
© 2014 John Wiley & Sons Ltd

This version available [http://nora.nerc.ac.uk/503494/](http://nora.nerc.ac.uk/503494/)

NERC has developed NORA to enable users to access research outputs wholly or partially funded by NERC. Copyright and other rights for material on this site are retained by the rights owners. Users should read the terms and conditions of use of this material at [http://nora.nerc.ac.uk/policies.html#access](http://nora.nerc.ac.uk/policies.html#access)

This document is the author's final manuscript version of the journal article, incorporating any revisions agreed during the peer review process. There may be differences between this and the publisher's version. You are advised to consult the publisher's version if you wish to cite from this article.

The definitive version is available at [http://onlinelibrary.wiley.com/](http://onlinelibrary.wiley.com/)

Contact CEH NORA team at [noraceh@ceh.ac.uk](mailto:noraceh@ceh.ac.uk)

The NERC and CEH trademarks and logos (‘the Trademarks’) are registered trademarks of NERC in the UK and other countries, and may not be used without the prior written consent of the Trademark owner.
Modelling the carbon cycle of Miscanthus plantations: existing models and the potential for their improvement

Modelling Miscanthus C cycles

Andy D. Robertson1,2,3*, Christian A. Davies2, Pete Smith3, Marta Dondini3 and Niall P. McNamara1

1 Centre for Ecology and Hydrology, Lancaster Environment Centre, Library Avenue, Bailrigg, Lancaster LA1 4AP, UK
2 Shell International Exploration and Production, Shell Technology Center Houston, 3333 Highway 6 South, Houston, TX 77082-3101, USA
3 Institute of Biological and Environmental Sciences, University of Aberdeen, 23 St Machar Drive, Aberdeen AB24 3UU, Scotland, UK

* Correspondence: Andy Robertson
Centre for Ecology and Hydrology,
Lancaster Environment Centre,
Library Avenue,
Bailrigg,
Lancaster
LA1 4AP
UK

Telephone: +44 (0) 7749 676 725
E-mail: andber58@ceh.ac.uk

Keywords: C cycle model; Miscanthus; soil carbon; WIMOVAC; DayCent; Agro-IBIS; BioCro; DNDC; ECOSSE; carbon;

Research Review
Abstract

The lignocellulosic perennial grass *Miscanthus* has received considerable attention as a potential bioenergy crop over the last 25 years, but few commercial plantations exist globally. This is partly due to the uncertainty associated with claims that land use change (LUC) to *Miscanthus* will result in both commercially viable yields and net increases in carbon (C) storage. To simulate what the effects may be after LUC to *Miscanthus*, six process-based models have been parameterised for *Miscanthus* and here we review how these models operate. This review provides an overview of the key *Miscanthus* soil organic matter models and then highlights what measurers can do to accelerate model development. Each model (WIMOVAC, BioCro, Agro-IBIS, DAYCENT, DNDC and ECOSSE) is capable of simulating biomass production and soil C dynamics based on specific site characteristics. Understanding the design of these models is important in model selection as well as being important for field researchers to collect the most relevant data to improve model performance. The rapid increase in models parameterised for *Miscanthus* is promising but refinements and improvements are still required to ensure model predictions are reliable and can be applied to spatial scales relevant for policy. Specific improvements, needed to ensure the models are applicable for a range of environmental conditions, come under two categories: 1) increased data generation and 2) development of frameworks and databases to allow simulations of ranging scales. Research into non-food bioenergy crops such as *Miscanthus* is relatively recent and this review highlights that there are still a number of knowledge gaps regarding *Miscanthus* specifically. For example, the low input requirements of *Miscanthus* make it particularly attractive as a bioenergy crop but it is essential that we increase our understanding of the crop’s nutrient re-mobilisation and ability to host N-fixing organisms in order to derive the most accurate simulations.
Introduction

Interest in the C4 perennial grass *Miscanthus* as a renewable energy source has grown significantly over the last two decades. *Miscanthus* has great potential for large scale deployment as a bioenergy crop, used either for electricity generation in power stations, or as a future renewable source of bioethanol (Heaton *et al*., 2008; Karp and Richter, 2011). It is native to South and East Asia, but has been shown to often produce high yields without fertiliser input across Europe and North America, and is tolerant to a range of climatic conditions (Clifton-Brown *et al*., 2004; Dohleman and Long, 2009; Strullu *et al*., 2011; Poeplau and Don, 2013). If *Miscanthus* plantations are established for bioenergy purposes, land use change (LUC) is inevitable and the influence that LUC has on soil carbon stocks and greenhouse gas (GHG) emissions is a key component of assessing sustainability within a bioenergy context. Carbon (C) accounting as a site management practice is becoming increasingly important (Borak *et al*., 2013) and consequently models need to become better in representing the full C cycle. Modelling of C dynamics in terrestrial ecosystems will undoubtedly become more comprehensive over time, but we are currently limited, not by our knowledge of the processes governing C transfer, but rather by the availability of reliable field data and high resolution large spatial datasets to test that understanding. This empirical data can be used to either validate the processes that govern a model’s simulation, or to verify the model's outputs and predictions. Therefore, a lack of validation data also means a lack of verification data, thereby creating questions about uncertainty quantification of current simulations.

Modelling studies consistently report an increase in soil C stocks after LUC from most different land uses into *Miscanthus* (Don *et al*., 2012; Mishra *et al*., 2012) but empirical
studies of LUC to Miscanthus show mixed results: some show a notable increase in topsoil soil C stocks each year (Hansen et al., 2004) whereas others report no significant increase (Schneckenberger and Kuzyakov, 2007; Zimmermann et al., 2012; Zatta et al., 2013) or various changes including reductions (Poeplau and Don, 2013). Consequently, this discrepancy needs to be addressed so accurate predictions can be made about the environmental impacts of the LUC and the C budgets associated with Miscanthus plantations. Such uncertainties have helped contribute to the limited establishment of Miscanthus and it is imperative that before wide-scale deployment is undertaken, model outputs are verified by robust and wide-ranging field data. These empirical data are also needed to underpin validation of the mechanistic aspects of models predicting biomass yields and ecosystem C dynamics, with improved validation leading to reduced uncertainty.

Validating model processes and verifying their outputs is particularly difficult for emerging non-native crops, like Miscanthus, as relevant field data is scarce. Current literature highlights that a lack of field data has limited 1) effective model parameterisation, 2) the quantification of model uncertainty, 3) inter-model comparisons and 4) the eventual application of Miscanthus models (Clifton-Brown et al., 2007; Miguez et al., 2009; Hastings et al., 2009b; VanLoocke et al., 2010; Cuadra et al., 2012; Miguez et al., 2012; Surendran Nair et al., 2012).

The last decade has seen a rise in the number and sophistication of mechanistic models capable of simulating C dynamics of Miscanthus plantations, but there are still aspects of these models that can be improved. These include, but are not limited to, the longevity of soil C stocks and specifically newly sequestered soil C (Dondini et al., 2009), C losses and gains from roots, shoots and leaves individually (Foereid et al., 2004; Amougou et al., 2011).
nutrient remobilisation and associated N-fixation (Eckert et al., 2001; Cadoux et al., 2012; Keymer and Kent, 2013) and the increased uncertainty when scaling up simulations to larger spatial scales (Pogson, 2011). A number of applicable models have recently been reviewed by Surendran Nair et al. (2012), providing detailed model descriptions of how the models simulate biomass production and soil water, nutrient and C cycle dynamics for bioenergy crops in general. However, in this review, we focus the discussion only on models that have been parameterised and validated for Miscanthus simulations and specifically what measurements are required to improve model performance regarding soil C aspects of the C cycle. In addition to two models (WIMOVAC and Agro-IBIS) also discussed by Surendran Nair et al. (2012), we review four other models suitable for Miscanthus C studies (BioCro, DayCent, DNDC and ECOSSE). Further, we briefly discuss the current databases available for use with the models described and possible frameworks that may encourage and accelerate model development.

