

# **The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project: Part 2 - Environmental driver data**

**Y. Wei<sup>1</sup>, S. Liu<sup>1</sup>, D. N. Huntzinger<sup>2</sup>, A. M. Michalak<sup>3</sup>, N. Viovy<sup>4</sup>, W. M. Post<sup>1</sup>, C. R. Schwalm<sup>5</sup>, K. Schaefer<sup>6</sup>, A. R. Jacobson<sup>7</sup>, C. Lu<sup>8</sup>, H. Tian<sup>8</sup>, D. M. Ricciuto<sup>1</sup>, R. B. Cook<sup>1</sup>, J. Mao<sup>1</sup>, and X. Shi<sup>1</sup>**

<sup>1</sup>Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, TN, USA

<sup>2</sup>School of Earth Sciences and Environmental Sustainability & the Department of Civil Engineering, Construction Mgmt., and Environmental Engineering, Northern Arizona University, Flagstaff, AZ, USA

<sup>3</sup>Department of Global Ecology, Carnegie Institution for Science, Stanford, CA, USA

<sup>4</sup>Laboratoire des Sciences du Climat et l'Environnement, Paris, France

<sup>5</sup>School of Earth Sciences and Environmental Sustainability, Northern Arizona University, Flagstaff, AZ, USA

<sup>6</sup>University of Colorado, National Snow and Ice Data Center, Boulder, CO, USA

<sup>7</sup>NOAA Earth System Research Lab, Global Monitoring Division, Boulder, CO, USA

<sup>8</sup>International Center for Climate and Global Change Research, School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL, USA

## **Abstract**

Ecosystems are important and dynamic components of the global carbon cycle, and terrestrial biospheric models (TBMs) are crucial tools in further understanding of how terrestrial carbon is stored and exchanged with the atmosphere across a variety of spatial and temporal scales. Improving TBM model skills, and quantifying and reducing their estimation uncertainties, pose significant challenges. The Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP) is a formal multi-scale and multi-model intercomparison effort set up to tackle these challenges. The MsTMIP protocol prescribes standardized environmental driver data that are shared among model teams to facilitate model-model and model-observation comparisons. This paper describes the global and North American environmental driver data sets prepared for the MsTMIP activity to both support their use in MsTMIP and make these data, along with the processes used in selecting/processing these data, accessible to a broader audience. Based on project needs and experiences learned from past model intercomparison activities, we compiled climate, atmospheric CO<sub>2</sub> concentrations, nitrogen deposition, land-use and land-cover change (LULCC), C3/C4 grasses fractions, major crops, phenology, and soil data into a standard format for global (0.5° × 0.5° resolution) and regional (North American, 0.25° × 0.25° resolution) simulations. In order to meet the needs of MsTMIP, improvements were made to several of the original environmental data sets, by changing the quality, the spatial and temporal coverage, resolution, or a combination of these. The resulting standardized model driver datasets are being used by over 20 different models participating MsTMIP. The data are archived at the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC, <http://daac.ornl.gov>) to provide long-term data management and distribution.

## 1 Introduction

The need to understand and quantify the role of terrestrial ecosystems in the global carbon cycle and its climate change feedbacks has been driving the development of global terrestrial biogeochemistry and biogeography models since the late 1980s (Foley, 1995). Since that time, the carbon cycle science modeling community has continued to improve understanding of terrestrial ecosystems in global and regional carbon cycling (US CCSP, 2011). Due to limited availability of direct observations that can be used to assess model results at their spatial and temporal scales, investigating variability of different model results and its attribution to inter-model differences becomes an important way for and diagnosing and improving models (Huntzinger et al., 2013). Since 1990s, there had been many multi-model intercomparison projects (MIPs) conducted. The Vegetation-Ecosystem Modeling and Analysis Project<sup>1</sup> (VEMAP) was an early pioneer MIP activity that focused on evaluating the sensitivity of terrestrial ecosystem and vegetation processes to altered climate forcings and elevated atmospheric CO<sub>2</sub> for the continental United States. It achieved this goal through intercomparison and sensitivity analysis of multiple ecosystem physiology models and biome distribution models (Schimel et al., 1997). The Potsdam NPP MIP was an intercomparison activity focusing on annual and seasonal fluxes of net primary productivity (NPP) for the land biosphere involving a wide range of 17 global TBMs (Cramer et al., 1999). GCP-TRENDY<sup>2</sup>, part of the Global Carbon Project (GCP), organized and performed a factorial set of model simulations over the historical period (1901-2010) to investigate further the global spatial trends in net land-atmosphere carbon exchange. These and many other multi-model intercomparison activities had been progressively improving the understanding of ecosystem dynamics.

Huge challenges still remain, however, especially in developing approaches for evaluating model predictions and assessing the uncertainties associated with model estimates (e.g., Randerson et al., 2009; USCCSP, 2011; Schwalm et al., 2013). The challenges associated with representing terrestrial ecosystem fluxes of carbon dioxide are illustrated by the huge variability in model predictions observed as part of the recent North American Carbon Program (NACP) regional and site interim synthesis activities (e.g., Huntzinger et al., 2012; Schaefer et al., 2012). The results

---

<sup>1</sup> The Vegetation-Ecosystem Modeling and Analysis Project. <http://www.cgd.ucar.edu/vemap/>.

<sup>2</sup> Trends in Net Land-Atmosphere Carbon Exchange. <http://dgvn.ceh.ac.uk/node/21>.

from these model intercomparison activities confirmed the large uncertainties associated with our ability to represent terrestrial ecosystem carbon fluxes, but the reliance of the regional synthesis on “off the shelf” simulations without a prescribed protocol or standardized driver data sets limited the degree to which the observed variability could be attributed to specific sources of uncertainty.

Four types of uncertainties drive differences between predictions of terrestrial carbon flux (e.g., Enting et al., 2012): uncertainty associated with (1) the choice of driver data, (2) parameter values, (3) initial conditions, as well as (4) the choice of processes to include and how these processes are represented within the model (i.e., structural uncertainty). Estimating and reducing these uncertainties are both critical to improving model performance, and consequently to understanding the role of terrestrial ecosystems in the global carbon cycle.

In response to this need, the Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP) was established to provide a consistent and unified modeling framework to interpret and address structural and parameter uncertainties (Huntzinger et al., 2013). Based on experiences and lessons learned from past and on-going MIPs, MsTMIP is a continuation and evolution of the multi-model intercomparison activities. MsTMIP prescribes standardized environmental driver data and a consistent spin-up protocol for all model simulations. This approach greatly reduces the inter-model variability caused by differences in environmental drivers, initial conditions, and the process used for defining steady-state conditions. This approach makes it possible to focus the analysis on the ways in which the structure of TBMs (i.e., their choice and formulation of ecosystem processes) and associated internal parameters impact a model’s estimates of terrestrial ecosystem carbon dynamics.

This paper describes the driver data needs of MsTMIP and outlines the environmental driver data sets compiled and synthesized for the MsTMIP activity. In doing so, this paper aims to address the needs of multiple communities and audiences. First, it provides the detailed background about environment driver data choices that is necessary for the scientific interpretation of modeling results coming out of the MsTMIP effort. As such, it addresses the needs of researchers focusing on the scientific interpretation of the MsTMIP results. Second, it provides the rationale for the choice of specific environmental driver data and the details associated with their processing. Thus, the paper also aims to address the needs of researchers who wish to

leverage the work reported here by using the driver data for follow-on studies or related applications. Third, this paper reports on the decision making and implementation process involved in putting together common driver data for large modeling studies and intercomparison efforts, including lessons learned that are independent of the specific applications addressed by MsTMIP. As such, this paper also aims to inform future efforts focused on assembling consistent data sets for use by multiple modeling teams.

The remainder of this paper is structured to address the needs of the three intended audiences described above. For each data category, we first provide a brief review of the data source chosen for MsTMIP and the rationale for the choice, along with a description of other similar data sources currently available. We then describe the processing and analysis completed to convert the original data source into a form meeting the needs of the MsTMIP activity, and in some cases to improve the quality of the original data source. We also provide a brief evaluation of standardized MsTMIP data products, and suggestions on how the data should be used in terrestrial biosphere modeling. Following the MsTMIP environmental driver data sets, we briefly summarize their connections to and differences from input data sets used in past and on-going MIPs. Finally, we introduce some lessons learned on data processing and management, to guide future data-intensive projects.

## **2 Driver data needs of MsTMIP**

The overarching goal of the MsTMIP activity is to provide a unified intercomparison framework that allows for the critical synthesis, benchmarking, evaluation, and feedback needed to improve TBMs (Huntzinger et al., 2013). To meet this goal, the MsTMIP activity is conducting a suite of simulations that can be used to quantify: (1) the impact of the scale and spatial resolution of model simulations on model estimates; and (2) the additive influence of a suite of time-varying environmental drivers or forcing factors on model estimates of carbon stocks and fluxes. As such, MsTMIP includes simulations over two spatial domains and resolutions: globally at  $0.5^\circ \times 0.5^\circ$  resolution and regionally over North America at  $0.25^\circ \times 0.25^\circ$  resolution. To evaluate the additive impacts of different types of forcing, e.g. climate, land-use and land-cover change (LULCC), atmospheric CO<sub>2</sub> concentrations, and nitrogen deposition, on model estimates of carbon fluxes and stocks, a series of sensitivity simulations are prescribed at both spatial scales

for a simulation period from 1801 to 2010 (Huntzinger et al., 2013). Inherent to MsTMIP's experimental design is the focus on controlling for as many sources of variability in TBM predictions as possible, to isolate and quantify the impact of the model itself (i.e., structural and parameter uncertainties) on estimates.

One source of variability in model estimates is the choice of (and uncertainty associated with) environmental driver and input data sets. Most uncoupled TBMs require, at a minimum, a land-water mask, climate forcing data, soil characteristics, and atmospheric CO<sub>2</sub> concentrations to simulate how carbon is exchanged between the land and atmosphere. Many models also require additional information such as LULCC, phenology, nitrogen deposition rates, and disturbance history. Ideally, the temporal resolution of drivers should be fine enough to enable prediction at sub-daily temporal resolution, thus making it possible to investigate the diurnal cycle of carbon and energy fluxes. To meet the objectives of MsTMIP's experimental design, the goal was to provide modeling teams, to the extent possible, with a complete and consistent set of environmental driver data. In addition to being of high quality, the environmental driving and input data chosen for MsTMIP also needed to meet the following requirements:

- Data sets must be compatible with over 20 different TBMs;
- Data sets must provide consistent spatial coverage for the land surface within the two simulation domains: (1) North American: 10–84° North latitude; 50–170° West longitude, and (2) global: all land surface areas excluding Antarctica;
- Spatial resolutions must be compatible with the two sets of simulations: (1) North American (0.25° × 0.25°) and (2) global (0.5° × 0.5°);
- Temporal resolution and extent must be compatible with the two sets of simulations: (1) North American (3-hourly, 1801–2010) and (2) global (6-hourly, 1801–2010);
- Data sets must provide smooth transitions in time, without any unrealistic spikes or discontinuities; and
- Data sets must be physically consistent with one another. For example, climate, soil, and land cover change history needed to represent the same land domain as indicated in the land-water mask, and the prescribed phenology data needed to be consistent with the time-varying land cover data for each time step.

The environmental driver and input data sets chosen for the MsTMIP activity are a reflection of these overall project needs and requirements.

### **3 Environmental driver data sets**

MsTMIP environmental driver and associated data products include data sets describing climatology, time-varying atmospheric CO<sub>2</sub> concentrations, time-varying nitrogen deposition, LULCC, C3 /C4 grasses fractions, major crop distribution, phenology, soil characteristics, and a land-water mask, all at 0.5° × 0.5° for the global domain and 0.25° × 0.25° for the North American domain (Table 1). All MsTMIP model driver data files are stored in Climate and Forecast (CF)<sup>3</sup> 1.4 convention compliant netCDF version 3 format, which is supported by a wide range of programming APIs (e.g., C, C++, Fortran, Java, Perl) and multiple operating systems (e.g., Linux, Unix, Mac OS X, Windows). All drivers are saved in Greenwich Mean Time (GMT) with all sub-monthly drivers (e.g., climate) including leap years.

For most data categories, the North American data sets are based on the same data sources as the global products. We did, however, choose different climatology and soil data products for the two domains. This decision was driven primarily by the availability of these drivers at the spatial and temporal resolution needed for the regional simulations. However, by holding the source of other drivers constant between the global and North American simulations, we are also creating an opportunity to test the impact of the choice of climate and soil characteristics on model estimates.

#### **3.1 Climate**

##### **3.1.1 Global climate: CRU-NCEP**

Several reanalysis and observation-based gridded global climatology data sets exist, including products produced by the Climate Research Unit (CRU) (Harris et al., 2013), the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research (NCAR) Reanalysis 1 (Kalnay et al., 1996), and the European Centre for Medium-Range

---

<sup>3</sup> NetCDF Climate and Forecast (CF) Metadata Conventions, version 1.4. <http://cf-pcmdi.llnl.gov/documents/cf-conventions/1.4/cf-conventions.html>.

Weather Forecasts (ECMWF) (Uppala, et al., 2005; Dee, et al., 2011). However, none of the available climatology data sets fully met the spatial and temporal requirements of MsTMIP. The CRU Time Series (TS) 3.2 product covers the time period from 1901 to present at a  $0.5^\circ$  spatial resolution, but only at a monthly temporal resolution. The NCEP/NCAR product, on the other hand, has a finer temporal resolution (6-hourly) but has a coarse spatial resolution ( $2.5^\circ$ ) and only provides climatology back to 1948. The ECMWF product similarly lacks the temporal coverage required for MsTMIP.

