Available online at www.sciencedirect.com

SCIENCE DIRECT

Remote Sensing of Environment 93 (2004) 246–258

Remote Sensing of Environment

www.elsevier.com/locate/rse

Spatial scaling of net primary productivity using subpixel information

Anita Simic[a,*], Jing M. Chen[a], Jane Liu[b], Ferko Csillag[a]

[a]Department of Geography, University of Toronto, 100 St. George Street, Toronto, Ontario, Canada M5S 3G3
[b]Department of Physics, University of Toronto, Toronto, Ontario, Canada M5S 3G3

Received 30 March 2004; received in revised form 5 July 2004; accepted 11 July 2004

Abstract

Spatial scaling is of particular importance in remote sensing applications to terrestrial ecosystems where spatial heterogeneity is the norm. Surface parameters derived at different resolutions can be considerably different even though they are derived using the same algorithms. This article addresses issues related to spatial scaling of net primary productivity (NPP). The main objective is to develop algorithms for spatial scaling of NPP using subpixel information. NPP calculations were performed using the Boreal Ecosystem Productivity Simulator (BEPS). The area of interest is near Fraserdale, Ontario, Canada. It is found from this investigation that lumped (coarse resolution) calculations can be considerably biased (by +14.9% on average) from the distributed (fine resolution) case. Based on these results, algorithms for removing these biases in lumped NPP are developed using subpixel land cover type information. The correlation between the distributed NPP and lumped NPP is improved from r^2=0.16 to r^2=0.59 after the correction. In addition, subpixel leaf area index (LAI) information is used to reduce the remaining biases. After the LAI correction, the correlation is further improved to r^2=0.90.

Keywords: Remote sensing; Net primary productivity; Spatial scaling; Subpixel information

1. Introduction

The concept of scale is a subject to many studies of modern science; however, the term has various meanings in literature. It is of importance to distinguish the term of scale related to data and to ecological processes (Csillag et al., 2000). The data are characterized by both the size of sampling (observational) units and the size of ecological entities (e.g., tree), while the processes depend on size of areas over which the ecological processes operate (Csillag et al., 2000). Remote sensing characterizes the ecological processes at certain spatial resolutions (Curran and Foody, 1994; Waring & Running, 1998). Inappropriate resolutions lead to errors in assessing the ecological processes and, through spatial scaling, these errors may be minimized.

Spatial scaling refers to a process of bridging gaps in quantitative information retrieved at different spatial resolutions. In the process of scaling, information at one spatial resolution deduces characteristics at another (Ehleringer & Field, 1993). Major causes for the errors in a surface parameter retrieved at coarse resolutions are associated with the averaging process of a radiative signal received by a sensor, and algorithms that define the relationship between the averaged radiative signal and the surface parameter. Transferring algorithms from one resolution to another without incurring considerable errors is one of the greatest challenges in remote sensing (Chen, 1999). Nonlinearity of the algorithms and their dependence on land cover types are two major issues in spatial scaling.

Spatial heterogeneity of the land surface introduces major uncertainties in large-scale analysis (Ehleringer & Field, 1993). It can, therefore, affect predictions of ecosystem functioning (Tian et al., 2003). Two types of measures can be used for quantifying spatial heterogeneity: texture that is based on the variability in brightness of pixels in an image (Hu & Islam, 1997), and contexture that is defined by the size and shape of features displayed on an image including the areas, distributions and patterns of the

* Corresponding author. Tel.: +1 416 978 7085; fax: +1 416 946 3886.
E-mail address: a.simic@rogers.com (A. Simic).

0034-4257/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.rse.2004.07.008

features (Chen, 1999). Image texture is often quantified by variance and covariance (Friedl et al., 1995; Hall et al., 1992), while contexture can be simply considered as the fractions of land cover types in a pixel. The later approach considers various land cover types as subpixel information. The contexture-based approach was found to be successful in improving leaf area index (LAI) mapping (Chen, 1999), where the texture-based approach provided just approximations of the scaling effect in the same study. In other studies (Bonan et al., 1993; Friedl, 1996; Rastetter et al., 1992), textural approaches are generally used for characterizing the surface heterogeneity.

The present research is an extension of work by Chen (1999) who recognized the usefulness of contextural information for spatial scaling of LAI. This extension addresses spatial scaling of net primary productivity (NPP), determined not only by remotely sensed parameters but also meteorological and soil conditions. NPP is defined as the net amount of carbon stock increase in living plant tissues per unit time and space, and, therefore, it is the foremost quantifier and predictor of the atmospheric carbon budget regionally and globally. Recent studies involve mapping NPP of large areas to understand the behavior of the boreal ecosystem, which makes up a considerable portion of the global terrestrial carbon stock (Solomon & Shugart, 1993). Terrestrial biosphere dynamics plays an important role in global climate and remote sensing can provide useful data for quantifying the dynamics.

Spatial scaling of NPP is necessary for two reasons: (1) the nonlinearity in the relationship between NPP and remotely sensed parameters such as LAI and (2) pixels at coarse resolution often contain more than one cover type. The focus of this study is on the effects of mixed cover types and the nonlinearity on spatial scaling. Algorithms for scaling from fine (30-m) to coarse (1-km) resolution, which consider land cover type heterogeneity as subpixel information, are developed and discussed. In addition, subpixel LAI information is used to further improve the algorithms. A Landsat TM image was used for this purpose. Fine scale models provide the best tool in validation of global models (Reich et al., 1999) since the collection of ground-based data for large areas is impractical, time-consuming, and often impossible. The area studied as an example is near Fraserdale, Ontario. Both NPP calculations at 30- and 1-km resolutions were performed using the Boreal Ecosystem Productivity Simulator (BEPS), a process model for estimating the carbon budget of terrestrial ecosystems (Liu et al., 1997).

2. Theory

2.1. Spatial scaling by contextural approach

A contextural approach of spatial scaling uses area fractions to derive surface parameters at different resolutions. Remote sensing generally identifies only a dominant land cover type in a pixel and ignores the impacts of other land cover types in the pixel on derived surface parameters. The fact that only one cover type is labeled per pixel at a coarse resolution results in major uncertainties of the final products (Chen et al., 2002). Ideally, the surface parameters at a coarse resolution should be calculated by averaging the same parameter from the fine resolution pixels. These values and algorithms are often considered to be ‘distributed’. However, remote sensing values generated at coarse resolutions are often the only available data. They are often named ‘lumped’ data and algorithms (Chen, 1999; Hu & Islam, 1997). Tian et al. (2002) discussed the mathematical approach of these two methods in calculating of LAI.

Fig. 1 schematically represents the distributed NPP (NPP_D) and lumped NPP (NPP_L) calculations for one coarse mixed pixel. The fine resolution calculation closely represents the reality, and the distributed NPP is assumed to be correct (distributed case A) in our scaling studies. The lumped case (B) is an actual approach simulating the values retrieved from coarse resolution satellite sensors such as AVHRR. We assume that a coarse pixel (1×1 km) in the figure contains 1089 fine pixels (30×30 m), which consist of different land cover types. In case A, unique algorithms are applied to each land cover type for both LAI and NPP calculations before the arithmetic averaging. Four sets of algorithms are needed for four cover types in this study. In case B, the averaging is performed before the application of the algorithms. The coarse pixel is labeled as one (dominant) land cover type and only one set of algorithms specific to that cover type is applied. Ignoring other land cover types and the related algorithms results in considerable errors in NPP_L. Ignoring the subpixel land cover type information is prone to either positive or negative biases. This depends exclusively on the type of dominant cover type assigned as a unique label to a lumped pixel. The differences between true parameter values, derived at the fine resolution, and parameter values derived at the coarse resolution are functions of subpixel cover type area fractions (Chen, 1999). The process of aggregation may maintain the mean values; however, the variance of NPP is considerably reduced when input data are averaged (Band & Moore, 1995). Chen et al. (2002) developed a methodology to use known cover type area fractions within a pixel in order to remove the biases in LAI calculations.

