Biomass Production in Switchgrass across the United States: Database Description and Determinants of Yield

S. D. Wullschleger,* E. B. Davis, M. E. Borsuk, C. A. Gunderson, and L. R. Lynd

ABSTRACT

Fundamental to deriving a sustainable supply of cellulosic feedstock for an emerging biofuels industry is understanding how biomass yield varies as a function of crop management, climate, and soils. Here we focus on the perennial switchgrass (Panicum virgatum L.) and compile a database that contains 1190 observations of yield from 39 field trials conducted across the United States. Data include site location, stand age, plot size, cultivar, crop management, biomass yield, temperature, precipitation, and information on land quality. Statistical analysis revealed the major sources of variation in yield. Frequency distributions of yield for upland and lowland ecotypes were unimodal, with mean (±SD) biomass yields of 8.7 ± 4.2 and 12.9 ± 5.9 Mg ha⁻¹ for the two ecotypes, respectively. We looked for, but did not find, bias toward higher yields associated with small plots or preferential establishment of stands on high quality lands. A parametric yield model was fit to the data and accounted for one-third of the total observed variation in biomass yields, with an equal contribution of growing season precipitation, annual temperature, N fertilization, and ecotype. The model was used to predict yield across the continental United States. Mapped output was consistent with the natural range of switchgrass and, as expected, yields were shown to be limited by precipitation west of the Great Plains. Future studies should extend the geographic distribution of field trials and thus improve our understanding of biomass production as a function of soil, climate, and crop management for promising biofuels such as switchgrass.

MANY COUNTRIES, including the United States, have launched ambitious research programs to accelerate development of domestic, renewable sources of energy. In particular, transportation fuels derived from cellulosic biomass offer an attractive bio-based alternative to conventional energy sources (Ragauskas et al., 2006; Schmer et al., 2008). Supplementing the use of fossil fuels with, for example, ethanol produced from energy crops, would benefit economic growth and energy security, as would sustainable cropping systems that both reduce greenhouse-gas emissions and promote energy independence (Tilman et al., 2009). Currently, biofuels derived from corn-based ethanol represent an important, but small, fraction of our domestic energy production. Thus, as society moves to displace an increasing proportion of our transportation fuels with biomass-derived biofuels, energy crops will need to be increasingly deployed.

Among the many agricultural crops screened as potential biofuels, the herbaceous energy crop switchgrass has been identified as a promising feedstock for conversion to biofuels (Sanderson et al., 1996; McLaughlin et al., 2002; Parrish and Fike, 2005). Switchgrass is a warm-season perennial that historically has been an important component of the North American tallgrass prairie. Ranging from northern Mexico to southern Canada, and from the Atlantic coast to the Rocky Mountains, switchgrass has broad adaptability, high growth rates, and tolerates a wide variety of climatic and edaphic conditions. Across its geographic range, two distinct forms or ecotypes are observed: a lowland type found in wetter and more southern habitats and an upland type found in drier mid- and northern latitudes (Porter, 1966; Sanderson et al., 1996; Casler et al., 2004). A variety of lowland and upland cultivars are available and cultivars of both ecotypes are being considered for biofuels; cultivar selection, crop management decisions, and expectations regarding biomass yield will depend to a great extent on geographic location (Parrish and Fike, 2005).

Although there has been a recent resurgence of interest in biofuels, research sponsored in the early 1990s by the Department of Energy and USDA-ARS provides a wealth of agronomic data for switchgrass (McLaughlin and Kszos, 2005; Sanderson et al., 2006). Goals of this effort were focused on indentifying the best cultivars and crop management practices to optimize biomass production for different geographic regions of the country, while developing sustainable cropping systems that could be used to support a future biofuels industry. This goal was largely achieved as a result of research conducted across a network of field sites, with a concerted effort to evaluate biomass yields for upland and lowland cultivars of switchgrass across a range of crop management practices, soils, and climates. Interestingly, despite the obvious value of such data to our current interest in biofuels, there has been only modest effort to quantitatively analyze results from studies conducted over the last 20 yr. Heaton et al. (2004) extracted data from several peer-reviewed publications and evaluated productivity of switchgrass as a function of


Published online 12 May 2010
doi:10.2134/agronj2010.0087
Available freely online through the author-supported open access option.

Copyright © 2010 by the American Society of Agronomy, 5585 Guilford Road, Madison, WI 53711. All rights reserved. No part of this periodical may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording, or any information storage and retrieval system, without permission in writing from the publisher.
N fertilization, growing degree days, and precipitation. Their analysis, while limited (i.e., 77 observations), showed that yield responded positively to water and N, but not temperature, and that biomass yields averaged 10.3 Mg ha\(^{-1}\) for stands harvested 3 yr or more after planting (Heaton et al., 2004). No attempt was made to identify ecotypic variation in patterns of response or to interpret these findings in a geographic context; however, the results illustrate clearly how biomass yield for this species could vary as a function of climate across the United States.

