

APEC기후센터

원격 탐사 자료를 활용한 가뭄 감시 · 예측 기술 개발

이진영 기후변화연구팀

APEC기후센터
연구보고서

원격 탐사 자료를 활용한 가뭄 감시 · 예측 기술 개발

Development of Drought Monitoring and Forecasting
Techniques based on Remotely Sensed Data

이진영 기후변화연구팀

연구보고서 2015-14

발간사

세계 곳곳에서 꾸준히 발생하는 가뭄으로 농작물 피해 및 사회·경제적인 피해에 대한 보도가 끊이지 않고 있습니다. 최근 우리나라에서도 2012년 백여 년 빈도의 전국적인 가뭄에 이어 2013년에는 경상도와 제주도에, 2014년에는 강원도에 극심한 가뭄이 발생하여 많은 경제적 손실이 발생하였습니다. 더욱이 기후변화로 인해 이러한 가뭄 발생의 공간적인 범위나 시간적인 빈도가 변화하고 있습니다.

먼 미래의 가뭄 발생을 전망하기에는 불확실성이 매우 크지만 현재 발생하거나 진행하고 있는 가뭄의 상태를 정확히 파악하고 몇 개월 이후의 상황을 예측하는 것은 가뭄으로 인한 피해를 최소화하기 위해서 반드시 필요한 일입니다. 전통적으로 가뭄의 감시는 기상 관측 자료를 이용하여 이루어졌고, 가뭄의 예측은 최근 들어 기후 모델을 통한 계절 예측 자료를 이용하여 시도되고 있습니다. 그러나 많은 경우 관측 지점이 조밀하지 못하여 공간적으로 연속적인 정보를 얻기 어렵습니다. 이런 경우 위성 영상 자료의 이용이 큰 도움이 됩니다.

본 연구에서는 강수, 토양 수분, 지표면 온도, 식생 지수, 증발산량, 유효 광합성 반사율, 엽면적 지수 등 여러 가뭄 관련 변수의 값을 위성 영상 자료를 통해 얻어내고 여기에 지면 모델의 산출물인 증발산량과 토양 수분 자료, 그리고 북극 진동 등의 대규모 기후 지수를 함께 고려하여 가뭄 감시 및 가뭄 예측 모델을 구축하였습니다. 모델의 구축은 상대적으로 관측 기간이 짧은 위성 영상 자료의 이용을 위해 지식 기반 전문가 시스템을 훈련시키는 기계학습을 활용하여 수행하였습니다.

본 연구에서 개발된 가뭄 감시 및 예측 모델이 미계측 지역의 가뭄 관리를 위해 유용하게 쓰일 수 있기를 기대합니다. 특히 각 모델 구축을 위해 이용된 위성 영상은 별도의 비용 없이 대부분의 지역에 대해 얻을 수 있으므로 대상 지역을 개발도상국으로 확장하여 미관개 지역의 농작물 수확량 예측, 산불 발생 빈도 예측 등에도 활용하여 가뭄 피해를 최소화하는데 도움이 될 수 있을 것으로 기대됩니다.

니다.

본 연구를 수행해 주신 이진영 박사의 노고에 감사를 포함합니다.

2015. 3월
APEC 기후센터 소장
정진승

ABSTRACT

Drought is a regularly occurring natural disaster that affects large numbers of people and causes serious economic damage. The frequency of severe drought is expected to increase in many regions due to the effect of climate change. Despite the high levels of uncertainty on the spatial extent and the temporal timing of the changes monitoring the current conditions and forecasting future drought several months ahead can be valuable to reduce the negative impacts. It is essential to develop a drought early warning system that performs drought monitoring and forecasting, delivers the information to decision-makers in a timely manner, and results in the reduction of the adverse impacts of drought.

Many hydro-meteorological variables should be included for comprehensive drought monitoring and forecasting. Since observation data of the variables are only available where there are dense networks of weather stations, remotely sensed data can be used instead. The purpose of this study is to develop drought monitoring and forecasting techniques using remotely sensed data to be used for regions where observation data is lacking or absent. In this study, hydrometeorological and biophysical variables such as precipitation, land surface temperature, actual and potential evapotranspiration, fractional photosynthetically active radiation (FPAR), leaf area index (LAI), vegetation index, and soil moisture were obtained from remotely sensed data. Several variables of actual and potential evapotranspiration and soil moisture were also derived from land surface model outputs. The large-scale climate indices Arctic Oscillation (AO) and Multivariate ENSO Index (MEI) were integrated with these variables to obtain drought information.

While drought monitoring can be traditionally performed based on drought indices such as Standardized Precipitation Index (SPI) and Palmer Drought Severity Index, a machine learning technique, random forest, was

used in this report in developing drought monitoring and forecasting models to overcome the limitation of short records. Drought-related variables are intertwined through complex feedbacks and couplings between variables, and the interactions between them present challenges in simulations using simple algorithms. Including limited number of variables may lead to inaccurate results of drought conditions. The variables derived from remotely sensed data may be less accurate than observation data since in many cases they are estimated from many algorithms which are physically based but far simpler than real-world interactions. In this case, machine learning techniques can be effectively used to develop drought monitoring and forecasting models since they are not required to explicitly provide or to make assumptions on physical relationships between variables.

In this report, random forest was used to develop drought monitoring and forecasting models with a lead time from 1-6 months. The random forest model is known to produce robust results against outliers and noises. As input variables, variables derived from remotely sensed data (day, night, and mean soil moisture; day, night and mean land surface temperature; actual and potential evapotranspiration; vegetation indices of Normalized Difference Vegetation Index and Enhanced Vegetation Index; precipitation; FPAR; LAI), variables derived from land surface model output (actual and potential evapotranspiration; soil moisture), and large-scale climate indices (AO and MEI) were used. As target variables of the models, SPI and Standardized Precipitation Evapotranspiration Index based on observation data as well as Normalized Difference Water Index (NDWI), Normalized Difference Drought Index (NDDI), and Normalized Multi-band Drought Index (NMDI) based on remotely sensed data were used.

Remotely sensed and observation data were compared for precipitation, land surface temperature, potential evapotranspiration, and soil moisture. For precipitation, TRMM monthly rainfall data produced relatively small

errors and high correlations in the regions with lower latitudes. There also exist high correlations between MODIS land surface temperature and observed air temperature except for daytime land surface temperature during summer. While the MODIS potential evapotranspiration appeared closer to the calculated values based on Penman-Monteith compared to the land surface model output, the soil moisture data from the land surface model output were closer to the KoFlux observation data than the remotely sensed data from AMSR-E sensor. A regression model was built between the AMSR-E derived data and the land surface output for soil moisture, and it was applied to AMSR-E derived soil moisture data to enhance the spatio-temporal resolutions of the remote sensing-based soil moisture data and to reduce estimation errors.

Drought monitoring models with several sets of input variables were developed using random forest, and their performances were evaluated. A drought monitoring model with all available variables performed the best. Random forest provides internal estimates of generalization errors and relative variable importance using out-of-bag procedures. Precipitation turned out to be exceptionally important for target variables based on observation data. For target variables based on remotely sensed data, the relative importance of large-scale climate indices of AO and MEI were very high. It indicates that drought conditions in the study area are very much affected by teleconnections. The correlations between input data and output of models for NDWI6 (NDWI with MODIS band 6) and NDWI7 (NDWI with MODIS band 7) were especially high, suggesting them as useful target variables for drought monitoring.

Drought forecasting models with a lead time from 1-6 months were developed using input variables previously selected for drought monitoring models, and their performances were evaluated. Forecasting errors tend to increase with longer lead times, while the errors tend to decrease with more

past months of the target variable used as input data. The number of past months of the target variable affects forecasting errors more than the lead time used. Target variables NDDI5 (NDDI with MODIS band 5) and NDDI6 (NDDI with MODIS band 6) were excluded because of their unstable error levels. Drought forecasting models developed performed well for other target variables. The target variable itself with past values showed the highest relative importance among input variables. Large-scale climate indices of AO and MEI also have large relative importance as with drought monitoring models. The relative importance of vegetation-related variables of FPAR, LAI, NDVI, and EVI were moderately low.

The drought forecasting models developed using target variables based on remote sensing were applied for the crop yield data of highland bok choy and radish as well as frequency and area data of forest fire. Rather than developing models for directly forecasting crop yields or forest fire frequencies, the statistical correlations between crop yields or forest fire frequencies and remote sensing-based target variables for administrative districts of the study area were obtained, and then the performance of the models for the administrative districts with high correlations was evaluated on the remote sensing-based target variables.

Among nine SI-DO and six SI-KUN-KU used for the crop yield analyses, crop yield data and remote sensing-based target variables showed high correlations in two SI-DO of Jeollabuk-do and Gyeongsangnam-do and four SI-KUN-KU of Yeongwol-gun, Yanggu-gun, Inje-gun, and Gangneung-si for highland bok choy. High correlations were observed for highland radish in three SI-DO of Gangwon-do, Gyeongsangbuk-do, and Gyeongsangnam-do, as well as five SI-KUN-KU of Taebaek-si, Samcheok-si, Yeongwol-gun, Inje-gun, and Gangneung-si. Relatively low correlations were observed in some cases since there are many other factors affecting crop yield. The remote sensing-based target variables, however, can be used as good proxies for

the cases previously listed.

The performance of a drought forecasting model using NDWI7 as a target variable with 6-month lead time was especially good in Yeongwol-gun for highland bok choy, as well as a model using NDWI7 with 1-month lead time in Gyeongsangnam-do, and a model using NDWI6 with 2-month and 6-month lead times respectively in Yeongwol-gun. For highland radish, models using NDWI6 with 6-month lead time in Gyeongsangnam-do and 2-month lead time in Gyeongsangbuk-do showed high correlations with crop yield.

The correlations between the remote sensing-based target variables and the forest fire frequency data, fire area data, and the fire area ratio to the total forest area were tested and only forest fire frequency showed good correlations. The high correlations were only observed when analyzed for SI-DO level, and for NDWI and NDDI among many target variables. The correlations were especially high in Busan, Ulsan, Chungcheongbuk-do, Jeollanam-do, Gyeongsangbuk-do, and Gyeongsangnam-do, and the performance of the drought forecasting models for those administrative districts for forecasting the target variables were excellent in all cases.

The developed drought monitoring and forecasting models based on a machine learning methodology, random forest, can be used by a variety of end users. The models developed with target variables based on observation data with time scales from 1-9 months can be applied to many different types of drought. The models are suited to provide general drought information for many drought-affected sectors by being integrated in the drought information system operated by Korea Meteorological Administration. Since the NDWI and NDDI were developed targeting drought impacts on vegetation, and the NMDI was developed for monitoring soil and vegetation water conditions, the models developed with target variables of remote sensing-based drought indices of NDWI, NDDI, and NMDI can provide the most valuable information

for agricultural drought (soil moisture drought). As applied with highland crop yield data or forest fire frequency data, the models can provide valuable information for farmers and forest managers by being integrated with drought information system operated by the Ministry of Agriculture, Food and Rural Affairs or the Korea Forest Service.

Since remote sensing has not yet been actively used for drought monitoring and forecasting, the integration of remote sensing techniques to existing drought information system or early warning systems operated by many government organizations offers a large potential. An environment where drought information derived from remote sensing can be used by end-users with confidence should be built soon, by performing a range of case studies for evaluating remote sensing data and for improving their reliability.

목차

원격 탐사 자료를 활용한 가뭄 감시·예측 기술 개발

발간사	i
ABSTRACT	iii
1. 서론	1
1.1 연구 배경 및 목적	1
1.1.1 가뭄의 정의	1
1.1.2 가뭄 피해 및 기후변화의 영향	2
1.1.3 연구의 목적	10
1.2 연구 범위	11
1.2.1 공간적 범위	11
1.2.2 시간적 범위	12
1.2.3 내용적 범위	13
2. 원격 탐사를 활용한 가뭄 감시	14
2.1 전통적 가뭄 감시	14
2.1.1 전통적 가뭄 지수	15
2.2 가뭄 관련 변수의 원격 탐사	19
2.2.1 강수량	20
2.2.2 증발산량	20
2.2.3 토양수분	20
2.2.4 식생	21
2.3 원격 탐사 기반 가뭄 지수	22
2.4 변수 융합을 통한 맞춤형 가뭄 감시	24
3. 원격 탐사를 활용한 가뭄 예측	27
3.1 가뭄 예측 모델링	27
3.1.1 회귀 모델	28

3.1.2 시계열 모델	29
3.1.3 인공 신경망 모델	30
3.1.4 의사결정분지도 모델	31
3.1.5 장기 예측 모델	32
3.2 가뭄 예측을 위한 원격 탐사의 활용	33
4. 자료 및 연구 방법론	34
4.1 위성 영상 자료	35
4.1.1 강수량	35
4.1.2 지표면온도	37
4.1.3 증발산량	39
4.1.4 광합성유효복사율	41
4.1.5 엽면적지수	41
4.1.6 식생지수	42
4.1.7 토양수분	42
4.2 원격 탐사 기반 가뭄 지수	45
4.2.1 표면반사율	45
4.2.2 원격 탐사 기반 가뭄 지수의 산정	46
4.3 관측 자료	46
4.3.1 정관기상관측자료	46
4.3.2 KoFlux 자료	46
4.3.3 기후 지수	51
4.3.4 고랭지 채소 수확량	52
4.3.5 산불 발생 빈도 및 피해 면적	53
4.4 관측 자료 기반 가뭄 지수	54
4.4.1 SPI의 산정	54
4.4.2 SPEI의 산정	55
4.5 지면모델 자료	55
4.5.1 지면모델 개요	55
4.5.2 NOAH-LSM	60
4.5.3 지면모델 결과와 위성 영상 자료의 결합	61
4.6 기계학습을 통한 가뭄 감시 및 예측 모델 개발	61
4.6.1 인공지능과 전문가시스템	61

4.6.2 기계학습	62
4.6.3 Random Forests	64
4.6.4 모델의 개발	65
4.6.5 모델의 적용	68
5. 결과 및 토의	75
5.1 위성 영상 추출 변수의 검증	75
5.1.1 강수량	75
5.1.2 지표면온도	81
5.1.3 증발산량	93
5.1.4 토양수분	103
5.2 가뭄 감시 및 예측 모델 선정	124
5.2.1 가뭄 감시 모델의 선정	124
5.2.2 가뭄 예측 모델의 선정	133
5.3 가뭄 예측 모델의 검증	147
5.3.1 고랭지 채소 수확량	147
5.3.2 산불 발생 빈도 및 피해 면적	151
6. 결론 및 시사점	155
6.1 연구 결과의 요약	155
6.2 연구의 결론 및 시사점	159
6.3 활용 제안	160
6.4 향후 가뭄 연구 제안	162
REFERENCES	163

표 목 차

Table 1-1. Drought events in South Korea from 1962~1994.	7
Table 2-1. Classification categories of SPI (unitless).	18
Table 4-1. Weather stations used in the Korean Peninsula.	47
Table 4-2. List of KoFlux sites used in this study.	50
Table 4-3. Input variables of random forest models for drought monitoring.	67
Table 5-1. Statistics between TRMM PRCP and observed monthly precipitation averaged over 106 weather stations.	76
Table 5-2. Statistics between MODIS LST and observed monthly temperature averaged over 106 weather stations.	82
Table 5-3. Statistics between MODIS PET and calculated PET based on Penman-Monteith averaged over 35 weather stations.	94
Table 5-4. Statistics between MODIS PET and calculated PET based on Penman-Monteith averaged over 35 weather stations.	98
Table 5-5. Statistics between monthly averaged AMSR-E A_SM and KoFlux Soil Moisture for two KoFlux Sites.	105
Table 5-6. Statistics between monthly averaged AMSR-E D_SM and KoFlux Soil Moisture for two KoFlux Sites.	106
Table 5-7. Statistics between monthly averaged AMSR-E SM_MEAN and KoFlux Soil Moisture for two KoFlux Sites.	107
Table 5-8. Statistics between daily AMSR-E A_SM and KoFlux Soil Moisture for two KoFlux Sites.	110
Table 5-9. Statistics between daily AMSR-E D_SM and KoFlux Soil Moisture for two KoFlux Sites.	111
Table 5-10. Statistics between daily AMSR-E SM_MEAN and KoFlux Soil Moisture for two KoFlux Sites.	112
Table 5-11. Statistics between monthly NOAA-LSM Soil Moisture and KoFlux Soil Moisture for two KoFlux Sites.	115
Table 5-12. Statistics between daily NOAA-LSM Soil Moisture and KoFlux Soil Moisture for two KoFlux Sites.	117
Table 5-13. Statistics between monthly AMSR-E SM_MEAN and NOAA-LSM Soil Moisture.	122

Table 5-14. Statistics between converted AMSR-E Soil Moisture and KoFlux Soil Moisture for two KoFlux Sites.	124
Table 5-15. RMSE of random forest models for target variables.	126
Table 5-16. Coefficient of determination (R^2) of random forest models for target variables.	127
Table 5-17. Relative importance of input variables for each target variable based on observed data (Unit: %).	128
Table 5-18. Relative importance of input variables for each target variable based on remotely sensed data (Unit: %).	129
Table 5-19. RMSE of random forest models for target variables based on observation data with various lead times and numbers of past months used as input.	135
Table 5-20. RMSE of random forest models for target variables based on remotely sensed data with various lead times and numbers of past months used as input.	136
Table 5-21. Coefficient of determination (R^2) values of random forest models for target variables based on observation data with various lead times and numbers of past months used as input.	138
Table 5-22. Coefficient of determination (R^2) values of random forest models for target variables based on remotely sensed data with various lead times and number of past months used as input.	139
Table 5-23. RMSE of random forest models for target variables based on observation data with various lead times and numbers of past months used as input.	141
Table 5-24. RMSE of random forest models for target variables based on remotely sensed data with various lead times and numbers of past months used as input.	142
Table 5-25. Coefficient of determination (R^2) values of random forest models for target variables based on observation data with various lead times and numbers of past months used as input.	143

Table 5-26. Coefficient of determination (R^2) values of random forest models for target variables based on remotely sensed data with various lead times and numbers of past months used as input.	144
Table 5-27. Relative importance of input variables for each target variable based on remotely sensed data (Unit: %).	146
Table 5-28. Statistics between crop yield and target variable data (Unit for RMSE and MAE is ka/10ha).	148
Table 5-29. Statistics between observed and estimated target variable data (RMSE and MAE are unitless).	150
Table 5-30. Statistics between forest fire frequency and target variable data (Unit for RMSE and MAE is event/month, all p-value < 0.001).	153
Table 5-31. Statistics between observed and estimated NDWI7 in Gyeongsangnam-do (RMSE and MAE are unitless; all p-value < 0.001).	154

그림 목차

Figure 1-1. Top ten drought disasters from 1900~2014 (Source: EM-DAT, 2014).	4
Figure 1-2. Top ten drought disasters from 1985~2014 (Source: EM-DAT, 2014).	5
Figure 1-3. Study area – the Korean Peninsula.	13
Figure 2-1. Processes contributing to (a) soil moisture-temperature, and (b) soil moisture-precipitation coupling and feedback loops (adapted from Seneviratne et al., 2010, recited in Rhee et al., 2015).	25
Figure 2-2. September P-PET percentiles in Upper Namhan-gang basin (APEC Climate Center, 2013; recited in Rhee et al., 2015).	25
Figure 3-1. Components for drought forecasting (adapted from Mishra and Singh, 2011).	29
Figure 3-2. Components of a typical Artificial Neural Network model (Jensen, 2005).	30
Figure 3-3. A decision-tree expert system with a rule and conditions to test a hypothesis (Jensen, 2005).	32
Figure 4-1. Diagram of the TRMM satellite and instruments (Source: pmm.nasa.gov).	36
Figure 4-2. Global Precipitation Measurement Mission Core Observatory Satellite (Source: gpm.nasa.gov).	37
Figure 4-3. Terra spacecraft with instruments (Source: NASA).	38
Figure 4-4. Sample Terra orbital tracks on 15 February 2005 (Source: NASA Distributed Active Archive Center at National Snow and Ice Data Center; nsidc.org).	39
Figure 4-5. Flow chart of the MODIS evapotranspiration algorithm (Source: Mu et al., 2011).	40
Figure 4-6. Aqua’s instruments (Source: NASA).	44
Figure 4-7. Sample Aqua orbital tracks on 15 February 2005 (Source: NASA Distributed Active Archive Center at National Snow and Ice Data Center; nsidc.org).	44

Figure 4-8. Two KoFlux sites used in this study.	50
Figure 4-9. Multivariate ENSO Index (Source: www.esrl.noaa.gov/psd/enso/mei).	52
Figure 4-10. Four SI-DO and 18 SI-KUN-KU provinces with crop yield data (highland bok choy and radish).	53
Figure 4-11. Illustration of a first-generation land surface model (Source: Sellers et al., 1997; recited in Pitman, 2003).	56
Figure 4-12. Illustration of a second-generation land surface model (Source: Sellers et al., 1997; recited in Pitman, 2003).	57
Figure 4-13. Illustration of a third-generation land surface model (Source: Pitman, 2003).	58
Figure 4-14. Levels of land surface modelling approaches (Source: Pitman, 2003).	58
Figure 4-15. VIC-3L model.	60
Figure 4-16. Components of rule-based expert system (Source: Jensen, 2005).	63
Figure 4-17. Locations of grids used: Input data for 5×5km grids located in highland areas are averaged for each SI-DO or SI-KUN-KU.	71
Figure 4-18. Forest Fire Risk Prediction System of Korea Forest Service.	73
Figure 4-19. Locations of grids used: Input data for 5×5km grids located in forest areas are averaged for each SI-DO or SI-KUN-KU.	74
Figure 5-1. RMSE between TRMM PRCP and observed monthly precipitation during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: mm/month).	77
Figure 5-1. (Continued). RMSE between TRMM PRCP and observed monthly precipitation during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: mm/month).	78
Figure 5-2. Pearson's r between TRMM PRCP and observed monthly precipitation during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	79

Figure 5-2. (Continued). Pearson's r between TRMM PRCP and observed monthly precipitation during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: mm/month).	80
Figure 5-3. Time-series of TRMM PRCP and observed monthly precipitation from during 2001~2012 in Daejeon station (ID: 133) (Unit: mm/month).	80
Figure 5-4. RMSE between LSTDAY and observed monthly maximum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: degree C).	83
Figure 5-4. (Continued). RMSE between LSTDAY and observed monthly maximum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: degree C).	84
Figure 5-5. Pearson's r between MODIS LSTDAY and observed monthly maximum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	85
Figure 5-5. (Continued). Pearson's r between MODIS LSTDAY and observed monthly maximum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	86
Figure 5-6. RMSE between LSTNIGHT and observed monthly minimum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: degree C).	87
Figure 5-6. (Continued). RMSE between LSTNIGHT and observed monthly minimum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: degree C).	88
Figure 5-7. Pearson's r between MODIS LSTNIGHT and observed monthly minimum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	89
Figure 5-7. (Continued). Pearson's r between MODIS LSTNIGHT and observed monthly minimum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	90
Figure 5-8. Time-series of MODIS LSTDAY and observed monthly maximum temperature in Nampo station (ID: 47060) (Unit: °C).	91

Figure 5-9. Time-series of MODIS LSTNIGHT and observed monthly minimum temperature in Seoul station (ID: 108) (Unit: °C).	92
Figure 5-10. Time-series of MODIS LST_MEAN and observed monthly mean temperature in Sangju station (ID: 137) (Unit: °C).	92
Figure 5-11. Pearson's r between MODIS PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	94
Figure 5-11. (Continued). Pearson's r between MODIS PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	95
Figure 5-12. Time-series of MODIS PET and calculated PET in Inje station (ID: 211) (Unit: mm/month).	97
Figure 5-13. RMSE between NOAH-LSM PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: mm/month).	100
Figure 5-13. (Continued). RMSE between NOAH-LSM PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: mm/month).	101
Figure 5-14. Pearson's r between NOAH-LSM PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	101
Figure 5-14. (Continued). Pearson's r between NOAH-LSM PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	102
Figure 5-15. Time-series of NOAH-LSM PET and calculated PET in Gangneung station (ID: 105) (Unit: mm/month).	103
Figure 5-16. Time-series of KoFlux Soil Moisture and monthly averaged AMSR-E (a) A_SM, (b) D_SM, and (c) SM_MEAN in Haenam station (Unit: g/cm ³).	108
Figure 5-16. (Continued). Time-series of KoFlux Soil Moisture and monthly averaged AMSR-E (a) A_SM, (b) D_SM, and (c) SM_MEAN in Haenam station (Unit: g/cm ³).	109
Figure 5-17. Time-series of KoFlux Soil Moisture and daily averaged AMSR-E (a) A_SM, (b) D_SM, and (c) SM_MEAN in Haenam station (Unit: g/cm ³).	113

Figure 5-18. Time-series of monthly NOAA-LSM Soil Moisture and KoFlux Soil Moisture in Haenam station (Unit: g/cm ³).	116
Figure 5-19. Time-series of daily aveaged NOAA-LSM Soil Moisture and KoFlux Soil Moisture in Haenam station (Unit: g/cm ³).	118
Figure 5-20. Pearson's r between AMSR-E Soil Moisture and NOAA-LSM Soil Moisture during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	120
Figure 5-20. (Continued). Pearson's r between AMSR-E Soil Moisture and NOAA-LSM Soil Moisture during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).	121
Figure 5-21. Scatterplot between observed SPI1 and estimated SPI1 based on a random forest model.	131
Figure 5-22. Scatterplot between observed SPEI1 and estimated SPEI1 based on a random forest model.	131
Figure 5-23. Scatterplot between observed NDWI6 and estimated NDWI6 based on a random forest model.	132
Figure 5-24. Scatterplot between observed NDWI7 and estimated NDWI7 based on a random forest model.	132
Figure 5-25. RMSE of random forest models for target variables based on observation data with various lead times and fixed number of past months (three months) used as input.	134
Figure 5-26. Coefficient of determination (R^2) values of random forest models for target variables based on observation data with various lead times and fixed number of past months (three months) used as input.	137
Figure 5-27. Scatterplot between observed July NDWI7 and estimated July NDWI7 based on a random forest model with 6-month lead time in Yeongwol-gun.	151
Figure 5-28. Scatterplot between observed NDWI7 and estimated NDWI7 for all months based on a random forest model with 6-month lead time in Gyeongsangnam-do.	154

1. 서론

1.1 연구 배경 및 목적

1.1.1 가뭄의 정의

우리나라 법령 중 본문에 가뭄에 대한 항목을 포함하고 있는 것은 모두 22개 법령으로, 이 중 법률은 안전행정부 소관의 「재난 및 안전관리 기본법」, 소방방재청 소관의 「자연재해대책법」, 환경부 소관의 「수도법」, 농림축산식품부, 해양수산부 공동 소관의 「농어업재해대책법」, 해양수산부 소관의 「내수면어업법」이 있다. 가뭄 정책에 있어 어느 나라보다 앞서 나가고 있는 미국의 경우 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 이루었다. 그러나 우리나라의 경우 아직 가뭄에 대한 개별 법령은 존재하지 않는다.

우리나라 가뭄 관련 법률 중 「재난 및 안전관리 기본법」과 「농어업재해대책법」은 각각 자연재난과 농업재해의 일종으로 가뭄을 포함시키고 있으며 이 밖의 법령에서는 가뭄 대책의 법적 근거가 마련되어 있지만 이들 법령에 가뭄에 대한 정의는 포함되어 있지 않다. 반면 미국의 「NIDIS Act」는 가뭄의 포괄적인 정의를 포함하고 있는데, 이는 일반적인 가뭄 연구에서의 정의와 일맥상통한다. 여기에서 가뭄은 “지표수나 지하수 공급(하천, 습지, 토양수분, 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”과 같이 정의되고 있다(NIDIS Act, 2006).

가뭄은 강수량의 부족에 기인하나, 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(1987)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 서술하였으

며, 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.

우리나라에서 법령으로 규정된 가뭄의 정의가 존재하지 않으므로, 본 연구에서는 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)에서 사용한 가뭄의 정의를 차용하였다. 이에 따라 가뭄은 “심각한 수문학적 불균형을 야기할 정도로 장기간 동안 지속되는 비정상적으로 건조한 날씨”로 정의할 수 있으며, 여기서 가뭄의 원인이 되는 비정상적인 강수량 부족 기간을 기상학적 가뭄으로 정의할 수 있다(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 토양수분 가뭄(농업적 가뭄) 및 하천, 호수, 지하수 등 수위가 비정상적으로 낮아지는 수문학적 가뭄이 발생할 수 있다(Seneviratne et al., 2012).

1.1.2 가뭄 피해 및 기후변화의 영향

1.1.2.1 세계 가뭄 피해

재난역학연구소(Center for Research on the Epidemiology of Disasters; CRED)의 국제 재난 데이터베이스(EM-DAT)를 이용하여 1900~2014년 동안 사망자수, 피해자수, 경제적 손실로 각각 상위 10건의 가뭄 재난 발생과(Figure 1-1) 최근 30년간 1985~2014년 동안의 상위 10건의 가뭄 재난 발생을(Figure 1-2) 살펴보았다.

가뭄으로 인한 사망자수를 보면, 1900~2014년 동안의 가장 극심한 10건의 가뭄 사상은 모두 1983년까지의 기간에 발생했다는 것을 알 수 있다(Figure 1-1a). 1985~2014년 동안의 가장 극심한 10건의 가뭄 사상으로 인한 사망자수는 그 이전과 비교

하여 크게 줄어들었는데(Figure 1-2a), 이 중 가장 많은 사망자를 기록한 1991년 중국 가뭄의 경우 사망자수는 2,000명으로 이는 1973년 에티오피아 가뭄으로 인한 100,000명 사망의 1/50에 해당하는 기록이다(Figures 1-1a, 1-2a). 그러나 가뭄으로 인한 피해자수는 1900~2014년 동안의 가장 극심한 10건의 가뭄 사상 중 7건이 1987년 이후에 발생하여(Figure 1-1b), 과거에 비해 훨씬 증가하였음을 알 수 있다. 이는 실제 피해자수의 증가를 나타낼 수도 있고 과거보다 피해 집계가 잘 이루어졌기 때문일 수도 있다. 최근 30년이나 전체 기간 모두 대부분의 극심한 가뭄 사상이 많은 인구가 분포하는 중국과 인도에서 발생하였으며 특히 인도의 피해자수가 매우 높음을 알 수 있다(Figures 1-1b, 1-2b).

가뭄으로 인한 경제적 피해도 피해자수와 비슷한 양상을 보이는데, 1900~2014년 동안 발생한 가장 경제적 피해가 심한 10건의 가뭄 사상 중 한 건을 제외하고는 모두 1990년 이후에 발생하였다(Figures 1-1c, 1-2c). 가장 극심한 경제적 피해를 기록한 10건의 가뭄 중 4건의 미국에서 발생하였으며, 경제적 피해로는 2012년 미국에서 발생한 가뭄에 이어 1994년 중국에서 발생한 가뭄의 피해가 크게 나타났다(Figures 1-1c, 1-2c).

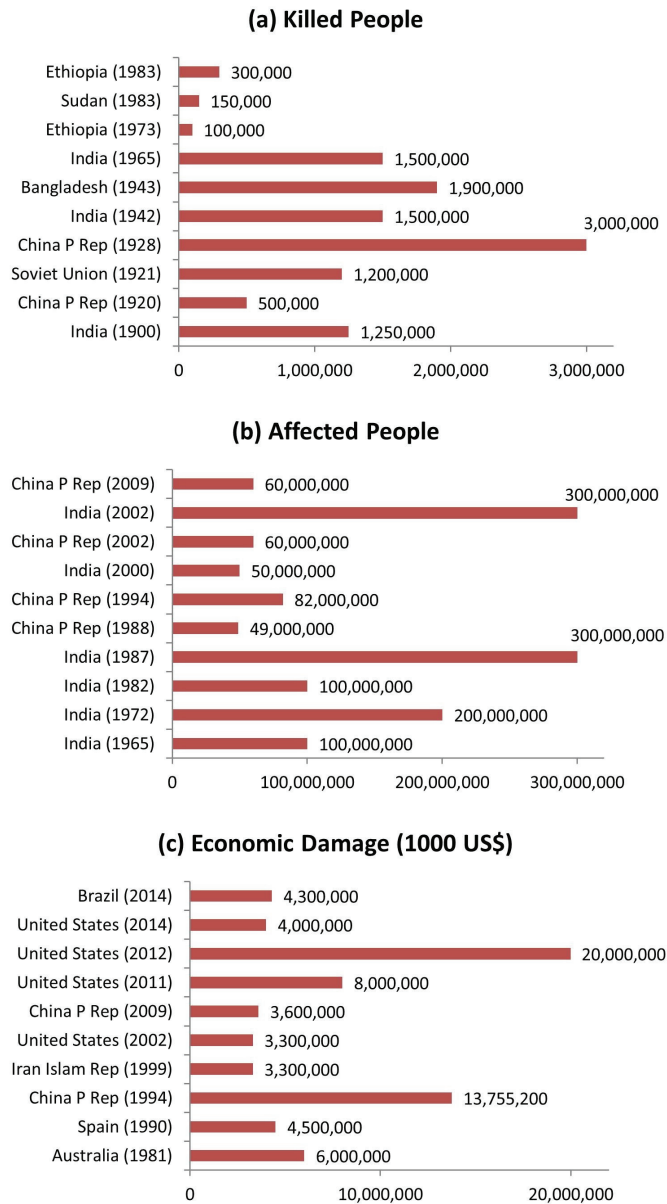


Figure 1-1. Top ten drought disasters from 1900~2014 (Source: EM-DAT, 2014).

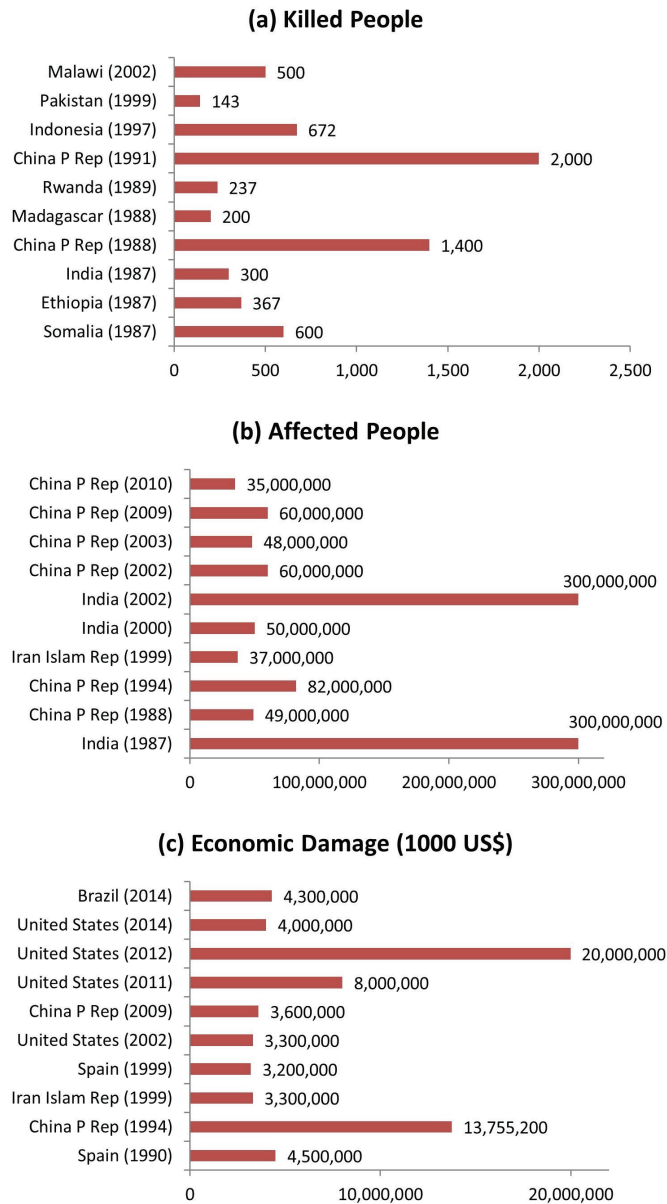


Figure 1-2. Top ten drought disasters from 1985~2014 (Source: EM-DAT, 2014).

세계적인 가뭄 재난 발생의 통계를 볼 때, 사망자수는 크게 감소하였으나 피해자수와 경제적 피해에 있어 최근 30년 동안의 가뭄 사상이 매우 높은 기록을 보였음을 알 수

있다. 특히 인구가 많은 중국과 인도 등의 국가에서 매우 많은 가뭄 피해자가 발생하였으며, 미국의 경우 매우 큰 경제적 피해를 보여 선진국도 가뭄으로 인한 경제적 손해를 피해갈 수 없음을 알 수 있다.

1.1.2.2 국내 가뭄 피해

가뭄은 「재난 및 안전관리 기본법」 제3조에 태풍, 홍수, 호우, 강풍, 풍랑, 해일, 대설, 낙뢰, 지진, 황사, 조류 대 발생, 조수 등과 함께 자연재난으로 지정되어 있다. 그러나 소방방재청에서 매년 발간하는 「재해연보」는 호우, 태풍, 폭풍, 폭풍설 등의 풍수해로 인한 인명피해와 재산피해 통계만을 제공하고 있어 국내 가뭄으로 인한 정량적 피해 통계 자료는 희소하다. 다만 과거의 문헌 자료와 기록을 바탕으로 가뭄이 발생한 시기와 심도, 피해 정도를 도출할 수 있는데, 대부분 농산물의 흉작으로 인한 피해에 해당한다.

1900년대의 우리나라 4대 가뭄으로 1939년, 1968년, 1978년, 1982년에 발생한 가뭄을 들 수 있는데(방재연구소, 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 추정할 수 있다(농림부, 2000). 이 중 1939년 가뭄의 심도가 가장 깊었으며 특히 낙동강 유역이 심한 물 부족을 나타내었고, 1968년의 가뭄은 영산강, 낙동강 유역의 물 부족을 보였다(방재연구소, 2009). 가뭄 피해에 대한 상세한 통계는 없으나 농림수산부에서 1978년과 1983년에 발간된 「한해극복지」를 바탕으로 가뭄 면적 및 가뭄 대책비 정보를 도출하고 1960년부터 1994년까지 발생한 9건의 가뭄에 대해 가뭄 피해액을 산정하였는데, 이를 통해 그 피해 규모를 가늠해볼 수 있다(Table 1-1).

2000년대에도 여러 차례 가뭄이 발생하였는데, 2001년 봄 중국에서 이동한 이동성 고기압의 영향을 받아 중부지방을 중심으로 심한 가뭄이 발생하였다(방재연구소, 2009). 이로 인하여 3~5월의 강수량이 평년의 20~40% 수준에 머물렀으며 경기, 강원, 충북, 충남, 경북의 19,000ha 면적의 논밭에 피해가 발생하였다(방재연구소, 2009). 2008~2009년에는 평년보다 적은 강수량으로 가뭄이 발생하였으며 특히 강원

도 태백시에 생활용수 부족으로 인한 피해가 있었다(방재연구소, 2009). 2012년에는 봄철 전국 평균 평년의 28.7%에 해당하는 강수량으로 전국적으로 심한 가뭄이 발생하여 농업용수 및 생활용수를 공급하는 등 가뭄 극복을 위한 노력이 있었다(한국환경정책평가연구원, 2013).

Table 1-1. Drought events in South Korea from 1962~1994.

Year	Monthly Precipitation (mm, May-July)	Reservoir Storage Ratio (%)	Affected Area (ha)	Economic Damage Cost (1,000,000 KRW)	Compensation Cost (1,000,000 KRW)
1967	307.4	5	420,547	626,615	5,758
1968	122.2	4	470,422	700,928	5,558
1976	368.7	37	28,218	42,044	2,548
1977	287.5	29	60,222	89,370	13,920
1978	717.6	13	-	-	37,201
1981	658.2	46	145,457	216,730	51,783
1982	300.8	27	231,244	344,533	48,257
1992	392.4	23	31,523	46,969	21,400
1994	231.3	15	231,569	249,281	61,866

※ Note: The base year for the cost estimation is the year 1992 for 1967~1992, and the year 1994 for 1994.

최근의 국내 가뭄 피해를 파악하고자 2013년 7월 29일부터 2014년 7월 22일까지 국내 주요 방송 및 일간지의 가뭄 관련 기사를 살펴보았다. 전체 기간 중 2013년 11월과 12월을 제외하고는 모든 달에 우리나라 곳곳에 심각한 가뭄의 발생과 피해에 대한

기사가 있었다. 2014년 1월에 실린 강원도 지역의 심각한 겨울 가뭄 해갈과 관련된 기사(연합뉴스, 2014년 1월 20일) 및 전라남도 도서 지역의 극심하게 지속되는 겨울 가뭄에 대한 기사(광주일보, 2014년 1월 27일; 광남일보 2014년 1월 27일)로 미루어보아 가뭄에 대한 기사가 없던 2013년 겨울에도 심한 겨울 가뭄을 겪는 지역이 있었음을 알 수 있다.

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 기록하여 폭염과 더불어 농축산 농가에 큰 피해가 있었다(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 22일까지 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 “마른 장마”로 중부지방을 중심으로 7월 중순까지 극심한 가뭄 피해가 발생하였다(한국일보, 2014년 7월 14일; 조선일보 2014년 7월 14일).

1.1.2.3 기후변화의 영향

기후변화로 인해 전지구 규모의 에너지 순환에 변화가 진행되고 있으므로 가뭄 사상의 발생에도 변화가 예상된다. IPCC 4차 보고서(AR4)에는 Palmer Drought Severity Index(PDSI)를 이용하여 1970년대 이후로 극심하게 건조한 지역의 면적이 두 배 이상

으로 확대되었다는 결과가 있었으나(Dai et al., 2004), 이는 PDSI가 강수 보다는 온도에 매우 민감하게 영향을 받기 때문이다. 관측 자료 기반의 지면모델을 활용하여 1950~2000년에 대해 토양수분을 살펴본 바에 의하면 전지구적으로 가뭄의 지속기간, 심도 및 강도가 감소할 것으로 예상되며 다만 지역적인 편차가 크게 나타날 것으로 보인다(Sheffield and Wood, 2008). 아시아 지역에 대해서는 공간적인 차이가 크게 나타나므로 가뭄 발생의 변화에 대한 불확실성이 높으나 여러 가뭄 지수를 이용한 많은 연구가 동아시아 지역에서 20세기 후반에 가뭄이 심화되었음을 보이고 있다(Seneviratne et al., 2012).

전지구 기후모델(Global Climate Model; GCM)을 이용하여 미래의 가뭄 발생에 대한 연구도 활발히 이루어지고 있다. Burke and Brown(2008)은 Coupled Model Intercomparison Project 3(CMIP3) 다중모델앙상블 자료를 기반으로 Standardized Precipitation Index(SPI), PDSI, Precipitation Potential Evaporation Anomaly(PPEA), Soil Moisture Anomaly(SMA) 지수를 이용하여 미래 변화를 분석하였는데, SPI를 제외한 다른 지수들은 모두 5~45%의 가뭄 면적의 증가를 전망하였다. 다른 지수들과 달리 SPI가 가뭄 면적의 변화를 전망하지 않은 것은 강수량만을 고려하여 대기의 수분 수요를 고려하지 않았기 때문이다. IPCC SREX에서는 CCD와 SMA를 이용하여 전지구 가뭄 변화를 분석하였는데, 유럽 남부 및 지중해 지역, 유럽 중부, 북아메리카 중부, 중앙아메리카 및 멕시코, 브라질 북동부, 남아프리카 등 넓은 지역에서는 기존의 연구와 마찬가지로 두 지수 모두 연간 규모로 가뭄의 증가를 전망하였다(Seneviratne et al., 2012). 반면 동북아시아에서는 두 지수가 서로 다른 전망을 보였는데, CDD는 가뭄 발생의 감소를 전망하였으나 SMA가 같은 결과를 보이지는 않았다(Seneviratne et al., 2012). 수문학적 가뭄은 기준점 이하의 유량으로 정의한 가뭄 일수로 살펴볼 수 있는데, 중국 내륙 등의 지역에서는 강수량 증가가 예상되나 다른 여러 요인의 변화와 함께 가뭄의 증가가 전망되었다(Hirabayashi et al., 2008).

전지구적인 규모로 수행된 여러 연구의 결과를 볼 때, 우리나라를 포함한 동아시아

지역에서는 20세기까지 가뭄의 발생이 심화되었음을 알 수 있었다. 그러나 미래 가뭄 발생의 전망에 있어서는 일치하는 결론을 도출할 수 없었다. 이는 동아시아 지역의 공간적인 편차가 매우 크기 때문이며(Seneviratne et al., 2012) 따라서 불확실성 또한 매우 크다고 할 수 있다. CMIP5와 같은 최근의 다중모델 결과물과 역학적 또는 통계적 규모내림(downscaling)을 결합한 동아시아 지역, 특히 우리나라에 중점을 둔 지역적 연구가 요구된다.

1.1.3 연구의 목적

앞서 살펴본 바와 같이 가뭄으로 인한 피해는 과거로부터 지속적으로 존재하며 우리나라를 포함한 동아시아 지역 가뭄 발생의 미래 전망은 매우 불확실하다. 따라서 구조적인 대책을 통해 가뭄 발생에 사전에 대비하는 것과 동시에 가뭄조기경보시스템과 같은 비구조적 수단을 통해 이미 발생하였거나 발생할 가능성이 있는 가뭄에 즉각적으로 대처할 수 있는 역량을 구축하는 것이 매우 중요하다. 가뭄조기경보시스템을 이용하여 가뭄 평가 및 감시를 수행, 정확하고 시의적절한 정보를 의사결정자에게 전달하여 적확하고 신속한 결정을 내리도록 도와주어 결과적으로 가뭄으로 인한 피해를 최소화할 수 있다. 이를 위해서는 가뭄 정보를 생산할 수 있는 기상 및 수문 자료가 필요하다.

가뭄 정보를 도출하기 위해 강수, 온도, 증발산, 토양수분 등의 자료를 사용하는데 이러한 자료는 일반적으로 관측 지점에서만 얻을 수 있다. 따라서 관측 지점이 조밀하게 분포하지 않거나 현장 관측 자료가 부재하는 경우 공간적으로 분포된 자료를 얻기 위해서는 내삽 등의 통계적 방법을 활용하거나 지면모델과 같은 모델링을 이용하게 되는데, 최근 인공 위성 및 관측 센서 기술의 발달로 가뭄의 특성을 파악하기 위한 여러 변수를 원격 탐사를 통해 도출하고 있다. 원격 탐사는 시공간적으로 연속적인 정보를 비용 효과적으로 제공하므로, 최근 가뭄 연구에 널리 활용되는 추세이다.

본 연구의 목적은 기상 관측 자료가 조밀하지 못한 지역에 대하여 원격 탐사 자료를 활용하여 가뭄을 감시하고 예측하는 기술을 개발하는 것이다. 위성 영상에서 도출된 강수, 토양수분, 지면온도, 식생지수, 증발산 등의 자료를 기계학습을 통해 모델링하여 공

간적으로 연속적인 가뭄 감시 및 예측 정보를 도출하고자 하였다. 가뭄 감시 및 예측 모델을 구축하려면 참값으로 삼을 수 있는 가뭄 상태의 지표가 필요한데, Standardized Precipitation Index(SPI), Standardized Precipitation Evapotranspiration Index(SPEI)의 두 관측 기반 가뭄 지수와 Normalized Difference Water Index(NDWI), Normalized Difference Drought Index(NDDI), Normalized Multiband Drought Index(NMDI)의 세 원격 탐사 기반의 가뭄 지수를 사용하였다.

SPI와 SPEI를 이용하여 구축한 모델의 경우는 1, 3, 6, 9개월의 시간척도(time scale)를 사용하였으므로 1개월의 짧은 시간척도를 이용한 경우 기상학적 가뭄에 적용하고 3, 6개월의 시간척도를 이용한 경우 농업적 가뭄에, 9개월의 긴 시간척도를 이용한 경우 수문학적 가뭄에 적용하여 활용할 수 있다. NDWI와 NDDI의 경우는 식생의 가뭄 피해를 중심으로 살펴보기 위해 개발되었으며 NMDI는 토양과 식생의 수분 상태를 함께 보기 위해 개발되었으므로 이들을 이용하여 구축한 모델은 농업적 가뭄(토양 수분 가뭄)에 가장 많은 정보를 제공할 수 있다.

1.2 연구 범위

1.2.1 공간적 범위

연구 대상 지역은 자료의 검증이 용이하도록 우리나라 남북한 지역으로 설정하였다(Figure 1-3). 남북한을 함께 보았을 때, 극동 지점은 경상북도 울릉군 울릉읍 독도리 동도(131°52'32'E), 극서 지점은 평안북도 용천군 비단섬(124°10'47'E), 극남 지점은 제주특별자치도 서귀포시 대정읍 마라도(33°06'32'N), 극북 지점은 함경북도 온성군 풍서리(43°00'36'N)이며, 남북한 지역의 면적은 각각 100,188km², 123,138km²이다(국가통계포털, 2014). 대략적인 토지 이용은 남한 지역의 19.4%가 논밭, 80.6%가 임야 등으로 이루어져 있으며 북한 지역의 17.3%가 논밭, 82.7%가 임야 등으로 이루어져 있어(국가통계포털, 2014) 전체 면적의 80% 가량이 산지로 이루어졌음을 알 수 있다.

남한 지역 주요 도시의 평균 기온은 12.3~17.2°C, 북한 주요 도시의 평균 기온은 3.9~12.1°C 범위에 분포하며 평균 강수량은 남한 지역 주요 도시는 859.1~1,403.8mm, 북한 지역 주요 도시는 635.2~1,454.3mm이다(국가통계포털, 2014).

