



HAL
open science

Soil water balance performs better than climatic water variables in tree species distribution modelling

Christian Piedallu, Jean-Claude Gégout, Vincent Perez, François F. Lebourgeois

► **To cite this version:**

Christian Piedallu, Jean-Claude Gégout, Vincent Perez, François F. Lebourgeois. Soil water balance performs better than climatic water variables in tree species distribution modelling. *Global Ecology and Biogeography*, 2013, 22 (4), pp.470-482. <10.1111/geb.12012>. <hal-00835755>

HAL Id: hal-00835755

<https://hal.science/hal-00835755v1>

Submitted on 19 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Soil water balance performs better than climatic water variables in tree species distribution modelling

Christian Piedallu^{1,2,*}, Jean-Claude Gégout^{1,2}, Vincent Perez^{1,2}, François Lebourgeois^{1,2}

¹*AgroParisTech, UMR1092, Laboratoire d'Étude des Ressources Forêt-Bois (LERFoB), ENGREF, 14 rue Girardet, CS14216, FR-54042 Nancy Cedex, France*

²*INRA, UMR1092, Laboratoire d'Étude des Ressources Forêt-Bois (LERFoB), Centre INRA de Nancy, F-54280 Champenoux, France*

*Corresponding author

AgroParisTech, UMR1092, Laboratoire d'Étude des Ressources Forêt-Bois (LERFoB), ENGREF, 14 rue Girardet, CS14216, FR-54042 Nancy Cedex, France

E-mail: christian.piedallu@engref.agroparistech.fr

Article type: research paper

Abstract

Aim: Soil water is essential for the physiological processes of plant growth and fitness. Owing to the difficulty of assessing wide variations in soil water reserves, plant distribution models usually estimate available water for plants through such climatic proxies as precipitation data (P) or climatic water balance (P minus potential evapotranspiration). We evaluated the ability of simple climatic proxies and soil water balance indices to predict the ecological niches of forest tree species.

Location: France

Methods: Soil water content and deficits were computed and mapped at a resolution of 1 × 1 km throughout France. The predictive abilities of these indices were compared to those of P and climatic water balance to model the distributions of 37 of the most common European tree species. We focused on two species with contrasting water tolerance, *Quercus robur* and *Quercus pubescens*, to illustrate the differences between climatic proxies and soil water balance in species response curves and distribution maps.

Results: Throughout France, soil water content was poorly correlated with P and climatic water balance, because low P in the lowlands can be compensated for by water provided by deeper soils, which is not the case in most mountainous areas. Soil water balance performed better than simple climatic water variables for explaining tree species distribution, improving 82% of the models for hygrophilous, meso-hygrophilous, meso-xerophilous, and meso-xerophilous species.

Main conclusions: Our results showed that simple climatic values do not accurately represent available water for trees and that soil water balance indices perform better than do climatic proxies for most species. This point is crucial in avoiding underestimating the importance of water in studies aiming to determine the ecological niches of plant species and their responses to climate change.

Key words: water balance; water availability; digital soil mapping; species distribution model; tree ecology; ecological niche; spatial biogeography; climate change; geographic information system

INTRODUCTION

Plant distribution models are used worldwide for a variety of applications in ecology and biogeography. Studies of this nature are increasing as research on the impacts of global change on vegetation continues to expand (Guisan and Thuiller, 2005). Such applications include identification of priority conservation areas, invasive species control and planning, and identification of native assemblages for restoration (Platts *et al.*, 2010; Siles *et al.*, 2010). Despite the recent development of remote sensing and geographic information system (GIS) techniques to obtain spatially distributed explanatory variables, the availability over large areas of ecological descriptors influencing plant physiology is limited (Synes and Osborne, 2011).

Soil water availability is recognized as one of the most important ecological factors for plant growth and development, influencing both gross primary production and CO₂ uptake (Davi *et al.*, 2005). The ecophysiological effects of water stress have been fully documented for plants, and include decreased transpiration and inhibited photosynthesis, resulting in plant death if droughts are severe or repeated (Breda *et al.*, 2006; Reichstein *et al.*, 2007). The water available to plants is commonly evaluated using soil water balance, the change in water resources stored in the rooted soil taking into account incoming and outgoing fluxes over a given time (Zhang *et al.*, 2002).

Although the relationship between available water and vegetation distribution was identified decades ago (Stephenson, 1990), soil water balance is rarely used in plant distribution studies. In recent years, only a few studies have characterized the responses of species to soil water content (SWC) to analyse their sensitivity to recorded or expected climate change (Engelbrecht *et al.*, 2007; Lutz *et al.*, 2010). Individual species response to soil water availability has been identified as a direct determinant of niche differentiation. The importance of available water in plant vulnerability to climate change has been illustrated by several researchers. In a study in California, Crimmins *et al.* (2011) attributed an observed shift in species distribution to changes in climatic water balance rather than to changes in temperature. Moreover, increased drought conditions have been shown to play a key role in tree mortality rates in the western United States (Van Mantgem *et al.*, 2009). Because ongoing climate change may strongly affect the water cycle (Holsten *et al.*, 2009), these studies emphasize the need for ecologists and land managers to evaluate water budgets more thoroughly to accurately determine the ecological niches of plants and their responses to climate warming.

Since the advent of GIS techniques in the early 1990s, a number of spatially distributed water balance models have been described. Most of these were developed for hydrological purposes and are unsuitable for regional scale simulations (Beven, 2001). They are generally not convenient for modelling species distribution, which require large ecological gradients for determining ecological niches. Digital spatially distributed soil data have only recently become available for large areas that can describe the maximum amount of water that can be stored by the soil, defined as the soil water-holding capacity (SWHC) (Dyer, 2009). However, plant ecologists commonly predict hydrological processes using easily available climatic proxies such as precipitation (P) or climatic water balance (CWB), computed as the difference between P and potential evapotranspiration (PET) (Kerkhoff *et al.*, 2004). The consequences of using climatic proxies instead of soil water balance have been poorly studied; therefore, the performance differences between these predictors of plant distribution remain unknown. Despite the consequences of drought on plant fitness, the role of available water in determining ecological niche and the effects of expected changes in water availability on species range shifts are poorly understood.

The aim of this study was to map soil water balance over a broad area and to compare its ability to describe the ecological niches of forest tree species with those of simple climatic proxies. Considering the annual sum of growing degree days >5°C (GDD₅), minimum winter temperature

(minWT), and various water availability indices, we ran distribution models for 37 tree species, including 20 of the 25 most common European tree species. To evaluate water availability, we compared 3 climatic proxies with 2 indices of soil water balance. For the climatic proxies, we used P and 2 CWB indices, one considering the entire range of differences between P and PET (CWB_w) and the other only negative values (CWB₋₀). The soil water balance indices used were the Thornthwaite SWC and soil water deficit (SWD). To illustrate the differences in the effects of these indices on species niche assessment and distribution maps, we focused on 2 species of the same genus that display contrasting drought tolerances, *Quercus robur* (a meso-hygrophilous species) and *Quercus pubescens* (a xerophilous species).

MATERIALS AND METHODS

Species dataset

We selected 37 species of the most common trees growing in Europe with various drought tolerances, excluding the most frequently planted species (Table 1) (Rameau *et al.*, 1989; Timbal and Aussenac, 1996). Presence and absence were extracted from 32,828 400-m² georeferenced plots from a phytocological database collected by the French National Forest Inventory (NFI) between 2005 and 2008, without minimum tree size (Drapier and Cluzeau, 2001). The data were recorded on a systematic sampling grid covering all of France that shifted every year, and were georeferenced with a precision of ±500 m. This database has the advantage of having homogeneous sample intensity within the forest area (about 1 plot per 5 km²), while avoiding spatial autocorrelation problems due to small distances between plots. The mean prevalence for all species was approximately 10% (range, 0.10–40%). With elevations ranging from 0 to 2500 m and 5 biogeographical regions represented, a wide range of European ecological conditions were included.