2.2. Algorithms for spatial scaling using land cover type area fractions and LAI subpixel information

The corrections of lumped NPP_L for a pixel labeled as cover type j (NPP_{Lj}) are based on the regression coefficients C_{ij} retrieved by correlating the correction factor R_j with each nondominant cover type fraction F_{ij} within the uniquely labeled pixel. Let R_j be the ratio of known NPP_{Dj}

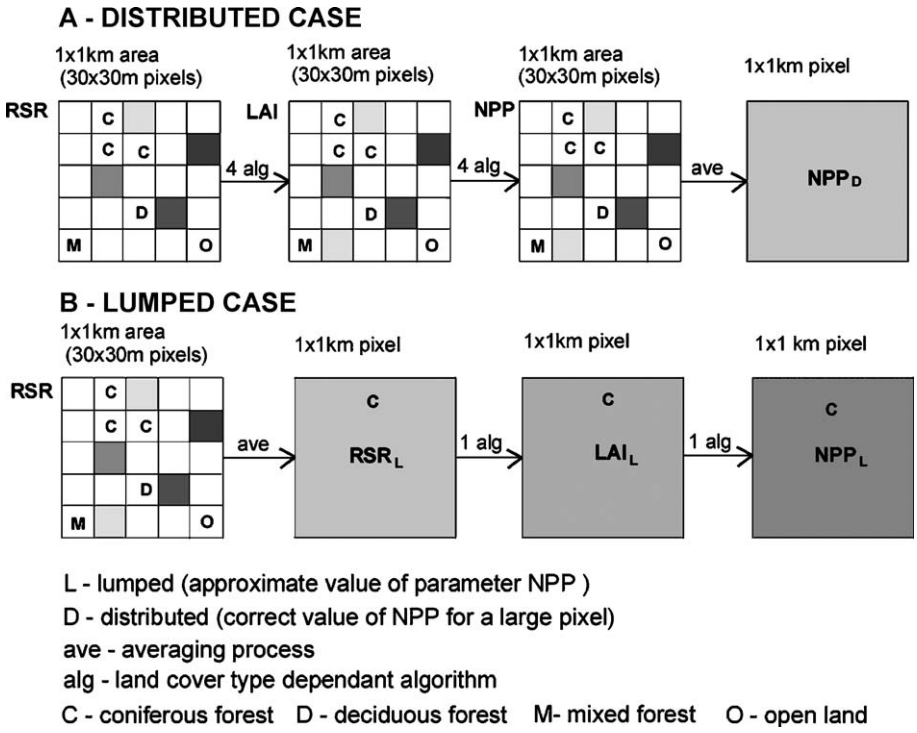

Fig. 1. Schematical (simplified) representation of the distributed NPP and lumped NPP calculations for one coarse pixel.

and NPP_{Lj} values. The relationships are theoretically developed as follows:

$$NPP_{Dj} = NPP_{Lj} * R_j \tag{1}$$

$$R_j = 1 - \sum_{i=1}^{n} C_{ij} F_{ij} \tag{2}$$

where F_{ij} is the fraction of a nondominant cover type i in a pixel labeled as cover type j; C_{ij} is a regression coefficient for a particular nondominant cover type i in a pixel labeled as cover type j; n is the number of nondominant types within the same pixel.

For each land cover type, the correction coefficient R_j can be plotted against all nondominant cover type fractions simultaneously and against each nondominant type separately. The later method, which involves multiple simple linear regression statistical approach, is used to derive the algorithms. To illustrate, in the case of a coniferous forest-designated pixel (j), the regression coefficients C_{ij} from the relationship between R_j and other than conifers fractions F_{ij} within the pixel are used in Eq. (2).

Based on the above relationships and known fractions of nondominant land cover types within each pixel, the following algorithm for correcting NPP_{Lj} can be developed:

$$NPP_{Lj\ \text{corrected}} = NPP_{Lj} * \left(1 - \sum_{i=1}^{n} C_{ij} F_{ij}\right) \tag{3}$$

Using the appropriate correction coefficients, the algorithm is applied to each land cover type (conifer-, deciduous-, mixed-, and open land-labeled pixels).

We found that the regression coefficients C_{ij} in Eqs. (2) and (3), determined through multiple simple linear regressions, are similar to the same coefficients determined through one-step multivariate regression analysis, suggesting that either multiple simple linear regression approach or multivariate regression methods could be used. One important assumption of the multiple simple linear regressions is that each regressors (nondominant fractions) must be linearly independent. When the method of multiple simple regression is used, the slopes in the regressions are of interest in the scaling process while the intercepts, which includes the effects of other cover types, could be ignored, e.g., in the case of $R_j = 1 - C_{1j}F_1 - C_{2j}F_{2j}$, the intercept in the $R_j - F_{2j}$ regression is $1 - C_{1j}F_{1j}$.

For the same cover type, LAI can vary greatly and cause errors in lumped NPP calculations because of the nonlinearity in the relationship between LAI and NPP. To correct for this nonlinearity effect, an additional correction is developed based on the relative deviation of LAI values of a dominant land cover type from the mean LAI value of a lumped pixel:

$$NPP_{Lj\ \text{corrected}} = NPP_{Lj} * \left(1 - \sum_{i=1}^{n} C_{ij} F_{ij} L_{ij}\right) * f(L_j) \tag{4}$$

$$L_{ij} = (LAI_i - LAI_L)/LAI_L \tag{5}$$

$$L_j = (LAI_j - LAI_L)/LAI_L \tag{6}$$

where LAI_i is the average of distributed LAI values for a nondominant cover type i within a lumped pixel; LAI_j is the average of distributed LAI value for a dominant cover type j; and LAI_L is the mean LAI value of a lumped pixel. In Eq. (4), the correction on the lumped NPP for LAI variability is done in two steps. One is a linear correction as an adjustment to the nondominant cover type fraction with an assumption that the NPP–LAI relationship is linear. The other considers the effect of the nonlinearity, which is based on the NPP–LAI relationship calculated for a dominant land cover type and includes a nonlinearity factor for the given relationship. The later step is expressed as:

$$f(L_j) = 1 + \alpha L_j \tag{7}$$

where α is a nonlinearity factor within a dominant land cover type. If the LAI–NPP relationships are linear, α is zero. It should generally be between 0 and 1.

3. Data processing and modeling

Approximately a 50×50-km boreal forest study area located near Fraserdale, Ontario, Canada was chosen in this study. The area mostly consists of conifer forest (59.4%) of high and medium crown density and open land (20.3%). In addition, some parts of the area are covered with mixed forest (12.3%), and to lesser extent with deciduous forest (4.5%) of low crown density. The open land consists of mostly regenerating areas that includes seedlings and fast-growing grass and shrubs as documented by aerial photographs. It also includes recently burned areas, barren soil/rock and wetlands.

The NPP calculations by BEPS comprise five steps (Liu et al., 1997). Soil water balance estimation in step 1 includes the calculations of rainfall input, snowmelt, canopy interception, evapotranspiration, and overflow. Soil water content is used in step 2 to calculate the mesophyll conductance and the stomatal conductance for sunlit and shaded leaves. In step 3, the daily gross photosynthesis and daily autotrophic respiration of leaves, stems, and roots are computed. Daily NPP was calculated in step 4 as the difference between daily gross photosynthesis and daily autotrophic respiration. The final step involves the derivation of annual NPP (Liu et al., 1997). BEPS input data consist of land cover, LAI, biomass, soil, and daily meteorological data.

The remote sensing data include surface reflectance images for three TM bands (R, NIR, and SWIR) for Fraserdale scene acquired on 25 August 1998. LAI and land cover images of 30 m resolution were generated from the TM images by CCRS (Chen et al., 2002). The LAI algorithms were developed by Chen et al. (2002) and are based on the reduced simple ratio (RSR). The RSR is a vegetation index based on three bands (red, NIR, and SWIR), developed by Brown et al. (2000) after Nemani et al. (1993) who made a similar modification to NDVI, and is defined as follows (Brown et al., 2000):

$$RSR = \frac{\rho_{\mathrm{NIR}}}{\rho_{\mathrm{red}}} * \left[1 - \frac{\rho_{\mathrm{SWIR}} - \rho_{\mathrm{SWIR}_{\mathrm{min}}}}{\rho_{\mathrm{SWIR}_{\mathrm{max}}} - \rho_{\mathrm{SWIR}_{\mathrm{min}}}}\right]$$

where ρ_{NIR}, ρ_{red}, and ρ_{SWIR} are the reflectances in NIR, red, and SWIR bands, respectively; $\rho_{\mathrm{SWIR}_{\mathrm{max}}}$ and $\rho_{\mathrm{SWIR}_{\mathrm{min}}}$ are the maximum and minimum SWIR reflectances determined from the images as the 1% low and high cutoff points in the histogram.

