Science of the Total Environment 861 (2023) 160618

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Robust spatial estimates of biomass carbon on farms

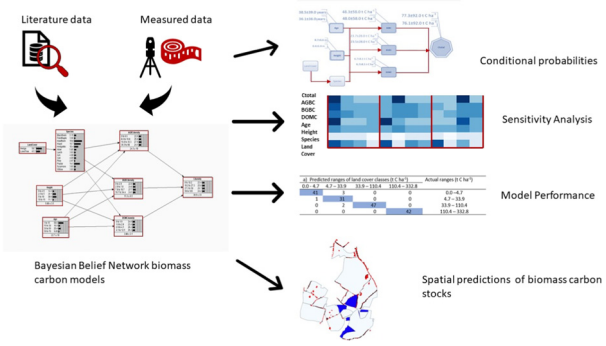
Styliani Beka *, Paul J. Burgess, Ron Corstanje

Cranfield University, Cranfield, Bedfordshire MK43 0AL, UK

HIGHLIGHTS

- Integrated spatial approach, for the quantification of total biomass carbon stocks of land cover and landscape features.
- Results demonstrate the possibility of developing and using remote integrated approach for farm-scale total biomass carbon.
- Highest achievable prediction accuracy attained by applying both current literature and measured information.
- Successful predictions of test values, with error rates of 6.7% and 4.3% for the land cover and landscape features, respectively.
- Spatial presentation and calculation of variable biomass carbon stocks along the length of hedges and within woodlands using lidar data.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Jacopo Bacenetti

Keywords:
Biomass carbon
Spatial variation
Integrated method
Land cover
Landscape features

ABSTRACT

The drive for farm businesses to move towards net zero greenhouse gas emissions means that there is a need to develop robust methods to quantify the amount of biomass carbon (C) on farms. Direct measurements can be destructive and time-consuming and some prediction methods provide no assessment of uncertainty. This study describes the development, validation, and use of an integrated spatial approach, including the use of lidar data, and Bayesian Belief Networks (BBNs) to quantify total biomass carbon stocks (C_{total}) of i) land cover and ii) landscape features such as hedges and lone trees for five case study sites in lowland England. The results demonstrated that it was possible to develop and use a remote integrated approach to estimate biomass carbon at a farm scale. The highest achievable prediction accuracy was attained from models using the variables AGBC, BGBC, DOMC, age, height, species and land cover, derived from measured information and from literature review. The two BBN models successfully predicted the test values of the total biomass carbon with propagated error rates of 6.7 % and 4.3 % for the land cover and landscape features respectively. These error rates were lower than in other studies indicating that the seven predictors are strong determinants of biomass carbon. The lidar data also enabled the spatial presentation and calculation of the variable C stocks along the length of hedges and within woodlands.

1. Introduction

Efforts to mitigate climate change have been enacted by many countries across the world with the aim of achieving net-zero greenhouse gas (GHG) emissions by 2050 (Brandão et al., 2013; Smith et al., 2012; van Soest et al., 2021). Achieving this will require management changes across each sector

* Corresponding author.
E-mail address: s.beka@cranfield.ac.uk (S. Beka).

http://dx.doi.org/10.1016/j.scitotenv.2022.160618
Received 10 August 2022; Received in revised form 27 November 2022; Accepted 27 November 2022
Available online 30 November 2022

of the economy. This is particularly the case in agriculture, which was the source of about 10.7 billion tonnes of carbon dioxide equivalent (Gt CO_2 eq) emissions in 2019, and in the long-term, it could provide a net sink (Food and Agriculture Organization of the United Nations, 2021).

Since management practices under the control of farmers can reduce GHG emissions at a farm level (Burbi et al., 2016), a key policy measure is to incentivise land managers to store more carbon (C) on their farms and reduce emissions (Ziegler et al., 2016; Petrokofsky et al., 2012). For these policies to be successful, robust and accurate information on the current C stocks and methods to quantify the amount of C that has been or will be stored are required. Ideally, the quantified baseline inventories will be developed according to industry-accepted best practices and will be auditable (World Business Council for Sustainable Development and World Resources Institute, 2020). Physical sampling of soil and biomass C might provide the most accurate assessments, yet such methods can be destructive, expensive and/or labour intensive (Rosenstock et al., 2016). Most approaches to estimate C stocks using either a combination of models (e.g. species specific allometric equations) and/or farm landscape features (e.g. woodland, hedges, fields, etc.). These methods can contain considerable uncertainty as features on farms are often heterogenous in composition (e.g. species) and spatial distribution.

There is a substantive body of research on approaches using generic C storage values (Adger and Subak, 1996; Cannell and Milne, 1995; Cantarello et al., 2011; Cruickshank et al., 2000; Dewar and Cannell, 1992; Gregory et al., 1995; IPCC, 2019; Jian et al., 2020) and highly detailed emission factor tools and calculators (Campos, 2020; Carlson et al., 2017; Janzen et al., 2006; Stöckle et al., 2001; Sustainable Food Lab et al., 2011; Ziegler et al., 2016). However, this generic approach tends to present a very broad range between the minimum and maximum values, often without defining the levels of uncertainty. Similarly, the use of farm C calculators can be labour, time and cost-consuming and too complicated for some farmers to use. Additionally, the residual uncertainty surrounding the validity of carbon calculators is sometimes hard to determine. Moreover, farmers need tools that can incorporate changing technical, social, and economic considerations while being able to specify the level of certainty in the C estimates. Here we explore a modelling approach that is capable of incorporating biological information (species, age, height) with landscape features which can be obtained through detailed (Lidar) and less detailed (hyperspectral remote sensed) methods.

Given the above, farmers need tools that can be adapted to changing technical, social, and economic environments while being able to specify the level of certainty in the carbon estimates. Additionally, ways to remotely measure the variables of interest are needed. This study determines the potential of accounting for farm biomass C stocks with a level of uncertainty and examines the consequences for policy, payments and land users. This was achieved by developing two Bayesian Belief Network (BBN) models to quantify biomass carbon stocks and their uncertainty at a farm scale.

2. Materials and methods

2.1. Methodology formulation

Probabilistic Bayesian Belief Networks (BBNs) provide a way of understanding social-environment interactions where complex sets of geospatially interdependent variables can affect environmental outcomes and consequently human decisions on landscape management (Grafius et al., 2019; Karimi et al., 2021). The capability of BBNs to describe complex systems (Borsuk et al., 2004; Heckerman, 1997; Pearl, 1988; Taalab et al., 2015a, 2015b) potentially provide an approach to quantifying biomass stocks and their certainty for a situation where we typically have incomplete knowledge and information (Hassall et al., 2019; Korb and Nicholson, 2010). Moreover, BBNs when combined with Geographical Information Systems (GIS) can be used to guide on-farm decisions by providing an idea of the probability of events, defining key variables and conditional dependencies, and identifying how these could affect the outcomes (Grafius et al., 2019). The C stored in a farm can be affected by many factors (Krauss et al., 2022; Branca et al., 2013), however, conditional dependencies can assist in narrowing these down to the ones affecting most the parameter that needs to be predicted (Taalab et al., 2015a, 2015b).

On-farm carbon storage can occur as soil carbon, plant debris and in the form of biomass of relatively homogenous land covers and landscape features. In this study, a BBN modelling approach was used for predicting the total amount of biomass carbon stocks of i) homogeneous land covers such as arable crops, woodland, and grassland, and ii) landscape features such as lone trees and hedges. According to the UK's National Forest Inventory (NFI) (National Forest Inventory Forest Research, 2017) lone trees are comprised of trees where their canopy does not touch any other tree and have a height of 2 or more meters. These features have proven to store significantly more C than managed hedges (Falloon et al., 2004) and can be found as part of agroforestry treatments planted either as line of trees (García de Jalón et al., 2018) or planted in a specified distance from each other (Upson et al., 2016).

The methodological approach involved five steps as shown in Fig. 1. Initially, key influence predictor variables were identified from a literature search, followed by the construction of the BBN models using those literature values. We then quantified the conditional probabilities before (the naive model) and after (the informed model) adding some additional measured values. Lastly, we used the informed model to spatially predict the total biomass carbon stocks across five case studies. A more detailed description of the steps is provided in the following sections.

2.2. Identification of 'state' values

In the first step, we analysed published studies to identify the variables that were reported to affect the amount of biomass carbon. The key variables identified were land cover, species, age, height, above-ground biomass carbon, below-ground biomass carbon, and dead organic matter carbon. The dead organic matter pool is comprised of dead wood and litter (IPCC, 2019).

Secondly, we used two search engines (Scopus, and Google Scholar) and a keyword approach to inform the BBNs and in total we retrieved 43 studies. The literature was focused on UK studies and we identified eight broad types of homogeneous land cover using the classification of UKCEH (Centre for Ecology and Hydrology, 2020). These were: arable land, improved grassland, semi-natural grassland, broadleaf woodland, coniferous woodland, mixed woodland, mountain heath and bog, and shrubland. Additionally, we completed a similar literature search for the landscape features which were classified as hedgerows or lone trees (Brewer et al., 2017).

Within the land cover types, subdivisions were sometimes included down to species levels (Appendix Table A.5 found in the Appendix). Arable land cover was subdivided into 9 categories, while the broadleaf and coniferous woodland species were separated into 10 and 8 categories. Respectively. The improved grassland comprised of 3 subdivisions while the semi-natural grassland, mixed woodland, shrubland, mountain heath and bog covers were not subdivided. Lastly, for the landscape features, the hedgerow class was subdivided into 6 categories and the lone trees were classified into 9.

For each land cover and landscape feature, values were derived for the species, age, height, aboveground biomass carbon stock (AGBC), belowground biomass carbon stock (BGBC), dead organic matter carbon stock (DOMC) and total biomass carbon stock (C_{total}). When a value for a variable was unknown, they were calculated using equations from Table A1 found in the appendix. All the values, their descriptive properties and their references were inserted into Microsoft Excel spreadsheets (Microsoft Office, 2022), and were subsequently converted into the same format in order to be used as 'cases' in the next step.