Climatic data

Monthly values for mean temperature, total P, and cloudiness were used for the 1961–1990 period in climate models based on GIS techniques developed by Ninyerola *et al.* (2000). These values were modelled and mapped over the entire French territory using hybrid methods that combined modelling with spatially distributed variables characterising the topography or distance to the sea and interpolation of model residuals. Data gathered from 237, 432, and 87 meteorological stations (source: Météo-France) throughout France were used to model mean temperature, P, and cloudiness, respectively. Cloudiness was then combined with clear sky solar radiation from the Helios model to produce solar radiation maps (Piedallu and Gegout, 2007). These models were validated using 493 independent meteorological stations for temperature, 471 for P, and 86 for solar radiation. The models explained 93%, 82%, and 78% of the variance in temperature, solar radiation, and P of the measured annual values (Table 2). Minimum temperatures were extracted from 1-km² resolution-gridded datasets provided by Météo-France and processed using a similar approach (Benichou and Le Breton, 1987).

This climatological dataset was used to derive various indices characterising the physiological constraints on plant development used in plant distribution studies (Marmion *et al.*, 2009). The effects of temperature on vegetation were defined by GDD₅, representing the yearly sum of temperatures above 5°C, and minWT, representing frost (Prentice *et al.*, 1992). For the water-related climatic parameters, CWB was used to complement P, calculated as the difference between P and PET. PET is the amount of water that can be evaporated and transpired by a standard crop when available water is unlimited. PET was estimated monthly using a combination of mean temperature and solar radiation using the Turc formula (Turc, 1961), which is recognized as one of the best

compromises between data availability and performance (Lu *et al.*, 2005). We used both the CWB_w , considering the entire gradient, and $CWB_{<0}$, the difference between P and PET when outgoing water is greater than incoming water. For $CWB_{<0}$, a null value was assigned when no deficit existed. These indices were mapped at a 1-km² resolution for the 1961–1990 period and monthly values were averaged for CWB_w and summed for P and $CWB_{<0}$ to determine spring, summer, autumn, and annual values.

Species	Pr	Water tolerance
<i>Acer pseudoplatanus</i>	13.8	Hygrophilous
<i>Alnus glutinosa</i>	4.9	
<i>Betula pubescens</i>	0.3	
<i>Fraxinus excelsior</i>	31.0	
<i>Prunus padus</i>	0.8	
<i>Salix alba</i>	0.6	
<i>Salix cinerea</i>	3.0	
<i>Acer platanoides</i>	2.6	Meso- hygrophilous
<i>Alnus incana</i>	0.3	
<i>Fraxinus angustifolia</i>	0.8	
<i>Populus tremula</i>	13.1	
<i>Quercus robur</i>	37.4	
<i>Salix caprea</i>	14.0	
<i>Salix fragilis</i>	0.1	
<i>Abies alba</i>	15.8	Mesophilous
<i>Acer campestre</i>	19.0	
<i>Carpinus betulus</i>	30.4	
<i>Fagus sylvatica</i>	39.6	
<i>Larix decidua</i>	1.7	
<i>Pinus cembra</i>	0.3	
<i>Pinus sylvestris</i>	18.6	
<i>Quercus petraea</i>	28.1	
<i>Sorbus aucuparia</i>	10.8	
<i>Sorbus torminalis</i>	14.6	
<i>Tilia cordata</i>	3.8	
<i>Tilia platyphyllos</i>	3.9	
<i>Acer opalus</i>	4.0	Meso- xerophilous
<i>Fraxinus ornus</i>	0.9	
<i>Pinus pinaster</i>	0.3	
<i>Quercus pubescens</i>	19.5	
<i>Quercus pyrenaica</i>	2.3	
<i>Quercus suber</i>	1.1	
<i>Sorbus aria</i>	16.9	
<i>Acer monspessulanum</i>	3.2	Xerophilous
<i>Pinus halepensis</i>	2.7	
<i>Prunus mahaleb</i>	3.6	
<i>Quercus ilex</i>	9.6	

Table 1: Prevalence (Pr, in %) of the 37 studied species organised into ecological groups by drought tolerance (Julve, 1998).

Soil water balance mapping

Soil water balance was calculated using climatic data and an SWHC map (cell size, 1 km²) (Piedallu *et al.*, 2011). The SWHC map was obtained through interpolation using kriging of SWHC values estimated for 100,307 plots over the 1985–2004 period, surveyed by the French NFI in quasi-systematic sampling over all forests in France, with a mean distance apart of approximately 1 km (Drapier and Cluzeau, 2001). SWHC was estimated for each plot by taking into account both the stone content at the soil pit scale and rock outcrops at the plot scale for a maximum soil depth of 1 m (Piedallu *et al.*, 2011). For each horizon, the difference in potential water content between –100 hPa and –15,000 hPa was evaluated using class pedotransfer functions. A comparison of the SWHC values provided by the map and that estimated for independent plots collected on a regular grid by the NFI over the 2005–2008 period resulted in a Pearson's $r = 0.59$ and a root mean square error of 33.9 mm for values ranging between 0 and 148 mm ($n = 20,595$, Table 2). This degree of correlation is likely underestimated, as the validation was affected by inconsistency of scale. SWHC map predictions represent a mean value over a 1-km² cell, while the validation dataset provided an estimate at the plot scale that was unrepresentative because of the substantial spatial heterogeneity. The map predicted the growth of three species (*Picea abies*, *Fagus sylvatica*, and *Quercus petraea*) with a precision close to that of the SWHC values estimated directly from the plots (Piedallu *et al.*, 2011). Smoothing during the interpolation procedure probably minimized the effects of field survey errors and unrepresentative conditions at the plot scale. These results illustrate the value of using SWHC extracted from GIS maps to study tree species ecology.

	n	Spring	Summer	Autumn	Annual
Temperature (°C)	493	0.96 (0.52)	0.96 (0.57)	0.97 (0.51)	0.96 (0.54)
Precipitation (mm)	471	0.89 (10.9)	0.93 (8.42)	0.89 (13.4)	0.88 (11.4)
Solar radiation (MJ/m²)	86	0.86 (24.4)	0.81 (42.1)	0.89 (20.0)	0.91 (227)
SWHC (mm)	20595	0.59 (33.9)			

Table 2: Performance of the models used to calculate the elementary variables required to estimate water balance expressed by Pearson's coefficient of correlation (r) and root mean square error (in parentheses) between predicted values and measurements at meteorological stations. n , number of meteorological stations used for validation; SWHC, soil water-holding capacity. For each period, the average of the monthly values was used for temperature and the sum of the monthly values for precipitation and solar radiation.

We chose the Thornthwaite formulas to compute soil water balance for the 1961–1990 period (Thornthwaite and Mather, 1955) because they are among the most commonly used methods and require a limited number of parameters for the studied area (Lutz *et al.*, 2010). The formulas allow estimation of monthly water content fluctuations driven by P and PET . When evapotranspirative water loss exceeds P , part of the water demand is supplied by soil moisture, which decreases SWC following a negative exponential curve. The actual evapotranspiration (AET) is the amount of water that can be evaporated and transpired given the level of soil water availability. AET becomes lower than PET when the amount of water remaining in the soil decreases. A SWD appears when the evaporative demand is not met by the available water. SWC is calculated based on the difference between P and PET for a given month (t):

- If $P_{(t)} \geq PET_{(t)}$, then $SWC_{(t)} = \text{minimum value between } SWC_{(t-1)} + P_{(t)} - PET_{(t)} \text{ and } SWHC$, and $AET_{(t)} = PET_{(t)}$
- If $P_{(t)} < PET_{(t)}$, then $SWC_{(t)} = SWC_{(t-1)} \times \exp\{(P_{(t)} - PET_{(t)})/SWHC\}$, and $AET_{(t)} = SWC_{(t-1)} + (P_{(t)} - SWC_{(t)})$
- $SWD_{(t)} = PET_{(t)} - AET_{(t)}$

We calculated classically used indices in vegetation studies to determine the spring, summer, autumn, and annual values for the 1961–1990 period by averaging monthly values for SWC, AET, AET/PET ratio, and SWC/SWHC ratio and by summing the SWD values (Granier *et al.*, 1999; Schwarzel *et al.*, 2009). Herein, we present only the results for SWC and SWD because they performed better than the other variables in explaining plant distribution.

