

GENERATING TIME SERIES OF MEDIUM-RESOLUTION TEMPERATURE VEGETATION DROUGHT INDEX IMAGES USING A KALMAN FILTER METHOD FOR SOIL MOISTURE CHANGE ANALYSIS

H. Zhai¹, F. Huang^{1,2,*}, H. Qi¹, Y. Ren¹, R. Liu¹, B. Li¹, S. Chang¹

¹ School of Geographical Sciences, Northeast Normal University, Renmin Street, Changchun, China – (zhaih669, huangf835, qih102, reny899, liur026, lib25, changs164)@nenu.edu.cn

² Key Laboratory of Geographical Processes and Ecological Security in Changbai Mountains, Ministry of Education, School of Geographical Sciences, Northeast Normal University, 130024, Changchun, China

KEY WORDS: Temperature Vegetation Drought Index, Data fusion, Time series, Kalman filter, MODIS; Landsat

ABSTRACT:

Soil moisture is one of key environmental variables that affect vegetation cover and energy exchange between the land surface and the atmosphere. Satellite remote sensing technology can provide information for monitoring large-scale soil moisture dynamics quickly. The temperature vegetation dryness index (TVDI) acts as an effective indicator of inferring soil moisture status which is calculated according to the empirical parameterization of composed of the land surface temperature (LST) and the normalized difference vegetation index (NDVI) characteristic space. In this paper, the MODIS TVDI was calculated based on MODIS LST product (MOD11A2, 1 km) and NDVI data (derived from MOD09A1, 500m). Meanwhile, LST and NDVI from Landsat8 OLI images were estimated to obtain Landsat-based TVDI. Then, a Kalman filter algorithm was used to simulate TVDI time series data with 30m resolution and a revisit period of 8 days combining TVDI derived from Landsat and MODIS data. We selected the west of the Songnen Plain, China as the test area and high quality cloudy-free images during growing season (April to October) of 2018 as the input data. The predicted TVDI time series data of medium resolution not only improved the temporal resolution to capture the changes at fine scale within a short period, but also made up for the deficiency of low spatial resolution MODIS data. The results show that it is feasible to generate medium or high resolution TVDI time series data by applying different remote sensing data by Kalman filtering algorithm.

1. INTRODUCTION

1.1 General Instructions

Soil moisture is one of the important indicators for monitoring land conditions and a decisive factor for crop growth. It has been widely applied in hydrology, climate, ecology and other field (Yang, 2009). Soil moisture not only acts as the link of energy exchange between atmosphere and land in hydrologic cycle, but also is closely related to the growth development of

crops and vegetation. It plays a decisive role in the spatial distribution of terrestrial ecosystem (Wu, 2018). Traditionally, the methods of monitoring soil moisture mainly rely on site-scale soil moisture content data to characterize soil drought conditions. Although this method has a high measurement accuracy, its spatial representativeness was poor (Garcia, 2014). Compared with traditional methods, remote sensing technology has the advantages of strong timeliness, wide monitoring range and rapid acquisition of large-scale ground information, which is of great significance to the dynamic monitoring of drought.

Temperature Vegetation Drought Index (TVDI) is an empirical parameterized calculation based on the feature space composed by land surface temperature (LST) and normalized difference vegetation index (NDVI). It integrates the information of vegetation index and land surface temperature, and is regarded as an effective indicator to infer soil moisture status. In recent years, TVDI has been widely used in drought monitoring and obtained many effective results. Wu et al. (Wu, 2007) used MODIS data to establish NDVI-LST characteristic space to obtain TVDI for drought warning and monitoring in mountainous. Qi et al. (Qi, 2003) adopted the NDVI and LST extracted by NOAA-AVHRR to construct the NDVI-LST feature space as a drought index and studied the drought from March to May in 2000 in China. Wu et al. (Wu, 2018) calculated the monthly TVDI of Songnen Plain from 2000 to 2015, and found a good correlation between TVDI and the relative soil moisture data of meteorological stations, which indicated that TVDI could be used as an indicator of soil moisture.

