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

CHRISTOPHER DALY[1] (iD), MICHAEL D. HALBLEIB[1], DAVID B. HANNAWAY[2] and LAURENCE M. EATON[3]

[1]PRISM Climate Group, Northwest Alliance for Computational Science and Engineering, 2000 Kelley Engineering Center, Oregon State University, Corvallis, OR, USA, [2]Department of Crop and Soil Science, Oregon State University, 125 Crop Science Building, Corvallis, OR, USA, [3]Bioenergy 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.

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.

Correspondence: Christopher Daly, tel. +1 541 737 2531, fax +1 541 737 6609, e-mail: Chris.Daly@oregonstate.edu

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 (Evans *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.

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 spontaneam* 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 conterminous 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 grown 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 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).
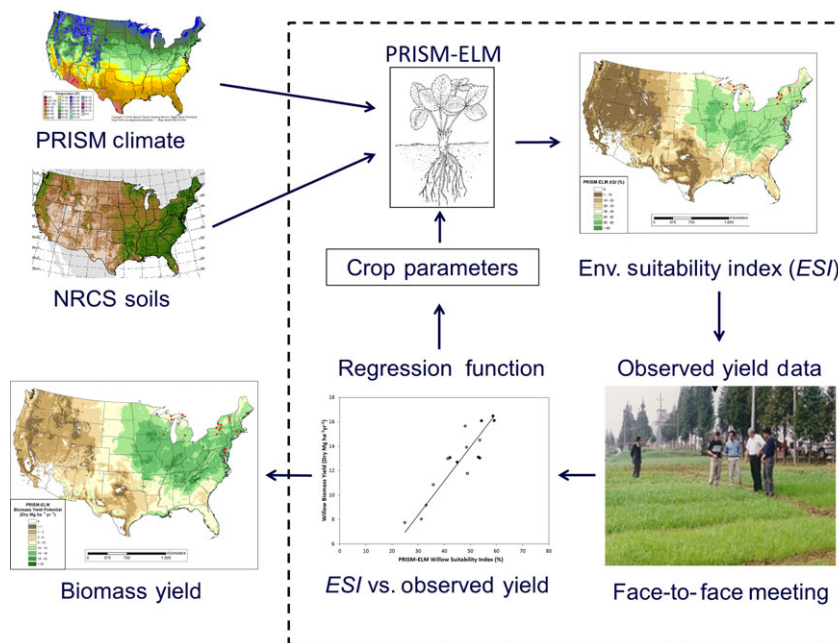
Fig. 1 Schematic of the Parameter-elevation Regressions on Independent Slopes Model Environmental Limitation Model (PRISM-ELM) workflow for mapping bioenergy resources. Inputs to PRISM-ELM were gridded climate and soils data, and a preliminary parameter file for the crop being modeled. An initial Environmental Suitability Index (ESI) grid was produced, and during a face-to-face meeting with agronomists, the ESI grid was evaluated against observed yield data to help understand data outliers and adjust model parameters. Once an agreement was reached on model parameters and yield data to be used, a final regression function was developed and applied to the ESI grid to produce a potential biomass yield grid.

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:

$$ESI = \min(S_w, S_c, S_h, S_p, S_s, S_d),$$

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 drainage. 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
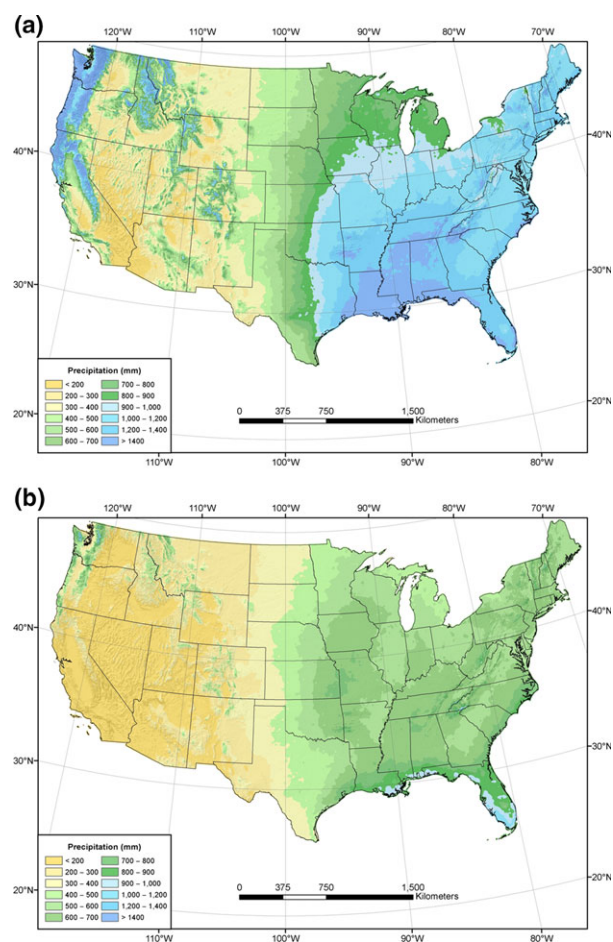
Fig. 2 Conterminous US 1981–2010 (a) mean annual precipitation and (b) mean April–September precipitation. Data source: PRISM Climate Group (http://prism.oregonstate/edu).

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 ($S = K_s T_r$). 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

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 (T_{opt}), 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 providing the best fit to the reported spatial patterns of production exhibited relatively low optimum temperature values (OptT = 18 °C for wheat and 21.5 °C for maize) (Table S3) (Fig. 3a, b). PRISM-ELM used 1981–2010 mean semi-monthly temperature to drive the response curves;

the photosynthetically active (daytime) temperature would be several degrees higher than this mean. Research has also suggested that diurnal mean temperature may be a better predictor of plant response than daytime highs alone (Peet & Willits, 1998). The difference between the two temperature response curves is primarily at lower temperatures, where maize growth is severely limited at temperatures below 10 °C (Fig. 3b).

Temperature response curves for biomass crops were ranked in comparison with those of wheat and maize, and with each other. The ranking of the species' temperature optimum, from cool to warm, was: winter wheat, CRP grasses, Miscanthus, upland switchgrass, maize, lowland switchgrass, biomass sorghum, and energycane. CRP grasses, which are a mixture of C3 and C4 species, had the next coolest temperature optimum after wheat. A relatively cool optimum temperature for
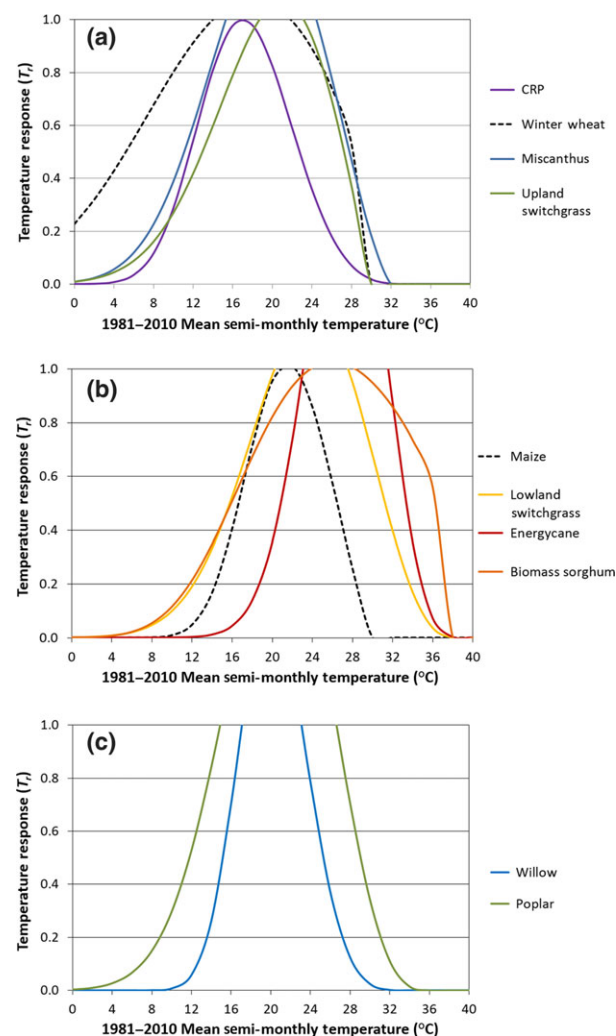

Fig. 3 Parameter-elevation Regressions on Independent Slopes Model Environmental Limitation Model (PRISM-ELM) growing season temperature response curves for (a, b) herbaceous and (c) woody biomass crops. Parameter values are defined in Table S2 and given in Tables S3 and S4. Response curve values exceeding 1.0 are set to 1.0.

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 optimums, 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; Inman-Bamber & McGlinchey, 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 °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 °C (Fig. S5a). 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 °C or greater (Fig. S5a). 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. S5a). Upland switchgrass, lowland switchgrass, and energycane were considered to have decreasing levels of cold tolerance, respectively (Fig. S5b). 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. S6a, 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 ontology (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 conterminous 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 S7; 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. Stavridou *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,
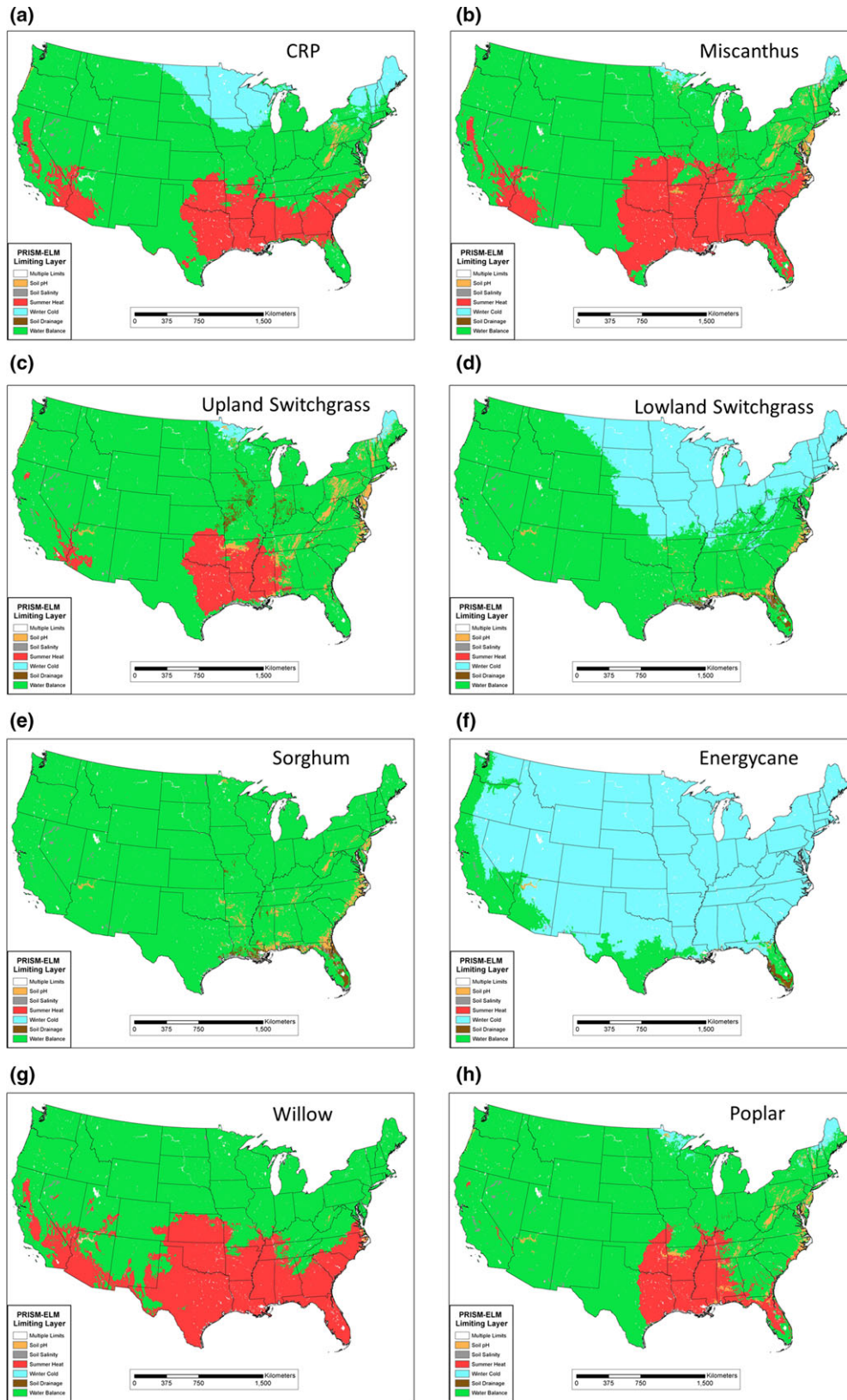
Fig. 4 Parameter-elevation Regressions on Independent Slopes Model Environmental Limitation Model (PRISM-ELM) limiting factor distributions for herbaceous and woody biomass crops.