1.2.2 시간적 범위

연구의 시간적 범위는 사용된 여러 위성 영상 자료의 시간 범위를 고려하여 2001년부터 2012년의 12년 동안의 기간으로 설정하였다. 다만 사용된 Aqua 위성의 Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) 센서 자료의 가용 범위가 2002년 7월부터 2011년 9월까지 이므로, 분석에 위성 토양수분 자료를 포함시키는 경우는 이 기간 동안에 대해 분석하였다.

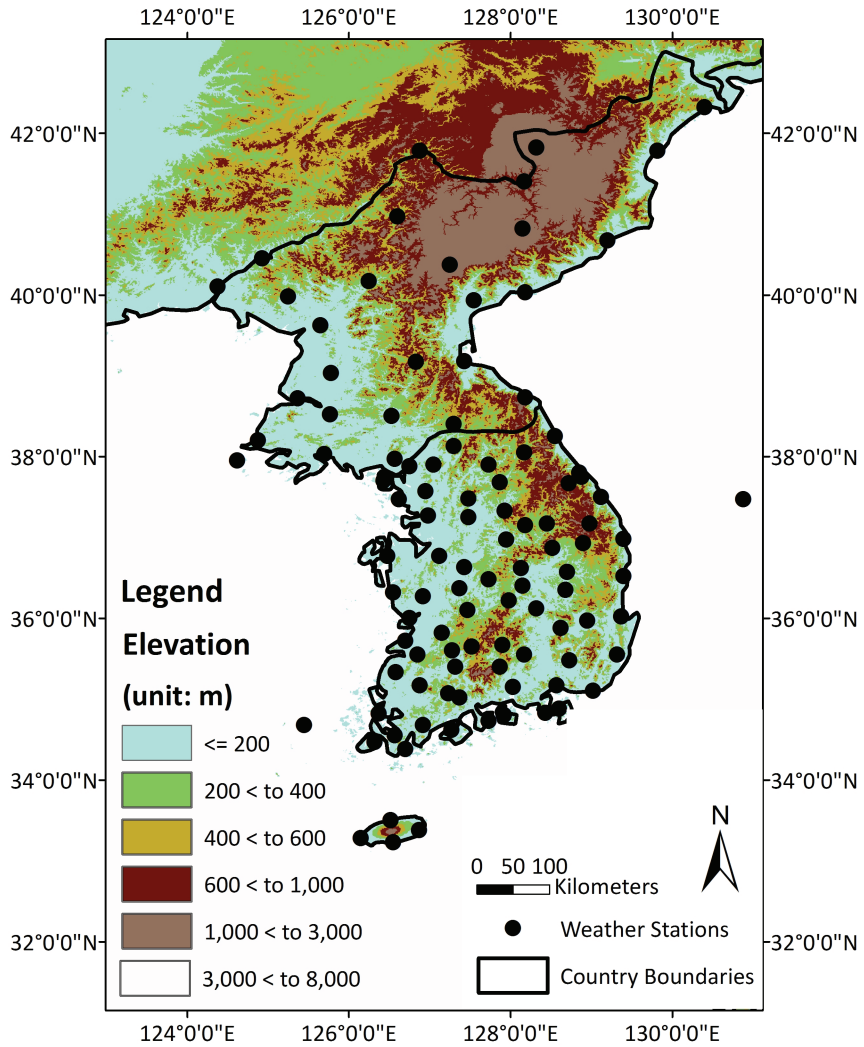


Figure 1-3. Study area – the Korean Peninsula.

1.2.3 내용적 범위

본 연구의 내용적 범위는 원격 탐사를 활용한 가뭄 감시 및 가뭄 예측이다. 가뭄 감시 및 가뭄 예측을 위한 여러 가지 연구 방법론이 존재하나 이 중 위성 영상 자료를 이

용한 기계학습 모델링 방법론만을 다루었다. 위성 영상 자료 또한 상업적으로 판매하는 것이 아닌 누구나 연구 및 활용 목적으로 별도의 비용 없이 획득하여 사용할 수 있는 자료만을 포함시켰다. 또한 위성 영상 자료의 해상도 또는 가뭄 예측성을 개선하기 위해 지면모델 결과와의 결합을 시도하였는데, 이는 토양수분 변수에 대해서만 적용하였다.

개발된 가뭄 감시 및 예측 모델을 고랭지 채소 수확량 및 산불 발생 빈도 예측에 적용해 보았는데, 이는 농업 생산성 예측이나 산불 발생 빈도 예측을 목표로 한 것이 아니며 고랭지 채소 수확량과 산불 발생에 미치는 가뭄의 영향을 보이기 위한 것이다. 배추나 무 등의 작물은 수분 함량이 높으므로 가뭄이 발생하면 생리현상의 장애를 수반하여 성장과 생육의 지연이 발생하고 품질이 저하되게 되는데(김기덕, 2014), 고랭지 지역은 기본적으로 관개 시설이 열악하여 강수량, 온도, 토양수분 등에 대한 가뭄의 영향이 수확량과 높은 상관관계를 가진다(한국농촌경제연구원, 1996). 산불의 발생에 영향을 미치는 주요 변수는 기온, 실효습도, 평균풍속, 강수량 등으로 산불 발생과 가뭄은 밀접한 관련이 있다(원명수, 2014).

2. 원격 탐사를 활용한 가뭄 감시

2.1 전통적 가뭄 감시

전통적인 가뭄 감시 방법론은 일반적으로 현장 관측 자료에 기반하고 있다. 앞 장에서 언급한 가뭄의 기술적인 정의인 가뭄의 시작(onset), 끝(termination), 심도(severity) 등과 가뭄 피해가 발생한 공간적 범위(extent), 동일 심도 가뭄의 재현 빈도(return period) 등의 특성을 파악하기 위해 강수, 온도 등의 현장 관측 자료를 그대로 이용하기도 하지만 대부분 가뭄 지수(drought index)를 계산하여 활용한다.

2.1.1 전통적 가뭄 지수

20세기 중반 이후 수많은 가뭄 지수들이 개발되었는데, 가장 많이 사용되는 가뭄 지수로는 미국에서 1960년대에 개발된 파머가뭄지수(PDSI; Palmer, 1965)와 1990년대에 개발된 표준강수지수(SPI; McKee et al., 1993) 등이 있다. 2000년대 들어 SPI가 사용하는 변수인 강수량 뿐 아니라 대기의 수분 수요를 고려하여 증발산량을 포함시킨 표준강수증발산지수(SPEI, Vicente-Serrano et al., 2010) 등이 개발되어 다양하게 활용되고 있다. 기상 관측 지점에 대해 시계열로 존재하는 기상 변수 자료를 이용하여 이들 가뭄 지수를 산정한 다음 가뭄 평가 및 감시를 수행한다.

2.1.1.1 파머가뭄지수

파머가뭄지수는 물수지 모델에 기반하고 있으며 Z-Index, PDSI, Palmer Hydrological Drought Index(PHDI)으로 구성되어 있다. PHDI를 현업화하여 운영할 수 있도록 수정한 Palmer Modified Drought Index(PMDI)도 제공한다. 이 중 Z-Index는 1개월의 시간 척도를 가지므로 단기간의 강수량 부족으로 인한 기상학적 가뭄 감시에 유용하다. 그밖에 PDSI는 대략 9개월에서 12개월의 시간 척도를 가지며 (Guttman, 1998), PHDI는 그 이상의 시간 척도를 가져 수문학적 가뭄 감시에 적합한 것으로 알려져 있다. 동일한 물수지 모델을 이용하여 1주일 시간 척도의 토양수분 수지를 산정하는 Crop Moisture Index(CMI)도 개발되었다(Palmer, 1968).

PDSI는 물수지 모델을 이용하므로 강수 이외에 온도, 증발산, 토양수분, 지표유출 등의 변수를 고려하게 된다. 따라서 가뭄 발생에 기여하는 다른 요인을 폭넓게 고려할 수 있다는 장점이 있다. 그러나 토양 수분 함량(Available Water Capacity) 등의 변수 값이 필요해서 토양 관련 데이터베이스가 구축되어 있지 않은 지역에 적용하기 어려운 단점이 있고, 또한 시공간적인 비교가 적합하지 않다는 지적이 있어 왔다(Alley, 1984). 시간 척도가 9개월에서 12개월로 고정되어 있기 때문에 다양한 형태의 가뭄에 융통성 있게 적용하기 어렵다는 단점도 있다.

2.1.1.2 표준강수지수

표준강수지수는 원하는 시간 척도에 따라 강수량을 누적시키고, 이를 정규화 하여 산정할 수 있다. 우선 강수량만을 사용하므로 많은 지역에 적용시킬 수 있고, 정규화 과정을 포함하므로 시공간적인 비교가 용이하다는 장점이 있다. 원하는 시간 척도에 대해 누적 강수량을 월별 등의 각 기간에 따라 정규화 하므로 특정 월 등의 강수량을 평년의 강수량 등과 비교할 수 있다. 개발된 이후 SPI는 많은 연구자들에 받아들여져 World Meteorological Organization(WMO)에 의해 기상학적 가뭄을 감시하기에 가장 적합한 지수로 선정되기도 하였다(Hayes et al., 2011). 사용자가 시간 척도를 지정할 수 있기 때문에 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, 우선 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 시간 척도에 따른 누적 강수량의 시계열을 구한 다음, 월별로 L-모멘트 값을 계산하고 이 값들을 이용하여 선택된 확률분포함수의 매개변수 값

을 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. SPI의 평균은 0, 표준편차는 1이 되는데, McKee et al.(1993)의 체계에 따라 가뭄의 심도를 분류할 수 있다(Table 2-1).

2.1.1.3 표준강수증발산지수

우리나라 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.

Table 2-1. Classification categories of SPI (unitless).

Classification	Index Value
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

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 2-2). 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 a , β , γ 는 각각 scale 변수, shape 변수, origin 변수이다.

2.2 가뭄 관련 변수의 원격 탐사

최근 수문기상 변수 값을 측정하거나 추정하게 하는 위성 관측 센서 기술의 발달로 원격 탐사를 통한 가뭄 감시 연구가 활발히 진행되고 있다. 가뭄 특성을 파악하기 위한 여러 변수의 값을 위성 영상으로부터 도출할 수 있는데, 위성 영상 자료는 시공간적으로 비교적 연속적으로 생산되므로 수문기상 변수의 관측 자료의 부재나 미흡함의 한계를 극복하고 자료가 존재하는 지역 뿐 아니라 미계측 지역에 대해서도 가뭄 정보를 시기적절하게 도출할 수 있게 되었다. 또한 수문기상모델이나 지면모델을 함께 활용하여 위성 영상 자료를 자료 동화 등의 기법을 통해 입력 자료로 활용하는 등, 정확한 수문기상 초기 조건을 제공하여 가뭄 예측의 정확도 향상에 기여하고 있다. 정지궤도 위성이 아니라 극궤도 위성을 이용하는 경우 연속적인 자료 수집이 이루어지는 것이 아니라 특정 시각의 자료만을 수집하게 되는 단점이 있으나, 이런 경우 해당 변수의 일변화 함수 등을 이용하여 최대, 최소, 또는 일평균 등의 원하는 값을 도출하는 연구가 있어왔다(APECC기후센터, 2013a).

기상학적 가뭄, 토양수분 가뭄(농업적 가뭄), 수문학적 가뭄을 파악하기 위한 변수들은 물수지 및 에너지 수지와 관련된 수문기상 주요 변수들로, 강수, 증발산, 토양수분, 유출, 저수위, 유량, 기온 등이 포함된다. 가뭄의 발달 정도와 관심 부문에 따라 다양한 형태의 가뭄이 존재하므로, 적절한 가뭄 감시를 위해서는 평가하고자 하는 가뭄을 정의하고 이를 정량화하여 가뭄 특성을 파악할 수 있게 하는 변수 또는 가뭄 지수를 도출하는 것이 필요하다. 강수량은 특히 기상학적 가뭄 감시에 적합한 변수이며, 수문학적 가뭄을 평가하고 감시하기 위해서는 유량이나 저수위 자료가 사용되어야 할 것이다. 토양수분 가뭄(농업적 가뭄)은 토양수분 상태나 농작물 생산량 등을 활용하여 평가할 수 있다(Mishra and Singh, 2011). 수문기상 변수들 중 강수나 증발산 등은 가뭄 상태를 걸

정짓는 변수(drivers)라고 생각할 수 있고 유량이나 저수위 등은 가뭄의 영향을 나타내는 변수(response variables)라고 볼 수 있다. 효율적인 가뭄 감시를 위해서는 이 두 형태의 변수들이 모두 필요하다.

2.2.1 강수량

위성 기반 기상 변수 값을 도출하려는 노력은 1960년대에 첫 위성인 Television Infrared Observation Satellite Program(TIROS)-1 이후로 계속 되어 왔다. 정지궤도(geostationary) 위성들은 가시광선(Visible; VIS) 영역과 적외선(Infrared; IR) 영역을 스캔하여 구름 상층의 온도와 알베도 자료를 알고리즘에 적용하여 강수 정보를 도출한다. 극궤도(polar-orbiting) 위성들은 수동마이크로파(passive microwave; PWM)를 통해 강수를 예측하는데, 같은 위치에 대해 하루에 1~2회 측정한다. 널리 이용되는 강수 알고리즘으로는 Tropical Rainfall Measuring Mission(TRMM) 3B42, Climate Prediction Center MORPHing(CMORPH), Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks(PERSIANN) 등이 있으며, 앞으로는 2014년 2월에 쏘아올린 Global Precipitation Measurement(GPM) Core Observatory Satellite을 중심으로 한 GPM 자료들이 더욱 널리 사용될 전망이다.

2.2.2 증발산량

증발산량은 물수지 또는 에너지 수지를 통해 추정할 수 있는데, 위성 영상 자료를 활용할 수 있는 여러 에너지 수지 기반 증발산량 모델이 존재한다. 널리 사용되는 모델로 SEBAL(Surface Energy Balance Algorithm for Land), METRIC(Mapping EvapoTRanspiration with Internalized Calibration), SSEB(Simplified Surface Energy Balance) 등이 있으며, 입력 변수로는 강수, 지표면 온도, 토양 자료 등 다양한 변수를 사용한다.

2.2.3 토양수분

토양수분은 수동형 마이크로파 복사계(radiometer)를 이용하여 지표로부터 에너지를

측정, 알고리즘에 적용하여 추정하거나, 능동형 마이크로파 Synthetic Aperture Radar(SAR) 또는 산란계(scatterometer)를 이용하여 도출할 수 있다. 복사계로는 Scanning Multichannel Microwave Radiometer(SMMR), Special Sensor Microwave/Imager(SSM/I), TRMM Microwave Imager(TMI), AMSR-E, Soil Moisture and Ocean Salinity sensor(SMOS) 등이 존재하며, SAR로는 Seasat, Spaceborne Imaging Radar-C(SIR-C), European Remote Sensing(ERS), RADARSAT, Environmental Satellite(Envisat), Advanced Land Observing Satellite(ALOS) 등이 있으며, 산란계로는 ERS, QuickSCAT(QSCAT) 등이 있다. 토양수분의 경우 위성 자료의 공간 해상도가 낮고 또한 시공간적인 변이가 매우 크기 때문에 관측 자료와의 검증이 용이하지 않은 어려움이 있다.

2.2.4 식생

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 정규식생지수(Normalized Difference Vegetation Index) 등이 개발되어 널리 사용되어 왔다(Tucker, 1979; Eq. 2-8).

$$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}} \quad (\text{Eq. 2-8})$$

이때, ρ 는 표면반사율이다. 그런데 각 지역별로 NDVI의 최저점과 포화점이 다르게 나타나므로, 환경적인 요인과 기상학적 요인을 구분하기 위하여 Kogan(1995)은 장기간의 시계열 자료의 최솟값과 최댓값을 이용하여 NDVI를 표준화시킨 Vegetation Condition Index(VCI)를 개발하였으며(Eq. 2-9), 역시 표준화된 지표온도인 Temperature Condition Index(TCI; Eq. 2-10)를 선형으로 조합하여 Vegetation Health Index(VHI)를 개발

Medium Resolution Imaging Spectrometer(MERIS), Moderate Resolution Imaging Spectroradiometer(MODIS) 등의 센서가 쓰이고 있다.

$$VCI = \frac{NDVI - NDVI_{MIN}}{NDVI_{MAX} + NDVI_{MIN}} \quad (\text{Eq. 2-9})$$

$$TCI = \frac{LST_{MAX} - LST}{LST_{MAX} + LST_{MIN}} \quad (\text{Eq. 2-10})$$

$$VHI = a \cdot VCI + (1 - a) \cdot TCI \quad (\text{Eq. 2-11})$$

이때, LST는 지면온도이며, a는 선형 조합을 위한 매개변수로 0.7의 값이 제안되었다.

2.3 원격 탐사 기반 가뭄 지수

관측 자료만을 이용하여 산정하는 가뭄 지수는 오로지 관측 자료가 존재하는 지점에 대해서만 가뭄 정보를 얻을 수 있는 단점이 있으므로, 시공간적으로 연속적인 자료를 생산하는 원격 탐사를 활용하여 가뭄 지수를 구하려는 노력이 계속되고 있다. 앞 절에서 언급한 NDVI, VCI, VHI는 식생 지수와 지면 온도를 이용한 가뭄 지수로, 농업적 가뭄 감시 및 평가에 유용하게 사용될 수 있으며, Anderson et al.(2010)이 실제 증발산량(Evapotranspiration; ET)와 잠재 증발산량(Potential Evapotranspiration; PET)의 비와 아노말리를 이용하여 개발한 Evaporative Stress Index(ESI; Eq. 2-12)는 과거 기상이나 토양에 대한 정보 없이 가뭄 정보를 얻을 수 있는 가뭄 지수이다.

$$ESI = 1 - f_{PET} = 1 - \frac{ET}{PET} \quad (\text{Eq. 2-12})$$

Rhee et al.(2015)가 정리한 최근에 개발된 원격 탐사 기반 가뭄 지수를 보면, Gao(1996)는 MODIS의 단파 적외선(Shortwave Infrared; SWIR)을 이용하여 NDWI

를 개발하였다(Eq. 2-13; ρ 는 표면반사율). NDWI는 $0.86\mu\text{m}$ 근적외선(NIR) 영역과 $1.24\mu\text{m}$ 단파 적외선(SWIR) 영역을 이용하여 산정할 수 있는데, 근적외선 영역에서는 식생에 존재하는 수분의 흡수율이 매우 작으나, 단파 적외선 영역에서는 어느 정도 나타나게 된다. 따라서 NDWI는 식생 임관의 수분 보유에 민감하게 반응한다. 또한 NDVI와 비교하여 NDWI는 대기 에어로졸 산란 효과에 덜 민감하게 반응한다(Gao, 1996).

$$NDWI = \frac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}} \quad (\text{Eq. 2-13})$$

Gu et al.(2007)은 NDVI와 NDWI를 이용하여 NDDI를 개발하였다(Eq. 2-14). RED, NIR, SWIR 영역의 표면반사율을 모두 결합하여 가뭄 심도에 대한 민감도를 개선하였으며, 미국의 캔자스와 오클라호마 지역의 초지에 대해 테스트한 결과 NDDI는 NDVI, NDWI와 비교하여 여름 가뭄 상태를 더 잘 평가하는 것으로 판단되었다(Gu et al., 2007).

$$NDDI = \frac{NDVI - NDWI}{NDVI + NDWI} \quad (\text{Eq. 2-14})$$

Wang and Gu(2007) 역시 MODIS의 NIR과 SWIR을 이용하여 NMDI를 개발하였다(Eq. 2-15). NMDI는 NIR 이외에 토양 및 식생에 민감한 두 $1.640\mu\text{m}$ 과 $2,130\mu\text{m}$ 두 SWIR 채널의 차를 이용함으로써 토양과 식생의 수분을 추정하여 가뭄 심도에 대한 민감도를 개선하였다.

$$NMDI = \frac{\rho_{NIR} - (\rho_{SWIR1} - \rho_{SWIR2})}{\rho_{NIR} + (\rho_{SWIR1} - \rho_{SWIR2})} \quad (\text{Eq. 2-15})$$

이때, ρ 는 표면반사율, SWIR1은 $1.640\mu\text{m}$, SWIR2는 $2,130\mu\text{m}$ 영역을 나타낸다.

2.4 변수 융합을 통한 맞춤형 가뭄 감시

가뭄이 발달함에 따라 가뭄 피해 지역도 확장되고 가뭄과 관련된 이해당사자도 변화하게 되며, 가뭄을 감시하는데 중요한 변수가 달라지기도 하는데 수문기상 변수들은 사실 육지와 대기간의 상호작용으로 인해 긴밀하게 연결되어 있다. 이들 변수들 간에는 커플링(coupling)과 피드백(feedback)이 존재하는데 이는 지역에 따라 다르게 나타날 수 있다. 예를 들어, 토양수분과 증발산 사이의 커플링은 건조한 지역에서는 토양수분에 의해 결정되지만 습윤한 지역에서는 다르게 작용한다(Seneviratne et al. 2010). 토양수분과 온도, 또는 토양수분과 강수 사이의 커플링과 피드백은 조금 더 복잡하게 나타난다(Seneviratne et al. 2010; Figure 2-1).

원격 탐사 기술은 가뭄과 관련된 변수 값을 추정하기 위해 사용될 수 있는데(Jensen, 2000), 종종 위성 센서는 얻고자 하는 변수 값을 직접 측정하기 보다는 다른 여러 변수를 이용하여 알고리즘을 통해 그 값을 추정하기도 한다. 이런 경우 각 변수의 추정 값이 실제 값과 차이가 있을 수 있다. 그러나 변수들 간에 위에 언급한 바와 같이 커플링과 피드백이 존재하므로 여러 변수를 조합하여 가뭄 상태 평가를 개선할 수 있다. 단 하나의 변수만을 고려하는 것보다는 여러 변수를 함께 고려한 가뭄 지수가 가뭄 감시에 더 유용할 수 있으며, 많은 변수는 원격 탐사를 통해 추정할 수 있다(Wardlow et al., 2012).

Rhee et al. (2010)은 TRMM으로부터의 강수량, MODIS 의 지표면 온도와 식생 지수 등 여러 변수를 조합하여 건조한 지역 뿐 아니라 습윤한 지역에서도 가뭄의 상태를 효과적으로 감시하는 농업적 가뭄 지수인 Scaled Drought Condition Index(SDCI)를 개발하였으며 이 지수가 농작물 수확량과 높은 상관관계가 있음을 보였다. APEC기후센터(2013a)는 위성 영상으로부터 강수량과 증발산량을 추정하여 각각의 백분율의 차로 이루어진 지수가 유량과 높은 상관관계를 보이며, 따라서 수문학적 가뭄 평가 및 감시에 유용하게 쓰일 수 있음을 보였다(Figure 2-2).

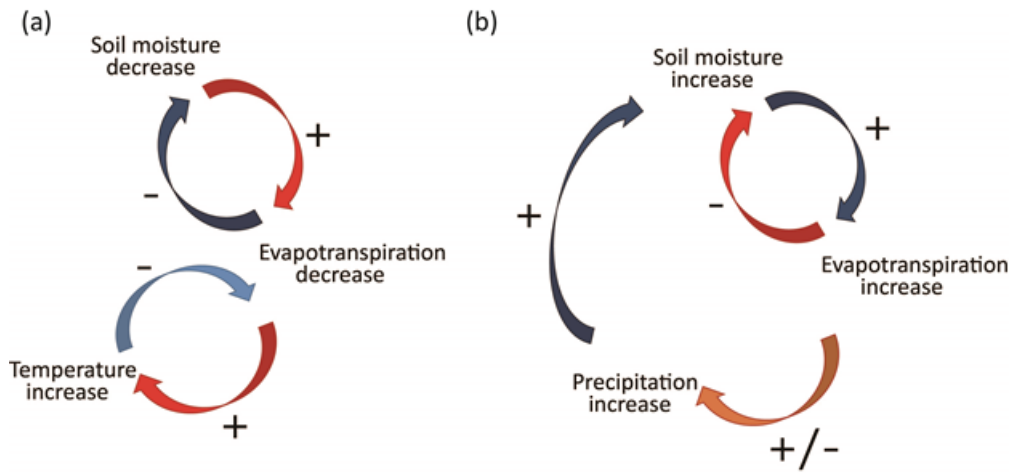


Figure 2-1. Processes contributing to (a) soil moisture-temperature, and (b) soil moisture-precipitation coupling and feedback loops (adapted from Seneviratne et al., 2010, recited in Rhee et al., 2015).

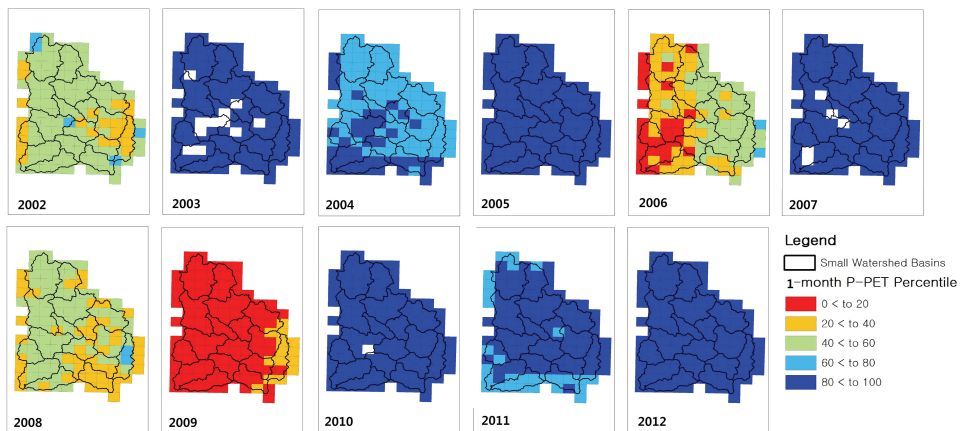


Figure 2-2. September P-PET percentiles in Upper Namhan-gang basin (APEC Climate Center, 2013; recited in Rhee et al., 2015).

Mu et al.(2013)은 MODIS의 실제 증발산량, 잠재 증발산량, NDVI을 이용해서 전 지구적으로 가뭄을 감시하기 위한 Drought Severity Index(DSI)를 제안하였으며, Zhang and Jia(2013)는 건조한 지역의 기상학적 가뭄과 같은 단기 가뭄을 감시하기 위해 TRMM 강수량 자료와 AMSR-E의 토양 수분 자료, 지면 온도 자료를 결합한 Microwave Integrated Drought Index(MIDI)를 개발하였다. 이 밖에도 위성 영상 자료를 이용한 많은 가뭄 지수가 개발되고 있다.

효율적인 가뭄 평가 및 감시를 위해서는 위성 영상 자료 뿐 아니라 가용한 자료를 되도록 다양하게 활용하는 것이 바람직하다. 위성 영상 자료, 관측 자료, 모델 결과 등 가용한 자료를 적극적으로 결합하여 가뭄 감시를 수행하는 사례들이 존재하는데, 미국의 United States Drought Monitor(USDM; <http://drought.unl.edu/dm>), 미국 기후예측센터(Climatic Prediction Center; CPC)의 Objective Blend Drought Index(OBDI), 미국 지질조사국(United States Geological Survey; USGS)의 Vegetation Drought Response Index(VegDRI; Brown et al., 2008) 등이 있다.

USDM은 미국 전역에 대해서 매주 가뭄의 심도와 공간적인 범위를 지도로 제공해주는 가뭄 감시 시스템이다. 가뭄 정보는 6개의 주요 물리적인 가뭄 지표로부터 도출하는데, 관측 자료 기반의 PDSI, 강수량의 평년 대비 비율(Percent of Normal Precipitation), SPI, USGS의 일간 유량 백분율, CPC의 모델 결과인 토양수분 백분율, 위성 영상에서 도출한 VHI가 이에 속한다. 보조 자료로는 CMI, 산불 지수인 Keetch-Byram Drought Index(KBDI; Keetch and Byram, 1968), 저수지 수위, 호수 수위, 지하수 수위, 관측한 토양수분 자료 등이 쓰이며, 미국 서부에 대해서는 산악 지역 적설의 물 등가량을 측정하기 위한 자연자원보전국(Natural Resources Conservation Service; NRCS)의 Snowpack Telemetry (SNOTEL) 관측 자료가 추가적으로 사용된다(Svoboda et al., 2002).

USDM은 위에서 언급한 객관적인 가뭄 지표 이외에 미국의 국가가뭄경감센터(National Drought Mitigation Center; NDMC), 농업부(United States Department

of Agriculture, USDA), CPC, 국가기후자료센터(National Climatic Data Center; NCDC)의 가뭄 전문가들의 주관적인 의견을 통합하여 완성된다. 사용되는 가뭄 지표들과 이들의 조합 시 활용하는 가중치는 이 전문가들에 의해서 결정되는데, 가중치가 미리 결정되어 있는 OBDI 도 완전히 객관적인 것은 아니다(Svoboda et al., 2002). 그러나 OBDI의 경우 가중치가 정해져 있기 때문에 추후에 다시 가뭄 지수를 생산하려 할 때 같은 값을 생산할 수 있다. 지표의 종류와 가중치는 평가하고자 하는 가뭄의 시간 척도에 따라 결정한다.

VegDRI는 가뭄으로 인한 식생의 스트레스를 측정하는데, 관측 자료 기반의 SPI, self-calibrating PDSI(Wells et al., 2004), USDM와 원격 탐사 기반의 식생 지수들인 NDVI, Percent of Average Seasonal Greenness (PASG), Start of Season Anomaly (SOSA)를 이용한다. 또한 토지 피복/토지 이용, 토양의 가용한 수분 함량(Available Water Capacity; AWC), 관개 면적 등의 생물리학적 정보도 활용한다. VegDRI는 모델링을 위해 Supervised Classification and Regression-Tree (CART) 분석 방법을 사용하고 있다.

3. 원격 탐사를 활용한 가뭄 예측

3.1 가뭄 예측 모델링

가뭄 감시를 통해 실시간(real-time) 또는 준실시간(near real-time)으로 생산한 가뭄 정보는 시기적절하고 즉각적인 가뭄 대처에 반드시 필요하다. 이에 앞서 가뭄 발생에 대한 정보가 사전에 주어진다면 가뭄 피해를 경감하기 위한 자원을 확보하고 효율적 배분을 위한 계획을 세워 가뭄 피해를 최소화할 수 있을 것이다. 가뭄 예측 모델링을 통해 가뭄이 발생하기 전에 가뭄의 발생과 종료 시점 등 여러 가뭄 특성을 예측하여 가뭄 관련 의사결정에 다양하게 활용하기 위한 기술이 개발되고 있다.

Mishra and Singh(2011)는 가뭄 예측 모델링의 구성 요소를 수문기상변수, 가뭄 지수, 기후 지수, 방법론, 결과물로 나누었는데(Figure 3-1), 이 중 주요 수문기상변수

는 예측하고자 하는 가뭄의 유형에 따라 결정된다. 서론에서 설명한 바와 같이 기상학적 가뭄의 감시에는 강수량이 중요한 변수이며, 토양수분 가뭄(농업적 가뭄) 감시를 위해서는 토양수분, 농작물 수확량 등의 변수가, 수문학적 가뭄의 감시를 위해서는 유량, 저수위 등의 변수가 중요하다. 가뭄 심도, 지속기간, 가뭄 영향의 공간적 범위 등의 가뭄 특성을 도출하기 위해서는 여러 변수를 결합하여 사용하거나 가뭄 지수를 활용하기도 하고 시공간적인 분석을 함께 수행한다. El Nino-Southern Oscillation(ENSO), 북극진동(Arctic Oscillation; AO) 등의 대규모 기후 지수(large-scale climate index)는 계절적 시간 척도 이상의 장기적인 가뭄 예측을 위해 활용될 수 있다. 가뭄 예측 모델링 방법론으로 회귀 모델, 시계열 모델, 확률 모델, 신경망 모델, 하이브리드 모델 등이 있는데(Figure 3-1), 각각의 방법론은 가뭄을 정량화하기 위한 여러 매개변수를 도출한다.

3.1.1 회귀 모델

회귀 모델링은 예측하려고 하는 가뭄 상태를 정량화하는 일종의 목표 변수를 설정하는데, 일반적으로 관측 값 기반의 SPI 등의 가뭄 지수, 가뭄의 영향을 잘 나타내 주는 농작물 수확량(crop yield), 위성 영상에서 도출한 NDVI 등과 같은 식생 지수 등을 사용한다. 이러한 목표 변수를 회귀 방정식의 종속 변수로 설정하며, 가뭄의 발생 및 진행에 영향을 주는 강수량, 증발산량, 토양수분 등과 같은 여러 수문기상 변수를 독립 변수로 설정하여 모델링을 수행한다. 한 예로, 가뭄 상태에서의 밀 수확량과 이를 결정짓는 여러 구성 요소들과의 관계를 규명하기 위해 Leilah and Al-Khateeb(2005)는 단순 상관(simple correlation), 경로 분석(path analysis), 다중 선형 회귀(multiple linear regression), 단계적 회귀(stepwise regression), 요인 분석(factor analysis), 주성분 분석(principal components analysis), 군집 분석(cluster analysis) 등의 여러 통계적 분석 방법론을 이용하여 비교하였는데, 이 중 다중 선형 회귀 분석을 통해 얻은 모델은 종속 변수 분산의 98%를 설명하였다.

3.1.2 시계열 모델

시계열 모델로는 Autoregressive Integrated Moving Average (ARIMA) 모델과 계절 내 또는 계절 간의 비정상성(nonstationarity)을 고려한 Seasonal Autoregressive Integrated Moving Average (SARIMA) 모델을 많이 사용한다(Mishra and Singh, 2011). 가뭄 예측을 위해 적용하는 경우 가뭄을 정량화하는 SPI 등과 같은 변수의 시계열을 이용하여 지난 기간의 관측 값을 사용하여 어느 정도의 선행 예측 시간의 값을 예측하게 된다. Han et al.(2010)의 연구를 예로 살펴보면, 위성 영상 자료로부터 Vegetation Temperature Condition Index (VTCI) 지수 값의 시계열을 36개 픽셀에 대해 도출한 다음 ARIMA 모델을 구축하여 가뭄 예측 모델링을 수행하였다.

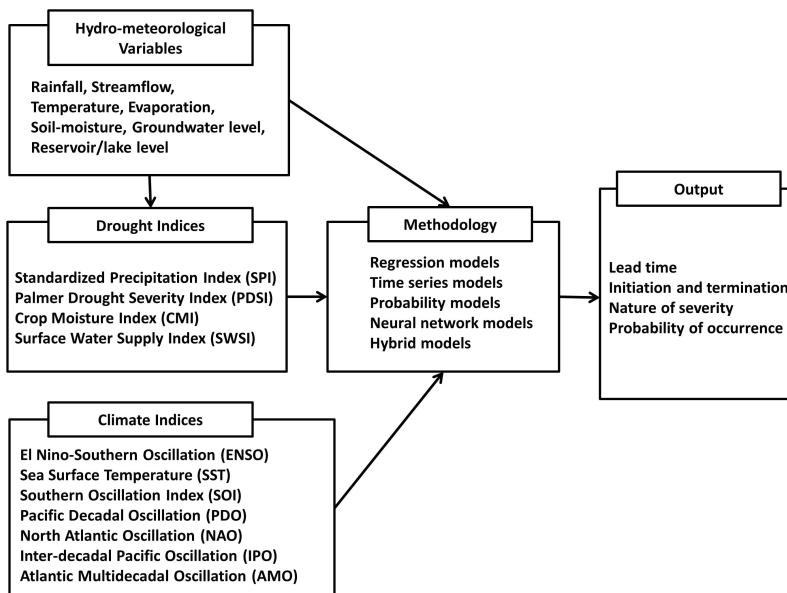


Figure 3-1. Components for drought forecasting (adapted from Mishra and Singh, 2011).

3.1.3 인공 신경망 모델

인공 신경망(Artificial Neural Networks) 모델은 인간의 뇌가 들어오는 정보를 처리하기 위해 복잡하게 연결되어 있는 뉴런들을 사용하는 사고 과정을 모의하여 개발된 모델로(Jensen, 2005; Figure 3-2), 생물학적 신경망의 결함 허용성(fault-tolerance) 및 학습 역량을 모방한 것이다. 각 뉴런은 전기화학적인 신호를 전달하기 위해 특화된 세포로, 정보가 입력되는 수상돌기(dendrite), 신경세포체, 정보가 출력되는 축삭돌기(axon)으로 구성되어 있으며 축삭돌기는 시냅스를 통해 다른 뉴런의 수상돌기와 연결된다. 인공 신경망 모델은 이러한 구조를 모방하여 각각의 뉴런이 활성화 여부를 결정하는 역치(threshold)를 가지며 받아들인 입력 정보가 각각의 가중치(weights)를 거쳐 결과물 생산으로 연결된다.

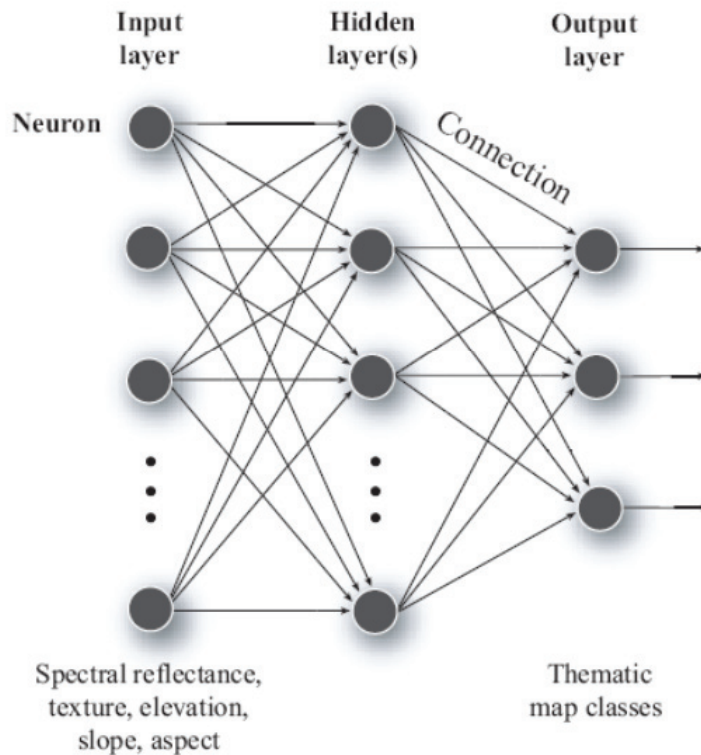


Figure 3-2. Components of a typical Artificial Neural Network model (Jensen, 2005)

Morid et al.(2007)은 인공 신경망을 이용한 가뭄 예측을 수행하였는데, 가뭄을 정량화하는 목표 변수로 유효 가뭄 지수(Effective Drought Index; EDI)와 SPI를 설정하고 독립 변수로는 과거 여러 기간의 강수량, Southern Oscillation Index(SOI)나 North Atlantic Oscillation(NAO)와 같은 대규모 기후 지수, 그리고 과거 기간의 EDI와 SPI를 사용하였다. 여러 기상관측소 위치에 대하여 다양한 구조의 인공 신경망 모델을 시험하였는데, 비교적 간단한 구조의 모델이 가뭄 예측을 우수하게 수행하였으며, 대규모 기후 지수는 그다지 큰 역할을 보이지 않았다.

3.1.4 의사결정분지도 모델

의사결정분지도(decision tree) 방법론은 지식기반(knowledge-based) 전문가 시스템(expert system)의 일종으로, 가설, 규칙, 조건을 이용하여 일련의 속성을 가진 변수 값을 입력으로 받아들여 결과를 생산해 낸다(Jensen, 2005; Figure 3-3). 가설을 검증하기 위해서 규칙과 조건을 사용하는데, 전문가 시스템에 대해서는 제4장에서 기계학습(machine learning)과 함께 자세히 언급하였다. 의사결정분지도의 목적은 분류(classification) 및 회귀(regression)인데, 이 모델을 통해 얻고자 하는 결과가 불연속인 경우 모델의 목적은 분류가 되며, 연속적인 결과를 얻고자 하는 경우 회귀가 된다.

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 시간차(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.

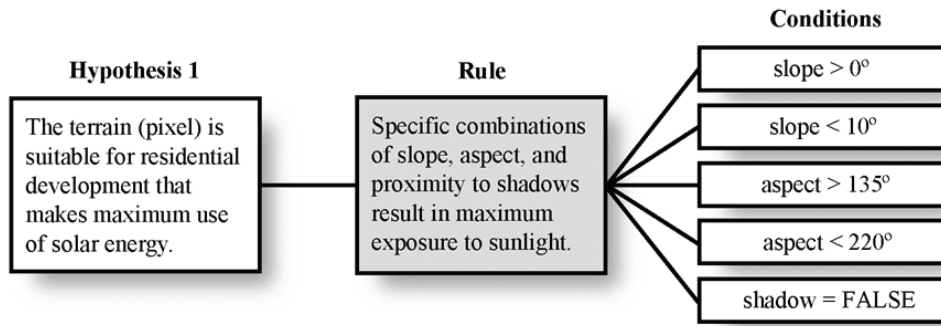


Figure 3-3. A decision-tree expert system with a rule and conditions to test a hypothesis (Jensen, 2005).

3.1.5 장기 예측 모델

기후 장기 예측 모델은 강수량, 대기 온도, 상대 습도 등의 다양한 변수에 대한 3~6 개월 선행 예측 시간의 기후 예측 값을 생산하므로 이 결과를 가뭄 예측에 활용할 수 있다. 국내에는 현업화되어 생산되고 있는 가뭄 예측 자료가 없으나, 미국 National Oceanic and Atmospheric Administration(NOAA)의 기후예측센터에서 제공하고 있는 계절 가뭄 전망(seasonal drought outlook)을 예를 들어 살펴볼 수 있다. 계절 가뭄 전망 결과를 생산하기 위해 사용되는 입력 자료는 기후예측센터의 1개월 온도 및 강수 전망, 3개월 장기 예측, 기상예측센터(Weather Prediction Center)와 기후예측센터의 단중기 예측, 토양수분 모델 결과물과 역학적 모델 결과, 4개월 파머가뭄지수의 종료 및 경감 확률, 기후값(climatology), 초기 조건(initial conditions) 등이 있다(CPC, 2014).

기후 장기 예측 모델 결과를 사용하여 생산한 가뭄 예측 자료는 최종 수요자가 쉽게 이해하고 사용하기 위해서 편리한 형태로 재가공 되기도 한다. Steinemann(2006)은 기후 과학과 사회적인 결정 간의 격차를 줄이기 위해 수자원 관리자들을 대상으로 설문 조사를 실시하였고, Forecast Precipitation Index(FPI)를 개발하여 가뭄 예측 결과가 실제 의사결정에 쉽게 사용될 수 있도록 하였다.

3.2 가뭄 예측을 위한 원격 탐사의 활용

앞 절에서 설명한 각각의 방법론에서 원격 탐사를 통해 획득한 위성 영상 자료는 자료 획득 시점의 대상의 상태와 조건을 보다 정확하게 측정하여, 관측 자료가 부재한 지역 등을 위한 입력 자료를 제공해 준다. 이를 통해 가뭄 정량화를 위한 목표 변수 추정이 개선될 수 있다.

가뭄 예측을 위해 역학적 모델을 사용하는 경우 개선된 수문기상 초기 조건을 제공하여 모델의 수행을 향상시킬 수 있다. Shukla and Lettenmaier(2011)는 수문학적 가뭄 감시를 위해 활용하는 목표 변수인 유량 및 토양수분을 생산하기 위해 계절 수문 예측 모델을 활용하였는데 이를 위한 초기 수문 조건과 기후 예측 자료의 상대적인 중요도를 비교·평가하였다. 대상지역에 대해 1개월 선행예측시간의 예측을 위해서는 초기 수문 조건이 특히 토양수분 예측에 대해 중요하다는 결론을 보였다.

가뭄 예측 모델링의 입력 자료로 쓰이는 수문기상 변수는 위성 영상 기반의 자료와 현장 관측 자료, 수문 기상 모델 결과물 등의 결합으로 도출될 수 있는데 이 중 위성 영상 기반 변수 추정의 매우 중요하다.

4. 자료 및 연구 방법론

본 연구에서는 위성 영상 자료를 이용하여 가뭄 감시 및 예측 모델을 구축하였다. 위성 영상에서 도출한 각각의 변수 값과, 관측 값이나 위성 영상 자료에서 도출한 자료를 이용하여 산정한 가뭄 지수를 이용하였으며, 위성 영상에서 도출한 변수 값들의 오차 범위 등을 살펴보기 위해 관측 자료와 지면모델 결과를 이용하였다. 또한 개발된 감시 및 예측 모델의 성능을 고랭지 채소 수확량 및 산불 발생 빈도를 통해 검증하고자 하였다. 관측 값과 추정 값은 피어슨 상관계수(Pearson's Correlation Coefficient) r 값과 이에 상응하는 비모수 통계 값인 켄달의 순위 상관계수(Kendall's rank correlation coefficient) τ 값, 오차를 파악할 수 있는 Root Mean Square Error(RMSE) 값, Mean Absolute Error(MAE) 값을 구하여 비교하였다.

$$r = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (\text{Eq. 4-1})$$

$$\tau = \frac{(\text{number of concordant pairs}) - (\text{number of discordant pairs})}{\frac{1}{2}n(n-1)} \quad (\text{Eq. 4-2})$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{Y}_i - Y_i)^2} \quad (\text{Eq. 4-3})$$

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

이 때, n 은 표본의 개수이고, X 또는 \hat{Y} 은 추정된 값, Y 는 관측 값을 의미한다.

4.1 위성 영상 자료

4.1.1 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) 자료를 사용하였다. 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(Figure 4-1).

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 0.25×0.25°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.

3B43 자료는 mm/hr 단위로 제공되므로 mm/month 단위로 전환되었으며 여러 시간 척도에 따라 누적된 강수량을 이용할 수 있도록 1, 3, 6, 9개월 동안 누적된 강수량 값을 구하였다. 또한 가뭄 감시 및 예측 모델 개발에 활용할 수 있도록 기상관측소 위치에 따른 시계열 값을 추출하였으며, 행정구역별 모델 검증에 활용하기 위해 남한 전역에 대하여 5×5km 격자에 대해 리샘플(resample) 하였다.

2014년 7월 8일 이후로 TRMM의 위치를 고정하기 위한 연료 부족으로 고도가 하강하여 10월 7일 이후로는 PR 자료의 제공이 중지되었다. 따라서 앞으로 TRMM 자료

의 생산은 불가할 것으로 보인다. 그러나 2014년 2월에 NASA와 JAXA에 의해 차세대 강수 및 강설 관측 값을 전지구적으로 생산하기 위한 Global Precipitation Measurement(GPM; Figure 4-2)의 핵심 위성인 Core Observatory Satellite이 발사됨에 따라 앞으로는 원격 탐사를 통한 강수량 관측에 있어 훨씬 개선된 자료를 얻을 수 있을 것이다. TRMM은 열대 지방과 그 주변에 중·고 강도의 강우에 집중한 반면 GPM은 고위도 지방으로까지 대상 지역을 확대하고 있다.

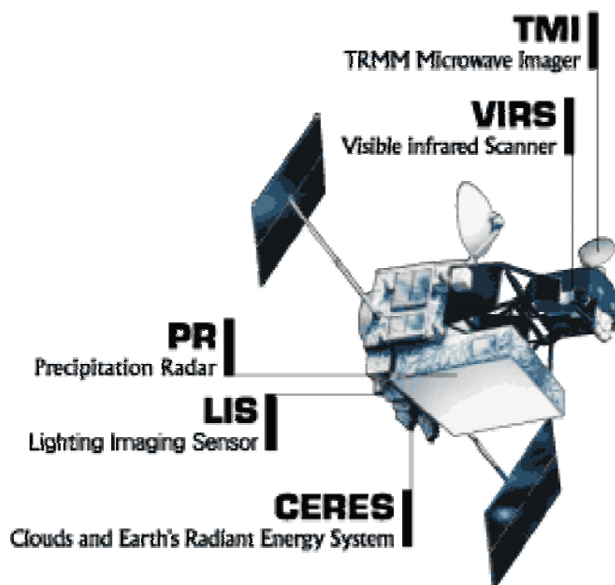


Figure 4-1. Diagram of the TRMM satellite and instruments (Source: pmm.nasa.gov).

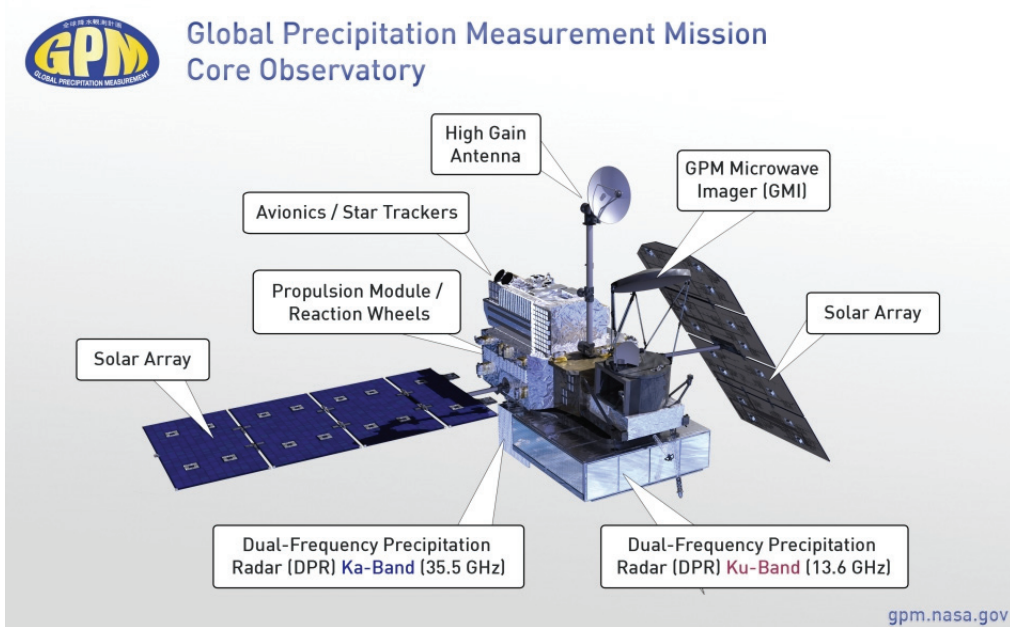


Figure 4-2. Global Precipitation Measurement Mission Core Observatory Satellite (Source: gpm.nasa.gov).

