



Kalman filter method for generating time-series synthetic Landsat images and their uncertainty from Landsat and MODIS observations



Fuqun Zhou*, Detang Zhong

Canada Center for Remote Sensing, Canada Centre for Mapping and Earth Observation, Natural Resources Canada, 6th floor, 560 Rochester Street, Ottawa, ON K1A 0E4, Canada

ARTICLE INFO

Editor: Jing M. Chen

Keywords:
Kalman filter
Landsat
MODIS
Image fusion
Uncertainty
Time-series

ABSTRACT

The Landsat program, since its commencement in 1972, has acquired millions of images of our planet. Those images are one of the most valuable Earth Observation resources for local, regional and global land surface monitoring and study due to their moderate spatial resolution and rich spectral information. However, their applications are impeded largely by their relatively low revisit frequency and cloud contamination on images. In order to improve their usability, a number of studies have been conducted to blend Landsat images with Moderate Resolution Imaging Spectroradiometer (MODIS) images to take merits of the two sensors. All blending models reported that they can predict synthetic Landsat images with various degrees of accuracy. However, only a couple of models reported that they can explicitly estimate uncertainty for their blended images.

In this study, we propose a new surface reflectance blending model based on a Kalman Filter algorithm (Kalman Filter Reflectance Fusion Model - KFRFM) to predict time-series synthetic Landsat images from Landsat and MODIS images, and simultaneously to estimate uncertainty of the predicted synthetic images to quantify the quality of the synthetic images. Using the model, we predicted a time-series of 38 synthetic Landsat images with a temporal interval of 4 days for a vegetation growing season spanning about 6 months from Nadir Bi-directional Reflectance Distribution Function adjusted MODIS product (MCD43A4), and their corresponding uncertainties. From this time-series, we calculated five vegetation indices involving all the spectral bands of the synthetic images, and compared them to those from Landsat observations. The results demonstrated that the proposed method is able to produce high quality synthetic Landsat images to meet various application demands for higher spatial and temporal resolution images. Uncertainty analysis reveals that cropland has the largest uncertainty followed by grassland while forests have the smallest uncertainties among the seven vegetation land cover types of the study area. For performance evaluation, we compared KFRFM to several published models. The comparison results reveal that KFRFM performs the best based on the assessed image quality indices.

1. Introduction

Since the launch of the first Landsat satellite in 1972, the satellite series from Landsat-1 to Landsat-8 have collected millions of images of our planet, which has become the legacy of Earth Observations. These images acquired over more than four decades through the course of the missions provide a unique and rich resource for study of global land surface. The data have been widely used for various applications in forestry, agriculture, geology, mapping, and regional planning, as well as for emergency and disaster response and management. The usage of Landsat images has got more momentum since 2008 due to United States Geological Survey (USGS) open data policy (Woodcock et al., 2008).

Although Landsat images possess many merits, for example,

moderate spatial resolution and rich spectral information, their applications are largely impeded by its relatively low revisit frequency. With a revisit cycle of every 16 days compounded with various unfavorable weather conditions, for example, in most regions of Canada, the chance to get clear-sky images for a full coverage is small within a vegetation growing season (Zhou and Zhang, 2013). To supplement the dataset of clear-sky Landsat images, a number of image blending methodologies have been developed to produce synthetic Landsat images by fusion of Landsat and MODIS (Moderate Resolution Imaging Spectroradiometer) images over the last decade or so.

Broadly, there are mainly three types of algorithms for image blending which are well summarised by Zhu et al. (2016). One of them is image blending based on weighted functions. The pioneer work on this is the spatial and temporal adaptive reflectance fusion model

* Corresponding author.

E-mail addresses: fuqun.zhou@canada.ca (F. Zhou), detang.zhong@canada.ca (D. Zhong).

(STARFM) by Gao et al. (2006), and then the spatial temporal adaptive algorithm for mapping reflectance change (STAARCH) by Hilker et al. (2009a), and the enhanced STARFM (ESTARFM) by Zhu et al. (2010). The weighted function based blending models assume that the reflectance images from the two sensors used for blending have a good correlation and that the reflectance changes on both images are consistent for homogenous pixels (which are covered by only one land cover type). Hence, the change of reflectance from a coarser resolution image (e.g., MODIS) acquired at a prediction date can be added to the finer resolution image (e.g., Landsat image) to predict the finer resolution image at the prediction date. To offset the constraint of the homogenous pixel condition, multiple pixels with similar spectral information from its surrounding pixels are selected and weighed for calculation of the temporal reflectance change.

The second type of the algorithms is based on linear spectral mixing theory which states that the reflectance at a pixel of an image is a linear combination of the reflectance of its endmembers (e.g., different land cover types) present within the pixel according to their individual reflectance and area proportion. According to this, the reflectance of the endmembers such as pixels in a finer spatial resolution image can be derived by unmixing the reflectance of a coarse spatial resolution pixel. Thereafter, the derived reflectance change for a prediction period is added to the finer resolution pixel to predict the finer resolution image at the prediction date. A variety of unmixing methods are developed to achieve this task. Examples of such kind of models can be found in the literatures (Zhukov et al., 1999; Zurita-Milla et al., 2008; Wu et al., 2012; Amorós-López et al., 2013; Zhang et al., 2013; Huang and Zhang, 2014; Gevaert and García-Haro, 2015; Ma et al., 2018; Zhu et al., 2016; Zhong and Zhou, 2018).

The third method is dictionary-pair learning based. A dictionary pair learning-based algorithm is, in general, to train dictionaries by establishing a correspondence between structures of two counterpart images (e.g., Landsat and MODIS) through sparse representation from some image patches, and to reconstruct finer-resolution images via the learned dictionaries and sparse coding. Huang and Song (2012) proposed the sparse representation-based spatiotemporal reflectance fusion model (SPSTFM) to predict phenological changes and land cover changes uniformly. To improve the model's computation efficiency, Song and Huang (2013) presented a continual two-stage framework based on one-pair image learning by employing a sparse representation procedure that includes dictionary training and sparse coding for derivation of the super-resolved MODIS images. Then the Landsat image at date t_1 and the two super-resolved MODIS images at t_1 and t_2 (the prediction date) were used to predict the synthetic Landsat image at t_2 by using a high pass modulation technique. Recently, Rao et al. (2015) developed a method based on dictionary-learning algorithm to predict a synthetic LISS III (Linear Imaging and Self-Scanning Sensors III) image from a single Advanced Wide-Field Sensor (AWIFS)-LISS III image pair. In their study, they proposed a novel approach to minimize spurious spatial discontinuities that were not dealt with by previous studies. Wu et al. (2015) advanced a spatiotemporal reflectance fusion model (EBSPTM) using an error-bound-regularized sparse coding to adapt the possible perturbations of an overcomplete dictionary and formulate the dictionary perturbations to be a sparse elastic net regression problem. They also proposed semi-coupled dictionary learning to address the differences between the high-spatial-resolution and low-spatial-resolution images. Their method has been validated by fusion of Landsat ETM+ and MODIS images.

All these methods are able to produce synthetic images with various advantages and disadvantages. However, a major shortcoming of these methods is lacking uncertainty estimation of their blended images. Most of these studies used some quality indices or a scatter plot against true Earth Observations to evaluate quality of the blended images, i.e., to illustrate how similar a blended image is to its corresponding real observation image. Recently, Wang and Huang (2017) presented a blending model - A Rigorously-Weighted Spatiotemporal Fusion Model

with Uncertainty Analysis (RWSTFM) based on geostatistical approach known as Kriging interpolation (Matheron, 1967) to blend the surface reflectance of MODIS and Landsat-5 Thematic Mapper (TM) imagery. For the first time, this model presented a variance estimation for its blended images as a quality indicator, which is a step forward in terms of quantitative quality assessment of the output. However, a drawback of the method is its low computation efficiency. To compute the semivariogram of every pixel for weight calculation is very time-consuming which may limit its practical applications for fusion exercise of time-series images over a large coverage or in near real time. Most recently, Zhong and Zhou (2018) developed a Prediction Smooth Reflectance Fusion Model (PRSFIM) for predicting synthetic Landsat images from MODIS observations. The model can incorporate uncertainties of input images for its prediction and estimate uncertainties of the prediction results. Based on the estimated uncertainties of forward and backward predictions, a smooth filter of variance weighted average is also developed to improve the model output. However, the uncertainty analysis of the predicted images is preliminary and limited, especially for its impacts on applications of the predicted images. Apparently, more studies on uncertainty of blending models are needed.

In this study, we developed a new image fusion model based on Kalman Filter method (Kalman, 1960; Welch and Bishop, 2006) for blending Landsat-8 Operational Land Imager (OLI) and MODIS images to produce a time-series of synthetic Landsat images and their uncertainty estimate, as well as to evaluate their applications for generation of vegetation indices. As a well-known sensor fusion and data fusion algorithm, Kalman filter has been widely used not only for guidance, navigation, and control of vehicles such as aircraft and spacecraft, etc., but also in time series data analysis and integration in many fields such as signal processing and econometrics, etc. In remote sensing, we can also find its applications in correction of atmospheric effects in satellite images (Arbel et al., 2004), image fusion of multisource images (Sun and Deng, 2004; Hu et al., 2013), integration of remote sensing data with ecosystem models to retrieve soil temperature (Huang et al., 2008), surface BRDF parameters (Samain et al., 2008), soil moisture (Huang et al., 2008), change detection (Kleynhans et al., 2011), and ecosystem productivity (Quaife et al., 2008), or monitor crop phenology (Vicente-Guijalba et al., 2014) and generate time series of Medium-Resolution NDVI (Normalized Difference Vegetation Index) Images (Sedano et al., 2014). However, all of these applications focused on different problems from the issues to be solved by our study. The most relevant study was conducted by Hilker et al. (2009b) for generating dense time series synthetic Landsat data using STARFM (Gao et al., 2006). However, in our study, we not only predict a time series of synthetic Landsat images using the proposed method, but also estimate their uncertainty. In addition to being as a quality indicator of the predicted images, the uncertainty is also used for smoothing the synthetic Landsat images for improved prediction.

Majority of the blending exercises explored fusion of Landsat and MODIS images as the two sensors share some similarity of the spectral information, which makes the blending of the two types of images very promising. Table 1 lists the spectral bands and bandwidths of Landsat-7 ETM+, Landsat-8 OLI and the corresponding MODIS bands for land study. As MODIS has a daily revisit capability, it has a much bigger chance to get enough clear-sky images for time-series image blending exercise. For these reasons, in the first case study, we use MODIS Nadir Bidirectional Reflectance Distribution Function (BRDF) adjusted reflectance daily product (MCD43A4) to blend Landsat-8 OLI images to demonstrate our algorithms and results. In the second study case, we use Landsat-7 ETM+ and MODIS (MOD09GA) images for image fusion model assessment and comparison in a heterogeneous landscape. The remainder of this paper is organized as follows. Following the introduction, in Section 2, we will briefly introduce Kalman Filter model, as well as Kalman Filter model formulation for image blending; Section 3 will describe the study areas and the data used for the experiment and test of the model. In Section 4, we will present the experiment results

Table 1

Band and bandwidth list of Landsat-8 OLI and MODIS (MCD43A4).

Landsat-7 ETM + band	Landsat-7 ETM + bandwidth (nm)	Landsat-8 OLI band	Landsat-8 OLI bandwidth (nm)	MODIS land band	MODIS bandwidth (nm)	Landsat band name (short name)
1	452–514	2	452–512	3	459–479	Blue
2	519–601	3	533–590	4	545–565	Green
3	631–692	4	636–673	1	620–670	Red
4	772–898	5	851–879	2	841–876	Near-infrared (NIR)
5	1547–1748	6	1566–1651	6	1628–1652	Shortwave infrared 1 (SWIR-1)
7	2065–2346	7	2107–2294	7	2105–2155	Shortwave infrared 2 (SWIR-2)

and discussion. Finally we will give conclusions in [Section 1](#).

2. Methodology

2.1. Overview of Kalman filter algorithm

In general, a Kalman filter can be considered as a recursive inference algorithm that integrates observations, models and their respective uncertainties to estimate the state of a process by minimizing the mean of the squared errors ([Welch and Bishop, 2006](#)). It works usually in a two-step process. The first step is “prediction” in which the Kalman filter produces estimates of the current state variables, along with their uncertainties. The second step is “update” in which these estimates are updated with new observed measurements (necessarily corrupted with some amount of error, including random noise) using a weighted average, with more weight being given to estimates with higher certainty.

In a mathematical expression, a Kalman filter is usually formulated in a time frame. It assumes that the true state at time k is evolved from the state at time $(k - 1)$ according to

$$\mathbf{x}_k = F_k \mathbf{x}_{k-1} + B_k \mathbf{u}_k + w_k \quad (1)$$

where F_k is a state transition model which is applied to the previous state \mathbf{x}_{k-1} ; B_k is a control-input model which is applied to the control vector \mathbf{u}_k ; w_k is process noise which is assumed to be drawn from a zero mean multivariate normal distribution N , with covariance, Q_k ; $w_k \sim N(0, Q_k)$.

At time k an observation (or measurement) \mathbf{z}_k of the true state \mathbf{x}_k is made according to

$$\mathbf{z}_k = H_k \mathbf{x}_k + v_k \quad (2)$$

where H_k is an observation model which maps the true state space into the observed space and v_k is observation noise which is assumed to be zero mean Gaussian white noise with covariance R_k ; $v_k \sim N(0, R_k)$.

Following the two-step process mentioned above, we have the predicted (a priori) state estimate $\hat{\mathbf{x}}_{k|k-1}$ and covariance $P_{k|k-1}$ in the prediction step:

$$\hat{\mathbf{x}}_{k|k-1} = F_k \hat{\mathbf{x}}_{k-1|k-1} + B_k \mathbf{u}_k \quad (3)$$

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_k \quad (4)$$

and the updated (a posteriori) state estimate $\hat{\mathbf{x}}_{k|k}$ and covariance $P_{k|k}$ in the update step:

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + K_k \tilde{\mathbf{y}}_k \quad (5)$$

$$P_{k|k} = P_{k|k-1} - K_k S_k K_k^T \quad (6)$$

where

$$\tilde{\mathbf{y}}_k = \mathbf{z}_k - H_k \hat{\mathbf{x}}_{k|k-1} \quad (\text{the innovation or measurement pre-fit residual}) \quad (7)$$

$$S_k = R_k + H_k P_{k|k-1} H_k^T \quad (\text{the innovation or pre-fit residual covariance}) \quad (8)$$

$$K_k = P_{k|k-1} H_k^T S_k^{-1} \quad (\text{the optimal Kalman gain}). \quad (9)$$

A great property of Kalman Filter is that a large observation or measurement noise R_k results in a low Kalman gain, which gives more weight to the model process. Conversely, large process covariance $P_{k|k-1}$ results in a high Kalman gain and more weight for the new measurements ([Sedano et al., 2014](#)).

2.2. Formulation of KFRFM

A Kalman filter requires a system model that describes the system's state variable transition, the state variables' covariance, and an observation model for updating the priori state variable estimate. For building the Kalman filter model to blend Landsat and MODIS images, we require two pairs of clear-sky Landsat and MODIS images, say at dates t_0 and t_n (t_n is later than t_0) and a time series of MODIS images at date t_k (t_k is between t_0 and t_n) for prediction of a time series of synthetic Landsat images at dates t_k ($k = 1, 2, \dots, n - 1$).

2.2.1. System model for reflectance variable transition

The KFRFM system model describes the reflectance transition from date k to $k + 1$ as captured by Landsat images. If we define surface reflectance of a Landsat image and its change rate as the system state variables, we can use either the Landsat image at the start date t_0 or the predicted synthetic Landsat image at t_k as the state estimate of the reflectance. Their covariance is the observation uncertainty at t_0 or the uncertainty estimate of the synthetic Landsat image at the prediction date t_k , respectively. The state of the reflectance change rate between dates t_{k-1} and t_k can be estimated from two MODIS images at t_{k-1} and t_k for every pixel. The estimated reflectance change rate plus the Landsat image at start date t_0 or the synthetic Landsat image at date t_{k-1} form a prediction of the reflectance state variable at t_k in the “prediction” step. In the update step, the predicted synthetic Landsat image (i.e., the state of the reflectance variable) will be updated by the output of the observation model as described in [Section 2.2.2](#). In general, when a new MODIS image at t_{k+1} is available, the new change rate from the reflectance difference of the two MODIS images at t_k and t_{k+1} with the updated reflectance state variable prediction at t_k become the new prediction of the state variable for date t_{k+1} . This process is progressed until a synthetic Landsat image at $t_{k=n-1}$ is predicted. At this point, t_n is taken as the new start date t_0 , and the date of a new pair of clear-sky Landsat and MODIS images later than t_n becomes the end date t_n for synthetic Landsat image prediction within the new period $[t_0, t_n]$ as described above. By this way, a time series of synthetic Landsat images can be generated. With the input uncertainty of Landsat and MODIS images as their observation accuracies, the covariance of the generated synthetic Landsat images can be estimated by the Kalman filter algorithm described in [Section 2.1](#).

A challenge with prediction of the system state variables is that the reflectance change rate from t_{k-1} to t_k for a synthetic Landsat pixel is not directly obtainable because the only measurement about it is limited to MODIS observations at a much lower spatial resolution. They are estimable if there are no extra measurements. To address this issue, it is

reasonable to assume that the reflectance change rate is the same for a classified region or cluster of a Landsat image. By introducing a reflectance change rate for each classified region or cluster, we can describe KFRFM system transition model for predicting the reflectance state variable \hat{r} from time t_{k-1} to t_k of band b at a pixel (x, y) belonging to a cluster c as follows:

$$\hat{r}_{k/k-1}(x, y, c, b) = r_{k-1}(x, y, c, b) + (t_k - t_{k-1})\dot{r}_{k-1}(x, y, c, b) \quad (10)$$

where $\dot{r}_{k-1}(x, y, c, b)$ is the reflectance change rate on the date t_{k-1} , and $r_{k-1}(x, y, c, b)$ is the true Landsat reflectance (when $k = 1$) or updated synthetic Landsat reflectance (when $k > 1$ and $< n$).

To simplify the mathematical expression, we omit the pixel's location (x, y) and the band number b , and sort observed Landsat or synthetic Landsat surface reflectance at date k into l clusters for band b as

$$\begin{bmatrix} \hat{r}_1 \\ \vdots \\ \hat{r}_c \\ \vdots \\ \hat{r}_l \end{bmatrix}_{k/k-1} = \begin{bmatrix} r_1 \\ \vdots \\ r_c \\ \vdots \\ r_l \end{bmatrix}_{k-1} + \begin{bmatrix} \Delta t & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \Delta t \end{bmatrix}_{n_l \times 1} \begin{bmatrix} \dot{r}_1 \\ \vdots \\ \dot{r}_c \\ \vdots \\ \dot{r}_l \end{bmatrix}_{k-1} \quad (11)$$

where \hat{r}_c is an $n_c \times 1$ vector that includes all pixel values belonging to cluster c ($c = 1, 2, \dots, l$). The total number of Landsat pixels of the processed image is $N = \sum_1^l n_c$. \dot{r}_c is reflectance change rate of cluster c . $\Delta t = t_k - t_{k-1}$ is the time interval between the two observation dates.