2.3. Construction of the Bayesian Belief Network models using literature values

We then developed predictive knowledge-based BBN probabilistic models using Netica software (Norsys Software Corp., 2022) for the

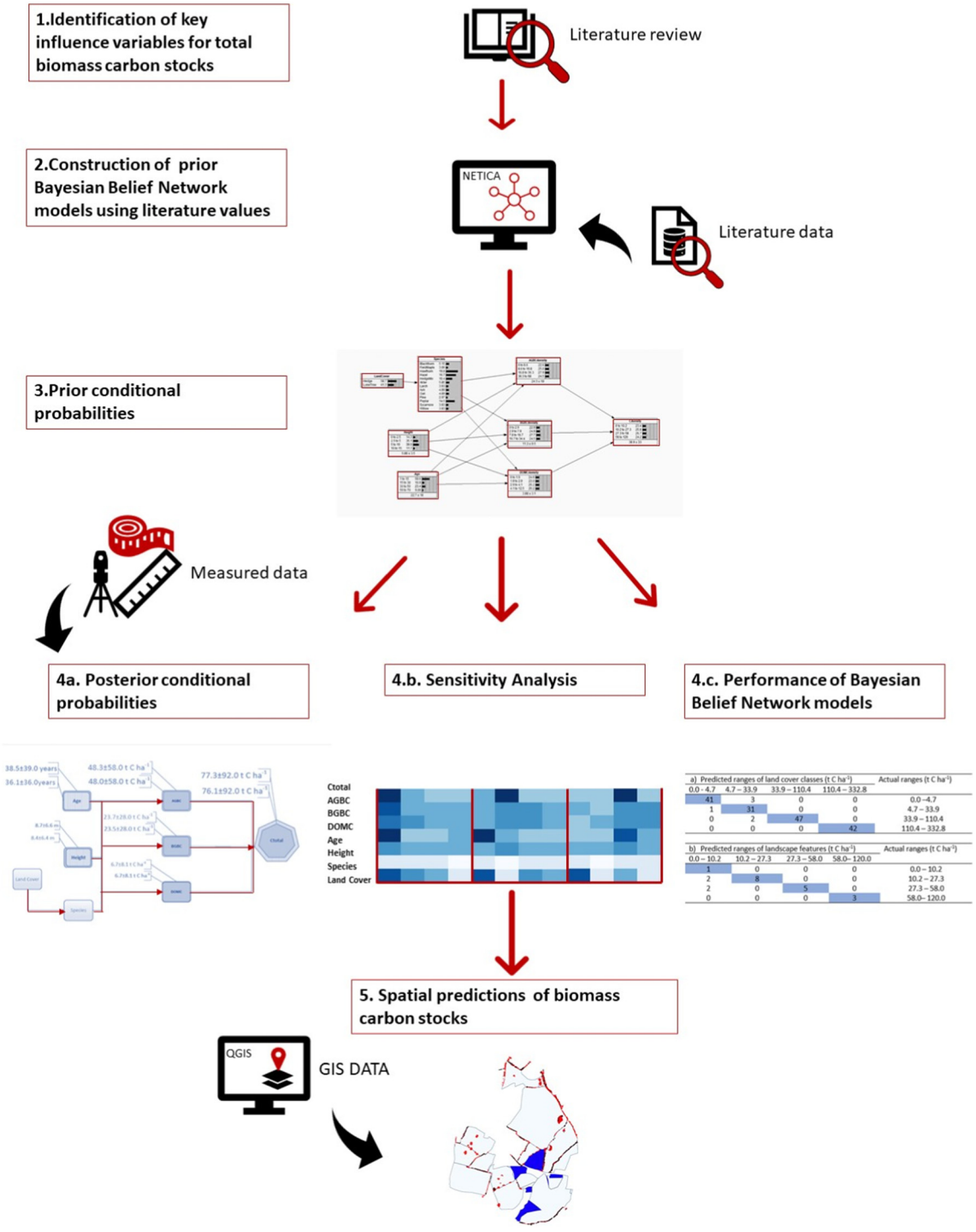


Fig. 1. Schematic representation of the methodological approach.

homogeneous land covers and the landscape features. These graphical models link a set of nodes called variables to a set of directed edges that represent the conditional probabilities between the nodes (Taalab et al., 2015a). In order to derive the structure of the probabilistic models, relationships between the variables were established as presented in Eqs. (1)–(5). The species variable was a function (f) of the land cover as seen in Eq. (1), while the AGBC, BGBC and DOMC were a function of the species, age and height (Eqs. (2)–(4)). Lastly, C_{total} was a function of the three forms of biomass carbon (Eq. (5)).

$$Species = f\,(Land\ cover) \tag{1}$$

$$AGBC = f\,(species, age, height) \tag{2}$$

$$BGBC = f\,(species, age, height) \tag{3}$$

$$DOMC = f\,(species, age, height) \tag{4}$$

$$C_{total} = f\,(AGBC, BGBC, DOMC) \tag{5}$$

The influence network for the homogeneous land covers contained seven variables as predictor nodes i.e., land cover type, species class, age, height, AGBC, BGBC, and DOMC. Each node contained four bins as it is

considered good practice (Constantinou et al., 2016; Marcot, 2012). The total biomass carbon stock (C_{total}) was selected as the response variable node. The influence network for the landscape features contained similar predictor and response variable nodes, although the land cover type node contained only two states (hedgerow or lone trees) instead of the CEH land cover classes.

Using only literature data, prior conditional probabilities were defined for "naive" models for the homogeneous land cover and the landscape features. For the naive models, 576 literature-derived values were used for land cover and 68 values for the landscape features model.

2.4. Addition of measured values

The next stage was to include additional measurement inputs of land cover and landscape features from five study areas (Fig. 2) to create an "informed" model. These included measurements of all seven predictors together with C_{total} measurements. The measured values were derived from five study areas all in lowland England identified by the names: Silsoe (S), Clapham Park (CP), Hanhill Manor Farm (HM), Crowmarsh Battle Farm (CB), and Elm Farm (E).

The site at Silsoe comprised an arable field planted with widely-spaced (6.4 m × 10 m) poplar trees, which 19 years after planting, measured in 2011, had a grass understorey (Upson and Burgess, 2013). The measurements at Clapham Park, taken in 2012, comprised a 14-year-old broadleaf woodland, a grassland area at least 50 years old, and a silvopasture agroforestry system with 14-year trees planted on grassland (Upson et al., 2016). At Harnhill Manor Farm measurements during the winter of 2013 were taken from three 12–18-year-old hawthorn and blackthorn hedges, and a 30 m adjacent area which included a grassland field margin and a small part of the arable land with a mix of forbs and legumes (Axe et al., 2017). At Crowmarsh Battle Farm, measurements were taken from five mixed 10-year-old hawthorn and mixed hedgerows in 2012 (Alexa Varah, personal communication 2021). Lastly, at Elm Farm, two blackthorn, two hawthorn and two hazel hedgerows aged 15–40 years were measured in 2014 (Westaway and Smith, 2020).

In total, there were an additional 5 measured values of land cover and 26 measured values for landscape features. The 31 measured values were used to update the parameter distribution and predict the posterior probabilities by the application of Bayes' Theorem (Eq. (6)).

$$p\left(\frac{\theta}{D}\right) = c\, p\left(\frac{D}{\theta}\right) p(\vartheta) \tag{6}$$

where $p\left(\frac{\theta}{D}\right)$ is the posterior probability density function (pdf) for the model's parameters ϑ after incorporating the new measured information D, $p(\vartheta)$ is the prior pdf from the literature information for ϑ before the arrival of the new information D. The likelihood of D for given values of ϑ is $p\left(\frac{D}{\theta}\right)$ and c is a proportionality constant (Koller and Friedman, 2009).

2.5. Performance of BBN models and sensitivity testing

For both informed models, the predictors' influence on the total biomass C stock (C_{total}) was examined by developing heat maps of the probabilistic associations. Output maps of the predicted total carbon across the five case studies were also generated using the QGIS 3.28 software environment (QGIS, 2022).

The performance of the models was then assessed using the "Test with cases" feature within Netica. The feature can be used to determine the performance of the two models by comparing the total carbon stored in the biomass predicted by the informed BBN models to the "true" value within a confusion matrix. An error rate was defined for each network as the percentage of the cases where the measured value occurred outside of the range of the allocated "bin" (Norsys Software Corp., 2022). Although this process does not validate the results in the same way as determined in modelling, due to the Bayesian probabilistic inference base (Grafius et al., 2019), it provides a measure of confidence similar to model predictions.

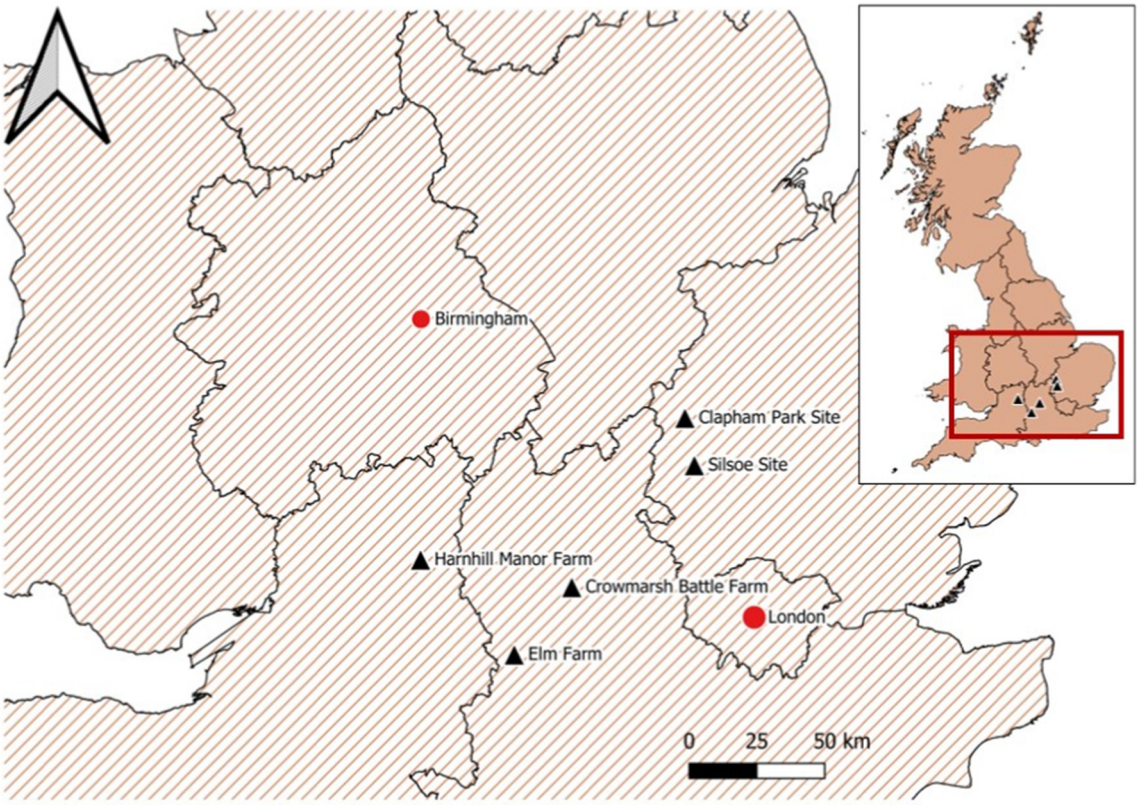


Fig. 2. Map showing the location of the five study areas ▲ where measurements of the carbon biomass, dry biomass, dimensions, age and height of landscape features were collected. Basemap presenting Great Britain downloaded from Digimap (2022). Sites created manually in QGIS using information provided by the studies.

