RESEARCH ARTICLE

Upscaling as ecological information transfer: a simple framework with application to Arctic ecosystem carbon exchange

Paul C. Stoy · Mathew Williams · Mathias Disney ·
Ana Prieto-Blanco · Brian Huntley · Robert Baxter ·
Philip Lewis

Abstract Transferring ecological information across scale often involves spatial aggregation, which alters information content and may bias estimates if the scaling process is nonlinear. Here, a potential solution, the preservation of the information content of fine-scale measurements, is highlighted using modeled net ecosystem exchange (NEE) of an Arctic tundra landscape as an example. The variance of aggregated normalized difference vegetation index (NDVI), measured from an airborne platform, decreased linearly with log(scale), resulting in a linear relationship between log(scale) and the scale-wise modeled NEE estimate. Preserving three units of information, the mean, variance and skewness of fine-scale NDVI observations, resulted in upscaled NEE estimates that deviated less than 4% from the fine-scale estimate. Preserving only the mean and variance resulted in nearly 23% NEE bias, and preserving only the mean

resulted in larger error and a change in sign from CO_2 sink to source. Compressing NDVI maps by 70–75% using wavelet thresholding with the Haar and Coiflet basis functions resulted in 13% NEE bias across the study domain. Applying unique scale-dependent transfer functions between NDVI and leaf area index (LAI) decreased, but did not remove, bias in modeled flux in a smaller expanse using handheld NDVI observations. Quantifying the parameters of statistical distributions to preserve ecological information reduces bias when upscaling and makes possible spatial data assimilation to further reduce errors in estimates of ecological processes across scale.

Keywords Abisko · Information content · Information theory · Leaf area index · Net ecosystem exchange · Normalized difference vegetation index · Skew-normal distribution · Tundra · Upscaling · Wavelet decomposition

P. C. Stoy (✉) · M. Williams
School of GeoSciences, University of Edinburgh,
Edinburgh EH9 3JN, UK
e-mail: paul.stoy@ed.ac.uk

M. Disney · A. Prieto-Blanco · P. Lewis
Department of Geography, University College London,
26 Bedford Way, London WC1H 0AP, UK

B. Huntley · R. Baxter
School of Biological and Biomedical Sciences, Durham
University, South Road, Durham DH1 3LE, UK

Introduction

Ecological information is often collected in plots whose spatial extent (e.g., 10^{-2}–10^1 m^2) do not overlap the spatial grain of satellite remote sensing products (e.g., 10^2–10^6 m^2). Upscaling (downscaling) ecological information thus relies on spatial extrapolation, often via aggregation (disaggregation), which can result in bias errors if the relationship between

observation and process is nonlinear (O'Neill and Rust 1979; Pelgrum 2000; Rastetter et al. 1992). It has been demonstrated that preserving the information content (IC) of fine-scale observations can minimize bias in upscaled estimates of land-surface features and processes (Chen and Blong 2002; Stoy et al. 2009), but general approaches for quantifying how much ecological information must be retained to minimize bias have rarely been explored [but see Rastetter et al. (1992)].

Problems of scale in ecology are fundamentally related to the transfer of information, for which a well-developed theory exists, information theory, that has paralleled the academic development of computer science for the communication, compression and quantification of information (Kullback 1997; Kullback and Leibler 1951; Reza 1994; Shannon 1948). It is implicit that not all information must be retained when transferring ecological information across scale (Rastetter et al. 1992), i.e., information can be compressed to minimize computational space and preserve only the critical features required for describing the system. Information compression can be either lossless (e.g., the .zip file format), which preserves IC, or lossy, in which case the amount of information lost must be quantified and analyzed to ensure that the outcome is desirable for the task at hand.

The proportion of data removed in lossy data compression algorithms is generally related to what our eyes and ears can resolve, rather than removing information that is not of ecological importance for the efficient and unbiased transfer of relevant information across scale. For example, the common JPEG (Joint Photographic Experts Group) image compression algorithm (the .jpg file format) reduces the amount of digital information contained in an image without appreciably altering its resolution to the human eye. It is an example of lossy data compression; digital information that contributes little to the eye's ability to resolve an image is removed. The degree to which ecological information can be compressed for accurate multi-scale process estimation is less clear and depends on the application.

Information theoretical concepts have found applications in ecological science (Ulanowicz 2001), but are infrequently applied to problems of scale (Brunsell et al. 2008; Brunsell and Young 2007; Chen and Blong 2002), especially scale in ecology (Stoy et al.

2009). Rather, IC has been used to describe ecological diversity (MacArthur 1955) and landscape patterning (Turner 1989), to discriminate amongst ecological models (Akaike 1974; Schwarz 1978), and in the creation of efficient ecological models (Burnham and Anderson 2002; Ulanowicz 1980). Here, we demonstrate how concepts from information theory can be applied to communicate information across ecological scales. As opposed to most investigations of aggregation in remote sensing of the environment, which tend to focus, logically, on methodology for the accurate estimation of surface features (Jiang et al. 2006), the minimization of bias in the output of an ecosystem model (Shaver et al. 2007) is used as our primary criterion for success when choosing an aggregation procedure. The purpose is therefore to quantify the minimum amount of 'ecologically relevant' information about the land surface that must be maintained when moving across scale, using tundra ecosystem C exchange as an example. Specifically, we address the question of how much fine-scale land-surface information, via the normalized difference reflectance index (NDVI), must be retained at multiple scales for unbiased estimates of CO_2 flux in an Arctic landscape.

We choose this example for its practical importance for upscaling estimates of C exchange given the spatial complexity of vegetation distribution (Spadavecchia et al. 2008) that results from, for example, patterned ground formations, microtopography and grazing of tundra ecosystems (Kessler and Werner 2003; Krantz 1990), and to highlight the care that must be taken when applying a highly nonlinear transfer function, in this case the relationship between NDVI and LAI in tundra (van Wijk and Williams 2005; Williams et al. 2008), to scale information. Whereas the results are specific to our study location, we envision that similar information conserving techniques can be used to determine an appropriate aggregation procedure in other landscapes that are likewise dominated by areas of relatively high and low ecological activity (e.g., 'hotspots' and 'dead zones').

We first quantify the errors in modeled CO_2 flux that accrue when scaling observations of the NDVI through a nonlinear transfer function to LAI (van Wijk and Williams 2005; Williams et al. 2008) for use in a tundra ecosystem model ['PLIRTLE', Shaver et al. (2007)]. The choice of NDVI scaling reflects the

LAI requirement of the PLIRTLE model, and the nonlinearity of the NDVI-LAI transfer function introduces an upscaling challenge that is representative of nonlinear scaling issues in ecology (O'Neill and Rust 1979; Pelgrum 2000; Rastetter et al. 1992). We then demonstrate relationships amongst bias and scale and suggest techniques to reduce this bias by preserving the IC of the original observations. The effects of reducing digital information, but not ecological information, via 2D wavelet image compression are explored to extend the methodology to cases where preserving spatial relationships matter. The discussion focuses on why the information required for scaling is formally related to the observed variance from the standpoint of information theory, how scaling laws in ecology (Enquist et al. 2003) may be employed to reduce scale-wise bias, and how information retained via the probability density function (pdf) of land surface properties can be used to assimilate data into ecological models further to reduce scale-wise bias errors.

Methods

Study site

The spatial aggregation of NDVI measurements from handheld and airborne instruments is investigated. Land surface reflectance and CO_2 exchange were measured in a sub-Arctic tundra landscape near Abisko, Sweden (68°18′N, 18°50′E). Vegetative and edaphic characteristics of the study area are described elsewhere (Fox et al. 2008; Shaver et al. 2007; Street et al. 2007; Williams et al. 2008).

