Landscape Ecology, in press

Linear Downscaling from MODIS to Landsat: Connecting Landscape Composition with Ecosystem Functions

Jiquan Chen[1, 2]*, Pietro Sciusco[1, 2], Zutao Ouyang[1, 2], Rong Zhang[2], Geoffrey M. Henebry[1, 2], Ranjeet John[3], and David. P. Roy[1, 2]

[1] Department of Geography, Environment, and Spatial Sciences, Michigan State University, MI 48824, USA

[2] Center for Global Change and Earth Observations, Michigan State University, MI 48823, USA

[3.] Department of Biology, University of South Dakota, Vermillion, SD 57069, USA

Corresponding Author: Dr. Jiquan Chen, Email: jqchen@msu.edu; Phone: 517-884-1884

Date of manuscript revision: 10/9/2019

Abstract

Context

The open and free access to Landsat and MODIS products have greatly promoted scientific investigations on spatiotemporal change in land mosaics and ecosystem functions at landscape to regional scales. Unfortunately, there is a major mismatch in spatial resolution between MODIS products at coarser resolution (≥250 m) and landscape structure based on classified Landsat scenes at finer resolution (30 m).

Objectives

Based on practical needs for downscaling popular MODIS products at 500 m resolution to match classified land cover at Landsat 30 m resolution, we proposed an innovative modelling approach so that landscape structure and ecosystem functions can be directly studied for their interconnections. As a proof-of-concept of our downscaling approach, we selected the watershed of the Kalamazoo River in southwestern Michigan, USA as the testbed.

Methods

MODIS products for three fundamental variables of ecosystem function are downscaled to ensure the approach can be extrapolated to multiple functional measurements. They are blue-sky albedo (0-1), evapotranspiration (ET, mm), and gross primary production (GPP, Mg C ha^{-1} yr^{-1}). An object-oriented classification of Landsat images in 2011 was processed to generate a land cover map indicating landscape structure. The downscaling model was tested for the five Level IV ecoregions within the watershed.

Results

We achieved satisfactory downscaling models for albedo, ET, and GPP for all five ecoregions. The adjusted R^2 was >0.995 for albedo, 0.915-0.997 for ET, and 0.902-0.962 for GPP. The estimated

albedo, ET, and GPP values appear different in the region. The estimated albedo was the lowest for water (0.076-0.107) and the highest for cropland (0.166-0.172). Estimated ET was the highest for the built-up cover type (525.6-687.1 mm) and the lowest for forest (209.7-459.7 mm). The estimated GPP was the highest for the build-up cover type (8.65-9.85 MgC ha^{-1} yr^{-1}) and the lowest for forest.

Conclusions

Estimated values for albedo, ET, and GPP appear reasonable for their ranges in the Kalamazoo River region and are consistent with values reported in the literature. Despite these promising results, the downscaling approach relies on strong assumptions and can carry substantial uncertainty. It is only valid at a spatial scale where similar climate, soil, and landforms exist (i.e., values in isolated patches of the same cover type are similar). Plausibly, the uncertainties associated with each estimation, as well as the model residuals, can be explored for other pattern-process relationships within the landscape.

Keywords: *Downscaling, MODIS, Landsat, GPP, ET, Albedo, Kalamazoo River watershed*

1. Introduction

The last two decades have witnessed a rapid increase in the application of remote sensing imagery in various fields of natural science and resource management, especially using data from the Landsat satellite series and from the MODIS (Moderate Resolution Imaging Spectroradiometer) sensors on board NASA's Terra and Aqua satellites. These images and derived value-added products provide the scientific community, practitioners, and natural resource managers with powerful tools for quantifying spatial and temporal changes in land surface properties. Landsat data, for example, have been widely used to create regional to global land cover maps and to describe the landscape structure, ecosystem dynamics, land use change, and their connections with landscape processes (e.g., fragmentation, succession) (Bresee et al. 2004; Hansen and Loveland 2012; Kennedy et al. 2014; Wulder et al. 2018). Systematically derived, quality assessed, validated, and reprocessed NASA MODIS products (Justice et al. 2002) meanwhile have been widely used to create global products of gross ecosystem production (GPP) (Zhao et al. 2005), land cover (Friedl et al. 2010), evapotranspiration (ET; Mu et al. 2007), land surface temperature (Wan 2013), albedo (Schaaf et al. 2002), disturbance (e.g., fire; Giglio et al. 2018), and land surface phenology (Zhang et al. 2003; Moon et al. 2019). The open and free access to both Landsat and MODIS products further promoted their use and influence in scientific investigations. Explorations of landscape structure and dynamics based on the Landsat archive, as another example, have seen a substantial increase in applications (Bresee et al. 2004; Roy et al. 2014; Krehbiel et al. 2017; Wulder et al. 2019).

Connecting ecosystem functions (e.g., ecosystem production, biological diversity) with landscape structure (e.g., patch type, characteristics and spatial configuration) and processes (e.g., fragmentation, movement of species and materials across the landscape) has long been an interest of landscape ecologists (Franklin and Forman 1987; Forman 1995; Wu and Hobbs 2007; Turner and Gardner 2015). As an example, John et al. (2016) examined the long-term changes of vegetation indices from MODIS and AVHRR (Advanced Very High Resolution Radiometer), ecosystem production, livestock, and land cover and land use changes on the Mongolian Plateau to tease apart the contributions of climatic change and human influences, which is a scientific frontier in global change science (Alberti et al. 2011; Chen et al. 2013; Mills et al. 2017). Their studies, along with many other similar ones, were often conducted at coarser spatial resolutions of 500-1000 m (i.e., the pixel size of sensor data such as MODIS and AVHRR). Meanwhile, Landsat images have been widely used to produce multiple versions of accurate land cover maps at 30 m resolution that better represents cover types (patches) and their spatial configurations at landscape to regional scales (Wulder et al. 2018). Higher spatial resolution is desirable because many patches in a landscape may be substantially smaller than the finest resolution of a MODIS pixel (250 m). Recently, the Landsat data has been complemented by twin satellites from the European Space Agency (Sentinel-2A and 2B) that provide Landsat-like imagery but at finer (10-20 m) spatial resolution (Drusch et al. 2012). Clearly, there exists a major mismatch in the spatial resolution between MODIS products at coarser resolution and landscape structure from Landsat scenes at finer resolution.

Downscaling to finer spatiotemporal resolution has been a research focus in remote sensing (see Atkinson 2013 for an overview). Both mechanistically and empirically based approaches have been attempted in landscape studies. One approach is to quantify ecosystem functions directly at Landsat resolution by integrating ecosystem models with patches (e.g., Zheng et al. 2004; Gitelson et al. 2012; Semmens et al. 2016; Yao et al. 2017). Another approach is to downscale coarse resolution data through spatial interpolation, such as block cokriging where the landscape is categorized into multiple classes (i.e., blocks). However, the first approach requires extensive and detailed measurements at near surface and ground levels for model parameterization and validation; whereas, the products from the second approach are often not matched with the patch boundaries and require an assumption of statistical stationarity rarely met in landscapes (Saunders et al. 2005). An alternative is to take the advantage of the available, coarser spatial resolution MODIS products and downscale them to the finer spatial resolution of Landsat imagery through dasymetric modeling, an approach to thematic mapping that uses land cover information as ancillary information and spatially disaggregates coarser resolution information (Fisher and Langford 1996; Petrov 2012). Dasymetric modeling has been widely applied in human geography (Nagle et al. 2014; Jia and Gaughan 2016), but not in landscape studies. Here we applied the principles of dasymetric modeling to estimate higher resolution (30 m) ecosystem functions from MODIS products by treating finer spatial resolution land cover data (i.e., the ancillary layer) that are accurately classified from Landsat imagery.

The idea is simple. Consider the value of gross primary productivity (GPP) defined by a MODIS pixel to be the linear sum of GPP_i of land cover type (i), expressed as (Fig. 1):

$$GPP = \sum (\kappa_i \times GPP_i) + \varepsilon \qquad [1]$$

where κ_i is the compositional portion (0-1) of land cover type *i* within a MODIS pixel where $\sum (\kappa_i) = 1$, and ε is the model residual errors to include uncertainties from both the land cover classification and the MODIS estimates of GPP. Through multiple linear regression constrained to a zero intercept, GPP_i can be estimated empirically for a landscape composed of multiple MODIS pixels and land cover types. The zero intercept constraint is necessary to ensure that the mathematical weight of each land cover type is allocated properly. A key assumption of this downscaling approach is that ecosystem functions of the same land cover type remain the same in different MODIS pixels. With this downscaling model, the uncertainty of GPP_i estimates from Eq. 1 (i.e., the standard error associated with GPP_i estimate in Eq. 1), as well as the residuals (**ε**) can be examined for their relationships with other landscape characteristics (e.g., interactions between patches, landscape composition, etc.) for further investigation of pattern-process relationships.

As a proof-of-concept of our downscaling approach, we selected the watershed of the Kalamazoo River in the southwestern Michigan as a testbed. MODIS products for three fundamental variables of ecosystem function—albedo, ET, and GPP—were downscaled to ensure the approach can be extrapolated to multiple functional measurements. To simplify our demonstration, we focused on the year of 2011 when we developed an accurate, Landsat-derived land cover map for the watershed. To meet the model assumptions, we based our model at the scale of a single Level IV ecoregion, within which climate, soils, and potential vegetation are similar (Omernik and Griffith 2014). Our modeling effort was then independently repeated for each of the five Level IV ecoregions within the watershed to demonstrate the validity of the approach.