To achieve the ambitious biofuel production goals set forth for the United States and other countries, more information is needed to characterize productivity of bioenergy crops in relation to soil, climate, and management practices. In this study, we compile published literature on biomass yield for lowland and upland ecotypes of switchgrass grown at 39 field sites across the United States. The resulting database contains information on cultivar-specific biomass yield, site location, plot size, stand age, harvest frequency, fertilizer application, soil texture, temperature, precipitation, and land quality. The primary objective of this study was to examine relationships between these factors and yield, and to develop an empirical model that describes biomass yield as a function of significant variables in the database. Our interest in conducting this analysis was motivated by several factors, one of which was the recent work of Johnston et al. (2009), who commented that in communicating potential biofuels production to an increasingly diverse audience, researchers often fail to emphasize that yield estimates are collected from numerous disparate sources and sometimes neglect how biomass production is influenced by geographic location, climate, soil type, or agricultural management regime for the crop in question. Compilation of a database for biomass yield in switchgrass, as we have undertaken here, attempts to address this concern for switchgrass. Issues of plot size and land quality as sources of potential bias are also examined. The empirical model derived from the database was used to provide a spatially explicit projection of biomass yield across the United States for upland and lowland ecotypes based on precipitation, temperature, and level of N fertilization. We conclude with a discussion of future research needs to better understand and then model biomass production for switchgrass.

### MATERIALS AND METHODS

#### Data Collection and Characterization

A survey of publications that reported switchgrass biomass yield was conducted with a focus on biofuels production in the United States. Yield data were compiled only for field trials, extracting data from text, tables, and graphs. As noted by Heaton et al. (2004), many studies have been published on switchgrass as a forage crop, but cultural practices between forage and biomass production can be quite different (e.g., harvest frequency). Therefore, only studies conducted with crop management practices typical of those used in bioenergy crop production were included in our analysis (see Data Analysis and Modeling section for details). Given these criteria, a total of 18 publications were identified, from which we compiled 1190 observations of biomass yield from 39 sites in 17 states (Table 1). Productivity estimates for 25 upland and 14 lowland cultivars, varieties, strains, or synthetic lines are included in this compilation. Rather than make a distinction, we will refer to all plant materials as cultivars. Readers should consult the original publication for details. For each observation, planting date, management practice, fertilization regime, harvest frequency, growing season precipitation (April–September) and plot size were characterized, preferably from the publication itself. When this was not possible, information was obtained through personal communication with the authors. Latitude and longitude for each site were also determined, which enabled us to estimate annual temperature and growing season precipitation.
Data Analysis and Modeling

In addition to descriptive statistics and bivariate plots, we also fit multivariate statistical models to the data. This began with an exploratory data analysis using a generalized additive model (GAM) framework (Hastie and Tibshirani, 1990). In a model of this type, the traditional linear or nonlinear parametric function relating response and predictor variables can be replaced by a non-parametric smoothing function. This provides a flexible method for visualizing relationships when the choice of a parametric function is not obvious. A GAM has the following model form:

\[
Y = \alpha + \sum_{i=1}^{p} f_i(X_i) + \varepsilon
\]

where \(Y\) is the response variable of interest, \(\alpha\) is the overall mean of \(Y\), \(f_i\) is a smoothing function for the \(i\)th of \(p\) possible predictor variables \(X_i\), and \(\varepsilon\) is a normally distributed error term. The smoothing functions, which are estimated iteratively using a back-fitting algorithm, describe locally persistent patterns in the data attributable to each predictor variable. Plots of smoothing functions together with partial model residuals can be used to assess the choice of a parametric function.

Terms in the model include an intercept to account for the fact that a nonzero yield is possible at extreme temperatures with no fertilization. We did not include an intercept term for precipitation. The full biomass yield model can thus be written as:

\[
y_{\text{pred}} = y_{\text{temp}} \times y_{\text{precip}} \times y_{\text{low}} \times y_{\text{fert}}
\]

where

\[
y_{\text{temp}} = \begin{cases} 1 + k_{\text{app}} \exp[-(x_t - t_{\text{app}})^2] & \text{for } x_t \leq t_{\text{app}} \\ 1 + k_{\text{app}} \exp[-(x_t - t_{\text{app}})^2] & \text{for } x_t > t_{\text{app}} \end{cases}
\]

\[
y_{\text{precip}} = k_p \sqrt{x_p}
\]

\[
y_{\text{fert}} = 1 + k_{\text{fert}} \times x_f
\]

\[
y_{\text{low}} = \begin{cases} \frac{1}{k_{\text{low}} - k_{\text{lat}} \times \max(0, \text{lat} - \text{lat}^*)} & \text{for upland} \\ \text{for lowland} \end{cases}
\]

and

\(x_t = \) annual average temperature (°C)  
\(x_p = \) growing season precipitation (mm)  
\(x_f = \) N fertilization (kg ha\(^{-1}\))  
\(\text{lat} = \) latitude (°N)

To estimate parameter values, we used the nonlinear least squares function \texttt{nls}, available for the statistical graphics and programming package \texttt{R} (Ikhaha and Gentleman, 1996). To limit \(k_1\) and \(k_2\) to strictly positive values, the logarithmic transformation of these coefficients was used as the fitted parameter, which was then re-exponentiated. Residual analysis revealed that the assumptions of regression were better met by fitting a logarithmic transformation of the yield equation. Initial observations confirmed the known association of low yields with first-year stands and four-cut harvest systems (data not shown). Yield data from first-year harvest and four-cut systems were therefore dropped from the regression analyses to better relate yield to variables of interest. Finally, some plots in one study were subjected to N fertilizer application rates of 448 and 896 kg ha\(^{-1}\) (Thomason et al., 2004). These values far exceeded those for other plots in our dataset and exerted disproportionate leverage on the model fit. We therefore excluded these points from our analysis.

