Near-future forest vulnerability to drought and fire varies across the western US

Polly C. Buotte[1*], Samuel Levis[2], Beverly E. Law[1], Tara W. Hudiburg[3], David E. Rupp[4], Jeffery J. Kent[3]

[1]Department of Forest Ecosystems, and Society, Oregon State University, Corvallis, OR, USA; [2]SLevis Consulting, LLC, Oceanside, CA, USA; [3]Department of Forest, Rangeland, and Fire Sciences, University of Idaho, Moscow, ID, USA; [4]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR, USA
[*]Corresponding author, pcbuotte@gmail.com, 406.589.6863

Introduction

The Community Land Model (CLM) requires time series of several meteorological variables as input. These variables include temperature, precipitation, vapor pressure deficit, incoming short-wave radiation, and wind speed. The desired temporal resolution of these variables is 3-hourly.

The CLM was driven with observation-based and simulated gridded meteorological data at 1/24 degree. The source for observational data was METDATA and the source for the simulated data was MACAv2-METDATA, or "MACA" for short hereon. Both METDATA and MACA are at a daily resolution, so the daily data were disaggregated to a 3-hourly resolution using the methodology described in this companion file (taken from the SI Appendix to the publication Buotte et al,).

Methodology for downscaling and bias correcting CRUNCEP climate data used to drive the Community Land Model

Data sources:
1. CRUNCEP ½ degree climate data (Mitchell & Jones, 2005)
2. Observation-based 4km climate data (Abatzoglou, 2013), referred to as METDATA

Methods:
1. Downscale the ½-degree CRUNCEP data (1901-2010) to our 1/24 degree grid using nearest neighbor
2. Temporally and spatially bias-correct the CRUNCEP to the Abatzoglou METDATA
3. Define the base climatological period (BCP; period of overlap between CRUNCEP and METDATA), 1979-2010.
4. Temperature

4.1 Calculate mean monthly values during climatological base period.

$T_G(month,BCP)$ = mean CRUNCEP temperature in a given month (January, February, …, December) over all years in base climatological period in a given pixel (12 values per pixel); averages are computed by summing all 6-hourly values within that month in those years, and dividing by N = 4 time steps per day * number of days in the month * 32 years.

$T_M(month,BCP)$ = mean METDATA temperature in a given month during the base climatology period for a given pixel (12 values per pixel); averages are computed summing all 3-hourly values within that month in those years, and dividing by N = 8 time steps per day * number of days in the month * 32 years.

4.2 Calculate the difference to be applied to CRUNCEP. This results in a mean difference between the two climatologies for each month (12 values per pixel).

$$\Delta T(month) = T_M(month,BCP) - T_G(month,BCP)$$

4.3 Bias correct the GSWP during 1901-1978 ($T_G(6\text{-}hour,month,year)$), where *6-hour* is the 6-hourly timestep within the *month*). $T_{Gbc}(6\text{-}hour,month,year)$ is the bias-corrected temperature.

$$T_{Gbc}(6\text{-}hour,month,year) = T_G(6\text{-}hour,month,year) + \Delta T(month)$$

Five Variables with lower bound at 0 (precipitation, radiation, humidity, wind speed, pressure)

5.1 Calculate average monthly values during climatological base period, precipitation rate used as the example

$P_G(month,BCP)$ = average CRUNCEP precipitation rate (mm/s) in a given month (January, February, …, December) over all years in base climatological period in a given pixel (12 values per pixel); sums are computed using all 6-hourly values within that month in those years, divided by N = 4 timesteps perday * number of days in month * 32 years

$P_M(month,BCP)$ = data from David = average METDATA precipitation rate (mm/s) in a given month during the base climatology period for a given pixel (12 values per pixel); sums are computed using all daily values within that month in those years (**ensure METDATA units are same as CRUNCEP for all other variables**), divided by N = 8 time steps per day * number of days in the month * 32 years

5.2 Calculate a ratio to be applied to CRUNCEP. This results in a scale factor between the two climatologies for each month (12 values per pixel).

$$P_{scale\_factor}(month) = P_G(month,BCP) / P_M(month,BCP)$$

5.3 Ensure there are no extreme ratios

Inspect the distribution of pixel values for each monthly ratio, further inspect any outliers.

