A new approach for assessing climate change impacts in ecotron experiments

Biogeosciences

January 30, 2020

Inne Vanderkelen[1], Jakob Zschleischler[2, 3], Lukas Gudmundsson [4], Klaus Keuler [5], Francois Rineau [6], Natalie Beenaerts [6], Jaco Vangronsveld [6, 7], Sara Vicca [8], Wim Thiery[1, 4]
inne.vanderkelen@vub.be

[1] Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium.
[2] Climate and Environmental Physics, University of Bern, Bern, Switzerland.
[3] Oeschger Center for Climate Change Research, University of Bern, Bern, Switzerland.
[4] Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland.
[5] Department of Environmental Meteorology, Brandenburg University of Technology Cottbus-Senftenberg, Cottbus, Germany.
[6] Centre for Environmental Sciences, UHasselt, Hasselt, Belgium.
[7] Department of Plant Physiology, Faculty of Biology and Biotechnology, Maria Curie-Sklodowska University, Lublin, Poland.
[8] Department of Biology, University of Antwerp, Wilrijk, Belgium.

Contents

Abstract

This response letter contains numbered figures and references to these figures. To prevent confusion, the figures embedded within this response letter are called illustrations. Finally, the following convention is applied to denote modification in the original manuscript: new text.

1 Reviewer 1

Reviewer 1 Comment 1

This review focuses mainly on the design and evaluation of the UHasselt Ecotron Experiment, as was requested by the editor. However, I did read through the entire manuscript.

Response

We thank Reviewer 1 for his/her review of the study and the UHasselt Ecotron Experiment and the important issues raised. We understand the confusion on the main objective of this paper, which is not describing the UHasselt Ecotron experiment, but to depict a new methodology to provide climate forcing to ecosystem experiments. To this extent, the UHasselt ecotron experiment serves as an example application of the new methodology. We now clarified the objectives in the manuscript:

> In this paper, we present a protocol for creating realistic climate forcing for manipulation experiments.
>
> [...]
>
> In our methodology, variability and co-variance between variables is 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.

Below, we carefully address every comment and explain the corresponding changes in the manuscript.

Reviewer 1 Comment 2

I am unclear about the three compartments that the authors refer to when describing their ecotron facility. They cite Rineau et al (in review) which apparently describes these, and other essential details (e.g., like which ecosystem processes will be measured and how they will be measured). I have no access to this paper.

The Rineau et al. (2019) paper appeared after this review in Nature Climate Change. The reference to this paper is updated throughout the manuscript. We also added the Rineau et al. paper and its supplementary material as an appendix to this authors' response. While this study includes a more detailed description of the facility, we summarise the most important features in our current study.

We updated figure 2 in our manuscript to include a schematic overview on the three different compartments (dome, chamber and lysimeter, illustration 1).

The ecosystem processes which will be measured are listed in figure S4 of the supplementary material of Rineau et al. (2019), and we copy the table here below for reference (illustration 6). We now added some examples of which ecosystem processes will be measured in the manuscript:

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, CH4 and N2O emissions.

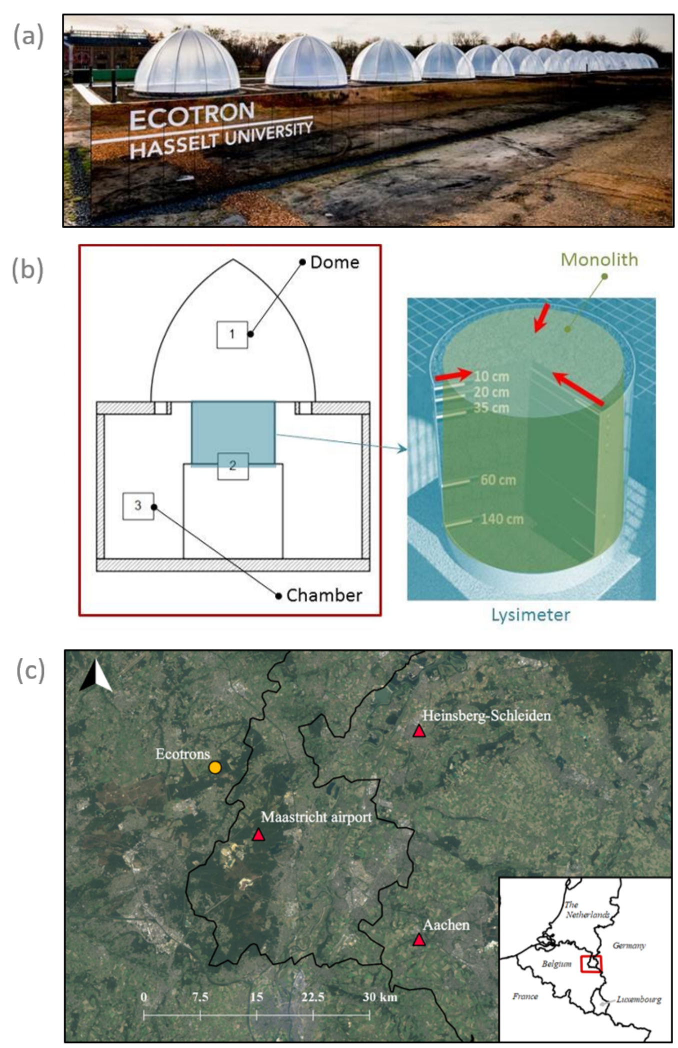

Illustration 1: The UHasselt Ecotron experiment (a; picture: Liesbeth Driessen), scheme of a unit with the three compartments and the lysimeter compartment in detail (b), and overview map with location of the infrastructure and reference weather observation stations (c).

Compartment	Measurement	Measurement frequency
Atmospheric	Air temperature *	1'
	Relative humidity *	1'
	Wind speed *	30'
	Precipitation *	10'
	CO_2 concentration *	30'
	CH_4 concentration	30'
	N_2O concentration	30'
	Air pressure	30'
	Incoming / reflected radiation	1'
	Photosynthetically active radiation	1'
Soil	Temperature *	10'
	Water tension *	10'
	Water content	10'
	Electrical conductivity	10'
	Pore CO2	10'
	Weight	1'
	Water leaching	1'
	(suction cups for water sampling)	2 weeks

Illustration 2: Controlled and measured parameters in one macrocosm. The asterix marked variables are both controlled and measured, the others only monitored. Directly taken from Rineau et al. (2019), supplementary materials.

Reviewer 1 Comment 3

Regardless, there is sufficient information in the methods of the manuscript to give me pause and concern. With 12 ecotron units, and what looks to be 12 individual treatment combinations, it appears that only one macrocosm will be used per treatment combination, with no experimental replication. This looks like a so-called "regression design". These designs are fine. However, the absence of spatial replication makes it essential to obtain robust baseline ecosystem response conditions under "control" conditions (i.e. the conditions under which the control macrocosm in the spatially-unreplicated experiment will be maintained). A robust baseline for a multi-year study would require using the first year of the study to obtain/quantify the particular "behavioral personalities" of each individual (each of the 12) macrocosm. Then, only once each "personality" has been measured, can a rigorous assessment of treatments be reliably measured in the following 4-5 years. Without such a pre-assessment, it will be impossible to know whether treatment responsesâevaluated against a single "control" macrocosmâare due to the treatment(s) or to an anomalous "control macrocosm" (analogous to a random "crazy personality"). This is a really critical need, and critical shortfall in the study design, as I understand it, and should be addressed.

Response

The experiment set-up follows indeed a "regression design". Like the reviewer commented, small initial differences in small-scale soil heterogeneity between different units can increase to the point that the unit becomes statistically different from the others (Rineau et al., 2019). As the reviewer correctly points out, this can prevent interpreting results from individual units like robust baseline ecosystem responses, as there is no replication of the experiment. Therefore, a cluster analysis has been conducted, which quantifies the natural variation of the 14 measured variables between the Ecotron units during 11 months prior to the experiment (illustration 3, from Rineau et al. (2019) supplementary material). Based on this analysis, the units are distributed in two gradient classes, minimizing the natural variance (noise). The resulting unit distribution over the two gradients is illustrated in Rineau et al. (2019) by figure 1 (see illustration 4 below).

We thank the reviewer for pointing out that our manuscript is lacking this information. To accommodate this, we updated the manuscript as follows:

> The climatology of the unit forced by +1°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).

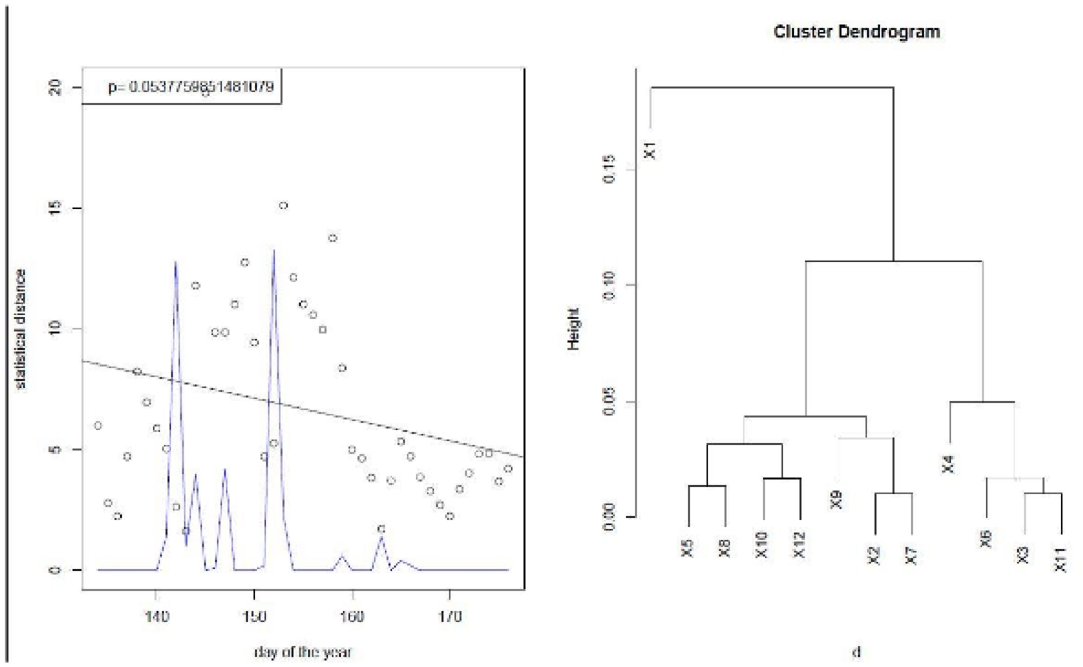

Figure S2. Cluster analysis of the ecotron units. The measurements in each unit were analyzed for 11 months before the start of the experiment to evaluate the level of noise (natural variation) between units and the evolution of this noise with time. For this purpose, the 14 variables measured in the Ecotron (see Fig. S1) were aggregated daily based on their average value in each unit. A unit-to-unit distance matrix was then computed and the statistical distance between units for every day was calculated as the sum of these distances, which was then plotted against time; the increase in distance with time was then tested through a linear model. A cluster analysis was run every day to check how units were grouped using complete linkage method. Clusters were very consistent throughout the testing period, so we only present the results of the endpoint. Statistical analyses were done using R. Results showed that statistical distance was not significantly increasing through time (left panel) and that the units were separated into several clusters of different sizes (right panel).

Illustration 3: Cluster analysis of the ecotron units and explanation, directly taken from Rineau et al. (2019), supplementary materials, figure S2.

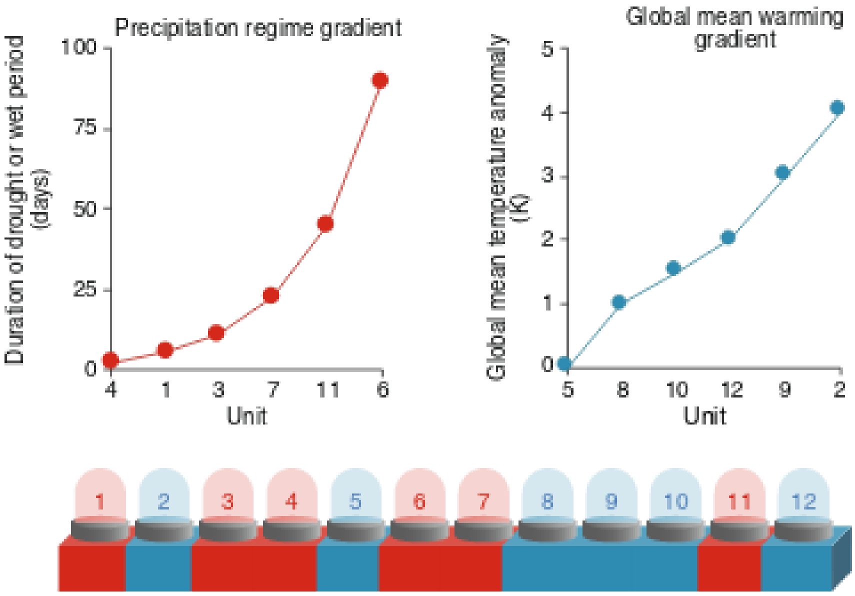

Illustration 4: Distribution of the ecotron units over the two gradients. Figure taken from Rineau et al. (2019)

Reviewer 1 Comment 4

Perhaps I missed this, but I also did not see any description of how the empirical data collected from the 12-ecotron experiment would be statistically analyzed, nor did I see any specific research questions or hypotheses articulated.

Response

The measurements of the ecotron experiment will be analysed by a broad interdisciplinary framework, ranging from plant ecologists, mathematicians, hydrologists, microbial and fauna ecologists to ecosystem ecologists and environmental economists (Rineau et al., 2019). Examples of where the Ecotron facility will contribute are the development of numerical models to describe water movement in one ecotron unit, calibration and prediction of a soil-carbon model by the carbon cycling measurements, investigation of soil organisms and their role in the soil bio-geochemistry, water quality regulation, carbon sequestration and quantification of ecosystem services by measurements of soil abiotic parameters (see illustration 5 for an overview of the integration of the different scientific disciplines). Rineau et al. (2019) also included some hypotheses on the possible outcomes and listed them in supplementary figure 3, here included as illustration 6.

We modified the manuscript to highlight the main research questions of the experiment, but did not elaborate this as this is not the main objective of this paper (see reviewer 2).

> 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). 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 forest or acid grassland and what the consequences are for ecosystem services (Rineau et al., 2019).

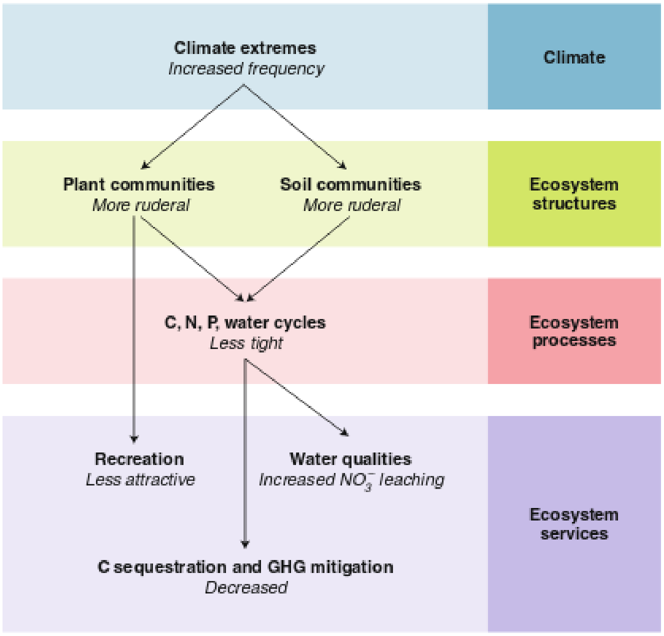

Illustration 5: Overview of hypotheses and integration of scientific disciplines, directly taken from Rineau et al. (2019).

Heathland ecosystems typically feature trajectories that lead to alternative stable states with contrasting functioning (most often birch or pine forest as natural transition, or acid grassland under high nitrogen (N) deposition,[1] and therefore distinct patterns of ecosystem services delivery. Therefore, a key question is how climate change will accelerate or slow down the transition to these alternative stable states, and what will be the consequences for ecosystem services.

We expect that future climate, characterized by more climate extremes, will result in an **increasing frequency of disturbances** for plants and soil organisms, **favoring fast-growing, ruderal** grassy species **at the expense of stress-tolerant** shrub ones, once their tolerance threshold is exceeded.[2] We also hypothesize the same process will affect community composition of fungi, bacteria and soil arthropods,[2,3] albeit at smaller spatiotemporal scales.

As ruderal species are reported to display **lower energy and nutrient-use efficiency** (Aerts & Heil, 1993), and increased frequency of drought/rewetting cycles is associated with nutrient losses,[4] we predict that climate change will lead to **less tight water, C and nutrient cycles** in heathlands, leading to lower C sequestration (grass biomass decomposes about three times faster than that of shrubs)[1] or increased N leakage from the system.

Illustration 6: Rationale and hypotheses of the UHasselt Ecotron Experiment, directly taken from Rineau et al. (2019).

Reviewer 1 Comment 5

I'm wondering whether the problem of the lack of spatial replication could be addressed by reducing the number of treatment combinations to six, so that there would be at least two replicate ecotron units per treatment combination.

Response

The scientific consortium of the UHasselt Ecotron experiment decided to run 2 types of gradients, each existing of 6 ecotron units: a precipitation regime gradient and a global mean warming gradient. As mentioned above, the issue of spatial replication has been partially addressed by comparing soil and plant measurements from all units exposed to the same climate conditions, and statistically cluster those in two homogeneous blocks of 6 units. Reducing the number of treatment combinations to six could indeed be a good compromise to increase statistical power, but is unfortunately practically not possible as two parallel experiments have been foreseen, each occupying 6 units (Rineau et al., 2019). One approach (six units) measures the effect of an altered single factor (here, precipitation regime), while maintaining the natural variation of other abiotic factors, and the other approach (six units) manipulates climate by jointly simulating all co-varying parameters, representing increasingly intense climate change. The second approach is the one described in this paper.

Reviewer 1 Comment 6

I do appreciate the approach of using data from downscaled climate models to guide which experimental treatments to include. I also like the use of real-time ICOS data to incorporate realistic climate variability to some of the treatments. It is my understanding that these models deliver daily (24 h means or sums) resolution data, that would not be suitable to understand sub-daily/diel climate/weather variability. Is that what the ICOS data will be used for? It would certainly be important to retain diel air T, RH, and precipitation patterns in the experiment.

Response

The precipitation gradient experiment uses the ICOS data directly to force all variables except precipitation as this is altered. We clarified the time resolution of the ICOS data coupling with the Ecotron infrastructure in the manuscript:

> 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 minutes (Rineau et al., 2019).

For the RCM data however, the highest temporal resolution available is used, which is 3-hourly. In this way the sub-daily weather variability is accounted for, which is crucial for providing realistic climate forcing. The 3-hourly data is interpolated to the 30-minute resolution to force the ecotrons. The detailed treatment of the different variables is described in the third paragraph of section 3.2.1.

To clarify this more, we adjusted the third paragraph of section 3.1.

> In the remaining six units, atmospheric conditions along the GMT anomaly gradient will be simulated as described in section 2. The 3-hourly RCM output is linearly interpolated to a 30-minute 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 temperatures in the chamber (despite being thermically 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.