4.1.2 지표면온도

지표면온도는 미국 NASA의 Terra 위성에 탑재된 MODIS 센서로부터 생산된 Level-3 표준 자료인 MOD11A2 Land Surface Temperature and Emissivity 8-day L3 Global 1km 자료에서 도출하였다. 본 자료의 시간 해상도는 8일, 공간 해상도는 1×1km이며 일 단위 자료인 MOD11A1에서 구름 없는 날의 자료 값을 평균하여 생산한다. 주간 지표면온도와 야간 지표면 온도를 모두 제공하며, Sinusoidal 도법으로 투영되어 있다. 자료는 미국 NASA의 Earth Observing System Data and Information System, <http://reverb.echo.nasa.gov> (EOSDIS)에서 획득하였다.

Terra 위성은 1999년 12월 18일에 발사되어 2000년 2월 24일부터 자료를 수집하

기 시작하였으므로 지표면온도는 2000년 3월 5일부터 제공된다. 강교점(descending node)에서의 적도 통과 시간은 오전 10시 30분, 승교점(ascending node)에서의 적도 통과 시간은 오후 10시 30분이며, 1-2일에 한 번씩 지구 전체를 지나게 된다(Figures 4-3, 4-4). 본 자료는 8일 시간 해상도를 가지므로 월별 분석에 사용하기 위해 각 월에 해당하는 날 수를 가중치로 사용하여 월별 자료로 전환하였다. 또한 강수량 자료와 마찬가지로 기름 감시 및 예측 모델 개발에 활용할 수 있도록 기상관측소 위치에 따른 시계열 값을 추출하였으며, 행정구역별 모델 검증에 활용하기 위해 남한 전역에 대하여 $5 \times 5\text{km}$ 격자에 대해 리샘플(resample) 하였다.

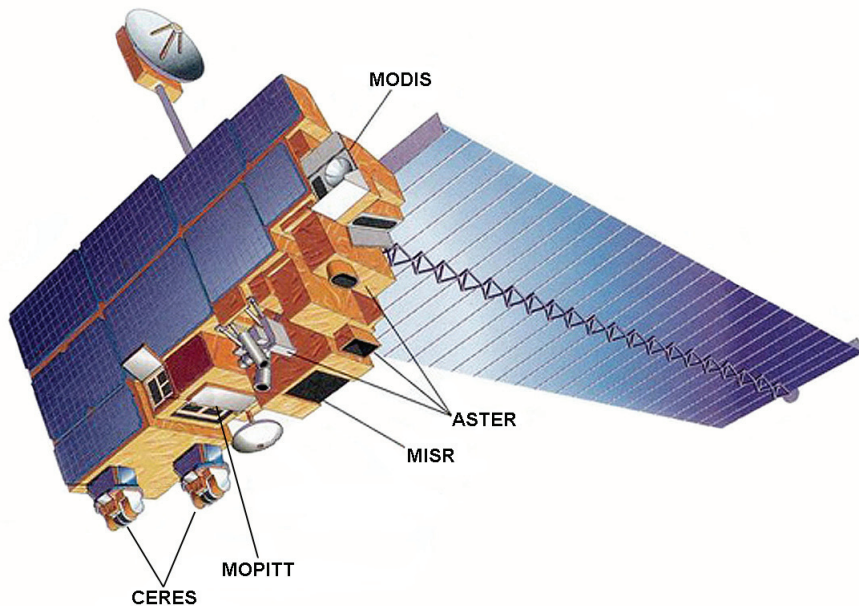


Figure 4-3. Terra spacecraft with instruments (Source: NASA).

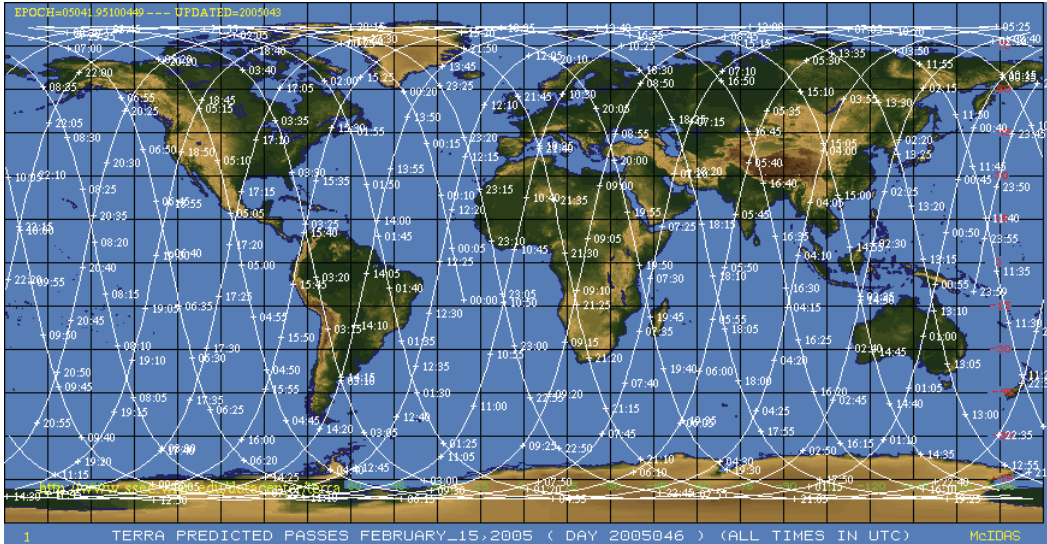


Figure 4-4. Sample Terra orbital tracks on 15 February 2005 (Source: NASA Distributed Active Archive Center at National Snow and Ice Data Center; nsidc.org).

Terra 위성의 사용 연한은 6년으로 계획되었으므로 조만간 수명이 다할 수 있다. 따라서 2011년 10월 28일 발사되어 관측을 시작한 NASA의 National Polar-orbiting Operational Environmental Satellite System Preparatory Project(NPP) 위성 자료 및 향후 계획된 NOAA의 Joint Polar Satellite System(JPSS) 프로그램과 연계하여 자료 생산 및 활용의 지속성을 유지하여야 할 것이다.

4.1.3 증발산량

실제 증발산량(Actual Evapotranspiration; ET)과 잠재 증발산량(Potential Evapotranspiration; PET) 자료는 Penman-Monteith 방정식을 기초로 한 Mu et al. (2011)의 알고리즘을 이용하여 미국 NASA의 Terra 위성에서 탑재된 MODIS 센서로부터 생산한 MODIS Level-4 자료인 MOD16 Global Terrestrial Evapotranspiration에서 도출하였다(Figure 4-5). MOD16의 8일, 월별, 연간 자료

Evapotranspiration에서 도출하였다(Figure 4-5). MOD16의 8일, 월별, 연간 자료 중 월별 자료를 사용하였는데, 모든 자료의 공간 해상도는 $1 \times 1\text{km}$ 이고 다른 MODIS 표준 생산물과 마찬가지로 Sinusoidal 도법으로 투영되어 있다. 자료는 미국 몬태나 주립대학교(University of Montana)의 Modeling and Monitoring Ecosystem Function at Multiple Scales: Numerical Terradynamic Simulation Group 홈페이지에서 획득하였다(<http://www.ntsg.umt.edu/project/mod16>).

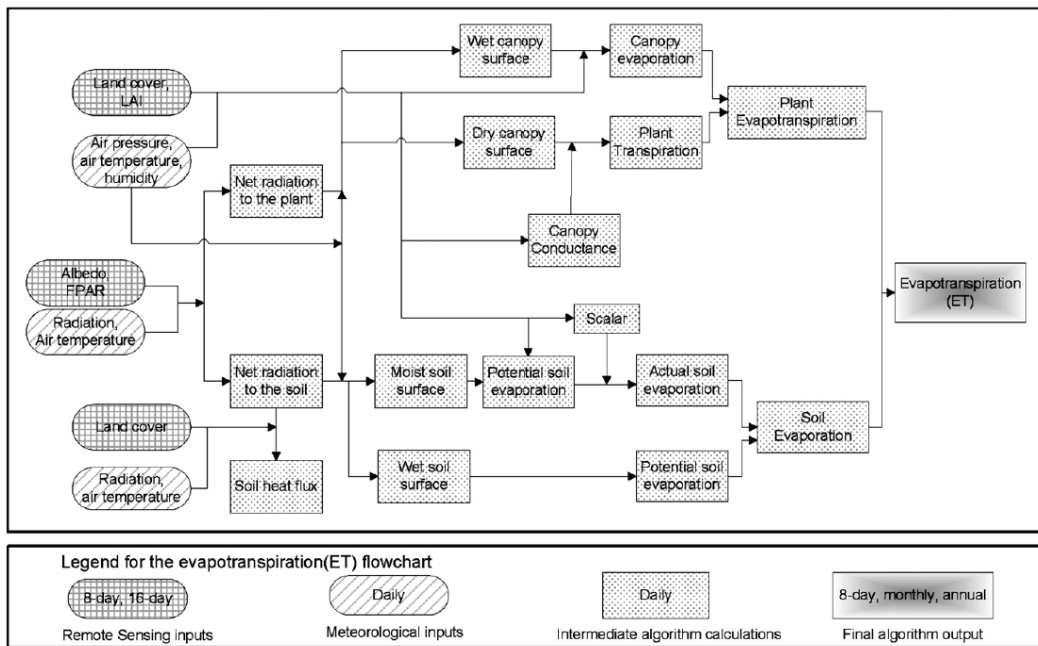


Figure 4-5. Flow chart of the MODIS evapotranspiration algorithm (Source: Mu et al., 2011).

MOD16 월별 자료는 mm/month 단위로 제공되는데, 강수량과 마찬가지로 여러 시간 척도에 따라 누적된 증발산량을 이용할 수 있도록 1, 3, 6, 9개월 동안 누적된 증발산량 값을 구하였다. 또한 가뭄 감시 및 예측 모델 개발에 활용할 수 있도록 기상관측소 위치에 따른 시계열 값을 추출하였으며, 행정구역별 모델 검증에 활용하기 위해 남한 전역에 대하여 $5 \times 5\text{km}$ 격자에 대해 리샘플(resample) 하였다.

4.1.4 광합성유효복사율

광합성유효복사율(Fractional Photosynthetically Active Radiation; FPAR)은 역시 미국 NASA의 Terra 위성에 탑재된 MODIS 센서로부터 생산된 Level-4 표준 자료인 MOD15A2 Leaf Area Index (LAI) and Fractional Photosynthetically Active Radiation (FPAR) 8-day L4 Global 1km에서 도출하였다. 지표면 온도와 마찬가지로 본 자료의 시간 해상도는 8일, 공간 해상도는 $1 \times 1\text{km}$ 이며, Sinusoidal 도법으로 투영되어 있다. 자료는 미국 NASA의 Earth Observing System Data and Information System, <http://reverb.echo.nasa.gov> (EOSDIS)에서 획득하였다.

FPAR는 식생의 임관(canopy)이 흡수하여 광합성 하는데 쓰이는 파장 영역($0.4 - 0.7\mu\text{m}$)의 비율을 측정하며 자료는 % 단위로 제공된다. FPAR 산정을 위한 알고리즘은 산림에 중점을 두어 개발되었다. 지표면 온도 자료와 마찬가지로 8일 시간 해상도의 본 자료는 월별 분석에 사용하기 위해 각 월에 해당하는 날 수를 가중치로 사용하여 월별 자료로 전환하였다. 그리고 가뭄 감시 및 예측 모델 개발에 활용할 수 있도록 기상관측소 위치에 따른 시계열 값을 추출하였으며, 행정구역별 모델 검증에 활용하기 위해 남한 전역에 대하여 $5 \times 5\text{km}$ 격자에 대해 리샘플(resample) 하였다.

4.1.5 엽면적지수

엽면적지수(Leaf Area Index; LAI)는 FPAR와 함께 MOD15A2 Leaf Area Index (LAI) and Fractional Photosynthetically Active Radiation (FPAR) 8-day L4 Global 1km에서 도출하였다. 따라서 시간 해상도는 8일, 공간 해상도는 $1 \times 1\text{km}$ 이며, Sinusoidal 도법으로 투영되어 있다. 자료는 미국 NASA의 Earth Observing System Data and Information System, <http://reverb.echo.nasa.gov> (EOSDIS)에서 획득하였다.

LAI는 바이오매스로 표현되는 FPAR에 상응하는 값으로, 지면 면적 중 엽면적 비율로 제공된다($\text{m}^2\text{plant}/\text{m}^2\text{ground}$). 월별 분석에 사용하기 위해 각 월에 해당하는 날 수를 가중치로 사용하여 8일 자료를 월별 자료로 전환하였으며, 가뭄 감시 및 예측 모델

개발에 활용할 수 있도록 기상관측소 위치에 따른 시계열 값을 추출하였으며, 행정구역별 모델 검증에 활용하기 위해 남한 전역에 대하여 5×5km 격자에 대해 리샘플(resample) 하였다.

4.1.6 식생지수

식생지수인 Normalized Difference Vegetation Index(NDVI)와 Enhanced Vegetation Index(EVI)도 Terra 위성에 탑재된 MODIS 센서로부터 생산되는 Level-3 표준 자료인 MOD13A3 Vegetation Indices Monthly L3 Global 1km 로부터 도출되었다. 공간 해상도는 1×1km이며 Sinusoidal 도법으로 투영되어 있고, 자료는 역시 미국 NASA의 Earth Observing System Data and Information System, <http://reverb.echo.nasa.gov> (EOSDIS)에서 획득하였다.

적색광과 근적외선을 이용하여 산정되는 정규식생지수(NDVI)에 대해서는 2장에서 언급하였다. EVI는 청색광, 적색광, 근적외선을 함께 사용하여 토양, 적설, 물 등으로 인해 나타나는 임관의 배경 효과를 최소화하고 식생이 우거진 지역에 대해 NDVI에서 종종 나타나는 포화 현상을 줄여 민감도를 유지한다. 또한 청색광을 이용함으로써 연기나 얼은 구름으로 인해 나타나는 대기오염 효과를 제거한다.

$$EVI = 2 \frac{(\rho_{NIR} - \rho_{RED})}{(L + \rho_{NIR} + C_1 \rho_{RED} + C_2 \rho_{BLUE})} \quad (\text{Eq. 4-5})$$

이때, ρ 는 대기 상층의 반사율 또는 지표에서의 방향성(directional) 반사율이며, L 는 임관의 배경 효과를 조절하기 위한 매개변수, C_1 과 C_2 는 적색광의 에어로졸 효과를 교정하기 위해 청색광을 함께 사용할 때 쓰이는 가중치이다(Liu and Huete, 1995).

4.1.7 토양수분

토양수분은 Aqua 위성에 탑재된 Advanced Microwave Scanning Radiometer-EOS(AMSR-E) 위성의 Level-3 표준 자료인 AE_Land3 Surface Soil Moisture 로

부터 도출하였다. 본 자료는 일별 자료로 약 $25 \times 25 \text{km}$ 의 공간 해상도를 가지고 있으며, 하루에 두 번 강교점 및 승교점에서 토양 수분을 측정한다. 자료는 Cylindrical Equal-Area Scalable Earth Grid (EASE-Grid) 도법으로 투영되어 있으며 g/cm^3 의 단위로 제공된다. 토양수분 자료는 미국 NASA의 Earth Observing System Data and Information System, <http://reverb.echo.nasa.gov> (EOSDIS)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), MODIS는 아직 자료를 잘 생산하고 있다(Figure 4-6). AMSR-E는 2011년 10월 4일 안테나 작동 중지로 인해 더 이상 자료를 생산하지 않는다. 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 (Figure 4-7).

본 자료는 일별 시간 해상도를 가지므로 월별 분석에 사용하기 위해 월별로 평균값을 계산하였다. 즉, 일별 오전 및 오후의 토양 수분 자료는 월별 오전, 오후, 그리고 평균 토양 수분 자료로 전환되었다. 또한 가뭄 감시 및 예측 모델 개발에 활용할 수 있도록 기상관측소 위치에 따른 시계열 값을 추출하였으며, 행정구역별 모델 검증에 활용하기 위해 남한 전역에 대하여 $5 \times 5 \text{km}$ 격자에 대해 리샘플(resample) 하였다.

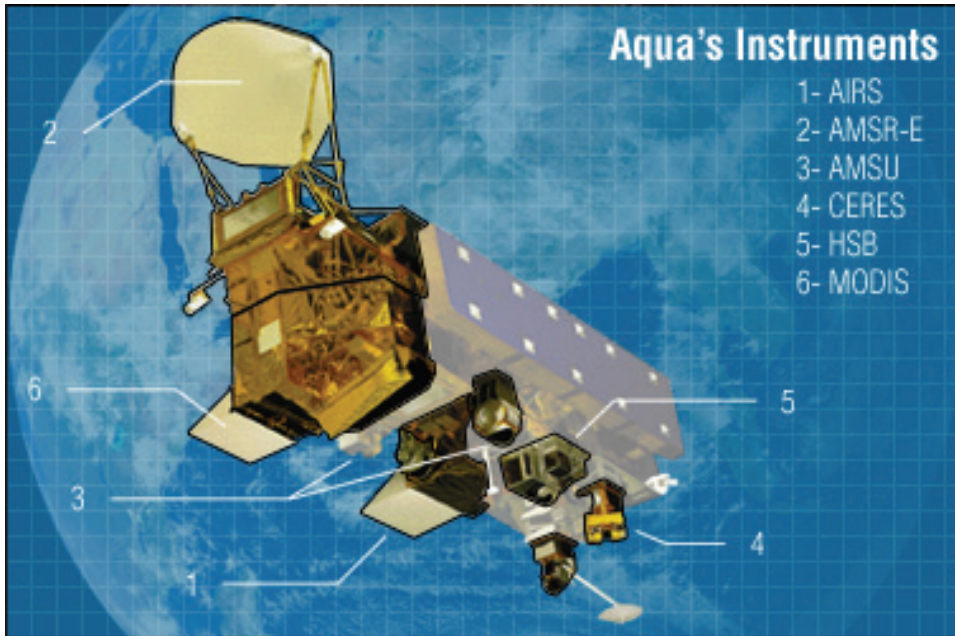


Figure 4-6. Aqua's instruments (Source: NASA).

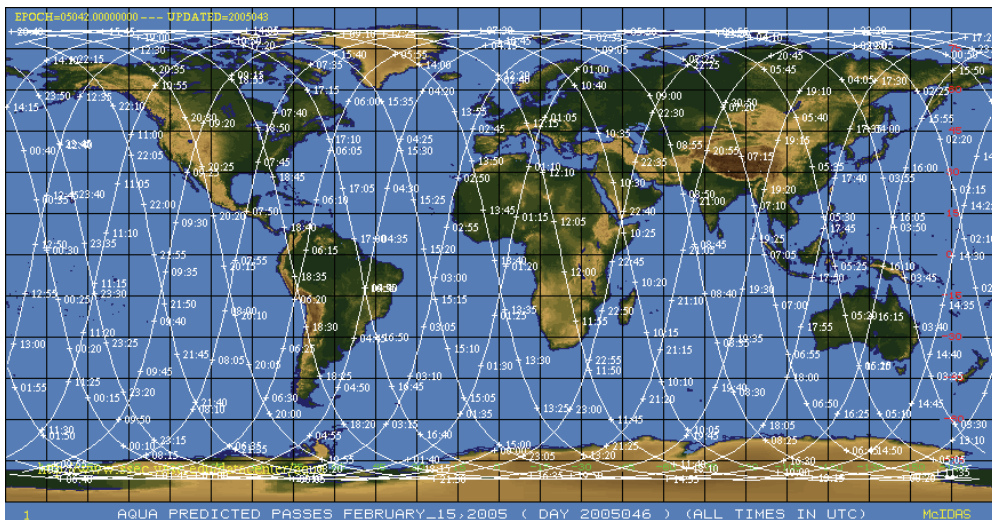


Figure 4-7. Sample Aqua orbital tracks on 15 February 2005 (Source: NASA Distributed Active Archive Center at National Snow and Ice Data Center; nsidc.org).

Aqua 위성의 사용 연한도 역시 6년으로 계획되어 있었으므로 Terra 위성과 마찬가지로 조만간 수명이 다할 수 있다. AMSR-E는 2011년 9월까지의 자료밖에 생산하지 못하고 멈추었으나 앞으로의 위성 영상 기반의 토양 수분 자료는 2015년 1월 29일 발사 예정인 Soil Moisture Active Passive(SMAP)로부터 얻을 수 있을 것이다. SMAP은 능동형 마이크로파 센서인 Synthetic Aperture Radar(SAR)와 수동형 마이크로파 센서인 복사계(radiometer)로부터의 자료를 결합하여 고해상도의 정확도 높은 측정 자료를 생산할 것이다. SMAP의 적도 통과 시간은 오전 6시와 오후 6시이다. SAR로부터의 자료는 3×3km, 복사계로부터의 자료는 대략 36×36km의 공간 해상도를 가지며 이들을 결합한 Level-3 토양 수분 자료의 해상도는 9×9km 가 될 것이다.

4.2 원격 탐사 기반 가뭄 지수

4.2.1 표면반사율

위성 영상 기반의 가뭄 지수를 산정하기 위해 Terra 위성 및 Aqua 위성에 탑재된 MODIS 센서로부터의 Level-3 표준 자료인 Nadir BRDF-Adjusted Reflectance 16-Day L3 Global 1km 로부터 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 1×1km 이다. 또한 자료는 Sinusoidal 도법으로 투영되었고, 미국 NASA의 Earth Observing System Data and Information System, <http://reverb.echo.nasa.gov> (EOSDIS)에서 획득하였다.

MODIS의 36개 밴드 중 7개의 표면반사율이 제공되며 16일 자료는 월별 분석에 사용하기 위해 각 월에 해당하는 날 수를 가중치로 사용하여 월별 자료로 전환하였다. 또한 강수량 자료와 마찬가지로 가뭄 감시 및 예측 모델 개발에 활용할 수 있도록 기상관측소 위치에 따른 시계열 값을 추출하였으며, 행정구역별 모델 검증에 활용하기 위해 남

한 전역에 대하여 5×5km 격자에 대해 리샘플(resample) 하였다.

4.2.2 원격 탐사 기반 가뭄 지수의 산정

위성 영상의 표면반사율을 이용하여 구할 수 있는 가뭄 지수인 NDWI, NDDI, NMDI의 식은 2장에서 소개하였다. Gao(1996)가 개발한 NDWI는 0.86 μ m 근적외선(NIR) 영역과 1.24 μ m 단파 적외선(SWIR) 영역을 이용하였는데, 이는 MODIS의 밴드 2와 밴드5에 해당한다.

Rhee et al.(2010)는 가뭄 상태를 평가하기 위해 NDWI를 사용하며 MODIS의 SWIR 밴드로 밴드5 이외에도 밴드6, 밴드7로 사용하여 시험하였는데, 때로 지역 및 비교하는 관측 기반의 가뭄 지수에 따라 밴드6과 밴드7을 사용한 경우 밴드5를 사용한 경우보다 가뭄의 상태가 잘 드러났다. 본 연구에서도 NDWI, NDDI의 산정에 있어 밴드5 이외에도 밴드6, 밴드7도 사용하였으며 각각 NDWI5, NDWI6, NDWI7 및 NDDI5, NDDI6, NDDI7로 표기하였다. NMDI는 MODIS의 밴드2, 밴드6 및 밴드7을 사용하였다.

4.3 관측 자료

4.3.1 종관기상관측자료

원격 탐사 기반 자료의 정확성을 살펴보고, 관측 자료 기반의 가뭄 지수를 산정하여 가뭄 평가의 척도로 사용하기 위해서 기상청에서 종관기상관측(Automatic Synoptic Observation System; ASOS) 자료를 획득하였다(Figure 1-3). 한반도 남한 지역의 79개 ASOS 관측 지점과 북한 지역의 27개 관측 지점에 대해 1980~2012년의 월별 강수량, 최대온도, 최소온도, 평균온도 자료를 이용하였다(Table 4-1). 북한 자료는 북한 측에서 기상청으로 제공한 3시간 단위의 자료를 일별로 전환한 것이다.

4.3.2 KoFlux 자료

본 연구에서는 위성 영상 기반 토양 수분 자료 및 지면 모델 결과로 산출된 토양 수

와 연세대학교 지구환경연구소가 2000년 기상청의 「Korea Monsoon Experiment (KORMEX) 지표 플럭스 자료의 품질관리 및 지면 모수화」사업을 수행하기 위해 광릉 수목원 침엽수림에 플럭스 관측 타워를 구축함으로써 시작되었으며, 2002년 KoFlux가 공식으로 출범하며 아시아 플럭스 관측망인 AsiaFlux 와 전지구 플럭스 관측망인 FLUXNET에 정식 등재되었다(권효정, 김준, 2010).

Table 4-1. Weather stations used in the Korean Peninsula.

Station ID	Station Name	Latitude (dd)	Longitude (dd)	Elevation (m)	Station ID	Station Name	Latitude (dd)	Longitude (dd)	Elevation (m)
90	Sokcho	38.25	128.55	18.1	232	Cheonan	36.77	127.12	21.3
95	Cheorwon	38.13	127.30	153.7	235	Boryeong	36.32	126.55	15.5
98	Dongducheon	37.90	127.05	109.1	236	Buyeo	36.27	126.92	11.3
99	Munsan	37.88	126.75	29.4	238	Geumsan	36.10	127.47	170.4
100	Daegwallyeong	37.67	128.72	772.6	243	Buan	35.72	126.70	12.0
101	Chuncheon	37.90	127.73	77.7	244	Imsil	35.60	127.28	247.9
102	Baengnyeongdo	37.95	124.62	144.9	245	Jeongeup	35.55	126.85	44.6
104	Bukgangneung	37.80	128.85	78.9	247	Namwon	35.40	127.32	90.3
105	Gangneung	37.75	128.88	26.0	248	Jangsu	35.65	127.52	406.5
106	Donghae	37.50	129.12	39.9	256	Juam	35.07	127.23	74.6
108	Seoul	37.57	126.95	85.8	260	Jangheung	34.68	126.92	45.0
112	Incheon	37.47	126.62	68.2	261	Haenam	34.55	126.57	13.0
114	Wonju	37.33	127.93	148.6	262	Goheung	34.62	127.27	53.1

Station ID	Station Name	Latitude (dd)	Longitude (dd)	Elevation (m)	Station ID	Station Name	Latitude (dd)	Longitude (dd)	Elevation (m)
115	Ulleungdo	37.47	130.88	222.8	271	Bongwhoa	36.93	128.90	319.8
119	Suwon	37.27	126.98	34.1	272	Yeongju	36.87	128.52	210.8
121	Yeongwol	37.17	128.45	240.6	273	Mungyeong	36.62	128.13	170.6
127	Chungju	36.97	127.95	115.1	277	Yeongdeok	36.52	129.40	42.1
129	Seosan	36.77	126.48	28.9	278	Uiseong	36.35	128.68	81.8
130	Ulsan	36.98	129.40	50.0	279	Gumi	36.12	128.32	48.9
131	Cheongju	36.63	127.43	57.2	281	Yeongcheon	35.97	128.95	93.6
133	Daejeon	36.37	127.37	68.9	284	Geochang	35.67	127.90	221.0
135	Chupungnyeong	36.22	127.98	244.7	285	Hapcheon	35.55	128.17	33.1
136	Andong	36.57	128.70	140.1	288	Miryang	35.48	128.73	11.2
137	Sangju	36.40	128.15	96.2	289	Sancheong	35.40	127.87	138.1
138	Pohang	36.02	129.37	2.3	294	Geoje	34.88	128.60	46.3
140	Gunsan	36.00	126.75	23.2	295	Namhae	34.80	127.92	45.0
143	Daegu	35.88	128.62	64.1	47003	Seonbong	42.32	130.40	3.0
146	Jeonju	35.82	127.15	53.4	47005	Samjiyeon	41.82	128.32	1386.0
152	Ulsan	35.55	129.32	34.6	47008	Cheongjin	41.78	129.82	43.0
155	Changwon	35.17	128.57	37.2	47014	Junggang	41.78	126.88	332.0
156	Gwangju	35.17	126.88	72.4	47016	Hyesan	41.40	128.17	714.0
159	Busan	35.10	129.02	69.6	47020	Gangkye	40.97	126.60	306.0

Station ID	Station Name	Latitude (dd)	Longitude (dd)	Elevation (m)	Station ID	Station Name	Latitude (dd)	Longitude (dd)	Elevation (m)
162	Tongyeong	34.83	128.43	32.7	47022	Pungsan	40.82	128.15	1206.0
165	Mokpo	34.82	126.37	38.0	47025	Gimchek	40.67	129.20	23.0
168	Yoesu	34.73	127.73	64.6	47028	Supung	40.45	124.93	83.0
169	Heuksando	34.68	125.45	76.5	47031	Jangjin	40.37	127.25	1081.0
170	Wando	34.38	126.70	35.2	47035	Sinuiju	40.10	124.38	7.0
172	Gochang	35.33	126.58	52.0	47037	Guseong	39.98	125.25	99.0
174	Suncheon	35.02	127.37	165.0	47039	Huicheon	40.17	126.25	155.0
175	Jindo	34.47	126.32	476.5	47041	Hamheung	39.93	127.55	38.0
184	Jeju	33.50	126.52	20.4	47046	Sinpo	40.03	128.18	19.0
185	Gosan	33.28	126.15	74.3	47050	Anju	39.62	125.65	27.0
188	Seongsan	33.38	126.87	17.8	47052	Yangdeok	39.17	126.83	279.0
189	Seogwipo	33.23	126.55	49.0	47055	Weonsan	39.18	127.43	36.0
192	Jinju	35.15	128.03	30.2	47058	Pyeongyang	39.03	125.78	38.0
201	Ganghwa	37.70	126.43	47.0	47060	Nampo	38.72	125.37	47.0
202	Yangpyeong	37.48	127.48	48.0	47061	Jangjeon	38.73	128.18	35.0
203	Icheon	37.25	127.48	78.0	47065	Sariwon	38.52	125.77	52.0
211	Inje	38.05	128.17	200.2	47067	Sinkyae	38.50	126.53	100.0
212	Hongcheon	37.68	127.87	140.9	47068	Yongyeon	38.20	124.88	5.0
216	Taeback	37.17	128.98	712.8	47069	Haeju	38.03	125.70	81.0
221	Jecheon	37.15	128.18	263.6	47070	Gaeseong	37.97	126.57	70.0
226	Boeun	36.48	127.73	175.0	47075	Pyeonggang	38.40	127.30	371.0

국내 플럭스 관측지로는 경기도 포천시 광릉 국립수목원내 활엽수림 및 침엽수림, 경기도 파주 설마천의 혼효림, 경기도 강화군 하리 수원과 여주 청미천의 논, 전라남도 해남의 농경지, 이어도 해양이 있다. 본 연구에서는 경기도 파주 설마천의 혼효림과 전라남도 해남의 농경지 지점에 대해서 토양 수분 자료를 도출하여 위성 영상 기반의 토양수분 자료 및 지면 모델 결과와 비교하였다(Table 4-2; Figure 4-8). 두 지점 모두 토양 수분 자료는 Time-Domain Reflectometer(TDR) Soil Water Content 센서로 측정된 상층 10cm까지의 값이며, 30분마다 % (vol/vol) 단위로 제공되었다.

Table 4-2. List of KoFlux sites used in this study.

Site	Ecosystem Type	Latitude (dd)	Longitude (dd)	Data Period Used	Research Institute
Seolmacheon	Mixed Forest	37.94	126.96	2008	Yonsei Univ.
Haenam	Farmland	34.55	126.56	2004~2006, 2008	Yonsei Univ.

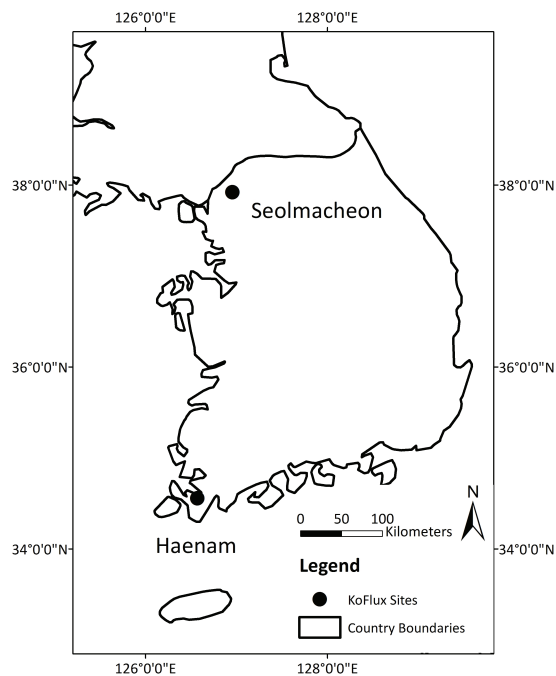


Figure 4-8. Two KoFlux sites used in this study.

4.3.3 기후 지수

지역에 따라 장기적인 가뭄 현상은 대규모 기후 지수와 상관관계를 보이기도 한다. 특히 ENSO는 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 북극진동(Arctic Oscillation; AO)을 가뭄 감시 및 예측에 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. 1950년부터의 MEI 자료는 미국 NOAA의 Earth System Research Laboratory 웹사이트에서 획득하였다(<http://www.esrl.noaa.gov/psd/enso/mei>; Figure 4-9).

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). 1950년부터의 AO 자료는 미국 NOAA의 기후예측센터 웹사이트에서 획득하였다(www.cpc.ncep.noaa.gov/products/precip/CWlink/daily_ao_index/monthly.ao.index.b50.current.ascii).

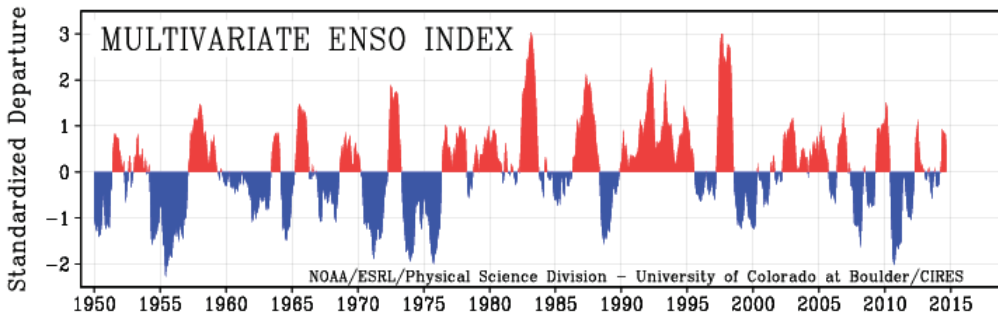


Figure 4-9. Multivariate ENSO Index (Source: www.esrl.noaa.gov/psd/enso/mei).

4.3.4 고랭지 채소 수확량

국내 주요 밭작물은 과채류, 엽채류, 근채류, 조미채소 등으로 배추, 무, 양파, 마늘 등이 이에 속한다. 위성 영상에서 도출한 변수를 결합하여 개발된 가뭄 감시 및 예측 모델의 검증을 위해 고랭지 배추 및 무의 수확량을 이용하였는데, 시도별 수확량은 국가통계포털에서(www.kosis.kr), 강원도의 시군구별 수확량은 강원도청 농식품유통과에서 획득하였다(전승택, 2014). 시도별 자료는 강원도, 전라북도, 경상북도, 경상남도에 대해서 2003-2011년 동안 kg/10a의 단위로 주어졌으며, 강원도의 시군구별 자료는 강릉시, 고성군, 동해시, 삼척시, 속초시 등 18개 시군구에 대해서 2007-2013년 동안 tonne/ha의 단위로 제공되었다(Figure 4-10).

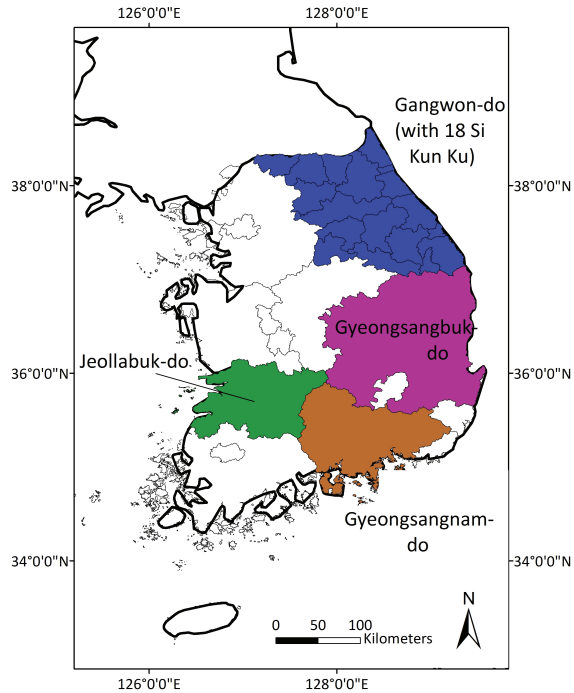


Figure 4-10. Four SI-DO and 18 SI-KUN-KU provinces with crop yield data (highland bok choy and radish).

4.3.5 산불 발생 빈도 및 피해 면적

위성 영상에서 도출한 변수를 결합하여 본 연구에서 개발된 가뭄 감시 및 예측 모에 적용해 보기 위해 연도별, 지역별 산불 통계를 산림청 「산불통계연보」 및 산림청 웹사이트(www.forest.go.kr)로부터 획득하였다. 2003년 1월부터의 산불 발생 통계 자료는 시도 및 시군구별 발생장소, 발생원인 및 피해면적 등으로 구성되어 있는데, 피해면적은 산불이 발생되어 지상입목, 관목, 시초 등을 연소시키면서 실제로 산불이 지나간 면적을 말하며, ha 단위로 제공된다.

시군구별로 피해면적을 합산하여 전체 산지 면적 대비 비율을 계산하였으며 월별로

최댓값, 최솟값을 이용하여 표준화하여 지역별, 월별로 표준화된 시계열 자료를 산정하였다. 시군구별 산지 면적은 「산지관리법」에 의하여 지정된 보전산지와 준보전산지 면적의 합을 이용하였는데, 이 자료는 산림청으로부터 획득하였다. 산지 면적으로 가장 최근의 2012년 자료 값을 사용하였는데, 1월에서 5월까지의 지역별 면적이 동일하게 조사되었으며 6월부터 12월까지 동일하게 조사되었다. 두 자료 사이에 큰 차이가 없으므로 산지 면적으로 2012년 5월의 값을 이용하였다.

산불 발생 통계 자료는 세종특별자치시 출범 이전의 자료도 현재의 행정구역으로 전환하여 제공하고 있으나, 산지 면적 자료는 그렇지 않으므로 세종특별자치시 행정구역으로 편입된 충청남도 공주시 일부와 충청북도 청원군 일부의 면적을 고려하지 않고 충청남도 연기군의 산지 면적을 세종특별자치시 산지 면적으로 사용하였다.

산불 발생의 원인은 대부분 입산자 실화, 담뱃불 실화, 성묘객 실화, 쓰레기 소각 등 사람에게 의한 것으로, 본 연구에서는 화재 요인이 발생하였을 때 실제 화재로 이어질 수 있는 건조한 가뭄 조건만을 고려하기 위해 발생원인 별로 분석하지 않고 전체적인 분석만을 수행하였다.

4.4 관측 자료 기반 가뭄 지수

가뭄 감시 및 예측 성능 평가의 척도로 사용하기 위해서 ASOS 자료를 이용하여 관측 자료 기반의 가뭄 지수를 산정하였다. 가뭄 지수로는 SPI와 SPEI를 이용하였는데, 이들의 산정 과정은 2장에 자세히 언급하였다.

4.4.1 SPI의 산정

Guttman(1999)이 제안한 대로 Pearson Type III 확률분포함수를 이용하여 1981년 1월부터 2012년 12월까지 월별 SPI 값을 산정하였다. 먼저 1개월, 3개월, 6개월, 9개월, 12개월의 시간척도에 따른 누적 강수량 시계열을 각각 계산하였고, 월별로 L-모멘트 값이 산정되었다. L-모멘트 값을 산정할 때는 강수량 값이 서로 독립적이라고 가정하였으며 이 값을 이용하여 Pearson Type III 확률분포함수의 매개변수 값이 산정되

었다. 매개변수 도출을 위한 기준 기간으로는 1981년에서 2010년까지 30년 동안을 사용하였다.

4.4.2 SPEI의 산정

SPEI 산정에는 SPEI 지수를 개발한 Vicente-Serrano et al.(2010)이 개발하여 배포한 SPEI 산정을 위한 R 패키지(버전 1.6; <http://cran.r-project.org/web/packages/SPEI/>)를 이용하였다. 먼저 강수량과 증발산량의 차를 구하여 1개월, 3개월, 6개월, 9개월, 12개월의 시간척도에 따른 누적 시계열을 계산하였는데, 이 패키지는 증발산량 계산에 최고 온도와 최저 온도를 바탕으로 미국 FAO(Food and Agriculture Organization)에서 활용하는 추정식인 Hargreaves 방정식을 이용하고 있다.

SPI 산정과 유사한 과정을 거쳐 SPEI를 계산하였는데, 단 누적 강수량 시계열 대신에 강수량과 증발산량 차의 누적 시계열이 사용되며, Pearson Type III 확률분포함수 대신에 Vicente-Serrano et al.(2010)이 추천한 Log-logistic 확률분포함수를 사용하였다. 역시 1981년에서 2010년까지 30년 동안을 매개변수 도출을 위한 기준 기간으로 사용하였고, 1981년 1월부터 2012년 12월까지 월별 SPEI 값을 산정하였다.

4.5 지면모델 자료

4.5.1 지면모델 개요

지면모델(Land Surface Model; LSM)은 일반적으로 기후 모델 안에서 지면의 토지 피복(Land Cover Change; LCC)의 초기 효과를 모의하는 모델을 뜻한다(Pitman, 2003). 지역 규모로 지면에서 일어난 변수들 간의 작용이 대륙 규모 기후의 변화를 유발할 수 있다는 여러 증거가 있으며, 또한 주요 지역의 지역 규모의 지면의 변화도 대기의 원격상관(teleconnection)을 통해 지리적으로 멀리 떨어진 지역에 심각한 변화를 미칠 수 있다는 증거들이 존재한다(Pitman, 2003). 지면의 변화는 물, 에너지, 탄소 등의 순환에 영향을 미치며 기후 모델은 여기에 민감하게 반응하므로 기후 모델 안에 지면모델이 반드시 필요하다(Pitman, 2003).

Manabe(1969)에 의해 개발된 지면모델이 제1세대 지면모델이라고 할 수 있는데, 단순하고 이상적인 해양과 대륙의 분포를 가지며, 계절 변화나 일변화를 고려하지 않았다(Figure 4-11). 제2세대 지면모델은 Deardorff(1997)에 의해 개발되었다고 볼 수 있는데, 두 층으로 이루어진 토양 모델과 한 층의 식생 모델을 포함하고 있다(Figure 4-12). 토양 온도, 토양 수분, 증발량의 모의나 식생의 매개변수 등을 개선시킴으로써 제2세대 지면모델은 대륙 지면의 물리적인 구현을 개선시켜 제1세대 지면모델보다 우수한 성능을 보였다.

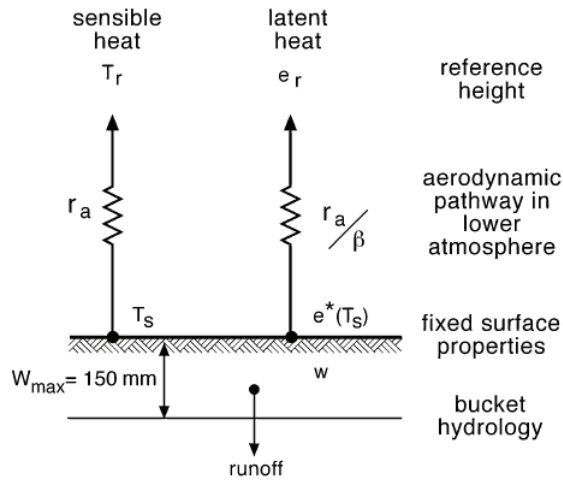


Figure 4-11. Illustration of a first-generation land surface model (Source: Sellers et al., 1997; recited in Pitman, 2003).

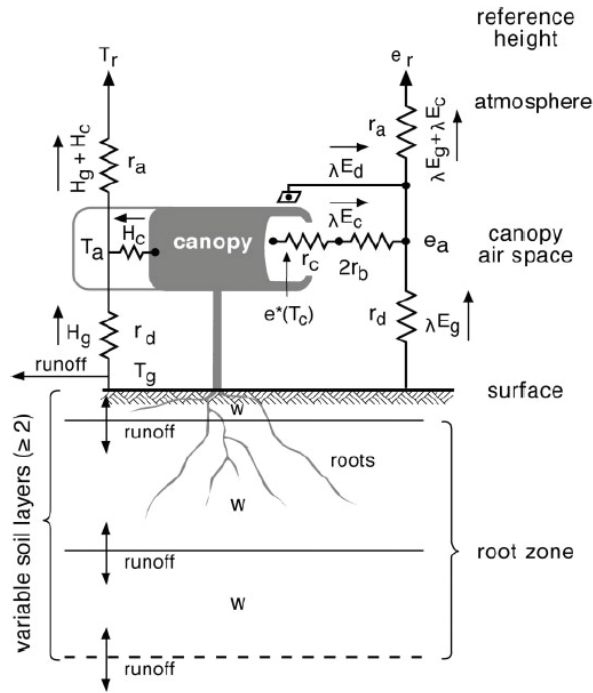


Figure 4-12. Illustration of a second-generation land surface model (Source: Sellers et al., 1997; recited in Pitman, 2003).

제2세대 지면모델은 식물과 환경적인 조건을 고려하여 경험적으로 임관의 전도도를 모델링하고 있는데 이를 단지 증산량의 모의에만 사용한다는 제한이 있었다. 1980년대 들어 임관의 전도도를 증발산량 모의를 개선시키고 식물의 탄소 흡수량을 설명하는데 이용하기 시작하였는데, 이렇게 식생의 Phenology를 결합시킨 제3세대 지면모델이 개발되었다(Figure 4-13). 그리고 여기에 생태 모델링이 결합되어 순탄소수지(net carbon balance)를 할당하는 등 전체 육상 탄소 순환을 고려 DGVM(Dynamic Global Vegetation Model)으로 발전하였다(Figure 4-14).

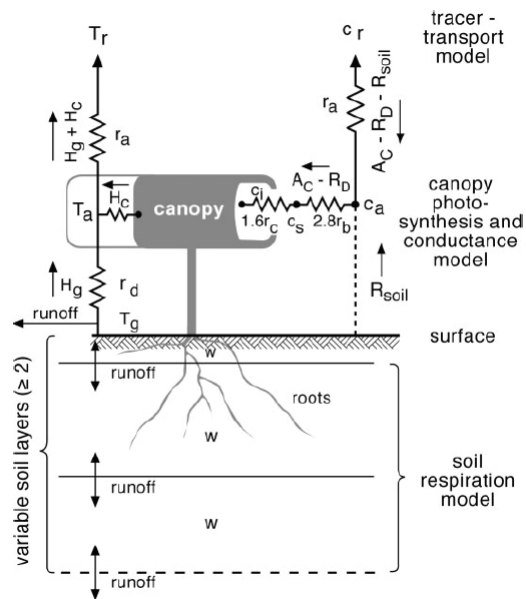


Figure 4-13. Illustration of a third-generation land surface model (Source: Pitman, 2003).

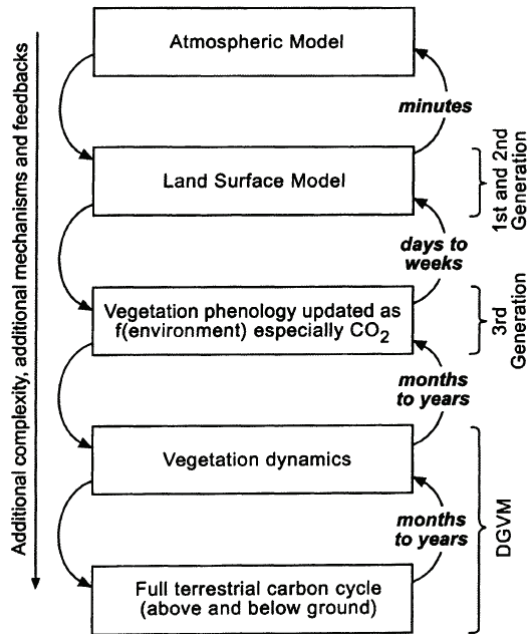


Figure 4-14. Levels of land surface modelling approaches (Source: Pitman, 2003).

지면모델의 하나인 Community Land Model(CLM)은 다양한 시공간적 규모로 육상생태계와 기후가 영향을 주고받는 물리적, 화학적, 생물학적 과정을 포함하고 있으며, 생지물리학(biogeophysics), 수문학적 순환(hydrologic cycle), 생지화학(biogeochemistry) 및 역학적 식생(dynamic vegetation)의 구성요소로 이루어진다. 현재 최신 버전은 4.0이며 3.5 버전 이후 MODIS에 기반한 지표 자료를 활용하는 등 원격 탐사를 융합시키고 있다.

