 피해를 최소화시킬 수 있는 가뭄예측 정보의 제공은 미흡한 실정이다. 의사결정자의 목적에 맞는 가뭄지표를 이용한, 지역에 쉽게 적용할 수 있는 고해상도 가뭄감시 및 예측 정보의 제공이 필요하다.

가뭄예측 정보가 제공된다고 하더라도, 의사결정자들에게 잘 이용되지 않을 수도 있다. 가뭄예측 정보는 가뭄 대응을 위한 사회·경제적 자산 운용에 대한 사전 계획 및 수정을 가능하게 하여 가뭄 취약성 감소에 큰 도움을 줄 수 있으나, 실제 활용에 있어서는 여러 장애가 존재할 수 있다. 특히 기후 정보를 수집하고 이해하는 과정, 제도, 인지와 관련된 장벽이 있을 수 있으며, 기후 정보를 받아들인다 할지라도 기존의 업무를 변경하고자 하는 역량이나 의사결정자의 의지가 부족할 수 있다(Cutter et al., 2012). 그리고 자료 생산자와 사용자간의 목적, 필요, 영역, 우선순위, 기관 환경이나 규범, 문화적인 가치, 이해의 차이와 불신으로 인해 공동의 문제 도출 및 협력을 저해하는 기능적, 구조적, 사회적인 요소들이 존재한다(Cutter et al., 2012). 지속적인 가뭄관리는 앞서 언급한 재해관리의 요소 뿐 아니라 가뭄감시·예측정보를 이용한 의사 결정, 가뭄 리스크 평가와 이를 기반으로 한 여러 대응책의 선정 및 평가, 이행 및 감시, 그리고 되먹임에 이르는 통합적 리스크 관리 체계를 통해 이루어질 수 있는데, 이를 위해 우선적으로 의사결정자의 가뭄 정보 수요 및 활용 의사 도출 등 연구자와 의사결정자 간의 의사소통이 필요하다.

1.2 연구의 목적 및 범위

1.2.1 연구의 목적

본 연구에서는 적합한 가뭄예측 정보 제공을 통해 앞으로 가뭄 취약성 및 실제 피해를 감소시킬 수 있도록 의사결정자의 목적에 맞는 가뭄지표를 이용한, 지역에 쉽게 적용할 수 있는 고해상도 가뭄예측 정보를 제공할 수 있는 모델을 개발하고자 한다. 가뭄예측은 미래에 발생할 수 있는 기상관련 사상을 예측하는 것이므로 기후모델의 예측 성능이 가장 중요한 역할을 하게 되지만 본 연구에서는 기후모델의 예측 성능을

향상시키는 것은 포함시키지 않고 위성 자료와 장기예측 자료를 이용하여 기계학습 모델링을 통해 가뭄예측을 하는 경우 그 성능을 파악하고 장기예측의 성능이 향상되는 경우 가뭄예측이 얼마나 향상되는지 그 향상 가능한 범위를 제시하고자 한다.

우리나라를 대상지역으로 한 최근의 연구가 보여주듯이(e.g., Rhee and Cho, 2016), 기후변화로 인해 이상가뭄이 빈발할 것으로 전망되므로 물 확보 및 관리 여건은 점점 더 어려워 질 것으로 보인다. 이러한 상황에서 「제4차 국토종합계획 수정계획(2011-2020)」에서는 국토형성의 기본목표로 기후변화로 인한 홍수·가뭄 등 재해에 안전한 국토 구현을 명시하고 있는데, 이의 달성을 위해 여러 구조적인 가뭄 대책이 수립되어 실행되고 있다. 본 연구에서는 대표적인 비구조적 가뭄 대책으로서 가뭄예측 정보 생산을 위한 모델 개발을 수행하였다. 이를 통해 차후 가뭄예측 정보를 실제 지역 규모에서 활용하도록 하는 방안을 도출하여 단발성이 아닌 지속적인 재난관리가 이루어지도록 할 것이다. 본 연구의 목적은 (1) 설문을 통해 의사결정자들의 가뭄정보에 대한 수요를 파악하고, (2) 이를 반영한 가뭄예측 모델을 구축하고 성능을 평가하며, (3) 장기예측의 성능이 향상되는 경우 가뭄예측이 얼마나 향상될 지 그 향상 가능한 범위를 제시하는 것이다. 본 연구의 흐름도를 Figure 1-2에 나타내었다.

1.2.2 연구의 범위

연구의 공간적 범위는 우리나라이다(Figure 1-3). 가뭄예측의 시간적 범위는 2003년 1월-2015년 8월이며 선행시간은 1-6개월로 정하였다. 우리나라 면적은 대략 100,284 km²로, 토지이용은 2014년 기준으로 19%가 논밭, 63.8%가 임야 등으로 이루어져 있다(국가통계포털, 2015a). 주요 도시 중 서울, 인천, 강릉, 목포, 부산의 연평균 기온은 1990년 기준으로 11.4-14.1 °C에 분포하며 평균 강수량은 1,107.5-1,444.7 mm이다(국가통계포털, 2015b).

가뭄 피해는 크게 용수공급 피해, 농수축산업 피해, 경제적 피해, 환경적 피해, 사회적 피해로 나눌 수 있는데(김연주 외, 2013), 위성 영상 자료의 공간적으로 상세한 정보는 식생의 수분 상태 등을 잘 도출하여 특히 가뭄의 환경적 피해를 감시하거나 예측하는데 유용하다. 본 연구는 가뭄의 환경 피해에 초점을 맞추어 수행되었다. 또한 연구의 목적에서 언급하였듯이 역학적 및 통계적 다운스케일링 등을 통한 기후모

6 | 위성 및 장기예측자료의 기계학습을 통한 가뭄예측

델의 예측 성능 향상은 본 연구의 내용적 범위에 포함시키지 않았으며, 단순한 장기 예측 결과를 이용했을 경우의 가뭄예측 성능을 의사결정자에게 제시하고 장기예측 성능이 향상되는 경우를 가정한 가뭄예측 성능과의 비교를 통해 향상 가능한 가뭄예측의 범위를 제시하고자 하였다.

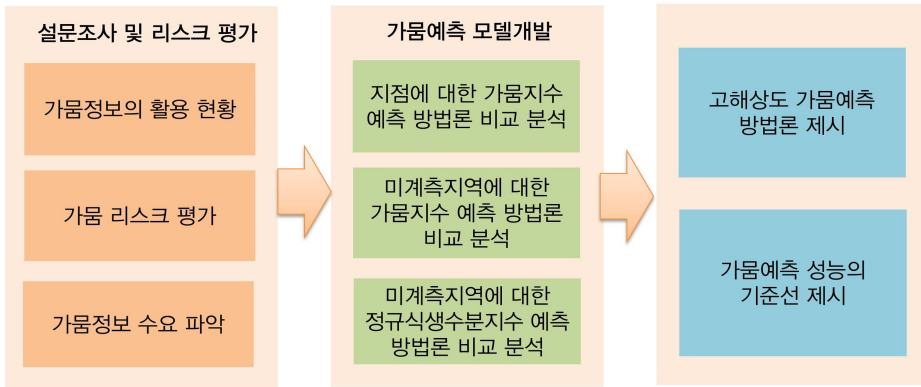


Figure 1-2 Flowchart of the study.

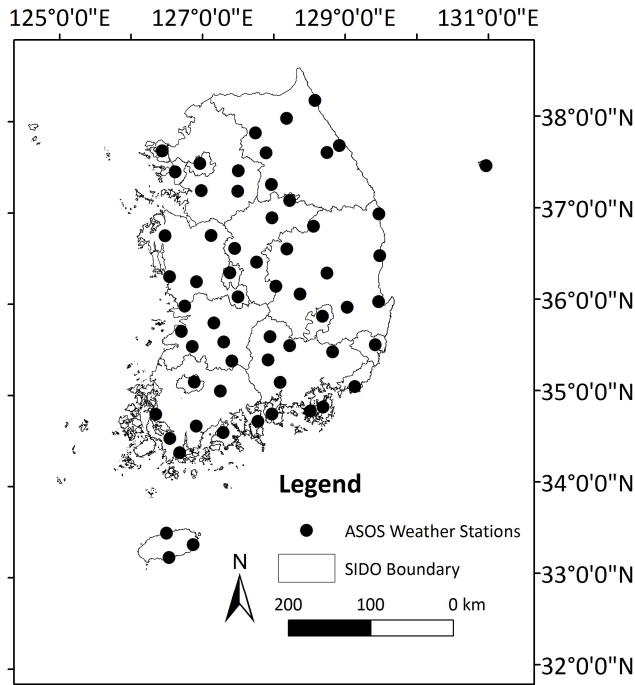


Figure 1-3 Study area.

1.3 연구 방법 및 자료

1.3.1 연구 방법

1.3.1.1 설문조사 및 리스크 평가

의사결정자의 가뭄 정보 수요 파악은 17개 광역지방자치단체의 가뭄 및 가뭄의 환경 피해와 관련이 있는 담당자를 대상으로 한 설문조사를 통해 수행되었다. 본 설문조사는 표본을 통해 모집단의 특성을 도출하려는 표본조사보다는 실제 가뭄 정보의 활용과 밀접한 광역지방자치단체 공무원을 대상으로 하는 인터뷰를 대체하도록 수행되었다. 두 차례에 걸친 설문조사는 전체 30문항으로 이루어져 있으며 이를 통해 가뭄 발생 인지 및 업무 관련성, 가뭄정보의 활용, 가뭄 대처 현황 및 수요, 가뭄정보의 정확성에 대한 수요 등을 도출하고 가뭄 리스크 평가를 수행하였다.

리스크에 대한 정의가 다양한데, 대부분 확률에 기반을 둔 개념이다(Table 1-2). 즉, 리스크는 재난(disaster), 피해(loss)의 확률인데 이 재난을 유발시키는 것이 위해요소(hazard)가 된다. Smith(1996), IPCC(2001), Adams(1995), Jones and Boer(2004) 등은 리스크를 위해요소의 발생확률과 그 피해 내지 결과의 곱으로 보았다. 반면 Downing et al.(2001)는 위해요소를 인류와 복지에 위협을 가하는 재난과 같은 개념으로 정의하여 리스크를 위해요소의 확률이라고 정의했다. Stenchion(1997)과 UNDRR(1992)는 리스크를 사회적 취약성과 연관시켜 위해요소의 발생 확률과 취약성의 곱으로 보았으며 Crichton(1999)는 위해요소, 취약성, 노출의 함수로 정의하였다. 본 연구에서는 Jones and Boer(2003)의 정의대로 리스크를 위해요소의 발생확률 또는 발생 빈도와 피해 강도의 곱으로 정의하였다.

$$\text{리스크} = \text{발생빈도} \times \text{피해강도} \quad (\text{Eq. 1-1})$$

Table 1-2 Definitions of risk (adopted from Brooks, 2003).

Author(s)	Risk definition
Smith, 1996	Probability × loss (probability of a specific hazard occurrence) Hazard = potential threat
IPCC, 2001	Function of probability and magnitude of different impacts
Morgan and Henrion, 1990; Random House, 1966	“Risk involves an ‘exposure to a chance injury or loss’”
Adams, 1995	“a compound measure combining the probability and magnitude of an adverse affect”
Jones and Boer, 2004; Helm, 1996	Probability × consequence Hazard: an event with the potential to cause harm, e. g. tropical cyclones, droughts, floods, or conditions leading to an outbreak of disease-causing organisms.
Downing et al., 2001	Expected losses (of lives, persons injured, property damaged, and economic activity disrupted) due to a particular hazard for a given area and reference period Hazard: a threatening event, or the probability of occurrence of a potentially damaging phenomenon within a given time period and area.
Downing et al., 2001	Probability of hazard occurrence Hazard = potential threat to humans and their welfare
Crichton, 1999	“Risk” is the probability of a loss, and depends on three elements, hazard, vulnerability and exposure.”
Stenchion, 1997	“Risk might be defined simply as the probability of occurrence of an undesired event [but might] be better described as the probability of a hazard contributing to a potential disaster...importantly, it involves consideration of vulnerability to the hazard.”
UNDHA, 1992	“Expected losses (of lives, persons injured, property damaged, and economic activity disrupted) due to a particular hazard for a given area and reference period. Based on mathematical calculations, risk is the product of hazard and vulnerability.”

일반적인 리스크 관리 체계는 문제점 및 목적의 파악, 의사 결정 기준의 수립, 리스크 평가, 대응책의 선정 및 평가, 이행, 감시로 이루어지는데 이 중 리스크 평가는 위해요소와 피해의 유형을 파악하고 각각에 대해 피해 정도와 확률을 고려하여 리스크의 크기를 도출하는 과정을 포함한다(Figure 1-4). 가뭄 리스크 평가 과정을 거치면 가뭄과 관련된 전체 피해 목록 중 본 지역에 해당하는 피해를 가려내고 각 피해

유형의 정도와 확률을 알아내기 위한 가뭄 정보 수요를 도출하는 데 도움이 된다.

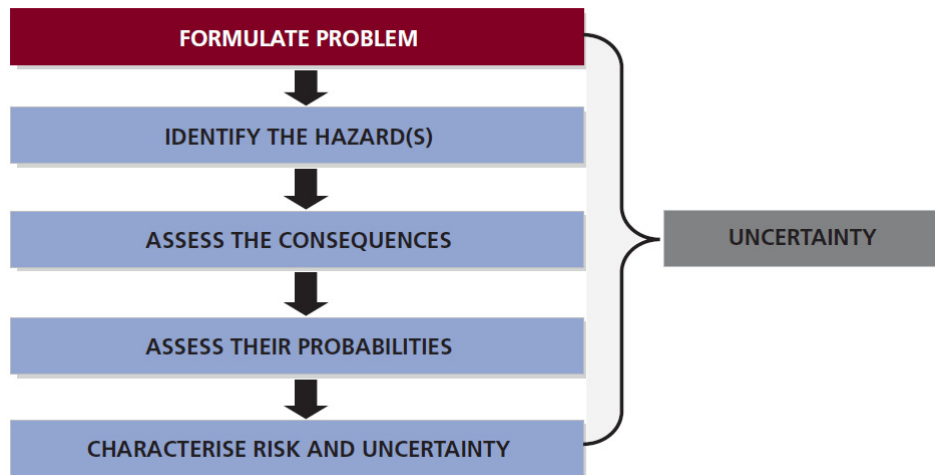


Figure 1-4 Risk Assessment Processes (adopted from DEFRA, 2011).

1.3.1.2 기계학습 모델링

고해상도 가뭄예측 모델 개발은 기계학습 모델링을 통해 수행하였다. 기계학습은 인공지능 및 전문가시스템 개념으로부터 출발하였는데, 인공지능은 현재는 사람이 더 잘하는 일을 컴퓨터가 더 잘할 수 있도록 연구하는 학문 분야로 1950년대 중반 시작되었다. 1960년대 과신되는 경향이 있었지만 사람이 하는 일을 컴퓨터가 모두 잘하도록 하는 일이 실제로 어렵다는 것이 판명되어 결과적으로는 상대적으로 좁은 학문 분야에 적용하도록 발전되어 왔다(Jensen, 2005). 이후 어떤 특정한 좁은 주제의 문제 해결에 있어 보통 인간의 지능이 요구되는 지식을 활용하는 시스템을 일컫는 전문가 시스템(expert system)의 개념이 등장하였는데, 전문가 시스템을 통해 좁은 영역의 문제를 효율적, 효과적으로 해결할 수 있다(Jensen, 2005).

전문가 시스템으로 문제를 해결하기 위해 활용하는 자료 및 규칙(rule)이 지식기반(knowledge base)이며, 해당 영역의 지식기반이 형성되면 동일 영역의 다른 문제에 대해서 새로운 프로그램 없이 지식기반을 통해 해답을 제시할 수 있다. 지식기반을 형성하는 방법에는 전통적인 컴퓨터 프로그램을 통한 알고리즘을 이용하는 방법과

자기발견적인(heuristic) 전문가 시스템을 이용하는 방법이 있다. 자기발견적인 전문가 시스템이 필요로 하는 규칙은 인간 전문가에 의해 개발되거나, 인간이 제공한 훈련 자료를 이용하여 기계가 도출할 수 있는데, 이러한 학습 과정의 컴퓨터 모델링을 기계학습이라고 정의한다(Jensen, 2005). 뒤에 언급한 Decision Tree(의사결정분지도 또는 의사결정트리) 모델은 원격 탐사 영상 분석을 위한 귀납적 기계학습 알고리즘의 하나이다. 기계학습 알고리즘은 인간 전문가의 간섭 없이 제공된 훈련 자료로부터 규칙과 조건을 생성하여 전문가 시스템을 훈련시켜 문제를 해결한다(Huang and Jensen, 1997).

기계학습을 원격 탐사 분야에 적용하는 경우 여러 장점이 있다. 원격 탐사 자료는 많은 경우 추정 알고리즘을 통해 도출되므로 관측 자료에 비해 부정확한 단점이 있다(Khaleghi et al., 2013). 또한 대기, 에어로졸 등 정확한 자료의 관측을 저해하는 많은 요인이 있어 불확실성을 내포하기도 한다. 이런 경우 기계학습의 사용은 도움이 될 수 있는데, 이러한 방법론에서는 변수들 간의 물리적 관계를 명시적으로 밝힐 필요가 없고 이상치(outlier)에 안정적인 결과를 생산하기 때문이다(Mishra and Desai, 2006). 여러 위성 영상에서 도출한 변수를 결합하여 가뭄을 예측하는 모델을 개발하는 본 연구에 특히 적합하다.

본 연구에서는 Decision Tree, Random Forest, Extremely Randomized Trees의 세 가지 기계학습 모델을 이용하였다. 각각의 특징을 살펴보면 다음과 같다.

Decision Trees(DT모델)

Decision Tree(DT모델)는 지식기반 전문가 시스템의 일종으로 가설, 규칙, 조건을 이용하여 일련의 속성을 가진 입력 변수의 벡터를 받아들여 목표 변수의 결과 값을 생산해 낸다(Jensen, 2005; Figure 1-5). DT모델은 기계학습 방법론 중 비매개 변수 감독 분류 방법론의 하나로 분류(classification) 및 회귀(regression) 모델 구축을 목적으로 한다. 모델을 통해 얻고자 하는 결과가 불연속인 경우 모델의 목적은 분류가 되며, 연속적인 결과를 얻고자 하는 경우 회귀가 된다. DT모델은 white box 모델로 훈련한 트리를 가시적으로 나타낼 수 있으며 트리를 통해 모델의 구조를 이해하기 용이하다는 장점이 있다.

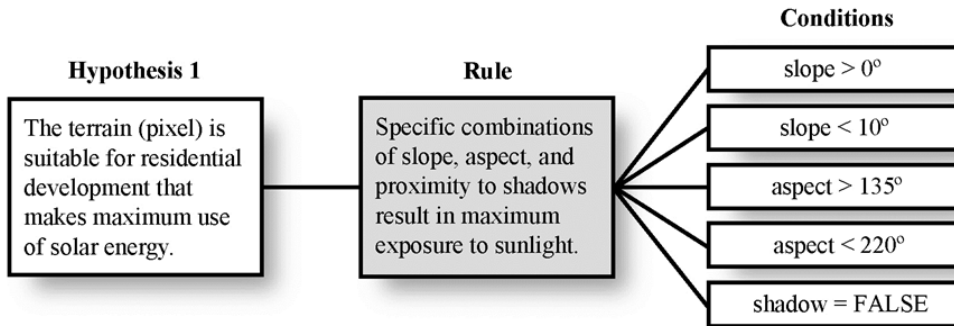


Figure 1-5 A decision-tree expert system with a rule and conditions to test a hypothesis [Jensen, 2005].

DT모델은 여러 잎(leaf node)과 가지(branch)로 구성되는데, 한 잎노드로는 하나의 가지만 들어올 수 있고 하나의 잎노드에서는 다수의(보통 두 개 정도) 가지가 뻗어나간다(Figure 1-6). 트리의 훈련 단계에서는 학습에 사용되는 입력 변수의 부분 집합을 각 잎에서 나누는 분할 함수(split function)의 매개변수를 최적화하는데, 이 과정은 잎에서의 부분 집합이 모두 같은 목표 변수 값을 가지거나 분할로 인해 예측이 더 이상 향상되지 않을 때까지 반복된다. 본 연구에서도 이러한 하향 귀납식(top-down induction) Decision Tree를 이용하였다.

각 잎에서 목표 변수의 동질성을 측정하는 기준으로는 분류의 경우 가장 일반적으로 사용되는 Gini 지수를 이용하였으며 회귀의 경우는 Mean Squared Error(MSE)를 이용하였다. Gini 지수는 부분 집합에서 샘플을 뽑아 부분 집합 내의 라벨 분포에 맞게 임의로 라벨을 주었을 때 틀리는 정도를 나타낸다. 훈련 입력 벡터 $x_i \in R^n, i = 1, \dots, l$ 과 목표 변수인 $y \in R^l$ 에 대해 N_m 개의 샘플(입력 벡터의 부분 집합)이 있는 영역 R_m 인 잎노드 m 에 대해서 Gini 지수와 MSE는 각각 다음과 같은 식으로 계산할 수 있다. 목표 변수를 0, 1, ..., K-1로 분류하는 것이 목표인 경우, p_{mk} 는 잎노드 m 에서 분류 k 의 비율이다. Decision Tree에서 변수들의 상대적인 중요도 또한 해당 변수로 인해 감소되는 Gini 지수 값으로 계산할 수 있다.

$$p_{mk} = \frac{1}{N_m} \sum_{x_i \in R_m} I(y_i = k) \quad (\text{Eq. 1-1})$$

$$Gini\ Impurity : H(X_m) = \sum_k p_{mk} (1 - p_{mk}) \tag{Eq. 1-2}$$

$$c_m = \frac{1}{Nm} \sum_{i \in N_m} y_i \tag{Eq. 1-3}$$

$$MSE : H(X_m) = \frac{1}{N_m} \sum_{i \in N_m} (y_i - c_m)^2 \tag{Eq. 1-4}$$

Decision Tree를 가뭄예측에 활용한 예를 보면 Tadesse et al.(2005)은 위성영상에서 추출한 식생 조건 자료를 SPI 등의 가뭄 지수 및 토지 이용, 토양 수분 최대 가능 함유량, 관개 지역 등과 같은 생물리적인 변수들과 함께 이용하여 2주, 4주, 6주의 다양한 지연시간(time lag)을 가지는 모델을 구축, 가뭄 상태 및 작물 수확량의 예측이 가능하게 하였다. 이를 위해 규칙 기반의 회귀 분지도(rule-based regression-tree) 모델을 사용하였다.

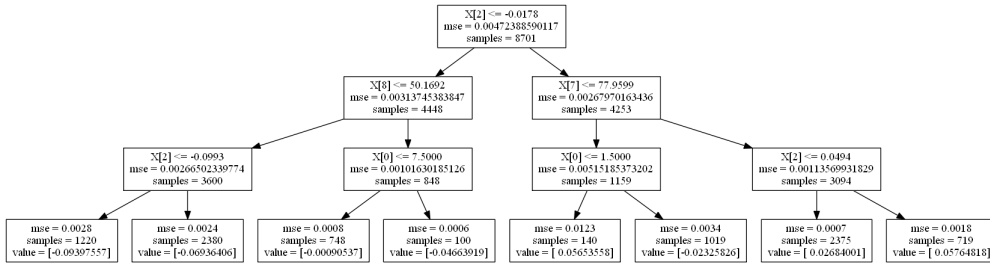


Figure 1-6 An example of a simple decision tree from a regression model tested in this study.

Random Forest(RF모델)

Random Forest는 Decision Tree에 기초한 기계학습 방법론으로, 마찬가지로 분류 및 회귀를 목적으로 한다. Decision Tree가 하나의 트리만을 이용하는데 비하여 Random Forest는 여러 개의 트리를 앙상블하여 이용하는데, 각각의 트리를 이루는 분류자(classifier) 또는 회귀자(regressor)가 모여 앙상블된 하나의 분류자 또는 회귀자를 구성하게 된다.

훈련 과정에서 여러 임의적인 요소(randomness)가 활용되는데, 배깅(bagging)이라는 과정을 통해 각각의 트리에 대하여 중복을 허용하여 입력 벡터의 임의의 부분 집합을 원래 입력 벡터 집합과 같은 크기로 생성하여 사용한다. 배깅은 Bootstrap

Aggregating의 약자로 각각의 트리의 임의의 입력 벡터는 서로 독립적이지만 동일한 확률분포를 가지고 있다. 분류의 경우 최종적으로 모든 트리가 가장 적합한 분류 클래스를 표결로 결정하거나 확률 예측 값을 평균하며, 회귀의 경우 각 트리의 목표 변수 값을 평균하거나 곱하여 최종 결과값을 생산한다. 이러한 과정을 거쳐 각 트리 간의 상관관계가 최소화되므로 정확도가 증가하여 분류 및 회귀 능력이 향상되게 된다. 각각의 잎노드에서 입력 벡터에 포함되어 있는 변수의 조합을 임의적으로 선택할 수 있는데, 이렇게 일부 변수(feature)만을 이용하는 것을 피쳐 배깅(feature bagging)이라고 한다. 이러한 임의적인 요소의 사용은 결과적으로 이상치(outlier)나 노이즈(noise)에 상대적으로 영향을 덜 받게 한다.

Random Forest는 Out-Of-Bag(OOB) 오차를 이용해 사용된 각 변수의 상대적인 중요도를 제공해 준다. OOB 오차는 훈련에 사용되지 않은 입력 벡터의 예측값과 실제 값의 뜻하는데, 어떠한 변수 A의 중요도를 측정하고자 할 때 각각의 트리에 대해 OOB 오차를 구하고 훈련 자료에서 이 변수의 값을 치환한 다음 다시 OOB 오차를 구하여 그 차를 계산하여 모든 트리에 대해 평균한다. 이를 통해 교차검증(cross-validation)과는 달리 훈련 자료의 일부를 따로 떼어 둘 필요가 없는 불편(unbiased) 오차 추정 값을 제공한다(Breiman, 2001). 그러나 본 연구에서는 자료의 기간이 짧고 분류 및 회귀 모델의 목적이 가뭄예측이므로 각 기간별 입력 벡터의 특징이 매우 다를 수 있다. 따라서 과적합(overfitting)을 피하기 위해 각 연도별로 교차검증을 수행하였다.

Extremely Randomized Trees(ET모델)

Extremely Randomized Trees는 각 잎에서 분할이 이루어지는 방식에 임의적인 요소가 추가되었다. 즉 Random Forest에서는 각 잎에서 분할이 이루어질 때 입력 벡터 부분 집합이 가장 차별적으로 나뉘는 기준을 사용하는데, Extremely Randomized Trees에서는 그 기준을 임의로 선택한 후에 그 중 가장 우수한 기준을 분할 규칙으로 선정한다. 변수들의 상대적 중요도는 RF모델과 같은 방식으로 구할 수 있다.

1.3.1.3 가뭄예측의 항상 범위 도출

장기예측의 예측성을 평가하고자 장기예측 결과를 전혀 사용하지 않은 경우와 단순한 장기예측 결과를 사용한 경우를 지점 위치에 대해 비교하였으며, 미세측지역의 가뭄예측을 위해 지점에 대한 가뭄예측 결과를 보간(interpolation)한 경우와 기계학습을 통해 훈련된 모델을 적용하는 경우를 비교하였다. 보간법은 주변의 지점 정보와 그 지점까지의 거리를 이용하여 정보를 도출하는 방법으로 전통적으로 많이 사용되어 왔으며, 기계학습은 가뭄예측의 경우 초기조건이 중요한 역할을 한다는 점을 인지하여 초기조건에 대한 정보를 보강하고 위성자료를 이용하여 공간적으로 분포된 정보를 활용하기 위해 적용되었다.

가뭄예측을 수행하는 시점과 목표 시점 사이의 아직 도래하지 않은 미래구간에 대하여 장기 기후예측 자료를 사용한 경우를 기준선과 비교하였는데, 기준선으로는 Lyon et al. (2012)의 방법론을 적용하였다. 이 방법은 관측 자료가 존재하지 않는 미래 기간에 대하여 같은 달에 해당하는 과거 관측 자료의 집합에서 Monte Carlo Simulation 방법을 이용하여 100회 정도의 샘플을 추출하고 이 들의 중간값(median)을 이용하는 방법이다. 이 방법을 통하여 강수량, 대기 온도 등 변수의 기후값을 도출하였다.

단순한 장기예측 결과 적용의 경우 Quan et al. (2012)의 방법론을 이용하였다. Quan et al. (2012)은 Global Climate Model(GCM)의 기후값과 예측값의 차인 아노말리를 구한 후 이 아노말리와 모델 기후값의 비인 증가백분율(percent increment)을 산정하고 이를 관측 자료의 기후값에 적용하여 보정된 예측값을 계산하였다.

$$P_a = P_T - P_c \quad (\text{Eq. 1-5})$$

$$P_a^* = P_a / P_c \quad (\text{Eq. 1-6})$$

$$P_{cal} = O_c(1 + P_a^*) \quad (\text{Eq. 1-7})$$

이 때, P_T, P_c, P_a 는 각각 모델의 예측값, 기후값, 아노말리이며 P_a^* 는 증가백분율, P_{cal}, O_c 는 각각 보정된 예측값 및 관측의 기후값이다.

미계측지역의 가뭄예측에 대해 가뭄예측의 향상 범위를 제시하기 위해 완벽한 장기예측이 존재하는 경우를 가정하여 과거에 대해 실제 관측한 지점 자료 또는 위성 자료에 약간의 노이즈를 추가하여 예측 성능을 평가하였다. 그리고 이를 장기예측을 이용한 경우와 비교하여 향상될 수 있는 범위를 도출하였다.

1.3.2 자료

1.3.2.1 종관기상관측자료

우리나라 61개의 ASOS 관측지점에서(Figure 1-1) 1973년 1월부터 2015년 8월까지의 월별 강수 및 대기온도 자료를 획득하였다. 대기온도 자료로부터 Thornthwaite 방법(1948)을 이용하여 잠재 증발산량을 산정하였다. Thornthwaite 방법은 기온이 섭씨 영하로 내려가는 경우 잠재 증발산량이 없다고 가정하므로 특히 우리나라 겨울철에 대하여 오차가 클 수 있으나 본 연구에서 사용하는 장기예측자료에서 제공하는 변수가 Penman-Monteith 방법이나 Hargreaves 방법을 적용할 만큼 다양하지 않으므로 일관성을 위하여 Thornthwaite 방법을 적용하였다.

1.3.2.2 기후지수

지역에 따라 장기적인 가뭄 현상은 대규모 기후 지수와 상관관계를 보이기도 한다. 특히 El Nino Southern Oscillation(ENSO)는 호주 등의 지역에서 강수량과 밀접한 관계를 나타내는데(Chiew and McMahon, 2002), 상관관계의 강도는 지구상의 각 지역에 따라 다르게 나타난다. 본 연구에서는 ENSO와 관련된 여러 변수를 결합한 다변수 ENSO 지수(Multivariate ENSO Index, MEI)와 북극진동지수(Arctic Oscillation Index, AOI)을 가뭄예측에 이용하였다.

대기와 해양의 상호작용으로 인해 2년에서 7년 주기로 동-중부 열대 태평양의 수온이 비정상적으로 높아지는 El Nino와 열대 태평양에서 동서 대기압의 차와 관련하여 나타나는 해수면 상태의 변화인 Southern Oscillation을 결합하여 ENSO로 칭한다. ENSO를 정량화하기 위한 여러 지수가 있으나, 본 연구에서는 열대 태평양의 여섯 변수를 결합한 MEI를 사용하였다. MEI에 사용된 여섯 변수는 해면 기압(Sea-Level Pressure), 동서

(zonal) 및 남북(meridional) 방향의 풍속, 해수면 온도(Sea Surface Temperature), 표면 대기 온도(Surface Air Temperature)와 총 운량(Total Cloudiness Fraction of the Sky)이며, 주성분 분석을 통해 첫 번째 비회전 주성분으로 MEI를 도출한다. 1950년부터 2015년 8월까지의 MEI 자료는 미국 NOAA의 Earth System Research Laboratory에서 획득하였다(<http://www.esrl.noaa.gov/psd/enso/mei>).

북극진동(AO)은 해면 온도의 온난하고 한랭한 상태의 주기적인 진동으로, 온난한 상태에 있을 때는 극지방에는 상대적으로 저기압이, 저위도 지방에는 평년보다 높은 기압이 분포하며, 한랭한 상태에 있을 때에는 반대로 극지방에 평년보다 높은 기압이, 중앙대서양 지역에는 저기압이 분포하게 된다. AO는 수십 년의 시간 척도를 가지고 비규칙적으로 변화한다(Aquado and Burt, 2007). 1950년부터 2015년 8월까지의 AOI 자료는 미국 NOAA의 기후예측센터에서 획득하였다(www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/monthly.ao.index.b50.current.ascii).

1.3.2.3 위성 자료

사용한 위성 자료를 Table 1-3에 나타내었다.

Table 1-3 Satellite data used in this study.

Product	Temporal Resolution	Spatial Resolution	Variable	Start Date	End Date	Latency
TRMM 3B43	Monthly	0.25 × 0.25°	Precipitation	Jan, 2001	Dec, 2012	Archived past data
GPM IMERG	Monthly	0.1 × 0.1°	Precipitation	Apr, 2014	Apr, 2015	4 Months
GPM IMERG	Daily	0.1 × 0.1°	Precipitation	May, 2015	Aug, 2015	18 hours (Late Run)
MCD43C4	16-day (Produced every 8 days with 16 days of acquisition)	0.05 × 0.05°	Bands1-7 Surface Reflectance	Mar, 2000	Aug, 2015	NA
MYD11C3	Monthly	0.05 × 0.05°	Land Surface Temperature (Day, Night)	Aug, 2002	Aug, 2015	NA
MYD13C2	Monthly	0.05 × 0.05°	NDVI	Jul, 2002	Aug, 2015	NA

강수량

강수량 자료는 미국 National Aeronautics and Space Administration(NASA)와 일본 Japan Aerospace Exploration Agency(JAXA)에 의해 공동으로 개발되어 1997년부터 자료를 생산한 Tropical Rainfall Measuring Mission(TRMM) 자료와 2014년 2월 27일에 발산된 강우와 강설의 차세대 지구관측 자료를 제공하는 국제 위성 네트워크인 Global Precipitation Measurement(GPM) 자료를 이용하였다.

TRMM의 Precipitation Radar(PR)은 강수의 여러 층을 투과하여 관측함으로써 태풍의 구조 등을 관찰할 수 있으며, TRMM Microwave Imager(TMI)는 지표면 및 대기에서 발산되는 마이크로파 에너지를 관측하여 수증기, 구름의 물 입자와 강우 강도를 정량화한다. 세 시간 단위 강수량 자료인 TRMM 3B42 자료와 미국 NOAA의 기후예측센터에서 생산한 Climate Anomaly Monitoring System(CAMS) 전지구 격자 형태의 강우 자료, Global Precipitation Climatology Center(GPCC)의 전지구 강우 관측 자료를 결합한 월 단위의 강수량 자료인 TRMM 3B43 자료를 이용하였다. 공간적 범위는 50°S-50°N, 180°W-180°E이며, 공간 해상도는 $0.25 \times 0.25^\circ$ 이다. 본 자료는 정방형 도법(Equirectangular projection or geographic projection), WGS84 데이텀(datum)으로 투영되어 있으며, 미국 Goddard Earth Sciences Data and Information Service Center(GES DISC)에서 획득하였다.

2014년 7월 8일 이후로 TRMM의 위치를 고정하기 위한 연료 부족으로 고도가 하강하여 10월 7일 이후로는 PR 자료의 제공이 중지되었으므로 2014년 4월부터는 GPM Integrated Multi-satellite Retrievals for GPM(IMERG) 자료를 이용하였다. IMERG 자료의 공간적 범위는 90°S-90°N, 180°W-180°E이며, 공간 해상도는 $0.1 \times 0.1^\circ$ 이다. TRMM 자료와 마찬가지로 정방형 도법(Equirectangular projection or geographic projection)으로 투영되어 있으며, NASA의 Precipitation Measurement Missions(PMM)에서 획득하였다(<http://pmm.nasa.gov/data-access/downloads/gpm>).

TRMM 3B43 및 GPM IMERG 자료는 mm/hr 단위로 제공되므로 mm/month 단위로 전환되었으며 여러 시간 척도에 따라 누적된 강수량을 이용할 수 있도록 1-6개월 동안 누적된 강수량 값을 구하였다. GPM IMERG 자료의 경우 월별 자료의 제공 대기 시간이 4개월이므로 2015년 5월부터 8월까지의 자료는 일별 자료(18시간 지연

의 Late Run)를 구하여 월별로 전환하였다.

지표면온도

지표면온도는 미국 NASA의 Aqua 위성에 탑재된 MODIS 센서로부터 생산된 표준 자료인 MYD11C3 Land Surface Temperature and Emissivity 자료에서 도출하였다. 본 자료의 시간 해상도는 월별, 공간 해상도는 $0.05 \times 0.05^\circ$ 이며 주간 지표면온도와 야간 지표면 온도를 모두 제공한다. 하며, 정방형 도법(Equirectangular projection or geographic projection)으로 투영되어 있으며 자료는 미국 NASA의 Earth Observing System Data and Information System(EOSDIS, <http://reverb.echo.nasa.gov>)에서 획득하였다.

Aqua 위성은 2002년 5월 4일에 발사되었으며 여섯 개의 센서를 탑재하고 있다. 이 중 Atmospheric Infrared Sounder(AIRS), Advanced Microwave Sounding Unit (AMSU), Clouds and the Earth's Radiant Energy System(CERES), MODIS 는 아직 자료를 잘 생산하고 있다. AMSR-E는 2011년 10월 4일 안테나 작동 중지로 인해 더 이상 자료를 생산하지 않으며 여섯 번째 센서인 Humidity Sounder for Brazil (HSB)는 2003년 2월 작동이 중지되었다. Aqua 위성의 강교점에서의 적도 통과 시간은 오전 1시 30분, 승교점에서의 적도 통과 시간은 오후 1시 30분이다.