Define the state variable vector r , its change rate, and its corresponding coefficient matrix as

$$\hat{r} = \begin{bmatrix} \hat{r}_1 \\ \vdots \\ \hat{r}_c \\ \vdots \\ \hat{r}_l \end{bmatrix}, \dot{r} = \begin{bmatrix} \dot{r}_1 \\ \vdots \\ \dot{r}_c \\ \vdots \\ \dot{r}_l \end{bmatrix}, F_r = \begin{bmatrix} \Delta t & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \Delta t \end{bmatrix}, \quad (12)$$

then we have following Kalman filter prediction by Eq. (13) and its corresponding uncertainty estimation by Eq. (14).

$$[\hat{r}]_{k/k-1} = [r]_{k-1} + [F_r \dot{r}]_{k-1} \quad (13)$$

$$[\hat{P}_{rr}]_{k/k-1} = [P_{rr}]_{k-1} + [F_r P_{rr} F_r^T]_{k-1} \quad (14)$$

where $[P_{rr}]_{k-1}$ is the uncertainty matrix of reflectance at date $k-1$, and $[P_{rr}]_{k-1}$ is the uncertainty of the reflectance change rates.

Assuming that M_{k-1} and M_k are two MODIS datasets observed on two different dates t_{k-1} and t_k , they are independent, and the uncertainty in variance of Landsat and MODIS observations are σ_L^2 and σ_M^2 , respectively, then we can calculate the reflectance change rate \dot{r}_M and its uncertainty $\sigma_{\dot{r}_M}^2$ for each of the MODIS pixels as follows:

$$\dot{r}_M = (r_{M_k} - r_{M_{k-1}})/(t_k - t_{k-1}) = (r_{M_k} - r_{M_{k-1}})/\Delta t \quad (15)$$

and

$$\sigma_{\dot{r}_M}^2 = \frac{2\sigma_M^2}{\Delta t^2} \quad (16)$$

According to the spectral linear mixing theory, the reflectance change of a MODIS pixel is weighted sum of the Landsat reflectance change Δr_c of all clusters within the MODIS pixel:

$$\Delta r_M(x_i, y_i) = \sum_{c=1}^l f_c(x_i, y_i) \times \Delta r_c \quad (17)$$

where $f_c(x_i, y_i)$ is the portion of the Landsat pixels belonging to cluster c within a MODIS coarse pixel. We can determine it through counting the number of Landsat pixels of each cluster for every MODIS pixel by Eq. (18):

$$f_c(x_i, y_i) = N_c(x_i, y_i)/N \quad (18)$$

where $N_c(x_i, y_i)$ is the number of the Landsat pixels belonging to cluster c within the MODIS pixel at (x_i, y_i) . N is the total number of Landsat pixels within a MODIS pixel.

If we divide the two sides of Eq. (17) by the time difference, we have the change rate equation; that is

$$\dot{r}_M(x_i, y_i) = \sum_{c=1}^l f_c(x_i, y_i) \times \dot{r}_c \quad (19)$$

Assuming that the reflectance change of the same cluster among all MODIS pixels is the same, we can apply all m ($m > l$) MODIS pixels to compose a system of linear mixture equations:

$$\begin{bmatrix} \dot{r}_M(x_1, y_1) \\ \vdots \\ \dot{r}_M(x_i, y_i) \\ \vdots \\ \dot{r}_M(x_n, y_n) \end{bmatrix} = \begin{bmatrix} f_1(x_1, y_1) & f_2(x_1, y_1) & \dots & f_l(x_1, y_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_i, y_i) & f_2(x_i, y_i) & \dots & f_l(x_i, y_i) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n, y_n) & f_2(x_n, y_n) & \dots & f_l(x_n, y_n) \end{bmatrix} \begin{bmatrix} \dot{r}_1 \\ \vdots \\ \dot{r}_c \\ \vdots \\ \dot{r}_l \end{bmatrix} \quad (20)$$

From Eq. (20), we can obtain an estimation of the Landsat reflectance change rate \dot{r} and its covariance matrix P_{rr} through a weighted Least-squares adjustment (Hirvonen, 1971; Wolf, 1995) as follows

$$\hat{r} = (A^T W_M A)^{-1} A^T W_M R_M \quad (21)$$

$$P_{rr} = \hat{\sigma}_r^2 (A^T W_M A)^{-1} \quad (22)$$

where

$$A = \begin{bmatrix} f_1(x_1, y_1) & f_2(x_1, y_1) & \dots & f_l(x_1, y_1) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_i, y_i) & f_2(x_i, y_i) & \dots & f_l(x_i, y_i) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n, y_n) & f_2(x_n, y_n) & \dots & f_l(x_n, y_n) \end{bmatrix}, R_M = \begin{bmatrix} \dot{r}_M(x_1, y_1) \\ \vdots \\ \dot{r}_M(x_i, y_i) \\ \vdots \\ \dot{r}_M(x_n, y_n) \end{bmatrix}, V_M = A \hat{r} - R_M$$

$$\hat{\sigma}_r^2 = \frac{V_M^T W_M V_M}{m-l}, W_M = Q_M^{-1}, Q_M = \sigma_{\dot{r}_M}^2 I.$$

Q_M is the variance matrix of the change rate measurements R_M calculated from the MODIS observation uncertainty σ_M^2 through the error propagation law. $\hat{\sigma}_r^2$ is the unit variance estimation from the residuals of the observation Eq. (20). I is a unit matrix.

When the Landsat observation at t_0 is used as the initial prediction of the state of the reflectance,

$$[P_{rr}]_0 = \sigma_L^2 I. \quad (23)$$

When the updated synthetic Landsat image at t_{k-1} is used as the first part of the reflectance transition model, P_{rr} is the uncertainty in variance of the updated synthetic Landsat image:

$$[\hat{P}_{rr}]_{k/k-1} = [P_{rr}]_{k-1} + [F_r P_{rr} F_r^T]_{k-1} \quad (24)$$

2.2.2. Observation model

KFRFM requires two pairs of clear-sky Landsat and MODIS images at t_0 and t_n for prediction of the period $[t_0, t_n]$. The system state variable transition model is based on the first pair of Landsat and MODIS images at t_0 while the observation model is based on the second pair of Landsat and MODIS images at t_n . Since there is no Landsat observation available at t_k for updating the predicted system state variable $\hat{r}_{k/k-1}$ but only MODIS observation is available at a much lower spatial resolution, we developed a pseudo-observation model which calibrates the MODIS measurement to Landsat spatial resolution at t_k by using the same prediction method described above to predict a synthetic Landsat image from the MODIS image at t_k with the pair of Landsat and MODIS images at t_n . Using this predicted synthetic Landsat image with its uncertainty estimate as a newly "observed" measurement of the Landsat image at t_k , we can update the state prediction at t_k from the system variable

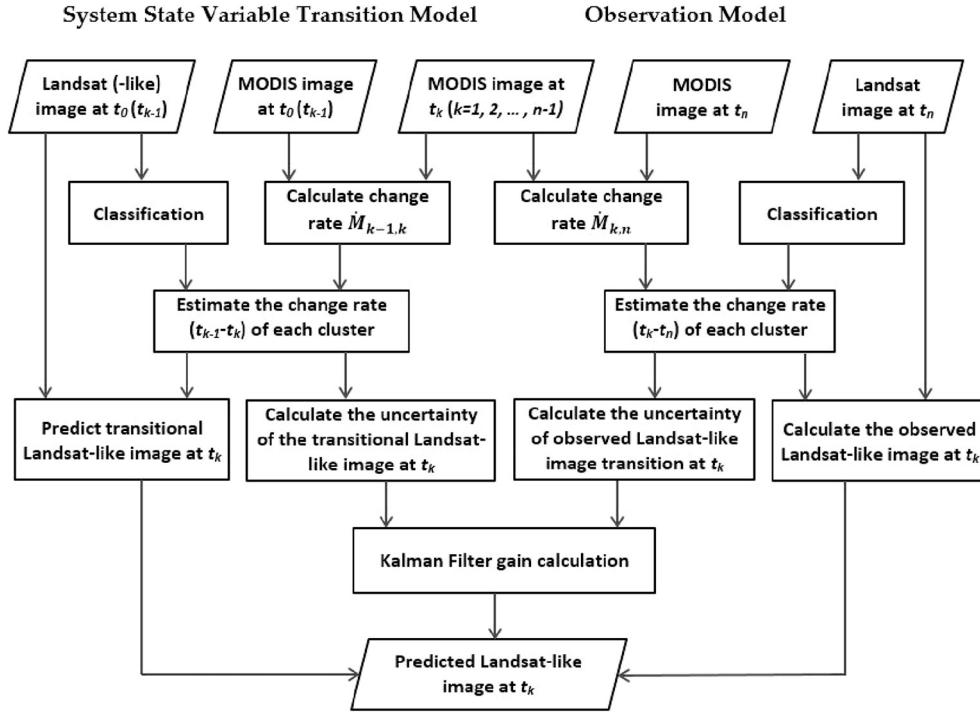


Fig. 1. Main process flowchart of Kalman Filter Reflectance Fusion Model (KFRFM).

transition model. For this processing we first calculate the change rates of the surface reflectance in the period $[t_k, t_n]$ using the reflectance differences of the two MODIS images at dates t_k and t_n , then use the change rate plus the Landsat image at t_n to estimate the synthetic Landsat image at t_k as the observation model output for updating the Landsat-like image at t_k predicted from the system state variable transition model.

According to the above described, similar to Eq. (10), we have the observed state variable equation as follows:

$$\hat{r}_{k/n}(x, y, c, b) = r_n(x, y, c, b) + (t_k - t_n)\dot{r}_n(x, y, c, b) \quad (25)$$

where $\dot{r}_n(x, y, c, b)$ is the reflectance change rate between dates t_k and t_n , and $r_n(x, y, c, b)$ is the true Landsat reflectance at t_n .

In the matrix form of Eq. (25), we have

$$[\hat{r}]_{k/n} = [r]_n + [F_r \dot{r}]_n \quad (26)$$

According to error propagation law, we have the uncertainty of $\hat{r}_{k/n}$ as follows:

$$[\hat{P}_{rr}]_{k/n} = [P_{rr}]_n + [F_r P_{rr} F_r^T]_n \quad (27)$$

where

F_r is specified by Eq. (14) with $\Delta t = t_k - t_n$

$[P_{rr}]_n$ is the uncertainty matrix of the Landsat image at t_n

P_{rr} is specified by Eq. (22).

From the formulation of the prediction model above, we can see that there are two issues that may affect the effectiveness of the prediction model. One is the number of clusters of the Landsat images used for deriving gradual seasonal changes. Another is land cover changes occurring during the period of fused images. To solve these issues, we take the optimization method in determination of the cluster numbers and the residual adjustment for land cover changes (Zhong and Zhou, 2019). The implementation procedures are shown in Section 2.2.4 below.

2.2.3. State variable updating

Using the output $[\hat{r}]_{k/n}$ of the observation model as the “observed” measurement, the prediction $[\hat{r}]_{k/k-1}$ of the system state of the

reflectance from the system variable transition model can be updated by calculating the Kalman gain K_k from the uncertainty information of the two images: $[\hat{P}_{rr}]_{k/n}$ and $[\hat{P}_{rr}]_{k/k-1}$.

As shown in Eqs. (8) and (9), since $H_k = I$ in our case, we have

$$S_k = [\hat{P}_{rr}]_{k/n} + [\hat{P}_{rr}]_{k/k-1}$$

$$K_k = [\hat{P}_{rr}]_{k/k-1} S_k^{-1}$$

$[\hat{r}]_{k/k} = [\hat{r}]_{k/k-1} + K_k([\hat{r}]_{k/n} - [\hat{r}]_{k/k-1})$ where $[\hat{P}_{rr}]_{k/k-1}$ is the variance matrix (uncertainty) of the predicted state of the reflectance variable from Eq. (24), and $[\hat{P}_{rr}]_{k/n}$ is the variance matrix of the pseudo-observations from Eq. (27).

The uncertainty of the updated synthetic Landsat image at t_k is

$$[P_{rr}]_{k/k} = (I - K_k)[\hat{P}_{rr}]_{k/k-1}(I - K_k)^T + K_k[\hat{P}_{rr}]_{k/n}K_k^T \quad (28)$$

2.2.4. Implementation of KFRFM

Assuming that we have acquired two pairs of clear-sky Landsat and MODIS images on start date t_0 , and end date t_n , and a time series of MODIS images on date t_k ($k = 1, 2, \dots, n - 1$), we implemented the KFRFM algorithm to predict synthetic Landsat images on date t_k using the two pairs of Landsat and MODIS images and the MODIS images at t_k .

Before we apply KFRFM, both the Landsat and MODIS images are required to be calibrated and co-registered to the same physical quantity, such as top-of atmosphere reflectance or surface reflectance. The co-registration process can be done in the following steps: re-projecting coordinates of coarse-resolution image to that of the fine-resolution image if they are different, resampling coarse resolution image to fine resolution by nearest neighbor algorithm, geo-referencing one image to another by selecting control points or maximizing correlation between the two images, and then cropping them to cover the same area (Gao et al., 2010; Gevaert and García-Haro, 2015). After we preprocessed all the Landsat and MODIS images to the same physical quantity, we can process the data by KFRFM in the following steps. The major process steps are also illustrated in Fig. 1.

1. For the system state variable transition from time t_{k-1} to t_k ($k = 1, \dots, n - 1$)

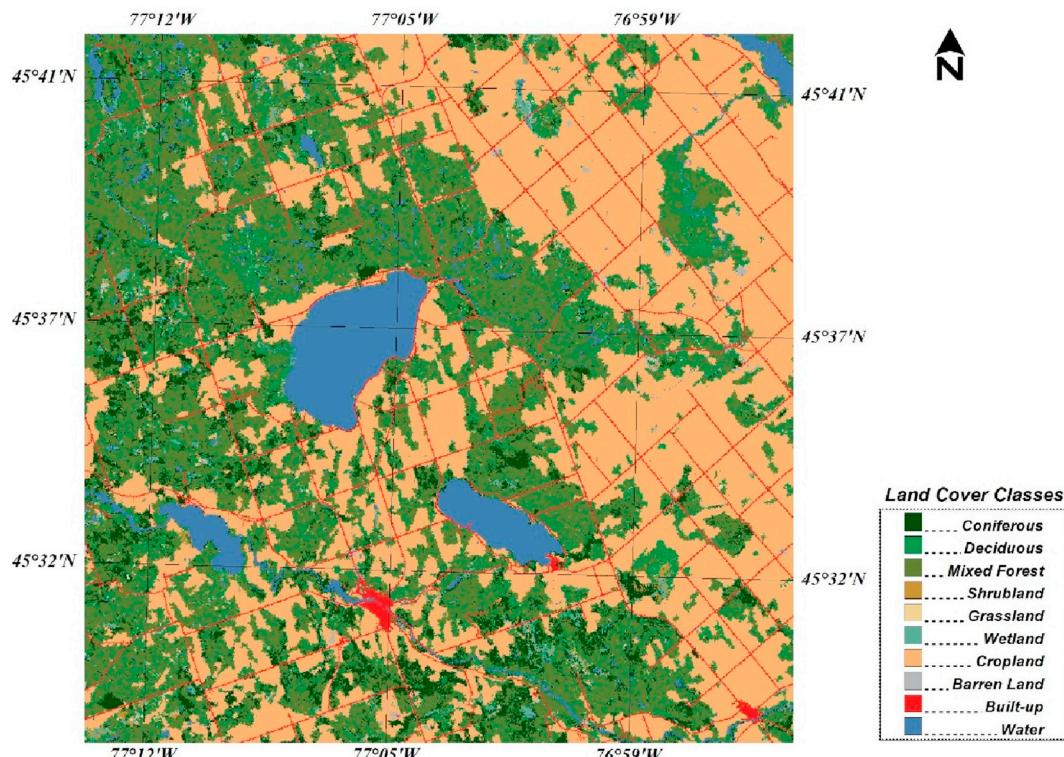


Fig. 2. Study area and its land cover map.

Table 2
Land cover types and their percentage within the study area.

	Land cover class	%
Primarily vegetated areas	Temperate or sub-polar needleleaf (coniferous) forest	6.4
	Temperate or sub-polar broadleaf deciduous forest	12.7
	Mixed forest	28.0
	Temperate or sub-polar shrubland	2.4
	Temperate or sub-polar grassland	0.5
	Wetland	1.4
	Cropland	38.0
Primarily non-vegetated areas	Barren land	0.9
	Urban and built-up	3.8
	Water	5.8

- 1.1. classify the Landsat image at t_0 into several clusters;
- 1.2. calculate the MODIS reflectance change using the MODIS observations at t_{k-1} and t_k ;
- 1.3. estimate the reflectance change rate for each cluster from the MODIS reflectance change calculated in step 1.2;
- 1.4. predict synthetic Landsat image at t_k using the reflectance change rate of all the clusters estimated in step 1.3 and the previously updated synthetic Landsat image at t_{k-1} (from step 3 blow) or Landsat image at t_0 if k equals to 1;
- 1.5. calculate the variance matrix (uncertainty) of the predicted Landsat-like image at t_k ;
- 1.6. compute the correlation coefficients of the differences of the MODIS images at t_0 (or t_k) and t_k (or t_{k+1}) and the differences of the Landsat image at t_0 (or synthetic Landsat image at t_k) and the blended synthetic Landsat image at t_k (or t_{k+1}) and sums of the residue squares of the MODIS image and the reversely computed MODIS image from the synthetic Landsat image at t_k ;
- 1.7. repeat the steps 1.1 to 1.6 for all the number of clusters and identify an optimized cluster number with the largest correlation coefficients and relatively small sums of the residue

- squares;
- 1.8. adjust the residues computed in the step 1.6 with a linear interpolation method to clusters of each pixel.
2. For the observation model
 - 2.1. Carry out the steps from 1.1 to 1.8 using the Landsat image at t_n , the MODIS images at t_k and t_n to derived the pseudo-observation at t_k and its corresponding variance matrix (uncertainty) for updating the image predicted from the system state variable transition model.
3. For the state variable update
 - 3.1. Calculate the Kalman gain from the uncertainty estimates from Steps 1 and 2;
 - 3.2. Derive the predicted synthetic Landsat image and estimate its uncertainty at t_k by using the Kalman gain and the two generated synthetic Landsat images above.
4. Switch the system state variable transition model and the observation model and derive a backward synthetic Landsat image at t_k . The uncertainty-weighted average of the forward and backward synthetic images becomes the final prediction at t_k .
5. For predicting time-series of synthetic Landsat images continue the processes from steps 1 to 4 to generate the time-series of synthetic Landsat images between the two time epochs $[t_0, t_n]$ using the two pairs of Landsat and MODIS images at t_0 and t_n and the clear-sky MODIS images between $[t_1, t_{n-1}]$. When the date reaches t_n , we use a new pair of clear-sky Landsat and MODIS images at t (t is later than t_n). The previous t_n then becomes t_0 , and the date of the new pair of clear-sky Landsat and MODIS images becomes t_n . Then the processes described above are repeated to generate a time-series of synthetic Landsat images. Fig. 1 shows the main process flowchart.

