Agricultural and Forest Meteorology 335 (2023) 109469

Contents lists available at ScienceDirect

Agricultural and Forest Meteorology

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

Fine scale assessment of seasonal, intra-seasonal and spatial dynamics of soil CO_2 effluxes over a balsam fir-dominated perhumid boreal landscape

Antoine Harel [a,b,*], Jean-Daniel Sylvain [b,*], Guillaume Drolet [b], Evelyne Thiffault [a], Nelson Thiffault [c,d], Sylvie Tremblay [b]

[a] Research Centre on Renewable Materials, Department of Wood and Forest Sciences, Laval University, Quebec City, QC G1V 0A6, Canada
[b] Direction de la recherche forestière, Ministère des Ressources naturelles et des Forêts, 2700, rue Einstein, Québec, QC G1P 3W8, Canada
[c] Canadian Wood Fibre Centre, Natural Resources Canada, 1055 du P.E.P.S., P.O. Box 10380, Sainte-Foy Stn., Québec, QC G1V 4C7, Canada
[d] Centre d'étude de la forêt, Department of Wood and Forest Sciences, Laval University, Quebec City, QC G1V 0A6, Canada

ARTICLE INFO

Keywords:
Soil CO_2 effluxes
Boreal perhumid landscape
Spatiotemporal dynamics
Topographic attributes
Recent weather
Canopy height model

ABSTRACT

Total soil CO_2 efflux (F_{CO2}) is the second most important carbon flux after photosynthesis in boreal forests. However, accurate modelling of F_{CO2} remains challenging because of its high variability, both temporally and spatially. Using an *Abies balsamea*-dominated boreal landscape in Quebec (eastern Canada) as a case study, we modelled seasonal, intra-seasonal and spatial variability of F_{CO2} using climate variables and topographic and canopy structure attributes derived from Light Detection and Ranging (LiDAR) and assessed their respective contributions to soil CO_2 emissions. Weekly point measurements of F_{CO2} at 99 sites were taken over an area of 122 ha between June and October 2020. The seasonal component of F_{CO2} was quantified and subtracted from F_{CO2} measurements to isolate the spatial and intra-seasonal components of the flux. The two components were then modelled using a *Random Forest Regression* model and studied using accumulated local effect plots (ALE plots). Our approach explained 81% of the variation in F_{CO2}: the seasonal pattern explained 36% of the variation in F_{CO2} measurements, while spatial and intra-seasonal patterns together explained 45%. The most important factors explaining spatial variation were vegetation height and the slope height. Average air temperature of the last two days before efflux measurements was the most important factor explaining intra-seasonal variation. The proposed methodology makes it possible to predict F_{CO2} from external factors derived from climate and remote sensing data and enables the decomposition of F_{CO2} into its seasonal, intra-seasonal and spatial components. Our results demonstrate the importance of spatial and intra-seasonal variations in F_{CO2} compared to seasonal variation, a finding that has implications for the measurement and modelling of F_{CO2} at landscape and global scales.

1. Introduction

The boreal forest is the largest terrestrial biome on Earth (Pan et al., 2011); it stores a significant amount of organic carbon, particularly in the soil (Bradshaw and Warkentin, 2015). In forest ecosystems, the carbon balance depends on carbon uptake via photosynthesis and carbon release through autotrophic and heterotrophic respiration (Lindroth et al., 1998; Malhi et al., 1999; Kurz et al., 2013). Soil respiration can account for between 48 and 71% of total ecosystem respiration in boreal forest stands (Lavigne et al., 1997). While the boreal forest is currently considered as a global carbon sink (Pan et al., 2011), it could become a carbon source under global warming, as increasing temperatures could lead to an increase in soil organic matter decomposition through heterotrophic soil respiration (Schindlbacher et al., 2012; Kurz et al., 2013; Bond-Lamberty et al., 2018; Tremblay et al., 2018; Marty et al., 2019).

Total soil CO_2 efflux (F_{CO2}) represents the amount of CO_2 moving from the soil to the atmosphere. It depends mainly on biological processes (i.e., respiration by autotrophic and heterotrophic organisms) (Kuzyakov, 2006) and on the physical characteristics of the soil that affect CO_2 diffusion, such as texture, tortuosity or water content (Currie, 1961; Kimball, 1983; Moldrup et al., 2001). At the very small scale, F_{CO2} depends on soil conditions controlled by environmental and biotic factors (i.e., "internal soil factors"): soil temperature (Lloyd and Taylor, 1994), soil water content (Orchard and Cook, 1983; Linn and Doran, 1984), soil organic matter content (Trumbore, 2000), the abundance and composition of the soil microbial community (Carney et al., 2007;

* Corresponding authors.
E-mail addresses: antoine.harel.1@ulaval.ca (A. Harel), jean-daniel.sylvain@mffp.gouv.qc.ca (J.-D. Sylvain).

https://doi.org/10.1016/j.agrformet.2023.109469
Received 20 January 2023; Received in revised form 29 March 2023; Accepted 11 April 2023
Available online 23 April 2023

Abbreviations and acronyms

A_i	Median predicted total soil CO_2 efflux for the study area at day i that reflects the seasonal component of total soil CO_2 efflux
ALE	Accumulated local effects
DEM	Digital elevation model
DOY	Day of the year
DSM	Digital surface model
F_{SOM}	Soil CO_2 efflux from decomposition of soil organic matter
F_{CO2}	Total soil CO_2 efflux
F_{CO2ij}	Total soil CO_2 efflux at time i and location j
$\widehat{F_{CO2}}$	Predicted total soil CO_2 efflux ($A_i + W_i + S_j$)
LiDAR	Light Detection and Ranging
MDI	Mean decrease in impurity
MRVBF	Multiresolution index of valley bottom flatness
Q_{10}	Rate of change of soil respiration for a temperature increase of 10 °C
R_{Fij}	Intra-seasonal and spatial components of total soil CO_2 efflux at time i and location j
RMSE	Root mean square error
R^2	Coefficient of determination
S_j	Spatial component of total soil CO_2 efflux at location j
SOM	Soil organic matter
W_i	Component of the total soil CO_2 efflux related to recent weather of the i^{th} day

Allison et al., 2010), and local root activity (the sum of root respiration and the amount of root exudates secreted into the soil).

Other biotic and abiotic factors (i.e., "external soil factors") linked to vegetation, topography and recent weather indirectly affect F_{CO2} by influencing internal factors. Vegetation influences F_{CO2} through its effect on soil temperature (Song et al., 2013), soil water content (Kergoat, 1998), quality and quantity of soil organic matter (Raich and Tufekciogul, 2000; Trumbore, 2000) and local root activity via root biomass (Han et al., 2007; Vargas and Allen, 2008), net and gross primary productivity (Raich and Schlesinger, 1992; Janssens et al., 2001), leaf area index (Reichstein et al., 2003) and photosynthetic activity (Kuzyakov and Cheng, 2004; Vargas et al., 2011a). For instance, a dense vegetation cover limits the amount of solar energy reaching the ground, impacting negatively soil temperature (Oliver et al., 1987), which could negatively influence soil respiration. In addition, topography and landscape attributes control the movement of water, energy and matter (MacMillan et al., 2000; McBratney et al., 2003), which directly affects soil water content, temperature and other properties (e.g., nutrient availability, texture), and thus, indirectly influences F_{CO2} (Brubaker et al., 1993; Hanson et al., 1993; Abnee et al., 2004; Tamai, 2010). For example, soil temperature is higher on sun-exposed slopes, as more energy reaches the ground (Prévost and Raymond, 2012), which could positively influence F_{CO2}. Finally, recent weather conditions such as air temperature changes and precipitations can affect F_{CO2} through changes in soil temperature and soil water content (Granier et al., 2007; Phillips et al., 2010). For example, rain can increase biological activity or decrease CO_2 diffusion in the soil and thus, increase or decrease F_{CO2} (Orchard and Cook, 1983; Linn and Doran, 1984; Skopp et al., 1990). Interactions between internal and external factors translate into temporal (i.e., seasonal, intra-seasonal and diel) and spatial patterns of F_{CO2}. As internal and external factors can both vary temporally and spatially, it can be challenging to isolate the specific contribution of each factor on F_{CO2} (Vargas et al., 2011b).

The availability of high-quality remote sensing data has increased the use of external factors (as opposed to internal factors) in soil CO_2 modelling. For example, Jian et al. (2022) used air temperature and monthly precipitation to model soil respiration at the site scale instead of using soil temperature and soil water content. Warner et al. (2019) used LiDAR-derived topographic attributes to model soil CO_2 and CH_4 fluxes at the landscape scale and Huang et al. (2015) relied on night-time land surface temperature to model annual soil respiration rates at the site scale. However, fractioning F_{CO2} into its different components (i.e., spatial and temporal variation patterns) remains challenging; studies usually tend to minimize the number of field F_{CO2} measurements used to model the seasonal, intra-seasonal, spatial and diel trends of F_{CO2} (Betson et al., 2007; Martin and Bolstad, 2009; Perez-Quezada et al., 2016).

Our main objective was to assess the seasonal, intra-seasonal and spatial variability of F_{CO2} using empirical quantitative relationships established between F_{CO2} measurements and spatially referenced environmental covariates. A 122-hectare balsam fir (*Abies balsamea* (L.) Mill.) – paper birch (*Betula papyrifera* Marshall) dominated forest located in boreal Quebec (Canada) was used as a case study, and we used a closed-dynamic-chamber system to measure F_{CO2} at a weekly interval. These relationships were applied to model the spatial and temporal variations of F_{CO2} over the study area. The effects on F_{CO2} of external factors related to recent weather, topography and vegetation were also studied. Our methodological approach is based on the decomposition of F_{CO2} into its temporal and spatial variation patterns and the use of a machine learning approach that allows to take into account the complex interactions between covariates. This approach allowed us to quantify the importance and isolate the specific effect of each variable on F_{CO2}, and to map the predictions of F_{CO2} in our study area. We hypothesized that seasonal variation in F_{CO2} does not explain a significant part of the local variation in F_{CO2}, while external factors related to recent weather, vegetation, soil water content and soil characteristics explain most of the spatial and intra-seasonal variations.

Table 1
Monthly means of daily mean, minimum and maximum air temperature and monthly total precipitation at Forêt Montmorency in 2020.

Month	Monthly means of daily mean air temperature (°C)	Monthly means of daily maximum air temperature (°C)	Monthly means of minimum daily air temperature (°C)	Monthly total precipitation (mm)
January	- 12.4	- 6.5	- 18.2	88.6
February	- 13.5	- 6.7	- 20.2	99.3
March	- 8.5	- 0.9	- 16.0	161.0
April	- 3.3	3.2	- 9.7	106.9
May	5.3	12.6	- 2.0	85.5
June	12.7	20.5	5.0	80.0
July	16.6	23.1	10.2	196.0
August	13.4	19.9	6.9	156.4
September	8.5	14.7	2.4	186.3
October	1.4	6.3	- 3.6	191.2
November	- 2.9	2.4	- 8.2	103.4
December	- 8.0	- 3.2	- 12.9	204.7

Data from Environment and Climate Change Canada (2021).