Airborne and handheld NDVI measurements

NDVI was calculated for 4×4 m pixels from an Azimuth Systems AZ-16 Airborne Thematic Mapper (ATM) on 17th July, 2005 using the 0.63–0.69 μm (visible red) and 0.76–0.90 μm (NIR) bands, equivalent to Landsat bands 3 and 4. These 16 m² grid cells represent the spatial grain (smallest spatial extent) of the remotely sensed NDVI measurements and are henceforth called the 'fine scale' measurements. A 65,536 m² slice of NDVI data, equivalent to 64×64 pixels, was chosen for this analysis. ATM

data were atmospherically corrected using a combination of empirical line correction (ELC) and FLAASH/MODTRAN5, ITT Visual Information Systems, and georeferenced using a 4 m resolution digital elevation model of the area and full ancillary imaging geometry. Handheld NDVI measurements were made using a SKR1800 two-channel sensor (Skye Instruments, Llandrindod Wells, UK) at the scale of 0.2×0.2 m in a regular 900 m² area as described by van Wijk and Williams (2005).

Spatial aggregation

ATM-measured NDVI was aggregated using a 'quad-tree' of square averaging operators of increasing size by doubling the length scale until a single pixel representing the average of all measurements was reached (Kolaczyk et al. 2005). For the handheld NDVI measurements, the 5,625 pixel area was averaged using 625, 225, 25 and 9 square pixels (i.e., multiples of $75 = 5625^{1/2}$).

In probabilistic terms, the average over the extent of observations represents a degenerate distribution in which the probability mass function has single-valued support. Preserving only the mean collapses the retained information, as formally quantified via the Shannon Entropy for any bin number N, to zero:

$$H(X) = -\sum_{i=1}^{N} p(x_i) \log p(x_i)$$
$$= -(N-1) \times 0 - 1 \times \log(1) = 0. \qquad (1)$$

It is proposed that maintaining more information about the underlying distribution can minimize aggregation error (Rastetter et al. 1992). This information is retained by the parameters of the distribution of fine-scale NDVI. The beta, normal and skew-normal distributions are explored for transferring information in fine-scale NDVI measurements to larger scales.

The choice of the beta distribution follows from its finite support; NDVI bounded between 0 and 1. The pdf of the beta distribution is:

$$\text{Beta}(\alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} \qquad (2)$$

where α and β are the shape parameters and $B(\alpha, \beta)$ is the beta function.

The common Normal distribution:

$$N(\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad (3)$$

can be expanded to account for non-zero skewness using the skew-normal distribution:

$$\begin{aligned} \text{SN}(\xi, \omega^2, \alpha_s) = &\frac{2}{\omega\pi}\exp\left(-\frac{(x-\xi)^2}{2\omega^2}\right) \\ &\times \int_{-\infty}^{\alpha_s\left(\frac{x-\xi}{\omega}\right)}\exp\left(-\frac{t^2}{2}\right)dt \end{aligned} \qquad (4)$$

with location (ξ), scale (ω^2) and shape (α_s) parameters (Azzalini 1985). Maximum likelihood parameter estimates were quantified for airborne and handheld NDVI using MATLAB (Natick, MA). Open-source code for maximum likelihood parameter estimates using the skew-normal distribution was provided by A. Azzalini.

Retaining spatial information: 2D wavelet compression

Image compression via wavelet thresholding can decrease file size dramatically while preserving IC (Chang et al. 1997). Changes in variance, IC, and NEE (i.e., bias) that result from image compression were explored using wavelet thresholding with the 2D Haar wavelet basis function (a 2D square wave). The 2D Haar basis function was chosen to resolve the sharp discontinuities in NDVI that may exist along the edge of vegetation patches (Laurance 2000), and to create a pixilated land surface representation of a tundra landscape that is non-stationary in space. Results were compared to compression using the Coiflet wavelet, a more common choice for image compression (Mandal et al. 1996). In the compression step, the original NDVI map was first decomposed into its spatial wavelet coefficients using fast wavelet transformation. The resulting wavelet coefficients were sorted and thresholded by setting a fraction of these coefficients to zero, thereby removing the spatial variance that these coefficients describe. A compressed version of the original NDVI map was then created by inverse fast wavelet transformation. The compression step was repeated, thresholding 5% more coefficients, until 95% compression was reached.

NDVI–LAI transfer functions

Measured fine-scale NDVI, spatially aggregated NDVI, and fitted distributions of NDVI were converted to LAI using the relationship described by both van Wijk and Williams (2005),

$$\text{LAI} = 0.00067e^{9.237\text{NDVI}} \qquad (5)$$

and the scale-dependent transfer functions described by Williams et al. (2008) in which the parameters of Eq. 5 vary as a function of scale [see Fig. 3 in Williams et al. (2008)]. The distributions of these LAI estimates, $p(\text{LAI})$ are used subsequently in the carbon flux modeling analysis. The coefficients of the equations used by van Wijk and Williams (2005) and Williams et al. (2008) are based on NDVI measures calculated from radiance (i.e., not accounting for spectral variations in down-welling radiance). The aircraft-measured NDVI are based on surface reflectance data, and so do account for spectral variations in down-welling radiance. To convert these to the basis used by van Wijk and Williams (2005) and Williams et al. (2008), it was noted that the ratio of incident near infrared to incident red radiation was a time-invariant 0.8 at a nearby flux tower using upward-pointing Skye NDVI sensors so:

$$\frac{1 + \text{NDVI}}{1 - \text{NDVI}} = 0.8\frac{1 + \text{NDVI}'}{1 - \text{NDVI}'} \qquad (6)$$

where NDVI′ is the reflectance-based value. ATM-measured NDVI was converted to this upwelling-only basis for the purposes of flux calculation.

The PLIRTLE biosphere-atmosphere CO_2 flux model

The net ecosystem exchange of CO_2 (NEE) was modeled using the observed relationship between photosynthesis, LAI and irradiation and between respiration, temperature and LAI ['PLIRTLE', Shaver et al. (2007)]. Shaver et al. (2007) demonstrated functional convergence of tundra vegetation by applying the PLIRTLE model with a single parameterization to measured C fluxes from globally distributed tundra ecosystems, including the Abisko tundra ecosystem investigated here, making it a logical choice for multi-scale modeling of C flux in tundra.

The photosynthesis [gross primary productivity (GPP)] model follows the aggregated canopy

photosynthesis model of Rastetter et al. (1992). GPP is assumed to follow saturating response to photo-synthetically active radiation (I). The distribution of GPP [p(GPP)] across pixels of the measurement domain was calculated using:

$$p(\text{GPP}) = \left(\frac{P_{\max}}{k} \ln \left(\frac{P_{\max} + E_o I}{P_{\max} + E_o I e^{-kp(\text{LAI})}} \right) \right) \qquad (7)$$

where k is the light extinction coefficient [assumed here to be 0.5, after Shaver et al. (2007)], E_o is the light sensitivity of photosynthesis, $P_{\max}$ is maximum photosynthesis and $p(\text{LAI})$ is the pdf of LAI from fine-scale or aggregated NDVI, or NDVI distribution.

Shaver et al. (2007) found that an ecosystem respiration (RE) model that assumes both LAI-dependent and LAI-independent sources of respiration R performs better than those that employ a single substrate pool such that:

$$p(\text{RE}) = p(\text{LAI}) R_0 e^{\beta T} + R_x \qquad (8)$$

where R_0 is base respiration at 0°C, R_x is the LAI and temperature (T)-insensitive component of ecosystem respiration and β is the temperature sensitivity of RE. Parameter values were likewise taken from the global parameter set determined by Shaver et al. (2007).

The net ecosystem exchange of CO_2 (NEE) is defined as

$$p(\text{NEE}) = p(\text{RE}) - p(\text{GPP}), \qquad (9)$$

and the total landscape-level NEE estimate was determined by summation across the pdf.

Equation 9 follows the meteorological convention where C flux from atmosphere to biosphere is denoted as negative; this convention was also chosen for consistency with Shaver et al. (2007). Continuous half-hourly measurements of T and I from the nearby ANS research station were used to drive PLIRTLE. The period from June 1 until August 31, 2007, is taken to be the peak growing season (gs) for the purposes of this analysis. NEE estimates that result from the airborne and handheld fine-scale NDVI measurements are taken to be the best estimates of C flux at their respective spatial grain sizes. NEE estimates created by NDVI aggregation, modeling the distribution of NDVI, and compressing information in the NDVI map by wavelet thresholding are compared to these fine-scale estimates.

