A set of benchmark tests for JULES

Eleanor Blyth[1*], Douglas Clark[1], Richard Ellis[1], Chris Huntingford[1], Sietse Los[3], Matt Pryor[2], Martin Best[2], Stephen Sitch[4]

[1] Centre for Ecology and Hydrology, Wallingford OX10 8BB, UK
[2] Hadley Centre for climate prediction and research, Met Office, Joint Centre for Hydro-Meteorological Research, Wallingford OX10 8BB, UK
[3] Swansea University
[4] Leeds University
** Corresponding author: Eleanor Blyth, emb@ceh.ac.uk*

1. Introduction

Climate and weather prediction models need to include a robust and reasonably accurate representation of the land surface. As a minimum, land-surface models need to portray the regional, seasonal variability of the carbon and water stores and fluxes. The land surface model in the UK Met Office Unified Model is JULES (Joint UK Land Environment Simulator, Blyth et al., 2006). JULES is based on the MOSES-TRIFFID model described by Cox et al., (1999) and includes mechanistic formulations of the physical, biophysical, and biochemical processes that control the radiation, heat, water and carbon fluxes in response to hourly conditions of the overlying atmosphere. JULES has integrated coupling of photosynthesis, stomatal conductance, and transpiration so that the biophysical processes in the vegetation interact with hydrological processes in the soil.

A set of benchmark datasets to quantify the performance of the land surface model run at the global scale has been developed which comprises a set of observations that imply the regional, seasonal carbon and water fluxes, a set of metrics to quantify performance and a modelling protocol. This abstract describes the data chosen for this task and the results of the JULES model. The metrics are in the process of being developed.

The first step in defining the benchmark tests is to identify dataset that fulfil the following criteria: they should be independent of any model (i.e. as far as is possible, not a derived data set), globally available, span for several years and at a frequency of at least a month.

2. Datasets chosen

2.1. Fluxnet data for fluxes of Water, Energy and Carbon Dioxide

For this benchmarking exercise, ten Fluxnet sites were selected to sample a range of climate zones (temperate, Mediterranean, tropical and boreal) and plant functional types and soils. Blyth et al (2010) describe this data more fully, covering the use of this data to test the energy (water and heat) fluxes of JULES.

2.2. Mean Monthly Flask CO_2 from 4 stations around the world

Using a simple atmospheric transport model, it is possible to locate the regional land-source contribution of carbon dioxide fluxes for location of the flask-stations. Applying this transport model, the modelled fluxes of carbon dioxide from the land surface can be integrated and transported to the location of the measurement. It is then possible to compare the seasonal variation of carbon dioxide release and uptake by the land surface models compared to the observations. The four stations chosen were: Azores (AZR), Ascension Islands (ASC), Barrow (BRW) and Manoa Loa (MLO), representing both the northern and southern hemispheres.

2.3. Monthly river flow records from 8 major rivers

For this study, a selection of rivers are chosen that represent the same north-south gradient that we have for the Fluxnet data: the Americas are represented by the Mackenzie, Mississippi, Amazon, Piranha and the Europe/Africa longitude by the Lena, Danube, Niger and Congo.

2.4. Monthly NDVI and LAI representing phenology for the 8 river catchments

The normalized difference vegetation index (NDVI) observed from satellite is related to leaf area index. Los et al (2007) have developed a data set of NDVI based on AVHRR and SeaWiFSfor 25 years. To ease comparison with the results of the water-balance information, the tests are made in the areas covered by the 8 rivers in the previous test. The total LAI from the model is scaled with the total NDVI for each catchment.

3. Model Setup and results

For the fluxnet runs, the model forcing was extracted from among the measured variables for each site. For the distributed runs, the JULES model was run with a resolution of 1 degree, using the GSWPII forcing data. The data is available for ten years: 1986-1995. It is necessary to spin up the soil carbon, so the model was run through the data 5 times and the results of the last decade were used in the analysis. Outputs needed to compare to the data are monthly values of carbon dioxide fluxes combined with the atmospheric transport matrices which integrate the fluxes up to the chosen flask stations, monthly values of leaf area index and the surface and sub-surface runoff. A routing model (TRIP) is used to translate the runoff generated at the grid-cell into an equivalent river-flow.

