Determination of Appropriate Remote Sensing Indices for Spring Wheat Yield Estimation in Mongolia

Battsetseg Tuvendorj 1,2, Bingfang Wu 1,2,* , Hongwei Zeng 1,2, Gantsetseg Batdelger 3 and Lkhagvadorj Nanzad 2,3,4

1 State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China; baku@radi.ac.cn (B.T.); zenghw@radi.ac.cn (H.Z.)
2 University of Chinese Academy of Sciences, Beijing 100049, China; lkhagya@radi.ac.cn
3 National Remote Sensing Center, Information and Research Institute of Meteorology, Hydrology, and Environment (IRIMHE), Ulaanbaatar 15160, Mongolia; gantsetseg@irimhe.namem.gov.mn
4 Key Laboratory of Digital Earth Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100194, China

* Correspondence: wubf@radi.ac.cn; Tel.: +86-10-64855689; Fax: +86-10-64858721

Received: 6 September 2019; Accepted: 29 October 2019; Published: 1 November 2019

Abstract: In Mongolia, the monitoring and estimation of spring wheat yield at the regional and national levels are key issues for the agricultural policy and food management as well as for the economy and society as a whole. The remote sensing data and technique have been widely used for the estimation of crop yield and production in the world. For the current research, nine remote sensing indices were tested that include normalized difference drought index (NDDI), normalized difference water index (NDWI), vegetation condition index (VCI), temperature condition index (TCI), vegetation health index (VHI), normalized multi-band drought index (NMDI), visible and shortwave infrared drought index (VSDI), and vegetation supply water index (VSWI). These nine indices derived from MODIS/Terra satellite have so far not been used for crop yield prediction in Mongolia. The primary objective of this study was to determine the best remote sensing indices in order to develop an estimation model for spring wheat yield using correlation and regression method. The spring wheat yield data from the ground measurements of eight meteorological stations in Darkhan and Selenge provinces from 2000 to 2017 have been used. The data were collected during the period of the growing season (June–August). Based on the analysis, we constructed six models for spring wheat yield estimation. The results showed that the range of the root-mean-square error (RMSE) values of estimated spring wheat yield was between 4.1 (100 kg ha\(^{-1}\)) to 4.8 (100 kg ha\(^{-1}\)), respectively. The range of the mean absolute error (MAE) values was between 3.3 to 3.8 and the index of agreement (d) values was between 0.74 to 0.84, respectively. The conclusion was that the best model would be \(R^2 = 0.55\) based on NDWI, VSDI, and NDVI out of the nine indices and could serve as the most effective predictor and reliable remote sensing indices for monitoring the spring wheat yield in the northern part of Mongolia. Our results showed that the best timing of yield prediction for spring wheat was around the end of June and the beginning of July, which is the flowering stage of spring wheat in this study area. This means an accurate yield prediction for spring wheat can be achieved two months before the harvest time using the regression model.

Keywords: MODIS; northern Mongolia; remote sensing indices; spring wheat; yield estimation
1. Introduction

Food security is an important topic for every country in the world [1]. Accurate and timely estimation of the spring wheat yield on regional and national scales is becoming absolutely essential for developing countries like Mongolia. In particular, crop yield estimation and the monitoring of crop production can provide fundamental information for crop producers, decision-makers in planning harvest and for agricultural development overall [2]. The agriculture sector is the second contributor to the Mongolian economy after mining [3]. However, only 13% of agricultural production is sourced from crops, mostly spring wheat, the remaining 87% is from the livestock [4] since the Mongolian climate is more suitable for extensive grazing, which covers more than 80% of the total land area. The spring wheat is below 1% of the total land area and around 1.35 million hectares of the total land is suitable for crop cultivation [5]. The northern part of Mongolia has the most favorable natural conditions and a more suitable area for rain-fed crops [6]. Hence, most of the spring wheat is grown in the northern provinces due to above-average precipitation. However, precipitation can only support the basic water requirement of spring wheat, and little variation in precipitation would cause a big fluctuation in crop yield. The vegetation cover, crop yields, and their growth are highly dependent on the amount of precipitation and the related soil moisture [7,8]. Mongolia has an extreme continental climate, with a short growing season, high evaporation, and low precipitation, which pose serious limitations for the Mongolian agriculture development. Because of the high altitude, our country’s climate is much colder than other countries in the same latitude. More than 80% of total spring wheat cultivation is rain-fed and only 5000 hectares is irrigated for spring wheat in Mongolia [4]. Therefore, agricultural production is particularly sensitive to climate variability and climatic conditions make agriculture very challenging. Due to the impacts of climate change, more extreme and continued droughts have occurred in many parts of Mongolia and have directly affected the vegetation and crop growth, biodiversity and socioeconomics in Mongolia [9]. Nanzad et al. [10] found that about 41–57% of Mongolia has been ravaged by mild to severe droughts for many of the last 17 years. A consecutive severe drought in 2002, 2005, 2007, 2010, 2013, and 2015 lowered spring wheat production severely as shown in Figure 1 and spring wheat had to be imported as local production declined due to weather conditions [4].

