Doctoral Thesis

Drought forecasts using satellite data based on deep learning over East Asia

Sumin Park

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

Ulsan National Institute of Science and Technology

2021

Drought forecasts using satellite data based on deep learning over East Asia

Sumin Park

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

Ulsan National Institute of Science and Technology

Drought forecasts using satellite data based on deep learning over East Asia

A thesis/dissertation submitted to
Ulsan National Institute of Science and Technology
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Sumin Park

12/17/2020

Approved by

Advisor
Jungho Im

Drought forecasts using satellite data based on deep learning over East Asia

Sumin Park

This certifies that the thesis/dissertation of Sumin Park is approved.

12/17/2020

Advisor: Jungho Im

Signature

Thesis Committee Member #1 : Myong-In Lee

Signature

Thesis Committee Member #2 : Dong-Hyun Cha

Signature

Thesis Committee Member #3 : Jonghun Kam

Signature

Thesis Committee Member #4 : Sunyurp Park

Abstract

This thesis/dissertation seeks to 1) forecast drought conditions effectively considering temporal patterns of drought indices and upcoming weather conditions through the deep learning approach, and 2) forecast drought by identifying the teleconnection effect based on the sea surface temperature through the deep learning approach.

In this thesis/dissertation, there are four chapters. Chapter 1 summarizes the background of the research and overviews of the thesis research. In Chapter 2, drought-forecasting models on a short-term scale (8 days) were developed considering the temporal patterns of satellite-based drought indices and numerical model outputs through the synergistic use of convolutional long short term memory (ConvLSTM) and random forest (RF) approaches over a part of East Asia. Through the combination of temporal patterns and the upcoming weather conditions (numerical model outputs), the overall performances of drought-forecasting models (ConvLSTM and RF combined) produced competitive results. Furthermore, our short-term drought-forecasting model can be effective regardless of drought intensification or alleviation. The proposed drought-forecasting model can be operationally used, providing useful information on upcoming drought conditions with high resolution (0.05°). In Chapter 3, the Drought forecasting model on a mid-and long-term scale (one-three lead time) over East Asia was developed using temporal patterns of drought indices and teleconnection phenomena of SST through the CNN. Reanalysis based drought index, SPI, were selected with a mid- and long-timescale (one to three months), and satellite-based variable, precipitation and SST across the Pacific Ocean. As the lead time increased, the accuracy tended to fall, but it showed good results compared to CFS. When compared to a drought case, the SST of 8 months ago influenced on the results. Chapter 4 provides a brief summary of these studies

Contents

Chapter 1

Chapter 2

Chapter 3

List of figures

List of Tables

Drought forecasts using satellite data based on deep learning over East Asia

Chapter 1

1. Introduction

1.1 Background

Drought is a major natural disaster that generally occurs due to the deficit of precipitation (West et al., 2019). Drought can result in substantial damage to vegetation and crop conditions, human communities and economies depending on its persistence. Low levels of precipitation for weeks to months cause meteorological drought (Zhang and Jia, 2013). When meteorological drought is prolonged and the amount of water required for plant growth is insufficient, agricultural drought occurs. Hydrological drought indicates a shortage of water resources, such as a significant reduction in groundwater, reservoir or stream levels (Thomas et al., 2017), while socio-economic drought refers to the impact on the environment when water demands for agricultural, industry and living exceed the available water supply (Wu et al., 2013; Park et al., 2017; Zhang et al., 2017; Tu et al., 2018). According to EM-DAT provided by the Centre for Research on the Epidemiology of Disasters (CRED, 2019), there have been 33 drought events with the economic loss of $18 billion worldwide from 2008 to 2018. The drought monitoring and forecasting are essential to appropriate response and minimize the damage. In particular, drought forecast is more helpful for securing the prevention time from drought events than drought monitoring. In other words, drought forecasting plays a vital role in risk management as a comprehensive preparation and mitigation of potential drought-caused damages in a timely manner (Zhang et al., 2019; Belayneh et al., 2016; Demisse et al., 2019; Rhee and Im, 2017).

For this reason, In the United States, information about upcoming drought conditions can obtain from National Oceanic and Atmospheric Administration (NOAA) Climate Weather Service Climate Prediction Center (CPC, https://www.cpc.ncep.noaa.gov, Figure 1.1) and the National Integrated Drought Information System's (NIDIS's) Drought Early Warning Systems (DEWS) (https://www.drought.gov). Those systems provide monthly and seasonal drought information by combining United States Drought Monitor (USDM) data, various satellite observations (e.g., products

or reflectance of Advanced Very-High-Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS)), satellite-based drought indices (e.g., Vegetation Drought Response Index (VegDRI)) and climate forecasting models (e.g., Climate Forecast System (CFS)).

In South Korea, drought forecasting information can be obtained from the National Drought Information-Analysis Center (www.drought.go.kr, Figure 1.2), which has been operated by the Ministry of Environment and Korea Water Resources Corporation. They provide present and upcoming drought conditions up to three months using a standardized precipitation index (SPI), the present conditions of water storage, and the available water capacity of the soil and water level of dams. China officially provides the present drought conditions using ground measurements and satellite data, while they provide only the forecast of each drought factor (e.g., precipitation and El Niño-Southern Oscillation (ENSO) from climate models) (Weather China, http://products.weather.com.cn). No drought forecast systems are blending in situ and satellite observations with climate model output, unlike in the United States.

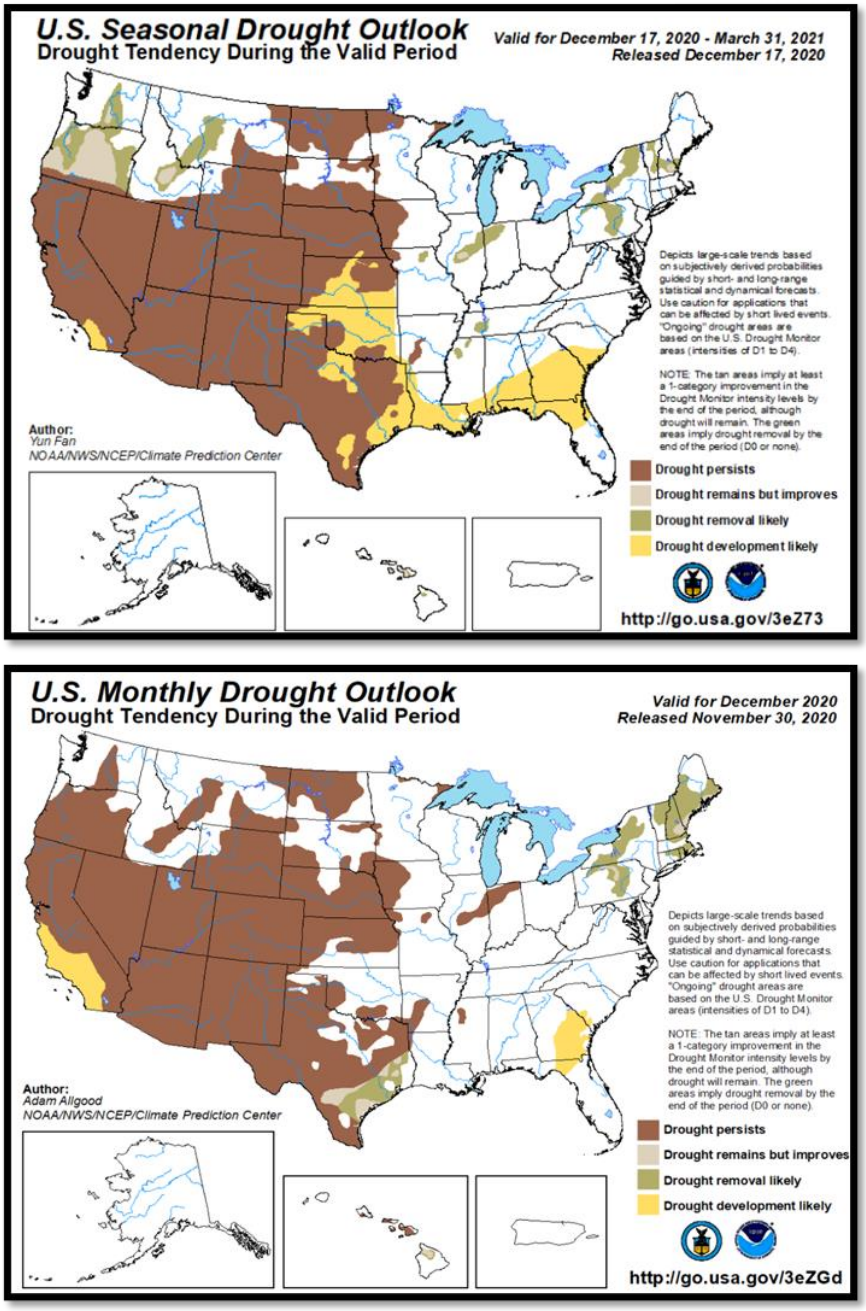

Figure 1.1 Seasonal(above) and Monthly (below) Drought Outlook from Climate Prediction Center (https://www.cpc.ncep.noaa.gov) across the United States

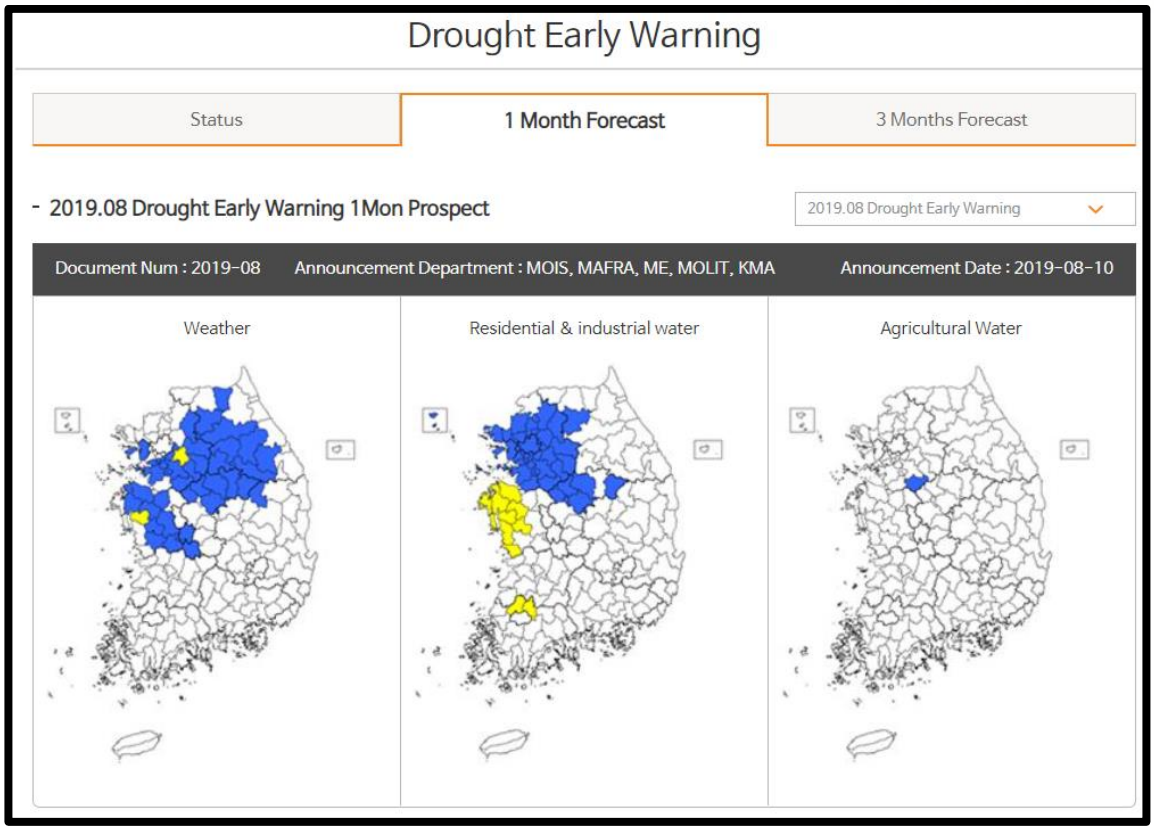

Figure 1.2 Monthly Drought Outlook from National Drought Information-Analysis Center (www.drought.go.kr) across South Korea

1.2 Previous research for drought forecasts

1.2.1 Drought indices for drought forecasts

A number of studies for drought forecast have been conducted using hydro-meteorological variables (e.g., precipitation, temperature, evapotranspiration, soil moisture and streamflow), drought indices (e.g., SPI, standardized precipitation evapotranspiration index (SPEI), evaporative stress index (ESI)) and climate indices (e.g., sea surface temperature (SST) and Atlantic Multidecadal Oscillation) (Figure 1.3, Mishra and Singh, 2011). Hydro-meteorological variables and drought indices were mainly used for short-term forecasts of drought (within two weeks), and climate indices were used for sub-seasonal or seasonal forecasts of drought in addition to variables used for short-term forecasts of drought. Among three components (i.e., hydro-meteorological variables, drought indices and climate indices), drought indices are effectively presenting different drought types and drought conditions (e.g., intensity, duration and severity) as fusing drought factors. For these reasons, most studies used three components (i.e., hydro-meteorological variables, drought indices, and climate indices) for independent variables and drought indices for dependent variables.

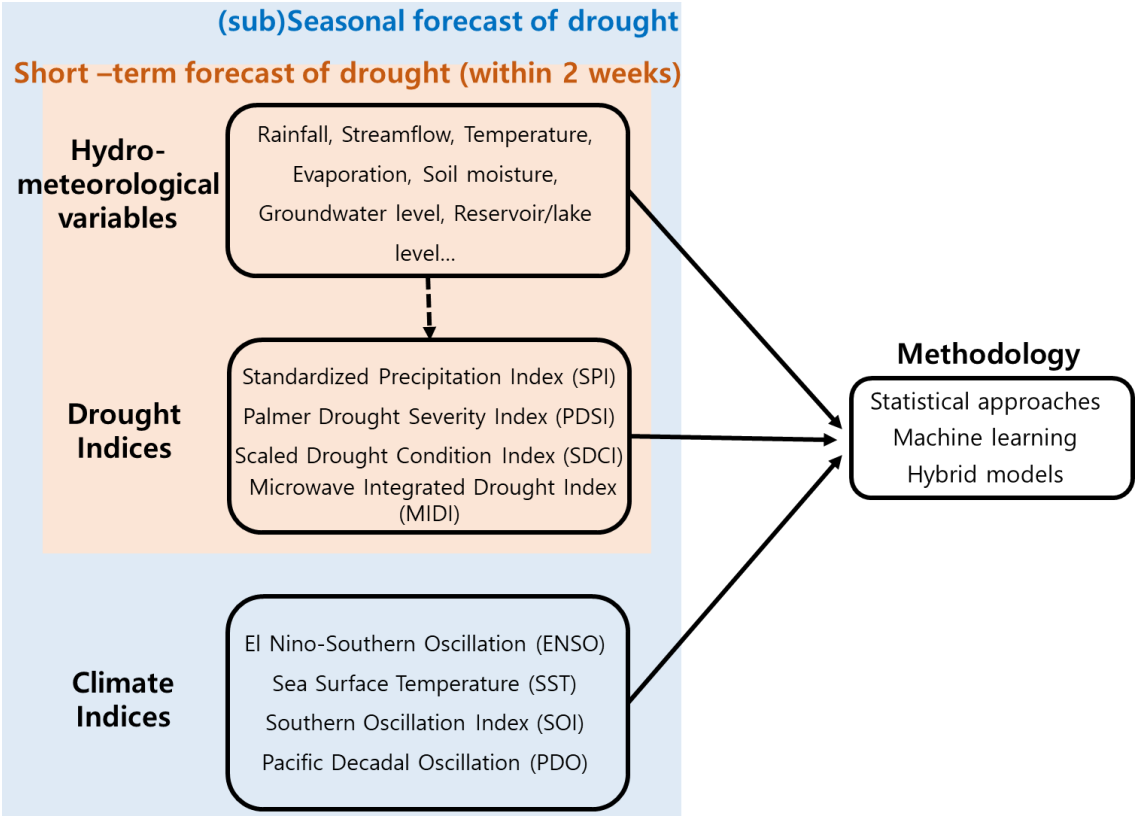

Figure 1.3 Component for drought forecast (Mishra and Singh, 2011)

Drought indices were generated by using drought factors (e.g., precipitation and temperature) obtained from ground stations, satellite, and model. Surface-based drought indices have been widely used for drought forecasts, including Standardized Precipitation Index (SPI), Palmer Drought Severity Index (PDSI), Standardized Precipitation Evapotranspiration Index (SPEI) and Palmer Moisture Anomaly Index (Z-index) (Mishra and Singh, 2011; Zhang and Jia., 2013; Fang et al., 2019). However, point-based observation stations are often sparse and these indices can suffer from high levels of uncertainty when interpolating an insufficient number of stations across vast areas (Jiao et al., 2019; Rhee et al., 2010). Remote sensing data can mitigate this problem by providing several key drought-related factors such as land surface temperature (LST), precipitation, vegetation indices (VI), evapotranspiration (ET), and soil moisture (Bayissa et al., 2019).

Researchers have proposed various satellite-based drought indicators for different study areas. Some have been developed at continental scales, such as the evapotranspiration stress index (ESI, Anderson et al., 2007), a vegetation index based on the universal pattern decomposition method based Vegetation condition index (VIUPD-VCI, Jiao et al., 2016), and the vegetation drought index (VDI, Sun et al., 2013). Others were developed for application at local or regional scales (e.g., one or two states for United States (US) and provinces for China) such as Multivariate Standardized Drought Index (MSDI) for California and North Carolina (Hao and Aghakouchak, 2013), Vegetation Drought Response Index (VegDRI) for north-central US (Brown et al., 2008), Microwave Integrated Drought Index (MIDI) for North China (Zhang and Jia, 2013), Optimized Meteorological Drought Index (OMDI) for Southwest China (Hao et al., 2015), and Vector Projection Index of Drought (VPID) for East Asia and United States (Son et al., 2021).

1.2.2 Methods for drought forecasts

A variety of methods were used for drought forecasting. In the early 2000s, the regression model was tested using Normalized Different Vegetation Index (NDVI), Oscillation Index (SOI), SST, ENSO (Kumar and Panu., 1997, Leilah and Al-Khateeb., 2005, Liu and Negron-Juarez., 2001). Time series models were applied, taking characteristics that droughts gradually intensified and alleviated in general (Mishra and Desai, 2005; Durdu., 2010; Modarres., 2007; Han et al., 2010, Fernandez et al., 2009). Recently, machine learning approaches were used in drought forecasting. Belayneh et al.(2014) used SPI for drought forecasting in the Awash River Basin by combining wavelet neural networks and wavelet support vector regression approaches. Park and Kim. (2019) developed Prediction of Severe Drought Area based on Random Forest using Landsat8 and SRTM based SMI and Topography Data. Park et al. (2018) used time-series of satellite-based drought indices

(SDCI, MIDI, VSDI) and climate indices (MJO) for developing a drought forecasting model through random forest.

1.3 Objectives

In previous studies, a variety of methods were applied to drought forecasts. However, there are some limitations: 1) Some drought forecasts were conducted only using ground-based drought indices in the local region, especially in short-term drought forecasts. The model based on ground-based drought indices cannot be applied to the ungauged area. 2) Most studies applied only time-series patterns or upcoming conditions of hydro-meteorological variables obtained from numerical models. 3) When generating drought forecasting models, geographical information was not considered. However, depending on the geographical characteristics, there are different characteristics in the intensification and alleviation of droughts (especially in the short-term forecast). 4) Improvements in forecasting skills are needed by applying new approaches to reflect the complex mechanisms of drought. 5) In East Asia, there are no drought forecast systems blending in situ and satellite observations and climate model output, unlike the United States.

The ultimate goal of this dissertation is the development of drought forecasting models over East Asia . The research hypothesis is that 1) drought can be forecasted well if the drought indices suitable for the research area are well predicted, 2) Blending temporal patterns of drought and upcoming weather conditions can improve drought forecasting skills, and 2) Deep learning approaches can enhance drought forecasting skills. This thesis focuses on 1) development of short-term forecast model of drought considering the temporal patterns of satellite-based drought indices and numerical model outputs through the synergistic use of convolutional long short term memory (ConvLSTM) and random forest (RF) approaches (Chapter 2), and 2) (sub) seasonal forecasts models of drought considering SST teleconnection phenomena through a convolutional neural network (CNN) (Chapter 3) (Figure 1.4).

