# Evaluation of the ECOSSE model to predict heterotrophic soil respiration by direct measurements

M. DONDINI[a], G. ALBERTI[b,c], G. DELLE VEDOVE[c], M. VENTURA[d], G. TONON[d], M. VIGER[b], Z.M. HARRIS[b], J.R. JENKINS[b], M. RICHARDS[a], M. POGSON[a,e], G. TAYLOR[c], J. U. SMITH[a] & P. SMITH[a]

[a]*Institute of Biological and Environmental Sciences, School of Biological Sciences, University of Aberdeen, 23 St Machar Drive, Aberdeen, UK,* [b]*Centre for Biological Sciences, Faculty of Natural & Environmental Sciences, University of Southampton, Highfield Campus, Southampton, UK,* [c]*Department of Agricultural, Food, Environmental and Animal Sciences, University of Udine, Italy,* [d]*Faculty of Science and Technology, Free University of Bolzano, Universitätsplatz 5 - piazza Università, 5, Bozen-Bolzano,* [e]*Department of Applied Mathematics, Liverpool John Moores University, Faculty of Engineering and Technology, 3 Byrom Street, Liverpool, UK.*

Correspondence: Marta Dondini. E-mail: marta.dondini@abdn.ac.uk.

*Running title: Evaluating the ECOSSE model by direct measurements*

## Summary

This paper aims to evaluate the suitability of the ECOSSE model to estimate soil heterotrophic respiration ($R_h$) from arable land, and short rotation coppices of poplar and willow. Between 2011 and 2013, we measured $R_h$ with automatic closed dynamic chambers on root exclusion plots at one site in the United Kingdom (willow, mixed commercial genotypes of *Salix* spp.) and two sites in Italy (arable and poplar, *Populus x Canadensis* Moench, Oudemberg genotype), and compared these measured fluxes to simulated values of $R_h$ with the ECOSSE model. Correlation coefficients (*r*) between modelled and measured monthly $R_h$ data were strong and significant with a range between 0.81 and 0.96 for all three types of vegetation. There was no significant error and bias in the model for any site. The model was able to predict seasonal trends in $R_h$ at all three sites even though it occasionally underestimated the flux values during warm weather in spring and summer. Because of the strong correlation between the measured and modelled values, it is unlikely that underestimation of the flux is the result of missing processes in the model. Therefore, further detailed monitoring of $R_h$ is needed to modify the model. In this research, a limited set of input data was used to simulate $R_h$ at the three sites. Nevertheless, overall results of the model evaluation suggest that the ECOSSE model simulates soil $R_h$ adequately under all land uses tested and that continuous and direct measurements (such as automatic chambers installed on root-exclusion plots) are a useful tool to test model performance to simulate $R_h$ at the site level.



*Keywords: soil process-based model, $CO_2$ emission, willow, poplar, arable, modelling*


## Highlights

- Model evaluation is crucial to predict soil carbon balance accurately.
- Modelled and measured heterotrophic respiration were compared for three land uses.
- The model performed well statistically for all three vegetation types.
- Modelled heterotrophic respiration should be evaluated by comparison to continuous measurements.

## Introduction

Globally, the soil releases around 60 Gt of carbon (C) to the atmosphere each year through soil-surface carbon dioxide ($CO_2$) efflux, which is a major component of the global fluxes of $CO_2$ (Giardina *et al*., 2014). It is, therefore, an important regulator of climate change as well as a determinant of the terrestrial C balance (Yan *et al*., 2015).

Soil respiration ($R_s$) is generally expressed as the sum of soil $CO_2$ efflux from both root respiration (autotrophic respiration, $R_a$) and organic C and the mineralization and decomposition of litter (heterotrophic respiration, $R_h$; Bowden *et al*., 1993). Several methods have been used to separate $R_a$ and $R_h$ from the overall $R_s$, under both laboratory and field conditions, and over a range of spatial and temporal scales (Subke *et al*., 2006). Separation of $R_s$ into $R_a$ and $R_h$ is important to understand the processes that underlie total $R_s$, and to enable predictions of soil C under changing environmental conditions such as climate and land-use type. Ryan & Law (2005) grouped the methods to separate autotrophic and heterotrophic contributions into four categories: (i) comparison of $R_s$ determined from soil with roots excluded (usually by trenching) and intact soil, (ii) summation of the individual components of root respiration and litter decomposition, (iii) stable or radioactive isotope methods to determine the origin of the C and (4) ring barking around a tree's circumference (girdling) of the cambium, which cuts off the supply of photosynthates to roots. Several authors have reviewed the advantages and disadvantages of all these approaches for determining autotrophic and heterotrophic contributions to $R_s$ (Kuzyakov, 2006; Subke *et al*., 2006). These authors showed that the most reliable methods for the separation of $R_s$ into its constituent parts are based on stable isotope techniques because they involve less disturbance to the soil–plant system than root exclusion or component integration techniques (Kuzyakov, 2006). The bomb-$^{14}$C approach allows $CO_2$ sources to be separated with the least disturbance, but the large costs of analysis and some uncertainties limit its application. In field experiments, where high costs limit the use of isotope approaches, the root exclusion techniques have been shown to produce accurate separation of $R_s$ into the plant and soil components (Rochette *et al*., 1999). Because of the considerable heterogeneity and inaccessibility of the soil medium and high cost of measurement instruments, $R_s$, and its subdivision into $R_a$ and $R_h$, remains the least well quantified component of the terrestrial C cycle (Trumbore, 2006). With these constraints, regional and global estimates of $R_s$ are imprecise, and modelling is critical to make progress in this area.

Several multi-pool models, such as RothC (Coleman & Jenkinson, 2005) and ECOSSE (Smith *et al*., 2010a) have been developed over the last decade to describe both short- and long-term responses of soil C to land use and changes in the climate. In general, all multi-pool models are conceptually similar: organic litter entering the soil is divided into pools of different decomposability. During decomposition of the litter pools, several C pools of organic matter are formed in the mineral soil with different turnover times. Decomposed soil C is either transferred into one or more pools or is released as $CO_2$. Decomposition of the C pools is typically described by first-order kinetics, which implies that the amount of heterotrophic biomass does not directly affect the decomposition rate of organic matter pools (Bauer *et al*., 2008).