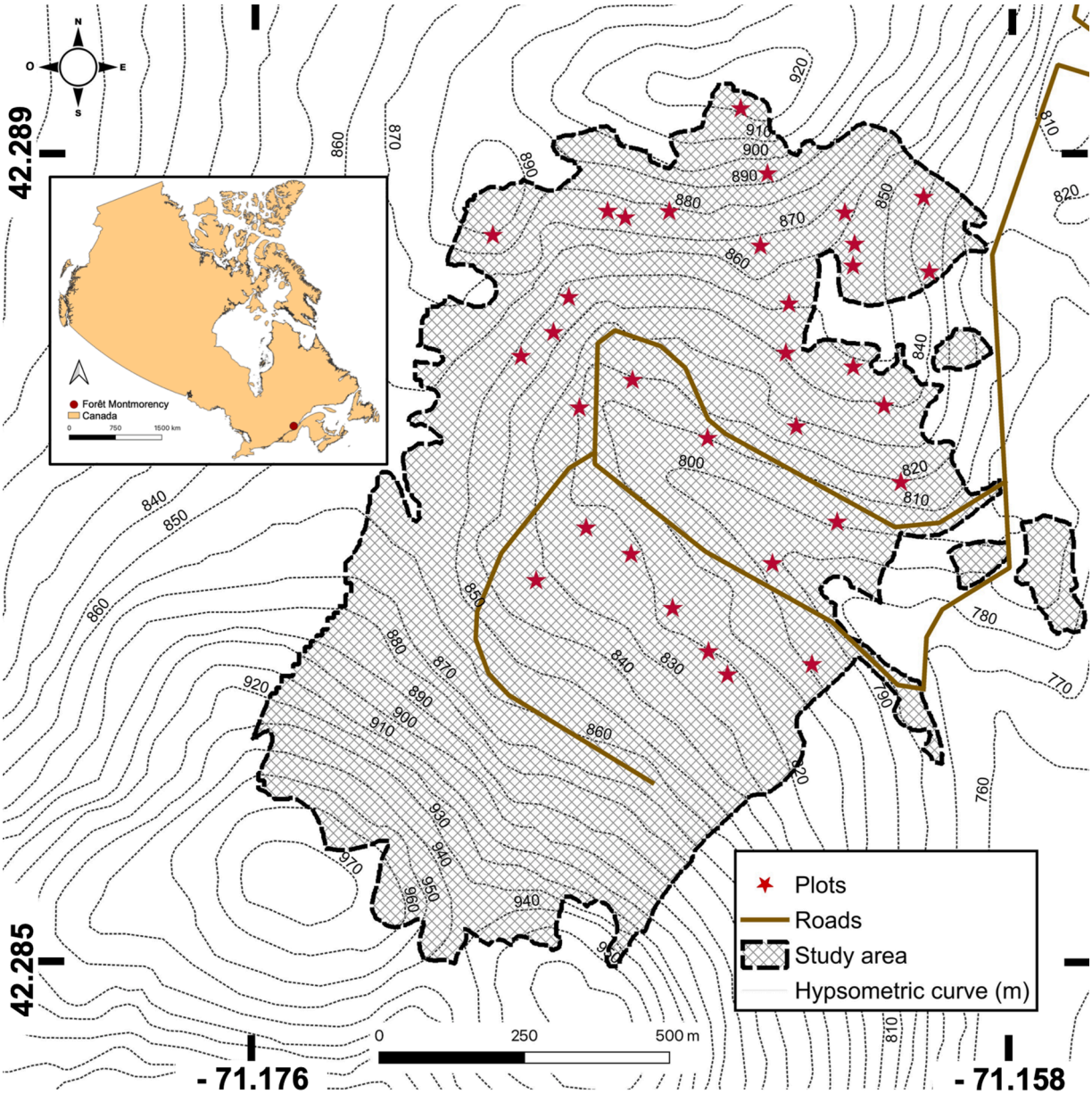


Fig. 1. Plots location within the "Bassin Expérimental du Ruisseau des Eaux-Volées" at Forêt Montmorency (Québec, Canada). Plots were distributed in forest stands that originated from a clearcut in 1993–1994 (shaded area), by considering topographic attributes and vegetation height derived from LiDAR datasets.

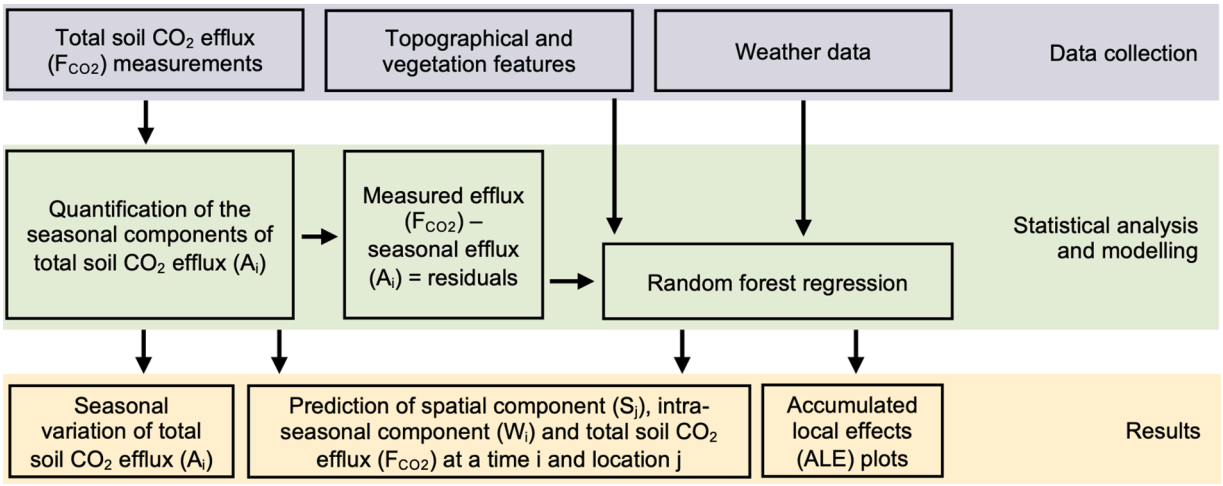


Fig. 2. Conceptual framework used to assess the seasonal, intra-seasonal and spatial variability of F_{CO2} in a balsam fir (*Abies balsamea* (L.) Mill.) – white birch (*Betula papyrifera* Marshall) dominated perhumid boreal watershed in Quebec (Canada).

2. Materials and methods

2.1. Study area

The study took place in a sub-watershed of the "Bassin Expérimental du Ruisseau des Eaux-Volées" (BEREV-7a, 47°17′27.1″ N; 71°09′55.9″ W) (Plamondon and Ouellet, 1980), located in Forêt Montmorency, an experimental forest managed by Université Laval (Quebec, Canada). Forêt Montmorency is part of the perhumid boreal region of eastern Canada (Clayden et al., 2011). It has a cold and wet climate with an mean annual temperature of 0.5 °C and annual precipitation of 1 583 mm, of which 620 mm falls as snow (1981–2010 mean from weather station #7,042,388) (Environment and Climate Change Canada, 2019). The climatic data for the site are presented in Table 1.

The study area (122 ha) (Fig. 1) consists of forest stands that originated from a clearcut with protection of advance regeneration and soils performed in 1993–1994 (27 years old in 2020) (Guillemette et al., 2005); trees were cut and delimbed at the stump, leaving branches and tops on the clearcut area. Forest stands in the study area are dominated by balsam fir (90% of the basal area). Other species include white spruce (*Picea glauca* (Moench) Voss), black spruce (*P. mariana* (Mill.) B.S.P.) and paper birch. Ground vegetation consists of mosses, including *Sphagnum* sp. in wetter areas. Information on tree height, number of stems per hectare and basal area for each plot is presented in the Supplementary Material section. The main soil types are ferro-humic and humo-ferric podzols (Soil Classification Working Group, 1998) associated with deep deposits of glacial origin dominated by loamy sands and sandy loams and a 5–15 cm mor humus on dry mineral soils. Wetter sites were characterized by a 5–40 cm organic layer mainly composed of *Sphagnum* sp. Soil organic carbon content in B horizons is relatively low, varying between 5 and 79 g/kg with a mean of 30.2 g/kg.

Within the study area, we distributed 33 plots according to a topographic gradient (mean [5^{th} – 95^{th} quantile]) based on altitude (m) (841.23 [798.40 – 890.19]), slope (%) (17 [8 – 24]) and vegetation height (m) (5.85 [2.60 – 9.42]) using conditioned Latin hypercube sampling; a sampling strategy which produces a sample design that respects the variance of covariates in the study area (Minasny and McBratney, 2006; Minasny et al., 2013). Each plot (20 m × 20 m) was divided into four 10 m × 10 m quadrants, which were further divided into four 5 m × 5 m sub-quadrants. Three quadrants were randomly selected in each plot; in each one, we randomly selected a sub-quadrant for F_{CO2} measurements using inserted cylinders. The cylinder was placed randomly within the sub-quadrant (with, on average, at least 1 metre distance from the nearest tree). Cylinders were made of sections of opaque white PVC pipes that were 10 cm in length and 10 cm in diameter. Cylinders were slightly (0–3 cm) inserted into the forest floor and secured to the ground using brackets to limit air exchange between the inside and ambient air during F_{CO2} measurements.

2.2. Data collection

Fig. 2 illustrates the main steps relative to data collection, statistical analyses and F_{CO2} modelling. Collected data are grouped into three categories: F_{CO2} measurements, recent weather data, and topographic and vegetation data.

2.2.1. Total soil CO_2 efflux measurements

We measured F_{CO2} with two LI-8100 portable infrared gas analyzers (IRGA) and two survey chambers (8100–102 10 cm) (LI-COR inc., Lincoln, Nebraska, USA). Both instruments were calibrated in the spring before the first measurements. F_{CO2} measurements were taken between mid-June and late October 2020, which typically corresponds to the period when the ground is snow-free at Forêt Montmorency. We performed F_{CO2} measurements on all 99 cylinders on a weekly basis in June, July and August, and every two weeks in September and October. Each week, F_{CO2} measurements were performed randomly (i.e., collars were measured randomly throughout the day). F_{CO2} measurements were made during the day, between 7:00 a.m. and 4:00 p.m. (with 95% of the measurements being taken between 8:00 a.m. and 3:00 p.m.). No F_{CO2} measurement was made during precipitation events. The observation time for each F_{CO2} measurement was 90 s with a dead band of 30 s following the closure of the survey chamber. Only one measurement cycle was performed per cylinder and per date. After each measurement, the instrument performed a 60-second post-purge sequence during which air was pumped into the system, thus preventing the accumulation of moisture in the instrument.

2.2.2. Recent weather

We used climate data from the Meteorological Service of Canada (Environment and Climate Change Canada, 2021) to assess recent weather features, based on a station located 3.5 km from the study area (ID 7,042,395; coordinates 47°19′22.000″ N; 71°08′54.000″ W; altitude 672.8 m). Three features were calculated to describe the effect of recent weather on the intra-seasonal variability on F_{CO2}: the mean air temperature during the two days (D_{-2} and D_{-1}) prior to the day of F_{CO2} measurements (D_0), the difference between the mean daily air temperature of D_{-1} and the mean daily air temperature of D_0, to account for short-term variations of air temperature that are not accounted for by

seasonal trends, and finally, the cumulative precipitation of the last 12 h prior to D_0 (see Table A1 for details). Several time steps were tested for each feature. We selected the 3 time steps that best improved model performance and that were described earlier.

2.2.3. Topography and vegetation

We used digital surface and elevation models (DSM and DEM) at a 1-m resolution that were derived from a 2016 aerial LiDAR survey of the study area (2.5 pt/m^2). The original DEM was first resampled to 15 m and filtered with multiple average filter windows to alleviate the effect of local noise in inducing spurious errors in topographical derivatives (MacMillan et al., 2000). The following topographic features were derived from the DEM using System for Automated Geoscientific Analyses (SAGA, 2.3.2) (Conrad et al., 2015): aspect, catchment area, catchment slope, convergence index, convexity, landforms, LS factor, mid-slope position, MRVBF, valley depth, slope height, slope length and the slope (Table A2). Vegetation height was calculated by subtracting the DEM from the DSM (1 to 1.5 m error) (Mielcarek et al., 2018) (see Table A2 and the Supplementary Material for details).

2.2.4. Feature selection

To reduce multicollinearity and enhance the interpretability of the model, we performed a step-by-step covariates (feature) selection using the variance inflation factor (VIF) calculated using the Python *statsmodels* (0.12.2) module (Seabold et al., 2010). At each step, we removed the feature with the highest VIF value. Then, we recalculated VIF for all remaining features. We repeated this operation until all features exhibited a VIF value below 10, as values between 5 and 10 are generally used to identify features that are moderately correlated with other features (Craney and Surles, 2002). VIF selection resulted in seven LiDAR-derived features for the model: 1) the vegetation height, 2) a multiresolution index of valley bottom flatness (MRVBF), 3) a topographical convergence index, 4) the slope height and 5) length, 6) the mid-slope position and 7) the aspect. The mean air temperature during D_{-2} and D_{-1}, the difference between the mean daily air temperature of D_{-1} and D_0, and the cumulative precipitation of the last 12 h prior to measurements were also selected for modelling (Appendix A).

2.3. Modelling total soil CO_2 effluxes

Time series of F_{CO2} can be broken down into three components that reflect the effects of variables related to 1) the seasonal trend (A_i), 2) daily/weekly patterns (W_i), and 3) spatial patterns (S_j) (Eq. (1)) (Appendix B).

$$F_{CO2_{ij}} = A_i + W_i + S_j + E_{ij} \quad (1)$$

Where $F_{CO2_{ij}}$ is the total soil CO_2 efflux at time i and location j, A_i is the total mean soil CO_2 efflux relative to the seasonal component at time i which makes it possible to disentangle the amount of variance that is associated with seasonal climatic conditions, W_i is the total soil CO_2 efflux relative to the intra-seasonal component (recent weather conditions) at time i, S_j is the total soil CO_2 efflux relative to the spatial component at location j (spatially explicit biotic and abiotic factors: topography, vegetation), and E_{ij} is the remaining components that could not be explained by the model at time i and at location j. The proposed approach can be considered as a specific case of seasonal trend decomposition for multiple time series, in which a spatial component is involved.

2.3.1. Seasonal trend

We removed outliers in F_{CO2} measurements on a per-plot basis using the moving average window (k = 18 F_{CO2} measurements, 3 days of measurements on both sides) function from *caTools* R package (1.18.0) (Tuszynski, 2020). Outliers were removed when they were outside the mean±2 standard deviations' interval (which represented about 0.7% of the measurements). We then used filtered values to simulate A_i for the entire study area using a double logistic function, which uses the day of the year (DOY) as input (Eq. (2)):

$$A_i = \min + (\max - \min) \times \left(\frac{1}{1 + \exp(-\mathrm{mS} \times (i - S))} + \frac{1}{1 + \exp(\mathrm{mA} \times (i - A))} - 1 \right) \quad (2)$$

Where A_i is the total soil CO_2 efflux (μmol CO_2 m^{-2} s^{-1}) related to the seasonal component for a specific DOY (i), min is the minimum predicted total soil CO_2 efflux, max is the maximum measured total soil CO_2 efflux in the dataset, S and A are times of inflection (DOY) during the early and late summer periods, respectively, and mS and mA are the rates of change at S and A, respectively.