2. Methods

2.1. Study area

The Kalamazoo River watershed (5261 km^2) includes portions of 10 counties (Allegan, Ottawa, Van Buren, Kent, Barry, Kalamazoo, Calhoun, Eaton, Jackson, and Hillsdale) in southwestern Michigan, USA. The watershed is dominated by cultivated crops, deciduous forest, mixed prairies, lakes and ponds, wooded wetlands, and urban areas. Prior to European settlement and land conversion to agriculture, this region had a mosaic of tallgrass prairies, savannas, and oak openings, including both C_3 and C_4 grasses as well as forbs (Chapman and Brewer 2008). Kalamazoo and Battle Creek are the two major urban centers within the watershed, and the Kalamazoo–Battle Creek–Portage Combined Statistical Area had a population of more than 524,000 in 2010.

The Kalamazoo River has a moderate stream gradient and drops 165 m in elevation from its headwaters, which are >300 m a.s.l. The River drains a landscape consisting of thick glacial deposits. Alfisols are the most common soil formation and reflect the dominance of deciduous forests in the past. The watershed is covered with prime agricultural soils, including 70% coarse soils that are permeable to rainwater and help in the recharge of groundwater (Schaetzl et al. 2009). The annual average precipitation is 890 mm, with ~65% returned as evapotranspiration (Fongers 2008). The recent growth in urban area can cause the local hydrology to become more “flashy”, owing to the sharp increase in impervious surface areas and could contribute to higher phosphorus loading from urban land uses (Bass 2009).

While 96% of the land is privately owned, there are major inclusions of public lands (e.g., Allegan State Game area, Fort Custer Recreation Area, Yankee Springs Recreation Area). The land use history in the watershed shows that row crop agriculture takes place in the more productive soils; whereas, croplands have been abandoned in marginal areas including land with steep slopes, lands that were excessively drained, or that had poor drainage (Schaetzl et al. 2009). Five Level IV ecoregions of the United States Environmental Protection Agency fall within the watershed (Omernik and Griffith 2014) (Fig. 2): Battle Creek Outwash Plain (56b); Michigan Lake Plain (56d); Lake Michigan Moraines (56f); Lansing Loamy Plain (56g); and Interlobate Dead Ice Moraines (56h). Detailed descriptions of these ecoregions can be found online at EPA: https://www.epa.gov/eco-research/ecoregion-download-files-state-region-5#pane-20.

2.2 Data sources and pre-processing

We used MODIS products with Landsat images to demonstrate the broader applications of the downscaling model (Eq. 1). Three MODIS products were used: blue-sky albedo (0-1), evapotranspiration (ET, mm), and gross primary production (GPP, Mg C ha^{-1} yr^{-1}). Ecosystem water use efficiency (WUE, g kg^{-1}), defined as the ratio between GPP and ET, was further calculated with estimated GPP and ET values by land cover type to expand from the downscaled predictions to other potential ecological applications. It is worth noting that ET in units of mm is converted to kg ha^{-1} yr^{-1} based on water density of 1.0 Mg m^{-3}.

We used Google Earth Engine (GEE) to process the most recent Collection 6 MODIS ET and GPP products. ET and GPP were provided by the MODIS 500m 8 day Evapotranspiration/Latent Heat Flux (MOD16A2-V6) (https://doi.org/10.5067/MODIS/MOD16A2.006) and Gross Primary Productivity (MOD17A2H-V6) (https://doi.org/10.5067/MODIS/MOD17A2H.006) products. We filtered the ET and GPP data using the MODIS product quality bands “NPP_QC” for ET, and “PSN_QC” for GPP, to ensure only the high quality (QC=0) data were retained. For albedo, we used the blue-sky albedo product available as a global annual mean for 2000-2015 (http://rslab.gr/downloads_blue_sky.html). The blue-sky albedo was estimated using the MCD43A1 product (https://doi.org/10.5067/MODIS/MCD43A1.006). This MODIS product was generated daily at 500 m by inversion of a Bidirectional Reflectance Distribution Function (BRDF) model against 16-day moving window of MODIS 500m observations and then the BRDF model is used to derive the black-sky and white-sky albedos (Wang et al. 2018). The blue-sky albedo was estimated every 8 days as a weighted average of both black- and white-sky albedos and using the MODIS aerosol Optical Thickness product (Chrysoulakis et al. 2018).

The three MODIS products defined from January 1st to Dec 31st 2011 were used to calculate annual mean blue-sky albedo (ranging 0 to 1), total GPP (Mg ha^{-1} yr^{-1}) and total annual ET (mm) at each 500 m MODIS pixel. The results were clipped spatially to the extent of the Kalamazoo River watershed and Level IV ecoregion boundaries. Zonal statistical analyses were performed in ArcMap 10.6 to extract values of albedo, ET, and GPP data by ecoregion and by each MODIS pixel into data tables. The extracted data tables were then analyzed in RStudio (RStudio 1.1.453, https://www.rstudio.com/) to compute mean, standard deviation, and probability densities.

A land cover map of the watershed was created using an object-oriented classification of Landsat Thematic Mapper (TM) images that were acquired on 6th and 31st July of 2011 over the three Landsat scenes (WRS-2 Path/Rows (21/30, 21/31, and 22/30) that completely cover the watershed. The most recent cloud-free Collection 1 Landsat TM images processed to surface reflectance were obtained from Earth Explorer (https://earthexplorer.usgs.gov/). The Collection 1 Landsat images are defined with per-pixel cloud and quality information (Dwyer et al. 2018) and for this study cloudy pixels were removed. Following the Level 1 classification scheme of Anderson et al. (1976), seven land cover types were identified at 30 m resolution: barren, built-up (mainly as urban and roads), cropland, forest, grassland, open water, and wetland (Table 1). The image classification was conducted using the eCognition software (eCognition, 2019, version 9.2). Our first step of the classification was to generate homogeneous image objects through segmentation. We applied a multi-resolution segmentation algorithm based on both the spectral and shape information that quantify homogeneity using a set of parameters including scale, band weights, shape smoothness and compactness. First, rule sets were used to identify the water and urban object/classes using the Modified Normalized Difference Water Index (MNDWI) (McFeeters 1996) and Normalized Difference built-up Index (NDBI) (Zha et al. 2003) (viz., MNDWI>0.0 and NDBI>0.2, respectively). The remaining object/classes were classified using the nearest neighborhood classifier with training samples carefully selected by expert visual interpretation of the Landsat images and with the training samples selected across the watershed.

Land cover classification accuracy assessment was conducted by stratified random sampling of 700 Landsat TM pixels, with ~100 samples per class. The land cover of the 700 references pixels was determined by examination of high resolution commercial and airborne true color imagery in Google Earth supplemented by field visits. Standard confusion matrices were derived as cross-tabulations of the classified versus the reference class and used to estimate the overall, user's, and producer's accuracies (Foody 2002). The overall classification accuracy was 86.4%, with user and producer's accuracy of individual classes from 64.8-93.3% and 82.4-91.4%, respectively (Table 1).

2.3 Linear downscaling of albedo, ET and GPP

We calculated the proportions of the seven land cover types (κ_i, i = 1, 2 … 7) in each 500 m MODIS pixel for each Level IV ecoregion (Fig. 3). The annual mean blue-sky albedo, total annual ET, and annual mean GPP (hereafter referred to as albedo, ET, and GPP) of each MODIS 500 m pixel were used as the dependent variables in Eq. 1 to empirically estimate for each land cover type $albedo_i$, ET_i, and GPP_i:

$$Albedo = \sum (\kappa_i \times albedo_i) + \varepsilon \qquad [2]$$

$$ET = \sum (\kappa_i \times ET_i) + \varepsilon \qquad [3]$$

$$GPP = \sum (\kappa_i \times GPP_i) + \varepsilon \qquad [4]$$

where ε is the residuals of the ordinary least squares linear regression. RStudio (RStudio 1.1.453, https://www.rstudio.com/) was used to perform the downscaling with zero intercept and κ_i is the compositional portion (0-1) of land cover type *i* within the MODIS pixel, and the sum of κ_i must equal 1. The estimated albedo, ET, GPP and WUE for each cover type were quantified by cover type and ecoregion with their means and standard errors (SE) reported through boxplots.

3. Results

3.1 Landscape Characterization

Over 41.2% of the Kalamazoo River watershed falls in the ecoregion "Battle Creek Outwash Plain – 56b", which is characterized as broad and flat post-glacial plain. The "Lake Michigan Moraines -56f" and "Michigan Lake Plain – 5fd" ecoregions cover 22.5% and 9.6% of the watershed, respectively, and the remaining two ecoregions occupy 13.4%. The four dominant classified land cover types are cropland (57.1%), forest (28.1%), built-up (16.8%), and wetland (13.7%). Water, barren and grasslands each account for 0.5% to 3.0% of the watershed. There are large differences in landscape composition among the five ecoregions. The portion of cropland (40.9%) and forest (21.5%) in ecoregion 56b are substantially lower than the watershed mean, due to large portion of built-up land (21.8%) that includes the two largest cities: Kalamazoo and Battle Creek. In contrast, 45.6% of ecoregion 56d is classified as forest, resulting in a much lower cropland coverage (21.0%). The portion of wetland is higher than the watershed mean in 56d (14.4%) and 56g (14.6%). Build-up land in four of the ecoregions accounts for 7.5%-11.0% of the watershed (Fig. 2).