The ECOSSE (estimation of carbon in organic soils–sequestration and emissions) model was developed to simulate the C and nitrogen (N) cycles and greenhouse gas (GHG) fluxes with minimal input data for both mineral and organic soil (Smith *et al*., 2010a,b). The ECOSSE model is based on principles used initially for mineral soil in the two ‘mother’ models, RothC and SUNDIAL (Smith & Glendining, 1996). The ECOSSE model follows these established models and uses a pool-type approach, which describe the soil organic matter (SOM) as pools of inert organic matter, humus, biomass, resistant plant material (RPM) and decomposable plant material (DPM; Smith *et al*., 2010a,b). During the decomposition process, material is exchanged between the SOM pools according to first-order rate equations, characterized by a specific rate constant for each pool that depends on temperature, moisture, vegetation cover and soil pH.

Previous evaluations have determined the accuracy of ECOSSE simulations to predict soil C after land-use change to short rotation forestry (Dondini *et al*., 2015), *Miscanthus* and short rotation coppice willow (Dondini *et al*., 2016a). The modelled C under short rotation forestry showed a strong correlation with the soil C measurements at both 0–30 cm (correlation coefficient, $r = 0.93$) and 0–100 cm soil depth ($r = 0.82$, Dondini *et al*., 2015). Dondini *et al*. (2016a) also reported a strong correlation between modelled and measured soil organic C (SOC) after transition to *Miscanthus* and short rotation coppice-willow at two soil depths (0–30 and 0–100 cm), as well as the absence of significant bias in the model.

The ECOSSE model was also evaluated against soil nitrous oxide ($N_2O$) emissions from cropland sites in Europe (Smith *et al*., 2010b; Bell *et al*., 2012; Khalil *et al*., 2013), $CO_2$ emissions from peatlands (Abdalla *et al*., 2014) and all GHG fluxes under bioenergy and conventional crops (Dondini *et al*., 2016b). Previous evaluations of simulated $CO_2$ emissions compared model outputs against the $R_h$ derived from soil chamber and eddy covariance (EC) measurements. There were strong correlations between modelled and measured $R_h$ at different sites in the UK (Dondini *et al*., 2016b) and Europe (Abdalla *et al*., 2014), but both of these approaches have their limitations. The $R_h$ derived from the soil chamber measurements was estimated from periodic measurements of $R_s$, therefore, the degree of coincidence between measured and modelled $R_h$ was also related to the $R_h$:$R_s$ ratio adopted (Dondini *et al*., 2016b). The $R_h$ derived from EC measurements was estimated from the measured ecosystem respiration ($R_{eco}$) during daytime, which is a modelled flux driven by air temperature and other environmental factors (Dondini *et al*., 2016b). Therefore, further evaluation by comparison of the model output with direct measurements of soil $R_h$ is needed to demonstrate further the ability of the ECOSSE model to predict such a flux adequately.

In this paper we evaluate the suitability of the ECOSSE model for estimating soil $R_h$ at three independent sites that represent three different vegetation types, namely willow, poplar and arable land. Measured input data were used to initialize the model. At each site, automatic dynamic (non-steady state through flow) closed chambers were installed on field plots where roots had been excluded by the trenching method. This measurement technique provides continuous and direct measurements of $R_h$ and therefore enables a more accurate evaluation of the performance of the model than methods that use discontinuous measurements. Our research hypothesis was that the soil $R_h$ estimated by the ECOSSE model is statistically comparable to the measured $R_h$ at the three study sites.

## Materials and methods

*ECOSSE model*

The ECOSSE model simulates soil C and N dynamics in both mineral and organic soil. All of the major processes of C and N turnover in soil are included in the model, but each of the processes is simulated by simple equations and using readily available input variables. This enables the model to be developed from a field based model to a national scale tool, without great loss of accuracy (Smith *et al*., 2010a,b,c).

The ECOSSE model describes SOM by the following five pools: inert organic matter, humus (HUM), biomass (BIO), RPM and DPM. Each pool decomposes with a specific rate constant, except for the inert organic matter which is not affected by decomposition. The rate constants used are those given in RothC: for HUM = 0.02 year$^{-1}$, for BIO = 0.66 year$^{-1}$, for RPM = 0.3 year$^{-1}$ and DPM = 10 year$^{-1}$.

The ECOSSE model simulates the soil profile to a depth of 3 m; it divides the soil into 5-cm layers to simulate soil processes accurately with depth. Plant C and N inputs are added monthly to the DPM and RPM pools. During the decomposition process, material is exchanged between the SOM pools according to first-order equations, characterized by a specific decomposition rate for each pool. The decomposition rate of each pool is modified by temperature, water content, plant cover and pH of the soil (with additional modifiers that depend upon soil bulk density and inorganic N concentration in the case of anaerobic decomposition; Smith *et al.*, 2010c). The decomposition process results in $R_h$ and gaseous losses of methane ($CH_4$); $R_h$ dominates under aerobic conditions and $CH_4$ losses under anaerobic conditions. In ECOSSE, $CH_4$ emissions are calculated as the difference between $CH_4$ production and oxidation. Methane production during anaerobic decomposition is simulated by a similar pool approach to that used for aerobic decomposition. The difference between the rates of aerobic and anaerobic decomposition is simulated by the different functions used to calculate the rate modifiers, which account for changes in soil moisture, temperature, pH and water availability. ECOSSE also simulates the oxidation of atmospheric $CH_4$, which, under aerobic conditions, can lead to the soil being a net consumer of $CH_4$ (Smith *et al.*, 2010c).