3. Experiments and validation

To verify effectiveness of the proposed KFRFM method, we have two study cases for our experiments. The first case attempted to generate a time series of synthetic Landsat images from a set of Landsat-8 OLI and MODIS images in a vegetation growing season. Then we assessed the

Table 3

Functions and cloud conditions of Landsat-8 OLI images used in the study.

Landsat 8 OLI (acquisition date)	Path/Row	Cloud condition	Usage
April 20, 2016	016/028	Clear-sky	Model input
April 27, 2016	017/028	Partially clear-sky	Vegetation index validation
May 6, 2016	016/028	Clear-sky	Vegetation index validation
May 22, 2016	016/028	Clear-sky	Surface reflectance assessment
May 29, 2016	017/028	Partially clear-sky	Model input
June 14, 2016	017/028	Clear-sky	Vegetation index validation
June 23, 2016	016/028	Partially clear-sky	Surface reflectance assessment
June 30, 2016	017/028	Clear-sky	Vegetation index validation
August 10, 2016	016/028	Clear-sky	Vegetation index validation
August 17, 2016	017/028	Partially clear-sky	Model input
August 26, 2016	016/028	Partially clear-sky	Vegetation index validation
October, 4, 2016	017/028	Clear-sky with cloud/haze residues	Vegetation index validation

Table 4

Input images of Landsat-8 OLI and MODIS (MCD43A4) for time-series synthetic Landsat images prediction.

Landsat-8 OLI (acquisition date)	MODIS MCD43A4 (acquisition date)	Predicted time series Day Of Year (DOY)
April 20, 2016	April 20, 2016	114 (calibration pair)
	April 24, 2016	115
	April 28, 2016	119
	May 2, 2016	123
	May 6, 2016	127
	May 10, 2016	131
	May 14, 2016	135
	May 18, 2016	139
	May 22, 2016	143 (calibration pair)
	May 26, 2016	147
May 22, 2016	May 30, 2016	151
	June 2, 2016	154
	June 6, 2016	158
	June 10, 2016	162
	June 14, 2016	166
	June 18, 2016	170
	June 22, 2016	174
	June 26, 2016	178
	June 30, 2016	182 (calibration pair)
	July 4, 2016	186
June 30, 2016	July 8, 2016	190
	July 12, 2016	194
	July 16, 2016	198
	July 20, 2016	202
	July 24, 2016	206
	July 28, 2016	210
	August 2, 2016	215
	August 6, 2016	219
	August 10, 2016	223 (calibration pair)
	August 14, 2016	227
August 10, 2016	August 18, 2016	231
	August 22, 2016	235
	August 26, 2016	239
	August 30, 2016	243
	September 2, 2016	246
	September 6, 2016	250
	September 10, 2016	254
	September 14, 2016	258
	September 18, 2016	262
	September 22, 2016	266
October 4, 2016	September 26, 2016	270
	September 30, 2016	274
October 4, 2016	October 4, 2016	278 (calibration pair)

time-series by producing several vegetation indices and compared them to those computed from Landsat-8 observations. We also analyzed the temporal uncertainty distribution of the blended images. The second case is especially used to compare our proposed KFRFM model to ESTARFM and FSDAF for image fusion performance in a heterogeneous region.

3.1. The first study case

The study area of the first study case is located around at 45°36' N and 77°04' W, about 140 km west of Ottawa, Canada. Fig. 2 shows the land cover map ([Latifovic et al., 2017](#)) of the study area with the same spatial resolution of the Landsat images. Table 2 lists the land cover types and their occupancy percentage within the study area. The region is mainly covered by Forest lands (Coniferous, Deciduous, and Mixed forest) (47.1%) and Cropland (39%). Water, Built-up, and Shrubland account for about 12%, with some small percentage of Wetland (1.4%), Barren Land (0.9%) and Grassland (0.5%). The vegetation growing season of the study area runs roughly from April to October every year.

3.2. Image data

For the modelling exercise of the first study case, two types of Earth Observation data are used. The first is Landsat-8 OLI images. In order to get maximum number of clear-sky images possible for either modelling exercise or model validation, the study area is chosen so that it is covered by two Landsat-8 scenes. The paths/rows of the Landsat-8 scenes are 16/28 and 17/28. The size of the study area is 24 km by 24 km which is equivalent to 800x × 800 Landsat pixels with a resolution of 30 m. The second type of images is daily MCD43A4 (Version 6), which is used as calibration pair with the Landsat images as well as for predicting a time-series of synthetic Landsat images with a temporal interval of 4 days. MCD43A4 is Nadir Bi-directional Reflectance Distribution adjusted reflectance from combined MODIS-Terra and MODIS-Aqua acquisitions with a 16-day retrieval moving window, which removes land surface reflectance anisotropy and is alike observation from nadir view angle ([Schaaf et al., 2002](#)). MCD43A4 has been used to monitor vegetation dynamics and land surface disturbance due to improved accuracy of change detection in intra- and inter-annual temporal dynamics ([Zhang et al., 2003](#)). The MODIS images have a spatial resolution of 500 m. The coverage of one MODIS image pixel is

Table 5

Correlation Coefficient (Eq. (29)) of synthetic Landsat and observation images on May 6, 2016.

Fusion model	Band name (see Table 1)					
	Blue	Green	Red	NIR	SWIR-1	SWIR-2
KFRFM	0.9539	0.9528	0.9601	0.9602	0.9722	0.9662
STARFM	0.8907	0.9180	0.8977	0.9568	0.9485	0.9141
ESTARFM	0.9308	0.9509	0.9372	0.9721	0.9529	0.9301
FSDAF	0.9423	0.9496	0.9523	0.9706	0.9674	0.9585

The bold indicates the largest value (CC, QI) or the smallest value (RMSE) among the 4 models compared.

Table 6

RMSE (Eq. (30)) of the synthetic Landsat image against observation image on May 5, 2016.

Fusion model	Band name					
	Blue	Green	Red	NIR	SWIR-1	SWIR-2
KFRFM	0.0067	0.0097	0.0122	0.0232	0.0214	0.0193
STARFM	0.0101	0.0129	0.0193	0.0248	0.0309	0.0306
ESTARFM	0.0081	0.0098	0.0151	0.0191	0.0281	0.0283
FSDAF	0.0074	0.0099	0.0133	0.0201	0.0230	0.0212

Table 7

QI (Eq. (31)) of the synthetic Landsat image against observation image on May 5, 2016.

Fusion model	Band name					
	Blue	Green	Red	NIR	SWIR-1	SWIR-2
KFRFM	0.9501	0.9506	0.9573	0.9587	0.9715	0.9633
STARFM	0.8781	0.9036	0.8876	0.9501	0.9441	0.9095
ESTARFM	0.9298	0.9492	0.9366	0.9708	0.9527	0.9294
FSDAF	0.9368	0.9469	0.9482	0.9684	0.9669	0.9562

Table 8

Correlation Coefficient (Eq. (29)) of synthetic Landsat and observation images on June 14, 2016.

Fusion model	Band name					
	Blue	Green	Red	NIR	SWIR-1	SWIR-2
KFRFM	0.9236	0.9251	0.9212	0.9391	0.9308	0.9328
STARFM	0.8933	0.8876	0.8875	0.9038	0.9104	0.9034
ESTARFM	0.9046	0.9052	0.8941	0.9308	0.9144	0.9028
FSDAF	0.9208	0.9209	0.9108	0.9464	0.9183	0.9203

Table 9

RMSE (Eq. (30)) of the synthetic Landsat image against observation image on June 14, 2016.

Fusion model	Band name					
	Blue	Green	Red	NIR	SWIR-1	SWIR-2
KFRFM	0.0081	0.0105	0.0155	0.0321	0.0285	0.0245
STARFM	0.0094	0.0128	0.0180	0.0393	0.0321	0.0273
ESTARFM	0.0090	0.0120	0.0181	0.0337	0.0318	0.0290
FSDAF	0.0083	0.0107	0.0165	0.0304	0.0299	0.0264

Table 10

QI (Eq. (31)) of the synthetic Landsat image against observation image on June 14, 2016.

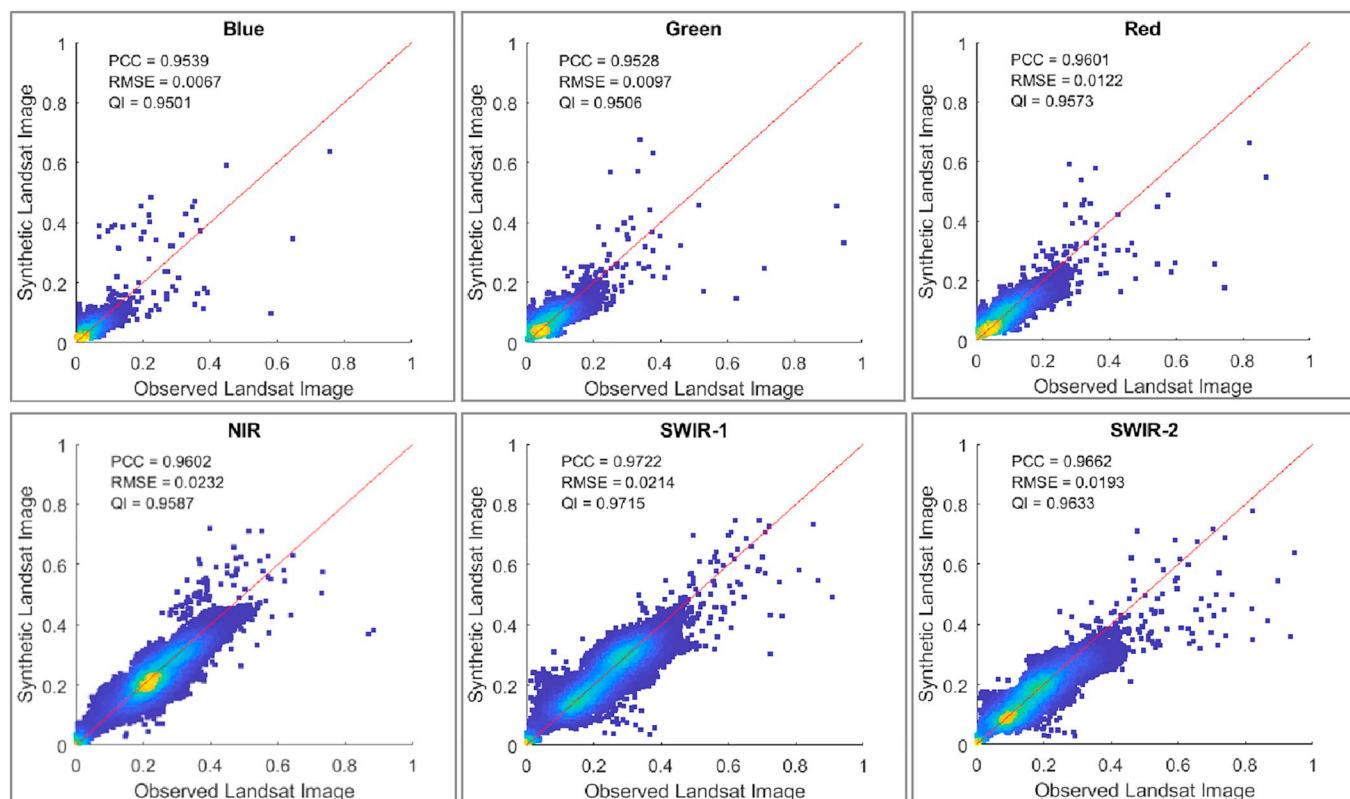
Fusion model	Band name					
	Blue	Green	Red	NIR	SWIR-1	SWIR-2
KFRFM	0.9195	0.9233	0.9166	0.9356	0.9277	0.9283
STARFM	0.8842	0.8822	0.8821	0.8976	0.9071	0.9011
ESTARFM	0.9017	0.9037	0.8903	0.9301	0.9092	0.8990
FSDAF	0.9142	0.9177	0.9044	0.9390	0.9160	0.9150

equivalent to those of 16 by 16 pixels of the Landsat image. The MODIS reprojection tool (MRT) (https://lpdaac.usgs.gov/tools/modis_reprojection_tool) is used to reproject and resample MODIS images to the map projection and spatial resolution of the Landsat images. All the images are surface reflectance from USGS (https://lta.cr.usgs.gov/get_data).

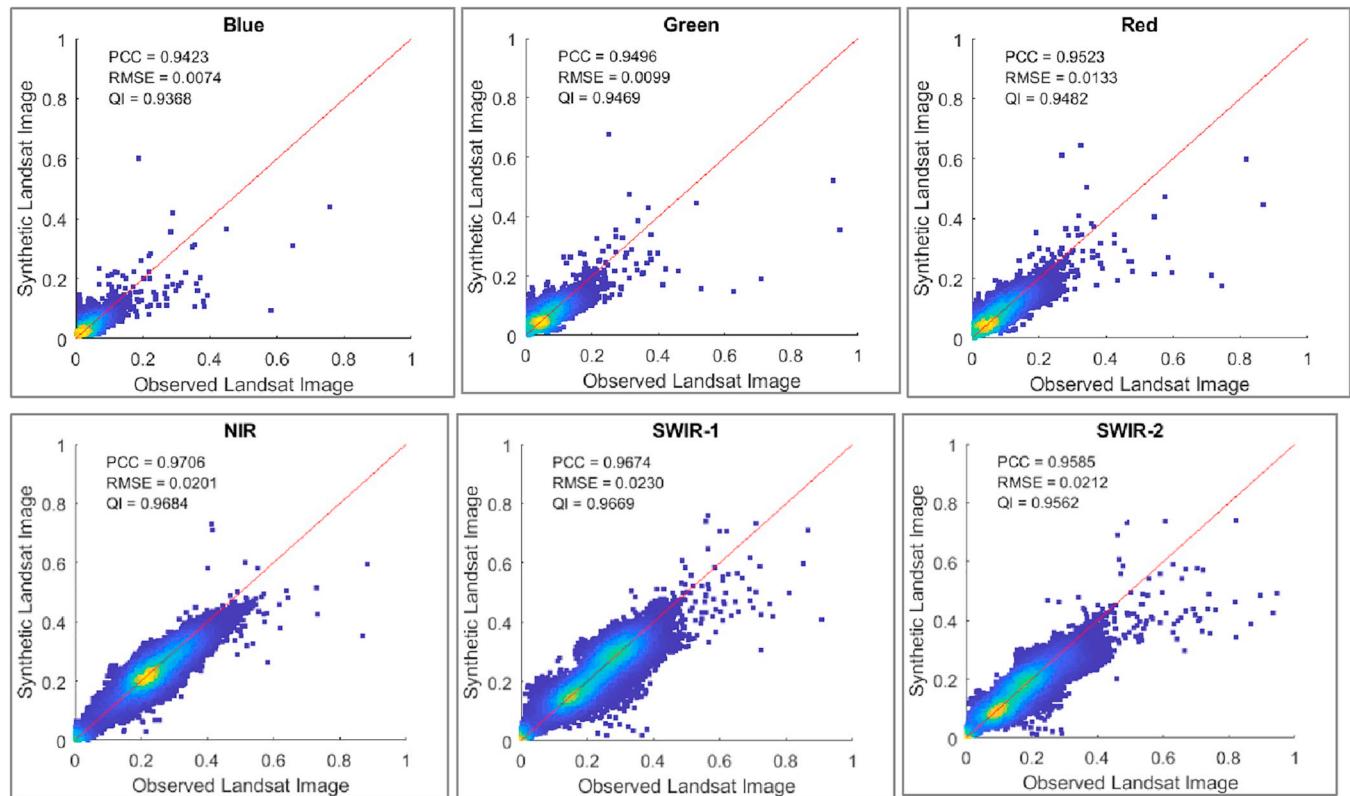
To cover the vegetation growing season, we searched all the clear-sky Landsat-8 OLI images of the study area in 2016. The first and the last available clear-sky Landsat-8 OLI images during the vegetation growing season were acquired at April 20, 2016 and October 4, 2016, respectively. The period of the acquisition dates of two images spans almost the entire vegetation growing season of the study area, therefore, the two images are used as the first and the last images for prediction of the time-series of synthetic Landsat images. As the time interval of the first and the last images lasts about six months, in order to control quality of outputs, we divided the period into multiple sub-periods based on the temporal distribution of all available clear-sky Landsat-8 OLI images. Within the period, in addition to the first and the last two images, there are another 5 clear-sky and 5 partially clear-sky Landsat-8 OLI images. Depending on the dates and cloud conditions of these images, they play different roles in the experiment. Majority of the clear-sky Landsat are used for model input for predicting time-series of synthetic Landsat images, and two clear-sky Landsat-8 OLI images (dated as May 06, and June 14, 2016) are used for model validation. All these images and the partially clear-sky Landsat images are utilized for vegetation index validation. The images and their roles in this study are described in Table 3.

The five pairs of clear-sky Landsat-8 OLI and MODIS images (MCD43A4) acquired on April 20, 2016, May 22, 2016, June 30, 2016, August 10, 2016 and October 4, 2016 (Table 4) are used for calibration image pairs. All these Landsat images are clear-sky images except the image acquired at October 4, 2016 has some haze and cloud residuals. Among them, every two consecutive images constitute two pairs of input images for predicting synthetic Landsat images between them, as such the entire study period from April 20, 2016 to October 4, 2016 is divided into 4 shorter periods as shown in Table 4.

Although the MODIS images (MCD43A4) have a daily coverage, to demonstrate the capabilities of KFRFM, we used the daily images one out of every 4 days for image fusion exercise and analysis from the start date to the end date of the study period as listed in Table 4. In total, there are 38 images over the study period (the third column of Table 4) with varied number of images within two pairs of Landsat 8 OLI and MODIS images depending on the length of the time interval of each two pairs of calibration images. For the synthetic Landsat image prediction, all the six bands of Landsat-8 OLI and MODIS images listed in Table 1 are blended.



(a)



(b)

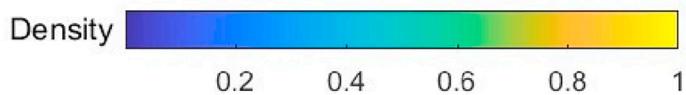


Fig. 3. Scatter plots with quality indices of surface reflectance by band of the synthetic Landsat image against observed Landsat image on May 6, 2016. (a) KFRFM vs. Observation; (b) FSDAF vs. Observation.

3.3. The second study case (Coleambally, Australia)

The second study case is based on the Coleambally Irrigation Area study site ('Coleambally' herein). Coleambally is located in southern New South Wales, Australia. Over Coleambally, 17 cloud-free Landsat 7 ETM+ and MODIS (MOD09GA) images were available for the 2001–2002 summer growing season. The images cover an area of 43 km by 51 km with a spatial resolution of 25 m. Due to the relatively small field sizes and large spatial variances in the irrigation area, Coleambally is considered as a spatially heterogeneous landscape [Emelyanova et al., 2013a]. Hence this dataset is used to assess model's capability of dealing with heterogeneous landscape situations. The 17 cloud-free Landsat and MODIS time-series images are downloaded and used for model evaluation and comparison in this study. The details of the dataset can be found in [Emelyanova et al., 2013a; Emelyanova et al., 2013b].