중규모 지면모델의 하나인 Variable Infiltration Capacity(VIC) 모델은 University of Washington과 Princeton University에서 개발되어 세계적으로 사용되고 있으며(Figure 4-15), 전구 규모의 CLM 등과는 다르게 강의 흐름을 고려하여 구역별로 분석이 가능하다(Gao et al. 2010). 입력 자료로는 수치고도모델, 토지피복도, 정밀토양도 및 기상 자료 등이며 입력 자료 및 매개변수에 원격 탐사 자료를 결합한 여지가 존재한다.

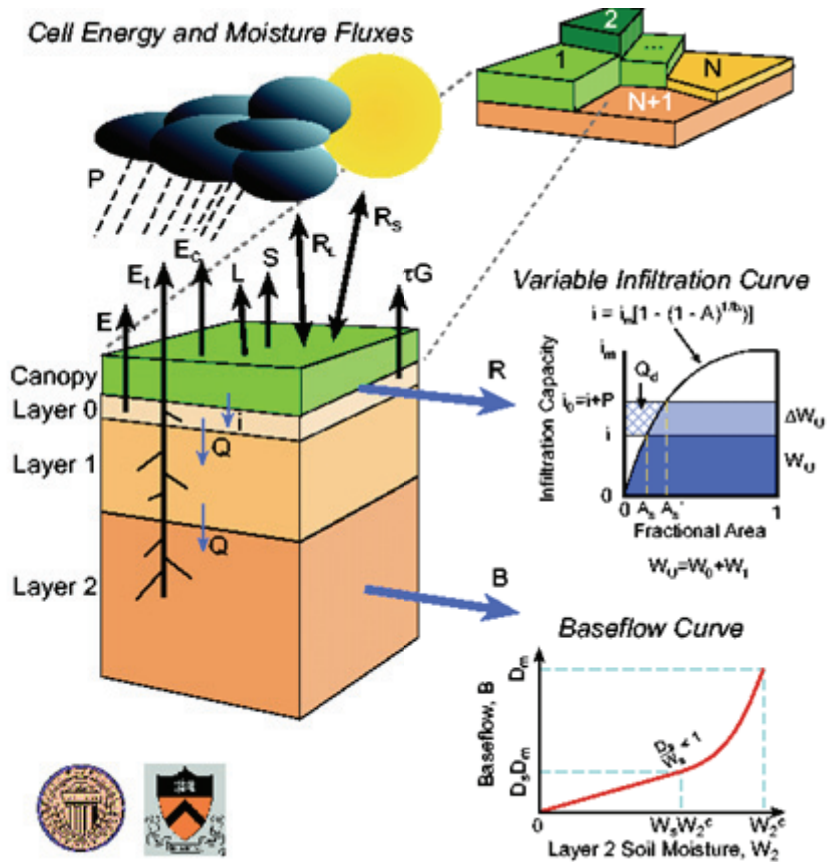


Figure 4-15. VIC-3L model (Source: www.hydro.washington.edu/Lettenmaier/Models/VIC/).

4.5.2 NOAH-LSM

토양수분 추정 개선을 시험하기 위해 본 연구에서는 APEC기후센터(2013b)에서 개선한 NOAH-LSM의 결과 자료를 활용하였다. NOAH-LSM은 미국의 National Centers for Environmental Prediction/Environmental Modeling Center(NCEP/EMC), 미국 오리건 주립대학교(Oregon State University), 미국 National Weather

Service Hydrology Laboratory(HRL), 미국 Air Force Weather Agency(AFWA), 미국 Air Force Research Laboratory(AFRL)에서 개발되었으며 토양수분, 토양온도, 지표면 온도, 적설심(snow depth), 적설 물당량(snowpack water equivalent), 증발산량 등을 모의한다.

본 연구에서는 남한 지역의 $5 \times 5\text{km}$ 격자에 대하여 NOAH-LSM 모델의 모의 결과로 생산되는 토양수분, 실제 증발산량, 잠재 증발산량의 일별 및 월별 자료를 획득하여 활용하였다. 토양수분 자료는 토양의 상층 10cm 에 대해서 kg/m^2 단위로 주어지며, 증발산량은 $\text{kg/m}^2\text{s}$ 단위로 제공된다.

4.5.3 지면모델 결과와 위성 영상 자료의 결합

AMSR-E 센서로부터의 토양수분 자료는 대략 $25 \times 25\text{km}$ 의 공간 해상도를 가지고 지면모델 결과는 $5 \times 5\text{km}$ 의 공간 해상도를 가지므로, 이 둘을 결합하기 위해 AMSR-E 토양수분 자료에 해당하는 $25 \times 25\text{km}$ 격자에 대해 지면모델 결과물인 토양수분 자료를 평균하였다. 결과적으로 각 $25 \times 25\text{km}$ 격자마다 위성 영상에서 추출한 토양수분 자료와 지면모델에서 얻은 토양수분 자료가 존재하게 된다.

4.6 기계학습을 통한 가뭄 감시 및 예측 모델 개발

4.6.1 인공지능과 전문가시스템

인공지능은 현재 사람이 더 잘하는 일을 컴퓨터가 더 잘하도록 연구하는 학문 분야로 1950년대 중반에 시작되었다. 1960년대 들어 인공지능이 과신되는 경향이 있었으나, 사람이 하는 일을 컴퓨터가 모두 잘하도록 하는 일이 실제로 매우 어렵다는 것이 판명되었으며 따라서 인공지능은 상대적으로 좁은 학문 분야에 적용하도록 발전되어 왔다(Jensen, 2005). 어떤 특정한 좁은 주제의 문제 해결에 있어 보통 인간의 지능이 요구되어 지는 지식을 활용하는 시스템을 일컫는 전문가 시스템(expert system)의 개념이

등장하였는데, 전문가 시스템을 통해 좁은 영역의 문제를 효율적, 효과적으로 해결할 수 있다(Jensen, 2005; Figure 4-16).

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 방법을 크게 셋으로 나누면, 전통적인 컴퓨터 프로그램을 통한 알고리즘을 이용하는 방법, 그리고 3장에서 설명한 인공 신경망을 이용하는 방법이 있으며, 자기발견적인(heuristic) 전문가 시스템을 이용하는 방법이 있다.

4.6.2 기계학습

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). 제3장에서 설명한 의사결정분지도 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).

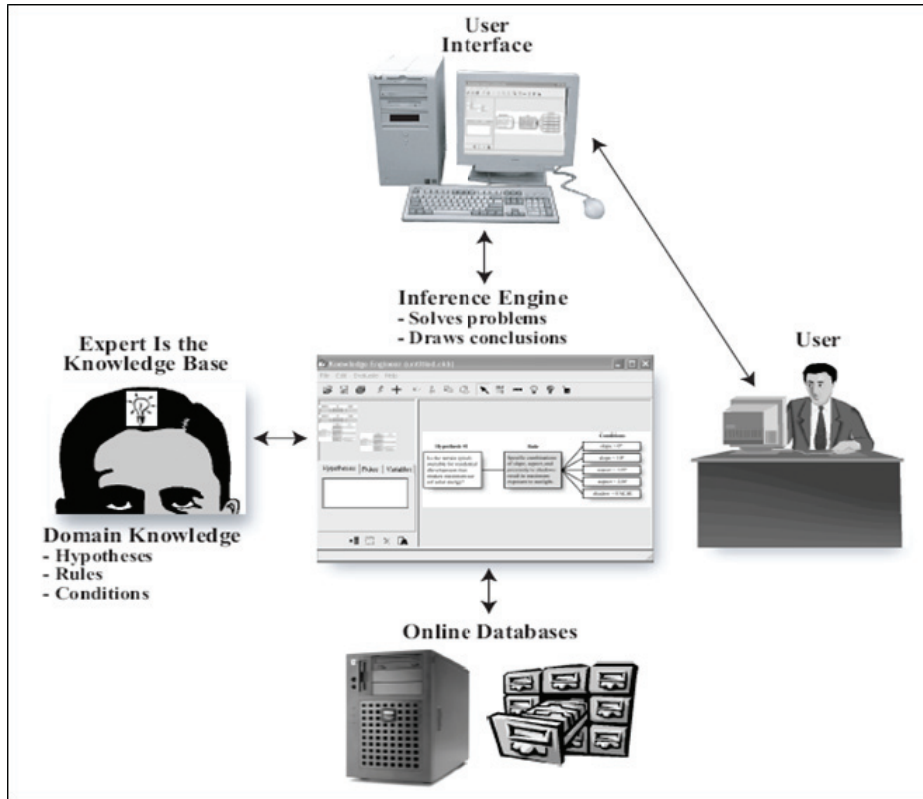


Figure 4-16. Components of rule-based expert system (Source: Jensen, 2005).

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 기계학습이나 인공 신경망 등의 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 감시하고 예측하는 모델을 개발하는 본 연구에 특히 적합하다.

4.6.3 Random Forests

Random forest는 제3장에 설명한 의사결정분지도(decision tree)에 기초한 기계학습 방법론으로, 의사결정분지도와 마찬가지로 분류(classification) 및 회귀(regression)을 목적으로 한다. 의사결정분지도는 하나의 트리(Tree; 분지도)만을 이용하는데 비하여 random forests는 여러 개의 트리를 앙상블하여 이용하므로 forest로 일컬으며, 그 과정에서 여러 임의적인 요소(randomness)를 삽입하여 random forest라고 명명되었다.

분류를 예를 들어 설명하면, random forest는 여러 트리 구조의 분류자(classifier)가 모여 앙상블된 하나의 분류자를 구성한다. 각각의 트리의 임의의 입력 벡터는 서로 독립적이지만 동일한 확률분포를 가지고 있으며, 모든 트리가 입력 벡터에 해당하는 가장 적합한 분류 클래스의 표결(vote)에 참여한다. 동일한 이론이 회귀에도 적용된다.

각 트리의 분류 및 회귀 능력(strength)은 향상시키고, 서로간의 상관관계(correlation)는 최소화하기 위해 여러 임의적인 요소가 들어가는데, 훈련 자료의 임의적인 부분집합(subset)을 이용하거나, 입력 벡터 또는 입력 벡터의 조합을 각 트리의 노드에서 임의적으로 선택하는 등의 방법이 포함된다. 이러한 임의적인 요소의 사용은 결과적으로 이상치(outlier)나 노이즈(noise)에 상대적으로 영향을 덜 받게 하고, 오차(generalization error), 분류자의 성능(strength), 트리 분류자 간의 상관관계(correlation), 또는 사용된 변수의 상대적인 중요도(variable importance) 등을 내부적으로 예측할 수 있게 한다.

주어진 훈련 자료의 집합에서 Bootstrap으로 임의로 1/3 가량의 새로운 훈련 자료 집합을 뽑아내어 트리를 키우는데(bagging), 이러한 방법을 통해 일반적으로 정확도가 증가하게 된다. 주어진 훈련 자료 집합에서 임의로 뽑아낸 자료의 집합들이 있으면 이에 따른 각각의 분류자들이 있는데, 각 분류자의 예측 변수(bagged predictor)를 평가할 때 이 입력 벡터가 포함되지 않은 나머지 분류자들로 표결하게 한다. 이를 out-of-bag 분류자라고 부른다. Out-of-bag(OOB) 과정을 통해 어떠한 변수가 분류 또는 회귀에

포함되지 않았을 때의 오차를 추정할 수 있으며, 또한 각 변수의 상대적인 중요도를 평가할 수 있다. 교차검증(cross-validation)과는 달리 훈련 자료의 일부를 따로 떼어 둘 필요가 없는 불편(unbiased) 오차 추정 값을 제공한다(Breiman, 2001).

4.6.4 모델의 개발

앞 절에서 설명한 random forest를 이용하여 가뭄 감시 및 예측 모델을 개발하였다. Random forest 방법은 이상치(outlier)와 노이즈(noise)에 영향을 덜 받으며 확고한(robust) 결과를 생산하고, 입력되는 여러 변수와 추정 및 예측하고자 하는 변수 사이의 물리적인 관계를 명확히 할 필요가 없으므로 특히 위성 영상 자료를 통해 가뭄 감시 및 예측을 수행하는 본 연구의 목적에 매우 잘 부합된다.

4.6.4.1 가뭄 감시 모델의 개발

가뭄의 상태를 정량화하는 목표 변수로 관측 자료를 이용하여 계산한 1, 3, 6, 9 개월 시간 척도의 SPI 및 SPEI와 위성 영상 자료인 표면반사도를 이용하여 얻은 가뭄 지수인 NDWI5, NDWI6, NDWI7, NDDI5, NDDI6, NDDI7, NMDI를 선정하였다.

관측 자료 기반의 SPI와 SPEI에 대해서는 한반도 106개 기상관측소 위치에 대해 자료를 추출하여 random forest 모델을 구축하였다. 위성 영상 기반의 NDWI, NDDI, NMDI에 대해서는 한반도를 대상으로 하는 전체 5×5km 격자에 대해 자료를 추출하여 모델을 구축하려 하였으나, 컴퓨팅 리소스의 제약으로 인해 모델링 시간이 너무 길어 역시 106개 기상관측소 위치에 대해 random forest 모델을 구축하였다.

입력 변수로는 위성 영상에서 도출한 입력 변수, 대규모 기후 지수, 지면모델 결과에서 도출한 입력 변수가 있다. 위성 영상에서 도출한 입력 변수는 AMSR-E의 승교점(A_SM), 강교점(D_SM) 및 평균 토양수분(SM_MEAN), MODIS 주간(LSTDAY), 야간(LSTNIGHT) 및 평균 지표면온도(LST_MEAN), EVI, NDVI, 실제 증발산량(ET), 잠재 증발산량(PET), TRMM의 강수량(PRCP), 유효광합성반사율(FPAR), 엽면적지수(LAI) 등이 있으며, 대규모 기후 지수는 MEI와 AO 지수를 사용하였고 지면모델 결과

는 실제 증발산량(EVAP), 잠재 증발산량(POTEVAP), 토양수분(SOILMOIST)가 있다.

토양수분 평균값의 일별 비교를 제외하고는 모델 구축은 월별 자료를 이용하였다. 각 지점 별로 환경 조건의 영향에 따라 도출되는 변수 값의 최솟값에서 최댓값에 이르는 폭이 다양하게 나타난다. 이 때, 각 지점에 대해 각각 모델을 구축한다면 상관없으나 본 연구에서는 모든 지점의 자료를 통합하여 대상 지역에 대해 단일한 모델을 구축하므로, 각 지점별로 최댓값과 최솟값을 이용하여 자료를 표준화하였다.

각 입력 변수 집합을 포함시킬지의 여부에 따라 네 유형의 모델을 구축하였다(Table 4-3). 위성 영상에서 도출한 입력 변수와 대규모 기후 지수는 한반도 전체에 대하여 자료를 획득하였으므로 106개 기상관측 지점에 대한 자료를 모두 이용하여 모델을 구축하였으나, 지면모델 결과를 이용하는 경우 남한 지역에 대해서만 자료가 존재하므로 79개 남한의 기상관측 지점에 대한 자료만을 이용하여 모델을 구축하였다.

$$D_t = f(SET_{x,t}) \quad (\text{Eq. 4-6})$$

이 때, D 는 목표 변수의 상태, t 는 현재 달을 나타내고 x 는 입력 변수 집합으로 SET1부터 SET4까지(Table 4-3)를 나타낸다. SPI와 SPEI를 이용한 목표 변수에 대해서는 입력 변수에서 강수량, 실제 증발산량, 잠재 증발산량의 누적 시간 척도를 목표 변수의 시간 척도와 일치시켰으며, 위성 영상에서 도출된 목표 변수에 대해서는 입력 변수에 강수량, 실제 증발산량, 잠재 증발산량을 1, 3, 6개월에 대하여 누적시킨 값을 각각 개별적인 변수로 사용하였다. Random forest는 입력 변수들 간의 독립성을 가정하지 않고 임의성을 이용하여 최적의 변수를 선택하므로 변수들 간의 다중공선성(multicollinearity)은 고려하지 않았다.

Table 4-3. Input variables of random forest models for drought monitoring.

Type	Input Variables	SET1 (RS+CI)	SET2 (RS+LSM)	SET3 (ALL)	SET4 (RS)
Remote Sensing (RS)	AMSR-E A_SM	O	O	O	O
	AMSR-E D_SM	O	O	O	O
	AMSR-E SM_MEAN	O	O	O	O
	MODIS LST_MEAN	O	O	O	O
	MODIS LSTDAY	O	O	O	O
	MODIS LSTNIGHT	O	O	O	O
	MODIS EVI	O	O	O	O
	MODIS NDVI	O	O	O	O
	MODIS ET	O	O	O	O
	MODIS PET	O	O	O	O
	TRMM PRCP	O	O	O	O
Climate Index (CI)	AO	O		O	
	MEI	O		O	
Land Surface Model (LSM)	LSM EVAP		O	O	
	LSM POTEVAP		O	O	
	LSM SOILMOIST		O	O	

4.6.4.2 가뭄 예측 모델의 개발

가뭄 예측 모델의 경우 가뭄 감시 모델과 유사하게 개발되었다. 우선 가뭄 감시 모델의 개발과 같이 많은 입력 변수를 포함시킨 경우를 살펴보고(SET3), 추가적으로 원격 탐사 기반 자료의 사용 가능성을 보기 위해 입력 변수 중 대규모 기후 지수와 지면모델 결과는 고려하지 않고 위성 영상에서 도출한 입력 변수만을 살펴보았다(SET4).

본 연구에서 시험한 선행 예측 시간은 1개월에서 6개월까지이며($t + n$; $n = 1 \sim 6$), 입력 변수와 예측하고자 하는 목표 변수의 과거 값을 입력 자료로 사용하였는데 현재달의 값만 입력 변수에 포함시키는 경우(t 만 사용; number of past months used, $pm = 1$), 현재 달과 전 달의 값을 입력 변수에 포함시키는 경우($t, t-1$ 을 사용; $pm = 2$), 현재 달, 전 달, 전전 달의 값을 입력 변수에 포함시키는 경우($t, t-1, t-2$ 을 사용; $pm = 3$)로 나누어 분석하였다.

$$D_{t+n} = f(SET_{3,t}, D_t) \quad (\text{Eq. 4-7})$$

$$D_{t+n} = f(SET_{3,t}, SET_{3,t-1}, D_t, D_{t-1}) \quad (\text{Eq. 4-8})$$

$$D_{t+n} = f(SET_{3,t}, SET_{3,t-1}, SET_{3,t-2}, D_t, D_{t-1}, D_{t-2}) \quad (\text{Eq. 4-9})$$

$$D_{t+n} = f(SET_{4,t}, D_t) \quad (\text{Eq. 4-10})$$

$$D_{t+n} = f(SET_{4,t}, SET_{4,t-1}, D_t, D_{t-1}) \quad (\text{Eq. 4-11})$$

$$D_{t+n} = f(SET_{4,t}, SET_{4,t-1}, SET_{4,t-2}, D_t, D_{t-1}, D_{t-2}) \quad (\text{Eq. 4-12})$$

이 때, D 는 목표 변수의 상태, t 는 현재 달을 나타내고 n 은 선행 예측 시간으로 1개월에서 6개월 까지, pm 은 사용된 과거 개월 수로 1개월에서 3개월 까지를 나타낸다.

4.6.5 모델의 적용

4.6.5.1 고랭지 채소 수확량

배추나 무 등의 작물은 90% 이상이 수분으로 이루어져 있으므로 기뭄 등의 영향으로 수분이 부족하면 모든 생리현상의 장애를 수반하여 성장과 생육의 지연이 발생하고 품질이 저하되게 된다(김기덕, 2014). 배추의 경우 기뭄이 발생하면 결구되지 않아 생장이 늦어지며, 칼슘 결핍증 등이 발생하여 수확이 불가능해 질 수도 있다(김기덕, 2014). 기뭄이 지속되면 기상청에서 제공하는 기상 정보를 기반으로 농진청에서 언론을

통해 농업인들에게 홍보하고 있는데, 이는 조기경보 보다는 대처의 성격을 가진다(김기덕, 2014). 가뭄의 피해를 줄이기 위해서는 고랑 높이를 낮추고 토양비옥도와 보습성을 향상시키는 경종적 방법, 스프링클러, 관정, 양수기 등을 이용하여 가뭄에 강한 기반을 조성하는 방법이 있으며 근본적으로 농업용수가 충분한 곳으로 재배지를 선정할 수 있다(김기덕, 2014).

가뭄과 관련된 요인들 중에 무와 배추 등 작물 생육에 가장 크게 관여하는 것은 온도와 강우량으로 결정되는 토양수분 요인이다. 무, 배추 재배 농업인을 대상으로 단위당 수확량에 미치는 요인에 대한 설문 조사가 있었는데, 기상 요인 중에서는 강우량(응답농가 구성비 17.2%), 기온(16.7%), 일조(15.2%) 요인이 상위를 차지했다(한국농촌경제연구원, 1996). 고랭지 무의 파종은 준고랭지(해발고도 400m 이상 지역)에서는 4월 하순에서 5월 중순에 이루어지고, 고랭지(해발고도 600m 이상 지역)에서는 6월에서 7월에 이루어지며, 수확은 준고랭지에서는 7월에서 8월, 고랭지에서는 8월에서 9월에 걸쳐 이루어진다(한국농촌경제연구원, 1996).

한국농촌경제연구원(1996)에서는 고랭지 무의 단위당 수확량에 대한 기온과 강우량의 영향을 분석하였는데, 기상 요인과 수확량의 자연 로그 값을 선형으로 조합하여 모델을 구축하였다. 이를 분석한 결과 기온의 경우 전라북도, 충청북도와 같이 특히 고도가 낮은 지역에서 7월의 기온 상승이 수확량을 감소시키는 것으로 나타났다. 이는 7월에 장마로 인해 습도가 높을 때 기온이 상승하면 병해가 발생할 확률이 증가하기 때문으로 분석하고 있다.

강우량의 경우에는 일반적으로 5월 강우량의 증가는 단위당 수확량을 증가시키며 6월과 특히 8월의 강우량 증가는 반대로 감소시키는 것으로 나타났다. 이는 파종기인 5월에 강우량이 증가하면 무의 생육에 도움이 되기 때문이다. 반면 8월에는 장마로 인해 강우량이 지나치게 증가하여 수확량이 감소하는 것으로 보인다(한국농촌경제연구원, 1996). 고랭지 배추도 고랭지 무와 비슷한 양상을 보여 7월의 기온 상승이 수확량을 감소시키고 5월의 강우 증가가 수확량을 증가시키는 것으로 알려져 있다. 다만 8월의 강

우량 증가는 수확량과 큰 관계가 없는 것으로 나타나며, 고랭지 지역의 특성으로 강원도 지역에서는 6월과 8월의 기온 증가도 수확량 감소 요인으로 작용한다(한국농촌경제연구원, 1996).

앞 고랭지 무·배추 수확량 자료 설명에서 언급했듯 고랭지 채소 수확량 자료는 강원도, 전라북도, 경상북도, 경상남도에 대해서 2003~2011년 동안, 강원도의 경우 강릉시, 고성군, 동해시, 삼척시, 속초시 등 18개 시군구에 대해서 2007~2013년 동안 제공된다. 수확량 자료는 연간 자료이며 시도 또는 시군구에 대해서만 제공되므로 실제 수확량 예측 모델을 구축하기에는 표본의 개수가 매우 부족하다. 따라서 본 연구에서는 앞서 가뭄 감시 및 가뭄 예측 모델 구축에 사용한 목표 변수 중 고랭지 무·배추의 수확량과 밀접한 상관관계를 가지는 변수를 선정하고, random forest를 이용하여 이미 구축한 모델을 이용하여 이 변수의 값을 예측해 보았다.

가뭄 감시 및 예측 모델이 각 5×5 km 격자들의 값을 훈련 자료로 이용하여 구축된 반면, 여기에서 이 구축된 모델의 입력 자료는 시도별 또는 시군구별로 입력하였다. 각 시도 또는 시군구에 대해서 준고랭지 및 고랭지 지역을 추출하고 이 지역에 위치한 5×5 km 격자를 도출한 다음 각 격자에 대한 입력 변수 값을 각 시도 또는 시군구에 대하여 평균하여 고랭지 무·배추의 수확량과 밀접한 관련을 가지는 목표 변수를 예측하기 위한 모델의 입력 값으로 사용하였다(Figure 4-17).

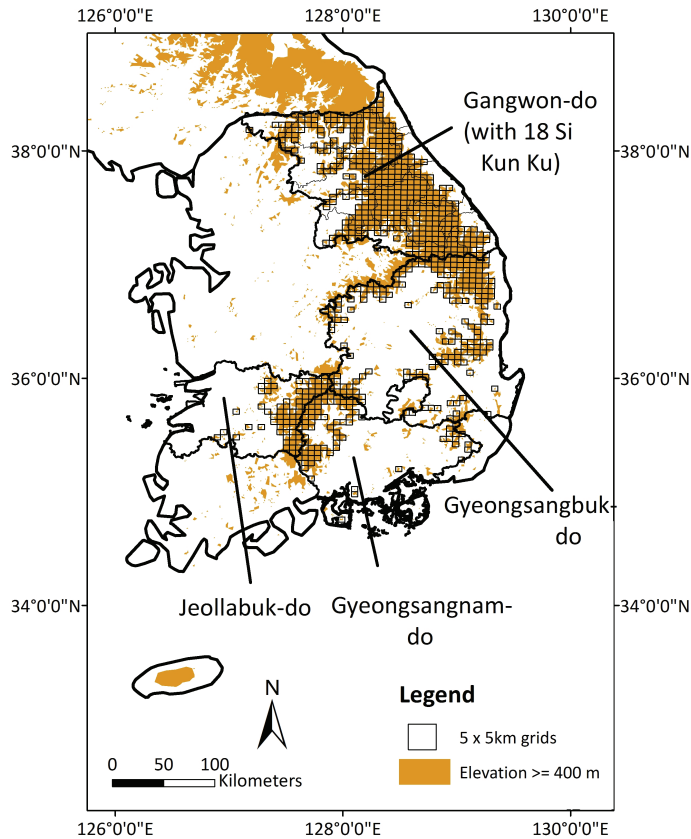


Figure 4-17. Locations of grids used: Input data for 5x5km grids located in highland areas are averaged for each SI-DO or SI-KUN-KU.

4.6.5.2 산불 발생 빈도 및 피해 면적

「산림보호법」제1장 제2조 7항에 의하면 산불은 산림이나 산림에 잇닿은 지역의 나무, 풀, 낙엽 등이 인위적으로나 자연적으로 발생한 불에 타는 것을 말한다. 산불 발생의 3요소는 열, 산소, 연료로 이러한 요소들이 갖추어 졌을 때 산불이 발생하며 산불의 형태는 바람, 연료, 지형 등의 영향을, 산불의 작용은 연료, 지형, 기상 등 환경인자의 영향을 많이 받는다(산림청, 2011). 우리나라의 경우 전체적으로 산지 면적이 넓으며 산

지 중 97%가 임목지로 풍부한 산불의 연료가 제공되며, 봄과 가을의 건조기와 풍부한 바람의 영향으로 산불의 발생에 좋은 조건을 갖추고 있다(산림청, 2011). 예를 들어, 2013년 같은 경우에는 산불이 다발하는 시기인 3월에서 4월의 기간 동안에 강수량이 감소하고 건조 일수가 증가하였으며, 봄철 저온이 지속되다가 갑작스런 기온 상승이 있어 산불의 위험이 증가하였다(산림청, 2014). 산불 통계를 보면, 산불 발생의 주원인은 논·밭두렁 및 쓰레기 소각, 입산자 실화이며 남부 지방에 집중적으로 발생한다.

산림청은 2000년대 초반에 산불위험지수를 개발하여 홈페이지를 통해 공간적으로 분포된 현재산불위험지수를 제공하고 있는데(Figure 4-18), 이 밖에 행정구역별 산불 위험등급, 상세산불위험정보 등도 제공하고 있다. 산불 발생에 영향을 미치는 것으로 고려되는 주요 변수는 기온, 실효습도, 평균풍속, 강수량으로 이들 기상 조건을 분석하여 산불위험지수를 도출해내고 있다. 국립산림과학원에서는 산불 발생 빈도와 가뭄 지수인 SPI의 상관관계에 대한 연구도 수행하였는데, 1990년대와 2000년대의 양상이 다르게 나타났다. 1990년대에는 특히 경기, 강원, 충남 지역에서는 SPI3이, 경남, 경북, 전남, 전북 지역에서는 SPI12가 산불 건수의 자연로그 값과 밀접한 관계를 나타냈으며, 2000년대에는 경기, 충남, 충북, 전남, 전북 지역에서는 SPI6이, 영서 지방에서는 SPI3, 영동, 경남, 경북 지역에서는 SPI1이 밀접한 관계를 보였다(원명수, 2014).

가뭄 지수로 널리 이용되고 있는 SPI는 특히 기상학적 가뭄 감시에 유용한 것으로 알려져 있는데, 강수량만을 입력 변수로 이용하고 있다. 본 연구에서는 위성 영상에서 도출한 다양한 변수와 지면모델의 산출물, 대규모 기후 지수를 이용하여 가뭄 감시 및 예측 모델을 구축하였는데, 이를 통한 산불 발생의 간접적 예측을 시험하였다. 다양한 변수의 포함이 산불 발생에 많은 영향을 미치는 기온, 실효습도 등 다른 기상 변수의 영향을 설명하는데 유용할 수 있으며, 특히 식생 관련 변수를 포함함으로써 산불 발생의 요소인 연료량을 고려할 수 있을 것으로 보인다.

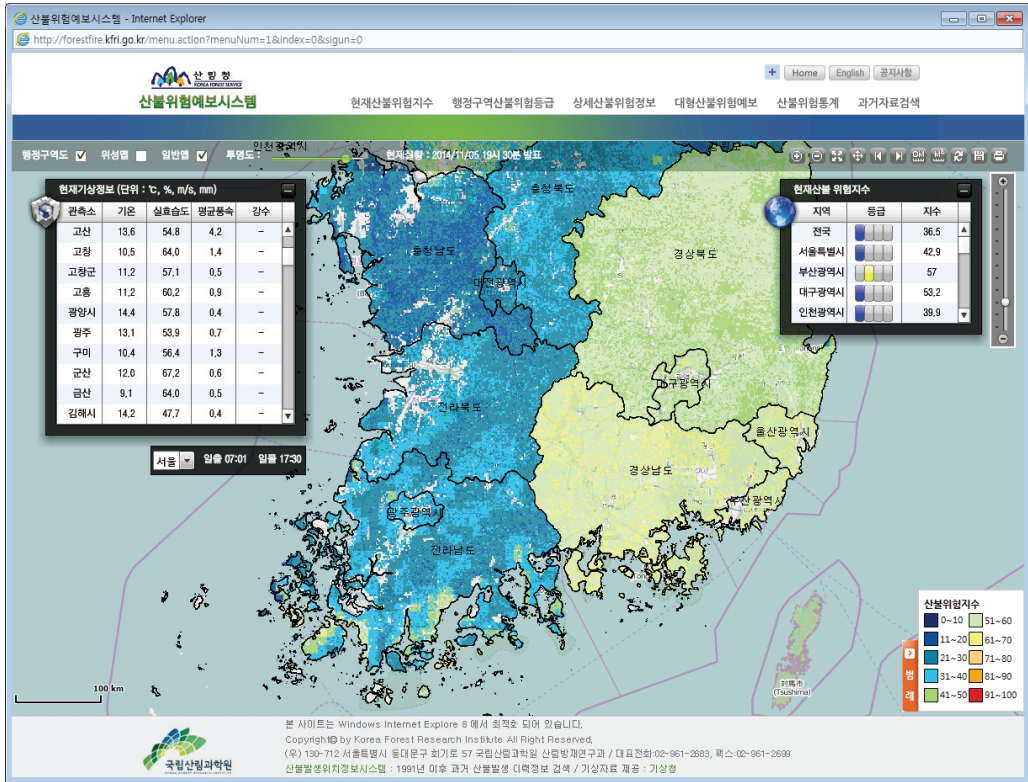


Figure 4-18. Forest Fire Risk Prediction System of Korea Forest Service (Source: <http://forestfire.kfri.go.kr/menu.action?menuNum=1&index=0&sigun=0>).

산불 발생 통계 자료가 발생한 산불의 위치와 시각, 피해 면적을 제공하므로 이를 공간적으로는 시도 및 시군구 단위로, 시간적으로는 월별로 전환하여 각 행정구역에 대한 월별 산불 발생 빈도, 산불 피해 면적, 산지 면적에 대한 산불 피해 면적 비율 자료를 도출하였다. 각 시도 또는 시군구에 대해서 토지피복도를 이용하여 산지 지역을 추출하고 이 지역에 위치한 5×5km 격자를 도출한 다음 각 격자에 대한 입력 변수 값을 각 시도 또는 시군구에 대하여 평균하여 월별 산불 발생 빈도, 산불 피해 면적, 산지 면적에 대한 산불 피해 면적 비율과 밀접한 관련을 가지는 목표 변수를 선정하였다. 그리고 이들 목표 변수를 추정하기 위한 가뭄 예측 변수의 입력 값으로 이 세 변수 값을 사용하였다(Figure 4-19).

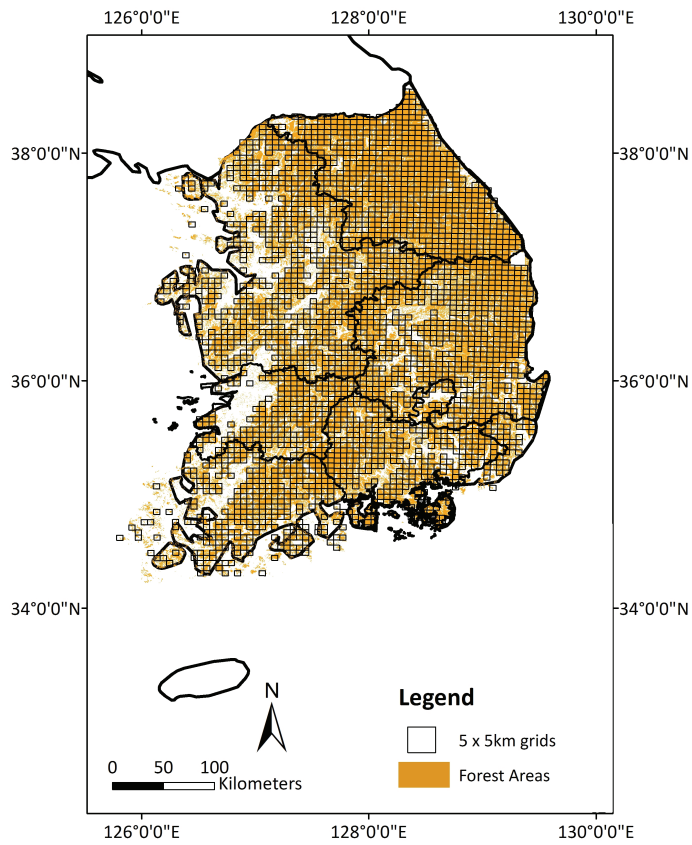


Figure 4-19. Locations of grids used: Input data for 5×5km grids located in forest areas are averaged for each SI-DO or SI-KUN-KU.

5. 결과 및 토의

5.1 위성 영상 추출 변수의 검증

위성 영상 자료에서 추출한 강수량, 지표면온도, 증발산량, 토양수분 변수 값을 관측 자료 값과 비교하여 모수 및 비모수 상관관계 통계 값과 오차 값을 살펴보았다.

5.1.1 강수량

5.1.1.1 TRMM PRCP vs. 관측 강수량

한반도의 106개 기상관측 지점 위치에 대하여 TRMM PRCP 월별 강수량 자료와 관측 강수량 자료를 mm/month 단위로 통일한 다음, 2001~2012년 기간에 대하여 비교하였다. TRMM 자료의 해상도가 $0.25 \times 0.25^\circ$ 이므로 여러 개의 기상관측 지점을 하나의 TRMM 격자와 비교하게 되는 경우도 있다. 통계 값의 비교는 전체 계절(ALL) 뿐만 아니라 봄철(March, April, May; MAM), 여름철(June, July, August; JJA), 가을철(September, October, November; SON), 겨울철(December, January, February; DJF)의 계절별로 수행되었다.

계절별로 각각의 기상관측점에 대해 수행된 분석 결과를 평균한 값을 Table 5-1에 나타내었다. TRMM의 해상도가 매우 낮음에도 불구하고, 피어슨 상관계수와 켄달 순위 상관계수의 경우 통계적인 의미(statistical significance)를 나타내는 p 값은 모두 매우 작아(p-value < 0.001; 자료 미포함) 위성 영상에서 도출한 강수량과 관측 강수량 간에 통계적으로 유의미한 관계가 있음을 보이고 있다. 위성 영상과 인공 신경망을 사용하여 강수량을 예측한 Hsu et al.(1996)의 연구에서 미국 플로리다 지역에서 1월 RMSE가 65.48mm, 일본 지역에서 6월에 55.36mm, 7~8월에 63.41mm로 분석되는 등의 결과와 비교하면, 여름철을 제외하고는 오차 폭이 그다지 크지 않으며, 상관관계도 전체적으로 상당히 높다.

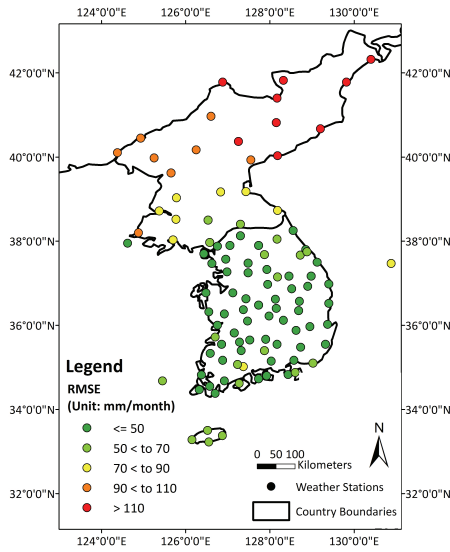
각 기상관측지점의 결과를 개별적으로 살펴보면, 높은 RMSE와 MAE 값은 대부분

한반도 북쪽 지역에 위치한 기상관측지점의 높은 값 때문인 것으로 보인다(Figure 5-1). 특히 이 지점들의 여름철 RMSE와 MAE는 남쪽 지역에 위치한 지점들에 비해 상당히 높다(Figure 5-1c). 봄철이나 가을철에도 상대적으로 큰 오차 값을 보인다(Figures 5-1b, 5-1d). 이는 당연한 결과로, TRMM은 원래 열대 지방에 중점을 두어 30° S에서 30° N 위도의 지역에 적용하도록 설계되어 위성이 적도 부근을 회전하고 있기 때문이다. 35° 부근 까지만 정확한 값을 얻을 수 있다. TRMM 3B43 자료 같은 경우 여러 전지구적 관측 자료와 결합하여 50° S에서 50° N 위도까지 값은 제공이 되나, 정확도는 떨어질 수밖에 없다.

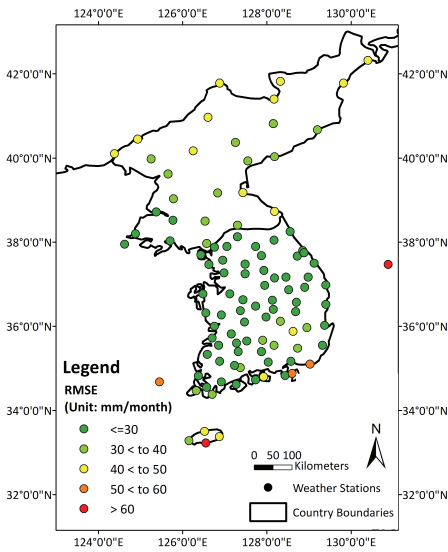
남쪽 지역에 위치한 기상관측지점의 RMSE, MAE 값은 대부분 북쪽의 지점들에 비해 낮는데, 간혹 몇몇 지점이 비이상적으로 매우 큰 값을 보인다. 예를 들어, 서귀포(189)의 봄철 RMSE는 75.32 mm/month로, 전체 평균인 29.69 mm/month 에 비해 상당히 높으며, 울릉도(115)의 겨울철 RMSE는 94.48 mm/month로 전체 평균인 18.23 mm/month에 비해 매우 높다. 이 지점들은 모두 섬 또는 남부 연안에 위치하고 있다(Figure 5-1).

Table 5-1. Statistics between TRMM PRCP and observed monthly precipitation averaged over 106 weather stations.

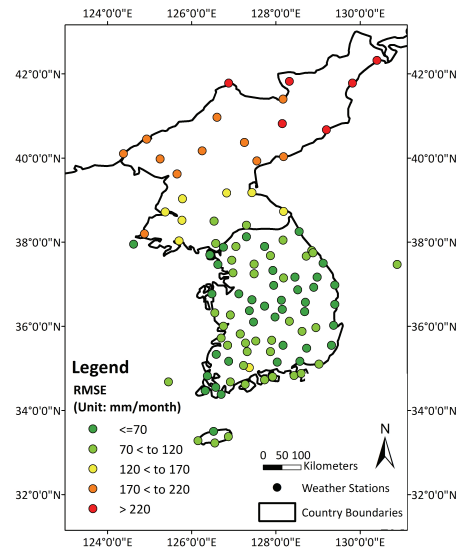
SEASON	RMSE (mm/month)	MAE (mm/month)	Obs. range (min~max)	r	tau
ALL	57.27	33.79	0~1244.5	0.92	0.8
MAM	29.69	22.14	1.8~516.5	0.84	0.7
JJA	98.43	72.92	0.1~1244.5	0.85	0.71
SON	39.38	25.53	0~880	0.89	0.73
DJF	18.23	13.6	0~321.1	0.84	0.67



(a)



(b)



(c)

Figure 5-1. RMSE between TRMM PRCP and observed monthly precipitation during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: mm/month).

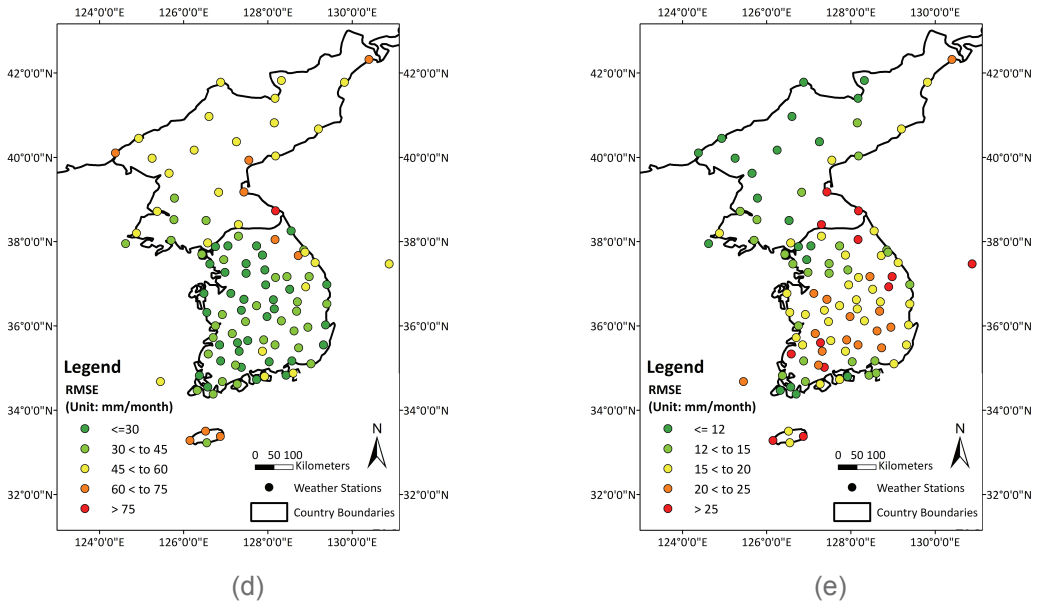
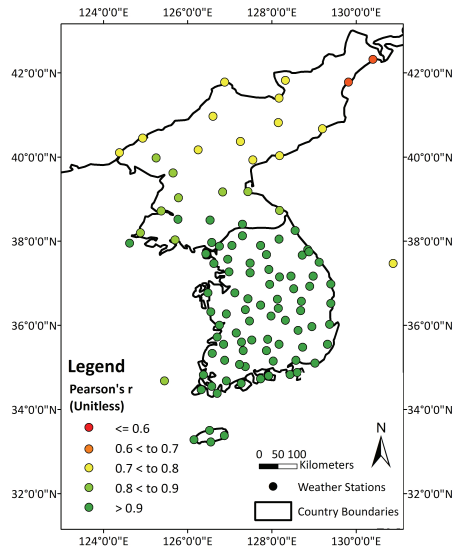


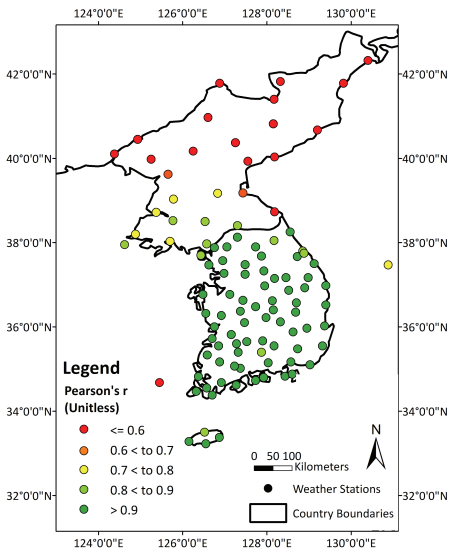
Figure 5-1. (Continued). RMSE between TRMM PRCP and observed monthly precipitation during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: mm/month).

상관계수 값은 봄철과 겨울철에 가장 낮게 나타나는데(Table 5-1), 전체적으로 북쪽 지역에 위치한 지점들의 상관계수가 낮으나 봄철의 경우 특히 낮게 나타났으며, 겨울철의 경우에는 북쪽 지역 중에서도 동해안을 따라 함경북도, 함경남도 및 황해남도에 위치한 지점에서 특히 낮게 나타났다(Figures 5-2b, 5-2e). 겨울철의 경우 남쪽 지역에서 고창(172)에서 매우 상관계수가 낮게 나타났으며($r = 0.10$), 순천(174)의 경우 유일하게 음의 상관계수($r = -0.56$)를 나타내었다(Figure 5-2e). 이 두 지점에서는 결측 값이 많아 표본 집합의 수가 매우 적기 때문으로 보인다.

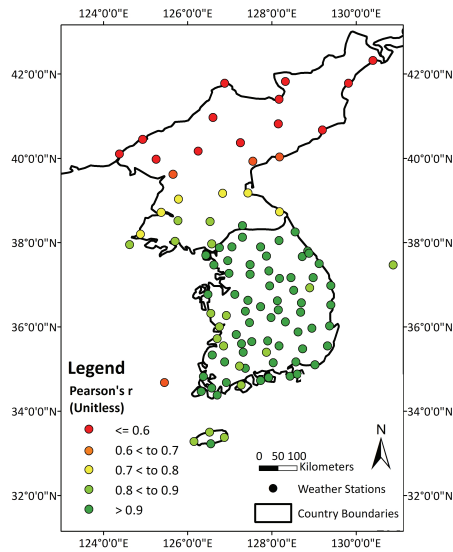
전체 계절에 대해 비교적 높은 상관계수($r = 0.97$)와 낮은 오차(RMSE = 26.06) 값을 보인 대전(ID: 133) 지점에 대해 2001~2012년 기간에 대해 시계열을 살펴보았다(Figure 5-3). 시기별로 TRMM PRCP가 관측 값을 과대 추정하는 부분이 종종 나타나기는 하나 비교적 가까운 값을 보이고 있다.



(a)



(b)



(c)

Figure 5-2. Pearson's r between TRMM PRCP and observed monthly precipitation during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).

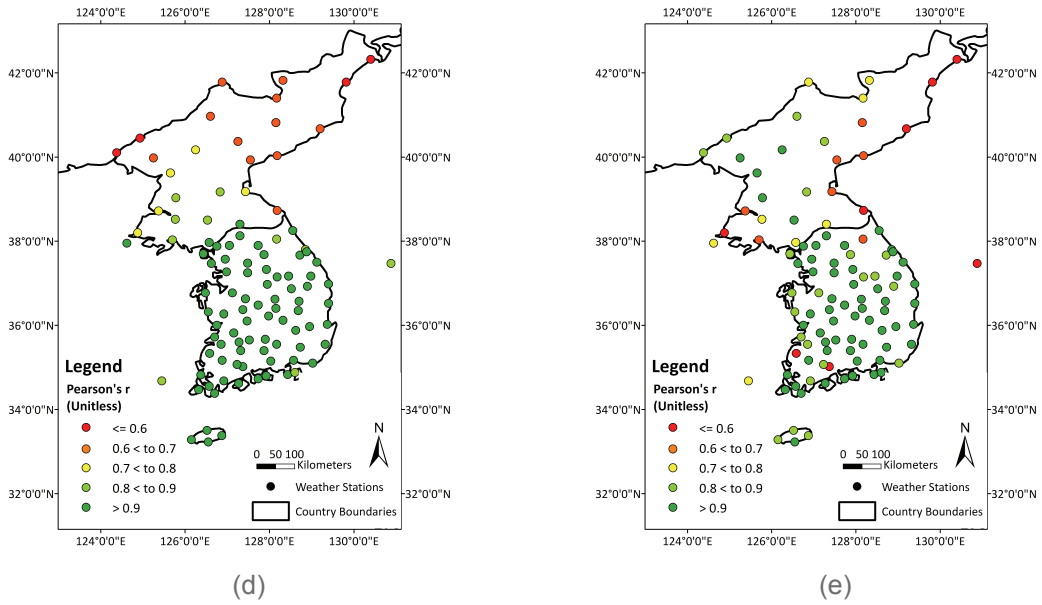


Figure 5-2. (Continued). Pearson's r between TRMM PRCP and observed monthly precipitation during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: mm/month).

