Site-Level North American Carbon Program (NACP) Interim Synthesis

1. Introduction

The site-level NACP data set is used to assess how well models simulate carbon process across vegetation types and environmental conditions at 47 eddy covariance flux tower sites in North America. This data set includes the measurements at the flux towers and the outputs of 24 terrestrial biosphere models (Schwalm *et al.*, 2010). The data set is divided into three separate sub-data sets, which are model driver data and observations, the processed model output, and the original model output. The model driver data and observations not only include gap-filled model driver data and flux observations with uncertainties but also include phenology data and site ancillary data. The results of the site-level NACP interim synthesis appeared and will appear in a series of papers, each focusing on a different aspect of model performance. A list of publications is in Appendix I. The authorship of different data products is in Appendix II. Supplemental information about sites is in Appendix III.

2. Data set 1: Model Driver Data and Observations with Uncertainty

2.1 Gap-Filled Model Driver Data

The gap-filled meteorology data include eight variables: air temperature (K), specific humidity (kg/kg), wind speed (m/s), precipitation (kg/m^2/s), surface pressure (Pa), surface incident shortwave radiation (W/m^2), surface incident longwave radiation (W/m^2), and CO_2 concentration (ppm). The files are in both ASCII format and NetCDF format (ALMA standard). The time resolution of these files matches the averaging period used by the site PIs, which is half-hourly for most sites but hourly for several sites. Time in each file is UTC where each timestamp marks the end of the observation period. The time resolution is either 30 minutes or 60 minutes depending on the averaging period used at the flux site. Filenames follow the 6-character tower codes used in the site lists (see Appendix III Tables 1, 2, and 3). Please note that these data are subject to the AmeriFlux fair-use data policy and the Fluxnet-Canada data policy. These data should not be used for purposes other than the NACP site-level interim synthesis without the permission of site investigators.

2.1.1 Filling Methodology

If there is another flux tower on the list within 30 km and within 150 m elevation, data from this nearby site are used to fill. Remaining gaps are filled using NCDC climate station data. NCDC climate stations within 50 km are available for most sites. About half of these NCDC stations had hourly measurements, generally from ASOS sites. The rest were usually coop sites reporting daily data. In addition, DAYMET Version 1 modeled fine-scale climate data are available for continental U.S. sites at a daily time-step through the year 2003. Hourly data were used first when available. When both DAYMET and NCDC daily data are available, the data set with the better correlation is used. When station data are not available, a 10-day running mean diurnal cycle is used. The table 4 in the appendix provides links to NCDC stations that were used to fill each site. Detailed descriptions for individual variables follow below.

Temperature (Tair)

If hourly data from a nearby NCDC station are available, these are used to perform the filling. Biases between the NCDC station and the tower are corrected by adding the difference in mean annual temperature

If daily data are available but not hourly, we use DAYMET or NCDC maximum and minimum temperature data. If both options are available, we choose the option that has the best linear correlation with available tower data. Biases are corrected by adding the difference in mean annual temperature between the tower and the climate data (separately for maximum and minimum temperature). Daily data are converted to hourly (or half-hourly) using a sine wave assuming the following: High temperature occurs at 15 LST, low temperature occurs at 3 LST. Additionally, a 6-hr running mean filter is applied on filled data between 18 LST and 06 LST to prevent large discontinuities in temperature across day boundaries. If neither hourly nor daily climate data are available, data is filled with diurnal running mean method described above.

Near surface specific humidity (Qair)

If hourly temperature and heterotrophic respiration data from a nearby NCDC station are available, convert these to Qair and use to fill. Potential biases between the tower and NCDC station are not corrected.

If hourly climate data are not available, data is filled with diurnal mean method described above.

Wind speed (Wind)

If hourly wind speed data from a nearby NCDC station are available, these data are used to fill gaps. Potential biases between the tower and NCDC station are not corrected.

If hourly climate data are not available, data is filled with diurnal mean method described above.

Precipitation (Rainf)

If hourly precipitation data from a nearby NCDC station are available, these data are used to fill gaps. Potential biases between the tower and NCDC station are not corrected.

If daily data are available and not hourly, use DAYMET or NCDC precipitation data. If both options are available, choose the option that has the best linear correlation with available tower data. Potential biases between the tower and NCDC station are not corrected. Daily data are converted to hourly (or half-hourly) by dividing the daily sum by 24 (or 48) and evenly distributing rainfall throughout the day.

Surface pressure (Psurf)

If hourly surface pressure data from a nearby NCDC station are available, these data are used to fill gaps. Potential biases between the tower and NCDC station are not corrected.

If hourly data are not available, data is filled using diurnal mean method described above.

If pressure is not measured at all (occurs for several sites), data is filled using hydrostatic balance given site elevation.

