



# A novel method for assessing climate change impacts in ecotron experiments

Inne Vanderkelen<sup>1</sup> · Jakob Zscheischler<sup>2,3</sup> · Lukas Gudmundsson<sup>4</sup> · Klaus Keuler<sup>5</sup> · Francois Rineau<sup>6</sup> · Natalie Beenaerts<sup>6</sup> · Jaco Vangronsveld<sup>6,7</sup> · Sara Vicca<sup>8</sup> · Wim Thiery<sup>1,4</sup>

Received: 27 March 2020 / Accepted: 4 June 2020 / Published online: 27 July 2020  
© The Author(s) 2020

## Abstract

Ecotron facilities allow accurate control of many environmental variables coupled with extensive monitoring of ecosystem processes. They therefore require multivariate perturbation of climate variables, close to what is observed in the field and projections for the future. Here, we present a new method for creating realistic climate forcing for manipulation experiments and apply it to the UHasselt Ecotron experiment. The new methodology uses data derived from the best available regional climate model projection and consists of generating climate forcing along a gradient representative of increasingly high global mean air temperature anomalies. We first identified the best-performing regional climate model simulation for the ecotron site from the Coordinated Regional Downscaling Experiment in the European domain (EURO-CORDEX) ensemble based on two criteria: (i) highest skill compared to observations from a nearby weather station and (ii) representativeness of the multi-model mean in future projections. The time window is subsequently selected from the model projection for each ecotron unit based on the global mean air temperature of the driving global climate model. The ecotron units are forced with 3-hourly output from the projections of the 5-year period in which the global mean air temperature crosses the predefined values. With the new approach, Ecotron facilities become able to assess ecosystem responses on changing climatic conditions, while accounting for the co-variation between climatic variables and their projection in variability, well representing possible compound events. The presented methodology can also be applied to other manipulation experiments, aiming at investigating ecosystem responses to realistic future climate change.

**Keywords** Regional climate model · Climate forcing · Controlled environment experiment · Global warming · Ecosystem response

## Introduction

Ecosystem climate change experiments are one of the key instruments to study the response of ecosystems to a change in climate. There are primarily four different factors that are altered in such experiments: air temperature, precipitation, CO<sub>2</sub> concentration, and nitrogen deposition (Curtis and Wang 1998; Rustad et al. 2001; Lin et al. 2010; Wu et al. 2011; Knapp et al. 2018). More recently, multi-factor

experiments are starting to emerge. In those experiments, different combinations of the four main drivers are altered (Kardol et al. 2012; Yue et al. 2017). What is common in the majority of climate change experiments is that while the drivers of interest are being altered, all other variables are being held equal between the different treatment groups. Consequently, differences in the response can be related to the change in the main driving factor (or multiple driving factors).

Altering only one or a limited number of climate change drivers allows for a straightforward analysis of the observed responses and has provided a wealth of mechanistic insights in ecosystem responses to environmental changes (e.g., Hovenden et al. 2014; Karlowsky et al. (Karlowsky et al. 2018; Terrer et al. 2018)). However, the resulting multivariate combination of climate variables may be physically unrealistic and may miss key aspects related to natural climate variability and the co-variance of multiple

---

**Electronic supplementary material** The online version of this article (<http://dx.doi.org/10.1007/s00484-020-01951-8>) contains supplementary material, which is available to authorized users.

---

✉ Inne Vanderkelen  
inne.vanderkelen@vub.be

Extended author information available on the last page of the article.

variables, linked to each other by synoptic conditions. This is particularly important for representing compound events, where the combination of non extreme drivers can lead to extreme impacts (Zscheischler and Seneviratne 2017; Zscheischler et al. 2018; Rineau et al. 2019). For example, droughts and heatwaves often co-occur (Zscheischler and Seneviratne 2017), and soil moisture conditions and precipitation occurrence are linked (Guilod et al. 2015; Moon et al. 2019). Incorporating the co-variability of key climate drivers is also important for the studied responses. For instance, heatwaves characterized by similar extreme air temperature can lead to different plant responses depending on the atmospheric conditions: under different shortwave radiation, relative humidity, and surface wind conditions, the leaf temperature and the potential for heat stress vary a lot (De Boeck et al. 2016).

Until recently, it was not possible to simulate plausible future climates in ecosystem climate change experiments (Korell et al. 2019), as these experiments require accurate manipulation of environmental variables to represent current and future climate conditions. Controlled environment facilities meet these requirements by providing systems to simultaneously manipulate as well as measure multiple parameters (e.g., Lawton 1993; Lawton 1996; Griffin et al. 1996; Steward et al. 2013; Clobert et al. 2018), especially in combination with an observation station in the field providing real-time observations of most of those parameters (Rineau et al. 2019). This approach is powerful especially when combined with a measurement station in the field providing real-time observations of most of these required parameters (Rineau et al. 2019). In such facilities, climate change experiments can be informed by meteorological forcing representing both present and future climatic conditions in a holistic manner. For instance, this forcing can include both realistic changes of climate variability as well as important drivers of changes in the frequency, intensity, and duration of meteorological extremes. This potential is especially interesting in gradient experiments covering a range of global warming levels, as this combination allows for the detection of non-linearities, thresholds, and possible tipping points in ecosystem responses to increasing climate change forcing (Kreyling et al. 2018; Rineau et al. 2019).

Sampling realistic climate information in a climate change context is challenging but can be achieved by using climate model output. Global climate models (GCMs) are generally used to assess the climate state and variability at global to continental scales with a resolution of 100 to 250 km. By dynamically downscaling GCMs, regional climate models (RCMs) typically resolve the climate on a regional scale with higher spatial resolutions of 1 to 50 km. As such, RCMs allow a more realistic representation of meso-scale atmospheric processes and processes related to orography and surface heterogeneities.

As climate models realistically simulate the atmospheric state under past, present, and future climatic conditions with a high temporal resolution, they are suited to provide a holistic and physically consistent climate forcing for ecosystem climate change experiments. Generally, ensemble climate projections show a large spread for future climate conditions (Keuler et al. 2016), especially for variables relevant for ecosystem experiments such as extreme air temperature, droughts, and intense precipitation (Sillmann et al. 2013; Orłowsky and Seneviratne 2013; Rajczak and Schär 2017; Greve et al. 2018). This spread is related to (i) different climate sensitivities of the GCMs, (ii) structural differences between the models, and (iii) natural variability within the climate system. The Coordinated Regional Climate Downscaling Experiment in the European domain (EURO-CORDEX) provides an ensemble of high-resolution dynamically downscaled RCMs (Kotlarski et al. 2014) and is therefore highly suitable to serve as a base for the selection of representative climate forcing for climate change experiments. With a suite of GCM/RCM combinations available, a well-informed choice on the most adequate RCM/GCM simulation can be made based on (i) the model skill in representing the observed climatology and (ii) the air temperature sensitivity to future increases in greenhouse gas concentrations.

So far, statistically downscaled GCM output has only rarely been used as climate forcing in ecosystem experiments. Thompson et al. (2013) describe a process for generating air temperature forcing for experiments in which they use daily air temperature output from a GCM (MIROC) and a stochastic weather generator to generate hourly weather. They validated their method against statistical characteristics of air temperature observations. Likewise, the Montpellier CNRS ecotron facility is driven by multivariate statistically downscaled GCM projections (using the ARPEGEv4 model; Roy et al. 2016). They force their experiment with climatic conditions of an average climatological year of the period 2040–2060. During the summer months, they artificially simulate an extreme event by including drought and heatwave by reducing the irrigation amount to zero and increasing the air temperature artificially. However, by using a climatological year, possible extreme events are dampened by averaging. Both studies lack a thorough evaluation procedure for selecting the used climate model. Moreover, to the best of our knowledge, no study accounts for the co-variance between climate variables.

In this paper, we present a new method for creating realistic climate forcing for manipulation experiments. From an ensemble of dynamically downscaled climate model simulations, we select one simulation that well represents present-day climate conditions for four key variables in the region of interest and is representative of the multi-model mean of these variables in future projections.

In this way, the new methodology accounts both for co-variance of climate parameters and for climate variability while naturally incorporating extreme events under present and future climate conditions. Furthermore, the method can be combined with a gradient approach. We apply the new methodology to generate climate forcing for the UHasselt Ecotron Experiment, an infrastructure consisting of 12 climate-controlled units, each equipped with a lysimeter containing a dry heathland soil monolith extracted from the National Park Hoge Kempen in Belgium (Rineau et al. 2019). In this experiment, six units are directly forced with regional climate model output along a global mean air temperature (GMT) gradient anomaly.

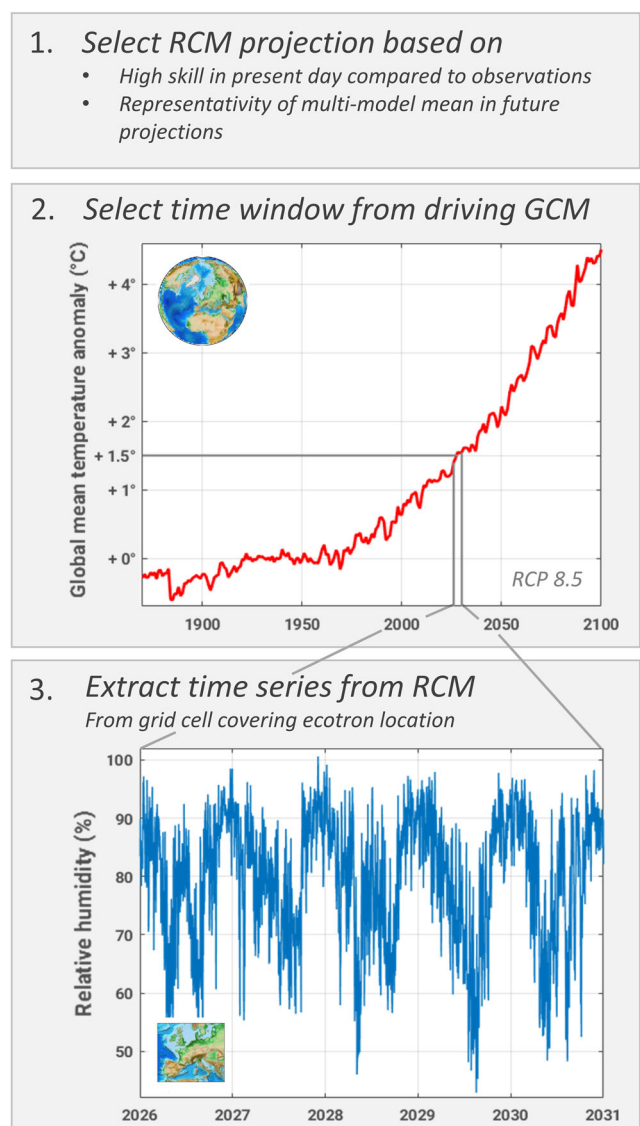
## Data and methods

### New methodology for generating climate forcing for ecosystem climate change experiments

In our methodology, variability and co-variance between variables are preserved by selecting the best-performing RCM simulation and subsequently extract the required variables from the grid cell covering the location of the experiment. By extracting a single grid cell of a single RCM simulation, climate extremes are not smoothed and the climate variability inherent to the model is fully preserved. The units in the ecosystem climate change experiments follow a gradient of increasing GMT anomalies. In this way, a given unit is forced with the climatic conditions consistent with, e.g., a 2 °C warmer world, and the units represent conditions associated with increasingly warmer climates.

The methodology presented here is deployed in three steps. First, the best-performing RCM projection needs to be selected based on two criteria: (i) the simulation should have high skill in reproducing mean and extreme present-day climatic conditions and (ii) the projected future air temperature anomalies should be close to the multi-model mean; that is, the selected simulation should be representative of the future mean projection (Fig. 1, step 1). To this end, the model performance is evaluated for four variables that are highly relevant for ecosystem climate change experiments: precipitation, air temperature, relative humidity, and surface wind speed. Precipitation is considered one of the most important variables, as water availability is likely to constrain plant growth the most.

