Environmental limitation mapping of potential biomass resources across the conterminous United States

CHRISTOPHER DALY1, MICHAEL D. HALBLEIB1, DAVID B. HANNAWAY2 and LAURENCE M. EATON3

1PRISM Climate Group, Northwest Alliance for Computational Science and Engineering, 2000 Kelley Engineering Center, Oregon State University, Corvallis, OR, USA, 2Department of Crop and Soil Science, Oregon State University, 125 Crop Science Building, Corvallis, OR, USA, 3Bioenergy Resource and Engineering Systems Group, Environmental Sciences Division, Oak Ridge National Laboratory, PO BOX 2008 MS6036, Oak Ridge, TN 37831-6036, USA

Abstract

Several crops have recently been identified as potential dedicated bioenergy feedstocks for the production of power, fuels, and bioproducts. Despite being identified as early as the 1980s, no systematic work has been undertaken to characterize the spatial distribution of their long-term production potentials in the United states. Such information is a starting point for planners and economic modelers, and there is a need for this spatial information to be developed in a consistent manner for a variety of crops, so that their production potentials can be intercompared to support crop selection decisions. As part of the Sun Grant Regional Feedstock Partnership (RFP), an approach to mapping these potential biomass resources was developed to take advantage of the informational synergy realized when bringing together coordinated field trials, close interaction with expert agronomists, and spatial modeling into a single, collaborative effort. A modeling and mapping system called PRISM-ELM was designed to answer a basic question: How do climate and soil characteristics affect the spatial distribution and long-term production patterns of a given crop? This empirical/mechanistic/biogeographical hybrid model employs a limiting factor approach, where productivity is determined by the most limiting of the factors addressed in submodels that simulate water balance, winter low-temperature response, summer high-temperature response, and soil pH, salinity, and drainage. Yield maps are developed through linear regressions relating soil and climate attributes to reported yield data. The model was parameterized and validated using grain yield data for winter wheat and maize, which served as benchmarks for parameterizing the model for upland and lowland switchgrass, CRP grasses, Miscanthus, biomass sorghum, energycane, willow, and poplar. The resulting maps served as potential production inputs to analyses comparing the viability of biomass crops under various economic scenarios. The modeling and parameterization framework can be expanded to include other biomass crops.

Keywords: biomass crop, biomass production potential, biomass resource map, biomass resources, biomass sorghum, energycane, miscanthus, PRISM-ELM, Sun Grant, switchgrass

Received 13 February 2017; revised version received 27 July 2017 and accepted 30 August 2017

Introduction

In 2005, the US Department of Energy (USDOE) released its Billion Ton Study (updated in 2011 and 2016), which envisioned an expansion of domestic bioenergy production to one billion tons per year as a way to increase and diversify the nation’s energy resources (USDOE, 2005, 2011, 2016). Presently, the US bioeconomy consumes roughly one million tons per day for the generation of power, fuels, and chemicals from agricultural, forestry, and waste resources (USDOE, 2016). To achieve a domestic billion-ton bioeconomy, an additional 635 million tons per year of biomass must be produced on an annual basis from US land resources. The near-term potential can be generated from agricultural and forestry residues and waste resources equal to approximately 345 million tons per year. Traditional agricultural crops such as wheat and maize provide residues that can serve as sources of biomass; these crops have long production histories and rich knowledge bases with regard to physiology, production, and spatial distribution. To fill the supply deficit, dedicated bioenergy crops have become a subject of national focus.
Several crops have been identified as potential dedicated bioenergy crops for the production of power, fuels, and bioproducts. Despite many crops being identified as potential feedstocks as early as the 1980s, they still have little commercial production history in the United States, and hence, relatively little is known about the spatial distribution of their long-term production potential across the United States (Evens et al., 2010). Such information is a starting point for planners and economic modelers tasked with assessing land requirements, management options, harvest and transportation methods, processing needs, and infrastructure for biomass crops. Equally important is the need for this spatial information to be developed in a consistent manner for a variety of crops, so that their production potential can be intercompared to support crop selection decisions (Miguez et al., 2012; Castillo et al., 2015).

Efforts to map the spatial distribution of biomass resources in the United States have focused on one or two biomass crops at a time, with several potential biomass crops receiving little attention. Two approaches to mapping biomass resources are empirical modeling and mechanistic plant growth modeling. Commonly used empirical approaches have involved statistical extrapolation of plot or field-level yield data to larger regions using climatic envelope methods (e.g., Jager et al., 2010; Wullschleger et al., 2010; Tulbure et al., 2012). The main drawback of empirical approaches has been a lack of suitable yield data (Miguez et al., 2012), and a limited ability to extrapolate beyond the range of the explanatory data (Jager et al., 2010). Relationships between yield data and environmental conditions can be masked and even misled by factors other than environment, such as fertilization, cutting rotation, supplemental irrigation, and other management practices and economic considerations, making it difficult to quantify what the actual environmental tolerances are (Jager et al., 2010). Information needed to control for these factors is not always available in the literature, and access to researchers who conducted the trials is often limited. In addition, yield histories can be as short as a single year and are thus affected by year-to-year variability in weather conditions, making it difficult to estimate long-term yield potentials (Lobell et al., 2009). Finally, yield data are typically collected from demonstration plots in areas where the crop is likely to succeed, and thus provide little guidance as to how environmental factors limit production near the edges of a crop’s range or across steep climatic gradients (Miguez et al., 2012). Despite these shortcomings, empirical approaches provide important assessment tools for planning activities and supply guidance for more mechanistic modeling approaches (Jager et al., 2010).

Mechanistic plant growth models attempt to simulate the important physiological processes that affect growth, development, and yield. Examples of simulation models that have been used to model biomass crops include ALMANAC (Kiniry et al., 2008), EPIC (Williams et al., 1984; Brown et al., 2000; Thomson et al., 2009; Balkovic et al., 2013), MISCANFOR (Hastings et al., 2009; Miguez et al., 2012), and STICS (Brisson et al., 2008; Strullu et al., 2015). These models have the potential to provide detailed information on crop performance and yield. However, they require significant inputs of environmental data and detailed knowledge of crop physiology (e.g., Brown et al., 2000). In addition, calibration and validation of models require detailed plot-level data, which is often scarce or poorly distributed for many new crops (Nair et al., 2012). Parameterization of some models to specific crop varieties and locations can make it difficult to generalize results over large areas (e.g., Miguez et al., 2012). As more information on bioenergy crops becomes available, mechanistic models will become increasingly useful in planning for a biobased economy.

The resource mapping approach described here stems from the need for many different biomass crops to be compared within the same modeling framework to avoid confounding model differences with biological differences (Miguez et al., 2012). It stems from the recognition that many biomass crops have insufficient yield data from which to spatially extrapolate and estimate long-term yields. In addition, little quantitative information is available on the tolerances of these crops to environmental conditions. Our approach, undertaken as part of the Sun Grant RFP, was to take advantage of the informational synergy realized when bringing together field trials, close interaction with expert agronomists, and spatial modeling into a single, collaborative effort. The first component consisted of a coordinated set of field trials of several of the most promising herbaceous and woody biomass options conducted over a 3- to 7-year period (Lee et al., 2017; Volk et al., 2017), plus other relevant trials. The spatial representativeness of the coordinated field trials was optimized whenever possible through adherence to consistent, best-practice management protocols, thus controlling for the effects of management on the responses of crops to basic environmental limitations created by climate and soils. The second component was face-to-face interactions between the modeling group and the agronomists conducting the RFP and other field trials. During these meetings, yield data from the field trials were evaluated for their quality and representativeness, published literature was examined, and qualitative information on a crop’s spatial distribution based on personal experience was provided.