Downwelling shortwave radiation (SWdown)

SWdown may be converted from PAR if PAR measurements are deemed to be more reliable.

If hourly cloud cover data from a nearby NCDC station are available, there are four categories: 0 (clear), 1 (scattered clouds), 2 (broken clouds) and 3 (overcast). The existing tower SWdown measurements are divided into these four categories. Diurnal mean averages are calculated for each day of the year for each cloud cover category. Gaps are then filled using this diurnal information combined with the NCDC cloud cover data.

If DAYMET is available but not hourly NCDC, use DAYMET estimated radiation data. DAYMET radiation is estimated from a simple model based on daily temperature range. DAYMET daily radiation is downscaled to hourly (or half-hourly) values using a mean diurnal cycle.

If neither of the above is available, data is filled using diurnal mean method described above.

Downwelling longwave radiation (LWdown)

Emissivity were estimated using LWdown data observed at Bondville flux tower data based on temperature, humidity and cloud cover from a nearby NCDC station. A lookup table of emissivity for a given temperature and cloud cover was created for use in gap-filling all sites.

If hourly cloud cover data from a nearby NCDC station are available, there are four categories: 0 (clear), 1 (scattered clouds), 2 (broken clouds) and 3 (overcast). This information was used in combination with temperature and humidity measurements to look up an emissivity. The emissivity was then used to calculate LWdown with air temperature using the Stefan-Boltzmann Law. If hourly cloud cover were not available, another version of the lookup table was used that uses an average over sky conditions.

Carbon dioxide concentration (CO2air)

As high frequency CO_2 measurements are often unreliable, no tower measurements were used to produce CO_2 concentrations. Instead, an algorithm selects the closest continental flask site out of the following list: Argyle, ME; Barrow, AK; Estevan Point, BC; Park Falls, WI; La Jolla, CA; Niwot Ridge, CO; Moody, TX; Grifton, NC. Globalview CO_2 data from these sites are then used. No temporal or spatial interpolation is performed. Flask measurements are generally every 14 days. Therefore, these CO_2 measurements will not capture the diurnal cycle and will miss some synoptic variability, but will characterize the seasonal cycle well.

2.2 Phenology

We provide remotely sensed standard phenology data for each site derived from GIMMS version g NDVI data set and MODIS. We define plant phenology as periodic or seasonal changes in Leaf Area Index (LAI) and absorbed fraction of Photosynthetically Active Radiation (fPAR). Each file represents an average seasonal cycle for one site. The 'phenology_readme' file in the companion file folder describes how the phenology files were created.

The gap-filled and smoothed MODIS LAI and fPAR time series data for 2001-2006 for each of the sites in the site-level synthesis. The MODIS data information for each site can be found in 'MODIS_LAI-fPAR_gap-filled_flux_tower_sites.xls' in the companion file folder. The gap-filled and smoothed LAI and fPAR data products were prepared by the MODIS for NACP Project (http://accweb.nascom.nasa.gov/). Information about the methods used can be found on the MODIS for NACP Web site.

We used the ORNL DAAC's global subsetting tool (http://daac.ornl.gov/MODIS/modis.html), which allowed us to input the site coordinates, MODIS product, and time period for Collection 4. The tool generates a set of time series graphs and some data files. For this interim synthesis, we've selected the 7 x 7 km area around the flux tower site, and we've pointed users to the summary statistics for those 7 x 7 km areas.

In addition to looking at the entire area (49 km^2), we also allow the user to look at only those pixels in the subset area that have the same land cover as the tower site. Please note that the land cover we use is the MODIS Land Cover for Collection 4 (the latest land cover from the MODIS team).

The Data Visualization page (Column G) in 'MODIS_LAI-fPAR_gap-filled_flux_tower_sites.xls' has a number of links to useful information about the products, sites, and our methods. The ASCII file (Column H) in 'MODIS_LAI-fPAR_gap-filled_flux_tower_sites.xls', which contains the summary statistics for each 8-day period, is posted on one of our servers. The description of the contents of that file can be found on the Data Visualization page:

http://daac.ornl.gov/MODIS/Download_Data_File_Description.html

2.3 Observations and Uncertainties

Quantified uncertainty and bias of the flux measurements are essential to the core objectives of the Site Synthesis. The basic strategy is to use published estimates of the various sources of uncertainty. If published estimates are lacking or incomplete for a particular source of uncertainty, we coordinated with analysis teams in AmeriFlux, Fluxnet Canada, and Fluxnet to develop estimates suitable for use in the Site Synthesis. We then combined the various estimates of uncertainty to obtain total uncertainty for the observed NEE, and the derived Gross Primary Productivity (GPP) and respiration (R). We estimated total flux uncertainty at the native time resolution of the observations (30 or 60-minute), as well as the diurnal, seasonal, and annual time scale. Components of uncertainty include uncertainties resulting from turbulence, gap-filling, flux partitioning and u* threshold determination. NEE is gap-filled using the standard approach of Fluxnet Canada. GPP and ecosystem respiration estimates with uncertainties are created using a standard flux partitioning. Random and u* threshold uncertainty data originally provided by Alan Barr were converted to netCDF format and updated to include gap-filling, partitioning, and total uncertainty by Kevin Schaefer. For the methodology of estimating uncertainty, please refer to Barr et al. (2009) and Schaefer et al. (2012).

