Development of earth observational diagnostic drought prediction model for regional error calibration: A case study on agricultural drought in Kyrgyzstan

Eunbeen Park, Hyun-Woo Jo, Woo-Kyun Lee, Sujong Lee, Cholho Song, Halim Lee, Sugyeong Park, Whijin Kim and Tae-Hyung Kim

*Department of Environmental Science and Ecological Engineering, Korea University, Seoul, Republic of Korea; †OJEong Resilience Institute (OJERI), Korea University, Seoul, Republic of Korea; ‡Information and Communications Technology and Disaster Risk Reduction Division (IDD), United Nations Economic and Social Commission for Asia and the Pacific (UNESCAP), Bangkok, Thailand

ABSTRACT

Drought is a natural disaster that occurs globally and is a main trigger of secondary environmental and socio-economic damages, such as food insecurity, land degradation, and sand-dust storms. As climate change is being accelerated by human activities and environmental changes, both the severity and uncertainties of drought are increasing. In this study, a diagnostic drought prediction model (DDPM) was developed to reduce the uncertainties caused by environmental diversity at the regional level in Kyrgyzstan, by predicting drought with meteorological forecasts and satellite image diagnosis. The DDPM starts with applying a prognostic drought prediction model (PDPM) to 1) estimate future agricultural drought by explaining its relationship with the standardized precipitation index (SPI), an accumulated precipitation anomaly, and 2) compensate for regional variances, which were not reflected sufficiently in the PDPM, by taking advantage of preciseness in the time-series vegetation condition index (VCI), a satellite-based index representing land surface conditions. Comparing the prediction results with the monitored VCI from June to August, it was found that the DDPM outperformed the PDPM, which exploits only meteorological data, in both spatiotemporal and spatial accuracy. In particular, for June to August, respectively, the results of the DDPM (coefficient of determination \( R^2 = 0.27, 0.36, \) and 0.4) root mean squared error \( \text{RMSE} = 0.16, 0.13, \) and 0.13 were more effective in explaining the spatial details of drought severity on a regional scale than those of the PDPM \( R^2 = 0.09, 0.10, \) and 0.11; \( \text{RMSE} = 0.17, 0.15, \) and 0.16. The DDPM revealed the possibility of advanced drought assessment by integrating the earth observation big data comprising meteorological and satellite data. In particular, the advantage of data fusion is expected to be maximized in areas with high land surface heterogeneity or sparse weather stations by providing observational feedback to the PDPM. This research is anticipated to support policymakers and technical officials in establishing effective policies, action plans, and disaster early warning systems to reduce disaster risk and prevent environmental and socio-economic damage.

1. Introduction

Drought, a natural disaster that causes environmental and socio-economic damage, maximizes its impact by imperceptibly progressing over a wide area (Hazaymeh and Hassan 2016; Mottaleb et al. 2015). In addition to the difficulties in drought specification, the uncertainty of drought prediction is becoming more severe owing to climate change (He et al. 2013). Thus, enhancing the adaptive capacity to drought according to each country’s circumstances is essential, and the predictive performance is a crucial factor (Raikes et al. 2019). In particular, some countries in arid and semi-arid regions that lack adaptive capacity, such as Kyrgyzstan, suffer from continuous drought damage in the agricultural sector (Kogan, Guo, and Yang 2019; Schwabe et al. 2013). The agriculture and food security sectors are considered to be top policy priorities because they are the most vulnerable to climate change (FAO 2018). Particularly, the occurrence of severe agricultural drought in the growing season (mostly June to August in Kyrgyzstan) is directly linked to declining crop yields (Frenken 2013; World Food Programme 2014).

CONTACT Woo-Kyun Lee leewk@korea.ac.kr

Current address: Institute for Integrated Management of Material Fluxes and of Resources (UNU-FLORES), Ammonstrasse 74, Dresden 01067, Germany (H. L.) Prediction Research Department, APEC Climate Center, Haeundae-gu, Busan 48058, Republic of Korea (S. P.)

These authors contributed equally

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
To manage the threat caused by drought and to improve national food security, earth observation data, various algorithm, and data fusion systems should be examined to develop drought prediction models with higher accuracy and precision, so that they can be used to prioritize drought-related responses (Diaz et al. 2020; Gu and Ning 2012). The existing models can be subdivided into two types based on data characteristics, which is one of the most intuitive and performance-affecting criteria proposed by Maisongrande et al. (1997). The first type, a prognostic model, is applicable to predict future disasters based only on climate data as independent variables (Ito and Oikawa 2002; Sasai et al. 2005). The second type, a diagnostic model, can more realistically simulate disasters by considering both land surface and climate conditions (Kim et al. 2012; Sasai et al. 2007).

The prognostic model uses historical climate data to clarify the interlinkage between precipitation deficiency and drought and then predicts drought with future climate data (Aghakouchak and Nakhjiri 2012; Thenkabail 2015). In previous studies, most researchers identified the spatio-temporal changes (Komuscu 1999; Li et al. 2017) in several meteorological drought indices (Nyatuame and Agodzo 2017; Park 2017; Shen et al. 2019) and applied various models, such as machine learning techniques and regression models, to predict drought (Belayneh and Adamowski 2012; Belayneh et al. 2014; Mesbahzadeh et al. 2020).

However, since these studies relied only on climate patterns, they could neither explain how the impacts of meteorological droughts have manifested at ground level as agricultural drought nor reflect regional differences such as ecosystem, infrastructure, and management (Hao et al. 2014; Sheffield et al. 2004; Thenkabail 2015). In the same context, because the predictions were based on historical climate patterns, the effect of dynamic changes on the land surface could not be reflected in the drought prediction. In addition, the performance of the prognostic model is likely to be decreased in countries with sparse weather station networks; the model must rely on estimated meteorological data rather than measured data. This process leads to a dilemma in which the low adaptive capacity prevents itself from being enhanced (Kogan, Guo, and Yang 2019; Wardlow et al. 2017).

