Article

A Prediction Smooth Method for Blending Landsat and Moderate Resolution Imagine Spectroradiometer Images

Detang Zhong * and Fuqun Zhou

Canada Center for Remote Sensing, Canada Centre for Mapping and Earth Observation,
Natural Resources Canada, 6th floor, 560 Rochester Street, Ottawa, ON K1S 5K2, Canada;
fuqun.zhou@canada.ca
* Correspondence: detang.zhong@canada.ca; Tel.: +1-613-759-6192

check for
updates

Abstract: Landsat images have been widely used in support of responsible development of natural resources, disaster risk management (e.g., forest fire, flooding etc.), agricultural production monitoring, as well as environmental change studies due to its medium spatial resolution and rich spectral information. However, its availability and usability are largely constrained by its low revisit frequency. On the other hand, MODIS (Moderate Resolution Imaging Spectroradiometer) images for land studies have much more frequent coverage but with a lower spatial resolution of 250–500 m. To take advantages of the two sensors and expand their availability and usability, during the last decade, a number of image fusion methods have been developed for generating Landsat-like images from MODIS observations to supplement clear-sky Landsat imagery. However, available methods are typically effective or applicable for certain applications. For a better result, a new Prediction Smooth Reflectance Fusion Model (PSRFM) for blending Landsat and MODIS images is proposed. PSRFM consists of a dynamic prediction model and a smoothing filter. The dynamic prediction model generates synthetic Landsat images from a pair of Landsat and MODIS images and another MODIS image, either forward or backward in time. The smoothing filter combines the forward and backward predictions by weighted average based on elapsed time or on the estimated prediction uncertainty. Optionally, the smooth filtering can be applied with constraints based on Normalized Difference Snow Index (NDSI) or Normalized Difference Vegetation Index (NDVI). In comparison to some published reflectance fusion methods, PSRFM shows the following desirable characteristics: (1) it can deal with one pair or two pairs of Landsat and MODIS images; (2) it can incorporate input image uncertainty during prediction and estimate prediction uncertainty; (3) it can track gradual vegetation phenological changes and deal with abrupt land-cover type changes; and (4) for predictions using two pairs of input images, the results can be further improved through the constrained smoothing filter based on NDSI or NDVI for certain applications. We tested PSRFM to generate a Landsat-like image time series by using Landsat 8 OLI and MODIS (MOD09GA) images and compared it to two reflectance fusion algorithms: STARFM (Spatial and Temporal Adaptive Reflectance Fusion Model) and ESTARFM (Enhanced version of STARFM). The results show that the proposed PSRFM is effective and outperforms STARFM and ESTARFM both visually and quantitatively.

Keywords: remote sensing; Landsat; MODIS; image blending; image fusion; reflectance fusion; spatiotemporal data fusion

1. Introduction

Regional land-cover mapping and environmental change detection and monitoring often depend on remote sensing data with both high spatial resolution and acquisition frequency.

However, currently no single remote sensing instrument can meet these requirements. A more practical solution is to "blend" the radiometry from lower resolution sensors with more frequent global observations, for example, Moderate Resolution Imaging Spectroradiometers (MODIS), with data from high/medium-resolution sensors but less frequent coverage, for example, Landsat images. In the last decade, a number of image fusion methods have been developed to predict synthetic Landsat-like images from frequent MODIS images and a limited set of Landsat images. Examples include the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) [1], the Enhanced version of STARFM (ESTARFM) [2], the Spatial Temporal Adaptive Algorithm for Mapping Reflectance Change (STAARCH) [3], the Unmixing-based Spatial-Temporal Reflectance Fusion Model (U-STFM) [4,5], the Flexible Spatiotemporal Data Fusion (FSDAF) model [6], the Spatiotemporal image-fusion model (STI-FM) for enhancing the temporal resolution [7], the Hybrid Color Mapping (HCM) Approach [8] and the Rigorously-Weighted Spatiotemporal Data Fusion Model (RWSTFM) [9] and so forth. However, studies have demonstrated that all of the above reflectance fusion methods have their advantages and disadvantages. Some of them are only effective or applicable for certain applications [1–9]. Interested readers can find concise reviews for these methods in [6,7]. For example, STARFM assumes no land-cover type changes between input and prediction date [1,2]. As a result, they can successfully predict pixels with gradual changes in attributes like vegetation phenology or soil moisture, as these changes are consistent within a spatial partition of the available imagery. However, it is not applicable over heterogeneous land-cover types [2,6]. Song and Huang [10] applied STARFM in an urbanized area and found that it failed to recover pixels with land-cover type changes. ESTARFM and STAARCH were developed to improve the fusion performance over heterogeneous land-cover types but they require two pairs of Landsat and MODIS images as input and are computationally expensive. Emelyanova et al. [11] conducted a comprehensive study to investigate the performance of STARFM and ESTARFM in two landscapes with contrasting spatial and temporal dynamics. Their results demonstrate that the performance of the data fusion methods is strongly associated with land-cover spatial and temporal variance. ESTARFM is better than STARFM in heterogeneous landscapes but it is even worse than STARFM for predicting abrupt changes in land-cover types. The unmixing based fusion methods, such as U-STFM, also require that no land-cover type change occurs between the input and prediction dates. They are computationally efficient but still face the same challenges or bear shortcomings in regions with abrupt land-cover type changes. The most recently developed methods FSDAF, STI-FM and HCM focus on forward prediction by using one pair of Landsat and MODIS images as input data and simplify algorithms for better computation efficiency but they did not explore benefit of combining forward and backward predictions from two pairs of Landsat and MODIS images through a proper smoothing filter. RWSTFM uses a Kriging estimator to blend Landsat and MODIS images. One of its main advantages is that it is able to present a variance estimation for the blended image as a quality indicator of the Kriging estimator, which is not available from all of other methods mentioned above. However, its computation efficiency is low. This limits its practical applications for fusion exercise of time series of images over a large coverage.

In order to better model both gradual vegetation phenological dynamics and abrupt land-cover type changes, we propose a new Prediction Smooth Reflectance Fusion Model (PSRFM) that takes advantages of the unmixing based fusion methods in computation efficiency and overcomes shortcomings in poor performance over regions with abrupt land-cover type changes and lack of uncertainty estimation through a model extension and optimization and introduction of uncertainty estimation based on the Least-Squares adjustment [12,13]. PSRFM combines a dynamic prediction model based on the linear spectral mixing algorithm [4–6] and a smoothing filter based on a weighted average using estimated uncertainties together with constraints on the Normalized Difference Snow Index (NDSI) [14,15] or Normalized Difference Vegetation Index (NDVI) [16,17]. Following the introduction, in Section 2, we describe PSRFM methodologies. In Section 3, we explain the experiment to test the model by generating a Landsat-like image time series with Landsat-8 OLI and MODIS

(MOD09GA) images. In Section 4, we present the experiment results and discuss our findings. At the end, in Section 5, we present the conclusions of the study.

2. Methodology

2.1. Problem Description

Assume available co-incident clear-sky Landsat and MODIS surface reflectance data on a start date t_0, followed by no clear-sky Landsat data for a period due to obscured land surface conditions such as cloud contamination or the limitation of the Landsat system revisiting capability [18] but some MODIS clear-sky observations. The first question to be addressed is how well we can predict Landsat-like surface reflectance at a later date t_1 based on the pair of Landsat and MODIS images at date t_0 and an additional MODIS observation at date t_1. The second question of the problem is how the prediction can be improved once a second pair of clear-sky Landsat and MODIS surface reflectance data become available on a later date t_2.

2.2. Modelling of Gradual Vegetation Phenological Changes

As mentioned in the introduction, a good fusion model should cover both gradual vegetation phenological dynamics and abrupt land-cover type changes. To model both of them for predicting Landsat-like image from MODIS observations, we can take the Landsat reflectance image R_0 at the start date t_0 as an initial estimate with a variance (i.e., its observation uncertainty) and then add two additional terms that represent gradual vegetation phenological change ΔR_{01} and abrupt land-cover type change Δ_{01} from the start date t_0 to a prediction date t_1, respectively. That is, we can predict a Landsat-like image R_1 at t_1 as

$$R_1 = R_0 + \Delta R_{01} + \Delta_{01}, \tag{1}$$

where the gradual vegetation phenological changes ΔR_{01} and the abrupt land-cover type change Δ_{01} are to be determined from MODIS observations at t_0 and t_1.

Gradual vegetation phenological changes are modelled as a linear surface reflectance change velocity. Abrupt land-cover type changes are modelled as a correction based on the analysis of spatial residual analysis and spatial interpolation as in [6]. A challenge here is that we cannot measure the reflectance change velocity for each Landsat pixel because the only new measurements about them are limited to MODIS observations at a much lower resolution. To address this issue, we assume that the reflectance change rate is the same for a classified region or cluster of Landsat pixels. If there are no land-cover type changes within the cluster, we assume the reflectance change velocity of the cluster will reflect reasonably well the phenological changes with time.

Assuming that the land-cover type change such as snowmelt, flooded and forest fire area at a Landsat-like image pixel is δ and the reflectance change velocity for a cluster c is $\dot{r}$, then we can predict reflectance r of band b at the pixel (x, y) belonging to cluster c at time t_1 as follows:

$$r_1(x, y, c, b) = r_0(x, y, c, b) + (t_1 - t_0)\dot{r}(x, y, \ c, b) + \delta(x, y, \ c, b). \tag{2}$$

where $r_0(x, y, \ c, b)$ is Landsat surface reflectance observation of band b at the pixel (x, y) belonging to cluster c at time t_0.

To simplify the mathematical expression, we omit the pixel's location (x, y) and the band number b and sort Landsat and Landsat-like surface reflectance into k clusters for band b as

$$
\begin{bmatrix} r_1 \\ {\scriptstyle n_1 \times 1} \\ \vdots \\ r_c \\ {\scriptstyle n_c \times 1} \\ \vdots \\ r_k \\ {\scriptstyle n_k \times 1} \end{bmatrix}_1 = \begin{bmatrix} r_{1,} \\ {\scriptstyle n_1 \times 1} \\ \vdots \\ r_c \\ {\scriptstyle n_c \times 1} \\ \vdots \\ r_k \\ {\scriptstyle n_k \times 1} \end{bmatrix}_0 + \begin{bmatrix} \Delta t & \cdots & 0 \\ {\scriptstyle n_1 \times 1} & & {\scriptstyle n_1 \times 1} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Delta t \\ {\scriptstyle n_k \times 1} & & {\scriptstyle n_k \times 1} \end{bmatrix} \begin{bmatrix} \dot{r}_1 \\ \vdots \\ \dot{r}_c \\ \vdots \\ \dot{r}_k \end{bmatrix}_0 + \begin{bmatrix} \delta_1 \\ {\scriptstyle n_1 \times 1} \\ \vdots \\ \delta_c \\ {\scriptstyle n_c \times 1} \\ \vdots \\ \delta_k \\ {\scriptstyle n_k \times 1} \end{bmatrix}_1 \tag{3}
$$

where r_c is an $n_c \times 1$ vector that includes all Landsat pixel values belonging to a cluster c ($c = 1, 2, \ldots$, k). The total number of Landsat pixels of the processed image is $n = \sum_1^k n_c$. $\dot{r}_c$ and δ_c are reflectance change velocity and land-cover change correction for cluster c, respectively. $\Delta t = t_1 - t_0$ is the time interval between the start and prediction dates.