© 2017 The Authors. GCB Bioenergy Published by John Wiley & Sons Ltd., 10, 717–734
The third component was a biogeographical modeling and mapping system called Parameter-elevation Regressions on Independent Slopes Model Environmental Limitation Model (PRISM-ELM). An early version of PRISM-ELM was first developed to estimate the potential suitability zones of US-grown perennial grass exports to China (Hannaway et al., 2005). PRISM is the name of the system used to generate high-quality, spatial climate datasets that drive the model (Daly et al., 1994, 2008). PRISM-ELM was designed to answer a basic question: How do climate and soil characteristics affect the spatial suitability and long-term production patterns of a given crop? It draws from both empirical and mechanistic approaches and therefore falls into a hybrid category that is becoming more powerful as high-quality climate, remote sensing, land use, and soils data become available (Song et al., 2015; Wightman et al., 2015; Richter et al., 2016). It employs a simple water balance model to simulate the correspondence, or lack thereof, between water availability (based on precipitation and soil moisture) and growing season timing (based on a temperature response curve). The model uses simplified metrics to represent complex processes. January mean minimum temperature and July mean maximum temperature are used to identify areas that have cold or warm-season temperature extremes that may limit meaningful crop production. Soil pH, salinity, and drainage response curves also serve as metrics for unsuitable soil conditions. The focus is on a general approach to modeling climatic and soil constraints on biomass production for any crop, rather than a detailed accounting of the particular phenology or other morpho-physiological features of a given species or genotype. Suitability maps estimated by PRISM-ELM are transformed into yield potential maps through statistical regressions between the level of environmental suitability and biomass yield data from the field trials. These maps serve as potential production inputs to analyses that compare the viability of biomass crops under various economic scenarios (USDOE, 2016).

The objective of this article was to present (1) a description of our biomass resource mapping process, including an overview of the work flow and interaction with RFP agronomists; (2) PRISM-ELM model underpinnings, structure and function; (3) model validation and parameterization; (4) environmental suitability mapping; (5) and transformation of environmental suitability to biomass yield potential. Dedicated herbaceous biomass crops included in the RFP evaluation, and in this article, were upland and lowland switchgrass (Panicum virgatum L.), Giant Miscanthus (Miscanthus × giganteus), energycane (Saccharum officinarum L. × Saccharum spontaneum L.), biomass sorghum (Sorghum bicolor), and mixed Conservation Reserve Program (CRP) grasses (Lee et al., 2017). Woody biomass crops included willow (Salix spp.) and poplar (Populus spp.) (Volk et al., 2017).

Materials and methods

Data and processing

Climate data. Climate inputs for PRISM-ELM were grids of daily maximum, mean, and minimum temperature (T_max, T, and T_min, respectively) and precipitation (P) from the PRISM AN81d dataset (PRISM Climate Group, 2015). PRISM climate datasets have been peer-reviewed and used in many agricultural and natural resource applications (Daly et al., 1994, 2008). The daily data were summarized at a semi-monthly time step for use in PRISM-ELM; temperature values were averaged and precipitation values summed twice each month for the period 1981–2010, resulting in 720 grids. Each semi-monthly grid was then averaged across each of the thirty grids representing that semi-monthly period (e.g., the first half of January) to obtain 30-year averages. The result was 24 semi-monthly averages representing a 1981–2010 climatological ‘year’. Spatial resolution of the gridded data was 30 arc-seconds, or approximately 800 m, across the continental United States.

PRISM-ELM required ET_0 and bare soil evaporation as inputs. Given that only temperature and precipitation were available from the PRISM climate dataset at the time of access, daily ET_0 was estimated using methods outlined by Hargreaves & Samani (1985), which requires T_max, T, T_min, and estimates of extraterrestrial radiation based on solar geometry. Daily ET_0 values were summed to semi-monthly totals. Soil evaporation (E_s) over each semi-monthly time step was estimated as a proportion of ET_0, which varies with rainfall frequency (Allen et al., 1998). Once calculated on a semi-monthly basis for each year, ET_0 and E_s were averaged over the 1981–2010 climatological period in the same manner as temperature and precipitation.

Soils data. Soil characteristics greatly influence the suitability of plants for a particular location and their potential production. Important factors include water holding capacity, pH, salinity, and drainage. Soils data were obtained from the USDA Natural Resources Conservation Service (NRCS) in the form of the U.S. General Soil Map Coverage (http://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/geo/?cid=nrcs142p2_053629). The data were available as shapefile polygons and related data tables. Using standard GIS tools to view and query the data, the NRCS ‘representative’ data fields were selected that contained the variables for available water holding capacity (AWC), soil pH, salinity, and drainage class for each polygon. The linked polygon data for each variable were cast to an 800-m grid that was coincident with the 800-m PRISM climate data. The smallest General Soil Map Coverage polygon is 1012 ha in size, which is an area equal to about 4 × 4 800-m grid cells. NRCS SSURGO data, while at a much higher spatial resolution, were
not used in this study, because at the time of access, the data were not yet complete and consistent over the entire modeling domain.

_Yield data._ County-level grain yield data from winter wheat and maize, commonly known as cool-season and warm-season crops, respectively, were used to initially calibrate and validate PRISM-ELM. These data are described in Supporting Information, Data S6.

Yield data from biomass crop field trials were used in the parameterization of PRISM-ELM and the transformation of PRISM-ELM suitability estimates into potential annual biomass production. The yield trials used are summarized in Table S1. Details on the RFP yield trials for herbaceous and woody crops are provided in Lee et al. (2017) and Volk et al. (2017), respectively. RFP yield trials were conducted in a coordinated fashion, which provided a unique opportunity to control for management practices across sites by selecting trials that were most internally consistent. Since management practices greatly influence yields (e.g., Wullschleger et al., 2010), controlling for these practices allowed the modeling work to focus on how climate and soil constraints influence potential yield production patterns. Management practices were designed to approximate those used in farm-scale production. Trials conducted outside the RFP were also evaluated in a similar manner, although information on management practices and other details was sometimes not as readily available as that from the RFP trials.

Each of the RFP yield trials was evaluated in face-to-face meetings with the agronomists that were directly responsible for the trial. This allowed insight into the data that was not obvious when examining yield values alone; examples included reports of damaging single day weather events, unusual field conditions, residual pesticides, or other management issues. This additional information about the yield data helped to determine if they met the inclusion criteria for this study. These criteria were developed based on producer needs for maps that represent long-term production potential at field scale, assuming best management practices, including minimal inputs of fertilizer and pesticides. It was understood that the field trials lacked a long history and consisted of only 3–7 years, thus reducing the strength of the relationship to long-term average yields. To be most useful, the field trials were selected to identify those that represented:

- Dryland conditions (nonirrigated).
- Absence of significant damaging weather events and field conditions.
- Yields from the best local cultivar available at the time of the yield trials.
- Once-per-year harvest frequency.
- Estimated mean annual volume increment (MAI) at maturity for woody perennials (defined as total increment divided by age).
- Field-scale yields, as opposed to test plot-scale yields.
- Yields of fully established crops, if perennials; establishment years not included.

- Best-practice fertilizer application using a combination of pre-establishment soil test recommendations and mass balance approach to replace only what is removed by the crop.
- Best-practice pesticide application, typically minimal inputs.