To rank the influence of the seven predictors on C_{total}, their sensitivities and correlations were assessed using the sensitivity feature of the Netica software (Norsys Software Corp., 2022).

2.6. Spatial accounting approach of biomass carbon stocks

The prior naïve model and the posterior informed model were used to estimate the total biomass carbon stocks of the study areas in order to identify the advantages of incorporating measured data.

2.6.1. Homogeneous land cover input data

A mixture of raster and vector input data was used to depict the estimated total carbon stored in the biomass of the homogeneous land cover features found in the five case study sites. In more detail, vector data in the form of field delineations from the Rural Payments Agency in the UK, the 2015 land cover dataset (Centre for Ecology and Hydrology, 2017) and raster data in the form of the 1 m resolution Digital Surface Model (DSM) and a Digital Terrain Model (DTM) (Department for Environment Food & Rural Affairs, 2022) were collected for each study area. During the first step, the field parcel information was combined in the QGIS 3.28 software environment (QGIS, 2022) with a land cover shapefile. This resulted in the field parcels being categorised in one of the land cover states of the BBN land cover node. The age and species of the homogeneous land cover species present at each land parcel were taken from the information provided from the sampling of the areas. The age of the arable crops was assumed to be 1 year. The age of the grassland (3–50 years) was determined using aerial images from previous years from Google Earth (Google Earth Pro, 2022). In order to derive the height data for land cover features such as the woodland, the DTM was subtracted from the DSM and a raster with the height was derived. This was then segmented using the ENVI Version 5.6 (ENVI, 2022) software and was overlayed to the land cover/parcel vectors to classify the different height objects in land covers. Following, the vector shapefile with the segments was clipped to the boundaries of each field parcel and the mean height of the segments that fell within each field parcel was derived (height BNN predictor variable). The relationship between the aboveground biomass for all woodland ($AGB_{allwood}$; t ha^{-1}) and the mean canopy height (H_{mean}; m), was determined using Eq. (7) from Lefsky et al. (2002) which had an R^2 of 84 %.

$$AGB_{allwood} = 0.378\ H_{mean}{}^2 \quad (7)$$

Lastly, the predicted range of C_{total} was spatially depicted in QGIS for each study area.

2.6.2. Landscape features input data

The height of landscape features such as hedges and lone trees and their spatial location were derived using the method described for homogeneous land covers. The height segments were overlaid with the parcel boundaries and everything on the edge of each parcel boundary was classified as a hedge, while everything located at the parcel interior was defined as a lone tree. The age, species, and C stored in the biomass were obtained using a similar method as in the homogeneous land cover features and the predicted total carbon range was again spatially depicted.

3. Results

3.1. Bayesian uncertainty quantification and predictor sensitivity

For the BBN of the homogeneous land cover classes, a sensitivity analysis showed that the most influential predictor variables in descending order were AGBC, DOMC, BGBC, height, species, age, and land cover. For the landscape features the three predictor variables with the most significant influence were the three forms of biomass C, followed by age, species, height, and land cover.

Fig. 3 presents the five continuous variables of the two BBN models. The prior probability values, indicated in red in Fig. 3, present the mean value of each one of the naive model's parameters together with the uncertainty expressed as standard deviation (sd) associated with each prediction using just literature values. After incorporating the measured information into the naive models, an updated informed model and newly updated uncertainties were derived as indicated in black in Fig. 3. A more detailed description of the values can be found in Tables A1 and A2 in the Appendix. The combination of literature and additional measured information seemed to marginally improve the precision of estimation of the mean value of the total biomass C stock, as the mean value increased and the standard deviation stayed the same. In terms of the landscape features, the addition of measured values also increased the precision of the estimated total C stock value, as the mean value increased and the standard deviation decreased. Similarly, the precision of the estimate of the three forms of biomass C stock also improved as the values increased but the standard deviation stayed the same.

3.2. Drivers of carbon density

Using the posterior models, heat maps were generated to describe the probabilistic associations between the predictor variables and the predicted total biomass C stock (C_{total}) for the homogeneous land cover classes (Fig. 4a) and the landscape features (Fig. 4b). The conditional probabilities of an outcome given a set of parent node states are presented with shades of blue, with low and high probabilities indicated in light and dark blue respectively. For the homogeneous land cover features, although there were more than four states on the land cover and species nodes, only the four most common land covers are presented in the heatmaps. The four probabilistic ranges that the C_{total} can take are referred to as “bins”.

For the homogeneous land covers, Bin 1 comprised 0–4.7 t C ha^{-1}, and Bins 2, 3, and 4 comprised 4.7–33.9, 33.9–110.4, and 110.4–332.8 t C ha^{-1} respectively. The heat map shows that low values of C_{total} were associated with arable land (AR) and that high values of C_{total} were associated with high values of aboveground biomass carbon and age and with broadleaf woodland (Fig. 4a). The influence effect Differences between the effect of the four most common species in the dataset (UA: unspecified arable; UG: unspecified grassland, PA: pasture, and SS: Sitka spruce) on C_{total} were relatively small. The low-level ‘bin’ of C_{total} (0.0–4.7 t C ha^{-1}) was dominated by short and young land covers that store low amounts of C as aboveground, below-ground, and dead organic matter (Fig. 4a). By contrast, the high level ‘bin’ of C_{total} (110.4–332.8 t C ha^{-1}) comprised tall land cover classes typically in the age range of 40–60 years and with high levels of aboveground biomass (AGBC).

For the heatmap created for the landscape features (4b), Bin 1 comprised 0–10.2 t C ha^{-1}, and Bins 2, 3, and 4 comprised 10.2–27.3, 27.3–58.0, and 58.0–120.0 t ha^{-1} respectively. As with the land cover, the highest values of C_{total} were associated with high values of aboveground biomass, belowground biomass and dead organic matter (Fig. 4b). High values of age and height were also associated with high values of C_{total}. Hence, young lone trees and hedges below a height of 5 m tended to have low values of C_{total}, and older tall features had high values. Lone trees like poplar and hedges like hawthorn and hazel and in an age group of greater than 30 years tended to have high values of C_{total}.

3.3. Model performance

The performance of the BBN models in determining C_{total} can be partly assessed using the error rate derived from a confusion matrix. The confusion matrices produced for each model contained four possible ranges for C_{total}. For each one of the 581 cases in the land cover dataset, the use of the BBN model in Netica generated beliefs for each possible state and the one with the highest probability value most likely was chosen as its prediction. This value was then compared with the measured value taken from the test file. The confusion matrix (Table 1) shows that 542 out of the 581 datum points were correctly allocated to the correct “bin”. A total of 37 values were not estimated correctly, 15 of which were predicted to have a higher value and the other 22 a lower value than their actual one. A similar matrix (Table 1) was produced for the landscape features, where a total of

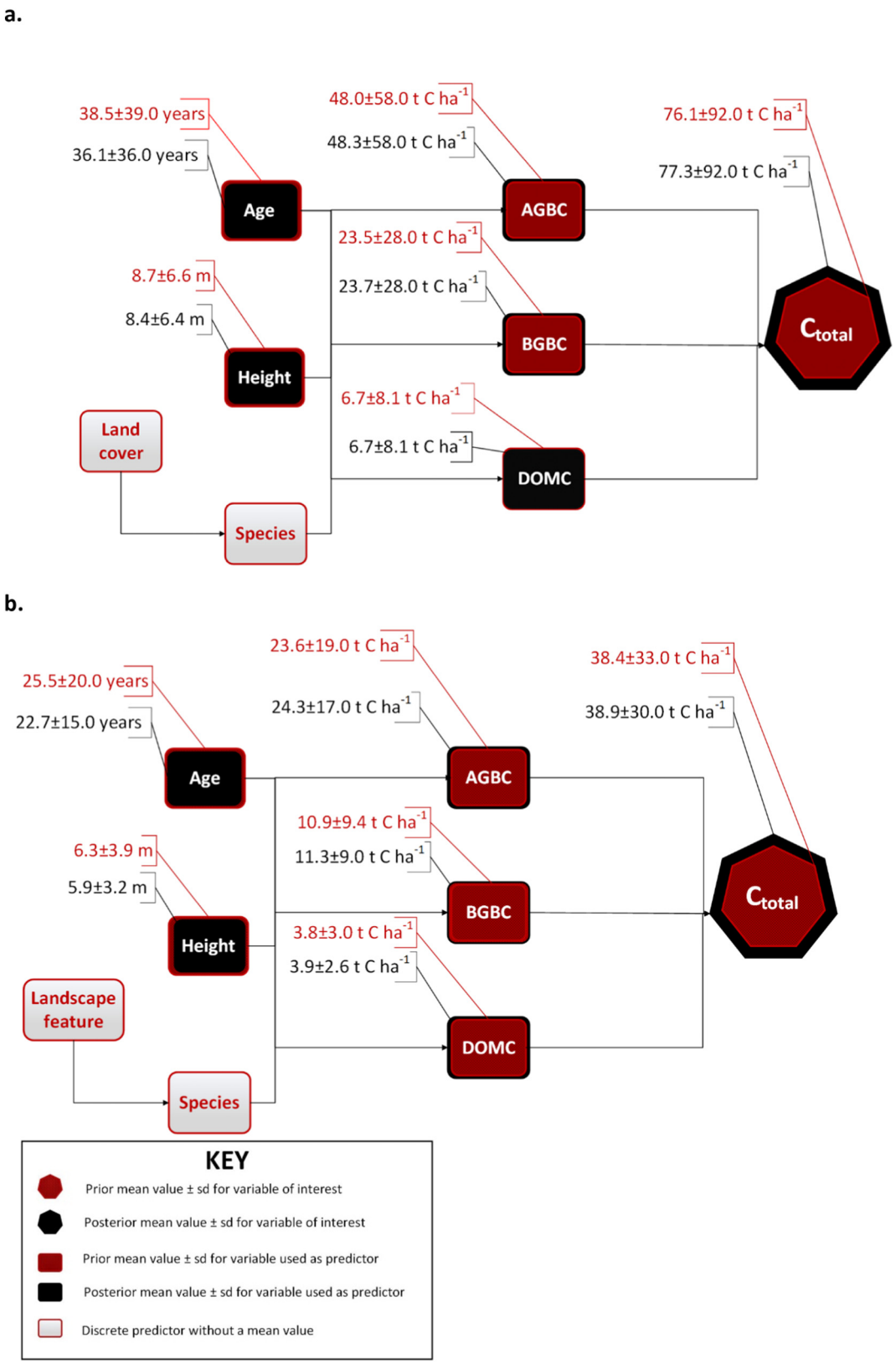


Fig. 3. Graphic representation for mean prior and posterior probability values and standard deviation for the parameters of the Bayesian Belief Network (BBN) model for a) the homogeneous land cover classes (n_{prior} = 576, $n_{posterio\ r}$ = 576 + 5) and b) landscape features (n_{prior} = 68, $n_{posterior}$ = 68 + 26).