Reviewer 1 Comment 7

Taken together, the paper on its own left me with many unanswered questions. These may be covered in the Rineau et al. manuscript. I would recommend placing the essentials of that paper in the next version of this paper, particularly items that address the questions and the issues I have identified above. Thus, based mainly on the section of the manuscript on which I was asked to focus, I feel compelled to rate the decision as "reject" at this stage of the manuscript. I would encourage improving the ms. and resubmitting, with the managing editor's approval.

Response

We understand that some essential details on describing the experiment were missing in the manuscript. As stated above, we added information to address the reviewers comments. In this way, we strive to make the manuscript more self-containing. We thank the reviewer for the critical and constructive review of the experiment.

2 Reviewer 2

Reviewer 2 Comment 1

Dear Dr. Vanderkelen,
it has been extremely difficult to find reviewers for your manuscript, so to make some progress at this point I have decided to provide a review myself.

For several decades controlled environment facilities have been a key approach for studying effects of climate change on plants and small stature ecosystems, and since the 1990ies ecotrons and their application have been repeatedly described (e.g. Lawton et al. 1993 and 1996, Griffin et al. 1996).

There is no doubt that phytotrons and ecotrons are state-of-the-art tools, whose technical capacities, including the controlled volume and the precision, have increased tremendously during the past 15 years. Such infrastructures provide an outstanding possibility to test for individual and interactive effects of multiple global change drivers, and / or to simulate specific scenarios projected by climate models, and there is no doubt that studies based an ecotrons will yield major novel scientific insights. However, from my perspective there is only limited novelty in the description of the facility itself. For this reason, such descriptions have previously been included in the supplements of papers reporting on the actual outcome of the climate manipulation experiments performed (e.g. Arnone et al. 2008, Roy et al. 2016).

From my own background I cannot judge the degree of novelty contained in your new methodology for generating climate forcing using a single Regional Climate Model Simulation, which is one of the reasons why I sent your manuscript out for review. One of the experts on the topic, who I trust, declined my invitation to review your manuscript with the comment "This paper does not look very interesting to me - it merely describes the plan to regulate controlled environments following some very specific climate change predictions.â

Taken together, I am therefore not convinced that your manuscript is advancing the field to a sufficient extent to be acceptable for publication as a full paper in Biogeosciences. I may nevertheless revise my opinion in case you manage to convince me otherwise in your author responses.

Best regards,
Michael Bahn (Editor)

Response

We thank the editor for the extensive search for suited reviewers for this study and the review provided. We acknowledge that the manuscript did not state clear enough its main objective: providing a new methodology to select the most appropriate forcing data for a specific region and research question from a large set of available climate simulations. Below we describe why we believe this methodology is advancing the field, and which changes we applied to the manuscript text to frame this more clearly. An overview of all changes compared to the original manuscript is provided in the *diff.pdf* file attached to the resubmission.

There is a need for realistic climate change experiments to better understand ecosystem responses (Korell et al., 2019; Song et al., 2019). This is not only necessary, but is also appealing and timely for more advanced controlled environment experiments, like the ecotron experiments. Climate forcing that represents realistic future climate conditions is characterized by a realistic co-variance of multiple variables, which are linked to each other by the synoptic condition and a realistic temporal sequence of weather events. In this way, we avoid the simulation of unrealistic situations, e.g. dry conditions with a high relative humidity. Additionally, it is important to correctly represent the evolution of future climate variability, thereby well representing climate extremes. In this way, compound events, a combination of meteorological drivers leading to an even more extreme event, are better represented, which is key for measuring ecosystem responses (Zscheischler and Seneviratne, 2017; Rineau et al., 2019). An example of such a compound event occurred in the summer of 2018: the combination of a drought and heatwave led to massive vegetation dieback in the macrocosms of the UHasselt experiment, which were driven by the ICOS observations from the field. The dieback in the ecotrons had the same extent as the vegetation in the field.

Classical climate change experiments do not consider this co-variance, as they alter one factor, while keeping everything else equal. While these experiments are useful and informative, they remain characterised by several limitations: first, ignoring this co-variance leads to an unrealistic forcing, e.g. simulating a drought without accounting for a change in temperature; while droughts and heatwaves often co-occur (Zscheischler and Seneviratne, 2017). Second, events focused on short-term, hourly variability are not captured in mean climatology or spatially averaged values. Third, it is important to account for potential future changes in climate variability to have a realistic representation of extremes, coherent with future climate projections. Finally, in many experiments the applied temperature, CO2 or precipitation perturbation is derived from coarse resolution (global) climate models, limiting their representativity for a given site (Thompson et al., 2013; Roy et al., 2016; Korell et al., 2019).

Our new methodology overcomes these four shortcomings by sampling forcing data from state-of-the-art regional climate simulations, which are more refined and better in solving meso-scale circulation (Kotlarski et al., 2014; Dosio et al., 2015). By using 3-hourly data for a range of variables from a single regional climate simulation, we strive to provide the most sound forcing possible. Furthermore, we verify that this simulation is representative for the multi-model mean in future simulations, ensuring a realistic temperature sensitivity to future increases in greenhouse gases. The resulting experiment is a prime example of interdisciplinary collaboration between experimentalists (providing the research questions and controlled environment facility) and climate modellers (providing forcing data) described by Rineau et al. (2019) and recently highlighted by Muller et al. (2019).

To better represent these arguments, we now substantially reworked the introduction:

> 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., 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 variables, linked to each other by synoptic conditions. This is particularly important for representing compound events, where a the combination of non extreme drivers can lead to extreme events (Rineau et al., 2019; Zscheischler and Seneviratne, 2017; Zscheischler et al., 2018).
>
> Until recently, it was not possible to simulate realistic 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, 1996; Stewart 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 under the most realistic conditions currently feasible (Rineau et al., 2019; Kreyling et al., 2018).
>
> [...]
>
> In this paper, we present a protocol for creating realistic climate forcing for manipulation experiments.

We also updated section 2 describing the new methodology:

> In our methodology, variability and co-variance between variables is 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 in the ecosystem climate change experiments follow a gradient of increasing Global Mean Temperature (GMT) anomalies.

Finally, we also adapted the conclusions to better reflect the main objective of the study:

> 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, changes in extreme events. To this extent, we presented a new

method for creating realistic climate forcing for manipulation experiments using a single Regional Climate Model (RCM) simulation, and subsequently applied it on the UHasselt Ecotron Experiment.

Our new methodology provides realistic climate forcing, accounting for covariances between climatic variables and their change in variability, well representing possible compound events. This is in particular 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 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.

We very much hope this response helps to show that this manuscript is worth to be considered for publication as full paper in Biogeosciences. If not, we are willing to propose the current manuscript as a technical note.

References

Clobert, J., Chanzy, A., Le Galliard, J.-F., Chabbi, A., Greiveldinger, L., Caquet, T., Loreau, M., Mougin, C., Pichot, C., Roy, J., and Saint-André, L. (2018). How to Integrate Experimental Research Approaches in Ecological and Environmental Studies: AnaEE France as an Example. *Frontiers in Ecology and Evolution*, 6:43.

Dosio, A., Panitz, H. J., Schubert-Frisius, M., and Lüthi, D. (2015). Dynamical downscaling of CMIP5 global circulation models over CORDEX-Africa with COSMO-CLM: evaluation over the present climate and analysis of the added value. *Climate Dynamics*, 44(9-10):2637–2661.

Hovenden, M. J., Newton, P. C., and Wills, K. E. (2014). Seasonal not annual rainfall determines grassland biomass response to carbon dioxide. *Nature*, 511(7511):583–586.

Karlowsky, S., Augusti, A., Ingrisch, J., Hasibeder, R., Lange, M., Lavorel, S., Bahn, M., and Gleixner, G. (2018). Land use in mountain grasslands alters drought response and recovery of carbon allocation and plant-microbial interactions. *Journal of Ecology*, 106(3):1230–1243.

Korell, L., Auge, H., Chase, J. M., Harpole, S., and Knight, T. M. (2019). We need more realistic climate change experiments for understanding ecosystems of the future. *Global Change Biology*, 00:1–3.

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., and Wulfmeyer, V. (2014). Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geoscientific Model Development*, 7(4):1297–1333.

Kreyling, J., Schweiger, A. H., Bahn, M., Ineson, P., Migliavacca, M., Morel-Journel, T., Christiansen, J. R., Schtickzelle, N., and Larsen, K. S. (2018). To replicate, or not to replicate â that is the question: how to tackle nonlinear responses in ecological experiments. *Ecology Letters*, 21(11):1629–1638.

Lawton, J. H. (1996). The ecotron facility at silwood park: The value of "Big Bottle" experiments. *Ecology*, 77(3):665–669.

Muller, L. A., Ballhausen, M. B., Lakovic, M., and Rillig, M. C. (2019). Response to the Editor: Assessing the robustness of communities and ecosystems in global change research. *Global Change Biology*.

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., and Vangronsveld, J. (2019). Supplementary materials: Towards more predictive and interdisciplinary climate change ecosystem experiments. *Nature Climate Change*, 9(11):809–816.

Roy, J., Picon-Cochard, C., Augusti, A., Benot, M.-L., Thiery, L., Darsonville, O., Landais, D., Piel, C., Defossez, M., Devidal, S., Escape, C., Ravel, O., Fromin, N., Volaire, F., Milcu, A., Bahn, M., and Soussana, J.-F. (2016). Elevated CO 2 maintains grassland net carbon uptake under a future heat and drought extreme. *Proceedings of the National Academy of Sciences*, 113(22):6224–6229.

Song, J., Wan, S., Piao, S., Knapp, A. K., Classen, A. T., Vicca, S., Ciais, P., Hovenden, M. J., Leuzinger, S., Beier, C., Kardol, P., Xia, J., Liu, Q., Ru, J., Zhou, Z., Luo, Y., Guo, D., Adam Langley, J., Zscheischler, J., Dukes, J. S., Tang, J., Chen, J., Hofmockel, K. S., Kueppers, L. M., Rustad, L., Liu, L., Smith, M. D., Templer, P. H., Quinn Thomas, R., Norby, R. J., Phillips, R. P., Niu, S., Fatichi, S., Wang, Y., Shao, P., Han, H., Wang, D., Lei, L., Wang, J., Li, X., Zhang, Q., Li, X., Su, F., Liu, B., Yang, F., Ma, G., Li, G., Liu, Y., Liu, Y., Yang, Z., Zhang, K., Miao, Y., Hu, M., Yan, C., Zhang, A., Zhong, M., Hui, Y., Li, Y., and Zheng, M. (2019). A meta-analysis of 1,119 manipulative experiments on terrestrial carbon-cycling responses to global change. *Nature Ecology and Evolution*, 3(9):1309–1320.

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., and Woodward, G. (2013). Mesocosm Experiments as a Tool for Ecological Climate-Change Research. *Advances in Ecological Research*, 48:71–181.

Terrer, C., Vicca, S., Stocker, B. D., Hungate, B. A., Phillips, R. P., Reich, P. B., Finzi, A. C., and Prentice, I. C. (2018). Ecosystem responses to elevated CO2 governed by plantâsoil interactions and the cost of nitrogen acquisition.

Thompson, R. M., Beardall, J., Beringer, J., Grace, M., and Sardina, P. (2013). Means and extremes: Building variability into community-level climate change experiments. *Ecology Letters*, 16(6):799–806.

Zscheischler, J. and Seneviratne, S. I. (2017). Dependence of drivers affects risks associated with compound events. *Science Advances*, 3(6):e1700263.

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., and Zhang, X. (2018). Future climate risk from compound events. *Nature Climate Change*, page 1.

A new approach for assessing climate change impacts in ecotron experiments

Inne Vanderkelen[1], Jakob Zschleischler[2,3], Lukas Gudmundsson[4], Klaus Keuler[5], Francois Rineau[6], Natalie Beenaerts[6], Jaco Vangronsveld[6,7], Sara Vicca[8], and Wim Thiery[1,4]

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

Correspondence: Inne Vanderkelen (inne.vanderkelen@vub.be)

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, preserving the co-variances between variables and the projected changes in variability. Here we present a new ~~experimental design for studying climate change impacts on terrestrial ecosystems~~ 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 (RCM) projection and consists of generating climate forcing along a gradient representative of increasingly high global mean temperature anomalies~~and uses data derived from the best available regional climate model (RCM) projection~~. We first identified the best performing regional climate model (RCM) simulation for the ecotron site from the Coordinated Regional Downscaling Experiment in the European Domain (EURO-CORDEX) ensemble with a 0.11° (12.5 km) resolution based on two criteria: (i) highest skill of the simulations compared to observations from a nearby weather station and (ii) representativeness of the multi-model mean in future projections. Our results reveal that no single RCM simulation has the best score for all possible combinations of the four meteorological variables and evaluation metrics considered. Out of the six best performing simulations, we selected the simulation with the lowest bias for precipitation (CCLM4-8-17/EC-EARTH), as this variable is key to ecosystem functioning and model simulations deviated the most for this variable, with values ranging up to double the observed values. The time window is subsequently selected from the RCM projection for each ecotron unit based on the global mean temperature of the driving Global Climate Model (GCM). The ecotron units are forced with 3-hourly output from the RCM projections of the five-year period spanning the year in which the global mean 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 gradient approach will allow to

identify possible ~~threshold~~ thresholds and tipping points. The presented methodology can also be applied to other manipulation experiments, aiming at investigating ecosystem responses to realistic future climate change.

Copyright statement.

1 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: temperature, precipitation, CO_2 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 ~~have become more common~~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).

~~In most cases, climate change experiments apply step changes to the studied drivers, that is, one factor is increased/decreased by a fixed amount. This makes it difficult to use the obtained results for model development, for which usually gradient responses need to be known. This insight has lead to the development of climate change experiments that alter the driving factors along a gradient, as implemented for instance in the Spruce and Peatland Responses Under Climatic and Environmental Change (SPRUCE, Krassovski et al., 2015) experiment. In SPRUCE spruce and peatland response to altered~~ ~~CO_2~~ ~~concentration and temperature gradient are studied (no change, +4, +8, +12, and +16 degrees Fahrenheit). Implementing climate change gradients substantially reduces the number of replica per treatment and can be analyzed by a regression approach instead of an ANOVA-type of approach (Kreyling et al., 2018).~~

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; Ka However, the resulting multivariate combination of climate variables may be physically unrealistic and may miss key aspects related to natural climate variability and ~~driven by land-atmosphere feedbacks. For instance~~the co-variance of multiple variables, linked to each other by synoptic conditions. This is particularly important for representing compound events, where a the combination of non extreme drivers can lead to extreme impacts (Rineau et al., 2019; Zscheischler and Seneviratne, 2017; Zscheischler et a For example, droughts and heatwaves often co-occur (Zscheischler and Seneviratne, 2017) and, soil moisture conditions and precipitation occurrence are linked (Guillod et al., 2015; Moon et al., 2019). Incorporating the ~~covariability~~ co-variability of key climate drivers is also important for the studied responses. For instance, heatwaves characterized by similar extreme air temperatures 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 varies a lot (De Boeck et al., 2016).

~~By focusing primarily on changes in mean climate conditions, projected change in climate variability is not taken into account in~~ Until recently, it was not possible to simulate realistic future climates in ecosystem climate change experiments ~~(Thompson et al., 2013). Changes in variability are important drivers of changes in the frequency, intensity and duration of extremes, which in turn are important drivers of ecosystem responses such as changes in community dynamics (Gutschick and BassiriRad, 2~~ ~~To capture the full range of changing climatic conditions, a holistic representation of the overall climate is necessary. This will be especially interesting also in gradient experiments covering a range of global warming levels, allowing for detection of non-linearities, thresholds and possible tipping points, as described in the novel approach by Rineau et al. (2019).~~

~~Climate change experiments require both extensive monitoring of the ecosystem processes at various spatio-temporal scales and~~ (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, 1996; Griffin et al., 1996; Stewart et al., 2013; Clobert et al., 2018)~~. They also allow to test the difference in response to an individual driver (e.g. one climate variable) and to simultaneous changes in multiple drivers, reflecting real-world conditions. Therefore, these types of infrastructures are very useful to perform climate change experiments, as they allow the control of a variety of climate variables with high accuracy~~ , especially in combination with an observation station in the field providing real time observations of most of those parameters (Rineau et al., 2019). This ~~implies that the experiments are driven by climate forcing that represents~~ 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 ~~realistic,~~ 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 (Rineau et al., 2019; Kreyling et al., 2018).

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 temperatures, droughts and intense precipitation

(Sillmann et al., 2013; Orlowsky and Seneviratne, 2013; Greve et al., 2018; Rajczak and Schär, 2017). 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 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 temperature forcing for experiments in which they use daily temperature output from a GCM (MIROC) and a stochastic weather generator to generate hourly weather. They validated their method against statistical characteristics of 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 a 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 experimental design for studying climate change impacts on terrestrial ecosystems~~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 ~~and naturally incorporates~~ while naturally incorporating extreme events under present and future climate conditions. Furthermore, the method can be ~~used in~~ 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 Temperature (GMT) gradient anomaly.

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

In our methodology, variability and co-variance between variables is 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 Global Mean Temperature (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. ~~With this approach, both the climatology and variability corresponding to these warming levels are represented. To preserve variability and co-variance between variables, we select 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 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 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, 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 worst-case scenario following an unabated greenhouse gas emissions pathway (Riahi et al., 2011). The experiments are running for 5 years. We choose time windows corresponding to the experimental period and centred around the year in which the climatological GMT anomaly (averaged with a 30-year period) crossed the pre-defined 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.

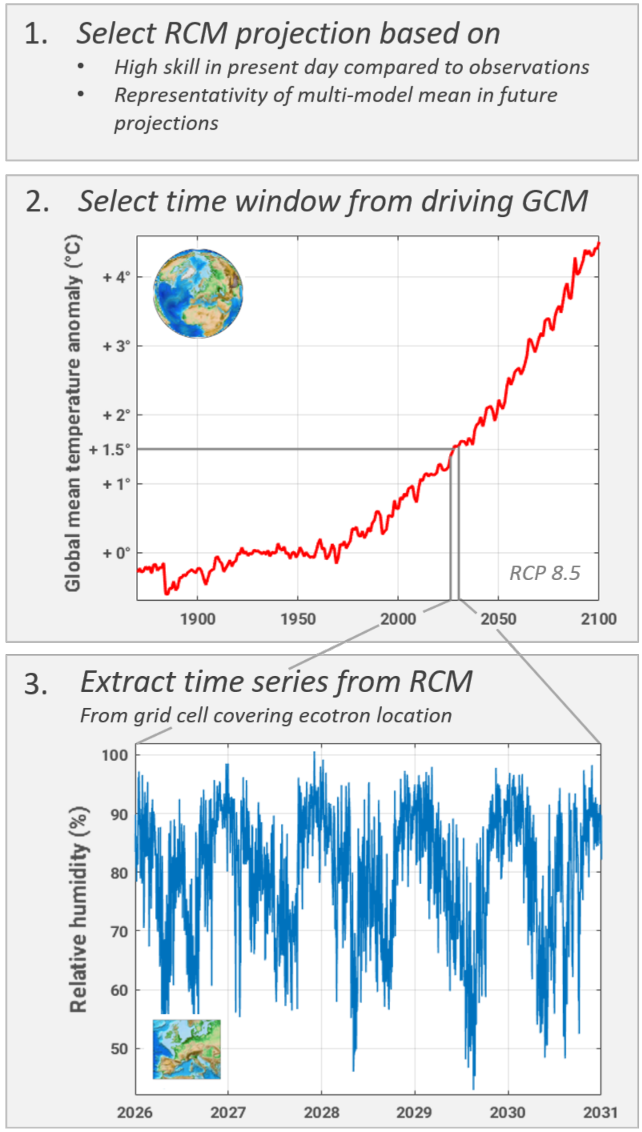

Figure 1. Methodology for generating climate forcing along the GMT anomaly gradient.

