Remote Sensing of Environment 150 (2014) 207–217

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

A framework for simulating map error in ecosystem models

Sean P. Healey [a,*], Shawn P. Urbanski [b], Paul L. Patterson [c], Chris Garrard [d]

[a] US Forest Service, Rocky Mountain Research Station, 507 25th St., Ogden, UT 84401, United States
[b] US Forest Service, Rocky Mountain Research Station, 5775 W U.S. Highway 10, Missoula, MT 59808, United States
[c] US Forest Service, Rocky Mountain Research Station, 240 W Prospect Road, Ft. Collins, CO 80526, United States
[d] Utah State University, Logan, UT 84322, United States

ARTICLE INFO

Article history:
Received 23 September 2013
Received in revised form 16 March 2014
Accepted 29 April 2014
Available online 12 June 2014

Keywords:
Monte Carlo
Landsat
Change detection
Carbon

ABSTRACT

The temporal depth and spatial breadth of observations from platforms such as Landsat provide unique perspective on ecosystem dynamics, but the integration of these observations into formal decision support will rely upon improved uncertainty accounting. Monte Carlo (MC) simulations offer a practical, empirical method of accounting for potential map errors in broader ecosystem assessments. However, unless steps are taken across simulations to vary the probability density functions (PDFs) that control simulated map error, the large number of map units to which those PDFs are applied may cause convergence of mean simulated conditions and an artificial reduction of MC estimates of uncertainty. For MC simulation of errors in categorical maps, we introduce a technique we call "PDF weaving" which both: 1) allows variation of PDFs across simulations; and, 2) explicitly aligns the resulting range of simulated populations with estimates and uncertainties identified by traditional monitoring methods such as design-based inventories. This approach is based on solving systems of linear equations and inequalities for each simulation. Each system incorporates linear constraints related to the unchanging distribution of area among classes in the original map (akin to the fixed longitudinal "warp" on a loom) with variable linear constraints related to the class distribution to be simulated in any one iteration (analogous to the perpendicular, variable fibers of the "weft"). Additional constraints specify how many map units to treat as "correct" based on validation exercises at the map unit level. Solution of these systems provides PDFs which will simulate error at both the map unit level and the population level in a way that is consistent with validation exercises and available population-level estimates. We illustrated this approach in an assessment of the effects of wildfire and harvest on carbon storage over 20 years on a forested landscape in the western United States (US). This assessment utilized the Forest Carbon Management Framework (ForCaMF) approach, which is being implemented by the US National Forest System (NFS). Results showed that simulating map error through the use of dynamic PDFs can contribute significant, realistic uncertainty in a Monte Carlo analysis, but that impacts of fire and harvest on carbon storage may nevertheless be clearly identified and differentiated using remotely sensed maps of vegetation and disturbance.

Published by Elsevier Inc.

1. Introduction

Forest management and natural disturbance can have a significant impact on storage or emission of greenhouse gases (Bond-Lamberty, Peckham, Ahl, & Gower, 2007; Kurz & Apps, 1999), and significant additional forest carbon storage is considered achievable with informed forest management (Birdsey, Pregitzer, & Lucier, 2006). The United States National Forest System (NFS), which manages approximately one fifth of the country's forestland (Smith, Miles, Perry, & Pugh, 2009), has a recognized need for information about the link between disturbance and carbon storage. The NFS Climate Change Performance Scorecard (U.S.F.S, 2011), referenced by the National Forest System Land Management Planning Rule (36 CFR Part 219), specifically calls for assessment of "how disturbance and management activities are influencing carbon stocks or carbon sequestration and emissions."

Sample-based field inventories underpin most NFS monitoring, and can generate point-in-time estimates of carbon storage with clear confidence intervals (Heath, Smith, Woodall, Azuma, & Waddell, 2011). Sampling approaches, however, are generally limited in their capacity to describe the effects of relatively rare processes such as disturbance (Masek & Healey, 2012). In the federal forest management context, "state and transition" models are an important alternative source of information about the effects of forest dynamics on forest structure. In these models, the co-occurrence of vegetation variables and disturbance processes is postulated for a particular landscape. Consequences of alternative disturbance trends on variables such as carbon storage are identified. However, the expert assumptions used to calibrate these models can vary significantly (Czembor, Morris, Wintle, & Vesk, 2011),

* Corresponding author. Tel.: +1 011 801 625 5770.
E-mail address: seanhealey@fs.fed.us (S.P. Healey).

http://dx.doi.org/10.1016/j.rse.2014.04.028
0034-4257/Published by Elsevier Inc.

and while such models have proven value in decision support, the uncertainty around a particular result for a given landscape can be difficult to quantify.

Remote sensing with Landsat and similar platforms provides the opportunity to monitor vegetation and disturbance patterns at an extent and resolution that cannot be matched with field measurements. Coupled with appropriate forest dynamics models, remote sensing may offer the decision support that distinguishes "state and transition" models with the fidelity to individual landscapes achieved through designed field samples. Broad application of such methods will largely depend upon their capacity to support assessment of uncertainty, which is indispensable context for any information used in the planning process.

Monte Carlo (MC) simulation methods are a promising approach for quantifying aggregate uncertainty of the many parameters and inputs that often inform ecosystem models. MC methods involve randomly altering the values of key inputs to a complex function such as an ecosystem model, and then characterizing the variation of outputs as an empirical measure of integrated system uncertainty. A new set of input values is "realized" many times to simulate the composite effects of input uncertainties. New values for each variable are chosen randomly using a Probability Density Function (PDF), which quantifies the probability distribution of potential alternative values over a particular range.

MC methods are common in the field of forest carbon modeling, particularly with non-spatial analyses where variables are simple equation parameters. For example, Williams, Collatz, Masek, and Goward (2012) employed MC methods to propagate uncertainty in constituent terms of inventory-derived age-biomass relationships used to model carbon fluxes and levels of biomass. Similarly, Meigs, Turner, Ritts, Yang, and Law (2011) probabilistically varied equation parameters used to calculate net ecosystem productivity at sites that had burned.

Inclusion of input map uncertainty is less common, however. Hudiburg et al. (2009) used maps of forest type to apply equations for biomass stores and net primary productivity across Oregon and Northern California, but MC error simulations were restricted to equation parameters and did not encompass potential map error. Wiedinmyer and Neff (2007) combined fire occurrence maps with 1-km^2 maps of fuel loading to determine combustion emissions for the United States. That study acknowledged the lack of a framework for considering spatial inputs in uncertainty calculations, and instead assigned a level of error based on author experience.

This paper proposes a general approach for MC simulation of map error to fill this gap. The technique presented, which we term "PDF weaving," was developed to accommodate simulation of error in maps with discrete classes, including maps of continuous variables that have been grouped into bins. We assert that the primary obstacle to accurately simulating error in mapped inputs involves the Law of Large Numbers (LLN) and the vast numbers of predictions that can compose both pixel- and polygon-based maps. The LLN states that the mean of a sample will tend toward its expected value as the sample number approaches infinity.

We illustrate potential problems in simulating error over a large number of map units with an example where the mean value of a mapped continuous variable, such as biomass/hectare, is used as a simple proxy for a map's contribution to an ecosystem model. A map's impact within a model may be complex in some cases, but it is likely to be correlated with the mean value (or proportion of population units per class in the categorical case). Consider the population mean ($\overline{P}_j$) realized in an MC simulation, *j*, that is the average of a PDF-based variation function (F_{ji}) applied to a population of pixel-level map predictions (α_i), where the index *i* indicates that the variation function may depend on the map prediction (α_i):

$$\overline{P}_j = \frac{1}{n}\sum_{i=1}^{n} F_{ji}(\alpha_i). \quad (1)$$

In an MC analysis, it is the variance of $\overline{P}_j$ that would quantify system uncertainty. Under reasonable assumptions on the stochastic structure of the variation function (F_{ji}), according to the LLN, the mean of the simulated values will tend toward the value expected from the structure of the PDF:

$$\lim_{n\xrightarrow{P}\infty} \overline{P}_j - E\left(\overline{F}_{ji}\right) = 0, \quad (2)$$

where $\xrightarrow{P}$ indicates convergence in probability. If F_i is constant across iterations, *j*, the central limit theorem suggests that under the above assumptions, estimates of $\overline{P}_j$ will be approximately normally distributed with mean $n^{-1}\sum_i E(F_i)$ and variance $n^{-2}\sum_i Var(F_i)$ (Serfling, 1980).