The geographic distribution of yield across the continental United States was determined using the multiplicative, parametric model. We applied equations and parameter coefficients to spatially distributed 30-yr climate averages from PRISM and generated a 400 × 400 m resolution thematic raster map using Albers Equal Area Conic projections and ESRI ArcInfo (Redlands, CA). To represent both ecotypes on a single map, lowland yield projections are represented across the country, up to 38.2°N
RESULTS AND DISCUSSION

Biomass Yield, Cultivar Selection, and Crop Management Factors

The compiled database covers 39 field sites across 17 states from Beeville, TX (Sanderson et al., 1999a) in the south, to Munich, ND (Schmer et al., 2008) in the Midwest, to Rock Springs, PA (Sanderson et al., 2004) in the northeast. Lowland cultivars were generally planted in the south, whereas upland cultivars were planted across the full range of latitudes. For the database compiled, annual switchgrass yields varied from less than 1 Mg ha\(^{-1}\) to almost 40 Mg ha\(^{-1}\) (Fig. 1a). As indicated by a histogram of piled, annual switchgrass yields varied from less than 1 Mg ha\(^{-1}\) planted across the full range of latitudes. For the database compiled, annual switchgrass yields varied from less than 1 Mg ha\(^{-1}\) to almost 40 Mg ha\(^{-1}\) (Fig. 1a). As indicated by a histogram of the data distribution, the most frequently observed yield classes across all cultivars, soils, climate, and crop management practices were between 10 and 14 Mg ha\(^{-1}\). The frequency distribution for the total database was unimodal and skewed to the left, with a long tail at higher yields. Biomass yields greater than 28 Mg ha\(^{-1}\) were uncommon, but were reported for lowland ecotypes (e.g., Kanlow and Alamo) planted at field sites in Alabama, Texas, and Oklahoma (Sladden et al., 1991; Kiniry et al., 2005; Thomason et al., 2004). The highest annual yield reported in this compilation was 39.1 Mg ha\(^{-1}\) for the cultivar Alamo under high fertilization (200 kg N ha\(^{-1}\)) and in a year with high precipitation and temperatures (Kiniry et al., 1999). Frequency distributions for upland (Fig. 1b) and lowland (Fig. 1c) ecotypes were similar in shape (i.e., unimodal, skewed to the left, with long tails), with mean (±SD) biomass yields of 8.7 ± 4.2 Mg ha\(^{-1}\) and 12.9 ± 5.9 Mg ha\(^{-1}\) for the two ecotypes, respectively. Differences between ecotypes were significant (\(p < 0.001\)).

Just as yield differences were observed between upland and lowland ecotypes, there was considerable variation in biomass yields between cultivars within an ecotype (Fig. 2). Within the lowland ecotype, Alamo, SL941, SL931, Kanlow, NL942 and SL932 were the highest yielding cultivars with median rates of annual biomass production that ranged from 12.2 to 14.8 Mg ha\(^{-1}\) (Fig. 2). Kanlow and Alamo were the two most reported lowland cultivars in this compilation with 240 and 277 yield observations, respectively. The high-yielding lowland cultivars SL931, SL932, SL941, and NL942 had only four observations each. According to Cassida et al. (2005b), the cultivars with NL and SL designations are northern and southern lowland synthetic lines from Oklahoma and southern Kansas (i.e., NL942) and from central and southern Texas (i.e., SL931, SL932, and SL941). Among upland ecotypes, Cave-in-Rock, NE Late, HDMDC3, the experimental strain Late Synthetic HY, Shelter, and NU94 were the highest yielding cultivars with median rates of annual biomass production that ranged from 9.6 to 11.4 Mg ha\(^{-1}\) (Fig. 2). For these cultivars, Cave-in-Rock had far more observations associated with it (\(n = 119\)) than did NE Late (\(n = 10\)), HDMDC3 (\(n = 3\)), Late Synthetic HY (\(n = 14\)), Shelter (\(n = 62\)), and NU94 (\(n = 3\)).