## Results

Airborne NDVI measurements

Aggregated NDVI maps and corresponding pdfs are shown in Fig. 1 (Note that the original NDVI image is shown in Fig. 5a). It is clear that extreme values of NDVI were 'averaged out' in the progressive spatial aggregation steps of the quad-tree upscaling procedure. Corresponding NEE estimates are listed in Table 1. The fine-scale NEE estimate of -371 kg C gs^{-1} (for the 65,536 m^2 study domain) changed to -249 kg C m^{-1} gs^{-1} (i.e., lower C sequestration) after the first aggregation step (64 m^2 pixels), a difference of nearly 33%. Averaging all NDVI values before conversion to LAI resulted in a flux estimate of $+355$ kg C m^{-1} gs^{-1} (Table 1) which represents a landscape-level estimate of C loss that is on the order of the C sequestration estimate that resulted from fine-scale NDVI observations. The substantial differences in NEE estimates, nearly 200%, that are incurred by spatial averaging of NDVI are the result of comparatively minor percent biases in GPP (28%) and RE (8%, Table 1). The absolute values of GPP and RE are an order of magnitude larger than NEE and relatively small biases in these component fluxes resulted in substantial NEE bias.

Changing the scale of the spatial averaging operators resulted in a decrease in the variance of NDVI (σ^2_{NDVI}, Fig. 2a) and a corresponding increase in the NEE estimate as aggregated NDVI was transferred through the NDVI-LAI transfer function (Eq. 5, Fig. 2b) to generate an LAI distribution for the PLIRTLE model. The result is an approximately linear relationship between log(scale) and the modeled NEE estimate (Fig. 2C, Table 1) due to the alteration in fine scale information, as quantified by the Kullback–Liebler divergence (D$_{KL}$, Fig. 2D) for normal distributions:

$$D_{\text{KL}}(F|A) = \frac{1}{2} \left(\log\left(\frac{\sigma^2_A}{\sigma^2_F} \right) + \frac{\sigma^2_F}{\sigma^2_A} + \frac{(\mu_A - \mu_F)^2}{\sigma^2_A} - 1 \right) \qquad (10)$$

where F and A represent the fine scale and aggregated NDVI distributions and μ and σ^2 represent the respective mean and variance.

Fig. 1 Spatial aggregation of ATM-measured NDVI in a tundra landscape near Abisko, Sweden at increasing pixel length scale from *top* to *bottom* (*Left panels*). The corresponding probability density functions of aggregated NDVI [*p*(NDVI)] versus the distribution of observed fine-scale (4 × 4 m) NDVI (*Right panels*)
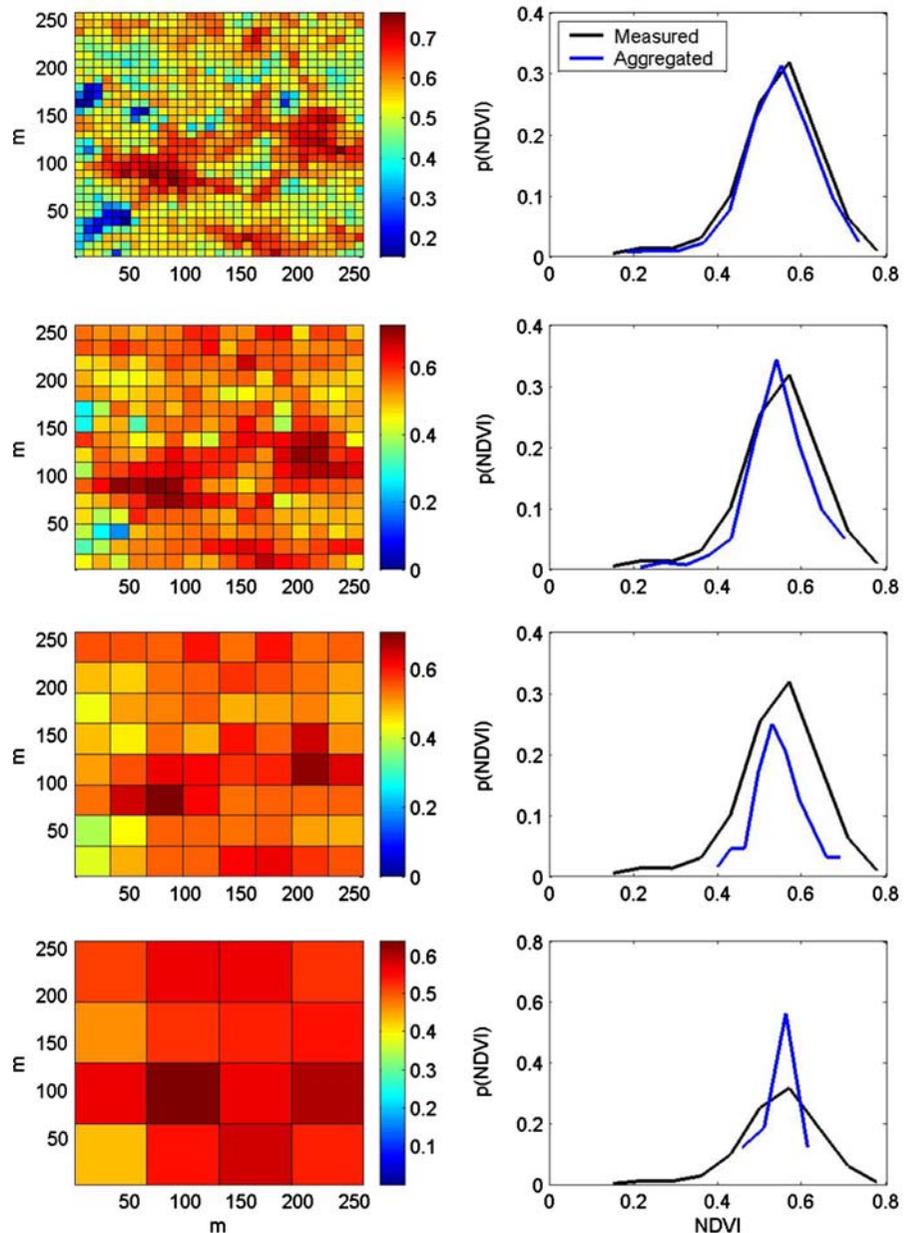

The analysis of spatial aggregation suggests that preserving the variance of the fine-scale NDVI distribution will minimize bias in C flux estimates (Fig. 2). This conclusion is similar to that of Liang (2001) for NDVI-LAI scaling, but is variance retention enough to ensure accuracy in the upscaled NDVI distribution and thereby the landscape-level NEE estimate? The pdfs of measured NDVI and corresponding beta, normal and skew-normal distributions

are shown in Fig. 3a and corresponding maximum-likelihood parameter estimates are listed in Table 2. Transferring the resulting NDVI distributions using Eq. 5 resulted in distributions of LAI estimates (Fig. 3b), and, combined with meteorological data, distributions of corresponding PLIRTLE model output (Fig. 3c), the sum of which represents the total flux for the growing season over the study domain (Table 3). It was demonstrated in the previous

Table 1 Deviation from fine-scale NEE estimates using ATM-observed NDVI after spatial aggregation of NDVI (see Fig. 1)

Pixel size (m^2)	NEE (kg C gs^{-1})	GEP (kg C gs^{-1})	RE (kg C gs^{-1})
16 (fine-scale)	−371	−3,482	3,112
64	−249	−3,316	3,067
256	−107	−3,122	3,015
1,024	50	−2,909	2,960
4,096	187	−2,724	2,910
16,384	325	−2,536	2,862
65,536	355	−2,496	2,851