We thus combined the strengths of the CRU and NCEP/NCAR Reanalysis products, fusing them to produce the “CRU-NCEP” global climate data set. This new data set provides a globally gridded ( $0.5^\circ \times 0.5^\circ$ ) and sub-daily (6-hourly) time-varying climatology product that spans the period between 1901 and 2010. CRU-NCEP contains 7 climatology variables, including downward longwave and shortwave radiations, pressure, air specific humidity, precipitation, temperature, and wind (Table 1). In the process of creating this new climatology product, we also corrected known biases in temperature and shortwave radiation in the NCEP/NCAR Reanalysis product. Zhao et al. (2006) showed that NCEP/NCAR Reanalysis climatology overestimates downward shortwave radiation, especially in non-tropical regions, and underestimates surface temperature for almost all latitudes. Biases in climatological variables can introduce substantial errors into Gross Primary Productivity (GPP) and Net Primary Productivity (NPP) estimates (Zhao et al., 2006). By fusing NCEP/NCAR with the CRU climatology, we forced the monthly amplitude of CRU-NCEP product to be consistent with the observation-based CRU climatology, while preserving the diurnal variability in the NCEP/NCAR Reanalysis product. A comparison between the zonal mean of long-term mean downward shortwave radiation for each  $0.5^\circ$  grid cell over land (Fig. 1) shows that CRU-NCEP has lower downward shortwave radiation than the original NCEP/NCAR data, except at  $0\text{--}10^\circ$  North and  $50\text{--}55^\circ$  South, where CRU-NCEP downward shortwave radiation is similar to NCEP/NCAR Reanalysis 1.

### **3.1.2 North American climate: NARR**

Several climatology products are available for North America at finer spatial and temporal resolutions than the new CRU-NCEP product. In addition to better addressing the resolution needs of MsTMIP regional simulations ( $0.25^\circ \times 0.25^\circ$  spatial and 3-hourly temporal resolution), using a different climate driver data product for the North American simulations: (1) makes it



possible to test the influence of the choice of climate drivers on model estimates, and (2) provides a closer linkage between model estimates and fine-scale ground-based observations. Both the Daymet (Thornton et al., 2012) and Parameter elevation Regression on Independent Slopes Model (PRISM)<sup>4</sup> products provide temperature and precipitation data at high spatial resolution (e.g., 1km) for North America. However, the temporal resolutions of these products (PRISM: monthly; Daymet: daily) do not meet the needs of MsTMIP, and these data products also do not cover the full spatial extent of the North American simulations (10–84° North; 50–170° West).

The NCEP North America Regional Reanalysis (NARR), on the other hand, provides long-term, high-resolution, high-frequency atmospheric and land surface meteorological data for the North American domain (Mesinger et al., 2006). The NARR climatology begins in 1979 and extends to 2012 at 3-hourly temporal and 32 km spatial resolutions. Although the temporal coverage is shorter than desired, the NARR product was selected for the MsTMIP activity, because it best matched the needs of the North American simulations, and the time covered by the data set was extended as described in Sect. 4. The original NARR data were provided by the NOAA/OAR/ESRL PSD<sup>5</sup>, available at <http://www.esrl.noaa.gov/psd/> (last access: 14 January 2011).

The NARR variables were regridded to a spatial resolution of  $0.25^\circ \times 0.25^\circ$ , from their original Lambert Conformal Conic Projection at 32km resolution using both area-weighted and distance-weighted averages. An area-weighted averaging method was used for precipitation and radiation flux variables in order to conserve their total magnitude for North America. For highly spatially auto-correlated state variables (e.g., air temperature, humidity), distance-weighted averaging was used because values for these variables tend to cluster together in space. The  $U$  direction (along longitude) and  $V$  direction (along latitude) wind speeds were combined into an overall surface wind velocity variable prior to the regridding process.

In a study of rain gauge and NARR data, Sun and Barros (2010) found that, although NARR reproduces the spatial patterns of precipitation, it underestimates the frequency and magnitude of large rainfall events. In addition, Xie et al. (2003) found that the Global Precipitation

---

<sup>4</sup> PRISM Climate Group, Oregon State University, <http://prism.oregonstate.edu>, created 4 Feb 2004.

<sup>5</sup> NOAA/OAR/ESRL PSD: National Oceanic & Atmospheric Administration / Oceanic and Atmospheric Research / Earth System Research Laboratory Physical Sciences Division, Boulder, Colorado, USA.

Climatology Project (GPCP) monthly gridded ( $2.5^\circ \times 2.5^\circ$ ) precipitation product, derived from satellite and gauge measurements, reproduced spatial patterns of total precipitation with relatively high quality especially over land. Thus, to remove biases in the precipitation, we rescaled the NARR 3-hourly precipitation using the GPCP v2.1 (Adler et al., 2003). Although the GPCP product has a relatively coarse spatial resolution of  $2.5^\circ$ , it has the advantage of including a correction to compensate for systematic biases in gauge measurements due to wind, gauge wetting, and gauge evaporation. Applying this rescaling allowed us to retain the advantages provided by the NARR data product, while also leveraging the information provided by GPCP. To rescale the NARR precipitation, for each month, precipitation of all 3-hourly  $0.25^\circ$  NARR grids within each  $2.5^\circ$  GPCP grids were summed up along time, averaged over space, and linearly rescaled to match the magnitude of total monthly GPCP precipitation. Figure 2 shows the difference map between the long-term mean (1979–2010) annual total precipitation from rescaled NARR and original NARR products. The rescaled product better represents the magnitude of extreme rainfall events at the coastline of Gulf of Alaska and Central America, while also preserving both the magnitude and spatial pattern in most other areas of North America. This rescaling, however, doesn't perform calibration to the frequency of rainfall events.

As mentioned previously, biases in shortwave radiation can have a strong impact on model estimates of GPP. Kennedy et al. (2010) showed that between 1999–2001 the NARR product overestimates downward shortwave radiation flux relative to the Atmospheric Radiation Measurement (ARM) Southern Great Plains (SGP) site observations by about 10% under clear-sky and by about 30% under all-sky conditions. We also compared NARR downward shortwave radiation flux with observations from 23 FLUXNET sites across North America. For the FLUXNET sites examined, NARR overestimates downward shortwave radiation by about 30%, with higher positive bias under cloudy conditions (Fig. 3). The weather simulation model MTCLIM version 4.3 was thus used to reduce the shortwave radiation bias in the NARR product. Given input data from one location, MTCLIM generates weather information for another location with potentially different elevation, slope, and aspect from the input location (Running et al., 1987; Thornton and Running, 1999). Bohn et al. (2013) found that MTCLIM performed reasonably well for estimating downward shortwave radiation under most climate conditions for most of the global land surface, except coastal areas with negative bias of about -26%. It also found that MTCLIM v4.3's snow correction significantly reduced the bias in snow-

covered areas. We calculated the total daily shortwave radiation for each grid cell using the MTCLIM model driven by gridded daily maximum and minimum temperature and total daily precipitation derived from the 3-hourly NARR original temperature and rescaled precipitation. The original 3-hourly NARR downward shortwave radiation values were then linearly rescaled to match the total daily downward shortwave radiation generated from MTCLIM model. This process was effective at reducing the overall positive bias in shortwave radiation (Fig. 4), such that the rescaled NARR product better matches observed radiation at FLUXNET sites (Fig. 3).

### **3.1.3 Comparison of global and North American climate data**

To address

## **3.2 Land-water mask**

The land-water mask specifies the land grid cells on which MsTMIP global and regional simulations are run, and needed to be consistent with the climate driver data. We therefore based the global land-water mask on the CRU-NCEP land-water mask, and the North American land-water mask on the original NARR mask regridded to a spatial resolution of  $0.25^\circ \times 0.25^\circ$  using an area-weighted method to preserve the total amount of land area. Since a regridding process was involved for the preparation of NA land-water mask, to preserve only those  $0.25^\circ$  grid cells covered primarily by land, a threshold of 50 % was then applied to define “land” grid cells.

## **3.3 Atmospheric CO<sub>2</sub> concentration**

Atmospheric CO<sub>2</sub> concentrations have risen more than 40 % over pre-industrial levels. Increased atmospheric CO<sub>2</sub> content influences global climate not only through its greenhouse radiative effect, but also through its physiological effect (Sellers et al., 1996a; Ainsworth and Long, 2005). Under elevated CO<sub>2</sub> concentration, plant stomata open less widely, leading to reduced plant transpiration (Cao et al., 2010; Shi et al., 2011). In natural ecosystems, this CO<sub>2</sub> fertilization effect is modulated by many other factors, including access to light, water, and other nutrients. Furthermore, the net terrestrial sink inferred from analysis of atmospheric CO<sub>2</sub> distributions (e.g., Gurney et al., 2002) is due not only to increased productivity of natural ecosystems but also to historical land use (e.g., Pacala et al., 2001). Models are useful for simulating the complex interplay of these factors, and studies have suggested that of the major factors affecting

simulated net carbon exchange between the atmosphere and the terrestrial biosphere, CO<sub>2</sub> fertilization may have the strongest decadal trend (e.g., Norby et al., 2005; Kicklighter et al., 1999; McGuire et al., 2001). A realistic CO<sub>2</sub> concentration history was therefore needed for the entire MsTMIP simulation period.

The atmospheric CO<sub>2</sub> concentration data prepared for the MsTMIP are consistent with the GLOBALVIEW-CO<sub>2</sub> (2011) data product (henceforth GV), the time series of historic atmospheric CO<sub>2</sub> from Antarctic ice cores (MacFarling Meure et al., 2006), fossil fuel emissions (Marland et al., 2008), and atmospheric CO<sub>2</sub> observations at Mauna Loa (MLO) and the South Pole (SPO). During the period 1979–2010, when the temporally- and meridionally-resolved GV product is available, atmospheric CO<sub>2</sub> concentrations are set directly to the GV marine boundary layer reference surface interpolated to the MsTMIP global and North American grids. Prior to 1979, we preserve the 1979–2010 mean annual cycle from GV, and impose onto it a modeled CO<sub>2</sub> surface that represents annual mean concentrations and a time-evolving meridional gradient. Following Conway and Tans (1999), the annual mean difference between MLO and SPO in the GV product is modeled as a linear function of fossil fuel (FF) emissions (Marland et al., 2008). Extrapolated to zero FF emissions, the pre-industrial MLO-SPO difference estimated in this manner is 0.3 ppm. Performing this same exercise using Scripps CO<sub>2</sub> program observations at MLO and SPO instead of GV yields a stronger dependence of the meridional gradient on FF emissions and a pre-industrial MLO-SPO difference of –1.2 ppm. While it is possible that pre-industrial Southern Hemisphere CO<sub>2</sub> values exceeded those in the Northern Hemisphere (Conway and Tans, 1999), we judge that it is more parsimonious to assume a small pre-industrial inter-hemispheric CO<sub>2</sub> gradient, which the GV-based scheme achieves natively. The MsTMIP atmospheric CO<sub>2</sub> product agrees well with Scripps CO<sub>2</sub> data before 1979 at SPO and MLO (Fig. 5a), and with Law Dome ice core data in Antarctica (MacFarling Meure et al., 2006; Fig. 5b) data. The MsTMIP atmospheric CO<sub>2</sub> product before 1979, however, does not represent inter-annual variability other than that derived from variability in FF emissions, and it does not include speculative changes in the magnitude or phase of annual cycles of CO<sub>2</sub> in the atmosphere.

### 3.4 Nitrogen deposition

Nitrogen enrichment, increasing atmospheric nitrogen deposition in particular, has been recognized as one of the most significant global changes since it could stimulate plant growth, enhance terrestrial carbon sequestration capacity, and thus mitigate global climate warming (e.g., Holland et al., 1997; Pregitzer et al., 2008; Reay et al., 2008; De Vries et al., 2009). Models failing to capture nitrogen input and nitrogen cycling may overestimate ecosystem carbon uptake (Hungate et al., 2003). Up to now, more and more TBMs include nitrogen deposition as an important driving force. However, few global and North American nitrogen deposition products are available over the full period required by MsTIMP. Monitoring networks of nitrogen deposition in the United States and Europe were launched in the late 1970s, while other countries began such nationwide observations later (Holland et al., 2005; Lu and Tian, 2007). The Dentener global nitrogen deposition data product was generated using a three-dimensional chemistry transport model that estimated atmospheric deposition of total inorganic nitrogen (N),  $\text{NH}_x$  ( $\text{NH}_3$  and  $\text{NH}_4^+$ ), and  $\text{NO}_y$  (all oxidized forms of nitrogen other than  $\text{N}_2\text{O}$ ) for the years 1860, 1993, and 2050 at a spatial resolution of  $5^\circ$  longitude by  $3.75^\circ$  latitude (Dentener, 2006; Galloway et al., 2004). Most TBMs that include nitrogen deposition as an input driver do so by linearly interpolating Dentener's three-year maps into annual time-series data, ignoring the different changing trends among different regions and different periods (Jain et al., 2009; Zaehle et al., 2010).

To address the above issue, we used a different approach as described in Tian et al. (2010) and Lu et al. (2012) to create a time-varying annual nitrogen deposition data set for both global ( $0.5^\circ \times 0.5^\circ$  resolution) and North American ( $0.25^\circ \times 0.25^\circ$  resolution) simulations based on Dentener's maps and introduce spatial and temporal variations from nitrogen emissions. This approach took the following assumptions. For the time period between 1890 and 1990, annual variations in nitrogen deposition rate ( $\text{NH}_x\text{-N}$  and  $\text{NO}_y\text{-N}$ ) were defined by assuming that temporal trends of N-deposition are consistent with EDGAR-HYDE 1.3 nitrogen emission data (Van Aardenne et al., 2001). The EDGAR-HYDE product provides gridded ( $1^\circ \times 1^\circ$  resolution) annual total emissions of  $\text{NH}_3$  and  $\text{NO}_x$  from 10 anthropogenic sources. Nitrogen deposition was assumed to change linearly over the remaining time periods (1860–1890 and 1990–2010).