5.4 Bias correct the GSWP during 1901-1978 ($P_G(6\text{-}hour,month,year)$), where *6-hour* is the 6-hourly timestep within the *month*). $P_{Gbc}(6\text{-}hour,month,year)$ is the bias-corrected precipitation.

$$P_{Gbc}(6\text{-}hour,month,year) = P_G(6\text{-}hour,month,year) * P_{scale\_factor}(month)$$

5.5 Bias correct the other 0-bounded variables analogously.

Data sources

i. Observation-based gridded meteorological data (previously called METDATA, now called GRIDMET), daily, 1/24-degree resolution ((Abatzoglou, 2013). Last accessed 2016-01-29. http://www.climatologylab.org/gridmet.html.

ii. North America Regional Reanalysis (NARR), three-hourly, 0.3-degree resolution (Mesinger *et al.*, 2006). Accessed 2013-09-18 (data through 2012) and 2016-03-21 (data from 2013 through 2015). https://www.esrl.noaa.gov/psd/data/gridded/data.narr.html.

iii. Statistically downscaled global climate model simulations using the method of Multivariate Adaptive Constructed Analogs v.2 with METDATA as the training data (MACAv2-METDATA), daily, 1/24-degree resolution. Last accessed 2014-11-09. https://climate.northwestknowledge.net/MACA/.

iv. Global climate model (GCM) output from the Coupled Model Intercomparison Project Phase 5 (CMIP5) archive, three-hourly, various spatial resolutions (Taylor *et al.*, 2012). Our study uses output from IPSL-CM5A-MR *r1i1p1* and MIROC5 *r1i1p1*, *historical* and *rcp85* experiments. Accessed 2011-11-05. https://cmip.llnl.gov/cmip5/data_portal.html.

Disaggregation of METDATA

To disaggregate the daily METDATA to a 3-hourly resolution, we made use of the 3-hourly NARR data. To disaggregate the downscaled daily MACA to a 3-hourly resolution, we made use of the 3-hourly data from the "raw" (i.e. not downscaled) CMIP5 GCM simulations.

Briefly, the method consists of 'rescaling' the 3-hourly GCM (or NARR) time series to be consistent with aggregate daily values, or maximum and minimum daily values, from MACA (or METDATA). Note from hereon, we use the example of MACA 3-hour disaggregation, though the METDATA 3-hourly disaggregation follows the identical method, other than the GCM data are used with MACA whereas NARR is used with METDATA.

This rescaling entails first converting the variables in the 3-hourly GCM datasets to dimensionless quantities by a standardization method (see 2.1 below-the particular method depends on the variable). Second, the downscaled daily data is disaggregated to 3-hourly by multiplying the daily data by the standardized quantities, in a sense 'de-standardizing' the data.

To apply the second step, each fine-grained cell in the downscaled dataset set must be mapped to its associated coarse-grained cell in the GCM dataset.

The strength of this method is that it maintains the covariance structure of, and therefore the physical consistency between, all the variables. It also maintains consistency with the source GCM itself. This latter feature may also be a drawback if the GCM poorly represents variability of some variable, precipitation for example, at the 3-hourly timescale. A potential weakness of the method is that the standardized time series for a given day is identical across the GCM cell which are relatively large at ~1 to 2.5-degrees. This means that a storm, for example, would peak at exactly the same time everywhere within the GCM cell, though magnitude would vary across the cell. This temporal uniformity at the GCM scale might be important if lateral fluxes between the CLM cells were an issue (for example, if we were simulating flood discharge in a river). However, in CLM the fluxes are only vertical, so for out purposes this drawback of the method is unimportant. (Note: this is less of an issue with the higher resolution NARR data.)

2.1. Creating a standardized 3-hourly gridded dataset

We first transform each variable in the 3-hourly GCM dataset so it becomes "standardized". To transform temperature *T*, we use the following standardization:

$$T_i^* = \frac{T_i - (T_{\max} + T_{\min})/2}{T_{\max} - T_{\min}} \qquad \text{for} \qquad T_{\max} \neq T_{\min} \tag{1a}$$

where the superscript '*' indicates the standardized quantity, *i* indexes the 3-hour time step within a 24 hour day (1, 2, …, 8), and $T_{\max}$ and $T_{\min}$ are the maximum and minimum temperatures, respectively, on that day. If $T_{\max} = T_{\min}$ then we apply a simple sine function:

$$T_i^* = \frac{1}{2}\sin\left[\frac{\pi}{12}\left(t_{local} - 9\right)\right] \tag{1b}$$

where t_{local} is local time (0 to 24 hours). Eq. (1b) sets the minimum and maximum temperatures to fall at 3:00 and 15:00 hours, local time, respectively.