Thus, if the number of mapped predictions is large and the PDFs underlying the error simulation function, F_i, do not change among realizations (*j*), population-level output parameters (e.g., average biomass) of *even a highly uncertain map may be fairly stable* in an MC analysis. This raises the possibility of understating uncertainty associated with mapped inputs. Working with mapped continuous variables in carbon-related models, French, Goovaerts, and Kasischke (2004) and Gonzalez et al. (2010) observed low levels of landscape-wide carbon output variance when simulating map errors using unvarying PDFs. In both of these analyses, it was observed that pixel-level variations tended to "cancel each other out" across a large landscape, supporting the idea of exhaustive PDF sampling leading to stable population parameters. In a case where categorical spatial inputs (species and age class) were assigned using an unvarying probability function, Xu et al. (2004) likewise observed little difference among MC landscape-level carbon outputs.

It is important to consider sources of error that might simultaneously affect all or most pixels in a map, as would be represented by varying F_{ji} across simulations. Problems with the calibration data (e.g., random inclusion of a disproportionate number of outliers, consistent ground measurement error) or with satellite data (uncorrected sensor degradation, atmospheric artifacts) may cause such problems. Harris et al. (2012) have provided one of the only approaches to date for accomplishing variation of F_{ji} across simulations of remotely sensed input error in carbon models. They used a 2-step function to vary pixel values in an input biomass map. In addition to simulating independent, pixel-level variation centered around the mapped biomass values, they added or subtracted from each pixel's biomass value a fixed quantity drawn for the simulation from a PDF developed using the spread of their model predictions. This latter step was taken to provide "an estimate of the confidence limits around the mean response." Even a small consistent change in simulated values, when propagated over a large number of pixels, can yield a significantly different aggregated result.

This paper presents a strategy for changing the MC variance function for categorical mapped variables in a way which results in controlled changes in the mean response. This challenge is more complex in categorical maps than maps of continuous variables, which can be simply modified by constants drawn randomly for each simulation (as introduced by Harris et al.). Varying the mean response in a deliberate way for categorical maps may require class-specific PDFs controlling the probability of each class changing to each other class (or remaining the same). The effects of those PDFs must "work together" in consideration of the original mapped distribution to allow population parameters (total area per class, for example) to vary in a way that represents those parameters' understood uncertainty (taken from design-based inventories in this paper). Ideally, the PDFs for categorical maps will also reflect validated pixel-level accuracy, altering only the proportion of pixels one might expect to be incorrect from validation exercises.

To identify PDFs that satisfy these requirements, we establish systems of linear equations and inequalities that reflect both the incoming (mapped) and outgoing (to be simulated) distributions of map units among categorical classes. As described later, these systems solve for variables identifying the probabilities of each input–output contingency,

allowing creation of a set of class-specific PDFs that will satisfy randomly selected levels of error in population totals as well as any other constraints placed upon the system.

In an MC analysis that propagates categorical map error, incoming class distributions are set by the original map and are therefore constant across simulations. In the process we propose, output distributions are variable, selected randomly for each iteration. PDFs developed for each simulation must reconcile both the fixed input and variable output distributions. We liken the process detailed in Section 2.4 to weaving because solving for the field of probabilities needed to develop appropriate PDFs involves integrating fixed linear elements (map distributions), similar to the "warp" on a loom, with perpendicular linear elements (targeted output distributions) that are constantly changing across simulations, analogous to the "weft." This is similar to the process of "raking", which begins with preliminary internal values (in this case analogous to individual probabilities for each class to class variation contingency) that do not sum to the desired marginals (e.g., total area per class in both the original map and the targeted distribution). The raking process derives a weighting matrix which is used to achieve additivity (Deville, Särndal, & Sautory, 1993). With weaving, there are no a priori internal values and no need for weights.

In this study, the range of variation in the simulated output distribution (the "weft") is primarily constrained using information from the US Forest Service's Forest Inventory and Analysis Program (FIA), which maintains a nationally comprehensive sample of forest plots and can be used to estimate forest conditions and uncertainties at a range of scales (Reams et al., 2005). The standard errors reported for FIA population estimates covering the entire landscape are used to determine how much variation in class area is to be realized among simulations, which then constrains the PDFs produced by the weaving process. In countries where national forest inventory estimates are not available, or in transnational activities, population parameters could also be derived from more focused field- or image-based inventories (Frescino et al., 2009; Olofsson et al., 2011; Stehman, 2009) or even through professional judgment if necessary.

We illustrate PDF weaving by tracking the uncertainty introduced by a number of remotely sensed maps as they are used in an MC-based ecosystem model called ForCaMF (the Forest Carbon Management Framework). ForCaMF was designed to apply regionally representative carbon dynamics (Raymond et al., in review) to remotely sensed forest dynamics pathways to identify the relative impact of management activities and natural disturbances upon landscape-level carbon stocks. Specifically, these pathways are determined across the landscape through categorical Landsat-based maps of starting volume and forest type, coupled with the Landsat historical record of subsequent disturbances of different types and magnitudes. The goal of PDF weaving is simply to create a framework where simulations involving these types of mapped inputs can accommodate PDFs which change across simulations and which, in aggregate, produce MC input variation that conforms to map validation results and FIA-derived assumptions about population parameters and uncertainties.

ForCaMF analyses are being conducted across NFS (76 million ha) to respond to information needs mentioned above. This paper focuses on the effects of mapped harvests and fires on carbon storage over time in a county in the US state of Montana. While ForCaMF integrates uncertainties from both remotely sensed inputs and the carbon model used, model uncertainties are not simulated in this illustration to focus on the impact of map error. Because the PDF weaving method is generic with respect to categorical input classes and the source of population-level constraints, this illustration may be relevant to a broad range of ecologically oriented MC analyses.

2. Methods

2.1. Study area

The study area is Ravalli County, Montana, a 621,000-ha landscape surrounding the Bitterroot Valley in western Montana, US (114.07° W, 46.14° N; Fig. 1). Forests cover approximately 76% of the area and are primarily coniferous. Elevations range from approximately 1000 m to

Fig. 1. The location of Ravalli County, Montana (gray polygon), the study area. The county is wholly contained within Landsat Path/Row 41/28 (framed in black).

almost 3000 m, with an average of 1790 m. Timber harvest levels have declined significantly since the early 1990s (Spoelma et al., 2008). Wildfire periodically affects the county, with several years since 1985 seeing no appreciable fire, others seeing up to 5000 ha burned, and one particularly acute fire year, 2000, seeing over 80,000 ha burned, almost 20% of the county's forest area.

2.2. ForCaMF carbon model

ForCaMF was designed for use in the US to determine the relative impacts of disturbance, growth, and management upon the amount of carbon stored in forested landscapes over time. ForCaMF may be described as an accounting system which, given categorical Landsat-based maps of starting vegetation conditions and subsequent disturbances, applies regionally average carbon dynamics to track carbon storage or release at the pixel level. As pixel-level carbon stores are aggregated (summed) to the landscape level to provide landscape-level insight, ForCaMF uses a Monte Carlo framework to vary the input map class values used to associate each pixel with particular carbon storage trajectories. Resulting variance in carbon storage model outputs provides an integrated, empirical measure of uncertainty.

Fig. 2 illustrates how ForCaMF tracks carbon accumulation for an individual map unit based upon mapped starting conditions and subsequent disturbances. Landsat-based maps of historical volume (binned into 4 classes) and forest type are used to assign each pixel to an appropriate initial carbon "trajectory" (green color gradient). This trajectory indicates the regionally average (derivation described below) annual change in the sum of major non-soil forest carbon pools: above- and below-ground live and standing dead trees; down dead wood; shrubs; herbs; litter and duff. Production and validation of vegetation and disturbance maps are discussed in Section 2.3. Trajectories are specific to mapped forest type and binned starting volume classes within different disturbance scenarios. In Ravalli County, there were 16 *forest type* × *initial volume* classes: Douglas fir forest types broken into 4 volume bins based upon quartiles of the FIA-measured cumulative distribution (thresholds were 70, 126, and 280 m^3/ha); 4 similar lodgepole pine volume bins (105, 210, 280 m^3/ha); 4 ponderosa/whitebark pine volume bins (84, 168, 280 m^3/ha); and 4 fir/spruce/hemlock/*other* bins (84, 196, 322 m^3/ha).