Since we have MODIS observations on both dates t_0 and t_1, we can calculate the reflectance change velocity and its uncertainty for each of the MODIS reflectance pixels as follows

$$
\dot{M} = (M_1 - M_0)/(t_1 - t_0), \sigma_{\dot{M}}^2 = \frac{2\sigma_M^2}{\Delta t^2} \tag{4}
$$

where M_0 and M_1 are MODIS observations on date t_0 and t_1, respectively. $\sigma_{\dot{M}}^2$ is the uncertainty of the velocity estimate calculated from the MODIS observation uncertainty σ_M^2 through the error propagation law.

According to spectral linear mixing theory, the reflectance change of a MODIS pixel is weighted sum of the Landsat reflectance change $\Delta R_c(x_i, y_i)$ of all clusters within the MODIS pixel:

$$
\Delta M(x_i, y_i) = \sum_{c=1}^k f_c(x_i, y_i) \times \Delta R_c(x_i, y_i) \tag{5}
$$

where $f_c(x_i, y_i)$ is the portion of the Landsat pixels belonging to cluster c within a MODIS pixel. It can be determined by Equation (6):

$$
f_c(x_i, y_{i,}) = N_c(x_i, y_i)/p_i \tag{6}
$$

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). p_i is the total number of Landsat pixels within the MODIS pixel i.

If we divide the two sides of Equation (5) by the time difference, it becomes the equation for determining the reflectance change velocity; that is

$$
\dot{m}(x_i, y_{i,}) = \sum_{c=1}^k f_c(x_i, y_i) \times \dot{r}_c \tag{7}
$$

where $\dot{m}(x_i, y_i)$ is a velocity estimation at MODIS pixel i. Assuming that the reflectance change of the same cluster among all MODIS pixels is the same, we can apply all p ($p > k$) MODIS pixels to compose a system of linear mixture equations:

$$
\begin{bmatrix} \dot{m}(x_1, y_1) \\ \vdots \\ \dot{m}(x_i, y_i) \\ \vdots \\ \dot{m}(x_p, y_p) \end{bmatrix} = \begin{bmatrix} f_1(x_1, y_1) & f_2(x_1, y_1) & \cdots & f_k(x_1, y_1) \\ \vdots & \vdots & & \vdots \\ f_1(x_i, y_i) & f_2(x_i, y_i) & \cdots & f_k(x_i, y_i) \\ \vdots & \vdots & & \vdots \\ f_1(x_p, y_p) & f_2(x_p, y_p) & \cdots & f_k(x_p, y_p) \end{bmatrix} \begin{bmatrix} \dot{r}_1 \\ \vdots \\ \dot{r}_c \\ \vdots \\ \dot{r}_k \end{bmatrix} \tag{8}
$$

From Equation (8), we can obtain an estimation of the reflectance change velocity $\dot{r}$ and its covariance matrix $C_{\dot{r}\dot{r}}$ through a typical Least-Squares adjustment [12,13] as follows

$$\hat{\dot{r}} = \left(A^T P A\right)^{-1} A^T P l, \ C_{\dot{r}\dot{r}} = \hat{\sigma}_{\dot{r}}^2 Q_{\dot{r}\dot{r}} = \hat{\sigma}_{\dot{r}}^2 \left(A^T P A\right)^{-1} \tag{9}$$

where

$$\hat{\dot{r}} = \begin{bmatrix} \hat{\dot{r}}_1 \\ \vdots \\ \hat{\dot{r}}_c \\ \vdots \\ \hat{\dot{r}}_l \end{bmatrix}, A = \begin{bmatrix} f_1(x_1,y_1) & f_2(x_1,y_1) & \cdots & f_k(x_1,y_1) \\ \vdots & \vdots & & \vdots \\ f_1(x_i,y_i) & f_2(x_i,y_i) & \cdots & f_k(x_i,y_i) \\ \vdots & \vdots & & \vdots \\ f_1(x_p,y_p) & f_2(x_p,y_p) & \cdots & f_k(x_p,y_p) \end{bmatrix}, l = \begin{bmatrix} \dot{m}(x_1,y_1) \\ \vdots \\ \dot{m}(x_i,y_i) \\ \vdots \\ \dot{m}(x_p,y_p) \end{bmatrix} \tag{10}$$

$P = C_{ll}^{-1}$ and $C_{ll} = \sigma_M^2 I$ is the variance matrix of the velocity measurements l. $\hat{\sigma}_{\dot{r}}^2$ is the unit variance estimation from the residuals v of the observation Equation (8):

$$\hat{\sigma}_{\dot{r}}^2 = \frac{v^T P v}{p - k} \tag{11}$$

where

$$v = l - A\hat{\dot{r}}. \tag{12}$$

Equations (3), (4), (9) and (10) are formulas required to predict gradual vegetation phenological change ΔR_{01}. According to the error propagation law, we can also estimate their corresponding uncertainties for each pixel belonging to cluster c as follows

$$\hat{\sigma}_{\Delta r_c}^2 = (t_1 - t_0)^2 \hat{\sigma}_{\dot{r}}^2 Q_{\dot{r}_c \dot{r}_c} \tag{13}$$

where $Q_{\dot{r}_c \dot{r}_c}$ is the covariance factor corresponding to the estimated surface reflectance change velocity $\hat{\dot{r}}_c$.

If Landsat observation uncertainty is σ_R^2, then the uncertainty of the predicted Landsat-like image pixels of cluster c with only gradual vegetation phenological changes can be calculated as

$$\hat{\sigma}_{r_c}^2 = \sigma_R^2 + \hat{\sigma}_{\Delta r_c}^2 = \sigma_R^2 + (t_1 - t_0)^2 \hat{\sigma}_{\dot{r}}^2 Q_{\dot{r}_c \dot{r}_c} \tag{14}$$

2.3. Modelling of Land-Cover Type Changes

The modelling of gradual vegetation phenological changes in Section 2.3 is based on the spectral linear mixing theory. Its accuracy depends on the classification of the Landsat image at t_0. This means that it is most effective when no land-cover type change occurs between the start date t_0 and the prediction date t_1. In reality, this may not be true. For example, a sudden flooding or forest fire may take place during a short period. Snowmelt may also play an important role on land surface dynamics in cold northern regions during the transition period from winter to spring season. To consider such land-cover changes, we need to further model the land-cover change item δ in Equation (2).

Since land-cover type changes will affect the estimation of the velocity parameter vector $\hat{\dot{r}}$ and the residual v in Equation (11) directly, we can model land-cover type changes based on analysis of the residuals. If we assume that land-cover type change happens in a relatively small portion of the image, then the residuals would be outliers at those MODIS pixels that contain land-cover type changes. If there is no land-cover change at all, the residuals should follow a normal distribution $N(0, C_{vv})$ and a correction for it is not necessary.

For individual MODIS pixel, the residual can be calculated from [6]

$$\delta_M(x_i, y_i, b) = \Delta M(x_i, y_i, b) - \frac{1}{p_i}\left[\sum_{j=1}^{p_i}\hat{r}_{t1}(x_{ij}, y_{ij}, b) - \sum_{j=1}^{p_i}r_{t0}(x_{ij}, y_{ij}, b)\right] \tag{15}$$

where $\delta_M(x_i, y_i, b)$ is the residual pixel value of band b at pixel (x_i, y_i), $\Delta M(x_i, y_i, b)$ is the MODIS observation differences between dates t_0 and t_1; $\hat{r}_{t1}(x_{ij}, y_{ij}, b)$ is the predicted Landsat-like pixel value on the prediction date t_1, $r_{t0}(x_{ij}, y_{ij}, b)$ is the Landsat pixel value at the start date t_0 and p_i is the number of Landsat pixels within a MODIS pixel i.

Equation (15) indicates that the residual of a MODIS pixel represents a bias of the estimated MODIS reflectance difference based on the predicted Landsat-like image in comparison to the observed MODIS reflectance at pixel (x_i, y_i). If the predicted temporal change reflects the gradual vegetation phenological change of a cluster of Landsat pixels well, the corresponding residuals should follow a zero mean normal distribution. Otherwise, land-cover type changes may have taken place at some of the MODIS pixels within the same cluster. Based on this fact, we may use one of the following strategies to deal with land-cover changes with the aim of improving the prediction result:

1. Use the residuals to identify spatial pattern changes within a cluster, then reclassify the pixels within the cluster based on the residual pattern for a new prediction. For instance, if some of the pixels are spatially located together within a cluster have different residuals from others, we may split the pixels of this cluster into two new clusters based on the residual values. One is for the pixels without land-cover change and another is for the pixels with land-cover change, or

2. Distribute the residuals to Landsat pixels within the MODIS pixel through an adjustment of the residuals similar to the TPS interpolation used by FSDAF.

PSRFM treats this residual adjustment as a model correction for land-cover type change as the residuals mainly come from land-cover type change and within-cluster variability [6]. As mentioned above, a residual adjustment is necessary only when significant land-cover changes are detected.

Further, to make the clusters more accurately reflect land-cover changes between t_0 and t_1, we may also perform clustering on MODIS images at t_0 and t_1 and thereafter overlay the two cluster maps to check the agreement between the two cluster areas. For the agreement regions, it is not necessary to add a correction while for the areas that do not agree some kind of corrections may be necessary for abrupt land-cover changes.

2.4. Optimization of Models

From the formulation of the prediction model in Sections 2.2 and 2.3, we can see that there are two issues that may affect the accuracy of the prediction model significantly. One is the number of clusters of the Landsat image used for predicting gradual vegetation phenological changes. Another is the residual adjustment method for land-cover type changes that depends on the number of clusters used for prediction. Therefore, optimization of the models, for example, determining the clusters optimally, becomes a key step of the prediction. This issue exists with all image fusion models based on the linear spectral mixing theory, for example U-STFM and FSDAF.