식생지수

식생지수인 Normalized Difference Vegetation Index(NDVI)도 Aqua 위성에 탑재된 MODIS 센서로부터 생산되는 준 자료인 MYD13C2 Vegetation Indices 로부터 도출되었다. 가뭄으로 인해 식물이 필요로 하는 물이 부족해지는 경우 식생 상태의 변화가 비교적 잘 관찰되므로 건강한 식생과 시들은 식생에서 나타나는 적색광(RED) 과 근적외선(Near Infrared; NIR) 영역의 반사도의 변화를 이용하여 NDVI 등이 개발되어 널리 사용되어 왔다(Tucker, 1979; Eq. 1-8).

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (\text{Eq. 1-8})$$

이 때, ρ 는 표면반사율이다. 공간 해상도는 $0.05 \times 0.05^\circ$ 이며 정방형 도법(Equirectangular projection or geographic projection)으로 투영되어 있고, 자료는 역시 미국 NASA의 EOSDIS에서 획득하였다.

표면반사율

위성 영상 기반의 가뭄 지수를 산정하기 위해 Terra 위성 및 Aqua 위성에 탑재된 MODIS 센서로부터의 표준 자료인 MCD43C4 Nadir BRDF-Adjusted Reflectance로부터 표면반사율을 도출하였다. Nadir는 위성에서 원격으로 지표를 관측할 때 비스듬하지 않은 정 아래쪽 방향을 일컫는 용어로, 본 자료는 Bidirectional Reflectance Distribution Function(BRDF)을 이용하여 원격 탐사의 관측이 Nadir 방향으로 이루어진 것처럼 모델링을 통해 표면반사율을 조정한 것이다. Terra 위성의 MODIS 센서 자료와 Aqua 위성의 자료가 함께 사용되었으며, 생산된 자료의 시간 해상도는 16일, 공간 해상도는 $0.05 \times 0.05^\circ$ 이다. 자료는 정방형 도법(Equirectangular projection or geographic projection)으로 투영되었고, 미국 NASA의 EOSDIS에서 획득하였다. MODIS의 36개 밴드 중 7개의 표면반사율이 제공되며 16일 자료는 월별 분석에 사용하기 위해 각 월에 해당하는 날 수를 가중치로 사용하여 월별 자료로 전환하였다.

1.3.2.4 장기예측 자료

사용한 장기예측 자료를 Table 1-4에 나타내었다. APEC기후센터(APCC)의 6개월 예측에 사용되는 여섯 가지 개별 모델인 MSC CanCM3(Meteorological Service of Canada Climate Centre third generation coupled global climate model), MSC CanCM4(Meteorological Service of Canada Climate Centre fourth generation coupled global climate model), POAMA(Predictive Ocean Atmosphere Model for Australia), NCEP CFSv2(National Centers for Environmental Prediction Coupled Forecast System version 2), PNU(Pusan National University), NASA(National Aeronautics and Space Administration)의 1-6개월 예측 선행시간의 강우 및 2 m 대기온도 예측자료를 사용하였다. 자료는 APCC Database로부터 획득하였다. 강우자료와 대기온도 자료 모두 시간해상도는 월별, 공간해상도는 2.5

× 2.5° 이고 단위는 각각 mm/day와 Kelvin(K)이다.

Table 1-4 Global Climate Model(GCM) data used in this study.

Product	Temporal Resolution	Spatial Resolution	Variable	Hindcast	Forecast
MSC_CanCM3	Monthly	2.5 × 2.5°	Precipitation, 2m Air Temperature	Jan. 1982 – Jan. 2011	Jan. 2013 - Aug. 2015
MSC_CanCM4				Feb. 1981 – Jan. 2011	Jan. 2013 - Aug. 2015
NASA				Dec. 1981 – Nov. 2012	Jan. 2013 - Aug. 2015
NCEP CFSv2				Mar. 1982 – Sep. 2009	Jan. 2013 - Aug. 2015
PNU				Mar. 1980 – Dec. 2012	Jan. 2013 - Aug. 2015
POAMA				Jan. 1983 – Dec. 2011	Jan. 2013 - Aug. 2015

1.4 선행연구와의 차별성

국내를 대상으로 한 가뭄예측 또는 가뭄전망 연구는 희소하지만 손수진(2012), 이진영(2015)의 연구 및 기상청의 종합가뭄정보시스템(2015)을 살펴볼 수 있다. 손수진(2012)은 CCSM3(Community Climate System Model), POAMA(Predictive Ocean Atmosphere Model for Australia), NCEP CFS(National Centers for Environmental Prediction Coupled Forecast System), PNU(Pusan National University), SNU(Seoul National University)의 다섯 개의 대기-해양 결합 전지구 기후 모델의 다중모델앙상블(Multi Model Ensemble; MME)를 이용하여 SPI와 표준강수증발산지수(Standardized Precipitation Evapotranspiration Index)를 산정하고, 남한의 60개 관측 지점의 6개월 선행예측기간의 5월 가뭄을 예측하였다. 개별모델의 역학적 모델 결과를 다시 통계적으로 다운스케일링 하였는데, 이 때 각 개별모델 별로 최적 예측 변수 선정을 통한 Coupled pattern projection method 방법을 활용하였다. 즉, 지점별로 대기온도와 강수에 대해 개별모델의 해수면압, 2 m 대기온도, 500 hPa Geopotential 고

도, 850 hPa 온도, 850 hPa 수평/연직 바람, 200 hPa 수평/연직 바람, 해수면온도 (Sea Surface Temperature, SST)의 9가지 예측 변수 중 가장 잘 맞는 것을 선정하여 이용하는 것이다. LEPS(Linear error in probability score) Skill Score로 평가하였을 때, 60개 지점 평균의 SPI와 표준강수증발산지수(Standardized Precipitation Evapotranspiration Index, SPEI)의 Skill Score의 경우 통계적 다운스케일링을 하였을 때가 각각 0.27, 0.25로 통계적 다운스케일링을 하지 않았을 때인 0.07, 0.05 보다 높게 나타났다.

기상청에서는 2015년 10월 현재 전지구 계절예측시스템인 Global Seasonal Forecasting System Version 5 (GloSea5) 모델을 이용한 60 × 60 km 해상도의 종합가뭄정보시스템을 구성하여 정부부처와 지방자치단체에 대하여 시범운영을 시작하였다. 이를 통해 1개월, 3개월, 6개월의 SPI와 파머가뭄지수(Palmer Drought Severity Index; PDSI)의 가뭄전망 값을 생산하여 제공한다.

이진영(2015)은 남북한 지역에 대하여 위성 자료로부터 얻은 강우, 지표면온도, 실제 및 잠재 증발산량, 유효광합성반사율(Fractional photosynthetically active radiation; FPAR), 엽면적지수(Leaf area index; LAI), 정규식생지수(Normalized Difference Vegetation Index; NDVI), EVI(Enhanced Vegetation Index), 토양 수분 자료와 대규모 대기순환 지수인 다변수ENSO지수(Multivariate El Nino Southern Oscillation Index; MEI)와 북극진동 지수(Arctic Oscillation Index; AOI), 그리고 지면모델 산출물에서 얻은 실제 및 잠재 증발산량과 토양 수분을 기계학습 모델의 한 종류인 Random Forest 모델의 입력 값으로 사용하여 SPI, SPEI, 그리고 정규수분지수(Normalized Difference Water Index; NDWI), 정규가뭄지수(Normalized Difference Drought Index, NDDI), 정규다중밴드가뭄지수(Normalized Multi-band Drought Index, NMDI)의 여러 지수를 예측하였다. 지면모델 산출물은 남한 지역에 대해서만 적용하였다.

본 연구에서는 기계학습 모델링 방법을 이용한 가뭄예측 모델 개발과 향상 가능한 범위 제시에 초점을 맞추고 있다. 따라서 역학적 기후모델의 결과의 개선 자체에는 중점을 크게 두지 않았으며, 개발된 가뭄예측 모델의 예측 성능을 비교하고자 하는 목적으로 사용하였다. 단일 모델이 아니라 앞서 언급한 APEC기후센터의 6개월 예측

에 사용되는 여섯 가지 개별 모델인 의 역학적 모의 결과를 이용하였다. 손수진 (2012)과 같이 개선된 다운스케일링 방법을 사용하지 않았으며 Quan et al. (2012)의 방법론대로 개별 모델의 아노말리를 실제 관측값의 기후값에 적용하여 대기 온도와 강수 값을 얻는 방법을 이용하였다.

이진영(2015)에 의해 개발된 모델과 비교하면, 기계 학습 모델의 종류를 Random Forest 단일 모델에서 Decision Tree, Random Forest, Extremely Randomized Trees 세 가지로 확장하여 적용하였다. Random Forest와 Extremely Randomized Trees는 black box 모델로 실제 예측되는 과정을 알기 어려우나 Decision Tree는 결정 규칙을 보이는 white box 모델이므로 비등한 결과를 생산하는 경우에는 유리한 면이 있다. 또한 기존 연구에서 제시한 입력 변수의 상대적 중요도에 따라 모델 입력 변수를 좀 더 줄인(parsimonious) 모델을 설계하였다. 또한 초기값과 목표값 사이의 기간에 대해 Lyon et al. (2012)에 제시된 기준값 생산 방법과 Quan et al. (2012)에서 제시된 역학적 모델 결과를 이용한 방법을 적용하였다. 위에 설명한 바와 같이 Quan et al. (2012)의 방법론을 적용할 경우 단일 모델이 아닌 여섯 가지 모델의 결과를 이용하였다. 또한 이진영(2015)는 모델의 입력 자료로 사용되는 강우 자료를 Tropical Rainfall Measuring Mission(TRMM)에서 얻은 값만을 사용하여 2012년까지의 자료만을 이용하였는데, 본 연구에서는 2014년 4월부터 자료를 생산하는 GPM(Global Precipitation Measurement) IMERG(Integrated Multi-satellite Retrievals for GPM) 자료를 이용하여 자료의 연속성을 보장하였다.

후속 연구를 통해 개선된 다운스케일링 방법론 등을 적용하면 본 연구를 통해 개발된 가뭄예측 모델을 개선시킬 수 있을 것이다. 또한 여러 다양한 방법론을 이용한 결과를 앙상블하여 의사결정에 활용하는 추세에 맞추어 본 연구를 통해 개발된 지역 규모의 활용을 위한 고해상도 가뭄예측 모델은 기상청의 종합가뭄정보시스템과 상호 보완적으로 활용하면 불확실성에 효과적으로 대응할 수 있을 것이다.

2. 가뭄의 정의 및 분류

2.1 가뭄의 정의와 가뭄 지표

2.1.1 가뭄의 정의

우리나라 가뭄 관련 법률 중 「재난 및 안전관리 기본법」과 「농어업재해대책법」은 각각 가뭄을 자연재난과 농업재해의 일종으로 포함시키고 있지만 법적으로 성립된 가뭄에 대한 정의는 아직 존재하지 않는다. 반면 미국의 경우에는 1998년의 「National Drought Policy Act」 및 2006년의 「National Integrated Drought Information System(NIDIS) Act」의 제정으로 가뭄에 대한 개별 법안의 입법화를 통한 가뭄 정책의 진일보를 이루었는데, 가뭄을 포괄적으로 “지표수나 지하수 공급(하천, 습지, 토양 수분, 저수지 공급, 호수 수위, 적설 포함)의 부족으로 이어져 심각한 경제적, 사회적 영향이나 개인과 재산 또는 환경에 심각한 물리적인 피해를 가져올 수 있는 강수 부족(NIDIS Act, 2006)”으로 정의하였다.

가뭄은 강수량의 부족으로부터 시작되지만, 지역에 따라 가뭄이 지속되거나 확대되는 양상과 피해를 가져오는 부문 및 범위가 모두 다르게 나타나므로 그 종류만큼이나 가뭄의 정의도 다르게 내릴 수 있다. Wilhite and Glantz(1985)은 가뭄의 정의를 일반적인 가뭄의 사전적 의미를 뜻하는 개념적 정의와 가뭄의 시작(onset), 심도(severity), 끝(termination) 및 빈도(frequency) 등을 포함하는 기술적(operational) 정의로 구별하여 서술하였으며, 150여 문헌의 가뭄 정의를 바탕으로 학문 분야에 따라 기상학적, 농업적, 수문학적, 사회경제학적 가뭄을 정의하였다. Mishra and Singh(2010)도 과거 여러 연구에서 나타난 가뭄의 개념 및 정의를 살펴보았다.

최근 기후변화에 관한 정부 간 협의체(Intergovernmental Panel on Climate Change; IPCC)에서 발간한 극한현상 및 재해의 위험관리 특별보고서(Special Report of the IPCC: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation; SREX)에서는 가뭄을 “심각한 수문학적 불균형을 야기할 정도로 장기간 동안 지속되는 비정상적으로 건조한 날씨”로 정의하였다. 여기서 가뭄의 원인이 되는 비정상적인 강수량 부족 기간을 기상학적 가뭄으로

정의할 수 있으며(Seneviratne et al., 2012). 강수량 부족 이외에 가뭄의 원인이 되는 일사량이나 풍속, 증기압차의 증가로 인한 증발산량의 증가가 토양수분, 호소 등의 기존 상태와 결합하여 주로 뿌리 층의 토양수분의 부족을 나타내는 농업적 가뭄(토양수분 가뭄)이나 하천, 호수, 지하수 등 수위가 비정상적으로 낮아지는 수문학적 가뭄이 발생할 수 있다(Seneviratne et al., 2012).

본 연구에서도 이와 같은 가뭄의 정의를 차용하였으며 기상학적 가뭄, 농업적 가뭄, 수문학적 가뭄 중 특히 기상학적 가뭄과 농업적 가뭄에 초점을 맞추었다. 수문학적 가뭄은 다루지 않았는데, 우리나라는 다목적 댐이나 용수공급댐, 광역상수도과 같은 수리시설이 잘 되어있어 강수량 부족에 기인한 물 공급과 인간의 물 사용과 관련된 물 수요를 모두 고려하였을 때 물 수요량이 물 공급량을 충족시키지 못하는 상태인 물부족과 가뭄을 개별적으로 분리하기 쉽지 않기 때문이다.

2.1.2 가뭄 지표

수없이 많은 가뭄 지수가 개발되어 각각의 장단점을 가지고 사용되고 있는데(Mishra and Singh, 2010), 본 연구에서는 기상학적 가뭄 지수 중 SPI와 SPEI, 그리고 위성 자료에서 도출할 수 있는 NDWI를 이용하였다. SPI는 World Meteorological Organization(WMO)에 의해 기상학적 가뭄을 감시하기에 가장 적합한 지수로 선정되었다(Hayes et al., 2011). SPEI는 SPI와 유사한 과정을 통해 산정될 수 있는데, SPI가 강수량만을 사용하는 데 비하여 강수량과 증발산량을 동시에 고려하여 물가용량을 좀 더 현실적으로 반영하고 있다. 위성 자료에서 도출할 수 있는 NDWI는 식생이 함유하고 있는 수분량을 추정하므로 토양수분 가뭄의 상태를 판단할 수 있는 대리변수(proxy)로서 사용되었다.

2.1.2.1 표준강수지수

표준강수지수는 원하는 기간(timescale)에 따라 강수량을 누적시키고, 이를 정규화하여 산정한다. 강수량만을 사용하므로 많은 지역에 적용 가능하며 정규화 과정을 포함하므로 시공간적인 비교가 용이하다. 원하는 기간에 대한 누적 강수량을 월별 등

으로 정규화하므로 특정 월 등의 강수량을 평년과 비교할 수 있다. 사용자가 누적 기간을 지정할 수 있으므로 가뭄의 진행 정도에 따라 다양하게 사용될 수 있다. 그러나 강수량만을 사용하고 대기의 수분 수요를 고려하지 않기 때문에 주의를 기울여 사용할 필요가 있다.

정규화 과정을 좀 더 자세히 설명하면 다음과 같다. 일반적으로 12개월 미만의 강수량 누적량은 정규분포를 따르지 않으므로 적절한 확률분포함수에 맞추게 된다. 누적 강수량에 적합한 확률분포함수로 McKee et al.(1993)은 매개변수가 2개인 Gamma 확률분포함수를 제안하였으며 Guttman(1999)은 여러 확률분포함수를 비교한 뒤 매개변수 3개의 Pearson Type III 확률분포함수를 가장 적합한 함수로 제안하였다.

$$f(x) = \frac{|x - \xi|^{\alpha-1} \exp(-|x - \xi|/\beta)}{\beta^\alpha \Gamma(\alpha)} \quad (\text{Eq. 2-1})$$

이 때, x 는 Pearson Type III 분포를 따르는 확률 변수이고, $f(x)$ 는 Pearson Type III 확률분포함수이다. α , β , ξ 는 이 확률분포함수의 매개변수로 L-모멘트인 평균 μ (location 변수), 표준편차 σ (scale 변수), 왜도 γ (shape 변수)와 다음의 관계에 있다.

$$\alpha = \frac{4}{\gamma^2}, \beta = \frac{1}{2}\sigma|\gamma|, \xi = \mu - 2\frac{\sigma}{\gamma}, \gamma \neq 0 \quad (\text{Eq. 2-2})$$

월별로 분석을 수행할 경우 기간에 따른 누적 강수량의 시계열을 구한 다음, 월별로 L-모멘트 값을 계산하고 이 값들을 이용하여 선택된 확률분포함수의 매개변수 값을 얻는다. 이에 따라 각 누적 강수량의 누적 확률(cumulative probability)을 얻을 수 있으며, 정규분포함수의 역함수를 사용하여 SPI 값을 산정하게 된다.

$$F = G\left(\frac{x - \mu + 2\frac{\sigma}{\gamma}}{\left|\frac{1}{2}\sigma\gamma\right|}, \frac{4}{\gamma^2}\right), \gamma > 0 \quad (\text{Eq. 2-3})$$

$$F = 1 - G\left(\frac{-\left(x - \mu + 2\frac{\sigma}{\gamma}\right)}{\left|\frac{1}{2}\sigma\gamma\right|}, \frac{4}{\gamma^2}\right), \gamma < 0 \quad (\text{Eq. 2-4})$$

$$G(x, \alpha) = \{\Gamma(\alpha)\}^{-1} \int_0^x t^{\alpha-1} e^{-t} dt \quad (\text{Eq. 2-5})$$

이 때, F는 Pearson Type III 누적분포함수이며, G는 불완전 Gamma 함수이다. 산정된 표준강수지수의 평균은 0, 표준편차는 1이 되는데, McKee et al.(1993)의 체계에 따라 가뭄의 심도를 분류할 수 있다(Table 1-1).

2.1.2.2 표준강수증발산지수

기상청이 생산한 1 × 1km 해상도의 남한상세 기후변화 시나리오를 이용하여 미래 기후변화를 전망한 국립기상연구소의 「기후변화 시나리오 보고서 2011」(기상청, 2011)에 의하면, 동아시아 지역에는 강수량의 증가가 예상된다. 그러나 강수 강도가 증가하고 연강수일수가 감소함에 따라 집중적인 우기가 존재하게 되고, 따라서 강도 높은 우기 사이의 건기가 길어지며 기온 증가로 인해 증발량이 증가할 것으로 전망하고 있다(기상청, 2011). 가뭄의 일차적인 원인은 강수량의 부족이나, 수문기상변수들의 상호 작용으로 인해 다른 요인의 영향을 많이 받으므로, 대기의 수분 수요를 고려하지 않고 강수량만을 고려하면 미래 가뭄 발생을 잘못 전망할 수 있다. 따라서 기후변화의 영향을 고려하여 우리나라의 가뭄 발생을 전망하려면 강수량 뿐 아니라 증발산량과 같은 변수를 동시에 고려하여야 한다.

강수량과 증발산량을 동시에 고려하여 Vicente-Serrano et al.(2010)은 SPEI를 개발하였으며, Tsakiris et al.(2007)는 Reconnaissance Drought Index(RDI), Ellis et al.(2010)는 Hydroclimatic Index(HI)를 개발하였다. RDI는 강수량과 증발산량의 비가 log-normal 분포를 따른다고 가정하여 이를 이용하였으며, HI는 강수량과 증발산량의 차를 누적하여 계산한다. SPEI는 강수량과 증발산량의 차의 누적량을 log-logistic 분포에 맞춘 후 정규화 과정을 거친 뒤 산정한다. SPEI는 개념적으로 강수량과 증발산량의 차가 유출 및 지하수 및 토양 수분으로의 수분 저장의 합과 같아지는 물수지를 고려하며 시공간적인 비교를 가능하게 하는 정규화의 장점을 가지고 있다. SPEI는 SPI의 분류 체계를 따른다(Table 1-2). Log-normal 분포의 확률밀도함수 및 누적분포함수는 아래 식과 같다.

$$f(x) = \frac{\beta}{\alpha} \left(\frac{x-\gamma}{\alpha} \right)^{\beta-1} \left(1 + \left(\frac{x-\gamma}{\alpha} \right)^{\beta} \right)^{-2} \quad (\text{Eq. 2-6})$$

$$F(x) = \left[1 + \left(\frac{\alpha}{x-\gamma} \right)^{\beta} \right]^{-1} \quad (\text{Eq. 2-7})$$

이 때, x 는 log-logistic 분포를 따르는 확률변수, $f(x)$ 는 log-logistic 확률밀도함수이며, $F(x)$ 는 log-logistic 누적분포함수이다. 세 매개변수 α , β , γ 는 각각 scale 변수, shape 변수, origin 변수이다.

2.1.2.3 정규수분지수

산림 및 농작물 등 식생에 대한 가뭄의 영향을 정량화하기 위한 대리변수로 위성 자료에서 얻은 표면반사율에서 산정한 정규수분지수를 이용할 수 있다. 이진영 (2014)은 NDVI, NDWI, NMDI, NDDI를 이용하였는데, 본 연구에서는 이러한 지수 들 중 전통적인 가뭄지수와 상관관계가 높고(Rhee et al., 2010) 식생의 수분 함량을 잘 나타내주는 NDWI를 이용하였다.

NDWI는 Gao(1996)에 의해 개발되었으며 근적외선 영역에 속하는 $0.86 \mu\text{m}$ 와 단 파적외선 영역에 속하는 $1.24 \mu\text{m}$ 의 두 채널의 표면반사율을 이용하여 다음 식과 같이 산정된다.

$$NDWI = \frac{\rho(0.86\mu\text{m}) - \rho(1.24\mu\text{m})}{\rho(0.86\mu\text{m}) + \rho(1.24\mu\text{m})} \quad (\text{Eq. 2-8})$$

이 두 채널은 서로 비슷한 식생에 의한 산란 수준을 가지고 있다. Moderate Resolution Imaging Spectroradiometer(MODIS) 센서는 네 개의 단파적외선 영역에 속하는 밴드를 가지고 있는데(밴드5 = $1.24 \mu\text{m}$, 밴드26 = $1.38 \mu\text{m}$, 밴드6 = $1.62 \mu\text{m}$, 밴드7 = $2.1 \mu\text{m}$), 이 중 밴드5, 6, 7의 표면반사율을 사용하기 용이한 표준 자료로 제공한다. NDWI에 사용되는 $1.24 \mu\text{m}$ 채널은 이 보다 긴 파장의 채널에 비해 식생에 존재하는 수분 함량에 민감하다고 알려져 있는데, 이는 $0.9-1.3 \mu\text{m}$ 영역에 비해 $1.5-2.5 \mu\text{m}$ 영역에서 식생 내 물에 의한 에너지 흡수가 훨씬 강하기 때문이다 (Gao, 1996). 또한 NDWI는 NDVI와 비교하여 대기효과에 덜 민감하며, 식생이 우거

진 지역에 대해 포화되는 정도가 적고 층층이 쌓인 잎사귀의 총괄적인 수분 함량을 측정하므로 전통적으로 식생의 활력도 감시를 통한 가뭄 연구에 많이 사용되어 온 NDVI와 함께 가뭄의 평가에 유용하게 사용될 수 있다.

NDWI에 사용하는 단파적외선 채널로 Gao(1996)는 1.24 μm 를 제안하였다. 다른 단파적외선 채널도 다양한 연구에 사용되었는데, Gu et al. (2007)와 Gu et al. (2008)은 MODIS의 밴드7에 해당하는 2.1 μm 를 이용하였다. Huang et al. (2009)은 식생의 수분 함량 추정에 있어 1.24 μm 를 이용한 NDWI(NDWI_1240nm; MODIS 이용), 1.64 μm 를 이용한 NDWI(NDWI_1640nm; Landsat Thematic Mapper/Enhanced Thematic Mapper+이용), 2.13 μm 를 이용한 NDWI(NDWI_2130nm; Landsat Thematic Mapper/Enhanced Thematic Mapper+ 이용)를 비교하였으며, 결과적으로 옥수수의 경우 1.64 μm , 콩의 경우에는 1.24 μm 채널을 이용한 경우가 식생의 수분 함량을 가장 잘 추정하였다. 본 연구에서는 Gao(1996)의 제안에 따라 근적외선 채널과 유사한 대기 투과율을 지닌 1.24 μm 채널을 이용하였다.

NDWI를 활용한 가뭄 연구가 다수 존재하는데, 국외에서 이루어진 연구를 보면 Anderson et al. (2010)은 NDWI를 아마존 유역의 2005년 가뭄과 관련하여 수목 고사율과 비교한 결과 유효한 부의 상관관계가 있음을 보였다. Chakraborty and Sehgal (2010)은 인도 Rajasthan 주의 농업가뭄을 평가하기 위해 가뭄해인 2002년 식물 성장기와 비가뭄해인 2003년의 NDVI와 NDWI를 강수량과 비교하였는데, NDWI는 가뭄의 강도 및 범위를 잘 감지하고 강수량과 높은 상관관계를 보인 반면 NDVI는 강수량과 지연효과를 보였다. NDVI가 강수와 지연효과를 보인 반면 NDWI가 같은 시기 강수와 높은 상관관계를 보인 점은 Gu et al. (2007)이 NDWI가 NDVI보다 가뭄에 더 빨리 반응한다고 밝힌 점과 일맥상통한다.

Gu et al. (2008)은 2002년에서 2006년 식물 성장기의 자료를 이용하여 NDVI와 NDWI 값과 토양수분을 비교하였는데, 비교적 높은 상관관계를 가짐을 보일 수 있었다. 특히 상대적으로 균일한 식생으로 이루어진 지역에 토양이 미사질양토(silt loam)인 경우 토양 수분과 가장 높은 상관관계를 보였다. 이 연구에서 흥미로운 점은 식생지수가 5-cm 층이나 60-cm 층보다 25-cm 층의 토양수분과 가장 큰 상관관계를 보인 것인데(Gu et al., 2008) 이는 위성에서 도출한 식생 지수가 중간 토양층의

수분에 크게 반응함을 보이는 것으로 가시광이나 적외선 영역의 파장이 지표의 가장 표층만을 탐지한다는 단점을 어느 정도 보완해 줄 수 있다. 시간적으로는 대략 1주에서 2주 정도의 지연이 나타났다.

국내에서는 이가람 외(2008)가 NDWI를 이용하여 1999년부터 2006년까지의 SPOT/Vegetation 위성 영상을 이용하여 중국과 몽골 사막주변의 식생수분상태를 모니터링하는 연구를 수행하였으며, 성노훈 외(2015)는 1999년에서 2013년 동안의 SPOT/Vegetation 영상 자료를 이용, 한반도 지역의 산림 캐노피의 수분 스트레스를 평가하였다. 성노훈 외(2015)는 또한 한반도의 대형 산불 발생과 NDWI 아노말리를 비교 검증하여 NDWI 아노말리가 클수록 대형 산불이 빈번히 발생하였음을 밝혔다.

2.2 가뭄 발생의 원인

가뭄이 발생하면 그 피해를 경감하기 위해 대비, 대응 및 복구에 집중하게 되는데 이와 더불어 가뭄 발생의 원인을 분석하여 가뭄 예측에 활용하고자 하는 노력이 있어 왔다(기상연구소, 1993). 지금까지도 가뭄 발생의 메커니즘이 확실히 밝혀졌다고 할 수는 없으나, 지구의 대규모 기후 시스템과 연관된 여러 요인의 상호작용이 가뭄을 유발한다고 할 수 있다. 이러한 요인들은 엘니노/남방진동(El Nino and Southern Oscillation: ENSO), 해수면 온도의 아노말리(Sea Surface Temperature Anomaly: SSTa), 대륙과 대기의 상호작용 등을 포함한다(기상연구소, 1993).

2.2.1 국외 가뭄 발생의 원인

국외에서는 가뭄 발생의 원인을 분석하려는 노력이 비교적 오래 전부터 있어 왔는데, 초기의 연구들을 살펴보면 예를 들어 Namias(1955)는 미국의 여름철 가뭄의 요인으로 준정상 행성파의 앙상블(quasi-stationary planetary wave ensemble)을 지목하여 미국 상층의 고기압 발달을 유발하거나 억제하는 대서양 및 태평양 고기압의 중요성을 제시하였다. Namias(1991), Janowiak(1988), Ropelewski(1988)는 미국의 1988년 여름철 가뭄을 분석하였는데, Namias(1991)는 이 가뭄이 여름 초반 북태평양, 북대서양, 미국 대륙의 상층에 세 개의 전형적인 고기압 아노말리가 형성되어 발

생하였으며, 4월과 5월 동안 건조된 토양과의 상호작용으로 고기압이 강화되고 이러한 세 지역 간의 원격상관이 지속되었다고 설명하였다.

미국 이외의 다른 지역에 대해서도 가뭄의 원인을 분석하였는데, Namias(1972)는 브라질 북동부 강수량의 경년변화를 북반구 겨울과 봄 동안의 뉴펀들랜드와 그린란드 지역의 저기압 활동의 정도 또는 블로킹에 의해 결정된다고 설명하였다. 즉, 북미와 북대서양의 강력한 블로킹이 북동부 브라질의 극심한 가뭄과 연결되어 있다고 보았다. Kanamitsu and Krishnamurti(1978)는 중앙 인도와 서부 아프리카의 1972년 가뭄이 다음과 같은 요인과 연관이 있다고 설명하였다. 적도 태평양의 높은 해수면 온도, 서태평양의 과도한 태풍 일수, 태풍으로부터의 상층 유출과 관련된 적도 동인도양의 강한 동북동풍, 약한 열대 편동풍 제트, 인도 상부의 약한 남북 기압차, 약한 티벳 고기압, 주요 순환 패턴과 여러 역학적 매개변수의 남동쪽 편향(southeastward shift), 서아프리카 상부의 약한 연직바람시어(vertical wind shear)와 약한 순압-경압불안정, 인도 상부의 강우생성요란(rain-producing disturbance)의 약한 서쪽 지향(westward steering) 등이다.

2.2.2 국내 가뭄 발생의 원인

2.2.2.1 대규모 대기/해양 순환

국내의 경우 가뭄 발생원인 규명에 도움이 되는 가뭄 관련 기상 변수들에 대한 연구가 20세기 후반부터 활발히 수행되었다. ENSO와 기상 변수와의 상관성 연구가 주를 이루었는데, 하경자(1995)는 1960년부터 1992년까지의 자료에서 얻은 서울의 겨울철 평균 기온의 경년 변동이 거의 같은 기간에 Nino 1+2 지역과 높은 양의 상관성을 보이는 것을 밝혔다. 안중배 외(1997)는 국내 5개 도시의 겨울철, 여름철 기온, 강수량의 아노말리와 적도 태평양의 계절별 SST 아노말리와의 관계를 살펴보았는데, 겨울과 여름의 강수량의 경우 각각 Nino-3 지역과 WP(Western Pacific)-2 지역의 6개월 전의 적도 태평양 SST의 영향을 받으며 높은 음의 상관성을 보였고, 겨울과 여름 기온의 아노말리는 각각 6개월 전 WP-2 지역의 SST 아노말리, 3개월 전 WP-1 지역의 SST 아노말리와 높은 양의 상관성을 보였다.

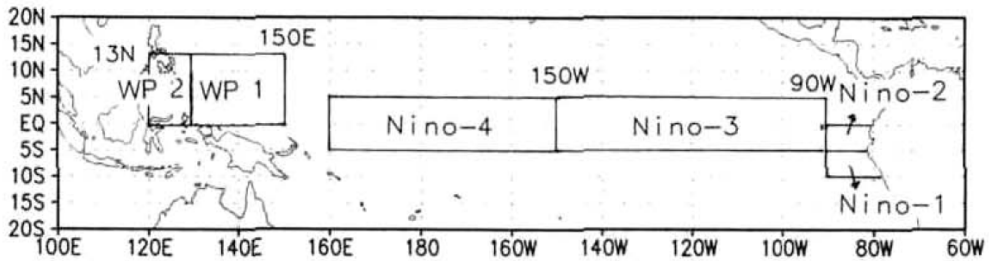


Figure 2-1 Locations of Niño-1, 2, 3, 4 and WP-1, 2 (Source: An et al., 1997).

강인식(1998)은 1935년부터 1996년의 61년간의 자료를 통해 ENSO 지수인 SOI (Southern Oscillation Index)와 한반도 기온 및 강수량의 경년 변동과의 상관관계를 통계적으로 분석하였는데, 기온의 경우 엘니뇨 해에는 겨울철 기온이 높고 여름철 기온이 낮게 나타났으며 엘니뇨 여름철에 강수량이 증가하는 것을 보였다. 분광 분석 결과 한반도 기온 변동 중 6년 주기의 변동이 엘니뇨 지수와 약 1년에서 1.5년의 시간 차이를 두고 연관되어 있으며, 강수량의 경우 5-7년 및 3년 주기의 변동이 엘니뇨와 연관되어 있음을 밝혔다. 차은정 외(1999)는 엘니뇨가 한반도 및 중위도 기후변화에 큰 영향을 미친다는 결론을 제시하였으며 기온의 경우 엘니뇨 해 여름철 저온 및 겨울철 고온과 라니냐 해의 여름철 고온 및 겨울철 저온을 밝혔다. 강수량의 경우 엘니뇨 해에는 북태평양 고기압의 발달이 평년보다 늦어지고 위치가 평년보다 남쪽/남동쪽으로 편향하기 때문에 장마 시작 시기와 종료 시기가 늦어지고 강수량이 증가하며 라니냐 해에는 장마 종료가 짧아 기간이 짧고 강수량이 감소하는 경향이 있음을 지적하였다.

가뭄지수를 이용하여 가뭄과 ENSO와의 관계를 규명하는 연구도 다수 수행되었다. 이동률(1999)은 PDSI, 강수량, 하천유량 자료와 9개 ENSO 지수와의 교차상관을 분석하였다. 분석 결과 PDSI 총 변량의 6% 정도만 ENSO에 의해 설명되어 가뭄을 ENSO 관련 지수만으로 예측하는 것은 어려운 것으로 밝혀졌으나, 우리나라의 가뭄은 엘니뇨가 끝나는 해의 12월부터 6개월 동안 평균 이하의 PDSI가 지속되어 가뭄 발생이 ENSO 지수에 13-19개월 지체되어 통계적으로 유의하게 나타나는 것을 보였다. 여러 지수 중 SOI, Niño 4와 Ship Track 6 지역의 SST가 한국의 가뭄과 가장 큰 연관성을 가지고 있는 것으로 나타났다. PDSI와 강수량, 하천유량의 합성백분위

수를 분석한 경우에는 PDSI와 유사하게 16개월 전후에 ENSO 발생과 유의하게 나타났습니다.

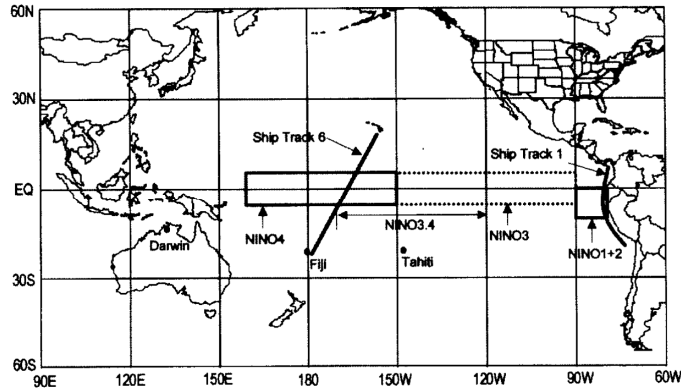


Figure 2-2 Locations of Ship Track 1, 6, Niño 1+2, 3, 4, and 3.4 [Source: Lee, 1999].

변희룡 외(2001)는 유효강수량지수(Effective Drought Index, EDI)와 Niño-3 지역의 SST 아노말리를 분석한 결과 엘니뇨/라니냐의 강도가 증가하는 동안에 아시아의 어떤 지점에서 발생하는 가뭄의 강도도 증가하고 같은 시기 또는 1개월 뒤에 가뭄이 극점에 도달하지만 그 이후 급격히 약화된다는 점을 밝혔다. 가뭄의 발생 지역은 인도와 만주를 잇는 축 상에 존재하며 이 축은 엘니뇨가 강화할 때는 북상하고 약화할 때는 남하하는 경향이 있으며, 가뭄의 강도는 엘니뇨/라니냐와 상관관계를 가지지 않는다.

ENSO 이외의 다른 원격상관과 가뭄의 관계를 분석한 연구도 존재하는데, 최기선 외(2009)는 1974년부터 2008년까지의 35년 동안 국내 가뭄지수는 EDI로 산정하고 북태평양진동(North Pacific Oscillation, NPO)은 NPI(North Pacific Index)를 이용하여 분석한 결과 두 지수 사이에 높은 음의 관계가 있음을 밝혔다. NPI는 베링해 부근의 한 지점(65°N, 170°E)과 아열대 중앙태평양의 한 지점(25°N, 165°E) 사이의 해면기압 차를 정규화하여 계산한 것이므로 음의 관계는 베링해 부근 지역에서의 해면 기압이 높을수록 국내의 가뭄은 강화되는 것을 의미한다. 이종화 외(2003)는 우리나라 봄 강수량이 북대서양진동(North Atlantic Oscillation, NAO) 패턴과 형태가 유사한 500hPa의 북대서양 고도 편차와 부의 상관관계가 있고 북대서양 SST 아노말리와의 부의 상관관계가 있다는 점으로부터 겨울철 북대서양의 대기 및 해양 패턴이

국내 봄 강수량에 영향을 주며 NAO 지수가 국내 봄 가뭄과 관계가 있을 수 있다는 점을 도출하였다. 서길수(2015)는 한반도의 봄 가뭄과 여러 원격상관과의 상관관계를 분석하였는데 이 중 NAO의 경우 봄철 강수량(3-5월)과 3개월의 지연을 가진 12-2월의 NAO 이상치와의 음의 상관성이 가장 큰 것으로 나타났다.