### 3.5 Land-use & land-cover change

LULCC has considerable influence on the biogeochemical cycling of carbon (e.g., Friedlingstein et al., 2010; Pielke Sr. et al., 2011; Sohl et al., 2012). Activities such as afforestation (Potter et al., 2007) or deforestation (Ramankutty et al., 2007) can alter carbon stocks. Similarly, biomass burning used in land clearing results in direct carbon emissions (Giglio et al., 2010). Despite its importance in carbon cycle dynamics, LULCC-caused CO<sub>2</sub> emissions are poorly constrained and highly uncertain with a global mean (2000–2009) value of  $1.0 \pm 0.5$  Pg C yr<sup>-1</sup> (Le Quéré, 2013).

Many global data products describing historical LULCC are available (e.g., Hurtt et al., 2011; Klein Goldewijk et al., 2011). In an effort to hold as many of the environmental drivers constant as possible in the MsTMIP activity, we chose to prescribe LULCC by merging a static satellite-based land cover product, SYNMA (Jung et al., 2006), with the time-varying land use harmonization (LUH) data for the fifth Assessment Report (AR5) of the Intergovernmental Panel on Climate Change (IPCC) (Hurtt et al., 2011). We chose the LUH product based on its global coverage, inclusion of land use change fractions (required for a subset of participating models), overlap with the time horizon of MsTMIP simulations, and use in the IPCC process. The LUH product was derived using a bookkeeping approach based on historical time series of crop and pasture data, national wood harvest, shifting cultivation, and population (Hurtt et al., 2011). LUH product provides mapped fractional coverages and underlying annual land use transitions for six land use classes (primary land, secondary land, cropland, pasture, urban, and barren) at  $0.5^\circ \times 0.5^\circ$  spatial resolution. The historical LUH data (1801–2005) were combined with a future projection (2006–2010) to match the time horizon of MsTMIP model simulations (1801–2010). This future projection was based on the Representative Concentration Pathway (RCP) (van Vuuren et al., 2011) 4.5 scenario, which hypothesizes a net radiative forcing of  $4.5 \text{ W m}^{-2}$  ( $\sim 650$  ppm CO<sub>2</sub> eq) by the end of the century based on a set of greenhouse gas emissions and concentrations as well as land-use trajectories.

As TBMs require a different land use/cover scheme than the six classes associated with the LUH we merged the 1801–2010 LUH with the static 2000/2001 SYNMAP land cover product (Jung et al., 2006). Although numerous land cover products exist, we chose SYNMAP due to its: (1) reconciliation of multiple global land cover products, i.e., Global Land Cover Characterization Database (GLCC) (Hansen et al., 2000; Loveland et al., 2000), GLC2000 (2003), and the 2001

MODIS land cover product (Friedl et al., 2002); (2) global coverage at 1km resolution; and (3) general definition of classes based on life form, leaf type, and leaf longevity which allowed for simple crosswalks to plant functional types (PFTs) used in different TBMs. Generality was a key concern as PFT schemes used in TBMs vary widely. The SYNMAP product contains 47 land cover classes such that a PFT scheme for a given TBM is a subset of SYNMAP classes based on a crosswalk between the two different schemes.

To provide annual maps of LULCC, LUH and SYNMAP were merged using a set of one-to-one and one-to-many mapping rules based on map intersection during their period of overlap, i.e., both products exist for 2000–2001. These invariant grid cell- specific mappings were then used to translate the six LUH classes to the 47 SYNMAP classes (Jung et al., 2006) for each annual LUH map. For example, assume a grid cell with LUH pasture at a fractional coverage of 0.5 for 2000–2001. In that same grid cell the SYNMAP product has only two eligible target classes: the shrubs and the grasses classes with fractional coverages of 0.2 and 0.4, respectively. This map intersection forms the basis of a one-to-many mapping, i.e., 0.5 LUH pasture is equivalent to 0.17 SYNMAP shrubs plus 0.33 SYNMAP grasslands, which preserve the original shrubs-grasslands ratio in SYNMAP for that grid cell. This scalable mapping rule is used for all other time steps for this grid cell and reflects the legacy of grid cell-specific changes in land use/cover through time.

Few models use these 47 SYNMAP classes directly in their simulations. For example, the Simple Biosphere (SiB) model uses 12 biome classes (Sellers et al., 1996b). In such instances, model teams developed crosswalks from the 47 SYNMAP classification scheme to their internal schemes. Given that many SYNMAP classes are mixed classes, e.g., “Shrubs and Crops” and “Trees and Crops”, which cannot be accommodated by some models, we created maps of “pure” biome classes by assuming each component in a mixed class was half the total area. Finally, as several models require information on photosynthetic the pathway in grasslands as well as crop types we also provided invariant maps for C3/C4 grass types (Sect. 3.6) and major crops (Sect. 3.7).

### **3.6 C3 and C4 grass fractions**

Because photosynthesis can vary significantly between species using the C3 and C4 photosynthetic pathways (Ehleringer and Cerling, 2002), most TBMs use separate algorithms for estimating GPP of C3 and C4 plant types. In order to provide the required spatial distribution of ecosystems dominating each of these pathways, we used an approach described in Still et al. (2003) based on growing season temperature. Since the C4 pathway is largely found in warm season grass species, we created a global gridded ( $0.5^\circ \times 0.5^\circ$ ) map of the relative fraction of C3 and C4 grasses using the “present” climate state based on the CRU-NCEP mean monthly precipitation and temperature data for 2000–2010. For grid cells characterized as grasslands (or containing grasslands) the relative fraction map defines the fraction of those grasses that are C3 or C4, so that in each of those grid cells the C3 and C4 grass fractions sum to 1 regardless of the total percentage of grassland contained in the grid cell.

SYNMAP contains 13 land cover classes that include grasses, with 12 of these mixtures of grasses with trees, shrubs, crops, or barren land. For the mixed classes, we assumed that grasses account for 50 % of the area of these mixed classes contained in each cell. The SYNMAP grass fraction in each cell was calculated as the sum of the grass fraction of all different classes, including both pure and mixed classes, in the cell. Figure 6 shows the relative fraction of C3 (top) and C4 (bottom) grassland globally ( $0.5^\circ$ ) under “present” (2000–2010) climate conditions. The actual C3 and C4 grassland fractions depend on the overall grass coverage and can be zero if no grass is present in a particular grid cell.

The North American ( $0.25^\circ \times 0.25^\circ$ ) C3 and C4 relative grassland fraction maps were created using the same approach, except that the NARR climate was used instead of CRU-NCEP. MsTMIP only provides a constant C3/C4 data product under “present” climate conditions. For models that need time-varying C3/C4 grass fractions, the same approach can be applied to historical land cover data and historical precipitation/temperature climate data to generate C3/C4 grassland maps for previous years.

### **3.7 Major crops**

The SYNMAP land cover map indicates which areas are predominantly crop but does not provide additional information about the crop types contained within each grid cell. This can be



important when, for example, a C4 crop like maize dominates a grid that would normally be covered by C3 vegetation, and vice-versa. Some models make use of such additional information to implement crop specific algorithms that capture some aspects of crop physiology and management including planting and harvesting phenology, fertilizer applications, irrigation, or tillage practices. We therefore identified and extracted the four globally significant crop types (maize, rice, soybean, and wheat) from the Monfreda et al. (2008) global crop database. The original Monfreda global crop product is a detailed database of global agricultural practices and describes the areas and yields of 175 different individual crops in 2000 at a 5min×5min (approximately 10 km × 10 km) spatial resolution. We resampled the original Monfreda crop data to 0.5° × 0.5° (global) and 0.25° × 0.25° (North American) spatial resolutions. These major crop designations do not provide detailed model simulation prescription, but rather guidance for models that need to specify crop types or cropping systems.

### **3.8 Phenology**

Some models do not have prognostic canopies and use remote sensing products to prescribe plant phenology to calculate GPP or NPP. Consequently, we constructed monthly maps of Normalized Difference Vegetation Index (NDVI), Leaf Area Index (LAI), and absorbed fraction of Photosynthetically Active radiation (fPAR) consistent with the MsTMIP LULCC data on both global and North American grids for 1801–2010. For NDVI data, we chose the Global Inventory Monitoring and Modeling System version g (GIMMSg) data set (Tucker et al., 2005), because it provides the longest global observation-based product. GIMMSg consists of 15-day maximum value composites at about 8 km spatial resolution from 1982 to 2010 adjusted for missing data, satellite orbit drift, sensor degradation, and volcanic aerosols (Tucker et al., 2005). We used the average seasonal cycle in NDVI for the entire time period from 1801 to 2010, since switching to observed values in 1982 would create abrupt changes in model output that would be difficult to interpret. The 15-day GIMMSg NDVI was first regridded to 0.5° × 0.5° (global) and 0.25° × 0.25° (North American) resolutions using area-weighted averaging. The NDVI data were fitted to the MsTMIP land masks using the nearest-neighbor technique to gap fill missing points. To minimize noise due to cloud and aerosol contamination, we converted the regridded 15-day GIMMSg NDVI to monthly maximum value composites and then calculated the average of all January maps, the average of all February maps, etc., to create the average NDVI seasonal cycle.

We calculated fPAR and LAI from the average seasonal cycle of GIMMSg NDVI using methods described in Sellers et al. (1996b) and Schaefer et al. (2002).

To harmonize phenology data with the LULCC used in MsTMIP, we assumed that a pixel would consist of tiles, each corresponding to a different land use/cover class with fractional areas set by the MsTMIP LULCC coverage maps as a function of year from 1801 to 2010. We first calculated maps of LAI and fPAR assuming the entire land surface was one of the 12 SiB biome classes (Sellers et al., 1986) resulting in 12 sets of LAI and fPAR maps corresponding to the 12 SiB biome classes, all calculated from the same NDVI values, but using different parameter values unique to each biome (Sellers et al., 1996b). We then mapped the 12 SiB biomes to the 47 SYNMAP land use/cover classes using one-to-one or one-to-many mapping, resulting in 47 sets of LAI and fPAR maps corresponding to the 47 SYNMAP classes. This two-step process was required because the parameters used to calculate LAI and fPAR are not available for each of the 47 SYNMAP types. Combining these 47 sets of LAI and fPAR maps and the yearly MsTMIP LULCC data, the time-evolving and land use/cover class explicit LAI and fPAR data products were created. If a grid cell did not contain a particular SYNMAP class in a specific year, a standard missing value was inserted into the corresponding LAI and fPAR maps. A model would then extract the LAI and fPAR values for a particular SYNMAP class in each year and use it for the corresponding tile.

### **3.9 Soil**

The Food and Agriculture Organization – United Nations Educational, Science and Cultural Organization (FAO-UNESCO) digitized soil map of the world (FAO, 1971– 1981, 1995, 2003), originally published in 1974, is commonly used in terrestrial biosphere modeling. Recently, however, significant improvements in soil mapping and databases of soil properties have led to a new generation of regional and global scale soil maps, such as the International Soil Reference and Information Centre (ISRIC) World Inventory of Soil Emission Potentials (ISRIC-WISE) (Batjes, 2008) and the Harmonized World Soil Database (HWSD) (FAO/IIASA/ISRIC/ISS-CAS/JRC, 2011). This new generation of soil products have increased details in the spatial distribution of soil types and more accurate characterizations of soil physical and chemical properties.

For MsTMIP, we selected and synthesized the HWSD v1.1 for global simulations because it was the most recent global soil database that incorporates updated soil data from Europe, Africa, and China. However, in both the ISRIC-WISE and HWSD databases, soil information for North America is based on an outdated FAO-UNESCO soil map from the 1970s. Thus, even in the most updated global soil databases, North America is less reliable than the other regions due to the use of an obsolete database (Batjes, 2005; FAO/IIASA/ISRIC/ISS-CAS/JRC, 2011). We therefore developed the Unified North American Soil Map (UNASM) by fusing the United States Department of Agriculture Natural Resources Conservation Services (USDA-NRCS) State Soil Geographic (STATSGO2) data set with both the Soil Landscapes of Canada (SLC) version 3.2 and 2.2 products, and the HWSD v1.1 (Liu et al., 2013).

Both data prepared for MsTMIP, the gridded 0.5° HWSD for global simulations and 0.25° UNASM for North American simulations, contain two standardized soil layers. The topsoil layer ranges from 0 to 30 cm and the subsoil layer ranges from 30 to 100 cm. For each soil layer, 8 physical and chemical soil properties, including clay/sand/silt fractions, pH, organic carbon, cation exchange capacity, reference bulk density, and gravel content, were compiled (Table 1). These variables are used by TBMs to calculate soil column hydrological characteristics that determine the dynamics of available soil water for plant transpiration and soil evaporation. Organic carbon content is provided for models that make use of an estimate for initialization.

### **3.9.1 Global soil: gridded HWSD**

The HWSD version 1.1 was used as the source for MsTMIP global soil data. The original HWSD is a 30arcsec raster database with over 16000 different soil mapping units that combines existing regional and national updates of the soil information worldwide, including the Soil and Terrain database (SOTER), European Soil Database (ESD), Soil Map of China, and WISE, with the information contained within the 1 : 5 000 000 scale FAO-UNESCO soil map of the world (FAO/IIASA/ISRIC/ISSCAS/JRC, 2011).

Each soil mapping unit in the HWSD is composed of several different soil units (or soil types) defined by major soil group code following a combined FAO-74/FAO-85/FAO-90 soil classification system. For the global simulations, the original HWSD was regridded to a spatial resolution of 0.5° × 0.5° by selecting the dominant soil type within each grid cell. Eight physical

and chemical soil properties associated with the dominant soil type in each soil layer were then selected. In addition to physical and chemical soil properties for each dominant soil type, we also provided modelers with the HWSD reference soil depth, as a proxy for mineral soil depth, even though this reference soil depth is not precise.

