PRIMARY PRODUCTION OF AN ARCTIC WATERSHED: AN UNCERTAINTY ANALYSIS

MATHEW WILLIAMS,1 EDWARD B. RASTETTER, GAUIS R. SHAVER, JOHN E. HOBIE, ELIZABETH CARPINO, AND BONNIE L. KWiatkowski
The Ecosystems Center, Marine Biological Laboratory, Woods Hole, Massachusetts 02543, USA

Abstract. We describe a scaling protocol that combines two hierarchically linked models with field surveys, spatially distributed weather data, and remotely sensed images to generate daily predictions of gross primary production (GPP) for a 9200-km² arctic watershed. A detailed process-based model of vegetation–atmosphere interactions, which has been tested in a variety of arctic ecosystems against independent hourly gas exchange data, forms the base of the hierarchy. This detailed model was used to construct a second and simpler, “big-leaf” model, which was calibrated for arctic conditions and which required many fewer parameters and input data. For landscape forcing data, we derived spatiotemporal data on weather conditions (maximum and minimum temperature, and irradiance) from weather stations throughout the watershed. Spatiotemporal descriptions of the biotic constraints on production, chiefly leaf area index (LAI) and total foliar nitrogen (N), were derived from field surveys, a land cover database, and normalized difference vegetation index (NDVI) data acquired from satellites. The scaling hierarchy avoided propagation of error via a compensation process, though the procedures involved still introduced uncertainty into daily GPP predictions averaging 16% of the growing season daily mean. The construction of the spatiotemporal temperature and irradiance fields introduced uncertainty of 1–2% at spatial and temporal resolutions of 1 km² and one day, respectively. The greatest uncertainty was introduced by assignment of LAI across the region, because of the highly heterogeneous landscape and the high sensitivity of production to changes in LAI at the low levels found in the Arctic. Uncertainty in vegetation properties introduced an uncertainty of ±15% in basin GPP predictions, a value commensurate with basin net ecosystem production (NEP). In conclusion, improved characterization of vegetation via remote sensing is required before any bottom-up approach to carbon budgeting can reduce uncertainty to a reasonable level.

Key words: arctic tundra watershed; big-leaf model; ecosystem model; gross primary productivity; leaf area index; net ecosystem production; normalized difference vegetation index; photosynthesis; scaling; total foliar nitrogen; uncertainty analysis; vegetation–atmosphere interactions.

INTRODUCTION

A major limit to understanding the earth as a system, and its responses to anthropogenic forcing, is that the scale at which understanding is required is far removed from the scale at which process knowledge is most advanced. Thus, regional and global problems have to be understood in terms of processes studied at a limited number of intensive, small-scale field sites. The goal of this paper is to bridge this mismatch of scales. First, we describe methods to scale up, extending detailed, process-based understanding across time and space, by using remote sensing and a number of linked models and data sets. And second, we undertake a detailed uncertainty analysis to identify the uncertainties in the predictions introduced by the scaling methods.

In this study, the spatial focus is on a 9200-km² arctic watershed in northern Alaska, the site of intensive field and modeling studies over past decades (Weller et al. 1995, Kane and Reeburgh 1998). The functional focus is on the terrestrial gross primary production (GPP), a key component of the global carbon cycle. The primary questions are (1) how does GPP vary both spatially and temporally within the watershed, and (2) how precisely can GPP be predicted? This research complements previous work examining spatial variation across the same watershed in carbon–nutrient interactions (Hobbie et al. 1998), active layer depth (Hinzman et al. 1998), and methane emissions (Reeburgh et al. 1998). Studies of the carbon balance of arctic tundra ecosystems have identified some as C sinks and some as sources over past decades (Oechel et al. 1993, Vourlitis and Oechel 1997, 1999, Walker et al. 1998), depending on the balance between C storage, governed by photosynthesis, and C losses via respiration. Because of the vast C stocks incorporated in arctic soils, measuring changes in stocks of organic C over annual time scales is not feasible; measurement errors are likely to be larger than interannual variability. The alternative is to

1 Present address: Institute of Ecology and Resource Management, University of Edinburgh, Edinburgh EH9 3JU, UK. E-mail: mat.williams@ed.ac.uk
measure net C fluxes via gas exchange at a variety of sites, the goal of the Arctic Flux Study (Kane and Reeburgh 1998). But, these data must then be scaled up to generate the regional picture. While the net C flux is ultimately of interest, a direct extrapolation is fraught with risks due to nonlinearities. Instead, regional predictions need to take account of how the two key processes, photosynthesis and respiration, vary over the landscape, and then determine net fluxes as the sum of the two.

Scaling protocols predict rates of a process across space and time according to variation in constraining factors. The scaling approach we applied used several steps. In the first step, we took process information at the stomatal and cellular levels to model hourly fluxes at the leaf scale, in multiple canopy layers. In the second step, we aggregated the results by constructing a daily productivity response surface, and fitted a simple function to mimic the shape of this surface. This process resulted in an aggregated-canopy model (ACM; Williams et al. 1997) of daily GPP that operates considerably faster than the leaf-scale model and has simpler input requirements. In the third step, described here, we implemented ACM in landscape mode, to generate predictions throughout the basin and across the 1995 growing season. To generate the required spatial fields to drive ACM, we used field sampling, surveys, remotely sensed data, and data from weather stations throughout the study watershed. We quantified uncertainty in GPP prediction based on precision of driving variables, scaling assumptions, and model accuracy.

Earlier field studies provided important underpinnings for the scaling protocol (Williams and Rastetter 1999). In a survey of 14 tundra sites, we found that although the dominant species varied, there was at each site a highly significant relationship for vascular plants between the leaf area index (LAI) and total foliar nitrogen per ground area, these being the critical biotic controls on production. The conservative nature of canopy LAI:N ratios simplifies the task of generating regional predictions of C assimilation by vascular plants, for which LAI and N are key requirements, because remote estimates of LAI can be used to infer N.

**STUDY AREA**

This study is an extension of the Arctic Flux Study (Weller et al. 1995, Kane and Reeburgh 1998), which undertook intensive field measurements during 1994–1996 along a north–south transect between 68° and 70° N, within, or near to, the 9219-km² Kuparuk River watershed in northern Alaska (Fig. 1). The range in elevation within the basin is from sea level in the north to 1489 m in the south. The vegetation within the watershed is characterized of a large part of the circumpolar tundra zone, containing five major physiographic categories (Walker et al. 1995, Walker and Walker 1996): heath tundra, tussock tundra, wetland tundra, and shrub tundra.

**MODELS**

*The soil–plant–atmosphere model (SPA)*

The soil–plant–atmosphere model (see Williams et al. (1996) for a full description) is a multilayer simulator of C₃ vascular plant processes operating at a 30-min time step. Modeled ecosystem structure is described by vertical variations among canopy layers in light absorbing area, photosynthetic capacity (related to foliar N), and plant hydraulic properties. We made only four changes to the model parametrization originally applied to temperate forests (Williams et al. 1996): (1) To account for acclimation to the cool arctic climate, we reset the optimum temperature for the photosynthesis model (Farquhar and von Caemmerer 1982) from 25°C to 18°C, close to measured optima for arctic plants, and ensured that photosynthetic rates at 5°C were 60–65% of their maxima (Semikhatova et al. 1992, Oberbauer et al. 1996). (2) We set the canopy minimum leaf water potential ($\psi_{min}$) at all sites to $-1.5$ MPa (Oberbauer and Dawson 1992). (3) Because of the small stature of tundra vegetation and relatively sparse canopies, we simulated two canopy layers, dividing leaf area index (LAI) and N equally between them. (4) We simplified the hydraulic scheme, applying a constant soil–root–shoot hydraulic conductance ($K_s$) to each layer (Williams et al. 2000).

