Implementation of an User-Friendly Interface of the Forest-BGC Model for Regional/Global-Scale Researches


*Ph.D. Student
****Professor
Grupo de Inteligencia Computacional. Universidad del Pais Vasco (UPV/EHU), Pº Manuel Lardizabal 1, 20018 San Sebastián, SPAIN

**Assistant
Escola Superior Agrária de Viseu/ISPV. Departamento de Produção Vegetal. Quinta da Alagoa, Estrada de Nelas – Ranhados, 3500-606 Viseu, PORTUGAL

***Auxiliar Professor
Departamento de Ciências Florestais. Universidade de Trás-os-Montes e Alto Douro, Apartado 1013, 5001-801 Vila Real, PORTUGAL

Abstract. Forest-BGC is a well-known and widely-used general ecophysiological model of forest ecosystem processes for regional applications. It's used to simulate the annual hydrologic balance and the Net Primary Production (NPP) of forests. The net flux of carbon between the atmosphere and terrestrial vegetation can be expressed on an annual basis in terms of NPP. This way, NPP is a key variable for a wide range of ecological studies, mainly for monitoring climate change impact of natural ecosystems. In order to be prepared for climate changes we need some means to predict how different ecosystems will be affected. The only feasible way of making this estimate on a global scale would be the regular coverage of world vegetation by satellite. Forest-BGC was designed to be driven eventually by remote sensing inputs of surface climate and vegetation structure, allowing its use for regional/global-scale ecological researches. One of the most important variables in Forest-BGC model is the Leaf Area Index (LAI), an annual quantification of the forest structure relevant for energy and mass exchange. LAI is a key simplification for regional/global-scale applications as it's possible to use LAI maps obtained from remote sensing data providers to estimate NPP maps through the Forest-BGC model. This paper presents a new implementation of the Forest-BGC model that allows feeding it with remote sensing input/output data. The user-friendly interface of this implementation eases the use of Forest-BGC model for regional/global ecological researches. This new implementation has been thoroughly validated against the original Coughlan's SIMLAT 5 implementation, obtaining same results and solving the latter's usability limitations.

Key words: FOREST-BGC; ecophysiological production model; automation

Corresponding Author E-mail: maveganzones@gmail.com
Implementação de uma Interface Amigável na Óptica do Utilizador para o Modelo Forest-BGC de Pesquisa a Nível Regional/Global

Sumário. A expressão Forest-BGC (Floresta - Ciclos Bio Geoquímicos) representa um modelo conhecido dos processos ecológicos do ecossistema florestal, utilizado em aplicações regionais. Este modelo usa-se para estimar o balanço hídrico anual e a produção primária líquida (PPL) das florestas. O fluxo de carbono entre a atmosfera e a vegetação terrestre pode expressar-se anualmente em termos de PPL. Deste modo, a PPL é uma variável chave para uma larga maioria de estudos ecológicos, sobretudo para monitorizar o impacto das alterações climáticas nos ecossistemas naturais. A predição do modo e intensidade com que os diferentes ecossistemas serão afectados pelas alterações climáticas, permitirá uma preparação prédia e uma resposta mais adequada ao fenómeno. A única forma de efectuar esta estimativa de forma global seria dispor de cobertura de satélite de toda a vegetação do mundo. O Forest-BGC foi concebido para funcionar com dados de detecção remota de clima e estrutura da vegetação, permitindo a sua utilização em pesquisas ecológicas à escala regional ou global. Uma das variáveis mais importantes do modelo Forest-BGC é o Índice de Área Foliar (IAF), uma quantificação anual da estrutura da floresta relevante para as trocas de massa e energia. IAF é fundamental para simplificar as aplicações à escala regional/global, e é possível utilizar mapas de IAF obtidos a partir de dados de detecção remota, para calcular a PPL através do modelo Forest-BGC. Este trabalho representa uma melhoria do modelo Forest-BGC que permite utilizar dados de detecção remota como input/output. A interface amigável para o utilizador desta aplicação, facilita o uso do modelo Forest-BGC para pesquisa ecológica à escala regional/global. Esta implementação foi totalmente validada através da implementação SIMLAT 5 de Coughlan, obtendo os mesmos resultados e resolvendo as limitações de utilização desta última.

