SUPPLEMENTARY DATA

Climate control of terrestrial carbon exchange across biomes and continents

Chuixiang Yi[1], Daniel Ricciuto[2], Runze Li[3], John Wolbeck[1], Xiyan Xu[1], Mats Nilsson[4], Luis Aires[5], John D Albertson[6], Christof Ammann[7], M Altaf Arain[8], Alessandro C de Araujo[9], Marc Aubinet[10], Mika Aurela[11], Zoltán Barcza[12], Alan Barr[13], Paul Berbigier[14], Jason Beringer[15], Christian Bernhofer[16], Andrew T Black[17], Paul V Bolstad[18], Fred C Bosveld[19], Mark S J Broadmeadow[20], Nina Buchmann[21], Sean P Burns[22], Pierre Cellier[23], Jingming Chen[24], Jiquan Chen[25], Philippe Ciais[26], Robert Clement[27], Bruce D Cook[28], Peter S Curtis[29], D Bryan Dail[30], Ebba Dellwik[31], Nicolas Delpierre[32], Ankur R Desai[33], Sabina Dore[34], Danilo Dragoni[35], Bert G Drake[36], Eric Dufrêne[32], Allison Dunn[37], Jan Elbers[38], Werner Eugster[21], Matthias Falk[39], Christian Feigenwinter[40], Lawrence B Flanagan[41], Thomas Foken[42], John Frank[43], Juerg Fuhrer[7], Damiano Gianelle[44], Allen Goldstein[45], Mike Goulden[46], Andre Granier[47], Thomas Grünwald[48], Lianhong Gu[2], Haiqiang Guo[49], Albin Hammerle[50], Shijie Han[51], Niall P Hanan[52], László Haszpra[53], Bernard Heinesch[10], Carole Helfter[54], Dimmie Hendriks[55], Lindsay B Hutley[56], Andreas Ibrom[57], Cor Jacobs[38], Torbjörn Johansson[58], Marjan Jongen[59], Gabriel Katul[60], Gerard Kiely[61], Katja Klumpp[62], Alexander Knohl[21], Thomas Kolb[34], Werner L Kutsch[63], Peter Lafleur[64], Tuomas Laurila[11], Ray Leuning[65], Anders Lindroth[58], Heping Liu[66], Benjamin Loubet[23], Giovanni Manca[67], Michal Marek[68], Hank A Margolis[69], Timothy A Martin[70], William J Massman[43], Roser Matamala[71], Giorgio Matteucci[72], Harry McCaughey[73], Lutz Merbold[74], Tilden Meyers[75], Mirco Migliavacca[76], Franco Miglietta[77], Laurent Misson[78,117], Meelis Mölder[58], John Moncrieff[27], Russell K Monson[79], Leonardo Montagnani[80,81], Mario Montes-Helu[34], Eddy Moors[82],Christine Moureaux[10,83], Mukufute M Mukelabai[84], J William Munger[85], May Myklebust[65], Zoltán Nagy[86], Asko Noormets[87], Walter Oechel[88], Ram Oren[89], Stephen G Pallardy[90], Kyaw Tha Paw U[39], João S Pereira[59], Kim Pilegaard[57], Krisztina Pintér[86], Casimiro Pio[91], Gabriel Pita[92], Thomas L Powell[93], Serge Rambal[94], James T Randerson[46], Celso von Randow[95], Corinna Rebmann[64], Janne Rinne[96], Federica Rossi[77], Nigel Roulet[97], Ronald J Ryel[98], Jorgen Sagerfors[4], Nobuko Saigusa[99], María José Sanz[100], Giuseppe-Scarascia Mugnozza[101], Hans Peter Schmid[102], Guenther Seufert[103], Mario Siqueira[89], Jean-François Soussana[62], Gregory Starr[104], Mark A Sutton[105], John Tenhunen[106], Zoltán Tuba[86,117], Juha-Pekka Tuovinen[11], Riccardo Valentini[107], Christoph S Vogel[108], Jingxin Wang[109], Shaoqiang Wang[110], Weiguo Wang[111], Lisa R Welp[112], Xuefa Wen[110], Sonia Wharton[113], Matthew Wilkinson[20], Christopher A Williams[114], Georg Wohlfahrt[50], Susumu Yamamoto[115], Guirui Yu[110], Roberto Zampedri[44], Bin Zhao[49] and Xinquan Zhao[116]

[1] School of Earth and Environmental Sciences, Queens College, City University of New York, New York 11367, USA
[2] Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831 USA
[3] Department of Statistics, Pennsylvania State University, University Park, Pennsylvania 16802, USA
[4] Department of Forest Ecology, The Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden
[5] CESAM and Department of Environmental Engineering, School of Technology and Management, Polytechnic Institute of Leiria, Portugal

[6] Department of Civil and Environmental Engineering, Duke University, Durham, North Carolina 22708-0287, USA
[7] Federal Research Station Agroscope Reckenholz-Tänikon, Reckenholzstr. 191, 8046 Zürich, Switzerland
[8] School of Geography and Earth Sciences, McMaster University, Hamilton, ON, L8S 4K1, Canada
[9] Instituto Nacional de Pesquisas da Amazonia , Programa LBA, Campus-II, Manaus—Amazonas 69060, Brazil
[10] University of Liege, Gembloux Agro-Bio Tech, Unit of Biosystem Physics, 2 Passage des Déportés, 5030 Gembloux, Belgium
[11] Finnish Meteorological Institute, Climate Change Research, FI-00101 Helsinki, Finland
[12] Department of Meteorology, Eötvös Loránd University, H-1117 Budapest, Pázmány sétány 1/A, Hungary
[13] Climate Research Division, Environment Canada, Saskatoon, SK, S7N 3H5, Canada
[14] INRA, UR1263 EPHYSE, Villenave d'Ornon F-33883, France
[15] School of Geography and Environmental Science, Monash University, Clayton, Victoria, 3800 Australia
[16] Institute of Hydrology and Meteorology, Dresden University of Technology, Pienner Str. 23, D-01737 Tharandt, Germany
[17] Land and Food Systems, University of British Columbia, Vancouver, BC, V6T 1Z4, Canada
[18] University of Minnesota, 115 Green Hall 1530 Cleveland Avenue N St Paul, Minnesota, 55108, USA
[19] Royal Netherlands Meteorological Institute, 3730 AE De Bilt, The Netherlands
[20] Forest Research, Alice Holt Lodge, Farnham, Surrey, GU10 4LH, UK
[21] ETH, Zurich, Institute of Plant Science, Universitaetsstrasse 2, Zuerich 8092, Switzerland
[22]National Center for Atmospheric Research Boulder, CO 80307-3000, USA
[23] UMR INRA-INA PG—Environment & Arable Crops Unit 78850 Thiverval-Grignon, France
[24] Department of Geography, University of Toronto, Toronto, ON, M5S 3G3, Canada
[25]Department of Environmental Sciences, University of Toledo, Toledo, OH 43606-3390, USA
[26] LSCE, UMR CEA-CNRS, Batiment 709, CE, L'Orme des Merisiers, F-91191 Gif-sur-Yvette, France
[27] School of GeoSciences, The University of Edinburgh, Mayfield Road, Edinburgh, EH9 3JU, UK
[28] Biospheric Sciences Branch, NASA's Goddard Space Flight Center, Greenbelt, MD 20771, USA
[29] Department of Evolution, Ecology, and Organismal Biology, Ohio State University, Columbus, OH 43210, USA
[30] Department of Plant, Soil, and Environmental Science, University of Maine, Orono, ME 04469, USA
[31] Wind Energy Division, Risø National Laboratory for sustainable Energy, Technical University of Denmark, PO 49, DK-4000 Roskilde, Denmark
[32] Université Paris-Sud Bâtiment 362, Ecologie, Systematique et Evolution Orsay Cedex, F-91405, France
[33] Department of Atmospheric and Oceanic Sciences, University of Wisconsin, Madison, WI 53706, USA
[34] Northern Arizona University, School of Forestry Northern Arizona University, Flagstaff, AZ 86001, USA
[35] Atmospheric Science Program, Department of Geography, Indiana University, Bloomington, IN 47405, USA
[36] Smithsonian Environmental Research Center, Edgewater, MD 21037, USA
[37] Department of Physical and Earth Science, Worcester State College, 486 Chandler Street Worcester, MA 01602, USA
[38] ESS-CC, Alterra Wageningen UR, 6700 AA Wageningen, The Netherlands
[39] Atmospheric Science Group, LAWR, UC Davis, Davis, CA 95616, USA
[40] Institute for Meteorology, Climatology and Remote Sensing, University of Basel, Klingelbergstrasse 27, CH-4056 Basel, Switzerland
[41] Department of Biological Sciences, University of Lethbridge, 4401 University Drive, Lethbridge, Alberta, T1K 3M4, Canada
[42] Department of Micrometeorology, University of Bayreuth, 95440 Bayreuth, Germany
[43] USDA Forest Service, Rocky Mountain Research Station, 240 West Prospect, Fort Collins, CO 80526, USA
[44] IASMA Research and Innovation Centre, Fondazione E Mach, Environment and Natural Resources Area, San Michele all'Adige, I38010 Trento, Italy
[45] Department of Environmental Science, Policy and Management, University of California, Berkeley, CA 94720, USA
[46] Department of Earth System Science, University of California, Irvine, CA 92697, USA