The LAI algorithms are cover type-dependant. The relationship between LAI and RSR is based on linear regression for coniferous forest and open land (cropland, grassland, tundra, barren), and nonlinear regression for deciduous and mixed forests (Chen et al., 2002). In addition to the LAI image, which represents the growing season, a LAI winter image, for the non-growing season, was created and incorporated in the calculations. Biomass calculations were based on the LAI values and algorithms developed by Liu et al. (2002). Available water-holding capacity of soil (AWC) and meteorological data were generated from the regional data of Canada (Liu et al., 1997). AWC is defined as the volume of water that is available to plants when the soil is at the field capacity. Since the meteorological data do not vary considerably (1% variations for air temperature, humidity, and radiation, and 6% for precipitation) across the study area and the topography is generally flat, the meteorological variables were assumed to be uniform across the site.

Methods for NPP spatial scaling were investigated on NPP images of 1 km resolution derived in two ways: (A) from distributed calculations (NPP_D), where NPP was calculated first at 30-m resolution and resampled (i.e. averaged) to 1-km resolution and (B) from lumped calculations (NPP_L), using input maps produced at 1-km resolution. The study area consists of 1650×1650 pixels at 30-m resolution and 50×50 pixels at the coarse resolution (≈1 km). All calculations were performed for a calendar year of 1998. The same algorithms were used for both the distributed and lumped calculations.

Step A. BEPS was run for 30-m resolution with daily meteorological, seasonal LAI and biomass, annual land cover and long-term AWC input data. The generated NPP image was further resampled to 990-m~1-km resolution. Arithmetic averaging was performed on the NPP map at the 30-m resolution to generate NPP_D.

Step B. The NPP_L map was based on the input maps produced at 1-km resolution. The lumped LAI calculations were based on the lumped values of RSR, which was generated by averaging RSR of fine resolution (30 m).

The land cover type for the study area was resampled by most dominant cover type into a 1×1-km resolution image.

The classes within the original land cover map (12 classes) were regrouped into four basic land cover types (coniferous, deciduous, and mixed forest, open land) and water. The generated NPP_L map was at 1-km resolution.

Simple linear regression analysis was used to investigate the relationship between NPP_D and NPP_L for all land cover types and for each land cover type separately. Since NPP_D has been assumed to represent the reality, any deviation of NPP_L from NPP_D is taken as an indication of the error produced at the coarse resolution modeling. Cover type area fractions within each 1-km pixel were calculated based on known cover type subpixel information from the 30-m resolution data. This sequence of the statistical analysis is the most vital in the scaling analysis, since the scaling algorithms could not be derived without knowledge of cover type area fractions. Since water percentages within the coarse pixels were generally low (3.4% on average), the impact of water was excluded from detailed analysis. Algorithms for correcting the NPP_L developed in this study (Eqs. (1)–(7)) were applied to NPP_L.

4. Results

Fig. 2a illustrates the distribution of NPP over the study area in 1998 calculated at fine resolution (30 m). Forest, which occupies most of the area (Fig. 2b), is most productive. Mixed and deciduous forests demonstrate slightly higher photosynthetic rates than conifers.

Maps of NPP_D and NPP_L over the study area are shown in Fig. 2c and d, respectively. Overall, NPP_L values are higher than NPP_D values. The NPP_L image appears to be more variable than the NPP_D image, particularly within open land.

4.1. Statistical characteristics of distributed and lumped NPP

Table 1 exemplifies the overall input (RSR, LAI, LAI winter, and biomass) and output data (NPP at 30-m resolution, NPP_D and NPP_L) statistics calculated at the fine and coarse resolutions.

The overall mean NPP_D for the entire study area is 212 g C m^{-2} $year^{-1}$ (Table 1). This value falls in the lower range

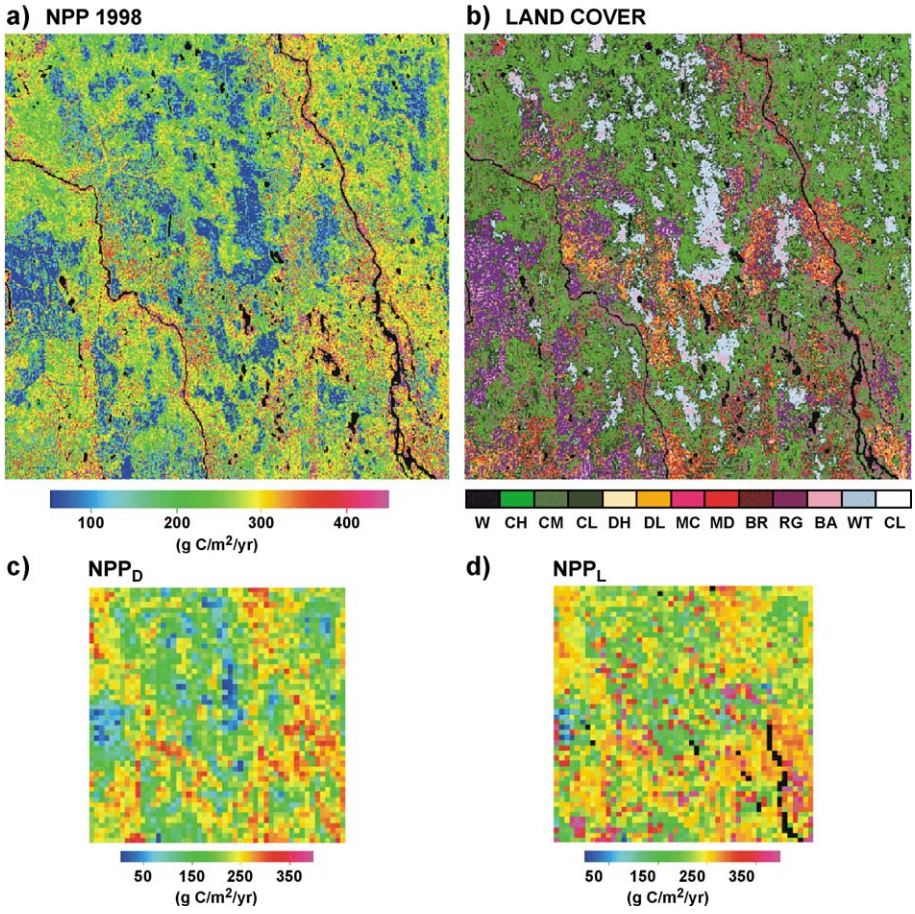

Fig. 2. (a) Net primary productivity map (30-m resolution). (b) Land cover map (CCRS, 1998). W denotes water; CH, CM and CL, conifers (high, medium and low crown density, respectively); DH and DL, deciduous forest (high and low crown density, respectively); MC and MD, mixed forest (coniferous and deciduous, respectively); BR, recent burns; RG, regeneration; BA, barren soil and rocks; WT, wetland; CL, clouds. (c) Spatial distribution of distributed NPP. (d) Spatial distribution of lumped NPP. Note: black color on the figures denotes water.

Table 1
Statistics of the RSR, LAI (for growing and non-growing season), biomass (BM), and NPP images of fine and coarse resolutions for entire study area

Parameters	Units	Mean	Min	Max	S.D.
Fine resolution					
RSR—30 m	–	4.1	0	25.5	2
LAI—30 m	–	2.9	0	10	1.7
LAI winter—30 m	–	2.5	0	10	1.9
BM—30 m	tons ha^{-1}	40	0	284.8	31.8
NPP—30 m	g C m^{-2} $year^{-1}$	213.5	0	740	117
NPP_D[a]—1 km	g C m^{-2} $year^{-1}$	211.77	10	360	62.62
Coarse resolution					
RSR—1 km	–	4.07	0.2	8.1	1.28
LAI—1 km	–	3.1	0	5.9	1
LAI winter—1 km	–	2.7	0	5.7	1.6
BM—1 km	tons ha^{-1}	38	0	103.4	22.9
NPP_L[a]—1 km	g C m^{-2} $year^{-1}$	243.32	15	610	61.67

NPP_D denotes distributed NPP for 1 km based on averages of 30-m resolution NPP values. NPP_L denotes lumped NPP.
[a] Water fraction excluded.

of NPP values for the globe (200–600 g C m^{-2} $year^{-1}$) summarized by Liu et al. (1997). Mixed forest is found to be the most productive with the mean value of 275 g C m^{-2} $year^{-1}$, and it is closely followed by deciduous (254 g C m^{-2} $year^{-1}$) and coniferous forest (227 g C m^{-2} $year^{-1}$). The mean value of conifers is similar to that reported by Liu et al. (1999). Due to the high latitude of the study area, deciduous and mixed forests in this study have much lower NPP_D values than the same cover types for the whole Canada, as found by Liu et al. (1999). Similarly, deciduous and mixed forests in the study area are unproductive by far compared with findings of Running and Coughlan (1988) and Reich et al. (1999) for temperate forests. A relatively high mean NPP_D (130 g C m^{-2} $year^{-1}$) and high range and variability of open land indicate that fast-growing grass and shrubs within the open land area are very productive.