The N content of the soil follows the decomposition of SOM, with a stable C:N ratio defined for each SOM pool at a given pH, and N is either mineralized or immobilized to maintain that ratio. Nitrogen is released from decomposing SOM as ammonium ($NH_4^+$) and may then be immobilized or nitrified to nitrate ($NO_3^-$). Carbon and N may be lost from the soil by the processes of leaching $NO_3^-$, dissolved organic C and dissolved organic N, nitrification and denitrification to nitric oxide (NO) and $N_2O$, volatilization of ammonia or plant assimilation of $NO_3^-$ and $NH_4^+$. Carbon and N may be returned to the soil by plant input, application of inorganic fertilizers, atmospheric deposition or organic amendments (e.g. manure, crop residues). More detail on the structure and parameters of the model are given in Smith *et al*. (2010a,c).

Vegetation inputs to the soil are estimated by a modification of the Miami model (Lieth, 1973), a simple model that links the net primary production (NPP) to annual mean temperature and total precipitation. For a full description of the ECOSSE model and the plant input estimates refer to Smith *et al*. (2010a) and Dondini *et al*. (2016a).

The minimum input requirements of the ECOSSE model for site-specific simulations are:

- 30-year average monthly rainfall (mm) and temperature (°C),
- Monthly rainfall (mm), temperature (°C) and potential evapotranspiration (PET; mm),
- Initial soil C content (kg $ha^{-1}$),

- Soil depth at which soil properties have been measured (cm),
- Soil sand, silt and clay content (%),
- Soil bulk density (g $cm^{-3}$),
- Soil pH,
- Crop type for each simulation year.

Initialization of the model is based on the assumption that the soil is at a steady state under the initial land use at the start of the simulation (Smith *et al.*, 2010a). Therefore, the model uses a 'spin-up' approach to adjust plant inputs until measured and simulated values of SOC converge. More detail on model initialization is given in Dondini *et al*. (2016b).

*Data and flux measurements*

In 2012–2013, one willow (mixed commercial genotypes of SRC willow, *Salix* spp.) site and one poplar (*Populus x Canadensis* Moench, Oudemberg genotype) site were chosen for sampling in the UK and Italy, respectively. The poplar trees were planted originally in 2010 and were last harvested in March 2012, a month before the start of the measurement period. The willow site was converted from grassland in 2008 and harvested in March 2009. An arable site was sampled in Italy in 2011–2012. The latter site had been under irrigated maize (*Zea mays* L.) monoculture for the previous 30 years, but in 2007 crop rotation was introduced with  three years (2007–2009) of alfalfa (*Medicago sativa* L.), one year of maize (2010), one year (2011) of soya beans (*Glycine max* Merr.) followed by  maize (2012). Management of the soil also changed in 2007 from ploughing to minimum tillage cultivation. The willow site and the measurements made there contribute to the ELUM (Ecosystem Land Use Modelling

& Soil Carbon GHG Flux Trial) project (Harris *et al*., 2014). The poplar site and measurements made there contributed to the EU-FP7 project EuroChar (Biochar for Carbon Sequestration and Large-Scale Removal of GHG from the Atmosphere; Ventura *et al*., 2015). The arable site and measurements made there contributed to the National Research Programme 'CarboItaly' (Alberti *et al*., 2010).

At the beginning of each experiment, three sampling plots per field were selected randomly, and three soil cores were taken within each sampling plot. At the poplar and arable sites, soil samples were collected to a depth of 40 and 60 cm, respectively, whereas soil samples at the willow site were collected to a depth of 1 m. All soil samples were sieved to pass through a 2-mm sieve; a subsample of the sieved soil was oven-dried (105 °C for 12 hours) and subsequently ball-milled (Fritsch Planetary Mill, Idar-Oberstein, Germany). The soil samples were analysed for percentage carbon (%C) with a LECO TruSpec CN analyser (Leco, TruSpec CN, St. Joseph, MI, USA), bulk density, particle-size distribution and pH (Table 1). The measurements of the soil properties of the three soil samples were averaged for each site and were used as inputs to the model.

At each sampling plot, the trenching method was used to measure $R_h$ as explained in Alberti *et al*. (2010) for the arable site and in Ventura *et al*. (2015) for the poplar and willow sites. At the poplar site, three trenched subplots (50 cm × 50 cm) were established by digging trenches 60–cm deep and 15-cm wide in the central part of each plot in February 2012, in the middle of two planted rows. Before the trenches were refilled with the original soil, each subplot was isolated with a geotextile canvas (Typar®, Dupont, Wilmington, DE, USA) to prevent root growth into the trenched subplot, but to  allow gas and water exchange. At the willow site, the trenched subplots were isolated in February 2012 by a root exclusion stainless-steel pipe (32-

cm diameter, 40-cm height). At the arable site, as part of a long-term monitoring experiment started in 2007 (Alberti *et al*., 2010), the trenched subplots were prepared every year with the same stainless-steel pipe used at the willow site; they were inserted into the soil before sowing and removed just before the crop was harvested.

At each site, $R_h$ was measured using six automated closed dynamic chambers (two per plot). Each chamber, placed over a collar inserted into the soil for 3–4 cm, has a base area of 196 $cm^2$ and a free headspace volume of around 2000 $cm^3$. To avoid a wind induced pressure difference between the inside and outside of the chamber, a pressure vent was built following Xu *et al.* (2006) and placed on the top of the chamber. The deployment time (i.e. after the chamber's lid closure) was 120 s. A pump circulated the air from the chamber to an infra-red gas analyser in a closed system (IRGA, SBA4 PP-Systems, Amesbury, MA, USA); $CO_2$ concentration, vapour partial pressure and total air pressure data were recorded every 1.6 s. The chambers were operated sequentially by a CR1000 (Campbell Scientific, Logan, UT, USA) data logger. More detail on the soil respiration systems and how $R_h$ fluxes were computed are described in Delle Vedove *et al.* (2007), Alberti *et al.* (2010) and Delle Vedove *et al.* (2015). At the willow and poplar sites, the sampling frequency was every 2 and 4 hours, respectively. At the arable site, the measurement frequency was every 2 hours.