[47] INRA, UMR 1137 Ecologie et écophysiologie Forestierès, F54280, Champenoux, France
[48] Technische Universität Dresden, Institute of Hydrology and Meteorology, Department of Meteorology, Piennerstrasse 9, 01737 Tharandtt, Germany
[49]Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Institute of Biodiversity Science, Fudan University, Shanghai 200433, People's Republic of China
[50] University of Innsbruck, Institute of Ecology Sternwartestr 15, Innsbruck 6020, Austria
[51] Institute of Applied Ecology, Chinese Academy of Sciences, Shenyang 110016, People's Republic of China
[52] Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO 80523, USA
[53] Hungarian Meteorological Service, H-1675 Budapest, PO Box 39, Hungary
[54] Centre for Ecology and Hydrology (Edinburgh) Bush Estate Penicuik, Midlothian, EH26 0QB, UK
[55] Department of Hydrology and Geo-Environmental Sciences, Boelelaan 1085, 1081 HV, VU University Amsterdam, The Netherlands
[56] School of Environmental and Life Sciences, Charles Darwin University, Darwin, NT 0909, Australia
[57] Biosystems Division, Risø National Laboratory for Sustainable Energy, Technical University of Denmark, PO 49, DK-4000 Roskilde, Denmark
[58] Geobiosphere Science Centre, Physical Geography and Ecosystems Analysis, Lund University, Sölvegatan 12, SE-223 62 Lund, Sweden
[59] Instituto Superior de Agronomia, Universidade Técnica de Lisboa, Tapada da Ajuda 1349-017 Lisboa, Portugal
[60] School of the Environment, Duke University, Durham, NC 27708-0328, USA
[61] Civil and Environmental Engineering Department, University College Cork, Cork, Republic of Ireland
[62] INRA, Unité d'Agronomie, 234, Avenue du Brézet, F-63000 Clermont-Ferrand, France
[63] Johann Heinrich von Thünen-Institut (vTI), Institut für Agrarrelevante Klimaforschung, Bundesallee 50, 38116 Braunschweig, Germany
[64] Department of Geography, Trent University, Peterborough, Ontario, K9J 7B8, Canada
[65] CSIRO Marine and Atmospheric Research, PO Box 3023, Canberra, ACT, 2601, Australia
[66] Department of Physics, Atmospheric Sciences & Geoscience, Jackson State University, Jackson, MS 39217, USA
[67] Rende Division, Institute for Atmospheric Pollution, Consiglio Nazionale delle Ricerche, 87036 Rende, Italy
[68] Institute of Systems Biology and Ecology, Division of Ecosystems Processes Lab. of Plants Ecological Physiology, Na Sadkach 7 370 050 Ceske Budejovice Czech Republic
[69] Centre d'études de la forêt Faculté de Foresterie et de Géomatique, Université Laval, Québec G1V 0A6, Canada
[70] University of Florida, Gainesville, FL 32611, USA
[71] Argonne National Laboratory, Biosciences Division, Argonne, IL 60439, USA
[72] National Research Council, Institute of Agroenvironmental and Forest Biology, 00015 Monterotondo Scalo (RM), Italy
[73] Department of Geography, Queen's University, Kingston, Ontario, K7L 3N6, Canada
[74] Max-Planck Institute for Biogeochemie, Jena, D-07745, Germany
[75] NOAA/ATDD, Oak Ridge, TN 37831-2456, USA
[76] Remote Sensing of Environmental Dynamics Laboratory, DISAT, Università degli Studi di Milano-Bicocca, Italy
[77] CNR-IBIMET, Istituto di Biometeorologia, via Giovanni Caproni 8, 50145 Firenze Italy
[78] CNRS-CEFE, 1919 route de Mende, 34293 Montpellier Cedex 5, France
[79] Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO 80309, USA
[80] Servizi Forestali, Agenzia per l'Ambiente, Provincia Autonoma di Bolzano, 39100 Bolzano, Italy
[81] Faculty of Sciences and Technologies, Free University of Bozen-Bolzano, Piazza Università 1, 39100, Bolzano, Italy
[82] Alterra Green World Research, Wageningen, NL 6700 AA, The Netherlands
[83] University of Liege, Gembloux Agro-Bio Tech, Unit of Crops Management, 2 Passage des Déportés, 5030 Gembloux, Belgium
[84] Zambian Meteorological Department, Western Province, Mongu, Zambia
[85] Division of Engineering and Applied Science, Department of Earth and Planetary Science, Harvard University, Cambridge, MA 02138, USA
[86] Institute of Botany and Ecophysiology, Agricultural University of Gödöllô, H-2103 Gödöllô, Páter Károly u. 1, Hungary