In order to overcome the limitation of the prognostic model, diagnostic models make predictions based on not only climate but fusing it with land surface observations, which are mostly acquired by satellite images. Satellite-based monitoring provides actual land surface information, which implies the integrated effect of climate, regional environment and complex drought mechanisms on the surface and thus enables the precise tracking of global long-term environmental changes over large inaccessible areas (Barrett et al. 2020). Most general approaches for fusing meteorological and satellite data use statistical and deep-learning models with the best-fitting parameters (AghaKouchak 2015; Mokhtari and Akhoondzadeh 2020), providing a fast and convenient way to increase predictive performance. However, the simple amalgamation of the data limits the exploitation of the characteristics of each dataset, which is a mandatory consideration for developing data fusion model (Morabito, Simone, and Cacciola 2008). Some of the other approaches have attempted to retain the characteristics of each dataset under the concept of drought phases (del Pilar Jiménez-Donaire, Tarquis, and Giraldez 2019). However, these studies still comprise basic algorithms such as thresholding and overlaying. Also, despite they have applied domain knowledge to diagnostic models exploiting the semantic feature of each data, they did not fully utilize the structural features of the data; for instance, satellite imaging provides a precise snapshot of the current surface, whereas climate data is forecastable. Therefore, there remains a need to develop a data fusion model to conduct accurate and precise agricultural drought prediction that reflects both the semantic and structural features of data and drought conditions on the land surface under predicted climate conditions (Karnieli et al. 2010; Mishra and Desai 2005).

Hence, in this study, a diagnostic drought prediction model (DDPM) was developed by proposing a novel data fusion method for the synergistic use of each earth observation data in the way of calibrating regional error of meteorological data with satellite images. The DDPM was examined to predict agricultural drought in Kyrgyzstan where the existing drought prediction models are struggling to return precise prediction and active discussions are ongoing by the national government.
to exploit land observations. In the DDPM progress, the prognostic drought prediction model (PDPM) is first performed to exploit the advantage of climate data predictability, and then compensates for the prediction error by considering the current land surface conditions monitored on satellite images.

2. Materials

2.1. Study area

Kyrgyzstan is a landlocked country located in Central Asia (39.146°–43.22° N, 69.297°–80.132° E; Figure 1). The country is dominated by mountains with permanently covered ice and snow where the average elevation of its territory is 2750 m, ranging from 401 m to 7439 m (Dzunusova, et al., 2016). The average annual precipitation is approximately 533 mm, and most of it is concentrated during the cold winter season (between October and April) (Frenken 2013). The absolute temperature greatly varies (from –18 to 28°C in winter to 43°C in summer) depending on the location.

Although irrigation systems are facilitated in more than 80% of cropland in Kyrgyzstan, they have deteriorated in quality since the 1990s, decreasing the water delivery efficiency (Asian Development Bank 2013). Therefore, according to the maintenance of the irrigation system and adjacency of the cropland to the supplement, drought severity can be more or less dependent on precipitation.

2.2. Meteorological data

Climate Hazards Center InFraRed Precipitation with Station (CHIRPS) was developed by the Climate Hazards Group (https://www.chc.ucsb.edu/data) at the University of California Santa Barbara (UCSB) and the U.S. Geological Survey (USGS). CHIRPS combines satellite imagery with an in-situ station dataset and provides gridded rainfall time-series data from 1981 to the present with a spatial resolution of 0.05° (Funk et al. 2015). The CHIRPS data have the advantage of high spatial-temporal resolution among estimated precipitation datasets, such as the Tropical Rainfall Measuring Mission (TRIMM), which has a 0.25° resolution from 1998–2019, and the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks–Climate Data Record (PERSIANN-CDR), which has a 0.25° resolution from 1983 to the present. In this study, monthly precipitation data from January 1984 to December 2018 were obtained from CHIRPS (version 2.0) to produce a meteorological drought index.

In this study, the standardized precipitation index (SPI), a meteorological drought index, was selected to predict drought conditions based on climate data. The SPI was developed by McKee, Doesken, and Kleist (1993) to reflect the impact of precipitation deficiency and identify precipitation shortages by comparing the total precipitation received during a certain period at a particular location with long-term rainfall distribution. Generally, to calculate the SPI, the cumulative sum of precipitation over a certain period is converted to cumulative probability under

![Figure 1. Map of the administrative districts of Kyrgyzstan and the spatial distribution of agricultural land.](image-url)
a selected gamma distribution function, and the inverse probability distribution function of a standard normal distribution is applied to the probability. In drought monitoring studies, the SPI is usually calculated using monthly precipitation data with time scales of 3, 6, 12, and 24 months. In addition, the SPI requires a minimum of 30 years of high-quality precipitation data to ensure statistical significance. The shorter SPI runs (e.g., SPI 3 and SPI 6) refer to meteorological drought and soil moisture, and the longer runs (e.g., SPI 12 and SPI 24) refer to hydrological drought (Svoboda, et al., 2012). In this study, a three-month SPI (SPI 3), which is closely related to agricultural impact, was selected and calculated from 1984 to 2018 (del Pilar Jiménez-donaire, Tarquis, and Giráldez 2019). In addition, the gamma probability density function (Equation 1) that best describes the precipitation distribution was fitted to a time-series precipitation, and the cumulative probability was converted to a Z-score of the standard normal. In order to integrate with satellite images, we selected the target period for drought analysis as 2000–2018. Therefore, Part of the SPI 3 dataset from 2000 to 2018 was utilized for drought prediction. The wet conditions are indicated as SPI values greater than 1.0, while dry conditions are indicated as SPI values less than −1.0.

\[
G(x) = \int_{0}^{x} g(x) \, dx = \frac{1}{\beta^\alpha \Gamma(\alpha)} \int_{0}^{x} x^{\alpha-1} \, e^{-\frac{x}{\beta}} \, dx, \tag{1}
\]

where \( g(x) \) is a gamma probability density function, \( \alpha > 0 \) is a shape parameter, \( \beta > 0 \) is a scale parameter, \( x \) is the amount of precipitation, and \( \Gamma(\alpha) \) is the gamma function.

