Article

Assessing a Multi-Platform Data Fusion Technique in Capturing Spatiotemporal Dynamics of Heterogeneous Dryland Ecosystems in Topographically Complex Terrain

Peter J. Olsoy [1,2], Jessica Mitchell [3], Nancy F. Glenn [1] and Alejandro N. Flores [1,*]

[1] Department of Geosciences, Boise State University, Boise, ID 83702, USA; peterolsoy@gmail.com (P.J.O.); nancyglenn@boisestate.edu (N.F.G.)
[2] School of the Environment, Washington State University, Pullman, WA 99164, USA
[3] Department of Geography and Planning, Appalachian State University, Boone, NC 28608, USA; mitchelljj@appstate.edu
[*] Correspondence: lejoflores@boisestate.edu; Tel.: +1-208-426-2903

Abstract: Water-limited ecosystems encompass approximately 40% of terrestrial land mass and play a critical role in modulating Earth's climate and provisioning ecosystem services to humanity. Spaceborne remote sensing is a critical tool for characterizing ecohydrologic patterns and advancing the understanding of the interactions between atmospheric forcings and ecohydrologic responses. Fine to medium scale spatial and temporal resolutions are needed to capture the spatial heterogeneity and the temporally intermittent response of these ecosystems to environmental forcings. Techniques combining complementary remote sensing datasets have been developed, but the heterogeneous nature of these regions present significant challenges. Here we investigate the capacity of one such approach, the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) algorithm, to map Normalized Difference Vegetation Index (NDVI) at 30 m spatial resolution and at a daily temporal resolution in an experimental watershed in southwest Idaho, USA. The Dry Creek Experimental Watershed captures an ecotone from a sagebrush steppe ecosystem to evergreen needle-leaf forests along an approximately 1000 m elevation gradient. We used STARFM to fuse NDVI retrievals from the MODerate-resolution Imaging Spectroradiometer (MODIS) and Landsat during the course of a growing season (April to September). Specifically we input to STARFM a pair of Landsat NDVI retrievals bracketing a sequence of daily MODIS NDVI retrievals to yield daily estimates of NDVI at resolutions of 30 m. In a suite of data denial experiments we compared these STARFM predictions against corresponding Landsat NDVI retrievals and characterized errors in predicted NDVI. We investigated how errors vary as a function of vegetation functional type and topographic aspect. We find that errors in predicting NDVI were highest during green-up and senescence and lowest during the middle of the growing season. Absolute errors were generally greatest in tree-covered portions of the watershed and lowest in locations characterized by grasses/bare ground. On average, relative errors in predicted average NDVI were greatest in grass/bare ground regions, on south-facing aspects, and at the height of the growing season. We present several ramifications revealed in this study for the use of multi-sensor remote sensing data for the study of spatiotemporal ecohydrologic patterns in dryland ecosystems.

Keywords: multispectral sensor data fusion; hillslope scale; vegetation; water-limited ecosystems; aspect; topography; STARFM

1. Introduction

Water-limited ecosystems cover approximately 40% of Earth's terrestrial surface and, despite lower primary productivity than forested systems on a per unit land area basis, exert significant controls on global water, energy, and biogeochemical cycles (e.g., [1,2]). Some drylands are among the most biodiverse areas in the world, and dryland biodiversity conservation is critical to sustainable development and food and water security [3]. Drylands are particularly susceptible to degradation through disturbances (acting independently or in conjunction) such as fire, invasive species (e.g., [4]), grazing (e.g., [5,6]), and climate change (e.g., [7]). Spatiotemporal patterns of vegetation in water-limited ecosystems are complex, and arise as a function of interacting abiotic and biotic processes that exert influence across a large range of spatial and temporal dimensions. Spatial variability in terrestrial vegetation at hillslope scales (e.g., 10 s to 100 s of m) in these ecosystems is both influenced by and can reveal important patterns in surface water, energy, erosion, and biogeochemical cycling (e.g., [8,9]). Patterns in vegetation can reflect and conform to gradients in abiotic controls like solar radiation, and topographic convergence, and soil water (e.g., [10,11]). A number of biotic factors can also influence patterns of terrestrial vegetation at these spatial scales. Herbivory has been shown to influence the distribution and abundance of vegetation at the relatively fine spatial scales such as hillslopes [5,12]. At the same time, biomass in water-limited ecosystems (or proxies such as remotely sensed greenness) can exhibit relatively rapid temporal variability controlled, among other things, by phenology; disturbances such as fire [13], water stress, and insect infestation [14]; and atmospheric teleconnections that can produce rare but large precipitation events [15].

The complexity, global significance, and sensitivity of dryland ecosystems motivates a need to develop improved capacities to monitor variability in terrestrial vegetation in reasonably fine spatial and temporal detail. The ability to accurately characterize vegetation properties across large regions of space and decades in time at spatial resolutions approaching hillslopes and temporal resolutions of weeks or days would provide the ability to understand the relative influence of various abiotic and biotic drivers of vegetation dynamics in space and time, and significantly enhance the ability to develop, parameterize, and verify spatial models of dryland ecosystems.

Of particular value to the ecohydrology community is the potential to produce historical datasets characterizing spatiotemporal variation in vegetation conditions in sufficiently fine detail to parameterize ecological and hydrologic models. These datasets are of particular interest for ecohydrologic process investigations because spatiotemporal patterns of vegetation are, in effect, an integrative and macroscopically observable manifestation of local patterns of water, energy, and nutrient cycling. Further, these datasets could also be used to condition models of vegetation dynamics, potentially improving the accuracy of model-predicted water, energy, and nutrient cycling and uncertainty quantification and propagation. Models with dynamic vegetation components, particularly those that prognostically simulate aboveground biomass or other complementary elements of phenology (e.g., [16–19]) require multitemporal vegetation remote sensing datasets in order to constrain values of model parameters and their uncertainty (e.g., [20,21]). In addition, while the vegetation remote sensing record has already led to significant advancement of these models, the ability to obtain reliable spatiotemporal datasets characterizing important attributes of terrestrial vegetation (i.e., Normalized Difference Vegetation Index, NDVI) at much higher spatial resolutions could significantly advance: (1) the ability to quantify spatiotemporal patterns at resolutions not resolved by global land models, (2) the development of parameterizations of sub-pixel resolution ecological processes within these global land models where needed, and (3) application of similar dynamic vegetation models at resolutions approaching individual hillslopes (e.g., [22,23]).

Satellites provide global observations of the Earth surface at predictable temporal revisit intervals and play a critical role in monitoring terrestrial ecosystems. Since the 1970s, for instance, the Landsat program has provided insight into global vegetation patterns at a 30 m resolution (e.g., [24]). The MODerate-resolution Imaging Spectroradiometer (MODIS) sensors on NASA's Terra and Aqua satellites provide global vegetation products at spatial resolutions of between 250–1000 m [25,26].

Until the recent launch of Sentinel-2, the configuration of these two platforms demonstrated the significant tradeoffs for remote sensing of dryland terrestrial ecosystems. Specifically, the higher resolution Landsat products are associated with a revisit of 16 days at best, while some MODIS vegetation data products are available at daily intervals. Both platforms suffer from the risk of clouds contaminating individual images because they rely on the visible and infrared portions of the electromagnetic spectrum. Recognizing the complementarity of the Landsat and MODIS platforms, however, previous studies have sought to use data fusion techniques to leverage the high spatial resolution of Landsat and fine temporal revisit of MODIS (e.g., [27–30]). These data fusion frameworks are particularly attractive for dryland ecosystems because remote sensing techniques in drylands rely on phenological changes to elucidate native from non-native vegetation (e.g., [31]). In addition, the heterogeneity and high soil albedo in drylands can require higher spatial resolution to properly identify vegetation community types [32]. The Sentinel-2 satellite with global coverage every 5 days (with two satellites) and multispectral bands similar to Landsat 8 OLI at 10 to 20 m (and three additional bands at 60 m) also have the potential to fill this data gap in dryland ecosystems.

In this study, we evaluate the ability of one such technique and software package, the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) [27], to accurately capture spatiotemporal variations in vegetation patterns in a water-limited ecosystem with steep environmental gradients. The STARFM is a freely-available data-driven algorithm that characterizes spatial patterns of atmospherically corrected radiance data from sequences of Landsat imagery and then synthesizes imagery at the resolution of Landsat (i.e., 30 m) by downscaling the MODIS imagery of that same landscape during the intervening periods of time. Like [33], we opted to use the STARFM algorithm rather than more complex variations such as the Enhanced STARFM (ESTARFM; [34]), which requires at least two pairs of training images (fine scale spatial resolution such as Landsat and coarse scale resolution such as MODIS) [34], or a modified version of ESTARFM referred to as mESTARFM [35] because testing is limited and some results are inconclusive about the use of ESTARFM in contrasting landscapes [36]. The usefulness of STARFM in heterogeneous landscapes and semi-arid environments has been demonstrated in several studies [29,34,37,38]. The [34] study found that ESTARFM improves the accuracy of predicted reflectance in three reflectance bands for heterogeneous forested landscapes, while the [29] study randomly sampled vegetation pixels for reflectance and determined that STARFM was feasible for time series compositing in dryland systems. The [38] study was conducted in a dryland ecosystem and evaluated the performance of red, NIR and NDVI vegetation pixels stratified by covertype and focused on comparing results among bands and vegetation type as well as the influence of image base pair selection and lag time across a growing season on model performance. The [37] study evaluated the use of STARFM for monitoring forest disturbance and regrowth using multi-date LiDAR for validation. The present study builds on these previous investigations by evaluating the influences of vegetation type, timing, topography and snow on STARFM performance with the integration of LiDAR and in the context of ecohydrologic modeling applications. Specifically, we focus on the accuracy of 30 m NDVI images synthesized with STARFM in a semiarid watershed that exhibits significant variation in environmental and physiographic conditions. The work is designed to address the following research questions: (1) How accurately can 30-m NDVI images obtained by fusing Landsat and MODIS NDVI retrievals using the STARFM algorithm reproduce corresponding, withheld, Landsat retrievals? (2) How do errors in STARFM-synthesized NDVI images vary with: (a) time of year, (b) vegetation functional type as captured by LiDAR canopy height models (CHM), (c) LiDAR-derived topographic aspect, and (d) presence/absence of snow as captured by the Normalized Difference Snow Index (NDSI)?