4. Result analysis and discussion

4.1. Model validation

The model validation is based on the first study case. The KFRFM model validation is based on two Landsat observations dated as May 06, 2016 and June 14, 2016, and their corresponding synthetic images predicted using the model. The two clear-sky Landsat images are left on purpose for model validation. In this study, we used three methods for model validation and quality assessment of the blended images. The first is to use quality indices of the predicted images against the Landsat observations. In this regard, KFRFM is also compared to three well-developed image fusion models: STARFM [Gao et al., 2006], ESTARFM [Zhu et al., 2010], and FSDAF [Zhu et al., 2016]. The second method is to use density scatter plot of the reflectance of the synthetic Landsat images and the observed Landsat images. The last method is visual assessment.

The first (May 06, 2016) of the two clear-sky images for model validation is within the first two pairs (April 20, 2016 and May 22, 2016) of Landsat and MODIS calibration images, and 16 days away from the two image pairs. The second image (June 14, 2016) is within the second two pairs (May 22, 2016 and June 30, 2016) of Landsat and MODIS calibration images, and 22 day and 16 days away from the first pair and the second pair of the images, respectively (Tables 3 and 4).

4.1.1. Quality indices and model performance comparison

There are a number of quality indices such as average differences (*AD*), Pearson correlation coefficient (*PCC*), root mean squared error (*RMSE*), relative dimensionless global error (*ERGAS*), and quality index (*QI*) or the structure similarity index (*SSIM*), etc. that can be used for quality assessment of the blended images against observations. (Huang and Song, 2012; Zhu et al., 2016; Zhong and Zhou, 2018; Wang and Bovik, 2002). In this study, we use three commonly used quality indices for the quality assessment and model comparison. The three quality indices include Pearson correlation coefficient (*PCC*), root mean square error (*RMSE*), and quality index (*QI*).

The formulas for the three quality indices (*PCC*, *RMSE*, and *QI*) are as follows:

1. Pearson correlation coefficient (*PCC*)

$$PCC = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (29)$$

2. The root mean squared error (*RMSE*)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2} \quad (30)$$

3. The quality index (*QI*)

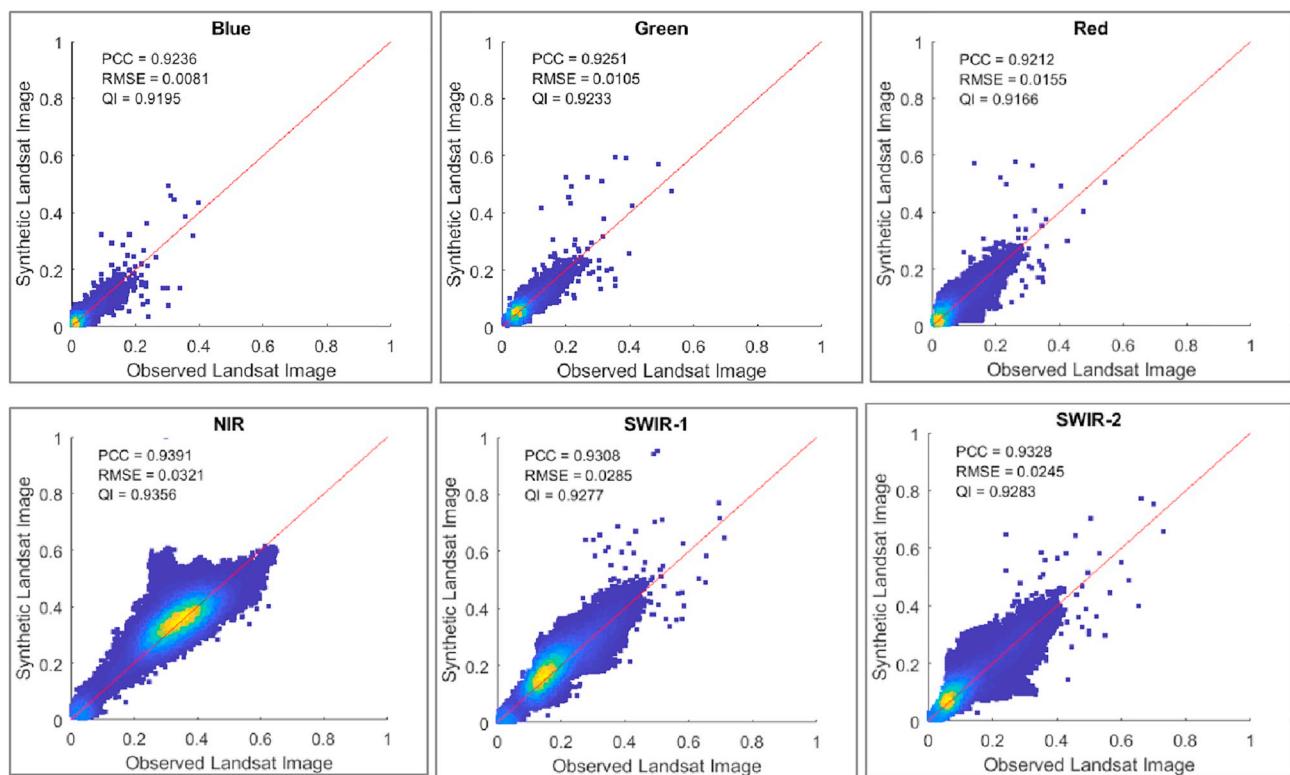
$$QI = \frac{4\sigma_{xy}\bar{x}\bar{y}}{(\sigma_x^2 + \sigma_y^2)(\bar{x}^2 + \bar{y}^2)} \quad (31)$$

where: x and y are surface reflectance; \bar{x} and \bar{y} are their mean values; σ_x^2 , σ_y^2 and σ_{xy} are their variances and covariance of a synthetic Landsat image (x) and the corresponding observed Landsat image (y). n is the total number of the image pixels.

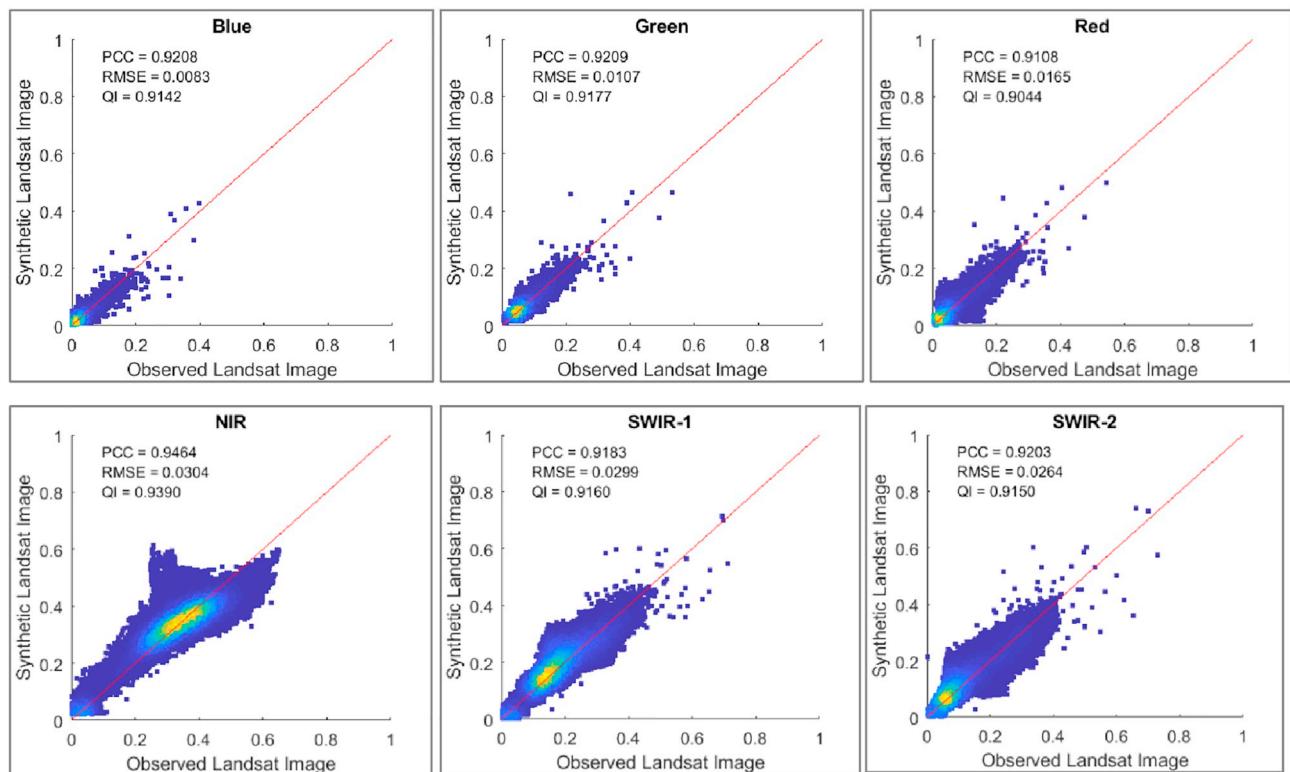
The *PCC* reflects the degree of correlation between the predicted and reference images. A value closer to one indicates a strong linear relationship between the two images. The *RMSE* is to assess the differences between the predicted and reference images. A smaller value indicates a closer match. The quality index *QI* is used to assess the similarity of the structures between the predicted and reference images. A value closer to one indicates structural similarity, and vice versa.

As discussed in Introduction section, there are a number of image fusion models published with various characteristics. Among those models, software implementation of STARFM, ESTARFM, and FSDAF is available online. STARFM is a pioneer such kind of surface reflectance blending model, and ESTARFM is an enhanced model based on STARFM and is able to deal with heterogeneous regions. FSDAF is designed for image fusion capable of handling land cover changes. We run these three models (with default parameter values of the software) with the same datasets used for the proposed KFRFM model validation described above. Tables 5 to 10 show the three quality indices (Pearson correlation coefficient (*PCC*), root mean squared error (*RMSE*), and quality index (*QI*) – Eqs. 29–31) of those models for the two validation images. Overall, the blended synthetic Landsat images have high correlation coefficient, small *RMSE* and large *QI* of all the bands across all the models with some variations against Landsat observations. Among these models, KFRFM performs consistently the best. For the six bands of the blended image on May 6, 2016, KFRFM has five bands better than other models (Tables 5 to 7). This holds true for the blended images on June 14, 2016 (Tables 8–10). ESTARFM has one band (Near Infrared) of May 6, 2016 better than other models, and FSDAF (temporal weighted average of forward and backward predictions) has one band (Near Infrared) of June 14, 2016 better than other models. The largest *PCC* and *QI* and the smallest *RMSE* values are highlighted in Tables 5 to 10 for each band.

All the *QI* values of the image bands of May 6, 2016 from KFRFM are > 0.95 , and those values of June 14, 2016 are bigger than 0.91, which indicates the structure similarity of the blended and observed images. The synthetic images predicted by KFRFM have a good correlation with the corresponding observed Landsat images by Pearson Correlation Coefficient (*PCC*) that are in the range of 0.9528 and 0.9722, and 0.9212 and 0.9391 for images of May 6, and June 14,



(a)



(b)

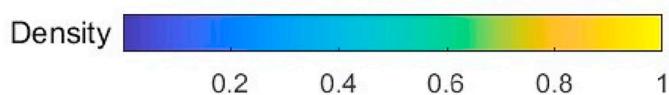


Fig. 4. Scatter plots with quality indices of surface reflectance by band of the synthetic Landsat image against observed Landsat image on June 14, 2016. (a) KFRFM vs. Observation; (b) FSDAF vs. Observation.

2016, respectively. In general, bands 2, 3, and 4 (Blue, Green, Red) have a smaller RMSE compared to those of bands 5, 6 and 7 (Near Infrared, Shortwave Infrared 1, Shortwave Infrared 2). The root mean square error (RMSE) of all bands of the image of May 6, 2016 are small from 0.0067 to 0.0232, and of the image of June 14, 2017 are from 0.0081 to 0.0321. These quality indices indicate that KFRFM is capable to produce high quality of synthetic Landsat images. Based on the quality indices assessment, among the four model assessed, KFRFM performed the best, followed by FSDAF, ESTARFM, and STARFM.

4.1.2. Scatter plots of reflectance

Figs. 3 and 4 show the scatter plots colored by density of the reflectance of the synthetic Landsat images and the reflectance of the Landsat observations with the quality index values displayed for individual band. Fig. 3 is for the images of May 06, 2016, and Fig. 4 is for the images of June 14, 2016. In addition to the scatter plots of KFRFM, outputs of FSDAF are also included for comparison (Figs. 3b and 4b)

The scatter plots of all the individual bands of the reflectance of the synthetic images against the reflectance of the observations reveal that very high-density points are along the 1:1 lines (in red), with some

points above and below the lines. That means that majority of the image pixels are predicted correctly. The good matches of the two images are also reflected by the three quality indices (PCC and QI are close to one, and small $RMSE$ for the most of the bands). These hold true for the two validation images (Figs. 3 and 4).

Although there are some variations of scattering points away from the 1:1 lines from band to band, these figures show the similarities of the two images predicted by KFRFM and FSDAF algorithms. It is noted that both the scatter plots of the Near Infrared band of the images on June 14, 2016 predicted by KFRFM and FSDAF have a spike with a larger $RMSE$ (0.0321 for KFRFM and 0.0304 FSDAF). The abnormal performance of the single band deserves further investigation.

4.1.3. Visual assessment

Figs. 5 and 6 show two false colour composites with different band combination of the two synthetic Landsat images from KFRFM and FSDAF and their corresponding observed Landsat images. According to the false colour composite comparison, overall, the two Landsat synthetic images from the two fusion models are very similar to their corresponding Landsat observations.

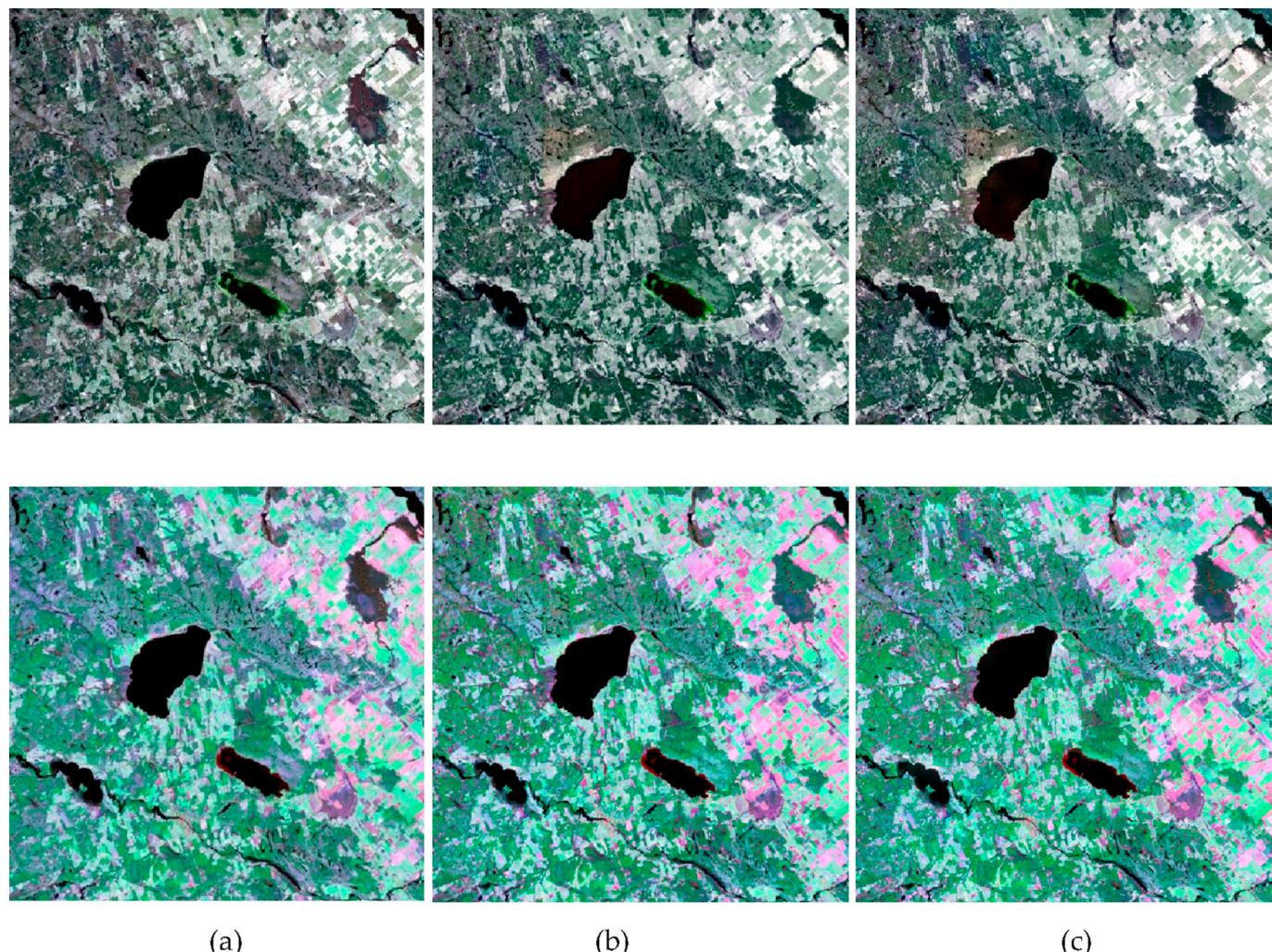


Fig. 5. False colour composites of Blue, Green and Red bands (RGB) (the upper row), and False colour composites of Red, Near-Infrared, and Shortwave Infrared 1 bands (RGB) (the second row) of the observed Landsat-8 OLI image (a) and the synthetic Landsat image of KFRFM (b) and FSDAF (c) on May 6, 2016.