The mean NPP_L for the study area was found to be 243 g C m^{-2} $year^{-1}$ (Table 1). The difference between the mean NPP_D and mean NPP_L is mostly due to the open land area, which has a mean NPP_L of 281 g C m^{-2} $year^{-1}$. A minor increase between the mean NPP_D and mean NPP_L is seen within coniferous and mixed forest. The mixed forest and open land are highly variable in the lumped calculations.

Generally, the mean values of the corresponding datasets shown in Table 1 are very similar, and the variability is reduced with decreasing resolution, as suggested by Tian et al. (2002). Different from other datasets, the mean LAI at the coarse resolution and NPP_L are slightly higher than the equivalent datasets at the fine resolution. Turner et al. (1996) found that although the mean values are generally similar, some variables result in reductions of the mean values with increasing resolution, and some behave differently (e.g., LAI). The ranges are reduced for all datasets at the coarse resolution, except for NPP_L, which, however, has a range smaller than that of NPP calculated at 30-m resolution, before the resampling to NPP_D.

4.2. Relationship of distributed and lumped NPP

It is found from this investigation that the lumped NPP values compare poorly with the distributed values (r^2=0.16) (Fig. 3), although the mean NPP values averaged for the whole area are reasonably close. For the total area, the lumped NPP is biased by +14.9%. There are about 10% pixels with a positive bias of larger than 100%. Fig. 4 shows the relationship between NPP_D and NPP_L for each cover type separately. The strongest relationship and similar ranges are observed within conifer-labeled pixels (r^2=0.84). Deciduous-labeled pixels also have a relatively strong relationship (r^2=0.68). More variability is found within mixed forest and open land with r^2=0.42 and r^2=0.47, respectively. The open land-labeled pixels show positive biases when they contain substantial percentage of conifer forests. This indicates the invasive characteristics of conifer species, and the open land as relatively old disturbed areas. High variability in NPP values, particularly for open land, can be related to the unique reflectance and, ultimately, different LAI values of conifer forests when combined with non-forest cover types (Tian et al., 2003). The bias is negative when conifer-labeled pixels (at 1-km resolution) contain considerable deciduous forests. Due to relatively high and variable NPP values of open land areas, and unproductive deciduous forest, the bias is often negative when deciduous-labeled pixels are mixed with open land.

As it is commonly assumed that open land is of low productivity, these trends are unforeseen, suggesting a different case from that studied by Chen et al. (2002). The overall relationship between NPP_D and NPP_L in this study is found to be weaker than the relationship between BEPS-derived and measured NPP values in the study of Liu et al. (1997) (r^2=0.40). This suggests the difficulty in using

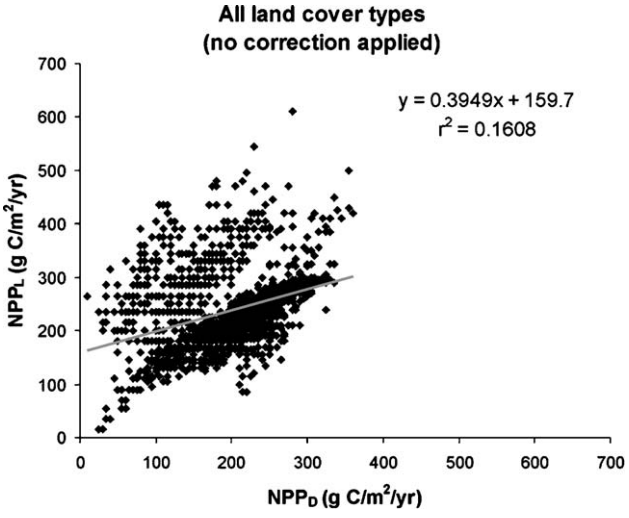

Fig. 3. Relationship between distributed NPP and lumped NPP for all land cover types before the correction.

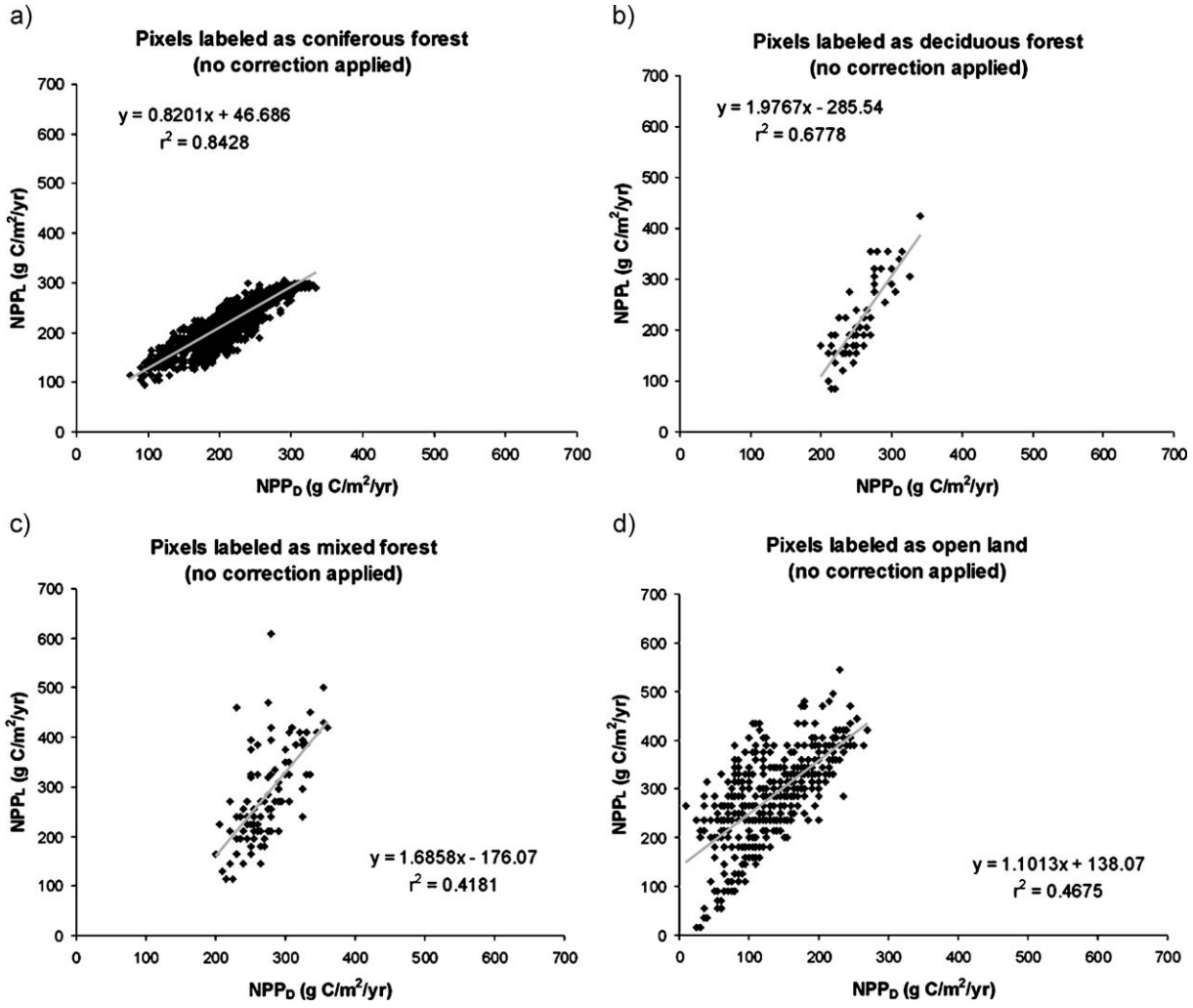

Fig. 4. Relationship between distributed NPP and lumped NPP before the correction for pixels designated as (a) coniferous forest, (b) deciduous forest, (c) mixed forest, and (d) open land.

ground NPP data for coarse resolution NPP map validations without rigorous spatial scaling, particularly within areas that have a high NPP within-class variability. Other studies also suggested bias of the productivity when calculated at coarser resolutions (Pierce & Running, 1995; Reich et al., 1999; Waring & Running, 1998).