Partitioning Uncertainty refers to the uncertainty in GPP and R estimated from observed NEE. Gap-filling Uncertainty refers to the structural uncertainty associated with the various techniques used to fill in missing values of observed NEE.

The uncertainty and gap-filled fluxes are generally expected to be stable for: CAGro, CALet, CAMer, CANS1, CAOas CAObs, CAOjp, CASJ1,2,3, CATP4, CAWP1, USARM (except 2006 and 2007), USHa1, USHo1, USIB1, USIB2, USLos, USMMS, USNR1, USNe1,2,3, USPFa, USSO2, USShd, USSyv.

Sites where there are unresolved questions about algorithm performance: CACa1, 2, 3 (significant low bias in NEP), USMe2, 3, 5, USTon, USVar, USUMB (may be OK), USWCr (may be OK).Sites where the product is most questionable: USAtq, USBrw, USDk2, USDk3.

There is more concern about the absolute value of the NEE fluxes and the partitioning of NEE into GPP and Re than about the uncertainty analysis or the characterization of inter-annual variability.

2.4 Ancillary Data

2.4.1 Cleaned Ancillary Data and Metadata

Investigators at each site provided ancillary information (in addition to the core micrometeorological measurements), describing site physical characteristics, disturbance history, and biological and ecological attributes of the vegetation, litter and soil. This information is recorded for each site using a standard data template (MS Excel spreadsheet) developed within the AmeriFlux network. The initial filled-in template provided by a participating site is referred to as an Ancillary Data and Metadata (ADM) record. These ADM records are archived at CDIAC and forwarded to the Berkeley Water Center for additional processing.

Following submission of the ADM record for a given site, the spreadsheet is processed at the Berkeley Water Center for basic QA/QC. This includes ingest of the ADM record into a database, standardization of date formats, unit conversions, consistent recording of missing data fields, and resolution of non-standard data field entries. A "cleaned" version of the data from the ADM record is then exported from the database in a standard format. These standard ancillary files are the official record of ADM information to be used by all project participants for modeling and analysis.

2.4.2 Site Location Table

The site location table (MS Excel spreadsheet) summarizes all site location information required as input to all models: latitude, longitude, elevation, instrument height, biome, start and stop years, time zone shift to local standard time, and the flux time averaging period. The site location table ('site_location_summary.xls') can be found in the companion file folder.

2.4.3 Soil Texture Summary Table

To save time for the modeling teams, we extracted data from the individual CADM files and constructed a soil texture summary table (MS Excel spreadsheet). Missing values were filled from the IGBP soil texture map. Sites with multiple observations as a function of depth were averaged to a single value. The 'soil_texture_summary.xls' can be found in the companion file folder.

3. Data set 2: The Standard Model Output Data

Standardized ASCII files have been created for gross primary productivity (GPP), net ecosystem exchange (NEE), ecosystem respiration (Re), latent heat flux (LE) and sensible heat flux (H). Each file contains output from 24 models for one variable at one site. These files also contain gap-filled observations and total uncertainty estimates from Barr *et al.* (2009) and Schaefer et al. (2012). All types of uncertainties are assumed to be independent and added in quadrature. These files are available at the native half-hourly time step, or in daily, monthly and annual averages. Standardized files are in local standard time (LST) with the timestamp representing the beginning of the averaging period. To match this standard, model output has been time-shifted when necessary. Included here is the latest version (mid-December 2009). GPP, NEE, and Re files have been further updated in November 2012 to provide total uncertainty estimated by Schaefer et al. (2012). However, LE and H have not been updated, and only contain the random and u* threshold uncertainty summed in quadrature. Plots representing diurnal and seasonal averages for these output variables are also provided in the companion file folder.

4. Data set 3: The Original Model Output

The original model output submitted by model teams was also archived with only format and units conversions. Generally there is one file per site per model, although several models also have separate files for each year. The model output format and size vary with different model teams. For a complete listing of model output variables and units, see the Site Synthesis Protocol in the companion file folder. It is strongly recommended that users work closely with the modeling teams when using these data to ensure accurate analysis. The access to the original model output will be sent by email upon request.