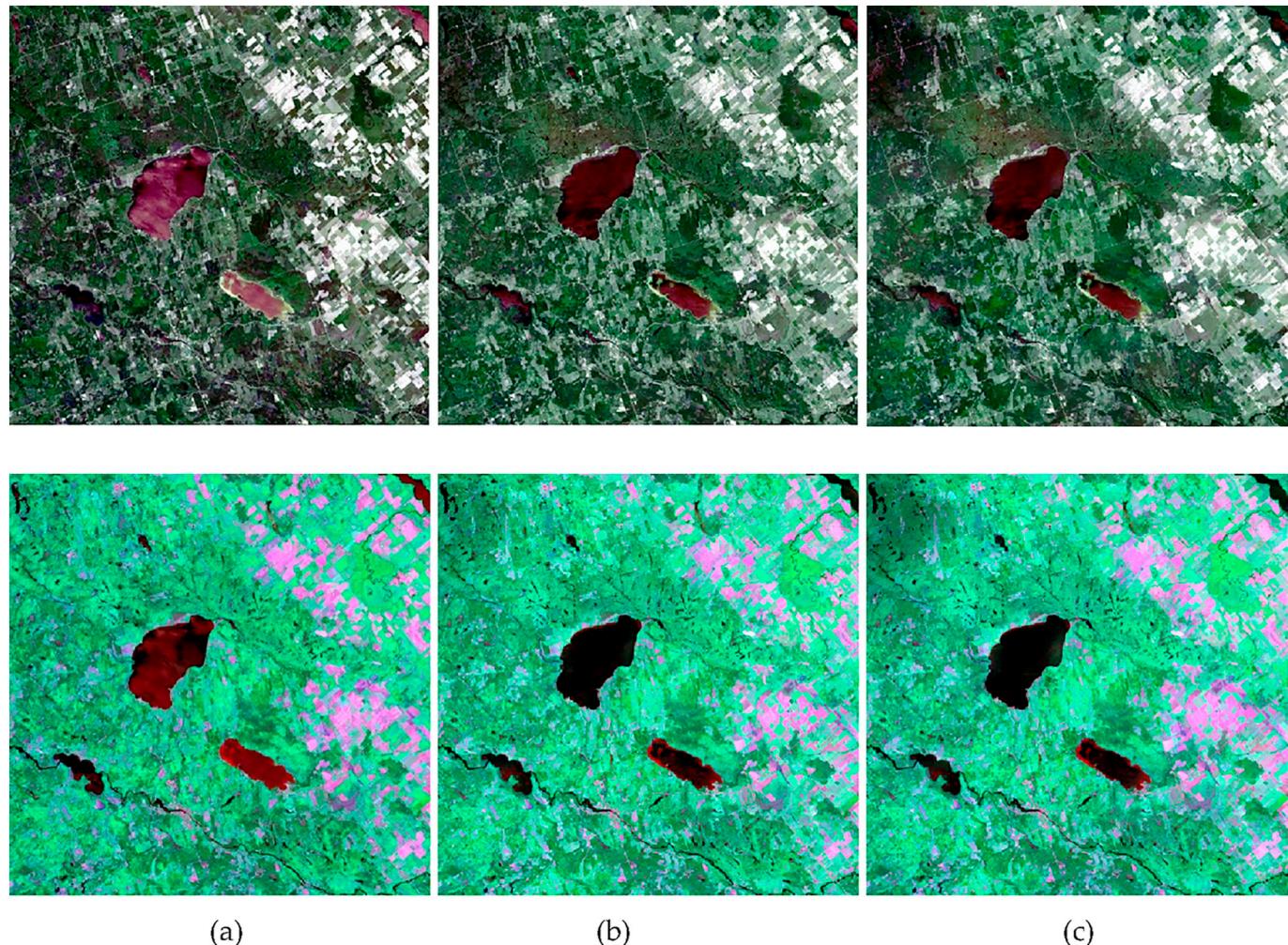


Fig. 6. False colour composites of Blue, Green and Red bands (RGB) (the upper row), and False colour composites of Red, Near-Infrared, and Shortwave Infrared bands (RGB) (the second row) of the observed Landsat-8 OLI image (a) and the synthetic Landsat images of KFRFM (b) and FSDAF (c) on June 14, 2016.

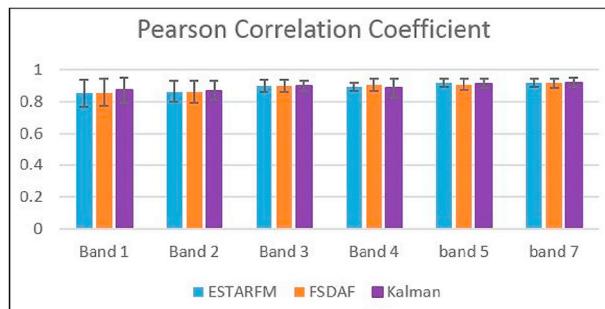


Fig. 7. Band average of Pearson correlation coefficient (PCC) with standard deviation

4.2. Model comparison of Coleambally, Australia

The model comparison is based on the time-series of fused images of the Coleambally Irrigation Area study site. We predicted 15 synthetic images (1800 by 1500 pixels) from the 17 image pairs, i.e., we use the image pairs immediate before and after the prediction date for reference images, and the MODIS image at the prediction date for prediction. The synthetic Landsat images are then assessed against its corresponding Landsat observations. The same as the first experiment, we computed the three quality indices (Eqs. (29)–(31)) of every band of all the predictions. As there are total 15 synthetic images, it is more

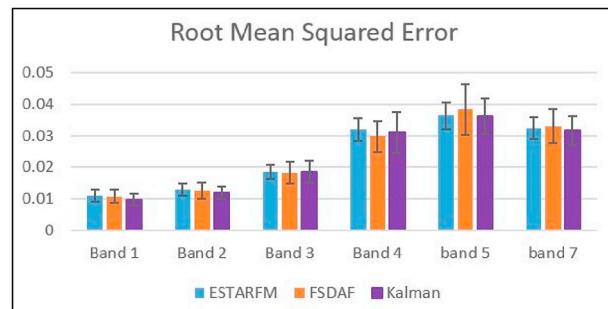


Fig. 8. Band average of root mean squared error (RMSE) with standard deviation

meaningful to use some statistic data for the comparison. Fig. 7 shows the band average of the 15 correlation coefficients (Eq. (29)) between the synthetic images and the observations of the three models, and Fig. 8 displays the band average of the 15 root mean squared errors (Eq. (30)) of the synthetic images against observations of the three models. Fig. 9 presents the band average of the quality index (Eq. (31)) of all the images. The error bars on the three graphs are the standard deviations of the corresponding quality index.

From these figures, it can be seen that all the three models performed well (large PCC and QI which are close to 1, and small RMSE) for the heterogeneous region although RMSE has relatively large

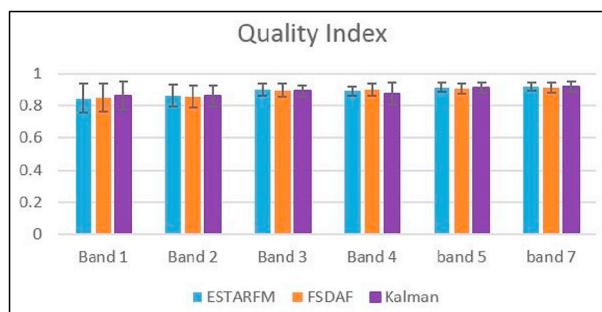


Fig. 9. Band average of quality index (QI) with standard deviation

difference among bands. Bands 4, 5 and 7 have a larger RMSE than bands 1 to 3. It is expected as reflectance of bands 4, 5 and 7 has large value and variation for vegetation. Among the three models, overall KFRFM performed better than ESTARFM and FSDAF. Based on all the quality indices, KFRFM has better average values (larger PCC and QI, and smaller RMSE) and smaller standard deviations about four out of the six bands.

Fig. 10 shows the Landsat observation acquired at 16 Oct. 2001 (Fig. 10.a) and the three fused images from KFRFM (Fig. 10.b), ESTARFM (Fig. 10.c), and FSDAF (Fig. 10.d). Visually, the fused image from KFRFM is more similar to the observation than other two models are.

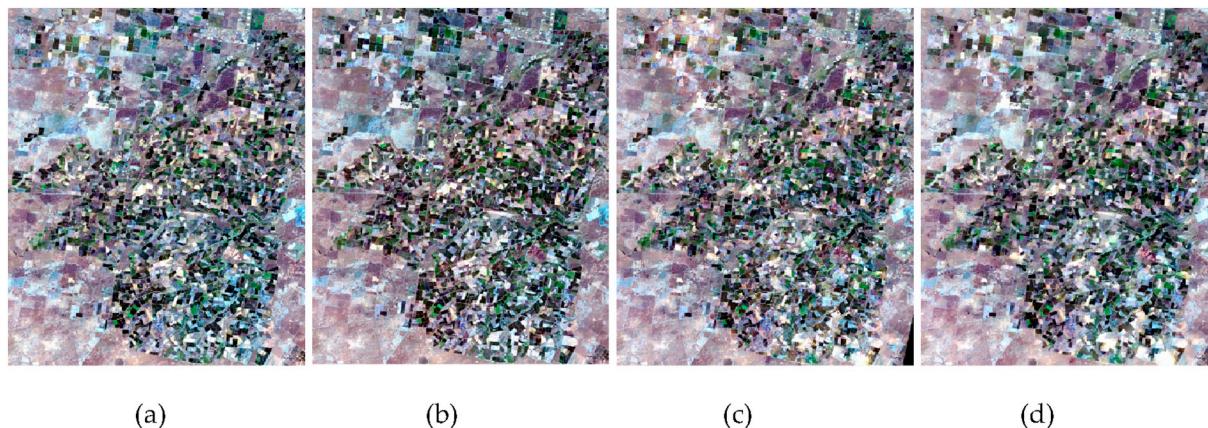


Fig. 10. Comparison of fused images against the observation (16 Oct. 2001). Colour composite of bands 1,2,3 for RGB. (a) Landsat 7 observation, (b) KFRFM; (c) ESTARFM, (d) FSDAF.

4.3. Assessment and analysis of the predicted time-series images of the first study case

4.3.1. Vegetation indices

The time-series of the produced synthetic Landsat-8 OLI images covers the vegetation growing season of the study area (from later April to early October 2016). With all the 38 synthetic images, we calculated a few vegetation indices as listed in Table 11 for all the vegetation types identified within the study area (the first 7 land cover classes in Table 2). At the same time, we also calculated the same indices using all available Landsat-8 OLI images (Table 4) of the same period. For partially cloud covered Landsat images, only those clear-sky pixels are used. Then we compared the two sets of indices with graphs (Figs. 11–15). As any of these indices uses only two or three bands, in order to evaluate all the 6 bands blended, we choose five (Chen et al., 2018) indices so that all the blended bands of Landsat-8 OLI image are used for computation of at least one vegetation index. Table 11 lists the vegetation indices and the bands and formulas used for calculation of those indices. The purpose of computing these time-series vegetation indices in this study is not for study of vegetation phenology but for

demonstration that the time-series of synthetic Landsat images produced by the model can be used to generate useful vegetation indices and they are compatible to those from actual Landsat observations.

4.3.1.1. NDVI. NDVI (Weier and Herring, 2000) is normalized difference of Near-Infrared and Red bands. Its value is between –1.0 and 1.0. Vegetated land surface usually has a bigger NDVI, and land surface with less vegetation has a smaller value due to vegetation's high reflectance in the Near-Infrared spectrum, and relative low reflectance in Red wavelength.

Fig. 11 shows the time-series of NDVI of the seven (Emelyanova et al., 2013) vegetation types within the study area. The square markers are the average values of all the pixels of each land cover type from the synthetic images, and the error bar is its one standard deviation. The red dots are the NDVI values computed from the observed Landsat-8 OLI images. Overall, the predicted NDVIs match well to those from Landsat observation within the entire vegetation growing season. All the NDVIs of all the vegetation types from the Landsat observations are well within the one standard deviation of the predicted values, and the time-series of the predicted NDVI values catch the phenology trend of all vegetation types well.

4.3.1.2. GNDVI. Green Normalized Difference Vegetation Index (Gitelson et al., 1996) is similar to NDVI but uses Green band to replace Red band of NDVI (Table 5). NDVI is sensitive to vegetation fraction and to rate of absorption of photosynthetic solar radiation, but

the green vegetation index can sense the concentration of chlorophyll, measure the rate of photosynthesis and monitor plant stress (Gitelson et al., 1996). The graphs in Fig. 12 show that the predicted values (square in light blue) are well aligned with the observed ones (red dots).

4.3.1.3. EVI. EVI is Enhanced Vegetation Index using three bands (Blue, Red, and Near-Infrared). Fig. 13 presents the EVI graphs of the 7 vegetation types of the study area. Compared to NDVI, EVI is more responsive to canopy variations, canopy type and architecture, and plant physiognomy. EVI can be associated with stress and changes related to drought (Huete et al., 2002). Same as NDVI and GNDVI, almost all EVI values from the predicted Landsat-like images (light blue) are well close to the EVIs from the observed Landsat images (red dots) at the same or close to the acquisition dates. Only one or two points are located on the edge of the standard deviation of the observed values.

4.3.1.4. GVMI. Global Vegetation Moisture Index (GVMI) (Ceccato et al., 2002a) is designed to detect water content in vegetation and uses the near-infrared and shortwave infrared 1 wavelengths for its

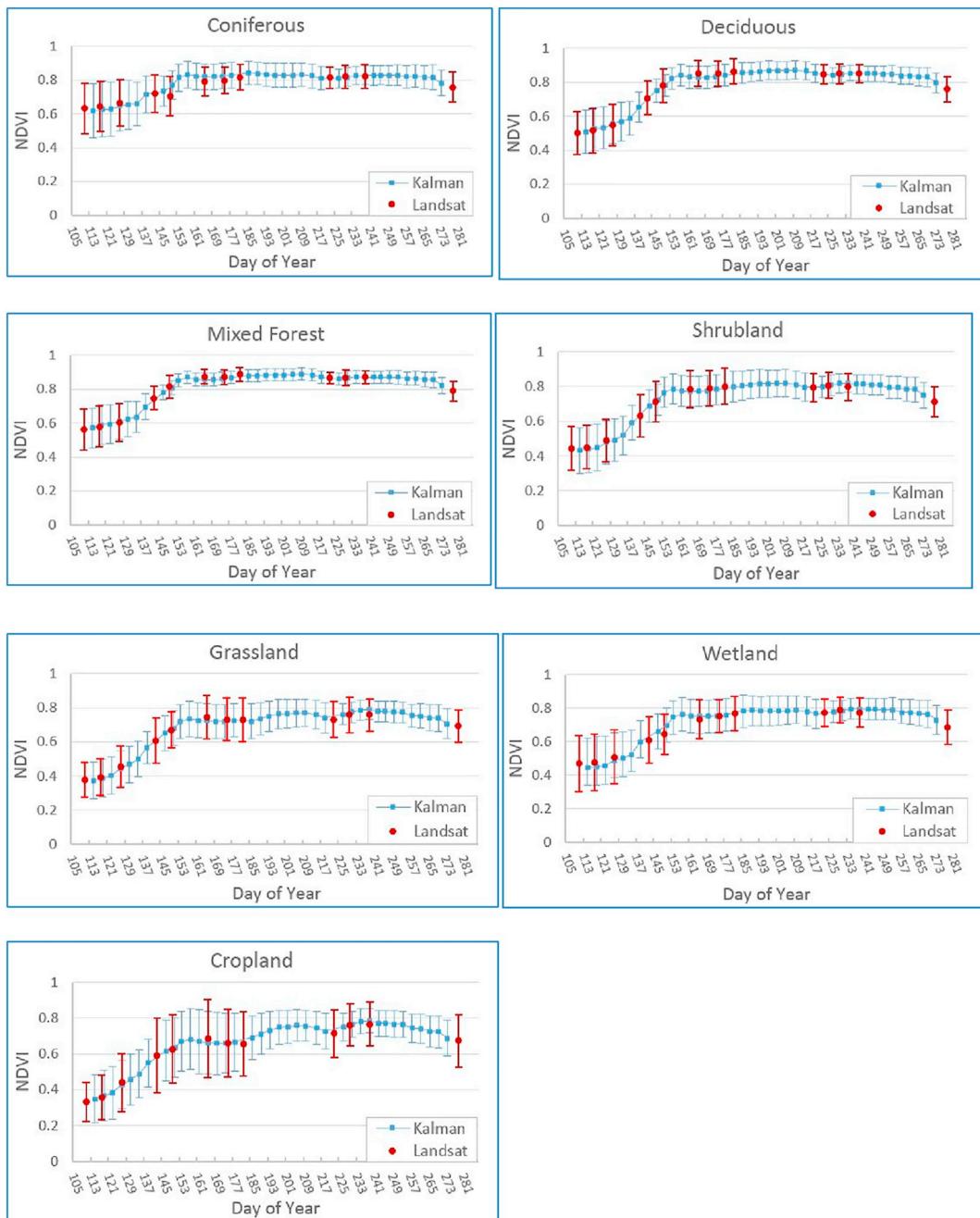


Fig. 11. Time-series of NDVI of vegetated land cover types of the study area.

calculation as listed in Table 5. GVMI-derived water content is consistent with water content expressed as a quantity of water per unit area. Therefore GVMI could be used to set appropriate calendars for managing early fires and to assess risk of fire occurrence (Ceccato et al., 2002b). The graphs in Fig. 14 show the time-series of GVМИs from the predicted synthetic Landsat images against those from the observed Landsat-8 OLI images. Like other vegetation indices, both values from the predicted synthetic Landsat images and observations are very close to each other for all the land cover types.

4.3.1.5. D_{1650nm} . Depth of bands within absorption features has been used to estimate various plant-based properties from hyperspectral remote sensing (van Niel et al., 2003). D_{1650nm} are related to environmental moisture. It involves Near-infrared band and the two Shortwave Infrared bands 1 and 2 (Table 5). As Shortwave Infrared 2 is not used for any index calculation above, therefore, using this index

aims at testing the band as well as the other two bands. Like other indices, D_{1650nm} values from the observed Landsat-8 OLI images are well matched to those from the predicted images as shown in Fig. 15.

In summary, as shown in Figs. 11 to 15, all the indices from the predicted images match well to those from the observed images and the time series of indices catches well the trajectory trends of indices presented by available Landsat images for all the vegetated land cover types. The observation is true for all the 5 different indices of all the 7 land cover types within the study area. It is worth noticing that not only the average values of indices of all the 7 vegetation types from the predicted time-series image match well to those from the observed images, their uncertainties (standard deviation) are also in the same magnitude of those of the observations at the corresponding dates, i.e., if the index uncertainty of the observed Landsat image is large, the index uncertainty from the predicted image at the same date or close to the date is large, and vice versa (Figs. 11 to 15).