Mapping process overview

The mapping process took advantage of the informational synergy realized when bringing together three components – field trials, close interaction with expert agronomists, and spatial modeling – into a single, collaborative effort. An overview of the process is shown in Figure 1. PRISM-ELM was provided with gridded climate and soils data, and a preliminary control file with crop-specific parameters was developed. PRISM-ELM produced an initial grid of the Environmental Suitability Index (ESI) from 0 to 100%, where 100 represented no climate or soil constraints on production and zero represented a full limitation.

For a given crop, yield data from field trials conducted by RFP agronomists and others were examined at face-to-face meetings with the modeling group. During this meeting, each yield data point was evaluated for adherence to the inclusion criteria presented previously. The initial PRISM-ELM ESI grid was also used to provide a framework for evaluating the yield data. The goal of each meeting was to come to an agreement on which yield data points would be included in a national regression function relating PRISM-ELM ESI to field trial yield. This nationwide regression function allowed the PRISM-ELM ESI grid to be transformed into a potential yield grid. The process of adjusting PRISM-ELM crop parameters and comparing the ESI map to the observed data was done iteratively during and subsequent to the meeting until a final solution was reached that was consistent with expert opinion, yield data, and published literature. Attempts were made to achieve the best agreement possible between PRISM-ELM and yield data, but within the constraints of model parameter values that were consistent with the type of crop being mapped (see Model parameterization section).

Model rationale

PRISM-ELM is based on the well-understood biogeographical tenant that long-term climate and soil conditions place limits on average plant production across the United States. On an annual average basis, precipitation, and hence dryland production, becomes increasingly limited as one moves from east to west across the Great Plains (Fig. 2a). The seasonality of precipitation determines the likelihood of successfully growing cool-season crops vs. warm-season crops. Over much of the eastern United States, average precipitation is sufficient for most crop production during the warm season, but in the West, very little precipitation falls during the warm season (Fig. 2b). Long-term average annual temperature largely determines the north–south and elevational range of crop species and varieties, and the timing of their production cycles (Fig. S1a). In addition, winter cold can limit the production of overwintering plants (Fig. S1b) and summer heat can limit production during the growing season (Fig. S1c).
In addition to climatic constraints, plant production is limited by soil characteristics, four of which are AWC, pH, salinity, and drainage. Shallow, sandy, or rocky soils have a low AWC, which limits their ability to store water, thus requiring greater precipitation inputs to maintain a water balance suitable for plant growth. Soil AWC is highly variable across the United States, but is greatest in parts of the Great Plains and Midwest (Fig. S2a). Very acid and alkaline soils decrease the solubility of many major plant nutrients and may also release toxic amounts of trace metals harmful to plant life. Soils are typically alkaline in arid areas of the West, acidic in parts of the east coast, and slightly acidic to neutral in the Midwest (Fig. S2b). Highly saline soils reduce the osmotic potential of the soil solution and may limit the uptake of some nutrients. High soil salinity is primarily found along coastlines and in arid areas of the western United States (Fig. S2c). Poorly drained soils can limit oxygen residing in soil pore spaces, necessary for healthy root activity. In contrast, water may leach rapidly through well-drained sandy soils, flushing nutrients in addition to storing little water. Soils are typically well drained in the western United States, but much of the eastern United States is poorly drained, especially in the Midwest (Fig. S2d).

Model organization

PRISM-ELM is composed of series of algorithms and metrics that evaluate the major climate and soil limiting factors to production discussed above. The PRISM-ELM ESI is the lowest suitability index resulting from the model response functions as follows:

\[
\text{ESI} = \min(S_w, S_c, S_h, S_p, S_s, S_d),
\]

where \(S_w\), \(S_c\), \(S_h\), \(S_p\), \(S_s\), and \(S_d\) are the suitability indexes from the water balance model, and response functions to winter low temperature, summer high temperature, soil pH, soil salinity, and soil drainage, respectively.

The water balance model contains generalized process-based algorithms that account for soil water availability, use and deficit, and works in concert with a temperature response curve. The other functions consist of response curves that serve as metrics for climate and soil processes that could limit plant production. These include the potential for low-temperature injury of overwintering crops, damage or growth reduction due to heat during the growing season, and plant responses to soil pH, salinity, and soil drainage, respectively. Each of these functions is summarized briefly below; model equations and further details are provided in Supporting Information, Data S3 and S4.

Water balance model. The PRISM-ELM water balance model is an Food and Agriculture Organization (FAO)-style function (Allen et al., 1998), operating on a semi-monthly time step, using 30-year average climate data described previously. Gridded inputs to the model are soil AWC, and semi-monthly average \(T\), \(P\), ET\(_0\) and \(E_s\). Crop-specific scalar inputs provided by the user are parameters defining the optimum temperature growth curve; average crop rooting depth \(D_{root}\); the crop coefficient \(K_c\), which encompasses canopy characteristics (e.g., height, coverage), stomatal control, and other factors that affect...
crop evapotranspiration; and the stress response factor \( p \), which is the fraction of soil water that a crop can extract from the root zone without suffering water stress. Gridded internal variables calculated by the model are the temperature response \( T_r \), actual evapotranspiration \( ET_a \), water stress coefficient \( K_s \), total available water in the root zone \( TAW \), readily available water in the root zone, and root zone water depletion \( D_r \). In concert with the water balance calculations, the temperature response of the crop is evaluated at each semi-monthly time step. User-defined parameters describe the mean daily temperature at which production is optimal, and the maximum and minimum temperatures at which production declines to zero.

At each time step \( t \), a water balance suitability index \( S_t \) is calculated as the product of the water stress coefficient and the temperature response \( S_t = K_s T_r \). The semi-monthly values of \( S_t \) are averaged to create monthly values \( S_m \). A Potential Suitability Window, the period within which a crop is expected to be in its most active production phase in an agricultural setting, is set by the user. This window is necessary because PRISM-ELM, being an environmental suitability model, does not simulate the timing of the life cycle stages of a crop. Within that window, the final water balance suitability \( S_w \) is calculated by the model as the average suitability during the Maximum Suitability Window, which is typically the three consecutive months for which the monthly suitability is highest (the number of consecutive months can be changed by the user). Water balance model equations are provided in Supporting Information, Data S3, and examples of its operation in two contrasting parts of the country are given in Supporting Information, Data S4.

Heat and cold temperature responses. The winter low-temperature response function is a metric for production limitations in overwintering crops that may occur because of injury or death caused by excessively low temperatures (Levitt, 1980; Beck et al., 2004). Conversely, in some species, low winter temperatures are required for induction of the plant’s flowering response (vernalization) through accumulation of chilling hours (Dennis, 1984). While low temperatures may be needed to maximize flowering and grain production in crops such as wheat, diversion of energy away from vegetative production and into flowering may reduce biomass yields (Schwartz et al., 2010). The summer high-temperature response function is a metric for production limitations that may occur because of stress caused by high temperatures during the growing season. Excessively high temperatures can cause direct damage to crops, and water stress in dryland crops, both of which can lead to reductions in performance. Crop-specific parameters for heat and cold injury are discussed in the Model parameterization section.

Soil pH response. The soil pH response function accounts for production limitations caused by excessively acidic (low pH) or alkaline (high pH) soils. Most plants prosper in the pH range from 5.6 to 7.3, classified as moderately acid to neutral (NRCS, 2003). As soils become more acidic the solubility of most major plant nutrients as well as some micronutrients, such as molybdenum, decrease. Nutrients must be soluble in water to be adsorbed by plant roots. Very acid soils may also release toxic amounts of aluminum, iron, and manganese. Alkaline soils can also decrease plant nutrient solubility, principally phosphorus, boron, copper, iron, manganese, and zinc. Often the largest problem with alkaline soils is their high salt content. Soil pH can be modified by addition of liming agents; this is discussed in greater detail in the Model parameterization section.