### 2.3. Land observation data

The normalized difference vegetation index (NDVI) time-series dataset was acquired from the Terra Moderate Resolution Imaging Spectroradiometer (MODIS) vegetation indices (MOD13A3) Version 6 (006) of the National Aeronautics and Space Administration Earth Observing System (https://eospso.gsfc.nasa.gov/mission-category/3). MOD13A3 Version 6 (006) has provided monthly level 3 data at 1 km spatial resolution from 2000 to the present (available online at: https://ladsweb.modaps.eosdis.nasa.gov/). In this study, the monthly MOD13A3 NDVI (h23v04 and h23v05) from March 2000 to December 2018 was utilized to generate a satellite-based agricultural index by monthly time step. Data with cloud cover and missing data were excluded from the MODIS quality assurance data.

In this study, the vegetation condition index (VCI), a satellite-based agricultural drought index, was selected to represent the actual drought conditions on the land surface. Kogan (1995) proposed a VCI based on relative NDVI changes for agricultural drought detection and tracking. This remote sensing-based agricultural drought index has been used to detect and monitor agricultural drought and is calculated using the time-series of NDVI by normalizing it between the maximum and minimum values within the same pixel (Equation 2). Therefore, the VCI contains both the current status and historical information of the NDVI (Xu et al. 2020) and represents the relative drought severity in the temporal dimension. A VCI value of 0 indicates dry conditions, whereas a value of 1 indicates wet conditions.

\[
VCI_i = \frac{NDVI_i - NDVI_{\text{min,}_i}}{NDVI_{\text{max,}_i} - NDVI_{\text{min,}_i}}, \tag{2}
\]

where \( VCI_i \) is the VCI at time-series \( i \), \( NDVI_i \) is the NDVI obtained at time-series \( i \), and \( NDVI_{\text{min,}_i} \) and \( NDVI_{\text{max,}_i} \) are the minimum and maximum NDVI computed within the same month at each pixel, respectively.

### 3. Methods

In a regression model, which has relatively simple structures, selecting the most persuasive independent variables is one of the most effective methods and should be prioritized for reducing fitting error (residuals). However, it is almost impossible to include all of the relevant independent variables that affect agricultural drought during the modeling process, and the omissions remain as the fitting error. As an alternative, the fitting error can themselves be used as independent variables instead of adding one (Hengl, Heuvelink, and Rossiter 2007; Piao et al. 2018). For example, regression-kriging uses spatially interpolated fitting error as independent variables by assuming that unknown independent variables remaining as the fitting error have spatial autocorrelation (Hengl, Heuvelink, and Rossiter 2007).
Assuming that the impacts of environmental factors that are not used as independent variables remain as temporally autocorrelated fitting error, the diagnostic prediction model predicts future fitting error based on its temporally autocorrelated characteristics. Subsequently, diagnostic prediction is performed by compensating for the estimated future fitting error in the prognostic prediction result. The process of diagnostic modeling, composed of prognostic model design and future fitting error estimation, can vary depending on the prediction target. However, the underlying principle is to develop a model with highly foreseeable variables and supplement it with time-series monitoring. In this study, based on the above concepts, the diagnostic prediction model was applied to agricultural drought prediction as a case study in Kyrgyzstan (Figure 2).

In the first stage of the research, the PDPM was developed to predict agricultural drought based on a meteorological drought index. In this study, VCI indicated agricultural drought, and SPI 3 was used as a meteorological drought index. In the PDPM, the agricultural drought index is a prediction target (it directly affects the agricultural sector), whereas the meteorological drought index is an input (it is a preface of disasters but more foreseeable). Based on the historical relationship between the two indices, the PDPM can return the agricultural drought index whether the input is observed or forecasted.

The DDPM was developed by exploiting the ratio between the monitored and predicted VCIs, expressed as the M/P ratio. The M/P ratio indicates the integrated effect of the unused variables in the PDPM. Assuming that the current M/P ratio is the best predictor of the M/P ratio in the near future, the M/P ratio is multiplied by the later time-series of the PDPM prediction, resulting in the DDPM prediction. In this approach, the present and future are relative concepts; the SPI that is used as the future value at a certain time is then reused as the present value at the next time step. Therefore, the time at which the M/P ratio was calculated is referred to as the present, and the time at which the PDPM was calibrated with the M/P ratio is referred to as the future. The process assumes the situation of predicting the future based on the most recent satellite observations, and the uncertainties of meteorological forecasts that occur in real cases are disregarded. After all of the future time-series were predicted, the results of the PDPM and DDPM were compared to those of the monitored agricultural drought index.

3.1. PDPM

The purpose of the PDPM is to predict agricultural drought represented by the VCI using meteorological drought represented by the SPI 3. To understand the general relationship between the two indices only in areas exposed to agricultural drought, the mean values of the monthly SPI 3 and VCI were acquired at a national scale in agricultural areas. The PDPM was designed as a nonlinear model because many crop and soil processes have nonlinear features (Archontoulis and Miguez 2015). Because both indices are calculated through relative distribution within a specific time period, the PDPM was developed at each month, especially in the major vegetation growing season (May to August) in Kyrgyzstan. The model explains the relationship between the monthly SPI 3 and monthly VCI with three parameters (a, b, and c), which were set by regression analysis between the indices each month. Through this model, a certain time i of the satellite-based agricultural drought index (VCI) is explained by the same time i of the meteorological drought index (SPI 3) (Equation 3). The model returns zero, the minimum possible value of VCI, when the sum of the SPI 3 and parameter b is negative, and e indicates the error term.

$$VCI_i = f_{prog}(SPI_3) + e_i = \begin{cases} SPI_3 + b > 0 & \alphaSPI_3 + b + e_i \\ SPI_3 + b < 0 & 0 + e_i \end{cases}$$

3.2. DDPM

The DDPM aims to minimize the uncertainty of the PDPM by utilizing the current fitting error to minimize future fitting error, as shown in Figure 3. The PDPM predicts future drought; the model is trained based on the relationship between the past SPI 3 and VCI data. However, the PDPM inevitably includes fitting error, as described in Equation 3, because not all environmental conditions – other than meteorology – are considered independent variables. Therefore, the actual
Agricultural drought could be more severe than that indicated by the PDPM result when the excluded surface-level environmental factors (such as geography, infrastructure, and agricultural activities) have a negative impact on drought, and the reverse impact is also possible. In this context, the DDPM diagnoses the current environment to estimate the impact of the excluded variables by calculating the M/P ratio. The diagnosed impact is then applied to the PDPM result of the next time step under the assumption that most environmental factors are temporally autocorrelated (Shahin et al. 2014; Shamsnia et al. 2011). In the actual application,
the PDPM of the next time step should use forecasted meteorology, but it has been replaced by observed meteorology in the future to separate the performance of the proposed models and the uncertainties in the forecast process.