Identifying existing models parameterised for Miscanthus

A literature search was performed to identify existing C budget models that fulfilled all of five criteria. Models were required to be: 1) mechanistic in design, 2) parameterised and validated for Miscanthus plantations, 3) capable of both predicting crop yields and soil C dynamics, 4) published in a peer-reviewed journal or conference proceedings and 5) report outputs validated against field data. Mechanistic models were specifically chosen to allow greater flexibility when simulating the impacts of future climatic scenarios and changing environmental conditions. Similarly, mechanistic models are more transferable to the larger geo-spatial scales that policy decisions are often created for. Consequently, mechanistic models can play an important role in deciding whether Miscanthus is an appropriate crop for
geographically distinct regions (i.e. their climate and soil properties). Additionally, the ability for the models to simulate both yields and soil C dynamics was required because assuring the commercial viability of a Miscanthus plantation, and assessing its impact on net C emissions, are essential parts of a landowner choosing to establish Miscanthus over a more conventional crop.

We identified five crop growth models parameterised for Miscanthus: WIMOVAC (Miguez et al., 2009), BioCro (Miguez et al., 2012), Agro-IBIS (VanLoocke et al., 2010), DayCent (Davis et al., 2010) and DNDC (Gopalakrishnan et al., 2012). In addition, the ECOSSE model (Smith et al., 2010a) is included as it is currently being adapted to simulate C budgets in a Miscanthus plantation (Jones et al., 2011). Of these models, WIMOVAC, BioCro and Agro-IBIS were originally created to simulate biomass production but have more recently had soil biochemistry and soil C incorporated in their simulations. Conversely, DayCent, DNDC and ECOSSE were all principally designed to simulate belowground nutrient cycling and have only included more complex plant growth routines recently.

**Considerations for Miscanthus model selection**

There are many factors to consider regarding model selection and here we highlight the key considerations for modelling Miscanthus C dynamics. We group these considerations into four categories: 1) model parameters, 2) scale, 3) inputs and 4) verification of model outputs. The models presented vary greatly in their approach to simulate the same ultimate outputs. For example, DayCent uses a continuous phonological development curve to allocate C assimilated into five pools of biomass (Davis et al., 2010), whereas Agro-IBIS applies three discrete development stages to calculate crop growth and partitions the C into three pools of
biomass (VanLoocke et al., 2010). Both approaches are valid but there are advantages and disadvantages to each which are discussed during Section 2 of this review. The general characteristics of each of the six models reviewed here are summarised in Table 1.

Model parameters

The research objective for a model may be to quantify a set variable but it is possible that the chosen model a) lacks the desired output (i.e. change in C stocks) or b) does not describe the output at the soil depth and/or timescale required (i.e. monthly change to the active soil C fraction/change in C at a certain soil depth increment). Therefore, early consideration of the models’ functionality is required to match expectations for use. As an example, WIMOVAC would be unsuitable to evaluate temporal variation between soil CO$_2$ emissions and soil C content as the model calculates C losses by assuming a fixed fraction is lost and, relative to total C accumulation, C losses will always be the same.

It is worth noting that the model parameters that determine many aspects of C transfer within the system (e.g. C assimilation, growth rates of the crop and SOM turnover rates) may change depending on the genetic variation of the Miscanthus species. The models discussed here have been parameterised and verified for Miscanthus giganteus and/or Miscanthus sinesis, but a number of genetic variants are being trialled in the UK to improve biomass production in less-than-optimal climatic conditions (Clifton-Brown et al., 2008; Robson et al., 2011).

Studies indicate that it is possible that different genetic variants may respond differently to abiotic stresses (Borzęcka-Walker et al., 2008). Similarly, it is plausible that litter quantity and quality will vary, influencing C inputs to the soil. Consequently, to accurately simulate
crop growth and associated ecosystem C dynamics from these new variants, the existing models may need to be re-parameterised when sufficient field data is available.

**Inputs**

As well as considering the desired outputs for a model it is also important to ensure that the required data inputs to the model are readily available. All of the six models reviewed require inputs of site soil properties (Table 2) meaning if this data is unavailable for the desired resolution (see section 1.2.3. Scale) the model outputs may not be reliable. Similarly, five of the six models are able to interpolate daily inputs of known site characteristics when additional data are unavailable, whilst DayCent can only interpolate monthly inputs. Consequently, certain models will be more appropriate than others depending on the research question. The input data available to a modeller is an important consideration when choosing a model just as is the importance of ensuring any new data collection meets the demands of the chosen model. Therefore, knowing the input requirements of relevant models is an essential part of selecting the most appropriate model. Table 2 lists both the essential and optional inputs of each model, therefore simplifying the model comparison when choosing. For a model to operate to the highest attainable degree of accuracy, all required and optional inputs are suggested.

**Scale**

If a model has been created to operate at a certain spatial scale, the assumptions made, and conclusions that can be drawn, may not be valid at different spatial scales. Similarly, if high resolution temporal outputs are required, some models are able to interpolate this data while
others are not. All of the models discussed in this review operate at high temporal resolutions
(daily time step or more frequently) but there is a greater variation among them with regards
to the geo-spatial scales they were originally designed to operate at (see Table 1). Many of
the models require site-specific conditions as inputs to drive the simulations, meaning that
using information from a single site to represent a larger area will likely be an oversimplified
approximation, resulting in unreliable projections. However, recent studies that attempt to
upscale the models to large geo-spatial regions indicate reliable simulations can be driven
with current databases. For example, MISCANFOR (Hastings et al., 2009a) - a mechanistic
model designed to simulate Miscanthus biomass production - has been successfully used to
simulate Miscanthus growth over 25 km grids using a number of different datasets for
meteorological inputs and soil characteristics (Hastings et al., 2009b; Pogson, 2011; Pogson
et al., 2012; Pogson et al., 2013). The model outputs were then related to geographic
information system (GIS) maps at European and global scales with good agreement against
empirical field data.

The MISCANFOR model uses many of the same inputs of soil and meteorological data that
the six models reviewed in this paper do, therefore indicating that each may cope with
upscaling with similarly accurate outputs using the same datasets. That said, at present there
is no robust dataset for soil C stocks at high spatial resolution and without this dataset models
can only provide part of the ecosystem C budget. It is worth noting that generally, the fewer
inputs required to drive the model, the easier the model will be to scale up to represent larger
areas due to the limited number of databases available, but the less mechanistic the model,
the less transferable its simulations are over large geo-spatial regions. Therefore, when
applying a model to large spatial scales, some trade off will always occur between the
datasets available and the inputs required by a model to achieve the highest accuracy
simulations. Most of the six models reviewed here have only been validated extensively at site scales. However Agro-IBIS is a dedicated dynamic global vegetation model (DGVM), and therefore is likely to be the most straightforward if the simulation of LUC, to and from other plant functional types, is required over ecosystem scales.