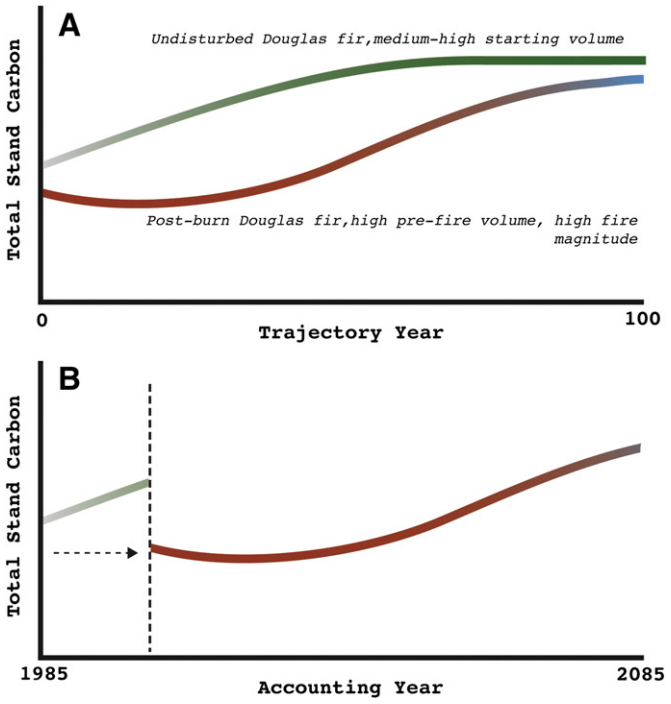

Fig. 2. Illustration of two regionally average 100-year carbon trajectories representing two different combinations of forest types, starting volumes, and disturbance histories (*A*), as well as an illustration of how ForCaMF would account for a change in trajectory from one to the other for a particular map unit due to a mapped fire (*B*). In panel B the undisturbed carbon trajectory has progressed from its initial medium–high volume status (year = 1985) to high volume status by the time of fire disturbance (year = 2000). Carbon storage for all map units is determined as a function of the time since its most recent trajectory assignment. Map unit-level carbon stores are summed within ForCaMF to get population-level estimates for each point in time. Trajectory changes are initiated according to Landsat-based maps depicting the timing, type, and magnitude of forest disturbances and management activities.

The Forest Vegetation Simulator (Crookston & Dixon, 2005) is an extensively calibrated growth model that is widely used for management and planning within the National Forest System (NFS: e.g., Henderson 2008) and is corporately maintained by the US Forest Service. It is an individual-tree, distance-independent forest growth model, and it allows simulation of a wide range of management and disturbance processes. One of FVS' outputs is carbon storage, broken into the pools mentioned above (Reinhardt & Crookston, 2003).

A regionally average carbon trajectory (such as that depicted by the green line in Fig. 2A) is derived for every *species* × *starting volume* bin by combining data from the FIA field sample and FVS. All regional FIA plots are submitted to FVS using the appropriate disturbance keywords, simulating carbon densities (megagrams/hectare, or Mg/ha) at 10-year intervals over 100 years for either undisturbed growth or different intensities of harvests and fires. Within each *starting condition* × *disturbance history* bin, equations are fit to all available projections. Equations are taken to be representative because simulations are based on field-measured tree lists gathered as part of FIA's simple random sample (Bechtold & Patterson, 2005). In this study, 623 FIA plots were available from the "Inland Empire" region which contains Ravalli County, and each bin was represented by a minimum of 5 simulations from independent tree lists. Raymond et al. (in review) describe derivation of both these equations and accompanying functions which can be used in ForCaMF to account for model error; these latter functions are ignored in this study to isolate the effects of map uncertainty.

A series of pixel-based maps is developed to identify areas where trajectories of undisturbed growth are interrupted by events such as harvests and fires. As detailed in Section 2.3, these maps indicate the location, type, timing, and magnitude of disturbances across the landscape. For map units showing a disturbance, carbon storage following that disturbance is altered in the accounting system by accessing the FVS-based carbon trajectory (e.g., the red line in Fig. 2a) associated with disturbance parameters matching the map. A map unit's *starting volume* value and bin can change at the time of disturbance to account for growth since the start of the study period. Volume is re-calculated as a function of the total stand carbon assigned to the map unit by ForCaMF at the time of disturbance, with this function representing the species-specific, plot-level relationship between the two variables, as observed on FIA's plot network.

Fig. 2b shows how ForCaMF accounts for a pixel that is burned during the study period. Up to the point of the burn, the area's carbon density is based upon regional average "undisturbed" trajectories associated with initial conditions in the forest type and volume maps. Following the fire, carbon storage for the pixel is assigned to a post-burn trajectory that represents the average carbon storage for stands affected by the mapped disturbance type and magnitude and containing the pre-burn level of volume calculated by ForCaMF.

Carbon accounting for each map unit, along with the probabilistic steps described in Section 2.4, is controlled with software written in the C# programming language. Model parameters and simulation results are stored in a PostgreSQL database. To reduce the computational burden within ForCaMF, both Monte Carlo variance of map values and the assignment of appropriate carbon accumulation functions were applied using aggregated 10-hectare simulation units (SUs). These units were composed of groups of 111 pixels sharing exactly the same

starting vegetation and disturbance map bins, but which may not have been spatially contiguous. This compromise in simulation complexity was imposed by local computing limits, but nevertheless resulted in approximately 50,000 SUs in Ravalli County.

Summing the projected carbon storage over all of the SUs making up a landscape (in this case, Ravalli County) produces a landscape-level estimate of carbon storage over time under the mapped disturbance scenario. Alternative (hypothetical) disturbance maps may be entered into ForCaMF to highlight the sensitivity of the system to patterns in phenomena such as harvest and wildfire. In this study, the overall impact of harvest and fire on carbon storage in Ravalli Country was assessed by sequentially running ForCaMF with altered disturbance data, showing results of: all observed disturbances (fires plus harvests); no disturbances; just harvests; and just fires. This was accomplished simply by changing disturbance codes in the database. In scenarios for which mapped disturbances were disregarded, affected SUs were kept on "undisturbed" carbon accumulation trajectories (e.g., green line in Fig. 2a). In this study, temporal resolution of disturbance patterns was limited by availability of clear, anniversary-date Landsat imagery (Table 1), so comparisons of carbon storage among disturbance scenarios were made only in years for which such imagery was available.

Map inputs were varied in this study across 1000 realizations, following protocols described in Section 2.4, and the standard deviation of outputs was taken as an empirical measure of the effect of map error on output uncertainty. One thousand simulations were considered sufficient for representing potential errors in ForCaMF because of results of an independent trial conducted using the nearby Flathead National Forest as a test case. That trial used increasingly large random subsets of total carbon storage simulations (increasing by intervals of 100 up to 1000) to investigate the effect of additional simulations on result stability. Little change in mean carbon storage output estimates was seen beyond using subsets of 400 simulations, at which point mean deviation from the average of 1000 simulations declined to about 0.10%.

2.3. Production and validation of input maps

This section describes both the layers derived from remotely sensed imagery used here as ForCaMF inputs and the validation activities used to constrain uncertainties built into the Monte Carlo operations described in the next section. A mostly biennial time series (Table 1) of Landsat imagery for World Reference System 2 Path/Row 41/28 from 1985 to 2005 was acquired and processed to surface reflectance (Masek et al., 2006), which allowed volume and cover models calibrated using recent (2001–2008) FIA ground data (described below) to be applied to all dates in the series (Nelson, Healey, Moser, Masek, & Cohen, 2011). Mapping of the two 1985 starting-point variables (volume and forest type group) occurred in two stages: creation of initial pixel-level estimates followed by post-process "stretching" of the maps' histograms to match the distribution of FIA measurements from the period.

The initial volume map was created using a non-parametric Random Forests (Breiman, 2001) modeling approach linking 94 recent (2001–2008) FIA plot measurements from Ravalli County with topographic predictors and contemporaneous Landsat satellite imagery. The resulting model was applied to 1985 imagery to create an historic volume map. This calibration of an historical map using contemporary field measurements was necessary because of the relatively low quality of plot coordinates measured in the 1980s and early 1990s (Hoppus & Lister, 2005). Details of this approach to historical mapping were described by Powell et al. (2010). The histogram of this map was then "stretched" to match the distribution of FIA plot measurements of volume from the period (nominal date: 1989; http://apps.fs.fed.us/fiadb-downloads/datamart.html; all web references last accessed February, 2014). This was done using the "histogram match" algorithm in ERDAS Imagine, which utilized a mathematically derived lookup table to alter pixel values so that their proportional frequency distribution matched the dimensions of the FIA sample. It must be emphasized that this stretching procedure was not required for PDF weaving; agreement with FIA estimates was simply a locally desirable attribute of these maps.