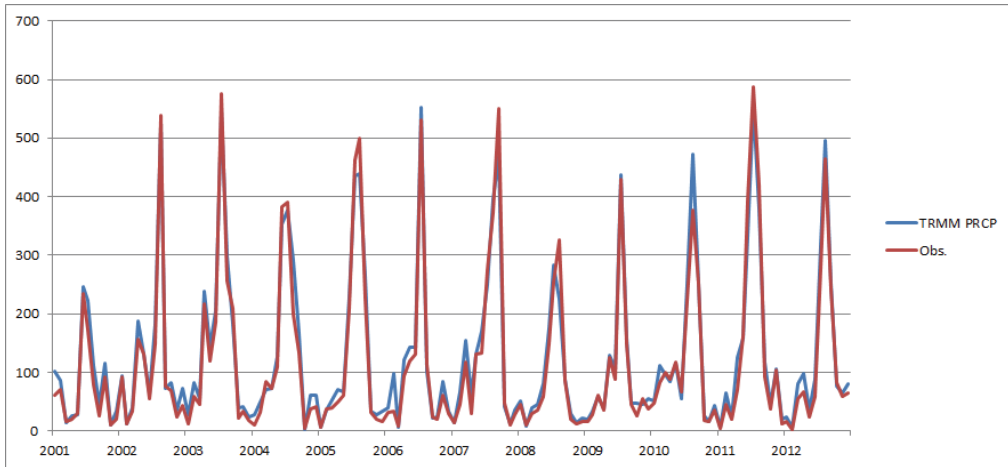


Figure 5-3. Time-series of TRMM PRCP and observed monthly precipitation from during 2001~2012 in Daejeon station (ID: 133) (Unit: mm/month).

5.1.2 지표면온도

5.1.2.1 MODIS LST vs. 관측 대기온도

한반도의 106개 기상관측 지점 위치에 대해 2001~2012년 기간에 대해 °C 로 통일하고 월별로 전환한 MODIS LST 자료와 월별 대기온도 자료를 비교하여 피어슨 상관계수 값과 켄달의 순위 상관계수 값, RMSE 및 MAE 값을 구하였다. MODIS 주간 LST(LSTDAY)와 월별 최고 온도를 비교하였으며, MODIS 야간 LST(LSTNIGHT)와 월별 최저 온도, MODIS 평균 LST(LST_MEAN)과 월별 평균 온도를 비교하였다. MODIS 자료가 주간 혹은 야간의 자료만 존재하는 경우도 있으므로, 이렇게 한쪽만 있는 경우는 그 기간의 평균 지표면온도 자료가 존재하지 않게 된다. 따라서 LSTDAY와 LSTNIGHT의 평균이 LST_MEAN과 같지 않을 수 있다. 모든 분석은 전체 계절 및 각 계절에 대해 수행하였다.

MODIS 지표면온도의 일반적인 정확도는 대략 1K 보다 나은 것으로 알려져 있는데 (<http://landval.gsfc.nasa.gov/ProductStatus.php?ProductID=MOD11>), 여기에서는 대기온도와 비교하였으므로 RMSE와 MAE 오차 값이 이보다는 크게 나타난다(Table 5-2). 전체적으로는 주간 온도에서의 오차가 크게 나타나는데, 계절별로 살펴보면 봄철과 여름철에서는 주간 온도의 오차 값이 더 크고, 가을철과 겨울철에는 야간 온도의 오차 값이 더 크다. 평균 온도를 살펴보면 주간 온도나 야간 온도만을 비교하였을 때보다 가을철과 겨울철에는 오차 값이 더욱 작아졌다(Table 5-2). 주간 온도는 위성 영상에서 도출한 자료와 관측 자료 간에 오차가 가을철과 겨울철에 가장 작은 반면, 야간 온도는 봄철과 여름철에 가장 작다(Table 5-2).

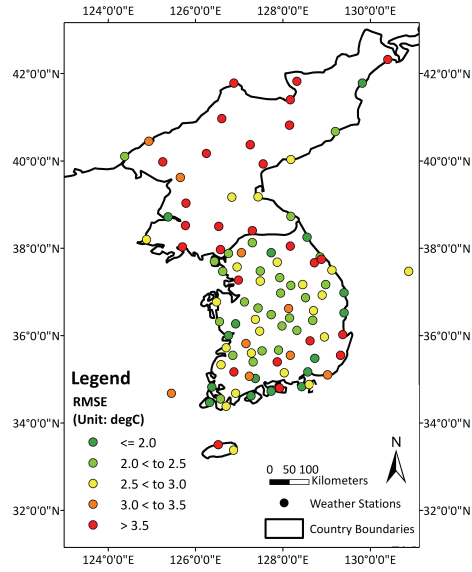
상관관계를 보면, 위성 영상에서 도출한 자료와 관측 자료 사이에 여름철 주간 온도의 상관관계는 매우 낮게 나타난다($r = 0.19$). 그 밖에는 겨울철을 제외하고 상관관계가 매우 높으며, 겨울철에도 비교적 높은 상관관계를 보인다(Table 5-2).

MODIS LSTDAY와 관측 최고 온도를 비교한 지점별 오차 결과를 살펴보면, 여름철

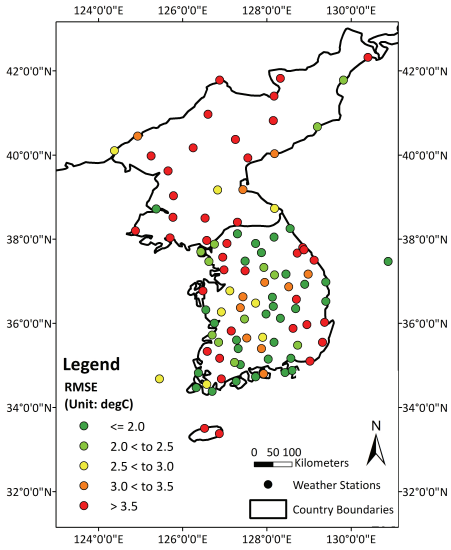
에는 한반도 전반적으로 오차가 크게 나타나지만(Figure 5-4c), 봄철과 가을철, 겨울철의 경우에는 북쪽 지역에 위치한 지점들의 오차가 더 높게 나타나는 편이다(Figures 5-4b, 5-4d, 5-4e). 특히 삼지연(47005) 지점의 경우 RMSE 값이 봄철에는 9.48°C, 가을철에는 10.45°C, 겨울철에는 14.87°C로 나타나는 등 오차 값이 매우 컸는데, 이는 이 지점의 특히 높은 해발고도(1386m) 때문인 것으로 보인다. 위성 영상에서 도출한 지표면온도와 관측한 대기온도를 비교하였으므로, 고도의 영향을 받을 수밖에 없다.

Table 5-2. Statistics between MODIS LST and observed monthly temperature averaged over 106 weather stations.

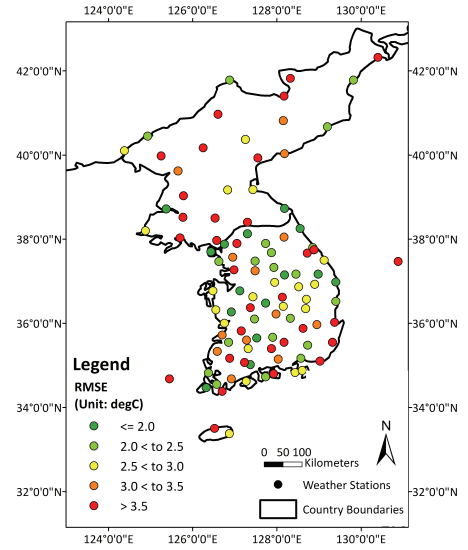
VARIABLES	SEASON	RMSE (°C)	MAE (°C)	Obs. range (min~max)	r	tau
LSTDAY vs. Monthly maximum temperature	ALL	3.09	2.54	-5.9~33.6	0.98	0.84
	MAM	3.27	2.89	2.8~28	0.96	0.83
	JJA	3.23	2.76	18.7~33.6	0.19	0.13
	SON	2.22	1.94	4.8~29.5	0.98	0.84
	DJF	2.97	2.62	-5.9~14.6	0.84	0.67
LSTNIGHT vs. Monthly minimum temperature	ALL	2.62	2.15	-19.1~26.4	0.98	0.89
	MAM	2.01	1.75	-7.4~17.0	0.97	0.81
	JJA	1.69	1.42	9.8~26.4	0.85	0.65
	SON	2.52	2.25	-5.9~23.2	0.98	0.85
	DJF	3.62	3.2	-19.1~8.0	0.68	0.51
LST_MEAN vs. Monthly mean temperature	ALL	2.23	1.83	-12.25~29.3	0.99	0.9
	MAM	2.21	1.96	-1.9~21.4	0.98	0.84
	JJA	2.13	1.82	14.2~29.3	0.63	0.47
	SON	1.76	1.53	-0.4~26.0	0.99	0.87
	DJF	2.34	2.03	-12.25~11.1	0.83	0.65



(a)



(b)



(c)

Figure 5-4. RMSE between LSTDAY and observed monthly maximum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: degree C).

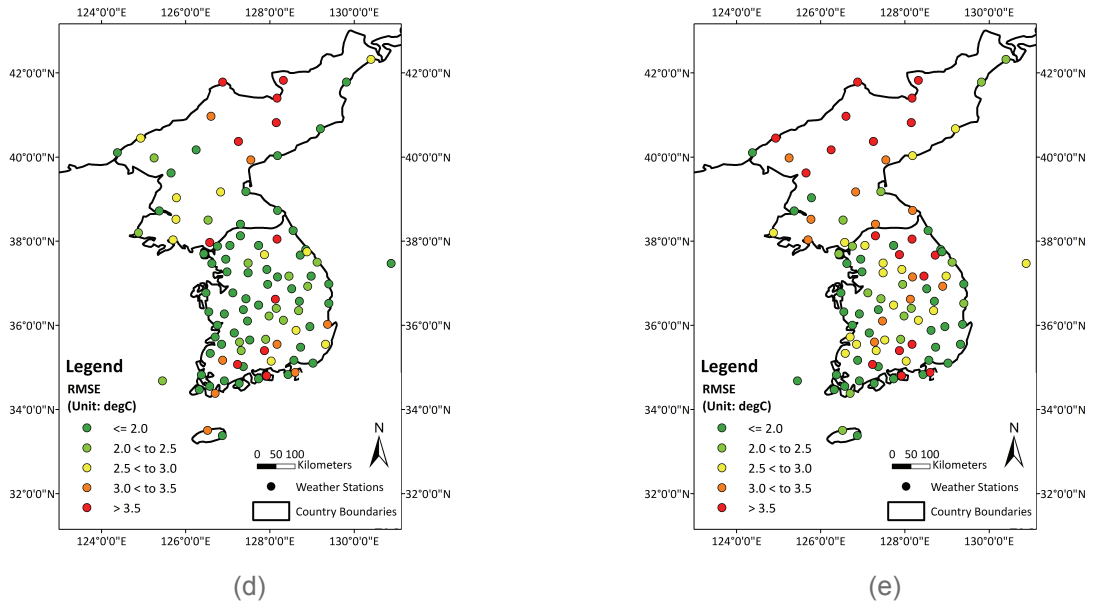
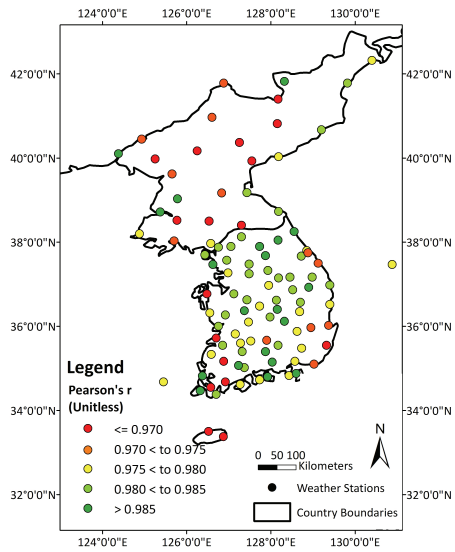


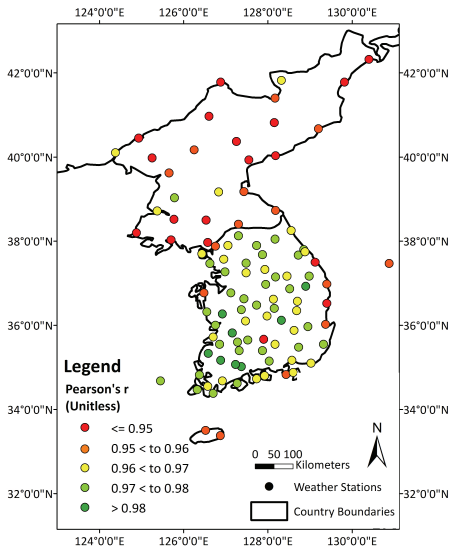
Figure 5-4. (Continued). RMSE between LSTDAY and observed monthly maximum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: degree C).

상관관계를 보면 봄철과 가을철은 전체 지점에 대해 전반적으로 상관계수가 높게 나타났으며(Figures 5-5b, 5-5d), 겨울철에는 북쪽 지역에 위치한 지점에서 상관계수가 약간 낮게 나타났다(Figure 5-5e). 여름철에는 전체적으로 위성 영상에서 도출한 값과 관측 값의 상관관계가 낮게 나타났는데, 북쪽 및 남쪽 지역에 위치한 많은 지점에서 음의 상관관계를 보였으며, 그 밖의 지점에도 상관계수 값이 매우 낮았다(Figure 5-5c). 앞서 설명한 바와 마찬가지로 지표면온도와 대기온도를 비교했기 때문에 이러한 결과가 나타났을 수 있으며, 북쪽에 위치한 지점들 중 해발고도가 높은 지점들 때문 일 수 있다. 그리고 북한 자료의 품질에서 야기된 문제일 가능성도 배제할 수는 없다. 북한의 기온 및 강수량 자료는 세계기상통신망(GTS)을 통해 수집된 값이다.

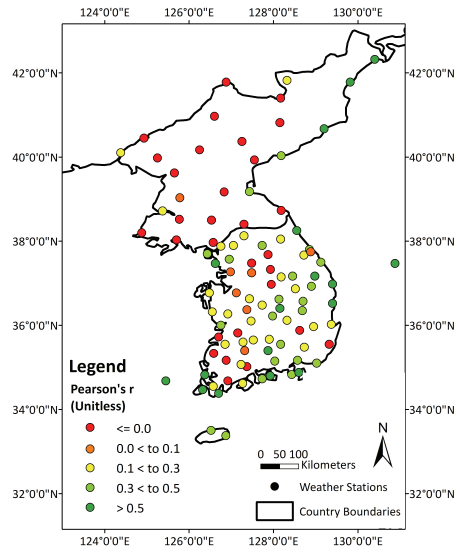


(a)

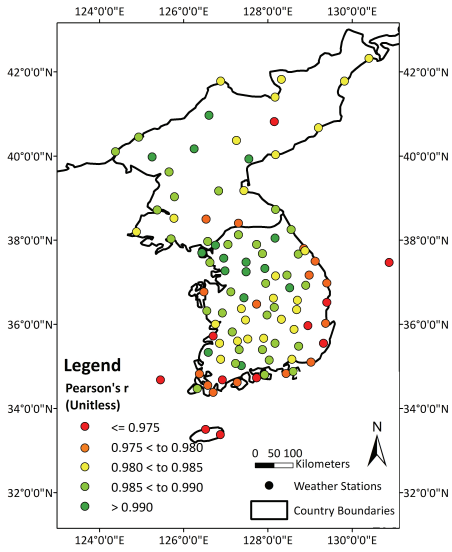
Figure 5-5. Pearson's r between MODIS LSTDAY and observed monthly maximum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).



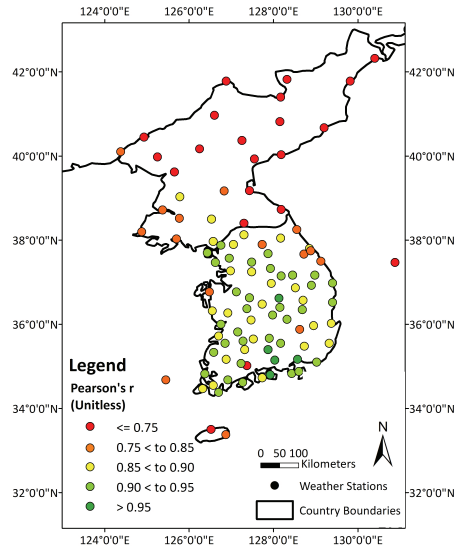
(b)



(c)



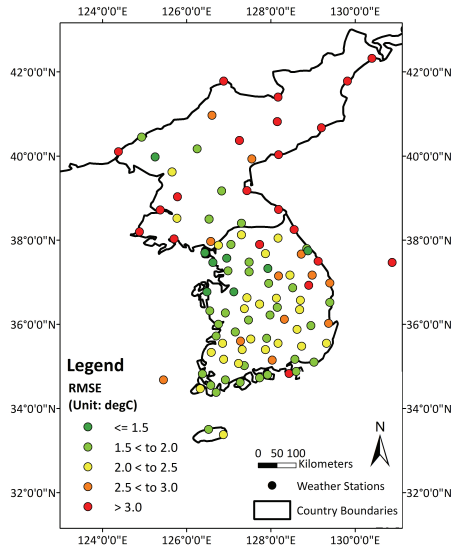
(d)



(e)

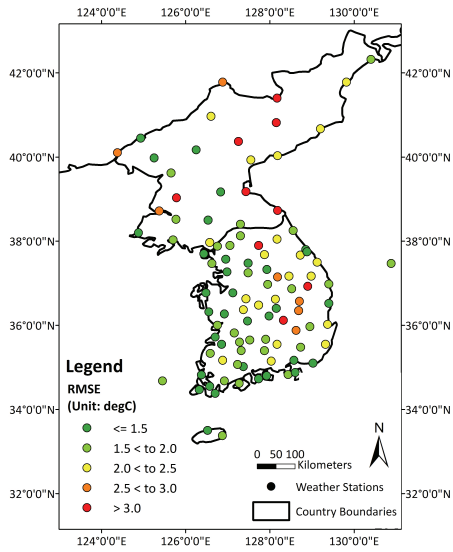
Figure 5-5. (Continued). Pearson's r between MODIS LSTDAY and observed monthly maximum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).

MODIS LSTNIGHT과 관측된 월별 최소 대기온도에 대해서도 비교 분석을 수행하였다. 모든 계절에서 북쪽 지역에 위치한 지점들에서 RMSE, MAE의 값이 크게 나타났다. 특히 고도가 매우 높은 삼지연(47005) 지점의 경우 오차가 봄철에 RMSE 값이 9.12°C, 여름철에 7.25°C, 가을철에 8.94°C, 겨울철에 11.29°C로 가장 크게 나타났다(Figure 5-6).

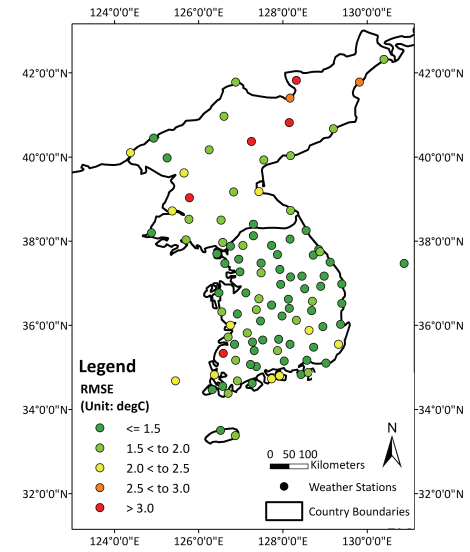


(a)

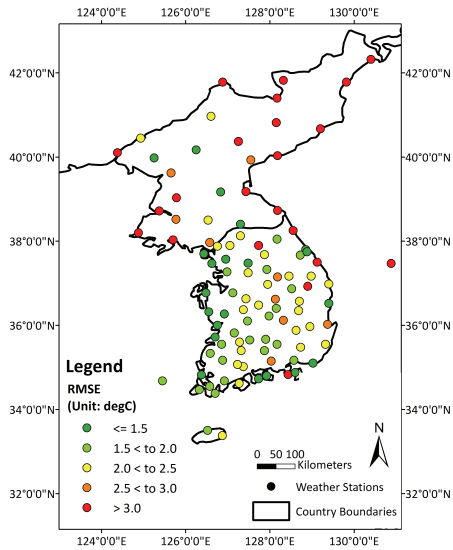
Figure 5-6. RMSE between LSTNIGHT and observed monthly minimum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: degree C).



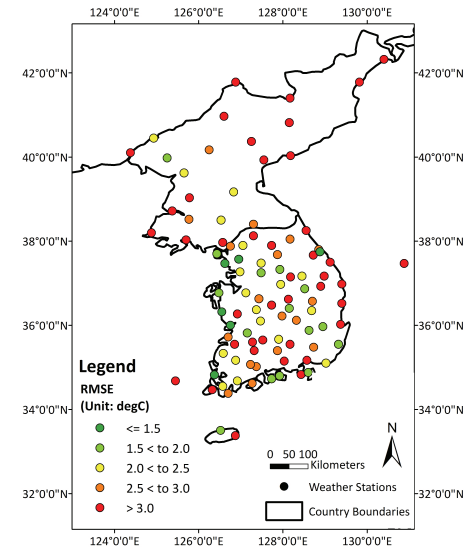
(b)



(c)



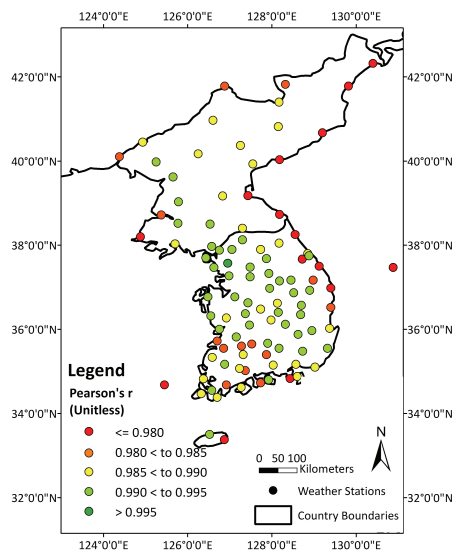
(d)



(e)

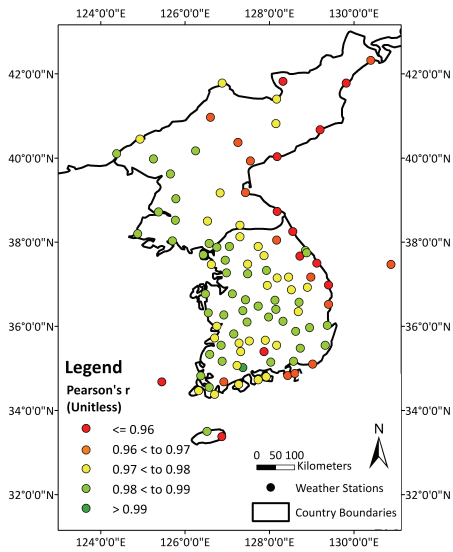
Figure 5-6. (Continued). RMSE between LSTNIGHT and observed monthly minimum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: degree C).

MODIS LSTNIGHT과 관측 월별 최소 대기온도와의 상관관계는 봄철과 가을철에는 매우 높은 반면 여름철과 겨울철, 특히 겨울철에 상대적으로 낮다(Table 5-2). 지점별로 살펴보면, 여름철에는 북쪽 지역과 남쪽 내륙에 위치한 지점들이 상대적으로 낮은 상관관계를 가지는 것을 볼 수 있는데(Figure 5-7c), 특별히 고도의 영향은 아닌 것으로 보인다. 겨울철에는 특히 동쪽 해안을 따라 위치한 지점들에 있어 상관관계가 낮게 나타났다(Figure 5-7e).

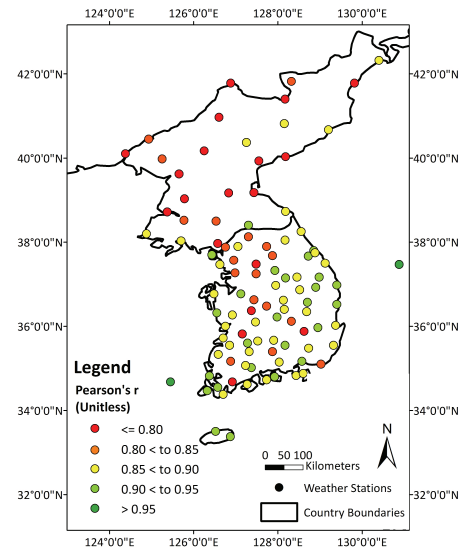


(a)

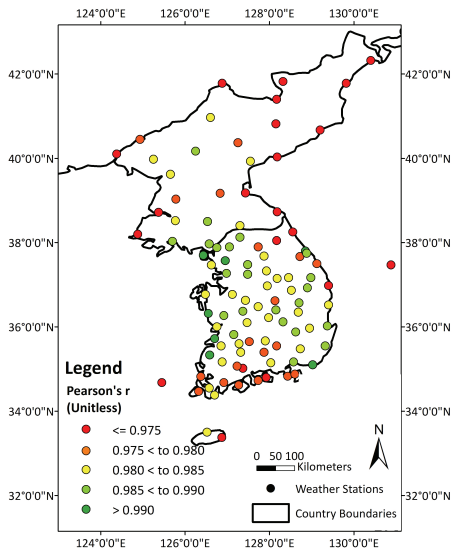
Figure 5-7. Pearson's r between MODIS LSTNIGHT and observed monthly minimum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).



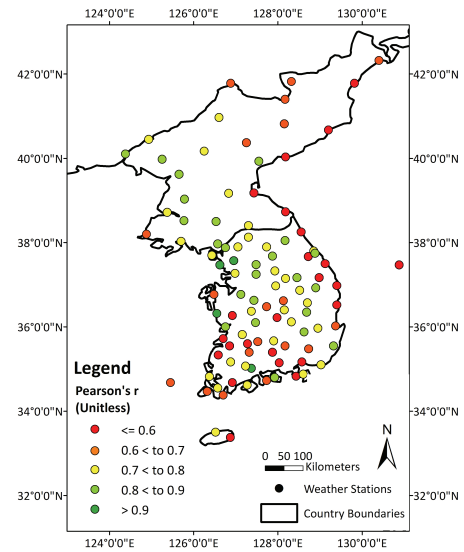
(b)



(c)



(d)



(e)

Figure 5-7. (Continued). Pearson's r between MODIS LSTNIGHT and observed monthly minimum air temperature during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).

주간, 야간, 평균 지표면온도에 대해서 각각 비교적 높은 상관계수와 낮은 오차 값을 보인 남포 관측지점(ID: 47060, $r = 0.99$, $RMSE = 1.49$), 서울 관측지점(ID: 108, $r = 0.99$, $RMSE = 1.23$), 상주 관측지점(ID: 137, $r = 0.99$, $RMSE = 1.05$)에 대해 2001~2012년 기간 동안의 시계열을 살펴보았다(Figures 5-8~10). 여기서 관찰한 세 지점의 경우 관측 값과의 차이가 여름철에 조금 더 두드러져 보이는 것을 볼 수 있다.

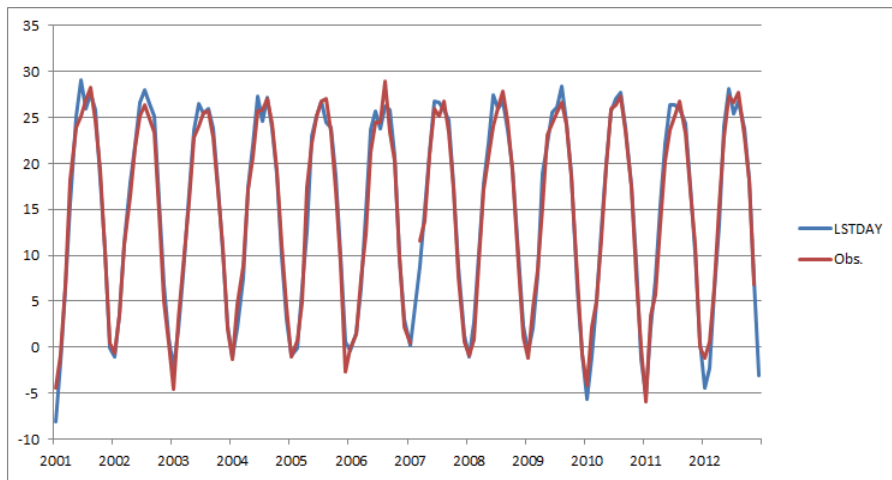


Figure 5-8. Time-series of MODIS LSTDAY and observed monthly maximum temperature in Nampo station (ID: 47060) (Unit: °C).

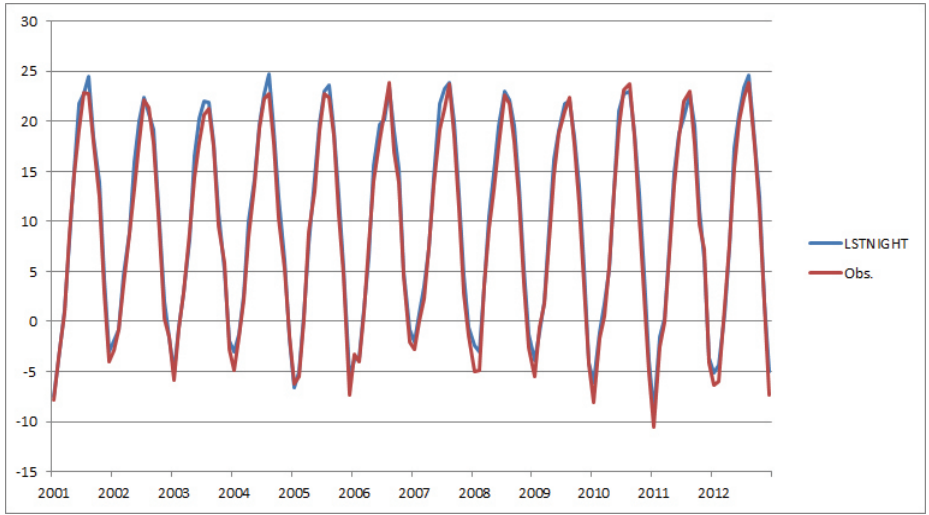


Figure 5-9. Time-series of MODIS LSTNIGHT and observed monthly minimum temperature in Seoul station (ID: 108) (Unit: °C).

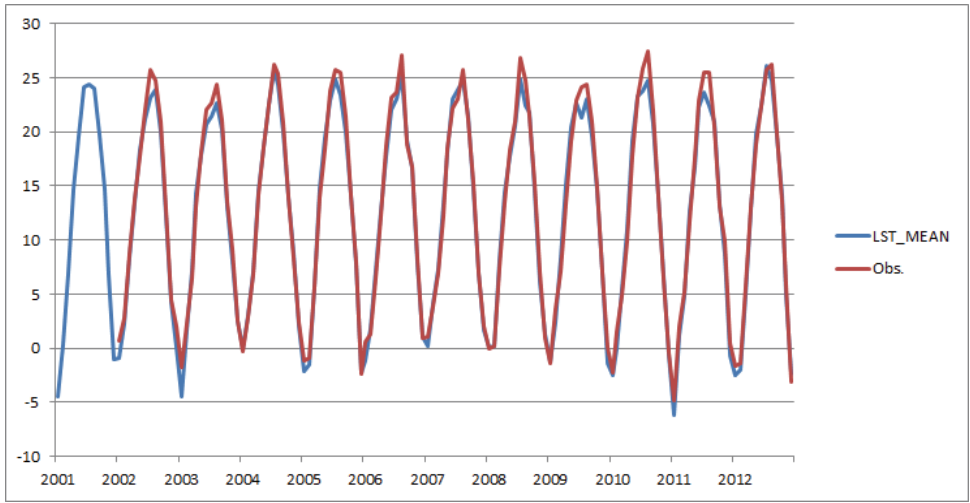


Figure 5-10. Time-series of MODIS LST_MEAN and observed monthly mean temperature in Sangju station (ID: 137) (Unit: °C).

5.1.3 증발산량

5.1.3.1 MODIS PET vs. Penman-Monteith PET

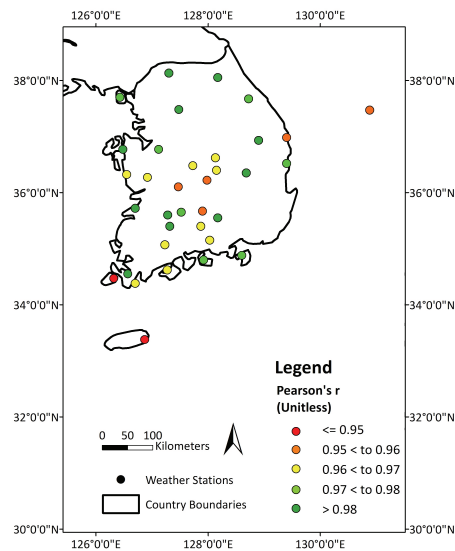
Penman-Monteith 식을 사용하기 위해 필요한 여러 변수 값을 갖추었으며, MOD16 PET 값이 존재하는 남한 지역의 35개 ASOS 기상관측 지점에 대해서 이 지점들의 관측 기상변수 값을 입력 값으로 사용하여 Penman-Monteith 식으로 계산한 잠재 증발산량을 MOD16 PET 값과 2002~2010년의 기간에 대해 비교하였다(Table 5-3). 단위는 mm/month 로 통일하였다.

RMSE와 MAE 로 나타낸 오차 값을 보면 특히 여름철에 비교적 큰 값을 나타낸다. 특히 APEC기후센터(2013a)에서 위성 영상에서 도출한 최고 및 최저 대기온도를 이용하여 Hargreaves 식으로 계산한 PET와 비교하였을 때, 오차 값이 상대적으로 큰 것을 알 수 있다. 그러나 상관관계를 나타내는 r 값과 τ 값을 보았을 때 여름철과 겨울철을 제외하고는 평균적으로 상관관계가 매우 높다. 여름철 상관관계가 상대적으로 낮기는 하지만 겨울철 값도 높은 편이라고 할 수 있다(Table 5-3). 가뭄 감시 및 예측 모델을 구축할 때 모든 기상관측 지점 값을 통합한 모델을 구축하기 위해 각 지점별로 최댓값과 최솟값을 이용하여 표준화하므로, 상대적인 가뭄의 상태를 평가하기는 여전히 유용한 정보를 얻을 수 있다.

지점별 오차 값도 계절별로 살펴보았는데, 모든 계절에 대해 뚜렷한 지역별 특성을 찾아볼 수 없었다(자료 미포함). 지점별 상관관계 값을 살펴보면, 일부 지점에서 상관관계가 특히 낮게 나왔는데 제주도 동부에 위치한 성산(188) 지점의 전체 계절에 대한 r 값이 0.64로 나타났다. 이 지점에서는 봄철의 피어슨 상관계수가 0.33으로 나타났는데, 유일하게 1% 유의수준에서 통계적으로 유의미하지 못한 것으로 드러났다(p -value = 0.09). 하지만 이 경우에도 켄달의 τ 값은 0.60이고 이 때 p 값은 0.001보다 작아 여전히 통계적으로 유의한 상관관계를 나타낸다(Figure 5-11).

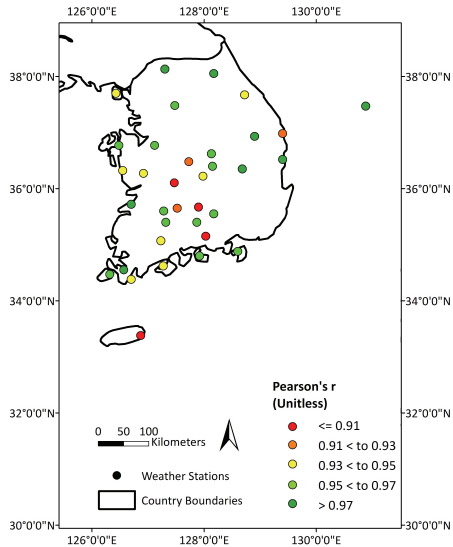
Table 5-3. Statistics between MODIS PET and calculated PET based on Penman-Monteith averaged over 35 weather stations.

SEASON	RMSE (mm/month)	MAE (mm/month)	Obs. range (min~max)	r	tau
ALL	47.42	42.07	0~172.61	0.96	0.82
MAM	46.76	44.99	0~151.21	0.93	0.81
JJA	67.87	66.23	0~172.61	0.56	0.4
SON	40.87	37.77	0~122.94	0.93	0.75
DJF	20.54	19.3	0~75.13	0.74	0.47

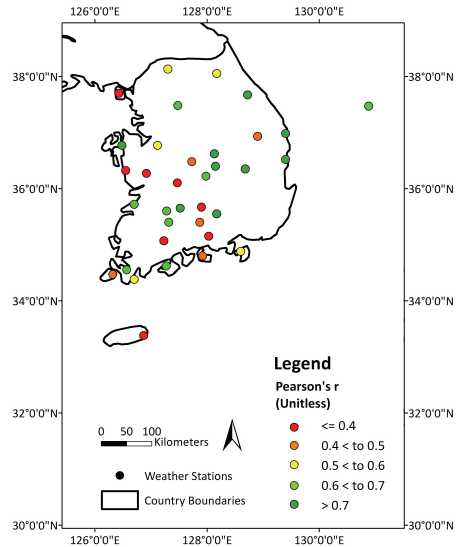


(a)

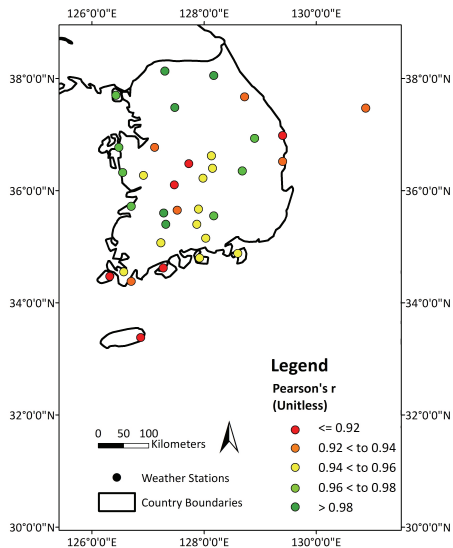
Figure 5-11. Pearson's r between MODIS PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).



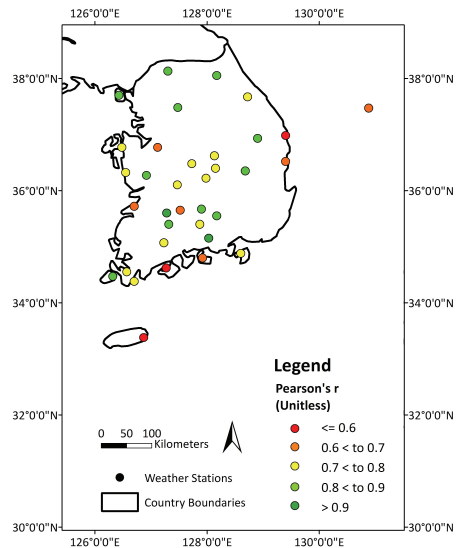
(b)



(c)



(d)



(e)

Figure 5-11. (Continued). Pearson's r between MODIS PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).

여름철에는 35개 ASOS 지점 중에서 12개의 지점에서 1% 유의수준에서 통계적으로 유의한 상관관계를 가지지 못하고 6개의 지점이 5% 유의수준에서 통계적으로 유의한 상관관계를 가지지 못하는 것으로 나타났다(자료 미삽입). 이 중 특히 진주(192), 부여(236) 등의 지점에서 상관관계가 0.2 보다도 작은 값으로 나타났다(각각 $r = 0.16, 0.19$). 비모수 통계 값에서도 유사하게 11개 지점에서 1% 유의수준에서 통계적으로 유의한 상관관계를 가지지 못하고 6개의 지점이 5% 유의수준에서 통계적으로 유의한 상관관계를 가지지 못하는 것으로 나타났다(자료 미삽입). 진주와 부여 뿐 아니라 보령(235), 금산(238) 에서도 매우 작은 켄달의 tau 값을 얻었다. 이들은 주로 남한의 남서부 내륙에 위치해 있다(Figure 5-11c).

가을철에는 35개 모든 지점이 1% 유의수준에서 통계적으로 유의미한 상관관계를 나타냈는데(Figure 5-11d), 겨울철에는 성산(188) 지점이 1% 유의수준에서 상관관계가 통계적으로 유의하지 못했고($r = 0.47, p\text{-value} = 0.01$), 울진(130) 지점이 5% 유의수준에서 상관관계가 통계적으로 유의하지 못했다($r = 0.19, p\text{-value} = 0.35$). 비모수 통계에서는 이 두 지점 이외에 남해(295), 영덕(277) 지점의 총 네 지점이 1% 유의수준에서 상관관계가 통계적으로 유의하지 못했다(울진, 성산, 남해, 영덕에 대해 각각 $\tau = 0.27, 0.29, 0.31, 0.33$). 이 네 지점들은 모두 동해안, 남해안, 제주도의 해안가에 위치하고 있다(Figure 5-11e).

인제 지점(ID: 211, $r = 0.99, RMSE = 4.35 \text{ mm/month}$)에 대하여 2002~2010년의 기간의 시계열을 나타내었다(Figure 5-12). MODIS PET와 Penman-Monteith 식을 이용해 계산한 PET 사이의 상관관계는 매우 높았으나, 시계열 그래프에서 볼 수 있듯 두 값 사이의 오차 값은 매우 크게 나타났으며, 35개 ASOS 지점 중 대부분의 경우 인제 지점보다 오차 값이 더 크게 나타났다. 여름철에는 MODIS PET가 잠재 증발산량을 크게 과대 추정하고 있으며, 겨울철에는 과소 추정하고 있다(Figure 5-12). MODIS의 잠재 증발산량 자료를 사용하려면 연구 지역에 대해 보정이 필요할 것으로 보인다.

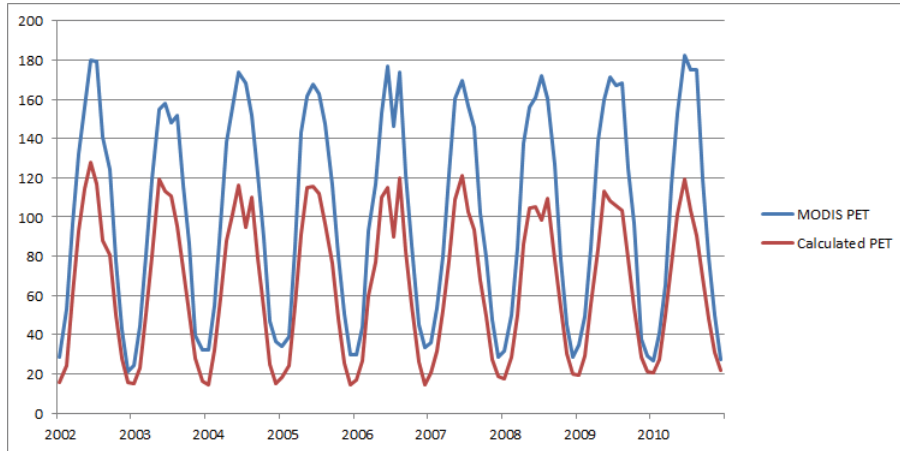


Figure 5-12. Time-series of MODIS PET and calculated PET in Inje station (ID: 211) (Unit: mm/month).

5.1.3.2 NOAH-LSM PET vs. Penman-Monteith PET

지면모델인 NOAH-LSM의 산출물 중 잠재 증발산량과, 앞서 설명한 관측 기상변수 값을 입력 값으로 사용하여 Penman-Monteith 식으로 계산한 잠재 증발산량을 mm/month로 전환하여 남한 지역의 72개 ASOS 기상관측 지점에 대해서 2002~2010년의 기간에 대해 비교하였다. 오차 값은 여름철을 제외하고는 더 크게 나타났으며, 상관계수도 여름철을 제외하고는 낮게 나타났다(Table 5-4). 이는 MOD16 자료의 공간 해상도(1km)가 지면모델의 공간 해상도(5km) 보다 높기 때문인 것으로 보인다. 여름철에는 지면모델을 사용한 경우가 오차가 조금 낮았으며 상관계수도 약간 높게 나타났다.

지점별 오차 분포를 보았을 때, MOD16을 계산한 PET와 비교한 경우와는 다르게 지면모델 산출물을 비교한 경우에는 공간적인 분포가 비교적 뚜렷이 드러났다(Figure 5-13). 봄철을 제외하고는 남부 연안을 따라 위치한 지점의 오차 값이 높게 나타났으며(Figures 5-13c, 5-13d, 5-13e), 봄철에는 남한 북부에 위치한 지점들의 오차 값

이 특히 높았다(Figure 5-13b).

Table 5-4. Statistics between NOAH-LSM PET and calculated PET based on Penman-Monteith averaged over 72 weather stations.

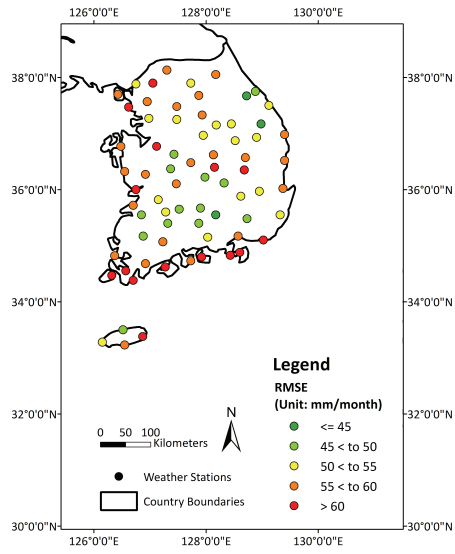
SEASON	RMSE (mm/month)	MAE (mm/month)	Obs. range (min~max)	r	tau
ALL	55.63	50.6	0~172.61	0.94	0.79
MAM	68.96	66.18	0~151.21	0.88	0.7
JJA	64.85	61.6	0~172.61	0.76	0.59
SON	48.44	46.27	0~122.94	0.89	0.69
DJF	30.4	28.33	0~75.13	0.63	0.42

지점별 상관계수를 살펴보면, 전체 계절에 대해서는 모두 1% 유의수준에서 통계적으로 유의미한 상관관계를 나타냈으나, 제주도 동부에 위치한 성산(188)에서 만이 0.90보다 작은 피어슨 상관계수 값을 보였다($r = 0.65$; Figure 5-14a). 봄철에는 성산(188) 지점에서만 1% 수준에서 통계적으로 유의한 상관관계를 보이지 못하였으나, 5% 유의수준에서는 유의미한 관계를 나타냈다($p\text{-value} = 0.02$; Figure 5-14b). 이 때, 피어슨 상관계수 값은 0.45이며, 켄달 상관계수 값을 보면 0.50으로 비교적 높고 1% 유의수준에서도 통계적으로 유의미한 것으로 나타났다. 성산(188) 지점을 제외하더라도 제주에 위치한 다른 지점들의 상관계수 값이 다른 지점들에 비해 상대적으로 낮았다($r < 0.80$; Figure 5-14b). 가을철도 봄철과 유사하게 나타났으나 다만 성산(188) 지점에서도 1% 유의수준에서 통계적으로 유의미한 상관관계가 있었고 r 값도 봄철에 비해서는 조금 높았다($r = 0.53$). 다른 지점에서는 모두 매우 높은 상관계수 값을 보였다(Figure 5-14d).

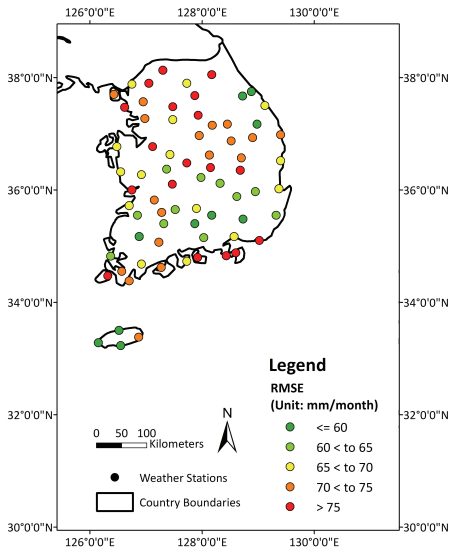
여름철에는 다섯 개의 지점에서 1% 유의수준에서 통계적으로 유의미한 상관관계를

나타내지 못하였는데, 이 둘 사이에 지리적인 공통점은 보이지 않았다(Figure 5-14c). 5% 유의수준에서는 성산(188)과 장흥(260)이 통계적으로 유의미한 상관관계를 보이지 못하였는데(성산과 장흥에 대하여 각각 $r = 0.38, 0.36$), 장흥은 남서부 해안과 가까운 곳에 위치하여 있다. 켄달 순위 상관계수로 평가하였을 때는 1% 유의수준으로 보았을 때 세 지점이 통계적으로 유의미한 상관관계를 보이지 못하였는데, 성산(188) 지점은 여기에 포함되지 않았다. 5% 유의수준에서는 장흥(260) 지점만이 통계적으로 유의미한 상관관계를 보이지 못하였다($\tau = 0.26, p\text{-value} = 0.05$).

