Tuning the Extended Kalman Filter for Rice Cultivation Date Estimation in Tropical Area using MODIS NDVI data

Rata Suwantong Panu Srestasathiern Preesan Rakwatin Saran Suwannachatkul
Geo-informatics and Space Technology Development Agency (GISTDA)
120 The Government Complex, Ratthaprasasanabhakti Building
6th and 7th Floor, Chaeng Wattana Road, Lak Si, Bangkok, Thailand 10210
E-mail: rata@gistda.or.th; panu@gistda.or.th; preesan@gistda.or.th; saran@gistda.or.th

KEY WORDS: Cultivation Date Estimation, Extended Kalman Filter (EKF), Moderate Resolution Imaging Spectroradiometer (MODIS), Normalized Difference Vegetation Index (NDVI), Tropical Area

ABSTRACT: Accurate crop cultivation date estimation is crucial for crop yield forecast, parasite attack trend forecast and for suitable government aid calculation in case of a natural disaster. The 8-day composite of the Normalized Difference Vegetation Index (NDVI) from the Moderate Resolution Imaging Spectroradiometer (MODIS) data is a good candidate to be used thanks to its global coverage and its free availability. Its spatial resolution is at 250 meter. However, in a tropical country as Thailand, the NDVI data is very noisy due to high cloud density. An appropriate estimation technique and tuning method must therefore be used to cope with the high measurement noises. An Extended Kalman Filter (EKF) was proposed in the literature for the rice cultivation date estimation in Thailand. The NDVI was modelled by a triply modulated cosine function with the mean, the amplitude and the initial phase as state variables. The estimated cultivation date was counted as the date where the seasonal part of the estimated NDVI, i.e. the estimated NDVI minus its mean, is higher than a certain threshold. However, the brute force algorithm in which the large numbers of tuning parameters were tested until the smallest mean error was obtained was used. The achieved tuning parameters were not only not consistent with the nature of the NDVI but might also be not appropriate for the study areas where the ground truths are not available. In this paper, we propose a method for tuning the EKF which is consistent with the nature of the NDVI data. The mean error of the estimated cultivation dates is at 21 days compared to 25 days of the existing method in the literature.

1. Introduction

Accurate crop cultivation date estimation is indispensable in an agricultural country. It is used not only to forecast crop yields and parasite attack trends, but also to calculate suitable government aids for the damage of the fields by natural disasters, such as flood, drought, and storm, which are frequent in a tropical area as the South-East Asia. A technique that enables the crop cultivation date estimation in wide-area is the satellite remote sensing technique. One of a good candidate to be the input for the crop cultivation date estimation is the Vegetation Index (VI) data from the Moderate Resolution Imaging Spectroradiometer (MODIS) which is available for free and has a global coverage. The MODIS is boarded on the satellites Terra and Aqua which have both a daily-repeated cycle. Terra passes from north to south across the equator in the morning, while Aqua passes south to north over the equator in the afternoon. Despite the daily-repeated cycle of Terra and Aqua, the MODIS data could be unavailable due to the presence of clouds and off-nadir sensor views. The data from Terra and Aqua is therefore processed independently at an 8-day interval. The data from the day with the least cloud and aerosol contamination and the lowest-view angle, is used to represent all the 8 days in the interval and is called the 8-day composite. Only the data from Terra is used in this studies by assuming that the data from Aqua is more noisy due to higher cloud density in the afternoon.

The vegetation indices (VI) are derived by examining spectral reflectance signatures given from leaves. The greener the plants are, the less reflected energy in the visible is due to high absorption by photosynthesis of the leaves. Among the visible bands, the blue band (470 nm) is the most absorbed and the red (670 nm) is the least absorbed. Nearly all of the near-infrared radiation (NIR) is scattered (reflected and transmitted) with very little absorption [Solano et al., 2010]. As a result, the greener the plants are, the higher the contrast between red and near-infrared responses. Two types of Vegetation Index from MODIS are available: the Enhanced Vegetation Index (EVI) and the Normalized Difference Vegetation Index (NDVI). The EVI, taking into account the NIR, the red and the blue bands has a greater sensitivity than that of the NDVI in high biomass area. Furthermore, the EVI is less affected by the atmospheric than the NDVI since it uses the difference in blue and red reflectances as an estimator of the atmosphere influence level. The EVI from MODIS has been used for the estimation of the number of crop cycles in China where the overall accuracy is shown to be of 91% [Li et al., 2014]. However, the spatial resolution of the EVI is at 500 meters which is too coarse to be used for most rice fields in Thailand where there are many small rice producers. The NDVI which is a normalized transform of the NIR to red reflectance ratio and has a spatial resolution of 250 meters will therefore be used in our studies. The NDVI varies from -1 to 1: 1 represents full vegetation and 0 represents no vegetation. However, the problem of the NDVI data from MODIS is that it is very noisy for a tropical study area such as Thailand due to high density of clouds. An efficient estimation method must therefore be chosen to cope with this problem. In [Chu et al., 2014], the NDVI from MODIS was used for the phenology detection of the winter wheat in the Yellow river delta. The Savitzky-Golay filter was selected to denoise the NDVI data. The ground truth of the phenology was not available in [Chu et al., 2014], hence the estimation error of their method is unknown.