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
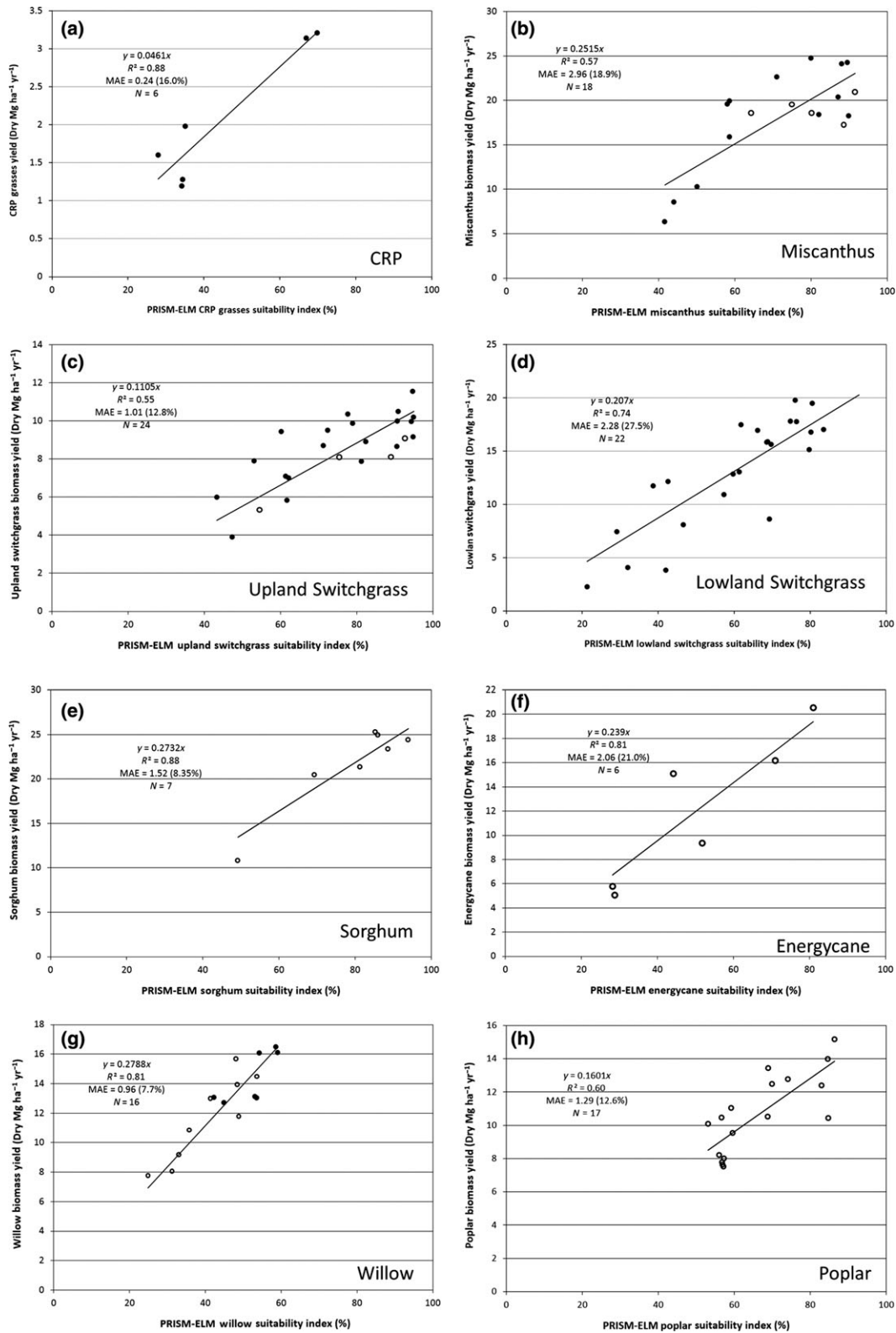
Fig. 5 Relationships between Parameter-elevation Regressions on Independent Slopes Model Environmental Limitation Model (PRISM-ELM) ESI and reported average biomass yields. Open circles are yields from Sun Grant Regional Feedstock Partnership (RFP) trials, and solid circles are average yields from other 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.
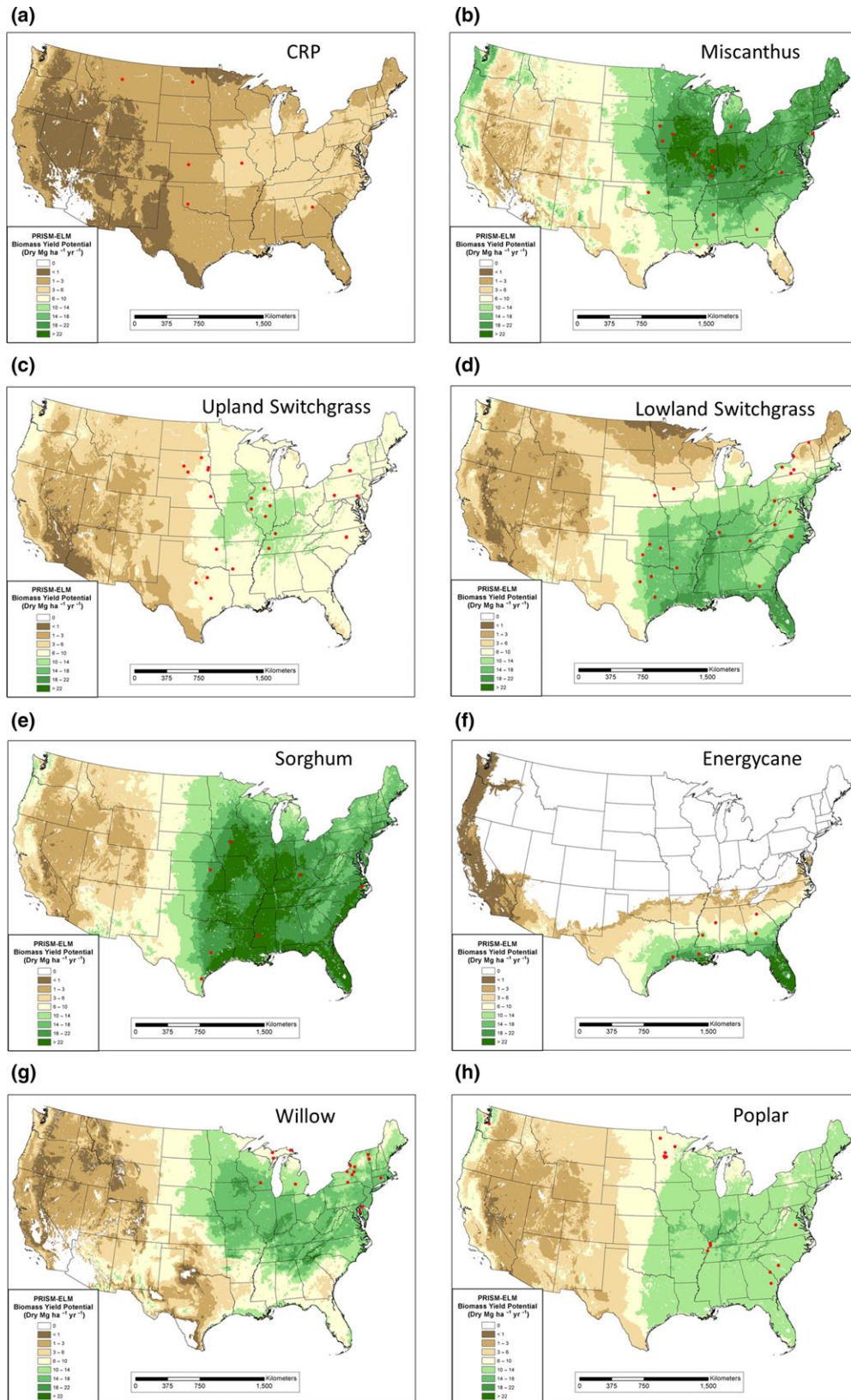
Fig. 6 Estimated average annual biomass yield potential maps, derived from regression functions relating PRISM-ELM ESI to reported yield (see Fig. 5).

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 Wullschleger *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 to 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 amendments 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

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) *Crop evapotranspiration – Guidelines for computing crop water requirements*. FAO Irrigation and Drainage Paper 56, Food and Agriculture Organization of the United Nations, Rome, Italy.

Ayers RS, Westcot DW (1985) *Water quality for agriculture*. FAO Irrigation and Drainage Paper No. 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**, 61–75.

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, Lingenfelter 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.

Brisson N, Launay M, Mary B, Beaudoin N (2008) *Conceptual Basis, Formalisations, 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.

Casler MD, Vogel KP, Taliaferro CM, Wynia RL (2004) Latitudinal adaptation of switchgrass populations. *Crop Science*, **44**, 293–303.

Castillo CP, Lavall C, Baranzelli C, Mubareka S (2015) Modelling the spatial allocation of second-generation feedstock (lignocellulosic crops) in Europe. *International Journal of Geographical Information Science*, **29**, 1807–1825.

Chen C-L, van der Schoot H, Dehghan S et al. (2017) Genetic diversity of salt tolerance in *Miscanthus*. *Frontiers in Plant Science*, **8**, 187.

Daly C, Neilson RP, Phillips DJ (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology*, **33**, 140–158.

Daly C, Halbleib M, Smith JI et al. (2008) Physiographically-sensitive mapping of temperature and precipitation across the conterminous United States. *International Journal of Climatology*, **28**, 2031–2064.

Daly C, Widrlechner MP, Halbleib MD, Smith JI, Gibson WP (2012) Development of a new USDA Plant Hardiness Zone Map for the United States. *Journal of Applied Meteorology and Climatology*, **51**, 242–264.

Dennis FG (1984) Flowering. In: *Physiological Basis of Crop Growth and Development* (ed. Tesar MB), pp. 237–264, American Society of Agronomy, Madison, WI.

Evans JM, Fletcher RJ, Alavalapati J (2010) Using species distribution models to identify suitable areas for biofuel feedstock production. *Global Change Biology: Bioenergy*, **2**, 63–78.

Fageria NK, Moreira A, Moraes LAC, Hale AL, Viator RP (2013) Sugarcane and energycane. In: *Biofuel Crops: Production, Physiology and Genetics* (ed. Singh BP), 151–171. CAB International, Wallingford, UK.

Fowler DB, Limin AE, Wang SY, Ward RW (1996) Relationship between low-temperature tolerance and vernalization response in wheat and rye. *Canadian Journal of Plant Science*, **76**, 37–42.

Guidi W, Piccioni E, Bonari E (2008) Evapotranspiration and crop coefficient of poplar and willow short-rotation coppice used as vegetation filter. *Bioresource Technology*, **99**, 4832–4840.

Gusta LV, Fowler DB, Tyler NJ (1982) Factors influencing hardening and survival in winter wheat. In: *Plant Cold Hardiness and Freezing Stress* (eds Li PH, Sakai A), pp. 23–40, Vol. II. Academic Press, New York, NY.

Hannaway DB, Daly C, Cao W et al. (2005) Forage species suitability mapping for China using topographic, climatic and soils spatial data and quantitative plant tolerances. *Scientia Agricultura Sinica*, **4**, 660–667.

Hargreaves GH, Samani ZA (1985) Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture*, **1**, 96–99.

Hassanein AHM, Azab AM (1993) Salt tolerance of grain sorghum. In: *Towards the Rational Use of High Salinity Tolerant Plants. Tasks for Vegetation Science* (eds Lieth H, Al Masoom AA), vol. 28, pp. 153–156. Springer, Dordrecht.

Hastings A, Clifton-Brown J, Wattenbach M, Mitchell CP, Smith P (2009) The development of MISCANFOR, a new *Miscanthus* crop growth model: towards more robust yield predictions under different climatic and soil conditions. *Global Change Biology: Bioenergy*, **1**, 154–170.

Hatfield J, Boote K, Fay P et al. (2008) Agriculture. In: *The Effects of Climate Change on Agriculture, Land Resources, Water Resources, and Biodiversity in the United States. A Report by the U.S. Climate Change Science Program and the Subcommittee on Global Change Research*. Washington, DC, USA.

Hu G, Liu Y, Zhang X et al. (2015) Physiological evaluation of alkali-salt tolerance of thirty switchgrass (*Panicum virgatum*) lines. *PLoS One*, **10**, e0125305.

Inman-Bamber NG, McGlinchey MG (2003) Crop coefficients and water-use estimates for sugarcane based on long-term Bowen ratio energy balance measurements. *Field Crops Research*, **83**, 125–138.

Jager HI, Baskaran LM, Brandt CC, Davis EB, Gunderson CA, Wullschleger SD (2010) Empirical geographic modeling of switchgrass yields in the United States. *Global Change Biology: Bioenergy*, **2**, 248–257.

Jaynes DB, James DE (2007) The extent of farm drainage in the United States. In: *Final Program and Abstracts, Soil and Water Conservation Society International Conference*, 21–25 July 2007, Tampa, FL. Available at: https://www.ars.usda.gov/SP2UserFiles/Place/50301500/TheExtentofFarmDrainageintheUnitedStates.pdf (accessed 4 January 2018).