TVDI data can be derived from many sensors including Moderate Resolution Imaging Spectroradiometer (MODIS), SPOT VEGETATION and Advanced Very High Resolution Radiometer (AVHRR) etc. These medium and low spatial resolution data (>300m) can only be used for large scale drought conditions monitoring (Cao, 2019; Garcia, 2014; Qi, 2003; Wang, 2014; Wu, 2007; Wu, 2018), but may not suitable for the areas at fine scale or with high heterogeneity. Since Landsat data has a spatial resolution of 30m TVDI data with fine resolution and spatial details can be obtained by calculating the feature space constructed by NDVI and LST. However, due to the cloud contamination and the limitations of instruments, continuous time series fine resolution images are not available most of the time, which makes it difficult for the acquisition of TVDI. Therefore, through combining the high spatial and low temporal resolution data of Landsat sensor with the high temporal and low spatial resolution data of MODIS, high spatial and temporal resolution TVDI data can be obtained to improve the ability of drought monitoring. Cao et al. (Cao, 2019) used Ensemble Kalman Filter method to integrate AMSR-E and ASCAT soil moisture data, which was helpful for the study of soil moisture data in Qaidam basin. Sedano et al. (Sedano, 2014) proposed a complete time series data assimilation method for the synthesis of medium-resolution images, and used Landsat NDVI images and MODIS NDVI products to generate the time series NDVI images of 16-day time step with a spatial resolution of 30 meters. The synthesized time series images can capture the seasonal surface dynamics and maintain the spatial structure of the landscape. Zhang et al. (Zhang, 2017) used Kalman filter algorithm to simulate 30m albedo images combined with Landsat ETM+ images and MODIS products for a year continuously. The results showed that Kalman filter method could reflect the variation tendency of MODIS data primarily.

In this paper, the Kalman filter (KF) algorithm is adopted to simulate TVDI data based on Landsat OLI images and MODIS products with a resolution of 30m at an 8-day time intervals. KF

algorithm is an optimal linear state estimation method which integrates observations, models and their respective uncertainties. The uncertainty information is included in the calculation process, which makes the method robust and flexible enough to process sub-optimal standard data (Huang, 2016; Zhang, 2017). KF method is applied to TVDI data for the first time in this study, which can provide reference for the research and application of high spatio-temporal resolution surface parameter simulation method.

2. DATA AND METHOD

2.1 Study Area

The study area in this paper is located in the western part of Songnen plain, between 122°56' E -122°14' 16" E and 47°16' 21" N - 47°0' 13" N, with the size of 48km*48km. As an agro-pastoral transitional area, the land cover types in the study area are mainly composed of forest, grassland, farmland and bare soil.

2.2 Satellite Data

The MODIS land surface reflectance data (MOD09A1) were used to create the NDVI time series in our study site which has 8-day revisiting period and 500m spatial resolution. MOD09A1 was atmospherically corrected and the high quality cloud-free data was used. MOD11A2 LST product data includes the day and night LST with the resolution of 1km. The images are the 8-day composite data of daily LST. MOD11A2 was resampled to 500m which was consistent with the NDVI. In this study, MOD09A1 and MOD11A2 LST products from April to October in 2018 were downloaded for the study area (23 images, h26 v04) from NASA LAADS web (<https://ladsweb.nascom.nasa.gov/>). Then, the 500m NDVI and the corresponding LST data was treated as input data to calculate TVDI with ENVI software

Landsat 8 Operational Land Imager (OLI) images processed at the Level 1 Generation System were used, which were accessed at the United States Geological Survey (USGS) <https://earthexplorer.usgs.gov/>. 7 cloud-free Landsat images was selected as the input data The Landsat imagery was radiation calibrated and atmospherically corrected. The reflectance at Band 3 red (0.63–0.69 μm) and Band 4 near-infrared (0.76–0.90 μm) were used to generate NDVI images for each Landsat image. Meanwhile, retrieving LST was a key step to obtain 30m TVDI images.