3.2 Land Surface Properties

The mean (SE) albedo of the five ecoregions is 0.154 (±0.016), with the highest value found in ecoregion 56g (0.163 ±0.010) and the lowest value in 56h (0.141 ± 0.016) (Table 2). Interestingly, frequency distributions of albedo in the five ecoregions are all left skewed ($\gamma < 0$), with the most skewed distribution found in 56g ($\gamma = -2.18$) and the least skewed distribution in 56f ($\gamma = -0.72$) (Fig. 3a2). For ET, the mean (SE) is 450.7 (±135.1) mm, but with a minimum of 365.9 (±179.4) in ecoregion 56d and a maximum of 477.5 (±94.5) mm in 56g. For GPP, the watershed mean (SE) of 7.27 (±2.36) MgC ha^{-1} yr^{-1} shows a deviation from a maximum of 7.71 (±1.71) MgC ha^{-1} yr^{-1} in ecoregion 56g and a minimum of 5.92 (±2.97) MgC ha^{-1} yr^{-1} in 56d (Table 2). As a result, WUE varies from 1.52 g kg^{-1} in ecoregion 56d to 1.63 g kg^{-1} in 56g (Table 2). Unlike the frequency distribution of albedo, both ET and GPP have bimodal density functions, except in ecoregion 56g (Fig. 3b2, c2). The first ET peak appears at ~210 mm and the second at ~515 mm. For GPP, similar peaks are at ~3 MgC ha^{-1} yr^{-1} and ~8 MgC ha^{-1} yr^{-1}. By comparing the spatial changes of ET, GPP, and land cover, the forest-dominated ecoregions exhibit lower ET and GPP than the cropland dominated ecoregions (Figs. 3, 4). For example, forest cover in 56d has the highest ET and GPP, which corresponds well with the lowest albedo. Whereas, in 56g, cropland accounts 61.6% of the landscape, and the corresponding albedo, ET, GPP, and WUE are the highest among the five Level IV ecoregions (Table 2, Fig. 4).

3.3 Downscaling

We achieved satisfactory downscaled results for albedo, ET, and GPP in all five ecoregions (Table 3). The adjusted R^2 is >0.995 for albedo, 0.915-0.971 for ET, and 0.902-0.962 for GPP. Values appear generally higher in ecoregions dominated by croplands (56g, 56h) than those dominated by forests (56d). Among the five ecoregions, 56f and 56d have higher variations (i.e., standard errors) in estimated albedo (Fig. 5a). The overall means for ET and GPP are the lowest in 56d compared with the four other ecoregions. Higher estimated variation is found in 56d and 56f, while lower variation in 56g appear in cropland dominant (Fig. 5b, c). Surprisingly, the WUE, based on estimated values of ET and GPP, appears very similar among the ecoregions, except in 56d where it is remarkably lower than other ecoregions and with high variation (Fig. 5d).

The downscaled estimates of albedo, ET, and GPP based on Eqs. 2-3 are reasonable (Table 3) and fall within the range of MODIS products (Fig. 3). The estimated albedo is the lowest for water (0.076-0.107) and the highest for cropland cover type (0.166-0.172). The albedo of the forest (0.136-0.153), wetland (0.141-0.155), and urban (0.136-0.152) cover types are similar, while grassland albedo (0.136-0.172) is slightly lower than that of cropland. Additionally, estimated albedo exhibits higher variation in barren, grasslands, and water than that of the four major cover types (Fig. 4a), likely because of their small sampling size (i.e., small portion of the landscape, Fig. 2) and large interannual variation. The high variation of albedo in the water cover type among the five ecoregions likely arises from seasonal variations in lake vegetation, sediment loading and waves.

Estimated ET (mm) is the highest for built-up cover type (525.6-687.1) and the lowest at forest (209.7-459.7), with a mean (SE) of 588.3 (47.7) and 318.3(69.9), respectively, among the five ecoregions (Table 3, Fig. 4b). Interestingly, the estimated ET for cropland land cover and water is not high as expected, especially from the water cover where ET source is sufficient (Fig. 4b). The variation associated with ET estimates for barren and grasslands are high, but not for water, possibly because of high similarity of evaporation among the open water bodies.

Estimated GPP (MgC ha^{-1} yr^{-1}) is the highest for the build-up cover type (8.65-9.85) and the lowest for the forest (4.81-7.42), with a mean (SE) of 9.71 (±0.76) and 5.13 (±1.09), respectively (Table 3, Fig. 4c). While a low GPP (6.42 ± 0.99) is reasonable for the wetland cover type, the high estimated GPP for barren cover (7.61 ± 3.67) and water (7.49 ± 0.86) need to be explored further to discover the source of the discrepancies in underlying mechanisms (e.g., small sampling size). Finally, WUE (g kg^{-1}) among the seven cover types are similar at an overall mean of 1.57, with the highest and the lowest WUE found in the built-up cover type (1.66) and barren (1.49) (Fig. 4d). Again, high variation in estimated WUE appears in grassland cover type, but not in the built-up and water cover type.

4. Discussion

Development of this downscaling approach was stimulated by the practical need to estimate ecosystem functioning at the Landsat spatial resolution of 30 m from coarser spatial resolution MODIS products at 500 m (Robinson et al. 2018). Alternative efforts had been made to derive or model functional measures directly from Landsat data (Zheng et al. 2004; Dieye et al. 2012; Gitelson et al. 2012; Semmens et al. 2016; Yao et al. 2017), or resample of Landsat within the MODIS frame (Wang et al. 2017; Trlica et al. 2017). Our approach offers a direct downscaling option by avoiding time-consuming processing or modeling efforts through integrating land cover maps with available MODIS products.

To demonstrate the concept, we applied the approach for blue-sky albedo, ET, and GPP in a watershed consisting of five Level IV ecoregions in southwestern Michigan. The estimated albedo, ET, and GPP values appear reasonable for the ecosystems in the region. Among the ecosystems, croplands and grasslands have higher albedos than the other cover types, with the water cover type showing the lowest values, as expected (Table 3). This pattern is consistent with snow-free albedo variations among ecosystems (Campbell and Norman 1998) and across the conterminous United States (Barnes and Roy 2010) and to widely reported values in similar ecosystems (e.g., Bonan 1997; Wang et al. 2017; Zhou et al. 2019). For example, Trlica et al. (2017) applied Landsat images and estimated albedo across the urban landscape of Boston. They reported a mean and range of 0.152

(0.112–0.187), which are comparable to our estimate of 0.136-0.152 (Table 3). Similarly, the ET and GPP estimates of our initial modeling appear in good agreements with those of *in situ* measurements (e.g., Papale et al. 2015), remote sensing modeling (Xiao et al. 2004; Yuan et al. 2010), or those from ecosystem models (Sun et al. 2011). For example, Abraha et al. (2015) reported the annual ET of measurements from eddy-covariance flux towers in seven bioenergy crops (2009-2012) that are within the 56b ecoregion of our watershed. Their annual ET ranged between 480 mm and 639 mm, which is similar to our estimate of the cropland in 2011 (512.2 mm).

4.1. Prerequisites of the downscaling approach

Our approach for downscaling available MODIS data to higher resolution products requires at least three strong assumptions. First, the dependent variable (e.g., functional measures albedo, ET, or GPP) is considered to be the linear sum of the independent values from each cover type. To estimate the contribution of each cover type empirically, its portion within a MODIS pixel is used in a multiple linear regression model with zero intercept (Eq. 1). A fundamental prerequisite is that the functional measure (e.g., albedo, ET, or GPP) is assumed to have the same value within each land cover type for every finer resolution pixel within among coarser resolution MODIS pixels. This constraint suggests that our approach can be only applied a spatial extent where the climate, soils, and landforms across the landscape are the same (or similar). The proper scale for model application will depend on the heterogeneity of the region studied (e.g., complexity of land forms, distribution of soils, climatic variation, etc.). For instance, strong terrain effects (elevation, aspect, slope) may require further subdivision beyond land cover type to portray spatial distributions effectively.

To satisfy this assumption, one would need to explore the changes of the functional measure with scale and spatial extent to determine an appropriate regionalization for the model application. In modeling of landscape that is stationary in the statistical sense (Saunders et al. 2005), geostatistical tools such as semivariance analysis (Atkinson 2013) can be applied to quantify the geostatistical range as the proper scale (Cressie and Wikle 2015). Lacunarity and wavelet analysis are alternative choices for non-stationary landscapes (Saunders et al. 2005). These preliminary analyses for identifying the minimal extent can be performed for MODIS products, and/or climate/soils data, if they are available prior to downscaling. Violations of this assumption will significantly reduce the model reliability and increase the uncertainty of downscaled estimates (see next section for discussion on uncertainty). An alternative approach to determine the minimal landscape extent for satisfying the model assumptions is to use ecoregions that have been delineated by "similar land forms, soils, and climate" criteria. We based our demonstrative modeling on the Level IV ecoregions for the State of Michigan (Omernik and Griffith 2014). Other similar products, such as maps of potential vegetation (Tucker et al. 2005; Fensholt and Proud 2012), climatic zones (Kottek et al. 2006), soils (Stoorvogel et al. 2017), or landforms (O'Loughlin et al. 2016) could be also considered in absence of quantitative identification of the landscape scale.

The second assumption of this downscaling approach is that every land cover type (i.e., the ancillary data used for the modeling) remains unchanged during the study period (e.g., within a year in this study), which is untrue if there are rapid changes of land use or land cover or strong seasonality in land cover characteristics. For example, many cropland-dominated landscapes are managed with rotation crops (i.e., corn – soybeans – wheat), different irrigation scheduling, conservation tillage and other land management practices (Plourde et al. 2013). Treating these croplands as a single cover type to estimate their ecosystem function may be misleading. An alternative is to apply the model during a shorter period so that this model assumption is not violated. For example, one could model albedo, ET, and GPP for each 16-day composite period of the MODIS product. The seasonal and annual values could then be summed from the series of values. One caveat: the spatial scale of the variable of interest may vary through time (cf., Henebry 1993; Goodin and Henebry 1998).