Plant distribution models

We modelled species distributions with seasonal predictors using a generalised additive model, a non-parametric method with a logistic link function and a binomial error distribution that fit a smoothing spline, allowing for consideration of complex response shapes (Hastie and Tibshirani, 1990). We allowed up to 4 degrees of freedom for the smoothing functions. To evaluate the importance of water on plant distribution and compare the efficiency of various moisture indices, we modelled species occurrence coded as the presence or absence of the 37 species with the various indices. For each species and water variable, a model was fitted using the spring, summer, and autumn mean values. The best model for each variable was used. Initially, we compared the ability of the best climatic proxy (P, CWB_w, or CWB_{<0}) and the best soil water balance index (SWC or SWD) to improve the models for each species based on temperature (GDD₅ and minWT). The performances of the different predictors were then averaged for various ecological groups classified by drought tolerance (Julve, 1998). The predictive ability of the water-related variables was compared between ecological groups and with temperature (GDD₅ or minWT). The area under the receiver operating characteristic curve (AUC) was the primary measure used to evaluate the models (Fielding and Bell, 1997), and the results were compared with those obtained using true skill statistics (TSS; Allouche *et al.*, 2006). For each predictor, we also determined the percentage of models for which they performed best.

Effects of soil water reserve on response curves for water and distribution maps: an illustration for *Q. robur* and *Q. pubescens*

We focused on the differences elicited by the choice between climatic proxies and soil water indices for two important European tree species with different water tolerances, *Q. robur*, a temperate oak, and *Q. pubescens*, a Mediterranean oak. *Quercus robur* is the most common meso-hygrophilous species in the studied area (Table 1). This species is distributed over a large range of soil nutritional conditions and is sensitive to summer drought (Timbal and Aussenac, 1996). It can be found over large areas of France other than the Mediterranean basin, and in forests up to 1300 m. *Quercus pubescens* is the most common meso-xerophilous species in France (Table 1). Because it tolerates a warm climate, it occupies a large portion of the southern territory, up to 1400 m in mountainous areas (Timbal and Aussenac, 1996). For each species, we selected the best climatic proxy—P, CWB_w, or CWB_{<0}—and the best soil water balance index—SWC or SWD—to improve the model based on GDD₅ and minWT. The performances of the models, water response curve shapes, and distribution maps were examined. To compare the predictions of the models with the observed distributions, we divided the continuous probability variable for presence in the presence-absence maps using a threshold defined by the point on the receiver operating characteristic curve (ROC curve) that had the shortest distance to the top left corner (Fielding and Bell, 1997).

RESULTS

Mapping and comparison of water availability indices

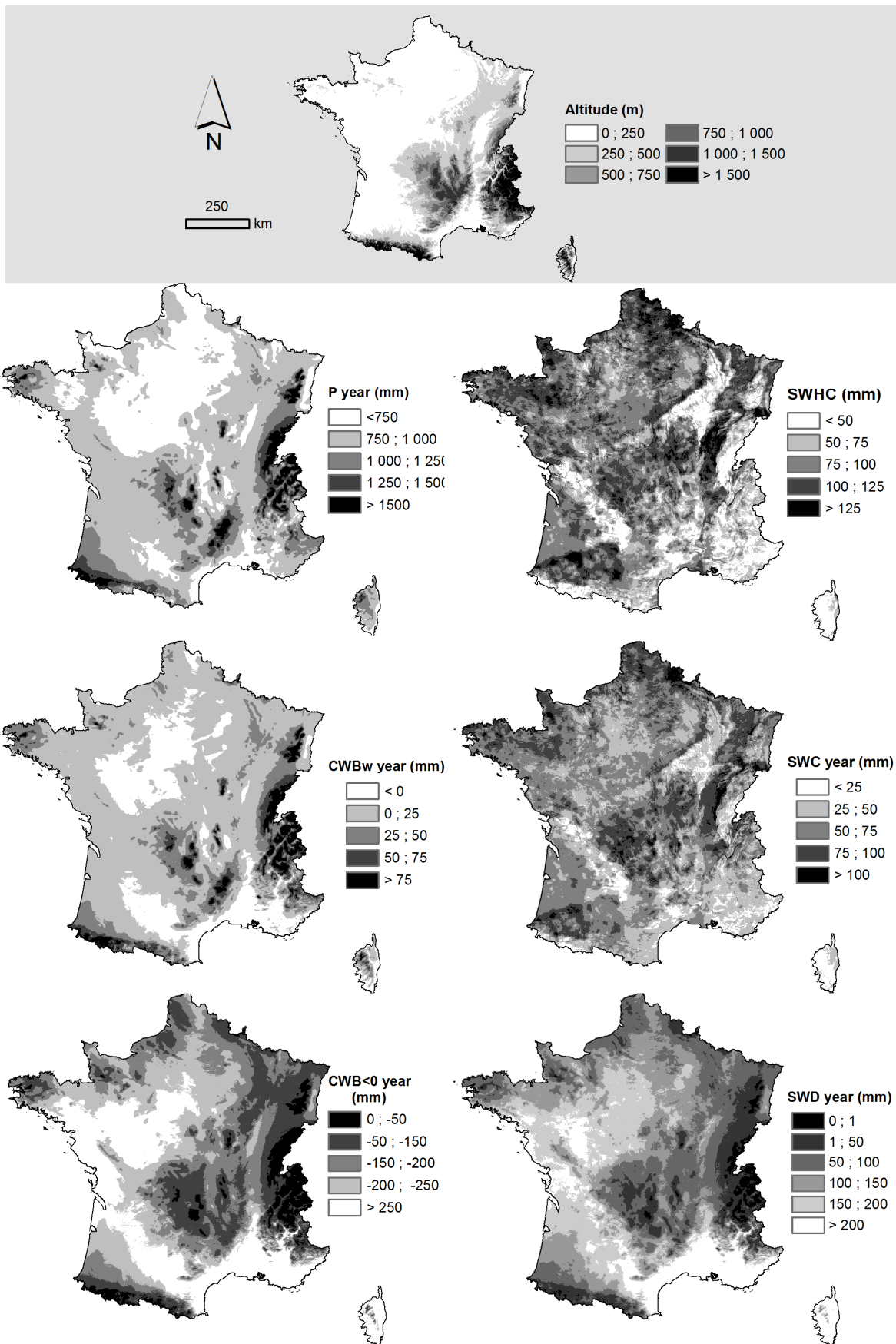


Fig. 1: Maps of the studied water availability indices. Top: spatial distribution of altitude; left: 1961–1990 mean annual values for precipitation (P), and climatic water balance calculated using all values (CWB_w) and only negative values (CWB<0); right: soil water-holding capacity (SWHC), and 1961–1990 mean annual values for soil water content (SWC) and deficit (SWD) derived from the soil water balance.

The studied water availability indices were calculated for the 1961–1990 period and mapped across France (Fig. 1). For annual values extracted from the map, SWD was better correlated than SWC with estimates from the plots (Table 3A). The various map-estimated water availability indices were not highly correlated with GDD₅ or minWT (Table 3B), the higher Pearson’s *r* reaching –0.74 between CWB_{<0} and GDD₅. P and CWB_w had a similar pattern overall (*r* = 0.96), with higher water content found in mountainous areas and lower content in the Mediterranean area and Parisian basin. CWB_{<0} is logically more strongly linked to CWB_w (*r* = 0.82) than to P (*r* = 0.67) owing to higher PET values in the southern and western parts of the country, indicating increased water deficits in these regions compared to those in mountainous areas and northern regions.