3 Data and methods

3.1 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 on-line 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 an 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 years 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), UVa and UVb. Here, wind and precipitation are measured and generated, and CO_2 , N_2, CH_4, 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, temperature, relative humidity, and CO_2 concentration are controlled (Rineau et al., 2019). 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 minutes (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, CH4 or N2O 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 forest or acid grassland and what the consequences are for ecosystem services (Rineau et al., 2019). The experiment will run uninterrupted for a period of at least five 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; Rineau et al., 2019). In the remaining six units, atmospheric conditions along the GMT anomaly gradient will be simulated as described in section 2. ~~Likewise,~~ The 3-hourly RCM output is linearly interpolated to a 30-minute 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 temperatures in the chamber (despite being thermically 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° (historical), +1° (present day), +1.5° (Paris Agreement), +2°C, +3°C and +4°C warmer world. The climatology of the unit forced by +1° can thereby be directly compared to the unit driven by the ICOS station and thus representing the present-day observed con-

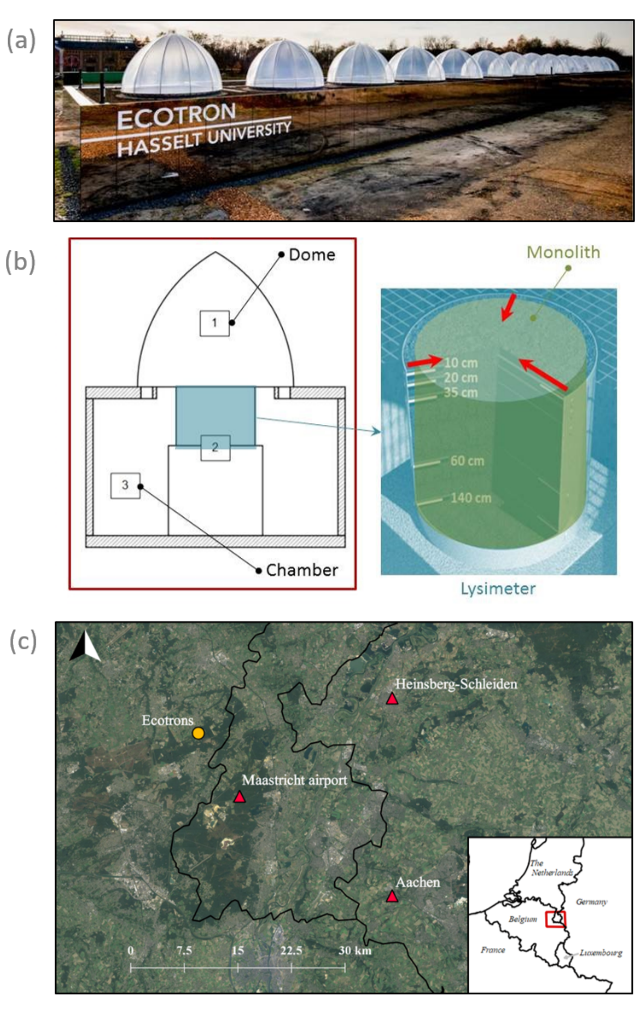

Figure 2. The UHasselt Ecotron experiment (a; picture: Liesbeth Driessen), scheme of a unit with the three compartments and the lysimeter compartment in detail (b), and overview map with location of the infrastructure and reference weather observation stations (~~b~~c).

ditions. 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).

3.2 Meteorological data

3.2.1 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 (Kotlarski et al., 2014; Jacob et al., 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 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 temperature and surface 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 are 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 analyse these variables for the respective simulations.

Once the EURO-CORDEX ensemble member is selected, the relevant variables (precipitation, mean temperature, surface 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 three-hourly values (except for surface up-welling latent heat flux and sensible heat flux) are then linearly interpolated to 30 minute 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 pressure is calculated from the mean sea level pressure using the altitude of the ecotron facility (43 m a.s.l.) and assuming hydrostatic equilibrium. The concentrations of the controllable greenhouse gases (CO_2, CH_4 and N_2O) are determined based on the annual values calculated by van Vuuren et al. (2011) according to RCP8.5. These correspond to the prescribed concentrations of the RCM simulations.

3.2.2 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 (11km), Aachen (37km) and Heinsberg-Schleiden (29 km; Fig. 2b). These weather stations provide daily observations from the end of the 19th century (Maastricht Airport and Aachen) or mid 20th century (Heinsberg-Schleiden) until the present-day, thereby covering both the EURO-CORDEX GCM and reanalysis downscaling periods. All stations record temperature [°C], precipitation [mm day^{-1}], relative humidity [%] and surface wind speed [m s^{-1}] 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 temperature, the curves overlay and for precipitation they are similar. Relative humidity 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 is 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).

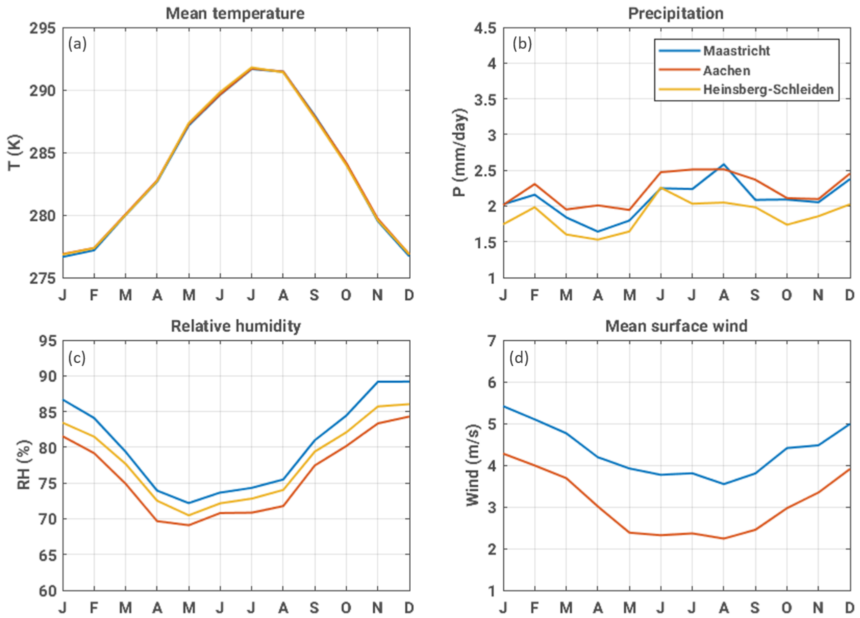

Figure 3. Seasonal cycles of observed mean temperature (a), precipitation (b), relative humidity (c) and mean surface wind (d) in the weather stations of Maastricht Airport, Aachen and Heinsberg-Schleiden (monthly averages based on daily data from 1963 to 2018). For Heinsberg-Schleiden no surface wind observations are available. The curves for temperature are overlaying.

3.3 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, ranging from 1-best performing model 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, can not 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 temperature and precipitation. To this extent, the average diurnal temperature range (DTR [K]; the difference between the daily maximum and minimum temperature) is calculated for the whole year, for the winter (December-January-February) and summer (June-July-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) and the number of frost days (days with a minimum temperature below 0°C) are calculated. Furthermore, the monthly maximum 1-day precipitation (Rx1day [mm day^{-1}]) 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]; 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 annual time scale and for the summer and winter seasonal averages, as correlations are expected to differ in sign and magnitude between the two seasons (e.g. negative correlation between 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) 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; Orlowsky and Seneviratne, 2013; Sillmann et al., 2013). The temperature indices are (i) ΔT [°C], the mean daily temperature change, (ii) ΔTXx [°C], the difference in annual maximum value of daily maximum temperature,(iii) ΔTNn [°C], the difference in annual minimum value of daily minimum temperature, (iv) Δ frost days, the difference in number of frost days (with a minimum temperature below 0°C), (v) Δ summer days, the difference in number of summer days (with the maximum temperature above 25°C), and finally (vi) ΔGSL [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 temperature higher than 5°C and the first span after July 1st of 6 days with a daily mean temperature lower than 5°C. The precipitation indices are (i) $\Delta PRCPTOT$ [mm], the difference in annual accumulated precipitation (as simulated over the five-year period), (ii) $\Delta Rx1day$ [mm] the difference in monthly maximum 1-day precipitation, (iii) $\Delta R10mm$ [days] the difference in number of days per year with more than 10 mm precipitation, (iv) ΔCDD [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)

ΔCWD [days] the maximum length of a wet spell (measured as the maximum number of consecutive days with more than 1 mm precipitation).

3.4 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 (Section 2, Fig. 1, table 2). Based on these time windows, we extract the three-hourly data for all necessary variables from the simulation for the 11 km by 11 km grid cell covering the location of the experiment.

4 Results

4.1 Identification of the best performing model simulation

4.1.1 First criterion: skill in present-day climate

Overall, model skill strongly varies across RCMs (Fig. 4). While the annual 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 temperature. CCLM4-8-17 is the best performing model based on the bias and total MAE metrics for temperature and precipitation, but is ranked in the mid range for the metrics related to the shape of its temperature distribution (PSS and percentile MAE). This can be attributed to an overestimation of the amplitude of the seasonal 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. 7a), 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

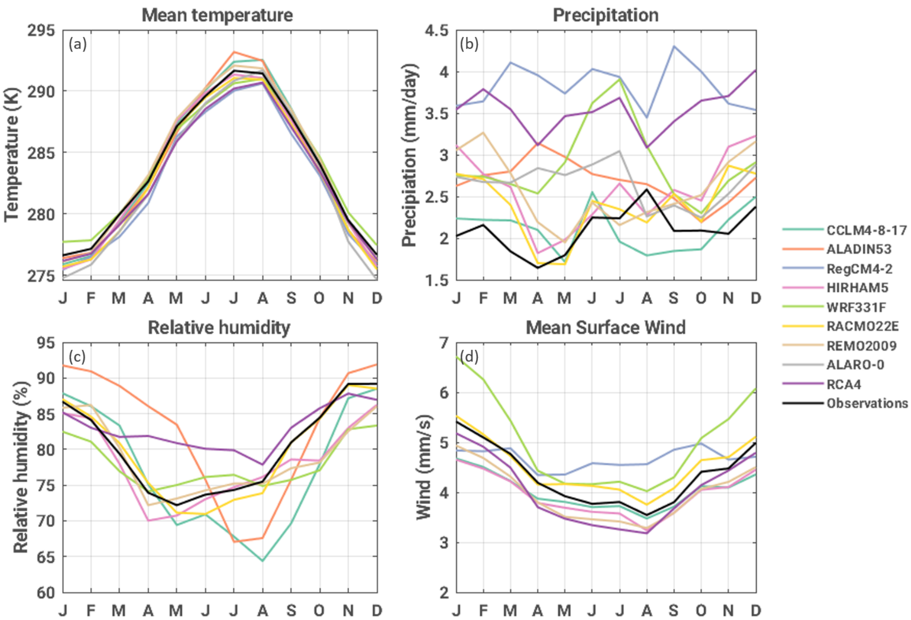

Figure 4. Seasonal cycle of the reanalysis downscalings for mean temperature (a), precipitation (b), relative humidity (c) and mean surface wind speed (d). (The RegCM4-2 and ALARO-0 simulations are not available for relative humidity and the ALADIN53 and ALARO-0 simulations are not available for surface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

skill for the annual, winter and summer diurnal temperature range. While RCA4 is highly ranked for 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 appendix Fig. A2). 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 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 appendix Fig. A1). The rankings show a mixed pattern for the different variables: there are no simulations which rank high for all considered variables (Fig. 6). For precipitation, the simulations with CCLM4-8-17, RACMO22E have better relative skill compared to the other simulations, 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 temperature climatology compared to the other simulations. For the climatological diagnostics, generally CCLM4-8-17 is scoring best for the

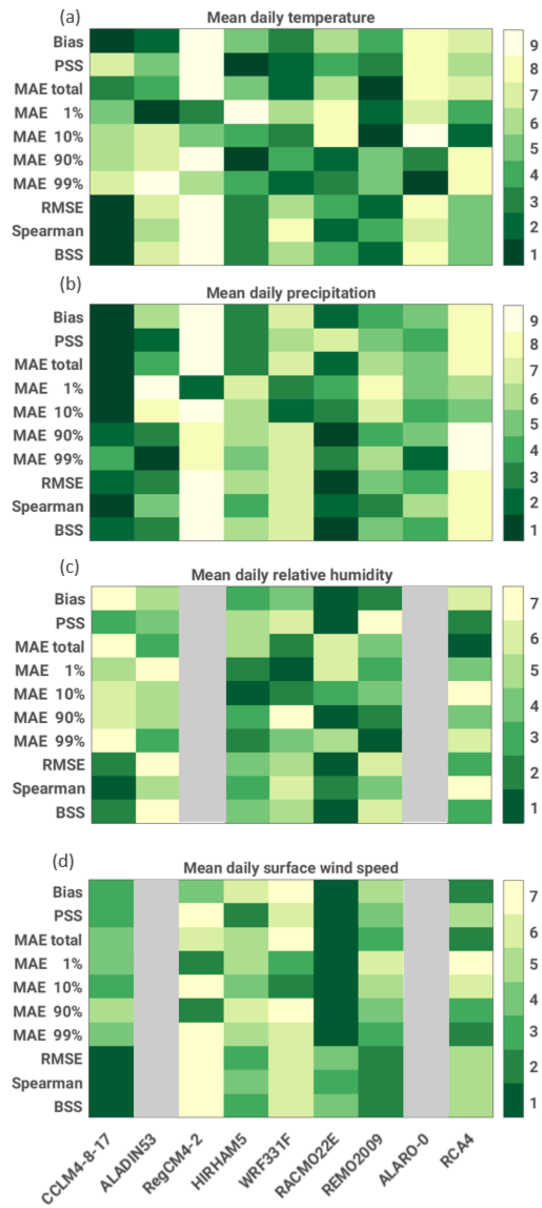

Figure 5. Ranking of the reanalysis downscalings based on performance on temperature (a), precipitation (b), relative humidity (c) and surface wind speed (d) compared to observations from Maastricht. The metrics shown are the Bias, Perkins Skill Score (PSS), Mean Absolute Error (MAE) for the entire time 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. Grey colors indicate that the variable is not available for the considered model.

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. 5,6 and 7). 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 temperature, relative humidity and surface wind speed. The precipitation biases for the considered simulations are +150 mm year^{-1} for CCLM4-8-17 driven by CNRM-CM5, +8 mm year^{-1} for CCLM4-8-17 driven by EC-EARTH, +24 mm year^{-1} for CCLM4-8-17 driven by MPI-ESM-LR, +323 mm year^{-1} for HIRHAM5 driven by EC-EARTH, 101 mm year-1 for HIRHAM5 driven by HadGEM2-ES and 35.51 mm year^{-1} for RACMO22E driven by HadGEM2-ES. Based on this, the CCLM4-8-17 EC-EARTH driven simulations 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.

4.1.2 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 temperature anomaly. For precipitation and relative humidity however, the CCLM4-8-17 EC-EARTH simulation show decreasing anomalies after 2050. 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 (Figs. A4,A5, A6, A7 and A8). For instance, the mean temperature anomaly of 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 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 temperature anomalies, up to the 90th percentile for 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.

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)	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

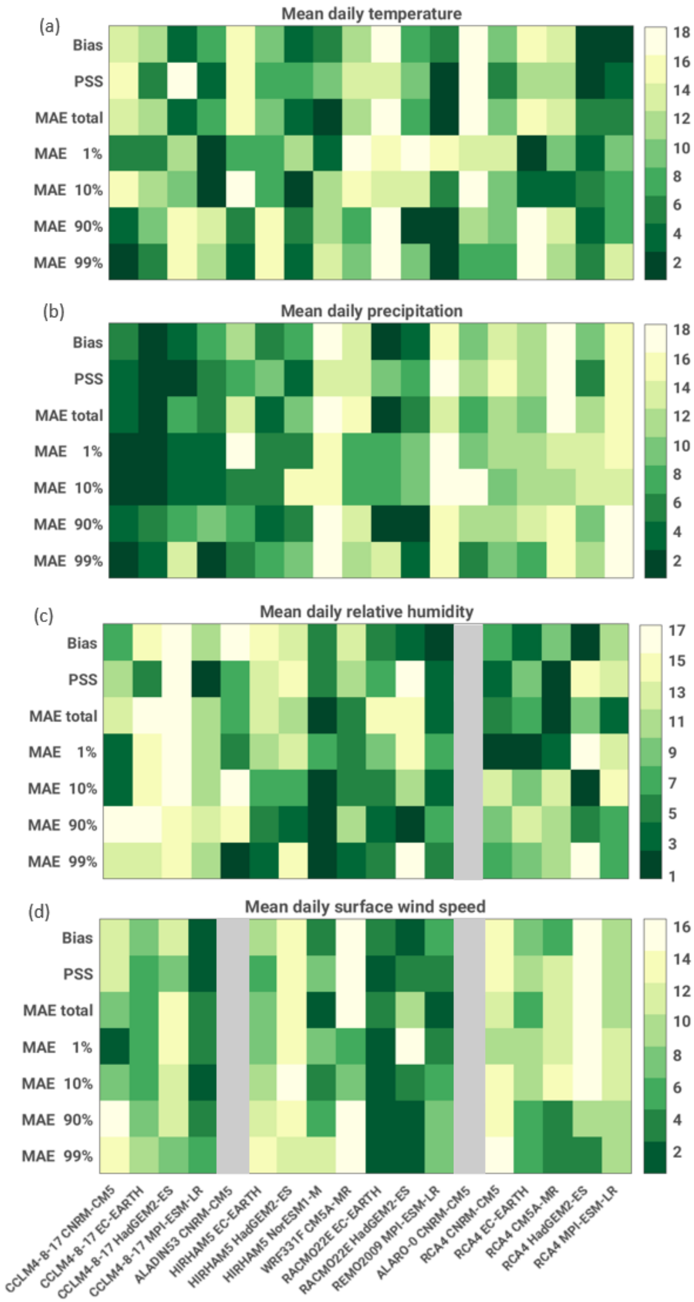

Figure 6. Ranking of the GCM downscalings based on performance on temperature (a), precipitation (b), relative humidity (c) and surface wind speed (d) compared to observations from Maastricht. The metrics showed are the bias, Perkins Skill Score (PSS), Mean Absolute Error (MAE) for the total and 1st, 10th, 90th and 99th percentile. Rankings are from 1-best to 16, 17 or 18-worst for surface wind speed, relative humidity, precipitation and temperature, respectively. Grey colors indicate that the variable is not available for the considered model.