The reason for the high productivity and NPP variability of the open land pixels could be investigated by exploring the reflectance, RSR and LAI values. In lieu of this need, we explored the LAI values for each land cover type derived from the fine and coarse resolutions (Table 2). The mean LAI at the coarse resolution, derived from lumped RSR values, is smaller than that at the fine resolution for all land cover types with the exception of open land, which is not in accordance with findings of Chen (1999) and Tian et al. (2002). Variations within LAI values are commonly affected by the changes in open land, which acts as the background and consists of different vegetation but mostly fast-growing grass and shrubs (Chen et al., 2002). Turner et al. (1996) found that shrubs significantly contributed to total LAI in the NPP simulations in one of their study areas, resulting in much higher NPP than in other forest areas. In this study, the difference in LAI between the lumped and distributed cases for open land contributes to the large difference in NPP between these two cases. In principle, one could separate the productive subclasses from the open land area according to the LAI value and possibly reduce the NPP estimation error. As NPP is calculated in the same way in BEPS for the same cover type regardless of the LAI values, it is theoretically

Table 2
Statistical values of LAI for each land cover type derived from fine and coarse resolutions

Variables	Statistics	Conifers	Deciduous forest	Mixed forest	Open land
LAI (fine resolution)	Mean	3.7	2.6	3	0.6
	S.D.	0.6	0.9	0.7	0.5
	Min	1.3	0.5	0.6	0
	Max	6.3	10	4.9	3
LAI (coarse resolution)	Mean	3.6	1.9	2.5	1.6
	S.D.	0.7	0.4	0.7	0.5
	Min	1.6	1.1	1.3	0
	Max	5.7	3.2	5.9	3.6

unnecessary to create these subclasses. This also allows us to demonstrate the importance of LAI variations within same cover type in the scaling process.

4.3. Correction of lumped NPP using land cover type area fractions

The relationships between the correction factor R and cover type area fractions demonstrate how NPP_L diverges from NPP_D in dependence of nondominant fractions. For conifer-labeled pixels, it is evident that the factor becomes more variable with increasing non-coniferous fractions, i.e. NPP_L diverges more from NPP_D with increasing heterogeneity (Fig. 5a). Differences between NPP_D and NPP_L of conifer pixels are less variable with increasing both deciduous (F_{DEC}) (Fig. 5b) and mixed forest area fractions (F_{MIX}) (Fig. 5c) than in the case for open land fraction (F_{OL}) (Fig. 5d). This suggests that open land fraction accounts for the variability seen in Fig. 5a. Similar to the conifer-labeled pixels, the greatest portion of the variability in R can be explained by high open land fractions within the deciduous-labeled pixels and by high conifer area fractions and/or open land area fractions within mixed-labeled pixels (not shown). For open land-labeled pixels, the variability in R decreases with the increase in non-open land fractions (Fig. 6a). The variability in R decreases with the increase in coniferous (F_{CON}) (Fig. 6b), deciduous (Fig. 6c), and mixed forest (Fig. 6d) fractions. Regression coefficients used in the algorithms for scaling in this study are listed in Table 3.

Overall, the variability in R is positively correlated to open land fractions within each land cover type. This is most apparent in Fig. 6a in particular. The trend of the relationship between R and all nondominant fractions indicates that R does not converge to 1 as the open land fraction increases toward 100%. This is because non-dominant cover types (mostly conifer and mixed) generally have considerably higher LAI than that in grass (open land). When the forest LAI is treated as grass LAI in the lumped calculations, the NPP of the lumped pixel can be positively biased. For the same LAI, NPP for grass is much higher than that in forest for the following reasons: (1) maintenance respiration in grass with small biomass is lower than that in

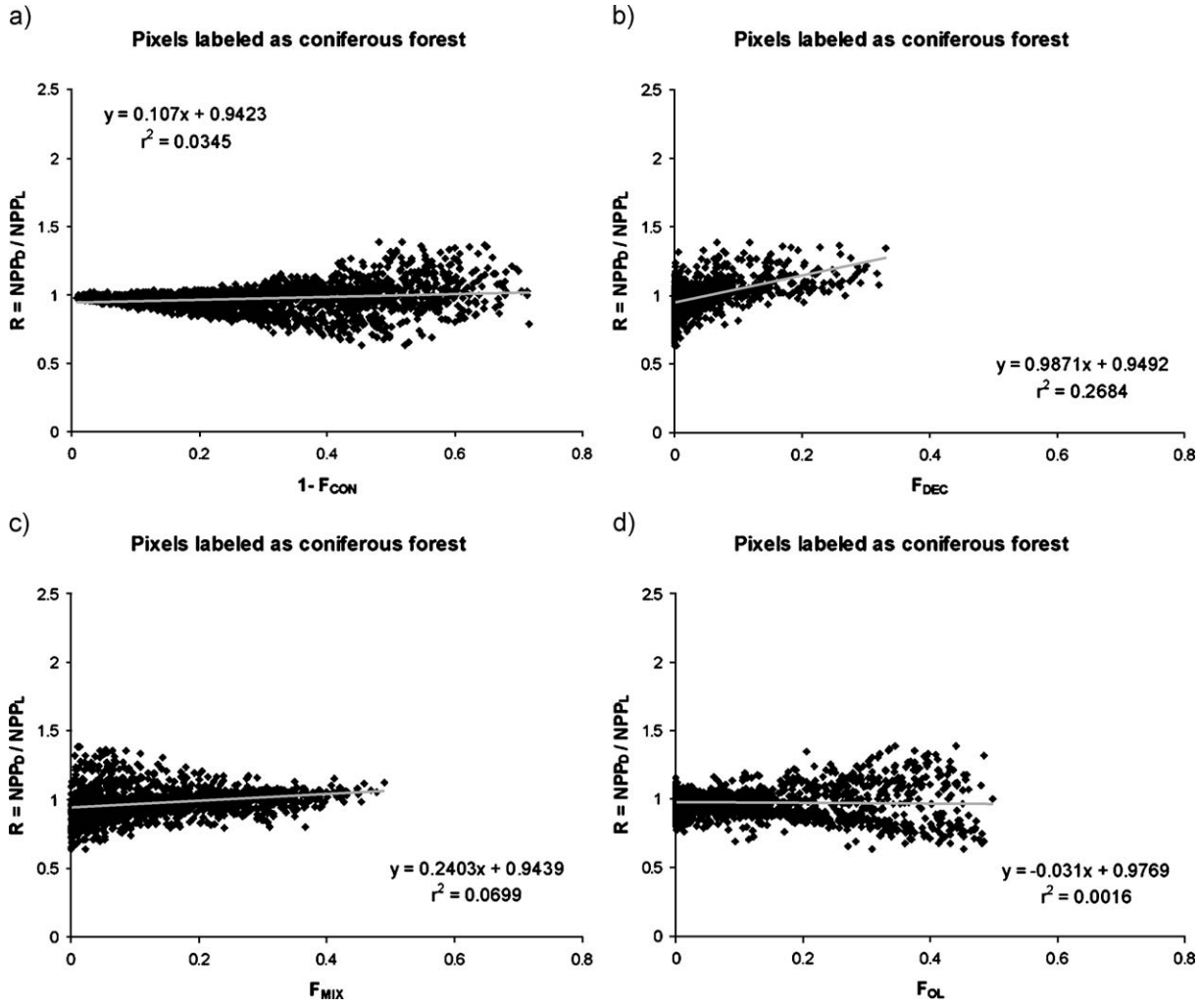

Fig. 5. Regression analysis for pixels labeled as coniferous forest: relationship between the correction factor R and (a) non-coniferous fraction ($1-F_{CON}$), (b) deciduous fraction (F_{DEC}), (c) mixed fraction (F_{MIX}), and (d) open land fraction (F_{OL}).