Under the concept of the diagnostic prediction model proposed in this study, the process of diagnosing the current environment and applying it to the next time step can vary according to the characteristics of the prediction target. To develop the diagnostic process with regard to drought prediction, the DDPM is derived from Equation 3 by dividing the equation of time \( i + 1 \) with that of time \( i \) (Equation 4). Equation 5 indicates that the DDPM calculates the M/P ratio and multiplies it with the next time step of the PDPM result. In this case, the fitting error of the DDPM is expressed as \( e_{i+1} \) subjected to the term \( e_i \), while the fitting error of the PDPM is only \( e_{i+1} \). Therefore, the relative advantage of the DDPM over the PDPM in drought prediction is verified by the fact that \( e_{i+1} \) is close to the term \( e_i \) (Equation 6). The similarity of the two values can be explained by the aforementioned temporal autocorrelation of environmental phenomena because the consecutive fitting error are the result of unused environmental factors. Furthermore, the term \( e_p \), which is proportional to the function of SPI 3, can be explained by water being a limiting factor for vegetation photosynthesis (Chang et al. 2018; Karnieli et al. 2010). Even though agricultural drought is a result of complex interactions between precipitation and various environmental factors, limited precipitation is a sufficient condition to promote vegetation deterioration, which results in a magnitude of fitting error variation proportional to the function of SPI 3 (Figure 4).

\[
\frac{VCl_{i+1}}{VCl_i} = \frac{f_{prog}(SPI 3_{i+1}) + e_{i+1}}{f_{prog}(SPI 3_i) + e_i}, \quad (4)
\]

\[
VCl_{i+1} = f_{diag}(SPI 3_i, VCl_i, SPI 3_{i+1}) + e_{i+1} - \frac{f_{prog}(SPI 3_{i+1})}{f_{prog}(SPI 3_i)} \frac{e_i}{f_{prog}(SPI 3_i)}, \quad (5)
\]

\[
e_{i+1} \approx e_i \frac{f_{prog}(SPI 3_{i+1})}{f_{prog}(SPI 3_i)}, \quad (6)
\]
3.3. Performance assessment in spatio-temporal dimension

In order to validate the performance of the proposed drought prediction models, it is important to consider the accuracy in both the spatial and temporal dimensions because the model should be able to differentiate between the drought severities in each area and time. Therefore, in this study, the monthly VCI predicted by the PDPM and DDPM was converted to a mean value of 100 km$^2$ grid units and compared to the monitored VCI values. The VCI was averaged over a 100 km$^2$ grid to understand the trend of spatial agreement between the prediction and observation; comparing the drought index at too small scale, such as at the pixel level, would ignore the overall pattern but only compare between the exact same locations, leading to an underestimation of the model performance. In contrast, comparing the drought indices at too large scale cannot estimate the spatial agreement or would do so roughly.

The model performance was estimated by comparing the monthly root mean squared error (RMSE) and coefficient of determination ($R^2$) for each model. To estimate the overall performance in both spatial and temporal dimensions, the RMSE and $R^2$ were calculated once using the spatially

![Figure 4](image-url). Scatter plot between the prognostic function of a three-month standardized precipitation index (SPI 3) and its fitting error.

![Figure 5](image-url). Model evaluation in the spatial and temporal dimensions.
averaged value for each 100 km$^2$ grid over the entire study period. In this case, since the difference of VCI between each year is much greater than that between each area, the evaluation indicators are more likely to be affected by the temporal dimension. Therefore, to estimate the performance with an emphasis on the spatial dimension, the indicators were calculated again at each time step from 2000 to 2018 and averaged (Figure 5).

4. Results

4.1. SPI 3 and VCI in Kyrgyzstan

Figure 6 shows the monthly SPI 3 and VCI data calculated from 2000 to 2018 in Kyrgyzstan. According to the results, severe droughts were recorded during the growing seasons of 2006, 2008, and 2014 in both the SPI 3 and VCI. In addition, the overall patterns of the monthly SPI 3 and VCI showed a similar tendency over the entire study period. These results indicate that the meteorological drought can directly affect agricultural drought.

4.2. PDPM and DDPM

The relationship between the monthly SPI 3 and VCI was investigated using the PDPM, as described in Equation 3. Through PDPM, the monthly VCI was explained by the SPI 3 at the national level through the nonlinear model composed of three coefficients for each month of the growing season. According to the results, the $R^2$ values of the PDPM for May, June, July, and August were 0.526, 0.764, 0.597, and 0.670, respectively (Figure 7).

According to the DDPM (Equation 5), the impact of current surface environments on future drought prediction was diagnosed, and the results of the PDPM were calibrated. Figures 8 and 9 show the time-series results of the PDPM and DDPM in representative years of the normal (2002) and severe (2008) drought condition. The figures represent the (a) SPI 3 in July, (b) SPI in August, (c) PDPM VCI in July, (d) PDPM VCI in August, (e) monitored VCI in July, (f) monitored VCI in August, and (g) DDPM VCI in August. Since the PDPM was designed based on the relationship between the SPI 3 and VCI, the VCIs predicted by the PDPM had similar patterns to those of the SPI 3s. However, the results of the PDPM occasionally had different patterns to those of the monitored VCIs, such as those in the western and central regions (Jalal-Abad, Osh, and Naryn; Figure 8) and western and northern regions (Jalal-Abad, Osh, and Bishkek; Figure 9). The difference between the PDPM result and the actual drought in July indicates the effect of environmental factors other than the SPI 3; thus, the difference ratio of July was multiplied to the PDPM of August in both figures to produce the DDPM result. As a result, the DDPM prediction showed a pattern much more similar to that of the monitored VCI than did the PDPM, especially in the aforementioned regions where the PDPM had high fitting error.