The reference bulk density values provided in HWSD v1.1 were calculated following the method developed by Saxton et al. (1986) that relates bulk density to soil texture. This method, although generally reliable, tends to overestimate the bulk density in soils that have a high porosity (e.g., Andosols) or that are high in organic matter content (e.g., Histosols). Therefore, the bulk density values of these two soil types were corrected using the corresponding depth-weighted average values from ISRIC-WISE, version 1.0. Figure 7 shows the globally gridded HWSD topsoil reference bulk density before and after correction. The correction mainly impacts the North American boreal region and a few places of southeastern Asia where Andosols and Histosols dominate.

### **3.9.2 North American soil: Unified North American Soil Map (UNASM)**

A new gridded database of harmonized soil physical and chemical properties for North America was created for MsTMIP by fusing the most recent regional soil information from US STATSGO2, SLC version 3.2 and 2.2, and the HWSD v1.1. The fused database was then harmonized into two standardized soil layers as for the HWSD. The top soil layer ranges from 0 to 30 cm and the sub soil layer ranges from 30 to 100 cm. The comparison with the subset of HWSD demonstrates the pronounced difference in the spatial distributions of soil properties and soil organic carbon mass between the UNASM and HWSD, but overall the UNASM provides more accurate and detailed information particularly in Alaska and central Canada. The methods used to develop the UNASM and the comparisons with HWSD are described in detail in Liu et al. (2013).

## **4 Spin-up data package**

A consistent spin-up data package shared among models eliminates any differences in prediction due to spin-up data choices. We created the spin-up data package using the standardized environmental driver data sets described above. MsTMIP requires that all simulations assume

steady-state initial conditions in 1801. The spin-up driver data package contains a 100-yr time series for each required environment driver data product (Table 2) that can be recycled until steady state is reached. For climatology, the 100-yr spin-up time series was created by randomly selecting from the first 30 (1901–1930, global) or 15 (1979–1993, North America) yr of climate driver data on the yearly time step. Using the first 30 or 15yr of climate driver data ensures a smooth transition from the spin-up to transient simulations, while preserving the seasonal cycle of the meteorological variables. A 100-yr period for the spin-up package was chosen to minimize any long-term trend in spin-up climate data, thus minimizing drift in reference simulations, which use constant driver data (Huntzinger et al., 2013). Nitrogen depositions were held constant at 1860 values and atmospheric CO<sub>2</sub> concentrations were held constant at 1801 values to represent near pre-industrial conditions and ensure a smooth transition between spin-up and transient simulations. Similarly, LULCC and phenology data were held constant at 1801 values so that near pre-industrial land cover characteristics and corresponding phenology could be captured in model spin-up. Soil data was assumed to be constant across the whole spin-up period.

All transient simulations defined by MsTMIP require driver data sets covering the period of 1801–2010 (Huntzinger et al., 2013). However, several of the environmental driver data sets, including climate, nitrogen deposition, and soil, do not cover the full period. The spin-up data package was thus recycled to fill these temporal gaps. For global climate data, the spin-up data were used directly to fill the gap between 1801 and 1900. For the NARR climate (North American) data, the full 100-yr time series plus the first 78 yr of the North American spin-up climate data were used to fill the gap between 1801 and 1978. The nitrogen deposition data in 1860 were repeated to fill the gap between 1801–1859 for nitrogen deposition driver data. Finally, constant soil data were used throughout the simulation period of 1801–2010.

## **5 Comparison of model driver data sets used in MsTMIP and other MIPs**

MsTMIP is an evolutionary continuation of the multi-model intercomparison research building on top of findings from past MIPs. MsTMIP uses the “best” existing environmental data available when MsTMIP started that matched its research needs. Some of these environmental data (or their earlier versions) had been used in other projects. For example, an earlier version of

the CRU-NCEP data was used in the GCP-TRENDY project, which focused on investigating the trends in NBP over the period 1980-2009. The ISI-MIP compiled 4 sets of global historical climate data, including one based on NCEP/NCAR Reanalysis 1. It also used HWSD as one of its soil driver data. Both GCP-TRENDY and ISI-MIP (phase 1) focused on global-scale intercomparison. ISI-MIP is unique in its multiple historical climate driver data sets to enable assessment of the influence of the choice of forcing data on the overall results. VEMAP was a pioneer research activity that started in the 1990s. Even though it focused on a regional scale of conterminous United States, it's one of the first to apply a consistent set of input data and boundary conditions to drive different models to minimize differences of model results caused by differences in inputs. MsTMIP was inspired by this idea and prepared its standardized environmental driver data sets. MsTMIP also followed VEMAP to ensure the temporal, spatial, and physical consistency among its environmental driver data sets. MsTMIP is unique in its two scales (global and continental) of driver data sets compiled in parallel that support one of MsTMIP's goals to test influence of both spatial resolution and changing driver data on model estimates. MsTMIP's  $0.25^\circ \times 0.25^\circ$  resolution North American driver data provide the necessary linkage with more well-characterized land-based observational data. By fusing data from multiple sources, MsTMIP provides the unified soil properties data for North American with finest spatial resolution. MsTMIP uses a strategy similar to what's used in GCP-TRENDY and ISI-MIP to prepare the spin-up data package.

## **6 Lessons learned**

Some of the lessons learned in the process of data preparation and distribution for MsTMIP have implications beyond the MsTMIP project. These are described here in order to provide some guidance for future data-intensive activities, especially those that involve assembling consistent data sets for use by multiple modeling teams.

### **1. Study the past**

Scientific discoveries rely heavily on findings from past researches. This is especially true for data-intensive, multi-partner modeling activities like MsTMIP. Since VEMAP in the 1990s till present, there have been a great many MIPs conducted, continuously advancing our

understanding of ecosystem dynamics and modeling techniques. The preparation of environmental driver data sets has been inspired by past MIPs, such as VEMAP, GCP-TRENDY, and NACP interim synthesis, in terms of data selection, processing, and using. This benefited the project significantly by avoiding repeating work and reducing data preparation time.

## 2. Resources for data planning, preparation, and management

Dedicated funding and expertise are needed to develop a plan with the modeling teams and to conduct the driver data compilation. The preparation of standardized model input driver data sets, especially for a project with many different collaborators, takes a significant amount of time and effort. Besides data processing, detailed documentation has to be compiled to capture all the processing steps and trace the origin of each data file. A long-term data management plan is needed to preserve and share the data after a project ends and maximize the value of the data products whenever they are used. Data centers should be identified for long-term data preservation, and the data center's requirements for metadata and documentation should be established at the beginning of the project.

## 3. Proper data formats and standards

Non-proprietary and standard data and metadata formats (e.g. netCDF, Comma-Separated Values (CSV), geotiff, CF metadata convention, or FGDC metadata standard<sup>6</sup>) should be used to maximize the interoperability of the data. Standards make data easier to understand and minimize the ambiguity and potential errors when using a given data product, especially beyond its original intended use.

Standards also help with the long-term preservation and usability of data (Hook et al., 2010). In addition, a data management effort should consider both current and future needs when choosing appropriate data and metadata formats.

## 4. Version control of data files

A controlled repository and versioning system should be used to control data files, not only for final data products to be released to modeling teams and the community, but also for intermediate data to be shared between different processing steps and among project

---

<sup>6</sup> Federal Geographic Data Committee geospatial metadata, <http://www.fgdc.gov/metadata>.

collaborators. When working with a large volume of data files with complicated data processing steps, version control is critical for ensuring that intermediate data files are self-consistent, that the provenance of data is correctly captured, and that final data products are properly distributed to data users.

#### 5. Workflow systems to improve reproducibility and collaboration among team members

Data processing is an error-prone activity. Even if every processing step is performed correctly, the processing algorithms themselves usually need adjustments to create better quality data products. Requirements on final data products sometimes change unexpectedly. In practice, therefore, similar data processing activities will usually be done multiple times before data products are finalized. In MsTMIP, a workflow system (e.g. VisTrails<sup>7</sup> and Kepler<sup>8</sup>) was not used, and as a result significant dedicated time was required to properly capture and adjust the settings and executing environment associated with each processing step. If a workflow system had been used, different data processing steps could have been packaged as individual modules and chained together as workflows, minimizing the time required to trace and reproduce processing steps (Santos et al., 2013). In addition, data reprocessing could have been automated.

#### 6. On-demand approach to distribute data

For a project such as MsTMIP that involves over 20 modeling teams, it is not possible to prepare a single set of data that meets the requirements of all models. TBMs have different native temporal resolutions, for example, and modelers may therefore need to regrid data. Similarly, if the products are used for future applications (outside the projects for which they were created), they may need to be subset to a smaller geographic region, rescaled to a different spatial resolution, or translated to a different geographic projection. On-demand data distribution systems, like the Thematic Realtime Environmental Distributed Data Services<sup>9</sup> (THREDDS) data server and Open Geospatial Consortium (OGC) Web Coverage Services (WCS), can perform spatial and temporal subsetting, as well as resampling, and can therefore help address the diverse needs of different research activities (Wei et al., 2009).

---

<sup>7</sup> VisTrails, <http://www.vistrails.org>.

<sup>8</sup> Kepler, <https://kepler-project.org>.

<sup>9</sup> Thematic Realtime Environmental Distributed Data Services (THREDDS), <http://www.unidata.ucar.edu/projects/THREDDS>.



## 7. “Better is the enemy of good enough”

There is constant pressure to create the “best” data sets possible, but this must be balanced against the overall priority of completing the simulations. If too much time is spent improving the driver data, the time available for model simulations and the evaluation of modeling results is compromised. Therefore, in order to maintain momentum, there comes a time when a decision has to be made to freeze data improvement activities and release a specific version of data products to modeling teams.

## 7 Conclusions

This paper presents the reasoning for, and a description of, driver data and spin-up procedures used in the setup of the global and North American simulations that are part of the MsTMIP activity. These data sets include climate, atmospheric CO<sub>2</sub> concentration, nitrogen deposition, LULCC, C3/C4 grasses fraction, major crop, phenology, soil data, and land-water mask information. In many cases, we found it necessary to develop new or enhanced data sets to serve the needs of long-term, high-resolution TBM simulations. In addition, the need for the data sets to be compatible with over 20 participating TBMs resulted in strict requirements for all data sets considered. These standardized drivers are designed to provide consistent inputs for models participating MsTMIP to minimize the inter-model variability caused by differences in environmental drivers and initial conditions. Thus, these consistent driver inputs, together with the sensitivity simulations defined by MsTMIP, enable better interpretation and quantification of structural and parameter uncertainties of model estimates.

In addition to serving the needs of the MsTMIP activity, this work is intended to serve the needs of researchers wishing to leverage the data products produced by MsTMIP for follow-on studies or related applications. Finally, we offer our experience with MsTMIP as a case study in the development of data sets for collaborative scientific use. The lessons learned from the work reported here, including the need for dedicated support for data development and sharing, for iterative product development, and for the generation of easily accessible and traceable products, among others, are thus broadly applicable. As such, we aim for this work to inform future efforts focused on assembling consistent data sets for use by multiple modeling teams.

All standardized model input driver data sets are archived in the ORNL DAAC to provide long-term data management, preservation, and distribution to the community.

***Acknowledgements.*** Funding for the Multi-scale Synthesis and Terrestrial Model Intercomparison Project (MsTMIP) was provided through NASA ROSES Grant # NNX11AO08A. Data management support for preparing, documenting, and distributing model driver was performed by the Modeling and Synthesis Thematic Data Center (MAST-DC) at Oak Ridge National Laboratory, with funding through NASA ROSES Grant #NNH10AN68I. MsTMIP environmental driver data and simulation output can be obtained from the MsTMIP product archive (<http://nacp.ornl.gov/MsTMIP.shtml>). This is MsTMIP contribution #2. We would like to thank the MsTMIP Modeling teams that participated in discussions of the requirements and characteristics of driver data needed for the simulations, as well as the protocol for running the simulations.

## References

- Adler, R. F., Huffman, G. J., Chang, A., Ferraro, R., Xie, P. P., Janowiak, J., Rudolf, B., Schneider, U., Curtis, S., Bolvin, D., Gruber, A., Susskind, J., Arkin, P., and Nelkin, E.: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979–present), *J. Hydrometeorol.*, 4, 1147–1167, doi:10.1175/1525-7541(2003)004<1147:TVGPCP>2.0.CO;2, 2003.
- Ainsworth, E. A. and Long, S. P.: What have we learned from 15 years of free-air CO<sub>2</sub> enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO<sub>2</sub>, *New Phytol.*, 165, 351–372, doi:10.1111/j.1469-8137.2004.01224.x, 2005.
- Batjes, N. H.: ISRIC-WISE Global Data Set of Derived Soil Properties on a 0.5 by 0.5 Degree Grid (Version 3.0), ISRIC-World Soil Information, Wageningen, 2005.
- Batjes, N. H.: ISRIC-WISE Harmonized Global Soil Profile Dataset (Version 3.1), ISRIC-World Soil Information, Wageningen, 2008.
- Bohn, T. J., Livneh, B., Oylere, W. J., Running, W. S., Nijssena, B., and Lettenmaier, P. D.: Global evaluation of MTCLIM and related algorithms for forcing of ecological and hydrological models, *Agric. For. Met.* 176, 38-49, doi: 10.1016/j.agrformet.2013.03.003, 2013.
- Cao, L., Govindasamy, B., Caldeira, K., Nemani, R., and Ban-Weiss, G.: Importance of carbon dioxide physiological forcing to future climate change, *P. Natl. Acad. Sci. USA*, 107, 9513–9518, doi:10.1073/pnas.0913000107, 2010.
- Conway, T. J. and Tans, P. P.: Development of the CO<sub>2</sub> latitude gradient in recent decades, *Global Biogeochem. Cy.*, 13, 821–826, doi:10.1029/1999GB900045, 1999.
- Cramer, W., D.W. Kicklighter, A. Bondeau, B. III Moore, G. Churkina, B. Nemry, A. Ruimy, A. L. Schloss, and Participants of "Potsdam'95": Comparing global models of terrestrial net primary productivity (NPP): Overview and key results. *Global Change Biology*, 5 (Suppl.1), 1-15., 1999.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Q. J. Roy. Meteor. Soc.*, 137, 553–597, doi:10.1002/qj.828, 2011.
- De Vries, W., Solberg, S., Dobbertin, M., Sterba, H., Laubhann, D., van Oijen, M., Evans, C., Gundersen, P., Kros, J., Wamelink, G. W. W., Reinds, G. J., and Sutton, M. A.: The impacts of

nitrogen deposition on carbon sequestration by European forest and heathlands, *Forest Ecol. Manag.*, 258, 1814–1823, ISSN 0378-1127, doi:10.1016/j.foreco.2009.02.034, 2009.