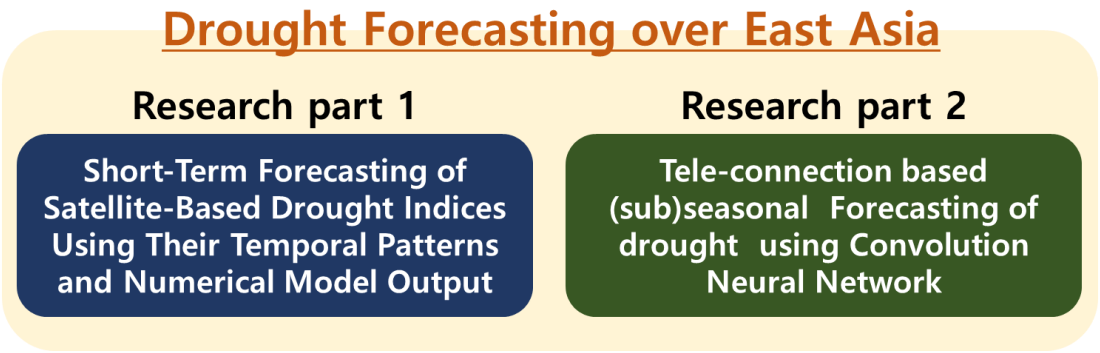

Figure 1.4 The structure of this dissertation

1.4 Overview of papers

Chapter 1 contains the part of the published papers (Son et al., 2021; Park et al., 2017), Chapter 2 has been published (Park et al., 2020) and chapter 3 is currently being drafted. Below are the summaries:

Chapter 1 :

Bokyung Son, **Sumin Park**, Jungho Im, Seohui Park, Yinghai Ke, Lindi J.Quackenbush. (2021). **A new drought monitoring approach: Vector Projection Analysis**. *Remote Sensing of Environment*, 252, 112145

Seonyoung Park, **Sumin Park**, Jungho Im, Jinyoung Rhee, Jinho Shin, Jun Dong Park. (2017). **Downscaling GLDAS Soil Moisture Data in East Asia through Fusion of Multi-Sensors by Optimizing Modified Regression Trees**. *Water*, 9(5), 332

Chapter 2 :

Sumin Park, Jungho Im, Daehyeon Han, Jinyoung Rhee. (2020). **Short-Term Forecasting of Satellite-Based Drought Indices Using Their Temporal Patterns and Numerical Model Output**. *Remote Sensing*, 12(21), 3499

Drought-forecasting models on a short-term scale (8 days) were developed considering the temporal patterns of satellite-based drought indices and numerical model outputs through the synergistic use of convolutional long short term memory (ConvLSTM) and random forest (RF) approaches over a part of East Asia. Two widely used drought indices—Scaled Drought Condition Index (SDCI) and Standardized Precipitation Index (SPI)—were used as target variables. Through the combination of temporal patterns and the upcoming weather conditions (numerical model outputs), the overall performances of drought-forecasting models (ConvLSTM and RF combined) produced competitive results in terms of r (0.90 and 0.93 for validation SDCI and SPI, respectively) and nRMSE (0.11 and 0.08 for validation of SDCI and SPI, respectively). Furthermore, our short-term drought-forecasting model can be effective regardless of drought intensification or alleviation. The proposed drought-forecasting model can be operationally used, providing useful information on upcoming drought conditions with high resolution (0.05°).

Chapter 3 :

Sumin Park, Juhyun Lee, Jungho Im, Sungmun Sim, Eunkyo Seo. (2021) **Tele-connection based (sub)seasonal Forecasting of drought using Convolution Neural Network**, to be submitted.

Drought-forecasting models on sub-seasonal and seasonal scales were developed considering temporal patterns of drought indices and teleconnection phenomena of SST through the CNN. Reanalysis based drought index, SPI, were selected with a sub-seasonal and seasonal timescale (one to three months), and satellite-based variable, precipitation and SST across the Pacific Ocean. The SPI-based drought forecasting models proposed in this study showed competitive results in terms of r (0.5-0.7 for validation SPI at one lead time) and nrmse (0.1-0.2 for validation SPI at one lead time) regardless of regions. As the lead time increased, the accuracy tended to fall, but it showed good results compared to CFS, which is a numerical model data. When compared to a drought case, the SST of 8 months ago was influenced by the results. We confirmed the consistency of existing studies through heat map.

Chapter 2

2. Short-Term Forecasting of Satellite-Based Drought Indices Using Their Temporal Patterns and Numerical Model Output

2.1 Introduction

Drought, one of the more extreme natural disasters observed in the world, is caused by complex mechanisms between the land surface, ocean and atmosphere (Bayissa et al., 2019; L. Han et al., 2016; NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters. Available online: https://www.ncdc.noaa.gov/billions/ (accessed on 06 August 2020).; Sheffield et al., 2014). Since drought can not only last for weeks, months or even years but also develop over large spatial extents, it causes considerable problems, such as the decrease of crop yield, shortage of water, desertification, wildfires and dust storms (Tadesse et al., 2017; Tran et al., 2017; Yan et al., 2017). Many studies have reported that ongoing global warming has increased the frequency of severe drought (Dai, 2011; Dai, 2013; Zhang et al., 2019). According to Emergency Events Database (EM-DAT) provided by Centre for Research on the Epidemiology of Disasters (Centre for Research on the Epidemiology of Disasters (CRED) Natural Disasters in 2018. Available online: https://cred.be/sites/default/files/adsr_2018.pdf (access on 06 August 2020)., 2019), there were 33 drought events worldwide between 2008 and 2018, creating an economic loss of $18 billion. For these reasons, drought monitoring and forecasting are essential for appropriately managing drought-related damage and providing relevant drought information to decision-makers (Belayneh et al., 2016; Zhang et al., 2019). Drought forecasting plays a particularly vital role in risk management as a comprehensive preparation and mitigation of potential drought-caused damage in a timely manner (Demisse et al., 2019; Rhee & Im, 2017; Yan et al., 2017).

In the United States, there are several drought forecasting systems, including the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center's (CPC's) Seasonal Drought Outlook (SDO) (Climate Prediction Center (CPC) US Seasonal Drought Outlook (SDO), Available online: https://www.cpc.ncep.noaa.gov/products/outreach/publications.shtml (accessed on 06 August 2020) ; Steinemann, 2006) and the National Integrated Drought Information System's (NIDIS's) Drought Early Warning Systems (DEWS) (https://www.drought.gov). These systems provide monthly and seasonal drought information that integrates United States Drought Monitor (USDM) data, various satellite observations (e.g., products or reflectance of Advanced Very-High-

Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS)), satellite-based drought indices (e.g., Vegetation Drought Response Index (VegDRI)) and climate forecasting models (e.g., Climate Forecast System (CFS)). In South Korea, drought forecasting information can be obtained from the National Drought Information-Analysis Center (www.drought.go.kr), which is operated by the Ministry of Environment and Korea Water Resources Corporation. The center provides present and upcoming drought conditions up to three months using the standardized precipitation index (SPI), the present conditions of water storage, and the available water capacity of the soil and the water level of dams. China officially provides the present drought conditions using ground measurements and satellite data, while they provide only the forecast of each drought factor individually such as precipitation and El Niño-Southern Oscillation (ENSO) from climate models (Weather China, http://products.weather.com.cn). However, in East Asia, there are no drought forecast systems blending in situ and satellite observations with climate model output, unlike in the United States.

There have been considerable research activities looking at drought forecasts using a number of drought-related variables—hydro-meteorological factors (e.g., precipitation, temperature, evapotranspiration, soil moisture and streamflow), drought indices (e.g., SPI, standardized precipitation evapotranspiration index (SPEI), and evaporative stress index (ESI)), and climate indices (e.g., sea surface temperature, and ENSO)—from ground stations, satellites, and reanalysis or numerical models (Fung et al., 2020; Mishra and Singh, 2011). Among them, drought indices have the advantage of being more readily useable for understanding drought conditions than hydro-meteorological factors (Zargar et al., 2011). This is because drought indices can reflect dryness anomalies and agricultural or hydrological impacts through using multiple drought factors. For this reason, drought indices have been used in many drought forecast studies.

In previous studies using drought indices, ground-based drought indices—SPI, SPEI, and Palmer Drought Severity Index (PDSI)—have been frequently used for forecasting drought (Belayneh et al., 2014; Cancelliere et al., 2007; Kim and Valdés, 2003; Lohani et al., 1998; Morid et al., 2007). However, those indices can cause high uncertainties over ungauged areas when calculated using spatial interpolation (Jiao et al., 2019; Rhee et al., 2010). For this reason, some studies have used remote sensing, reanalysis or numerical model data in their drought forecasts. Otkin et al. (Otkin et al., 2015) introduced a new index to forecast drought, named Rapid Change Index (RCI). RCI was developed to show the change in drought conditions using ESI (generated through Atmosphere-Land Exchange Invers (ALEXI) using satellite-based LST), soil moisture (obtained from North American Land Data Assimilation System (NLDAS)), and SPI (obtained from the CPC unified analysis of daily precipitation reports) across the United States. Park and Kim (2019) forecasted an area of severe

agricultural drought using a Landsat 8 and Shuttle Radar Topography Mission (SRTM)-based soil moisture index (SMI) and topographic data. Park et al. (2018) used the time-series of satellite-based drought indices (Scaled Drought Condition Index (SDCI), Microwave Integrated Drought Index (MIDI), and Very Short-term Drought Index (VSDI)) and climate indices (i.e., Madden-Julian Oscillation (MJO)) in order to develop a drought forecasting model for East Asia. Lorenz et al. (2018) conducted drought forecasting studies using USDM data, the anomaly of drought factors (i.e., precipitation, evapotranspiration and soil moisture), and forecasted drought factors from the Climate Forecasting System (CFS) focusing only on drought intensification with two-, four- and six-week time periods.

Although there has been an effort to forecast drought by integrating multiple source data, drought forecasts are still challenging. This is due to the inherent complexity of drought and the spatio-temporal variability of drought-related variables associated with the global hydrologic cycle (Mishra and Singh, 2011). Various methods have been applied to the development of drought forecasting models, e.g., regression (Li et al., 2016; Lorenz et al., 2017; Meng et al., 2017; Otkin et al., 2014), autoregressive integrated moving average (ARIMA) (Han et al., 2010; Han et al., 2012), and machine learning (Borji et al., 2016; Özger et al., 2012; Park and Kim, 2019; Park et al., 2018) models. However, most models have limited performance, especially with short-term (within two weeks) forecasting of drought showing lower predictive skills than long-term (e.g., seasonal) forecasts (Park et al., 2018). This is due to the complexity of the hydrological process related to droughts on the short-term scale (e.g., the relationship between soil moisture, temperature, and evapotranspiration after precipitation events).

Another reason for the inaccuracy of short-term forecasting is that most drought forecast studies on a short-term scale have only considered the time-series pattern of drought (Abebe and Foerch, 2008; Han et al., 2010; Modarres, 2007; Otkin et al., 2015). Only a few papers have used the numerical forecast model data and climatic indices to reflect the impact of upcoming weather conditions. Lorenz et al. (Lorenz et al., 2018) proved that the use of numerical predictors in the short-term (two weeks in this study) drought forecasting through logistic regression could increase predictive skills when the drought intensified. Park et al. (Park et al., 2018) combined the time-series of satellite-derived drought indices and climate indices for drought forecasting at a pentad scale. Although the use of climate indices improved the drought forecasting skills, the model performance was saturated with $r \sim 0.7$ regardless of the drought indices used. Therefore, it is necessary to improve drought forecasting skills at the short-term scale considering the alleviation and intensification of drought by combining forecasted climate factors (e.g., precipitation and temperature).

In this research, we aimed to propose a drought forecasting model on a short-term scale through the integration of numerical model outputs, topographic characteristics (i.e., climate zone, digital elevation model (DEM), and landcover), and satellite-based drought indices (i.e., Scaled Drought Condition Index (SDCI) and SPI) using convolutional long short term memory (ConvLSTM) and random forest (RF) approaches. The assumption of this study is that drought indices (SDCI and SPI for this study) perfectly represent drought. Therefore, if the drought indices are well forecasted, then the drought is also well predicted. The measurable objectives of this study were to 1) forecast drought on a short-term scale (8 days) over a part of East Asia and to 2) analyze the system's forecasting skills when only considering temporal patterns of drought conditions and when combining numerical model outputs, topographic information, and temporal patterns of drought conditions in the short-term forecasting of drought. The reasons for the 8-day time scale were 1) because satellite products are provided every eight days, and 2) because drought is not a rapidly changing (e.g., daily) phenomenon. The novelty of this study can be summarized in two aspects: 1) ConvLSTM was used to develop the drought forecasting model, which has the advantage of being able to predict time-series data by considering spatial characteristics. Because droughts have both spatial and temporal patterns, ConvLSTM can be useful in forecasting drought because it learns time-series and spatial patterns simultaneously. To our knowledge, this is the first attempt to use ConvLSTM in a drought forecasting model. 2) The proposed drought forecasting model considers not only drought patterns but also predicted climatological factors (i.e., precipitation, temperature) in both drought intensifications and alleviations, while the recent literature has only focused on drought intensification (Lorenz et al., 2018).

2.2 Study area and data

2.2.1 Study area

The study area is a part of East Asia (latitude: 25.17◦ N–45.72◦ N; longitude: 114.05◦ E–133.25◦ E), including east China, southeast Russia, Korea, and part of Japan (Figure 2.1). According to the EM-DAT (https://www.emdat.be), the study area suffered from severe drought 13 times between 2000 and 2018 with economic losses totaling over $20 million. The study area consists of diverse landcover types (i.e., water, forest, cropland, built-up, grassland, and savannas; Figure 2.1a) and climate zones (i.e., snow, warm temperature, and arid climates; Figure 2.1c). Cropland is located at low altitudes while grasslands and forests are located at relatively high altitudes. In terms of climatic characteristics (Figure 2.1c), the study area is generally hot and humid in summer caused by the East Asia monsoon (Park et al., 2019), while it is dry and cold in winter. Southern and central-eastern China, Japan and Southern and central-eastern China, Japan and the coastlines of South Korea have warm temperate climates, while north-east China, North Korea and inland South Korea have snow climates.

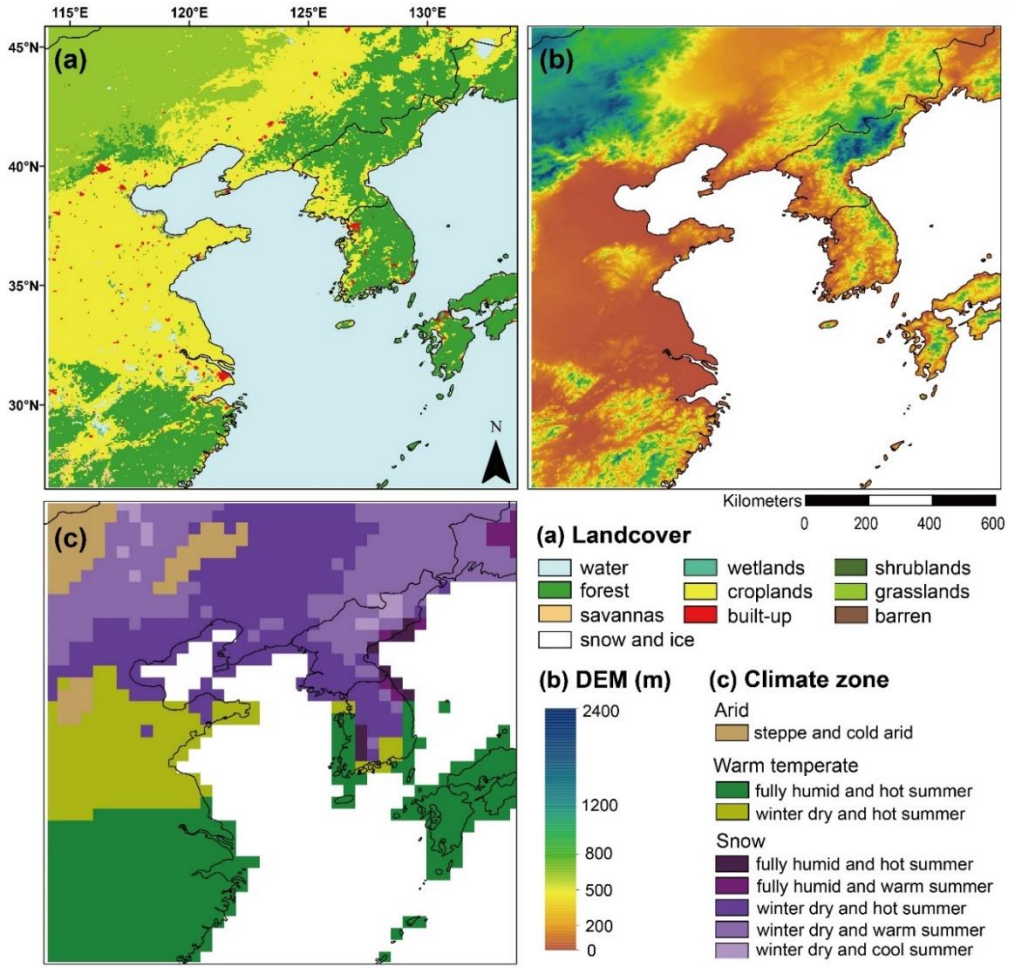

Figure 2.1 The study area of this research with (a) Moderate Resolution Imaging Spectroradiometer (MODIS) landcover, (b) Shuttle Radar Topography Mission (SRTM) digital elevation model (DEM), and (c) climate classification map (Kottek et al. (Kottek et al., 2006)

2.2.2 Data

Our drought forecasting model was developed using 1) satellite-based drought indices for documenting the historical patterns and current conditions of drought, 2) numerical forecasting model outputs for considering upcoming weather phenomena, and 3) the spatially distributed geographic characteristics of the study area.

satellite-based drought indices

In this study, we selected two satellite-based drought indices, SDCI and SPI, for drought forecasting over the study area. The non-vegetated area (i.e., urban area) was extracted when calculating both drought indices because drought indices are valid only over vegetated areas. The irrigation effect was not considered as the irrigation information of the study areas was not provided in detail on the spatial domain. The reasons why SDCI and SPI were selected included that 1) these

two indices have been proved useful in drought monitoring in previous studies and Chinese and South Korean drought management (National Bureau of Statistics of China (http://data.stats.gov.cn); National Drought Information Portal (http://www.drought.go.kr/english/), (Han et al., 2020; Park et al., 2018)), and 2) they are calculated using different drought factors. SDCI is designed to incorporate multiple drought factors (i.e., vegetation health, temperature, and precipitation), while SPI uses a single factor (i.e., precipitation). Thus, they have somewhat different spatial and temporal patterns, which can affect forecasting skills when using the same upcoming weather conditions. Note that this study aimed to forecast each drought index, not to compare two indices as we assumed the indices perfectly represent drought.

SDCI (Rhee et al., 2010) is designed to be not only applicable to both arid and humid regions but also flexible in terms of the multiple timescales of precipitation. It combines thermal stress (Temperature Condition Index (TCI), (Kogan, 1995a)), water stress (Precipitation Condition Index, PCI) and vegetation stress (Vegetation Condition Index, VCI, (Kogan, 1995b)), which can be used as drought indicators:

$$TCI = \frac{\mathrm{LST}_{\max} - \mathrm{LST}}{\mathrm{LST}_{\max} + \mathrm{LST}_{\min}} \quad (1)$$

$$PCI = \frac{\mathrm{precipitation} - \mathrm{precipitation}_{\min}}{\mathrm{precipitation}_{\max} + \mathrm{precipitation}_{\min}} \quad (2)$$

$$VCI = \frac{\mathrm{NDVI} - \mathrm{NDVI}_{\min}}{\mathrm{NDVI}_{\max} + \mathrm{NDVI}_{\min}} \quad (3)$$

$$SDCI = 0.25 \times TCI + 0.5 \times PCI + 0.25 \times VCI \quad (4)$$

Similar to USDM, SDCI has six categories: no drought, abnormally dry, moderate drought, severe drought, extreme drought, and exceptional drought (Table 2.1, Rhee et al., 2010). To obtain factors for SDCI, we used Terra MODIS products (i.e., MOD 13C1 NDVI and MOD11C2 LST, EARTHDATA (https://search.earthdata.nasa.gov) for VCI and TCI, and Tropical Rainfall Measuring Mission (TRMM) daily precipitation 3B42 (https://disc.gsfc.nasa.gov) for PCI (Table 2.1). PCIs with 1-, 3-, 6-, 9-month time scales were calculated at 0.05 degree for co-locating with MODIS products.

Table 2.1 Drought categories based on Scaled Drought Condition Index (SDCI, Rhee et al. (Rhee et al., 2010)) and Standardized Precipitation Index (SPI, McKee et al. (McKee et al., 1993)).

