Technical note: 3-hourly temporal downscaling of monthly global terrestrial biosphere model net ecosystem exchange

Joshua B. Fisher¹,* , Munish Sikka¹, Deborah N. Huntzinger², Christopher Schwalm³, Junjie Liu¹

¹ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA, 91109, USA
² School of Earth Sciences and Environmental Sustainability, Northern Arizona University, 527 S. Beaver St., Flagstaff, AZ, 86011-5694, USA
³ Woods Hole Research Center, Falmouth, MA, 02540, USA

* Corresponding author. E-mail: jbfisher@jpl.nasa.gov

Author contributions: JBF, DNH, and CS formulated idea; JBF and MS designed research; MS performed research; DNH and CS provided data; all authors contributed to the writing of the paper.

The authors declare no conflict of interest.

Running title: 3-hourly temporal downscaling of monthly NEE
Abstract

The land surface provides a boundary condition to atmospheric forward and flux inversion models. These models require prior estimates of CO$_2$ fluxes at relatively high temporal resolutions (e.g., 3-hourly) because of the high frequency of atmospheric mixing and wind heterogeneity. However, land surface model CO$_2$ fluxes are often provided at monthly time steps, typically because the land surface modeling community focuses more on time steps associated with plant phenology (e.g., seasonal) than on sub-daily phenomena. Here, we describe a new dataset created from 15 global land surface models and 4 ensemble products in the Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP), temporally downscaled from monthly to 3-hourly output. We provide 3-hourly output for each individual model over 7 years (2004-2010), as well as an ensemble mean, a weighted ensemble mean, and the multi-model standard deviation. Output is provided in three different spatial resolutions for user preferences: 0.5° x 0.5°, 2.0° x 2.5°, and 4.0° x 5.0° (latitude/longitude). These data are publicly available from: ftp://daac.ornl.gov/data/cms/CMS_NEE_CO2_Fluxes_TBMO/data.

Keywords: CO$_2$ flux; downscale; land surface; NEE; sub-daily; hourly
This technical note describes the methodological approach employed with temporally downscaling monthly terrestrial biosphere model (TBM) net ecosystem exchange (NEE) (i.e., net CO₂ flux between the land and atmosphere) output to 3-hourly time steps (Fisher et al., 2014). These data were created initially for NASA’s Carbon Monitoring System (CMS), and are useful to the broader scientific community (Fisher et al., 2011; Fisher et al., 2012). The general downscaling approach follows Olsen and Randerson (2004) with modifications. The logic takes the components of NEE, i.e., gross primary production (GPP) and ecosystem respiration (Re), and links them with incident shortwave solar radiation (I) and surface air temperature (T_a), respectively. I and T_a are provided at 6-hourly time steps from CRU-NCEP (Wei et al., 2014a; Wei et al., 2014b), which we interpolated to 3-hourly time steps following cosines of solar zenith angle for I and linear interpolation for T_a. Hence, GPP and Re are temporally downscaled to 3-hourly, and re-combined to form NEE at 3-hourly time steps.

The 6-hourly to two 3-hourly time steps from the solar zenith angle cosine interpolation follows this equation:

\[ I_{t1} = \frac{I_t \times \cos(z_{t1})}{\cos(z_{t1})} \]
\[ I_{t-1} = \frac{I_{t-1} \times \cos(z_{t-1})}{\cos(z_{t-1})} \]

where \( z \) is solar zenith angle and \( I \) is in units of W m\(^{-2}\). As an example, if the 0-6 hour \( I_t \) was 100 W m\(^{-2}\), and the 0-3 hour \( z_{t1} \) was 0 (i.e., \( \cos(z_{t1}) = 1 \)) and the 4-6 hour \( z_{t-1} \) was 60 (i.e., \( \cos(z_{t-1}) = 0.5 \)), then the 0-3 hour \( I_{t1} \) would be 133.3 W m\(^{-2}\), and the 4-6 hour \( I_{t-1} \) would be 66.7 W m\(^{-2}\).