![Fig. 1. A map of the Kuparuk Basin, Alaska. Measurement stations for temperature data only are marked as closed circles, remaining stations are marked with open squares. Locations of the 14 vegetation survey sites are marked V.](image-url)
Independent eddy covariance data collected in and around the Kuparuk Basin at nine sites during 1994–1996 (Eugster et al. 1997) was used to test the SPA model (Williams et al. 2000). Based on differences in LAI and $N_f$ (total foliar N, measured in grams N per square meter ground area) among sites, and on local meteorology, the SPA model was able to explain the major proportion of the variance in net ecosystem production (NEP) among sites; most of the site-to-site variation in NEP resulted from variation in GPP. The daily amplitude of NEP (and thus the magnitude of GPP) was closely related to LAI and $N_f$ at each site. During the periods investigated, we found that nonvascular plants seemed to play a minor role in NEP, and the flux data suggested that photosynthesis was less sensitive to cold temperatures than leaf-level studies suggested.

**The aggregated-canopy model (ACM)**

The aggregated-canopy model (Williams et al. 1997) is a “big-leaf” model of daily GPP derived directly from the SPA model by aggregation. The ACM is designed for extrapolation and regional assessments of GPP; it captures the detailed behavior of SPA, but at a coarser scale, and at much higher computational speeds—three orders of magnitude. Requirements for data inputs are modest (daily rather than 30-min inputs). But, because it is a calibrated model, ACM can only be applied within the bounds of drivers used to generate the calibration sets with SPA. For this arctic application, we altered the structure of ACM and developed a new calibration that required fewer calibrated parameters, nine in total (see the Appendix for ACM equations and parameter values).

The strategy to develop the arctic ACM required the following steps: (1) define the nature and ranges of the model drivers at a daily time scale; (2) for a comprehensive combination and broad range of the daily driving variables, generate disaggregated diurnal time series of 30-min driving variables required by the SPA model; (3) use the fine-scale SPA model, driven with the 30-min time series, to estimate daily, whole-canopy GPP for the full range of drivers; (4) use SPA estimates of daily GPP to calibrate the aggregated model (ACM).

The first step in producing an ACM calibration was to define the cumulative or daily average driving variables, and their minimum and maximum values (Table 1). The span of meteorological drivers was derived from examination of growing season weather data from the Toolik Lake (Long-term Ecological Research [LTER]) site. The spans of LAI and $N_f$ were derived from field survey data (Williams and Rastetter 1999). The parameters that were constrained least well are the hydraulic conductance, $K_s$, and $\Psi_d$ (the difference between soil and leaf water potentials) (Table 1). For this reason, variability around a reasonable mean value was included in the construction of the response surface for both.

The accuracy of the ACM calibration is improved if the domain of the calibration variables is restricted to the range observed in the field. For example, based on a field study (Williams and Rastetter 1999), LAI and $N_f$ can be related by

$$N_f = 0.31 + 1.29 \text{ LAI} \\
+ (0.08) \\
+ (0.10) \tag{1}$$

where standard errors are indicated in parentheses. We therefore restricted our calibration to LAI and $N_f$ values fitting this relationship, though to account for some variability around this the relationship, we randomly varied each estimate of $N_f$ by up to ±30%. Likewise, the long-term climate record at the Toolik Lake LTER site confirmed a common climatic correlation between daily irradiance and the daily temperature range (Bristow and Campbell 1984). We used these correlations to estimate irradiance from randomly assigned maximum and minimum daily temperatures (Table 2) and accounted for variability in the relationship by randomly varying irradiance by up to ±4 MJ·m$^{-2}$·d$^{-1}$ (with a minimum daily irradiance of 2 MJ·m$^{-2}$·d$^{-1}$).

### Table 1. Maximum and minimum values of driving variables for gross primary production (GRP) in tundra ecosystems used to construct the aggregated-canopy model (ACM).

<table>
<thead>
<tr>
<th>Factor</th>
<th>Symbol</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leaf area index (m$^2$/m$^2$)</td>
<td>$L$</td>
<td>0.1</td>
<td>2.5</td>
</tr>
<tr>
<td>Maximum daily temperature ($^\circ$C)</td>
<td>$T_{\text{max}}$</td>
<td>3.0</td>
<td>20.0</td>
</tr>
<tr>
<td>Daily temperature range ($^\circ$C)</td>
<td>$T_r$</td>
<td>2.0</td>
<td>16.0</td>
</tr>
<tr>
<td>Mean foliar N concentration (g N/m$^2$ leaf area)</td>
<td>$N_f$</td>
<td>$^{\dagger}$</td>
<td>$^{\dagger}$</td>
</tr>
<tr>
<td>Irradiance (MJ·m$^{-2}$·d$^{-1}$)</td>
<td>$I$</td>
<td>2</td>
<td>30</td>
</tr>
<tr>
<td>Day of year</td>
<td>$D$</td>
<td>173</td>
<td>270</td>
</tr>
<tr>
<td>Ambient CO$_2$ concentration (µmol/mol)</td>
<td>$C_a$</td>
<td>250</td>
<td>700</td>
</tr>
<tr>
<td>Leaf water–soil water potential difference (MPa)</td>
<td>$\Psi_d$</td>
<td>−2.5</td>
<td>−0.5</td>
</tr>
<tr>
<td>Leaf specific hydraulic conductance (mmol·m$^{-2}$·s$^{-1}$·MPa$^{-1}$)</td>
<td>$K_s$</td>
<td>0.1</td>
<td>3.0</td>
</tr>
</tbody>
</table>

$^{\dagger}$ See Eq. 1 and associated text.
incident radiation was improved to take account of the variation in optical path length according to changes in solar elevation. This is particularly important in the Arctic because the solar elevation tends to be so low. The disaggregation was solved numerically to find, for a given date and latitude, the atmospheric transmittance that resulted in the specified integrated daily incident radiation at the ground surface.

Using 4177 predictions of GPP from SPA, the best calibration of the nine-parameter ACM (see the Appendix), by nonlinear least squares, explained >96% of the variation in SPA predictions ($Y = 0.068(\pm 0.011) + 0.961(\pm 0.003)X$). The slope of the fit was slightly, but significantly ($P < 0.01$), less than unity, and the intercept was also slightly but significantly greater than zero ($P < 0.01$). The mean discrepancy between associated SPA and ACM predictions was 12% of the SPA prediction.

**Spatiotemporal Weather and Vegetation Data**

The time period 13 May–30 September 1995 (days 133–273) was chosen for the assessment of seasonal basin productivity, based on the availability of detailed weather data across the basin for that year. Phenology indicative of vegetation dynamics was evident in satellite imagery for the period mid-May through late September.