Verification of model outputs

At the field or plot scale, verifying a model’s reliability requires independent experimental verification data from the same location (Smith & Smith, 2007). However, since one of the main purposes of modelling is to upscale beyond measurement capacities (space and time), model verification through field data is often not possible. Therefore it is important to ensure that model outputs have been verified within a range of conditions to reflect the sites and land uses under consideration. Although the models reported in this review have been validated for various site conditions, there may be additional variables not included, or those described ineffectively. For example, Davis et al. (2010) found that parameters controlling the DayCent model’s N-cycling simulation were insufficient for simulating Miscanthus growth at sites in Illinois, USA. Further experimentation concluded that the plant, or microbial symbioses, were fixing N which was, in turn, influencing yields. Consequently the DayCent model was calibrated to describe N-fixation differently, accounting for the experiment’s findings. Overall, since Miscanthus propagation is relatively recent, there is likely to still be much we do not know, particularly with regards to the retranslocation of nutrients after the crop’s growth phase (Beale and Long, 1997; Beuch et al., 2000) and potential N-fixation from the crop’s association with N-fixing organisms (Keymar and Kent, 2013). This emphasises the importance of researchers collecting targeted field data that can be used to accelerate model development and validation (Zhang et al., 2010; Surendran Nair et al., 2012).
Models parameterised for Miscanthus

The six models identified share a number of similar internal components and represent many of the same ecosystem processes, albeit in different ways or with different levels of complexity (Table 1). For example, all of the models simulate plant growth, but only DayCent calculates Net Primary Productivity (NPP). The other models estimate C assimilation based on solar interception and Miscanthus-specific photosynthetic efficiency. Although the outputs from the models are similar, there are differences in how the C dynamics are simulated and allocated to plant biomass or soil C pools. In this section we give an overview of the six Miscanthus models and suggest the field data most relevant for improving model performance. Much of the discussion of model improvement is focussed on validating predictions of changes to soil C, nutrient translocation and N-fixation — the aspects of Miscanthus modelling where increased empirical data and improved understanding will have the greatest benefit on future simulations.

Selected model descriptions

WIMOVAC and BioCro

First created as a generic crop growth model (Humphries and Long, 1995), WIMOVAC was parameterised for a Miscanthus plantation in England, UK and shown to realistically simulate biomass production at a number of discrete sites with varying climatic conditions across Europe (Miguez et al., 2009). More recently, a new version of WIMOVAC named BioCro has been developed; this is written in the programming language C, with an interface to work with it using the statistical package R. BioCro has since been used to simulate Miscanthus
biomass production across the contiguous United States (Miguez et al., 2012). Both versions represent biophysicochemical processes in detail and operate at an hourly time step, with some processes and state variables updated sub-hourly (e.g. C assimilation) (Humphries and Long, 1995). Parameterised for Miscanthus, the models follow a phenological growth routine comprising six development stages in accordance with the typical growth phases of grasses reported by Cao and Moss (1997): emergence, juvenile, induction, post-induction, flowering and post-flowering (Miguez et al., 2009). Progression through these stages is controlled by thermal time, or degree days (DD) and the growing season is defined as being from the last frost of the spring to the first frost of the autumn in accordance with Price et al. (2004).

Both models are calibrated for C4 photosynthesis and use the biochemical approach of Collatz et al. (1992) for simulating the photosynthetic potential of the crop. Here, they calculate gross photosynthesis as a function of leaf temperature, intercellular CO₂ partial pressure, and incident solar radiation. WIMOVAC and BioCro define the light extinction coefficient (k) using the sunlit and shade classes defined by Norman (1980) but applying an additional parameter to describe the ellipsoid arrangement of leaves (Miguez et al., 2009; Miguez et al., 2012). Further, the models divide the canopy into ten layers and compute the radiative conditions for each. The canopy function in BioCro can also be used to obtain information about assimilation, transpiration and conductance at different canopy levels. Biomass production is then simulated and allocated to four pools (leaf, stem, rhizome and root) according to fixed partitioning parameters specific to the phenological stage. Although WIMOVAC and BioCro are similar in many ways, biomass partitioning coefficients are handled differently, with BioCro capable of applying negative coefficients for all biomass pools. Additionally, one specific difference concerns the magnitude of C translocation away from the rhizome during the emergence and juvenile stages: WIMOVAC uses coefficients of
-0.1 and -0.08, respectively, whereas BioCro uses -0.0008 and -0.0005 for the earliest two stages, respectively. Further, BioCro has been modified to ensure positive and negative allocation coefficients sum to zero. Both models also separate the respiration associated with crop growth from respiration associated with crop maintenance, using a constant related to the plant structure (Miguez et al., 2009).

In addition to C allocation within the crop biomass, C dynamics are also simulated for soil pools. The models separate soil C into active, slow and passive pools according to those of CENTURY (Parton et al., 1993). C losses from gaseous emissions and leaching are also accounted for by both models (Humphries and Long, 1995; Miguez et al., 2012). The models only account for mineralisation and immobilisation of nitrogen (N) thereby providing necessary information about N availability, but not a full account of the N budget. Similarly, soil water routines and the hydrological sub-model include only variables for the most important processes (Table 1). Since Miscanthus biomass production has been shown to be influenced greatly by water availability (Clifton-Brown and Lewandowski, 2000; Heaton et al., 2004; Richter et al., 2008; Oliver et al., 2009), it is worth noting that both WIMOVAC and BioCro use an empirical water stress response function based on that of Boyer (1970) but also accounting for the system’s energy balance and present growth phase. This function reduces stomatal conductance through a linear relationship to leaf water potential and soil moisture content, then alters biomass partitioning to roots when the Miscanthus-specific average daily plant water potential is below a fixed threshold value (Long et al., 1998). However, neither of the models account for stem death during periods of extended water stress (Miguez et al., 2009).
WIMOVAC and BioCro are very detailed mechanistic models, defining biophysiochemical interactions by the underlying processes. This makes both models advantageous in their transferability since the underlying processes do not change over spatial or temporal scales but disadvantageous when noting the substantial amount of data required to initialise the model for a specific site (Miguez et al., 2012). As a result, both models rely heavily on data published in only a few studies (e.g. Beale and Long, 1995; Beale et al., 1996; Naidu et al., 2003). Consequently, to improve WIMOVAC or BioCro the most beneficial data would be that which could verify model outputs under environmental conditions discrete to those already reported. Intra-annual measurements of CO₂ uptake rates and leaf area index (LAI), along with biomass accumulation in the four pools represented by the models would ensure validation of the key parameters influencing C assimilation as well as verification of the model’s outputs for biomass production and partitioning. A greater understanding of this partitioning can help us to know when is the right time to harvest aboveground biomass, maximising the landowner’s profits for a given year and acknowledging that inter-annual variability can influence Miscanthus yields by 10-25% (Price et al., 2004; Christian et al., 2008). Typically, Miscanthus is harvested in early spring as the combustion quality of the biomass is higher and specifically the moisture content of the biomass is lower (Lewandowski and Kicherer, 1997; El Bassam and Huisman, 2001; Lewandowski and Heinz, 2003; Lewandowski et al., 2003a). Unfortunately, since neither WIMOVAC nor BioCro currently simulate moisture content of the crop biomass, they cannot be used to predict the optimum harvest time with regards to moisture content. However, using a sorption model, such as the modified Oswin model used by Arabhosseini et al. (2010), coupled with the air temperature and relative humidity parameters already included within WIMOVAC or BioCro, moisture content for Miscanthus could be estimated depending on the environmental conditions of a specific site for a specific year. Unlike the issue of biomass moisture content,
quantifying *Miscanthus*’ unique properties regarding nutrient retranslocation and N-fixation
is not so easily resolved within the models. That said, both models have had multisite
validation and been verified against a number of sites across the USA, indicating that the
current parameterisation and calibration of the models is at least accurate within the range of
environmental conditions present in this region. However, empirical data regarding soil C
stocks are still scarce and therefore to date there is no literature citing WIMOVAC or BioCro
being used to simulate changes in soil C from sites discrete to those used to parameterise or
 initialise the models.