We optimized the parameters of the double logistic curves with the Levenberg-Marquardt Nonlinear Least-Squares algorithm (Levenberg, 1944; Marquardt, 1963) using the *minpack.lm* R package (1.2.1) (Elzhov et al., 2016). The uncertainty in the parameter estimates was assessed using a bootstrap resampling approach (B = 1000). Mean and standard deviations of each parameter were retained to describe the seasonal dynamics of F_{CO2}; we further retained median values of all daily simulated effluxes to assess the seasonal trend (A_i).

2.3.2. Intra-seasonal and spatial components

We assessed the intra-seasonal and spatial component (R_{Fij}), which contains W_i and S_j, by subtracting A_i from F_{CO2ij} using Eq. (3):

$$R_{Fij} = F_{CO2ij} - A_i = W_i + S_j + E_{ij} \quad (3)$$

Where R_{Fij} is the calculated spatial and intra-seasonal component, F_{CO2ij} is the measured total soil CO_2 efflux at time i and location j, A_i is the total mean soil CO_2 efflux relative to the seasonal component at time i, W_i is the total soil CO_2 efflux relative to the intra-seasonal component at time i, S_j is the total soil CO_2 efflux relative to the spatial component at location j, and E_{ij} is the remaining components that could not be explained by the model at time i and at location j. All terms in Eq. (3) are in μmol CO_2 m^{-2} s^{-1}.

W_i and S_j were estimated using a *Random Forest Regression* model (Breiman, 2001) implemented in the Python *scikit-learn* library (0.24.1) (Pedregosa et al., 2011). This non-parametric approach is relatively robust to multicollinearity and allows the integration of categorical and numerical features, and the identification of features that contribute to the model's accuracy (Hastie et al., 2009).

Six hyperparameters were optimized to improve the accuracy of the *Random Forest Regression* model (Appendix C). We performed the optimization using a grid search algorithm based on the performance of the algorithm in *cross-validation* (Hastie et al., 2009). We first divided the complete dataset (n = 1331) in two sub-datasets: one for training (70% of the dataset) and one for testing (out-of-bag observations) (30% of the dataset) based on cylinder ID, i.e., observations from a given cylinder were either all used for training or all used for testing; they were never split between the datasets. This allows to alleviate potential overfitting during the training stage. We further divided the training dataset into two sub-datasets: calibration (70% of the training dataset) and validation (30% of the training dataset), once again considering cylinder ID as a grouping factor for distributing observations between datasets. We used the validation dataset for hyperparameter optimization and for assessing the expected error; the test dataset was used to assess the generalization error. During the hyperparameter optimization stage, the best of 500 random hyperparameter combinations (out of 11 000 combinations) was tested with a k-fold *cross-validation* (n = 5). The top-5 of all combinations of hyperparameters were retained based on R^2 values (validation dataset). This operation led to a total of 125 models resulting from the combinations of 5 sets of hyperparameters, 5 training datasets and 5 test datasets. These combinations made it possible to study the

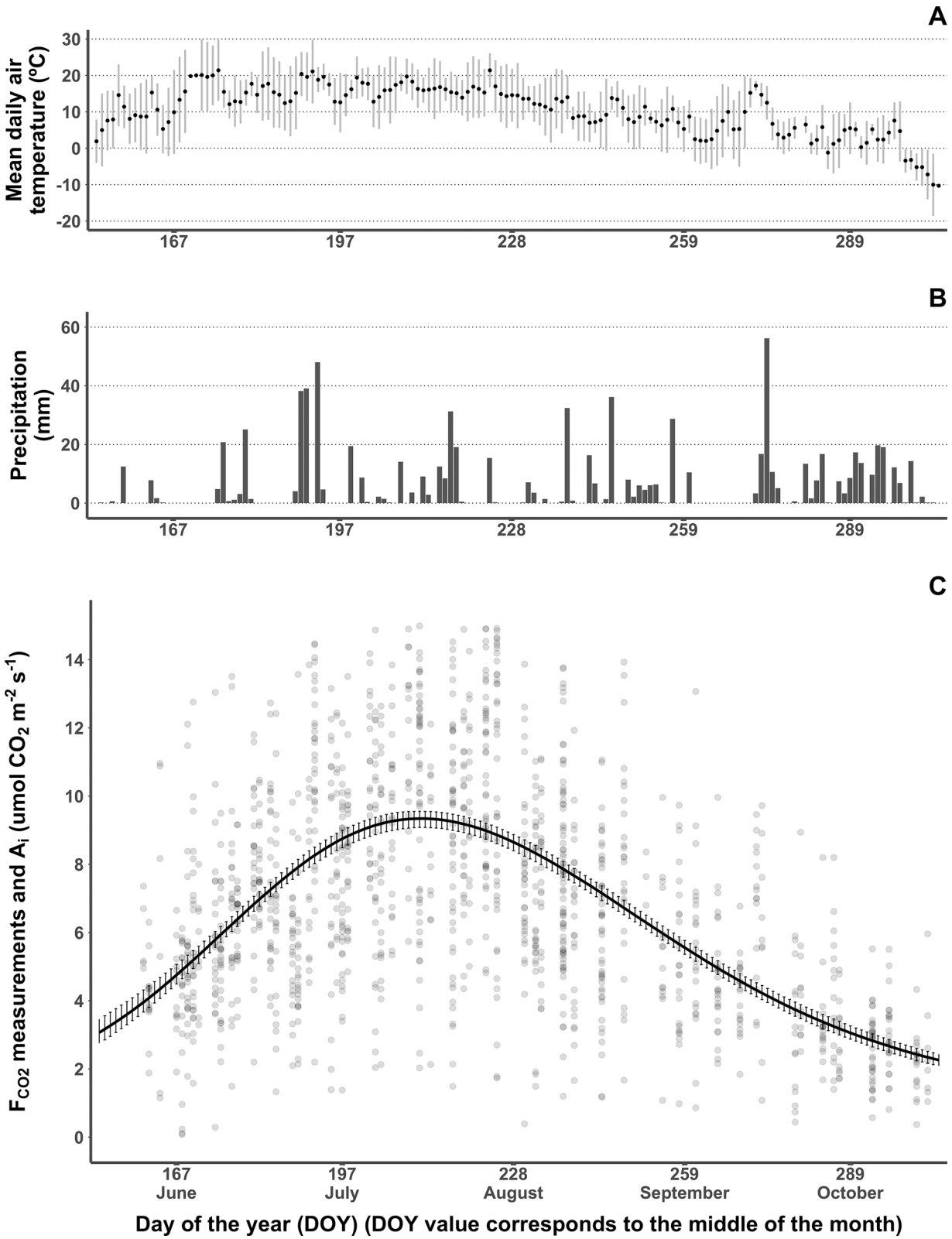


Fig. 3. Seasonal variations of daily mean (dots) and min/max (vertical bars) air temperature (A), total daily precipitation (B), total soil CO_2 efflux (F_{CO2}) measurements (dots) and median prediction of the seasonal trend (line) (A_i) with the related uncertainty (error bars) resulting from a bootstrap resampling approach ($B = 1000$) (for each day, 90% of the 1000 predictions of daily median F_{CO2} are within the error bar) (C) for the study area between June and October 2020.

Table 2
Parameters of the double logistic function used to characterize the seasonal trend (A_i) in total soil CO_2 efflux (F_{CO2}), with and without bootstrap (mean [5th percentile – 95th percentile]).

data	n	min	max	S	mS	A	mA
all dataset	1 431	1.04	15.88	180	0.05	238	0.04
with bootstrap	1 431	1.04	15.86	181	0.05	239	0.04
		[1.04 – 1.4]	[15.85 – 15.88]	[178 – 183]	[0.05 – 0.06]	[235 – 241]	[0.03 – 0.04]

Note: All total soil CO_2 efflux (F_{CO2}) measurements were used in "all dataset". "with bootstrap" refers to the bootstrap resampling approach used to quantify the uncertainty related to the time interval between the F_{CO2} measurements. n is the number of F_{CO2} measurements used to estimate the parameters: min and max are the minimum predicted and the maximum measured values of F_{CO2}, respectively, S and A are times of inflection points on the double logistic curve (DOY), and mS and mA are the rates of change at times S and A, respectively.

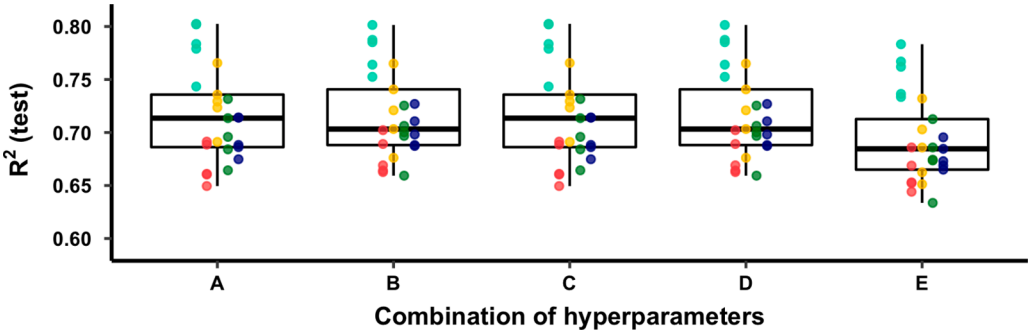


Fig. 4. Predictive performance of the 125 models grouped by the 5 best combinations of hyperparameters tested (boxplots) and the 5 test datasets (colour of the dots). Each boxplot thus represents 25 models. R^2 (test) (Y-axis) is the R^2 values on the test dataset. The combinations of hyperparameters used to fit these models are presented in Appendix C. For the boxplots, 50% of the values fall within the box (interquartile range (IQR), 25th – 75th quantile) with the thick horizontal line representing the median. The vertical line (on each edge of the box) represents the range of values that fall in the interval between −1.5 times the 25th quantile (q_{25}) and the 25th quantile (q_{25} – 1.5 IQR) or in the interval between the 75th quantile (q_{75}) and 1.5 times the 75th quantile (q_{75} + 1.5 IQR). Black dots represent the outliers (> 1.5 times and <3 times the interquartile range, in either direction from the box). coloured dots are used to identify each sub-dataset used in test.

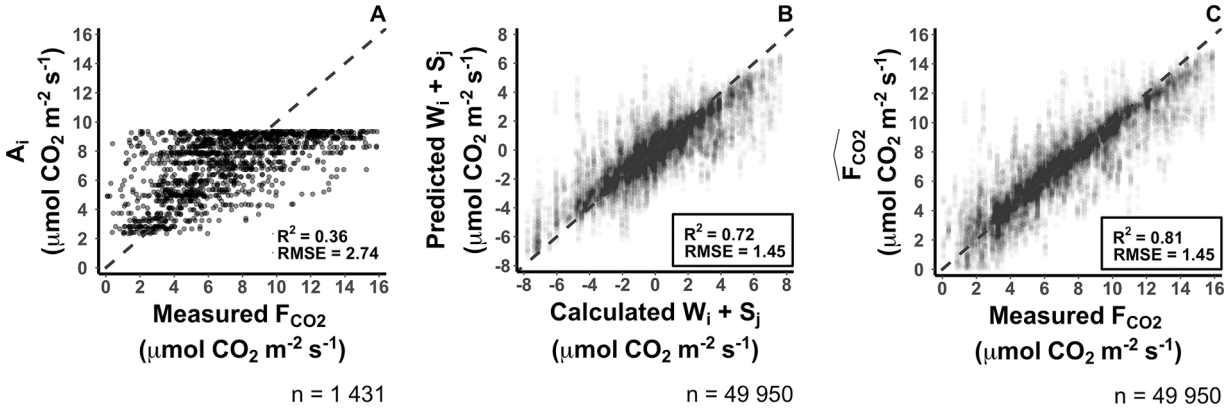


Fig. 5. Proportion of the variation (R^2 and RMSE) in total soil CO_2 efflux (F_{CO2}) measurements explained by the seasonal component of F_{CO2} (A_i) (A), proportion of the variation in calculated $W_i + S_j$ explained by the predicted $W_i + S_j$ (B), proportion of the variation in F_{CO2} explained by the predicted total soil CO_2 efflux ($\widehat{F_{CO2}}$) (C). For B and C, all 125 test datasets were used (n = 49 950). The dotted line represents the 1:1 line.

effect of sampling (training and testing datasets) and hyperparameter combinations on model performance. We used two indicators to evaluate the performance of the model: the coefficient of determination (R^2) was used as an indicator of the standardized variance that exists between measured and simulated values, and the root mean squared error (RMSE) was used to quantify the magnitude of the errors.

