Environmental Modelling & Software 66 (2015) 110–130

Contents lists available at ScienceDirect

Environmental Modelling & Software

journal homepage: www.elsevier.com/locate/envsoft

Understanding the DayCent model: Calibration, sensitivity, and identifiability through inverse modeling

Magdalena Necpálová [a, *], Robert P. Anex [a], Michael N. Fienen [b], Stephen J. Del Grosso [c], Michael J. Castellano [d], John E. Sawyer [d], Javed Iqbal [d], José L. Pantoja [d], Daniel W. Barker [d]

[a] Department of Biological System Engineering, University of Wisconsin–Madison, Madison, WI, USA
[b] United States Geological Survey, Wisconsin Water Science Center, Middleton, WI, USA
[c] Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO, USA
[d] Dept. of Agronomy, Iowa State University, Ames, IA, USA

ARTICLE INFO

Article history:
Received 11 August 2014
Received in revised form
12 December 2014
Accepted 12 December 2014
Available online 21 January 2015

Keywords:
DayCent model
Inverse modeling
PEST
Sensitivity analysis
Parameter identifiability
Parameter correlations

ABSTRACT

The ability of biogeochemical ecosystem models to represent agro-ecosystems depends on their correct integration with field observations. We report simultaneous calibration of 67 DayCent model parameters using multiple observation types through inverse modeling using the PEST parameter estimation software. Parameter estimation reduced the total sum of weighted squared residuals by 56% and improved model fit to crop productivity, soil carbon, volumetric soil water content, soil temperature, N_2O, and soil NO_3^- compared to the default simulation. Inverse modeling substantially reduced predictive model error relative to the default model for all model predictions, except for soil NO_3^- and NH_4^+. Post-processing analyses provided insights into parameter–observation relationships based on parameter correlations, sensitivity and identifiability. Inverse modeling tools are shown to be a powerful way to systematize and accelerate the process of biogeochemical model interrogation, improving our understanding of model function and the underlying ecosystem biogeochemical processes that they represent.

Software and data availability section

Software DayCent model
Developers W. J. Parton, S. J. Del Grosso, S. Ogle, K. Paustian,
Contact address Natural Resource Ecology Laboratory, Colorado State University, Fort Collins, CO, USA
Telephone (970) 491-2195
Email address century@colostate.edu
Year first available 1998
Hardware required PC with at least 512 K of RAM. A graphics adapter (CGA, EGA, VGA, or Hercules monographic) and 2 Mb of disk space are recommended.
Software required Windows
Availability and cost Available on request; Free
Software PEST version 13.0
Developer John Doherty
Contact address Watermark Numerical Computing, 336 Cliveden Avenue, Corinda 4075, Australia
Telephone 07 3379 1664
Email address johndoherty@ozemail.com.au

Abbreviations: ANPP, aboveground net primary productivity; ARS, Agricultural Research Service; C, carbon; CEC, cation-exchange capacity; CH_4, methane; C/N ratio, carbon to nitrogen ratio; d, index of agreement; DEFAC, decomposition factor; DNDC, denitrification decomposition model; EPA, Environmental Protection Agency; GHG, greenhouse gas; GML, Gauss–Marquardt–Levenberg; NH_4^+, ammonium cation; J, Jacobian matrix; N, nitrogen; N_2O, nitrous oxide; NPP, net primary productivity; NO_3^-, nitrate anion; PEST, parameter estimation software; MB, mean bias; RMSE, root mean square error; rRMSE, relative root mean square error; SOC, soil organic carbon; SOM, soil organic matter; SVD, singular value decomposition; SWSR, sum of weighted squared residuals; VSWC, volumetric soil water content; UAN, urea ammonium nitrate.

* Corresponding author. Sustainable Agroecosystems, ETH Zürich, Tannenstrasse 1, 8092 Zürich, Switzerland. Tel./fax: +41 (0)446338607.
E-mail addresses: magnec@gmail.com (M. Necpálová), anex@wisc.edu (R.P. Anex), mnfienen@usgs.gov (M.N. Fienen), delgro@warnercnr.colostate.edu (S.J. Del Grosso), castelmj@iastate.edu (M.J. Castellano), jsawyer@iastate.edu (J.E. Sawyer), jiqbal@iastate.edu (J. Iqbal).

http://dx.doi.org/10.1016/j.envsoft.2014.12.011

Year first available 1994
Hardware required Desktop or Laptop
Software required Windows or Linux
Availability and cost down load from: http://www.pesthomepage.org/Downloads.php; Free.

1. Introduction

Greenhouse gases (GHG) released from the soils of terrestrial ecosystems are highly variable in space and time due to the interaction of climatic drivers and ecosystem processes involved in carbon (C) and nitrogen (N) transformation associated with production and consumption of GHGs (Müllera et al., 2002; Rahn et al., 2012; Wrage et al., 2001). Field measurements that capture the high temporal and spatial variability of N_2O fluxes (Bouwman et al., 2002; Parkin, 2008; Snyder et al., 2009) or the high spatial variability of soil organic carbon (SOC; Conant and Paustian, 2002; Kravchenko and Robertson, 2011) are expensive, time intensive, and unable to capture the full range of ecological and environmental conditions. When properly informed by field observations, ecosystem process-based models are a powerful way to investigate the effects of management practices on GHG emissions or SOC from different ecosystems, soils, and climates.