김성(2004)은 한반도 봄철 강수량과 연관된 북반구 대기순환 패턴을 조사하기 위해 500 hPa 고도장에 대한 REOF(Rotated Empirical Orthogonal Function) 분석을 수행하였는데, 첫 번째 모드는 SOI와 강한 음의 상관관계를 보이는 ENSO 모드로 나타났으며 두 번째 모드는 AOI와 강한 음의 상관관계를 보이는 AO(Arctic Oscillation) 모드로 나타났다. 세 번째와 네 번째 모드는 한반도 봄철 강수량과 주변 지역의 종관 패턴을 나타내는 것으로 1984년 이전에는 유라시아 대륙의 시그널과 연관된 세 번째 모드가, 1985년 이후에는 북태평양에서 변동하는 시그널과 연관된 네 번째 모드가 상관성이 높은 것으로 나타났다. Guan and Yamagata(2003)는 우리나라를 포함한 동아시아 지역에 발생했던 1994년 여름 가뭄을 분석하며 그 원인의 하나를 IOD(Indian Ocean Dipole)로 보았는데, 이 특정 해 외에도 1961년과 같이 전형적인 IOD 해 동안에 동아시아 지역에 유사한 여름철 기후가 나타났음을 언급하였다.

2.2.2.2 종관규모 대기순환

앞서 언급한 김성(2004)의 연구에서 500 hPa 고도장에 대한 REOF의 첫 번째 및 두 번째 모드는 ENSO, AO와 상관관계를 가지는 것으로 나타난 반면 세 번째 및 네 번째 모드는 그 밖의 한반도 봄철 강수량과 연관된 시그널을 가지는 것을 볼 수 있었다. 이렇듯 국내의 가뭄 발생은 주변 지역의 종관 규모의 대기순환을 통해 설명할 수 있다. 변희룡(1996)은 여름철의 각 월별로 가뭄의 원인을 달리 규명하였는데, 6, 7월의 가뭄은 북서태평양 일부의 아열대 기단이 한반도까지 북상하지 못하여 발생하는 것으로 분석하였다. 북상하지 못하는 원인은 북서태평양지역의 고기압이 덜 발달되는 것과 6월 만주지방 및 7월 오토츠크해 주변의 기압이 평년보다 높은 것을 들었다. 8월의 가뭄은 한반도보다 고위도에서 발달하는 동쪽의 키 큰 고기압에 의한 것으로, 동서류가 약해지고 기압계 전체의 이동이 남북성분이 강해지며 가뭄 발생의 가능성이 증가한다고 보았다. 그 밖에 아열대 고압대의 북상이 저지되거나, 한반도 주변의

저기압의 이동경로가 남쪽으로 치우치거나 제트류의 평균적 중심이 한반도보다 남쪽에 위치하는 경우에도 가뭄이 발생한다고 밝혔다.

박중서, 전종갑(2003)은 우리나라의 계절 가뭄 중 가장 심각한 봄 가뭄에 대하여 1960년에서 2001년의 여러 가뭄을 21개 강우관측소 자료로부터 PDSI를 산정하여 도출하여 대기 순환장과의 상관관계를 구하였다. 우리나라와 동아시아 지역의 가뭄은 높은 상관관계를 보였으며 순환장 분석을 통해 동아시아 지역의 봄 가뭄은 상층 대기 흐름의 northerly에 의한 찬 공기의 남하로 발생한 것임을 밝혔다. 그 원인은 가뭄 발생 기간 동안 대기 순환장이 수직적으로는 barotropic한 구조를 보이고 수평적으로는 아래로는 알류시안 저기압이 우리나라와 중국 남동해안까지 강화되고 위로는 고기압성 회전이 있는 dipole 구조가 나타났기 때문으로 보았다. 이 밖에 최기선 외(2003), 김기훈 외(2003)는 각각 온도장과 해빙을 이용하여 한국의 5월 가뭄을 예측할 수 있는지 분석하였는데, 각각 EDI를 이용하여 가뭄 해를 도출하고 온도장과 해빙을 이용하여 상관관계가 높은 지역과 시기를 도출해 내었다. 전갑영 외(2008) 역시 통계적 방법을 이용하여 한반도의 가뭄과 주변 기상장과의 상관관계를 분석하였는데, 한반도의 가뭄은 여름에 시작하는 경향이 있으며 심한 가뭄은 그해 겨울 또는 다음 해 여름 우기 전까지 유지되는 것을 밝혔으며 5월부터 8월까지의 가뭄 발생과 상관관계가 높은 주변 지역을 도출하였다.

2.2.3 국내 과거 가뭄 발생

관측 자료가 존재하는 시기에 발생한 국내 과거 가뭄은 일반적으로 평년에 비해 부족한 강수량, 저수율, 하천 유량 등으로 평가할 수 있다. 국립방재연구소(1998)는 전국 연평균 강수량을 기준으로 1900년 이후의 가뭄 기간을 1927-1929년, 1937-1939년, 1942-1944년, 1967-1968년, 1976-1977년, 1982-1983년, 1988년 그리고 1994-1995년으로 선정하였다. 이 중 1939, 1968, 1978, 1982년의 가뭄을 우리나라 4대 한발로 꼽았다(국립방재연구소, 1998).

이대희 외(2002)는 1990년대 이전은 국회 도서관의 신문 기사 스크랩을 이용하여 앞서 언급된 대가뭄으로 인식되는 4대 한발을 조사하고 1990년대 이후로는 인터넷 검색을 통하여 가뭄과 관련된 모든 기사를 검색하여 정리하여 과거 주요 가뭄년을 분

류하고 피해를 조사하였다. 이를 통해 과거의 주요 가뭄년을 1939년 낙동강 유역의 가뭄, 1967-1968년 영산강, 낙동강 유역의 가뭄, 1973년 가뭄, 1976-1978년 영산강, 낙동강 유역의 가뭄, 1981-1982년 낙동강 유역의 가뭄, 1994-1995년 충청도 이남 지역의 가뭄으로 선정하였다(이대희 외, 2002). 후에 한국수자원공사(2002)는 이 목록에 2001년 한강 유역의 가뭄을 추가하였다.

권원태, 오재호(2011)는 근대 기상관측을 시작한 1907년 이래로 서울의 봄철 강수량을 살펴보았는데, 평년 225 mm인 서울의 봄철 강수량이 100 mm 이하인 해로 1910, 1917, 1950, 1965, 1978, 1984년의 6회를 선정하였다. 이 중 1965년의 봄철 강수량이 38.8 mm에 그쳤으며 그 다음으로 2001년 봄철 강수량이 46.9 mm이었다(권원태, 오재호, 2011). 서울의 연강수량의 경우에는 평년의 60% 이하인 800 mm 이하인 해는 1939, 1943, 1949, 1988년이 있었으며 이 중 1939년의 연 강수량이 638 mm로 가장 적었다(권원태, 오재호, 2011).

장기호 외(2011)는 소방방재청(2009)의 「2008-2009년 가뭄극복추진 성과보고서」의 강수량 자료를 바탕으로 1971년 이후 우리나라에 극심한 가뭄 피해가 발생한 연도를 1977, 1978, 1980, 1982, 1988, 1994, 1997, 2000, 2001, 2006, 2008-2009년으로 선정하였다. 또한 2001년에서 2010년 사이의 제한급수 상황을 살펴보았는데, 2003년을 제외하고는 매년 강수부족으로 인한 제한급수가 지역적으로 발생하였다(장기호 외, 2011). 2000년에는 2-5월의 강수량이 10년 평균의 40% 정도로 전국적으로 심한 봄 가뭄이 발생하였고(장기호 외, 2011; 기상청, 2000), 2001년에는 5-6월 경기, 강원, 충북, 경북 지역에서, 2006년에는 10월 충남, 전남북, 경남 지역에서, 2008년 9월-2009년 2월 전남, 경남, 강원 지역에 가뭄이 발생하였다(장기호 외, 2011).

또한 심현준 외(2013)는 자연용수공급 지역에서 강수량으로부터 산정한 표준강수지수와 용수수요량을 이용하여 용수부족량을 타나내는 RDI(Real-time Drought Index) 이용을 제안하였는데, RDI 산정을 위해 특성 매개변수를 도입하였다. 특성 매개변수 추정을 위해 여러 가뭄관련 국가 보고서 및 보도 자료를 이용하여 제한 급수 지역을 도출하였는데, 2000년대 이전에는 전국적으로는 1988, 1997년과 지역적으로는 1982, 1992, 1993, 1994, 1995년에 가뭄이 발생하였으며 2000년대 이후에는

전국적으로는 2001, 2008년, 지역적으로는 2002, 2004, 2005, 2006년에 가뭄이 발생하였다고 설명하였다. 국내에 발생하였던 최근 가뭄의 원인과 발달 등을 각 발생 가뭄 별로 문헌조사를 통해 아래와 같이 상세히 살펴보았다.

2.2.3.1 1994-1995년 가뭄

1994년에는 장마 시작 시기는 평년과 비슷하였으나 북태평양 고기압의 이상발달로 장마가 조기 종료되고 강력한 북태평양 고기압이 장기간 우리나라 상공에 정체되어 이상고온이 지속되어 심각한 가뭄이 전국적으로 확산되었다(송병현 외, 1995; Park and Shubert, 1997; 최영진, 1995). 1994년의 여름 강수량은 평년의 30-70% 수준에 그쳤으며 특히 경북내륙과 전남내륙, 남부 해안지방의 강수량이 매우 부족했던 것으로 알려졌으며 장마 기간 중 강우일수도 적었다(송병현 외, 1995). 5개 하천의 유량의 경우에도 한강을 제외하고는 평년의 20-60% 수준으로 매우 적었으며 특히 낙동강의 경우 매우 심각한 가뭄이 이어졌다(김승 외, 1995).

2.2.3.2 2001년 가뭄

1994-1995년의 가뭄이 장마의 조기종료로부터 시작된 것과는 다르게 2001년에는 3월에서 6월 중순까지 평년의 10-30%에 불과한 단기간의 강수량의 부족으로 심한 가뭄을 겪었다(김휘린 외, 2004; 어민선, 2001; 이동률 외, 2002). 2001년 봄철 가뭄은 중국 북부 내륙으로부터 우리나라에 이르는 지역에 이례적으로 발달한 상층 고기압과 중위도 지역의 동서 평균 성분의 상층 고압대의 형성으로 인한 우리나라 북쪽을 지나는 기압골의 세력 약화, 그리고 다소 남쪽으로 치우쳐 발달한 서태평양의 아열대 고기압의 영향으로 인한 수증기 유입의 현저한 감소 등으로 인해 발생한 것으로 보인다(이종화 외, 2001).

그러나 이 시기의 가뭄심도를 PDSI의 가뭄지수를 이용하여 살펴보면 과거 1968년, 1977년, 1982년, 1988년, 1994-1995년의 가뭄과 비교하여 상대적으로 미미한데, 이는 PDSI의 시간 스케일이 대략 9-12개월 정도로 3개월의 강수량부족으로 나타난 2001년의 가뭄이 잘 표현되지 않은 것으로 분석된다(이동률 외, 2002; 이동률,

김웅태, 2001). SPI를 이용하여 살펴본 경우 2001년의 가뭄이 1900년도 이후 가장 큰 가뭄으로 나타났다(이동률 외, 2002). 특히 경기, 강원, 충북, 경기북부 등 중부지방과 서남부 해안지역에 심한 가뭄이 발생하였다(오국렬, 유인상, 2014).

2.2.3.3 2008-2009년 가뭄

이 시기에는 2008년 여름과 가을의 강수량 부족이 시작되어 강원도 내륙지방, 전라남도, 경상남도 등 남부지방에 가뭄이 발생하였다(문장원, 이동률, 2009). 유출량을 2007년과 비교하면 2008년 여름철 유출량은 80% 수준이나 가을철 유출량은 25% 정도로 매우 적었으며 특히 낙동강은 20% 미만, 섬진강과 영산강은 10% 미만으로 나타났다(황석환, 이재형, 2009).

2.2.3.4 2012년 가뭄

2012년의 가뭄은 5월과 6월에 가장 심각하게 발생한 단기간 강수량 부족으로 인한 가뭄으로 시작하였으며, 서울, 인천, 경기, 대전, 충남 지역의 강우량이 가장 부족하였다(오국렬, 유인상, 2014). PDSI로 분석하였을 때 강화지역이 극한 가뭄, 만경 및 동진강 유역이 심한 가뭄을 보였으며, SPI의 경우 서해안 지역, 만경 및 동진강 유역이 심한 가뭄을 나타냈다(안정훈 외, 2012). 짧은 가뭄에 이어 장마 시기가 되었는데 6월 하순 평년보다 강한 오호츠크해 고기압의 영향으로 북태평양 고기압이 우리나라 까지 확장하지 못하여 장마가 늦게 시작하였으며 태풍의 폭상으로 일찍 종료되어 8월 하순에 많은 비가 내리기 전까지 다시금 가뭄이 이어졌다(차은정, 2012).

2.2.3.5 2013-2014년 가뭄

2013-2014년 가뭄에 대해서는 학술적인 논문 등의 기록으로 남은 것은 희소하나 여러 언론 보도를 통해 가뭄 피해를 살펴볼 수 있었다. 2013년 여름에는 여러 지역에서 가뭄이 발생하였는데, 특히 부산, 동해안 및 제주 지역이 극심한 가뭄에 시달렸다. 부산 지역은 2013년 7월의 강수일수가 10일, 강수량은 131mm로 평년의 41%를 기록

하여 폭염과 더불어 농축산 농가에 큰 피해가 있었으며(부산일보, 2013년 8월 7일), 포항, 영덕, 울진 등 경북 동해안 지역에서는 심한 가뭄과 폭염으로 일부 제한급수를 실시하고 급수 및 생수를 지원하는 등 용수 공급에 차질을 빚었다(한국일보, 2013년 8월 22일). 제주 지역의 가뭄이 특히 심각하였는데, 제주 전역의 7월 강우량은 14.7mm로 평년의 6%에 지나지 않았으며 1923년 이래 90년만의 가뭄으로 일컬어졌으며 가뭄 기간도 사상 최고를 기록했다(조선일보, 2013년 8월 20일). 이로 인해 농작물 파종시기를 놓치거나 농작물이 말라죽고 축산 농가의 가축이 집단 폐사하는 등 큰 피해가 발생하였다(동아일보, 2013년 8월 20일). 주요 농작물 피해액은 8월 22일 까지 총 1,698억6,700만원 상당의 소득 감소가 발생한 것으로 추산되었다(한국일보, 2013년 8월 22일).

2014년 봄에는 충청남도, 경상남도, 동해안 지역 등에서 봄 가뭄이 발생하였는데 7월에는 특히 북태평양 고기압이 북상을 못해 충분한 장맛비를 얻지 못한 이른바 “마른장마”로 중부지방을 중심으로 7월 중순까지 극심한 가뭄 피해가 발생하였다(한국일보, 2014년 7월 14일; 조선일보 2014년 7월 14일).

2.2.3.6 2015년 가뭄

서론에서 언급하였듯이 2015년 10월 현재 극심한 가뭄이 우리나라에 지속되고 있으며 특히 인천, 경기, 강원 등 중부 지역에서는 2013년에서부터 가뭄 상태가 계속 악화되고 있다(Figure 1-1). 2015년 1월과 2월에 서울·경기를 비롯한 중부지방에는 고기압의 영향으로, 또한 동해안 지방에는 동풍의 영향으로 눈이 적게 내려 심한 가뭄이 시작되었으며, 3월 이후 이동성 고기압의 영향으로 강수량이 적어 극심한 가뭄이 지속되었다(기상청 보도자료, 2016. 6. 3.). 2015년 6월 기준으로 19개 시군구의 65개 마을 3,803 세대가 생활용수 부족을 겪고 있으며 2,451 ha의 논 물마름, 3,359 ha의 밭작물 시듦 발생 등 농업피해지역이 계속적으로 확산되고 있다(국민안전처 보도자료, 2015. 6. 12).

이에 정부는 농축산식품부에 저수지 준설비 50억원, 특별교부세 가뭄대책비 141억 원을 지원하는 등 가뭄 대응에 총력을 기울이고 있으며(국민안전처 보도자료, 2015. 6. 22.), 2015년 10월에는 특히 충청남도 서부 지역의 물부족에 대응하여 보령댐 도

수로 사업을 조기 착수하고(국토교통부 보도자료, 2015. 10. 11), 제한급수지역 주민들의 자발적인 물 절약을 유도하기 위해 절수지원제도를 최초로 시행하였다(국무총리실 보도자료, 2015. 10. 11.).

2.3 가뭄의 영향과 분류

2.3.1 정규수분지수를 이용한 가뭄 영향 분석

위성영상자료가 가뭄연구에 활용된 것은 근래의 일로 특히 2000년 이후 Terra 및 Aqua 위성이 발사되며 양 위성에 탑재된 MODIS 센서의 자료가 다양하게 활용되고 있다. 2000년 이전에는 Kogan(1995) 등에 의해 National Oceanic and Atmospheric Administration(NOAA)의 Advanced Very High Resolution Radiometer(AVHRR) 위성의 자료에서 추출한 NDVI를 통해 식생 활력도가 가뭄에 의해 받는 영향을 주로 살펴보았다. 국내에서 AVHRR에서 도출한 NDVI를 통해 가뭄을 연구하는 문헌을 살펴보면, 신사철, 어민선(2004)은 NDVI를 통해 가뭄 분석을 수행하였는데 1997년에서 2002년까지 남한지역의 VCI(Vegetation Condition Index; 정규화된 NDVI)를 살펴본 결과 1997년(충청북도, 강원도, 경상북도)과 2000년(경상남북도, 전라남도)에 지역적으로 식생 스트레스가 발생하였으며 2001년에 전국적으로 극심한 스트레스가 발생하였음을 보였다. 신수현 외(2004)는 1994-2002년의 식물 성장기인 3-10월 동안의 NDVI로 한반도의 식생 현황을 살펴보고 강수량과의 상관관계를 구하였으며 NDVI가 1-2개월 전 강수량의 영향을 많이 받으며 농지 지역에서 강수량과의 상관관계가 높게 나타나 가뭄 모니터링에 NDVI 사용이 적합하다고 결론지었으나 이를 이용하여 특정 가뭄을 분석하지는 않았다.

본 연구에서는 MODIS 센서 자료에서 추출한 NDWI를 이용하여 가뭄을 살펴보고 있는데, 국내에서도 MODIS 센서에서 추출한 식생 지수를 이용한 가뭄 연구의 사례가 다수 존재한다. 신사철 외(2006)은 2001년의 3-5월에 대하여 MODIS에서 추출한 NDVI와 기상자료 중 기온자료를 이용하여 증발산량을 산정하고 이로부터 기후학적 물수지를 통해 부족수분량을 계산하여 물부족 심도를 추정하였다. 정수, 신사철(2006)은 2001년도와 2002년도의 3-11월까지의 낙동강 권역에 대하여 NDVI와 온도

자료를 통해 증발산량을 산정하여 가뭄을 평가하였다. 박정술, 김경탁(2009)은 남한 전체 지역에 대하여 2000-2007년 기간의 4월에 대해 NDVI를 PDSI, SPI3, SPI6과의 상관성을 분석하였는데, 평균 상관계수가 아주 크지는 않았으나 NDVI가 SPI6과 특히 식생피복을 갖는 내륙지역의 관측소에서 가장 높은 상관성을 가짐을 확인하였다. 서찬양 외(2014)는 MODIS에서 산정한 ESI(Evaporative Stress Index)를 이용하여 2013년 7-8월의 남부지방 가뭄을 평가하였다.

국내 과거 가뭄 사상을 MODIS 센서 자료를 이용하여 산정한 NDWI의 2000년 3월에서 2015년 8월까지의 시계열과 비교하여 보았다(Figure 2-3). 각 대권역에 대하여 식생 토지피복을 가지는 모든 픽셀을 평균한 값을 계산하였는데, 공간적·시간적으로 비교가 가능하고 각 픽셀의 평균이 유의미하도록 Kogan(1995), Rhee et al. (2010)의 방법론을 이용하여 NDWI 값을 각 픽셀별, 월별로 최댓값과 최솟값을 이용하여 다음과 같이 정규화하고 식생의 생장 시기인 3월에서 10월까지의 값을 평균하였다.

$$Scaled\ NDWI = \frac{NDWI - NDWI_{min}}{NDWI_{max} - NDWI_{min}} \quad (Eq. 2-9)$$

우리나라의 4대 대권역에 대하여 평균한 정규화한 정규수분지수(Scaled NDWI)의 시계열, 같은 기간에 대해 평균한 SPI3 및 SPI6의 시계열, 그리고 2000년 이후 우리나라의 과거 가뭄 기록을 비교하여 보았다(Figure 2-3).

장기호 외(2011)에서 언급된 2000년 전국적인 봄가뭄과 여러 문헌에서 공통적으로 지적하고 있는 2001년의 서울, 경기, 강원, 충북, 경북 지역 등 한강 권역 등지의 봄가뭄은 모든 권역에서 SPI3와 SPI6를 통해 공통적으로 나타났다(Figure 2-3). 그러나 NDWI 값은 2000년에는 상대적으로 매우 낮았으나 2001년에는 크게 증가하였다(Figure 2-3). 이는 2001년의 봄가뭄이 PDSI에서 잘 나타나지 않았듯이(이동률 외, 2002; 이동률, 김웅태, 2001) 식생의 반응이 3-6개월의 시간보다 느리게 반응했을 수도 있으며, 또는 3월-6월의 강수량 부족에 기인했던 2001년 봄 가뭄이 종료된 6월 이후 여러 요인에 의해 식생의 상태가 크게 호전되어 3월-10월을 평균한 값에서 심각한 가뭄의 상태가 잘 나타나지 않은 것으로 보인다.

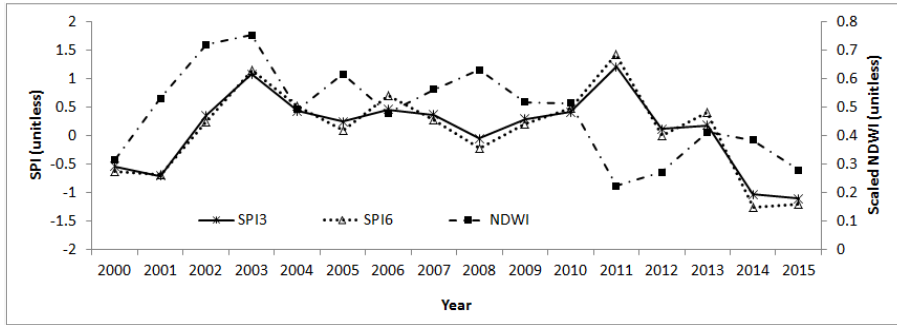
2006년의 경우 장기호 외(2011)에서 가뭄해로 언급하였으나, SPI3와 SPI6을 보면 한강, 낙동강, 영산·섬진강 권역에서 모두 전 해에 비하여 값이 증가하였으며 금강

권역에서만이 값이 감소하였음을 알 수 있다(Figure 2-3c). 반면 NDWI의 값은 네 권역에서 모두 감소하였다(Figure 2-3). 2006년은 전국적으로 태풍 에위니아와 빌리스, 개미의 피해를 많이 입은 시기로(기상청, 2007), 식생이 가뭄보다는 집중 호우와 태풍의 영향을 크게 받은 것으로 보인다.

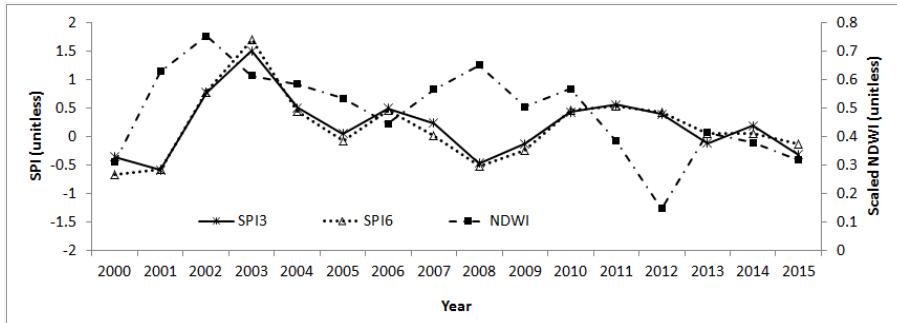
2008-2009년 전남, 경남, 강원 의 가뭄은 NDWI 변화에서는 잘 나타나지 않았다(Figure 2-3). 2008-2009년의 가뭄은 2008년 가을에 시작되었으므로 2008년 생장 시기의 NDWI 값에는 드러나지 않으나 2009년에는 특히 한강과 낙동강 권역의 값이 감소하는데 영향을 주었을 것이다.

NDWI가 가장 낮은 값을 보인 시기는 한강 권역에서는 2011년과 2012년, 낙동강, 금강, 영산·섬진강 권역에서는 2012년으로, 단기적으로 발생한 가뭄으로(오국렬, 유인상, 2014) SPI3, SPI6 값에는 크게 가뭄으로 나타나지 않았으나 식생의 수분 상태는 극심한 영향을 받은 것으로 보인다(Figure 2-3). 2011년은 6, 7월 집중호우와 8월 태풍 무이파의 영향을 받은 해로(기상청, 2012) 특히 한강 권역의 피해가 컸으므로 이 시기 한강 권역의 식생의 낮은 수분 함량은 집중 호우 및 태풍의 피해를 반영한 것으로 보인다.

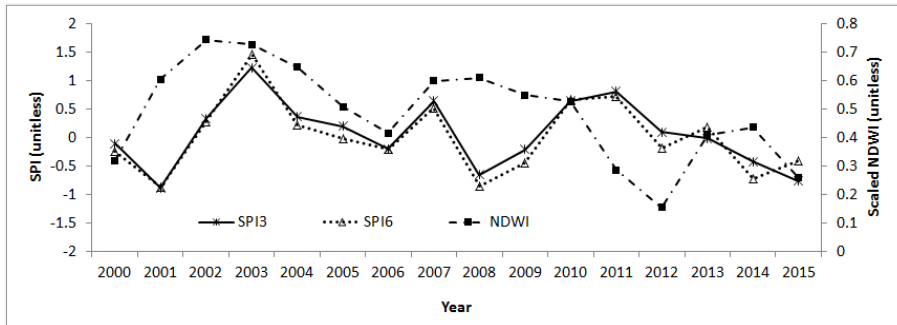
2015년에 지속되고 있는 가뭄의 상태도 볼 수 있다. 다른 해의 값은 3월-10월의 값을 평균하였으나 여기에 2015년의 값은 3월-8월의 값만 평균되어 표시하였는데, 영산·섬진강 권역을 제외하고는 전년도와 비교하여 식생의 수분 함량이 상당히 감소하였음을 볼 수 있다(Figure 2-3).



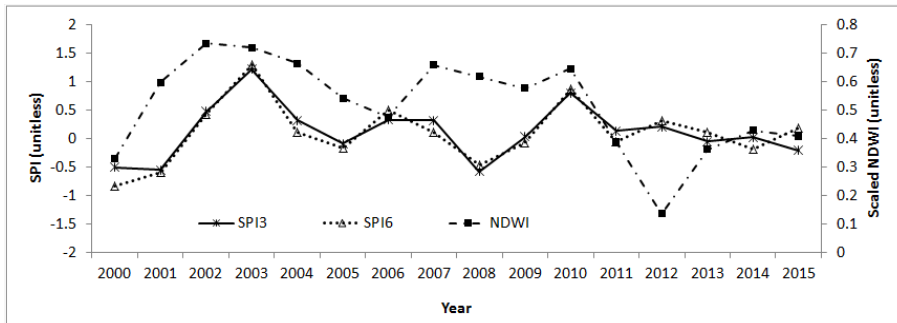
(a) Han-River Basin



(b) Nakdong-River Basin



(c) Geum-River Basin



(d) Yeongsan & Sumjin-River Basin

Figure 2-3 Scaled NDWI averaged for each river basin during March-October.

2.3.2 가뭄 피해 유형 분류

가뭄은 앞서 인용한 바와 같이 기상학적 가뭄, 토양수분 가뭄, 수문학적 가뭄 등으로 나누어 정의할 수 있다. 이는 가뭄이 진행되는 정도에 따른 것인데, 효과적인 가뭄 관리를 위해서는 이러한 단순한 분류 체계를 세분화하고 명확히 할 필요가 있다. 기존의 분류도 가뭄과 관련된 이해당사자(stakeholders)를 고려한 것이지만(Seneviratne et al., 2012) 가뭄정보 제공을 통해 다양한 가뭄 피해를 최소화하기 위해서는 이를 다시 가뭄의 피해 유형으로 나누는 것이 유용할 것이다.

가뭄 피해 유형이 인지되고 있는 범위를 대략적으로 파악하고자 2012년 수립된 16개 광역지방자치단체의 기후변화 적응대책 세부시행계획 중 가뭄과 관련된 사항을 검토하였다. 가뭄의 피해로 공통적으로 인지되고 있는 수자원 확보 문제를 제외하고 살펴보면, 충청북도는 미세먼지 농도 증가, 충청남도는 농작물 및 가축물 피해, 전라북도는 소나무 고사, 전라남도는 농작물 및 양식장의 생산량 저하를 언급하였으며 광주광역시에는 근래의 가뭄 피해 사례로 농수산물 작황부진, 산불, 간척지 염해를 꼽았다. 제주특별자치도는 밭작물 생육 불량, 수목 고사, 경상남도는 대형 산불의 우려를 언급하였으며 부산광역시는 수목 고사 및 생육 불량과 산불 위험성 증대를 가뭄 영향으로 들었다. 대구광역시는 송이 작황 부진, 수질 악화, 산불 위험을 가뭄 피해로 꼽았는데 일반추진분야 2순위에 가뭄에 의한 산림식생이 포함되는 등 가뭄으로 인한 피해에 민감한 것으로 나타났다. 경상북도는 가뭄으로 인한 작황부진, 소나무 고사, 산불 위험도 증가, 수질 저하 등을 언급하였으며 기후변화로 인한 중점 추진분야를 선정할 때 1순위가 산불, 소나무, 가뭄에 의한 산림 식생 등이 포함된 산림분야, 2순위가 가뭄 관련 이수, 3순위가 폭염, 한파, 4순위가 비생산성 등으로 가뭄의 영향이 큰 분야의 우선순위를 높게 선정하였다.

본 연구에서는 가뭄의 피해 유형으로 김연주 외(2013)의 분류(Table 2-1)를 차용하였는데, 이는 앞서 살펴본 여러 광역지자체가 인식하고 있는 가뭄의 피해를 광범위하게 포함하고 있다. 김연주 외(2013)는 농경 및 경제적 피해에 집중했던 과거에서 점차 다양한 부문에 대한 가뭄의 파급 효과가 고려되고 있는 추세를 고려하여 가뭄 피해 유형을 다섯 가지로 나누었다(Table 2-1). 여기에서 용수공급 피해는 생활용수와 발전용수 부족으로 인한 피해를 나타내며 농수축산업 피해는 농업, 축산업, 수산

업 등 1차 산업에 대한 가뭄 피해를 나타낸다(김연주 외, 2013). 경제적 피해는 경제적 피해 중에서 용수공급과 농수축산업 분야를 제외한 가뭄으로 인한 경제적 피해를 의미하며 사회적인 피해는 물관련 분쟁이나 공공보건과 관련된 피해를 포함한다(김연주 외, 2013). 본 연구에서는 산림에 대한 피해나 산불 위험성의 증가 등 가뭄의 환경적 피해에 특히 초점을 맞추고 있는데, 가뭄의 환경적 피해는 가뭄으로 인해 대기질, 수질, 생태 등에 피해가 발생한 것을 포함한다.

Table 2-1 Classifications of Drought Impacts (adopted from Kim et al., 2013).

피해 유형	세부 분류	피해 내용 예시
용수공급 피해	생활용수	절수, 잔디 관개 제한
	발전용수	발전량 감소, 수력발전 비용 증대
농수축산업 피해	농업	농업용수 피해, 농작물 품질 및 수확량 저하, 해충 증가
	축산업	축산물 품질 저하, 가축 질병 증대
	수산업	어류 품질 저하, 어류 서식지 손상
	임업	화재위험 증가, 병충해 만연, 삼림지 손상
경제적 피해	국가재정기관 부담	담보권 행사, 신용리스크 증가, 자본손실
	수자원 개발 비용	신규/보충 수자원 개발
	산업 및 여가사업	여가지 손실, 가뭄 관련 전력삭감
	물류	하천 및 운하 손상, 수운/물 수송 비용
환경적 피해	생태	생물 다양성 손실, 포식 증대
	수질	염수화, 자정능력 저감
	대기질	먼지, 공해
	수문	습지 소실, 수위 저하, 지하수 고갈, 지반 침강
	가시적 조경	먼지, 식생 면적, 심미성 저하
사회적 피해	갈등 증대	물 이용자 간 충돌, 관리당국 충돌
	건강 문제	정신 및 육체적 스트레스, 호흡기 질병 증가

3. 가뭄 정보의 활용

의사결정자의 가뭄 정보 수요를 파악하고 활용 의사 등을 도출하기 위하여 17개 광역지방자치단체의 가뭄 및 가뭄의 환경 피해와 관련이 있는 담당자를 대상으로 설문조사를 수행하였다. 설문 대상 공무원은 각 광역지자체의 홈페이지에 게시되어 있는 직원 정보를 바탕으로 가뭄 환경피해 관련 업무인 가뭄·폭염, 산불, 재난·재해, 도시숲(도시녹화) 담당자를 선정하여 시도별 2-4명으로 구성하였으며 복수의 대상자가 있을 경우 담당 업무 목록에 상이한 분야가 적고 해당 분야 업무에 집중되어 있는 대상자를 우선적으로 선정하였다. 총 53명을 대상으로 20개 문항의 1차 가뭄감시 및 예측정보 활용현황 및 인식조사를 수행하였다(부록 1).

제2차 설문조사는 본래 1차 설문조사를 통해 선정된 대상 광역지자체를 대상으로 심화인터뷰를 통해 상세 가뭄피해 목록을 작성하여 가뭄 리스크를 평가하는 것으로 계획되었으나, 1차 설문조사 결과를 분석한 결과 선정 지자체에 대한 다양한 가뭄 피해 인벤토리 작성 및 상세 리스크 평가보다는 여러 지자체에 대한 개략적인 가뭄 리스크 평가 및 가뭄 지표의 활용과 가뭄 정확성에 대한 인식 등에 대한 조사가 선행되어야 한다는 판단을 내렸다. 따라서 1차 설문조사에 응답한 27명(회수율 53%)의 공무원을 대상으로 10개 문항의 2차 설문조사를 수행하여 5개 광역지자체로부터 답변을 회수하였다(회수율 19%; 부록 2).

가뭄 및 가뭄의 환경 피해 관련 담당 공무원들의 응답을 모두 받은 것은 아니지만 본 설문조사는 표본을 통해 모집단의 특성을 도출하려는 표본조사보다는 실제 가뭄 정보의 활용과 밀접한 광역지자체 공무원을 대상으로 한 전수조사의 성격을 가지도록 수행되었다. 따라서 회수율은 저조하나 회신한 광역지자체에 대해서는 유의하다고 하겠다. 두 차례에 걸쳐 수행된 설문조사는 전체 30문항으로 이루어져 있으며 이를 통해 가뭄 발생 인지 및 업무 관련성, 가뭄 정보의 활용, 가뭄 대처 현황 및 수요, 가뭄 정보의 정확성에 대한 수요 등을 도출하고 간략한 가뭄 리스크 평가를 수행하였다.

3.1 가뭄 정보의 활용 현황

3.1.1 과거 가뭄 피해

가뭄 정보의 활용 여부는 과거 가뭄 피해 여부와 관련이 있을 수 있다. 1차 설문에서 대부분의 응답자가 본인의 담당 업무가 가뭄의 발생의 영향을 받는다고 답변하였는데 (70%인 19명이 매우 영향을 받는다, 26%인 7명이 약간 영향을 받는다; Figure 3-1), 담당 업무 분야에 대한 과거 가뭄 피해에 대해서는 52%(14명)는 매년 또는 빈번히 발생한다, 19%(5명)는 지난 5년 안에 발생한 적이 있다, 7%(2명)는 지난 10년 안에 발생한 적이 있다고 답하였다(Figure 3-2). 15%(4명)의 응답자는 발생한 적이 없다고 답하였는데, 이렇게 답한 응답자의 75%(3명)는 해당 업무 경력이 1년 미만, 25%(1명)는 1년-2년으로 답하여 관련 업무 수행 시 가뭄 피해를 입지 않았을 가능성이 있다(해당 업무 경력에 대한 자료 미제시).

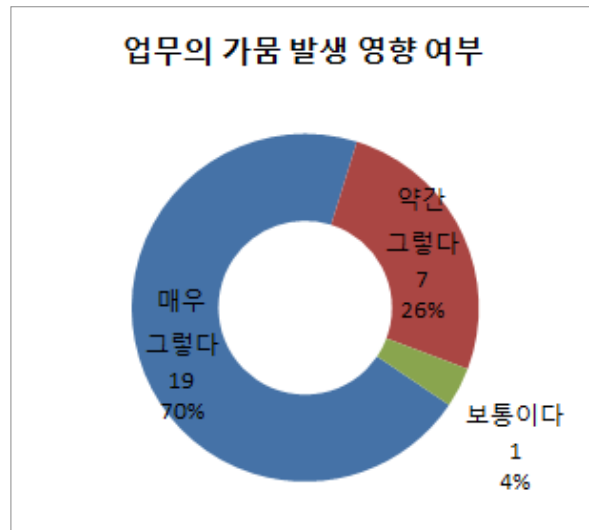


Figure 3-1 Percentages of responses on whether one's duties are affected by drought events.