SDCI value (unitless)	SDCI category	SPI value (unitless)	SPI category
0.0 to < 0.1	Exceptional Drought	-2.0 and less	Extreme Drought
0.1 to < 0.2	Extreme Drought	-1.99 to < -1.5	Severe Drought
0.2 to < 0.3	Severe Drought	-1.5 to < -1.0	Moderate Drought
0.3 to < 0.4	Moderate Drought	-1.0 to < 0	Mild Drought
0.4 to < 0.5	Abnormally Dry	0 or more	No Drought
0.5 to <= 1	No Drought		

Table 2.2 Summary of data used to develop drought forecasting model in this study. The entire period of this study is from 2003 to 2018 and the specific period of each variable is given in the table below.

Variables (collected period)		**Products**	**Spatial resolution**	**Temporal resolution**
Drought indices (from 2003 to 2018)	**SDCI**	Terra MODIS Normalized Difference Vegetation Index (NDVI, MOD13C1)	0.05º	16 days
		Terra MODIS Land Surface Temperature (LST, MOD11C2)		8 days
	SPI	TRMM precipitation (3B42)	0.25º	daily
Numerical model (from 2015 to 2018)		GFS air temperature	0.5º	3 hours (to 240 hours)
		GFS precipitation		
Static data		Terra MODIS landcover (MCD12C1)	0.05º	yearly
		SRTM digital elevation model (DEM)	90 m	-
		Climate zone (Kottek et al. (Kottek et al., 2006))	0.5º	

SPI is one of the ground-based drought indices, which is based on precipitation deficit or surplus with multiple time scales of accumulation (McKee et al., 1993). It is calculated by a gamma probability density function using more than 30 years of precipitation data and categorized by eight classes (Table 2.1). It has the advantage of determining drought types, such as meteorological and agricultural drought, based on the time scales of accumulation (Livada & Assimakopoulos, 2007; Zhu et al., 2019). However, ground measurement-based SPI does not provide spatially continuous drought information. In this study, similar to PCI, TRMM daily precipitation (3B42) was used for calculating SPI, which was applied to accumulated precipitation, 3-month (SPI3). Although SPI requires a long period (over 30 years) of data, some research has proved the capability of TRMM-based SPI for drought monitoring (De Jesús et al., 2016; Tan et al., 2017; G. Yan et al., 2018; Zhu et al., 2019). SPI was produced in 0.05 degree resolution using bilinear-resampled TRMM precipitation through SPI function (T. Lee) in Matlab 2019a.

Numerical model outputs

Global Forecasts System (GFS) is a widely used weather forecast model which blends atmosphere, ocean, land/soil, and sea ice models. It provides atmospheric (e.g., air temperature, precipitation) and land-soil variables (e.g., soil moisture) for 240-hour forecasts, four times a day, with 0.5° spatial resolution for the National Centers for Environmental Prediction (NCEP, https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/global-forcast-system-gfs) (Table 2.2). In this study, we used averaged air temperature and accumulated precipitation predicted for eight days. The reason for using upcoming weather conditions is that the current drought status may dramatically change depending on upcoming weather conditions, which cannot be expected only using the historical patterns of drought indices. To minimize the differences in temperature and precipitation data between satellite products and numerical model outputs, linear fitting was conducted between the two (i.e., GFS air temperature and precipitation were converted to LST and TRMM precipitation, respectively). The mean (max and min) slopes and intercepts were 1.28 (4.5 and 0.07) and 9.62 (50 and -6) for precipitation and 1.1 (1.33 and 0.76) and 6 (71 and -91) for temperature, respectively. Although air temperature is not always highly correlated with LST, many studies have proved a relatively strong positive relationship between them (Benali et al., 2012; Vogt et al., 1997). Fitted LST and precipitation were normalized to TCI (TCIGFS) and PCI (PCIGFS) using the minimum and maximum values when calculating satellite-based TCI and PCI.

Static data

In this study, landcover, elevation and climate zones were used as additional predictors, considering the environmental and topographic characteristics of the study area. MODIS landcover (MCD12C1, Majority land cover type 1) was used after simplifying classes through reclassification (Figure 1a). For elevation, 90m SRTM DEM was obtained from the United States Geological Survey (USGS) elevation products site (http://eros.usgs.gov/elevation-products). It was resampled to 5km using the mean aggregation for co-locating with MODIS products (Figure 2.1b).

Köppen-Geiger climate classification maps have been frequently used by researchers across a wide range of disciplines for the climatic regionalization of environmental variables (Kottek et al., 2006; Park et al., 2019). The map was developed through rule-based methods using 50 years (from 1950 to 2000) of temperature and precipitation from reanalysis data. The map has 31 classes (e.g., equatorial, arid, warm temperate, snow, and polar) in 0.5 degrees (http://koeppen-geiger.vu-wien.ac.at/present.htm). The study region was cropped from the map produced by Kottek et al. (2006) and converted to 0.05 degrees so that it matches MODIS datasets using the nearest neighbor method (Figure 2.1c).

2.3 Methodology

Figure 2.2 shows the process flow diagram of this study. The proposed approach is divided into two steps: step 1 uses ConvLSTM to obtain temporal patterns from historical drought conditions (i.e., SPI and SDCI) and step 2 uses RF to feed static variables (i.e., landcover, elevation and climate zone) and forecasted climate factors (i.e., temperature (TCIGFS) and precipitation (PCIGFS)) provided from the numerical model into the output of step 1. There are two reasons why the final drought forecasting model combines two machine learning approaches. First, the model structure becomes complex when all independent variables are used as input variables for ConvLSTM. This requires much more memory and processing time than statistical and basic machine learning (e.g., support vector regression or RF) approaches (He and Sun, 2015). Therefore, the temporal patterns of each drought index and spatial information were used in ConvLSTM and RF, respectively, to enable drought forecasting even in a memory limited environment. The second reason is that ConvLSTM is a model optimized for analyzing temporal patterns such as the drought indices used as predictors in a spatial context in this study. Numerical model outputs and static data were used to help improve the model's forecasting skills as they can provide information which was not included in the temporal patterns of the drought indices (i.e., SDCI and SPI).

During the entire study period, the ConvLSTM model was produced by real-time learning (step 1, 2003-2018 (2003-2014 for optimizing parameters)) to reflect the most recent drought condition. The RF model was applied to produce the final drought forecasts using data from 2015 to 2017, using the output of ConLSTM, static data, and numerical model outputs. The model performance was evaluated for 2018. The reason why study periods were divided into 2003-2018 and 2015-2017 is to reserve enough data for the tuning parameters of ConvLSTM and to obtain the number of reasonable samples required for the models (i.e., ConvLSTM and RF) considering the annual phenology of drought factors (i.e., temperature, vegetation, and precipitation). After obtaining drought indices (i.e., SDCI and SPI) as an 8-day interval, the ConvLSTM model for each drought index (SDCI model and ConvLSTM-SPI model for step 1) was generated by training the previous one-month (8-day x 4) of drought conditions in step 1. In step 2, the final drought forecasting model for each drought index (SDCI model and SPI model for step 2) was developed using forecasted weather conditions, climate zone, landcover, DEM, and outputs of step 1 through RF, which reduces errors caused by training only the temporal patterns of the drought indices.

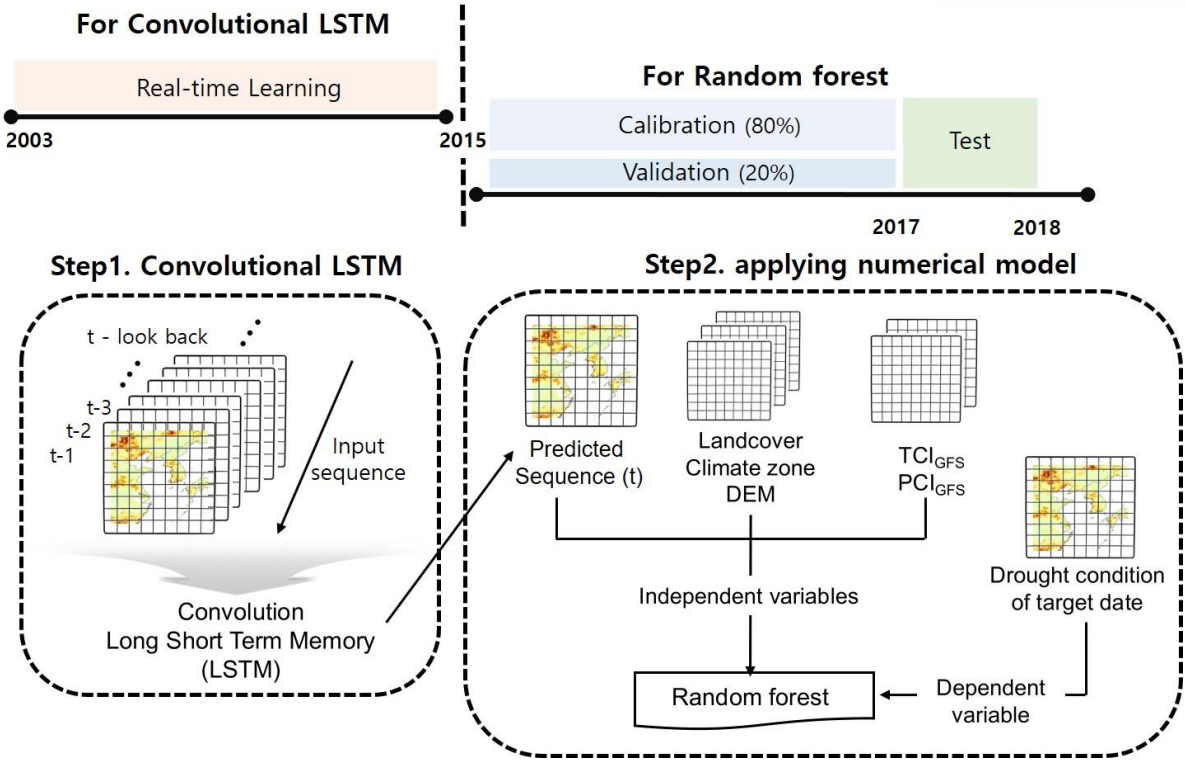

Figure 2.2 The process flow diagram of this study. Step1 used convolutional long short term memory (ConvLSTM) from 2003 to 2018 by real-time learning. Step 2 used random forest (RF) from 2015 to 2017). The test of the final drought forecasting model was conducted for 2018.

2.3.1. Step 1 : Convolutional long short term memory (ConvLSTM)

Shi et al. (2015) developed ConvLSTM (Figure 2.3) which combines convolutional neural networks (CNN) and long short term memory (LSTM). The two algorithms have been widely used in image classification and time-series forecasting, respectively. ConvLSTM has recently been applied to various research tasks that need to consider both time-series patterns and spatial information, such as segmentation, change detection, forecasting video frames, forecasting sea surface temperature, and air pollution research (Ghimire et al., 2019; Mateo-García et al., 2019; Mu et al., 2019; Petrou and Tian, 2019; Song et al., 2018; Wen et al., 2019). ConvLSTM models space-time structures through encoding spatial information, which can overcome the major limitations of LSTM, namely the loss of spatial information (Ma et al., 2019). The structure of ConvLSTM is similar to that of LSTM, which consists of memory cells and three gates (i.e., forget, input, and output gates). The three gates play roles in maintaining or discarding memory (temporal information) (Cruz and Bernardino, 2019). The main difference between LSTM and ConvLSTM is that the internal matrix multiplications of LSTM are replaced with convolution operations in ConvLSTM. As a result,

ConvLSTM can produce 2-D output while the result of LSTM is only a 1-D vector. The main equations are as follows (Shi et al., 2015):

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \circ C_{t-1} + b_i) \quad (5)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \circ C_{t-1} + b_f) \quad (6)$$

$$C_t = f_t \circ C_{t-1} + i_t \circ \tanh(W_{xc} * x_t + W_{hc} * H_{t-1} + b_c) \quad (7)$$

$$O_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \circ C_t + b_0) \quad (8)$$

$$H_t = O_t \circ \tanh(C_t) \quad (9)$$

i_t and O_t are input and output gates. C_t is the memory cell that has accumulated the state information using W (weight). When new input is entered, i_t is activated and C_t accumulates the state information. The past status can be forgotten when f_t is on. O_t can be obtained through the final state H_t, which is propagated by C_t (Figure 3). A more detailed explanation of ConvLSTM can be found in (Shi et al., 2015).

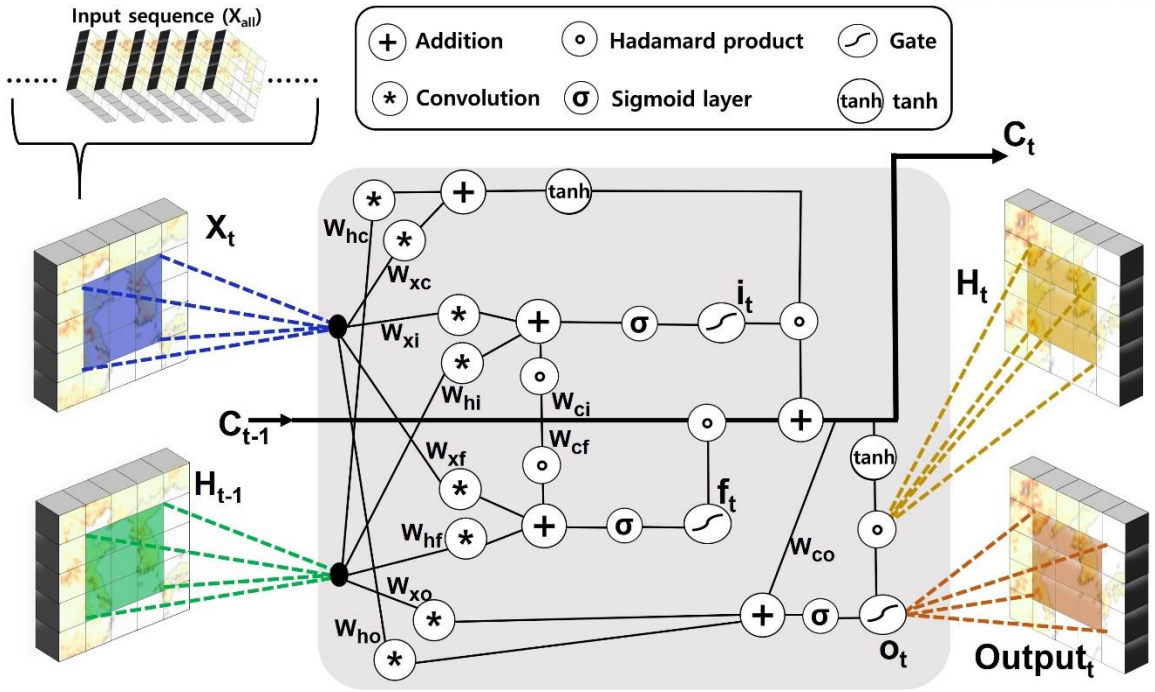

Figure 2.3 The structure of convolution long short term memory (ConvLSTM) model used in this research. Three layers used in the ConvLSTM model have the same structure shown in a gray shading box. X, H, C, i, f, and o are input sequence, hidden state, memory cell, input gate, forget gate and output gate, respectively.

Recently, Mu et al. (Mu et al., 2019) proposed a ConvLSTM-Rolling Mechanism (ConvLSTM-RM, named "real-time learning" in the present study), which utilizes the most recent data (named "stride period" in this study) to develop a forecasting model. The RM method helped improve forecasting performance (Akay and Atak, 2007; Mu et al., 2019; Zhao et al., 2012). Therefore, in this study, four temporally consecutive data with an 8-day interval that stride for the recent three years were applied to forecast drought conditions in the next 8-day interval through real-time learning from 2003 to 2017. For example, to forecast drought conditions on 9 January 2018, the ConvLSTM model is updated through striding the four consecutive data with the 8-day interval from 9 January 2015 to 1 January 2018 (for three years). The stride period of three years was determined considering the computational efficiency and the impact of annual phenology. The "convlstm2d" function provided by TensorFlow Core r2.0 (https://www.tensorflow.org/api_docs/python/tf/keras/layers/ConvLSTM2D) was used in this study. After testing various combinations of parameters using data from 2003 to 2014, the ConvLSTM structure was determined to have three layers with 16, 16, and 1 filters of 3x3, 3x3, and 1x1 size (Figure 2.3). Mean square error (MSE) was used as a loss function, and the model was optimized by

the adaptive moment estimation (Adam), which is a popular optimization algorithm in neural networks (Kingma & Ba, 2014). All experiments were run on a computer with Intel(R) Xeon(R) CPU E5-2680 v2 @ 2.80GHz and NVidia Titan Black GPU (6GB of memory). About 24 days were taken to train the ConvLSTM model in step 1 with 50 epochs and 2 batch size.

2.3.2. Step 2 : Random forest (RF)

RF is an ensemble approach based on classification and regression trees (CART), which overcomes the major limitations of CART, such as its sensitivity to training data configuration and the overfitting problem, by aggregating multiple independent trees (Breiman, 2001). RF consists of a variety of CARTs ("decision trees") that have the same probability distributions through bootstrapping-based randomization approaches. All decision trees are aggregated using (weighted) majority voting for classification and (weighted) averaging for regression. RF provides the relative importance of independent variables, which has been widely used in previous studies, even though it is of local importance, not global (Cho et al., 2020; McLaren et al., 2019; Park et al., 2017; Park et al., 2019; Rhee and Im, 2017). It can be obtained as a percentage of the increased MSE for each variable using out-of-bag (OOB) data. Having a variable with a relatively large percentage of increased MSE means that it made a relatively more significant contribution to our model than to other variables.

In this study, upcoming weather conditions and terrestrial information were used as additional inputs to RF to further improve forecasting skills. The predicted drought conditions from ConvLSTM (output from step 1), static variables (i.e., landcover, climate zone, and DEM) and normalized climate factors (i.e., TCI_{GFS} and PCI_{GFS}) were used as independent variables, while the drought conditions in the next eight days were used as a dependent variable in step 2 (Figure 2.2). From 2015 to 2017, the calibration and validation data set were divided into 80% and 20% respectively after randomly extracting samples to ensure that landcover types, climate zones and all ranges of drought index values (e.g., 0-1 for SDCI) were uniformly included. The RF package of R statistic software (http://www.r-project.org) was used with default settings except for the number of trees (500).

2.3.3. Accuracy assessment

In order to evaluate the performances of the forecasting model for each drought index (i.e., SDCI and SPI), three statistical metrics were used: correlation coefficient (r), normalized root mean square error (nRMSE), and mean absolute scaled error (MASE; Hyndman and Koehler (Hyndman and Koehler, 2006)).

$$r = \frac{n(\sum y\hat{y}) - (\sum y)(\sum \hat{y})}{\sqrt{[n\sum y^2 - (\sum y)^2][n\sum \hat{y}^2 - (\sum \hat{y})^2]}} \quad (10)$$

$$RMSE = \sqrt{\frac{\sum(y-\hat{y})^2}{n}},\ nRMSE = \frac{RMSE}{y_{max}-y_{min}} \quad (11)$$

$$MASE = \frac{1}{m}\sum\frac{|y_t-\hat{y}_t|}{(\sum|y_t-y_{t-1}|)/(n-1)} \quad (12)$$

n and m are the number of samples and y and $\hat{y}$ are the values of reference and predicted drought indices, respectively. The forecasting skills are useful when r (nRMSE) is closer to 1 (0). MASE is one of the statistical indices used to evaluate the time-series forecasting model, which has an advantage when there are very different scales (i.e., negative, positive and zero) (Hyndman and Koehler, 2006). MASE is less than 1 (MASE <1) when the forecasting error is better than the averaged variation of the time-series data.

2.4 Results and Discussion

2.4.1. The performance of drought forecasting model

Figure 4 depicts the model performance of the SDCI model for step 1 (ConvLSTM, Figure 2.4a) and step 2 (RF, Figures 2.4b-c). Based on the period of the three-year stride, the averaged r and nRMSE are presented from 2006 to 2017 (Figure 2.4a). The r and nRMSE values ranged from 0.23 to 0.92 (mean of 0.73) and from 0.07 to 0.2 (mean of 0.12), respectively. Despite the generally high accuracy, some dates had low r and high nRMSE due to sudden changes in drought conditions based on SDCI. For example, on 7 April 2006 (Day-Of-Year (DOY) 2006097) in Northeast China (Liaoning and Jilin provinces), drought conditions rapidly changed in eight days from abnormally dry to extreme drought, which was caused by little precipitation and unusually high temperatures. The region remained in normal drought conditions for about two months (DOY: 2006033-2006089), making it difficult to forecast serious drought outbreaks using temporal patterns. In other words, drought forecasting through ConvLSTM cannot reflect the impact of upcoming precipitation and temperature because ConvLSTM only considers historical patterns. Some previous studies that considered historical patterns were similarly vulnerable to sudden weather changes (Belayneh et al., 2014; Park et al., 2018).