Second, the time windows for the different units along the GMT anomaly gradient are defined based on the annual GMT projection of the driving GCM of the chosen RCM simulation (Fig. 1, step 2). To span a large range of climate change scenarios, we use projections following the Representative Concentration Pathway (RCP) 8.5, a



**Fig. 1** Methodology for generating climate forcing along the GMT anomaly gradient

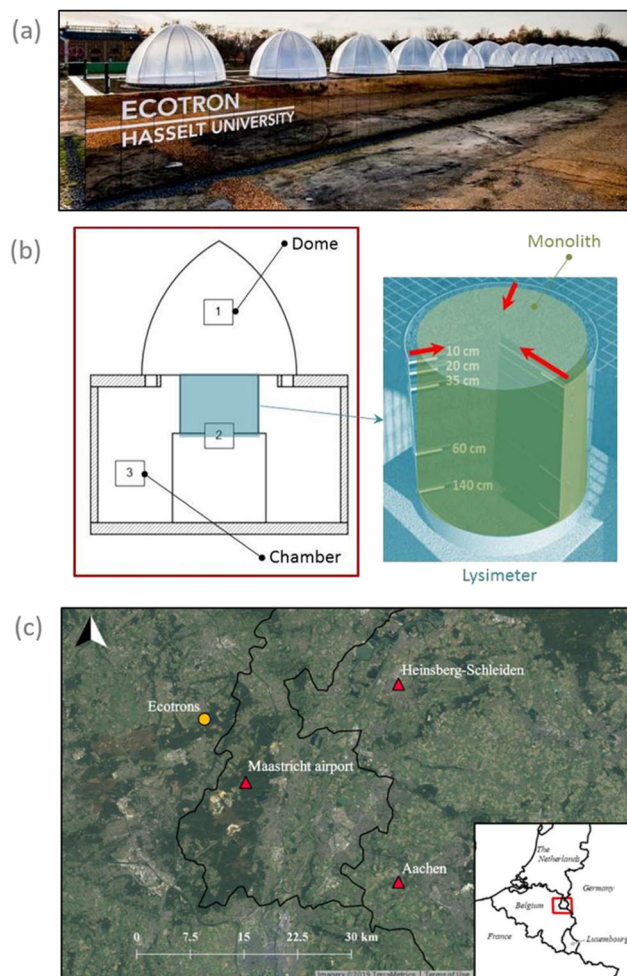
worst-case scenario following an unabated greenhouse gas emissions pathway (Riahi et al. 2011). The UHasselt Ecotron experiment, including all units, is running for 5 years. We choose time windows corresponding to the experimental period and centered around the year in which the climatological GMT anomaly (averaged with a 30-year period) crossed the predefined thresholds for the first time. In the third step, the values of all necessary variables are extracted from the chosen RCM projection based on the defined time windows for the grid cell covering the experiment location (Fig. 1, step 3). These time series are then directly used to force the ecotron units, in the highest available temporal resolution.

## The UHasselt Ecotron experiment

The UHasselt Ecotron experiment is an ecotron infrastructure consisting of replicated experimental units in which ecosystems are confined in enclosures. By allowing the simultaneous control of environmental conditions and the online measurement of ecosystem processes, the ecotron units are suited for experiments with highly controlled climate change manipulation of large intact parts of the ecosystem. The infrastructure allows intensive monitoring and control of key abiotic parameters on 12 large-scale ecosystem replicas, called “macrocosms.” These macrocosms had been extracted without disruption nor reconstitution of the soil structure from the same dry 6- to 8-year-old heathland plot in the National Park Hoge Kempen (50° 59′ 02.1″ N, 5° 37′ 40.0″ E) in November 2016.

The infrastructure is a W-E oriented, 100 m by 10 m wide, and 6 m tall building (Fig. 2a). Only 12 of the 14 units are used, excluding the outermost to avoid boundary effects. Each unit consists of three compartments in which the abiotic environmental variables are controlled: the dome, the macrocosm, and the chamber. The dome is transparent for photosynthetic active radiation (PAR) and long- and medium-wave ultraviolet radiation (UVA and UVB, respectively). Here, wind and precipitation are measured and generated, and CO<sub>2</sub>, N<sub>2</sub>O, CH<sub>4</sub>, PAR, and net radiation (NR; i.e., the difference in incoming and outgoing short- and longwave radiation) are measured. The second compartment, the macrocosm, contains the extracted soil column (the ecosystem) enclosed in a lysimeter. In this compartment, the soil water content, soil water tension, soil electrical conductivity, and soil temperature are measured and controlled. The chamber, the third compartment, the air pressure, air temperature, relative humidity, and CO<sub>2</sub> concentration are controlled. The ecotron infrastructure is linked with an Integrated Carbon Observation System (ICOS) ecosystem station, which provides real-time information on local weather and soil conditions. These data are used to simulate the current weather conditions within the ecotron units with a frequency of at least once every 30 min (Rineau et al. 2019).

The aim of the UHasselt Ecotron experiment is to study the ecological and societal impacts of climate change, by manipulating climatic variables alone or in combination and, across a wide range of predicted values, while monitoring as many soil biota and processes as possible and to translate them into socio-economic values using heathland as a case study (Rineau et al. 2019). Examples of measured ecosystem processes are evapotranspiration, net ecosystem exchange, CH<sub>4</sub> or N<sub>2</sub>O emissions. The main research questions of this multi-disciplinary experiment are how climate change will affect the transitioning of the heathland ecosystem to alternative stable states like pine



**Fig. 2** The UHasselt Ecotron experiment. **a** (picture: Liesbeth Driessen). **b** Scheme of a unit with the three compartments (1) denoting the dome; (2) the lysimeter, shown in detail on the right; and (3) the chamber. **c** An overview map with location of the infrastructure and reference weather observation stations

forest or acid grassland and what the consequences are for ecosystem services. The experiment will run uninterrupted for a period of at least 5 years. Six units will be used to simulate a gradient of increasing variability in precipitation regime. They are driven by the ICOS station and a perturbed precipitation time series following a gradient of increasingly long periods with no precipitation (2, 6, 11, 23, 45, and 90 days). In the remaining six units, atmospheric conditions along the GMT anomaly gradient will be simulated as described in “[New methodology for generating climate forcing for ecosystem climate change experiments.](#)” The 3-hourly RCM output is linearly interpolated to a 30-min time resolution to force the ecotron units. For soil temperature and soil water tension however, the 30-min ICOS data is used. This is because leaving the lysimeter uncontrolled would lead to (i) an overestimation of soil temperature



variability as the lysimeter is exposed to air temperature in the chamber (despite being thermally insulated), and (ii) accumulation of water at the bottom of the lysimeter, hence considerably overestimating soil water level, as soil water movements are mimicked by suction from the bottom. Following the gradient design, each ecotron unit represents the local climate conditions of a globally 0 °C (historical), +1 °C (present day), +1.5 °C (Paris Agreement), +2 °C, +3 °C, and +4 °C warmer world. The climatology of the unit forced by +1 °C can thereby be directly compared to the unit driven by the ICOS station and thus representing the present-day observed conditions. In this regression design, there is no experiment replication. To minimize the noise in initial ecosystem responses, the units are allocated to the two gradient experiments based on a cluster analysis of the variance of the 14 variables measured during a test period of 11 months (Rineau et al. 2019).

## Meteorological data

### EURO-CORDEX

The best-performing RCM simulation compared to observations is selected from the Coordinated Regional Climate Downscaling Experiment in the European domain (EURO-CORDEX), an ensemble of high-resolution dynamically downscaled simulations available at a horizontal resolution of 12 km (0.11° on a rotated grid; Jacob et al. 2014; Kotlarski 2014). The simulations, hereafter referred to as GCM downscalings, cover the historical period (1951–2005) and the three RCP scenarios (RCP 2.6, 4.5, and 8.5, for the period 2006–2100) by using GCMs as initial and lateral boundary conditions. Additionally, for each RCM, a reanalysis downscaling is provided in which the RCM is driven by the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA-Interim as initial and lateral boundary conditions for the period 1990–2008 (hereafter referred to as reanalysis downscalings). These reanalysis-driven simulations allow to evaluate the skill of the RCMs themselves by comparing them to observations (Kotlarski et al. 2014).

In this study, we use the variables for daily mean, minimum, and maximum air temperature, precipitation, mean surface wind, and relative humidity of all available simulations (Table 1). We consider the values of the 12 km by 12 km pixel covering the location of the reference station providing the observations. As relative humidity is not directly available for all simulations, we converted specific humidity to relative humidity using the mean air temperature and surface air pressure for every simulation. Comparing the applied conversion with the simulations for which relative humidity is available proves this conversion is applicable. Neither specific nor relative humidity is publicly available for the simulations with RegCM4-2 and

ALARO-0 and the mean surface wind speed variable is not available for ALADIN53 and ALARO-0; therefore, we do not analyze these variables for the respective simulations.

Once the EURO-CORDEX ensemble member is selected, the relevant variables (precipitation, mean air temperature, surface air pressure, surface up-welling latent heat flux and sensible heat flux, wind speed, and relative humidity) are extracted from the 3-hourly RCP 8.5 simulation for the pixel covering the ecotron location for the time windows in which the GMT anomalies are crossed for each dome. These 3-hourly values (except for surface up-welling latent heat flux and sensible heat flux) are then linearly interpolated to 30-min resolution and used to drive the climate controllers in the ecotron units. For precipitation, one additional step was added where drizzle (precipitation of less than 1 mm) was postponed and accumulated until it reached 1 mm to start a rain event in the ecotron. The surface air pressure is calculated from the mean sea level pressure using the elevation of the ecotron facility (43 m a.s.l.) and assuming hydrostatic equilibrium. The concentrations of the controllable greenhouse gases (CO<sub>2</sub>, CH<sub>4</sub>, and N<sub>2</sub>O) are determined based on the annual values calculated by van Vuuren et al. (2011) according to RCP 8.5. These correspond to the prescribed concentrations of the RCM simulations.

### Weather station observations

Reference station data is obtained from the European Climate Assessment and Dataset (Klein Tank et al. 2002). The three operational weather stations closest to the UHasselt Ecotron experiment are Maastricht Airport (11 km), Aachen (37 km) and Heinsberg-Schleiden (29 km; Fig. 2c). These weather stations provide daily observations from the end of the 19<sup>th</sup> century (Maastricht Airport and Aachen) or mid 20<sup>th</sup> century (Heinsberg-Schleiden) until the present-day, thereby covering both the EURO-CORDEX GCM and reanalysis downscaling periods. All stations record air temperature (°C), precipitation (mm day<sup>-1</sup>), relative humidity (%) and surface wind speed (m s<sup>-1</sup>) at daily resolution, except for the Heinsberg-Schleiden station where there are no surface wind observations available.

The seasonal cycles of the observations for the different stations follow a similar annual course (Fig. 3). For air temperature, the curves overlay and for precipitation they are similar. Relative humidity (RH) has a small offset between the three stations, possibly owing to the differences in absolute height and local topography. The difference in surface wind speed between Maastricht Airport and Aachen is considerable but plausible considering the large spatial variability in wind speed. Given that the model evaluation showed very little sensitivity to the choice of the reference station, we hereafter present the results with the reference station closest to the ecotron facility (Maastricht Airport).

**Table 1** Bias in annual precipitation (P bias) and rank based thereof (from 1-best to 18-worst) for the EURO-CORDEX GCM downscalings for the period 1951–2005 over Maastricht Airport

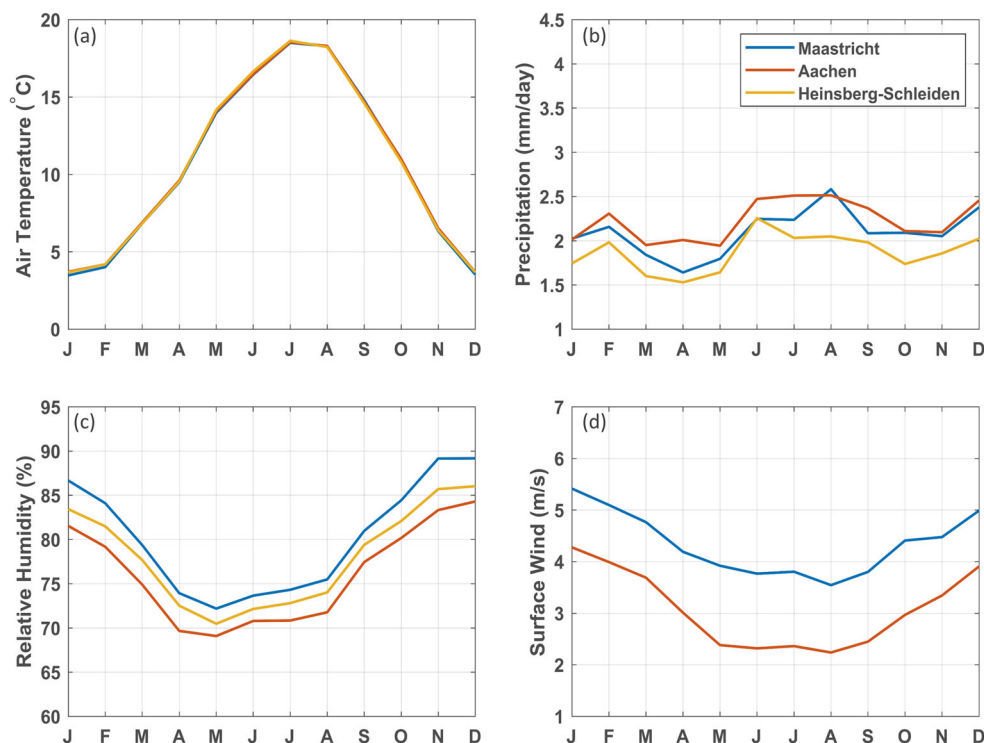
RCM	GCM	P bias (mm year <sup>-1</sup> )	Rank
CCLM4-8-17	CNRM-CERFACS-CNRM-CM5	145	8
CCLM4-8-17	ICHEC-EC-EARTH	8	1
CCLM4-8-17	MOHC-HadGEM2-ES	-174	9
CCLM4-8-17	MPI-M-MPI-ESM-LR	24	2
ALADIN53	CNRM-CERFACS-CNRM-CM5	550	14
HIRHAM5	ICHEC-EC-EARTH	323	12
HIRHAM5	MOHC-HadGEM2-ES	101	6
HIRHAM5	NCC-NorESM1-M	571	16
WRF331F	IPSL-IPSL-CM5A-MR	726	18
RACMO22E	ICHEC-EC-EARTH	99	5
RACMO22E	MOHC-HadGEM2-ES	36	3
REMO2009	MPI-M-MPI-ESM-LR	225	10
ALARO-0	CNRM-CERFACS-CNRM-CM5	560	15
RCA4	CNRM-CERFACS-CNRM-CM5	319	11
RCA4	ICHEC-EC-EARTH	386	13
RCA4	IPSL-IPSL-CM5A-MR	691	17
RCA4	MOHC-HadGEM2-ES	111	7
RCA4	MPI-M-MPI-ESM-LR	70	4