References:

Barr, A.G., D. Hollinger, and A. D. Richardson 2009. CO_2 flux measurement uncertainty estimates for NACP. Eos Transactions SGU, 90(52), Fall Meeting, Abstract B54A-04.

Schaefer, K., C. Schwalm, C. Williams, M.A. Arain, A. Barr, J. Chen, K.J. Davis, D. Dimitrov, T.W. Hilton, D.W. Hollinger, E. Humphreys, B. Poulter, B.M. Raczka, A.D. Richardson, A. Sahoo, P.E. Thornton, R. Vargas, H. Verbeeck, R. Anderson, I. Baker, T.A. Black, P. Bolstad, Jiquan Chen, P. Curtis, A.R. Desai, M. Dietze, D. Dragoni, C. Gough, R.F. Grant, L. Gu, A. Jain, C. Kucharik, B. Law, S. Liu, E. Lokipitiya, H.A. Margolis, R. Matamala, J.H. McCaughey, R. Monson, J.W. Munger, W. Oechel, C. Peng, D.T. Price, D. Ricciuto, W.J. Riley, N. Roulet, H. Tian, C. Tonitto, M. Torn, E. Weng, X. Zhou 2012. A Model-Data Comparison of Gross Primary Productivity: Results from the North American Carbon Program Site Synthesis. *Journal of Geophysical Research- Biogeosciences*, 117, doi:10.1029/2012JG001960

Schwalm, C. R., 2010. A model-data intercomparison of CO_2 exchange across North America: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research*, 115, doi:10.1029/2009JG001229.

Appendix I – List of Publications

Published

Dietze, M.C., R. Vargas, A.D. Richardson, P.C. Stoy, A.G. Barr, R.S. Anderson, M.A. Arain, I.T. Baker, T.A. Black, J.M. Chen, P. Ciais, L.B. Flanagan, C.M. Gough, R.F. Grant, D. Hollinger, C. Izaurralde, C.J. Kucharik, P. Lafleur, S. Liu, E. Lokupitiya, Y. Luo, J.W. Munger, C. Peng, B. Poulter, D.T. Price, D.M. Ricciuto, W.J. Riley, A.K. Sahoo, K. Schaefer, A.E. Suyker, H. Tian, C. Tonitto, H. Verbeeck, S.B. Verma, W. Wang, and E. Weng. 2012. Characterizing the performance of ecosystem models across time scales: A spectral analysis of the North American Carbon Program site-level synthesis. *Journal of Geophysical Research: Biogeosciences* 116: G04029. doi: 10.1029/2011JG001661

Keenan, T.F., I. Baker, A. Barr, P. Ciais, K. Davis, M. Dietze, D. Dragoni, C.M. Gough, R. Grant, D. Hollinger, K. Hufkens, B. Poulter, H. McCaughey, B. Rackza, Y. Ryu, K. Schaefer, H. Tian, H. Verbeeck, M. Zhao, and A.D. Richardson. 2012. Terrestrial biosphere model performance for inter-annual variability of land-atmosphere CO_2 exchange. *Global Change Biology* 18(6): 1971–1987. doi: 10.1111/j.1365-2486.2012.02678.x

Li H., M. Huang, M.S. Wigmosta, Y. Ke, A.M. Coleman, L.Y.R. Leung, A. Wang, and D.M. Ricciuto. 2011. Evaluating runoff simulations from the Community Land Model 4.0 using observations from flux towers and a mountainous watershed. *Journal of Geophysical Research: Atmospheres* 116: D24120. doi: 10.1029/2011JD016276

Richardson, A.D., R.S. Anderson, M.A. Arain, A.G. Barr, G. Bohrer, G. Chen, J.M. Chen, P. Ciais, K.J. Davis, A.R. Desai, M.C. Dietze, D. Dragoni, S.R. Garrity, C.M. Gough, R. Grant, D.Y. Hollinger, H.A. Margolis, H. McCaughey, M. Migliavacca, R.K. Monson, J.W. Munger, B. Poulter, B.M. Raczka, D.M. Ricciuto, A.K. Sahoo, K. Schaefer, H. Tian, R. Vargas, H. Verbeeck, J. Xiao, and Y. Xue. 2012. Terrestrial biosphere models need better representation of vegetation phenology: Results from the North American Carbon Program site synthesis. *Global Change Biology* 18(2): 566-584. doi: 10.1111/j.1365-2486.2011.02562.x