Kiniry JR, Schmer MR, Vogel KP, Mitchell RB (2008) Switchgrass biomass simulation at diverse sites in the northern great plains of the U.S. *Bioenergy. Research*, **1**, 259–264.

Larsen S, Jaiswal D, Bentsen NS, Wang D, Long SP (2016) Comparing predicted yield and yield stability of willow and *Miscanthus* across Denmark. *GCB Bioenergy*, **8**, 1061–1070.

Lee DK, Baldwin B, Anderson B et al. (2017) Herbaceous energy crops: management, yield and expectation by region. *Global Change Biology Bioenergy*, (accepted).

Levitt J (1980) *Responses of Plants to Environmental Stresses. Vol. 1. Chilling, Freezing, and High Temperature Stresses*. Academic Press, New York.

Lobell DB, Cassman KG, Field CB (2009) Field crop yield gaps: their importance, magnitudes, and causes. *Annual Review of Environment and Resources*, **3**, 179–204.

Luo Q (2011) Temperature thresholds and crop production: a review. *Climatic Change*, **109**, 583–598.

Maas EV (1993) Testing crops for salinity tolerance. *Proceeds of Workshop on Adaptation of Plants to Soil Stresses* (eds Maranville JW, BaAZigar BV, Duncan RR, Yohe JM), pp. 234–247. Institute for Sorghum and Millet, Pub. No. 94-2, University of Nebraska, Lincoln, NE, August l–4 1993.

Madramootoo CA, Johnston WR, Lyman SW (eds) (1997) *Management of agricultural drainage water quality*. FAO Water Report 13, Food and Agriculture Organization of the United Nations, Rome, Italy. Available at: http://www.fao.org/docrep/w7224e/w7224e00.htm#Contents (accessed 4 January 2018).

Miguez FE, Maughan M, Bollero GA, Long SP (2012) Modeling spatial and dynamic variation in growth, yield and yield stability of the bioenergy crops *Miscanthus × giganteus* and *Panicum virgatum* across the conterminous USA. *Global Change Biology Bioenergy*, **4**, 509–520.

Mola-Yudego B, Rahlf J, Astrup R, Dimitriou I (2016) Spatial yield estimates of fast-growing willow plantations for energy based on climatic variables in northern Europe. *Global Change Biology Bioenergy*, **8**, 1093–1105.

Munns R (2002) Comparative physiology of salt and water stress. *Plant Cell Environment*, **25**, 239–250.

Nair SS, Kang S, Zhang X *et al.* (2012) Bioenergy crop models: descriptions, data requirements, and future challenges. *Global Change Biology Bioenergy*, **4**, 620–633.

Nippert JB, Wieme RA, Ocheltree TW, Craine JM (2011) Root characteristics of C4 grasses limit reliance on deep soil water in tallgrass prairie. *Plant and Soil*, **355**, 385–394.

NRCS (1993) *Soil Survey Manual*. Soil Conservation Service, Soil Survey Division Staff. Handbook 18, US Department of Agriculture. Available at: http://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/ref/?cid=nrcs142p2_054253 (accessed 4 January 2018).

NRCS (1999) *Liming to improve soil quality in acid soils*. Soil Quality Institute Technical Note 8, Natural Resources Conservation Service, US Department of Agriculture, Auburn, AL, USA.

NRCS (2003) Chapter 3: Ecological sites and ecological forage suitability groups. In: *National Range and Pasture Handbook*. US Department of Agriculture. Available at: http://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/landuse/rangepasture/?cid=STELPRDB1043084 (accessed 4 January 2018).

Pavelis GA (ed.) (1987) *Farm Drainage in the United States: History, Status, and Prospects*. USDA Economic Research Service Miscellaneous Publication 1455, Washington, DC.

Peet MM, Willits DH (1998) The effect of night temperature on greenhouse grown tomato yields in warm climate. *Agricultural and Forest Meteorology*, **92**, 191–202.

PRISM Climate Group (2015) *AN81 gridded climate dataset*. Oregon State University. Available at: http://prism.oregonstate.edu/ (accessed 4 January 2018).

Richter GM, Agostini F, Barker A, Costomiris D, Qi A (2016) Assessing on-farm productivity of *Miscanthus* crops by combining soil mapping, yield modelling and remote sensing. *Biomass and Bioenergy*, **85**, 252–261.

Scherer T, Sands G, Kandel H, Hay C (2015) *Frequently asked questions about subsurface (tile) drainage*. Extension publication AE1690 (Revised), North Dakota State University, Fargo, North Dakota. Available at: https://www.ag.ndsu.edu/pubs/ageng/irrigate/ae1690.pdf (accessed 4 January 2018).

Schlenker W, Roberts MJ (2009) Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences*, **106**, 15594–15598.

Schwartz CJ, Doyle MR, Manzaneda AJ, Rey PJ, Mitchell-Olds T, Amasino RM (2010) Natural variation in flowering time and vernalization responsiveness in *Brachypodium distachyon*. *Bioenergy Research*, **3**, 38–46.

Song Y, Jain AK, Landuyt W, Kheshgi HS, Khanna M (2015) Estimates of biomass yield for perennial bioenergy grasses in the USA. *Bioenergy Research*, **8**, 688–715.

Stavridou E, Hastings A, Webster RJ, Robson PRH (2017) The impact of soil salinity on the yield, composition and physiology of the bioenergy grass *Miscanthus × giganteus*. *GCB Bioenergy*, **9**, 92–104.

Steduto P, Hsiao TC (1998) Maize canopies under two soil water regimes. II. Seasonal trends of evapotranspiration, carbon dioxide assimilation and canopy conductance, and as related to leaf area index. *Agricultural and Forest Meteorology*, **89**, 185–200.

Steduto P, Hsiao TC, Fereres E, Raes D (2012) *Crop yield response to water*. Food and Agriculture Organization, Irrigation and Drainage Paper 66. Available at: http://www.fao.org/docrep/016/i2800e/i2800e.pdf (accessed 4 January 2018).

Strullu L, Ferchaud F, Yates N *et al.* (2015) Multisite yield gap analysis of *Miscanthus × giganteus* using the STICS model. *Bioenergy Research*, **8**, 1735–1749.

Thomson AM, Izarrualde RC, West TO, Parrish DJ, Tyler DD, Williams JR (2009) *Simulating Potential Switchgrass Production in the United States*. Pacific Northwest National Laboratory, Richland, WA.

Triana F, o Di Nasso N, Ragaglini G, Roncucci N, Bonari E (2015) Evapotranspiration, crop coefficient and water use efficiency of giant reed (*Arundo donax* L.) and miscanthus (*Miscanthus × giganteus* Greef et Deu.) in Mediterranean environment. *Global Change Biology Bioenergy*, **7**, 811–819.

Tulbure MG, Wimberly MC, Boe A, Owens VN (2012) Climatic and genetic controls of yields of switchgrass, a model bioenergy species. *Agriculture, Ecosystems & Environment*, **146**, 121–129.

United States Department of Energy (USDOE) (2005) *Biomass as Feedstock for a Bioenergy and Bioproducts Industry: The Technical Feasibility of a Billion-Ton Annual Supply* (Leads Perlack RD, Wright LL, Turhollow AF, Graham RL, Stokes BJ, Erbach DC), ORNL/TM-2005/66. Oak Ridge National Laboratory, Oak Ridge, TN.

United States Department of Energy (USDOE) (2011) *U.S. Billion-Ton Update: Biomass Supply for a Bioenergy and Bioproducts Industry*. (Leads Perlack RD, Stokes BJ), ORNL/TM-2011/224. Oak Ridge National Laboratory, Oak Ridge, TN.

United States Department of Energy (USDOE) (2016) *2016 Billion-Ton Report: Advancing Domestic Resources for a Thriving Bioeconomy, Volume 1: Economic Availability of Feedstocks* (Leads Langholtz MH, Stokes BJ, Eaton LM), ORNL/TM-2016/160. Oak Ridge National Laboratory, Oak Ridge, TN.

Volk TA, Berguson B, Daly C *et al.* (2017) Poplar and Shrub Willow Energy Crops in the United States: Field Trial Results from the Multiyear Regional Feedstock Partnership and Yield Potential Maps Based on the PRISM-ELM Model. *Global Change Biology Bioenergy*, https://doi.org/10.1111/gcbb.12498

Wahid A, Gelani S, Ashraf M, Foolad MR (2007) Heat tolerance in plants: an overview. *Environmental and Experimental Botany*, **61**, 199–233.

Wang D, Jaiswal D, LeBauer DS, Wertin TM, Bollero GA, Leakey ADB, Long SP (2015) A physiological and biophysical model of coppice willow (*Salix* spp.) production yields for the contiguous USA in current and future climate scenarios. *Plant, Cell and Environment*, **38**, 1850–1865.

Wightman J, Ahmed ZU, Volk TA *et al.* (2015) Assessing sustainable bioenergy feedstock production potential by integrated geospatial analysis of land use and land quality. *Bioenergy Research*, **8**, 1671–1680.

Williams JR, Jones CA, Dyke PT (1984) A modeling approach to determining the relationship between erosion and soil productivity. *Transactions of the ASAE*, **27**, 129–144.

Wullschleger SD, Davis EB, Borsuk ME, Gunderson CA, Lynd LR (2010) Biomass production in switchgrass across the United States: database description and determinants of yield. *Agronomy Journal*, **102**, 1158–1168.

Yimam YT, Ochsner TE, Kakani VG (2015) Evapotranspiration partitioning and water use efficiency of switchgrass and biomass sorghum managed for biofuel. *Agricultural Water Management*, **155**, 40–47.

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.

Environmental Limitation Mapping of

Potential Biomass Resources across the Conterminous United States

Christopher Daly[1]*, Michael D. Halbleib[1], David B. Hannaway[2], and Laurence M. Eaton[3]

[1]PRISM Climate Group, Northwest Alliance for Computational Science and Engineering, 2000 Kelley Engineering Center, Oregon State University, Corvallis, Oregon, USA

[2] Department of Crop and Soil Science, 125 Crop Science Building, Oregon State University, Corvallis, Oregon, USA

[3] Bioenergy Resource and Engineering Systems Group, Environmental Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

* Corresponding author: Christopher Daly, PRISM Climate Group, Northwest Alliance for Computational Science and Engineering, 2000 Kelley Engineering Center, Oregon State University, Corvallis, Oregon, USA, tel. 541-737-2531, fax 541-737-6609, e-mail: Chris.Daly@oregonstate.edu

Supporting Information

Contents

5. 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.

6. 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.

