GCB-BIOENERGY
BIOPRODUCTS FOR A SUSTAINABLE BIOECONOMY  Open Access

WILEY

# Uncertainty of modelled bioenergy with carbon capture and storage due to variability of input data

Anita Shepherd [iD] | Mike Martin | Astley Hastings [iD]

Institute of Biological and Environmental
Sciences, School of Biological Sciences,
University of Aberdeen, Aberdeen, UK

**Correspondence**
Anita Shepherd, Institute of Biological
and Environmental Sciences, School
of Biological Sciences, University of
Aberdeen, 23 St. Machar Drive, Aberdeen
AB24 3UU, UK.
Email: anita.shepherd@abdn.ac.uk

## Abstract

Uncertainty is inherent in modelled projections of bioenergy with carbon capture and storage (BECCS), yet sometimes treated peripherally. One source of uncertainty comes from different climate and soil inputs. We investigated variations in 70-year UK projections of *Miscanthus × giganteus* ($M × g$), BECCS and environmental impacts with input data. We used cohort datasets of UKCP18 RCP8.5 climate projections and Harmonized World Soil Database (HWSD) soil sequences, as inputs to the MiscanFor bioenergy model. Low annual yield occurred 1 in 10 years as a UK-average but yield uncertainty varied regionally, especially south and east England. BECCS projections were similar among cohorts, with variation resulting from climate cohorts of the same database ensemble ($3.99 ± 0.14$ t C ha$^{-1}$ year$^{-1}$) larger than uncertainty resulting from soil sequences in each grid block ($3.96 ± 0.03$ t C ha$^{-1}$ year$^{-1}$). This is supported by annual time series, displaying variable annual climate and a close yield–BECCS–climate relationship but partial correspondence of yield and BECCS with maximal soil variability. Each HWSD soil grid square contains up to 10 ranked soil types. Predominant soil commonly has over 50% coverage, indicating why BECCS from combined soil sequences were not significantly different from BECCS using the dominant soil type. Mean BECCS from the full climate ensemble combined with the full soil sequences, over the current area of cropping limits in England and Wales, is $3.98 ± 0.14$ t C ha$^{-1}$ year$^{-1}$. The bioenergy crop has a mean seasonal soil water deficit of $65.79 ± 4.27$ mm and associated soil carbon gain of $0.22 ± 0.03$ t C ha$^{-1}$ year$^{-1}$, with bioenergy feedstock calculated at 131 GJ t$^{-1}$ y$^{-1}$. The uncertainty is specific to the input datasets and model used. The message of this study is to ensure that uncertainty is accounted for when interpreting modelled projections of land use impacts.

**KEYWORDS**
bioenergy, carbon capture, climate change, cohorts, HWSD soil sequences, MiscanFor, miscanthus, modelling, projections, uncertainty

# 1 | INTRODUCTION

Evidence from global Integrated Assessment Models (IAMs) suggest that some level of Greenhouse Gas Removal is needed to achieve a limit of global warming to below 2°C (POST, 2020). If deployed on a large scale, bioenergy with carbon capture and storage (BECCS) is a negative emission technology with the potential to remove carbon dioxide ($CO_2$) from the atmosphere (Smith et al., 2016). Policymakers make decisions on renewable energy and BECCS with the aid of modelled projections for the future.

Modelling is a key tool to predict yields and the inter-annual variability of bioenergy crop production and BECCS and evaluating uncertainty is a key part of environmental modelling (Smith & Smith, 2007). This study focusses on the uncertainty inherent in data, but the models also contribute to variability and uncertainty (Aodha & Edmonds, 2017; Helinga, 1998; Uusitalo et al., 2015). A literature search for the model details and testing may reveal inherent variability or bias of the model itself: in particular, take note of how field data were extracted to parameterize the model, the sensitivity analysis, how accurately the modules have been calibrated and how well the overall model performed in validation. Details for the version of the model used in this study, MiscanFor, can be found in Shepherd et al. (2020).

The impact of different meteorological and soil inputs on model results has significant interest (Pogson et al., 2012). Internal climate variability from within climate projections of the same database source (the random nature of the climate, climate model response and climate forcings) are a source of uncertainty and yet are seldom quantified for crop production predictions (Deser et al., 2012; Qian et al., 2020).

Input data commonly available for climate and soil vary, even from within an ensemble database from the same source. Ensemble databases are thought to be the best basis for estimating projection uncertainties (Mauritzen et al., 2017), these are supplied as multi-member datasets of gridded climate with equally likely occurrence.

UKCP18 RCP8.5 ensemble climate projections are in common use. RCP8.5 are worst-case climate projections reflecting business-as-usual socio-economic scenarios (Riahi et al., 2011). Rogelj et al. (2012) determined that during 2090–2099, RCP 8.5 temperature is projected to rise 3.8–5.7 degree C above pre-industrial temperatures.

A multi-member soil dataset, The Harmonized World Soil Database (HWSD), is also commonly used (Fischer et al., 2008) where multiple soils types are ranked in 'sequences' by the predominance of soil type in grid square area coverage. For the UK, there are eight HWSD soil data sequence groups. The soil types' percentage share of the grid square coverage could vary if the gridded boundaries were to change so we can either consider the difference between outputs of soil sequence groups, or outputs combined as provision intended.

Our study determines what effect the variability from input data cohorts has on reporting the BECCS projection in the UK and its uncertainty. To evaluate this we use a published, validated crop growth and bioenergy model, MiscanFor (Hastings et al., 2009) using *Miscanthus × giganteus* ($M \times g$) as a feedstock, and use the UKCP18 12-member RCP8.5 climate projections, for the variable climate input and the HWSD soil sequences as the variable soil input.

MiscanFor was used because it is applicable to a study of bioenergy, BECCS and environmental impacts, it has proved accurate and is continually updated (Hastings et al., 2009; Shepherd, Littleton, et al., 2020). The model contains many parameters and requires a minimal amount of commonly available datasets. In a literature comparison of multiple bioenergy models including MiscanFor, Surendran Nair et al. (2012) stated that models which simulate soil water, nutrient and carbon cycle dynamics make them especially useful for assessing the environmental consequences. Also, that field trials that address the influence of genetic, environmental and crop management on biomass production will provide valuable data for the development and calibration of bioenergy crop models. Plus, that future research should explore an integrated framework for efficient execution of large-scale simulations and processing of input and output data. MiscanFor has all three of these capabilities, the more recent version has since incorporated soil carbon sequestration and soil water deficit, the model is be used for different bioenergy crops and genetic varieties, and the model integrates a Java front end with Fortran processing capability and Python visualization scripts to create global bioenergy and environmental output.

Soil and climate are two of the most important inputs whose variability influences elements of a bioenergy crop: water availability, crop growth and, via yield and leaf litter, soil C.

We have calculated the potential for carbon capture and storage (CCS) following Albanito et al. (2019), who assumed 90% $CO_2$ capture post-combustion at biomass electricity plants, being broadly similar across plants with varying efficiency:

$$CCS = [DM \times 0.5] \times 0.9, \qquad (1)$$

where CCS is the annual $CO_2$ captured and transferred into geological storage expressed in terms of units of C (not $CO_2$), DM is the dry matter $M \times g$ biomass, and 0.5 assumes 50% C in biomass and 0.9 refers to 90% CCS efficiency.

We have referred to the CCS in units of C as BECCS_C throughout this text. It can be converted to units of $CO_2$ by multiplying by the ratio of 44/12 (ratio of the molecular weight of carbon dioxide to that of carbon).

Estimates by the Committee on Climate Change (CCC, 2018), also quoted in the Net Zero report (CCC, 2020, p. 142), estimate between 20 and 65 Mt $CO_2$e year$^{-1}$ could

be sequestered through BECCS in the UK up to 2050. These ranges indicate a substantial range of uncertainty. The land area projected to be given over to bioenergy crops is constantly changing, and some IAMs do not model bioenergy land area projections for the UK very well (Shepherd et al., 2020a) which is a separate source of uncertainty confusing the uncertainty from soil and climate. An assessment of modelling uncertainty is required arising from model inputs of climate and soil, and also from modelled bioenergy land area. In this study, we keep the area of land constant to quantify uncertainty for mean BECCS per hectare projections of the UK resulting from ensemble climate inputs and from soil sequence inputs.

## 2 | MATERIALS AND METHODS

MiscanFor simulates BECCS_C projections using $M \times g$ as a feedstock. $M \times g$ is a bioenergy crop with relatively high yields under a range of conditions (Pogson et al., 2012). It has a higher energy output/input ratios than other bioenergy crops and a lower carbon (C) cost of energy production than fossil fuels (McCalmont et al., 2017; Sims et al., 2006), with a consistent increase in grower uptake (Shepherd, Clifton-Brown, et al., 2020).

MiscanFor is a model which provides projections of bioenergy crop growth and power generation, along with a number of environmental variables such as soil water and soil organic carbon (SOC). It can be tailored to various crops, including $M \times g$. Annual or seasonal outputs are averaged over the years simulated.

Pogson et al. (2012) noted the model being moderately sensitive to temperature and precipitation, and a soil's field capacity and wilting point, as expected of any crop model.

Hastings et al. (2009) identified the model's main sensitivities as the photoperiod sensitivity in addition to drought resistance and frost tolerance. The model has since been updated to correct the sensitivity of yields to drought and newer climate and soil databases have underpinned the improvements (Shepherd, Littleton, et al., 2020); we update the sensitivity for the current model and input data used in this study.

In MiscanFor, dry matter assimilation is calculated from the fraction of radiation intercepted by the canopy (dependant on leaf area index, an extinction coefficient and photosynthetically active radiation), modified by radiation use efficiency and an overheating factor. Both the increase and senescent decline of LAI are linearly related to the degree day accumulation. Average annual crop yields of dry matter biomass are output. MiscanFor calculates a soil water balance and incorporates a reduction on crop growth by soil moisture deficit via photosynthesis and the reduction of evapotranspiration. A soil C module has been incorporated in MiscanFor (Dondini et al., 2009; Shepherd, Littleton, et al.,

2020), this module is based on a proposed generic theory for the dynamics of C and nitrogen (Bosatta & Agren, 1985, 1991). Algorithms simulate the input of crop litter as unique pools of soil organic matter with exponential rates for decomposition. The amount of crop litter input to the soil is the difference between $M \times g$ peak yield and harvest yield.

Projections of carbon capture and storage (CCS) are calculated from crop yield modelling multiplied by known efficiency factors for specific bioenergy supply chains. Following the methodology in Hastings et al. (2009), the MiscanFor model simulates a recommended $M \times g$ crop scenario for the UK: local use of the non-irrigated $M \times g$ as feedstock for electricity generation within 20 km.

