Soil depth affects simulated carbon and water in the MC2 dynamic global vegetation model

Wendy Peterman *, Dominique Bachelet, Ken Ferschweiler, Timothy Sheehan

Conservation Biology Institute, 136 SW Washington Avenue, Suite 202, Corvallis, OR 97333, USA

Article history:
Received 16 June 2014
Received in revised form 2 September 2014
Accepted 28 September 2014
Available online xxx

Keywords:
Soil
Correlation analysis
Climate change
Vegetation model
Carbon
DGVM

Abstract

Climate change has significant effects on critical ecosystem functions such as carbon and water cycling. Vegetation and especially forest ecosystems play an important role in the carbon and hydrological cycles. Vegetation models that include detailed subground processes require accurate soil data to decrease uncertainty and increase realism in their simulations. The MC2 DGVM uses three modules to simulate biogeography, biogeochemistry and fire effects, all of which use soil data either directly or indirectly. This study includes a correlation analysis of the MC2 model to soil depth by comparing a subset of the model's carbon and hydrological outputs using soil depth data of different scales and qualities. The results show that the model is very sensitive to soil depth in simulations of carbon and hydrological variables, but competing algorithms make the fire module less sensitive to changes in soil depth. Simulated historic evapotranspiration and net primary productivity show the strongest positive correlations (both have correlation coefficients of 0.82). The strongest negative correlation is streamflow (−0.82). Ecosystem carbon, vegetation carbon and forest carbon show the next strongest correlations (0.78, 0.74 and 0.74, respectively). Carbon consumed by forest fires and the part of each grid cell burned show only weak negative correlations (−0.24 and −0.0013 respectively). In the model, when the water demand is met (deep soil with good water availability), production increases and fuels build up as more litter gets generated, thus increasing the overall fire risk during upcoming dry periods. However, when soil moisture is low, fuels dry and fire risk increases. In conclusion, it is clear climate change impact models need accurate soil depth data to simulate the resilience or vulnerability of ecosystems to future conditions.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Climate change is an important driver of forest dieback and species migration with increases in drought, early snow melt, reduced snow depth, pest outbreaks, and fire risk (McKenzie et al., 2004; Mote et al., 2005; van Mantgem et al., 2009; Allen et al., 2010). In the North Pacific landscape of the USA, precipitation as rainfall is projected to increase in winter and spring, and decrease in summer, while temperatures rise from 2 to 5°C by 2080 (Mote and Salathé, 2010). Vegetation models suggest that forest cover may increase at high elevations and latitudes in response to wetter winters, and dramatically decrease at lower elevations and latitudes due to severe competition for water from shrubs and grasses, even without consideration of future water needs from human land use (Climate Impacts Group (CIC), 2011). However, some vegetation models suggest possible vegetation shifts to lower elevations where water might be more readily available as higher elevations become drier (Crimmins et al., 2011).

Climate-related stress can also affect forests indirectly by increasing their vulnerability to pests and pathogens. Littell et al. (2009) projected a reduction of climate suitability for Douglas fir in the Puget Trough as well as increases in wildfires and mountain pine beetle outbreaks, which would affect tree growth and survival in the region. Lodgepole pines in British Columbia, Oregon, Washington and California have also shown increased vulnerability to climate change in recent decades and have been subject to well-documented beetle attacks (e.g., Raffa et al., 2008). Vegetation models indicate that lodgepole may disappear from most of its current range by the end of this century (Coops and Waring, 2011). Further North, Alaskan Yellow Cedar decline in southeast Alaska and portions of British Columbia has also been connected to warming air that melts snow and exposes roots to lethal subfreezing temperatures (D’Amore and Hennon, 2006).

* Corresponding author. Tel.: +1 541 757 0677x117.
E-mail addresses: wendy@consbio.org (W. Peterman), dominique@consbio.org (D. Bachelet), ken@consbio.org (K. Ferschweiler), tim@consbio.org (T. Sheehan).