**Spatiotemporal weather data**

We generated spatiotemporal weather patterns for 1995 using data collected at nine stations operated by D. Kane and L. Hinzman (University of Alaska, Fairbanks, Alaska, USA), 31 stations operated by M. Sturm (Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire, USA) and the long-term station operated by the Toolik Lake LTER program. Of these 41 stations, 32 are within the basin (Fig. 1). Data collected 1990–1998 at the Toolik Lake LTER indicated that the 1995 growing season was slightly cooler and duller than the decadal average.

**Temperature**—We generated daily maximum and minimum temperature fields at 1-km² resolution, using an inverse distance-weighting interpolation routine. The interpolation determines surface values using a linearly weighted combination of ≤12 of the nearest sample points, and derives the weight as the inverse distance from each point raised to the power 3.5. The protocol we followed was as follows:

1) Gather daily maximum and minimum temperature data from a number of stations dispersed basin-wide (Fig. 1); we rejected ~4% of station data as outliers.

2) Convert temperature data at these stations to estimates at sea level using a lapse rate of 6°C/km for maximum temperatures and 3°C/km for minimum temperatures.

3) Generate basin-wide estimates of maximum and minimum temperature, adjusted to sea level, via interpolation of the modified station data from part (2).

4) Derive basin-wide estimates of topographically adjusted temperature by modifying sea level spatial estimates from part (3), using basin elevation data and lapse rates from part (2).

**Radiation**—The spatiotemporal coverage of radiation data in the Kuparuk Basin for 1995 was too sparse for any reliable interpolation. We used data collected during 1988–1997 at the Toolik Lake LTER site to generate a three-parameter ($j_1$, $j_2$, and $j_3$) empirical relationship between daily temperature range ($T_{max} - T_{min}$), estimated extraterrestrial radiation ($k_0$) and daily irradiance ($I$), using a modified version of the procedure of Bristow and Campbell (1984)

$$I = k_0 j_1 (1 - \exp[-j_2(T_{max} - T_{min})^{0.5}]).$$

(2)

To reduce the impact of changes in solar elevation, we generated separate parameter sets for each month of the growing season (Table 2), by minimization of least-square differences. Eq. 2 provided unbiased estimates of irradiance at Toolik Lake and was able to explain 35–52% of day-to-day variance (Table 2).

The three-step protocol to generate spatiotemporal fields of radiation was as follows:

**Table 2. Monthly parameter sets ($j_1$, $j_2$, $j_3$) and statistics ($r^2$, and slope and intercept [mean (1 SE)]) of linear regression of modeled vs. measured irradiance used to predict daily irradiance via a calibrated relationship with daily temperature range (Eq. 2).**

<table>
<thead>
<tr>
<th>Parameter/statistic</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
</tr>
</thead>
<tbody>
<tr>
<td>$j_1$</td>
<td>2.37</td>
<td>1.36</td>
<td>1.52</td>
<td>1.50</td>
<td>2.59</td>
</tr>
<tr>
<td>$j_2$</td>
<td>0.1</td>
<td>0.1</td>
<td>0.06</td>
<td>0.06</td>
<td>0.07</td>
</tr>
<tr>
<td>$j_3$</td>
<td>0.44</td>
<td>0.63</td>
<td>0.76</td>
<td>0.73</td>
<td>0.48</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.35</td>
<td>0.41</td>
<td>0.52</td>
<td>0.44</td>
<td>0.48</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.88 (0.08)</td>
<td>1.00 (0.09)</td>
<td>1.04 (0.06)</td>
<td>0.96 (0.07)</td>
<td>0.94 (0.06)</td>
</tr>
<tr>
<td>$E$</td>
<td>2.65 (1.6)</td>
<td>0.61 (1.67)</td>
<td>−0.29 (1.07)</td>
<td>0.87 (0.81)</td>
<td>0.39 (0.49)</td>
</tr>
<tr>
<td>Monthly mean irradiance (MJ·m⁻²·d⁻¹)</td>
<td>20.5</td>
<td>19.2</td>
<td>17.3</td>
<td>11.3</td>
<td>6.9</td>
</tr>
<tr>
<td>$n$</td>
<td>239</td>
<td>219</td>
<td>248</td>
<td>223</td>
<td>210</td>
</tr>
</tbody>
</table>

Notes: Parameter sets ($j_1$, $j_2$, $j_3$) were derived by nonlinear regression on $n$ daily data collected during 1988–1997 at the Toolik Lake LTER site.

† Mean error of prediction of the irradiance estimates ($E = [(\Sigma (M - P)^2)/n]^{0.5}$; $M =$ measured value, $P =$ modeled value).
FIG. 2. Predicting irradiance in the Kuparuk Basin. The left panel shows the comparison of modeled and measured daily radiation data at three sites (Fig. 1), 13 May–19 September 1995. Also shown are the 1:1 line (thin) and regression line (thick). $E$ is the mean error of prediction. The top right panel shows the frequency distribution of the prediction error and a fitted Gaussian curve; the bottom right panel shows the frequency distribution of the ratio of errors (modeled:measured) and a fitted lognormal curve.

1) Construct daily surfaces of extraterrestrial radiation ($k_0$) for the basin at 1-km resolution (Gates 1980).

2) Create daily irradiance ($I$) surfaces at 1-km resolution (Eq. 2), using $k_0$, elevation-corrected maximum and minimum temperature surfaces, and parameters for the appropriate month (Table 2).

3) Correct irradiance for slope and aspect of the landscape, at 1-km resolution (Swift 1976).

Independent irradiance data were available from three stations within the Kuparuk Basin (Imnavait Basin, West Kuparuk, and Sagwon Hill; Fig. 1), collected during 13 May–19 September 1995, for a total of 388 site-days of data. Predictions of daily irradiance from daily maximum and minimum temperature data in the cells containing these sites explained 65% of day-to-day variability (Fig. 2), and a frequency distribution of errors was distributed normally around zero (Fig. 2). The total irradiance on all sites over the full measurement period was 5644 MJ/m²; the modeled value was 5637 MJ/m², a difference of 0.1%.

Spatiotemporal vegetation data

The key vegetation data required for the extrapolation are leaf area index (LAI) and foliar N ($N_f$). To estimate their spatial distributions, we combined three sources of land surface data for the Kuparuk Basin: direct measurements of LAI and $N_f$, a land cover database, and remotely sensed normalized difference vegetation index (NDVI) data. For the determination of LAI, these data sets each have their own set of strengths and weaknesses.

1) Direct LAI and $N_f$ measurements were undertaken via destructive harvests at 14 sites within the Kuparuk River Basin (KRB) in late July 1997 (Williams and Rastetter 1999). Each of the main vegetation cover types was sampled at least once (Table 3). However, these data do not describe how LAI varies over time, and the sites cover a minute fraction of the basin, which may not be representative of the whole (Fig. 1).

2) The 50-m resolution land cover database, derived from remote sensing and field surveys, provides high-resolution information on the distribution of eight land

### Table 3. Vegetation types, mean one-sided vascular plant leaf area index (LAI), and basin coverage for each of five vegetation types.