### 2.1 | Pre-processing steps

*Step 1*: The climate projection dataset used was UKCP18 RCP8.5 (Met Office Hadley Centre, 2018), which contains 12 equally relevant climate projections at 0.19 degree spatial resolution of the British Isles (see Figure 1).

This dataset is produced by the 12 km Met Office Hadley Centre HadREM3-RA11 M regional model. The model spans the UK and is driven by the Met Office Unified Model Global Atmosphere GA7 model (HadREM3-GA705) at 12 km resolution. The HadREM3-GA705 model is driven by perturbed
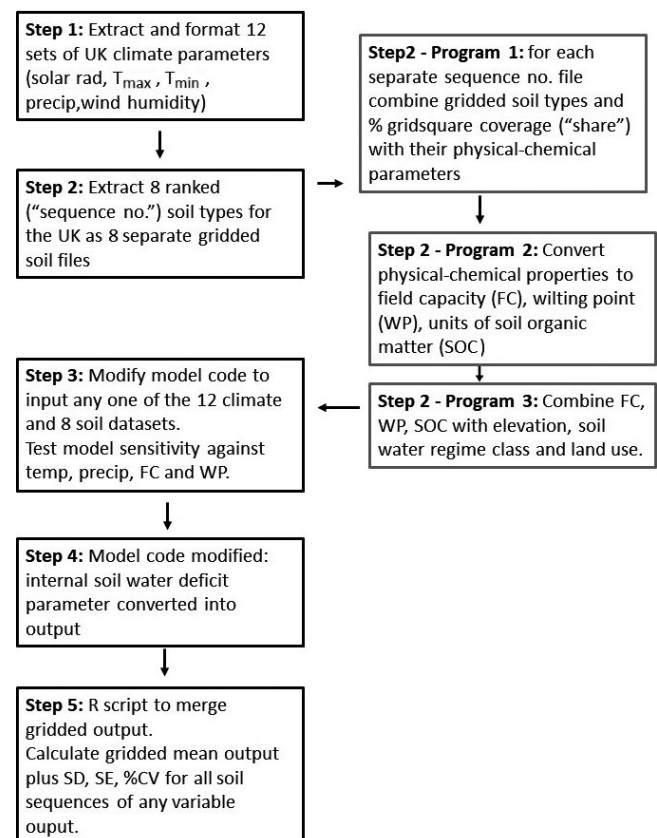


**Step 1:** Extract and format 12 sets of UK climate parameters (solar rad, $T_{max}$, $T_{min}$, precip, wind humidity)

**Step 2:** Extract 8 ranked ("sequence no.") soil types for the UK as 8 separate gridded soil files

**Step 3:** Modify model code to input any one of the 12 climate and 8 soil datasets. Test model sensitivity against temp, precip, FC and WP.

**Step 4:** Model code modified: internal soil water deficit parameter converted into output

**Step 5:** R script to merge gridded output. Calculate gridded mean output plus SD, SE, %CV for all soil sequences of any variable ouput.

**Step2 - Program 1:** for each separate sequence no. file combine gridded soil types and % gridsquare coverage ("share") with their physical-chemical parameters

**Step 2 - Program 2:** Convert physical-chemical properties to field capacity (FC), wilting point (WP), units of soil organic matter (SOC)

**Step 2 - Program 3:** Combine FC, WP, SOC with elevation, soil water regime class and land use.

**FIGURE 1** Pre-processing schedule

variants of the global climate model, HadGEM3-GC3.05 (Sexton et al., 2019). The 12 projections take the name of the perturbed-physics ID (e.g. p00000, p01935, p02868, etc.) for Met Office Hadley Centre models used in the CEDA archive from which they were downloaded. Sexton et al. (2019) explain that these perturbed physics in GA7 models may relate to variation in model parameters relating to land surface, snow, cloud, aerosol, convection, or gravity wave.

Individual parameters for net shortwave solar radiation, maximum and minimum temperature, precipitation, wind and humidity were extracted and reformatted for model input. RCP8.5 climate data are supplied as monthly data and the MiscanFor model is designed to partition these into daily inputs within the model.

*Step 2*: We used the HWSD gridded soil data for the UK with sequences of soil types and soil parameters per grid square (Wieder et al., 2014), at 0.00833 degree spatial resolution. Each grid point in the HWSD global dataset has up to 10 dominant soil sequence types with the percentage area of each within the grid block, although for the UK coverage there were a maximum of eight soil sequence types. The soil sequence groups were ranked in order of coverage per grid square. The maximum number of soil types a grid square in the UK contained was eight. Eight alternative UK soil sequence files for model input were produced of HWSD parameters, from the soil types with most coverage in sequence 1 to the soil types with least coverage in sequence 8.

To convert physical soil data from the HWSD database to the soil parameters we required, the data undergo transformation in a three-step soil data pre-processing:

Program 1: Extracted from GIS data, the latitude, longitude and MUGLOBAL ID (which identifies a specific combination of soil types for a grid square) are combined with the soil parameters for those soil types. Among the soil parameters affiliated with the MUGLOBAL ID is SEQ, the sequence or ranking of the soil in the soil mapping 8-unit composition and SHARE, the % coverage of the grid square for each soil sequence in the soil mapping unit. The latitude, longitude, MUGLOBAL ID and soil data are output into separate files based on the sequence number. This results in a complete soil data file for the UK containing the most predominant soil type of the grid squares (SEQ 1), and a file containing the next most predominant soil type parameters (SEQ 2) and so on. UK files run up to SEQ 8, and the higher the sequence number the lower the coverage of the UK as most of the UK grid does not have a combination of over five types of soils.

Program 2: This program originally created for MiscanFor (Hastings et al., 2009) is based on Campbell (1985) pedotransfer functions. It transforms the output of Program 1, physical soil data (soil depth, gravel, sand, silt, clay, bulk density, calcium carbonate for top and subsoil layers) to parameters of field capacity (FC) and permanent wilting point (PWP), and converts units of soil organic matter (SOC). The program uses the criteria of shallow soil less than or equal to 30 cm with a high percent of calcium carbonate in topsoil to identify soils on a chalk substrate. The model then considers the topsoil characteristics to extend to 4 m to enable the capillary water supplied by the chalk to be used by the plant and avoid water stress (Hastings et al., 2014).

Program 3: This program is a later addition (Shepherd, Littleton, et al., 2020), which combines the output of Program 2 with elevation from the climate dataset for adiabatic lapse correction to temperature and the SWR (Soil Water Regime) class (from the HWSD database) to indicate relative groundwater support. The program also adds in land use data (from Rounsevell et al., 2006).

All of the above programs involve the program scanning one set of data for the nearest neighbour within the resolution in the other set of data, to coordinate and merge the data.

*Step 3*: The latest version of the spatial MiscanFor bioenergy model (Shepherd, Littleton, et al., 2020) was modified to use any one of the eight sequences of HWSD soil data, and the UKCP18 RCP8.5 climate (one of any 12 cohort datasets). The sensitivity of the model output of BECCS_C (considering $M \times g$ as the feedstock) was tested against modifications of temperature, precipitation, field capacity and wilting point (Figure 2), varying one parameter while keeping others fixed (Pogson et al., 2012; Smith & Smith, 2007).

*Step 4*: The MiscanFor model produces soil water deficit as an internal intermediate variable. The code was modified for this study to output the aggregated seasonal water deficit for the UK between May and September, average the seasonal deficit over all years simulated for that grid square, and output the result.

*Step 5*: Gridded output from different HWSD soil sequence files varies in length, it varies in the number and order of grid squares, from the highest coverage soil type, to the lowest coverage soil type. A 4th program was written in R to read and merge eight outputs modelled using the different soil files, to collate output values to the same grid square, adding a null code if the soil sequence does not cover the grid square. Outputs resulting from the eight soil inputs were processed to calculate the mean, standard deviation, standard error and % coefficient of variation (%CV), then output these with the longitude and latitude so that they could be mapped. A separate version of the program was also developed for outputs using the 12-member climate cohorts.

A summary of the input datasets is shown in Table 1.