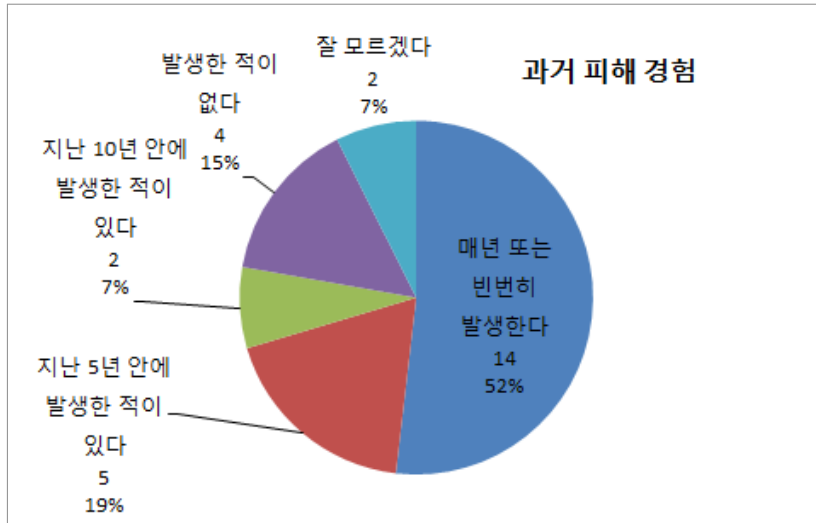


Figure 3-2 Frequencies of past experiences of drought impacts.

1차 설문지의 응답자의 담당 업무와 과거 피해 경험과의 연관성을 분석하기 위해 담당 업무를 산불군, 도시숲군, 가뭄·폭염군, 재난·재해군으로 나누어 일원 분산분석 (One-way ANOVA, Analysis of Variance)을 수행하였다. 선택지의 과거 가뭄 피해 경험은 정량적인 값이 아니나 응답자가 순차적인 단계로 인식하였을 것을 가정하여 “매년 또는 빈번히 발생한다”를 1의 값으로, “지난 5년 안에 발생한 적이 있다”를 2의 값, “지난 10년 안에 발생한 적이 있다”를 3의 값, “발생한 적이 없다”를 4의 값으로 하여 “잘 모르겠다”로 답한 응답자를 제외한 25명의 응답자를 담당 업무 군집으로 나누어 군집별 통계 값을 구하였으며(Table 3-1; 담당 업무에 대한 군집별 통계값), 각 군집별 평균에 차이가 없다는 귀무가설을 검정하였다. 일원 분산분석을 통한 검정 결과, 95% 신뢰수준에서 응답자의 담당 업무에 대한 군집별 응답 결과 F 통계량이 7.16으로 F 기각치인 3.07보다 커 귀무가설을 기각하였다(p -값 = 0.0017; Table 3-2). 즉, 담당 업무별로 과거 가뭄 피해 경험이 다르게 나타났다고 볼 수 있다.

일원 분산분석을 통해 응답자의 담당 업무별로 과거 가뭄 피해 경험의 평균이 모두 같지는 않다는 것을 알았으며, 이번에는 각각 담당 업무 군집 쌍에 대하여 평균이 다른지 검정하기 위해 양측 t -검정을 수행하였다. 각 담당 업무 쌍에 대하여 수행한 결과 95% 신뢰수준에서 재난·재해군과 산불군에서 통계적으로 유의미한 차이가 나타나(t 통계량 2.573 > 양측 검정 t 기각치 2.570, p -값 = 0.0498; 표 3) 두 군집의

평균이 같다는 귀무가설을 기각하였다.

각 담당 업무 쌍에 대하여 평균의 크고 작음을 평가하기 위해 단측 t-검정도 수행하였는데, 95% 신뢰수준에서 재난·재해군과 산불군, 재난·재해군과 도시숲군, 가뭄·폭염군과 산불군 사이에 차이가 있었으며 재난·재해군과 가뭄·폭염군이 산불군이나 도시숲군과 비교하여 과거 가뭄 피해의 빈도를 낮게 응답하였다(Table 3-3).

앞서 언급하였듯이 담당 업무를 재난·재해로 답한 응답자군과 가뭄·폭염으로 답한 응답자군은 분리하여 생각하기 어려운 면이 있으므로 두 응답자군을 통합하여(Table 3-4) 일원 분산분석과 t-검정을 앞서와 동일하게 실시하였다(Table 3-5). 일원 분산분석 결과 F 통계량(11.25)이 F 기각치(3.44)를 크게 넘어서 마찬가지로 귀무가설을 기각하였으며(p-값 = 0.00043), t-검정의 경우 양측 및 단측 검정 결과 재난·재해와 가뭄·폭염을 통합한 통합군과 산불군, 통합과 도시숲군 사이에 차이를 볼 수 있었다. 재난·재해와 가뭄·폭염 통합군이 가뭄으로 인한 과거 피해 경험의 빈도를 낮게 응답하였다.

이 결과는 중요한 시사점을 가지는데, 가뭄 발생의 영향 및 피해에 대한 조사 등을 수행할 때 일반적인 가뭄이나 재난재해 담당 공무원의 답변만을 참고할 경우 가뭄에 의한 피해가 과소평가될 우려가 있다. 특히 제2장에서 나열한 가뭄 피해 유형 중 용수공급 피해, 농수축산업 피해 등만이 일반적으로 인식되어 있으며 실제 광역상수도 등의 설비가 잘 되어 있어 해당 피해 유형에 대한 가뭄 피해를 인식하기 쉽지 않으므로 특히 가뭄의 환경적 피해는 간과되기 쉽다.

Table 3-1 Statistics for each duty groups.

Group	Responses	Sum	Mean	Variance
Disaster	5	15	3	2
Drought, Heatwave	3	9	3	1
Wildfire	8	10	1.25	0.5
Urban Greening	9	12	1.333	0.25

Table 3-2 ANOVA results for each duty group (95% significance level).

Source of Variation	Sum of Squares	Degree of Freedom	Mean Square	F	P-value	F crit
Treatment	15.86	3	5.287	7.163	0.0017141	3.072
Residual	15.5	21	0.738			
Total	31.36	24				

Table 3-3 Two-tail and one-tail t-test statistics (95% significance level).

Statistics	Disaster vs. Drought/Heatwave	Disaster vs. Wildfire	Disaster vs. Urban Greening	Drought/Heatwave vs. Wildfire	Drought/Heatwave vs. Urban Greening	Wildfire vs. Urban Greening
t Stat	0.000	2.573	2.548	2.782	2.774	-0.277
P(T<=t) one-tail	0.500	0.025	0.026	0.034	0.055	0.393
t Critical one-tail	1.943	2.015	2.015	2.353	2.920	1.782
P(T<=t) two-tail	1.000	0.050	0.051	0.069	0.109	0.786
t Critical two-tail	2.447	2.571	2.571	3.182	4.303	2.179

Table 3-4 Statistics for each duty groups (disaster group and drought/heatwave group were merged).

Group	Responses	Sum	Mean	Variance
Disaster/Drought/Heatwave	8	24	3	1.429
Wildfire	8	10	1.25	0.5
Urban Greening	9	12	1.333	0.25

Table 3-5 Two-sided and one-sided t-test statistics (95% significance level; disaster group and drought/heatwave group were merged).

Statistics	Disaster/Drought/Heat wave vs. Wildfire	Disaster/Drought/Heat wave vs. Urban Greening	Wildfire vs. Urban Greening
t Stat	3.564	3.669	-0.277
P(T<=t) one-tail	0.002	0.003	0.393
t Critical one-tail	1.796	1.833	1.782
P(T<=t) two-tail	0.004	0.005	0.786
t Critical two-tail	2.201	2.262	2.179

3.1.2 가뭄 정보의 활용

담당 업무에 기상청 또는 다른 기관에서 제공하는 가뭄 정보를 사용하는지에 대한 질문에 대해 70%(19명)의 응답자가 기상청 또는 다른 기관에서 제공하는 가뭄 정보를 활용하고 있는 것으로 나타났다(Figure 3-3). 현재는 가뭄 정보를 사용하고 있지 않지만 앞으로 활용할 의사를 밝힌 응답자도 19%(5명)에 이르렀다(Figure 3-3). 가뭄 정보의 활용 여부를 과거 가뭄으로 인한 피해 경험에 대한 답변과 연관지어 분석하면, 실제 가뭄으로 인한 피해가 담당 업무 분야에 매년 또는 빈번히 발생한다고 답한 응답자 52%(14명) 중에 79%(11명)가 가뭄 정보를 활용하고 있다고 답했으며 14%(2명)은 현재 활용하고 있지는 않지만 앞으로 활용하고 싶다고 답했다(Figure 3-4). 또한 지난 5년 안에 피해가 발생한 적이 있다고 답한 응답자 19%(5명) 중 21%(4명)는 가뭄 정보를 활용하고 있는 것으로 응답하였다(Figure 3-4). 가뭄으로 인한 피해가 발생한 적이 없다고 응답한 15%(4명) 중 50%(2명)는 가뭄 정보를 활용하고 있으며 가뭄 피해에 대해 잘 모르겠다고 응답한 7%(2명)는 현재 활용하고 있지는 않지만 앞으로 활용하고 싶다고 답했다(Figure 3-4).

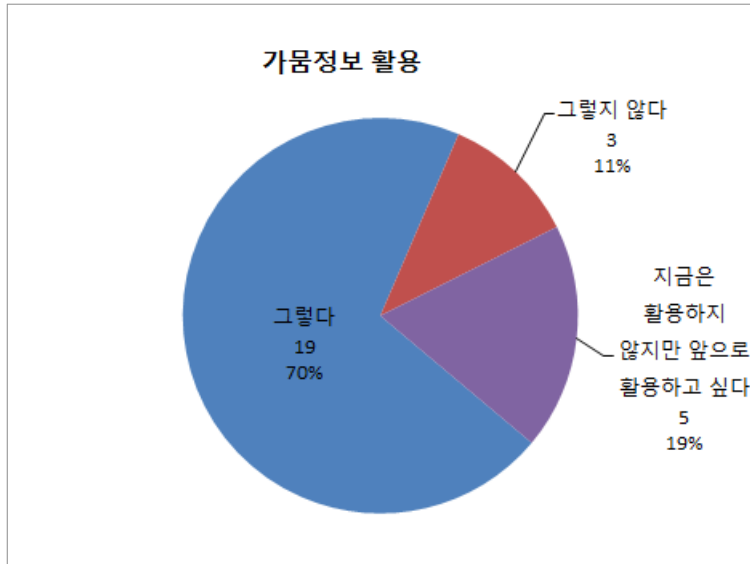


Figure 3-3 Percentages of responses on whether drought information is used for one's duties.

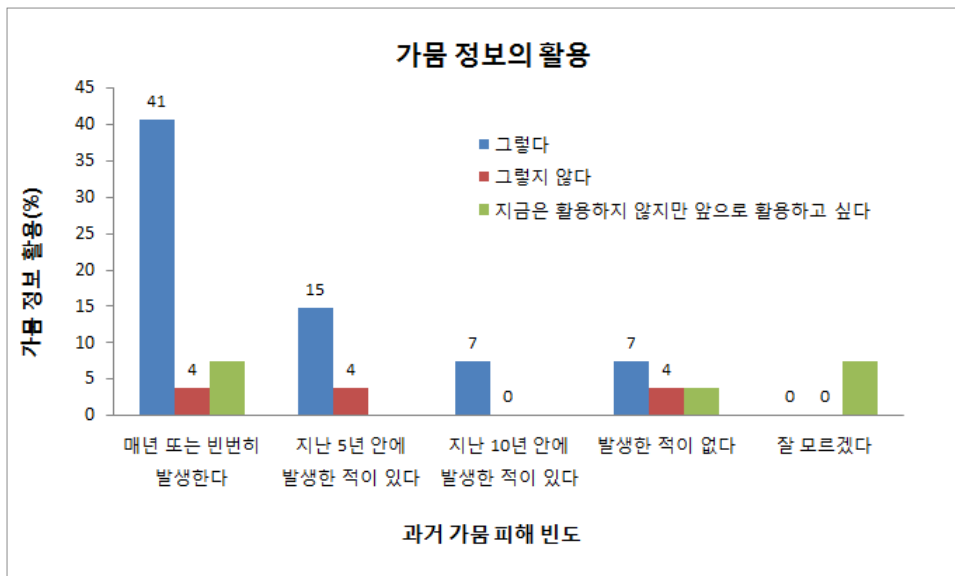


Figure 3-4 Percentages of responses on whether drought information is used for one's duties, for each response on the frequencies of past experiences of drought impacts.

가뭄 정보를 현재 활용하고 있는 70%(19명)의 응답자 중 58%(11명)은 기상청 자료를 활용하는 것으로 나타났다. 활용하는 가뭄 정보의 종류는 기상청에서 제공하는 표

준강수지수, 파머가뭄지수, 가뭄판단지수(현재는 제공되고 있지 않음), 지역별 관측 자료, 누적 강수량, 건조특보 등이었으며 그 밖에 응답된 기관 및 자료는 산림청의 산불위험예보시스템 정보, 수자원공사의 댐 정보, 농어촌공사의 저수율 정보, 일본기상청 및 미국기상청의 자료와 뉴스 등의 언론 등이 있었다. 신뢰도에 대해 응답한 89%(17명) 중 88%(15명)는 해당 가뭄 정보를 “매우 신뢰” 또는 “약간 신뢰” 하는 것으로 답하였다.

가뭄 정보를 활용하지 않는다고 답한(Figure 3-3) 11%(3명)의 응답자는 각각 “해당 업무가 가뭄과 관련이 없어서”, “가뭄 정보를 쉽게 얻을 수 없어서”, “가뭄 정보는 있으나 업무와 연결 지어 활용하기가 쉽지 않아서”로 답하였는데, 이 중 “해당 업무가 가뭄과 관련이 없어서”로 답한 응답자의 경우 업무의 가뭄 연관성을 묻는 앞선 질문에는 “약간 그렇다”로 답하였으나 가뭄 과거 발생 피해 경험을 묻는 질문에 “발생한 적이 없다”고 답한 것으로 보아 지금까지의 업무에 가뭄 피해 경험의 부재가 이러한 답변으로 연결되었다고 볼 수 있다.

제2차 설문조사를 통해 위성 자료로부터 도출한 가뭄 정보의 활용에 대해 물었다. 특히 기상 관측 자료가 없는 지역에 대해 위성 자료에서 도출한 NDVI, NDWI 등이 활용될 수 있음을 제시하였는데, 응답자 중 60%(3명)은 위성에서 도출한 가뭄 정보에 대해 들어본 적이 없다고 답변하였고 40%(2명)은 이미 위성 기반의 가뭄 정보를 활용하고 있다고 답하여 위성 기반 자료에 있어서는 정보 활용의 범위가 넓게 나타났다. 위성 기반 정보의 출처는 방재기상청의 자료(방재기상정보시스템으로 생각됨), 국립산림과학원 자료, 지방기상청자료, 그리고 직장동료로 나타났다.

3.2 가뭄 리스크 평가

제1차 설문조사에서 앞서 가뭄으로 인한 과거 피해 경험에 대해 “잘 모르겠다” 또는 “발생한 적이 없다” 등으로 답변한 응답자 등 26%(7명)을 제외하고 가뭄의 피해에 대해 구체적으로 서술하였다. 이를 담당 업무별로 나열하면 Table 3-6과 같다. 재난·재해와 가뭄·폭염은 통합하여 재난재해통합, 산불, 도시숲의 세 항목으로 나타내었으나 실제 담당 업무는 중복된 경우가 존재하므로 재난재해통합 분류의 항목 중 산불발생이 있는 등 내용상 교차되는 항목이 많다.

이상적으로는 이러한 각각의 담당 업무에 대해 상세한 가뭄 피해 인벤토리를 작성하고 그에 대한 리스크 평가가 수행되어야 하나 1차 설문조사 결과를 분석한 결과 선정 지자체에 대한 다양한 가뭄 피해 인벤토리 작성 및 상세 리스크 평가보다는 여러 지자체에 대한 개략적인 가뭄 리스크 평가 및 가뭄 지표의 활용과 가뭄 정확성에 대한 인식 등에 대한 조사가 선행되어야 한다는 판단을 내렸다. 따라서 제2장에서 언급한 가뭄피해유형 분류(Table 2-1)를 이용하여 가뭄피해의 발생빈도와 가뭄피해 강도를 각각 5단계로 선택하게 하였으며, 이를 통해 개략적인 가뭄 리스크 평가를 수행하였다.

가뭄 리스크는 가뭄피해 발생빈도(1~5의 값)와 가뭄피해 강도(1~5의 값)로부터 제1장의 Eq. 1-1에 따라 산정하였다. 응답자A~E의 경우 각 가뭄피해 분류에 따른 가뭄리스크가 매우 상이하게 나타났는데, 응답자의 수가 적으므로 각 응답자별 응답의 통계값을 제시하기 보다는 실제 계산값을 나타내었다(Table 3-7). 응답자A는 전반적인 항목에서 가뭄의 리스크를 매우 크게 평가한 반면 응답자C는 작게 평가하였다. 응답자B와 응답자D가 평가한 가뭄리스크는 그 범위에서 유사하게 나타났는데, 실제 가뭄의 발생빈도와 가뭄의 피해강도를 살펴보면 응답자B는 전반적으로 저중도의 빈도 및 강도로 평가한 반면 응답자D는 낮은 편인 가뭄의 발생빈도에 반하여 약간 높은 가뭄의 피해강도를 평가하였다(자료 미제시).

앞서 t-검정의 결과로 재난·재해 통합군의 경우 산불이나 도시숲 담당자군에 비하여 가뭄 피해를 크게 인식한다고 하였는데(Table 3-5), 본 결과에서는 표본의 수가 작아 같은 경향을 파악할 수는 없었다. 대부분 자신의 관련 업무 분야의 가뭄 리스크를 높게 평가하였지만, 응답자들이 가뭄의 발생빈도와 피해강도를 평가함에 있어 그 간 간과되어 온 가뭄의 환경적 피해와 사회적 피해를 다른 피해 유형과 마찬가지로 높게 평가한 점은 작은 표본 수에도 불구하고 주목할 만 하다. 본 결과를 기반으로 각 피해 유형의 상세한 리스크 평가가 수행되어야 할 것이다.

Table 3-6 Major impacts of drought for each duty group.

담당 업무 그룹	과거 가뭄 피해
재난재해통합	각종 안전사고 발생빈도 증가
	농업용수 부족으로 모내기 이양 작업 지연, 발작물 시듦 발생
도시숲	공원 숲, 띠녹지, 녹지대, 중앙분리대, 가로수, 조경수 및 초화류 고사
	고온건조로 도시숲 조성사업 공사 하자 발생을 증가
산불	산불 발생 증가
	더위로 인한 산림 내 취사행위 증가
	수목 식재 적기 가뭄 발생에 따른 식재 및 유지관리 제한

Table 3-7 Drought risk for each drought impact category and respondent.

피해 유형	세부 분류	응답자A	응답자B	응답자C	응답자D	응답자 E
용수공급 피해	생활용수	20	4	1	8	12
	발전용수	25	4	1	8	-
농수축산업 피해	농업	25	4	1	12	-
	축산업	-	12	1	6	-
	수산업	-	9	1	4	-
경제적 피해	임업	25	9	4	4	12
	국가재정기관 부담	25	2	1	8	12
	수자원 개발 비용	25	6	4	6	-
환경적 피해	산업 및 여가사업	12	4	1	6	-
	물류	20	6	1	2	-
	생태	25	4	1	3	-
사회적 피해	수질	25	6	4	6	-
	대기질	20	12	4	6	-
	수문	25	4	1	4	-
사회적 피해	가시적 조경	16	12	4	8	12
	갈등 증대	25	6	1	8	-
	건강 문제	25	4	4	8	-

※ 응답자A는 산불 담당, 응답자B, E는 도시숲 담당, 응답자C는 가뭄, 폭염, 응답자D는 재난, 재해 담당임.

3.3 가뭄 정보 수요

3.3.1 가뭄감시 정보의 수요

제1차 설문문의 가뭄 정보 활용 여부에 대한 질문에서, 앞으로 가뭄 정보를 활용하고 싶다고 답한 19%(4명)은 가뭄의 미래 활용 의사에 대하여 현재는 가뭄과 관련하여 무슨 자료가 제공되는지 잘 모르거나 가뭄을 판단하는 수치가 정확하지 않고 각 중앙부처별로 가뭄 판단 기준이 제각각이어서 활용하지 않지만 앞으로는 가뭄 정보를 활용하고 싶다고 밝혔다. 한 응답자는 지금은 기상청의 정보보다 수목 식재 시 토양의 상태나, 순찰을 통한 수목 및 식물의 상태를 확인하여 급수 필요성이 인지될 경우 급수 작업을 하고 있으나 기상청의 빨래지수나 야외지수처럼 좀 더 구체화된 정보가 나온다면 활용할 의사가 있다고 밝혀 맞춤형 정보의 필요성을 보였다. 이들은 현재 기상청 등에서 제공하는 정보가 업무에 직접적으로 활용될 만큼 구체적이지 않다고 여겨져 사용하고 있지 않은 것으로 보이며, 따라서 가뭄 정보를 실제 사용자에게 유용한 형태로 제공하게 되면 활용도를 높일 수 있을 것으로 판단된다. 담당 업무별로 업무에 가장 도움이 될 것으로 응답한 가뭄감시 정보의 종류 및 형태를 시간해상도/공간해상도/활용의 순으로 나타내었다(Table 3-8).

3.3.2 가뭄예측 정보의 수요

제1차 설문에서 가뭄에 대비하는 시기에 대한 질문을 하였는데, 가뭄에 대비하는 시기에 대해서는 “가뭄이 상승적으로 발생하는 시기에 대하여 평상시 대비”를 선택한 응답자가 70%(19명)으로 가장 많았다(Figure 3-5). 이 경우 가뭄이 상승적으로 발생하는 시기라는 것은 과거 기후값에 의지한다는 것을 의미한다. 기후변화로 인해 가뭄의 발생 추이가 달라지고 있으나, 과거의 사상에 의지해서라도 평상시 대비한다는 것은 적극적인 예방 활동에 해당한다. 또 22%(6명)의 응답자는 “가뭄 발생을 인지한 순간부터 대비”한다고 답변하였는데(Figure 3-5), 이는 적극적인 가뭄 감시 행동을 나타낸다. 그러나 8%(2명)의 응답자는 “가뭄이 심화되어 해당 업무에 영향/피해가 발생한 순간부터 대처”한다고 답해(Figure 3-5) 일부의 경우 예방 및 적절한 감시 및 대처가 부재함을 나타내었다.

가뭄 예측 정보 활용 의사에 대해서는 81%(22명)의 응답자가 활용할 의사가 있다고 답했는데, 15%(4명)의 응답자는 활용할 것이지만 기상청 등의 중앙기관에서 제공하는 경우에 활용할 것이라고 답하였다(Figure 3-6). 이는 가뭄 예측 정보를 제공할 경우 자료의 공신력이 매우 중요함을 나타낸다.

Table 3-8 Preferred drought monitoring information.

담당 업무	가뭄 정보 수요
재난재해통합	<ul style="list-style-type: none"> 주별/시군구/폭염, 가뭄 대비 일, 주, 월별/NA/가뭄의 유형 및 피해 상황분석, 소관분야별 지원대책 추진
산불	<ul style="list-style-type: none"> 주별/NA/NA
도시숲	<ul style="list-style-type: none"> 주별/NA/폭염 주별/NA/NA 주별/NA/농작물의 상태 등 식물이나 소형 저수지 일, 주별/NA/도심지의 공원과 녹지의 토양수분 자료 주, 월별/NA/NA

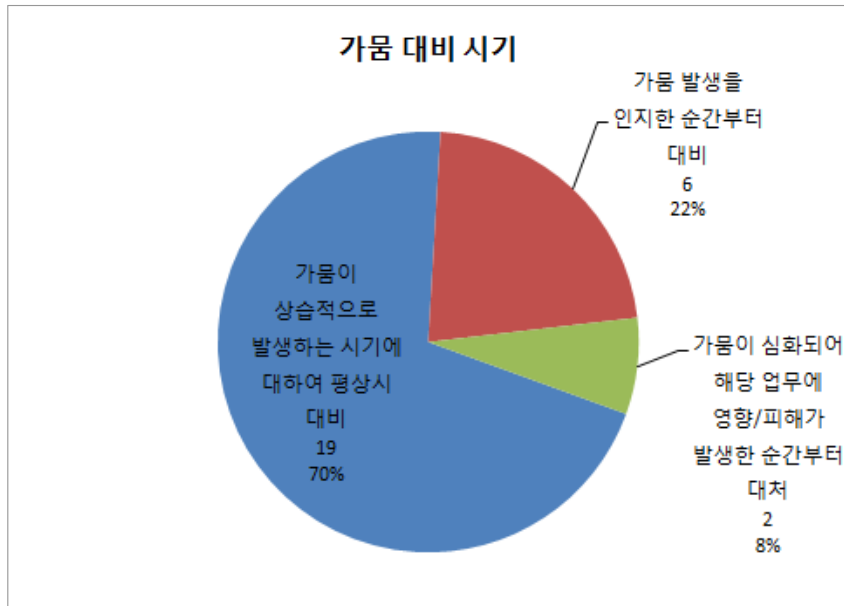


Figure 3-5 Timing of drought preparedness.

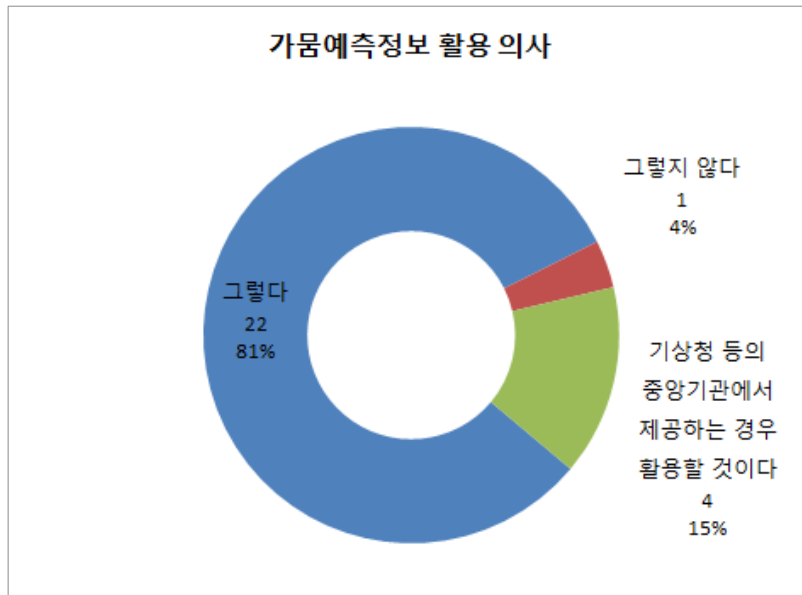


Figure 3-6 Percentages of responses on whether to use drought forecasting information.

가뭄예측 정보의 유용성을 파악하기 위하여 가뭄예측 정보가 사전에 주어진다면 활용할 수 있는 대처 방안을 요구하였다. 예측정보의 선행 시간으로는 주일 전, 1개월 전, 3개월 전, 6개월 전, 1년을 제시하였다. 재난관리의 예방-대비-대응-복구 단계에 따르면 장기적 관점에서 행해지는 관리 활용은 예방에 해당하고 비상시에 효과적인 대응을 하도록 하는 사전준비 활용은 대비, 재난 발생 직전, 직후, 또는 재난 진행 중에 행해지는 일련의 활동이 대응에 해당한다. 가뭄 예측 정보를 활용하여 담당 업무 분야에서 활용할 수 있는 대처 방안은 일반적으로 대비 활동에 해당하지만 여기서 가뭄정보 제공의 시간 단계와 연결지어 생각하면 1주일과 같이 가뭄 발생이 예측되는 시기와 가까운 경우 대비, 대응 활동이 함께 취해지게 되며 1년과 같이 장기적인 관리 활동이 가능한 시간의 경우 대비와 예방 활용이 함께 수행될 수 있게 된다. 응답자들의 답변을 묶어 Tables 3-9에서부터 3-13까지 제시하였다.

Table 3-9 Drought measures available for the 1-week forecast lead time.

담당 업무	대처 방안
재난재해통합	<ul style="list-style-type: none"> • 대책 및 지원 <ul style="list-style-type: none"> - 유관기관 대책 회의 - 시군구 지원반 운영 및 급수 지원 • 장비 및 인력: 응급 복구 장비 지원 가능 업체 연락 • 홍보: 안내 방송 등을 통한 절수 홍보
산불	<ul style="list-style-type: none"> • 대책 및 지원 <ul style="list-style-type: none"> - 산불 방지 대책 수립, 예방 대책 강화 - 시군구 산불 예방 특별 대책 수립 지시 - 시군구 비상근무 등 강화 지시 • 정보생산 및 제공 <ul style="list-style-type: none"> - 가뭄 예측 정보 문자 발송 - 지역별 실시간 가뭄정보 검색 시스템 구축 • 장비 및 인력 <ul style="list-style-type: none"> - 산불 감시 인력 증원 및 집중 배치 - 순찰 활동 강화
도시숲	<ul style="list-style-type: none"> • 대책 및 지원 <ul style="list-style-type: none"> - 가뭄 피해 예방 계획 실행 - 관수 계획 공문 발송 - 가로수 비상 급수 • 용수 확보: 관정 개발 • 장비 및 인력: 장비 임차, 인력 확보 및 배정 • 홍보: 시민대상 절수 홍보 강화

Table 3-10 Drought measures available for the 1-month forecast lead time.

담당 업무	대처 방안
재난재해통합	<ul style="list-style-type: none"> • 대책 및 지원 <ul style="list-style-type: none"> - 주민 불편 해소 및 피해 경감 대책 추진 - 갈수기 수질오염사고 예방 대책 추진 • 용수확보 <ul style="list-style-type: none"> - 용수 개발 추진 계획 수립 - 응급 용수 확보 지원 확인 • 장비 및 인력: 가뭄 대책 장비 점검 및 정비 • 홍보: 절수 운동 추진, 가뭄 대처 요령 책자 보급
산불	<ul style="list-style-type: none"> • 대책 및 지원 <ul style="list-style-type: none"> - 산불 예방 활동 강화, 산불 방지 대책 수립 - 산림 연접지 인화물질 제거 사업 - 입산 통제구역 및 등산로 폐쇄 구간 지정 - 수목 물주머니 설치

담당 업무	대처 방안
	<ul style="list-style-type: none"> • 정보생산 및 제공: 지역별 실시간 가뭄정보 검색 시스템 구축 • 용수확보: 취수원 확보 • 장비 및 인력: 산불 감시, 진화 인력 총원 및 집중 배치, 순찰 활동 강화 • 홍보: 절수 홍보
도시숲	<ul style="list-style-type: none"> • 대책 및 지원 <ul style="list-style-type: none"> - 도시숲, 가로수 관리 대책 방안 수립 후 시군에 시달 - 가뭄 피해 예방 실행 • 용수확보: 지하수 개발 등 사전 준비 • 장비 및 인력: 정기적 점검 및 장비 임차, 인력 동원 사항 최종 점검 • 홍보: 시민 대상 절수 홍보

Table 3-11 Drought measures available for the 3-month forecast lead time.

담당 업무	대처 방안
재난재해통합	<ul style="list-style-type: none"> • 대책 및 지원 <ul style="list-style-type: none"> - 대책 계획 수립 - 가뭄 대책 T/F 구성 운영 - 관련 부서간 역할 분담 및 공동대응 협력 체계 구축 • 용수확보 <ul style="list-style-type: none"> - 관정 개발, 기존 수자원 시설 활용 등 용수 확보 계획 수립 - 농업분야 및 용수 공급지하수 등을 위한 사업 발주 준비 • 정보생산 및 제공: 가뭄 지도 작성 및 보급 • 장비 및 인력: 장비 점검 정비 • 홍보: 홍보 강화 및 절수 대책 수립
산불	<ul style="list-style-type: none"> • 대책 및 지원 <ul style="list-style-type: none"> - 산불 예방 활동 강화, 산불 방지 대책 수립 - 수목 식재시기 판단, 조림사업 대상지 점검 • 정보생산 및 제공: 지역별 실시간 가뭄정보 검색 시스템 구축 • 장비 및 인력: 산불 감시, 진화 인력 총원 • 홍보: 산불 캠페인 활동
도시숲	<ul style="list-style-type: none"> • 대책 및 지원 <ul style="list-style-type: none"> - 가이드라인 수립에 따른 긴급대책 점검반 편성 - 계획 수립하여 시군구에 송부 - 사전 종합 점검 - 비상 급수 지원 체계 구축 • 용수확보 <ul style="list-style-type: none"> - 급수원 파악 - 저수지 가뭄대책용 저수 • 장비 및 인력: 급수 장비 및 인력 동원, 가뭄 해소용 장비 활용 계획 수립 • 홍보: 시민참여 홍보

Table 3-12 Drought measures available for the 6-month forecast lead time.

담당 업무	대처 방안
재난재해통합	<ul style="list-style-type: none"> • 예산: 가뭄 대비 예산 관련 운영 계획 수립 • 대책 및 지원: 가뭄 대처 계획 수립 • 용수확보 <ul style="list-style-type: none"> - 관정 등 수자원 확보 대책 수립 - 농업분야 및 용수 공급(지하수 등)을 위한 사업 발주 준비 • 장비 및 인력: 비상시 대비 양수 장비 점검 • 홍보: 절수 홍보 및 가뭄 피해 방지 홍보
산불	<ul style="list-style-type: none"> • 대책 및 지원 <ul style="list-style-type: none"> - 수목 식재 시기 조절, 조림 사업 추진 계획 작성 - 산불 예방 활동 강화 • 정보생산 및 제공 <ul style="list-style-type: none"> - 지역별 실시간 가뭄정보 검색 시스템 구축 - 예년 기상정보 등 자료 수집 • 장비 및 인력: 산불 관련 장비 구입(임차 헬기 등) 결정 • 홍보: 산불 예방 언론 홍보 강화
도시숲	<ul style="list-style-type: none"> • 예산: 다음 해 예산 편성, 가뭄 기금 편성 • 대책 및 지원: 상하반기 가이드라인 및 종합대책, 실행 계획 등 수립하여 시군구에 송부 • 용수확보: 저수지에 가뭄대책용 저수

Table 3-13 Drought measures available for the 1-year forecast lead time.

담당 업무	대처 방안
재난재해통합	<ul style="list-style-type: none"> • 예산: 가뭄 기금 편성 • 대책 및 지원: 상습가뭄지역에 대한 대책 수립 및 추진 • 정보생산 및 제공 <ul style="list-style-type: none"> - 매년 기상 자료 분석 - 상습 가뭄지역 도출 • 홍보: 절수 홍보
산불	<ul style="list-style-type: none"> • 예산: 다음 해 예산 편성 • 대책 및 지원 <ul style="list-style-type: none"> - 산불방지 종합대책에 가뭄에 따른 대책 수립 - 조림사업 식재 수종 선택 • 정보생산 및 제공 <ul style="list-style-type: none"> - 지역별 실시간 가뭄정보 검색 시스템 구축 - 예년 기상정보 등 자료 수집 - 산불발생 취약지 조사

담당 업무	대처 방안
	<ul style="list-style-type: none"> • 시설: 산불방지 시설 확충 • 홍보: 가뭄 예측 정보를 산불 예방 홍보에 활용
도시숲	<ul style="list-style-type: none"> • 예산: 다음 해 예산 편성 • 대책 및 지원: 가로수 관리, 산불 예방 등 대책 및 계획 수립하고 시군구에 송부 • 용수 확보: 저수지 및 소형댐(사방댐) 조성 • 정보생산 및 제공: 가뭄 예측 지도 작성 및 보급

가뭄감시의 경우와 마찬가지로 담당 업무에 가장 도움이 될 것으로 응답한 예측 정보의 종류 및 형태를 시간해상도/공간해상도/선행예측시간/활용의 순서로 나타내었다(Table 3-14). 가뭄 정보 제공 형태에 대해서는 인터넷 대표 매체, 스마트폰, 지역별 실시간 가뭄 정보 검색 가능한 시스템 등의 의견이 있었다.

Table 3-14 Demand for drought forecast information.

담당 업무	가뭄 정보 수요
재난재해통합	<ul style="list-style-type: none"> • NA/시군구/1개월, 2개월/NA • 주별/시군구/NA/NA
산불	<ul style="list-style-type: none"> • 주별/NA/1개월, 3개월, 6개월, 12개월/산불 방지 대책 • NA/NA/1개월/NA • 주, 월별/NA/NA/NA
도시숲	<ul style="list-style-type: none"> • NA/NA/1개월/수목에 급수작업 필요여부 판단(급수차량 확보 등 수목 급수에 대한 대비) • NA/NA/3개월/저수율에 따라 가뭄대처 • NA/NA/6개월, 1년/가뭄 및 폭염의 장기예측을 통해 계획 수립 • NA/NA/3개월/가로수 유지관리 추진

3.3.3 미계측지역의 가뭄 정보 수요

관측 자료 기반의 가뭄 정보를 생산할 경우 가뭄 정보는 지점 위치에서만 가용하게 되는데, 미계측지역에 대한 가뭄 정보를 제공하는 방식으로 공간 보간과 위성 기반 자료의 활용을 제시하였다. 2차 설문조사를 통해 이 경우 미계측지역을 포함한 관찰 시도의 가뭄 정보 제공에 대하여 원하는 정보 유형을 질문하였다. 원하는 가뭄 정보의 형태로 관찰 시도에 위치한 몇 개의 관측 지점에 대한 정보로 충분한지, 시군구와 같이 행정구역 단위의 가뭄 정보가 필요한지, 공간적으로 연속적으로 분포된 격자 등의 정보가 필요한지 질문하였는데, 응답자의 60%(3명)가 시군구와 같은 행정구역

단위의 가뭄 정보를 선호하였으며 40%(2명)는 공간적으로 연속적으로 분포된 격자 등의 형태의 정보를 선호하였다. 공간적으로 연속된 격자를 선호한 응답자의 경우 격자의 해상도에 대해서는 1×1 km 및 500×500 m 이상을 선택하여 매우 고해상도의 가뭄 정보에 대한 수요를 보였다.

3.3.4 가뭄 정보의 정확성 수용

미계측지역에 대한 정보를 제공할 때 지점에 대해 가용한 관측 자료를 보간하는 경우의 정확도와 위성 자료에서 얻은 가뭄 정보의 정확도가 각각 0.1~0.7 범위로 분포할 때 두 정보의 각각의 정확도 범위에 대해 사용하고자 하는 정보를 S 또는 N으로 표시하도록 질문하였다. 그러나 질문자의 의도가 충분히 전달되지 못하여 응답자들은 N 또는 S로 하나의 답변만을 제공하였다.