### Metrics and diagnostics

The evaluation of the EURO-CORDEX ensemble members is performed using different metrics accounting for

performance of representing the climatic means, distributions, and extremes.

A ranking is made of the reanalysis downscalings, assigning the lowest ranks to the best-performing models



**Fig. 3** Seasonal cycles of observed mean air temperature (a), mean daily precipitation (b), mean relative humidity (c), and mean surface wind speed (d) in the weather stations of Maastricht Airport, Aachen,

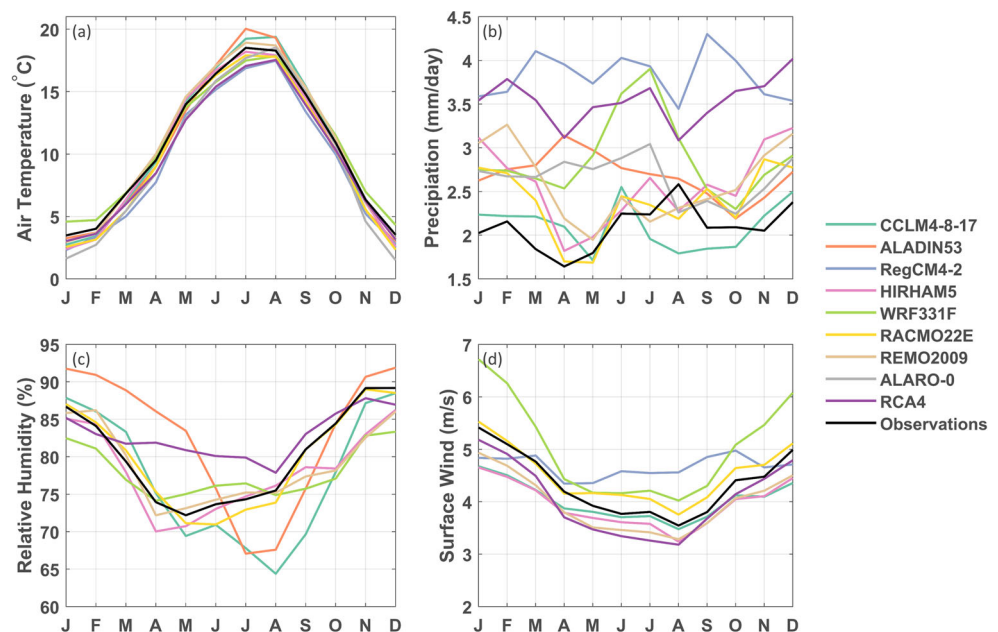
and Heinsberg-Schleiden (monthly averages based on daily data from 1963 to 2018). For Heinsberg-Schleiden no surface wind observations are available. The curves for air temperature are overlaying

and higher ranks to the least-performing models (1-best to 9-worst). First, the bias is calculated as the difference between the averages of the daily modelled and observed variables. The second metric, the Perkins skill score (PSS), is a quantitative measure of how well each simulation resembles the observed probability density functions by measuring the common area between two probability density functions (Perkins et al. 2007). The mean absolute error (MAE) is calculated by taking the means of the absolute differences between the modelled and observed seasonal cycles, calculated based on the whole series. This is done for the whole series and to capture the potential errors in the extremes, also for the 1st, 10th, 90th, and 99th percentiles which are calculated based on the daily time series of both observed and modelled time series. Next, the root mean square error (RMSE) is calculated by taking the root of the squared errors. The Spearman rank correlation (hereafter referred to as Spearman) coefficient shows the correlation of the observed and modelled series, calculated based on daily values. Finally, the Brier skill score (BSS) is calculated, which gives an indication of the improvement of the Brier score (an index to validate probability forecasts) compared to a background climatology in which each event has an equal occurrence probability (Brier 1950; Murphy 1973). For the GCM downscalings, we use the same ranking method and scores, except for the RMSE, Spearman rank correlation, and BSS because the internal variability, inherent to individual simulations with a coupled climate model, cannot be predicted on multi-decal timescales, and

can therefore not be compared to observations on a day-by-day basis (Fischer et al. 2014; Meehl et al. 2014).

In addition to the performance metrics computed on the actual time series, the RCM performance is also evaluated based on the bias in climatological diagnostics related to air temperature and precipitation. To this extent, the average diurnal air temperature range (DTR; K; the difference between the daily maximum and minimum air temperature) is calculated for the whole year, for the winter (December–February) and summer (June–August) season. Next, the number of wet days (defined as days during the year for which precipitation is larger than 0.1 mm or larger than 1 mm, to account for differences in drizzle; (Casanueva et al. 2016) and the number of frost days (days with a minimum air temperature below 0 °C) are calculated. Furthermore, the monthly maximum 1-day precipitation (Rx1day; mm day<sup>-1</sup>) and the number of consecutive dry days (CDD; days); the annual maximum number of days for which precipitation is below 1 mm and consecutive wet days (CWD; days); and the annual maximum number of days for which precipitation is equal to or more than 1 mm are included in the analysis. All indices are calculated for the simulated and observed time series, and consequently the ranking is established based on the difference between the model and observed diagnostic. Next, the correlation between the different variables is evaluated by comparing them to the observed correlation. This is done both on the annual time scale and for the summer and winter seasonal averages, as correlations are expected to differ in sign

**Fig. 4** Seasonal cycle of the reanalysis downscalings for mean air temperature (a), mean daily precipitation (b), mean relative humidity (c), and mean surface wind speed (d). (The RegCM4-2 and ALARO-0 simulations are not available for relative humidity and the ALADIN53 and ALARO-0 simulations are not available for surface wind speed)

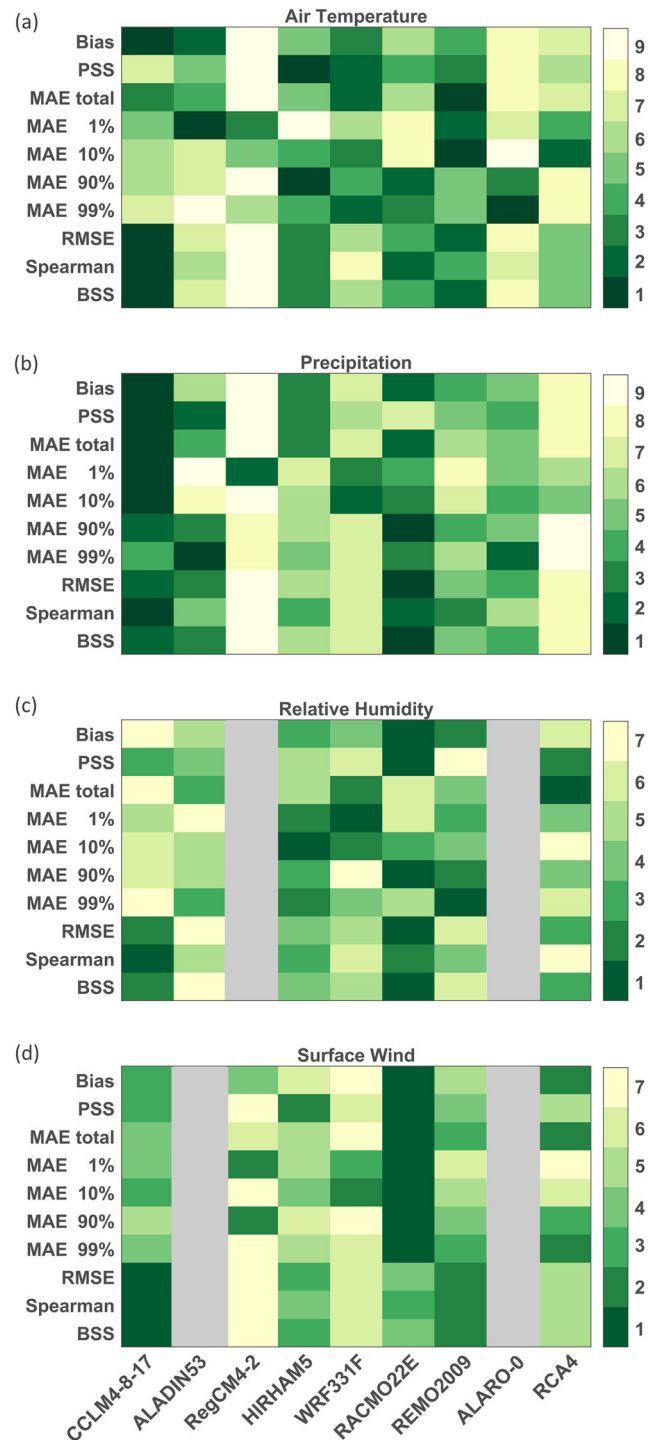


and magnitude between the two seasons (e.g., negative correlation between air temperature and relative humidity in summer reflecting heatwave conditions, and a positive correlation between wind speed and precipitation in winter reflecting storm conditions).

After choosing the best-performing simulation based on the evaluation of both the reanalysis and GCM downscalings, the climate change signals for this simulation are investigated by calculating changes in various climate change indices, based on the Expert Team on Climate Change Detection and Indices (ETCCDI; see [http://etccdi.pacificclimate.org/list\\_27\\_indices.shtml](http://etccdi.pacificclimate.org/list_27_indices.shtml)) for the 5-year periods defined by the GMT anomalies relative to the reference period (1951–1955). These indices are widely used for analyzing changes in extremes (e.g., Zhang et al. 2009; Orłowsky and Seneviratne 2013; Sillman et al. 2013). The air temperature indices are (i)  $\Delta T$  ( $^{\circ}\text{C}$ ), the mean daily air temperature change; (ii)  $\Delta \text{TXx}$  ( $^{\circ}\text{C}$ ), the difference in the annual maximum value of daily maximum air temperature; (iii)  $\Delta \text{TNn}$  ( $^{\circ}\text{C}$ ), the difference in the annual minimum value of daily minimum air temperature; (iv)  $\Delta$  frost days, the difference in the number of frost days (with a minimum air temperature below  $0^{\circ}\text{C}$ ); (v)  $\Delta$  summer days, the difference in the number of summer days (with the maximum air temperature above  $25^{\circ}\text{C}$ ); and finally (vi)  $\Delta$  GSL (days), the difference in growing season length, defined as the annual count between the first span of at least 6 days with a daily mean air temperature higher than  $5^{\circ}\text{C}$  and the first span after July 1st of 6 days with a daily mean air temperature lower than  $5^{\circ}\text{C}$ . The precipitation indices are (i)  $\Delta \text{PRCPTOT}$  (mm), the difference in annual accumulated precipitation (as simulated over the 5-year period); (ii)  $\Delta \text{Rx1day}$  (mm), the difference in monthly maximum 1-day precipitation; (iii)  $\Delta \text{R10mm}$  (days), the difference in the number of days per year with more than 10 mm precipitation, (iv)  $\Delta \text{CDD}$  (days), the difference in the maximum length of a dry spell (measured as the maximum number of consecutive days with less than 1 mm precipitation); and finally, (v)  $\Delta \text{CWD}$  (days), the maximum length of a wet spell (measured as the maximum number of consecutive days with more than 1 mm precipitation).