The $R_h$ data presented in this study were collected at the willow site from May 2012 to September 2013, at the poplar site from April 2012 to November 2013 and at the arable site from January 2012 to December 2013. Because of a technical malfunction of the chamber equipment, $R_h$ data were not collected in October 2012–February 2013 and in July 2013 at the willow site, in June–July 2013 at the poplar site and in March–April 2011 at the arable site.

At each location, monthly air temperature and precipitation for the 30 years before measurements started were used to calculate long-term averages (Table 2), which were used as input to the model. Air temperature and precipitation data were extracted from the E-OBS gridded dataset from the EU-FP6 project ENSEMBLES, provided by the ECA&D project (Haylock *et al*., 2008). This dataset is known as E-OBS and is publicly available (http://eca.knmi.nl/). At each site, air temperature and precipitation were recorded during the entire study period and monthly values were used as input to the model. The arable site was irrigated between June and August 2011 (276 mm) and in the same period of 2012 (269 mm); irrigation was included in the model by adding the water used for irrigation to the monthly precipitation. No irrigation was used at the other two sites. Monthly PET was estimated by the Thornthwaite method (Thornthwaite, 1948), which has been used in other modelling studies when directly observed data have not been available (e.g. Smith *et al*., 2005; Dondini *et al*., 2015).

*Model evaluation and statistical analysis*

The aim of this research was to evaluate the ability of the ECOSSE model to predict $R_h$ under different vegetation types; therefore, no model parameters or processes were implemented with the measurements taken at the three experimental sites. Instead, the model was evaluated with field data, i.e. independent data not used for developing the model.

At each site, measured soil C, bulk density, particle-size distribution, pH and meteorological data were used as inputs to run the ECOSSE model (see above for input details). Values of soil variables were available for different soil depths at the three sites (Table 1); therefore, the modelled $R_h$ values represent fluxes released at the soil surface from the upper 40-cm depth at the poplar site, from the upper 60-cm depth at the arable site and from 100-cm depth at the willow site.

Monthly simulations of soil $R_h$ fluxes at the soil surface were evaluated against mean monthly chamber measurements, also recorded at the soil surface.

The Shapiro–Wilk's test for normality was used to test the distribution of the measured $R_h$ values at each site with the IBM SPSS Statistics software, Version 24.0. This test failed to reject the null hypothesis of normality for the willow data ($P$ = 0.614), but it did reject the null hypothesis of normality for the poplar and arable data ($P$ = 0.021 and $P$ = <0.0001, respectively; Figure 1a). For each dataset, a general linear model was used to determine the residuals of the difference between the measured $R_h$ values and the sample mean. These residuals were also tested for normality by the Shapiro–Wilk's test, and the null hypothesis of normality was again rejected for the arable and poplar data ($P$ = 0.021 and $P$ = <0.0001, respectively; Figure 1b). Therefore, the arable and poplar data were transformed with the Box–Cox transformation. This transformation (Box & Cox, 1964) represents a family of power transformations that incorporates and extends the traditional options (e.g. square root, cube root, fourth root, natural logarithm, reciprocal square root transformations) to find the optimal normalizing transformation for each variable. The procedure identifies an appropriate exponent, Lambda, to transform data to a normal distribution. The Lambda value indicates the power to which all data should be raised. To do this, the Box–Cox power transformation searches for Lambda from –5 to +5 until the best value is found. In our study, this transformation suggested a Lambda value of 0.5 (i.e. the square root of the original data) and 0 (i.e. the natural logarithm of the original data) for transformation of $R_h$ values at the poplar and the arable sites, respectively. The Shapiro–Wilk's test for normality was again used to test the distribution of the transformed data and of the residuals of the difference between the transformed data and the sample mean. For both datasets (i. e. poplar and arable), the tests failed to reject the null hypothesis of normality for the transformed data and residuals ($P$ = 1.0 for all datasets analysed; Figure 1c,d). On the basis of these results, the statistical evaluation of the model performance to simulate $R_h$ was done on the transformed $R_h$ data for the poplar and arable sites and on non-transformed $R_h$ data for the willow site.

A quantitative statistical analysis was undertaken to determine the degree of coincidence and association between measured and modelled $R_h$ values, following the approach described in Smith *et al*. (1997) and Smith & Smith (2007). The analysis of association defines how well trends in the measured values relate to those that are simulated, and the analysis of coincidence determines the differences between the simulated and measured values.

The degree of association between modelled and measured $R_h$ values was determined with the sample correlation coefficient, $r$ (Chatfield, 1983). The significance of the association between simulated values and measurements was determined by the $F$-test (Armitage *et al*., 2002). The value of $F$ was calculated by:

$$F = \frac{(n-2) \times r^2}{(1 - r^2)}, \qquad (1)$$

where $n$ is the number of measured and simulated pairs being compared and $r$ is the sample correlation coefficient (Smith & Smith, 2007). The value of $F$ was related to the probability that the measured and simulated values were not associated by comparing to the $P$-values ($P$ = 0.05) of the $F$ distribution. If $F$ > $F$-value at ($P$ = 0.05) the association between modelled and measured values was considered statistically significant.

The analysis of coincidence between the simulated and measured values was determined from the total difference, the bias in the total difference and the goodness-of-fit between simulated and measured values. The total difference between the simulated and measured values was calculated as the root mean squared error (RMSE; Loague & Green, 1991). The statistical significance of the total difference between the simulated and measured $R_h$ was assessed by comparing the RMSE to the value obtained assuming a deviation corresponding to the 95% confidence interval of the replicated measurements ($RMSE_{95}$). If the relative error $RMSE < RMSE_{95}$ indicates that the simulated values fall within the 95% confidence interval of the measurements, the model cannot be improved further with these data (Smith & Smith, 2007).

The bias in the total difference between simulated and measured values was determined by calculating the relative error, $E$ (Addiscott & Whitmore, 1987):

$$E = \frac{100}{\bar{O}} \times \frac{\sum_{i=1}^{n}(O_i - P_i)}{n}, \qquad (2)$$

where $\bar{O}$ is the average of all measurements, $O_i$ is the $i$th measured value, $P_i$ is the $i$th simulated value and $n$ is the total number of values being compared.