The remainder of the paper is organized as follows: Section 2 overviews the methods, including a description of the study area, experimental setup, and error metrics. Section 3 details the results of the analyses in the context of the research questions posed above. Section 4 discusses these results, focusing on how errors in STARFM predictions could conceivably propagate to ecohydrologic simulations of

water, energy, and nutrient cycling. Section 5 presents conclusions of the study and highlights new potential avenues of inquiry identified by this work.

2. Methods

2.1. Study Area

We tested the STARFM model at Dry Creek Experimental Watershed (DCEW), a 27 km^2 site in southwestern Idaho, USA (Figure 1; 43.716°N, 116.127°W; elevation: ~1000–2100 m). The climate is highly dependent on elevation with mean annual temperature ranging from 10.0 °C at lower elevations to 8.8 °C at intermediate elevations, and 6.4 °C at upper elevations, while annual precipitation varies from approximately 370 to 1000 mm. Soils consist of Argixerolls, Haploxerolls, and Haplocambids, with the upper watershed classified as sandy loam and lower watershed classified as loam [39]. Sagebrush and grass communities dominate the lower elevations, which transitions into ponderosa pine (*Pinus ponderosa*) and Douglas fir (*Pseudotsuga menziesii*) at intermediate and upper elevations.
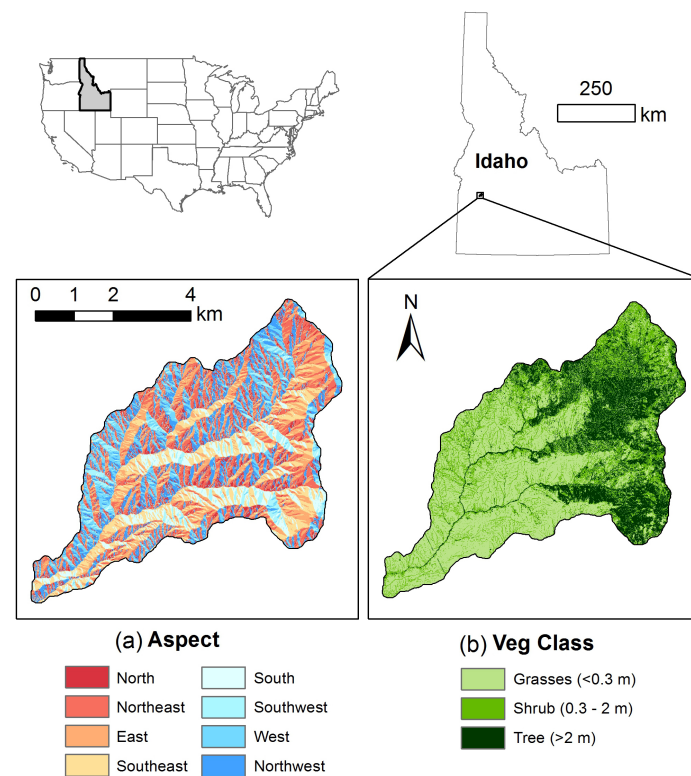

Figure 1. Location of Dry Creek Experimental Watershed (DCEW) in Idaho, USA. We created a LiDAR-derived aspect layer (**a**), and a layer of dominant vegetation functional type (**b**). DCEW has a strong vegetation gradient with tree-dominated areas largely contained to the eastern half of the watershed.

2.2. Remote Sensing Datasets

We selected the 2007 growing season (March–September 2007) because it contained the most cloud-free consecutive sequence of Landsat scenes as screened on Glovis (http://glovis.usgs.gov/), and coincided with airborne light detection and ranging (LiDAR) collection for our study area. The Landsat-5 (L5) Thematic Mapper (TM) scenes were processed to reflectance with LEDAPS [40]. We downloaded corresponding Nadir Bidirectonal Reflectance Distribution Function-Adjusted Reflectance (NBAR) MODIS daily 16-day composite data (MCD43A4) Version 5 from the USGS Land Processes Distributed Active Archive Center (LP DAAC) Data Pool. The NBAR MODIS composite

product is generated at a 500-m pixel resolution and combines data from both the Aqua and Terra satellites. The imagery was delivered in a sinusoidal projection then re-projected to a geographic coordinate system using the USGS MODIS Reprojection Tool. Individual Landsat and MODIS bands for red, green, and near-infrared (NIR) were extracted and saved, along with an NDVI composite, for input into the STARFM model. We also subset images from both products to a common spatial extent encompassing DCEW and surrounding areas.

Airborne LiDAR was flown 10–18 Nov 2007 (Watershed Sciences, Corvallis, OR, USA) with a Leica ALS50 Phase II instrument (Leica Geosystems. Inc. Norcross, GA, USA; 1064 nm wavelength) in a Cessna Caravan 208B aircraft flown approximately 900 m above ground level resulting in an average point density of 4.09 to 8.68 points m^{-2}. We classified the LiDAR points as vegetation or ground with the BCAL LiDAR Tools [41] to produce a canopy height model (CHM) and bare earth digital elevation model (DEM) at 5-m pixel resolution.

2.3. STARFM Algorithm

The STARFM algorithm was developed to quantify changes, sometimes rapid, in vegetation phenology during the growing season [27]. An exhaustive review of the STARFM algorithm is beyond the scope of this paper, and the reader is referred to [27] for a more thorough description of the algorithm, its assumptions and limitations, as well as other details. Here we describe relevant features of STARFM and how they were applied in this study. The STARFM algorithm uses one or more fine spatial resolution images (e.g., Landsat) paired with high frequency temporal resolution images (e.g., MODIS) to predict daily surface reflectance at the fine spatial resolution (e.g, 30 m) for dates where no Landsat images were acquired. The model assumes aggregation from finer scale homogeneous pixels into coarser heterogeneous pixels and is therefore sensitive to landscape patch size and degrades when applied to landscapes with fine scale heterogeneity. STARFM synthesizes a Landsat image using a temporally coincident pair of Landsat and MODIS images, and an additional MODIS image on the prediction date. The user has an option of including more matched pairs (k) to further constrain the prediction. For example, the weighting of surrounding pixels can be variable and dependent on cloud cover. The STARFM algorithm includes information from similar neighboring pixels (i.e., pixels of the same spectral class), giving the equation for the central pixel:

$$L\left(x_{w/2}, y_{w/2}, t_0\right) = \sum_{i=1}^{w} \sum_{j=1}^{w} \sum_{k=1}^{n} W_{ijk} \cdot \left(M\left(x_i, y_j, t_0\right) + L\left(x_i, y_j, t_k\right) - M\left(x_i, y_j, t_k\right) \right) \tag{1}$$

where L is the surface reflectance of a Landsat image, a pixel location is denoted (x_i, y_j), w is the moving window size, $(x_{w/2}, y_{w/2})$ is the central pixel of the moving window, prediction date is t_0, acquisition date is t_k, and $W_{i,j,k}$ is a weight function. None of the algorithm parameters, such as w were adjusted. We tested synthetic images created with a single coincident pair of Landsat and MODIS images ($k = 1$), and synthetic images created with two pairs of coincident Landsat and MODIS images ($k = 2$). All predicted dates for the synthetic images had a matching Landsat image collected within one to four days to assess prediction error.

2.4. Error Analysis

We assessed model accuracy with a pixel-based differencing of the Landsat images synthesized by STARFM and the withheld observation of NDVI and reflectance in the individual green, red, and near-infrared bands. Linear regression analysis was performed for all pixels in the image to determine bias (slope and intercept) and overall prediction accuracy (r^2). We calculated mean absolute error (MAE), mean signed error (MSE), standard deviation of the error, and root mean square error (RMSE). Finally, we provided the Nash-Sutcliffe Efficiency index (NSE) as an alternative to traditional regression-based analyses by comparing the prediction error with a mean value [42]. The NSE, which is most commonly used to assess the performance of hydrologic models (e.g., [43,44]), varies

from $-\infty$ to 1, where a value of 0 indicates no improvement over the mean, negative values signify degraded prediction with respect to the mean, and 1 is a perfect match to observed data.

$$\text{NSE} = 1 - \frac{\sum (\hat{y}_i - y_i)^2}{\sum (y_i - \bar{y})^2} \tag{2}$$

where $\hat{y}_i$ is the simulated value, y_i is the observed value, and $\bar{y}$ is the mean value across the observed scene.

In order to better determine the effectiveness of STARFM relative to seasonal changes, we compared the synthesized Landsat images with the two temporally closest cloud-free Landsat images. Our rationale for this comparison is that the basis for determining the value added by using the STARFM algorithm should be the extent to which the synthesized Landsat image offers a significant improvement over the temporally nearest cloud-free Landsat image (which we term a "null" model). This comparison is particularly important in water-limited ecosystems where vegetation response can be rapid in time and highly heterogeneous in space. For example, low errors in the null model indicate little change is occurring, while high errors indicate larger phenological changes. These errors provide context for the STARFM predictions in these water-limited environment. For example, cases in which the STARFM prediction has low error and the null model has high error demonstrates a greater gain of information. When the null model has low error and the STARFM prediction also has low error, there is little temporal change in vegetation for STARFM to predict and the low errors can be attributed to a lack of significant vegetation change between the two time intervals of interest. Finally, circumstances in which the STARFM prediction might indicate large errors while the null model is associated with low error might suggest the STARFM algorithm is actually producing spurious patterns of vegetation change.