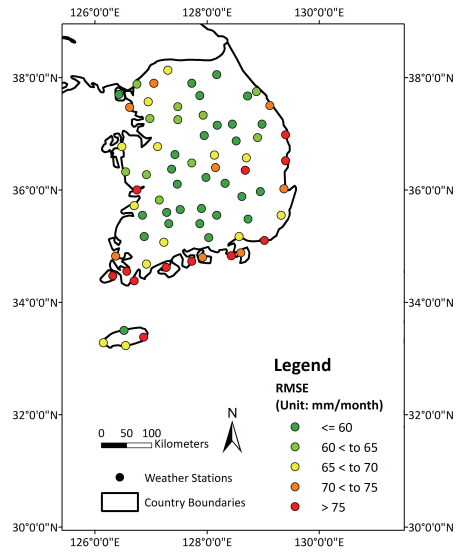
겨울철에는 72개 지점 중 13개 지점이 1% 유의수준에서 통계적으로 유의미한 상관관계를 나타내지 못하였으며, 7개의 지점이 5% 유의수준에서 통계적으로 유의미한 상관관계를 나타내지 못하였는데, 이 7개의 지점 중 두 지점은 제주도에, 네 지점은 남해 연안에 위치한다(Figure 5-14e). 비모수 통계에서는 17개 지점과 12개 지점이 각각 1%와 5%의 유의수준에서 통계적으로 유의미한 상관관계를 나타내지 못하였다. 특히 성산(188) 지점과 강원도에 위치한 태백(216) 지점에서 상관관계가 낮게 나타났다(각각 $r = 0.01, 0.14$).



(a)



(b)



(c)

Figure 5-13. RMSE between NOAH-LSM PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: mm/month).

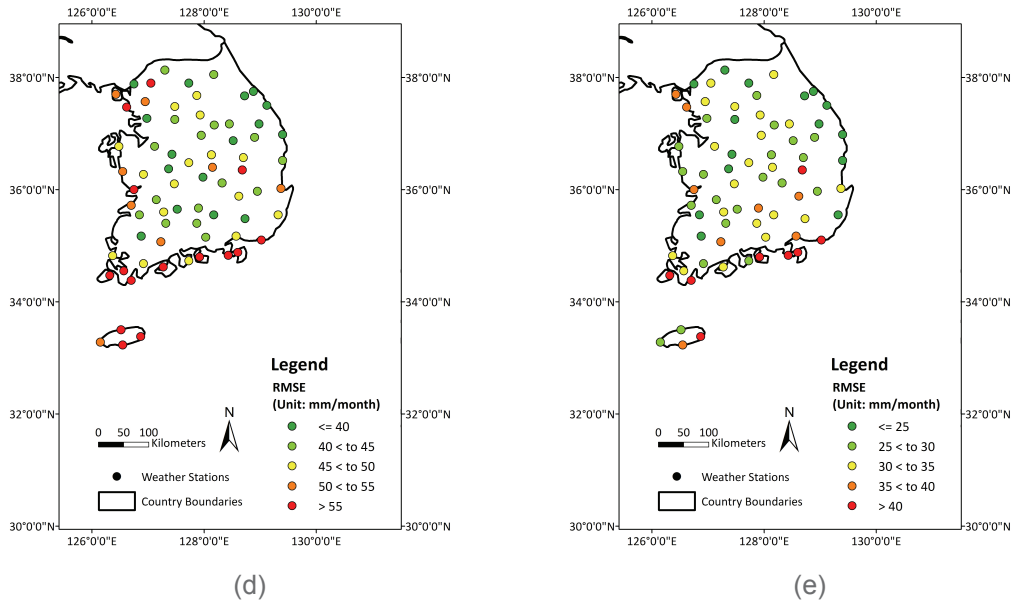


Figure 5-13. (Continued). RMSE between NOAH-LSM PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unit: mm/month).

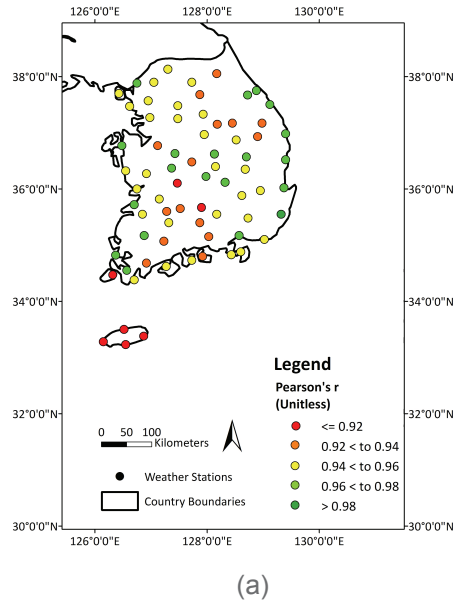
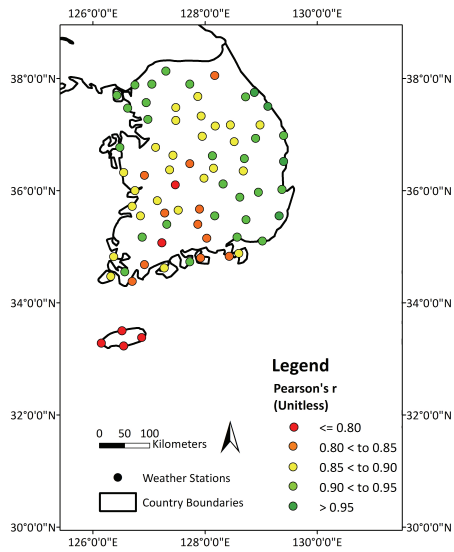
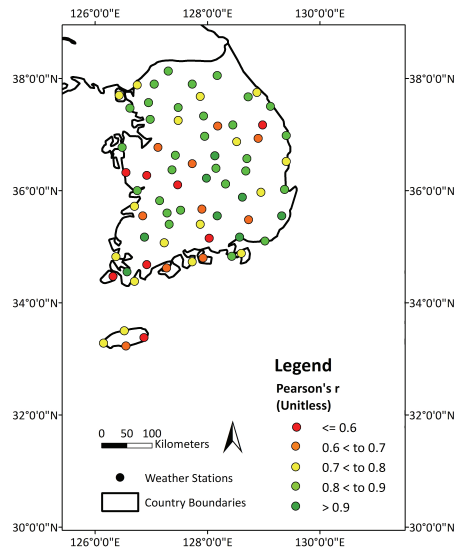


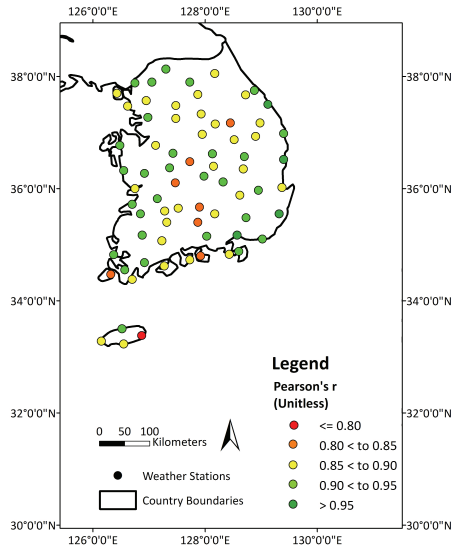
Figure 5-14. Pearson's r between NOAH-LSM PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).



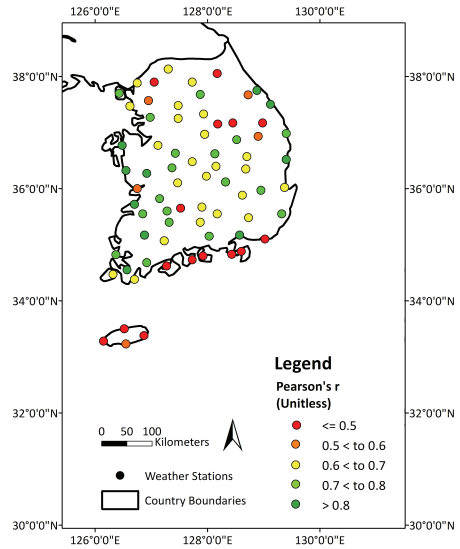
(b)



(c)



(d)



(e)

Figure 5-14. (Continued). Pearson's r between NOAH-LSM PET and calculated PET during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).

상대적으로 높은 상관계수 값과 작은 오차 값을 가지는 강릉 지점에 대하여(ID: 105, $r = 0.87$, $RMSE = 9.92$ mm/month) NOAH-LSM PET와 Penman-Monteith 식을 이용하여 계산한 PET의 시계열을 살펴보았다(Figure 5-15). 여름철에는 MODIS PET와 마찬가지로 NOAH-LSM이 잠재 증발산량을 과대추정하고 있음을 볼 수 있는데, 그 밖의 계절 특히 겨울철에는 MODIS PET와는 다르게 Penman-Monteith 식을 이용하여 계산한 PET를 매우 잘 추정하고 있음을 알 수 있다(Figure 5-15).

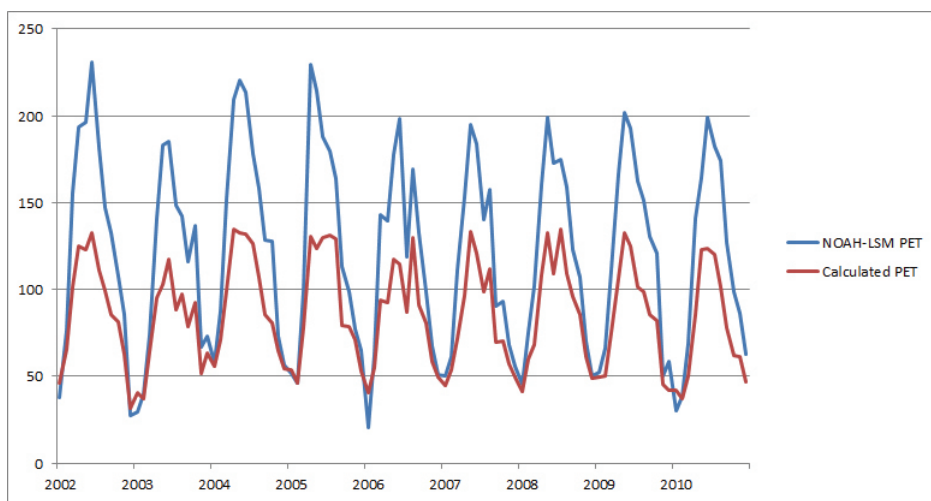


Figure 5-15. Time-series of NOAH-LSM PET and calculated PET in Gangneung station (ID: 105) (Unit: mm/month).

5.1.4 토양수분

제4장에서 설명한대로, 경기도 파주 설마천의 혼효립과 전라남도 해남의 농경지 지점에 대하여 관측된 토양수분 자료를 AMSR-E 센서로부터 도출한 토양수분 및 NOAH-LSM 산출물인 토양수분과 비교하였다. 단위는 %(vol/vol) 으로 통일하였다. 설마천 지점은 2008년 자료가, 해남 지점은 2004, 2005, 2006년 자료와 2008년 자료

가 가용하다.

5.1.4.1 월별 AMSR-E 토양수분 vs. KoFlux 토양수분

AMSR-E 토양수분 자료는 하루에 Aqua 위성의 승교점과 강교점의 두 시점의 값이 주어진다. 승교점(ascending mode) 및 강교점(descending mode)의 토양수분을 월별로 평균하거나(각각 A_SM, D_SM), 그 평균(SM_MEAN)을 월별로 평균하여 사용하였다. KoFlux 자료는 30분마다 주어지므로, AMSR-E 센서의 승교점(1:30 pm) 및 강교점(1:30 am)에 해당되는 시각의 자료를 도출하여 같은 방식으로 평균하였다. AMSR-E 자료가 승교점 혹은 강교점의 자료만 존재하는 경우도 있으므로, 이렇게 한쪽만 있는 경우는 그 날의 일별 평균 토양수분이 존재하지 않게 된다. 따라서 A_SM과 D_SM의 평균이 SM_MEAN과 같지 않을 수 있다.

사용한 KoFlux 지점과 가장 가까운 AMSR-E 격자의 값을 비교하였으며, 모든 분석은 전체 계절 및 각 계절에 대해 수행하였다. 월별 토양수분을 비교한 결과를 Tables 5-5~5-7에 나타내었다. 월별 자료를 이용하여 전체 계절에 대해 분석한 경우에는 해남 지점의 경우 표본의 개수가 40에 이르므로 모수 통계를 통한 피어슨 상관관계수 r 값을 보는 것이 적합하나, 설마천 지점의 경우라든지 계절별로 분석한 경우에는 표본의 수가 매우 작으므로 켄달의 순위 상관관계수인 τ 값을 보는 것이 더욱 적합하다.

해남 지점의 경우 A_SM, D_SM, SM_MEAN 의 경우 모두 각각 1%, 5%, 그리고 5% 유의수준에서 통계적으로 유의미한 상관관계를 나타냈다. 그러나 이 세 경우 모두 위성 영상에서 도출한 토양수분과 KoFlux 지점에서 TDR 센서를 이용하여 관측한 토양수분 자료가 부(-)의 상관관계를 보였다. 계절별로 살펴보면, A_SM 만이 봄철에 관측값과 5% 유의수준에서 통계적으로 유의미한 상관관계를 나타냈는데, 이때는 정(+)의 상관관계를 보였으며, 상관계수도 비교적 컸다($\tau = 0.45$). 설마천 지점의 경우에는 켄달의 순위 상관계수로 볼 때 모든 변수와 계절에 대해서 통계적으로 유의미한 상관관계를 보이지 않았다(Tables 5-5~5-7).

해남 지점의 KoFlux 토양 수분 관측 자료와 AMSR-E 토양 수분 자료를 AMSR-E의 승교점, 강교점에 해당하는 시각 자료의 월별 평균과 두 시각 자료 평균의 월별 평균의 시계열로 살펴보았다(Figure 5-16). 위 표에 나타난 통계 값에서 보았듯이 KoFlux 관측 값과 AMSR-E 값의 사이에 큰 차이가 있으며 경향도 전혀 맞지 않음을 알 수 있다. 그 원인은 여러 가지가 있을 수 있는데, 첫 번째로 토양 수분의 공간적 변이가 매우 크다는 점으로 AMSR-E의 25×25 km 공간해상도로는 그 변이를 관찰하기가 매우 어렵다. 두 번째 원인으로서는 해남 KoFlux 관측 지점이 농경지이므로 물 논일 가능성이 매우 높는데 이 경우 많은 경우 물이 차 있는 상태이기 때문에 실제 토양 수분 값을 제대로 얻기 힘들다. 셋째로 AMSR-E의 토양 수분 자료는 토양 상층 1cm 깊이까지의 토양 수분을 추정하는데, KoFlux 지점의 관측은 TDR 센서로 상층 10cm 깊이까지의 토양 수분을 측정한다.

Table 5-5. Statistics between monthly averaged AMSR-E A_SM and KoFlux Soil Moisture for two KoFlux Sites.

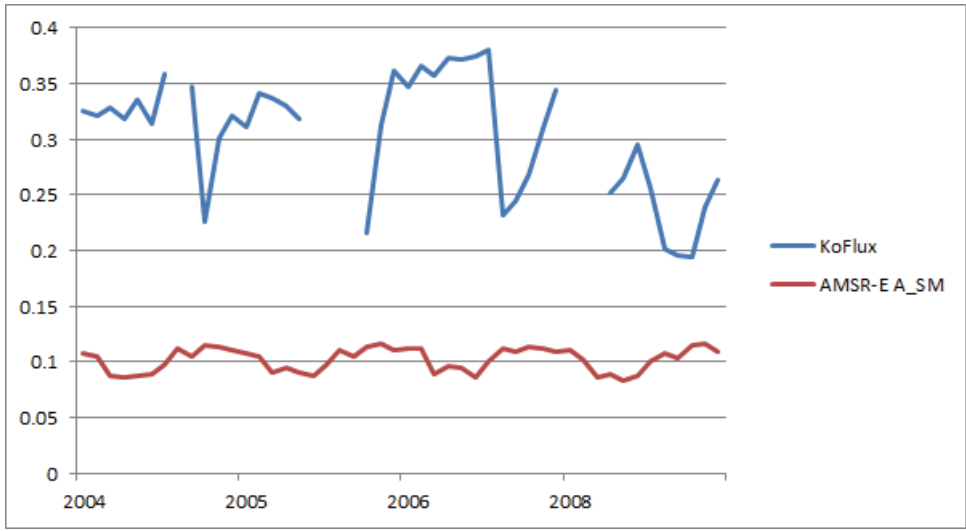
SITE	SEASON	n	RMSE	MAE	Obs. range (min~max)	r	r_p	tau	tau_p
HAENAM	ALL	40	0.2099	0.2017	0.1942~0.3809	-0.34	0.03	-0.22	0.04
	MAM	11	0.2385	0.2359	0.2524~0.3733	0.62	0.04	0.45	0.05
	JJA	8	0.2158	0.2044	0.2019~0.3809	-0.61	0.11	-0.36	0.22
	SON	11	0.1557	0.1475	0.1942~0.3467	-0.09	0.78	-0.09	0.7
	DJF	10	0.2231	0.2214	0.12636~0.3655	0.34	0.34	0.33	0.18
SEOJMA-CHEON	ALL	12	0.1023	0.0925	0.1090~0.2553	-0.34	0.28	-0.18	0.41
	MAM	3	0.1078	0.1078	0.1941~0.2168	1	0.02	1	0.12
	JJA	3	0.1434	0.1431	0.2283~0.2553	0.67	0.54	0.33	0.6
	SON	3	0.0383	0.0303	0.1091~0.1688	-1	0.04	-1	0.12
	DJF	3	0.0906	0.0889	0.1557~0.2095	0.84	0.37	1	0.12

Table 5-6. Statistics between monthly averaged AMSR-E D_SM and KoFlux Soil Moisture for two KoFlux Sites.

SITE	SEASON	n	RMSE	MAE	Obs. range (min~max)	r	r_p	tau	tau_p
HAENAM	ALL	40	0.2215	0.2136	0.1952~0.3807	-0.52	< 0.001	-0.34	< 0.001
	MAM	11	0.2509	0.2483	0.2569~0.3749	0.12	0.73	-0.05	0.82
	JJA	8	0.2176	0.2064	0.1995~0.3807	-0.66	0.07	-0.29	0.32
	SON	11	0.1687	0.1611	0.1952~0.3559	-0.27	0.42	-0.16	0.48
	DJF	10	0.2406	0.2390	0.2620~0.3671	0.27	0.45	0.24	0.33
SEOJMA- CHEON	ALL	12	0.0923	0.0833	0.1083~0.2488	-0.21	0.51	-0.09	0.68
	MAM	3	0.0999	0.0998	0.1922~0.2146	0.99	0.09	0.33	0.6
	JJA	3	0.1278	0.1276	0.2280~0.2488	0.56	0.62	0.33	0.6
	SON	3	0.0289	0.0245	0.1083~0.1672	0.39	0.75	0.33	0.6
	DJF	3	0.0834	0.0815	0.1562~0.2084	0.98	0.14	1	0.12

Table 5-7. Statistics between monthly averaged AMSR-E SM_MEAN and KoFlux Soil Moisture for two KoFlux Sites.

SITE	SEASON	n	RMSE	MAE	Obs. range (min ~ max)	r	r_p	tau	tau_p
HAENAM	ALL	40	0.2157	0.2077	0.1947~0.3808	-0.44	< 0.001	-0.31	< 0.001
	MAM	11	0.2447	0.2421	0.2546~0.3741	0.39	0.23	0.27	0.24
	JJA	8	0.2167	0.2054	0.2007~0.3808	-0.64	0.09	-0.43	0.14
	SON	11	0.1622	0.1543	0.1947~0.3513	-0.2	0.55	-0.16	0.48
	DJF	10	0.2318	0.2302	0.2628~0.3663	0.32	0.36	0.24	0.33
SEOLMA- CHEON	ALL	12	0.0973	0.0879	0.1087~0.2521	-0.29	0.36	-0.18	0.41
	MAM	3	0.1038	0.1038	0.1932~0.2157	1	0.04	1	0.12
	JJA	3	0.1356	0.1354	0.2281~0.2521	0.62	0.57	0.33	0.6
	SON	3	0.0336	0.0274	0.1087~0.1680	-0.73	0.48	-0.33	0.6
	DJF	3	0.0870	0.0852	0.1559~0.2089	0.91	0.27	1	0.12

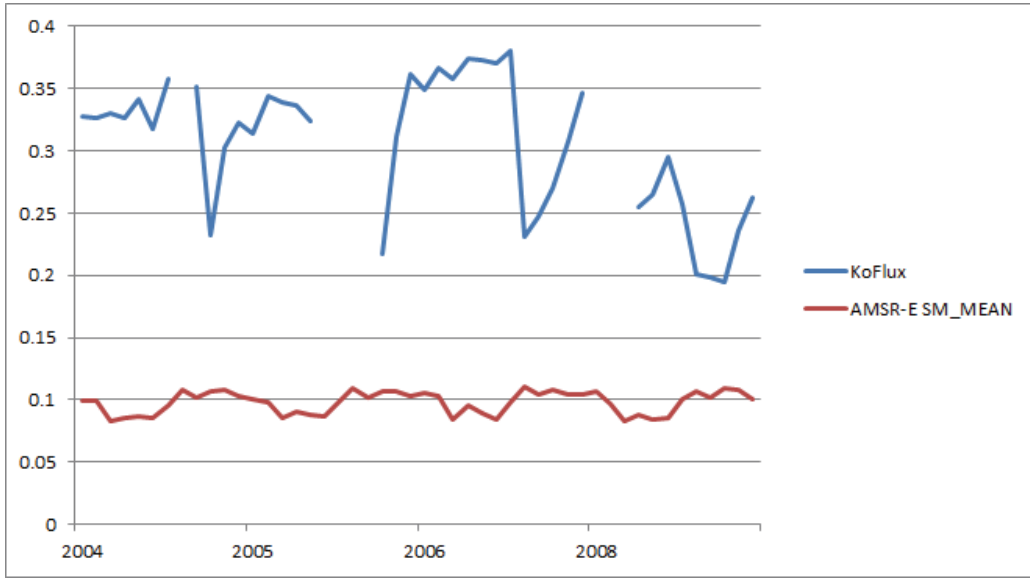


(a)



(b)

Figure 5-16. Time-series of KoFlux Soil Moisture and monthly averaged AMSR-E (a) A_SM, (b) D_SM, and (c) SM_MEAN in Haenam station (Unit: g/cm³).



(c)

Figure 5-16. (Continued). Time-series of KoFlux Soil Moisture and monthly averaged AMSR-E (a) A_SM, (b) D_SM, and (c) SM_MEAN in Haenam station (Unit: g/cm³).

5.1.4.2 일별 AMSR-E 토양수분 vs. KoFlux 토양수분

월별 비교 분석의 경우 표본의 개수가 작아 상관관계를 제대로 알 수 없을 수도 있으므로, 일별 자료를 이용하여 비교, 상관관계 분석을 수행하였다(Tables 5-8~5-10). 자료는 Aqua 위성의 승교점과 강교점에 해당하는 시각의 자료를 도출하여 분석에 이용하였다.

해남 지점의 경우 A_SM, D_SM, SM_MEAN 에 대해서 연중 전체 계절의 경우와 봄철, 여름철의 경우 통계적으로 유의미한 상관관계를 나타냈다. 그러나 전체 계절과 여름철의 경우 위성 영상에서 도출한 토양수분과 KoFlux 지점에서 TDR 센서를 이용하여 관측한 토양수분 자료가 부(-)의 상관관계를 보였다. 봄철에는 A_SM, D_SM,

SM_MEAN 에 대해서 월별 분석의 경우와 마찬가지로 정(+)의 상관관계를 볼 수 있었다. 설마천 지점의 경우 A_SM의 경우에는 봄철 및 겨울철의 경우 5% 유의수준에서 통계적으로 유의미한 정의 상관관계를 나타냈다. D_SM과 SM_MEAN의 경우에는 봄철, 여름철, 겨울철에 유의미한 정의 상관관계를 보였다.

해남 지점의 KoFlux 토양 수분 관측 자료와 AMSR-E 토양 수분 자료를 AMSR-E 의 승교점, 강교점에 해당하는 시각 자료와 두 시각 자료 평균의 시계열로 살펴보았다 (Figure 5-17). 월별 평균의 비교와 마찬가지로 KoFlux 관측 값과 AMSR-E 값의 사이에 큰 차이가 있으며 경향도 전혀 맞지 않음을 알 수 있다.

Table 5-8. Statistics between daily AMSR-E A_SM and KoFlux Soil Moisture for two KoFlux Sites.

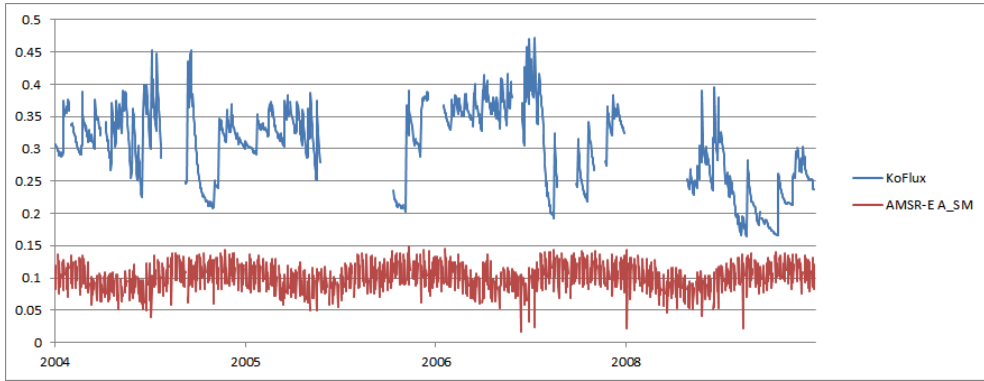
SITE	SEASON	n	RMSE	MAE	Obs. range (min~max)	r	r_p	tau	tau_p
HAENAM	ALL	770	0.2150	0.2031	0.1643~0.4712	-0.16	<0.001	-0.09	<0.001
	MAM	216	0.2477	0.2442	0.2294~0.4158	0.24	<0.001	0.16	<0.001
	JJA	158	0.2176	0.1999	0.1643~0.4712	-0.27	<0.001	-0.17	<0.001
	SON	194	0.1591	0.1441	0.1653~0.4534	-0.03	0.693	0.03	0.508
	DJF	202	0.2217	0.2184	0.2373~0.3886	0.07	0.312	0.05	0.271
SEOLMA-CHEON	ALL	264	0.1082	0.0954	0.0850~0.3440	-0.09	0.154	-0.04	0.348
	MAM	69	0.1128	0.1088	0.1390~0.2760	0.37	0.002	0.25	0.002
	JJA	68	0.1479	0.1438	0.1660~0.3440	0.22	0.068	0.19	0.020
	SON	65	0.0551	0.0372	0.0850~0.3240	-0.12	0.350	0.02	0.850
	DJF	62	0.0923	0.0884	0.1310~0.2520	0.27	0.035	0.20	0.022

Table 5-9. Statistics between daily AMSR-E D_SM and KoFlux Soil Moisture for two KoFlux Sites.

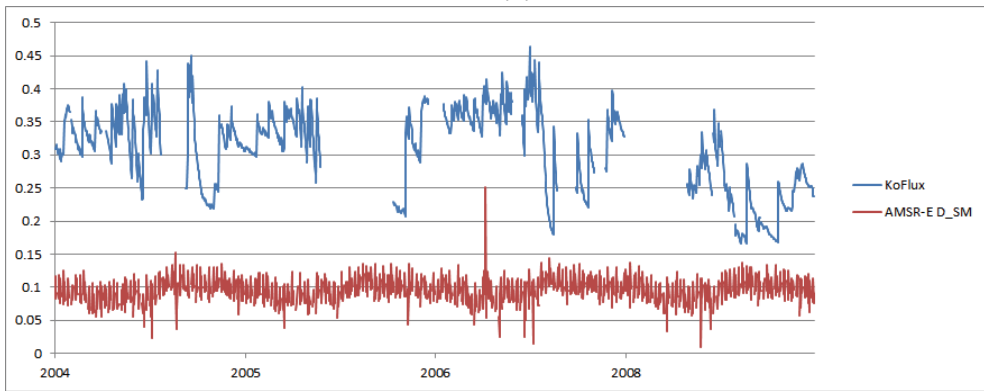
SITE	SEASON	n	RMSE	MAE	Obs. range (min~max)	r	r_p	tau	tau_p
HAENAM	ALL	771	0.2260	0.2153	0.1668~0.4646	-0.15	<0.001	-0.10	<0.001
	MAM	216	0.2598	0.2564	0.2343~0.4254	0.16	0.019	0.12	0.010
	JJA	157	0.2195	0.2035	0.1668~0.4646	-0.23	0.004	-0.14	0.009
	SON	195	0.1693	0.1572	0.1679~0.4511	0.02	0.741	0.03	0.542
	DJF	203	0.2394	0.2364	0.2372~0.3982	0.08	0.234	0.05	0.257
SEOJMA- CHEON	ALL	260	0.0968	0.0851	0.0850~0.3290	0.01	0.885	0.03	0.442
	MAM	68	0.1035	0.0995	0.1360~0.2820	0.35	0.003	0.21	0.011
	JJA	69	0.1321	0.1286	0.1670~0.3290	0.32	0.007	0.22	0.008
	SON	67	0.0443	0.0319	0.0850~0.3140	0.03	0.811	0.02	0.772
	DJF	56	0.0814	0.0778	0.1310~0.2360	0.34	0.009	0.24	0.009

Table 5-10. Statistics between daily AMSR-E SM_MEAN and KoFlux Soil Moisture for two KoFlux Sites.

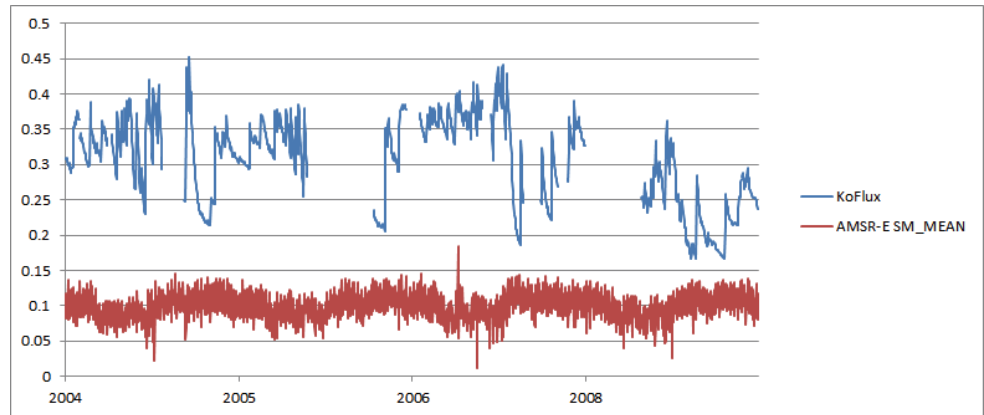
SITE	SEASON	n	RMSE	MAE	Obs. range (min~max)	r	r_p	tau	tau_p
HAENAM	ALL	1042	0.2184	0.2075	0.1668~ 0.4523	-0.15	<0.001	-0.09	<0.001
	MAM	292	0.2518	0.2485	0.2319~ 0.4162	0.21	<0.001	0.14	<0.001
	JJA	213	0.2154	0.1992	0.1668~ 0.4415	-0.25	<0.001	-0.16	<0.001
	SON	264	0.1625	0.1497	0.1679~ 0.4511	0.00	0.948	0.03	0.543
	DJF	273	0.2293	0.2261	0.2372~ 0.3905	0.08	0.172	0.06	0.126
SEOLMA- CHEON	ALL	359	0.1015	0.0896	0.0850~ 0.3125	-0.05	0.383	-0.01	0.875
	MAM	92	0.1071	0.1035	0.1400~ 0.2715	0.38	<0.001	0.22	0.002
	JJA	92	0.1382	0.1349	0.1690~ 0.3125	0.35	<0.001	0.25	<0.001
	SON	89	0.0489	0.0335	0.0850~ 0.3140	-0.05	0.626	0.02	0.816
	DJF	86	0.0886	0.0845	0.1310~ 0.2520	0.28	0.008	0.19	0.010



(a)



(b)



(c)

Figure 5-17. Time-series of KoFlux Soil Moisture and daily averaged AMSR-E (a) A_SM, (b) D_SM, and (c) SM_MEAN in Haenam station (Unit: g/cm³).

5.1.4.3 월별 NOAH-LSM 토양수분 vs. KoFlux 토양수분

NOAH-LSM 지면모델 산출물인 토양수분과 KoFlux 관측 토양수분 값을 비교하였다(Table 5-11). NOAH-LSM 산출물은 3시간 간격으로 생산되는데, 이를 월별로 모두 평균한 값과 30분마다 생산되는 KoFlux 자료를 월별로 모두 평균한 값을 비교하였다. 그러나 상관관계 분석의 p 값에서 알 수 있듯 두 KoFlux 지점에서 어느 계절에 대해 분석한 결과도 통계적으로 유의한 상관관계를 보이지 못하였다. 다만 RMSE, MAE 오차 값은 앞 절의 AMSR-E와 KoFlux 관측 값을 비교한 결과와 비교하여 볼 때 해남 지점의 경우 매우 낮게 나타났다.

해남 지점의 월별 평균한 NOAH-LSM 토양 수분 산출물과 KoFlux 관측 자료를 비교하였다(Figure 5-18). AMSR-E 토양 수분 자료의 경우와 비교하여 오차의 폭이 매우 작은 것을 관찰할 수 있었으나, 경향은 잘 맞지 않았다.

Table 5-11. Statistics between monthly NOAH-LSM Soil Moisture and KoFlux Soil Moisture for two KoFlux Sites.

SITE	SEASON	n	RMSE	MAE	Obs. range (min~max)	r	r_p	tau	tau_p
HAENAM	ALL	40	0.0701	0.0539	0.1947~0.3808	0.17	0.28	0.06	0.58
	MAM	11	0.0465	0.0315	0.2546~0.3741	-0.15	0.65	-0.05	0.82
	JJA	8	0.0937	0.0764	0.2007~0.3808	0.61	0.11	0.43	0.14
	SON	11	0.0898	0.0793	0.1947~0.3513	0.51	0.11	0.16	0.48
	DJF	10	0.0371	0.0325	0.2628~0.3663	-0.07	0.86	-0.42	0.09
SEOLMA- CHEON	ALL	12	0.1471	0.1368	0.1087~0.2521	0.25	0.43	0.15	0.49
	MAM	3	0.0994	0.0869	0.1932~0.2157	-0.68	0.52	-0.33	0.6
	JJA	3	0.1428	0.1426	0.2281~0.2521	0.82	0.38	0.33	0.6
	SON	3	0.2117	0.2111	0.1087~0.1680	0.77	0.44	0.33	0.6
	DJF	3	0.1072	0.1068	0.1559~0.2089	0.91	0.27	0.33	0.6

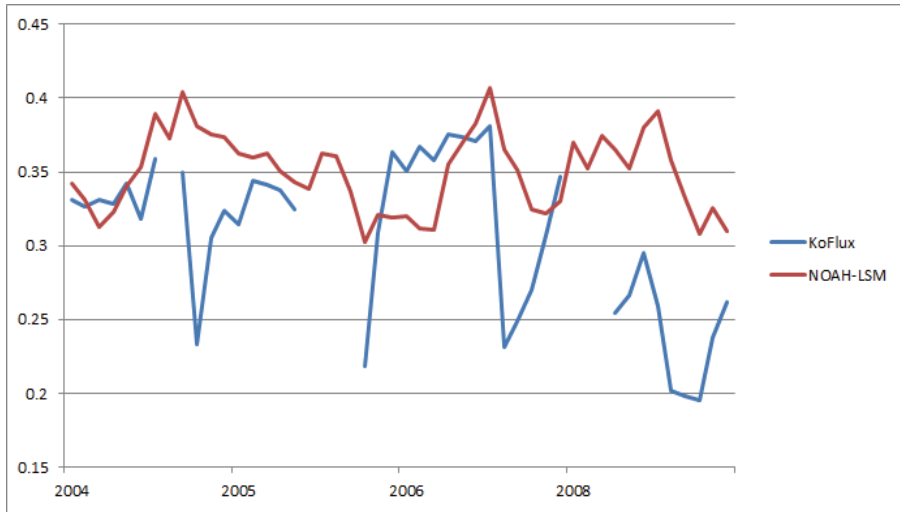


Figure 5-18. Time-series of monthly NOAH-LSM Soil Moisture and KoFlux Soil Moisture in Haenam station (Unit: g/cm³).

5.1.4.4 일별 NOAH-LSM 토양수분 vs. KoFlux 토양수분

NOAH-LSM의 경우도 역시 월별 비교 분석의 경우 표본의 개수가 작아 상관관계를 제대로 알 수 없을 수도 있으므로, 일별 자료를 이용하여 비교, 상관관계 분석을 수행하였다. 자료는 NOAH-LSM의 경우 3시간 산출물을 모두 평균하여 일별 자료로 전환하였고, KoFlux 관측 자료는 30분 자료를 모두 평균하여 일별 자료로 전환하였다.

일별로 분석한 결과는 월별로 비교 분석한 결과와 매우 다르게 나타났다. 해남 지점의 경우 겨울철을 제외한 경우 모두 1% 유의수준에서 통계적으로 유의미한 상관관계를 보였으며, 상관계수도 특히 가을철에 비교적 높게 나타났다($r = 0.53$; Table 5-12). 설마천의 경우에는 봄철을 제외하고는 모두 1% 유의수준에서 통계적으로 유의미한 상관관계를 보였으며, 상관계수도 여름철, 가을철, 겨울철에 매우 높게 나타났다(각각 $r = 0.6, 0.58, 0.75$; Table 5-12). 오차 값은 설마천 지점의 경우에는 비슷한 수준이거나

약간 크게 나타났으나, 해남 지점의 경우에는 AMSR-E 의 경우보다 상대적으로 적은 값을 보였다(Table 5-12).

해남 지점의 일별 평균한 NOAH-LSM 토양 수분 산출물과 KoFlux 관측 자료를 비교하였다(Figure 5-19). 월별 비교와 마찬가지로 AMSR-E 토양 수분 자료의 경우와 비교하여 오차의 폭이 매우 작은 것을 관찰할 수 있었으나 몇몇 구간을(2006년 봄철 등) 제외하고 경향은 잘 맞지 않았다.

Table 5-12. Statistics between daily NOAH-LSM Soil Moisture and KoFlux Soil Moisture for two KoFlux Sites.

SITE	SEASON	n	RMSE	MAE	Obs. range (min~max)	r	r_p	tau	tau_p
HAENAM	ALL	1050	0.0745	0.0576	0.1668~0.4523	0.26	< 0.001	0.16	< 0.001
	MAM	293	0.0414	0.0313	0.2319~0.4162	0.27	< 0.001	0.23	< 0.001
	JJA	214	0.1032	0.0847	0.1668~0.4415	0.6	< 0.001	0.45	< 0.001
	SON	270	0.0972	0.0844	0.1679~0.4511	0.53	< 0.001	0.38	< 0.001
	DJF	273	0.0428	0.0379	0.2372~0.3905	0.02	0.8	-0.05	0.23
SEOJMA-CHEON	ALL	361	0.1501	0.1392	0.0850~0.3125	0.29	< 0.001	0.23	< 0.001
	MAM	92	0.1051	0.095	0.1400~0.2715	0.08	0.45	0.08	0.29
	JJA	92	0.145	0.1426	0.1690~0.3125	0.6	< 0.001	0.45	< 0.001
	SON	91	0.2141	0.2113	0.0850~0.3140	0.58	< 0.001	0.27	< 0.001
	DJF	86	0.1084	0.1066	0.1310~0.2520	0.75	< 0.001	0.4	< 0.001

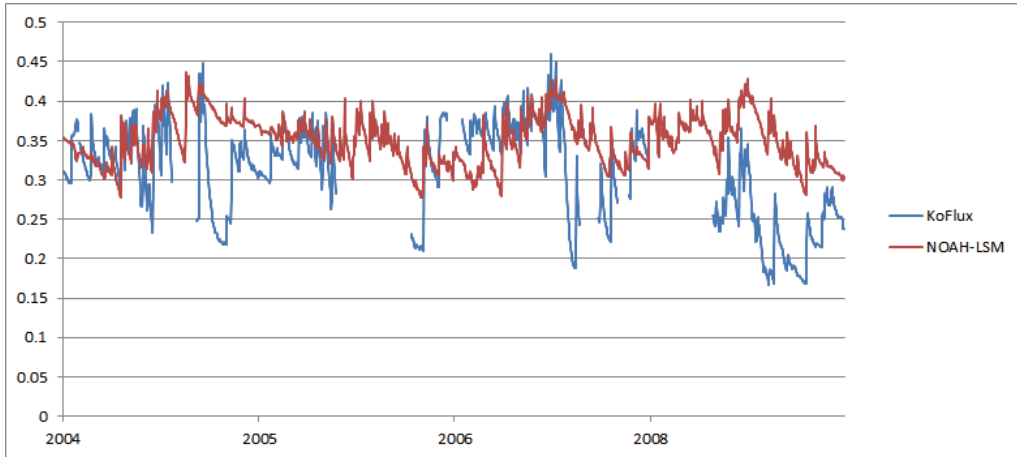


Figure 5-19. Time-series of daily averaged NOAH-LSM Soil Moisture and KoFlux Soil Moisture in Haenam station (Unit: g/cm³).

5.1.4.5 AMSR-E 토양수분과 NOAH-LSM 토양수분의 결합

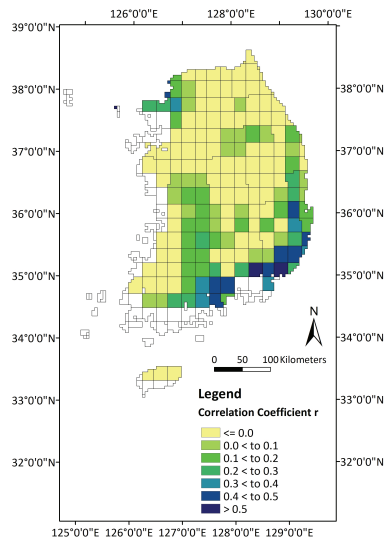
AMSR-E 토양수분 자료와 KoFlux 관측 값을 일별로 비교한 경우와(Table 5-8), NOAH-LSM 토양수분 산출물과 KoFlux 관측 값을 일별로 비교한 경우를(Table 5-12) 살펴볼 때, 설마천 지점에서 봄철의 경우 AMSR-E 토양수분 자료를 이용한 경우 1% 유의수준에서 통계적으로 유의미한 상관관계를 보였으나($r = 35$), NOAH-LSM의 경우 그렇지 못하였다. 그러나 그 밖에 경우에는 NOAH-LSM이 KoFlux 관측 값에 그나마 가까운 결과를 생산하였다.

본 연구에서는 지면모델을 이용한 경우가 위성 영상만을 이용한 경우보다 대부분 관측 값에 가까운 토양수분 값을 제공하는 것을 볼 수 있었다. 그러나 봄철 설마천 지점의 자료와 마찬가지로 지면모델의 결과가 항상 위성 영상에서 도출한 자료보다 우수한 것은 아니므로, 위성 영상에서 도출한 토양수분 자료와 지면모델의 산출물인 토양수분 자료를 결합하여 토양수분 추정의 정확도를 개선시킬 수 있을지 가능성을 시험하였다. 앞

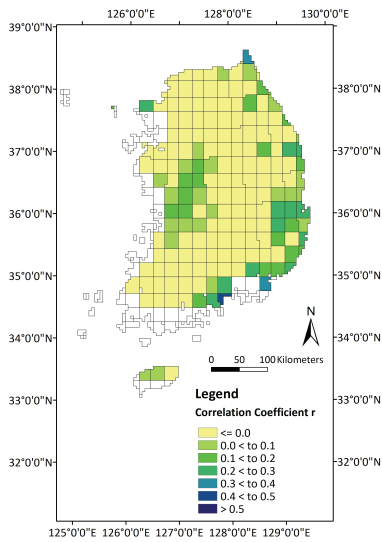
절에서 보았듯 상관관계는 월별 자료의 비교가 훨씬 좋았으나, 본 연구에서 개발한 가뭄 감시 및 예측 모델이 모두 월별 자료를 사용하고 있으므로 AMSR-E 토양수분 자료와 NOAH-LSM 토양수분 산출물의 결합도 월별 자료를 이용하여 수행하였다.

우선 AMSR-E 위성 영상 자료의 25×25km 격자에 맞추어 NOAH-LSM 토양수분 산출물을 리샘플 하였고, 각각의 격자 위치에서 회귀 분석을 통해 상관관계를 구하였다 (Figure 5-20). 표본의 개수는 2002년 7월에서 2011년 9월까지 두 자료가 가용한 월들로, 전체 계절, 봄철, 여름철, 가을철, 겨울철에서 각각 111, 27, 29, 28, 27 개월 이 사용되었다. 계절별 분석에서 표본의 개수가 모수 통계 값을 보기에 약간 작으나, 피어슨 r 값과 켄달 τ 값이 거의 비슷하게 도출되어 그림에서는 피어슨 r 값을 살펴보았다.

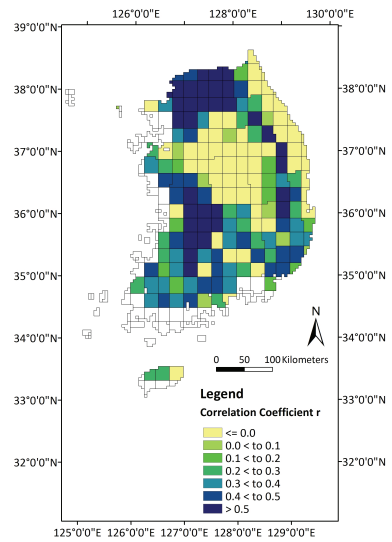
봄철의 경우에는 동해안을 따라 강원도와 경상남북도 지역과 충청남도 지역이 비교적 상관관계가 높게 나타났으며(Figure 5-20b), 가을철에는 동해안 지역과 경기도 및 전라남도 일부 지역(Figure 5-20d), 겨울철에는 강원도와 경상남도 지역 등에서 높은 상관관계를 보였다(Figure 5-20e). 여름철에는 전체적으로 AMSR-E 토양수분과 NOAH-LSM 토양수분 산출물의 상관관계가 높았는데, 내륙 지방과 동해안을 제외한 지역에서 비교적 높은 상관계수를 보였다(Figure 5-20c).



(a)



(b)



(c)

Figure 5-20. Pearson's r between AMSR-E Soil Moisture and NOAA-LSM Soil Moisture during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).

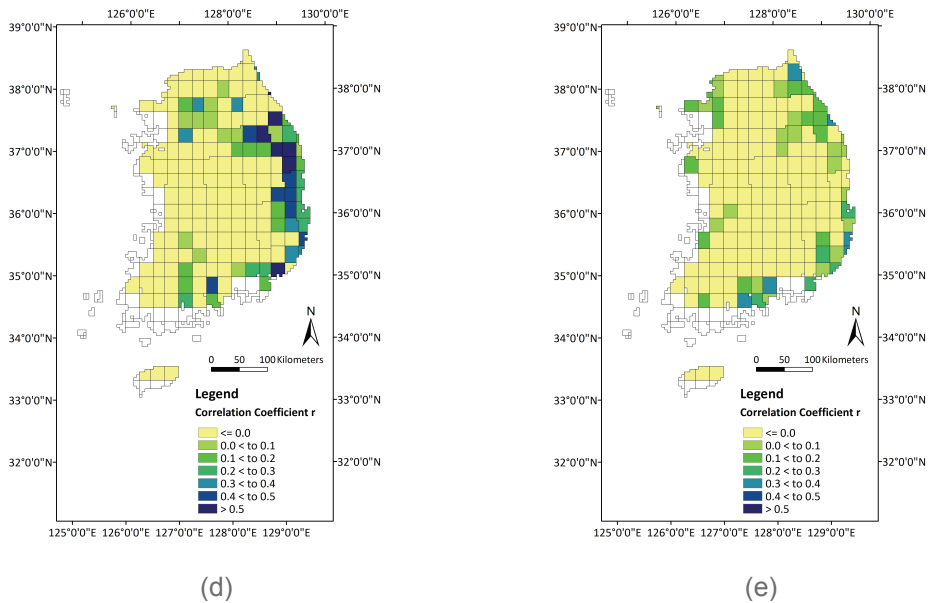


Figure 5-20. (Continued). Pearson's r between AMSR-E Soil Moisture and NOAA-LSM Soil Moisture during (a) all season, (b) MAM, (c) JJA, (d) SON, and (e) DJF (Unitless).

공간해상도 $25 \times 25\text{km}$ 격자들을 이용하여 회귀 모델을 구축하였다. 이 때, 전체 격자를 모두 포함시키는 경우 상관관계가 1% 유의수준에서 통계적으로 유의하기는 하나, 모두 부(-)의 상관관계를 나타냈다. 따라서 격자들 중에서 AMSR-E 토양수분과 NOAA-LSM 토양수분 산출물의 피어슨 상관관계수 값이 양인 격자만 사용한 경우와 어스 상관관계수 값이 0.5 이상인 격자만 사용한 경우에 대하여 모델을 구축하였다(Table 5-13).

양의 피어슨 상관관계수를 기준으로 삼은 경우는 전체 계절의 경우는 통계적으로 유의하지 못했으며 봄철과 여름철, 겨울철은 전체 격자를 사용한 경우와 같이 부의 상관관계를 나타냈다. 가을철은 5% 유의수준에서 통계적으로도 유의미하고 정의 상관관계를 보였으나 상관관계수 값이 그다지 크지 못하였다. 피어슨 상관관계수 0.5 이상인 격자만 포함시킨 경우는 봄철과 겨울철의 경우에는 해당하는 격자가 존재하지 않았다. 그러나 전

체 계절과 가을철의 경우는 1% 유의수준에서 통계적으로 유의미한 상관관계를 보였으며 상관계수 값도 비교적 높았다.

Table 5-13. Statistics between monthly AMSR-E SM_MEAN and NOAA-LSM Soil Moisture.