94 cases were processed, and in this case, 90 features were correctly allocated, but four features were incorrectly allocated to Bin 1. Hence, the error rate for the BBN model was 6.7 % for the homogeneous land cover classes and 4.3 % for the landscape features indicating that, in general, the BBNs were able to match the allocated C_{total} value to the same bin as the measured C_{total} test values.

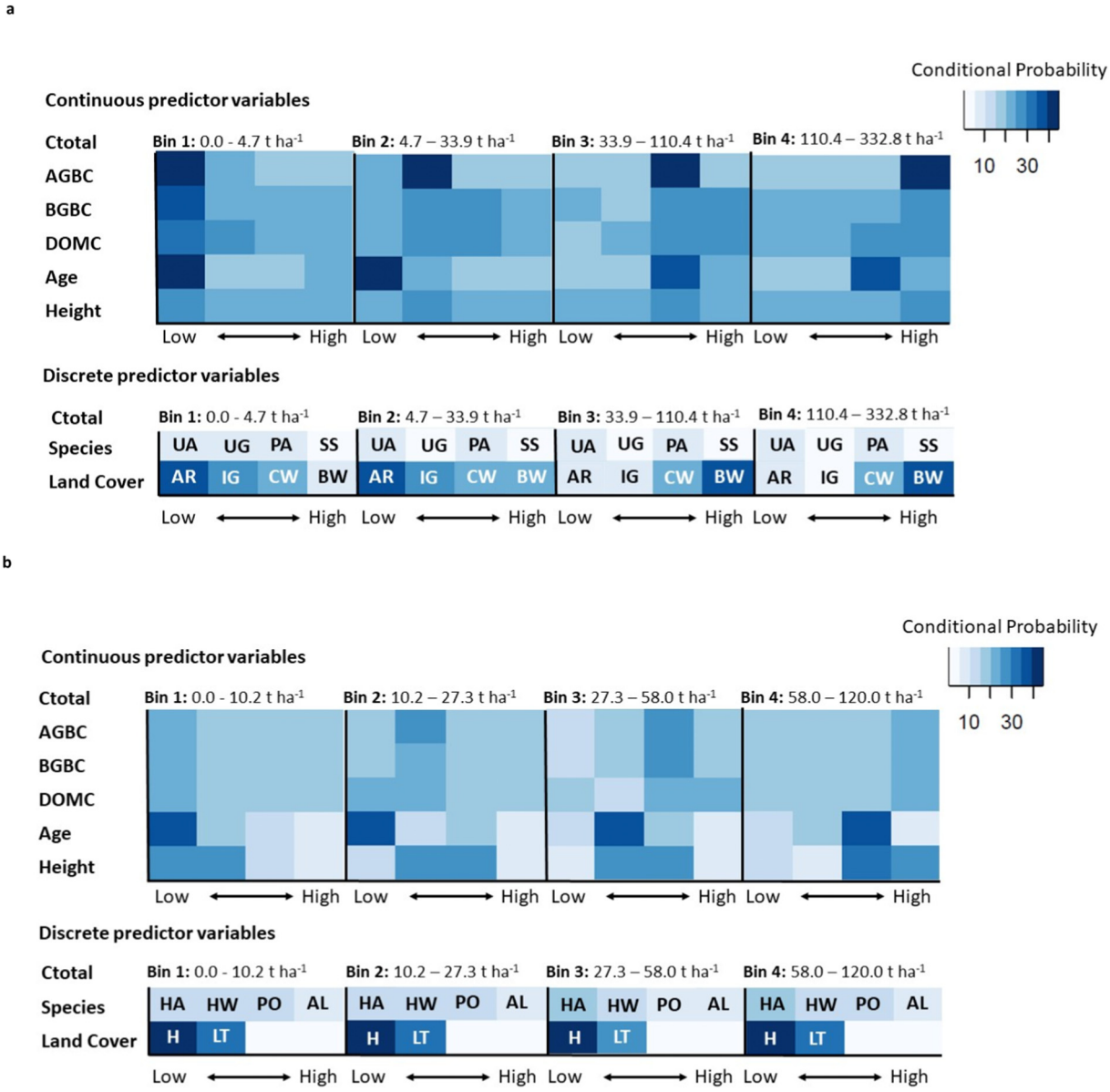


Fig. 4. a. Informed heat map presenting the four bins of the total biomass carbon stock (C_{total}; $n = 581$ of homogeneous land cover), and the conditional probabilities of five continuous predictor variables, and two discrete predictor variables showing the four most common species (UA: unspecified arable; UG: unspecified grassland, PA: pasture, and SS: Sitka spruce) and four most common land cover classes (AR: arable, IG: intensive grassland, CW: coniferous woodland, and BW: broadleaf woodland) b. Informed heat map presenting four bins of the total biomass carbon stock (C_{total}) of landscape features ($n = 94$), the conditional probabilities of five continuous predictor variables, and two discrete predictor variables showing the four most common species (HA: hazel; HW: hawthorn; PO: poplar, and AL: alder) and whether the landscape feature is a hedge (H) or lone tree (LT).

Table 1
Confusion matrix demonstrating the success of the posterior Bayesian Belief Network (BNN) to predict the total biomass carbon stock (C_{total}) values for a) the homogeneous land cover classes ($n = 581$) as inputs and b) landscape features as inputs ($n = 94$). The coloured values are those correclty predicted by the BBNs.

a) Predicted carbon stock (t C ha^{-1})				Actual carbon stock (t C ha^{-1})
0.0 - 4.7	4.7 – 33.9	33.9 – 110.4	110.4 – 332.8	
130	15	0	0	0.0 –4.7
18	127	0	0	4.7 – 33.9
0	2	142	1	33.9 – 110.4
0	0	1	143	110.4 – 332.8

b) Predicted carbon stock (t C ha^{-1})				Actual carbon stock (t C ha^{-1})
0.0 – 10.2	10.2 – 27.3	27.3 – 58.0	58.0– 120.0	
9	0	0	0	0.0 – 10.2
2	24	0	0	10.2 – 27.3
2	0	29	0	27.3 – 58.0
0	0	0	28	58.0– 120.0

3.4. Quantity and value of total carbon density

The next step in the analysis was to compare the measured values of C_{total} with the predicted results of C_{total} using the informed BBNs (Table 2). More details about the measured values can be found in Table A4 in the Appendix. For each land cover and landscape feature type, the most probable out of the four C_{total} bins is presented together with the weighted median value. This was calculated as the weighted median of all probable bins with their probabilities assigned as weights. The highest predicted value of C_{total} (72 t C ha^{-1}) was for an ash broadleaf woodland, compared to a measured value of 58 t C ha^{-1} (Table 2). The lowest predicted value of C_{total} of 2 t C ha^{-1} was for an arable field, compared to the measured value of 3 t C ha^{-1}. For the landscape features, the mean measured C_{total} of all hedges was 65 t C ha^{-1} and that for the lone trees was 100 t C ha^{-1}. The mean predicted

Table 2
Measured and predicted by the informed Bayesian Belief Network models total biomass carbon stock (C_{total}) for homogeneous land cover classes and landscape features at five study sites.

Farm	Land cover/feature	n	Total biomass carbon stock (t ha^{-1})		
			Measured	Most probable bin range and probability	Weighted median value and sd[a]
Homogeneous land cover					
HM	Arable: cereal	1	3.0	0.0–5.0 (96 %)	2.0 ± 2.7
S	Arable: stockfeed	1	5.0	0.0–5.0 (57 %)	4.0 ± 7.1
CP	Improved grassland	1	8.0	5.0–34.0 (96 %)	15.0 ± 9.9
HM	Improved grassland	1	7.0	5.0–34.0 (67 %)	15.0 ± 25.9
CP	Broadleaf: ash	1	58.0	34.0–110.0 (97 %)	72.0 ± 29.1
		5			
Landscape features					
CB	Hedge: hawthorn	1	19.0	10.0–27.0 (67 %)	20.0 ± 24.8
CB	Hedge: hawthorn	1	42.0	27.0–58.0 (84 %)	42.0 ± 17.4
CB	Hedge: hawthorn	1	44.0	27.0–58.0 (75 %)	41.0 ± 20.8
CB	Hedge: hawthorn	1	51.0	27.0–58.0 (75 %)	41.0 ± 20.8
CB	Hedge: mix	4	34.0–40.0	27.0–58.0 (84 %)	42.0 ± 17.4
E	Hedge: hazel	1	39.0	27.0–58.0 (84 %)	42.0 ± 17.4
CB	Hedge: mix	1	87.0	58.0–120.0 (90 %)	85.0 ± 27.0
E	Hedge: hawthorn	2	61.0–10.50	58.0–120.0 (75 %)	85.0 ± 36.5
E	Hedge: hazel	1	97.0	58.0–120.0 (75 %)	85.0 ± 36.5
E	Hedge: blackthorn	2	99.0–106.0	58.0–120.0 (75 %)	85.0 ± 36.5
HM	Hedge: hawthorn	1	79.0	58.0–120.0 (75 %)	85.0 ± 36.5
HM	Hedge: blackthorn	1	94.0	58.0–120.0 (75 %)	85.0 ± 36.5
HM	Hedge: mix	1	103.0	58.0–120.0 (75 %)	85.0 ± 36.5
	Mean		65.0		62.0
CP	Lone tree: ash	1	69.0	58.0–120.0 (90 %)	85.0 ± 27.0
S	Lone tree: poplar	7	98.0–111.0	58.0–120.0 (75 %)	85.0 ± 36.5
	Mean		100.0		85.0
		26			

HM: Harnhill Manor Farm; S: Silsoe; E: Elm Farm; CP: Clapham Park; CB: Crowmarsh Battle.
[a] sd = standard deviation.

value using the informed model was 62 t C ha^{-1} for the hedges and 85 t C ha^{-1} for the lone trees.