Palavras-chave: FOREST-BGC; modelo ecológico de produção; automatização

Mise en Œuvre d'une Interface Conviviale du Modèle Forest-BGC à l'Échelle Régionale/Mondiale

Résumé. Forest-BGC est bien connu et largement utilisé et est un modèle éco-physiologique des écosystèmes forestiers pour les applications régionales. Il est utilisé pour simuler la production primaire nette (PPN) des forêts. Le flux net de carbone entre l'atmosphère et la végétation terrestre peut être exprimé sur une base annuelle en termes de PPN. De cette façon, la PPN est une variable clé pour un large éventail d'études écologiques, principalement pour le contrôle des incidences des changements climatiques dans les écosystèmes naturels. Pour nous préparer aux changements climatiques, il nous faut un moyen de prédire comment les différents écosystèmes seront affectés. Le seul moyen pratique de faire cette estimation à l'échelle mondiale serait la couverture régulière de la végétation mondiale par satellite. Forest-BGC a été conçu pour être conduit à terme par les apports de la télédétection de climat en surface et la structure de la végétation, ce qui permet son utilisation pour des recherches écologiques à l'échelle régionale / mondiale. L'une des variables les plus importantes dans le modèle Forest-BGC est l’indice de la surface foliaire (ISF), une quantification annuelle de la structure de la forêt pertinente pour l'énergie et l’échange de masse. ISF est une simplification clé pour applications à l'échelle régionale / mondiale et il est possible d'utiliser des cartes ISF obtenues auprès des fournisseurs de données de télédétection pour estimer les cartes NPP à travers le modèle Forest-BGC. Ce document présente une nouvelle implémentation du modèle Forest-BGC qui permet de le nourrir avec l'apport de la télédétection. L'interface conviviale de cette application facilite l'utilisation du modèle Forest-BGC pour des recherches écologiques à l'échelle régionale / mondiale. Cette nouvelle implémentation a été entièrement validée par rapport à l’original SIMLAT 5 de Coughlan, permettant d'obtenir les mêmes résultats et de résoudre les limitations d'utilisation de ce dernier.
Introduction

Many of the most pressing ecological questions being asked today address ecosystem processes that take place at regional and global scales. These questions concern the effects of global climate change, global CO$_2$ increase, regional air pollution, and regional shifts in vegetation cover on forest processes; all require calculations of ecosystem processes at large scales (RUNNING, and COUGHLAN, 1988). In AGREN, et al., 1991 it is asserted that in order to be prepared for climate changes, we need some means to predict how different ecosystems will be affected; and, that the only feasible way of making ecosystem processes estimates on a global scale would be the regular coverage of world vegetation by satellite.

The carbon content of the atmosphere is less than the cumulative input from fossil fuel combustion and any additional sources that result from carbon exchanges between the atmosphere and other reservoirs, particularly, the oceans and terrestrial ecosystems (RUNNING and NEMANI, 1988). Thus, the "greenhouse effect" may lead to significant changes in the Earth's climate if there is a substantial increase in the concentrations of carbon dioxide and other greenhouse gases (EMANUEL and KILLOUGH, 1984; AGREN et al., 1991). The terrestrial biosphere can sequester significant amounts of atmospheric CO$_2$ (FAN et al., 1998; WOFSY et al., 1993), so it is important to reach a good understanding of carbon uptake magnitude and cause. Doing it will require spatial and temporal monitoring of Earth surface processes.

Net Primary Production (NPP), which can be defined as the time integral of positive increments to plant biomass, is becoming a key variable for climate change studies (GOETZ and PRINCE, 1996) and can be laborious, destructive, time-consuming and expensive to measure; making them inappropriate for global ecological researches (WOFSY et al., 1993; CLARK et al., 2001; GOWER et al., 2001; RANDERSON et al., 2002). Ecophysiolog-ical process models are also an option for the estimation of NPP. Most ecosystem process models have been built from and used to simulate the dynamics of small study plots. The seminal papers (RUNNING and COUGHLAN, 1988; RUNNING and GOWER, 1991) present the Forest-BGC model, that calculates key processes involved in the carbon, water and nitrogen cycles for forests, such as NPP.

SIMLAT-5 is the Forest-BGC implementation by Joseph C. Coughlan. It's a MS-DOS program implemented in Turbo Pascal, that runs Forest-BGC model for single-point NPP estimate. It requires daily climate data as well as data from the study site. The model should be run once for each sampling plot, changing only the Leaf Area Index (LAI) value, the ratio of leaf area per unit ground area, for each sampling point and keeping all the remaining data for the study site. From an usability point of view, the major limitation of SIMLAT-5 is the lack of automation in the use of LAI data obtained by remote sensing in order to be able to compute NPP maps.
for large areas. The Forest-BGC implementation reengineering presented in this work overcomes this major drawback, presenting some additional improvements that facilitate the use of Forest-BGC model for regional and global ecological researches.