The initial forest type map was subset from a national-scale map (Ruefenacht et al., 2008), aggregated to 4 principal forest type groups (Douglas fir, lodgepole pine, ponderosa pine, and an "other" category primarily composed of subalpine fir, Engelmann spruce, and aspen). This map was then re-sampled from 250-m resolution to 30-m and was submitted to a GIS process which iteratively re-classified isolated and boundary pixels using a 3×3 majority filter, such that the mapped area of each forest type areas matched FIA survey results (http://apps.fs.fed.us/Evalidator/evalidator.jsp).

Initial mapping of the location and timing of fires and harvests on the landscape was achieved with the above-mentioned Landsat time series using supervised classification under a multi-temporal composite approach (sensu Coppin & Bauer, 1996), as described by Healey et al. (2008). Intensive, post-process manual correction of these maps was carried out, using the Landsat time series itself and historical images served through Google Earth to digitize omitted disturbances and to remove false-positive changes. At the same time, separation of fires from harvests in the disturbance map was achieved visually and using a national fire database (Monitoring Trends in Burn Severity; (Schwind, Brewer, Quayle, & Eidenshink, 2010)).

Magnitudes of fires and harvests were assessed separately. Estimation of harvest magnitude in terms of timber volume removal was derived through a 2-step process. First, the Random Forests model used to initiate the starting-point volume map was applied to all Landsat images in the time series. Modeled differences in volume from before and after events mapped as harvests in the initial disturbance map were used as a first approximation of harvest intensity (following (Healey, Yang, Cohen, & Pierce, 2006; Nelson, Healey, Moser, & Hansen, 2009)). These initial estimates of volume removal were then scaled to match administrative harvest records for the county (reported nationally by FIA's Timber Product Output unit: www.fia.fs.fed.us/program-features/tpo/), such that the sum of mapped removals equaled the independent county-level harvest record in each year.

For fires, event magnitude was represented by mapped percent canopy cover loss. Cover change was seen as more responsive to fire intensity than volume loss because much of a stand's timber volume remains even following the most destructive fires. As with volume, percent cover was modeled and mapped using Random Forests at every date in the image time series. Models were calibrated at locations of 514 FIA plots using cover estimates derived from the inventory tree list by the FVS percent cover algorithm (Crookston & Stage, 1999). Because image values were needed for this modeling, only the subset of FVS "Inland Empire" plots within the p41/r28 footprint could be used. This included plots not within Ravalli County. Disturbance magnitude measures of both volume removal and cover loss were derived by dividing the difference between mapped pre- and post-event values by the pre-event value. Results were binned into quartiles of percentage loss: 0–25%; 25–50%; 50–75%; >75%.

As described in the next section, PDF weaving develops class-specific probability functions that produce MC realizations conforming to both assessed map unit-level confusion metrics and randomly drawn overall

Table 1
Date and sensor of Landsat imagery (Path 41/Row 28) used in this study.

Date	Sensor	Date	Sensor
June 19, 1985	TM	August 21, 1997	TM
August 7, 1986	TM	August 8, 1998	TM
July 27, 1988	TM	July 26, 1999	TM
July 20, 1991	TM	August 8, 2001	ETM+
August 10, 1993	TM	July 21, 2003	TM
July 31, 1995	TM	August 11, 2005	TM

distributions of the population among classes (constrained by inventory-derived standard errors). Pixel-level accuracies of the starting (1985) volume map, the forest type map, and the disturbance magnitude maps were assessed at FIA plot locations not used in building the models that generated the maps.

Pixel-level accuracy of the 1985 volume map was assessed with an independent test set (n = 44) of FIA plots from Ravalli County. As with calibration of the volume model, the lack of high-quality geo-location of FIA plots in the 1980s led to validation being carried out using FIA measurements collected between 2001 and 2008. Both predicted and observed volume values were binned using the species-specific thresholds listed earlier to align the assessment with how the layer was to be used within ForCaMF. Pixel-level validation of the forest type group map was conducted at the locations of 89 available FIA plots in Ravalli County. More FIA plots were available for validation of the type map than the volume map because the map's external production required no segregation of model-building and validation plots. As with the volume map, the matching of map and FIA estimates (in this case, using the majority filter process described above) allowed FIA standard errors (from the above FIA web interface) to be taken as standard errors of map estimates.

The map of disturbance timing and location was treated as error-free because of the intensive post-process manual map correction that made use of data sources commonly used to create reference data in map error assessments (e.g., (Cohen, Yang, & Kennedy, 2010; Thomas et al., 2011)). The percentage cover loss disturbance magnitude maps, however, were treated as variables in the MC analysis. Unfortunately, FIA plots in the western half of the U.S. have not all been re-measured; in fact, at the time of this study, not all of FIA's plots in Montana (of which 10% are measured each year) had been measured even once. Options for assessing pixel-level error rates when no case-specific reference data exist may include appealing to validations of similar maps (e.g., (Healey et al., 2006)) or to expert opinion.

The option selected here was to evaluate predicted versus observed differences in cover (binned into the percentage quartiles used in the maps and FVS work) over different locations at the same time instead of tracking the same locations at different times. Fifty plots were left out of the cover-modeling process, from which 1225 total pairwise comparisons were possible ($n! / k!(n - k)!$, where n = 50 test plots and k = groups of 2). Here, the fraction of pairwise differences predicted correctly (after binning into percentage removal quartiles) was taken as an estimate of overall accuracy. Taking the single-date pairwise comparisons as a proxy for true multiple-date assessments of cover change in the calculation of error rates depends upon an implicit assumption that the distribution of differences among classes in the single-date case is representative of the class distribution in the population of the disturbed pixels. We acknowledge that this assumption may not be justified, and future applications of ForCaMF may instead rely upon expectations of accuracy drawn from the literature.

In addition to the pixel-level accuracies needed for the PDF weaving technique described below, estimates and uncertainties of the distribution of area among classes in each map are required. For the volume map, the bins used to convert our continuous predictions of volume into categorical mapped classes are not strata for which standardized FIA variables are available. However, since our mapped predictions were "stretched" to the distribution of values measured in the historical FIA sample, mapped areas of each binned class were taken as a reasonable proxy for an FIA design-based estimate. These map-based estimates, however, did not provide for estimation of standard errors. Per-class standard errors, as a percentage of estimated area, were instead taken from FIA's estimates of forest area by four stand size classes: below 12.7 cm; 12.7 to 22.6 cm; 22.6 to 50.5 cm; and larger than 50.5 cm. Stand size is determined by plurality class (using the above ranges) of diameter at breast height (dbh) measurements for non-overtopped live trees. FIA's standard error estimates for these classes in Ravalli County (http://apps.fs.fed.us/Evalidator/evalidator.jsp, accessed 8 April, 2013) were: 19%; 18%; 6%; and 30%. These values reflected the general level of sampling error of the FIA sample, and their use demonstrated the flexibility of the PDF weaving process. It is acknowledged, however, that since the thresholds for our volume map bins were intended to evenly distribute the FIA sample into quartiles (unlike the Stand Size Class variable), an average standard error from the above values may be more representative of the actual class-level uncertainty. Further, use of 2013 uncertainties may not reflect class distributions at the time of the mapped distribution in the early 1990s, although the overall sample number has remained fairly constant.

Deriving population-level uncertainty around the mapped ("stretched") estimates for the forest type map was more straightforward since forest types were easily linked to FIA variables. Historical (nominal date: 1989) FIA values used were (again, in standard error as a percentage of the estimated area): 14% for the fir/spruce/hemlock/other class; 15% for the lodgepole pine class; 21% for the ponderosa pine class; and 10% for the Douglas fir class. Because FIA plots have not been re-measured in this part of the country, no inventory-based population estimates of harvest or disturbance magnitude were available. For purposes of this analysis, standard errors of 35% were assigned to the area of each disturbance magnitude class. This arbitrary level of uncertainty was likely conservatively large, particularly since harvest magnitude maps were explicitly linked to TPO records, which are considered authoritative.

2.4. Linear systems for linking validation data and MC simulations

Three ForCaMF input maps – forest type, starting volume, and disturbance magnitude – were treated as variable for the purposes of MC simulation. Methods are described above for deriving area estimates and standard errors associated with the 4 classes in each of these maps. Normally distributed PDFs were created for each class in these maps, with the mapped area estimate (as a proportion of the landscape) defining the mean, and the standard error of the population estimate (derived above) defining the standard deviation. In the absence of the "stretching" process linking FIA estimates and map totals, the estimates themselves could have been used to define these PDFs. At the beginning of each MC realization, a value was randomly drawn from the PDFs of the 3 least common (by pixel count) classes, with the area of the largest class in each map determined by the total area minus the sum of the other 3. The latter step ensured simulation of the same amount of forest in every realization.