The third assumption is the availability of many finer resolution pixels as well as an adequate number of coarser resolution pixels. The modeled landscape needs to include finer spatial resolution pixels (e.g., Landsat) that are many times the number of cover types in order to estimate the empirical coefficients, with more pixels providing higher estimation confidence. In the limit, mathematically, the coefficients in Eq. 1 can be calculated with the number of coarser spatial resolution pixels (e.g., MODIS) the same as the number of land cover types. For this study, a minimum of seven pixels were needed, which implies a minimal landscape size of 1.75 km^2 when using 500-m resolution MODIS products. It is worth noting that most MODIS pixels do not include all seven cover types (see Fig. 1), which will complicate estimation of the minimum extent. However, the model confidence can be significantly improved with larger sample sizes. For example, if the study landscape is 10-15 km^2, one would have 400-600 MODIS pixels at 500 m. An alternative to assure large sampling size could be done with the "sliding box" approach (Rodriguez-Iturbe et al. 1998; Saunders et al. 2005), where the study landscape with minimal size can be slid across a large region by allowing overlaps. However, each additional sample is a pseudo-replicate, likely with positive spatial autocorrelation, suggesting that substantially more effort is needed to conduct posterior analyses for model uncertainty (Hargrove and Pickering 1992). Nevertheless, the balance between sample size and assuring similar values by cover type needs also be analyzed prior to model application.

4.2 Uncertainties of the downscaling predictions

The conceptual downscaling model (Eq. 1) would be flawless if: (1) all model assumptions were met, and (2) both dependent and independent variables were truly representative measures of the ecosystem properties. The soundness of this modeling approach is demonstrated with high model R^2 values in each Level 4 ecoregion in the Kalamazoo River watershed (Table 3) with reasonable predictions of albedo, ET and GPP (Figs. 4). Unfortunately, this situation may not always be the case in practice, since uncertainties can arise from multiple sources. First, both the dependent and independent variables in Eqs. 2-3 are produced from satellite observations and carry uncertainties associated with the algorithm used for the retrieval, including calibration, geometric, and atmospheric corrections (Vermote et al. 2002; Wolfe et al. 2002; Helder et al. 2018). From a systems modeling perspective, uncertainty generally increases as more steps or more variables or more algorithms are involved. Here MODIS blue-sky albedo is retrieved from surface reflectance products corrected for BRDF and using retrieved atmospheric aerosols, and ET and GPP are modeled with algorithms with multiple inputs (e.g., climate scalars, coefficients of light use efficiency, surface roughness, etc.) (Schaaf et al. 2002; Zhao et al. 2005, 2006; Chrysoulakis et al. 2018). These processing chains may be a potential reason for the higher confidence level in predicting albedo than ET or GPP, which carry greater uncertainty. However, we note that albedo is a bounded value, unlike ET or GPP, which also constrains variation. The accuracy of MODIS products are also often influenced by atmospheric conditions, in particular undetected sub-pixel clouds, resulting in different predictions among pixels even though their landscape composition may be the same (Schaaf et al. 2002; Yang et al. 2006). These differences may also vary among the images of different times, which can further propagate uncertainty, regardless of quality control and data screening process, when multiple products are applied for calculating a value at longer temporal scale (e.g., annual values in this study).

The Landsat classified land cover type and distributions across the landscape are also not without uncertainty, since classification accuracy varies by cover type (Table 1). In this pilot study, the user's accuracies were lower for barren, water, and build-up cover types than for other cover types. It is well established that urban areas are difficult to classify reliably because they encompass such a variety of building types and land uses, so even at the Landsat scale of 30 m, the results are often mixed spatially (Zhang and Roy 2017). Similarly, water and barren are quite broad thematic definitions that can be highly variable in space and in time. Thus, the number of cover types as well as the variation within a single cover type can generate additional uncertainty. For example, forests within the watershed are treated as one cover type, although their species composition, stand structure, age, disturbance history (including management), etc. may vary substantially across the landscape. Similar variation in composition and structure exist among patches of the same cover type within a MODIS pixel.

Low classification accuracy may be partially responsible for the high standard errors in predicting albedo, ET, and GPP within an ecoregion (Table 2) and among the ecoregions (Fig. 5). The standard errors also vary by cover type (Fig. 4), with the barren cover type having the highest value and croplands having the lowest. While it is likely that the prevalence of nearly homogeneous surfaces is higher in cropland than in barren and grassland covers, providing a partial explanation for the differences in SE, we speculate that classification accuracy is also partially responsible for the uncertainty in downscaling albedo, ET, and GPP.

The scale at which the modeling is applied provides another source of uncertainty (Levin 1992; LeMoine and Chen 2003; Saunders et al. 2005). We based our modeling example here using the Level IV ecoregions without a quantitative exploration of the “right scale” (sensu Levin 1992). Future efforts are strongly recommended to include scale identification prior to the modeling. More importantly, neither MODIS nor Landsat pixels represents a homogeneous ecosystem (i.e., a cover type). Across the Kalamazoo River watershed, there exist smaller patches and edges (<30 m), suggesting that the composition of land cover within a Landsat pixel is an approximation or generalization. The uncertainty will likely be higher in landscapes with more cover types, finer patch sizes, and more edges.

4.3 Transformative applications of the downscaling approach

Our proposed downscaling approach decomposes the property of a spatially nested hierarchical system when the structure of the lower hierarchical level is known. This linear downscaling approach can be applied to nested spatial, temporal, or organizational hierarchies. A key advantage of this approach is that it enables the use of coarser spatial resolution MODIS products for predicting finer spatial resolution values that are directly connected with landscape composition. Although this paper focuses on downscaling MODIS products by using classified Landsat data, other finer resolution land cover data (e.g., derived from Sentinel-2) could be used. Similarly, products from the NOAA's operational MODIS follow-on sensor VIIRS (VIsible Infrared Radiometer Suite) could be used in place of MODIS products (Zhang et al. 2018). The downscaled estimates provide us the opportunity to explore ecosystem functioning at the level of spatial detail afforded by Landsat and comparable sensors.

Connecting patterns and processes, for example, has been a major interest within the landscape ecology community (Turner and Gardner 2015). While substantial knowledge has been gained over the past 30 years to improve the understanding of the empirical and theoretical relationships between ecosystem functioning and landscape structure, there remains a major disconnect between the disciplines of ecosystem ecology and landscape ecology. For example, many landscape studies have been published on landscape structure and dynamics, often based on Landsat images. While modeling ecosystem production and other functions at 30 m have been reported (e.g., Zheng et al. 2004), such efforts require much ground level data for model parametrization, calibration, and validation and intensive computations, despite often limited availability of cloud-free data. With availability of MODIS products, one can more readily relate ecosystem functions with landscape composition and structure.

Another promising direction is to examine the underlying landscape processes from using the residuals (ε in Eqs. 1-4), which may point to potential driving mechanisms. For example, the residuals from Eqs. 2-4 could be examined for their latent relationships using quantitative metrics describing landscape structure to address the question: What landscape structure and processes are responsible for this unexplained variation in ecosystem functioning? Here we hypothesize that the patch patterns, variation in soils, climate and management practices within a landscape may all contribute to the variation in residuals, especially for the high uncertainties found in ET and GPP.

The downscaling model could be further expanded to explore other landscape processes. For example, the GPP model (Eq. 4) could be refined to include interactive terms among the cover types:

$$GPP = \sum (\kappa_i \times GPP_i) + (\kappa_{i*j} \times GPP_{i*j}) + \varepsilon \qquad [5]$$

where the estimated GPP_{i*j} reflects the interactive contribution of two cover types (*i* and *j*) that could be related to edge effects between them. Once the model predictions are validated, it may provide a great opportunity for understanding the ecosystem consequences of spatial fragmentation and disturbances across the landscape (Franklin and Forman 1987; Chen 1991; Di Giulio et al. 2009).

5. Conclusions

Based on practical needs for downscaling popular MODIS products at 500 m resolution to match classified land cover from Landsat data at 30 m resolution, we have proposed an innovative modelling approach so that landscape structure and ecosystem functioning could be directly studied for their interconnections. As a proof-of-concept, we tested this model in the five landscapes of Level IV ecoregions found within the Kalamazoo River watershed using three key ecosystem functional attributes: albedo, ET, and GPP. An object-oriented classification of Landsat imagery in 2011 was processed to generate a land cover map indicating landscape stucture. Each downscaling model exhibited high fit (R^2 >90%), with higher confidence levels for albedo than for GPP or ET. The estimated values for albedo, ET, and GPP appear reasonable and within their ranges reported for the region and consistent with values reported in the literature. Despite these promising results, this approach relies on strong assumptions that can be difficult to characterize. It is only valid at a scale where similar climate, soils, and landforms exist (i.e., value of the same cover type in isolated patches are the similar). Plausibly, the uncertainties associated with each estimation, as well as the model residuals, can be explored further to detect other pattern-process relationships within the landscapes.

Acknowledgements: This study was supported, in part, by the NASA Carbon Cycle & Ecosystems program (NNX17AE16G), the Great Lakes Bioenergy Research Center funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-SC0018409 and DE-FC02-07ER64494; and the Long-term Ecological Research Program (DEB 1637653) at the Kellogg Biological Station, and the NASA Science of Terra and Aqua program (NNX14AJ32G). We thank the fruitful discussion at LEES Lab meetings where several members made fruitful suggestions for model development. Isabel Arroca assisted in formatting the references. The reviews from two anonymous reviewers helped improving the quality of this manuscript.

References

Abraha M, Chen J, Chu H, et al (2015) Evapotranspiration of annual and perennial biofuel crops in a variable climate. *G*lob Chang Biol *Bioenergy* 7(6): 1344-1356.

Alberti M, Asbjornsen H, Baker LA, et al (2011) Research on coupled human and natural systems (CHANS): Approach, challenges, and strategies. *Bull Ecol Soc Am* 92: 218–228.

Anderson JR, Hardy EE, Roach JT, et al (1976) A land use and land cover classification system for use with remote sensor data. *Professional Paper* 964: doi: 10.3133/pp964. 28p.

Atkinson PM (2013) Downscaling in remote sensing. *Int J Appl Earth Obs Geoinf 22:* 106-114.