Thus, the main goal of this paper is to create an interface for the FOREST-BGC which allows its use in a more practical way for studies at different levels of application.

Forest-BGC

The model

The Forest-BGC (Bio Geochemical Cycles) is an ecophysiological model developed by RUNNING and COUGHLAN, 1988; RUNNING and GOWER, 1991, which calculates key processes involved in the carbon, water and nitrogen cycles in forest ecosystems. The model treats canopy interception and evaporation, transpiration, photosynthesis, growth and maintenance respiration, carbon allocation above and below ground, litterfall, decomposition and nitrogen mineralization mechanistically; but, in a general way, incorporating minimal species-specific data.

The model has a mixed-time resolution. Hydrologic, photosynthetic and maintenance respiration processes are computed daily. The other carbon processes and all nitrogen processes are computed yearly (NEMANI and RUNNING, 1989). Figure 1 shows a flow diagram of Forest-BGC model showing the submodels corresponding to daily and yearly time resolutions. The paper RUNNING and COUGHLAN, 1988, presents a complete description of the daily half part of the model, while the dynamic carbon allocation and nitrogen budgets ruling the yearly half part of the model are presented in RUNNING and GOWER, 1991.

Figure 1 - Flow diagram of the Forest-BGC model, and the daily/annual time resolution (from [running_s.w._general_1988])
Hydrologic balances, plant water availability and canopy gas exchange processes are treated daily because meteorological data are routinely summarized as daily averages or totals, and these processes react on a daily basis to environmental conditions (RUNNING and COUGHLAN, 1988). The model requires so, daily climate input data of incident shortwave radiation, maximum and minimum air temperature, dew point and precipitation. Carbon allocation, litterfall and decomposition processes cannot be meaningfully calculated daily because the minimum routine measurement frequency of these processes is each month (RUNNING and COUGHLAN, 1988). Forest-BGC splits time resolution in daily and yearly processes because these time scales seem to be optimal for the model dynamics. Definition of important site and vegetation constants such as soil water capacity and LAI must be provided as inputs to the model. The full model requires the measurement of more than a hundred input variables to work. Forest-BGC was conceived so most of these variables could be injected eventually to the model from remote sensing data acquisition methods.

LAI represents the ratio of leaf area per unit ground area and it’s used by the model for measuring vegetation structure over large areas (for application at regional/global scale), and relating it to energy and mass exchange. It’s also used for calculating canopy interception, transpiration, respiration, photosynthesis, carbon allocation and litterfall (RUNNING and COUGHLAN, 1988; NEMANI and RUNNING, 1989). Most ecosystem process models that simulate carbon and hydrologic cycles require LAI as an input variable (GOWER et al., 1999).

In the Forest-BGC, LAI is calculated by the following equation:

\[
\text{LAI} = \frac{\text{LeafArea} \times \text{LeafCarbon}}{\text{SurfaceArea}}
\]

where Specific Leaf Area, Leaf Carbon and Ground Surface Area are input parameters introduced to the model.

However, according to the same authors, LAI can be easily obtained by remote sensing. As an example, MOD15A2 is the level-4 MODIS global Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation (FPAR) product, composited every 8 days at 1-kilometre resolution by the MODIS instrument operating on the Terra and Aqua NASA’s spacecrafts (KNYAZIKHIN et al., 1999). LOPES, 2005, provides more different ways to estimate LAI values from remote sensing data, i.e. from LandSat ETM+ imagery. According to this, LAI is a key simplification for regional/global applications, targeting Forest-BGC as a clear potential model for the study of forest ecosystem processes at regional and global scales.