3.1. Model results

Not all the results are complete at the time of writing. The results that are complete are shown here. For the selected FLUXNET sites, Figure 1 shows comparisons between the monthly average modelled evaporation, normalised observed evaporation, Figure 2 shows the comparison of the model with the observed atmospheric CO2, Figure 3 shows the comparison of the river flows and Figure 4 shows the comparison of NDVI and LAI for the 8 river catchments,

4. Results and Identified model weaknesses

A detailed analysis of the performance of the model against all these datasets is not possible or appropriate in this overview paper. More analysis will be forthcoming. However, it is good to note that a full assessment of the combined global carbon and water balance can be made with the current data-holdings. Much progress has already been made in recent JULES development to address the short-comings highlighted here. This series of tests acts as a benchmark of the performance of JULES in its operational mode, and will allow us to compare model performance following model changes with reference to the invaluable sub-daily observations.

It is to be noted that several issues are raised by this study: however, there is no consensus yet of how to set the metrics for model performance and how to define a 'pass mark'.

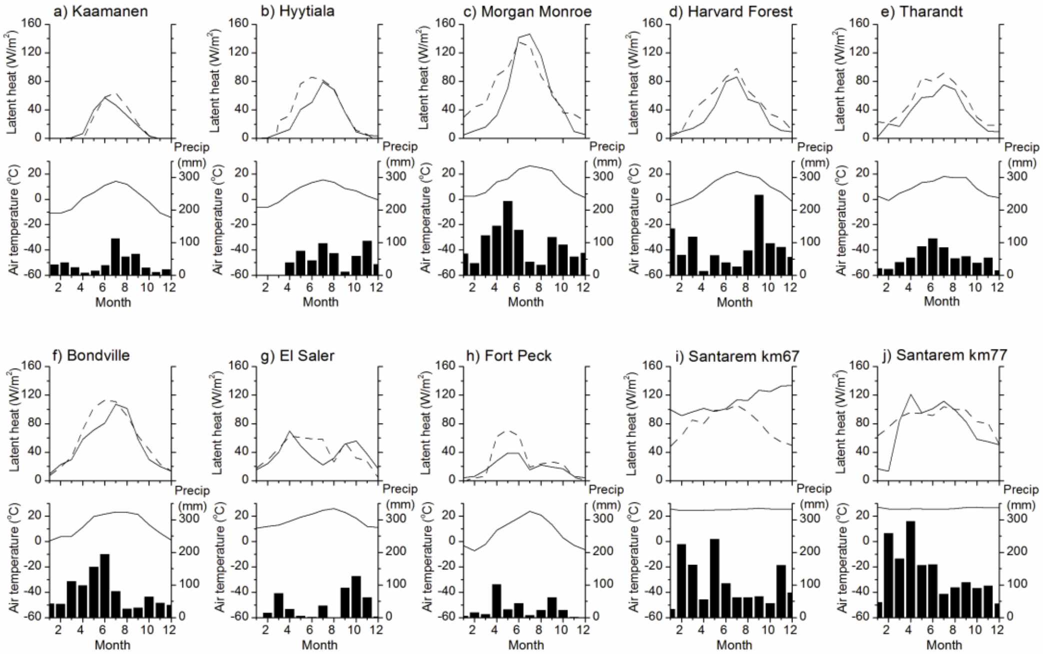

Figure 1: Modelled and observed evaporation fluxes at 10 fluxnet sites.