![Figure 1. Whole Mongolian grain production with study area (National Statistics Office, 2019).](image-url)
Weather information is normally used to forecast crop yield, but there is a lack of continuous measurement among others due to cost factors. Using Earth observation satellite imagery for monitoring temporal and spatial variation, combined with the point observation as a co-monitoring has advantages. Furthermore, satellite imagery is produced at a lower cost than the traditional way and is more easily accessible for use [11,12]. The use of remote sensing data helps to assess crop conditions in different fields at regional and whole country levels, even in remote areas, as it gives a timely and accurate measurement. Therefore, there have been many attempts in the applications of remote sensing in crop yield estimation, monitoring and mapping and most of these work streams indicated that remote sensing technology was prospective and promising [13–21]. A number of field studies have shown that models based on remote sensing data enable to estimate crop yield in many countries. Usually, remote sensing derived indices are connected to crop yield using empirical regression-based models [22,23]. During the past decades, remote sensing has been broadly used in forecasting crop yield. The Advanced Very-High-Resolution Radiometer (AVHRR) is the most popular sensor, the most widely used in terms of crop monitoring and yield forecasting since the early 1980s for a large scale [24,25]. In recent years, satellite-derived data such as Moderate Resolution Imaging Spectroradiometer (MODIS), Landsat, and Sentinel data were used for the yield prediction and monitoring and meaningful results have been obtained [11,12,26–28]. Lewis et al., [29] used AVHRR-NDVI data for maize production forecasts and correlated results showed that forecasts could be obtained one month before the harvest. In Spain, Vicente-Serrana et al. [30] combined AVHRR-NDVI data and drought indices and were able to predict wheat and barley yield four months before harvest. Moreover, Peterson [12] found the best timing to predict crop yield was from two to four months before the harvesting using NDVI, EVI, and NDWI of MODIS for different crops in Africa. Recently, some remote sensing indices such as the normalized multiband drought index (NMDI), vegetation supply water index (VSWI), and visible and shortwave infrared drought index (VSDI) were utilized in a number of studies for drought and crop monitoring and crop yield estimation according to previous studies [31–34]. The more promising method is using crop growth modeling that incorporates updated crop biophysical parameters such as leaf area index (LAI) and a fraction of absorbed photosynthetically active radiation (fPAR) retrieved from satellite imagery and by using survey information of crops throughout the growing season in local to regional areas. For example, Huang et al. [35] found that more accurate county-level winter wheat estimation was obtained using the WOFOST-PROSAIL model. Furthermore, many researchers have developed crop growth models to estimate crop yields [36–40]. However, the crop growth models require more specific information, such as daily weather data, soil properties, and crop growth determining factors, which would make analytical costs excessive. It is obvious that no general indicator can be used to predict crop yields in all regions. The applicability of the indicator will vary with the region, crop type, and crop growth stage.

Some recent studies in Mongolia were conducted to monitor the cropland cover changes, to assess land degradation for the agricultural region. Erdenee et al. [5] have used Landsat TM and ETM data in the detection of changes cropland over Tsagaannuur, Selenge provinces from 1989 and 2000. Otgonbayar et al. [41] investigated to prepare a cropland suitability map of Mongolia using Landsat and MODIS (MOD13, MOD15, and MOD17). Furthermore, Enkhjargal et al. [42] used MODIS and SPOT time-series remotely sensed data from 2000–2013 to estimate long-term soil moisture content in agricultural regions of Mongolia. Ariya [43] used Landsat images from 2000 and 2015 to assess land degradation for the agricultural area of Mongolia. Nevertheless, to date, no studies of crop yield estimation using remote sensing indicators have been done yet in Mongolia. Recently, remote sensing indicators are employed to monitor the drought across the pasture lands of Mongolia [44] and provide valuable information for drought management and reduction. Compared to drought, more attention should be paid to crop yield, while these indices were not getting enough attention in the field of crop yield in Mongolia. The small variation of precipitation would cause the big fluctuation of crop yield so that it is very important to forecast spring wheat yield early for food security in Mongolia. Although spring wheat accounts for a small proportion of Mongolia’s land area, the size of the spring wheat
field is large enough that it provides the possibility of predicting spring wheat yields based on remote sensing technology in Mongolia.

Therefore, our analysis focuses on Selenge and Darkhan provinces of Mongolia due to the availability of high-quality spring wheat yield data for those regions. It is the first attempt to estimate the spring wheat yield using space observation technology in Mongolia. The main objectives of this study were as follows: (i) to evaluate the potential of using remote sensing nine indices to estimate spring wheat yield; (ii) to choose the more suitable remote sensing indices for predicting spring wheat yield; (iii) to identify the best timing and more accurate model to estimate spring wheat yield in Northern Mongolia.

2. Study Area and Data

2.1. Study Area

Mongolia is divided into five different agro-ecological regions, which reflect distinct geographical patterns of agricultural production and climate. The study was conducted in Darkhan and Selenge provinces, which is located in Selenge-Onon agro-ecological region (N48–51 °C and E104–108 °C) (Figure 2). Selenge and Darkhan provinces are the principle cropping area for and account for more than 50% of national total grain production (Figure 1). In this region, most of the spring wheat cultivated area is rain-fed cropland. Thus, determining the fluctuation of spring wheat yield is highly dependent on the weather condition. The region averages between 90 to 110 frost-free days and has annual mean precipitation between 250 and 400 mm. In addition, crop growing duration is short (90–140 days) in this region, and depends on location and altitude. Approximately 90% of the nationwide precipitation is lost to evapotranspiration, which is associated with a continental climate. The remaining 10% of the total precipitation has been unable to evaporate, and 37% contribute to soil and underground reserves and streams, while 63% is surface runoff. In other words, 3–4% of total precipitation becomes potentially available as a water resource in the form of soil moisture or groundwater [6]. Additionally, the mean annual temperature is between 0.0 °C and 2.5 °C with cold temperature in January to –20 °C and warm temperature in July to 19 °C. Average elevations in this region have ranged between 1500 and 2000 m [45].