Rice cultivation date estimation in Thailand using NDVI data from MODIS has been proposed in [Chumkesornkulkit et al., 2013] where a triply-modulated cosine function was used to model the measured NDVI data and the Extended Kalman Filter (EKF) was chosen to estimate the mean, amplitude and phase parameters of the cosine function. The cultivation dates were estimated as the date where the estimated seasonal variation of NDVI, which is the estimated NDVI minus the estimated mean, is greater than an arbitrary threshold. The estimated cultivation date of single crop rice was compared to the ground truths from the rice department of Thailand in 40 study areas. The mean error of the cultivation date estimates of single crop rice was shown to be of 17 days [Suwannachatkul et al., 2014]. However, to achieve this accuracy, the brute force algorithm where large numbers of tuning parameters of the EKF and threshold were tested until the smallest mean error was obtained. This method is therefore not practical. The tuning parameters obtained by this method may therefore suit only for the study areas in consideration and can provide high errors when it is applied to other areas in which the ground truth is not available. Furthermore, the appropriate standard deviation of the NDVI measurement noise in [Chumkesornkulkit et al., 2013] was found to be 20 while the range of the NDVI is between -1 and 1 which is not possible physically.

In this paper, we propose a method for tuning the EKF to replace the brute force algorithm in [Chumkesornkulkit et al., 2013], our novel method gives the standard deviation of the measurement noise which is possible physically. The paper is organized as follow. In section 2, we describe the NDVI from MODIS, the pre-filtering method of the NDVI, the state and the measurement equations and the estimated cultivation date computation method. In section 3, the EKF equations and the methods for the filter initialization and parameter tuning are provided. In section 4, we present the result analysis where the estimated cultivation date will be compared to the real cultivation date provided by the rice department of Thailand. The paper ends with concluding remarks.

2. NDVI, State and Measurement Definitions

2.1. NDVI Measurements

2.1.1. 8-day composite NDVI Data from MODIS

As in [Chumkesornkulkit et al., 2013], the 8-day MODIS/Terra data (MOD09Q1) from NASA for the period from the 18th of February 2000 until November 2012 is used for the studies. The NDVI is a normalized transform of the NIR reflectance ρ_{NIR} to red reflectance ρ_{red} ratio, designed to standardize VI values to between 1 and $+1$. It is commonly expressed as [Solano et al., 2010]:

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \tag{1}$$

The greener the plants are, the less reflected energy in the visible is due to high absorption by photosynthesis of the leaves. Therefore, the greener the plants are, the higher the NDVI is. Theoretically, the range of the NDVI is from -1 to 1 where $NDVI = 1$ represents full vegetation and $NDVI = 0$ represents zero vegetation. The NDVI has a spatial resolution of 250 meters.

The NDVI data is started to be collected from the 1st of January of each year followed by 8 days interval for each measurement. In this paper, 38 pixels in the north eastern part of Thailand where the single crop rice (raind-fed) is cultivated, are studied. The NDVI data for each pixel from 18th of February 2000 to the 11th of December 2012 which give the total of 587 NDVI data for each pixel, are used.

In figure 1, we compare the NDVI of single crop rice in Thailand in 2007 to that of winter wheat in the Yellow river delta between 2012 and 2013 in China (right). We observe the higher noise affecting the NDVI in Thailand compared to China. This may be due to the tropical climate of Thailand which induces more clouds.

We can remark that the NDVI is between 0 and 1 which implies that the case of no vegetation is not present. This may due to some other crops on the fields.

