



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

Doctoral Thesis

Drought monitoring and forecasting
through multi-sensor satellite data fusion
using machine learning approaches

Seonyoung Park

Department of Urban and Environmental Engineering
(Environmental Science and Engineering)

Graduate School of UNIST

2018

Drought monitoring and forecasting through multi-sensor satellite data fusion using machine learning approaches

Seonyoung Park

Department of Urban and Environmental Engineering
(Environmental Science and Engineering)

Graduate School of UNIST

Drought monitoring and forecasting through multi-sensor satellite data fusion using machine learning approaches

A dissertation
submitted to the Graduate School of UNIST
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Seonyoung Park

05/29/2018

Approved by



Advisor

Jungho Im

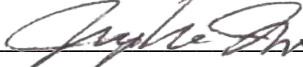
Drought monitoring and forecasting through multi-sensor satellite data fusion using machine learning approaches

Seonyoung Park

This certifies that the dissertation of Seonyoung Park is approved.

05/29/2018

signature



Advisor: Jungho Im

signature



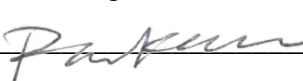
Thesis Committee Member: Myong-In Lee

signature



Thesis Committee Member: Kyung Hwa Cho

signature



Thesis Committee Member: Sunyurp Park

signature



Thesis Committee Member: Jaeil Cho

Abstract

This thesis seek to 1) better understand the characteristics of drought factors in different climate regions (arid and humid regions), 2) monitor drought effectively using improved soil moisture to high resolution data (~1 km) by developing drought index, 3) predict drought by identifying the relationship between drought and climate variability.

In this dissertation, there are five chapters. Chapter 1 summarizes the background of research and overviews of the thesis research. In Chapter 2, the characteristics of drought factors in different climate regions were understood based on relative variable importance which is provided from machine learning approaches. Three machine learning approaches (Random forest, Boosted regression trees, and Cubist) were used to target Standardized Precipitation Index (SPI) and crop yield, and Random forest outperformed the other approaches. The variable importance targeting SPI and crop yield was applied as weight factors to produce drought maps. Those drought maps well monitored drought comparing U.S. Drought Monitor (USDM). In Chapter 3, Soil moisture with coarse spatial resolution (25 km) was downscaled to high resolution (1 km), and drought index (High resolution Soil Moisture Drought Index; HSMDI) was developed using high resolution soil moisture. HSMDI was evaluated for three types of droughts (meteorological drought, agricultural drought, and hydrological drought). The index well monitored meteorological and agricultural drought, but there is limitation in monitoring hydrological drought. In Chapter 4, short-term (one pentad) prediction of drought conducted over East Asia by integrating remote sensing data and climate variability. The real-time multivariate (RMM) Madden–Julian oscillation (MJO) indices were used because the MJO is intraseasonal climate variability. MJO improved the performance of prediction model. Chapter 5 provides a brief summary of this study. Chapter 6 discusses future work of this study.

CONTENSTS

Abstract	i
Contents	iii
List of Figures	vi
List of Tables	viii

Chapter 1

1. Introduction	1
1.1 Background	1
1.2 Remote sensing in drought	3
1.2.1 Satellite-based drought indices	3
1.2.2 Drought Condition Index	3
1.2.2 Limitation of remote sensing data due to cloud	4
1.3 Overview of papers	4
1.4 Research highlight	6

Chapter 2

2. Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions	8
2.1 Background	8
2.2 Study area and data	10
2.2.1 Study area	10
2.2.2 Remote sensing data	11
2.2.3 Reference data	13
2.3 Methods	14
2.3.1 Machine learning approaches	14
2.3.2 Analyses of the variables for drought monitoring	15
2.4 Results and discussion	16
2.4.1 Monitoring meteorological drought	16
2.4.2 Monitoring agricultural drought	23
2.4.3 Comparison with USDM maps	25
2.5 Conclusion	29

Chapter 3

3. Drought monitoring using high resolution soil moisture through multi-sensor satellite data fusion over the Korean peninsula	30
3.1 Background	30
3.2 Study area and data	32

3.2.1 Study area	32
3.2.2 Satellite data	33
3.2.3 Observation data	34
3.3 Methods	35
3.3.1 Downscaling soil moisture	35
3.3.2 High resolution Soil Moisture Drought Index (HSMDI)	36
3.4 Results and discussion	37
3.4.1 Assessment of downscaled soil moisture <i>in situ</i> data	37
3.4.2 Monitoring meteorological drought	40
3.4.3 Monitoring agricultural drought	42
3.4.4 Monitoring hydrological drought	45
3.4.5 Novelty, opportunities, and limitations	46
3.5 Conclusion	47

Chapter 4

4. Very short-term prediction of drought using remote sensing data and MJO index through random forest over east Asia	49
4.1 Background	49
4.2 Study area and data	51
4.2.1 Study area	51
4.2.2 Data	51
4.2.2.1 Satellite data	51
4.2.2.2 Numerical Land Surface data	53
4.2.2.3 Madden-Julian Oscillation (MJO) index	53
4.3 Methods	53
4.4 Results and discussion	55
4.4.1 Meteorological climatology associated with MJO	55
4.4.2 Comparison of drought prediction model performance	57
4.5 Conclusion	64

Chapter 5

5. Conclusions	65
-----------------------	-----------

Chapter 6

6. Outlook and Future works	66
------------------------------------	-----------

Chapter 7

7. References	67
----------------------	-----------

List of Figures

Figure 1.1 Drought propagation in the Anthropocene (Source: National Drought Mitigation Center, University of Nebraska-Lincoln, U.S.A.)

Figure 1.2 Numbers of people affected by water-related disasters (1995-2015) (Source: UNISDR / CRED)

Figure 2.1 Study area includes the arid region of Arizona and New Mexico (left) and the humid region of North Carolina and South Carolina. There are 48 counties in the arid region and 146 counties in the humid region.

Figure 2.2 Flow chart of the processes used in this study.

Figure 2.3 Model performance of three machine learning approaches (random forest, boosted regression trees and Cubist) for 1-, 3-, 6-, 9-, and 12-month SPIs in the arid, humid, and combined regions.

Figure 2.4 Scatter plots between predicted SPI12 based on random forest, boosted regression trees and Cubist and in situ SPI12 in the arid (a-c), humid (d-f), and combined regions (g-i).

Figure 2.5 The relative importance (%) of the sixteen drought factors using random forest for 1-, 6-, and 12-month SPI.

Figure 2.6 The relative importance (%) of SPIs with 1- to 12-month time scales using random forest for estimating crop yield and Pearson's correlation coefficients between each SPI and crop yield in the non-irrigated region (humid region).

Figure 2.7 Year-to-year change between two types of remote sensing based drought maps (3-month SPI and crop yield) and USDM maps in the arid region for May from 2000 to 2012.

Figure 2.8 Year-to-year change between two types of remote sensing based drought maps (3-month SPI and crop yield) and USDM maps in the humid region for September from 2000 to 2012.

Figure 2.9. Validation drought monitoring map (2013 and 2014) in the arid (irrigated) and humid (non-irrigated) regions.

Figure 3.1. Study area with land cover, elevation, and in situ reference data locations.

Figure 3.2. Data processing flow diagram to develop the High resolution Soil Moisture Drought Index (HSMDI) based on a multi-sensor downscaling approach.

Figure 3.3. Scatterplots between in situ and satellite-derived soil moisture ($p < 0.05$) at two flux sites (CMK and SMK).

Figure 3.4. Boxplots of correlation coefficient values between HSMDI and SPIs with precipitation by month from April to October.

Figure 3.5. Time series of 1- and 3-month SPIs and HSMDI (monthly) from April to October for 2003-2011 by station within homogeneous land cover domains. SPI1 = 1-month SPI; and SPI3 = 3-month SPI.

Figure 3.6. Scatterplots between yearly crop yield (sesame, highland radish, and highland napa cabbage) and HSMDI (0.05 significant level) in August.

Figure 3.7. Correlation coefficients between SPIs and highland crop yields (radish and napa cabbage) at two administrative districts ($p < 0.05$).

Figure 3.8. Scatterplots between scaled streamflow and scaled HSMDI ($p < 0.05$) in five watersheds (September).

Figure 4.1 Study area with MODIS Land cover 2010 and sixteen land cover classes were aggregated to eleven representative classes.

Figure 4.2 Flowchart of this study. 1-, 2- and 3-pentad before the date RMM MJO indices and drought indices, latitude and longitude were used to predict drought

Figure 4.3 Composites of 30-90 days filtered surface air temperature based on each of eight MJO phases during 22 spring seasons (April-May) of 1991–2012

Figure 4.4 Variance of (a) surface air temperature, (b) top-level volumetric soil moisture, and (c) precipitation bandpass filtered for 30-90 days during 22 springs between 1991–2012, and their ratios to total variance in (d-f).

Figure 4.5 Modeling performances of two drought prediction models (with MJO and without MJO) for three drought indices. Scatter plots, r , and RMSE were produced through validation.

Figure 4.6 leave-one-out validation of VSDI prediction model. Each pixel includes 9 pentads x 17 years data. Each pixel was processed as no data when the number of data was less than 50.

Figure 4.7 Comparison of change of drought conditions between actual VSDI and predicted VSDIs from two models (with MJO and without MJO) in 2010.

Figure 4.8 Comparison of change of drought conditions between actual VSDI and predicted VSDIs from two models (with MJO and without MJO) in 2011.

List of Tables

Table 2.1 Remote sensing based drought factors and their formula.

Table 2.2 Irrigation rates (%) in arid, humid and temperate regions.

Table 2.3. The Relative importance (%) of the most important five drought factors among the sixteen variables for meteorological drought (SPI) using random forest (RF) and boosted regression trees (BRT) in the arid, humid and combined regions for 1-, 3-, 6-, 9-, and 12-month SPI.

Table 2.4. The variable usage information in conditions and models (%) of the most important five drought factors among the sixteen variables for meteorological drought (SPI) using Cubist in the arid, humid and combined regions for 1-, 3-, 6-, 9-, and 12-month SPI.

Table 2.5. The scaled relative importance of the selected six variables for 1-, 3-, 6-, 9-, and 12-month SPI and model performance (R² and RMSE) in the arid (A), humid (H) and combined (C) regions. The sum of the relative importance is 1.

Table 2.6 The scaled relative importance of the selected six variables for corn and soybean in the arid (irrigated) and humid (non-irrigated) regions and model performance (R², RMSE and rRMSE). The sum of the relative importance is 1.

Table 3.1. Soil moisture downscaling model performance and variable importance (%).

Table 3.2. Validation of the downscaled soil moisture with the AMSR-E soil moisture and in-situ soil moisture from March to November 2010.

Table 3.3 Correlation coefficients (r) between HSMDI in five watersheds and scaled streamflow at five stations during the growing season.

Table 4.1 Variable importance of two drought prediction models (with MJO and without MJO) for three drought indices.

Drought monitoring and forecasting through multi-sensor satellite data fusion using machine learning approaches

Chapter 1

1. Introduction

1.1 Background

Drought occurs from lack of precipitation due to climate variability (meteorological drought), and prolonged meteorological drought leads to soil moisture drought (agricultural drought) and low water level of dam and stream flow (hydrological drought) (figure 1.1). Drought has caused human and ecological damages such as reduced crop yield, water shortage, heat wave, forest fires (socioeconomic drought). According to United Nations International Strategy for Disaster Reduction (UNISDR), 1.1 billion people (around 26% of people affected by water-related disasters) have suffered from drought. Therefore, it is important to mitigate drought for developing adequate policy by monitoring and forecasting drought.

Persistency of drought differs for each region under the same condition of precipitation because each region has its own characteristics including land use and infrastructures (e.g., irrigation rate, dam building). In addition, depending on the characteristics of the climate zone (e.g., humid, arid, and temperate), the mechanism of occurrence of drought and the characteristics of drought are also different. While soil moisture droughts due to high temperature occur more frequently in arid regions than in humid regions, humid regions are more vulnerable to precipitation deficit than arid regions. Therefore, it is important to diagnose and monitor drought at local or regional scale. Drought monitoring at local scale makes it possible to know the starting and ending point of drought which is useful to develop efficient drought index and to devise a drought policy.

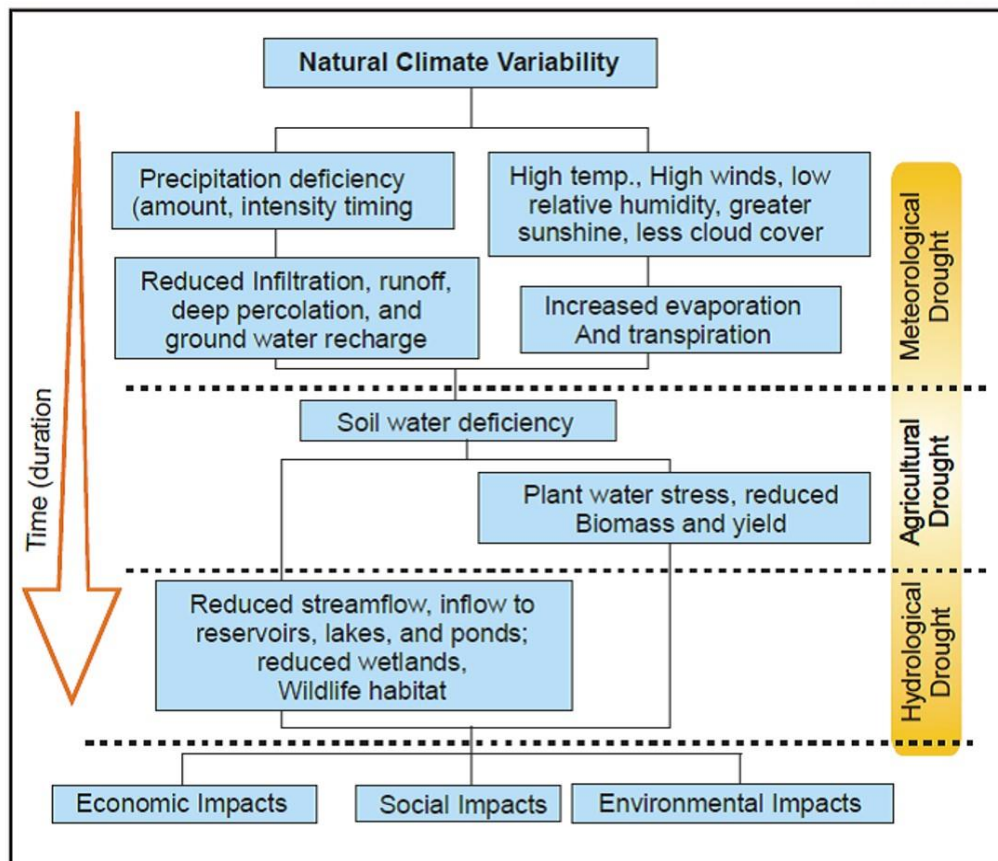


Figure 1.1 Drought propagation in the Anthropocene (Source: National Drought Mitigation Center, University of Nebraska-Lincoln, U.S.A.)

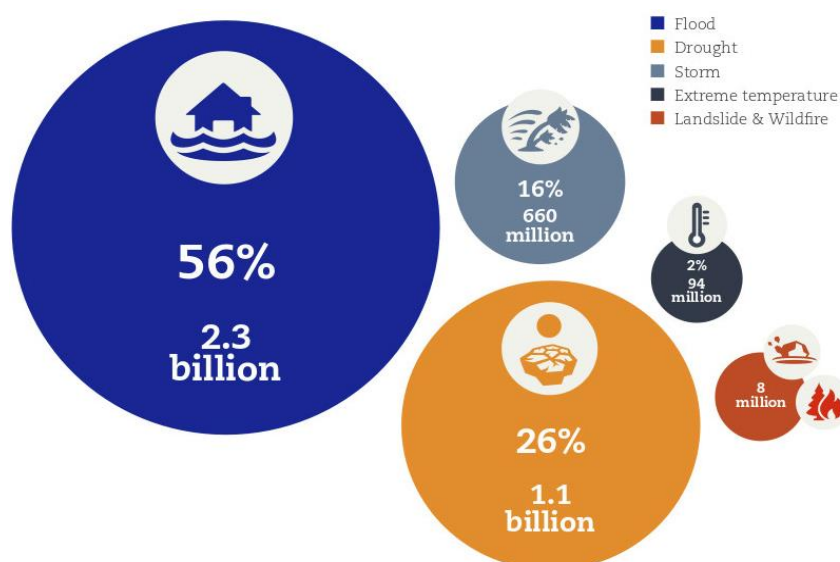


Figure 1.2 Numbers of people affected by water-related disasters (1995-2015) (Source: UNISDR / CRED)

1.2 Remote sensing in drought

1.2.1 Satellite-based drought indices

Satellite-based drought monitoring is very useful to monitor drought at local and regional scale because satellite-based remote sensing data provides space-time continuous information with high resolution (~1 km), which is economic and effective. Many satellites and climate models have provides drought related variables such as land surface temperature (LST), normalized difference vegetation index (NDVI), surface reflectance, evapotranspiration (ET), precipitation, and soil moisture. Many satellite-based drought indices have been developed to monitor and predict drought.

The Evaporative Stress Index (ESI) proposed by Anderson et al. (2011) is one of the widely used drought indices. ESI is calculated using standardized anomalies of the fraction between ET and Potential ET (PET). The index well captures early signals of drought that are detected by rapid soil moisture depletion. Rhee et al. (2011) developed Scaled Drought Condition Index (SCDI) by blending LST, NDVI, and Precipitation. SDCI has been reported that it is useful in monitoring drought in recent studies (Park et al., 2016; Du et al., 2013). Vegetation Drought Response Index (VegDRI) proposed by Brown et al. (2008) is a weekly national drought tool. VegDRI integrate various input data including satellite-based vegetation condition, climate data, land cover, irrigation, elevation, and soils. The index is not sensitive to short-term drought accompanied by high temperature and only focuses on vegetation condition. U.S. Drought Monitor (USDM; Svoboda et al., 2002) incorporates climate, hydrology and soil conditions measurements. In addition, USDM is based on reported impacts and observations from more than 350 contributors around the country. USDM provide the information of drought to policymakers and media. Most of satellite based drought indices and incorporated drought systems have been developed in developed countries including Europe, U.S. and China. Recently, drought monitoring have been studied using satellite based drought indices over developing countries (e.g., Africa, Southeast Asia).

Remote sensing data have been used in drought prediction. Vegetation Outlook (VegOut; Tadesse et al., 2010) incorporates satellite-based vegetation conditions data and climate information with other biophysical information, and predicts drought using Rule-based regression-tree modeling. Rhee et al. (2017) developed a drought forecast model combining remote sensing data and long-range forecast data using machine learning. Although satellite-based drought systems have been developed in recent decades, drought prediction is still challenging. In addition, most of drought monitoring systems are being implemented in developed countries including U.S. and Europe. Remote sensing data can be applied to any region where remote sensing data are available even with limited in situ data availability. Therefore, remote sensing could be useful to monitor and predict drought in developing countries including Africa and Southeast Asia.

1.2.2 Drought Condition Index

Many studies have normalized drought related variables in a given period to develop a drought index because what is considered a drought is relative. Drought conditions should be determined based on the climate of each region. Normalization using the max-min scaling (from 0 to 1) can take into account the potential

maximum and minimum values of the ecosystem. Therefore, this study normalized drought factors by using the max-min scaling approach. The data of each time scale (e.g. pentad, week, and month) at each pixel were scaled using max-min scaling (from 0 to 1) during a given period as discussed in Kogan (1995). As stated in many studies, including Kogan (1995) and Rhee (2010), the values between 0 and 0.4 were considered drought conditions. The criteria of the drought class is shown in Table 1.1.

Table 1.1 The criteria of drought class

Drought class	Values
Extreme Drought	< 0.1
Severe Drought	< 0.2
Moderate Drought	< 0.3
Mild Drought	< 0.4
No Drought	>0.4

1.2.3 Limitation of remote sensing data due to cloud

Remote sensing data has been used in various fields including for studies on the land, atmosphere, and ocean. However, there is a critical limitation to using remote sensing data in surface monitoring due to factors like clouds, haze, and dust, which can make it difficult to obtain valid information because of contaminated/no data pixels. The usability of remote sensing data therefore depends on the weather conditions. Since the drought index is calculated using historical data at each pixel in a given period, no data due to clouds can undermine the validity of the index. There have been many efforts to overcome the no data problem, e.g. by using interpolation, data fusion with climate model data, different sensors (microwave) or time (near time) data. Microwaves, which have a longer wavelength, can penetrate through the cloud cover. Therefore, data fusion with synthetic-aperture radar (SAR) sensor data, which is not affected by weather conditions, has been done to overcome the problem.

1.3 Overview of papers

This thesis seek to 1) better understand the characteristics of drought factors in different climate regions (arid and humid regions), 2) monitor drought effectively using improved soil moisture to high resolution data (~1 km) by developing drought index, 3) predict drought by identifying the relationship between drought and climate variability. Two core chapters (Chapter 2 and 3) have been published and the other core chapters (Chapter 4) is currently being drafted. Below are the full references and summaries:

Chapter 2: **Seonyoung Park**, Jungho Im, Eunna Jang, Jinyoung Rhee. 2016. Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate

regions. *Agricultural and Forest Meteorology*, 216, 157-169.

Drought triggered by a deficit of precipitation, is influenced by various environmental factors such as temperature and evapotranspiration, and causes water shortage and crop failure problems. In this study, sixteen remote sensing based drought factors from the Moderate Resolution Imaging Spectroradiometer (MODIS) and Tropical Rainfall Measuring Mission (TRMM) satellite sensors were used to monitor meteorological and agricultural drought during 2000-2012 growing seasons for different climate regions in the USA. Standardized Precipitation Index (SPI) with time scales from 1 to 12 months and crop yield data were used as reference data of meteorological and agricultural drought, respectively. The relationship between sixteen remote sensing based drought factors and *in situ* reference data was modeled through three machine learning approaches: random forest, boosted regression trees, and Cubist, which have proved to be robust and flexible in many regression tasks. Results showed that random forest produced the best performance ($R^2=0.93$, RMSE=0.3) for SPI prediction among the three approaches. Land surface-related drought factors, e.g., Land Surface Temperature (LST) and Evapotranspiration (ET) showed higher relative importance for short-term meteorological drought while vegetation-related drought factors, e.g., Normalized Difference Vegetation Index (NDVI) and Normalized Multi-band Drought Index (NMDI) showed higher relative importance for long-term meteorological drought by random forest. Six drought factors were selected based on the relative importance by their category to develop drought indicators that represent meteorological and agricultural drought by using the relative importance as weights. While TRMM showed higher relative importance for meteorological drought, LST and NDVI showed higher relative importance for agricultural drought in the arid and humid regions, respectively. Finally, drought distribution maps were produced using the drought indicators and compared with the U.S. Drought Monitor (USDM) maps, which showed a strong visual agreement.

Chapter 3: **Seonyoung Park**, Jungho Im, Sumin Park, Jinyoung Rhee. 2017. Drought monitoring using high resolution soil moisture through multi-sensor satellite data fusion over the Korean peninsula. *Agricultural and Forest Meteorology*, 237, 257-269.

Droughts, typically caused by the deficiencies of precipitation and soil moisture, affect water resources and agriculture. As soil moisture is of key importance in understanding the interaction between the atmosphere and Earth's surface, it can be used to monitor droughts. In this study, a High resolution Soil Moisture Drought Index (HSMDI) was proposed and evaluated for meteorological, agricultural, and hydrological droughts. HSMDI was developed using the 1 km downscaled soil moisture data produced from the Advanced Microwave Scanning Radiometer on the Earth Observing System (AMSR-E) from 2003 to 2011 (March to November) over the Korean peninsula. Seven products from the Moderate Resolution Imaging Spectroradiometer (MODIS) and Tropical Rainfall Measuring Mission (TRMM) satellite sensors were used to downscale AMSR-E soil moisture based on random forest machine learning. The downscaled 1 km soil moisture was correlated well with both *in situ* and AMSR-E soil moisture with the mean coefficient of determination (R^2) of 0.29 and 0.59, respectively. The Standardized Precipitation Index (SPI) with time scales

from 1 to 12 months, crop yields (for sesame, highland radish, and highland napa cabbage) and streamflow data were used to validate HSMDI for various types of droughts. The results showed that HSMDI depicted meteorological drought well, especially during the dry season, with a similar pattern with the 3-month SPI. However, the performance fluctuated a bit during the wet season possibly due to the limited availability of optical sensor data and heterogeneous land covers around the stations. HSMDI also showed high correlation with crop yield data, in particular the highland radish and napa cabbage cultivated in non-irrigated regions with a mean R^2 of 0.77. However, HSMDI did not monitor streamflow well for hydrological drought presenting a various range of correlations with streamflow data (from 0.03 to 0.83).

Chapter 4: **Seonyoung Park**, Eukyo Seo, Deahyun Kang, Jungho Im, Myong-In Lee. 2018. Very short-term prediction of drought using remote sensing data and MJO index through Random forest over East Asia, *Remote sensing of Environment*, to be submitted.

Rapidly developing droughts for the short-period of time including flash drought have frequently occurred over China in recent years, causing significant damage to agricultural ecosystems. Although many drought systems have been developed in recent decades, short-term prediction of drought (within 10 days) is still challenging. This study developed a drought prediction model for the short-period of time (one pentad) using remote sensing data and climate variability index through Random forest over East Asia (20°-50°N, 90°-150°E). Satellite-based drought indices were calculated using the European Space Agency (ESA) Climate Change Initiative (CCI) soil moisture, Tropical Rainfall Measuring Mission (TRMM) precipitation, Moderate Resolution Imaging Spectroradiometer (MODIS) Land Surface Temperature (LST) and Normalized Difference Vegetation Index (NDVI). The real-time multivariate (RMM) Madden-Julian oscillation (MJO) indices were used because the MJO that is a short time scale climate variability and has important implications for the drought in East Asia. The performance of the drought prediction model was evaluated using Root Mean Square Error (RMSE) and correlation coefficient (R). Random forest was performed well when the MJO index was included in the model ($R=0.4$ without the index and $R=0.7$ with the index for validation averaged) for drought prediction. The predicted drought index maps showed similar spatial distribution with actual drought index maps that well captured the sudden changes of drought conditions from normal/wet to dry or dry to normal/wet. Since the developed drought prediction model has not only high resolution (5 km) but a very short time scale (one pentad), it is expected to provide decision makers with accurate information of drought.

1.4 Research highlight

Chapter 2

- Multi-sensor data are used to assess meteorological and agricultural drought.
- Machine learning approaches were used to examine the importance of drought factors.
- Drought indicators are developed to monitor meteorological and agricultural drought.

- Satellite-based drought monitoring is assessed for different climate regions.

Chapter 3

- Surface variables were integrated to produce high resolution soil moisture (1 km).
- Machine learning was used to model high resolution soil moisture.
- High resolution Soil Moisture Drought Index (HSMDI) was proposed.
- HSMDI was evaluated for three types of droughts.
- HSMDI provided good performance to monitor meteorological and agricultural droughts.

Chapter 4

- Drought short-term prediction model was developed using Random forest.
- Very short-term Drought Index (VSDI) was proposed.
- Climate variability, Madden-Julian Oscillation (MJO) has improved the performance of prediction model.

Chapter 2

2. Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions

2.1 Background

Drought, one of the major worldwide natural hazards, causes water shortage which not only increases the vulnerability of the agricultural sector and economic loss, but also greatly threatens human life (Quiring and Papakryiakou, 2003; Wu et al., 2001). Many countries have suffered from drought-induced crop failure and water shortage problems. Unlike other natural disasters such as floods and earthquakes, drought is a slow developing event and it is hard to define its spatial extent and temporal starting and ending points because of its latency (Wilhite, 2000).

Drought can manifest in various ways (Wilhite et al., 2007; Wilhite and Glantz, 1985); meteorological drought occurs due to a lack of precipitation, and prolonged meteorological drought leads to the decrease of soil moisture content that triggers agricultural drought. Hydrologic drought occurs when streamflow, lake, and/or groundwater levels are significantly lower than normal. Factors triggering drought and factors affected by drought are interconnected through couplings and feedbacks in land-atmosphere processes (Rhee et al., 2015). Understanding these factors is very important not only because they can be used to monitor current condition of drought, but also they can be used to predict drought (Tadesse et al., 2005). Drought monitoring and prediction have received attention from policy makers because the results help their responses to drought events (Luo and Wood, 2007). Continuous efforts on drought monitoring and prediction are therefore essential for the sustainable management of water resources and various decision making processes in drought-related risk management (Gao et al., 2014; Zhang et al., 2013).

Drought has been monitored using various drought indices that provide information such as moisture supply conditions and quantitative severity of drought (Alley, 1984; Quiring and Papakryiakou, 2003). Drought monitoring using indices can contribute to various decision making processes and the development of drought early warning systems (Lohani and Loganathan, 1997; Quiring and Papakryiakou, 2003). Thus, development of effective drought indices that reflect various environmental characteristics by climate region is important. There are many drought indices developed using *in situ* precipitation data, such as Standardized Precipitation Index (SPI; McKee et al., 1995), Standardized Precipitation Evapotranspiration Index (SPEI; Vicente-Serrano et al., 2010), Palmer Drought Severity Index (PDSI) and Moisture Anomaly Index (Z-index; Palmer, 1965), and China-Z index (CZI; Wu et al., 2001) to name a few. They generally provide drought condition information using data measured at each weather station. The SPI was recommended as a standard worldwide

meteorological drought index at the 2009 World Meteorological Organization (WMO) meeting (Wardlow et al., 2012). However, these indices are point-based and limited in covering vast areas to show the spatial distribution of drought. Spatial interpolation is required to estimate spatial distribution of drought from station-based data, which often produces high uncertainty in interpolated areas (Brown et al., 2008; Swain et al., 2011). Drought assessment and monitoring using satellite remote sensing can be effective, as satellite data cover large areas at high temporal resolutions (e.g., daily) (Swain et al., 2011). Since satellite remote sensing has been used to provide drought-related products such as Land Surface Temperature (LST), Normalized Difference Vegetation Index (NDVI), and evapotranspiration (ET), it is possible to monitor drought using not only *in situ* measurements at weather stations but also satellite-based drought factors (Anderson et al., 2011).

There is no single indicator that can fully explain the complexity and diversity of drought because drought is caused by a multitude of factors. Blending various indices is thus useful to monitor drought (Hayes et al., 2005; Mizzell, 2008; Wardlow et al., 2012). The blending approaches started in the 1990s (Heim Jr., 2002) and many blended hybrid indices have been developed. Vegetation Health Index (VHI; Kogan, 1995) and Soil Wetness Deficit Index (SWDI; Keshavarz et al., 2014) were developed using LST and NDVI from Moderate Resolution Imaging Spectroradiometer (MODIS). VHI is calculated as a weighted combination of scaled NDVI and LST has been widely used as it provides information of drought conditions compared to the historical range (minimum and maximum) of input data in a given period (Kogan, 1995; Quiring and Ganesh, 2010; Tadesse et al., 2005). SWDI was developed using Soil Wetness Index (SWI) that uses a triangle space concept of hot and cold edges between NDVI and LST (Keshavarz et al., 2014). Microwave Integrated Drought Index (MIDI; Zhang and Jia, 2013) was developed using precipitation from Tropical Rainfall Measuring Mission (TRMM) and soil moisture and LST from Advanced Microwave Scanning Radiometer for EOS (AMSR-E) to monitor meteorological drought. Rhee et al. (2010) developed Scaled Drought Condition Index (SDCI) to monitor agricultural drought using LST, NDVI from MODIS and precipitation from TRMM through linear combination of the products. Some drought indices use not only satellite data but also climate, biophysical and oceanic data to more accurately monitor drought: Vegetation Drought Response Index (VegDRI; Brown et al., 2008), Vegetation Outlook (VegOut; Tadesse et al., 2010), North American Drought Monitor (NADM; Lawrimore et al., 2002) and U.S. Drought Monitor (USDM; Svoboda et al., 2002). However, it is hard to apply these approaches to any region because many regions have limited data availability. Therefore, drought indices that can effectively assess and monitor drought conditions for different climate regions should be developed through blending satellite-based drought factors and parameters. When combining multiple variables, it is crucial to decide what weighting scheme to use. Equal weighting and linear combination approaches have been commonly used because of their simplicity (Kogan, 1995; Rhee et al., 2010; Zhang and Jia, 2013). However, as various factors affect drought differently by region, time, and drought type, more advanced and adaptive weighting approaches should be examined for blending multiple factors to better monitor drought.