<table>
<thead>
<tr>
<th>Vegetation type†</th>
<th>LAI</th>
<th>No. sites</th>
<th>No. quadrats</th>
<th>Percent cover in basin§</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wetland</td>
<td>0.40</td>
<td>2</td>
<td>16</td>
<td>6</td>
</tr>
<tr>
<td>MAT</td>
<td>0.75</td>
<td>5</td>
<td>36</td>
<td>24</td>
</tr>
<tr>
<td>MNT/heath</td>
<td>0.61</td>
<td>4</td>
<td>28</td>
<td>45</td>
</tr>
<tr>
<td>Shrub</td>
<td>1.71</td>
<td>2</td>
<td>16</td>
<td>18</td>
</tr>
<tr>
<td>Barrens</td>
<td>0.11</td>
<td>1</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>

† MAT, moist acidic tundra; MNT, moist non-acidic tundra.
‡ Standard errors are calculated from the total number of sampled quadrats.
§ Percent cover data from Walker et al. (1994); remaining cover is water, clouds, and ice.
cover types across the basin (Walker et al. 1994, 1998, Muller et al. 1998). However, this database provides no information on temporal variation in cover characteristics, and data is qualitative rather than quantitative.

3) The satellite-derived NDVI database provides biweekly composites of spectral reflectance data, and is the only data set to describe temporal variability in a land cover characteristic (Stow et al. 1993). Advanced very high resolution radiometer (AVHRR) data used to calculate NDVI products were collected using the National Oceanic and Atmospheric Administration (NOAA) 14 satellite. Maximum value compositing was used to minimize effects of temporal differences in atmospheric and illumination conditions. AVHRR images collected over a half-month period were geographically registered, and the maximum observed NDVI was assigned to each pixel location for the final composite. However, the assumption that the maximum values were free of cloud contamination and best represent the true radiance values is questionable. Furthermore, NDVI data are relatively coarse-scale (1 km) compared to (1) and (2), and must be calibrated to predict LAI.

There is some uncertainty about the spatial resolution associated with off-nadir views, and the signal-to-noise ratio of AVHRR data decreases dramatically at high latitudes/large solar zenith angles. Further degradation of the NDVI signal may result because mosses and lichens are common in the arctic landscape (Williams and Rastetter 1999), and they exhibit distinctly different spectral characteristics from vascular plants (Bubier et al. 1997).

**Predicting LAI at 14 survey sites within the basin.**—We made direct comparisons between the three data sets by comparing respective descriptions of land surface properties at the locations of the 14 field surveys (Fig. 3). Because spatial errors in the land cover data set (2) are on the order of 50 m, with an accuracy of 87% (Muller et al. 1998), we estimated land cover at each site by sampling the 16 cells within 100 m of each study site and the 400 cells within 500 m of each study site. We converted the land cover data to an estimate of LAI, using relationships from site surveys for the five vegetation cover types (Table 3). Each cell was assigned the mean LAI of the corresponding vegetation

---

**Fig. 3.** Intercomparison of three land surface data sets at locations corresponding to the 14 sites surveyed in detail in July 1997 (Fig. 1). The data sets compared are (1) field measurements of leaf area index (LAI); (2) biweekly composites of normalized difference vegetation index (NDVI) for the period 22 July–4 August 1995; and (3) LAI estimated from land cover data at 50-m resolution, which was aggregated to either the 16 or 400 pixels surrounding each survey site. The numbers in parentheses in the regression equation are the standard errors for the coefficients.
type; a mean site LAI was determined for the 16- and 400-cell collations.

**Comparing measured LAI and NDVI.**—There were weak or insignificant relationships between site-specific NDVI measurements covering 22 July–4 August 1995 and the LAI determined from field surveys in late July 1997 (Fig. 3a). However, because the data were gathered on such different scales (1 km² vs. 0.04 m²), and because the tundra is so spatially heterogeneous across these scales, a poor relationship was not unexpected.

**Comparing NDVI and land-cover estimated LAI.**—For the same 1-km² cells, there was a highly significant, curvilinear relationship between NDVI and LAI estimated from the 400 neighboring land cover cells (Fig. 3b). An exponential equation explained 70% of the variance among sites, confirming that NDVI is sensitive to differences in vegetation cover found within the Kuparuk Basin. However, uncertainties in assignment of LAI to land cover types must be transferred into any calibration of an NDVI–LAI relationship. The nonlinearity of the relationship is unlikely to be a saturation response of the relatively sparse tundra canopies. It is more likely a result of the effective increase in path length traveled by photons relative to sensor viewing geometry when viewing off-nadir, which is a common situation at high latitudes.

**Comparing measured LAI and land-cover estimated LAI.**—This intercomparison indicates the degree to which the vegetation classification of Walker et al. (1995) matches that for the field data. The results show that there was no significant relationship between LAI estimated from 400 cells (covering 1 km²) and field measurements (Fig. 3c), most likely because of the difference in measurement scale and the high degree of landscape heterogeneity. There was a significant linear relationship (P < 0.05) between measured LAI and estimates using 16 neighboring cells (Fig. 3d). Though the explanatory power of the regression was weak, and confidence limits broad, the relationship suggests that the land cover database corresponds reasonably with our assignment of vegetation types to each site at this finer scale, except in those sites where a large fraction of land cover data was assigned to shrubs. The separation into land cover types is qualitative, and this intercomparison indicates that there is a difference between qualifying criteria for shrub lands used by Walker et al. (1995) and by this study.

**Generating LAI surfaces for the entire basin**

We extended analyses of land-surface characteristics to the entire basin. A comparison of peak season NDVI with LAI estimates derived from the land cover database and field surveys indicated a close correlation across the watershed (Fig. 4). Some sites were clear outliers to the overall relationship, where the cell fraction occupied by water was >30%; cloud, >30%; or shadow, >10%. Open water, clouds, and shadows have different reflectance characteristics from vegetation, and we removed these cells (n = 520, 5.4% of total) from further analyses. The LAI–NDVI relationship for the 9020 remaining cells was similar to that for the 14 survey site cells (Fig. 3), saturating beyond an NDVI of 0.5. We fitted an exponential equation to the data and explained 70% of variability in NDVI (Table 4).

Using this equation, the mean error of prediction of LAI from NDVI (0.13) is ~16% of the mean estimated LAI (0.81) for the basin at peak season, but there are large errors at high LAI because NDVI saturates near an LAI of 0.5.

Because there is no clear optimal approach to constructing daily LAI surfaces, we investigated and contrasted two methods to develop an estimate of uncertainty (Fig. 5).

**Method 1.**—We used NDVI data to predict LAI, via the calibrated equation between land cover estimated LAI and NDVI (Table 4). Leaf area index for the 520 outlier cells was derived using NDVI data from nearest neighbor, nonoutlier cells with similar dominant land cover. The strength of this method is that the NDVI data is directly correlated to temporal and spatial variation in land cover; however, it must be calibrated to predict LAI, nonvascular plants may confuse the spectral signal, and signal-to-noise ratio in AVHRR data is low at high latitudes. Because the calibration with land cover data was for peak season only, predicting LAI from NDVI at low values relies on extrapolating the relationship outside the calibration range.