Simulation of the randomly selected output areas for each class (and by extension, a particular level of population-level error) required integration of those population-level targets with the PDFs which govern MC alteration of mapped values. The fundamental assumption in our approach is that the probability functions we derive, controlling the iterative shifting of map values, will produce their long-term expected distribution of values when applied over a large number of map units. This assumed equivalency between probabilities and proportional distributions allows us to implement pre-determined shifts among classes by altering PDF parameters.

With both the input (mapped) and desired distribution of the 4 classes across the landscape known for a given realization, we needed to determine the probability $P(X_r \mid X_m)$ of any pixel in the map having mapped value X_m and then being "realized" as possible class X_r in a given simulation. After the relative probability of each of these possible transitions is determined, development of PDFs becomes straightforward.

Fig. 3 shows the 16 relative probabilities (equivalent to proportions of the entire map) that need to be determined for a 4-class mapped variable. Row totals are determined by the map's histogram, while column totals represent the randomly drawn area of each class to be represented in the simulation. The challenge is to derive the 16 internal probabilities while both maintaining known row totals and targeting particular column totals. It may also be advantageous to constrain probabilities using pixel-based map validation metrics such as overall accuracy

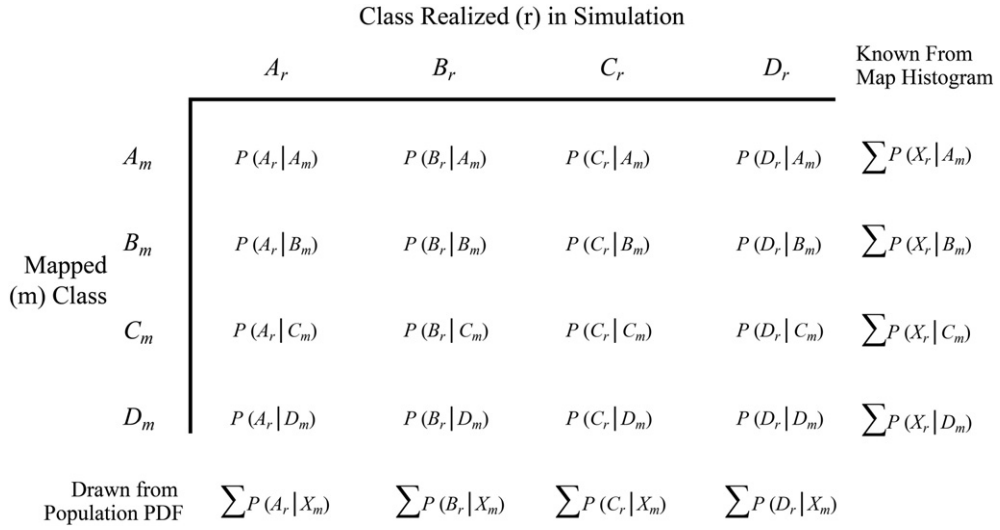

Fig. 3. Schematic layout of a system of linear equations used to solve for contingency probabilities needed in the PDFs developed in Fig. 4. The 4-class case is illustrated, in which 16 variables describe distribution of all map units among contingencies ($X_r \mid X_m$) among mapped values (X_m) and "realized" values (X_r). Probabilities are assumed to be equivalent to proportional distributions across large populations of map units. All probabilities sum to 1.0. Row sums are set equal to the mapped distribution among classes, and column sums are set equal to totals representing a random re-distribution among classes.

(setting, for example, a target for the fraction of map units ending up along the diagonal in Fig. 3).

The needed probabilities may be specified as variables to be solved for as part of a system of linear equations and linear inequalities. The expressions defining this system may be visualized from Fig. 3. Row totals (e.g., $P(A_r|A_m) + P(B_r|A_m) + P(C_r|A_m) + P(D_r|A_m)$) must equal the proportional distribution among classes from the map. Since the row totals, expressed as a fraction of the entire map, sum to 1.0, this indirectly ensures that the 16 variables will sum to 1.0 as well. Column sums (e.g., $P(A_r|A_m) + P(A_r|B_m) + P(A_r|C_m) + P(A_r|D_m)$) represent another set of equations defining the system. These sums are set equal to the randomly drawn proportional areas for each class in the given MC realization. In addition to row and column sums (8 separate equations in the 4-variable case), the diagonal sum (e.g., $P(A_r|A_m) + P(B_r|B_m) + P(C_r|C_m) + P(D_r|D_m)$) may be used to specify an assumed overall level of accuracy, which we took from pixel-level validation activities for each map.

Statements of inequality may also be used to further refine the set of possible inputs for the simulations. We introduced inequalities relating assumptions that small errors would be more common than large errors in maps having ordinal classes (volume and disturbance magnitude, but not forest type). Specifically, we built ordinal relationships between classes into the probability structure by including inequations such as: $P(B_r|A_m) > P(C_r|A_m)$.

A feasible solution to the above systems of linear equations and inequalities was found by invoking the hybrid local search (HLS) solver provided in the Microsoft Solver Foundation 3.1 within ForCaMF's C# code. This solver also accommodates non-linear systems, and it is possible that other approaches may outperform HLS in solving these linear systems, particularly in terms of processing time. In cases where HLS

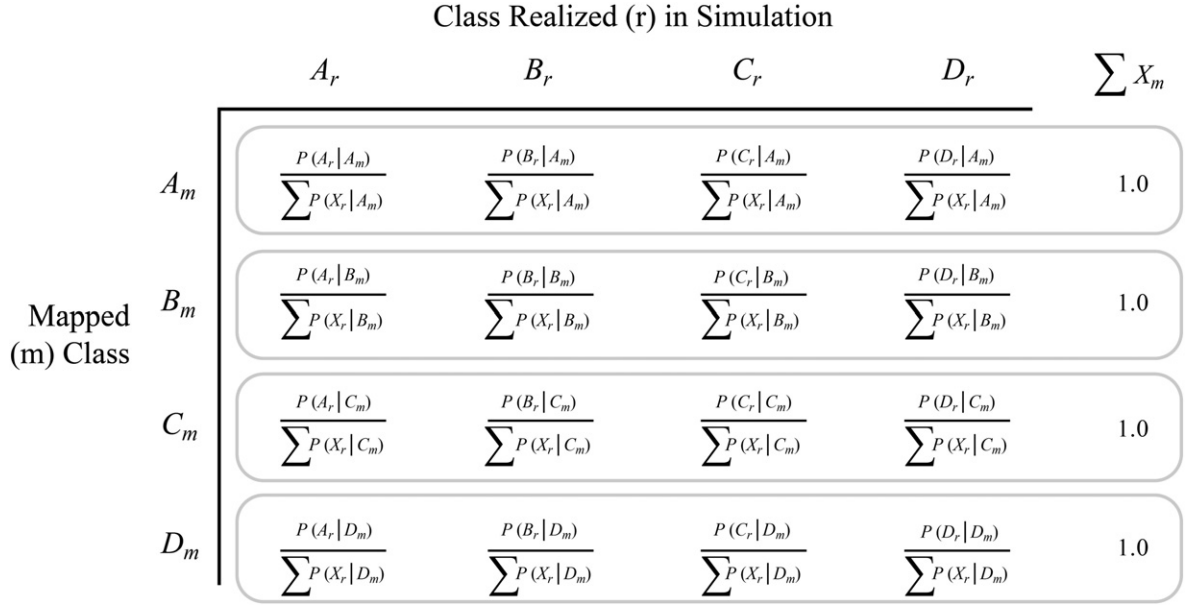

Fig. 4. Schematic showing determination of probabilities for 4 class-specific PDFs following solution of the linear system described in Fig. 3. Probability distributions are discrete among the 4 potential map classes.

Table 2
Assessment of mapped volume class values in Ravalli County, as measured at 44 sites not used for model-making. Per-hectare volumes were grouped according to thresholds described herein.

		Observed volume class				
		Low	Medium low	Medium high	High	% correct
Predicted volume class	Low	10	0	0	1	91%
	Medium Low	2	2	2	0	33%
	Medium High	1	2	6	2	55%
	High	1	2	4	9	56%
	% correct	71%	33%	50%	75%	61%

found no solution to the system as specified in a given MCA realization, the ForCaMF software incrementally relaxed the system's diagonal sum (i.e. the simulated overall map accuracy) by turning the linear equation for that sum into two linear inequalities bounding a range of gradually increasing acceptable values. This range first targeted diagonal sums within 1.0% of the original value, then 2.0%, 4.0%, and 8%. If no solution was found within an 8.0% relaxation of the diagonal sum, the randomly drawn target distribution was discarded. Metrics were kept about the frequency of this "relaxation" process and the rate of discarded sets of equations.