Fig. 2 The relationship between **a** spatial averaging operator length scale and NDVI variance (σ^2_{NDVI}), **b** NDVI variance and landscape-level NEE estimates over the 65,536 m^2 study domain, **c** pixel length scale and NEE, and **d** the Kullback-Liebler divergence (D_{KL}, Eq. 10) between measured and aggregated NDVI and corresponding NEE estimates for an upscaling exercise using aircraft-measured NDVI and PLIRTLE (Shaver et al. 2007)-modeled NEE in a tundra ecosystem near Abisko, Sweden
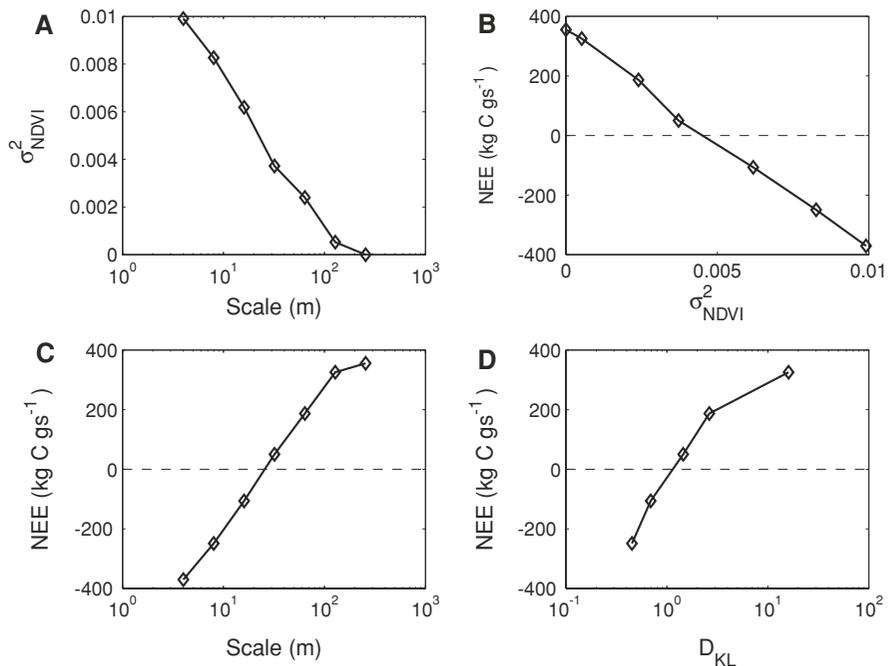
analysis that merely preserving the mean measured NDVI resulted in a large bias error and a sign shift in the flux estimates from C sink to source (Table 1). Preserving both the mean and variance using the normal distribution resulted in ca. 23% NEE error (Table 3), but adding the skewness via the skew-normal distribution reduced NEE error to <4%. GEP and RE were biased by only ca. 3 and 1% by assuming that NDVI was normally distributed, but modeling NDVI using the skew-normal distribution resulted in GEP and RE estimates that were biased by 0.5% or less. The selection of the beta distribution did not result in an improvement over the normal or skew-normal distributions (Table 3).

Handheld sensor NDVI measurements

Introducing the spatially variable NDVI-LAI transfer function parameters (Williams et al. 2008) to the NEE estimates that resulted from applying the PLIRTLE model to handheld NDVI measurements resulted in lower bias at larger scales (Fig. 4, using variance as a surrogate for scale after Fig. 3). The 'best' NEE estimate from the fine-scale observed NDVI is −38 g C gs^{-1} for the smaller 900 m^2 study domain. The NEE estimate using the static NDVI-LAI equation (i.e., Eq. 5) was biased by 17% at the coarsest scale. The coarsest-scale NEE estimate when using the NDVI-LAI transfer functions with scale-variant
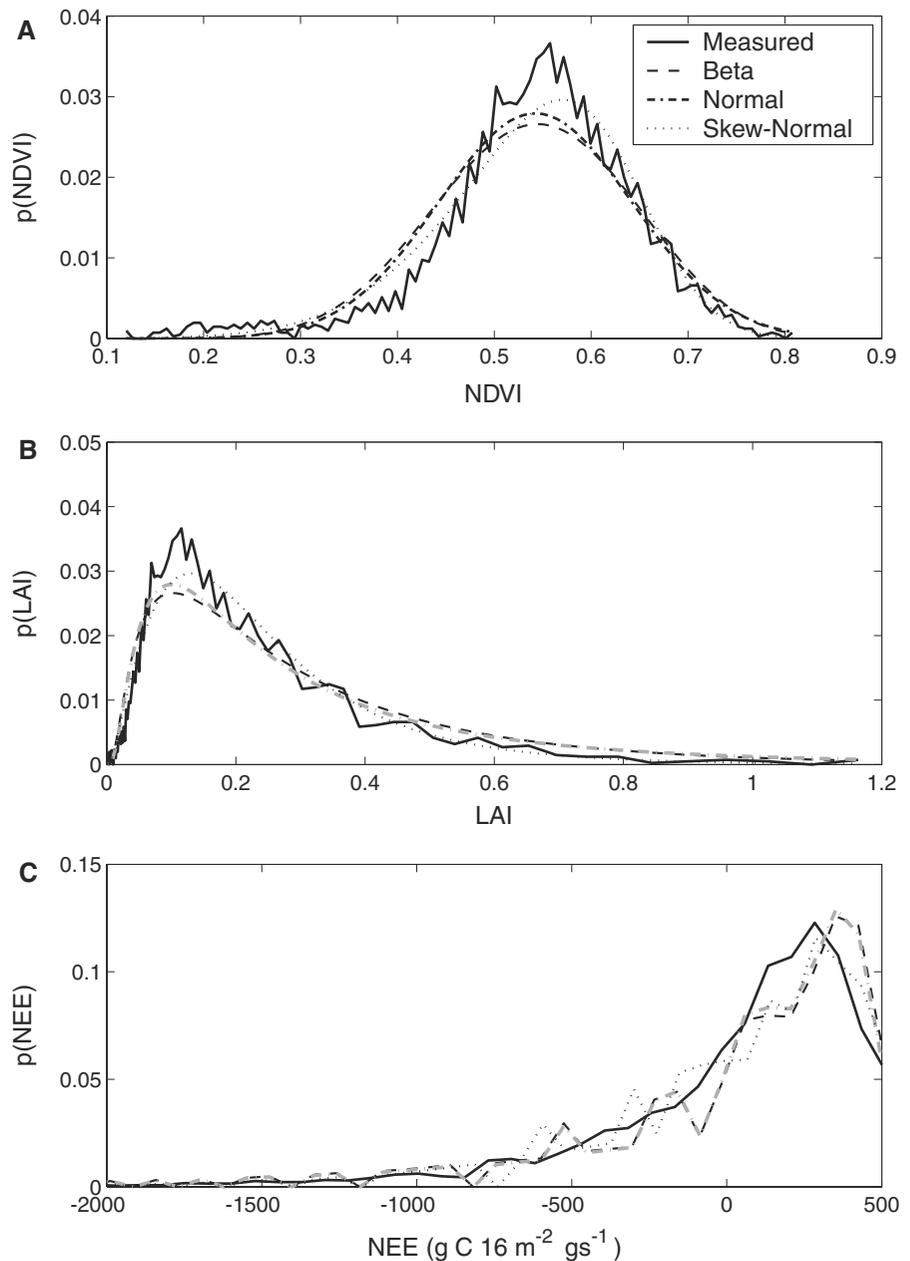
parameters was biased by <10%. Less than 3% error in NEE, GPP and RE estimates was incurred by fitting the fine-scale handheld NDVI data using beta, normal, or skew-normal distributions (data not shown).

NDVI image compression

Selected outputs from the 2D wavelet thresholding analysis (image compression) are shown in Fig. 5.