2.1.2. Pre-filtered 8-day composite NDVI Data

Seeing that the 8-day composite NDVI data acquired from MODIS is very noisy, we pre-filter this 8-day composite signal using the wavelet decomposition function (wavedec) available in MATLAB. Denote N as the total number of the instants. The wavelet decomposition, also called the Discrete Wavelet Transform (DWT), converts the

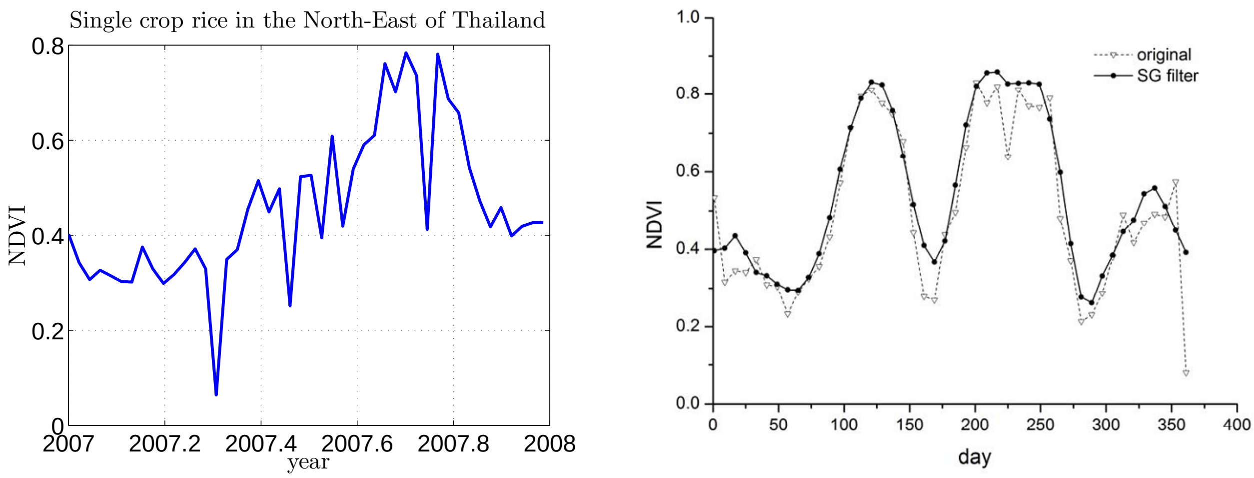

Figure 1: Example of NDVI data from MODIS produced by single crop rice in Thailand in 2007 (left) and that by winter wheat in the Yellow river delta between 2012 and 2013 in China from [Chu et al., 2014] (right). We observe the higher noise affecting the NDVI in Thailand compared to China. This is due to the tropical climate of Thailand which induces more clouds.

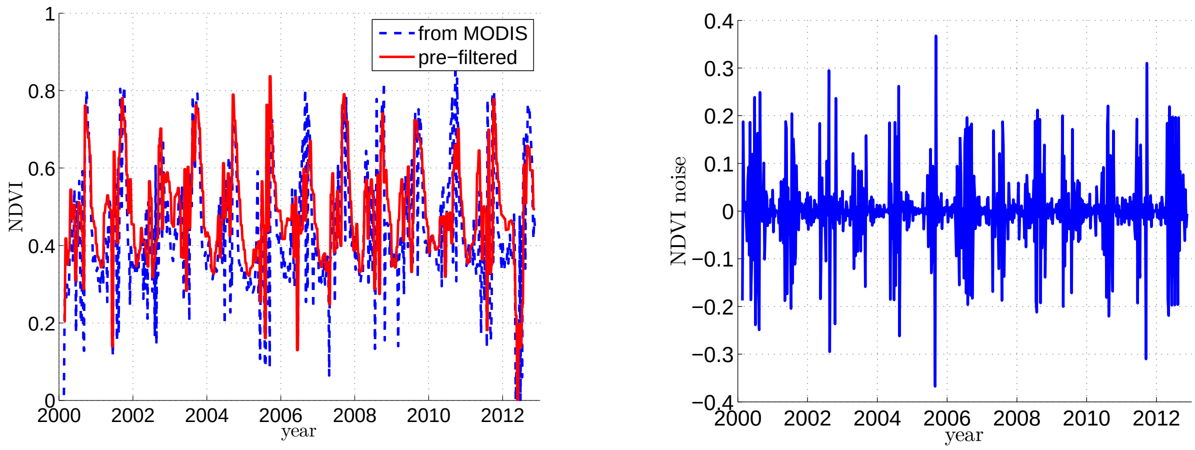

Figure 2: The pre-filtered NDVI which is the reconstruction from the low-pass coefficient series from the wavelet decomposition. This pre-filtered NDVI are given as the input (measurements) to the Extended Kalman Filter (left). The noise level on the NDVI data which is the reconstruction from the high-pass coefficient series from the wavelet decomposition (right).

signals, here the 8-day composite NDVI, into one high-pass wavelet coefficient series and one low-pass coefficient series of length $N/2$ each. The Haar wavelet which is equivalent to the Daubechies db1 wavelet is chosen as the mother wavelet.