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

7. Environmental Suitability Mapping for Biomass Crops

8. References

1.   Basic Environmental Constraints on Production
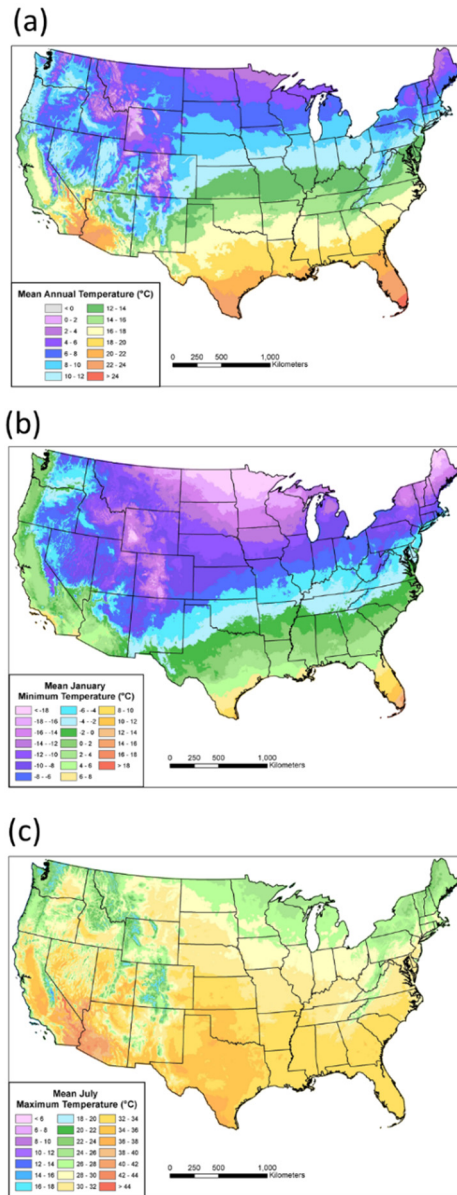
Figure S1. Conterminous US 1981-2010 (a) mean annual temperature, (b) mean January minimum temperature, and (c) mean July maximum temperature. Data source: PRISM Climate Group (http://prism.oregonstate/edu).
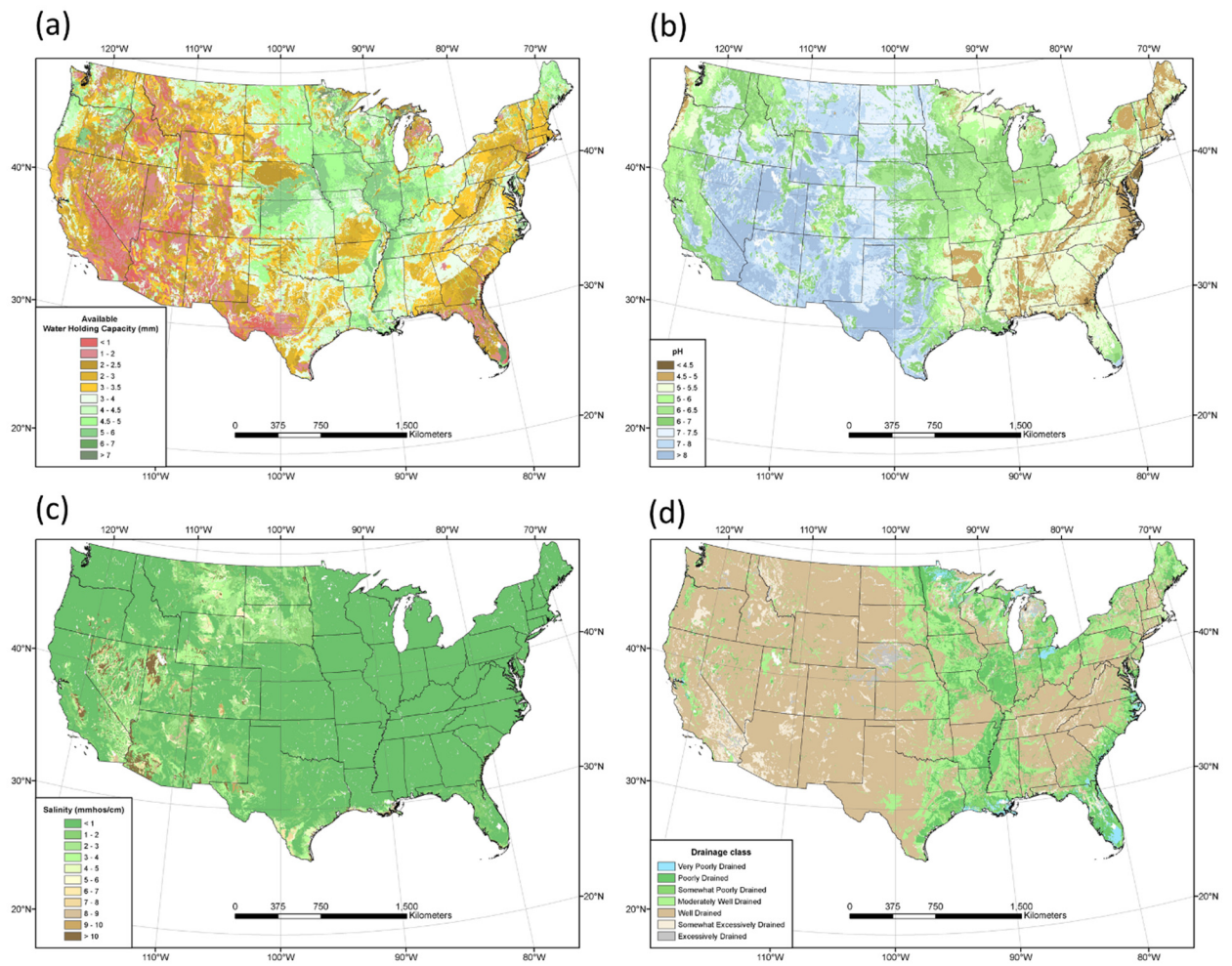
Figure S2. Conterminous US soil (a) available water capacity (b) pH, (c) salinity, and (d) drainage. Data source: USDA-NRCS.

2.   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. Details on the herbaceous and woody field trial data are available from Lee et al. (2017) and Volk et al. (2017), respectively. In the source column, RFP refers to the Sun Grant Regional Feedstock Partnership.

Crop	State	Longitude	Latitude	Source
CRP Grasses				
	GA	-83.4482	33.78938	RFP
	KS	-99.4106	38.84192	RFP
	MO	-92.1863	38.97347	RFP
	MT	-110.011	47.08814	RFP
	ND	-99.2282	47.48365	RFP
	OK	-99.2998	34.73266	RFP
Energycane				
	GA	-83.54335	33.873133	RFP
	TX	-94.293861	30.067667	RFP
	MS	-88.792661	33.422828	RFP
	MS	-90.51865	32.222247	RFP
	LA	-91.105739	30.267056	RFP
	GA	-83.916372	31.866922	RFP
Miscanthus				
	LA	-91.1031	30.4111	Arundale (2012)
	GA	-83.591947	31.439648	T. Voigt (Pers. Comm.)
	MS	-88.7953	33.4245	Arundale (2012)
	MS	-88.79746	33.43798	B. Baldwin (Pers. Comm.)
	MS	-88.79716	33.43895	B. Baldwin (Pers. Comm.)
	OK	-97.046	35.9917	Arundale (2012)

	VA	-79.395449	36.933481	RFP
	IL	-88.722914	37.453577	Arundale (2012)
	KY	-84.4971	38.12778	RFP
	IL	-88.389409	38.381002	Arundale (2012)
	IL	-90.819545	39.806104	Arundale (2012)
	IL	-88.190582	40.064943	RFP
	NJ	-74.2475	40.2245	RFP
	IA	-95.139	41.339	E. Heaton (Pers. Comm.)
	IA	-93.643	42.003	E. Heaton (Pers. Comm.)
	MI	-85.3765	42.3948	Arundale (2012)
	IA	-95.534	42.907	E. Heaton (Pers. Comm.)
	NE	-96.4656	41.1737	RFP
Lowland Switchgrass				
	NY	-76.1132	43.78954	R. Crawford (Pers. Comm.)
	NY	-76.95448	42.13432	R. Crawford (Pers. Comm.)
	VA	-80.4167	37.1833	Fike et al. (2006a, 2006b)
	NY	-73.47333	44.88524	R. Crawford (Pers. Comm.)
	OK	-97.91416	35.03166	Fuentes and Taliaferro (2002)
	NC	-78.4567	35.6506	Fike et al. (2006a, 2006b)
	TX	-96.3342	30.6278	Cassida et al. (2005)
	TX	-96.8039	32.7828	Cassida et al. (2005)
	NY	-78.05889	42.95017	R. Crawford (Pers. Comm.)
	OK	-95.6392	35.7425	Fuentes and Taliaferro (2002)
	AR	-93.5914	33.6669	Cassida et al. (2005)
	NY	-76.447	42.45	R. Crawford (Pers. Comm.)

	TN	-83.95	35.8833	Fike et al. (2006a, 2006b)
	NE	-96.4173	41.2222	R. Crawford (Pers. Comm.)
	WV	-79.9167	39.6167	Fike et al. (2006a, 2006b)
	VA	-78.1167	38.2167	Fike et al. (2006a, 2006b)
	KY	-87.8167	37.1	Fike et al. (2006a, 2006b)
	NC	-78.6667	35.7167	Fike et al. (2006a, 2006b)
	TX	-98.2	32.2167	Cassida et al. (2005)
	OK	-97.0581	36.1156	Aravindhakshan et al. (2010)
	IA	-93.75222	41.918888	Wilson et al. (2014)
	GA	-83.5086	31.0543	Knoll et al. (2012)
Upland Switchgrass				
	IA	-93.742783	42.008967	RFP
	SD	-96.7731	44.2538	Hong et al. (2013)
	IL	-88.959	38.9537	Anderson et al. (2013)
	TX	-96.3342	30.6278	Cassida et al. (2005)
	SD	-96.848	44.0012	Zilverberg et al. (2014)
	TX	-96.8039	32.7828	Cassida et al. (2005)
	SD	-97.8358	45.269	RFP
	IL	-88.8527	41.8418	Anderson et al. (2013)
	SD	-99.75	43.717	Mulkey et al. (2006)
	AR	-93.5914	33.6669	Cassida et al. (2005)
	NY	-76.382	42.456	RFP
	TN	-88.8333	35.6167	Fike et al. (2006a, 2006b)
	PA	-79.2378	40.2431	Adler et al. (2006)
	IL	-90.7204	40.9364	Anderson et al. (2013)
	OK	-95.6392	35.7425	RFP

	IL	-90.8221	39.8057	Anderson et al. (2013)
	SD	-100.34	44.3403	Hong et al. (2013)
	KY	-87.8167	37.1	Fike et al. (2006a, 2006b)
	NC	-78.6667	35.7167	Fike et al. (2006a, 2006b)
	NC	-78.6389	35.7719	Burns et al. (2002)
	PA	-76.1544	39.7276	Adler et al. (2006)
	TX	-98.2019	32.2206	Cassida et al. (2005)
	NY	-76.4606	42.4624	R. Crawford (Pers. Comm.)
	IL	-88.2239	40.0405	Anderson et al. (2013)
Biomass Sorghum				
	IA	-93.742783	42.008967	RFP
	KS	-96.597048	39.204116	RFP
	KY	-84.48827	38.129834	RFP
	MS	-90.521145	32.211111	RFP
	NC	-76.6597	35.85015	RFP
	TX	-96.445061	30.548953	RFP
	TX	-97.560149	27.777898	RFP
Willow				
	WI	-89.37	43.29	T. Volk (Pers. Comm.)
	NY	-76.11273	43.78984	RFP
	NY	-76.955931	42.135356	RFP
	MI	-84.46734	46.39951	RFP
	VT	-73.215637	44.497822	Liu (2013)
	NY	-75.772169	43.050882	Liu (2013)
	NY	-75.52596	43.55778	RFP
	MI	-87.19967	45.76824	RFP

	MI	-84.38503	42.82359	Liu (2013)
	VT	-73.19583	44.00944	RFP
	DE	-75.730289	39.326517	Liu (2013)
	MD	-76.15439	38.91376	Liu (2013)
	MI	-87.24657	46.3619	RFP
	CT	-72.231239	41.797333	RFP
	NY	-76.1164	42.79416	RFP
	NY	-76.738778	43.244695	Liu (2013)
Poplar				
	IL	-89.163944	36.97675	RFP
	GA	-81.966627	32.138428	RFP
	WA	-122.3	47.2	RFP
	SC	-80.720733	33.864867	RFP
	GA	-81.902633	33.29955	RFP
	MN	-95.057799	46.129283	RFP
	MN	-95.119014	46.18593	RFP
	KY	-89.162889	36.796722	RFP
	MN	-95.760589	48.151061	RFP
	MN	-95.145939	46.272252	RFP
	MN	-93.490207	47.249223	RFP
	TN	-89.585472	36.212083	RFP
	MN	-94.861783	46.195214	RFP
	MN	-95.008075	46.616042	RFP
	MO	-89.154667	36.757361	RFP
	MN	-95.145169	46.264689	RFP
	VA	-77.5764	37.702088	RFP

3. PRISM-ELM Model Formulation

A summary of PRISM-ELM structure and function is included in the main paper. Below are mathematical details on model formulation and operation.

Water Balance

At the start of the simulation, the total available water in the root zone is calculated, which is a product of the user-define rooting depth and the NRCS-defined available water capacity:

$$TAW = D_{root}\,AWC \tag{1}$$

Readily available water in the root zone is a product of TAW and the user-defined stress response factor, typically set to 0.5 (Allen et al., 1998):

$$RAW = p\,TAW \tag{2}$$

The simulation begins in the first half of January. At each semi-monthly time step t, the water stress coefficient of the crop (K_s) is calculated as:

$$K_{s(t)} = (TAW\text{-}D_{r(t)}) \,/\, (TAW\text{-}RAW) \tag{3}$$

$D_{r(t)}$, the root zone moisture depletion, is initialized to zero at the first time step (January). At subsequent time steps, it is a function of the precipitation and actual evapotranspiration at the previous time step $t\text{-}1$:

$$D_{r(t)} = D_{r(t-1)} + (ET_{a(t-1)} - P_{t-1}) \qquad (4)$$

where ET_a is a function of the reference evapotranspiration, water stress coefficient, and crop coefficient if the crop is present, and the water stress coefficient and soil evaporation if the crop is not present:

$$ET_{a(t-1)} = \text{If crop on: } ET_{0(t-1)}\, K_{S(t-1)}\, K_c;$$

$$\text{If crop off: } K_{S(t-1)}\, E_s \qquad (5)$$

A crop is assumed to be present if the time-step falls within the potential growth period, discussed below. A positive D_r indicates a moisture deficit, while a negative D_r indicates a moisture surplus.

The water balance calculations are run for a complete "spin up" year to allow the soil water stores to equilibrate. The year is then run again to obtain the final results, which are output at a monthly time step.