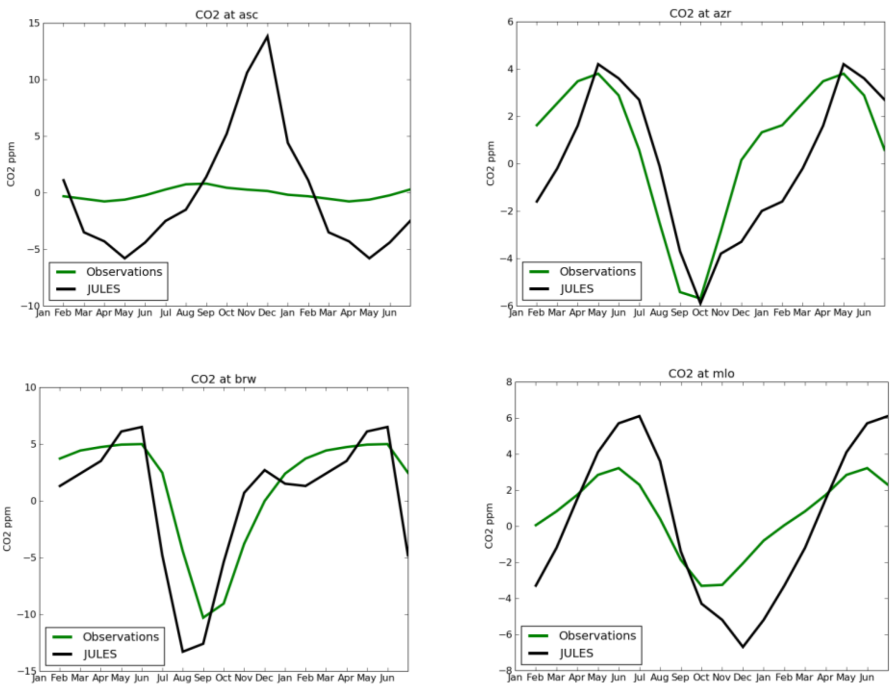

Figure 2: Modelled and observed seasonal variation of atmospheric carbon dioxide at four stations: Ascension (asc), Azores (azr), Barrow 9brw) and Manoa Loa (mlo)