*Agro-IBIS*

Agro-IBIS is a DGVM and unlike the other models in this review, is calibrated to use grid-based simulation techniques to function at large spatial scales. The original model (Foley *et al*., 1996) was designed to apply the same agroecological rationales from models operating at a site scale but draw conclusions about the growth and management of crops at scales often required for policy making (Kucharik, 2003; Kucharik and Byre, 2003). VanLoocke *et al*. (2010) parameterised Agro-IBIS for *Miscanthus* and reported both the default and new values for any altered internal parameters. The model does not employ *Miscanthus*-specific phenological development stages but rather the model retains three stages that have been used successfully to simulate maize biomass production: budburst, senescence and dormancy. Although this is a generalisation of the phenology for the given plant functional type, many aspects of biomass production are similar for maize and *Miscanthus* (Dohleman and Long, 2009), and Agro-IBIS adjusts parameters to affect the timing of these events accordingly (VanLoocke *et al*., 2010).
The Miscanthus version of Agro-IBIS uses the specific leaf area (SLA) approach (Adam et al., 2011) to estimate total leaf area for the crop, assigning a constant value according to Dohleman and Long (2009). The light extinction coefficient ($k$) is then estimated according to incident radiation, surface albedo, leaf area and leaf orientation (Kucharik and Byre, 2003; Cuadra et al., 2012). Further, the incidence of diffuse and direct radiation is determined at canopy level for near-IR and visible wavelengths (Foley et al., 1996; Kucharik et al., 2000). Gross C assimilation is then calculated using the approaches described by Farquhar et al. (1980) and later simplified and adapted for C4 photosynthesis by Collatz et al. (1992). C is then partitioned into stem, root and leaf pools according to variable partitioning parameters within the model: initially, 80%, 10% and 10% of the C assimilated is allocated to the leaf, stem and root pools, respectively, but by the end of the growing season this changes to 10%, 80% and 10%, respectively (VanLoocke et al., 2010).

Agro-IBIS contains two major sub-models besides those to simulate land-atmosphere interactions and vegetation dynamics: a belowground C and N module and a solute transfer module. The model represents soil C in three discrete pools and includes both gaseous C losses and those from leaching. These are coupled with N cycling, including four processes to define N availability (Table 1). Soil water availability is controlled by a number of factors including canopy interception, surface runoff and evapotranspiration. Each is calculated for the area based on climate inputs and soil properties with the largest spatially explicit dataset being the limiting factor (Kucharik et al., 2000). The model allows for stress from leaf temperature extremes and limited water availability; when given thresholds are exceeded, gross C assimilation and stomatal conductance are modified to simulate the impact of that stress (Mu et al., 2007).
As a DGVM, Agro-IBIS was designed to operate at large spatial scales to predict the impacts of LUC on C budgets but, to date, published simulations are confined to sites in Illinois, USA where good agreement between simulated and observed values for leaf photosynthesis, LAI and latent heat flux were observed (VanLoocke et al., 2010). Although the Miscanthus-specific values used for parameterising Agro-IBIS have been validated successfully in Illinois, additional field data from a wider range of environmental conditions would ensure that the parameter values used are universal when Miscanthus is grown in other locations. Datasets that report LAI, SLA, maximum rubisco activity and measures of maturity in DD would be the most beneficial to verify accompanying yield data given these are key drivers of simulated biomass production. That said, as long as conditions of simulated sites are within the limits of those found at the validated sites in Illinois, the new predictions made by Agro-IBIS are likely to have similar degrees of uncertainty as those reported by VanLoocke et al., 2010. Published literature (e.g. Kucharik, 2003; Twine and Kucharik, 2009; Sacks and Kucharik, 2011) for Agro-IBIS has focussed on evaluating its accuracy at simulating harvestable biomass production and a lack of field data has limited verifying model predictions of belowground biomass and soil C sequestration. DGVMs and next generation earth system models such as JULES (Alton et al., 2007, Hughes et al., 2010) will play an increasingly important part in deciding whether a crop is appropriate for a location, as policy decisions are often based on regions or whole countries, and not individual sites.

DayCent

DayCent is an adaptation of the comprehensive ecosystem model, CENTURY, but with a daily timestep. The DayCent version was originally reported by Parton et al. (1998) as a way of producing more reliable simulations of GHG fluxes because they respond rapidly to
abiotic factors, and since has been applied to different native and managed systems encompassing a wide range of environmental conditions (Del Grosso et al., 2005; Del Grosso et al., 2008; Gathany and Burke, 2012). The DayCent model has been calibrated and parameterised for Miscanthus, altering N-fixation routines according to the findings by Davis et al. (2010). Model outputs from this version of DayCent were verified against measurements in the work. The model does not represent the phenological development stages specific to Miscanthus but rather uses a growth response routine according to abiotic factors and water/nutrient availability (Del Grosso et al., 2001). This allows DayCent to generate accurate simulations without needing additional inputs about the processes involved in C assimilation.

NPP is estimated by DayCent according to species-specific relationships of plant growth to soil and air temperature, soil water availability and nutrient availability (Del Grosso et al., 2001). For Miscanthus, Davis et al. (2010) parameterised DayCent using empirical data from Europe and Illinois, USA, (Beale and Long, 1995; Clifton-Brown and Lewandowski, 2000; Heaton et al., 2004; Cosentino et al., 2007; Heaton et al., 2008) to form plant growth parameters. This resulted in simulations being possible in alternate locations where only data for climate and basic soil properties are available (Table 2). Net growth is then partitioned into five plant components (leaves, branches, large wood, fine roots and large roots) as a function of soil water and soil nutrient functions with the most limiting factor (temperature, water, nutrients) having a direct constraint on biomass production (Parton et al., 1993). By avoiding measurements of radiation use efficiency (RUE), light extinction coefficient ($k$) and other factors controlling C assimilation, DayCent reduces the number of potentially sensitive parameters that can influence net growth. This simplicity can be considered an advantage over more mechanistic models and because the processes that influence water, nutrient and
Separate from the plant production sub-model, DayCent features routines to describe coupled soil water and soil temperature modules, plant decomposition, SOM and trace gas emissions (Table 1). Modelled processes and outputs of DayCent have been validated and verified a number of times since its inception, but Davis et al. (2010) reported significant findings regarding additional N-fixation, required when the model was parameterised for Miscanthus. These results highlight the importance of a Miscanthus model including aspects of the N cycle to ensure the annual demands of the plant are simulated accurately. Following biomass production in the five plant biomass pools, DayCent predicts the amount and quality of plant residue supplemented to the surface and soil, also simulating the plant’s influence on the soil environment (Lee et al., 2012). This makes DayCent the most comprehensive of the models described here when it comes to C dynamics representing plant and soil interactions. Within the same simulation plant biomass decomposition is calculated, adding C to and transferring C between three conceptual pools in the SOM module: active, slow and passive. Each of these represent different turnover times of the SOM ranging from months and years to centuries and millennia. In addition to these represented C dynamics, nitrogen, phosphorus and sulphur exchange is also cycled through the model, accounting for a number of key processes (Del Grosso et al., 2001; Table 1).