The significance of $E$ was determined again by comparing its value to that obtained assuming a deviation corresponding to the 95% confidence interval of the measurements ($E_{95}$). If $E < E_{95}$ it indicates that the bias in the simulation is less than the 95% confidence interval of the measurements, and the model bias cannot be reduced further with these data (Smith & Smith, 2007).

The lack of fit statistic, *LOFIT* (Whitmore, 1991), was used to assess the goodness-of-fit between simulated and measured values. Assuming experimental errors to be random, this statistic enables the experimental errors to be distinguished from the failure of the model. The significance of *LOFIT* was determined with an $F$-test; in accord with statistical convention, a value of $F$ greater than the critical 5% $F$-

value was taken to indicate that the total error in the simulated values was significantly greater than the error inherent in the measured values.

## Results and discussion

### *Model evaluation*

The ECOSSE model was evaluated by comparing the output from the model to the measured $R_h$ fluxes from the three sites, which represent the following land uses: willow, poplar and arable (soya bean–maize rotation). The modelled $R_h$ was strongly and significantly correlated with the measured values at all sites, with *r* values of 0.81 (willow), 0.96 (poplar) and 0.83 (arable) (Table 3). The model evaluation also showed no significant difference between measured and modelled values (RMSE < $RMSE_{95}$), no bias in the total difference ($E < E_{95}$) and no significant model bias for all three types of vegetation (Table 3).

The model was able to predict seasonal trends in $R_h$ at all of the sites (Figure 1); at the poplar and arable sites, it occasionally underestimated the flux values during the warm weather in spring and summer compared to the measured $R_h$. At the poplar site, the modelled $R_h$ was estimated to be 2134 kg C $ha^{-1}$ from May to October 2012, against a measured $R_h$ value of 4676 kg C $ha^{-1}$ for the same period. At the arable site, the model estimated an $R_h$ of 1336 kg C $ha^{-1}$ from May to October 2011, whereas the $R_h$ measured at the same time was 3071 kg C $ha^{-1}$. The model predicts the $R_h$ that occurs only from the soil depth at which the soil characteristics have been measured, which were used as inputs to the model. The soil characteristics used to run the model for the poplar and arable sites were available at depths of 40 and 60 cm only, respectively. Therefore, the $R_h$ efflux that the model simulates at the soil surface is that which comes from these specific depths. On the other hand, the measured $R_h$

represents the flux from the whole soil profile; therefore, we expected the modelled $R_h$ to be underestimated compared to the measured values. For the willow site, measured values used as inputs to the model were from a depth of 1 m and so the model values of $R_h$ were underestimated less because they were related to fluxes from 1-m depth (2989 kg C $ha^{-1}$ modelled $R_h$ against 3858 kg C $ha^{-1}$ measured $R_h$ from April to September 2012).

Another possible explanation for the underestimated $R_h$ fluxes is that the soil might not have been in a steady state at the start of the simulation, which was assumed. If SOM was being lost from the soil instead of being in a steady state, then the rate of SOM decomposition would be underestimated, which means that the simulations would also underestimate $R_h$. Unfortunately, we do not have historical data to reject or accept this hypothesis. However, because there was no significant error between the simulated and measured values of $R_h$ and no model bias, it is unlikely that underestimation of the flux is due to missing processes in the model. If a model is evaluated against independent data, the evaluation could show an error, exposing the effect of the missing process. It is important to note the large variability in the measured values, which led to large $RMSE_{95}$ and $E_{95}$ values at the poplar and arable sites (Table 3), resulted in the calculated RMSE and $E$ values not being statistically significant. To reduce uncertainties in the evaluation of the model, it is advisable that $R_h$ is measured on more field plots than we used (i.e. $n > 3$). A larger number of field plots will lead to a greater accuracy in the measured $R_h$, less variation in the measured values and consequently a more accurate representation of the values against with the model will be evaluated.

The evaluation of a process-based model, such as ECOSSE, depends strictly on the quality, type and frequency of the measured values used to test the model. Therefore, it is a procedure that is in constant development. The first evaluation studies on the ability of ECOSSE to simulate $R_h$ were done with $R_h$ data from two different sampling methods, EC (Abdalla *et al.*, 2014; Dondini *et al*., 2016b) and chamber methods (Dondini *et al*., (2016b). Dondini *et al*. (2016b) evaluated the suitability of the ECOSSE model to estimate soil GHG fluxes from short rotation coppice willow, short rotation forestry (*Pinus sylvestris* L.) and *Miscanthus* after land-use change from conventional systems (grassland and arable). The $R_h$ was simulated at four paired sites in the UK and compared to estimates of $R_h$ derived from the ecosystem respiration estimated from EC and $R_h$ determined from monthly chamber (IRGA) measurements. The correlations between modelled and measured $R_h$ were weak when model values were compared with the values from the chambers (Dondini *et al*., 2016b). The discrepancy between modelled- and chamber-derived $R_h$ appeared to be due to the nature of the chamber-derived $R_h$, which was not related to the soil processes described in the model. The chamber-derived $R_h$ was estimated from direct measurements of total soil respiration, therefore the degree of correlation between measured and modelled $R_h$ was also related to the $R_h$:$R_s$ ratio adopted. In addition to this, the chamber-derived $R_h$ was estimated from a single data point which was taken to represent monthly total soil respiration. Dondini *et al*. (2016b) suggested that direct and continuous measurements of $R_h$ would be needed to test these hypotheses and to evaluate the ECOSSE model further. The results from the current study for the willow site can be compared directly to the aforementioned study by Dondini *et al*. (2016b). At the willow site the correlations between EC-derived $R_h$ and chamber-derived $R_h$ were 0.77 and 0.75, respectively, whereas the correlation coefficient from the present study at this site was stronger ($r$ = 0.81) with direct and continuous measurements of

$R_h$. The present study, therefore, reinforces former findings and improves on previous evaluations of the ECOSSE model.