Schaefer, K., C. Schwalm, C. Williams, M.A. Arain, A. Barr, J. Chen, K.J. Davis, D. Dimitrov, T.W. Hilton, D.W. Hollinger, E. Humphreys, B. Poulter, B.M. Raczka, A.D. Richardson, A. Sahoo, P.E. Thornton, R. Vargas, H. Verbeeck, R. Anderson, I. Baker, T.A. Black, P. Bolstad, Jiquan Chen, P. Curtis, A.R. Desai, M. Dietze, D. Dragoni, C. Gough, R.F. Grant, L. Gu, A. Jain, C. Kucharik, B. Law, S. Liu, E. Lokipitiya, H.A. Margolis, R. Matamala, J.H. McCaughey, R. Monson, J.W. Munger, W. Oechel, C. Peng, D.T. Price, D. Ricciuto, W.J. Riley, N. Roulet, H. Tian, C. Tonitto, M. Torn, E. Weng, X. Zhou 2012. A model-data comparison of gross primary productivity: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research: Biogeosciences* 117: G03010. doi: 10.1029/2012JG001960

Schwalm, C. and 47 coauthors. 2010. A model-data intercomparison of CO_2 exchange across North America: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research: Biogeosciences* 115: G00H05. doi: 10.1029/2009jg001229

Sulman, B.N., A.R. Desai, N.M. Schroeder, D. Ricciuto, A. Barr, A.D. Richardson, L.B. Flanagan, P.M. Lafleur, H. Tian, G. Chen, R.F. Grant, B. Poulter, H. Verbeeck, P. Ciais, B. Ringeval, I.T. Baker, K. Schaefer, Y. Luo, and E. Weng. 2012. Impact of hydrological variations on modeling of peatland CO_2 fluxes: Results from the North American Carbon Program site synthesis. *Journal of Geophysical Research: Biogeosciences* 117: G01031. doi: 10.1029/2011JG001862

Complete drafts circulated to authors, to be submitted soon:

Stoy, P. and 32 coauthors. 201x. Evaluating the agreement between measurements and models and different times and time scales using wavelet coherence. To be submitted to *Biogeosciences* or *Journal of Geophysical Research: Biogeosciences.*

Raczka, B.M., B.M., K.J. Davis, D. Huntzinger, D. Ricciuto, A. Richardson, J. Xiao, K. Schaefer, E. Tomellieri, H. Tian, H. Verbeek, and others. Evaluation of continental carbon cycle simulations with North American flux tower observations.

Appendix II

Table 1. List of authors for model driver data and observations with uncertainty (Data set 1).

Contact	Email
Alan G. Barr	alan.barr@ec.gc.ca
Dan M. Ricciuto	ricciutodm@ornl.gov
Kevin Schaefer	kevin.schaefer@nsidc.org
Andrew Richardson	arichardson@oeb.harvard.edu
Deb Agarwal	daagarwal@lbl.gov
Peter E. Thornton	thorntonpe@ornl.gov
Ken Davis	davis@meteo.psu.edu
Barbara Jackson	
Robert B. Cook	cookrb@ornl.gov
David Y. Hollinger	dhollinger@fs.fed.us
Catharine van Ingen	vaningen@microsoft.com
Brian Amiro	brian_amiro@umanitoba.ca
Arlyn Andrews	arlyn.andrews@noaa.gov
M. Altaf Arain	arainm@mcmaster.ca
Dennis Baldocchi	Baldocchi@berkeley.edu
T. Andy Black	andrew.black@ubc.ca
Paul Bolstad	pbolstad@umn.edu
Peter Curtis	curtis.7@osu.edu
Ankur Desai	desai@aos.wisc.edu
Danilo Dragoni	ddragoni@indiana.edu
Lawrence Flanagan	larry.flanagan@uleth.ca
Lianhong Gu	lianhong-gu@ornl.gov
Gabriel Katul	gaby@duke.edu
Beverly Law	bev.law@oregonstate.edu
Peter Lafleur	plafleur@trentu.ca
Hank Margolis	hank.margolis@sbf.ulaval.ca
Roser Matamala	matamala@anl.gov
Tilden Meyers	Tilden.Meyers@noaa.gov
Harry McCaughey	mccaughe@queensu.ca
Russell Monson	russmonson@email.arizona.edu
J. William Munger	jwmunger@seas.harvard.edu
Walter Oechel	oechel@sunstroke.sdsu.edu
Ram Oren	ramoren@duke.edu
Nigel Roulet	nigel.roulet@mcgill.ca
Margaret Torn	mstorn@lbl.gov
Shashi Verma	sverma1@unl.edu

Table 2. List of authors for model outputs (Data set 2).

