Available online at www.sciencedirect.com
ScienceDirect
Environmental Modelling & Software 23 (2008) 1468–1473
Environmental Modelling & Software
www.elsevier.com/locate/envsoft

Parallelization of a grid-oriented model on the example of a biogenic volatile organic compounds emission model☆

Gerhard Smiatek

Institute for Meteorology and Climate Research (IMK-IFU), Forschungszentrum Karlsruhe, 82467 Garmisch-Partenkirchen, Germany

Received 24 August 2006; received in revised form 28 July 2007; accepted 29 July 2007
Available online 15 October 2007

Abstract

A general fast parallelization approach of a grid-oriented model is described on the example of the semi-empirical biogenic volatile compound (BVOC) emission model (seBVOC). It uses distributed memory parallel (DMP) model relying on the MPI (message passing interface) library. The parallel version achieves nearly linear decrease in execution time as the number of processors is increased. However, with some numbers of processors the efficiency suffers from unequal load balance caused by different number of calculations within the grid cells. Application of variable domain decomposition improves the load balance adding up to 35% additional decrease of processing time.

Keywords: BVOC; Biogenic emissions; MPI; Parallel processing

1. Introduction

Gridded environmental models provide very favorable conditions for parallelization (D'Ambrosio et al., 2007; Giunta et al., 2007; San Hose et al., 2007; Venevsky and Maksyutov, 2007). The modeling area can easily be subdivided into smaller areas. Calculations in those areas are performed in parallel following the single program multiple data (SPMD) principle. Often the modelled processes are also spatially independent and do not require information from neighboring cells, making the parallelization quite simple. On the other hand environmental features vary in space and time. This influences the number of required calculations within the grid cells and thus, can lead to load imbalance and limit the performance of the parallel code.

The semi-empirical biogenic volatile organic compound (BVOC) model seBVOC is a grid-oriented model. It has been developed to assess regional BVOC emissions from the vegetation in high temporal and spatial resolution. Due to their high reactivity VOCs from biogenic sources play an important role in the chemistry of the atmosphere even in densely populated areas where VOCs from anthropogenic sources (AVOC) dominate. Therefore, they must be included in any air quality simulation using numerical chemistry transport models (CTM). The implementation of a parallel version of seBVOC was motivated by the increasing need of highly resolved BVOC emission inventories as well as the substantial requirement of computational power in the uncertainty assessment of the BVOC estimates with a Monte Carlo Study (MCS).

The article describes the seBVOC parallelization in the distributed memory parallel (DMP) approach using the standard message passing interface (MPI) (Snir et al., 1996) version 1.x and standard Fortran 90. Both are known to produce a portable code with good scalability. The article introduces in broad terms the model's background, describes the parallelization approach and discusses the load balance problem.

2. Numerical model

The semi-empirical BVOC emission model (seBVOC) is based on the algorithm presented by Guenther et al. (1993),

☆ Revised version of a paper presented at the OICMS 2005.
E-mail address: gerhard.smiatek@imk.fzk.de

1364-8152/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.envsoft.2007.07.009

Guenther (1997) and Shao et al. (2001) and discussed in detail by Stewart et al. (2003).

In general the emission E (in $\mu g\,m^{-2}\,h^{-1}$) of a distinct BVOC species k from a plant species l can be quantified as:

$$E_{kl} = A_l D_l \epsilon_{kl} \gamma \tag{1}$$

where A_l is the area covered by the plant species l (m^2), D_l is the foliar biomass ($g\,m^{-2}$), ϵ_{kl} is an average emission factor (in $\mu g\,(g\text{-}dw)^{-1}\,h^{-1}$) and γ is an environmental correction factor. It corrects for effects of the actual temperature and solar radiation on the emission as the emission factor is standardized for 30 °C and 1000 $\mu mol\,m^{-2}\,s^{-1}$ photosynthetically active radiation (PAR).

BVOC consists of a large number of compounds. The main focus, however, is on isoprene and a class of monoterpene compounds. Isoprene emissions depend on temperature and solar radiation (synthesis emission) while monoterpene emissions are often temperature driven (pool emission). Monoterpene emissions of some trees, i.e. *Picea abies* (L.) Karst (Norway spruce), depend on both temperature and radiation (Shao et al., 2001). The total emission can therefore be described as:

$$E_{kl} = E_{kl-\mathrm{pool}} + E_{kl-\mathrm{synth}} \tag{2}$$

where $E_{kl-\mathrm{pool}}$ is the pool emission and $E_{kl-\mathrm{synth}}$ is the synthesis emission. For synthesis emission the environmental correction factor is:

$$\gamma_{\mathrm{synth}} = C_{\mathrm{L}} C_{\mathrm{T}} \tag{3}$$

where C_{L} is the factor describing the response of the emission to PAR and C_{T} is the response to the temperature. With empirical coefficients α (=0.0027) and c_{L1} (=1.066) as well as with the PAR flux I ($\mu mol\,m^{-2}\,s^{-1}$) C_{L} is:

$$C_{\mathrm{L}} = \frac{\alpha c_{\mathrm{L1}} I}{\sqrt{1+\alpha^2 I^2}} \tag{4}$$

and C_{T}:

$$C_{\mathrm{T}} = \frac{\exp\frac{c_{\mathrm{T1}}(T-T_{\mathrm{S}})}{RT_{\mathrm{S}}T}}{1+\exp\frac{c_{\mathrm{T2}}(T-T_{\mathrm{M}})}{RT_{\mathrm{S}}T}} \tag{5}$$

where R is the ideal gas constant (=8.314 $J\,K^{-1}\,mol^{-1}$), C_{T1} (=95,000 $J\,mol^{-1}$), C_{T2} (=230,000 $J\,mol^{-1}$), T_{M} (=314 K) are empirical coefficients, T (K) is the leaf temperature, and T_{S} (=303 K) is an empirical normalizing temperature.

The correction factor for the pool emission is:

$$\gamma_{\mathrm{pool}} = \exp(\beta(T - T_{\mathrm{S}})) \tag{6}$$

where T_{S} is the standard temperature (303 K), T is the leaf temperature (K) and β (=0.09 K^{-1}) is a constant.

In seBVOC the light extinction within the canopy is taken into account by dividing the canopy into sunlit and shade fractions following the model presented by de Pury and Farquhar (1997). The irradiance absorbed by the canopy I_{c} ($\mu mol\,m^{-2}\,s^{-2}$) is divided into sunlit I_{cSun} and shaded fraction I_{cSh}. The sunlit fraction is:

$$I_{\mathrm{cSun}} = I_{\mathrm{lbSun}} + I_{\mathrm{ldSun}} + I_{\mathrm{lbsSun}} \tag{7}$$

where lb denotes the direct beam, ld denotes the diffuse irradiance and lbs denotes the scattered beam irradiance. The shaded fraction is calculated as:

$$I_{\mathrm{cSh}} = I_{\mathrm{c}} - I_{\mathrm{cSun}} \tag{8}$$

The Eqs. (7) and (8) are solved twice for visible and NIR (near infrared) part of the irradiance yielding I_{c} and I_{ir}. PAR I is assumed to be $0.45 \times I_{\mathrm{c}}$ (Pinker and Laszlo, 1992).

The leaf temperatures are determined by the canopy energy budget (Dai et al., 2004):

$$C_{\mathrm{c}} \frac{\partial [T_l]_j}{\partial t} = 0 = [I_{\mathrm{c}}]_j + [I_{\mathrm{ir}}]_j - [H_{\mathrm{c}}]_j - L[E_{\mathrm{c}}]_j \tag{9}$$

where I_{c} is the summed net solar radiation absorbed by sunlit/shaded fraction of canopy ($W\,m^{-2}$), I_{ir} is net NIR radiation absorbed by sunlit/shaded fractions of the canopy ($W\,m^{-2}$), H_{c} is the sensible heat flux from foliages to canopy air ($W\,m^{-2}$), and $L[E_{\mathrm{c}}]$ is the latent heat flux from leaves to canopy air ($W\,m^{-2}$). The integration for sunlit ($j=1$) and shaded fractions of the canopy ($j=2$) are denoted by j. C_{c} is the canopy heat capacity ($J\,m^{-2}\,K^{-1}$) which is neglected for the steady state case. The numerical solution for leaf temperatures is calculated by application of the iterative Quasi-Newton–Raphson method.

The BVOC emissions are estimated for the sunlit and shadowed fractions of the canopy and then added. The seBVOC model can be run in a grid mode with mesh specification and meteorology input provided by the Fifth Generation NCAR/Penn State Mesoscale Model (MM5) (Dudhia, 1993). In the gridded approach the domain grid consist of n rows and m columns with a km by a km cell size.

SeBVOC has been coded in Fortran 90. The basic structure is (see Fig. 1):

(1) initialize the 2D, modes and options;
(2) read required tabular input data (land cover, foliar biomass, leaf area index (LAI) and emission factors);
(3) read gridded meteorology data for the term 0 (temperature, wind speed, solar radiation and humidity);
(4) loop over columns, rows and land cover categories and calculate the light extinction (Eq. (7)), leaf temperature (Eq. (9)) and loop over the BVOC compounds and calculate the BVOC emission in each grid cell;
(5) iterate steps 3–4 for all required time terms, time step is 1 h; and
(6) write final 2D grid containing the BVOC emissions.

3. Distributed memory parallelism

In the distributed memory parallel model each processor has its own dedicated computer memory. The data are exchanged between the processes via messages. Messages are sent and