Although variability was high, there was a distinct response of switchgrass yield to N fertilizer for both ecotypes (Fig. 3). For lowland ecotypes (Fig. 3a), there was a hint of an optimum around 100 kg N ha\(^{-1}\), but in many cases the zero fertilizer plantings did as well as fertilized stands. In upland ecotypes (Fig. 3b), yields appeared to respond to total rates of N application up to approximately 100 kg ha\(^{-1}\) and decreased above those rates, although there were fewer observations at higher application rates and none above 160 kg ha\(^{-1}\), making comparison to the lowland responses difficult. High levels of fertilization did not guarantee increased biomass production. This analysis does not take into account timing of N application, whether or not other nutrients were added, or the N availability in native soil. Thomason et al. (2004) did not see a strong response of biomass yield to variation in either rate or timing of N application, whereas Muir et al. (2001) saw yields increase with N application from zero to 170 or 224 kg ha\(^{-1}\). Parrish and Fike (2005) summarize the issue of N application in switchgrass as “unsettled” and suggest that the range of recommendations for N management “is not narrowing, nor is a central tendency developing.” They describe switchgrass as a plant that is inherently N-thrift y, especially when managed for biomass production. Although the biomass yield summaries presented here do not indicate an unqualified dependence on fertilization, and certainly not a strong response to high levels.
of application, a positive response to N at moderate application rates is likely an important management observation.

There were no apparent patterns of biomass yield in relation to plot size or row spacing across these experiments. Across the entries contained in our database, plot size varied from 2.88 m² (Casler and Boe, 2003) to 9.5 ha (Schmer et al., 2008). Yields ranged from less than 5 to more than 20 Mg ha⁻¹ across all plot sizes (Table 2). Although greater than 97% of the total number of yield observations came from plots less than 50 m² in size, there was no significant relationship between biomass yield and plot size (Fig. 4). For both lowland and upland ecotypes, the slope of the line was not different from zero. Schmer et al. (2008) recently reported biomass yields for cultivars established in field plots from 3 ha to almost 10 ha. Fields of the upland cultivars Cave-in-rock, Trailblazer, Shelter, and Sunburst were planted on 10 farms across a broad precipitation and temperature gradient in the midcontinental United States. Annual biomass yields for upland cultivars planted on marginal lands ranged from 5.2 to 11.1 Mg ha⁻¹ for the years in which harvests were conducted, with a site-wide average close to 7.2 Mg ha⁻¹ (Schmer et al., 2008). Such a yield was only slightly lower than that calculated for all upland cultivars in this compilation. Therefore, we find that across the many observations contained in the database, no plot size bias was observed for either lowland or upland ecotypes. This suggests that yields were not influenced by undue management attention to small individual plots.

Similar conclusions can be drawn about row spacing, which in our database varied from 15 cm (Casler and Boe, 2003) to 61 cm (Thomason et al., 2004). It appears that biomass yield in switchgrass, a perennial that spreads vegetatively, is not particularly sensitive to row spacing or planting density. Muir et al. (2001) reported that tiller density and, more importantly tiller mass, both compensate for variable row spacing in switchgrass. These authors noted that 18- to 25-cm rows were preferred especially in the establishment year since narrow rows allowed quicker canopy closure and greater weed control than did wide rows (i.e., 36–102 cm). Beyond the establishment year, however, row spacing was not a crucial determinant of biomass yield.

Information in the database was used to assess the potential response of biomass yield to soil characteristics. Switchgrass is native in regions representing a wide variety of soil types, and it may be adapted to perform well in a variety of soil textures,

![Box plot of biomass yields for upland and lowland cultivars of switchgrass.](image1)

**Fig. 2.** Box plot of biomass yields for upland and lowland cultivars of switchgrass. Bold horizontal lines indicate the median for each cultivar. Boxes represent the interquartile range (IQR, or middle 50%) of the yield values for each cultivar. Vertical whiskers extend to the furthest values up to 1.5 IQR. Additional points are shown as outliers. Box widths are proportional to the number of observations for each cultivar. Cultivars are grouped by ecotype, with gray horizontal lines indicating the median yield across cultivars within each ecotype.

![Biomass yield as a function of total N applied during the growing season.](image2)

**Fig. 3.** Biomass yield in (a) lowland and (b) upland ecotypes as a function of total N applied during the growing season.
in soils low in nutrients, and in otherwise marginal growing conditions. Both soil type and soil acidity have been discounted as major determinants of switchgrass productivity (Parrish and Fike, 2005), yet spatial variation in soil parameters may impact switchgrass yields within a single field (Di Virgilio et al., 2007). In our analysis, there was not an obvious response of biomass yield to variation in soil texture, i.e., no clear pattern as a function of percentage sand, silt, or clay (data not shown), at least at the available resolution of soil texture data. An interaction between precipitation and soil texture might impact yield, as relative amounts of sand, silt, and clay affect soil water-holding capacity, with implications for seedling survival and yield (Evers and Parsons, 2003; Parrish and Fike 2005). Likewise soil texture could influence rooting depth and nutrient availability, but these relationships were not within the scope of our analysis.

As an extension to this analysis, the response of biomass yield to land quality was also examined. In total, we were able to associate land capability classes as defined by the NRCS with the vast majority of sites. The only exception to this was one field trial in West Virginia (Fike et al., 2006). Across the 39 sites for which data were compiled, statistical analysis showed no significant relationship between productivity and land capability class for either lowland or upland ecotypes analyzed together or separately (Table 3). Although careful application of this and other yield models is warranted, and new data will further inform our understanding, we see no evidence of a high-yield bias due to land quality in our database. This may be in part because many of the switchgrass field trials established in the 1990s tended to be on marginal lands since a common expectation at the time was that less-than-prime agricultural lands would be used for woody and herbaceous energy crop production (Wright, 1994).