*Use of direct measurements as a tool to test model simulation*

In the present study, the trenching method was applied to measure $R_h$ at three experimental sites, and subsequently to compare its value to the ECOSSE output. This technique to separate soil $CO_2$ flows has been used successfully before to measure $R_h$ under different vegetation types and climatic conditions (Saiz *et al*., 2006; Ventura *et al*., 2015). Kuzyakov (2006) reviewed the existing approaches to estimate the contribution of individual sources to total soil $CO_2$ efflux, but he found no single satisfactory partitioning method. The study reported that the most reliable methods for the separation of root-derived from SOM-derived $CO_2$ are based on isotopes. However, in situations where high costs or the lack of appropriate expertise or both might limit the use of isotope approaches, future investigators might consider the root exclusion techniques. In a comparative study of root exclusion and isotopic approaches, Rochette *et al*. (1999) found that $^{13}C$ isotopic labelling and root exclusion methods produced similar values for root respiration, and concluded that both approaches were useful to partition total soil respiration. The main concern with the trenching technique is that it results in a considerable increase in dead root biomass in the treated plots, which can lead to an increase in the measured $R_h$ (Subke *et al*., 2006). This issue is generally acknowledged by authors and the root decay in trenched plots is often measured, estimated or derived from other published studies to correct the measured $R_h$. In a review of partitioning methods, Subke *et al*. (2006) reported that, if the additional root decay in trenched plots is taken into account, the $R_h$ contribution to $R_s$ would be reduced by, on average, 12%. The considerable range of decay constants observed by Subke *et al*. (2006) indicates that root decay depends strongly on C lost as $CO_2$, which suggests that these variables depend on local conditions (e.g. soil type, climate or litter quality). The authors therefore recommended that the fine and coarse root biomass should be measured for each area at the beginning and at the end of any root exclusion experiment, and that root decay should be measured independently. Because of cost limitations in the present study, it was not possible to measure the rate of root decay. Nevertheless, we can exclude any possible effect of roots within the root exclusion plots at the arable site because the trenched plots were set up before sowing. At the willow and poplar sites the root exclusion plots were placed between tree rows, therefore root respiration should be minimal. Despite this aspect, the model was able to simulate soil $R_h$ with a good degree of accuracy at all three sites.

## Conclusions

We used a limited set of input data to simulate $R_h$ at three sites in Europe with the ECOSSE model, and the output predicted seasonal trends in $R_h$ at all of the sites. The correlation between measured and modelled values was strong ($r$ ranged from 0.81 to 0.93) and statistically significant. The total difference between the simulated and measured values and the 'lack-of fit' statistical analyses showed no significant differences between modelled and measured $R_h$, suggesting that the ECOSSE model can simulate soil $R_h$ adequately under all land uses tested (willow, poplar and arable).

The overall results of the present study also emphasized that continuous and direct measurements (such as automatic chambers installed on root-exclusion plots) are a useful tool to test the model's simulation of $R_h$ at the site level. Furthermore,

more chambers and experimental plots should be used to monitor $R_h$ where soil conditions are very variable.

## Acknowledgements

This work contributes to the ELUM (Ecosystem Land Use Modelling & Soil Carbon GHG Flux Trial) project, which was commissioned and funded by the Energy Technologies Institute (ETI), and to Carbo-BioCrop (www.carbobiocrop.ac.uk; a NERC funded project; NE/H010742/1), UKERC Phase II and III (NERC; NE/H013237/1), MAGLUE (www.maglue.ac.uk; an EPSRC funded project; EP/M013200/1) and as part of the Seventh Framework For Research Programme of the EU, within the EUROCHAR project (N 265179) and EXPEER within WU FP7-Infrastructures. We acknowledge the use of the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu). We thank two anonymous reviewers and Dr. William van Dijk for their valuable suggestions.

## References

Abdalla, M., Hastings, A., Bell, M.J., Smith, J.U., Richards, M., Nilsson, M.B. *et al*. 2014. Simulation of $CO_2$ and attribution analysis at six European peatland sites using the ECOSSE model. *Water, Air, and Soil Pollution*, **225**, 2182–2196.

Addiscott, T.M. & Whitmore, A.P. 1987. Computer simulation of changes in soil mineral nitrogen and crop nitrogen during autumn, winter and spring. *Journal of Agricultural Science*, **109**, 141–157.

Alberti, G., Delle Vedove, G., Zuliani, M., Peressotti, A., Castaldi, S. & Zerbi, G. 2010. Changes in $CO_2$ emissions after crop conversion from continuous maize to alfalfa. *Agriculture, Ecosystems & Environment*, **136**, 139–147.

Armitage, P., Berry, P.J. & Matthews, J.N.S. 2002. *Statistical Methods in Medical Research*. Blackwell Publishing, Oxford.

Bauer, J., Herbst, M., Huisman, J.A., Weihermüller, L. & Vereecken, H. 2008. Sensitivity of simulated soil heterotrophic respiration to temperature and moisture reduction functions. *Geoderma*, **145**, 17–27.

Bell, M.J., Jones, E., Smith, J., Smith, P., Yeluripati, J., Augustin, R. *et al*. 2012. Simulation of soil nitrogen, nitrous oxide emissions and mitigation scenarios at 3 European cropland sites using the ECOSSE model. *Nutrient Cycling in Agroecosystems*, **92**, 161–181.

Bowden, R.D., Nadelhoffer, K.J., Boone, R.D., Mellilo, J.M. & Garrison, J.B. 1993. Contributions of aboveground litter, belowground litter, and root respiration to total soil respiration in a temperate mixed hardwood forest. *Canadian Journal of Forest Research*, **23**, 1402–1407.

Box, G.E.P. & Cox, D.R. 1964. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, **26**, 211–252.

Chatfield, C. 1983. *Statistics for Technology: a Course in Applied Statistics. Vol. 3.* CRC Press, Boca Raton, FL.

Coleman, K. & Jenkinson, D.S. 2005. *ROTHC-26.3. A Model for the Turnover of Carbon in Soil. Model Description and Windows Users Guide* [WWW document].