The r, nRMSE, and slope of RF calibration and validation are 0.98 and 0.90, 0.05 and 0.11, and 0.89 and 0.78, respectively (Figure 2.4b-c). In terms of r and nRMSE, the forecasting skill was improved when integrating upcoming weather and spatial information with ConvLSTM outputs. There was a tendency for an overestimation in exceptional and extreme drought conditions (low values of SDCI) and underestimation in no drought conditions (high values of SDCI). This was probably because 1) the samples in extreme values (0-0.1 and 0.9-1) were smaller (4% (for 0-0.1) and 5% (for 0.9-1) than the other samples (Park et al., 2019), 2) random forest produces trees with reducing errors, which leads to the values trending around the mean value, especially when there are not many extreme samples (Im et al., 2016; Park et al., 2016), and 3) the numerical model outputs did not reflect weather well in some regions and on some dates due to low correlation with the satellite products (i.e., GPM precipitation and MODIS LST) . However, compared to the validation results on Park et al. (2018), who developed RF-based drought forecasting models with the pentad interval using SDCI and climate index (MJO), our model showed larger dynamic ranges in the output, possibly due to the incorporation of numerical model outputs. Lorenz et al. (Lorenz et al., 2018) also demonstrated that drought forecasting models using weather forecasts (when the drought intensified) performed better than those that only used the past and present drought conditions. The accuracy of the model without using upcoming weather conditions decreased, resulting in r of 0.85 and nRMSE of 0.13 for

validation, not shown here), which supports the incorporation of upcoming weather conditions into the model.

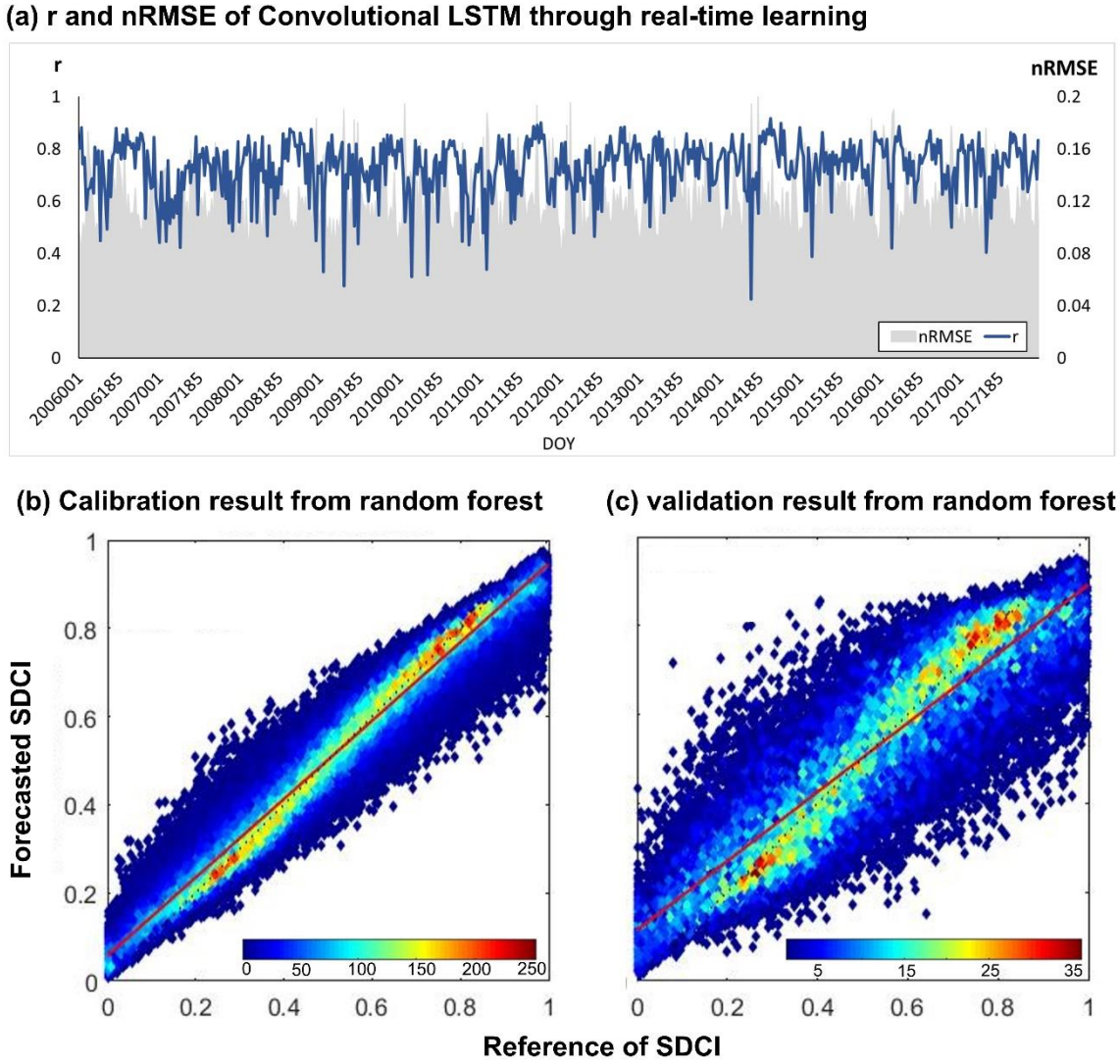

Figure 2.4 Results of step1 and step2 using SDCI. (a) r (blue line) and RMSE (gray shading) through the real-time training of convolutional long short term memory (ConvLSTM) from 2006 to 2017, (b) calibration results from random forest (RF) and (c) validation results from random forest (RF)

Figure 2.5 shows the performance of the SPI model through step 1 (ConvLSTM, Figure 2.5a) and step 2 (RF, Figures 2.5b and 2.5c). The r and nRMSEs of the SPI ConvLSTM model ranged from 0.22 to 0.96 (mean of 0.84) and from 0.04 to 0.2 (mean of 0.09), respectively (Figure 2.5a). Performance varied between SPI- and SDCI- based ConvLSTM model in terms of r and nRMSE (Figure 2.4). This is probably because of the temporal variabilities according to type and number of

drought factors when calculating drought indices. There are three cases: (1) If there is a drastic change in precipitation only, SPI rapidly changes dry to wet or wet to dry, unlike SDCI that can have lower temporal variability caused by other factors (i.e., temperature or vegetation condition). (2) If there is an abrupt change in temperature, SDCI undergoes drastic changes. However, such a temperature change does not impact SPI because it is a function of precipitation only. (3) If both precipitation and temperature change dramatically, the result depends on the intensity of the change in each factor.

In the RF model, the r, nRMSE, and slope of RF calibration and validation were 0.98 and 0.93, 0.05 and 0.08, and 0.90 and 0.84, respectively (Figure 2.5b-c). In terms of r and nRMSE, both RF-based SDCI and SPI models showed better results when using both temporal patterns and numerical model outputs compared to when only temporal data was used.

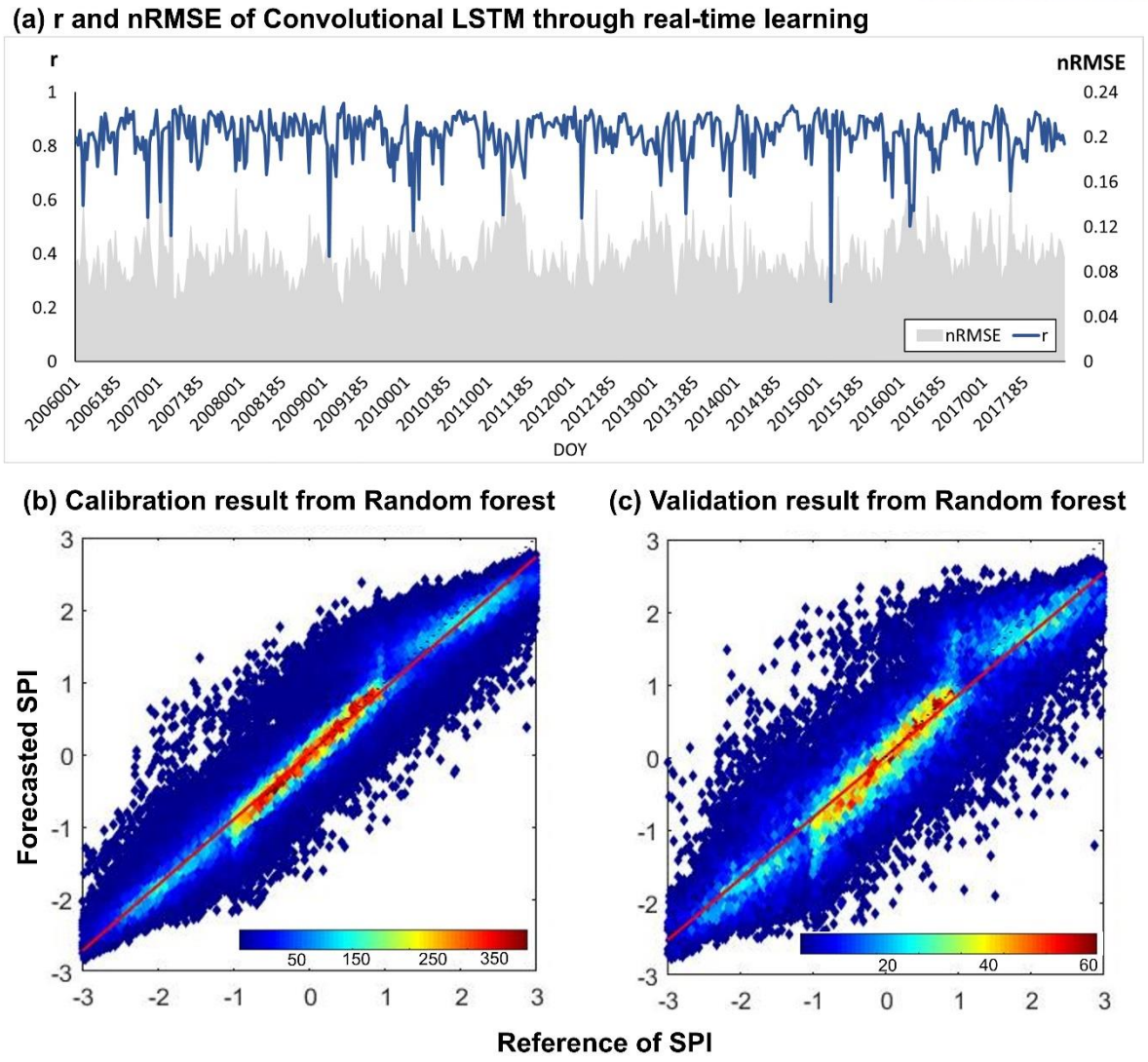

Figure 2.5 Results of step 1 and step 2 for SPI. (a) r (blue line) and RMSE (gray shading) through the real-time training of convolutional long short term memory (ConvLSTM) from 2006 to 2017, (b) calibration results from random forest (RF), and (c) validation results from random forest (RF)

Figure 2.6 describes the spatial and time-series distribution of the performance metrics of the SDCI and SPI model in 2018 (validation year). High correlation values (mean r of 0.62 and 0.77) for the SDCI and SPI model are shown in most areas in Figures 2.6a and 2.6d, respectively. There are common distributions in the SDCI and SPI models (see r graph in middle of Figures 2.6a and 2.6d). Some regions (e.g., Shandong and Jiangxi province) and dates (e.g., DOY 2018193 and 2018337) showed lower correlations than other regions and dates for the SDCI model, which indicates that the numerical model did not have much impact on the improvement of the forecasting skills through RF in those specific cases. This is because there is a significant difference between upcoming

precipitation and temperature from the numerical model and the precipitation and temperature in SDCI that were produced by satellite products. Besides, upcoming vegetation stress was not considered as an independent variable in the RF model due to its absence in the numerical model, which could cause a decrease in correlations. The spatial distributions and time-series distributions of nRMSE (Figures 2.6a and 2.6e) also show generally low values (means of 0.19 and 0.17 for SDCI and SPI, respectively), which have similar patterns to r.

In Figures 2.6c and 6f, MASE shows that the forecasting ability is good (MASE<1) when the forecasting error is less than the average time-series variation of the drought indices. Unlike the other metrics (i.e., r and nRMSE), the MASE maps and graphs have different distributions, which were caused by considering the time-series variation of each index. While most SDCI MASE values were less than 1 (mean of 0.93), SPI MASE were not (mean of 1.15). The fact that the MASE value is over 1 indicates that the error between forecasts and actual values is greater than the average of the time-series variation at each pixel. There are two reasons why SPI has relatively high MASE. One is a limitation of the MASE metric, which is sensitive to outliers (Davydenko and Fildes, 2016). Since SPI values fluctuate depending on precipitation only, the variation of the SPI is generally greater than that of SDCI. The other is that drought forecasting skills decrease when drought conditions change rapidly. Therefore, the MASE of the SPI can be higher than SDCI even if the average of the time-series variation in the two drought indices is similar.

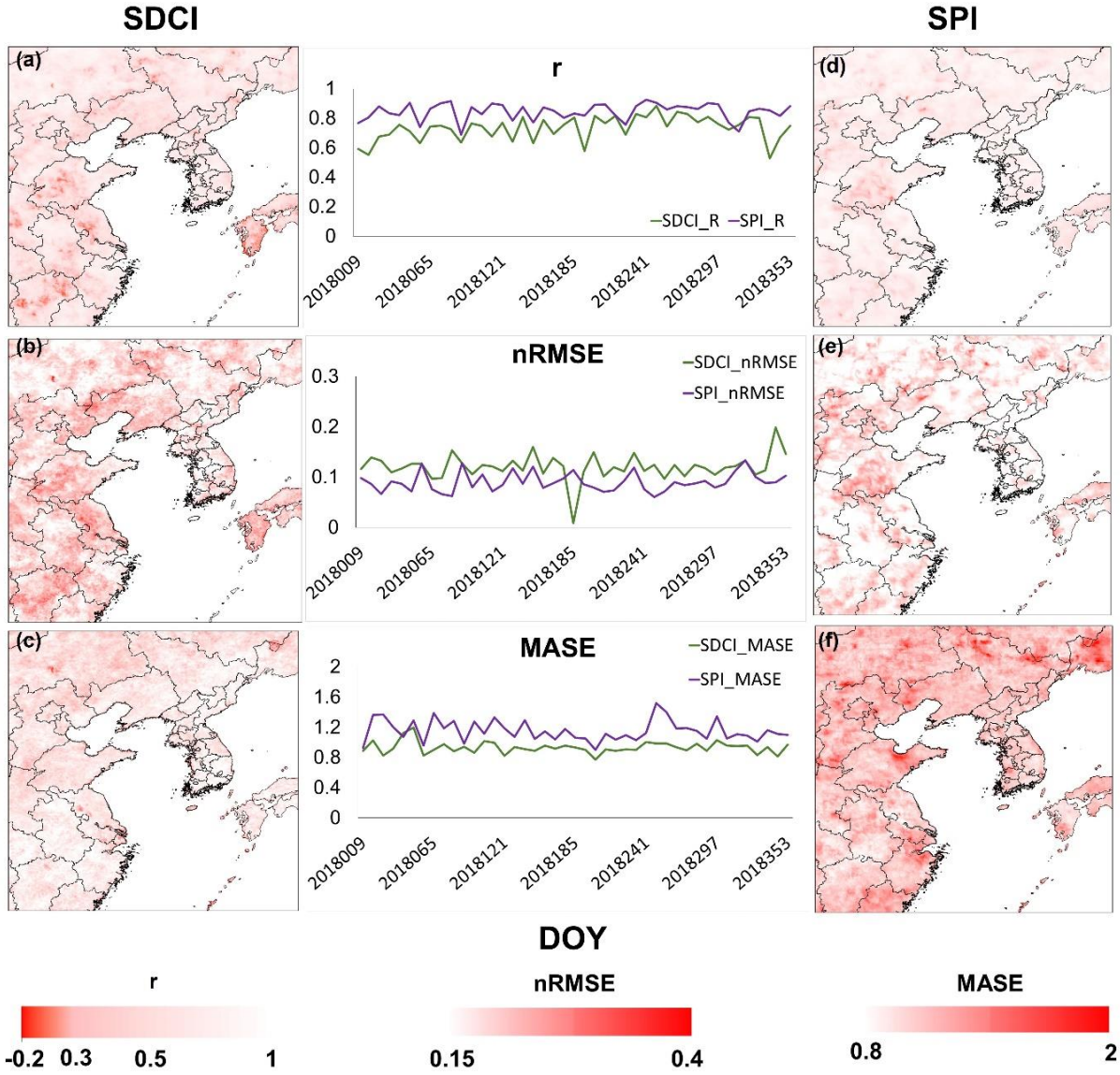

Figure 2.6 Spatial distribution and time-series patterns of r, nRMSE and MASE from random forest model results for SDCI and SPI. The vivid red in the six maps indicates areas of relatively high errors (a-f). The green and purple in three time-series graphs indicate the time-series patterns of SDCI and SPI, respectively.

2.4.2. The spatial distribution of the drought forecasting model

Figure 2.7 depicts spatial distributions of reference SDCI (SDCI_o), outputs of forecasting model (i.e., ConvLSTM (SDCI_s1) and RF (SDCI_s2)), and forecasted climate factors (i.e., precipitation and temperature) from 1 May to 2 June in 2018. The spatial patterns of drought and no-drought conditions were well detected regardless of outputs of SDCI_s1 or SDCI_s2 (e.g., alleviation in the north-western region (i.e., Inner Mongolia) from 1 to 17 May and intensification in Korean

Peninsula from 17 May to 2 June). This indicates that the spatial distribution of drought can be forecasted only using historical patterns of drought when droughts gradually intensify or alleviate (Choi et al., 2017). However, in the case of sudden droughts, such as in the central- and north-western regions (i.e., Inner Mongolia and Henan Provinces) on DOY 2018045, the forecasting skills were limited, especially in terms of drought intensity. According to the relative importance of the independent variables used in RF, two upcoming weather data have a significant impact on the RF-based drought forecasting model (after the ConvLSTM output). However, on that day, the quality of the TCI_{GFS} and PCI_{GFS} was lower in that 8-day interval, which lowered forecasting skills.

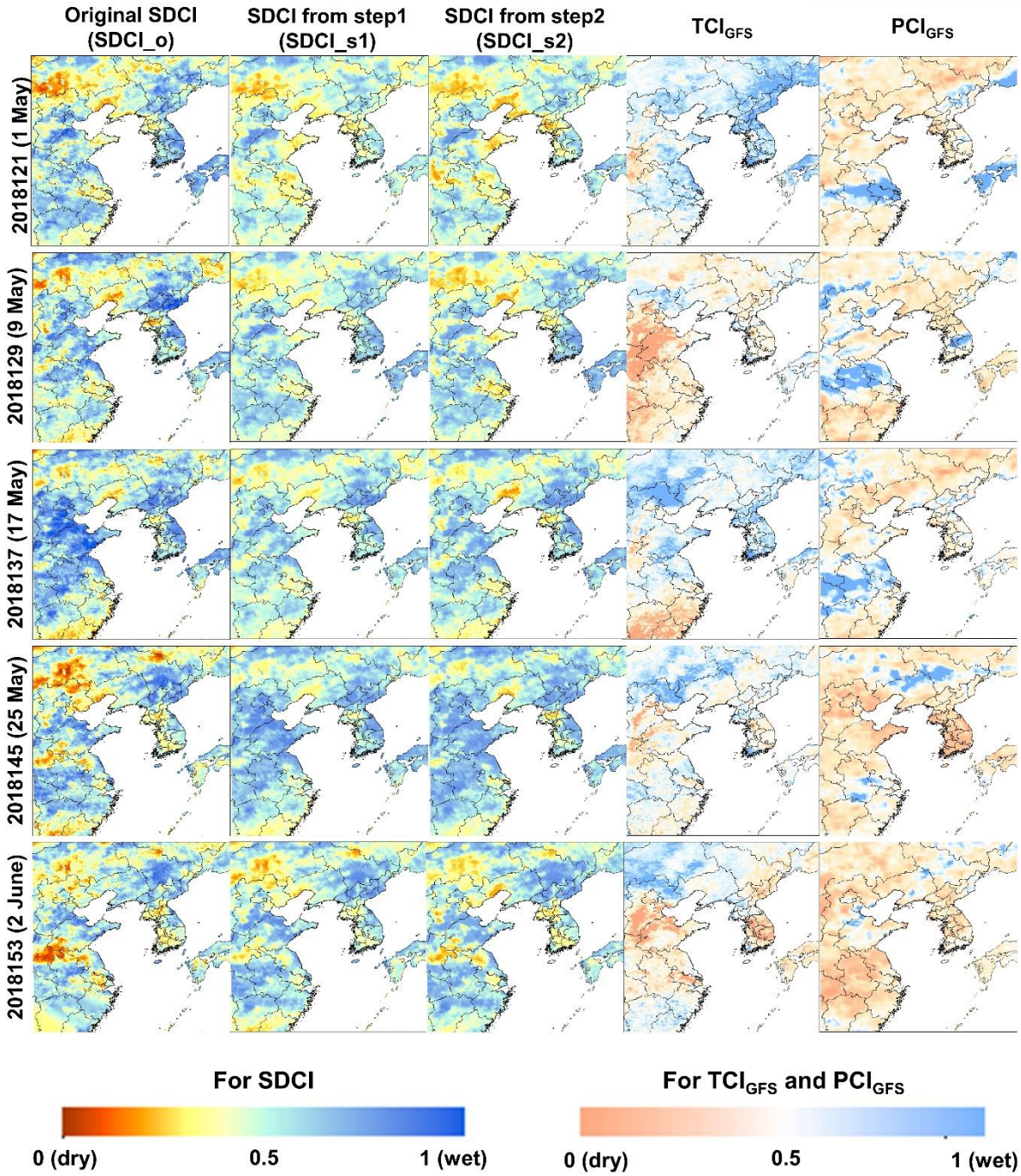

Figure 2.7 Spatial distribution of forecasted SDCI from 1 May to 2 June. The vivid red and blue present dry and wet conditions, respectively, in SDCI maps (i.e., reference SDCI (SDCI_o), SDCI from step 1 (SDCI_s1) and SDCI from step 2 (SDCI_s2). The light red and blue present dry and wet conditions, respectively, caused by temperature (TCIGFS) and precipitation (PCIGFS) from Global Forecasts System (GFS).