Soil salinity response. Highly saline soils increase the osmotic potential of the soil solution, requiring plants to expend more energy to absorb water from the soil. High soil salinity can also limit the uptake of certain nutrients, such as nitrate, manganese, and calcium, leading to nutrient imbalances in the plant (Bano & Fatima, 2009). Very slightly saline soils (2–4 mmhos cm\(^{-1}\)) can restrict the performance of sensitive plants. Slightly saline soils (4–8 mmhos cm\(^{-1}\)) restrict the performance of most plants except the most tolerant. Moderately saline soils (8–16 mmhos cm\(^{-1}\)) depress the performance of even salt tolerant plants. Strongly saline soils (>16 mmhos cm\(^{-1}\)) will not produce acceptable results from

© 2017 The Authors. GCB Bioenergy Published by John Wiley & Sons Ltd., 10, 717–734
any agronomic plant (Munns, 2002; NRCS, 2003). Crop-specific parameters for soil salinity response are discussed in the Model parameterization section.

**Soil drainage response.** Soil drainage deals with water supply issues that affect crop production and management. Oxygen residing in soil pore spaces, necessary for healthy root activity, is limited in poorly drained soils. In addition, poorly drained soils experience limited leaching and flushing of salts left from soil evaporation, which may result in increased salinity. In contrast, water may leach too rapidly through excessively drained soils, leading to premature drought stress and excessive flushing of soil nutrients (NRCS, 1993; Madramootoo et al., 1997; Scherer et al., 2015). Soil drainage response is not a continuous function, but instead is handled categorically, in keeping with NRCS soil drainage categories. Each of seven drainage categories, ranging from very poorly drained to excessively well drained, is assigned a suitability value. Crop-specific parameters for soil drainage are discussed in the Model parameterization section.

**Model parameterization**

County-level grain yield data from winter wheat and maize, commonly grown cool-season and warm-season crops, respectively, were used to initially calibrate and validate PRISM-ELM (Fig. S9). This exercise served two purposes: (1) assess the ability of PRISM-ELM to provide reasonable environmental suitability estimates for two well-known crops that have very different biophysical characteristics; and (2) provide important ‘anchor’ model parameter settings to aid in ranking the settings for biomass crops, which have poorly known environmental tolerances. Data processing and validation details are given in Supporting Information, Data S6.

PRISM-ELM input parameters are defined in Table S2 and crop-specific values given in Tables S3 and S4. The process of setting parameters drew on several sources of information in an iterative fashion: (1) ranking of the species for optimum temperature (Topt), using wheat and maize values as guides; (2) the degree of adherence of resulting PRISM-ELM ESI maps to known spatial patterns of biomass production based on expert review and published literature; (3) and relationships with biomass yield trial data. Taken together, these sources of information allowed the model to be parameterized with greater confidence than using any one source alone. No biomass crop had sufficient data for a purely statistical validation exercise to be performed; therefore, model results were evaluated based on their level of consistency with accumulated knowledge for each crop.

Since winter wheat is a cool-season crop and maize is a warm-season crop, their relationships between temperature and growth differ, especially at lower temperatures. The ranges of air temperature for optimum growth has been reported to be 15–23 °C for wheat (e.g., Steduto et al., 2012), and 25–28 °C for maize (e.g., Schlenker & Roberts, 2009). The PRISM-ELM response curves provided the best fit to the reported spatial patterns of biomass production based on expert review and published literature; (3) and relationships with biomass yield trial data. Taken together, these sources of information allowed the model to be parameterized with greater confidence than using any one source alone. No biomass crop had sufficient data for a purely statistical validation exercise to be performed; therefore, model results were evaluated based on their level of consistency with accumulated knowledge for each crop.

Since winter wheat is a cool-season crop and maize is a warm-season crop, their relationships between temperature and growth differ, especially at lower temperatures. The ranges of air temperature for optimum growth has been reported to be 15–23 °C for wheat (e.g., Steduto et al., 2012), and 25–28 °C for maize (e.g., Schlenker & Roberts, 2009). The PRISM-ELM response curves provided the best fit to the reported spatial patterns of biomass production based on expert review and published literature; (3) and relationships with biomass yield trial data. Taken together, these sources of information allowed the model to be parameterized with greater confidence than using any one source alone. No biomass crop had sufficient data for a purely statistical validation exercise to be performed; therefore, model results were evaluated based on their level of consistency with accumulated knowledge for each crop.

Since winter wheat is a cool-season crop and maize is a warm-season crop, their relationships between temperature and growth differ, especially at lower temperatures. The ranges of air temperature for optimum growth has been reported to be 15–23 °C for wheat (e.g., Steduto et al., 2012), and 25–28 °C for maize (e.g., Schlenker & Roberts, 2009). The PRISM-ELM response curves provided the best fit to the reported spatial patterns of biomass production based on expert review and published literature; (3) and relationships with biomass yield trial data. Taken together, these sources of information allowed the model to be parameterized with greater confidence than using any one source alone. No biomass crop had sufficient data for a purely statistical validation exercise to be performed; therefore, model results were evaluated based on their level of consistency with accumulated knowledge for each crop.
Miscanthus (a C4) consistently gave the best fit to the available yield data, and best matched the expectations of the agronomists conducting the field trials (Fig. 3a). Upland switchgrass had a best-fit temperature response curve that is similar to that of Miscanthus. Temperature response curves for lowland switchgrass, energycane, and biomass sorghum were relatively warm, which also provided reasonable fits to the yield data and matched the expectations of the RFP agronomists (Fig. 3b). Temperature response curves for willow and poplar had similar temperature optima, but poplar was adapted to a wider range of temperature conditions, based on the larger number of available species and varieties tested (Volk et al., 2017).

The Potential Suitability Window, the period within which a crop is expected to be in its most active production phase within potential growing regions of the United States, was set to March–July for winter wheat, a cool-season crop, and April–September for maize, a warm-season crop. This window was set to March–September for CRP, which contains a mixture of cool- and warm-season grasses. Miscanthus, upland and lowland switchgrass, and biomass sorghum, all warm-season crops, were given the same April–September window as maize. Poplar and willow were set to a March–September window. Energycane, which is confined to the extreme southeastern United States, was given a somewhat wider window (March–November) to match that region’s longer growing season. The floating averaging period for peak biomass production was set to 3 months for all crops, which represents the approximate period of time when biomass production is typically most active during the growing season.

$K_c$, the crop coefficient, encompasses canopy characteristics (e.g., height, coverage), stomatal control, and other factors that affect crop evapotranspiration. $K_c$ varies during the life cycle of a crop and water stress situation, ranging from values $<0.5$ early and late in the life cycle, to above 1.0 during mid-season (Allen et al., 1998). Here we used a single $K_c$ value for the entire growing season, which in most cases is a value bracketed by early, mid, and late-season values. Mid-season $K_c$ values have been reported to be 1.0–1.2 for both wheat and maize (e.g., Allen et al., 1998; Steduto & Hsiao, 1998). Single $K_c$ values providing the best fit to the RMA yield data were 1.0 for winter wheat and 0.9 for maize. The few water consumption studies of biomass crops reported large variations in $K_c$, depending on fertilization rate and water stress. Triana et al. (2015) reported Miscanthus $K_c$ values ranging from 0.6 early in the season to 1.6 mid-season. $K_c$ values reported by Guidi et al. (2008) for unfertilized willow and poplar ranged from 0.6 to 1.2 during the growing season. Faced with a lack of $K_c$ estimates for biomass sorghum and switchgrass in the literature, Yimam et al. (2015) approximated them using FAO $K_c$ values (Allen et al., 1998) for sweet sorghum (1.05 mid-season, 1.2 late season) and Sudan grass (0.5 early season, 1.15 mid-season, 1.1 late season), respectively, as surrogates. For biomass crops, PRISM-ELM $K_c$ was set to single values between 1.0 and 1.3, with refinements made which gave the best fit to the yield data. The greatest $K_c$ was 1.3 for energycane, which is slightly lower than mid-season values reported for sugarcane (Allen et al., 1998; Imman-Bamber & McGlone, 2003).