As an effort to account for an additional confounding error, we calculated the normalized difference snow index (NDSI) to determine if errors early and late in the growing season could be attributed to snowmelt or accumulation between dates.

$$\text{NDSI} = \frac{\rho_{green} - \rho_{SWIR}}{\rho_{green} + \rho_{SWIR}}, \tag{3}$$

where ρ_{green} and ρ_{SWIR} are the reflectances in the green and shortwave infrared wavelengths, respectively. NDSI was calculated with the observed Landsat image and used to identify areas with snow (NDSI $\geq$ 0.4; [45]) and without snow (NDSI $<$ 0.4). The RMSE for STARFM predictions in snow-covered areas was compared with non-snow areas to determine if errors were larger.

We also investigated the extent to which errors in STARFM predictions are sensitive to the dominant vegetation type in each Landsat pixel. We accomplished this by using the extant LiDAR dataset to classify vegetation into dominant functional types and quantifying errors within each type. Specifically, we investigated the absolute and relative magnitude of STARFM prediction errors within three dominant vegetation classes obtained by screening the LiDAR-derived CHM mentioned above. The dominant vegetation class within each Landsat pixel is classified as "tree" when the average CHM height within that pixel is greater than 2 m, "shrub" when average CHM height is between 0.3 m and 2.0 m, and "grasses" when average CHM height is less than 0.3 m.

In DCEW, topographic aspect is a strong predictor of presence/absence of different plant functional types at the same elevation and the distribution of these functional types can vary with correlation lengths that are larger than those resolved by Landsat but beneath that resolved by MODIS [46–48]. Similar studies in water-limited ecosystems have also found similar aspect-dependent distributions of vegetation type [49–52]. For that reason, there is some reason to suspect that errors in STARFM-derived images might be correlated with topographic aspect both in the study area and more broadly in the water-limited ecosystems being studied. We, therefore, used a LiDAR-derived DEM to calculate the local topographic aspect in all eight cardinal directions, North, Northeast, East, Southeast,

South, Southwest, West, and Northwest. We investigated both the absolute and relative RMSE and MSE for each aspect class, as well.

3. Results

3.1. STARFM Algorithm Performance

The majority of green-up for tree and shrub pixels occurs in April and May (27 April 2007–29 May 2007), where mean NDVI values in the observed Landsat scenes increase from 0.33 to 0.51 for tree pixels and from 0.24 to 0.37 for shrub pixels (Table 1). Peak NDVI values are maintained until approximately early August, then decline into September (0.41 for trees and 0.25 for shrubs). Pixels with vegetation less than 0.3 m tall, most likely dominated by grasses, contained observed mean NDVI values comparable to shrub pixels early in the growing season (0.24 in April), with maximum production in May (0.27), followed by a steady decline to 0.10 in September.

Table 1. Mean actual Landsat-5 Normalized Difference Vegetation Index (NDVI) for each date and vegetation type in Dry Creek Experimental Watershed (DCEW), Idaho, USA. This was used to normalize/standardize the vegetation type and aspect figures.

Date	Trees	Shrubs	Grasses
27 April 2007	0.3317	0.2380	0.2406
13 May 2007	0.4281	0.3374	0.2654
29 May 2007	0.5068	0.3728	0.2436
14 June 2007	0.4734	0.3417	0.1983
30 June 2007	0.5255	0.3567	0.1780
1 August 2007	0.4367	0.2714	0.1090
17 August 2007	0.4141	0.2514	0.1009
2 September 2007	0.4110	0.2419	0.0984

Inputting two pairs of training data (Table 2, Figure 2), one pair before the predicted date and one pair after, into STARFM resulted in a more accurate synthetic NDVI image than using only one pair of training data (Table 3). In both the one-pair and two-pair cases, early spring had the highest model errors, particularly the 27 April 2007 image. Despite poorer results early in the growing season, all synthetic images in the two-pair model contained new information (NSE > 0) and images beginning in late May through the end of the growing season explained over 94% of the variation in NDVI values (Table 2). However, the one-pair model underperformed early in the growing season compared to the two-pair model, with the April synthetic image not adding additional information with an NSE of -1.06, an r^2 of 0.20, and a slope of 0.44 (Table 3). Further, both synthetic images in May (5 May 2007 and 29 May 2007) from the one-pair model performed worse than the two-pair model in all error and correlation metrics (Tables 2 and 3). Both sets of models stabilize in the middle and late growing season and show only marginal differences in error metrics (Tables 2 and 3), suggesting that the amount of training data is less important when little phenological change is occurring.

The null model that used the temporally nearest cloud-free Landsat image, rather than a synthetic STARFM-predicted image, performed even poorer in the early growing season (April and May), but stabilized to provide similar results in summer and slightly better results in September (Table 4). The negative NSE values for the first two dates (Table 4) shows a high level of phenological change that is at least partially modeled with STARFM (Table 2).

Table 2. Error and correlation statistics for full scene difference of observed and predicted Normalized Difference Vegetation Index (NDVI) values. The STARFM-predicted synthetic image was created using two pairs of training data.

Date	Data Set	MAE	MSE	RMSE	Error SD	NSE	Intercept	Slope	r^2
27 April 2007	NDVI	0.060	−0.023	0.077	0.073	0.197	0.094	0.561	0.369
13 May 2007	NDVI	0.029	0.008	0.041	0.040	0.810	0.018	0.971	0.838
29 May 2007	NDVI	0.036	−0.032	0.044	0.030	0.909	0.023	0.851	0.973
14 June 2007	NDVI	0.040	0.039	0.048	0.029	0.899	−0.002	1.124	0.984
30 June 2007	NDVI	0.036	−0.033	0.045	0.031	0.943	0.005	0.894	0.981
1 August 2007	NDVI	0.021	0.009	0.029	0.028	0.971	0.012	0.991	0.974
17 August 2007	NDVI	0.031	0.017	0.043	0.040	0.931	0.024	0.973	0.943
2 September 2007	NDVI	0.022	−0.007	0.030	0.029	0.966	0.001	0.969	0.968

MAE = mean absolute error; MSE = mean signed error; RMSE = root mean squared error; NSE = Nash-Sutcliffe efficiency index; r^2 = coefficient of determination.
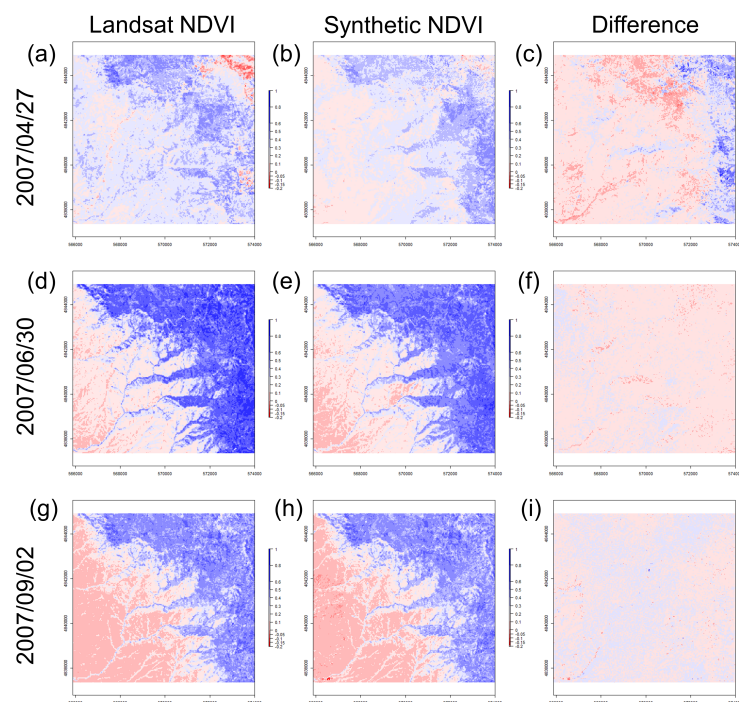

Figure 2. Landsat NDVI retrievals (NDVI$_L$) (**a**,**d**,**g**), STARFM-predicted synthetic $NDVI$ (NDVI$_S$) (**b**,**e**,**h**), and the differenced errors (NDVI$_L$ − NDVI$_S$) (**c**,**f**,**i**) for three dates throughout the 2007 growing season. Images cover the full extent of Dry Creek Experimental Watershed (DCEW), Idaho, USA.

Table 3. Error and correlation statistics for full scene difference of observed and predicted Normalized Difference Vegetation Index (NDVI) values. The STARFM-predicted synthetic image was created using one pair of training data.

Date	Data Set	MAE	MSE	RMSE	Error SD	NSE	Intercept	Slope	r^2
27 April 2007	NDVI	0.095	−0.085	0.123	0.089	−1.063	0.066	0.435	0.199
13 May 2007	NDVI	0.043	−0.015	0.060	0.058	0.604	0.018	0.901	0.693
29 May 2007	NDVI	0.037	−0.016	0.048	0.045	0.892	0.064	0.784	0.927
14 June 2007	NDVI	0.043	0.042	0.050	0.027	0.894	0.014	1.083	0.979
30 June 2007	NDVI	0.039	−0.033	0.050	0.037	0.929	0.021	0.846	0.979
1 August 2007	NDVI	0.041	0.037	0.052	0.036	0.912	0.023	1.053	0.966
17 August 2007	NDVI	0.018	0.002	0.024	0.024	0.979	0.003	0.995	0.979
2 September 2007	NDVI	0.019	−0.006	0.025	0.025	0.975	0.001	0.973	0.977

Table 4. Error and correlation statistics for full scene difference of Normalized Difference Vegetation Index (NDVI) values for observed and preceding Landsat scenes (i.e., the "null" model). Higher error indicates larger phenological changes between the given dates and serves as a baseline for determining model prediction accuracy in Tables 2 and 3.