Modes of spatial variance that contribute little to the total variance of the original image are progressively filtered out such that the fundamental spatial patterns of the image that contribute disproportionally to the total spatial variance are retained. The p(NDVI) that results from 50, 85 and 95% NDVI image compression for Haar and Coiflet wavelet compression (i.e., those thresholds shown in Fig. 5) are displayed in Fig. 6, where it is evident that 95%

Table 2 Maximum likelihood parameter estimates for the beta, normal, and skew-normal distributions used to fit the distribution of observed normalized difference vegetation index (NDVI) shown in Figs. 1, 3a, and 6

Beta	
α	12.6
β	10.7
Normal	
μ	0.542
σ^2	0.100
Skew-normal	
ξ	0.645
ω^2	0.0203
α_S	−2.19

compression with the Haar basis function deviates substantially from the shape of measured p(NDVI), noting also the removal of high NDVI values (Fig. 6a). Compression using the Coiflet basis function generates a distribution with progressively higher kurtosis than the measured NDVI distribution,

concentrating the distribution toward the mean (Fig. 6b). The alteration of σ^2_{NDVI}, D_{KL}, NEE, and NEE bias with progressive image compression using both wavelet basis functions are demonstrated in Fig. 7. Compression using the Coiflet wavelet basis function resulted in lower bias per percent spatial variance removed than Haar wavelet compression; e.g., a bias of 50 kg C gs^{-1} (13%) for the study area could be incurred by removing ca. 70% of the Haar wavelet coefficients and 75% of the Coiflet wavelet coefficients (Fig. 7d).

Discussion

Scaling information

The relationship between scale and variance resembles a logarithmic scaling relationship across the range of scales observed (Fig. 2c). This result suggests that scaling laws in biology (Enquist et al. 2007; West et al. 1997) and ecology (Enquist

Table 3 NEE estimates from the PLIRTLE model (Shaver et al. 2007) after modeling the aircraft-measured NDVI distribution using beta, normal or skew-normal distributions, the parameters for which are listed in Table 2

Source	NEE (kg C gs^{-1})	NEE error (%)	GEP (kg C gs^{-1})	GEP error (%)	RE (kg C gs^{-1})	RE error (%)
Fine-scale	−371		−3,482		3,112	
Beta	−461	24.5	−3,609	3.6	3,148	1.2
Normal	−456	22.9	−3,601	3.4	3,146	1.1
Skew-normal	−356	−3.7	−3,465	−0.5	3,108	−0.12

Fig. 4 Net ecosystem exchange (NEE) estimates for a 900 m^2 expanse of tundra vegetation near Abisko, Sweden after spatial averaging with either static or spatially variable parameters of the NDVI-LAI transfer function (van Wijk and Williams 2005; Williams et al. 2008) using handheld NDVI measurements. The variance of measured NDVI is 7.1×10^{-3}
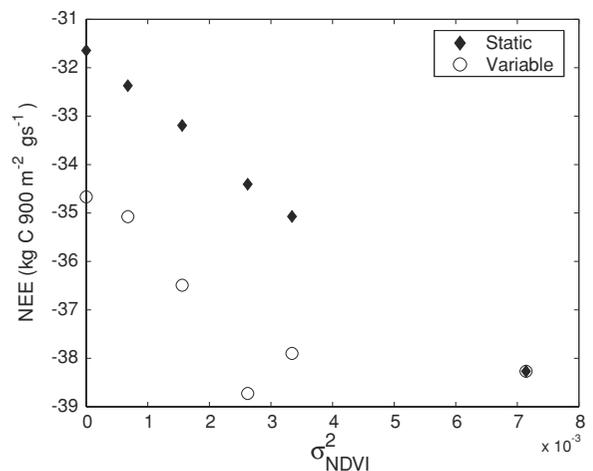
Fig. 5 Two-dimensional Haar (**a–d**) and Coiflet (**e–h**) wavelet compression of a map of normalized difference vegetation index (NDVI) measured using an Airborne Thematic Mapper (ATM) in a tundra landscape near Abisko, Sweden
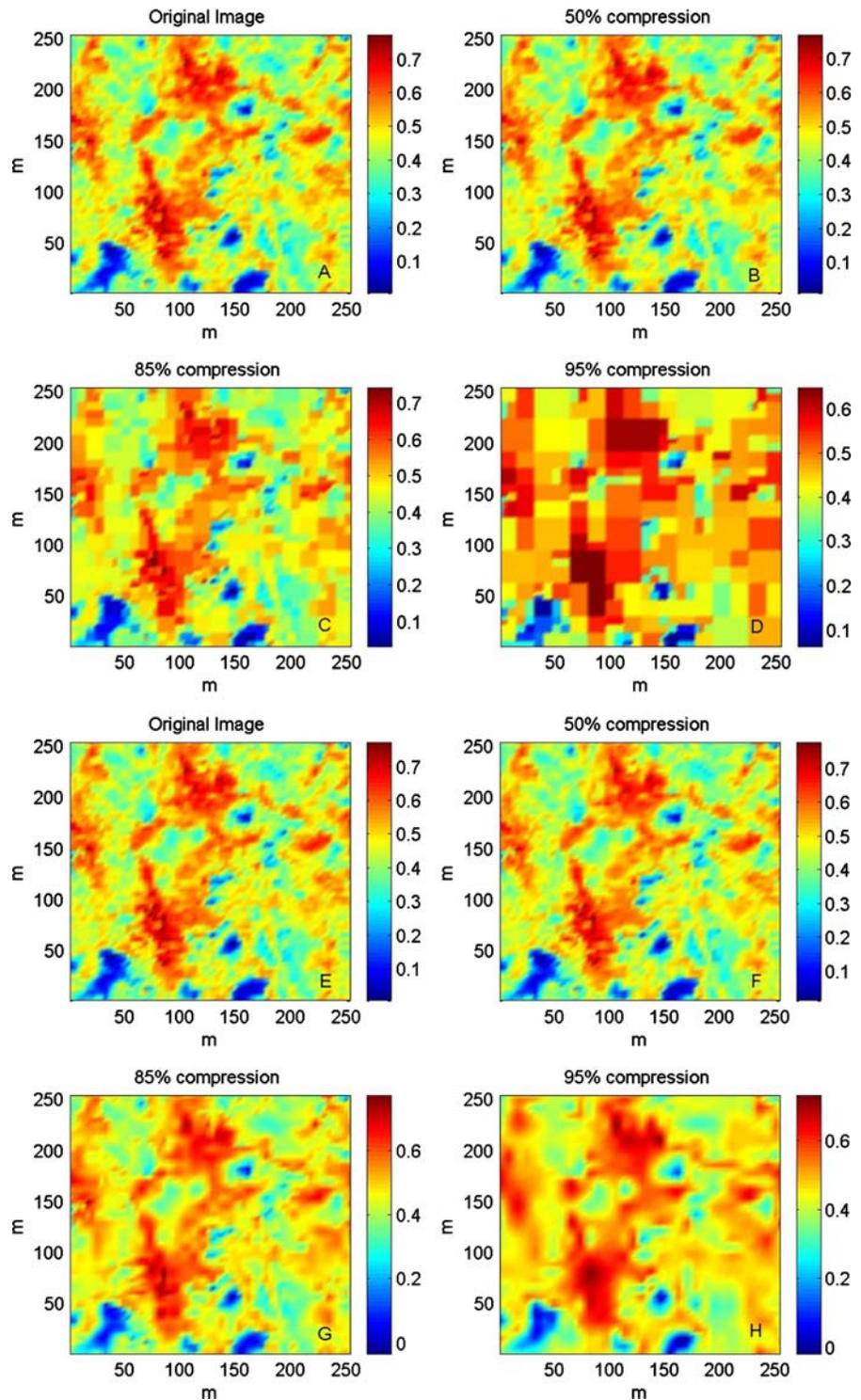
et al. 2003) may be applied to reduce bias when upscaling or downscaling estimates of land-surface processes in our case. If the spatial grain, representing

the minimum important ecological length scale, is known, an improved estimate of the unit of interest (in this case NEE, GEP and RE) might be found.