2.3.3. *Mapping total soil CO_2 effluxes*

We mapped predicted total soil CO_2 efflux ($\widehat{F_{CO2}}$) daily at a 10 m spatial resolution by combining the seasonal trend (A_i) with the intra-seasonal (W_i) and spatial (S_j) components (Eq. (4)). We divided the study area into 10 m × 10 m cells for which we derived topographical attributes and vegetation height. We then calculated the three recent weather features for each day between 1 June and 31 October 2020. For each pixel, $W_i + S_j$ were assessed using the median value from the 125 models. We also calculated the uncertainty associated with the prediction of $W_i + S_j$ by subtracting, for each pixel, the 5th percentile value from the 95th percentile value of the 125 models' predictions.

$$\widehat{F_{CO2}} = A_i + W_i + S_j \quad (4)$$

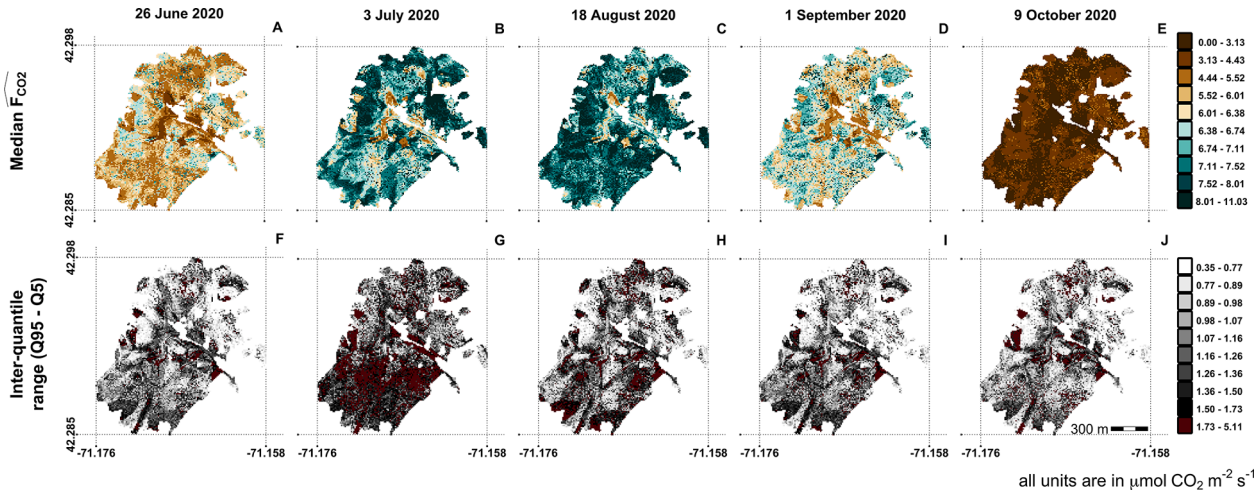


Fig. 6. Median prediction of total soil CO_2 effluxes ($\widehat{F_{CO2}}$) (daytime) for randomly selected days of June (A), July (B), August (C), September (D) and October (E) and corresponding inter-quantile ranges associated with the predictions (F-J) (calculated by subtracting the 5th percentile from the 95th percentile) for the study area. Recent weather features values used for the prediction of $W_i + S_j$ are presented as Supplementary Material.

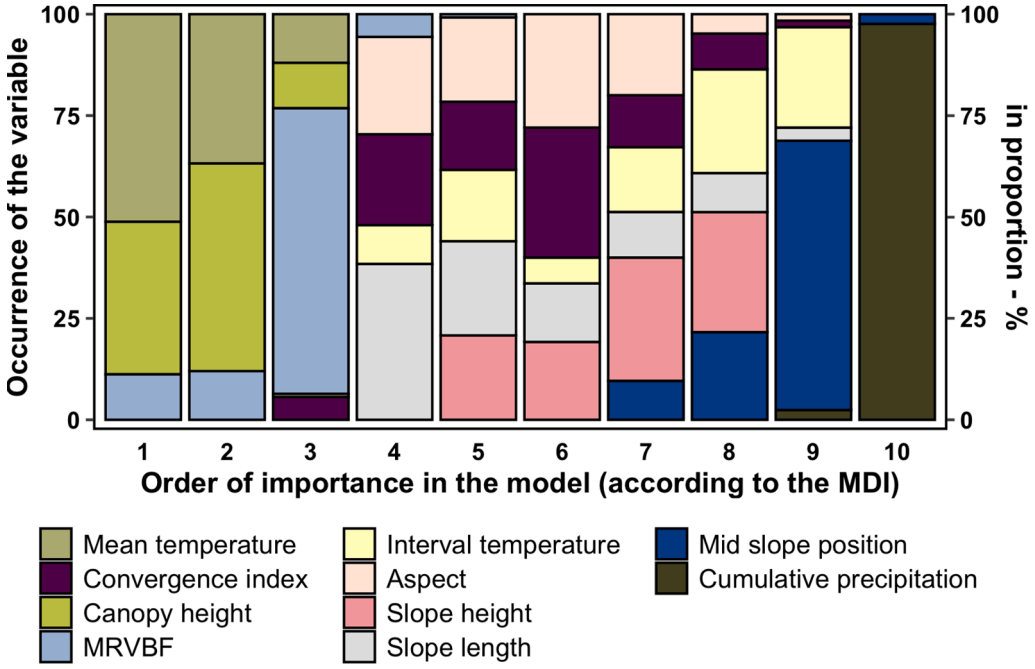


Fig. 7. Proportion (%) of occurrence of each variable by order of importance for the 125 models. The X-axis represents the position of importance of each variable according to the mean decrease in impurity (MDI). The variables in position 1 are the most important and those in position 10 are the least important. Each bar represents the proportion of variable that occurs in that position. MRVBF: Multiresolution Valley Bottom Flatness Index.

2.4. *External factors and total soil CO_2 effluxes*

We evaluated the importance of each variable from the *Random Forest Regression* models using the impurity-based feature importance (mean decrease in impurity, MDI) in *scikit-learn* library (Pedregosa et al., 2011). MDI varies between 0 and 1, with values closer to 1 associated with more important features in the model.

We produced main effect accumulated local effects plots (ALE plots) to isolate the individual effect of each variable on the prediction of the intra-seasonal and spatial components. ALE plots describe how variables influence the mean prediction of a machine learning model while keeping others fixed, even if the predictors are correlated (Molnar, 2019; Apley and Zhu, 2020). This is achieved by plotting, for each input variable, the differences in model predictions as a function of intervals (called "bins") of input variable values. The use of differences avoids making mean predictions from artificial variable values that are unlikely to be observed, especially when variables are correlated (Molnar, 2019). We produced ALE plots with the *alepython* module (0.1) (Jumelle et al.,

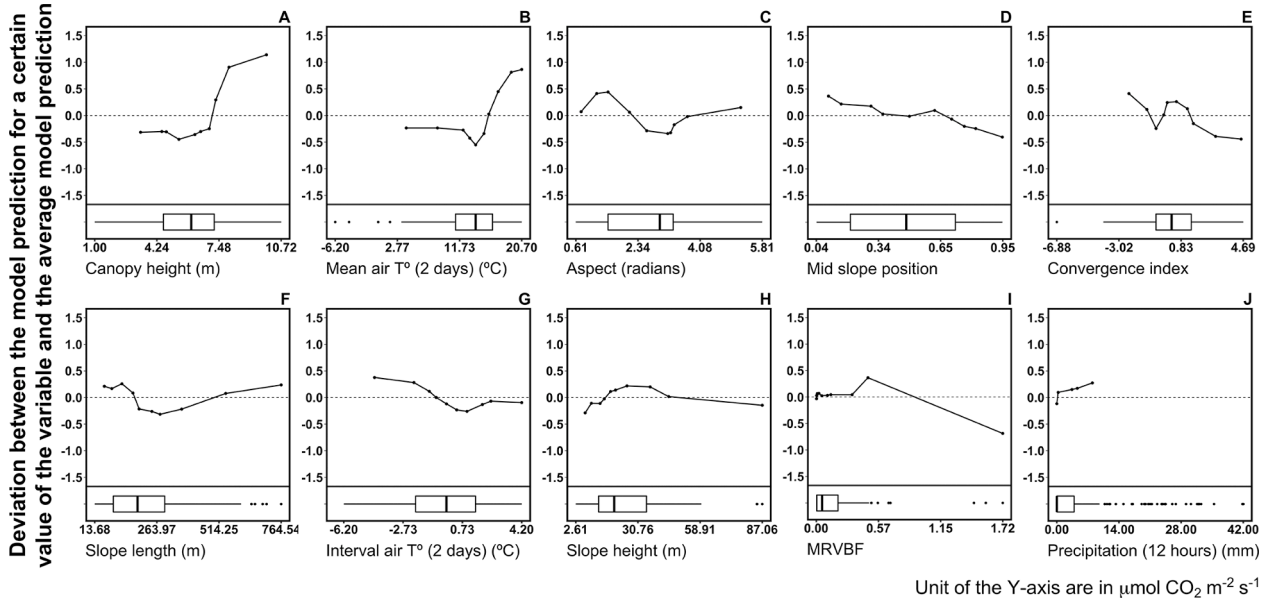


Fig. 8. Accumulated local effects (ALE) plots for the 10 variables. For a single variable, the Y-axis of the ALE plot represents the deviation between the model prediction for a given range of values (i.e., bin) of the variable and the mean model prediction (with the Y-axis = 0 representing the mean prediction of the model). The model with the test dataset R^2 value that was closest to the median R^2 value (test dataset) of 125 models was used to build the ALE plots. See Fig. 4 for the meaning of the boxplots. The values of the different graphs are presented in the Supplementary Material section.

2020) using the model whose R^2 on the test dataset was equal to the median of the R^2 values of the 125 models.

We performed all modelling and statistical analyses in R (4.0.3) (R Core Team, 2020) and Python (3.7.10) (Python Core Team, 2015) programming languages. Graphs and maps were produced using the *ggplot2* package in R (Hadley et al., 2020).

3. Results

3.1. Seasonal component

Fig. 3 illustrates the intra-annual variation of daily mean air temperature (3A), total daily precipitation (3B) and total soil CO_2 efflux (F_{CO2}) measurements and median prediction of the seasonal trend (A_i, 3C) at the scale of the study area and the uncertainty associated with this prediction. The mean [5^{th} – 95^{th} quantile] measured F_{CO2} was 7.04 [2.01 – 13.30] μmol CO_2 m^{-2} s^{-1} (n = 1431) and the predicted A_i was 6.13 [2.60 – 9.30] μmol CO_2 m^{-2} s^{-1} between June and October 2020 (n = 153 days) across the study area. Measured F_{CO2}, and its variance, were higher during July and August – the months with the highest monthly average air temperature of 16.6 °C and 13.4 °C, respectively. Conversely, lower temperatures were observed in June, September and October, with monthly average temperature of 12.7 °C, 8.5 °C and 1.4 °C, respectively (Fig. 3C). The variance of measured F_{CO2} showed a strong relationship with daily mean air temperature variation (Fig. 3A). June's cumulative precipitation was much lower (80.0 mm) than the other months (196.0 mm, 156.4 mm, 186.3 mm, and 191.2 mm for July, August, September and October, respectively) (see Appendix A for more details regarding climate data).

The parameters of the dual logistic equation used to model A_i, which can be interpreted in order to study the seasonal trend, are presented in Table 2. The absolute value of spring rate of change in A_i (mS) was higher than the autumn rate of change (mA) (Table 2), suggesting that F_{CO2} increases slightly more rapidly in the spring and gradually decreases in the fall. A_i rose from 3.97 μmol CO_2 m^{-2} s^{-1} (June 9, first day of measurement) to 9.35 μmol CO_2 m^{-2} s^{-1} (July 29) and then fell back to 2.34 μmol CO_2 m^{-2} s^{-1} (October 29, last day of measurement). The low standard deviation (SD) (Table 2) and error bar values (Fig. 3C) indicate a low uncertainty associated with the estimation of these parameters and thus, on the estimation of A_i.

3.2. Hyperparameter combinations used for modelling

No sub-daily pattern or seasonal variation was observed in the predicted residuals ($W_i + S_j$). Fig. 4 shows the predictive performance (R^2) of the 125 *Random Forest Regression* models used to predict $W_i + S_j$.

Both the combination of hyperparameters and the training-testing division had an influence on the performance of the model. For example, the combinations A, C, and D or divisions blue and yellow consistently produced higher R^2 values on the testing datasets (Fig. 4). For the same combination of hyperparameters and training-testing division, there was little variability in the models (n = 5) (Fig. 4). Therefore, we used 125 models with different hyperparameter combinations and training-testing divisions to consider these sources of variability. In addition, using 125 models instead of just one allowed us to not randomly choose a model or pick the best or worst model in terms of prediction performance.

3.3. Partitioning of total soil CO_2 efflux variance

Predicted A_i explained 36% of the variation in F_{CO2} measurements (R^2 = 0.36, RMSE = 2.74, Fig. 5A). The increasing variance of the residuals with F_{CO2} measurements (Fig. 5A) provides an indication of the strong dependency of F_{CO2} to recent weather (Wi) and spatial component (Sj). This dependency is also supported by the predicted $W_i + S_j$, which explained 72% of the calculated $W_i + S_j$ (R^2 = 0.72, RMSE = 1.45, Fig. 5B). The combination of the predictions of A_i, W_i and S_j allowed to explain 81% of F_{CO2} measurements (R^2 = 0.81, RMSE = 1.45, Fig. 5C).

By subtracting the variance of A_i from the variance of the global model, we observed that W_i and S_j explained almost half of the variance in F_{CO2} measurements (45%).