Dentener, F. J.: Global maps of atmospheric nitrogen deposition, 1860, 1993, and 2050, data set, Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, doi:10.3334/ORNLDAAC/830, available at: <http://daac.ornl.gov/>, last access: 12 December 2011, 2006.

Ehleringer, J. R. and Cerling, T. E.: C3 and C4 photosynthesis, in: *Encyclopedia of Global Environmental Change*, vol. 2, *The Earth System: Biological and Ecological Dimensions of Global Environmental Change*, edited by: Mooney, H. A. and Canadell, J. G., John Wiley & Sons, Ltd, Chichester, ISBN 0-471-97796-9, 186–190, 2002.

Enting, I. G., Rayner, P. J., and Ciais, P.: Carbon Cycle Uncertainty in REgional Carbon Cycle Assessment and Processes (RECCAP), *Biogeosciences*, 9, 2889–2904, doi:10.5194/bg-9- 2889-2012, 2012.

FAO: The FAO-UNESCO Soil Map of the World, Legend and 9 Volumes, UNESCO, Paris, 1971–1981.

FAO: The Digitized Soil Map of the World Including Derived Soil Properties (version 3.5), FAO Land and Water Digital Media Series #1, FAO, Rome, 1995, 2003.

FAO/IIASA/ISRIC/ISS-CAS/JRC: Harmonized World Soil Database (version 1.1), FAO, Rome, Italy and IIASA, Laxenburg, Austria, 2011.

Foley, J. A.: Numerical models of the terrestrial biosphere, *J. Biogeogr.*, 22, 837–842, doi:10.2307/2845984, 1995.

Friedl, M. A., McIver, D. K., Hodges, J. C. F., Zhang, X. Y., Muchoney, D., Strahler, A. H., Woodcock, C. E., Gopal, S., Schneider, A., Cooper, A., Baccini, A., Gao, F., and Schaaf, C.: Global land cover mapping from MODIS: algorithms and early results, *Remote Sens. Environ.*, 83, 287–302, doi:10.1016/S0034-4257(02)00078-0, 2002.

Friedlingstein, P., Houghton, R. A., Marland, G., Hackler, J., Boden, T. A., Conway, T. J., Canadell, J. G., Raupach, M. R., Ciais, P., and Le Quéré, C.: Update on CO<sub>2</sub> emissions, *Nat. Geosci.*, 3, 811–812, doi:10.1038/ngeo1022, 2010.

Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C., Green, P., Holland, E., Karl, D. M., Michaels, A. F., Porter, J. H., Townsend, A., and Vörösmarty, C.: Nitrogen cycles: past, present, and future, *Biogeochemistry*, 70, 153–226, doi:10.1007/s10533-004-0370-0, 2004.

Giglio, L., Randerson, J. T., van der Werf, G. R., Kasibhatla, P. S., Collatz, G. J., Morton, D. C., and DeFries, R. S.: Assessing variability and long-term trends in burned area by merging multiple satellite fire products, *Biogeosciences*, 7, 1171–1186, doi:10.5194/bg-7-1171-2010, 2010.

Global Land Cover 2000 Database, European Commission, Joint Research Centre, available at: <http://bioval.jrc.ec.europa.eu/products/glc2000/glc2000.php>, last access: 28 August 2011, 2003.

GLOBALVIEW-CO<sub>2</sub>: Cooperative Atmospheric Data Integration Project – Carbon Dioxide, NOAA ESRL, Boulder, Colorado, available at: <http://www.esrl.noaa.gov/gmd/ccgg/globalview/> (last access: 21 April 2011), 2011.

Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D., Bousquet, P., Bruhwiler, L., Chen, Y.-H., Ciais, P., Fan, S., Fung, I. Y., Gloor, M., Heimann, M., Higuchi, K., John, J., Maki, T., Maksyutov, S., Masarie, K., Peylin, P., Prather, M., Pak, B. C., Randerson, J., Sarmiento, J., Taguchi, S., Takahashi, T., and Yuen, C.-W.: Towards robust regional estimates of CO<sub>2</sub> sources and sinks using atmospheric transport models, *Nature*, 415, 626–630, doi:10.1038/415626a, 2002.

Hansen, M. C., DeFries, R. S., Townshend, J. R. G., and Sohlberg, R.: Global land cover classification at 1 km spatial resolution using a classification tree approach, *Int. J. Remote Sens.*, 21, 1331–1364, doi:10.1080/014311600210209, 2000.

Holland, E. A., Braswell, B. H., Lamarque, J. F., Townsend, A., Sulzman, J., Müller, J. F., Dentener, F., Brasseur, G., Levy, H., Penner, J. E., and Roelofs, G.-J.: Variations in the predicted spatial distribution of atmospheric nitrogen deposition and their impact on carbon uptake by terrestrial ecosystems, *J. Geophys. Res.*, 102, 15849–15866, doi:10.1029/96JD03164, 1997.

Holland, E. A., Braswell, B. H., Sulzman, J., and Lamarque, J. F.: Nitrogen deposition onto the United States and Western Europe: synthesis of observation and models, *Ecol. Appl.*, 15, 38–57, doi:10.1890/03-5162, 2005.

Hook, L. A., Santhana-Vannen, S., Beaty, T. W., Cook, R. B., and Wilson, B. E.: Best Practices for Preparing Environmental Data Sets to Share and Archive, Oak Ridge National Laboratory Distributed Active Archive Center, available at: <http://daac.ornl.gov/PI/BestPractices-2010.pdf>, last access: 21 May 2012, 2010.

Hungate, B. A., Dukes, J. S., Shaw, M. R., Luo, Y. Q., and Field, C. B.: Nitrogen and climate change, *Science*, 302, 1512–1513, doi:10.1126/science.1091390, 2003.

Huntzinger, D. N., Post, W. M., Wei, Y., Michalak, A. M., West, T. O., Jacobson, A. R., Baker, I. T., Chen, J. M., Davis, K. J., Hayes, D. J., Hoffman, F. M., Jain, A. K., Liu, S., McGuire, A. D., Neilson, R. P., Potter, C., Poulter, B., Price, D., Raczka, B. M., Tian, H. Q., Thornton, P., Tomelleri, E., Viovy, N., Xiao, J., Yuan, W., Zeng, N., Zhao, M., and Cook, R.: North American

Carbon Program (NACP) regional interim synthesis: terrestrial biospheric model intercomparison, *Ecol. Model.*, 232, 144–157, doi:10.1016/j.ecolmodel.2012.02.004, 2012.

Huntzinger, D. N., Schwalm, C., Michalak, A. M., Schaefer, K., King, A. W., Wei, Y., Jacobson, A., Liu, S., Cook, R. B., Post, W. M., Berthier, G., Hayes, D., Huang, M., Ito, A., Lei, H., Lu, C., Mao, J., Peng, C. H., Peng, S., Poulter, B., Riccuto, D., Shi, X., Tian, H., Wang, W., Zeng, N., Zhao, F., and Zhu, Q.: The North American Carbon Program Multi-scale synthesis and Terrestrial Model Intercomparison Project – Part 1: Overview and experimental design, *Geosci. Model Dev. Discuss.*, 6, 3977–4008, doi:10.5194/gmdd-6-3977-2013, 2013.

Hurtt, G. C., Chini, L., Frohling, S., Betts, R., Edmonds, J., Feddema, J., Fisher, G., Goldewijk, K. K., Hibbard, K., Houghton, R., Janetos, A., Jones, C., Kinderman, G., Konoshita, T., Riahi, K., Shevliakova, E., Smith, S. J., Stefest, E., Thomson, A. M., Thornton, P., van Vuuren, D., and Wang, Y.: Harmonization of land-use scenarios for the period 1500–2100: 600years of global gridded annual land-use transitions, wood harvest, and resulting secondary lands, *Climatic Change*, 109, 117–161, doi:10.1007/s10584-011-0153-2, 2011.

Jain, A., Yang, X., Kheshgi, H., McGuire, A. D., Post, W., and Kicklighter, D.: Nitrogen attenuation of terrestrial carbon cycle response to global environmental factors, *Global Biogeochem. Cy.*, 23, GB4028, doi:10.1029/2009GB003519, 2009.

Jung, M., Henkel, K., Herold, M., and Churkina, G.: Exploiting synergies of global land cover products for carbon cycle modeling, *Remote Sens. Environ.*, 101, 534–553, doi:10.1016/j.rse.2006.01.020, 2006.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., and Reynolds, R.: The NCEP/NCAR 40-year reanalysis project, *B. Am. Meteorol. Soc.*, 77, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996.

Kennedy, A., Dong, X., Xi, B., Xie, S., Zhang, Y., and Chen, J.: A comparison of MERRA and NARR reanalysis with the DOE ARM SGP continuous forcing data, Abstract #A53E-0296, AGU fall meeting 2010, San Francisco, California, USA, 13–17 December 2010.

Kicklighter, D. W., Bruno, M., Donges, S., Esser, G., Heimann, M., Helfrich, J., Ift, F., Joos, F., Kaduk, J., Kohlmaier, G. H., McGuire, A. D., Melillo, J. M., Meyer, R., Moore III, B., Nadler, A., Prentice, I. C., Sauf, W., Schloss, A. L., Sitch, S., Wittenberg, U., and Wurth, G.: A first-order analysis of the potential role of CO<sub>2</sub> fertilization to affect the global carbon budget: a comparison of four terrestrial biosphere models, *Tellus B*, 51, 343–366, doi:10.1034/j.1600-0889.1999.00017.x, 1999.

Klein Goldewijk, K., Beusen, A., van Drecht, G., and de Vos, M.: The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years, *Global Ecol. Biogeogr.*, 20, 73–86, doi:10.1111/j.1466-8238.2010.00587.x, 2011.

Le Quéré, C., Andres, R. J., Boden, T., Conway, T., Houghton, R. A., House, J. I., Marland, G., Peters, G. P., van der Werf, G. R., Ahlström, A., Andrew, R. M., Bopp, L., Canadell, J. G., Ciais, P., Doney, S. C., Enright, C., Friedlingstein, P., Huntingford, C., Jain, A. K., Jourdain, C., Kato, E., Keeling, R. F., Klein Goldewijk, K., Levis, S., Levy, P., Lomas, M., Poulter, B., Raupach, M. R., Schwinger, J., Sitch, S., Stocker, B. D., Viovy, N., Zaehle, S., and Zeng, N.: The global carbon budget 1959–2011, *Earth Syst. Sci. Data*, 5, 165–185, doi:10.5194/essd-5-165-2013, 2013.

Liu, S., Wei, Y., Post, W. M., Cook, R. B., Schaefer, K., and Thornton, M. M.: The Unified North American Soil Map and its implication on the soil organic carbon stock in North America, *Biogeosciences*, 10, 2915–2930, doi:10.5194/bg-10-2915-2013, 2013.

Loveland, T. R., Reed, B. C., Brown, J. F., Ohlen, D. O., Zhu, Z., Yang, L., and Merchant, J. W.: Development of a global land cover characteristics database and IGBP DISCover from 1 km AVHRR data, *Int. J. Remote Sens.*, 21, 1303–1330, doi:10.1080/014311600210191, 2000.

Lu, C., Tian, H., Liu, M., Ren, W., Xu, X., Chen, G., and Zhang, C.: Effects of nitrogen deposition on China's terrestrial carbon uptake in the context of multiple environmental changes, *Ecol. Appl.*, 22, 53–75, doi:10.1890/10-1685.1, 2012.

Lu, C. and Tian, H.: Spatial and temporal patterns of nitrogen deposition in China: synthesis of observational data, *J. Geophys. Res.*, 112, D22S05, doi:10.1029/2006JD 007990, 2007.

MacFarling Meure, C., Etheridge, D., Trudinger, C., Steele, P., Langenfelds, R., van Ommen, T., Smith, A., and Elkins, J.: Law Dome CO<sub>2</sub>, CH<sub>4</sub> and N<sub>2</sub>O ice core records extended to 2000 years BP, *Geophys. Res. Lett.*, 33, L14810, doi:10.1029/2006GL026152, 2006.

Marland, G., Boden, T. A., and Andres, R. J.: Global, regional, and national fossil fuel CO<sub>2</sub> emissions, in: *Trends: a Compendium of Data on Global Change*, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tenn., USA, available at: <http://cdiac.ornl.gov/trends/emis/overview>, last access: 21 April 2011, 2008.

McGuire, A. D., Sitch, S., Clein, J. S., Dargaville, R., Esser, G., Foley, J., Heimann, M., Joos, F., Kaplan, J., Kicklighter, D. W., Meier, R. A., Melillo, J. M., Moore III, B., Prentice, I. C., Ramankutty, N., Reichenau, T., Schloss, A., Tian, H., Williams, L. J., and Wittenbe, U.: Carbon balance of the terrestrial biosphere in the twentieth century: analyses of CO<sub>2</sub>, climate and land use effects with four process-based ecosystem models, *Global Biogeochem. Cy.*, 15, 183–206, doi:10.1029/2000GB001298, 2001.

- Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., Jović, D., Woollen, J., Rogers, E., Berbery, E. H., Ek, M. B., Fan, Y., Grumbine, R., Higgins, W., Li, H., Lin, Y., Manikin, G., Parrish, D., and Shi, W.: North American regional reanalysis, *B. Am. Meteorol. Soc.*, 87, 343–360, doi:10.1175/BAMS-87-3-343, 2006.
- Mitchell, T. D. and Jones, P. D.: An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int. J. Climatol.*, 25, 693–712, doi:10.1002/joc.1181, 2005
- Monfreda, C., Ramankutty, N., and Foley, J. A.: Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000, *Global Biogeochem. Cy.*, 22, GB1022, doi:10.1029/2007GB002947, 2008.
- Norby, R. J., DeLucia, E. H., Gielen, B., Calfapietra, C., Giardina, C. P., King, J. S., Ledford, J., McCarthy, H. R., Moore, D. J. P., Ceulemans, R., De Angelis, P., Finzi, A. C., Karnosky, D. F., Kubiske, M. E., Lukac, M., Pregitzer, K. S., Scarascia-Mugnozza, G. E., Schlesinger, W. H., and Oren, R.: Forest response to elevated CO<sub>2</sub> is conserved across a broad range of productivity, *P. Natl. Acad. Sci. USA*, 102, 18052–18056, doi:10.1073/pnas.0509478102, 2005.
- Pacala, S. W., Hurtt, G. C., Baker, D., Peylin, P., Houghton, R. A., Birdsey, R. A., Heath, L., Sundquist, E. T., Stallard, R. F., Ciais, P., Moorcroft, P., Caspersen, J. P., Shevliakova, E., Moore, B., Kohlmaier, G., Holland, E., Gloor, M., Harmon, M. E., Fan, S.-M., Sarmiento, J. L., Goodale, C. L., Schimel, D., and Field, C. B.: Consistent land- and atmosphere-based US carbon sink estimates, *Science*, 292, 2316–2320, doi:10.1126/science.1057320, 2001.
- Pielke Sr., R. A., Pitman, A., Niyogi, D., Mahmood, R., McAlpine, C., Hossain, F., Goldewijk, K., Nair, U., Betts, R., Fall, S., Reichstein, M., Kabat, P., and de Noblet-Ducoudré, N.: Land use/land cover changes and climate: modeling analysis and observational evidence, *WIREs Clim. Change*, 2, 828–850, doi:10.1002/wcc.144, 2011.
- Potter, C., Klooster, S., Hiatt, S., Fladeland, M., Genovese, V., and Gross, P.: Satellite-derived estimates of potential carbon sequestration through afforestation of agricultural lands in the United States, *Climatic Change*, 80, 323–336, doi:10.1007/s10584-006-9109-3, 2007.
- Pregitzer, K. S., Burton, A. J., Zak, D. R., and Talhelm, A. F.: Simulated chronic nitrogen deposition increases carbon storage in northern temperate forests, *Glob. Change Biol.*, 14, 142–153, 2008.
- Ramankutty, N., Gibbs, H. K., Achard, F., DeFries, R., Foley, J., and Houghton, R. A.: Challenges to estimating carbon emissions from tropical deforestation, *Glob. Change Biol.*, 13, 51–66, doi:10.1111/j.1365-2486.2006.01272.x, 2007.
- Randerson, J. T., Hoffman, F. M., Thornton, P. E., Mahowald, N. M., Lindsay, K., Lee, Y.-H., Nevison, C. D., Doney, S. C., Bonan, G., Stöckli, R., Covey, C., Running, S. W., and Fung, I.



- Y.: Systematic assessment of terrestrial biogeochemistry in coupled climate–carbon models, *Glob. Change Biol.*, 15, 2462–2484, 2009.
- Reay, D. S., Dentener, F., Smith, P., Grace, J., and Feely, R. A.: Global nitrogen deposition and carbon sinks, *Nat. Geosci.*, 1, 430–437, 2008.
- Running, S. W., Nemani, R. R., and Hungerford, R. D.: Extrapolation of synoptic meteorological data in mountainous terrain and its use for simulating forest evaporation and photosynthesis, *Can. J. Forest Res.*, 17, 472–483, doi:10.1139/x87-081, 1987.
- Santos, E., Poco, J., Wei, Y., Liu, S., Cook, B., Williams, D. N., and Silva, C. T.: UV-CDAT: analyzing climate datasets from a user’s perspective, *Comput. Sci. Eng.*, 15, 94–103, doi:10.1109/MCSE.2013.15, 2013.
- Saxton, K. E., Rawls, W. J., Romberger, J. S., and Papendick, R. I.: Estimating generalized soil-water characteristics from texture, *Soil Sci. Soc. Am. J.*, 50, 1031–1036, doi:10.2136/sssaj1986.03615995005000040039x, 1986.
- Schaefer, K., Denning, A. S., Suits, N., Kaduk, J., Baker, I., Los, S., and Prihodko, L.: Effect of climate on interannual variability of terrestrial CO<sub>2</sub> fluxes, *Global Biogeochem. Cy.*, 16, 49-1–49-12, doi:10.1029/2002GB001928, 2002.
- Schaefer, K., Schwalm, C. R., Williams, C., et al.: A model-data comparison of gross primary productivity: results from the North American Carbon Program site synthesis, *J. Geophys. Res.*, 17, G03010, doi:10.1029/2012JG001960, 2012.
- Schimel, D.S., VEMAP Participants, and B.H. Braswell.: Continental scale variability in ecosystem processes: Models, data, and the role of disturbance. *Ecological Monographs* 67: 251-271. 1997.
- Schwalm, C. R., Huntinzger, D. N., Michalak, A. M., Fisher, J. B., Kimball, J. S., Mueller, B., Zhang, K., and Zhang, Y.: Sensitivity of inferred climate model skill to evaluation decisions: a case study using CMIP5 evapotranspiration, *Environ. Res. Lett.*, 8, 024028, doi:10.1088/1748-9326/8/2/024028, 2013.
- Sellers, P. J., Mintz, Y., Sud, Y. C., and Dalcher, A.: A simple biosphere model (SiB) for use within general circulation models, *J. Atmos. Sci.*, 43, 505–531, doi:10.1175/1520-0469(1986)043<0505:ASBMFU>2.0.CO;2, 1986.
- Sellers, P. J., Bounoua, L., Collatz, G. J., Randall, D. A., Dazlich, D. A., Los, S. O., Berry, J. A., Fung, I., Tucker, C. J., Field, C. B., and Jensen, T. G.: Comparison of radiative and physiological effects of doubled atmospheric CO<sub>2</sub> on climate, *Science*, 271, 1402–1406, 1996a.

Sellers, P. J., Los, S. O., Tucker, C. J., Justice, C. O., Dazlich, D. A., Collatz, G. J., and Randall, D. A.: A revised land surface parameterization (SiB2) for atmospheric GCMs, Part II: The generation of global fields of terrestrial biophysical parameters from satellite data, *J. Climate*, 9, 706–737, doi:10.1175/1520-0442(1996)009<0706:ARLSPF>2.0.CO;2, 1996b.

Shi, X., Mao, J., Thornton, P. E., Hoffman, F. M., and Post, W. M.: The impact of climate, CO<sub>2</sub>, nitrogen deposition and land use change on simulated contemporary global river flow, *Geophys. Res. Lett.*, 38, L08704, doi:10.1029/2011GL046773, 2011.

Sohl, T. L., Sleeter, B. M., Zhu, Z., Saylor, K. L., Bennett, S., Bouchard, M., Reker, R., Hawbaker, T., Wein, A., Liu, S., Kanengieter, R., and Acevedo, W.: A land-use and land-cover modeling strategy to support a national assessment of carbon stocks and fluxes, *Appl. Geogr.*, 34, 111–124, doi:10.1016/j.apgeog.2011.10.019, 2012.

Still, C. J., Berry, J. A., Collatz, G. J., and DeFries, R. S.: Global distribution of C<sub>3</sub> and C<sub>4</sub> vegetation: carbon cycle implications, *Global Biogeochem. Cy.*, 17, 1–14, doi:10.1029/2001GB001807, 2003.

Sun, X. and Barros, A. P.: An evaluation of the statistics of rainfall extremes in rain gauge observations, and satellite-based and reanalysis products using universal multifractals, *J. Hydrometeorol.*, 11, 388–404, doi:10.1175/2009JHM1142.1, 2010.

Thornton, P. E. and Running, S. W.: An improved algorithm for estimating incident daily solar radiation from measurements of temperature, humidity, and precipitation, *Agr. Forest Meteorol.*, 93, 211–228, doi:10.1016/S0168-1923(98)00126-9, 1999.

Thornton, P. E., Thornton, M. M., Mayer, B. W., Wilhelmi, N., Wei, Y., and Cook, R. B.: Daymet: Daily surface weather on a 1 km grid for North America, 1980–2012, acquired online (<http://daymet.ornl.gov/>) on 30 November 2012 from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA, doi:10.3334/ORNLDAAAC/Daymet\_V2, 2012.

Tian, H., Xu, X., Liu, M., Ren, W., Zhang, C., Chen, G., and Lu, C.: Spatial and temporal patterns of CH<sub>4</sub> and N<sub>2</sub>O fluxes in terrestrial ecosystems of North America during 1979–2008: application of a global biogeochemistry model, *Biogeosciences*, 7, 2673–2694, doi:10.5194/bg-7-2673-2010, 2010.

Tucker, C. J., Pinzon, J. E., Brown, M. E., Slayback, D. A., Pak, E. W., Mahoney, R., Vermote, E. F., and El Saleous, N.: An extended AVHRR 8-km NDVI dataset compatible with MODIS and SPOT vegetation NDVI data, *Int. J. Remote Sens.*, 26, 4485–4498, doi:10.1080/01431160500168686, 2005.

Uppala, S. M., Kållberg, P. W., Simmons, A. J., Andrae, U., da Costa Bechtold, V., Fiorino, M., Gibson, J. K., Haseler, J., Hernandez, A., Kelly, G. A., Li, X., Onogi, K., Saarinen, S., Sokka,

N., Allan, R. P., Andersson, E., Arpe, K., Balmaseda, M. A., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Caires, S., Chevallier, F., Dethof, A., Dragosavac, M., Fisher, M., Fuentes, M., Hagemann, S., Hólm, E., Hoskins, B. J., Isaksen, L., Janssen, P. A. E. M., Jenne, R., McNally, A. P., Mahfouf, J.-F., Morcrette, J.-J., Rayner, N. A., Saunders, R. W., Simon, P., Sterl, A., Trenberth, K. E., Untch, A., Vasiljevic, D., Viterbo, P., and Woollen, J.: The ERA-40 re-analysis, *Q., J. Roy. Meteor. Soc.*, 131, 2961–3012, doi:10.1256/qj.04.176, 2005.

United States Carbon Cycle Science Program (US CCSP): A U. S. Carbon Cycle Science Plan, August 2011.

van Aardenne, J. A., Dentener, F. J., Olivier, J. G. J., Klein Goldewijk, C. G. M., and Lelieveld, J.: A  $1^\circ \times 1^\circ$  resolution dataset of historical anthropogenic trace gas emissions for the period 1890–1990, *Global Biogeochem. Cy.*, 15, 909–928, doi:10.1029/2000GB001265, 2001.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Matsui, T., Hurtt, G., Lamarque, J.-F., Meinshausen, M., Smith, S., Grainer, C., Rose, S., Hibbard, K. A., Nakicenovic, N., Krey, V., and Kram, T.: Representative concentration pathways: An overview, *Climatic Change*, 109, 5–31, doi:10.1007/s10584-011-0148-z, 2011.

Wei, Y., Santhana-Vannan, S. K., and Cook, R. B.: Discover, visualize, and deliver geospatial data through OGC standards-based WebGIS system, in: 2009 17th International Conference on Geoinformatics, IEEE, 12-14 August 2009, Fairfax, VA, USA, 1–6, doi:10.1109/GEOINFORMATICS.2009.5293520, 2009.

Xie, P., Janowiak, J. E., Arkin, P. A., Adler, R., Gruber, A., Ferraro, R., Huffman, G. J., and Curtis, S.: GPCP pentad precipitation analyses: an experimental dataset based on gauge observations and satellite estimates, *J. Climate*, 16, 2197–2214, doi:10.1175/2769.1, 2003.

Zaehle, S., Friend, A. D., Friedlingstein, P., Dentener, F., Peylin, P., and Schulz, M.: Carbon and nitrogen cycle dynamics in the O-CN land surface model: 2. Role of the nitrogen cycle in the historical terrestrial carbon balance, *Global Biogeochem. Cy.*, 24, GB1006, doi:10.1029/2009GB003522, 2010.

Zhao, M., Running, S. W., and Nemani, R. R.: Sensitivity of Moderate Resolution Imaging Spectroradiometer (MODIS) terrestrial primary production to the accuracy of meteorological reanalyses, *J. Geophys. Res.*, 111, G01002, doi:10.1029/2004JG000004, 2006.