Barnes CA, Roy DP (2010) Radiative forcing over the conterminous United States due to contemporary land cover land use change and sensitivity to snow and interannual albedo variability, *J Geophys Res* 115: G04033, doi:10.1029/2010JG001428.

Bass DG (2009) *Inferring Dissolved Phosphorus Cycling in a TMDL Watershed Using Biogeochemistry and Mixed Linear Models*. PhD Disertation, Michigan State University, East Lansing, Michigan.

234p.

Bonan GB (1997) Effects of land use on the climate of the United States. *Clim Change* 37(3): 449–486.

Bresee MK, Le Moine J, Mather S, et al. (2004) Disturbance and landscape dynamics in the Chequamegon National Forest Wisconsin, USA, from 1972 to 2001. *Landsc Ecol* 19(3): 291-309.

Campbell GS, Norman JM (1998) *Introduction to Environmental Biophysics*. Springer-Verlag, New York, USA. 286p.

Chapman KA, Brewer R (2008) Prairie and Savanna in southern lower Michigan: History, Classification, Ecology. *Michigan Bot* 47(1). doi: hdl.handle.net/2027/spo.0497763.0047.101. 40p.

Chen J (1991) *Edge Effects: Microclimatic Pattern and Biological Responses in Old-Growth Douglas-fir Forests*. PhD Dissertation, University of Washington, Seattle, WA, USA.

Chen J, Wan S, Henebry G, et al (eds) (2013) *Dryland East Asia: Land Dynamics amid Social and Climate Change*. DE GRUYTER, Berlin, Boston, 470p.

Chrysoulakis N, Mitraka Z, Gorelick N (2018) Exploiting satellite observations for global surface albedo trends monitoring. *Theor Appl Climatol* 1–9. doi: 10.1007/s00704-018-2663-6

Cressie N, Wikle CK (2015) *Statistics for Spatio-Temporal Data*. John Wiley & Sons. New Jersey, USA. 624p.

Di Giulio M, Holderegger R, Tobias S (2009) Effects of habitat and landscape fragmentation on humans and biodiversity in densely populated landscapes. *J Environ Manage* 90: 2959-2968.

Dieye AM, Roy DP, Hanan NP, et al (2012), Sensitivity analysis of the GEMS soil organic carbon model to land cover land use classification uncertainties under different climate scenarios in Senegal. *Biogeosciences* 9: 631-648.

Drusch M, Del Bello U, Carlier S, et al (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sens Environ* 120: 25-36.

Dwyer JL, Roy DP, Saue B, et al (2018) Analysis ready data enabling analysis of the Landsat archive. *Remote Sensing* 10 (1363): 19. doi.org/10.3390/rs10091363

Fensholt R, Proud SR (2012) Evaluation of earth observation based global long term vegetation trends—

Comparing GIMMS and MODIS global NDVI time series. *Remote Sens Environ* 119: 131-147.

Fongers D (2008) *Kalamazoo River Watershed Hydrologic Study*, Michigan Department of Environmental Quality, Lansing, Michigan. 67p.

Foody GM (2002). Status of land cover classification accuracy assessment. *Remote Sens Environ* 80(1): 185-201.

Forman RTT (1995) *Land mosaics : The Ecology of Landscapes and Regions*. Cambridge University Press, Cambridge, UK, 217p.

Fisher, F. Peter and Mitchel Langford. 1996. Modeling sensitivity to accuracy in classified imagery: A study of areal interpolation by dasymetric mapping. *Prof Geog* 48(3):299-309.

Franklin JF, Forman RTT (1987) Creating landscape patterns by forest cutting: Ecological consequences and principles. *Landsc Ecol* 1: 5-18.

Friedl MA, Sulla-Menashe D, Tan B, et al (2010) MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens Environ* 114: 168–182.

Giglio L, Boschetti L, Roy DP, et al (2018) The collection 6 MODIS burned area mapping algorithm and product, *Remote Sens Environ* 217: 72-85.

Gitelson AA, Peng Y, Masek JG, et al (2012) Remote estimation of crop gross primary production with Landsat data. *Remote Sens Environ* 121: 404–414.

Goodin DG, Henebry GM (1998). Seasonality of finely-resolved spatial structure of NDVI and its component reflectances in tallgrass prairie. *Int J Remote Sens 19*(16): 3213-3220.

Hansen MC, Loveland TR (2012) A review of large area monitoring of land cover change using Landsat data. *Remote Sens Environ* 122: 66–74.

Helder D, Markham B, Morfitt R, et al (2018) Observations and recommendations for the calibration of Landsat 8 OLI and Sentinel 2 MSI for improved data interoperability. *Remote Sens* 10(9): 1340. doi.org/10.3390/rs10091340.

Hargrove WW, Pickering J (1992). Pseudoreplication: A sine qua non for regional ecology. *Landsc Ecol* 6(4): 251-258.

Henebry GM (1993) Detecting change in grasslands using measures of spatial dependence with Landsat TM data. *Remote Sen Environ 46*(2): 223-234.

Jia P, Gaughan AE (2016) Dasymetric modeling: A hybrid approach using land cover and tax parcel data for mapping population in Alachua County, Florida. *Appl Geogr* 66: 100-108.

John R, Chen J, Kim Y, et al (2016) Differentiating anthropogenic modification and precipitation-driven change on vegetation productivity on the Mongolian Plateau. *Landsc Ecol* 31: 547-566.

Justice CO, Giglio L, Korontzi S, et al (2002) An overview of MODIS Land data processing and product status. *Remote Sens Environ* 83: 3-15.

Kennedy RE, Andréfouët S, Cohen WB, et al (2014) Bringing an ecological view of change to Landsat-based remote sensing. *Front Ecol Environ* 12: 339–346.

Kottek M, Grieser J, Beck C, et al (2006) World Maps of Köppen-Geiger climate classification updated. *Meteorol Zeitschrift* 15(3): 259-263

Krehbiel CP, X Zhang, GM Henebry (2017) Impacts of thermal time on land surface phenology in urban areas. *Remote Sensing* 9(5): 499. doi.org/10.3390/rs9050499

LeMoine JM, Chen J (2003) Placing our research objectives and results in time and space. *Acta Phytoecol Sinica* 27: 1–10.

Levin SA (1992) The problem of pattern and scale in ecology: The Robert H. MacArthur Award Lecture. *Ecology* 73(6): 1943-1967.

McFeeters S (1996) The use of Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int J Remote Sens* 17: 1425-1432.

Mills K, Schillereff D, Saulnier-Talbot É, et al (2017) Deciphering long-term records of natural variability and human impact as recorded in lake sediments: A palaeolimnological puzzle. *Wiley Interdiscip Rev Water* 4: e1195. doi: 10.1002/wat2.1195.

Moon M, Zhang X, Henebry GM, et al. (2019) Long-term continuity in land surface phenology measurements: A comparative assessment of the MODIS land cover dynamics and VIIRS land surface phenology products. *Remote Sens Environ 226:* 74-92.

Mu Q, Heinsch FA, Zhao M, et al (2007) Development of a global evapotranspiration algorithm based on MODIS and global meteorology data. *Remote Sens Environ* 111: 519–536.

Nagle NN, Buttenfield BP, Leyk S, Spielman S (2014) Dasymetric modeling and uncertainty. *Ann Am Assoc Geogr 104*(1): 80-95.

O'Loughlin FE, Paiva RCD, Durand M, et al (2016) A multi-sensor approach towards a global vegetation corrected SRTM DEM product. *Remote Sens Environ* 182: 49-59.

Omernik JM, Griffith GE (2014) Ecoregions of the conterminous United States: Evolution of a hierarchical spatial framework. *Environ Manage* 54(6): 1249-1266.

Papale D, Black TA, Carvalhais N, et al (2015) Effect of spatial sampling from European flux towers for estimating carbon and water fluxes with artificial neural networks. *J Geophys Res Biogeosciences* 120(10): 1941-1957.

Petrov A (2012) One hundred years of dasymetric mapping: Back to the origin. *Cartogr J* 49 (3): 256–264.

Plourde, J. D., Pijanowski, B. C., Pekin, B. K., 2013. Evidence for increased monoculture cropping in the Central United States. *Agric Ecosyst Environ* 165: 50-59.

Robinson NP, Allred BW, Smith WK, et al. (2018) Terrestrial primary production for the conterminous United States derived from Landsat 30 m and MODIS 250 m. *Remote Sens Ecol Conserv 4*(3): 264-280.

Rodriguez-Iturbe I, D'Odorico P, Rinaldo A (1998) Configuration entropy of fractal landscapes. *Geophys Res Lett* 25(7): 1015-1018.

Roy DP, Wulder MA, Loveland TR, et al (2014) Landsat-8: Science and product vision for terrestrial global change research. *Remote Sens Environ* 145: 154–172.

Saunders SC, Chen J, Drummer TD, et al (2005) Identifying scales of pattern in ecological data: A comparison of lacunarity, spectral and wavelet analyses. *Ecol Complex 2*(1): 87-105.

Schaaf CB, Gao F, Strahler AH, et al (2002) First operational BRDF, albedo nadir reflectance products from MODIS. *Remote Sens Environ* 83(1-2): 135-148

Schaetzl RJ, Darden JT, Brandt DS (2009) *Michigan Geography and Geology.* Pearson Learning Solutions. 672p.

Semmens KA, Anderson MC, Kustas WP, et al (2016) Monitoring daily evapotranspiration over two California vineyards using Landsat 8 in a multi-sensor data fusion approach. *Remote Sens Environ* 185: 155–170.

Stoorvogel JJ, Bakkenes M, Temme AJAM, et al (2017) S-World: A global soil map for environmental modelling. *L Degrad Dev* 28: 22–33.

Sun G, Alstad K, Chen J, et al (2011) A general predictive model for estimating monthly ecosystem evapotranspiration. *Ecohydrology* 4(2): 245-255.