Fig. 6 The probability density function of measured and **a** 2D Haar wavelet and **b** 2D Coiflet wavelet compressed NDVI [p(NDVI)] for the images shown in Fig. 5
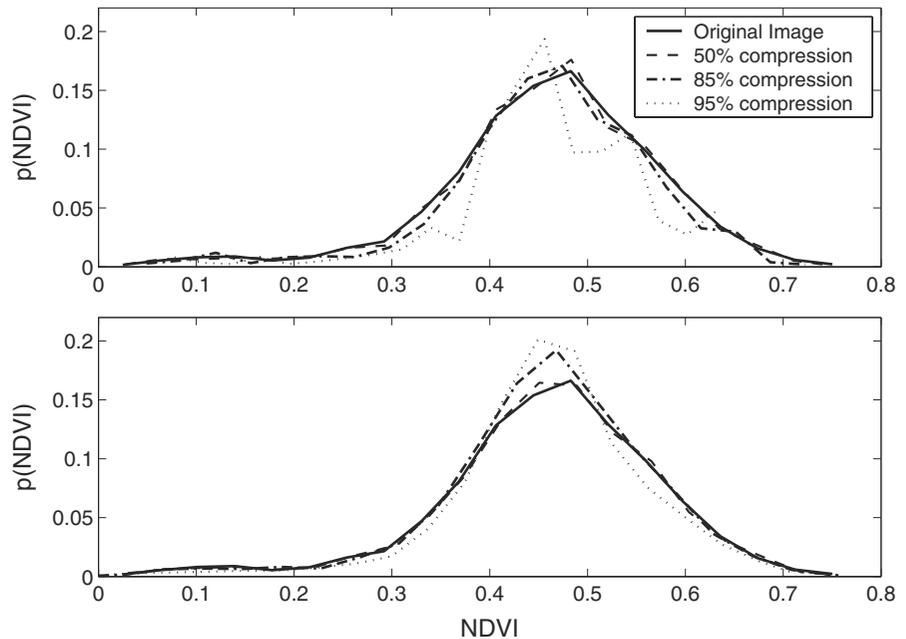
Stoy et al. (2009), using data from van Wijk and Williams (2005) and Williams et al. (2008), demonstrated a rapid change in the D_{KL} of aggregated tundra LAI at a mean pixel size of ca. 10 m^2, i.e., a scale of $10^{1/2} = 3.16$ m. The grain size of the ATM image is 4 m. Downscaling the NEE estimates from 4 to 3.16 m using the empirical relationship in Fig. 2c, supported from the standpoint of information theory via the D_{KL} (Fig. 2d), results in a NEE estimate of −429 kg C gs^{-1} for the airborne study domain, a further increase in sequestration of 16% from the −371 kg C gs^{-1} estimate derived from NEE derived from ATM-measured NDVI (Table 1).

Is the downscaled estimate an improvement over that derived from the ATM data at a scale of 4 m? This modeling analysis cannot answer that question, but the length scale of the study domain, 256 m (Fig. 2), overlaps with the range of length scales of the turbulent eddy covariance flux footprint at a nearby flux tower (Fox et al. 2008). A bias of ca. 16% is on the order seasonal eddy covariance flux error estimates across a range of ecosystems (Goulden et al. 1996; Moncrieff et al. 1996; Stoy et al. 2006), such that the modeled discrepancy between 3.16 and 4 m grain sizes can arguably be ignored given that potential differences approximate the range of eddy covariance flux measurement error.

Scale-wise information preservation

If information is altered by aggregation across scale [Chen and Blong (2002); Stoy et al. (2009), Fig. 2c], it follows that preserving information via the parameters of the distribution that describes this information (Rastetter et al. 1992) will reduce bias. Parameters of statistical distributions are the elements by which information can be formally quantified [e.g., Eq. 10, Kullback (1997)]. For our example of NDVI upscaling, the first three moments of the normal distribution are required to minimize the scale-dependence of the flux estimate. The result that <4% NEE error was incurred by choosing the three parameter skew-normal distribution suggests that incorporating higher order moments (Rastetter et al. 1992) may add minimally to the skill of the upscaling technique here. Despite the fact that NDVI is defined on a bounded interval between 0 and 1, choosing the Beta distribution did not improve the estimate because it poorly replicated the tails of the measured NDVI distribution compared to the skew-normal distribution (Fig. 3a). The higher-order moments of ATM-measured NDVI departed from normality as revealed by a Jarque-Bera test: the null hypothesis that p(NDVI) is drawn from a normal distribution with unit skewness and zero excess kurtosis can be rejected at the 95% confidence level, suggesting that

Fig. 7 The alteration of **a** total NDVI variance σ^2_{NDVI}, **b** the Kullback–Liebler divergence for the normal distribution (D_{KL}, Eq. 10), **c** NEE estimates from the PLIRTLE model (Shaver et al. 2007) and **d** NEE bias that result from progressive NDVI image compression using the 2D Haar wavelet basis function
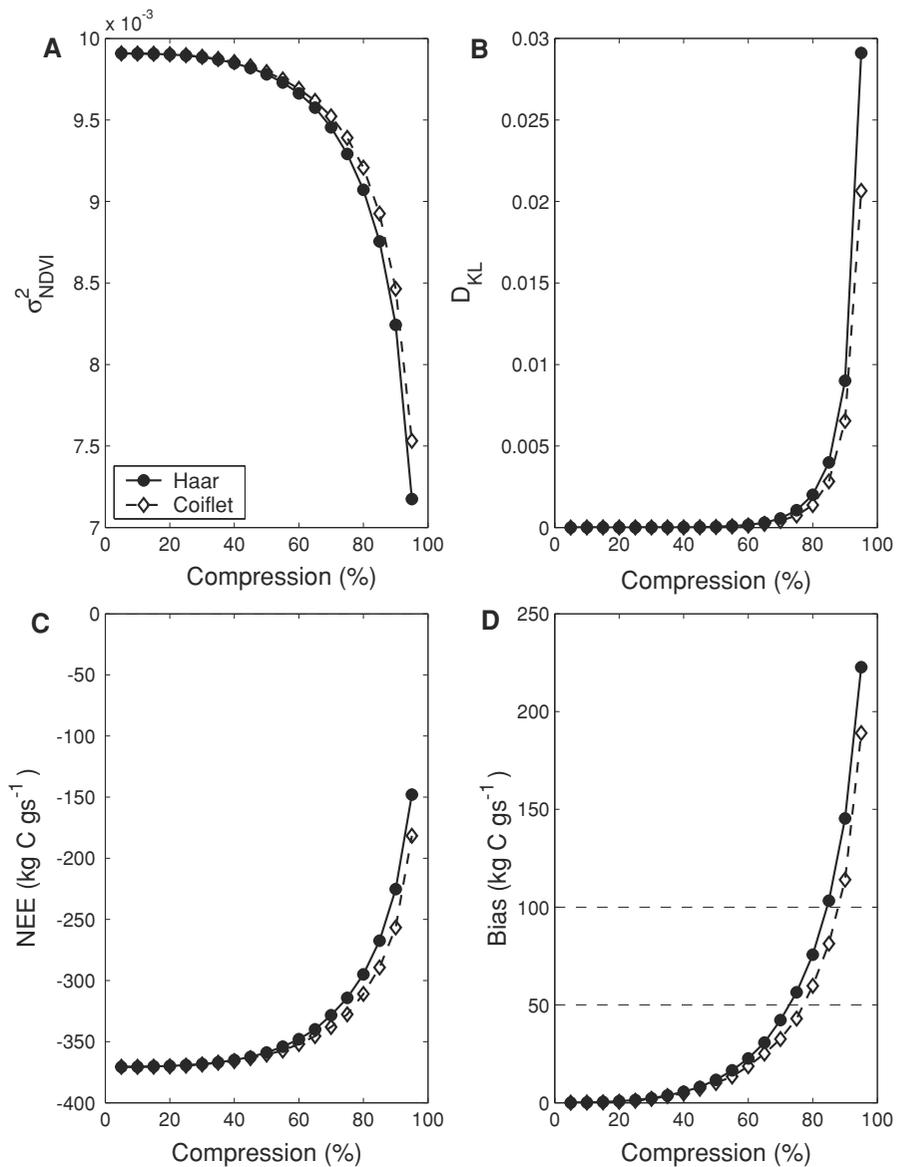
the normal distribution is insufficient for modeling observed NDVI. The choice of a statistical distribution that departs from normality is logical for modeling the distribution of ATM-measured NDVI here.