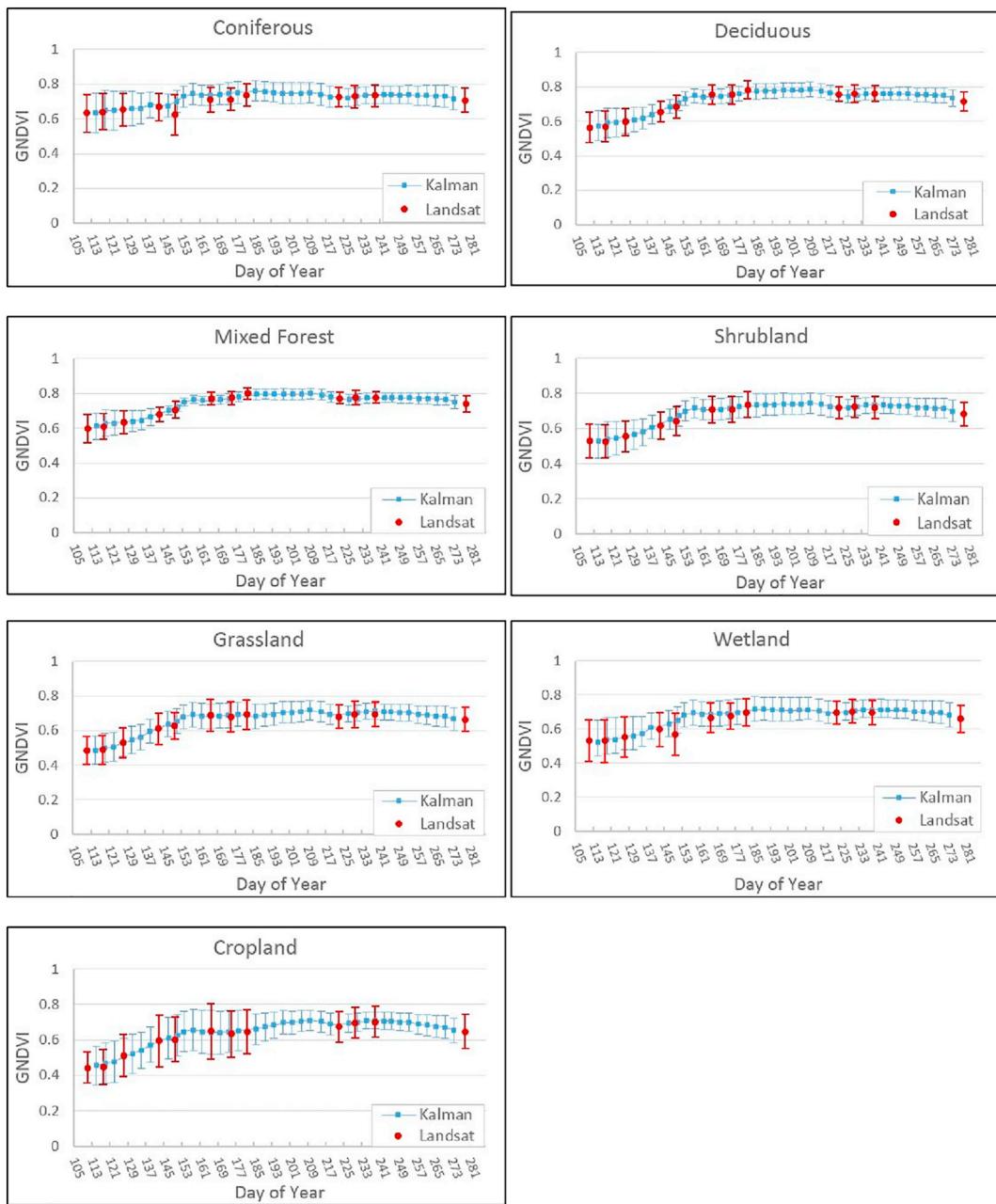


Fig. 12. Time-series of GNDVI of vegetated land cover types of the study area.

It is also worth noticing that different land cover types have different index uncertainties. For example, NDVIs of forested land cover types have smaller uncertainty than Cropland. This is largely reflected in the uncertainties of the blended images and is discussed in the following section.

In addition to quality validation and visual assessment using two images presented in Section 4.2, the purpose of the vegetation index computation is, from another angle, to verify whether all the bands of the images are equally well blended for the entire period of the time-series by using all available clear-sky Landsat observations. It is noticed that, in general, there are two approaches to produce blended indices: “Index-then-Blend” or “Blend-then-Index” [Jarihani et al., 2014; Chen et al., 2018]. The former is to generate index first, then blend the index; and the latter is to blend images first, and then to generate indices from the blended images. According to the study [Jarihani et al., 2014], Index-Then-Blend approach could produce better output than Blend-Then-Index approach; However, the recent study [Chen et al., 2018]

suggested both approaches can be considered equivalent.

From the comparison of the vegetation indices generated from our blended time-series images vs. those from the Landsat observations, our study suggests that, accurate vegetation indices can be achieved if the blended images resemble true earth observations. On the one hand, if an image fusion model can blend image accurately for all the bands, “Blend-then-Index” should be more efficient as the blended images can be used not only for index generation, but also for other applications such as land surface monitoring. On the other hand, “Index-then-Blend” can be more efficient and accurate, as only index needs to be blended once an index is computed (vs. two or three bands blending required for an index generation). However, due to limited studies on this topic, an agreement on which approach performs better has not been reached yet. Therefore, the advantages and disadvantages of the two approaches need further investigation.

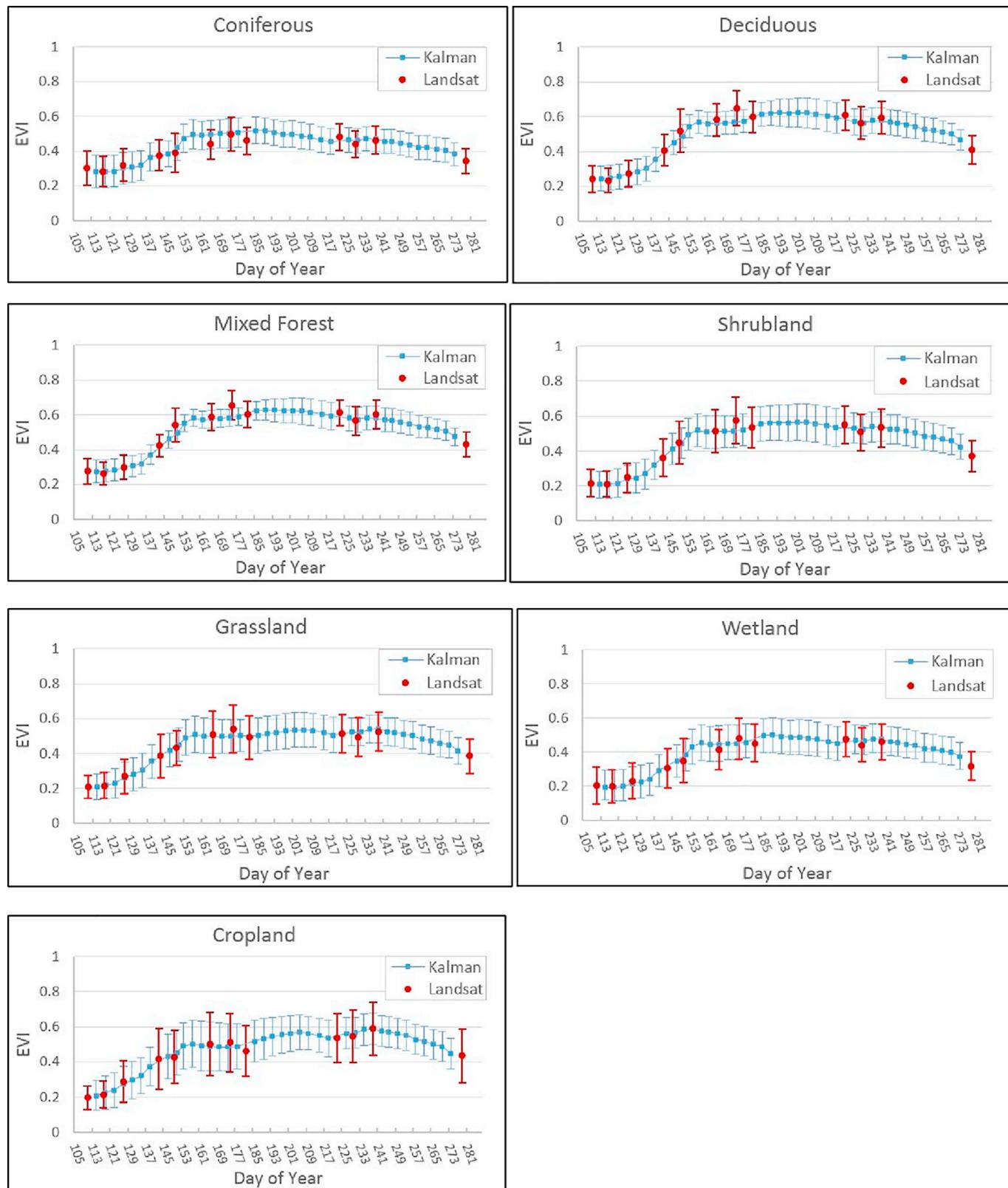


Fig. 13. Time-series of EVI of vegetated land cover types of the study area.

4.3.2. Uncertainty analysis

KFRFM is able to incorporate uncertainty of input images during its process and estimate uncertainty for its produced synthetic images. Uncertainty of a predicted synthetic image comes mainly from two source groups: The first is the uncertainty of the input images, and the

second is the modelling process which propagates the uncertainty of the input images to the processed output. As the input Landsat and MODIS surface reflectance data do not have explicit uncertainty values attached to them, we use an empirical value of a standard deviation 0.004 for both Landsat-8 OLI and MODIS image pixels which was also used by

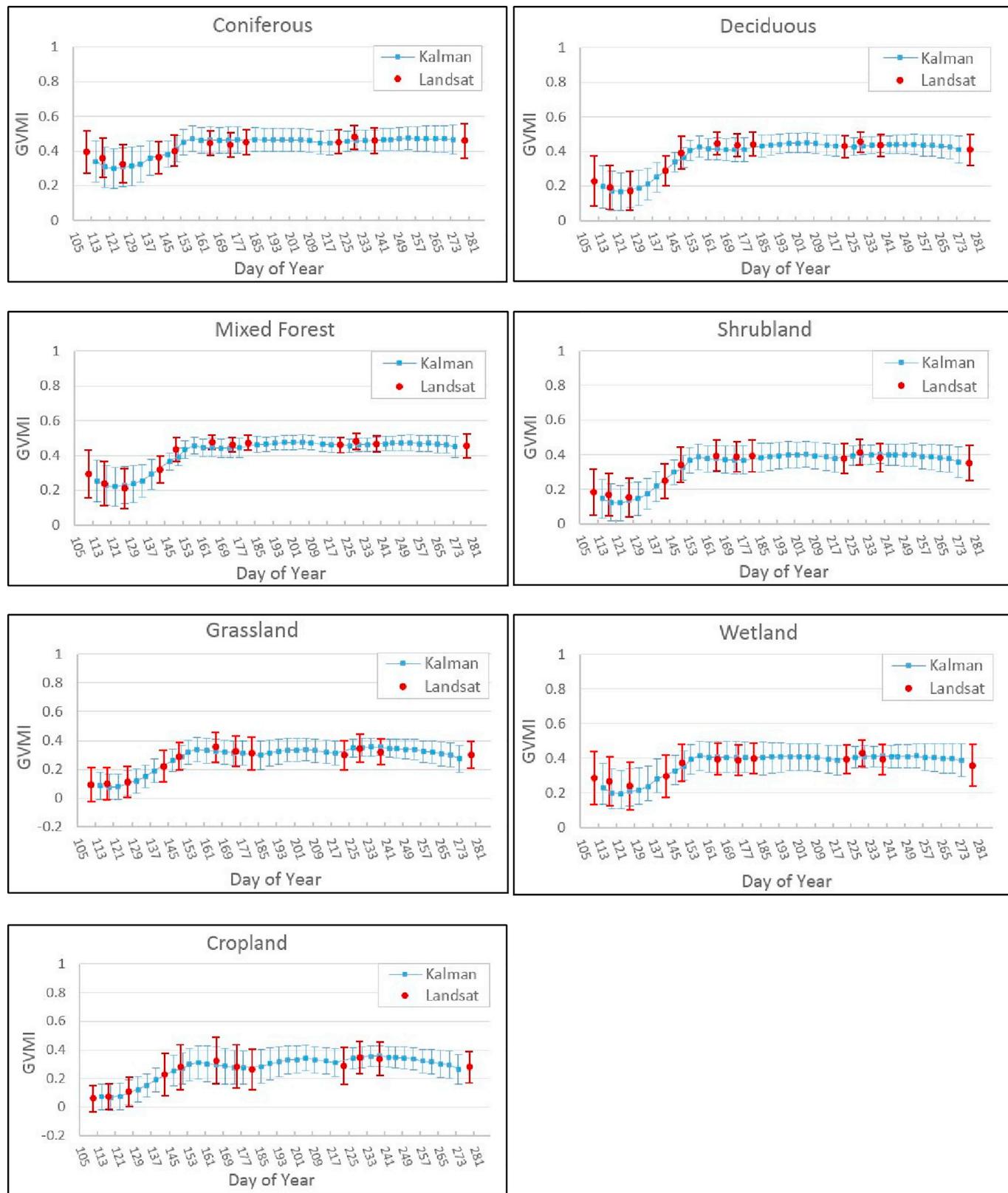


Fig. 14. Time-series of GVMI of vegetated land cover types of the study area.

other studies (Gao et al., 2006; Zhong and Zhou, 2018). Therefore the uncertainty of a synthetic images produced by the model is based on the empirical values of the input images.

In this study, the uncertainty of the predicted synthetic images is assessed in two ways. First, we analyze the overall uncertainty of a

synthetic image by individual band; second, we conduct the uncertainty analysis for each and all the vegetation types within the study area to evaluate if all the vegetation types are equally well predicted or to find any abnormal prediction variations.

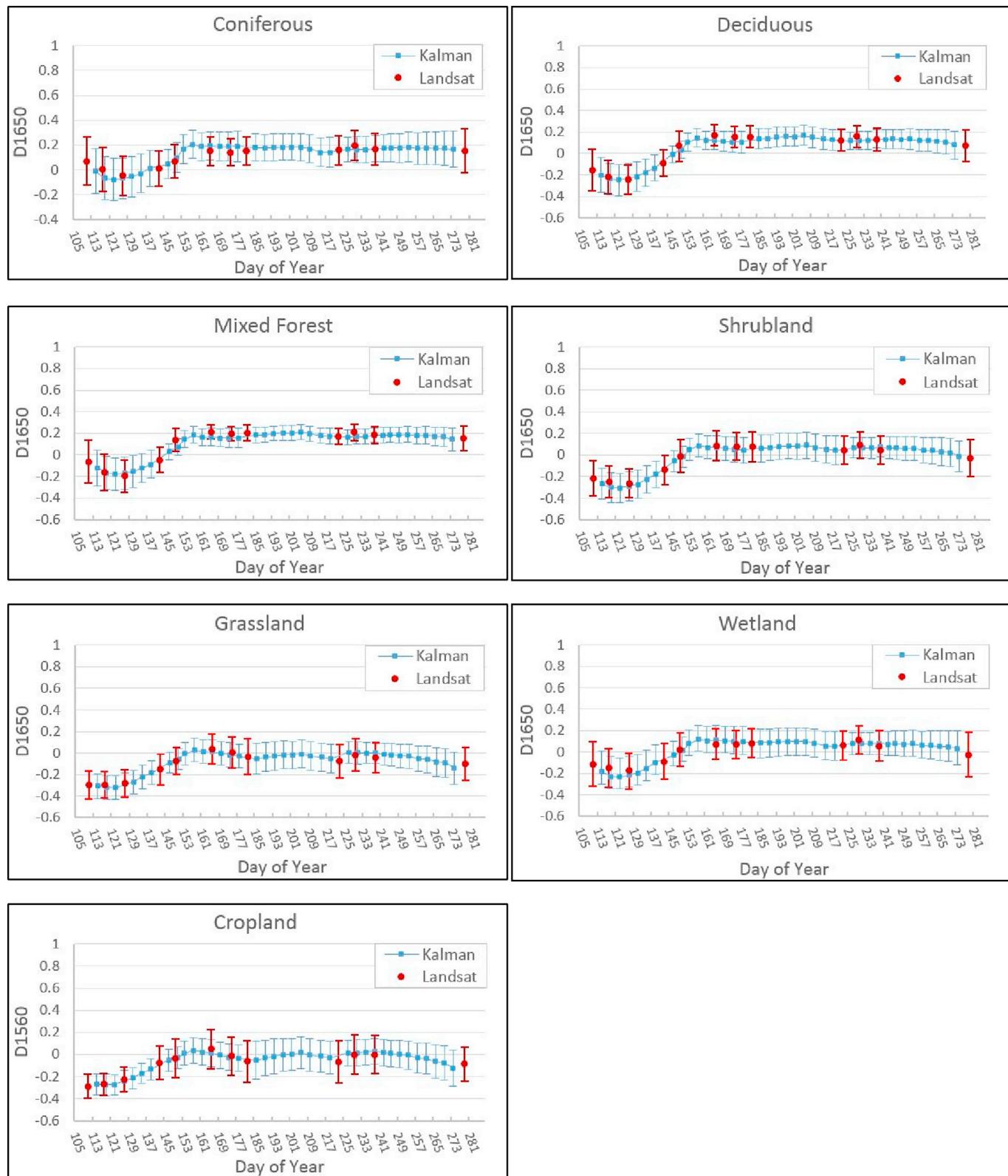


Fig. 15. Time-series of D_{1650} of vegetated land cover types of the study area.

4.3.3. Overall uncertainty analysis

The uncertainty of a synthetic Landsat image is expressed as standard deviation of every pixel according to Eq. (28). The first attempt for uncertainty analysis in this study is to examine the overall uncertainty distribution of every band of the predicted time-series of 38 synthetic Landsat images and their temporal variation. The overall uncertainty of

each band of an image is the average of the standard deviation of all the pixels of the band regardless the land cover types, which is used to check if all the blended bands are equally well predicted and their differences.

Fig. 16 shows the time-series of the averaged variance of all the pixels of all the 6 bands blended. Overall, although there are some

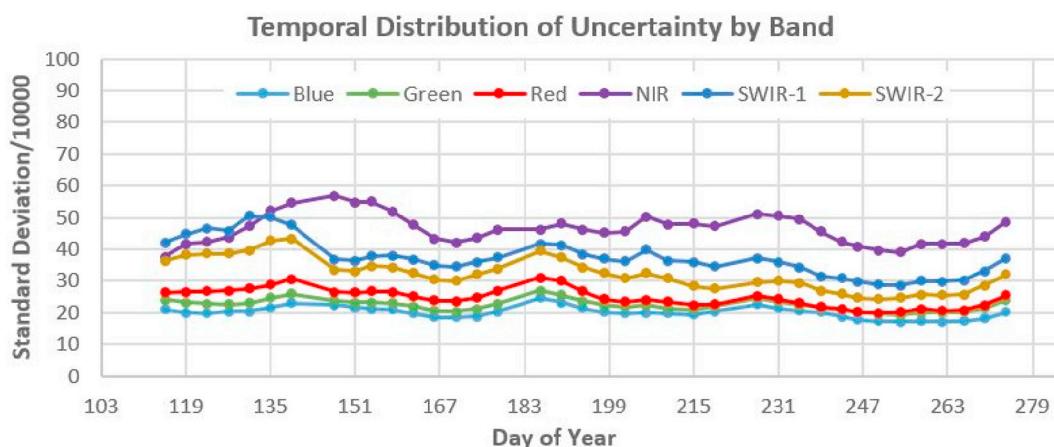


Fig. 16. Temporal distribution of uncertainty by band.

temporal variations of the uncertainty, there is no obvious trend of the uncertainty in the blending period which spans about 6 months. In general, the uncertainties of bands 2 (Blue), 3 (Green), and 4 (Red) are smaller than those of bands 5 (NIR), 6 (SWIR-1), and 7 (SWIR-2). Since band 5 is Near Infrared, and bands 6 and 7 are in the spectrum of shortwave Infrared, their reflectance values and variations are larger than those of bands 2, 3, and 4 for vegetation land covers. Among the six bands, NIR band has the largest uncertainty throughout the study period except a few dates at the beginning, SWIR-1 has the largest uncertainty. This can be confirmed by the validation of the two predicted synthetic Landsat images presented in Section 4.2. Tables 6 and 9 show that Blue, Green, and Red bands have a smaller RMSE than NIR, SWIR-1 and SWIR-2 have.