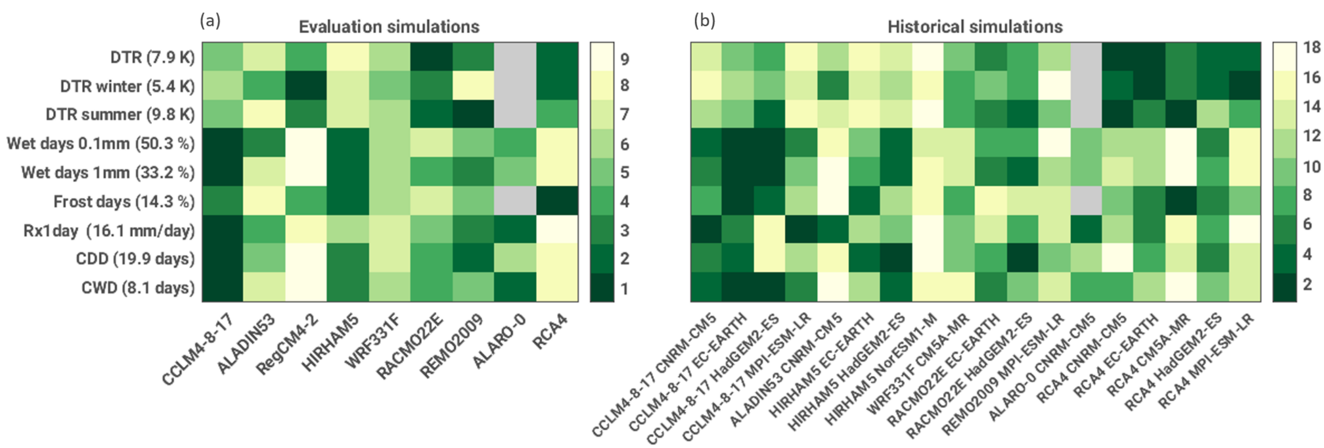

Figure 7. Ranking of the reanalysis (a) and GCM (b) downscalings for the historical period based on climatological diagnostics. Diurnal temperature range (DTR) in summer (July-August) and winter (December-February), number of wet days defined as days with precipitation > 0.1 mm and precipitation > 1 mm, number of frost days defined as days with mean temperature < 273.15 K, 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 downscalings, respectively.

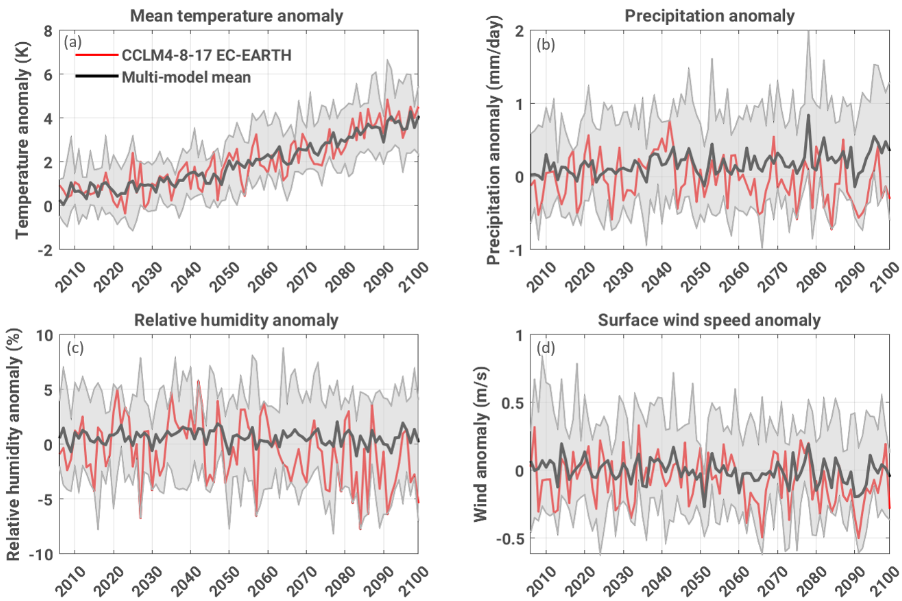

Figure 8. Anomalies for the CCLM4-8-17 EC-EARTH simulation following RCP 8.5 at the ecotron site for temperature (a), precipitation (b), relative humidity (c) and surface wind speed (d). The reference period is 1977 to 2006, the anomalies of the CLM4-8-17 EC-EARTH simulation are calculated compared to its own values in the reference period. In gray the envelope of all EURO-CORDEX RCP8.5 simulations is showed.

4.2 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. The overall trend of the local 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 temperature anomalies which are both increasing (Fig. 9). The mean daily temperature is increasing at a similar rate compared to GMT anomaly, and minimum and maximum temperature show a larger increase (table 2). None of the 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 a 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 favourable 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^{-1} +1.92 mm day^{-1} and +2.34 mm day^{-1}, respectively). These +2°C and +3°C units also knows 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.

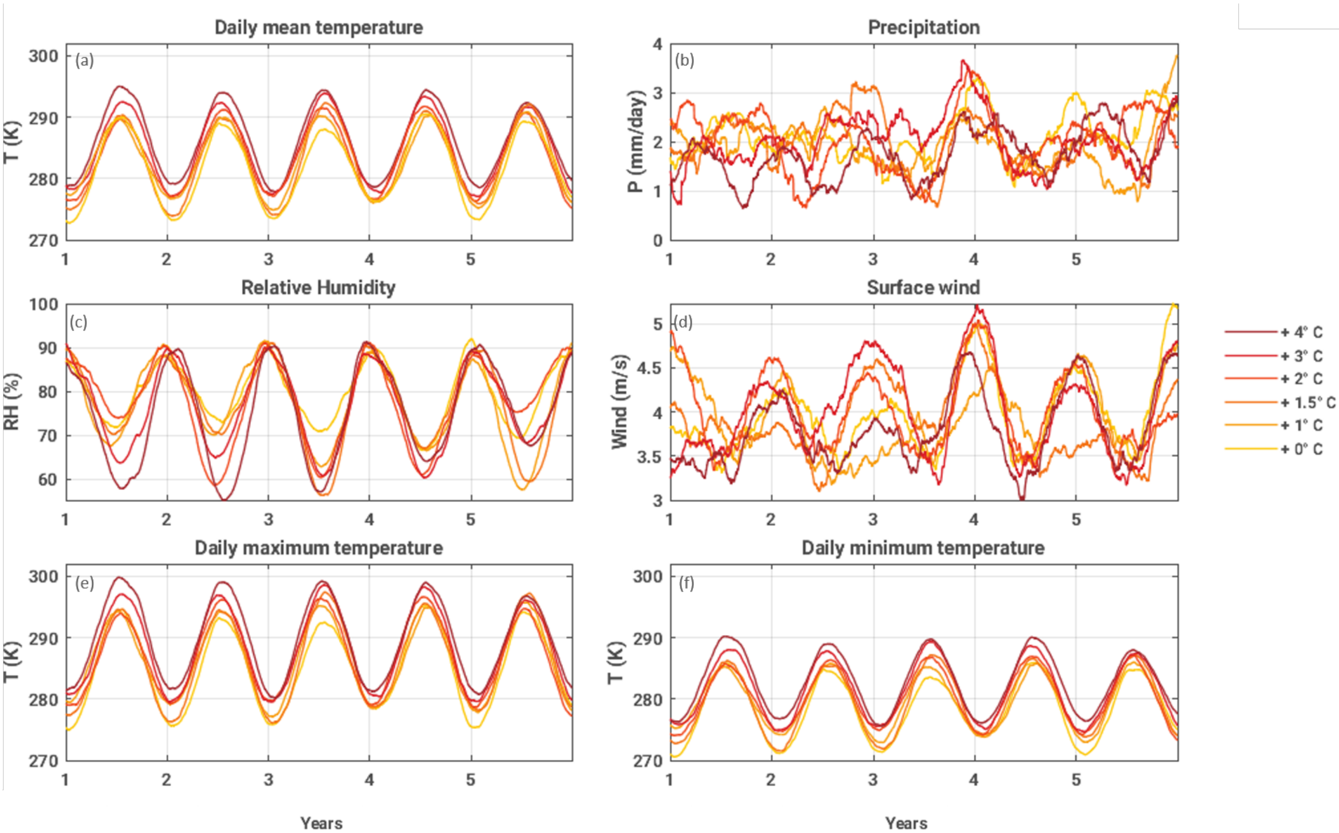

Figure 9. Annual cycles of the CCLM4-8-17 EC-EARTH ecotron unit forcing for the +1°C, +1.5°C, +2°C, +3°C and +4°C units compared to the 0°C reference period. Curves were smoothed using Savitzky-Golay filtering (order = 2 frame = 301; Savitzky and Golay (1964)

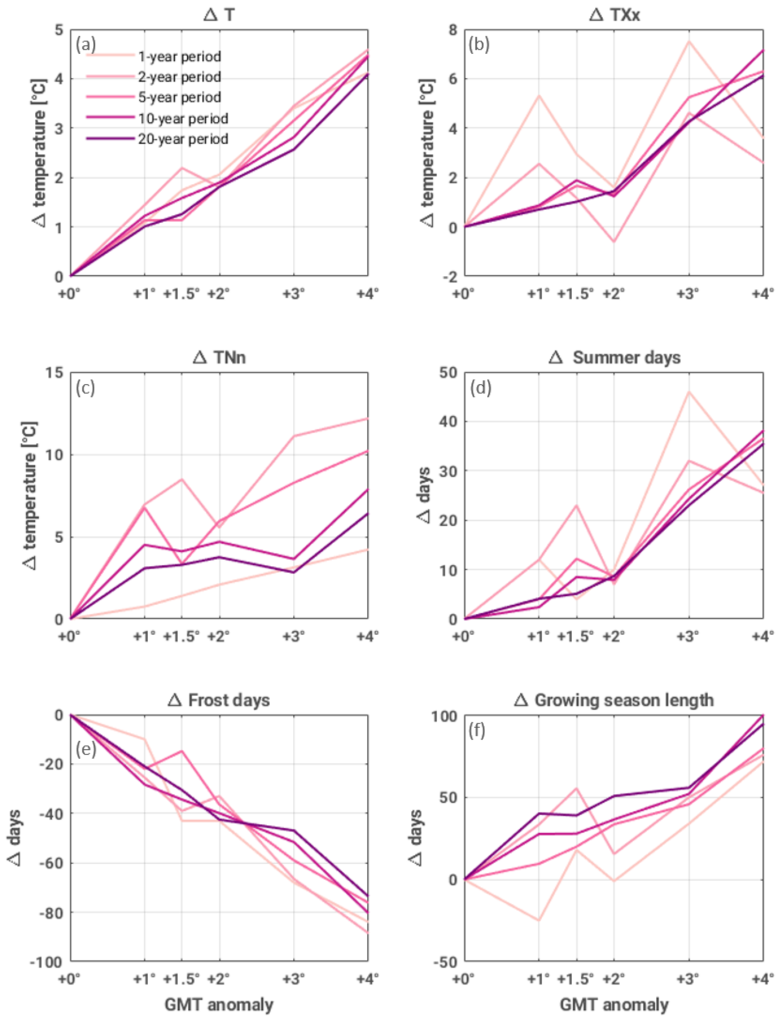

Figure 10. Annual anomalies per GMT anomaly for increasing time window lengths (ranging from a 1-year period to a 20-year period) of the CCLM4-8-17 EC-EARTH simulation following RCP 8.5 for temperature indices: mean temperature anomaly (ΔT; a), annual maximum temperature (ΔTXx; b), annual minimum temperature (ΔTNn; 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.

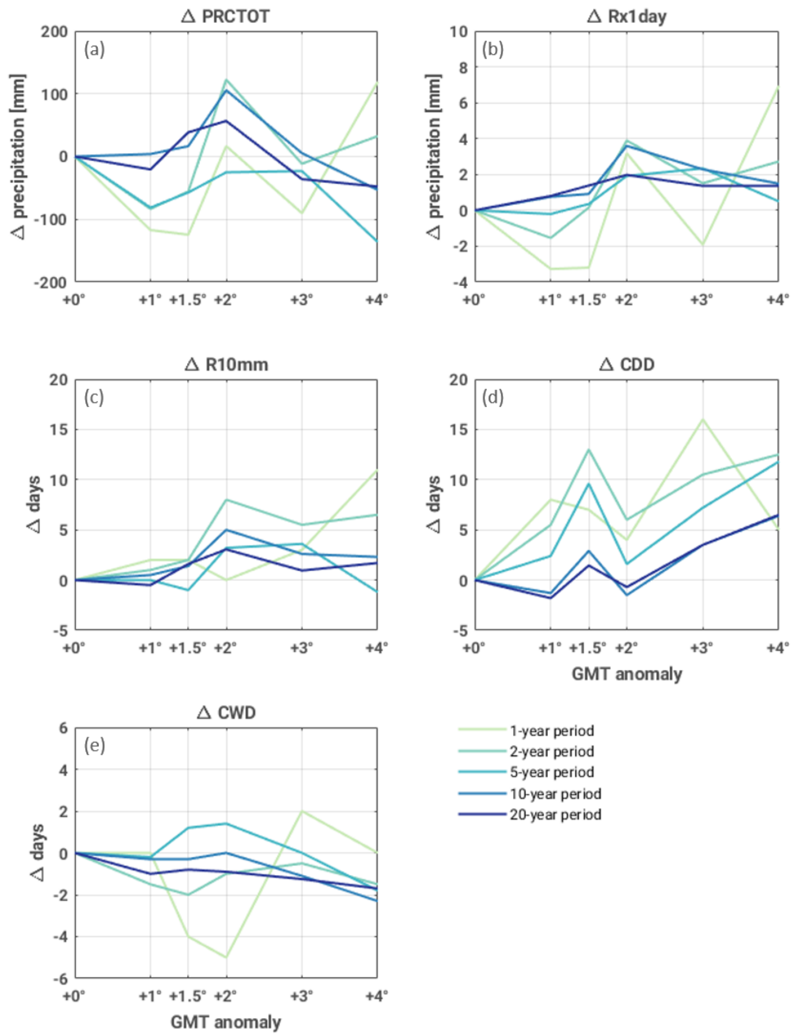

Figure 11. Same as Fig. 10, but now for precipitation indices: the annual accumulated precipitation anomaly ($\Delta PRCPTOT$; a), anomaly of monthly maximum 1-day precipitation ($\Delta Rx1day$; b), anomaly of annual number of days with more than 10 mm precipitation ($\Delta R10mm$; c), anomaly of annual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

Table 2. Extracted 5-year periods and temperature and precipitation indices based on ETCCDI for the CCLM4-8-17 EC-EARTH simulation at the ecotron location. The 0°C column gives the absolute reference values. The periods are calculated based on the 30-year averaged global mean temperature (GMT) anomaly calculated from EC-EARTH.

	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
ΔT [°C]	*8.17*	+1.13	+1.14	+1.81	+3.15	+4.49
ΔTXx [°C]	*30.98*	+0.82	+1.66	+1.34	+5.24	+6.30
ΔTNn [°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
ΔGSL [days]	*225.6*	+9.6	+20	+33.6	+45.8	+80
$\Delta PRCPTOT$ [mm]	*771.09*	-81.32	-57.2	+25.12	-23.14	-136.05
$\Delta Rx1day$ [mm]	*14.38*	-0.2	+0.35	+1.92	+2.34	+0.5
$\Delta R10mm$ [days]	*14.6*	0	-1	+3.2	+3.6	-1.2
ΔCDD [days]	*17.2*	+2.4	+9.6	+1.6	+7.2	+11.8
ΔCWD [days]	*9.6*	-0.2	+1.2	+1.4	0	-1.8

5 Discussion

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

We extract all climate variables from one grid cell of the RCM simulation to conserve a realistic, non smoothed signal. 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 neighbouring grid cells. The grid-cell values also reflect strong interannual 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., 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 multi-variate drivers as compound controls. For instance, it will provide a unique opportunity to study the impact from 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 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 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 set-up.

In the different ecotron units, we assume that the controlled variables (CO_2 and CH_4 concentration, 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 et al. (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, 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).

~~Finally, the~~ The set-up 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 temperature and moisture is 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 (**?**). Although climate variables are prescribed, ecosystem-climate feedbacks originating from interactions between the biosphere and atmosphere can by 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 (Guillod et al., 2015) may be assessed by calculating imbalances in the energy budget.

6 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 ~~natural variability and~~ co-variance between climate drivers, natural variability, changes in extreme events. To this extent, we presented a new ~~methodology for generating climate forcing~~ method for creating realistic climate forcing for manipulation experiments using a single Regional Climate Model (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 (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 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 ~~ecotron units~~ units of the UHasselt Ecotron Experiment are forced with climate conditions along a Global Mean 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 temperature thresholds. Subsequently, the ecotron forcing is extracted from the 3-hourly RCM simulation according to the time windows.

~~The UHasselt Ecotron experiment allows to quantify and assess the ecosystem responses on changing climatic conditions, thereby accounting for the~~ Our new methodology provides realistic climate forcing, accounting for co-variances between climatic variables and their change in variability, well representing possible compound events. ~~By applying a gradient approach, thresholds and possible tipping points can be identified~~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 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.

Code and data availability. 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/. The scripts used in the analysis are available on github: https://github.com/VUB-HYDR/2020_Vanderkelen_etal_BG.

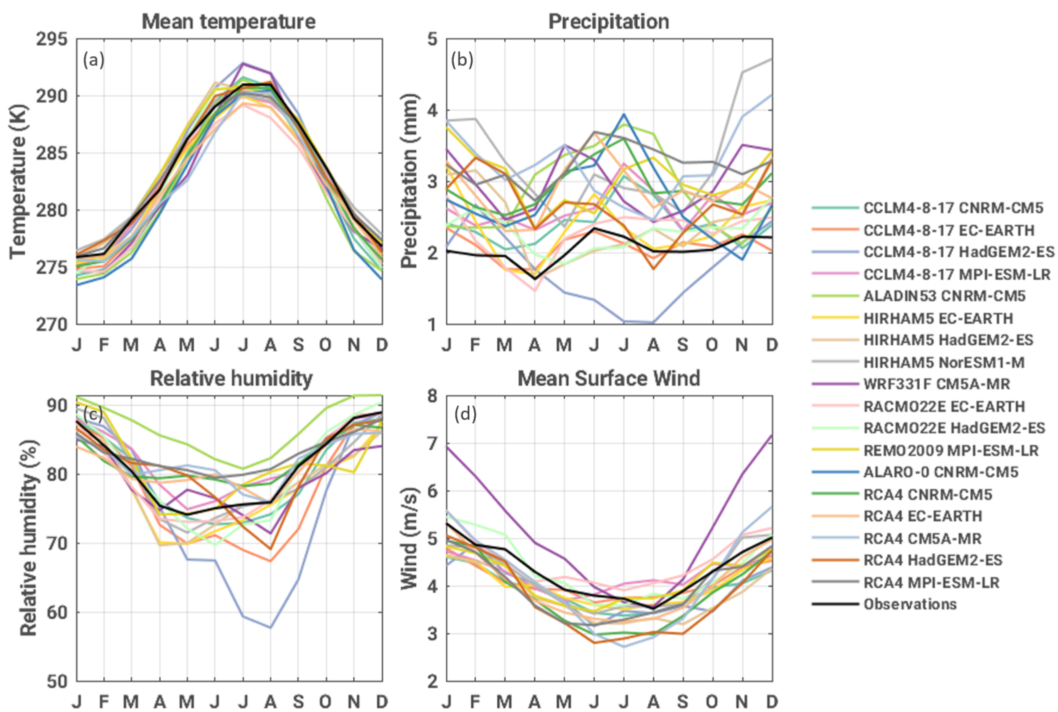

Figure A1. Seasonal cycle of the GCM downscalings for mean temperature (a), precipitation (b), relative humidity (c) and mean surface wind speed (d).