Abbreviations: ASW, available soil water storage capacity; DGVM, dynamic global vegetation model.

http://dx.doi.org/10.1016/j.ecolmodel.2014.09.025
0304-3800 © 2014 Elsevier B.V. All rights reserved.
Changes in available soil moisture are increasing tree vulnerability across many systems, and available soil water storage capacity (ASW) thresholds have now been documented beyond which forest decline starts to occur during multi-year droughts (Peterman et al., 2013; Mathys et al., 2014). Soil physical characteristics are important for assessing ASW, an essential component of ecosystem functions, including carbon and nutrient cycling, as well as succession through seedling establishment in post-disturbance forests (USDA NRCS Soil Survey Division Staff, 1993; Neilson and Drapek, 1998; Dale et al., 2001; Allen et al., 2010; Puhlick et al., 2012). Simulation results from vegetation models are used in global and regional assessments in an attempt to forecast ecosystem responses to climate change (Cramer et al., 2001; IPCC, 2007; Handler et al., 2013). The complex interactions between plants and pests or pathogens, often constrained by ASW can only be simulated if reliable soil data are available. Soil data have historically been a primary source of uncertainty for modelers who simulate aboveground-processes such as root growth and decomposition as well as hydrological processes (Allen et al., 2010; Coops et al., 2012). In this paper, we report results from a correlation analysis of a dynamic global vegetation model (MC2) that demonstrates the importance of soil inputs in simulations of vegetation dynamics in the 21st century.

1.1. Background and model description

The MC1 dynamic global vegetation model (DGVM) was developed for the vegetation/ecosystem modeling and analysis (VEMAP) project (Bachelet et al., 2001). It consists of three component modules (Fig. 1): (1) a biogeography module derived from the static biogeography model MAPSS (Neilson, 1995), (2) a biogeochemistry module, derived from the CENTURY model (Parton et al., 1987), and (3) a dynamic fire module called MCFire (Lenihan et al., 1998). The MAPSS model is used solely to determine the potential life forms and vegetation types present on the landscape, using a twelve-month long-term average climate to characterize each grid cell during the equilibrium phase of the model (Bachelet et al., 2001). A modified version of the CENTURY model is then called to simulate the carbon and nitrogen pools associated with the potential vegetation types (Bachelet et al., 2001). These initial conditions are used to start the “spinup” phase during which the full DGVM simulates: (1) biogeography, using a set of climate and biomass threshold rules, (2) carbon and nitrogen cycling, using a modified version of CENTURY version 4 and (3) fire occurrence and effects, using the dynamic fire module. The DGVM is run iteratively for 600 years using a de-trended historical monthly climate until net ecosystem productivity nears zero and the fire return interval (FRI) nears historical estimates (Leenhouts, 1998). Once this “spinup” phase is completed, the DGVM is run with historical climate and future climate projections.

In the MC2 DGVM, the hydrology algorithms from CENTURY are used to calculate hydrological flows. The model uses soil depth, texture, rock fragment content and bulk density to estimate monthly available soil moisture. Because the MC2 DGVM uses these soil characteristics to regulate the water fluxes that directly affect plant growth and decomposition, we expect changes in these inputs to result in changes in simulated carbon and hydrology. However, to date, no formal analysis of the relationship between soil characteristics and model simulations has been performed. Conklin (2009) used MC1 to simulate vegetation shifts in Yosemite National Park and observed that the model was overestimating carbon pools and simulating closed-canopy forests at the top of the Sierras. He found that the STASGO-based US soils map that had been used (e.g., Bachelet et al., 2008; Lenihan et al., 1998) included overestimated soil depths, especially at higher elevations (Conklin, 2009). In the original data, he found deep soils at the top of Half Dome, where there should be no soil or vegetation. He used a modified soil dataset based on expert opinion for Yosemite and simulated the more realistic bare rock at the crest of the Sierras.

The NATSOG soil dataset (1:7.5 Million scale), originally used in MC1, was replaced by the STASGO (1:250,000 scale) national soil dataset for the USA (personal communication Kern, 1994). Since then, the finer scale State Soil Geodatabase (SSURGO – average 1:24,000 scale) has been expanded to cover large areas at the state and county level (USDA NRCS, 2014), although the data do not yet provide full coverage of the U.S. For this paper, we conducted our correlation analysis using the MC2 model, the most recent C++ version of MC1, to evaluate whether substituting the soil depth layer from STASGO with SSURGO data, where available, would result in significant changes in carbon cycling and hydrological flows.