Date	Previous Date	Data-Set	MAE	MSE	RMSE	Error SD	NSE	Inter-cept	Slope	r^2
27 April 2007	26 March 2007	NDVI	0.142	−0.142	0.152	0.056	−2.180	−0.012	0.515	0.586
13 May 2007	27 April 2007	NDVI	0.077	−0.069	0.103	0.077	−0.189	0.073	0.578	0.410
29 May 2007	13 May 2007	NDVI	0.060	−0.035	0.076	0.067	0.731	0.111	0.605	0.870
14 June 2007	29 May 2007	NDVI	0.042	0.038	0.047	0.027	0.905	0.057	0.942	0.968
30 June 2007	14 June 2007	NDVI	0.038	−0.019	0.048	0.044	0.936	0.050	0.804	0.976
1 August 2007	30 June 2007	NDVI	0.077	0.077	0.086	0.037	0.756	0.061	1.059	0.963
17 August 2007	1 August 2007	NDVI	0.023	0.017	0.030	0.025	0.966	0.004	1.051	0.982
2 September 2007	17 August 2007	NDVI	0.017	0.005	0.023	0.022	0.980	0.003	1.006	0.982

3.2. Sources of Error

Aside from bias due to differences between Landsat and MODIS processing techniques, bandwidth, acquisition time and geolocational errors, differences in prediction error were observed across bands and also varied by vegetation type. The near-infrared (NIR) band was predicted with less accuracy than the green and red bands except for 27 April 2007 where NIR was best (Table 5). On all dates for all bands, NSE was positive indicating added information from the model. While all three bands were not predicted well on 27 April 2007, it is clear that most of the error for NDVI on 13 May 2007 is due to the high error in NIR (r^2 = 0.65; RMSE = 226; NSE = 0.57) and not from the red band (r^2 = 0.96; RMSE = 97; NSE = 0.93). A similar pattern emerges throughout the rest of the growing season 29 May 2007–2 September 2007 with all error metrics being lower for the red band (NSE = 0.91–0.98; r^2 = 0.98–0.99) compared to the NIR band (NSE = 0.79–0.88; r^2 = 0.86–0.93) (Table 5).

Table 5. Error and correlation statistics for full scene difference of observed and predicted green, red, and near-infrared (NIR) reflectance values. The STARFM-predicted synthetic image was created using two pairs of training data.

Date	Data Set	MAE	MSE	RMSE	Error SD	NSE	Intercept	Slope	r^2
27 April 2007	Green	229.172	148.714	333.796	298.840	0.053	99.360	0.358	0.270
	Red	232.321	122.325	377.638	357.280	0.329	87.230	0.474	0.410
	NIR	227.374	101.759	339.601	323.999	0.405	126.700	0.517	0.464
13 May 2007	Green	63.016	−53.908	77.160	55.205	0.875	−16.910	1.083	0.954
	Red	76.917	−58.629	96.561	76.725	0.929	−14.560	1.060	0.964
	NIR	157.154	−85.866	225.754	208.788	0.570	65.180	0.744	0.647
29 May 2007	Green	68.304	54.570	82.069	61.299	0.892	30.580	0.817	0.961
	Red	87.665	56.253	106.767	90.747	0.947	27.110	0.854	0.977
	NIR	128.410	−94.798	168.274	139.031	0.822	62.970	0.769	0.897
14 June 2007	Green	104.656	−102.948	119.241	60.168	0.782	−23.220	1.086	0.961
	Red	117.969	−107.972	141.261	91.087	0.907	−27.540	1.104	0.978
	NIR	107.866	42.687	139.203	132.498	0.878	−37.650	1.133	0.932
30 June 2007	Green	71.448	60.893	85.524	60.054	0.932	22.670	0.885	0.975
	Red	91.783	63.634	111.525	91.589	0.964	22.670	0.897	0.983
	NIR	133.999	−92.914	176.863	150.493	0.867	57.930	0.790	0.923
1 August 2007	Green	71.702	−69.025	83.267	46.572	0.939	−6.799	0.999	0.981
	Red	83.710	−77.200	101.627	66.093	0.971	−47.127	0.982	0.988
	NIR	125.724	−91.562	159.062	130.067	0.842	35.800	0.841	0.898
17 August 2007	Green	44.814	−14.633	58.288	56.421	0.971	1.951	0.976	0.973
	Red	72.434	−55.137	91.863	73.476	0.976	−4.255	0.992	0.985
	NIR	119.299	−13.138	152.453	151.887	0.854	52.380	0.796	0.859
2 September 2007	Green	106.265	86.322	121.152	85.008	0.871	16.870	0.935	0.936
	Red	109.260	97.836	128.196	82.839	0.951	11.960	0.986	0.980
	NIR	164.839	131.660	197.753	147.554	0.785	44.210	0.871	0.880

Prediction error in the synthetic images was largely segregated by vegetation type. Tree pixels showed higher error on all dates, followed by shrub pixels, while ground or grass-dominated pixels had the lowest RMSE (Figure 3). After accounting for the mean NDVI for the given vegetation functional type on the given date, $RMSE/\mu_{\text{NDVI}}$ is higher for shrub pixels in the early growing season (Figure 4; 27 April 2007, 13 May 2007), but transitions to being much higher in grass pixels later in the growing season (Figure 4; 30 June 2007–2 September 2007). Further analysis of sensitivity to both vegetation and aspect indicated that prediction error is likely higher on south-facing slopes (Figures 5–8; 13 May 2007, 1 August 2007, 17 August 2007). However, prediction error was not influenced by aspect on some dates (Figure 5; 27 April 2007, 29 May 2007, 30 June 2007), suggesting that these errors may be due to a seasonal or phenological effect. We found snow to be a significant cause of error in the first early spring synthetic image with snow-covered pixels having a mean RMSE of 0.146, while snow free pixels had mean RMSE of 0.075. Snow did not appear to be a source of error in the final synthetic image (2 September 2007), with snow-covered pixels having mean RMSE of 0.029 and snow free pixels having mean RMSE of 0.0301.
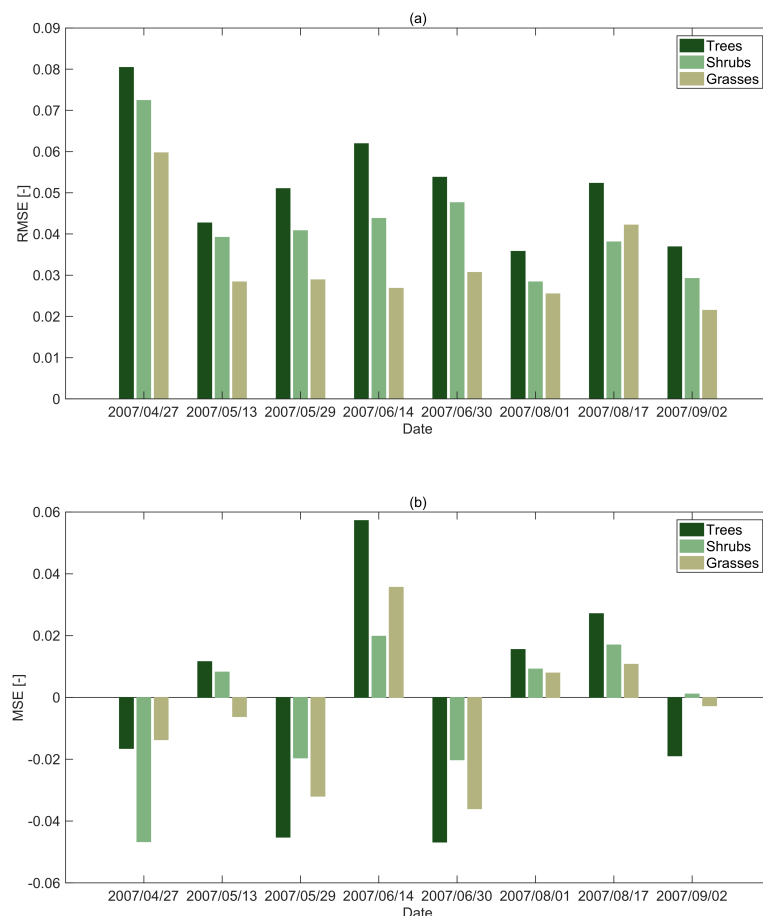
Figure 3. Root mean square error (RMSE) (**a**) and mean signed error (MSE), (**b**) separated by dominant vegetation functional type over the 2007 growing season at Dry Creek Experimental Watershed (DCEW), Idaho, USA.
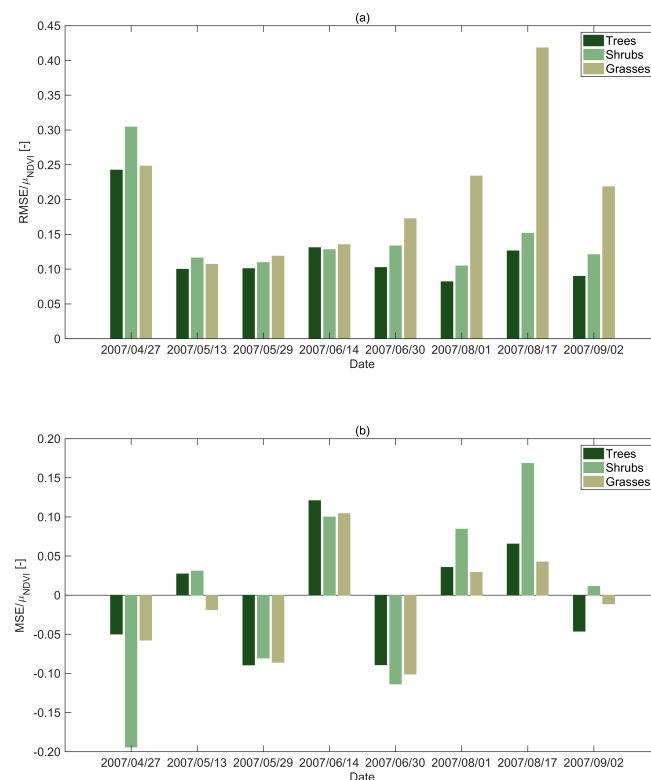
Figure 4. Normalized root mean square error ($RMSE/\mu_{\mathrm{NDVI}}$) (**a**) and normalized mean signed error (MSE/μ_{NDVI}), (**b**) separated by dominant vegetation functional type over the 2007 growing season at Dry Creek Experimental Watershed (DCEW), Idaho, USA. Errors were normalized by dividing RMSE and MSE by the mean normalized difference vegetation index (NDVI) for the given vegetation functional type on the given date (see Table 1).
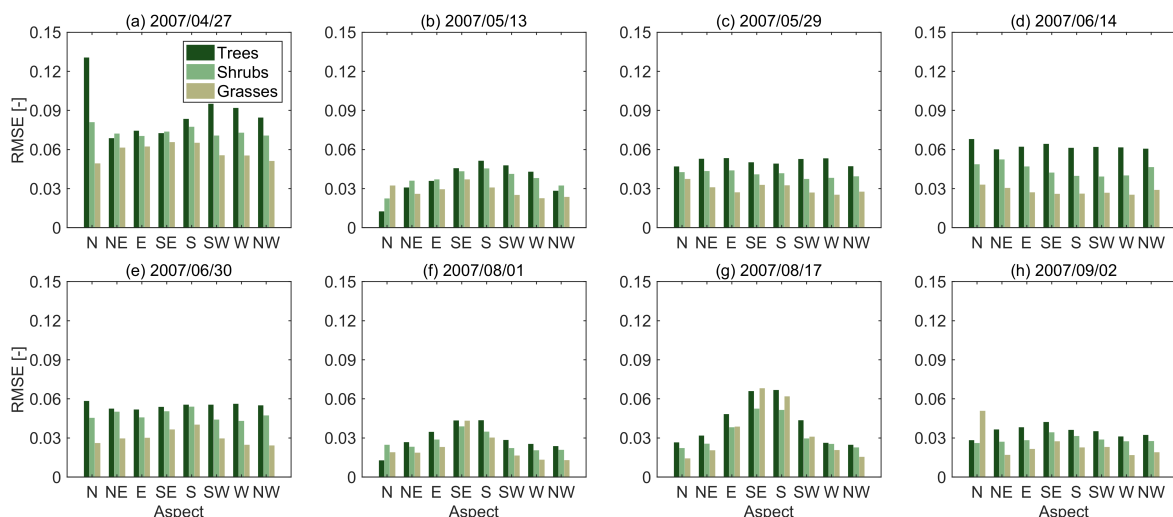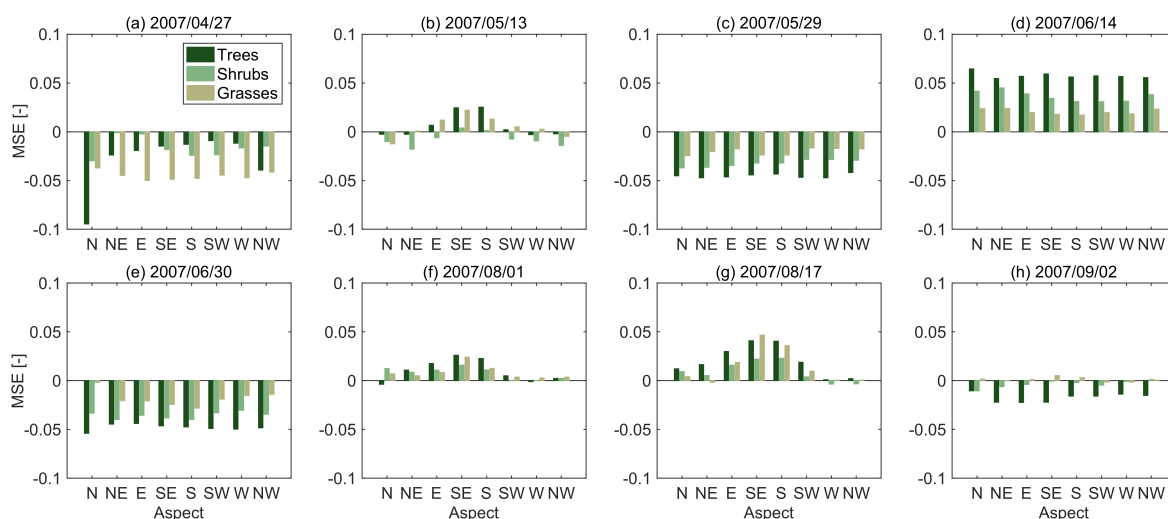