### Biomass Yield and Climate

Biomass yield varied as a function of average annual temperature for all ecotypes taken together as well as for individual ecotypes (Fig. 5). In general, yields increased with increasing temperature up to a point and then decreased. Curves of the same shape were observed whether yield was taken as a function of growing season temperature, annual temperature, maximum summer temperature, or minimum temperature during the preceding winter (data not shown). All of the temperature variables were highly correlated, except that maximum summer temperature was only weakly correlated with winter minimum temperature.

Regardless of the temperature variables used, the response patterns suggest a broad optimal temperature range for biomass yield, below and above which yields are diminished. Because the climate data were correlated, it cannot be determined whether conditions during all or part of the growing season, or even winter conditions, are key to understanding the temperature response of yields. Muir et al. (2001) suggested a southern limit of adaptation for Alamo between their sites in Stephenville, TX (32.2°N) and Beeville, TX (28.4°N) that was unrelated to rainfall or soil type; soils at both sites were described as shallow. Although high temperature cannot be singled out as the

### Table 2. Mean biomass yield (dry mass, Mg ha⁻¹) for lowland and upland ecotypes of switchgrass in represented plot size classes (m²). Standard deviation, number of observations, and mean plot size within a given plot size class are also given.

<table>
<thead>
<tr>
<th>Plot size class (range, m²)</th>
<th>0–3</th>
<th>3–6</th>
<th>6–9</th>
<th>9–12</th>
<th>12–15</th>
<th>15–18</th>
<th>27–30</th>
<th>36–39</th>
<th>79</th>
<th>147</th>
<th>175.5</th>
<th>&gt;30,000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean yield (Mg ha⁻¹)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lowland</td>
<td>4.2</td>
<td>14.8</td>
<td>17.6</td>
<td>13.8</td>
<td>12.7</td>
<td>12.5</td>
<td>11.8</td>
<td>8.2</td>
<td>na</td>
<td>19.1</td>
<td>14.6</td>
<td>na</td>
</tr>
<tr>
<td>(SD)</td>
<td>(0.6)</td>
<td>(4.9)</td>
<td>(4.3)</td>
<td>(6.6)</td>
<td>(3.5)</td>
<td>(5.2)</td>
<td>(6.0)</td>
<td>(4.3)</td>
<td>(11.2)</td>
<td>(7.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>2</td>
<td>62</td>
<td>22</td>
<td>188</td>
<td>30</td>
<td>228</td>
<td>30</td>
<td>71</td>
<td>10</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean plot size (m²)</td>
<td>2.4</td>
<td>5.9</td>
<td>7.0</td>
<td>9.7</td>
<td>13.2</td>
<td>17.6</td>
<td>27.3</td>
<td>37.2</td>
<td>147.0</td>
<td>175.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean yield (Mg ha⁻¹)</td>
<td>11.0</td>
<td>9.2</td>
<td>9.3</td>
<td>6.1</td>
<td>7.1</td>
<td>10.8</td>
<td>na</td>
<td>na</td>
<td>8.3</td>
<td>na</td>
<td>na</td>
<td>7.2</td>
</tr>
<tr>
<td>(SD)</td>
<td>(3.4)</td>
<td>(4.4)</td>
<td>(3.3)</td>
<td>(2.9)</td>
<td>(2.7)</td>
<td>(4.9)</td>
<td>(2.7)</td>
<td>(2.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>26</td>
<td>100</td>
<td>42</td>
<td>86</td>
<td>78</td>
<td>135</td>
<td>10</td>
<td>29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean plot size (m²)</td>
<td>2.8</td>
<td>5.6</td>
<td>7.0</td>
<td>10.9</td>
<td>13.8</td>
<td>17.4</td>
<td>79.0</td>
<td>67,862</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 4.** Relationship between biomass yield and plot size for lowland and upland ecotypes of switchgrass. Plot sizes were binned according to a given range and associated with biomass yield estimates contained within the database. Standard deviations for both plot size and biomass yield are shown. Plot size is graphed on a log scale.

### Table 3. Mean biomass yield (dry mass, Mg ha⁻¹) for lowland and upland switchgrass ecotypes in each of seven nonirrigated land capability classes (LCC). Standard deviation, median, number of observations, and standard error of the mean are also given.