Fig. 1. Graphical representation of MC2 DGVM. The biogeography component, uses rules derived from the static biogeography model MAPSS (Neilson, 1995), the biogeochemistry component uses algorithms from a modified version of the biogeochemical model CENTURY (Parton et al. 1987) and the dynamic fire component includes both fire occurrence and effects (Lenihan et al. 1998).
2. Methods

2.1. Study area

Our study area is the portion of the North Pacific Landscape Conservation Cooperative (NPLCC, 2014) within the conterminous United States. This includes the states of Washington and Oregon, west of the Cascade Mountains, and a fraction of northern California, west of the Sierra Nevada mountains (Fig. 2), between the latitudes 39 and 49°. The climate is maritime with moderate winter temperatures and precipitation falling mostly as rain during winter and spring seasons. The region is largely forested with some grasslands and oak savannas in lowland regions, reflecting the “rain-shadow effect” of the Coast Range to the west and the high Cascades to the east. Soils range from multi-colored Spodosols on the coast to ancient Alfisols and Ultisols in the foothills, and fertile Mollisols in the valleys. High, steep mountain slopes may have less developed Inceptisols and Entisols or even bare rock outcrops.

2.2. Soil data preparation

To generate the most reliable soil dataset, we attempted to maximize the use of state soil data. We downloaded SSURGO and STATSGO data from the NRCS Geospatial Data Gateway website (USDA NRCS, 2012) in spatial and tabular forms. First, all of the spatial data were aggregated from the county and national soil datasets into a single soil map covering the entire study area. Then, for every polygon, a Python script was used to extract: (1) attributes directly from the polygon tables of SSURGO and STATSGO databases, (2) attributes for the dominant component within the polygon from the “component” table (the component with the highest representative percent of the polygon), and (3) attributes for the soil layers for all recorded components. For each component, the arithmetic mean of all non-zero values across all soil layers was recorded. Components within each polygon were combined using an area-weighted algorithm based on the normalized percent of each component relative to the others.

The soil datasets include various gaps and errors. While we prefer to use SSURGO soil data because of their finer spatial scale, they often include gaps over large landscapes. Firstly, where SSURGO data were not available, STATSGO data were used to fill these gaps. Secondly, we used the “select by attributes” tool in ArcMap (ESRI, 2011) to locate water bodies, marshes and dams and gave those units a “missing value” of −9999. We also selected rock outcrops, badlands and urban lands, and set the depths to zero. Thirdly, where there were NULL values for minimum depth, we used the polygon name(s) to look up the official soil series descriptions on the NRCS website (USDA NRCS, 2014) and entered the minimum depth recorded for that series. For associations of several soil series within one polygon, we entered the dominant

![Fig. 2. Study area. Map of the study area for the continental portion of the North Pacific landscape conservation cooperative (shaded area).](image-url)
soil series value for minimum depth to bedrock. After processing all polygons this way, there were 8008 total polygons with 4606 null values replaced by soil depth values ranging from 0 to 3860 mm. Where the values could not be filled, the value was set to −9999. The polygons were converted to raster with a 30” grid. The final soil depth grid used in the simulations contained a total of 5676 cells, of which 88.4% held SSURGO data, and 11.6% held STATSGO data.

Differences in soil depth between the Kern dataset and our dataset (Fig. 3) ranged from −1530 to 1700 mm. A comparison between soil depth and elevation for both soil datasets (Fig. 4) illustrates that the coarse resolution STATSGO-based depth dataset (depth Kern) shows less overall variation (standard deviation = 161 mm) than the finer-scale SSURGO–STATSGO hybrid depth dataset (depth Peterman, standard deviation = 247 mm). The ‘depth Kern’ soils are on average 394 mm deeper than the ‘depth Peterman’ soils. Both depth datasets show a general downward trend in depth with increasing elevation, however the slope for depth Kern versus elevation is −0.03 ($R^2 = 0.018$), and the slope for the depth Peterman versus elevation is −0.23 ($R^2 = 0.35$).