**Table 1. The MsTMIP environmental driver data summary**

Category	Name	Spatial Extent & Resolution	Native Temporal Period, Resolution	Extended Temporal Period, Resolution <sup>a</sup>	Variables
<b>Climate</b>	CRU-NCEP <sup>b</sup>	Global (0.5°)	1901-2010, 6-hourly	1801-2010, 6-hourly	<ul style="list-style-type: none"> <li>- precipitation</li> <li>- air temperature</li> <li>- air specific humidity</li> <li>- air relative humidity (NA only)</li> <li>- pressure</li> <li>- downward longwave radiation</li> <li>- downward shortwave radiation</li> <li>- wind speed</li> </ul>
	NARR	NA (0.25°)	1979-2010, 3-hourly	1801-2010, 3-hourly	
<b>Land Water Mask</b>	CRU-NCEP	Global (0.5°)			binary land vs. water map
	NARR	NA (0.25°)	constant	constant	
<b>CO<sub>2</sub></b>	Extended GLOBALVIEW-CO <sub>2</sub>	Global (0.5°), NA (0.25°)	1801-2010, monthly	1801-2010, monthly	atmospheric CO <sub>2</sub> concentration
<b>Nitrogen Deposition</b>	Enhanced Dentener	Global (0.5°), NA (0.25°)	1860-2010, annual	1801-2010, annual	<ul style="list-style-type: none"> <li>NH<sub>x</sub>-N deposition</li> <li>NO<sub>y</sub>-N deposition</li> </ul>
<b>Land Cover Change</b>	SYNMAP+ Hurtt	Global (0.5°), NA (0.25°)	1801-2010, annual	1801-2010, annual	land cover state maps
<b>C3/C4 Grass</b>	C3/C4 grass fraction	Global (0.5°), NA (0.25°)	constant	constant	relative fractions of C3/C4 grasses
<b>Major Crops</b>	Monfreda et al. 2008	Global (0.5°), NA (0.25°)	constant	constant	fraction of harvest area in each grid cell for maize, rice, soybean, and wheat
<b>Phenology</b>	GIMMSg	Global (0.5°), NA (0.25°)	1801-2010, monthly	1801-2010, monthly	NDVI, LAI, and fPAR
	HWSD v1.1	Global (0.5°)	constant	constant	<ul style="list-style-type: none"> <li>- soil layers</li> <li>- dominant soil type</li> <li>- reference soil depth</li> <li>- clay/sand/silt fractions</li> <li>- pH</li> <li>- organic carbon</li> <li>- cation exchange capacity</li> <li>- reference bulk density</li> <li>- gravel content</li> </ul>
<b>Soil</b>	STATSGO2 (US)	NA (0.25°)	constant	constant	<ul style="list-style-type: none"> <li>- organic carbon</li> <li>- cation exchange capacity</li> <li>- reference bulk density</li> <li>- gravel content</li> </ul>
	SLC 3.2&2.2 (CA) HWSD 1.1 (Other)				

<sup>a</sup> Native temporal periods of environmental driver data sets compiled for MsTMIP are extended to be compatible with the simulation time period (1801-2010) defined by MsTMIP. Please refer to section 4. Spin-up Data Package to see how data with shorter native temporal period are extended back to 1801 to address the needs of MsTMIP simulations.

<sup>b</sup> **CRU-NCEP**: Climate Research Unit, National Centers for Environmental Prediction; **NARR**: North American Regional Reanalysis; **SYNMAP**: SYNergetic land cover MAP; **GIMMSg**: Global Inventory Monitoring and Modeling System version g; **NDVI**: Normalized Difference Vegetation Index; **LAI**: Leaf Area Index; **fPAR**: fraction of Photosynthetically Active Radiation; **HWSD**: Harmonized World Soil Database; **STATSGO2**: State Soil Geographic data version 2; **SLC**: Soil Landscapes of Canada

**Table 2. The MsTMIP spin-up environmental driver data summary**

<b>Category</b>	<b>Global</b>	<b>Regional (North American)</b>
<b>Climate</b>	A 100-year time series with no significant trend by randomizing CRU-NCEP in 1901-1930 (30 years)	A 100-year time series with no significant trend by randomizing NARR in 1979-1993 (15 years)
<b>Atmospheric CO<sub>2</sub> concentration</b>	A 100-year time series by repeating atmospheric CO <sub>2</sub> concentration driver data in 1801	
<b>Nitrogen deposition</b>	A 100-year time series by repeating nitrogen deposition driver data in 1860	
<b>Land cover and land cover change</b>	A 100-year time series by repeating harmonized Hurtt-SYNMAP land cover change driver data in 1801	
<b>Phenology</b>	A 100-year time series by repeating phenology driver data in 1801	
<b>Soil</b>	Constant gridded HWSD	Constant UNASM
<b>Land-water mask</b>	Constant global land-water mask	Constant North American land-water mask

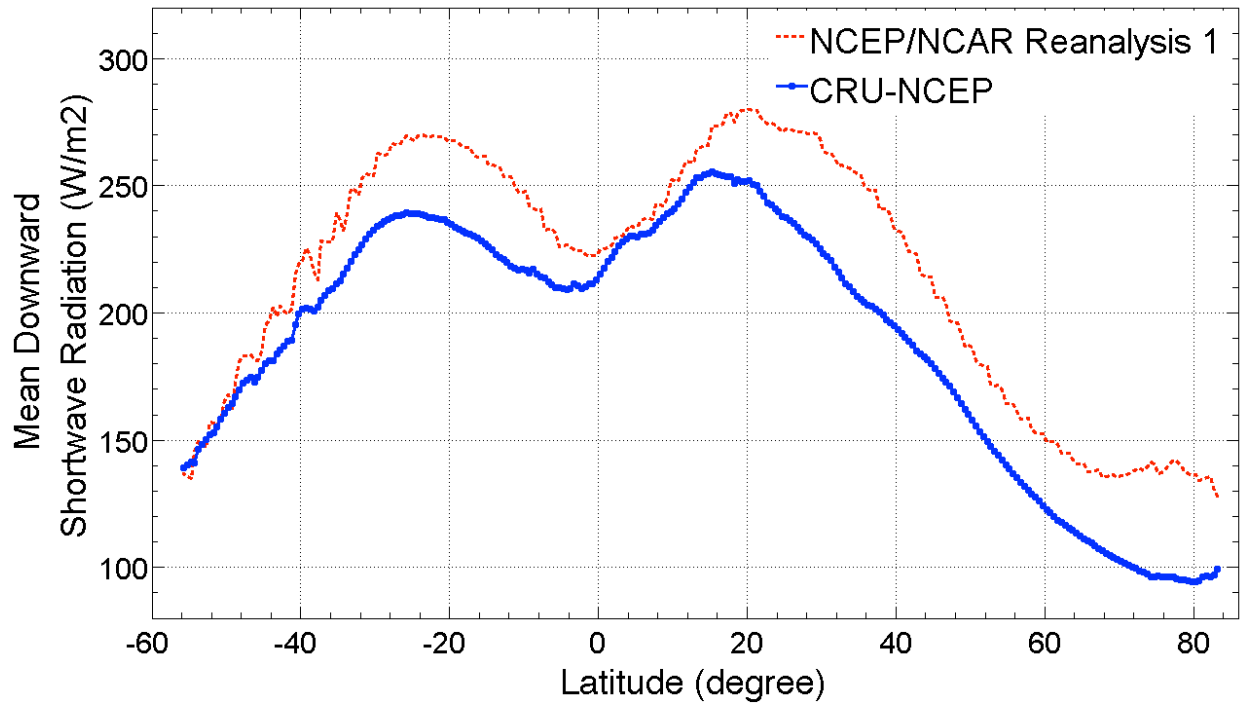


Figure 1. Comparison of the mean of long-term mean downward shortwave radiation (1948-2010) on land surface for each 0.5-degree latitudinal band from NCEP/NCAR Reanalysis 1 and CRU-NCEP data sets

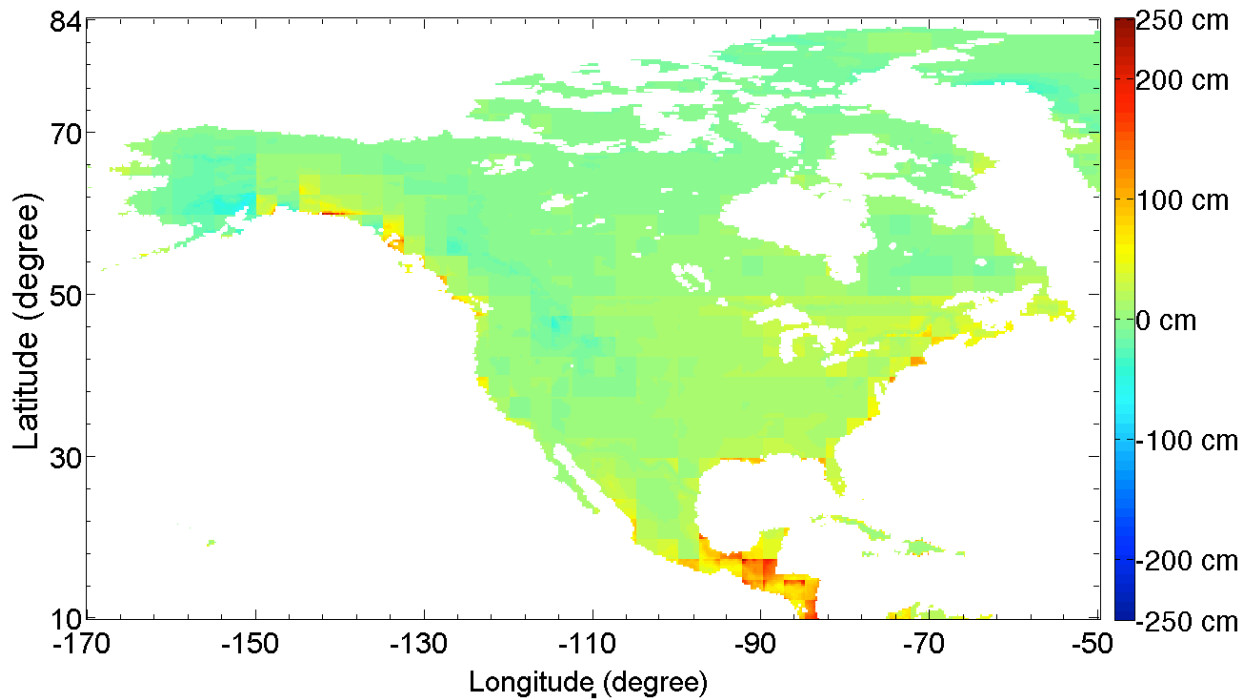


Figure 2. Difference Map between the Long-term Mean (1979-2010) Annual Total Precipitation from Rescaled NARR and Original NARR (*Rescaled NARR Precipitation – Original NARR Precipitation*)

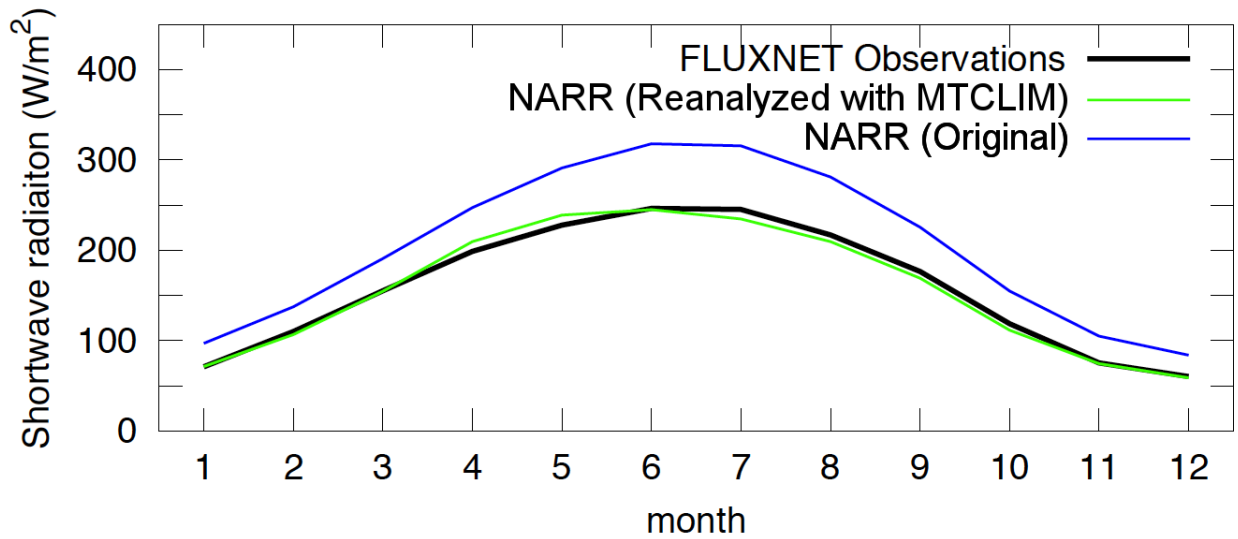


Figure 3. Comparison of shortwave radiation from original and reanalyzed NARR against observations averaged over 23 FLUXNET sites across North America