VARIABLE	SEASON	n	RMSE	MAE	LSM range (min ~max)	slope	inter cept	r	r_p
For pixels with r > 0.0	ALL	8880	0.2096	0.2037	0.1658~ 0.4649	0.01	0.32	0	0.71
	MAM	1566	0.2055	0.1981	0.1658~ 0.4546	-0.13	0.33	-0.12	< 0.001
	JJA	3480	0.233	0.2283	0.2178~ 0.4149	-0.04	0.35	-0.04	0.03
	SON	1512	0.2186	0.215	0.2104~ 0.4121	0.23	0.31	0.18	< 0.001
	DJF	1296	0.198	0.188	0.1862~ 0.4624	-0.13	0.32	-0.11	< 0.001
For pixels with r > 0.5	ALL	333	0.2353	0.2334	0.1908~ 0.3980	1.09	0.22	0.45	< 0.001
	JJA	957	0.2205	0.2177	0.2339~ 0.4049	0.18	0.32	0.12	< 0.001
	SON	196	0.2061	0.2034	0.2833~ 0.4070	0.37	0.30	0.4	< 0.001

전체 계절을 포함하여 분석하였을 때 AMSR-E 토양수분 자료와 NOAA-LSM 토양수분 산출물의 상관계수가 0.5 이상인 격자만을 선정하여, 이 상관관계에서 도출한 기울

기(1.09)와 편차(0.22)를 이용하여 AMSR-E 자료 값을 전환시키고, 이를 다시 KoFlux 관측 토양수분 값과 비교하였다(Table 5-14). AMSR-E 자료의 전환은 선형 이동이므로 두 변수 간의 상관계수 등은 변하지 않았으나, RMSE, MAE의 오차 값은 변화하였다. 해남 지점의 경우, 오차 값이 모든 계절에 대해 감소하였으며 그러나 설마 천 지점에서는 여름철에 대하여 오차 값이 감소하였으나 다른 계절에 대해서는 약간 증가하였다. 결과적으로 얻어진 해남 지점의 오차 값은 지면모델 산출물을 KoFlux 관측 토양수분과 비교하였을 때 오차 값보다 더 작다.

여기에서는 월별 자료를 이용하여 구축된 모델을 활용하였는데, 일별 자료의 상관관계가 더 높으므로 일별 자료를 이용하여 모델을 구축하면 더 나은 결과를 얻을 수 있을 것이다. 이렇게 AMSR-E 토양수분 자료와 NOAH-LSM 토양수분 추출물을 결합하여 위성 영상 도출 자료의 토양수분 추정 값을 개선시키고 공간·시간 해상도를 높일 수 있다.

Table 5-14. Statistics between converted AMSR-E Soil Moisture and KoFlux Soil Moisture for two KoFlux Sites.

SITE	SEASON	n	RMSE	MAE	Obs. range (min~max)	r	r_p	tau	tau_p
HAENAM	ALL	40	0.0638	0.0483	0.1947~0.3808	-0.44	0	-0.31	0.01
	MAM	11	0.0371	0.0317	0.2546~0.3741	0.39	0.24	0.27	0.24
	JJA	8	0.0746	0.0604	0.2007~0.3808	-0.64	0.09	-0.43	0.14
	SON	11	0.0931	0.0816	0.1947~0.3513	-0.2	0.57	-0.16	0.48
	DJF	10	0.0281	0.0204	0.2628~0.3663	0.32	0.36	0.24	0.33
SEOLMA- CHEON	ALL	12	0.1542	0.1475	0.1087~0.2521	-0.29	0.35	-0.18	0.41
	MAM	3	0.1296	0.1296	0.1932~0.2157	0.99	0.07	1	0.12
	JJA	3	0.0987	0.0983	0.2281~0.2521	0.6	0.59	0.33	0.6
	SON	3	0.2148	0.2133	0.1087~0.1680	-0.73	0.48	-0.33	0.6
	DJF	3	0.1496	0.1486	0.1559~0.2089	0.9	0.28	1	0.12

5.2 가뭄 감시 및 예측 모델 선정

5.2.1 가뭄 감시 모델의 선정

입력 변수의 집합(SET1, SET2, SET3, SET4) 별로 각 목표 변수에 대해 random forest에서 내부적인 오차 추정 방법인 out-of-bag(OOB)을 통해 RMSE 및 각 변수의 상대적인 중요도를 산정하였다. 또한 목표 변수의 시계열 값과 random forest 모델 결과물의 시계열 값을 이용하여 이 둘 사이의 결정계수(R^2) 값을 구하였다. 상대적인 중요도는 해당 변수가 포함되지 않았을 경우 증가하는 Mean Square Error(MSE)의 백분율(%IncMSE)로 제공되는데, 여기에서 목표 변수별로도 비교가 용이하도록 사용된

입력 변수들의 중요도 합이 100%가 되도록 백분율로 전환하였다.

각 입력 변수의 집합 별로 목표 변수에 대한 RMSE 값을 보면 대부분의 경우 $RMSE(SET4) > RMSE(SET1) > RMSE(SET2) > RMSE(SET3)$ 으로 나타나는 것을 알 수 있는데, 이는 많은 변수를 사용할수록 오차가 작아지는 것을 나타낸다(Table 5-15). NDDI5와 NDDI6의 오차 값이 간혹 상대적으로 크게 나타나는데, 이는 계산 과정에서 분모를 이루는 NDWI와 NDVI의 합이 매우 작은 값이 되어 NDDI5, NDDI6의 값 자체가 매우 커지는 경우가 종종 있기 때문이다. 목표 변수 값과 모델 결과 값 사이의 결정계수 값을 보아도 $R^2(SET3) > R^2(SET2) > R^2(SET1) > R^2(SET4)$ 로 SET3의 경우 가장 결정계수가 높게 나타나 가장 우수한 성능을 보인다(Table 5-16).

TRMM PRCP 등의 경우에는 낮은 위도 지역에서 훨씬 작은 오차가 나타났으므로, 입력 변수 집합들 간의 비교 분석 결과는 이러한 지역별 오차의 차이가 드러났을 수 있다. 위성 영상에서 도출한 자료나 대규모 기후 지수만을 사용하는 경우에는(SET1과 SET4) 한반도 전체의 기상 관측 지점 위치에 대해서 분석을 수행한 반면, 지면모델 결과까지 포함한 경우에는(SET2와 SET3) 남한 지역에 위치한 지점 위치에 대해서만 분석을 수행했기 때문이다. 이러한 지역적 범위의 효과를 시험하기 위해 SET1과 SET4 중에 우수한 결과를 보인 SET1을 남한 지역에만 적용하여 SET2 및 SET3과 비교하여 보았는데, 모델의 성능이 SET2와 SET3 사이에 위치했다. 즉, 남한 지역의 지점들만 포함한 경우 확실히 모델의 성능이 우수했다. 그러나 결과적으로 같은 지역을 대상으로 한 경우에 역시 많은 변수를 포함한 경우가 모델의 성능이 우수함을 알 수 있었다.

Table 5-15. RMSE of random forest models for target variables.

Target Variable	SET1	SET2	SET3	SET4
SPI1	0.5536	0.5044	0.4784	0.617
SPI3	0.5993	0.5417	0.5196	0.6782
SPI6	0.6118	0.5417	0.5192	0.678
SPI9	0.6067	0.571	0.547	0.709
SPEI1	0.5526	0.5008	0.4764	0.6126
SPEI3	0.6054	0.5582	0.5338	0.6821
SPEI6	0.6189	0.5494	0.5276	0.6845
SPEI7	0.6189	0.5821	0.5547	0.7267
NDWI5	0.06	0.0532	0.0438	0.0692
NDWI6	0.1162	0.1148	0.0954	0.1366
NDWI7	0.1463	0.1576	0.1339	0.1772
NDDI5	1.7041	0.7937	0.7186	1.7415
NDDI6	4.2703	2.227	2.1758	4.2826
NDDI7	0.3335	0.2603	0.2362	0.3532
NMDI	0.0848	0.0695	0.0666	0.0884

Table 5-16. Coefficient of determination (R^2) of random forest models for target variables.

Target Variable	SET1	SET2	SET3	SET4
SPI1	0.65	0.71	0.74	0.56
SPI3	0.58	0.65	0.68	0.45
SPI6	0.56	0.67	0.69	0.45
SPI9	0.57	0.64	0.67	0.39
SPEI1	0.65	0.71	0.74	0.56
SPEI3	0.61	0.68	0.71	0.50
SPEI6	0.59	0.69	0.72	0.49
SPEI7	0.60	0.67	0.70	0.43
NDWI5	0.44	0.43	0.61	0.25
NDWI6	0.50	0.44	0.62	0.28
NDWI7	0.56	0.45	0.62	0.31
NDDI5	0.06	0.23	0.37	0.02
NDDI6	0.01	0.05	0.10	0.01
NDDI7	0.25	0.27	0.41	0.14
NMDI	0.23	0.26	0.33	0.14

가장 작은 RMSE와 큰 결정계수를 보인 모든 입력 변수를 다 포함한 집합인 SET3의 경우에 대하여 각 입력 변수별 상대적 중요도를 살펴보았다(Table 5-17). SPI, SPEI의 관측 기반의 목표 변수에 대해서는 모든 시간 척도의 목표 변수에 대해서 TRMM PRCP가 중요도가 압도적으로 큰 것으로 나타났다(Table 5-17). 이는 실제 SPI의 산정에 강수량만을 사용하며 SPEI도 강수량이 가장 중요한 변수로 쓰이기 때문에 당연한 결과라고 할 수 있다. 이 밖에 다른 변수들의 상대적 중요도를 보면, 위성 영상에서 도출한 입력 변수들 중 MODIS PET가 중요도가 가장 크며, 그 다음으로는 대규

모 기후 지수의 중요도가 매우 높은 것으로 나타났다(Table 5-17). 그리고 지면모델의 산출물인 LSM POTEVAP와 LSM SOILMOIST, LSM EVAP 의 중요도가 높게 나타났다(Table 5-17). 다른 위성 영상에서 도출한 입력 변수들의 중요도는 비슷한데, 식생지수인 EVI와 NDVI의 중요도가 그 중 가장 작은 편이다(Table 5-17).

Table 5-17. Relative importance of input variables for each target variable based on observed data(Unit: %).

Input Variable	SPI1	SPI3	SPI6	SPI9	SPEI1	SPEI3	SPEI6	SPEI9
AMSR-E A_SM	3.60	3.40	2.70	2.66	3.65	2.95	2.35	2.56
AMSR-E D_SM	3.56	4.04	3.45	2.59	3.48	4.33	3.69	2.69
AMSR-E SM_MEAN	4.15	4.06	3.68	2.95	3.80	3.76	3.53	3.22
MODIS LST_MEAN	4.69	4.43	4.11	4.49	3.97	4.51	3.76	4.81
MODIS LSTDAY	4.08	3.80	3.85	4.68	4.12	4.38	3.68	4.14
MODIS LSTNIGHT	4.42	3.99	3.78	3.10	3.46	4.03	3.44	3.12
EVI	1.58	2.03	2.54	2.50	1.92	1.95	2.15	1.93
NDVI	1.65	2.58	1.94	2.16	1.40	1.83	1.31	2.26
FPAR	2.89	2.17	2.64	2.95	2.78	2.30	2.41	2.30
LAI	2.58	2.13	2.65	2.67	2.69	2.06	2.37	2.68
MODIS ET	3.57	2.84	3.41	4.20	3.10	2.28	2.78	2.77
MODIS PET	6.48	6.43	7.27	7.95	9.14	8.11	8.48	9.11
TRMM PRCP	26.8	26.6	25.8	26.2	25.4	26.2	28.0	26.9
LSM EVAP	3.94	4.15	5.57	5.37	4.37	3.32	5.40	5.21
LSM POTEVAP	5.62	5.92	6.04	5.35	6.58	6.38	6.41	5.84
LSM SOILMOIST	5.63	6.92	5.82	5.14	5.79	6.43	5.46	4.90
AO	8.32	7.30	6.48	7.27	7.71	7.73	6.98	7.94
MEI	6.42	7.20	8.30	7.80	6.66	7.48	7.75	7.66

위성 영상에서 도출한 목표 변수에 대해서는 대부분에 대하여 대규모 기후 지수인 AO와 MEI의 상대적 중요도가 가장 큰 것으로 나타났다(Table 5-18). 선정한 시간 척도에 따라 다양하게 활용될 수 있는 SPI, SPEI 목표 변수와는 다르게 위성 영상에서 도출한 NDWI, NDDI, NMDI의 목표 변수는 모두 현 가뭄의 상태를 감시하는 동일한 목적을 가지고 있으므로 이 중 결정계수 값이 가장 큰 NDWI6를 중심으로 입력 변수들의 중요도를 살펴보았다. NDWI6의 경우 AO, MEI 다음으로 상대적인 중요도가 큰 입력 변수는 MODIS ET1, TRMM PRCP6, LSM POTEVAP1 등이다(Table 5-18). 반면, FPAR, LAI 등은 가장 작은 중요도를 나타냈다(Table 5-18).

Table 5-18. Relative importance of input variables for each target variable based on remotely sensed data (Unit: %).

Input Variable	NDWI5	NDWI6	NDWI7	NDDI5	NDDI6	NDDI7	NMDI
AMSR-E A_SM	2.39	2.34	2.46	2.24	2.54	2.34	2.79
AMSR-E D_SM	2.76	2.66	2.54	2.42	2.11	2.39	2.56
AMSR-E SM_MEAN	2.08	2.39	2.48	2.43	2.20	2.07	2.45
MODIS LST_MEAN	3.34	2.88	3.04	3.31	4.45	2.73	3.85
MODIS LSTDAY	4.47	3.70	3.67	3.74	4.72	3.97	3.56
MODIS LSTNIGHT	2.23	2.88	2.64	3.43	5.48	3.04	2.95
EVI	2.17	2.85	3.04	2.34	1.43	1.88	1.84
NDVI	1.75	2.72	3.23	1.25	0.68	1.40	1.49
FPAR	1.91	2.13	2.32	1.95	1.02	2.44	2.23
LAI	1.61	1.83	2.05	2.25	2.48	2.20	2.12
MODIS ET1	4.12	4.79	4.90	3.17	3.93	3.71	2.98
MODIS ET3	2.87	3.24	3.55	3.84	3.95	3.73	3.38
MODIS ET6	2.96	2.65	2.46	2.75	1.42	3.00	4.07
MODIS PET1	3.10	4.13	4.25	3.51	3.34	3.48	3.75

Input Variable	NDWI5	NDWI6	NDWI7	NDDI5	NDDI6	NDDI7	NMDI
MODIS PET3	3.82	4.17	4.04	3.18	1.85	3.66	3.46
MODIS PET6	3.58	3.73	4.07	4.17	5.45	4.56	4.50
TRMM PRCP1	4.22	4.02	3.98	4.10	5.76	3.61	4.03
TRMM PRCP3	3.98	3.40	3.28	4.69	4.57	4.90	5.38
TRMM PRCP6	4.69	4.41	4.36	5.35	2.40	4.79	6.22
LSM EVAP1	3.42	3.34	3.01	3.10	3.67	3.90	3.23
LSM EVAP3	2.78	3.25	3.68	2.99	2.74	3.43	3.39
LSM EVAP6	3.17	3.07	3.02	2.67	3.35	2.98	3.56
LSM POTEVAP1	3.98	4.18	3.62	3.06	6.12	3.76	3.63
LSM POTEVAP3	4.51	4.14	4.23	4.29	3.72	3.95	4.31
LSM POTEVAP6	3.80	3.41	3.99	3.10	3.93	3.02	3.10
LSM SOILMOIST	3.23	2.78	2.89	4.37	5.40	3.90	3.80
AO	8.27	7.56	6.78	8.00	6.07	8.12	5.09
MEI	8.79	7.33	6.40	8.31	5.23	7.03	6.27

결과적으로 가뭄 감시 모델은 위성 영상에서 도출한 입력 변수, 대규모 기후 지수, 지면모델 결과에서 도출한 입력 변수를 모두 포함한 SET3의 입력 변수를 사용하는 random forest 모델로 선정되었다. SPI, SPEI의 관측 자료 기반의 변수를 목표 변수로 사용하였을 때 모델은 각각의 시간 척도에 따라 기상학적 가뭄, 농업적 가뭄 등 적합한 가뭄 형태에 대해 활용할 수 있는데 특히 기상학적 가뭄 감시에 적합하다고 여겨지는 SPI1과 SPEI1에 대해서 RMSE 값이 가장 작다. 위성 영상 기반의 목표 변수는 식생의 상태에 중점을 두어 활용할 수 있는데, 이 경우 결정계수 값이 큰 NDWI6나 NDWI7 등의 목표 변수가 유용할 것으로 보인다. 목표 변수 SPI1, SPEI1, NDWI6, NDWI7의 관측 값과 random forest 모델을 통해 추정된 값을 산포도로 나타내었다(Figures 5-21 ~5-24).

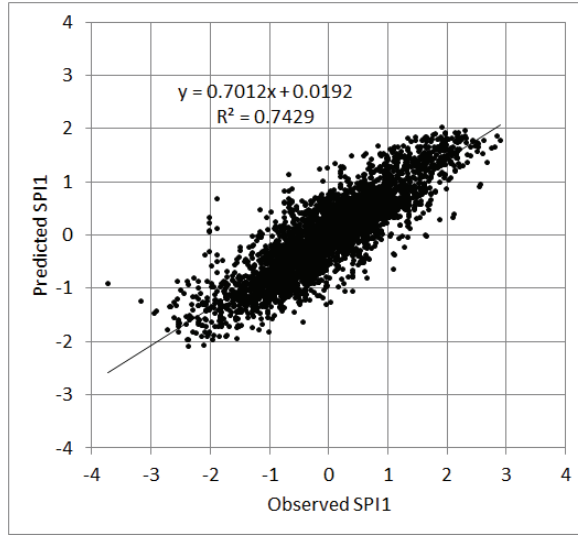


Figure 5-21. Scatterplot between observed SPI1 and estimated SPI1 based on a random forest model.

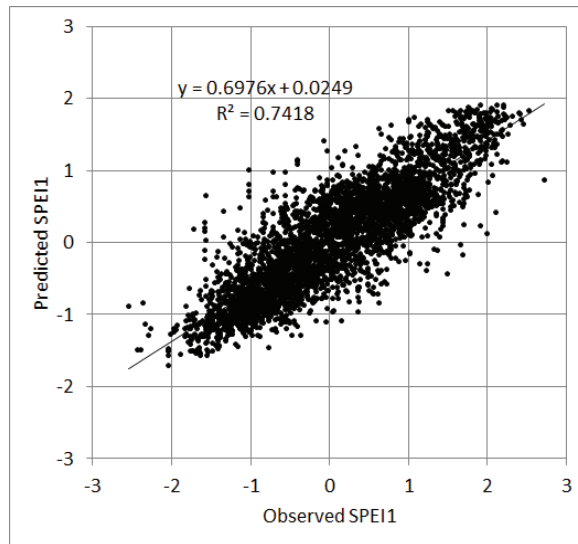


Figure 5-22. Scatterplot between observed SPEI1 and estimated SPEI1 based on a random forest model.

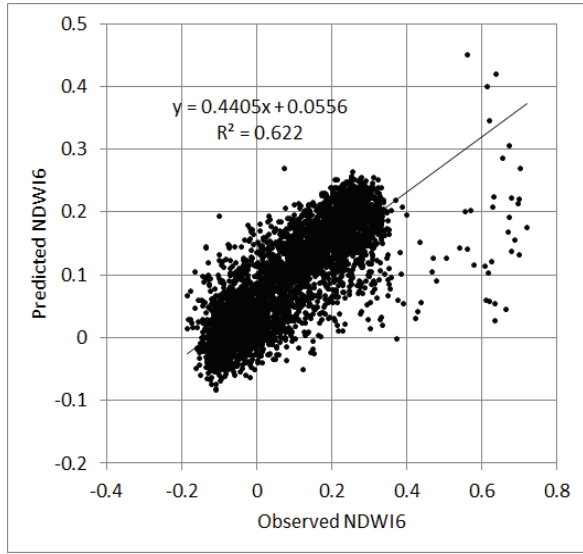


Figure 5-23. Scatterplot between observed NDWI6 and estimated NDWI6 based on a random forest model.

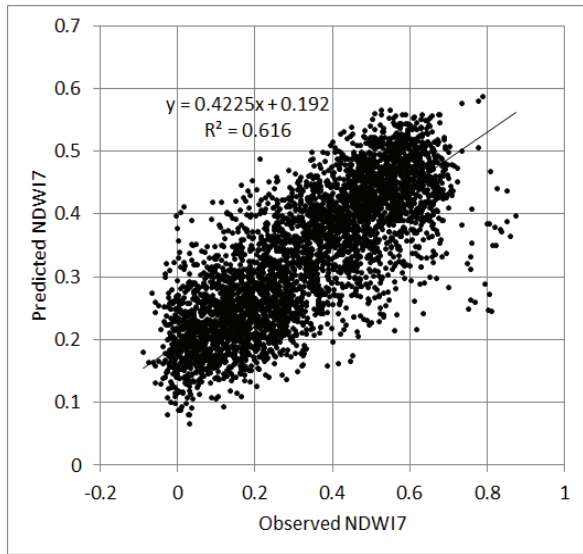


Figure 5-24. Scatterplot between observed NDWI7 and estimated NDWI7 based on a random forest model.

5.2.2 가뭄 예측 모델의 선정

가뭄 감시 모델 개발 시 가장 작은 오차 값과 가장 큰 결정계수 값을 보였던 입력 변수의 집합 SET3과 위성 영상에서 도출한 입력 변수만을 사용한 SET4에 대해서 가뭄 예측 성능을 시험하였다. 목표 변수는 가뭄 감시의 경우와 동일하게 1, 3, 6, 9개월 시간 척도의 SPI, SPEI 및 위성 영상에서 도출한 가뭄 지수인 NDWI, NDDI, NMDI를 사용하였다. 제4장 방법론에서 설명하였듯이 1개월에서 6개월까지의 선행 예측 시간($n = 1 \sim 6$)에 대해 분석하였다. 미래의 가뭄 상태에 대한 예측을 위한 모델이므로 입력 변수와 목표 변수의 과거 값이 입력 자료로 들어가게 되는데, 각각의 선행 예측 시간에 대해서 입력 변수와 목표 변수의 과거 개월 수를 1개월에서 3개월까지 다양하게 사용하여 분석하였다($p_m = 1 \sim 3$).

먼저 위성 영상에서 도출하는 입력 변수만을 사용하는 SET4의 선행 예측 시간 n 및 사용 과거 개월 수 p_m 별 RMSE를 살펴보았다. 선행 예측 시간이 길어질수록 오차 값이 점차 증가하며(Figure 5-25) 사용하는 입력 변수의 과거 개월 수가 늘어나면 오차 값이 감소하는 것으로 나타났다(Tables 5-19, 5-20).

그러나 선행 예측 시간이 1개월에서 6개월로 증가할수록 증가하는 오차 폭 보다 오히려 포함하는 과거 개월 수에 따라 RMSE 값이 크게 변화하는 것을 관찰할 수 있었다(Tables 5-19, 5-20). 위성 영상에서 도출한 목표 변수들에 대해 살펴보았을 때는 NDDI5의 경우 5개월 선행 예측 시간까지 RMSE가 점진적으로 증가하다가 6개월일 때 큰 폭으로 상승하였으며 NDDI6의 경우 RMSE가 다른 변수들에 비해 매우 크게 나타났다(Table 5-20) 이 두 변수는 안정적인 목표 변수로 사용하기에 부적합한 것으로 드러났다. 결정계수 값을 살펴보면, 선행 예측 시간이 커질수록 R^2 값이 감소하는 것을 알 수 있으며(Figure 5-26), 사용한 입력 변수 및 목표 변수의 과거 값의 개월 수가 커질수록 R^2 값이 증가한다(Tables 5-21, 5-22).

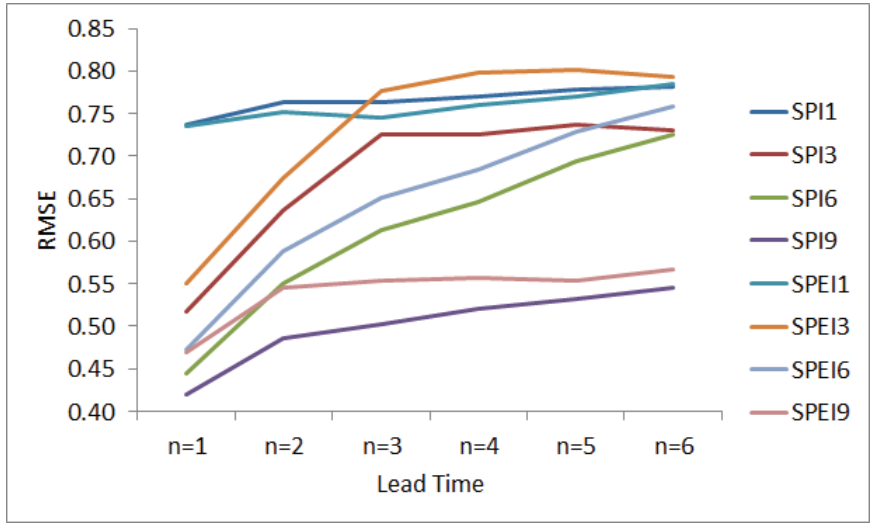


Figure 5-25. RMSE of random forest models for target variables based on observation data with various lead times and fixed number of past months (three months) used as input.

Table 5-19. RMSE of random forest models for target variables based on observation data with various lead times and numbers of past months used as input.

	Target Variable	n=1	n=2	n=3	n=4	n=5	n=6
pm = 1	SPI1	0.8590	0.8478	0.8604	0.8834	0.8523	0.8644
	SPI3	0.6136	0.7500	0.8175	0.8176	0.8200	0.8262
	SPI6	0.5186	0.6421	0.7043	0.7454	0.7827	0.8093
	SPI9	0.4747	0.5668	0.6006	0.6185	0.6314	0.6426
	SPEI1	0.8551	0.8455	0.8453	0.8714	0.8409	0.8578
	SPEI3	0.6549	0.8009	0.8623	0.8693	0.8717	0.8794
	SPEI6	0.5543	0.6881	0.7451	0.7910	0.8215	0.8485
	SPEI9	0.5215	0.6204	0.6565	0.6680	0.6692	0.6680
pm = 2	SPI1	0.7683	0.7844	0.8126	0.8016	0.7920	0.8301
	SPI3	0.5626	0.6664	0.7626	0.7577	0.7641	0.7706
	SPI6	0.4682	0.5708	0.6385	0.6820	0.7187	0.7671
	SPI9	0.4331	0.5187	0.5459	0.5592	0.5754	0.5850
	SPEI1	0.7699	0.7774	0.7946	0.7900	0.7816	0.8247
	SPEI3	0.5994	0.7078	0.8093	0.8199	0.8290	0.8291
	SPEI6	0.5054	0.6163	0.6819	0.7283	0.7551	0.8010
	SPEI9	0.4787	0.5758	0.6032	0.6064	0.6030	0.6080
pm = 3	SPI1	0.7363	0.7634	0.7631	0.7697	0.7784	0.7816
	SPI3	0.5164	0.6361	0.7257	0.7258	0.7361	0.7303
	SPI6	0.4448	0.5502	0.6126	0.6458	0.6944	0.7257
	SPI9	0.4195	0.4859	0.5024	0.5198	0.5323	0.5454
	SPEI1	0.7354	0.7510	0.7448	0.7602	0.7694	0.7845
	SPEI3	0.5496	0.6748	0.7767	0.7974	0.8017	0.7932
	SPEI6	0.4729	0.5886	0.6512	0.6838	0.7288	0.7588
	SPEI9	0.4696	0.5458	0.5530	0.5574	0.5532	0.5663

※ n: lead time, pm: number of past months used.

Table 5-20. RMSE of random forest models for target variables based on remotely sensed data with various lead times and numbers of past months used as input

	Target Variable	n=1	n=2	n=3	n=4	n=5	n=6
pm = 1	NDWI5	0.0564	0.0672	0.0701	0.0711	0.0723	0.0709
	NDWI6	0.1117	0.1339	0.1383	0.1407	0.1430	0.1413
	NDWI7	0.1296	0.1645	0.1727	0.1785	0.1812	0.1774
	NDDI5	1.7216	1.7522	1.8025	2.0011	2.0445	19.5122
	NDDI6	3.4360	3.3498	4.1827	4.4421	4.4467	5.6319
	NDDI7	0.2956	0.3256	0.3348	0.3590	0.3642	0.3600
	NMDI	0.0688	0.0783	0.0827	0.0867	0.0893	0.0887
pm = 2	NDWI5	0.0525	0.0633	0.0667	0.0682	0.0678	0.0677
	NDWI6	0.1029	0.1250	0.1314	0.1345	0.1352	0.1330
	NDWI7	0.1147	0.1495	0.1632	0.1683	0.1686	0.1651
	NDDI5	1.7511	1.7862	1.8274	2.0069	2.0693	20.3156
	NDDI6	3.2567	3.4169	4.2848	4.5435	4.5343	5.7741
	NDDI7	0.2715	0.3057	0.3233	0.3421	0.3456	0.3453
	NMDI	0.0673	0.0771	0.0820	0.0858	0.0863	0.0843
pm = 3	NDWI5	0.0519	0.0618	0.0652	0.0652	0.0652	0.0660
	NDWI6	0.1008	0.1231	0.1288	0.1299	0.1288	0.1297
	NDWI7	0.1118	0.1467	0.1587	0.1597	0.1580	0.1584
	NDDI5	1.7887	1.8176	1.8581	2.0369	2.1044	20.9565
	NDDI6	3.3341	3.4910	4.3708	4.6458	4.6300	5.9065
	NDDI7	0.2600	0.2990	0.3131	0.3298	0.3333	0.3375
	NMDI	0.0672	0.0772	0.0819	0.0842	0.0831	0.0810

※ n: lead time, pm: number of past months used.

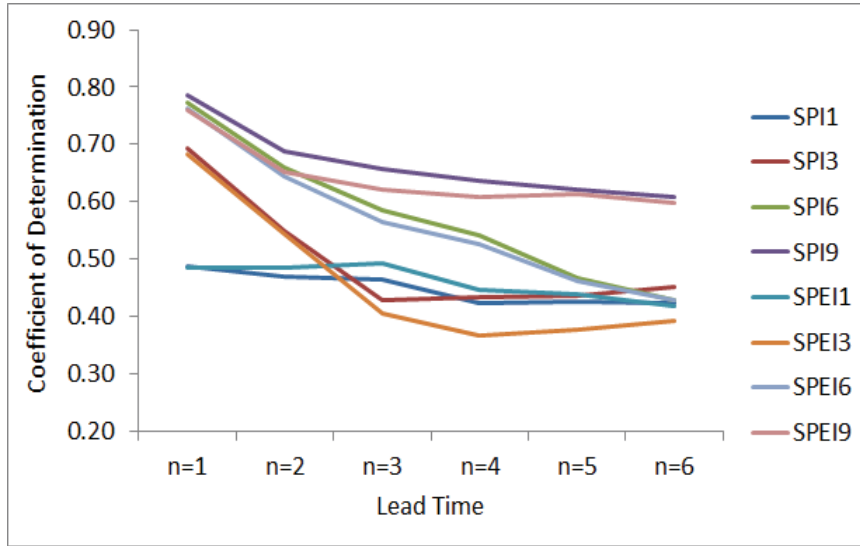


Figure 5-26. Coefficient of determination (R^2) values of random forest models for target variables based on observation data with various lead times and fixed number of past months (three months) used as input.

Table 5-21. Coefficient of determination (R^2) values of random forest models for target variables based on observation data with various lead times and numbers of past months used as input.

	Target Variable	n=1	n=2	n=3	n=4	n=5	n=6
pm = 1	SPI1	0.17	0.20	0.17	0.14	0.19	0.17
	SPI3	0.55	0.33	0.21	0.20	0.21	0.19
	SPI6	0.69	0.52	0.42	0.34	0.27	0.23
	SPI9	0.73	0.61	0.54	0.49	0.47	0.44
	SPEI1	0.16	0.19	0.19	0.13	0.19	0.16
	SPEI3	0.54	0.30	0.20	0.18	0.19	0.17
	SPEI6	0.67	0.49	0.40	0.32	0.26	0.21
	SPEI9	0.71	0.57	0.51	0.47	0.45	0.43
pm = 2	SPI1	0.38	0.38	0.35	0.35	0.35	0.30
	SPI3	0.63	0.49	0.35	0.35	0.35	0.35
	SPI6	0.75	0.63	0.54	0.47	0.41	0.33
	SPI9	0.78	0.66	0.60	0.58	0.56	0.54
	SPEI1	0.38	0.39	0.37	0.37	0.36	0.29
	SPEI3	0.62	0.48	0.33	0.31	0.30	0.31
	SPEI6	0.73	0.60	0.51	0.44	0.40	0.33
	SPEI9	0.75	0.63	0.57	0.55	0.54	0.53
pm = 3	SPI1	0.49	0.47	0.46	0.42	0.43	0.42
	SPI3	0.69	0.55	0.43	0.43	0.44	0.45
	SPI6	0.77	0.66	0.58	0.54	0.47	0.43
	SPI9	0.78	0.69	0.66	0.64	0.62	0.61
	SPEI1	0.49	0.48	0.49	0.45	0.44	0.42
	SPEI3	0.68	0.54	0.41	0.37	0.38	0.39
	SPEI6	0.76	0.64	0.57	0.53	0.46	0.43
	SPEI9	0.76	0.65	0.62	0.61	0.61	0.60

※ n: lead time, pm: number of past months used.

Table 5-22. Coefficient of determination (R^2) values of random forest models for target variables based on remotely sensed data with various lead times and numbers of past months used as input.

	Target Variable	n=1	n=2	n=3	n=4	n=5	n=6
pm = 1	NDWI5	0.47	0.26	0.23	0.22	0.21	0.24
	NDWI6	0.49	0.28	0.25	0.24	0.23	0.24
	NDWI7	0.61	0.38	0.33	0.30	0.28	0.31
	NDDI5	0.06	0.03	0.02	0.01	0.04	0.00
	NDDI6	0.06	0.01	0.00	0.00	0.00	0.00
	NDDI7	0.37	0.20	0.18	0.14	0.14	0.17
	NMDI	0.47	0.32	0.26	0.20	0.15	0.17
pm = 2	NDWI5	0.53	0.36	0.31	0.29	0.30	0.30
	NDWI6	0.56	0.39	0.34	0.33	0.32	0.33
	NDWI7	0.69	0.50	0.43	0.41	0.41	0.44
	NDDI5	0.04	0.02	0.03	0.04	0.05	0.00
	NDDI6	0.04	0.00	0.00	0.00	0.01	0.00
	NDDI7	0.44	0.27	0.24	0.21	0.22	0.24
	NMDI	0.49	0.35	0.28	0.21	0.21	0.23
pm = 3	NDWI5	0.54	0.39	0.35	0.34	0.35	0.35
	NDWI6	0.58	0.41	0.38	0.37	0.39	0.38
	NDWI7	0.70	0.52	0.47	0.48	0.50	0.49
	NDDI5	0.03	0.03	0.04	0.04	0.06	0.00
	NDDI6	0.03	0.01	0.00	0.00	0.01	0.01
	NDDI7	0.45	0.30	0.28	0.26	0.28	0.28
	NMDI	0.50	0.35	0.28	0.24	0.25	0.30

※ n: lead time, pm: number of past months used.

오차 값과 결정계수 값을 가뭄 감시 모델의 경우와 비교하였는데, 이 때 선행 예측 시간이 짧은 경우에는 가뭄 예측 모델의 RMSE 값이 가뭄 감시 모델보다 더 작은 경우도 발생하였다. 또한 SPI9, SPEI9처럼 시간 척도가 큰 목표 변수나 NDDI5, NDDI6을 제외한 다른 원격 탐사 기반의 목표 변수의 경우, 과거 3개월의 값을 입력 자료로 사용한 경우에는 6개월처럼 긴 선행 예측 시간 조건에 조차 가뭄 감시 모델에 비해 작은 RMSE 값을 보였다(Tables 5-19, 5-20). 결정계수의 경우에도 선행 예측 시간이 짧은 경우에는 가뭄 감시 모델보다 높은 결정계수 값을 보이는 경우가 많았으며, SPI9와 SPEI9의 경우와 NDDI5, NDDI6을 제외한 원격 탐사 기반의 목표 변수의 경우 RMSE에서 관찰하였듯 6개월의 선행 예측에 조차 가뭄 감시 모델에 비해 우수한 R^2 값을 보였다(Tables 5-21, 5-22).

원격 탐사 기반의 입력 변수 뿐 아니라, 대규모 지후 지수와 지면모델 산출물을 모두 입력 변수 집합에 포함시킨 SET3에 대해서도 가뭄 예측 모델을 random forest를 이용하여 구축하였다. 원격 탐사 기반의 입력 변수만을 포함시킨 SET4와 비교하여, 가뭄 예측 모델이 개선되었음을 알 수 있었다. RMSE 값을 살펴보면 모든 목표 변수에 대해 대부분의 경우 가뭄 예측 모델의 RMSE가 SET4에 비해 감소하였으며, 가뭄 감시 모델과 비교하는 경우에도 선행 예측 시간이 6개월이 되는 경우에도 더 작은 RMSE 값을 보였다(Tables 5-23, 5-24). 입력 변수 및 목표 변수의 과거 2개월 또는 3개월의 값을 입력 자료로 사용하는 경우는 모든 경우에서 가뭄 감시 모델보다 작은 RMSE를 보였으며, 1개월의 값을 입력 자료로 사용하는 경우에도 SPI1과 SPEI1을 제외하고는 선행 예측 시간 6개월까지 더 작은 RMSE를 보였다(Tables 5-23, 5-24).

결정계수의 경우, 선행 예측 시간이 증가할수록 값이 감소하였으나 SET4와 비교할 때 모든 경우에 매우 높은 결정계수 값을 보였다(Tables 5-25, 5-26). NDDI5와 NDDI6가 가뭄 예측을 제대로 수행하지 못하는 경우를 제외하면, 단지 1개월의 과거 값을 입력 자료로 사용한 경우에도 6개월의 선행 예측 시간까지 상대적으로 높은 결정계수 값을 생산하여, 가뭄 감시 모델보다 높은 예측력을 보였다(Tables 5-25, 5-26).

Table 5-23. RMSE of random forest models for target variables based on observation data with various lead times and numbers of past months used as input.

	Target Variable	n=1	n=2	n=3	n=4	n=5	n=6
pm = 1	SPI1	0.6437	0.6336	0.6597	0.6208	0.6271	0.6479
	SPI3	0.5283	0.6436	0.6648	0.6575	0.6465	0.6431
	SPI6	0.4538	0.55	0.5806	0.6214	0.65	0.6541
	SPI9	0.4504	0.5252	0.5443	0.562	0.5799	0.5908
	SPEI1	0.6453	0.6297	0.6518	0.6139	0.6201	0.6332
	SPEI3	0.569	0.6705	0.6904	0.6853	0.6746	0.6646
	SPEI6	0.4848	0.5754	0.6098	0.6496	0.6695	0.6634
	SPEI9	0.4999	0.5755	0.5873	0.5844	0.604	0.6115
pm = 2	SPI1	0.5438	0.5596	0.5577	0.547	0.5565	0.5732
	SPI3	0.4974	0.5834	0.594	0.5855	0.579	0.5804
	SPI6	0.4197	0.4992	0.5406	0.5821	0.5939	0.6141
	SPI9	0.4263	0.4938	0.5081	0.5219	0.5416	0.549
	SPEI1	0.549	0.5598	0.5524	0.5452	0.5499	0.5624
	SPEI3	0.5267	0.6053	0.616	0.6019	0.5852	0.5876
	SPEI6	0.4446	0.5213	0.5659	0.6057	0.6042	0.6159
	SPEI9	0.4749	0.5398	0.5392	0.5489	0.5559	0.5672
pm = 3	SPI1	0.5162	0.5175	0.5243	0.5199	0.5262	0.5334
	SPI3	0.4789	0.5594	0.561	0.5488	0.5509	0.5499
	SPI6	0.4003	0.4823	0.5268	0.5561	0.57	0.5852
	SPI9	0.4271	0.4729	0.4921	0.5012	0.5139	0.523
	SPEI1	0.5234	0.5192	0.5223	0.5157	0.519	0.5226
	SPEI3	0.5038	0.5774	0.5734	0.5545	0.5517	0.5563
	SPEI6	0.4199	0.5024	0.5484	0.5729	0.5743	0.5741
	SPEI9	0.4739	0.5107	0.5154	0.5139	0.5268	0.533

※ n: lead time, pm: number of past months used.

Table 5-24. RMSE of random forest models for target variables based on remotely sensed data with various lead times and numbers of past months used as input.

	Target Variable	n=1	n=2	n=3	n=4	n=5	n=6
pm = 1	NDWI5	0.0394	0.0431	0.0447	0.0454	0.0467	0.0450
	NDWI6	0.0801	0.0900	0.0939	0.0960	0.0993	0.0973
	NDWI7	0.0959	0.1159	0.1234	0.1294	0.1341	0.1298
	NDDI5	0.5884	0.6212	0.7074	0.8694	1.0265	22.3050
	NDDI6	2.1324	2.1350	2.1481	2.2078	2.2187	4.5410
	NDDI7	0.2054	0.2199	0.2170	0.2322	0.2371	0.2422
	NMDI	0.0529	0.0589	0.0604	0.0624	0.0643	0.0644
pm = 2	NDWI5	0.0365	0.0407	0.0419	0.0435	0.0433	0.0421
	NDWI6	0.0732	0.0832	0.0863	0.0900	0.0905	0.0878
	NDWI7	0.0859	0.1047	0.1121	0.1183	0.1176	0.1130
	NDDI5	0.5500	0.6127	0.6937	0.8603	1.0065	22.8468
	NDDI6	2.1367	2.1409	2.1795	2.1993	2.2456	4.6314
	NDDI7	0.1948	0.2040	0.2044	0.2135	0.2234	0.2274
	NMDI	0.0510	0.0559	0.0578	0.0598	0.0609	0.0599
pm = 3	NDWI5	0.0357	0.0401	0.0418	0.0419	0.0412	0.0409
	NDWI6	0.0709	0.0812	0.0858	0.0859	0.0837	0.0842
	NDWI7	0.0838	0.1011	0.1094	0.1085	0.1060	0.1048
	NDDI5	0.5459	0.6127	0.6920	0.8567	1.0236	23.6171
	NDDI6	2.1674	2.1942	2.2192	2.2266	2.2635	4.7508
	NDDI7	0.1900	0.1992	0.1975	0.2045	0.2116	0.2206
	NMDI	0.0500	0.0548	0.0564	0.0583	0.0573	0.0565

※ n: lead time, pm: number of past months used.

Table 5-25. Coefficient of determination (R^2) values of random forest models for target variables based on observation data with various lead times and numbers of past months used as input.

	Target Variable	n=1	n=2	n=3	n=4	n=5	n=6
pm = 1	SPI1	0.61	0.61	0.58	0.61	0.62	0.59
	SPI3	0.68	0.54	0.52	0.52	0.56	0.55
	SPI6	0.77	0.68	0.65	0.59	0.56	0.55
	SPI9	0.78	0.69	0.66	0.62	0.59	0.58
	SPEI1	0.59	0.61	0.58	0.61	0.61	0.60
	SPEI3	0.68	0.56	0.55	0.54	0.57	0.58
	SPEI6	0.76	0.68	0.64	0.59	0.57	0.59
	SPEI9	0.76	0.67	0.65	0.64	0.60	0.58
pm = 2	SPI1	0.69	0.69	0.68	0.69	0.69	0.68
	SPI3	0.72	0.63	0.61	0.63	0.64	0.64
	SPI6	0.81	0.74	0.70	0.65	0.64	0.60
	SPI9	0.80	0.73	0.69	0.67	0.65	0.64
	SPEI1	0.68	0.69	0.69	0.69	0.69	0.68
	SPEI3	0.72	0.65	0.63	0.65	0.67	0.67
	SPEI6	0.80	0.74	0.70	0.66	0.66	0.64
	SPEI9	0.78	0.71	0.70	0.67	0.66	0.65
pm = 3	SPI1	0.72	0.72	0.71	0.71	0.72	0.72
	SPI3	0.75	0.66	0.65	0.67	0.67	0.67
	SPI6	0.83	0.76	0.72	0.69	0.66	0.64
	SPI9	0.80	0.73	0.71	0.70	0.69	0.68
	SPEI1	0.71	0.72	0.71	0.71	0.72	0.72
	SPEI3	0.75	0.68	0.68	0.69	0.70	0.69
	SPEI6	0.83	0.76	0.72	0.69	0.69	0.69
	SPEI9	0.79	0.73	0.72	0.71	0.70	0.69

※ n: lead time, pm: number of past months used.

Table 5-26. Coefficient of determination (R^2) values of random forest models for target variables based on remotely sensed data with various lead times and numbers of past months used as input.

	Target Variable	n=1	n=2	n=3	n=4	n=5	n=6
pm = 1	NDWI5	0.66	0.60	0.58	0.58	0.54	0.57
	NDWI6	0.70	0.63	0.61	0.60	0.56	0.57
	NDWI7	0.78	0.70	0.67	0.64	0.61	0.62
	NDDI5	0.55	0.48	0.39	0.31	0.30	0.00
	NDDI6	0.11	0.08	0.10	0.08	0.06	0.01
	NDDI7	0.52	0.43	0.44	0.40	0.37	0.35
	NMDI	0.55	0.44	0.42	0.39	0.36	0.36
pm = 2	NDWI5	0.70	0.64	0.62	0.59	0.59	0.62
	NDWI6	0.75	0.68	0.67	0.63	0.62	0.65
	NDWI7	0.82	0.75	0.73	0.69	0.69	0.71
	NDDI5	0.56	0.47	0.40	0.31	0.34	0.00
	NDDI6	0.10	0.08	0.09	0.08	0.06	0.01
	NDDI7	0.55	0.49	0.49	0.44	0.39	0.40
	NMDI	0.58	0.50	0.46	0.43	0.41	0.42
pm = 3	NDWI5	0.71	0.65	0.62	0.61	0.63	0.64
	NDWI6	0.76	0.69	0.66	0.65	0.67	0.68
	NDWI7	0.83	0.77	0.73	0.74	0.75	0.75
	NDDI5	0.55	0.46	0.39	0.31	0.33	0.00
	NDDI6	0.08	0.06	0.07	0.08	0.08	0.02
	NDDI7	0.55	0.49	0.48	0.44	0.43	0.43
	NMDI	0.59	0.51	0.47	0.44	0.45	0.48

※ n: lead time, pm: number of past months used.

결과적으로 NDDI5, NDDI6을 제외한 나머지 목표 변수는 모두 비교적 우수한 가뭄 예측 성능을 보였다. 특히 입력 변수 및 목표 변수의 과거 3개월 값을 입력 자료로 사용한 경우($pm = 3$) 작은 오차 값과 높은 결정계수 값을 보였으며, 이러한 결과는 1개월에서 6개월까지의 선행 예측 시간의 경우에 모두 해당하였다.

가장 우수한 예측 성능을 보인 SET3의 과거 3개월 값을 입력 자료로 사용한 경우에 대하여, 선행 예측 시간 1개월과 6개월에 대해 입력 변수 간 상대적 중요도를 살펴 보았다(Table 5-27; NDWI6에 대한 값만 나타냄). 사용된 변수들의 과거 기간 중에 2개월 전($t-2$)과 1개월 전($t-1$)에 비해서 현재 달(t)의 중요도가 더 높을 것으로 생각할 수 있으나, 이는 변수마다 다르게 나타났다. 목표 변수 NDWI6 예측에 있어 선행 예측 시간 6개월의 경우 목표변수인 NDWI6의 현재 달 값의 중요도가 가장 크게 나타났으며, 이어 대규모 지수인 AO의 1개월 전($t-1$)과 2개월 전($t-2$) 값, NDWI6의 1개월 전 값 등의 중요도가 높다. 반면 식생 관련된 변수인 LAI, FPAR 등의 중요도가 가장 낮게 나타났다(Table 5-17). 선행 예측 시간 1개월의 경우에는 현재 달, 1개월 전, 2개월 전 NDWI6의 중요도가 압도적으로 크게 나타났으며 AO와 MEI가 뒤따랐다. 또한 FPAR, LAI 뿐 아니라 NDVI, EVI의 중요도가 특히 낮게 나타났다.

Table 5-27. Relative importance of input variables for each target variable based on remotely sensed data (Unit: %).

Lead Time Input Variable	n = 1			n = 6		
	t	t-1	t-2	t	t-1	t-2
AMSR-E A_SM	0.53	0.70	0.66	0.58	1.09	0.85
AMSR-E D_SM	0.63	0.36	0.71	0.95	0.86	0.59
AMSR-E SM_MEAN	0.71	0.68	0.82	0.76	0.94	0.43
MODIS LST_MEAN	1.22	1.16	0.79	1.43	1.42	1.27
MODIS LSTDAY	1.35	1.21	1.00	1.44	1.71	1.35
MODIS LSTNIGHT	0.85	0.85	0.69	0.79	1.08	0.91
EVI	0.65	0.46	0.18	0.44	0.63	0.56
NDVI	0.54	0.40	0.11	0.59	0.54	0.52
FPAR	0.51	0.56	0.29	0.60	0.59	0.40
LAI	0.55	0.19	0.23	0.47	0.31	0.32
MODIS ET1	1.12	0.70	0.82	0.74	1.12	0.96
MODIS ET3	0.80	0.72	0.65	0.94	0.80	0.93
MODIS ET6	1.03	0.60	0.48	0.69	0.99	1.04
MODIS PET1	1.20	1.55	0.91	1.07	1.25	1.27
MODIS PET3	1.05	0.93	0.94	1.15	1.13	1.23
MODIS PET6	1.04	1.32	1.22	0.74	0.76	1.17
TRMM PRCP1	1.45	1.50	1.73	1.50	1.71	1.21
TRMM PRCP3	1.25	1.15	1.18	1.59	1.17	1.07
TRMM PRCP6	1.37	1.37	0.98	1.12	1.36	1.16
LSM EVAP1	0.73	0.75	0.79	0.80	0.93	1.04
LSM EVAP3	0.70	0.73	0.66	0.87	0.82	0.81
LSM EVAP6	0.45	0.49	0.63	0.63	0.76	0.83
LSM POTEVAP1	1.07	1.43	1.04	1.51	1.29	1.15
LSM POTEVAP3	1.17	1.07	1.00	1.22	1.10	0.87
LSM POTEVAP6	0.93	0.97	1.06	1.11	1.21	1.10
LSM SOILMOIST	1.24	1.22	0.81	1.14	1.08	0.78
AO	2.74	2.09	2.34	2.52	2.85	2.80
MEI	2.59	2.11	2.30	1.90	2.44	1.97
NDWI6	10.51	3.72	4.01	5.24	2.59	2.35

※ t: current month.