## 2.2 | Processing steps

*Step 1. Same climate database, different cohorts.* 12-member climate set 2008–2080: We calculated the uncertainty between 12 sets of mean annual output (averaged over 2008–2080) simulated from using the full 12-member
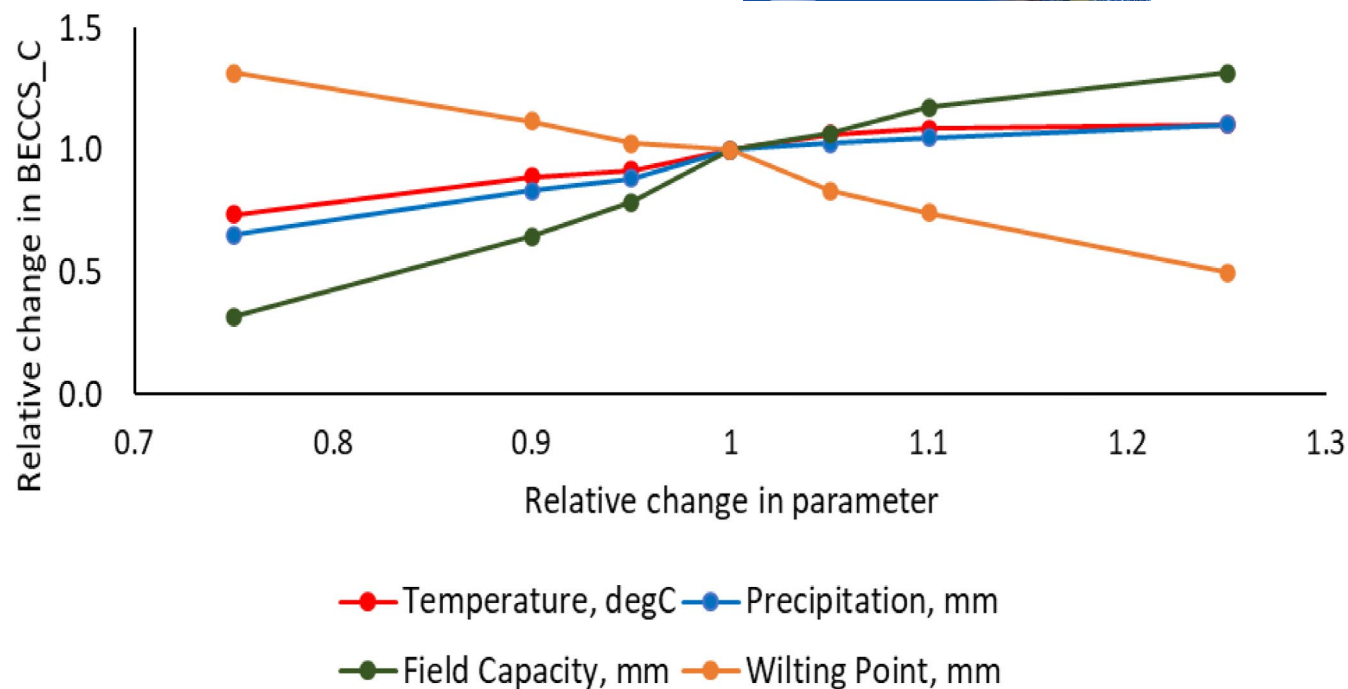
**FIGURE 2**   Sensitivity of model-produced BECCS_C to input parameters (using RCP8.5 climate 2010–2080 and HWSD soil parameters)
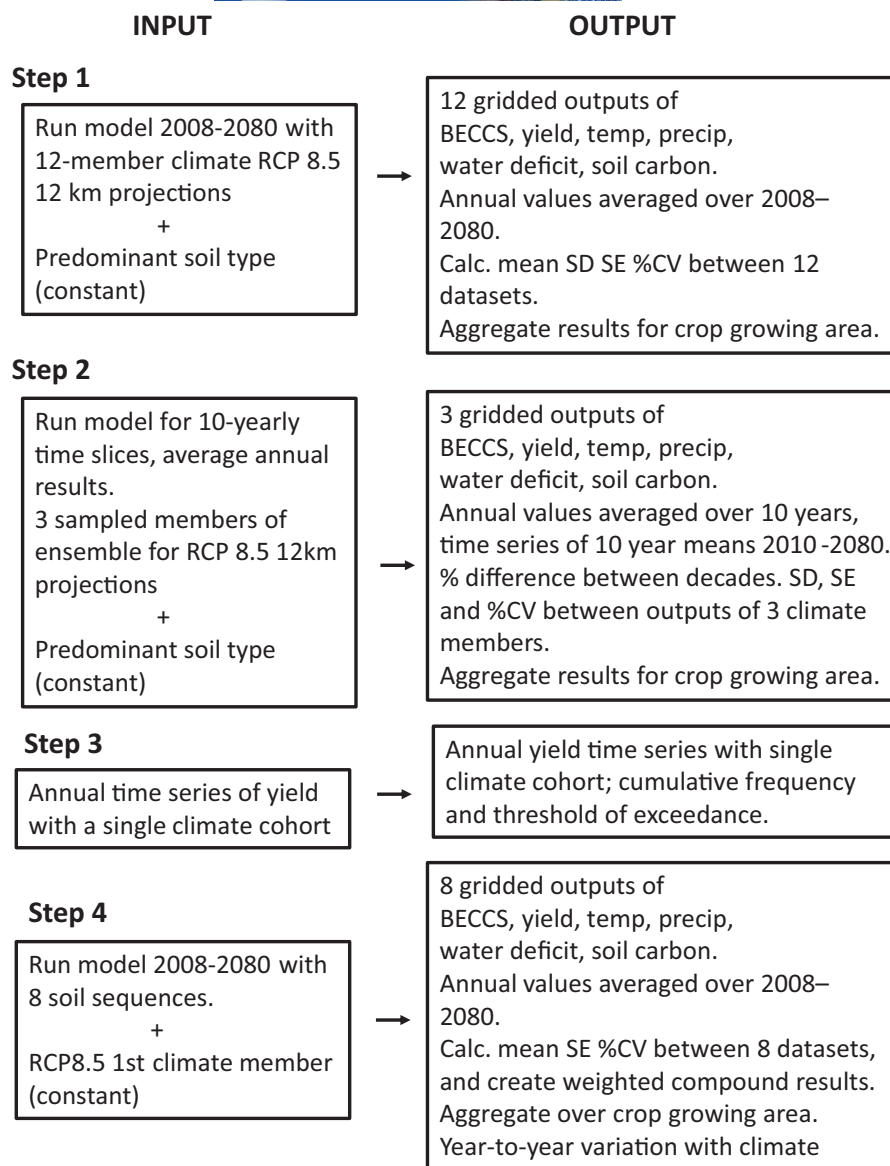
**TABLE 1**   Summary of input datasets

| Data source | Format | Resolution |
| --- | --- | --- |
| UKCP18 RCP8.5 (Met Office Hadley Centre, 2018) | 12 cohorts of a climate ensemble.<br>Files: incoming shortwave radiation, precipitation, temperature max and min, wind, relative humidity, elevation associated with dataset.<br>Data boundary: UK<br>Data in ascii text files | Spatial resolution: 0.19 degrees<br>Temporal resolution: monthly data, interpolated to daily in model |
| The Harmonized World Soil Database (HWSD) (Fischer et al., 2008) | 8 UK cohorts (globally 10 cohorts) of ranked % share of soil type within a single grid cell area.<br>Files of parameters associated with soil type: contents of sand, silt, clay, gravel, chalk, organic C (which are used in the model in pedo-transfer functions and organic matter decomposition functions).<br>Data boundary: UK<br>Data in ascii text files | Spatial resolution: 1 km<br>Temporal resolution: fixed data |

climate RCP 8.5 12 km projections (Hadley Centre model HadREM3-GA705 using physics perturbations from the global circulation model HadGEM3 GC3.05, Met Office Hadley Centre, 2018; see Figure 3). $M \times g$ growth was simulated, while all other parameters were kept constant, using soil sequence 1, the predominant soil type of the grid squares. Variables output were total annual precipitation and mean annual temperature, seasonal water deficit, BECCS_C, dry matter yield and soil C, all averaged over the period 2008–2080. These outputs are absolute values resulting from different cohorts of the same climate database. This is different to the next section which shows

mean decadal change in output resulting from using a sample of those same climate cohorts.

*Step 2. Same climate database, different cohorts, and variation in decadal change*: Decadal mean outputs 2011–2020 through to 2071–2080 were produced (total annual precipitation and mean annual temperature, seasonal water deficit, BECCS_C, dry matter yield and soil C) resulting from running the MiscanFor model for 10 years using three cohorts of the monthly RCP8.5 climate projections to reduce the amount of processing required. The cohorts of the RCP8.5 climate ensemble are given IDs referring to their climate physics perturbations in HadGEM3 GC3.05. We used cohorts p00000,

**INPUT**                    **OUTPUT**                    **FIGURE 3**  Processing schedule

**Step 1**

| Run model 2008-2080 with 12-member climate RCP 8.5 12 km projections  +  Predominant soil type (constant) | → | 12 gridded outputs of BECCS, yield, temp, precip, water deficit, soil carbon. Annual values averaged over 2008–2080. Calc. mean SD SE %CV between 12 datasets. Aggregate results for crop growing area. |

**Step 2**

| Run model for 10-yearly time slices, average annual results. 3 sampled members of ensemble for RCP 8.5 12km projections  +  Predominant soil type (constant) | → | 3 gridded outputs of BECCS, yield, temp, precip, water deficit, soil carbon. Annual values averaged over 10 years, time series of 10 year means 2010 -2080. % difference between decades. SD, SE and %CV between outputs of 3 climate members. Aggregate results for crop growing area. |

**Step 3**

| Annual time series of yield with a single climate cohort | → | Annual yield time series with single climate cohort; cumulative frequency and threshold of exceedance. |

**Step 4**

| Run model 2008-2080 with 8 soil sequences.  +  RCP8.5 1st climate member (constant) | → | 8 gridded outputs of BECCS, yield, temp, precip, water deficit, soil carbon. Annual values averaged over 2008–2080. Calc. mean SE %CV between 8 datasets, and create weighted compound results. Aggregate over crop growing area. Year-to-year variation with climate |

p01935 and p02868 as named in the UKCP18 database (named here as 1, 6 and 12) to view changes of temperature and precipitation through the century and associated changes in BECCS_C projection and environmental impact.

*Step 3. Annual variation and yield risk.* For higher temporal resolution underlying the decadal results, a cumulative frequency analysis was performed on annual dry matter yield limits under climate cohort p00000. The inverse of cumulative frequency is the threshold of exceedance or non-exceedance, which is the return period of low and high yields, a statistic applied in flood modelling (e.g. Shepherd et al., 2017).

*Step 4. Same soil dataset and different cohorts*: All eight soil sequence files were used to simulate $M \times g$ growth while all other factors including climate projections were kept constant (using the first ensemble member p00000 and running MiscanFor for all years 2008–2080, which produced mean annual outputs). Variables output were field capacity,

permanent wilting point, seasonal water deficit, BECCS_C, dry matter yield and soil C.

In steps 1, 2 and 4, the mean, standard deviation, standard error and %CV between the outputs for each variable were calculated and ANOVAs performed to compare datasets.

The whole of the UK was modelled for $M \times g$ growth, but only the area of England and Wales below 54.5 degrees North was aggregated for statistical results, as this is the current limit of viable $M \times g$ crop yields, (the most northern UK widely known $M \times g$ crop being at ADAS High Mowthorpe farm on the Yorkshire Wolds). There is a temperature limit for viable yields of $M \times g$ and a risk of frost kill, and it is very difficult to predict where growers will choose to invest in $M \times g$ in the future despite projected temperature increases, so we have based our crop growing area within the current cropping limit. In addition, this keeps aggregated area fixed along with all other variables while we vary only the climate cohorts or the soil cohorts. The full simulation period chosen

is 2008–2080 which starts when grower investment in $M \times g$ for bioenergy increased and extends to the end year of the UKCP18 RCP8.5 12-member dataset, decadal data used are within this period. Steps 1 and 4 simulate the full period on a daily timestep and average output for mean annual values, while step 2 simulates each separate decade between 2010 and 2080 on a daily timestep (with an initial 2-year spin-up 2008–2009), and averages the 10-year output for mean annual values to determine the size of decadal changes and differences between the cohorts. Step 3 simulates the yield annually 2010–2080 using the first climate cohort as an example of annual variation and risk.

# 3 | RESULTS

## 3.1 | Same climate database, variability between cohorts for spatial mean

Table 2 displays the mean, standard deviation, standard error and %CV values of total annual precipitation and mean annual temperature, seasonal water deficit, dry matter yield, BECCS_C and soil C change, corresponding to inputs of 12 RCP8.5 climate cohorts while other inputs remain static.

Table 2 shows a relatively low standard deviation compared to the mean for all variables, indicating a low spread of data around the mean of the 12 climate cohorts over the 73-year simulation period. Table 2 shows the BECCS_C standard error of the mean over the gridded area in England and Wales (0.14) to be lower than that for temperature (0.19) and precipitation (18.52). Precipitation always has a larger uncertainty relative to other climate variables, influencing

the uncertainty of the water deficit (3.39). The standard error of soil C is negligible.