Once the needed distribution of map units among transition contingencies ($X_r \mid X_m$) is known, class-specific ("row-wise") PDFs can be derived to control the alteration of map unit values. Specifically, discrete probability functions are developed for each mapped class, describing probability of transition to each alternative class, as illustrated in Fig. 4. These were the PDFs applied in each MC realization to the starting forest type and volume values of each SU, as well as to the disturbance magnitude bins of disturbed SUs within the ForCaMF accounting system.

3. Results

3.1. Validation results

Pixel-level error rates were assessed for the 3 maps submitted to MC analysis (volume, forest type, magnitude change). Tables 2-4 present pixel-level agreement matrices comparing these 3 maps with FIA ground measurements. The MC analysis we describe used the most basic accuracy parameter from these assessments: the fraction of pixels mapped correctly, shown on the tables' top-left to bottom-right diagonal. These overall accuracies for these 4-class maps were: 0.74 for forest type group; 0.60 for cover change; and 0.61 for initial volume class. The population-level uncertainty assessed for each class in these maps in the PDF weaving process was described as background in Section 2.3.

3.2. Simulation Results

The potential pixel- and population-level errors simulated in each realization of the MC analysis were parameterized by the validation activities above. Before presenting results of the simulations, it is important to consider how often a solution – e.g., values for the 16 probabilities identified in Fig. 3 – could actually be found for the linear systems we have described. Specifically, these solutions were required to represent specific randomly drawn levels of population-level error as well as constraints related to overall accuracy (diagonal total) and, for ordinal maps, constraints requiring that big errors be less common than small errors.

As described above, ForCaMF was programmed to gradually "relax" the constraint on overall accuracy (represented by the diagonal sum) until a solution was found. For the forest type map, 94.4% of realizations required no parameter relaxation, and the rest (5.6%) were solved with relaxation of the diagonal total by 1%. Error functions were derived for the volume map with no relaxation 98.5% of the time, and with relaxation by 1% in 1.5% of the realizations. Solutions for the linear systems used for the magnitude maps were less easily solved: 84.4% required no relaxation; 13.2% required relaxation of the diagonal total constraint by 1%; 0.2% by 2%; and 0.2% by up to 8%, after which 2.0% of systems were deemed "unsolvable" and discarded. This higher difficulty of solving systems for change magnitude may have derived from the conservatively high standard error (35% of the mapped area for each class) used to develop the PDFs controlling the target distribution (the "weft") in each simulation.

The first of two MC analyses simply summed simulated carbon stocks for all forested map units under alternative disturbance scenarios. These scenarios included: 1) all mapped fires and harvests; 2) just mapped fires; 3) just harvest; and 4) no disturbances. Results of these simulations are presented in Fig. 5, with the error bars representing the standard deviation of landscape-level non-soil forest carbon storage over 1000 realizations.

The "harvest only" simulations show greater immediate deviation from the "no disturbance" dynamics than the "fire only" simulations. However, the "fire only" stocks abruptly drop below the "harvest only" levels of storage following the widespread fires in the year 2000. The standard deviation intervals of both the "fire-only" and "harvest-only" scenarios overlap with the "no-disturbance" interval in all time periods. The interval of the "fire-plus-harvest" scenario does separate from the no-disturbance scenario, but only after the fires of the year 2000. This relative inseparability was addressed in the second MC analysis, which was based upon the same set of simulations. Instead of plotting the distribution of overall carbon storage estimates for each scenario, we plotted the distribution of *differences* among scenarios across the 1000 sets of simulated errors.

Table 3
Assessment of mapped forest type group classification in Ravalli County using field data from FIA. Types include: Douglas fir (PSME); Fir/Spruce/Hemlock/other (F/S/H/other); lodgepole pine (PICO); and ponderosa/white pine (PIPO/PIAL).

		Observed forest type				
		PSME	F/S/H/other	PICO	PIPO/PIAL	% correct
Predicted forest type	PSME	25	4	2	0	81%
	F/S/H/other	1	12	5	0	67%
	PICO	2	3	9	0	64%
	PIPO/PIAL	1	0	0	6	86%
	% correct	86%	63%	56%	100%	74%

Table 4
Differences between mapped and observed percent cover values, as measured at all possible pairwise comparisons of an independent test set of 50 cover estimates. "Observed" cover was generated through FVS for tree lists measured by FIA.

		Observed cover difference				
		0–25%	25–50%	50–75%	75–100%	% correct
Predicted cover difference	0–25%	279	62	30	15	72%
	25–50%	150	65	51	20	23%
	50–75%	65	32	42	53	22%
	75–100%	13	3	0	345	96%
	% correct	55%	40%	34%	80%	60%

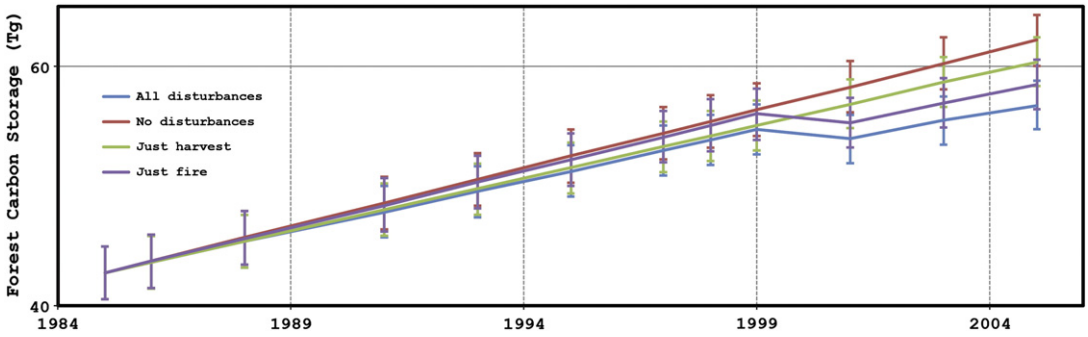

Fig. 5. Estimated carbon storage in all non-soil forest pools in Ravalli County, 1985–2005. Lines represent alternative disturbance scenarios, as implemented in ForCaMF. Error bars represent the standard deviation of estimates over 1000 simulations.

Simulated error in the volume, forest type, and disturbance magnitude maps caused significant variation in predicted levels of carbon storage, as shown in the error bars in Fig. 5. However, across levels of simulated error, there were relatively constant differences in projected storage among disturbance scenarios. This is manifested in the relatively low standard deviation of differences among scenarios in carbon storage across simulations (with one PDF weaving solution per simulation applied to all disturbance scenarios). Fig. 6 represents a simple reframing of the analysis presented in Fig. 5; estimates of absolute carbon storage are abandoned in favor of identifying the difference that alternative disturbance scenarios make in the context of map uncertainties. From that point of view, there is a high level of discrimination among disturbance scenarios from the beginning of the study period.

4. Discussion

Remote sensing can and does provide critical monitoring information for ecological models, commonly including surfaces related to biological productivity and representing vegetation structure and change. While a few efforts have been made with MC methods to account for map error in the uncertainty of model outputs, there has to date been no generic approach proposed for doing so, particularly for maps of categorical variables. PDF weaving allows simulation of map error that accommodates two important considerations: variance of PDFs across simulations and alignment of the range of simulated conditions with reference data at both the population and map unit levels.

As discussed earlier, there is reason to believe that use of an invariant PDF over large numbers of map units can suppress variance in the mean response across simulations, possibly leading to understatement of MC uncertainty. Our results indicate that use of variable PDFs can have a significant impact on assessment results. For example, there is frequent overlap in simulated landscape carbon levels (Fig. 5) between a "no disturbance" scenario and one which includes historical fires — even following a year of extreme fire activity. This overlap is entirely due to input map uncertainty; ForCaMF mechanisms for simulating model error were deactivated for this study. Effects of map error are likely to vary by algorithm and landscape, but this analysis suggests that ignoring such error may result in substantial understatement of output uncertainty.

However, if the added MC variance were either arbitrary of uncontrolled, it would not serve to realistically simulate map uncertainties any more than the artificially reduced variance that may result from the use of an invariant PDF. Map validation activities generally occur at the map unit level, with predicted versus observed values compared at a selection of reference points. Such information (e.g., Tables 2–4) is highly informative regarding the probability of confusion among different mapped classes, but by themselves such assessments are not typically a framework for directly inferring map uncertainties at the population level.