3.4. Mapping total soil CO_2 efflux

The pattern of spatial variation in $\widehat{F_{CO2}}$ was consistent from month to month (Fig. 6, A-E). The uncertainty in $\widehat{F_{CO2}}$ was rather low (< 1.00 μmol CO_2 m^{-2} s^{-1}), although some portions of the study area had a higher uncertainty (> 1.40 μmol CO_2 m^{-2} s^{-1}) (Fig. 6, F-J).

3.5. External factors and total soil CO_2 effluxes

Fig. 7 shows the importance of each feature amongst the 125 models. According to the impurity importance metric (i.e., mean decrease in impurity), mean air temperature of the last two days and canopy height were the two variables that most influenced the estimation of R_F (Fig. 7). They were followed by variables related to topography (in descending order of importance: valley bottom flatness index (MRVBF), convergence index, slope length, aspect, slope height, mid-slope position). Finally, the cumulative precipitation of the last 12 h was consistently the least important variable of the model (Fig. 7).

Of the 10 variables, canopy height had the most influence on $\widehat{F_{CO2}}$. When canopy height increased from 5.39 m to 9.95 m, $\widehat{F_{CO2}}$ increased on average by 1.58 μmol CO_2 m^{-2} s^{-1} (Fig. 8A). Mean air temperature of the last two days was the second most important variable affecting $\widehat{F_{CO2}}$. Below 14 °C, air temperature had no effect on $\widehat{F_{CO2}}$, but above this threshold, an increase in air temperature from 14.05 °C to 20.70 °C led to an average increase in $\widehat{F_{CO2}}$ of 1.41 μmol CO_2 m^{-2} s^{-1} (Fig. 8B). Northeast facing slopes have on average higher $\widehat{F_{CO2}}$ than south facing slopes (0.77 and 0.49 μmol CO_2 m^{-2} s^{-1} respectively) (Fig. 8C). Moving away from the middle of the slope (down or up the slope) favors lower $\widehat{F_{CO2}}$ on average (−0.77 μmol CO_2 m^{-2} s^{-1}) (Fig. 8D). Diverging slopes and converging slopes (Fig. 8E) favoured lower $\widehat{F_{CO2}}$ compared to regular slopes (−0.69 and −0.50 μmol CO_2 m^{-2} s^{-1} respectively). Between 52 m and 277 m, increasing slope length generally decreased $\widehat{F_{CO2}}$ (−0.52 μmol CO_2 m^{-2} s^{-1}), while it increased slightly between 277 m and 765 m (0.52 μmol CO_2 m^{-2} s^{-1}) (Fig. 8F). Interval air temperature, slope height, MRVBF (between 0.00 and 0.31, indicating it is not located in a valley bottom) and total precipitation over the past 12 h all had a small or negligible effect on $\widehat{F_{CO2}}$ (Fig. 8G, H, I and J).

4. Discussion

Our study simulated the spatial and temporal behaviour of F_{CO2} over a perhumid boreal watershed dominated by balsam fir forest stands. We used external factors related to topography, vegetation height and recent weather to account for spatial and intra-seasonal variability in F_{CO2}. The high sampling intensity allowed us to measure the importance of spatial and temporal heterogeneity in F_{CO2} at landscape-scale.

In our study, the mean measured F_{CO2} was just below 7 μmol CO_2 m^{-2} s^{-1}; some rare F_{CO2} measurements reached 16 μmol CO_2 m^{-2} s^{-1}. Overall, these values may seem high compared to other studies conducted in boreal ecosystems. For example, Laganière et al. (2012) and Khomik et al. (2006) measured total soil respiration rates of up to 7 μmol CO_2 m^{-2} s^{-1} in boreal stands using a closed and dynamic chamber method, with cylinders inserted at 5 cm depth. However, Gaumont-Guay et al. (2006) (using a closed and dynamic chamber method, cylinders inserted at 3–4 cm depth) and Rayment and Jarvis (2000) (using an open chamber system, cylinders inserted at 2–7 cm depth) reported rates of up to 9.2 and 14 μmol CO_2 m^{-2} s^{-1}, respectively. This could be explained by the young age of the stands in this study (here, 33 years since clearcutting). For example, Payeur-Poirier et al. (2012) observed that juvenile boreal stands (33 years old) can have higher rates of soil respiration than pre-harvest (105 years old) or recently harvested (8 years old) stands while still being moderate to strong carbon sinks. Younger stands (10–15 years) have higher rates of heterotrophic soil respiration (related to greater availability of organic carbon at the soil surface) and autotrophic soil respiration (related to greater root biomass) than older stands (31–47 years) (Saiz et al., 2006). Microclimatic conditions vary with stand age, with higher air temperatures and more daily air temperature variation in young stands (Lindenmayer et al., 2022), which influences soil conditions and the soil respiration rate. Stand age also regulates the response of autotrophic and heterotrophic soil respiration to changes in temperature, with a higher sensitivity (Q_{10}) in mature sites (45 years) compared to young sites (10 and 20 years) (Ma et al., 2014).

4.1. Seasonal trend

We used a double logistic curve to estimate A_i as a function of DOY. This curve allows to decompose F_{CO2} into four stages on a yearly basis: minimum values in winter, a gradual increase in spring, a plateau in summer and a gradual decrease in autumn. This corresponds to the pattern of annual variation in soil respiration that has been reported in other studies in boreal (e.g., Morén and Lindroth, 2000; Rayment and Jarvis, 2000; Khomik et al., 2006) and temperate forests (e.g., Curiel Yuste et al., 2005). Consequently, parameters of double logistic curves could be easily used to compare F_{CO2} between years or studies.

The relationship between air temperature and variance of F_{CO2} (Fig. 3) suggests that spring and fall F_{CO2} are largely dependant on available energy and the effect of recent weather and local conditions is rather limited. Previous studies have shown that soil temperature and water content were the main variables responsible for the temporal variation of F_{CO2} (Rayment and Jarvis, 2000; Subke et al., 2003; Gaumont-Guay et al., 2006; Khomik et al., 2006). According to A_i, F_{CO2} increased rapidly during spring but decreased more slowly during late summer and fall. This phenomenon was observed in deciduous stands: for the same soil temperature in spring and autumn, the soil respiration rate was slightly higher in autumn due to the addition of fresh organic matter, which coincides with a decrease in the vegetation leaf area index due to shedding (Curiel Yuste et al., 2005). Another possibility is the seasonal variation in the allocation of photosynthetic products in the tree (Davidson and Holbrook, 2009). After mid-August, photosynthates are no longer distributed to the aerial parts of the trees; they are rather allocated to roots, where they support growth and the development of ectomycorrhizal fungi (Fortin and Lamhamedi, 2009), which could result in increased rhizosphere respiration. Lavigne et al. (2004) observed higher root respiration rates in early fall compared to late spring in a similar forest ecosystem.

4.2. Spatial and intra-seasonal variation

Our study demonstrated the usefulness of external factors to simulate spatial and intra-seasonal components of F_{CO2}. F_{CO2} is linked to external factors through their influence on soil conditions (i.e., internal factors). Therefore, the use of external factors makes it possible to avoid using internal factors (e.g., soil temperature, soil water content, soil organic matter quantity and quality and microbial and root activity) for modelling purposes. The use of *Random Forest Regression* model and ALE

plots allowed us to quantify the individual effects of biotic and abiotic factors on F_{CO2}. Further research is needed to establish how soil or air temperature ALE-plots compare to traditional Q_{10} approaches. In our study, the air temperature ALE-plot provided an efficient way to assess the effect of temperature on F_{CO2} with respect to other external factors. $\widehat{F_{CO2}}$ showed higher values and variance during summer, which supports that the effects of recent weather and local abiotic factors on F_{CO2} are at a maximum during summer and less important during other periods.

Vegetation can influence soil respiration rates in multiple ways through root respiration, which can account for a significant portion of total soil respiration in forest ecosystems (Ekblad and Högberg, 2001; Högberg et al., 2001, 2009) and respiration by heterotrophic organisms through litterfall (Davidson et al., 2002a; Sayer, 2006), dead organic matter and root exudates. In our study area, our model predicts higher F_{CO2} in areas with taller trees and vice versa. Since tree height can be used to characterize site productivity (Bélanger et al., 1995), we hypothesize that F_{CO2} are higher in the more productive sites in our study area. Tree height could be correlated to soil respiration through shading and its effect on microclimate (e.g., Greiser et al., 2018), as well as through photosynthetic activity (e.g., Högberg et al., 2001) and its effect on root activity (root respiration and rhizodeposition). Several studies have looked at the correlation between soil respiration and above-ground productivity (e.g., Raich, 1998; Reichstein et al., 2003); future studies should look at the specific effect of tree height on both compartments of soil respiration i.e., respiration by autotrophic or heterotrophic organisms.

The valley bottom flatness index (MRVBF) was an important external factor in the model (Fig. 7). In our study, the valley bottom areas corresponded to areas of poor drainage, covered with sphagnum moss and with small (5–7 m), densely packed fir trees. According to the mapping of the model's F_{CO2} predictions, low F_{CO2} are predicted in valley bottoms. However, on its own, MRVBF explained little variation in the prediction of $W_i + S_j$ (Fig. 7). Thus, taken individually, MRVBF was likely not responsible for the low F_{CO2} values in these areas; it was rather a combined effect of the different topography-specific factors. In addition, the fact that only a small proportion of the study area was composed of valley bottoms (2 plots out of 33 for which MRVBF > 1.5, i.e., 90 of the 1 428 F_{CO2} measurements) could distort the use of the mean decrease in impurity (MDI) to quantify the importance of MRVBF in the model. Either way, the fact that F_{CO2} was lower in valley bottom areas is consistent with other empirical studies (Davidson et al., 1998; Phillips et al., 2010). High soil water content can reduce the rate of soil respiration by reducing soil aeration (Linn and Doran, 1984) or limiting CO_2 diffusion (Skopp et al., 1990).

Slope orientation and convergence index are also important factors according to the model. The model predicted higher average F_{CO2} on northeast-facing slopes than on south-facing slopes, which is consistent with the empirical measurements made by Abnee et al. (2004). Abnee et al. (2004) attributed these differences to wetter and more fertile soil conditions (higher pH, higher total N, and higher concentration of plant-extractable nutrients) favoring net primary productivity and microbial and root activity. According to the model, F_{CO2} was higher on regular slopes than on divergent slopes, a result consistent with those of Warner et al. (2018), who called these regions "transition zones". Previous studies have shown that transition zones can be hotspots of soil respiration because of their optimal soil water content conditions and large amount of labile carbon pool (Webster et al., 2008).

Air temperature of the last two days before the F_{CO2} measurement was the most important factor in the intra-seasonal trend. Soil temperature is the most important factor influencing soil respiration (Lloyd and Taylor, 1994) and air temperature is a key factor in predicting soil temperature under forest cover, although other factors such as leaf area index, soil depth and litter mass are also involved (Kang et al., 2000; Paul et al., 2004). Our results do not show a clear relationship between F_{CO2} and a 2-day interval of air temperature. This could be because temperature changes generally follow a seasonal pattern, so this effect on F_{CO2} was removed when A_i was subtracted from the F_{CO2} measurements. Moreover, the relationship between air and soil temperatures is complex and non-linear (Jungqvist et al., 2014; Lembrechts et al., 2022). However, rapid changes in soil temperature can alter CO_2 production in the soil (Vargas et al., 2010). The use of air temperature seems promising and has, for example, already been used to model soil respiration on a regional scale (Jian et al., 2022) or to calculate Q_{10} (i.e., the rate of change of soil respiration for a temperature increase of 10 °C) (Bond-Lamberty and Thomson, 2010; Wang et al., 2010). Compared to soil temperature, air temperature data have the advantage of being available on a very large scale and with a very high accuracy (e.g., Oyler et al., 2015; Hooker et al., 2018; Osborn et al., 2021).

Cumulative precipitation of the last 12 h had little or no effect on predictions of F_{CO2}. Generally, precipitation influences the water content of the soil, causing a short and sudden increase in the soil respiration rate (Orchard and Cook, 1983; Savage and Davidson, 2003), which is particularly important in ecosystems with a drier climate (Kelliher et al., 2004). Given the humid climate of our study region, water is likely not a limiting factor for soil microbial processes. Jian et al. (2022) showed that using monthly precipitation instead of soil moisture content was not effective at the regional scale and that other factors, such as soil texture, should be considered. We also assume that the sampling strategy for soil respiration measurements could influence the correlation between soil respiration and precipitation. In our case, there were no soil respiration measurements during precipitation events (although, it could have rained in the 12 h that preceded F_{CO2} measurements).

4.3. Partitioning total soil CO_2 effluxes

Our methodology makes it possible to partition F_{CO2} into its three components of variation (i.e., seasonal, intra-seasonal and spatial) based on punctual and repeated F_{CO2} measurements. amongst the different methods available to measure F_{CO2} (Norman et al., 1997; Janssens et al., 2000), closed-dynamic-chamber systems allow for a higher spatial resolution in the measurement of F_{CO2} (Savage and Davidson, 2003). Our results demonstrate the importance of intra-seasonal variation (effect of recent weather) and spatial variation (topography and vegetation height) on F_{CO2}, two sources of variation that may be difficult to model from small F_{CO2} datasets.