In essence, an upscaling technique that preserves the tails of the NDVI distribution is disproportionately important in our case, because these values correspond to patches with extremely high or near zero LAI which represent hotspots or dead zones for C flux. Relatively high NDVI translates to extremely high LAI in tundra owing to the nonlinearity of the transfer function (van Wijk and Williams 2005) and

therefore large fluxes because the relationship between LAI and NEE in PLIRTLE is approximately linear [Fig. 8a in Stoy et al. (2009)]. This nonlinear upscaling challenge is best met at this site by modeling the measured NDVI distribution as skew-normal to preserve the important information about the distribution while minimizing data storage requirements with only three parameters.

Image compression

If preserving spatial relationships is required for accurate upscaling, it follows that retaining all

elements of the spatial variance that fully describe pattern across scale is probably not necessary. In our case, ca. 70–75% of the spatial wavelet coefficients, depending on wavelet basis function, could be removed from the measured NDVI map to arrive at a NEE estimate that deviated by less than 50 kg C gs^{-1}, i.e., 13%, approximating eddy covariance measurement error, from the original. Image compression applications are primarily concerned with reducing digital information to retain an image that appears unaltered to human senses; lossy compression of digital data typically amounts to reducing file size while maintaining sufficient information content for a given application, where the sufficiency criterion is often an arbitrary, or heuristic, choice. Here, we demonstrate that ecological or land-surface information can likewise be compressed, and that models can be used to determine how much information is required for process-based studies. We envision that ecological information preservation will be useful for designing field experiments and sampling protocols in addition to simultaneously minimizing both bias and computational load for multi-scale studies.

Assimilating remotely sensed data into ecological models

We demonstrated that preserving the variance of measured land-surface observations can increase modeling skill over merely preserving the mean. Preserving variance, higher order moments, or parameters of other, non-normal statistical distributions makes it possible to assimilate land surface observations into ecological models (Quaife et al. 2008) because data assimilation entails the combination of information via pdfs. The added benefit is that assimilating data into ecological models can further reduce errors in process estimates (Williams et al. 2005). Measurement error, often expressed in terms of variance, is just as important as a measurement when employing Bayesian techniques for ecosystem modeling (Clark 2005). For example, recursive Bayesian estimation via the Kalman filter assumes normally distributed error terms on all measurements (Williams et al. 2005), and simply incorporating NDVI variance reduces error by nearly 175% over merely preserving the mean, a priori, before the data assimilation step. An emerging challenge in ecology

is to incorporate a broader set of probability distributions into the data assimilation step, preferably classes of conjugate distributions that enable the use of Hierarchical Bayes for comprehensive multi-scale models of ecological function (Ogle and Barber 2008).

Techniques for reducing nonlinearity

The NDVI-LAI transfer functions of van Wijk and Williams (2005) and Williams et al. (2008) are extreme examples of nonlinearity in an upscaling procedure, but it is the best available technique for relating remotely sensed indices to land-surface properties in tundra at the present (Boelman et al. 2003; van Wijk and Williams 2005; Williams et al. 2008). The as-yet unexplored relationship between enhanced vegetation index EVI and LAI (Boegh et al. 2002; Jiang et al. 2008; Liu and Kafatos 2005) in tundra may be a logical way to decrease bias if the degree of nonlinearity in the EVI–LAI relationship is lower than that of NDVI-LAI (Rocha and Shaver 2009). Transfer functions to estimate LAI from EVI, or other indices (Verstraete and Pinty 1996) such as the scaled difference vegetation index (Jiang et al. 2006), have yet to be developed for tundra using direct LAI measurements (but see Rocha and Shaver 2009), but may be critical for reducing bias in future studies of multi-scale C flux estimates in tundra ecosystems that are characterized by pronounced spatial heterogeneity (Spadavecchia et al. 2008; Walker et al. 1995). Transforming remote sensing observations to reduce nonlinearity with respect to measured LAI may also be a logical step for reducing bias.

Applicability to other ecological studies

Our argument is that by conceptualizing upscaling as a transfer of information across scales enables mathematical techniques from Information Theory to be applied to the upscaling problem of interest. In this sense our approach, if not the specific study results, should be generalizable across larger (or smaller) scales and in different ecosystem types. A critical test for any novel application is its ability to pass some criteria for success. In our example, conserving information in an observation (NDVI) that must be converted to a model input (LAI) reduced bias in the model output (NEE). We

anticipate that other problems of nonlinear scaling in Ecology across different ecosystem types or spatial scales may likewise benefit by conserving information in cases where aggregation alters information content.

Conclusions

It was demonstrated that preserving three parameters that define the distribution of aircraft-measured NDVI, the mean, variance, and skewness, were required to minimize bias in upscaled tundra NEE to <4%. Given the amount of information about surface features that is able to be extracted (Verstraete et al. 1996), preserving more information about the land surface than merely the mean was shown to be highly advantageous in this example of multi-scale C flux modeling. Introducing scale-dependent NDVI-LAI transfer functions were shown to additionally reduce bias in an example at smaller spatial scales, but may be difficult to parameterize at large scales (Williams et al. 2008). A large percentage of spatial variance can be removed from the measured NDVI map without appreciably biasing the resulting NEE estimate from the PLIRTLE model, as evidenced by the wavelet-based image compression analysis, but bias increased sharply as the amount of spatial variance removed approached unity. Given that NEE bias empirically resembles a scaling law in our case, NEE estimates were generated at the characteristic length scales of spatial variance for the study tundra ecosystem as a means to improve multi-scale flux estimates, which require further validation. Such relationships may be useful when scaling NEE from leaf to plot (e.g., eddy covariance NEE estimates) to region (e.g., aircraft or tall tower NEE estimates). In addition to minimizing scale-wise bias in flux estimates, preserving information content additionally enables the application of spatial data assimilation techniques (Quaife et al. 2008) to further improve multi-scale estimates of CO_2 flux in an information-conserving upscaling framework.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Contr 19:716–723. doi:10.1109/TAC.1974.1100705

Azzalini A (1985) A class of distributions which includes the normal ones. Scand J Stat 12:171–178

Boegh E, Soegaard H, Broge N, Hasager CB, Jensen NO, Schelde K (2002) Airborne multispectral data for quantifying leaf area index, nitrogen concentration and photosynthetic efficiency in agriculture. Remote Sens Environ 81:179–193. doi:10.1016/S0034-4257(01)00342-X

Boelman NT et al (2003) Response of NDVI, biomass, and ecosystem gas exchange to long-term warming and fertilization in wet sedge tundra. Oecologia 135:414–421

Brunsell NA, Young CB (2007) Land surface response to precipitation events using MODIS and NEXRAD data. Int J Remote Sens 29:1965–1982. doi:10.1080/01431160701373747

Brunsell NA, Ham JM, Owensby CE (2008) Assessing the multi-resolution information content of remotely sensed variables and elevation for evapotranspiration in a tallgrass prairie environment. Remote Sens Environ 112:2977–2987. doi:10.1016/j.rse.2008.02.002

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach, 2nd edn. Springer, New York

Chang SG, Yu B, Vetterli M (1997) Image denoising via lossy compression and wavelet thresholding. In: Proceedings of the International Conference of Image Processing, Santa Barbara

Chen K, Blong R (2002) Integrating remotely sensed images and areal census data for building new models across scales. Geoscience and Remote Sensing Symposium, 2002. IGARSS '02. 2002 IEEE International vol 4, pp. 2385–2387