ANOVA was performed on 12 outputs simulated from the climate cohorts, each one with 73 annual values. Outputs averaged over England and Wales were temperature, precipitation and BECCS_C, and all showed that the variables are not statistically similar between all 12-members of the climate ensemble (Table 2). However, post-hoc comparison using a t test with Bonferroni correction indicated similar groups 1, 3, 5, 6, 9 and 12 to have no significant differences for BECCS_C ($M = 4.37$, 4.45, 4.15, 4.08, 4.43, 4.36; SD = 1.14, 0.93, 0.90, 1.10, 1.10, 1.14, respectively). ANOVA on BECCS_C values from these groups confirms this, $F(5,432) = 1.58$, $p = 0.16$ (Fcrit 2.23, α 0.05).

## 3.2 | Same climate database, different cohorts, comparing variability in decadal change

Table 3 shows the mean decade-to-decade change between 2011 and 2080, sampling three climate cohorts out of 12 (sampled due to the intensive processing required). Mean decadal changes between the three climate cohorts have a low standard deviation relative to the mean, except for precipitation. In the case of precipitation, this is due to the mean change being clustered around zero, showing slight increases or slight decreases, producing a SD as large as the mean, and hence a relatively large %CV.

Temperature projections in the RCP8.5 database did not change equally decade to decade, the largest increases were 8% and 6% increase on the previous decadal mean

**TABLE 2** Comparison of output from all 12 climate cohorts, full bioenergy simulation period of 2008–2080, spatially averaged values below 54.5 deg. N, current limit of $M \times g$ growth

| | Annual mean temp (°C) | Annual total precipitation (mm) | Dry matter yield[a] (t ha$^{-1}$ year$^{-1}$) | Annual BECCS in units of carbon[a] (C storage t ha$^{-1}$ crop year$^{-1}$) | Water deficit, annual May to Sept (mm) | Annual soil C change[a] (t ha$^{-1}$ year$^{-1}$) |
|---|---|---|---|---|---|---|
| 2008–2080 mean | 10.86 | 1019 | 8.87 | 3.99 | 71.38 | 0.22 |
| 2008–2080 Std Dev | 0.67 | 64.18 | 1.10 | 0.49 | 11.77 | 5.8E−17 |
| 2008–2080%CV | 6.18 | 6.30 | 12.68 | [b] | 16.48 | 2.6E−14 |
| 2008–2080 Std Error | 0.19 | 18.52 | 0.32 | 0.14 | 3.39 | 1.7E−17 |
| ANOVA 12 groups of 73 annual BECCS_C 2008–2080 | Temperature: $F(11,864) = 26.4$, $p = 1.2 \times 10^{-47}$ (Fcrit 1.8, α 0.05) Precipitation: $F(11,864) = 21.3$, $p = 1.4 \times 10^{-38}$ (Fcrit 1.8, α 0.05) BECCS_C: $F(11,864) = 13.5$, $p = 5.1 \times 10^{-24}$ (Fcrit 1.8, α 0.05) The temperature, precipitation and BECCS_C resulting from the 12 cohorts is significantly different However paired t tests between groups of the above showed groups 1, 3, 5, 6, 9, 12 to have no significant difference, ANOVA on BECCS_C for climate cohorts 1, 3, 5, 6, 9, 12 produced $F(5,432) = 1.58$, $p = 0.16$ (Fcrit 2.23, α 0.05) | | | | | |

[a]Soil C only arable soils (<15% SOC), not including high carbon soils which average extreme values with the majority to give a spurious result.

[b]%CV is the same as for DM Yield, since BECCS is DMYield multiplied by a factor of 50% carbon and 90% efficiency.

**TABLE 3** Percentage decadal data changes (2011–2080) of RCP8.5 climate output (averaged over UK's 2020 $M \times g$ cropping area, below 54.5 degrees N; averaged over three climate cohorts)

| | Cohort | Annual mean temp (°C) | Annual total precip. (mm) | Dry matter yield (t ha⁻¹ year⁻¹) | Annual BECCS in units of carbon (C storage t ha⁻¹ crop year⁻¹) | Water deficit, annual May to Sept (mm) | Soil C change (t ha⁻¹ crop year⁻¹) |
|---|---|---|---|---|---|---|---|
| 2011–2020 (absolute value) | 1 | 8.90 | 1036 | 9.66 | 4.35 | 68.14 | 0.85 |
| | 6 | 9.20 | 1052 | 9.52 | 4.28 | 71.77 | 0.82 |
| | 12 | 10.00 | 1143 | 11.33 | 5.10 | 75.34 | 1.05 |
| | Mean | 9.37 | 1077 | 10.17 | 4.58 | 71.75 | 0.90 |
| Mean consecutive decade-to-decade change (2011–2080) | 1 | 4.63% | −0.9% | −3.55%[a] | [b] | 9.55% | −22.54% |
| | 6 | 4.04% | 0.54% | −3.75% | | 6.51% | −19.14% |
| | 12 | 2.95% | 0.85% | −2.01% | | 5.70% | −20.00% |
| | Mean | 3.87% | 0.16% | −3.10% | | 7.25% | −20.56% |
| Std Dev of mean decadal change between cohorts | | 0.85 | 0.93 | 0.95 | [b] | 2.02 | 1.77 |
| %CV of mean decadal change between cohorts | | 21.93 | 568 | −30.74 | [b] | 27.97 | −8.61 |
| Std Error of mean decadal change between cohorts | | 0.49 | 0.54 | 0.55 | [b] | 1.17 | 1.02 |
| ANOVA between three sampled cohort means, 7 decades | | F(2,18) = 1.19, p = 0.32, Fcrit 3.55, α 0.05 no signif difference between cohorts for decadal changes | F(2,18) = 0.41, p = 0.67, Fcrit 3.55, α 0.05 no signif difference between cohorts for decadal changes | F(2,18) = 0.15, p = 0.86, Fcrit 3.55, α 0.05 no signif difference between cohorts for decadal changes | F(2,18) = 0.15, p = 0.86, Fcrit 3.55, α 0.05 no signif difference between cohorts for decadal changes | F(2,18) = 0.36, p = 0.70, Fcrit 3.55, α 0.05 no signif difference between cohorts for decadal changes | F(2,18) = 0.95, p = 0.41, Fcrit 3.55, α 0.05 no signif difference between cohorts for decadal changes |

[a]Consecutive decade to decade changes in DM Yield and BECCS_C hide that initial decades experience an increase and post-2051 experiences decreases.

[b]% decadal change and %CV is the same as for DM Yield, since BECCS is DMYield multiplied by a factor of 50% carbon and 90% efficiency.

(2021–2030 and 2061–2070, respectively) displayed by cohorts 1 and 6.

Early decades experience increased precipitation, 2041–2050 and 2061–2070 display the largest decreases of 5.6–7% and 5–6% decrease on the previous decade, over all cohorts.

Climate and climate effects are variable over the decades with an overall resulting decade-to-decade decrease in yield and C storage and slight decadal increase in water deficit using all three sampled climate cohorts.

## 3.3 | Annual variation and yield risk

Underlying the decadal results are annual variation which would interest growers, so we give some information on the annual variation and likelihood of low or high yields, which are linearly related to BECCS_C by a factor of 0.45. Between 2010 and 2080, under the HWSD majority soil type and RCP8.5 climate projection using cohort p00000, the projected mean annual yield for the UK is 12.2 t ha⁻¹ year⁻¹, with a SD of 1.9, %CV of 15.8 and SE of 0.23. Figure 4 shows the variability of the mean UK yield closely follows the precipitation and accumulated degrees during growing season (May–October), severe drought stress also influences the mean yield and can be seen to rise towards the end of the projected period. All simulations assume a non-irrigated crop.

Averaged over the whole of the UK, the mean yield has a normal temporal distribution 2010–2080 (Figure 5a), this is because the yields in the north are lower with cooler climate

FIGURE 4  Mean UK annual dry matter yield under RCP8.5 climate scenario with (a) annual precipitation, (b) accumulated May–June degrees and (c) % grid cells with extreme drought stress
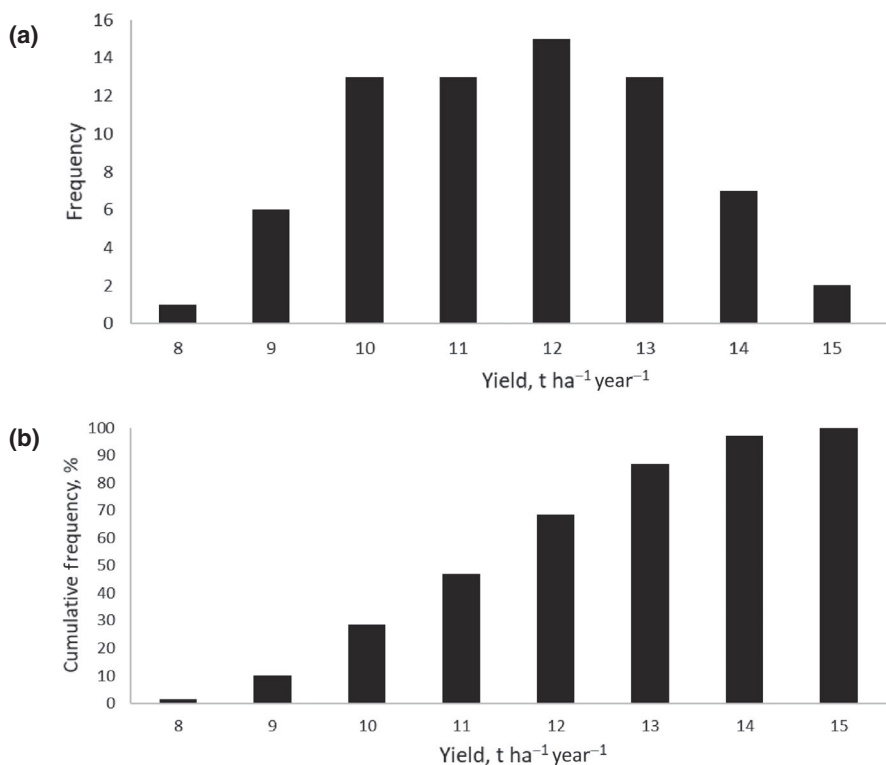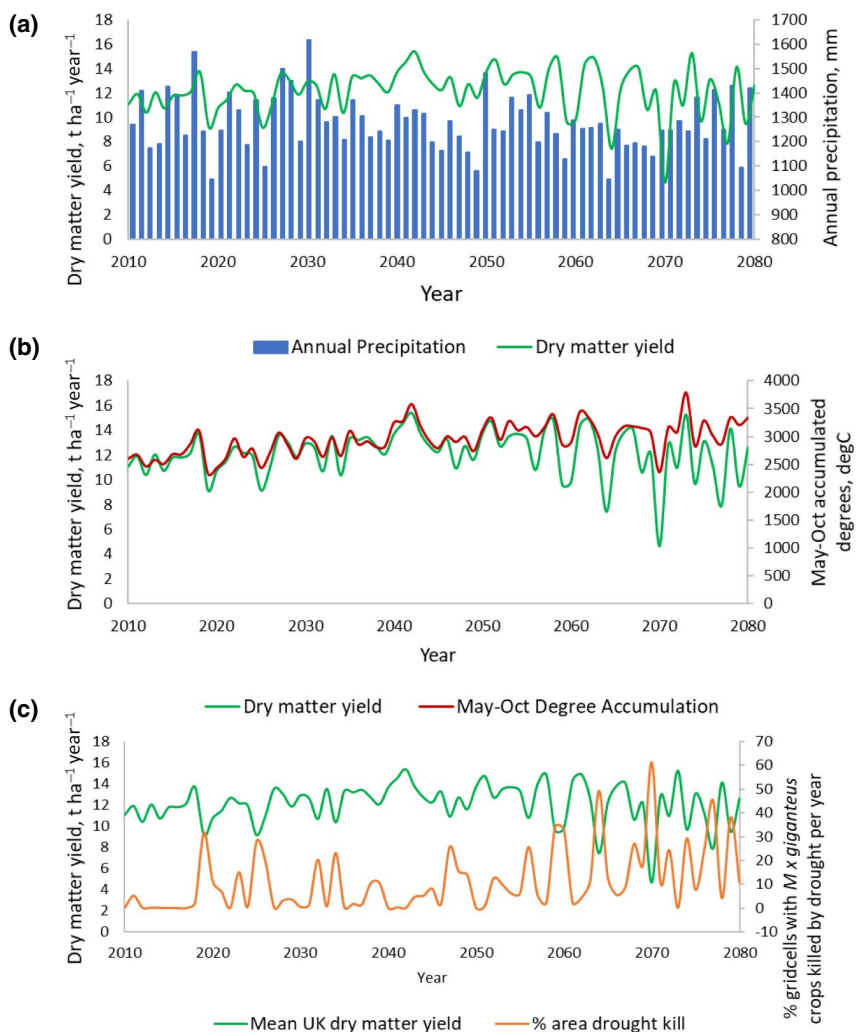