Then, the low-pass wavelet coefficient series are reconstructed using the waverec function in MATLAB. The reconstructed signal, which is less noisy, is then given to the Extended Kalman Filter (EKF) as the measurements (inputs) of the filter. We call this reconstructed signal the pre-filtered measurement. The high-pass wavelet coefficient series are also reconstructed using the waverec function to visualize the measurement noise level. Figure 2 shows the reconstruction from two series.

2.2. State and Measurement Definitions

As in [Chumkesornkulkit et al., 2013], the NDVI is modeled by a triply modulated cosine function. Denote μ_k, α_k and ϕ_k, the mean, the amplitude and the initial phase of the triply modulated cosine function at instant k respectively. Recall that the number of the pixels of interests is 38. The state of the system for the n^{th} pixel x_k^n at instant k is defined as

$$x_k^n = \begin{pmatrix} \mu_k^n & \alpha_k^n & \phi_k^n \end{pmatrix}^T \tag{2}$$

The dynamic of the state is supposed to be governed by a zero mean white Gaussian noise w_k with process noise covariance matrix Q, i.e. $w_k^n \sim \mathcal{N}(0, Q)$, $\forall k \in [1, 2, \ldots, 587]$ and $\forall n \in [1, 2, \ldots, 38]$.

$$x_{k+1}^n = x_k^n + w_k^n \tag{3}$$

In section 3, we will explain how to choose the process noise covariance matrix Q.

The estimated state at instant k for the n^{th} pixel is defined as

$$\hat{x}_k^n = \begin{pmatrix} \hat{\mu}_k^n & \hat{\alpha_k}^n & \hat{\phi}_k^n \end{pmatrix}^T \tag{4}$$

Denote y_k^n the NDVI measurement at instant k for the n^{th} pixel which is modelled by

$$y_k^n = \mu_k^n + \alpha_k^n cos(\omega k + \phi_k^n) + v_k^n \tag{5}$$

where $\omega = 2\pi f$ is the angular frequency where f is the frequency of the measurement acquisition. For the sake of simplification $f = 8/365$ is used here seeing that the rain-fed rice has a single crop cycle. v_k^n is the measurement noise at instant k. We assume that v_k^n is a zero mean white Gaussian noise with standard deviation of σ_v, i.e. $v_k^n \sim \mathcal{N}(0, R = \sigma_v^2)$, $\forall k \in [1, 2, \ldots, 587]$ and $\forall n \in [1, 2, \ldots, 38]$. R is the measurement noise covariance matrix. In section 3.3.2, we will explain how to choose σ_v.

The estimated NDVI at instant k for the n^{th} pixel $\hat{y}_k^n$ is calculated from the estimated state $\hat{x}_k^n$ as

$$\hat{y}_k^n = \hat{\mu}_k^n + \hat{\alpha}_k^n cos(\omega k^n + \hat{\phi}_k^n) \tag{6}$$

2.3. Estimated Cultivation Date Computation

In [Chumkesornkulkit et al., 2013], the estimated mean of the triply modulated cosine function μ_k^n was considered to represent long term vegetation changes such as deforestation and climate changes. Only the term $\alpha_k^n cos(\omega k + \phi_k^n)$, called the seasonal term was considered to be related to the growth of the rice. The estimated cultivation date was set to be equal to the date for which the seasonal term is greater than a predefined threshold τ. In [Chumkesornkulkit et al., 2013], the minimum mean estimation error was achieved when the threshold $\tau = -0.02$ was chosen for the single crop rice. A negative threshold may represent when the farmer cut off other crops on the field, before the rice cultivation.

In this paper, the trend of the optimal thresholds, which are the values of the threshold giving zero estimation error, will be studied for the first time using the real estimation dates from the ground truths. The detail will be given in section **??**.

3. The Extended Kalman Filter: Equations, Tuning and Initialization

3.1. The Extended Kalman Filter Equations

The Extended Kalman Filter (EKF) makes use of the formulations of the Kalman Filter (KF) which is the best linear estimator when process noise w_k and measurement noise v_k are uncorrelated and white even though they are non Gaussian [Simon, 2010]. The KF is the best estimator in the sense that it gives the minimum achievable error variance for a linear system. The proof of the KF is provided in [Kalman, 1960].