The last stage of the analysis was to predict the total biomass carbon stock (C_{total}) found across all of the land covers and landscape features on each of the five case study farms. As presented in Table 3 and presented visually in Fig. 5 together with the age and height predictors, Crowmarsh Battle, Harnhill Manor and Elm Farm had quite similar levels of C_{total} per hectare stored in the biomass of the homogenous land covers, ranging from 11.4 to 13.5 t ha^{-1}, while Silsoe had the lowest levels of just 2.4 t ha^{-1} (Table 3). Clapham Park had the highest stored biomass C_{total} of 21.7 t ha^{-1}. For the landscape features Crowmarsh Battle and Elm Farm had the highest per hectare values ranging from 59.4 to 60.6 t ha^{-1} and Harnhill Manor together with Clapham Park had the lowest, ranging from 32.1 to 35.9 t ha^{-1}. The level of C_{total} stored at the poplar trees of Silsoe was 52.3 t ha^{-1}.

4. Discussion

The results are discussed in terms of the key components and drivers determining C_{total}, the development of an integrated spatial model, the importance of landscape features for carbon storage, the benefits of using lidar data, and the benefits and disadvantages of using Bayesian Belief Network models.

4.1. Key drivers

The identification of important predictor variables influencing C_{total} and the structure of the BBN models were informed by published literature. The creation of heatmaps (Fig. 4a and b) is one method to illustrate the influence of each predictor on C_{total}. The developed heatmaps highlight that aboveground biomass carbon (AGBC) is the most influential predictor of C_{total} for both homogenous land covers and landscape features. This is a useful output of this research, as AGBC is often the easiest component to measure (Petrokofsky et al., 2012). Additionally, as described in Section 2.2 and as seen in Table A1 found at the Appendix, it can be used to calculate the C stored in the other forms of the biomass. For arable land and woodlands, the aboveground biomass carbon was typically 2–5 times higher than the belowground biomass carbon, as also indicated by algorithms in Table A1 in the Appendix. High values of C_{total} were also positively associated with height and large values for belowground and dead organic biomass. This is very important particularly for woody vegetation, as it can assist in understanding how efficiently trees store C throughout their lifetime. The least influential predictor, of those considered, was the species for land cover classes and underlying land cover for the landscape features, presumably because these predictors were less directly linked to C_{total}.

Table 3
Mean per hectare and total biomass carbon stock (C_{total}) present across all the land cover and landscape features of five case study farms, together with the proportion of the area and the total biomass carbon stored in landscape features.

Farm	Total area (ha)	Total area landscape features (ha)	Mean carbon of land cover (t ha^{-1})	Mean carbon of landscape feature (t ha^{-1})	Total Carbon (t)	Total carbon of land cover (t)	Total carbon of landscape feature (t)
Crowmarsh	42.46	5.44 (13 %)	11.4	59.4	807	484	323 (40 %)
Elm Farm	80.06	4.46 (6 %)	13.5	60.6	1354	1084	270 (20 %)
Harnhill Manor	394.86	7.50 (2 %)	11.8	32.1	4882	4641	241 (5 %)
Clapham Park	15.25	0.67 (4 %)	21.7	35.9	354	330	24 (7 %)
Silsoe	4.71	2.66 (57 %)	2.4	52.3	150	11	139 (93 %)
Mean	107.47	4.15 (4 %)	12.2	48.1	1509	1310	199 (13 %)

4.2. Development of an integrated spatial method for biomass carbon

The approach described in this paper shows that it is possible to determine both the biomass C stock of homogeneous land covers and landscape features in an integrated logical format using remotely sensed data. This is something relatively new as previous studies (Adger and Subak, 1996; Cannell and Milne, 1995; Cantarello et al., 2011; Cruickshank et al., 2000; Dewar and Cannell, 1992; Gregory et al., 1995; IPCC, 2019; Jian et al., 2020) have focused on just the land cover and missing a big part of farm C that can be found in the landscape features. Similarly, GHG emission factor calculators (Campos, 2020; Carlson et al., 2017; Janzen et al., 2006; Shrestha, 2014; Stöckle et al., 2001; Sustainable Food Lab et al., 2011; Ziegler et al., 2016) do not separately account for landscape features.

Within our approach, we assumed that a lone tree is an additive feature, and the C_{total} of the tree is added to the C_{total} of the underlying land cover. By contrast, we considered that the C_{total} of a hedgerow was not additive but that it effectively represented a different type of land cover. This integrated method also incorporates the spatial aspect, as we have demonstrated that the BBNs can be successfully combined with a GIS interface across the five study areas to spatially apply the influence network of the C_{total}.

The development of such robust farm-scale biomass carbon storage inventories is useful to farmers who wish to achieve net-zero GHG emissions

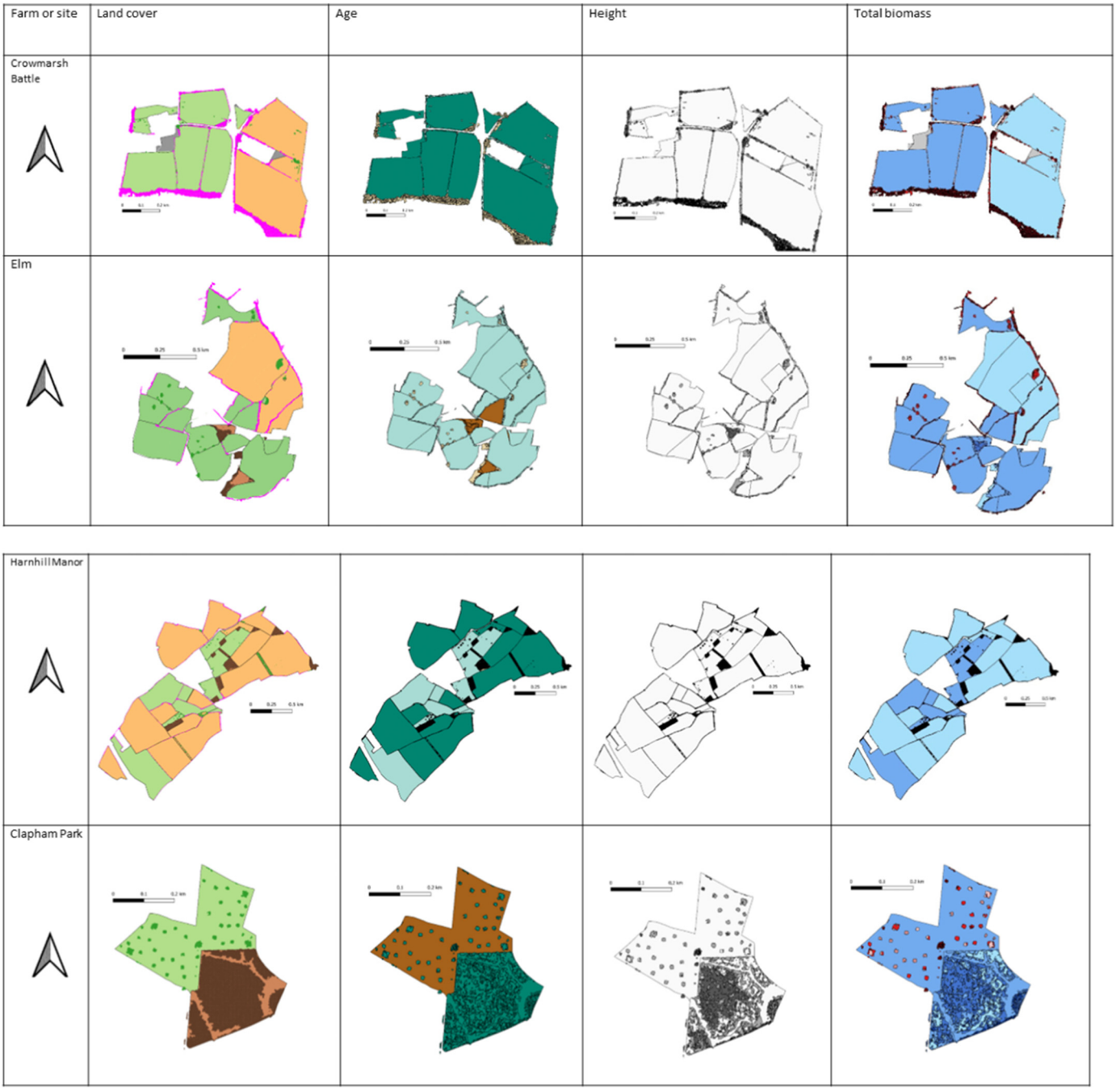


Fig. 5. Maps presenting the land cover types and species together with their age, height and the simulated levels of the total biomass carbon stock (C_{total}) across five studies.

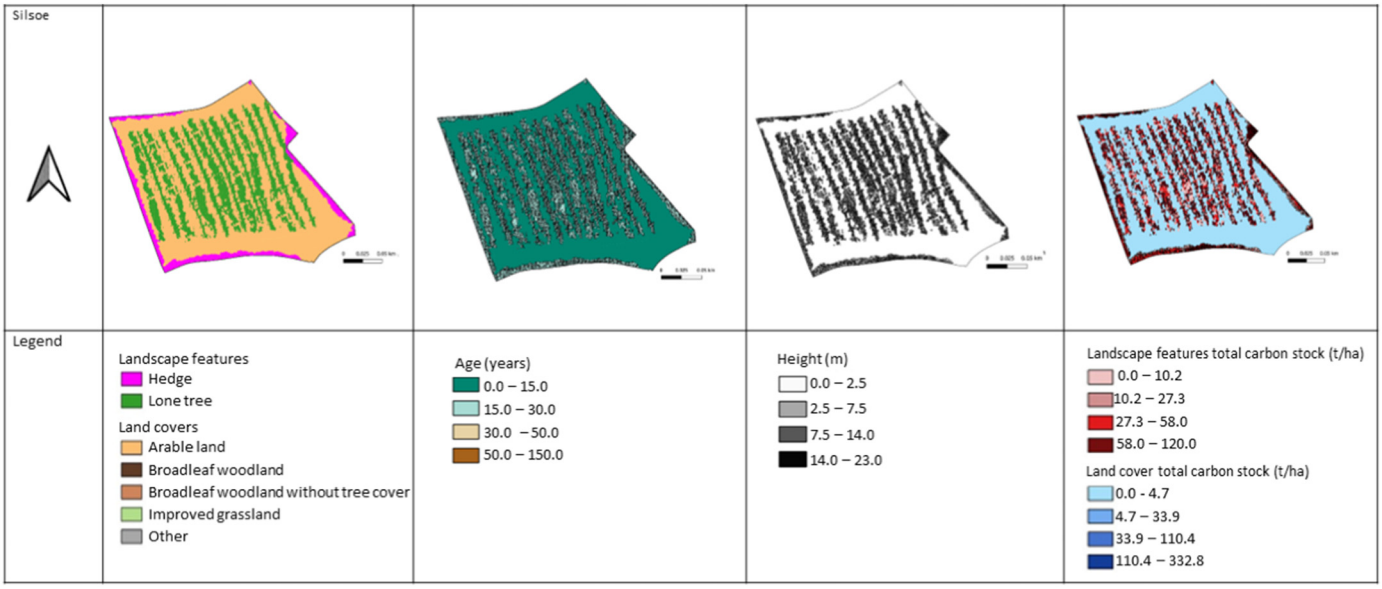


Fig. 5 (*continued*).