In concert with the water balance calculations above, the temperature response of the crop is evaluated at each time-step. The temperature response function is controlled by five user-defined parameters, *OptT, MaxT, Mag, F1, and F2,* based on controlled environment experimental data (e.g., Baker and Jung, 1968), crop simulation modeling (e.g., Brown et al., 1986), and species experts' knowledge [e.g., tall fescue growth response function (http://forages.oregonstate.edu/tallfescuemonograph/suitability/creating/growth)]. *OptT* and *MaxT* and are the mean daily temperatures at which crop growth is optimal and the maximum temperature at which crop growth declines to zero, respectively. *Mag* controls the maximum

value the response curve can reach, and *F1* and *F2* define the shape of the curve.   The

temperature response T_r of the modeled crop is calculated as a function of the mean daily

temperature at time step t (T_t):

$$C_t = \frac{(MaxT - T_t)}{(MaxT - OptT)}; \ C_t = max(C_t, 0) \tag{6}$$

$$T_{r(t)} = min(1, C_t^{F1} e^{\left(\frac{F1}{F2}(Mag - C_t^{F2})\right)}) \tag{7}$$

Once K_s and T_r are calculated for time step t, the suitability (S_t) for that time step is

$$S_t = K_{s(t)} \ T_{r(t)} \tag{8}$$

Once all time steps have been simulated in the year, the two semi-monthly S_t values in

each month are averaged to obtain monthly values. (S_m). Example S_m values for winter wheat

in the Willamette Valley, Oregon are shown in Figure S1. A Potential Suitability Window,

established for the crop by the user, defines the beginning month (M_{beg}) and ending month (M_{end})

of the period within which crop production is likely to be most active somewhere within the

modeling region, accounting for varying planting and harvest dates and regional climatic

differences. In the Figure S3 example, the Potential Suitability Window is defined as March-

July ($M_{beg} = 3$, $M_{end} = 7$). The final water balance suitability (S_w) for the crop is then calculated

as the average of the highest consecutive monthly S_m values within the Potential Suitability

Window of M_{avg} months in length, termed the Maximum Suitability Window. Experience with

PRISM-ELM simulations has shown a M_{avg} of three months as being optimal for most biomass

14

crops. In the Figure S1 example, the three greatest consecutive S_m values occurred during April-

June. The average of these three values is 84, which is the final S_w.
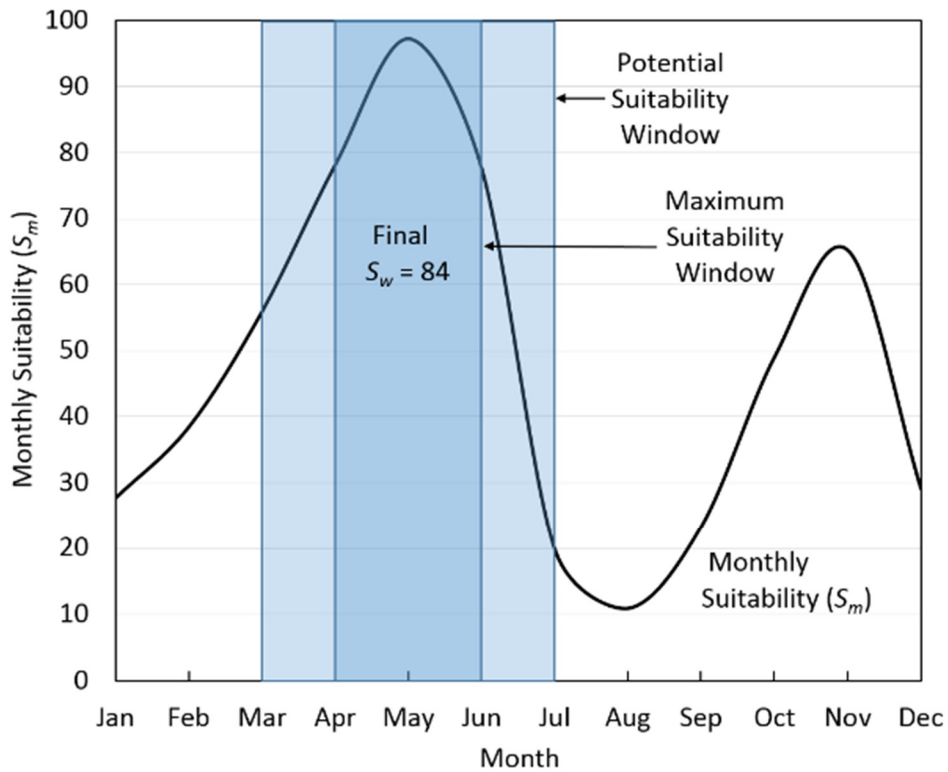

Figure S3. Illustration of method used to obtain final suitability (S_w) of winter wheat in the Willamette Valley, Oregon from the PRISM-ELM water balance model. In this example, monthly S_m values are shown in the black line. The Potential Suitability Window (shaded in light blue) is set to start in March and end in July. S_m values across the Maximum Suitability Window (shaded in dark blue), the highest yielding consecutive monthly span of three months, are then averaged to obtain the final S_w of 84.

Winter Low Temperature Response

The winter low temperature response curve is a metric for suitability reductions in perennial crops that may occur because of injury or death caused by excessively low winter temperatures. The two-tailed response curve relates the winter cold suitability index (S_c) to the 1981-2010 mean minimum temperature for the month of January:

$$if\ (x < Tmin_{min}\ or\ x > Tmin_{max})\ S_c = 0;\ else$$

$$S_c = Tmin_{mag}\, e^{\left[Tmin_{w1}\left(\frac{x - Tmin_{opt}}{Tmin_{w2}}\right)^2\right]};\ 0 \le S_c \le 100 \tag{9}$$

where x is the 1981-2010 January mean minimum temperature. User-defined parameters for the curve are the lower $Tmin$ at which S_c is zero ($Tmin_{min}$); upper $Tmin$ at which S_c is zero ($Tmin_{max}$); $Tmin$ at which S_c is at maximum ($Tmin_{opt}$); curve magnitude ($Tmin_{mag}$), which is the S_c value at $Tmin_{opt}$; and curve shape factors $Tmin_{w1}$ and $Tmin_{w2}$.

Summer High Temperature Response

The summer high temperature response curve is a metric for suitability reductions that may occur because of stress caused by high temperatures during the growing season. The one-

tailed response curve relates the summer high temperature suitability index (S_h) to the 1981-2010 mean maximum temperature for the month of July:

$$S_h = c_0 + c_1 x + c_2 x^2 + c_3 x^3 \qquad (10)$$

where x is mean July maximum temperature, and $c0$, $c1$, $c2$, and $c3$ are coefficients calculated by fitting a third-order polynomial to the user-defined x values that cause no yield reduction ($Tmax_{opt}$), 50% yield reduction ($Tmax_{mid}$) and 100% yield reduction ($Tmax_{max}$).

Soil pH Response

The soil pH response curve within PRISM-ELM accounts for suitability reductions caused by excessively acidic (low pH) or alkaline (high pH) soils. The two-tailed response curve relating the soil pH suitability index (S_p) to soil pH is of the same form as the winter low temperature response curve:

$$if\,(x < pH_{min}\,or\,x > pH_{max})\ S_p = 0;\ else$$

$$S_p = pH_{mag}\,e^{[pH_{w1}(\frac{x - pH_{opt}}{pH_{w2}})^2]}; 0 \le S_p \le 100 \qquad (11)$$

where x is the grid soil pH value. User-defined parameters for the curve are the lower pH threshold at which S_p is zero (pH_{min}); the upper pH threshold at which S_p is zero (pH_{max}); the pH

value at which S_p is at maximum (pH_{opt}); the curve magnitude (pH_{mag}), which is the S_p value at pH_{opt}; and curve shape factors pH_{w1} and pH_{w2}.

Soil Salinity Response

The soil salinity response curve accounts for suitability reductions that may occur because of excessive soil salinity. The one-tailed response curve for soil salinity is of the same form as the summer high temperature response, and relates the soil salinity suitability index (S_s) to soil salinity:

$$S_s = c_0 + c_1 x + c_2 x^2 + c_3 x^3 \tag{12}$$

where x is soil salinity, and $c0$, $c1$, $c2$, and $c3$ are coefficients calculated by fitting a third-order polynomial to the user-defined x values that result in an S_s of 100 (SS_{opt}), an S_s of 50 (SS_{mid}) and an S_s of 0 (SS_{max}).

Soil Drainage Response

Soil drainage deals with water supply issues that affect crop production and management. 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 is assigned an S_d value: very poorly drained (*VPD*), poorly drained (*PD*), somewhat poorly drained (*SPD*),

18

moderately well drained (*MWD*), well drained (*WD*), somewhat excessively drained (*SED*), and excessively drained (*ED*).

PRISM-ELM Final Suitability

Once all of the environmental suitability indexes described above have been evaluated for a grid cell, the PRISM-ELM final environmental suitability index (*ESI*) is:

$$ESI = min (S_w, S_c, S_h, S_p, S_s, S_d) \qquad (13)$$

4. Examples of PRISM-ELM Operation

Contrasting examples of the operation of the PRISM-ELM water balance model for wheat and maize are illustrated for the Willamette Valley, Oregon (45N, 123W; mean annual precipitation = 1037 mm) and extreme southeastern Indiana (39N, 85W; mean annual precipitation = 1097 mm) in Figure S4. While wheat and maize were not the focus of the biomass mapping effort, they were chosen as examples based on their contrasting, but relatively well-known, behaviors, and as aids to model parametrization for biomass crops (see Model Parameterization section in the main text). Wheat is a cool-season crop which typically reaches maximum production in spring, while maize is a warm-season crop that reaches maximum production during the warmer, summer months.

The climate regime in the Willamette Valley is Mediterranean, characterized by wet winters and dry summers (Figure S4a).   Wheat water usage increases in spring, as the temperature response and actual evapotranspiration rise (Figure S4b). In response to reduced

moisture replenishment as the dry summer months approach, root zone moisture drops below zero (indicating moisture deficit), causing evapotranspiration to be restricted. The Potential Suitability Window is set to March-July for wheat (shaded in light blue in Figure S4c). Within that window, the final water balance suitability, S_w, is calculated as the average suitability during the three-month Maximum Suitability Window (shaded in dark blue). In this example, the greatest suitability months are April-June, and $S_w = 84$, or 84% of optimum. S_m increases again in the fall, as moisture conditions improve and the temperature response is still relatively high. This is the period when winter wheat varieties are planted, but is not included in the Potential Suitability Window because above-ground productivity is typically low (Figure S4c).

In contrast to wheat, maize encounters greater environmental limitation in the Willamette Valley, because temperatures are lower than optimum, and precipitation is lowest during mid-summer, the period of maximum moisture demand. Maize begins to draw soil water in the spring, as temperature response and actual evapotranspiration rise (Figure S4d). In response to increased water usage and reduced replenishment from precipitation, root zone water becomes quickly depleted. Relatively cool temperatures limit the peak temperature response to mid-summer, too late for suitable moisture conditions. The Potential Suitability Window for maize is April-September (Figure S4e). The Maximum Suitability Window for this location is May-July, and $S_w = 25$, or only 25% of optimum.
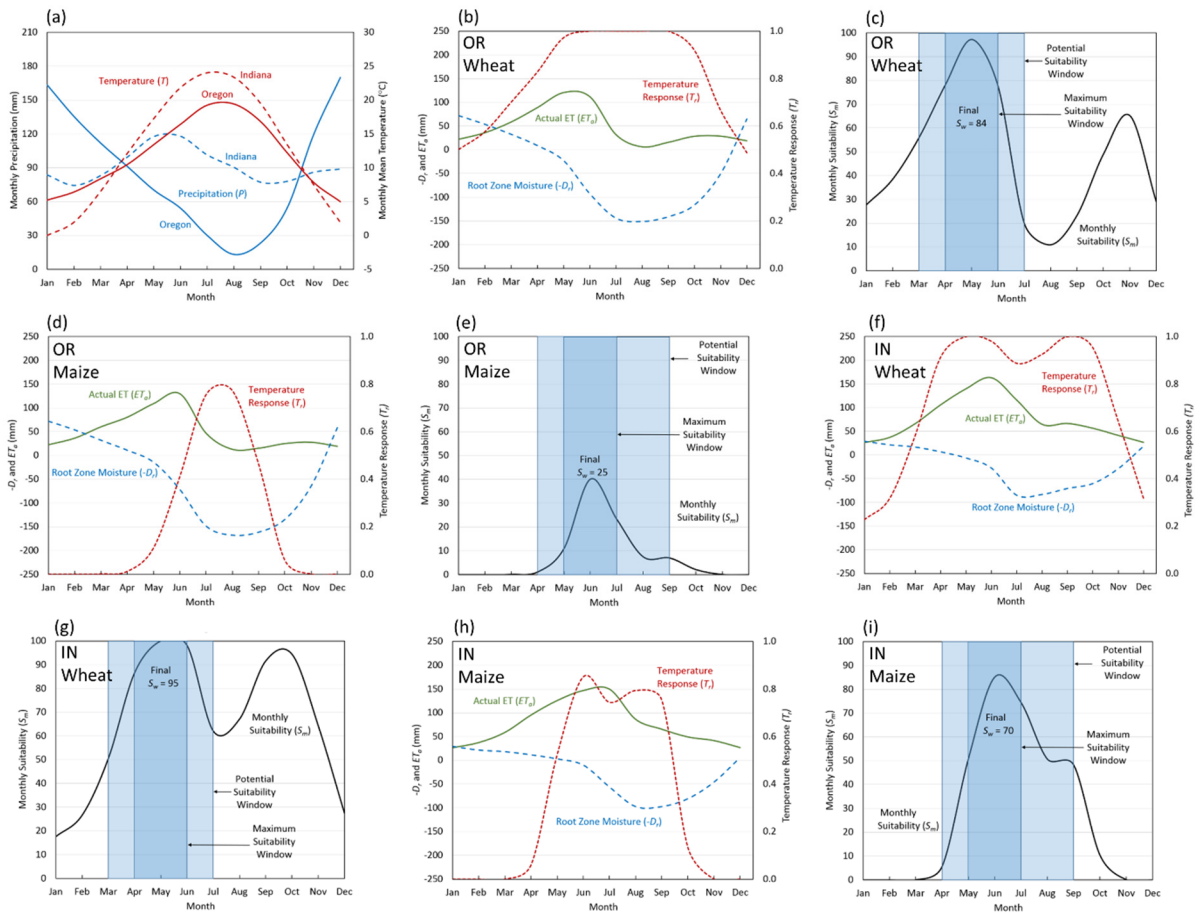
Figure S4. Comparison of PRISM-ELM water balance operation for wheat (cool-season crop) and maize (warm-season crop) in western Oregon, characterized by dry, mild summers, and southeastern Indiana, where summers are warm and moist. Shown are: (a) 1981-2010 mean monthly and temperature and precipitation; Oregon water balance model operation for wheat (b, c) and maize (d, e); and Indiana water balance model operation for wheat (f, g) and maize (h, i). In both locations, evapotranspiration of wheat and maize increases from spring into summer in response to increasing temperatures (b, d, f, h). Available moisture during spring is suitable for the production of wheat in both locations. Summer moisture is also suitable for maize production in Indiana, but summer drought limits its suitability in the Willamette Valley. See text for details.