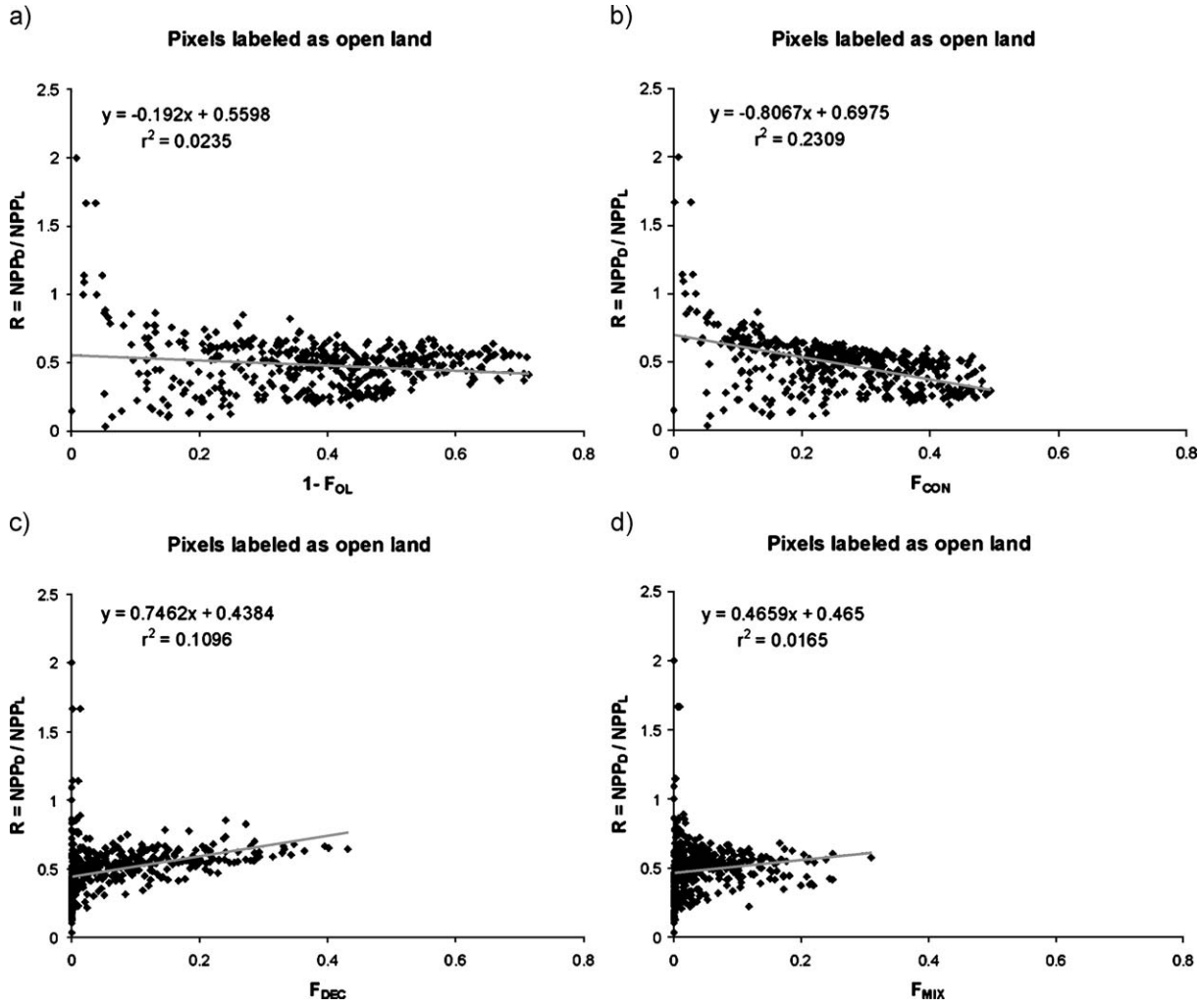

Fig. 6. Regression analysis for pixels labeled as open land: relationship between the correction factor R and (a) non-open land fraction ($1-F_{OL}$), (b) coniferous fraction (F_{CON}), (c) deciduous fraction (F_{DEC}), and (d) mixed fraction (F_{MIX}).

forest with large biomass (Reich et al., 1999) and (2) leaves in grass canopies are more evenly distributed than that in forests allowing for more radiation absorption and higher NPP (Liu et al., 2002).

The outcome of the corrections for all land cover types is shown in Fig. 7. The figure illustrates that the corrected NPP_L values are reduced substantially and the correlation between the corrected NPP_L and NPP_D values is considerably improved (r^2=0.59) from the non-corrected case (r^2=0.16) shown in Fig. 3. Although the mean corrected NPP_L (240.4 g C m^{-2} $year^{-1}$) is not considerably reduced from NPP_L (243.3 g C m^{-2} $year^{-1}$) as expected, the major corrections can be observed within the highest values. The range is reduced from 15–610 to 14–522 g C m^{-2} $year^{-1}$, and is closer to the NPP_D range (10–360 g C m^{-2} $year^{-1}$). The overall standard deviation is reduced from 61.7 to 56.5 g C m^{-2} $year^{-1}$ in the corrected results. The r^2 value of the NPP_D–NPP_L relationship for each land cover type ranges from 0.36 to 0.91 after applying the correction (Fig. 8), while it ranges from 0.42 to 0.84 before the correction (Fig. 4). The correction results in better agreement for conifer-labeled (r^2=0.91) (Fig. 8a) and deciduous forest-labeled pixels (r^2=0.74) (Fig. 8b). The correction is less effective for mixed forest-labeled pixels. Although most points fall closely to the regression line, a number of points, containing high conifer fractions in particular, are widely scattered. This reduces the overall correlation of mixed forest-labeled pixels from r^2=0.42 to r^2=0.38 (Fig. 8c). The correction is most obvious for open land- labeled pixels with a marked improvement in the correlation from r^2=0.47

Table 3
Regression coefficients C_{ij} used in the algorithms for scaling in this study (F_{CON} denotes coniferous nondominant fraction, F_{DEC} deciduous nondominant fraction, F_{MIX} mixed nondominant fraction, and F_{OL} denotes open land nondominant fraction)

Nondominant fractions	Labels of lumped pixels			
	Coniferous	Deciduous[a]	Mixeda[a]	Open land
F_{CON}	–	−0.7374	−0.5019	−0.8067
F_{DEC}	0.9871	–	1.2156	0.7462
F_{MIX}	0.2403	−0.6027	–	0.4659
F_{OL}	−0.031	1.688	2.0985	–

[a] Statistical approach was used to reduce the impact of extreme values.

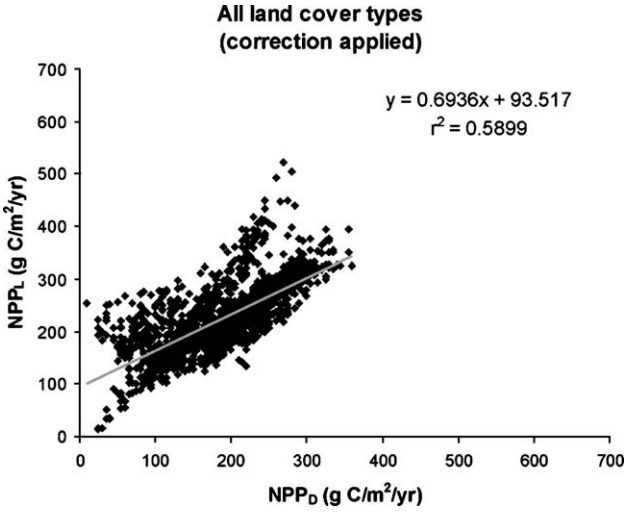

Fig. 7. Distributed NPP and lumped NPP after the correction for land cover type area fractions as subpixel information (all land cover types).

to r^2=0.65. The regression line also becomes close to the 1:1 line (Fig. 8d). Even though the correction appears to be most effective for the open land-labeled pixels, a majority of all rebelliant points after the correction are seen within this land cover type. This indicates that the effectiveness of the spatial scaling based on the land cover type fractions may be suppressed by within-class NPP variability, which is, as discussed above, related to LAI variability of the class.

4.4. Application of LAI correction

Fig. 9 shows the relationship between NPP_D and NPP_L after the additional LAI correction (Eq. (4)) for all land cover types together. The results are substantially improved mostly due to the LAI correction applied to open land-labeled pixels. The nonlinearity factor α is based on the NPP–LAI relationship and according to BEPS simulations it