![Figure 2. Study area: (a) Agro-ecological regions of Mongolia; (b) Digital Elevation Model data (DEM), Spatial distribution of agrometeorological stations in Selenge and Darkhan provinces and cropland.](image-url)
2.2. MODIS Data and Processing

This study applied the use of 1 km spatial resolution, daily intervals MODIS/MOD1B time-series data to evaluate spring wheat yield. The MODIS data provided a high temporal resolution and wide coverage areas but a low spatial resolution. MODIS time-series data during the growing season (June to August) were obtained for the study area from the Atmosphere Archive and Distribution System (LAADS) of (NASA) for 2000–2017. Available online https://ladsweb.modaps.eosdis.nasa.gov (accessed on 20-10-2019). The study area covered over three tiles of granules of MODIS data such as an H24V03, H24V04, and H25V04 (H is horizontal and V is vertical coordinate). All downloaded MODIS 18 years’ images were re-projected from sinusoidal to the Albers equal area projection. The reflectance bands (NIR, red, blue, and SWIR) were calculated using MOD021KM level 1b calibrated data. In the processing section, all the collected images were re-projected, mosaicked and calibrated for atmospheric and geometric correction using ENVI IDL software. In this study, the nine vegetation and drought indices (NDVI, NMDI, NDWI, VCI, TCI, VHI, NDDI, VSDI, and VSWI) are calculated from cloud-free and corrected reflectance bands.

2.3. Crop Data

In this study, we have utilized annual spring wheat yield data for eight agrometeorological stations from 2000 to 2017 in Northern Mongolia (Table 1). The sowing stage of spring wheat is generally in the first decade of May. All the crops present an important vegetative development in the June–August period and spring wheat harvest occurs generally in September. The statistics of spring wheat yield were obtained from the agrometeorological division of Information and Research Institute of Meteorology, Hydrology, and Environment (IRIMHE) in Mongolia. Agrometeorological stations measure the crop phenology stage, growth condition and damage, crop density and height for every 10 days from May to September. Finally, the spring wheat yield sown from sampling surveys at the end of the growing season was also measured. Spring wheat yield was collected in 50 × 50 cm plots, in four repeated samplings at each agrometeorological station. From the homologies sample plots, four samples of spring wheat were taken through crop cutting from a different area at equal distance and their average was taken to minimize random errors. Before thrashing and weighting the spring wheat grain yield, the sample plot was placed in the oven for 5–10 min and at 20°C–25°C to easily split the grain yield and the straw. The final spring wheat grain weight of each station was converted to 100kg ha⁻¹ unit. As shown in (Figure 3), the phenological stages of the normal growth cycle of spring wheat and mean climate variables in the study area from May to September (2000–2017). In order to obtain the crop remote sensing indices values correctly, we had to solve image masking for regional spring wheat yield estimation. Applying a cropland mask to select remote sensing indices values as input to a crop yield model significantly improves the accuracy of the crop yield estimation [18,46]. A copy of the crop cover mask was obtained from the land cover map of Mongolia, provided by Elbegjargal et al. [47], and used to reduce the influence of non-agricultural areas on the remote sensing indices signal [48]. Finally, all areas with non-agricultural land were masked out and the regional annual yield was estimated and mapped for only cropland areas yield estimated maps were produced by model (M4) in the study area from 2000 to 2017.
Table 1. Information of meteorological stations with location (spring wheat yield available stations).

<table>
<thead>
<tr>
<th>N</th>
<th>Province Name</th>
<th>Station Name</th>
<th>Station ID</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Crop Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Selenge</td>
<td>Tsagaannuur</td>
<td>209</td>
<td>50.0886</td>
<td>105.3252</td>
<td>wheat</td>
</tr>
<tr>
<td>2</td>
<td>Selenge</td>
<td>Baruunkharaa</td>
<td>241</td>
<td>48.7856</td>
<td>106.2624</td>
<td>wheat</td>
</tr>
<tr>
<td>3</td>
<td>Selenge</td>
<td>Orkhon</td>
<td>242</td>
<td>49.0039</td>
<td>105.4208</td>
<td>wheat</td>
</tr>
<tr>
<td>4</td>
<td>Selenge</td>
<td>Eruu</td>
<td>243</td>
<td>49.6842</td>
<td>106.6008</td>
<td>wheat</td>
</tr>
<tr>
<td>5</td>
<td>Selenge</td>
<td>Orkhontuul</td>
<td>245</td>
<td>48.7208</td>
<td>105.0358</td>
<td>wheat</td>
</tr>
<tr>
<td>6</td>
<td>Darkhan</td>
<td>Tsaidam</td>
<td>2443</td>
<td>49.3221</td>
<td>105.9826</td>
<td>wheat</td>
</tr>
<tr>
<td>7</td>
<td>Darkhan</td>
<td>6th Brigad</td>
<td>2444</td>
<td>49.3602</td>
<td>106.0782</td>
<td>wheat</td>
</tr>
<tr>
<td>8</td>
<td>Darkhan</td>
<td>Altangadas</td>
<td>2447</td>
<td>49.2328</td>
<td>105.9466</td>
<td>wheat</td>
</tr>
</tbody>
</table>

3. Methodologies

The purpose of this study is to create a predictive measure of spring wheat yield computed from satellite data. The methodology has three main parts. Firstly, we calculate the nine remote sensing indices using MODIS data. Secondly, nine remote sensing indices are used for the relationship between actual spring wheat yield as input for the testing of the estimation model of spring wheat yield. Thirdly, we develop spring wheat yield regression models and map estimated spring wheat yield for the regional level to the 18 years. The general flowchart of this research, the processing method, and the individual steps are illustrated in Figure 4.
3.1. Calculation of Remote Sensing Nine Indices:

Crop yield is markedly influenced by the growth conditions in each crop stage. Nine typical vegetation and drought indices were computed and compared with spring wheat yield. These indices are commonly used for crop yield estimation, drought monitoring, and only require optical spectral bands for calculation. For example, the NDVI has been most commonly used for vegetation monitoring, crop yield assessment and forecasting [27,49,50]. Kogan [51] has been developed the VCI to study the response of vegetation to drought conditions worldwide. Unganai et al. [52] conducted that the vegetation condition index (VCI) derived from AVHHR-NDVI correlated significantly with maize yield in Zimbabwe. The growth condition and crop yield are strongly correlated at the pixel level. Each pixel’s value of remote sensing indices was directly taken at point locations of the eight stations. We examined the relationship between NDVI, NDWI, NMDI, TCI, VCI, VHI, NDDI, VSDI, and VSWI with actual spring wheat yield during the growing season (June–August) for 2000–2017 across Northern Mongolia. The 10 day and monthly nine remote sensing indices (NDVI, NMDI, NDWI, VCI, TCI, VHI, NDDI, VSDI, and VSWI), which are derived from transformations of the red, NIR, blue, and SWIR spectral bands, was used to continuous time series of data the represented the spring wheat growth indices temporal curve for each pixel in the study area. We used Table 2 for remote sensing indices calculation. Active plants are intensely absorbed by red and blue bands and reflected by the near-infrared band (NIR) [53]. Depending on the type of plant and the crop yield stage, the absorption of red and blue bands and the degree of the NIR band is different. When vegetation is stressed by lack of water, and also at the end of the growing period, the chlorophyll absorption declines and the ratio of NIR to RED or visible reflectance decreases. Therefore, vegetation index can be defined as the difference between RED and NIR bands reflections measured from satellites. The SWIR-1 and SWIR-2 bands are sensitive to the soil and vegetation moisture content. Hence, we calculated nine remote sensing indices utilized NIR, red, blue, and SWIR bands for spring wheat yield monitoring. Furthermore, we determined which indices from nine remotely sensed indices were more suitable to estimate spring wheat yields. To reduce the impact of atmosphere and cloud, the 10-day remote sensing-based indices derived from the daily data using maximum value composition (MVC) [54] were considered. Besides, we have retrieved monthly remote sensing indices from 10 days MVC indices by a simple moving average method. Particularly, each monthly remote sensing index was calculated by averaging three temporally 10 days MVC indices.

Figure 4. Flowchart of processing method.
3.2. Sensitive Analysis between Remote Sensing Indicators and Crop Yield

In this study, Pearson’s correlation coefficient (R) between remote sensing indices and spring wheat yield was calculated for every 10 days and every month of the growing season from June to August for the northern part of Mongolia for 2000–2017 using following (Equation (1)). Pearson’s correlation coefficient (R) represents the degree and direction of the linear regression between two continuous variables that are measured on the equal interval. The range of values for the R is from −1 to 1 (R > 0 it is a positive linear relationship, R = 0 it is indicated that there is no relationship and R < 0 it is a negative linear relationship).

\[
R = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}
\]  

(1)

where \(x_i\) and \(y_i\) presents remote sensing indices and the value of spring wheat yield at different time periods, \(n\) is the number of samples, \(\bar{x}\) and \(\bar{y}\) are the average values of \(x_i\) and \(y_i\).

3.2.1. Crop Yield Estimation Model

The process of crop production is directly influenced by many biological, physiological and biophysical laws that are directly responsible for plants, and theoretically, it is possible to model the whole process of vegetative growth by mathematically describing the physical processes of these processes. Nowadays, dynamic modeling techniques are based on simple statistical equations, based on complex differential equations systems, in modeling the events of agroecosystems. In this study, using the remote sensing nine indices, we tested and attempted to develop the spring wheat yield estimation model. The relationship between spring wheat yield and remote sensing indices was observed through the linear regression model, where the independent variable was represented by remote sensing nine indices and the dependent variable was spring wheat. The estimation model is multilinear and includes slope and an interception constant coefficient. The empirical regression methods based on spectral indices have commonly used for modeling crop yield in many studies \cite{49,60,61}. Bolton et al. \cite{62} found a good linear correlation between different crop yields with MODIS-NDVI, EVI, and NDWI data at the county levels. Sui et al. \cite{11} developed the estimation model for winter wheat production based on the environmental factors derived from satellite at a regional level, with errors of <12% for winter wheat yield, respectively. To determine when the correlation between remote sensing indices and spring wheat yield is strongest, we estimated the coefficient of determination (\(R^2\)) using data from every 10 days and monthly in the growing period of spring wheat. In order to make the final predicted results more accurate and stable, we considered it necessary to select the critical growth stage and remote sensing indices from all indices. We used a stepwise regression model in order to select the best candidate indices for yield estimation using SPSS 12 software. Stepwise regression provides a strong

<table>
<thead>
<tr>
<th>N</th>
<th>Remote Sensing Based Indices</th>
<th>Equation</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Normalized Difference Vegetation Index</td>
<td>(NDVI = \frac{NIR - RED}{NIR + RED})</td>
<td>\cite{55}</td>
</tr>
<tr>
<td>2</td>
<td>Normalized Difference Water Index</td>
<td>(NDWI = \frac{NIR - SWIR}{NIR + SWIR})</td>
<td>\cite{56}</td>
</tr>
<tr>
<td>3</td>
<td>Temperature Condition Index</td>
<td>(TCI = (NDVI1-NDVI2)/100)</td>
<td>\cite{57}</td>
</tr>
<tr>
<td>4</td>
<td>Normalized Multi-Band Drought Index</td>
<td>(NMDI = \frac{NIR-(SWIR1-SWIR2)}{NIR-(SWIR1-SWIR2)})</td>
<td>\cite{58}</td>
</tr>
<tr>
<td>5</td>
<td>Vegetation Supply Water Index</td>
<td>(VSWI = \frac{Ts}{NDVI})</td>
<td>\cite{59}</td>
</tr>
<tr>
<td>6</td>
<td>Normalized Multi-Band Drought Index</td>
<td>(NDDI = \frac{NIR+NDVI}{NIR+NDVI})</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Visible and Shortwave Infrared Drought Index</td>
<td>(VSDI = 1 - [(SWIR2 - BLUE) + (RED - BLUE)])</td>
<td>\cite{63}</td>
</tr>
</tbody>
</table>

NIR-near-infrared band, Band2 (841–876 nm); RED-red band, Band1 (620–670 nm); SWIR1 - shortwave infrared band, Band6 (1628–1652 nm); SWIR2-shortwave infrared band, Band7 (2105–2155 nm); Blue–Band 3 (459–479 nm); Tmax, Tmin–surface temperature (current, maximum, and minimum).
mean between the one or more independent variables and a dependent variable that conforms to the general equation for a multidimensional flat.

$$\hat{Y} = b_0 + b_1X_1 + b_2X_2 + \ldots + b_nX_n,$$  \hspace{1cm} (2)

where, $\hat{Y}$ is the dependent variable and predicted spring wheat yield, $X_1, X_2\ldots X_n$ are the independent variables and MODIS remote sensing indices, and $b_0, b_1, b_2 \ldots b_n$ are the regression coefficients and $n$ is the number of independent variables.