2.3. Soil depth in the MC2 model

MC2, the latest C++ version of the MC1 dynamic global vegetation model, uses static soil and monthly climate inputs to simulate plant dynamics, carbon pools and fluxes as well as fire occurrence and effects. All variables in the MC2 model are in a rectangular gridded format with a geographic projection of WGS 1984 and cell size of 30”. Climate variables include precipitation, minimum and maximum temperatures, vapor pressure and wind speed. Other inputs include elevation and eleven soil inputs representing bulk density, depth to bedrock, sand and clay content, as well as rock fragment content. Percent sand, clay and rock fragment data are each provided for surface (0–0.5 m), intermediate (0.5–1.5 m) and deep (>1.5 m) soil layers. MC2 uses the depth layer to determine how many of these soil layers to use for each grid cell. For example, if a cell has a maximum depth of 0.1 m, MC2 would only look at the surface grids for percent sand, clay and rock fragment data. On the other hand, if the maximum depth of a cell is 1.7 m, MC2 would look at each of the three grid layers for percent sand, clay and rock fragment data. For this analysis, we only modified the minimum depth to bedrock variable. Depth is used to determine the total number of soil layers (Fig. 5), including rooted layers, and is therefore critical to simulating available soil moisture content and hydrological fluxes. The top four soil layers are each 0.15 m thick, and the bottom six soil layers are

---

**Fig. 3.** Soil depth maps. Soil depth (left-A) from Jeff Kern derived from STATSGO data (Kern 1994, 1995, 5) and soil depth derived from SSURGO and STATSGO data (right-B) by Wendy Peterman (this paper).

**Fig. 4.** Graph of elevation versus mean soil depth. Comparison between the relationship of soil depth to elevation for both the STATSGO-based soil depth data provided by Kern and the SSURGO-based soil depth data provided by Peterman.
each 0.3 m thick (Bachelet et al., 2001). The rooting depth for grasses is constrained to the top six soil layers, but tree roots may use all ten soil layers.

ASW is calculated using the difference between “field capacity” and “wilting point” (Gupta and Larson, 1979). “Field capacity” is the moisture content of the soil following a precipitation or snowmelt event after the excess water has drained through gravity flow, and “wilting point” is the moisture content of the soil when the tension required for plant roots to extract the water from the soil is too great for their survival. It is important to recognize that the standard wilting point (the soil water content held at 15 bars of tension) and field capacity (the soil water content held at 1/3 bar of tension) are derived for farm crops that have fine, shallow root systems. In reality, plants and trees have a wide range of wilting points that are not individually represented in models of vegetation water balances.

In this research, maximum ASW is used as a limit to the amount of water that can be received, stored and redistributed in a soil profile. MC2 uses separate inputs for mineral soil depth and bulk density to calculate ASW for each cell in a spatial grid (Bachelet et al., 2001). In the model, soil depth does not influence runoff, but it does influence streamflow and baseflow (Bachelet et al., 2001). Using a simple bucket approach, monthly precipitation drains through the soil profile until it reaches bedrock, and then it is lost through belowground flow (Bachelet et al., 2001). There are no algorithms to simulate unsaturated flow in the current version of the model.

2.4. Model correlation analysis

For the LandCarbon Project (USGS, 2011), MC2 was run for the conterminous USA under nine climate and emissions scenarios at

<table>
<thead>
<tr>
<th>Variable</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecosystem carbon</td>
<td>0.7835</td>
</tr>
<tr>
<td>NPP</td>
<td>0.8170</td>
</tr>
<tr>
<td>Stream flow</td>
<td>-0.8166</td>
</tr>
<tr>
<td>Forest carbon</td>
<td>0.7442</td>
</tr>
<tr>
<td>Carbon consumed by fire</td>
<td>0.2389</td>
</tr>
<tr>
<td>Part of cell burned by fire</td>
<td>-0.0013</td>
</tr>
<tr>
<td>Actual evapotranspiration</td>
<td>0.8161</td>
</tr>
</tbody>
</table>
30°. Since large runs of MC2 are processor-intensive and time-consuming, “strided” runs using subsamples of the input datasets were first used to calibrate the model for the historical time period (1895–2005) (Bachelet et al., 2013). To create the “strided” data, we resampled the data for every tenth row and column of the gridded datasets. To quickly perform our correlation analysis, we decided to use the same method, so we generated a new strided soil data set and ran the model with strided 30° climate data for our study area. All model runs were configured with fire suppression turned on. Historical vegetation simulations for the LandCarbon project were compared to Kuchler’s map of potential natural vegetation for the United States (Bachelet et al., 2013). Initial comparisons

Fig. 6. Correlation maps for carbon and hydrology variables. Percent change in 30 year means of a subset of MC2 simulated ecosystem processes with strong positive correlations (0.7–1.0) with the difference in soil depth: (A) total ecosystem carbon, (B) forest carbon, (C) vegetation carbon, (D) net primary productivity and (E) actual evapotranspiration.
showed an over-estimation of grasslands in the northern Great Plains of the US, so the model was calibrated to adjust the biogeography thresholds and create better agreement between modeled and mapped vegetation types (Bachelet et al., 2013).