A number of biogeochemical models have been developed and used to quantify GHG emissions and SOC at both plot and landscape scales, e.g., Century (Parton et al., 1994; Parton, 1996), DayCent (Del Grosso et al., 2005; Parton et al., 1998), denitrification–decomposition (DNDC) (Li et al., 2000), ecosys (Grant et al., 1993) and EPIC (Wang, 2005). These models are mathematical representations of our understanding of the complicated, coupled biogeochemical soil processes that allow us to test our understanding through comparison of model results with observations, and predict responses to conditions that have not yet been observed, such as ecosystem responses to changing climate. Thus these models have become important tools in the study of biogeochemical cycles. Model development is based on a quantitative understanding of the interactions among physical, chemical and biological processes that is critical for predicting the ecosystem response to land use or climate change. The individual underlying processes are represented by sets of equations in component models that are coupled together to describe a full system (Wallach et al., 2014). Models usually have a mechanistic structure that reflects our understanding of the processes governing the system behavior. Many ecosystem models utilize several hundred parameters representing individual physical quantities or combinations of physical quantities that may not be observable through direct measurement. It is thus impossible to measure the sensitivity of system behavior to each of these parameters and information on their identifiability through field observations is often not available. Yet for model users and particularly model developers, an understanding of how model parameters influence the simulation of target ecosystem processes and which field observations are most useful in defining parameter values is essential.

The DayCent model is a widely used terrestrial biogeochemical process-based model of intermediate complexity (Del Grosso et al., 2001, 2002; Parton et al., 1998). It has been used to simulate ecosystem responses to changes in climate and agricultural management practices in crop, grassland, forest and savanna ecosystems (Brilli et al., 2013; Cheng et al., 2014; Del Grosso et al., 2008a, 2009; Hartman et al., 2009; Parton et al., 2007; Parton and Rasmussen, 1994). In the USA, it has been used to quantify N_2O emissions from agricultural soils for the US National Greenhouse Gas Inventory compiled by the EPA (Olander and Haugen-Kozyra, 2011) and reported annually to the UN Framework Convention on Climate Change (US EPA, 2014). DayCent consists of sub-models for soil water content and temperature by layer, plant production and allocation of net primary production (NPP), decomposition of litter and soil organic matter (SOM), mineralization of nutrients, N gas emissions from nitrification and denitrification, and CH_4 oxidation in unsaturated soils.

The accuracy with which a model represents the natural system observed in the field depends on how completely the underlying biophysical processes are represented in the model and how well the model parameters are calibrated to field observations. Like other biogeochemical process-based models, DayCent is typically calibrated manually by adjusting one parameter at a time, thus the calibrated parameters are adjusted in an iterative fashion in multiple stages (Wallach et al., 2014). At each stage, specific processes are targeted (e.g., plant growth and yield, SOC), and the most influential parameters are adjusted to match simulated to observed values (Del Grosso et al., 2011). This approach, however, does not guarantee full extraction of information from the field observations and it is difficult to know when calibration correctly balances the performance of all model components (Nolan et al., 2011). It is generally accepted that manual calibration of complex ecosystem models does not necessarily yield optimal parameter estimates, is somewhat arbitrary, and results in high uncertainty in model parameters and simulated variables (Schwarz et al., 2006). Inverse modeling, based on an objective statistical method and mathematical techniques for stable parameter estimation, has become a widely accepted way to enhance the transfer of information contained in field observations to model parameters (Doherty, 2003; Doherty and Hunt, 2010a; Hunt et al., 2007). Despite mathematical objectivity, some subjectivity is unavoidable: through defining the conceptualization of the inverse problem and making a set of decisions related to regularization, parameter bounds, observation weighting strategy, etc. (Fienen, 2013). The inverse modeling tool PEST (Doherty, 2010) uses an iterative, nonlinear regression approach that involves simultaneous adjustment of multiple model parameters and evaluation of model fit by the sum of weighted squared residuals between field observations and simulated values. In addition to providing sophisticated estimates of the parameter values that provide the best possible fit for a given calibration problem, inverse modeling provides a method for comprehensive model analysis through statistical measures such as the variance/covariance matrix, parameter correlations, confidence intervals, sensitivities, identifiability, and predictive uncertainty analysis (Moore and Doherty, 2005, 2006).

These tools can help users recognize model problems that are difficult to identify with manual calibration methods (Hill and Tiedeman, 2007; Poeter and Hill, 1997). For example, it has been repeatedly observed that only a small number of the many parameters used in most environmental models are uniquely estimable with most datasets (Beck and Halfon, 1991; Beven and Freer, 2001; Doherty and Hunt, 2009). The inability to uniquely identify certain model parameters can be the result of their high correlation with other parameters, or lack of sensitivity of the model outputs to these parameters. This sort of problem is extremely difficult to recognize without specialized tools and can lead to misidentification of parameter values, model over-fitting, and inaccurate model projections for conditions outside the range of the calibration dataset. Applying inverse modeling tools provides valuable insight about parameter dependencies, which parameters are exerting the most influence on the simulated values, whether the field observations contain enough information to estimate the model parameters, and the uncertainty associated with the predictions based on the estimated parameter values.