For the entire period, we used 5 pairs of Landsat and MODIS images to predict the time-series synthetic Landsat images. Between every two pairs of the input images a number of synthetic images are predicted (Table 3). In addition to the modelling process, uncertainty of the predicted synthetic images is affected by the given empirical (a priori) uncertainty values of the input images as well as by the actual quality of the images themselves. Due to cloud/cloud shadow/haze residues of some pixels of input images, these non-clear-sky pixels would have some negative effects on quality of the predicted images.

The temporal distribution and variations of the uncertainty therefore not only largely indicate the quality of the predicted images, but also give us some clues of the quality of the input images. In summary, the uncertainty of a predicted synthetic image, in general, reflects the quality of a predicted image, and at the same time, the uncertainty

information also can be used to identify quality issues of input images.

As shown in Fig. 16, the time-series of uncertainty of each band is consistent throughout the entire study period although there are some variations. These variations likely come from the varied qualities of the input images. The visible bands have smaller values as well as smaller fluctuations, while the near infrared band has the largest values and temporal variation. This is reasonable since the near infrared bands have the largest temporal values and range for vegetation land covers. Fig. 16 also reveals that KFRFM performs equally well for any temporal interval between two pairs of calibration images (Table 4).

4.3.4. Uncertainty by land cover type

The land cover map used in the analysis is the Circa 2010 Land Cover of Canada by Latifovic et al. (2017). The 30 m land cover map of Canada for the base year 2010 is generated from Thematic Mapper and Enhanced Thematic Mapper Landsat sensor observations. The land cover map is the most current such kind of product of the country with 30 m spatial resolution. As the study area is a rural region and no significant land cover change occurs since the mapping, the land cover map still well presents its reality of the study area. The land cover classification system level II of the North American Land Change Monitoring System has 19 land cover types. There are ten land cover types within the study area as listed in Table 2. Among the ten land cover types, seven of them are vegetation types. Uncertainty of all the 7 vegetated land covers is analyzed for each and every band. Figs. 17–22 show the temporal behaviors of the uncertainty of the land cover types by band.

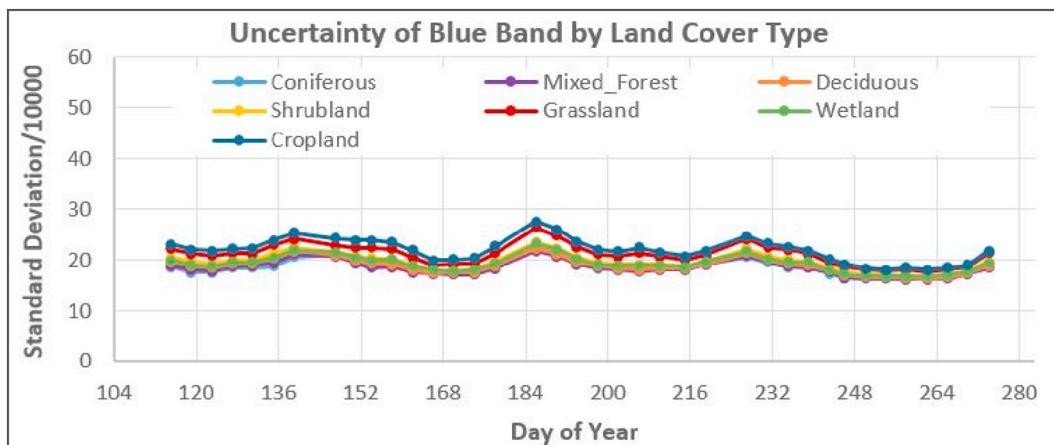


Fig. 17. Uncertainty of Blue band by land cover types. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

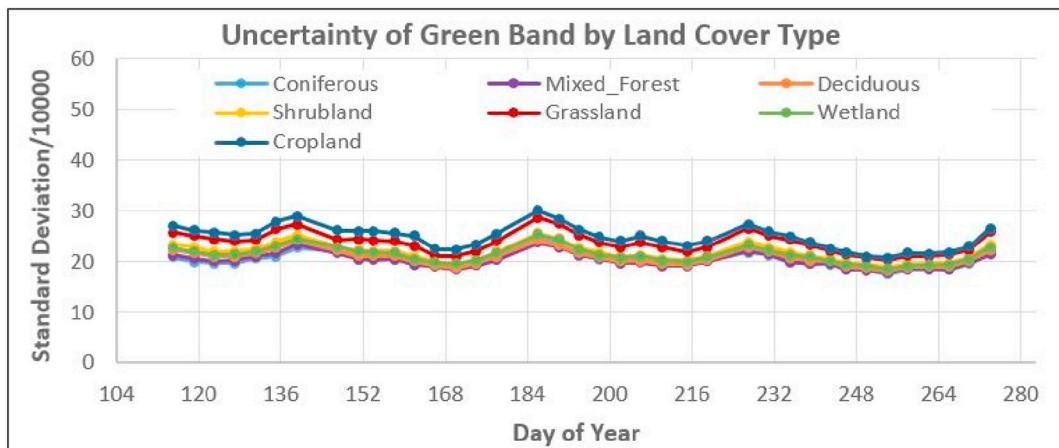


Fig. 18. Uncertainty of Green band by land cover types. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

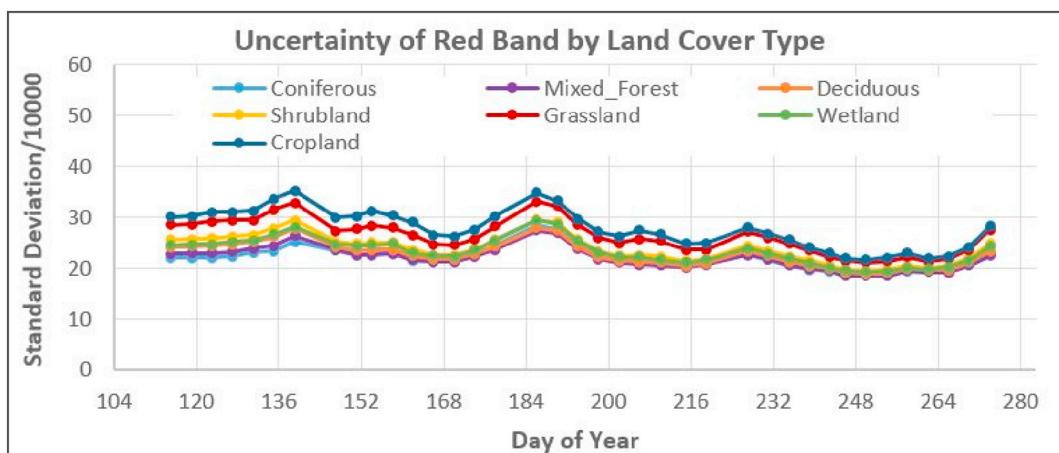


Fig. 19. Uncertainty of Red band by land cover types. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

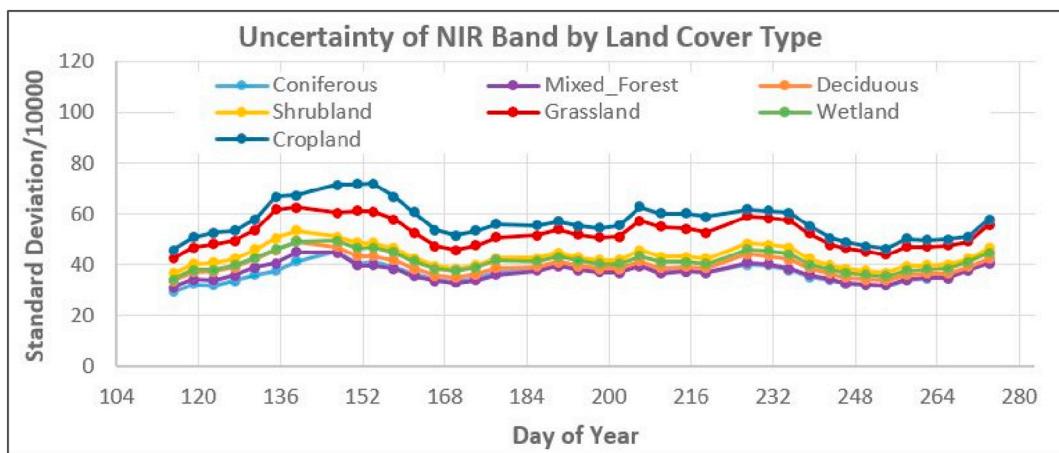


Fig. 20. Uncertainty of Near Infrared band by land cover types.

Overall, the temporal variation of the vegetation uncertainty of every band (Figs. 17–22) is very similar to the average uncertainty (Fig. 16) of all the land cover types: vegetated land covers have a smaller uncertainty of Blue, Green and Red bands than that of NIR, SWIR-1, and SWIR-2 bands. Among the 7 vegetated land types, almost for all the bands, cropland and grassland have a larger uncertainty, and forest land cover types (deciduous, coniferous, and mixed forest) have a

smaller uncertainty. It is likely due to growing environment of cropland and grassland, which is mostly affected by weather conditions and human interference such as fallow fields of cropland and grazing on grassland. Therefore cropland and grassland are usually less accurately predicted. In addition, if the overall uncertainty is larger at some dates, the uncertainty of all the vegetated land cover types is larger at the same dates although there are some variations among these land cover

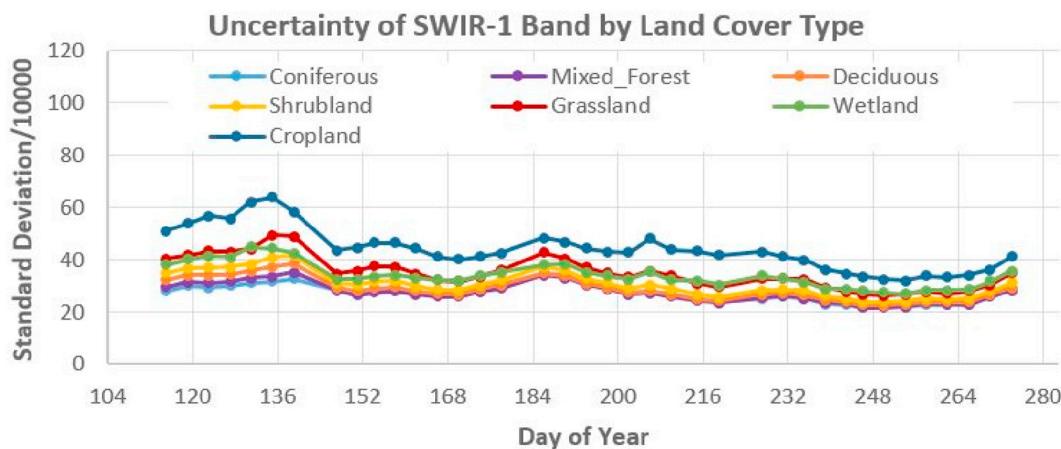


Fig. 21. Uncertainty of Shortwave Infrared 1 band by land cover types.

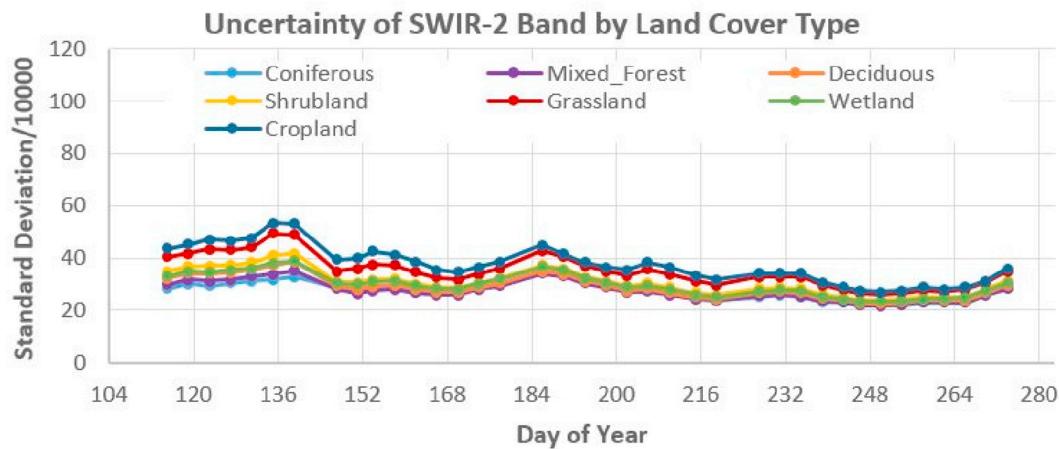


Fig. 22. Uncertainty of Shortwave Infrared 2 band by land cover types.

Table 11

Vegetation indices and index formulas - Normalization Difference Vegetation Index (NDVI), Green Normalization Difference Vegetation Index (GNDVI), Enhanced Vegetation Index (EVI), Global Vegetation Moisture Index (GVMI), and Depth of 1650 nm (D_{1650}).

Index name	Bands used	Equation
NDVI	Red, NIR	$(NIR - Red) / (NIR + Red)$
GNDVI	Green, NIR	$(NIR - Green) / (NIR + Green)$
EVI	Blue, Red, NIR	$2.5 * (NIR - Red) / (NIR + 6 * Red - 7.5 * Blue + 1)$
GVMI	NIR, SWIR1	$((NIR + 0.1) - (SWIR1 + 0.02)) / ((NIR + 0.1) + (SWIR1 + 0.02))$
$D_{1650\text{nm}}$	NIR, SWIR1, SWIR2	$1 - (SWIR1 / (NIR * (1 - 0.59359) + 1 - 0.59359 * SWIR2))$

types.

In general, a larger uncertainty implies a larger reflectance range of a land cover type, which could result in a larger variation of vegetation indices, and a smaller uncertainty means a smaller reflectance range, and a smaller variation of vegetation indices. For example, in our study, Cropland has largest uncertainties throughout the entire study period for all the 6 bands (Figs. 17–22) among all the land cover types, and it has the largest uncertainty (standard deviation) of all the vegetation indices (Figs. 11 to 15), while forest land covers (Coniferous, Deciduous, and Mix-forest) have the smallest uncertainties (Figs. 17–22), they have the smallest standard deviations of their vegetation indices (Figs. 11 to 15). Therefore it can be seen that the uncertainty can largely reveal the quality of the blended images as well as the by-products produced from them.

The uncertainty of a synthetic image is from multi-sources. In detail, the first is the empirical uncertainty values of the input Landsat and MODIS images. A larger uncertainty of the input images, the larger

uncertainty of the predicted images (Zhong and Zhou, 2018). Second, the quality of the input images affects the quality of the predicted images. In order to predict a high quality synthetic image, clear-sky images are required. If any input Landsat or MODIS image (at date t_0 or t_2) has cloud/haze residues, it would have a negative impact on all the predicted images at date t_k between the two dates (t_0 or t_2) although the Kalman Filter algorithm would greatly reduce the impact through Kalman gain based on the uncertainties of the state variable estimate and the observation estimate. For predicting a synthetic Landsat image at t_k , a clear-sky MODIS image at t_k is required. Similarly, if the input MODIS image is of poor quality, the predicted image resulted from the MODIS image would have a larger uncertainty. Third, to predict a time-series of synthetic Landsat images, MCD43A4 images at a regular time interval of 4 days are used. MCD43A4 is produced by using 16 days retrieval period for BRDF inversion. During a fast vegetation growing/withering stage, phenology and spectral information of such kind of vegetation pixels is quite different. Reflectance variations from these

situations would also have some impact on the quality of the predicted image. In addition, pixels with full BRDF inversion have a higher quality adjustment than those pixels with magnitude BRDF inversion (Wang et al., 2018). Without distinguishing them in uncertainty computation the uncertainty outputs may not accurately reflect the reality. KFRFM computes reflectance change rate between the calibration dates to the prediction date by unmixing coarse resolution images. Accurate clusters of land cover types would result in a better reflectance change computation. In this regard, using more advanced clustering or unmixing algorithm such as proposed by Du et al. (2019) would produce a more accurate blending output.

The last is the spectral compatibility of Landsat-8 OLI and MODIS sensors although their spectral information is similar.

4.4. Issues of heterogeneous regions and land cover changes for image blending

4.4.1. Heterogeneous regions

One of the main questions that a fusion model of blending coarse and fine resolution images needs to address is heterogeneous regions (mixed pixels) of the coarse resolution image. Some models such as ESTARFM and STAARCH explicitly address heterogeneous regions in their algorithm. For example, ESTARFM introduces a conversion coefficient to each endmember of a coarse resolution pixel to handle mixed pixels. In our Kalman filter reflectance fusion model, we address this issue by unmixing a coarse resolution pixel (i.e., a MODIS pixel) based on the corresponding input fine resolution image (i.e., Landsat) classification. We first cluster the input Landsat image into multiple reflectance similar groups. A mixed pixel of a coarse resolution image can contain multiple land cover types, therefore multiple clusters. Assuming the same cluster within all the pixels has the same change rate for a temporal period from t_k to t_{k+1} , based on these clusters and through linear systems of equations (Eq. (11)), we derive the change rate of the endmembers (clusters) of a coarse resolution pixel. Then the change rate of a cluster is added to the corresponding observation Landsat pixel to form the blending output. The number of clusters can be predetermined in a certain range [C_{min}, C_{max}] according to the land surface conditions of the study region and the study period. The final optimized number of clusters can be determined with the model optimization process (Section 2.2.4).

The model assessment and comparison in a heterogeneous landscape of Coleambally, Australia of a time-series of Landsat-7 ETM+ and MODIS (MOD09GA) imagery fusion indicated that, statistically, our proposed model KFRFM performed better than the two compared models, including ESTARFM.

4.4.2. Land cover change

Another question that a fusion model should deal with is land cover change, especially for time-series coarse and fine resolution image blending practice. This issue is more challenging for a fusion model with only one pair of input images as any land cover change occurring at date t_k (t_k is later than t_0) is only captured by the coarse resolution image at t_k or later date. A fusion model with input of two pairs of coarse and fine resolution images has some advantages over a fusion model with only one pair of input images since land cover change information can be recorded by both the coarse resolution image at time t_k and the image pair at time t_n (t_n is later than t_k). The land cover change information logged in the images at the two time stamps, especially in the image pair at t_n , largely contribute to the blended fine resolution image at t_k .