Trlica A, Hutyra LR, Schaaf CL, et al (2017) Albedo, land cover, and daytime surface temperature variation across an urbanized landscape. *Earth's Futur* 5: 1084–1101.

Tucker CJ, Pinzon JE, Brown ME, et al (2005) An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data. *Int J Remote Sens* 26(20): 4485-4498.

Turner MG, Gardner RH (2015) *Landscape Ecology in Theory and Practice: Pattern and Process*, second edition, Springer, New York, USA. 482p.

Vermote EF, El Saleous NZ, Justice CO (2002) Atmospheric correction of MODIS data in the visible to middle infrared: First results. *Remote Sens Environ* 83: 97-111.

Wan Z (2013) *MODIS Land Surface Temperature Products Users' Guide. Institute for Computational Earth System Science*, University of California, Santa Barbara, CA. 77p.

Wang Z, Schaaf CB, Sun Q, et al (2017) Monitoring land surface albedo and vegetation dynamics using high spatial and temporal resolution synthetic time series from Landsat and the MODIS BRDF/NBAR/albedo product. *Int J Appl Earth Obs Geoinf 59: 104-117.*

Wang Z, Schaaf CB, Sun Q, et al (2018). Capturing rapid land surface dynamics with Collection V006 MODIS BRDF/NBAR/Albedo (MCD43) products. *Remote Sens Environ 207:* 50-64.

Wolfe R, Nishihama M, Fleig A, et al (2002) Achieving sub-pixel geolocation accuracy in support of MODIS land science. *Remote Sens Environ* 83: 31-49.

Wu J, Hobbs RJ (eds) (2007) *Key Topics in Landscape Ecology*. Cambridge University Press, Cambridge, UK. 297p.

Wulder MA, Coops NC, Roy DP, et al (2018) Land Cover 2.0. *Int J Remote Sens* 39(12): 4254-4284.

Wulder MA, Loveland TR, Roy DP, et al (2019) Current status of Landsat program, science, and applications, *Remote Sens Environ* 224: 127-147.

Xiao X, Hollinger D, Aber J, et al (2004) Satellite-based modeling of gross primary production in an evergreen needleleaf forest. *Remote Sens Environ* 89: 519–534.

Yang W, Shabanov N V., Huang D, et al (2006) Analysis of leaf area index products from combination of MODIS Terra and Aqua data. *Remote Sens Environ* 104(3): 297-312.

Yao Y, Liang S, Li X, et al (2017) Estimation of high-resolution terrestrial evapotranspiration from Landsat data using a simple Taylor skill fusion method. *J Hydrol* 553: 508-526.

Yuan W, Liu S, Yu G, et al (2010) Global estimates of evapotranspiration and gross primary production based on MODIS and global meteorology data. *Remote Sens Environ* 114(7): 1416-1431.

Zha Y, J Gao, Ni S (2003) Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *Inl J Remote Sens* 24(3): 583-594.

Zhang, HK, Roy DP (2017) Using the 500 m MODIS land cover product to derive a consistent continental scale 30 m Landsat land cover classification. *Remote Sens Environ* 197: 15-34.

Zhang X, Friedl MA, Schaaf CB, et al (2003). Monitoring vegetation phenology using MODIS. *Remote Sens Environ* 84(3): 471-475.

Zhang XY, Liu L, Liu Y et al (2018) Generation and evaluation of the VIIRS land surface phenology product. *Remote Sens Environ* 216: 212-229.

Zhao M, Heinsch FA, Nemani RR, et al (2005) Improvements of the MODIS terrestrial gross and net primary production global data set. *Remote Sens Environ* 95(2): 164-176.

Zhao M, Running SW, Nemani RR (2006) Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses. *J Geophys Res Biogeosciences 111(G1)*. doi: 10.1029/2004JG000004

Zheng D, Rademacher J, Chen J, et al (2004) Estimating aboveground biomass using Landsat 7 ETM+ data across a managed landscape in northern Wisconsin, USA. *Remote Sens Environ* 93(3): 402-411.

Zhou H, Liang S, He T, et al (2019) Evaluating the spatial representativeness of the MODerate Resolution Image Spectroradiometer albedo product (MCD43) at AmeriFlux sites. *Remote Sens 11*(5), 547, doi.org/10.3390/rs11050547.

Table 1. Brief description of seven land cover types in the Kalamazoo River watershed, the number of validation samples by cover type, and classification accuracies from user's and producer's perspective.

Barren: Areas where not covered by vegetation or construction

Built-up: built-up land with a mixture of constructed materials and vegetation

Cropland: annual crops and/or pastoral land

Forest: land dominated by high tree cover

Grassland: herbaceous vegetation, including lawns

Water: open water (e.g., lakes, reservoir, and large rivers)

Wetland: mostly forested wetlands

	Built-up	Cropland	Grassland	Forest	Water	Wetland	Barren	User's accuracy
Built-up	46	9	3	11	0	0	2	64.8%
Cropland	2	191	2	11	0	0	2	91.8%
Grassland	1	2	57	1	0	1	2	89.1%
Forest	2	10	3	154	0	1	0	90.6%
Water	0	0	0	6	53	9	0	77.9%
Wetland	0	1	0	1	3	70	0	93.3%
Barren	0	3	1	3	2	1	34	77.3%
Producer's Accuracy	90.2%	88.4%	86.4%	82.4%	91.4%	85.4	85.0%	**86.4%**

Table 2. Statistics for blue-sky albedo, evapotranspiration (ET, mm), gross primary production (GPP, MgC ha^{-1} yr^{-1}) and water use efficiency (g kg^{-1}) in the five Level IV ecoregions of the Kalamazoo Watershed in Southwest Michigan, USA. MED is Median; γ is skewness; SE is standard error.

Ecoregion	**Blue-sky albedo**			**ET**			**GPP**			**WUE**
	Mean (SE)	MED	γ	Mean (SE)	MED	γ	Mean (SE)	MED	γ	Mean (SE)
Battle -56b	0.150 (±0.017)	0.151	-1.85	455.41 (±131.79)	498.80	-0.91	7.35 (±2.30)	8.09	-0.83	1.60 (±0.01)
Michigan - 56d	0.141 (±0.016)	0.140	-1.19	365.86 (±179.41)	270.50	0.41	5.92 (±2.97)	4.44	0.43	1.52 (±0.13)
Lake - 56f	0.161 (±0.014)	0.161	-0.72	470.30 (±130.23)	508.60	-1.02	7.63 (±2.33)	8.28	-0.96	1.58 (±0.01)
Lansing - 56g	0.163 (±0.010)	0.165	-2.18	477.47 (±94.49)	492.00	-1.41	7.71 (±1.71)	7.98	-1.23	1.63 (±0.03)
Interlobate - 56h	0.155 (±0.012)	0.156	-0.99	444.89 (±132.44)	486.60	-0.87	7.13 (±2.34)	7.80	-0.78	1.59 (±0.03)
Overall	**0.154 (±0.016)**	**0.156**	**-1.49**	**450.65 (±135.13)**	**492.70**	**-0.84**	**7.27 (±2.36)**	**7.98**	**-0.76**	**1.57 (±0.11)**

Table 3. Estimated mean (standard error, SE) of blue-sky albedo, evapotranspiration (ET, mm) and gross primary production (GPP, MgC ha^{-1} yr^{-1}) by cover type and Level IV ecoregions in the Kalamazoo River watershed in southwestern Michigan, USA.

Cover Type	Battle (56b)			Michigan (56d)			Lake (56f)			Lansing (56g)			Interlobate (56h)		
	Albedo	**ET**	**GPP**	**Albedo**	**ET**	**GPP**	**Albedo**	**ET**	**GPP**	**Albedo**	**ET**	**GPP**	**Albedo**	**ET**	**GPP**
Built-up	0.1419	555.6	9.25	0.1362	687.1	9.75	0.1490	597.0	9.85	0.1518	525.6	8.65	0.1436	576.0	9.47
	(0.0004)	(5.9)	(0.10)	(0.0018)	(23.9)	(0.16)	(0.0011	(15.4)	(0.28)	(0.0014)	(17.7)	(0.33)	(0.0006)	(9.3)	(0.17)
Cropland	0.1664	515.7	8.27	0.1685	600.6	9.72	0.1728	512.2	8.35	0.1699	479.7	7.71	0.1673	511.0	8.21
	(0.0003)	(2.7)	(0.05)	(0.0007)	(9.1)	(1.01)	(0.0003	(3.3)	(0.06)	(0.0003)	(3.3)	(0.06)	(0.0002)	(3.0)	(0.05)
Grassland	0.1606	489.4	7.49	0.1356	766.6	3.60	0.1723	713.3	11.63	0.1706	442.7	7.15	0.1654	510.6	8.08
	(0.0035)	(39.1)	(0.70)	(0.0046)	(58.4)	(0.12)	(0.0066	(84.2)	(1.53)	(0.0085)	(92.6)	(1.73)	(0.0033)	(43.9)	(0.79)
Forest	0.1380	300.5	4.81	0.1359	209.7	5.64	0.1425	334.3	5.33	0.1530	459.7	7.42	0.1429	287.2	4.49
	(0.0005)	(5.2)	(0.09)	(0.0005)	(6.9)	(0.58)	(0.0007	(8.8)	(0.16)	(0.0010)	(11.0)	(0.21)	(0.0004)	(5.5)	(0.10)
Water	0.0764	514.1	8.28	0.0807	408.6	4.64	0.1072	472.9	7.49	0.1024	541.4	8.06	0.1000	489.2	7.37
	(0.001)	(17.1)	(0.30)	(0.0018)	(33.4)	(0.20)	(0.0017	(26.5)	(0.48)	(0.0020)	(33.7)	(0.63)	(0.0015)	(22.8)	(0.41)
Wetland	0.1482	391.7	6.23	0.1408	307.3	5.51	0.1456	399.9	6.31	0.1553	496.2	8.10	0.1490	425.5	6.81
	(0.0007)	(7.7)	(0.14)	(0.0009)	(11.6)	(0.68)	(0.0011	(14.1)	(0.26)	(0.0008)	(9.0)	(0.17)	(0.0005)	(6.9)	(0.12)
Barren	0.1535	524.3	8.50	0.1362	376.2	9.85	0.1107	201.4	2.37	0.1471	795.7	12.51	0.1546	571.8	9.18
	(0.0023)	(25.1)	(0.45)	(0.0018)	(39.2)	(0.28)	(0.0055	(69.3)	(1.26)	(0.0101)	(114.9)	(2.15)	(0.0043)	(56.0)	(1.01)
Adj. R^2	**0.995**	**0.944**	**0.933**	**0.996**	**0.915**	**0.902**	**0.997**	**0.947**	**0.935**	**0.998**	**0.971**	**0.962**	**0.997**	**0.943**	**0.929**

Figure Captions

Figure 1. A snapshot of land cover overlaid with the 500 m MODIS grids within the Kalamazoo River Watershed. The portion of each cover type in each MODIS grid is calculated as κ_i that ranges from 0 to 1 for downscaling grid-level GPP to cover type GPP_i (Eq. 1).