[87] Department of Forestry and Environmental Resources, North Carolina State University, NC 29695, USA
[88] Department of Biology, San Diego State University, San Diego, CA 92182-4614, USA
[89] Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC 27708-0328, USA
[90] Department of Forestry, University of Missouri, Columbia, Missouri 65211, USA
[91] CESAM and Department of Environment, University of Aveiro, Aveiro 3810-193, Portugal
[92] Instituto Superior Tecnico, Mechanical Engineering Department, 1049-001 Lisboa, Portugal
[93] The Department of Organismic and Evolutinary Biology, Harvard University, Cambridge, MA 02138, USA
[94] DREAM, CEFE, CNRS, UMR5175, 1919 route de Mende, F-34293 Montpellier Cedex 5, France
[95] Earth System Science Center, National Institute of Space Research, Cachoeira Paulista, SP 12630, Brazil
[96] Department of Physics, FI-00014, University of Helsinki, Finland
[97] Department of Geography, McGill University 805, Sherbrooke Street West Montréal, Québec, H3A 2K6, Canada
[98] Department of Wildland Resources, Utah State University, Logan, UT 84322-5230, USA
[99] Center for Global Environmental Research, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba 305-8506, Japan
[100] Centro de Estudios Ambientales del Mediterraneo, Parque Tecnologico, Charles H Darwin 14, E-46980 Paterna, Spain
[101] Agricultural Research Council, Department of Agronomy, Forestry and Land Use, 00184, Rome, Italy
[102] Atmospheric Environmental Research Institute of Meteorology and Climate Research, Forschungszentrum Karlsruhe, Garmisch-Partenkirchen, Germany
[103] Institute for Environment and Sustainability, Joint Research Center European Commission, TP 280, I-21020 Ispra, Italy
[104] Department of Biological Sciences, University of Alabama, Tuscaloosa, AL 35487-0206 USA
[105] Atmospheric Sciences Centre for Ecology and Hydrology (CEH), Bush Estate, Penicuik, Midlothian, EH26 0QB, UK
[106] Department of Plant Ecology, University of Bayreuth, 95440 Bayreuth, Germany
[107] Department of Forest Environment and Resources, University of Tuscia, I-01100 Viterbo, Italy
[108] The University of Michigan Biological Station, Pellston, MI 49769, USA
[109] School of Mathematics, Liaoning Normal University, Dalian 116039, People's Republic of China
[110] Institute of Geographic Sciences and Natural Resource Research, Chinese Academy of Science, Beijing 100101, People's Republic of China
[111] IMSG@National Center for Environmental Predictions, NOAA, Camp Springs, MD 20746, USA
[112] Geosciences Research Division, Scripps Institution of Oceanography, University of California, La Jolla, CA 92093, USA
[113] Atmospheric, Earth and Energy Division, Lawrence Livermore National Laboratory, Livermore, CA 94551, USA
[114] Graduate Degree Program in Geography, Clark University, Worcester, MA 01610-1477, USA
[115] Okayama University, Okayama 700-8530, Japan
[116] Northwest Plateau Institute of Biology, Chinese Academy of Sciences, Xining 810001 Qinghai, People's Republic of China

[117] Deceased.

The authors from seventh to the end are listed alphabetically and contributed equally to this work.

Methods

1. Meteorological data gap filling

Producing reliable estimates of site-average temperature, radiation and precipitation requires comprehensive gap-filling techniques because of the sporadic data collection outages that occur at eddy covariance sites. Without gap filling, the distribution of these gaps can bias long-term averages (e.g., if there are more gaps in summer, the site's mean temperature will have a low bias). Although gap-filled meteorological data are available from the FLUXNET database, these are problematic because they do not account for missing precipitation data. We developed an algorithm to locate the nearest flux tower or climate station in the National Climatic Data Center (NCDC in Asheville North Carolina) database to provide daily temperature and precipitation data. If data from a nearby tower were available, these were used to fill missing meteorological data. When alternate towers were not available within a 30 km radius, daily NCDC data from the nearest station were downscaled to hourly or half-hourly resolution and used to fill the gaps. Temperature data were downscaled by using the daily maximum and minimum information to construct a sine wave with the appropriate amplitude (assuming daily maximum at 15 LST and daily minimum at 3 LST), and precipitation data were downscaled by dividing daily totals by the number of daily time steps (24 or 48 depending on the site). Differences in annual averages between the eddy covariance site and the climate stations were adjusted using linear regression so that the inclusion of station data did not alter long-term temperature or precipitation averages.

Net radiation data were not available from NCDC. If no alternate tower was available, gaps in these data were filled with the diurnal average values for the given hour and day of year. Diurnal averages were calculated for each hour or half-hour and day of the year using all available years and a 20-day moving window. Similarly, if NCDC temperature and precipitation data were not available to fill data gaps, diurnal average values of the site were also used.

The accuracy of our empirical findings are limited by eddy flux measurements in the following aspects: (1) the flux sites probably do not represent true random samples of biome types; a number of biomes, like tropical rain forests and savannas, are underrepresented; and (2) potential biases exist in the eddy covariance method as a result of advection errors, energy imbalance errors, and errors associated with the data integration approach.

2. Segregation method

2.1 Posterior probability and prototype subgroups

We first employed mixtures of a third-order polynomial regression (Goldfeld and Quandt 1976) with two subpopulations, one for a temperature-limited group (TG) and the other for a dryness-limited group (DG). The mixtures regression provides us the posterior probabilities of each site belonging to TG and DG. Supplementary table S1 lists the posterior probability of each site belonging to the temperature limited group PP(TG) and to the dryness limited group PP(DG). From this table the initial temperature and dryness limited prototypes can be defined based on probability of belonging to a specific group. Using only sites that have a larger than 99% probability of belonging either to the T Group or to the D Group, a set of prototype subgroups can be selected. From the 125 site population 26 sites meet this objective criteria of being highly temperature limited and 21 sites meet the criteria for being highly dryness limited (the filled circles with mango colour in figures 2(a) and 2(b), respectively. Also see supplementary table S1).

Analysis of the 26 highly temperature limited sites (>99% confidence) and the 21 highly dryness limited sites (>99% confidence) allows a set of prototype equations to be developed, which will predict the NEE of any site based on their mean annual temperature or their dryness. These two prototype equations are:

$$NEE_T = -0.001T^3 - 0.0143T^2 + 0.0271T + 0.2399, \qquad (1)$$

$$NEE_D = -0.5726D^3 + 0.7323D^2 + 5.7007D - 9.9968, \qquad (2)$$

where T is mean annual temperature in ^{o}C, $D = R_n/(\lambda P)$ is dryness, R_n is mean annual net radiation MJ m^{-2} yr^{-1}, P is mean annual precipitation mm yr^{-1}, and λ (=2.5 MJ kg^{-1}) is the enthalpy of vaporization, NEE_T and NEE_D are the site-average NEE predicted by the prototype model (1) and (2).

2.2 Residual index

Having defined the criteria equations for temperature and dryness prediction of net ecosystem exchange of carbon, a further statistical analysis of the residual error between the predicted and observed NEE values can be performed. From this residual error analysis, a dimensionless residual index (RI) is given by:

$$RI = \frac{RE_D - RE_T}{RE_D + RE_T}, \qquad (3)$$

where $RE_D = \left| \left(NEE_D - NEE_O \right) / NEE_O \right| \times 100\%$ is a percent error in NEE_D prediction by the equation (2) for a site, NEE_O is the observed mean annual NEE at the site,

$RE_T = \left| \left(NEE_T - NEE_O \right) / NEE_O \right| \times 100\%$ is a percent error in NEE_T prediction by the equation (1) for the site. The values of RI were calculated for all the 125 sites and listed in supplementary table S1.

2.3 Grouping by the residual index

The residual index value (RI) is useful in the classification of different response functions of ecosystem carbon exchanges. A positive RI indicates a temperature-limited site while a negative RI indicates a dryness-limited site. However, how shall we interpret sites that have a RI value near zero? A low RI value indicates that the predictive ability of the NEE_T and NEE_D equations each have similar outcomes. Sites with a low RI appear to be equally limited by both mean annual temperature and dryness. Given this result we can classify a third type of sub group called the B group since they are approximately equally sensitive to both of the meteorological parameters of temperature and dryness. The B-group sites are defined with RI values between +30% and -30% (figure 3, supplementary table S1). The monotonic function of the T-group with temperature and the D-group with dryness are cross-verified by an independent nonparametric analysis (figure 1), as well as the bi-variable function of the B-group with temperature and dryness (details see the text).