Average rooting depth ($D_{root}$), in a programmatic sense, specifies the fraction of the soil’s available water capacity (AWC) accessible by the crop (see Supporting Information, Eq. 1). Since the root density of most plants gradually decreases with depth, the value of $D_{root}$ is thought of as the depth to which approximately half of the active root biomass extends. This depth varies with crop type, crop development, soil structure and moisture conditions. Most water extraction typically occurs in the upper meter for many crops, declining exponentially with increasing soil depth to maximum rooting depths of up to several meters (Nippert et al., 2011). In PRISM-ELM, $D_{root}$ was set to a single value per crop that ranges from 0.6 to 1.3 m, with refinements made which gave the best fit to the yield data. The stress response factor ($p$), which is the fraction of soil water that a crop can extract from the root zone without suffering water stress, was set to 0.5 for all crops, which is a typical value for most agricultural crops (Allen et al., 1998).

A winter low-temperature response curve was not applicable to maize, which is an annual crop. For winter wheat, the curve defined a two-tailed response: cold tolerance on one tail and vernalization (chilling) requirements needed for grain production on the other. There is a high degree of correlation between winter survival and vernalization requirements. Those varieties that have higher vernalization requirements also tolerate lower winter temperatures with the lowest thresholds being $-15$ to $-23\, ^{\circ}C$, depending on exposure duration (Gusta et al., 1982; Fowler et al., 1996). In the northern United States, the presence of an insulating snowpack can protect wheat plants from damage at ambient temperatures well below the normal range that causes injury, but our modeling system does not have an explicit snow cover component. Because there is typically some snow cover present, the PRISM-ELM response curve does not show significant relative yield reduction until the mean January minimum temperature reaches $-15$ to $-20\, ^{\circ}C$ (Fig. 55a).

Wheat varieties that have relatively low vernalization requirements are now available, reducing the need to lower the relative yield dramatically until the 1981–2010 mean January minimum temperature reaches $10\, ^{\circ}C$ or greater (Fig. 55a). Areas having such warm winter temperatures are confined to the extreme southern tier of states.

For biomass crops, Miscanthus was the only biomass species evaluated that required a two-tailed winter temperature response curve. It has been reported that the dependence of Miscanthus on winter temperature and photoperiod for life stage timing can cause early flowering in southern locations, leading to reduced biomass yields (T. Voigt, personal communication) (Fig. 55a). Upland switchgrass, lowland switchgrass, and energycane were considered to have decreasing levels of cold tolerance, respectively (Fig. 55b). The northern distribution of energycane is confined to the southeastern United States mainly by its susceptibility to freezing injury in winter. Lowland switchgrass is more tolerant of winter cold, but yields are significantly reduced in the northern half of the United States due to cold injury (Casler et al., 2004). Upland switchgrass tolerates lower winter temperatures than lowland switchgrass and thus performs better in northern latitudes (Casler & Vogel, 2014). The willow cultivars in this study, having been selected from breeding programs in southern Ontario and central New York, are also relatively tolerant to winter cold (Volk et al., 2017). The several hybrid genotypes of poplar tested in this
study span a wide range of environments (Volk et al., 2017), some having a high tolerance to low winter temperatures.

While maize has a higher optimum growth temperature than wheat, both suffer from heat-related yield reductions at temperatures which rise above their optimum temperature development thresholds (Wahid et al., 2007) (Fig. 56a, b). Both transitory and constantly high temperatures can lead to heat stress and loss of biomass or yield, and heat stress will affect plant growth throughout its ontogeny (Abrol & Ingram, 1996; Wahid et al., 2007; Hatfield et al., 2008; Luo, 2011). Although much of the wheat crop matures or is harvested before the hottest part of the summer arrives, July mean maximum temperature is still used as a metric for areas that may be at risk of incurring heat damage.

CRP grasses, being a mixture of cool- and warm-season varieties, and Miscanthus, which is best adapted to the Midwest, were assigned relatively similar heat tolerances to wheat. Since the core production area for upland switchgrass extends somewhat further south than that of maize, upland switchgrass was expected to be slightly more heat tolerant than maize. Summer temperatures experienced in the contiguous United States were not expected to be limiting for lowland switchgrass, energycane, and biomass sorghum. Willow was assigned a relatively low heat tolerance, as the cultivars in this study were not selected for lower latitudes in the southern United States (Volk et al., 2017). The broad geographic range of hybrid poplar that results from the mix of a variety of genotypes (Volk et al., 2017) suggests that poplar is relatively tolerant to a wide range of summer temperature conditions.

Response curves to soil pH, salinity, and drainage are shown in Figure 57; parameter values are given in Tables S2 and S3. The practice of liming soils to adjust pH has been well established, and within the last 50 years the intensification of agriculture has driven this practice to the point that most agricultural lands that tend toward natural acidity are being amended to adjust the pH for specific crops (NRCS, 1999; Beegle & Lingenfelter, 2001). The NRCS soil pH data used in this study are representative values of pH for large areas of soil (soil types) in an un-amended (natural/native) state. Given the practice of annual lime applications, and based on the distribution of winter wheat yields, we found it necessary to broaden the pH constraints in the model for all crops (Fig. S7a).

Wheat and maize are classified as moderately tolerant to soil salinity (Ayers & Westcot, 1985; Maas, 1993). Relative crop yield is only slightly affected at salinity levels below 5 mmhos cm$^{-1}$, is reduced by about 50% at 10 mmhos cm$^{-1}$, and falls to zero at about 16 mmhos cm$^{-1}$ (Maas, 1993; Steduto et al., 2012). Tolerances of biomass crops to soil salinity have begun to be studied only recently. Stavriou et al. (2017) found a 50% reduction in biomass yield of Miscanthus × giganteus at 10.65 mmhos cm$^{-1}$, which can be classified as moderate. However, salinity tolerance among Miscanthus genotypes has been found to vary widely (Chen et al., 2017). Wide variations among genotypes have also been reported for grain sorghum (Hassanein & Azab, 1993), switchgrass (Hu et al., 2015), and energycane (Fageria et al., 2013). Given that the PRISM-ELM biomass resource maps are intended to reflect a combination of what often are many ‘best local varieties’, the PRISM-ELM salinity tolerance curves for biomass crops were set to moderate values (Fig. S7b).

It has long been a practice to drain water from low lying lands for agricultural purposes. In North America, this process accelerated with the passage of the swamplands acts of 1849, 1850, and 1860. This practice of draining lands has evolved greatly from the 1800s with added technology, plastic pipe, and GIS planning systems contributing to modernize the practice today (Pavelis, 1987). The USDA provides estimates of drained cropland by county with ranges from 0 to 100%, with large portions of Ohio, Indiana, Illinois, and Iowa having >25% of the land drained (Jaynes & James, 2007). The level of drainage has affected the native soil productivity and increased production, necessitating assigning relatively high soil drainage suitability values to even poorly drained NRCS soil categories (Fig. S7c). A comparison of maps of NRCS soil drainage category and RMA county-level average yields of winter wheat in Ohio illustrates the benefits of soil drainage in the Midwest, where yields are much higher than would be predicted assuming un-amended soil drainage conditions (Fig. S8).