**Method 2.**—We used the land cover database peak-season LAI estimates to generate daily LAI surfaces directly. Phenology was calculated by modifying the peak-season LAI in each cell by the ratio of cell NDVI on the target day to the cell NDVI at peak season. Thus, we assume that the relative changes in NDVI for each cell are identical to the relative changes in LAI over the same period (Fig. 5), but peak LAI is independent of NDVI. This approach allowed us to avoid the saturation problem in the NDVI data. The land cover approach is justified by its direct connection to measurement data, and its incorporation of fine spatial resolution, but is weakened by its qualitative and discontinuous scaling strategy.

Both methods combine the land surface and NDVI data sets, and so they are not strictly independent. But the important difference in LAI assignment is revealed by frequency analysis; there is a distinct bimodality for both the peak-season land cover-estimated LAI and the NDVI data (Fig. 4). But in the NDVI approach (method 1), the mode LAI is in the upper section of the range; whereas in the land cover approach (method 2), the mode is in the lower part of the range. The NDVI distribution shows a sharp cutoff at its upper end, while the land cover LAI estimates decline more gradually. The NDVI approach ensures a strict upper limit in predictions at an LAI of ~1.1, while in the
land cover approach some sites are assigned an LAI of $\sim 1.4$ (Fig. 4).

**Applying the aggregated-canopy model (ACM) spatiotemporally**

Five spatiotemporal fields were used to drive the aggregated-canopy model (ACM): maximum and minimum temperature, daily irradiance, LAI and $N_l$. However, $N_l$ data were derived directly from LAI, and irradiance data were derived from the temperature fields, so the only independent data were maximum and minimum temperature and LAI. The gross primary production (GPP) was determined for each of the 9542 1-km² cells that encompass the basin for each day during 13 May–30 September. We defined a baseline GPP estimate as that derived with elevation-corrected temperature surfaces, slope-corrected irradiance surfaces, and with NDVI linearly interpolated to produce daily LAI surfaces. For the baseline, both methods of assigning LAI were tested. No productivity occurred in cells where maximum daily temperature failed to rise above freezing.

**Model construction and the Propagation of Error**

**Disaggregation routines**

We examined how radiation disaggregation introduced error into predictions of daily and hourly irradiance, and propagation of this error into gross primary production (GPP) estimates. Using data gathered during 141 days in the summer of 1995 at Toolik Lake, we compared measured hourly irradiance with predictions derived from disaggregations of daily total irradiance, which were themselves predicted from daily temperature differentials (see Spatiotemporal weather and vegetation data; Spatiotemporal weather data; Radi-

<table>
<thead>
<tr>
<th>LAI assignment</th>
<th>$P_1$</th>
<th>$P_2$</th>
<th>$P_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.165</td>
<td>0.280</td>
<td>-0.049</td>
</tr>
<tr>
<td>Mean + 1 se</td>
<td>0.144</td>
<td>0.258</td>
<td>0.081</td>
</tr>
<tr>
<td>Mean - 1 se</td>
<td>0.085</td>
<td>0.228</td>
<td>0.020</td>
</tr>
</tbody>
</table>

*Notes: Three different calibrations were developed: one using mean measured LAI values for each vegetation type, and the others by increasing or reducing LAI by one standard error.*
FIG. 5. Temporal variation in predictions of mean temperature, irradiance, leaf area index (LAI), and gross primary production (GPP) for the Kuparuk Basin, 13 May–30 September 1995. For the temperature and irradiance plots, dashed lines indicate the maximum and minimum values predicted within the basin on that day, whereas the solid line indicates mean values. For both LAI and GPP, the solid line represents predictions using method 1 (NDVI), and the dashed line indicates method 2 (land cover).

ation). Using prediction and disaggregation routines, we accounted for 74% of variation in measured hourly data, and effectively tracked diurnal dynamics on many days (Fig. 6). Major discrepancies occurred on days with variable cloudiness; these conditions invalidated the assumption of isotropic skies used in the disaggregation routines. Measured mean hourly irradiance over the 141-d period was 0.32 mmol·m⁻²·s⁻¹, while the modeled mean value was 0.37 mmol·m⁻²·s⁻¹. Mean error of prediction for each of the 3384 hourly values over the full growing season was 0.23 mmol·m⁻²·s⁻¹, 70% of the measured mean value, and linear regression analysis indicated a bias towards overprediction of hourly values.

We examined how error in photosynthetically active radiation (PAR) predictions and disaggregation propagated into errors in GPP predictions of the soil–plant–atmosphere (SPA) model. We made hourly predictions of GPP using measured hourly insolation data, and then with estimated hourly values. Apart from the difference in PAR, simulations were undertaken using the same measured meteorological data, peak-LAI and foliar-N data measured at the site in 1997, and a simple phenology that ramped up to and down from these values. Hourly GPP predictions using simulated PAR matched 86% of variability in predictions made using measured hourly data (Fig. 6). Mean GPP prediction for the entire growing season using measured PAR was 1.55 μmol·m⁻²·s⁻¹, compared to 1.66 μmol·m⁻²·s⁻¹ using simulated PAR. Mean error of hourly predictions was 0.71 μmol·m⁻²·s⁻¹, 45% of mean GPP prediction with measured PAR data. Thus, the error in predicting hourly GPP was relatively less than the error in predicting PAR.

These comparisons show that, while there is appreciable error in predicting hourly irradiance given only daily maximum and minimum temperature, only a fraction of this error is propagated when estimated irradiance is used to predict photosynthesis. The propagation of error is controlled by the nature of the transforming function. If the function is linear, the relative error will be unchanged by the transformation. But, if the function is nonlinear, the error will either be amplified or reduced by the transformation. In the case of a saturating function, like the response of photosynthesis to irradiance, relative error is always reduced by the transformation, an action called error compensation (Fig. 6). Because of error compensation, GPP predictions are relatively insensitive to errors in PAR, especially at higher PAR values.

Estimates of daily irradiance showed a significant linear relationship with measured values \((P < 0.0001)\), as expected. Mean daily measured shortwave radiation at Toolik Lake was 12 MJ·m⁻²·d⁻¹ over the growing season, whereas the value predicted from temperature differentials was 14 MJ·m⁻²·d⁻¹. Mean error of daily predictions was 3.4 MJ·m⁻²·d⁻¹, equal to 28% of mean daily irradiance. Thus, normalized error of the daily irradiance predictions was less than half the magnitude of the normalized mean error of hourly predictions.

We used the SPA model to compare daily predictions of GPP using hourly measured irradiance data vs. simulated hourly irradiance, derived from disaggregations of both measured and estimated daily irradiance (Fig. 7). Using estimated daily irradiance, the mean error of prediction of GPP was 19% of the growing season mean daily value, and the full growing season difference in predicted GPP was 7%. Thus, a smaller portion of the large hourly errors in PAR predictions, averaging 70% of the mean values, has been propagated through to daily and annual prediction of GPP. Errors have been damped by two factors. First, the saturation response of carboxylation to irradiance compensates for errors, especially at high PAR values. As long as both measured and predicted PAR are high, even relatively large differences between values will translate into a small difference in GPP predictions because of saturation. Second, through most days, the hourly prediction errors
have both positive and negative components, so that when predictions are aggregated to daily values, total relative error is reduced. In summary, while we are cautious about the accuracy of individual hourly predictions of PAR and GPP, daily predictions show considerably reduced errors and bias. Although these errors must still be seriously considered in any set of predictions, these analyses suggest that over daily and longer time scales, the disaggregation approach employed will produce predictions at a reduced level of error.