To determine the clusters optimally, we suggest a rule of Maximizing Correlation but Smaller Residuals (MCSR). The basic idea of this rule is to select a cluster number from a set of pre-defined cluster numbers that will maximize the correlation between the differences $(\Delta R = \hat{R}_1 - R_0)$ of the Landsat images and the differences $(\Delta M = M_1 - M_0)$ of the MODIS images but will not increase the sum of the residual squares $(sum = \sum \delta_M^2)$ of the predicted Landsat-like image significantly. We select the differences for the comparison because they may cancel same systematic biases of the sensors at different observation times. Assuming that we have already a well-developed unsupervised automatic classification method available, for example, the ISODATA classification method implemented in Geomatica 2016 [http://www.pcigeomatics.com/], we can optimize our models in two steps:

1. Conduct predictions for a predetermined cluster number range from k_{min} to k_{max} and calculate their corresponding correlation coefficients between the differences of the Landsat image at t_0 and the predicted Landsat-like image at t_1 and the differences of the MODIS images at t_0 and t_1 and sums of the residual squares for all predicted Landsat-like images, respectively;

2. Compare the correlation coefficients and the sums of residual squares and select the prediction that meets the MCSR rule as the optimized prediction.

2.5. Smoothing of Forward and Backward Predictions

The prediction based on our model formulation can be performed both forward and backward if another pair of Landsat and MODIS surface reflectance data at time t_2 ($t_2 > t_1 > t_0$) are available. In such a case, we can combine the forward and the backward predictions for the final Landsat-like predictions on date t_1.

One combination method is to smooth the forward and backward predictions through a weighted average based on their estimated uncertainties. Another option is to average the forward and backward predictions by weights derived from the elapsed time between the observation dates of the input images and the predicted date. As phenology change is a function of time, in general, the closer the dates of reflectance images, the more similar they are. Therefore, we may use the elapsed time for calculation of weights for smoothing. A risk with weighting by elapsed time is that quality of the predicted image with the shorter elapsed time may be much worse than quality of the predicted image with longer elapsed time. Therefore, if we can estimate the uncertainties properly, weighting by the uncertainties is favorable over weighting by the elapsed time.

Both weighting methods considered cannot cover situations when significant land-cover type changes occurred between the start date t_0 and the end date t_2. For instance, snow onset, snowmelt and flooding may happen at any time during the period between t_0 and t_2. To consider such land-cover changes, we recommend applying Normalized Difference Snow Index (NDSI) [14,15] or Normalized Difference Vegetation Index (NDVI) [16,17] as a constraint for the combination of the forward and backward predictions for certain situations. Recommendation of applying NDSI for the smoothing process is based on the observations that the two pairs of Landsat images used for blending are acquired at least 16 days apart ($t_2 - t_0 \geq 16$). In a transition period from snow cover to snow free, the two images may have the same or different snow cover conditions than those of the MODIS image at date t_1. It can be expected that the predicted image (either forward or backward) would have a higher accuracy if the images used for the blending have the same snow cover status and vice versa. NDVI represents the general conditions of vegetation coverage. For a region with only one vegetation growing season (e.g., over natural areas in temperate, boreal and arctic zones), yearly time-series of NDVI of vegetation without disturbance has a bell shape with one maximum [19]. When a forest fire or flooding on vegetated land occurs, the affected area would have a much lower NDVI. In the situation of flooding, negative NDVI value can be observed. Similar to NDSI, using NDVI as a constraint to the smoother would help identify some types of land-cover changes such as forest fire or flooding and improve the accuracy of the smoothed images.

This process consists of the following steps:
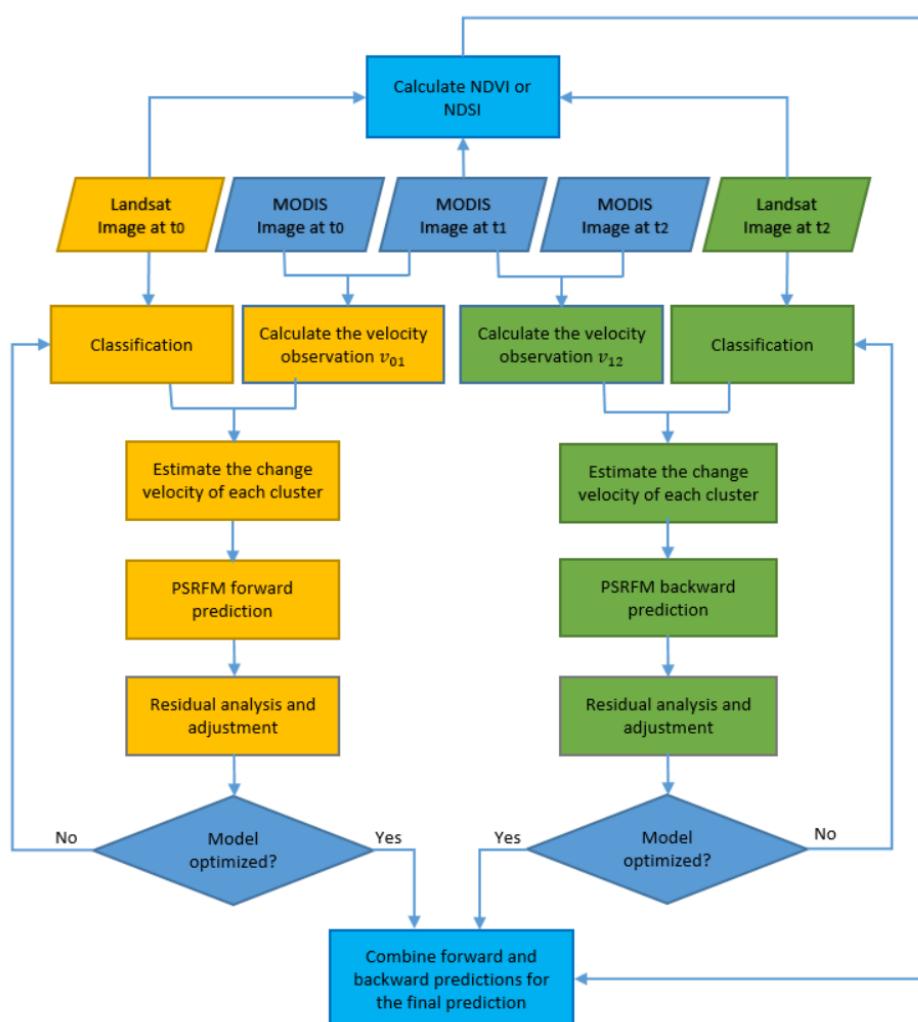
1. Calculate indices (NDSI for snow onset and snowmelt season or NDVI for vegetation growing season) of the input Landsat images at t_0 and t_2 and the MODIS observation at t_1, namely I_{fwd}, I_{bwd} and I_{mod} for forward prediction, backward prediction and the MODIS, respectively. NDSI and NDVI are used separately. When there is snow cover present on at least one image, NDSI is used. NDVI is used for those images in vegetation growing season.

2. Check if there are any invalid pixel values in the forward and backward predictions. If one of them is invalid, select the valid pixel as the final pixel;

3. Set up an index boundary I_0. For example, the boundary I_0 is set to 0.4 for snow cover ($\geq$0.4) and snow free (<0.4) situation as it is the commonly used threshold. And then compare the boundary value to the index value I_{fwd}, I_{bwd} and I_{mod} to determine the final pixel as follows:

a.   If $I_{fwd} \geq I_0$ and $I_{mod} \geq I_0$ but $I_{bwd} < I_0$, select the forward prediction;

b.   If $I_{fwd} \leq I_0$ and $I_{mod} \leq I_0$ but $I_{bwd} > I_0$, select the forward prediction;

c.   If $I_{fwd} < I_0$ but $I_{mod} \geq I_0$ and $I_{bwd} \geq I_0$, select the backward prediction;

d.   If $I_{fwd} > I_0$ but $I_{mod} \leq I_0$ and $I_{bwd} \leq I_0$, select the backward prediction;

e.   For all other cases, apply weighted average based on the estimated uncertainties or their time intervals between the observation dates and the prediction date.

In this study, we choose a study period from winter to spring and have incorporated NDSI in the modeling exercise to demonstrate its effectiveness in dealing with situation when land surface is transition from snow cover to snow free. However, the use of NDSI or NDVI may not be applicable to some applications. For instance, in urban areas, it may be less effective to apply NDSI or NDVI for change detection. In such cases, the methods described in Section 2.3 for change detection are more general and effective.

2.6. Implementation

Based on the formulation above, we implemented the PSRFM algorithm to predict Landsat-like image on date t_1 forward and backward from the input data (Figure 1).


Figure 1. Processing flowchart of Prediction Smooth Reflectance Fusion Model (PSRFM).

Before we apply PSRFM, 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 bilinear or nearest neighbor interpolation 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 [20,21]. To further reduce the difference between coarse and fine-resolution data caused by sensor configuration and processing chains, a radiometric normalization can be applied by assuming a linear relationship between both data sets [21,22]. After we preprocessed both the Landsat and MODIS images to the same physical quantity, we can process the data by PSRFM in the following steps:

1. Perform forward predictions for predetermined cluster number range from 4 to 16 and determine the optimized forward prediction. This process includes:

    a. classify the Landsat image at t_0 for a specified cluster number;
    b. calculate the MODIS reflectance change velocity using the MODIS observations at t_0 and t_1;
    c. estimate the reflectance change velocity;
    d. predict Landsat-like image at t_1 using the reflectance change velocity of all clusters estimated in step c;
    e. calculate the correlation coefficient between the differences of the predicted Landsat-like image and Landsat image at t_0 and the MODIS image differences;
    f. calculate residuals of the predicted Landsat-like image and their sum of squares;
    g. adjust residuals by interpolating them to each Landsat-like image pixels to obtain a new Landsat-like image prediction;
    h. calculate the correlation coefficient between the differences of the newly predicted Landsat-like image and Landsat image at t_0 and the MODIS image differences;
    i. compare the two correlation coefficients and check the change of the sum of residual squares and select the prediction with bigger correlation coefficient and the sum change of residual squares less than 5% if it increases.
    j. repeat steps from a to i for all cluster numbers in the range from 4 to 16;
    k. compare all predictions and choose the prediction with the maximal correlation coefficient as the final optimal prediction.

2. Perform the same processing as step 1 for backward prediction.
3. Compute NDSI or NDVI of Landsat images acquired on dates t_1 and t_2 and of MODIS image observed on prediction date t_1, then combine the optimized forward and backward predictions as the final prediction of the Landsat-like image.

3. Validations and Experiments

To verify effectiveness of the proposed PSRFM method and compare it to some published methods such as STARFM and ESTARFM, validation experiments were carried out to generate a Landsat-like image time series between start date t_0 and end date t_2 of two pairs of Landsat and MODIS images and a time series of MODIS images within the two dates. The predicted Landsat-like image time series are evaluated against their observed Landsat images (as reference) and compared to images predicted by STARFM and ESTARFM both visually and quantitatively. Additional to the visual and quantitative evaluations, the estimated uncertainties of the Landsat-like images by proposed PSRFM and their variations with different priori uncertainties used for the Landsat and MODIS input images are also analyzed and discussed.