Figure 5. Root mean square error (RMSE) separated by dominant vegetation functional type and aspect over the 2007 growing season at Dry Creek Experimental Watershed (DCEW), Idaho, USA for (**a**) 27 April 2007, (**b**) 13 May 2007, (**c**) 29 May 2007, (**d**) 14 June 2007, (**e**) 30 June 2007, (**f**) 1 August 2007, (**g**) 17 August 2007, and (**h**) 2 September 2007.
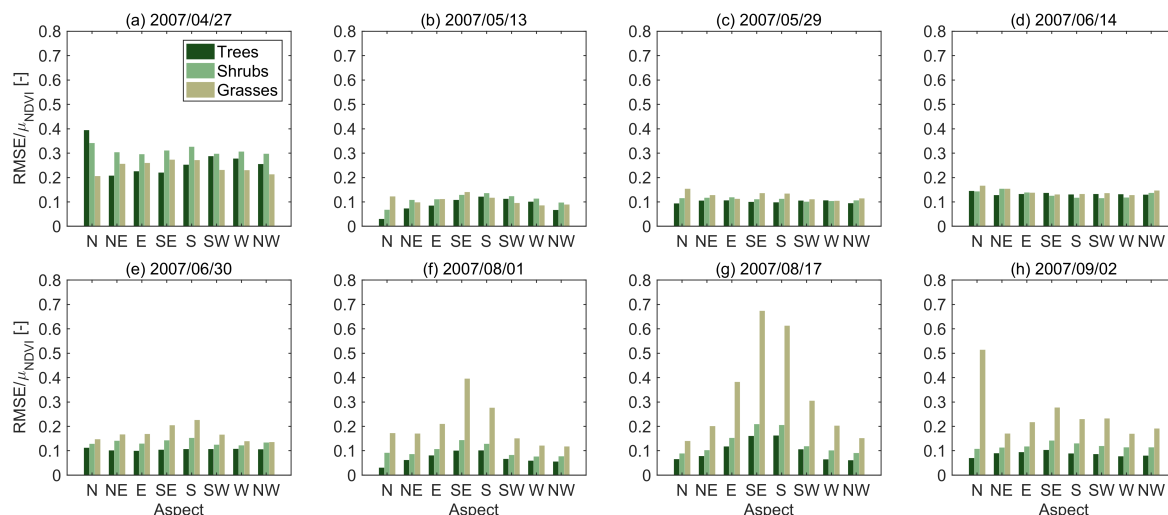
Figure 6. Normalized root mean square error ($RMSE/\mu_{\mathrm{NDVI}}$) separated by dominant vegetation functional type and aspect over the 2007 growing season at Dry Creek Experimental Watershed (DCEW), Idaho, USA for (**a**) 27 April 2007, (**b**) 13 May 2007, (**c**) 29 May 2007, (**d**) 14 June 2007, (**e**) 30 June 2007, (**f**) 1 August 2007, (**g**) 17 August 2007, and (**h**) 2 September 2007. Errors were normalized by dividing RMSE by the mean normalized difference vegetation index (NDVI) for the given vegetation functional type on the given date (see Table 1).

Figure 7. Mean signed error (MSE) separated by dominant vegetation functional type and aspect over the 2007 growing season at Dry Creek Experimental Watershed (DCEW), Idaho, USA for (**a**) 27 April 2007, (**b**) 13 May 2007, (**c**) 29 May 2007, (**d**) 14 June 2007, (**e**) 30 June 2007, (**f**) 1 August 2007, (**g**) 17 August 2007, and (**h**) 2 September 2007.
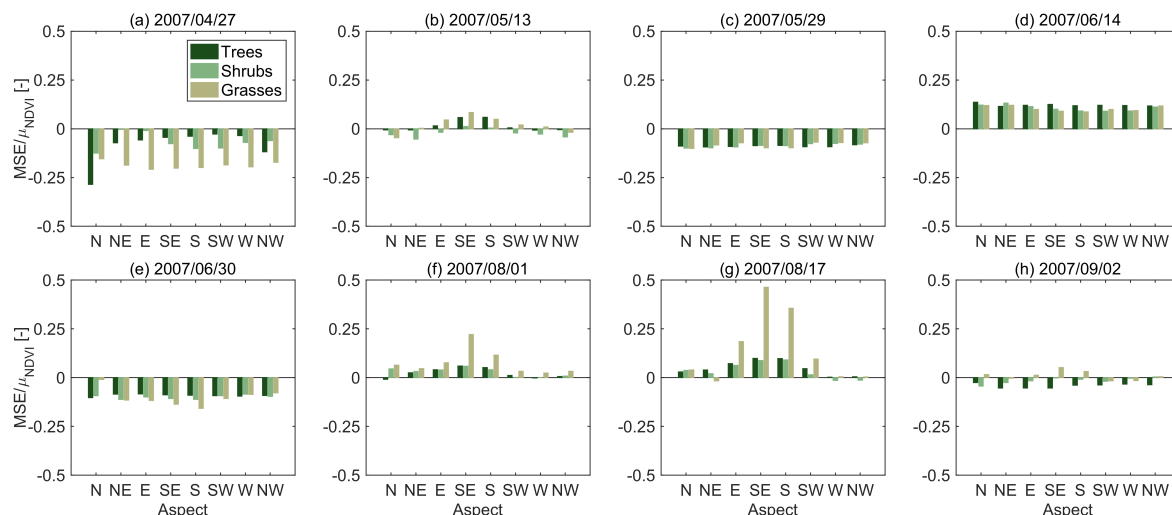
Figure 8. Normalized mean signed error (MSE/μ_{NDVI}) separated by dominant vegetation functional type and aspect over the 2007 growing season at Dry Creek Experimental Watershed (DCEW), Idaho, USA for (**a**) 27 April 2007, (**b**) 13 May 2007, (**c**) 29 May 2007, (**d**) 14 June 2007, (**e**) 30 June 2007, (**f**) 1 August 2007, (**g**) 17 August 2007, and (**h**) 2 September 2007. Errors were normalized by dividing MSE by the mean normalized difference vegetation index (NDVI) for the given vegetation functional type on the given date (see Table 1).