3. Sensitivity analysis

Lengths (durations) of site data sets are different (supplementary table S1). To test the potential influence of different data set length on the results, we conducted sensitivity analysis in five cases: (1) removing all data that were before 2000; (2) removing all single-year sites; (3) removing all sites with less than three years of data; (4) removing all sites with less than four years of data; and (5) removing all sites with less than five years of data. The sensitivity analysis indicates that the relationships between site-average NEE and climate controls found in this paper are stable to the perturbation of difference of sampling years. This result raises the question why the results are insensitive to temporal perturbations. To answer this question, we conducted a comparison analysis between spatial and temporal variability of NEE. The fundamental reason for the relative insensitivity to variation in length of the data sets is that spatial variability of NEE is 2.5 times greater than temporal variability of NEE.

Mixture regression

Here we explain why the commonly used clustering methods, including K-mean, multiple discriminate analysis, mixture models, may not work well here. These methods cluster or partition the sample space of (T, D, NEE), where T stands for temperature, D for dryness. As an illustration under what condition the commonly used cluster methods work , we generate a random sample of size 200 from a mixture of normal distribution (x,y), 50% sample from bivariate normal distribution with mean (-1.5,0) and covariance matrix being a diagonal matrix with diagonal elements 1 and 1, and 50% sample from bivariate normal distribution with mean (1.5,0) and covariance matrix being a diagonal matrix with diagonal elements 1 and 1. The scatter plot of the generated sample is depicted in supplementary figure S3. The vertical line x=0 is the theoretic optimal line to partition the sample space into two parts with a certain misclassification rate. If the purpose is to partition the sample space, then one should be clustering methods to group data.

It is worth to noting that what we are interested in is the regression relation between (T,D) and NEE, and what we want to do is to group the data by the regression function of NEE on (T,D) rather than partition the sample space into some several subspaces. Therefore, we conduct mixture regression, which is different from a mixture model in that the mixture regression is to group data by taking into account the regression relation between response and predictors. As an illustration under what conditions the mixture regression may work better than the commonly used clustering method, we generate a random sample of size 200 from a mixture regression model:

with 50% probability, y= x + e, where the random error e follows N(0,1);
with 50% probability, y= -x + e, where the random error e follows N(0,1).

The scatter plot of (x,y) is depicted in supplementary figure S4, from which it is easy to imagine that most clustering method won't be able to partition the sample space into two parts with low misclassification rate, while mixture regression can be used to identify the two different regression relations.

Uncertainties in NEE gap-filling

Uncertainty about the u* threshold is the largest contributor, with annual uncertainties between 0.15 and 1.00 t C ha^{-1} yr^{-1}. The 90% confidence interval generally ranges between 10-20% of annual NEE. Gap-filling uncertainties can be estimated by comparing results from different methodologies.

Based on a survey of 18 different gap filling methods, Moffat *et al* (2007) concluded that most methods produced estimates of annual integrated NEE that were within 0.25 t C ha^{-1} y^{-1} of the mean of the other methods. Accumulated random error caused by turbulent sampling uncertainties are around 0.20 t C ha^{-1} yr^{-1} at Howland forest (Richardson *et al* 2006), and are expected to be of similar magnitude at other sites with some variation caused by differences in micrometeorological conditions (Wohlfahrt *et al* 2008a, 2008b). Total errors in annual estimates of NEE typically range between 0.3 and 1 t C ha^{-1} yr^{-1}. The total error is certainly below the value of 2 t C ha^{-1} yr^{-1} tested conservatively by a Monte-Carlo analysis. For the purpose of this study, we conservatively estimate the 90% confidence interval of site NEE by adding the three major sources of error in quadrature, assuming that the sources of error are independent and that u* uncertainty is 20% of annual NEE:

$$\sigma_{NEE}(\mathrm{t\ C\ ha^{-1}\ yr^{-1}}) = \sqrt{0.20^2 + 0.25^2 + (0.2 * NEE_{av})^2}$$

where NEE_{av} is the site average NEE. We conclude that these errors do not significantly affect the outcome of our analysis because the spatial variability in NEE among sites is much larger than the random error.

References

Goldfeld S M and Quandt R E 1976 A Markov model for switching regression *J. Econom.* **1** 3–16

Moffat A M *et al* 2007 Comprehensive comparison of gap filling techniques for eddy covariance net carbon fluxes *Agric. Forest Meteorol.* **147** 209–32

Richardson A D *et al* 2006 A multi-site analysis of random error in tower-based measurements of carbon and energy fluxes *Agric. Forest Meteorol.* **136** 1–18

Wohlfahrt G, Fenstermaker L F and Arnone J A III 2008a Large annual net ecosystem CO_2 uptake of a Mojave Desert ecosystem *Glob. Change Biol.* **14** 1475–87

Wohlfahrt G *et al* 2008b Seasonal and inter-annual variability of the net ecosystem CO_2 exchange of a temperate mountain grassland: effects of weather and management *J. Geophys. Res.* **113** D08110

Supplementary table S1. Main site characteristics, climatic index, posterior probability, residual index, group classification, and carbon flux of terrestrial ecosystems observed in this analysis.

Site Code	Latitude (^oN)	Longitude (^oE)	Elevation (m)	Vegetation type	T (^oC)	Dryness	C-flux (t C ha^{-1} yr^{-1})	PP (TG)	PP (DG)	RI	Group	Years of data
US-Atq	70.47	-157.41	15	WET	-10.60	4.87	-0.45	100.0%	0.0%	100%	TG	2003-2006
IE-Dri	51.99	-8.75	187	GRA	9.64	0.51	-1.85	99.9%	0.1%	95%	TG	2003
CA-Mer	45.41	-75.52	70	WET	6.21	1.05	-0.53	98.0%	2.0%	92%	TG	1999-2006
IT-Cpz	41.71	12.38	68	EBF	14.90	1.68	-5.60	100.0%	0.0%	90%	TG	1997, 2001-2006
CA-NS4	55.91	-98.38	260	ENF	-2.08	1.56	0.05	77.9%	22.1%	90%	TG	2003-2004
CA-NS7	56.64	-99.95	273	OSH	-1.70	1.41	0.29	92.2%	7.8%	89%	TG	2003-2004
IT-MBo	46.02	11.05	1550	GRA	5.65	0.97	-0.47	99.4%	0.6%	89%	TG	2003
AT-Neu*	47.12	11.32	970	GRA	6.50	0.67	-0.10	100.0%	0.0%	88%	TG	2001-2008
FI-Kaa	69.14	27.30	155	WET	-1.10	0.64	-0.20	100.0%	0.0%	88%	TG	2000-2007
CA-TP4	42.71	-80.36	184	ENF	8.55	1.08	-1.36	88.2%	11.8%	87%	TG	2003-2007
FI-Sod*	67.36	26.64	180	ENF	-0.70	0.80	0.62	100.0%	0.0%	87%	TG	2000-2001, 2003-2007
IT-PT1	45.20	9.06	60	DBF	14.27	1.82	-4.86	99.9%	0.1%	85%	TG	2003
US-WBW*	35.96	-84.29	283	DBF	14.92	0.95	-5.74	72.7%	27.3%	84%	TG	1995-1998
DK-Sor	55.49	11.65	40	DBF	8.25	0.75	-0.63	99.9%	0.1%	83%	TG	1997-2006
US-Wrc*	45.82	-121.95	371	ENF	8.92	0.54	-0.79	100.0%	0.0%	83%	TG	1999-2002, 2004