In Indiana, precipitation is relatively evenly distributed throughout the year, and summer temperatures are warmer than those in the Willamette Valley (Figure S4a). The timing of optimal temperature response for wheat coincides with suitable moisture conditions in spring to allow relatively high evapotranspiration rates (Figure S4f). As a result, the final $S_w = 95$, or 95% of optimum (Figure S4g). As seen in the Willamette Valley, wheat S_m increases again in the fall, as moisture conditions improve and temperature response is still relatively high. Maize fares better in Indiana than in the Willamette Valley, due to ample precipitation in summer (Figure S4a). Soil moisture does not reach a minimum until late summer, allowing a longer period of production when temperatures are closer to the optimum range (Figure S4h). The result is a final S_w of 70, or 70% of optimum (Figure S4i).

5. PRISM-ELM Input Parameters

PRISM-ELM was parameterized for six herbaceous and two woody biomass feedstocks. Parameter symbols and descriptions are given in Table S2. Parameter values are listed for herbaceous biomass crops in Table S3 and woody biomass crop in Table S4. Details on model parameterization are given in the main text.

Table S2. PRISM-ELM input parameters and descriptions.

Parameter	Description
Water Balance	
M_{beg}	First month in potential suitability window
M_{end}	Last month in potential suitability window
M_{avg}	Length of maximum suitability window (months)
K_c	Crop coefficient
D_{root}	Average rooting depth (m)
p	Water stress response factor
$MaxT$	Temperature response maximum (ºC)
$OptT$	Temperature response optimum (ºC)
Mag	Temperature response magnitude
$F1, F2$	Temperature response shape factors 1, 2
Winter Temperature	**January T_{min} Response Function**
$Tmin_{opt}$	Optimum (ºC)
$Tmin_{max}$	Maximum (ºC)
$Tmin_{min}$	Minimum (ºC)
$Tmin_{mag}$	Magnitude
$Tmin_{w1}, Tmin_{w2}$	Shape factors 1, 2
Summer Temperature	**July T_{max} Response Function**
$Tmax_{opt}$	Optimum (ºC)
$Tmax_{mid}$	Mid - 50% yield (ºC)
$Tmax_{max}$	Max – 0% yield (ºC)
Soil pH	**Soil pH response Function**
pH_{opt}	Optimum

pH_{max}	Maximum
pH_{min}	Minimum
pH_{mag}	Magnitude
pH_{w1}, pH_{w2}	Shape factors 1, 2
Soil Salinity	**Soil Salinity Response Function**
SS_{opt}	Optimum (mmhos/cm)
SS_{mid}	Mid - 50% yield (mmhos/cm)
SS_{max}	Max – 0% yield (mmhos/cm)
Soil Drainage	**Soil Drainage Classes**
VPD	Very poorly drained
PD	Poorly drained
SPD	Somewhat poorly drained
MWD	Moderately well drained
WD	Well drained
SED	Somewhat excessively drained
ED	Excessively drained

Table S3. PRISM-ELM input parameters for herbaceous biomass crops, including winter wheat and maize, ordered left to right from low to high optimum temperature (*OptT*). See Table S2 for parameter descriptions.

Parameter	CRP	Winter Wheat*	Miscanthus	Upland Switchgrass	Maize*	Lowland Switchgrass	Biomass Sorghum	Energy-cane
Water Balance								
M_{beg} (month)	3	3	4	4	4	4	4	3
M_{end} (month)	9	7	9	9	9	9	9	11
M_{avg} (months)	3	3	3	3	3	3	3	3
K_c	1.0	1.0	1.1	1.05	0.9	1.1	1.0	1.2
D_{root} (m)	1.0	1.0	1.0	1	1.1	1.0	1.0	1
p	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
$MaxT$ (°C)	34	30	32	30	30	38	38	38
$OptT$ (°C)	17	18	20	21	21.5	24	26	27
Mag	1	1.4	1.5	1.25	1.05	1.2	1.25	1.3
$F1, F2$	3.4, 3.8	0.5, 2.7	1.4, 2.6	1.0, 2.2	1.2, 3.3	2.2, 1.8	0.4, 3.8	1.5, 2.2
Winter Temperature								
$Tmin_{opt}$ (°C)	-6	-3.85	-9.6	-12	NA	2	NA	12
$Tmin_{max}$ (°C)	-8	25	14	-12	NA	2	NA	12
$Tmin_{min}$ (°C)	-22	-30	-23	-25	NA	-24	NA	-3
$Tmin_{mag}$ (°C)	100	100	100	100	NA	100	NA	140
$Tmin_{w1}, Tmin_{w2}$	-0.7, 9.0	-0.71, 16	-0.6, 12	-0.7, 10.0	NA	-0.7, 10.0	NA	-0.8, 8.0
Summer Temperature								
$Tmax_{opt}$ (°C)	28	27.5	30.5	31	31	N/A	NA	N/A
$Tmax_{mid}$ (°C)	34	33.9	33.5	35.5	34	N/A	NA	N/A

$Tmax_{max}$ (°C)	38	39	42	43	40	N/A	NA	N/A
Soil pH								
pH_{opt}	6.5	6.5	6.5	6.5	6.5	6.5	6.5	6.5
pH_{max}	11	11	11	11	11	11	11	11
pH_{min}	3	3	3	3	3	3	3	3
pH_{mag}	150	150	150	150	150	150	150	150
pH_{w1}, pH_{w2}	-0.5, 1.6	-0.5, 1.6	-0.5, 1.6	-0.5, 1.6	-0.5, 1.6	-0.5, 1.6	-0.5, 1.6	-0.5, 1.6
Soil Salinity								
SS_{opt} (mmhos/cm)	0	0	0	0	0	0	0	0
SS_{mid} (mmhos/cm)	10	10	10	10	10	10	10	10
SS_{max} (mmhos/cm)	16	16	16	16	16	16	16	16
Soil Drainage								
VPD	90	90	90	90	90	90	90	90
PD	95	95	95	95	95	95	95	95
SPD	100	100	100	100	100	100	100	100
MWD	100	100	100	100	100	100	100	100
WD	100	100	100	100	100	100	100	100
SED	95	95	95	95	95	95	95	95
ED	90	90	90	90	90	90	90	90

*"Anchor" species used to guide model parameterization of biomass feedstock species.

Table S4. PRISM-ELM input parameters for woody biomass crops, ordered from left to right by optimum temperature (*OptT*).   See Table 1 for parameter descriptions.

Parameter	Willow	Poplar
Water Balance		
M_{beg} (month)	3	3
M_{end} (month)	9	9
M_{avg} (months)	3	3
K_c	1.05	1.2
D_{root} (m)	0.9	1.1
p	0.5	0.5
$MaxT$ (°C)	35	36
$OptT$ (°C)	22	21
Mag	1.35	1.45
$F1, F2$	2.7, 3.6	3.0, 2.4
Winter Temperature		
$Tmin_{opt}$ (°C)	NA	-1
$Tmin_{max}$ (°C)	NA	-1
$Tmin_{min}$ (°C)	NA	-30
$Tmin_{mag}$ (°C)	NA	120
$Tmin_{w1}, Tmin_{w2}$	NA	-0.8, 20
Summer Temperature		
$Tmax_{opt}$ (°C)	30	30
$Tmax_{mid}$ (°C)	32	37
$Tmax_{max}$ (°C)	36	45
Soil pH		
pH_{opt}	6.5	6.5
pH_{max}	11	11
pH_{min}	3	3
pH_{mag}	150	150
pH_{w1}, pH_{w2}	-0.5, 1.6	-0.5, 1.6
Soil Salinity		

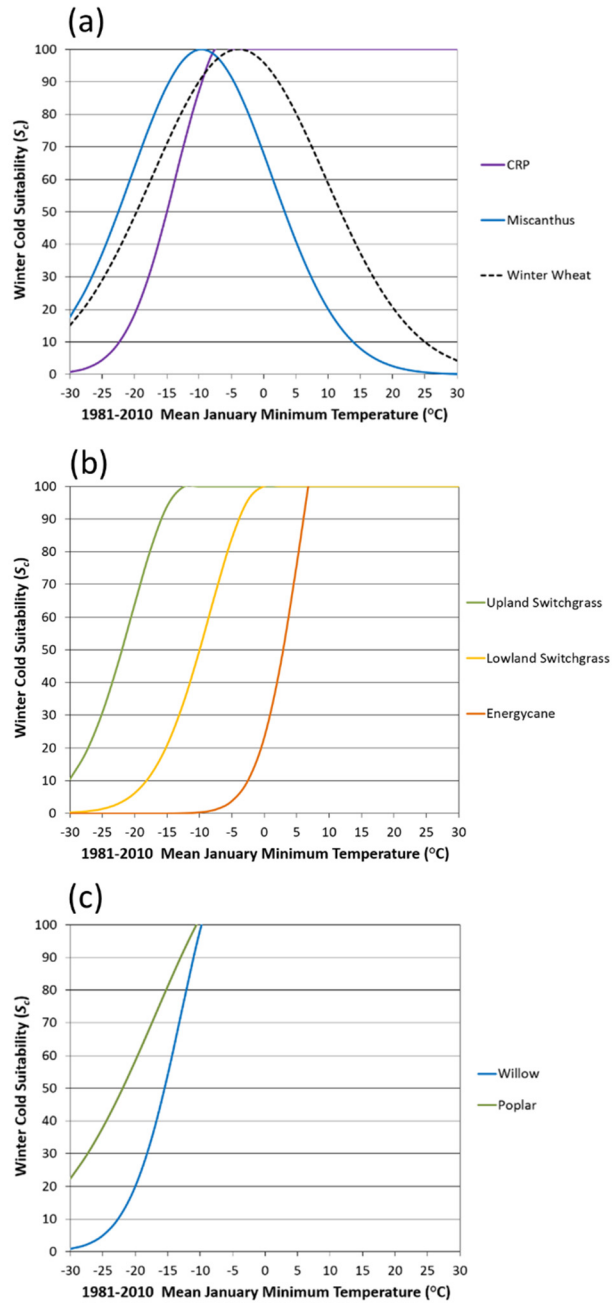|		
SS_{opt} (mmhos/cm)	0	0
SS_{mid} (mmhos/cm)	10	10
SS_{max} (mmhos/cm)	16	16
Soil Drainage		
VPD	90	90
PD	95	95
SPD	100	100
MWD	100	100
WD	100	100
SED	95	95
ED	90	90