### Transforming ESI to yield potential

PRISM-ELM grids of ESI values for each biomass species were transformed into yield potential grids through linear least-squares regression functions between average reported yield and ESI. Each function was forced through a zero $y$-intercept to avoid cases where a positive yield is predicted when the ESI value is zero, but in all cases, the unforced regression was very close to a zero intercept.

### Results

#### Environmental suitability mapping for biomass crops

ESI maps for each biomass crop are shown in Figure S12. An advantage to expressing suitability as a dimensionless number is that crop-to-crop variations in biomass yield are controlled for, leaving environmental limitations as the dominant predictor of the spatial patterns. Maps of the limiting factor, that is, the lowest suitability index of the PRISM-ELM submodels, are shown in Figure 4. All ESI maps show a more or less consistent dividing line in the middle of the country, representing the transition from wetter climates to the east and drier climates to the west. In general, the cooler the temperature optimum of a crop, the more likely it is to have an ESI maximum on the West Coast, especially in the Pacific Northwest, where precipitation during the cooler spring months is sufficient to support crop production before summer drought sets in. The importance of a favorable water balance is illustrated in the limiting factor maps, where the water balance suitability index is the greatest limitation over large areas.

Summer heat limits the southern distribution of crops adapted to cooler temperatures, such as CRP,
Fig. 4  Parameter-elevation Regressions on Independent Slopes Model Environmental Limitation Model (PRISM-ELM) limiting factor distributions for herbaceous and woody biomass crops.
Miscanthus, and willow. Winter cold limits the northward distribution of warmer adapted perennial crops, most notably lowland switchgrass, and energycane, which is restricted to the extreme southeastern United States because of damage due to freezing temperatures. Soil pH and drainage play relatively minor roles as limiting factors in these simulations, because of our assumption of widespread limed and drained soils. Soil salinity is the limiting factor primarily along coastlines and in parts of the arid West.

Potential biomass production mapping

Scatterplots and linear regression equations used to convert PRISM-ELM ESI grids to potential biomass production are shown in Figure 5. $R^2$ values ranged from 0.55 for upland switchgrass to 0.88 for CRP and biomass sorghum. Mean absolute errors (MAEs) ranged from 0.24 Mg ha$^{-1}$ yr$^{-1}$ for CRP to 2.96 Mg ha$^{-1}$ yr$^{-1}$ for Miscanthus. On a percentage basis, MAEs ranged from 7.7% for willow to 27.5% for lowland switchgrass. In Figure 5, open circles represent locations of RFP field trials, and closed circles represent other trials. For most crops, RFP trials supplied most, if not all, of the data points used in the regressions. However, non-RFP trials played a major role in the Miscanthus and upland switchgrass regressions, and were the only source of suitable yield data for lowland switchgrass. The range of environmental conditions represented by the yield data can be seen by viewing the distribution of data points along the $x$-axis. A number of biomass crops, including Miscanthus, upland switchgrass, sorghum, and poplar, did not have yield trial data for areas with ESI values of $<40$, suggesting that the full range of environmental conditions was not well represented by the field trials.

Estimated average annual biomass yield potential maps, derived from regression functions relating PRISM-ELM ESI to reported yield, are shown in Figure 6. CRP grasses have estimated yields of up to 3–6 Mg ha$^{-1}$ yr$^{-1}$ in the eastern United States, with lower values in the West. Maximum yields for Miscanthus were estimated to exceed 22 Mg ha$^{-1}$ yr$^{-1}$ in the Midwest, decreasing to $<10$ Mg ha$^{-1}$ yr$^{-1}$ in the extreme southern United States. Yields reach a secondary maximum of 14–18 Mg ha$^{-1}$ yr$^{-1}$ in the wetter areas of the Pacific Northwest. Upland and lowland switchgrass have very different yield distributions; upland switchgrass yields reach a maximum of 10–14 Mg in the Midwest, and maintain a fairly wide north-to-south swath of 6–10 Mg ha$^{-1}$ yr$^{-1}$ across the eastern United States, and also in the wetter areas of the Pacific Northwest. In contrast, lowland switchgrass yields exceed 18 Mg ha$^{-1}$ yr$^{-1}$ along the Gulf Coast and into the lower Mississippi Valley and decrease northward to $<3$ Mg ha$^{-1}$ yr$^{-1}$ along the northern tier of states. Lowland switchgrass production in the West is generally $<6$ Mg ha$^{-1}$ yr$^{-1}$. Biomass sorghum yields exceed 22 Mg ha$^{-1}$ yr$^{-1}$ across the southern and central portions of the eastern United States decreasing to 10–14 Mg ha$^{-1}$ yr$^{-1}$ in the upper Midwest. Yields are generally $<10$ Mg ha$^{-1}$ yr$^{-1}$ in the West. Energycane is confined to the extreme southeastern United States, with yields estimated to exceed 18 Mg ha$^{-1}$ yr$^{-1}$ in Florida and along the Gulf Coast.

Yield estimates for willow show a maximum of 14–20 Mg ha$^{-1}$ yr$^{-1}$ in the Midwest, and extending into central New England and southward into the southern Appalachians. Yields are estimated to be low in the southern states and throughout most of the West. Poplar yields also reach a maximum in the Midwest, with an extensive area of $>10$ Mg ha$^{-1}$ yr$^{-1}$ across the eastern United States and in wetter areas of the Pacific Northwest.

Discussion

Most previous work to map the production potential of biomass crops in the United States has focused on switchgrass and to a lesser extent Miscanthus and willow, both of which have been more extensively researched in Europe (e.g., Hastings et al., 2009; Larsen et al., 2016; Mola-Yudego et al., 2016). Looking at modeled biomass yield maps from the literature for the same crop, we find a wide variation in yield estimates which appear to stem from differences among models. This illustrates one of the inherent difficulties in comparing biomass estimates between crops that do not use the same modeling approach. Thomson et al. (2009) used the EPIC mechanistic simulation model to estimate 30-year average switchgrass yield on a large watershed scale. They simulated lowland ecotypes south of 41°N and upland ecotypes to the north and then combined the two simulations into one map. The spatial patterns of production are roughly similar to those of the PRISM-ELM lowland and upland maps (if combined), with maxima in the Midwest and South, but the watershed-scale simulation does not resolve topographic variations well. Using the ALMANAC mechanistic simulation model, Behrman et al. (2013) estimated local biomass potential in the eastern United States at 0.25° resolution, with lowland and upland ecotypes combined. The resulting map shows biomass maxima ($>18$ Mg ha$^{-1}$ yr$^{-1}$) along the Gulf Coast, and high production extending up into the Midwest. The simulation shows a tongue of lower biomass potential in eastern Oklahoma and Arkansas (6–10 Mg ha$^{-1}$ yr$^{-1}$), but the PRISM-ELM lowland switchgrass map has higher...
Fig. 5 Relationships between Parameter-elevation Regressions on Independent Slopes Model Environmental Limitation Model (PRISM-ELM) ESI and reported average biomass yields. Open circles are yields from Sun Grant Regional Feedstock Partnership (RFP) trials, and solid circles are average yields from other trials. Linear regression functions were developed using all data shown and forced through a zero y-intercept. See Table S1 for a listing of the yield data used.