SDCI_s1 was slightly underestimated or overestimated when droughts alleviated or intensified, respectively, whereas SDCI_s2 reduced the differences between SDCI_s1 and SDCI_o using forecasted weather data (i.e., TCI_{GFS} and PCI_{GFS}). SDCI_s2 improved 51% of the total pixels on average (up to 68%) in terms of the value of |forecasted-actual|. In other words, the GFS data were used to improve the SDCI_s1 through the RF model. For example, the drought on 17 May was alleviated compared to the eight days before in the central-western region (e.g., Shandong and Hebei Provinces). The output of SDCI_s2 is closer to SDCI_o than SDCI_s1 (the error of 60% of the whole pixel decreased on 17 May). When compared to Lorenz et al. (Lorenz et al., 2018), who focused only on drought intensification, our models are applicable for both drought intensification and alleviation.

However, the improvement is relatively low when there are large gaps between the satellite data (i.e., TRMM Precipitation, MODIS LST) and the GFS data (i.e., TCIGFS and PCIGFS). For example, on 9 May, SDCI_s2 improved only about 45% of the pixels due to an underestimation of GFS precipitation in Shandong province. Kumar et al. (Kumar et al., 2016) compared GFS and TRMM in terms of monthly precipitation and pointed out that GFS precipitation was overestimated in south and northeast China and underestimated in central-eastern China. Another reason for the discrepancy between SDCI_o and SDCI_s2 is the degradation of forecasting skills in the numerical model as lead time increases (Park et al., 2017).

The outputs from our forecasting model for SPI (i.e., ConvLSTM (SPI_s1) and RF (SPI_s2)) are depicted in Figure 2.8 with reference SPI (SPI_o) and forecasted precipitation (PCIGFS). SPI and SDCI models showed the spatial patterns. However, some regions have different drought severity because vegetation and temperature stresses were not considered (e.g., the central-western region and Korean Peninsula on 2 June).

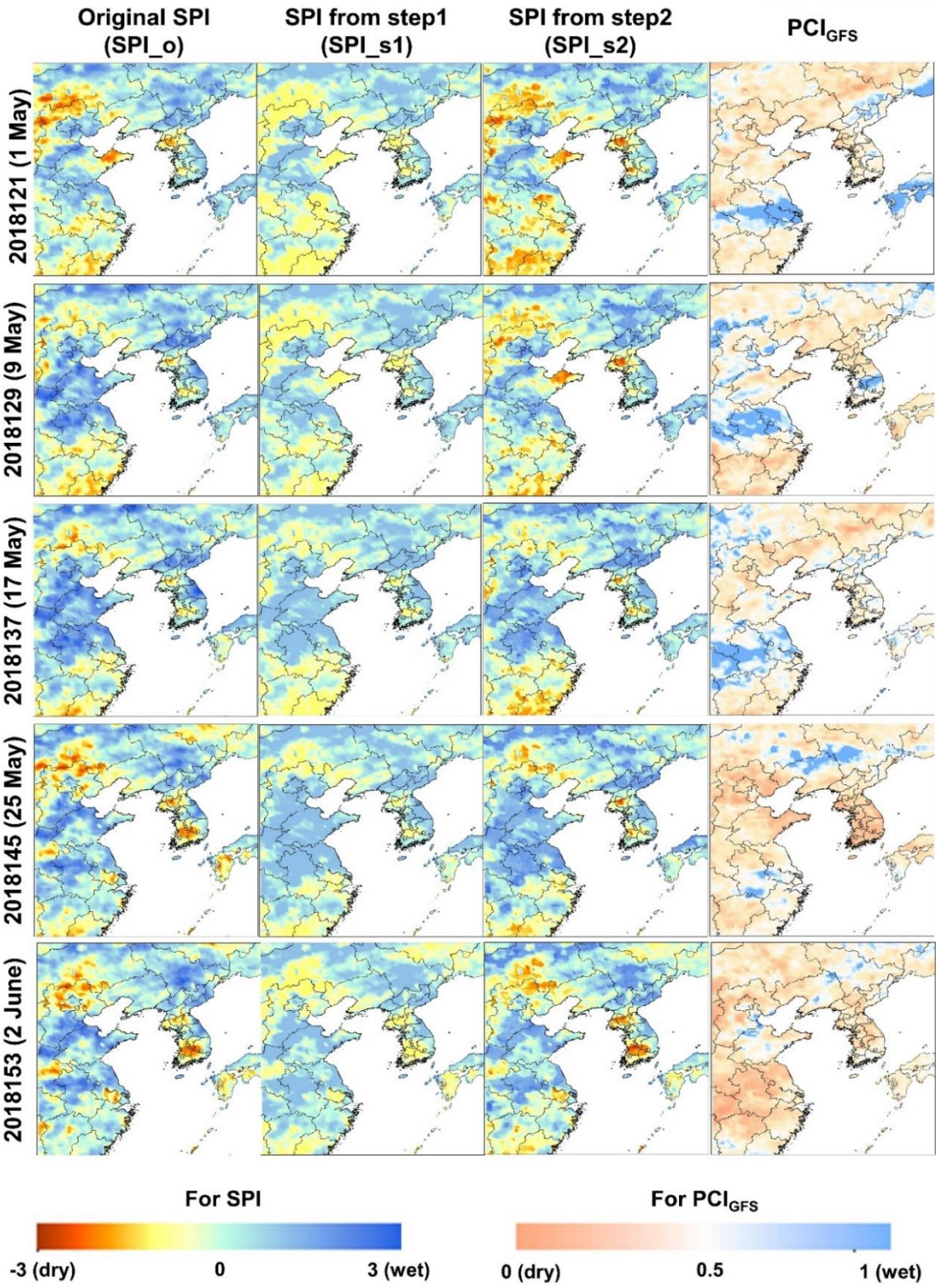

Figure 2.8 Spatial distribution of forecasted SPI from 1 May to 2 June. The vivid red and blue present dry and wet conditions, respectively, in the SPI maps (i.e., reference SPI (SPI_o), SPI from step1 (SPI_s1) and SDCI from step2 (SPI_s2)). The light red and blue present dry and wet conditions, respectively, using precipitation (PCIGFS) from Global Forecasts System (GFS).

Similar to the results of SDCIs, the spatial patterns of drought and no-drought conditions were well detected in SPI_s1 and SPI_s2. The outputs improved when using forecasted precipitation data (PCI_{GFS}) (up to 62% of pixels). For example, SPI_s1 on 9 May was underestimated in the central-western regions (e.g., Jiangsu and Anhui Provinces), although the wet conditions were well detected. SPI_s2 captured wetter conditions than SPI_s1 using $PCI_{GFS,}$ which was helpful to improve the forecasting skills (60% of the pixels were improved on 9 May, Figure 8). Another example is that the drought on 25 May was well captured when considering the dry conditions in PCI_{GFS}, which helped improve about 51% of pixels from SPI_s1 to SPI_s2, especially in the Korean Peninsula. In other words, if there are well-forecasted weather data, the accuracy of the drought forecasting model can be improved. Lorenz et al. (Lorenz et al., 2018) also found that the weather forecasting model is responsible for improving short-term forecasting of drought. In contrast, sometimes the drought forecasting skills were degraded when integrating forecasted data, due to the discrepancy between satellite products and numerical model outputs (e.g., Shandong Province on 9 May and Jiangxi and Fuzhou Province on 2 June) (Kumar et al., 2016). This has already been described in Lorenz et al. (2018).

2.4.3. Novelty and Limitations

In this study, we proposed a new drought forecasting model on a short-term scale using time-series patterns and upcoming weather conditions through ConvLSTM and RF combined. Many previous studies have developed drought forecasting models aimed at improving the accuracy of forecasting skills (e.g., drought area and intensity) (Jiao et al., 2019; Rhee et al., 2010; Otkin et al., 2015; Park and Kim. 2019; Park et al., 2018; Lorenz et al., 2018). However, they still show limitations in forecasting skills: they have relatively simple statistical approaches (e.g., logistic regression and RF) and use historical data. To improve forecasting skills on a short-term scale we combined two models, ConvLSTM and RF, blending temporal patterns of drought and upcoming weather conditions. There are two novelties in our study: 1) the ConvLSTM approach used in this study well reflected both spatial and temporal patterns in drought forecasting. To the best of our knowledge, this is the first study in which ConvLSTM is used in the drought forecasting. 2) Our model fuses time-series patterns and upcoming weather conditions by blending approaches (i.e., ConvLSTM + RF and satellite products + numerical model outputs), which has not been tried much in previous studies. Lorenz et al. (Lorenz et al., 2018) also fused past drought conditions and upcoming weather conditions, but they only investigated drought intensifications. The present study, on the other hand, has examined both drought intensifications and alleviations.

However, despite the novelties of the proposed model, there are still some limitations. First, computational demand is a common problem in deep learning-based models (Lee et al., 2020). Although ConvLSTM reflects time-series patterns well, it requires a significant computational demand in terms of memory and running time when optimizing parameters because there are 26 parameters in the Keras "convlstm2d" function. Second, machine learning approaches are generally known as black box models which give results that are hard to interpret in terms of the causal relationships between and specific importance of variables. Third, the forecasting skills were not good when there were sudden changes in drought conditions. This is because of the discrepancies between satellite products and numerical model outputs and the degradation of forecasting skills in the numerical model with increasing lead time (Park et al., 2017).

To overcome these limitations, several plans can made for future studies: 1) auto-parameterization tools (e.g., Keras-tuner and AutoKeras) should be adopted for cost-effective parameterizing, 2) heatmaps should be generated to interpret the effect of each input variable on model performance (Ye et al., 2019), 3) other machine or deep learning approaches that can reflect complex drought mechanisms should be tested to further improve drought forecasting skills, and 4) an ensemble of various numerical models should be tested in order to reduce the gap between the satellite products and numerical model outputs.

2.5 Conclusions

Short-term forecasting of drought is crucial to reduce the damage to agriculture caused by drought, especially during critical crop yield development stages. Many studies have been conducted for drought forecast, but they still have limited forecasting skills. In this study, a drought forecasting model on a short-term scale was developed using temporal patterns of drought indices and upcoming weather conditions (numerical model outputs) through the synergistic use of ConvLSTM and RF. Two satellite-based drought indices—SDCI and SPI—were selected with a short-time scale (eight days), and the GFS numerical model was used to improve drought forecasting skills, considering upcoming weather conditions. The SDCI- and SPI-based drought forecasting models proposed in this study (ConvLSTM and RF combined) showed competitive results in terms of r (0.90 and 0.93 for validation SDCI and SPI respectively) and RMSE (0.11 and 0.08 for validation of SDCI and SPI, respectively). Furthermore, our drought forecasting model on a short-term scale can be applicable regardless of drought intensification or alleviation. While ConvLSTM resulted in good performance, the combined model showed better results by feeding upcoming weather conditions and topographic characteristics. The proposed drought forecasting model can be operationally used, providing useful information on upcoming drought conditions with high resolution (0.05°).

Chapter 3

3. Tele-connection based (sub)seasonal forecasting of drought using Convolution Neural Network

3.1 Introduction

Drought is a major natural disaster that generally occurs due to the deficit of precipitation (West et al., 2019). Drought can result in substantial damage to vegetation and crop conditions, human communities and economies depending on its persistence. Low levels of precipitation over a period of weeks to months cause meteorological drought (Zhang and Jia, 2013). When meteorological drought is prolonged and the amount of water required for plant growth is insufficient, agricultural drought occurs. Hydrological drought indicates a shortage of water resources, such as a significant reduction in groundwater, reservoir or stream levels (Thomas et al., 2017), while socio-economic drought refers to the impact on the environment when water demands for agricultural, industry and living exceed the available water supply (Wu et al., 2013; Hao et al., 2015; Park et al., 2017; Zhang et al., 2017; Tu et al., 2018).

There is a relatively strong relationship between SST and rainfall in the long-range precipitation forecast (Palmer and Anderson, 1994). Several studies have shown that one factor (e.g., SST) can have varying effects from one region to another (Chiew and McMahon, 2002; Morid et al. (2006), affecting the timing of ocean-atmosphere forcing caused by SST, and the duration and magnitude of continental areas are largely dependent on land-atmosphere feedbacks (Ferguson et al., 2010). Barros and Bowden (2008) improved drought forecasting models in Australia up to 12 months in advance based on precipitation, SST anomaly patterns over the Indian and Pacific Oceans, and the far western Pacific wind-stress anomaly. According to the warming and cooling of SST, the distribution of precipitation can be changed. SST anomaly can make the chance of droughts across East Asia on altering the tropical and subtropical circulation anomalies and carrying the influence far to the mid and high latitudes by teleconnection. The SST anomalies have related to large-scale based drought (Hoerling and Kumar, 2003) by influence on temperature and precipitation.

In this research, we aimed to propose a drought forecasting model on a mid and long-term scale through considering SST teleconnection phenomena using a convolutional neural network (CNN) approaches. The measurable objectives of this study were to 1) forecast drought on a mid and long-term scale (i.e., 1-3month) over a part of East Asia and to 2) analyze the forecasting skills as lead

time. The novelty of this study can be summarized in two aspects: 1) CNN was used to develop the drought forecasting model, which has the advantage of being able to recognize both temporal patterns when using time-series data and (Lee et al., 2020) and spatial characteristics, which is suitable for drought forecasting (Park et al., 2020). To our knowledge, using CNN is the first attempt in drought forecast modeling. 2) The proposed drought forecasting model considers drought patterns and teleconnection phenomena using SST across the Pacific Ocean. 3) “Heat map” which is one of the visualization methods for understanding the CNN models, was used for analyzing the teleconnection phenomena (Zeiler and Fergus, 2014).

3.2 Study area and data

3.2.1 Study area

The study area is a part of East Asia (latitude: 25.17◦ N–45.72◦ N; longitude: 97°E–134°E, East Asia monsoon region) for drought forecast (Figure 3.1). East Asia is generally hot and humid in summer caused by the East Asia monsoon system, while it is dry and cold in winter (Park et al., 2019). Southern and central-eastern China, Japan and Southern and central-eastern China, Japan and the coastlines of South Korea have warm temperate climates, while north-east China, North Korea and inland South Korea have snow climates.

In this study, the study area was divided into eight zones, which have similar drought characteristics on a large scale. In teleconnection-related research, north and south Korea unified for drought forecast (Ham et al., 2016; Yoon and Lee, 2016; Son et al., 2014). Li et al. (2015) assessed drought trend and risk across China and conducted drought climate division using station-based SPI through clustering approach. Referring to Li et al.(2015) and previous research, we finally divided the study area as shown in Figure 3.1.

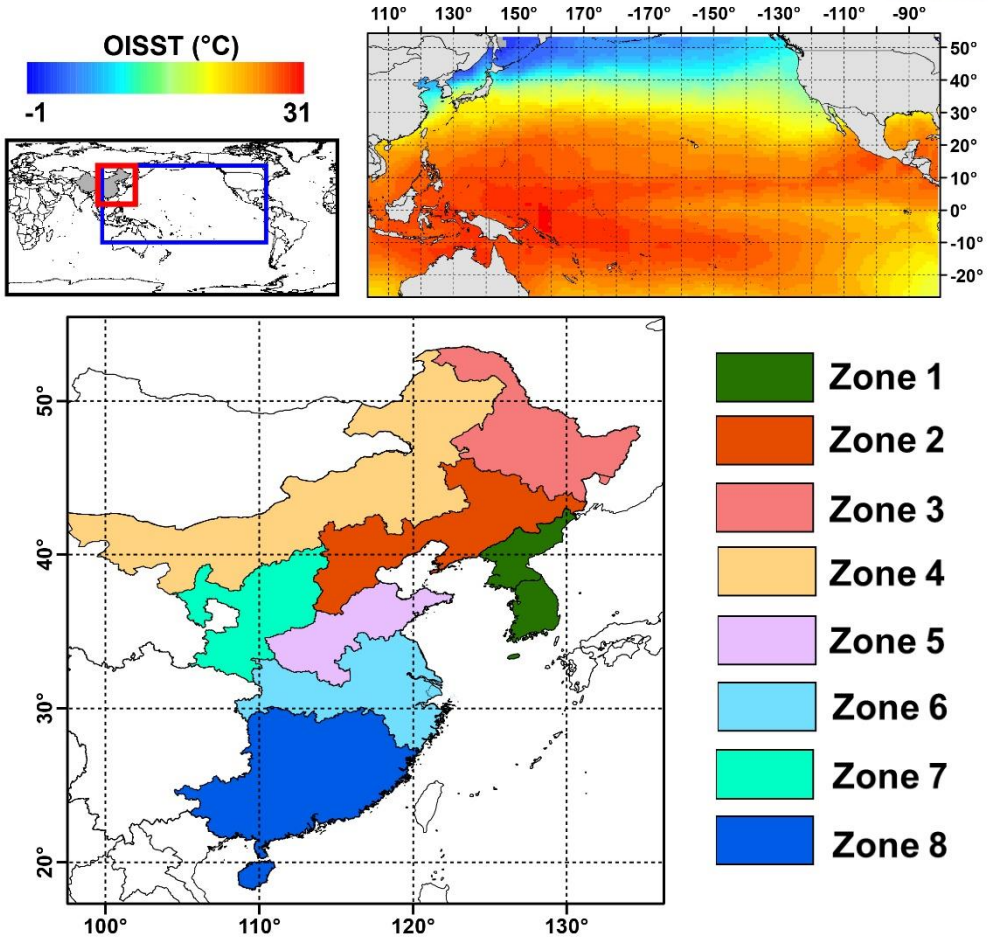

Figure 3.1 The study area of this research with optimal interpolation sea surface temperature (OISST, upper) across the Pacific Ocean and eight zones over East Asia (below)

3.2.2 Data

Our drought forecasting model was developed using 1) reanalysis data based drought index (i.e., Standardized Precipitation Index (SPI)) for documenting the historical patterns across the study area, 2) satellite products as independent variables for considering the possible operational forecast, and 3) SST and Niño indices for reflecting teleconnection phenomena from the Pacific Ocean.

Dependent variables : Standardized Precipitation Index (SPI)

This study used a surface-based drought index, SPI, as a dependent variable to forecast drought across the study areas. The World Meteorological Organization (WMO) recommends that SPI is s suitable drought monitoring index using precipitation rate (Hayes et al., 2011) (Table 3.1). SPI is calculated using the gamma probability distribution of precipitation accumulated at various time scales,

such as one month (McKee et al., 1993; Rhee and Im, 2017). However, ground measurement-based SPI does not provide spatially continuous drought information. It is challenging to explain drought conditions in areas where ground stations are skewed in some areas or low dense (Park et al., 2020). Global Precipitation Climatology Centre provides grid-based gauge-analysis products after conducting controlled station data at 1° spatial resolution. In this study, 3-month SPI (SPI3) from 1984 to 2016 was produced in 1° spatial resolution using GPCC precipitation through SPI function (T. Lee, n.d.) in Matlab 2020a. In order to extract the representative value of SPI3 from each zone in the study area (Figure 3.1), the SPI of 25 percentile at each zone was used.

Table 3.1 Drought categories based on Standardized Precipitation Index (SPI, McKee et al. (McKee et al., 1993)).