Previous MC1 and MC2 simulations used soil data for the conterminous USA originally provided by Jeff Kern for the VEMAP project (Kern 1994, 1995, 5). Melillo et al. (1995) projected and resampled the soils data to a finer scale at 10 km, based on a gridded soil conservation service national level (NATSGO) database. Cluster analysis was used to group the 10 km subgrid elements into 1–4 dominant ("modal") soil types for each 0.5° cell. Using this approach, cell soil properties were represented by one or more dominant soil profile that might not correspond to realistic soil characteristics in the region.

For our correlation analysis, we substituted the standard soil depth that had also been used for the LandCarbon runs (depth Kern) with a new depth layer developed specifically for this study (depth Peterman) and left the other 10 soil variables (representing bulk density and soil texture) unchanged. We ran MC2 with strided input data for our study area for the historical period using the same configuration settings used in the LandCarbon project. Once the runs were completed, we correlated the differences in depths between the two soil input datasets (depth Kern–depth Peterman) with differences in the 30-year means of carbon, fire and hydrology model results (LandCarbon 30-year mean – this study's 30-year mean) using Pearson's correlation coefficients. The model was run for 111 years (1895–2005), but we used the average of the most recent 30-years in our comparisons. Differences between any two variables were calculated using the "ncdiff" command in the NCO toolbox (Zender, 2014). Correlation coefficients were calculated using a Python script that eliminated outliers greater than three standard deviations from the mean for each variable's 30-year average differences. Vegetation cover is a discrete variable, so we used the modal value of the last 30-years of the historical period to compare differences between the simulations. In ArcMap we calculated the difference in area simulated for each vegetation type and compared that with the total area simulated.

3. Results

All but three of the simulated variables chosen for this study showed strong correlations to differences in mineral soil depth (Table 1). Ecosystem carbon, net primary production, forest carbon, evapotranspiration and vegetation carbon (Fig. 6) all showed strong positive correlations (0.7–1.0) with changes in depth, while carbon consumed by wildfire showed a weak positive correlation (0–0.3). Correlations were most evident in northern California (Fig. 7). Stream flow values showed a strong negative correlation (–1.0 to –0.7) with differences in mineral depth, and the difference in the fraction of the grid cell burned by wildfire showed a weak negative correlation (–0.3 to 0) (Table 1). We found some small differences between the two vegetation maps simulated with the 2 soil datasets (Fig. 8) at this scale. For approximately 92% of the study area, there was no difference in simulated vegetation. Approximately 3% of the area that was simulated as temperate needle leaf forest with the Kern soil data was simulated as temperate needle leaf woodland with the new soil dataset. This means that the original runs predicted the cells as having a biomass greater than 1150 gm⁻², while this study's data predicted those cells as having less than 1150 gm⁻² of biomass. Approximately 4% of the area went from being predicted with the original data to a null value with the new soil data, probably due to zero values in the new soil depth dataset. The Kern soil dataset had no zero values for soil depth.

4. Discussion

We expected that soil depth would affect MC2 carbon simulations, since its three modules each use soil depth either directly or indirectly. Biogeography rules use biomass thresholds that are indirectly linked to soil depth through soil water availability for plant growth. The biogeochemistry algorithms from CENTURY use soil depth to calculate water availability and nutrient cycling. The fire module of MC2 uses soil moisture as a proxy for fuel moisture, which is used as one of the thresholds for fire risk. The strong correlations between the change in soil depth

Fig. 7. Carbon consumed by wildfire map. Difference (g C m⁻²) in 30-year mean carbon consumed by wildfire between the two MC2 simulations, one using Kern (1994, 1995,) and the other using Peterman (this paper) soil depth.
and resulting differences in carbon pools and streamflow show that the biogeochemistry module is sensitive to soil depth. Decreases in soil depth lead to decreases in net primary productivity and ecosystem, forest and vegetation carbon as expected, since it reduces overall soil water availability. The strong positive correlation to simulated evapotranspiration shows that vegetation in the model is able to respond to the greater or lesser amount of available soil water provided by the change in soil depth to meet the evaporative demand. The strong negative correlation to streamflow indicates that the shallower soils were less rooted than deeper soils. The related transpiration fluxes, calculated for each soil layer, were better able to use the available soil moisture and thus reduce the amount of water draining through the profile.