FIGURE 5  (a) Temporal distribution of mean UK dry matter yield 2010–2080 (b) cumulative frequency of mean UK dry matter yield 2010–2080

in the early century, and the yields in the south and east are reduced due to drought conditions, increasingly in the later period of simulation after the 2070s, both of which create near-equal and opposite tails of the distribution. The cumulative frequency (non-exceedance; Figure 5b) shows the cumulative percentage of annual data which fall under a threshold yield. Conversely, an exceedance plot would show the cumulative percentage of data over a threshold yield.

Based on the cumulative frequency, the return period or threshold of non-exceedance and of exceedance (Table 4)

**TABLE 4** Threshold of exceedance (return period) of low or high yields

| Under yield threshold, t ha$^{-1}$ year$^{-1}$ | Return period (years) |
|---|---|
| 8 | 70.0 |
| 9 | 10.0 |
| 10 | 3.5 |

| Over yield threshold, t ha$^{-1}$ year$^{-1}$ | Return period (years) |
|---|---|
| 15 | 35.0 |
| 14 | 7.8 |
| 13 | 3.2 |

provides a risk assessment of low or high yield re-occurrence during 2010–2080, respectively.

## 3.4 | Same soil dataset, different cohorts, compared and combined

Table 5 shows the difference between variables using the different ranked soil type coverage, known as soil sequence groups. Seven out of eight soil sequence groups included coverage for England and Wales under 54.5 degrees N. In the MiscanFor model, the soil parameters that affect output are the ones which contribute towards field capacity, wilting point and soil C (see Program 2 in Section 2). Although the standard error of field capacity and wilting point was 64.9 and 48.6, respectively, outputs are not sensitive to these changes (standard error water deficit 2.6, BECCS_C and soil C both 0.03), as also reported on DM yield by Pogson et al. (2012) using different soil and climate inputs. This reflects a dampening effect of the variability, smoothed by the modelled growth processes over time. The exception to this would be when risks converge, such as low precipitation occurring on soil of low water holding capacity.

The only output parameter that does not work well for the combined calculation is Soil C. Over the UK, central

**TABLE 5** (a and b) Spatially averaged results under 54.5 degN parameters for output from all soil sequences

| (a) | SEQ1 | SEQ2 | SEQ3 | SEQ4 | SEQ5 | SEQ6 | SEQ7 | SEQ8 |
|---|---|---|---|---|---|---|---|---|
| Field Capacity, mm | 663.7 | 566.7 | 471.2 | 618.2 | No data[a] | 611.7 | 299.6 | 169.7 |
| Wilting Point, mm | 376.6 | 345.2 | 269.2 | 378.9 | No data[a] | 292.7 | 91.22 | 33.1 |
| May–September Water Deficit, mm | 66.3 | 64.9 | 65.2 | 61.0 | No data[a] | 70.8 | 83.6 | 68.1 |
| Dry Matter Yield, t ha$^{-1}$ year$^{-1}$ | 9.71 | 9.70 | 9.42 | 9.86 | No data[a] | 10.60 | 10.98 | 9.87 |
| Annual BECCS, t C. ha$^{-1}$ year$^{-1}$ | 4.37 | 4.38 | 4.28 | 4.28 | No data[a] | 4.77 | 4.77 | 4.77 |
| Soil C change, t ha$^{-1}$ year$^{-1}$ | 0.23 | 0.26 | 0.19 | 0.25 | No data[a] | 0.28 | 0.32 | 0.06 |

| (b) | Mean of 7 soil sequences | Std Error of 7 soil sequences | %CV of 7 soil sequences | Weighted compound value: $\sum_1^7$ (DMYield. fraction coverage) |
|---|---|---|---|---|
| Field Capacity, mm | 485.8 | 64.9 | 35.3 | 556.8 |
| Wilting Point, mm | 255.3 | 48.6 | 50.4 | 322.8 |
| May–September Water Deficit, mm | 68.6 | 2.6 | 9.8 | 60.2 |
| Soil C change, t ha$^{-1}$ year$^{-1}$ | 0.23 | 0.03 | 34.6 | 0.21 |
| Dry Matter Yield, t ha$^{-1}$ year$^{-1}$ | 10.03 | 0.06 | 5.27 | 8.80 |
| Annual BECCS, t C. ha$^{-1}$ year$^{-1}$ | 4.52 | 0.03 | [b] | 3.96 |

| ANOVA: 7 groups of 73 annual BECCS_C 2008–2080 | Comparing area mean annual BECCS_C output from 7 soil sequence groups existing for the area below 54.5 degN (separate annual values, 2008–2080): at $\alpha = 0.05$, $F(6,504) = 2.818$, $p = 0.01$, Fcrit 2.12, significant difference. However paired t tests between soil sequences showed groups 1, 2, 3, 4 to have no significant difference, ANOVA on BECCS_C for soil cohorts 1, 2, 3, 4 at $\alpha$ 0.05 produced $F(3,292) = 0.166$, $p = 0.92$ (Fcrit 2.64). |
|---|---|

[a]There are no soil grid squares below 54.5 degN containing data for a 5th ranking soil cover.

[b]%CV is the same as for DM Yield, since BECCS is DMYield multiplied by a factor of 50% carbon and 90% efficiency.

Scotland has a high number of soil types per grid square. It also has moorland with rich peaty soils high in C, and these would not be used to grow $M \times g$, modelled output using the majority peat soils shows a reduction in soil C. When other non-peat arable soils which increase in soil C are aggregated with the C reduction occurring on the peat soils, the effects cancel each other out and very little change on soil C is seen. Due to the opposing effects on soil C from $M \times g$ growth, the peat and non-peat soils must be treated separately. Since we would only consider non-peaty soils for crop growth, our results reflect only the non-peat soils.

Unlike the climate cohorts, the soil sequences are meant to be combined, so instead of looking at outputs from all cohorts as we did with climate cohorts, we can assess if the combined soil sequences and percentage coverage result in a dataset of BECCS_C that is significantly different to the modelled output from merely using the majority soil coverage. BECCS_C is modelled for soil types from all soil sequence groups and these are combined with the percentage coverages of the soil type in the grid square. BECCS_C is dependent on dry matter yield and the dominant soil characteristic influential on yield is water carrying capacity. Lower coverage soil types from other soil sequence groups can result in a higher or lower BECCS_C projection than that resulting from the majority soil type; however, the majority soil accounts for over 50% of the coverage of grid squares and will be the dominant influence on the BECCS_C.

ANOVA revealed that the different soil sequence groups result in BECCS_C projections which, at $\alpha = 0.05$, is statistically different between groups (at $\alpha = 0.05$, $F(6,504) = 2.818$, $p = 0.01$, Fcrit 2.12). However post-hoc comparison using a $t$ test with Bonferroni correction indicated similar BECCS_C between soil sequences groups 1, 2, 3, 4 to have no significant difference (means 4.37, 4.38, 4.28, 4.28; SD 1.14, 1.22, 1.24, 1.24, respectively). Confirming this ANOVA on BECCS_C for climate cohorts 1, 2, 3, 4 at $\alpha = 0.05$ produced $F(3,292) = 0.166$, $p = 0.92$ (Fcrit 2.64).

# 4 | DISCUSSION

In Table 3, no significant differences were found between the three climate cohorts for the decadal changes, which is more pertinent to being members of a climate ensemble since dynamic change of climate rather than absolute values is more relevant to climate projections. The mean decade to decade change hides that initial decades experience an increase and post-2051 experiences decreases in dry matter yield and BECCS_C. This relates to an increasing temperature but relatively constant precipitation, which gives rise to a decade-to-decade increase in water deficit.

The weather is a driver of crop variance via water deficit (Frieler et al., 2017). Increased temperature influences increased biological turnover in the soil and together with decreased yields, soil C increases less over the consecutive decades from the initial increase 2011–2020. The combination of increasing water deficit and the slowing of soil C increase is a dangerous one. To retain water holding capacity, the soil requires increases in soil C which improve texture.

The decade-to-decade uncertainties of the variables are higher than the uncertainty between climate cohorts, corresponding to Pogson et al. (2012) using different datasets, who determined that inter-annual variation of met data was higher than between datasets.

Investigating higher temporal resolution annual yield revealed a close relationship with precipitation, also influenced by accumulated growing season temperature and severe drought stress. The risk of a lower or higher yield under the projected climate is relatively the same, with the probability of yields of under 10 t ha$^{-1}$ year$^{-1}$ or over 13 t ha$^{-1}$ year$^{-1}$ being about 1 in 3 years return period, but below or above these values the probability decreases exponentially.