and move to being "carbon" positive. Theoretically, it would be possible to organise on-site measurements of aboveground, belowground, and dead organic carbon of each land cover or landscape feature on a farm, but this would be costly in terms of time and resources and could be destructive. There are also potential accessibility issues (Kumar and Saran, 2014; Timothy et al., 2016). For woodlands, species-specific allometric equations have been developed to derive tree biomass from measurements such as stem wood, stem bark, branches, and the cross-sectional area (Zianis et al., 2005; Zianis, 2008; Vargas-Larreta et al., 2017). However, in practice woodlands on many farms comprise a mixture of species.

Farmers who wish to sell "carbon credits" from the accumulation of biomass carbon need to demonstrate both the additionality and the permanence of carbon storage. For example, within the Woodland Carbon Code in the UK, it is necessary to demonstrate that increases in biomass carbon arising from tree planting have only occurred because of carbon credits; that is, it would not have occurred without the credits. In addition, it is necessary to ensure the permanence of carbon storage. In contrast to woodland establishment, which can involve a redesignation of land use from agriculture to woodland, it is more difficult for a farmer to demonstrate the permanence of carbon storage in landscape features. Hence, in many cases, the drive for increasing biomass carbon storage on farms through landscape features is likely to be more useful in demonstrating that an individual farm or group of farm businesses is achieving net-zero greenhouse gas emissions rather than to profit from the sale of carbon credits.

4.3. Importance of landscape features

Landscape features such as hedgerows and lone trees are well-recognised for their landscape, biodiversity, and cultural value. However, this paper also highlights that they can store large quantities of carbon similar, on a per hectare basis, to woodlands and substantially higher than the mean biomass carbon levels of farms (Table 3). Hence increasing the area of hedgerows and agroforestry systems can be a useful on-farm strategy to increase the storage of carbon biomass (Axe et al., 2017; Drexler et al., 2021). Although the case study farms were not randomly selected, the results demonstrate that landscape features could comprise 5–93 % of the biomass carbon storage on farms. On average, the mean value was 13 % across the five case studies.

4.4. Benefits of using lidar data

A useful innovation described in this paper, is the capacity to spatially present the variation in carbon stocks along the length of hedges and within woodlands using remote sensing and lidar data (Fig. 5) (Petrokofsky et al., 2012). Airborne lidar data can both reduce the costs and effort required to conduct whole-farm surveys (Arenas-corraliza et al., 2020; Timothy et al., 2016) in addition to providing a remote screenshot of vegetation heights. Lidar data have been used in forest studies to estimate timber volumes (Naesset, 1997), canopy cover (Arenas-corraliza et al., 2020; Lefsky et al., 2002), forest aboveground carbon content (Patenaud et al., 2004; Urbazaev et al., 2018), and above-ground biomass (Lu et al., 2012; Timothy et al., 2016). The combined method of lidar and BBN models is relatively new. Woodland studies such as Wang et al. (2019), and Zapata-Cuartas and Sierra (2014) have proved the potential of improving the estimation of above-ground biomass through lidar data by incorporating a BBN method instead of using the traditional allometric equations. In this study, we have extended this work by using the BBN method and estimating the total biomass carbon in additional land cover categories together with landscape features. Moreover, a useful attribute of the methodology described in this paper is that the required variables (such as age and height) can potentially be determined remotely using historic land cover and lidar height data.

The use of lidar data allows the derivation of different levels of biomass within land units that have a single designated "land cover". For example, within a single parcel designated as "woodland" at Clapham Park (Fig. 5), the incorporation of height data from lidar enabled the BBN for land cover to distinguish between areas of trees and grass walkways. The vegetation height in the walkways was less than the surrounding trees, with corresponding lower values of C_{total}. The process was also able to infer different values of C_{total} along the length of the individual hedges across four of the farms because of differences in the lidar height data (e.g. Crowmarsh Battle in Fig. 6). This process of including variable heights within woodlands and along hedges can be more informative to land owners than previous studies in the UK (Cannell and Milne, 1995; Milne and Brown, 1997), Ireland (Cannell et al., 1996; Cruickshank et al., 2000), and in other parts of the world (Vargas-Larreta et al., 2017) where one value is given to the whole woodland with the same age.

4.5. Advantages and disadvantages of the Bayesian approach

This paper shows that it was possible to develop an approach which could quantify biomass carbon stocks and their certainty. Such a Bayesian approach can be particularly useful for farmer or land manager practitioners. It integrates information and measured values to derive relationships between parameters (McCandless and Gustafson, 2017), aggregates

uncertainty of information across complex networks when information is limited (De Campos et al., 2004; Friedman and Koller, 2003; Jiang et al., 2011), and prevents overfitting (Borsuk et al., 2004; Melody et al., 2011). Moreover a novelty of this study is the development of a theoretical framework in this area.

The way that BBN models quantify uncertainty is different from deterministic models as the uncertainty is directly embedded in the predictions and the generated error boundaries (Taalab et al., 2015b). The two informed BBN models developed in this paper were able to predict successfully the test values with propagated error rates of 6.7 % for C_{total} for the homogeneous land features model and 4.3 % for the landscape features. These error rates were lower than in studies where BBN models have been applied to soil systems (Taalab et al., 2015b) where the error ranged from 15 % to 36 % and in environmental systems of ecosystem service trade-offs and synergies where the error rate ranges from 41 to 45 % (Karimi et al., 2021). Despite the reliance on only seven predictors, the relatively lower error rates of BBN models reported in this paper suggests that the selected predictors are strong determinants of C_{total}.

The developed BBN models generate distributions of C_{total} that were not based on pre-determined relations and correlations between parameters. This can be an advantage when the relationships between parameters are complex. One example of a potential non-intuitive response in terms of biomass carbon is that the land cover model (Fig. 4a) suggests that expanding the continuity of a land cover beyond 40 years did not necessarily result in a higher value of C_{total}. Such relationships could be a result of the carbon biomass of woodlands plateauing once a woodland has reached 40 years or the confounding effects of tree harvests.

The models uses the predictions from the BBN models and classifies each land cover/feature found in the study sites into one of the four C_{total} bins, where the median value of the bin range is given as seen in Table 2. This lack of precision can result in overestimation or underestimation of values. The mean value for arable land at the two sites ranged between 2.0 and 4.0 t ha^{-1} and is quite similar to values of 2.0–6.57 t ha^{-1} reported for mean arable crop categories in the UK GHG Inventory (Brown et al., 2020), and studies by Cruickshank et al. (2000), Cantarello et al. (2011) and Adger and Subak (1996). The derived mean value for improved grassland of 15.0 t ha^{-1} (Table 2) at the two study sites, is also higher than values of 4.2–8.4 t ha^{-1} in the aforementioned studies. By contrast, the value of 72.0 t ha^{-1} for woodland is close to values of 52.9–74.0 ha^{-1} provided by Milne and Brown (1997) and Morison et al. (2012).

It is difficult to compare the C stored in the total biomass of the hedgerows with other studies, as studies such as Falloon et al. (2004), Black et al. (2014, page 24) have only measured the AGBC and Biffi et al. (2022) focused on AGBC and BGBC without any DOMC measurements. Similarly, for the lone trees, the comparison of the mean value of 85.0 t ha^{-1} is a challenge to be compared with other agroforestry studies, as the type of system has important implications on the stored carbon.

Within our approach, we developed models that provide a "snapshot" of the values of C_{total} within a given year. The model does not account for temporal variation in C_{total} within the year. However, in practice, the biomass of a growing arable crop can change from effectively zero before the seeding of cultivated soil to maximum values just before harvest. This temporal limitation is a common feature of most carbon accounting methods. For example, the standard practice in the IPCC guidelines (IPCC, 2019; Chapter 5 page 7), is to assume the C density value of arable crop after harvest, i.e. the biomass is primarily related to the weight of the roots and the stubble of the crop. In using the model, we also assumed a uniform species type and constant age within a land cover area. Although in practice, there may be changes in tree species within a woodland or a hedge, the logic generally follows the methodology described in the UK's Woodland Carbon Code Assessment Protocol (Jenkins et al., 2011).

Looking forward, we anticipate two avenues where the model could be further developed or used. Firstly, there is the potential to use the model to predict the effect of management interventions on C_{total} at a farm level. For instance, a farm manager could use it to identify the height and species mix of a specific hedgerow or woodland that would tend to achieve a specific level of C_{total}. Additionally, the model can be used in farm scenario testing, agroforestry and forestry studies. Secondly, for this study, we used the Netica BBN software (Norsys Software Corp., 2022) that categorised the outputs into discrete "bins". The advantage of this categorisation is that the computing of the relationships becomes far more feasible on a standard desktop computer or laptop. However, more computationally demanding software programmes exist (Thomas et al., 2003), where the outputs can be expressed as continuous variables. As the computational capacity of computers increases, future farm, or forestry studies may be able to apply the described in this paper techniques in order to derive continuous predictions of C_{total}.

5. Conclusions

This study supports the hypothesis that a probabilistic approach for accounting for farm biomass carbon stocks is feasible. A particularly useful attribute of the demonstrated approach is that the estimation of the total biomass carbon stocks (C_{total}) of both land cover and landscape features is performed in an integrated logical format. The described BBN approach, which seeks constantly to improve prediction accuracy, provided a method for applying current literature and measured information about the variables that most influence the C_{total} values of land cover classes and landscape features.