In this study, meteorological and agricultural drought was investigated using seven types of satellite-derived drought factors from 2000 to 2012, including LST, NDVI, Normalized Difference Drought Index (NDDI; Gu et al., 2007), Normalized Multi-band Drought Index (NMDI; Wang and Qu, 2007), Normalized Difference Water Index (NDWI; Gao, 1996), and actual ET from MODIS and precipitation data from TRMM. LST, precipitation, and water-related indices are closely related to meteorological drought, while vegetation-related indices, ET, and soil moisture are more related to agricultural drought (Lambert et al., 2013; Long et al., 2013). In particular, ET is highly related with drought conditions because it reflects energy and water exchanges among vegetation, soil, and atmosphere, and considers the characteristics of soil moisture (Allen et al., 2011; Cooke et al., 2012; Rhee et al., 2015). Since those satellite-derived (i.e., MODIS and TRMM) drought factors provide global data (50N-50S), they can be applied to any region which has limited *in situ* data availability.

This study aims at investigating various satellite-derived drought factors to determine how effectively they can be used to monitor meteorological and agricultural drought in different climate regions. The objectives of this research were to 1) predict drought conditions based on the sixteen drought factors using machine learning approaches, 2) identify the relative importance of the drought factors, which varies by region, time, and drought type, 3) explore the feasibility of the development of optimum drought indicators by drought type and climate region, and 4) compare the spatial distribution of satellite-based drought indicators with corresponding USDM drought maps. Multi-time scale SPI (1-month to 12-month) and crop yield (corn and soybean) were used as reference data that represent drought conditions. SPI has been used to represent meteorological drought, which can monitor drought up to 72-month broad time scale and recognize moisture deficiency more rapidly than PDSI (Alley, 1984; Edwards, 1997; Shahabfar and Eitzinger, 2013). Crop yield was used as a surrogate variable representing agricultural drought conditions (Handler, 1990; Lobell et al., 2014; Quiring and Papakryiakou, 2003; Yagci et al., 2015). Three rule-based machine learning approaches—random forest, boosted regression trees, and Cubist—were used to predict drought conditions.

2.2 Study area and data

2.2.1 Study area

Two different climate regions—arid and humid regions—were selected for this study. These two regions have different characteristics in terms of land surface and climate factors such as land cover, precipitation, temperature, and the frequency of drought. While the arid region includes 48 counties of Arizona and New Mexico (Fig. 2.1a), the humid region includes 146 counties of North and South Carolina (Fig. 2.1b). The annual precipitation of the arid region is 323mm. Due to the low amount of precipitation, agricultural lands in the arid region are very limited and less than 4 percent of the area is under cultivation. Major cultivated agricultural crops are corn, cotton, lettuce, pecan, and hay (USDA, 2014). On the other hand, the humid region has high average annual precipitation up to 1,105mm and rich soil ideal for cultivation. Major crops include

soybean, corn, cotton, tobacco, and wheat (USDA, 2014).

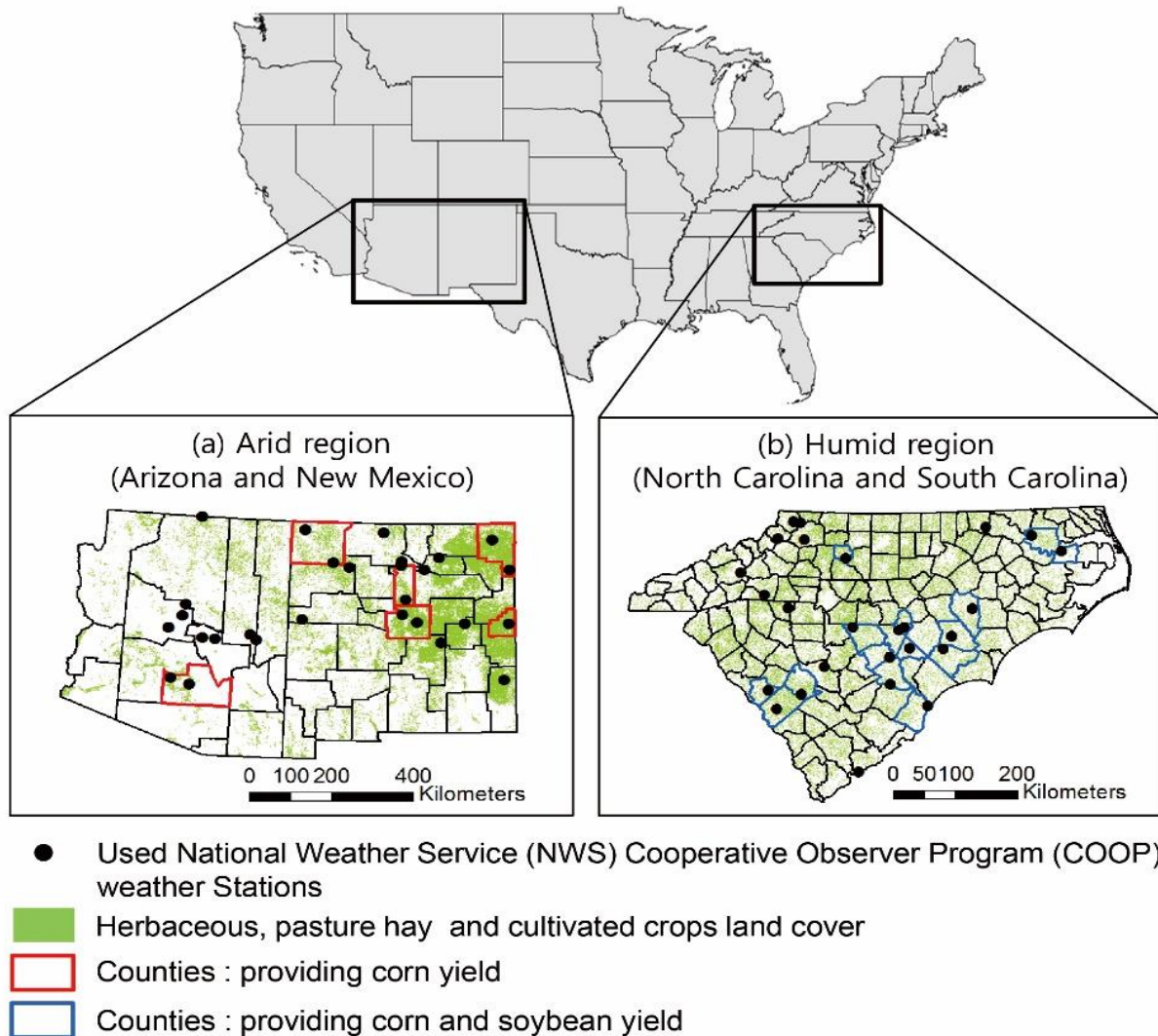


Figure 2.1 Study area includes the arid region of Arizona and New Mexico (left) and the humid region of North Carolina and South Carolina. There are 48 counties in the arid region and 146 counties in the humid region.

2.2.2 Remote sensing data

MODIS data

Drought factors were obtained from MODIS on the Terra platform, launched in December 1999. MODIS provides a multitude of environmental and biophysical products, which are available at [reverb echo](http://reverb.echo.nasa.gov/reverb/) (<http://reverb.echo.nasa.gov/reverb/>). MODIS ET was obtained from Numerical Terradynamic Simulation Group (NTSG) of University of Montana. Monthly data from 2000 to 2012 in growing season (May to September) were used. The monthly MOD16A2 ET and MOD13A3 NDVI with 1km spatial resolution were used. Since the MOD11A2 LST with 1km spatial resolution and the MOD09A1 surface reflectance with 500m spatial resolution were obtained as 8-day composite data, they were converted into monthly information

considering the number of days used in the composite data. Using MODIS surface reflectance variables, NDVI500 (at 500m resolution), NDWI, NDMI, and NDDI were obtained (Table 2.1). To calculate NDWI, each of Shortwave Infrared (SWIR) bands 5, 6 and 7 was used (i.e., NDWI5, NDWI6, and NDWI7).

Table 2.1 Remote sensing based drought factors and their formula.

Drought factors	Formula
NDVI 500	$(\rho_{\text{band2}} - \rho_{\text{band1}}) / (\rho_{\text{band2}} + \rho_{\text{band1}})$
NDWI5,6&7	$(\rho_{\text{band2}} - \rho_{\text{band5(or 6 or 7)}}) / (\rho_{\text{band2}} + \rho_{\text{band5 (or 6 or 7)}})$
NMDI	$(\rho_{\text{band2}} - (\rho_{\text{band6}} - \rho_{\text{band7}})) / (\rho_{\text{band2}} + (\rho_{\text{band6}} - \rho_{\text{band7}}))$
NDDI5,6&7	$(\text{NDVI} - \text{NDWI}) / (\text{NDVI} + \text{NDWI})$
Scaled LST	$(\text{LST}_{\text{max}} - \text{LST}) / (\text{LST}_{\text{max}} - \text{LST}_{\text{min}})$
Scaled ET	$(\text{ET} - \text{ET}_{\text{min}}) / (\text{ET}_{\text{max}} - \text{ET}_{\text{min}})$
Scaled NDVI	$(\text{NDVI} - \text{NDVI}_{\text{min}}) / (\text{NDVI}_{\text{max}} - \text{NDVI}_{\text{min}})$
Scaled NMDI	$(\text{NMDI}_{\text{max}} - \text{NMDI}) / (\text{NMDI}_{\text{max}} - \text{NMDI}_{\text{min}})$; arid region $(\text{NMDI} - \text{NMDI}_{\text{min}}) / (\text{NMDI}_{\text{max}} - \text{NMDI}_{\text{min}})$; humid region
Scaled NDWI	$(\text{NDWI} - \text{NDWI}_{\text{min}}) / (\text{NDWI}_{\text{max}} - \text{NDWI}_{\text{min}})$
Scaled NDDI	$(\text{NDDI}_{\text{max}} - \text{NDDI}) / (\text{NDDI}_{\text{max}} - \text{NDDI}_{\text{min}})$
Scaled TRMM	$(\text{TRMM} - \text{TRMM}_{\text{min}}) / (\text{TRMM}_{\text{max}} - \text{TRMM}_{\text{min}})$

TRMM rainfall data

TRMM provides rainfall data with 0.25 degree resolution. In this paper, the 3B43 monthly product from TRMM and other source data between 2000 and 2012 were obtained from Goddard Earth Sciences Data and Information Center (GES DIC; <http://mirador.gsfc.nasa.gov/>). Data are provided as a precipitation rate with mm/hr and converted into inch/month. In addition, 1-, 3-, 6-, 9-, and 12-month time-scale accumulated TRMM data were used to consider the lagged responses between precipitation and drought.

A total of 16 variables—LST, ET, NDVI, NDVI500, NDMI, NDWI5-7, NDDI5-7, and TRMM1-12—were used as input variables. The remote sensing images were re-projected to the Universal Transverse Mercator Coordinate System. The data of each month from 2000 to 2012 were scaled using max-min scaling (from 0 to 1) by considering the potential maximum and minimum values of the ecosystem at each pixel during the given period (Kogan, 1995). The formula for scaling (Table 1) were established by the correlation result between variables and SPI from Rhee et al. (2010). Scaled values of 0 represent the driest condition, while *values of 1 represent the wettest*. Since LST and NDDI showed negative correlation with SPI in both arid and humid regions and NMDI showed negative correlation in the arid region, the inverse was applied accordingly.

Randomly selected 80% of the samples were used as training data and the remaining 20% were used for validation.

2.2.3 Reference data

Standardized Precipitation Index (SPI)

SPI was used as reference data for monitoring meteorological drought. Monthly total precipitation from 1975 to 2012 for 54 stations (28 stations in the arid region and 26 stations in the humid region)(Fig. 1) excluding Global Historical Climatology Network (GHCN) stations were obtained from National Climate Data Center (NCDC; <http://www.ncdc.noaa.gov/>) because GHCN data is used as input of TRMM (Adler et al., 2003; Rhee et al., 2010). Growing season (May to September) SPIs from 2000 to 2012 were calculated based on the monthly total precipitation. SPI was calculated using a SPI program provided by National Drought Monitoring Center of the University of Nebraska Lincoln (NDMC; <http://drought.unl.edu/MonitoringTools/DownloadableSPIProgram.aspx>). To consider lag time effect between precipitation and drought, different time scales (i.e., from short term to long term scales) of SPI (accumulated 1-, 3-, 6-, 9- and 12- month SPI data) were used.

Crop yield data

Crop yield data from 2000 to 2012 of each county in the humid and arid regions were obtained from the National Agricultural Statistics Service (NASS; <http://www.nass.gov>) of United States Department of Agriculture (USDA). Total crop yield data from survey data were used. Corn yield data were used for the arid region, while corn and soybean yield data were used for the humid region to validate the performance of drought factors to assess agricultural drought. Corn is planted around April, and harvest begins around September. Soybean planting begins in mid-May and harvest runs from November. Irrigation information was identified from 2012 census of agriculture (<http://www.agcensus.usda.gov>). The arid region including Arizona and New Mexico shows a very high irrigation rate (> 94 %), while the humid and temperate regions show a relatively low irrigation rate (< 50%) (Table 2.2). Especially, the irrigation rate of North and South Carolina (i.e., the humid region in the present study) area is under 10%. Therefore, in this study, the arid and humid regions were considered as irrigated and non-irrigated regions, respectively. Fig. 2.1 summarizes counties that provide crop yield data (corn and soybean) in both arid and humid regions.

Table 2.2 Irrigation rates (%) in arid, humid and temperate regions.

	State	Irrigation rate (%)
Arid region	Arizona	94.64
	New Mexico	97.61
	California	97.34
Humid region	Louisiana	40.37
	Mississippi	48.89
	Alabama	6.33
	North Carolina	2.16
	South Carolina	9.59
	Nebraska	49.92
Temperate region	Ohio	0.20
	New York	0.23
	Pennsylvania	0.37
	Colorado	66.36

2.3 Methods

2.3.1 Machine learning approaches

Many statistical methods such as simple and multiple linear regressions have been used to analyze the relationship between drought-related factors and drought conditions (Narasimhan and Srinivasan, 2005; Rhee et al., 2010; Zhang and Jia, 2013). However, it is hard to derive the relationships between a lot of drought-related factors and drought conditions only using linear regression due to their complexity. In this study, three rule-based machine learning approaches—random forest (RF), boosted regression trees (BRT), and Cubist—which have proved to be robust and flexible for many regression applications were employed to examine the relationship between drought factors and drought conditions (i.e., SPI and crop yield), and to identify the relative importance of the variables to develop drought indicators. The three approaches are summarized in the following sections.

Random forest (RF)

RF is based on classification and regression trees (CART; Breiman, 2001) and produces numerous independent trees to reach a final decision through two randomization approaches to the selection of training samples and the selection of variables at each node of a tree. This randomness alleviates typical drawbacks of CART such as an overfitting problem and sensitivity to the training sample configuration (Breiman, 2001). Using out-of-bag (OOB) data from random selection, RF provides internal cross-validation and relative importance of a variable when samples are held in OOB (Kim et al., 2014; Long et al., 2013; Maxwell et al., 2014; Stumpf and Kerle, 2011). RF was implemented by R statistic software (<http://www.r-project.org>), through the random forest package with default settings except the number of trees as 1000 (Kim et al., 2015; Li et al., 2013; Rhee et al., 2014).

Boosted regression trees (BRT)

Similar to RF, BRT is a CART-based ensemble approach. BRT improves the performance of a single CART model by fitting and combining many models. BRT combines the strengths of two algorithms—regression trees and boosting. Especially boosting is used to improve performance by combining a multitude of models to find accurate prediction rules (Elith et al., 2008). While BRT uses the entire training samples to compute the relative importance of variables, RF computes the relative importance of variables using OOB (Li et al., 2014). The relative importance of a variable is scaled from 0 to 100 and estimated by averaging the difference of mean square error (MSE) calculated between before and after permuting a variable at each tree (Froeschke and Froeschke, 2011; Friedman, 2001).

Cubist

Cubist (RuleQuest, 2012), a commercial rule-based machine learning software tool that adopts a modified regression tree, has been used for various drought studies (Brown et al., 2008; Tadesse et al., 2010). Since Cubist produces rules and their associated multivariate regression equations to estimate target values, the result is more interpretable than RF and BRT (Kim et al., 2014; Riggins et al., 2009). As Cubist is not mutually exclusive, final estimation of a target is averaged when multiple rules are applied. Cubist provides information about the variable usage in rules and multivariate regression, which can be used to examine the relative importance of variables (Kuhn et al., 2012).

Comparison of model performance

A total of sixteen drought-related factors and 1-, 3-, 6-, 9- and 12-month SPIs were used as independent and dependent variables in the machine learning models, respectively. This modeling scheme was applied to each of the arid, humid, and combined regions. The performance of each machine learning model was evaluated using the coefficient of determination (R^2) measuring the fitness between actual and predicted values and Root-Mean-Square Error (RMSE) which means sample standard deviation calculated by measuring the difference between actual and predicted values.

2.3.2 Analyses of the variables for drought monitoring

The relative importance of sixteen drought factors for meteorological drought (1-, 3-, 6-, 9- and 12-month SPIs) was identified using three machine learning approaches in the arid, humid and combined regions. As some factors are highly inter-related, a few factors—LST, ET, NDVI, NDWI, NMDI and TRMM precipitation—were selected as those with the highest relative importance of each group of similar factors for further analysis. The relative importance of the selected six factors for meteorological (i.e., uses SPIs as reference) and agricultural (i.e., uses crop yield as reference) drought was again identified using three machine learning approaches in the arid, humid, and combined regions. Drought maps were created using the relative importance of the factors as weights in both arid and humid regions. Spatial patterns of drought conditions were examined through a series of drought maps and their interannual variations were compared with USDM maps which have been widely used in the United States. Fig. 2.2 summarizes the process flow of the proposed method.

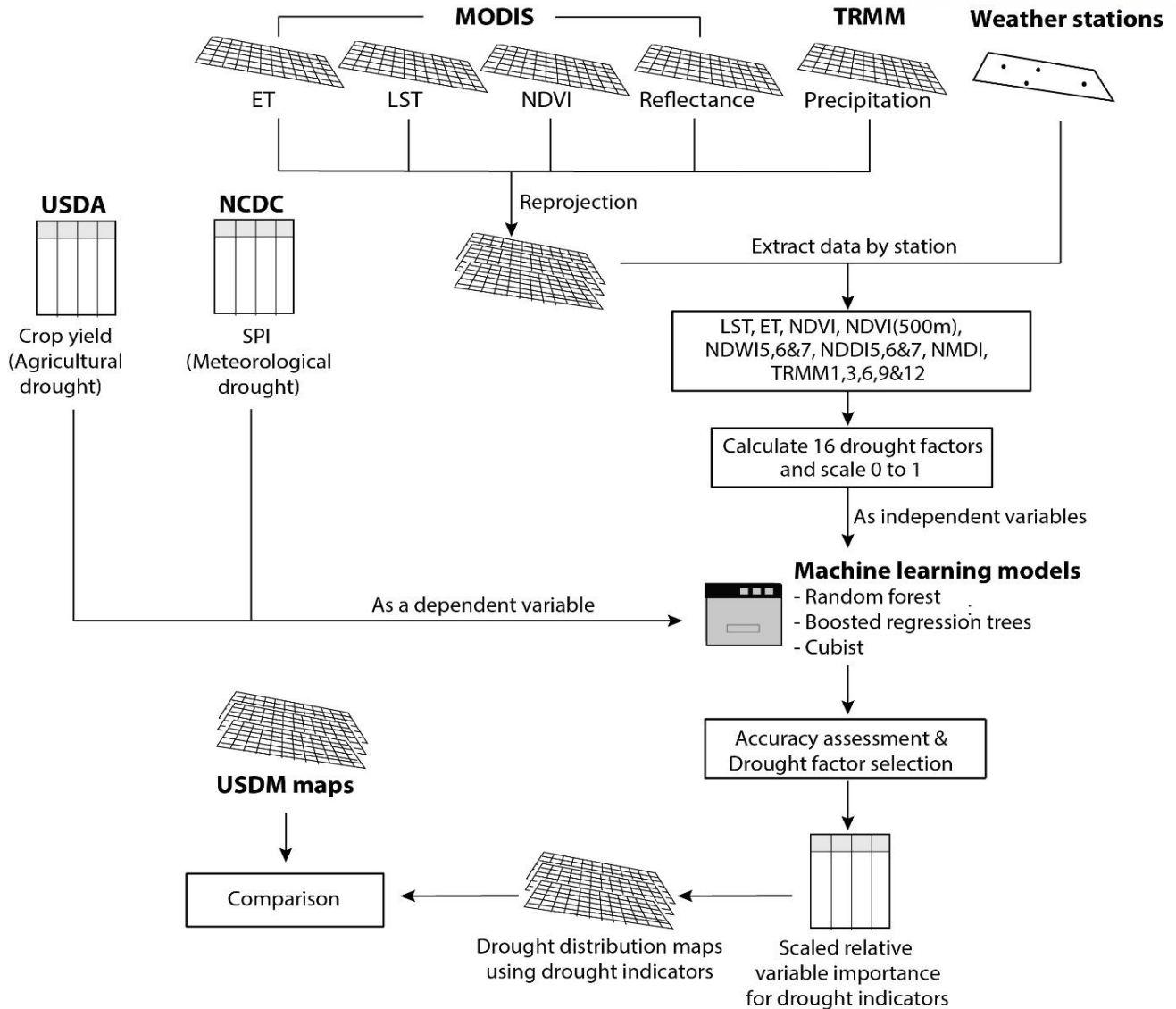


Figure 2.2 Flow chart of the processes used in this study.

2.4 Results and discussion

2.4.1 Monitoring meteorological drought

RF among the three machine learning models showed the best performance with $R^2 > 0.9$ and RMSE between 0.2 and 0.4. Both BRT and Cubist produced similar R^2 (ranging from 0.3 to 0.6) and RMSE (ranging from 0.6 to 0.8) (Fig. 2.3). Fig. 4 shows the relationships between the actual 12-month SPI values and the predicted ones by RF, BRT and Cubist in the arid, humid and combined regions. The range of predicted 12-month SPI from RF was similar with the range of 12-month SPI, while the range of predicted 12-month SPI from BRT and Cubist was much narrower than the range of 12-month SPI, which implies that RF was better trained than the other two. Compared to Cubist, RF uses an ensemble approach for preventing an overfitting problem (Prasad et al., 2006).

Drought conditions of longer time scales (corresponding to 9- and 12-month SPIs) were better

modeled than shorter time scales by all three machine learning approaches (Fig. 2.3) because drought factors including vegetation indices tend to represent the effect of a long-term precipitation shortage (Gessner et al., 2013; Nezlin et al., 2005). There is a reciprocal causal relationship between drought factors and SPI, especially for a long-term meteorological drought.

The model performance of the machine learning approaches varied in the arid, humid, and combined regions (Fig. 2.3). All three machine learning approaches worked better for the arid region than the humid and combined regions to assess drought conditions (i.e., SPIs), as the drought factors representing environmental characteristics of the regions were more separable in the arid region in terms of SPIs. However, there was not a large difference in RF performance among the regions; the RF performed well in all three regions due to its strength from the randomness and ensemble.

Validation was conducted using the remaining 20% of the samples (randomly selected 80% were used as training data). In addition, we also evaluated the models using year-based (i.e., at longer timescale) cross validation (i.e., 13-fold cross validation) as drought is a process dominated by relatively slow variations in precipitation and soil moisture. The RMSE of 13-fold cross validation was calculated by using 12-year data for training (calibration) and remaining 1-year data for test (validation) for 13 years (2000-2012). The 13-fold cross validation results (RMSE ~ 0.84) were not significantly different, but slightly higher than the validation results (RMSE ~ 0.8) based on the split sampling, which implies that overfitting was not an issue for the split sampling.

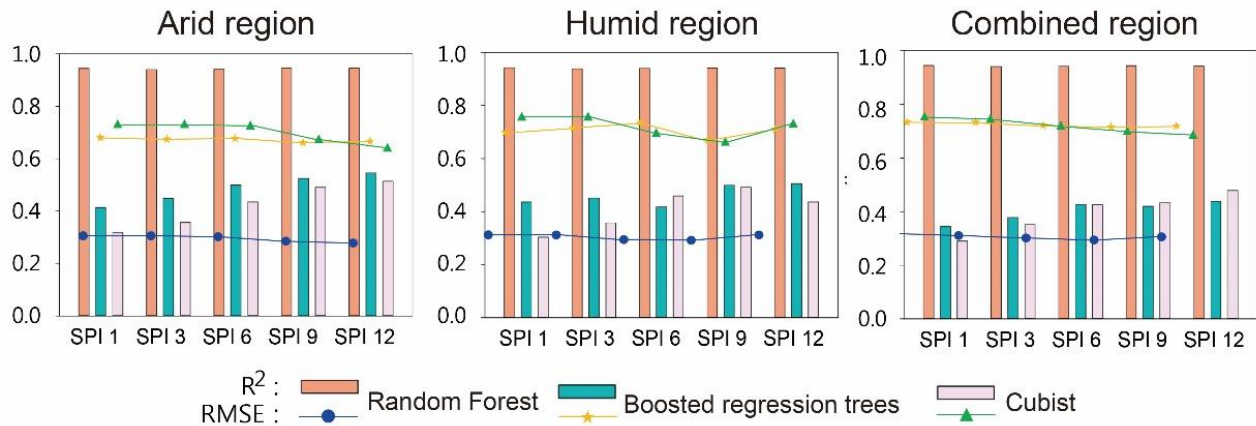


Figure 2.3 Model performance of three machine learning approaches (random forest, boosted regression trees and Cubist) for 1-, 3-, 6-, 9-, and 12-month SPIs in the arid, humid, and combined regions.

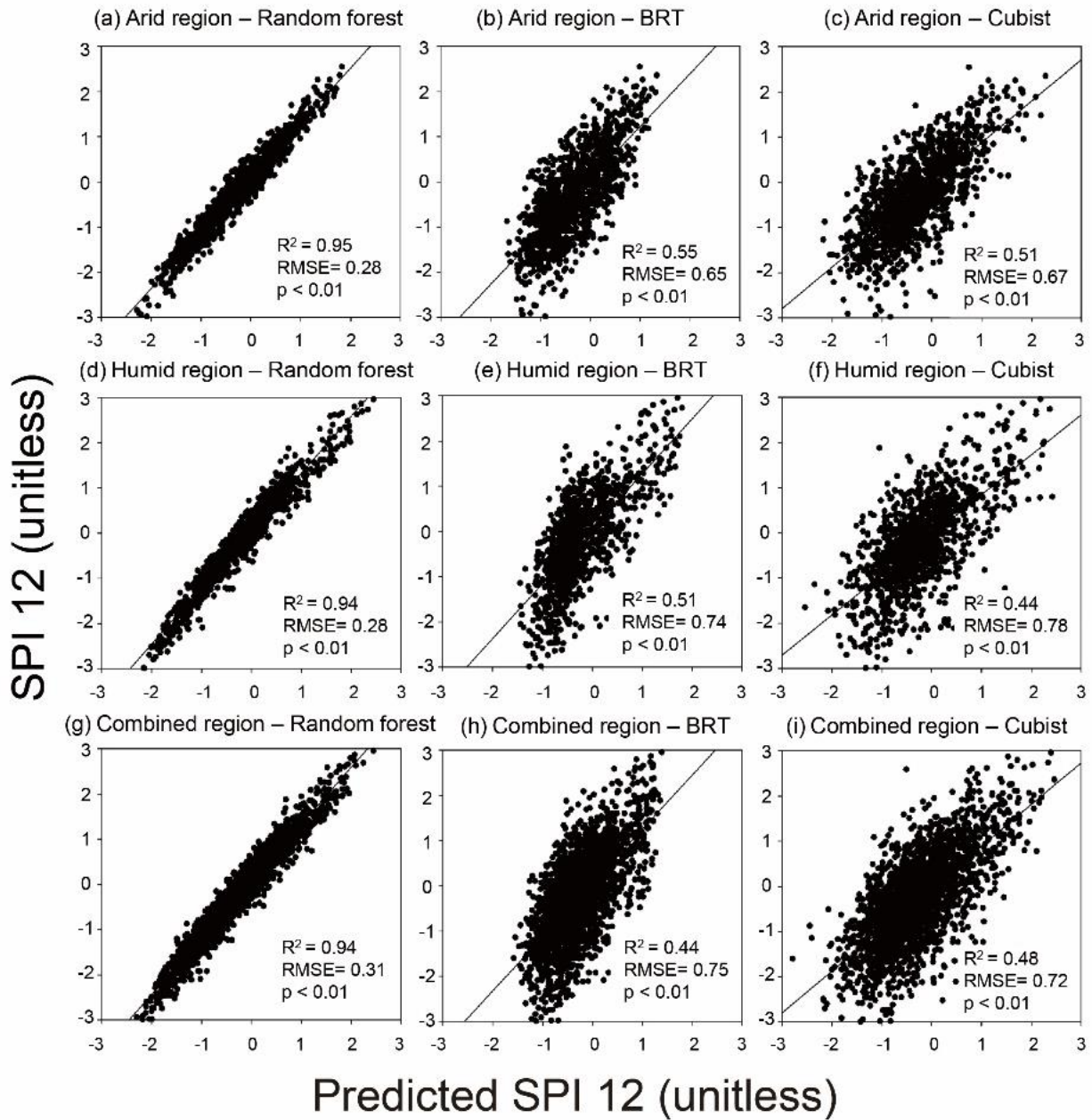


Figure 2.4 Scatter plots between predicted SPI12 based on random forest, boosted regression trees and Cubist and *in situ* SPI12 in the arid (a-c), humid (d-f), and combined regions (g-i).

Importance

While RF provides the relative importance of variables using the increased percentage of mean squared error (MSE) when each variable is held in OOB, the relative importance of variables is given by BRT as the frequency of variables selected for splitting. Cubist provides variable usage information in rules and associated multivariate regression models, which can be used as the relative importance of variables. The relative importance of variables indicates which variables (i.e., drought factors) are more important for estimating SPI

(i.e., meteorological drought) at different time scales in each region. Since each machine learning approach identified relative variable importance in its own way, the important variables identified by approach were slightly different. Except for the large importance of TRMM variables which are highly related to SPI because SPI is calculated based on precipitation, the relative importance of the other variables varied by approach (Tables 2.3 and 2.4). The variable importance by RF and BRT was similar (Table 2.3) due to the similarity of the methods for determining relative variable importance. On the other hand, Cubist provided variable usage information, which was not easy to directly compare with variable importance by RF and BRT.