### Applying the new methodology for the UHasselt Ecotron experiment

The best-performing RCM simulation is identified by elimination based on expert judgment based on the performance of the two selection criteria. Next, we define the time windows for the different units along the gradient based on the 30-year averaged GMT anomaly of the driving GCM under RCP8.5 relative to 1951–1955. Based on these



**Fig. 5** Ranking of the reanalysis downscalings based on performance on mean air temperature (a), mean daily precipitation (b), mean relative humidity (c), and mean surface wind speed (d) compared to observations from Maastricht. The metrics shown are the bias, Perkins skill score (PSS), mean absolute error (MAE) for the entire time series and the 1st, 10th, 90th, and 99th percentiles, root mean square error (RMSE), Spearman rank correlation (Spearman), and Brier skill score (BSS). Rankings are from 1-best to 9-worst. Gray colors indicate that the variable is not available for the considered model



time windows, we extract the 3-hourly data for all necessary variables from the simulation for the 12-km by 12-km grid cell covering the location of the experiment.

## Results

### Identification of the best-performing model simulation

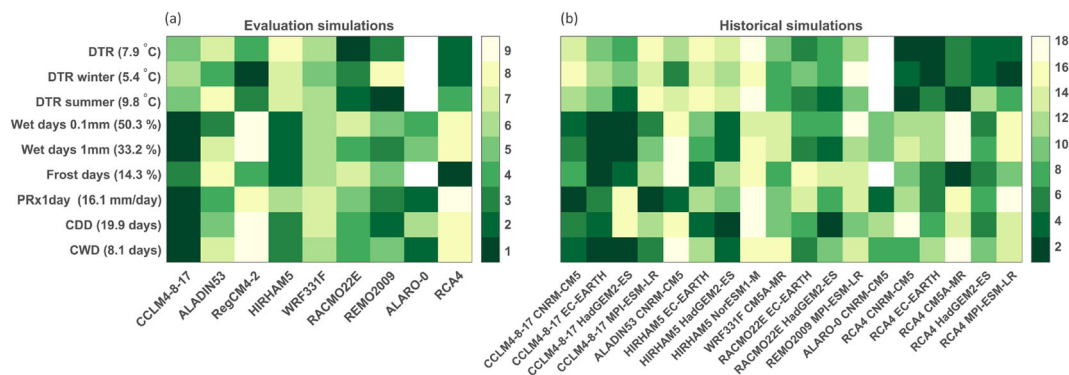
#### First criterion: skill in present-day climate

Overall, model skill strongly varies across RCMs (Fig. 4). While the annual air temperature cycle is generally well represented by all RCMs, biases may reach up to 2 degrees in individual months for some RCMs. The biases in precipitation are generally positive (up to factor 2.4) and vary across RCMs. Only CCLM4-8-17 simulates precipitation in the same range as the observed climatology (nearly no bias (100.22 %) on annual mean precipitation amounts), while the other RCMs overestimate the total precipitation amounts from 114 % up to 182 %. For relative humidity and surface wind speed, all RCMs generally succeed in representing the seasonal cycle, but exhibit deviations in amplitude and absolute values (e.g., amplitude biases of RCA4 (−37.8 %), ALADIN53 (23.3 %) and CCLM4-8-17 (+16.3 %) for relative humidity, and annual mean biases for WRF331F (+15.6 %) and HIRHAM5 (−9.1 %) for surface wind speed). Overall, these seasonal cycles indicate that for all simulations, the relative bias in precipitation is large compared to biases in other variables.

The rankings of the reanalysis downscalings for the four variables (Fig. 5) indicate that, overall, CCLM4-8-17, RACMO22E, REMO2009, and HIRHAM5 are performing best. CCLM4-8-17 and RACMO22E show the highest relative skill for precipitation, while REMO2009 and

HIRHAM5 demonstrate high skill for air temperature. CCLM4-8-17 is the best-performing model based on the bias and total MAE metrics for air temperature and precipitation but is ranked in the mid range for the metrics related to the shape of its air temperature distribution (PSS and percentile MAE). This can be attributed to an overestimation of the amplitude of the seasonal air temperature cycle in this model (too cold in winters, too hot in summers; Fig. 4a; Kotlarski et al. 2014). For relative humidity and surface wind speed, RACMO22E generally demonstrates the highest skill. Considering the climatological diagnostics (Fig. 6a), CCLM4-8-17 shows the highest relative skill for precipitation-related diagnostics (wet days, monthly maximum 1-day precipitation, length of dry and wet spells), while RACMO22E and RCA4 show higher relative skill for the annual, winter, and summer diurnal air temperature range. While RCA4 is highly ranked for air temperature-related diagnostics, it is one of the models with the lowest relative skill for precipitation-related diagnostics. The correlation ranking shows a more scattered image, for the annual correlation as well as summer and winter correlations (see Suppl. Fig. A1). Overall, as the reanalysis-driven simulations with ALADIN53, RegCM4-2, WRF331F and ALARO-0 show the lowest skill compared to the other RCMs, we take them out of consideration to serve as ecosystem forcing.

Second, we evaluate the GCM downscalings for the period 1951–2005. The seasonal cycles of the air temperature, precipitation, relative humidity, and surface wind speed show a similar pattern as the reanalysis downscalings, with again a strong wet bias for precipitation in most models (see Suppl. Fig. A2). The rankings show a mixed pattern for the different variables: there are no simulations which rank high for all considered variables (Fig. 7). For precipitation, the simulations with CCLM4-8-17 and RACMO22E have better relative skill compared to the other simulations,



**Fig. 6** Ranking of the reanalysis (a) and GCM (b) downscaling for the historical period based on climatological diagnostics. Diurnal air temperature range (DTR) in summer (July–August) and winter (December–February), number of wet days defined as days with precipitation > 0.1 mm and precipitation > 1 mm, number of frost days defined as days with mean air temperature < 0 ° C, monthly

maximum 1-day precipitation (Rx1day), consecutive dry days (CDD), the maximum length of a dry spell, and consecutive wet days (CWD), the maximum length of a wet spell. Next to the diagnostic name, its value as observed in Maastricht Airport is shown. Rankings are from 1-best to 9 or 18-worst for the reanalysis and GCM downscaling, respectively

which is in line with the high ranking of these models in the reanalysis downscalings. Furthermore, it is remarkable that the simulations which show a high skill for precipitation, typically show lower skill for relative humidity and vice versa, e.g., CCLM4-8-17 driven by HadGEM2-ES (high ranking in precipitation, lowest in relative humidity) and REMO2009 driven by MPI-ESM-LR (high ranking in relative humidity and lower in precipitation). The three MPI-ESM-LR-driven simulations appear to be better in reproducing the air temperature climatology compared to the other simulations. For the climatological diagnostics, generally CCLM4-8-17 is scoring best for the precipitation-related diagnostics, whereas simulations with RCA4 are ranked the best for DTR (annual, summer, and winter).

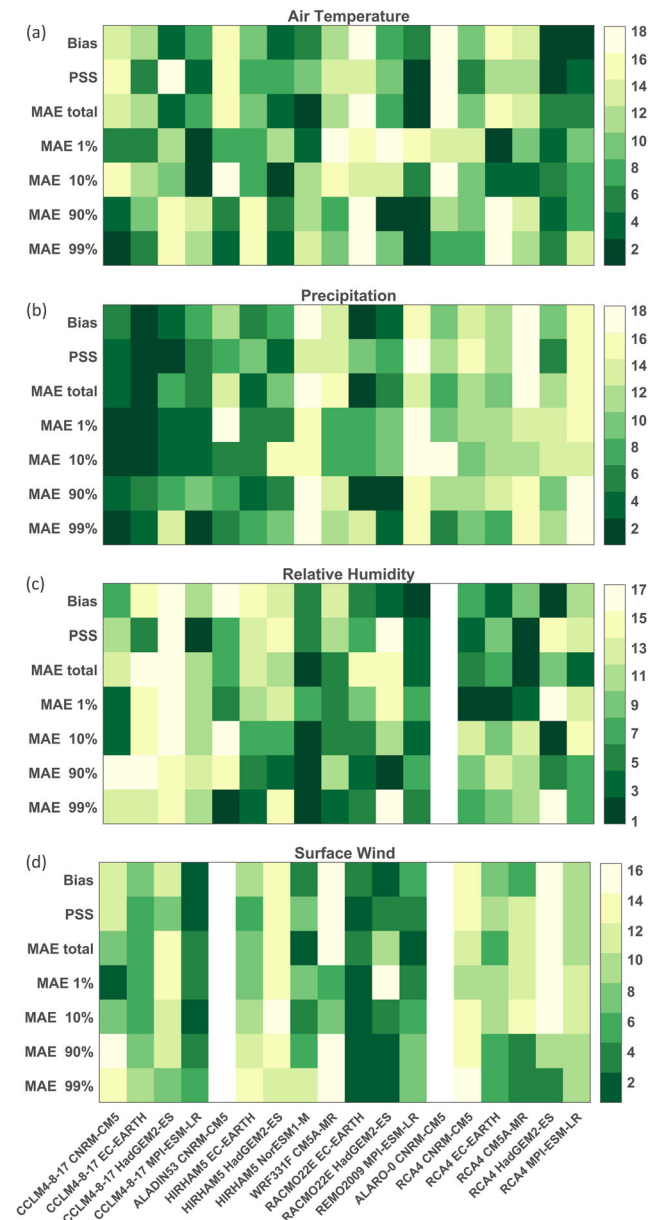
Based on the ranking of the GCM downscalings, the following simulations are considered potential candidates to serve as climate forcing: CCLM4-8-17 driven by CNRM-CM5, EC-EARTH and MPI-ESM-LR, HIRHAM5 driven by EC-EARTH and HadGEM2-ES, and RACMO22E driven by HadGEM2-ES (Figs. 5, 7, and 6). Since precipitation biases strongly differ among RCMs (Table 1), and since precipitation is a critical variable for the ecosystem experiments (Van der Molen et al. 2011; Vicca et al. 2014; Estiarte et al. 2016), we prioritize a minimum relative bias for precipitation over a lower bias for air temperature, relative humidity, and surface wind speed. The precipitation biases for the considered simulations are  $+150 \text{ mm year}^{-1}$  for CCLM4-8-17 driven by CNRM-CM5,  $+8 \text{ mm year}^{-1}$  for CCLM4-8-17 driven by EC-EARTH,  $+24 \text{ mm year}^{-1}$  for CCLM4-8-17 driven by MPI-ESM-LR,  $+323 \text{ mm year}^{-1}$  for HIRHAM5 driven by EC-EARTH,  $101 \text{ mm year}^{-1}$  for HIRHAM5 driven by HadGEM2-ES, and  $36 \text{ mm year}^{-1}$  for RACMO22E driven by HadGEM2-ES. Based on this, the CCLM4-8-17 EC-EARTH-driven simulation has the best chance to be chosen as forcing, followed by the CCLM4-8-17 MPI-ESM-LR, and the RACMO22E HadGEM2-ES-driven simulation.