Figure S5. PRISM-ELM January minimum temperature response curves (a metric for winter cold injury) for (a, b) overwintering herbaceous and (b) woody biomass crops. Winter wheat is shown for comparison. Parameter values are defined in Table S2 and listed in Tables S3 and S4.
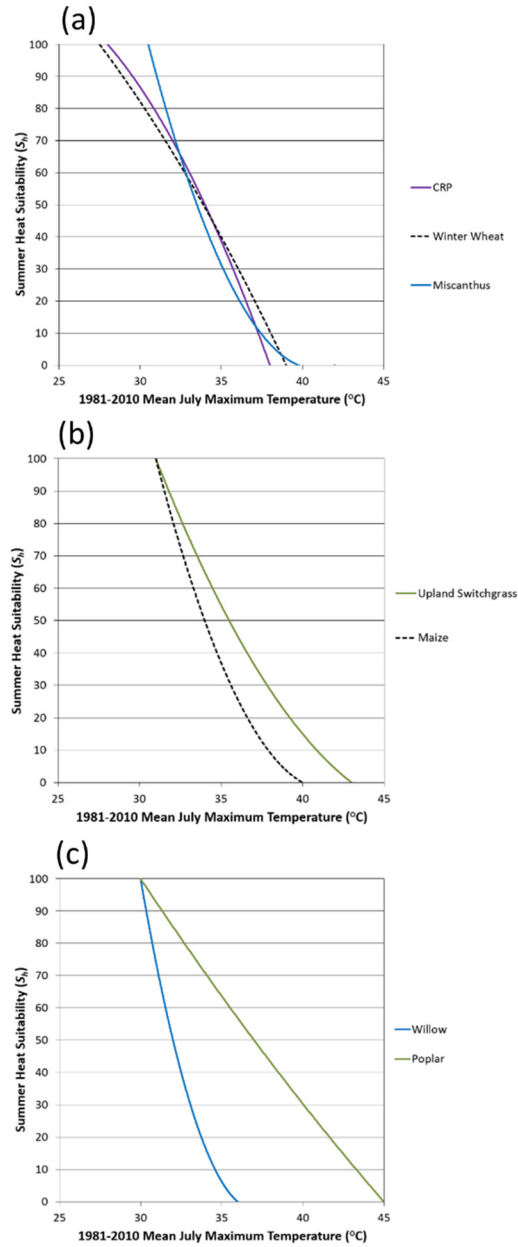
Figure S6. PRISM-ELM July maximum temperature response curves (a metric for heat injury) for (a, b) herbaceous and (c) woody biomass crops. Lowland switchgrass, biomass sorghum, and energycane had no temperature limits for heat injury. Parameter values are defined in Table S2 and given in Tables S3 and S4.
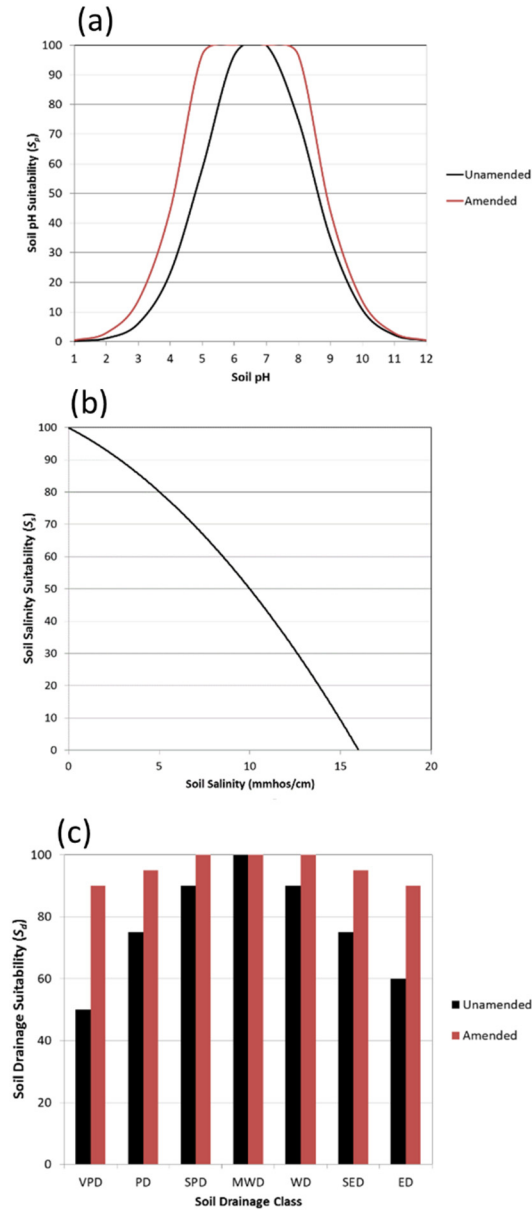
Figure S7. PRISM-ELM relative yield response curves to: (a) soil pH; (b) salinity; and (c) drainage, for all biomass crops. The pH response curve was broadened to account for soil liming practices, which modified the soil pH from NRCS native soil values. The drainage responses were also broadened to account for extensive field tiling to improve drainage, which are also not reflected in the NRCS native soil values. Parameter values are defined in Table S2 and given in Tables S3 and S4.
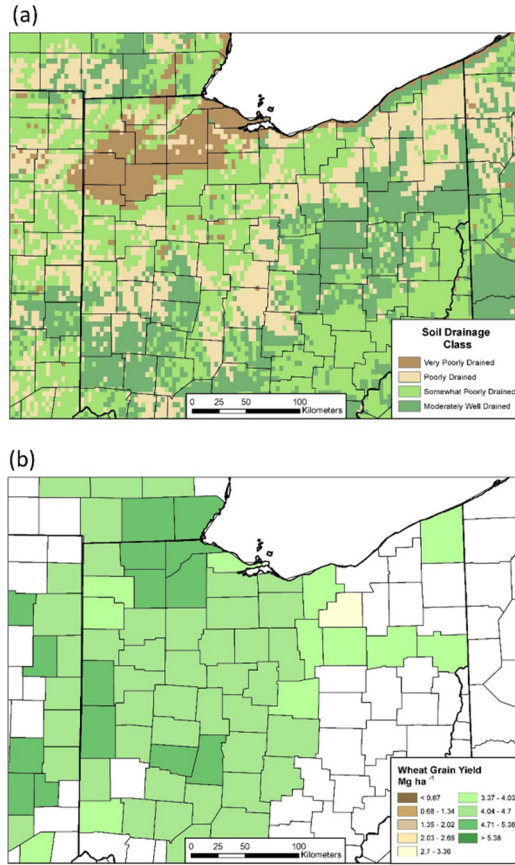
Figure S8. Comparison of (a) USDA NRCS soil drainage class and (b) RMA county-level winter wheat yields for northern Ohio. Relatively high yields are obtained through drainage tiling in areas of otherwise poorly drained soils.

6. Winter Wheat and Maize Model Validation

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 (Figure S9).   Yield data were obtained from the Risk Management Agency (RMA) of the United States Department of Agriculture (USDA) as part of a cooperative agreement between

RMA and Oregon State University. Producers participating in the federal crop insurance program are required to report acreage planted and harvested yield, and these data are aggregated by RMA to produce county-level statistics.   Winter wheat was chosen over spring wheat because of the relatively large amount of yield data available, and this choice also allowed the evaluation of PRISM-ELM's winter temperature component. The RMA winter wheat yield data were available in several broadly defined categories: Winter wheat (no practice specified), durum, irrigated, non-irrigated, summer fallow, continuous, and spring. We selected winter wheat (no practice specified) and continuous, which together comprised most of the yield records. For maize (corn), irrigated and non-irrigated categories were available, and we selected the non-irrigated data. Yield data were available for the period 2000-2014. To avoid unrepresentative county-level data, only those counties that had at least thirty wheat yield records or ninety maize yield records per county in at least eight of the fifteen years were selected.

The patterns of PRISM-ELM *ESI* match those of the USDA RMA reported yields reasonably well where there are data (Figures S9 and S10). Both indicate maximum maize yields in the Iowa-Illinois-Indiana corridor, and relatively high yields in the northeast. Modeled and reported yields decrease along the east-to-west precipitation gradient in the plains in a similar fashion. Winter wheat data coverage is not as extensive, but maximum yields are also reproduced in the Midwest and east, as is a similar gradient in the plains. PRISM-ELM's relatively high *ESI* values in the Pacific Northwest are also corroborated by RMA data.
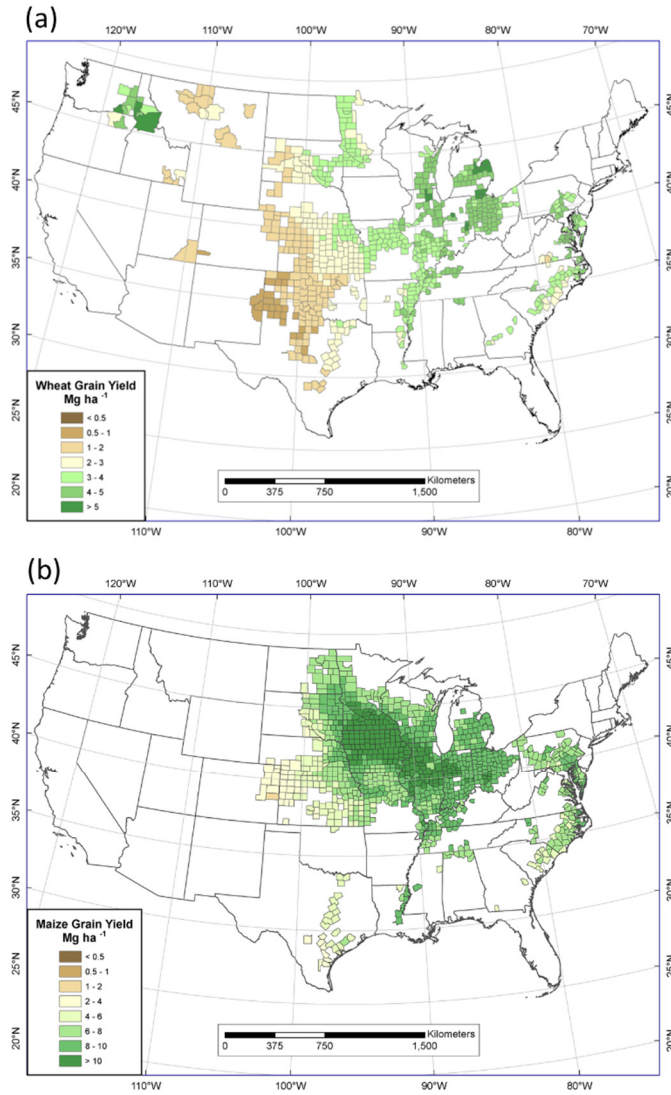
Figure S9. USDA RMA county-level grain yield data for non-irrigated (a) winter wheat and (b) maize. Counties shown had at least thirty reports for at least eight years during the period 2000-2015 for wheat, and ninety reports for eight years during the same period for maize.
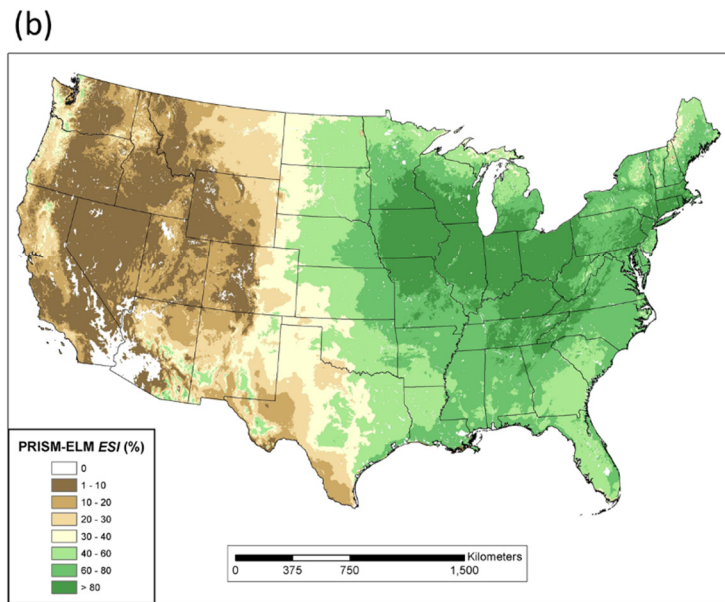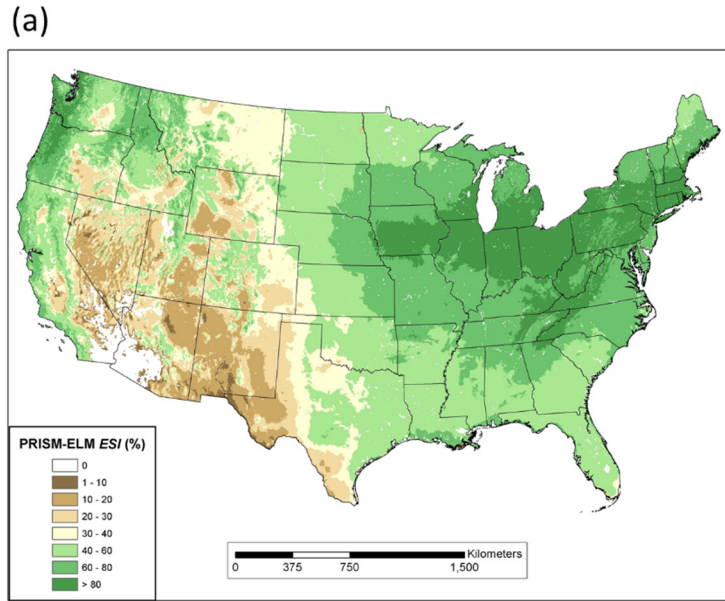
Figure S10. PRISM-ELM *ESI* maps for (a) winter wheat and (b) maize.