<table>
<thead>
<tr>
<th>LCC</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowland</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>na</td>
<td>11.4</td>
<td>16.6</td>
<td>12.3</td>
<td>na</td>
<td>16.2</td>
<td>na</td>
</tr>
<tr>
<td>(SD)</td>
<td>(5.1)</td>
<td>(5.8)</td>
<td>(4.2)</td>
<td>(4.1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>median</td>
<td>11.1</td>
<td>15.7</td>
<td>13.1</td>
<td>16.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>283</td>
<td>191</td>
<td>58</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEM</td>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Upland |
| Mean | 10.7 | 5.9 | 10.3 | 7.5 | na | 9.6 | 6.3 |
| (SD) | (3.6) | (2.8) | (4.1) | (3.7) | (4.4) | (3.1) |    |
| median | 10.2 | 5.5 | 9.9 | 6.9 | 9.3 | 6.6 |    |
| n | 27 | 102 | 163 | 36 | 98 | 48 |    |
| SEM | 0.7 | 0.3 | 0.3 | 0.6 | 0.4 | 0.4 |    |
limiting factor, mean April to September temperatures were 2 to 3°C higher at the more southern site, averaging 26°C during the study. Such high temperatures could be limiting in themselves if they exceed a physiological optimum, or they might interact by increasing evapotranspiration and reducing soil moisture.

Growing season temperatures seem like an obvious correlate for yield, growth rate being expected to be temperature-dependent to some degree (Parrish and Fike, 2005). Spring temperatures, however, might also be key in some environments, for example, North Dakota (Berdahl et al., 2005) and eastern Canada (Madakadze et al., 1998b), as cool temperatures and short growing seasons limit switchgrass growth potential. A focus on spring temperatures would thus isolate temperature requirements for initiation of new growth. In southwestern Quebec, for example, initial spring growth was hastened by up to 35 d by a warmer spring (Madakadze et al., 1998a). Leaf area duration also correlated with yield, and this varied across cultivars, highlighting the interaction between genetics and environment. Conversely, minimum winter temperatures play a role in determining winter survival. Vogel et al. (2002) found genetic variation in winter survival and subsequent sward recovery within selections from a single field population. The duration of low winter temperatures could affect the length of the growing season and could impact survival, particularly of lowland ecotypes. The northern distribution limit of switchgrass is assumed to be at least partially a function of cold winter temperatures (Vogel et al., 2002; Berdahl et al., 2005; Casler, 2005; Parrish and Fike, 2005).

Switchgrass yield as a function of precipitation was variable, and, unlike the pattern seen for biomass yield in relation to temperature, low yields were observed across all values of precipitation, with no strong correlation between yield and precipitation in either ecotype (Fig. 6). If the upper boundaries of the relationship represent the maximum potential yield, however, low precipitation during the growing season did appear to limit yield. Particularly in the upland ecotype, yield increased with an increase in April to September precipitation, up to a threshold of approximately 600 mm. Above 600 mm, precipitation did not limit growth in either ecotype.

Total precipitation is only one factor contributing to soil moisture availability. The timing and size of rainfall events are important modifiers; in particular, sufficient rainfall must occur during more critical portions of the growing season. Sanderson et al. (1999b) reported that high yields at five east Texas locations were associated with years when April to September precipitation was high, and Berdahl et al. (2005) found that low April to September precipitation severely limited yield for eight upland cultivars in North Dakota. Narrowing the sensitive time frame in their study, Muir et al. (2001) correlated March to August precipitation with yield at Stephenville, TX. Over a 4-yr period in South Dakota, biomass production was best explained by a linear relationship with April to May precipitation (Lee and Boe, 2005). Because cultivars differ in phenology (Berdahl et al., 2005; Casler, 2005; Parrish and Fike, 2005) and locations differ in daylength and growing season, it is not surprising that the period of sensitivity might vary among studies. Management factors play a role as well; low precipitation in August to September reduced yields in a two-cut, but not in a one-cut system (Reynolds et al., 2000).

**Empirical Modeling and Spatial Extrapolation**

Plots of GAM smoothing functions and partial residuals (Fig. 7) reveal that in a multivariate model accounting for climate and N fertilization, lowland ecotypes produce a consistently higher yield than upland ecotypes. Yield is also seen to increase steadily with increasing precipitation, more so than in bivariate plots (Fig. 6). Yield increases with increasing temperature up to an annual average of about 14°C, at which point it declines, similar to the observed bivariate relationship.
(Fig. 5). With ecotype, precipitation, and temperature included simultaneously in the model, an approximately linear N fertilization effect becomes apparent. All four predictors are highly significant, and the overall model explains about 34% of the variation in the natural logarithm of biomass yield.

The nonlinear parametric model provides a more quantitative description of the effects of the variables influencing biomass yield (Table 4). For example, parameter estimates indicate that lowland cultivars produce approximately 1.5 times the yield of upland cultivars and that the annual average temperature optimum is about 14.14°C. Biomass yield drops at annual average temperatures above this optimum faster than it increases at temperatures below this value. Precipitation has a significant positive effect on yield, as does N fertilization. At latitudes greater than 38.2°N, the yield of lowland cultivars declines at a rate of about 12.5% per degree latitude. This yield decline at northerly latitudes is consistent with the results of Casler et al. (2004). Such a rate implies that the productivity advantage of lowland relative to upland cultivars disappears at latitudes greater than 38.2° N, consistent with the findings of Casler et al. (2004). Overall, the parametric model fits the measured yield data about as well as the nonparametric GAM, with a R² value of 0.34 (Fig. 8).