Figure 2. Landscape cover of the Kalamazoo River Watershed in 2011 in southwestern Michigan, USA. The table includes the total land area (km^2) of each cover type and its portion (%) within the Level IV Ecoregion.

Figure 3. Spatial distributions of blue-sky albedo (%), evapotranspiration (ET, mm) and gross primary production (GPP, MgC ha^{-1} yr^{-1}) in 2011, as well as their frequency distributions described with probability density function (PDF). The PDF distributions were generated with the density function in RStudio with the bandwidth (*bw*) of standard deviation.

Figure 4. Boxplots of: (a) blue-sky albedo, (b) evapotranspiration (ET, mm), (c) gross primary production (GPP, MgC ha^{-1} yr^{-1}), and (d) ecosystem water use efficiency WUE (g kg^{-1}) within the Kalamazoo River Watershed in southwestern Michigan, USA. The statistics were calculated among the five Level IV Ecoregions in 2011 (see Fig. 5). The horizontal line and the number inside the circle are the grand mean of the watershed.

Figure 5. Boxplots of: (a) albedo, (b) evapotranspiration (ET, mm), (c) gross primary ecosystem production (GPP, MgC ha^{-1} yr^{-1}), and (d) ecosystem water use efficiency (WUE, g kg^{-1}) at five Level IV Ecoregions within the Kalamazoo River Watershed in southwestern Michigan, USA. The statistics were calculated among the seven land cover types in 2011 (see Fig. 4). The horizontal line and the number inside the circle are the grand mean of the watershed.

Fig. 1.

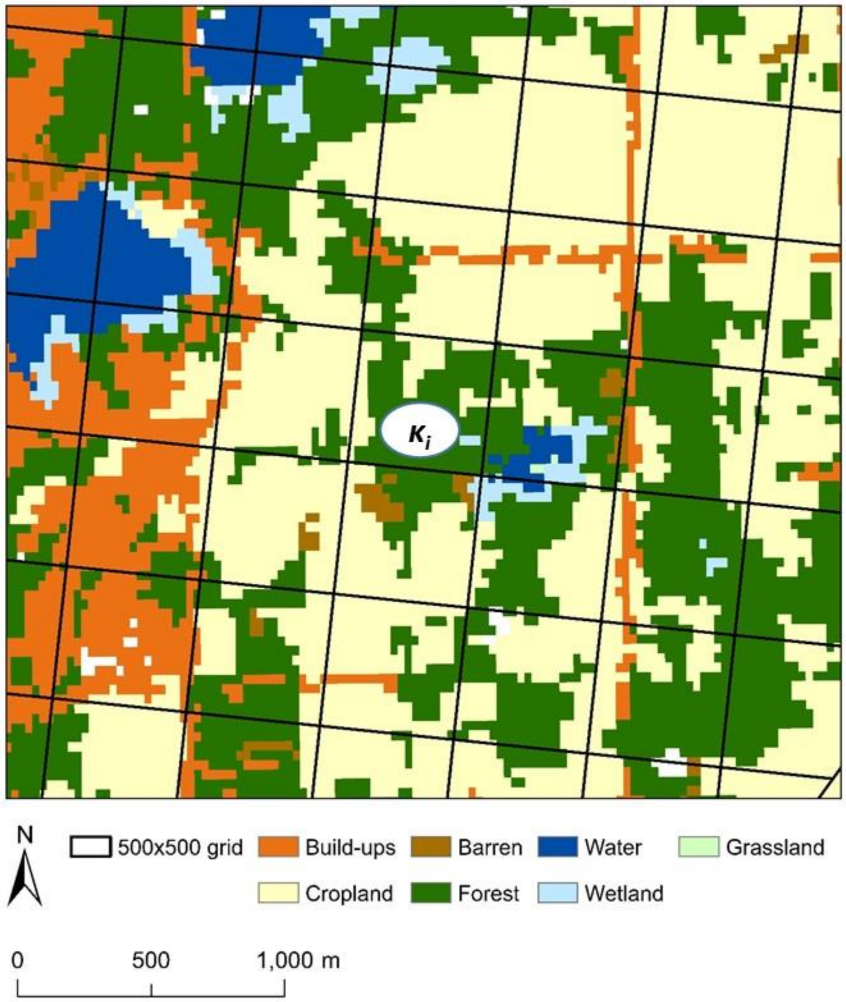
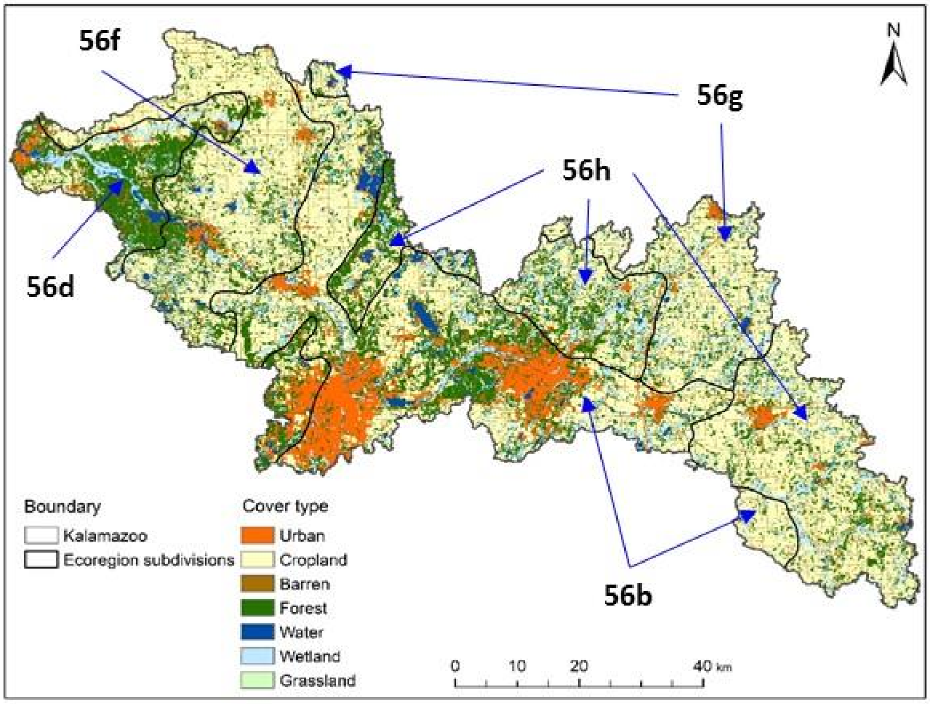
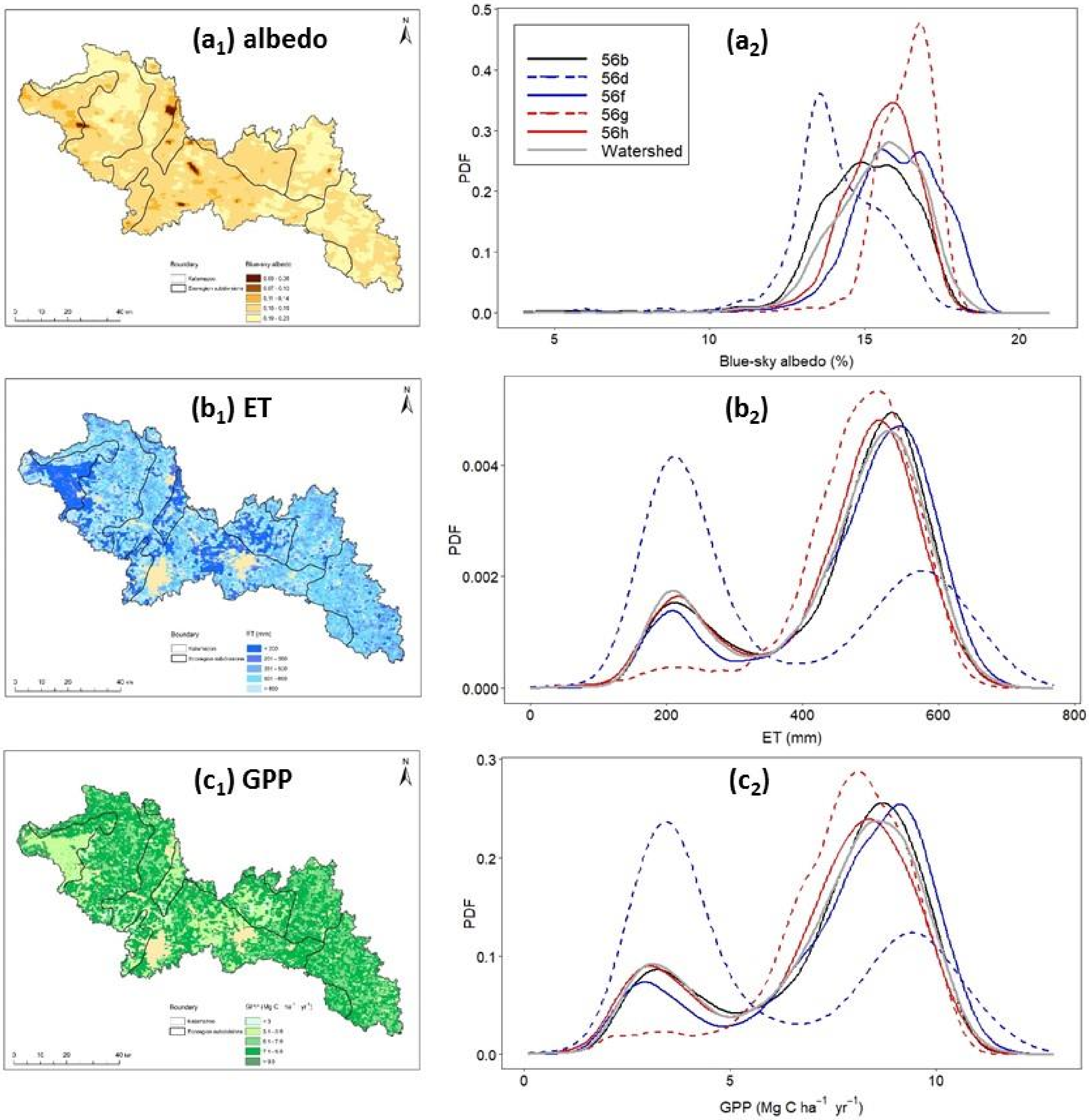