The harmonization of this map-unit-level validation information with population parameters and uncertainties from sources such as FIA may be the defining utility of PDF weaving. The "histogram-stretching" map-building step practiced here ensured that mapped population estimates agreed with FIA in this study. However, even if significant differences existed between the FIA- and map-based proportional distributions of areas (i.e. even if there were large differences between class-specific row and column totals), PDF weaving would ensure that MC simulations agreed with FIA at the population level. Since designed field samples such as FIA's form the basis of traditional monitoring efforts around the world (Brown, 1997), congruence with

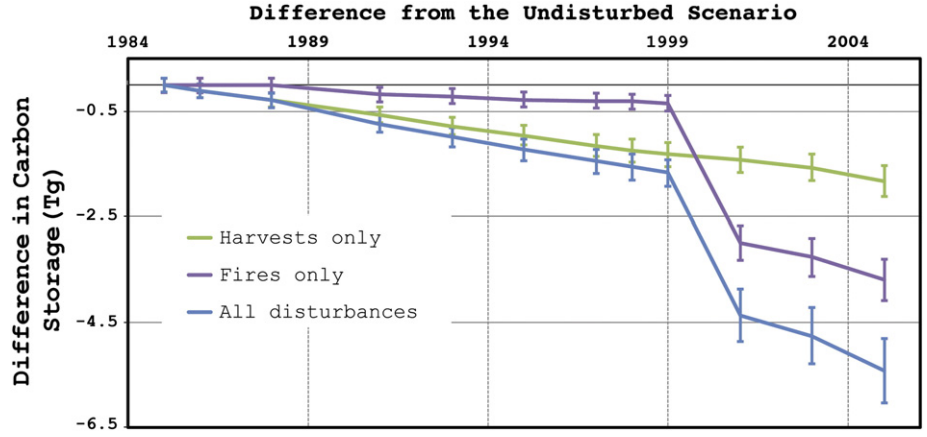

Fig. 6. Differences in landscape-level carbon storage between alternative disturbance scenarios in Ravalli County and a "no disturbance" scenario. Differences were calculated at the level of the MC realization within ForCaMF, keeping simulated error rates constant among scenarios within each of 1000 repetitions.

inventory data at this level is likely to enhance MC analysis credibility and increase compatibility with assessments of other resources such as timber or habitat.

Despite the uncertainty introduced into the ForCaMF analysis by potential map error, assessment of the relative impact of harvest and fire was possible (Figs. 5 and 6). Harvests in Ravalli County since 1985 have had a more consistent impact on landscape-level carbon storage than fires, with "harvest only" storage patterns showing more immediate and steady divergence from "no disturbance" storage levels than "fire only" scenarios. This consistent effect matches administrative records; Spoelma et al. (2008) reported sustained and significant levels of harvest in Ravalli County during the study period, despite some year-to-year variation and a reduction in the cut starting in the mid-1990s. While typical impacts of fire are lower (with multiple periods with no registered wildfire), extreme fire events such as those observed in the year 2000 can have a greater long-term impact on carbon storage than the cumulative effect of removal by harvest.

In general, the FVS simulations from which ForCaMF's carbon dynamics are derived show a more gradual release of carbon from fire-killed trees than from harvested trees. This is consistent with field studies showing that while high levels of carbon in litter and small trees can be released through initial fire consumption, the majority of carbon in trees is not immediately consumed (Campbell, Donato, Azuma, & Law, 2007; Meigs, Donato, Campbell, Martin, & Law, 2009). Unlike harvest, fires can leave significant amounts of newly dead material on site, decaying and mitigating the effect of subsequent re-growth on overall stand carbon for decades (Kashian, Romme, Tinker, Turner, & Ryan, 2006). In this context, even if fire ceased to occur past 2005, ForCaMF's "fire only" carbon storage trajectory may continue to diverge from the "undisturbed" scenario.

Our results show that, rather than trying to infer differences among scenarios on the basis of independent carbon storage estimates (Fig. 5), it may be more effective to tabulate differences among scenarios directly at the level of the MC realization. For instance, simulated population-level error in starting volume and forest type maps may cause significant variation in the amount of carbon predicted for the landscape, as is apparent in Fig. 5. For each realized level of volume error, though, the *difference* in predicted carbon storage among alternate disturbance scenarios may remain relatively stable. Fig. 6 shows that such differences (framed in the figure in terms of difference from a "no disturbance" baseline) are indeed consistent across realizations in Ravalli County, resulting in better separation of scenario effects. This approach effectively abandons estimation of total carbon in favor of greater acuity in comparisons among disturbance scenarios and management approaches.

In many applications, it is this understanding of relative differences among scenarios that is paramount. The concept of "additionality" built into agreements such as the Kyoto Protocol requires that benefits of a particular management strategy be demonstrated in relation to some "business as usual" baseline (Moura Costa, Stuart, Pinard, & Phillips, 2000). Similarly, NFS monitoring guidance regarding carbon dynamics discusses the relative effects of alternative management and disturbance processes (U.S.F.S, 2011),

5. Conclusion

We illustrated how MC simulations of map error may be aligned with reference data at both the map unit and population levels. The importance of map error as a source of uncertainty was underscored by the significant model output variability resulting from simulation of map errors. Despite this variability, however, Landsat-derived model inputs provided unique information which allowed the quantitative effects of different types of disturbance to be identified and compared. Such comparisons, along with the uncertainty assessments provided by Monte Carlo error simulations, address critical information needs that have been articulated by forest managers such as the US National Forest System.

Acknowledgments

This work was made possible by a grant (10-CARBON10-0082) from NASA's Applied Sciences Program. Jim Morrison of the National Forest System provided clear guidance related to needed resource information, as did David Cleaves, Greg Kujawa, and Elizabeth Reinhardt of the Forest Service Climate Change Office. Extensive support was also provided by the Forest Service Interior West Forest Inventory and Analysis Program. Many thanks are also due Douglas Ramsey of Utah State University.

References

Bechtold, W. A., & Patterson, P. L. (Eds.). (2005). *The Enhanced Forest Inventory and Analysis Program — National Sampling Design and Estimation Procedures*. Asheville, NC: USDA Forest Service, Southern Research Station.

Birdsey, R., Pregitzer, K., & Lucier, A. (2006). Forest carbon management in the United States: 1600–2100. *Journal of Environmental Quality*, *35*, 1461–1469.

Bond-Lamberty, B., Peckham, S. D., Ahl, D. E., & Gower, S. T. (2007). Fire as the dominant driver of central Canadian boreal forest carbon balance. *Nature*, *450*, 88–92.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Brown, S. (1997). Estimating biomass and biomass change of tropical forests. *A Primer*. Rome, Italy: Food and Agriculture Organization of the United Nations (FAO).

Campbell, J., Donato, D., Azuma, D., & Law, B. (2007). Pyrogenic carbon emission from a large wildfire in Oregon, United States. *Journal of Geophysical Research: Biogeosciences*, *112*, G04014.

Cohen, W. B., Yang, Z., & Kennedy, R. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — tools for calibration and validation. *Remote Sensing of Environment*, *114*, 2911–2924.

Coppin, P. R., & Bauer, M. E. (1996). Digital change detection in forest ecosystems with remote sensing imagery. *Remote Sensing Reviews*, *13*, 207–234.

Crookston, N. L., & Dixon, G. E. (2005). The forest vegetation simulator: a review of its structure, content, and applications. *Computers and Electronics in Agriculture*, *49*, 60–80.

Crookston, N. L., & Stage, A.R. (1999). Percent canopy cover and stand structure statistics from the forest vegetation simulator. *General Technical Report RMRS-GTR-24*.

Czembor, C. A., Morris, W. K., Wintle, B.A., & Vesk, P. A. (2011). Quantifying variance components in ecological models based on expert opinion. *Journal of Applied Ecology*, *48*, 736–745.

Deville, J. -C., Särndal, C. -E., & Sautory, O. (1993). Generalized raking procedures in survey sampling. *Journal of the American Statistical Association*, *88*, 1013–1020.

French, N. H. F., Goovaerts, P., & Kasischke, E. S. (2004). Uncertainty in estimating carbon emissions from boreal forest fires. *Journal of Geophysical Research*, *109*, 1–12.

Frescino, T. S., Moisen, G. G., Megown, K. A., Nelson, V. J., Freeman, E. A., Patterson, P. L., et al. (2009). *Nevada photo-based inventory pilot (NPIP) photo sampling procedures*. RMRS-GTR-222.