SPI value (unitless)	SPI category
-2.0 and less	Extreme Drought
-1.99 to < -1.5	Severe Drought
-1.5 to < -1.0	Moderate Drought
-1.0 to < 0	Mild Drought
0 or more	No Drought

Independent variables

National Oceanic and Atmospheric Administration (NOAA) Physical Sciences Laboratory (PSL) provides Optimally Interpolated SST version 2 (OISSTv2) products, which combines in-situ measurement, satellite SST product, and simulated SST based on sea ice cover (Reynolds et al., 2002; Shukla and Shin, 2020). In this study, monthly OISSTv2 monthly data (1° spatial resolution) derived by linear interpolation of the weekly OISSTv2 product (Reynolds and Banzon 2008) was used from 1983 to 2016 to reflect teleconnection phenomena (https://psl.noaa.gov/data/gridded/data.noaa.oisst.v2.html). We masked the Pacific Ocean (longitude: 105°E to 280°E, latitude: 30°S to 50°N) for modeling and additionally obtained Niño indices (i.e., Niño1+2 and Niño3.4) based on OISSTv2. Niño index— Niño1+2 and Niño3.4— is an anomaly of SST the part of the pacific ocean(i.e., 0-10S, 90W-80W and 5N-5S, 170W-120W, respectively) which is useful for explaining teleconnection based drought conditions.

Global Precipitation Mission (GPM), a follow-up to the Tropical Rainfall Measuring Mission (TRMM), was launched in 2014, which has monitored precipitation including rainfall and snow

worldwide based on Du/Ka-band Dual-frequency Precipitation Radar (DPR) and multi-channel GPM Microwave Imager (GMI) (Draper et al., 2015; Casella et al., 2017; Son et al., 2021). In order to calculate GPM-based SPI, we accumulated 3-month of GPM IMERG precipitation (Final Run, 3B-MO, version 06) from 2000 to 2018 provided by Earthdata, which was the same way with GPCC-based SPI3. Unlike dependent variables, the reason for using satellite-based precipitation is to consider the possibility of real-time operation. For adopting independent variables on CNN model, precipitation data were normalized using minimum and maximum values in each grid.

3.3 Methodology

Figure 3.2 shows the model structure of this study. In this study, we adopted combined CNN structures, called “the integrated CNN model”. It consists of one 3D CNN model and three 1D CNN models due to the different size(or dimension) of independent variables. In the 3D CNN model, normalized SST over the Pacific Ocean was used to reflect teleconnection phenomena on drought conditions across East Asia. A series of past 12 months of normalized SST data (from t-11 to t) has been applied as independent variables in 3D CNN. In terms of 1D CNN model, there are three types of independent variables—monthly precipitation from GPM, Niño1+2, and Niño3.4. The series of past 12 months of three types of independent variables (from t-11 to t) are entered into each 1D CNN model. The reason for applying the values of the past 12 months for each independent variable is to determine how many months earlier data would help forecast future drought conditions (i.e., one to three months). The temporal patterns of precipitation can especially explain how the drought conditions have been changing for about a year in each region. During the entire study period, the CNN model was generated by using information from 1984 to 2013 and model performance was evaluated for 2014-2016.

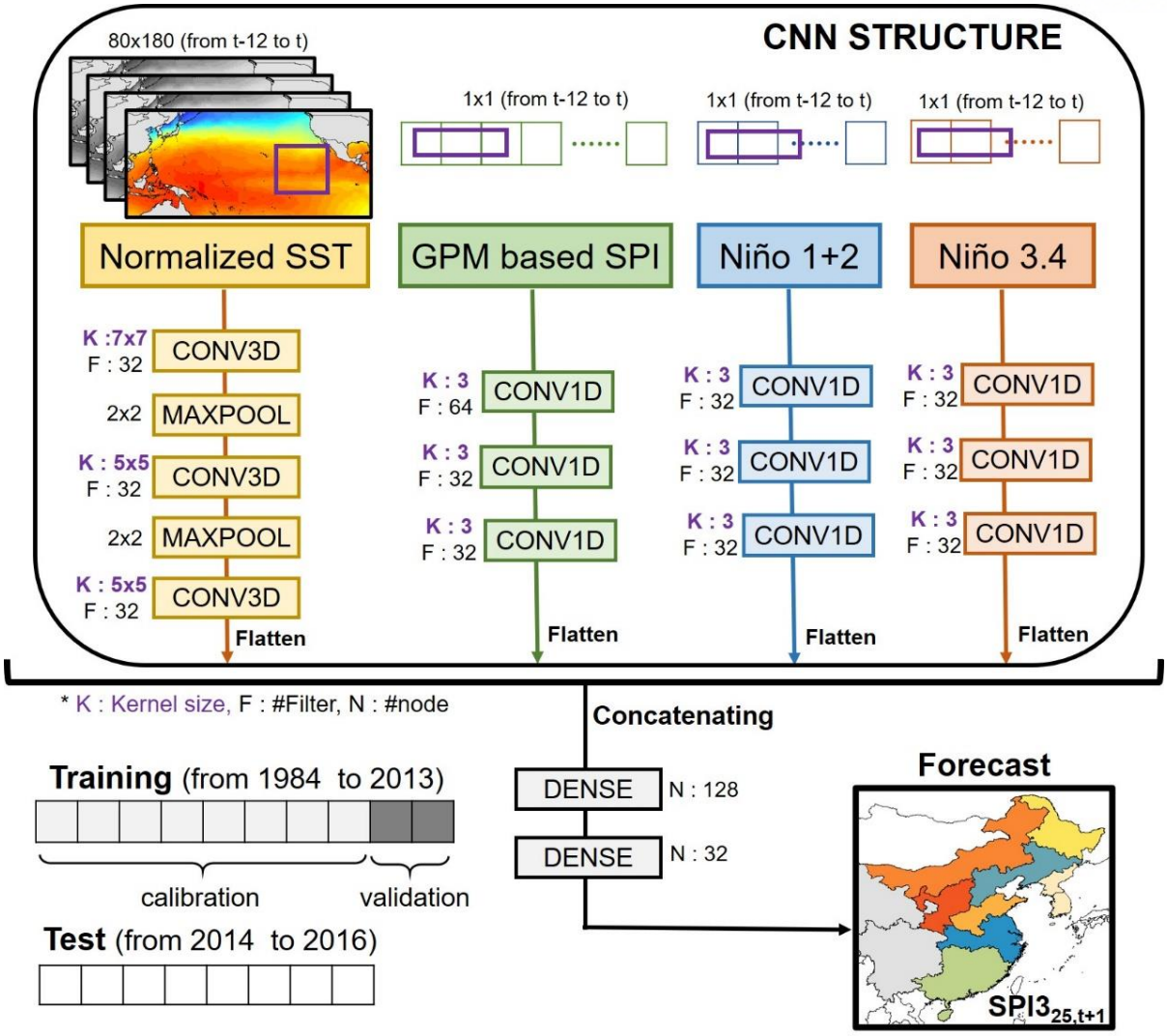

Figure 3.2 Model structure for developing drought forecast model in this study.

3.3.1 Convolutional Neural Networks (CNN)

CNN is one of the artificial neural network approaches, which is widely applied in recognizing visualization data such as handwriting, photos, and medical images (LeCun et al., 2015; Lee et al., 2019; Krizhevsky et al., 2017). A major characteristic of CNN is extracting the spatial patterns using convolutional layers (Yoo et al., 2019). CNN modeling proceeds as follows :1) The input data pattern is extracted using the convolutional layer. 2) The pooling layers make dimension decrease to avoid overfitting problems. 3) The extracted feature determines the output. 4) the last convolutional layer is flattened so that it can be applied to the fully connected layer.

In this study, CNN models were constructed using the Keras library in Python. In order to find an optimal model in 3D CNN (for SST), 32, 64 and 128 filters and 3x3, 5x5, 7x7 kernels at convolutional layers were tested. The final CNN structure is shown in Figure 3.2. The ReLU and linear activation function was adopted at 3D CNN and 1D CNN, respectively. In the last steps, a linear function was used to predict drought conditions with 128 and 32 nodes.

3.3.2 Accuracy assessment

In order to evaluate the performances of the forecasting model, two statistical metrics were used: correlation coefficient (r) and normalized root mean square error (nrmse).

$$r = \frac{n(\sum y\hat{y}) - (\sum y)(\sum \hat{y})}{\sqrt{[n\sum y^2 - (\sum y)^2][n\sum \hat{y}^2 - (\sum \hat{y})^2]}}$$

$$rmse = \sqrt{\frac{\sum(y-\hat{y})^2}{n}}, \qquad nrmse = \frac{RMSE}{y_{max} - y_{min}}$$

n and m are the number of samples and y and $\hat{y}$ are the values of reference and predicted drought indices, respectively. r is the strength of a linear relationship between predicted and real values, which is useful when r is closer to 1. nrmse indicates the relative standard deviation of the forecasting errors. The model can be evaluated for having an excellent, good, fair, or poor performance when nrmse lower than 0.1, higher than 0.1 but lower than 0.2, higher than 0.2 but lower than 0.3, or higher than 0.3 (Dettori et al., 2011; Nouri and Homaee, 2018; Feng et al., 2020).

3.4. Results and discussion

3.4.1 The performance of the drought forecasting model

Drought Accuracy values were compared between lead times of 1–3 months (Figure. 3.3). In the result of calibration, averaged r(nrmse) are 0.96(0.09), 0.94(0.06) and 0.92(0.07) in lead time 1, 2, and 3 respectively. Regardless of lead time, the model has excellent performance (r is over 0.94 and nrmse is under 0.1). In the result of validation, averaged r(nrmse) are 0.63(0.17), 0.62(0.16) and 0.67(0.15) in lead time 1, 2, and 3 respectively. Although validation performance was lower than calibration, the model still has good performance (nrmse is under 0.2). However, for test, the value of r(nrmse) (Figure 3.3, below) was rapidly reduced(increased). The reasons are that 1) the distribution of the calibration data used in model training and the test data is different, and 2) SST across the Pacific Ocean is not cover SPI due to other drought factors (e.g., Atlantic or Indian Ocean and snow depth). Nevertheless, it showed fair performances (under 0.3 in nrmse). As lead time increase from 1 to 3, the accuracy in r is rapidly decreased in the test, unlike in calibration and validation. It is consistent with Rhee and Im (2017) results, but there is relatively little decreasing accuracy.

There is no difference in each zone's calibration and validation accuracy at one lead time (orange color in Figure 3.3). However, there is a large difference in accuracy for each zone at two and three lead time (purple and green color respectively, in Figure 3.3). For example, there are low correlations in zone 1, 4 and 6 caused by inconsistency between the predicted data and the temporal pattern of the reference.

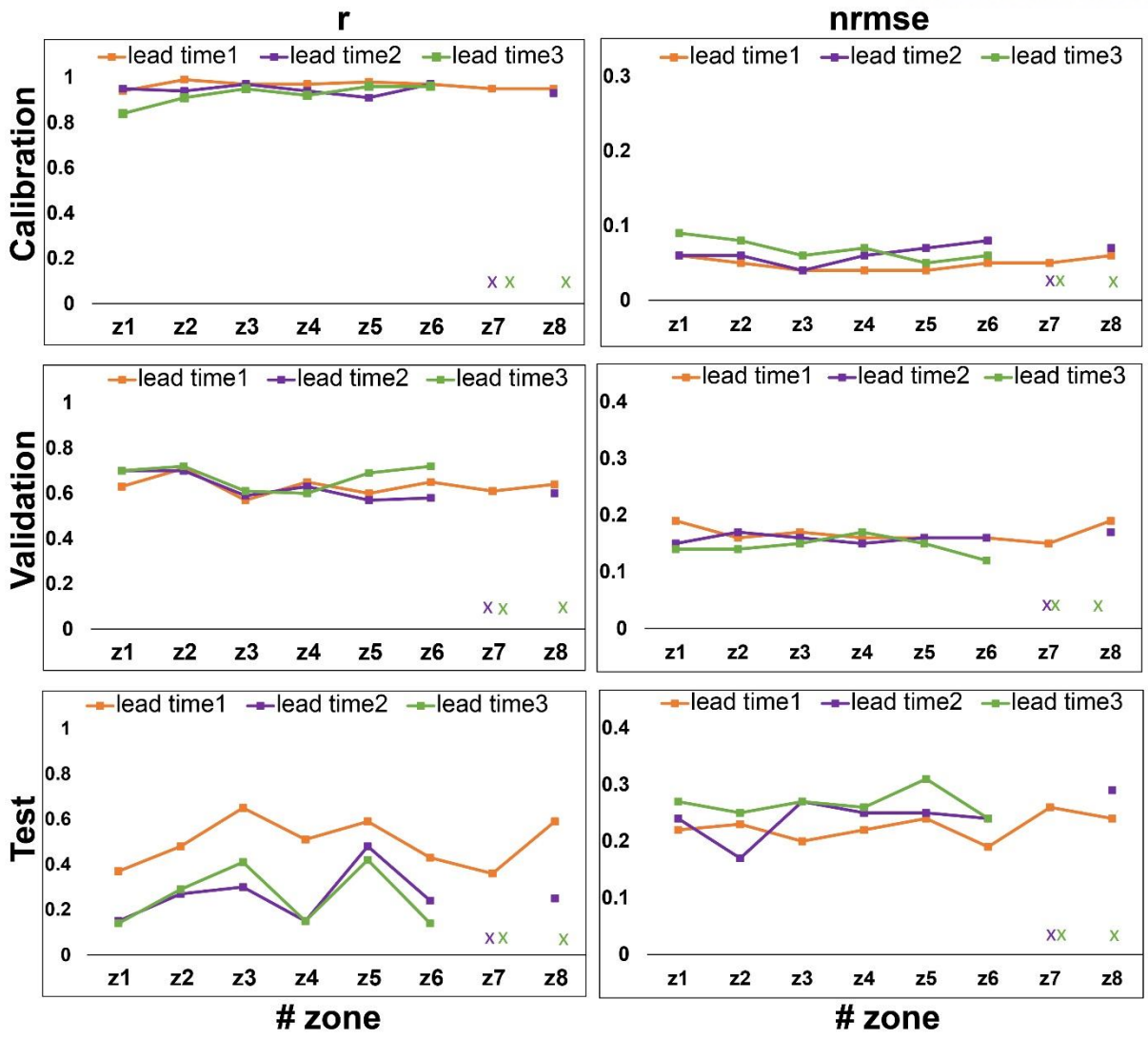

Figure 3.3 The model performance in terms of calibration, validation and test. The performances of each lead time model (i.e., lead time1-3) were shown as the orange, purple and green line, respectively. There are no models in zone 7 for lead time 2-3 and zone 8 for lead time 3 (described as "x").

Figure 3.4 shows the temporal patterns of GPCC based SPI (reference), CNN-based forecasted SPI (forecasted), and Climate Forecast System (CFS) based SPI for one lead time. CFSv2 operational forecasts model was used in order to compare to our model performance. CFS has been providing data since 2011, we only used 2016 data due to data missing in the past period. Our results showed that the pattern tends to match the time series, and some periods have tended to be not good. Due to the lack of precipitation by 2014 and early 2015, GPCC-based SPIs are showing drought conditions, which is consistent with the results of our model. Regardless of zones, The predicted values showed the tendency to be averaged, and the maximum value (underestimated) and the minimum value (overestimated) were not good (e.g., zone3, zone4 and zone8). Except for zone8, CFS based SPI tended to match GPCC-based SPIs in summer, fall and winter. The reason that zone 8 has a different pattern is the CFS model overestimated precipitation considering regional characteristics.

Figure 3.5 shows the temporal patterns for two lead time. Our model showed a tendency not to be very sensitive to changes in temporal patterns despite the increase in lead time. The forecasted values showed the tendency to be averaged, and the maximum value (underestimated) and the minimum value (overestimated) were not good (e.g., zone3, zone4 and zone8). In CFS based SPI, generally, spring precipitation is more overestimated than one lead time. CFS has better temporal patterns (except for spring 2016), but there are relatively large errors in the intensity.

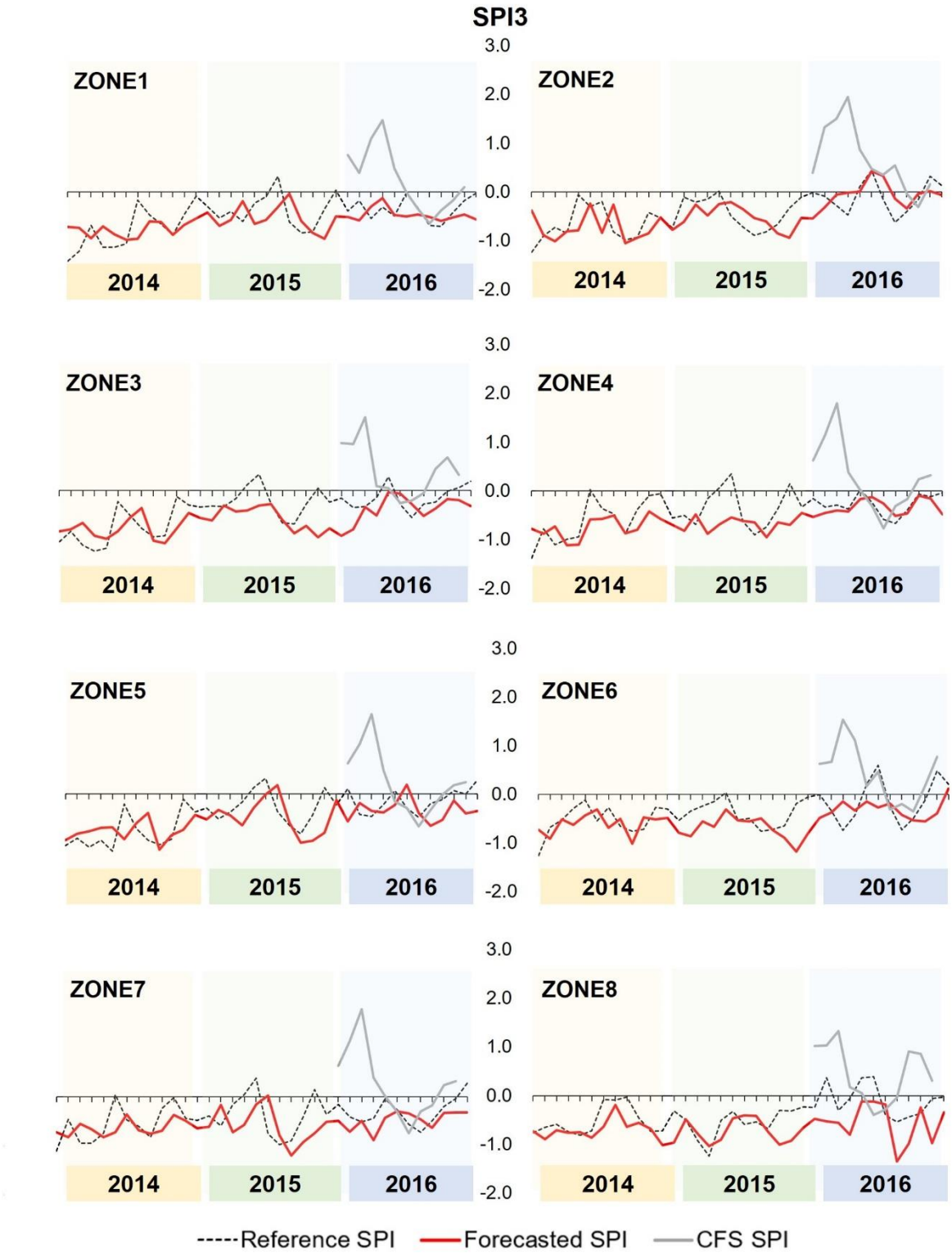

Figure 3.4 The time-series patterns of SPI in one lead time from 2013 to 2016. The reference SPI, Forecasted SPI, and CFS SPI were described by black dach, red line, and gray line, respectively.