Several studies have investigated the minimum number of measurements of F_{CO2} required (per day, per year or per site) to properly characterize the daily and annual (e.g., Betson et al., 2007; Wang et al., 2010; Perez-Quezada et al., 2016) and spatial variability (e.g., Davidson et al., 2002b; Herbst et al., 2009) of F_{CO2}. However, little mention is made of the intra-seasonal variation in F_{CO2}, a normal approach when the objective is to quantify annual rates of soil respiration. In any case, when modelling is based on an empirical approach (i.e., punctual and repeated measurements of soil CO_2 flux), adequate characterization of spatial and intra-seasonal variations in soil respiration depends on the balance between the spatial and temporal distribution of soil CO_2 flux measurements (Savage and Davidson, 2003). We hypothesize that when F_{CO2} is partitioned into its different patterns of variation, the proportion of each source of variation could be influenced by the temporal and spatial intensity of soil respiration measurements. In this study, we used topographic attributes and a canopy height model in combination with conditional Latin hypercube sampling for sampling design (Minasny

et al., 2013). Therefore, the selection and number of sites to be measured is important. More research is needed to test this hypothesis and its impact on soil respiration modelling. Soil respiration varies as much on a very fine scale (<1 m) (± 1.28 μmol CO_2 m^{-2} s^{-1}) than on a small scale (<10 m), and almost as much on a medium scale (>100 m) (Martin and Bolstad, 2009). While weekly or biweekly measurements are sufficient to estimate total annual soil respiration (Savage et al., 2008), determining the minimum number of soil respiration measurements needed to properly quantify the intra-seasonal and spatial trend for a specific area requires further research.

Conclusion

In this study, we proposed a new methodology to model total soil CO_2 effluxes (F_{CO2}) at the landscape scale using external factors derived from remote sensing and climate data. Our approach is based on the partitioning of F_{CO2} into three sources of variation i.e., seasonal, intra-seasonal and spatial. The *Random Forest Regression* model, using 10 external factors related to recent weather, topography and vegetation height, explained 81% of the variation in F_{CO2} measurements. While spatial and intra-seasonal variation explained 45% of the variation in F_{CO2} measurements, seasonal variation only explained 36%. In our study area, F_{CO2} was highest in stands with tall trees or when air temperature was high. Although there were only a few valley bottoms characterized by dense and small stands of balsam fir on sub-hydric drainage over our study area, these sites had lower F_{CO2} than more productive, moderately drained areas with higher tree heights.

For the quantification of A_i, the use of the day of the year (DOY) as a predictor allows for low uncertainty, but it also means that the predicted A_i is specific to current climatic conditions and the area covered by our study area, which was relatively small (122 ha). Further research is needed to predict A_i as a function of external factors that may be available at high temporal resolution, such as air temperature, rainfall or LAI, e.g., from Earth observation satellites. Nevertheless, our research highlights the high spatial and temporal variability of F_{CO2} in the area. F_{CO2} exhibits a high level of variability amongst and within sites, recalling the importance of increasing our understanding of the impact of external factors on F_{CO2}. We used ALE plots to understand and quantify the impact of external factors on F_{CO2}, but this tool could also be used to better understand their effect on different soil respiration processes or to predict their response to changes in biotic or abiotic variables.

F_{CO2} measurement campaigns should ensure that both intra-seasonal and spatial variations are considered, as they are as important as seasonal variation. This will allow F_{CO2} to be fully modelled from external factors without having to consider very small-scale soil conditions (i.e., temperature, moisture content, soil organic matter, soil microbial community and root activity).

Authors' contributions

J.D. Sylvain, G. Drolet, and N. Thiffault designed the study and acquired the funding. E. Thiffault also contributed to the funding related to field work. J.D. Sylvain conceptualized the study and designed the methodology. A. Harel carried out the soil respiration measurements, data curation, statistical analyses and visualization. A. Harel wrote a first draft of the manuscript. All authors reviewed and edited the manuscript.

Declaration of Competing Interest

The authors declare no competing interests.

Data Availability

The data underlying this article are available at https://doi.org/10.6084/m9.figshare.21865383.v1.

Acknowledgments

The authors thank L. Papillon, F. Wiseman, J. Monffette, S. Marcouiller, M. Rhéaume, P-L Déchêne, M. Bélec, M. Prudhommeaux, J.G Yee Paré and members of the Team Carbone for technical support. We also thank the anonymous reviewers and the Editor, whose comments greatly improved the quality of this manuscript. The authors also thank Fairmont Le Château Frontenac (3568 G) and J. André Fortin (3395 G) for the merit scholarships (Université Laval) awarded to Antoine Harel during his graduate studies.

Funding

This work was conducted and funded as part of the Evap-for project of Ministère des Ressources naturelles et des Forêts (Quebec, Canada) (project #142332139). Part of the field work was paid by E. Thiffault (NSERC Discovery grant RGPIN-2018-05755).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.agrformet.2023.109469.

Appendix A. Variables used in the model

For the cumulative precipitation of the last 12 h, the mean was calculated using a gamma distribution, with mean = α / β, where α (α > 0) is the shape parameter and β (β > 0) is the rate parameter. The package *fitdistrplus* (Delignette-Muller et al., 2020) was used.

"The VIF values were calculated taking into account 10 variables, 3 weather variables (Table A1) + 7 topographic and vegetation variables (Table A2)."

The VIF values were calculated taking into account 10 variables, 3 weather variables (Table A1) + 7 topographic and vegetation variables (Table A2).

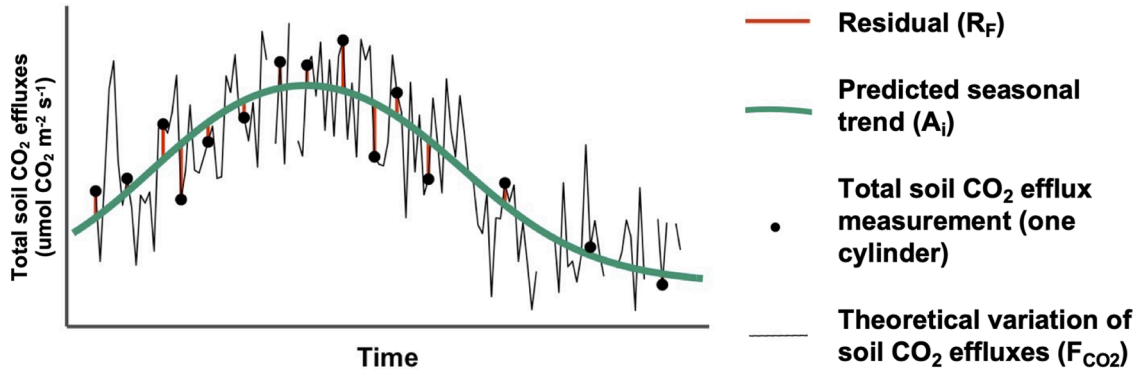
Table A1
Recent weather feature values.

Variable type	Name	Definition	For the prediction period (June to October 2020)		Scale of the measurements collected (n = 1464)		Variance inflation factor (VIF)
			Mean [5% and 95% quantile]				
Recent weather	Mean temperature (last 2 days) (°C)	Average air temperature during the two days preceding the day of the flux measurement (day of the flux measurement is not included)	June	12.50 [4.71 – 19.98]	June	15.27 [8.58 – 20.7]	6.13
			July	16.57 [13.08 – 19.98]	July	16.22 [12.65 – 20.35]	
			August	13.56 [7.68 – 17.48]	August	13.89 [7.45 – 19.20]	
			September	8.44 [2.27 – 14.89]	September	8.47 [3.65 – 16.20]	
			October	1.87 [-7.4 – 6.33]	October	2.06 [-4.20 – 5.60]	
	Interval temperature (last 2 days) (°C)	Average air temperature of the day of the flux measurement minus the mean air temperature of the day before the flux measurement	June	0.47 [-5.03 – 5.52]	June	0.81 [-3.40 – 4.20]	1.15
			July	0.05 [-3.50 – 2.80]	July	-0.39 [-4.90 – 2.30]	
			August	-0.27 [-3.25 – 2.30]	August	-0.27 [-3.25 – 2.3]	
			September	0.16 [-4.31 – 4.66]	September	-1.39 [-6.20 – 2.30]	
			October	-0.81 [-6.52 – 3.42]	October	1.84 [-2.00 – 3.70]	
	Sum precipitation (last 12 h) (mm)	The cumulative of the total precipitation of the 12 h preceding the flux measurement	June	1.54 [0.00 – 11.63]	June	4.45 [0.00 – 24.30]	1.30
			July	3.09 [0.00 – 19.49]	July	2.70 [0.00 – 5.90]	
			August	2.58 [0.00 – 12.67]	August	5.51 [0.00 – 31.90]	
			September	2.12 [0.00 – 10.94]	September	1.21 [0.00 – 4.50]	
			October	4.13 [0.00 – 14.47]	October	1.28 [0.00 – 9.60]	

Table A2
Topographic and vegetation feature values.

Variable type	Name	Definition	Experimental design level	Cylinders (n = 99) level	Variance inflation factor (VIF)
			Mean [5% and 95% quantile]		
Topography	Aspect	The orientation of the slope in radians was assessed from north (with 0=north-facing, 1.57=east-facing, 3.14=south-facing and 4.71=west-facing)	1.81 [0.40 – 3.58]	2.51 [0.75 – 3.77]	7.54
	Convergence index (Koethe and Lehmeier, 1996)	Convergence index for overland flow (0 corresponds to a regular slope, a negative value to a convergence situation and a positive value to a divergence situation)	-0.01 [-3.45 – 3.32]	0.28 [-2.56 – 3.44]	2.13
	Mid slope position (Böhner and Selige, 2006)	Value from 0 to 1, with 0 representing the middle of the slope and 1 representing the maximum vertical distance from the middle of the slope (either up the slope or down)	0.46 [0.05 – 0.90]	0.47 [0.06 – 0.92]	4.33
	Multiresolution Index of Valley Bottom Flatness (MRVBF) (Gallant and Dowling, 2003)	Identifies valley bottoms using flatness (inverse of the slope) and lowness (ranking of elevation) on different range of scale. Values below 0.5 are not valley bottom areas. Values above 1.5 are valley bottoms (the higher the value, the flatter and larger the valley bottom)	0.15 [0.03 – 0.58]	0.17 [0.00 – 0.58]	2.40
	Slope height (Böhner and Selige, 2006)	The height of the back slope	33.42 [4.12 – 88.92]	23.67 [5.16 – 56.26]	4.70
	Slope length	The length of the back slope	218.77 [17.16 – 553.30]	233.25 [24.61 – 607.05]	5.31
Vegetation	Canopy height model	Height of the canopy (1 m) (LiDAR 2016)	5.25 [0.81 – 8.59]	5.84 [2.60 – 9.42]	6.97

Appendix B. Modelling of total soil CO_2 effluxes



Scheme of the different components of the total soil CO_2 efflux (F_{CO2}) modelled from the F_{CO2} measurements taken from one cylinder. The use of the spatial component (not shown here) implies multiple point measurements.

Appendix C. Hyperparameters of the model

Table C1 and C2

Random Forest Regression model consists of combining a large number (*n*) of *regression decision trees*, where each regression decision *tree* is built from a bootstrap sample of the training data set and with a random subset of feature (predictive variables) (Hastie et al., 2009). The final output is obtained by averaging the outputs of all *regression decision trees*. Random Forest allows building a large ensemble of decorrelated trees that improves the overall performance compared to other algorithms. Random Forest trees are fast, easy to parametrize and mostly robust to overfitting (Hastie et al., 2009) when trained adequately.

The construction of an individual regression decision tree is controlled by a set of hyperparameters. A decision tree consists of recursive splits of the data set (mother) into two data sets (daughter) until the maximum size of the tree is reached (max *depth*) or if there are not enough samples (here, F_{CO2} measurements) in the mother data set to continue making splits (min *sample split* and min *sample leaf*). Splitting a mother data set into two daughter data sets is done by randomly selecting a subset of features (this number of features is called max *features*) amongst all the features and choosing the one (amongst the subset) that minimizes the sum of the square error of the two daughter data sets (Hastie et al., 2009). This operation is repeated to fit n regressions trees; the output of all regression trees is then averaged together and used as the final output of the model.

Table C1
Hyperparameters and range of tested values.

Hyperparameters		Range of tested values
n	number of decision trees in the forest	200:2000 (by 10)
min sample split	minimum number of samples required to split an internal node	2, 3, 4, 5, 10
min sample leaf	minimum number of samples required to be at a leaf node	1, 2, 3, 4, 5
max features	number of features to consider when looking for the best split	all, sqrt
max depth	maximum depth of the tree	10:110 (by 10)
bootstrap	Random point selection method (with or without replacement)	true, false

Table C2
Hyperparameter combinations based on R^2 (validation).

ID[a]	n	min sample split	min sample leaf	max feature	max depth	bootstrap	Mean R^2 (validation)	SD R^2 (validation)
A	200	5	1	sqrt	60	False	0.55	0.02
B	400	2	1	sqrt	50	True	0.55	0.03
C	200	5	1	sqrt	60	False	0.55	0.03
D	400	2	1	sqrt	50	True	0.56	0.03
E	400	5	1	auto	100	True	0.57	0.05

[a] ID used in Fig. 4.