Map-estimated variables						
A	P _{year}	CWB _w year	CWB _{<0} year	SWHC	SWC _{year}	SWD _{year}
SWHC _{plot}	0.19	0.16	0.07	0.61	0.56	0.07
SWC _{year plot}	0.03	0.01	0.14	0.57	0.60	0.30
SWD _{year plot}	0.35	0.51	0.57	0.30	0.47	0.91

B	Climatic water balance			Soil water balance			
	GDD ₅	MinWT	P _{year}	CWB _w year	CWB _{<0} year	SWC _{year}	SWD _{year}
SWHC	0.12	0.22	-0.28	-0.23	-0.02	0.78	-0.19
GDD ₅		0.77	-0.41	-0.61	-0.74	-0.08	0.61
MinWT			-0.45	-0.57	-0.66	0.01	0.52
P _{year}				0.96	0.67	0.01	-0.47
CWB _w year					0.82	0.06	-0.70
CWB _{<0} year						0.23	-0.96
SWC _{year}							-0.47

Table 3: Pearson’s coefficients of correlation (*r*) between the various indices used in the plant distribution models (*n* = 32,828). A: Correlations between plot-based and map-estimated water variables. B: Correlations between the various map-estimated indices used in the plant distribution models. SWHC, soil water-holding capacity; GDD₅, growing degree days above 5°C; minWT, minimum winter temperature; P, precipitation; CWB_w and CWB_{<0}, climatic water balance using all values and only negative values, respectively; SWC and SWD, Thornthwaite soil water content and deficit, respectively.

The annual SWC map was, as expected, correlated with SWHC (*r* = 0.78), but surprisingly not with the annual mean climatic proxies (*r* = 0.01–0.23 with P, CWB_w, and CWB_{<0}; Table 3B). The greatest amount of water available to plants is located in the sedimentary plains, with less available in mountainous areas, in calcareous plateaus, and in southern France (Fig. 1). We observed different spatial distributions for water inputs and SWHC. High SWHC values were found in areas with relatively dry and warm climates in which the water stored during winter provides a reserve for plants over much of the year. In contrast, the wettest areas are mainly located in the mountains where the SWHC is generally lower, with moderate SWC. Thus, the lower P in the lowlands compared with that in the mountains can be compensated for by deeper soils. On a more local scale, the spatial variability of SWC is greater than that of the climatic water indices because SWHC is strongly influenced by geology. The annual water deficits calculated using soil (SWD) and climate (CWB_{<0}) show similar patterns for France overall (*r* = –0.96; Table 3) despite water deficits being estimated as lower in sedimentary plains using SWD than those estimated using CWB_{<0} (Fig. 1).

Available water estimated using the mapped indices differs with the season (Fig. 2). The mean P in France is roughly constant throughout the year despite small summertime decreases, whereas the CWB indices show substantial losses during this period because the PET is elevated. When the SWHC is high, the water stored in the soil provides an important complementary reserve, limiting SWD even in dry climates. In contrast, the interaction between low SWHC and low P decreases water availability for both the climatic and the soil water balance indices.

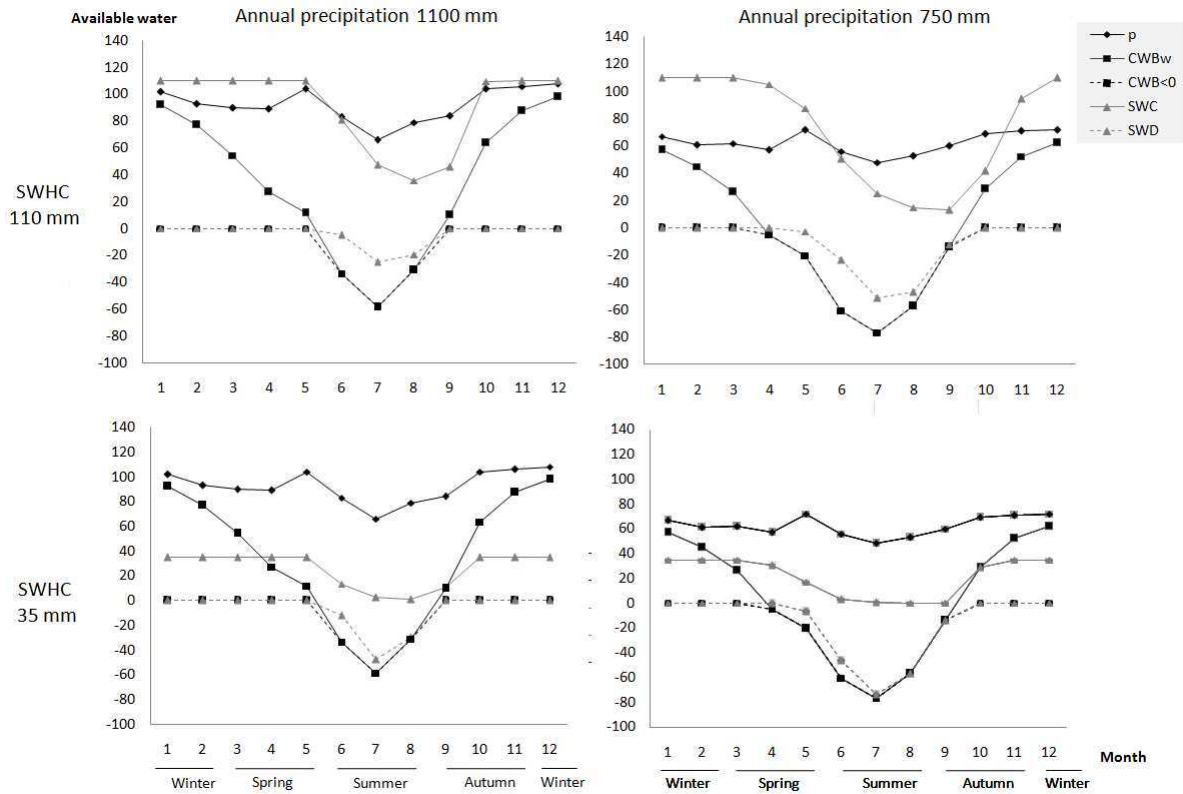


Fig. 2: Monthly values of precipitation, climatic indices, and soil water balance indices for contrasting conditions of soil water-holding capacity (SWHC) and annual precipitation (P). Climatic water balance was characterised by indices based on all values (CWB_w) and only the negative values ($CWB_{<0>}$), whereas soil water balance was estimated through soil water content (SWC) or deficit (SWD). Values of SWHC and P correspond to the first and ninth decile of the observed values ($n = 32,828$), and mean monthly potential evapotranspiration values for France were used in the water balance calculations.

Performance of climatic proxies versus soil water balance indices in predicting plant distribution

We compared the performances of the most efficient climatic proxies (among P, CWB_w , and $CWB_{<0>}$) and the most efficient soil water balance indices (between SWC and SWD) in improving distribution models based on GDD_5 and minWT for the 37 studied species (Fig. 3). Models employing the soil water reserve were statistically more accurate than those using water-related climatic variables (mean difference in AUC, +0.019; $n = 37$, Student's paired t-test, $t = 3.68$, degrees of freedom = 36; $P = 0.0008$). The mean AUC was 0.76 when only the GDD_5 and minWT values were taken into account, increasing to 0.80 when the best water-related climatic proxy was added or to 0.82 when the best soil water balance index was added. SWC or SWD performed better than climatic proxies for 14 of the 16 species whose AUC increased by a value of >0.05 (Fig. 3). For the species whose distributions were best predicted by soil water balance, SWC and SWD performed similarly. For the remaining species, P performed better than CWB or $CWB_{<0>}$. Despite a strong correlation between $CWB_{<0>}$ and

SWD for the annual values, seasonal $CWB_{<0}$ values were always less efficient than SWD values in the tree species distribution models. These results can probably be attributed to stronger correlations between $CWB_{<0}$ and GDD_5 and $minWT$ than SWD had (Table 3B), and the greater predictive ability of spring and autumn SWD values compared to $CWB_{<0}$ (see Appendix S1 in Supporting Information).