The EKF generates estimates for nonlinear systems by linearizing the system at the current estimate and then applies the Kalman Filter equations to the linearized system. Denote $\hat{x}_k^-$ the a priori estimate at instant k, P_k^- the a priori estimation error covariance matrix, $\hat{x}_k$ the a posteriori estimate and P_k the a posteriori error covariance matrix. The algorithm of the EKF can be divided into two stages: prediction and update as follow [Suwantong, 2014]:

Prediction

$$\hat{x}_k^- = f(\hat{x}_{k-1}) \tag{7}$$

$$P_k^- = Q_{k-1} + \hat{F}_{k-1} P_{k-1} \hat{F}_{k-1}^T \tag{8}$$

$\hat{F}_{k-1}$ is the Jacobian of f evaluated at $\hat{x}_{k-1}$, i.e.

$$\hat{F}_{k-1} = \left.\frac{\partial f}{\partial x}\right|_{x=\hat{x}_{k-1}} \tag{9}$$

Update

$$\hat{x}_k = \hat{x}_k^- + K_k(y_k - h(\hat{x}_k)) \tag{10}$$

$$P_k = P_k^- - K_k S_k K_k^T \tag{11}$$

where

$$S_k = \hat{H}_k P_k^- \hat{H}_k^T + R_k \tag{12}$$

$$K_k = P_k^- \hat{H}_k^T S_k^{-1} \tag{13}$$

$\hat{H}_k$ is the Jacobian of h evaluated at $\hat{x}_k^-$, i.e.

$$\hat{H}_k = \left.\frac{\partial h}{\partial x}\right|_{x=\hat{x}_k} \tag{14}$$

3.2. Filter Initialization

3.2.1. A priori Initial State

Denote the a priori initial estimated state for the n^{th} pixel as $\hat{x}_1^{n-} = \begin{pmatrix}\hat{\mu}_1^{n-} & \hat{\alpha}_1^{n-} & \hat{\phi}_1^{n-}\end{pmatrix}^T$.

As in [Chumkesornkulkit et al., 2013], the a priori initial estimated mean $\hat{\mu}_1^{n-}$ and amplitude $\hat{\alpha}_1^{n-}$ for the n^{th} pixel are computed using the NDVI measurements y_k^n, $\forall k \in [1, 2, \ldots, 587]$, i.e. for the n^{th} pixel

$$\hat{\mu}_1^{n-} = \sum_{k=1}^{N} \frac{y_k^n}{N} \tag{15}$$

$$\hat{\alpha}_1^{n-} = \frac{\max_k y_k^n - \min_k y_k^n}{2} \tag{16}$$

The a priori initial estimated $\hat{\phi}_1^{n-} = 120°$ is chosen $\forall n \in [1, 2, \ldots, 38]$ supposing that the start date which is on the 18th of February 2000 represents the stage after the rice harvest.

3.2.2. A priori Initial Error Covariance Matrix

For the sake of simplicity, the a priori initial estimation error $\tilde{x}_1^{n-} = \hat{x}_1^{n-} - x_1^n$ is modelled by a zero-mean white Gaussian noise $\tilde{x}_1^{n-} \sim \mathcal{N}(0, P_1^-)$, $\forall n \in [1, 2, \ldots, 38]$. P_1^- is called the a priori initial error covariance matrix. Each component of the a priori initial estimate is supposed to be uncorrelated, hence P_1^- is diagonal. High values of the standard deviation of the a priori initial estimation error of each component are chosen to let the EKF explore large domain of possible values until it finds ones that correspond to the measurements the most.

$$P_1^- = \begin{pmatrix} 1^2 & 0 & 0 \\ 0 & 1^2 & 0 \\ 0 & 0 & (\frac{10}{365} \cdot 2\pi)^2 \end{pmatrix} \tag{17}$$

3.3. Filter Tuning

3.3.1. Process Noise Covariance Matrix

We suppose that the process noise of each component is uncorrelated. Denote σ_{w_μ} the standard deviation of the noise on the mean, σ_{w_α} the standard deviation of the noise on the amplitude and σ_{w_ϕ} the standard deviation of the noise on the initial phase. The process noise covariance matrix for the n^{th} pixel is therefore described as