4.3. Model evaluation

To verify the relative advantage of using DDPM over PDPM, the model performance was evaluated in both spatial and temporal dimensions in a 100 km$^2$ grid unit from 2000 to 2018 (Figure 10). According to the spatio-temporal evaluation comparing the entire grid and time period at the same time, DDPM had higher

Figure 6. Patterns of meteorological drought (SPI 3) and agricultural drought (VCI) in the period 2000–2018.
and lower RMSE values than PDPM in every growing season, with $R^2$ values of 0.310, 0.482, and 0.561 and RMSE values of 0.161, 0.133, and 0.137 for June to August, respectively, while the PDPM had $R^2$ values of 0.257, 0.305, and 0.327 and RMSE values of 0.171, 0.150, and 0.163, respectively. The results of the spatio-temporal evaluation are depicted as black lines. Both the PDPM and DDPM point in the upper-right direction, indicating that although they both succeed to predict drought in general, the DDPM showed a clearer agreement between the prediction and observation. Through the spatial evaluation performed within the grids of each year, the strength of the DDPM is highlighted as it shows the extent to which the model can differentiate between drought severities among regions. The red lines (representing the spatial evaluation) of the PDPM are almost horizontal, indicating that it hardly differentiates between drought severities at the regional scale. Unlike the trend lines of the PDPM, those of the DDPM show a relatively consistent upper-right direction for both the spatial and spatio-temporal evaluations, which implies that the model can differentiate detailed vegetation conditions not only between each year but also between each area within a year. The same effect of evaluation indicators can also be identified on scatter plots between the monitored and predicted values, which show a dense upper-right directional point cloud for the DDPM results and a sparse declining point cloud for the PDPM results.

5. Discussion

In this study, the VCI prediction results of PDPM and DDPM were compared with monitored VCI values to validate the effect of the DDPM diagnostic process, which complements the PDPM with the latest fitting error information. In the case of the PDPM, a nonlinear regression model was used to explain the VCI using

![Figure 7. Relationship between the three-month standardized precipitation index (SPI 3) and vegetation condition index (VCI) in the prognostic drought prediction model (PDPM).](image-url)
Figure 8. Process for predicting the diagnostic drought prediction model (DDPM) vegetation condition index (VCI) in August 2002, an example year for normal condition, and its intermediate results.
Figure 9. Process for predicting the diagnostic drought prediction model (DDPM) vegetation condition index (VCI) in August 2008, an example year for severe conditions, and its intermediate results.
SPI 3 on a national scale. Even though the PDPM uses only meteorological data, it recorded $R^2$ values of between 0.5 and 0.77 throughout the growing period. This result indicates that meteorological drought is related to agricultural drought and that rainfall deficit is one of the most important factors responsible for drought, which is consistent with the results of previous studies (Haile et al. 2020; Ghazaryan et al. 2020; Phuong et al. 2021). However, when the PDPM was evaluated in 100 km$^2$ grids, the performance was significantly decreased owing to the inability to differentiate between drought severities in different areas within a year. The reason for such limited prediction in the spatial dimension is that even though the PDPM exploits meteorological information that is forecasted using multiple advanced models (Cho et al. 2019), the meteorological impacts can vary because they are linked to environmental circumstances at the ground level, such as topography, ecosystems, man-made structures, and policy management (Feng et al. 2019; Park, Kim, and Lee 2019). Consequently, the smoothed pattern of the PDPM, originating from that of the SPI, did not coincide with the actual agricultural drought, which is diversified at the surface level. In contrast, the drought prediction results obtained by DDPM were similar to the observed agricultural drought conditions on the land surface. The DDPM diagnostic process was designed by considering the drought reaction to SPI 3 to maximize the calibration performance. Under the premise that future fitting error is proportional to the current fitting error, DDPM diagnoses the current fitting error and calibrates the results of the PDPM. Therefore, the most highlighted difference in the DDPM can be observed in areas where the PDPM has high fitting error, such as the
cities of Bishkek and Osh, which are highly populated areas where socio-economic systems, such as irrigation, are relatively well sustained (USAID 2018), and agricultural drought is more affected by factors other than meteorology.

The relative advantage of the DDPM over PDPM, especially in the spatial dimension, is highlighted when the prediction results are used as basic data for disaster reduction. Since decision making is a matter of using limited labor force and commodities to maximize the effect, preparing for a disaster needs an order of priority. As the DDPM can differentiate between the detailed drought severity of different areas, capacity building can be focused on the most vulnerable areas. For example, the DDPM approach could be used effectively in Central Asia in the mid-latitude region, where there is insufficient field data and an urgent drought early warning system is required. The DDPM can exploit satellite images to compensate for relatively imprecise meteorological data derived from sparsely distributed weather stations. Through the combination of drought prediction with drought-related decision making, adaptive systems, and disaster early warning systems, the DDPM could achieve significant risk mitigation, set drought risk management strategies, and help to improve food security (Kogan, Guo, and Yang 2019). Recently, United Nations (UN) organizations, such as the UN Conventions to Combat Desertification and UN Economic and Social Commission for Asia and the Pacific, have been aiming to support Central Asian countries in their efforts to develop drought decision-making systems (United Nations Convention to Combat Desertification (UNCCD) 2021; United Nations Economic and Social Commission for Asia and the Pacific (UNESCAP) 2021), and the DDPM could be one of the possible solutions.