## Second criterion: Representativeness of multi-model mean in future projections

To verify the second requirement, we look at anomalies from the mean signal of the four variables for the future period of the simulations under RCP 8.5. The EC-EARTH-driven CCLM4-8-17 simulation is representative of the multi-model mean for all four variables (Fig. 8), and even the median simulation for the mean air temperature anomaly. For precipitation and relative humidity, however, the CCLM4-8-17 EC-EARTH simulation shows decreasing anomalies after 2050 and underestimates the multi-model mean anomaly. The other selected simulations have a larger positive bias in precipitation for their GCM downscalings. A possible reason is that these simulations overestimate

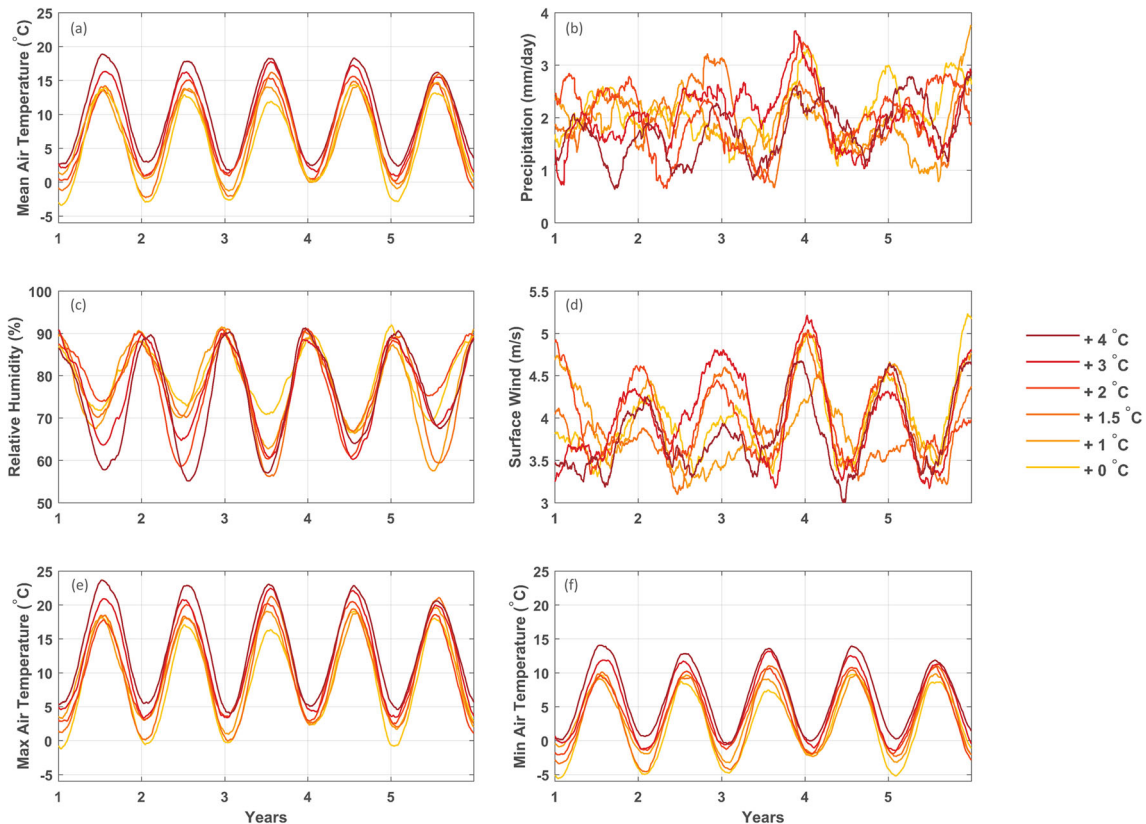
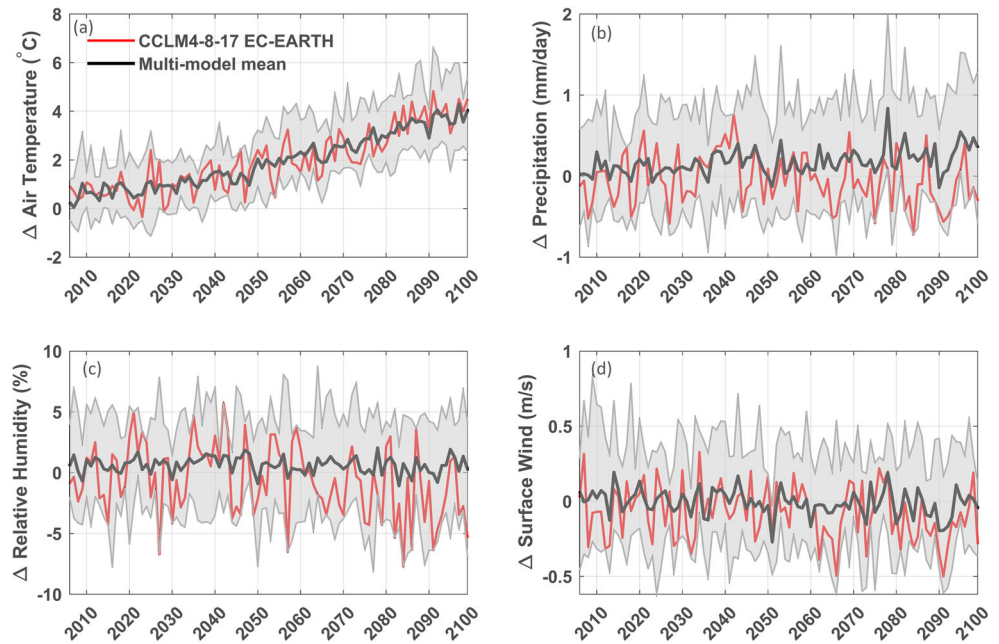
precipitation and simulate a more intensive hydrologic cycle, which also implies stronger changes in the future.

The remaining five simulations from step 1 (CCLM4-8-17 driven by MPI-ESM-LR, HIRHAM5, and RACMO22E driven by HadGEM2-ES) all systematically underestimate or overestimate other variables (Suppl. Figs. A4, A5, A6, A7 and A8). For instance, the mean air temperature anomaly of



**Fig. 7** Ranking of the GCM downscalings based on performance on mean air temperature (a), mean daily precipitation (b), mean relative humidity (c), and mean surface wind speed (d) compared to observations from Maastricht. The metrics showed are the bias, Perkins skill score (PSS), mean absolute error (MAE) for the total and 1st, 10th, 90th, and 99th percentile. Rankings are from 1-best to 16, 17, or 18-worst for surface wind speed, relative humidity, precipitation, and air temperature, respectively. Gray colors indicate that the variable is not available for the considered model

**Fig. 8** Anomalies for the CCLM4-8-17 EC-EARTH simulation following RCP 8.5 at the ecotron site for mean air temperature (a), mean daily precipitation (b), mean relative humidity (c), and mean surface wind speed (d). The reference period is 1977 to 2006, the anomalies of the CLM4-8-17 EC-EARTH simulation are calculated compared to its own values in the reference period. In gray, the envelope of all EURO-CORDEX RCP 8.5 simulations is showed



**Fig. 9** Annual cycles of the CCLM4-8-17 EC-EARTH ecotron unit forcing for the +1 °C, +1.5 °C, +2 °C, +3 °C, and +4 °C units compared to the 0 °C reference period with a daily mean air temperature, b mean daily precipitation, c mean relative humidity, d mean surface

wind speed, e daily maximum air temperature, and f daily minimum air temperature. Curves were smoothed using Savitzky-Golay filtering (order = 2 frame = 301; Savitzky and Golay 1964)

CCLM4-8-17 driven by MPI-ESM-LR simulation (1.46 °C) is lower than the 10th percentile of all simulations (1.51 °C) and the air temperature anomaly for CCLM4-8-17 driven by CNRM-CM5 is the 30th percentile (1.67 °C). HIRHAM5 driven by HadGEM2-ES overestimates relative humidity anomalies compared to the multi-model mean, with a mean value (1.26 %) around the 80th percentile. Finally, the HadGEM2-ES-driven RACMO22E simulation overestimates relative humidity and air temperature anomalies, up to the 90th percentile for air temperature. Overall, we conclude that the EC-EARTH-driven CCLM4-8-17 simulation is the most appropriate candidate for serving as climate forcing for the UHasselt Ecotron experiment.

### Characterization of the selected meteorological forcing

Based on the selection criteria, we single out the EC-EARTH (ensemble member r12i1p1) driven CCLM4-8-17 simulation as climate forcing for the UHasselt Ecotron experiment. The climatic conditions in the six units along the gradient represent an increasing signal of climate change, representing the local climatic conditions of 6 future climate states corresponding to an increasing GMT. The overall trend of the local air temperature anomaly compared to the reference period (0 °C) increases monotonically with the corresponding GMT anomalies (Fig. 9a). No clear trends are visible for precipitation, relative humidity, and surface wind speed anomalies, but very clear for the minimum and maximum air temperature anomalies which are both increasing (Fig. 9). The mean daily air temperature is increasing at a similar rate compared to GMT anomaly, and minimum and maximum air temperature show a larger increase (Table 2). None of

the air temperature indices show a linear increase, reflecting the difference between global and local climatic conditions and the influence of decadal internal variability. The ecotron unit representing a +4 °C world is the most extreme case, with increases of TXx of +6.30 °C and an increase of TNn with +10.21 °C (Table 2). The number of frost days decreases with about −76.2, while the number of summer days with air temperature above 25 °C increases with about 36.6 days. The annual growing season length is extended with 80 days on average, leaving only 59.4 days of the year not favorable for growth. The indices for precipitation show a less clear trend (Table 2). The total precipitation amount varies for the five units, without any trend and shows a substantial decadal variability in all seasons (see Fig. 9). Rx1day has positive anomalies for the +1.5 °C, +2 °C, and +3 °C units (+0.35 mm day<sup>−1</sup>, +1.92 mm day<sup>−1</sup>, and +2.34 mm day<sup>−1</sup>, respectively). These +2 °C and +3 °C units also know an increase in R10mm (+3.2 and +3.6 days) compared to the other units. Finally, there is no clear trend in CWD, but there is an increase in CDD up to +11.8 days for the +4 °C unit. The +1.5 °C unit spans a drier time window, with an average CDD of +9.6 days. Figure 9 further shows a systematic decrease of relative humidity during summer with increasing warming and a strong decadal variability of surface wind speed especially in winter.

### Discussion

The presented methodology exhibits some challenges, which are addressed in the following section.

We extract all climate variables from one grid cell of the RCM simulation to conserve the most realistic,

**Table 2** Extracted 5-year periods and air temperature and precipitation indices based on ETCCDI for the CCLM4-8-17 EC-EARTH simulation at the ecotron location

	0 °C (ref value) 1951–1955	+1 °C 2011–2015	+1.5 °C 2028–2032	+2 °C 2043–2047	+3 °C 2067–2071	+4 °C 2091–2095
Δ T (°C)	8.17	+1.13	+1.14	+1.81	+3.15	+4.49
Δ TXx (°C)	30.98	+0.82	+1.66	+1.34	+5.24	+6.30
Δ TNn (°C)	−12.73	+6.75	+3.34	+5.94	+8.27	+10.21
Δ Frost days	103	−22	−14.8	−36.4	−59	−76.2
Δ Summer days	11.4	+4	+12.2	+8.6	+26.2	+36.6
Δ GSL (days)	225.6	+9.6	+20	+33.6	+45.8	+80
Δ PRCPTOT (mm)	771.09	−81.32	−57.2	+25.12	−23.14	−136.05
Δ Rx1day (mm)	14.38	−0.2	+0.35	+1.92	+2.34	+0.5
Δ R10mm (days)	14.6	0	−1	+3.2	+3.6	−1.2
Δ CDD (days)	17.2	+2.4	+9.6	+1.6	+7.2	+11.8
Δ CWD (days)	9.6	−0.2	+1.2	+1.4	0	−1.8

The 0 °C column gives the absolute reference values. The periods are calculated based on the 30-year averaged global mean air temperature (GMT) anomaly calculated from EC-EARTH