응답자의 40%(2명)는 위성 기반 가뭄 정보를 활용하겠다고 답변하였으며 그 이유로 위성 자료로 연속해서 가뭄의 기간을 추측할 수 있기 때문에, 위성 자료에 온도, 기후 등이 색상으로 표현될 것 같고 가뭄 정보를 파악하는 데 도움이 될 것 같아서라고 답하였다. 한 응답자는 위성 기반 가뭄 정보(N)의 정확도가 0.4 이상인 경우, 지점에서 공간보간한 가뭄 정보(S)의 정확도가 0.6 이상인 경우 사용하겠다고 응답하였는데, 그 이유로 관측 지점의 가뭄 예측 정보보다 위성 자료로부터 가뭄 예측 정보를 더 믿는다고 답하였다. 공간보간을 통한 지수를 사용하겠다고 답한 응답자는 지역의 정보를 이용한 것이 지역별 특성이 반영되어 적합하다고 답하였다. 표본의 크기는 작으나 5명의 응답자 중 3명이 위성 자료에 기반을 둔 가뭄 정보의 활용 의사를 밝힌 점은 주목할 만한 점이다. 질문 시 의도하였던 응답자의 정확도 수용 범위에 대해서는 정보를 얻지 못하였다.

3.4 소결

가뭄정보의 시간해상도에 대하여 의사결정자들은 일별, 월별 가뭄 정보가 필요하다고 응답한 경우도 있었으나 대체적으로 주별로 가뭄감시 및 예측 정보가 제공되기를 원하였다. 공간해상도는 시군구와 같은 행정구역 단위의 가뭄 정보와 공간적으로

연속된 고해상도 격자에 대한 요구가 함께 있었으며, 가뭄예측 정보의 선행 시간은 1개월에서 1년까지 다양하게 응답하였다.

따라서 주별 시간해상도 및 1×1 km 이상 공간해상도의 가뭄 정보 생산을 위한 모델 구축이 뒤따르는 것이 바람직하나, 본 연구에서 사용한 위성 자료 중 가장 중요한 강수 자료의 공간해상도가 $0.1 \times 0.1^\circ$ (약 10×10 km; Table 1-3)로 이 역시 고해상도지만 다운스케일링 없이 1×1 km 에는 미치지 못하며 또한 미래 기간에 대해 사용한 GCM 장기예측자료의 시간해상도는 월별이므로 본 연구에서는 본 연구에서는 우선 MODIS 센서의 지표면온도 및 식생지수 자료의 해상도인 약 $0.05 \times 0.05^\circ$ 공간해상도와 월별 시간해상도의 가뭄예측모델을 구축하기로 하였다.

제2장에서 과거 가뭄사상을 통해 살펴보았듯이 가뭄은 넓은 지역에 대해 발생하는 특성이 있으므로 일정 공간해상도 이상으로는 큰 의미가 없을 수도 있다. 가뭄정보의 적절한 공간해상도를 정하기 위해서는 의사결정자들의 수요에 대한 고려와 더불어 가뭄 발생의 공간적 변이에 대한 분석이 차후 연구를 통해 병행되어야 할 것이다. 지점 관측 자료를 이용한 가뭄 정보와 위성 자료를 이용한 가뭄 자료에 대한 수요가 모두 있었다.

4. 가뭄예측 모델의 개발

고해상도 가뭄예측 모델을 개발하고 향상 가능한 예측 성능 범위를 제시하기 위해 지점 및 미세측지역에 대한 가뭄정보 생산 방법을 제시하고 예측 성능을 비교 분석하였다. 지점에 대한 정보만을 이용하는 경우, 가뭄정보는 가뭄지수 등을 통해 얻을 수 있다. 가뭄지수를 예측하는 방법에 대해서는 제1장에서 언급한대로 미래 기간에 대한 값을 기후값의 집합에서 Monte-Carlo Simulation을 통해 표본을 추출한 후 중간값으로 도출하는 방법(Lyon et al., 2012; 이하 Climatology, C방법)을 기준선으로 정하여 장기에측에서 얻는 방법(Quan et al., 2012; 이하 Long-Range Forecast, F방법)의 성능을 평가하였다. 단, 장기에측을 이용하는 경우 단일모델을 사용한 것이 아니라 6개 개별모델에 대해 각각 증가백분율을 구한 후 이를 앙상블하여 적용하였다.

미세측 지점에 대한 가뭄정보는 기계학습을 통해 수행되었는데, 기계학습은 가뭄예측의 경우 초기조건이 중요한 역할을 하므로 초기조건에 대한 정보를 보강하고 위성자료를 이용하여 공간적으로 분포된 정보를 활용하기 위해 적용되었다. 기계학습의 성능을 비교할 기준 방법론으로는 지점 정보의 공간보간을 이용하였다. 보간법은 주변의 지점 정보와 그 지점까지의 거리를 이용하여 정보를 도출하는 방법으로 전통적으로 많이 사용되어 왔다.

따라서, C방법 및 F방법을 통해 얻은 정보를 공간보간하는 방법(이하 Climatology-Interpolation, C-I방법 및 Long-Range Forecast-Interpolation, F-I방법)과 기계학습 모델링을 통해 예측하는 방법(이하 Climatology-Machine Learning, C-ML방법 및 Long-Range Forecast-Machine Learning, F-ML방법)을 비교 분석하였다. 그리고 장기에측 성능의 향상에 따른 가뭄예측의 향상 가능 범위를 도출하기 위하여 완벽한 예측을 가정한 경우(이하 Perfect Forecast-Interpolation, PF-I방법 및 Perfect Forecast-Machine Learning, PF-ML방법)와 비교하였다.

식생의 수분 함량과 관련된 가뭄 정보가 유용하게 쓰일 수 있는 가뭄의 환경피해와 관련된 부분의 의사결정자가 활용할 수 있도록 가뭄지수를 예측하지 않고 정규수분지수를 기계학습 모델링을 통해 직접 예측하는 경우(이하 Vegetation-Climatology-Machine Learning, V-C-ML 방법 및 Vegetation-Long-Range Forecast-Machine Learning, V-F-ML 방법)를 함께 살펴보았으며 이 경우에도 예측의 향상 가능 범위를

도출하기 위해 완벽한 예측을 가정한 경우와 비교하였다(이하 Vegetation-Perfect Forecast-Machine Learning, V-PF-ML방법).

4.1 지점에 대한 가뭄지수 예측

4.1.1 가뭄지수의 선정

분석을 위한 가뭄지수로는 6개월 SPI(SPI6)와 SPEI(SPEI6)를 선정하였다. SPI와 SPEI는 강수량만을 또는 강수량과 잠재 증발산량을 이용하여 산정하는 기상학적 가뭄지수인데, 6개월의 시간척도를 가지게 되면 전통적인 가뭄분류에 대해 토양수분 가뭄(농업적 가뭄) 및 수문학적 가뭄 감시에 모두 사용될 수 있다(Mo and Lyon., 2015). 본 연구에서 사용한 GCM 모델의 예측 선행시간이 1개월에서 6개월이므로 선행시간이 길어질수록 의존할 수 있는 관측 자료의 기간이 짧아지고 미래 기간이 길어져 방법론간의 차이가 점차 드러난다는 점에서 본 연구에 적합하다. Mo and Lyon(2015)도 미국의 NMME(North American MME)를 Lyon et al. (2012)의 방법론과 비교 평가하기 위해 6개월의 시간척도를 사용하였다.

4.1.2 성능 평가 기준

가뭄예측은 분류 및 회귀에 대해 수행되었다. 분류의 경우 정확도분석(Accuracy Assessment)을 통해 성능을 평가하게 되는데, 일반적으로 총 정확도(Total Accuracy), 사용자 정확도(User's Accuracy), 생산자 정확도(Producer's Accuracy)를 이용한다. 총 정확도는 전체 샘플 중에 제 클래스에 맞게 분류된 샘플 수의 비를 나타낸다(Eq. 4-1). 어떤 Class A에 대해서 사용자 정확도는 분류를 통해 Class A로 분류된 샘플 수와 그 중에 실제로 Class A에 속하는 샘플 수의 비를 나타내며(Eq. 4-2), 생산자 정확도는 실제 Class A에 속하는 샘플 수와 그 중 분류를 통해 Class A로 분류된 샘플 수의 비를 나타낸다(Eq. 4-3). 모의 자료를 예를 들어 설명하면(Table 4-1), 이 경우:

$$Total Accuracy = \frac{2 + 3 + 11 + 54 + 11 + 4 + 2}{114} = 0.76 \quad (Eq. 4-1)$$

$$User's Accuracy(SD) = \frac{3}{7} = 0.43 \quad (\text{Eq. 4-2})$$

$$Producer's Accuracy(SD) = \frac{3}{5} = 0.6 \quad (\text{Eq. 4-3})$$

사용자 정확도는 해당 클래스도 분류 되지 말았어야 하는데 분류되어서 발생한 오차에 해당하는 Error of Commission과 통하는 개념이며 생산자 정확도는 분류되어야 하는데 분류되지 않아서 발생한 오차에 해당하는 Error of Omission과 상응하는 개념이다. 본 연구의 가뭄예측에 있어 만일 심한 가뭄으로 예측하였는데 그렇게 심한 가뭄이 오지 않은 경우 발생하는 비용과, 심한 가뭄이 아니라고 예측하였는데 극심한 가뭄이 발생한 경우에 발생하는 비용을 비교하여보면 후자가 훨씬 크게 될 것이다. 따라서 Error of Omission이 Error of Commission보다 더 중요하게 평가되어야 하므로 총 정확도나 사용자 정확도 보다는 생산자 정확도를 사용하였다.

지점에 대해 가뭄지수를 예측하는 경우, 총 정확도와 가뭄 사상에 대한 정확도를 함께 살펴보았으며, 그 결과를 바탕으로 그 밖의 성능 평가에 대해서는 가뭄 사상에 대한 정확도만을 고려하였다. 가뭄 사상에 대한 정확도는 가뭄 카테고리 중 Extreme drought(ED), Severe drought(SD), Moderate drought(MD)의 세 카테고리에 대해서 생산자 정확도 개념을 적용하여 아래와 같이 산정하였다.

$$Accuracy = \frac{C_{ED} + C_{SD} + C_{MD}}{N_{ED} + N_{SD} + N_{MD}} \quad (\text{Eq. 4-4})$$

이 때, N은 실제로 각 클래스에 속하는 샘플의 수이고 C는 그 중 그 클래스로 맞게 분류된 샘플의 수이다.

회귀의 경우에는 Mean Absolute Error(MAE) 값을 이용하여 각 방법의 성능을 평가하였다. 이 경우에도 지점에 대해 가뭄지수를 예측하는 경우 전체 자료에 대한 MAE 값과 가뭄 사상에 대한 MAE를 함께 살펴보았으며 그 결과를 바탕으로 그 밖의 성능 평가에 대해서는 가뭄 사상에 대한 MAE만을 고려하였다. 분류의 생산자 정확도 개념을 적용하여, 실제 미래에 발생한 가뭄상태가 ED, SD, MD인 경우에 한하여 예측 성능을 평가하였다.

Table 4-1 An example of classification results.

	Class types determined from reference data								
	# Samples	ED	SD	MD	NN	MW	SW	EW	Total
Class types determined from the classification	ED	2	1						3
	SD	1	3	3					7
	MD		1	11	10				22
	NN			2	54	1			57
	MW				4	11	2		17
	SW					2	4		6
	EW							2	2
	Total	3	5	16	68	14	6	2	114

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{Y}_i - Y_i| \quad (\text{Eq. 4-5})$$

이 때, n은 ED, SD, MD에 속하는 표본의 개수이고, \hat{Y} 또는 \hat{Y}_i 는 추정된 값, Y 는 관측 값을 의미한다.

성능을 평가할 때 ASOS 지점의 위치에 대해 예측하여 실제 값과 비교하였고, 미계측지역에 대한 예측을 평가할 때도 역시 지점 위치에 대하여 Leave-One-Out 교차검증(cross-validation) 방식으로 평가하였다. 그러나 각 지점에 대해 개별적으로 평가하거나 특정 시군구나 유역 단위로 지점을 묶어 평가하지 않았으며 우리나라 전체 지역을 하나로 보고 평균적인 성능을 평가하였고, 시간적으로도 월별 평가를 개별적으로 수행하지 않고 목표로 삼는 월이 1월에서 12월까지 모두 포함되도록 평균적인 성능을 평가하였다.

4.2 미계측지역에 대한 가뭄지수 예측

4.2.1 공간보간법

미계측지역에 대한 가뭄지수 예측 성능을 평가하기 위해 기준 방법론으로서 공간보간을 사용하는 경우(C-I방법, F-I방법) Leave-One-Out 교차검증 방법을 적용하

여 각 지점 위치에 대하여 그 지점의 자료가 없다고 가정하고 주변 지점의 값을 이용하여 공간보간한 후에 그 지점의 값과 비교하도록 하였다. C-I방법은 C방법으로 지점에 대해 산정된 가뭄지수를 공간보간하고 F-I방법은 F방법으로 지점에 대해 산정된 가뭄지수를 공간보간하였다. 공간보간 방법으로는 표면의 총 곡률을 최소화하는 방식으로 거리에 의존하는 보간법인(Cressie, 1993) Radial Basis Function(RBF)를 이용하였다(Broomhead and Lowe, 1988). 이 방법은 Inverse Distance Weighted(IDW) 방법과 같이 자료가 있는 지점을 정확히 지나가는 Exact Interpolator로 IDW와는 다르게 주어진 자료의 최솟값과 최댓값을 벗어나는 범위를 외삽할 수 있으며 크리깅(Kriging)에 비해 생성이 쉽고 계산이 빠른 장점이 있다. 기저함수로는 일반적으로 사용되는 Multiquadric 함수를 이용하였다(multiquadric spline, Eq. 4-6).

$$\phi(r) = \sqrt{1 + (\epsilon r)^2} \quad (\text{Eq. 4-6})$$

이 때, r 는 거리이고 ϵ 은 평활화(smoothing)를 위한 매개변수이다. ϵ 값으로서는 지점 사이의 평균 거리를 이용하였다.

4.2.2 기계학습 모델링

4.2.2.1 입력 변수

기계학습을 이용하는 경우(C-ML방법, F-ML방법) 이진영(2015)에서 사용한 방법과 유사하게 제1장에서 열거한 다양한 위성 자료를 입력 변수로 사용하였다. 기계학습의 경우 변수들의 다중공선성(multicollinearity)에 영향을 크게 받지 않으므로 서로 상관관계가 높은 변수들을 함께 활용할 수 있다. 그러나 모델의 이해를 돕고 결과의 분석을 용이하게 하기 위해 상관관계가 높은 변수를 최대한 줄여 사용하였으며, 이진영(2015)에서 사용한 토양수분의 경우 실제 KoFlux에서의 관측값과의 차이가 매우 크고 변수의 상대적 중요도도 높지 않았으므로 본 연구에서는 제외하였다. 따라서 제1장에서 나열한 NDVI, NDWI, 주간지표면온도(LST_DAY), 야간지표면온도(LST_NIGHT), 강수량(PRCP), 잠재증발산량(PET)와 대규모 대기순환 기후지수인 MEI, AOI를 입력 변수로 사용하였으며 월별로 따로 분석하지 않고 전체 월을 포함시켰으므로 월 또한 추가적인 입력 변수로 사용하였다(MONTH).

NDVI나 NDWI의 경우 식생과 관련된 지수이므로 토지피복이 식생인 지점의 자료만을 이용하는 것이 유의하나, 이들 위성 자료의 공간해상도가 $0.05 \times 0.05^\circ$ 로 여러 토지피복이 혼재하게 된다. 따라서 토지피복이 식생인 지점만 따로 선별하지는 않고 전체를 사용하였다. 공간해상도가 500×500 m 와 같이 높아지는 경우는 다르게 선정하여야 할 것이다.

지점이나 지역별로 분석을 따로 하지 않고 우리나라 전체에 대해 같은 모델을 구축하였으므로 여러 지점의 자료가 한꺼번에 입력 자료로 들어가게 된다. 토지피복 등의 각 지역별 특징에 따라 같은 가뭄 상태의 변이에 대해서도 반응하는 각 변수 값의 폭이 달라지므로, Kogan(1995)이 NDVI에서 환경적인 요소를 제외시키고자 지점별로 최댓값과 최솟값을 이용하여 정규화시킨 것처럼 위성에서 도출한 NDVI, NDWI, LST_DAY, LST_NIGHT, PRCP, PET에 대해서는 각각 지점별, 월별 최댓값과 최솟값을 이용하여 정규화시켰다(Feature Scaling). 정규화시키지 않은 경우와 정규화시킨 경우의 차이를 평가하기 위해서 모델의 입력 변수로 정규화한 변수와 정규화하지 않은 변수를 모두 포함시켜 변수의 상대적 중요도를 검토하였으나 미미한 차이만 있었다.

기후지수인 MEI와 AOI의 지연 효과를 고려하고자 제2장에서 우리나라의 가뭄과 대규모 대기순환 기후지수와의 상관관계를 살핀 기존 문헌을 살펴보았다. 그러나 각 지수의 값과 실제 가뭄발생간의 발생 지연에 있어 많은 연구 문헌이 일관적인 결과를 보이지 않았다. 따라서 61개 ASOS에 대하여 평균한 SPI6, SPEI6과 MEI, AOI의 지연을 고려한 상관관계를 구하여 보았다(Table 4-2). 지연시간을 0개월에서 24개월까지 테스트하였는데 모든 경우에 피어슨상관계수가 매우 낮게 나타났다(Table 4-2). 그러나 낮은 상관관계에도 90% 신뢰도 수준에서 통계적으로 유의미한 경우를 도출하여 MEI와 AOI의 지연 효과는 1년으로 가정하였으며 그게 맞추어 모델의 입력 자료를 생성하였다.

C-ML방법과 F-ML방법의 차이가 드러나는 변수가 PRCP와 PET이다. 각 입력 변수의 값이 목표시점까지의 6개월 누적 강수량 또는 잠재증발산량의 합을 정규화한 것이므로 완벽한 예측을 가정한 관측된 위성 자료 값을 이용하여 기계학습 모델을 구축하였다. 완벽한 예측은 해당 위치에서 장기예측의 증가백분율을 이용해 구한 보정

된 예측값(Eq. 1-7) 대신에 그 위치의 관측 자료를 이용해서 산정한 가뭄지수 값을 완벽한 예측으로 가정하여 이용하는 것을 뜻하는데, 관측 자료를 이용해서 구한 가뭄지수를 그대로 사용하지는 않고 각 지점의 월별로 형성된 강수량이나 잠재증발산량의 시계열의 표준편차에 아주 작은 실수값을 곱한($2.22e-16$) 만큼의 정규분포를 가지는 작은 노이즈를 시뮬레이션하여 원래 값에 더하여 주었다.

C-ML방법과 F-ML방법의 성능을 비교하기 위해 앞서 C방법과 F방법에서 사용하였듯이 이 6개월에 대해 관측 자료가 있는 부분은 관측 자료를 합산하고 미래 기간에 대해서는 C-ML방법은 기후값에서 Monte-Carlo Simulation을 통해서, F-ML방법은 장기예측모델 결과에서 얻은 값을 사용하였다(Eq. 1-5 에서 1-7). 단, 이 때 사용하는 기후값은 지점 관측 자료의 기후값이 아니라 위성 자료에 기반을 둔 PRCP과 PET의 기후값이 된다.

Table 4-2 Statistics for lagged correlations between climate indices(MEI, AOI) and drought indices(SPI6, SPEI6).

Climate Index	Drought Index	Lag (months)	Number of samples	r	p-value	Standard Error
MEI	SPEI6	12	500	-0.08	0.066	0.057
MEI	SPEI6	13	499	-0.08	0.082	0.057
MEI	SPI6	12	500	-0.08	0.088	0.058
MEI	SPI6	13	499	-0.07	0.096	0.058
AOI	SPEI6	13	499	0.14	0.002	0.058
AOI	SPEI6	12	500	0.10	0.025	0.058
AOI	SPEI6	14	498	0.09	0.035	0.058
AOI	SPEI6	2	507	-0.09	0.047	0.057
AOI	SPI6	13	499	0.14	0.002	0.059
AOI	SPI6	12	500	0.09	0.049	0.059
AOI	SPI6	14	498	0.09	0.057	0.060

4.2.2.2 모델의 매개변수

본 연구에서는 Decision Tree(DT모델), Random Forest(RF모델), Extremely Randomized Trees(ET모델)의 세 가지 기계학습 방법론을 사용하였는데, 각각의 중

요 매개변수 값에 따라 결과가 개선될 수 있다. DT모델의 경우에는 트리의 최대심도(maximum depth, MAX_DEPTH)와 최종 잎노드에 최소한 배정되어야 하는 샘플수(minimum number of samples for each leaf, MIN_SAMPLES)의 설정이 중요하다. 예를 들어, Figure 1-6와 같은 경우는 MAX_DEPTH가 3으로 설정된 경우로 시작 잎노드로부터 아래로 세 층까지 전개되었으며 MIN_SAMPLES는 100으로 설정되어 최종 잎노드에서 최소한 100개의 샘플을 배정받도록 하였다. MAX_DEPTH를 지정하지 않으면 트리는 최대한으로 성장하게 되는데 이렇게 최대한 성장한 트리의 경우 훈련 자료에 과적합(overfitting)되는 경우가 있어 새로운 자료를 가지고 검증하는 경우 예측을 잘 수행하지 못할 수 있다. 이 두 매개변수는 결과적으로 유사한 작용을 하는데, MAX_DEPTH를 제한하는 경우에도 트리의 성장을 제한할 수 있지만 MIN_SAMPLES를 어느 정도 크게 잡으면 이를 만족시키기 위해 또한 트리의 성장이 제한되기 때문이다. 따라서 본 연구에서는 DT모델의 경우 MAX_DEPTH의 변화만을 살펴보았다.

RF모델과 ET모델의 경우에는 트리의 최대심도(maximum depth, MAX_DEPTH)와 트리의 개수(NUM_TREES)가 중요하다. RF모델과 ET모델의 경우 DT모델과는 다르게 여러 개의 트리를 사용하여 앙상블 방식으로 결과값을 얻는다. 따라서 DT모델의 경우에는 트리를 최대한 성장시켰을 경우 과적합이 일어나는 경우가 많지만 RF모델과 ET모델은 각각의 트리를 최대 성장시켜도 여러 트리를 각각 상관관계가 적도록 추출된 입력 벡터의 부분 집합을 통해 훈련하므로 과적합을 방지할 수 있다. 그러나 여전히 MAX_DEPTH에 영향을 받을 수 있다. 트리의 개수는 많으면 많을수록 과적합을 피해갈 수 있으나, 모델의 구동시간이 길어지게 된다.

DT모델

DT모델의 경우 MAX_DEPTH의 값을 [3, 5, 10, 완전히 성장시키는 경우]에 대해서 변화시키며 가장 나은 예측 성능을 보이는 매개변수의 조합을 찾았다.

분류의 경우 1개월에서 6개월까지의 모든 선행시간에 대해 SPI6의 경우 트리를 완전히 성장시킨 경우에 분류 정확도가 가장 뛰어났으며 SPEI6의 경우에는 선행시간 1개월과 6개월의 경우에는 5개의 MAX_DEPTH가, 그 밖의 선행시간에서는 트리를

완전히 성장시킨 경우가 분류 정확도가 가장 뛰어났다. 트리를 완전히 성장시킨 경우를 보통 15-20의 심도로 자라났다. DT모델에서 트리를 완전히 성장시키는 경우 과적합의 우려가 있다고 언급하였는데, 이 경우에는 교차검증을 했음에도 이러한 결과가 나타나 트리를 완전히 성장시키는 경우에도 성장을 제한하는 경우보다 대부분 모델의 성능이 뛰어났음을 알 수 있다. 각 MAX_DEPTH, 선행시간별로 분류 정확도를 비교한 결과를 Figure 4-1에 나타내었다.

회귀의 경우는 MAX_DEPTH를 3개 또는 5개까지만 트리를 성장시킨 경우 가장 작은 오차(MAE)를 보였다. SPEI6의 경우 모든 선행시간에서 5개의 MAX_DEPTH가 가장 우수한 결과를 생산하였으며 SPI6의 경우 선행시간 1개월과 2개월 및 5개월의 경우는 3개, 나머지 경우는 5개의 MAX_DEPTH의 결과가 좋았다. MAX_DEPTH가 5개인 경우에 대하여 각 선행시간별로 회귀의 MAE 값을 비교한 결과를 Figure 4-2에 나타내었다.

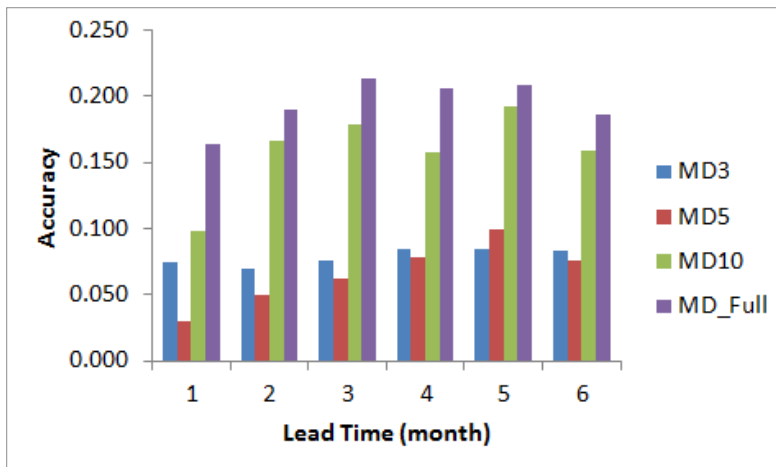


Figure 4-1 Classification accuracy of DT model for each MAX_DEPTH and lead time (based on SPI6 forecast).

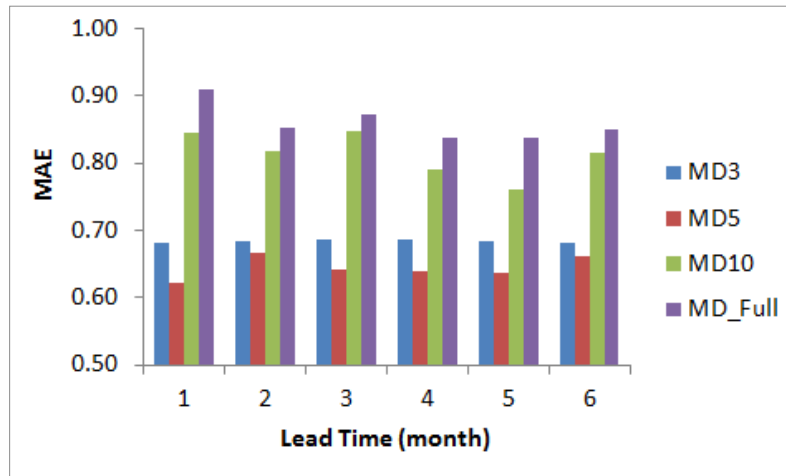


Figure 4-2 Regression MAE of DT model for each lead time (based on SPEI6 forecast).

RF모델과 ET모델

RF모델과 ET모델의 경우에는 DT모델과 같은 범위의 MAX_DEPTH와 [10개, 50개, 100개, 200개]의 트리의 개수인 NUM_TREES에 대해 민감도 분석을 수행하였다.

분류를 수행하는 경우 거의 모든 경우에(1개월에서 6개월까지의 선행시간, SPI6, SPEI6의 두 가뭄지수의 12가지 중 10가지의 경우) 10개의 트리를 사용하는 경우 가장 분류 정확도가 높았으며 MAX_DEPTH는 제한을 두지 않고 트리를 완전히 성장시키는 경우 성능이 가장 뛰어난 것으로 나타났다. 즉 DT모델의 경우와 같이 분류를 위해서는 트리를 크게 성장시키는 경우가 모델의 성능이 우수하게 나타났으며, 트리의 개수를 많이 두는 것보다 적게 유지하는 것이 정확도가 상대적으로 높게 나타났다. RF모델에 대하여 트리를 완전히 성장시킨 경우, 각 NUM_TREES, 선행시간별로 분류 정확도를 비교한 결과를 Figure 4-3에 나타내었다.

회귀의 경우에도 RF모델의 경우에는 DT모델에서 살펴본 것과 비슷하게 MAX_DEPTH를 5로 정한 경우가 대부분 분류 성능이 가장 우수하였다. NUM_TREES는 다양하게 나타났는데, 이에 따른 모델 성능의 차이는 크지 않았다. 트리의 개수가 작아도 큰 경우와 거의 같은 정도의 오차를 나타내었다. MAX_DEPTH가 5로 고정된 경우 각 NUM_TREES, 선행시간별로 MAE 값을 비교한 결과를 Figure 4-4에 나타내었다.

ET모델의 경우에는 RF모델과 약간 다르게 나타났는데, MAX_DEPTH 값 10이 가장 적합한 경우로 나타났으며 이때 NUM_TREES는 RF모델과 마찬가지로 다양하게 나타났다. MAX_DEPTH가 10으로 고정된 경우 각 NUM_TREES, 선행시간별로 MAE 값을 비교한 결과를 Figure 4-5에 나타내었다. 이 경우 트리의 개수가 10개인 경우의 오차가 약간 높았지만 큰 차이는 없었다. 각각의 경우에 대해 가장 나은 성능을 보이는 매개변수의 조합을 이용하여 모델을 구축하였다.

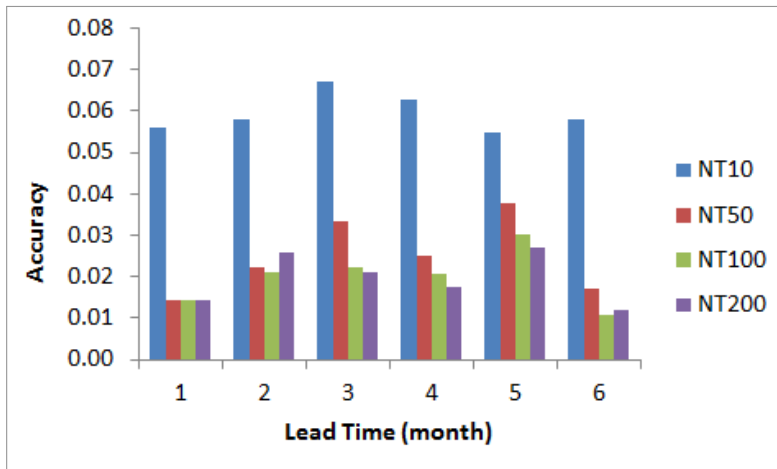


Figure 4-3 Classification accuracy of RF model for each NUM_TREES and lead time with fixed fully developed trees (based on SPI6 forecast).

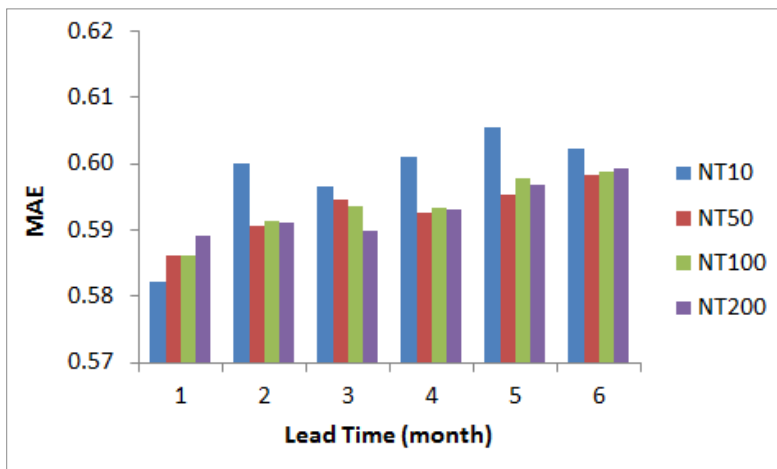


Figure 4-4 Regression MAE of RF model for each NUM_TREES and lead time with fixed MAX_DEPTH = 5 (based on SPEI6 forecast).

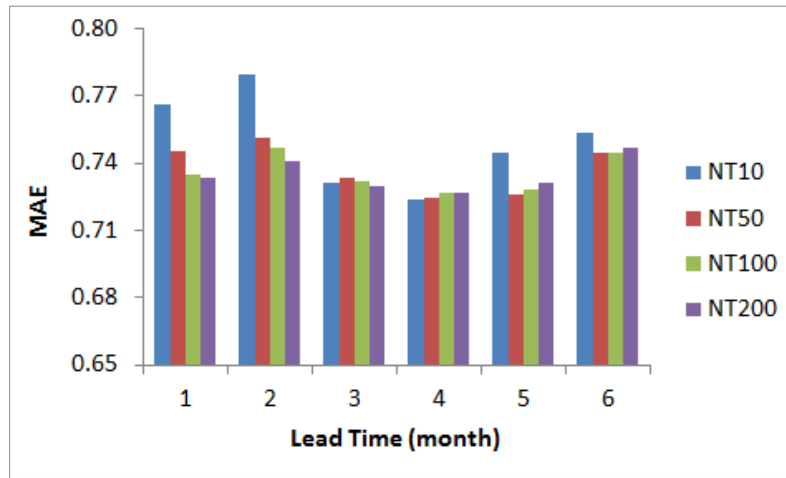


Figure 4-5 Regression MAE of ET model for each NUM_TREES and lead time with fixed MAX_DEPTH = 10 (based on SPI6 forecast).

4.2.2.3 교차검증

사용되는 위성 자료의 기간이 짧고 분류 및 회귀 모델의 목적이 가뭄예측이므로 각 기간별 입력 벡터의 특징이 매우 다를 수 있다. 따라서 과적합(overfitting)을 피하기 위해 각 연도별로 Leave-One-Year-Out 교차검증을 수행하였다. SPI6, SPEI6를 예측하는 모델과 마찬가지로 NDWI를 예측하는 모델도 Leave-One-Year-Out 교차검증을 수행하였으며 그 결과를 샘플 하나씩을 제외하는 Leave-One-Out 교차검증, 전체를 무작위로 섞은 후 k개의 임의의 부분집합으로 나누는 k-fold with shuffle 교차검증(2-fold와 10-fold 사용), 그리고 공간적인 차이를 비교하기 위해 각각의 지점을 하나씩 제외시킨 Leave-One-Station-Out 교차검증을 각각 수행하여 비교하였다 (자료 미제시). 결과적으로 가뭄지수를 예측하는 경우에는 모든 경우 Leave-One-Year-Out 교차검증보다 10-fold 교차검증, Leave-One-Station-Out 교차검증의 경우가 모델의 예측 성능이 매우 뛰어났는데(자료 미제시), 이는 더더욱 Leave-One-Year-Out 을 제외하고는 모델이 과적합되어 지금까지의 자료를 가지고 모델을 훈련시키고 가뭄 지수를 예측하기 위해 아직 위성 자료가 가용한 기간에 대해 모델이 훈련되지 못한 패턴이 앞으로 나타난다면 모델의 성능이 떨어질 가능성이 있다는 것을 의미한다.

NDWI 예측의 경우도 대부분의 경우 Leave-One-Year-Out 교차검증의 예측 성

능이 가장 낮았지만 가뭄지수 예측처럼 정확도 차이가 크지는 않았다(자료 미제시). 즉, 모델이 훈련되지 못한 미래의 자료에 대해 조금 더 안정적이라고 할 수 있다.

4.3 정규수분지수 예측

4.3.1 기계학습 모델링

4.3.1.1 입력 변수

NDWI를 예측하기 위한 모델에서는 SPI6나 SPEI6 예측의 경우와는 달리 기계학습 모델의 입력 변수 중 PRCP와 PET의 누적 기간을 6개월로 정하지 않았다. NDWI의 초기값 이전의 강수량과 잠재증발산량은 이미 NDWI에 영향을 주었다고 가정하고, 초기값이 주어지고 난 후의 기간에 대해서만 누적하였다. 이 때 선행시간 1개월의 경우에는 기후값 또는 장기예측 자료의 누적기간이 1개월, 선행시간 2개월에 대해서는 누적기간이 2개월, 같은 방식으로 선행시간이 6개월인 경우에는 누적 기간이 6개월이 된다.

SPI6, SPEI6 예측의 경우와 마찬가지로 NDVI, NDWI, LST_DAY, LST_NIGHT, PRCP, PET, MEI, AOI, MONTH를 입력 변수로 사용하였다. 이 경우에도 위성에서 도출한 변수들에 대해서 각각 위치별, 월별 최댓값과 최솟값을 이용하여 정규화시켜 정규화시키지 않은 경우와 비교하였다. 입력 변수로 정규화한 변수와 정규화하지 않은 변수를 모두 포함시킨 후 변수들의 상대적 중요도를 검토하였는데, 초기값을 입력 변수로 사용하기도 하면서 목표변수이기도 한 NDWI를 제외하고 다른 변수들에 대해서는 SPI6, SPEI6 예측의 경우와는 다르게 정규화하지 않은 변수들의 상대적 중요도가 더 크게 나타났다. MAX_DEPTH의 제한을 두지 않고 완전히 성장시킨 트리들을 이용한 RF모델을 통한 예를 Figure 4-6에서 살펴보았다. 앞서 언급한대로 Leave-One-Year-Out 교차검증을 이용하였으므로 제외시킨 각 연도별로 구해진 변수의 상대적 중요도를 전체에 대해 평균하였다. 여기서 나타낸 중요도 비율(Importance Ratio)은 각 변수에 대해 정규화시킨 경우의 중요도를 정규화시키지 않은 경우의 중요도로 나눈 값으로 나타낼 수 있으므로 1보다 작은 경우 정규화시키지 않은 변수의 중요도가 더 높음을 나타낸다. 선행시간 5개월의 PRCP과 선행시간 6개월의 PET를 제외하면

모두 정규화시키지 않은 변수의 상대적 중요도가 높았다. 따라서 NDWI 예측에 대해서는 변수들의 값을 정규화시키지 않았다.