Clark JS (2005) Why environmental scientists are becoming Bayesians. Ecol Lett 8:2–14. doi:10.1111/j.1461-0248.2004.00702.x

Enquist BJ, Economo EP, Huxman TE, Allen AP, Ignace DD, Gillooly JF (2003) Scaling metabolism from organisms to ecosystems. Nature 423:639–642. doi:10.1038/nature01671

Enquist BJ, Kerkhoff AJ, Stark SC, Swenson NG, McCarthy MC, Price CA (2007) A general integrative model for scaling plant growth, carbon flux and functional trait spectra. Nature 449:218–222. doi:10.1038/nature06061

Fox AM, Huntley B, Lloyd CR, Williams M, Baxter R (2008) Net ecosystem exchange over heterogeneous Arctic tundra: scaling between chamber and eddy covariance measurements. J Geophys Res 22:GB2027

Goulden ML, Munger JW, Fan S, Daube BC, Wofsy S (1996) Measurements of carbon sequestration by long-term eddy

covariance: methods and a critical evaluation of accuracy. Glob Change Biol 2:169–182. doi:10.1111/j.1365-2486.1996.tb00070.x

Jiang Z et al (2006) Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction. Remote Sens Environ 101:366–378. doi:10.1016/j.rse.2006.01.003

Jiang Z, Huete AR, Didan K, Miura T (2008) Development of a two-band enhanced vegetation index without a blue band. Remote Sens Environ 112:3833–3845. doi:10.1016/j.rse.2008.06.006

Kessler MA, Werner BT (2003) Self-organization of sorted patterned ground. Science 299:380–383. doi:10.1126/science.1077309

Kolaczyk ED, Ju J, Gopal S (2005) Multiscale, multigranular statistical image segmentation. J Am Stat Assoc 100:1358–1369. doi:10.1198/016214505000000385

Krantz WB (1990) Self-organization manifest as patterned ground in recurrently frozen soils. Earth Sci Rev 29:117–130

Kullback S (1997) Information theory and statistics, 2nd edn. Dover Publications, Mineola

Kullback S, Leibler RA (1951) On information and sufficiency. Ann Math Stat 22:79–86. doi:10.1214/aoms/1177729694

Laurance WF (2000) Do edge effects occur over large spatial scales. Trends Ecol Evol 15:134–135. doi:10.1016/S0169-5347(00)01838-3

Liang S (2001) Numerical experiments on spatial scaling of land surface albedo and leaf area index. Remote Sens Rev 19:225–242

Liu X, Kafatos M (2005) Land-cover mixing and spectral vegetation indices. Int J Remote Sens 26:3321–3327. doi:10.1080/01431160500056907

MacArthur R (1955) Fluctuations of animal populations, and a measure of community stability. Ecology 36:533–536. doi:10.2307/1929601

Mandal MK, Panchanathan S, Aboulnasr T (1996) Choice of wavelets for image compression. LNCS 1133:239–249

Moncrieff JB, Mahli Y, Leuning R (1996) The propagation of errors in long-term measurements of land-atmosphere fluxes of carbon and water. Glob Change Biol 2:231–240. doi:10.1111/j.1365-2486.1996.tb00075.x

O'Neill RV, Rust B (1979) Aggregation error in ecological models. Ecol Model 7:91–105. doi:10.1016/0304-3800(79)90001-2

Ogle K, Barber JJ (2008) Bayesian data-model integration in plant physiological and ecosystem ecology. In: Lüttge U, Beyschlag W, Murata J (eds) Progress in botany, vol 69. Springer, Berlin

Pelgrum H (2000) Aggregation of a nonlinear land surface model for heterogeneous terrain. In: Owe M, Brubaker K, Ritchie J, Rango A (eds) Remote sensing and hydrology, vol 267. IAHS, Santa Fe, pp 181–186

Quaife T et al (2008) Assimilating canopy reflectance data into an ecosystem model with an ensemble Kalman filter. Remote Sens Environ 112:1347–1364. doi:10.1016/j.rse.2007.05.020

Rastetter EB, King AW, Cosby BJ, Hornberger GM, O'Neill RV, Hobbie JE (1992) Aggregating fine-scale ecological knowledge to model coarser-scale attributes of ecosystems. Ecol Appl 2:55–70. doi:10.2307/1941889

Reza FM (1994) An introduction to information theory. Dover Publications, Mineola, NY, pp. 496

Rocha AV, Shaver GR (2009) Advantages of a two band EVI calculated from solar and photosynthetically active radiation fluxes. Agric For Meteorol. doi:10.1016/j.agrformet.2009.03.016

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464. doi:10.1214/aos/1176344136

Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423 and 623–656

Shaver GR, Street LE, Rastetter EB, van Wijk MT, Williams M (2007) Functional convergence in regulation of net CO_2 flux in heterogeneous tundra landscapes in Alaska and Sweden. J Ecol 95:802–817. doi:10.1111/j.1365-2745.2007.01259.x

Spadavecchia L, Williams M, Bell R, Stoy PC, Huntley B, van Wijk MT (2008) Topographic controls on the leaf area index of a Fennoscandian tundra ecosystem. J Ecol 96:1238–1251. doi:10.1111/j.1365-2745.2008.01424.x

Stoy PC, Katul GG, Siqueira MBS, Juang J-Y, Novick KA, Oren R (2006) An evaluation of methods for partitioning eddy covariance-measured net ecosystem exchange into photosynthesis and respiration. Agric For Meteorol 141:2–18. doi:10.1016/j.agrformet.2006.09.001

Stoy PC et al. (2009) Using information theory to determine optimum pixel size and shape for ecological studies: application to leaf area index aggregation in arctic ecosystems. Ecosystems 12:574–589

Street LE, Shaver GR, Williams M, van Wijk MT (2007) What is the relationship between changes in canopy leaf area and changes in photosynthetic CO2 flux in arctic ecosystems? J Ecol 95:139–150. doi:10.1111/j.1365-2745.2006.01187.x

Turner MG (1989) Landscape ecology: the effect of pattern on process. Annu Rev Ecol Syst 20:171–197. doi:10.1146/annurev.es.20.110189.001131

Ulanowicz RE (1980) An hypothesis on the development of natural communities. J Theor Biol 85:223–245. doi:10.1016/0022-5193(80)90019-3

Ulanowicz RE (2001) Information theory in ecology. Comput Chem 25:393–399. doi:10.1016/S0097-8485(01)00073-0

van Wijk MT, Williams M (2005) Optical instruments for measuring leaf area index in low vegetation: application in arctic ecosystems. Ecol Appl 15:1462–1470. doi:10.1890/03-5354

Verstraete M, Pinty B (1996) Designing optimal spectral indexes for remote sensing applications. IEEE Trans Geosci Remote Sens 34:1254–1265. doi:10.1109/36.536541

Verstraete M, Pinty B, Myneni RB (1996) Potential and limitations of information extraction on the terrestrial biosphere from satellite remote sensing. Remote Sens Environ 58:201–214. doi:10.1016/S0034-4257(96)00069-7

Walker DA, Auerbach NA, Lewis BE, Shippert MM (1995) NDVI, biomass, and landscape evolution of glaciated terrain in northern Alaska. Polar Rec (Gr Brit) 31:169–178

West GB, Brown JH, Enquist BJ (1997) A general model for the origin of allometric scaling laws in biology. Science 276:122–126. doi:10.1126/science.276.5309.122

Williams M, Schwarz PA, Law B, Irvine J, Kurpius MR (2005)
       An improved analysis of forest carbon dynamics using
       data assimilation. Glob Change Biol 11:89–105. doi:
       10.1111/j.1365-2486.2004.00891.x

Williams M, Bell R, Spadavecchia L, Street LE, van Wijk MT
       (2008) Upscaling leaf area index in an Arctic landscape
       through multi-scale observations. Glob Change Biol
       14:1517–1530. doi:10.1111/j.1365-2486.2008.01590.x