Since CENTURY was developed in the late 1980s, the model or parts of its structure have been used frequently to simulate C and N dynamics, but only recently has the model been parameterised for Miscanthus. The only published study validating the Miscanthus iteration of DayCent is contained within the inaugural paper and is constrained to field data from
Illinois, USA, due to the article’s scope (Davis et al., 2010). However, employing the framework and parameter values used by Davis et al. (2010), there is no reason why model outputs could not be verified for different plantations at numerous other sites (e.g., those reported in Clifton-Brown et al., 2001a; Danalatos et al., 2007; Christian et al., 2008). Data mining from these studies could prove to be very useful when considering model development, and although the framework of DayCent has been validated numerous times, there is still room for improvement regarding Miscanthus simulations (Davis et al., 2010): the models were not initially designed to simulate the nutrient retranslocation from aboveground to belowground plant biomass (e.g. Heaton et al., 2009) that makes Miscanthus a particularly attractive bioenergy crop. Similarly, Davis et al. (2010) estimated that Miscanthus can host N-fixing organisms that can provide up to 25 g N m$^{-2}$ each year - a significant amount for any given land use (Stewart, 1975). Having empirical data to validate this estimate, and the degree of retranslocation, such as the studies by Keymar and Kent (2013) and Heckathorn and DeLucia (1994), respectively, can help model development to ensure simulations are accurate based on specific site conditions.

**DNDC**

Originally designed to simulate trace gas emissions and soil C and N dynamics (Li et al., 1992), the DNDC model was later calibrated to represent crop growth routines as well as the soils they grow on (named Crop-DNDC; Zhang et al., 2002). More recently, DNDC was parameterised for Miscanthus (Borzecka-Walker et al., 2012; Gopalakrishnan et al., 2012). Most aspects of the DNDC model run a daily time step but due to considerable diurnal variation the soil climate and denitrification sub-models operate at hourly time steps. All adaptations of the DNDC model include a plant growth module but the more basic versions
simply calculate biomass accumulation according to a generalised crop growth curve using thermal time units. However, the versions parameterised for *Miscanthus* use the mechanistic Crop-DNDC routines allowing crop growth to respond to climatic conditions and soil biogeochemistry. Crop-DNDC simulates crop growth using nine phenological development stages based on those included in CERES models (Ritchie, 1991). Although the original Crop-DNDC structure was applicable for *Miscanthus*, Gopalakrishnan *et al.* (2012) noted the conclusions drawn about N-fixation from Davis *et al.* (2010) and calibrated the DNDC model accordingly, changing the model’s default N fixation index from 1 to 3. It is worth noting that within the model, the fixation index is intended to represent N-fixation of the crop directly and not associated organisms, although successful use of this model parameter suggest it may at least help provide accurate yield simulations of *Miscanthus* (Borzecka-Walker *et al.*, 2012; Gopalakrishnan *et al.*, 2012).

Crop-DNDC calculates gross photosynthesis using LAI according to Spitters (1986) and gross crop respiration according to McCree (1979) and Penning de Vries *et al.* (1989). The respiration is then subtracted from gross photosynthesis to estimate net C assimilation available for growth. Atmospheric CO₂ concentration, air temperature and the canopy profile are all considered when calculating photosynthetic rates, as are water and nitrogen stress factors (Zhang *et al.*, 2002). C is then allocated to stem, leaf, grain and root state variables based on the phonological stage, with more C allocated to stems later in the growing season. Although not initially designed for accurate estimates of crop growth, the adaptations to DNDC now allow accurate simulations after parameterisation using the model’s ‘Crop Creator’ module. Using this module, Gopalakrishnan *et al.* (2012) was able to apply *Miscanthus*-specific parameters to achieve good agreement of model outputs with measured field data. The changed parameters and values used are reported in their paper.
The DNDC model comprises of six sub-models to simulate crop growth, soil climate, nitrification, denitrification, decomposition and fermentation. Each of these sub-models are coupled so each effects the other and specifically allowing C and N dynamics to be balanced within the system. Although DNDC has most regularly been used to simulate trace gas emissions (e.g. Cai et al., 2003; Levy et al., 2007; Abdalla et al., 2010), the processes governing C dynamics into and between the state variables are equally detailed. Stems, leaves and roots senesced from the crop growth sub-model enter the decomposition sub-model and are allocated to one of three pools: very labile litter, labile litter and resistant litter. The C is moved to microbial and then humad pools before eventually reaching the passive humus pool. For each of these transfers, specific decomposition rates are applied and at each step trace gas emissions are also calculated (Li et al., 1994; Li, 2000).

The DNDC model has been used frequently since its inception but only recently has the model been parameterised and tested to simulate the C and N dynamics of a Miscanthus plantation. There are two publications reporting DNDC used to simulate Miscanthus growth but both focus mainly on the nitrogen losses and trace gas emissions associated (Borzecka-Walker et al., 2012; Gopalakrishnan et al., 2012). Consequently it is difficult to review the model performance for ecosystem C budgets, but both papers report good agreement between modelled and measured values of crop yield. Confidence in this agreement can also be drawn from the different locations simulated; the Gopalakrishnan et al. (2012) paper simulated yields in Illinois, USA, whereas Borzecka-Walker et al. (2012) simulated Miscanthus plantations in Poland, where annual precipitation was roughly half that of the sites in Illinois. The DNDC model represents soil processes mechanistically and in a lot of depth, and is therefore particularly good at improving our understanding of how Miscanthus plantations
may influence key issues regarding the crop’s sustainability criteria (e.g. net change in soil C, GHG emissions and N dynamics). Each of these were simulated and reported by Borzecka-Walker et al. (2012) with interesting results that suggest soil type has a significant impact on potential soil C storage, net global warming potential and soil N balance, but not on yield. However, these simulations were only valid for the site modelled as there is not sufficient field data to initialise the model at different Miscanthus sites. The model requires inputs of each of the SOM pools in kg C kg$^{-1}$ soil and few datasets include this information; soil fractionation to relate measured soil C fractions to conceptual modelled pools is a relatively recent aspect of this research. However as fractionation data becomes more readily available, DNDC’s SOM module outputs can be validated for a range of sites and conditions with added confidence given the successful simulations performed by Borzecka-Walker et al. (2012). An additional aspect of the DNDC model that makes it particularly attractive is its integration with GIS databases (e.g. Pathak et al., 2005; Tang et al., 2006). This allows many of the model outputs to be scaled up across larger regions as long as there is sufficient high quality input data to drive the simulations.

**ECOSSE**

The ECOSSE model (Smith et al., 2010a) was developed to simulate highly organic soils from concepts originally derived for mineral soils in the RothC (Jenkinson and Rayner, 1977; Jenkinson et al. 1987; Coleman and Jenkinson, 1996) and SUNDIAL (Bradbury et al. 1993; Smith et al. 1996) models. Following these established models, ECOSSE is primarily designed to simulate belowground C and N dynamics but has more recently been coupled with an updated version of the MIAMI model (Leith, 1972) to calculate NPP. ECOSSE uses a pool type approach, describing soil organic matter (SOM) as pools of inert organic matter
(IOM), humus (HUM), biomass (BIO), resistant plant material (RPM) and decomposable plant material (DPM). All of the major processes of C and N turnover in the soil are included in the model, but each of the processes is simulated using only simple equations driven by readily available input variables, allowing it to be developed from a field based model to a national scale tool, without high loss of accuracy. ECOSSE differs from RothC and SUNDIAL in the addition of descriptions of a number of processes and impacts that are important in organic soils, but not relevant in the mineral arable soils that these models were originally developed for. More importantly, ECOSSE differs from RothC and SUNDIAL in the way that it makes full use of the limited information that is available to run models at national scale. In particular, measurements of soil C are used to interpolate the activity of the SOM and the plant inputs needed to achieve those measurements. Any data available describing soil water, plant inputs, nutrient applications and timing of management operations are used to drive the model and so better apportion the factors determining the interpolated activity of the SOM. However, if any of this information is missing, the model can still provide accurate simulations of SOM turnover, although the impact of changes in conditions will be estimated with less accuracy due to the reduced detail of the inputs (Smith et al., 2010b).