Model sensitivities (Figure 2) highlight that the model is more sensitive to reduction in available water capacity, rather than climate, and it is likely that UK climate parameters do not create a water deficit at a critical level for crop growth as much as the water carrying capacity of the soil. The modelling process for chalk soil as described in the introduction will enhance uncertainty of water carrying capacity between chalk and other soils which is inherent in the uncertainty between applications of soil sequences. The MiscanFor model is sensitive to field capacity and wilting point, but it is the difference between the two, the water capacity, which affects growth processes, both may rise or fall together across different soils and water capacity variance between different soils groups has a lower spread of data around the mean (standard error 22.4) than its field capacity and wilting point.

Soil sequence 1 (soil type with majority coverage) gave a mean gridded BECCS_C of 4.37 t C ha$^{-1}$ year$^{-1}$ and a combined soil sequence output gave a mean gridded BECCS_C of 3.96 t C ha$^{-1}$ year$^{-1}$ (from BECCS_C projections of all soil types multiplied by their % coverage of the gridded area). ANOVA performed on these two groups of 73 annual BECCS_C 2008–2080 (area-means for England and Wales) revealed no significant difference, for an $\alpha$ of 0.05, $F(1,144) = 1.658$, $p = 0.2$, Fcrit = 3.9. Therefore, based on the annual area-means, it is statistically acceptable to use solely the majority soil data instead of all files when modelling output, which could save considerable time in processing and analysis.

Maps of BECCS_C resulting from the majority soil coverage and the combination of all ranked coverage soil types (Figure 6) indicate that the range of values are similar, except
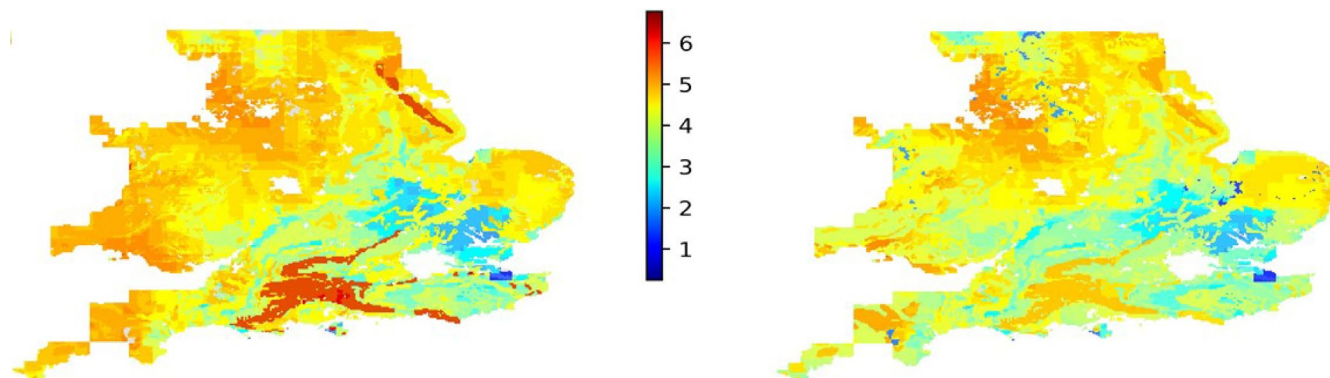
**FIGURE 6** BECCS_C projection 2008–2080 (t C ha$^{-1}$ year$^{-1}$), resulting from HWSD majority coverage soil (left), combined yield from percentage share of all HWSD ranked coverage soil types (right)

for small areas in sequence 1 which have majority chalk soil coverage. BECCS in these areas is improved by larger crop yields resulting from improved water storage. The combined results of BECCS_C will provide more spatial detail. The choice to use majority soil output or run all ranked soils and combined output comes down to the application of the BECCS_C, whether for locational detail or area-average projections.

We could have attempted an aggregated total for specific $M \times g$ growing areas, but this is difficult to do since we used future climate projections and the areas of $M \times g$ growth change with time. We used the average of England and Wales which includes some infeasible areas (industrial, urban, national parks), this lowers BECCS_C projections slightly erring on the cautious side and gives a higher standard error, which is suitable for worst-case scenario climate. This study could be repeated using a GIS mask to extract results under specific areas. Rather, the focus here was on changing members of the same source of climate and soil data while areas remained fixed.

## 4.1 | Uncertainty budget and combined uncertainty

Table 6 is a summary related to the various data options we have tested in this study, and serves as a summary of variability in the data we have used, and a reminder to be aware of the variability of modelled BECCS projections. The annual data support the finding that the climate cohorts create more variability between BECCS projections than soil cohorts. BECCS is related to yield, annual yield time series showed a close relationship with precipitation and accumulated degrees, whereas yield maps only partially coincided with areas of higher soil variability.

The %CV and the standard error are uncertainty measures; standard error also standardizes uncertainty for different size datasets. The values tell us about the uncertainty of an average and the different processes involved in producing them. To combine the uncertainties, we use

**TABLE 6** Uncertainty Budget of BECCS_C 2008–2080

| Source of BECCS_C, Uncertainty of the mean between groups | CV |
|---|---|
| RCP8.5 climate projection, 12 cohorts | 12.68% |
| HWSD soil dataset, 8 different cohorts | 5.27% |

summation in quadrature, which is also known as the root sum of the squares.

$$\sqrt{\left(a^2 + b^2 + \dots\right)}. \tag{2}$$

The %CVs are combined in quadrature, and %CVs can be compared and combined from different size datasets (US Dept. of Commerce—Bureau of Standards, 1961).

We can combine the uncertainty of the mean of the 12 cohort simulations (12.68%) and the uncertainty of the mean of the 8 soil cohort simulations (5.27%) by summation in quadrature to give us a CV of 13.73%, resulting from 12 inputs in mean temperature (10.86 degree C, 6.18% CV) and precipitation (1019 mm, 6.30% CV), and 8 inputs in field capacity (combined 556.8 mm, 35.3% CV) and wilting point (combined 322.8 mm, 50.4% CV).

Alternately, we can state the uncertainty in terms of the standard error. If the majority cover soil sequence is not statistically different from the combined values, the majority cover can be input to a model together with the 12 climate cohorts which would be a BECCS_C value averaged over England and Wales 2008–2080 of $3.99 \pm 0.14$ t C ha$^{-1}$ year$^{-1}$. Combining the results from soil files plus all 12 climate files using a summation in quadrature for standard errors, $3.99 \pm 0.14$ t C ha$^{-1}$ year$^{-1}$ for climate and $3.96 \pm 0.03$ t C ha$^{-1}$ year$^{-1}$ for weighted compound soil is a BECCS_C of $3.98 \pm 0.143$ t C ha$^{-1}$ year$^{-1}$. Applying the same method, combined projections for impacts on associated environmental effects are a mean seasonal water deficit of $65.79 \pm 4.27$ mm and mean soil C increase of $0.22 \pm 0.03$ t C ha$^{-1}$ year$^{-1}$.

Figure 7 shows the maximum no. of soil sequences within a 1 km grid square, displaying local soil variability for the UK. The straight line of the data across the south of England originates from the original HWSD dataset, external to our data processing. Regional standard error between yields simulated from the different soil sequences (Table 7) was calculated using a UK regions NUTS Level 1 shapefile (Office for National Statistics, 2018). The variability of yield between local soils sequences shows the south-east and east of England and the south-west having the highest variability, and corresponds to areas of high soil variability in Figure 7.

BECCS, water deficit and soil C standard errors resulting from combined climate and soil cohort variation are shown in
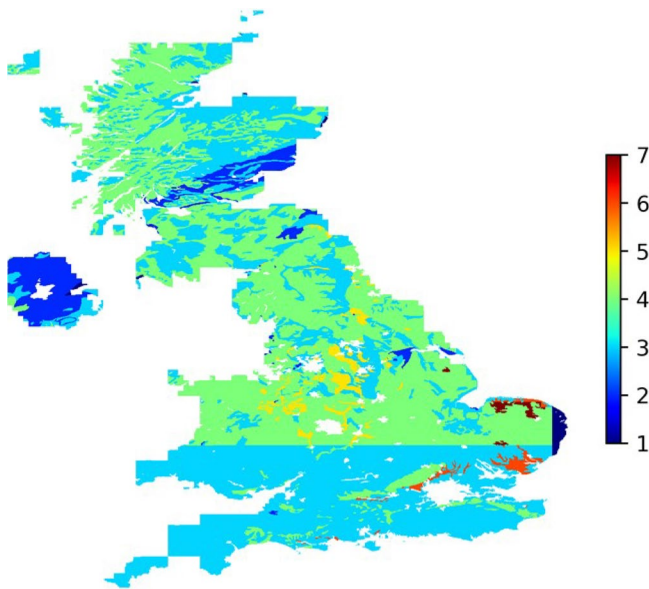


**FIGURE 7** Maximum no. of UK HWSD soil sequences in a 1 km grid square

**TABLE 7** Regional standard error of dry matter yield between local soil sequences

| UK regions | Regional mean Std Error of yield from local area-composite soil sequences |
| --- | --- |
| NE England | 0.140 |
| NW England | 0.099 |
| Yorkshire & The Humber | 0.278 |
| E Midlands | 0.600 |
| W Midlands | 0.352 |
| E England | 0.600 |
| SE England | 0.881 |
| SW England | 0.556 |
| Wales | 0.096 |
| Scotland | 0.088 |

Figure 8. Variation in BECCS (Figure 8a, dependant on dry matter yield variation) is concurrent with some areas showing a greater variability in soils shown in Figure 7. However, most parameters are likely to respond to local variation between loam soils and chalk soils which influence results via water capacity, rather than the maximum variability of soil types, and additionally to climate variation.

The energy of $M \times g$ bales is calculated as an energy yield of 18 GJ t$^{-1}$ of dry matter yield, minus latent heat of vaporization at 2.72 GJ t$^{-1}$ of moisture content (30% of dry matter biomass), minus fixed energy cost (5.64 GJ year$^{-1}$) of crop establishment, minus energy input 0.61 GJ t$^{-1}$ per dry matter yield, incorporating fertilizer, harvesting and transport (Hastings et al., 2017). This gives a feedstock bioenergy of 131 GJ t$^{-1}$ associated with a BECCS of 3.98 t C ha$^{-1}$ year$^{-1}$.