FR-Lq1	45.64	2.74	1040	GRA	7.66	0.32	-1.51	100.0%	0.0%	82%	TG	2004-2006
IT-SRo	43.73	10.28	4	ENF	14.20	1.59	-4.76	99.1%	0.9%	81%	TG	1999-2007
SE-Deg*	64.18	19.55	270	WET	2.56	0.45	-0.53	100.0%	0.0%	81%	TG	2001-2002, 2004-2005
US-Ivo	68.49	-155.75	570	WET	-9.37	1.38	-0.22	86.1%	13.9%	80%	TG	2004-2006
DE-Bay*	50.14	11.87	775	ENF	6.20	0.64	0.44	100.0%	0.0%	78%	TG	1997-1999
CA-Qfo*	49.69	-74.34	382	ENF	1.11	0.97	-0.33	99.6%	0.4%	74%	TG	2004-2006
FR-Lq2	45.64	2.74	1040	GRA	7.66	0.32	-1.86	100.0%	0.0%	72%	TG	2004-2006
CA-Qcu	49.27	-74.04	392	ENF	1.26	0.81	1.41	100.0%	0.0%	70%	TG	2002-2006
CA-SJ3	53.88	-104.64	488	ENF	2.17	2.06	0.31	59.2%	40.8%	69%	TG	2005
CA-TP1	42.66	-80.56	265	ENF	8.73	0.82	-0.38	99.8%	0.2%	68%	TG	2003-2007
CA-Man	55.88	-98.48	259	ENF	-1.23	1.91	0.09	63.4%	36.6%	68%	TG	1994-2006
IT-Amp	41.90	13.61	884	GRA	9.52	1.20	-1.28	73.5%	26.5%	65%	TG	2003-2006
PT-Esp	38.64	-8.60	95	EBF	16.02	2.17	-5.76	100.0%	0.0%	62%	TG	2002-2004, 2006-2007
CA-SJ2	53.94	-104.65	580	ENF	0.42	1.08	1.48	100.0%	0.0%	62%	TG	2003-2006
DE-Wet*	50.45	11.46	785	ENF	6.52	0.87	-1.32	98.8%	1.2%	61%	TG	2002-2007
US-FPe	48.31	-105.10	634	GRA	5.75	1.41	0.32	83.6%	16.4%	61%	TG	2000-2006
SE-Abi	68.36	18.79	TBD	DBF	0.10	0.42	-1.30	100.0%	0.0%	60%	TG	2005
CA-Ca3	49.53	-124.90	165	ENF	8.75	0.53	0.63	100.0%	0.0%	59%	TG	2001-2006
IT-Non	44.69	11.09	25	DBF	13.80	1.04	-5.04	68.2%	31.8%	57%	TG	2001-2003, 2006
SE-Nor	60.09	17.48	43	EBF	6.25	1.07	0.96	99.8%	0.2%	56%	TG	1996-1997, 1999, 2003, 2005

FI-Sii*	61.83	24.19	162	WET	3.99	1.35	-0.51	83.7%	16.3%	53%	TG	2005
IT-Ro2	42.39	11.92	224	DBF	14.88	1.42	-7.52	100.0%	0.0%	52%	TG	2002-2006
CA-NS6	55.92	-98.96	276	OSH	-0.35	1.51	-0.23	75.8%	24.2%	51%	TG	2002-2004
US-WCr	45.81	-90.08	520	DBF	5.27	1.21	-0.90	87.6%	12.4%	49%	TG	1999-2006
SE-Fla	64.11	19.46	226	ENF	2.69	1.27	-0.57	88.8%	11.2%	48%	TG	1997-1998. 2001-2002
DK-Lva	55.68	12.08	15	GRA	9.33	0.77	-2.57	93.4%	6.6%	48%	TG	2006-2007
JP-TAK*	36.15	137.42	1420	DBF	6.53	0.47	-2.28	99.8%	0.2%	46%	TG	1994-2004
US-Syv	46.24	-89.35	540	MF	4.20	1.01	-1.16	95.8%	4.2%	42%	TG	2002-2003, 2005
US-IB2	41.84	-88.24	227	GRA	10.46	2.14	-3.97	99.2%	0.8%	37%	TG	2005
US-PFa	45.95	-90.27	470	MF	4.99	1.24	-1.02	83.4%	16.6%	35%	TG	1997-2000, 2003
CA-Gro	48.22	-82.16	300	MF	3.36	1.30	-0.83	81.7%	18.3%	30%	TG	2004-2006
US-Me3*	44.32	-121.61	1005	ENF	8.49	2.76	-1.76	60.5%	39.5%	28%	BG	2004-2005
US-Ha1*	42.54	-72.17	340	DBF	7.88	0.78	-2.53	91.7%	8.3%	28%	BG	1992-2007
FR-LBr	44.72	-0.77	61	ENF	14.03	1.29	-4.12	77.7%	22.3%	27%	BG	1997-1998
HU-HH2*	46.96	16.65	248	GRA	8.90	1.10	-2.20	73.7%	26.3%	25%	BG	1999-2000, 2007
CA-Ojp	53.92	-104.69	579	ENF	1.52	1.69	-0.25	65.8%	34.2%	23%	BG	2000-2006
US-NC2*	35.80	-76.67	12	ENF	15.80	0.94	-5.91	79.8%	20.2%	22%	BG	2005-2008
CA-Let*	49.71	-112.94	960	GRA	6.41	2.12	-1.30	69.6%	30.4%	17%	BG	1999-2006
US-MOz	38.74	-92.20	219	DBF	13.52	1.47	-3.40	74.5%	25.5%	17%	BG	2005-2006
FR-Fon*	48.48	2.78	90	DBF	11.50	0.84	-3.80	63.6%	36.4%	13%	BG	2006

US-UMB*	45.56	-84.71	234	DBF	5.50	1.19	-1.51	76.8%	23.2%	11%	BG	1999-2003
US-OHO*	41.55	-83.84	230	DBF	10.40	1.42	-2.67	64.3%	35.7%	9%	BG	2004-2008
CH-Oe1*	47.29	7.73	450	GRA	9.57	0.65	-3.72	70.5%	29.5%	8%	BG	2002-2007
US-ME4*	44.44	-121.57	1183	ENF	7.89	2.77	-2.06	59.5%	40.5%	7%	BG	2001-2002
NL-Loo*	52.17	5.74	25	ENF	10.30	1.00	-3.07	65.3%	34.7%	4%	BG	1997-2007
US-Ho1*	45.20	-68.74	60	ENF	6.61	1.17	-1.88	70.6%	29.4%	-1%	BG	1996-2004
DE-Hai*	51.08	10.45	430	DBF	8.31	0.89	-2.94	69.1%	30.9%	-1%	BG	2000-2007
US-MLT*	42.50	-113.41	1370	GRA	8.75	2.90	-0.26	83.1%	16.9%	-1%	BG	2005
CA-Ca1	49.87	-125.33	300	ENF	8.69	0.73	-3.59	58.1%	41.9%	-6%	BG	1998-2006
US-Me2*	44.45	-121.56	1253	ENF	7.61	2.91	-4.71	5.9%	94.1%	-7%	BG	2002-2008
AU-Wac*	-37.43	145.19	545	EBF	10.10	0.80	-3.76	57.1%	42.9%	-8%	BG	2006
CN-Cha*	42.40	128.10	761	MF	4.80	1.90	-2.50	65.9%	34.1%	-9%	BG	2003-2004
US-Dk3*	35.98	-79.09	163	ENF	14.73	1.10	-4.54	69.4%	30.6%	-9%	BG	2001-2005
DE-Gri	50.95	13.51	385	GRA	7.99	0.97	-2.83	62.9%	37.1%	-12%	BG	2005-2006
CN-Do1	31.52	121.96	2-5	WET	15.64	0.58	-6.23	52.3%	47.7%	-17%	BG	2005
US-BN1*	63.92	-145.38	518	ENF	0.15	1.99	-1.40	50.6%	49.4%	-18%	BG	2002-2004
CA-WP1	54.95	-112.47	540	MF	1.87	1.85	-2.21	42.0%	58.0%	-20%	BG	2004-2007
CN-Do2	31.58	121.90	2-5	WET	15.56	0.70	-4.37	60.1%	39.9%	-21%	BG	2005
CA-SJ1	53.91	-104.66	580	ENF	0.68	2.08	-0.73	58.4%	41.6%	-24%	BG	2004-2005
US-Bar	44.06	-71.29	272	DBF	7.54	0.76	-3.71	30.3%	69.7%	-24%	BG	2004-2006