The total plant inputs of C are assumed to be given by the NPP. If this is known, it can be entered as an input by the user, otherwise, NPP is entered as zero and the plant inputs are estimated using the MIAMI model (Leith, 1972). The N inputs from the plant to the soil are calculated using standard C:N ratios for the different land uses. The C:N ratios of simple land use classes, such as arable, grassland, forestry and semi-natural, are initially all set to 10. As an alternative, in the site specific version of the model, the plant inputs of C and N can be calculated from the expected yield as described by Bradbury et al. (1993). The plant input of
C is given as a function of the crop yield modified by empirical parameters, specific to each crop or plant type. The plant input of N is calculated from the amount of N taken up in above ground plant biomass.

The rates of decomposition modelled by ECOSSE are important determinants in the modelling of soil N and N$_2$O emissions, as the soil N content follows the decomposition of SOM (Bell et al., 2012). The rate of SOM decomposition in the model is modified by temperature, soil water content, plant cover and soil pH. Decomposition rate is also dependent on how SOM is proportioned into the different SOM pools. Following the approach used in the RothC model (Coleman and Jenkinson, 1996), the IOM pool does not undergo decomposition; the C in this pool does not take part in soil processes either due to its inert chemical composition or its protected physical state. The HUM pool decomposes slowly, representing material that has undergone stabilization due to earlier decomposition processes. The BIO pool decomposes more rapidly and represents material that has undergone some decomposition but is still biologically active. The DPM and RPM pools are composed of undecomposed plant material, the DPM pool being readily decomposable while the RPM pool is more recalcitrant. The ratio of DPM to RPM defines the decomposability of the plant material that is added to the soil. Values for the ratio of DPM to RPM for the different land uses are standard as used in RothC, although these can be changed within ECOSSE for a specific instance of a land use type (e.g. Miscanthus).

In ECOSSE, soil layers are divided into 5 cm layers. Each layer is filled with water until it reaches field capacity, when it then drains to the layer below, or evaporates from the uppermost layer. ECOSSE differs from many other soil models, as it is able to simulate how decomposition will respond to soil water contents above field capacity. The model assumes
that decomposition is at its maximum at field capacity, but is slowed by water limitation
below field capacity, and oxygen limitation above field capacity. When the water content is
above field capacity the decomposition rate falls linearly, and at saturation it is only 20% of
that at field capacity (Smith et al., 2010a). Under aerobic conditions, the decomposition
process results in gaseous losses of CO₂; under anaerobic conditions losses as methane (CH₄)
dominate. The N content of the soil follows the decomposition of the SOM, with a stable C:N
ratio defined for each pool at a given pH, and N being either mineralised or immobilised to
maintain that ratio. Nitrogen released from decomposing SOM as ammonium (NH₄⁺) or
added to the soil may be nitrified to nitrate (NO₃⁻). C and N may be lost from the soil by the
processes of leaching (NO₃⁻, dissolved organic C (DOC), and dissolved organic N (DON)),
denitrification, volatilisation or crop offtake, or C and N may be returned to the soil by plant
inputs, inorganic fertilizers, atmospheric deposition or organic amendments.

The ECOSSE model has already been validated and applied spatially to simulate land-use
change impacts on SOC and GHG emissions (Smith et al., 2010a,b) and to simulate soil N
and N₂O emissions in cropland sites in Europe (Bell et al., 2012). Smith et al. (2010a,b)
reported the estimate in Scottish soil C stocks and changes using ECOSSE. The results of this
work reported that, despite the uncertainties in the input data and the measurements used to
evaluate the model, the simulated values show a high degree of association with the
measurements in both total C and change in C content of the soil. Over all sites where land-
use change occurred, the average deviation between the simulated and measured values of
percentage change in soil C was less than the experimental error (11% simulation error, 53%
measurement error). This suggests that the uncertainty in using this model for the national-
scale simulations will be ~11%. Bell et al. (2012) reported the first test of the ECOSSE
model at predicting N₂O emissions from arable soils in Europe, indicating that although
further modifications are required in the form of predictions on a daily time-step, the model is currently predicting such fluxes with a greater degree of accuracy than other available methods of quantification which can then be used to estimate emissions on a large scale. The parameterisation and evaluation of ECOSSE to simulate soil C and GHG emissions under *Miscanthus* and short-rotation forestry is currently on-going and the results are in preparation for publication.

**Alternatives to mechanistic crop models**

In addition to the models that fulfil our original selection criteria there are other relevant models which satisfy a subset of our criteria. Alternatives include empirical models (Heaton *et al.*., 2004; Richter *et al.*., 2008), soil-only models (Dondini *et al.*., 2009), plant growth only models with no soil C components (Clifton-Brown *et al.*., 2000; Hastings *et al.*., 2009a) and those where no peer-reviewed literature exists supporting the models’ validities for *Miscanthus* specifically (e.g. EPIC (Williams *et al.*., 1989) reported by Zhang *et al.* (2011) and JULES (Hughes *et al.*., 2010).

Empirical models are excellent simple predictors when estimating how a known set of conditions will influence a specific variable, and are likely to provide more accurate predictions than mechanistic alternatives, since they are built solely around field observations and not the biogeochemical processes that define plant growth or C transfer. However, this also means empirical models require a large amount of field data from individual sites, and their simulations cannot be used to predict the system’s response to environmental variation. Heaton *et al.* (2004) and Richter *et al.* (2008) used observed *Miscanthus* yields and accompanying site conditions (i.e., soil type, growing degree days, average precipitation) to
create empirical models capable of simulating different scenarios of available N, air
temperature and water availability. Their model outputs correlated well with observed values
and using known site conditions across larger spatial areas each was able to predict potential
yields and how they would be affected by changes in N-availability, air temperature and
water availability. While informative for this purpose, the models are less suitable for
understanding why yields vary. Furthermore, creating similar empirical models to simulate
the response of other variables of interest (e.g. soil C) is unlikely to be successful, due to the
number of factors that drive change in those variables and the lack of field data for
Miscanthus.

Soil-only models that describe the belowground C dynamics of Miscanthus plantations can
predict changes to soil C stocks over time but by definition do not include any plant growth
routines and so lack information when considering the total ecosystem C budgets. Primarily,
RothC (Coleman and Jenkinson, 1999) has been the chosen model to simulate soil C
dynamics beneath Miscanthus plantations (Dondini et al., 2009; Hillier et al., 2009; Poeplau
and Don, 2013). This process-based model is similar to the SOC sub-models included in
DayCent, WIMOVAC, BioCro and Agro-IBIS, and like these models, their outputs come
with considerable uncertainty due to limited validation data for Miscanthus plantations under
a wide range of environmental conditions. That said, the work done by Dondini et al. (2009)
and Poeplau and Don (2013) is taking great steps to help reduce the associated uncertainty
and continued investigation into the size of measurable soil C pools using fractionation
techniques will provide the information regarding initial pool sizes that is essential for
accurate simulations of soil C dynamics.
Just as soil-only models do not include plant-growth routines, models like MISCANMOD (Clifton-Brown et al., 2000) and MISCANFOR (Hastings et al., 2009a) do not include sub-models to describe soil C dynamics. Both MISCANMOD and MISCANFOR have been used to simulate Miscanthus productivity across the UK and Europe using a number of databases for required inputs (Clifton-Brown et al., 2004; Hastings et al., 2009b; Pogson et al., 2012). More recently, MISCANFOR was also used to predict Miscanthus yields globally and relate them to potential energy generation accounting for land use, cost and carbon restrictions (Pogson et al., 2013). The crop growth routines used in these Miscanthus-specific models have regularly been proven to be accurate within a wide range of environmental conditions but the model's lack of soil C pools make it of limited use when requiring information about the full ecosystem C budget. Consequently, coupling such models with those that describe other aspects of the C cycle may allow for accurate simulations without developing a whole new model.