Contact	Model	Email
Dan M. Ricciuto		ricciutodm@ornl.gov
Kevin Schaefer	SIBCASA	kevin.schaefer@nsidc.org
Peter E. Thornton		thorntonpe@ornl.gov
Ken Davis		davis@meteo.psu.edu
Robert B. Cook		cookrb@ornl.gov
Shishi Liu		lius2@ornl.gov
Ryan Anderson	BIOME-BGC	ryan.anderson@ntsg.umt.edu
M. Altaf Arain	CN-CLASS	arainm@mcmaster.ca
Ian Baker	SiB3	baker@atmos.colostate.edu
Jing Ming Chen	BEPS	chenj@geog.utoronto.ca
Michael Dietze	ED2	mdietze@life.uiuc.edu
Robert Grant	ecosys	Robert.Grant@afhe.ualberta.ca
Cesar Izaurralde	EPIC	cesar.Izaurralde@pnl.gov
Atul K. Jain	ISAM	jain@atmos.uiuc.edu
Anthony W. King	LoTEC	kingaw@ornl.gov
Chris Kucharik	Agro-IBIS	kucharik@wisc.edu
Shuguang Liu	ECLUEEDCM	sliu@usgs.gov
Erandi Lokupitiya	SiBcrop	erandi@atmos.colostate.edu
Yiqi Luo	TECO	yluo@ou.edu
Changhui Peng	CLASS-CTEM (TRIPLEX-Flux)	peng.changhui@uqam.ca
Ben Poulter	LPJ wsl	ben.poulter@pik-potsdam.de
David Price	Can-IBIS	David.Price@NRCan-RNCan.gc.ca
William Riley	ISO-LSM	wjriley@lbl.gov
Alok Sahoo	SSIB2	aksahoo2004@gmail.com
Hanqin Tian	DLEM	tianhan@auburn.edu
Christina Tonitto	DNDC	ct244@cornell.edu
Hans Verbeeck	ORCHIDEE	Hans.Verbeeck@UGent.be

Table 3. List of authors for model outputs (Data set 3).

Contact	Model	Email
Dan M. Ricciuto		ricciutodm@ornl.gov
Kevin Schaefer	SIBCASA	kevin.schaefer@nsidc.org
Peter E. Thornton		thorntonpe@ornl.gov
Robert B. Cook		cookrb@ornl.gov
Ryan Anderson	BIOME-BGC	ryan.anderson@ntsg.umt.edu
M. Altaf Arain	CN-CLASS	arainm@mcmaster.ca
Ian Baker	SiB3	baker@atmos.colostate.edu
Jing Ming Chen	BEPS	chenj@geog.utoronto.ca
Michael Dietze	ED2	mdietze@life.uiuc.edu
Robert Grant	ecosys	Robert.Grant@afhe.ualberta.ca
Cesar Izaurralde	EPIC	cesar.Izaurralde@pnl.gov
Atul K. Jain	ISAM	jain@atmos.uiuc.edu
Anthony W. King	LoTEC	kingaw@ornl.gov
Chris Kucharik	Agro-IBIS	kucharik@wisc.edu
Shuguang Liu	ECLUEEDCM	sliu@usgs.gov
Erandi Lokupitiya	SiBcrop	erandi@atmos.colostate.edu
Yiqi Luo	TECO	yluo@ou.edu
Changhui Peng	CLASS-CTEM (TRIPLEX-Flux)	peng.changhui@uqam.ca
Ben Poulter	LPJ wsl	ben.poulter@pik-potsdam.de
David Price	Can-IBIS	David.Price@NRCan-RNCan.gc.ca
William Riley	ISO-LSM	wjriley@lbl.gov
Alok Sahoo	SSIB2	aksahoo2004@gmail.com
Hanqin Tian	DLEM	tianhan@auburn.edu
Christina Tonitto	DNDC	ct244@cornell.edu
Hans Verbeeck	ORCHIDEE	Hans.Verbeeck@UGent.be

Appendix III

Table 1 NACP Site-Level Synthesis – First-priority Site List: Primary sites with complete data sets (36 sites).