Our proposed KFRFM has advantages for better carrying the land cover change information to the blended images from t_k up to t_n , compared to other blending methods. In general, for a fusion model with two pairs of input images, the blended final resolution image at t_k is composed of the forward blended image t_k from the images at t_0 and t_k and the backward blended image at t_k from the images at t_k and t_n .

That is, the input image at t_0 without land cover change information always contributes to the forward blended image at t_k ($k = 1, 2, \dots, n-1$), which would reflect less accurately the land cover change information. However, our proposed KFRFM is a continuous process of the blended images, i.e., the fine resolution image blended at t_k directly contributes to the next blended image at t_{k+1} , i.e., the forward blended image is based on the blended image at t_k and the input image at t_{k+1} instead of images at t_0 and t_{k+1} . As the blended image at t_k is from the system model (Section 2.2.1) as well as the observation model (Section 2.2.2), the blended image at t_k contains the land cover change information from the images at t_k and t_n . Therefore, all blended images after t_k can reproduce more accurately the land cover change situation.

To further improve the model's capability for land cover change, a residual adjustment method for land cover changes such as used by FSDAF and PSRFM is implemented. The residue of a pixel is the difference of the reversely computed MODIS pixel value from the predicted synthetic Landsat image at t_k and the observed MODIS pixel value at t_k . The residue is mainly resulted from un-captured land cover change information. The residue adjustment is detailed in Section 2.2.4.

5. Conclusions

In this study we developed a new surface reflectance fusion model based on a Kalman Filter algorithm (Kalman Filter Reflectance Fusion Model - KFRFM) for blending images with different spatial resolutions. KFRFM has capability to handle two challenging issues of land cover change and heterogeneous regions that fusion models face. In addition, compared to other fusion models, one significant advantage of the model is that it can quantify the uncertainty of the blended images, and smooth the synthetic image prediction based on the uncertainty information.

Two synthetic Landsat images predicted by KFRFM are evaluated against actual Landsat observations visually and quantitatively. Comparison results show that the model is able to predict high quality of synthetic Landsat images from MODIS observations. Performance of KFRFM is also assessed against three well-developed reflectance fusion models (STARFM, ESTARFM, and FSDAF) using the same images. Comparison results reveal that KFRFM performs the best among all the compared models based on the assessed quality indices. KFRFM is also assessed and compared to ESTARFM and FSDAF in a heterogeneous landscape by fusion of a time-series Landsat-7 ETM+ and MODIS images. Based on the validation and comparison of the 6 bands of the 15 fused synthetic Landsat images against observations, it can be concluded that the three models can predict Landsat images well in a heterogeneous environment, and statistically, KFRFM performed better than the two compared models.

For further evaluation, we used the model to predict a time-series of synthetic Landsat images for a vegetation growing season spanning about 6 months from five pairs of Landsat-8 OLI and MODIS images (MCD43A4) and a time-series of Nadir Bi-directional Reflectance Distribution Function (BRDF) adjusted MODIS images (MCD43A4). The time-series of synthetic Landsat images predicted consists of 38 images by a temporal interval of 4 days, from which five time-series of vegetation indices that involve all the 6 spectral bands (Blue, Green, Red, Near-Infrared, Shortwave Infrared 1, and Shortwave Infrared 2) are calculated and compared to those from Landsat-8 OLI observations of the same period. The results demonstrate that all the indices are well aligned with those from Landsat observations for all the vegetated land cover types over the entire period of the time-series.

Companied with the time-series of blended images, uncertainty of every pixel and every band is computed. The uncertainty reflects largely the quality of the blended images, which is well shown with the vegetation indices and their standard deviation. Uncertainty analysis of the predicted time-series of synthetic Landsat images band by band for all the vegetation types of the study area reveals that no obvious abnormal behavior of all the spectral bands and vegetation types was

discovered. This indicates that KFRFM performs equally well for all the spectral bands blended, all the vegetation types, and for the entire time period. On the one hand, in general, the uncertainty magnitude indicates relative accuracy of the synthetic images at both temporal and thematic dimensions, and on the other hand, it helps identify possible quality issues of input images such as cloud and haze residues on the images, and interpret and use the data more accurately.

Therefore it can be concluded that KFRFM is able to produce time-series synthetic Landsat images of high quality from MODIS observations to supplement clear-sky Landsat images for applications that require more frequent, medium spatial resolution images which are usually unobtainable due to cloud coverage and/or low revisit frequency of Landsat platform. Hence, to fill the temporal gap of Landsat observations and other sensors, KFRFM can be used effectively to generate dense time-series synthetic Landsat images for timely monitoring dynamic environmental changes such as vegetation growing conditions at a local scale.

Despite the advantages of the model, more study for the model improvement is needed. For example, how to define and quantify the uncertainty of each pixel based on the pixel's quality information of both Landsat and MODIS input images is a question to be investigated further. It is also worth to explore if the quality information, such as full BRDF inversion and magnitude BRDF inversion for Nadir BRDF adjustment of MCD43A4 images, can be quantified as a part of uncertainty information of the input images for uncertainty computation of the model outputs. KFRFM can also be improved by merging cloud-free Landsat observations between the two pairs of images when new observations are available but partially cloud covered during the update process of KFRFM. This potential improvement deserves further investigation. One disadvantage of KFRFM is that it requires input of two pairs of images. Therefore it is not able to perform extrapolation prediction.

Author contributions

Zhou Fuqun: Conceptualization, Methodology, Investigation, Validation, Formal analysis, Writing- Original draft preparation & editing, Funding acquisition, Supervision.

Zhong Detang: Conceptualization, Methodology, Software, Investigation, Writing - Review & Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study is a part of the research and development project “Big Data Analytics of Earth Observation Data in Support of Evidence-Based Decision Making for Climate Change” funded by Natural Resources Canada and Canadian Space Agency’s GRIP program.

Authors would like to thank the Editor Dr. Junhua Li and Dr. Najib Djamai, and the anonymous reviewers for their valuable comments and suggestions for improvement of the manuscript. Authors are also grateful to the researchers of STARFM, ESTARFM and FSDAF for providing free software of their models for our model comparison.

Author contributions

Fuqun Zhou and Detang Zhong jointly conducted all the work presented in this manuscript, including concept development, system development, experiment implementation, data analysis, as well as manuscript preparation.

Declaration of competing interest

The authors declare no conflict of interest.

References

- Amorós-López, J., Gómez-Chova, L., Alonso, L., Guanter, L., Zurita-Milla, R., Moreno, J., Camps-Valls, G., 2013. Multitemporal fusion of Landsat/TM and ENVISAT/MERIS for crop monitoring. *Int. J. Appl. Earth Obs. Geoinf.* 23, 132–141. <https://doi.org/10.1016/j.jag.2012.12.004>.
- Arbel, D., Cohen, E., Citroen, M., Blumberg, D.G., Kopeika, N.S., 2004. Landsat TM satellite image restoration using Kalman filters. *Photogramm. Eng. Remote. Sens.* 70, 91–100 2004.
- Ceccato, P., Gobron, N., Flasse, S., Pinty, B., Tarantola, S., 2002a. Designing a spectral index to estimate vegetation water content from remote sensing data: part 1: theoretical approach. *Remote Sens. Environ.* 82, 188–197 2002.
- Ceccato, P., Flasse, S., Grégoire, J.-M., 2002b. Designing a spectral index to estimate vegetation water content from remote sensing data: part 2. Validation and applications. *Remote Sens. Environ.* 82 (2002), 198–207.
- Chen, X.H., Liu, M., Zhu, X.L., Chen, J., Zhong, Y., Cao, X., 2018. “Blend-then-index” or “index-then-blend”: a theoretical analysis for generating high-resolution NDVI time series by STARFM. *Photogramm. Eng. Remote. Sens.* 84, 65–73. <https://doi.org/10.14358/PERS.84.2.65>.
- Du, B., Wei, Q.C., Liu, R., 2019. An improved quantum-behaved particle swarm optimization for endmember extraction. *IEEE Trans. Geosci. Remote Sens.* <https://doi.org/10.1109/TGRS.2019.2903875>. (2019).
- Emelyanova, I.V., McVicar, T.R., Van Niel, T.G., Li, L.T., van Dijk, A.I.J.M., 2013a. Assessing the accuracy of blending Landsat–MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: a framework for algorithm selection. *Remote Sens. Environ.* 133 (2013), 193–209. <https://doi.org/10.1016/j.rse.2013.02.007>.
- Emelyanova, I., McVicar, T., Van Niel, T., Li, L., Van Dijk, A., 2013b. Landsat and MODIS Data for the Coleambally Irrigation Area. v1. CSIRO. *Data Collection*. <https://doi.org/10.4225/08/5111AC0BF1229>.
- Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* 44 (8), 2207–2218 2006.
- Gao, F., Masek, J., Wolfe, R., Huang, C., 2010. Building consistent medium resolution satellite data set using moderate resolution imaging spectroradiometer products as reference. *J. Appl. Remote. Sens.* 2010 (4), 43526. <https://doi.org/10.1117/1.3430002>.
- Gevaert, C.M., García-Haro, F.J., 2015. A comparison of STARFM and an unmixing based algorithm for Landsat and MODIS data fusion. *Remote Sens. Environ.* 156, 34–44. <https://doi.org/10.1016/j.rse.2014.09.012>. 2015.
- Gitelson, A.A., Kaufman, Y., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing Environ.* 58, 289–298.
- Hilker, T., Wulder, M.A., Coops, N.C., Linke, J., McDermid, G., Masek, J.G., Gao, F., White, J.C., 2009a. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sens. Environ.* 113, 1613–1627 2009.
- Hilker, T., Wulder, M.A., Coops, N.C., Seitz, N., White, J.C., Gao, F., Masek, J.G., Stenhouse, G., 2009b. Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sensing Environ.* 113, 1988–1999.
- Hirvonen, R.A., 1971. Adjustments by Least Squares in Geodesy and Photogrammetry. 1971 Ungar, New York 0804443971 261 p. ISBN 978-0804443975.
- Hu, J., Ding, X.L., Li, Z.W., Zhu, J.J., Sun, Q., Zhang, L., 2013. Kalman-filter-based approach for multisensor, multitrack, and multitemporal InSAR. *IEEE Trans. Geosci. Remote Sens.* 51, 4226–4239 2013.
- Huang, B., Song, H., 2012. Spatiotemporal reflectance fusion via sparse representation. *IEEE Trans. Geosci. Remote Sens.* 50 (10), 3707–3716. <https://doi.org/10.1109/TGRS.2012.2186638>.
- Huang, B., Zhang, H., 2014. Spatio-temporal reflectance fusion via unmixing: accounting for both phenological and land-cover changes. *Int. J. Remote Sens.* 35, 6213–6233 2014.
- Huang, C., Li, X., Lu, L., Gu, J., 2008. Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. *Remote Sens. Environ.* 112, 888–900 2008.
- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83 (1), 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2).
- Jarihani, A.A., McVicar, T.R., Van Niel, T.G., Emelyanova, I.V., Callow, J.N., Johansen, K., 2014. Blending Landsat and MODIS data to generate multispectral indices: a comparison of “Index-then-Blend” and “Blend-then-Index” approaches. *Remote Sens.* 6, 9213–9238. <https://doi.org/10.3390/rs6109213>.
- Kalman, R.E., 1960. A new approach to linear filtering and prediction problems 1. *J. Basic Eng.* 82, 35–45 1960.
- Kleynhans, W., Olivier, J.C., Wessels, K.J., Salmon, B.P., van den Bergh, F., Steenkamp, K., 2011. Detecting land cover change using an extended Kalman filter on MODIS NDVI time-series data. *IEEE Geosci. Remote Sens. Lett.* 8, 507–511 2011.
- Latifovic, R., Pouliot, D., Olthof, I., 2017. Circa 2010 land cover of Canada: Local optimization methodology and product development. *Remote Sens.* 9 (11) 1 November

- 2017.
- Ma, J., Zhang, W., Marinoni, A., Gao, L., Zhang, B., 2018. An improved spatial and temporal reflectance unmixing model to synthesize time series of Landsat-like images. *Remote Sens.* 10 (9), 1388 2018.
- Matheron, G., 1967. Kriging or polynomial interpolation procedures—a contribution to polemics in mathematical geology. *Can. Min. Metall. Bull.* 60, 240–244 1967.
- van Niel, T.G., McVicar, T.R., Fang, H., Liang, S., 2003. Calculating environmental moisture for per-field discrimination of rice crops. *Int. J. Remote Sens.* 2003 (24), 885–890.
- Quaife, T., Lewis, P., Dekauwe, M., Williams, M., Law, B., Disney, M., Bowyer, P., 2008. Assimilating canopy reflectance data into an ecosystem model with an ensemble Kalman filter. *Remote Sens. Environ.* 2008 (112), 1347–1364.
- Rao, J.M., Rao, C.V., Kumar, A.S., Krishna, B.G., Dadhwal, V.K., 2015. Spatiotemporal data fusion using dictionary learning and temporal edge primitives. *I.J. Image, Graphics and Signal Processing* 2015 (10), 47–59. <https://doi.org/10.5815/ijigsp.2015.10.06>.
- Samain, O., Roujean, J., Geiger, B., 2008. Use of a Kalman filter for the retrieval of surface BRDF coefficients with a time-evolving model based on the ECOCLIMAP land cover classification. *Remote Sens. Environ.* 2008 (112), 1337–1346.
- Schaaf C.B., F.Gao, A.H. Strahler, W. Lucht, X.Li, T. Tsang, N.C. Strugnell, X. Zhang, Y. Jin, J.P. Muller, P. Lewis, M. Barnsley, P. Hobson, M. Disney, G. Roberts, M. Dunderdale, C. Doll, R.P. D'Entremont, B. Hu, S. Liang, J.L. Privette, D. Roy. 2002. First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sens. Environ.*, 83 (2002), pp.135–148, [https://doi.org/10.1016/S0034-4257\(02\)00091-3](https://doi.org/10.1016/S0034-4257(02)00091-3).
- Sedano, F., Kempeneers, P., Hurtt, G., 2014. A Kalman filter-based method to generate continuous time series of medium-resolution NDVI images. *Remote Sens.* 2014 (6), 12381–12408. <https://doi.org/10.3390/rs61212381>.
- Song, H., Huang, B., 2013. Spatiotemporal satellite image fusion through one-pair image learning. *IEEE Trans. Geosci. Remote Sens.* 51 (4), 1883–1896 2013.
- Sun, S.-L., Deng, Z.L., 2004. Multi-sensor optimal information fusion Kalman filter. *Automatica* 2004 (40), 1017–1023.
- Vicente-Guijalba, F., Martínez-Marin, T., Lopez-Sánchez, J.M., Member, S., 2014. Crop phenology estimation using a multitemporal model and a Kalman filtering strategy. *IEEE Geosci. Remote Sens. Lett.* 2014 (11), 1081–1085.
- Wang, J., Huang, B., 2017. A rigorously-weighted spatiotemporal fusion model with uncertainty analysis. *Remote Sens.* 2017 (9). <https://doi.org/10.3390/rs9100990>.
- Wang, Z., Bovik, A.C., 2002. A universal image quality index. *IEEE Signal Processing Letters* 9, 81–84 (Mar. 2002).
- Wang, Z., Schaaf, C.B., Sun, Q., Shuai, Y., Román, M.O., 2018. Capturing rapid land-surface dynamics with collection V006 MODIS BRDF/NBAR/albedo (MCD43) products. *Remote Sens. Environ.* 207, 50–64. <https://doi.org/10.1016/j.rse.2018.02.001>.
- Weier J., Herring, D., 2000. Measuring Vegetation (NDVI & EVI), https://earthobservatory.nasa.gov/Features/MeasuringVegetation/measuring_vegetation_1.php, vol. 2000.
- Welch, G., Bishop, G., 2006. *An Introduction to the Kalman Filter*. 2006 University of North Carolina, Chapel Hill, NC, USA.
- Wolf, P.R., 1995. Survey measurement adjustments by least squares. In: *The Surveying Handbook*. 1995. pp. 383–413. https://doi.org/10.1007/978-1-4615-2067-2_16.
- Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., Gao, F., Goward, S.N., Helder, D., Helmer, E., Nemani, R., Oreopoulos, L., Schott, J., Thenkabail, P.S., Vermote, E.F., Vogelmann, J., Wulder, M.A., Wynne, R., 2008. Free access to Landsat imagery. *Science*. 23 May 2008: Vol. 320, Issue 5879, pp. 1011 DOI: <https://doi.org/10.1126/science.320.5879.1011a>.
- Wu, B., Huang, B., Zhang, L., 2015. An error-bound-regularized sparse coding for spatiotemporal reflectance fusion. *Geoscience and Remote Sensing, IEEE Transactions on* 53, 6791–6803. <https://doi.org/10.1109/TGRS.2015.2448100>.
- Wu, M., Wang, C., Wang, L., 2012. Use of MODIS and Landsat time series data to generate high-resolution temporal synthetic Landsat data using a spatial and temporal reflectance fusion model. *J. Appl. Remote. Sens.* 6. <https://doi.org/10.1111/1.JRS.6.063507>.
- Zhang, W., Li, A., Jin, H., Bian, J., Zhang, Z., Lei, G., Qin, Z., Huang, C., 2013. An enhanced spatial and temporal data fusion model for fusing Landsat and MODIS surface reflectance to generate high temporal Landsat-like data. *Remote Sens.* 2013 (5), 5346–5368. <https://doi.org/10.3390/rs5105346>.
- Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C., Huete, A., 2003. Monitoring vegetation phenology using MODIS. *Remote Sens. Environ.* 84, 471–475. [https://doi.org/10.1016/S0034-4257\(02\)00135-9](https://doi.org/10.1016/S0034-4257(02)00135-9).
- Zhong, D., Zhou, F., 2018. A prediction smooth method for blending Landsat and moderate resolution imagine Spectroradiometer images. *Remote Sens.* 10 (9), 1371. <https://doi.org/10.3390/rs10091371>. 2018.
- Zhong, D., Zhou, F., 2019. Improvement of clustering method for Modelling abrupt land surface changes in satellite image fusions. *Remote Sens.* 2019 (11), 1759. <https://doi.org/10.3390/rs11151759>.
- Zhou, F., Zhang, A., 2013. Methodology for estimating availability of cloud-free image composites: a case study for southern Canada. *Int. J. Appl. Earth Obs. Geoinf.* 21, 17–31.
- Zhu, X., Chen, J., Gao, F., Chen, X., Masek, J.G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* 114, 2610–2623 2010.
- Zhu, X., Helmer, E.H., Gao, F., Liu, D., Chen, J., Lefsky, M.A., 2016. A flexible spatio-temporal method for fusing satellite images with different resolutions. *Remote Sens. Environ.* 2016 (172), 165–177.
- Zhukov, B., Oertel, D., Lanzl, F., Reinhackel, G., 1999. Unmixing-based multisensor multiresolution image fusion. *IEEE Trans. Geosci. Remote Sens.* 37, 1212–1226 1999.
- Zurita-Milla, R., Clevers, J.G.P.W., Schaepman, M.E., 2008. Unmixing-based Landsat TM and MERIS FR data fusion. *IEEE Geosci. Remote Sens. Lett.* 5, 453–457.