### 3.2.2. Model Performance Evaluation

Then we selected the month that produced the highest coefficients to the determination to develop multilinear regression models based on all 18 years of data. Generally, the comprehensive method to validate models is to correlate the measured values against the predicted values [48]. We used model fitting and performance statistics such as the coefficient of determination ($R^2$), root mean square error (RMSE), mean absolute error (MAE), bias and an index of agreement ($d$) agreement.

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2}{n}}, \hspace{1cm} (3)$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |\hat{Y}_i - Y_i|, \hspace{1cm} (4)$$

$$d = 1 - \frac{\sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n} (|Y_i - \bar{Y}| + |\hat{Y}_i - \bar{Y}|)^2}, \hspace{1cm} (5)$$

where $Y$ is observed values, $\hat{Y}$ is modeled values, $\bar{Y}$ is an average of observed values, and $n$ is a number of yield and RS data.

### 4. Results

#### 4.1. Temporal Climate Variables and Remote Sensing Indices Profiles for Spring Wheat

Figure 5 describes annual spring wheat yield with the amount of precipitation and average temperature throughout each growing season from 2000 to 2017. The temperature change, precipitation, and soil moisture have a significant impact on wheat yield. Mainly in the period May–September which accounts for 85% of annual precipitation. Especially in June–August, about 50%–60% of the annual precipitation occurs in Mongolia [63]. According to the [64] results show that the break of rainy season caused a similar reduction in soil moisture around mid-July in Mongolia. The highest mean precipitation of the growing months at the eight meteorological stations was 352 mm in 2013, 309 mm in 2009, 293 mm in 2008, 290 mm in 2012 and lowest annual precipitation occurred in 143 mm in 2002, 172 mm in 2001, and 199 mm in 2005 for the study area. In summer, especially growing season temperature and precipitation were negatively correlated. From (Figure 5) the lowest yield of spring wheat harvested in 2002(4.2 kg ha$^{-1}$) and the highest yield in 2014(21.9 kg ha$^{-1}$) in these two provinces. The main reason is that high temperatures and low precipitation led to soil moisture deficits, which is a significant impact on wheat yield. The climate of Mongolia is dry and semi-arid, and the growing vegetation cover and crop yields and development are highly dependent on the amount of precipitation and the related soil moisture [7,8].

Processed growing season (June–August) Moderate Resolution Imaging Spectroradiometer (MODIS-Terra) daily reflectance bands for study area for the years to 2000 to 2017. Based on the corrected surface reflectance NIR, red, blue, SWIR bands and LST of MODIS data, we have computed nine remote sensing indices and listed in (Table 2). We extracted time series values of NDVI, NDVI,
NMDI, NDWI, VCI, TCI, VHI, NDDI, VSDI, and VSWI interpolated 10 day and monthly intervals in growing season (June–August) from 2000 to 2017. As shown in (Figure 6), the long-term annual remote sensing nine indices variables for spring wheat yield in Northern Mongolia.

Figure 5. Long-term trend of average spring wheat yield and climate variables (2000–2017).

4.2. Sensitivity Analysis between Remote Sensing Indicators and Crop Yield

The Pearson’s correlation analysis examined between multi-year spring wheat yield data of eight stations and NDVI, NMDI, NDWI, TCI, VCI, VHI, NDDI, VSDI, and VSWI for June-August from 2000 to 2017, for a total of about 135 data points. To test the regional effectiveness of remote sensing indices and determine the best index during the growing period (June to August) for spring wheat yield estimation, we compared the 10 days and monthly remote sensing indices with the spring wheat yield. In general, in terms of the sown stage in the middle of May, emergency in early June, flowering in late June and early July, milk, and dough in August, maturity in early September, and harvesting in late September in the study area (each growing stage shown in Figure 3). For each of the remote sensing indices, the correlation values and significant p-values were produced by eight stations from 2000 to 2017 (Table 3).

The results show that most of the indices had found a higher correlation with spring wheat yield in June and July. This period covers the heading and flowering phenological stage of spring wheat. To see all of the correlations between June remote sensing indices and spring wheat yield in more detail, refer to Figure 7. About 50–60% of the annual precipitation occurred in the growing period, especially most of the precipitation occurred in June and the beginning of July. Because of that reason, there is sufficient moisture and suitable weather condition. Slightly lower correlation was found in August since crops are harvested in September. We can see that results of correlation (R) among the nine indices, NDDI, VSDI, VSWI, and crop yield were negatively correlated with spring wheat yield, while other NDVI, VCI, TCI, VHI, NDWI, and NMDI were positively correlated. The highest correlation coefficient (R) values of between 10 day remote sensing 9 indices with spring wheat yield were NDVI
(0.51) in first 10 days of July, NMDI (0.18) in first 10 days of August, NDWI (0.48) in first 10 days of June, TCI (0.57) in third 10 days of June, VCI (0.31) in first 10 days of July, VHI (0.46) in third 10 days of June, NDDI (−0.38) in second 10 days of July, VSDI (−0.56) in first 10 days of July and August, and VSWI (−0.36) in first 10 days of June, respectively. These were statistically significant values with a yield of \( p < 0.05 \). Indices including NDDI, VSDI, and VSWI were negatively correlated with spring wheat yield, this implies when the values of these indices increase, the spring wheat yield decreased.