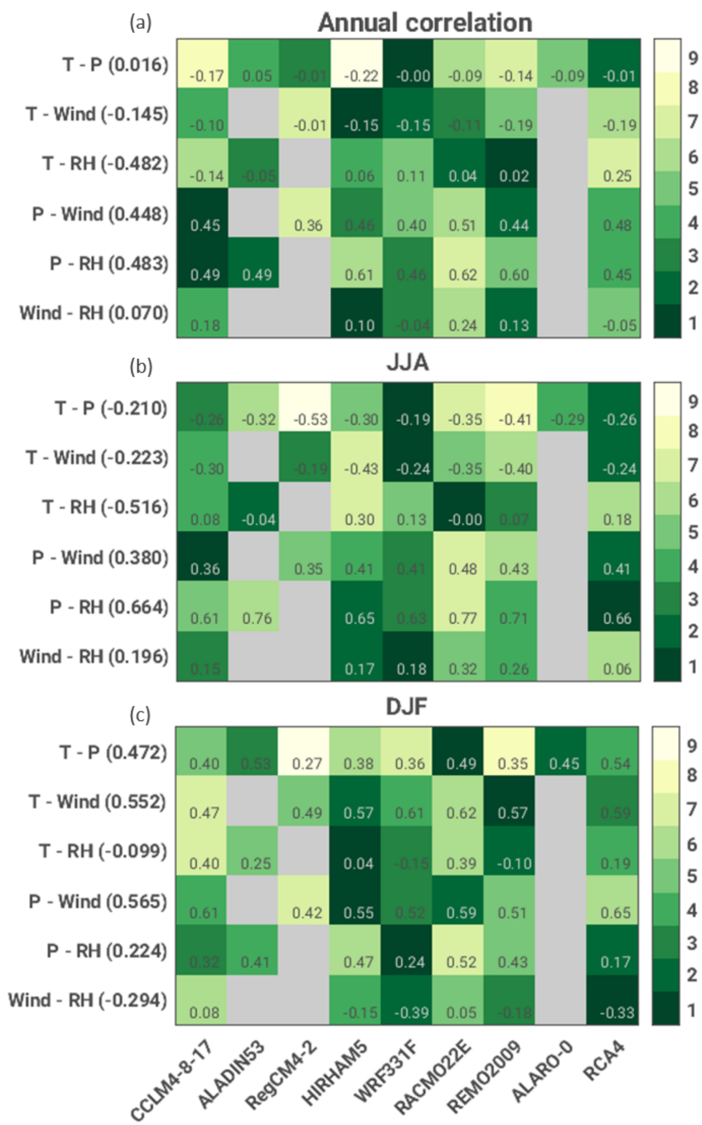

Figure A2. Correlations for the reanalysis downscalings (1990-2008): Annual correlations (a), correlations in June, July and August (JJA; b) and correlations in December, January and February (DJF; c). T stands for temperature, P for precipitation and RH for relative humidity. The values in the y-axis labels are the observed correlations, and the other values correlations between the simulated variables. Rankings are from 1-best to 9-worst.

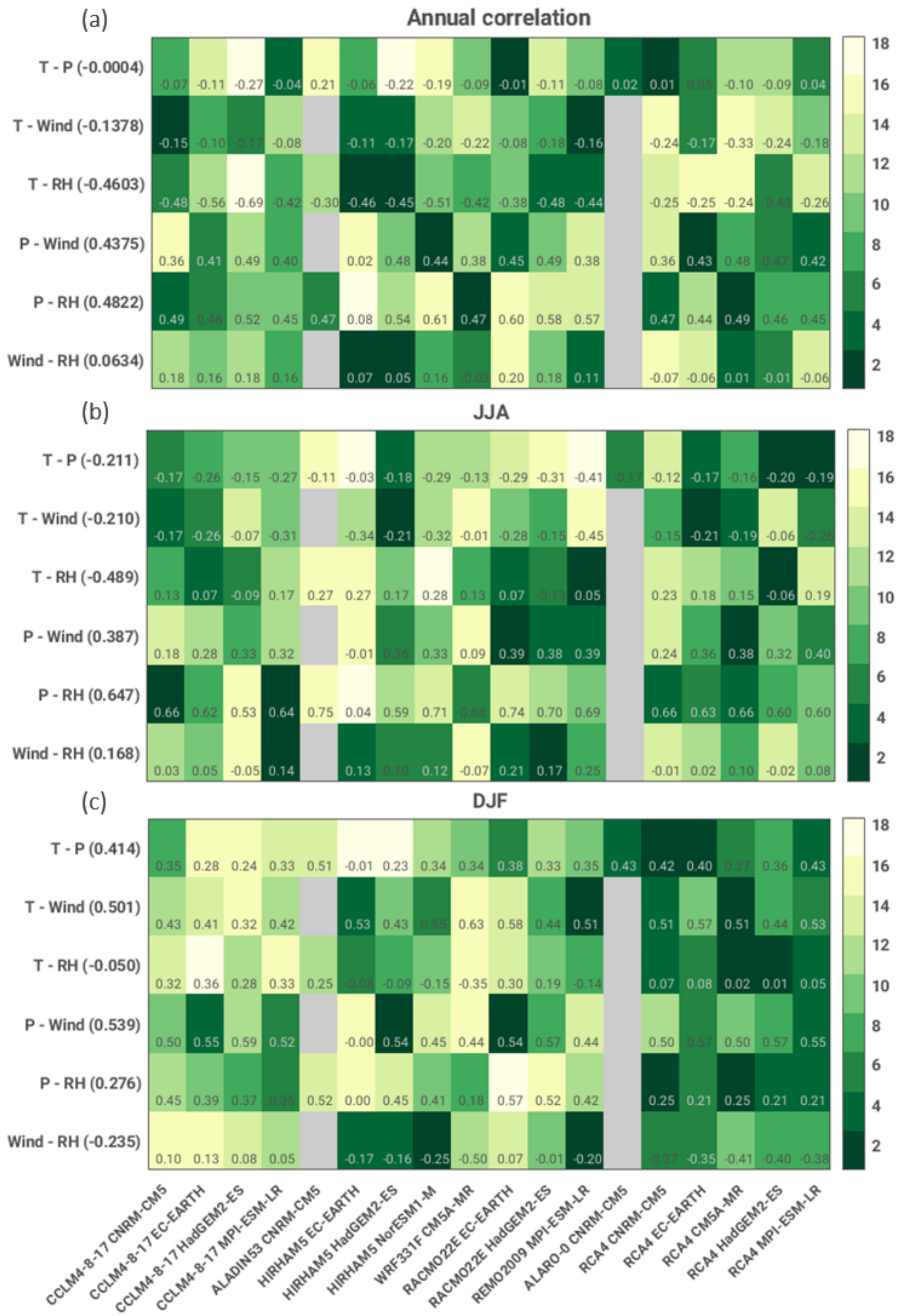

Figure A3. Correlations for the GCM downscalings (1951-2005): Annual correlations (a), correlations in June, July and August (JJA; b) and correlations in December, January and February (DJF; c). T stands for temperature, P for precipitation and RH for relative humidity. The values in the y-axis labels are the observed correlations, and the other values correlations between the simulated variables. Rankings are from 1-best to 9-worst.

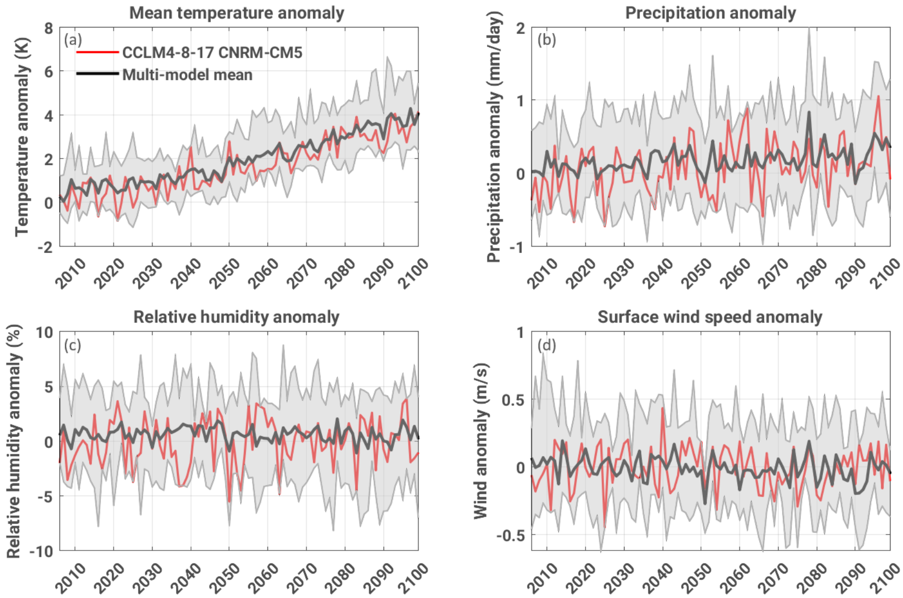

Figure A4. Same as Fig. 8, but now for CCLM4-8-17 CNRM-CM5.

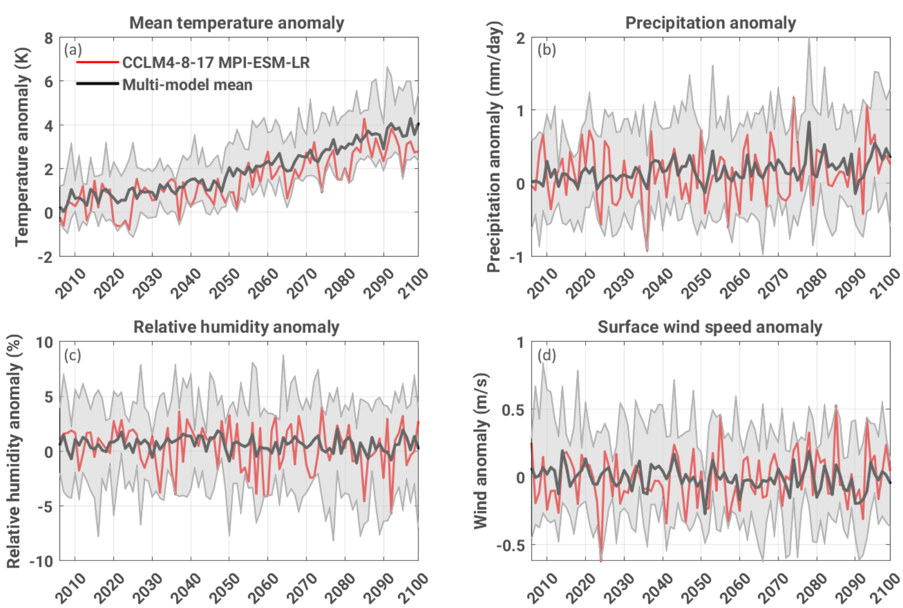

Figure A5. Same as Fig. 8, but now for CCLM4-8-17 MPI-ESM-LR.

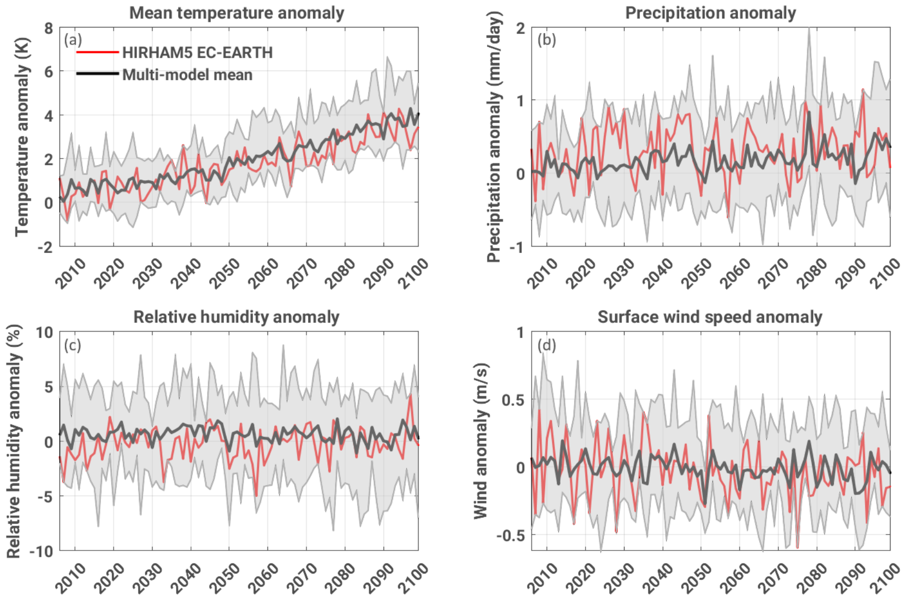

Figure A6. Same as Fig. 8, but now for HIRHAM5 EC-EARTH.

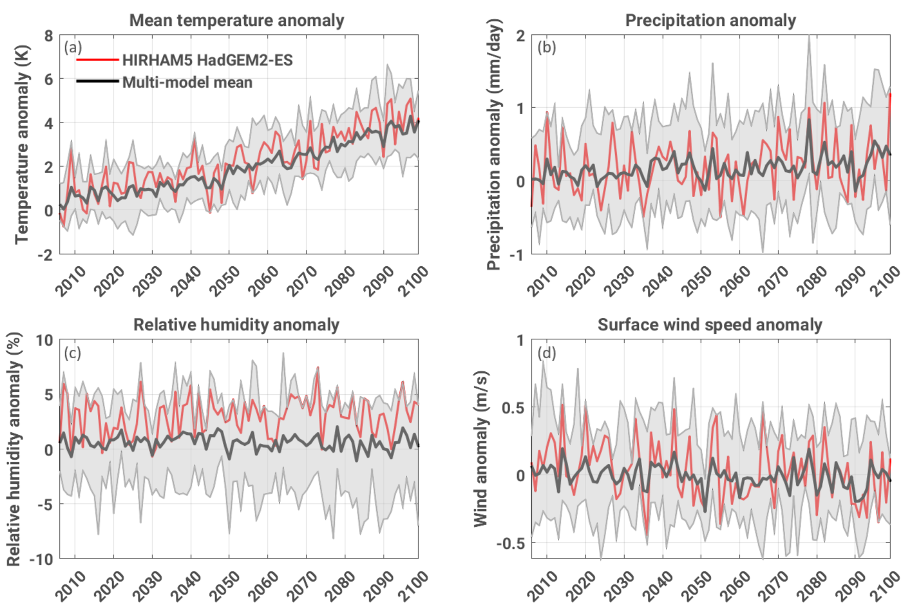

Figure A7. Same as Fig. 8, but now for HIRHAM5 HadGEM2-ES.

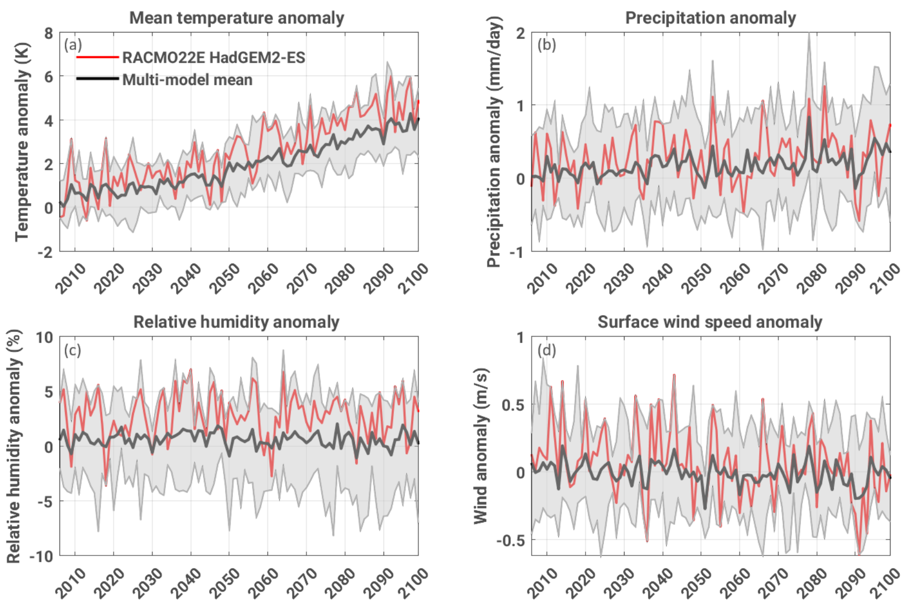

Figure A8. Same as Fig. 8, but now for RACMO22E HadGEM2-ES.

Author contributions. JZ, LG, FR, WT, NB and JV conceived the ideas and designed the methodology, KK provided the 3-hourly simulation data. IV and WT led the writing of the manuscript, with major contributions from SV, JK, FR and input from all other authors. All authors critically revised the draft and gave final approval for publication.

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

Acknowledgements. 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. We are grateful to the World Climate Research Programme (WRCP) for initiating and coordinating the EURO-CORDEX initiative, to the modelling centres 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. 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. We also thank RLKM (Regional Landscape Kempen and Maasland) for its collaboration and support."

References

Berendse, F., Schmitz, M., and de Visser, W.: Experimental manipulation of succession in heathland ecosystems, Oecologia, 100, 38–44, https://doi.org/10.1007/BF00317128, 1994.

Brier, G. W.: Verification of forecasts in terms of probability, Monthly Weather Review, 78, 1–3, 1950.

Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G., and Saba, V.: Observed fingerprint of a weakening Atlantic Ocean overturning circulation, Nature, 556, 191–196, https://doi.org/10.1038/s41586-018-0006-5, 2018.

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., and Unnikrishnan, A. S.: Sea Level Change, Cambridge University Press, Cambridge, 2013.

Clobert, J., Chanzy, A., Le Galliard, J.-F., Chabbi, A., Greiveldinger, L., Caquet, T., Loreau, M., Mougin, C., Pichot, C., Roy, J., and Saint-André, L.: How to Integrate Experimental Research Approaches in Ecological and Environmental Studies: AnaEE France as an Example, Frontiers in Ecology and Evolution, 6, 43, https://doi.org/10.3389/fevo.2018.00043, 2018.

Coumou, D., Di Capua, G., Vavrus, S., Wang, L., and Wang, S.: The influence of Arctic amplification on mid-latitude summer circulation, https://doi.org/10.1038/s41467-018-05256-8, 2018.

Curtis, P. S. and Wang, X.: A meta-analysis of elevated CO 2 effects on woody plant mass, form, and physiology, Oecologia, 113, 299–313, https://link.springer.com/content/pdf/10.1007{%}2Fs004420050381.pdf, 1998.

De Boeck, H. J., Van De Velde, H., De Groote, T., and Nijs, I.: Ideas and perspectives: Heat stress: More than hot air, Biogeosciences, 13, 5821–5825, https://doi.org/10.5194/bg-13-5821-2016, 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., and Janssens, I. A.: Few multiyear precipitation-reduction experiments find a shift in the productivity-precipitation relationship, Global change biology, 22, 2570–2581, https://doi.org/10.1111/gcb.13269, 2016.

Fischer, E. M. and Knutti, R.: Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes, Nature Climate Change, 5, 560–564, https://doi.org/10.1038/nclimate2617, 2015.

Fischer, E. M., Sedláček, J., Hawkins, E., and Knutti, R.: Models agree on forced response pattern of precipitation and temperature extremes, Geophysical Research Letters, 41, 8554–8562, https://doi.org/10.1002/2014GL062018, 2014.