The MC2 fire module uses soil moisture as a proxy to determine coarse fuel moisture. Below a set fuel moisture threshold, conditions are met to start a fire in a given cell as long as other conditions (fuel load and climate) allow it. A shallower soil depth reduces soil moisture available to plant growth thereby limiting fuel production and thus limiting fire risk. Such a combination of competing thresholds (shallow soil depth means low fuel moisture but also low fuel load) may explain the low correlation between soil depth and the amount of biomass consumed by wildfire, and an even lower correlation for area burned.

Most of the large differences in simulated carbon, hydrology and fire effects are seen in northern California (see Figs. 7 and 8). This suggests the dominance of climate over soils as the determining factor in productivity and nutrient and water cycling in the Pacific Northwest. In western Oregon and Washington, where there is high annual rainfall, available soil water storage capacity is not as limiting to plant growth. In California, however, the drier climate makes the influence of soil characteristics more important. Not only does soil depth limit the growth of root systems in the model, the higher evaporative demand of the more arid environment combined with shallower soils places greater constraints on plant productivity in the region.

5. Conclusion

Concerns that MC2 was simulating closed-canopy forests on the peaks of Yosemite National Park raised questions about the
model's representations of soil characteristics. Conklin (2009) found that the problem was attributed to inaccuracies in the soil input data. In that case, by revising the soil dataset with local knowledge, the problem was resolved. To better evaluate the model's sensitivity to soil depth, we performed a correlation analysis over the North Pacific LCC domain evaluating the relationship of a subset of model results to two soil datasets, one used previously and originally created for the VEMAP project at coarse resolution and a new dataset created with finer scale soil information more relevant for regional analysis.

Correlations between changes in soil depth and resulting differences in simulated carbon and water fluxes were strong and demonstrated the expected responsiveness of the model to soil depth. Weak correlations between changes in soil depth and changes in fire effects reflect competing drivers of fire occurrence and effects. Fire occurrence and effects depend on high fuel loads driven by high plant productivity and high soil water availability, as well as by dryness of fuels driven by low soil water associated with shallow depth. The small change in simulated vegetation cover reflects the climatic variability of the study area. Only in northern California, where winter precipitation is less abundant than in western Oregon and Washington, does soil depth affect the potential vegetation type simulated by the model.

This modeling exercise demonstrates the need for detailed and accurate soils data for more realistic simulations of ecosystem processes. Carbon sequestration potential and the amount and timing of water provision are highly valued ecosystem functions that can only be projected with the best soil data available. Therefore, efforts to map soils with better accuracy and precision are essential to decreasing uncertainty and increasing the "realism" of climate change impact models.

Some interesting projects that could develop from this research are: (1) correlation analysis of how MC2 simulations are affected by finer scale texture input data, (2) greater development of the role of soils in the MC2 hydrology algorithms, especially the calculation of runoff and unsaturated flow, and (3) more exploration into the details of how soils are used in the MC2 fire module and how this can be improved.

Acknowledgments

We would like to acknowledge the North Pacific Landscape Conservation Cooperative and the US Department of the Interior, for providing funding for this research. In addition, we would like to thank individuals who provided technical support and expertise. Brendan Ward (Conservation Biology Institute) created the Python script for extracting and aggregating soils data from the NRCS soils database. David Conklin (Common Futures), Myrica McCune (Institute for Natural Resources), Jeff Kern (Oregon Department of Fish and Wildlife) and Raymond Drapek (USDA Forest Service) provided background on previous methods and uses of soil data and MC1/2. David King provided insight about the relationships between soil depth data and the MC2 fire module. Story Edison (Oregon State University) and Gwynne Corrigan (Conservation Biology Institute) provided editorial advice.

References