The EPIC model (Williams et al., 1989) is a mechanistic model that is reported to have been parameterised for Miscanthus and showed a good correlation between field data and model output (Zhang et al., 2011). However, no further detail is given by the Zhang et al. (2011) paper. Similarly the JULES model (Met Office, 2013) was used by Hughes et al. (2010) to simulate Miscanthus production and calculate payback times for different regions globally, based on the assumption that 50% of NPP is available to offset fossil fuel emissions. The JULES model is essentially a land-atmosphere energy transfer model but it includes TRIFFID, a DGVM much like Agro-IBIS. Although the Hughes et al. (2010) study reports parameterisation of the JULES simulation, and the values used, there is no discussion of its validation against field data. Therefore the reliability of its predictions cannot be assessed or the uncertainty quantified. That said, the study integrates the JULES model with the global
climate impacts model, IMOGEN, to simulate the impact of IPCC Special Report Emissions Scenarios (Hughes et al., 2010). After validation from field data encompassing a range of environmental conditions, this is the right approach to predicting how beneficial Miscanthus plantations may be in different regions.

A simple semi-mechanistic model created to estimate C sequestration rates under Miscanthus plantations was reported by Matthews and Grogan (2001) over a decade ago but has not since been further developed. The model by Matthews and Grogan (2001) uses solar radiation, the light extinction coefficient ($k$) and LAI to estimate biomass production and partitions this according to the values reported by Himken et al. (1997). In this way it is very similar to a number of mechanistic models that have been developed for bioenergy crops since (see Surendran Nair et al., 2012). Each of the plant and soil values required to parameterise Matthews and Grogan's model were taken from published field data and decomposition was estimated using decay rates that adhere to first-order kinetics. The model was based on mass-balance principles but did not have data from a Miscanthus plantation to verify its predictions. Instead, it used parameters specific to a woodland site to verify outputs against field data, reporting a good correlation between observed and simulated values. Since the model does not include environmental variables such as precipitation or air temperature, it is not capable of simulating different climatic scenarios. However, if model predictions can be verified, a sensitivity analysis could reveal how changing LAI or the proportion of plant biomass allocated belowground (i.e. changes that may occur in genetic variants) may impact soil C sequestration rates.
Future modelling of Miscanthus plantations

Empirical data beneficial for improving model performance

Each of the models described have their own advantages and disadvantages dependent on the research objective, input data availability and the required outputs (at appropriate temporal and spatial scales). Overall, there is an increasing number of datasets becoming available to aid model improvement, inter-model comparisons and uncertainty quantification. That said, certain aspects of the Miscanthus C cycle remain poorly quantified due to a lack of experimental data and in particular, parameters describing belowground biomass and root turnover. These scarcely-quantified model parameters are the crux of defining how much C is sequestered in the soil and also of changes in stable soil C stocks that can help offset the C emissions associated with energy generation from Miscanthus biomass. Another important area which requires attention relates to the assigned splitting ratios in SOM modules. These dictate how much C is allocated to the stable or labile soil pools, which in turn determines how long the C is predicted to stay in the soil. To relate these conceptual labile or stable pools of C used by the models to measurable pools, soil fractionation is possible using physiochemical procedures, such as that suggested by Zimmermann et al. (2007). To date there are only two published studies relating soils beneath Miscanthus plantations to SOM pools (Dondini et al., 2009; Poeplau and Don., 2013). Both studies suggest that an increase in total SOC during the lifetime of a Miscanthus plantation is likely, but their estimates of change in stable soil C pools vary considerably. If full C accounting is desired to satisfy cradle-to-grave life-cycle analyses (LCAs) it is essential to accurately predict soil C sequestration rates using stable isotope techniques as used by Dondini et al. (2009) and Poeplau and Don (2013).
Validating model predictions has principally been achieved through comparison with young commercial Miscanthus plantations (e.g. Case et al., 2013; Zimmermann et al., 2013) or field trials (e.g. Christian et al., 2008; Borzecka-Walker et al., 2012). However these do not address issues of optimal crop rotation length, nor do they necessarily reflect the same site conditions experienced by an old (10+ years) commercial plantation of considerable size. This is particularly true for field trials where continued disturbance and different plant densities can influence final yield (Lewandowski et al., 2000; Lewandowski et al., 2003b). Consequently, the importance of continuing to monitor existing commercial plantations cannot be underestimated if we are to ensure model predictions are accurate throughout the entire lifecycle of a Miscanthus plantation. The few larger and older plantations that have been studied indicate that yields can vary greatly depending on climatic conditions during each year and that Miscanthus plantation yields become less commercially viable after 20 years of continual harvesting (Clifton-Brown et al., 2001b; Khanna et al., 2008). It is therefore reasonable to assume other aspects of the C cycle (i.e. soil C sequestration and GHG emissions) may be interacted upon by changes in crop productivity. Since Miscanthus is a relatively recent addition to land owner's establishment options there are few studies in Europe or North America that report how the LUC to a commercial sized plantation affects the site's C budgets (e.g. Zimmermann et al., 2012). To ensure the mechanistic relationships represented by the model simulations are accurate for the LUC and management for both normal and extreme conditions, we need additional field data.

Databases and frameworks for Miscanthus model development
Comprehensive datasets from single field sites are needed to 1) better quantify model uncertainty and 2) allow for model comparison. With the advances in database technology and ease of electronic communication for data sharing, such exchanges are becoming more achievable and recently online databases have been created (e.g. LeBauer et al., 2010). Sensitivity and uncertainty analyses of the models can help identify where data collection needs focus to best reduce uncertainty; however, for all process-based models, uncertainty can come from inputs, model structure or observations (Smith et al., 2012a). Due to the number of sources of uncertainty and limited data availability at present, uncertainty quantification is rarely addressed by the literature that report Miscanthus simulations. That said, LeBauer et al. (2013) were able to provide estimates of how much model uncertainty is due to individual model parameter values by using a Bayesian meta-analysis with available species-specific data. Although, their study uses the specific example of switchgrass, the tool they report and use (the Predictive Ecosystem Analyzer; PEcAn) is just as applicable to Miscanthus. PEcAn can therefore be used with the six models reported here to identify which model parameters contribute most to uncertainty, providing much more information than simple sensitivity analyses.