4. Discussion

Comparisons between STARFM predictions of NDVI using two pairs of training data (Table 2, Figure 9) and the null model (Table 4) suggest the Landsat-MODIS data fusions led to information gain for the five scenes from 13 May 2007 to 1 August 2007, where errors from STARFM predictions were lower than the corresponding errors in the null model. This is particularly evident in the May scene, where the null model was associated with an r^2 equal to 0.410 and the STARFM predictions were associated with an r^2 equal to 0.838. In this case STARFM is seemingly able to accurately apportion changes in NDVI captured by MODIS in space during a period characterized by rapid changes in phenology during spring green-up. By contrast, prediction differences in the null model were lower than corresponding differences obtained via the STARFM algorithm for the 27 April 2007, 17 August 2007, and 2 September 2007 scenes. Differences between the null model and the STARFM algorithm were large in April (r^2 of 0.586 and 0.369, respectively), and minor in August (r^2 of 0.982 and 0.943, respectively) and September (r^2 of 0.982 and 0.968, respectively). The large discrepancy in the April scene can be partially attributed to snow covered pixels, which is consistent with similar studies (e.g., [53]). When taking the NSE error metric into consideration, all synthetic images in the two-pair model contained new information NSE > 0, which is consistent with studies that found STARFM can capture phenological timing more precisely, even in areas where there are patterns of relatively fine-grained spatial heterogeneity [38]. It is possible that application of the STARFM algorithm to DCEW (and water-limited, heterogeneous regions more generally) could be improved, especially in the "shoulder months" of the growing season, by using a MODIS product with higher temporal resolution than the 16-day NBAR composite. A study by [29] tested differences in MODIS daily surface reflectance, 8-day composite and 16-day NBAR on STARFM performance in drylands and found that the 16-day NBAR was the optimal imagery for fusion with Landsat-5 TM; however, study observations highlighted the inherent temporal constraints of composite datasets (i.e., 8-day composite of 16-day NBAR) during times of rapid phenological change, such as green-up or senescence.
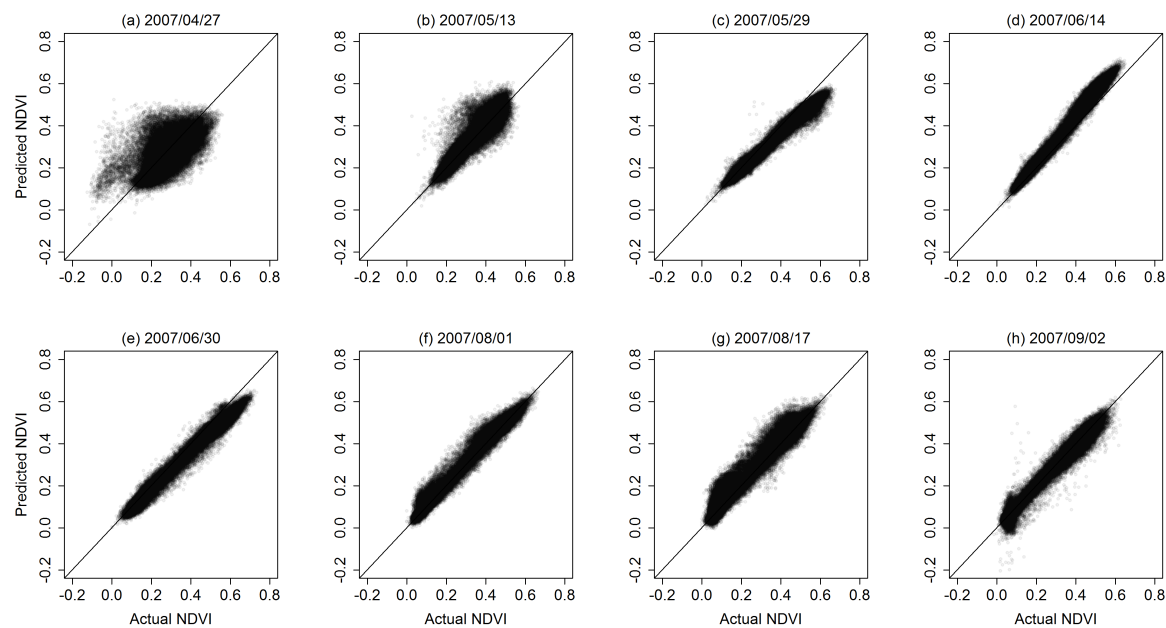
Figure 9. Regression graphs for actual (Landsat) normalized difference vegetation index (NDVI) and STARFM-predicted NDVI over the 2007 growing season at Dry Creek Experimental Watershed (DCEW), Idaho, USA or (**a**) 27 April 2007, (**b**) 13 May 2007, (**c**) 29 May 2007, (**d**) 14 June 2007, (**e**) 30 June 2007, (**f**) 1 August 2007, (**g**) 17 August 2007, and (**h**) 2 September 2007.

With the exception of the April scene, STARFM prediction error was consistently greatest in the NIR band compared to red and green bands. To date, several studies have compared STARFM performance as it relates to differences between predicted and observed values in the NIR and visible bands, with mixed results [29,34,38,53,54]. Studies by [53,54] found that errors in STARFM-derived predictions in the NIR may be attributable to atmospheric contamination at shorter wavelengths, which has been reported to affect the prediction accuracy for other fusion techniques [28]. In the study by [28] the intercept of the relationship between observed and predicted images was positive in all cases, which can be interpreted as a noise signal likely due to atmospheric and BRDF effects. Prediction differences in the study by [54] could also be related to atmospheric influence, as the Landsat images were corrected to top of atmosphere reflectance not apparent reflectance [29]. Similar to our study, both [29] and [38], which were conducted in dryland systems, found lower correlations in the NIR. Reference [38] findings differed from previous studies that attributed better NIR [28,53] and shortwave infrared [55] predictions to greater influence of atmospheric contamination on shorter wavelengths. Importantly, that study was set in semi-arid dryland forests and yielded similar results to [29]. Reference [38] recommend that all synthetic reflectance products be evaluated prior to use as the influence of site specific conditions might be variable across the visible and NIR portions of the spectrum. Using a modified version of the STARFM algorithm, reference [34] found that the ESTARFM algorithm performed slightly better than STARFM in a region of relatively homogeneous vegetation, but significantly better in a region of heterogeneous vegetation.

The above discussion reveals the difficulty in comparing across studies using STARFM due to the complexities in the region being studied, the data sets being fused, and the details of the satellite platforms from which those data are collected. In our study domain, absolute errors in STARFM predictions of NDVI tended to be largest in the tree cover type, followed by shrubs, and then grasses. At the same time, however, this conclusion does not hold in its entirety when considering the prediction errors scaled to average NDVI (μ_{NDVI}). The relative error (RMSE in NDVI scaled by μ_{NDVI}) tends to be largest in the grass cover type, and this impact is greatest in mid-August, when RMSE in the grass regions is approximately 40% of the average NDVI. This suggests that a degree of caution

be exercised in applying STARFM, and potentially other data fusion algorithms, in savanna-like ecosystems with significant grass coverage. This result suggests that particular attention is warranted in grass-dominated regions within heterogeneous ecosystems where errors in STARFM-predicted NDVI may be small relative to other regions within a study area in an absolute sense, but significant relative to the average NDVI of the grass-dominated region.

Correspondingly, there do not appear to be significant, generalizable conclusions about the role of topographic aspect in STARFM-derived predictions of NDVI across all vegetation functional types and time periods. Related to the above discussion of the magnitude of higher relative error in grass-dominated regions, this effect is particularly pronounced in south- and southeast-facing hillslopes in DCEW. We caution against over-interpreting this conclusion, however, because south- and southeast-facing slopes found in the middle elevations of DCEW tend to be dominated by grasses. Hence, the observed trend in Figure 5 seems to largely echo the previously discussed conclusions about the role of vegetation functional type as a predictor of STARFM performance. Interestingly, though, when we look at the magnitude of the relative error in NDVI in grass-covered regions we find that south- and southeast-facing pixels demonstrate an increase in relative error in the 1 August 2007 prediction. Relative error in NDVI in grass-dominated regions is again larger in the 17 August 2007 prediction on south- and southeast-facing pixels, but there is also an increase in relative error in east- and southwest-facing pixels relative to the previous 1 August 2007 prediction. By the 2 September 2007 prediction, relative error has decreased on east-, southeast-, south-, and southwest-facing pixels, but risen significantly on north-facing pixels in grass-dominated regions. Thus, topographic aspect may not play as large an influence as vegetation type in defining the magnitude of STARFM prediction errors in water-limited ecosystems, but it may impact the timing of errors when STARFM is being used to synthesize a time series of images.