Site Code	Full Name	Period of Record[1]	State/ Prov	Type[2]
US-ARM	ARM – Southern Great Plains	2000-2006	OK	CRO
US-Ne1	Mead – Irrigated maize	2001-2006	NE	CRO
US-Ne2	Mead – Irrigated maize/soybean	2001-2006	NE	CRO
US-Ne3	Mead – Rainfed maize/soybean	2001-2006	NE	CRO
US-IB1	Fermi Lab – Maize/soybean rotation	2005-2007	IL	CRO
CA-Let	Lethbridge Grassland	1997-2006	AB	GRA
US-Var	Vaira Ranch	2001-2007	CA	GRA
US-Shd	Shidler	1997-2001	OK	GRA
US-IB2	Fermi Lab – Prairie	2004-2007	IL	GRA
CA-Oas	BERMS – Old Aspen	1997-2006	SK	DBF
US-Ha1	Harvard Forest – EMS Tower	1991-2006	MA	DBF
US-Dk2	Duke Forest – Hardwood	2003-2005	NC	DBF
US-UMB	University of Michigan Biological Station (UMBS)	1998-2006	MI	DBF
US-MMS	Morgan Monroe State Forest	1999-2006	IN	DBF
US-WCr	Willow Creek	1998-2006	WI	DBF
US-MOz	Missouri Ozark	2004-2007	MO	DBF
CA-Man	BOREAS – Northern Study Area, Old Black Spruce	1994-2006	MB	ENFB
CA-Obs	BERMS – Old Black Spruce	2000-2006	SK	ENFB
CA-Ojp	BERMS – Old Jack Pine	2000-2006	SK	ENFB
CA-Qfo	Quebec – Mature Black Spruce	2004-2006	QB	ENFB
CA-Ca1	Campbell River – Mature Douglas-fir	1998-2006	BC	ENFT
US-Dk3	Duke Forest – Loblolly Pine	1998-2005	NC	ENFT
US-Ho1	Howland Forest – Main Tower	1996-2004	ME	ENFT
US-Me2	Metolius – Intermediate-aged Ponderosa Pine	2002-2007	OR	ENFT
US-NR1	Niwot Ridge	1998-2007	CO	ENFT
CA-TP4	Turkey Point – Mature	2002-2007	ON	ENFT
US-PFa	Park Falls / WLEF	1997-2005	WI	MF
US-Syv	Sylvania Wilderness Area	2001-2006	MI	MF
CA-Gro	Groundhog River Station	2004-2006	ON	MF
US-Ton	Tonzi Ranch	2001-2007	CA	WSA

US-SO2	Sky Oaks – Old	1998-2006	CA	SHR
US-Brw	Barrow	1998-2006	AK	TUN
US-Atq	Atqasuk	1999-2006	AK	TUN
CA-Mer	Eastern Peatland – Mer Bleue	1999-2006	ON	WET
US-Los	Lost Creek	2000-2006	WI	WET
CA-WP1	Western Peatland – LaBiche River	2003-2007	AB	WET

Table 2 NACP Site-Level Synthesis – Second-priority Site List:
Secondary (chronosequence) sites with complete data sets (11 sites)

Site Code	**Full Name**	**Period of Record**[1]	**State/ Prov**	**Type**[2]
CA-SJ1	BERMS – Jack Pine, 1994 harvest	2002-2005	SK	ENFB
CA-SJ2	BERMS – Jack Pine, 2002 harvest	2003-2006	SK	ENFB
CA_SJ3	BERMS – Jack Pine, 1975 harvest	2004-2005	SK	ENFB
CA-Ca2	Campbell River – Douglas-fir clearcut	2001-2006	BC	ENFT
CA-Ca3	Campbell River – Douglas-fir juvenile	2002-2006	BC	ENFT
US-Me3	Metolius – Ponderosa Pine, young #2	2004-2005	OR	ENFT
US-Me4	Metolius – Ponderosa Pine, old-growth	1996-2000	OR	ENFT
US-Me5	Metolius – Ponderosa Pine, Young #1	1999-2002	OR	ENFT
CA-TP1	Turkey Point – Young	2003-2007	ON	ENFT
CA-TP2	Turkey Point – Seedling	2003-2007	ON	ENFT
CA-TP3	Turkey Point – Middle-aged	2003-2007	ON	ENFT

Table 3 NACP Site-Level Synthesis – Third-priority Site List:
Sites which lack only ancillary and biological data templates (11 sites)

Site Code	Full Name	Period of Record[1]	State/ Prov	Type[2]
US-Bo1	Bondville	1996-2006	IL	CRO
US-FPe	Fort Peck	2000-2007	MT	GRA
CA-NS1	UCI Chronosequence – 1850 burn	2001-2005	MB	ENFB
CA-NS2	UCI Chronosequence – 1930 burn	2001-2005	MB	ENFB
CA-NS3	UCI Chronosequence – 1964 burn	2001-2005	MB	ENFB
CA-NS4	UCI Chronosequence – 1964 burn (wet)	2001-2005	MB	ENFB
CA-NS5	UCI Chronosequence – 1981 burn	2001-2005	MB	ENFB
CA-NS6	UCI Chronosequence – 1989 burn	2001-2005	MB	ENFB
CA-NS7	UCI Chronosequence – 1998 burn	2001-2005	MB	ENFB
US-Blo	Blodgett Forest	1997-2006	CA	ENFT
US-Ivo	Ivotuk	2003-2006	AK	TUN

Notes (apply to all tables):

[1] Start-end years in the gap-filled weather data. Partial years (from flux data record) have been extended to complete years of surface weather data to simplify model forcing.