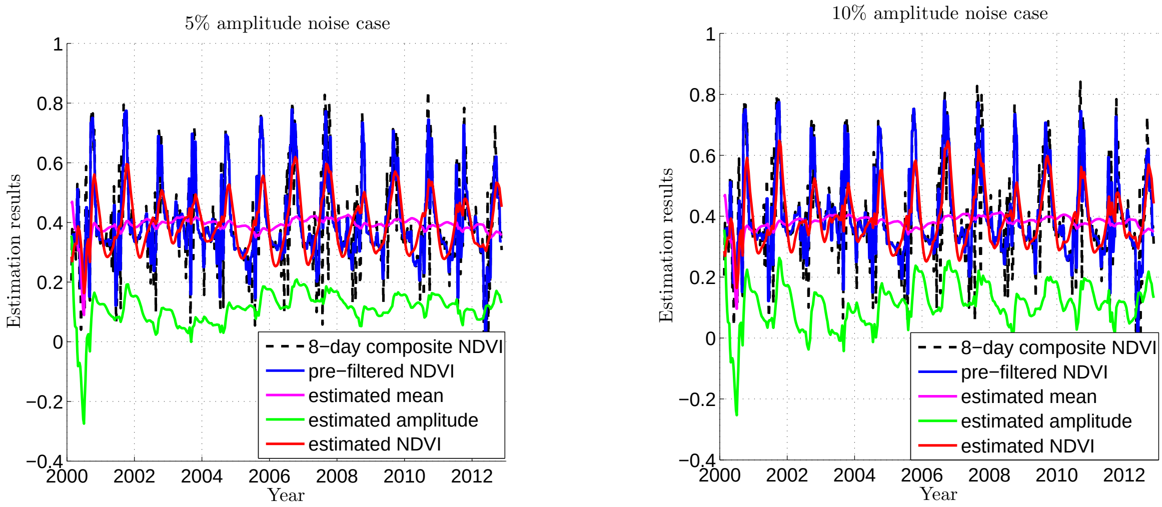

Figure 3: Examples of the estimated mean $\hat{\mu}$, amplitude $\hat{\alpha}$ and NDVI $\hat{y}$ in time for the 5% amplitude noise case (left) and for the 10% amplitude noise case (right).

$$Q^n = diag\left((\sigma^n_{w_\mu})^2 \quad (\sigma^n_{w_\alpha})^2 \quad (\sigma^n_{w_\phi})^2\right) \tag{18}$$

We suppose that $\sigma^n_{w_\mu}$ is at 2% of the a priori initial estimated mean $\hat{\mu}^{n-}_1$, i.e. $\sigma_{w_\mu} = 0.02 \cdot \hat{\mu}^{n-}_1$. Two choices of $\sigma^n_{w_\alpha}$ are tested: $\sigma^n_{w_\alpha} = 0.1 \cdot \hat{\alpha}^{n-}_1$ and $\sigma^n_{w_\alpha} = 0.05 \cdot \hat{\alpha}^{n-}_1$. The standard deviation of the noise on the a priori estimate of the initial phase is chosen to be equivalent to two days, i.e. $\sigma^n_{w_\phi} = (2/365) \cdot 2\pi$ in radian, $\forall n \in [1, \ldots, 38]$.

The study case in which $\sigma^n_{w_\alpha} = 0.05 \cdot \hat{\alpha}^{n-}_1$ will be called as the 5% amplitude noise case and the one in which $\sigma^n_{w_\alpha} = 0.1 \cdot \hat{\alpha}^{n-}_1$ will be called as the 10% amplitude noise case.

3.3.2. Measurement Noise Covariance Matrix

We observe in figure 2 that the noise level depends on the season. The noise is high during the middle of the year which corresponds to the raining season where there are lots of clouds. For the sake of simplicity, in this study a constant value of the standard deviation of the measurement noise $\sigma_v = 0.3^2$ is chosen.

4. Result Analysis

4.1. Estimated States and Measurements

Before plotting the estimate states and measurements, we define the year vector t_{year} which is constructed with the same number of instants than the number of the measurements for a pixel, i.e. 587 instants. The start date is on the 18th of February. Therefore, the value of the first component of the year vector $t_{year}(1) = 2000 + (31+18)/366$ since 2000 is a leap year. When the NDVI measurements are collected all year long, the total number of the measurement for the year will be 46. The start date of the NDVI measurement acquisition from MODIS is on the 1st of January of each year.

Example of the estimated states and measurements are presented in figure 3 for the 5% amplitude noise case (left) and for the 10% amplitude noise case (right). For the 10% amplitude noise case, we can observe high variation of the estimated amplitude $\hat{\alpha}$ to represent the fact that the amplitude of the measured NDVI is lower when the measured NDVI is below the mean than when it is above the mean. The estimated amplitude in the 10% case fits the estimated of the measured NDVI more but it also induces more noisy estimated NDVI.

4.2. Cultivation Date Estimation

For each pixel, the estimated cultivation date for the n^{th} pixel $\hat{t}^n_{cult}$ is computed from a given threshold τ and the estimated seasonal NDVI as shown in figure 4 (left). First, we search for the first year vector instant where the estimated seasonal term $\hat{y}^n_{sea} > \tau$. Denote, k_τ as this first instant. Then, we compute the equation of the line between