Results from this study have important implications for the use of STARFM and other data fusion algorithms for creating value-added spatiotemporal vegetation remote sensing datasets that may be used to inform land models. These implications are related to the previous points we have discussed above. Specifically, results suggest that STARFM and other data fusion algorithms can serve to compensate for the tradeoffs between spatial and temporal resolution common to many remote sensing platforms used to characterize terrestrial vegetation, even in water-limited ecosystems with significant spatial heterogeneity in vegetation functional types. This suggests that it is possible to create historical reconstructions of spatiotemporal variation in variables like NDVI, which provide an important window into local dynamics of water, energy, and biogeochemical cycling, with reasonable confidence, over extended periods of time, and in relatively high spatial and temporal detail. When STARFM, specifically, can be applied in a way that makes use of Landsat images bracketing the dates of interest, it is possible to create accurate imputations of NDVI and other variables at the spatial resolution of Landsat and temporal resolution of MODIS. Historical reconstructions with these characteristics would be of significant value for constraining key land model parameters related to the vegetation canopy (e.g., LAI, albedo, etc.). Hence, data fusion algorithms like STARFM may have an important role to play in development, application, calibration, and verification of land models in water-limited ecosystems.

Along with previous studies, this work highlights the importance of continuity in remote sensing datasets with complementary characteristics. Algorithms like STARFM are useful only because of the simultaneous existence and availability of Landsat and MODIS data. STARFM and algorithms like it are sufficiently generic that other complementary pairs of remote sensing platforms could be used to develop historical reconstructions of spatiotemporal vegetation characteristics. While some studies have already begun to explore application of STARFM and other data fusion algorithms to constrain other satellite datasets to each other, additional work on this topic is warranted (e.g., [2,56–59]). Landsat and the Advanced High-Resolution Radiometer (AVHRR) have an even longer period of operational overlap. In addition, although AVHRR is associated with a much coarser resolution than MODIS, it may be worthwhile to explore the performance of AVHRR-Landsat data fusion through STARFM in a variety of ecosystems of interest. The Visible Infrared Imaging Radiometer Suite (VIIRS) is a potential

asset with observational characteristics similar to MODIS and complementary to Landsat. With the potential to derive fine spatial and temporal scales, data fusion of Landsat and Sentinel-2 could provide significant advances in monitoring the phenophases of dryland ecosystems.

5. Conclusions

In this study, we investigated (1) how accurately the STARFM algorithm can be used to synthesize reflectance measurements from MODIS and Landsat within a semi-arid watershed with complex spatial heterogeneity. (2) We evaluated how STARFM prediction error varied with (a) the rate of phenological change in a growing season, (b) vegetation functional type (c) topographic aspect, and (d) the presence of seasonal snowpacks. Specifically, STARFM was used to synthesize images of NDVI as well as reflectance in the green, red, and NIR bands using MODIS NBAR and Landsat images bracketing the dates when we applied the STARFM algorithm. We repeated this application of STARFM for eight dates during the 2007 growing season 27 April 2007–2 September 2007. We found that the images synthesized using STARFM yielded accurate predictions of NDVI, when compared with a null model consisting of the temporally nearest Landsat observation, for five of the eight dates. Two of the remaining dates were associated with negligible differences in accuracy between the STARFM and null model predictions of NDVI, while the null model exhibited superior performance on the remaining date (27 April 2007). The presence/absence of snow was a significant factor on the performance of STARFM on that date. Absolute errors in STARFM-derived predictions of NDVI are highest in regions covered by trees, while relative errors are highest in grass-dominated regions of the study watershed. Topographic aspect was not a strong predictor of absolute errors across all vegetation types, although errors in grass-dominated regions on south- and southeast-facing slopes tended to be highest in the late-growing season. This study suggests that data fusion algorithms like STARFM can, when carefully used, be applied to create historical reconstruction of the spatiotemporal dynamics of key vegetation characteristics like NDVI in even highly heterogeneous shrub-dominated systems. Future applications of STARFM in DCEW using Landsat and MODIS satellite pairs consider refining results by using MODIS daily reflectance products during periods of rapid greenup rather than the 16-day composite products that were evaluated in this investigation. This study uniquely demonstrates that availability of LiDAR datasets in a study region can be used to examine how errors in the STARFM downscaling algorithm vary by dominant vegetation type. Because the spatiotemporal patterns of NDVI provide a reflection of the associated patterns in water, energy, and biogeochemical cycling, the ability to develop similar reconstructions is important from the perspective of ecohydrologic process study and modeling.

Author Contributions: A.N.F. and N.F.G. conceived and designed the numerical experiments; P.J.O. performed the numerical experiments; P.J.O. and J.M analyzed the output of the STARFM algorithm; P.J.O, J.M., N.F.G., and A.N.F. wrote the paper.

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

References

1. Poulter, B.; Frank, D.; Ciais, P.; Myneni, R.B.; Andela, N.; Bi, J.; Broquet, G.; Canadell, J.G.; Chevallier, F.; Liu, Y.Y.; et al. Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature* **2014**, *509*, 600–603.
2. Huang, J.; Ji, M.; Xie, Y.; Wang, S.; He, Y.; Ran, J. Global semi-arid climate change over last 60 years. *Clim. Dyn.* **2016**, *46*, 1131–1150.
3. Davies, J.; Poulsen, L.; Schulte-Herbrüggen, B.; MacKinnon, K.; Crawhall, N.; Henwood, W.D.; Dudley, N.; Smith, J.; Gudka, M. *Conserving Dryland Biodiversity*; IUCN: Nairobi, Kenya, 2012.

4.	Balch, J.K.; Bradley, B.A.; D'Antonio, C.M.; Gómez-Dans, J. Introduced annual grass increases regional fire activity across the arid western USA (1980–2009). *Glob. Chang. Biol.* **2013**, *19*, 173–183.

5.	Adler, P.; Raff, D.; Lauenroth, W. The effect of grazing on the spatial heterogeneity of vegetation. *Oecologia* **2001**, *128*, 465–479.

6.	Van de Koppel, J.; Rietkerk, M.; van Langevelde, F.; Kumar, L.; Klausmeier, C.A.; Fryxell, J.M.; Hearne, J.W.; van Andel, J.; de Ridder, N.; Skidmore, A.; et al. Spatial heterogeneity and irreversible vegetation change in semiarid grazing systems. *Am. Nat.* **2002**, *159*, 209–218.

7.	Backlund, P.; Janetos, A.C.; Schimel, D.S. *The Effects of Climate Change on Agriculture, Land Resources, Water Resources, and Biodiversity in the United States*; Nova Science Pub Incorporated: New York, NY, USA, 2008.

8.	Puigdefábregas, J. The role of vegetation patterns in structuring runoff and sediment fluxes in drylands. *Earth Surf. Process. Landf.* **2005**, *30*, 133–147.

9.	Foti, R.; Ramírez, J.A. A mechanistic description of the formation and evolution of vegetation patterns. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 63–84.

10.	Zou, C.B.; Barron-Gafford, G.A.; Breshears, D.D. Effects of topography and woody plant canopy cover on near-ground solar radiation: Relevant energy inputs for ecohydrology and hydropedology. *Geophys. Res. Lett.* **2007**, *34*, doi:10.1029/2007GL031484.

11.	Schlaepfer, D.R.; Lauenroth, W.K.; Bradford, J.B. Ecohydrological niche of sagebrush ecosystems. *Ecohydrology* **2012**, *5*, 453–466.

12.	Asner, G.P.; Levick, S.R.; Kennedy-Bowdoin, T.; Knapp, D.E.; Emerson, R.; Jacobson, J.; Colgan, M.S.; Martin, R.E. Large-scale impacts of herbivores on the structural diversity of African savannas. *Proc. Nat. Acad. Sci. USA* **2009**, *106*, 4947–4952.

13.	D'Odorico, P.; Okin, G.S.; Bestelmeyer, B.T. A synthetic review of feedbacks and drivers of shrub encroachment in arid grasslands. *Ecohydrology* **2012**, *5*, 520–530.

14.	Kurz, W.A.; Dymond, C.C.; Stinson, G.; Rampley, G.J.; Neilson, E.T.; Carroll, A.L.; Ebata, T.; Safranyik, L. Mountain pine beetle and forest carbon feedback to climate change. *Nature* **2008**, *452*, 987–990.

15.	Stark, S.C.; Breshears, D.D.; Garcia, E.S.; Law, D.J.; Minor, D.M.; Saleska, S.R.; Swann, A.L.S.; Villegas, J.C.; Aragão, L.E.O.C.; Bella, E.M.; et al. Toward accounting for ecoclimate teleconnections: Intra- and inter-continental consequences of altered energy balance after vegetation change. *Landsc. Ecol.* **2016**, *31*, 181–194.

16.	Moorcroft, P.R.; Hurtt, G.C.; Pacala, S.W. A method for scaling vegetation dynamics: The ecosystem demography model (ED). *Ecol. Monogr.* **2001**, *71*, 557–585.

17.	Best, M.J.; Pryor, M.; Clark, D.B.; Rooney, G.G.; Essery, R.L.H.; Ménard, C.B.; Edwards, J.M.; Hendry, M.A.; Porson, A.; Gedney, N.; et al. The Joint UK Land Environment Simulator (JULES), model description—Part 1: Energy and water fluxes. *Geosci. Model Dev.* **2011**, *4*, 677–699.

18.	Lawrence, D.M.; Oleson, K.W.; Flanner, M.G.; Thornton, P.E.; Swenson, S.C.; Lawrence, P.J.; Zeng, X.; Yang, Z.L.; Levis, S.; Sakaguchi, K.; et al. Parameterization improvements and functional and structural advances in Version 4 of the Community Land Model. *J. Adv. Model. Earth Syst.* **2011**, *3*, doi:10.1029/2011MS00045.