In the case of precipitation *P*, the precipitation total within a 3-hourly period is divided by the total precipitation on that day:

$$P_i^* = \frac{P_i}{\sum_{i=1}^{8} P_i} \qquad \text{for} \qquad \sum_{i=1}^{8} P_i > 0 \tag{2a}$$

else

$$P_i^* = 1/8 \tag{2b}$$

Specific humidity *SH* is standardized by the mean of *SH* over the day:

$$SH_i^* = \frac{SH_i}{\frac{1}{8}\sum_{i=1}^{8} SH_i} \qquad \text{for} \qquad \sum_{i=1}^{8} SH > 0 \tag{3a}$$

else

$$SH_i^* = 1 \tag{3b}$$

Downward short-wave radiation *SW* is standardized similarly:

$$SW_i^* = \frac{SW_i}{\frac{1}{8}\sum_{i=1}^{8} SW_i} \qquad \text{for} \qquad \sum_{i=1}^{8} SW > 0 \tag{4a}$$

else

$$SW_i^* = 1 \tag{4b}$$

In the case of wind, anomalies of the zonal and meridional velocity components, *U* and *V*, respectively are generated as differences from the daily mean:

$$\tilde{U}_i = U_i - \overline{U} \tag{5a}$$

$$\tilde{V}_i = V_i - \overline{V} \tag{5b}$$

where

$$\overline{U} = \frac{1}{8}\sum_{i=1}^{8} U_i \, , \tag{6a}$$

$$\overline{V} = \frac{1}{8}\sum_{i=1}^{8} V_i \tag{6b}$$

2.2. Disaggregating the downscaled daily dataset

To disaggregate the daily data to 3-hourly, each fine-grained cell in the downscaled MACA dataset must be first be mapped to its 'parent' coarse-grained cell in the GCM. This mapping simply consists of assigning each cell in the downscaled grid the 'address' of its parent cell as a set of (j, k) indices.

Once the address is known, the disaggregation consists of inverting the above standardization procedure, with the difference that the standardization parameters are now taken from the daily

MACA data instead of the GCM data. In the case of temperature, the standardization parameters are t_{max} and t_{min} from MACA data and the inversion is as follows:

$$t_i = t_{max}\left(T_i^* + \frac{1}{2}\right) - t_{min}\left(T_i^* - \frac{1}{2}\right) \qquad (7)$$

where t_i denote the fine-grained (downscaled) temperature at time step indexed by i as before.

For precipitation, the parameter is the total precipitation p_{tot} on the day in question:

$$p_i = P_i^* p_{tot} \qquad (8)$$

For specific humidty, the standardized time series is multiplied by the mean specific humidity $\overline{sh}$ on a given day:

$$sh_i = SH_i^* \overline{sh} \qquad (9)$$

Short-wave radiation is calculated likewise from the mean short-wave radiation $\overline{sw}$:

$$sw_i = SW_i^* \overline{sw} \qquad (10)$$

For the wind vectors, we simply add the 3-hourly wind vector anomalies from the GCM to the daily downscaled mean wind components $\overline{u}$ and $\overline{v}$:

$$u_i = \overline{u} + \widetilde{U}_i \qquad (11a)$$

$$v_i = \overline{v} + \widetilde{V}_i \qquad (11b)$$

This preserves the downscaled daily wind speed calculated using the mean wind components $\overline{u}$ and $\overline{v}$.

Methodology for generating harvest scenarios used in the Community Land Model to mimic historical, "business-as-usual" harvest rates

Timber harvest was calibrated at the state level to reproduce historic harvest totals in each state (Berner *et al.*, 2017), with grid cell selection respecting a 60-year rotation length. Grid cells were considered for harvest only if they were not protected public lands and had soils with at least 120 mm available water holding capacity (Peterman *et al.*, 2013), at least 300 trees per hectare and at least a 40-year fire return interval (Hudiburg *et al.*, 2013).

References

Abatzoglou JT (2013) Development of gridded surface meteorological data for ecological applications and modelling. International Journal of Climatology, **33**, 121-131.
Berner LT, Law BE, Meddens AJH, Hicke JA (2017) Tree mortality from fires, bark beetles, and timber harvest during a hot and dry decade in the western United States (2003-2012). Environmental Research Letters, **12**.
Hudiburg TW, Luyssaert S, Thornton PE, Law BE (2013) Interactive Effects of Environmental Change and Management Strategies on Regional Forest Carbon Emissions. Environmental Science & Technology, **47**, 13132-13140.
Mesinger F, Dimego G, Kalnay E *et al.* (2006) North American regional reanalysis. Bulletin of the American Meteorological Society, **87**, 343-+.
Mitchell TD, Jones PD (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International Journal of Climatology, **25**, 693-712.
Peterman W, Waring RH, Seager T, Pollock WL (2013) Soil properties affect pinyon pine - juniper response to drought. Ecohydrology, **6**, 455-463.
Taylor KE, Stouffer RJ, Meehl GA (2012) An Overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society, **93**, 485-498.