To scale GPP and Re to 3-hourly time steps, we followed Olsen and Randerson (2004) with modifications starting first with the calculation of scale factors based on I and T_a:

\[ Q10_{3hr} = 1.5^\left(\frac{T_a_{3hr}-30}{10}\right) \]
\[ T_{scale} = Q10_{3hr} \sum_{3day} Q10_{3hr} \]
\[ I_{scale} = I_{3hr} \sum_{3day} I_{3hr} \]

where \( Q10 \) is the temperature dependency of Re, and T_a is in degrees Celsius (converted from Kelvin, as provided by CRU-NCEP). Note that Olsen and Randerson (2004) originally used time integral periods of calendar months, but we observed that this caused unrealistic distinct shifts between months. Instead, we modified the integral period to a 30-day moving window (Figure 1).

For the first 15 days of January of the record and the last 15 days of December of the record, we used the last 15 days of December and the first 15 days of January, respectively, within the first (2004) and last (2010) years to complete the 30-day window.

The 3-hourly resolution scale factors are then multiplied by GPP and Re, respectively, for each 3-hourly time step each month:

\[ Re_{3hr} = T_{scale} \times Re_{month} \]
\[ GPP_{3hr} = I_{scale} \times GPP_{month} \]

We modified \( Re_{month} \) and \( GPP_{month} \) from Olsen and Randerson (2004) to be given at a 3-hourly time step, linearly interpolated to 3-hourly time steps based on the present, previous, and subsequent month, maintaining the original units (g C m\(^{-2}\) mo\(^{-1}\)). \( Re_{3hr} \) and \( GPP_{3hr} \) are in units of g C m\(^{-2}\) 3hr\(^{-1}\). This modification avoided using the same monthly value for the multiplier for all 3-hourly time steps per month as per Olsen and Randerson (2004), and instead provided a
smooth transition from one month to the next. The result of this modification was to eliminate a
“ramping” effect whereby values would, for example, increase steadily within a month, then
suddenly shift to a new starting point at the beginning of the next month (Figure 1). Note that the
original nomenclature of Olsen and Randerson (2004) used \((2 \times NPP_{\text{month}}) - N\) in
place of \(R_{\text{month}}\), and \((2 \times NPP_{\text{month}})\) in place of \(GPP_{\text{month}}\), where \(NPP\) is net primary production
\((GPP\) minus autotrophic respiration) and \(NEP\) is net ecosystem production (approximately
equivalent to the inverse sign of \(NEE\), with caveats (Hayes and Turner, 2012)). The assumption
here, therefore, is that \(GPP = 2 \times NPP\) and \(R = (2 \times NPP) - NEP\). The \(Re\) assumption misses
\(CO_2\) emissions other than respiration, e.g., fire, which we correct for at a later step.

The initial \(NEE\) calculation simply subtracts \(GPP\) from \(Re\):
\[
NEE_{3\text{hr}} = R_{3\text{hr}} - GPP_{3\text{hr}}
\]
where \(NEE_{3\text{hr}}\) is calculated in units of g C m\(^{-2}\) 3hr\(^{-1}\). However, we applied an additional units
conversion for the publicly available data to kg C km\(^{-2}\) s\(^{-1}\), as these units are more readily
ingestible by atmospheric inversion models (Deng et al., 2014).

Because the downscaling approach uses \(Re\) as the primary \(CO_2\) efflux term, other ecosystem \(CO_2\)
loss components, such as fire and other disturbances (Hayes and Turner, 2012), are excluded in
the downscale. Hence, the sum of the downscaled 3-hourly \(NEE\) fluxes in a given month did not
necessarily equal the original monthly \(NEE\) flux. So, we included a per-pixel correction whereby
we: I) calculated the difference between the sum of the downscaled 3-hourly \(NEE\) in a given
month and the original monthly \(NEE\); II) divided that difference by the total 3-hourly time steps
in the month, and III) added that difference to each 3-hourly \(NEE\) flux. In so doing, the sum of
the downscaled 3-hourly \(NEE\) fluxes subsequently summed exactly to the original monthly \(NEE\).