URL www.rothamsted.ac.uk/sustainable-soils-and-grassland-systems/rothamsted-carbon-model-rothc [accessed on 13 January 2017].

Delle Vedove, G., Alberti, G., Peressotti, A., Inglima, I., Zuliani, M. & Zerbi, G. 2007. Automated monitoring of soil respiration: an improved automatic chamber system. *Italian Journal of Agronomy*, **2**, 377–382.

Delle Vedove, G., Grignani, C. & Bertora, C. 2015. Greenhouse gas emissions from soils. In: *Technical Report – Distributed Infrastructure for Experimentation in Ecosystem Research (ExpeER)*, pp. 65–86. European Commission Seventh Framework Programme. Grant Agreement Number 262060, France.

Dondini, M., Jones, E.O., Richards, M., Pogson, M., Rowe, R.L., Keith, A.M. *et al*. 2015. Evaluation of the ECOSSE model for simulating soil carbon under short rotation forestry energy crops in Britain. *GCB Bioenergy*, **7**, 527–540.

Dondini, M., Richards, M., Pogson, M., Jones, E.O., Rowe, R.L., Keith, A.M. *et al*. 2016a. Evaluation of the ECOSSE model for simulating soil carbon under *Miscanthus* and short rotation coppice - willow crops in Britain. *GCB Bioenergy*, **8**, 790–804.

Dondini, M., Richards, M., Pogson, M., McCalmont, J., Drewer, J., Marshall, R. *et al*. 2016b. Simulation of greenhouse gases following land-use change to bioenergy crops using the ECOSSE model. A comparison between site measurements and model predictions. *GCB Bioenergy*, **8**, 925–940.

Giardina, C.P., Litton, C.M., Crow, S.E. & Asner, G.P. 2014. Warming-related increases in soil $CO_2$ efflux are explained by increased below-ground carbon flux. *Nature Climate Change*, **4**, 822–827.

Harris, Z.M., McNamara, N.P., Rowe, R., Dondini, M., Finch, J., Perks, M. *et al*. 2014. Research Spotlight: The ELUM project: Ecosystem Land-Use Modeling and Soil Carbon GHG Flux Trial. *Biofuels*, **5**, 111–116.

Haylock, M.R., Hofstra, N., Klein Tank, A.M.G., Klok, E.J., Jones, P.D. & New, M. 2008. A European daily high-resolution gridded dataset of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research (Atmospheres)*, **113**, D20119.

Khalil, M.I., Richards, M., Osborne, B., Williams, M. & Müller, C. 2013. Simulation and validation of greenhouse gas emissions and SOC stock changes in arable land using the ECOSSE model. *Atmospheric Environment*, **81**, 616–624.

Kuzyakov, Y. 2006. Sources of $CO_2$ efflux from soil and review of partitioning methods. *Soil Biology & Biochemistry*, **38**, 425–448.

Lieth, H. 1973. Primary production: terrestrial ecosystems. *Human Ecology*, **1**, 303–332.

Loague, K. & Green, R.E. 1991. Statistical and graphical methods for evaluating solute transport models: overview and application. *Journal of Contaminant Hydrology*, **7**, 51–73.

Rochette, P., Flanagan, L.B. & Gregorich, E.G. 1999. Separating soil respiration into plant and soil components using analyses of the natural abundance of carbon-13. *Soil Science Society of America Journal*, **63**, 1207–1213.

Ryan, M.G. & Law, B.E. 2005. Interpreting, measuring, and modeling soil respiration. *Biogeochemistry*, **73**, 3–27.

Saiz, G., Byrne, K.A., Butterbach-Bahl, K., Kiese, R., Blujedea, V. & Farrell, E.P. 2006. Stand age related effects on soil respiration in a first rotation Sitka spruce chronosequence in central Ireland. *Global Change Biology*, **12**, 1007–1020.

Smith, J.U. & Glendining, M.J. 1996. A decision support system for optimising the use of nitrogen in crop rotations. Rotations and cropping systems. *Aspects of Applied Biology*, **47**, 103–110.

Smith, J.U., Gottschalk, P., Bellarby, J., Chapman, S., Lilly, A., Towers, W. *et al*. 2010a. Estimating changes in national soil carbon stocks using ECOSSE – a new model that includes upland organic soils. Part I. Model description and uncertainty in national scale simulations of Scotland. *Climate Research*, **45**, 179–192.

Smith, J.U., Gottschalk, P., Bellarby, J., Chapman, S., Lilly, A., Towers, W. *et al*. 2010b. Estimating changes in national soil carbon stocks using ECOSSE-a new model that includes upland organic soils. Part II. Application in Scotland. *Climate Research*, **45**, 193–205.

Smith, J.U., Gottschalk, P., Bellarby, J., Richards, M., Nayak, D., Coleman, K. *et al*. 2010c. *ECOSSE User Manual*. University of Aberdeen, Aberdeen. www.abdn.ac.uk/staffpages/uploads/soi450/ECOSSE%20User%20manual%20310810.pdf [accessed on 29 February 2016].

Smith, J.U. & Smith, P. 2007. *Environmental Modelling. An Introduction*. Oxford University Press, Oxford.

Smith, J.U., Smith, P., Wattenbach, M., Zaehele, S., Hiederer, R., Jones, R.J.A. *et al.* 2005. Projected changes in mineral soil carbon of European croplands and grasslands, 1990–2080. *Global Change Biology*, **11,** 2141–2152.

Smith, P., Smith, J.U., Powlson, D.S., McGill, W.B., Arah, J.R.M., Chertov, O.G. *et al*. 1997. A comparison of the performance of nine soil organic matter models using datasets from seven long-term experiments. *Geoderma*, **81**, 153–225.

Subke, J.-A., Inglima, I. & Cotrufo, M.F. 2006. Trends and methodological impacts in soil $CO_2$ efflux partitioning: A metaanalytical review. *Global Change Biology*, **12**, 921–943.