3.1. Quality Indices for Validation and Comparison

To quantify the closeness of the predicted Landsat-like images to the actual Landsat images, five commonly used quality indices are used for the validation.

Let x be the predicted Landsat-like image, y be the observed Landsat image as reference with n pixels, $\bar{x}$ and $\bar{y}$ be their mean values, σ_x^2, σ_y^2 and σ_{xy} be their variances and covariance, then we can evaluate the quality of x against y by using the following 5 commonly-used quality indices [23,24]:

1.　Average absolute difference (AAD)

$$AAD = \frac{1}{n} \sum_{i=1}^{n} (|x_i - y_i|) \tag{16}$$

2.　The root mean squared error (RMSE)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - y_i)^2} \tag{17}$$

3.　Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS)

$$ERGAS = 100 \frac{h}{l} \sqrt{\frac{1}{n} \sum_{k=1}^{N} \frac{RMSE(B_k)^2}{\bar{x}_k^2}} \tag{18}$$

4.　Correlation coefficient (CC)

$$CC = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \tag{19}$$
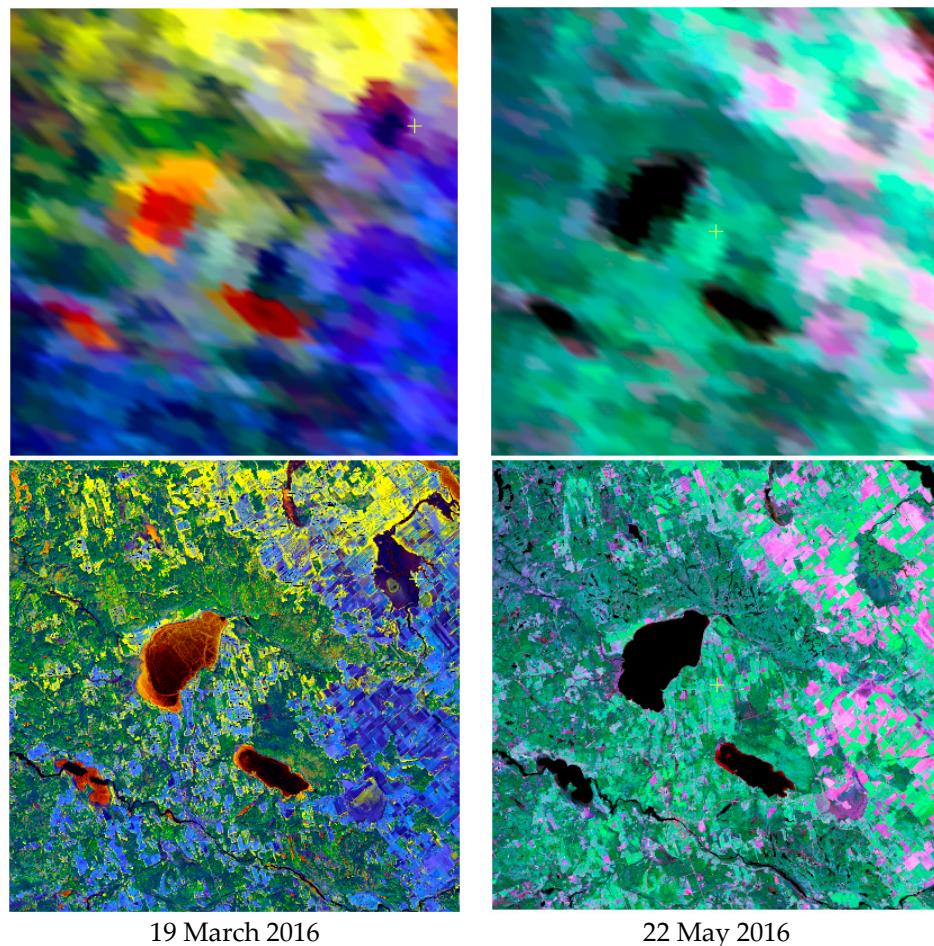
5.　The quality index (QI) [25]

$$QI = \frac{4\sigma_{xy}\overline{xy}}{\left(\sigma_x^2 + \sigma_y^2\right)\left(\bar{x}^2 + \bar{y}^2\right)} \tag{20}$$

Among the quality indices, the AAD is used to measure the prediction bias. The RMSE is a measure of total uncertainty (bias plus precision errors) between the predicted and reference images. The ERGAS evaluates the quality of a fused image in terms of normalized error compared to the reference image. A higher value of ERGAS indicates a bigger distortion in the predicted image and a lower value indicates that the predicted image is similar to the reference image [26]. The CC reflects the degree of correlation (i.e., similarity) between the predicted and reference images. 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 more structural similarity and vice versa.

3.2. Study Area and Dataset

The study area is shown in Figure 2. It is located around 45° N and 76° W (Ontario, Canada) and within the Landsat-8 OLI scene of 16/28 (path/row). The area is dominated by agricultural and forest land-covers. A portion of the Landsat scene (800 by 800 pixels) is used for the model evaluation. In this study, five pairs of clear-sky Landsat-8 OLI and MODIS (MOD09GA) surface reflectance images are acquired from 19 March 2016 to 22 May 2016 (Table 1) with a time interval of Landsat revisiting cycle of 16 days. One of the reasons to choose this period is that beside two pairs of images are used for modelling, there are still three clear-sky Landsat images available for model evaluation as shown in Table 1.

19 March 2016   22 May 2016

Figure 2. Red-NIR-SWIR1 (R-G-B) false color composites of the two pairs of Landsat-8 OLI surface reflectance (**lower row**) and MODIS surface reflectance bands 1-2-6 false color composites (**upper row**) used for a prediction of Landsat-like image time series between 19 March 2016 and 22 May 2016. The yellow and reddish colors on the images of the left show snow and ice coverage.

Table 1. Images, acquisition dates and their function.

Image Acquisition Date	Landsat-8 OLI	MODIS (MOD09GA)
19 March 2016	Model input (start date)	Model input (start date)
4 April 2016	For model validation	Model input (for prediction)
20 April 2016	For model validation	Model input (for prediction)
6 May 2016	For model validation	Model input (for prediction)
22 May 2016	Model input (end date)	Model input (end date)

The first (19 March 2016) and the last (22 May 2016) pairs of Landsat and MODIS images were used for the start date (t_0) and the end date (t_2) of the model, respectively. The time interval between the two dates was 63 days. For the demonstration and evaluation, a time series of Landsat-like images were predicted only at three dates (4 April, 20 April and 6 May 2016) within the period at when actual clear-sky Landsat images are available; therefore, the predicted images could be evaluated against the observed Landsat images. As these images were temporally distributed every 16 days within the study period, the predictions and their evaluation against the true Landsat images can reveal the model's performance using an evenly temporal sampling. Another reason to choose these images is that the study period covers the transition from the winter to spring, when snowmelt is one of the main causes of land surface changes. Significant snow and ice coverage is visible on the images of

19 March 2016 while it is not present on 22 May 2016 imagery (Figure 2). Hence, the PSRFM algorithms can be fully evaluated.

The MODIS Reprojection Tool (https://lpdaac.usgs.gov/tools/modis_reprojection_tool) is used for reprojection and resample of MODIS images to the map projection and resolution of the Landsat images. Three bands of 4 (Red), 5 (NIR—Near Infrared) and 6 (SWIR 1—Shortwave Infrared 1) of Landsat 8 OLI images, whose corresponding bands of MODIS bands are 1, 2 and 6, are used in the modeling exercise

4. Results and Discussion
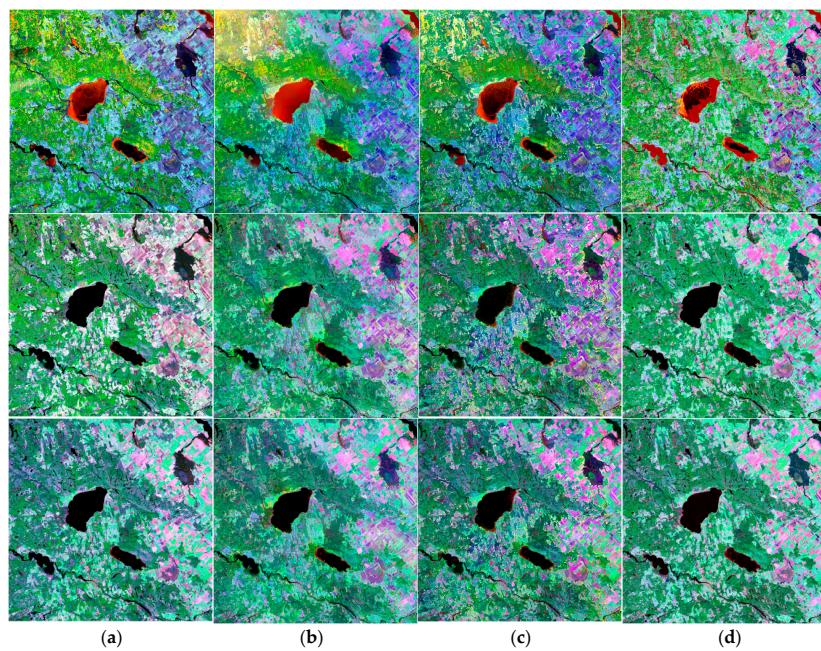
We evaluated the performance of the proposed PSRFM against the Landsat images in the test dataset and subsequently compared our results, visually and quantitatively using the five (5) quality indices listed in Section 3.1, to two published and well-studied methods: STARFM [1] and ESTARFM [2]. We chose these two methods for comparison to our model mainly due to their wide acceptance within the remote sensing community and especially the availability of their processing software. For the comparison results, we applied STARFM with a max search distance of 25 pixels and other default values set by the program (available from https://www.ars.usda.gov/research/software/download/?softwareid=432#downloadForm) and ESTARFM with a half-window size of 25 pixels and other default values set by the program (available from https://xiaolinzhu.weebly.com/open-source-code.html). With PSRFM, we used unsupervised ISODATA classification method and its default parameters implemented in Geomatica 2016 except that the cluster number was determined by the model optimization algorithm (Section 2.4). The priori uncertainty used for all PSRFM processing was $\sigma_R = 40$ for Landsat images and $\sigma_M = 10$ for MODIS images (with reflectance scaled by 10,000), respectively, as suggested in Section 4.3. For combination of the forward and backward prediction, the NDSI constraint with an index boundary $I_0 = 0.4$ was applied.