2.3 LST Retrieval from Landsat Images

Mono-window algorithm for retrieving LST from Landsat data proposed by Qin (Qin, 2001) was used in this paper. In this method, land surface emissivity, atmospheric transmissivity and

effective mean atmospheric temperature are used to estimate surface temperature. For Landsat 8, the brightness temperature of each pixel can be estimated by formula (1):

$$T_L = 1321.0789 / \ln \left(1 + \frac{774.8853}{b1} \right) \quad (1)$$

where $b1$ is band 10 of Landsat 8 after radiation correction. The actual land surface temperature is computed as following formula:

$$LST = (-67.355351 * (1 - C - D) + T_L * (0.458606 * (1 - C - D) + C + D) - D * T_a) / C \quad (2)$$

In formula (2), T_a is the effective mean atmospheric temperature, C and D is the intermediate variable which can be calculated by land surface emissivity and atmospheric transmissivity. More details can refer to the literature (Qin, 2001).

2.4 The Kalman Filter Method for Generating Continuous Time Series TVDI Images

The Kalman filter is a recursive algorithm which integrates observations, a linear state-transition models, and their respective uncertainties to simulate the state of a process with minimization of the root mean square error (Huang, 2019):

$$x_k = Ax_{k-1} + w_{k-1} \quad (3)$$

$$z_k = Hx_k + v_k \quad (4)$$

where x_{k-1}, x_k = the model estimates in the previous and present state individually

A = the transition matrix links x_{k-1} and x_k

w, v = Gaussian random variables, stands for the process and measurement noise respectively

z_k = a given observation

H = links x_k to z_k

Moreover, the implementation of KF algorithm includes two steps: time update and measurement update. Firstly, the linear transition model provides the estimate of the previous state and its uncertainty to produce prior estimates of the present state by equation (5) and (6):

$$x_k^- = A\hat{x}_{k-1} \quad (5)$$

$$P_k^- = AP_{k-1} + w_{k-1} \quad (6)$$

where x_k^- = the prior estimate of the present state

\hat{x}_{k-1} = the posterior estimate of the previous state;

P_k^- = the prior uncertainty for the present state

P_{k-1} = the posterior uncertainty of the previous state

Secondly, the prior estimates (7) and their uncertainties (8) are updated by new observations throw a linear combination of the prior model estimate and a weighted difference between the observations and the prior estimates. And the Kalman gain (K) is calculated by equation (9):

$$\hat{x}_k = x_k^- + K_k(z_k - Hx_k^-) \quad (7)$$

$$P_k = (1 - K_kH)P_k^- \quad (8)$$

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \quad (9)$$

Where \hat{x}_k = the posterior estimate of the state

R, P = measuring noise and process covariance respectively

Large value of R leads to low Kalman gain, which provides more weight to the model process. On the contrary, large P results in high K and greater weight for new measurements (Sedano, 2014).

In our study, a Kalman filter recursive algorithm was used to produce synthetic 30m resolution TVDI time series data with 8-day time step using TVDI derived from Landsat 8 OLI and MODIS TVDI time series as input data. The TVDI data calculated from available Landsat images within the period of study were defined as the observations. The transition model was defined as an ensemble of two submodels. The first submodel provides a general trajectory of the TVDI values to characterize the temporal dynamics in our study area. The priori estimation was presented as equation (10):

$$x_{k1}^- = b + a \times \hat{x}_{k-1} \quad (10)$$

where x_{k1}^- = the priori estimation at status k from submodel

\hat{x}_{k-1} = the posteriori estimation at status $k-1$

a, b = the slope and intercept of the linear regression of MODIS values between time steps $k-1$ and k

The second submodel reveals the relationship between MODIS and Landsat data and gives a further restraint to the temporal tracks of the first model. Then, 10000 pixels of each pair of MODIS products and Landsat data were selected randomly to establish the regression model. A priori estimation was fulfilled using slope and intercept of the linear regressions according to the following equation (11), and the meaning of each symbol could be referred to equation (10).

$$x_{k2}^- = c + d \times \hat{x}_{k1} \quad (11)$$

The initial state of model was critical to establishing an accurate time series TVDI data. The available high resolution TVDI image was used as the initial data of the model. If the Landsat image was available for the first state (x_0), it was treated as the initial state. If not, the first MODIS TVDI image was defined as the initial state within the period of study, and the initial uncertainty factor (P_0) was defined as the standard deviation of the MODIS pixels.