Intercomparison of predictions from the aggregated-canopy (ACM) and soil–plant–atmosphere (SPA) models

For the same period and with the same conditions as for the disaggregation analysis, a comparison of GPP predictions from the SPA model (driven with hourly meteorological data) against the ACM model (driven with predicted daily irradiance from temperature data) showed that typical daily prediction discrepancies were 16% of the mean (Fig. 7), −0.25 g C m⁻² d⁻¹. However, over the whole season, the annual GPP differed by just 1%, and a linear regression indicated that ACM predictions were unbiased. The aggregated-canopy model tended to underpredict GPP on bright days and overpredict on dull days. But overall, given estimated daily irradiance as a driver, the ACM was as accurate a method of predicting GPP as disaggregating irradiance and using SPA.

Prediction Uncertainty Associated with Driving Data

The basin simulations suggest that the most productive parts of the basin are in its southern and central portions (Fig. 8) and nearly 90% of annual production occurred during June, July, and August, with nearly 40% in July alone (Fig. 5). We undertook a series of
Fig. 7. Examination of the error in daily GPP predictions associated with radiation disaggregation routines (top panels) and ACM calibration (bottom panels). Meteorological data are from the Toolik Lake LTER site, collected during 1995. The top left panel shows the time series of daily GPP predictions using the SPA model and PAR forcing derived from a number of sources: (1) measured 30-min PAR data (squares), (2) measured total daily PAR, disaggregated to 30-min estimates (solid line), and (3) daily total PAR estimated from temperature data, and then disaggregated to 30-min estimates (dashed line). The top right panel compares the daily GPP predictions using 30-min measured PAR data with those derived using estimated daily irradiance, disaggregated to 30 min. The bottom left panel shows the time series of predicted daily photosynthesis (gross primary production; GPP) using the soil–plant–atmosphere (SPA) model forced with measured 30-min irradiance data (squares) vs. the aggregated-canopy model (ACM) forced with daily irradiance estimated from temperature data (solid line). The bottom right panel compares SPA and ACM predictions for each day. In the right-hand panels, 1:1 lines are dashed, and regression lines are solid. Equations of the linear regression, with standard errors in parentheses, are shown. $E$ is the mean error of comparison; see Table 2 footnotes for defining equation.

analyses to quantify uncertainty in overall estimates of basin production and in spatial and temporal patterns suggested by the simulations.

Vegetation properties

The key biotic constraints on production are leaf area index (LAI) and foliar N ($N_f$). There are three main causes of uncertainty in predicting how LAI and $N_f$ vary over the arctic landscape. First, field sampling was limited to a few sites, causing uncertainty in the mean LAI and $N_f$ values assigned to vegetation types. Second, distributing LAI and $N_f$ across the landscape and through time relies on remotely sensed data that are associated with uncertainty. And third, the correlation between LAI and $N_f$ derived from fieldwork, only approximates the actual relationship.

Normalized difference vegetation index (NDVI) and land cover data

The two methods used to assign LAI across the basin relied primarily on remotely sensed data, using a combination of frequent, coarse-scale data (normalized difference vegetation index; NDVI), and a fine-scale land cover database (collected by Walker et al. 1995). An analysis of their frequency distributions indicated important differences between the two databases (Fig. 4), which have ramifications for assignment of land cover properties. Predictions of annual gross primary production (GPP) within the basin indicated that the overall productivity of the basin was 6% greater using land cover data (method 2) to estimate LAI rather than NDVI data (method 1) (Table 5). Throughout the season, there were clear differences in predicted spatial
patterns, especially in the central and southern basin (Fig. 8). Areas with a high concentration of shrub tundra were particularly sensitive to the method employed, and this explains the hot spots of activity following the river valleys in method 2. However, temporal differences in whole-basin activity occurred only over a brief and early portion of the growing season (Fig. 5). During the peak season (late July), there was little difference in the total basin predictions of GPP, largely because the mean LAI for the basin was similar in both methods; on July 26 it was 0.79 for method 1 and 0.80 for method 2. The difference in overall basin GPP occurs in the early season and is attributable to differences in phenological predictions between the two methods. The assumption of a linear relationship between changing NDVI and LAI in the land cover method (method 2) means that LAI rises more rapidly and declines more slowly than when LAI is related to NDVI nonlinearly (method 1).

At peak season there is little difference (<2%) in predictions of total watershed GPP between methods, though there are clear differences in the predicted distribution of GPP. While the NDVI method has lower peak LAI predictions, this is compensated for by greater spatial weighting towards the upper end of the LAI distribution (Fig. 8). Statistical comparison of the two methods indicated that the largest relative error of prediction was associated with spatial rather than temporal predictions. For individual locations, the mean difference in predicted annual GPP between the two methods is ~16% of the mean value of method 1 (Table 5). For the temporal comparison, whole basin daily prediction uncertainty is 8% of the mean daily value; most of this difference occurs in the early part of the growing season (Fig. 5).

Estimates of land cover leaf area index (LAI)

We examined how uncertainty in assigning mean LAI to each land cover type contributed to uncertainty in GPP predictions, first by increasing and then decreasing the assigned LAI by one standard error of the mean measurement (Table 3). Other assumptions were left unchanged, so that $N_i$ increased according to Eq. 1. In the case of method 2, new land cover estimates
Table 5. Error analysis of predictions of basin GPP, comparing both spatial and temporal errors in predicted gross primary production (GPP).