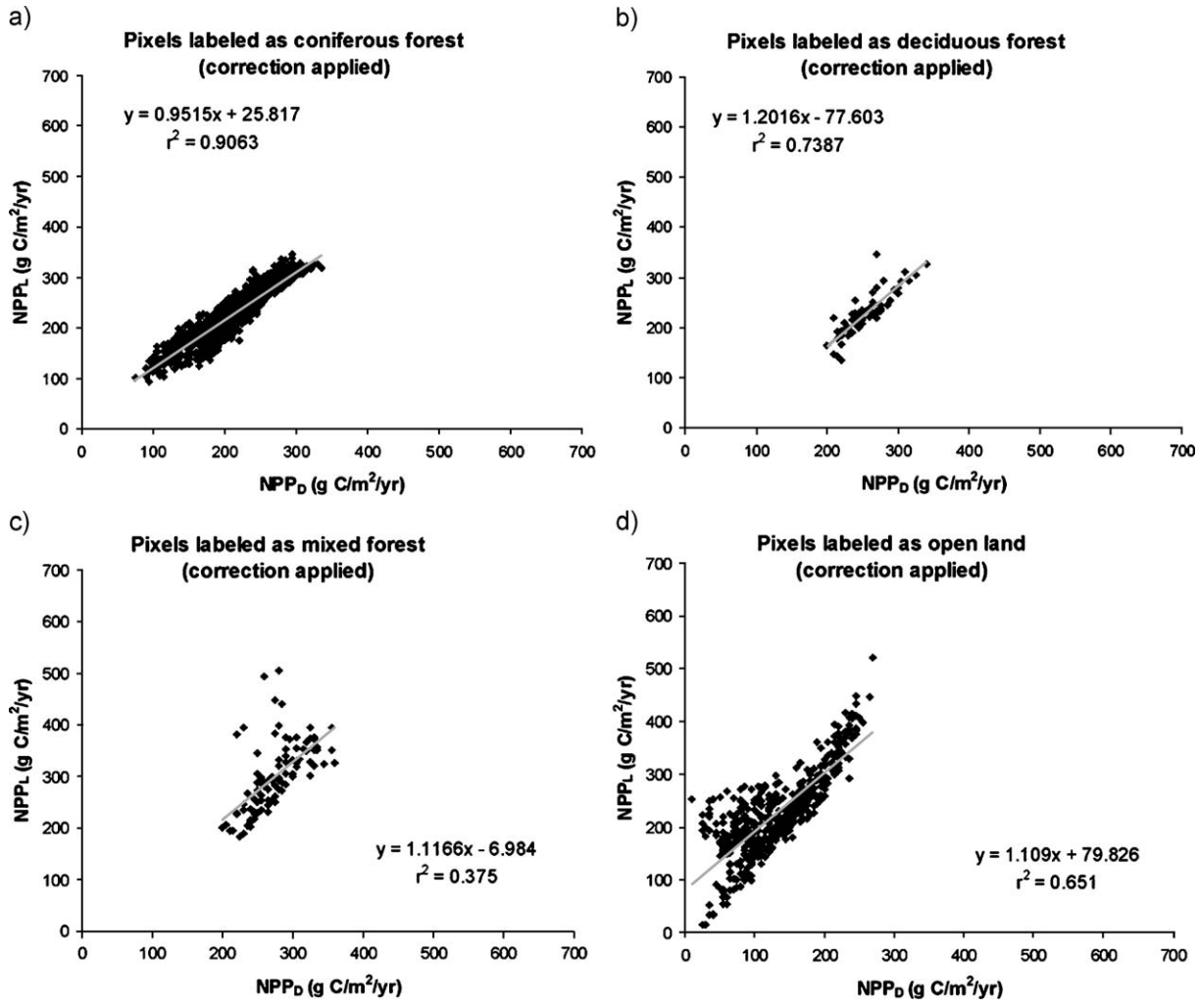

Fig. 8. Relationship between distributed NPP and lumped NPP after the correction for pixels designated as (a) coniferous forest, (b) deciduous forest, (c) mixed forest, and (d) open land.

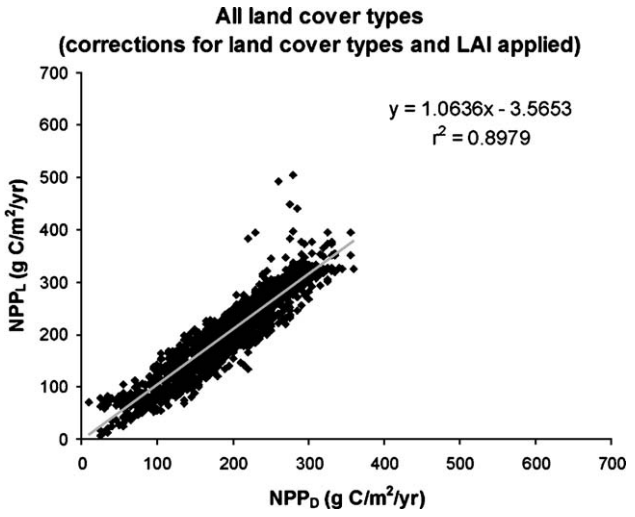

Fig. 9. Distributed NPP and lumped NPP after the correction for land cover type area fractions and LAI (all land cover types).

was found to be 0.78. The effect of the NPP–LAI nonlinearity within open land pixels is also apparent in Fig. 6a, when linear regression is compared with data. As the differences in LAI within the other land cover types are not substantial, no significant improvements are found for them. The correction for the open land LAI alone results in a strong agreement between the NPP_D and corrected NPP_L values (r^2=0.90). All points in Fig. 9 are located close to 1:1 line with the exception of some points of mixed-labeled pixels. The results suggest that the high LAI variability of open land plays an important role in scaling process for this study area.

5. Discussion

The spatial scaling process here is implemented as corrections to lumped values. Although the meaning of spatial scaling goes far beyond changes in spatial resolution, the scaling procedure proposed here is nevertheless a practical solution to improving the accuracy of mapping a complex surface parameter NPP. The importance of spatial scaling for NPP, a parameter of much concern in climate change and carbon cycle studies, is seen from (i) the large variability in the difference between distributed and lumped NPP values; (ii) the seemingly random errors in lumped pixels do not cancel each other when the average of the whole study area is taken and the overall bias in the lumped calculations is considerable (about 15%); and (iii) simple corrections using subpixel land cover and LAI information are effective in reducing the random and bias errors.

Two main issues arise when the scaling procedure is considered for application in other areas and/or for other sensor types: (i) is the scaling procedure derived between two fixed resolutions applicable to images of other resolutions? and (ii) how can we obtain the needed subpixel information?

To address issue (i), we should realize that the contextural approach of scaling is only sensitive to the fractions of cover types within lumped pixels and, in theory, it should be independent of the scale of concern. In other words, the methodology should be applicable to images of any resolutions. However, the basic assumption in using contextural information, as the major scaling vehicle, is that variability within a given cover type is smaller than the differences among cover types. While this assumption is generally true, it can be violated in some cases, especially when several cover types are combined into one, such as the case in our present study where open land refers to a wide range of residual cover types (bare soils, new burned scars, old burned areas with productive grasses and shrubs, etc.). We have therefore made a further step to reduce the impact of the variability within the same cover type between lumped pixels. This second step of correction is useful when the application is limited to only a small number of cover types and can address the issues of within-cover type variability.

Issue (ii) may be most critical to make the scaling procedure practical. If we only have coarse resolution images for an area, we are severely limited in our ability to perform meaningful spatial scaling. Although procedures of scaling using textural information have been proposed assuming the scale-invariance among resolutions (Hall et al., 1992; Hu & Islam, 1997), these can only result in weak accuracy improvement (Chen, 1999). However, to meet the scaling requirement, the traditional practice in land classification with ‘hard labeling’, i.e. forcing a pixel to a unique cover type, may be replaced with ‘soft labeling’ approaches, i.e. giving the percentage of major cover types within each pixel. This soft classification approach has been successfully demonstrated by DeFries et al. (1997). We realize that it is generally difficult to do the soft classification for multiple cover types through spectral unmixing as the unique dimensions of optical remote sensing are generally smaller than the dimensions of surface variability (Verstrate et al., 1996). We therefore suggest that we should pay a great attention to regional and global land cover mapping at high resolutions, i.e. at a resolution which is comparable to surface variability, if we take the spatial scaling as a serious issue in quantitative remote sensing applications. This is in agreement with the suggestion by Chen (1999) that we at least need a high-resolution water area mask for spatial scaling of surface parameters.

We realize that to some extent the results presented here can suffer from sources of errors due to treatments of data, assumptions used and inaccuracy in the NPP model. High-resolution LAI data were acquired one time in the growing season and the values are assumed to be constant for each conifer pixel and follow a square wave (i.e. zero before and after the growing season but constant during the growing season) for each deciduous pixel. Mixed and open land pixels were treated as the intermediate cases between conifers and deciduous. These simple treatments can cause

biases in both distributed and lumped calculations, and it is believed that these biases are greatly reduced when the analysis is based on the ratio of distributed and lumped values as the biases are in the same direction in these two cases. The major assumption that could potentially affect the scaling analysis is that fine-resolution pixels truly capture the surface heterogeneity and vegetation is uniform within each pixel. The degree to which this assumption is true is unknown, but we believe that the mean errors should be much smaller than 15% found between 30 m and 1 km resolutions and are mostly removed in the process of taking the ratio between distributed and lumped values. The accuracy of the BEPS model in simulating the differences in NPP among cover types can have considerable effects on determination of the actual values of the scaling coefficients. As BEPS NPP results have been carefully calibrated for conifer and deciduous boreal species against tower flux data (Chen et al., 1999), we believe that the model used here has captured the main differences among the major cover types in the study area.