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., and Rummukainen, M.: Evaluation of Climate Models, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., pp. 741–866, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

Greve, P., Gudmundsson, L., and Seneviratne, S. I.: Regional scaling of annual mean precipitation and water availability with global temperature change, Earth System Dynamics, 9, 227–240, https://doi.org/10.5194/esd-9-227-2018, 2018.

Griffin, K. L., Ross, P. D., Sims, D. A., Luo, Y., Seemann, J. R., Fox, C. A., and Ball, J. T.: EcoCELLs: Tools for mesocosm scale measurements of gas exchange, Plant, Cell and Environment, 19, 1210–1221, https://doi.org/10.1111/j.1365-3040.1996.tb00437.x, 1996.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations - A comparison of methods, Hydrology and Earth System Sciences, 16, 3383–3390, https://doi.org/10.5194/hess-16-3383-2012, 2012.

Guillod, B. P., Orlowsky, B., Miralles, D. G., Teuling, A. J., and Seneviratne, S. I.: Reconciling spatial and temporal soil moisture effects on afternoon rainfall, Nature Communications, 6, 6443, https://doi.org/10.1038/ncomms7443, 2015.

Gutschick, V. P. and BassiriRad, H.: Extreme events as shaping physiology, ecology, and evolution of plants: Toward a unified definition and evaluation of their consequences, https://doi.org/10.1046/j.1469-8137.2003.00866.x, 2003.

Hovenden, M. J., Newton, P. C., and Wills, K. E.: Seasonal not annual rainfall determines grassland biomass response to carbon dioxide, Nature, 511, 583–586, https://doi.org/10.1038/nature13281, 2014.

IPCC 2013: Summary for Policymakers, in: Climate Change 2013 - The Physical Science Basis, vol. 1542, pp. 1–30, Cambridge University Press, Cambridge, https://doi.org/10.1017/CBO9781107415324.004, https://www.cambridge.org/core/product/identifier/CBO9781107415324A009/type/book{_}part, 2013.

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., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J. F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: EURO-CORDEX: New high-resolution climate change projections for European impact research, Regional Environmental Change, 14, 563–578, https://doi.org/10.1007/s10113-013-0499-2, 2014.

Kardol, P., De Long, J., and Sundgvist, M.: Crossing the threshold: the power of multi-level experiments in identifying global change responses, New Phytologist, 196, 323–326, 2012.

Karlowsky, S., Augusti, A., Ingrisch, J., Hasibeder, R., Lange, M., Lavorel, S., Bahn, M., and Gleixner, G.: Land use in mountain grasslands alters drought response and recovery of carbon allocation and plant-microbial interactions, Journal of Ecology, 106, 1230–1243, https://doi.org/10.1111/1365-2745.12910, http://doi.wiley.com/10.1111/1365-2745.12910, 2018.

Keuler, K., Radtke, K., Kotlarski, S., and Lüthi, D.: Regional climate change over Europe in COSMO-CLM: Influence of emission scenario and driving global model, Meteorologische Zeitschrift, 25, 121–136, https://doi.org/10.1127/metz/2016/0662, 2016.

Klein Tank, A. M., Wijngaard, J. B., Können, G. P., Böhm, R., Demarée, G., Gocheva, A., Mileta, 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., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L. V., and Petrovic, P.: Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment, International Journal of Climatology, 22, 1441–1453, https://doi.org/10.1002/joc.773, 2002.

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., and Smith, M. D.: A reality check for climate change experiments: Do they reflect the real world?, Ecology, 99, 2145–2151, https://doi.org/10.1002/ecy.2474, 2018.

Knutti, R. and Rogelj, J.: The legacy of our CO2 emissions: a clash of scientific facts, politics and ethics, Climatic Change, 133, 361–373, https://doi.org/10.1007/s10584-015-1340-3, 2015.

Korell, L., Auge, H., Chase, J. M., Harpole, S., and Knight, T. M.: We need more realistic climate change experiments for understanding ecosystems of the future, Global Change Biology, 00, 1–3, https://doi.org/10.1111/gcb.14797, 2019.

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., and Wulfmeyer, V.: Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble, Geoscientific Model Development, 7, 1297–1333, https://doi.org/10.5194/gmd-7-1297-2014, 2014.

Krassovski, M. B., Riggs, J. S., Hook, L. A., Nettles, W. R., Hanson, P. J., and Boden, T. A.: A comprehensive data acquisition and management system for an ecosystem-scale peatland warming and elevated CO 2 experiment, Geosci. Instrum. Method. Data Syst, 4, 203–213, https://doi.org/10.5194/gi-4-203-2015, www.geosci-instrum-method-data-syst.net/4/203/2015/, 2015.

Kreyling, J., Schweiger, A. H., Bahn, M., Ineson, P., Migliavacca, M., Morel-Journel, T., Christiansen, J. R., Schtickzelle, N., and Larsen, K. S.: To replicate, or not to replicate – that is the question: how to tackle nonlinear responses in ecological experiments, Ecology Letters, 21, 1629–1638, https://doi.org/10.1111/ele.13134, 2018.

Lawton, J. H.: The Ecotron: a controlled environmental facility for the investigation of population and ecosystem processes, Philosophical Transactions - Royal Society of London, B, 341, 181–194, https://doi.org/10.1098/rstb.1993.0102, 1993.

Lawton, J. H.: The ecotron facility at silwood park: The value of "Big Bottle" experiments, Ecology, 77, 665–669, https://doi.org/10.2307/2265488, http://doi.wiley.com/10.2307/2265488, 1996.

Lin, D., Xia, J., and Wan, S.: Climate warming and biomass accumulation of terrestrial plants: A meta-analysis, New Phytologist, 188, 187–198, https://doi.org/10.1111/j.1469-8137.2010.03347.x, 2010.

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., and Yeager, S.: Decadal climate prediction an update from the trenches, Bulletin of the American Meteorological Society, 95, 243–267, https://doi.org/10.1175/BAMS-D-12-00241.1, 2014.

Moon, H., Guillod, B. P., Gudmundsson, L., and Seneviratne, S. I.: Soil Moisture Effects on Afternoon Precipitation Occurrence in Current Climate Models, Geophysical Research Letters, 46, 1861–1869, https://doi.org/10.1029/2018GL080879, 2019.

Murphy, A. H.: A New Vector Partition of the Probability Score, 1973.

Orlowsky, B. and Seneviratne, S. I.: Elusive drought: Uncertainty in observed trends and short-and long-term CMIP5 projections, Hydrology and Earth System Sciences, 17, 1765–1781, https://doi.org/10.5194/hess-17-1765-2013, 2013.

Perkins, S. E., Pitman, A. J., Holbrook, N. J., and McAneney, J.: Evaluation of the AR4 climate models' simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions, Journal of Climate, 20, 4356–4376, https://doi.org/10.1175/JCLI4253.1, 2007.

Rajczak, J. and Schär, C.: Projections of Future Precipitation Extremes Over Europe: A Multimodel Assessment of Climate Simulations, Journal of Geophysical Research: Atmospheres, 122, 10,773–10,800, https://doi.org/10.1002/2017JD027176, https://doi.org/10.1002/, 2017.

Riahi, K., Rao, S., Krey, V., Cho, C., Chirkov, V., Fischer, G., Kindermann, G., Nakicenovic, N., and Rafaj, P.: RCP 8.5-A scenario of comparatively high greenhouse gas emissions, Climatic Change, 109, 33–57, https://doi.org/10.1007/s10584-011-0149-y, 2011.

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., and Vangronsveld, J.: Supplementary materials: Towards more predictive and interdisciplinary climate change ecosystem experiments, Nature Climate Change, 9, 809–816, https://doi.org/10.1038/s41558-019-0609-3, 2019.

Roy, J., Picon-Cochard, C., Augusti, A., Benot, M.-L., Thiery, L., Darsonville, O., Landais, D., Piel, C., Defossez, M., Devidal, S., Escape, C., Ravel, O., Fromin, N., Volaire, F., Milcu, A., Bahn, M., and Soussana, J.-F.: Elevated CO 2 maintains grassland net carbon uptake under a future heat and drought extreme, Proceedings of the National Academy of Sciences, 113, 6224–6229, https://doi.org/10.1073/pnas.1524527113, 2016.

Rustad, L., Campbell, J., Marion, G., Norby, R., Mitchell, M., Hartley, A., Cornelissen, J., and Gurevitch, J.: A meta-analysis of the response of soil respiration, net nitrogen mineralization, and aboveground plant growth to experimental ecosystem warning, Oecologia, 126, 543–562, https://doi.org/10.1007/s004420000544, https://link.springer.com/content/pdf/10.1007{%}2Fs004420000544.pdf, 2001.

Savitzky, A. and Golay, M. J.: Smoothing and Differentiation of Data by Simplified Least Squares Procedures, Analytical Chemistry, 36, 1627–1639, https://doi.org/10.1021/ac60214a047, 1964.

Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., and Teuling, A. J.: Investigating soil moisture-climate interactions in a changing climate: A review, https://doi.org/10.1016/j.earscirev.2010.02.004, 2010.

Seneviratne, S. I., Donat, M. G., Pitman, A. J., Knutti, R., and Wilby, R. L.: Allowable CO2 emissions based on regional and impact-related climate targets, https://doi.org/10.1038/nature16542, 2016.

Sillmann, J., Kharin, V. V., Zwiers, F. W., Zhang, X., and Bronaugh, D.: Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections, Journal of Geophysical Research: Atmospheres, 118, 2473–2493, https://doi.org/10.1002/jgrd.50188, 2013.

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., and Woodward, G.: Mesocosm Experiments as a Tool for Ecological Climate-Change Research, Advances in Ecological Research, 48, 71–181, https://doi.org/10.1016/B978-0-12-417199-2.00002-1, 2013.

Terrer, C., Vicca, S., Stocker, B. D., Hungate, B. A., Phillips, R. P., Reich, P. B., Finzi, A. C., and Prentice, I. C.: Ecosystem responses to elevated CO2 governed by plant–soil interactions and the cost of nitrogen acquisition, https://doi.org/10.1111/nph.14872, 2018.

Thompson, R. M., Beardall, J., Beringer, J., Grace, M., and Sardina, P.: Means and extremes: Building variability into community-level climate change experiments, Ecology Letters, 16, 799–806, https://doi.org/10.1111/ele.12095, 2013.

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., Doubková, M., Friedl, M. A., Jung, M., van den Hurk, B. J., de Jeu, R. A., Kruijt, B., Ohta, T., Rebel, K. T., Plummer, S., Seneviratne, S. I., Sitch, S., Teuling, A. J., van der Werf, G. R., and Wang, G.: Drought and ecosystem carbon cycling, https://doi.org/10.1016/j.agrformet.2011.01.018, 2011.

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., and Rose, S. K.: The representative concentration pathways: An overview, Climatic Change, 109, 5–31, https://doi.org/10.1007/s10584-011-0148-z, 2011.

Vanderkelen, I., van Lipzig, N. P. M., and Thiery, W.: Modelling the water balance of Lake Victoria (East Africa) - Part 2: Future projections, Hydrology and Earth System Sciences, 22, 5527–5549, https://doi.org/10.5194/hess-22-5527-2018, 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., Escolar, 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., and Janssens, I. A.: Can current moisture responses predict soil CO2 efflux under altered precipitation regimes? A synthesis of manipulation experiments, Biogeosciences, 11, 2991–3013, https://doi.org/10.5194/bg-11-2991-2014, 2014.

Wu, Z., Dijkstra, P., Koch, G. W., Peñuelas, J., and Hungate, B. A.: Responses of terrestrial ecosystems to temperature and precipitation change: A meta-analysis of experimental manipulation, Global Change Biology, 17, 927–942, https://doi.org/10.1111/j.1365-2486.2010.02302.x, 2011.

Yue, K., Fornara, D. A., Yang, W., Peng, Y., Peng, C., Liu, Z., and Wu, F.: Influence of multiple global change drivers on terrestrial carbon storage: additive effects are common, Ecology Letters, 20, 663–672, https://doi.org/10.1111/ele.12767, 2017.

Zekollari, H., Huss, M., and Farinotti, D.: Modelling the future evolution of glaciers in the European Alps under the EURO-CORDEX RCM ensemble, Cryosphere, 13, 1125–1146, https://doi.org/10.5194/tc-13-1125-2019, 2019.

Zhang, X., Zwiers, F. W., and Hegerl, G.: The influences of data precision on the calculation of temperature percentile indices, in: International Journal of Climatology, vol. 29, pp. 321–327, https://doi.org/10.1002/joc.1738, 2009.

Zscheischler, J. and Seneviratne, S. I.: Dependence of drivers affects risks associated with compound events, Science Advances, 3, e1700 263, https://doi.org/10.1126/sciadv.1700263, 2017.

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., and Zhang, X.: Future climate risk from compound events, Nature Climate Change, p. 1, https://doi.org/10.1038/s41558-018-0156-3, 2018.

Zscheischler, J., Fischer, E. M., and Lange, S.: The effect of univariate bias adjustment on multivariate hazard estimates, Earth System Dynamics, 10, 31–43, https://doi.org/10.5194/esd-10-31-2019, 2019.

nature climate change

PERSPECTIVE
https://doi.org/10.1038/s41558-019-0609-3

Towards more predictive and interdisciplinary climate change ecosystem experiments

Francois Rineau[1]*, Robert Malina[2], Natalie Beenaerts[1], Natascha Arnauts[1], Richard D. Bardgett[3], Matty P. Berg[4,5], Annelies Boerema[6], Liesbeth Bruckers[7], Jan Clerinx[1], Edouard L. Davin[8], Hans J. De Boeck[9], Tom De Dobbelaer[10], Marta Dondini[11], Frederik De Laender[12], Jacintha Ellers[4], Oscar Franken[4], Lucy Gilbert[13], Lukas Gudmundsson[8], Ivan A. Janssens[9], David Johnson[3], Sebastien Lizin[2], Bernard Longdoz[14], Patrick Meire[6], Dominique Meremans[6], Ann Milbau[10], Michele Moretti[2,15], Ivan Nijs[9], Anne Nobel[2], Iuliu Sorin Pop[16], Thomas Puetz[17], Wouter Reyns[1,12], Jacques Roy[18], Jochen Schuetz[16], Sonia I. Seneviratne[8], Pete Smith[11], Francesca Solmi[7], Jan Staes[6], Wim Thiery[8,19], Sofie Thijs[1], Inne Vanderkelen[19], Wouter Van Landuyt[10], Erik Verbruggen[9], Nele Witters[2], Jakob Zscheischler[8,20,21] and Jaco Vangronsveld[1,22]

Despite great advances, experiments concerning the response of ecosystems to climate change still face considerable challenges, including the high complexity of climate change in terms of environmental variables, constraints in the number and amplitude of climate treatment levels, and the limited scope of responses and interactions covered. Drawing on the expertise of researchers from a variety of disciplines, this Perspective outlines how computational and technological advances can help in designing experiments that can contribute to overcoming these challenges, and also outlines a first application of such an experimental design.

Climate change is expected to have an impact on ecosystem communities and ecosystem functioning[1]. Crop yields[2], carbon sequestration in soil[3] and pollination rate[4] are generally predicted to decrease, while land evapotranspiration[5] and tree mortality, especially in the boreal region, are expected to increase[6]. At the same time, the redistribution of species will increase opportunities for pest and pathogen emergence[1].

Ecosystem functions are crucial for human well-being, and impacts on them will have important consequences for society[7]. However, refining the estimations of societal cost remains a challenge, partly because of large gaps in our knowledge of the amplitude and dynamics of these responses that make it difficult to plan for climate adaptation. Specifically designed climate change experiments are necessary to address these issues.

The goal of this Perspective is fourfold. First, while acknowledging the great advances achieved so far by experiments on ecosystem responses to climate change, we identify the challenges that many of them currently face: high complexity of climate change in terms of environmental variables, constraints in the number and amplitude of climate treatment levels, and the limited scope with regard to responses and interactions covered. Second, to overcome these challenges, we propose an experimental design that can make use of improvements in computational and technological capabilities to capture more accurately the complexity of climate change in experiments; increase the number and range of climate treatment levels; and use an interdisciplinary approach to broaden the range of responses and interactions covered. Third, we outline an experiment that applies these design recommendations to demonstrate how it can enhance our capacity to understand and predict ecosystem responses to climate change. We describe the technical infrastructure used in this experiment, the climate manipulations, and the analysis pathway all the way to the evaluation of the changes in

[1]Environmental Biology, Centre for Environmental Sciences, Hasselt University, Diepenbeek, Belgium. [2]Environmental Economics, Centre for Environmental Sciences, Hasselt University, Diepenbeek, Belgium. [3]Department of Earth and Environmental Sciences, University of Manchester, Manchester, UK. [4]Department of Ecological Science, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands. [5]Groningen Institute for Evolutionary Life Science, Community and Conservation Ecology Group, Groningen University, Groningen, the Netherlands. [6]Ecosystem Management Research Group (ECOBE), University of Antwerp, Antwerp, Belgium. [7]Data Science Institute, Hasselt University, Diepenbeek, Belgium. [8]Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland. [9]Plants and Ecosystems (PLECO), University of Antwerp, Wilrijk, Belgium. [10]Research Institute for Nature and Forest (INBO), Brussels, Belgium. [11]Institute of Biological and Environmental Sciences, University of Aberdeen, Aberdeen, UK. [12]Namur Institute of Complex Systems, and the Institute of Life, Earth, and Environment, Research Unit in Environmental and Evolutionary Biology, Namur University, Namur, Belgium. [13]Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK. [14]Gembloux Agrobio Tech, TERRA, University of Liege, Gembloux, Belgium. [15]Faculty of Business and Economics, University of Antwerp, Antwerp, Belgium. [16]Faculty of Sciences, Hasselt University, Diepenbeek, Belgium. [17]Institute for Bio- and Geosciences IBG-3, Forschungszentrum Juelich, Juelich, Germany. [18]Ecotron Européen de Montpellier, Centre National de la Recherche Scientifique, Montferrier-sur-Lez, France. [19]Department of Hydrology and Hydraulic Engineering, Vrije Universiteit Brussel, Brussels, Belgium. [20]Climate and Environmental Physics, University of Bern, Bern, Switzerland. [21]Oeschger Centre for Climate Change Research, University of Bern, Bern, Switzerland. [22]Department of Plant Physiology, Faculty of Biology and Biotechnology, Maria Curie-Sklodowska University, Lublin, Poland. *e-mail: francois.rineau@uhasselt.be

ecosystem services. Fourth, this design is placed within the larger context of climate change experiments, and we pinpoint its complementarity to other designs.

Challenges of climate change experiments

Climate change experiments face three types of challenge: limitations in addressing the complexity of climate change in terms of control of environmental variables; constraints in the number and range of climate level treatments; and restrictions in scope.