Tables 2.3 and 2.4 summarize relative importance of the sixteen drought factors for different time scales (i.e., from 1-month to 12-month SPI) when using RF, BRT and Cubist in the arid (Tables 2.3a and 2.4a), humid (Tables 2.3b and 2.4b), and combined (Tables 2.3c and 2.4c) regions. The relative importance of LST and ET decreased with the increasing time scale (Table 2.3 and Fig. 2.5), implying that the surface conditions of LST and ET are more affected by short-term meteorological drought, consistent with the finding of Zhang and Jia (2013). Vegetation index-based variables such as NDVI, NDWI, NDDI, and NMDI showed increasing variable importance as the time scale increases with some variations (Tables 2.3 and 2.4). It implies that there is a time-lag relationship between drought and vegetation responses as discussed in Gessner et al. (2013) and Piao et al. (2003). Lagged responses of vegetation to drought depend on the characteristics of the region (Gessner et al., 2013). The vegetation index-based variables showed high relative importance for long-term meteorological drought (9- and 12-month SPI) in the arid region, while they appeared otherwise in the humid region. Vegetation reacts to water deficits and adapts to limited water availability rapidly in the arid region during short-term droughts, the influence of meteorological drought only appeared for long time scales (more than six month) as discussed by Vicente-Serrano et al. (2013).

Table 2.3. The Relative importance (%) of the most important five drought factors among the sixteen variables for meteorological drought (SPI) using random forest (RF) and boosted regression trees (BRT) in the arid, humid and combined regions for 1-, 3-, 6-, 9-, and 12-month SPI.

a) Arid region

Rank	Relative importance % (random forest and boosted regression trees)									
	SPI1		SPI3		SPI6		SPI9		SPI12	
1	TRMM1 (21.3)	TRMM1 (38.0)	TRMM6 (12.3)	TRMM6 (19.8)	TRMM6 (17.8)	TRMM1 (21.3)	TRMM1 (38.0)	TRMM6 (12.3)	TRMM6 (19.8)	TRMM6 (17.8)
2	TRMM6 (10.3)	TRMM3 (12.0)	TRMM3 (9.8)	NDWI6 (13.0)	NDWI6 (9.6)	TRMM6 (10.3)	TRMM3 (12.0)	TRMM3 (9.8)	NDWI6 (13.0)	NDWI6 (9.6)
3	TRMM3 (9.6)	TRMM6 (11.6)	NDWI6 (8.1)	TRMM3 (12.6)	NDWI7 (9.3)	TRMM3 (9.6)	TRMM6 (11.6)	NDWI6 (8.1)	TRMM3 (12.6)	NDWI7 (9.3)
4	ET (8.7)	LST (8.5)	NDWI7 (7.2)	ET (9.0)	NDDI6 (9.1)	ET (8.7)	LST (8.5)	NDWI7 (7.2)	ET (9.0)	NDDI6 (9.1)
5	LST (6.4)	ET (6.3)	NDDI6 (7.1)	NDVI (8.7)	NDVI (7.8)	LST (6.4)	ET (6.3)	NDDI6 (7.1)	NDVI (8.7)	NDVI (7.8)

b) Humid region

Rank	Relative importance % (random forest and boosted regression trees)									
	SPI1		SPI3		SPI6		SPI9		SPI12	
1	TRMM1 (25.4)	TRMM1 (50.0)	TRMM3 (13.2)	TRMM3 (27.2)	TRMM6 (15.0)	TRMM6 (40.7)	TRMM9 (15.2)	TRMM9 (47.8)	TRMM12 (21.3)	TRMM12 (55.4)
2	LST (10.1)	LST (14.1)	TRMM1 (11.5)	TRMM1 (16.9)	TRMM9 (10.6)	TRMM9 (11.9)	TRMM12 (11.9)	TRMM12 (10.8)	TRMM9 (8.6)	TRMM1 (7.8)
3	TRMM3 (7.4)	TRMM6 (4.8)	TRMM6 (9.4)	LST (11.0)	TRMM3 (8.4)	TRMM1 (9.8)	TRMM6 (9.6)	TRMM1 (9.3)	TRMM6 (7.1)	LST (5.7)
4	TRMM6 (7.0)	NDVI500 (4.0)	TRMM9 (8.8)	TRMM9 (9.1)	TRMM12 (8.1)	LST (8.8)	TRMM1 (8.2)	LST (5.9)	NDDI6 (6.8)	NDVI (4.3)
5	TRMM9 (6.4)	TRMM3 (2.9)	TRMM12 (8.4)	NDWI7 (5.3)	TRMM1 (8.0)	NDVI (3.8)	TRMM3 (6.7)	NDVI (3.8)	TRMM1 (6.8)	NMDI (3.9)

c) Combined region

Rank	Relative importance % (random forest and boosted regression trees)									
	SPI1		SPI3		SPI6		SPI9		SPI12	
1	TRMM1 (13.2)	TRMM1 (52.7)	TRMM6 (14.1)	TRMM3 (28.2)	TRMM6 (12.2)	TRMM6 (45.9)	TRMM9 (15.2)	TRMM9 (47.1)	TRMM12 (13.2)	TRMM12 (54.3)
2	TRMM3 (12.3)	LST (13.1)	TRMM3 (10.3)	TRMM6 (14.8)	TRMM9 (10.0)	TRMM9 (9.7)	TRMM12 (11.9)	TRMM12 (9.1)	ET (12.1)	TRMM6 (7.4)
3	TRMM6 (9.5)	TRMM3 (7.9)	NDWI7 (8.1)	NDWI6 (9.4)	TRMM12 (7.6)	LST (7.0)	TRMM6 (9.6)	TRMM6 (8.1)	TRMM6 (7.6)	NDWI7 (5.8)
4	LST (8.5)	TRMM6 (5.4)	NDWI6 (7.4)	TRMM1 (9.3)	NDWI7 (7.4)	NDWI7 (6.4)	TRMM1 (8.2)	NDWI6 (4.7)	NDVI (6.6)	NDVI (5.7)
5	ET (6.9)	ET (2.8)	ET (6.9)	LST (9.3)	ET (6.9)	NDWI6 (5.4)	NDDI7 (5.5)	NDVI (4.7)	NDWI6 (6.4)	LST (4.2)

Table 2.4. The variable usage information in conditions and models (%) of the most important five drought

factors among the sixteen variables for meteorological drought (SPI) using Cubist in the arid, humid and combined regions for 1-, 3-, 6-, 9-, and 12-month SPI.

a) Arid region

Rank	Cubist (usage in conditions and models, %)									
	SPI1		SPI3		SPI6		SPI9		SPI12	
1	TRMM6 (89)	TRMM6 (100)	NDWI6 (100)	NDWI6 (100)	TRMM6 (100)	TRMM6 (79)	TRMM9 (100)	TRMM9 (100)	TRMM12 (100)	TRMM12 (100)
2	TRMM3 (63)	TRMM3 (89)		NDWI7 (100)		NDWI7 (100)	NDVI500 (33)	NDVI500 (93)	TRMM6 (19)	TRMM6 (100)
3	TRMM1 (51)	TRMM1 (89)		TRMM3 (100)		LST (100)	TRMM6 (26)	TRMM6 (88)		NDWI6 (100)
4		TRMM9 (100)		TRMM6 (100)		NDWI6 (79)		NDWI6 (100)		NDDI6 (100)
5		LST (89)		TRMM9 (100)		NDDI6 (79)		TRMM3 (100)		NDVI500 (86)

b) Humid region

Rank	Cubist (usage in conditions and models, %)									
	SPI1		SPI3		SPI6		SPI9		SPI12	
1		LST (100)	TRMM3 (100)	TRMM3 (100)	TRMM6 (82)	TRMM6 (100)	TRMM9 (97)	TRMM9 (100)	TRMM12 (100)	TRMM12 (100)
2		NDVI500 (100)		LST (100)	TRMM9 (53)	TRMM9 (65)	TRMM1 (51)	TRMM1 (100)	TRMM1 (53)	TRMM1 (53)
3		NDWI7 (100)		NDVI500 (100)		LST (100)	NDDI6 (13)	NDDI6 (95)		NDDI5 (100)
4		TRMM1 (100)		NDWI7 (100)		NDWI7 (100)	TRMM6 (9)	TRMM6 (93)		NDVI500 (100)
5		TRMM6 (100)			TRMM1 (47)	TRMM1 (47)		NDDI7 (93)		NMDI (100)

c) Combined region

Rank	Cubist (usage in conditions and models, %)									
	SPI1		SPI3		SPI6		SPI9		SPI12	
1	TRMM1 (100)	TRMM1 (100)	TRMM3 (90)	TRMM3 (100)	TRMM6 (88)	TRMM6 (100)	TRMM9 (100)	TRMM9 (98)	TRMM12 (73)	TRMM12 (97)
2	TRMM3 (54)	TRMM3 (100)	NDWI6 (88)	NDWI6 (100)	NDWI7 (69)	NDWI7 (100)	NDWI6 (83)	NDWI6 (83)	NDWI6 (87)	NDWI6 (69)
3		LST (100)	TRMM1 (69)	TRMM1 (59)	TRMM1 (66)	TRMM1 (54)	TRMM3 (74)	TRMM3 (83)	TRMM9 (41)	TRMM9 (100)
4		NDVI500 (100)		NDVI500 (100)	LST (19)	LST (88)		NDDI6 (100)	NDDI6 (74)	NDDI6 (49)
5		NDWI7 (100)		TRMM6 (100)		NDVI500 (96)		NDWI7 (98)	LST (17)	LST (74)

Fig. 2.5 shows the variations of the relative importance of the sixteen drought factors in the combined region based on RF with multiple time scales of meteorological drought (i.e., 1-, 6-, and 12-month SPI). A lack

of precipitation is instantaneously connected to surface conditions with increasing LST and decreasing ET. It can trigger the loss of water content in soil, decreasing soil moisture for a long period that harms the growth of vegetation (Rhee et al., 2015). TRMM precipitation-derived variables are highly related to meteorological drought that results from a lack of precipitation. The relative importance of TRMM1 precipitation was the highest for 1-month SPI and decreased with increasing time scales of SPI consistent with Zhang and Jia (2013). The relative importance of TRMM6 precipitation had a peak for 6-month SPI. TRMM12 precipitation was also the most important variable for 12-month SPI. In estimating SPI, the TRMM precipitation with the same time scale was the most suitable.

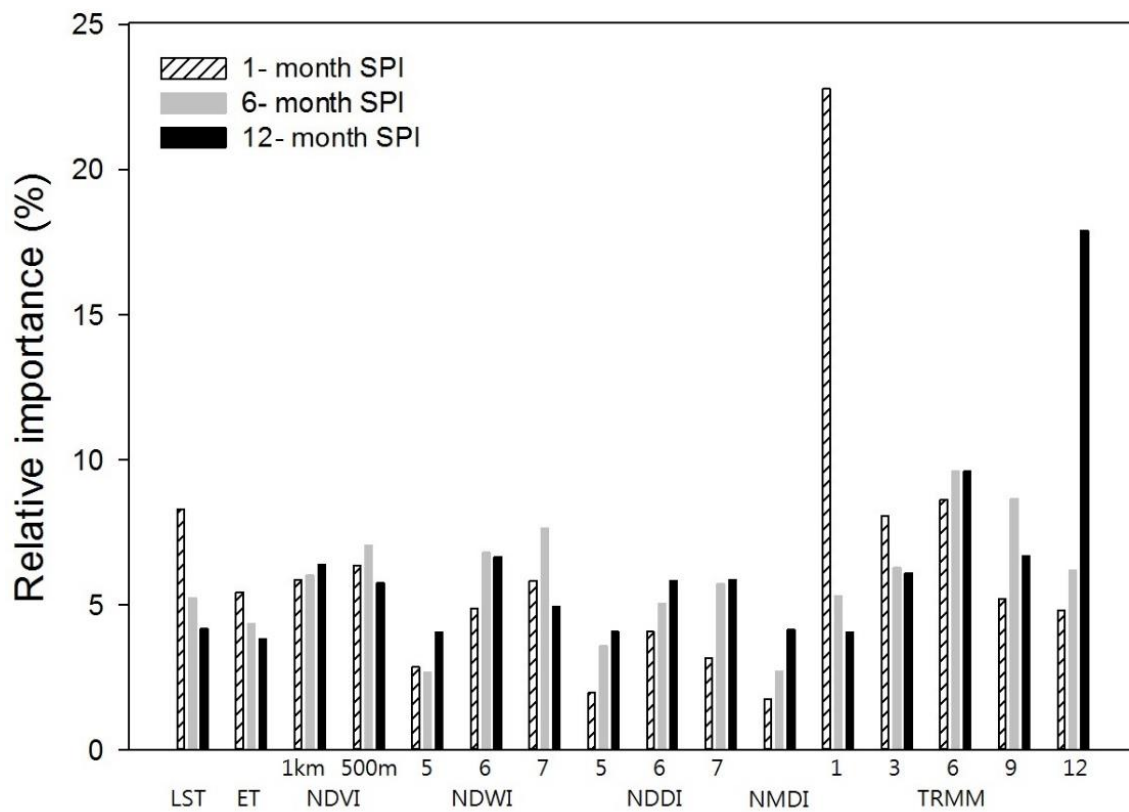


Figure 2.5 The relative importance (%) of the sixteen drought factors using random forest for 1-, 6-, and 12-month SPI.

Since some variables among the sixteen drought factors are inter-related, redundant variables are better to be avoided when developing optimum drought indicators through blending of the variables. All three MODIS SWIR channels (i.e., bands 5, 6, and 7) were used to calculate NDWI and NDDI (Table 2.1). While NDWI6 and NDWI7 were generally more important than NDWI5, NDDI6 was more important than NDDI5 and NDDI7 (Tables 2.3 & 2.4 and Fig. 2.5). Thus, NDWI6 and NDDI6 were selected as representative NDWI and NDDI to monitor drought, respectively. Similarly, NDVI (1km) product was selected to represent NDVI (Tables 2.3 & 2.4 and Fig. 2.5). Since NDDI was calculated using NDWI and NDVI (Table 2.1) and the relative importance of NDDI was lower than NDWI and NDVI (Fig. 2.5), NDDI was excluded. Finally six variables,

LST, NDVI, NDWI, NMDI, ET, and TRMM were selected to develop drought indicators for monitoring both meteorological and agricultural drought.

Table 2.5 presents the scaled relative importance of the selected six variables for meteorological drought using RF that produced the best performance among three machine learning models. The selected six drought factors show similar patterns of relative importance with Tables 2.3 and 2.4. The relative importance of vegetation-related drought factors had a peak for 3-month SPI. The relative importance of vegetation-related drought factors in the arid region was higher than in the humid region for each time scale, while the relative importance of LST and TRMM in the humid region was higher than in the arid region for each time scale.

Table 2.5. The scaled relative importance of the selected six variables for 1-, 3-, 6-, 9-, and 12-month SPI and model performance (R^2 and RMSE) in the arid (A), humid (H) and combined (C) regions. The sum of the relative importance is 1.

Drought factors	Target feature														
	SPI 1			SPI 3			SPI 6			SPI 9			SPI 12		
	C	A	H	C	A	H	C	A	H	C	A	H	C	A	H
LST	0.17	0.13	0.21	0.14	0.14	0.16	0.12	0.12	0.14	0.10	0.12	0.12	0.08	0.06	0.08
NDVI	0.12	0.14	0.06	0.16	0.16	0.10	0.11	0.13	0.04	0.12	0.15	0.07	0.13	0.13	0.10
NDWI	0.12	0.11	0.07	0.17	0.21	0.10	0.13	0.17	0.08	0.11	0.17	0.06	0.11	0.17	0.06
NMDI	0.04	0.04	0.02	0.05	0.08	0.07	0.06	0.09	0.04	0.06	0.10	0.06	0.07	0.12	0.06
ET	0.08	0.19	0.07	0.04	0.13	0.04	0.09	0.11	0.06	0.07	0.10	0.07	0.07	0.10	0.05
TRMM	0.47	0.39	0.57	0.44	0.28	0.53	0.49	0.38	0.65	0.54	0.36	0.62	0.54	0.42	0.65
R^2	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92	0.92
RMSE	0.37	0.37	0.36	0.37	0.36	0.38	0.36	0.36	0.35	0.36	0.35	0.35	0.36	0.34	0.37

2.4.2 Monitoring agricultural drought

To monitor agricultural drought, the relationships between the selected six variables and crop yield data (corn and soybean) were identified using RF. The impact of drought on corn and soybean can be serious, and the most affected crops by severe droughts during July 2012 in Indiana, USA were corn and soybean (more than 70%) (Mallya et al., 2013). The relative importance of drought factors for corn and soybean yield can be very useful to monitor agricultural drought. Table 2.6 presents the scaled relative importance of the six variables for crop yield with the R^2 , RMSE, and relative RMSE (rRMSE) values. The rRMSE divides RMSE by the average of crop yield in the arid (irrigated) and humid (non-irrigated) regions. Since the crop yield depends on the region and the type of crop, rRMSE can provide relatively unbiased error estimation. Based on R^2 and rRMSE, the modeling result for corn yield in the arid region was better compared to the other cases (Table 2.6). In the arid region, NDVI was the most important variable for estimating corn yield, while LST was the most important variable for estimating corn and soybean yield in the humid region (Table 2.6). Whereas the TRMM variables showed higher relative importance than the other variables for meteorological

drought (Table 2.5), the most important variables for crop yield were NDVI (in the arid region) and LST (in the humid region). The results were consistent with Johnson (2014) that predicted corn and soybean yield using MODIS LST and NDVI.

Table 2.6 The scaled relative importance of the selected six variables for corn and soybean in the arid (irrigated) and humid (non-irrigated) regions and model performance (R^2 , RMSE and rRMSE). The sum of the relative importance is 1.

Drought factors	Target feature	
	Arid (Irrigated)	Humid (Non-irrigated)
	Corn	Corn Soybean
LST	0.15	0.29 0.29
NDVI	0.22	0.07 0.13
NDWI	0.05	0.11 0.07
NMDI	0.22	0.29 0.16
ET	0.20	0.11 0.15
TRMM 3	0.16	0.13 0.20
R^2	0.94	0.94 0.94
RMSE	9.87	14.95 3.04
rRMSE	0.05	0.18 0.12

The relationships between crop yield and SPIs (from 1-month to 12-month) were also examined, and the relative importance of SPIs for estimating crop yield was identified using RF (Fig. 2.6 for the humid (non-irrigated) region). Unlike our expectation that long-term meteorological drought (9- and 12-month SPI) would be more highly related to crop yield, 3-month SPI (19.59%) showed the highest relative importance for estimating crop yield (Fig. 2.6). Mallya et al. (2013) used short term (1-month SPI) for monitoring Midwest drought in the United States to analyze its agricultural and socioeconomic impact. Quiring and Ganesh (2010) demonstrated that 6-month SPI had a significant relationship with vegetation condition indices, showing that between 1- and 6-month SPIs were more important to agricultural drought than SPIs with longer time scales.

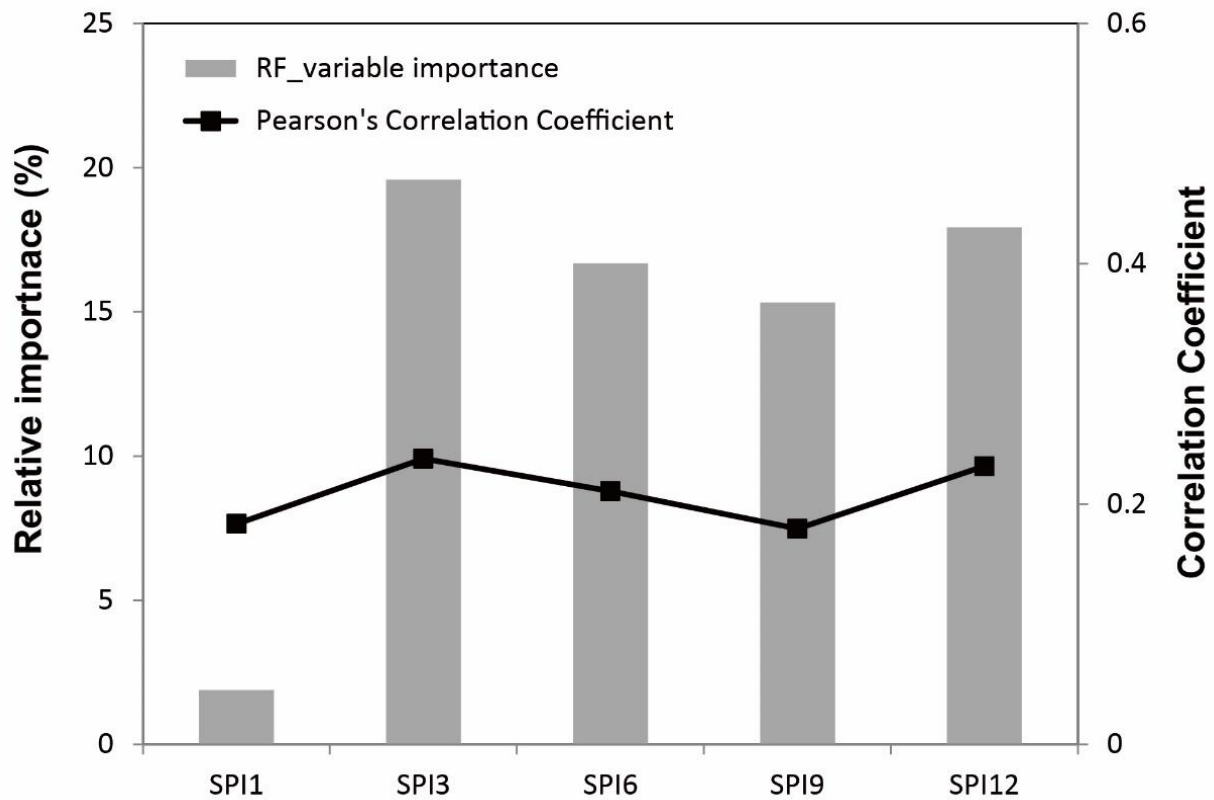


Figure 2.6 The relative importance (%) of SPIs with 1- to 12-month time scales using random forest for estimating crop yield and Pearson's correlation coefficients between each SPI and crop yield in the non-irrigated region (humid region).

The relationship between crop yield and remote sensing-derived factors varied with the crop phenology. For both corn and soybean, remote sensing-based factors two- to three-month before harvesting produced the highest correlation with crop yield (August and September for soybean and July for corn), which implies that flowering season is the most important period that determines crop yield. The scaled relative importance values of the six variables for corn and soybean in the arid and humid regions (Table 2.6) were used as weights, and the weighted combinations of the six variables were calculated and used as drought indicators to monitor agricultural drought.

2.4.3 Comparison with USDM maps

The spatial distribution of drought conditions was mapped over the arid and humid regions based on RF-derived weights of the selected six drought factors. Two types of drought maps were created: one type of maps shows the distribution of meteorological drought conditions based on the scaled relative importance of the six drought factors for estimated 3-month SPI (Table 2.5). 3-month SPI was used because it showed the highest correlation with agricultural drought (crop yield) (Fig. 2.6). The other type of maps presents the distribution of agricultural drought (crop yield) conditions estimated using the scaled relative importance of the six drought factors for estimating crop yield in the arid (irrigated) and humid (non-irrigated) regions (Table 2.6). Averaged weights from the results of corn and soybean yield were used for the humid region.

The spatial distribution of drought conditions were compared to 2000-2012 USDM maps that integrated PDSI, SPI, soil moisture, and satellite-based VHI (Svoboda et al., 2002). Since drought factors used in this study are based on global remote sensing data (e.g., MODIS and TRMM) which covers from 50S to 50N, the proposed approach can be applied to any vegetated region in the world. Figs. 2.7 and 2.8 show drought maps based on the six drought factors on May and September except for rainy season (June, July, and August) from 2000 to 2012 in the arid and humid regions, respectively.

In the arid region, spatial pattern of the two types of remote sensing based drought maps (3-month SPI and crop yield) and USDM maps were similar in most cases except 2001, 2008, 2009, and 2010 (Fig. 2.7). During the two years of 2001 and 2010, two remote sensing based maps present the drought status more severely compared to the USDM maps. Since the drought factors were scaled using the maximum and minimum values for the 13 years to represent relative conditions of drought for the given period, the difference of the base period between remote sensing based maps and USDM maps as well as the difference of input data may have caused the differences. Two remote sensing based maps (3-month SPI and crop yield) agreed each other with minor differences in 2007 and 2011 (Fig. 2.7).

In the humid region, two types of remote sensing-based maps caught well the most extreme drought in 2002 and 2007 (Fig. 2.8). The spatial distributions of drought in 2002, 2005, and 2007-2011 were rather consistent with the USDM maps, while the spatial distributions of drought in 2000, 2003-2004 and 2006 were quite different from the USDM maps (Fig. 2.8) with estimating more severe drought status. Both types of remote sensing based drought maps (3-month SPI and crop yield) showed similar spatial distributions of drought with minor differences in 2001, 2005, 2008, 2010, and 2011 (Fig. 2.8). Compared to the estimated 3-month SPI maps, the estimated crop yield maps resulted in more similar spatial distributions with the USDM maps in 2001, 2003-2006, and 2010-2011 (Fig. 2.8); it is possibly because that the estimation of SPI is dominantly affected by precipitation while the estimation of crop yield is affected by many other factors including vegetation indices and temperature.

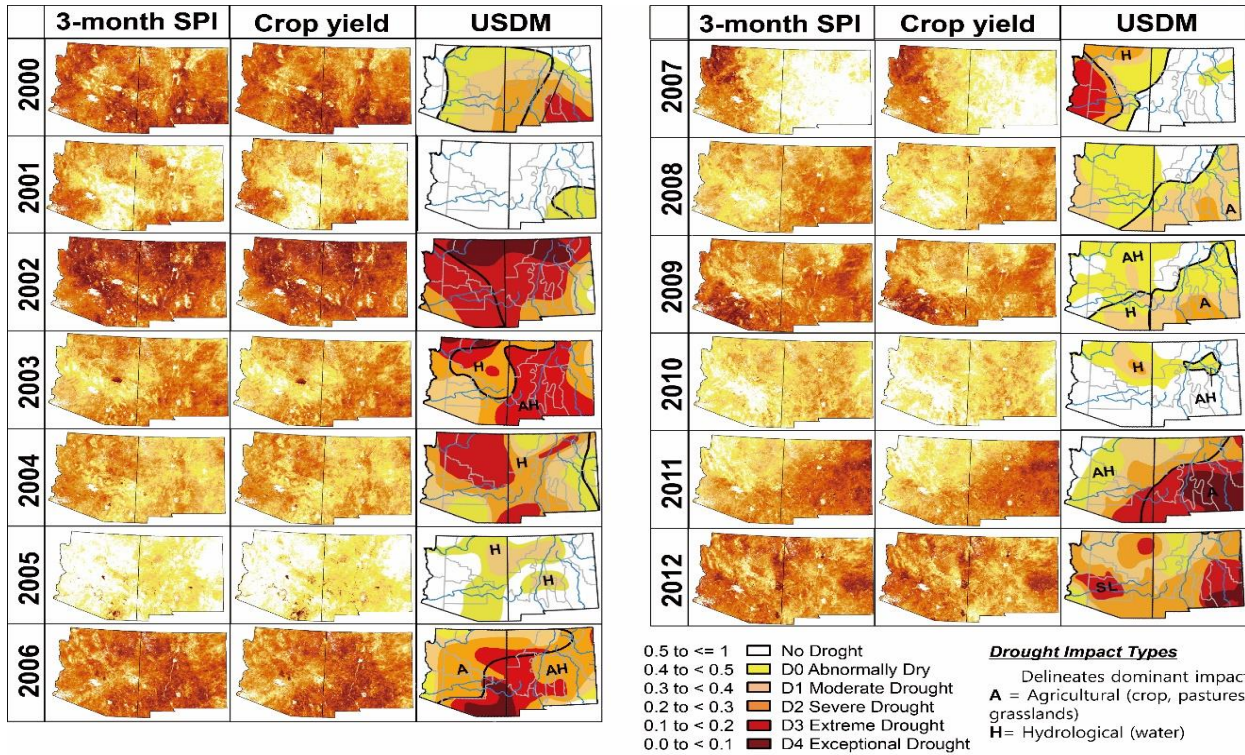


Figure 2.7 Year-to-year change between two types of remote sensing based drought maps (3-month SPI and crop yield) and USDM maps in the arid region for May from 2000 to 2012.

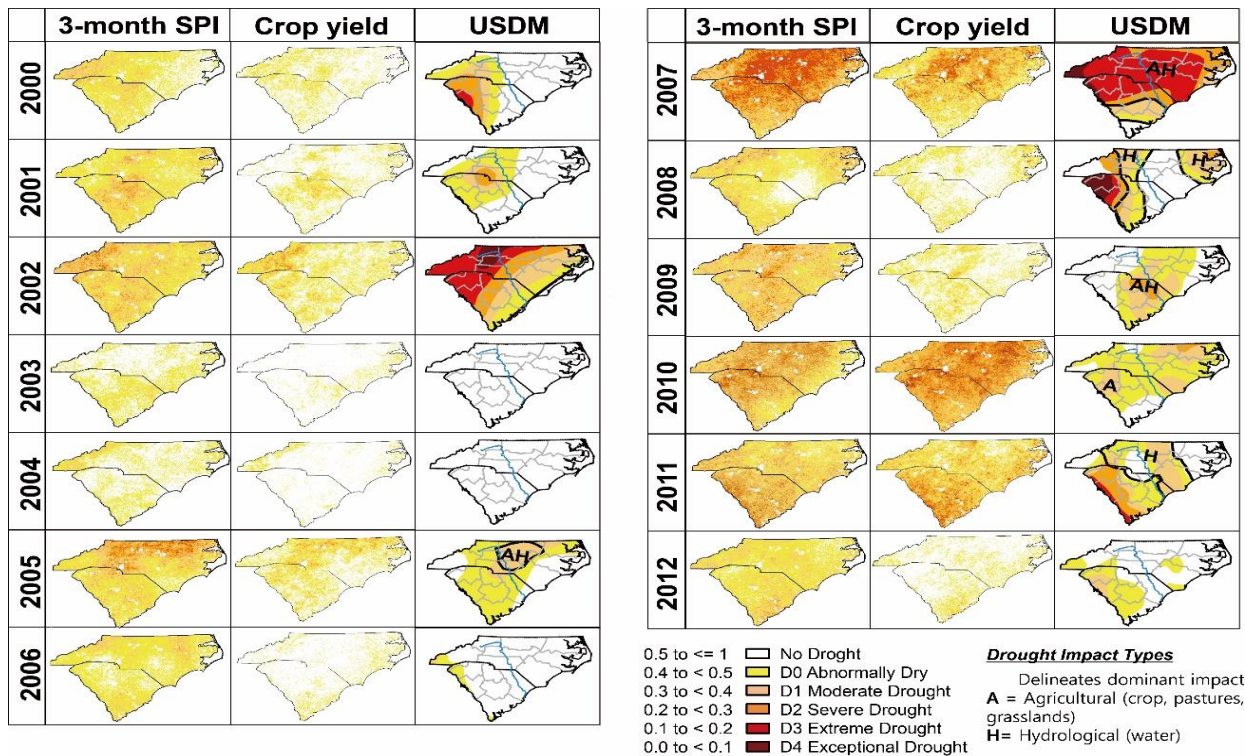


Figure 2.8 Year-to-year change between two types of remote sensing based drought maps (3-month SPI and crop yield) and USDM maps in the humid region for September from 2000 to 2012.