[2] Types assigned for convenience in this project, to identify combination of vegetation type and climate zone as an aid in site selection. These type names are not intended to match the IGBP classification assigned in other databases. CRO = crop, GRA = grassland, ENFB = Evergreen needleleaf forest – boreal, ENFT = evergreen needle leaf forest – temperate, DBF = deciduous broadleaf forest, MF = mixed (deciduous/evergreen) forest, WSA = woody savanna, SHR = shrubland, TUN = tundra, WET = wetland.

Table 4 NCDC stations that were used to fill each site.

Site Code	Priority	Class	Latitude	Longitude	Elevation (m)
US-Atq	1b		70.4696	-157.4089	16
US-Brw	1b	11	71.3225	-156.6259	1
US-Ivo	3		68.4865	-155.7503	570
CA-Ca1	1a	1	49.8673	-125.3336	300
CA-Ca2	2		49.8705	-125.2909	180
CA-Ca3	2		49.5346	-124.9004	165
US-Me4	2	1	44.4992	-121.6224	915
US-Me3	2		44.3157	-121.6078	1,005
US-Me5	2		44.4372	-121.5668	1,183
US-Me2	1a		44.4524	-121.5572	1,253
US-Ton	1b	8	38.4316	-120.966	177
US-Var	1b	10	38.4067	-120.9507	129
US-Blo	3	1	38.3953	-120.6328	1,315
US-SO2	1b	8	33.3739	-116.6229	1,392
CA-Let	1a	10	49.7093	-112.9402	960
CA-WP1	1b	5	54.9538	-112.467	540
CA-Oas	1a	4	53.6289	-106.1978	530
US-NR1	1	1	40.0329	-105.5464	3,050
CA-Obs	1a	1	53.9872	-105.1178	629
CA-Ojp	1b	1	53.9163	-104.692	579
CA-SJ3	2		53.908	-104.656	511
CA-SJ2	2		53.945	-104.649	518
CA-NS7	3	7	56.6358	-99.9483	273
CA-NS6	3	7	55.9168	-98.9644	276
CA-NS2	3	1	55.9058	-98.5247	268
CA-NS5	3	1	55.8631	-98.485	265
CA-NS1	3	1	55.8792	-98.4839	260
CA-Man	1b	1	55.8796	-98.4808	259
CA-NS3	3	1	55.9117	-98.3822	260
CA-NS4	3	1	55.9117	-98.3822	260
US-ARM	1b	12	36.605	-97.4884	310
Us-Shd	1b	10	36.9333	-96.6833	350
US-Ne1	1b	12	41.1651	-96.4766	361
US-Ne2	1b		41.1649	-96.4701	361
US-Ne3	1a	12	41.1797	-96.4396	361
US-PFa	1b	5	45.9459	-90.2723	485
US-WCr	1b	4	45.8059	-90.0799	520
US-Los	1b	11	46.0827	-89.9792	480
US-Syv	1b	5	46.242	-89.3477	540
US-Bo1	3	12	40.0061	-88.2919	300
US-IB2	1b	10	41.8406	-88.241	227
US-IB1	1b		41.8593	-88.2227	227
US-MMS	1	4	39.3232	-86.4131	275
US-UMB	1a	4	45.5598	-84.7138	234
CA-Gro	1b	5	48.2167	-82.1556	300

CA-TP1	2		42.6606	-80.5595	190
CA-TP2	2		42.7744	-80.4588	230
CA-TP4	1b		42.7098	-80.3574	219
CA-TP3	2		42.7068	-80.3483	219
US-Dk2	1b	4	35.9736	-79.1004	160
US-Dk3	1b	1	35.9782	-79.0942	163
CA-Mer	1a	11	45.4094	-75.52	70
CA-Qfo	2		49.6925	-74.3421	382
US-Ha1	1a	4	42.5378	-72.1715	303
US-Ho1	1a	1	45.2041	-68.7403	60

Priorities:
1a = first priority list – top ten site
1b = first priority list - rest of sites
2 = second priority list
3 = third priority list

Classes:
1 = Evergreen needleleaf (ENF)
2 = Evergreen broadleaf (EBF)
3 = Deciduous needleleaf (DNF)
4 = Deciduous broadleaf (DBF)
5 = Mixed forest (MF)
6 = Closed shrubland (CSH)
7 = Open shrubland (OSH)
8 = Woody savanna (WSA)
9 = Savanna
10 = Grassland (GRA)
11 = Permanent wetland (WET)
12 = Cropland (CRO)

Availability
0 = not available
1 = meteorology and gap-filled fluxes available
2 = meteorology and fluxes available, but no gap-filled fluxes
5 = partial year available
9 = unprocessed 2007 data