An advantage of the approach described in this paper is that the required variables (such as age and height) can potentially be determined remotely using historic land cover and lidar height data. Additionally, in this paper, we have demonstrated that the BBNs can be successfully combined with a GIS interface across the five study areas to spatially apply the influence network of the C_{total}. A useful attribute of the methodology described in the paper, including the use of lidar data, is that it was possible to predict differences in C_{total} across hedgerows and land parcels. As the computational capacity of computers increases, there may be future opportunities to use forms of BBN analysis that produces continuous rather than discrete estimates of C_{total}.

CRediT authorship contribution statement

Styliani Beka: Conceptualization, Methodology, Formal analysis, Data curation, Writing – original draft, Visualization. **Paul J. Burgess:** Conceptualization, Supervision, Project administration, Writing – review & editing. **Ron Corstanje:** Conceptualization, Supervision, Writing – review & editing.

Data availability

Data underlying this study can be accessed through the Cranfield University repository at: https://doi.org/10.17862/cranfield.rd.20375055.v2.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The research was supported by Cranfield University and NERC through the CENTA2 PhD programme (NE/L002493/1). We are also thankful for the input of Alexa Verah for providing us with some unpublished hedgerow measurements.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2022.160618.

References

Adger, W.N., Subak, S., 1996. Estimating Above-Ground Carbon Fluxes From UK Agricultural Land. 162, pp. 191–204.

Arenas-corraliza, I., Nieto, A., Moreno, G., 2020. Automatic mapping of tree crowns in scattered-tree woodlands using low-density LiDAR data and infrared imagery. Agrofor. Syst. 2. https://doi.org/10.1007/s10457-020-00517-2.

Axe, M.S., Grange, I.D., Conway, J.S., 2017. Carbon storage in hedge biomass—a case study of actively managed hedges in England. Agric. Ecosyst. Environ. 250, 81–88. https://doi.org/10.1016/j.agee.2017.08.008.

Biffi, S., Chapman, P.J., Grayson, R.P., Ziv, G., 2022. Soil carbon sequestration potential of planting hedgerows in agricultural landscapes. J. Environ. Manage. 307 (2022). https://doi.org/10.1016/j.jenvman.2022.114484.

Black, K., Green, S., Mullooley, G., Poveda, A., 2014. Carbon Sequestration by Hedgerows in the Irish Landscape. Towards a National Hedgerow Biomass Inventory for the LULUCF Sector Using LiDAR Remote Sensing. . https://www.epa.ie/publications/research/climate-change/ccrp-32-for-webFINAL.pdf.

Branca, G., Lipper, L., McCarthy, N., et al., 2013. Food security, climate change, and sustainable land management.A review. Agron. Sustain. Dev. 33, 635–650. https://doi.org/10.1007/s13593-013-0133-1.

Brandão, M., Levasseur, A., Kirschbaum, M.U.F., Weidema, B.P., Cowie, A.L., Jørgensen, S.V., Hauschild, M.Z., Pennington, D.W., Chomkhamsri, K., 2013. Key issues and options in accounting for carbon sequestration and temporary storage in life cycle assessment and carbon footprinting. Int. J. Life Cycle Assess. 18, 230–240. https://doi.org/10.1007/s11367-012-0451-6.

Brewer, A., Ditchburn, B., Cross, D., Whitton, E., Ward, A., 2017. Tree Cover Outside Woodland in Great Britain Edinburgh.

Brown, P., Cardenas, L., Choudrie, S., Jones, L., Karagianni, E., MacCarthy, J., Passant, N., Richmond, B., Smith, H., Thistlethwaite, G., Thomson, A., Turtle, L., Wakeling, D., 2020. UK Greenhouse Gas Inventory 1990 to 2018: Annual Report for submission under the Framework Convention on Climate Change. https://naei.beis.gov.uk/reports/reports?report_id=998.

Borsuk, M.E., Stow, C.A., Reckhow, K.H., 2004. A bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. Ecol. Model. 173, 219–239. https://doi.org/10.1016/j.ecolmodel.2003.08.020.

Burbi, S., Baines, R.N., Conway, J.S., 2016. Achieving successful farmer engagement on greenhouse gas emission mitigation. Int. J. Agric. Sustain. 14, 466–483. https://doi.org/10.1080/14735903.2016.1152062.

Campos, I., 2020. Common Agricultural Policy Post 2020-Farm Sustainability Tool for Nutrients - FaST Farm Sustainability Tool for Nutrients.

Cannell, M.G., Cruickshank, M.M., Mobbs, D., 1996. Carbon storage and sequestration in the forests of Northern Ireland. Forestry 69, 155–165. https://doi.org/10.1093/forestry/69.2.155.

Cannell, M.G.R., Milne, R., 1995. Carbon pools and sequestration in forest ecosystems in Britain. Forestry 68, 361–378. https://doi.org/10.1093/forestry/68.4.361.

Cantarello, E., Newton, A.C., Hill, R.A., 2011. Potential effects of future land-use change on regional carbon stocks in the UK. Environ. Sci. Policy 14, 40–52. https://doi.org/10.1016/j.envsci.2010.10.001.

Carlson, B.R., Carpenter-Boggs, L.A., Higgins, S.S., Nelson, R., Stöckle, C.O., Weddell, J., 2017. Example OFoot. Comput. Electron. Agric. 142, 211–223. https://doi.org/10.1016/j.compag.2017.09.007.

Centre for Ecology and Hydrology, 2020. The UKCEH Land Cover Maps for 2017, 2018 and 2019.

Centre for Ecology and Hydrology, 2017. Land Cover Map 2015 Dataset Documentation. https://doi.org/10.5285/bb15e200-9349-403c-bda9-b430093807c7. Crowmarsh Gifford, Wallingford Oxfordshire, OX10 8BB United Kingdom.

Constantinou, A.C., Fenton, N., Neil, M., 2016. Integrating expert knowledge with data in bayesian networks: preserving data-driven expectations when the expert variables remain unobserved. Expert Syst. Appl. 1 (56), 197–208. https://doi.org/10.1016/j.eswa.2016.02.050.

Cruickshank, M.M., Tomlinson, R.W., Trew, S., 2000. Application of CORINE land-cover mapping to estimate carbon stored in the vegetation of Ireland. J. Environ. Manag. 58, 269–287. https://doi.org/10.1006/jema.2000.0330.

De Campos, M., M Fernández-Luna, J., Huete, J.F., 2004. Bayesian networks and information retrieval: an introduction to the special issue. Info. Process. Manag. 40, 727–733. https://doi.org/10.1016/j.ipm.2004.03.001.

Department for Environment Food & Rural Affairs, 2022. National Lidar Programme. https://www.data.gov.uk/dataset/f0db0249-f17b-4036-9e65-309148c97ce4/national-lidar-programme.

Dewar, R.C., Cannell, M.G.R., 1992. Carbon sequestration in the trees, products and soils of forest plantations: an analysis using UK examples. Tree Physiol. 11, 49–71. https://doi.org/10.1093/treephys/11.1.49.

Digimap, 2022. GB National Outlines [SHAPE geospatial data], Scale 1:250000, Tiles: GB, Updated: 8 June 2005, Ordnance Survey (GB), Using: EDINA Digimap Ordnance Survey Service. https://digimap.edina.ac.uk.

Drexler, S., Gensior, A., Don, A., 2021. Carbon sequestration in hedgerow biomass and soil in the temperate climate zone. Reg. Environ. Chang. 21. https://doi.org/10.1007/s10113-021-01798-8.

ENVI, 2022. ENVI. (Version 5.6). [Computer software]. https://www.l3harrisgeospatial.com/Software-Technology/ENVI.

Falloon, P., Powlson, D., Smith, P., 2004. Managing field margins for biodiversity and carbon sequestration: a Great Britain case study. Soil Use Manag. 20, 240–247.

Food and Agriculture Organization of the United Nations, 2021. World Food and Agriculture – Statistical Yearbook 2021.

Friedman, N., Koller, D., 2003. Being Bayesian about Bayesian network structure: a Bayesian approach to structure discovery in Bayesian networks. Mach. Learn. 50 (1–2), 95–126. https://doi.org/10.4060/cb4477en.

García de Jalón, S., Graves, A., Palma, J.H.N., Williams, A., Upson, M., Burgess, P.J., 2018. Modelling and valuing the environmental impacts of arable, forestry and agroforestry systems: a case study. Agrofor. Syst. 92 (4), 1059–1073. https://doi.org/10.1007/s10457-017-0128-z.

Google Earth Pro, 2022. Version 7.3.4. Available to download: https://earth.google.com/web/.

Grafius, D.R., Corstanje, R., Warren, P.H., Evans, K.L., Norton, B.A., Siriwardena, G.M., Pescott, O.L., Plummer, K.E., Mears, M., Zawadzka, J., Richards, J.P., Harris, J.A., 2019. Using GIS-linked Bayesian belief networks as a tool for modelling urban biodiversity. Landsc. Urban Plan. 189, 382–395. https://doi.org/10.1016/j.landurbplan.2019.05.012.

Gregory, P.J., Palta, J.A., Batts, G.R., 1995. Root systems and root:mass ratio-carbon allocation under current and projected atmospheric conditions in arable crops. Plant Soil 187, 221–228. https://doi.org/10.1007/BF00017089.

Hassall, K.L., Dailey, G., Zawadzka, J., Milne, A.E., Harris, J.A., Corstanje, R., Whitmore, A.P., 2019. Facilitating the elicitation of beliefs for use in Bayesian belief modelling. Environ. Model. Softw. 122, 104539. https://doi.org/10.1016/j.envsoft.2019.104539.

Heckerman, D., 1997. Bayesian networks for data mining. Data Min. Knowl. Disc. 1, 79–119. https://doi.org/10.1023/A:1009730122752.

IPCC, 2019. In: Calvo Buendia, E., Tanabe, K., Kranjc, A., Baasansuren, J., Fukuda, M., Ngarize, S., Osako, A., Pyrozhenko, Y., Shermanau, P., Federici, S. (Eds.), 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories. IPCC, Switzerland.

Janzen, H.H., Angers, D.A., Boehm, M., Bolinder, M., Desjardins, R.L., Dyer, J.A., Ellert, B.H., Gibb, D.J., Gregorich, E.G., Helgason, B.L., Lemke, R., Massé, D., McGinn, S.M., McAllister, T.A., Newlands, N., Pattey, E., Rochette, P., Smith, W., Vandenbygaart, A.J., Wang, H., 2006. A proposed approach to estimate and reduce net greenhouse gas emissions from whole farms. Can. J. Soil Sci. 86, 401–418. https://doi.org/10.4141/S05-101.