19.	Niu, G.Y.; Yang, Z.L.; Mitchell, K.E.; Chen, F.; Ek, M.B.; Barlage, M.; Kumar, A.; Manning, K.; Niyogi, D.; Rosero, E.; et al. The community Noah land surface model with multiparameterization options (Noah-MP): 1. Model description and evaluation with local-scale measurements. *J. Geophys. Res. Biogeosci.* **2011**, *116*, doi:10.1029/2010JD015140.

20.	Rodell, M.; Houser, P.R.; Jambor, U.; Gottschalck, J.; Mitchell, K.; Meng, C.J.; Arsenault, K.; Cosgrove, B.; Radakovich, J.; Bosilovich, M.; et al. The Global Land Data Assimilation System. *Bull. Am. Meteorol. Soc.* **2004**, *85*, 381–394.

21.	Medvigy, D.; Wofsy, S.C.; Munger, J.W.; Hollinger, D.Y.; Moorcroft, P.R. Mechanistic scaling of ecosystem function and dynamics in space and time: Ecosystem Demography model version 2. *J. Geophys. Res. Biogeosci.* **2009**, *114*, doi:10.1029/2008JG000812.

22.	Tague, C.; McMichael, C.; Hope, A.; Choate, J.; Clark, R. Application of the RHESSys model to a California semiarid shrubland watershed. *JAWRA J. Am. Water Resour. Assoc.* **2004**, *40*, 575–589.

23.	Ivanov, V.Y.; Bras, R.L.; Vivoni, E.R. Vegetation-hydrology dynamics in complex terrain of semiarid areas: 1. A mechanistic approach to modeling dynamic feedbacks. *Water Resour. Res.* **2008**, *44*, doi:10.1029/2006WR005588.

24.	Wulder, M.A.; Masek, J.G.; Cohen, W.B.; Loveland, T.R.; Woodcock, C.E. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sens. Environ.* **2012**, *122*, 2–10.

25. Huete, A.; Didan, K.; Miura, T.; Rodriguez, E.P.; Gao, X.; Ferreira, L.G. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* **2002**, *83*, 195–213.

26. Myneni, R.B.; Hoffman, S.; Knyazikhin, Y.; Privette, J.L.; Glassy, J.; Tian, Y.; Wang, Y.; Song, X.; Zhang, Y.; Smith, G.R.; et al. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sens. Environ.* **2002**, *83*, 214–231.

27. 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.

28. Roy, D.P.; Ju, J.; Lewis, P.; Schaaf, C.; Gao, F.; Hansen, M.; Lindquist, E. Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sens. Environ.* **2008**, *112*, 3112–3130.

29. Walker, J.J.; de Beurs, K.M.; Wynne, R.H.; Gao, F. Evaluation of Landsat and MODIS data fusion products for analysis of dryland forest phenology. *Remote Sens. Environ.* **2012**, *117*, 381–393.

30. Wang, Q.; Zhang, Y.; Onojeghuo, A.O.; Zhu, X.; Atkinson, P.M. Enhancing Spatio-Temporal Fusion of MODIS and Landsat Data by Incorporating 250 m MODIS data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, doi:10.1109/JSTARS.2017.2701643.

31. Bradley, B.A.; Mustard, J.F. Characterizing the landscape dynamics of an invasive plant and risk of invasion using remote sensing. *Ecol. Appl.* **2006**, *16*, 1132–1147.

32. Okin, G.S.; Roberts, D.A.; Murray, B.; Okin, W.J. Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sens. Environ.* **2001**, *77*, 212–225.

33. Semmens, K.A.; Anderson, M.C.; Kustas, W.P.; Gao, F.; Alfieri, J.G.; McKee, L.; Prueger, J.H.; Hain, C.R.; Cammalleri, C.; Yang, Y.; et al. Monitoring daily evapotranspiration over two California vineyards using Landsat 8 in a multi-sensor data fusion approach. *Remote Sens. Environ.* **2016**, *185*, 155–170.

34. 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.

35. Fu, D.; Chen, B.; Wang, J.; Zhu, X.; Hilker, T. An improved image fusion approach based on enhanced spatial and temporal the adaptive reflectance fusion model. *Remote Sens.* **2013**, *5*, 6346–6360.

36. 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.

37. Schmidt, M.; Udelhoven, T.; Gill, T.; Röder, A. Long term data fusion for a dense time series analysis with MODIS and Landsat imagery in an Australian Savanna. *J. Appl. Remote Sens.* **2012**, *6*, doi:10.1117/1.JRS.6.063512.

38. Olexa, E.M.; Lawrence, R.L. Performance and effects of land cover type on synthetic surface reflectance data and NDVI estimates for assessment and monitoring of semi-arid rangeland. *Int. J. Appl. Earth Obs. Geoinf.* **2014**, *30*, 30–41.

39. Yenko, M.K. Hydrometric and Geochemical Evidence of Streamflow Sources in the Upper Dry Creek Experimental Watershed, Southwestern Idaho, Ph.D. Thesis, Boise State University, Boise, ID, USA, 2003.

40. Masek, J.G.; Vermote, E.F.; Saleous, N.E. A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 68–72.

41. Streutker, D.R.; Glenn, N.F. LiDAR measurement of sagebrush steppe vegetation heights. *Remote Sens. Environ.* **2006**, *102*, 135–145.

42. McCuen, R.H.; Knight, Z.; Cutter, A.G. Evaluation of the Nash–Sutcliffe efficiency index. *J. Hydrol. Eng.* **2006**, *11*, 597–602.

43. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE* **2007**, *50*, 885–900.

44. Jain, S.K.; Sudheer, K.P. Fitting of hydrologic models: A close look at the Nash–Sutcliffe index. *J. Hydrol. Eng.* **2008**, *13*, 981–986.

45. Hall, D.K.; Frei, A.; Drey, S.J. *Remote Sensing of Snow Cover. Section: Snow Extent*; NASA: Washington, DC, USA, 2012.

46. Geroy, I.J.; Gribb, M.M.; Marshall, H.P.; Chandler, D.G.; Benner, S.G.; McNamara, J.P. Aspect influences on soil water retention and storage. *Hydrol. Process.* **2011**, *25*, 3836–3842.

47. Kunkel, M.L.; Flores, A.N.; Smith, T.J.; McNamara, J.P. A simplified approach for estimating soil carbon and nitrogen stocks in semi-arid complex terrain. *Geoderma* **2011**, *165*, 1–11.

48. Smith, T.J.; McNamara, J.P.; Flores, A.N.; Gribb, M.M.; Aishlin, P.S.; Benner, S.G. Small soil storage capacity limits benefit of winter snowpack to upland vegetation. *Hydrol. Process.* **2011**, *25*, 3858–3865.

49. Yetemen, O.; Istanbulluoglu, E.; Vivoni, E.R. The implications of geology, soils, and vegetation on landscape morphology: Inferences from semi-arid basins with complex vegetation patterns in Central New Mexico, USA. *Geomorphology* **2010**, *116*, 246–263.

50. Gutiérrez-Jurado, H.A.; Vivoni, E.R.; Cikoski, C.; Harrison, J.B.J.; Bras, R.L.; Istanbulluoglu, E. On the observed ecohydrologic dynamics of a semiarid basin with aspect-delimited ecosystems. *Water Resour. Res.* **2013**, *49*, 8263–8284.

51. Pelletier, J.D.; Barron-Gafford, G.A.; Breshears, D.D.; Brooks, P.D.; Chorover, J.; Durcik, M.; Harman, C.J.; Huxman, T.E.; Lohse, K.A.; Lybrand, R.; et al. Coevolution of nonlinear trends in vegetation, soils, and topography with elevation and slope aspect: A case study in the sky islands of southern Arizona. *J. Geophys. Res. Earth Surf.* **2013**, *118*, 741–758.

52. Zhou, X.; Istanbulluoglu, E.; Vivoni, E.R. Modeling the ecohydrological role of aspect-controlled radiation on tree-grass-shrub coexistence in a semiarid climate. *Water Resour. Res.* **2013**, *49*, 2872–2895.

53. Hilker, T.; Wulder, M.A.; Coops, N.C.; Seitz, N.; White, J.C.; Gao, F.; Masek, J.G.; Stenhouse, G. Generation of dense time series synthetic Landsat data through data blending with MODIS using a spatial and temporal adaptive reflectance fusion model. *Remote Sens. Environ.* **2009**, *113*, 1988–1999.

54. Watts, J.D.; Powell, S.L.; Lawrence, R.L.; Hilker, T. Improved classification of conservation tillage adoption using high temporal and synthetic satellite imagery. *Remote Sens. Environ.* **2011**, *115*, 66–75.

55. Liu, H.; Weng, Q. Enhancing temporal resolution of satellite imagery for public health studies: A case study of West Nile Virus outbreak in Los Angeles in 2007. *Remote Sens. Environ.* **2012**, *117*, 57–71.

56. Tewes, A.; Thonfeld, F.; Schmidt, M.; Oomen, R.J.; Zhu, X.; Dubovyk, O.; Menz, G.; Schellberg, J. Using RapidEye and MODIS data fusion to monitor vegetation dynamics in semi-arid rangelands in South Africa. *Remote Sens.* **2015**, *7*, 6510–6534.

57. Wu, M.; Li, H.; Huang, W.; Niu, Z.; Wang, C. Generating daily high spatial land surface temperatures by combining ASTER and MODIS land surface temperature products for environmental process monitoring. *Environ. Sci. Process. Impacts* **2015**, *17*, 1396–1404.

58. Zheng, Y.; Zhang, M.; Zhang, X.; Zeng, H.; Wu, B. Mapping winter wheat biomass and yield using time series data blended from PROBA-V 100- and 300-m S1 products. *Remote Sens.* **2016**, *8*, 824.

59. Wang, Q.; Blackburn, G.A.; Onojeghuo, A.O.; Dash, J.; Zhou, L.; Zhang, Y.; Atkinson, P.M. Fusion of Landsat 8 OLI and Sentinel-2 MSI data. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 3885–3899.