Analysis of model residuals (data not shown) indicated compliance with the assumptions of regression, including approximate independence, normality, and homoscedasticity with respect to the predictor variables. Additional linear regressions of residuals against other variables in the data set, including plot size, year harvested, and stand age after the establishment year did not reveal any significant remaining relationships at the p = 0.05 level. Regressions against measures of land quality (e.g., land capability classes) did not reveal any evidence of higher residual yields associated with higher quality lands. Regressions of residuals against cultivar did show some significant relations, with cultivar accounting for an additional 5% of the variability in yield.

While our empirical model does account for broad interannual differences in weather, we expect that there are many other weather-related factors that can account for some of the 60 to 65% unexplained variability that remains. For example, the particular seasonal distribution of precipitation and temperature creates unquantified variability, as do their interactions. This is because yield can be affected by not only total precipitation and annual average temperature, but also by minimum winter temperature, an unusually cold, wet spring, or by a partial season drought, either in early and mid-summer, or directly after the first harvest in a two-harvest system. Differences in irradiance due to cloud cover (confounded with precipitation and temperature), or seasonal damage from insects or pathogens may also cause variability across yield measurements made at a single location. Such occurrences have been documented, but not frequently enough to be statistically useful in our analysis.

As our mapping exercise focuses on representing variation across, rather than within, locations, we re-estimated the R² value of our model after removing within-site variability. We did this by calculating the average observed yield for each ecotype at each location and comparing it to the yield of each ecotype predicted for the average conditions at each location in the year for which yield measurements were made. This calculation produced a significantly higher R² of 0.58 and a lower root mean square error (RMSE) of 3.03 Mg ha⁻¹. We believe these statistics, based on average yields and average climate at a site, to be more analogous to the mapping exercise, and more representative of its accuracy than the statistics reported in Table 4.

The variation in yield across locations that remains, beyond the unquantified within-season weather patterns, is likely due to factors not included in our dataset. For example, few data were

<table>
<thead>
<tr>
<th>Parameter†</th>
<th>Estimate</th>
<th>Lower 90% CL</th>
<th>Upper 90% CL</th>
</tr>
</thead>
<tbody>
<tr>
<td>k_{low}</td>
<td>1.51</td>
<td>1.44</td>
<td>1.59</td>
</tr>
<tr>
<td>k_{opt}</td>
<td>0.927</td>
<td>0.795</td>
<td>1.060</td>
</tr>
<tr>
<td>k_1</td>
<td>0.028</td>
<td>0.016</td>
<td>0.050</td>
</tr>
<tr>
<td>k_2</td>
<td>0.118</td>
<td>0.078</td>
<td>0.176</td>
</tr>
<tr>
<td>t_{opt}</td>
<td>14.14</td>
<td>13.50</td>
<td>14.78</td>
</tr>
<tr>
<td>k_p</td>
<td>0.196</td>
<td>0.177</td>
<td>0.215</td>
</tr>
<tr>
<td>k_{fert}</td>
<td>0.0025</td>
<td>0.0017</td>
<td>0.0032</td>
</tr>
<tr>
<td>k_{lat}</td>
<td>0.125</td>
<td>0.080</td>
<td>0.171</td>
</tr>
<tr>
<td>lat*</td>
<td>38.18</td>
<td>36.60</td>
<td>39.75</td>
</tr>
<tr>
<td>R²</td>
<td>0.34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSE</td>
<td>4.62 Mg ha⁻¹</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
available for edaphic factors, including differences in soil type, structure, water-holding capacity, or fertility, including N availability before fertilization. If these factors are measured and added to future analyses, there may be potential for further reducing the variability that remains unexplained across locations.

Mapped output from the model (Fig. 9) was relatively consistent with the natural geographic range of switchgrass, although significant yields were also predicted along the Pacific coast states. This prediction is consistent with results from USDA field trials (Fransen et al., 2006) The map displays both predicted and observed yields of switchgrass as background colors and circles, respectively. As shown by the observed values displayed on the map, lowland cultivars (red circles) outperformed the upland cultivars (black circles) in mid- to southern latitudes. Maximum biomass yields are projected in a corridor westward from the mid-Atlantic coast region to Kansas and Oklahoma. Beyond the Great Plains region, our model predicts that switchgrass yields are limited by low precipitation and to a lesser extent, low mean annual temperature.

Switchgrass production estimates in some areas are extrapolated outside the climatic conditions captured by our model and the original dataset. These are outlined in black in Fig. 9. Specifically, at some locations in the Rocky Mountains, annual average temperatures are below the lowest value in our dataset.

Fig. 8. Measured vs. predicted yield from the parametric model. Gray points represent upland cultivars; black points represent lowland cultivars. Size of points is proportional to experimental plot size. Solid line represents a 1:1 fit, and dashed lines represent the limits of the 80% predictive interval. Five plots had measured yields greater than 30 Mg ha\(^{-1}\) and therefore fall outside the range of the plot.