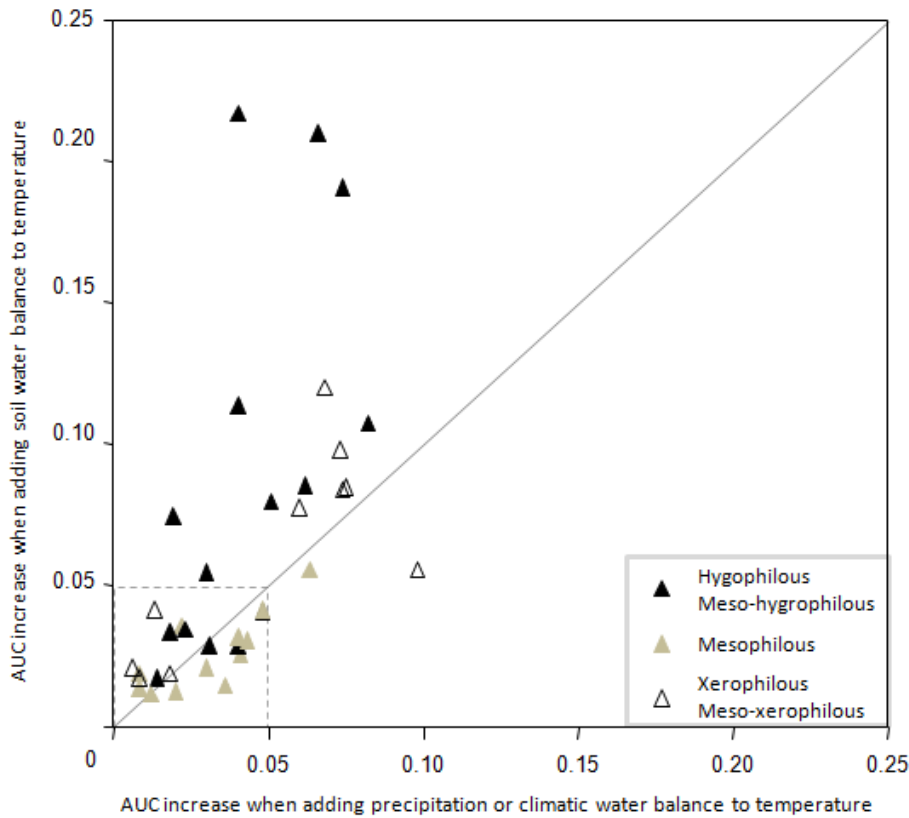


Fig. 3: Area under the curve (AUC) increases for species distribution models determined using the most efficient index among either climatic proxies or soil water balance, when added to growing degree days above 5°C and minimum winter temperature ($n = 37$ species). Climatic proxies were represented by precipitation and climatic water balance based on all values or only the negative values (CWB_w and $CWB_{<0}$), whereas soil water balance was characterised by soil water content or deficit (SWC and SWD). Spring, summer, and autumn values were used for all of the indices. The mean AUC increase for models including climatic proxies was 0.041, whereas that for models including soil water balance was 0.060.

We examined the predictive capabilities of the various water availability indices for species classified into indicator groups according to drought tolerance. Soil water balance produced better models than those using climatic proxy values for most hygrophilous, meso-hygrophilous, meso-xerophilous, and xerophilous species (Fig. 3, Table 4), both when the predictors were considered alone and when they were used in addition to GDD_5 and $minWT$ (Table 4). Modelling using TSS produced similar results (see Appendix S2 in Supporting Information). Models using only temperature information performed less robustly than did those based on soil water balance indices (SWC or SWD) for hygrophilous, meso-hygrophilous, and xerophilous species (Table 4A). In contrast, temperature-related variables explained much of the distributions of mesophilous and meso-xerophilous species, and simple climatic proxies and soil water indices performed nearly equally in characterizing available water for plants.

A					Climatic	Soil	% Species improved	
	n	SWHC	GDD ₅	MinWT	P	water balance	water balance	by soil water balance
Hygrophilous	7	0.63	0.63	0.64	0.66	0.66	0.74	71
Meso-hygrophilous	7	0.63	0.65	0.68	0.68	0.67	0.76	71
Mesophilous	12	0.61	0.73	0.74	0.71	0.70	0.70	33
Meso-xerophilous	7	0.73	0.80	0.82	0.80	0.81	0.83	57
Xerophilous	4	0.76	0.74	0.68	0.80	0.77	0.83	75

B					T + climatic	T + soil	% Species improved
	n	T	T + P		water balance	water balance	by soil water balance
Hygrophilous	7	0.68	0.73		0.72	0.78	71
Meso-hygrophilous	7	0.72	0.74		0.74	0.79	100
Mesophilous	12	0.77	0.80		0.79	0.80	25
Meso-xerophilous	7	0.86	0.89		0.89	0.90	71
Xerophilous	4	0.77	0.85		0.82	0.86	75

Table 4: Performance of water availability indices in predicting plant distributions of species grouped by drought tolerance (n = 37). A: Average areas under the curve (AUCs) for models using soil water-holding capacity (SWHC), growing degree days above 5°C (GDD₅), minimum winter temperature (minWT), precipitation (P), climatic water balance estimated using all values (CWB_w) or only the negative values (CWB₋₀), and soil water balance represented by soil water content (SWC) or deficit (SWD). B: AUCs for models using temperature alone (T, GDD₅, and minWT used together), and with the addition of P, climatic variables, or soil water balance. The most efficient index for spring, summer, and autumn values was used for all the water availability indices. The percentage (%) of species improved by soil water balance is the percentage of the models for which soil water balance indices are more efficient than P or climatic water balance indices in improving the AUCs of the models, when these indices are used alone (A) or combined with T (B). n, number of species.

Comparison of distribution models using climatic proxies or soil water for *Q. robur* and *Q. pubescens*

Summer CWB_w and SWC were the most efficient climatic proxy and soil water balance index, respectively, for improving the *Q. robur* and *Q. pubescens* distribution models based on GDD₅ and minWT. The AUC improved by 0.05 and 0.06 using CWB_w and by 0.08 and 0.09 using SWC for *Q. robur* and *Q. pubescens*, respectively. The use of soil water balance indices rather than climatic proxies improved the performances of the models and produced important differences in the water response curve shape (Fig. 4). The choice between CWB_w and SWC also strongly modified the spatial distribution predictions for the species, as illustrated for a region of contrasting geomorphological units in southwestern France (Fig. 5). *Q. robur* was observed in field surveys over most of this area, except on a calcareous plateau extending from the southeast to the northwest (Fig. 5A); *Q. pubescens* showed the opposite pattern, being absent over most of the area other than on the calcareous plateau. The models using CWB_w predicted the presence of *Q. robur* and the absence of *Q. pubescens* over the entire studied area (Fig. 5B). After substituting summer SWC for CWB_w, the models highlighted the absence of *Q. robur* and the presence of *Q. pubescens* on the calcareous plateau (Fig. 5C). The observed distributions of both species were clearly better linked to soil water availability than to climatic indices, available soil water being a key factor in discriminating the distributions of these species in this area.