Moreover, the DDPM can be easily improved under the concept of diagnostic prediction models. The premise of diagnostic prediction models is that the fitting error are temporally autocorrelated because they are the result of environmental factors that barely change over short periods. In keeping with this premise, the diagnostic process for estimating the impact of current fitting error on the next prediction can be improved by applying time-series analyses, such as seasonal autoregressive integrated moving average (SARIMA), recurrent neural networks (RNNs), and long short-term memory (LSTM), and by considering drought domain knowledge. In addition, the importance of improving the inputs and design of the PDPM cannot be ignored because the PDPM is not only the basis of the DDPM but also a changing force from the current status. Although the diagnostic process of the DDPM is crucial for achieving accurate and precise prediction results, the DDPM only uses data from the near past, while the PDPM outlines the drought severity based on the weather forecast. Increasing the number of weather stations to acquire more detailed and long-term meteorological information is essential to minimizing the overall uncertainties in the PDPM and thus enhancing the DDPM performance.

6. Conclusion

In this study, a DDPM was developed to predict agricultural drought in Kyrgyzstan by combining a series of earth observation big datasets under the concept of diagnosis that exploits fitting error information. Therefore, the application of DDPM was confirmed to be more effective in improving the spatio-temporal predictive performance than the PDPM. In particular, while the PDPM recorded an acceptable performance in the temporal dimension, for June to August, respectively, the performance of the DDPM ($R^2 = 0.27$, $0.36$, and $0.4$; RMSE = $0.16$, $0.13$, and $0.13$) overwhelmed that of the PDPM ($R^2 = 0.09$, $0.10$, and $0.11$; RMSE = $0.17$, $0.15$, and $0.16$) in the spatial dimension, differentiating between detailed drought severity patterns in different regions. This study is significant in that it presents a new data fusion methodology for predicting future drought that combines the advantages of predictability in climate data and precision in satellite data. As the DDPM is able to compensate for the limitation of sparsely distributed weather stations by integrating satellite images, it is expected to contribute to the drought preparation decision-making system for areas where infrastructure is insufficient. In addition, the concept of diagnostic prediction models can not only be applied to drought prediction – which results in the DDPM – but can also be applied to predict diverse environmental variables with advanced model structures and
diagnostic processes using varied inputs. Therefore, the results of this study will not only contribute to agricultural drought adaptation but also provide more comprehensive support to address environmental and sustainability issues, such as establishing policies and action plans relevant to disaster risk reduction and climate change reactions for policymakers and technical officials. However, when the DDPM is used for climate change reactions, it should be considered that it is a specialized tool for short-term prediction because it relies on temporal autocorrelation between adjacent time-series.

Therefore, enhancing short-term adaptation capacity with the DDPM and preparing long-term adaptation or mitigation plans using climate change scenarios should be conducted at the same time for a more comprehensive reaction. In addition, since the DDPM in this study has been applied only in Kyrgyzstan using past meteorological data, the effectiveness of the DDPM should be verified under more diverse circumstances, such as different temperate zones, sub-monthly intervals, and using forecasted meteorology.

Acknowledgements

This work was inspired by the “Enhancing the capacity of the developing countries in Central Asia on effective use of space applications for drought monitoring and early warning through the Regional Drought Mechanism” project by the United Nations Economic and Social commission for Asia and the Pacific (UNESCAP), International Research & Development Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (2018K1A3A7A03089842) and Korea University Grant.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the International Research and Development Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT [2021K1A3A1A78097879].

ORCID

Eunbeen Park http://orcid.org/0000-0002-0442-7621
Hyun-Woo Jo http://orcid.org/0000-0001-6127-883X
Woo-Kyun Lee http://orcid.org/0000-0002-2188-359X
Sujong Lee http://orcid.org/0000-0002-6383-805X
Cholho Song http://orcid.org/0000-0002-8491-9545
Halim Lee http://orcid.org/0000-0001-5026-9604
Sugyeong Park http://orcid.org/0000-0002-2065-041X
Whijin Kim http://orcid.org/0000-0002-7093-7312
Tae-Hyung Kim http://orcid.org/0000-0002-5808-170X

Data availability statement

The data that support the findings of this study are openly available from “LAADS DAAC-NASA” at http://doi.org/10.5067/MODIS/MOD13A3.006 and “CHIRPS: Rainfall Estimates from Rain Gauge and Satellite Observations” at http://doi.org/10.15780/G2RP4Q and https://www.chc.ucsb.edu/data/chirps.

Author contribution

E.P. and H.W.J. designed the study, analyzed and interpreted the data, and wrote the manuscript. W.K.L. contributed to the methodology and review. S.L., C.S., and S.P. contributed to the software and validation. W.K. and H.L. participated in data investigation. All authors, including T.H.K., provided comments and approved the final manuscript. All authors have read and agreed to the published version of the manuscript.

References

Aghakouchak, A., and N. Nakhjiri. 2012. “A near Real-Time Satellite-Based Global Drought Climate Data Record.” Environmental Research Letters 7 (4): 044037. doi:10.1088/1748-9326/7/4/044037.
AghaKouchak, A. 2015. “A Multivariate Approach for Persistence-Based Drought Prediction: Application to the 2010–2011 East Africa Drought.” Journal of Hydrology 526: 127–135. doi:10.1016/j.jhydrol.2014.09.063.
Archontoulis, S. V., and F. E. Miguez. 2015. “Nonlinear Regression Models and Applications in Agricultural Research.” Agronomy Journal 107 (2): 786–798. doi:10.2134/agronj2012.0506.
Asian Development Bank. 2013. “Kyrgyz Republic: Developing Water Resources Sector Strategies in Central and West Asia.” Vol. 1 (June).
Barrett, S., N. Brooks, N. Quadrianto, S. Anderson, and B. Nebus. 2020. “Measuring Climate Resilience by Linking Shocks to Development Outcomes.” Climate and Development 12 (7): 677–688. doi:10.1080/17565529.2019.1676689.
Belhaj, A., and J. Adamowski. 2012. “Standard Precipitation Index Drought Forecasting Using Neural Networks, Wavelet Neural Networks, and Support Vector Regression.” *Applied Computational Intelligence and Soft Computing* 2012: 1–13. doi:10.11095/2012/794061.