The Forest-BGC model provides different outputs, where the more relevant are the Gross Primary Production (GPP), which represents the carbon fixed during photosynthesis; and, the Autotrophic Respiration \( R_a \) (FIELD et al., 1995; GOWER et al., 1999). Net Primary Production (NPP), is the central carbon related variable summarizing the interface between plant and other processes (FIELD et al., 1995), and it can be measured in terms of \( \text{CO}_2 \) exchange as:

\[
NPP = GPP + R_a
\]

when GPP and \( R_a \) have opposite signs.
It is broadly accepted that NPP is a key variable for monitoring climate changes. Species/soil interrelationships, climate and stand age influence NPP of forests. The importance of NPP creates a great need for process models that accurately estimate this value. In a global scale, terrestrial NPP is one of the most modelled ecological parameters, with models that differ markedly in approach and complexity and often yielding comparable estimates. NPP is sensitive to many controls, including aspects of climate, topography, soils, plant and microbial characteristics, disturbance regime and anthropogenic impacts; and these mechanisms operate over different time scales.

A number of global models have been developed, particularly for global carbon, but they are primarily static estimates that do not provide dynamical explanations of the mechanisms underlying the observed processes, and which are not driven by data for specific sites and conditions (RUNNING and COUGHLAN, 1988; GOETZ and PRINCE, 1996). Forest-BGC allows the accurate estimate of NPP values for stands and regional/global areas, on a yearly time scale, providing a mechanistic process that uses minimal species-specific information. Because of this, Forest-BGC can be a very useful tool for climate changing monitoring at global scale (NUNES, 2008).

Simlat-5

Simlat-5 V3.6 is a Forest-BGC implementation by Joseph C. Coughlan (Copyright 1988, J.C. Coughlan). It was implemented in Turbo Pascal V.4 (Copyright 1994-2009 Borland International) and it works on a MS-DOS platform (Copyright 1981-1999 Microsoft Corp). Figure 2 shows the Simlat-5 interface asking for the input files.

Simlat-5 requires two input files to execute a simulation: a climate file with ".clm" extension, and a header file with extension ".dat" including site parameters and constants, and input-output processing control (Figure 3). Two optional files allow any one of the input site parameters to change automatically with a specified year day (files with ".sea" extension) or year number of simulation (files with ".lif" extension). All the files are simple ASCII files and should have the same filename as the header file.

![Figure 2 - Simlat-5 interface](image-url)
Simlat-5 allows the user to specify the desired output file(s) in the header file. Three output files are possible: a daily file with ".day" extension providing daily figures for soil water, outflow, evaporation, transpiration, water potential, photosynthesis and maintenance respiration; an annual file with ".grw" extension providing annual figures for transpiration, photosynthesis, maintenance respiration and leaf, stem and root carbon (Figure 4); and a long file with extension ".lng" providing detailed annual output of any variable specified in the header file.

The major limitation of Simlat-5 is the lack of automation to repeat simulations for each sampling plot. Before each simulation runs on each sampling plot, the header input file must be edited to change the specific plot values, making this approach useless for simulating large areas. The need for this automation together with the use of image data from satellites and/or airbones, as LAI products, has been the main reason to reengineering the implementation of Forest-BGC.
The reengineered implementation of Forest-BGC model is based on the SIMLAT-5 implementation presented in section Simlat-5. A new user friendly interface has been implemented, facilitating the use of Forest-BGC model, the data visualization and the management of results outputs. The new program also permits the feeding of Forest-BGC model by remote sensing data, allowing users the automatic and easy use of Forest-BGC model for regional and global applications. This can be done currently in three different ways:

- As an ASCII file with extension ".lai". This file must include for each line the spatial horizontal and vertical coordinates and a LAI value. The spatial coordinates are used to present the estimate results as visual maps. Figure 5 shows an example of LAI file (".lai").
- As a MODIS MOD-15 product (Knyazikhin et al., 1999). The MOD-15 product is a LAI image obtained by using the MODIS sensors installed in AQUA and TERRA satellites.
- As a LANDSAT ETM+ image. This allows the estimate of LAI maps using mathematical deterministic models (Lopes, 2005).

The code has been developed using common modern Object Oriented (OO) software development paradigm and implemented using Java software language. It is platform independent so it can be used under any operative system running a Java Virtual Machine version 1.5 or superior. The code is freely available and distributed under an open source GPL license.

Figure 6 shows the new user-friendly interface running under a Linux O.S. The interface contains three main areas: the control area (A, B, C and D), the data and results view area (E) and the console area (F). We explain them in more detail at continuation:

- The control area allows the user to load and control the model parameters. It is divided in four subareas:
  - (A) In this control panel the user can load the climatic data, the site data and the specific yearday or season information from text files. The files formats are the same ASCII files used by Simlat-5 program: ".clm", ".dat", ".lif" and ".sea" files.
Figure 5 - LAI ASCII file (.lai) example

Figure 6 - Interface for the Forest-BGC implementation reengineering. (A) File selection control box, (B) LAI data control box, (C) Output configuration control box, (D) Model running control box, (E) Data visualization panels, (F) Console for showing information
(B) This is the LAI regional data control panel. The user can set here different LAI values to automatically run Forest-BGC model for different stands or regional/global areas. It can be done by selecting a LAI file, a MODIS MOD-15 product file or a LANDSAT ETM+ image.