In addition to data availability limiting model development, a comprehensive framework is lacking around which the different crop models can be developed from site-scales up to regional, or even continental, scales. This is arguably the most important aspect of future model development, since model outputs are the best predictions we have to inform policy decisions. Since governmental policies are often a major driver of LUC, and occur over large geographical scales, a key requirement before Miscanthus plantations are established is the implementation of an integrative computational framework (Zhang et al., 2010). This framework should 1) compile comparable, and spatially explicit, input data; 2) execute model
simulations and report outputs that are consistent with each other; 3) prepare unambiguous visualisations of the findings. Surendran Nair et al. (2012) suggest that such a framework should include a GIS for reprocessing spatial datasets (e.g., Zhang et al., 2010) and an efficient platform upon which to perform model simulations and powerful post-processing analysis of model predictions (e.g., Nichols et al., 2011) — thereby allowing simple comparisons between biomass production/C dynamics and geographic features/climate data to be visualised. Model adaptations of MISCANFOR to simulate Miscanthus biomass production over GIS datasets for Europe (Pogson, 2011; Pogson et al., 2012) and globally (Pogson et al., 2013) are good examples of how upscaling existing models may be the fastest way of forming reliable predictions over larger spatial areas. Similarly, it is worth specifically mentioning the framework outlined by Zhang et al. (2010). This was able to extract input information from text files for use in the EPIC model (Williams et al., 1989) and overlay results against GIS maps. Validated outputs from simulations running under frameworks such as this are promising indicators of how model development may progress in the near future.

Implementing an ecosystem model over large geo-spatial regions requires the most appropriate input databases for site characteristics, such as soil type and climatic variables. Most ecosystem models use similar driving data. Recent applications of Miscanthus models at UK (Hastings et al., 2013), European (Hastings et al., 2009) and global levels (Pogson et al., 2013) demonstrate that appropriate datasets exist at a range of spatial scales to run energy crop soil models. Smith et al., (2012b) recently reviewed spatial datasets available for modelling soil C change at regional to global scales.
Conclusions

In the coming years, a drive toward renewable sources of energy and commitments to reduce national CO₂ emissions are likely to increase interest in bioenergy crops such as Miscanthus, and therefore may induce considerable LUC around the world. With land resources diminishing and concerns for food security increasing, it is a major concern that any LUC that does occur is appropriate and the most effective land use for a given area. Process-based models are valuable tools for addressing this issue and it is therefore essential that they operate reliably for a wide range of environmental conditions. Here we reviewed six process-based crop models that have been parameterised for Miscanthus, reported the current extent of their application and described the possible uses of these models. The models differ in both their design and computational power but none is vastly superior; selecting one over another depends mainly on the particular research question to be answered.

For model development to occur, and therefore to improve the reliability of model predictions, high-quality experimental field data are essential. As interest has grown in the viability of Miscanthus as a bioenergy crop, experimental data from its non-native countries has increased considerably, but not all is freely available or accompanied by the supporting information essential for model development. In addition, to better validate the internal model parameters and rate modifiers, increased data availability through data sharing is also key to addressing the issues of uncertainty quantification and model comparisons, which is necessary for ensuring realistic model outputs. The scientific modelling community is sufficiently motivated to develop Miscanthus crop models and therefore it is highly likely that as data availability and computational modelling power increases, there will be a significant improvement in Miscanthus model performance.
Acknowledgements

We thank the Centre for Ecology and Hydrology and Shell for providing a joint PhD studentship grant award to Andy Robertson (CEH project number NEC04306). The authors are also grateful to Emily Bottoms and Sean Case at CEH Lancaster, Alice Massey at Aberystwyth University, Amy Thomas at the University of East Anglia and Mark Pogson at the University of Aberdeen for helpful discussions. Pete Smith is a Royal Society-Wolfson Research Merit Award holder.
References


Tool (SWAT) to Simulate Its Water Quality Effects As a Bioenergy Crop”. *Environmental Science and Technology*, **45**: 6211-6212. dx.doi.org/10.1021/es201463x


<table>
<thead>
<tr>
<th>Model</th>
<th>Spatial Scale</th>
<th>Biomass Production Submodel</th>
<th>Hydrological Submodel</th>
<th>Soil Quality Submodel</th>
<th>Stress factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Phenological development</td>
<td>Biomass calculation method</td>
<td>Biomass partitioning</td>
<td>Processes simulated</td>
</tr>
<tr>
<td>WIMOVAC and BioCro</td>
<td>Site/Ecosystem</td>
<td>6 stages</td>
<td>Biochem.</td>
<td>4 pools</td>
<td>R, Et</td>
</tr>
<tr>
<td>Agro-IBIS</td>
<td>Ecosystem</td>
<td>3 stages</td>
<td>Biochem.</td>
<td>3 pools</td>
<td>C, R, Et</td>
</tr>
<tr>
<td>DAYCENT</td>
<td>Site</td>
<td>Curve</td>
<td>ARF</td>
<td>5 pools</td>
<td>C, Et, Sf</td>
</tr>
</tbody>
</table>

Table 2: Essential and optional input parameters required by six process-based carbon models parameterised for *Miscanthus*.

<table>
<thead>
<tr>
<th>Model</th>
<th>Weather input parameters</th>
<th>Site input parameters</th>
<th>Crop input parameters</th>
<th>Optional input parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>WIMOVAC and BioCro</td>
<td>Air temperature, Solar radiation, Relative humidity, Wind speed, Precipitation</td>
<td>Sand, silt and clay contents, Initial soil C pools, Soil pH, Latitude and altitude, Field capacity, Wilting point, Irrigation, sowing, harvest dates</td>
<td>Maximum rooting depth, Maximum carboxylation rate, Quantum efficiency, Dark respiration, Thermal periods for growth stages, Dry matter partitioning coefficients</td>
<td>Parameters easily changed through R function calls (generic WIMOVAC model has user interface front end), Rate and timing of fertiliser application, Tillage timing and technique</td>
</tr>
<tr>
<td>Agro-IBIS</td>
<td>Air temperature, Solar radiation, Relative humidity, Wind speed, Precipitation</td>
<td>Sand, silt and clay contents, Initial soil C pools for all layers, Soil pH, Soil bulk density, Initial soil C pools, Irrigation, sowing, harvest dates</td>
<td>Maximum carboxylation rate, Quantum efficiency, Dark respiration</td>
<td>Additional climate inputs (e.g. days frost per year) edited via text file inputs, Rate and timing of fertiliser application, Tillage timing and technique</td>
</tr>
<tr>
<td>DAYCENT</td>
<td>Max./Min. Air Temperature, Precipitation</td>
<td>Sand, silt and clay contents, Soil pH, Initial soil C pools for all layers, Latitude and longitude</td>
<td>Respiration partitioning coefficients, Biomass partitioning coefficients, Crop growth temperature thresholds, Thermal periods for growth stages, N and lignin content</td>
<td>Relative humidity and wind speed, Soil C for each 15 cm layer, Rate and timing of fertiliser application, Tillage timing and technique, Water and nutrient stress modifiers, Solar radiation</td>
</tr>
<tr>
<td>DNDC</td>
<td>Max./Min. Air temperature, Precipitation, N rainfall concentration</td>
<td>Sand, silt and clay content, Soil pH, Soil bulk density, Initial soil C pools, Initial NO₃⁻/NH₄⁺ soil contents, Latitude, longitude and slope</td>
<td>Leaf area index, Maximum crop height, Thermal degree days, Biomass partitioning coefficients, Thermal periods for growth stages</td>
<td>Atmospheric CO₂ &amp; NH₃ concentrations, Rate and timing of fertiliser application, Tillage and harvest timings, Soil C for each 5 cm layer, Solar radiation</td>
</tr>
<tr>
<td>ECOSSE</td>
<td>Air temperature, Potential evapotranspiration, Precipitation</td>
<td>Sand, silt and clay contents, Initial soil C, Soil pH, Soil bulk density, Water Table depth, Latitude</td>
<td>Vegetation cover type <em>(Miscanthus parameters are already included within the basic version)</em></td>
<td>Rate and timing of fertiliser application, Annual crop yield, Tillage and harvest timings, Soil C for each 5 cm layer, Solar radiation</td>
</tr>
</tbody>
</table>