K_i
N
500x500 grid
Build-ups
Barren
Water
Grassland
Cropland
Forest
Wetland
0
500
1,000 m

Fig. 2



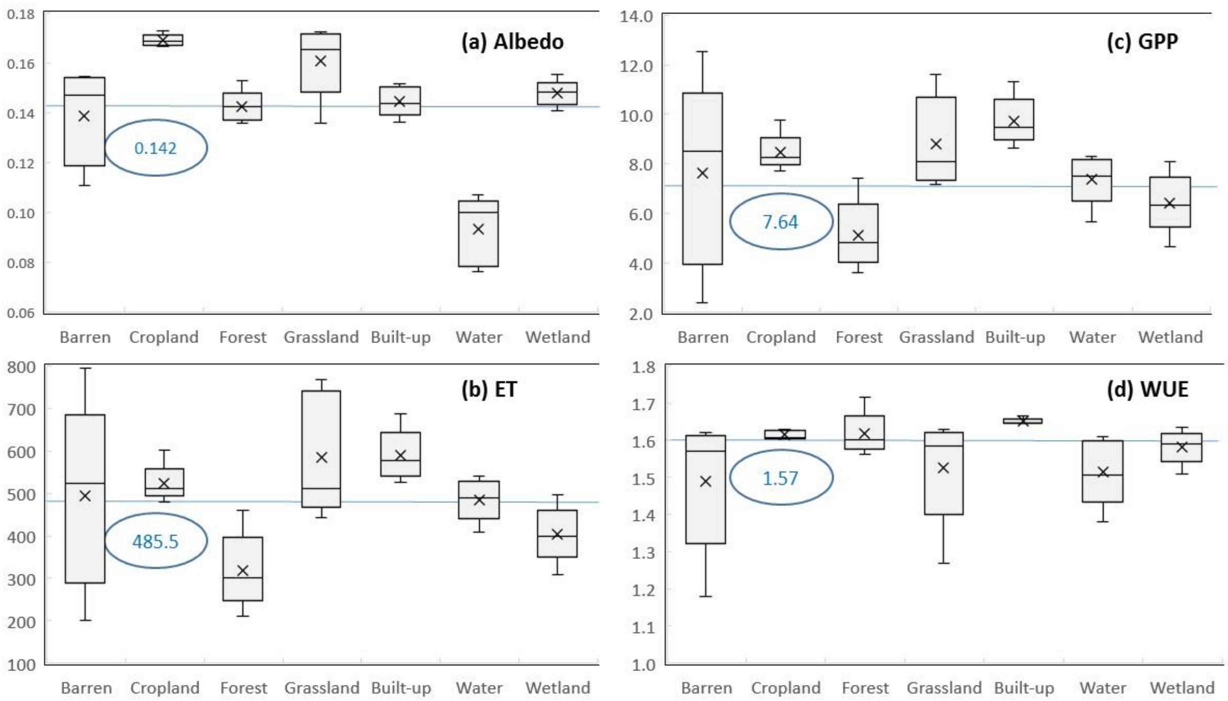
Cover Type	Level IV Ecoregion				
	56b	56d	56f	56g	56h
Built-up	389.3 (21.8)	38.0 (9.1)	91.8 (9.4)	43.6 (7.5)	164.7 (11.0)
Cropland	729.1 (40.9)	87.2 (21.0)	583.9 (59.8)	356.8 (61.6)	719.3 (47.8)
Grassland	24.6 (1.4)	18.9 (4.5)	10.1 (1.0)	3.0 (0.5)	13.8 (0.9)
Forest	383.0 (21.5)	189.5 (45.6)	189.1 (19.4)	81.9 (14.2)	376.0 (25.0)
Water	61.1 (3.4)	15.4 (3.7)	18.5 (1.9)	9.0 (1.6)	24.7 (1.7)
Wetland	184.7 (10.4)	59.7 (14.4)	76.5 (7.8)	80.8 (14.0)	194.1 (12.9)
Barren	8.7 (0.5)	3.4 (0.8)	4.0 (0.4)	1.3 (0.2)	6.2 (0.4)
Total	**1785**	**416**	**977**	**579**	**579**

Fig. 3



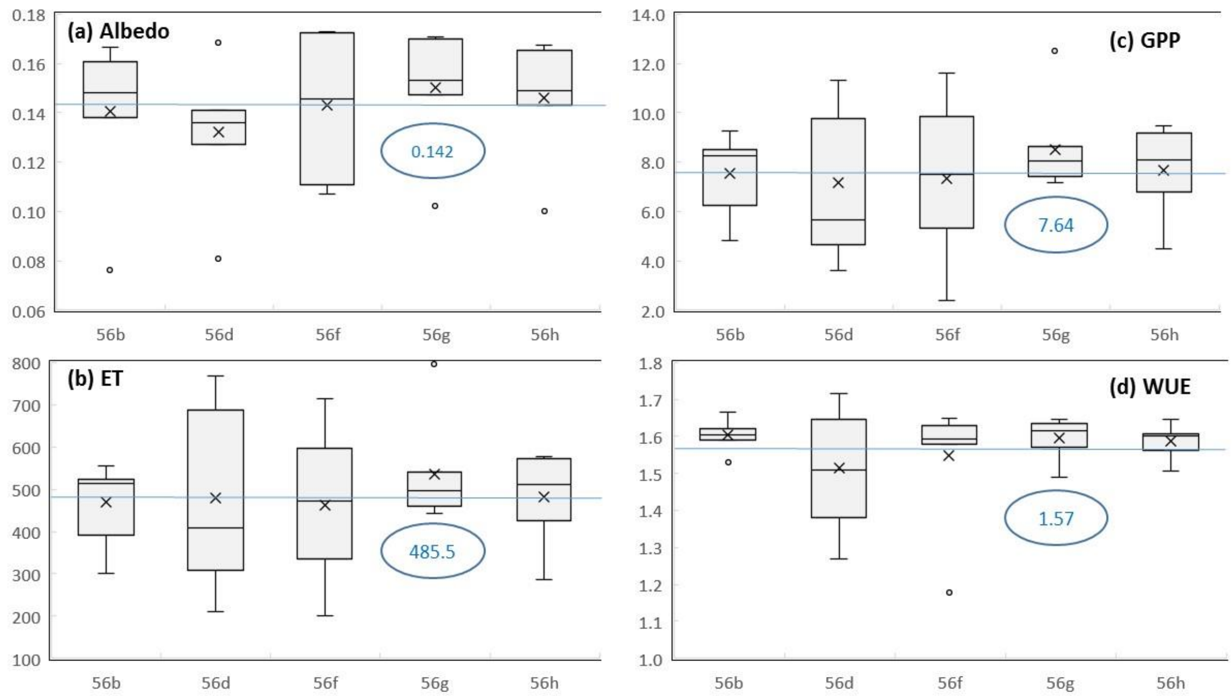
(a$_1$) albedo
(a$_2$)
56b
56d
56f
56g
56h
Watershed
PDF
Blue-sky albedo (%)
(b$_1$) ET
(b$_2$)
PDF
ET (mm)
(c$_1$) GPP
(c$_2$)
PDF
GPP (Mg C ha^{-1} yr^{-1})

Fig. 4.



(a) Albedo
0.142
0.18
0.16
0.14
0.12
0.10
0.08
0.06
Barren
Cropland
Forest
Grassland
Built-up
Water
Wetland
(c) GPP
7.64
14.0
12.0
10.0
8.0
6.0
4.0
2.0
(b) ET
485.5
800
700
600
500
400
300
200
100
(d) WUE
1.57
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1.0

Fig. 5



(a) Albedo
0.18
0.16
0.14
0.12
0.10
0.08
0.06
0.142
56b
56d
56f
56g
56h
(c) GPP
14.0
12.0
10.0
8.0
6.0
4.0
2.0
7.64
(b) ET
800
700
600
500
400
300
200
100
485.5
(d) WUE
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1.0
1.57