© 2017 The Authors. GCB Bioenergy Published by John Wiley & Sons Ltd., 10, 717–734
Fig. 6  Estimated average annual biomass yield potential maps, derived from regression functions relating PRISM-ELM ESI to reported yield (see Fig. 5).
estimated yields at 14–18 Mg ha\(^{-1}\) yr\(^{-1}\). In addition, biomass potential in the extreme upper Midwest (e.g., northern Minnesota) decreases to 2–6 Mg ha\(^{-1}\) yr\(^{-1}\), while PRISM-ELM maintains 6–10 Mg ha\(^{-1}\) yr\(^{-1}\), contributed by upland switchgrass.

Using a statistical model, Jager et al. (2010) estimated annual yield potential for upland and lowland switchgrass at 4-km resolution. Areas not within the range of the explanatory data used in the statistical model were not simulated. Surprisingly high yields for both ecotypes extended into western Texas and southern New Mexico and Arizona, which are unlikely locations for favorable production. The issue of unusually high yields extending into the Southwest is also seen in the switchgrass map produced with a statistical model developed by Wullscherleger et al. (2010); relatively high yields were estimated throughout much of Colorado, northwestern Texas and New Mexico, where precipitation is generally inadequate and moisture deficits are likely to restrict production. Tulbure et al. (2012) used a statistical model to map the yield potential for upland and lowland switchgrass in the eastern United States at 1-km resolution. The general patterns of the upland map are similar to those of PRISM-ELM, but the lowland switchgrass map is very different than all others reviewed. Yield maxima are located in the Midwest and Appalachians, with anomalously high values along the edge of the statistical range of the data in the southeastern United States.

Song et al. (2015) used the mechanistic Integrated Science Assessment Model (ISAM) to estimate biomass yield potential in the eastern United States for Miscanthus, as well as an upland switchgrass cultivar (Cave-in-Rock) and a lowland switchgrass cultivar (Alamo) at coarse (0.5°) resolution. Yield patterns are similar to those of PRISM-ELM in the major crop production areas of the United States, but production declines to zero north of the boundaries of USDA plant hardiness zones: zone 4 for Miscanthus, zone 3 for upland switchgrass, and zone 5 for lowland switchgrass. The USDA plant hardiness statistic is defined as the mean annual extreme minimum temperature and thus is a measure of the potential for winter cold (Daly et al., 2012). PRISM-ELM shows yield reductions near these limits as well, but the decreases are more gradual. Miguez et al. (2012) used BioCro to estimate long-term biomass productivity of Miscanthus and switchgrass at 32-km resolution. BioCro is a process-based model for plant growth which simulates plant biochemistry and biophysics. Estimated yields of Miscanthus have two main maxima, one in the Midwest with values higher than PRISM-ELM (30–35 vs. >22 Mg ha\(^{-1}\) yr\(^{-1}\)), and another in the far south, which is considerably higher than PRISM-ELM (30–35 vs. 10–14 Mg ha\(^{-1}\) yr\(^{-1}\)). A small area of extremely high yield (40–45 Mg ha\(^{-1}\) yr\(^{-1}\)) is located along the southern Washington coast; this area is also a maximum in PRISM-ELM, but yield values are much lower (14–18 Mg ha\(^{-1}\) yr\(^{-1}\)). The southern portion of the switchgrass map is similar to PRISM-ELM lowland switchgrass map, with maximum production extending up the lower Mississippi Valley and along the southeast Atlantic coastline. In the north, yields are somewhat higher than those shown in the PRISM-ELM upland switchgrass map, but the regional patterns are similar.

The PRISM-ELM potential production map for willow is roughly similar to that produced with BioCro by Wang et al. (2015). Both maps exhibit production maxima in the Midwest and northeast, and a limited production maximum in the Pacific Northwest. The PRISM-ELM map is more conservative in estimating less production along the Gulf Coast, which is likely on the edge of the distributions of willow cultivars tested in the RFP field trials.

While there are clearly large differences among the potential yield maps reviewed, the PRISM-ELM maps exhibit regional patterns that are more similar to those produced with mechanistic models than with statistical models. The reason for this may be PRISM-ELM’s treatment of the water balance, which is the dominant control on modeled production potential over much of the country (see Fig. 4). Statistical relationships between precipitation and yield can be misleading unless the timing of water demand, which is controlled largely by temperature, is accounted for simultaneously. The 800-m resolution PRISM-ELM maps show more detail in complex terrain than those produced with mechanistic models, mainly because PRISM-ELM is driven with high-resolution, topographically sensitive PRISM climate data. Detailed input data needed to drive mechanistic models are not always available at fine scales, and the models themselves are computationally intensive.

Not surprisingly, the areas of greatest disagreement among models are located along the edges of a crop’s distribution, where field trials are largely absent. An example of this is a lack of switchgrass field trials in the relatively arid southwestern United States, where statistical methods overpredicted yield potential. In fact, nearly all other biomass crops lack yield data in areas estimated by PRISM-ELM to have low environmental suitability. That said, it can be challenging to conduct trials where a crop is likely to fail or do poorly, and these trials may not be financially feasible to perform.

A significant source of uncertainty in our modeling system is difficulties in characterizing on-farm soil conditions using NRCS data. The NRCS soils data used in this study are representative values for large areas of soil (soil types) in an un-amended (natural/native)
state. This makes an accurate accounting of adjustments such as drainage systems and liming applications problematic. Tile drainage systems and other measures to ameliorate poor soil drainage were not reflected in the NRCS soils data, and detailed information on the locations of drainage systems was unavailable. Our remedy was to make the PRISM-ELM soil drainage response function less restrictive during model validation with wheat and maize, and this parameterization was carried over to simulations for all biomass crops. As a result, yields may have been overpredicted in poorly drained areas that do not have drainage systems in place. Most agricultural lands that tend toward natural acidity have pH adjustments made through lime applications. These adjustments were also not reflected in the NRCS data, and the spatial distribution of liming practices was poorly known. In response, pH constraints in PRISM-ELM were made less restrictive during simulations for wheat and maize, and again carried over into all biomass simulations. Unfortunately, the key question of whether biomass crops will be grown on amended fields or relegated to marginal, un-amended lands requires economic considerations that are outside the scope of this study. However, one approach may be to bracket the range of possible outcomes by producing several biomass potential maps based on differing assumptions of soil improvements such as drainage and liming.

PRISM-ELM modeling has so far used soils data derived from the NRCS U.S. General Soils Maps; representative values were extracted from this map and averaged to an 800-m grid cell. Recently completed, higher resolution soils data from the NRCS can be used to improve soil characterizations, but the level of spatial detail may still be insufficient to capture conditions at the field trials, which are often conducted on small plots. Options to improve the spatial accuracy of soil representations are to apply PRISM-ELM to native NRCS soil polygons rather than summarizing conditions over arbitrarily defined grid cells, or re-cast the model to run in ‘field’ mode for specific trial locations, and have soil characteristics specified based on data collected by agronomists conducting the trials. Including a metric in PRISM-ELM for characterizing soil fertility and productivity should also be considered. Soil organic matter is a good choice, as it plays key roles in soil health by increasing carbon content, acting as a buffer for soil acidification, and contributing to soil structure and water holding capacity.