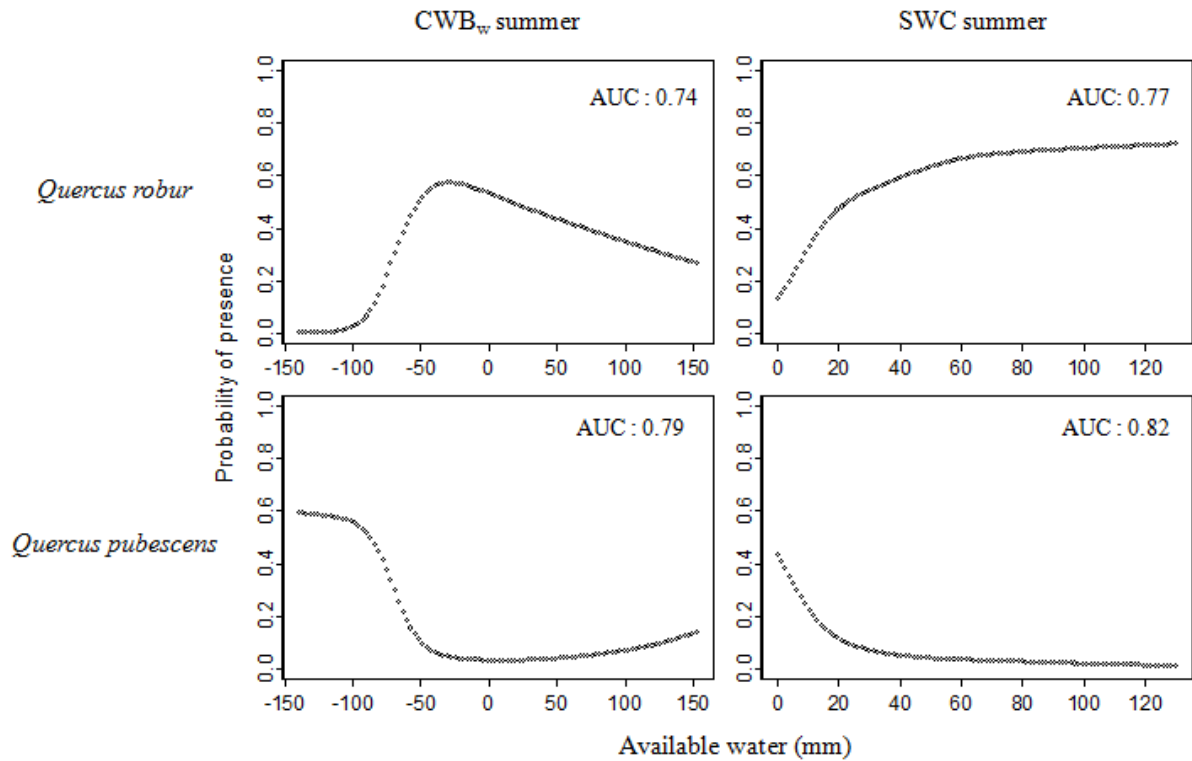


Fig. 4: Probability of the presence of *Quercus robur* and *Quercus pubescens* and areas under the curve (AUCs) for summer climatic water balance calculated using all values (CWB_w summer, mm) or soil water content (SWC summer, mm) to complement growing degree days above 5°C and minimum winter temperature.

DISCUSSION

Soil water balance is a physiologically meaningful predictor, but its calculation requires information that can be difficult to obtain. As a result, plant ecologists mainly use simple climatic water indices and variables that are more readily available. Here, we showed that these simpler variables can be inadequate for representing the soil water available for plants. We also demonstrated that they perform less robustly than do soil water balance values in explaining the distribution of most European tree species. Our results clearly show that simple climatic proxies can lead to underestimates of the importance of water in plant distribution studies, a conclusion that agrees with numerous studies demonstrating the key role of soil water balance in tree physiology and growth (Lebourgeois *et al.*, 2005; Granier *et al.*, 2007). Thus, our study provides new insights into the importance of soil water balance modelling for accurate estimation of plant distribution. These results complement temperate forest research identifying the importance of soil moisture in shaping plant functional types on a continent scale and individual species on a local scale in tropical forests (Stephenson, 1990; Engelbrecht *et al.*, 2007).

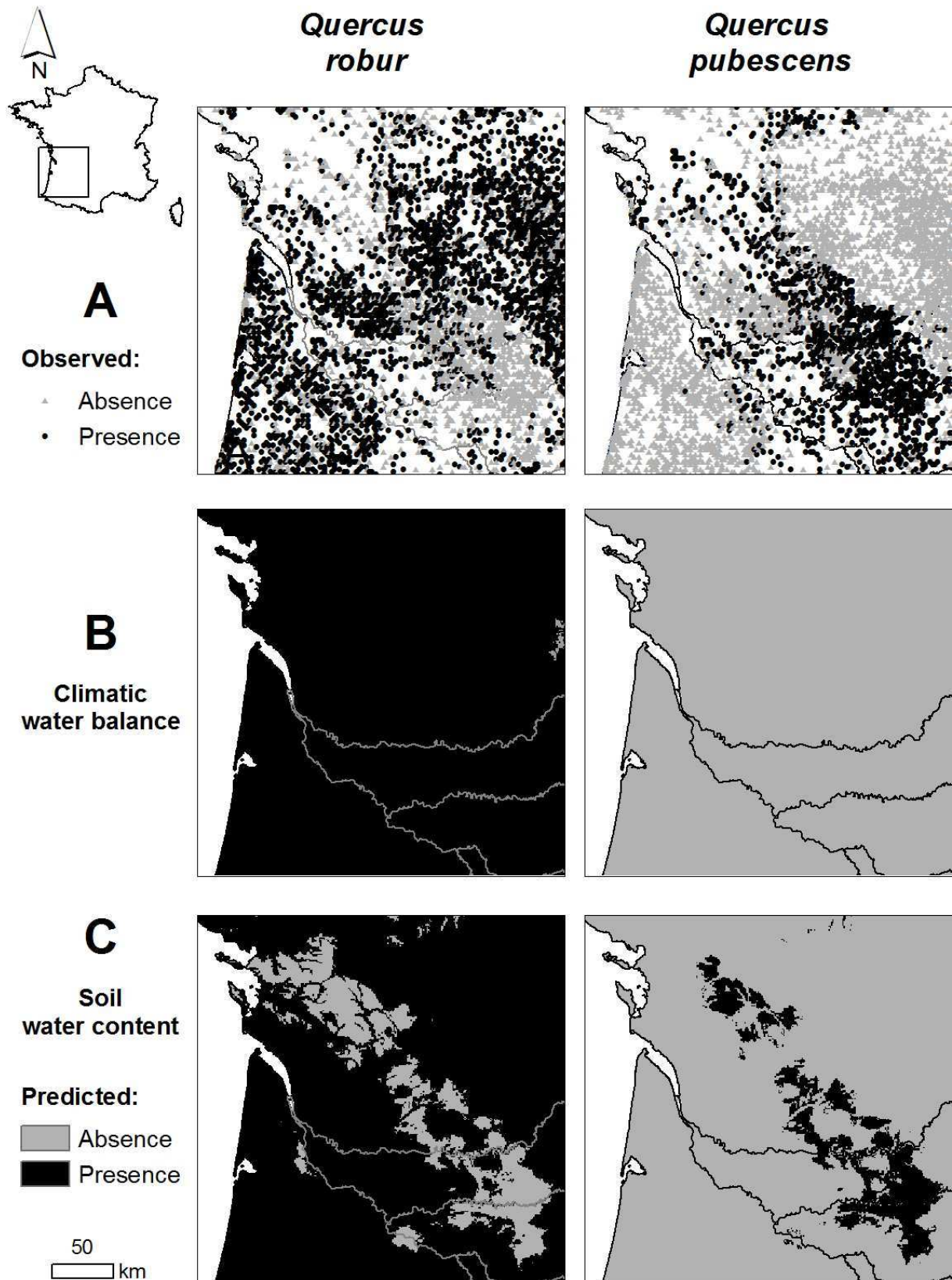


Fig. 5: Observed distributions (A) and maps predicting the distributions of *Quercus robur* and *Quercus pubescens* in southwestern France. Predictions were made using summer climatic water balance calculated using all values (CWB_w ; B) or summer soil water content (SWC; C) combined with growing degree days above 5°C and minimum winter temperature. Presence-absence maps were created from continuous probabilities of presence using the most efficient thresholds for each model (0.39 and 0.41 for summer CWB_w or summer SWC, respectively, for *Q. robur*, and 0.26 and 0.20 for summer CWB_w or summer SWC, respectively, for *Q. pubescens*).

Although poor relationships between climatic water variables and SWC are expected, to our knowledge, no comparison of these two approaches has been conducted over broad areas. Water modelling based on climatic predictors shows important differences in geographical patterns and drought intensity compared with results based on SWC. Water stored in soil acts as a reserve that limits losses related to summer PET, highlighting fine-scale variations due to geomorphological attributes, as has been observed in local-scale studies based on field data (Lookingbill and Urban, 2004; Dyer, 2009). In most temperate countries, lowlands are warmer than highlands, because evaporative demand is higher and P is generally lower. Most lowlands are also sedimentary plains with substantial SWHC, however, and use of climatic water indices in these areas can greatly overestimate water stress. Conversely, P does not always translate to significant water reserves for plants if the soil is too shallow to store water. This relationship has been demonstrated for both a wide range of soils (from podzols in siliceous substrates to calcareous soils) and climates (with mean annual P of 550–2500 mm and mean annual temperatures of 5–15°C); thus, our findings are relevant for most ecological conditions in temperate latitudes.