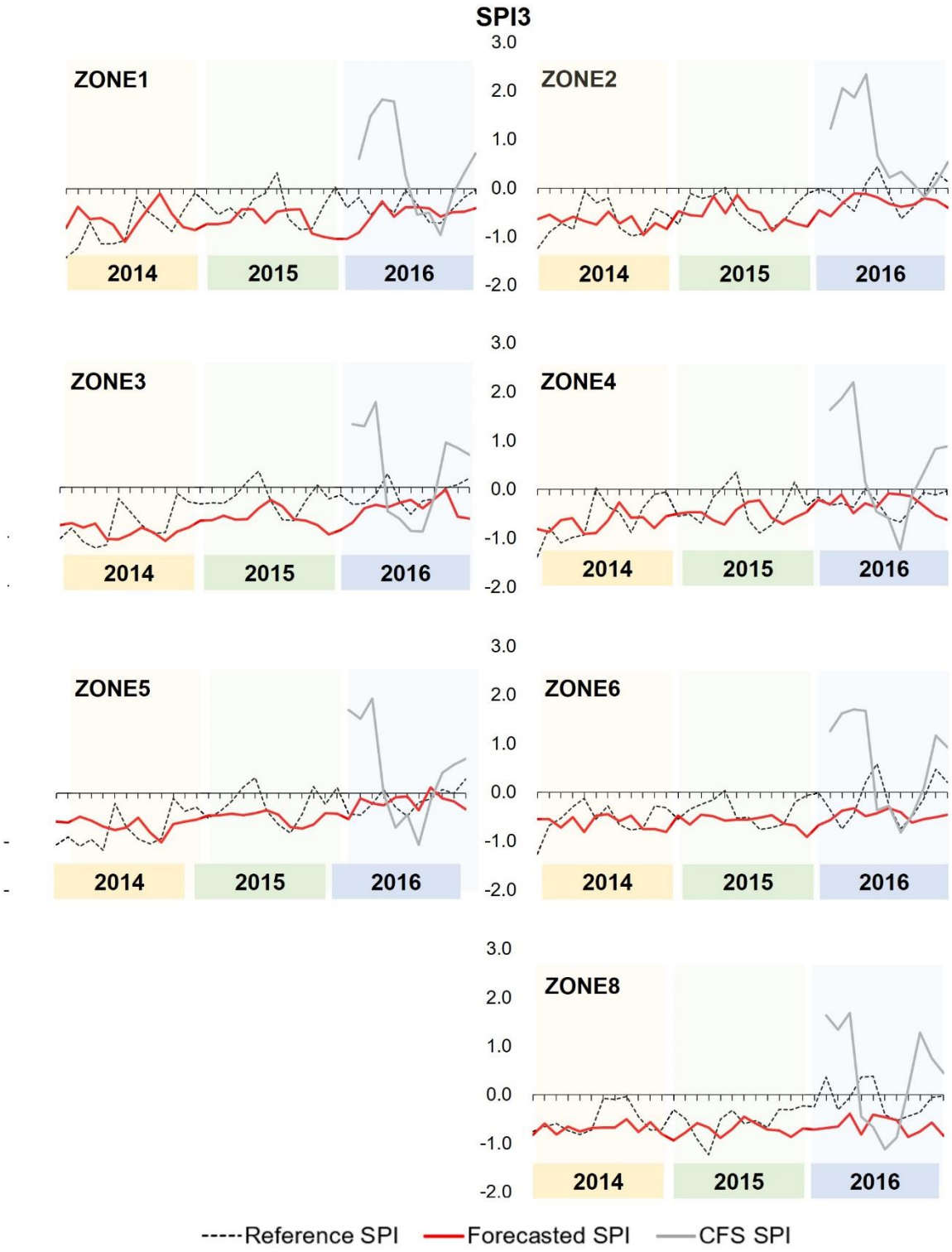

Figure 3.5 The time-series patterns of SPI in two lead time from 2013 to 2016. The reference SPI, Forecasted SPI, and CFS SPI were described by black dash, red line, and gray line, respectively. There is no model in zone 7.

3.4.2 The spatial distribution of heat map

CNN is called “Black box” because they struggle to identify the general relationship between the variables and the model parameters (Lee et al., 2019). However, the part of the CNN model can be interpreted through the heat map (Figure 3.6). Since June 2010, the SST in the Niño 3.4 area was recorded under -0.5 and continued until April 2011 (La Nina). In May 2011, South-East China (i.e., zone 6) suffered from drought due to a deficiency of precipitation. Jin et al.(2013) analyzed SST anomaly in the western and eastern equatorial Pacific from Jan to May of 2011. The positive signal in the western Pacific and the negative signal in the eastern Pacific (Figure 3.6). La Nina facilitate a less than general winter and spring. Figure 3.6 described the SST anomaly and heat map from CNN model from May 2010 to April 2011. In the heat map, there are similar patterns to SST anomaly. The forecasted SPI in zone6 was derived from Niño1+2 and Niño3.4 regions. At least 8 months ago, the feature began to appear. There is a similar case in spring 1997 (drought case in the same region(zone6)). The heat map also indicates a strong signal in Niño zones 8 months ago.

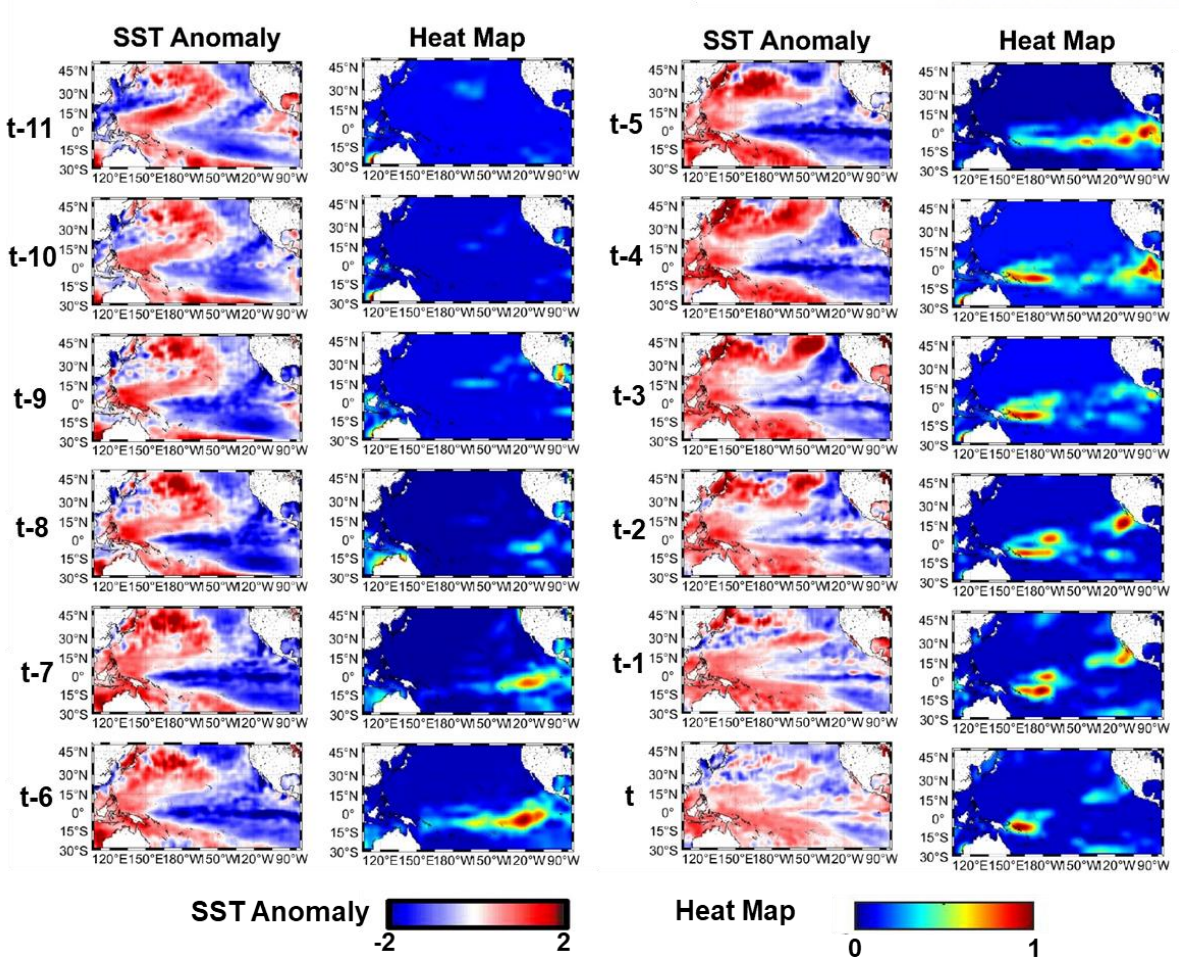

Figure 3.6 The SST anomaly and heat map from CNN model for May 2011 (for zone 6). The positive(negative) anomaly of SST was presented red(blue) color, and the strong signals of the heat map were presented by red color.

In May 2014, North-East China (i.e., zone 2 and 3) suffered from drought due to a deficiency of precipitation. In the heat map 9 months ago, a strong signal was found in SST near Australia, and a strong positive signal was also observed in SST anomaly map (Figure 3.7). The positive SST anomaly makes air mass sink faster, which causes the southward shift of subtropical westerly jet (Zhao et al., 2020). Furthermore, it leads to drier conditions in northern and wetter conditions in central and southeastern East Asia. Strong signals were shown in the eastern Pacific three to seven months ago and in the western Pacific one to two months ago.

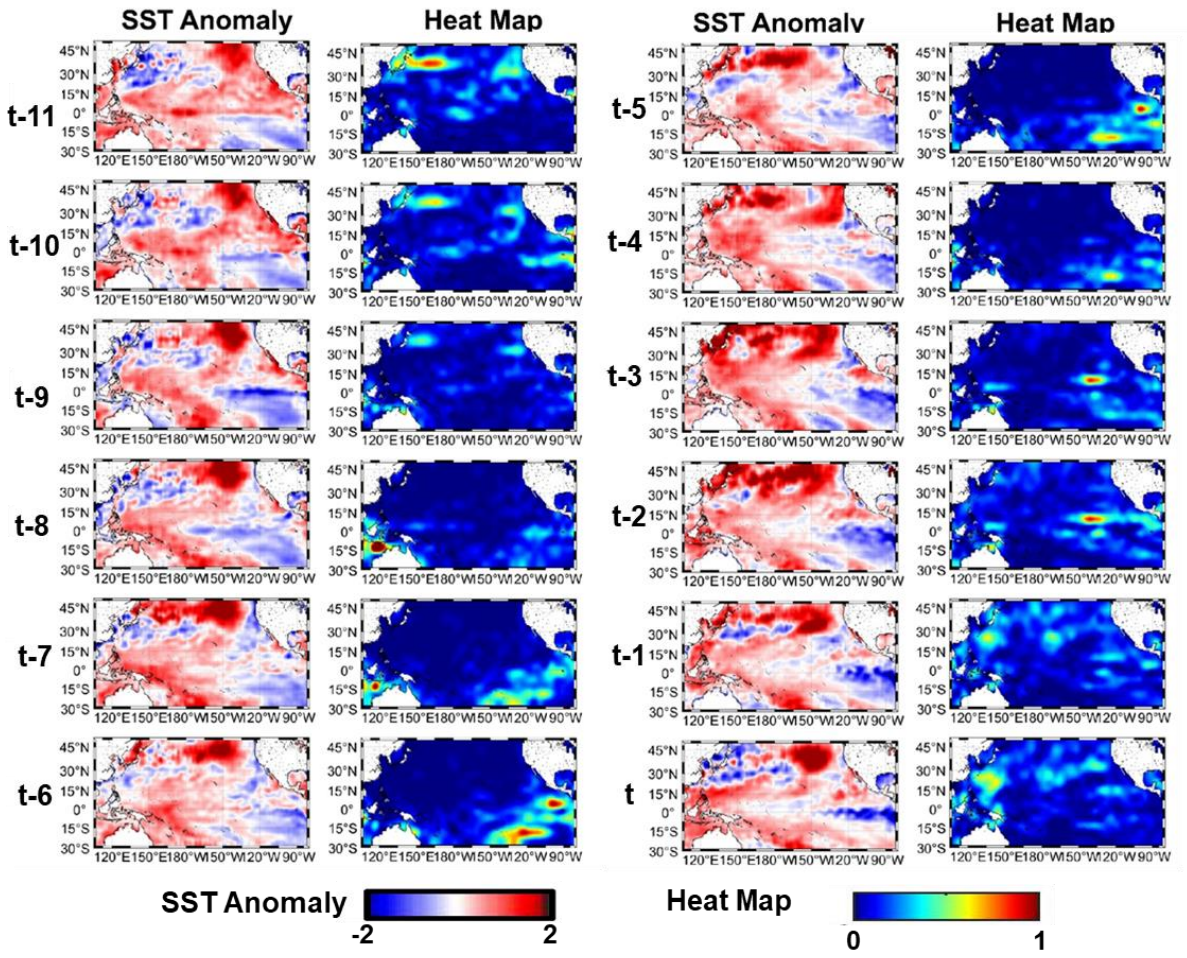

Figure 3.7 The SST anomaly and heat map from CNN model for May 2014 (for zone 2 and 3). The positive(negative) anomaly of SST was presented red(blue) color, and the strong signals of the heat map were presented by red color.

3.5. Conclusions

Sub-seasonal and seasonal forecasting of drought is crucial to reduce the damage caused by drought and manage drought-related polices. Many studies have been conducted for drought forecast, but they still have limited forecasting skills. In this study, a drought forecasting model on a mid-and long-term scale (one-three lead time) was developed using temporal patterns of drought indices and teleconnection phenomena of SST through the CNN. Reanalysis based drought index, SPI, were selected with a sub-seasonal and seasonal timescale (one to three months), and satellite-based variable, precipitation and SST across the Pacific Ocean. The SPI-based drought forecasting models proposed in this study showed competitive results in terms of r (0.5-0.7 for validation SPI at one lead time) and nrmse (0.1-0.2 for validation SPI at one lead time) regardless of regions. As the lead time increased, the accuracy tended to fall, but it showed good results compared to CFS, which is a numerical model data. When compared to a drought case, the SST of 8 months ago was influenced by the results. We confirmed the consistency of existing studies through Heat Map. The proposed drought forecasting model can be operationally used, providing useful information on upcoming drought conditions.

Chapter 4

4. Conclusions

The objective of this is to develop a short- mid-and long-term drought forecasting model by integrating satellite products and model outputs over East Asia through the Deep learning approaches. Drought forecasting is essential for effectively managing drought-related damage and providing relevant drought information to decision-makers so they can make appropriate decisions in response to drought. Although there have been great efforts in drought-forecasting research, drought forecast is still difficult.

Drought-forecasting models on a short-term scale (8 days) were developed considering the temporal patterns of satellite-based drought indices and numerical model outputs through the synergistic use of convolutional long short term memory (ConvLSTM) and random forest (RF) approaches over a part of East Asia. Two widely used drought indices—Scaled Drought Condition Index (SDCI) and Standardized Precipitation Index (SPI)—were used as target variables. Through the combination of temporal patterns and the upcoming weather conditions (numerical model outputs), the overall performances of drought-forecasting models (ConvLSTM and RF combined) produced competitive results in terms of r and nrmse. Furthermore, our short-term drought-forecasting model can be useful regardless of drought intensification or alleviation. The proposed drought-forecasting model can be operationally used, providing useful information on upcoming drought conditions with high resolution.

Drought forecasting model on sub-seasonal and seasonal scales (one-three lead time) was developed using temporal patterns of drought indices and teleconnection phenomena of SST through the CNN. Reanalysis based drought index, SPI, were selected (one to three months lead time), and satellite-based variable, precipitation and SST across the Pacific Ocean. The SPI-based drought forecasting models proposed in this study showed competitive results in terms of r (0.5-0.7 for validation SPI at one lead time) and nrmse (0.1-0.2 for validation SPI at one lead time) regardless of regions. As the lead time increased, the accuracy tended to fall, but it showed good results compared to CFS, which is a numerical model data. When compared to a drought case, the SST of 8 months ago was influenced on the results. we confirmed the consistency of existing studies through Heat Map.

Deep learning-based drought forecasting models over East Asia will be helpful to effectively manage drought-related damage and provide relevant drought information to decision-makers so they can make appropriate decisions in response to drought.

Chapter 5

5. Outlook and Future works

Although great efforts have been made in drought forecast, there are still some limitations. Despite the presence of many drought indices, the drought forecasting performance based on various drought indices was not evaluated. Unlike the United States, there is no official drought index in East Asia. Recently, Son et al. (2021) developed a drought index that reflects the climatic characteristics of East Asia. Since the index contains several drought factors, it is expected to be helpful in drought forecasts in East Asia if the characteristics of each drought factor are well understood.

Compared to previous studies, forecasting skills have been improved, but the predictive ability was still low for the rapid change (flash drought). To overcome this limitation, several plans can try for future studies : 1) machine or deep learning approaches that can reflect complex drought mechanisms should be tested to further improve drought forecasting skills, and 2) an ensemble of various numerical models should be tested in order to reduce the gap between the satellite products and numerical model outputs.

Another limitation is the efficiency of the drought forecasting model proposed in this dissertation. Although models were divided according to the forecasting scale (e.g., short-term, seasonal), it is not efficient for decision-makers who need drought forecasting information. Additionally, there is no information affected area and drought persistency. In order to minimize damage caused by drought, information (e.g., timing, the affected area, duration as well as the drought intensity) should be forecasted.

Reference

Abebe, A., & Foerch, G. (2008). Stochastic simulation of the severity of hydrological drought. *Water and Environment Journal*, 22(1), 2–10.

Akay, D., & Atak, M. (2007). Grey prediction with rolling mechanism for electricity demand forecasting of Turkey. *Energy*, 32(9), 1670–1675.

Anderson, M. C., Norman, J. M., Mecikalski, J. R., Otkin, J. A., & Kustas, W. P. (2007). A climatological study of evapotranspiration and moisture stress across the continental United States based on thermal remote sensing: 2. Surface moisture climatology. *Journal of Geophysical Research: Atmospheres*, 112(D11).

Bayissa, Y. A., Tadesse, T., Svoboda, M., Wardlow, B., Poulsen, C., Swigart, J., & Van Andel, S. J. (2019). Developing a satellite-based combined drought indicator to monitor agricultural drought: a case study for Ethiopia. *GIScience & Remote Sensing*, 56(5), 718–748.

Belayneh, A., Adamowski, J., & Khalil, B. (2016). Short-term SPI drought forecasting in the Awash River Basin in Ethiopia using wavelet transforms and machine learning methods. *Sustainable Water Resources Management*, 2(1), 87-101.

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.

Benali, A., Carvalho, A. C., Nunes, J. P., Carvalhais, N., & Santos, A. (2012). Estimating air surface temperature in Portugal using MODIS LST data. *Remote Sensing of Environment*, 124, 108–121.

Borji, M., Malekian, A., Salajegheh, A., & Ghadimi, M. (2016). Multi-time-scale analysis of hydrological drought forecasting using support vector regression (SVR) and artificial neural networks (ANN). *Arabian Journal of Geosciences*, 9(19), 725.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

Brown, J. F., Wardlow, B. D., Tadesse, T., Hayes, M. J., & 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.

Casella, D., Panegrossi, G., Sanò, P., Marra, A. C., Dietrich, S., Johnson, B. T., & Kulie, M. S. (2017). Evaluation of the GPM-DPR snowfall detection capability: Comparison with CloudSat-CPR.

Atmospheric Research, 197, 64-75.

Cancelliere, A., Di Mauro, G., Bonaccorso, B., & Rossi, G. (2007). Drought forecasting using the standardized precipitation index. *Water Resources Management*, 21(5), 801–819.

Centre for Research on the Epidemiology of Disasters (CRED) Natural Disasters in 2018. Available online: https://cred.be/sites/default/files/adsr_2018.pdf (access on 06 August 2020). (2019). https://cred.be/sites/default/files/adsr_2018.pdf

Chiew, F. H., & McMAHON, T. A. (2002). Global ENSO-streamflow teleconnection, streamflow forecasting and interannual variability. *Hydrological Sciences Journal*, 47(3), 505-522.

Cho, D., Yoo, C., Im, J., Lee, Y., & Lee, J. (2020). Improvement of spatial interpolation accuracy of daily maximum air temperature in urban areas using a stacking ensemble technique. *GIScience & Remote Sensing*, 57(5), 1–17.

Choi, J. W., Cha, Y., & Kim, J. Y. (2017). Prediction Experiment of Regional Drought over Korea Using the Similarity of Spatiotemporal Patterns of Past Droughts. *J Climatol Weather Forecasting*, 5(190), 2.

Climate Prediction Center (CPC) US Seasonal Drought Outlook (SDO), Available online: https://www.cpc.ncep.noaa.gov/products/outreach/publications.shtml (accessed on 06 August 2020).

Cruz, G., & Bernardino, A. (2019). Learning Temporal Features for Detection on Maritime Airborne Video Sequences Using Convolutional LSTM. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9), 6565–6576.

Dai, A. (2011). Drought under global warming: a review. Wiley Interdisciplinary Reviews: *Climate Change*, 2(1), 45–65.

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature Climate Change*, 3(1), 52–58.3

Dettori, M., Cesaraccio, C., Motroni, A., Spano, D., & Duce, P. (2011). Using CERES-Wheat to simulate durum wheat production and phenology in Southern Sardinia, Italy. *Field crops research*, 120(1), 179-188.