Gonzalez, P., Asner, G. P., Battles, J. J., Lefsky, M.A., Waring, K. M., & Palace, M. (2010). Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. *Remote Sensing of Environment*, *114*, 1561–1575.

Harris, N. L., Brown, S., Hagen, S.C., Saatchi, S. S., Petrova, S., Salas, W., et al. (2012). Baseline map of carbon emissions from deforestation in tropical regions. *Science*, *336*, 1573–1576.

Healey, S. P., Cohen, W. B., Spies, T. A., Moeur, M., Pflugmacher, D., Whitley, M. G., et al. (2008). The relative impact of harvest and fire upon landscape-level dynamics of older forests: lessons from the northwest forest plan. *Ecosystems*, *11*, 1106–1119.

Healey, S. P., Yang, Z. Q., Cohen, W. B., & Pierce, D. J. (2006). Application of two regression-based methods to estimate the effects of partial harvest on forest structure using Landsat data. *Remote Sensing of Environment*, *101*, 115–126.

Heath, L. S., Smith, J. E., Woodall, C. W., Azuma, D. L., & Waddell, K. L. (2011). Carbon stocks on forestland of the United States, with emphasis on USDA Forest Service ownership. *Ecosphere*, *2*, 1–21.

Henderson, E. B. (2008). Development of state and transition model assumptions used in national forest plan revision. In N. L. Crookston (Ed.), *Third Forest Vegetation Simulator Conference;* 2007 February 13-15; Fort Collins. 89–97. U.S. Department of Agriculture, Forest Service: Rocky Mountain Research Station.

Hoppus, M., & Lister, A. (2005). The status of accurately locating Forest Inventory and Analysis plots using the Global Positioning System. In R. E. McRoberts, & Others (Eds.), *Proceedings of Fourth Annual Forest Inventory and Analysis Symposium (pp. Gen. Tech Rep. NC-252)* (pp. 291–297).

Hudiburg, T., Law, B., Turner, D. P., Campbell, J., Donato, D., & Duane, M. (2009). Carbon dynamics of Oregon and Northern California forests and potential land-based carbon storage. *Ecological Applications*, *19*, 163–180.

Kashian, D.M., Romme, W. H., Tinker, D. B., Turner, M. G., & Ryan, M. G. (2006). Carbon storage on landscapes with stand-replacing fires. *BioScience*, *56*, 598–606.

Kurz, W. A., & Apps, M. J. (1999). A 70-year retrospective analysis of carbon fluxes in the Canadian forest sector. *Ecological Applications*, *9*, 526–547.

Masek, J. G., & Healey, S. P. (2012). Monitoring US forest dynamics with Landsat. In F. Achard, & M. H. Hansen (Eds.), *Global Forest Monitoring*. : CRC Press.

Masek, J. G., Vermote, E. F., Saleous, N. E., Wolfe, R., Hall, F. G., Huemmrich, K. F., et al. (2006). A Landsat surface reflectance dataset for North America, 1990–2000. *Geoscience and Remote Sensing Letters, IEEE*, *3*, 68–72.

Meigs, G., Donato, D., Campbell, J., Martin, J., & Law, B. (2009). Forest fire impacts on carbon uptake, storage, and emission: the role of burn severity in the Eastern Cascades, Oregon. *Ecosystems*, *12*, 1246–1267.

Meigs, G., Turner, D., Ritts, W., Yang, Z., & Law, B. (2011). Landscape-scale simulation of heterogeneous fire effects on pyrogenic carbon emissions, tree mortality, and net ecosystem production. *Ecosystems*, *14*, 758–775.

Moura Costa, P., Stuart, M., Pinard, M., & Phillips, G. (2000). Elements of a certification system for forestry-based carbon offset projects. *Mitigation and Adaptation Strategies for Global Change*, *5*, 39–50.

Nelson, M.D., Healey, S. P., Moser, W. K., & Hansen, M. H. (2009). Combining satellite imagery with forest inventory data to assess damage severity following a major blowdown event in northern Minnesota, USA. *International Journal of Remote Sensing*, *30*, 5089–5108.

Nelson, M.D., Healey, S. P., Moser, W. K., Masek, J. G., & Cohen, W. B. (2011). Consistency of forest presence and biomass predictions modeled across overlapping spatial and temporal extents. *Mathematical and Computational Forestry & Natural Resources Sciences*, *3*, 0–11.

Olofsson, P., Kuemmerle, T., Griffiths, P., Knorn, J., Baccini, A., Gancz, V., et al. (2011). Carbon implications of forest restitution in post-socialist Romania. *Environmental Research Letters*, *6*, 10.

Powell, S. L., Cohen, W. B., Healey, S. P., Kennedy, R. E., Moisen, G. G., Pierce, K. B., et al. (2010). Quantification of live aboveground forest biomass dynamics with Landsat time-series and field inventory data: a comparison of empirical modeling approaches. *Remote Sensing of Environment*, *114*, 1053–1068.

Raymond, C. L., Healey, S. P., Patterson, P. L., & Peduzzi, A. (2014w). Representative regional trajectories of post-disturbance forest carbon: integration of inventory data and a growth and yield model. *Forest Ecology and Management* (In review).

Reams, G. A., Smith, W. D., Hansen, M. H., Bechtold, W. A., Roesch, F. A., & Moisen, G. G. (2005). The forest inventory and analysis sampling frame. *The Enhanced Forest Inventory and Analysis Program — National Sampling Design and Estimation Procedures, General Technical Report SRS-80*.

Reinhardt, E., & Crookston, N. L. (2003). *The Fire and Fuels Extension to the Forest Vegetation Simulator, RMRS-GTR-116, 1-209.*

Ruefenacht, B., Finco, M. V., Nelson, M.D., Czaplewski, R., Helmer, E. H., Blackard, J. A., et al. (2008). Conterminous U.S. and Alaska forest type mapping using Forest Inventory and Analysis data. *Photogrammetric Engineering and Remote Sensing*, *74*, 1379–1388.

Schwind, B., Brewer, K., Quayle, B., & Eidenshink, J. C. (2010). Establishing a nationwide baseline of historical burn-severity data to support monitoring of trends in wildfire effects and national fire policies. In J. M. Pye, H. Rauscher, Y. Sands, D. C. Lee, & J. S. Beatty (Eds.), *Advances in threat assessment and their application to forest and rangeland management* (pp. 381–396). Portland, OR: USDA Forest Service, Pacific Northwest Research Station.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics: John Wiley.*

Smith, W. B., Miles, P. D., Perry, C. H., & Pugh, S. A. (2009). Forest Resources of the United States, 2007. In F.S. Department of Agriculture (Ed.), *Mathermatical and Computational Foresty & Natural Resources Sciences* (pp. 1–336) (Washington, DC).

Spoelma, T. P., Morgan, T. A., Dillon, T., Chase, A. L., Keegan, C. E., & DeBlander, L. T. (2008). *Montana's Forest Products Industry and Timber Harvest, 2004. In (p. 36).* Fort Collins, CO: USDA Forest Service.

Stehman, S. V. (2009). Model-assisted estimation as a unifying framework for estimating the area of land cover and land-cover change from remote sensing. *Remote Sensing of Environment*, *113*, 2455–2462.

Thomas, N. E., Huang, C., Goward, S. N., Powell, S., Rishmawi, K., Schleeweis, K., et al. (2011). Validation of North American forest disturbance dynamics derived from Landsat time series stacks. *Remote Sensing of Environment*, *115*, 19–32.

U.S.F.S (2011). *Navigating the Climate Change Performance Scorecard: A Guide for National Forests and Grasslands (Version 2, August 2011).* In. http://www.fs.fed.us/climatechange/advisor/scorecard/scorecard-guidance-08-2011.pd. United States Department of Agriculture, Forest Service.

Wiedinmyer, C., & Neff, J. C. (2007). Estimates of CO_2 from fires in the United States: implications for carbon management. *Carbon Balance and Management*, *2*, 10.

Williams, C. A., Collatz, G. J., Masek, J., & Goward, S. N. (2012). Carbon consequences of forest disturbance and recovery across the conterminous United States. *Global Biogeochemical Cycles*, *26*, GB1005.

Xu, C., He, H. S., Hu, Y., Chang, Y., Larsen, D. R., Li, X., et al. (2004). Assessing the effect of cell-level uncertainty on a forest landscape model simulation in northeastern China. *Ecological Modelling*, *180*, 57–72.