Complexity of climate change. The complex manner in which global climate change will affect local weather presents challenges for research into ecosystem responses. To mimic a future climate, factors such as air temperature, atmospheric CO_2 and precipitation need to be manipulated in combination, and this can be both conceptually and technologically challenging[8]. Therefore, a high proportion of climate change experiments have focused on measuring the effects of specific combinations of climate factors (such as warming plus drought), manipulated using technology that was available or affordable at that time (such as passive night-time warming and rain exclusion curtains)[9]. Although these experiments have led to many invaluable outcomes, such approaches cannot fully cover the complexity of climate projections or the covariance of meteorological variables. As such, they may, for example, under- or overestimate the effects on ecosystem functioning of changes in the frequencies of frosts and heatwaves, drought–heatwave reinforcements[10], interactions between soil moisture conditions and subsequent precipitation occurrence[11], increased frequencies of mild droughts (including in spring and autumn) and increased frequency of heavy precipitation events[12]. These climate alterations can have a strong influence on ecosystem functioning: for example, decreased frost frequency may have a considerable impact on plant mortality[13], and more frequent mild droughts can trigger plant acclimation and hence resistance to drought stress[14].

Many climate change experiments did not simulate an extreme event instead of a change in the mean for a given single factor; regimes of events instead of a single event for a given single factor; or complex coupling between multiple factors. This lack of refinement in climate manipulations is likely to have compromised the reliability of the estimation of ecosystem responses. Some steps have already been taken to address this, by applying treatments of precipitation regime or heatwaves as observed in the field[15,16] and by using translocation experiments, in which macrocosms are displaced across geographical gradients to expose them to other climates that match possible future conditions at the location of origin (the 'space for time' approach)[17]. However, such an issue cannot be solved by modelling alone, because it requires testing of too many possible interactions between factors, as well as changing regimes of single factors.

Number and range of climate treatment levels. The cost of specialized infrastructure often limits the number of experimental units that scientists can set up within a given experiment. Hence, climate factors are often applied at only two levels: ambient, and future projections[9]. This provides useful estimations of the direction of ecosystem responses but does not provide insights into the shape of the responses to these factors or how far away current conditions are from potential tipping points to alternative stable states[18]. Moreover, ecosystem responses to multifactor drivers of global change are regulated by complex, nonlinear processes[19], which makes modelling difficult with experimental data that come only from the two-level manipulation of environmental factors[20].

Also stemming from high equipment costs is the narrow range of climate treatments. Most experiments have kept this range within conservative boundaries[21], presumably because more extreme (although realistic) climate treatments may have a catastrophic impact on a studied ecosystem, potentially leading to the loss of expensively equipped replicates. The truncation of more extreme climate conditions has, in turn, led to a lack of evidence of their effects on ecosystem functioning.

Finally, low temporal resolution is an issue. A substantial proportion of climate change experiments have only measured the ecosystem dynamics or trajectories annually or seasonally. Such experiments may fail to detect short-term dynamics of ecosystem responses[22] or trajectories leading to a transition to an alternative stable state[23,24]. However, trends related to ecosystem dynamics often appear on decadal timescales, because of the time needed to alter biogeochemical cycles and the properties of soil organic matter. Therefore, the duration of the monitoring should be prioritized over its frequency if the set-up does not allow good coverage of both.

Integration among disciplines. The very nature of climate change and its impacts is discipline-spanning and therefore requires an integrated approach[25]. Although the number of interdisciplinary studies related to climate change is increasing steadily[26], there are still many challenges. These include establishing common terminology, concepts and metrics[25,27,28], a consistently lower funding success for interdisciplinary research projects[29] and a general lack of interdisciplinary research positions[25]. The barriers depend largely on the purpose, forms and extent of knowledge integration, and their combination[30]. Although climate change research developed from multidisciplinarity to interdisciplinarity, and further to transdisciplinarity[31], most collaborative work in environmental research is small-scale rather than large-scale interdisciplinary work[30]. Small-scale integration refers to collaborations between similar partners (for example, different natural science disciplines), whereas large-scale integration crosses broader boundaries (such as between natural and social science)[30]. Currently, ecosystem services studies are mostly limited to either the natural science aspects or the socio-economic science aspects and rarely cover the entire ecosystem services cascade[32]. This lack of large-scale knowledge integration results in errors along this cascade, both when moving from biodiversity and ecosystem functions to ecosystem services, and when moving from ecosystem services to societal values.

Recommendations

Here we present potential ways to address these challenges: improving computational and technological capabilities, increasing the number and range of climate treatment levels, and using an interdisciplinary approach.

Using climate model outputs and technology to refine treatments. A first option to prescribe changes in weather dynamics is to alter one environmental parameter in line with future predictions (such as drought duration or heatwave intensity), while keeping other climatic variables identical between treatments using high-frequency data of ambient weather conditions. The advantage of this method is that atmospheric conditions can be modified with high-quality field data instead of relying on less-precise regional climate model (RCM) outputs with lower spatial and temporal resolution. Moreover, if used to manipulate one climate factor at a time, such an approach aids a mechanistic understanding of ecosystem responses that can be further extrapolated through modelling. This design may combine two or more factors to provide information about interactions between climate parameters.

Incorporating the complexity of projected changes can also be achieved by using outputs of state-of-the-art climate models. Because of model biases, the appropriate model must be selected very carefully. Global climate models (GCMs) are useful tools for assessing climate variability and change on global to continental scales, typically with a spatial resolution of 100–250 km. To estimate climate variability at more local scales, GCMs are dynamically downscaled using RCMs, which resolve the climate at higher resolutions (typically 10–50 km). The GCM/RCM combinations can then be chosen on the basis of how well models perform against local climate and weather characteristics in the studied ecosystem, and how representative future projections are of the multimodel mean. In this case, one can simulate an ecosystem response to a given climate set-up with higher accuracy. However, unlike with a full factorial experiment, it is not possible to attribute an ecosystem response to a given climate factor. Nevertheless, the model-output approach does aid the application of increasingly high warming levels by using a global mean temperature gradient (see section on the Hasselt University ecotron experiment below). It also addresses the issues of covarying variables, and it can be directly linked with a scenario from the Intergovernmental Panel on Climate Change, which would represent a major step towards bridging the gap between climate and ecosystem science.

To implement these options, however, it is necessary to control climate conditions and atmospheric composition with high frequency and high accuracy. This can be achieved only with dedicated and advanced equipment. Ecotron infrastructures, which consist of a set of replicated experimental units in which environmental conditions are tightly controlled and multiple ecosystem processes are automatically monitored, are well suited to fulfil these needs[33]. Such infrastructures have been historically limited to a handful across the world[9] but are becoming increasingly widespread[34–36]. They also offer the opportunity to monitor ecosystem responses at sub-hourly frequencies, making it possible to discriminate between short- and long-term ecosystem responses.

Increasing the number and range of climate treatment levels. A gradient design, in which one or several climate factors are applied at increasing levels, can substantially increase the resolution of a climate change experiment. This is better suited to quantitatively describing the relationship between a response variable and a continuous climate factor than the more traditional approach of testing ambient versus a single future projection, and it allows the collection of quantitative data for ecological models[37]. It also makes it possible to detect nonlinearity, thresholds and tipping points, and to interpolate and extrapolate ecosystem responses[18]. Although such gradient designs should ideally be replicated, unreplicated regression designs can be a statistically powerful way of detecting response patterns to continuous and interacting environmental drivers, provided that the number of levels in the gradient is large enough[37].

To ensure appraisal of the largest possible range of ecosystem responses, the gradient should be as long as possible, even extending beyond the most extreme conditions expected. Broader treatment modalities can also inform us where a specific ecosystem response is situated relative to its upper or lower tolerance limit. In addition, the levels of the gradient may be spread nonlinearly to achieve the highest resolution in the range where the strongest ecosystem responses are expected.

Using an interdisciplinary approach to capture responses and interactions. We argue that an overarching objective of climate change experiments is to contribute to the understanding of the impacts that climate change has on nature and society, as well as to enlarge our potential for adaptation. However, as outlined above, the lack of large-scale knowledge integration can result in errors along the ecosystem services cascade, first in the step from biodiversity and ecosystem functions to ecosystem services, and second from ecosystem services to societal values.

Regarding the first step, thorough quantification of ecosystem services should be based on specific data on how the ecosystem is functioning. Many studies take land use as an indicator of ecosystem service delivery[32], but land-use classification often cannot capture differences between abiotic conditions and ecological processes that explain differences in service delivery[38]. Therefore, using land use as a simple indicator will result in inappropriate management decisions[38].

Regarding the second step, economists need to be involved early in the process. There are many ways in which ecosystem function changes can affect the provision of ecosystem services to society[39], but budget constraints necessitate the selection of those functions and services that are considered most important. A common selection approach is to consider the potential impact of ecosystem changes in terms of human welfare endpoints, often by means of monetary valuation. Ecologists and economists must interact across disciplinary boundaries if ecological experiments are intended to predict these endpoints within an ecosystem services context. Hence, economists need to be involved during the design of ecological experiments to ensure that the ecosystem service changes most relevant for human welfare are measured and predicted.

We suggest that the desired large-scale integration can be achieved in several steps, organized in a top-down approach. The first step is to identify the key ecosystem services to value, based on welfare endpoints[40]. For most terrestrial ecosystems, this would imply assessing services from the following list: food and raw material production and quality, water supply and quality, carbon sequestration, depollution, erosion prevention, soil fertility, pest and pathogen control, pollination, maintenance of biodiversity and recreation. The second step consists of identifying the set of variables that best describes the ecosystem functions, processes and structures associated with these services. Based on the literature[41], we suggest the following measures (see also Fig. 3): (1) vegetation variables (plant community structure, above/belowground biomass, litter quality); (2) atmospheric parameters (net ecosystem exchange, greenhouse gas emissions); (3) soil abiotic (pH, texture, electrical conductivity, macro- and micronutrient and pollutant content) and biotic (fauna and microbial community structure, mineralization rates, respiration and biomass) variables; and (4) all parameters that describe movements of water in the soil–plant–atmosphere continuum (precipitation, leaching, air relative humidity, evapotranspiration, water potential). Air and soil temperatures should also be monitored, as they determine biogeochemical reaction rates. Finally, ecosystem processes, structures and functions need to be translated into services and ultimately into societal value by expressing them in monetary and non-monetary terms. Measuring all of these variables, integrating them in an ecosystem service framework, and estimating the societal value of these services would require expertise from plant ecologists and ecophysiologists, hydrologists, soil biogeochemists, animal ecologists, microbiologists, pedologists and climatologists, as well as modellers and environmental economists[42].

The UHasselt Ecotron as an initial application

Here we describe the proposed interdisciplinary approach in the context of a climate change manipulation using the proposed Hasselt University ecotron experiment (UHasselt Ecotron).

Ecotron infrastructure. The UHasselt Ecotron facility consists of tightly controlled climate change manipulations of 12 macrocosms (soil–canopy columns of 2 m in diameter and 1.5 m depth), extracted without significant disruption of the soil structure from a dry heathland plot in the 'Hoge Kempen' National Park (50° 59′ 02.1″ N, 5° 37′ 40.0″ E) in November 2016, and placed in 12 separate ecotron units. The plot was managed for restoration 6 years before the sampling. The design of this infrastructure benefited from exchanges through the AnaEE (Analysis and Experimentation on Ecosystems)/ESFRI (European Strategy Forum on Research Infrastructure) project. Some of its features were inspired by the Macrocosms platform of the CNRS Montpellier Ecotron[16]. Each UHasselt Ecotron unit consists of three compartments: the dome, the lysimeter and the chamber. The shell-shaped dome is made

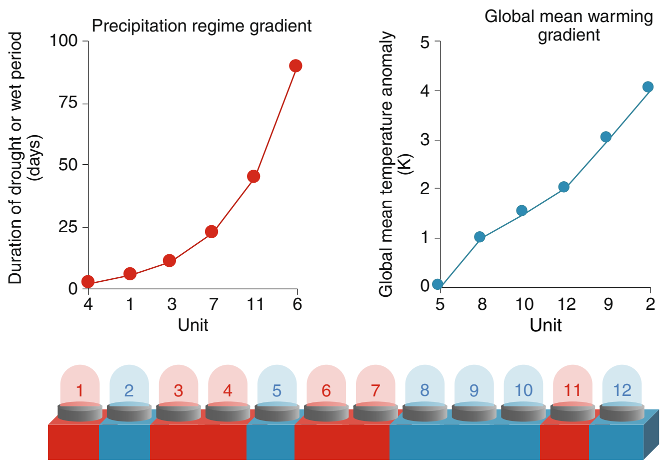

Fig. 1 | Overview of the two climate change gradient designs in the UHasselt Ecotron experiment. The 12 units shown here (each consisting of a dome, a lysimeter and a gastight chamber) have been spatially redistributed to maximize statistical similarity within a climate gradient before the treatment. Global mean temperature anomalies are computed with respect to the reference period 1951–1955.

of material that is highly transparent to photosynthetically active radiation. Within the dome, wind and precipitation are generated and measured, and the concentration of greenhouse gases (CO_2, N_2O, CH_4), photosynthetic photon flux density and difference between incoming and outgoing short- and long-wave radiation are measured. The lysimeter (which measures hydrological variations undergone by a body of soil under controlled conditions) contains the soil–canopy column, where soil-related parameters are controlled (including the vertical gradient of soil temperature and water tension) and measured, and is weighed every minute. Suction cups and soil sensors are installed following a triplicated five-depth design (Supplementary Fig. 1). The chamber is a gastight room that encloses the lysimeter, where air pressure, air temperature, relative humidity and CO_2 concentration are controlled, and key variables measured in each unit (Supplementary Fig. 1). The ecotron is linked with a nearby Integrated Carbon Observation System (ICOS) ecosystem tower (https://www.icos-ri.eu/home), which provides real-time data on local weather and soil conditions, with a frequency of at least once every 30 minutes.

Climate manipulations. A double-gradient approach is adopted: one approach (involving six of the ecotron units) measures the effect of an altered single factor (here, precipitation regime) while maintaining the natural variation of other abiotic factors; the other approach (six units) manipulates climate by jointly simulating all covarying parameters, representing increasingly intense climate change. The two approaches are described below. Because they sit isolated in an enclosed facility, it is possible that small initial differences in the soil–canopy core in a given unit will increase with time to the point where the unit becomes statistically different from the others. Therefore, the units were first distributed within the two gradients using a cluster analysis to minimize the noise in ecosystem responses measured during a test period (see Supplementary Fig. 2) due to small-scale soil heterogeneity. This clustering was used to distribute the units according to the pattern shown in Fig. 1.

Climate change projections for the northwest Europe region predict higher probability of both heavier precipitation and longer droughts, without a significant change in yearly precipitation[43]. The precipitation regime gradient uses real-time input from the ecosystem tower nearby, and only alters precipitation events: across the gradient, increasingly long periods (2, 6, 11, 23, 45 and 90 days, based on local climate records from Maastricht[44]) in which precipitation is withheld (dry period) are followed by increasingly long periods in which precipitation is increased (wet period), with the duration of the two periods kept equal within a unit (Fig. 1). Precipitation events during the wet period are increased twofold and are adjusted at the end of the period to avoid altering the yearly precipitation amount.

To drive the second gradient of the UHasselt Ecotron experiment, we use the climate variables produced by an RCM following Representative Concentration Pathway 8.5, a high-emission scenario[45]. The gradient itself is based on global mean temperature anomalies. In the six units, climates corresponding to a +0 °C to +4 °C warmer world (projected for periods ranging from 1951–1955 to 2080–2089) are simulated (Fig. 1, Supplementary Fig. 3), by extracting local climate conditions from the RCM for periods consistent with these warming levels (Supplementary Fig. 3)[46]. This setup also aids comparison of the 'present-day' climate as simulated by the RCM (the +1 °C unit) with the unit driven by ICOS field observations. Moreover, the climate simulated in the +1.5 °C unit is reasonably consistent with the lower end of the long-term temperature goals set by the Paris Agreement[47].

Integrating scientific disciplines for an interdisciplinary approach. As outlined in Recommendations, climate change experiments require large-scale knowledge integration to enable more useful estimates of climate change effects on ecosystem functioning and on society. The UHasselt Ecotron facility makes it possible to extend the degree of interdisciplinarity by investigating the entire cascade from climate changes to ecosystem functions, ecosystem services and, finally, societal values. As such, the facility contributes towards large-scale knowledge integration on climate change. Consequently, the ecotron experiment brings together several disciplines in an interdisciplinary framework (Fig. 2). With input from other involved disciplines, climatologists design the protocols for climate manipulations and plant ecologists monitor plant communities in each ecotron unit. Numerical models for water movement within one unit are developed by mathematicians and hydrologists. Ecotron output on carbon cycling is fed into a soil-carbon model[48], both for calibration and prediction purposes. Community modellers improve the power of this model by accounting for the soil community structure and species interactions (food web). The specific role of soil organisms in soil biogeochemistry is investigated by microbial and soil fauna ecologists. This is inferred from variation in responses of different functional groups such as nitrogen fixers, mycorrhizal fungi and different feeding guilds of soil fauna, combined with additional separate experiments, both in the field and in vitro. The outputs of the measurements above (see Fig. 3) allow experts in ecosystem ecology to quantify ecosystem services. Environmental economists express the change in ecosystem services provided, using best-practice monetization approaches[49]. For example, water quality regulation is assessed as the prevented cost of intensified water treatment or use of other water resources. Measurements of vegetation, soil abiotic parameters and the water balance make it possible to quantify this benefit. Carbon sequestration is assessed as the prevented cost from increased global temperature, which can be quantified based on measurements of vegetation, air parameters and soil abiotic parameters. Maintenance of biodiversity and recreation can be assessed from measurements of vegetation.

We note that (monetary) estimates from an individual study often cannot be applied directly for generating policy recommendations[50], especially for complex and spatially heterogeneous problems such as climate change impacts on ecosystems. However, meta-analyses need to rely on data generated by primary studies that estimate the societal cost (or benefit) of changes in specific services provided by a specific ecosystem at specific location(s). In this regard, the

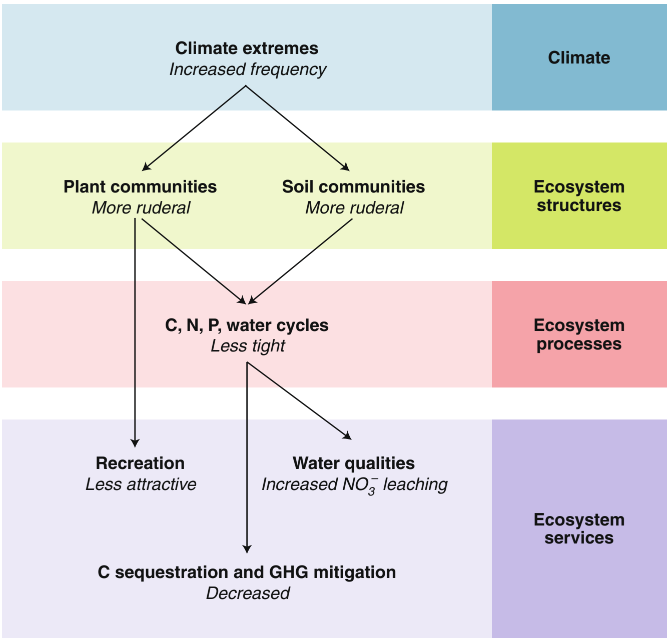

Fig. 2 | Impact pathway showing the reasoning behind the integration of scientific disciplines in the UHasselt Ecotron experiment. Ruderal species are fast-growing species colonizing disturbed environments. 'Less tight' indicates less internal recycling and more losses from the environment. C, carbon; GHG, greenhouse gas. The research hypotheses are given in italics and described in more detail in Supplementary Fig. 4.

UHasselt Ecotron experiment can also provide valuable input data for dedicated policy-guiding analyses[51].