Fig. 9. Map of switchgrass biomass yields across the United States predicted by the empirical model. Simulations assume annual application of 100 kg N ha\(^{-1}\). Thin circles represent observed yield as reported in the literature, black for upland and red for lowland ecotypes. Diameter of circle is proportional to yield, from 1 to 39 Mg ha\(^{-1}\). Apparent thick circles are the result of multiple yield observations too similar to be resolved at this scale. Regions of the United States where extrapolations of switchgrass production fall outside the precipitation and temperature conditions used to parameterize our model are outlined in black. Small enclosed areas may appear as solid colors.
In southern Florida and Texas, annual average temperatures exceed our highest value of 21.8°C. Arid regions west of the Rockies and east of the Sierra and Cascade ranges receive growing season precipitation lower than the minimum value in our dataset (188 mm). Conversely, the very high precipitation found on the Olympic Peninsula is above our highest observed value (1210 mm). Finally, areas of New England and along the Appalachian mountains are within the temperature range of our dataset, but have higher rainfall for a given temperature than any location included in our dataset.

Refinement of the current statistical model with additional data from the regions not currently represented in our study might improve the prediction of switchgrass yields with respect to climate and interactions with cultivar and management practices. This would assist growers in choosing high-yielding cultivars within the context of local environmental growing conditions. Moreover, geographically distributed models of bioenergy crops, such as this one, can play an important role when investigating issues related to land use, potential regional or national biofuels production and other considerations related to facility siting and supply logistics.

Evaluation of Bias and the Possibility of Yield Overestimation

In their study of biofuel yields from agricultural crops, Johnston et al. (2009) mention that a number of variables can exert a considerable influence on yield including geographic location, climate, soils, and management practice. In our analysis, these factors have been evaluated and their impact on biomass yield considered to the extent possible based on available data for switchgrass grown in the United States. Temperature, precipitation, N application, ecotype, and latitude exhibit a statistically significant impact on yield, and thus a strong case can be made for including these in an empirical model. However, a surprising number of variables did not exhibit a significant impact on yield for upland or lowland ecotypes: plot size, row spacing, stand age (after establishment), soil texture, and land quality. Especially interesting is the observation that no evidence was found to support an effect of plot size and land quality on biomass yield. Small plot size and preferential placement of field trials on better than average agricultural land have been cited as factors that contribute to overestimation of yields that might otherwise be realized on typical farm-scale fields. For example, Johnston et al. (2009) cited several factors commonly responsible for overestimating yield of bioenergy crops, at regional to continental scales, including uncritical application of site-specific yield estimates, failure to use regionally specific and spatially explicit agricultural data, and use of potentially unrealistic yield estimates that might come from the location of field trials on agricultural research stations. It is indeed important to give these factors due consideration and future studies should take great care to fully characterize the climatic, edaphic, and crop management practices that contribute to reported yields. One of the lessons learned from the current analysis is that more yield data would be desirable over a broader range of conditions (e.g., northern colder and western dryer sites). Although we acknowledge these and other limitations of available data, and the impact they may have on our analysis of yield, we find no evidence for overestimation of yield with respect to the soil, crop management, or climatic variables considered. We systematically looked for, but did not find, evidence of bias toward higher yields associated with small plots or preferential establishment of stands on high quality land. Thus, we conclude that yield estimates reported in the literature and summarized here for switchgrass are at this time the best available predictors of the yield for this bioenergy crop grown under large-scale cultivation. Biomass yield is, however, just one component of biofuels production. Another is the potential land base available for energy crop production in the United States (Graham, 1994). We believe this to be the greater uncertainty as it includes the relative economics of biofuels production and alternative uses of agricultural land. These uncertainties must be considered, but are well beyond the scope of this analysis.

CONCLUSIONS

Here we report a large compilation and analysis of empirical biomass yield data from across the United States for the biofuels crop switchgrass. Each observation has been identified by latitude and longitude, ecotype, cultivar, plot size, stand age, harvest frequency, N fertilization, and year of harvest. Other sources of information were used to associate observations of biomass yield with soil characteristics, temperature and precipitation, and land quality. Of the relationships established between biomass yield and these variables, ecotype, temperature, precipitation, and N fertilization were identified as the most important predictors of yield. Statistical modeling revealed patterns of response that could be used to predict biomass production associated with a given climate. In addition, no systematic bias was observed for an overestimation of yield due to small plot size or preferential establishment of stands on high quality land. Results thus suggest that field trials conducted and reported in the current literature can be used, as we have, to extrapolate observed biomass yields to larger spatial scales. We acknowledge, however, that this should be done with care and with a full understanding of the uncertainties that accompany such extrapolations.

ACKNOWLEDGMENTS

The authors wish to thank those site investigators who provided additional information and data for the analyses presented. The authors would also like to thank Dr. Jonathan Chipman, the director of the Applied Spatial Analysis Laboratory at Dartmouth College, for advice, consultation, and access to the lab facilities. This research was sponsored by the U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Biomass Program. Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the U.S. Department of Energy under contract DE-05-00OR22725. E.B Davis was supported by a grant from the Morgan Family Fund.

REFERENCES