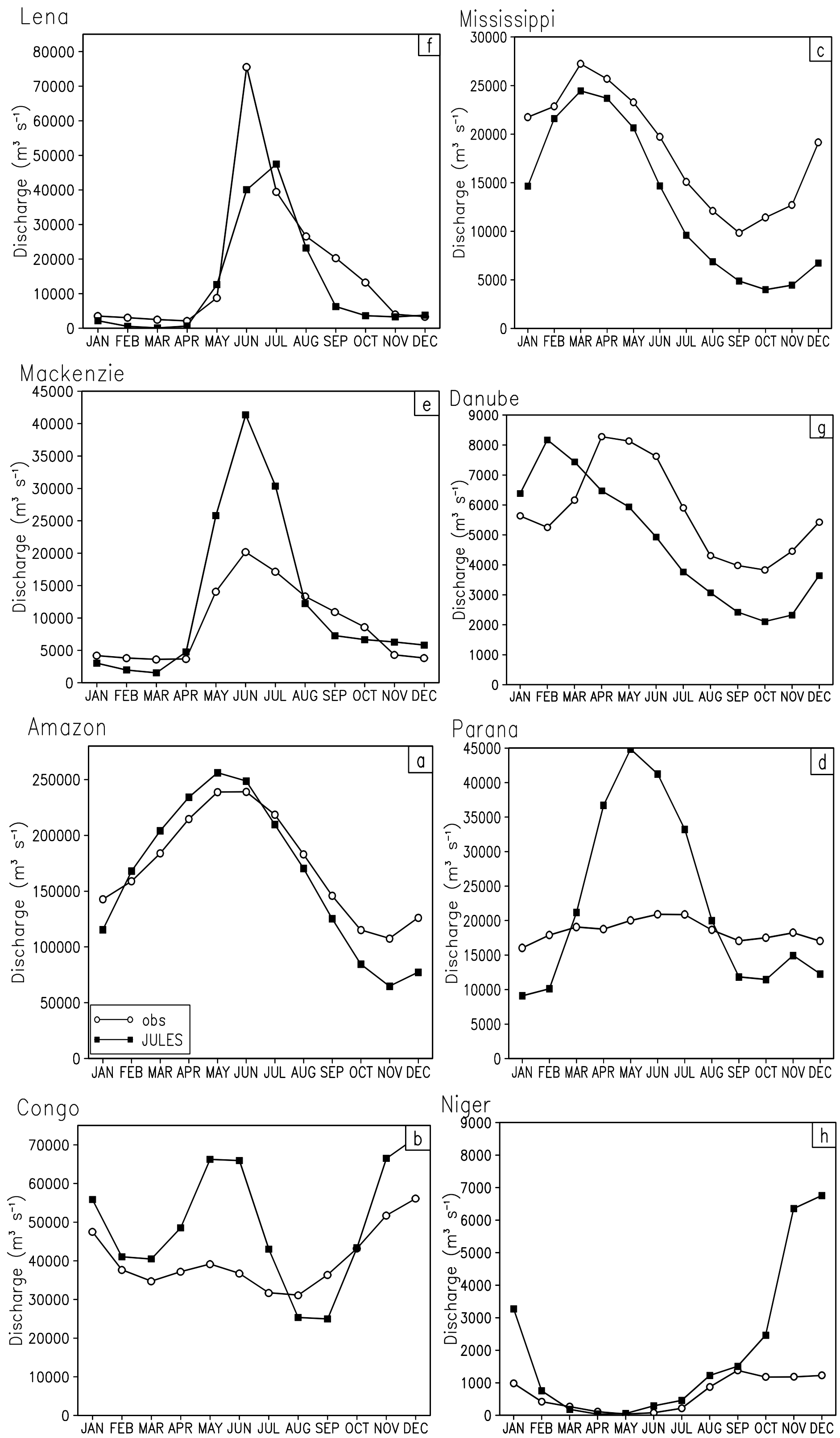

Figure 3: Comparison of modelled and observed seasonal river flow from 8 river catchments