Simple climatic water variables remain better indices than soil water balance variables for predicting the distributions of 35% of the studied tree species. These species were mainly mesophilous, and their distributions are less influenced by water availability than are those of the other studied trees. Their range appeared to be influenced mainly by temperature, whereas the other studied variables had little influence on their distributions, allowing soil or climatic water proxies to be equally used in the models. Conversely, we showed that water availability is probably a more important determinant than temperature in determining the distributions of continental tree species living at the extremes of the water balance gradient. For more water-dependent species, soil water balance performs better than climatic proxies, as illustrated by *Q. robur* and *Q. pubescens* distributions in southwestern France. For plants located in the warmest areas, the greater effect of available water compared to other climatic predictors has previously been reported for plant richness (Hawkins *et al.*, 2003).

Soil water balance improves the performance of the models by refining the prediction of the spatial distributions of species compared to the predictions made using simple climatic water proxies. Soil water balance allows differentiation of the probabilities of the presence of a species in areas with similar climatic conditions—predicting, for example, the absence of a drought-sensitive species when the soil water reserve is low. Conversely, drought-resistant species can be excluded from areas with deeper soils despite dry climatic conditions, probably owing to increased competition from species that require more water. The observed differences in sensitivities between species provide new insights into plant distribution responses to water resources. The resulting response curves allow determination of thresholds for the presence of a species relative to available water, refining empirical plant functional classifications (Ellenberg *et al.*, 1992) and complementing ecophysiological studies of the drought tolerance of forest tree species (Bréda *et al.*, 2006). Further investigations are nevertheless necessary to compare species responses to water stress using plant distribution modelling and an ecophysiological approach.

In this study, we used the AUC to measure the performance of the various models. Several authors have pointed out limitations on the accuracy of such measures when applied to distribution modelling. For the AUC, these limitations are, mainly, dependence on the geographical extent and number of predictors used; problems increase when pseudo-absences are generated (Lobo *et al.*, 2008). These drawbacks are important when the AUC is used to compare the performance of models with different species, particularly when the ratio of the extent of occurrence to the size of the studied territory differs between species. In this study, the AUC was used to compare the predictive abilities of the studied water indices for the same species, without variations in extent or number of predictors and using phytocological surveys with recorded presence and absence. Thus, the

differences observed between the various predictors cannot be attributed to the reported limitations of the AUC. Moreover, the results obtained using the AUC were corroborated by similar results using TSS.

Data from the soil water balance maps were processed using simple formulas at a monthly scale. SWHC introduces an important source of uncertainty compared to climatic information owing to the difficulty of surveying soil parameters and the substantial spatial variability of SWHC on a local scale (Piedallu *et al.*, 2011). To estimate SHWC, we included soil depths of up to 1 m based on studies showing that 95% of the roots are present within the top meter (Jackson *et al.*, 1996) and because the site index prediction was not improved by using SWHC estimated for soil profiles exceeding 1 m (Berges and Balandier, 2010). The importance of water uptake by the small fraction of fine roots located deeper in the soil is well known and influences tree resistance to drought (Breda *et al.*, 1995). The use of 1 m soil depth probably led to consistent underestimates of SWC owing to the difficulty of determining the volume of soil available to store water for trees, particularly in fragmented bedrock or areas with rock outcroppings. Moreover, snow, lateral runoff, and percolation effects were excluded from the water balance calculation, despite the likelihood that taking them into account would improve the quality of the estimates (Redding and Devito, 2008; Viviroli *et al.*, 2009). Improvements in soil water balance mapping may provide greater insights into the importance of soil water variables in future distribution models.

The development of efficient approaches to estimating available water over broad areas is a major step in evaluating the effects of water stress on plants. The use of soil water maps improves our understanding of the mechanisms related to plant distribution patterns and allows better determination of the ecological niches of species. Soil water balance indices can be averaged over long periods to characterize mean conditions for a site, or estimated over shorter periods to determine the effects of extreme conditions. In the current context of climate change, use of relevant predictors is crucial for evaluating changes in available water for plants and consequences in terms of vegetation shifts. These predictors can also contribute to identification of potential refugia or opportunities for local expansion in the context of species migration (Keppel *et al.*, 2012), aid in the selection of alternative tree species in forestry, and identify areas that are potentially more sensitive to changes in tree productivity or mortality.

Acknowledgements

This study was supported by the GIP ECOFOR, the RMT AFORCE, the Région Lorraine, and the Direction Régionale de l'Alimentation, de l'Agriculture et de la Forêt de la région Lorraine. We also thank the referees for their helpful review of the manuscript.

References

- Allouche, O., Tsoar, A., Kadmon, R. (2006) Assessing the accuracy of species distribution models: prevalence, kappa and the true skill statistic (TSS). *Journal of Applied Ecology*, **43**, 1223-1232.
- Austin, M.P. (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling*, **157**, 101-118.
- Benichou, P., Le Breton, O. (1987) Prise en compte de la topographie pour la cartographie des champs pluviométriques statistiques. *La Météorologie*, **7**, 23-24.
- Berges, L., Balandier, P. (2010) Revisiting the use of soil water budget assessment to predict site productivity of sessile oak (*Quercus petraea* Liebl.) in the perspective of climate change. *European Journal of Forest Research*, **129**, 199-208.
- Beven, K. (2001) How far can we go in distributed hydrological modelling? *Hydrology and Earth System Sciences*, **5**, 1-12.
- Bréda, N., Granier, A., Barataud, F., Moyne, C. (1995) Soil-water dynamics in an oak stand. 1. Soil-moisture, water potentials and water-uptake by roots. *Plant and Soil*, **172**, 17-27.
- Bréda, N., Huc, R., Granier, A., Dreyer, E. (2006) Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences. *Annals of Forest Science*, **63**, 625-644.
- Crimmins, S.M., Dobrowski, S.Z., Greenberg, J.A., Abatzoglou, J.T., Mynsberge, A.R. (2011) Changes in climatic water balance drive downhill shifts in plant species' optimum elevations. *Science*, **331**, 324-327.
- Davi, H., Dufrêne, E., Granier, A., Le Dantec, V., Barbaroux, C., François, C., Bréda, N. (2005) Modelling carbon and water cycles in a beech forest Part II: Validation of the main processes from organ to stand scale. *Ecological Modelling*, **185**, 387-405.
- Drapier, J., Cluzeau, C. (2001) La base de données écologiques de l'IFN. The NFI's ecological data base. *Revue Forestière Française*, **53**, 365-371.
- Dyer, J.M. (2009) Assessing topographic patterns in moisture use and stress using a water balance approach. *Landscape Ecology*, **24**, 391-403.
- Ellenberg, H., Weber, H.E., Düll, R., Wirth, V., Werner, W., Paulißen, D. (1992) Zeigerwerte von Pflanzen in Mitteleuropa. *Scripta geobotanica*, **18**, 1-248.
- Engelbrecht, B.M.J., Comita, L.S., Condit, R., Kursar, T.A., Tyree, M.T., Turner, B.L., Hubbell, S.P. (2007) Drought sensitivity shapes species distribution patterns in tropical forests. *Nature*, **447**, 80-U82.
- Fielding, A.H., Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38-49.
- Granier, A., Breda, N., Biron, P., Villette, S. (1999) A lumped water balance model to evaluate duration and intensity of drought constraints in forest stands. *Ecological Modelling*, **116**, 269-283.
- Granier, A., Reichstein, M., Breda, N., Janssens, I.A., Falge, E., Ciais, P., Grunwald, T., Aubinet, M., Berbigier, P., Bernhofer, C., Buchmann, N., Facini, O., Grassi, G., Heinesch, B., Ilvesniemi, H., Keronen, P., Knohl, A., Kostner, B., Lagergren, F., Lindroth, A., Longdoz, B., Loustau, D., Mateus, J., Montagnani, L., Nys, C., Moors, E., Papale, D., Peiffer, M., Pilegaard, K., Pita, G., Pumpanen, J., Rambal, S., Rebmann, C., Rodrigues, A., Seufert, G., Tenhunen, J., Vesala, I., Wang, Q. (2007) Evidence for soil water control on carbon and water dynamics in European forests during the extremely dry year: 2003. *Agricultural and Forest Meteorology*, **143**, 123-145.
- Guisan, A., Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993-1009.
- Hastie, T.J., Tibshirani, R. (1990) *Generalized additive models*. Chapman and Hall, London.
- Hawkins, B.A., Field, R., Cornell, H.V., Currie, D.J., Guegan, J.F., Kaufman, D.M., Kerr, J.T., Mittelbach, G.G., Oberdorff, T., O'Brien, E.M., Porter, E.E., Turner, J.R.G. (2003) Energy, water, and broad-scale geographic patterns of species richness. *Ecology*, **84**, 3105-3117.