Comparing our findings against other studies, Hastings et al. (2014) modelled mean $M \times g$ yield using MiscanFor for all UK regions using Met Office UKCP09 A1B scenario climate projections (Jenkins et al., 2009) for medium emission and the HWSD sequence 1 dominant soil dataset. They obtained mean values of dry matter yield for England and Wales south of what they termed 'northern English counties' which gave values convertible to a BECCS_C of 4.43 t C ha$^{-1}$ year$^{-1}$ (2011) and 5.26 t C ha$^{-1}$ year$^{-1}$ (2051). This compares with our BECCS_C values of 4.51 t C ha$^{-1}$ year$^{-1}$ (2011) and 4.97 t C ha$^{-1}$ year$^{-1}$ (2051). We have used RCP8.5 worst-case climate projections which increase values of BECCS rapidly in the early part of the century, then slow the rate of increase mid-century as water deficits increase. Pogson et al. (2012) also found that when using different meteorological and soil datasets to determine $M \times g$ yields, it highlighted the significance of soil water parameters, and commented that this could become an issue in areas affected by climate change. They also made the important point that if datasets vary widely, and a model is calibrated while using a particular soil or climate dataset, it would be therefore be biased towards using that dataset so it is important to note the datasets used during calibration. In our study, the model sensitivity highlighted the importance of soil water capacity, and the climate projections showed a static precipitation with increasing temperature, leading to an increasing water deficit.

Hoffman et al. (2016) state that there is a bias inherent in aggregating crop yields over large areas. This results from aggregating input data which may be generated by averaging and sampling. In this study, yield (upon which BECCS_C is dependant) shows a distinctive east-west divide following higher precipitation totals trending to the west of the UK. Yawson et al. (2016) comment that UKCP09 climate influenced yields to increase more east to west then north to south, this was a consistent trend for this study for all climate cohorts.

Folberth et al. (2016) point out that estimated climate change effects on yield can be either negative or positive depending on the chosen soil type, and therefore soils have the
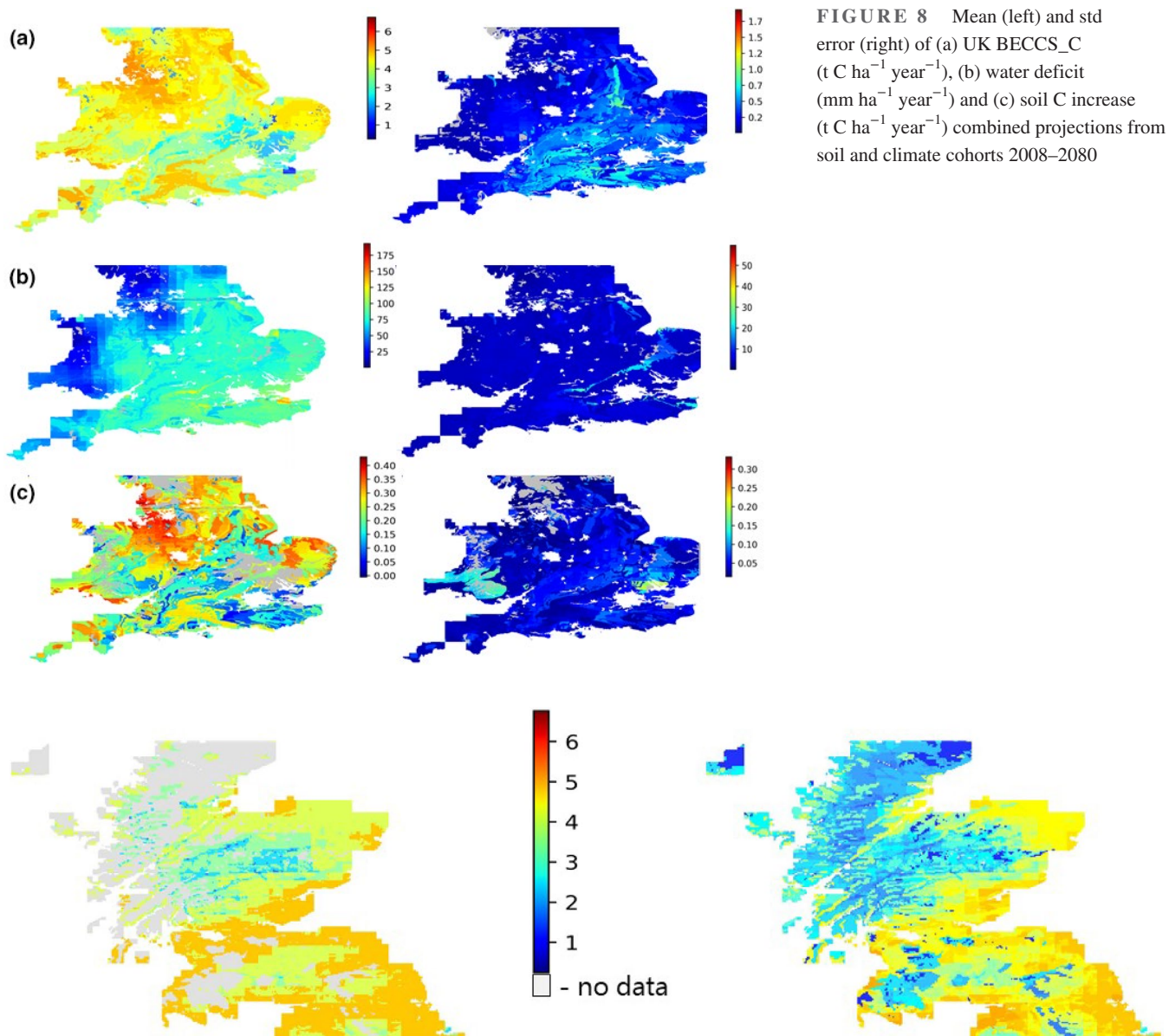
**FIGURE 8** Mean (left) and std error (right) of (a) UK BECCS_C (t C ha$^{-1}$ year$^{-1}$), (b) water deficit (mm ha$^{-1}$ year$^{-1}$) and (c) soil C increase (t C ha$^{-1}$ year$^{-1}$) combined projections from soil and climate cohorts 2008–2080



**FIGURE 9** BECCS_C projection 2008–80 (t C ha$^{-1}$ year$^{-1}$) north of current UK 54.5°N latitude threshold; HWSD majority coverage soil (left), combined yield from percentage share of all HWSD ranked coverage soil types (right)

capacity to either buffer or amplify these impacts. We found the buffering (or cancelling) effect to be particularly the case with changes in soil C on peat and non-peat soils, so did not use peat soils which would not be used to grow crops due to the loss of carbon.

This study is based on the current $M \times g$ growing area to give mean value parameters per grid square comparable with present crop growing conditions. It is not realistic to determine the uncertainty among Scottish areas when so many are inhospitable to crop growth. We recognize however that temperature projections increase throughout the 21st century, so we have included Scotland and England north of 54.5 degrees in the map of mean 2008–2080 BECCS projections (Figure 9) for the majority soil and all soil sequences to show that.

1. As the RCP8.5 projected climate warms, areas further north (northern England and southern Scotland lowlands show promise for $M \times g$ growth and associated BECCS, despite simulations incorporating loss for winter crop kill.
2. BECCS_C in Figure 9 shows a good correspondence with Figure 7, soil type variation. North and west Scotland is the Scottish region showing the largest difference between using a majority soil and the full soil sequence. The dominant soil is often classed as no soil or bare rock which produces a large area of missing data. When all soil sequences are included, we gain improved coverage of results.
3. Scotland has regions with complex soil combinations per grid square.

In summary, BECCS_C uncertainty from climate cohorts of the same climate ensemble database was larger than the uncertainty from soil cohorts of the same database. This result is supported in literature (Waha et al., 2015), and was also supported by the annual time-series data, displaying a close relationship with precipitation, accumulated temperature and drought events and a partial correspondence with maximal soil variability. This contrasts with the model sensitivity which is greater for available water capacity than the climate but also indicates the UK climate is not as limiting as the water capacity of the soil. BECCS_C projections from soil cohorts differed, although four out of seven were not significantly different. The predominant soil coverage was over 50% so the combined BECCS_C output of all soil sequences was not significantly different from the BECCS_C projection resulting from the soil of majority coverage.

The uncertainty of BECCS_C in this study using RCP8.5 climate ensemble and HWSD soil sequence data is relatively low in terms of standard error. Mean BECCS_C for England and Wales averaged over the current $M \times g$ production area is $3.98 \pm 0.14$ t C ha$^{-1}$ year$^{-1}$.

Policymakers and managers will review BECCS projections resulting from various models, climates and soils databases. Ensure that you know the uncertainty and its sources when reporting a modelled BECCS projection. The uncertainty of each should be included with the mean BECCS projection and the model, climate and soil database used. Uncertainty varies between models inherent in model parameterization, sensitivity, calibration bias, validation, but it also varies between climate databases and soils databases.

While this study analyses projected BECCS uncertainty resulting from modelled output, we recognize that there is a great deal of uncertainty surrounding BECCS related to socio-economics, financial viability and the aggregation in area grown, we have studied but one part of the whole. The message from this study is for non-modellers to be aware of variation in a modelled BECCS projection even from climate and soil databases with the same source.

## ACKNOWLEDGEMENTS

## DATA AVAILABILITY STATEMENT
The data that support the findings of this study are available from the corresponding author upon reasonable request.

## ORCID
*Anita Shepherd* https://orcid.org/0000-0003-1902-5147
*Astley Hastings* https://orcid.org/0000-0001-9863-7613

## REFERENCES
Albanito, F., Hastings, A., Fitton, N., Richards, M., Martin, M., MacDowell, N., Bell, D., Taylor, S. C., Butnar, I., Li, P.-H., Slade, R., & Smith, P. (2019). Mitigation potential and environmental impact of centralized versus distributed BECCS with domestic biomass production in Great Britain. *GCB Bioenergy*, *11*, 1234–1252. https://doi.org/10.1111/gcbb.12630

Aodha, L., & Edmonds, B. (2017). Some pitfalls to beware when applying models to issues of policy relevance. In B. Edmonds & R. Meyer (Eds.), *Simulating social complexity. Understanding complex systems*. Springer. https://doi.org/10.1007/978-3-319-66948-9_29

Bosatta, E., & Agren, G. I. (1985). Theoretical analysis of decomposition of heterogeneous substrates. *Soil Biology and Biochemistry*, *17*, 601–610.