The implementation of KF algorithm contains two modes: filtering and smoothing. The smoothing mode estimated the time series medium-resolution TVDI images. The process combined the forward and backward KF recursions:

$$\hat{x}_{FBk} = \hat{x}_{Fk} \left[\frac{P_{Fk}}{P_{Fk} + P_{Bk}} \right] + \hat{x}_{Bk} \left[\frac{P_{Bk}}{P_{Fk} + P_{Bk}} \right] \quad (12)$$

$$\frac{1}{P_{FBk}} = \left(\frac{1}{P_{Fk}} \right) + \left(\frac{1}{P_{Bk}} \right) - \left(\frac{1}{R_k} \right) \quad (13)$$

where $\hat{x}_{Fk}, \hat{x}_{Bk}, \hat{x}_{FBk}$ = the posterior estimates at state k in forward, backward and combined mode, respectively

P_{Fk}, P_{Bk}, P_{FBk} = the uncertainties at state k

R_k = the measurement uncertainty at state k

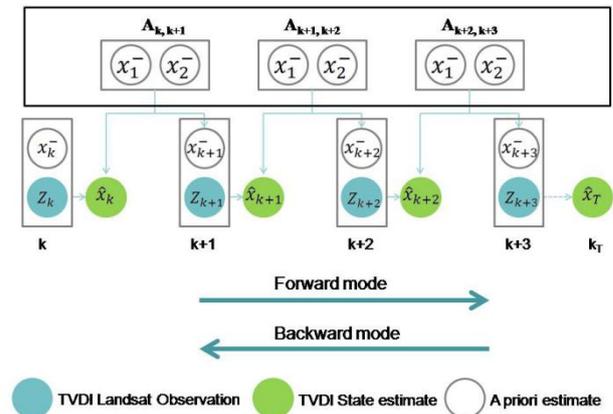


Figure 1. Flow chart of the Kalman filter approach

Figure 1 shows the process of Kalman filter algorithm for generating time series TVDI data. The transition model A projected the estimate from a previous state (time update). The priori estimate X_k^- was produced by the combination of submodels 1 and 2, X_{k2}^-, X_{k1}^- , respectively. If a Landsat observation is available, Z_k provides a new estimate (measurement update). \hat{x}_k is the final estimate which is the weighted average of x_k^- and Z_k . If not available, the estimate is solely the result of the measurement update from the previous step. The model runs taking turns to forward and backward mode (filter mode), and the corresponding estimates are then combined together (smoothing mode).

3. RESULT

3.1 Spatial Pattern of the Predict TVDI Images

Figure 2-5 shows the images of the Landsat retrieved TVDI, the MODIS TVDI (calculated by data of MOD09A1 and MOD11A2) and the Kalman filter simulation TVDI at the corresponding date. For each image, there is a significant gain in spatial detail of the synthetic images and the 500m TVDI MODIS images. Visually, the retrieval Landsat and synthetic TVDI images looked quite similar in spatial pattern. MODIS data has a lot of missing values due to the influence of clouds and get rid of the illegal values exceeding range (0-1). However, the missing data could be supplement by the temporal trend of MODIS data which composed of the forward and the backward Landsat data. Moreover, at the pixel level, time series of synthetic 30m resolution TVDI scenes kept the same general phenology trends of MODIS images. And the predicted TVDI data generating by Kalman filter algorithm could preserve the effect of topography fluctuation on TVDI.