4.1. Visual Comparison of Predicted Landsat-Like Image Time Series against True Landsat Images
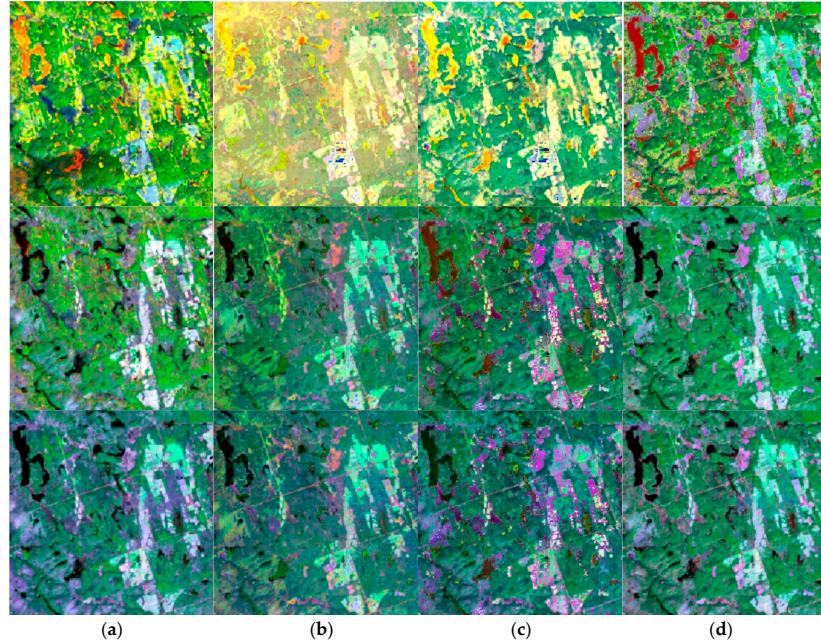
As stated earlier, we predicted Landsat-like images at three dates (4 April, 20 April and 6 May 2016) when a true Landsat image is available by PSRFM, ESTARFM and ESTARFM for PSRFM model validation and comparison among these three models. In reality, clear-sky Landsat images are available on consecutive 5 revisiting days is rare. Figure 3 compares the observed Landsat images and the predicted Landsat-like image time series by the three different methods.

From visual comparison of the images predicted by the three methods we can see that all of the methods are able to capture the general land-cover changes in the test area during this observation period. However, compared to the observed Landsat images, the variations from model to model are clearly visible, for instance, in the top-left area (see their enlarged image comparison of the first 200 × 200 pixels in Figure 4). In contrast to STARFM and ESTARFM, PSRFM retains more details and clear boundaries for small objects like water bodies. Therefore, the images predicted by PSRFM have a higher fidelity to the true Landsat-8 OLI images than those predicted by STARFM and ESTARFM, especially for the prediction dates after snow is gone due to NDSI involvement by PSRFM.
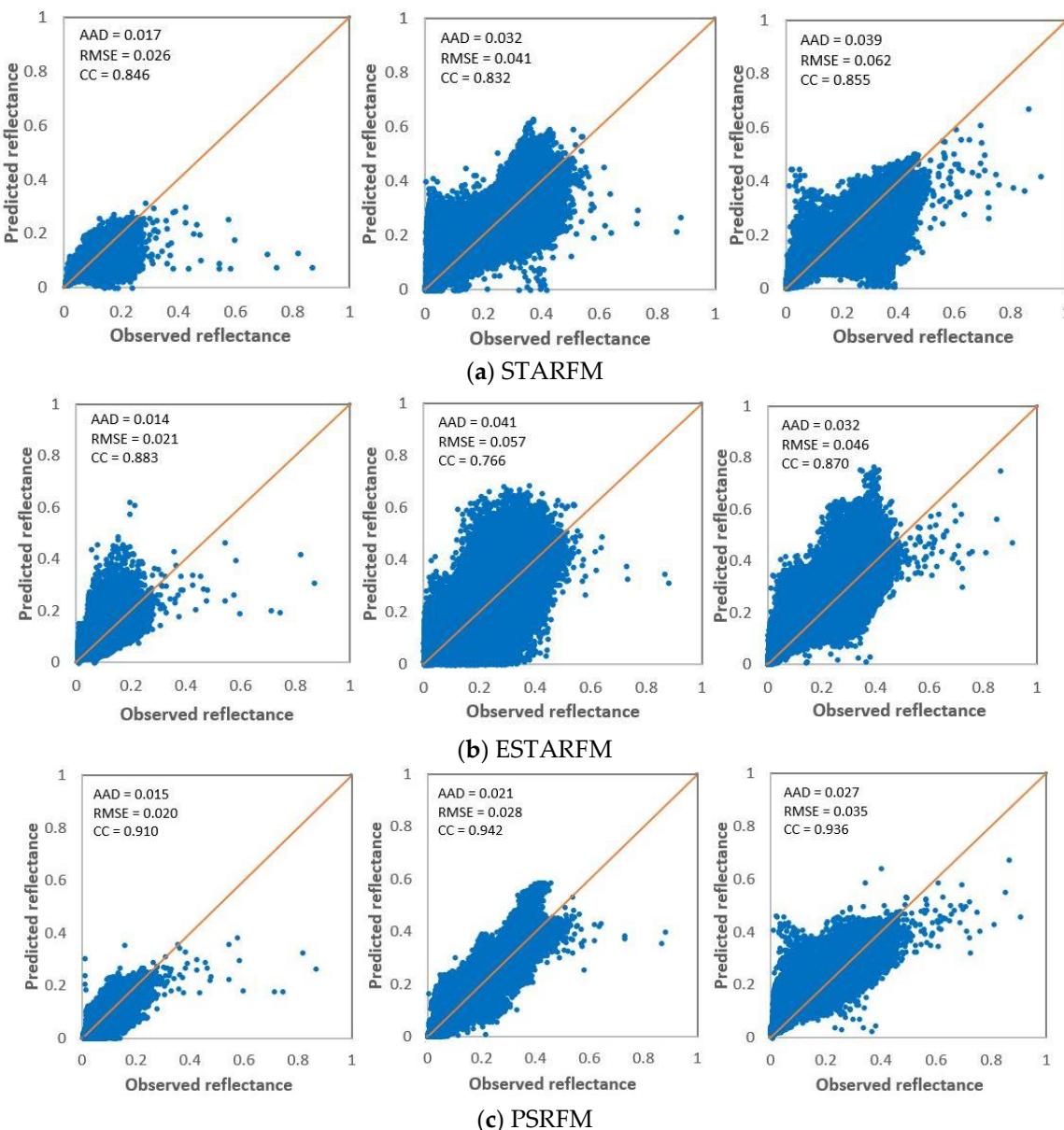
To further confirm the performance of PSRFM, we compared the surface reflectance of the predicted Landsat-like image time series against that of the true Landsat images for the prediction date 6 May 2016 using scatter plots (Figure 5). The scatter plots indicate that PSRFM results in better agreement with true Landsat reflectance than STARFM or ESTARFM.

Figure 3. Visual comparison of Landsat-like image series predicted by STARFM, ESTARFM and the proposed PSRFM methods. From the left to the right, Red-NIR-SWIR1 (R-G-B) composites of (**a**) true Landssat-8 OLI images; (**b**) predicted images from STARFM; (**c**) predicted images from ESTARFM; and (**d**) predicted images from PSRFM. From the top to the bottom, the dates of the image time series are 4 April 2016, 20 April 2016 and 6 May 2016.


Figure 4. Enlarged visual comparison of the first 200×200 pixels of the Landsat-like image series predicted by STARFM, ESTARFM and the proposed PSRFM method. From the left to the right, Red-NIR-SWIR1 (R-G-B) composites of (**a**) the true Landssat-8 OLI images; (**b**) the predicted images from STARFM; (**c**) the predicted images from ESTARFM; and (**d**) the predicted images from PSRFM. From the top to the bottom, image time series for dates 4 April, 20 April and 6 May 2016.

(**a**) STARFM

(**b**) ESTARFM

(**c**) PSRFM

Figure 5. Scatter plots of predicted reflectance against observed reflectance using (**a**) STARFM, (**b**) ESTARFM and (**c**) PSRFM for prediction date 6 May 2016.

4.2. Quantitative Comparison of the Quality Indices

For quantitative comparisons of the predicted Landsat-like image time series, we calculated the five quality indices (Equations (16)–(20)) for all predicted Landsat-like images dated as 4 April 2016 (day of year 95), 20 April 2016 (day of year 111) and 6 May 2016 (day of year 127) against the corresponding observed Landsat images, respectively. Table 2 shows three error estimates (AAD, REMS and ERGAS) of the predicted Landsat-like image time series by the three models (STARFM, ESTARFM and PSRFM). Table 3 gives their corresponding correlation coefficients (CC) and the quality index (QI). Figure 6 compares all of the five quality indices band by band graphically.

Table 2. Quality indices AAD, REMS and ERGAS of the predicted Landsat-like image time series by the three models STARFM, ESTARFM and PSRFM.