CN-HaM	37.37	101.18	3250	GRA	-1.53	2.48	-0.49	57.8%	42.2%	-25%	BG	2003-2005
IT-Ren*	46.59	11.43	1730	ENF	4.75	1.20	-2.00	54.7%	45.3%	-28%	BG	1999,2001-2007
US-BN3*	63.92	-145.74	469	MF	0.15	1.99	-0.09	61.2%	38.8%	-32%	DG	2002-2003
US-Blo	38.90	-120.63	1315	ENF	11.23	0.99	-5.76	14.2%	85.8%	-33%	DG	2000-2006
US-MMS	39.32	-86.41	275	DBF	12.36	1.05	-4.23	58.2%	41.8%	-33%	DG	1999-2005
US-Dk2*	35.97	-79.10	168	DBF	15.06	1.07	-4.44	63.6%	36.4%	-38%	DG	2001-2005
US-Fuf*	35.09	-111.76	2180	ENF	9.15	2.04	-0.58	32.0%	68.0%	-39%	DG	2007
US-Goo	34.25	-89.87	87	GRA	16.31	0.95	-2.13	2.1%	97.9%	-40%	DG	2003-2006
CA-NS2	55.91	-98.52	260	ENF	0.85	1.70	-1.91	33.3%	66.7%	-42%	DG	2002, 2004
HU-Bug	46.69	19.60	140	GRA	9.99	1.63	-0.74	27.6%	72.4%	-42%	DG	2003-2007
BE-Vie	50.31	6.00	450	MF	8.18	1.10	-5.17	2.8%	97.2%	-45%	DG	1997-2006
US-SP3*	29.75	-82.16	50	ENF	20.06	1.03	-6.40	79.5%	20.5%	-47%	DG	2001-2004
FR-Hes*	48.67	7.06	300	DBF	9.99	0.97	-3.71	46.3%	53.7%	-49%	DG	1997-1999, 2001-2007
DE-Tha*	50.96	13.57	380	ENF	8.79	0.94	-6.00	0.4%	99.6%	-51%	DG	1997-2007
AU-TUM	-35.66	148.15	1200	EBF	9.50	1.26	-3.37	52.3%	47.7%	-51%	DG	2002-2007
NL-Hor*	52.03	5.07	-2.2	GRA	10.98	1.11	-3.29	59.3%	40.7%	-52%	DG	2004-2005
IT-Col	41.85	13.59	1550	DBF	7.36	0.96	-5.87	0.1%	99.9%	-55%	DG	1997-1998, 2000-2001, 2005
CA-Oas	53.63	-106.20	530	DBF	2.27	1.67	-1.61	45.1%	54.9%	-56%	DG	1997-2006
US-Ton	38.43	-120.97	177	WSA	16.29	2.11	-1.71	0.2%	99.8%	-56%	DG	2002-2006
CA-NS1	55.88	-98.48	260	ENF	0.37	1.83	-0.94	53.3%	46.7%	-56%	DG	2004

CA-TP3	42.71	-80.35	184	ENF	8.81	1.10	-4.42	15.1%	84.9%	-58%	DG	2003-2007
US-Fmf*	35.14	-111.73	2160	ENF	9.99	2.07	0.51	2.7%	97.3%	-58%	DG	2007
IT-Ro1	42.41	11.93	234	DBF	15.37	1.38	-3.04	20.5%	79.5%	-61%	DG	2001-2006
UK-Gri	56.61	-3.80	340	ENF	7.38	0.86	-6.12	0.0%	100.0%	-63%	DG	1997-1998, 2000-2001
CN-Do3	31.52	121.97	2-5	WET	15.67	0.77	-5.12	60.0%	40.0%	-65%	DG	2005
FR-Pue	43.74	3.60	270	EBF	13.67	1.23	-2.60	31.7%	68.3%	-71%	DG	2001-2007
UK-Ham	51.12	-0.86	80	DBF	10.50	0.59	-5.88	1.6%	98.4%	-71%	DG	2004
US-Aud	31.59	-110.51	1469	GRA	16.12	1.94	0.97	0.0%	100.0%	-72%	DG	2003-2005
CA-NS5	55.86	-98.49	260	ENF	-1.76	1.69	-1.25	41.7%	58.3%	-72%	DG	2002, 2004
US-SO3	33.38	-116.62	1429	CSH	14.50	2.03	-0.89	0.2%	99.8%	-73%	DG	2005-2006
UK-EBu	55.87	-3.21	190	GRA	9.08	0.42	-6.73	0.0%	100.0%	-74%	DG	2004
CZ-BK1	49.50	18.54	908	ENF	8.26	0.64	-7.09	0.0%	100.0%	-74%	DG	2004-2006
PT-Mi1	38.54	-8.00	250	EBF	15.86	2.46	-0.89	0.0%	100.0%	-76%	DG	2003-2005
ZM-MON*	-15.43	23.25	1053	SAV	22.00	1.42	-0.01	0.0%	100.0%	-78%	DG	2007
US-Var	38.41	-120.95	129	GRA	15.94	1.60	-0.58	0.0%	100.0%	-79%	DG	2001-2006
ES-LMa	39.94	-5.77	260	SAV	16.16	1.46	-1.28	0.0%	100.0%	-82%	DG	2004-2006
US-GLE*	41.36	-106.24	3190	ENF	0.09	0.97	-3.90	0.2%	99.8%	-83%	DG	2005-2008
US-NR1	40.03	-105.55	3050	ENF	2.46	1.86	-0.49	60.5%	39.5%	-83%	DG	1999-2000, 2002-2003
CA-NS3	55.91	-98.38	260	ENF	-2.43	1.71	-0.89	49.9%	50.1%	-85%	DG	2002-2004
US-KS2	28.61	-80.67	3	CSH	22.11	1.31	-3.60	0.0%	100.0%	-85%	DG	2002, 2004-2006