(C) This control panel allows the user to select the way results are reported. The "resume", "day" and "long" reports correspond to the analogous from Simlat-5, but formatted in a way they can be easily exported to other data analysing tools as spreadsheet-applications. The "map" report export results as visual maps that could be imported from common image analysis tools as ENVI or MATLAB. Each report can be configured to present the specific data required by the user and to be exported to different formats.

(D) The last control panel contains some important Forest-BGC model control variables: the start and stop days, the amount of days in a season, and the ratio of carbon fixation respect to NPP estimates.

- The data and results view area allows the easily visualization of input data and model estimates. There is a pane for each kind of input data: climatic data, state variables, site constants, LAI values and ontogeny and phylogeny replacing constants. There is also a pane for each simulation presenting the NPP and carbon fixation estimates and important variables.

- The console area presents system logs to the user such as simulation running feedback, errors, warnings or configuration advices.

Experiments

In order to clarify the use of the proposed Forest-BGC reengineered implementation, we expose here an example of its use to elaborate NPP maps using MOD-15 LAI products. The goal here is to obtain an NPP estimates map of the north of Portugal by mean of the Forest-BGC model, given the climate and site specific data, and feeding the LAI variable by mean of the MOD-15 LAI product for the corresponding location and date, freely available through the NASA's Warehouse Inventory Search Tool (WIST)\(^1\) in HDF5\(^2\) format.

MOD-15 LAI image is a dimensionless ratio (m\(^2\)/m\(^2\)) of leaf area covering a unit of ground area. It is computed on a daily and on a 8-day basis at 1km from MODIS spectral reflectances for all vegetated land surface globally. The MODIS/Terra Leaf Area Index (LAI) product is a value between zero and 8.

The climatic and site specific used data corresponds to the 2002 season. The MOD-15 image we have used corresponds to measures realized between the May 17\(^{th}\) and 24\(^{th}\) of 2002 season, for the area given by the polygon:

- Long: -15.5572 Lat: 50.0
- Long: 0.0216 Lat: 50.0062
- Long: 0.0107 Lat: 39.8191

Figures 7 and 8 show the visualization of the used MOD-15 product and estimated NPP map in false color, with low values being blue and high values being red.

\(^1\) http://wist.echo.nasa.gov/
\(^2\) http://www.hdfgroup.org
Figure 7 - Snapshot of the reengineered Forest-BGC interface showing the MOD-15 LAI product used to estimate the NPP map for the North of Portugal

Figure 8 - Snapshot of the reengineered Forest-BGC interface showing the NPP estimate map for the North of Portugal
Conclusions and discussion

We report the realization of reengineered implementation of the Forest-BCG model which is platform independent, can be fed by remote sensing data and has a user friendly interface that eases its use and the realization of large experimental designs running the simulation at several time and spatial scales. The implementation overcomes basic shortcomings of the previous one which had become obsolete in many respects. Being platform-tied was a serious difficulty, because equipment obsolescence forced to search for emulation solutions which were both uncomfortable and inefficient. The growing availability of remote sensing data, and the growing scope and scale of the research project asked for the easy use of this data in the implementation.

This implementation is being used at the Departamento Florestal for their current research projects. The proposal of this implementation to the community of ecological and agricultural studies has two main objectives. First, to help the reproducibility of results by independent research groups. Reproducibility is a basic requirement for scientific knowledge, and the nowadays trend to simulation and computational experiments imposes the need of a shared implementation where all the research parties can test their hypothesis and show the results in a common framework. We hope that this implementation will perform this kind of role in this community. Second is the need of open source implementations as an additional guarantee of scientific quality. Open source means that the scientific community can examine and criticize the implementation and identify shortcomings and shortcuts. Open source is, therefore, a growing requisite for scientific computational tools to be accepted and validated.

Acknowledgements

We would like to thank PTDC/AGR-CFL/69733/2006 and PTDC/AGR-CFL/68186/2006 for their financial support.

References