All input data were given in a spatial resolution of 0.5° x 0.5° (latitude/longitude); hence, we
provide the 3-hourly \(NEE\) output in 0.5° x 0.5° (Figure 2). We also provide two additional sets of
spatially upscaled \(NEE\) output in 2.0° x 2.5° and 4.0° x 5.0°. These resolutions are used by the
atmospheric modeling community, i.e., the GEOS-Chem atmospheric \(CO_2\) transport model in the
NASA CMS (Liu et al., 2014). To generate the coarser resolution data we: I) multiplied each
pixel value by the land area of that pixel; II) summed the flux from all pixels that represent one
defined in coarser resolution (e.g., 8 x 10 pixels from 0.5° x 0.5° comprise 1 pixel in 4.0° x 5.0°)
III) calculated the total area covered by the pixels summed in step II; and, IV) divided the value
in step II by the value in step III. The regridding preserved the total sum flux of the finer grid
cells as well as the total global flux. We provide a file containing the land area contained in each
latitude band for each of the 3 resolutions (folder name: ‘latitude_area’). We provide two
versions of the 2.0° x 2.5° and 4.0° x 5.0° resolution products—one version with consistent
global resolution, and another that conforms to the GEOS-Chem setup whereby the northern and
southern most latitudinal bands for the 2.0° x 2.5° resolution are 1.0° x 2.5°, and for the 4.0° x
5.0° they are 2.0° x 5.0°. The orientation of the global grid in the NetCDF files is transposed (i.e.,
90°S x 180°W at top-left). The time vector represents the mid-point of each 3-hourly period.

Processing time in R, un-parallelized, on a standard PC for a single year for the forcing data was
as follows:
- Interpolation of 6-hourly \(I\) and \(T_a\) to 3-hourly time step: 1 hr per variable
- 30-day moving window for \(I\): 48 hr
30-day moving window for \( T_a \): 68 hr

Total time to process forcing data for 7 years: \( 7 \times (1 \times 2 + 48 + 68) = 826 \) hr

Processing time for the application of the modified Olsen and Randerson (2004) downscaling approach for a single model for a single year was:

- Monthly interpolation to 3-hourly time steps for GPP: 1 hr
- Monthly interpolation to 3-hourly time steps for Re: 1 hr
- GPP and Re downscaling: 2 hr
- Monthly NEE closure correction: 1 hr
- NetCDF generation with additional spatial resolutions: 2 hr

Total time to process all 19 products for 7 years: \( 7 \times 19 \times (1 + 1 + 2 + 1 + 2) = 931 \) hr

The total storage size of the final NetCDF data products for all 19 products (15 models + 4 ensemble products) for all 7 years is: 374 GB at 0.5° x 0.5°, 38 GB at 2.0° x 2.5°, and 10 GB at 4.0° x 5.0°.

We provide the data in NetCDF with a separate file for each day per product at ftp://daac.ornl.gov/data/cms/CMS_NEE_CO2_Fluxes_TBMO/data. Each file contains the global gridded data with the eight 3-hourly intervals in the day. Open water pixels are set to 0, as this was desired by the atmospheric modeling community. However, we realize that NEE values can conceivably be 0 (though unlikely as our precision is to 16 decimal places); nonetheless, there are some pixels over land that are calculated as 0, but this is due to missing forcing data (e.g., \( I \) in the high latitudes during winter). Our code is set up that we can easily provide a different file output structure and missing value mask by request (contact the corresponding author: jbfisher@jpl.nasa.gov).

Model output (GPP, Re, and NEE) was from the Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP) (Huntzinger et al., 2013; Huntzinger et al., 2016), version 1. 15 models were included: 1) BIOME_BGC, 2) CLM, 3) CLM4VIC, 4) CLASS_CTEM, 5) DLEM, 6) GTEC, 7) ISAM, 8) LPJ-wsl, 9) ORCHIDEE, 10) SIB3, 11) SIBCASA, 12) TEM6, 13) TRIPLEX-GHG, 14) VEGAS2.1, and 15) VISIT (Table 1). Additionally, 4 ensemble products were included: 1) un-weighted (naïve) ensemble mean, 2) un-weighted (naïve) ensemble standard deviation, 3) weighted (optimal) ensemble mean, and 4) weighted (optimal) ensemble standard deviation. Weights for model ensemble integration were derived based on model skill in reproducing GPP and biomass (Schwalm et al., 2015). Model output was obtained from: ftp://nacp.ornl.gov/synthesis/2009/reutlingen/CMS/20141006/.