<table>
<thead>
<tr>
<th>Land cover method†</th>
<th>Analysis‡</th>
<th>Mean GPP</th>
<th>Mean discrepancy between predictions</th>
<th>Standard error of the regression</th>
<th>Slope of the regression</th>
<th>Intercept of the regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Space</td>
<td>Time</td>
<td>Space</td>
<td>Time</td>
<td>Space</td>
</tr>
<tr>
<td>Method 1 baseline</td>
<td>Analysis</td>
<td>216</td>
<td>14 185</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Method 2 baseline</td>
<td></td>
<td>230</td>
<td>15 109</td>
<td>34 1192</td>
<td>...</td>
<td>7 0.70 0.96</td>
</tr>
<tr>
<td>Method 1 +1 SE LAI</td>
<td></td>
<td>248</td>
<td>16 330</td>
<td>33 2425</td>
<td>2 294</td>
<td>0.93 0.89</td>
</tr>
<tr>
<td>Method 2 +1 SE LAI</td>
<td></td>
<td>259</td>
<td>17 013</td>
<td>30 2267</td>
<td>3 72</td>
<td>0.91 0.89</td>
</tr>
<tr>
<td>Method 1 +1 SE NI1</td>
<td></td>
<td>185</td>
<td>12 184</td>
<td>31 2365</td>
<td>1 94</td>
<td>1.14 1.16</td>
</tr>
<tr>
<td>Method 2 -1 SE LAI</td>
<td></td>
<td>200</td>
<td>13 161</td>
<td>31 2317</td>
<td>4 87</td>
<td>1.11 1.15</td>
</tr>
<tr>
<td>Method 1 +1 SE NI1</td>
<td></td>
<td>234</td>
<td>15 391</td>
<td>19 1435</td>
<td>0 124</td>
<td>0.93 0.92</td>
</tr>
<tr>
<td>Method 2 +1 SE NI1</td>
<td></td>
<td>249</td>
<td>16 355</td>
<td>20 1479</td>
<td>1 113</td>
<td>0.93 0.92</td>
</tr>
<tr>
<td>Method 1 -1 SE NI1</td>
<td></td>
<td>196</td>
<td>12 856</td>
<td>21 1561</td>
<td>1 173</td>
<td>1.08 1.10</td>
</tr>
<tr>
<td>Method 2 -1 SE NI1</td>
<td></td>
<td>209</td>
<td>13 751</td>
<td>21 1601</td>
<td>1 140</td>
<td>1.08 1.10</td>
</tr>
<tr>
<td>Method 1 radiation</td>
<td></td>
<td>213</td>
<td>13 977</td>
<td>4 249</td>
<td>2 28</td>
<td>1.01 1.01</td>
</tr>
<tr>
<td>Method 1 temperature</td>
<td></td>
<td>218</td>
<td>14 309</td>
<td>2 155</td>
<td>1 90</td>
<td>0.99 1.00</td>
</tr>
<tr>
<td>Method 1 meteorological station</td>
<td></td>
<td>214</td>
<td>14 073</td>
<td>11 1678</td>
<td>9 1623</td>
<td>1.15 1.05</td>
</tr>
<tr>
<td>Method 2 meteorological station</td>
<td></td>
<td>207</td>
<td>13 619</td>
<td>17 1756</td>
<td>11 1382</td>
<td>0.84 1.12</td>
</tr>
<tr>
<td>Method 1 4 meteorological station</td>
<td></td>
<td>212</td>
<td>13 954</td>
<td>9 749</td>
<td>8 669</td>
<td>1.02 1.03</td>
</tr>
</tbody>
</table>

Notes: The spatial comparison examines differences among annual predictions for each of the 9542 1-km² grid cells comprising the Kuparuk Basin. The temporal comparison examines differences among 141 daily predictions of whole basin production during the 1995 growing season. Statistical comparisons of predictions use: (1) the mean discrepancy between predictions, $E = \frac{(\sum (P_i - P_j))^2}{n}^{1/2}$, $P_i$ = predicted value from analysis i (similarly for analysis j); n = 9542 or 141; (2) the standard error of the predicted $P_i$ value vs. predicted $P_j$ in a linear regression between the two sets of predictions; (3) the slope of the linear regression between $P_i$ and $P_j$; and (4) the intercept of the regression between $P_i$ and $P_j$. Units for spatial predictions are Mg km⁻² yr⁻¹, and for daily predictions are Mg/d for the entire basin.

†Two methods are used for generating leaf area index (LAI) estimates: method 1 is primarily governed by the normalized difference vegetation index (NDVI), and method 2 by land cover data.

‡Methods 1 and 2 are first compared directly. In succeeding cases the statistics are for the comparison of the particular analysis with the baseline predictions using the identical method of LAI assignment.

of LAI were derived with the altered LAI assignment, and, for method 1, we calculated new relationships between estimated land cover LAI and NDVI. We found that increasing LAI by one standard error raised total basin GPP by 15% or 13%, while decreasing LAI by one standard error reduced basin GPP by 14% or 13%, respectively, for methods 1 and 2. While mean differences of prediction were as high as or higher than in the baseline comparison between methods (Table 5), the much-reduced standard error of prediction indicates that the changes induced by altered LAI assignments are scalar and linear for each method. Thus, the overall spatial and temporal patterns are simply shifted by the change, rather than significantly rearranged.

Foliar N estimates

We examined the prediction uncertainty related to the assumption that foliar N ($N_f$) is linearly related to LAI (Eq. 1). We first increased both slope and intercept of the relationship by one standard error, and then reduced both, calculating the implications on GPP for both methods of assigning land cover characteristics, increasing $N_f$ raised production by 8–9%, while decreased $N_f$ reduced production by 9%. The low standard errors of prediction indicate, again, a relatively simple scalar response in GPP to changes in $N_f$ (Table 5).

Meteorological drivers

The routines used to predict daily total irradiance, while unbiased, do introduce error into their estimates (Fig. 2). From frequency analysis of the ratio of modeled to measured irradiance, we constructed a probability density function (PDF) describing the likely distribution of errors. We used a random number generator to propagate this error into irradiance predictions for each cell and every day, and examined the effect on GPP predictions. Overall, total basin GPP was reduced by 1% (Table 5). The alteration to mean spatial and temporal predictions was 2% of respective means.

The interpolation routines used to construct temperature estimates (which also determine irradiance predictions) were also associated with errors (Fig. 9). Using interpolations constructed from 20 stations, we con-
FIG. 9. Comparison of measured and interpolated maximum and minimum daily temperature data at 21 sites (left panels) or 39 sites (right panels) in the Kuparuk Basin during the summer of 1995. Left panels: interpolations were constructed from 20 stations (Fig. 2, open squares). Right panels: two stations were used (Toolik Lake and Betty Pingo). Also shown are the 1:1 line (dashed) and regression lines (solid). $E$ is the mean error of prediction ($E = \frac{(\sum(M - P)^2)}{n}^{0.5}$; $M$, measured value; $P$, modeled value).

We examined how a sparser network of meteorological stations affected prediction bias and accuracy. With a single station (Toolik Lake), overall basin GPP was reduced by 2%. There was a greater rearrangement in temporal vs. spatial predictions (Table 5); mean discrepancies and standard errors of temporal prediction were 12% and 11% of the means, compared to 5% and 4%, respectively, for spatial discrepancy/error. Using constructed PDFs of the absolute differences between interpolated and measured temperature, and combined these with a random-number generator to introduce realistic error into the predicted surfaces. We found that the effect of this error was to increase basin GPP by 1%. The effect on mean spatial and temporal predictions was to change GPP in individual cells or on specific days by ~1% also (Table 5).
two stations, Toolik Lake and Betty Pingo in the north of the basin (Fig. 9), total basin GPP was reduced by 4%, and mean discrepancies of temporal and spatial predictions were 12% and 8% of mean values, respectively. Thus, driving data from the extreme north of the basin (Betty Pingo) was less representative of the basin than Toolik Lake data, and degraded predictive accuracy overall. With four stations, adding Sagwon Hill (central east) and West Kuparuk (central west), accuracy was improved; overall, the basin prediction was reduced 2%, and mean discrepancies of spatial and temporal prediction were 4% and 5%, respectively.