Bosatta, E., & Agren, G. I. (1991). Dynamics and carbon and nitrogen in the organic matter of the soil: A generic theory. *The American Naturalist*, *138*, 227–245. https://doi.org/10.1016/j.geoderma.2017.11.025

Campbell, G. S. (1985). *Soil physics with BASIC, transport models for soil-plant systems* (Vol. 14, 1st ed.), Elsevier. eBook ISBN: 9780080869827. https://www.elsevier.com/books/soil-physics-with-basic/campbell/978-0-444-42557-7

Committee on Climate Change. (2018). Biomass in a low carbon economy. © Committee on Climate Change Copyright. https://www.theccc.org.uk/wpcontent/uploads/2018/11/Biomass-in-a-low-carbon-economy-CCC-2018.pdf

Committee on Climate Change. (2020). Net Zero - The UK's contribution to stopping global warming. © Committee on Climate Change Copyright https://www.theccc.org.uk/publication/net-zero-the-uks-contribution-to-stopping-global-warming/

Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2012). Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*, *38*(3–4), 527–546. https://doi.org/10.1007/s00382-010-0977-x

Dondini, M., Hastings, A., Saiz, G., Jones, M. B., & Smith, P. (2009). The potential of Miscanthus to sequester carbon in soils: Comparing field measurements in Carlow, Ireland to model predictions. *GCB Bioenergy*, *1*, 413–442. https://doi.org/10.1111/j.1757-1707.2010.01033.x

Fischer, G. F., Nachtergaele, F. O., Prieler, S., van Velthuizen, H. T., Verelst, L., & Wiberg, D. (2008). *Global Agro-ecological Zones Assessment for Agriculture (GAEZ 2008)*. FAO, IIASA.

Folberth, C., Skalský, R., Moltchanova, E., Balkovič, J., Azevedo, L. B., Obersteiner, M., & van der Velde, M. (2016). Uncertainty in soil data can outweigh climate impact signals in global crop yield simulations. *Nature Communication*, *7*, 11872. https://doi.org/10.1038/ncomms11872

Frieler, K., Schauberger, B., Arneth, A., Balkovič, J., Chryssanthacopoulos, J., Deryng, D., Elliott, J., Folberth, C., Khabarov, N., Müller, C., Olin, S., Pugh, T. A. M., Schaphoff, S., Schewe, J., Schmid, E., Warszawski, L., & Levermann, A. (2017). Understanding the weather signal in national crop-yield variability. *Earth's Future*, *5*, 605–616. https://doi.org/10.1002/2016EF000525

Hastings, A., Clifton-Brown, J., Wattenbach, M., Mitchell, C. P., & 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. https://doi.org/10.1111/j.1757-1707.2009.01007.x

Hastings, A., Mos, M., Yesufu, J. A., McCalmont, J., Schwarz, K., Shafei, R., Ashman, C., Nunn, C., Schuele, H., Cosentino, S., Scalici, G., Scordia, D., Wagner, M., & Clifton-Brown, J. (2017). Economic and environmental assessment of seed and rhizome propagated Miscanthus in the UK. *Frontiers of Plant Science*, *8*, 1058. https://doi.org/10.3389/fpls.2017.01058

Hastings, A., Tallis, M. J., Casella, E., Matthews, R. W., Henshall, P. A., Milner, S., Smith, P., & Taylor, G. (2014). The technical potential of Great Britain to produce ligno-cellulosic biomass for bioenergy in current and future climates. *GCB Bioenergy*, *6*, 108–122. https://doi.org/10.1111/gcbb.12103

Hellinga, B. R. (1998) *Requirements for the calibration of traffic simulation models. Proceedings of the Canadian Society for Civil Engineering* (Vol. *IVB*, pp. 211–222). http://www.civil.uwaterloo.ca/bhellinga/publications/Publications/CSCE-1998-Calibration.PDF

Hoffmann, H., Zhao, G., Asseng, S., Bindi, M., Biernath, C., Constantin, J., Coucheney, E., Dechow, R., Doro, L., Eckersten, H., Gaiser, T., Grosz, B., Heinlein, F., Kassie, B. T., Kersebaum, K.-C., Klein, C., Kuhnert, M., Lewan, E., Moriondo, M., … Ewert, F. (2016). Impact of spatial soil and climate input data aggregation on regional yield simulations. *PLoS One*, *11*(4), e0151782. https://doi.org/10.1371/journal.pone.0151782

Jenkins, G., Murphy, J., Sexton, D., Lowe, J., Jones, P., & Kilsby, C. (2009). *UK climate projections. Briefing report*. Met Office Hadley Centre. ISBN 978-1-906360-04-7.

Mauritzen, C., Zivkovic, T., & Veldore, V. (2017). On the relationship between climate sensitivity and modelling uncertainty. *Tellus A: Dynamic Meteorology and Oceanography*, *69*(1). https://doi.org/10.1080/16000870.2017.1327765

McCalmont, J. P., Hastings, A., McNamara, N. P., Richter, G. M., Robson, P., Donnison, I. S., & Clifton-Brown, J. (2017). Environmental costs and benefits of growing Miscanthus for bioenergy in the UK. *GCB Bioenergy*, *9*(3), 489–507. https://doi.org/10.1111/gcbb.12294

Met Office Hadley Centre. (2018). UKCP18 Regional Projections on a 12km grid over the UK for 1980–2080. Centre for Environmental Data Analysis. https://catalogue.ceda.ac.uk/uuid/589211abeb844070a95d061c8cc7f604

Office for National Statistics, NUTS Level 1 shapefile. (January 2018). Boundaries. Reference date - publication: 2017-07-31. https://geoportal.statistics.gov.uk/datasets/nuts-level-1-january-2018-generalised-clipped-boundaries-in-the-united-kingdom

Pogson, M., Hastings, A., & Smith, P. (2012). Sensitivity of crop model predictions to entire meteorological and soil input datasets highlights vulnerability to drought. *Environmental Modelling and Software – ENVSOFT*, *29*, 37–43. https://doi.org/10.1016/j.envsoft.2011.10.008

POST. (2020). Postnote No. 618. Bioenergy with carbon capture and storage (BECCS). House of Parliament, Parliamentary Office of Science & Technology. https://post.parliament.uk/research-briefings/post-pn-0618/

Qian, B., Jing, Q., Smith, W., Grant, B., Cannon, A., & Zhang, X. (2020). Quantifying the uncertainty introduced by internal climate variability in projections of Canadian crop production. *Environmental Research Letters*, *15*(7). https://doi.org/10.1088/1748-9326/ab88fc

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., & Rafaj, P. (2011). RCP 8.5 – A scenario of comparatively high greenhouse gas emissions. *Climatic Change*, *109*(1-2), 33–57. https://doi.org/10.1007/s10584-011-0149-y

Rogelj, J., Meinshausen, M., & Knutti, R. (2012). Global warming under old and new scenarios using IPCC climate sensitivity range estimates. *Nature Climate Change*, *2*, 248–253. https://doi.org/10.1038/nclimate1385

Rounsevell, M., Reginster, I., Araújo, M. B., Carter, T. R., Dendoncker, N., Ewert, F., House, J. I., Kankaanpää, S., Leemans, R., Metzger, M. J., Schmit, C., Smith, P., & Tuck, G. (2006). A coherent set of future land use change scenarios for Europe. *Agriculture, Ecosystems and Environment*, *114*, 57–68.

Sexton, D. M. H., Karmalkar, A., Murphy, J., Booth, B., Regayre, L., & Johnson, J. (2019). The elicitation of distributions of parameters in HadGEM3 versions GA4 and GA7 for use in perturbed parameter ensembles, Technical Note 101, Met Office. https://digital.nmla.metoffice.gov.uk/digitalFile_70d009c8-e04e-449b-98a2-ada3be97167d/

Shepherd, A., Atuhaire, W., Wu, L., Hogan, D., Dunn, R., & Cardenas, L. (2017). Historic record of pasture soil water and the influence of the North Atlantic Oscillation in south-west England. *Hydrology Research*, *48*, 277–294. https://doi.org/10.2166/nh.2016.195

Shepherd, A., Clifton-Brown, J., Kam, J., Buckby, S., & Hastings, A. (2020). Commercial experience with Miscanthus crops: Establishment, yields and environmental observations. *GCB Bioenergy*, *12*(7), 510–523. https://doi.org/10.1111/gcbb.12690

Shepherd, A., Littleton, E., Clifton-Brown, J., Martin, M., & Hastings, A. (2020). Projections of global and UK bioenergy potential from *Miscanthus × giganteus* – Feedstock yield, carbon cycling and electricity generation in the 21st century. *GCB Bioenergy*, *12*, 287–305. https://doi.org/10.1111/gcbb.12671

Sims, R. E. H., Hastings, A., Schlamadinger, B., Taylor, G., & Smith, P. (2006). Energy crops: current status and future prospects. *Global Change Biology*, *12*, 2054–2076. https://doi.org/10.1111/j.1365-2486.2006.01163.x

Smith, J., & Smith, P. (2007). *Introduction to environmental modelling*. Oxford University Press. 180 pp. ISBN 978-0-19-927206-8.

Smith, P., Davis, S. J., Creutzig, F., Fuss, S., Minx, J., Gabrielle, B., Kato, E., Jackson, R. B., Cowie, A., Kriegler, E., van Vuuren, D. P., Rogelj, J., Ciais, P., Milne, J., Canadell, J. G., McCollum, D., Peters, G., Andrew, R., Krey, V., … Yongsung, C. (2016). Biophysical and economic limits to negative $CO_2$ emissions. *Nature Climate Change*, *6*(1), 42–50. https://doi.org/10.1038/nclimate2870

Surendran Nair, S., Kang, S., Zhang, X., Miguez, F. E., Izaurralde, R. C., Post, W. M., & Wullschleger, S. D. (2012). Bioenergy crop models: Descriptions, data requirements, and future challenges. *GCB Bioenergy*, *4*(6), 620–633. https://doi.org/10.1111/j.1757-1707.2012.01166.x

United States department of Commerce. (1961). Publications of the National Bureau of Standards July 1, 1957, to June 30, 1960, National Bureau of Standards. Miscellaneous Publication 240. Issued April 3, 1961. Published by United States Government printing office, Washington DC.

Uusitalo, L., Lehikoinen, A., Helle, I., & Myrberg, K. (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software*, *63*, 24–31. https://doi.org/10.1016/j.envsoft.2014.09.017

Waha, K., Huth, N., Carberry, P., & Wang, E. (2015). How model and input uncertainty impact maize yield simulations in West Africa. *Environmental Research Letters*, *10*(2), 024017. https://doi.org/10.1088/1748-9326/10/2/024017

Wieder, R., Boehnert, J., Bonan, G. B., & Langseth, M. (2014) Regridded harmonized world soil database v1.2. data set. http://daac.ornl.gov from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA. 2019. https://doi.org/10.3334/ORNLDAAC/1247

Yawson, D. O., Ball, T., Adu, M. O., Mohan, S., Mulholland, B. J., & White, P. J. (2016). Simulated regional yields of spring barley in the United Kingdom under projected climate change. *Climate*, *4*(4), 54. https://doi.org/10.3390/cli4040054