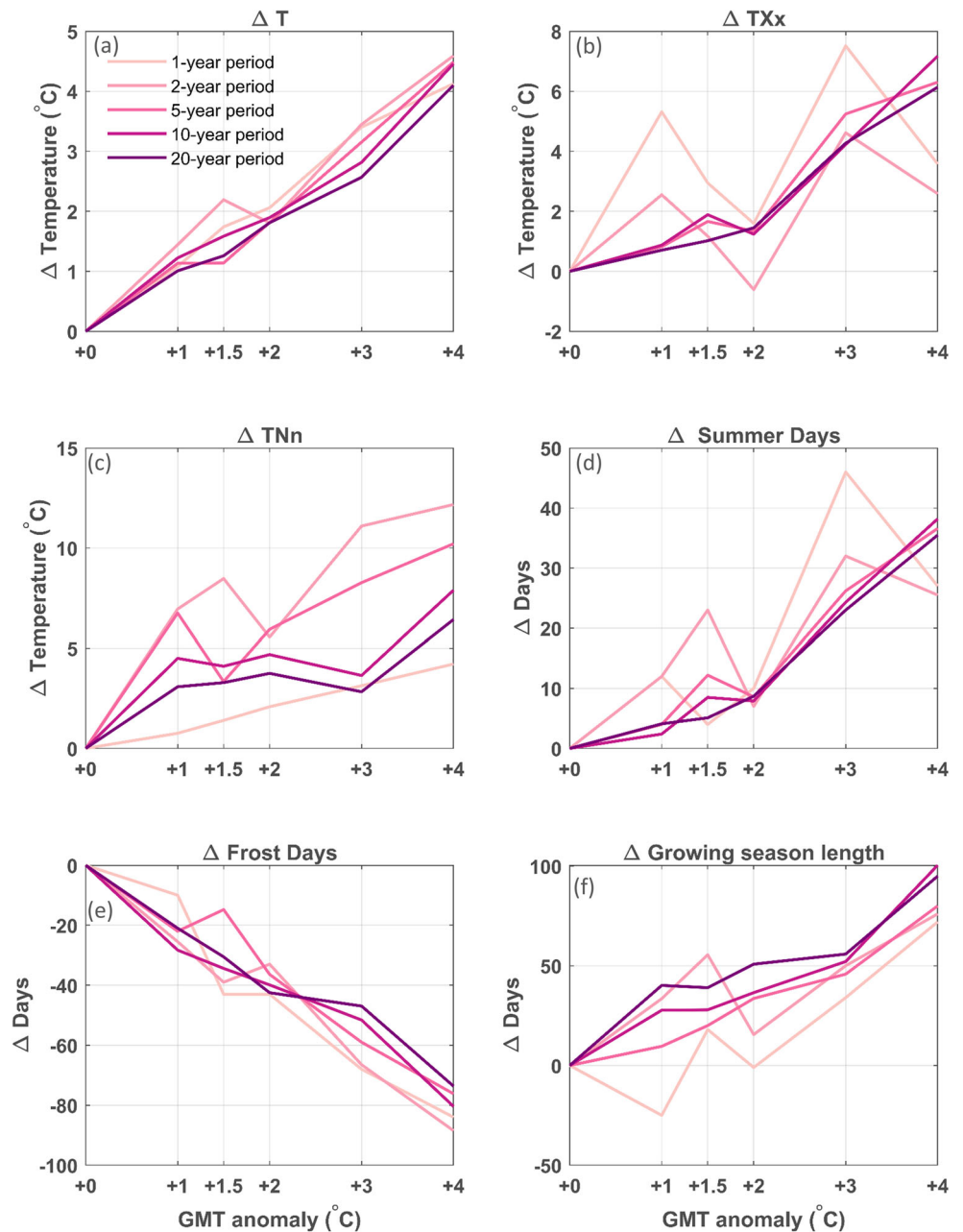


non smoothed signal of the climate models. However, the extracted time series of the grid cell can differ a lot between different models and time periods, reflecting the natural climate variability. GCMs and RCMs provide robust signals when aggregated over a larger spatial area (Seneviratne et al. 2016; Fischer and Knutti 2015). By taking the spatial mean, a more robust estimate of the mean climate is obtained, including robust signals of climate change. This explains the difference in local climate change signals (Fig. 8; Table 2) and non-linearities compared to the GMT anomaly obtained by global averaging (Seneviratne et al. 2016). It is however necessary to use actual time series from a single grid cell

to capture, e.g., the extreme precipitation event occurring in the considered grid cell, but not in the neighboring grid cells. The grid cell values also reflect strong inter-annual to decadal variability which is of high relevance for a realistic forcing of the ecosystem.

Climate model simulations are often biased, which is mostly related to structural model deficiencies (Flato et al. 2013). Applying bias adjustment is a standard way to deal with biases (Gudmundsson et al. 2012; Vanderkelen et al. 2018), but such methods face several challenges and need to be chosen carefully to not increase biases in the co-variability of variables (Zscheischler et al.

**Fig. 10** Annual anomalies per GMT anomaly for increasing time window lengths (ranging from a 1-year period to a 20-year period) of the CCLM4-8-17 EC-EARTH simulation following RCP 8.5 for air temperature indices: mean air temperature anomaly ( $\Delta T$ ; a), annual maximum air temperature ( $\Delta TXx$ ; b), annual minimum air temperature ( $\Delta TNn$ ; c), anomaly in annual number of summer days (d), frost days (e), and the anomaly in growing season length (f). Note the different y-axis scales

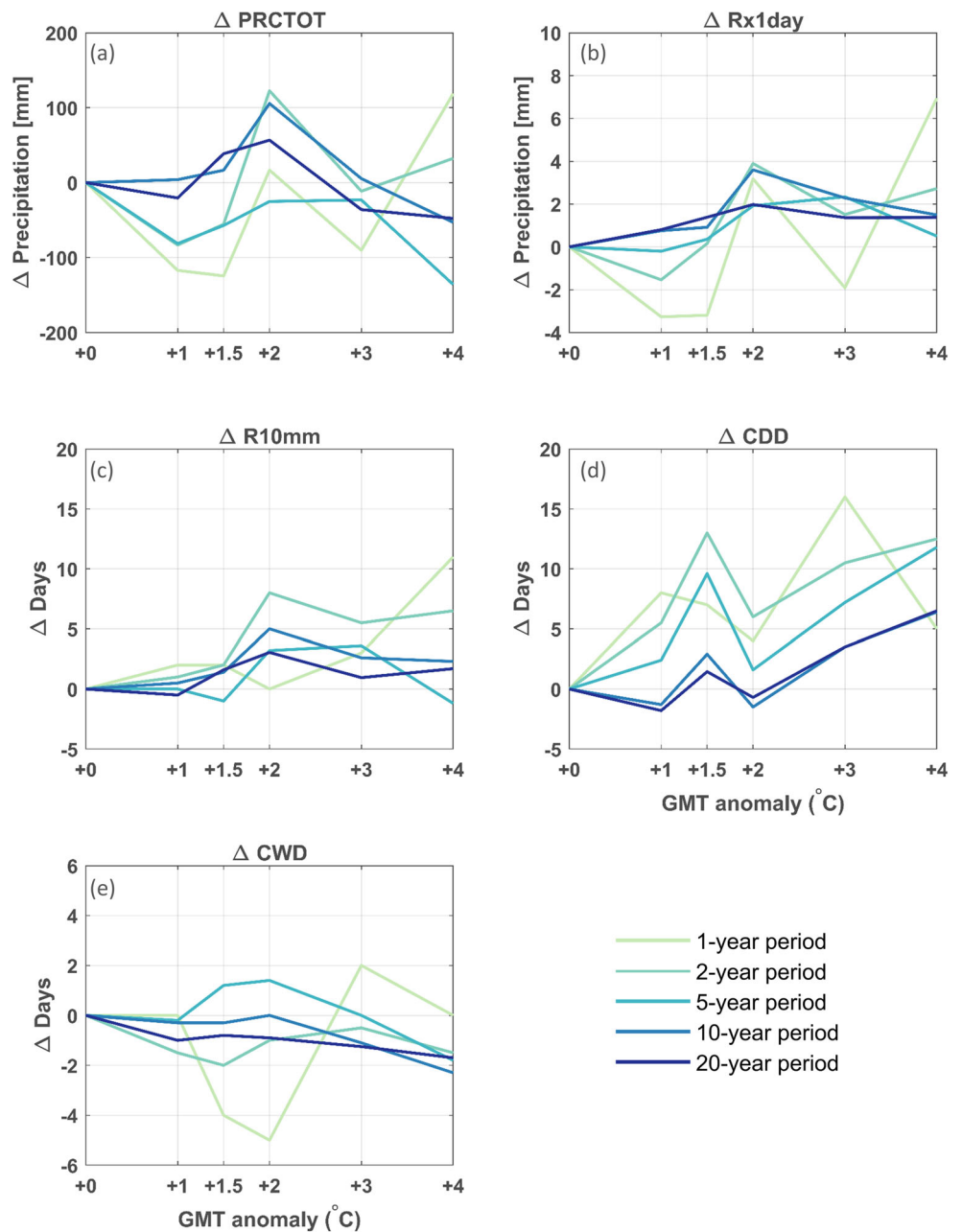


2019). In the proposed method, we therefore directly use the “raw” model output, as such preserving climate variability and the physically consistent co-variance of the different meteorological variables. In this way, the Ecotron experiment will study ecosystem responses to multivariate drivers as compound controls. For instance, it will provide a unique opportunity to study the impact of realistic compound events (Zscheischler et al. 2018), e.g., events similar to the drought-heat event of 2018, which caused massive heather die-off both in the field and in the ecotrons, forced by conditions like they happened in the field.

The gradient for the different ecotron units does not follow a monotonic trend for some of the key indicators

(Fig. 9 and Table 2), due to the high local and inter-annual natural climate variability of the climate system. This issue could be alleviated by running the experiment for a longer period. Comparing different time frames, all extracted based on 30-year averaged GMT anomaly thresholds, shows that choosing longer time windows of 10 or 20 years leads to more clear monotonic trends (Figs. 10 and 11), which is more pronounced for air temperature-derived indices than for precipitation-derived indices. For shorter time windows of 1 to 2 years, the inter-annual and local natural variability leads to larger variations in trend for the different GMT anomaly levels. Therefore, the experiment would have to run for a long period, but the experimental time frame is

**Fig. 11** Same as Fig. 10, but now for precipitation indices: the annual accumulated precipitation anomaly ( $\Delta$  PRCTOT; **a**), anomaly of monthly maximum 1-day precipitation ( $\Delta$  Rx1day; **b**), anomaly of annual number of days with more than 10 mm precipitation ( $\Delta$  R10mm; **c**), anomaly of annual maximum length of a dry spell ( $\Delta$  CDD; **d**) and anomaly of maximum length of a wet spell ( $\Delta$  CWD; **e**). Note the different y-axis scales



constrained by the experimental setup and possible renewal. As a compromise, here, we use a 5-year experimental period. Ideally, the entire gradient should be replicated several times with different climate trajectories to average out the natural climate variability. This approach is however constrained by the high cost of the experimental setup.

In the different ecotron units, we assume that the controlled variables ( $\text{CO}_2$  and  $\text{CH}_4$  concentration, air temperature, precipitation, atmospheric humidity, wind, ...) are in equilibrium with the warming level, by extracting the 5-year period in which the GMT anomaly in the driving GCM is reached. While this is a reasonable assumption, several components in the climate system will not yet be in equilibrium with the GMT anomaly at the time of simulation (e.g., glaciers, ice sheets, sea level; Zekollari 2019; Church et al. 2013). Therefore, we cannot rule out that changes in these slower components may still affect the meteorological conditions until these reach equilibrium too. For instance, a delayed melting of sea ice could alter the polar circulation and thereby affecting the mid-latitude circulation (Coumou et al. 2018), whereas ice sheet melting may affect oceanic pole-ward heat transport (Caesar et al. 2018). However, to select the time windows, we follow the same approach as the transient response to cumulative emissions (TRCE) as presented in the Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (IPCC 2013). This concept describes the warming per unit of carbon emissions, which largely follows a linear relationship independent of the emission scenario (Knutti and Rogelj 2015).

The setup of the UHasselt Ecotron experiment implies that the incoming shortwave radiation will follow current weather conditions and not the weather conditions as prescribed by the RCM forcing. It is thus possible to have, for instance, clear-sky conditions and associated high incoming shortwave radiation in the field, while in the ecotron unit, a heavy precipitation event is simulated consistent with the RCM forcing. In this example, the system receives more incoming shortwave radiation than in the simulated climate. Likewise, the surface fluxes will be higher, but the resulting air temperature and moisture are corrected within the ecotron unit by the controlling devices to fully follow the boundary layer conditions as they are prescribed by the RCM.

The UHasselt Ecotron experiment allows to investigate ecosystem responses to different levels of climate change. This allows to study subtle changes in ecosystem responses such as impacts of decreased frost frequency on plant mortality (Berendse et al. 1994) and the interactions between the occurrence of mild droughts and plant acclimation for longer droughts (Backhaus et al. 2014). Although climate variables are prescribed, ecosystem-climate feedbacks originating from interactions between the

biosphere and atmosphere can be partially diagnosed. For instance, heatwave reinforcements by occurring droughts (Seneviratne et al. 2010; Zscheischler and Seneviratne 2017) as well as soil moisture effects on precipitation events (Guilod et al. 2015) may be assessed by calculating imbalances in the energy budget.

## Conclusions

Ecosystem experiments investigating climate change responses require a holistic, realistic climate forcing, reflecting not only the changes in the mean climate, but also representing physically consistent co-variance between climate drivers, natural variability, and changes in extreme events. To this extent, we presented a new method for creating realistic climate forcing for manipulation experiments using a single RCM simulation, and subsequently applied it on the UHasselt Ecotron experiment. To account for co-variances between variables and to fully capture the climate variability including extreme events, we selected an RCM simulation from the EURO-CORDEX ensemble based on the following criteria: (i) high skill in the local present-day climate and (ii) representative of local changes in the multi-model mean.

Based on a thorough evaluation of four key variables (air temperature, precipitation, relative humidity, and wind speed), we found that there is no single RCM-GCM combination outperforming all others for all considered variables and metrics. We made a selection of the six best-performing simulations as potential candidates and verified whether they represent the multi-model mean for the considered variables. As precipitation is considered, the most important variable in ecosystem experiments, and as most GCM downscalings have a large bias for this variable, we use the precipitation bias as the decisive factor to single out the simulation which will serve as forcing: CCLM4-8-17 driven by EC-EARTH.