6. Conclusions

It is demonstrated that NPP mapping at coarse resolutions can have large random errors and considerable bias errors when the subpixel heterogeneity is not considered. In the current study, the relative random errors in NPP are found to exceed 100% for about 10% pixels at 1-km resolution. The relative bias errors for the whole study area at 1-km resolution are about 15%. These errors may be common in our current regional and global NPP maps. These errors can be greatly reduced when a spatial scaling procedure is used based on the subpixel information. The subpixel information required in the scaling procedure includes the area fractions and LAI of major cover types within each pixel at a coarse resolution. The results of the scaling process suggest the following: (i) contextural-based spatial scaling based on subpixel cover type area fractions is an effective approach to reduce the errors in coarse resolution NPP estimation; and (2) The use of subpixel LAI values within a coarse-resolution pixel can further improve the estimation of the NPP value of the pixel.

Advances in remote sensing technology results in multiple sensors with various spatial and temporal resolutions. This accentuates the importance of the spatial scaling process in quantitative analysis of remote sensing. The availability of space-based imagery with a high spatial resolution is in constant increase, which will allow further exploration of the relationship between measurements at different spatial resolutions. Since land cover is generally stable with time, a long-term investment to create land cover maps for a given region and for the globe using high-resolution images such as from Landsat would be useful. However, constant attention should be given to the regions with natural disasters such as forest fires where land cover maps should be generated at a frequency comparable to the speed of surface changes. The most practical conclusion of this study is that we need to make efforts to produce regional and global land cover maps at a high resolution as masks for spatial scaling of surface parameters between various spatial resolutions.

Acknowledgements

The authors are indebted to Mr. Goran Pavlic at CCRS for assistance in the initial Landsat TM image analysis. J. Chen thanks Dr. Josef Cihlar of CCRS for stimulating and encouraging discussions.

References

Band, L. E., & Moore, I. D. (1995). Scale: Landscape attributes and geographical information systems. *Hydrological Processes*, *9*, 401–422.

Bonan, G. B., Pollard, D., & Thomson, S. L. (1993). Influence of subgrid-scale heterogeneity in leaf area index, stomatal resistance, and soil moisture on grid-scale land–atmosphere interactions. *Journal of Climate*, *6*, 1882–1896.

Brown, L., Chen, J. M., Leblanc, S. G., & Cihlar, J. (2000). A shortwave infrared modification to the simple ratio for LAI retrieval in boreal forests: An image and model analysis. *Remote Sensing of Environment*, *71*, 16–25.

Chen, J. M. (1999). Spatial scaling of a remote sensed surface parameter by contexture. *Remote Sensing of Environment*, *69*, 30–42.

Chen, J. M., Liu, J., Cihlar, J., & Goulden, M. L. (1999). Daily canopy photosynthesis model through temporal and spatial scaling for remote sensing applications. *Ecological Modelling*, *124*, 99–119.

Chen, J. M., Pavlic, G., Brown, L., Cihlar, J., Leblanc, S. G., White, H. P., Hall, R. J., Peddle, D. R., King, D. J., Trofymow, J. A., Swift, E., Van der Sanden, J., & Pellikka, P. K. E. (2002). Derivation and validation of Canada-wide coarse-resolution leaf area index maps using high-resolution satellite imagery and ground measurements. *Remote Sensing of Environment*, *80*, 165–184.

Csillag, F., Fortin, M., & Dungan, J. L. (2000). On the limits and extensions of the definition of scale. *Bulletin of the Ecological Society of America*, *81*(3), 230–232.

Curran, P. J., & Foody, G.M (1994). The use of remote sensing to characterize the regenerative states of tropical forests. In G. Foody, & P. Curran (Eds.), *Environmental remote sensing from regional to global scales*. Chichester, England: John Wiley and Sons.

Defries, R., Hansen, M., Steininger, M., Dubayah, R., Sohlberg, R., & Townshend, J. (1997). Subpixel forest cover in central Africa from multisensor, multitemporal data. *Remote Sensing of Environment*, *60*, 228–246.

Ehleringer, J. R., & Field, C. B. (1993). Scaling physiological processes: Leaf to globe. Boston: Academic.

Friedl, M. A. (1996). Examining the effects of sensor resolution and subpixel heterogeneity on spectral vegetation indices: Implications for biophysical modeling. In D. A. Quattrochi, & M. F. Goodchild (Eds.), *Scaling of remote sensing data for GIS* (pp. 113–139). New York: Lewis Publishers.

Friedl, M. A., Davis, F. W., Michaelsen, D. J., & Moritz, M. A. (1995). Scaling and uncertainty in the relationship between the NDVI and land surface biophysical variables: An analysis using a scene simulation model and data from FIFE. *Remote Sensing of Environment*, *54*, 233–246.

Hall, F. G., Huemmrich, K. F., Goetz, S. J., Sellers, P. J., & Nickeson, J. E. (1992). Satellite remote sensing of surface energy balance: Success, failures, and unresolved issues in FIFE. *Journal of Geophysical Research*, *97*(D17), 19061–19089.

Hu, Z., & Islam, S. (1997). A framework for analyzing and designing scale invariant remote sensing algorithm. *IEEE Transactions on Geoscience and Remote Sensing*, *35*, 747–755.

Liu, J., Chen, J. M., Cihlar, J., & Chen, W. (1999). Net primary productivity distribution in the BOREAS region from a process model using satellite and surface data. *Journal of Geophysical Research*, *104*(D22), 27735–27754.

Liu, J., Chen, J. M., Cihlar, J., & Chen, W. (2002). Net primary productivity mapped for Canada at 1-km resolution. *Global Ecology and Biogeography*, *11*, 115–129.

Liu, J., Chen, J. M., Cihlar, J., & Parl, W. M. (1997). A process-based boreal ecosystem productivity simulator using remote sensing inputs. *Remote Sensing of Environment*, *62*, 158–175.

Nemani, R. R., Pierce, L., Band, L. E., & Running, S. W. (1993). Forest ecosystem processes at the watershed scale: Sensitivity to remotely sensed leaf area index observations. *International Journal of Remote Sensing*, *14*, 2519–2534.

Pierce, L. L., & Running, S. W. (1995). The effects of aggregating sub-grid land surface variation on large-scale estimates of net primary production. *Landscape Ecology*, *10*, 239–253.

Rastetter, E. B., King, A. W., Cosby, B. J., Hornberger, G. M., O'Neill, R. V., & Hobbie, J. E. (1992). Aggregating fine-scales ecological knowledge to model coarser-scale attributes of ecosystems. *Ecological Applications*, *2*(1), 55–70.

Reich, P. B., Turner, D. P., & Bolstad, P. (1999). An approach to spatially distributed modeling of net primary productivity (NPP) at the landscape scale and its application in validation of EOS NPP products. *Remote Sensing of Environment*, *50*, 69–81.

Running, S. W., & Coughlan, J. (1988). A general model of forest ecosystem processes for regional applications: I. Hydrological balance, canopy gas exchange and primary production processes. *Ecological Modelling*, *42*, 125–154.

Solomon, A. M., & Shugart, H. H. (1993). Vegetation dynamics and global change. New York: Chapman and Hall.

Tian, Y., Wang, Y., Zhang, Y., Knyazikhin, Y., Bogaert, J., & Myneni, R. B. (2003). Radiative transfer based scaling of LAI retrievals from reflectance data of different resolution. *Remote Sensing of Environment*, *84*, 143–159.

Tian, Y., Woodcock, C. E., Wang, Y., Privette, J. L., Shabanov, N. V., Zhou, L., et al. (2002). Multiscale analysis and validation of the MODIS LAI product: II. Sampling Strategy. *Remote Sensing of Environment*, *83*, 431–441.

Turner, D. P., Dodson, R., & Marks, D. (1996). Comparison of alternative spatial resolution in the application of a spatially distributed biochemical model over complex terrain. *Ecological Modelling*, *90*, 53–67.

Verstrate, M. M., Pinty, B., & Myneni, R. B. (1996). Potential and limitations of information extraction on the terrestrial biosphere from satellite remote sensing. *Remote Sensing of Environment*, *58*, 201–214.

Waring, R. H., & Running, S. W. (1998). Forest ecosystems: Analysis at multiple scales. San Diego, CA: Academic Press.