Validation

Weight factors for the selected six variables (LST, NDVI, NDWI, NMDI, ET, and TRMM3) were determined by relative importance using RF. The relative importance of the variables was identified using thirteen-year data (2000-2012). Those weight factors were applied to 2013 and 2014 data and the results were compared with USDM (Fig. 2.9). In the arid region, both types of remote sensing-based maps matched well extreme drought in 2013 and 2014 when compared to the USDM data (Fig. 2.9). The estimated crop yield maps showed a bit more similar spatial distribution of drought with USDM compared to the three-month SPI maps (Fig. 2.9). In the humid region, both types of remote sensing based drought maps estimated the drought status more severely as in Fig 8, while the distribution of drought in 2014 was consistent with the USDM map (Fig. 2.9). Similar to the arid region, the estimated crop yield maps showed more similar drought status with the USDM data than the 3-month SPI maps did (Fig. 2.9). Thus, the proposed weights for the factors based on the importance are considered reasonable and can be used to monitor drought effectively.

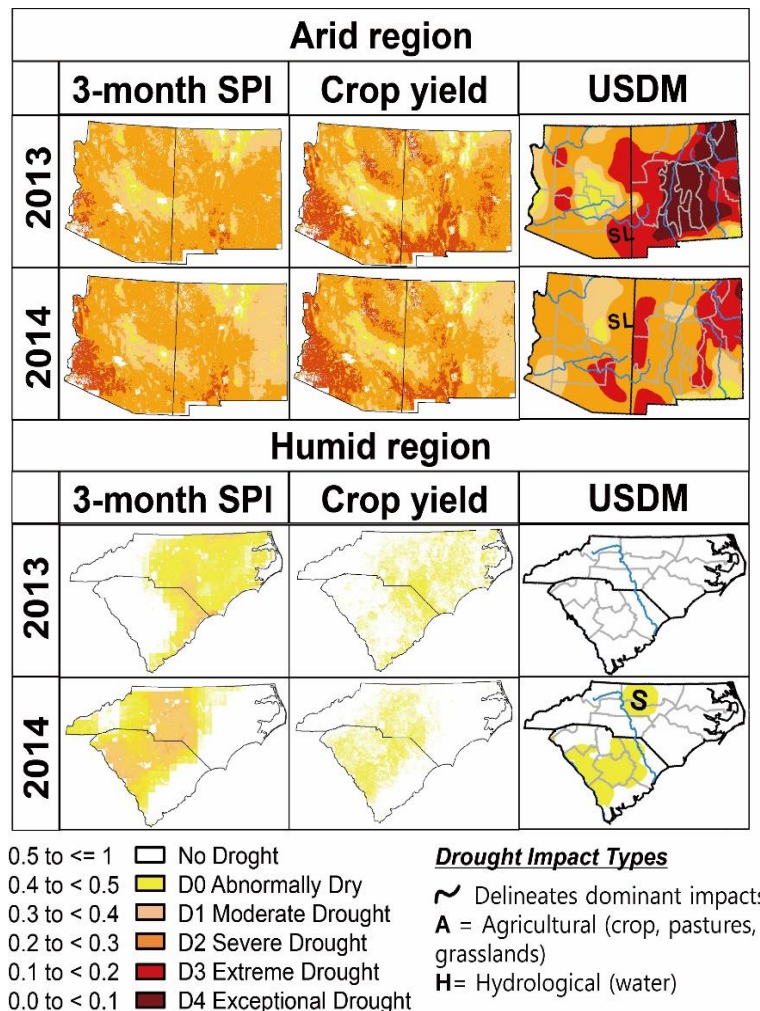


Figure 2.9. Validation drought monitoring map (2013 and 2014) in the arid (irrigated) and humid (non-irrigated) regions.

2.5 Conclusion

In this study, sixteen drought factors were analyzed using three machine learning approaches targeting a meteorological drought index (SPI) and an agricultural drought index (crop yield). The characteristics of drought factors for meteorological and agricultural drought were examined using relative importance of variables derived from the machine learning approaches. Among the three machine learning methods, RF outperformed the other approaches and was used to select the most important six variables (LST, NDVI, NDWI, NMDI, ET, and TRMM) with the consideration of the inter-relation of the variables based on relative variable importance. The modeling results based on the relative importance of the six variables in terms of 3-month SPI and crop yield were shown in the maps to compare and validate spatial distributions of drought to USDM maps. The crop yield based maps showed more consistent distributions with USDM maps compared to the 3-month SPI based maps. The approach proposed in this study can be applied to any vegetated region where remote sensing data are available even with limited *in situ* data availability.

However, there are some limitations which should be improved in the future. Crop yield samples were limited due to the study area selection. In future study, agricultural drought modeling will be tested in regions with more dominant corn and soybean agriculture (i.e., the Midwest) and with different dominant crops (e.g., wheat) to more generalize parameterizations in the proposed approach. As the current thresholds (e.g., 0.5 for normal condition) to determine drought conditions are arbitrary, specific standards and guidelines on thresholds should follow. Future research includes the incorporation of soil moisture and evapotranspiration into development of drought indicators, and the development of remote sensing-based optimized approaches to different types of drought (i.e., agricultural drought and hydrological drought).

Chapter 3

3. Drought monitoring using high resolution soil moisture through multi-sensor satellite data fusion over the Korean peninsula

3.1 Background

Droughts occur almost every year, and drought-induced losses in the US are estimated to be 6-8 billion dollars per year (Witt, 1997). Droughts are driven by an absence or deficiency of precipitation and their impact on the agricultural sector is crucial for the economy at local and country levels (IPCC, 2007; Park et al., 2016; Lessel et al., 2016; Ryu et al., 2014). Since it is difficult to define the spatial and temporal starting and ending points of droughts, unlike other disasters including typhoons and earthquakes (Park et al., 2016; Wilhite, 2000), drought monitoring and assessment at multiple time scales with high resolution are essential for drought mitigation. In particular, high-resolution soil moisture data are necessary to identify and assess the geographical extent of drought and to monitor flash drought (Otkin et al., 2016). This provides decision makers with accurate information of drought conditions especially for agricultural uses (Zipper et al., 2014; Luo and Wood, 2007; Tadesse et al., 2005). Droughts are broadly categorized as meteorological, agricultural, hydrological, and socioeconomic droughts (Heim, 2002; Wilhite et al., 2007). Meteorological droughts occur due to lack of precipitation, and perennial meteorological droughts cause soil moisture reduction which triggers agricultural droughts that lead to a decrease of crop yields caused by the lag time responding to insufficient precipitation (Park et al., 2016; Potop et al., 2012; Park et al., 2005). Hydrological droughts refer to low stream flows and reservoir levels as well as a deficiency of water supply when compared to normal conditions. Thus, hydrological droughts have an even longer lag time in response to meteorological conditions, which act more slowly than meteorological and agricultural droughts (Choi et al., 2013; Logan et al., 2010). Socioeconomic droughts are driven by an insufficient supply to satisfy societal water demands along with the above three drought forms (Zhang et al., 2013).

Soil moisture is a key factor in drought monitoring because of its relationship with precipitation, evapotranspiration (ET), and crop yield (Zhu et al., 2017; AghaKouchak et al., 2015; Fleisher et al., 2015; Bolten et al., 2010), which are crucial information associated with droughts (Im et al., 2016; Choi et al., 2013; Hogg et al., 2013). Soil moisture is essential for understanding the interaction between the Earth's surface and atmosphere as well as the hydrological cycle (Molero et al., 2016; Finn et al., 2011; Salvucci et al., 2002). Soil moisture interacts with many environmental variables such as land ET, albedo, land surface temperature (LST), and precipitation, which are key processes of the climate system (Im et al., 2016; Djamai et al., 2015; Dobriyal et al., 2012). Soil moisture highly influences plant growth, including agriculture and natural vegetation, as the dynamics of soil moisture determines water resources available in the agroecosystem (Gao et al., 2014; Hunt

et al., 2014; Dobriyal et al., 2012;).

Soil moisture is generally obtained through in situ measurements which provide accurate readings with high temporal resolution. Soil moisture can be acquired from the International Soil Moisture Network (ISMN; <http://www.ipf.tuwien.ac.at/insitu>), which provides global in situ soil moisture data (Dorigo et al., 2011). However, it is difficult to represent the spatial distribution of soil moisture from point-based measurements and the cost of field measurements is high (Grayson and Western, 1998; Zhu et al., 2017; Im et al., 2016). Remote sensing approaches, such as satellite passive and active microwave sensor data, provide spatiotemporally continuous soil moisture (top soil moisture, 3-5 cm). Soil moisture has been retrieved from many satellite instruments since the 1970s, including Scanning Multichannel Microwave Radiometer (SMMR), Advanced Microwave Scanning Radiometer on the Earth Observing System (AMSR-E), Soil Moisture and Ocean Salinity sensor (SMOS), the Advanced Scatterometer (ASCAT), and Soil Moisture Active Passive (SMAP) (Martínez-Fernández et al., 2015; Choi and Hur, 2012; Li et al., 2015; Champagne et al., 2012; Yee et al., 2016; Mallick et al., 2009).

Some drought indices have been developed through satellite-based soil moisture because drought is closely related with soil moisture and satellites provide spatiotemporally continuous soil moisture data. Hao et al. (2013) proposed the Multivariate Standardized Drought Index (MSDI). MSDI is produced by applying the copula function to the Standardized Precipitation Index (SPI) and Standardized Soil moisture Index (SSI) produced using 0.5 degree soil moisture and precipitation data acquired from National Oceanic and Atmospheric Administration (NOAA) National Weather Service (NWS) Climate Prediction Center (CPC). Global drought monitoring has been conducted using MSDI through the Global Integrated Drought Monitoring and Prediction System (GIDMaPS) from University of California (<http://drought.eng.uci.edu/>). The Microwave Integrated Drought Index (MIDI), which monitors meteorological drought in semiarid regions, was proposed by Zhang et al. (2013). MIDI is calculated by combining the normalized Tropical Rainfall Measuring Mission (TRMM) precipitation, normalized AMSR-E soil moisture, and normalized MODIS LST. MIDI showed the highest correlation with 1-month SPI, although the correlation decreased as the time scale of SPI increased. The U.S. Drought Monitor (USDM; Svoboda et al., 2002) is based on six key physical indicators including soil moisture and supplementary indicators (e.g., the Palmer Crop Moisture Index (CMI; Palmer, 1968) and the Keetch Bryam Drought Index (KBDI; Keetch and Byram, 1968)). USDM is a well-developed drought index which can detect meteorological, agricultural, and hydrological droughts (Svoboda et al., 2002).

Soil moisture based drought indices are very useful to monitor drought. However, it is difficult to monitor drought at a local or regional scale using global soil moisture data with a coarse spatial resolution (~10 to 40 km) because even a pixel (~10 to 40 km) typically includes complex topography and heterogeneous land cover types (Werbylo et al., 2014). In addition, coarse resolution soil moisture data are not much useful for local or regional applications of agriculture and water resources since they do not provide detailed variations of soil moisture (Djamai et al., 2016). Although global soil moisture-based drought indices may provide useful

information for drought over vast areas, it is necessary to use high resolution soil moisture for identifying spatially detailed drought conditions and for making relevant decisions from the perspective of managers and decision makers. There are several downscaling approaches including DISaggregation based on Physical And Theoretical scale CHange (DISPATCH; Merlin et al., 2013), the smoothing filter-based intensity modulation (SFIM) downscaling technique (Parinussa et al., 2013), and machine learning approaches (Im et al., 2016). This study uses the soil moisture downscaling approach proposed by Im et al. (2016), which generated 1 km soil moisture using MODIS surface products (LST, ET, vegetation indices, and albedo) and TRMM precipitation through machine learning.

The purpose of this study is to develop a drought index named the High resolution Soil Moisture Drought Index (HSMDI) using downscaled soil moisture through multi-sensor satellite data fusion for monitoring meteorological, agricultural, and hydrological droughts. The specific objectives of this study are to 1) downscale AMSR-E soil moisture based on six MODIS indicators and TRMM precipitation using machine learning, 2) validate the downscaled soil moisture using in situ data, 3) develop HSMDI by normalizing downscaled soil moisture for a given period, and 4) evaluate HSMDI through comparison with station-based multi-time scale SPI for meteorological drought, crop yield (i.e., sesame, highland radish, and napa cabbage) as reference data for agricultural drought, and streamflow data which represents hydrological drought conditions. Since the soil moisture downscaling in this study was conducted based on surface variables such as vegetation indices, LST, precipitation, albedo, and ET, this new drought index can contain various characteristics related to drought.

3.2 Study area and data

3.2.1 Study area

The Korean Peninsula, including both North and South Korea (latitude 34°N to 43°N and longitude 124.5°E to 131°E), was selected for this study (Figure 3.1). These two contiguous regions have slightly different climate characteristics by latitude. The annual mean temperature is about 10.7°C in North Korea and 14.2°C in South Korea. North Korea, with an average annual precipitation of 596.3 mm is drier than South Korea (1196.4 mm). Most of the rain falls in summer during the Asian monsoon. The major land cover categories of the two combined regions are forest (57.4%) and croplands (35.1%) (Figure 3.1).

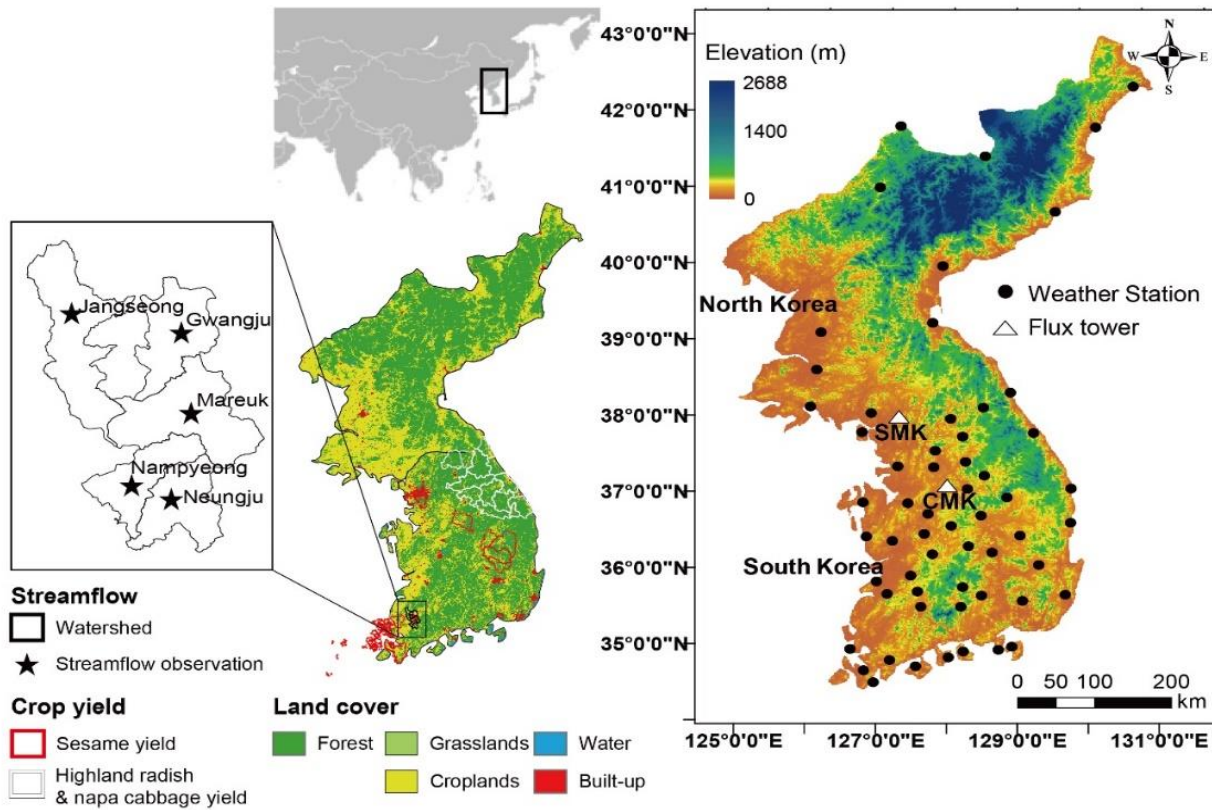


Figure 3.1. Study area with land cover, elevation, and *in situ* reference data locations.

3.2.2 Satellite data

AMSR-E soil moisture data

The AMSR-E on the Aqua satellite, developed by the Japanese Aerospace Exploration Agency (JAXA), provides global soil moisture data derived from a land-surface radiative transfer model using microwave brightness temperature. There are many algorithms used to retrieve soil moisture data: NASA National Snow and Ice Data Center (NSIDC; Nijoku et al., 2003), JAXA (Koike et al., 2004), the United States Department of Agriculture (USDA; Jackson, 1993) and the Vrije Universiteit Amsterdam (VUA) in collaboration with NASA (VUA–NASA; Owe et al., 2001). In this study, the X-band AMSR-E JAXA Level3 soil moisture (g/cm^3) with daily ascending (13:30) overpass (0.25 degree resolution) was used from 2003 to 2011 (March to November) because Korea has similar climate conditions with Japan and both are affected by westerlies. The dataset was obtained from JAXA (<https://gcom-w1.jaxa.jp/>).

MODIS data

MODIS onboard both Terra and Aqua provides numerous environmental and biophysical products that can be used in various applications. Six MODIS products—the 8-day LST (MYD11A2), evapotranspiration (ET; MOD16A2), albedo (MCD43B3), Leaf Area Index (LAI; MCD15A2), and 16-day

NDVI & EVI (MYD13A2)—with 1 km resolution were obtained from reverb echo (<http://reverb.echo.nasa.gov/reverb/>) and the Numerical Terradynamic Simulation Group (NTSG) at the University of Montana for the three tiles (h27v04, h27v05, h28v05) covering the study area. MODIS Aqua products having the same pass time (13:30) as AMSR-E were used to model soil moisture except for ET product which is only available from Terra. NDVI and EVI (16-day) were converted into 8-day data by simply averaging the data for two 16-day data. These 8-day six factors from 2003 to 2011 (March to November) were used to downscale soil moisture.

TRMM rainfall data

TRMM, which launched in November 1997, provides rainfall data with a 0.25 degree spatial resolution. In this study, daily rainfall data (3B42) between 2003 and 2011 (March to November) were downloaded from the Goddard Earth Sciences Data & Information Center (GES DIC; <http://mirador.gsfc.nasa.gov/>). These daily rainfall data were converted into 8-day rainfall data through combining daily rainfall for 8 days.

3.2.3 Observation data

Ground soil moisture data were obtained from two Korean regional flux monitoring networks (KoFlux), Seolmacheon (SMC) and Cheongmicheon (CMC), for the period from March to November in 2010 (drought year). KoFlux hourly soil moisture (m^3/m^3) at 10 cm depth data were obtained from the Hydrological Survey Center (<http://www.hsc.re.kr/>). The KoFlux data at 13:00 were used to validate the downscaled soil moisture and the AMSR-E soil moisture. Both the SMK site (126.95°E and 37.93°N) and the CMK site (126.65°E and 37.15°N) are located over a mixed forest (Figure 3.1). A total of 19 stations from SMC and 17 stations from CMK were averaged. KoFlux soil moisture is measured using Time Domain Reflectometry (TDR), which is used all over the world to measure soil moisture through travel time analysis (Lin, 2003).

Standardized Precipitation Index (SPI; McKee et al., 1995) is a widely used meteorological drought index based on precipitation measurements at stations at a given time-scale (e.g., 1 month), using more than 30 years of historical data. SPIs for March to November (2003-2011) were calculated using precipitation data measured at 29 stations located over vegetated areas from 1973 to 2011 which were obtained from the Korea Meteorological Administration (KMA; <https://data.kma.go.kr/>) (Figure 1). In this study, multi-scale SPIs—1-, 3-, 6-, 9-, and 12-month SPIs—were used to consider the time lag effect between precipitation and drought.

Crop yield data was used as reference data for evaluating HSMDI in terms of agricultural drought. This study used the yield data of sesame (2006-2011), and highland radish and highland napa cabbage (2007-2011) for each administrative district. Yearly crop yield data were obtained from Statistics Korea (KOSTAT; <http://kostat.go.kr/>). Figure 3.1 shows the locations of the administrative districts. Sesame is planted mid-May and harvested at the end of August. Highland radish and cabbage are planted between May and June, and harvesting begins at the end of August.

Streamflow was used as reference data for hydrological drought. Streamflow data measured at five hydrometric stations (Jangseong, Gwangju, Mareuk, Nampyeong, and Neungju), which are located near

agricultural regions and operated by the South Korean government were used (Figure 3.1). Monthly averaged streamflow data during the study period (from 2003 to 2011) was downloaded from the Korean Water Resource Management Information System (WAMIS; <http://www.wamis.go.kr>). Since the range of streamflow varies by month, each month would most likely have different mean and standard deviation values of streamflow. Therefore, it will be very difficult to compare streamflow for different months in terms of drought. Streamflow in each month from 2003 to 2011 was converted to normalized anomaly using the min-max scaling. A scaled value of one represents the maximum streamflow value of each month during those nine years, while a value of zero represents the minimum streamflow value of each month during the nine years.

3.3 Methods

A flowchart for producing HSMDI is shown in Figure 3.2. The process consists of two parts: 1) downscaling soil moisture using 1 km MODIS and TRMM variables through the random forest (RF) machine learning approach, and 2) developing HSMDI by scaling the downscaled soil moisture.

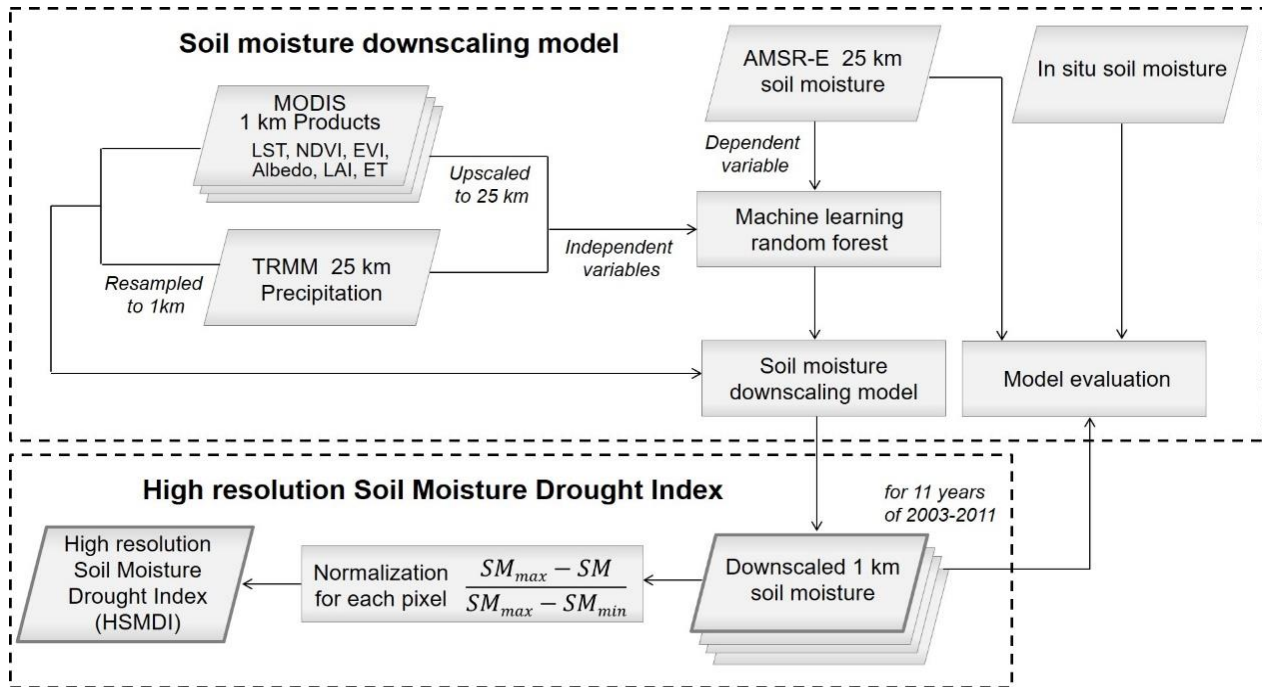


Figure 3.2. Data processing flow diagram to develop the High resolution Soil Moisture Drought Index (HSMDI) based on a multi-sensor downscaling approach.

3.3.1 Downscaling soil moisture

In this study, the soil moisture downscaling approach proposed by Im et al. (2016) was used. Im et al. (2016) downscaled 25 km AMSR-E soil moisture to 1 km soil moisture using 1 km MODIS variables through machine learning approaches. Six MODIS variables (LST, NDVI, EVI, LAI, albedo, and ET) from 2003 to 2011 were upscaled by averaging pixels to 25 km, which is the same resolution as AMSR-E soil moisture. The upscaled 25 km MODIS variables were extracted for the corresponding AMSR-E soil moisture grids and used

as training samples for developing downscaling models. Machine learning approaches—RF, Cubist, and Boosted Regression Trees (BRTs)—were used to model upscaled 25 km MODIS variables (independent variables) and AMSR-E soil moisture (the dependent variable). Six MODIS 1 km variables were applied to the developed machine learning models, and 1 km soil moisture was then produced. Im et al. (2016) showed the highest model prediction accuracy when using the RF approach, and the downscaled soil moisture showed overall good agreement with *in-situ* measurements, resulting in R^2 of 0.52 and Relative Mean Square Difference (RMSD) of 0.09 m³/m³. In the present study, TRMM precipitation was added to the six MODIS variables and the RF method was used to model soil moisture (Figure 3.2). Three seasonal models (spring, summer, and autumn) were developed because Korea has four distinct seasons, and soil moisture has very high uncertainty when the surface is frozen or covered with snow in winter (Holmes et al., 2009).

The RF is based on classification and regression trees (CART; Breiman, 2001), one of decision tree algorithms. Rule-based machine learning such as decision trees and RF has been widely used in remote sensing applications (Im et al., 2009; 2012; Jensen and Im, 2007; Kim et al., 2015; Li et al., 2013; Lu et al., 2014; Rhee et al., 2008; Torbick and Corbiere, 2015). The RF uses an ensemble approach from multiple decision trees (typically 500 to 1000) is used to make a decision for prediction. The name “random forest” refers to data prediction accomplished using many independent decision trees (a “forest”) through randomly selected training samples and variables at each node, which alleviates the well-known problems of CART such as overfitting and sensitivity to training data. A randomly selected subset of training samples is used to produce a tree. A final decision from multiple trees is made through the aggregation of individual tree results based on an averaging approach for regression or a majority voting for classification. RF calculates the increased percentage of mean square error (MSE) using out-of-bag (OOB) data when a variable is permuted with a random value (Ke et al., 2016; Kim et al., 2014; Rhee et al., 2014; Tang et al., 2015). Based on this information, the relative importance of a variable (i.e., contribution of the variable to predict a target variable) can be identified.

3.3.2 High resolution Soil Moisture Drought Index (HSMDI)

HSMDI was developed using the downscaled 1 km soil moisture data produced using MODIS and TRMM products and the random forest machine learning approach. An HSMDI is calculated through normalization using the maximum and minimum soil moisture values at each pixel (j_{th}) from the same day of year (DOY), every 8-day, during the given period (2003-2011) using Eq.1. This scaling approach is based on McVicar and Jupp (1998), which refined the VCI proposed by Kogan (1995) to develop a monthly vegetation condition index (MVCI). A normalized soil moisture anomaly can consider potential minimum and maximum values of the ecosystem and express the status of observed values (Kogan 1995; McVicar and Jupp 1998; Rhee et al., 2010; Park et al., 2016). The HSMDI identifies areas which are drier (represented by 0) or wetter (represented by 1) than their usual condition. The HSMDI values below 0.5 were considered being in drought because 0.5 indicates the median status for 9 years. Although the proposed HSMDI is similar with soil moisture-based drought indices introduced by Zhang et al. (2013) and Sánchez et al. (2016), HSMDI provides

high resolution (1 km) soil moisture-based drought information integrated from various related surface variables.

$$\text{HSMDI}_{j,\text{DOY}} = (\text{SM}_{j,\text{DOY}} - \text{SM}_{\min,\text{DOY}}) / (\text{SM}_{\max,\text{DOY}} - \text{SM}_{j,\min}) \quad (3.1)$$

This study aims to develop and validate the drought index, HSMDI, using meteorological, agricultural, and hydrological reference data based on observations. The relationships and the time series patterns between HSMDI and multi-timescale SPIs were analyzed for meteorological drought. The relationship between HSMDI and the yearly crop yield was examined using regression analysis for agricultural drought. Monthly HSMDI in the selected watersheds was also validated for hydrological drought using the scaled stream flow data.

3.4 Results and Discussion

3.4.1 Assessment of downscaled soil moisture *in situ* data

Table 3.1 shows the results of the three seasonal downscaling models with calibration coefficient of determination (R^2) and root-mean-squared-error (RMSE). The three seasonal models produced similar R^2 values (ranging from 0.91 to 0.95), RMSEs (ranging from 0.032 to 0.038 m^3/m^3), and slopes (ranging from 0.60 to 0.67). In the spring and autumn models, LST (178%) and albedo (143%) appeared as more important variables, while ET (122%) as well as albedo (132%) and LST (131%) appeared as more contributing variables than the others in the summer model. Spring and autumn have similar meteorological characteristics in North and South Korea including mean seasonal precipitation and temperature. Thus, it is not surprising to have similar results from the spring and autumn models (Table 3.1). Since the seasonal characteristics are similar in spring and autumn, the two models showed the same high-ranked variables of importance (Table 3.1). The reason why LST and albedo showed higher variable importance among the seven variables is that they are the most important factors for surface energy flux; they are thus often used in studies to downscale soil moisture (Im et al., 2016; Choi et al., 2012; Kim and Hogue, 2012; Chauhan et al., 2003). High surface temperature and albedo result in a decrease of soil moisture (Im et al., 2016; Sugathan et al., 2014). ET (122%) was identified as an important variable by the summer model along with LST (131%) and albedo (132%). This is because summer has higher energy flux, higher temperatures, and more rainfall than the other seasons, which can lead to larger variations in soil moisture and evapotranspiration (Cleverly et al., 2013; Rosenbaum et al., 2012; Otkin et al., 2013).

Table 3.1. Soil moisture downscaling model performance and variable importance (%).

Downscaling model	R^2	RMSE (m^3/m^3)	a (slope)	b (intercept)	Important variables (%)*
Spring	0.91	0.032	0.674	0.041	LST (178%), albedo (143%)
Summer	0.95	0.038	0.601	0.068	albedo (132%), LST (131%), ET (122%)
Autumn	0.91	0.032	0.665	0.043	albedo (135%), LST (97%)

*The value in the parenthesis means the increased percentage of mean square error (MSE) when the variable is perturbed by random forest. The larger the value, the more contributing the variable was.