Belhaj, A., J. Adamowski, B. Khalil, and B. Ozga-Zielinski. 2014. “Long-Term SPI Drought Forecasting in the Awash River Basin in Ethiopia Using Wavelet Neural Networks and Wavelet Support Vector Regression Models.” *Journal of Hydrology* 508: 418–429. doi:10.1016/j.jhydrol.2013.10.052.

Chang, K. Y., L. Xu, G. Starr, and K. T. Paw U. 2018. “A Drought Indicator Reflecting Ecosystem Responses to Water Availability: The Normalized Ecosystem Drought Index.” *Agricultural and Forest Meteorology* 250–251: 102–117. doi:10.1016/j.agrformet.2017.12.001.

Cho, J., J. Jung, B.-G. Song, Y. Jung, J. Han, E.-J. Lee, and D. Kang. 2019. “Platform Development for Statistical Downscaling of APCC Seasonal Forecasting and Climate Change Scenarios II.” Republic of Korea: APEC Climate Center.

del Pilar Jiménez-donaire, M., A. Tarquis, and J. V. Giráldez. 2019. “Evaluation of a Combined Drought Indicator and Its Predictive Potential for Agricultural Droughts in Southern Spain.” *Natural Hazards and Earth System Sciences Discussions* 2018: 1–22. doi:10.5194/nhess-2019-135.

Díaz, V., G. A. C. Perez, H. A. J. Van Lanen, D. Solomatine, and E. A. Varouchakis. 2020. “An Approach to Characterise Spatio-Temporal Drought Dynamics.” *Advances in Water Resources* 137 (July 2019): 103512. doi:10.1016/j.advwatres.2020.103512.

Dzunusova, M., Apasov, R., and Mammadov, A. 2016. “Country Report on the State of Plant Genetic Resources for Food and Agriculture.” Rome, Italy: FAO. [https://www.fao.org/3/i5000e/Kyrgyzstan.pdf](https://www.fao.org/3/i5000e/Kyrgyzstan.pdf)

FAO. 2018. “Country Fact Sheet on Food and Agriculture Policy Trends: The Kyrgyz Republic.” *Food and Agriculture Organization of the United Nations*, no. 1, 6.

Feng, P., B. Wang, D. Li Liu, and Q. Yu. 2019. “Machine Learning-Based Integration of Remotely-Sensed Drought Factors Can Improve the Estimation of Agricultural Drought in South-Eastern Australia.” *Agricultural Systems* 173 (December 2018): 303–316. doi:10.1016/j.agsy.2019.03.015.

Frenken, K. 2013. “Irrigation in Central Asia in Figures.” *AQUASTAT Survey*. 2012. FAO. 39. http://www.fao.org/3/i1467-6346.2007.00685x.

Funk, C., P. Peterson, M. Landsfeld, D. Pederson, J. Verdin, S. Shukla, G. Husak, et al. 2015. “The Climate Hazards Infrared Precipitation with Stations - A New Environmental Record for Monitoring Extremes.” *Scientific Data* 2: 150066. doi:10.1038/sdata.2015.66.

Ghazaryan, G., O. Dubovyk, V. Graw, N. Kussul, and J. Schellberg. 2020. “Local-Scale Agricultural Drought Monitoring with Satellite-Based Multi-Sensor Time-Series.” *GIScience and Remote Sensing* 57 (5): 704–718. doi:10.1080/15481603.2020.1778332.

Gu, X., and Ning, L. 2012. “Study of Droughts and Floods Predicting System Based on Spatial-Temporal Data Mining.” *Proceedings of 2012 6th International Conference on New Trends in Information Science, Service Science and Data Mining* (NISS, ICMA and NASNIT), ISSDM 2012, Oct 23, 2012 - Oct 25, 2012. Taipei, Taiwan. IEEE, 248–253.

Halle, G. G., Q. Tang, L. Wen-hong, X. Liu, and X. Zhang. 2020. “Drought: Progress in Broadening Its Understanding.” *WIREs Water* 7 (2): 1–25. doi:10.1002/wat2.1407.

Hao, Z., A. AghaKouchak, N. Nakhjiri, and A. Farahmand. 2014. “Global Integrated Drought Monitoring and Prediction System.” *Scientific Data* 1: 140001. doi:10.1038/sdata.2014.1.

Hazynem, K., and Q. K. Hassan. 2016. “Remote Sensing of Agricultural Drought Monitoring: A State of Art Review.” *AIMS Environmental Science* 3 (4): 604–630. doi:10.3934/enviros.2016.4.604.

He, B., W. Jianjun, A. Lü, X. Cui, L. Zhou, M. Liu, and L. Zhao. 2013. “Quantitative Assessment and Spatial Characteristic Analysis of Agricultural Drought Risk in China.” *Natural Hazards* 66 (2): 155–166. doi:10.1007/s11069-012-0398-8.

Hengl, T., G. B. M. Heuvelink, and D. G. Rossiter. 2007. “About Regression-Kriging: From Equations to Case Studies.” *Computers and Geosciences* 33 (10): 1301–1315. doi:10.1016/j.cageo.2007.05.001.

Ito, A., and T. Oikawa. 2002. “A Simulation Model of the Carbon Cycle in Land Ecosystems (Sim-cycle): A Description Based on Dry-Matter Production Theory and Plot-Scale Validation.” *Ecological Modelling* 151 (2–3): 143–176. doi:10.1016/S0304-3800(01)00473-2.

Kamieli, A., R. T. Nurit Agam, M. A. Pinker, M. L. Imhoff, G. G. Gutman, N. Panov, and A. Goldberg. 2010. “Use of NDVI and Land Surface Temperature for Drought Assessment: Merits and Limitations.” *Journal of Climate* 23 (3): 618–633. doi:10.1175/2009JCLI2900.1.