![Figure 6. Annual growth season (June–August) average VSWI (a), VSDI (b), NDDI (c), VHI (d), VCI (e), TCI (f), NDWI (g), NMDI (h), NDVI (i), and Spring wheat yield (j).](image-url)
Table 3. Multi-year correlation between 10-day indices and spring wheat yield in June to August (2000–2017).

<table>
<thead>
<tr>
<th>Month</th>
<th>Decade</th>
<th>Index</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>NDVI</td>
</tr>
<tr>
<td>June</td>
<td>First 10 days</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>Second 10 days</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>Third 10 days</td>
<td>0.31</td>
</tr>
<tr>
<td>July</td>
<td>First 10 days</td>
<td>0.51</td>
</tr>
<tr>
<td></td>
<td>Second 10 days</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Third 10 days</td>
<td>0.27</td>
</tr>
<tr>
<td>August</td>
<td>First 10 days</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>Second 10 days</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>Third 10 days</td>
<td>0.3</td>
</tr>
</tbody>
</table>

Significant at highlighted values = $p < 0.05$, Number of samples = 126 – 134 ((Due to the clouds or error data of satellite images there were some 10 days of some indices are missed).

Figure 7. The correlations between spring wheat yield with June (a) NDVI, (b) NDWI, (c) NMDI, (d) TCI, (e) VCI, (f) VHI, (g) NDDI, (h) VSDI, and (i) VSWI. Number of samples $N = 135$.

Also, we did a correlation between monthly nine remote sensing indices and spring wheat yield at each station in the monitoring period. The results show that the absolute value of correlation coefficients (R) of each month as shown in Table 4. These were statistically significant $p < 0.01$ with yield, except NMDI.

<table>
<thead>
<tr>
<th>N</th>
<th>Index</th>
<th>June</th>
<th>July</th>
<th>August</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NDVI</td>
<td>0.38***</td>
<td>0.47***</td>
<td>0.28***</td>
</tr>
<tr>
<td></td>
<td>NMDI</td>
<td>0.1</td>
<td>0.12</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>NDWI</td>
<td>0.51***</td>
<td>0.29***</td>
<td>0.35***</td>
</tr>
<tr>
<td></td>
<td>TCI</td>
<td>0.31***</td>
<td>0.35***</td>
<td>0.32***</td>
</tr>
<tr>
<td></td>
<td>VCI</td>
<td>0.45***</td>
<td>0.35***</td>
<td>0.38***</td>
</tr>
<tr>
<td></td>
<td>VHI</td>
<td>-0.33***</td>
<td>-0.39***</td>
<td>-0.37***</td>
</tr>
<tr>
<td></td>
<td>NDDI</td>
<td>-0.33***</td>
<td>-0.57***</td>
<td>-0.33***</td>
</tr>
<tr>
<td></td>
<td>VSDI</td>
<td>-0.51***</td>
<td>-0.37***</td>
<td>-0.33***</td>
</tr>
<tr>
<td></td>
<td>VSWI</td>
<td>-0.38***</td>
<td>-0.57***</td>
<td>-0.25**</td>
</tr>
</tbody>
</table>

Significant at * = p < 0.05, ** = p < 0.01, *** = p < 0.001; Number of samples=134.

The relationship between NMDI and spring wheat yield showed was the lowest results (0.10–0.15), which is indicating that this index not suitable for spring wheat yield estimation in this region. From Table 4 we can see that monthly NDWI was the highest correlated indices with yield in June (0.51) and monthly VSDI was the highest correlated with yield in July (-0.57) and these were statistically significant p < 0.001, respectively. It indicates the soil and crops moisture and water content are most important for crop yield.

4.3. Yield Estimation Model

In this study, we tested nine remote sensing indices to develop the best and most accurate estimation models for spring wheat yield for Northern Mongolia. The yield was estimated at station level based on remote sensing indices for the spring wheat-growing season (June to August) and ground crop yield data. Each model has used eight stations crop yield data and nine remote sensing indices (10 daily and monthly) from 2000 to 2017. We used stepwise regression as a technique for choosing independent nine remote sensing indices for a multiple linear regression equation from a list of candidate indices. The results of regression analysis and best-fitted models are summarized in Table 5.

Table 5. Best-fit estimation models for spring wheat yield during the growing period (June–August).

<table>
<thead>
<tr>
<th>Month</th>
<th>Model</th>
<th>Equations</th>
<th>R²</th>
<th>SEM</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>Model 1</td>
<td>y = 44.837 - 41.661 × VSDI₆ˢ + 37.745 × NDWI₆</td>
<td>0.53</td>
<td>4.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Model 2</td>
<td>y = 35.041 - 36.358 × VSDI₆ˢ + 24.621*NDWI₆ + 13.668 × VHI₆</td>
<td>0.57</td>
<td>4.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>July</td>
<td>Model 3</td>
<td>y = 44.721 - 39.502*VSDI₇ + 21.296 × NDWI₇</td>
<td>0.44</td>
<td>4.6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Model 4</td>
<td>y = 34.492 - 37.189 × VSDI₇ + 28.571 × NDWI₆ + 19.61 × NDVI₇</td>
<td>0.55</td>
<td>4.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>August</td>
<td>Model 5</td>
<td>y = 52.224 + 12.774 × NDWI₈ˢ - 46.207 × VSDI₈</td>
<td>0.39</td>
<td>4.9</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>Model 6</td>
<td>y = 62.527 + 25.492 × NDWI₈ˢ - 50.254 × VSDI₈ + 23.552 × NDVI₈</td>
<td>0.44</td>
<td>4.8</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Number of samples (N = 135), VSDI₆ˢ-third 10 days of June, VSDI; NDWI₆-June, NDWI; VSDI₇-July, VSDI; NDVI₇-first 10 days, NDVI; VSDI₈, NDWI₈, and NDVI₈-first 10 days of August, VSDI, NDWI, and NDVI.