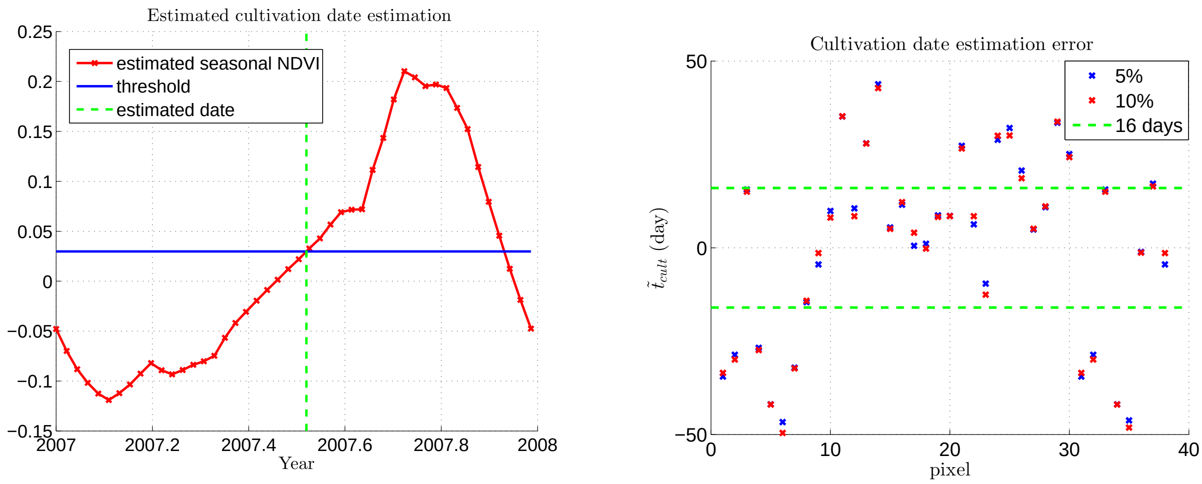

Figure 4: The estimated cultivation date $\hat{t}^n_{cult}$ in green is defined as the x-component of the first intersection between the estimated seasonal NDVI in red and the threshold line in blue (left). Cultivation Date Estimation Error in day (right).

Estimated Cultivation Date	5% amplitude noise	10% amplitude noise	Tuning in [Chumkesornkulkit et al., 2013]
Threshold τ	-0.005	-0.001	-0.02
Mean error	0.13	-0.12	7.4
Standard deviation of error	25.55	25.65	27.2
Mean of the absolute error	20.98	20.90	24.6

Table 1: Cultivation date estimation errors for each study case.

the coordinates $(t_{year}(k_\tau), \hat{y}^n_{sea,k_\tau})$ and $(t_{year}(k_\tau - 1), \hat{y}^n_{sea,k_\tau - 1})$. The estimated cultivation date $\hat{t}^n_{cult}$ is defined as the x-component of the intersection between the threshold line and the constructed line.

Figure 4 (right) shows the errors of the estimated cultivation date $\tilde{t}^n_{cult} = \hat{t}^n_{cult} - t^n_{cult}$ for each pixel n for both amplitude noise cases. The cultivation date estimation errors in mean, standard deviation and in mean of the absolute value are given in table 1. Three cases are studied: the case of 5% amplitude noise where the threshold $\tau = -0.005$ is chosen, the case of 10% amplitude noise where the threshold $\tau = -0.001$ is chosen and the case when the same tuning and threshold as in [Chumkesornkulkit et al., 2013] is used.

The cultivation date estimation errors in absolute value for the 5% and the 10% amplitude noise are both 21 days which is better than the one obtained using the tuning method in [Chumkesornkulkit et al., 2013] which gives 25 days of error.

5. Conclusions and Perspectives

In this paper, we propose a method for estimating the cultivation date of single crop rice using 8-day composite normalized difference vegetation index (NDVI) from the Moderate Resolution Imaging Spectroradiometer (MODIS). The studied areas are in Thailand, whose climate is tropical, and hence the measured NDVI data from MODIS in this area is very noisy due to high cloud density. A good estimation technique is therefore needed to cope with this high measurement noise. In this study, the NDVI data from early 2000 until late 2012 are used.

The NDVI is modelled by a triply modulated cosine function with the mean, the amplitude and the initial phase as the state of the system which is dynamic. We use the NDVI measurements as inputs for the Extended Kalman Filter to estimate the state of the system. The dynamics of the state is governed by a zero-mean white Gaussian noise. Contrary to the existing work in [Chumkesornkulkit et al., 2013] where the brute force algorithm where large numbers of tuning parameters of the EKF and threshold were tested until the smallest mean error was obtained, we propose a method for tuning the parameters of the EKF. A formulation for the a priori initial error covariance matrix is also provided in this paper. The estimated cultivation dates are compared to the real cultivation dates collected by the rice government of Thailand which are in 2004 and 2007. There are 38 study areas (pixels) in total.

The a priori initial estimated mean for a specific pixel is chosen as the mean of all the NDVI data we have on the pixel. The a priori initial estimated amplitude is chosen as half of the difference between the maximum NDVI and the minimum NDVI data we have on the pixel. The a priori initial estimated initial phase is chosen equal to $120°$ by supposing that the start date is around the harvest time.