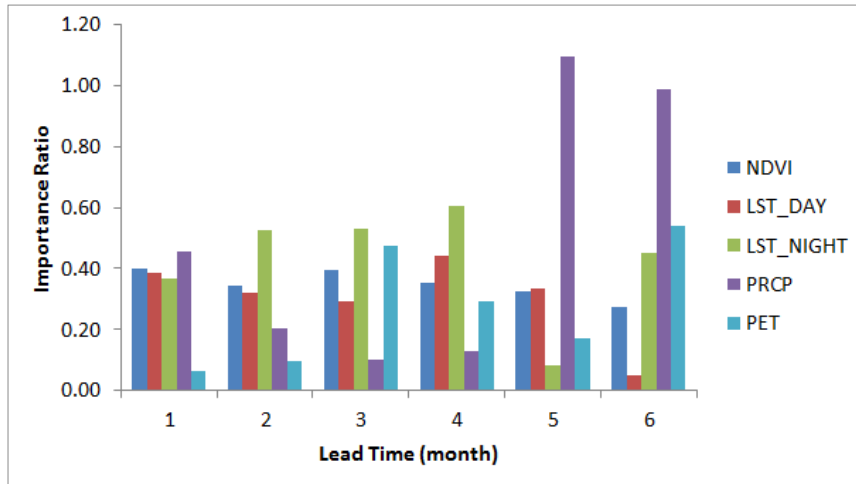


Figure 4-6 Importance ratio (normalized variable / raw variable) for each variable and lead time.

4.3.1.2 모델의 매개변수

NDWI 예측의 경우에도 SPI6, SPEI6 예측과 마찬가지로 기계학습 모델의 매개변수에 대한 민감도 분석을 수행하였다.

DT모델

DT모델의 경우 역시 MAX_DEPTH의 값을 [3, 5, 10, 완전히 성장시키는 경우]에 대해서 변화시키며 가장 나은 예측 성능을 보이는 매개변수의 조합을 찾았다.

분류의 경우 적합한 매개변수는 선행시간에 따라 다르게 나타났는데, 선행시간이 짧은 경우는 MAX_DEPTH를 10으로 트리의 성장을 제한한 경우가 분류 성능이 높은 것으로 나타났으나 선행시간이 긴 경우는 최대로 트리를 성장시킨 경우가 우수하게 나타났다. 이 경우 MAX_DEPTH와 선행시간에 따른 분류 정확도를 살펴보면(Figure 4-7) 일반적으로 MAX_DEPTH가 증가할수록 모델의 성능이 우수함을 알 수 있다. 따라서 트리를 많이 성장시키는 것이 적합할 것으로 여겨진다.

회귀의 경우에는 적합한 MAX_DEPTH가 다양하게 나타났으나 많은 경우 MAX_DEPTH가 10인 경우가 적합하였다(Figure 4-8).

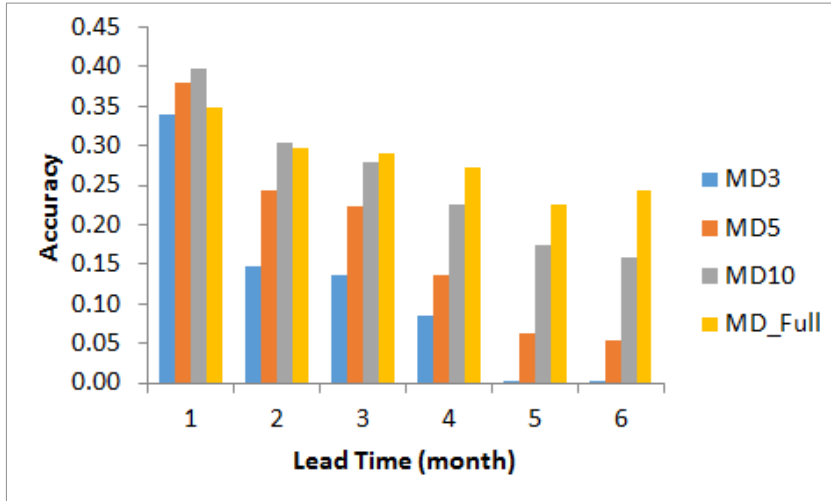


Figure 4-7 Classification accuracy of DT model for each MAX_DEPTH and lead time (based on NDWI forecast).

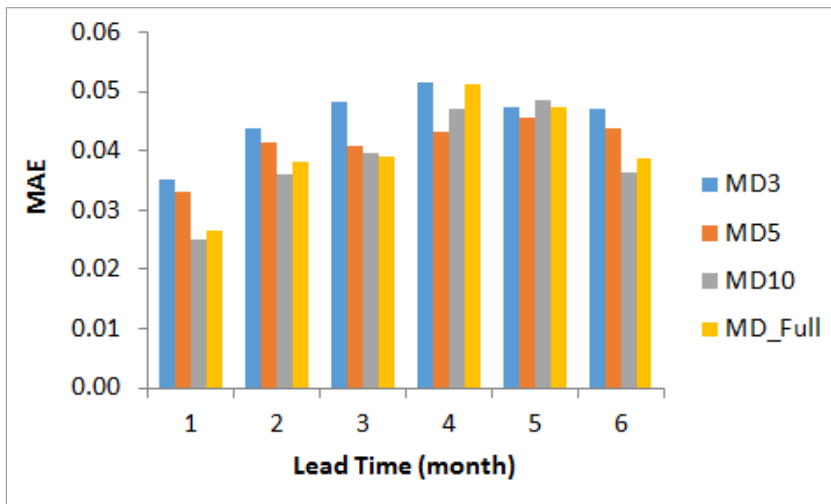


Figure 4-8 Regression MAE of DT model for each MAX_DEPTH and lead time (based on NDWI forecast).

RF모델과 ET모델

RF모델과 ET모델의 경우에는 DT모델과 같은 범위의 MAX_DEPTH와 [10개, 50개, 100개, 200개]의 트리의 개수인 NUM_TREES에 대해 민감도 분석을 수행하였다.

분류의 경우 MAX_DEPTH는 최대로 트리를 성장시킨 경우가 분류 정확도가 높은 것으로 나타났다. 적합한 트리의 개수 NUM_TREES는 다양하게 나타났는데 성능에 큰 차이가 없었다(Figure 4-9). ET모델의 경우에는 5개월 및 6개월 선행시간에서는 NUM_TREES가 50개 이상인 경우 모델 성능에 큰 차이가 없었으나 그보다 짧은 선행시간에는 그보다 큰 트리 개수가 더 뛰어난 분류 성능을 보이는 경우가 있었다(자료 미제시).

회귀의 경우 RF모델은 대부분의 경우 MAX_DEPTH가 10 이거나 최대로 트리를 성장시키는 경우가 가장 오차가 작게 나타났고, ET모델은 트리를 최대로 성장시키는 경우 오차가 가장 작았다. 이에 따른 트리의 개수는 다양하게 나타났는데 큰 차이는 없었으나 테스트한 최대수인 200개인 경우가 가장 작거나 비등하게 작은 오차를 보였다(Figure 4-10). 각각의 경우에 대하여 가장 우수한 성능을 보이는 매개변수 값을 이용하였다.

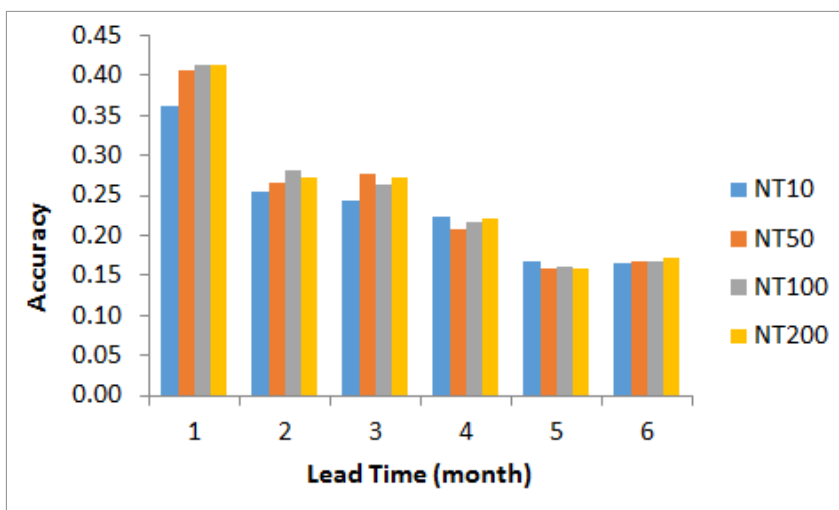


Figure 4-9 Regression MAE of RF model for each NUM_TREES and lead time with fully developed trees (based on NDWI forecast).

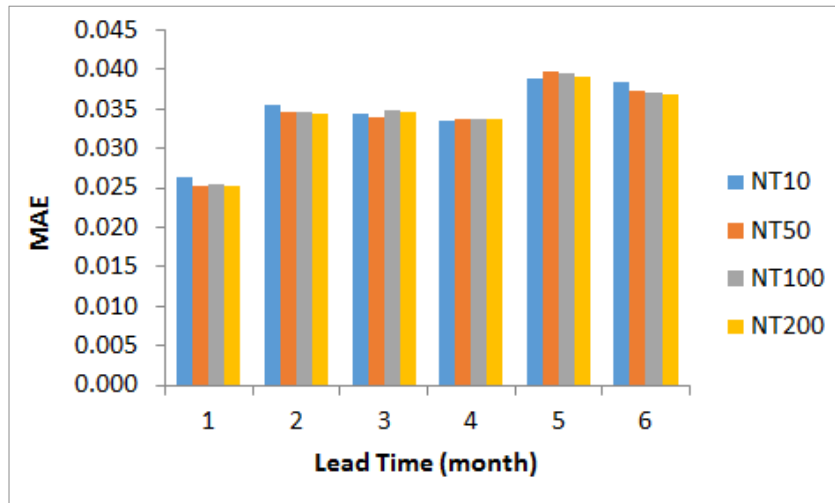


Figure 4-10 Regression MAE of ET model for each MAX_DEPTH and lead time with fully developed trees (based on NDWI forecast).

5. 모델 간 비교 결과

5.1 지점에 대한 가뭄지수 예측 결과

5.1.1 분류 정확도

전체 클래스를 모두 포함한 총 정확도(Eq. 4-1)를 비교한 결과 1-6개월까지의 선행시간에 대하여 C방법과 F방법 사이에 예측 성능의 차이가 거의 없었다(Figures 5-1a, b). 그러나 ED, SD, MD의 가뭄사상만을 고려한 분류 정확도(Eq. 4-4)를 비교한 결과 1-6개월까지의 선행시간에 대하여 C방법의 정확도가 F방법보다 우수하게 나타났다(Figures 5-1c, d; Table 5-1). C방법의 경우 선행시간이 6개월이 되는 경우 정확도가 0으로 떨어졌으며 F방법의 경우에는 선행시간이 5개월이 되었을 때 6개월의 누적기간 중에서 관측 자료가 한 달을 차지하고 있음에도 불구하고 정확도가 0에 가깝도록 수렴하였다(Figures 5-1c, d; Table 5-1).

결과적으로 F방법의 경우 선행시간 1개월에서 4개월에 대해 나타난 분류 정확도는 6개월의 기간 동안에 포함되어 있는 관측 자료로 인한 메모리에 의한 것으로 장기예측으로 인한 예측력은 없는 것으로 보이며, 기준선으로 지정한 기후값을 이용한 C방법보다 정확도가 낮은 것으로 나타났다. 두 가뭄지수의 경우 큰 차이가 없었다. 이후의 결과에서는 ED, SD, MD의 가뭄사상만을 고려한 분류의 정확도를 살펴보았다.

5.1.2 회귀 성능

회귀의 경우 가뭄 상태와 상관없이 전체 사상을 비교·분석했을 때 짧은 선행시간에서 MAE의 큰 차이는 없으며 선행시간이 길어질수록 F방법의 오차가 약간 작게 나타났다(Figures 5-2a, b; Table 5-1). 두 가뭄지수를 비교한 경우 큰 차이는 존재하지 않았다(Figures 5-2a, b; Table 5-1). 가뭄사상만을 고려하였을 경우에는 분류의 경우와 같은 결과를 얻었는데, 즉 F방법의 오차가 C방법보다 크게 나타났으며 선행시간이 길수록 오차의 차이도 커졌다(Figures 5-2c, d; Table 5-1). 이후의 결과에서는 가뭄사상만을 고려한 회귀 성능을 살펴보았다.

종합적으로 지점에서 C방법과 F방법의 가뭄지수 예측 성능을 비교한 결과 분류와 회귀에 있어 모두 C방법이 우수하였다. 그리고 앞서 언급하였듯이 선행시간이 길어질수록 예측 성능이 크게 떨어졌으며 F방법의 낮은 성능에서 알 수 있듯이 장기예측자료가 예측에 기여하는 바가 거의 없는 것으로 나타났다.

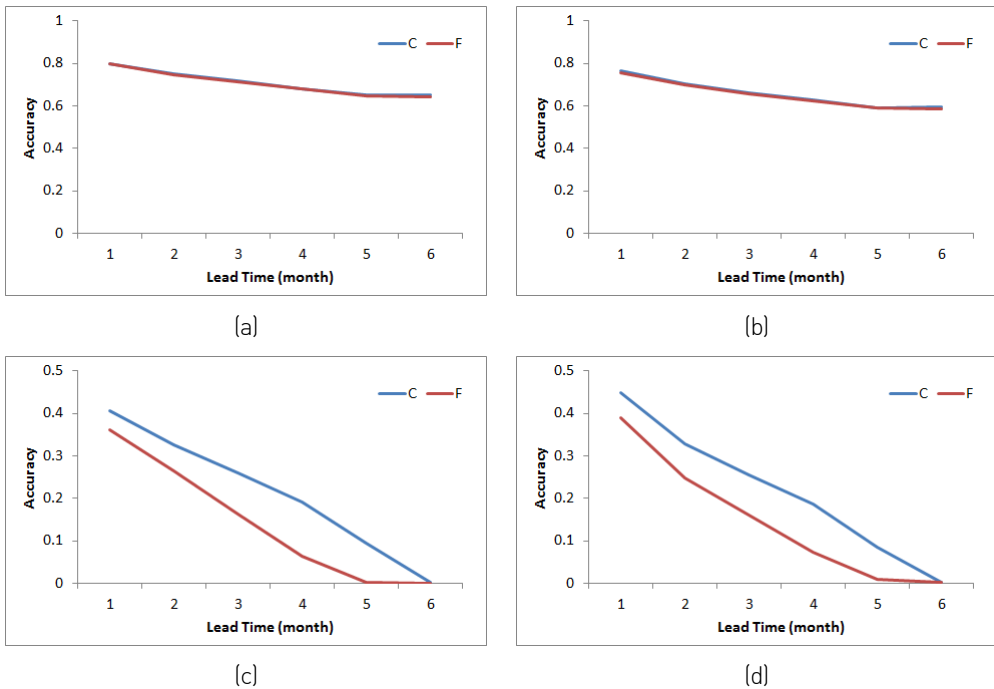


Figure 5-1 Classification accuracy of C and F methods based on (a) all events for SPI6, (b) all events for SPEI6, (c) drought events only for SPI6, and (d) drought events only for SPEI6.

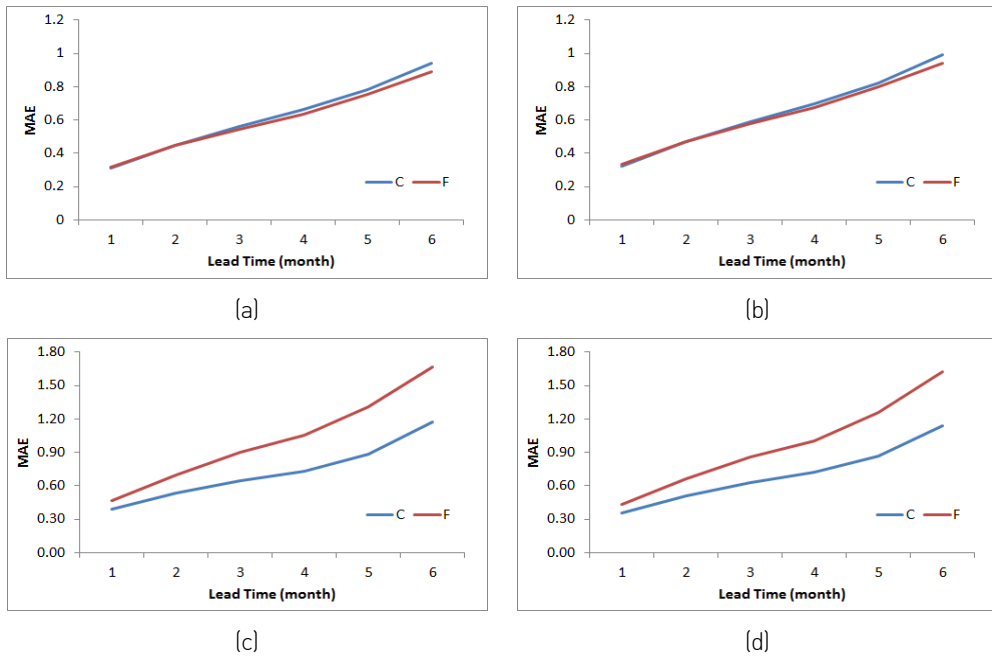


Figure 5-2 Regression MAE of C and F methods based on (a) all events for SPI6, (b) all events for SPEI6, (c) drought events only for SPI6, and (d) drought events only for SPEI6.

Table 5-1 Model performance measures for gauges considering only drought events of ED, SD, and MD.

	LT	Accuracy		MAE	
		C	F	C	F
SPI6	1	0.41	0.36	0.39	0.47
	2	0.33	0.26	0.53	0.70
	3	0.26	0.16	0.64	0.90
	4	0.19	0.06	0.73	1.05
	5	0.09	0.00	0.89	1.31
	6	0.00	0.00	1.17	1.67
SPEI6	1	0.45	0.39	0.36	0.43
	2	0.33	0.25	0.51	0.66
	3	0.26	0.16	0.63	0.86
	4	0.19	0.07	0.72	1.00
	5	0.08	0.01	0.87	1.26
	6	0.00	0.00	1.14	1.62

5.2 미계측지역에 대한 가뭄지수 예측 결과

5.2.1 분류 정확도

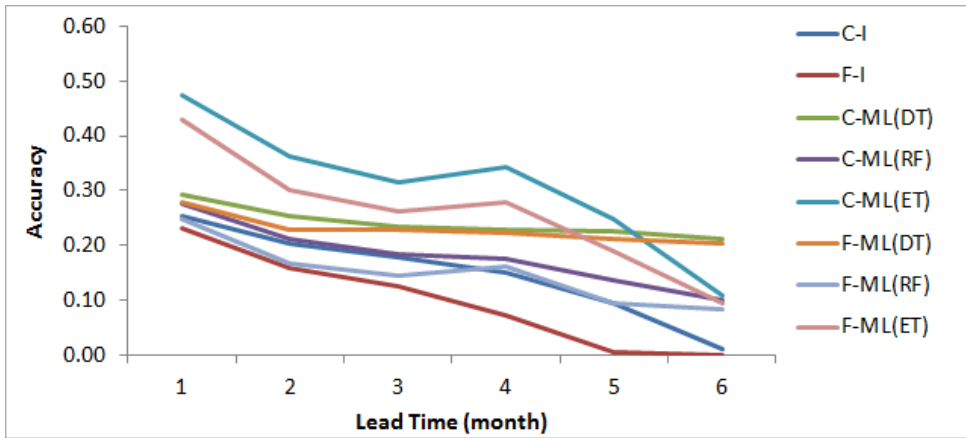
기계학습 모델을 이용한 가뭄 카테고리의 분류의 경우, 직접 분류 모델을 구축하여 분류를 수행할 수도 있고 회귀 모델을 구축하여 회귀를 수행한 다음 가뭄 카테고리(Table 1-1)에 맞게 분류할 수 있다. 본 연구에서 DT, RF, ET모델을 이용한 C-ML방법과 F-ML방법의 경우 모두 후자의 경우가 분류 정확도가 높았다. 따라서 본 장에서는 회귀를 수행한 다음 분류를 수행한 경우의 분류 정확도만을 살펴보았다. 물론 이 경우는 회귀를 수행한 다음 분류를 수행했을 때의 정확도가 가장 높도록 매 개변수가 선정되었다.

분류 정확도를 C-I방법, F-I방법, C-ML방법(DT, RF, ET모델), F-ML방법(DT, RF, ET모델)을 적용한 경우에 대해 비교하였다. 기계학습을 이용한 경우 중 높은 분류 정확도를 보인 모델의 결과는 보간법을 이용한 경우보다 모든 선행시간에서 대략 0.14-0.22 정도 높았다(Table 5-2). 대부분 기후값을 이용한 경우가 장기예측 자료를 이용한 경우보다 우수하였으며 결과적으로 C-ML방법의 성능이 가장 우수하였다(Figure 5-3, Table 5-2).

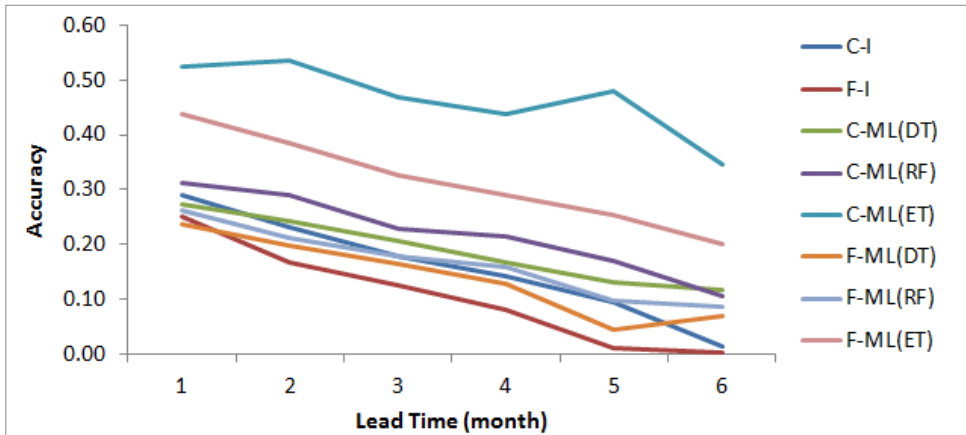
기계학습 모델의 성능간 차가 나타났는데, C-ML방법에서 SPI6 예측의 경우 대부분의 선행시간에서는 ET모델이 우수하였으며 6개월 선행시간에는 DT모델이 높은 분류 정확도를 나타냈다(Figure 5-3a, Table 5-2). SPEI6 예측의 경우에는 모두 ET 모델이 가장 좋은 결과를 보였다(Figure 5-3b, Table 5-2).

5.2.2 회귀 성능

보간법을 이용한 경우 분류와 마찬가지로 기후값을 이용한 C-I방법이 F-I방법보다 작은 오차를 보였으며, 기계학습을 이용한 경우가 보간법을 이용한 경우보다 우수한 성능을 나타냈다(Figure 5-4, Table 5-2). 두 가뭄지수 예측의 경우 대부분 C-ML방법(RF모델)의 성능이 가장 우수하였다(Figure 5-4, Table 5-2).

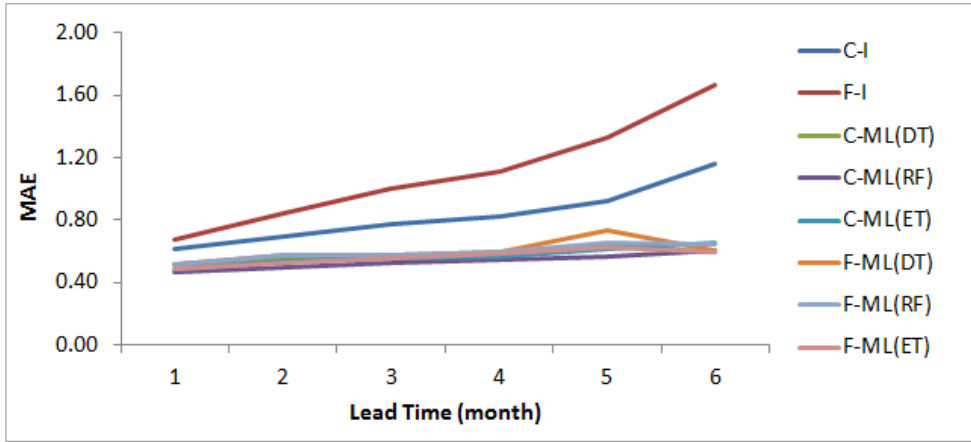


(a)

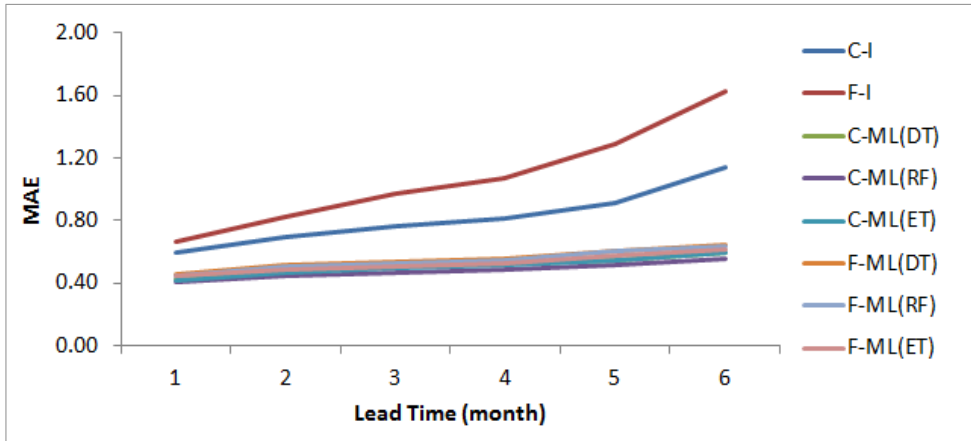


(b)

Figure 5-3 Classification accuracy based on C-I and F-I methods, C-ML and F-ML methods using the DT, RF, and ET models respectively for (a) SPI6 and (b) SPEI6.



(a)



(b)

Figure 5-4 Regression MAE based on C-I and F-I methods, C-ML and F-ML methods using the DT, RF, and ET models respectively for (a) SPI6 and (b) SPEI6.

Table 5-2 Model performance measures for ungauged locations.

Accuracy									
	LT	C-I	F-I	C-ML (DT)	C-ML (RF)	C-ML (ET)	F-ML (DT)	F-ML (RF)	F-ML (ET)
SPI6	1	0.25	0.23	0.29	0.28	0.47	0.28	0.25	0.43
	2	0.20	0.16	0.25	0.21	0.36	0.23	0.17	0.30
	3	0.18	0.12	0.23	0.18	0.32	0.23	0.14	0.26
	4	0.15	0.07	0.23	0.17	0.34	0.22	0.16	0.28

Accuracy									
	LT	C-I	F-I	C-ML (DT)	C-ML (RF)	C-ML (ET)	F-ML (DT)	F-ML (RF)	F-ML (ET)
	5	0.10	0.01	0.23	0.14	0.25	0.21	0.09	0.19
	6	0.01	0.00	0.21	0.10	0.11	0.20	0.08	0.09
SPEI6	1	0.29	0.25	0.27	0.31	0.52	0.24	0.26	0.44
	2	0.23	0.17	0.24	0.29	0.53	0.20	0.21	0.39
	3	0.18	0.12	0.21	0.23	0.47	0.16	0.18	0.33
	4	0.14	0.08	0.17	0.22	0.44	0.13	0.16	0.29
	5	0.10	0.01	0.13	0.17	0.48	0.04	0.10	0.25
	6	0.01	0.00	0.12	0.11	0.35	0.07	0.09	0.20
MAE									
	LT	C-I	F-I	C-ML (DT)	C-ML (RF)	C-ML (ET)	F-ML (DT)	F-ML (RF)	F-ML (ET)
SPI6	1	0.61	0.67	0.51	0.47	0.49	0.52	0.51	0.49
	2	0.69	0.84	0.55	0.50	0.54	0.58	0.57	0.52
	3	0.77	1.00	0.56	0.52	0.55	0.58	0.58	0.56
	4	0.82	1.11	0.57	0.54	0.57	0.59	0.59	0.58
	5	0.93	1.33	0.63	0.57	0.61	0.74	0.65	0.62
	6	1.16	1.67	0.59	0.60	0.66	0.60	0.65	0.60
SPEI6	1	0.60	0.66	0.43	0.41	0.42	0.45	0.44	0.45
	2	0.69	0.82	0.48	0.44	0.46	0.52	0.50	0.48
	3	0.77	0.97	0.50	0.47	0.49	0.53	0.53	0.51
	4	0.81	1.07	0.52	0.48	0.51	0.55	0.55	0.52
	5	0.92	1.29	0.54	0.51	0.54	0.61	0.60	0.57
	6	1.14	1.62	0.60	0.56	0.60	0.64	0.64	0.62

5.2.3 향상 가능한 범위

미계측지역에서 두 가뭄지수 예측성능을 비교한 앞의 결과에서 기후값을 이용하여 기계학습 모델링을 적용한 C-ML방법이 거의 모든 경우 가장 우수한 성능을 보였다. 기후값을 이용한 경우를 기준선으로 잡았으나 현재의 장기예측 성능에서는 가뭄예측의 성능이 이보다 오히려 낮았다. 하지만 장기예측의 성능이 향상된다면 가뭄예측 또한 향상될 것이므로 어느 시점에서 기후값을 이용한 기준선보다 높은 성능을 보일 것이다.

완벽한 예측을 가정한 PF-I방법과 PF-ML방법의 분류 정확도 및 회귀 오차를 F-I방법 및 F-ML방법과 비교하여 항상 가능한 범위를 도출하였다. 이 비교는 미계측지역을 위한 것이며 지점에서 완벽한 예측이 있다면 정확도는 1에 수렴하고 오차는 0에 수렴할 것이므로 고려하지 않았다.

분류 정확도는 보간법을 이용한 PF-I방법의 경우 대략 0.34-0.38 정도의 값을 가지는 것으로 나타났다(Figures 5-5a, b). 즉, 장기에측의 예측 성능이 완벽해 진다고 하더라도 미계측지역에 대한 분류 정확도는 최대가 0.4 이하에 머무른다는 의미이다. F-I방법의 성능은 기준선인 C-I방법보다 낮았으나 PF-I방법의 성능은 C-I방법보다 우수하므로 예측 성능이 증가한다면 장기에측을 활용하는 경우의 가뭄예측 성능이 더 우수해질 수 있음을 알 수 있다. 기계학습 모델링을 이용한 PF-ML방법은 0.39-0.56에 이르는 더욱 높은 정확도를 보였는데, SPI6와 SPEI6 예측의 경우 모두 ET모델이 가장 우수한 성능을 보였다(Figure 5-5).

회귀의 경우 PF-I방법의 오차가 완벽한 예측으로 인해 0.47-0.53 수준으로 가장 낮아졌으며 이는 기준선인 C-I방법과 비교하였을 때 매우 작았다. 기계학습을 이용한 PF-ML방법의 경우는 이보다 더 낮은 0.35-0.42의 오차 수준을 보였으며 특히 RF모델과 ET모델의 성능이 가장 우수했다(Figure 5-6).

결과적으로 분류나 회귀 모두 장기에측의 성능이 향상되는 경우 F-I, F-ML방법의 성능이 PF-I, PF-ML방법에 가까워짐에 따라 가뭄예측의 성능이 C-I, C-ML방법을 능가하게 될 것으로 예상된다.

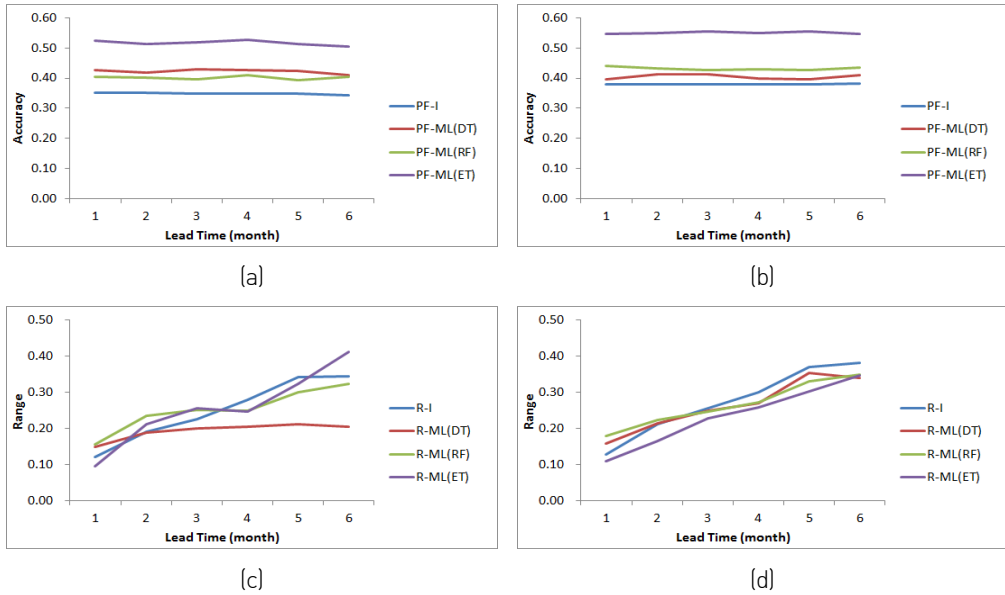


Figure 5-5 Classification accuracy based on perfect forecast for (a) SPI6 and (b) SPEI6 and the improvable ranges obtained as the difference between F-I and PF-I or between F-ML and PF-ML for (c) SPI6 and (d) SPEI6.

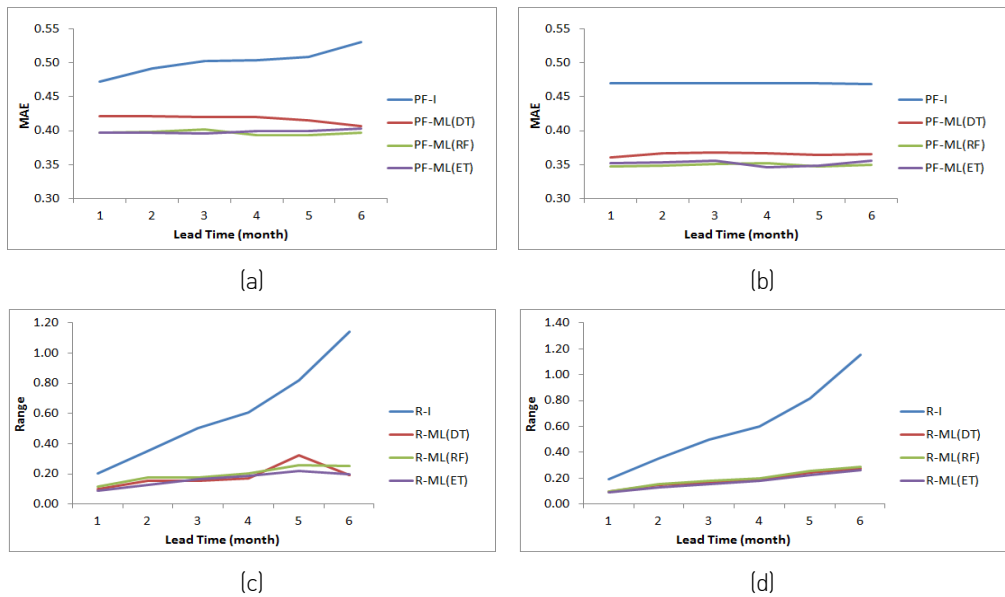
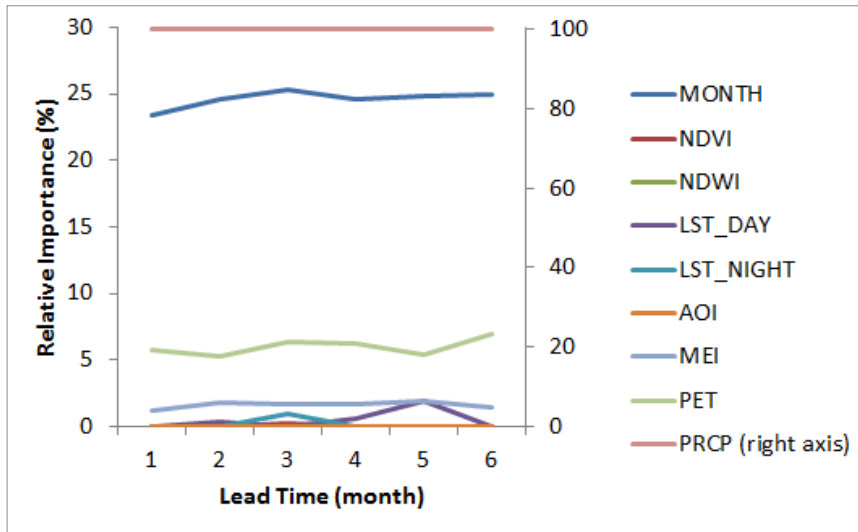


Figure 5-6 Regression MAE based on perfect forecast for (a) SPI6 and (b) SPEI6 and the improvable ranges obtained as the difference between F-I and PF-I or between F-ML and PF-ML for (c) SPI6 and (d) SPEI6.

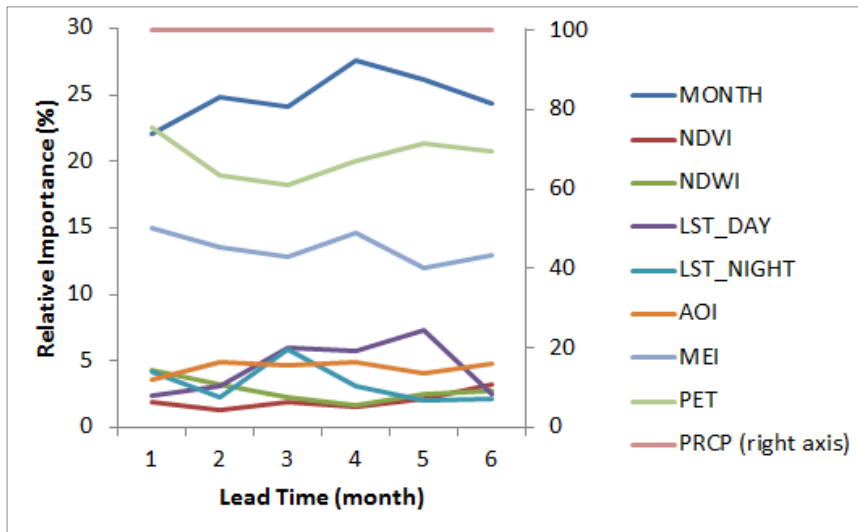
5.2.4 변수의 상대적 중요도

기계학습 모델링은 앞서 설명한 바와 같이 각 입력 변수의 상대적 중요도를 제공하는데, 가뭄지수 예측모델에 대해 변수의 상대적 중요도를 SPI6와 SPEI6의 두 가뭄지수, DT모델, RF모델, ET모델의 세 기계학습 모델, 그리고 1개월에서 6개월까지의 선행시간에 대해 검토하였다. Leave-One-Year-Out 교차검증 하였으므로 2003년부터 2015년까지 검증하는데 사용된 각 연도별로 도출된 변수의 상대적 중요도를 모든 검증 연도에 대해 평균하였다. 여기에서 가장 중요한 변수의 상대적 중요도가 100%이다.