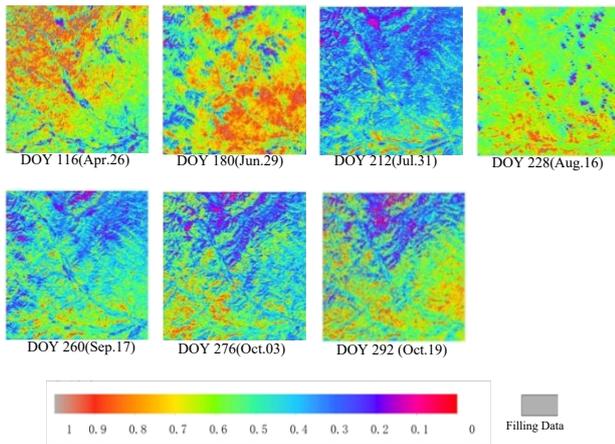


Figure 2. The retrieval TVDI from Landsat 8 OLI data

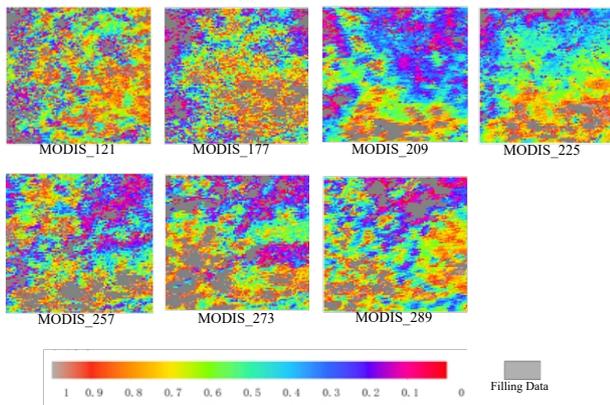


Figure 3. MODIS TVDI data in the same timeperiod with Landsat data

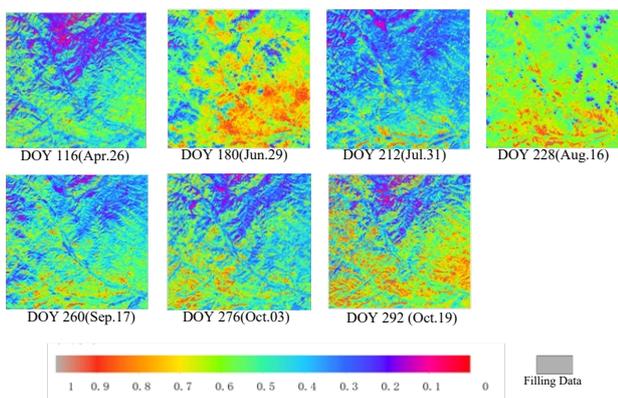


Figure 4. Kalman filter TVDI data corresponding with Landsat

3.2 Evaluation of Synthetic TVDI Data

Figure 5 shows the coefficients of determination (R^2) between the simulated TVDI images at 30m spatial resolution and the

retrieval TVDI from Landsat OLI images (observation data) of the same dates. The scatterplot was built by randomly selected from the predicted image and the observed image ($N = 10,000$). The Kalman filter method accurately predicted TVDI in the range of 0.3 to 0.7. However, the KF TVDI might be underestimated when TVDI was higher than 0.7. From Figure 5, MODIS TVDI data were slightly lower than the Landsat estimations in high value area and higher than Landsat in low value area. Moreover, the areas with filling values differed from the actual values. Therefore, these reasons may affect the predicted KF values for these pixels having low or high TVDI values.

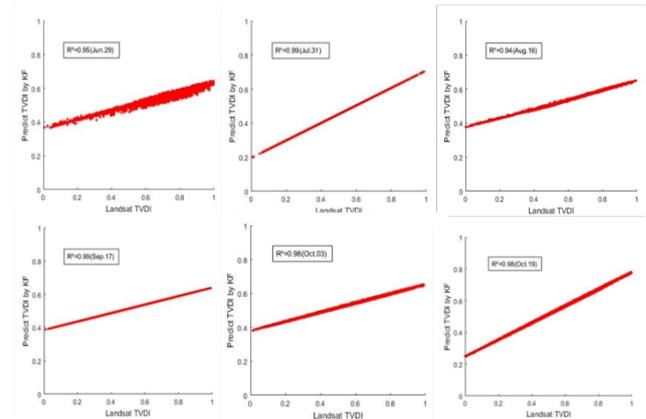


Figure 5. Scatter plot of 30m synthetic KF TVDI and Landsat inversion TVDI.