Band	Method	AAD			RMSE			ERGAS		
Date (Day of Year)		**95**	**111**	**127**	**95**	**111**	**127**	**95**	**111**	**127**
RED	STARFM	0.040	0.023	0.017	0.060	0.031	0.026	3.397	2.688	2.529
	ESTARFM	0.039	0.023	**0.014**	0.062	0.035	0.021	**3.162**	2.656	1.824
	PSRFM	**0.037**	**0.019**	0.015	**0.059**	**0.025**	**0.020**	3.361	**2.077**	**1.792**
NIR	STARFM	0.034	0.031	0.032	**0.046**	0.041	0.045	**1.286**	1.286	1.277
	ESTARFM	0.043	0.047	0.041	0.062	0.065	0.057	1.705	2.049	1.622
	PSRFM	**0.034**	**0.023**	**0.021**	0.049	**0.031**	**0.028**	1.339	**0.906**	**0.743**
SWIR1	STARFM	0.055	0.047	0.039	0.070	0.062	0.055	2.557	2.051	1.805
	ESTARFM	0.057	0.050	0.032	0.074	0.069	0.046	2.406	1.922	1.257
	PSRFM	**0.049**	**0.031**	**0.027**	**0.064**	**0.043**	**0.035**	**2.095**	**1.320**	**1.081**
AVERAGE	STARFM	0.043	0.033	0.029	0.059	0.045	0.042	2.413	2.008	1.870
	ESTARFM	0.046	0.040	0.029	0.066	0.056	0.042	2.425	2.209	1.568
	PSRFM	**0.040**	**0.024**	**0.021**	**0.057**	**0.033**	**0.027**	**2.265**	**1.434**	**1.205**
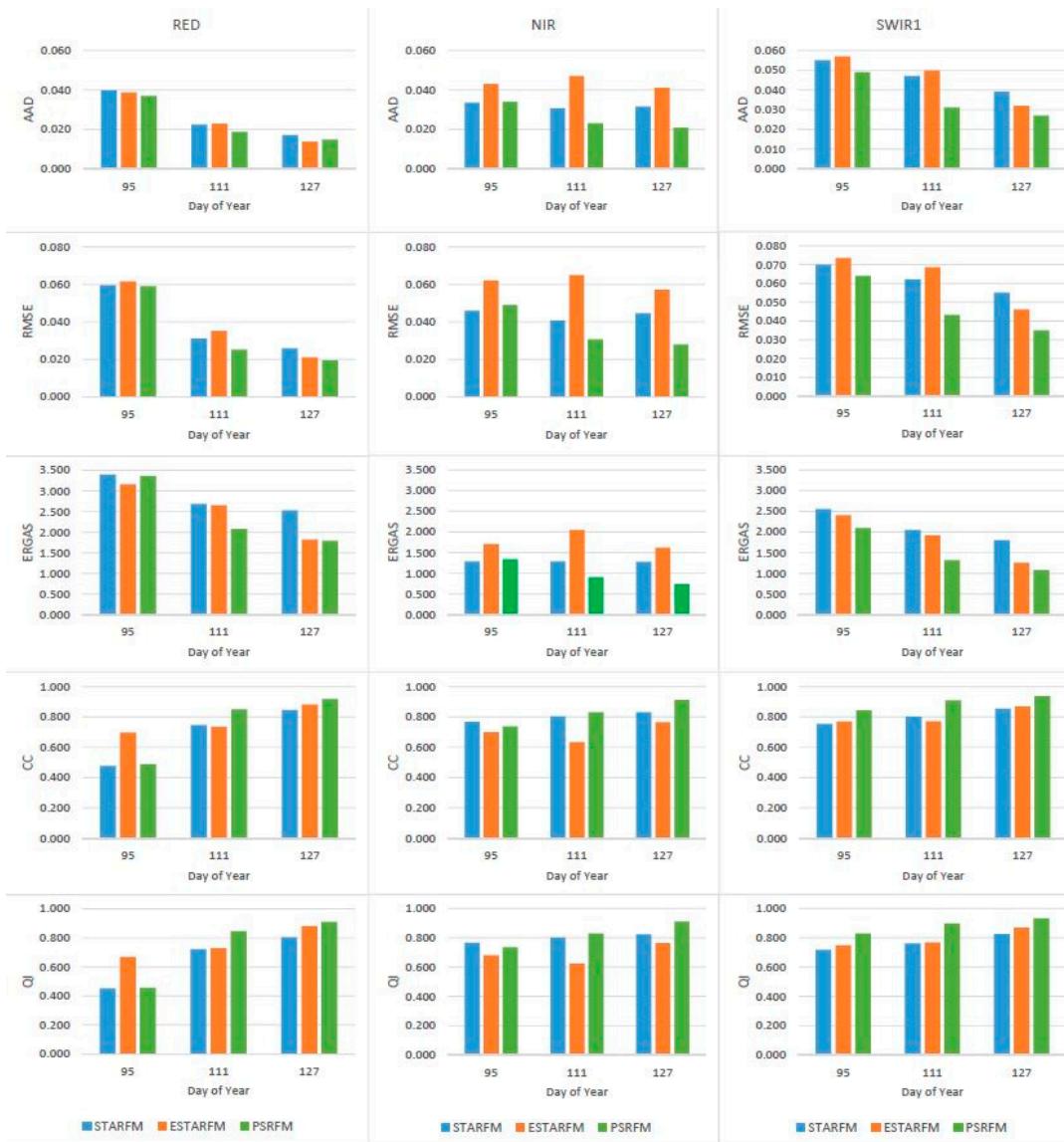
Table 3. Quality indices CC and QI of the predicted Landsat-like image time series by the three models STARFM, ESTARFM and PSRFM.

Band	Method	CC			QI		
Date (Day of Year)		**95**	**111**	**127**	**95**	**111**	**127**
RED	STARFM	0.478	0.746	0.846	0.453	0.722	0.805
	ESTARFM	**0.697**	0.735	0.883	**0.669**	0.729	0.882
	PSRFM	0.487	**0.850**	**0.917**	0.456	**0.847**	**0.911**
NIR	STARFM	**0.767**	0.804	0.832	**0.767**	0.801	0.823
	ESTARFM	0.700	0.633	0.766	0.680	0.623	0.764
	PSRFM	0.738	**0.830**	**0.911**	0.736	**0.830**	**0.910**
SWIR1	STARFM	0.755	0.802	0.855	0.717	0.762	0.826
	ESTARFM	0.769	0.772	0.870	0.749	0.768	0.869
	PSRFM	**0.844**	**0.909**	**0.938**	**0.829**	**0.899**	**0.934**
AVERAGE	STARFM	0.667	0.784	0.844	0.646	0.761	0.818
	ESTARFM	**0.722**	0.713	0.840	0.699	0.706	0.838
	PSRFM	0.690	**0.863**	**0.922**	**0.674**	**0.859**	**0.918**

Tables 2 and 3 indicate that out of the total 45 quality index values (5 indices × 3 bands × 3 dates), PSRFM outperforms other two methods for 37 data values, which counts for about 82.2% of the total quality index data values; the other 8 better quality values (18.2%) are equally shared by STARFM and ESTARFM. Overall, PSRFM outperforms STARFM and ESTARFM in predicting Landsat-like image time series. This can be seen clearly from graphical comparison of the five quality indices in Figure 6.

In addition, computation speed of the proposed PSRFM is more efficient than STARFM and ESTARFM too. Based on the current implementation of PSRFM in Matlab on a desktop PC with an Intel i5 3.2-GHz processor, 8 GB RAM and 1 TB hard disk, for a predetermined cluster number, the processing times for the test case with an image size of 800 × 800 pixels was 25 s, 2 min and 11 min for PSRFM, STARFM and ESTARFM, respectively. Even for a model optimization with a predetermined cluster number range of 4 to 16, the total processing time of PSRFM is still less than 6 min.

From the visual and quantitative comparisons of the predicted Landsat-like image time series against the true Landsat images above and the performance comparisons among the three methods PSRFM, STARFM and ESTARFM, we can conclude that the proposed PSRFM is effective and outperforms STARFM and ESTARFM.

Figure 6. Quality indices comparison of the predicted Landsat-like image time series by the three models STARFM, ESTARFM and PSRFM. Three columns present the three bands RED, NIR and SWIR1, respectively. (The smaller of AAD, REMS and ERGAS, the better quality of the predicted image; and the bigger of CC and QI, the better quality of the image).
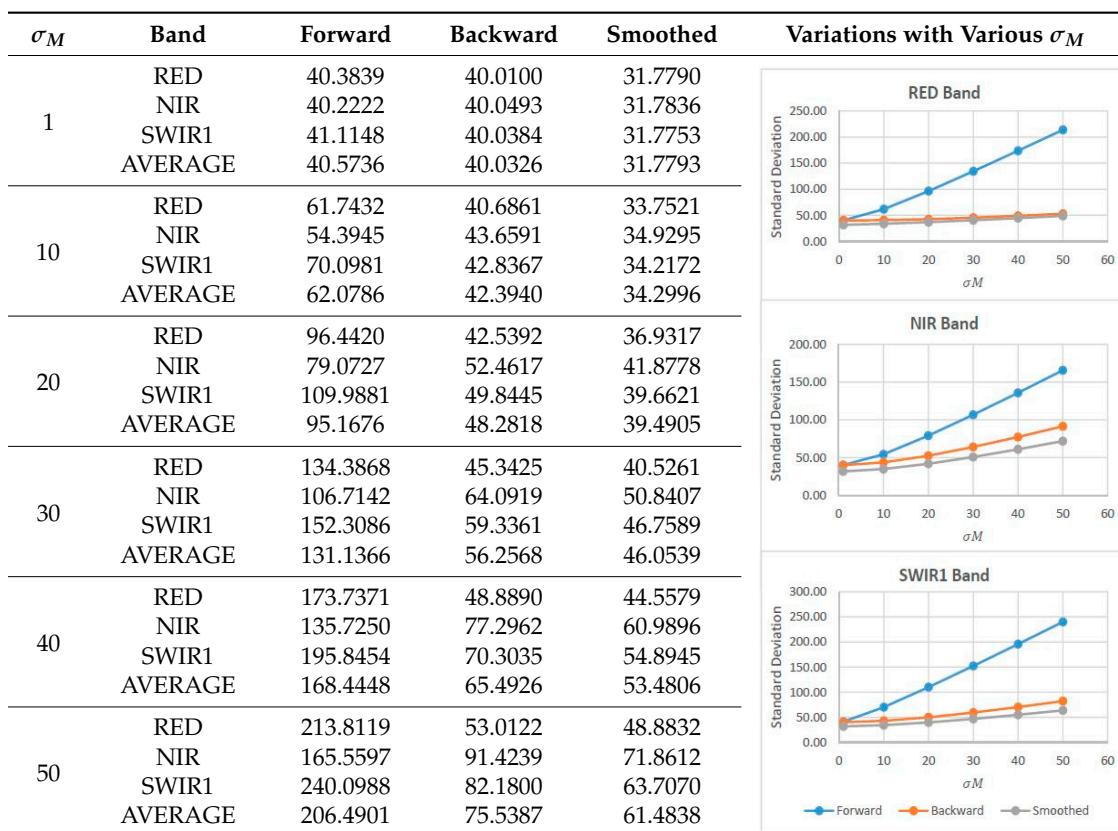
4.3. Uncertainties of the Predicted Landsat-Like Images

Another advantage of PSRFM is that it can incorporate uncertainties of the input Landsat and MODIS images and quantify uncertainty information for its predicted synthetic Landsat-like images. The uncertainty is propagated from the uncertainties of the input Landsat and MODIS images via the blending process in the form of standard deviation of each and every pixel. The system can produce uncertainty data for both forward and backward predictions as the smoother could reduce uncertainty of the smoothed image by uncertainty-weighted average. In a general sense, a smaller uncertainty value of a pixel represents a more precise prediction of the pixel value.