The models had R² values ranging from 0.39 to 0.57 and all models had statistically significant p < 0.001, respectively. The models with high R² values, low RMSE, and MAE values indicate the best model for spring wheat yield estimation. The highest R² values were 0.57 in June and 0.55 in July. Final all models include NDWI and VSDI (VHI and NDVI in some models) from June to August were good predictors of spring wheat yield.

In order to test the estimate performance of the method, we used the coefficient of determination (R²), root means square error (RMSE), mean absolute error (MAE), bias and index of agreement (d) to
evaluate the estimated spring wheat yield in regional level. We compared the predicted yield with the actual yield of eight stations for 2000–2017, the results showed in (Figure 8).

![Figure 8. Comparison between actual spring wheat yield with estimated spring wheat yield for each model; model 1 (a), model 2 (b), model 3 (c), model 4 (d), model 5 (e), model 6 (f). Each point indicates the estimated yield versus the actual yield for a single station and year.](image)

The best timing and more accurate model for spring wheat yield estimation was found at the end of June and beginning of July. Model 4 was selected as the best estimation model for spring wheat in Northern Mongolia. Model 4 has combined variables of VSDI, NDWI, and NDVI, and index of agreement (d) value was 0.84, the relationship between estimated and actual yield was $R^2 = 0.55$, mean absolute error was $MAE = 3.3$ and root mean square error was $RMSE = 4.1(100\, \text{kg}\, \text{ha}^{-1})$, respectively.

### 4.4. Evaluation of Spring Wheat Yield at the Regional Scale

Using a cropland mask, the calibrated model was applied in the study area. The spring wheat yield was evaluated based on the best model 4 from 2000 to 2017 in the Northern part of Mongolia (Selenge and Darkhan provinces) and the results are shown in (Figure 9). The temporal and spatial spring wheat yield maps produced based on the best model using NDWI (June), VSDI (third decade of June), and NDVI (first decade of July) indices, at the regional level have been generated in cropland area from 2000 to 2017. The spatial patterns of estimated spring wheat yield ranged from 1.3–35(100 kg ha$^{-1}$). In generated spring wheat yield maps, green colors indicating the normal and favorable condition and highest values of crop yield; red color shows unfavorable conditions and less amount of crop yield, respectively. From this result, we can see that spring wheat yield was low in drought years (2000–2002, 2005, 2006, 2015, and 2017). The irrigation system is limited in Mongolia, most of the spring crops at their vegetative and reproductive stages suffer water stress due to recurrent drought. Drought stress influences the water supply to vegetation and reduces accumulated biomass and production of crops [65].
water stress due to recurrent drought. Drought stress influences the water supply to vegetation and reduces accumulated biomass and production of crops [65].

**Figure 9.** Estimated temporal and spatial spring wheat yield map at the regional level.
5. Discussions

The study paves a new way for crop monitoring in Northern Mongolia. We have explored nine remote sensing indices in decade and month intervals. We found that there are 4 indices (VHI, VSDI, NDWI, and NDVI) that are more relevant than other indices for spring wheat yield estimation. This study has found that the NDWI and VSDI are the best indices for Mongolian crop monitoring. The NDWI was mainly indicated as an effective tool for water stress, soil, and vegetation moisture conditions and water content in vegetative areas, which was determined by the NIR and SWIR bands [11,12,33,48,62]. The supply of moisture to the north-central cropping region of Mongolia comes out as the main factor that clearly demonstrates the results of findings [66]. Water deficiency causes a physiological disorder that can inhibit cell division and differentiation, leading to the reduction of plant size and yield [67]. Our best results were obtained through Model 4 that showed \( R^2 = 0.55 \), respectively.

The results of the relationship between indices showed that MSAVI obtained the best wheat yield estimation model \( (R^2 = 0.63) \), which was slightly higher than our result. Indices (SAVI, MSAVI, NDVI, and EVI) selected for wheat yield estimation in irrigated Indus Basin of Pakistan [68] are the difference from ours, because at irrigation area, water is not stress issue. Water is a stressed factor in rainfed spring wheat in Mongolia. Dempewolf et al. [49] developed the wheat yield forecasting model in Punjab province of Pakistan using time-series MODIS and Landsat derived vegetation indices (NDVI, WDRVI, EVI2, SANDVI). The final results show that a forecasted wheat yield was within 0.2% and 11.5% of actual values, which was lower than our result. Bolton and Friedl et al. [62] compared the accuracy of different indices (EVI2, NDVI, and NDWI) in different zones for maize and soybean yield in the Central United States, and the EVI2 obtained best accuracy result with \( R^2 = 0.73 \), respectively. This result was higher than our result.

We also find the highest correlation between indices with spring wheat yield and peaked at the flowering stage. The peak period of vegetative for spring wheat yield is June and July in this region. This implies that vegetation has its strongest response to moisture availability during this period. The growth condition of spring wheat during flowering stage might have more yield information than other stages at all growth stages, which means the developed regression model can be predicted two months before harvesting the crop and the correlation of estimated and actual yield from heading to flowering periods is higher than other crop growth stages. These results are in agreement with previous studies showing this to be the most suitable time to predict yield [27,48,60,61,69].

Furthermore, we find that later June is the most critical time for spring wheat yield formation. Indices in later June are used in every equation. We examined the relationship between actual spring wheat yield with 10 days and monthly indices for the regression model in the growing period (June–August). Similarly, we tried out the relationship between accumulated and relative values of nine indices with spring wheat yield from 2000 to 2017. The results of the correlation between accumulated and relative values of remote sensing nine indices with actual yield were lower than 10 day and monthly indices value. Juan Sui et al. [11] developed the dry aboveground mass and wheat yield estimation model using several remote sensing indices derived from MODIS and Himawari-8 sensor. A dry aboveground mass and yield errors of <10% and 12% were reported in Hengshui city of Hebei province, which was slightly lower than our results. Lopresti et al. [27] performed based on time-series MODIS-NDVI data for wheat yield estimation model obtained a higher correlation with estimated and actual yield \( (R^2 = 0.75) \), which was higher than our results. Also, several regression models for crop yield estimation based on MODIS, NDVI, are presented [18,48,61]. Moriondo et al. [61] have carried out the NDVI data to estimate wheat yield in the Grosseto and Foggia provinces of Italy. The results of the correlation coefficient between simulated and actual yield were 0.73–0.77, with corresponding RMSE were 0.44Mg/ha and 0.47Mg/ha, respectively.