BR-Ma2*	-2.61	-60.21	120	EBF	25.85	0.77	-3.87	0.0%	100.0%	-87%	DG	1999-2002
PT-Mi2*	38.48	-8.02	190	GRA	14.37	1.63	-0.93	0.3%	99.7%	-87%	DG	2005-2007
CN-QYZ*	26.74	115.07	100	MF	18.59	1.30	-3.07	0.1%	99.9%	-89%	DG	2003-2004
NL-Ca1*	51.97	4.93	0.7	GRA	10.93	0.97	-4.40	36.9%	63.1%	-90%	DG	2003-2004, 2006-2007
ZA-KRU*	-25.02	31.50	300	SAV	21.78	2.72	0.25	0.0%	100.0%	-90%	DG	2001-2005
US-SO2	33.37	-116.62	1394	CSH	14.36	1.97	-0.54	0.1%	99.9%	-91%	DG	2004-2005
AU-How*	-12.49	131.15	38	WSA	26.21	0.93	-3.60	0.0%	100.0%	-92%	DG	2001-2005
US-SP1*	29.74	-82.22	50	ENF	20.25	1.34	-1.99	0.0%	100.0%	-93%	DG	2001, 2003, 2005-2006
CA-Obs	53.99	-105.12	628	ENF	1.65	1.85	-0.55	59.6%	40.4%	-97%	DG	2000-2006
FI-Hyy	61.85	24.29	181	ENF	4.25	1.41	-2.09	39.7%	60.3%	-97%	DG	1997-1999, 2001-2004, 2006

The vegetation is coded according to the IGBP classification: CSH, closed shrublands; DBF, deciduous broad-leaf forests; EBF, evergreen broad-leaf forests; ENF, evergreen needle-leaf forests; GRA, grassland; MF, mixed forests; OSH, open shrublands; SAV, savannas; WET, permanent wetlands; WSA, woody savannas.

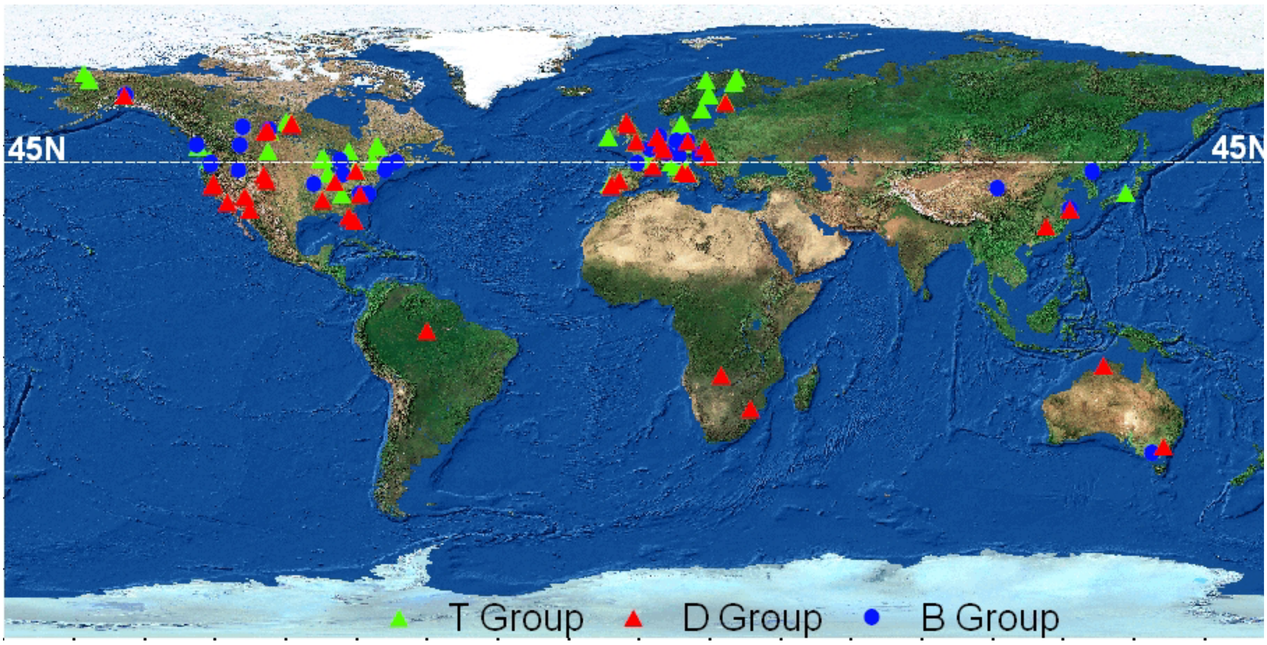
PP(TG) indicates the posterior probability of each site belonging to the temperature group.

PP(DG) indicates the posterior probability of each site belonging to the dryness group.

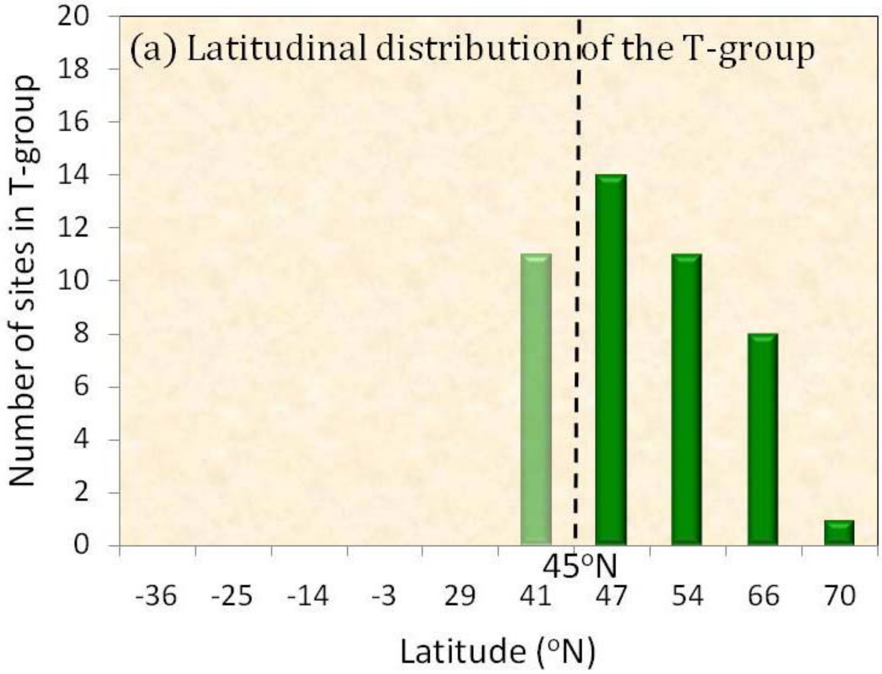
RI refers to the residual index defined by the equation (3).

BG stands for B group, TG for temperature group, and DG for dryness group.

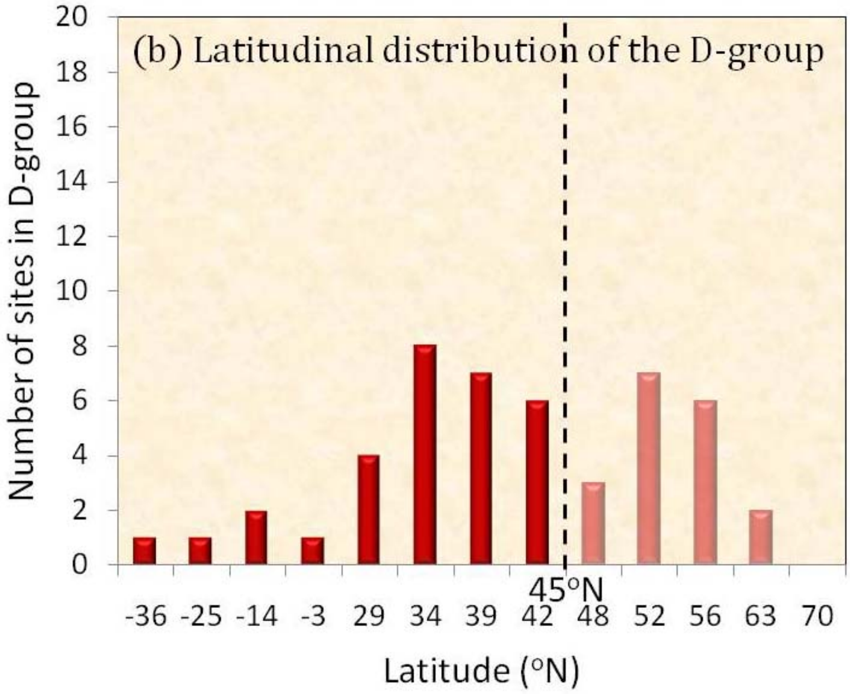
* indicates that NEE data was provided by the site P.I..