The uncertainty of a predicted image is from two sources: one is the PSRFM blending process and the other is from the priori uncertainty of the input images for the blending. Since actual priori uncertainty of the Landsat and MODIS images is unavailable, to demonstrate how the priori uncertainties influence the uncertainty estimation of the output images and find out the significance

of the uncertainty to the predicted images, we predicted a number of synthetic Landsat-like images using the same input Landsat and MODIS images but with different priori uncertainties for the input images. This is to simulate different accuracy of input images due to various reasons such as cloud/haze residual and so forth. Table 4 gives the average of the pixel prediction uncertainties (standard deviations) estimated by the forward, backward and smoothed predictions of PSRFM with an empirically fixed priori pixel uncertainty $\sigma_R = 40$ for the input Landsat images and various priori pixel uncertainties σ_M ranged from 1 to 50 for the input MODIS images. The date of the prediction Landsat-like image is 20 April 2016 and of the input pairs of Landsat and MODIS images are 19 March 2016 (t_0) and 22 May 2016 (t_2), respectively.

Table 4. Mean of prediction uncertainties (standard deviations, pixel value scaled by 10,000) estimated by PSRFM with an empirically fixed priori pixel uncertainty of $\sigma_R = 40$ for input Landsat images and various priori pixel uncertainties σ_M ranged from 1 to 50 for input MODIS images.

σ_M	Band	Forward	Backward	Smoothed	Variations with Various σ_M
1	RED	40.3839	40.0100	31.7790	
	NIR	40.2222	40.0493	31.7836	
	SWIR1	41.1148	40.0384	31.7753	
	AVERAGE	40.5736	40.0326	31.7793	
10	RED	61.7432	40.6861	33.7521	
	NIR	54.3945	43.6591	34.9295	
	SWIR1	70.0981	42.8367	34.2172	
	AVERAGE	62.0786	42.3940	34.2996	
20	RED	96.4420	42.5392	36.9317	
	NIR	79.0727	52.4617	41.8778	
	SWIR1	109.9881	49.8445	39.6621	
	AVERAGE	95.1676	48.2818	39.4905	
30	RED	134.3868	45.3425	40.5261	
	NIR	106.7142	64.0919	50.8407	
	SWIR1	152.3086	59.3361	46.7589	
	AVERAGE	131.1366	56.2568	46.0539	
40	RED	173.7371	48.8890	44.5579	
	NIR	135.7250	77.2962	60.9896	
	SWIR1	195.8454	70.3035	54.8945	
	AVERAGE	168.4448	65.4926	53.4806	
50	RED	213.8119	53.0122	48.8832	
	NIR	165.5597	91.4239	71.8612	
	SWIR1	240.0988	82.1800	63.7070	
	AVERAGE	206.4901	75.5387	61.4838	

From Table 4, we can see that the priori uncertainties of the input Landsat and MODIS images will influence the uncertainty estimations of the output images significantly and as expected, the larger the priori uncertainty of the input images, the larger the estimated uncertainty of the predicted images and vice versa. Depending on qualities of the input MODIS images and the priori uncertainty values used for them, the estimated uncertainties of the forward and backward predictions may be considerably different. Such different uncertainties of the forward and backward images would affect the quality of the smoothed image. To limit such difference or better balance their effects on the smoothed image, a smaller priori uncertainty value should be assigned to the MODIS input images. This can be seen clearly from the uncertainty variations of the forward, backward and smoothed predictions in Table 4.
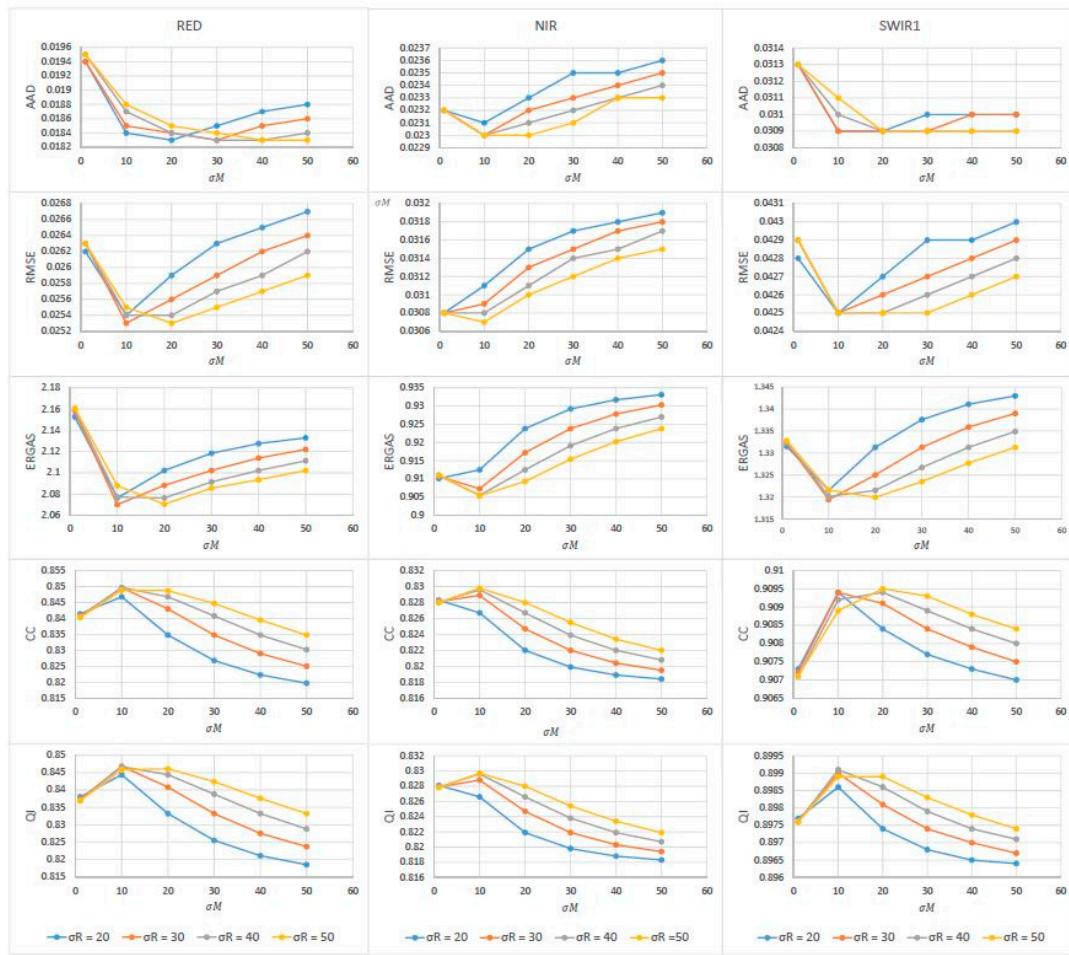
To explore the quality of the predicted images with different priori uncertainties of the input images, we computed the five quality indices of those predictions against the observed image. The quality indices comparison in Table 5 and Figure 7 shows that the impacts of different priori uncertainty sets on the final prediction are not very significant although some minor differences of

those index values exist. This means that PSRFM prediction system is stable and not affected much by the priori uncertainty of the input image even if the uncertainties of the predicted images vary. From the variations of the quality indices with the different priori uncertainties sets, we can see that a small priori uncertainty less than 10 for the input MODIS images is not recommended because it will decrease the quality of the predicted images. However, overall, a priori uncertainty between 10 and 20 for the input MODIS images and a priori uncertainty above 30 for the input Landsat images are helpful in improving quality of the smoothed image from the forward and backward predictions.
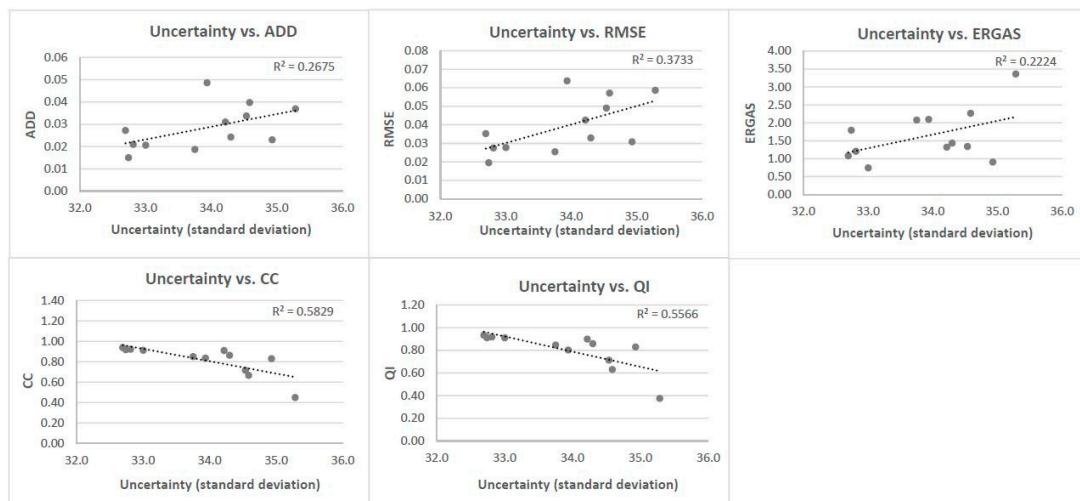
Table 5. Quality indices of predicted Landsat-like images by PSRFM with an empirically fixed priori pixel uncertainty of $\sigma_R = 40$ for the input Landsat images and different priori pixel uncertainties σ_M ranged from 1 to 50 for the input MODIS images for prediction dated 20 April 2016.

Band	σ_M	AAD	RMSE	ERGAS	CC	QI
RED	1	0.0195	0.0263	2.1600	0.8404	0.8370
	10	0.0187	**0.0254**	2.0772	**0.8497**	**0.8467**
	20	0.0184	0.0254	**2.0763**	0.8468	0.8444
	30	0.0183	0.0257	2.0914	0.8408	0.8388
	40	0.0183	0.0259	2.1022	0.8348	0.8332
	50	0.0184	0.0262	2.1113	0.8302	0.8288
NIR	1	0.0232	0.0308	0.9109	0.8280	0.8279
	10	**0.0230**	**0.0308**	**0.9055**	**0.8296**	**0.8296**
	20	0.0231	0.0311	0.9125	0.8267	0.8266
	30	0.0232	0.0314	0.9191	0.8239	0.8238
	40	0.0233	0.0315	0.9238	0.8220	0.8219
	50	0.0234	0.0317	0.9270	0.8208	0.8207
SWIR1	1	0.0313	0.0429	1.3326	0.9071	0.8976
	10	0.0310	**0.0425**	**1.3201**	0.9092	**0.8991**
	20	**0.0309**	0.0425	1.3215	**0.9094**	0.8986
	30	0.0309	0.0426	1.3267	0.9089	0.8979
	40	0.0309	0.0427	1.3313	0.9084	0.8974
	50	0.0309	0.0428	1.3349	0.9080	0.8971
AVERAGE	1	0.0247	0.0333	1.4679	0.8585	0.8542
	10	0.0242	**0.0329**	**1.4343**	**0.8628**	**0.8585**
	20	**0.0241**	0.0330	1.4368	0.8610	0.8565
	30	0.0242	0.0332	1.4458	0.8579	0.8535
	40	0.0242	0.0334	1.4524	0.8551	0.8508
	50	0.0243	0.0335	1.4578	0.8530	0.8488

To further explore the meaning of the uncertainty to the predicted images, relationships between the estimated uncertainties and the quality of predicted images are studied. Figure 8 shows the correlations between the uncertainty values (averaged standard deviation) and the five quality indices listed in Tables 2 and 3, that is, calculated by using priori uncertainty $\sigma_M = 10$ for input MODIS images and $\sigma_R = 40$ for input Landsat images. From these graphs, it can be seen that, statistically, the uncertainty does reflect the quality of the predicted images—a smaller uncertainty indicates a smaller ADD, RMSE, ERGAS and a larger CC and QI and vice versa. Therefore, in general, with priori uncertainties of the input images specified, a smaller uncertainty gives us indication of higher quality of the prediction and a larger uncertainty implies poorer quality of the prediction.