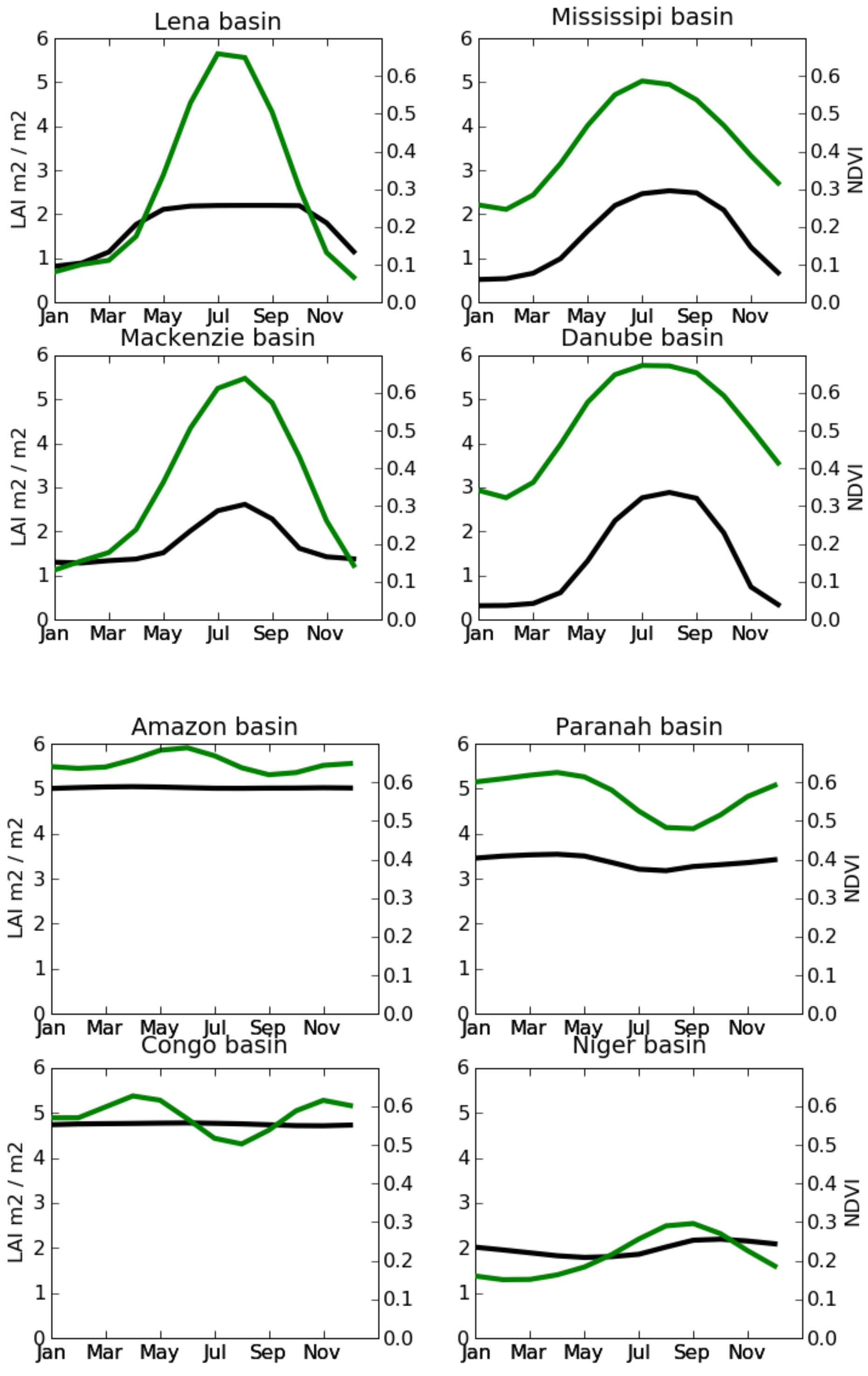

Figure 4: Comparison of modelled (black, LAI) and observed (green, NDVI) phenology from the 8 river catchments.

It is clear from this study that such a large-scale benchmarking system is a useful tool for identifying problems in the model performance. In particular, it allows the relative importance of different problems to be assessed, for instance soil moisture control over the carbon flux, over leaf growth control. But it is not possible to use the data and tests to solve the problem of improving process representation. It is not possible, for instance, to distinguish the reason for the seasonal drop in evaporation in the Santarem 67 fluxnet data, without more information concerning the soil moisture. This benchmarking system is not a replacement for more detailed process-based field data. It is recommended that regional field data is still used for model improvement and calibration.

Acknowledgements

We gratefully acknowledge the FLUXNET PIs at all the sites from which data were taken for this study and the institutions and agencies that support data collection at these sites, including Dr. Laurila of the Finnish Meteorological Institute, Dr. Vesala of the University of Helsinki, Dr. Monroe and Dargoni of the Indiana University, Dr. Wofsy of Harvard University, Dr. Bernhofer of the IHM Technical University of Dresden, Dr. Meyers of NOAA/ARL, Dr. Sanz of the Fundacion CEAM Parque Technogogico, and Dr. Fitzgerrald of the University of Albany.

M.Pryor. was supported by the Joint DECC, Defra and MoD Integrated Climate Programme – DECC/Defra (GA01101), MoD (CBC/2B/0417_Annex C5).

References

Los S.O., North, P.R.J. and Grey W. M. F., 2007, Fused AVHRR SeaWiFS Inetrannual Reanalysis (FASIR) 10-day (3 time mothly) data 1982-2006 (version 5.0) Earth system atlas, http://earthatlas.sr.unh.edu/maps/metadata/Atlas_fasir-ndvi_v5.pdf

Blyth, E. M. Best, P. Cox, R. Essery, O. Boucher, R. Harding, C. Prentice, P.L. Vidale and I. Woodward, 2006. JULES: a new community land surface model. IGBP newsletter

Blyth, E.M, Gash, J.H.C., Lloyd, A., Pryor, M., Weedon, G.P. and Shuttleworth, J.W. 2010. Evaluating the JULES model with Fluxnet data. Journal of Hydrometeorology. In press

Cox, P.M., Betts, R.A., Bunton, C.B., Essery, R.L.H., Rowntree, P.R. Smith, J., 1999. The impact of new land surface physics on the GCM sensitivity of climate and climate sensitivity. Clim Dyn. 15, 183-203.