The standard deviation of the process noise on the mean is set to be 2% of the a priori initial estimated mean for the pixel. For the standard deviation of the process noise on the amplitude, two study cases are considered, in one case it is set to be 5% of the a priori initial estimated amplitude for the pixel. In the other case, it is set to be 10% of the a priori initial estimated amplitude for the pixel. The standard deviation of the noise on the a priori estimate of the initial phase is chosen to be equivalent to two days.

The standard deviation of the measurement noise is chosen by converting the NDVI measurements using the wavelet decomposition to filter the NDVI into the high-pass part, which represents the noise, and the low-pass part, which represents the signal. We reconstructed the high-pass part to visualize the noise level. The low-pass-part is reconstructed and call, the pre-filtered NDVI, which is given to the Extended Kalman Filter (EKF) for the cultivation date estimation. In this study, we choose the standard deviation of the measurement noise to be 0.3 which is realistic compared to the range of the NDVI which is between 0 and 1 unlike in [Chumkesornkulkit et al., 2013] where it was chosen to be 20.

The cultivation date estimation errors in absolute value for the 5% and the 10% amplitude noise are found to be of 21 days which is better than the one obtained using the tuning method in [Chumkesornkulkit et al., 2013] which gives 25 days of error.

For future works, the trend of the optimal threshold should be derived using more study areas. We have observed in this paper that for the year where the estimated amplitudes of the NDVI are high, the optimal thresholds tend to be high and noisy. This may due to the higher level of the rice greenness resulting from more rainfalls. Higher noise therefore induces higher cloud density. Moreover, the period of the single crop rice is set to be equal to 365 in this study. One could also consider the period of the crop as state variable for further development. Other measurements such as EVI from MODIS, dual polarization image from TerraSAR-X as used in [Vicente-Guijalba et al., 2014] for crop phenology estimation and polarimetric radar data from Radarsat-2 as used in [Vicente-Guijalba et al., 2015] for real time crop phenology estimation should also be combined with the NDVI data to increase the estimated cultivation date accuracy.

References

[Chu et al., 2014] Chu, L., Liu, G.-h., Huang, C., and Liu, Q.-s. (2014). Phenology detection of winter wheat in the yellow river delta using modis ndvi time-series data. In *Agro-geoinformatics (Agro-geoinformatics 2014), Third International Conference on*, pages 1–5. IEEE.

[Chumkesornkulkit et al., 2013] Chumkesornkulkit, K., Kasetkasem, T., Rakwatin, P., Eiumnoh, A., Kumazawa, I., and Buddhaboon, C. (2013). Estimated rice cultivation date using an extended kalman filter on modis ndvi time-series data. In *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2013 10th International Conference on*, pages 1–6. IEEE.

[Kalman, 1960] Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1).

[Li et al., 2014] Li, L., Friedl, M. A., Xin, Q., Gray, J., Pan, Y., and Frolking, S. (2014). Mapping crop cycles in china using modis-evi time series. *Remote Sensing*, 6(3):2473–2493.

[Simon, 2010] Simon, D. (2010). Kalman filtering with state constraints. A survey of linear and nonlinear algorithms. *Control Theory Applications, IET*, 4(8).

[Solano et al., 2010] Solano, R., Didan, K., Jacobson, A., and Huete, A. (2010). Modis vegetation index users guide (mod13 series). *Vegetation index and phenology lab*.

[Suwannachatkul et al., 2014] Suwannachatkul, S., Kasetkasem, T., Chumkesornkulkit, K., Rakwatin, P., Chanwimaluang, T., and Kumazawa, I. (2014). Rice cultivation and harvest date identification based on a hidden markov model. In *Electrical Engineering/Electronics, Computer, Telecommunications, and Information Technology (ECTI-CON), 2014 11th International Conference on*, page 16.

[Suwantong, 2014] Suwantong, R. (2014). *Development of the Moving Horizon Estimator with Pre-Estimation (MHE-PE). Application to Space Debris Tracking during the Re-Entries.* PhD thesis, Supélec.

[Vicente-Guijalba et al., 2014] Vicente-Guijalba, F., Martinez-Marin, T., and Lopez-Sanchez, J. (2014). Crop phenology estimation using a multitemporal model and a kalman filtering strategy. *Geoscience and Remote Sensing Letters, IEEE*, 11(6):10811085.

[Vicente-Guijalba et al., 2015] Vicente-Guijalba, F., Martinez-Marin, T., and Lopez-Sanchez, J. (2015). Dynamical approach for real-time monitoring of agricultural crops. *Geoscience and Remote Sensing, IEEE Transactions on*, 53(6):32783293.