Place of the design within the experimental landscape

A comprehensive understanding of ecosystem responses to climate change can only be achieved through a broad range of different, complementary experimental designs, all of which can be integrated through modelling. The experimental design suggested here exhibits a set of advantages and drawbacks that makes it suited to tackle specific needs within the landscape of climate change experiments.

Strengths and limitations of the design. The strengths of the suggested design comprise (1) high-performance microclimate conditioning, both above- and belowground, which makes it possible to approximate field conditions while maintaining control, (2) high-frequency automated measurements of ecosystem functions and thus of treatment impact thereon, and (3) a large-scale interdisciplinary approach. The first two strengths are inherent to the ecotron research infrastructure, whereas the large-scale integration could theoretically be implemented in any climate change experiment. However, we consider ecotron infrastructures to be particularly suitable for such an interdisciplinary approach, because of the high-end climate control and the broad range of functions monitored at a high frequency.

With respect to (1), studies focusing on ecosystem functions, processes and structures that are highly sensitive to soil temperature and soil water potential would benefit most from being conducted in ecotrons (for example, soil CO_2 exchange and carbon sequestration, growth and activity of soil microbes and soil fauna), as the lysimeter component can generate very precise lower boundary conditions and thus realistic vertical soil profiles of temperature and soil water status. With respect to (2), studies in which the high-resolution temporal pattern of ecosystem functions and their coupling is important would also benefit from ecotron infrastructures, as it is difficult to measure these parameters manually across long timescales. For example, simultaneous automated measurement of the carbon, water and mineral nutrient cycles makes it possible to disentangle their interactions in a range of climate conditions, and to feed control mechanisms into models.

A first set of constraints in the usefulness of the experimental design described here stems from the scale limitation of the experimental units. Ecotrons can accommodate plants of only small stature (less than 2 m in height), which excludes forests and tall crops. For the same reason, the impact of megafauna such as grazers or top predators cannot be tested. Results obtained in macrocosms integrate only small-scale (less than 1 m) variability, which leads to a lack of accuracy when scaling up to ecosystem.

Second, it may be difficult to financially support this type of experiment on the timescale of ecosystem responses (10 years or more)[52]. Ecosystem shifts to alternative stable states may remain undetected if the funding period is shorter than the period required for the shift. A partial solution for this would be to adopt a gradient design with increasingly late endpoints of projected climate change; this would allow for some extrapolation of ecosystem response in time (trajectories), which is possibly enough to estimate ranges of this response in the longer term.

Third, macrocosms in ecotron facilities are isolated from their ecosystem of origin. Hence, genetic input from propagules or pollination probably differs significantly from the field, which can be an issue, especially in long-term experiments. This could be mitigated in two ways. The first is by replacing soil sampling cores in the lysimeter by cores taken from the same ecosystem. If microbes and soil fauna are sampled not more than twice a year, using soil cores of 10 cm in diameter, this would account for disturbance of only 1.5% of total lysimeter surface annually. The second way is to use field traps to collect airborne propagules, which can be collected yearly and their contents spread on the enclosed surface of the soil–canopy columns. These solutions would at least ensure fresh genetic input into the system, even though this input may be different in the field in future conditions.

Finally, radiation in ecotron enclosures sometimes differs from that in the field. Artificial LED-lighting allows radiation to be controlled precisely but is yet not able to reach the same radiation level as in the field, while ambient lighting can disrupt its synchronization with temperature or precipitation. This may be an issue while simulating heatwaves and droughts, which have more sunshine hours than wet periods[53].

Complementarity with other climate change experiments. The weaknesses of the proposed design (small spatial scale, potentially insufficient timescale, lack of interaction with the surrounding environment) can be mitigated further through the use of complementary experiments, which might even be partially integrated into the overarching approach. For example, owing to small spatial scale, the results might have limited validity as a predictor of ecosystem responses at other sites and in other habitats. Running experiments in parallel across multiple climates and locations with the same methodology, also known as 'coordinated distributed experiments' (CDEs), would be better suited for this purpose as these experiments allow extrapolation and generalization of results while correcting for effect size[54]. For example, such a design makes it possible to study plant response to nutrient addition and herbivore exclusion[55], and ecological responses to global change factors across 20 eco-climate domains using a set of observatory sites[56]. In fact, a CDE using the UHasselt Ecotron design presented above and testing the same climate gradient in different ecosystems across several ecotron facilities would combine the high generalization potential of CDEs with the precision of ecotrons.

A second area for potential complementarity and integration is translocation experiments. These experiments are well suited for long-term observations, owing to their relatively low funding

Food	Raw materials	Water quality	C sequestration	Erosion prevention	Maintenance of biodiversity	Recreation	Climate regulation	Water retention	Soil fertility	Depollution	Pathogen control	Variable category	Variable	Frequency of measurement
●	●	●	●	●	●	●						Vegetation	Plant community structure	6 months
●	●	●	●	●	●								Shoot and root biomass	6 months
			●									Air parameters	Net ecosystem exchange	30 min
							●						Temperature	2 min
			●										GHG emissions (CH_4, N_2O)	2 min
				●				●	●			Soil abiotic parameters	Texture	1 year
							●	●					Temperature	2 min
		●							●				pH	1 year
			●						●				Macro and micronutrient concentration	1 year
		●						●					Electrical conductivity	30 min
		●						●					Soil pore water chemistry	2 weeks
										●			Available pollutant concentration	1 year
									●		●	Soil biotic parameters	Fauna community structure	6 months
				●					●		●		Microbial community structure	6 months
									●				Mineralization rate	1 year
		●		●				●				Water balance	Precipitation	30 min
		●						●	●	●			Leaching	30 min
		●						●					Relative humidity	30 min
		●						●					Evapotranspiration	30 min
		●						●					Soil water potential	30 min

Ecosystem service	Economic valuation
Water quality	Prevented cost of intensified water treatment or use of other water resources
C sequestration	Prevented damage cost from increased global temperature
Maintenance of biodiversity	Non-use value of continued existence of biodiversity
Recreation	Use value of recreational enjoyment

Fig. 3 | Measured variables in the UHasselt Ecotron experiment and links with ecosystem functions, services and values. Left-hand side of the table: ecosystem services. Right-hand side: variables measured in the ecotron experiment. Lower part of the table: illustration of how the societal value of four of the ecosystems services will be assessed.

requirements and ease of implementation, and the soil macrocosms used in these experiments are still connected to their surrounding environment[17]. However, the functioning of the ecosystem is monitored less comprehensively and frequently within these types of experiments, and the influence of different climate factors on ecosystem functioning cannot be disentangled. Consequently, running an ecotron and a translocation experiment in parallel on the same ecosystem with similar climate treatments would make it possible to estimate the effect size of the connection with the surrounding environment on ecosystem response to climate change. This information could then, in turn, be used to correct the outputs of future ecotron experiments by accounting for the isolation factor.

Usefulness of suggested design for modelling ecosystem response. Although ecosystem models can be evaluated and calibrated using a range of data sources, including sites in different climate zones and long-term experiments without climate manipulation[57], data from well-controlled, replicated and highly instrumented facilities such as those described here are invaluable for testing the process understanding encapsulated in the models, and for testing model behaviour against detailed, multiparameter observations[36]. Models that are tested and, where necessary, calibrated against such data can then be evaluated against data from other sites. If the outputs do not prove to be generalizable, the information derived from testing the model could be used to refine the experimental design and explain variation in the measured values. If the outputs prove generalizable, the models can be used across larger temporal and spatial scales to project potential impacts of future climate change[58,59].

Conclusion

The effects of climate change on ecosystem functioning have far-reaching consequences for society. Here we present a type of experiment that is designed to estimate the amplitude and dynamics of ecosystem responses to climate change, and the consequences for ecosystem services. We foresee that the holistic approach outlined in this Perspective could yield more reliable, quantitative predictions of terrestrial ecosystem response to climate change, and could improve knowledge on the value of ecosystem services and their links with ecosystem processes. We expect these results to be of interest for society beyond just scientists: they provide nature managers with predictions on ecosystem responses to help them decide on ecosystem management practices in the mid- and long-term, and they will explain to policymakers and the wider public the societal impact of ecosystem changes induced by climate change at a more detailed, ecosystem-specific level.

Received: 20 December 2018; Accepted: 25 September 2019;
Published online: 28 October 2019

References

1. Scheffers, B. R. et al. The broad footprint of climate change from genes to biomes to people. *Science* **354**, aaf7671 (2016).
2. Zhao, C. et al. Temperature increase reduces global yields of major crops in four independent estimates. *Proc. Natl Acad. Sci. USA* **114**, 9326–9331 (2017).
3. Allen, C. D. et al. A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Ecol. Manag.* **259**, 660–684 (2010).
4. Hatfield, J. L. & Prueger, J. H. Temperature extremes: effect on plant growth and development. *Weather Clim. Extremes* **10**, 4–10 (2015).
5. Collins, M. et al. in *Climate Change 2013: The Physical Science Basis* (eds. Stocker, T. F. et al.) 1029–1136 (IPCC, Cambridge Univ. Press, 2013).
6. Pecl, G. T. et al. Biodiversity redistribution under climate change: impacts on ecosystems and human well-being. *Science* **355**, eaai9214 (2017).
7. Millenium Ecosystem Assessment. *Ecosystems and Human Well-being: Synthesis* (Island, 2005); https://doi.org/10.1196/annals.1439.003
8. Leuzinger, S. et al. Do global change experiments overestimate impacts on terrestrial ecosystems? *Trends Ecol. Evol.* **26**, 236–241 (2011).
9. Stewart, R. I. A. et al. Mesocosm experiments as a tool for ecological climate-change research. *Adv. Ecol. Res.* **48**, 71–181 (2013).
10. Zscheischler, J. & Seneviratne, S. I. Dependence of drivers affects risks associated with compound events. *Sci. Adv.* **3**, e1700263 (2017).
11. Guillod, B. P., Orlowsky, B., Miralles, D. G., Teuling, A. J. & Seneviratne, S. I. Reconciling spatial and temporal soil moisture effects on afternoon rainfall. *Nat. Commun.* **6**, 6443 (2015).
12. Thiery, W. et al. Hazardous thunderstorm intensification over Lake Victoria. *Nat. Commun.* **7**, 12786 (2016).
13. Berendse, F., Schmitz, M. & Visser, W. De Experimental manipulation of succession in heathland ecosystems. *Oecologia* **100**, 38–44 (1994).
14. Backhaus, S. et al. Recurrent mild drought events increase resistance toward extreme drought stress. *Ecosystems* **17**, 1068–1081 (2014).
15. Verburg, P. S. J. et al. Impacts of an anomalously warm year on soil nitrogen availability in experimentally manipulated intact tallgrass prairie ecosystems. *Glob. Change Biol.* **15**, 888–900 (2009).
16. Roy, J. et al. Elevated CO_2 maintains grassland net carbon uptake under a future heat and drought extreme. *Proc. Natl Acad. Sci. USA* **113**, 6224–6229 (2016).
17. Cantarel, A. M. & Bloor, J. M. G. Four years of simulated climate change reduces above-ground productivity and alters functional diversity in a grassland ecosystem. *J. Veg. Sci.* **24**, 113–126 (2013).
18. Kreyling, J. et al. To replicate, or not to replicate — that is the question: how to tackle nonlinear responses in ecological experiments. *Ecol. Lett.* https://doi.org/10.1111/ele.13134 (2018).
19. Zhou, X., Weng, E. & Luo, Y. Modeling patterns of nonlinearity in ecosystem responses to temperature, CO_2, and precipitation changes. *Ecol. Appl.* **18**, 453–466 (2008).
20. Luo, Y. et al. Modeled interactive effects of precipitation, temperature, and [CO_2] on ecosystem carbon and water dynamics in different climatic zones. *Glob. Change Biol.* **14**, 1986–1999 (2008).
21. Kayler, Z. E. et al. Experiments to confront the environmental extremes of climate change. *Front. Ecol. Environ*. https://doi.org/10.1890/140174 (2015).
22. Svenning, J. C. & Sandel, B. Disequilibrium vegetation dynamics under future climate change. *Am. J. Bot.* **100**, 1266–1286 (2013).
23. Harris, R. M. B. et al. Biological responses to the press and pulse of climate trends and extreme events. *Nat. Clim. Change* **8**, 579–587 (2018).
24. Scheffer, M., Carpenter, S., Foley, J. A., Folke, C. & Walker, B. Catastrophic shifts in ecosystems. *Nature* **413**, 591–596 (2001).
25. Hein, C. J. et al. Overcoming early career barriers to interdisciplinary climate change research. *WIREs Clim. Change* **9**, 1–18 (2018).
26. Xu, X., Goswami, S., Gulledge, J., Wullschleger, S. D. & Thornton, P. E. Interdisciplinary research in climate and energy sciences. *WIREs Energy Environ.* **5**, 49–56 (2016).
27. Sievanen, L., Campbell, L. M. & Leslie, H. M. Challenges to interdisciplinary research in ecosystem-based management. *Conserv. Biol.* **26**, 315–323 (2012).
28. Abiven, S. et al. Integrative research efforts at the boundary of biodiversity and global change research. *Curr. Opin. Environ. Sustain.* **29**, 215–222 (2017).
29. Bromham, L., Dinnage, R. & Hua, X. Interdisciplinary research has consistently lower funding success. *Nature* **534**, 684–687 (2016).
30. Turner, L. M. et al. Transporting ideas between marine and social sciences: experiences from interdisciplinary research programs. *Elem. Sci. Anth.* **5**, 14 (2017).
31. Hellsten, I. & Leydesdorff, L. The construction of interdisciplinarity: the development of the knowledge base and programmatic focus of the journal *Climatic Change,* 1977–2013. *J. Assoc. Inf. Sci. Technol.* **67**, 2181–2193 (2016).
32. Boerema, A., Rebelo, A. J., Bodi, M. B., Esler, K. J. & Meire, P. Are ecosystem services adequately quantified? *J. Appl. Ecol.* **54**, 358–370 (2017).
33. Clobert, J. et al. How to integrate experimental research approaches in ecological and environmental studies: AnaEE France as an example. *Front. Ecol. Evol.* https://doi.org/10.3389/fevo.2018.00043 (2018).
34. Mougin, C. et al. A coordinated set of ecosystem research platforms open to international research in ecotoxicology, AnaEE-France. *Environ. Sci. Pollut. Res.* **22**, 16215–16228 (2015).
35. Eisenhauer, N. & Türke, M. From climate chambers to biodiversity chambers. *Front. Ecol. Environ.* **16**, 136–137 (2018).
36. Milcu, A. et al. Functional diversity of leaf nitrogen concentrations drives grassland carbon fluxes. *Ecol. Lett.* **17**, 435–444 (2014).
37. Cottingham, K. L., Lennon, J. T. & Brown, B. L. Knowing when to draw the line: designing more informative ecological experiments. *Front. Ecol. Environ.* **3**, 145–152 (2005).
38. Van der Biest, K. et al. Evaluation of the accuracy of land-use based ecosystem service assessments for different thematic resolutions. *J. Environ. Manag.* **156**, 41–51 (2015).
39. Polasky, S. & Segerson, K. Integrating ecology and economics in the study of ecosystem services: some lessons learned. *Annu. Rev. Resour. Econ.* **1**, 409–434 (2009).
40. Costanza, R. et al. Changes in the global value of ecosystem services. *Glob. Environ. Change* **26**, 152–158 (2014).
41. Braat, L. C. & de Groot, R. The ecosystem services agenda: bridging the worlds of natural science and economics, conservation and development, and public and private policy. *Ecosyst. Serv.* **1**, 4–15 (2012).
42. Plaas, E. et al. Towards valuation of biodiversity in agricultural soils: a case for earthworms. *Ecol. Econ.* **159**, 291–300 (2019).
43. Brouwers, J. et al. *MIRA Climate Report 2015: About Observed and Future Climate Changes in Flanders and Belgiu*m (Flanders Environment Agency, 2015); https://doi.org/10.13140/RG.2.1.2055.8809
44. Klein Tank, A. M. G. et al. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Int. J. Climatol.* **22**, 1441–1453 (2002).
45. van Vuuren, D. P. et al. The representative concentration pathways: an overview. *Clim. Change* **109**, 5–31 (2011).
46. Seneviratne, S. I., Donat, M. G., Pitman, A. J., Knutti, R. & Wilby, R. L. Allowable CO_2 emissions based on regional and impact-related climate targets. *Nature* **529**, 477–483 (2016).
47. UNFCCC. Paris Climate Change Conference November 2015, COP 21. *Adoption of the Paris Agreement. Proposal by the President* **21932**, 32 (UNFCCC, 2015).
48. Smith, J. et al. Estimating changes in Scottish soil carbon stocks using ECOSSE. II. Application. *Clim. Res.* **45**, 193–205 (2010).
49. Schaubroeck, T. et al. Environmental impact assessment and monetary ecosystem service valuation of an ecosystem under different future environmental change and management scenarios; a case study of a Scots pine forest. *J. Environ. Manag.* **173**, 79–94 (2016).
50. Hunter, J. E. & Schmidt, F. L. Cumulative research knowledge and social policy formulation: the critical role of meta-analysis. *Psychol. Public Policy Law* **2**, 324–347 (1996).
51. Gerstner, K. et al. Will your paper be used in a meta-analysis? Make the reach of your research broader and longer lasting. *Methods Ecol. Evol.* **8**, 777–784 (2017).
52. Knapp, A. K. et al. Past, present, and future roles of long-term experiments in the LTER network. **62**, 377–389 (2012).
53. De Boeck, H., Dreesen, F. E., Janssens, I. A. & Nijs, I. Climatic characteristics of heat waves and their simulation in plant experiments. *Glob. Change Biol.* **16**, 1992–2000 (2010).
54. Fraser, L. H. et al. Coordinated distributed experiments: an emerging tool for testing global hypotheses in ecology and environmental science. *Front. Ecol. Environ*. https://doi.org/10.1890/110279 (2013).
55. Lind, E. M. et al. Life-history constraints in grassland plant species : a growth-defence trade-off is the norm. *Ecol. Lett.* **16**, 513–521 (2013).
56. Keller, M., Schimel, D. S., Hargrove, W. W. & Hoffman, F. M. A continental strategy for the National Ecological Observatory Network. *Front. Ecol. Environ.* **6**, 282–284 (2008).
57. Smith, P. et al. Towards an integrated global framework to assess the impacts of land use and management change on soil carbon: current capability and future vision. *Glob. Change Biol.* **18**, 2089–2101 (2012).
58. Richards, M. et al. High-resolution spatial modelling of greenhouse gas emissions from land-use change to energy crops in the United Kingdom. *GCB Bioenergy* **44**, 627–644 (2017).
59. Song, J. et al. A meta-analysis of 1,119 manipulative experiments on terrestrial carbon-cycling responses to global change. *Nat. Ecol. Evol.* **3**, 1309–1320 (2019).

Acknowledgements

We thank the Flemish government (through Hercules Stichting big infrastructure and the Fund for Scientific Research Flanders project G0H4117N) and 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 comments on the experimental design. We also thank Regional Landscape Kempen and Maasland for its collaboration and support. N.W., S.L., A.N. and I.V. are funded by Research Foundation-Flanders (FWO).

Author contributions

F.R. and R.M. took the lead in writing the manuscript and received input from all co-authors. The initial conceptualization of this manuscript was discussed during a consortium meeting. All authors proofread and provided their input to different draft versions and gave their final approval for submission.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41558-019-0609-3.

Correspondence should be addressed to F.R.