Since the AMSR-E, downscaled, and *in-situ* soil moisture data have different spatial representativeness, it is difficult to directly compare them. As in the study of Im et al. (2016), the satellite-based soil moisture data were compared with *in situ* soil moisture by assuming that all points (i.e., locations) within each pixel have the same value. Table 3.2 and Figure 3.3 show the validation results of downscaled soil moisture with AMSR-E and *in-situ* soil moisture at the two flux towers (Cheongmi-cheon (CMK) and Seolma-cheon (SMK)) from March to November 2010 (one of the driest years). The downscaled soil moisture was well correlated with the *in-situ* soil moisture, showing a mean R^2 of 0.334 and mean RMSE of $0.06 \text{ m}^3/\text{m}^3$. The correlation and RMSE results from SMK are better than the results from CMK. This is because it is more difficult to estimate soil moisture as the soil property of the CMK (i.e., silty clay, sandy clay, clay loam, silty clay loam, silt loam, loam, silt, sandy loam, and sand), area is more heterogeneous than the soil property of SMK (i.e., sandy loam dominant) (HSC, 2011). The range of the downscaled soil moisture is closer to the range of the *in-situ* measurement than the AMSR-E soil moisture, showing a mean slope of 0.891 (Figure 3.3). Although the correlation coefficient between the downscaled soil moisture and *in situ* soil moisture is lower than that between the AMSR-E soil moisture and *in situ* soil moisture, the downscaled soil moisture shows a reasonable correlation coefficient for *in situ* soil moisture and the RMSE was also improved through downscaling. Despite 2010 is one of the driest years, AMSR-E soil moisture ranged from 10-40 %, while the downscaled soil moisture ranged from 10-25%, which shows a similar range with *in situ* soil moisture. The downscaled soil moisture also shows a strong correlation with the AMSR-E soil moisture, with a mean R^2 of 0.612. Since the downscaled soil moisture was well correlated with both the *in-situ* and AMSR-E soil moisture, it can therefore be considered that the downscaled soil moisture is reasonably represented at the spatial scale of 1 km.

Table 3.2. Validation of the downscaled soil moisture with the AMSR-E soil moisture and *in-situ* soil moisture from March to November 2010.

(a) Cheongmi-cheon (CMK)				
	R^2	RMSE (m^3/m^3)	a (slope)	b (intercept)
Downscaled SM (1 km) vs <i>In-situ</i> SM	0.11	0.088	0.611	1.933
AMSR-E SM (25 km) vs <i>In-situ</i> SM	0.24	0.100	1.246	-14.89
Downscaled SM (1 km) vs AMSR-E SM (25 km)	0.44	0.049	0.497	8.191
(b) Seolma-cheon (SMK)				
	R^2	RMSE (m^3/m^3)	a (slope)	b (intercept)
Downscaled SM (1 km) vs <i>In-situ</i> SM	0.56	0.040	1.171	-0.102
AMSR-E SM (25 km) vs <i>In-situ</i> SM	0.56	0.158	3.268	-17.59
Downscaled SM (1 km) vs AMSR-E SM (25 km)	0.79	0.122	0.369	6.202

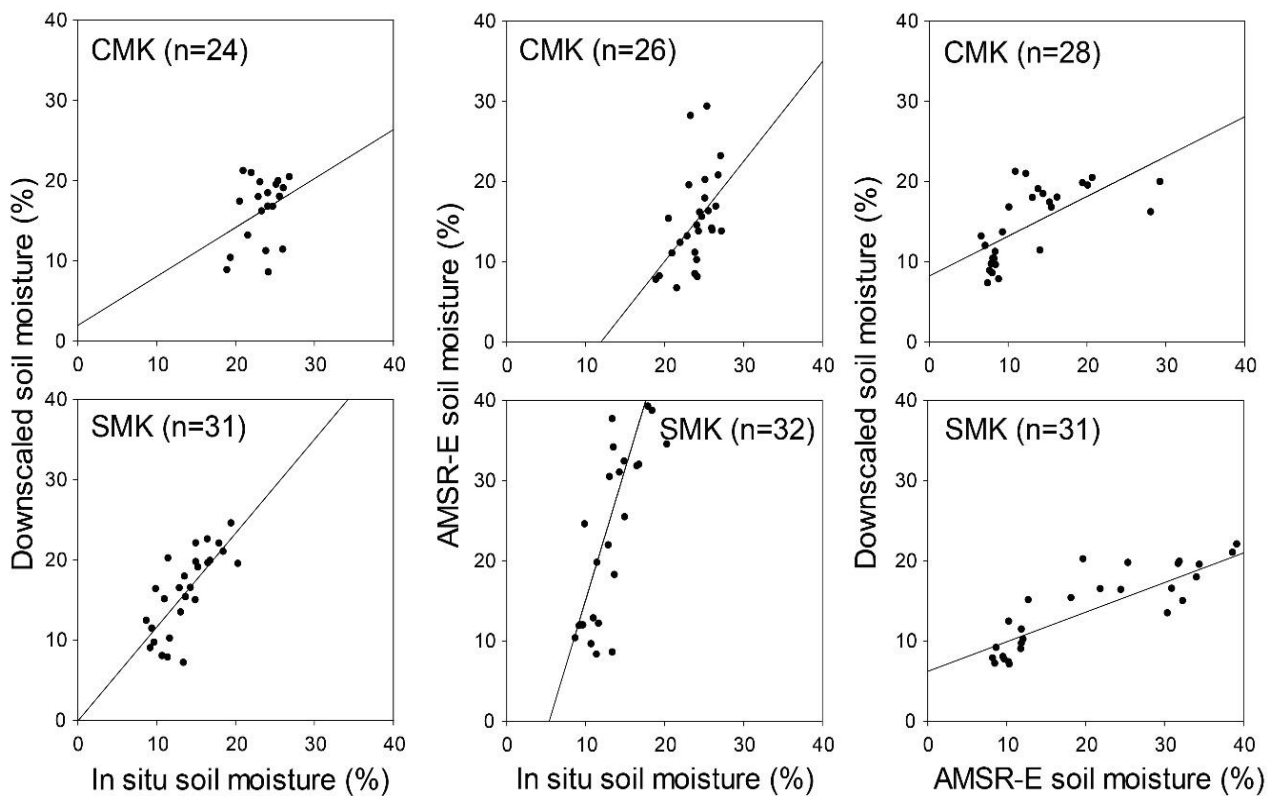


Figure 3.3. Scatterplots between *in situ* and satellite-derived soil moisture ($p < 0.05$) at two flux sites (CMK and SMK).

3.4.2 Monitoring meteorological drought

HSMDI was validated using the 1-, 3-, 6-, 9-, and 12-month SPIs for meteorological drought which were calculated using precipitation data measured at the weather stations located within vegetation areas. Figure 4 shows the correlation coefficient values between HSMDI and the multi-scale SPIs at 29 stations located over vegetated areas (refer to Figure 3.1) from April to October along with monthly precipitation on the right axis (mm/month). Spring generally begins in March and ends in May in Korea, and drought often occurs in spring due to the limited amount of rainfall. Summer is from June to August, which is monsoon season, and most of the rain falls during the summer monsoon in Korea. Autumn begins in September and ends in November, and the amount of rainfall is similar to spring. The temperature increases from March onwards, reaches its highest in July and August, and decreases from August until February of the following year. Most of the main crops are planted in mid-April and harvested during September and October.

HSMDI shows relatively high correlation coefficients with 1-month and 3-month SPIs among the five different timescales of SPIs (Figure 3.4). This is because top soil moisture (< 10 cm) is more sensitive to the cumulative amount of rainfall for a shorter term timescale (e.g., 1 and 3 months) (Zhang et al., 2013). We also examined the correlation coefficients between HSMDIs and 1-/3-month SPIs that were calculated with different time lags (1-2 months) because of known 1- to 3-month time lag effect between meteorological drought and soil moisture drought (AghaKouchak, 2014). However, such time lag effect was not clearly identified in the results. One possible reason is that many of the 29 ground stations are located over mixed land covers (i.e., vegetation + impervious built-up areas) within the 1 km spatial domain, which made HSMDI less sensitive to actual soil moisture. In addition, 24 among the 29 stations are located over cropland with high irrigation rates, implying that a soil moisture-based drought index might not be able to provide consistent meteorological drought information.

The correlation between HSMDI and SPI is higher during the dry season (i.e., April and May) than during the wet season (June to August). On the other hand, HSMDI and SPIs produced low correlation during the wet season (i.e., June, July, and August), which is consistent with the results of Zhang et al. (2013) that showed lower correlation between the soil moisture-based drought index and SPIs over the wetter area. It is regarded that HSMDI is more sensitive to topsoil moisture because of its generation using c-band microwave observations. In addition, since HSMDI was calculated using LST, there was a limited data availability during the wet season due to cloud (no data). Thus, it should be noted that the relationship between HSMDI and SPIs was not well reflected during the wet season. Consequently, HSMDI can be useful to monitor drought effectively during spring and autumn when drought often occurs.

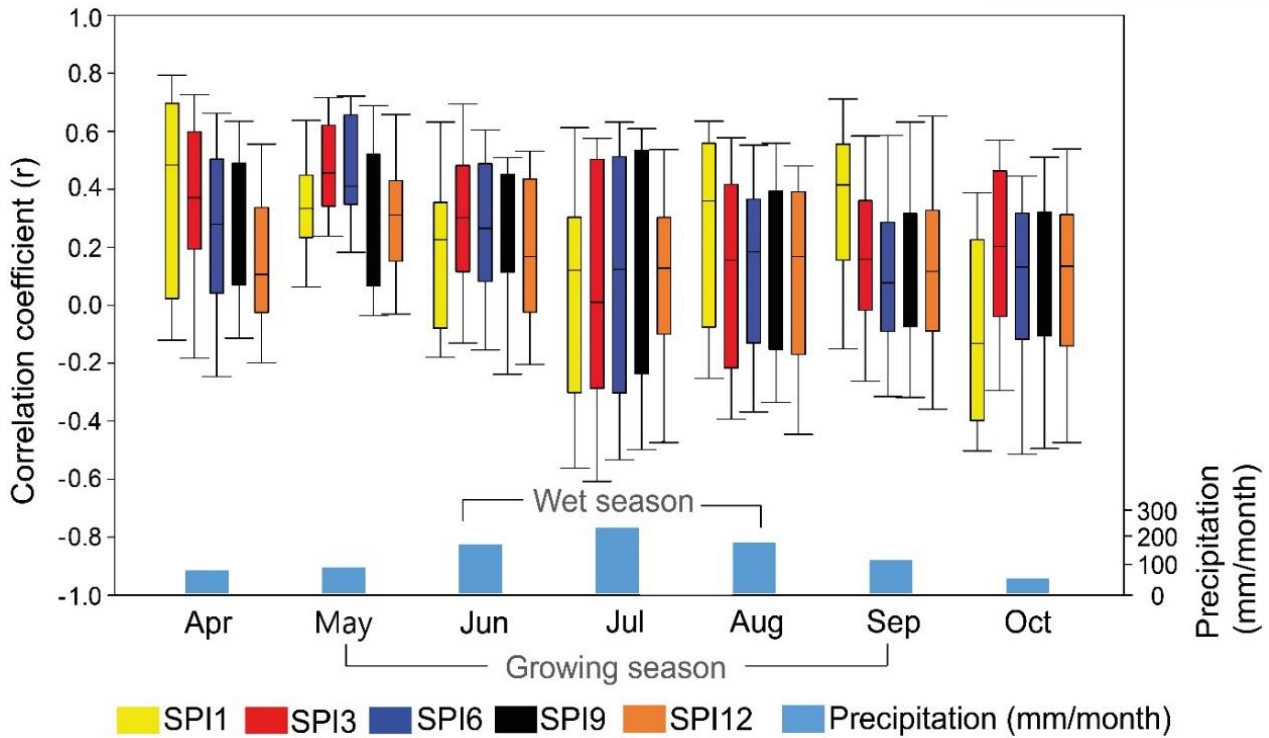


Figure 3.4. Boxplots of correlation coefficient values between HSMDI and SPIs with precipitation by month from April to October.

Figure 3.5 shows the time series of the HSMDIs and the 1-month and 3-month SPIs at six stations which are located within homogeneous land cover domains from April to October between 2003 and 2011. Compared to the results within heterogeneous domains, the temporal trends and patterns of drought (less than 0 for SPI and less than 0.5 for HSMDI) and no drought conditions (greater than 0 for SPI and greater than 0.5 for HSMDI) between HSMDI and SPIs for the homogeneous land cover domains were more similar. The HSMDI at Geumsan and Yangpyeong depicted meteorological drought status well except in 2008. The HSMDI at Geochang and Hongcheon showed very similar temporal variability of SPIs (i.e., decreasing from 2003 to 2005; increasing from 2007 to 2008). The HSMDI at Sancheong also depicted a similar pattern with SPIs except for during 2004 and 2006. At Yeongcheon, meteorological drought conditions were either over- or underestimated. HSMDI generally showed a greater variation than SPIs because SPIs were calculated over a long time period (30 years) through fitting the probability distribution of data while HSMDI was simply standardized without considering the data distribution for relatively short period of time (2003-2011). In addition, there were some periods showing large differences between HSMDI and SPI; these were often lag-time effects for one to three months. This is because the two drought indices have different characteristics: SPI was based solely on precipitation, and HSMDI was based on surface variables such as vegetation conditions, surface temperature, and evapotranspiration as well as precipitation.

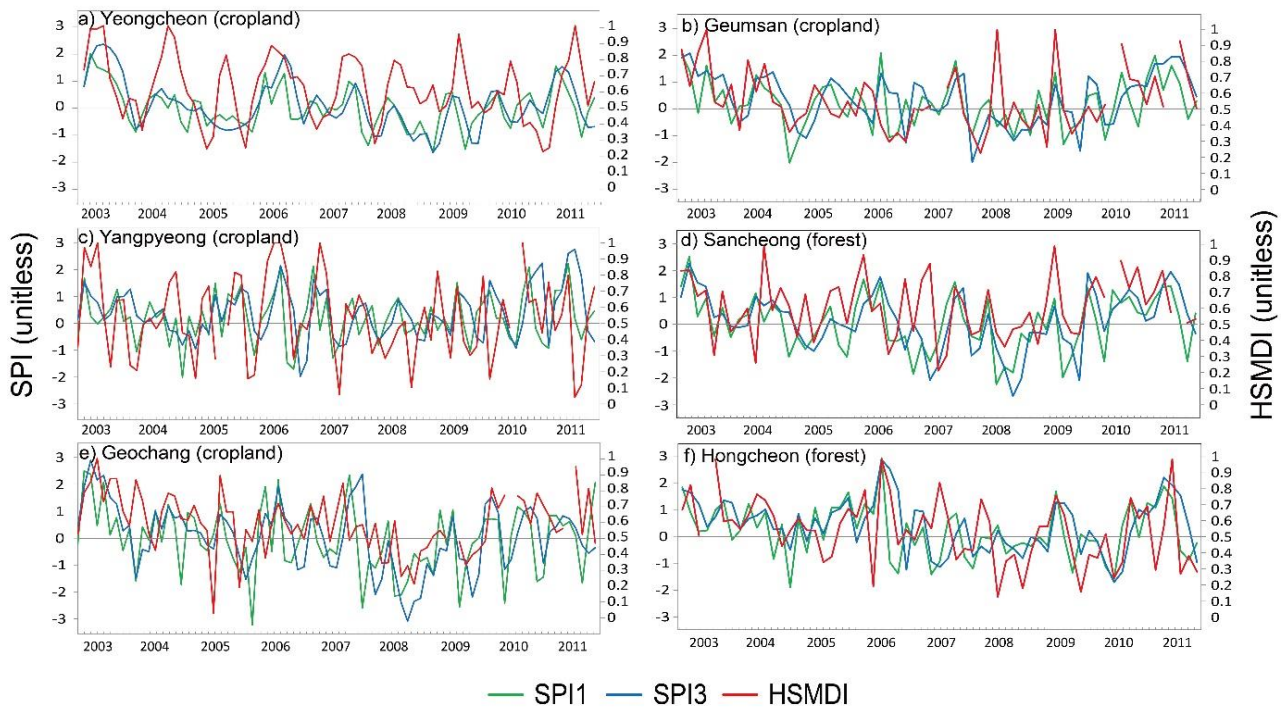


Figure 3.5. Time series of 1- and 3-month SPIs and HSMDI (monthly) from April to October for 2003-2011 by station within homogeneous land cover domains. SPI1 = 1-month SPI; and SPI3 = 3-month SPI.

3.4.3 Monitoring agricultural drought

Crop yield data have been used to evaluate the drought indices for agricultural drought monitoring (Mishra et al., 2015; Park et al., 2016; Quiring and Papakryiakou, 2003; Rhee et al. 2010). Sesame yield was obtained at ten administrative districts in Jeolla Province, and highland radish and napa cabbage yield were obtained at twelve administrative districts in Gangwon Province (refer to Figure 3.1). The correlation coefficients between HSMDI and crop yield from May to September at all stations were 0.36 for sesame, 0.36 for highland radish, and 0.64 for highland napa cabbage. The correlation between HSMDI and crop yield by month was the highest in August among five months, which is the hottest month. Heat stress affects crop yield because crop yield decreases if temperature increases above a critical temperature threshold (Gourdji et al., 2013; Teixeira et al., 2013; Wheeler et al., 2000). The high correlations in August imply that the deficiency of soil moisture due to heat stress also affects crop yield. The thermal-sensitive period (TSP), when crops are only sensitive to heat stress, is the flowering season (August) for sesame and is the reproductive period (June to August) for highland radish and napa cabbage (Wheeler et al., 2000; RDA report). Sesame, highland radish, and highland napa cabbage are vulnerable to high temperature. Since HSMDI incorporates LST information, it is not surprising to observe that HSMDI and crop yield are closely related in August. This result is consistent with previous studies' findings that the flowering season for corn and soybean is vulnerable to drought and important to determine crop yield (Park et al., 2016; Rhee et al., 2010; Johnson et al. 2014; Otkin et al. 2016).

Figure 3.6 shows the scatterplots between yearly crop yield (kg/ha) and HSMDI of one month before

harvesting (August) with R^2 values. In spite of the small number of samples from five-year data, promising results were shown with high correlation coefficient values. Sesame ($R^2 = 0.60, 0.56$, and 0.52 at Shinan, Haenam and Muan, respectively) shows relatively lower correlation coefficient values with HSMDI than the highland radish ($R^2 = 0.79, 0.85$, and 0.81 at Taebaek, Inje, and Hongcheon, respectively) and napa cabbage ($R^2 = 0.66, 0.94$, and 0.58 at Taebaek, Inje, and Hongcheon, respectively). This implies that the effect of drought varies by crop because sesame is cultivated in high irrigation-rate regions, while highland radish and napa cabbage are cultivated in low irrigation-rate regions. Jeolla Province where sesame is cultivated has 9,948 irrigation facilities, while Gangwon Province where highland crops are cultivated has 6,171 irrigation facilities (MAFRA 2010). This suggests that highland crops are useful in evaluating agricultural drought, due to their vulnerability to drought.

Figure 3.7 shows the correlation coefficient values between crop yield (i.e., highland radish and napa cabbage) and SPIs at two administrative districts (Inje and Hongcheon) in August. Two highland crops (radish and napa cabbage) show the similar pattern of the correlation coefficients for SPIs. The relationship between highland crops and SPIs shows a stronger relationship in Hongchen than Inje, despite both administrative districts being mountainous regions. This is because the accessibility for cultivation including irrigation is more limited in Hongcheon than Inje as the mean elevation of Hongcheon (641 m) is higher than Inje (489 m). There is no positive relationship between 1-month SPI and highland crops. This implies that very short-term drought (e.g., 1 month) doesn't affect crop yield seriously except for the month right before harvest. In addition, crop yields and 1-month SPI show a negative relationship, which implies that lots of rainfall can result in a decrease of crop yield because August is a rainy season. Long-term meteorological drought is closely related to crop yield and 3-month SPI is highly correlated with crop yield as well as HSMDI (Figures 3.4 and 3.7). Therefore, it can be considered that HSMDI is responding to similar drought conditions as 3-month SPI or SPI at a longer timescale.

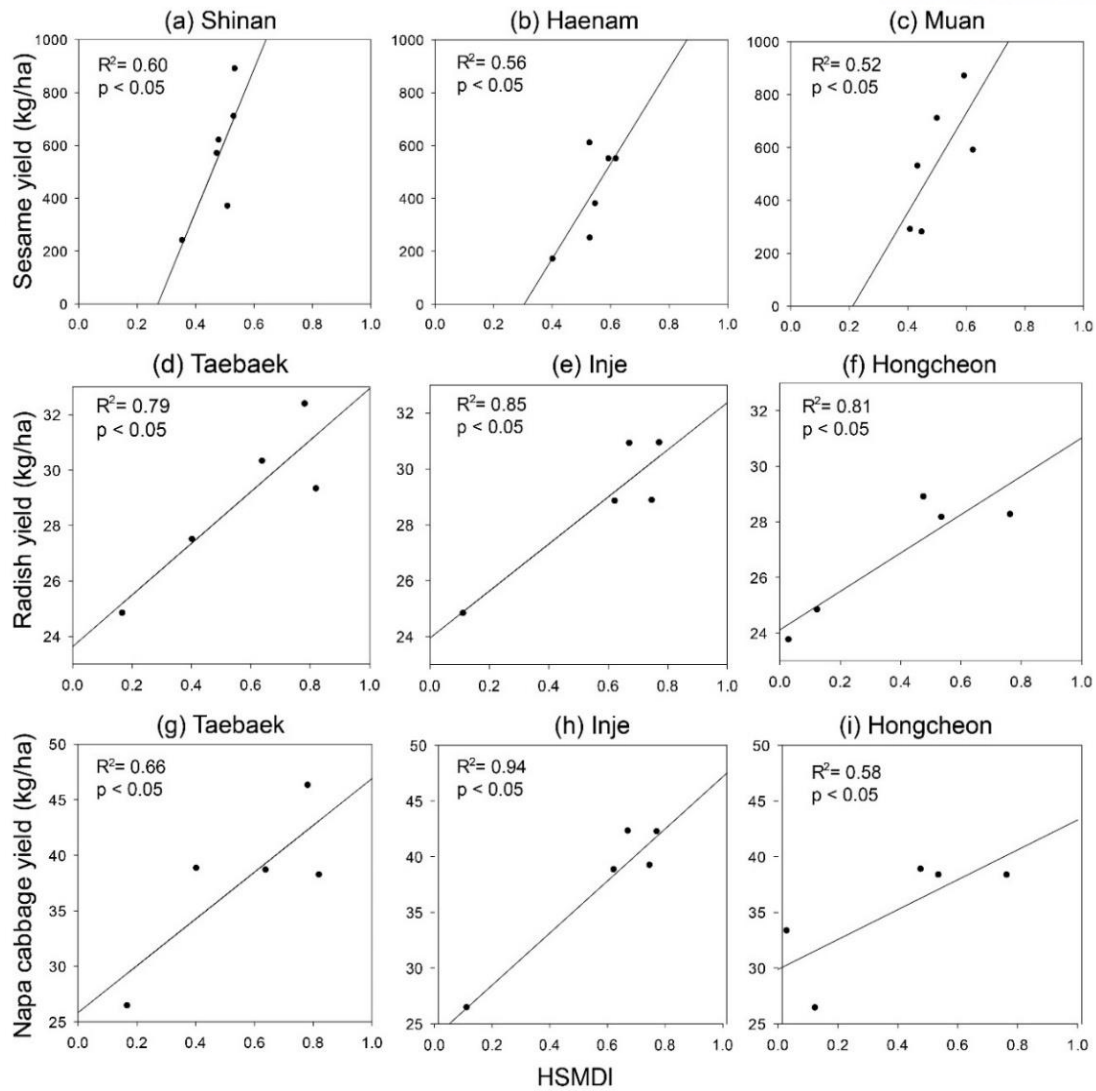


Figure 3.6. Scatterplots between yearly crop yield (sesame, highland radish, and highland napa cabbage) and HSMDI (0.05 significant level) in August.

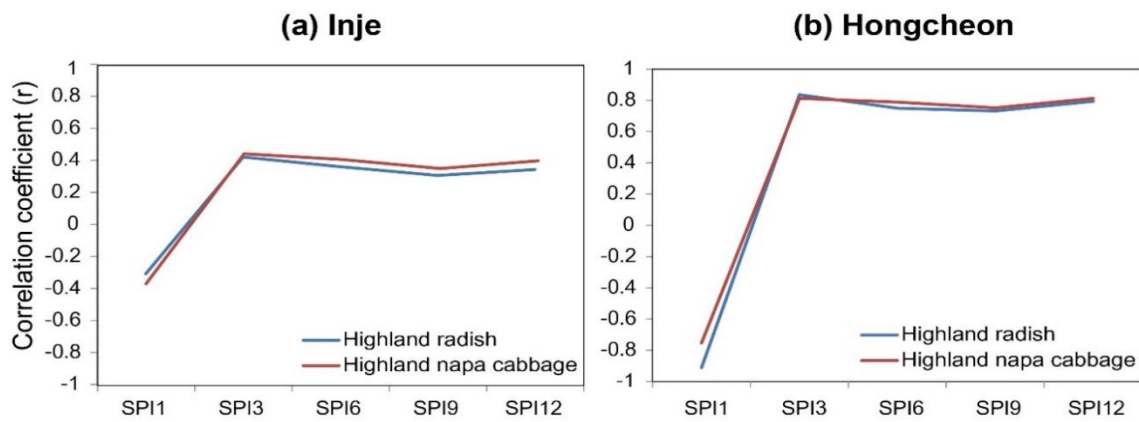


Figure 3.7. Correlation coefficients between SPIs and highland crop yields (radish and napa cabbage) at two administrative districts ($p < 0.05$).

3.4.4 Monitoring hydrological drought

The limited amount of precipitation and soil moisture lead to streamflow deficiency, which develops into hydrological drought. HSMDI was evaluated for hydrological drought monitoring using normalized streamflow anomaly. Monthly streamflow data were obtained from five stations with similar surface transport and geomorphic properties (refer to Figure 3.1). The HSMDIs were averaged within five watersheds and were converted to normalized anomalies using the min-max scaling as streamflow data. The correlation coefficient values between HSMDI and normalized streamflow anomaly were determined by month during the growing season (May to September) from 2003 to 2011.

Figure 3.8 shows the correlations between the scaled streamflow and scaled HSMDI in September. Since they show a range of R^2 (~ 0.09 to 0.69), it is regarded that HSMDI is not able to monitor hydrological drought consistently well. In addition, when HSMDI shows a drought condition, the scaled streamflow shows a normal or wet condition. This is because the drought condition is sensitive to many factors including the given period, their maximum and minimum values (soil moisture and streamflow), and characteristics of variables (e.g., time lag effect of precipitation deficiency, surface variables such as LST, NDVI, EVI, and ET). As the value of 0.5 was set for the normal condition in this study, the HSMDI and scaled streamflow have different dynamic ranges (Figure 8), which implies the limited ability of HSMDI to monitor hydrological drought. The correlation coefficient values between the scaled streamflow and HSMDI during the growing season at all stations are shown in Table 3.3. The correlation coefficients at all stations range from 0.03 to 0.63 in May, from -0.21 to 0.51 in June, from 0.42 to 0.72 in July, from 0.31 to 0.74 in August, and from 0.3 to 0.83 in September.

Table 3.3 Correlation coefficients (r) between HSMDI in five watersheds and scaled streamflow at five stations during the growing season.

Watershed	Month				
	May	June	July	August	September
Gwangju	0.33	0.04	0.46	0.63	0.56
Nampyeong	0.09	0.33	0.42	0.74	0.73
Neungju	0.03	-0.21	0.72	0.31	0.83
Mareuk	0.53	0.51	0.68	0.70	0.30
Jangseong	0.63	0.46	0.70	0.50	0.76

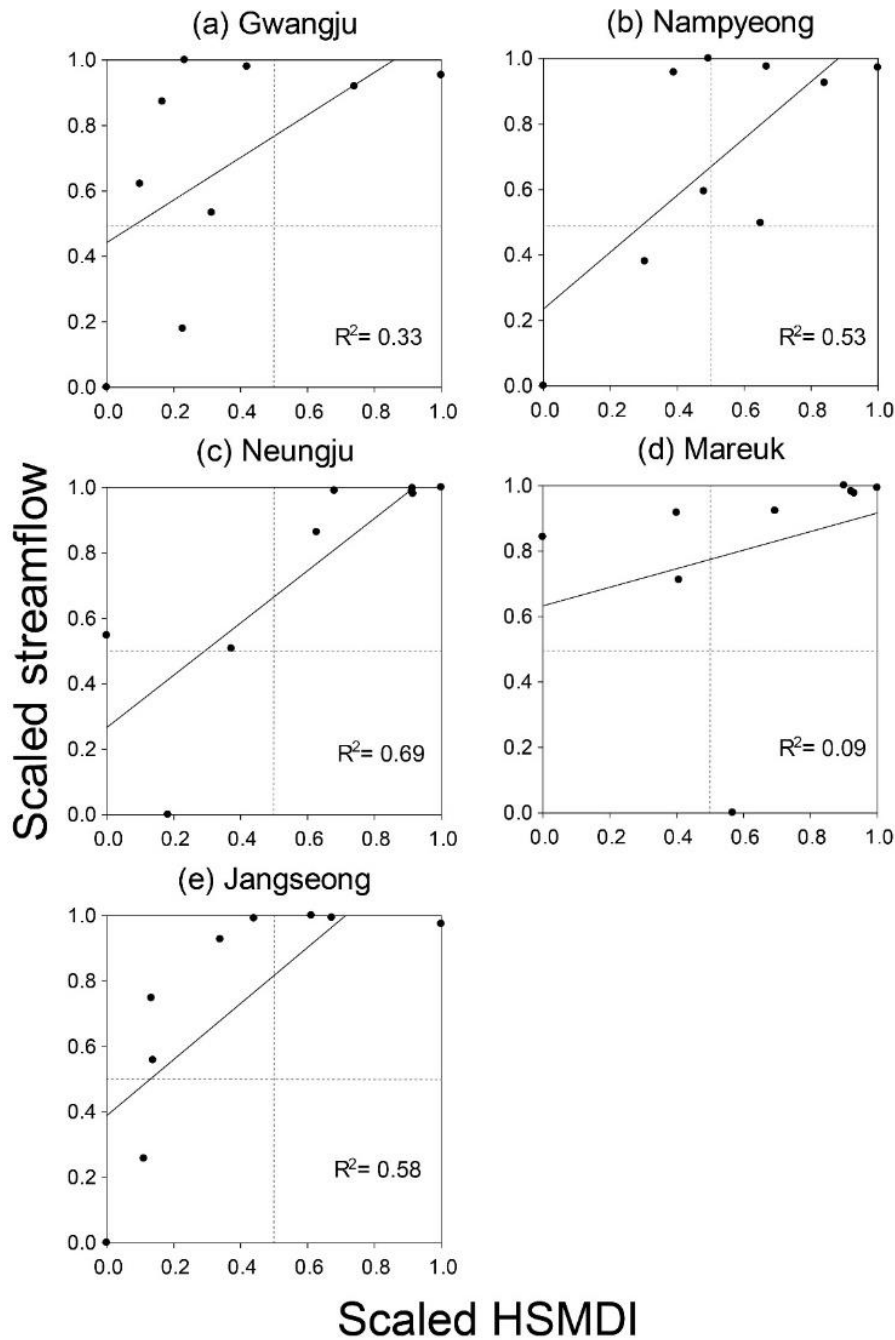


Figure 3.8. Scatterplots between scaled streamflow and scaled HSMDI ($p < 0.05$) in five watersheds (September).