References

Abnee, A.C., Thompson, J.A., Kolka, R.K., 2004. Landscape modeling of in situ soil respiration in a forested watershed of southeastern Kentucky, USA. Environ. Manage. 33 (1), S168–S175.

Allison, S.D., Wallenstein, M.D., Bradford, M.A., 2010. Soil-carbon response to warming dependent on microbial physiology. Nat. Geosci. 3 (5), 336–340.

Apley, D.W., Zhu, J., 2020. Visualizing the effects of predictor variables in black box supervised learning models. J. R. Stat. Soc.: Ser. B (Stat. Methodol.) 82 (4), 1059–1086.

Bélanger, L., Paquette, S., Morel, S., Bégin, J., Meek, P., Bertrand, L., Pineau, M., 1995. Indices de qualité de station du sapin baumier dans le sous-domaine écologique de la sapinière à bouleau blanc humide. Forestry Chronicle 71 (3), 317–325.

Betson, N.R., Göttlicher, S.G., Hall, M., Wallin, G., Richter, A., Högberg, P., 2007. No diurnal variation in rate or carbon isotope composition of soil respiration in a boreal forest. Tree Physiol. 27 (5), 749–756.
Böhner, J., Selige, T., 2006. Spatial prediction of soil attributes using terrain analysis and climate regionalisation. SAGA - Analysis and Modelling Applications. Goettinger Geographische Abhandlungen. Goettingen, pp. 13–28.
Bond-Lamberty, B., Bailey, V.L., Chen, M., Gough, C.M., Vargas, R., 2018. Globally rising soil heterotrophic respiration over recent decades. Nature 560 (7716), 80–83.
Bond-Lamberty, B., Thomson, A., 2010. Temperature-associated increases in the global soil respiration record. Nature 464 (7288), 579–582.
Bradshaw, C.J.A., Warkentin, I.G., 2015. Global estimates of boreal forest carbon stocks and flux. Glob. Planet. Change 128, 24–30.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Brubaker, S.C., Jones, A.J., Lewis, D.T., Frank, K., 1993. Soil properties associated with landscape position. Soil Sci. Soc. Am. J. 57 (1), 235–239.
Carney, K.M., Hungate, B.A., Drake, B.G., Megonigal, J.P., 2007. Altered soil microbial community at elevated CO_2leads to loss of soil carbon. Proc. Natl Acad. Sci. 104 (12), 4990–4995.
Clayden, S.R., Cameron, R.P., McCarthy, J.W., 2011. Perhumid boreal and hemiboreal forests of Eastern Canada. In *Temperate and Boreal Rainforests of the World: Ecology and Conservation. Dominick A. DellaSala.* Washington, DC, pp. 111–131.
Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1.4. Geosci. Model Dev. 8 (7), 1991–2007.
Craney, T.A., Surles, J.G., 2002. Model-dependent variance inflation factor cutoff values. Qual. Eng. 14 (3), 391–403.
Curiel Yuste, J., Nagy, M., Janssens, I.A., Carrara, A., Ceulemans, R., 2005. Soil respiration in a mixed temperate forest and its contribution to total ecosystem respiration. Tree Physiol. 25 (5), 609–619.
Currie, J., 1961. Gaseous diffusion in porous media. Part 3-Wet granular materials. Br. J. Appl. Phys. 12 (6), 275.
Davidson, E.A., Belk, E., Boone, R.D., 1998. Soil water content and temperature as independent or confounded factors controlling soil respiration in a temperate mixed hardwood forest. Glob. Chang. Biol. 4 (2), 217–227.
Davidson, E.A., Holbrook, N.M., 2009. Is temporal variation of soil respiration linked to the phenology of photosynthesis? In *Phenology of Ecosystem Processes: Applications in Global Change Research. Asko Noormets.* New York, NY, pp. 187–199.
Davidson, E.A., Savage, K., Bolstad, P., Clark, D.A., Curtis, P.S., Ellsworth, D.S., Zak, D., 2002a. Belowground carbon allocation in forests estimated from litterfall and IRGA-based soil respiration measurements. Agric. For. Meteorol. 113 (1), 39–51.
Davidson, E.A., Savage, K., Verchot, L.V., Navarro, R., 2002b. Minimizing artifacts and biases in chamber-based measurements of soil respiration. Agric. For. Meteorol. 113 (1), 21–37.
Delignette-Muller, M.L., Dutang, C., Pouillot, R., Denis, J.B., Siberchicot, A., 2020. Help to fit of a parametric distribution to non-censored or censored data (fitdistrplus). 1.1-3 Ed.
Ekblad, A., Högberg, P., 2001. Natural abundance of 13C in CO2 respired from forest soils reveals speed of link between tree photosynthesis and root respiration. Oecologia 127 (3), 305–308.
Environment and Climate Change Canada. 2019. Normales et moyennes climatiques de 1981-2010 - Forêt Montmorency. https://climat.meteo.gc.ca/climate_normals/results_1981_2010_f.html?searchType=stnName&txtStationName=montmorency&searchMethod=contains&txtCentralLatMin=0&txtCentralLatSec=0&txtCentralLongMin=0&txtCentralLongSec=0&stnID=5682&dispBack=1 last update: 2019-12-04 (accessed on February 5, 2022), 2020).
Environment and Climate Change Canada. 2021. Rapport de données quotidiennes (ID climatologique 7042395). https://climat.meteo.gc.ca/climate_data/daily_data_f.html?hlyRange=2003-11-10%7C2022-09-09&dlyRange=2003-11-10%7C2022-09-09&mlyRange=%7C&StationID=32413&Prov=QC&urlExtension=_f.html&searchType=stnName&optLimit=yearRange&StartYear=1840&EndYear=2022&selRowPerPage=25&Line=1&searchMethod=contains&txtStationName=montmorency&timeframe=2&Day=9&Year=2020&Month=1#.
Fortin, J.A., Lamhamedi, M.S., 2009. Ecophysiology of sporocarp development of ectomycorrhizal basidiomycetes associated with boreal forest gymnosperms. In Advances in mycorrhizal science and technology. D. P. Khasa, Y. Piche and A. P. Coughlan. pp. 161–173.
Elzhov, T.V., Mullen, K.M., Spiess, A.N., Bolker, B., 2016. In: Interface to the Levenberg-Marquardt Nonlinear Least-Squares Algorithm Found in MINPACK. Plus Support for Bounds, p. minpack.lm.
Gallant, J.C., Dowling, T.I., 2003. A multiresolution index of valley bottom flatness for mapping depositional areas. Water Resour. Res. 39 (12), 1–13.
Gaumont-Guay, D., Black, T.A., Griffis, T.J., Barr, A.G., Jassal, R.S., Nesic, Z., 2006. Interpreting the dependence of soil respiration on soil temperature and water content in a boreal aspen stand. Agric. For. Meteorol. 140 (1), 220–235.
Granier, A., Reichstein, M., Bréda, N., Janssens, I.A., Falge, E., Ciais, P., Wang, Q., 2007. Evidence for soil water control on carbon and water dynamics in European forests during the extremely dry year: 2003. Agric. For. Meteorol. 143 (1), 123–145.
Greiser, C., Meineri, E., Luoto, M., Ehrlén, J., Hylander, K., 2018. Monthly microclimate models in a managed boreal forest landscape. Agric. For. Meteorol. 250, 147–158.
Guillemette, F., Plamondon, A.P., Prévost, M., Lévesque, D., 2005. Rainfall generated stormflow response to clearcutting a boreal forest: peak flow comparison with 50 world-wide basin studies. J. Hydrol. (Amst) 302 (1), 137–153.
Hadley, W., Winston, C., Lionel, H., Thomas Lin, P., Kohske, T., Claus, W., Dewey, D., 2020. ggplot2 (R 4.0.2). Title: create elegant data visualisations using the grammar of graphics. 3.3.2 Ed., pp. Depends: r (>= 3.2).
Han, G., Zhou, G., Xu, Z., Yang, Y., Liu, J., Shi, K., 2007. Biotic and abiotic factors controlling the spatial and temporal variation of soil respiration in an agricultural ecosystem. Soil Biol. Biochem. 39 (2), 418–425.
Hanson, P.J., Wullschleger, S.D., Bohlman, S.A., Todd, D.E., 1993. Seasonal and topographic patterns of forest floor CO2 efflux from an upland oak forest. Tree Physiol. 13 (1), 1–15.
Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. springer, New York, pp. 1–758. Vol. 2.
Herbst, M., Prolingheuer, N., Graf, A., Huisman, J., Weihermuller, L., Vanderborght, J., 2009. Characterization and understanding of bare soil respiration spatial variability at plot scale. Vadose Zone J. 8 (3), 762–771.
Högberg, P., Bhupinderpal, S., Löfvenius, M.O., Nordgren, A., 2009. Partitioning of soil respiration into its autotrophic and heterotrophic components by means of tree-girdling in old boreal spruce forest. For. Ecol. Manage. 257 (8), 1764–1767.
Högberg, P., Nordgren, A., Buchmann, N., Taylor, A.F.S., Ekblad, A., Högberg, M.N., Read, D.J., 2001. Large-scale forest girdling shows that current photosynthesis drives soil respiration. Nature 411 (6839), 789–792.
Hooker, J., Duveiller, G., Cescatti, A., 2018. A global dataset of air temperature derived from satellite remote sensing and weather stations. Sci. Data 5 (1), 180246.
Huang, N., Gu, L., Black, T.A., Wang, L., Niu, Z., 2015. Remote sensing-based estimation of annual soil respiration at two contrasting forest sites. J. Geophys. Res.: Biogeosci. 120 (11), 2306–2325.
Janssens, I.A., Kowalski, A.S., Longdoz, B., Ceulemans, R., 2000. Assessing forest soil CO2 efflux: an in situ comparison of four techniques. Tree Physiol. 20 (1), 23–32.
Janssens, I.A., Lankreijer, H., Matteucci, G., Kowalski, A.S., Buchmann, N., Epron, D., Valentini, R., 2001. Productivity overshadows temperature in determining soil and ecosystem respiration across European forests. Glob. Chang. Biol. 7 (3), 269–278.
Jian, J., Steele, M.K., Zhang, L., Bailey, V.L., Zheng, J., Patel, K.F., Bond-Lamberty, B.P., 2022. On the use of air temperature and precipitation as surrogate predictors in soil respiration modelling. Eur. J. Soil Sci. 73 (1), e13149.
Jumelle, M., Rajaratnam, R., Kuhn-Regnier, A., 2020. Python accumulated local effects (ALEPyhton) (Apache-2.0). https://github.com/blent-ai/ALEPython.
Jungqvist, G., Oni, S.K., Teutschbein, C., Futter, M.N., 2014. Effect of climate change on soil temperature in Swedish boreal forests. PLoS One 9 (4).
Kang, S., Kim, S., Oh, S., Lee, D., 2000. Predicting spatial and temporal patterns of soil temperature based on topography, surface cover and air temperature. For. Ecol. Manage. 136 (1), 173–184.
Kelliher, F.M., Ross, D.J., Law, B.E., Baldocchi, D.D., Rodda, N.J., 2004. Limitations to carbon mineralization in litter and mineral soil of young and old ponderosa pine forests. For. Ecol. Manage. 191 (1), 201–213.
Kergoat, L., 1998. A model for hydrological equilibrium of leaf area index on a global scale. J. Hydrol. (Amst.) 212-213, 268–286.
Khomik, M., Arain, M.A., McCaughey, J.H., 2006. Temporal and spatial variability of soil respiration in a boreal mixedwood forest. Agric. For. Meteorol. 140 (1), 244–256.
Kimball, B.A., 1983. Canopy gas exchange: gas exchange with soil. Limitations to Efficient Water Use in Crop Production, pp. 215–226.
Koethe, R., Lehmeier, F., 1996. System zur automatischen relief-analyse. User Manual, 2. *Edition [Dept. of Geography, University of Goettingen, unpublished].*
Kurz, W.A., Shaw, C.H., Boisvenue, C., Stinson, G., Metsaranta, J., Leckie, D., Neilson, E. T., 2013. Carbon in Canada's boreal forest—a synthesis. Environ. Rev. 21 (4), 260–292.
Kuzyakov, Y., 2006. Sources of CO2 efflux from soil and review of partitioning methods. Soil Biol. Biochem. 38 (3), 425–448.
Kuzyakov, Y., Cheng, W., 2004. Photosynthesis controls of CO2 efflux from maize rhizosphere. Plant Soil 263 (1), 85–99.
Laganière, J., Paré, D., Bergeron, Y., Chen, H.Y.H., 2012. The effect of boreal forest composition on soil respiration is mediated through variations in soil temperature and C quality. Soil Biol. Biochem. 53, 18–27.
Lavigne, M.B., Foster, R.J., Goodine, G., 2004. Seasonal and annual changes in soil respiration in relation to soil temperature, water potential and trenching. Tree Physiol. 24 (4), 415–424.
Lavigne, M.B., Ryan, M.G., Anderson, D.E., Baldocchi, D.D., Crill, P.M., Fitzjarrald, D.R., Striegl, R.G., 1997. Comparing nocturnal eddy covariance measurements to estimates of ecosystem respiration made by scaling chamber measurements at six coniferous boreal sites. J. Geophys. Res.: Atmos. 102 (D24), 28977–28985.
Lembrechts, J.J., van den Hoogen, J., Aalto, J., Ashcroft, M.B., De Frenne, P., Kemppinen, J., Lenoir, J., 2022. Global maps of soil temperature. Glob. Chang. Biol. 00, 1–35.
Levenberg, K., 1944. A method for the solution of certain non-linear problems in least squares. Quarterly of applied mathematics 2 (2), 164–168.
Lindenmayer, D., Blanchard, W., McBurney, L., Bowd, E., Youngentob, K., Marsh, K., Taylor, C., 2022. Stand age related differences in forest microclimate. For. Ecol. Manage. 510, 120101.
Lindroth, A., Grelle, A., Morén, A.-.S., 1998. Long-term measurements of boreal forest carbon balance reveal large temperature sensitivity. Glob. Chang. Biol. 4 (4), 443–450.
Linn, D.M., Doran, J.W., 1984. Effect of water-filled pore space on carbon dioxide and nitrous oxide production in tilled and nontilled soils. Soil Sci. Soc. Am. J. 48 (6), 1267–1272.
Lloyd, J., Taylor, J.A., 1994. On the temperature dependence of soil respiration. Funct. Ecol. 8 (3), 315–323.
Ma, Y., Piao, S., Sun, Z., Lin, X., Wang, T., Yue, C., Yang, Y., 2014. Stand ages regulate the response of soil respiration to temperature in a Larix principis-rupprechtii plantation. Agric. For. Meteorol. 184, 179–187.