The units of the UHasselt Ecotron experiment are forced with climate conditions along a global mean air temperature (GMT) anomaly gradient, representing conditions of a 0 °C (historical), +1 °C (present day), +1.5 °C, +2 °C, +3 °C, and +4 °C warmer world. Five-year time windows corresponding to these warming levels are defined based on when the 30-year averaged GMT anomaly of EC-EARTH, the driving GCM, crosses these air temperature thresholds. Subsequently, the ecotron forcing is extracted from the 3-hourly RCM simulation according to the time windows.

Our new methodology provides realistic climate forcing, accounting for co-variances between climatic variables and their change in variability, well representing possible compound events. This is particularly interesting for controlled environment facilities, as their setup allows

to realistically simulate future climate by controlling and measuring multiple parameters. Other controlled environment facilities could also benefit from the proposed methodology, depending on the posed research questions. The protocol for selecting a suitable regional climate simulation and extracting time series for the needed variables based on the time window defined by a global mean air temperature threshold provides a framework for different types of manipulation experiments aiming to investigate ecosystem responses to a realistic future climate change, even without a gradient approach.

**Acknowledgments** We also thank RLKM (Regional Landscape Kempen and Maasland) for its collaboration and support. We are grateful to the World Climate Research Programme (WRCP) for initiating and coordinating the EURO-CORDEX initiative, to the modelling centers for making their downscaling results publicly available through ESGF. Computational resources and services were provided by the Shared ICT Services Centre funded by the Vrije Universiteit Brussel, the Flemish Supercomputer Center (VSC) and FWO.

**Author contributions** Jakob Zschleischler, Lukas Gudmundsson, Francois Rineau, Wim Thiery, Natalie Beenaerts, and Jaco Vangronsveld conceived the ideas and designed the methodology. Klaus Keuler provided the 3-hourly simulation data. Inne Vanderkelen performed the analysis. Inne Vanderkelen and Wim Thiery led the writing of the manuscript, with major contributions from Sara Vica, Jakob Zschleischler, and Francois Rineau and input from all other authors. All authors critically revised the draft and gave final approval for publication.

**Funding information** Inne Vanderkelen is a research fellow at the Research Foundation Flanders (FWOTM920). Wim Thiery was supported by an ETH Zurich postdoctoral fellowship (Fel-45 15-1). The Uniscientia Foundation and the ETH Zurich Foundation are thanked for their support to this research. The authors thank the Flemish government (through Hercules stichting big infrastructure and the Fund for Scientific Research Flanders project G0H4117N), LSM (Limburg Sterk Merk, project 271) for providing funds to build the UHasselt Ecotron; Hasselt University for both funding and policy support (project BOF12BR01 and Methusalem project 08M03VGRJ); and the ecotron research committee for useful comments on the experimental design.

## Compliance with ethical standards

**Competing interests** The authors declare that they have no conflict of interest.

**Availability of data and material** Reference station data of the European Climate Assessment and Dataset is publicly available at <https://www.ecad.eu/>. The greenhouse gas concentrations as prescribed by RCP 8.5 are available at <https://tntcat.iiasa.ac.at/RcpDb/>. Data from the Coordinated Regional Climate Downscaling Experiment (CORDEX) Africa framework is available at <http://cordex.org/data-access/esgf/>.

**Code availability** The scripts used in the analysis are available on Github: [https://github.com/VUB-HYDR/2020\\_Vanderkelen.etal.IJBM](https://github.com/VUB-HYDR/2020_Vanderkelen.etal.IJBM).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as

long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Backhaus S., Kreyling J., Grant K., Beierkuhnlein C., Walter J., Jentsch A. (2014) Recurrent mild drought events increase resistance toward extreme drought stress. *Ecosystems* 17(6):1068–1081. <https://doi.org/10.1007/s10021-014-9781-5>
- Berendse F., Schmitz M., de Visser W. (1994) Experimental manipulation of succession in heathland ecosystems. *Oecologia* 100(1-2):38–44. <https://doi.org/10.1007/BF00317128>
- Brier G. W. (1950) Verification of forecasts in terms of probability. *Mon Weather Rev* 78:1–3
- Caesar L., Rahmstorf S., Robinson A., Feulner G., Saba V. (2018) Observed fingerprint of a weakening Atlantic Ocean overturning circulation. *Nature* 556(7700):191–196. <https://doi.org/10.1038/s41586-018-0006-5>
- Casanueva A., Kotlarski S., Herrera S., Fernández J, Gutiérrez JM, Boberg F., Colette A., Christensen O. B., Goergen K., Jacob D., Keuler K., Nikulin G., Teichmann C., Vautard R. (2016) Daily precipitation statistics in a EURO-CORDEX RCM ensemble: added value of raw and bias-corrected high-resolution simulations. *Clim Dyn* 47(3-4):719–737. <https://doi.org/10.1007/s00382-015-2865-x>
- Church J. A., Clark P. U., Cazenave A., Gregory J. M., Jevrejeva S., Levermann M. A., Milne G. A., Nerem R. S., Nunn P. D., Payne A. J., Pfeffer W. T., Stammer D., Unnikrishnan A. S. (2013) Sea level change. Cambridge University Press, Cambridge
- Clobert J., Chanzy A., Le Galliard J. F., Chabbi A., Greiveldinger L., Caquet T., Loreau M., Mouglin C., Pichot C., Roy J., Saint-André L. (2018) How to integrate experimental research approaches in ecological and environmental studies: AnaEE France as an example. *Front Ecol Evol* 6:43. <https://doi.org/10.3389/fevo.2018.00043>
- Coumou D., Di Capua G., Vavrus S., Wang L., Wang S. (2018) The influence of arctic amplification on mid-latitude summer circulation. <https://doi.org/10.1038/s41467-018-05256-8>
- Curtis P. S., Wang X. (1998) A meta-analysis of elevated CO<sub>2</sub> effects on woody plant mass, form, and physiology. *Oecologia* 113:299–313. <https://link.springer.com/content/pdf/10.1007%2Fs004420050381.pdf>
- De Boeck H. J., Van De Velde H., De Groote T., Nijs I. (2016) Ideas and perspectives: heat stress: more than hot air. *Biogeosciences* 13(20):5821–5825. <https://doi.org/10.5194/bg-13-5821-2016>
- Estiarte M., Vicca S., Peñuelas J, Bahn M., Beier C., Emmett B. A., Fay P. A., Hanson P. J., Hasibeder R., Kigel J., Kröel-Dulay G, Larsen K. S., Lellei-Kovács E, Limousin J. M., Ogaya R., Ourcival J. M., Reinsch S., Sala O. E., Schmidt I. K., Sternberg M., Tielbörger K, Tietema A., Janssens I. A. (2016) Few multiyear precipitation-reduction experiments find a shift in the productivity-precipitation relationship. *Glob Chang Biol* 22(7):2570–2581. <https://doi.org/10.1111/gcb.13269>
- Fischer E. M., Knutti R. (2015) Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nat Clim Chang* 5(6):560–564. <https://doi.org/10.1038/nclimate2617>



- Fischer E. M., Sedláček J., Hawkins E., Knutti R. (2014) Models agree on forced response pattern of precipitation and temperature extremes. *Geophys Res Lett* 41(23):8554–8562. <https://doi.org/10.1002/2014GL062018>
- Flato G., Marotzke J., Abiodun B., Braconnot P., Chou S. C., Collins W., Cox P., Driouech F., Emori S., Eyring V., Forst C., Gleckler P., Guilyardi E., Jakob C., Kattsov C., Reason V., Rummukainen M. (2013) Evaluation of climate models. In: Stocker TF, Qin D, Plattner GK, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) *Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge University Press, Cambridge, pp 741–866
- Greve P., Gudmundsson L., Seneviratne S. I. (2018) Regional scaling of annual mean precipitation and water availability with global temperature change. *Earth Syst Dynam* 9(1):227–240. <https://doi.org/10.5194/esd-9-227-2018>
- Griffin K. L., Ross P. D., Sims D. A., Luo Y., Seemann J. R., Fox C. A., Ball J. T. (1996) EcoCELLs: tools for mesocosm scale measurements of gas exchange. *Plant Cell Environ* 19(10):1210–1221. <https://doi.org/10.1111/j.1365-3040.1996.tb00437.x>
- Gudmundsson L., Bremnes J. B., Haugen J. E., Engen-Skaugen T. (2012) Technical note: downscaling RCM precipitation to the station scale using statistical transformations - a comparison of methods. *Hydrol Earth Syst Sci* 16(9):3383–3390. <https://doi.org/10.5194/hess-16-3383-2012>
- Guillod B. P., Orlowsky B., Miralles D. G., Teuling A. J., Seneviratne S. I. (2015) Reconciling spatial and temporal soil moisture effects on afternoon rainfall. *Nat Commun* 6:6443. <https://doi.org/10.1038/ncomms7443>
- Hovenden M. J., Newton P. C., Wills K. E. (2014) Seasonal not annual rainfall determines grassland biomass response to carbon dioxide. *Nature* 511(7511):583–586. <https://doi.org/10.1038/nature13281>
- IPCC (2013) Summary for policymakers. In: *Climate change 2013 - the physical science basis, vol 1542*. Cambridge University Press, Cambridge, pp 1–30. <https://doi.org/10.1017/CBO9781107415324.004>. [https://www.cambridge.org/core/product/identifier/CBO9781107415324A009/type/book\\_part](https://www.cambridge.org/core/product/identifier/CBO9781107415324A009/type/book_part), arXiv:1011.1669v3
- Jacob D., Petersen J., Eggert B., Alias A., Christensen O. B., Bouwer L. M., Braun A., Colette A., Déqué M., Georgievski G., Georgopoulou E., Gobiet A., Menut L., Nikulin G., Haensler A., Hempelmann N., Jones C., Keuler K., Kovats S., Kröner N., Kotlarski S., Kriegsmann A., Martin E., van Meijgaard E., Moseley C., Pfeifer S., Preuschmann S., Radermacher C., Radtke K., Reich D., Rounsevell M., Samuelsson P., Somot S., Soussana J. F., Teichmann C., Valentini R., Vautard R., Weber B., Yiou P. (2014) EURO-CORDEX: new high-resolution climate change projections for European impact research. *Reg Environ Chang* 14(2):563–578. <https://doi.org/10.1007/s10113-013-0499-2>
- Kardol P., De Long J., Sundgqvist M. (2012) Crossing the threshold: the power of multi-level experiments in identifying global change responses. *New Phytol* 196(2):323–326
- Karlowsky S., Augusti A., Ingrisch J., Hasibeder R., Lange M., Lavorel S., Bahn M., Gleixner G. (2018) Land use in mountain grasslands alters drought response and recovery of carbon allocation and plant-microbial interactions. *J Ecol* 106(3):1230–1243. <https://doi.org/10.1111/1365-2745.12910>
- Keuler K., Radtke K., Kotlarski S., Lüthi D. (2016) Regional climate change over Europe in COSMO-CLM: influence of emission scenario and driving global model. *Meteorol Z* 25(2):121–136. <https://doi.org/10.1127/metz/2016/0662>
- Klein Tank A. M., Wijngaard JB, Können GP, Böhm R, Demarée G, Gocheva A., Miletta M., Pashiardis S., Hejkrlik L., Kern-Hansen C., Heino R., Bessemoulin P., Müller-Westermeier G, Tzanakou M., Szalai S., Pálsdóttir T, Fitzgerald D., Rubin S., Capaldo M., Maugeri M., Leitass A., Bukantis A., Aberfeld R., Van Engelen A. F., Forland E., Miletus M., Coelho F., Mares C., Razuvayev V., Nieplova E., Cegnar T., Antonio López J, Dahlström B, Moberg A., Kirchhofer W., Ceylan A., Pachaliuk O., Alexander L. V., Petrovic P. (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Int J Climatol* 22(12):1441–1453. <https://doi.org/10.1002/joc.773>
- Knapp A. K., Carroll C. J., Griffin-Nolan R. J., Slette I. J., Chaves F. A., Baur L. E., Felton A. J., Gray J. E., Hoffman A. M., Lemoine N. P., Mao W., Post A. K., Smith M. D. (2018) A reality check for climate change experiments: do they reflect the real world? *Ecology* 99(10):2145–2151. <https://doi.org/10.1002/ecy.2474>
- Knutti R., Rogelj J. (2015) The legacy of our CO2 emissions: a clash of scientific facts, politics and ethics. *Clim Chang* 133(3):361–373. <https://doi.org/10.1007/s10584-015-1340-3>
- Korell L., Auge H., Chase J. M., Harpole S., Knight T. M. (2019) We need more realistic climate change experiments for understanding ecosystems of the future. *Glob Chang Biol* 00:1–3. <https://doi.org/10.1111/gcb.14797>
- Kotlarski S., Keuler K., Christensen O. B., Colette A., Déqué M, Gobiet A., Goergen K., Jacob D., Lüthi D, Van Meijgaard E., Nikulin G., Schär C, Teichmann C., Vautard R., Warrach-Sagi K., Wulfmeyer V. (2014) Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci Model Dev* 7(4):1297–1333. <https://doi.org/10.5194/gmd-7-1297-2014>
- Kreyling J., Schweiger A. H., Bahn M., Ineson P., Migliavacca M., Morel-Journel T., Christiansen J. R., Schtickzelle N., Larsen K. S. (2018) To replicate, or not to replicate – that is the question: how to tackle nonlinear responses in ecological experiments. *Ecol Lett* 21(11):1629–1638. <https://doi.org/10.1111/ele.13134>
- Lawton J. H. (1993) The Ecotron: a controlled environmental facility for the investigation of population and ecosystem processes. *Philos Trans R Soc Lond B* 341(1296):181–194. <https://doi.org/10.1098/rstb.1993.0102>
- Lawton J. H. (1996) The ecotron facility at silwood park: the value of “big bottle” experiments. *Ecol* 77(3):665–669. <https://doi.org/10.2307/2265488>
- Lin D., Xia J., Wan S. (2010) Climate warming and biomass accumulation of terrestrial plants: a meta-analysis. *New Phytol* 188(1):187–198. <https://doi.org/10.1111/j.1469-8137.2010.03347.x>
- Meehl G. A., Goddard L., Boer G., Burgman R., Branstator G., Cassou C., Corti S., Danabasoglu G., Doblas-Reyes F., Hawkins E., Karspeck A., Kimoto M., Kumar A., Matei D., Mignot J., Msadek R., Navarra A., Pohlmann H., Rienecker M., Rosati T., Schneider E., Smith D., Sutton R., Teng H., Van Oldenborgh G. J., Vecchi G., Yeager S. (2014) Decadal climate prediction an update from the trenches. *Bull Am Meteorol Soc* 95(2):243–267. <https://doi.org/10.1175/BAMS-D-12-00241.1>
- Moon H., Guillod B. P., Gudmundsson L., Seneviratne S. I. (2019) Soil moisture effects on afternoon precipitation occurrence in current climate models. *Geophys Res Lett* 46(3):1861–1869. <https://doi.org/10.1029/2018GL080879>
- Murphy A. H. (1973) A new vector partition of the probability score. *J Appl Meteorol* 12:595–600
- Orlowsky B., Seneviratne S. I. (2013) Elusive drought: uncertainty in observed trends and short-and long-term CMIP5 projections. *Hydrol Earth Syst Sci* 17(5):1765–1781. <https://doi.org/10.5194/hess-17-1765-2013>
- Perkins S. E., Pitman A. J., Holbrook N. J., McAneney J. (2007) Evaluation of the AR4 climate models’ simulated daily maximum temperature, minimum temperature, and precipitation