Figure 7. Quality indices comparison of the predicted Landsat-like images by PSRFM with different priori uncertainties set ranged from 20 to 50 and from 1 to 50 for the input Landsat and MODIS images, respectively. The different colors represent the priori uncertainties used for the input Landsat images σ_R. The graphic curves represent changes of quality indices with different priori uncertainties (σ_M) used for the input MODIS images. The columns present the three bands RED, NIR and SWIR1, respectively.

Figure 8. Uncertainty vs. Quality indices.

In summary, although the estimated uncertainty of the predicted Landsat-like images by PSRFM varies with different priori uncertainties of the input images, its quality, compared to its counterpart observed image, is not significantly affected. Hence, with an input set of priori uncertainty of the input images, the estimated uncertainty will give us an overall quality estimation of the forward, backward and the smoothed predictions relatively. That is, a smaller uncertainty means a higher quality of the predicted pixels. To balance the effects between the forward and backward predictions on the smoothed image, from Figure 7, we can see that a smaller priori uncertainty from 10 to 20 for the MODIS images but a bigger priori uncertainty from 30 to 50 for the Landsat images are recommended. However, whether these recommendations are valid for applications in other locations and landscapes (e.g., an urban area) or different seasons is still worth to be verified by more and different test cases.

5. Conclusions

This paper proposed a new Prediction Smooth Reflectance Fusion Model (PSRFM) to improve current Landsat and MODIS reflectance fusion models for both gradual vegetation phenological and abrupt land-cover type changes. PSRFM combines a dynamic prediction model based on the linear spectral mixing algorithm with a smoothing filter based on weighted average of uncertainties of the predicted images with an option of under constraints of Normalized Difference Snow Index (NDSI) or Normalized Difference Vegetation Index (NDVI). PSRFM can predict synthetic Landsat-like images from one pair or two pairs of Landsat and MODIS images, and, at the same time, generate the uncertainty information of the predicted images. Further, to optimize the modelling of both gradual vegetation phenological change and land-cover changes, a model optimization method based on Maximizing Correlation but Smaller Residuals (MCSR) is introduced.

The proposed PSRFM possesses following desirable characteristics: (1) it works with one pair or two pairs of Landsat and MODIS images; (2) it can incorporate input image uncertainty during prediction and estimate prediction uncertainty; (3) it can track gradual vegetation phenological changes through the dynamic prediction model and deal with abrupt land-cover type changes through a model correction based on residuals of the predicted images; and (4) for predictions using two pairs of input images, the results can be further improved through the constrained smoothing filter based on weighted average of uncertainties and on NDSI or NDVI for certain applications.

The proposed model PSRFM was tested to produce a Landsat-like image time series from an observed satellite imagery dataset. Comparisons of the predicted versus actual Landsat images show that the proposed PSRFM is effective and outperforms STARFM (Spatial and Temporal Adaptive Reflectance Fusion Model) and ESTARFM (Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model) in terms of visual and quantitative evaluations.

Notwithstanding the improvement that PSRFM represents over current methods, there is room for improvement. First, modelling abrupt land-cover changes based on advanced analysis of the prediction residuals including applying proper interpolations of the residuals is still a challenge task. Second, the prediction model for gradual vegetation phenological changes is based on the spectral linear mixing theory and depends on the classification of Landsat images. The clustering optimization implemented in this study is a time-consuming process and ignores the land-cover change information from MODIS observations. It would be beneficial to test and integrate newly developed clustering approaches (e.g., [7,8]). Third, complete clear-sky Landsat images for input to PSRFM are difficult to find for an entire scene. For practical applications, an improved algorithm that can deal with Landsat images with cloud-contaminated pixels and with additional data such as quality assessment (QA) band is also required.

Author Contributions: D.Z. and F.Z. jointly conceived and designed the experiments; D.Z. developed the processing program and performed the experiments; F.Z. prepared the test dataset and computed the NDSI and so forth, D.Z. and F.Z. jointly analyzed the data and wrote the paper.

References

1. Gao, F.; Masek, J.; Schwaller, M.; Hall, F. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2207–2218.
2. Zhu, X.; Chen, J.; Gao, F.; Chen, X.; Masek, J.G. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* **2010**, *114*, 2610–2623. [CrossRef]
3. Hilker, T.; Wulder, M.A.; Coops, N.C.; Linke, J.; McDermid, G.; Masek, J.G.; Gao, F.; White, J.C. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sens. Environ.* **2009**, *113*, 1613–1627. [CrossRef]
4. Zhukov, B.; Oertel, D.; Lanzl, F.; Reinhackel, G. Unmixing-based multisensor multiresolution image fusion. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1212–1226. [CrossRef]
5. Huang, B.; Zhang, H. Spatio-temporal reflectance fusion via unmixing: Accounting for both phenological and land-cover changes. *Int. J. Remote Sens.* **2014**, *35*, 6213–6233. [CrossRef]
6. Zhu, X.; Helmer, E.H.; Gao, F.; Liu, D.; Chen, J.; Lefsky, M.A. A flexible spatiotemporal method for fusing satellite images with different resolutions. *Remote Sens. Environ.* **2016**, *172*, 165–177. [CrossRef]
7. Hazaymeh, K.; Hassan, Q.K. Spatiotemporal image-fusion model for enhancing temporal resolution of Landsat-8 surface reflectance images using MODIS images. *J. Appl. Remote Sens.* **2015**, *9*, 096095. [CrossRef]
8. Kwan, C.; Budavari, B.; Gao, F.; Zhu, X. A Hybrid Color Mapping Approach to Fusing MODIS and Landsat Images for Forward Prediction. *Remote Sens.* **2018**, *10*, 520. [CrossRef]
9. Wang, J.; Huang, B. A Rigorously-Weighted Spatiotemporal Fusion Model with Uncertainty Analysis. *Remote Sens.* **2017**. [CrossRef]
10. Song, H.; Huang, B. Spatiotemporal satellite image fusion through one-pair image learning. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1883–1896. [CrossRef]
11. Emelyanova, I.V.; McVicar, T.R.; Van Niel, T.G.; Li, L.T.; van Dijk, A.I.J.M. 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.* **2013**, *133*, 193–209. [CrossRef]
12. Hirvonen, R.A. *Adjustments by Least Squares in Geodesy and Photogrammetry*; Ungar: New York, NY, USA, 1971; 261p, ISBN1 0804443971, ISBN2 978-0804443975.
13. Wolf, P.R. Survey Measurement Adjustments by Least Squares. In *The Surveying Handbook*; Springer: Boston, MA, USA, 1995; pp. 383–413.
14. Dozier, J. Spectral signature of alpine snow cover from the Landsat Thematic Mapper. *Remote Sens. Environ.* **1989**, *28*, 9–22. [CrossRef]
15. Hall, D.K.; Riggs, G.A.; Salomonson, V.V. Development of methods for mapping global snow cover using Moderate Resolution Imaging Spectroradiometer (MODIS) data. *Remote Sens. Environ.* **1995**, *54*, 127–140. [CrossRef]
16. Weier, J.; Herring, D. Measuring Vegetation (NDVI & EVI). 2000. Available online: https://earthobservatory. nasa.gov/Features/MeasuringVegetation/measuring_vegetation_1.php (accessed on 1 December 2017).
17. Yengoh, G.T.; Dent, D.; Olsson, L.; Tengberg, A.E.; Tucker, C.J. *The Use of the Normalized Difference Vegetation Index (NDVI) to Assess Land Degradation at Multiple Scales: A Review of the Current Status, Future Trends, and Practical Considerations*; Lund University Center for Sustainability Studies (LUCSUS); The Scientific and Technical Advisory Panel of the Global Environment Facility (STAP/GEF): Lund, Sweden, 2014.
18. Zhou, F.; Zhang, A. Methodology for estimating availability of cloud-free image composites: A case study for southern Canada. *Int. J. Appl. Earth Observ. Geoinf.* **2012**, *21*, 17–31. [CrossRef]
19. Zhou, F.; Zhang, A.; Townley-Smith, L. A data mining approach for evaluation of optimal time-series of MODIS data for land cover mapping at a regional level. *ISPRS J. Photogramm. Remote Sens.* **2013**, *84*, 114–129. [CrossRef]

20.   Emelyanova, I.; McVicar, T.; Van Niel, T.; Li, L.; Van Dijk, A. *Landsat and MODIS Data for the Lower Gwydir Catchment Study Site, v3*; Data Collection; CSIRO: Canberra, Australia, 2013.

21.   Gevaert, C.M.; García-Haro, F.J. A comparison of STARFM and an unmixingbased algorithm for Landsat and MODIS data fusion. *Remote Sens. Environ.* **2015**, *156*, 34–44. [CrossRef]

22.   Gao, F.; Masek, J.; Wolfe, R.; Huang, C. Building consistent medium resolution satellite data set using moderate resolution imaging spectroradiometer products as reference. *J. Appl. Remote Sens.* **2010**, *4*, 043526. [CrossRef]

23.   Vaiopoulos, A.D. Developing Matlab scripts for image analysis and quality assessment. In *Proc. SPIE 8181, Earth Resources and Environmental Remote Sensing/GIS Applications II*; International Society for Optics and Photonics: Bellingham, WA, USA, 2011; p. 81810B. [CrossRef]

24.   Zhou, W.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612.

25.   Pa, J.; Hegdeb, A.V. A Review of Quality Metrics for Fused Image. *Aquat. Procedia* **2015**, *4*, 133–142.

26.   Zhang, Y. Methods for Image Fusion Quality Assessment—A Review, Comparison and Analysis. In Proceedings of the International Achieves of the Photogrammetry, Remote Sensing Information Sciences, Beijing, China, 11–13 July 2008; pp. 1101–1109.