MacMillan, R.A., Pettapiece, W.W., Nolan, S.C., Goddard, T.W., 2000. A generic procedure for automatically segmenting landforms into landform elements using DEMs, heuristic rules and fuzzy logic. Fuzzy Sets Syst. 113 (1), 81–109.
Malhi, Y., Baldocchi, D., Jarvis, P., 1999. The carbon balance of tropical, temperate and boreal forests. Plant Cell Environ. 22 (6), 715–740.
Marquardt, D.W., 1963. An algorithm for least-squares estimation of nonlinear parameters. Journal of the society for Industrial and Applied Mathematics 11 (2), 431–441.
Martin, J.G., Bolstad, P.V., 2009. Variation of soil respiration at three spatial scales: components within measurements, intra-site variation and patterns on the landscape. Soil Biol. Biochem. 41 (3), 530–543.
Marty, C., Piquette, J., Morin, H., Bussières, D., Thiffault, N., Houle, D., Paré, M.C., 2019. Nine years of in situ soil warming and topography impact the temperature sensitivity and basal respiration rate of the forest floor in a Canadian boreal forest. PLoS One 14 (12), e0226909.
McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117 (1), 3–52.
Mielcarek, M., Stereńczak, K., Khosravipour, A., 2018. Testing and evaluating different LiDAR-derived canopy height model generation methods for tree height estimation. Int. J. Appl. Earth Obs. Geoinf. 71, 132–143.
Minasny, B., McBratney, A.B., 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Comput. Geosci. 32 (9), 1378–1388.
Minasny, B., McBratney, A.B., Malone, B.P., Wheeler, I., 2013. Chapter one - digital mapping of soil carbon. In Advances in Agronomy. Donald L. Sparks. pp. 1–47.
Moldrup, P., Olesen, T., Komatsu, T., Schjønning, P., Rolston, D.E., 2001. Tortuosity, diffusivity, and permeability in the soil liquid and gaseous phases. Soil Sci. Soc. Am. J. 65 (3), 613–623.
Molnar, C., 2019. *Interpretable machine learning. A guide for making black box models explainable*. https://christophm.github.io/interpretable-ml-book/.
Morén, A.-S., Lindroth, A., 2000. CO2 exchange at the floor of a boreal forest. Agric. For. Meteorol. 101 (1), 1–14.
Norman, J., Kucharik, C., Gower, S., Baldocchi, D., Crill, P., Rayment, M., Striegl, R.G., 1997. A comparison of six methods for measuring soil-surface carbon dioxide fluxes. J. Geophys. Res.: Atmos. 102 (D24), 28771–28777.
Oliver, S.A., Oliver, H.R., Wallace, J.S., Roberts, A.M., 1987. Soil heat flux and temperature variation with vegetation, soil type and climate. Agric. For. Meteorol. 39 (2), 257–269.
Orchard, V.A., Cook, F.J., 1983. Relationship between soil respiration and soil moisture. Soil Biol. Biochem. 15 (4), 447–453.
Osborn, T.J., Jones, P.D., Lister, D.H., Morice, C.P., Simpson, I.R., Winn, J., Harris, I.C., 2021. Land surface air temperature variations across the globe updated to 2019: the CRUTEM5 data set. J. Geophys. Res.: Atmos. 126 (2), e2019JD032352.
Oyler, J.W., Ballantyne, A., Jencso, K., Sweet, M., Running, S.W., 2015. Creating a topoclimatic daily air temperature dataset for the conterminous United States using homogenized station data and remotely sensed land skin temperature. Int. J. Climatol. 35 (9), 2258–2279.
Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Canadell, J.G., 2011. A large and persistent carbon sink in the world's forests. Science 333 (6045), 988–993.
Paul, K.I., Polglase, P.J., Smethurst, P.J., O'Connell, A.M., Carlyle, C.J., Khanna, P.K., 2004. Soil temperature under forests: a simple model for predicting soil temperature under a range of forest types. Agric. For. Meteorol. 121 (3), 167–182.
Payeur-Poirier, J.-L., Coursolle, C., Margolis, H.A., Giasson, M.-A., 2012. CO2 fluxes of a boreal black spruce chronosequence in eastern North America. Agric. For. Meteorol. 153, 94–105.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Duchesnay, E., 2011. Scikit-learn: machine Learning in {P}ython. J. Mach. Learn. Res. 12, 2825–2830.
Perez-Quezada, J.F., Brito, C.E., Cabezas, J., Galleguillos, M., Fuentes, J.P., Bown, H.E., Franck, N., 2016. How many measurements are needed to estimate accurate daily and annual soil respiration fluxes? Analysis using data from a temperate rainforest. Biogeosciences 13 (24), 6599–6609.
Phillips, S.C., Varner, R.K., Frolking, S., Munger, J.W., Bubier, J.L., Wofsy, S.C., Crill, P. M., 2010. Interannual, seasonal, and diel variation in soil respiration relative to ecosystem respiration at a wetland to upland slope at Harvard Forest. J. Geophys. Res.: Biogeosci. 115 (G2).
Plamondon, A.P., Ouellet, D.C., 1980. Partial clearcutting and streamflow regime of ruisseau des Eaux-Volees experimental basin. The influence of man on the hydrological regime with special reference to representative and experimental basins., Helsinki Symposium, June 1980 (International Association of Hydrological Sciences Washington DC; IAHS-AISH Publication 1130). pp. 1129–1136.
Prévost, M., Raymond, P., 2012. Effect of gap size, aspect and slope on available light and soil temperature after patch-selection cutting in yellow birch–conifer stands, Quebec, Canada. For. Ecol. Manage. 274, 210–221.
Python Core Team, 2015. Python: A Dynamic, Open Source Programming Language. 3.7.10 Ed. Python Software Foundation.
R Core Team, 2020. R: A Language and Environment for Statistical Computing. R version 4.0.3 (2020-10-10) – "Bunny-Wunnies Freak Out" Ed. R Foundation for Statistical Computing, Vienna, Austria.
Raich, J.W., 1998. Aboveground productivity and soil respiration in three Hawaiian rain forests. For. Ecol. Manage. 107 (1), 309–318.
Raich, J.W., Schlesinger, W.H., 1992. The global carbon dioxide flux in soil respiration and its relationship to vegetation and climate. Tellus B 44 (2), 81–99.
Raich, J.W., Tufekciogul, A., 2000. Vegetation and soil respiration: correlations and controls. Biogeochemistry 48 (1), 71–90.
Rayment, M.B., Jarvis, P.G., 2000. Temporal and spatial variation of soil CO2 efflux in a Canadian boreal forest. Soil Biol. Biochem. 32 (1), 35–45.
Reichstein, M., Rey, A., Freibauer, A., Tenhunen, J., Valentini, R., Banza, J., Yakir, D., 2003. Modeling temporal and large-scale spatial variability of soil respiration from soil water availability, temperature and vegetation productivity indices. Glob. Biogeochem. Cycles 17 (4), 1–15.
Saiz, G., Byrne, K.A., Butterbach-Bahl, K., Kiese, R., Blujdea, V., Farrell, E.P., 2006. Stand age-related effects on soil respiration in a first rotation Sitka spruce chronosequence in central Ireland. Glob. Chang. Biol. 12 (6), 1007–1020.
Savage, K., Davidson, E., Richardson, A., 2008. A conceptual and practical approach to data quality and analysis procedures for high-frequency soil respiration measurements. Funct. Ecol. 22 (6), 1000–1007.
Savage, K.E., Davidson, E.A., 2003. A comparison of manual and automated systems for soil CO2 flux measurements: trade-offs between spatial and temporal resolution. J. Exp. Bot. 54 (384), 891–899.
Sayer, E.J., 2006. Using experimental manipulation to assess the roles of leaf litter in the functioning of forest ecosystems. Biol. Rev. 81 (1), 1–31.
Schindlbacher, A., Wunderlich, S., Borken, W., Kitzler, B., Zechmeister-Boltenstern, S., Jandl, R., 2012. Soil respiration under climate change: prolonged summer drought offsets soil warming effects. Glob. Chang. Biol. 18 (7), 2270–2279.
Seabold, Skipper, Perktold, J., 2010. statsmodels: econometric and statistical modeling with python. In: 9th Python in Science Conference (ed.). 0.12.2 Ed.
Skopp, J., Jawson, M., Doran, J., 1990. Steady-state aerobic microbial activity as a function of soil water content. Sci. Soc. America J. 54 (6), 1564–1571.
Soil Classification Working Group, 1998. The Canadian system of soil classification. Agric. Agri-Food Canada. Publication 1646, 187.
Song, Y., Zhou, D., Zhang, H., Li, G., Jin, Y., Li, Q., 2013. Effects of vegetation height and density on soil temperature variations. Chin. Sci. Bull. 58 (8), 907–912.
Subke, J.-A., Reichstein, M., Tenhunen, J.D., 2003. Explaining temporal variation in soil CO2 efflux in a mature spruce forest in Southern Germany. Soil Biol. Biochem. 35 (11), 1467–1483.
Tamai, K., 2010. Effects of environmental factors and soil properties on topographic variations of soil respiration. Biogeosciences 7 (3), 1133–1142.
Tremblay, S.L., D'Orangeville, L., Lambert, M.-C., Houle, D., 2018. Transplanting boreal soils to a warmer region increases soil heterotrophic respiration as well as its temperature sensitivity. Soil Biol. Biochem. 116, 203–212.
Trumbore, S., 2000. Age of soil organic matter and soil respiration: radiocarbon constraints on belowground c dynamics. Ecol. Appl. 10 (2), 399–411.
Tuszynski, J., 2020. caTools. Tools: moving Window Statistics, GIF, Base64, ROC AUC, etc. 1.18.0 Ed.
Vargas, R., Allen, M.F., 2008. Environmental controls and the influence of vegetation type, fine roots and rhizomorphs on diel and seasonal variation in soil respiration. New Phytol. 179 (2), 460–471.
Vargas, R., Baldocchi, D.D., Bahn, M., Hanson, P.J., Hosman, K.P., Kulmala, L., Yang, B., 2011a. On the multi-temporal correlation between photosynthesis and soil CO2 efflux: reconciling lags and observations. New Phytol. 191 (4), 1006–1017.
Vargas, R., Carbone, M.S., Reichstein, M., Baldocchi, D.D., 2011b. Frontiers and challenges in soil respiration research: from measurements to model-data integration. Biogeochemistry 102 (1), 1–13.
Vargas, R., Detto, M., Baldocchi, D.D., Allen, M.F., 2010. Multiscale analysis of temporal variability of soil CO2 production as influenced by weather and vegetation. Glob. Chang. Biol. 16 (5), 1589–1605.
Wang, H., McConkey, B., Curtin, D., Cutforth, H., 2010. Estimation of daily soil CO2 flux using a single-time-point measurement. Can. J. Soil Sci. 90 (3), 517–522.
Warner, D.L., Guevara, M., Inamdar, S., Vargas, R., 2019. Upscaling soil-atmosphere CO2 and CH4 fluxes across a topographically complex forested landscape. Agric. For. Meteorol. 264, 80–91.
Warner, D.L., Vargas, R., Seyfferth, A., Inamdar, S., 2018. Transitional slopes act as hotspots of both soil CO2 emission and CH4 uptake in a temperate forest landscape. Biogeochemistry 138 (2), 121–135.
Webster, K., Creed, I., Bourbonniere, R., Beall, F., 2008. Controls on the heterogeneity of soil respiration in a tolerant hardwood forest. J. Geophys. Res.: Biogeosci. (G3), 113.