Kim, S.-N., W.-K. Lee, Y. Son, Y. Cho, and M.-S. Lee. 2012. “Applicability of Climate Change Impact Assessment Models to Korean Forest.” *Journal of Korean Society of Forest Science* 98 (1): 33–48.

Kogan, F. N. 1995. “Application of Vegetation Index and Brightness Temperature for Drought Detection.” *Advances in Space Research* 15 (11): 91–100. doi:10.1016/0273-1177(95)00079-T.

Kogan, F., W. Guo, and W. Yang. 2019. “Drought and Food Security Prediction from NOAA New Generation of Operational Satellites.” *Geomatics, Natural Hazards and Risk* 10 (1): 651–666. doi:10.1080/19475705.2018.1541257.

Komoscu, A. U. 1999. “Using the SPI to Analyze Spatial and Temporal Patterns of Drought in Turkey Using the SPI to Analyze Spatial and Temporal Patterns of Drought in Turkey.” *Drought Network News* 11 (1): 7–13.

Li, Z., Y. Chen, G. Fang, and L. Yupeng. 2017. “Multivariate Assessment and Attribution of Droughts in Central Asia.” *Scientific Reports* 7 (1): 1316. doi:10.1038/s41598-017-01473-1.
Mai, B. T., Huang, N. T., & Pham, X. T. (2020). “The Impact of Drought on Crop Yields in Vietnam.” *Environmental Research: Ecological Applications* 20 (6): 1–15. doi:10.1007/s41748-020-00211-w.

Sasai, T., K. Ichii, Y. Yamaguchi, and R. Nemani. 2005. “Simulating Terrestrial Carbon Fluxes Using the New Biosphere Model ‘Biosphere Model Integrating Eco-Physiological and Mechanistic Approaches Using Satellite Data’ (BEAMS).” *Journal of Geophysical Research: Biogeosciences* 110 (G2): n/a–n/a. doi:10.1029/2005JG000045.

Sasai, T., K. Okamoto, T. Hiyama, and Y. Yamaguchi. 2007. “Comparing Terrestrial Carbon Fluxes from the Scale of a Flux Tower to the Global Scale.” *Ecological Modelling* 208 (2–4): 135–144. doi:10.1016/j.ecolmodel.2007.05.014.

Schwabe, K., J. Albic, J. D. Connor, and R. M. Hassan. 2013. “Drought in Arid and Semi-Arid Regions.” *Springer*. doi:10.1007/978-94-007-6636-5.

Shahin, M. A., M. A. Ali, and A. S. Ali. 2014. “Vector Autoregression (VAR) Modeling and Forecasting of Temperature, Humidity, and Cloud Coverage.” In *Computational Intelligence Techniques in Earth and Environmental Sciences*, edited by Islam, T., Srivastava, P. K., Gupta, M., Zhu, Z., and Mukherjee, S. P. K., 29–51. Dordrecht: Springer Netherlands. doi:10.1007/978-94-017-8642-3_2.

Shamsnia, S. A., N. Shahidi, A. Liaghat, A. Sarraf, and S. F. Vahdat. 2011. “Modeling of Weather Parameters Using Stochastic Methods (Arima Model)(Case Study: Abadeh Region, Iran).” *IPCBEE* 12. International Conference on Environment and Industrial Innovation, 282–285. http://www.ipcbee.com/vol12/55-C30028.pdf.

Sheffield, J., G. Goteti, F. Wen, and E. F. Wood. 2004. “A Simulated Soil Moisture Based Drought Analysis for the United States.” *Journal of Geophysical Research* 109 (D24): 1–19. doi:10.1029/2004JD005182.

Shen, Z., Z. Zhang, V. P. Singh, P. Sun, C. Song, and Y. Huijjan. 2019. “Agricultural Drought Monitoring across Inner Mongolia, China: Model Development, Spatiotemporal Patterns and Impacts.” *Journal of Hydrology* 571: 793–804. doi:10.1016/j.jhydrol.2019.02.028.

Svoboda, M., Hayes, M., and Wood, D.A. 2012. “Standardized Precipitation Index User Guide.” Geneva, Switzerland: WMO. 978-92-63-11090-9

Thenkabail, P. S. 2015. “Remote Sensing Handbook: Remote Sensing of Water Resources, Disasters, and Urban Studies.” *Remote Sensing of Water Resources, Disasters, and Urban Studies* 3. doi:10.1201/b19321.

United Nations Convention to Combat Desertification (UNCCD). 2021. “Regional Approaches to Combat Drought, Sand and Dust Storms in Central Asia.” Accessed 19 January 2021. https://www.unccd.int/conventionregions/regional-approaches-combat-drought-sand-and-dust-storms-central-asia.
United Nations Economic and Social Commission for Asia and the Pacific (UNESCAP). 2021. “Enhancing the Capacity of Countries in Central Asia on Effective Use of Space Applications for Drought Monitoring and Early Warning through the Regional Drought Mechanism.” Accessed 22 August 2019. https://www.unescap.org/our-work/ict-disaster-risk-reduction/regional-space-cooperation-mechanisms-drought-monitoring-and-sustainable-development

USAID. 2018. “Climate Risk Profile: Kyrgyz Republic.” Washington, D.C., USA: USAID. Retrieved 25 Nov 2020 https://www.climatelinks.org/resources/climate-risk-profile-kyrgyz-republic

Wardlow, B., M. Anderson, C. Hain, W. Crow, J. Otkin, T. Tadesse, and A. Amir. 2017. “Advancements in Satellite Remote Sensing for Drought Monitoring.” CRC Press 2. doi:10.1201/9781315265551-14.

World Food Programme. 2014. “Kyrgyz Republic: An Overview of Climate Trends and the Impact on Food Security.”

Xu, Z., L. Cao, S. Zhong, G. Liu, Y. Yang, S. Zhu, X. Luo, and D. Liping. 2020. “Trends in Global Vegetative Drought from Long-Term Satellite Remote Sensing Data.” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 13: 815–826. doi:10.1109/JSTARS.2020.2972574.