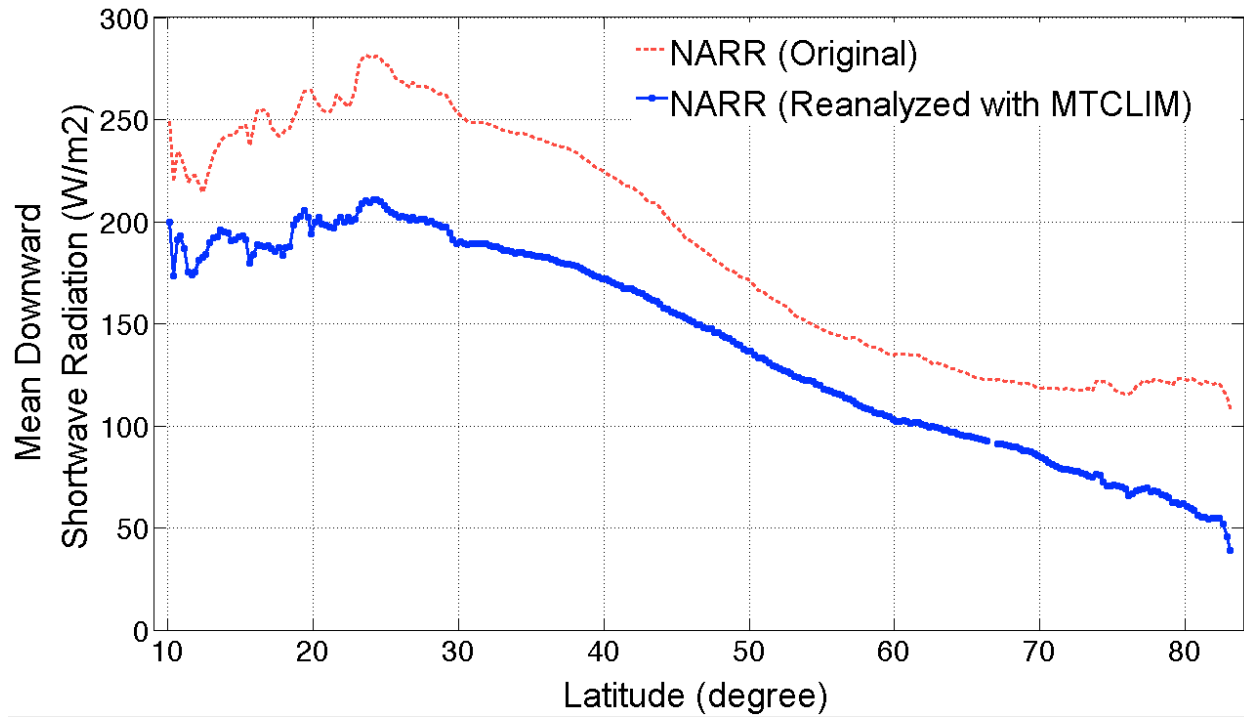


Figure 4. Comparison of the zonal mean latitudinal zonal ( $0.25^\circ$ ) mean of long-term mean downward shortwave radiation (1979-2010) on land surface from original NARR and reanalyzed NARR data sets



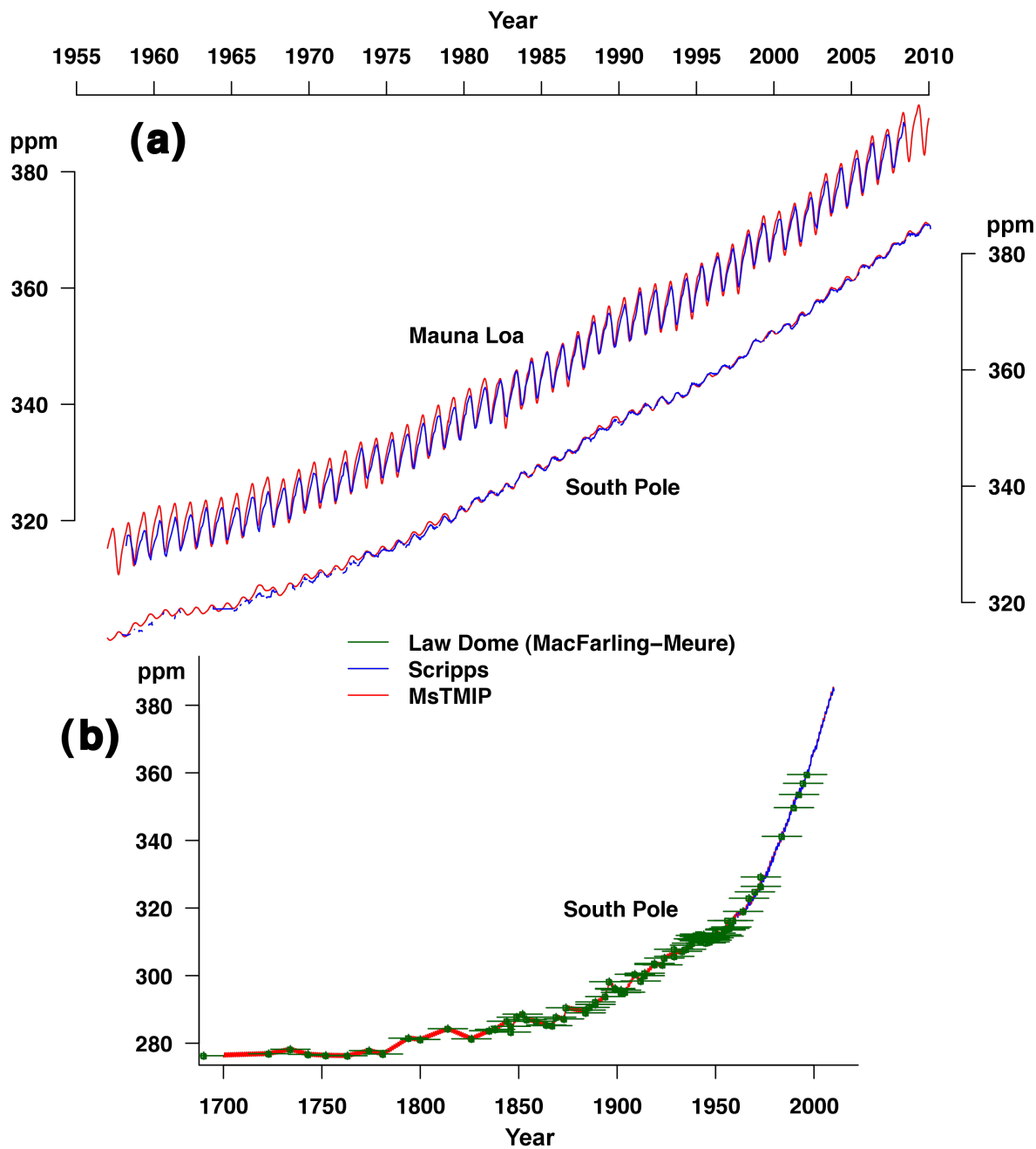


Figure 5. Comparison of MsTMIP driver data atmospheric CO<sub>2</sub> with independent data. (a) Comparison of MsTMIP and Scripps CO<sub>2</sub> program data at Mauna Loa and South Pole from 1958-2010, and (b) comparison with Law Dome ice core records of atmospheric CO<sub>2</sub> (MacFarling Meure et al., 2006).

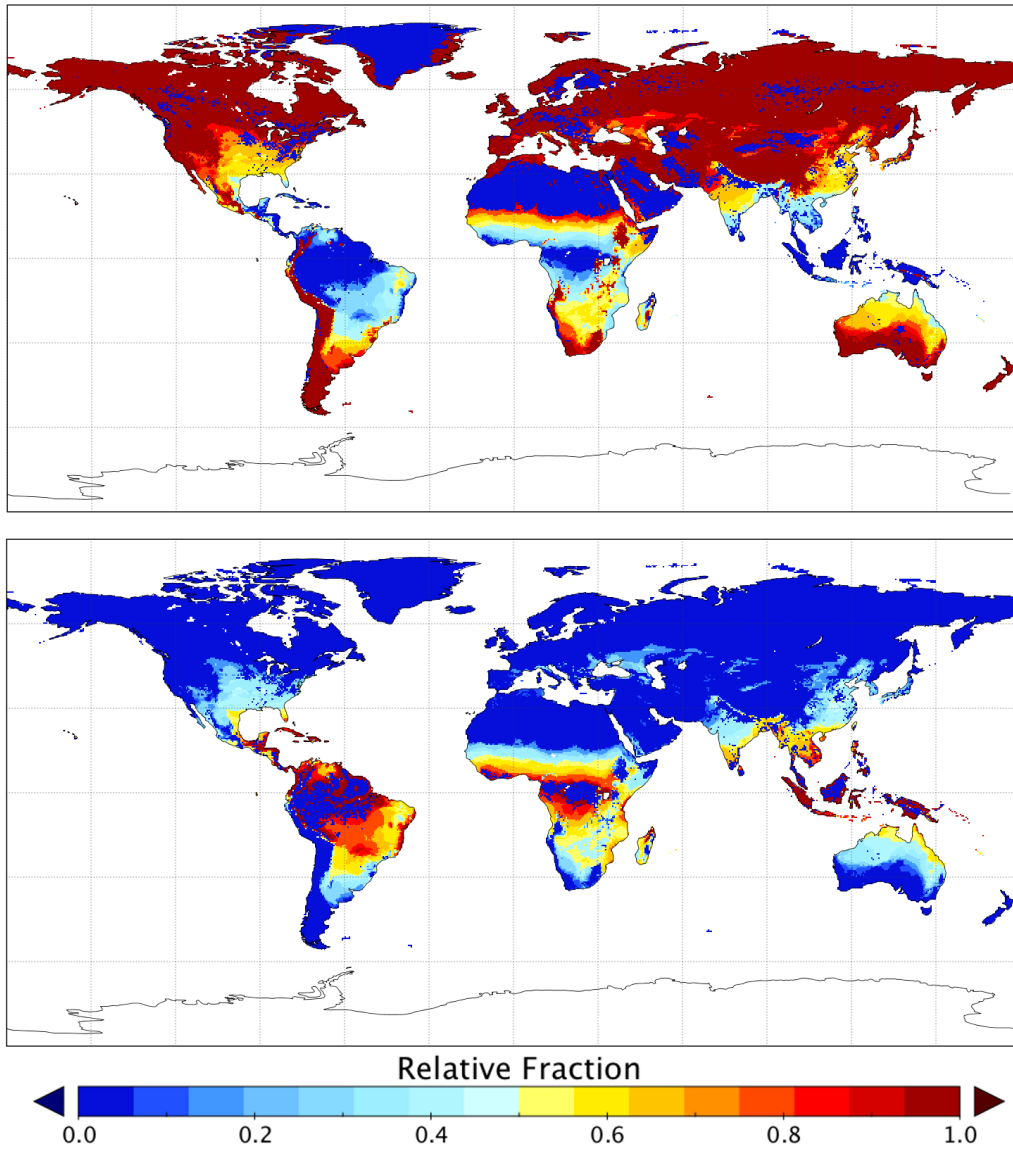


Figure 6. Relative fractions of C3 (top) and C4 (bottom) grassland on global 0.5° scale under “present” climate (2000-2010)

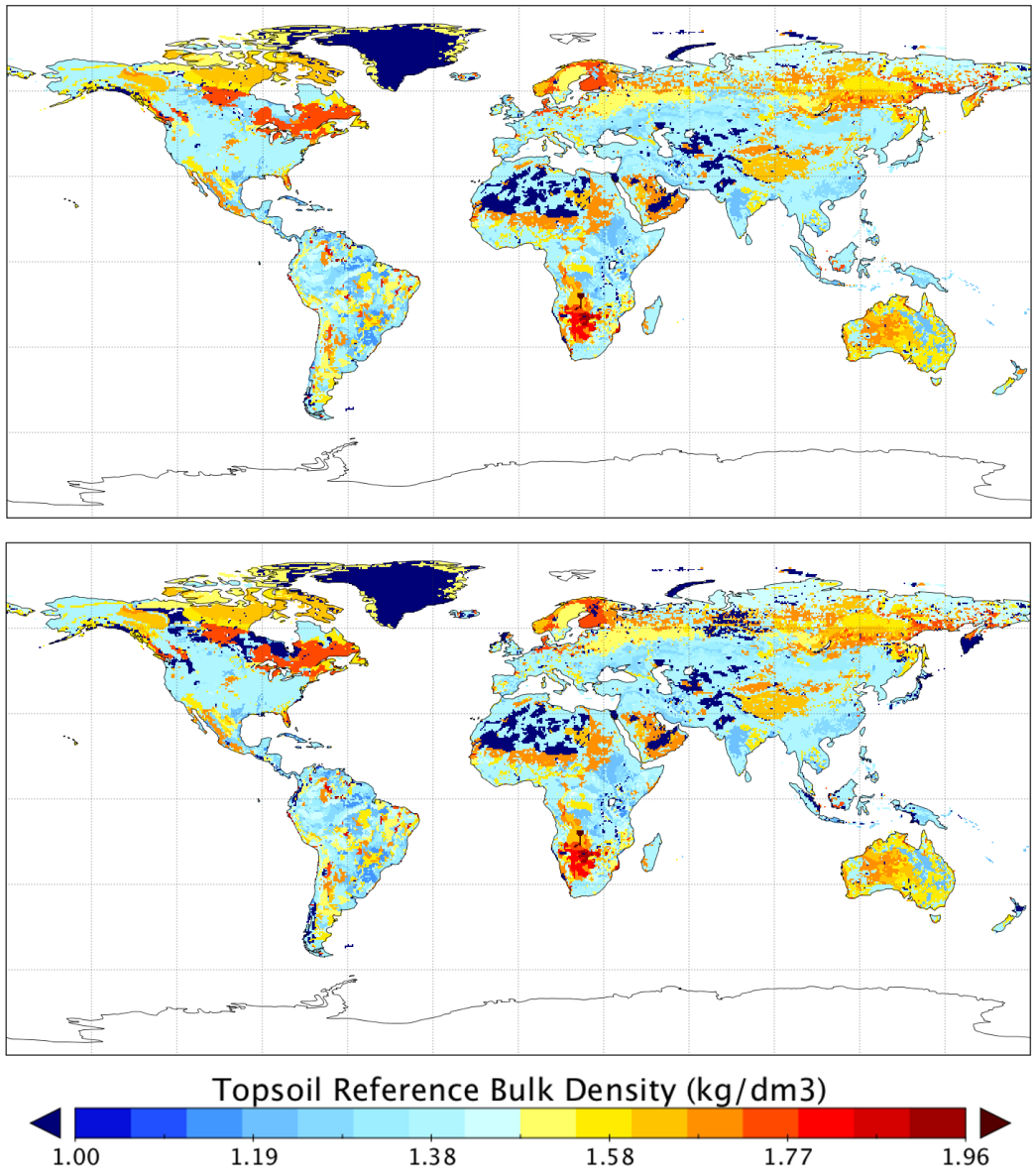


Figure 7. HWSD topsoil reference bulk density before (top) and after (bottom) correction at 0.5° resolution