Geosci. Model Dev. Discuss., 6, C2006–C2010, 2013
www.geosci-model-dev-discuss.net/6/C2006/2013/

Geoscientific Model Development Discussions

Interactive comment on "The North American Carbon Program Multi-scale Synthesis and Terrestrial Model Intercomparison Project – Part 2: Environmental driver data" *by* Y. Wei et al.

D. Schimel (Referee)

david.schimel@jpl.nasa.gov

Received and published: 22 November 2013

Review of Wei et al "The North American .... driver data".

This is a useful paper and in one sense hard to review. While I could comment on specific technical choices used in data set construction, the point is moot since these data sets have been produced and used in (outstanding) follow-on activities. A detailed critique of methods reported here would serve no particular purpose except possible to inform subsequent activities, such as preparing new data sets. In any case, as the authors quote, one shouldn't let the "best be the enemy of good enough". I'll focus mainly on presentation as a result-

To respond to the authors with another quote "those who do not study history are condemned to repeat it". I know this is a typical "old timer" review but this paper follows closely in the footsteps of earlier work without acknowledging that. The first data sets specifically for ecosystem model intercomparison were probably developed first by Wolfgang Cramer. I wouldn't focus on the historical so much except for the concluding section on lessons learned, instead of implying that this is a new development. The lessons learned would be stronger if placed in the context of the evolution of terrestrial carbon and ecosystem modeling: one of those lessons could be to examine earlier studies and see what had worked and what did not!

In particular, this paper does not acknowledge or even cite any of the very similar precursor VEMAP papers. The manuscript paper is strikingly similar in approach to the two VEMAP data set description papers (Kittel et al 1995 and 2004) even to striking parallelisms in the statements of requirements between this paper and Kittel et al 2004. The lessons learned are also remarkably parallel, with Kittel et al speaking to similar needs for formats, informatics etc.

Since eventual ecosystem model validation is enabled by having low driver data error, it is critical that the driver data be first, as good as possible, and second, that potential bias errors and uncertainty be well-documented. The Kittel papers set a high standard for this, in terms of both quality and documentation, and this paper does not rise to that standard, especially for analysis of error and uncertainty. Without that documentation and analysis, it is very hard to separate model error and driver error is subsequent studies. This paper is deficient is identifying the potential nature and magnitude of such problems.

The VEMAP data could even serve as a benchmark for sampling bias, correction for topography etc , as they was prepared from the underlying high density station data using QA/QC, geostatistical interpolation and orographic techniques that remain at or close to the state-of-the-art for the period 1895-1995. The VEMAP climate data record almost certainly has fewer biases than reanalysis or CRU.

The VEMAP papers present extensive validation and cross-validation information so that the contribution of error in the driver data sets can be propagated through to the model output (Fuentes et al 2006). VEMAP requirements were, in their time, drivers for the evolution of THREDDS and its precursor tools, and so that project had a lasting impact on informatics infrastructure. While the VEMAP models had less sophisticated photosynthesis and transpiration algorithms than today, in many ways their nitrogen cycles were more complete than many of today's models.

The VEMAP exercise also resulted in extensive and detailed model error analysis, using techniques for isolating model errors to specific process parameterizations that also would still seem to be relevant (Schimel et al 1997). Note that last paper clearly indicated the need for land use and disturbance data, as many systematic model errors were linked to insufficient resolution of stand agres in regrowing forests and other historical land use effects. Finally, the VEMAP models were used to present an estimate of US carbon flux dependence on CO2 and climate that are the natural benchmark against which to compare both drivers and responses (Pan et al 1998, Schimel et al 2000).

Of the Wei et al concluding points, 1-4 were also directly addressed in VEMAP and are reported in the Kittel papers. With respect to 5, VEMAP made use of early workflow tools and processes to allow reprocessing. 6 is extensively documented in the literature for the VEMAP data sets and in fact, extraordinary efforts were made to harmonize $\sim$8,000 station records and grid them. 7 was utilized and as noted influenced the evolution of NetCDF and THREDDS tools subsequently. VEMAP suffered badly from 8, with accompanying schedule delays!

There are some major advances from the VEMAP era to the MsTMIP data products and these should be specifically highlighted as evolution. These include the advances in models that allow net surface fluxes of CO2 to be simulated, transported and used in model evaluation and the accompanying requirement for CO2, sub-daily meteorology, and the use of a harmonized phenology data set as driver or for validation of prognostic

phenology models, as well as the parallel development of regional and global scaled data (a goal of VEMAP that was never realized).

In a similar time period to VEMAP, there were other terrestrial ecosystem model comparisons, sponsored by the IGBP, including the NPP intercomparison and CCMLP. None of these efforts are referenced at all, suggesting that while the authors have compiled lessons learned for future researchers, they have not heeded their own advice and sought lessons from prior studies.

Kittel, T.G.F., N.A. Rosenbloom, T.H. Painter, D.S. Schimel, and VEMAP Modelling Participants. 1995. The VEMAP integrated database for modeling United States ecosystem/vegetation sensitivity to climate change. Journal of Biogeography 22: 857-862. Kittel, T.G.F., N.A. Rosenbloom, J.A. Royle, C. Daly, W.P. Gibson, H.H. Fisher, P. Thornton, D. Yates, S. Aulenbach, C. Kaufman, R. McKeown, D. Bachelet, D.S. Schimel, and VEMAP2 Participants. 2004. The VEMAP Phase 2 bioclimatic database. I: A gridded historical (20th century) climate dataset for modeling ecosystem dynamics across the conterminous United States. Climate Research 27:151-170. Fuentes, M., T.G.F. Kittel, and D. Nychka. 2006. Sensitivity of ecological models to their climate drivers: Statistical ensembles for forcing. Ecological Applications, 16:99-116 Schimel, D.S., VEMAP Participants and B.H. Braswell. 1997. Continental scale variability in ecosystem processes: Models, data, and the role of disturbance. Ecological Monographs 67: 251-271 Pan, Y.D., J.M. Melillo, A.D. McGuire, D.W. Kicklighter, L.F. Pitelka, K. Hibbard, L.L. Pierce, S.W. Running, D.S. Ojima, W.J. Parton, and D.S. Schimel 1998. Modeled responses of terrestrial ecosystems to elevated atmospheric CO2: A comparison of simulations by the biogeochemistry models of the Vegetation/Ecosystem Modeling and Analysis Project (VEMAP). Oecologia 114:389-404. Schimel, D., J. Melillo, H. Tian, A.D. McGuire, D. Kicklighter, T. Kittel, N. Rosenbloom, S. Running, P. Thornton, D. Ojima, W. Parton, R. Kelly, M. Sykes, R. Neilson, and B. Rizzo. 2000. Contribution of increasing CO2 and climate to carbon storage by ecosystems of the United States. Science 287:2004-2006

Interactive comment on Geosci. Model Dev. Discuss., 6, 5375, 2013.