Davydenko, A., & Fildes, R. (2016). Forecast error measures: critical review and practical recommendations. Business Forecasting: Practical Problems and Solutions. *Wiley*, 34.

Demisse, G. B., Tadesse, T., Wall, N., Haigh, T., Bayissa, Y., & Shiferaw, A. (2019). Linking seasonal drought product information to decision makers in a data-sparse region: A case study in the Greater

Horn of Africa. *Remote Sensing Applications: Society and Environment*, 14, 200-206.

Draper, D. W., Newell, D. A., Wentz, F. J., Krimchansky, S., & Skofronick-Jackson, G. M. (2015). The global precipitation measurement (GPM) microwave imager (GMI): Instrument overview and early on-orbit performance. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(7), 3452-3462.

Durdu, Ö. F. (2010). Application of linear stochastic models for drought forecasting in the Büyük Menderes river basin, western Turkey. *Stochastic Environmental Research and Risk Assessment*, 24(8), 1145-1162.

Fang, W., Huang, S., Huang, Q., Huang, G., Wang, H., Leng, G., ... & Guo, Y. (2019). Probabilistic assessment of remote sensing-based terrestrial vegetation vulnerability to drought stress of the Loess Plateau in China. *Remote Sensing of Environment*, 232, 111290.

Fernández, C., Vega, J. A., Fonturbel, T., & Jiménez, E. (2009). Streamflow drought time series forecasting: a case study in a small watershed in North West Spain. *Stochastic Environmental Research and Risk Assessment*, 23(8), 1063.

Ferguson, I. M., Dracup, J. A., Duffy, P. B., Pegion, P., & Schubert, S. (2010). Influence of SST forcing on stochastic characteristics of simulated precipitation and drought. *Journal of Hydrometeorology*, 11(3), 754-769.

Feng, P., Wang, B., Luo, J. J., Li Liu, D., Waters, C., Ji, F., ... & Yu, Q. (2020). Using large-scale climate drivers to forecast meteorological drought condition in growing season across the Australian wheatbelt. *Science of The Total Environment*, 138162.

Fung, K. F., Huang, Y. F., Koo, C. H., & Soh, Y. W. (2020). Drought forecasting: A review of modelling approaches 2007--2017. *Journal of Water and Climate Change*, 11(3), 771–799.

Ghimire, S., Deo, R. C., Raj, N., & Mi, J. (2019). Deep solar radiation forecasting with convolutional neural network and long short-term memory network algorithms. *Applied Energy*, 253, 113541.

Ham, Y. G., Kug, J. S., Yeh, S. W., & Kwon, M. (2016). Impact of two distinct teleconnection patterns induced by western Central Pacific SST anomalies on Korean temperature variability during the early boreal summer. *Journal of Climate*, 29(2), 743-759.

Han, L., Zhang, Q., Ma, P., Jia, J., & Wang, J. (2016). The spatial distribution characteristics of a comprehensive drought risk index in southwestern China and underlying causes. *Theoretical and Applied Climatology*, 124(3–4), 517–528.

Han, P., Wang, P., Tian, M., Zhang, S., Liu, J., & Zhu, D. (2012). Application of the ARIMA models in drought forecasting using the standardized precipitation index. *International Conference on Computer and Computing Technologies in Agriculture*, 352–358.

Han, P., Wang, P. X., Zhang, S. Y., & others. (2010). Drought forecasting based on the remote sensing data using ARIMA models. *Mathematical and Computer Modelling*, 51(11–12), 1398–1403.

Han, Y., Li, Z., Huang, C., Zhou, Y., Zong, S., Hao, T., Niu, H., & Yao, H. (2020). Monitoring Droughts in the Greater Changbai Mountains Using Multiple Remote Sensing-Based Drought Indices. *Remote Sensing*, 12(3), 530.

Hao, Z., & AghaKouchak, A. (2013). Multivariate standardized drought index: a parametric multi-index model. *Advances in Water Resources*, 57, 12-18.

Hao, C., Zhang, J., & Yao, F. (2015). Combination of multi-sensor remote sensing data for drought monitoring over Southwest China. *International Journal of Applied Earth Observation and Geoinformation*, 35, 270-283.

Hayes, M., Svoboda, M., Wall, N., & Widhalm, M. (2011). The Lincoln declaration on drought indices: universal meteorological drought index recommended. *Bulletin of the American Meteorological Society*, 92(4), 485-488.

He, K., & Sun, J. (2015). Convolutional neural networks at constrained time cost. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5353–5360.

Hoerling, M., & Kumar, A. (2003). The perfect ocean for drought. Science, 299(5607), 691-694.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679–688.

Im, J., Park, S., Rhee, J., Baik, J., & Choi, M. (2016). Downscaling of AMSR-E soil moisture with MODIS products using machine learning approaches. *Environmental Earth Sciences*, 75(15), 1120.

Jiao, W., Tian, C., Chang, Q., Novick, K. A., & Wang, L. (2019). A new multi-sensor integrated index for drought monitoring. *Agricultural and forest meteorology*, 268, 74-85.

Jiao, W., Zhang, L., Chang, Q., Fu, D., Cen, Y., & Tong, Q. (2016). Evaluating an enhanced vegetation condition index (VCI) based on VIUPD for drought monitoring in the continental United States. *Remote Sensing*, 8(3), 224.

Jin, D., Guan, Z., & Tang, W. (2013). The extreme drought event during winter–spring of 2011 in East

China: combined influences of teleconnection in midhigh latitudes and thermal forcing in maritime continent region. *Journal of climate*, 26(20), 8210-8222.

Kim, T.-W., & Valdés, J. B. (2003). Nonlinear model for drought forecasting based on a conjunction of wavelet transforms and neural networks. *Journal of Hydrologic Engineering*, 8(6), 319–328.

Kogan, F. N. (1995a). Application of vegetation index and brightness temperature for drought detection. *Advances in Space Research*, 15(11), 91–100.

Kogan, F. N. (1995b). Droughts of the late 1980s in the United States as derived from NOAA polar-orbiting satellite data. *Bulletin of the American Meteorological Society*, 76(5), 655–668.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, 15(3), 259–263.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84-90.

Kumar, P., Kishtawal, C. M., & Pal, P. K. (2016). Skill of regional and global model forecast over Indian region. *Theoretical and Applied Climatology*, 123(3–4), 629–636.

Kumar, V., & Panu, U. (1997). PREDICTIVE ASSESSMENT OF SEVERITY OF AGRICULTURAL DROUGHTS BASED ON AGRO-CLIMATIC FACTORS 1. *JAWRA Journal of the American Water Resources Association*, 33(6), 1255-1264.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, 521(7553), 436-444.

Lee, J., Im, J., Cha, D.-H., Park, H., & Sim, S. (2020). Tropical cyclone intensity estimation using multi-dimensional convolutional neural networks from geostationary satellite data. *Remote Sensing*, 12(1), 108.

Lee, C. S., Sohn, E., Park, J. D., & Jang, J. D. (2019). Estimation of soil moisture using deep learning based on satellite data: a case study of South Korea. *GIScience & Remote Sensing*, 56(1), 43-67.

Lee, T. (2020). Standardized Precipitation Index. Available online: https://www.mathworks.com/matlabcentral/fileexchange/26018-standardized-precipitation-index (accessed on 06 August 2020).

Leilah, A. A., & Al-Khateeb, S. A. (2005). Statistical analysis of wheat yield under drought conditions. *Journal of Arid environments*, 61(3), 483-496.

Li, X., Zhou, W., & Chen, Y. D. (2015). Assessment of regional drought trend and risk over China: A

drought climate division perspective. *Journal of Climate*, 28(18), 7025-7037.

Li, J., Zhou, S., & Hu, R. (2016). Hydrological drought class transition using SPI and SRI time series by loglinear regression. *Water Resources Management*, 30(2), 669–684.

Livada, I., & Assimakopoulos, V. D. (2007). Spatial and temporal analysis of drought in Greece using the Standardized Precipitation Index (SPI). *Theoretical and Applied Climatology*, 89(3–4), 143–153.

Lohani, V. K., Loganathan, G. V, & Mostaghimi, S. (1998). Long-term analysis and short-term forecasting of dry spells by Palmer Drought Severity Index. *Hydrology Research*, 29(1), 21–40.

Lorenz, D J, Otkin, J. A., Svoboda, M., Hain, C. R., & Zhong, Y. (2018). Forecasting rapid drought intensification using the Climate Forecast System (CFS). J*ournal of Geophysical Research: Atmospheres*, 123(16), 8365–8373.

Lorenz, David J, Otkin, J. A., Svoboda, M., Hain, C. R., Anderson, M. C., & Zhong, Y. (2017). Predicting the US Drought Monitor using precipitation, soil moisture, and evapotranspiration anomalies. Part II: Intraseasonal drought intensification forecasts. *Journal of Hydrometeorology*, 18(7), 1963–1982.

Liu, W. T., & Juárez, R. N. (2001). ENSO drought onset prediction in northeast Brazil using NDVI. *International Journal of Remote Sensing*, 22(17), 3483-3501.

Ma, C., Li, S., Wang, A., Yang, J., & Chen, G. (2019). Altimeter Observation-Based Eddy Nowcasting Using an Improved Conv-LSTM Network. *Remote Sensing*, 11(7), 783.

Mateo-García, G., Adsuara, J. E., Pérez-Suay, A., & Gómez-Chova, L. (2019, July). Convolutional Long Short-Term Memory Network for Multitemporal Cloud Detection Over Landmarks. *In IGARSS 2019-2019 IEEE International Geoscience and Remote Sensing Symposium* (pp. 210-213). IEEE.

McKee, T. B., Doesken, N. J., Kleist, J., & others. (1993). The relationship of drought frequency and duration to time scales. *Proceedings of the 8th Conference on Applied Climatology*, 17(22), 179–183.

McLaren, K., McIntyre, K., & Prospere, K. (2019). Using the random forest algorithm to integrate hydroacoustic data with satellite images to improve the mapping of shallow nearshore benthic features in a marine protected area in Jamaica. *GIScience & Remote Sensing*, 56(7), 1065–1092.

Meng, L., Ford, T., & Guo, Y. (2017). Logistic regression analysis of drought persistence in East China. *International Journal of Climatology*, 37(3), 1444–1455.

Mishra, A. K., & Desai, V. R. (2005). Drought forecasting using stochastic models. *Stochastic Environmental Research and Risk Assessment*, 19(5), 326-339.

Mishra, A. K., & Singh, V. P. (2011). Drought modeling–A review. *Journal of Hydrology*, 403(1-2), 157-175.

Modarres, R. (2007). Streamflow drought time series forecasting. *Stochastic Environmental Research and Risk Assessment*, 21(3), 223–233.

Morid, S., Smakhtin, V., & Moghaddasi, M. (2006). Comparison of seven meteorological indices for drought monitoring in Iran. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 26(7), 971-985.

Morid, S., Smakhtin, V., & Bagherzadeh, K. (2007). Drought forecasting using artificial neural networks and time series of drought indices. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 27(15), 2103–2111.

Mu, B., Peng, C., Yuan, S., & Chen, L. (2019). ENSO Forecasting over Multiple Time Horizons Using ConvLSTM Network and Rolling Mechanism. *2019 International Joint Conference on Neural Networks (IJCNN),* 1–8.

Nouri, M., & Homaee, M. (2018). On modeling reference crop evapotranspiration under lack of reliable data over Iran. *Journal of Hydrology*, 566, 705-718.

NOAA National Centers for Environmental Information (NCEI) U.S. Billion-Dollar Weather and Climate Disasters. Available online: https://www.ncdc.noaa.gov/billions/ (accessed on 06 August 2020). (n.d.). Retrieved 22/10/18 from https://www. ncdc. noaa. gov/billions. https://doi.org/10.25921/stkw-7w73

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.

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, T. N., & Anderson, D. L. T. (1994). The prospects for seasonal forecasting—A review paper. *Quarterly Journal of the Royal Meteorological Society*, 120(518), 755-793.

Park, H., & Kim, K. (2019). Prediction of severe drought area based on random forest: Using satellite image and topography data. *Water*, 11(4), 705.

Park, S., Kim, D. J., Lee, S. W., Lee, K. W., Kim, J., Song, E. J., & Seo, K. H. (2017). Comparison of extended medium-range forecast skill between KMA ensemble, ocean coupled ensemble, and GloSea5. *Asia-Pacific Journal of Atmospheric Sciences*, 53(3), 393-401.

Park, S., Shin, M., Im, J., Chang-Keun, S., Choi, M., Kim, J., ... & Kim, S. K. (2019). Estimation of ground-level particulate matter concentrations through the synergistic use of satellite observations and process-based models over South Korea. *Atmospheric Chemistry and Physics*, 19(2), 1097-1113.

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.

Park, S., Im, J., Park, S., & Rhee, J. (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.

Park, S., Seo, E., Kang, D., Im, J., & Lee, M. I. (2018). Prediction of drought on pentad scale using remote sensing data and MJO index through random forest over East Asia. *Remote Sensing*, 10(11), 1811.

Park, S., Park, S., Im, J., Rhee, J., Shin, J., & Park, J. D. (2017). Downscaling gldas soil moisture data in east asia through fusion of multi-sensors by optimizing modified regression trees. *Water*, 9(5), 332.

Park, S., Im, J., Han, D., & Rhee, J. (2020). Short-Term Forecasting of Satellite-Based Drought Indices Using Their Temporal Patterns and Numerical Model Output. *Remote Sensing*, 12(21), 3499.

Park, S., Park, H., Im, J., Yoo, C., Rhee, J., Lee, B., & Kwon, C. (2019). Delineation of high resolution climate regions over the Korean Peninsula using machine learning approaches. *PloS one*, 14(10), e0223362.

Petrou, Z. I., & Tian, Y. (2019). Prediction of Sea Ice Motion With Convolutional Long Short-Term Memory Networks. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9), 6865–6876.

Reynolds, R. W., Rayner, N. A., Smith, T. M., Stokes, D. C., & Wang, W. (2002). An improved in situ and satellite SST analysis for climate. *Journal of climate*, 15(13), 1609-1625.

Reynolds, R. W., & Banzon, V. F. (2008). NOAA Optimum Interpolation 1/4 Degree Daily Sea Surface Temperature (OISST) Analysis, Version 2. NOAA National Centers for Environmental Information. doi,

10, V5SQ8XB5.

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.

Rhee, J., Im, J., & 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.

Sheffield, J., Wood, E. F., Chaney, N., Guan, K., Sadri, S., Yuan, X., Olang, L., Amani, A., Ali, A., Demuth, S., & others. (2014). A drought monitoring and forecasting system for sub-Sahara African water resources and food security. *Bulletin of the American Meteorological Society*, 95(6), 861–882.

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., & Woo, W. (2015). Convolutional LSTM network: A machine learning approach for precipitation nowcasting. *Advances in Neural Information Processing Systems*, 802–810.

Shukla, R. P., & Shin, C. S. (2020). Distinguishing spread among ensemble members between drought and flood Indian summer monsoon years in the Past 58 Years (1958–2015) reforecasts. *Geophysical Research Letters*, 47(4), e2019GL086586.

Son, B., Park, S., Im, J., Park, S., Ke, Y., & Quackenbush, L. J. (2021). A new drought monitoring approach: Vector Projection Analysis (VPA). *Remote Sensing of Environment*, 252, 112145.

Son, H. Y., Park, J. Y., Kug, J. S., Yoo, J., & Kim, C. H. (2014). Winter precipitation variability over Korean Peninsula associated with ENSO. *Climate dynamics*, 42(11-12), 3171-3186.

Song, A., Choi, J., Han, Y., & Kim, Y. (2018). Change detection in hyperspectral images using recurrent 3D fully convolutional networks. *Remote Sensing*, 10(11), 1827.

Steinemann, A. C. (2006). Using climate forecasts for drought management. *Journal of Applied Meteorology and Climatology*, 45(10), 1353–1361.

Sun, H., Zhao, X., Chen, Y., Gong, A., & Yang, J. (2013). A new agricultural drought monitoring index combining MODIS NDWI and day–night land surface temperatures: A case study in China. *International Journal of Remote Sensing*, 34(24), 8986-9001.

Tadesse, T., Champagne, C., Wardlow, B. D., Hadwen, T. A., Brown, J. F., Demisse, G. B., Bayissa, Y. A., & Davidson, A. M. (2017). Building the vegetation drought response index for Canada (VegDRI-Canada) to monitor agricultural drought: First results. *GIScience & Remote Sensing*, 54(2), 230–257.

Tan, M. L., Tan, K. C., Chua, V. P., & Chan, N. W. (2017). Evaluation of TRMM product for monitoring drought in the Kelantan River Basin, Malaysia. *Water*, 9(1), 57.

Thomas, B. F., Famiglietti, J. S., Landerer, F. W., Wiese, D. N., Molotch, N. P., & Argus, D. F. (2017). GRACE groundwater drought index: Evaluation of California Central Valley groundwater drought. *Remote Sensing of Environment*, 198, 384-392.

Tu, X., Wu, H., Singh, V. P., Chen, X., Lin, K., & Xie, Y. (2018). Multivariate design of socioeconomic drought and impact of water reservoirs. *Journal of Hydrology*, 566, 192-204.

Vogt, J. V, Viau, A. A., & Paquet, F. (1997). Mapping regional air temperature fields using satellite-derived surface skin temperatures. International Journal of Climatology: *A Journal of the Royal Meteorological Society*, 17(14), 1559–1579.

Wen, C., Liu, S., Yao, X., Peng, L., Li, X., Hu, Y., & Chi, T. (2019). A novel spatiotemporal convolutional long short-term neural network for air pollution prediction. *Science of the Total Environment*, 654, 1091–1099.

West, H., Quinn, N., & Horswell, M. (2019). Remote sensing for drought monitoring & impact assessment: Progress, past challenges and future opportunities. *Remote Sensing of Environment*, 232, 111291.

Wu, J., Zhou, L., Liu, M., Zhang, J., Leng, S., & Diao, C. (2013). Establishing and assessing the Integrated Surface Drought Index (ISDI) for agricultural drought monitoring in mid-eastern China. *International Journal of Applied Earth Observation and Geoinformation*, 23, 397-410.

Yan, H., Moradkhani, H., & Zarekarizi, M. (2017). A probabilistic drought forecasting framework: A combined dynamical and statistical approach. *Journal of Hydrology*, 548, 291–304.

Ye, W., Cheng, J., Yang, F., & Xu, Y. (2019). Two-Stream Convolutional Network for Improving Activity Recognition Using Convolutional Long Short-Term Memory Networks. *IEEE Access*, 7, 67772–67780.

Yoo, C., Han, D., Im, J., & Bechtel, B. (2019). Comparison between convolutional neural networks and random forest for local climate zone classification in mega urban areas using Landsat images. I*SPRS Journal of Photogrammetry and Remote Sensing,* 157, 155-170.

Yoon, S., & Lee, T. (2016). Investigation of hydrological variability in the Korean Peninsula with the ENSO teleconnections. *Proceedings of the International Association of Hydrological Sciences*, 374, 165.

Zargar, A., Sadiq, R., Naser, B., & Khan, F. I. (2011). A review of drought indices. *Environmental Reviews*, 19(NA), 333–349.

Zhao, L., Liu, H., Hu, Y., Cheng, H., & Xiao, Z. (2020). Extratropical extended-range precursors near the tropopause preceding persistent strong precipitation in South China: a climatology. *Climate Dynamics*, 55(11), 3133-3150.

Zhang, R., Chen, Z. Y., Xu, L. J., & Ou, C. Q. (2019). Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, China. *Science of The Total Environment*, 665, 338-346.

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.

Zhang, Q., Zhang, J., & Wang, C. (2017). Risk assessment of drought disaster in typical area of corn cultivation in China. *Theoretical and Applied Climatology*, 128(3-4), 533-540.

Zeiler, M. D., & Fergus, R. (2014, September). Visualizing and understanding convolutional networks. In European conference on computer vision (pp. 818-833). Springer, Cham.

Zhu, Q., Luo, Y., Zhou, D., Xu, Y.-P., Wang, G., & Gao, H. (2019). Drought monitoring utility using satellite-based precipitation products over the Xiang River Basin in China. *Remote Sensing*, 11(12), 1483.

Acknowledgement

대학원 면접보던 때가 엊그제 같은데 벌써 6 년이라는 시간이 흘러 졸업을 하게 되었습니다. 아직 배워야할 점이 많지만 이번 졸업이 더 좋은 연구를 하기 위한 한 걸음이라고 생각합니다. 믿고 응원해주신 가족, 지도교수님과 네 분의 교수님, 우리 IRIS 연구실 식구, 그리고 친구들을 포함하여 긴 학위 기간 동안 도움을 주신 모든 분들께 감사드립니다.