- Holsten, A., Vetter, T., Vohland, K., Krysanova, V. (2009) Impact of climate change on soil moisture dynamics in Brandenburg with a focus on nature conservation areas. *Ecological Modelling*, **220**, 2076-2087.
- Jackson, R.B., Canadell, J., Ehleringer, J.R., Mooney, H.A., Sala, O.E., Schulze, E.D. (1996) A global analysis of root distributions for terrestrial biomes. *Oecologia*, **108**, 389-411.
- Julve, P. (1998) Baseflor. *Index botanique, écologique et chorologique de la flore de France*. <http://perso.wanadoo.fr/philippe.julve/catminat.htm>.
- Keppel, G., Van Niel, K., Wardell-Johnson, G., Yates, C., Byrne, M., Mucina, L., Schut, A., Hopper, S., Franklin, S. (2012) Refugia: identifying and understanding safe havens for biodiversity under climate change. *Global Ecology and Biogeography*, **21**, 393-404.
- Kerkhoff, A.J., Martens, S.N., Shore, G.A., Milne, B.T. (2004) Contingent effects of water balance variation on tree cover density in semiarid woodlands. *Global Ecology and Biogeography*, **13**, 237-246.
- Lebourgeois, F., Bréda, N., Ulrich, E., Granier, A. (2005) Climate-tree-growth relationships of European beech (*Fagus sylvatica* L.) in the French Permanent Plot Network (RENECOFOR). *Trees: Structure and Function*, **19**, 385-401.
- Lobo, J.M., Jimenez-Valverde, A., Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145-151.
- Lookingbill, T., Urban, D. (2004) An empirical approach towards improved spatial estimates of soil moisture for vegetation analysis. *Landscape Ecology*, **19**, 417-433.
- Lu, J.B., Sun, G., McNulty, S.G., Amatya, D.M. (2005) A comparison of six potential evapotranspiration methods for regional use in the southeastern United States. *Journal of the American Water Resources Association*, **41**, 621-633.
- Lutz, J.A., van Wagtenonk, J.W., Franklin, J.F. (2010) Climatic water deficit, tree species ranges, and climate change in Yosemite National Park. *Journal of Biogeography*, **37**, 936-950.
- Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K., Thuiller, W. (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, **15**, 59-69.
- Ninyerola, M., Pons, X., Roure, J.M. (2000) A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology*, **20**, 1823-1841.
- Piedallu, C., Gégout, J.C. (2007) Multiscale computation of solar radiation for predictive vegetation modelling. *Annals of Forest Science*, **64**, 899-909.
- Piedallu, C., Gégout, J.C., Bruand, A., Seynave, I. (2011) Mapping soil water holding capacity over large areas to predict potential production of forest stands. *Geoderma*, **160**, 355-366.
- Platts, P.J., Ahrends, A., Gereau, R.E., McClean, C.J., Lovett, J.C., Marshall, A.R., Pellikka, P.K.E., Mulligan, M., Fanning, E., Marchant, R. (2010) Can distribution models help refine inventory-based estimates of conservation priority? A case study in the Eastern Arc forests of Tanzania and Kenya. *Diversity and Distributions*, **16**, 628-642.
- Prentice, I.C., Cramer, W., Harrison, S.P., Leemans, R., Monserud, R.A., Solomon, A.M. (1992) A global biome model based on plant physiology and dominance, soil properties and climate. *Journal of Biogeography*, **19**, 117-134.
- Rameau, J.C., Mansion, D., Dumé, G. (1989) *Flore forestière française - guide écologique illustré - tome 1: plaines et collines*. Institut pour le Développement, Paris, 1785 p.
- Redding, T.E., Devito, K.J. (2008) Lateral flow thresholds for aspen forested hillslopes on the Western Boreal Plain, Alberta, Canada. *Hydrological Processes*, **22**, 4287-4300.
- Reichstein, M., Ciais, P., Papale, D., Valentini, R., Running, S., Viovy, N., Cramer, W., Granier, A., Ogee, J., Allard, V., Aubinet, M., Bernhofer, C., Buchmann, N., Carrara, A., Grunwald, T., Heimann, M., Heinesch, B., Knohl, A., Kutsch, W., Loustau, D., Manca, G., Matteucci, G., Miglietta, F., Ourcival, J.M., Pilegaard, K., Pumpanen, J., Rambal, S., Schaphoff, S., Seufert, G., Soussana, J.F., Sanz, M.J., Vesala, T., Zhao, M. (2007) Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: a joint flux tower, remote sensing and modelling analysis. *Global Change Biology*, **13**, 634-651.

- Schwarzal, K., Feger, K.H., Hantzschel, J., Menzer, A., Spank, U., Clausnitzer, F., Kostner, B., Bernhofer, C. (2009) A novel approach in model-based mapping of soil water conditions at forest sites. *Forest Ecology and Management*, **258**, 2163-2174.
- Siles, G., Alcantara, J.M., Rey, P.J., Bastida, J.M. (2010) Defining a target map of native species assemblages for restoration. *Restoration Ecology*, **18**, 439-448.
- Stephenson, N.L. (1990) Climatic control of vegetation distribution: the role of the water balance. *The American Naturalist*, **135**, 649-670.
- Synes, N.W., Osborne, P.E. (2011) Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. *Global Ecology and Biogeography*, **20**, 904-914.
- Swets, J.A. (1988) Measuring the accuracy of diagnostic systems. *Science*, **240**, 1285-1293.
- Thornthwaite, C.W., Mather, J.R. (1955) *The water balance*. In: *Publications in Climatology VIII*, Drexel Institute of Climatology Laboratory, Centerton NJ, pp. 1-104.
- Timbal, J., Aussenac, G. (1996) An overview of ecology and silviculture of indigenous oaks in France. *Annals of Forest Science*, **53**, 649-661.
- Turc, L. (1961) Evaluation des besoins en eau d'irrigation et évaporation potentielle. *Annales agronomiques*, **12**, 13-49.
- Van Mantgem, P.J., Stephenson, N.L., Byrne, J.C., Daniels, L.D., Franklin, J.F., Fule, P.Z., Harmon, M.E., Larson, A.J., Smith, J.M., Taylor, A.H., Veblen, T.T. (2009) Widespread increase of tree mortality rates in the western United States. *Science*, **323**, 521-524.
- Viviroli, D., Mittelbach, H., Gurtz, J., Weingartner, R. (2009) Continuous simulation for flood estimation in ungauged mesoscale catchments of Switzerland. Part II: Parameter regionalisation and flood estimation results. *Journal of Hydrology*, **377**, 208-225.
- Zhang, L., Walker, G.R., Dawes, W.R. (2002) *Water balance modelling: concepts and applications*. in: McVicar, T., Rui, L., Walker, J., Fitzpatrick, R.W., and Liu, C. (eds), *Regional water and soil assessment for managing sustainable agriculture in China and Australia*, (ACIAR Monograph; 84), Canberra :ACIAR: 31-47.

Biosketch

Christian Piedallu is a GIS engineer at the *Laboratoire d'Étude des Ressources Forêt-Bois of AgroParisTech*. His main research interests include derivation of high-resolution ecological variables and indices and their testing in plant species predictive distribution models.