Da y of yea r	Landsat TVDI			MODIS TVDI			Synthetic TVDI		
	Mea n	SD	Max	Mea n	SD	Max	Mea n	SD	Max
121	0.630	0.171	0.999	0.583	0.248	0.997	0.375	0.006	0.359
177	0.630	0.176	0.999	0.540	0.264	0.990	0.649	0.005	0.528
209	0.382	0.132	0.999	0.526	0.248	0.995	0.7092	0.068	0.392
225	0.597	0.114	0.999	0.527	0.217	0.998	0.652	0.032	0.530
257	0.443	0.143	0.999	0.478	0.261	0.998	0.639	0.036	0.498
273	0.462	0.171	0.999	0.5077	0.270	0.998	0.660	0.046	0.510
289	0.512	0.191	0.999	0.5364	0.250	0.997	0.785	0.100	0.521

Table 1. The summary statistics: time series level mean values (Mean), standard deviation (SD) and maximum values of each image (Max).

R^2 is an indicator which could characterize the similarity between the predicted TVDI data by KF algorithm and the Landsat retrieval TVDI. It was found that R^2 of the predicted and measured values of each image was greater than 0.9. The scatter plot on April 26 was missing because the image of Landsat TVDI was acted as the original image for Kalman filter algorithm, so there was a large difference between the synthetic TVDI and the Landsat observations. Some larger and lower observed values were caused by the weather conditions and the bare soil land cover type. As it shown in Table 1 above, the mean value of MODIS data was a little higher than the observation Landsat data and the standard deviation were greater than both Landsat and synthetic TVDI data. As for the KF simulation TVDI, the standard deviation was the lowest compared to other data for the same date. In addition, the TVDI images predicted by Kalman filter could repaired the fluctuations of time series data to a certain extent.

3.3 Seasonal Changes of Soil Moisture Based on the Synthetic KF TVDI

Figure 5 shows the carves of the MODIS TVDI, the simulated TVDI by Kalman filter algorithm and the points of Landsat retrieval TVDI at the observing time. The regional mean values of TVDI during the peroid of our study ranged from 0.37-0.78. The highest TVDI was occurred at Day 185. The value of Day 209 was on the low side obviously because the influence of some cloud in the image. In the early, the synthetic data was lower than MODIS on account of the lack of initial images. Onerall, the 30m simulated TVDI data predicted by the Kalman filter captured the seasonal changes of MODIS data. As can be sean in the Figure 5, the TVDI values were high in April to September. TVDI decreased gradually and dropped to the lowest after the 289th day. The temporal change of TVDI was consistent with the seasonal changes in vegetation growth in the period of our study.

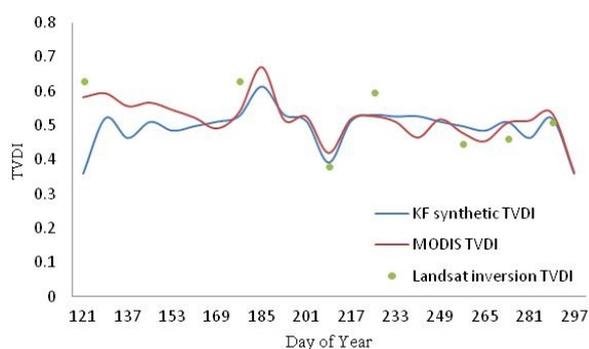


Figure 3. Comparison of MODIS TVDI, the simulated TVDI by Kalman filter algorithm and the Landsat retrieval TVDI

4. CONCLUSIONS

Accurate knowledge of soil moisture is useful for plant growth modeling, estimation of vegetation phenology and drought monitoring. In order to meet the demand for more informative soil moisture data and support a wider range of monitoring applications, this paper performed a data fusion algorithm among different sensors to generate soil moisture data with high spatio-temporal resolution. We adopted a the Kalman filter algorithm to simulate TVDI time series data with 30m resolution and a revisit period of 8 days combining TVDI derived from Landsat and MODIS data. The results revealed that the coefficients of determination (R^2) between the synthetic TVDI images and the Landsat retrieval TVDI were greater than 0.9. The average predicted residuals between the simulated TVDI images using KF model and the Landsat TVDI images were mostly between 0.02 and 0.25. The accuracy of synthetic TVDI data may meet the demands on input parameters of land surface process and ecological models. The predicted TVDI time series data of medium resolution not only improved the temporal resolution to capture the changes at fine scale within a short period, but also made up for the deficiency of low spatial resolution MODIS data. KF method introduces uncertain information in the process of calculation, which makes it flexible to deal with the uncertainties and different characteristics caused by long time series in TVDI datasets, with good flexibility and robustness. More spatial features derived from Landsat images were preserved. It was found that the dense time series TVDI data could reveal the seasonal variability of soil moisture in the study area, which would improve the timeliness of drought monitoring.