PRISM-ELM *ESI* output was expressed as a grid with 800-m spatial resolution, while the yield data reported by RMA were expressed as county averages. To allow a statistical comparison of obsevered and modeled values, the PRISM-ELM grid values were summarized at the county level. Since only a portion of a county may have been in agricultural production, the 2013 USDA NASS National Cultivated Layer (https://www.nass.usda.gov/Research_and_Science/Cropland/Release/) was used to identify areas under cultivation. The National Cultivated Layer (NCL) identified areas that were cultivated during the most recent five years at the time of access (2008-2012). The data layer was available as a grid at 30-m resolution. A grid cell was considered cultivated if there was agricultural activity in at least two out of the five years. The cultivated layer was resampled to an 800-m grid that was coincident with the PRISM-ELM climate and soils grids. An 800-m grid cell that contained at least one 30-m cultivated grid cell was considered cultivated. Only cultivated 800-m grid cells were included in the county averages.

The RMA yield datasets were randomly divided into a training half and an evaluation half (fifty percent data withholding). The training half was used in the initial parameterization of PRISM-ELM and model performance assessed. Using the same parameter settings, PRISM-ELM was then applied to the evaluation half of the data, and finally to the entire dataset, with model performance assessed at each step. The statistical relationship between PRISM-ELM *ESI* and RMA average winter wheat and maize yields was fairly consistent for the training, evaluation, and full RMA datasets (Table S5) (Figure S11). Winter wheat correlation

36

coefficients (for a least-squares regression forced through zero) declined from 0.75 for the training dataset to 0.71 for the evaluation dataset; the MAE changed little, averaging about 16 percent. Correlation coefficients for maize were lower, ranging from 0.59 for the training dataset to 0.57 for the evaluation dataset. Percent MAEs were slightly lower than those of winter wheat, averaging 14.1 percent for the training dataset and 14.7 percent for the evaluation dataset.

There are obvious drawbacks to using grain yield as validation datasets for PRISM-ELM. Reported grain yield from farms across the country represent a sampling of production outcomes that reflect a myriad of interrelated management decisions and economic forces that can partially or completely mask the basic environmental constraints simulated in PRISM-ELM. Hence, the relationships between PRISM-ELM suitability estimates and observed yields were expected to have significant scatter.

Table S5. PRISM-ELM performance statistics for RMA winter wheat and maize yield.
Comparisons were made after the PRISM-ELM *ESI* was transformed into actual yield using
national regression equations forced through zero. The dataset was divided into training and
evaluation halves, with performance statistics calculated for each, as well as for the entire
dataset. Model parameter settings were the same for both the training and evaluation datasets.

	Regression Equation	R^2	MAE (Mg ha^{-1}yr^{-1} / %)
Winter Wheat			
RMA Training Dataset (N=358)	y=0.0501x	0.75	0.49 / 15.8
RMA Evaluation Dataset (N=358)	y=0.0508x	0.71	0.49 / 15.9
RMA Full Dataset (N=716)	y=0.0505x	0.73	0.49 / 15.6
Maize			
RMA Training Dataset (N=495)	y = 0.1057x	0.59	1.05 / 14.1
RMA Evaluation Dataset (N=495)	y = 0.1039x	0.57	1.07 / 14.7
RMA Full Dataset (N=990)	y = 0.1048x	0.58	1.06/ 14.4

(a)

$$y = 0.0505x$$
$$R^2 = 0.73$$
$$MAE = 0.49\ (15.6\%)$$
$$N=716$$

(b)

$$y = 0.1048x$$
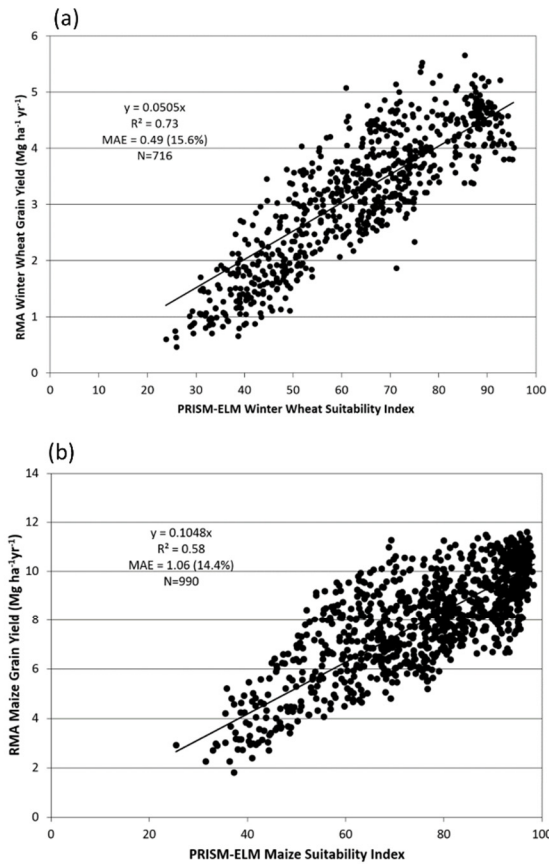$$R^2 = 0.58$$
$$MAE = 1.06\ (14.4\%)$$
$$N=990$$

Figure S11. Scatterplots and least-squares linear regressions forced through zero between county-level PRISM-ELM *ESI* and the full RMA dataset of 2000-2015 average annual reported grain yields for (a) winter wheat and (b) maize.

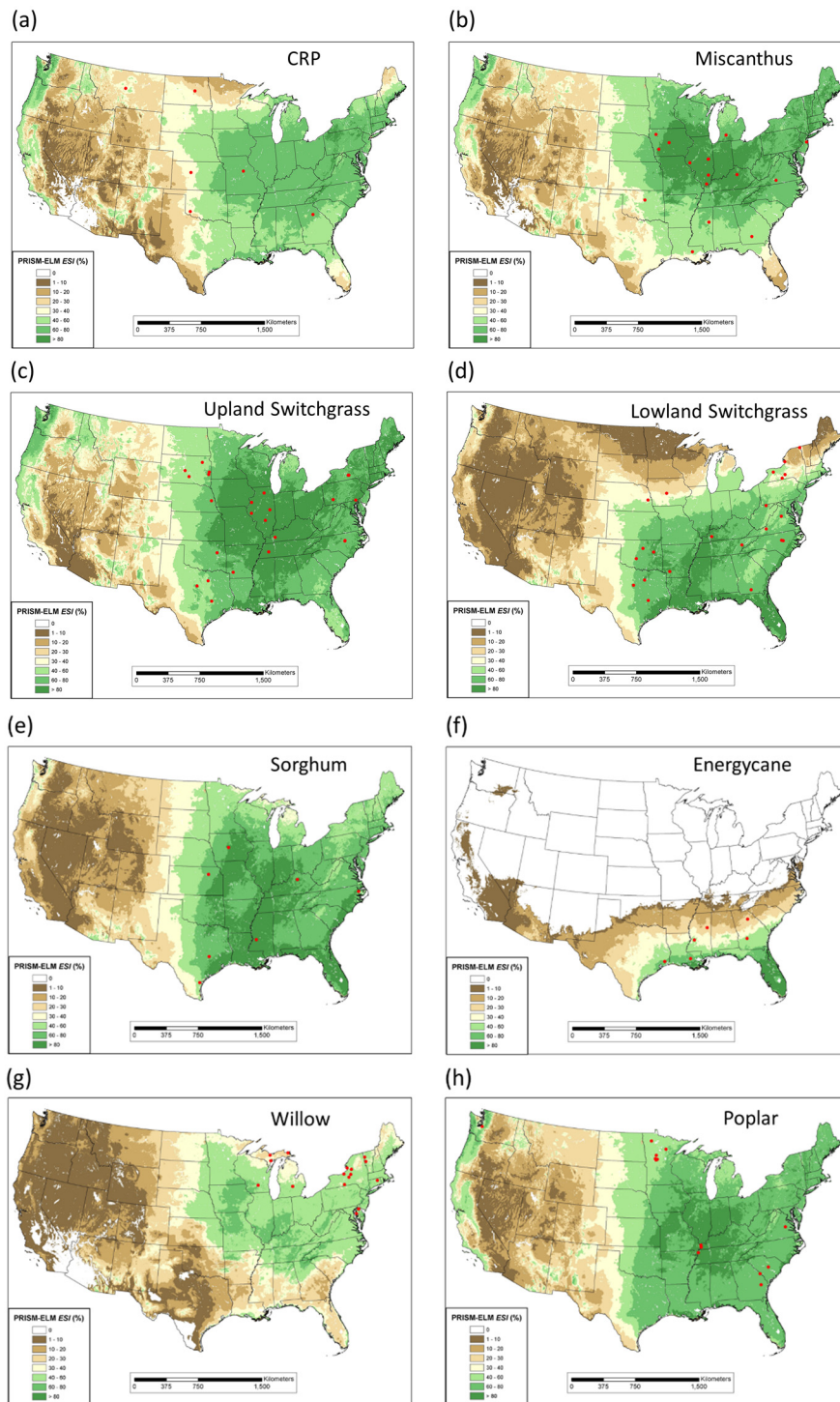7.   Environmental Suitability Mapping for Biomass Crops

Figure S12. PRISM-ELM *ESI* distributions for herbaceous and woody biomass crops. Locations of field trials used in the regression functions relating *ESI* to potential yield are shown as red dots.

8. References

Adler PR, Sanderson MA, Boateng AA, Weimer PJ, Jung, H-JG (2006) Biomass yield and biofuel quality of switchgrass harvested in fall or spring. Agronomy Journal, **98** 1518–1525.

Allen RG, Pereira LS, Raes D, Smith M (1998) *Crop evapotranspiration - Guidelines for computing crop water requirements*. FAO Irrigation and Drainage Paper 56, Food and Agriculture Organization of the United Nations, Rome, Italy.

Anderson EK, Parrish AL, Voigt TB, Owens VN, Hong CH, Lee DK (2013) Nitrogen fertility and harvest management of switchgrass for sustainable bioenergy feedstock production in Illinois. Industrial Crops and Products **48,** 19-27.

Aravindhakshan SC, Epplin FM, Taliaferro CM (2010) Economics of switchgrass and Miscanthus relative to coal as feedstock for generating electricity. Biomass and Bioenergy **34,** 1375-1383.

Arundale R, (2012) The higher productivity of the bioenergy feedstock Miscanthus x giganteus relative to Panicum virgatum is seen both into the long term and beyond Illinois. Unpublished Ph.D. dissertation, University of Illinois Urbana-Champaign, Urbana, IL.

Baker BS, Jung GA (1968) Effect of environmental conditions on the growth of four perennial grasses. I. Response to controlled temperature. Agronomy Journal, **60**, 155-158.

Burns JC, Godshalk EB, and Godshalk DH (2002) Registration of 'Colony' Lowland Switchgrass J. Plant Registrations **4**, 189-194.

Brown WF, Moser LE, Klopfenstein TJ (1986) Development and validation of a dynamic model of growth and quality for cool season grasses. Agricultural Systems, **20**, 37-52.

Cassida KA, Muir JP, Hussey MA, Read JC, Venuto BC, Ocumpaugh WR (2005) Biofuel component concentrations and yields of switchgrass in South Central U.S. environments. Crop Science, **45,** 682-692.

Fike JH, Parrish DJ, Wolf DD, Balasko JA, Green JT, Rasnake M, Reynolds JH (2006a) Long-term yield potential of switchgrass-for-biofuel systems. Biomass & Bioenergy, **30,** 198–206.

Fike JH, Parrish DJ, Wolf DD, Balasko JA, Green JT, Rasnake M, Reynolds JH (2006b) Switchgrass production for the upper southeastern USA: influence of cultivar and cutting frequency on biomass yields. Biomass & Bioenergy, **30,** 207–213.

Fuentes RG and Taliaferro CM (2002) Biomass yield stability of switchgrass cultivars. p. 276–282. In: J Janick and A Whipkey (eds.) Trends in new crops and new uses. ASHS Press, Alexandria, VA.

Hong CO, Owens VN, Lee DK, and Boe A (2013) Switchgrass, big bluestem, and indiangrass monocultures and their two- and three-way mixtures for bioenergy in the northern Great Plains. Bioenergy Research, **6**, 229-239.

Knoll JE, Anderson WF, Strickland, TC, Hubbard, RK, Malik R, (2012) Low-input production of biomass from perennial grasses in the Coastal Plain of Georgia, USA. BioEnergy Research, **5** 206-214.

Liu, B (2013) Biomass production of willow short-rotation coppice across sites and determinants of yields for SV1 and SX61. Master of Science thesis, State University of New York College of Environmental Science and Forestry, Syracuse, New York December 2013.

Mulkey VR, Owens VN, Lee DK (2006) Management of switchgrass-dominated Conservation Reserve Program lands for biomass production in South Dakota. Crop Science, **46**, 712-720.

Wilson DM, Heaton EA, Schulte LA, et al. (2014) Establishment and short-term productivity of annual and perennial bioenergy crops across a landscape gradient. BioEnergy Research, **7**, 885-898.

Zilverberg CJ, Johnson WC, Owens V, et al. (2014) Biomass yield from planted mixtures and monocultures of native prairie vegetation across a heterogeneous farm landscape. Agric., Ecosystems and Environment, **186**, 148-159.