Jenkins, T.A.R., Mackie, E.D., Matthews, R.W., Miller, G., Randle, T.J., White, M.E., 2011. Forestry Commission Woodland Carbon Code Assessment Protocol. Forestry Commission.

Jian, J., Du, X., Reiter, M.S., Stewart, R.D., 2020. A meta-analysis of global cropland soil carbon changes due to cover cropping. Soil Biol. Biochem. 143, 107735. https://doi.org/10.1016/j.soilbio.2020.107735.

Jiang, X., Neapolitan, R.E., Barmada, M.M., et al., 2011. Learning genetic epistasis using Bayesian network scoring criteria. BMC Bioinformatics 12, 89. https://doi.org/10.1186/1471-2105-12-89.

Karimi, J.D., Harris, J.A., Corstanje, R., 2021. Using Bayesian belief networks to assess the influence of landscape connectivity on ecosystem service trade-offs and synergies in urban landscapes in the UK. Landsc. Ecol. 36, 3345–3363. https://doi.org/10.1007/s10980-021-01307-6.

Koller, D., Friedman, N., 2009. Probabilistic Graphical Models-Principles and Techniques, Italica. The MIT Press https://doi.org/10.2307/478142.

Korb, K.B., Nicholson, A.E., 2010. Bayesian Artificial Intelligence. CRC Press.

Kumar, K., Saran, S., 2014. Web based geoprocessing tool for coverage data handling. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives https://doi.org/10.5194/isprsarchives-XL-8-1139-2014 Edinburgh.

Krauss, M., Wiesmeier, Don, M.A., Cuperus, F., Gattinger, A., Gruber, S., Haagsma, W.K., Peigné, J., Chiodelli-Palazzoli, M., Schulz, F., Van der Heijden, M.G.A., Vincent-Caboud, L., Wittwer, R.A., Zikeli, S., Steffens, M., 2022. Reduced tillage in organic farming affects soil organic carbon stocks in temperate Europe. Soil Tillage Res. (ISSN: 0167-1987) 216, 105262. https://doi.org/10.1016/j.still.2021.105262.

Lefsky, M.A., Cohen, W.B., Harding, D.J., Parker, G.G., Acker, S.A., Gower, S.T., 2002. Lidar remote sensing of above-ground biomass in three biomes. Glob. Ecol. Biogeogr. 11, 393–399. https://doi.org/10.1046/j.1466-822x.2002.00303.x.

Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., Vaglio Laurin, G., Saah, D., 2012. Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates. Int. J. For. Res. 2012, 1–16. https://doi.org/10.1155/2012/436537.

Marcot, B.G., 2012. Metrics for evaluating performance and uncertainty of Bayesian network models. Ecol. Model. 230, 50–62. https://doi.org/10.1016/j.ecolmodel.2012.01.013.

Melody, N., Phillips, L.D., George, H., 2011. Exploring Bayesian belief networks using Netica. Evidence Synthesis in Healthcare: A Practical Handbook for Clinicians, pp. 1–326 https://doi.org/10.1007/978-0-85729-206-3.

McCandless, L.C., Gustafson, P., 2017. A comparison of Bayesian and Monte Carlo sensitivity analysis for unmeasured confounding. Stat. Med. 36, 2887–2901. https://doi.org/10.1002/sim.7298.

Microsoft Office, 2022. Microsoft Excel 365 (computer software). https://www.microsoft.com/en-us/microsoft-365/excel.

Milne, R., Brown, T.A., 1997. Carbon in the vegetation and soils of Great Britain. J. Environ. Manag. 49, 413–433. https://doi.org/10.1006/jema/1995/0118.

Morison, J., Matthews, R., Miller, G., Perks, M., Randle, T., Vanguelova, E., White, M., Yamulki, S., 2012. Understanding the carbon and greenhouse gas balance of forests in Britain. . https://cdn.forestresearch.gov.uk/2022/02/fcrp018-3.pdf.

National Forest Inventory Forest Research, 2017. Tree Cover Outside Woodland in Great Britain, Forestry Commission Statistical Report. Edinburgh.

Naesset, E., 1997. Estimating timber volume of forest stands using airborne laser scanner data. Remote Sens. Environ. 61 (2), 246–253. https://doi.org/10.1016/S0034-4257(97)00041-2.

Norsys Software Corp., 2022. Netica TM, Application for Belief Networks and Influence Diagrams: User's Guide, Quality. Norsys Software Corp.

Patenaud, G., Hill, R.A., Milne, R., Gaveau, D.L.A., Briggs, B.B.J., Dawson, T.P., 2004. Quantifying forest above ground carbon content using LiDAR remote sensing. Remote Sens. Environ. 93, 368–380. https://doi.org/10.1016/j.rse.2004.07.016.

Pearl, J., 1988. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Petrokofsky, G., Kanamaru, H., Achard, F., Goetz, S.J., Joosten, H., Holmgren, P., Lehtonen, A., Menton, M.C.S., Pullin, A.S., Wattenbach, M., 2012. Comparison of methods for the measurement and assessment of carbon stocks and carbon stock changes in terrestrial carbon pools. Environ. Evid. 1–21.

QGIS, 2022. QGIS: A Free and Open Source Geographic Information System. (Version 3.28). [Computer software]. https://www.qgis.org/en/site/index.html.

Rosenstock, T.S., Rufino, M.C., Butterbach-Bahl, K., Wollenberg, E., Richards, M., 2016. Methods for Measuring Greenhouse Gas Balances and Evaluating Mitigation Options in Smallholder Agriculture, Methods for Measuring Greenhouse Gas Balances and Evaluating Mitigation Options in Smallholder Agriculture. https://doi.org/10.1007/978-3-319-29794-1.

Smith, P., Davies, C.A., Ogle, S., Zanchi, G., Bellarby, J., Bird, N., Boddey, R.M., McNamara, N.P., Powlson, D., Cowie, A., van Noordwijk, M., Davis, S.C., Richter, D.D.B., Kryzanowski, L., van Wijk, M.T., Stuart, J., Kirton, A., Eggar, D., Newton-Cross, G., Adhya, T.K., Braimoh, A.K., 2012. Towards an integrated global framework to assess the impacts of land use and management change on soil carbon: current capability and future vision. Glob. Chang. Biol. 18, 2089–2101. https://doi.org/10.1111/j.1365-2486.2012.02689.x.

Stöckle, C.O., Nelson, R., Mccool, D., 2001. Cropping Systems Simulation Model User's Manual CropSyst Preface. Washingt. State Univ. Biol. Syst. Eng. Dep.

Sustainable Food Lab, University of Aberdeen, Unilever, 2011. The Cool Farm Tool, a User's Guide.

Taalab, K., Corstanje, R., Mayr, T.M., Whelan, M.J., Creamer, R.E., 2015a. The application of expert knowledge in Bayesian networks to predict soil bulk density at the landscape scale. Eur. J. Soil Sci. 66, 930–941. https://doi.org/10.1111/ejss.12282.

Taalab, K., Corstanje, R., Zawadzka, J., Mayr, T., Whelan, M.J., Hannam, J.A., Creamer, R., 2015b. On the application of Bayesian networks in digital soil mapping. Geoderma 259–260, 134–148. https://doi.org/10.1016/j.geoderma.2015.05.014.

Thomas, A., Best, N., Way, R., 2003. WinBUGS User Manual.

Timothy, D., Onisimo, M., Cletah, S., Adelabu, S., Tsitsi, B., 2016. Remote sensing of aboveground forest biomass: a review. Trop. Ecol. 57, 125–132.

Upson, M.A., Burgess, P.J., 2013. Soil organic carbon and root distribution in a temperate arable agroforestry system. Plant Soil 373, 43–58. https://doi.org/10.1007/s11104-013-1733-x.

Upson, M.A., Burgess, P.J., Morison, J.I.L., 2016. Soil carbon changes after establishing woodland and agroforestry trees in a grazed pasture. Geoderma 283, 10–20. https://doi.org/10.1016/j.geoderma.2016.07.002.

Urbazaev, M., Thiel, C., Cremer, F., Dubayah, R., Migliavacca, M., Reichstein, M., Schmullius, C., 2018. Estimation of forest aboveground biomass and uncertainties by integration of field measurements, airborne LiDAR, and SAR and optical satellite data in Mexico. Carbon Balance Manag. 13. https://doi.org/10.1186/s13021-018-0093-5.

Van Soest, H.L., den Elzen, M.G.J., van Vuuren, D.P., 2021. Net-zero emission targets for major emitting countries consistent with the Paris Agreement. Nat. Commun. 12, 1–9. https://doi.org/10.1038/s41467-021-22294-x.

Varah, A., 2021. Hedgerow Biomass Measurements.

Vargas-Larreta, B., López-Sánchez, C.A., Corral-Rivas, J.J., López-Martínez, J.O., Aguirre-Calderón, C.G., Álvarez-González, J.G., 2017. Allometric equations for estimating biomass and carbon stocks in the temperate forests of North- Western Mexico. Forests 8, 1–20. https://doi.org/10.3390/f8080269.

Wang, M., Liu, Q., Fu, L., Wang, G., Zhang, X., 2019. Airborne LIDAR-derived aboveground biomass estimates using a hierarchical bayesian approach. Remote Sens. 11. https://doi.org/10.3390/rs11091050.

Westaway, S., Smith, J., 2020. Elm Farm: Integrating Productive Trees and Hedges Into a Lowland Livestock Farm.

World Business Council for Sustainable Development and World Resources Institute, 2020. The Greenhouse Gas Protocol. A Corporate Accounting and Reporting Standard. . https://ghgprotocol.org/corporate-standard.

Zapata-Cuartas, M., Sierra, C.A., 2014. Alleman, L. Probability distribution of allometric coefficients and Bayesian estimation of aboveground tree biomass. Available to download: For. Ecol. Manag. 2012 (277), 173–179. https://www.sciencedirect.com/science/article/pii/S0378112712002484.

Zianis, D., 2008. Predicting mean aboveground forest biomass and its associated variance. For. Ecol. Manag. 256, 1400–1407. https://doi.org/10.1016/j.foreco.2008.07.002.

Zianis, D., Muukkonen, P., Mäkipää, R., Mencuccini, M., 2005. Biomass and stem volume equations for tree species in Europe. Silva Fennica Monographs.

Ziegler, J., Easter, M., Swan, A., Brandle, J., Ballesteros, W., Domke, G., Chambers, A., Eve, M., Paustian, K., 2016. A model for estimating windbreak carbon within COMET-FarmTM. Agrofor. Syst. 90, 875–887. https://doi.org/10.1007/s10457-016-9977-0.