3.4.5 Novelty, opportunities, and limitations

This study developed a drought index, HSMDI, using high resolution soil moisture through downscaling AMSR-E soil moisture based on a machine learning approach. HSMDI is calculated using a min-max scaling at each pixel during a given period and thus ranges from 0 to 1. Based on the potential maximum and minimum at each pixel, it is identified how dry the pixel is, which allows informing relative drought status

for short or long periods by setting a period of interest. The soil moisture-based drought index is very useful to monitor meteorological drought that occurs due to the lack of precipitation and agricultural drought that is triggered by soil moisture deficiency. In addition, since the HSMDI was developed through integrating various factors (i.e., LST, vegetation indices, ET, albedo, and precipitation), it is useful to understand the characteristics of meteorological and agricultural droughts. This study can be applied to any region including ones where *in situ* data has limited availability because all input data are satellite-based. In this study, the HSMDI was evaluated for meteorological, agricultural, and hydrological droughts. The results show that the HSMDI properly monitors meteorological and agricultural droughts. In comparison with multi-timescale SPIs, HSMDI showed a similar ability to identify drought conditions as the 3-month SPI. The HSMDI and scaled streamflow data showed relatively low correlations, and their dynamic ranges of drought conditions are a bit different.

There are some limitations in this study. Korea has rugged topography and heterogeneous land covers, which makes it difficult to directly compare HSMDI with *in situ* data due to the different spatial scale. In addition, the limited optical sensor data availability during the wet season and relatively high irrigation rates over cropland made the performance of HSMDI for monitoring drought conditions to be somewhat unstable. In particular, the correlation coefficients between HSMDI and SPIs were relatively low at the stations where the spatial domain (i.e., 1 km x 1 km) contains built-up areas. In addition, although many surface factors were used to develop the HSMDI, it is difficult to monitor various types of droughts using only a soil moisture-based drought index. Since the original AMSR-E soil moisture product can only represent surface soil moisture (< 10 cm), it is not surprising that HSMDI can monitor meteorological and agricultural droughts reasonably, but has a limited capability to model hydrological drought. If deep soil moisture (e.g., 1 m) can be incorporated into HSMDI, the modeling ability of hydrological drought may be greatly enhanced. There is a limitation for acquiring validation data (i.e., crop yield and streamflow). This study uses streamflow observations from a single small river basin due to topography and data quality problems. In future research, the portion of the analysis (i.e., number of validation sites) will be expanded to verify drought index more accurately. Another limitation is the short period (9 years) of the satellite data used. Considering the operation period of typical satellite sensors, it might be difficult to have data from a single sensor for a long time (e.g., 20 years). Thus, advanced data/product fusion of multiple satellite sensors should continue to be conducted in order to secure long-term data for effective and reliable drought monitoring and forecasts.

3.5 Conclusion

In this study, seven MODIS and TRMM surface factors were used to downscale AMSR-E soil moisture using random forest. Downscaled 1 km soil moisture yielded high correlations with both *in situ* and AMSR-E soil moisture. The downscaled soil moisture produced lower RMSE when compared with *in situ* soil moisture than the AMSR-E soil moisture did. A novel soil moisture-based drought index, HSMDI, was proposed in this study. HSMDI was evaluated for meteorological, agricultural, and hydrological droughts using multi-timescale SPIs (1 to 12-month), crop yield, and streamflow data. HSMDI and SPIs showed high correlation during the dry season and, in particular, HSMDI was highly correlated with the 3-month SPI.

Although HSMDI showed unstable performance to monitor meteorological drought during the wet season, it should be noted that HSMDI well monitored meteorological drought (i.e., 1-/3-month SPI) over the areas with homogeneous land covers. Crop yield data (i.e., sesame, highland radish, and highland napa cabbage) were used to validate HSMDI for agricultural drought. HSMDI is more closely correlated with highland crop yield, which has a relatively lower irrigation rate than sesame.

HSMDI monitored both meteorological and agricultural droughts well and has characteristics similar to the drought conditions with 3-month precipitation deficiency. However, HSMDI did not model the streamflow very well for hydrological drought, which deserves future research. Factors associated with hydrological drought might be incorporated into the downscaling process or the HSMDI generation to improve the performance of drought monitoring. Future research includes 1) enhancing HSMDI to better model hydrological drought through additional data/product fusion, and 2) evaluating HSMDI over areas with different climate and environmental characteristics.

Chapter 4

4. Very short-term Prediction of drought using remote sensing data and MJO index through Random forest over East Asia

4.1 Background

Drought is one of the complex disasters because it is difficult to identify the start and ending points of drought unlike other disasters such as typhoons, landslides, and floods. Drought associates a meteorological process to agricultural and hydrological processes. Meteorological drought occurs due to the deficit of precipitation, which brings about a shortage of available soil water for plant growth (Zhang et al., 2017). Agricultural drought is caused by the shortage of soil water, which results in significant damage to agricultural ecosystems (e.g., crop yield) (Martínez-Fernández et al., 2016). Hydrological drought refers to the lack of surface water, which affects water resources for allocation (Van Loon et al., 2016). Numerous indices such as Standardized Precipitation Index (SPI; McKee et al., 1995), Scaled Drought Condition Index (SDCI; Rhee et al., 2010), Microwave Integrated Drought Index (MIDI; Zhang and Jia 2013) and many drought monitoring systems including U.S. Drought Monitor (USDM; Svoboda et al., 2002), European Drought Observatory (EDO; <http://edo.jrc.ec.europa.eu/>), and Vegetation Drought Response Index (VegDRI; Brown et al., 2008) have been developed to monitor drought. While drought monitoring has been relatively well advanced in many countries, short- to long-term predictions of drought are still difficult due to the inherent complexity of drought. Timely prediction of drought provides valuable information to decision makers, which can mitigate drought.

For those reasons, drought predictions are of an increasing interest, and various studies for drought predictions have been conducted. Several stochastic models have been developed for the prediction of meteorological and agricultural drought based on 1) empirical methodology such as Multiple Linear Regression (Lorenz et al., 2017; Otkin et al., 2014), 2) the autoregressive moving-average (ARMA)/Autoregressive integrated moving average models (ARIMA)(Mishra & Singh 2011; Han et al., 2010), 3) the Ensemble Streamflow Prediction (ESP) method (AghaKouchak 2015; Mo et al., 2012; Lyon et al., 2012), and 4) machine learning (Belayneh et al., 2014; Belayneh & Adamowski 2012; Özger et al., 2012). Those models well predicted the gradually intensified drought. However, it was hard to predict the rapid changes of drought conditions such as wet to dry and dry to wet due to delayed signal because the variables tend to respond slowly to sudden changes of drought conditions (Otkin et al, 2015; Otkin et al., 2013). Therefore, it is necessary to understand the regional meteorological or climate factors as well as historical patterns to improve the prediction models.

There are many climate factors affect the occurrence of drought. The East Asian monsoon, which is closely related to drought, is prominently affected by the variations from the Pacific Ocean. The dynamical

connection between East Asia and the El Niño-Southern Oscillation (ENSO) is identified in Wang et al. 2000. The East Asian monsoon is also affected by Sea Surface Temperature (SST) variations in the Indian Ocean as well as the tropical Pacific Ocean (Yang et al. 2007). In this regard, there have been many attempts to use the climate variability to forecast regional precipitation and East Asian monsoon. For example, Wu et al. (2009) developed a statistical model for predicting the East Asian summer monsoon using ENSO and the North Atlantic Oscillation (NAO). From these potential relationships at an interannual time scale, the climate indices such as ENSO, Arctic Oscillation (AO), and NAO were used for drought forecasting in the previous studies (Deo, Kisi, & Singh 2017; Rhee & Im 2017; Morid, Smakhtin, & Bagherzadeh 2007).

Although many drought systems have been developed in recent decades, short-term prediction of drought (within 10 days) is still challenging. The historical pattern of short-term drought prediction is less valid than long-term drought prediction because the drought factors respond slowly to sudden changes of drought conditions and have delayed signal (Otkin et al, 2015; Otkin et al., 2013). The causative factors of short-term prediction of drought are also more complex than long-term prediction of drought, and it is difficult to understand the process. While drought prediction for the long-period of time is mainly based on the lack of precipitation, drought prediction for the short-period of time is based on other factors including temperature and evapotranspiration rather than precipitation. However, short-term prediction of drought is important. Some droughts such as flash drought caused by a rapid rate of intensification has increased in recent years, and it has resulted in damage to agricultural systems (Wang et al., 2016) such as economic losses reported in the billions of \$US from the 2010/11 flash drought in U.S. (Wallander et al., 2012). Therefore, an early warning of short-term drought to the agricultural community is necessary to mitigate the related losses (Mo et al., 2015).

Short-term prediction of drought can be affected by atmospheric variability at intraseasonal time scale. A representative variability is Madden-Julian oscillation (MJO: Madden and Julian 1971) which leads to changing the teleconnection pattern that affects extratropical circulation over East Asia during the boreal winter season (Jeong et al., 2005). The anomalous teleconnection pattern associated with MJO also leads to the anomalous precipitation which subsequently affects hydrological land surface conditions (Jeong et al., 2008). Additionally, Peng et al. (2017) mainly explored the relationship between the MJO and land surface soil moisture across the world especially over monsoon regions, where the relationship through the change of the atmospheric teleconnection pattern was identified using satellite-based soil moisture and precipitation data. The soil moisture changes over MJO phases could be a useful information to predict short-term drought event since the variable has persistence for a few days.

This study aims at developing a drought prediction model for the short-period of time (one pentad) using random forest machine learning. The proposed drought prediction model considered the real-time multivariate (RMM) MJO indices because the MJO that is a short time scale climate variability and has important implications for the drought in East Asia (Chen et al., 2017; Lu et al., 2012; Jeong et al., 2008). Three satellite based drought indices, the SDCI, MIDI, and Very Short-term Drought index (VSDI) were used because satellite-based drought indices are able to detect sudden changes at relatively high spatial resolution and

increase the awareness of flash droughts (Mo et al., 2015). The objectives of this study are to 1) suggest a modified drought index, Very Short-term Drought Index (VSDI) combining surface soil moisture, Land Surface Temperature (LST), and Normalized Difference Vegetation Index (NDVI), 2) develop drought prediction models by drought index using random forest, 3) validate the drought prediction models, and 4) compare the spatial distributions of drought evolution from original and predicted drought indices.

4.2 Study area and data

4.2.1 Study area

This study mainly focused on East Asia region (10° - 50° N and 90° - 150° E) including east China, southeast Russia, Taiwan, Korea, and Japan. East Asia suffers from droughts especially during the spring (from March to May) due to the transient season of East Asian monsoon onset. East Asia has diverse land cover types. Figure 3.1 shows MODIS (MCD12Q1) land cover distribution with eleven land cover classes. While Korea, Taiwan, and Japan mostly consists of forest and cropland, China composed of forest, croplands, shrublands, grasslands, and barren. Most of the region over Southern East China, South Korea, and Japan is covered with vegetation including forest and croplands where annual rainfall is concentrated, while North Central China including desert is the barren. In this regard, the land surface conditions on the East Asia is strongly associated with atmospheric phenomenon in diverse time scale.

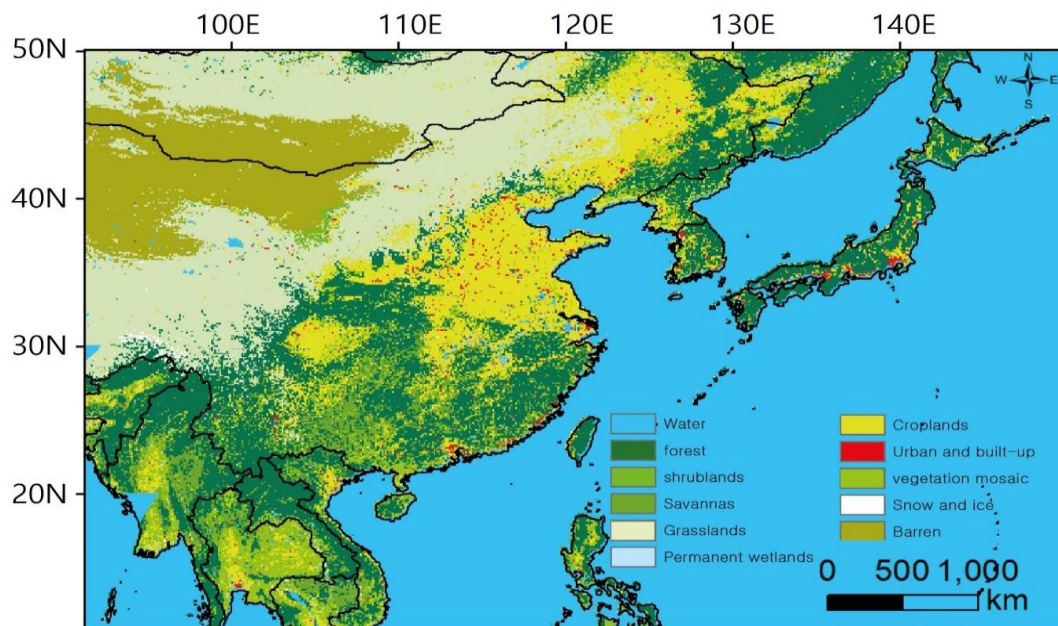


Figure 4.1 Study area with MODIS Land cover 2010 and sixteen land cover classes were aggregated to eleven representative classes.

4.2.2 Data

4.2.2.1 Satellite data

In this study, three satellite-based drought indices were used because satellite images have showed excellent performance in drought monitoring (Lorenz et al 2017; Park et al 2017; Zhang et al 2017; Park et al

2016; Martínez-Fernández et al 2016; Rhee et al 2010; Brown et al. 2008). It is expected that it is useful in short-term prediction of drought. Four variables, LST, NDVI, precipitation, and soil moisture, were used to produce three drought indices. LST and NDVI were calculated using MODIS daily LST (MOD11C1) and daily surface reflectance (MOD09CMG) with 0.05° spatial resolution which were obtained from EARTH DATA (<https://earthdata.nasa.gov/>). In this study, pentad mean (5 days) data were used to capture the short-term change of drought conditions, and thus, daily products were converted into pentad information. To calculate NDVI, bands 1 (red) and 2 (near-infrared) were used, and the Maximum Value Composite approach (MVC; Holben et al., 1986) was applied to produce pentad NDVI. This study used TRMM 3B42 daily precipitation data with 0.25° spatial resolution obtained from Goddard Earth Sciences Data and Information Center (GES DIC; <https://mirador.gsfc.nasa.gov/>). Daily precipitation were converted to monthly with a 5-day interval (pentad). European Space Agency (ESA) – Climate Change Initiative (CCI) satellite-based soil moisture dataset collaborating active and passive measurements provides daily volumetric soil moisture at surface level with 0.25° spatial resolution since 1979 (Liu et al., 2012; Dorigo et al., 2017; Gruber et al., 2017). The data is available from ESA-CCI (<http://www.esa-soilmoisture-cci.org>), and version 3.3 soil moisture data were used in this study. TRMM precipitation and ESA CCI soil moisture were resampled to 0.05° grid size using the bilinear interpolation.

SDCI proposed by Rhee et al. (2010) was calculated using Precipitation Condition Index (PCI), Temperature Condition Index (TCI), and Vegetation Condition Index (VCI) (equation 4.1). MIDI proposed by Zhang and Jia (2013) were calculated using PCI, Soil Moisture Condition Index (SMCI), and TCI (equation 4.2). Those condition indices were calculated by normalizing from 0 to 1 through max-min scaling at each pixel, which considers the potential maximum and minimum of ecosystem as discussed in Kogan et al. (1995). The values of 0 (1) represents the driest (wettest) condition. Although the growing season over East Asia is from April to September, drought frequently occurs from April to May when people start farming. Thus, the temporal scope of this study is from April to May between 2000 and 2016. The indices were computed using pentad means for 17 years (twelve pentads per year). Each drought index has twelve pentads per year. This study proposed a new drought index, Very Short-term Drought Index (VSDI), which modified SDCI because precipitation in SDCI is inadequate for monitoring and predicting change in drought conditions for the short-period of time. It is difficult for the lack of precipitation within one pentad to be the standard of drought, and other factors such as temperature and evapotranspiration are highly related with drought rather than precipitation for the short-period of time. Since soil moisture reflects precipitation as discussed in Wang et al. 2016, VSDI was developed by replacing PCI with SMCI (equation 4.3).

$$\text{SDCI} = 0.5 \cdot \text{PCI} + 0.25 \cdot \text{TCI} + 0.25 \cdot \text{VCI} \quad (\text{equation 4.1})$$

$$\text{MIDI} = 0.5 \cdot \text{PCI} + 0.3 \cdot \text{SMCI} + 0.2 \cdot \text{TCI} \quad (\text{equation 4.2})$$

$$\text{VSDI} = 0.5 \cdot \text{SMCI} + 0.25 \cdot \text{TCI} + 0.25 \cdot \text{VCI} \quad (\text{equation 4.3})$$

4.2.2.2 Numerical Land Surface data

A land surface model offline simulation with Joint UK Land Environment Simulator (JULES: Best et

al., 2011) was implemented. The model simulation produced land surface variables with a 0.5° spatial and 4 layers (0.1, 0.25, 0.65, and 2 meters) vertical resolution in the land. Near surface atmospheric forcing provided by the Terrestrial Hydrology Research Group (Sheffield et al. 2006) was used to drive the land surface model. The atmospheric variables are based on the National Centers for Environmental Prediction–National Center for Atmospheric Research reanalysis (Kalnay et al. 1996) and their mean climatology was corrected by observational datasets (e.g., Global Precipitation Climatology Project (GPCP) and TRMM 3B42RT). Climatic Research Unit (CRU) time-series (TS) 2.0 monthly data (Mitchell et al. 2004) was used to correct air temperature, and NASA Langley surface radiation budget products (Stackhouse et al. 2004; monthly) was adopted to correct surface radiative fluxes.

4.2.2.3 Madden-Julian Oscillation (MJO) index

To define MJO phases, we adopted the method proposed in Wheeler and Hendon (2004), which is based on a pair of empirical orthogonal functions (EOFs) of the combined fields of 850-hPa, 200-hPa zonal wind anomalies and outgoing longwave radiation (OLR) averaged over the tropics (15°S – 15°N). The two leading principal components (PC1 and PC2) of the EOFs are referred to the Real-time Multivariate MJO series 1 (RMM1) and 2 (RMM2), respectively. Based on RMMs, MJO phases one to eight are determined by the location of convection center and its intensity is calculated by the root of the sum of the squared RMM1 and RMM2.

4.3. Methodology

A total of eleven input variables, 1-, 2- and 3- pentads before the target prediction dates RMM1 MJO, RMM2 MJO, and drought indices, latitude, and longitude, were used in the drought prediction models. The target variable is each drought index. For example, to predict SDCI on a target date, SDCIs computed 1-, 2- and 3-pentad before the date were used with other variables including 1-, 2- and 3-pentad before RMM MJO indices, latitude, and longitude of each pixel over study area. Figure 4.2 shows the process flow diagram of the approach proposed in this study. All variables drought—indices, RMM MJO indices, latitude, and longitude—were resampled to 100 km to consider the characteristics of variables over the whole area of study region by using all pixels. Then, a total of eleven independent variables and the dependent variable (i.e., each drought index for target dates), were fed into random forest to develop drought prediction model. Randomly selected 80% of the samples were used as training data, and the remaining 20% were used for validation.

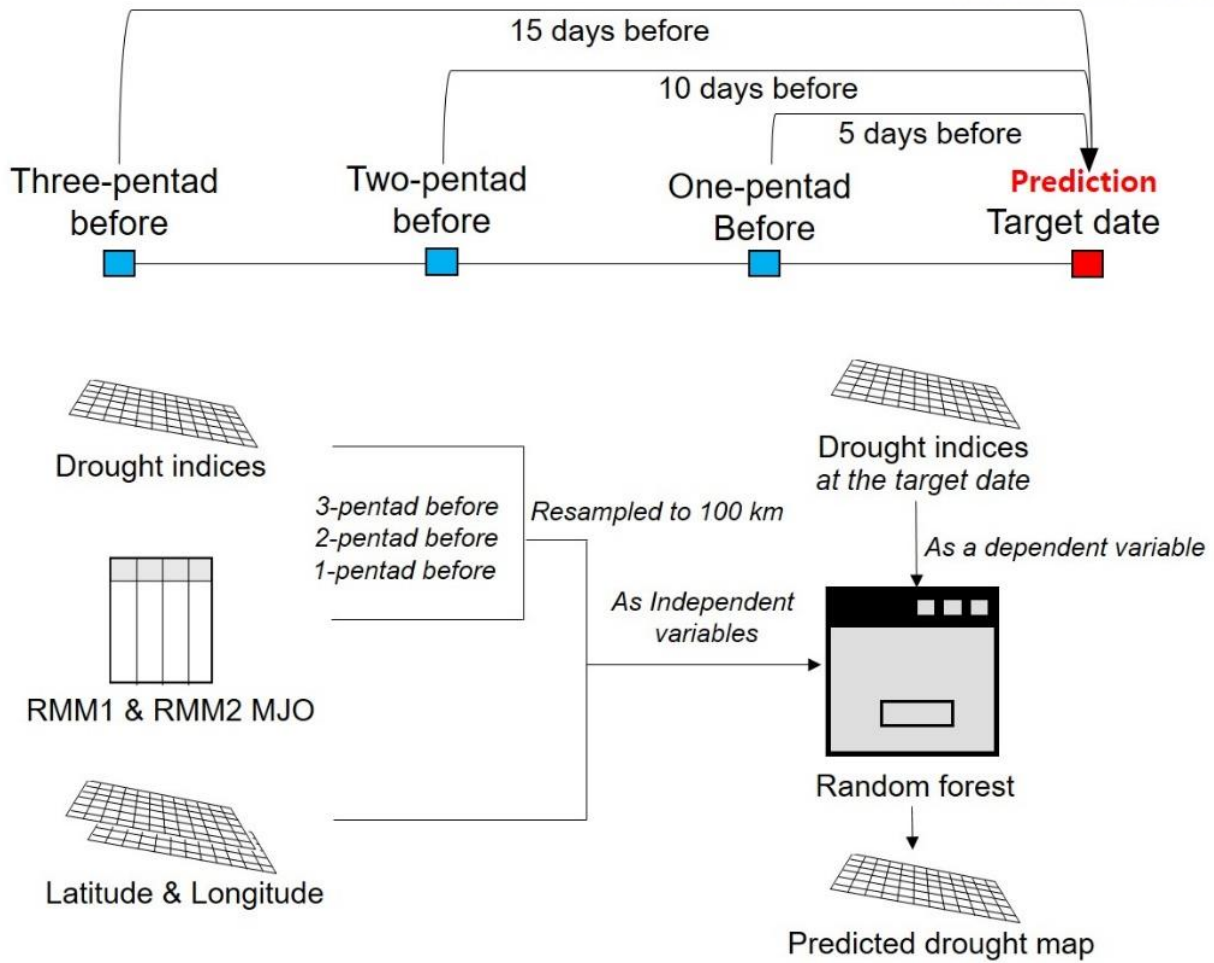


Figure 4.2 Flowchart of this study. 1-, 2- and 3-pentad before the date RMM MJO indices and drought indices, latitude and longitude were used to predict drought.

Random forest (RF) machine learning was adopted to develop drought prediction models because RF has been used in many remote sensing field and showed relatively better performance than other machine learning approaches such as decision tree, boosted regression trees, Support Vector Regression (Yinghai et al. 2017; Im et al. 2016; Park et al. 2016). RF consists of many (typically 500-1000) classification and regression trees (CARTs). There are two major randomization processes in RF. A subset of training samples (2/3) was randomly selected to develop each tree, and a random subset of variables was used at each node of a tree to determine a splitting variable. This study used squared root of the number of variables. Each tree is fully grown through repeated splits. The best split at a node is determined and the split is determined in the lowest residual sum of squares (RSS). Finally, the prediction of an unknown pixel is determined by averaging of the results of all trees. RF also provides relative variable importance (%IncMSE) that implies the contribution of the variable for predicting a target variable. It is calculated by permuting variable. The prediction error (Mean Squared Error; MSE) is computed for each tree, and the same error is also computed after permuting each predictor variable. The difference between the two MSEs from all trees are averaged and then standardized. High

difference means the variable is important. R software (<https://www.r-project.org/>) was used to implement RF through Random forest package. Default settings were applied except for the number of trees (1,000 trees).

In this study, the intraseasonal variability of MJO over East Asia was analyzed. Two prediction models for each drought index were compared to understand the effect of RMM MJO indices in drought prediction. The original prediction model used five variables (drought indices three pentads before, latitude and longitude) excluding RMM MJO indices to predict each drought index, while the other MJO-based prediction model used eleven variables including RMM MJO indices. The developed drought prediction models were validated through leave a year out cross validation. This study used 17-fold cross validation (from 2000 to 2016). The performance of the drought prediction models was evaluated using Root Mean Square Error (RMSE) and correlation coefficients (r), relative RMSE (rRMSE), and p-value. The change of spatial distribution of drought from normal to dry and dry to normal conditions was also compared in this study.

4.4 Results and discussion

4.4.1 Meteorological climatology associated with MJO

Figure 4.3 shows the composited air temperature by MJO phases 1 to 8 during the spring season (April–May) for 22 years. To extract the variability of air temperature at intraseasonal time scale, the anomaly of air temperature generated by JULES offline simulation, which was filtered using the 30-90 days bandpass, were used. The spatial anomaly of air temperature according to the MJO phase clearly appears over East Asia. During MJO phases 1–3, there is a distinct warm anomaly pattern, while the opposite anomaly (i.e., cold pattern) is shown for phases 5–7 over East Asia and central China. However, the MJO-related temperature variation is not clear in southern China. The result is not consistent with the analysis in Jeong et al. (2005) during boreal winter season, when MJO intensity is strong, and thus there is a significant cooling in phases 2–4 while a warming in phases 6–8 over the East Asia region.

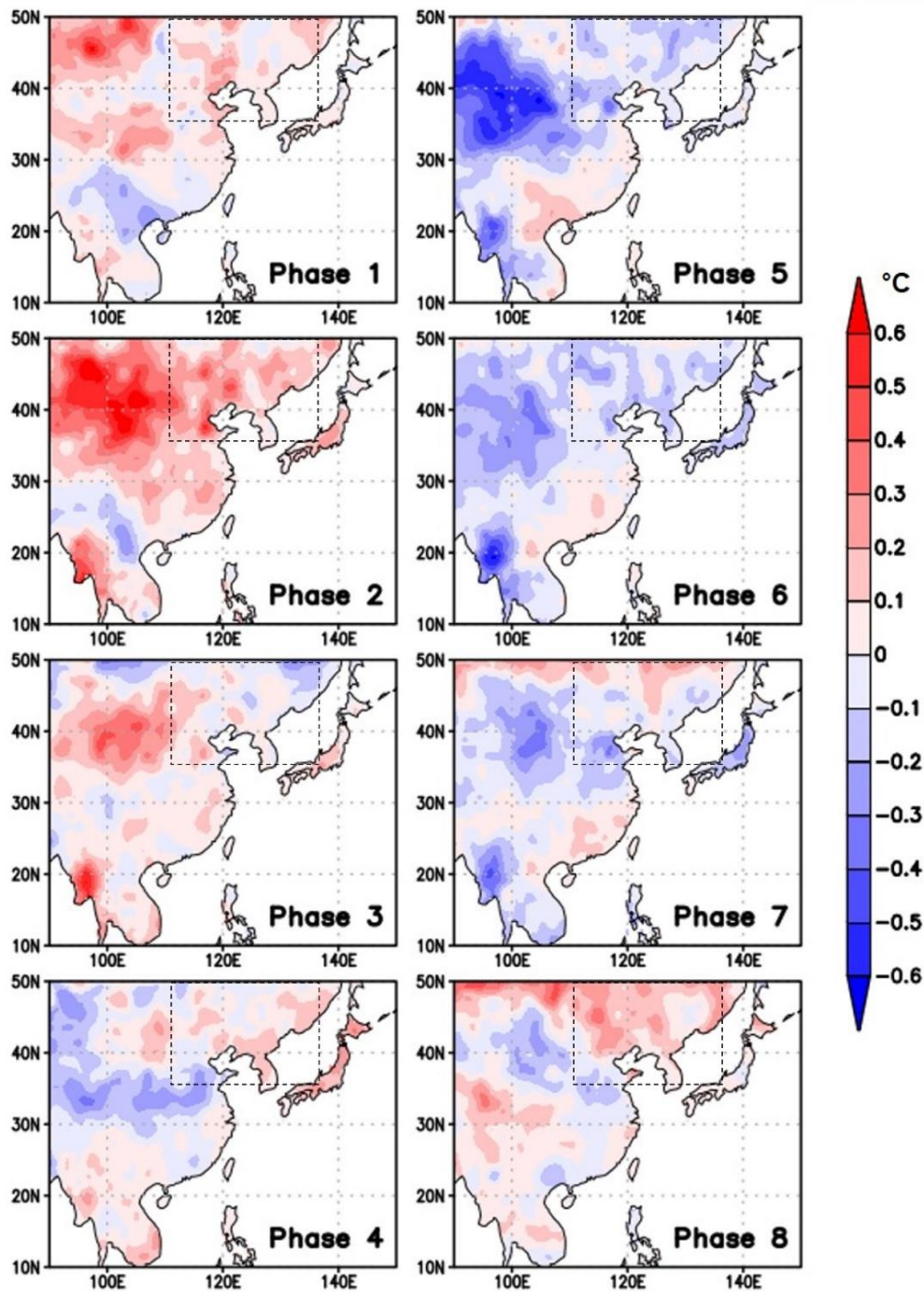


Figure 4.3 Composites of 30-90 days filtered surface air temperature based on each of eight MJO phases during 22 spring seasons (April-May) of 1991–2012

Figure 4.4 represents the daily variance of the air temperature, top-level (~10cm) volumetric soil moisture, and precipitation at intra-seasonal time scale. The variation of the air temperature is active in central China where the land surface condition is relatively dry. The variability of soil moisture is dominant over Southeast Asia and Northeast China. Precipitation variability at subseasonal time scale of 30–90 days is high in the main rainfall regions including Southeast Asia, South China, Korea, and Japan. However, the distribution

of the ratio explained by the intra-seasonal variation versus total variance is quite different. The degree of explanation of soil moisture variability on the range of 30–90 days compared to its total variance is practically high about 40% over central and northeast China where soil moisture memory is relatively large in East Asia (Seneviratne et al., 2006). The ratio of the filtered temperature variance to its total variance is about 10–15% over the research region, and the ratio of the precipitation is less than 10%. Based on the introduced variables, this study examines their role to forecast short-range drought over East Asia.