**Physiological simulation**

Hydraulic conductance of vegetation and distribution of soil moisture deficits throughout the basin are poorly defined, and are not incorporated in the up-scaling. We varied both the hydraulic conductance and the leaf water-soil water potential difference ($\Psi_d$) to examine the effect on patterns of production. Cutting conductance by 50% reduced overall production by 3%. The low standard error of prediction (Table 5) indicated a scalar response both temporally and spatially, suggesting that there are no clear correlations between hydraulic limitation and other driving variables. Production was raised 1% by doubling conductance, indicating that this parameter is not a key constraint on production. There are similar patterns in changing $\Psi_d$; reduction by 33% cut overall production by 2% across the whole basin, whereas an increase by 33% raised production by 1%.

We examined the effect of an alternative parametrization of the photosynthetic response to low temperatures. Instead of setting photosynthetic rates to 65% of maxima at 5°C, we developed an SPA calibration with rates at 5°C set to 85% of maxima. The effect was to increase overall basin production by 11%, and was temporally, rather than spatially, distinct, with greatest impact on cold days during the growing season.

**Discussion**

The uncertainty in model predictions introduced by aggregating from a 30-min multilayer model (soil-plant-atmosphere; SPA) to a daily big-leaf model (aggregated-canopy model; ACM) is of the same magnitude as that introduced in SPA alone by applying disaggregated daily estimated irradiance in place of measured 30-min driver data (Fig. 7). Thus, there is a similar degree of uncertainty attached to daily prediction using either SPA or ACM, if only daily total irradiance is known. Summed over longer periods and larger areas, the unbiased relationship of SPA and ACM ensures that uncertainties related to the aggregation procedure decline, perhaps to as low as 1%. Uncertainty related to radiation disaggregation remains relatively high at annual time scales (7%); improved methods to disaggregate daily radiation to hourly estimates are required, especially for intermittently cloudy days. Perhaps remote sensing can be used to provide detailed spatial cloudiness and irradiance data. Adequately quantifying the uncertainty in spatial extrapolation of temperature data was complicated by the clustering of stations along a transect through the basin. When we used one to four stations to generate temperature surfaces, predictions of gross primary production (GPP) decreased by 2–4%, and the timing and location of photosynthetic activity was altered. With 20 stations distributed across the basin, the likely error introduced by assumptions in interpolation routines was 1–2%. If this scaling approach was extended across larger areas of the Arctic with sparser meteorological stations, a breakdown in predictive accuracy is likely, although careful siting of stations may be able to minimize bias. From an analysis of radiation predictions, we found that uncertainty in GPP predictions resulting from the construction of irradiance surfaces was small, ~1%, although the lack of corroborative irradiance data on the coastal plain is problematic. Furthermore, this analysis is only valid at the scale of investigation (1 km²), and the effects of microtopographic variation have not been investigated.

The greatest potential uncertainty and bias in GPP predictions relates to assumptions in estimating spatiotemporal patterns of leaf area index (LAI) and foliar N. Alternative methods of applying remotely sensed data identified uncertainty in prediction of canopy development in the early season, and also spatial uncertainty in the central and southern basin, especially in the river valleys dominated by shrub tundra, reflecting uncertainty in assigning LAI to shrub tundra land covers. Phenology is clearly a critical control on basin GPP, and studies should focus on detailed phenological measurements at sites within the basin. An effort to improve the calibration of relationships between the normalized difference vegetation index (NDVI) and the LAI would also be useful.

Equally important to predictive uncertainty was the assignment of mean LAI values to particular land cover types. In the sparse canopies of arctic vegetation, production is extremely sensitive to changes in LAI (Williams and Rastetter 1999). Variability in mean vegetation type LAI by a single standard error, determined from field surveys, translated into 14–15% uncertainty in overall basin production. We used an empirical relationship to predict foliar N from LAI, but variation in this relationship by a standard error altered basin production by 8–9%. Future surveys of canopy properties must collect LAI and N data to test variability in their relationship in other locales, and reduce or quantify the uncertainty in the relationship.

The mean basin GPP via the two methods of LAI assignment was 216 and 230 g C·m⁻²·yr⁻¹, approximately twice the mean net primary production (NPP) (92 g C·m⁻²·yr⁻¹) predicted by a biogeochemical model (Hobbie et al. 1998) for the same area; this is a realistic ratio between the two quantities (Waring et al. 1998).
Vourlitis and Oechel (1999) summarize the extant estimates of seasonal net ecosystem production (NEP) for wet and moist tundra ecosystems of the North Slope; in 10 studies undertaken between 1971–1995, tundra ecosystems were on average net sinks for CO₂ during the growing season, with a seasonal mean of 22 g C/m², although the standard deviation was large (± 38 g C/m²). Over the past 25 yr, Vourlitis and Oechel (1999) conclude that tundra ecosystems have alternated between net C sinks and sources. Net ecosystem production is the difference between GPP and ecosystem respiration, both large fluxes. Predictions of NEP rely on very accurate assessment and prediction of the component fluxes. This study suggests that over large landscape units, current models and available data sets introduce uncertainty on the order of 20–30 g C/m² yr⁻¹ into predictions (Table 5). Most of this uncertainty is associated with assignment of LAI and foliar N, rather than in estimates of meteorology, or in propagation of error through the model hierarchy. However, minimization of meteorological uncertainty relies on a dense network of stations that are peculiar to this area of the Arctic. And while the model hierarchy is generally compensatory and unbiased, there is still uncertainty at finer temporal scales. So, construction of daily NEP estimates for a landscape would have to take account of model uncertainty.

A degree of predictive uncertainty arose through other components of the photosynthesis model. We found that the model parametrization of the sensitivity of GPP to low temperatures was a critical control on GPP predictions for the whole basin; a 50% reduction in sensitivity to low temperatures increased basin production by 10%. The strength of the productivity response to rising temperatures induced by climate change will be strongly affected by the nature of this temperature sensitivity. If photosynthesis is less sensitive to low temperatures, then current production is likely greater, but the response to warming may be less than expected. Data on hydraulic conductivity and water stress of arctic plant species, such as decomposition, respiration, or plant allocation. We have not accounted for interactions with basin hydrology. What we have demonstrated is that scaling photosynthesis in arctic ecosystems requires a focus on several important sets of data. Most important at this scale of study is characterizing the spatiotemporal coverage of LAI and N, This underlines the need to improve remote sensing of LAI and to investigate mechanisms of vegetation change. Functionally, the most critical parts of the model parametrization that invite refinement are the sensitivity of photosynthesis to low temperatures and the activity of moss and lichens. Improved methods to construct hourly time series of irradiance from daily values will reduce uncertainty in daily predictions of GPP. Without this information, precision of predictions of basin GPP under current and altered environments will be impaired.

**Acknowledgments**

We are very grateful to D. Walker for the land cover and digital elevation data; A. Hope and D. Stow for the normalized difference vegetation index (NDVI) data and the Toolik Lake Long-term Ecological Research (LTER) project; and D. Kane, L. Hinzman, and M. Sturm for weather station data. National Science Foundation grants (NSF OPP 9318529, 9614038, and 9732281) and the Andrew W. Mellon Foundation provided funding support for this work. Spatiotemporal data used and generated in this study are archived ARCSS Data Co-ordination Center. We thank Dave Schimel and two anonymous reviewers for helpful comments.

**Literature Cited**


APPENDIX

Equations and parameter values for the aggregated canopy model may be found online in ESA's Electronic Data Archive: Ecological Archives A011-022.