5.3 가뭄 예측 모델의 검증

5.3.1 고랭지 채소 수확량

제4장에서 설명한 대로 자료 개수의 부족으로 수확량을 예측하는 모델을 직접 구축하기는 쉽지 않으므로, 고랭지 채소 수확량과 가장 밀접한 상관관계를 가지는 원격 탐사 기반의 목표 변수를 선정하여 대신 이 변수에 대한 예측 성능을 시험하기로 하였다. 준고랭지 및 고랭지에 위치한 5×5km 격자에 대해서 SET3의 입력 변수 집합에 해당하는 입력 변수 값을 추출하였으며, 이를 각 시도, 시군구에 대하여 각각 평균하였다. 이렇게 얻어진 준고랭지 및 고랭지 지역을 포함하는 시도 및 시군구 시계열 입력 자료를 NDDI5, NDDI6을 제외한 원격 탐사 기반 목표 변수와 회귀하여 상관관계를 분석하였다. 네 개의 시도와 강원도에 위치한 18개의 시군구 별로 회귀 분석을 수행하였는데, 이때 각각의 표본 개수는 시도의 경우 9개, 시군구의 경우에는 6개가 된다(Table 5-28).

표본개수가 매우 작으므로 비모수 순위 상관계수인 켄달 τ 값을 이용하였는데, 많은 경우 상관관계가 낮게 나타났다. 이는 제4장에서 서술하였듯이 비가 너무 많이 오는 경우 생장이 저해되는 등 시기별로 여러 요인이 다르게 작용하기 때문이다. 본 검증에서는 단위당 수확량이 목표 변수와 높은 상관관계를 가지는 경우에 대해서만($\tau > 0.5$ 이거나 $\tau < 0.5$ 이고 $p\text{-value} < 0.05$) 앞서 구축된 가뭄 예측 모델을 적용해 보았다. 일반적으로 NDVI, NDWI 등의 지수는 가뭄 상태와 부(-)의 상관관계를 가지며(즉, 가뭄이 심화되면 NDVI, NDWI 등의 값은 작아짐), NDDI 등의 지수는 가뭄 상태와 정(+)의 상관관계를 가지고, NMDI는 지역의 습윤한 정도에 따라 다르게 나타나는 것으로 알려져 있다(Rhee et al., 2010).

앞 절에서 구축한 가뭄 예측 모델 중 과거 3개월의 입력 변수 및 목표 변수를 입력 자료로 이용하는 모델($pm = 3$)에 아래 22가지의 작물, 지역과 목표 변수 조합에 대해 시도 또는 시군구 별로 시계열 입력 자료를 구축하여 대입, 고랭지 채소 수확량과 상관관계가 큰 대리변수(원격 탐사 기반의 가뭄 지수; Table 5-28의 목표 변수들)를 통한 고랭지 채소 수확량 예측 성능을 시험하였다.

1개월에서 6개월까지의 선행 예측 시간에 대해서 앞서 구축된 random forest 가뭄

예측 모델에 시도 및 시군구별 입력 변수 자료를 대입하여 결과물을 산출하였고, 같은 행정구역에 대한 목표 변수 값과 비교하여 모델의 예측력을 켄달의 순위 상관계수를 통해 살펴보았다(Table 5-28). 결과 중 $\tau > 0.5$ 이거나 $\tau < 0.5$ 이고 5% 유의수준에서 통계적으로 유의미한 경우(p-value < 0.05)만을 살펴보았다.

Table 5-28. Statistics between crop yield and target variable data(Unit for RMSE and MAE is ka/10ha).

CROP	CASE	SI-DO or SI-KUN-KU	Target Variable	Month	n	RMSE	MAE	tau	tau_p	Crop yield range (min~max)
BOK CHOY	1	Jeollabuk-do	NDWI5	9	9	4014	4012	0.56	0.037	3831~4281
	2	Gyeongsangnam-do	NDWI7	8	9	4199	4181	0.56	0.037	3792~5088
	3	Gyeongsangnam-do	NDDI7	9	9	4200	4181	-0.56	0.037	3792~5088
	4	Gangneung-si	NDDI7	7	6	3516	3484	-0.73	0.039	2640~4004
	5	Hoengseong-gun	NDWI7	6	6	3501	3474	-0.83	0.020	2644~3886
	6	Yeongwol-gun	NDWI6	7	6	3597	3566	0.83	0.020	2644~3976
	7	Yeongwol-gun	NDWI7	7	6	3597	3566	0.83	0.020	2644~3976
	8	Yanggu-gun	NDDI7	7	6	3319	3262	0.83	0.020	2469~3887
	9	Inje-gun	NDWI6	9	6	3835	3797	0.83	0.020	2643~4230
	10	Inje-gun	NMDI	5	6	3835	3797	0.83	0.020	2643~4230
RADISH	11	Gangwon-do	NDDI7	5	9	2917	2908	-0.56	0.037	2484~3339
	12	Gyeongsangbuk-do	NDWI5	7	9	2973	2912	-0.61	0.022	2302~4342
	13	Gyeongsangbuk-do	NDWI6	7	9	2973	2912	-0.67	0.012	2302~4342
	14	Gyeongsangbuk-do	NMDI	7	9	2973	2912	-0.61	0.022	2302~4342
	15	Gyeongsangnam-do	NDWI7	6	9	3903	3873	0.56	0.037	2986~4617
	16	Gyeongsangnam-do	NDDI7	6	9	3903	3873	-0.67	0.012	2986~4617
	17	Gangneung-si	NDDI7	5	6	2877	2865	-0.83	0.020	2490~3310
	18	Taebaek-si	NDWI6	6	6	2882	2872	0.73	0.039	2483~3239
	19	Taebaek-si	NDWI7	6	6	2882	2872	0.87	0.015	2483~3239
	20	Samcheok-si	NMDI	9	6	3452	3317	0.73	0.039	2483~5380
	21	Yeongwol-gun	NDWI7	6	6	2754	2744	0.73	0.039	2433~3106
	22	Inje-gun	NDWI5	5	6	2894	2887	0.83	0.020	2483~3093

앞서 목표변수와 수확량의 상관관계를 보았을 때는 고랭지 배추에 대해서는 시도로는 전라북도와 경상남도 두 곳, 시군구로는 강릉시, 횡성군, 영월군, 양구군, 인제군의 다섯 곳에서 수확량이 원격 탐사 기반 가뭄 지수와 높은 상관관계를 보였으며, 고랭지 무에 대해서는 시도로는 강원도와 경상북도, 경상남도 세 곳, 시군구로는 강릉시, 태백시, 삼척시, 영월군, 인제군의 다섯 곳이 높은 상관관계를 보였다(Table 5-28). Random forest 가뭄 예측 모델에 따른 해당 목표 변수의 예측 결과(Table 5-29)를 함께 살펴 보면, 선행 예측 시간이나 목표 변수, 해당 월에 대해서 어떤 특정한 경향을 이끌어내는 것이 어려웠다. 고랭지 배추의 경우 순위 상관계수 값을 중심으로 살펴보면, 특히 시군구 단위로 영월군의 NDWI7이 6개월 선행 예측 시간에 대해 예측 성능이 가장 뛰어났으며(Figure 5-17), 경상남도의 NDWI7이 1개월 선행 예측 시간에 대해서, 영월군의 NDWI6이 2개월과 6개월 선행 예측 시간에 대해, 인제군의 NDWI6가 1개월 선행 예측 시간에 대해 잘 예측되었다(Table 5-29). 고랭지 무의 경우, 경상남도에서 NDWI6가 6개월 선행 예측 시간에 대해 예측 성능이 뛰어났으며 경상북도에서 NDWI6가 2개월 선행 예측 시간에 대해 수확량과 높은 상관관계를 보였다(Table 5-29).

Table 5-29. Statistics between observed and estimated target variable data (RMSE and MAE are unitless).

Lead Time	CASE	n	RMSE	MAE	tau	tau_p
4	1	10	0.0365	0.0214	0.64	0.009
5	1	10	0.0238	0.0191	0.56	0.025
1	2	9	0.0258	0.0230	0.78	0.004
5	2	9	0.0526	0.0473	0.61	0.022
1	3	10	0.0150	0.0134	0.60	0.016
2	3	10	0.0714	0.0607	0.60	0.016
3	3	10	0.0955	0.0845	0.69	0.006
4	3	10	0.0517	0.0440	0.69	0.006
3	4	9	0.0408	0.0374	0.72	0.007
2	6	9	0.0535	0.0520	0.78	0.004
3	6	9	0.0257	0.0190	0.67	0.012
6	6	9	0.0519	0.0408	0.78	0.004
3	7	9	0.0473	0.0416	0.56	0.037
6	7	9	0.0271	0.0200	0.83	0.002
3	8	9	0.0247	0.0202	0.72	0.007
4	8	9	0.2967	0.2895	0.61	0.022
1	9	10	0.0667	0.0655	0.78	0.002
6	11	9	0.1102	0.0941	0.61	0.022
1	12	9	0.0063	0.0047	0.61	0.022
6	12	9	0.0241	0.0202	0.56	0.037
2	13	9	0.0566	0.0555	0.78	0.004
3	13	9	0.0563	0.0529	0.72	0.007
6	13	9	0.0204	0.0179	0.56	0.037
1	14	9	0.0095	0.0080	0.72	0.007
3	14	9	0.0181	0.0145	0.56	0.037
6	14	9	0.0135	0.0106	0.61	0.022
3	16	9	0.0277	0.0265	0.72	0.007
5	16	9	0.0519	0.0445	0.67	0.012
6	16	9	0.0740	0.0624	0.83	0.002
1	18	9	0.0373	0.0363	0.56	0.037
4	21	9	0.0707	0.0664	0.56	0.037

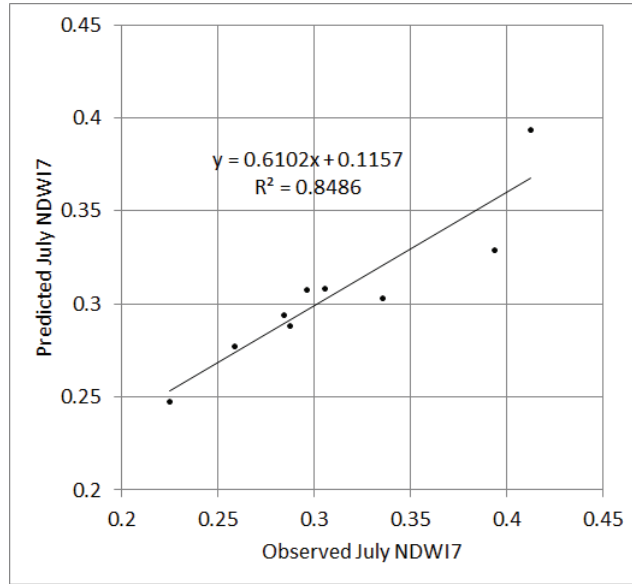


Figure 5-27. Scatterplot between observed July NDWI7 and estimated July NDWI7 based on a random forest model with 6-month lead time in Yeongwol-gun.

5.3.2 산불 발생 빈도 및 피해 면적

고랭지 채소 수확량과 마찬가지로 산불 관련 변수 값을 예측하는 모델을 직접 구축하지 않고 이들과 가장 밀접한 상관관계를 가지는 원격 탐사 기반의 목표 변수를 대리 변수로 선정하여 해당 행정구역에서의 이 변수에 대한 예측 성능을 시험하기로 하였다. 산지 지역에 위치한 5×5km 격자에 대해서 SET3의 입력 변수 집합에 해당하는 입력 변수 값을 추출하였으며, 이를 각 시도, 시군구에 대하여 각각 평균하였다. 그리고 시도 또는 시군구의 월별 산불 발생 빈도(frequency), 산불 발생으로 인한 피해 면적(area), 해당 행정구역의 산지 면적에 대한 피해 면적의 비율(ratio)를 NDDI5, NDDI6을 제외한 원격 탐사 기반의 가뭄 지수와 회귀 분석하여 상관관계를 살펴보았다.

산불 발생 빈도(frequency), 산불 발생으로 인한 피해 면적(area), 해당 행정구역의 산지 면적에 대한 피해 면적의 비율(ratio) 중 오직 산불 발생 빈도만이 원격 탐사 기반

가뭄 지수와 높은 상관관계를 보였다(Table 5-30). 산불 피해 면적이나 산지 면적 대비율보다 발생 빈도만이 가뭄 지수와 밀접한 관계를 가지는 것은, 산불 피해 면적은 일단 산불이 발생하고 난 뒤 산불의 조기 발견과 진화가 어떻게 이루어지는지에 영향을 많이 받는데, 산불 발생 빈도는 산불이 발생하기 위한 건조한 조건이 많은 영향을 미치므로 가뭄과 상관관계가 높기 때문이다. 행정 구역으로는 시군구 단위로는 높은 상관관계를 관찰할 수 없었으며, 오로지 시도 단위에서만 높은 상관관계를 볼 수 있었다. 가뭄 지수도 NDWI와 NDDI 만이 상관관계가 좋은 것으로 드러났다(Table 5-30).

앞 표에 나열된 여섯 곳의 시도와 각각의 행정구역에서 산불 발생 빈도와 밀접한 상관관계를 보이는 목표 변수에 대해서, 이미 앞 절에서 구축된 가뭄 예측 모델 중 과거 3개월의 입력 변수 및 목표 변수를 입력 자료로 이용하는 모델($pm = 3$)에 시도 또는 시군구 별로 시계열 입력 자료를 구축하여 대입하여 해당 행정구역에서 목표 변수가 잘 예측될 수 있는지 모델의 예측 성능을 시험하였다. 위 13가지 조합에 대하여 1개월에서 6개월까지의 선행 예측 시간에 대해 random forest 가뭄 예측 모델은 모두 매우 높은 상관관계를 나타냈다(모든 경우에 대하여 $r > 0.5$, $p\text{-value} < 0.001$; 전체 자료 미삽입; 경상남도 지역의 NDWI7 예측에 대해 Table 5-31, Figure 5-28). 이 중 굳이 상관관계가 상대적으로 낮은 경우를 꼽으면, 전라남도 지역에서 NDDI7을 예측하는 경우 모든 선행 예측 시간에서 r 값이 0.51~65로 나타났으며, 다른 모든 경우에는 $r > 0.89$ 를 보였다(Table 5-31).

Table 5-30. Statistics between forest fire frequency and target variable data (Unit for RMSE and MAE is event/month, all p-value < 0.001).

CASE	SIDO	Target Variable	n	RMSE	MAE	Frequency range (min~max)	r
1	Busan	NDWI7	120	2.95	1.74	0~24	-0.55
2	Busan	NDWI6	120	3.05	1.73	0~11	-0.54
3	Ulsan	NDWI7	120	2.26	1.39	0~34	-0.51
4	Chungcheongbuk-do	NDDI7	120	3.99	2.14	0~24	0.52
5	Jeollanam-do	NDWI7	120	7.96	3.97	0~11	-0.55
6	Jeollanam-do	NDWI6	120	8.07	3.98	0~13	-0.52
7	Jeollanam-do	NDDI7	120	7.96	3.90	0~34	0.56
8	Gyeongsangbuk-do	NDWI7	120	10.09	5.93	0~44	-0.51
9	Gyeongsangbuk-do	NDDI7	120	10.06	5.86	0~24	0.53
10	Gyeongsangnam-do	NDWI7	120	5.91	3.69	0~17	-0.57
11	Gyeongsangnam-do	NDWI6	120	6.06	3.78	0~34	-0.56
12	Gyeongsangnam-do	NDWI5	120	6.14	3.82	0~44	-0.51
13	Gyeongsangnam-do	NDDI7	120	5.92	3.64	0~24	0.55

Table 5-31. Statistics between observed and estimated NDWI7 in Gyeongsangnam-do (RMSE and MAE are unitless; all p-value < 0.001).

Lead Time	n	RMSE	MAE	r
1	109	0.0372	0.0292	0.98
2	109	0.0494	0.0394	0.98
3	109	0.0613	0.0475	0.98
4	109	0.0563	0.0461	0.98
5	109	0.0523	0.0431	0.98
6	109	0.0500	0.0427	0.99

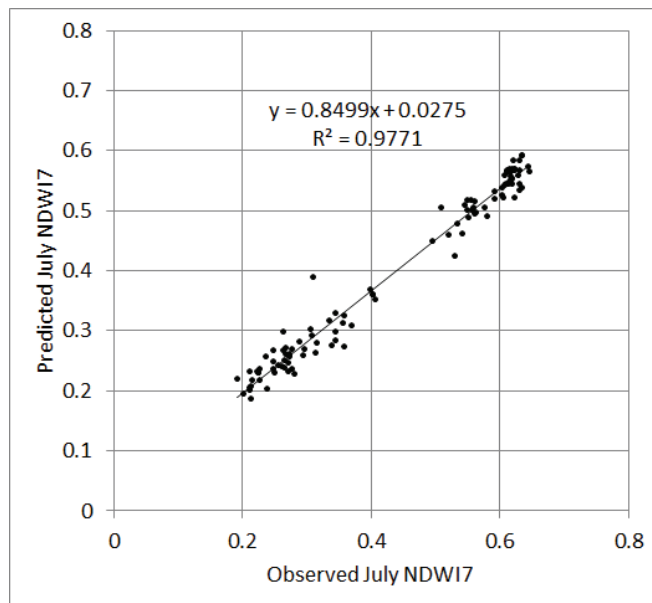


Figure 5-28. Scatterplot between observed NDWI7 and estimated NDWI7 for all months based on a random forest model with 6-month lead time in Gyeongsangnam-do.

6. 결론 및 시사점

6.1 연구 결과의 요약

여러 자연재난 가운데 사망자 수는 적으나 실제로 피해를 입는 인구수와 경제적인 피해가 막대한 재난인 가뭄은 현재까지도 꾸준히 발생하였으나 앞으로는 기후변화로 인하여 가뭄 발생의 공간적인 범위 및 시간적인 빈도 등의 불확실성이 증가하고 있는 상황이다. 먼 미래의 가뭄 발생을 전망하기는 매우 어렵고 이에 대한 여러 대책을 미리 세우는 일 또한 실천하기 쉽지 않지만, 현재의 가뭄 상태를 정확히 파악하고 몇 개월 이후의 가뭄 진행 상황을 예측하는 것은 가뭄으로 인한 피해 발생을 최소화하는데 매우 필요하다. 이에 가뭄 감시 및 예측을 수행, 정확하고 시의적절한 정보를 의사결정자에게 전달하여 정확하고 신속한 결정을 내리도록 도와주어 결과적으로 가뭄으로 인한 피해를 최소화할 수 있게 하는 가뭄조기경보시스템이 필수적이다.

가뭄의 감시와 예측을 위해서는 여러 기상 및 수문 변수들의 값이 필요한데, 이러한 자료의 관측 값은 일반적으로 기상 관측 지점에서만 얻을 수 있다. 따라서 관측 지점이 조밀하게 분포하지 않거나 현장 관측 자료가 부재하는 경우 위성 영상 자료가 유용하게 쓰이고 있다. 본 연구의 목적은 기상 관측 자료가 조밀하지 못한 지역에 대하여 위성 영상 자료를 활용하여 가뭄을 감시하고 예측하는 기술을 개발하는 것이다. 이를 위해 한반도 지역을 대상으로 여러 인공 위성 영상으로부터 강수, 토양수분, 지표면온도, 식생지수, 실제 및 잠재증발산량, 유효광합성반사율, 엽면적지수 등의 자료를 도출하였으며, 여기에 지면모델의 산출물인 증발산량과 토양수분 자료와 대규모 기후 지수를 함께 고려하여 가뭄 감시 및 예측 모델을 구축하였다.

전통적인 가뭄 감시는 SPI, PDSI 등의 가뭄 지수를 통해 관측 지점에서 이루어지는데, 본 연구에서는 관측 기간이 상대적으로 짧은 위성 영상 자료의 이용을 위해서 지식 기반 전문가 시스템을 훈련시키는 기계학습을 활용하였다. 가뭄과 관련 있는 수문 및 기상 변수들은 서로간의 커플링과 피드백으로 긴밀하게 연결되어 있는데, 이러한 변수들의

상호 작용을 단순한 알고리즘으로 설명하기 어려우며 종종 일부 변수만을 감시하여 부정확한 결론을 내리기도 한다. 특히 원격 탐사로 얻은 위성 영상 자료에서 도출한 수문 기상변수는 많은 경우 추정 알고리즘에 의해 도출되므로 관측 자료보다 정확하지 못할 수 있다. 이런 경우에 변수들 간의 물리적 관계를 명시하거나 여러 기본 가정을 할 필요가 없는 기계학습은 매우 적합한 방법론으로, 가뭄 감시 및 예측 모델 구축에 효과적으로 쓰일 수 있다.

본 연구에서는 기계학습 방법 중 random forest 방법론을 이용하여 가뭄 감시 및 선행 예측 시간 1개월에서 6개월까지의 예측 모델을 구축하였다. Random forest 모델은 이상치와 노이즈에 강한 것으로 알려져 있는데, 입력 변수로는 토양수분(주간, 야간, 평균), 지표면온도(주간, 야간, 평균), 실제 및 잠재 증발산량, 식생 지수인 EVI와 NDVI, 강수량, 유효광합성반사율, 엽면적지수 등 위성 영상에서 도출한 자료와 지면모델 산출물인 실제 및 잠재 증발산량, 토양수분과 대규모 기후 지수인 AO와 MEI가 사용되었다. 모델의 목표 변수로는 가뭄의 상태를 평가하기 위해 개발되었으며 널리 사용되고 있는 관측 자료 기반의 SPI, SPEI와 위성 영상에서 도출된 반사율을 이용하여 산정하도록 개발된 NDWI, NDDI, NMDI를 선정하여 사용하였다.

모델에 입력 변수로 사용하기에 앞서 도출된 위성 영상 기반 변수 중 강수량, 지표면 온도, 잠재 증발산량, 토양수분에 대하여 관측 값과 비교해 보았는데, 강수량의 경우 TRMM 위성이 신뢰할 만한 정보를 제공하는 위도 구간에 대해서는 오차 폭이 그다지 크지 않았으며 상관관계도 매우 높게 나타났다. 지표면 온도의 경우에는 여름철 주간 온도를 제외하고는 관측 값과 높은 상관관계를 보였다. 잠재 증발산량의 경우에는 MODIS 자료에서 도출한 경우가 지면모델 산출물보다 Penman-Monteith 식으로 산정한 값에 더 가까운 것으로 나타났으나, 토양수분의 경우에는 지면모델 산출물이 AMSR-E 위성 자료에서 도출한 값보다 KoFlux 관측 값에 더 가까웠다. 위성 영상 자료와 지면모델 산출물 사이에 월별 자료를 이용한 회귀 모델을 구축하여 두 자료를 결합하였는데, 이 경우 토양수분 추정 값의 오차를 줄이고 시공간 해상도를 높일 수 있다.

가뭄 감시 모델은 random forest를 이용하여 구축하였으며 여러 입력 변수의 집합에 대하여 목표 변수를 얼마나 잘 추정하는지 평가하였다. 결과적으로 모든 입력 변수를 다 포함한 경우 모델의 성능이 가장 우수한 것으로 판단되었다. Random forest 는 OOB 과정을 통해 내부 오차를 예측하며 변수별 중요도를 제공하는데, 관측 기반의 목표 변수의 경우 강수량의 중요도가 압도적으로 높으며 원격 탐사 기반의 목표 변수의 경우 대규모 기후 지수인 AO와 MEI의 상대적 중요도가 가장 크게 나타났다. 원격 탐사 기반의 목표 변수 중 특히 NDWI6와 NDWI7이 입력 값과 출력 값 사이의 상관관계가 가장 높아 가뭄 감시에 유용할 것으로 보인다.

가뭄 감시 모델에서 선정된 모든 입력 변수를 포함한 입력 변수 집합에 대하여 1개월에서 6개월까지의 선행 예측 시간과 1개월에서 3개월까지의 사용 과거 자료 개월 수에 대해 가뭄 예측 모델을 구축하여 성능을 평가하였다. 일반적으로 선행 예측 시간이 길어질수록 오차 값이 점차 증가하며 사용하는 입력 변수의 과거 개월 수가 늘어나면 오차 값이 감소하는 것으로 나타났는데, 선행 예측 시간이 1개월에서 6개월로 증가할수록 증가하는 오차 폭 보다 오히려 포함하는 과거 개월 수에 따라 오차 값이 크게 변화하는 것을 관찰할 수 있었다. 원격 탐사 기반의 목표 변수 중에서는 NDDI5와 NDDI6은 불안정한 오차 변이를 나타내었으나 나머지 목표 변수들은 우수한 가뭄 예측 성능을 보였다. 변수들의 상대적 중요도를 보면, 목표 변수 자신의 과거 값의 중요도가 가장 높은 것으로 나타났으며 가뭄 감시 모델과 마찬가지로 대규모 기후 지수인 AO, MEI의 중요도가 높았다. 식생과 관련된 변수인 LAI, FPAR, EVI, NDVI의 중요도는 상대적으로 낮게 나타났다.

고랭지 채소 수확량과 산불 발생 빈도를 이용하여 개발된 가뭄 예측 모델을 적용하였다. 고랭지 채소 수확량 및 산불 발생 빈도를 예측하는 모델을 따로 구축한 것이 아니라 이들 변수에 대해서 각각의 값과 높은 상관관계를 가지는 원격 탐사 기반 가뭄 지수를 선정하고 이러한 밀접한 관계를 보이는 행정구역에 대하여 이 목표 변수가 잘 예측될 수 있는지 성능을 평가하였다.

고랭지 채소 수확량의 경우 분석에 사용된 아홉 곳의 시도와 여섯 곳의 시군구 중 고랭지 배추에 대해서는 시도로는 전라북도와 경상남도 두 곳, 시군구로는 강릉시, 횡성군, 영월군, 양구군, 인제군의 다섯 곳에서 수확량이 원격 탐사 기반 가뭄 지수와 높은 상관관계를 보였으며, 고랭지 무에 대해서는 시도로는 강원도와 경상북도, 경상남도 세 곳, 시군구로는 강릉시, 태백시, 삼척시, 영월군, 인제군의 다섯 곳이 높은 상관관계를 보였다. 고랭지 배추와 무의 성장에는 본 모델 구축에 활용된 수문기상변수 이외에도 많은 요인이 작용하므로, 많은 경우 상관관계가 낮게 나타났으나 앞서 나열한 경우에는 원격 탐사 기반의 가뭄 지수를 고랭지 지역의 채소 수확량 예측에 활용할 수 있는 대리변수로 사용할 수 있을 것으로 보인다.

해당 행정구역에서 목표 변수의 예측 성능을 보면, 고랭지 배추의 경우 특히 시군구 단위로 영월군의 NDWI7이 6개월 선행 예측 시간에 대해 예측 성능이 가장 뛰어났으며 강릉시의 NDDI7이 3개월 선행 예측 시간에 대해서, 영월군의 NDWI6이 2개월과 6개월 선행 예측 시간에 대해 잘 예측되었다. 고랭지 무의 경우 강원도, 태백시, 삼척시, 영월군, 강릉시에서 각각 NDDI7, NDWI7, NMDI, NDWI7, NDDI7이 수확량과 높은 상관관계를 보였지만 이 지역에서 이 목표 변수들의 예측 성능은 그다지 뛰어나지 않았다.

산불 발생 빈도(frequency), 산불 발생으로 인한 피해 면적(area), 해당 행정구역의 산지 면적에 대한 피해 면적의 비율(ratio) 중 오직 산불 발생 빈도만이 원격 탐사 기반 가뭄 지수와 높은 상관관계를 보였다. 오로지 시도 단위에서만 높은 상관관계를 볼 수 있었으며, 가뭄 지수는 NDWI와 NDDI 만이 상관관계가 좋은 것으로 드러났다. 산불 발생 빈도와 목표 변수와의 사이에 높은 상관관계를 보인 부산, 울산, 충청북도, 전라남도, 경상북도, 경상남도의 지역에서 모두 해당 목표 변수 예측 성능이 매우 우수한 것으로 나타났다.

6.2 연구의 결론 및 시사점

본 연구에서 기계학습 방법론인 random forest를 이용하여 구축한 가뭄 감시 및 가뭄 예측 모델을 오차 값과 결정계수 값을 이용하여 평가한 결과 가뭄 상태의 대리 변수로 선정된 SPI, SPEI, NDWI, NDDI, NMDI를 매우 잘 추정 및 예측하는 것으로 나타났다. 특히 위성 영상에서 도출한 자료, 지면모델 산출물, 대규모 기후 지수를 모두 포함한 경우 가뭄 감시 모델의 성능이 우수하게 나타났다. 변수의 상대적 중요도에 있어 가뭄 감시 모델 및 가뭄 예측 모델에서 AO, MEI의 기후 지수의 중요도가 높게 나온 것은 우리나라의 가뭄 상태 역시 원격상관(teleconnection)의 영향을 크게 받는다는 사실을 시사하고 있다. 적도 부근의 지역이나 동남아시아, 북미 등의 지역에 비해 원격상관의 영향을 그다지 받지 않는다고 알려져 있었던 것과는 다른 결과로, 가뭄 감시 및 예측에 원격상관을 활용하여 추정 및 예측 성능을 높일 수 있다.

일별 자료로 분석하였을 때 토양수분의 경우 지면모델의 산출물이 위성 영상에서 도출한 자료보다 더 관측 값에 가까운 결과를 생산하였으며, 가뭄 감시 모델에 있어 지면모델 산출물을 포함한 경우 모델의 성능이 더 우수하였다. 이는 위성 영상 자료만을 사용한 경우 보다 지면모델 산출물이 가용하다면 이를 함께 사용하는 것이 유용하다는 것을 보여주는데, 이 때 모델 오차 값을 줄일 수 있을 뿐 아니라 지면모델 결과는 입력 자료에 따라 다양한 공간해상도로 생산될 수 있으므로 가뭄 감시 및 예측 자료의 시공간 해상도를 높이는 데 기여할 수 있다.

가뭄 감시 및 예측 모델 구축에 있어 위성 영상에서 도출한 변수만을 포함한 경우에도 관측 기반 목표 변수인 SPI, SPEI와 원격 탐사 기반 목표 변수 중 NDWI의 경우에는 모델 성능이 상당히 우수하게 나타났으며, 위성 영상에서 도출한 변수와 지면모델 산출물, 대규모 기후 지수를 함께 쓴 경우에는 더욱 우수하게 나타났는데, 이는 본 연구의 궁극적인 목표로 삼은 관측 지점이 조밀하게 분포하지 않거나 현장 관측 자료가 부재하는 경우의 가뭄 감시 및 가뭄 예측이 random forest를 이용해 구축한 모델을 이용하여 잘 이루어질 수 있음을 나타낸다.

고랭지 채소 수확량과 산불 발생 빈도에 본 연구에서 구축한 가뭄 예측 모델을 적용하였는데, 본 모델을 통하여 직접 수확량과 발생 빈도를 예측하지는 않았지만 고랭지 채소 수확량과 산불 발생 빈도와 높은 상관관계를 가지는 행정구역 별 목표 변수를 선정하고 이 행정구역에서 해당 목표 변수의 예측 성능을 평가하였다. 이 때 모든 행정구역에서는 아니지만, 상당히 많은 시도 또는 시군구에서 고랭지 채소의 수확량 및 산불 발생 빈도와 여러 목표 변수가 높은 상관관계를 보였으며 이 지역에서 목표 변수의 예측 성능이 1개월에서 6개월의 선행 예측 시간에 걸쳐 비교적 높게 나타났다. 고랭지 채소의 수확량이나 산불 발생 빈도에는 본 가뭄 감시 모델이나 가뭄 예측 모델에서 고려한 입력 변수들 외에도 많은 요인이 작용하는데, 본 연구에서는 이들을 모두 고려하여 심도 깊은 분석을 수행하지는 못하였다. 그러나 앞서 설명한 대로 우수한 성능을 가진 random forest 기반의 가뭄 감시 모델 및 가뭄 예측 모델이 대리 변수로 사용된 목표 변수를 우수하게 추정하고 예측하고 있음을 알 수 있다.

6.3 활용 제안

홍수나 태풍과 같은 재난과는 다르게 가뭄은 정의부터가 가뭄의 피해를 입는 부문 및 대상을 중심으로 한다. 제1장에서 가뭄의 정의를 소개하였는데, 일반적으로는 심각한 수문학적 불균형을 야기할 정도로 장기간 동안 지속되는 비정상적으로 건조한 날씨로 정의하나, 가뭄에 대한 모든 연구는 여기에서 한걸음 더 나아가 가뭄의 피해가 발생하는 대상을 바탕으로 기상학적 가뭄, 농업적 가뭄, 수문학적 가뭄 등으로 세분화되어 이루어진다. 이렇게 가뭄의 연구가 가뭄 피해를 입는 대상을 위주로 이루어지다 보니 가뭄 연구는 당연히 사용자 중심으로 이루어져야 한다.

본 연구를 통해 가뭄 감시 모델 및 가뭄 예측 모델을 구축하였으며, 목표 변수로 선정한 다양한 시간척도의 SPI, SPEI와 원격 탐사 기반의 여러 가뭄 지수인 NDWI, NDDI, NMDI 등에 대하여 감시 및 예측 성능이 상당히 우수함을 보였다. 이 중 특히 SPI와 SPEI의 경우에는 시간 척도를 1개월에서 9개월로 다양하게 선정하였는데, 이에 따라 가뭄 피해를 입는 부문 및 대상이 달라질 수 있다. 따라서 이러한 가뭄 감시 및 가

물 예측 정보는 가뭄에 대한 일반적인 정보를 전달하는 기상청의 가뭄정보시스템에 통합된다면 가장 유용하게 배포 및 활용될 수 있을 것이다.

원격 탐사 기반의 가뭄 지수를 목표 변수로 선정한 경우는 가뭄 피해의 대상이 정해져 있는데, NDWI와 NDDI의 경우는 식생의 가뭄 피해를 중심으로 살펴보기 위해 개발되었으며 NMDI는 토양과 식생의 수분 상태를 함께 보기 위해 개발되었다. 따라서 이들 목표 변수를 대상으로 구축된 가뭄 감시 모델 및 가뭄 예측 모델은 농업적 가뭄(토양 수분 가뭄)에 가장 많은 정보를 제공할 수 있다. 본 연구에서 고랭지 채소 수확량과 산불 발생 빈도를 통해 모델을 적용해 보았듯 농업인을 위한 정부부처인 농림축산식품부, 농어촌공사 등이나 산림청 등의 가뭄정보시스템에 통합하여 정보를 제공하거나 직접적으로 농업인, 산림 관리인을 대상으로 하여 가뭄 감시 및 예측 정보를 생산하는 것이 본 연구 결과의 활용에 가장 적합할 것이다.

지면 모델 결과가 부재한 경우에도 위성 영상 자료와 대규모 기후 지수를 이용하여 관측 기반 및 원격 탐사 기반 목표 변수를 잘 모사하는 가뭄 감시 및 예측 모델을 구축하였으므로, 이러한 결과는 미계측 지역의 가뭄 관리를 위해 활용될 수 있을 것으로 보인다. 본 연구에서 활용한 위성 영상은 별도의 비용 없이 대부분의 지역에 대하여 획득할 수 있으므로 NDWI, NDDI, NMDI와 같은 원격탐사 기반 지수를 이용하여 미계측 지역을 대상으로 가뭄 감시나 예측, 미관개지역의 채소 수확량 예측, 산불 발생 빈도 예측 모델을 구축할 수 있을 것이다.

현재로서는 위성 영상 기반 정보가 충분히 활용되고 있지 못하는 만큼, 각 기관별로 구축되고 있는 가뭄 감시 시스템 및 가뭄 조기 경보 시스템 등에 지상관측 자료와 위성 영상 기반 정보를 통합하여 구성하여야 할 것이다. 위성영상 기반 자료의 평가 및 검증을 통해 이를 통한 가뭄 정보의 신뢰도를 제고하고 다양한 사례 연구 및 분석을 수행하여 최종 사용자가 신뢰를 가지고 자신 있게 위성영상 기반의 가뭄 정보를 의사 결정에 활용할 수 있는 환경 또한 만들어져야 한다. 본 연구에서 개발한 기술을 이용하여 모델을 구축, 활용하게 된다면 향후 GPM으로 TRMM을 대체하는 문제, SMAP이나

AMSR-2를 이용하여 AMSR-E를 대체하는 문제 등을 고려하여야 현업적으로 사용하는데 무리가 없을 것이다.

6.4 향후 가뭄 연구 제안

본 연구에서는 기계학습의 일종인 random forest 모델을 활용하여 가뭄 감시 모델 및 가뭄 예측 모델을 구축하고 모델의 성능을 살펴보았다. 원격 탐사 자료의 활용에 있어 위성 간의 연속성을 보장할 수 있도록 국내 위성을 포함하는 다양한 위성 자료의 활용을 추후 연구에서 살펴볼 것이다. 가뭄 감시 뿐 아니라 가뭄 예측의 방법론은 기계학습만 있는 것이 아니라 제3장에서 살펴본 바와 같이 시계열 모델, 확률 모델, 신경망 모델, 하이브리드 모델 등 다양한 방법론이 존재하므로 각각의 모델을 검토하여 보다 적합한 모델을 선정하거나 결합하여 활용하여야 할 것이며, 이렇게 생산한 가뭄 예측 정보를 가뭄 피해를 최소화하기 위해 사용하여야 할 것이다.

가뭄 예측 정보는 가뭄 대응을 위한 사회·경제적 자산 운용에 대한 사전 계획 및 수정을 가능하게 하여 가뭄 취약성 감소에 큰 도움일 줄 수 있으나, 실제 활용에 있어서는 여러 장애가 존재한다. 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). 또한 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).

따라서 본 연구에 따른 후속 연구로, APEC 기후센터의 다중모델앙상블(Multi-Model Ensemble; MME) 자료 등을 이용한 가뭄 예측 정보를 도출하고 본 연구에서 수행된 기계학습을 활용한 가뭄 예측과 함께 활용하여 불확실성에 효과적으로 대응할 수 있는 가뭄 예측 정보를 생산할 것이며 또한 이를 지방자치단체 단위의 의사결정에 활용할 수 있는 방안을 연구할 것이다.

REFERENCES

<국문 자료>

- APEC기후센터(이진영), 2013a: 강수와 증발산량 추정을 통한 수문학적 가뭄의 원격
APCC Research Report 2013-04. pp. 47-90.
- APEC기후센터(전종안), 2013b: LSM-Biogeochemistry 연계모형을 이용한 토지이용
별 온실가스 배출 추정. APCC Research Report 2013-05. pp. 137-173.
- 국가통계포털, 2014: <http://kosis.kr>.
- 권효정, 김준, 2010: KoFlux 역정: 배경, 현황 및 향방, *한국농림기상학회지*, **12(4)**,
241-263.
- 기상청, 2011: 「기후변화 시나리오 보고서 2011」. 117 pp.
- 김기덕, 2014, personal communication, 고려지농업연구센터.
- 농림부, 2000: 2000년 「가뭄·풍수해 극복 사례집」. 713 pp.
- 방재연구소(심기오), 2009: 「국한가뭄 대응방안에 관한 연구」. 170 pp.
- 산림청, 2010: 「산불통계 분석을 통한 산불정책의 변천 및 대응방안」. 302 pp.
- 산림청, 2014: 「2013년 산불통계 연보」. 215 pp.
- 원명수, 2014, personal communication, 국립산림과학원.
- 전승택, 2014, personal communication, 강원도청.
- 한국농촌경제연구원(이중웅), 1996: 무·배추의 생산결정요인 분석에 관한 연구. 77 pp.
- 한국환경정책평가연구원(김연주), 2013: 「가뭄 유형별 기후변화 적응정책 연구」. 235 pp.

<영문 자료>

- Aguado, E., and J. E. Burt, 2007: *Understanding weather and climate*. 4th Ed., Pearson rentice Hall, 562 pp.
- Alley, W. M., 1984: The Palmer Drought Severity index: limitations and assumptions, *J.Clim. Appl. Meteorol.*, **23**, 1100-1109.
- Anderson, L., Y. Malhi, L. Aragão, R. Ladle, E. Arai, N. Barbier, and O. Phillips, 2010: Remote sensing detection of droughts in Amazonian forest canopies. *New Phytol.*, **187**, 733-750.
- Breiman, L., 2001: Random forests. *Mach. Learn.*, **45**, 5-32.
- Brown, J. F., B. D. Wardlow, T. Tadesse, M. J. Hayes, and B. C. Reed, 2008: The Vegetation Drought Response Index (VegDRI): A New Integrated Approach for Monitoring Drought Stress in Vegetation. *GIScience & Rem. Sens.*, **45**, 16-46.
- Burke, E. J., and S. J. Brown, 2008: Evaluating uncertainties in the projection of future drought. *J. of Hydrometeorol.*, **9(2)**, 292-299.
- Chiew, F. H. S., and T. A. McMahon, 2002: Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability. *Hydrol. Sci. J.*, **47**, 505-522.
- CPC(Climate Prediction Center), 2014: Seasonal DroughtOutlook, http://www.cpc.ncep.noaa.gov/products/expert_assessment/sdo_summary.html.
- Cutter, S., B. Ossman-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. In: *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, pp. 291-338.
- Dai, A.G., K. E. Trenberth, and T. T. Qian, 2004: A global dataset of Palmer Drought Severity Index for 1870-2002: Relationship with soil moisture and effects of surface warming. *J. of Hydrometeorol.*, **5(6)**, 1117-1130.
- Deardorff, J. W., 1977: A parameterization of ground-surface moisture content for use in atmosphere prediction models. *J. Appl. Meteorol.*, **16**, 1182-1185.
- 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(www.em-dat.net). Université Catholique de Louvain, Brussels, Belgium.
- 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.
- Gao, H., Q. Tang, X. Shi, C. Zhu, T. Bohn, F. Su, J. Sheffield, M. Pan, D. Lettenmaier, and E. F. Wood, 2009: *Water budget record from Variable Infiltration Capacity (VIC) Model Algorithm Theoretical Basis Document (Version 1.2)*.
- Guttman, N. B., 1998: Comparing the Palmer Drought Index and the Standardized Precipitation Index. *J. Am. Water Resour. As.*, **34**, 113-121.
- 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.

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.

Hirabayashi, Y., S. Kanae, S. Emori, T. Oki, and M. Kimoto, 2008: Global projections of changing risks of floods and droughts in a changing climate. *Hydrolog. Sci. J.*, **53(4)**, 754-772.

Hsu, K.-L., X. Gao, S. Sorooshian, and H. V. Gupta, 1996: Precipitation estimation from remotely sensed information using artificial neural networks. *J. Appl. Meteorol.*, **36**, 1176-1190.

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.

Jensen, J. R., 2000: *Remote sensing of the environment: An earth resource perspective*. New Jersey: Upper Saddle River. 592 pp.

Jensen, J. R., 2005: *Introductory digital image processing: A remote sensing perspective*. 3rd Ed., Prentice Hall. 544 pp.

Keetch, J., and G. Byram, 1968: *A drought index for forest fire control*. Forest Service Research Paper SE-38, U. S. Department of Agriculture, 32 pp.

- 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: Drought of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data. *Bull. Ameri. Meteo. Soci.*, **76**, 655-668.
- Leilah, A. A., and S. A. Al-Khateeb, 2005: Statistical analysis of wheat yield under drought conditions. *J. Arid Environ.*, **61**, 483-496.
- Liu, H. Q., and A. R. Huete, 1995: A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE T. Geosci. Remote*, **33**, 457-465.
- Manabe, S., 1969: Climate and the ocean circulation: 1, the atmospheric circulation and the hydrology of the Earth's surface. *Mon. Weather Rev.*, **97**, 739-805.
- McKee, T. B., N. J. Doesken, and J. Kleist, 1993: The relationship of drought frequency and duration of time scales. In Proc. of the 8th Conf. of Applied Climatology, Anaheim, CA. 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.
- Mu, Q., M. Zhao, and S. W. Running, 2011: Improvements to a MODIS global terrestrial evapotranspiration algorithm. *Rem. Sens. Environ.*, **115**, 1781-1800.

- Mu, Q., M. Zhao, J. S. Kimball, N. G. McDowell, and S. W. Running, 2013: A remotely sensed global terrestrial drought severity index. *Bull. Ameri. Meteo. Soci.*, 94(1), 83-98.
- NIDIS (National Integrated Drought Information System) Act, 2006. United States Public Law 109-430.
- Palmer, W. C., 1965: *Meteorological drought*. U.S. Weather Bureau Research Paper 45, 58 pp.
- Palmer, W. C., 1968: Keeping track of crop moisture conditions, nationwide: The new crop moisture index. *Weatherwise*, **21**, 156-161.
- Pitman, A. J., 2003: Review: The evolution of, and revolution in, land surface schemes designed for climate models. *Int. J. Climatol.*, **23**, 479-510.
- Rhee, J., J. Im, and G. J. Carbone, 2010: Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data. *Rem. Sens. Environ.*, **114**, 2875-2887.
- Rhee, J., J. Im, and S. Park, 2015: *Regional drought monitoring based on multi-sensor remote sensing*, P. S. Thenkabail, Ed., Remote Sensing Handbook Vol. III, 125-161.
- Sellers, P. J., R. E. Dickinson, D. A. Randall, A. K. Betts, F. G. Hall, J. A. Merry, A. S. Denning, H. A. Mooney, C. A. Nobre, N. Sato, C. B. Field, and A. Henderson-Sellers, 1997: Modelling the exchanges of energy, water and carbon between continents and the atmosphere. *Science*, **275**, 502-509.
- Seneviratne, S. I., T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orłowsky, and A. J. Teuling, 2010: Investigating soil moisture-climate interactions

in a changing climate: A review. *Earth-Sci. Rev.*, **99**, 125-161.

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, Ed., A Special Report of Working Group I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, 109-230.

Sheffield, J., and E. F. Wood, 2008: Global trends and variability in soil moisture and drought characteristics, 1950-2000, from observation-driven simulations of the terrestrial hydrologic cycle. *J. of Climate*, **21(3)**, 432-458.

Shukla, S., and D. P. Lettenmaier, 2011: Seasonal hydrologic prediction in the United States: Understanding the role of initial hydrologic conditions and seasonal climate forecast skill, *Hydrol. Earth Syst. Sc.*, **15**, 3529-3538.

Steinemann, A. C., 2006: Using climate forecasts for drought management. *J. Appl. Meteorol. Clim.*, **45**, 1353-1361.

Svoboda, M., D. LeComte, M. Hayes, R. Heim, K. Gleason, J. Angel, B. Rippey, R. Tinker, M. Lalecki, D. Stooksbury, D. Miskus, and S. Stephens, 2002: The Drought Monitor. *Bull. Ameri. Meteo. Soci.*, **83**, 1181-1190.

Tadesse, T., J. F. Brown, and M. J. Hayes, 2005: A new approach for predicting drought-related vegetation stress: Integrating satellite, climate, and biophysical data over the U.S. central plains. *ISPRS J. Photogramm.*, **59**, 244-253.

- Tsakiris, G., D. Pangalou, and H. Vangelis, 2007: Regional drought assessment based on the Reconnaissance Drought Index(RDI). *Water Resour. Manag.*, **21**, 821-833.
- Tucker, C. J., 1979: Red and photographic infrared linear combinations for monitoring vegetation. *Rem. Sens. Environ.*, **8**, 127-150.
- 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. of Climate*, **23**, 1696-1718.
- Wang, L., and J. J. Qu, 2007: NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophys. Res. Lett.*, **34**, L20405, doi:10.1029/2007GL031021.
- Wardlow, B. D., M. C. Anderson, and J. P. Verdin, *Remote sensing of drought: innovative monitoring approaches*. CRC Press. Boca Raton, FL. 422 pp.
- Wells, N., S. Goddard, and M. J. Hayes, 2004: A Self-Calibrating Palmer Drought Severity Index. *J. Climate*, **17**, 2235-2351.
- Wilhite, D., and M. R. Glantz, 1987: Understanding the drought phenomenon-The role of definitions. Planning for drought, D. Wilhite, W. Easterling, and D. Wood, Ed., Westview Press, 11-27.
- Zhang, A., and G. Jia, 2013: Monitoring meteorological drought in semiarid regions using multi-sensor microwave remote sensing data. *Rem. Sens. Environ.*, **134**, 12-23.

연구보고서 2015-14

원격 탐사 자료를 활용한 가뭄 감시·예측 기술 개발

Development of Drought Monitoring and Forecasting Techniques based on Remotely Sensed Data

이진영 기후변화연구팀



APEC Climate Center

612-020 부산광역시 해운대구 센텀7로 12

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

www.apcc21.org

 www.facebook.com/apcc21

 www.youtube.com/APECClimateCenter21

 www.twitter.com/apcc21

 www.plus.google.com/+APECClimateCenter21

 www.flickr.com/apcc21