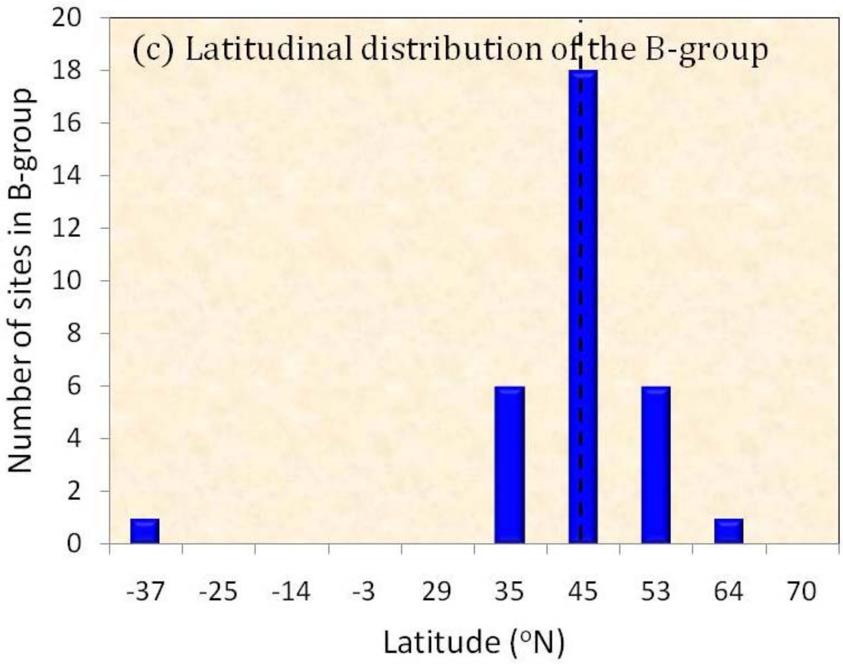
Supplementary figure S1. Geographical distribution of the sites in the three groups: temperature group, dryness group, and the B group.



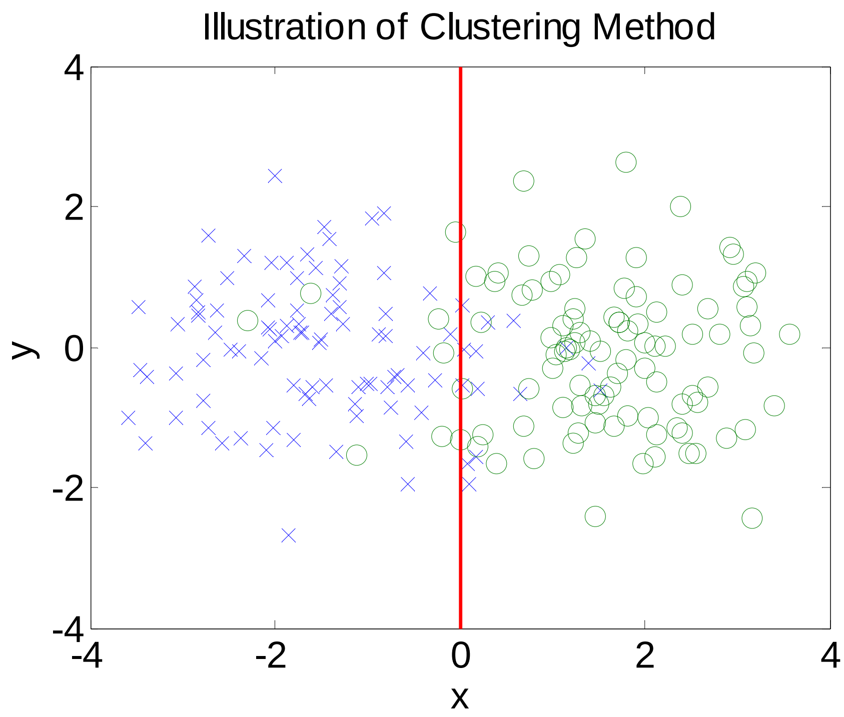
Supplementary figure S2. The latitudinal distribution of: (a) the T-group; (b) the D-group; and (c) the B-group.



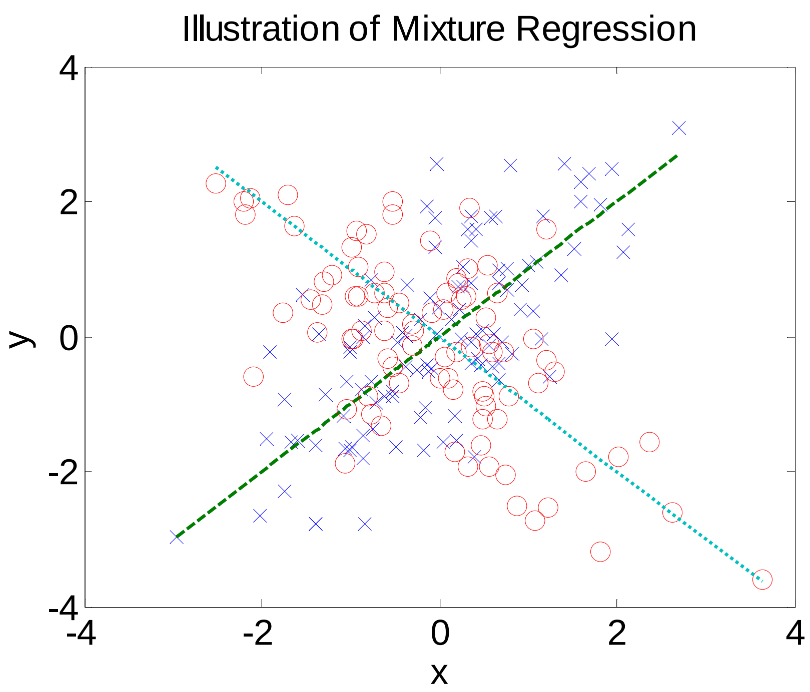
Supplementary figure S2. (Contnued.)



Supplementary figure S2. (Contnued.)



Supplementary figure. S3. Scatter plot of (x,y), 'o' stands for the samples from a bivariate normal distribution with mean (1.5,0) and covariance matrix being a diagonal matrix with diagonal elements 1 and 1, while 'x' stands for the samples from a bivariate normal distribution with mean (-1.5,0) and covariance matrix being a diagonal matrix with diagonal elements 1 and 1.



Supplementary figure S4. Scatter plot of (x,y). 'o' stands for the samples from y=-x + e, while 'x' stands for the samples from y=x+e.. The dashed line is the line of y=x, and the dotted line is the line of y=-x.