The results show that it is feasible to generate medium or high resolution TVDI time series data by applying different remote sensing data by Kalman filtering algorithm. Continuous time series of high resolution remote sensing data with great precision generated through data fusion would provide new opportunities for understanding and studying environmental processes with higher levels of spatio-temporal detail.

ACKNOWLEDGEMENTS

This research was supported by the National Natural Science Foundation of China (Grant No. 41571405 and 41630749).

REFERENCES

Cao J. W., W.J.H. 2019. Fusion of remote sensing data of soil moisture by ensemble kalman filtering. *Journal of Qinghai Normal University(Natural Science Edition)*, 35, 72-76

- Garcia, M., Fernandez, N., Villagarcia, L., Domingo, F., Puigdefaábregas, J., & Sandholt, I. 2014. Accuracy of the Temperature-Vegetation Dryness Index using MODIS under water-limited vs. energy-limited evapotranspiration conditions. *Remote Sensing of Environment*, 149, 100-117
- Huang, J.X., Ma, H.Y., Sedano, F., Lewis, P., Liang, S., Wu, Q.L., Su, W., Zhang, X.D., & Zhu, D.H. 2019. Evaluation of regional estimates of winter wheat yield by assimilating three remotely sensed reflectance datasets into the coupled WOFOST-PROSAIL model. *European Journal of Agronomy*, 102, 1-13
- Huang, J.X., Sedano, F., Huang, Y.B., Ma, H.Y., Li, X.L., Liang, S.L., Tian, L.Y., Zhang, X.D., Fan, J.L., & Wu, W.B. 2016. Assimilating a synthetic Kalman filter leaf area index series into the WOFOST model to improve regional winter wheat yield estimation. *Agricultural and Forest Meteorology*, 216, 188-202
- Qi S. H., W.C.Y., Niu Z. 2003. Evaluating Soil Moisture Status in China Using the Temperature/Vegetation Dryness Index(TVDI). *Journal of Remote Sensing*, 7, 420-427
- Qin Z. H., Z.M.H., Karnieli A., Berliner P. 2001. Mono-window Algorithm for Retrieving Land Surface Temperature from Landsat TM6 data. *Acta Geographica Sinica*, 68, 456-466
- Sedano, F., Kempeneers, P., & Hurtt, G. 2014. A Kalman Filter-Based Method to Generate Continuous Time Series of Medium-Resolution NDVI Images. *Remote Sensing*, 6, 12381-12408
- Wang Y., W.J.S., Yao Y. B., Zhao F. N. 2014. Dynamic monitoring of drought based on Temperature Vegetation Dryness Index in Guangdong Province. *Acta Prataculturae Sinica*, 23, 98-107
- Wu M. Q., C.W.H., Li J. G. 2007. Moniotring drought in mountainous area based on temperature/vegetation dryness index(TVDI). *Arid Land Geography*, 01, 30-35
- Wu X. R., N.X.D., Zang S. Y. Z. 2018. Application of temperature vegetation dryness index in estimation of soil moisture of the Songnen Plain. *Acta Ecologica Sinica*
- Yang, X., Wu J. J., Yan F. Zhang J. 2009. Assessment of regional soil moisture status based on characteristics of surface temperature/vegetation index space. *Acta Ecologica Sinica*, 29, 1205-1216
- Zhang G., H.F., Zheng M. L. 2017. Generating Time Series of Medium-Resolution Albedo Images by Kalman Filtering Algorithm. *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*