- over Australia using probability density functions. *J Clim* 20(17):4356–4376. <https://doi.org/10.1175/JCLI4253.1>
- Rajczak J., Schär C. (2017) Projections of future precipitation extremes over Europe: a multimodel assessment of climate simulations. *J Geophys Res Atmos* 122(20):10,773–10,800. <https://doi.org/10.1002/2017JD027176>
- Riahi K., Rao S., Krey V., Cho C., Chirkov V., Fischer G., Kindermann G., Nakicenovic N., Rafaj P. (2011) RCP 8.5-A Scenario of comparatively high greenhouse gas emissions. *Clim Chang* 109(1):33–57. <https://doi.org/10.1007/s10584-011-0149-y>
- Rineau F., Malina R., Beenaerts N., Arnauts N., Bardgett R. D., Berg M. P., Boerema A., Bruckers L., Clerinx J., Davin E. L., De Boeck H. J., De Dobbelaer T., Dondini M., De Laender F., Ellers J., Franken O., Gilbert L., Gudmundsson L., Janssens I. A., Johnson D., Lizin S., Longdoz B., Meire P., Meremans D., Milbau A., Moretti M., Nijs I., Nobel A., Pop I. S., Puetz T., Reyns W., Roy J., Schuetz J., Seneviratne S. I., Smith P., Solmi F., Staes J., Thiery W., Thijs S., Vanderkelen I., Van Landuyt W., Verbruggen E., Witters N., Zscheischler J., Vangronsveld J. (2019) Supplementary materials: towards more predictive and interdisciplinary climate change ecosystem experiments. *Nat Clim Chang* 9(11):809–816. <https://doi.org/10.1038/s41558-019-0609-3>
- Roy J., Picon-Cochard C., Augusti A., Benot M. L., Thiery L., Darsenville O., Landais D., Piel C., Defossez M., Devidal S., Escape C., Ravel O., Fromin N., Volaire F., Milcu A., Bahn M., Soussana J. F. (2016) Elevated CO<sub>2</sub> maintains grassland net carbon uptake under a future heat and drought extreme. *Proc Natl Acad Sci* 113(22):6224–6229. <https://doi.org/10.1073/pnas.1524527113>
- Rustad L., Campbell J., Marion G., Norby R., Mitchell M., Hartley A., Cornelissen J., Gurevitch J. (2001) A meta-analysis of the response of soil respiration, net nitrogen mineralization, and aboveground plant growth to experimental ecosystem warming. *Oecologia* 126:543–562. <https://doi.org/10.1007/s004420000544s004420000544>. <https://link.springer.com/content/pdf/10.1007%2Fs004420000544.pdf>
- Savitzky A., Golay M. J. (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem* 36(8):1627–1639. <https://doi.org/10.1021/ac60214a047>
- Seneviratne S. I., Corti T., Davin E. L., Hirschi M., Jaeger E. B., Lehner I., Orlowsky B., Teuling A. J. (2010) Investigating soil moisture-climate interactions in a changing climate: a review. <https://doi.org/10.1016/j.earscirev.2010.02.004>
- Seneviratne S. I., Donat M. G., Pitman A. J., Knutti R., Wilby R. L. (2016) Allowable CO<sub>2</sub> emissions based on regional and impact-related climate targets. <https://doi.org/10.1038/nature16542>
- Sillmann J., Kharin V. V., Zwiers F. W., Zhang X., Bronaugh D. (2013) Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *J Geophys Res Atmos* 118(6):2473–2493. <https://doi.org/10.1002/jgrd.50188>
- Stewart R. I., Dossena M., Bohan D. A., Jeppesen E., Kordas R. L., Ledger M. E., Meerhoff M., Moss B., Mulder C., Shurin J. B., Suttle B., Thompson R., Trimmer M., Woodward G. (2013) Mesocosm experiments as a tool for ecological climate-change research. *Adv Ecol Res* 48:71–181. <https://doi.org/10.1016/B978-0-12-417199-2.00002-1>
- Terrer C., Vicca S., Stocker B. D., Hungate B. A., Phillips R. P., Reich P. B., Finzi A. C., Prentice I. C. (2018) Ecosystem responses to elevated CO<sub>2</sub> governed by plant–soil interactions and the cost of nitrogen acquisition. <https://doi.org/10.1111/nph.14872>
- Thompson R. M., Beardall J., Beringer J., Grace M., Sardina P. (2013) Means and extremes: building variability into community level climate change experiments. *Ecol Lett* 16(6):799–806. <https://doi.org/10.1111/ele.12095>
- Van der Molen M. K., Dolman A. J., Ciais P., Eglin T., Gobron N., Law B. E., Meir P., Peters W., Phillips O. L., Reichstein M., Chen T., Dekker S. C., Doubkova M., Friedl M. A., Jung M., van den Hurk B. J., de Jeu R. A., Kruijft B., Ohta T., Rebel K. T., Plummer S., Seneviratne S. I., Sitch S., Teuling A. J., van der Werf G. R., Wang G. (2011) Drought and ecosystem carbon cycling. <https://doi.org/10.1016/j.agrformet.2011.01.018>
- Vanderkelen I., van Lipzig N. P. M., Thiery W. (2018) Modelling the water balance of Lake Victoria (East Africa) - part 2: future projections. *Hydrol Earth Syst Sci* 22(10):5527–5549. <https://doi.org/10.5194/hess-22-5527-2018>
- Vicca S., Bahn M., Estiarte M., Van Loon E. E., Vargas R., Alberti G., Ambus P., Arain M. A., Beier C., Bentley L. P., Borken W., Buchmann N., Collins S. L., De Dato G., Dukes J. S., Escobar C., Fay P., Guidolotti G., Hanson P. J., Kahmen A., Kröel-Dulay G., Ladreiter-Knauss T., Larsen K. S., Lellei-Kovacs E., Lebrija-Trejos E., Maestre F. T., Marhan S., Marshall M., Meir P., Miao Y., Muhr J., Niklaus P. A., Ogaya R., Peñuelas J., Poll C., Rustad L. E., Savage K., Schindlbacher A., Schmidt I. K., Smith A. R., Sotta E. D., Suseela V., Tietema A., Van Gestel N., Van Straaten O., Wan S., Weber U., Janssens I. A. (2014) Can current moisture responses predict soil CO<sub>2</sub> efflux under altered precipitation regimes? A synthesis of manipulation experiments. *Biogeosciences* 11(11):2991–3013. <https://doi.org/10.5194/bg-11-2991-2014>
- van Vuuren D. P., Edmonds J., Kainuma M., Riahi K., Thomson A., Hibbard K., Hurtt G. C., Kram T., Krey V., Lamarque J. F., Masui T., Meinshausen M., Nakicenovic N., Smith S. J., Rose S. K. (2011) The representative concentration pathways: an overview. *Clim Chang* 109(1):5–31. <https://doi.org/10.1007/s10584-011-0148-z>. 9605103
- Wu Z., Dijkstra P., Koch G. W., Peñuelas J., Hungate B. A. (2011) Responses of terrestrial ecosystems to temperature and precipitation change: a meta-analysis of experimental manipulation. *Glob Chang Biol* 17(2):927–942. <https://doi.org/10.1111/j.1365-2486.2010.02302.x>
- Yue K., Fornara D. A., Yang W., Peng Y., Peng C., Liu Z., Wu F. (2017) Influence of multiple global change drivers on terrestrial carbon storage: additive effects are common. *Ecol Lett* 20(5):663–672. <https://doi.org/10.1111/ele.12767>
- Zekollari H., Huss M., Farinotti D. (2019) Modelling the future evolution of glaciers in the European Alps under the EURO-CORDEX RCM ensemble. *Cryosphere* 13(4):1125–1146. <https://doi.org/10.5194/tc-13-1125-2019>
- Zhang X., Zwiers F. W., Hegerl G. (2009) The influences of data precision on the calculation of temperature percentile indices. *Int J Climatol* 29(3):321–327. <https://doi.org/10.1002/joc.1738>
- Zscheischler J., Seneviratne S. I. (2017) Dependence of drivers affects risks associated with compound events. *Science Advances* 3(6):e1700263. <https://doi.org/10.1126/sciadv.1700263>
- Zscheischler J., Westra S., van den Hurk B. J. J. M., Seneviratne S. I., Ward P. J., Pitman A., AghaKouchak A., Bresch D. N., Leonard M., Wahl T., Zhang X. (2018) Future climate risk from compound events. *Nat Clim Change* 8:469–477. <https://doi.org/10.1038/s41558-018-0156-3>
- Zscheischler J., Fischer E. M., Lange S. (2019) The effect of univariate bias adjustment on multivariate hazard estimates. *Earth Syst Dynam* 10(1):31–43. <https://doi.org/10.5194/esd-10-31-2019>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Affiliations

Inne Vanderkelen<sup>1</sup>  · Jakob Zscheischler<sup>2,3</sup> · Lukas Gudmundsson<sup>4</sup> · Klaus Keuler<sup>5</sup> · Francois Rineau<sup>6</sup> · Natalie Beenaerts<sup>6</sup> · Jaco Vangronsveld<sup>6,7</sup> · Sara Vicca<sup>8</sup> · Wim Thiery<sup>1,4</sup>

- <sup>1</sup> Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussels, Belgium
- <sup>2</sup> Climate and Environmental Physics, University of Bern, Bern, Switzerland
- <sup>3</sup> Oeschger Center for Climate Change Research, University of Bern, Bern, Switzerland
- <sup>4</sup> Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland
- <sup>5</sup> Department of Environmental Meteorology, Brandenburg University of Technology Cottbus-Senftenberg, Cottbus, Germany
- <sup>6</sup> Centre for Environmental Sciences, UHasselt, Hasselt, Belgium
- <sup>7</sup> Department of Plant Physiology, Faculty of Biology and Biotechnology, Maria Curie-Skłodowska University, Lublin, Poland
- <sup>8</sup> Department of Biology, University of Antwerp, Wilrijk, Belgium