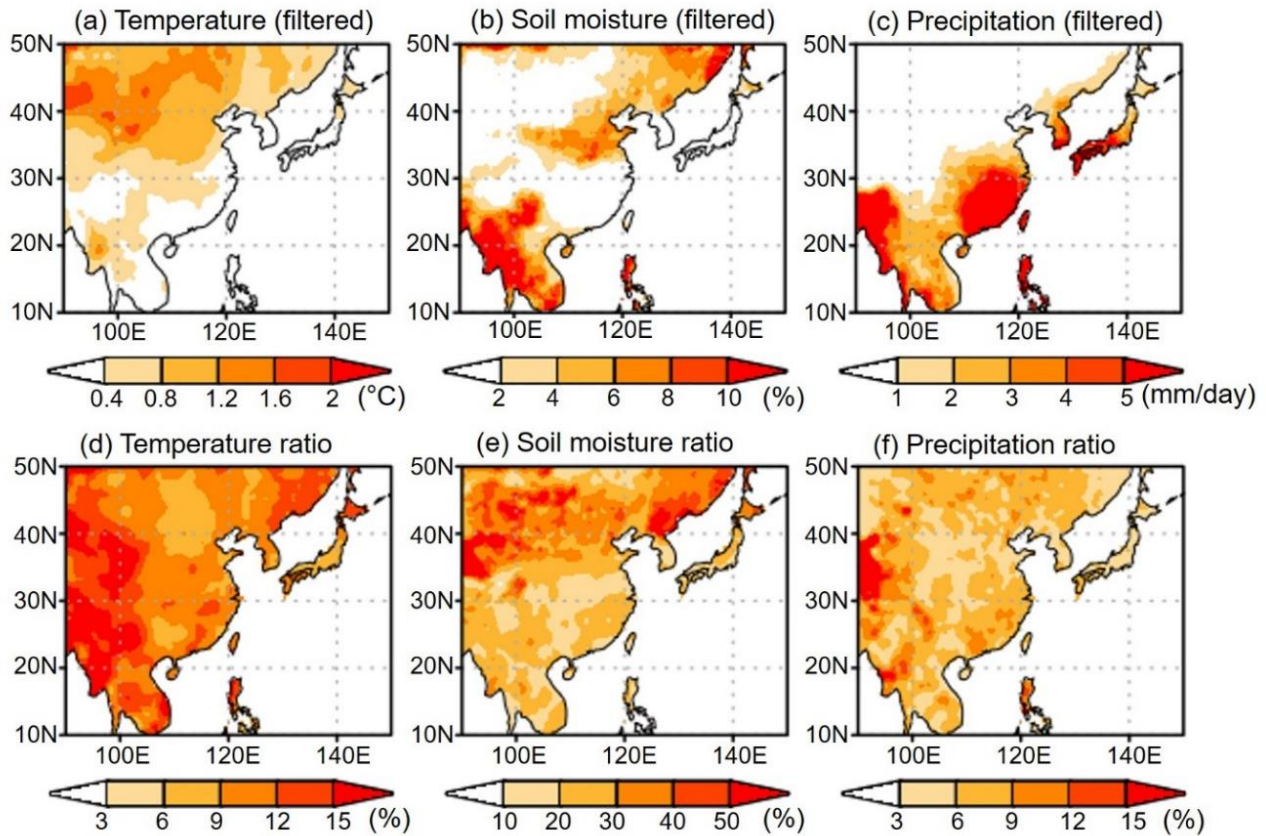


Figure 4.4 Variance of (a) surface air temperature, (b) top-level volumetric soil moisture, and (c) precipitation bandpass filtered for 30-90 days during 22 springs between 1991–2012, and their ratios to total variance in (d-f).

4.4.2 Comparison of drought prediction model performance

The MJO-based prediction models showed better performance than the original prediction models (Figure 4.5). The correlation coefficients and RMSEs were greatly improved by around 0.3 (r) and 10 % (RMSE) in all three drought indices. Figure 4.5 also shows that the predicted drought indices in MJO-based prediction models are matched well with the actual drought indices. For SDCI and MIDI, the values of predicted drought indices are concentrated on mean values (~ 0.35) in the original prediction models, which resulted in a smaller dynamic range (0.2-0.5). Such a tendency has been reported one of limitations of empirical models in many studies. However, predicted drought indices well catch the range of original drought indices (0.1-0.7) in the MJO-based prediction models. In particular, the driest and wettest conditions in SDCI and

MIDI were well predicted, which implies that RMM MJO indices provide useful information for prediction.

The improvement of performance through the application of RMM MJO indices to the drought prediction is significantly dominant for SDCI, which does not contain a soil moisture component (Figure 4.5). The ratio of soil moisture variability at 30–90 days time scale is dominantly higher than those of the temperature and precipitation (refer to Figure 4.4). It means that predictions of MIDI and VSDI without using MJO indices as input already have a MJO-related intraseasonal variation component due to soil moisture. The modeling performance of SDCI in the MJO-based prediction models increased (from 0.29 to 0.70 in r and from 42% to 32% in RMSE). Inclusion of MJO indices shows the substantially improved prediction of SDCI, which consists of PCI, TCI, and VCI, because the MJO-related intraseasonal variability is included less in PCI and TCI than SMCI (Figure 4.4). In VSDI where the soil moisture accounts for half of this index, it shows 0.57 prediction skill without using MJO information to the drought prediction because the surface condition itself already possesses the MJO-related intraseasonal memory.

The performance of the MJO-based prediction model was also better than that of the original prediction model for VSDI, although the original prediction model works well unlike the other drought indices. Since precipitation is more fluctuating during the short-term period than other variables such as soil moisture, LST, and NDVI, it is considered that precipitation is not adequate for monitoring short-term prediction of drought. In addition, most of drought conditions in SDCI and MIDI were dry (less than 0.4) because most of PCI values are less than normal (0.5) due to a large gap between maximum and minimum values of precipitation during one pentad. Therefore, it is regarded that VSDI is adequate for monitoring and predicting drought for the short-period of time. In addition, the performance of three prediction models were saturated around $r \approx 0.7$, which suggests that there are still possibility to improve the drought prediction through further understanding the physical drought processes over East Asia using the climate variabilities at other time scale.

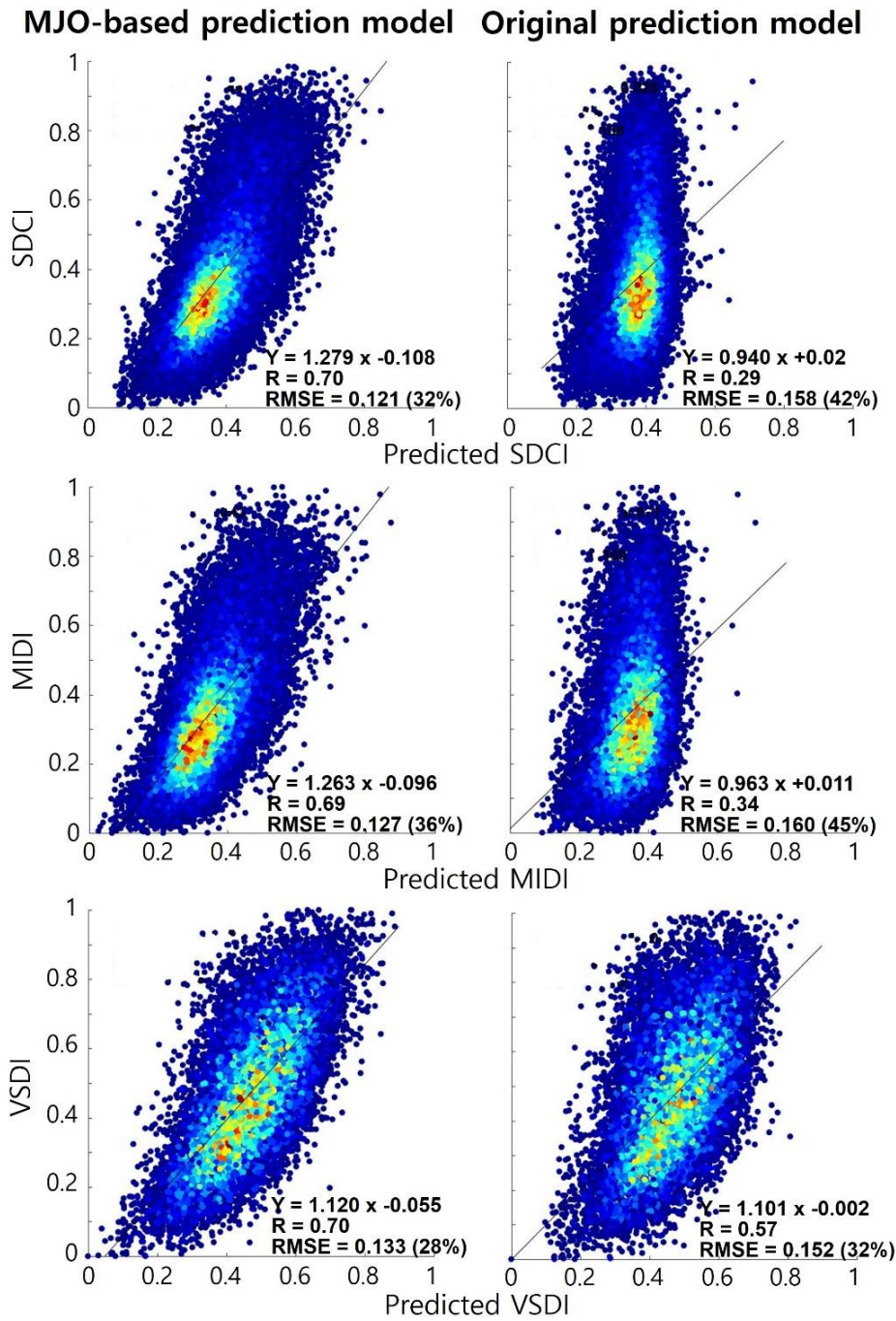


Figure 4.5 Modeling performances of two drought prediction models (with MJO and without MJO) for three drought indices. Scatter plots, r , and RMSE were produced through validation.

The relative variable importance from two models for three drought indices are appeared in Table 1. As expected the latest drought index (1-P before drought index) showed the highest variable importance in both MJO-base model and original model for three drought indices, and there is no noticeable difference in variable importance between 2-pentad before drought index and 3-pentad before drought index. When compared the variable importance of RMM1 MJO and RMM2 MJO, the contribution of the RMM1 MJO was

higher than the RMM2 MJO. As the each of RMM's value is determined by each of EOF leading modes, the RMM1 is more important in the drought forecast than the RMM2 due to the highest variance explanation of EOF1 (Wheeler and Hendon 2004). The variable importance of latitude was higher in original model, while the variable importance of longitude was higher in MJO-based model. MJO variability that primarily tends to eastward propagation in the equatorial belt influences on the horizontal anomaly of teleconnection pattern. The geographical MJO propagation would cause that longitude was more important than latitude in MJO-based model.

Table 4.2 Variable importance of two drought prediction models (with MJO and without MJO) for three drought indices.

Variables	SDCI		MIDI		VSDI	
	MJO-based model	Original model	MJO-based model	Original model	MJO-based model	Original model
1-P before drought index	189.21	126.74	208.45	151.64	267.95	227.17
2-P before drought index	144.38	66.42	146.00	81.43	128.91	107.89
3-P before drought index	146.22	64.62	141.75	73.25	127.34	111.08
Latitude	205.26	117.25	150.36	85.28	92.67	79.32
Longitude	215.02	76.10	177.90	67.37	106.38	74.55
1-P before RMM1	144.23		133.19		130.05	
1-P before RMM2	139.66		117.60		88.32	
2-P before RMM1	110.21		110.71		127.19	
2-P before RMM2	106.79		107.22		81.07	
3-P before RMM1	113.44		114.50		103.56	
3-P before RMM2	97.37		97.84		82.79	

Figure 4.6 shows the spatial distribution of 17-fold cross validation results of the MJO-based VSDI prediction model. Original VSDI and predicted VSDI during planting season (April to May) from 2000 to 2016 (153 pentads) were compared using r , p -value, RMSE, and rRMSE (Figure 4.6). Positive correlation (0.424 averaged) and low RMSE (<0.148) and rRMSE ($< 30\%$) are seen in most areas. Some regions such as southern and northwestern China showed lower r and higher p -value than other regions. It implies that the drought prediction model cannot work well to predict VSDI in the areas. It is because the limited number of samples due to clouds (LST and NDVI) and limited satellite paths (CCI soil moisture). The northeastern China near Gobi desert shows higher RMSE (~ 0.18) and rRMSE ($\sim 40\%$) with relatively low r (~ 0.23) compared to other areas. Because this region is sparsely vegetated and has a relatively dry condition (0.411 averaged in VSDI) regardless of the season unlike the other areas (0.523 averaged in VSDI) (Park & Park 2017). The drought prediction model performed well over densely vegetated regions, which consist of forest, shrublands, savannas and croplands (refer to Figure 4.1) with relatively high r and low RMSE and rRMSE (Figure 4.6).

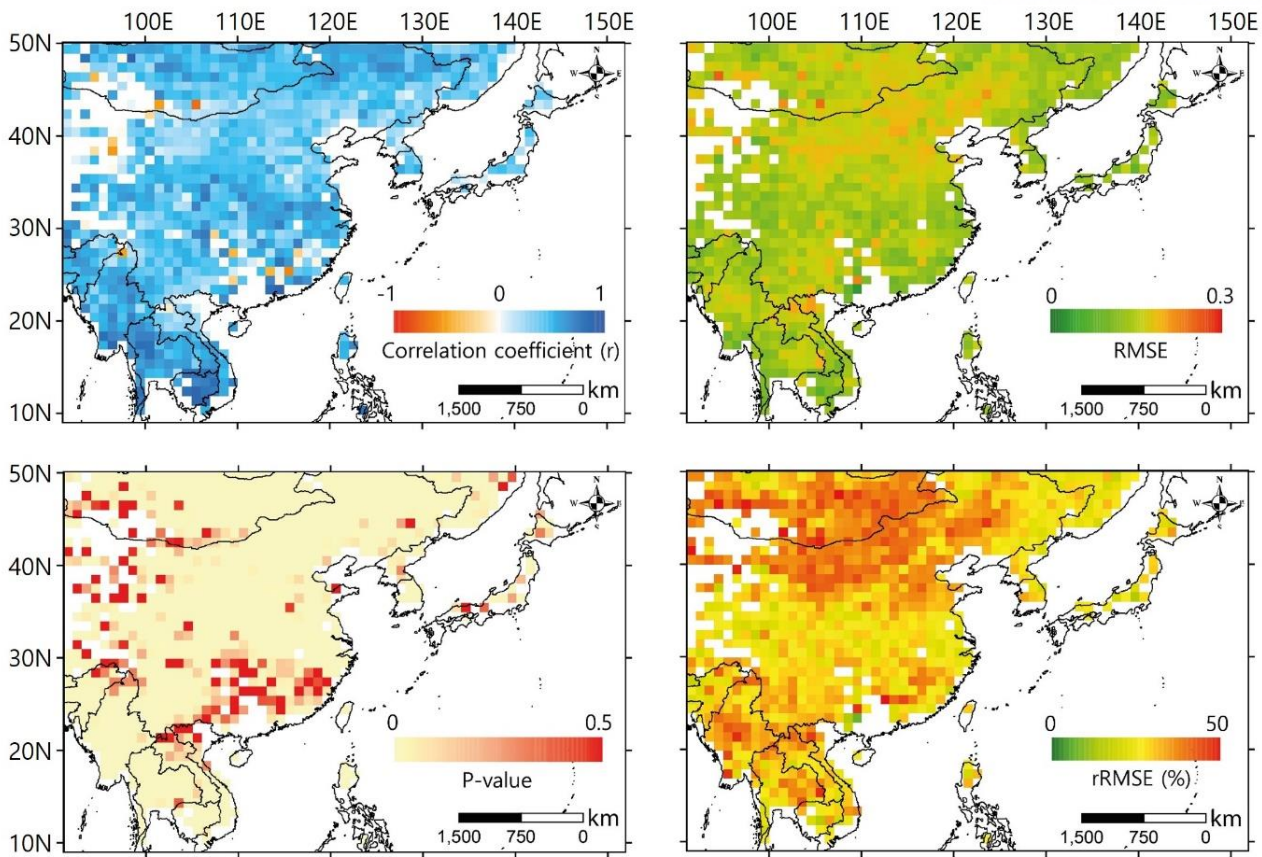


Figure 4.6 leave-one-out validation of VSDI prediction model. Each pixel includes 9 pentads x 17 years data. Each pixel was processed as no data when the number of data was less than 50.

The spatial distribution of drought conditions from original VSDI and predicted VSDI from two models (with MJO and without MJO) were compared with pentad scale in 2010 and 2011 (Figure 4.7 & 4.8). In 2010, drought was eased from April 21 to 26, and drought was intensified from April 26 to May 1. Then, drought was eased again from May 1 to May 6. The MJO-based prediction model well predicted drought conditions in both eased and intensified drought conditions during four pentads (Figure 4.7). It is noticeable that the predicted VSDIs from MJO model can catch the inverse patterns that was not learned from previous drought conditions because the predicted VSDIs from original model (without MJO) cannot catch the pattern especially May 1. April 26 was at the phase 1 and May 1 was at the phase 2, and this region can be dry from phase 1 to phase 2 (refer to Figure 4.3). The MJO-based prediction model can predict the drought near 40°N and 120°E, but the original prediction model cannot predict the drought (Figure 4.7). In Figure 4.8, Drought was evolved from May 11 to May 21 in 2011, and predicted VSDIs (MJO model) were also well matched with actual VSDI. May 11 was at the phase 7 and May 16 was at the phase 8, and Figure 4.3 shows the dryer change from phase 7 to phase 8. As expected, while the original prediction model cannot predict the intensified drought, the MJO-based prediction model can predict the drought (Fig. 4.8).

Although the spatial pattern of predicted VSDI was well matched with actual VSDI, the drought condition of the predicted VSDI was underestimated. The dynamic range of the predicted VSDI was slightly

smaller than the actual VSDI because random forest tends to produce results reducing errors, which leads to the values toward mean values (Im et al. 2016; Park et al. 2016). Although it is difficult to predict the sudden change of drought conditions (wet to dry or dry to wet) considering only previous patterns of droughts (Hao et al. 2016; AghaKouchak 2015; Belayneh et al. 2016), the MJO-based prediction models can improve the limitation. Therefore, it can be seen that MJO-based prediction models are very useful in predicting drought with high resolution (5 km).

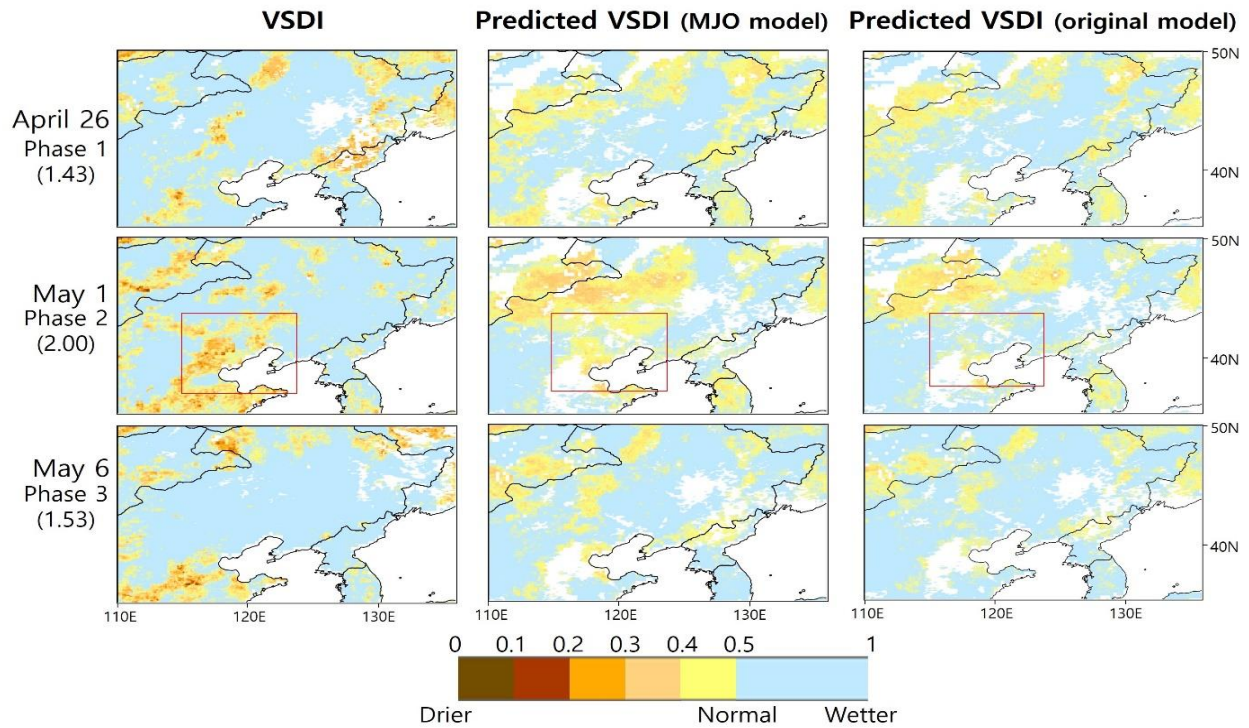


Figure 4.7 Comparison of change of drought conditions between actual VSDI and predicted VSDIs from two models (with MJO and without MJO) in 2010.

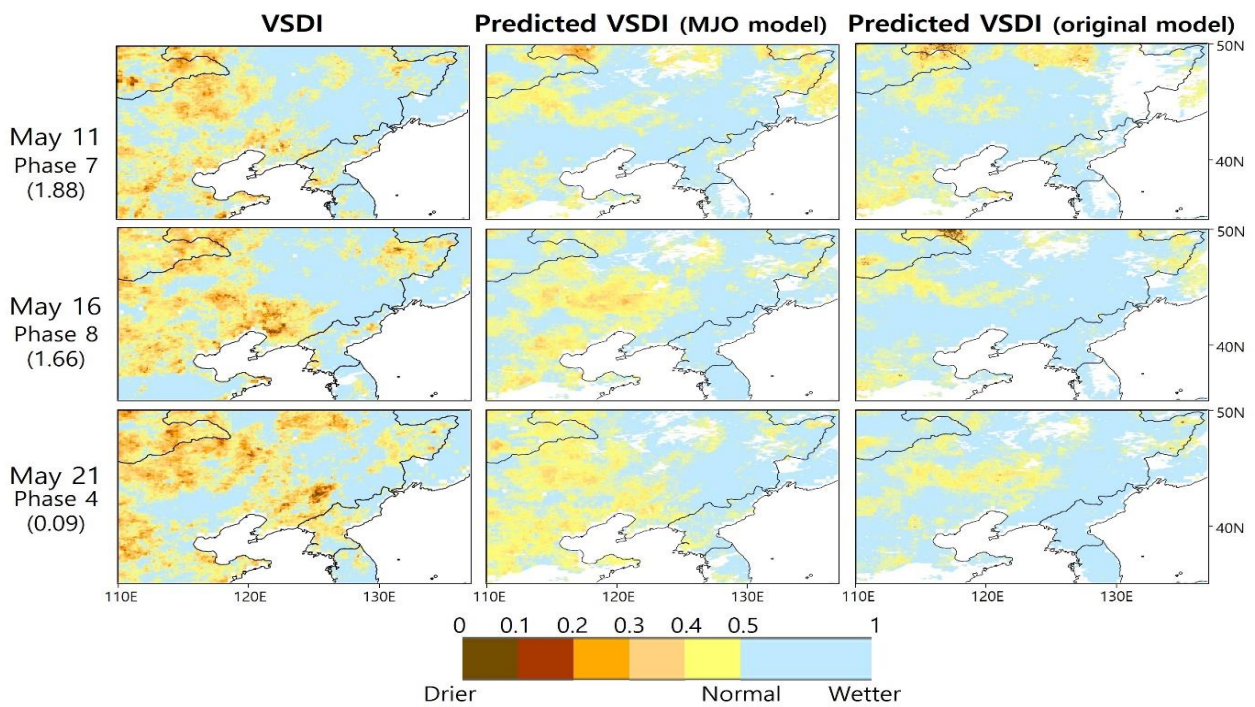


Figure 4.8 Comparison of change of drought conditions between actual VSDI and predicted VSDIs from two models (with MJO and without MJO) in 2011.

4.5. Conclusion

The frequency of rapidly developing droughts for the short-period of time (e.g. flash drought) has increased, and it has damaged on agricultural system. Short-term prediction of drought is important to provide accurate information of drought to decision makers. In this study, short-term prediction model considering climate variability was developed using random forest. Three satellite-based drought indices—SDCI, MIDI, and VSDI—were predicted with very short time scale (one pentad), and RMM MJO indices were used to improve drought predictability because MJO has short time scale variability and closely related to drought factors including precipitation, temperature, and soil moisture. The effect of MJO was evaluated by comparing two models. One included RMM MJO indices and the other excluded RMM MJO indices. The performance of drought prediction model including RMM MJO indices was improved in all drought indices. However, there are differences in improvement of performance among three drought indices. The performance of MJO-based drought prediction model was improved the most in SDCI and the least in VSDI. As stated above, since soil moisture includes intraseasonal variability of MJO, VSDI comprising soil moisture with a high portion (0.5) showed less improvement than MIDI (0.3) and SDCI (0). VSDI which modified SDCI and MIDI was regarded as the most acceptable drought index to predict and monitor drought for the short-period of time because soil moisture is more adequate than precipitation for monitoring drought in a short time scale. Although RMM MJO indices contributed to enhancement of drought prediction, there is still a limitation. The performances of drought prediction models were saturated to 0.7 in correlation in three drought indices. Therefore, in further study, other factors including interannual climate variability such as ENSO, AMO and local or regional characteristics including topography and land use should be investigated to develop more accurate prediction model.

Chapter 5

5. Conclusion

The characteristics of drought factors depending on the duration of drought were identified. Total sixteen drought factors were analyzed using three machine learning approaches targeting a meteorological drought index (SPI) and an agricultural drought index (crop yield), and relative importance of variables derived from the machine learning approaches. RF outperformed the other approaches and was used to select the most important six variables (LST, NDVI, NDWI, NMDI, ET, and TRMM) with the consideration of the inter-relation of the variables. The crop yield based maps showed more consistent distributions with USDM maps compared to the 3-month SPI based maps.

AMSR-E soil moisture with coarse resolution (25 km) was downscaled to 1 km soil moisture using seven surface factors through random forest. Downscaled 1 km soil moisture yielded high correlations with both in situ and AMSR-E soil moisture and produced lower RMSE in situ soil moisture than AMSR-E soil moisture. A novel soil moisture-based drought index, HSMDI, was proposed. HSMDI monitored both meteorological and agricultural droughts well and has characteristics similar to the drought conditions with 3-month precipitation deficiency. However, HSMDI did not model the streamflow very well for hydrological drought.

Short-term prediction model considering climate variability was developed using random forest. Three satellite-based drought indices were predicted with very short time scale (one pentad), and RMM MJO indices were used to improve drought predictability. The performance of drought prediction model including RMM MJO indices was improved in all drought indices. VSDI which modified SDCI and MIDI was proposed to predict and monitor drought for the short-period of time because soil moisture is more adequate than precipitation for monitoring drought in a short time scale.

The approaches proposed in this thesis can be applied to any vegetated region where remote sensing data are available even with limited in situ data availability.

Chapter 6

6. Outlook and future works

There are some limitations which should be improved in the future. Although the each subject of thesis is interconnected (understanding drought factors, drought monitoring, and drought prediction), each study region is different. In future study, those approaches will be applied to one study region (e.g., Africa), and some limitations reported previous studies will be improved.

Testing and Evaluating of developed drought indices will be conducted in broad regions with various reference data to more generalize parameterizations and validate in the proposed approach.

As the current thresholds (e.g., 0.5 for normal condition) of developed drought index to determine drought conditions are arbitrary, specific standards and guidelines on thresholds should follow.

Remote sensing-based optimized approaches to different types of drought (i.e., agricultural drought and hydrological drought) will be developed. Especially, hydrological drought was not well monitored. Future research enhances HSMDI to better model hydrological drought through additional data/product fusion. Factors associated with hydrological drought (e.g., GRACE satellite data, root zone soil moisture) might be incorporated into the downscaling process or the HSMDI generation to improve the performance of drought monitoring.

Although RMM MJO indices contributed to enhancement of drought prediction, the performances of drought prediction models were saturated to 0.7 in correlation in three drought indices. More accurate prediction model will be developed considering various factors including interannual climate variability (e.g., ENSO, AMO) and local or regional characteristics (e.g., topography, land cover).

7. Reference

- Adler, R.F. et al., 2003. The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). *Journal of Hydrometeorology*, 4(6): 1147-1167.
- AghaKouchak, A., 2015. A multivariate approach for persistence-based drought prediction: Application to the 2010–2011 East Africa drought. *Journal of Hydrology*, 526, 127-135.
- AghaKouchak, A. et al., 2015. Remote sensing of drought: Progress, challenges and opportunities. *Reviews of Geophysics*, 53(2): 452-480.
- AghaKouchak, A., 2014. A baseline probabilistic drought forecasting framework using standardized soil moisture index: application to the 2012 United States drought. *Hydrology and Earth System Sciences*, 18(7): 2485-2492.
- Allen, R. et al., 2011. Satellite-based ET estimation in agriculture using SEBAL and METRIC. *Hydrological Processes*, 25(26): 4011-4027.
- Alley, W.M., 1984. The Palmer drought severity index: limitations and assumptions. *Journal of climate and applied meteorology*, 23(7): 1100-1109.
- Anderson, M.C. et al., 2011. Evaluation of drought indices based on thermal remote sensing of evapotranspiration over the continental United States. *Journal of Climate*, 24(8): 2025-2044.
- Belayneh, A., Adamowski, J., Khalil, B., & Ozga-Zielinski, B., 2014. Long-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet neural network and wavelet support vector regression models. *Journal of Hydrology*, 508, 418-429.
- Belayneh, A., Adamowski, J., Khalil, B., & Quilty, J., 2016. Coupling machine learning methods with wavelet transforms and the bootstrap and boosting ensemble approaches for drought prediction. *Atmospheric Research*, 172, 37-47.
- Belayneh, A., & Adamowski, J., 2012. Standard precipitation index drought forecasting using neural networks, wavelet neural networks, and support vector regression. *Applied Computational Intelligence and Soft Computing*, 2012, 6.
- Best, M., Pryor, M., Clark, D., Rooney, G., Essery, R., Ménard, C., Gedney, N., 2011. The Joint UK Land Environment Simulator (JULES), model description–Part 1: energy and water fluxes. *Geoscientific Model Development*, 4(3), 677-699.
- Bolten, J.D., Crow, W.T., Zhan, X., Jackson, T.J., Reynolds, C.A., 2010. Evaluating the utility of remotely sensed soil moisture retrievals for operational agricultural drought monitoring. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of*, 3(1): 57-66.
- Breiman, L., 2001. Random forests. *Machine learning*, 45(1): 5-32.