The impacts of global warming are already confronted in Mongolia, visible from records between 1940 and 2013 from 48 meteorological stations. According to Dagvadorj et al. [70], the temperature has increased by 2.07°C compare to mean. Due to the impacts of climate change, more extreme and continued droughts have occurred in many parts of Mongolia and which directly affect the
vegetation and crop growth, biodiversity and socioeconomics in Mongolia [9]. Reportedly, [10,44,71] 2000–2002, 2004, 2005, 2007, and 2009 years were extremely affected by mild to severe drought and slight drought-hit Mongolia in 2003 and 2011. An additional notable finding of this study is that the spatial regional spring wheat yield distributions shown that the spring wheat yield was high in 2011, 2012, 2013, and 2016 and was low from 2000 to 2003 and 2015. It was statistically significant ($p < 0.001$), respectively and confirmed our result that during the drought years’ spring wheat yield was low. Perhaps, in the year 2003, we had the highest precipitation in our monitoring period. The amount of precipitation, soil type, soil moisture, and changes in air temperature have a significant impact on wheat yield. Particularly, drought and soil moisture deficit influence the most reduced crop yield and vegetation size [72]. Thus, from our results, we recommend developing an irrigation system for spring wheat cultivation and increase the number of crop yield observation samples in this region. These results obviously show the promising application of NDWI and VSDI data in crop yield assessment at relatively cheap cost and timely.

6. Conclusions

In Mongolia, the application of remote sensing methodology in agricultural policy and practices is in its nascent stage. This was the first time a multi-regression model based on remote sensing indicators was used to estimate crop yield in Northern Mongolia which is the main spring wheat-producing region. For this purpose, the best and most suitable indices were first defined through the testing of correlations between the nine indices and the actual spring wheat yield. Our results show that NDVI, NDWI, VCI, TCI, VHI, and NMDI indices with spring wheat yield were positively correlated (0.47, 0.51, 0.38, 0.47, 0.45, and 0.15), respectively and NDDI, VSWI, and VSDI with spring wheat yield were negatively correlated ($-0.39$, $-0.57$, and $-0.38$), respectively. Furthermore, the results confirmed the importance of the integration of both satellite and ground data for crop yield estimation. Consequently, we selected the NDWI, VSDI, and NDVI as the most suitable indices out of the nine indices, which are NDVI, NDWI, VCI, TCI, VHI, and NMDI. The highest negatively and positively correlated indices are a combination of NIR, red, blue, and SWIR bands. SWIR and red bands are found more sensitive to moisture variation and water stress of crops and soils [33]. Among nine indices, NDWI (0.51) in June and VSDI ($-0.57$) in July show the highest correlated indices with actual spring wheat yield, which indicates that the soil and crop moisture, as well as water content, are very important factors for crop yield.

As next step, we refined and developed the regression model using the above three selected indices in order to estimate crop yield. In total, six models were elaborated to be used for the growing period which is June to August. Timeline observation showed a higher correlation between indices and spring wheat yield during the flowering stage in June and July. Therefore, it is suggested that a suitable time to estimate spring wheat yield is at this stage of one to two months before the harvest. Whereas the results for the month of August showed a lower correlation indicating the lateness to estimate. The best results were obtained through Model 4 that used a combination of indicators from the period of third 10 days of June of VSDI, average of June NDWI and first 10 days of July of NDVI. Therefore, the Model 4 is the most effective predictor for crop yield monitoring in the northern part of Mongolia.

In this paper, we could estimate only 74% of the actual yield. This was due to several reasons that possibly could be ascribed to the different spatial resolution between MODIS data (1 km) and the ground measured spring wheat yield data (station-based measurements). Also, phenomena such as the different soil structures and the amount of precipitation have a big influence on the yield. However, the application of remote sensing regression model results enormously enrich the ground station collected data by providing large scale, region-wide data for the decision-makers to better manage food security challenges. In the future, work needs to be carried out to apply more consistently high-resolution images, such as Landsat and Sentinel for more accurate estimation of crop yield. In general, a comprehensive and systematic use of remote sensing technology in the agriculture sector of
Mongolia is to be considered, including broader policy for research and development, the introduction of the latest technology and equipment and targeted capacity-building activities.

**Author Contributions:** B.T. contributed to the research experiments, methodology, software, data analysis, and the writing-original draft of the manuscript; B.W. conceived the experiments, and was responsible for the research analysis; H.Z. contributed to experimental design and the manuscript revision; G.B. contributed to the collection of crop yield data and the manuscript revision; L.N. contributed the manuscript revision.

**Funding:** This paper was funded by the National Natural Science Foundation of China (41561144013, 41601464), National Key R&D Program of China (2016YFA0600304), the International Partnership Program of Chinese Academy of Sciences (131C11KYSB20160061,121311KYSB20170004), the Queensland-Chinese Academy of Sciences (Q-CAS) Collaborative Science Fund (grant number: 131211KYSB20170008 & 201700257).

**Acknowledgments:** We are very grateful to colleagues of the Information and Research Institute of Meteorology, Hydrology, and Environment (IRIMHE) of Mongolia. Special thanks to Elbegjargal Nasanbat for his comments and suggestions. We thank the anonymous reviewers for reviewing the manuscript and providing comments to improve the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).