To test and confirm the accuracy of our downscaling approach, we applied our method on a set of ground-truth data of measured NEE (and forcing variables) from the FLUXNET database (Baldocchi et al., 2001). We show, for example, a single year for a single site (3-hourly in background with daily-moving window overlaid) (Figure 3) and the scatterplot of calculated versus observed NEE values at the 3-hourly time step for that site and year (Figure 4). A full uncertainty analysis of the approach is beyond the scope of this technical note intended to describe the methodological detail of the downscaling.
Acknowledgements

Funding for this work was provided by NASA’s Carbon Monitoring System (CMS) and NASA’s Carbon Cycle Science (CARBON) programs. We thank the MsTMIP modeling teams for providing the model output. Access and information about MsTMIP model output can be found at http://nacp.ornl.gov/mstmipdata/, along with model and model team participant information.

Funding for the MsTMIP activity was provided through NASA ROSES Grant #NNX10AGO01A. Data management support for preparing, documenting, and distributing MsTMIP model driver and output data was performed by the Modeling and Synthesis Thematic Data Center at Oak Ridge National Laboratory with funding through NASA ROSES Grant #NNH10AN681. We thank Dennis Baldocchi and Siyan Ma for providing the Tonzi Ranch AmeriFlux/FLUXNET data; funding for AmeriFlux data resources and core site data was provided by the U.S. Department of Energy’s Office of Science. The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. Government sponsorship acknowledged.

Copyright 2016. All rights reserved.
References


Table 1. Global terrestrial biosphere models from the Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP) downscaled in this activity.

<table>
<thead>
<tr>
<th>Model</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIOME_BGC</td>
<td>Thornton et al. (2002)</td>
</tr>
<tr>
<td>CLM</td>
<td>Mao et al. (2012)</td>
</tr>
<tr>
<td>CLM4VIC</td>
<td>Li et al. (2011)</td>
</tr>
<tr>
<td>CLASS_CTEM</td>
<td>Huang et al. (2011)</td>
</tr>
<tr>
<td>DLEM</td>
<td>Tian et al. (2012)</td>
</tr>
<tr>
<td>GTEC</td>
<td>Ricciuto et al. (2011)</td>
</tr>
<tr>
<td>ISAM</td>
<td>Jain and Yang (2005)</td>
</tr>
<tr>
<td>LPJ-wsl</td>
<td>Sitch et al. (2003)</td>
</tr>
<tr>
<td>ORCHIDEE</td>
<td>Krinner et al. (2005)</td>
</tr>
<tr>
<td>SIB3</td>
<td>Baker et al. (2008)</td>
</tr>
<tr>
<td>SIBCASA</td>
<td>Schaefer et al. (2008)</td>
</tr>
<tr>
<td>TEM6</td>
<td>Hayes et al. (2011)</td>
</tr>
<tr>
<td>TRIPLEX-GHG</td>
<td>Peng et al. (2002)</td>
</tr>
<tr>
<td>VEGAS2.1</td>
<td>Zeng et al. (2005)</td>
</tr>
<tr>
<td>VISIT</td>
<td>Ito (2010)</td>
</tr>
</tbody>
</table>
Figure 1. The original downscaling approach of Olsen and Randerson (2004) used monthly fixed values, which led to a “stair-stepping” behavior between months (red). This was eliminated by using a 30-day moving window and interpolating monthly input values to 3-hourly time steps (black). Example shown for LPJ model global mean year 2005.
Figure 2. Vegetation productivity (e.g., blues/greens) follows the course of the sun for a single day of net ecosystem exchange (NEE or net CO₂ flux; g C m⁻² 3hr⁻¹) for each 3-hourly period. Shown here, for example, is July 1, 2007 for the weighted ensemble mean product.
Figure 3. The observed net ecosystem exchange (NEE) (blue) and reproduced NEE (red) shown at the 3-hourly time step with daily moving window overlaid for a single year from the Tonzi Ranch AmeriFlux/FLUXNET site (Baldocchi and Ma, 2013).
Figure 4. Observed versus reproduced net ecosystem exchange (NEE) at the 3-hourly time step for a single year at the Tonzi Ranch AmeriFlux/FLUXNET site (Baldocchi and Ma, 2013).