모든 경우 PRCP가 가장 중요한 변수로 나타났다(Figure 5-7). 가뭄지수인 SPI의 경우 강수량만을 이용하여 산정되며 SPEI의 경우 강수량과 증발산량을 이용하여 산정되므로 이는 당연한 결과이다. PRCP 다음으로 상대적 중요도가 높은 변수는 MONTH로 각 시기에 따라 모델의 형태 및 결과에 차이가 있음을 보여주었다. 그리고 SPEI6 예측모델의 경우 세 번째로 SPEI 입력 변수로 사용되는 PET의 중요도가 높았다. 그 밖의 다른 변수의 상대적 중요도는 10% 안팎으로 비교적 낮았는데, 특이한 점은 SPEI6 예측모델 중 ET모델의 경우 대규모 대기순환 지수인 MEI의 중요도가 다른 변수들보다 약간 높게 나타났다는 점이다(Figure 5-7).



(a)



(b)

Figure 5-7 Relative importance of input variables for SPEI6 forecast (regression) using (a) DT model, and (b) EF model.

5.3 미계측지역에 대한 정규수분지수 예측 결과

위성 자료에서 도출하는 NDWI는 각 격자에 대해 값이 존재하므로 보간법을 이용할 필요가 없다. 따라서 미래 기간에 대한 값을 기후값에서 도출하는 경우와 장기예

측 자료를 이용하는 경우인 V-C-ML방법과 V-F-ML방법을 비교하였다. 이 경우 각각의 격자는 기상 자료로 치면 지점에 해당하므로 모델의 성능은 앞서 지점에서의 가뭄을 예측한 경우인 C방법 및 F방법과 비교할 있으며 또한 가뭄지수 값이 없는 미계측지역에 대해서도 위성 기반의 식생 지수는 값을 가지므로 C-I방법, C-ML방법, F-I방법, F-ML방법과 비교할 수 있다.

5.3.1 분류 정확도

SPI6와 SPEI6의 분류에 대해서는 McKee et al. (1993)에서 제안한 분류 기준을 그대로 사용하였으나 NDWI의 분류 기준은 다른 변수와 융합된 형태로 제시된 적이 있을 뿐(e.g., Rhee et al., 2010) 가뭄지수와 상응하게 제안된 바를 찾기 어렵다. 따라서 본 연구에서는 여기서 사용한 9596개의 NDWI 자료의 경험적 누적분포함수(Empirical Cumulative Distribution Function, ECDF), SPI6 및 SPEI6의 분류 기준에 해당하는 누적분포에 맞는 Inverse ECDF를 도출하여 기준으로 사용하였다(Figure 5-8, Table 5-3).

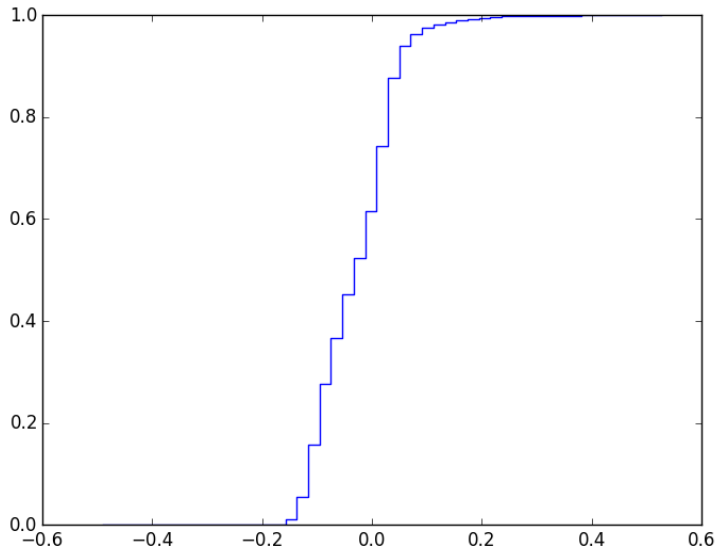


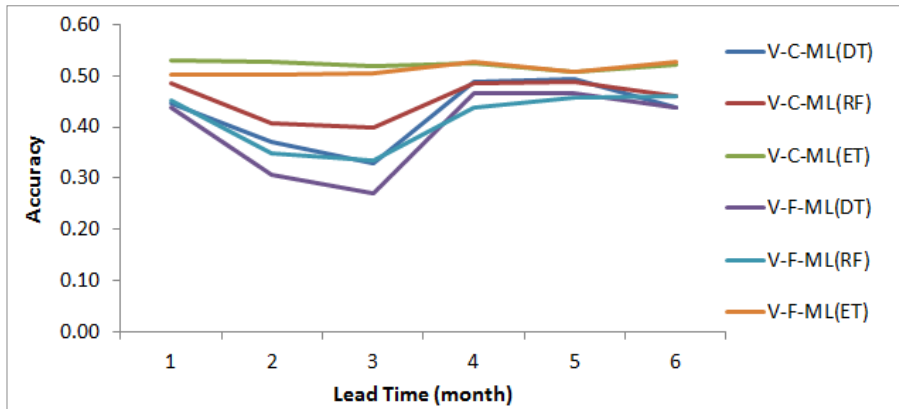
Figure 5-8 Empirical Cumulative Distribution Function of NDWI used in this study.

Table 5-3 Classification categories of NDWI (unitless).

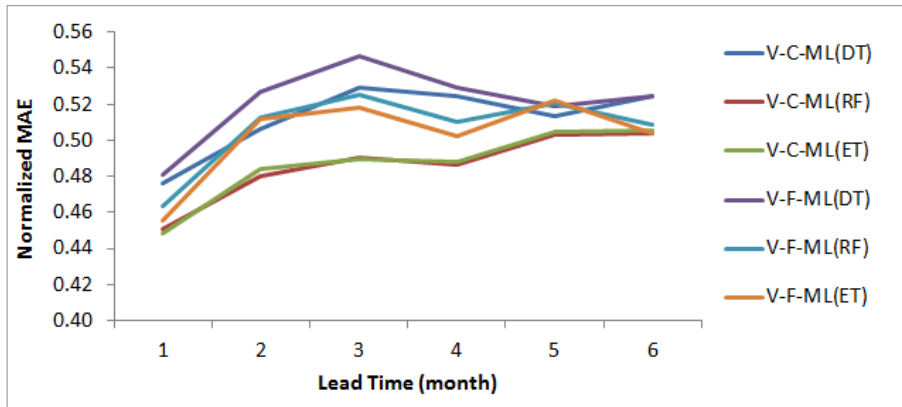
Classification	Index Value
Extremely wet(EW)	≥ 0.1173
Very wet(VW)	0.0673 to 0.1172
Moderately wet(MW)	0.0434 to 0.0672
Near Normal(NN)	-0.0952 to 0.0432
Moderate drought(MD)	-0.1137 to -0.0953
Severe drought(SD)	-0.1283 to -0.1138
Extreme drought(ED)	≤ -0.1284

가뭄지수 예측과 마찬가지로 분류자를 직접 구축하여 모델링을 수행한 경우보다 회귀자를 구축한 후 분류를 수행한 경우가 분류 정확도가 더 높았으므로 이 경우의 결과만을 분석하였다. NDWI 예측의 경우 V-C-ML방법의 분류 정확도가 0.50-0.53으로 매우 높았는데, 이는 가뭄지수 예측의 경우 완벽한 예측을 가정한 경우 가장 높은 성능을 보인 SPEI6 예측을 위한 PF-ML방법(ET모델)의 성능과 유사한 값이다(Table 5-4).

주목할 점은 앞서 지점에서의 가뭄지수 예측의 경우 선행시간이 지나며 분류 정확도가 계속 감소하여 C방법의 경우는 선행시간이 6개월, F방법의 경우는 5개월에는 예측력이 전혀 없이 정확도가 0에 이르렀고(Figure 5-1), 미계측지역의 가뭄지수 예측을 위한 다른 방법의 경우에도 0.16-0.21 수준으로 상당히 감소하였었는데(Figure 5-3), NDWI 예측의 경우 선행시간에 따라 분류 정확도가 유지되어 상대적으로 매우 큰 정확도를 보였다는 점이다(Figure 5-9).



(a)



(b)

Figure 5-9 (a) Classification accuracy and (b) normalized regression MAE for NDWI based on V-C-ML and V-F-ML methods using DT, RF, and ET models.

5.3.2 회귀 성능

앞서 C방법과 F방법의 경우는 가뭄예측 성능 평가의 기준인 SPI6, SPEI6는 평균 0, 표준편차 1의 정규분포 곡선을 따르는 값으로 NDWI 예측의 회귀 오차를 함께 비교하기 위해서는 식생지수 값을 그대로 사용할 수는 없다. 본 연구에서 사용한 9596개의 식생지수 자료는 대부분이 -0.2에서 0.2의 값을 가지며 테스트 결과 정규분포 곡선을 따른다고 할 수는 없지만 비교를 위해서 오차를 위해 구한 MAE에서 평균(-0.02)을 빼고 표준편차(0.069)로 나눈 값으로 정규화하였다(Table 5-4). SPI6과 SPEI6의 지점에서의 예측의 MAE와의 절대적인 비교는 지양하도록 하고 선행시간에 대해 변화하는 양

상만을 관찰하였는데, 앞서 가뭄지수 예측을 위한 C방법과 F방법의 경우 지점에서의 회귀 성능의 경우 MAE가 0.36-0.47 정도의 값에서 점차 증가하여 1.14-1.67 정도에 이르렀다(Figure 5-2). 그러나 NDWI 예측에서는 선행시간에 대한 오차증가의 폭이 그다지 크지 않았다(Figure 5-10). 선행시간 1-5개월의 경우 기후값을 이용한 V-C-ML방법이 장기예측 자료를 이용한 V-F-ML방법보다 오차가 조금 낮았으나 선행시간 6개월의 경우에는 거의 같은 값을 보였다(Figure 5-9, Table 5-4).

가뭄지수를 통해 가뭄에 대한 정보를 얻듯이 NDWI를 통해 가뭄 피해에 대한 정보를 얻을 수 있도록 활용한다면, 분류의 정확도를 위해서는 사용된 방법들이 큰 차이는 없었지만 선행시간 1-3개월인 경우는 V-C-ML방법(ET모델)이 가장 적합하며 4, 6개월에는 V-F-ML방법(ET모델)이 가장 적합하였다(Table 5-4). 선행시간 5개월인 경우는 이 두 모델이 같은 성능을 보였다(Table 5-4). 회귀의 경우 선행시간 1-5개월의 경우 V-C-ML방법(RF모델 및 ET모델), 6개월의 경우 V-C-ML방법(RF모델)과 V-F-ML방법(ET모델)이 가장 적합하였다(Table 5-4).

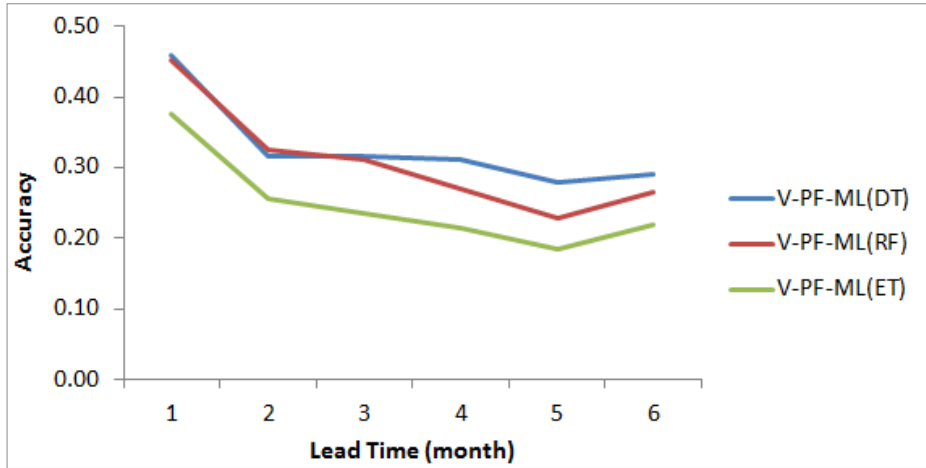
Table 5-4 Model performance measures for ungauged locations based on NDWI.

Accuracy						
LT	V-C-ML(DT)	V-C-ML(RF)	V-C-ML(ET)	V-F-ML(DT)	V-F-ML(RF)	V-F-ML(ET)
1	0.45	0.49	0.53	0.44	0.45	0.50
2	0.37	0.41	0.53	0.31	0.35	0.50
3	0.33	0.40	0.52	0.27	0.34	0.51
4	0.49	0.49	0.52	0.47	0.44	0.53
5	0.49	0.49	0.51	0.47	0.46	0.51
6	0.44	0.46	0.52	0.44	0.46	0.53
MAE						
LT	V-C-ML(DT)	V-C-ML(RF)	V-C-ML(ET)	V-F-ML(DT)	V-F-ML(RF)	V-F-ML(ET)
1	0.012	0.011	0.010	0.013	0.011	0.011
2	0.014	0.013	0.013	0.016	0.015	0.015
3	0.016	0.013	0.013	0.017	0.016	0.015
4	0.016	0.013	0.013	0.016	0.015	0.014
5	0.015	0.014	0.014	0.015	0.015	0.016
6	0.016	0.014	0.014	0.016	0.015	0.014

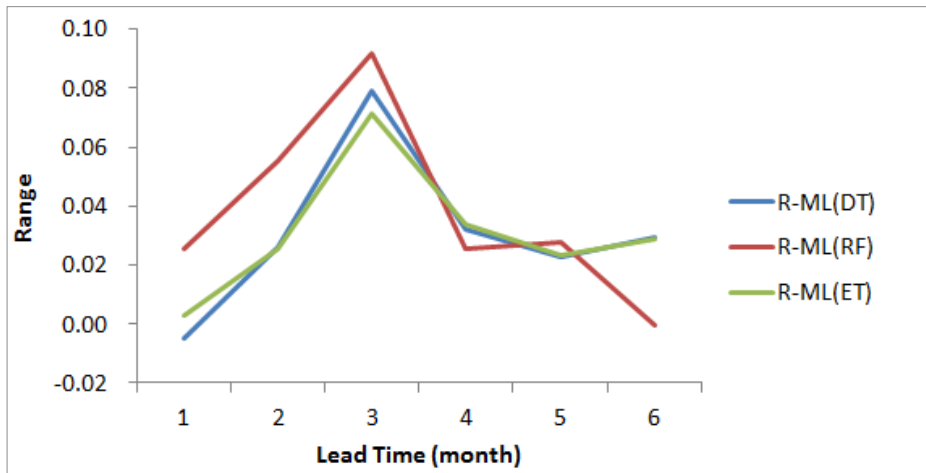
Normalized MAE						
LT	V-C-ML(DT)	V-C-ML(RF)	V-C-ML(ET)	V-F-ML(DT)	V-F-ML(RF)	V-F-ML(ET)
1	0.48	0.45	0.45	0.48	0.46	0.46
2	0.51	0.48	0.48	0.53	0.51	0.51
3	0.53	0.49	0.49	0.55	0.53	0.52
4	0.52	0.49	0.49	0.53	0.51	0.50
5	0.51	0.50	0.50	0.52	0.52	0.52
6	0.52	0.50	0.51	0.52	0.51	0.50

5.3.3 향상 가능한 범위

기계학습 모델은 완벽한 예측 자료를 이용하여 구축되었으나 NDWI 예측의 경우 기후값이나 장기예측값을 입력하여 NDWI 예측의 성능을 평가하였을 때 많은 경우 특히 기준선인 기후값을 이용한 V-C-ML방법의 경우와 큰 차이가 없었다(자료 미제시). 장기예측을 이용한 V-F-ML방법의 경우에 대하여 완벽한 예측의 경우 분류 정확도와 향상 가능한 범위(Figure 5-10), 정규화된 회귀오차 MAE와 향상 가능한 범위를 제시하였다(Figure 5-11).

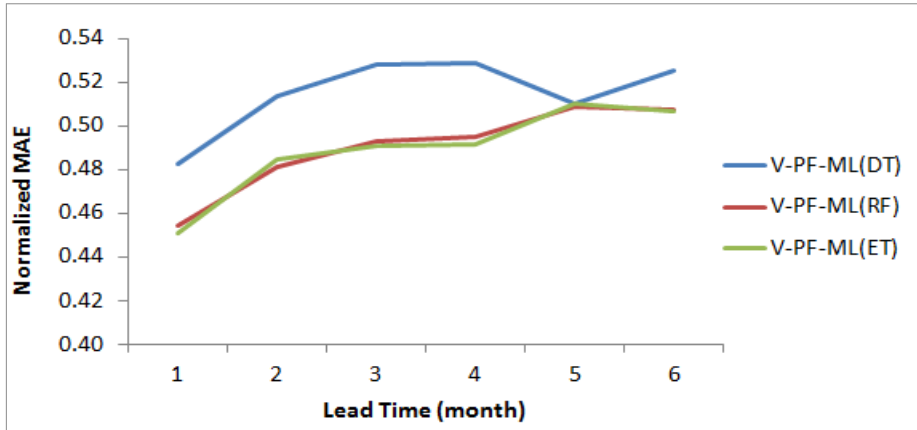


(a)

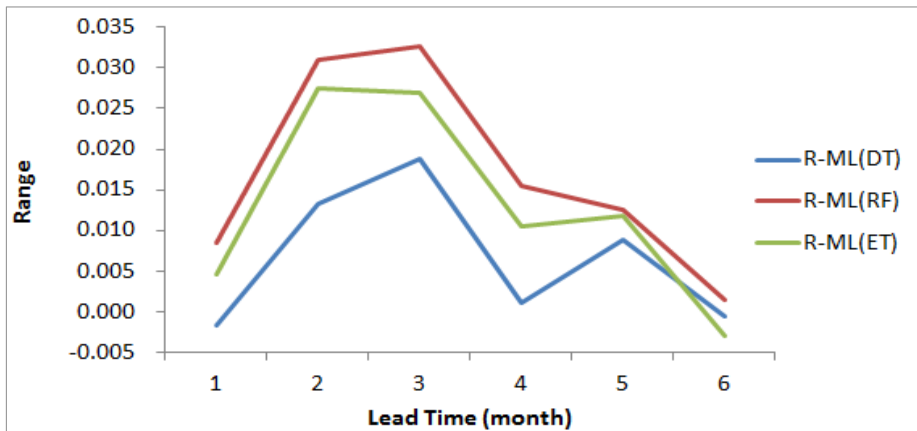


(b)

Figure 5-10 Classification accuracy based on (a) perfect forecast V-PF-ML for NDWI, and (b) the improvable ranges obtained as the difference between V-F-ML and V-PF-ML.



(a)



(b)

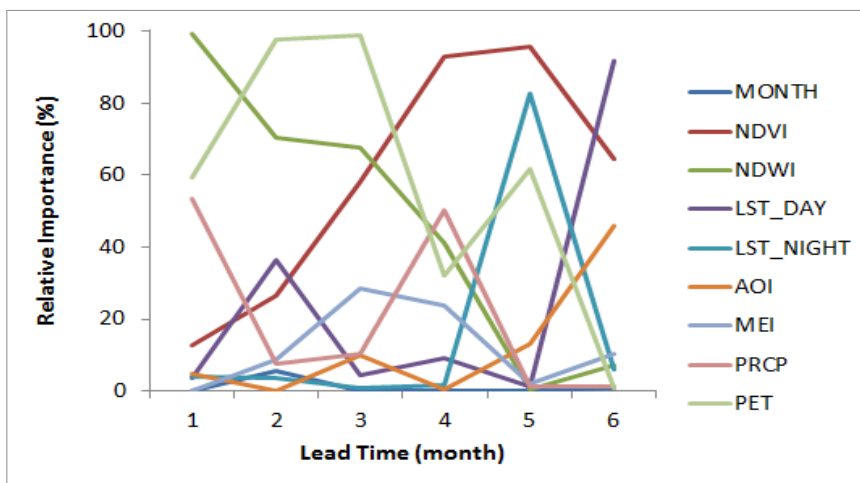
Figure 5-11 Normalized regression MAE based on (a) perfect forecast V-PF-ML for NDWI, and (b) the improvable ranges obtained as the difference between V-F-ML and V-PF-ML.

5.3.4 변수의 상대적 중요도

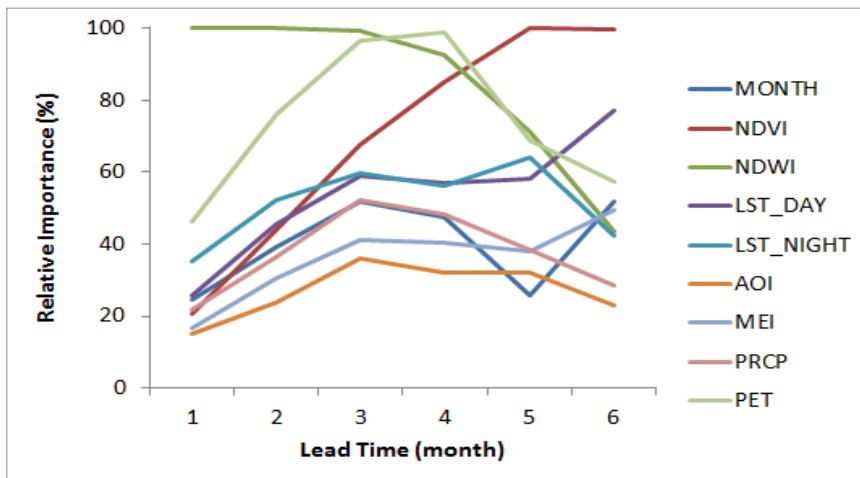
가뭄지수 예측의 경우 가장 중요한 변수가 PRCP로 일관적으로 나타난 것(Figure 5-7)에 비해 NDWI 예측의 경우에는 PRCP의 상대적 중요도가 다른 변수와 비교하여 그렇게 크지 않다. DT모델과 RF모델의 경우 선행시간 1개월인 경우는 NDWI가 가장 중요한 변수로 나타났으며 2-3개월인 경우 PET, 4-5개월인 경우 NDVI, 그리고 6개월인 경우 LST_DAY가 중요하게 나타났으며, ET모델의 경우는 1-3개월의 경우까지 NDWI가 가장 중요하고 4개월인 경우 PET, 5-6개월인 경우 NDVI의 중요도

가 높게 나타났다(Figure 5-12).

선행시간이 짧은 경우에는 NDWI의 초기값이 가장 큰 영향을 미치는데 그 메모리가 그리 오래 가지 않는 것으로 보인다. 그 이후에는 식생과 대기 상호 작용의 중요한 변수인 PET의 중요도가 높다. NDVI는 식생의 수분 함량을 나타내주는 NDWI와는 달리 식생의 활력도를 나타내는데, 이는 선행시간이 긴 경우 결과적으로 식생의 활력도 초기값이 미래의 식생 수분 함량에 영향을 미치기 때문으로 보인다.



(a)



(b)

Figure 5-12 Relative importance of input variables for NDWI forecast (regression) based on V-F-ML method using (a) DT model, and (b) ET model.

6. 토론 및 결론

6.1 가뭄 정보 수요 및 활용 의사

의사결정자의 가뭄정보 수요 파악 및 활용의사 도출을 위해 17개 광역지방자치단체의 가뭄의 환경피해와 관련이 있는 가뭄, 폭염, 재난, 재해, 산불, 도시숲 관련 담당자 51명을 대상으로 설문조사를 수행하였다. 제1차 설문조사에서는 이 중 27명이 응답하였으나(회수율 53%) 1차 응답자 27명을 대상으로 한 제2차 설문조사에서는 5명의 응답만을 얻을 수 있었다(회수율 19%). 표본수는 크지 않으나 실제 담당 공무원을 대상으로 하는 인터뷰 대신에 수행된 설문조사이므로, 개개의 응답을 의미있게 받아들이며 연구의 설계에 참고하였다. 두 차례에 걸쳐 수행된 설문은 총 30개의 문항으로 이루어져 있으며 이를 통해 가뭄 발생 인지, 업무 관련성, 가뭄 정보의 활용 현황 및 수요, 가뭄 대처 현황 및 수요, 가뭄 정보의 정확성에 대한 수용성 등을 도출하였으며 일반적인 가뭄피해유형에 대한 가뭄리스크 평가를 수행하였다.

가뭄정보의 활용에 대한 응답을 보면 1차 설문 응답자의 70%가 기상청 또는 다른 기관에서 제공하는 가뭄 정보를 사용하고 있었으며 2차 설문 응답자의 40%는 위성 기반의 가뭄정보도 활용하고 있었다. 또한 업무에 직접적으로 활용될 수 있는 구체적인 가뭄정보가 공신력 있는 기관에 의해 제공된다면 활용도를 높일 수 있을 것으로 보인다. 가뭄정보의 수요에 있어 일별, 월별 가뭄정보가 필요하다고 응답한 경우도 있었으나 대체적으로 주별로 가뭄정보가 제공되기를 원하였다. 공간해상도는 시군구와 같은 행정구역 단위의 가뭄정보에 대한 요구가 가장 많았으며 공간적으로 연속된 고해상도 격자(1×1 km 이상)에 대한 선호도 있었다. 가뭄예측 정보의 선행 시간은 1개월에서 1년까지 다양하게 응답하였다. 이에 따르면 주별 시간해상도 및 상당히 높은 공간해상도의 가뭄 정보 생산을 위한 모델 구축이 뒤따라야 하나, 본 연구에서는 우선 장기예측자료의 월별 시간해상도와 위성기반 식생 자료의 $0.05 \times 0.05^\circ$ 공간해상도에 맞추어 가뭄예측 모델을 개발하도록 하였다. 상세한 시간해상도 및 공간해상도 모델은 후속연구에서 개발할 것이다.

미계측지역의 가뭄정보 공간유형에 있어서는 지점 관측 자료에 기반을 둔 가뭄정보와 위성 자료에 기반을 둔 가뭄정보에 대한 수요가 모두 있었으며, 가뭄지표의 선

정에 있어 2차 설문 응답자의 60%가 위성 기반 가뭄 지표의 활용 의사를 밝혔다. 정확도에 있어 위성 기반 가뭄지표의 정확도가 지점의 관측 기반 정보보다 낮은 경우에도 위성 기반 가뭄지표를 사용하겠다고 답한 응답자도 있었다. 일반적인 가뭄피해유형에 대한 가뭄리스크 평가가 수행되었는데, 응답자가 적어 유의한 결론을 이끌어내지는 못하였다. 그러나 응답자들이 가뭄의 발생빈도와 피해강도를 평가함에 있어 그 간 간과되어 온 가뭄의 환경적 피해와 사회적 피해를 다른 피해 유형과 마찬가지로 높게 평가한 점은 작은 표본 수에도 불구하고 주목할 만 하다. 본 결과를 기반으로 각 피해 유형의 상세한 리스크 평가가 수행되어야 할 것이다.

6.2 기후값과 장기예측

분류 정확도에 있어, 가뭄지수 예측의 경우 모든 선행시간에서 기준선인 기후값을 이용한 경우(C방법, C-I방법, C-ML방법)가 장기예측 자료를 이용한 경우(F방법, F-I방법, F-ML방법)보다 우수한 결과를 생산했다. NDWI 예측의 경우도 선행시간이 짧은 경우에는 V-C-ML방법이 V-F-ML방법보다 분류 정확도가 높았다. 회귀 오차의 경우에도 마찬가지로 기준선인 기후값을 이용한 경우(C방법, C-I방법, C-ML방법)가 장기예측 자료를 이용한 경우(F방법, F-I방법, F-ML방법)보다 낮은 오차를 생산했으며 NDWI 예측의 경우도 역시 같은 결과를 보였다. 기후값을 이용한 경우를 기준선으로 놓고 장기예측 자료를 이용했을 때 예측성이 어느 정도인지를 평가해야 하지만 본 연구에서 장기예측 자료를 단순한 방법론만을 이용하여 적용시켰으므로 큰 예측성을 기대할 수 없었다.

그러나 점차 장기예측의 예측성이 향상되거나 통계적 다운스케일링의 결합 등 발달된 방법론을 통해 적용한다면 완벽한 예측을 가정한 경우에서 볼 수 있었던 가뭄 예측 성능에 가깝게 다가갈 수 있을 것으로 기대된다. 완벽한 예측을 가정한 경우인 PF-ML방법의 분류 정확도는 기후값을 이용한 C-ML방법의 성능을 훨씬 능가하며 심지어 지점에서의 C방법의 분류 정확도보다도 높은 값을 보이게 되므로, 장기예측의 성능이 향상됨에 따라 장기예측을 이용한 가뭄예측이 향상될 수 있을 것으로 보인다.

6.3 보간법과 기계학습 모델링

보간법과 기계학습 모델링은 미계측지역의 SPI6, SPEI6 두 가뭄지수의 예측을 통해 비교하였는데, 결론적으로 모든 경우 보간법보다 기계학습 모델링의 성능이 우수했다. 완벽한 예측을 가정한 결과를 보면 역시 기계학습 모델링의 분류 정확도가 높고, 회귀 오차가 낮았다. C-ML방법의 성능이 가장 좋은데, C-I방법과 F-ML방법을 비교하더라도 많은 경우 F-ML방법이 더 우수한 결과를 보였다. F-ML방법 중 분류의 경우에는 특히 ET모델이 C-I방법에 비해 훨씬 우수한 성능을 보이며 회귀의 경우 DT, RF, ET모델 모두가 C-I방법에 비해 높은 예측 성능을 나타냈다. 가뭄예측에 있어 장기예측의 성능이 가지는 한계를 기계학습 모델링을 통해 많은 경우 극복할 수 있다고 할 수 있다.

6.4 가뭄지수와 정규수분지수

가뭄지수인 SPI6와 SPEI6는 강수량과 잠재증발산량을 이용하여 산정하므로 입력 변수의 상대적 중요도 평가에서도 볼 수 있었듯이 이러한 변수들이 모델에 큰 영향을 미친다. 특히 강수량은 모든 경우에 대해 가장 큰 중요도를 지니는데 따라서 두 가뭄지수의 예측성은 장기예측 자료의 예측성에 더 큰 영향을 받을 수 있다. NDWI의 예측도 이 두 변수의 영향을 크게 받으나 짧은 선행시간에서 NDWI 초기값의 중요도가 매우 크고 선행시간이 길어질수록 PET, NDVI의 초기값 등 다른 변수의 중요도가 크게 나타났다. NDWI가 추정하는 것은 식생의 수분 함량으로 이는 여러 변수와의 상호작용에 따라 결정되므로 장기예측 자료에 덜 민감하며 모든 선행시간에 대해 높은 분류 정확도와 작은 회귀 오차를 보였다.

6.5 결론 및 제언

본 연구에서는 장기예측 자료를 활용한 가뭄예측을 기후값을 이용하는 기준선과 비교하였으며, 미계측지역에서 기계학습을 이용하여 가뭄예측을 수행한 경우를 보간법과 비교하였다. 기후값을 이용하여 기계학습 모델링을 적용시킨 경우가 대부분의

경우 가장 가뭄지수 예측 성능이 우수하게 나타났다. 장기예측 자료가 가뭄예측에 기여하는 바는 크지 않았으나 위성 자료를 이용한 기계학습 모델링을 통해 가뭄예측의 성능 향상에 기여할 수 있었다.

결과적으로 $0.05 \times 0.05^\circ$ 해상도에서 SPI6, SPEI6 가뭄지수와 NDWI 예측을 수행하는 경우 장기예측 자료를 사용하지 않고 위성자료의 기후값을 이용하여 기계학습 모델링을 수행, 공간적으로 연속적으로 분포된 가뭄정보를 생산하는 것이 가장 적합하다는 결론을 얻을 수 있었다. 따라서 가뭄 카테고리의 분류를 중심으로 다음과 같은 가뭄예측 모델이 선정되었다.

- SPI6 예측의 경우, 선행시간 1-5개월: C-ML방법(ET모델)
- SPI6 예측의 경우, 선행시간 6개월: C-ML방법(DT모델)
- SPEI6 예측의 경우, 모든 선행시간: C-ML방법(ET모델)
- NDWI 예측의 경우, 선행시간 1-3개월: V-C-ML방법(ET모델)
- NDWI 예측의 경우, 선행시간 4-6개월: V-F-ML방법(ET모델)

C-ML방법의 경우 미계측지역에서 분류의 정확도는 1개월 선행시간의 경우 0.47-0.52에서 선행시간이 길어질수록 감소하여 선행시간 6개월의 경우에는 0.21-0.35 정도의 값을 가지게 되며, 회귀 오차 MAE의 경우 1개월 선행시간에서 0.41-0.47에서 선행시간이 길어질수록 증가하여 6개월 선행시간에서 0.56-0.59 정도가 된다. 장기예측의 성능이 향상되면 장기예측을 이용하는 경우가 기후값을 이용하는 경우보다 예측 성능이 더 커질 수 있는데, 이 경우 장기예측 자료와 위성자료를 이용하여 기계학습 모델링을 수행하는 경우에 완벽한 예측을 가정한 가뭄예측 성능의 상한은 분류 정확도가 0.50-0.56, 회귀 오차 MAE는 0.35-0.40 수준이 될 것이다.

기계학습 모델링을 이용한 NDWI의 예측은 가뭄지수 예측과는 달리 장기예측 자료의 예측성에 크게 민감하지 않으면서 모든 선행시간에 대해 비교적 높은 분류 정확도와 작은 회귀 오차를 보였다. 가뭄지수 예측의 경우 선행시간이 길어짐에 따라 예측 성능이 감소하는 단점을 보완할 수 있으며, 식생의 수분 상태를 나타내는 지수로서 가뭄의 환경피해와 연관지어 가뭄지표로서 활용할 수 있다. 따라서 장기예측 자료

의 예측성을 개선시켜 의사결정자에게 익숙한 여러 가뭄지표의 예측 성능을 함께 향상시키는 노력을 지속함과 동시에 NDWI와 같이 가뭄의 환경피해를 나타내는 가뭄지표의 예측을 통한 가뭄정보를 생산하고 제공하여야 한다. V-C-ML방법의 경우, 분류의 정확도는 0.51-0.53 정도의 값을 가지게 되며 회귀 오차는 정규화된 MAE를 보았을 때 0.45-0.50 정도의 값을 가진다. 본 연구에서는 정규수분지수인 NDWI를 그 자체로 가뭄지표로 삼아 예측하였으나 이러한 위성 기반의 지수를 다른 변수와 융합하여 가뭄지수를 구성하는 SDCI(Scaled Drought Condition Index) 등(Rhee et al., 2010) 여러 가지 가뭄의 환경피해와 관련된 가뭄지표를 활용할 수 있다.

본 연구에서 기후값 및 단순한 방법론을 이용한 장기예측 자료를 통해 가뭄지수 예측의 성능을 도출하였는데, 이를 우리나라에서 수행하는 가뭄지수 예측 연구에 기준선으로서 활용할 수 있을 것이다. 또한 의사결정에 있어 불확실성에 효과적으로 대응하고자 본 연구를 통해 개발된 지역 규모의 활용을 위한 고해상도 가뭄예측 방법을 기상청의 가뭄정보시스템 등과 상호보완적으로 결합하여 활용하길 기대한다.

본 연구가 지니는 여러 한계는 후속 연구를 통해 개선될 것이다. 본 연구에서 사용된 기계학습 모델에 공간지리적인 정보와 초기조건을 보강하는 추가적인 변수의 고려 등 개선이 요구된다. 시간적, 공간적으로 차별화된 연구를 진행하지 못하였으므로 월별, 지역별 분석을 수행할 것이며 위성 자료에 기반을 둔 융합 가뭄지수를 활용하여 위성 자료를 이용해서도 의사결정자에게 친숙한 가뭄지수를 예측하는 연구가 뒤따를 것이다. 또한 아시아태평양 지역을 대상으로 적용하여 대상 지역에 적합한 가뭄지표 및 고해상도 가뭄정보 생산을 위한 연구를 수행할 것이다.