Thornthwaite, C.W. 1948. An approach toward a rational classification of climate. *Geographical Review*, **38**, 55–94.

Trumbore, S.E. 2006. Carbon respired by terrestrial ecosystems–recent progress and challenges, *Global Change Biology*, **12**, 141–153.

Vargas, R., Baldocchi, D.D., Allen, M.F., Bahan, M., Black, A., Collins, S.L. *et al*. 2010. Looking deeper into the soil: biophysical controls and seasonal lags of soil $CO_2$ production and efflux. *Ecological Applications*, **20**, 1569–1582.

Ventura, M., Alberti, G., Viger, M., Jenkins, J.R., Girardin, C., Baronti, S. *et al.* 2015. Biochar mineralization and priming effect on SOM decomposition in two European short rotation coppices. *GCB Bioenergy*, **7**, 1150–1160.

Whitmore, A.P. 1991. A method for assessing the goodness of computer simulations of soil processes. *Journal of Soil Science*, **42**, 289–299.

Xu, L., Furtaw, M.D., Madsen, R.A., Garcia, R.L., Anderson, D.J. & McDermitt, D.K. 2006. On maintaining pressure equilibrium between a soil $CO_2$ flux chamber and the ambient air. *Journal of Geophysical Research: Atmospheres*, **111**, D08S10.

Yan, M., Guo, N., Ren, H., Zhang, X. & Zhou, G. 2015. Autotrophic and heterotrophic respiration of a poplar plantation chronosequence in northwest China. *Forest Ecology and Management*, **337**, 119–125.

## FIGURE CAPTIONS

**Figure 1** Histograms of (a) $R_h$ data and (b) $R_h$ residuals from for the three experimental sites, and distribution of (c) the Box–Cox transformed $R_h$ data and (d) $R_h$ residuals after Box–Cox transformation for the arable and poplar sites. Line represents a normal distribution.

**Figure 2** Measured (filled triangle) and modelled (solid line with circle markers) monthly heterotrophic respiration ($R_h$) under (a) willow, (b) poplar and (c) arable during the measurement periods. Vertical bars are 95% confidence interval of the measured values. The $R_h$ data were not measured in October 2012–February 2013 and in July 2013 at the willow site, in June–July 2013 at the poplar site and in March–April 2011 at the arable site.

## TABLES

**Table 1** Land-use type, coordinates and soil characteristics of the study sites.

| Land-use, location | Latitude, longitude | Soil depth /cm | Soil bulk density /g cm$^{-3}$ | pH | Clay /% | Silt | Sand | Soil carbon /t C ha$^{-1}$ |
|---|---|---|---|---|---|---|---|---|
| Willow, West Sussex UK | 50.9 N, 0.4 E | 100 | 1.2 | 6.0 | 10 | 60 | 30 | 292 |
| Poplar, Prato Stesia IT | 45.6 N, 8.4 E | 40 | 1.4 | 5.4 | 12 | 34 | 54 | 88 |
| Arable, Beano IT | 46.0 N, 13.0 E | 60 | 1.1 | 7.1 | 15 | 58 | 27 | 72 |

**Table 2** Long-term (30 years) average precipitation, potential evapotranspiration (PET) and temperature at the study sites.

| | Arable | | | Poplar | | | Willow | | |
|---|---|---|---|---|---|---|---|---|---|
| | Precipitation /mm | PET /mm | Temperature °C | Precipitation /mm | PET /mm | Temperature /°C | Precipitation /mm | PET /mm | Temperature /°C |
| January | 46 | 6 | 4 | 45 | 4 | 2 | 80 | 16 | 16 |
| February | 42 | 10 | 5 | 37 | 10 | 4 | 54 | 18 | 18 |
| March | 64 | 27 | 9 | 64 | 30 | 8 | 55 | 30 | 30 |
| April | 87 | 55 | 13 | 102 | 53 | 12 | 46 | 48 | 48 |
| May | 89 | 96 | 18 | 125 | 89 | 16 | 47 | 73 | 73 |
| June | 91 | 127 | 21 | 98 | 121 | 20 | 48 | 95 | 95 |
| July | 73 | 146 | 24 | 74 | 140 | 23 | 49 | 110 | 110 |
| August | 78 | 135 | 23 | 83 | 128 | 22 | 52 | 103 | 103 |
| September | 100 | 92 | 19 | 97 | 88 | 18 | 60 | 79 | 79 |
| October | 98 | 52 | 14 | 93 | 49 | 13 | 99 | 51 | 51 |
| November | 93 | 22 | 9 | 95 | 19 | 7 | 88 | 29 | 29 |
| December | 83 | 8 | 5 | 48 | 6 | 3 | 86 | 18 | 18 |

**Table 3** Evaluation of the ECOSSE model to simulate heterotrophic respiration ($R_h$) at the study sites. Association is significant if *F*-value > *F*-value at ($P$ = 0.05). Error between measured and modelled values is not significant for RMSE < $RMSE_{95}$. Relative error is not significant for $E < E_{95}$. Lack of fit is significant if *F*-value > *F*-value at ($P$ = 0.05).

| Statistic | Willow | Poplar* | Arable* |
|---|---|---|---|
| *r* (Correlation Coefficient) | 0.8 | 0.96 | 0.8 |
| *F*-value | 4.2 | 175.2 | 43.4 |
| *F*-value at ($P$ = 0.05) | 2.3 | 4.5 | 4.4 |
| RMSE (Root mean square error of model)/% | 26 | 62 | 59 |
| $RMSE_{95}$ (95% Confidence Limit)/% | 54 | 104 | 217 |
| E (Relative Error) | 18 | 56 | 48 |
| $E_{95}$ (95% Confidence Limit). | 50 | 88 | 196 |
| *LOFIT* (Lack-of-fit) | | | |
| *F*-value | 0.03 | 0.6 | 0.4 |
| *F*-value at ($P$ = 0.05) | 2 | 1.7 | 1.7 |
| Number of values (months) | 11 | 18 | 22 |

*Statistical analysis of poplar and arable sites was done on transformed data