The potential biomass yield maps produced by PRISM-ELM represent estimates of annual average yields based on a 30-year (1981–2010) average climate, while the RFP field trials used here ran for 3–7 years from the late 2000s to the early 2010s. It is likely that the long-term climate data did not fully represent weather conditions experienced during these relatively short trials. For example, severe drought across the Midwest in 2012 reduced yields at several trial sites. The logical next step is to apply the model on a year-by-year basis to obtain a distribution of potential annual yields that can be used to develop risk assessments. This type of analysis would help answer questions about the long-term stability of expected biomass yields over time, given the historical variability in weather conditions, and potentially improve relationships between suitability estimates and observed yields. However, it may be difficult to evaluate model performance because of the limited duration of the biomass yield trials, and the confounding, noneconomic factors that affect yield variability in crops with longer histories such as wheat and maize.

The work described here represents a first, coordinated look at the potential long-term yield distribution of several important biomass crops in the United States. Use of a consistent modeling framework avoids the danger of confusing differences in model structure with biological differences. The resulting maps are intercomparable, allowing crop selection decisions to be made with increased confidence. These maps represent a patchwork of the best local varieties that would have been available to a producer at the time the field trials were conducted. A list of the field trial locations used in the modeling effort is provided in Table S1, and details on all of the RFP field trials are available from companion articles in this issue (Lee et al., 2017; Volk et al., 2017). As can be seen in Figure 6, the maps are based on a very small number of yield trial data points. Therefore, caution is advised when using these maps in regions that are distant from trial locations.

Providing spatial yield estimates for such a wide range of biomass crops required a simplified modeling framework that was generalizable over many species with different life cycles and environmental tolerances. Therefore, the focus was on a modified biogeographical approach to modeling climatic and soil constraints on biomass production for any crop, rather than a detailed accounting of the particular phenology and physiological features of a given species or genotype. The poor state of knowledge regarding the environmental tolerances of most biomass crops led to a model parameterization strategy that took advantage of the synergy realized by combining information from crops with long production histories, coordinated field trials, close interaction with expert agronomists, and spatial modeling. As such, this modeling and parameterization framework can be expanded and updated to include other biomass crops and varieties.

The potential production maps presented in this article are accessible via the USDOE Knowledge Discovery Framework (KDF) (https://www.bioenergykdf.net).
Acknowledgements

The authors would like to thank the many individuals who were part of the RFP, especially those who provided data for and input about the national maps. The RFP included personnel from numerous universities, federal laboratories, federal agencies, and private industry as well as producers who allowed field trials to be conducted on their land. Thanks also go to three anonymous reviewers who provided important guidance that helped the authors find the right ‘voice’ for this work.

This research was supported by funding from the North Central Regional Sun Grant Center at South Dakota State University through a grant provided by the US Department of Energy Bioenergy Technologies Office under award number DE-FC36-05GO85041 and by a cooperative agreement between the US Department of Agriculture Risk Management Agency and Oregon State University under PO 450073924.

This manuscript has been co-authored (Eaton) by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the US Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

References

Abrol YP, Ingram KT (1996) Effects of higher day and night temperatures on growth and yields of some crop plants. In: Global Climate Change and Agricultural Production: Direct and Indirect Effects of Changing Hydrological, Pedological and Plant Physiological Processes (eds Bazzaz F, Sombroek W). Food and Agriculture Organization, Rome, Italy. Available at http://www.fao.org/docrep/w5183e/w5183e08.htm (accessed 4 January 2018).

Allen RG, Pereira LS, Raes D, Smith M (1998) Guidelines for Computing Crop Water Requirements. FAO Irrigation and Drainage Paper 56, FAO and Agriculture Organization of the United Nations, Rome, Italy.

Ayers RS, Westcot DW (1985) Water Quality for Agriculture. FAO Irrigation and Drainage Paper 29. (Rev. 1). Food and Agriculture Organization of the United Nations, Rome, Italy.

Balkovic J, van der Velde M, Schmid E et al. (2013) Pan-European crop modelling with EPIC: implementation, up-scaling and regional crop yield validation. Agricultural Systems, 120, 61–79.

Bano A, Fatima M (2009) Salt tolerance in Zea mays (L) following inoculation with Rhizobium and Pseudomonas. Biology and Fertility of Soils, 45, 405–413.

Beck EH, Heim R, Hansen J (2004) Plant resistance to cold stress: mechanisms and environmental signals triggering frost hardening and dehardening. Journal of Biosciences, 29, 449–459.

Beegle DB, Lingenfelder DD (2001) Soil Acidity and Lime. Agronomy Facts, 3. Pennsylvania State University Extension, State College, PA.

Behrman KD, Kiniry JR, Winchell M, Juenger TE, Keitt TH (2013) Spatial forecasting of switchgrass productivity under current and future climate change scenarios. Ecological Applications, 23, 73–85.

Brison N, Launay M, Mary B, Beaudoin N (2008) Conceptual Basis, Formalizations, and Parameterization of the STICS Crop Model. French National Institute for Agronomic Research, Versailles, France.

Brown RA, Rosenberg NJ, Hays CJ, Easterling WE, Mearns LO (2000) Potential production and environmental effects of switchgrass and traditional crops under current and greenhouse-altered climate in the central United States: a simulation study. Agriculture, Ecosystems and Environment, 78, 31–47.

Casler MD, Vogel KP (2014) Selection for biomass yield in upland, lowland, and hybrid switchgrass. Crop Science, 54, 626–636.
Supporting Information

Additional Supporting Information may be found online in the supporting information tab for this article:

Data S1. Basic environmental constraints on production:
Figure S1. Conterminous US 1981–2010 (a) mean annual temperature, (b) mean January minimum temperature, and (c) mean July maximum temperature.
Figure S2. Conterminous US soil (a) available water capacity (b) pH, (c) salinity, and (d) drainage.

Data S2. Biomass yield trial data:
Table S1. Biomass yield field trials used in the parameterization of PRISM-ELM and the transformation of PRISM-ELM ESI into estimated annual biomass production.

Data S3. PRISM-ELM model formulation:
Figure S3. Illustration of method used to obtain final suitability ($S_w$) of winter wheat in the Willamette Valley, Oregon from the PRISM-ELM water balance model.

Data S4. Examples of PRISM-ELM operation:
Figure S4. Comparison of PRISM-ELM water balance operation for wheat (cool-season crop) and maize (warm-season crop) in western Oregon, characterized by dry, mild summers, and southeastern Indiana, where summers are warm and moist.

Data S5. PRISM-ELM input parameters:
Table S2. Descriptions of PRISM-ELM input parameters.
Table S3. PRISM-ELM input parameters for herbaceous biomass crops.
Table S4. PRISM-ELM input parameters for woody biomass crops.
Figure S5. PRISM-ELM January minimum temperature response curves (a metric for winter cold injury).
Figure S6. PRISM-ELM July maximum temperature response curves (a metric for heat injury).
Figure S7. PRISM-ELM relative yield response curves to soil pH, salinity; and drainage, for all biomass crops.
Figure S8. Comparison of (a) USDA NRCS soil drainage class and (b) RMA county-level winter wheat yields for northern Ohio.

Data S6. Winter wheat and maize model validation:
Figure S9. USDA RMA county-level grain yield data for non-irrigated winter wheat and maize.
Figure S10. PRISM-ELM ESI maps for (a) winter wheat and (b) maize.
Figure S11. Scatterplots and least-squares linear regressions forced through zero between county-level PRISM-ELM ESI and RMA 2000–2015 average annual reported grain yields for (a) winter wheat and (b) maize.

Table S5. PRISM-ELM performance statistics for RMA winter wheat and maize yield.

Data S7. Environmental suitability mapping for biomass crops:
Figure S12. PRISM-ELM ESI distributions for herbaceous and woody biomass crops.