- Brown, J.F., Wardlow, B.D., Tadesse, T., Hayes, M.J. and Reed, B.C., 2008. The Vegetation Drought Response Index (VegDRI): A new integrated approach for monitoring drought stress in vegetation. *GIScience & Remote Sensing*, 45(1): 16-46.
- Champagne, C., Berg, A., McNairn, H., Drewitt, G. and Huffman, T., 2012. Evaluation of soil moisture extremes for agricultural productivity in the Canadian prairies. *Agricultural and forest meteorology*, 165: 1-11.
- Chen, X., Li, C., Ling, J., & Tan, Y., 2017. Impact of East Asian winter monsoon on MJO over the equatorial western Pacific. *Theoretical and Applied Climatology*, 127(3-4), 551-561.
- Choi, M., Hur, Y., 2012. A microwave-optical/infrared disaggregation for improving spatial representation of soil moisture using AMSR-E and MODIS products. *Remote Sensing of Environment*, 124: 259-269.
- Choi, M., Jacobs, J.M., Anderson, M.C., Bosch, D.D., 2013. Evaluation of drought indices via remotely sensed data with hydrological variables. *Journal of Hydrology*, 476: 265-273.
- Cleverly, J. et al., 2013. Aerodynamic resistance and Penman-Monteith evapotranspiration over a seasonally two-layered canopy in semiarid central Australia. *Journal of Hydrometeorology*, 14(5): 1562-1570.
- Cooke, W.H., Mostovoy, G.V., Anantharaj, V.G. and Jolly, W.M., 2012. Wildfire potential mapping over the state of Mississippi: A land surface modeling approach. *GIScience & Remote Sensing*, 49(4): 492-509.
- Deo, R. C., Kisi, O., & Singh, V. P., 2017. Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. *Atmospheric Research*, 184, 149-175.
- Dobriyal, P., Qureshi, A., Badola, R., Hussain, S.A., 2012. A review of the methods available for estimating soil moisture and its implications for water resource management. *Journal of Hydrology*, 458: 110-117.
- Dorigo, W. et al., 2011. The International Soil Moisture Network: a data hosting facility for global in situ soil moisture measurements. *Hydrology and Earth System Sciences*, 15(5): 1675-1698.
- Dorigo, W., Wagner, W., Albergel, C., Albrecht, F., Balsamo, G., Brocca, L., Gruber, A., 2017. ESA CCI Soil Moisture for improved Earth system understanding: state-of-the art and future directions. *Remote Sensing of Environment*, 203, 185-215.
- Djamai, N. et al., 2016. A combination of DISPATCH downscaling algorithm with CLASS land surface scheme for soil moisture estimation at fine scale during cloudy days. *Remote Sensing of Environment*, 184: 1-14.
- Djamai, N. et al., 2015. Disaggregation of SMOS soil moisture over the Canadian Prairies. *Remote Sensing of Environment*, 170: 255-268.
- Du, L. et al. A comprehensive drought monitoring method integrating MODIS and TRMM data. *International Journal of Applied Earth Observation and Geoinformation* 23, 245-253 (2013).
- Edwards, D.C., 1997. Characteristics of 20th century drought in the United States at multiple time scales, DTIC Document.

- Elith, J., Leathwick, J.R. and Hastie, T., 2008. A working guide to boosted regression trees. *Journal of Animal Ecology*, 77(4): 802-813.
- Finn, M.P. et al., 2011. Remote sensing of soil moisture using airborne hyperspectral data. *GIScience & Remote Sensing*, 48(4): 522-540.
- Fleisher, D.H., Dathe, A., Timlin, D. and Reddy, V., 2015. Improving potato drought simulations: Assessing water stress factors using a coupled model. *Agricultural and Forest Meteorology*, 200: 144-155.
- Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. *Annals of statistics*: 1189-1232.
- Froeschke, J.T. and Froeschke, B.F., 2011. Spatio-temporal predictive model based on environmental factors for juvenile spotted seatrout in Texas estuaries using boosted regression trees. *Fisheries Research*, 111(3): 131-138.
- Gao, B.-C., 1996. NDWI—a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote sensing of environment*, 58(3): 257-266.
- Gao, X., Wu, P., Zhao, X., Wang, J., Shi, Y., 2014. Effects of land use on soil moisture variations in a semi-arid catchment: implications for land and agricultural water management. *Land Degradation & Development*, 25(2): 163-172.
- Gessner, U. et al., 2013. The relationship between precipitation anomalies and satellite-derived vegetation activity in Central Asia. *Global and Planetary Change*, 110: 74-87.
- Gourdji, S.M., Mathews, K.L., Reynolds, M., Crossa, J., Lobell, D.B., 2013. An assessment of wheat yield sensitivity and breeding gains in hot environments. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1752): 20122190.
- Grayson, R.B., Western, A.W., 1998. Towards areal estimation of soil water content from point measurements: time and space stability of mean response. *Journal of Hydrology*, 207(1): 68-82.
- Gruber, A., Dorigo, W. A., Crow, W., & Wagner, W., 2017. Triple Collocation-Based Merging of Satellite Soil Moisture Retrievals. *IEEE Transactions on Geoscience and Remote Sensing*, 55(12), 6780-6792.
- Gu, Y., Brown, J.F., Verdin, J.P. and Wardlow, B., 2007. A five-year analysis of MODIS NDVI and NDWI for grassland drought assessment over the central Great Plains of the United States. *Geophysical Research Letters*, 34(6).
- Han, P., Wang, P. X., Zhang, S. Y., & Zhu, D. H., 2010. Drought forecasting based on the remote sensing data using ARIMA models. *Mathematical and Computer Modelling*, 51(11-12), 1398-1403.
- Handler, A., 1990. USA corn yields, the El Niño and agricultural drought: 1867–1988. *International Journal of climatology*, 10(8): 819-828.
- Hao, Z., AghaKouchak, A., 2013. Multivariate standardized drought index: a parametric multi-index model. *Advances in Water Resources*, 57: 12-18.

- Hao, Z., Hao, F., Singh, V. P., Sun, A. Y., & Xia, Y., 2016. Probabilistic prediction of hydrologic drought using a conditional probability approach based on the meta-Gaussian model. *Journal of Hydrology*, 542, 772-780.
- Hayes, M., Svoboda, M., Le Comte, D., Redmond, K.T. and Pasteris, P., 2005. Drought monitoring: New tools for the 21st century. Dekker, New York.
- Heim Jr, R.R., 2002. A review of twentieth-century drought indices used in the United States. *Bulletin of the American Meteorological Society*, 83(8): 1149-1165.
- Hoell, A., & Funk, C., 2014. Indo-Pacific sea surface temperature influences on failed consecutive rainy seasons over eastern Africa. *Climate dynamics*, 43(5-6), 1645-1660.
- Holben, B. N., 1986. Characteristics of maximum-value composite images from temporal AVHRR data. *International journal of remote sensing*, 7(11), 1417-1434.
- Holmes, T., De Jeu, R., Owe, M., Dolman, A., 2009. Land surface temperature from Ka band (37 GHz) passive microwave observations. *Journal of Geophysical Research: Atmospheres*, 114(D4).
- Hogg, E., Barr, A. and Black, T., 2013. A simple soil moisture index for representing multi-year drought impacts on aspen productivity in the western Canadian interior. *Agricultural and forest meteorology*, 178: 173-182.
- HSC, 2011. 2010 Hydrology Research Report. In: Korean Hydrological Survey Center. <http://www.hsc.re.kr/> (accessed May 2, 2016).
- Hunt, E.D. et al., 2014. Monitoring the effects of rapid onset of drought on non-irrigated maize with agronomic data and climate-based drought indices. *Agricultural and Forest Meteorology*, 191: 1-11.
- Im, J., Jensen, J.R., Coleman, M. and Nelson, E., 2009. Hyperspectral remote sensing analysis of short rotation woody crops grown with controlled nutrient and irrigation treatments. *Geocarto International*, 24, 293-312.
- Im, J., Jensen, J.R., Jensen, R.J., Gladden, J., Waugh, J. and Serrato, M., 2012. Vegetation cover analysis of hazardous waste sites in Utah and Arizona using hyperspectral remote sensing. *Remote Sensing*, 4(2): 327-353.
- Im, J., Park, S., Rhee, J., Baik, J. and Choi, M., 2016. Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches. *Environmental Earth Sciences*, 75(15): 1120.
- IPCC, 2007. Climate change 2007: the physical science basis. *Agenda*, 6(07): 333.
- Jackson, T.J., 1993. III. Measuring surface soil moisture using passive microwave remote sensing. *Hydrological processes*, 7(2): 139-152.
- Jeong, J. H., Kim, B. M., Ho, C. H., & Noh, Y. H. 2008. Systematic variation in wintertime precipitation in East Asia by MJO-induced extratropical vertical motion. *Journal of climate*, 21(4), 788-801.
- Jeong, J.-H., C.-H. Ho, B.-M. Kim, and W.-T. Kwon, 2005, Influence of the Madden-Julian Oscillation on wintertime surface air temperature and cold surges in east Asia, *Journal of Geophysical Research*, 110, D11104, doi:10.1029/2004JD005408.

- Jensen, J.R., Im, J., 2007. Remote sensing change detection in urban environments, In *Geo-Spatial Technologies in Urban Environments (Second Edition): Policy, Practice, and Pixels*. Springer, pp. 7-31.
- Johnson, D.M., 2014. An assessment of pre-and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sensing of Environment*, 141: 116-128.
- Jones, C., Carvalho, L. M., Wayne Higgins, R., Waliser, D. E., & Schemm, J. E. 2004. A statistical forecast model of tropical intraseasonal convective anomalies. *Journal of climate*, 17(11), 2078-2095.
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Woollen, J., 1996. The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, 77(3), 437-471.
- Ke, Y., Im, J., Park, S., Gong, H., 2016. Downscaling of MODIS One Kilometer Evapotranspiration Using Landsat-8 Data and Machine Learning Approaches. *Remote Sensing*, 8(3): 215.
- Keetch, John J., Byram, George M. 1968. A Drought Index for Forest Fire Control. Res. Pap. SE-38. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southeastern Forest Experiment Station, 35 p.
- Keshavarz, M.R., Vazifiedoust, M. and Alizadeh, A., 2014. Drought monitoring using a Soil Wetness Deficit Index (SWDI) derived from MODIS satellite data. *Agricultural Water Management*, 132: 37-45.
- Kim, J., Hogue, T.S., 2012. Improving spatial soil moisture representation through integration of AMSR-E and MODIS products. *Geoscience and Remote Sensing, IEEE Transactions on*, 50(2): 446-460.
- Kim, M. et al., 2015. Landfast sea ice monitoring using multisensor fusion in the Antarctic. *GIScience & Remote Sensing*, 52(2): 239-256.
- Kim, Y.H., Im, J., Ha, H.K., Choi, J.-K., Ha, S., 2014. Machine learning approaches to coastal water quality monitoring using GOCI satellite data. *GIScience & Remote Sensing*, 51(2): 158-174.
- Kogan, F., 1995. Application of vegetation index and brightness temperature for drought detection. *Advances in Space Research*, 15(11): 91-100.
- Koike, T. et al., 2004. Development of an Advanced Microwave Scanning Radiometer (AMSR-E) algorithm of soil moisture and vegetation water content. *PROCEEDINGS OF HYDRAULIC ENGINEERING*, 48: 217-222. DOI:10.2208/prohe.48.217
- Kuhn, M., Weston, S., Keefer, C. and Coulter, N., 2012. Cubist Models For Regression.
- Lambert, J., Drenou, C., Denux, J.-P., Balent, G. and Cheret, V., 2013. Monitoring forest decline through remote sensing time series analysis. *GIScience & Remote Sensing*, 50(4): 437-457.
- Lawrimore, J., Heim Jr, R.R., Svoboda, M., Swail, V. and Englehart, P.J., 2002. Beginning a new era of drought monitoring across North America. *Bulletin of the American Meteorological Society*, 83(8): 1191-1192.
- Lessel, J., Sweeney, A. and Ceccato, P., 2016. An agricultural drought severity index using quasi-climatological anomalies of remotely sensed data. *International Journal of Remote Sensing*, 37(4): 913-925.

- Li, M., Im, J., Beier, C., 2013. Machine learning approaches for forest classification and change analysis using multi-temporal Landsat TM images over Huntington Wildlife Forest. *GIScience & Remote Sensing*, 50(4): 361-384.
- Li, Y. et al., 2015. Integrating soil moisture retrieved from L-band microwave radiation into an energy balance model to improve evapotranspiration estimation on the irrigated oases of arid regions in northwest China. *Agricultural and Forest Meteorology*, 214: 306-318.
- Lin, C.-P., 2003. Frequency domain versus travel time analyses of TDR waveforms for soil moisture measurements. *Soil Science Society of America Journal*, 67(3): 720-729.
- Liu, Y., Dorigo, W. A., Parinussa, R., de Jeu, R. A., Wagner, W., McCabe, M. F., Van Dijk, A., 2012. Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sensing of Environment*, 123, 280-297.
- Lobell, D.B. et al., 2014. Greater sensitivity to drought accompanies maize yield increase in the US Midwest. *Science*, 344(6183): 516-519.
- Logan, K., Brunsell, N., Jones, A., Feddema, J., 2010. Assessing spatiotemporal variability of drought in the US central plains. *Journal of Arid Environments*, 74(2): 247-255.
- Lohani, V.K. and Loganathan, G., 1997. An early warning system for drought management using the Palmer drought index. *Journal of the American Water Resources Association*, 33(6): 1375-1386.
- Long, J.A., Lawrence, R.L., Greenwood, M.C., Marshall, L. and Miller, P.R., 2013. Object-oriented crop classification using multitemporal ETM+ SLC-off imagery and random forest. *GIScience & Remote Sensing*, 50(4): 418-436.
- Lorenz, D. J., Otkin, J. A., Svoboda, M., Hain, C. R., Anderson, M. C., & Zhong, Y., 2017. Predicting US Drought Monitor states using precipitation, soil moisture, and evapotranspiration anomalies. Part I: Development of a nondiscrete USDM index. *Journal of Hydrometeorology*, 18(7), 1943-1962.
- Lu, Z., Im, J., Rhee, J., Hodgson, M., 2014. Building type classification using spatial and landscape attributes derived from LiDAR remote sensing data. *Landscape and Urban Planning*, 130: 134-148.
- Lü, J., Ju, J., Ren, J., & Gan, W., 2012. The influence of the Madden-Julian Oscillation activity anomalies on Yunnan's extreme drought of 2009–2010. *Science China Earth Sciences*, 55(1), 98-112.
- Luo, L., Wood, E.F., 2007. Monitoring and predicting the 2007 U.S. drought. *Geophysical Research Letters*, 34(22): L22702. DOI:10.1029/2007GL031673
- Lyon, B., Bell, M. A., Tippet, M. K., Kumar, A., Hoerling, M. P., Quan, X. W., & Wang, H., 2012. Baseline probabilities for the seasonal prediction of meteorological drought. *Journal of Applied Meteorology and Climatology*, 51(7), 1222-1237.
- MAFRA, 2011. Preliminary Results of the 2010 Agriculture & Fishery Corporation Survey. In: Ministry of Agriculture Food and Rural Affairs. <http://library.mafra.go.kr/skyblueimage/5075.pdf> (accessed May 2, 2016).

- Mallick, K., Bhattacharya, B.K. and Patel, N., 2009. Estimating volumetric surface moisture content for cropped soils using a soil wetness index based on surface temperature and NDVI. *Agricultural and Forest Meteorology*, 149(8): 1327-1342.
- Mallya, G., Zhao, L., Song, X., Niyogi, D. and Govindaraju, R., 2013. 2012 Midwest Drought in the United States. *Journal of Hydrologic Engineering*, 18(7): 737-745.
- Martínez-Fernández, J., González-Zamora, A., Sánchez, N., Gumuzzio, A. and Herrero-Jiménez, C., 2016. Satellite soil moisture for agricultural drought monitoring: Assessment of the SMOS derived Soil Water Deficit Index. *Remote Sensing of Environment*, 177: 277-286.
- Maxell, A., Strager, M., Warner, T., Zegre, N., and Yuill, C. 2014. Comparison of NAIP orthophotography and RapidEye satellite imagery for mapping of mining and mine reclamation. *GIScience and Remote Sensing*, 51: 301-320.
- McKee, T.B., Doesken, N.J. and Kleist, J., 1995. Drought monitoring with multiple time scales, Ninth Conference on Applied Climatology. American Meteorological Society, Boston, pp. 233-236.
- McVicar, T.R. and Jupp, D.L., 1998. The current and potential operational uses of remote sensing to aid decisions on drought exceptional circumstances in Australia: a review. *Agricultural systems*, 57(3): 399-468.
- Merlin, O. et al., 2013. Self-calibrated evaporation-based disaggregation of SMOS soil moisture: An evaluation study at 3km and 100m resolution in Catalunya, Spain. *Remote sensing of environment*, 130: 25-38.
- Mishra, A.K. et al., 2015. Anatomy of a local-scale drought: Application of assimilated remote sensing products, crop model, and statistical methods to an agricultural drought study. *Journal of Hydrology*, 526: 15-29.
- Mishra, A. K., & Singh, V. P., 2011. Drought modeling—A review. *Journal of Hydrology*, 403(1-2), 157-175.
- Mizzell, H., 2008. Improving drought detection in the Carolinas: evaluation of local, state, and federal drought indicators, University of South Carolina.
- Mitchell KE, Lohmann D, Houser PR, Wood EF, Schaake JC, Robock A, Cosgrove BA, Sheffield J, Duan Q, Luo L, 2004. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *Journal of Geophysical Research: Atmospheres* 109(D7)
- Mo, K. C., Shukla, S., Lettenmaier, D. P., & Chen, L. C., 2012. Do Climate Forecast System (CFSv2) forecasts improve seasonal soil moisture prediction?. *Geophysical Research Letters*, 39(23).
- Mo, K. C., & Lettenmaier, D. P., 2015. Heat wave flash droughts in decline. *Geophysical Research Letters*, 42(8), 2823-2829.
- Morid, S., Smakhtin, V., & Bagherzadeh, K., 2007. Drought forecasting using artificial neural networks and time series of drought indices. *International Journal of Climatology*, 27(15), 2103-2111.

- Molero, B. et al., 2016. SMOS disaggregated soil moisture product at 1km resolution: Processor overview and first validation results. *Remote Sensing of Environment*, 180: 361-376.
- Narasimhan, B. and Srinivasan, R., 2005. Development and evaluation of Soil Moisture Deficit Index (SMDI) and Evapotranspiration Deficit Index (ETDI) for agricultural drought monitoring. *Agricultural and Forest Meteorology*, 133(1): 69-88.
- Nezlin, N., Kostianoy, A. and Li, B.-L., 2005. Inter-annual variability and interaction of remote-sensed vegetation index and atmospheric precipitation in the Aral Sea region. *Journal of Arid Environments*, 62(4): 677-700.
- Otkin, J. A., Anderson, M. C., Hain, C., & Svoboda, M., 2014. Examining the relationship between drought development and rapid changes in the evaporative stress index. *Journal of Hydrometeorology*, 15(3), 938-956.
- Otkin, J. A., Anderson, M. C., Hain, C., & Svoboda, M., 2015. Using temporal changes in drought indices to generate probabilistic drought intensification forecasts. *Journal of Hydrometeorology*, 16(1), 88-105.
- Otkin, J.A. et al., 2013. Examining rapid onset drought development using the thermal infrared-based evaporative stress index. *Journal of Hydrometeorology*, 14(4): 1057-1074.
- Otkin, J.A. et al., 2016. Assessing the evolution of soil moisture and vegetation conditions during the 2012 United States flash drought. *Agricultural and Forest Meteorology*, 218: 230-242.
- Özger, M., Mishra, A. K., & Singh, V. P., 2012. Long lead time drought forecasting using a wavelet and fuzzy logic combination model: a case study in Texas. *Journal of Hydrometeorology*, 13(1), 284-297.
- Palmer, W.C., 1965. Meteorological drought. Research paper No. 45. United States Department of Commerce. Weather Bureau, Washington, DC.
- Palmer, W.C., 1968. Keeping Track of Crop Moisture Conditions, Nationwide: The New Crop Moisture Index. *Weatherwise*, 21(4): 156-161. DOI:10.1080/00431672.1968.9932814
- Parinussa, R., Yilmaz, M., Anderson, M., Hain, C. and Jeu, R., 2014. An intercomparison of remotely sensed soil moisture products at various spatial scales over the Iberian Peninsula. *Hydrological Processes*, 28(18): 4865-4876.
- Park, S., Feddema, J., Egbert, S., 2005. MODIS land surface temperature composite data and their relationships with climatic water budget factors in the central Great Plains. *International Journal of Remote Sensing*, 26(6): 1127-1144.
- Park, S., Im, J., Jang, E., Rhee, J., 2016. Drought assessment and monitoring through blending of multi-sensor indices using machine learning approaches for different climate regions. *Agricultural and Forest Meteorology*, 216: 157-169.
- Peng, J., Loew, A., and Crueger, T., 2017, The relationship between the Madden-Julian oscillation and the land surface soil moisture. *Remote Sensing of Environment*, 203, 226-239.

- Piao, S. et al., 2003. Interannual variations of monthly and seasonal normalized difference vegetation index (NDVI) in China from 1982 to 1999. *Journal of Geophysical Research: Atmospheres* (1984–2012), 108(D14).
- Potop, V., Možný, M. and Soukup, J., 2012. Drought evolution at various time scales in the lowland regions and their impact on vegetable crops in the Czech Republic. *Agricultural and Forest Meteorology*, 156: 121-133.
- Prasad, A.M., Iverson, L.R. and Liaw, A., 2006. Newer classification and regression tree techniques: bagging and random forests for ecological prediction. *Ecosystems*, 9(2): 181-199.
- Quiring, S.M., Papakryiakou, T.N., 2003. An evaluation of agricultural drought indices for the Canadian prairies. *Agricultural and forest meteorology*, 118(1): 49-62.
- Quiring, S.M. and Ganesh, S., 2010. Evaluating the utility of the Vegetation Condition Index (VCI) for monitoring meteorological drought in Texas. *Agricultural and Forest Meteorology*, 150(3): 330-339.
- RDA. The schedule of agricultural works. In: Korean Rural Development Administration. <http://www.nongsaro.go.kr/portal/ps/psb/psbl/workScheduleLst.ps?menuId=PS00087#1> (accessed May 2, 2016).
- Rhee, J., Im, J., Carbone, G.J., Jensen, J.R., 2008. Delineation of climate regions using in-situ and remotely-sensed data for the Carolinas. *Remote Sensing of Environment*, 112(6): 3099-3111.
- Rhee, J., Im, J. and Carbone, G.J., 2010. Monitoring agricultural drought for arid and humid regions using multi-sensor remote sensing data. *Remote Sensing of Environment*, 114(12): 2875-2887.
- Rhee, J., Park, S., Lu, Z., 2014. Relationship between land cover patterns and surface temperature in urban areas. *GIScience & Remote Sensing*, 51(5): 521-536.
- Rhee, J., Im, J. and Park, S., 2015. Regional Drought Monitoring Based on Multi-Sensor Remote Sensing. In: P.S. Thenkabail (Editor), *Remote Sensing of Water Resources, Disasters and Urban Studies*. CRC Press.
- Rhee, J., & Im, J., 2017. Meteorological drought forecasting for ungauged areas based on machine learning: Using long-range climate forecast and remote sensing data. *Agricultural and Forest Meteorology*, 237, 105-122.
- Riggins, J., Tullis, J. and Stephen, F., 2009. Per-segment aboveground forest biomass estimation using LIDAR-derived height percentile statistics. *GIScience and Remote Sensing*, 46(2): 232-248.
- Rosenbaum, U. et al., 2012. Seasonal and event dynamics of spatial soil moisture patterns at the small catchment scale. *Water Resources Research*, 48(10).
- RuleQuest, 2012. <http://www.rulequest.com>.
- Ryu, J.H., Sohrabi, M., Acharya, A., 2014. Toward mapping gridded drought indices to evaluate local drought in a rapidly changing global environment. *Water resources management*, 28(11): 3859-3869.
- Sánchez, N., González-Zamora, Á., Piles, M. and Martínez-Fernández, J., 2016. A New Soil Moisture Agricultural Drought Index (SMADI) Integrating MODIS and SMOS Products: A Case of Study over the Iberian Peninsula. *Remote Sensing*, 8(4): 287.

- Salvucci, G.D., Saleem, J.A., Kaufmann, R., 2002. Investigating soil moisture feedbacks on precipitation with tests of Granger causality. *Advances in water Resources*, 25(8): 1305-1312.
- Seneviratne, S. I., Koster, R. D., Guo, Z., Dirmeyer, P. A., Kowalczyk, E., Lawrence, D., & Verseghy, D., 2006. Soil moisture memory in AGCM simulations: analysis of global land-atmosphere coupling experiment (GLACE) data. *Journal of Hydrometeorology*, 7(5), 1090-1112.
- Shahabfar, A. and Eitzinger, J., 2013. Spatio-Temporal Analysis of Droughts in Semi-Arid Regions by Using Meteorological Drought Indices. *Atmosphere*, 4(2): 94-112.
- Sheffield, J., Goteti, G., & Wood, E. F., 2006. Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, 19(13), 3088-3111.
- Stackhouse P, Gupta S, Cox S, Mikovitz J, Zhang T, Chiacchio M, 2004. 12-year surface radiation budget data set. *Gewex News* 14(4):10-12
- Stumpf, A. and Kerle, N., 2011. Object-oriented mapping of landslides using Random Forests. *Remote Sensing of Environment*, 115(10): 2564-2577.
- Svoboda, M. et al., 2002. The drought monitor. *Bulletin of the American Meteorological Society*, 83(8): 1181-1190.
- Swain, S., Wardlow, B.D., Narumalani, S., Tadesse, T. and Callahan, K., 2011. Assessment of Vegetation Response to Drought in Nebraska Using Terra-MODIS Land Surface Temperature and Normalized Difference Vegetation Index. *GIScience & Remote Sensing*, 48(3): 432-455.
- Sugathan, N., Biju, V., Renuka, G., 2014. Influence of soil moisture content on surface albedo and soil thermal parameters at a tropical station. *Journal of Earth System Science*, 123(5): 1115-1128.
- Tadesse, T., Brown, J.F., Hayes, M.J., 2005. A new approach for predicting drought-related vegetation stress: Integrating satellite, climate, and biophysical data over the US central plains. *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(4): 244-253.
- Tadesse, T., Wardlow, B.D., Hayes, M.J., Svoboda, M.D. and Brown, J.F., 2010. The vegetation outlook (VegOut): A new method for predicting vegetation seasonal greenness. *GIScience & Remote Sensing*, 47(1): 25-52.
- Tadesse, T., Demisse, G.B., Zaitchik, B. and Dinku, T., 2014. Satellite-based hybrid drought monitoring tool for prediction of vegetation condition in Eastern Africa: A case study for Ethiopia. *Water Resources Research*, 50(3): 2176-2190.
- Tang, H., Zhang, T., Yang, X., Li, H., 2015. Classification of different types of slag samples by laser-induced breakdown spectroscopy (LIBS) coupled with random forest based on variable importance (VIRF). *Analytical Methods*, 7(21): 9171-9176.
- Teixeira, E.I., Fischer, G., van Velthuizen, H., Walter, C., Ewert, F., 2013. Global hot-spots of heat stress on agricultural crops due to climate change. *Agricultural and Forest Meteorology*, 170: 206-215.
- Torbick, N. and Corbiere, M., 2015. Mapping urban sprawl and impervious surfaces in the northeast United States for the past four decades. *GIScience and Remote Sensing*, 52, 746-764.

- USDA, 2014. Census of Agriculture State and County Profiles. In: U.S.D.o. Agriculture (Editor).
- Van Loon, A. F., Gleeson, T., Clark, J., Van Dijk, A. I., Stahl, K., Hannaford, J., ... & Hannah, D. M. (2016). Drought in the Anthropocene. *Nature Geoscience*, 9(2), 89-91.
- Vicente-Serrano, S.M., Beguería, S. and López-Moreno, J.I., 2010. A multiscalar drought index sensitive to global warming: the standardized precipitation evapotranspiration index. *Journal of Climate*, 23(7): 1696-1718.
- Vicente-Serrano, S.M. et al., 2013. Response of vegetation to drought time-scales across global land biomes. *Proceedings of the National Academy of Sciences*, 110(1): 52-57.
- Wallander, S., Aillery, M., Hellerstein, D., & Hand, M., 2013. The role of conservation programs in drought risk adaptation.
- Wang, B., Wu, R., & Fu, X., 2000. Pacific–East Asian teleconnection: how does ENSO affect East Asian climate?. *Journal of Climate*, 13(9), 1517-1536.
- Wang, L. and Qu, J.J., 2007. NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophysical Research Letters*, 34(20).
- Wang, L., Yuan, X., Xie, Z., Wu, P., & Li, Y., 2016. Increasing flash droughts over China during the recent global warming hiatus. *Scientific Reports*, 6, 30571.
- Wardlow, B.D., Anderson, M.C. and Verdin, J.P., 2012. Remote sensing of drought: innovative monitoring approaches. CRC Press.
- Werbylo, K.L. and Niemann, J.D., 2014. Evaluation of sampling techniques to characterize topographically-dependent variability for soil moisture downscaling. *Journal of Hydrology*, 516: 304-316.
- Wheeler, M. C., & Hendon, H. H., 2004. An all-season real-time multivariate MJO index: Development of an index for monitoring and prediction. *Monthly Weather Review*, 132(8), 1917-1932.
- Wheeler, T.R., Craufurd, P.Q., Ellis, R.H., Porter, J.R., Prasad, P.V., 2000. Temperature variability and the yield of annual crops. *Agriculture, Ecosystems & Environment*, 82(1): 159-167.
- Wilhite, D.A. and Glantz, M.H., 1985. Understanding: the drought phenomenon: the role of definitions. *Water international*, 10(3): 111-120.
- Wilhite, D.A., 2000. Drought as a natural hazard: concepts and definitions. *Drought, a global assessment*, 1: 3-18.
- Wilhite, D.A., Svoboda, M.D., Hayes, M.J., 2007. Understanding the complex impacts of drought: a key to enhancing drought mitigation and preparedness. *Water Resources Management*, 21(5): 763-774.
- Williams, A. P., & Funk, C., 2011. A westward extension of the warm pool leads to a westward extension of the Walker circulation, drying eastern Africa. *Climate Dynamics*, 37(11-12), 2417-2435.
- Witt, J.L., 1997. National mitigation strategy: partnerships for building safer communities. Diane Publishing.
- Wu, H., Hayes, M.J., Weiss, A. and Hu, Q., 2001. An evaluation of the Standardized Precipitation Index, the China-Z Index and the statistical Z-Score. *International journal of climatology*, 21(6): 745-758.

- Wu, Z., Wang, B., Li, J., & Jin, F. F., 2009. An empirical seasonal prediction model of the East Asian summer monsoon using ENSO and NAO. *Journal of Geophysical Research: Atmospheres*, 114(D18).
- Yagci, A., Di, L. and Deng, M., 2015. The effect of corn-soybean rotation on the NDVI-based drought indicators: a case study in Iowa, USA, using Vegetation Condition Index. *GIScience and Remote Sensing*, 52(3): 290-314.
- Yang, J., Liu, Q., Xie, S. P., Liu, Z., & Wu, L., 2007. Impact of the Indian Ocean SST basin mode on the Asian summer monsoon. *Geophysical Research Letters*, 34(2).
- Yee, M.S., Walker, J.P., Monerris, A., Rüdiger, C. and Jackson, T.J., 2016. On the identification of representative in situ soil moisture monitoring stations for the validation of SMAP soil moisture products in Australia. *Journal of Hydrology*, 537: 367-381.
- Zhang, L., Jiao, W., Zhang, H., Huang, C., & Tong, Q., 2017. Studying drought phenomena in the Continental United States in 2011 and 2012 using various drought indices. *Remote sensing of environment*, 190, 96-106.
- Zhang, T., Simelton, E., Huang, Y., Shi, Y., 2013. A Bayesian assessment of the current irrigation water supplies capacity under projected droughts for the 2030s in China. *Agricultural and Forest Meteorology*, 178: 56-65.
- Zhang, A., Jia, G., 2013. Monitoring meteorological drought in semiarid regions using multi-sensor microwave remote sensing data. *Remote Sensing of Environment*, 134: 12-23.
- Zhu, Q. et al., 2017. Integrating real-time and manual monitored data to predict hillslope soil moisture dynamics with high spatio-temporal resolution using linear and non-linear models. *Journal of Hydrology*, 545: 1-11.
- Zipper, S.C. and Loheide II, S.P., 2014. Using evapotranspiration to assess drought sensitivity on a subfield scale with HRMET, a high resolution surface energy balance model. *Agricultural and Forest Meteorology*, 197: 91-102.