참고 문헌

<국문 자료>

- 강인식, 1998: 엘니노와 한반도 기후변동의 관련성. *한국기상학회지*, 34(3), 390-396.
- 국가통계포털, 2015a: 지적통계: 행정구역별·지목별 국토이용현황. [Available online at <http://kosis.kr>.]
- 국가통계포털, 2015b: 국제통계연감: 기후. [Available online at <http://kosis.kr>.]
- 국립방재연구소, 1998: 가뭄 피해 원인의 조사 분석. 연구보고서 NIDP-98-04, 98 pp.
- 권진주, 안재현, 김태웅, 2013: 가뭄심도 분류기준의 개선방안 제시. *한국수자원학회지*, 46(2), 195-204.
- 권원태, 오재호, 2001: 가뭄에 대한 기상학적 고찰. *한국수자원학회지*, 34(4), 81-86.
- 기상연구소, 1993: 우리나라 가뭄에 관한 연구 - 강수량 변동에 관한 기후학적 특성 연구. 연구보고서 MR 93A-005, 80 pp.
- 기상청, 2001: 2000년도 기상연감. 376 pp.
- 기상청, 2007: 2006년도 기상연감. 445 pp.
- 기상청, 2011: 기후변화 시나리오 보고서 2011. 117 pp.
- 기상청, 2012: 2011년도 기상연감. 449 pp.
- 김기훈, 변희룡, 이영임, 황호성, 2003: 해빙을 이용한 한국 5월 가뭄의 장기예측. *대기*, 13(3), 382-383.
- 김성, 2004: 한반도 봄철 강수량과 연관된 북반구 대기순환의 패턴 변화. 공주대학교 대학원 대기과학과 대기과학 전공 석사학위논문, 61 pp.
- 김승, 정성원, 김현준, 김형섭, 2005: '94-'95 가뭄 실태. *'94-'95 가뭄 심포지엄*, 한국건설기술연구원·한국수문학회, 5-38.
- 김연주, 이진영, 장평화, 송재열, 2013: 가뭄 유형별 기후변화 적응정책 연구. 한국환경정책평가연구원. 235 pp.
- 김휘린, 박무중, 윤용남, 2004: 2001년 가뭄 특성에 관한 비교 연구. 한국수자원학회 2004년도 학술발표회, 1-5.
- 농림부, 2000: 2000년 가뭄·풍수해 극복 사례집. 713 pp.
- 문장원, 이동률, 2009: 2008-2009년 가뭄 현황 분석. 한국수자원학회 2009년도 학술발표회 초록집, 1514-1519.
- 박정술, 김경탁, 2009: 가뭄모니터링을 위한 MODIS NDVI의 활용성 평가: 가뭄지수와의 비교

- 를 중심으로. *한국GIS학회지*, 17(1), 117-129.
- 박종서, 전종갑, 2003: 1960~2001동안의 동아시아 봄가뭄 특성 및 발생 원인. *대기*, 13(3), 236-237.
- 변희룡, 1996: 한반도에 가뭄을 초래하는 대기순환. *한국기상학회지*, 32(3), 455-468.
- 변희룡, 강경아, 김기훈, 2001: Nino-3 아노말리와 아시아 가뭄·홍수의 상관. *한국기상학회지*, 37(5), 453-464.
- 배덕효, 손경환, 2012: 국내가뭄예보 현황 및 발전방향. *물과 미래*, 45(5), 46-54.
- 서길수, 2015: 원격상관 패턴과 한반도 봄가뭄의 시공간적 변화특성 분석. 중부대학교 대학원 토목공학과 석사학위논문, 66 pp.
- 서찬양, 김경준, 최우정, 심재현, 최민하, 2014: 인공위성 가뭄지수를 활용한 한반도 가뭄 평가: 2013년 남부지방 가뭄을 대상으로. *한국방재학회지*, 14(3), 127-131.
- 성노훈, 서민지, 이경상, 이창석, 김현지, 최성원, 한경수, 2015: NDWI를 활용한 한반도 지역의 산림 캐노피에 대한 water stress 평가. *대한원격탐사학회지*, 31(2), 77-83.
- 소방방재청, 2009: 2008~2009 가뭄극복추진 성과보고서. 560 pp.
- 손수진, 2012: Long-lead multi-model ensemble prediction of a drought index sensitive to global warming. APCC Technical Report 2012-01, 111-135.
- 송병현, 홍성길, 최영진, 1995: 1994년 고온 및 가뭄의 기상학적 특성과 산업에의 영향. 한국수자원학회 1995년도 학술발표회 논문집, 541-547.
- 신사철, 김정탁, 김주훈, 박정술, 2006: NDMI를 이용한 가뭄발생지역 검출 및 부족수분량 산정. *한국지리정보학회지*, 9(2), 102-114.
- 신사철, 어민선, 2004: 인공위성 영상 정보를 이용한 가뭄상황 및 징후분석. *한국방재학회논문집*, 4(2), 61-69.
- 신수현, 국민정, 이규성, 2004: 한반도의 다중시기 NDVI를 이용한 가뭄지수 적용성 분석. 한국 GIS학회 2004년도 GIS/RS 공동 춘계학술대회 논문집, 203-208.
- 심기오, 2009: 국한가뭄 대응방안에 관한 연구. 방재연구소, 170 pp.
- 심현준, 류정수, 안재현, 김정숙, 김상단, 2013: 자연공급지역에서 가뭄상황판단을 위한 실시간가뭄지수. *한국방재학회지*, 13(5), 365-373.
- 안정훈, 황필선, 강신욱, 2012년 봄 가뭄의 상황 및 대응. *물과 미래*, 45(9), 84-91.
- 안중배, 류정희, 조익현, 박주영, 류상범, 1997: 한반도 기온 및 강수량과 적도 태평양 해면온도와의 상관관계에 관한 연구. *한국기상학회지*, 33(3), 487-495.
- 어민선, 2001: 2001년 가뭄과 대책. *한국방재학회지*, 1(2), 4-8.
- 오국렬, 유인상, 2014: 우리나라 봄철 가뭄의 피해현황과 대책. *한국방재학회지*, 14(2), 4-9.

- 이기람, 김영섭, 한경수, 이창석, 염종민, 2008: 위성자료를 이용한 중국과 몽골 사막주변의 식생수분상태 모니터링. *한국지리정보학회지*, 11(4), 94-100.
- 이동률, 1999: 엘니뇨/남방진동과 한국의 가뭄과 관계. *한국수자원학회지*, 32(2), 111-120.
- 이동률, 김웅태, 2001: 2001년 봄 가뭄의 심도. *한국수자원학회지*, 34(4), 68-73.
- 이동률, 김웅태, 이대희, 2002: 2001년 가뭄의 특성과 진행과정 분석. *한국수자원학회 2002년도 학술발표회 논문집(III)*, 898-903.
- 이대희, 이동률, 김웅태, 2002: 과거 가뭄의 현상 및 가뭄지수 분석. *한국수자원학회 2002년도 학술발표회 논문집(III)*, 787-792.
- 이종화, 변영화, 박정규, 2001: 2001년 봄 가뭄 특성 분석. *대기*, 11(3), 342-345.
- 이종화, 정현숙, 박정규, 2003: 한반도 봄 가뭄 특성 및 원인 분석. *대기*, 13(1), 370-373.
- 이진영, 2015: 원격 탐사 자료를 활용한 가뭄 감시·예측 기술 개발. APEC기후센터 연구보고서 2015-14, 170 pp.
- 장기호, 정진임, 자영민, 양하영, 최영진, 권원태, 2011: 최근 강수 추세 및 가뭄년도의 특징. *한국방재학회지*, 11(1), 54-66.
- 정수, 신사철, 2006: 광역의 가뭄 분석을 위한 위성영상의 활용. *한국지형공간정보학회지*, 14(2), 55-62.
- 중앙재난안전대책본부 관계부처합동. 2015년 가뭄대비 관계부처 합동지침. 78 pp.
- 차은정, 전종갑, 정효상, 1999: 엘니뇨/라니냐 해의 우리 나라 기후 특성에 관한 연구. *한국기상학회지*, 35(1), 98-117.
- 차은정, 2012: 2012년 여름을 되돌아보다: 짧은 장마와 가뭄, 폭염과 열대야, 집중호우와 태풍. *한국방재학회지*, 12(3), 4-9.
- 최기선, 변희룡, 박수정, 김아정, 2003: 온도장을 이용한 한국 5월 가뭄의 장기예측. *대기*, 13(3), 380-381.
- 최기선, 김도우, 이지선, 변희룡, 2009: 한국 5월 가뭄과 북태평양진동의 연관성. *대기*, 19(1), 67-78.
- 최영진, 1995: 가뭄의 원인과 예측. *한국수자원학회지*, 28(1), 16-20.
- 하경자, 1995: 겨울철 서울 기온의 경년 변동과 태평양 해면 온도와의 상관성 연구. *한국기상학회지*, 31(4), 313-323.
- 한국수자원공사, 2002: 가뭄관리 종합대책 수립연구. 보고서 번호 IPD-2002-3, 378 pp.
- 황석환, 이재형, 2009: 2008년 유출과 가뭄현황. *하천과 문화*, 5(1), 92-95.

<영문 자료>

- Adams, J., 1995: *Risk*. University College London Press, London, UK, 228 pp.
- Anderson, L. O., Y. Malhi, L. E. O. C. Aragao, R. Ladle, E. Arai, N. Barbier, and O. Philips, 2010: Remote sensing detection of droughts in Amazonian forest canopies. *New Phytol.*, 187, 733-750.
- Brooks, N., 2003: Vulnerability, risk and adaptation: A conceptual framework. Tyndall Centre for Climate Change Research, Working Paper 38, 20 pp.
- Broomhead, D. S., and D. Lowe, 1988: Multivariable functional interpolation and adaptive networks. *Compl. Syst.*, 2, 321-355.
- Chakraborty, A., and V. K. Shegal, 2010: Assessment of agricultural drought using MODIS derived Normalized Difference Water Index. *J. Agr. Phys.*, 10, 28-36.
- Chiew, F. H. S., and T. A. McMahon, 2002: Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability. *Hydrol. Sci. J.*, 47, 505-522.
- Cressie, N. A. C., 1993: *Statistics for Spatial Data*, Revised Ed. John Wiley & Sons Inc., 928 pp.
- Cutter, S., B. Owsman-Elasha, J. Campbell, S.-M. Cheong, S. McCormick, R. Pulwarty, S. Supratid, and G. Ziervogel. 2012: Managing the risks from climate extremes at the local level. *Managing the Risks of Extreme Events and Disasters to Advanced Climate Change Adaptation*, Field, C. B., V. Barros, T. F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, M. D. Mastrandrea, K. J. Mach, G.-K. Plattner, S. K. Allen, M. Tignor, and P. M. Midgley, Eds., A Special Report of Working Group I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK, and New York, NY, USA, 291-338.
- DEFRA (Department of Environment, Food, and Rural Affairs), 2011: Guidelines for environmental risk assessment and management – Revised departmental guidance. Defra and Cranfield University, London, UK, 84 pp.
- Downing, T. E. , R. Butterfield, S. Cohen, S. Huq, R. Moss, A. Rahman, Y. Sokona, and L. Stephen, 2001: *Vulnerability Indices: Climate Change Impacts and Adaptation*. UNEP Policy Series 3, UNEP, Nairobi. Kenya, 91 pp.
- Ellis, A. W., G. B. Goodrich, and G. M. Garfin, 2010: A hydroclimatic index for examining patterns of drought in the Colorado River Basin. *Int. J. Climatol.*, 30(2), 236-255.
- EM-DAT, 2014: The OFDA/CRED International Disaster Database. Université Catholique de Louvain, Brussels, Belgium. [Available online at: <http://www.em-dat.net>.]

- Gao, B.-C., 1996: NDWI - A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water From Space, *Remote Sens. Environ.*, 58, 257-266.
- Gu, Y., J. F. Brown, J. P. Verdin, and B. Wardlow, 2007: A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. *Geophys. Res. Lett.*, 34, L06407, doi: 10.1029/2006GL029127.
- Gu, Y., E. Hunt, B. Wardlow, J. B. Basara, J. F. Brown, and J. P. Verdin, 2008: Evaluation of MODIS NDVI and NDWI for vegetation drought monitoring using Oklahoma Mesonet soil moisture data. *Geophys. Res. Lett.*, 35, L22401, doi:10.1029/2008GL035772.
- Guan, Z., and T. Yamagata, 2003: The unusual summer of 1994 in East Asia: IOD teleconnections. *Geophys. Res. Lett.*, 30(1), 1544, doi:10.1029/2002GL016831.
- Guttman, N. B., 1999: Accepting the Standardized Precipitation Index: A Calculation Algorithm. *J. Am. Water Resour. As.*, 35, 311-322.
- Hayes, M., M. Svoboda, N. Wall, and M. Widhalm, 2011: The Lincoln Declaration on Drought Indices: Universal Meteorological drought Index Recommended. *Bull. Ameri. Meteo. Soci.*, 92(4), 485-488.
- Huang, X., and J. R. Jensen, 1997: A machine-learning approach to automated knowledge-base building for remote sensing image analysis with GIS data. *Photogramm. Eng. Rem. S.*, 63(10), 1185-1194.
- Huang, J., D. Chen, and M. H. Cosh. 2014: Sub-pixel reflectance unmixing in estimating vegetation water content and dry biomass of corn and soybeans cropland using normalized difference water index (NDWI) from satellites. *Int. J. Remote Sens.*, 30(8), 2075-2104.
- IPCC (Intergovernmental Panel on Climate Change), 2001: Summary for policymakers. *Climate Change 2001: Impacts, Adaptation and Vulnerability*, McCarthy, J. J., O. F. Canziani, N. A. Leary, D. J. Dokken, and K. S. White, Eds., Cambridge University Press, Cambridge, UK, 1-18.
- Janowiak, J. E., 1988: The Global climate for March-May 1988: The end of the 1986-1987 Pacific Warm Episode and the onset of widespread drought in the United States. *J. Climate*, 1, 1019-1040.
- Jensen, J. R., 2005: Introductory digital image processing: A remote sensing perspective. 3rd Ed., Prentice Hall, Upper Saddle River, New Jersey, USA, 544 pp.
- Jones, R., and R. Boer, 2004: Assessing current climate risks. *Adaptation Policy Framework for Climate Change: Developing Strategies, Policies and Measures*, Lim, B., and E. Spanger-Siegfried, Eds., UNDP, Cambridge University Press, New York, USA,

91-117.

- Kanamitsu, M., and T. N. Krishnamurti, 1978: Northern summer tropical circulations during drought and normal rainfall months. *Mon. Wea. Rev.*, 106, 331-347.
- Khaleghi, B., A. Khamis, F. O. Karray, and S. N. Razavi, 2013: Multisensor data fusion: A review of the state-of-the-art. *Inform. Fusion*, 14(1), 28-44.
- Kogan, F. N., 1995: Application of Vegetation Index and Brightness Temperature for Drought Detection. *Adv.Space Res.*, 15, 91-100.
- Lyon, B., M. A. Bell, and M. K. Tippett. 2012: Baseline probabilities for the seasonal prediction of meteorological drought. *J. Appl. Meteor. Climatol.*, 51, 1222-1237.
- McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration of time scales. *Proc. of the 8th Conf. of Applied Climatology*, Anaheim, CA, USA, Amer. Meteor. Soc., 179-184.
- Mishra, A. K., and V. R. Desai, 2006: Drought forecasting using feed-forward recursive neural network. *Ecol. Model.*, 198, 127-138.
- Mishra, A. K., and V. P. Singh, 2010: A review of drought concept. *J. Hydrol.*, 391, 202-216.
- Mishra, A. K., and V. P. Singh, 2011: Drought modeling – A review. *J. Hydrol.*, 403, 157-175.
- Mo, K., and B. Lyon, 2015: Global meteorological drought prediction using the North American Multi-Model Ensemble. *J. Hydrometeor.*, 16, 1409-1424.
- Morgan, M. G., and M. Henrion, 1990: Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Pennsylvania, USA, 332 pp.
- Namias, J., 1955: Some meteorological aspects of drought: With special reference to the summers of 1952-54 over the United States. *Mon. Wea. Rev.*, 83, 199-205.
- Namias, J., 1972: Influence of Northern Hemisphere general circulation on drought in Northeast Brazil. *Tellus A*, 24(4), 336-343.
- Namias, J., 1991: Spring and summer 1988 drought over the contiguous United States - Causes and prediction. *J. Climate*, 4, 54-65.
- NIDIS (National Integrated Drought Information System) Act, 2006. United States Public Law 109-430. [Available online at: <http://www.gpo.gov/fdsys/pkg/PLAW-109publ430/pdf/PLAW-109publ430.pdf>.]
- Park, C.-K., and S. D. Schubert, 1997: On the nature of the 1994 East Asia summere drought. *J. Climate*, 10(5), 1056-1070.
- Quan, X.-W., M. P. Hoering, B. Lyon, A. Kumar, M. A. Bell, M. K. Tippett, and H. Wang,

- 2012: Prospects for dynamic prediction of meteorological drought. *J. Appl. Meteor. Climatol.*, 51, 1238-1252.
- Rhee, J., and J. Cho. 2016: Future changes in drought characteristics: Regional analysis for South Korea under CMIP5 projections. *J. Hydrometeor.*, 17, 437-451.
- Rhee, J., J. Im, and G. J. Carbone, 2010: Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data. *Remote Sens. Environ.*, 114, 2875-2887.
- Ropelewski, C.F., 1988: The global climate for June-August 1988: A swing to the positive phase of the Southern Oscillation, drought in the United States, and abundant rain in monsoon areas. *J. Climate*, 1, 1153-1174.
- Seneviratne, S.I., N. Nicholls, D. Easterling, C.M. Goodess, S. Kanae, J. Kossin, Y. Luo, J. Marengo, K. McInnes, M. Rahimi, M. Reichstein, A. Sorteberg, C. Vera, and X. Zhang, 2012: Changes in climate extremes and their impacts on the natural physical environment. *Managing the Risks of Extreme Events and Disasters to Advanced Climate Change Adaptation*, C. B. Field, V. Barros, T. F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, M. D. Mastrandrea, K. J. Mach, G.-K. Plattner, S. K. Allen, M. Tignor, and P. M. Midgley, Eds., A Special Report of Working Group I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, 109-230. [Available online at: <http://www.ipcc-wg2.gov/SREX/>].
- Smith, K., 1996: Environmental Hazards, Routledge, London, UK, 389 pp.
- Thorntwaite, C. W., 1948: An approach toward a rational classification of climate. *Geogr. Rev.*, 38(1), 5594. doi:10.2307/210739.
- Tsakiris, G., D. Pangalou, and H. Vangelis, 2007: Regional drought assessment based on the Reconnaissance Drought Index (RDI). *Water Resour. Manag.*, 21, 821-833.
- UNDHA (United Nations Department of Humanitarian Affairs), 1992: Internationally Agreed Glossary of Basic Terms Related to Disaster Management. UNDHA, Geneva, Switzerland, 81 pp.
- Vicente-Serrano, S. M., S. Begueria, and J. I. Lopez-Moreno, 2010: A multiscalar drought index sensitive to global warming: The Standardized Precipitation Evapotranspiration Index. *J. Climate*, 23, 1696-1718.
- Wilhite, D., and M. R. Glantz, 1985: Understanding the drought phenomenon: The role of definitions. *Water Int.*, 10, 111-120.

부록 1. 설문조사지

가뭄감시 및 예측정보 활용현황 및 인식조사

안녕하십니까? 귀한 시간 내어 설문에 응해 주셔서 대단히 감사합니다.

아태지역 이상기후 감시 및 예측의 중추적 역할을 수행하고 있는 공공연구기관인 APEC 기후센터(<http://www.apcc21.org/kor/index.jsp>)에서는 가뭄 취약성 감소를 통한 가뭄 대응 제고를 위해 「**지속가능한 재난관리를 위한 가뭄감시·예측정보 활용 연구**」를 수행 중에 있습니다.

가뭄의 환경적 피해에 대한 대응과 관련된 의사결정이 이루어지는 전국 17개 광역자치체를 대상으로 가뭄감시 및 예측정보 활용현황 및 인식조사를 진행하오니 적극적인 참여를 부탁드립니다.

답변해 주신 자료는 본 연구에 소중하게 사용될 것이며 응답하신 내용은 통계법 제33조(비밀의 보호 조항)에 따라 철저히 비밀이 보장되며 순수한 학문적 연구 목적으로만 사용됩니다. 본 조사에 문의사항이 있으시면 아래 연락처로 연락주시기 바랍니다.

설문에 소요되는 시간은 약 **5분**으로 예상되며, 바쁘시더라도 국내 가뭄 취약성 감소 및 대응 제고를 위하여 잠시 시간을 내시어 설문에 응해 주시면 대단히 감사하겠습니다.

2015년 4월

주관기관	APEC 기후센터
연구진	이진영 선임연구원 (051-745-3959)
이메일	이진영 선임연구원 (jyrhee@apcc21.org)
팩스	051-745-3999
우편	부산광역시 해운대구 센텀7로 12 APEC 기후센터, 612-020

※ 본 설문지를 **5월 4일 (월)**까지 작성하여 위의 이메일, 팩스, 또는 우편으로 회신하여 주시기 바랍니다.

〈설문조사 작성요령〉

- 선택형 문항의 경우 괄호 안에 해당되는 항목을 기입하여 주시고 (필요시 복수응답가능, 예: (1, 4))
- 서술형 문항의 경우 괄호 또는 빈칸에 길이에 상관없이 자유롭게 서술하여 주시면 됩니다.

가뭄 발생 인지 및 업무 관련성

1. 해당 시도에 가뭄이 발생했다는 것을 주로 어떤 경로로 인지하십니까? ()
- ① 신문, 뉴스 등의 언론보도를 통해
 - ② 생활용수, 농업용수 부족 등으로 인한 민원을 통해
 - ③ 소속기관 내 가뭄 담당자의 공지를 통해
 - ④ 기상청 등의 중앙행정기관에서 제공하는 가뭄정보를 통해
 - ⑤ 개인의 경험을 통해
 - ⑥ 기타 _____
2. 귀하께서 현재 맡고 계시는 업무가 가뭄 발생의 영향을 받습니까? ()
- ① 매우 그렇다
 - ② 약간 그렇다
 - ③ 보통이다
 - ④ 그렇지 않은 편이다
 - ⑤ 전혀 그렇지 않다
 - ⑥ 잘 모르겠다
3. 귀하의 해당 업무 분야에 과거 가뭄으로 인한 피해가 발생하였습니까? ()
- ① 매년 또는 빈번히 발생한다
 - ② 지난 5년 안에 발생한 적이 있다
 - ③ 지난 10년 안에 발생한 적이 있다
 - ④ 발생한 적이 없다
 - ⑤ 잘 모르겠다
4. 위 3번 문항에서 과거 가뭄으로 인한 피해가 발생하였다면 어떠한 피해가 발생했는지 구체적으로 서술하여 주시기 바랍니다. (예: 심한 가뭄 기간 동안 산불 발생 빈도 증가, 가로수 고사 등)

※ 길이에 상관없이 자유롭게 서술하여 주시기 바랍니다.

5. 앞으로 귀하의 해당 업무 분야에 가뭄으로 인한 피해 발생을 예상하십니까? ()

- ① 매년 또는 빈번히 발생할 것이다
- ② 향후 5년 안에 발생할 것이다
- ③ 향후 10년 안에 발생할 것이다
- ④ 발생하지 않을 것이다
- ⑤ 잘 모르겠다

가뭄 정보의 활용

※ 현재 기상청에서는 시군구 단위로 표준강수지수, 파머가뭄지수, 강우량 평균치비율 등을 이용하여 가뭄정보를 제공하고 있습니다(<http://www.kma.go.kr/weather/lifenindustry/drought.jsp>).

※ 아래 문항에서 언급한 가뭄 정보는 기상청 뿐 아니라 소속기관, 미디어, 사설 날씨기업 등에서 제공하는 강수량 부족 및 가뭄과 관련된 정보를 모두 포함한 개념입니다.

6. 귀하의 해당 업무에 기상청 또는 다른 기관에서 제공하는 가뭄 정보를 활용하십니까? ()

- ① 그렇다
- ② 그렇지 않다
- ③ 과거에는 활용했으나 현재는 그렇지 않다
- ④ 지금은 활용하지 않지만 앞으로 활용하고 싶다

7. 위 6번 문항에서 1번 가뭄 정보를 활용한다고 답변하신 경우 어느 기관에서 제공하는 정보를 활용하고 또한 해당 정보를 얼마나 신뢰하십니까?

(① 매우 신뢰 - ② 약간 신뢰 - ③ 보통 - ④ 약간 불신 - ⑤ 매우 불신 을 기준으로 적어주시기 바랍니다. 예: 기상청에서 제공하는 파머가뭄지수 - ②)

※ 길이에 상관없이 자유롭게 서술하여 주시기 바랍니다.

8. 위 6번 문항에서 2번 (그렇지 않다)으로 답변하신 경우 가뭄 정보를 활용하지 않는 가장 큰 이유는 무엇입니까? ()

- ① 해당 업무가 가뭄과 관련이 없어서
- ② 가뭄 정보를 쉽게 얻을 수 없어서
- ③ 가뭄 정보는 있으나 업무와 연결 지어 활용하기가 쉽지 않아서
- ④ 가뭄 정보는 있으나 믿을 수 없어서
- ⑤ 기타 _____

9. 위 6번 문항에서 3번 (과거에는 활용했으나 현재는 그렇지 않다) 또는 4번 (지금은 활용하지 않지만 앞으로 활용하고 싶다)으로 답변하신 경우 그 이유는 무엇인지 기술하여 주십시오.

※ 길이에 상관없이 자유롭게 서술하여 주시기 바랍니다.

10. 귀하의 소속기관에 가뭄 담당자가 있습니까? ()

- ① 그렇다
- ② 그렇지 않다
- ③ 기타 _____

11. 위 10번 문항에서 가뭄 담당자가 있다고 답변하신 경우 가뭄 담당자와 가뭄의 영향을 받는 다른 분야(예: 수자원, 농업, 폭염, 산불, 조경 등)의 담당자 간에 의사소통이 이루어지고 있습니까?
()

- ① 매우 그렇다
- ② 약간 그렇다
- ③ 보통이다
- ④ 그렇지 않은 편이다
- ⑤ 전혀 그렇지 않다

12. 위 11번 문항에서 4번 (그렇지 않은 편이다) 또는 5번 (전혀 그렇지 않다)로 답변하신 경우 의사소통이 이루어지고 있지 않은 이유는 무엇이라고 생각하십니까?

※ 길이에 상관없이 자유롭게 서술하여 주시기 바랍니다.

가뭄 대처 현황 및 수요

13. 귀하의 소속기관에 가뭄대책이 수립되어 있습니까? ()

- ① 그렇다
- ② 그렇지 않다
- ③ 기타 _____

14. 위 13번 문항에 가뭄대책이 수립되어 있다고 답변하신 경우 해당 가뭄대책 정보를 얻을 수 있는 인터넷 웹사이트나 가뭄대책 담당자에 대한 정보(성명, 부서, 전화, 이메일 등)를 기술하여 주시기 바랍니다(관련 문서를 이메일 회신과 함께 송부하여 주셔도 좋습니다).

※ 길이에 상관없이 자유스럽게 서술하여 주시기 바랍니다.

15. 귀하의 해당 업무에 가뭄의 영향이 있는 경우 가뭄 발생 정보에 따라 얼마나 미리 대비하십니까?
()

- ① 가뭄이 상습적으로 발생하는 시기에 대하여 평상시 대비
- ② 가뭄 발생을 인지한 순간부터 대비
- ③ 가뭄이 심화되어 해당 업무에 영향/피해가 발생한 순간부터 대처
- ④ 해당사항 없음

16. (귀하가 가뭄 담당자인 경우) 가뭄 발생을 인지한 경우 어떤 대응이 주로 이루어집니까?

- ① 소속기관의 가뭄대책 가이드라인을 따름
- ② 중앙행정기관의 가이드라인을 따름
- ③ 민원이 있는 부분만 해결
- ④ 해당사항 없음
- ⑤ 기타 _____

17. 가뭄 발생을 미리 알 수 있는 가뭄예측 정보가 주어진다면 귀하의 해당 업무에 활용할 의사가 있으십니까? ()

- ① 그렇다
- ② 그렇지 않다
- ③ 법적으로 활용이 요구되는 경우에만 활용할 것이다
- ④ 기상청 등의 중앙기관에서 제공하는 경우 활용할 것이다
- ⑤ 대학/연구원/공공기관에서 제공하는 경우 활용할 것이다
- ⑥ 기타 _____

18. 위 17번 문항에 가뭄예측 정보를 활용할 의사가 있다고 답변하신 경우 다음과 같이 1주일에서 1년의 시간 단위로 가뭄발생 가능성을 미리 알 수 있다면 귀하의 해당 업무에서 영향/피해를 최소화하기 위해 활용할 수 있는 대처 방안은 무엇이 있을 지 구체적으로 제시해 주시기 바랍니다. (예: 가뭄지도 작성 및 보급, 시민 대상 홍보 강화, 기금 재편성 등)

1주일 전	
1개월 전	
3개월 전	
6개월 전	
1년 전	

19. 귀하의 해당 업무에 가장 도움이 될 것으로 생각되는 가뭄감시 및 가뭄예측 정보의 종류 및 형태를 자유롭게 서술하여 주시기 바랍니다.

(예: 폭염을 가중시킬 수 있는 1주일 단위의 가뭄감시정보, 3개월 전의 가뭄예측정보 등)

※ 길이에 상관없이 자유롭게 서술하여 주시기 바랍니다.

20. 가뭄 위험의 정도와 함께 가뭄에 대한 취약성을 함께 분석하는 가뭄 리스크 평가를 수행하면 해당 지역의 가뭄대책을 수립 또는 개선하는데 도움이 되며, 가뭄으로 인해 영향/피해가 심화되는 각종 분야의 취약성 및 리스크도 함께 저감시킬 수 있습니다.

가뭄 리스크 평가를 통해 지역특성을 고려한 맞춤형 가뭄정보(감시 및 예측)가 제공된다면 해당 업무에 활용하실 의사가 있으십니까?

- ① 매우 그렇다
- ② 약간 그렇다
- ③ 보통이다
- ④ 약간 그렇지 않다
- ⑤ 전혀 그렇지 않다

설문에 응해주셔서 대단히 감사합니다.

응답자 기본 정보

※ 다음은 응답자 기본 정보에 대한 문항입니다.

1. 성명

2. 소속 (부서/직위/직책)

3. 귀하의 연령을 선택해 주십시오.

- ① 만 20대 이하
- ② 만 30대
- ③ 만 40대
- ④ 만 50대
- ⑤ 만 60대 이상

3. 귀하의 담당 업무 분야를 간략히 기술하여 주십시오.

4. 해당 분야의 경력 기간을 선택해 주십시오.

- ① 1년 미만
- ② 1년~2년
- ③ 2년~5년
- ④ 5년~10년
- ⑤ 10년 이상

부록 2. 설문조사지

가뭄감시 및 예측정보 활용 2차 설문조사

안녕하십니까? 귀한 시간 내어 설문에 응해 주셔서 대단히 감사합니다.

아태지역 이상기후 감시 및 예측의 중추적 역할을 수행하고 있는 공공연구기관인 APEC 기후센터(<http://www.apcc21.org/kor/index.jsp>)에서는 가뭄 취약성 감소를 통한 가뭄 대응 제고를 위해 「**지속가능한 재난관리를 위한 가뭄감시·예측정보 활용 연구**」를 수행 중에 있습니다.

지난 4월의 1차 설문에 이어 가뭄의 환경적 피해에 대한 대응과 관련된 의사결정이 이루어지는 전국 17개 광역자치체를 대상으로 가뭄감시 및 예측정보 활용 관련 2차 설문조사를 진행하오니 적극적인 참여를 부탁드립니다.

답변해 주신 자료는 본 연구에 소중하게 사용될 것이며 응답하신 내용은 통계법 제33조(비밀의 보호 조항)에 따라 철저히 비밀이 보장되며 순수한 학문적 연구 목적으로만 사용됩니다. 본 조사에 문의사항이 있으시면 아래 연락처로 연락주시기 바랍니다.

설문에 소요되는 시간은 약 **5분**으로 예상되며, 바쁘시더라도 국내 가뭄 취약성 감소 및 대응 제고를 위하여 잠시 시간을 내시어 설문에 응해 주시면 대단히 감사하겠습니다.

2015년 10월

주관기관	APEC 기후센터
연구진	이진영 선임연구원 (051-745-3959)
이메일	이진영 선임연구원 (jyrhee@apcc21.org)
팩스	051-745-3999
우편	부산광역시 해운대구 센텀7로 12 APEC 기후센터, 612-020

※ 본 설문지를 **10월 30일(금)**까지 작성하여 위의 이메일, 팩스, 또는 우편으로 회신하여 주시기 바랍니다.

〈설문조사 작성요령〉

- 선택형 문항의 경우 괄호 안에 해당되는 항목을 기입하여 주시고 (필요시 복수응답가능, 예: (1, 4))
- 서술형 문항의 경우 괄호 또는 빈칸에 길이에 상관없이 자유롭게 서술하여 주시면 됩니다.

가뭄 리스크 (위험도) 평가

1. (가뭄피해 발생빈도) 가뭄은 그 피해유형에 따라 용수공급 피해, 농수축산업 피해, 경제적 피해, 환경적 피해, 사회적 피해로 나눌 수 있습니다. 귀하의 관할 시도에 대하여 가뭄 피해유형별 세부 분류의 발생 빈도를 업무와 관련하여 알고 계시거나 체감하시는대로 표시하여 주시기 바랍니다. (예시)

피해 유형	세부 분류	피해 내용 예시	매우낮음 (30년에 1회)	약간낮음 (10년에 1회)	보통(3년 에 1회)	약간높음 (1년에 1회)	매우높음 (1년에 다회)
피해 유형 예시	세부 분류 사례	피해 내용 예시 사례			0		
피해 유형	세부 분류	피해 내용 예시	매우낮음 (30년에 1회)	약간낮음 (10년에 1회)	보통(3년 에 1회)	약간높음 (1년에 1회)	매우높음 (1년에 다회)
용수공급 피해	생활용수	절수, 잔디 관개 제한					
	발전용수	발전량 감소, 수력발전 비용 증대					
농수축산 업 피해	농업	농업용수 피해, 농작물 품질 및 수확량 저하, 해충 증가					
	축산업	축산물 품질 저하, 가축 질병 증대					
	수산업	어류 품질 저하, 어류 서식지 손상					
	임업	화재위험 증가, 병충해 만연, 삼림지 손상					
경제적 피해	국가재정기관 부담	담보권 행사, 신용리스크 증가, 자본손실					
	수자원 개발 비용	신규/보충 수자원 개발					
	산업 및 여가사업	여가지 손실, 가뭄 관련 전력삭감					
	물류	하천 및 운하 손상, 수운/물 수송 비용					
환경적 피해	생태	생물 다양성 손실, 포식 증대					

피해 유형	세부 분류	피해 내용 예시	매우낮음 (30년에 1회)	약간낮음 (10년에 1회)	보통(3년 에 1회)	약간높음 (1년에 1회)	매우높음 (1년에 다회)
	수질	염수화, 자정능력 저감					
	대기질	먼지, 공해					
	수문	습지 소실, 수위 저하, 지하수 고갈, 지반 침강					
	가시적 조경	먼지, 식생 면적, 심미성 저하					
사회적 피해	갈등 증대	물 이용자 간 충돌, 관리당국 충돌					
	건강 문제	정신 및 육체적 스트레스, 호흡기 질병 증가					

※ 피해 유형 출처: 김연주 외, 2013. 「가뭄 유형별 기후변화 적응정책 연구」. 한국환경정책·평가연구원. 235 p.

2. (가뭄피해 강도) 귀하의 관할 시도의 가뭄과 관련하여 일반적으로 편성되어 집행되는 예산 규모를 고려하여 아래의 각 가뭄 피해유형별 세부 분류에 대해 피해 강도를 업무와 관련하여 알고 계시거나 체감하시는데로 표시하여 주시기 바랍니다.

피해 유형	세부 분류	피해 내용 예시	매우 낮음	약간 낮음	보통	약간 높음	매우 높음
용수공급 피해	생활용수	절수, 잔디 관개 제한					
	발전용수	발전량 감소, 수력발전 비용 증대					
농수축산업 피해	농업	농업용수 피해, 농작물 품질 및 수확량 저하, 해충 증가					
	축산업	축산물 품질 저하, 가축 질병 증대					
	수산업	어류 품질 저하, 어류 서식지 손상					
	임업	화재위험 증가, 병충해 만연, 삼림지 손상					
경제적 피해	국가재정기관 부담	담보권 행사, 신용리스크 증가, 자본손실					
	수자원 개발 비용	신규/보충 수자원 개발					
	산업 및 여가사업	여가지 손실, 가뭄 관련 전력삭감					

피해 유형	세부 분류	피해 내용 예시	매우 낮음	약간 낮음	보통	약간 높음	매우 높음
환경적 피해	물류	하천 및 운하 손상, 수운/물 수송 비용					
	생태	생물 다양성 손실, 포식 증대					
	수질	염수화, 자정능력 저감					
	대기질	먼지, 공해					
	수문	습지 소실, 수위 저하, 지하수 고갈, 지반 침강					
사회적 피해	가시적 조경	먼지, 식생 면적, 심미성 저하					
	갈등 증대	물 이용자 간 충돌, 관리당국 충돌					
	건강 문제	정신 및 육체적 스트레스, 호흡기 질병 증가					

3. 앞서 1번과 2번 문항에서 제시한 가뭄의 피해유형과 세부 분류 중 귀하의 업무와 가장 관련 깊은 항목을 적어주시기 바랍니다. 제시된 항목에 존재하지 않는 경우 새로이 적어주셔도 됩니다.

가뭄 정보의 정확성

9. 관측 자료가 없는 지점에 대하여 가뭄 예측 정보를 얻으려고 할 경우 7번 문항에서 언급한대로 관측 지점의 가뭄 예측 정보를 얻은 후 공간 내삽을 이용하거나, 위성 자료로부터 가뭄 예측 정보를 얻을 수 있습니다. 공간 내삽을 통한 지수를 S라고 하고, 위성 자료로부터 추출한 지수를 N이라고 할 때, 각각 값에 따라 극심한 가뭄, 심한 가뭄, 약한 가뭄 등의 카테고리를 나누고 여러 방법론을 이용하여 가뭄 상태를 사전에 (1개월~6개월) 예측할 수 있습니다.

귀하의 관심 지역에 대하여 두 지수의 예측 정확도가 아래와 같다면 어떤 가뭄 지표를 사용하시겠습니까? S 또는 N으로 표시하여 주시기 바랍니다.

정확도 = 정확히 카테고리를 예측한 경우의 수 / 전체 예측한 경우의 수

N의 정확도

	$0.1 < \leq$	0.2	0.3	0.4	0.5	0.6	≥ 0.7
$0.1 < \leq$							
0.2							
0.3							
S의 정확도	0.4						
0.5							
0.6							
≥ 0.7							

10. 9번 문항과 같이 답한 이유를 적어주시기 바랍니다.

설문에 응해주셔서 대단히 감사합니다.

응답자 기본 정보

※ 다음은 응답자 기본 정보에 대한 문항입니다.

1. 성명

2. 소속 (부서/직위/직책)

3. 귀하의 연령을 선택해 주십시오.

- ① 만 20대 이하
- ② 만 30대
- ③ 만 40대
- ④ 만 50대
- ⑤ 만 60대 이상

3. 귀하의 담당 업무 분야를 간략히 기술하여 주십시오.

4. 해당 분야의 경력 기간을 선택해 주십시오.

- ① 1년 미만
- ② 1년~2년
- ③ 2년~5년
- ④ 5년~10년
- ⑤ 10년 이상

연구보고서 2015-16

위성 및 장기예측자료의 기계학습을 통한 가뭄예측

Drought forecasting based on machine learning of remote sensing and long-range forecast data

이진영



APEC 기후센터

48058 부산광역시 해운대구 센텀7로12

Tel: 051-745-3900 Fax: 051-745-3949

www.apcc21.org



ISBN 979-11-5698-113-8