1 Response to Referee 1: Stone, Daithi

This manuscript examines seasonal conditions in 2015-2016 in northern France using a type of stochastic weather generator. Technically I think everything in terms of the analysis is probably good, worth publishing, and not in need of modification. However, there were a number of aspects of the write-up that did not quite make sense to me, so I would request some work on the write-up.

We thank the reviewer for carefully reading our manuscript and for all the constructive comments. We think that they considerably improved the manuscript!

There are two main issues. The first is that the logical order is often confusing. An example is in the abstract, where one sentence says that extreme weather caused a certain low wheat yield, with the following sentence declaring that the connection has not been demonstrated yet. This is a contradiction. But if the sentences switched order, and with the addition of a "however" and some qualifiers in the second sentence, the contradiction would be lost, and I think the result would state what the authors intend. I have highlighted specific examples below.

The second issue is that this is supposed to be a demonstration of a certain technique to the problem of compound weather extremes. I am not clear though how the "compound" aspect really enters. The manuscript examines two types of events independently. The only connection is that the real-world motivating events occurred within a year of each other and that they may have been involved in the poor wheat yield that year (although it is not quite clear how). How is this "compound" and not just two events? If the interest is on seasonal weather events that affect wheat yield then why not also examine dry springs, hot summers, etc.? I do not see any connection between the two types of events as they are examined within the paper (i.e. ignoring that the motivating events happened to occur in the same year). In the hypothetical situation that the manuscript were split in two, each with one analysis per event type, would anything be lost from the current manuscript? What is it? This seems a salient concern for a special issue on compound events.

We agree with the reviewer, that from a technical point of view our study does not advance the understanding of compound events. As the reviewer pointed out, we study two events (warm December and wet spring) that appear to have happened independently from each other. We would however argue that in combination, these two events are indeed a compound event as they happened in the same place and likely lead to an impact in crop yields. Also it is expected that the observed crop loss would not have been as severe if either of the two components of the compound event would not have happened. In that sense the combination of an abnormally warm December followed by a wet spring is a compound event irrespective of whether there is a link between the two components. Furthermore, we would want to note that it isn't trivial that these two components of the compound event are independent of each other. This is a part of our results which justifies the rather simple technical approach of simulating both components (warm December and wet spring) individually.
We reframed the introduction section to clarify what exactly we are studying and why the observed event is indeed a compound event (lines 33-45).
It would have been interesting to study how this compound event actually led to a crop failure. This would however require a totally different analysis focusing more on the plant phenology and potential pest conditions. We fully agree that this would be a pertinent study. Nevertheless, we think that our analysis of potential worst cases of the meteorological compound event that is linked to the crop loss is a relevant contribution to a special issue on compound events.

One additional note that may or may not be relevant. My understanding is that ESD requests interdisciplinary submissions. This manuscript presents an analysis of climate data using a new statistical tool, but beyond that I do not see any interdisciplinarity.

We see the reviewer's point. Our paper is part of a special issue, which is the outcome of a European COST action (DAMOCLES) on compound events. The Editor of ESD went ahead with sending the manuscript to review, so there was no perceived problem with this. The paper is about the application of a novel statistical model to simulate *climate variability (interdisciplinarity stems from the underlined terms). One could argue that this justification is weak, but browsing through the last 20 papers published in ESD, we count no less than 14 papers that are probably no more interdisciplinary than ours! We also agree that this jurisprudence argument is debatable, but leave this decision to the Editor.*

Specific comments:
lines 1-4 It would seem more logical to me to switch the order of these two sentences. "The cause of this extreme event.... However, this event was likely in part due.... Here we focus on a compound...."

We agree with the reviewer and changed the manuscript accordingly.

lines 9-13 These sentences seem more appropriate for a methods section within the main text. While I can see a possible way in which they address the second question, these lines do not indicate how the method might address the first question.

For the first question we need to estimate the worst case event that could possibly happen under current climate conditions. In that way the questions that we presented in the original submission were a bit redundant. We therefore changed the manuscript and wrote about one question in the revised version (line 41). Nevertheless, we think that this is the right place to introduce the method we want to use to estimate the worst case event.

line 19 "of 2016" -> "of 2016's"

Done

line 22 "producer and exporter" -> "producers and exporters"

Done

line 23 France's trade balance in 2016 was -14.95 billion USD, which was actually higher than any year since 2006, except barely 2015. So -2.3 billion USD does not seem a "dramatic impact" to me.

We agree with the reviewer and changed "can have a dramatic impact on the national economy" to "can impact the national economy".

line 26 What is "oec"?

OEC is the Observatory of Economic Complexity. The way we managed the reference in Bibtex made it appear this way. We corrected this in the revised manuscript (line 26).

line 35 "gained a lot of attention" from whom? The reference is six years old.

We agree with the reviewer that this statement was not sufficiently backed by the reference. In the revised manuscript we reframed the introduction to address some other comments of the reviewers and the statement does not appear as such anymore.

line 44 Why is reanalysis data relevant here? There is good in situ monitoring of seasonal temperature and precipitation in France going back many more decades.

We thank the reviewer for pointing out this inconsistency: We indeed don't need reanalysis data to analyze precipitation and surface temperature in northern France. Furthermore, as we explain in the paragraph there is an intrinsic challenge with estimating a worst case event based on one realisation of our climate (which is to some extent irrespective of the length of our observations). We therefore removed this sentence.
We still use reanalysis data for the classification of analogues. We agree that reanalysis data is not optimal. However, it is not very easy to get access to the Meteo France observational data and for this study, which is mostly a proof of concept we felt that using reanalysis data was sufficient. If we were to extend this study and feed our generated events into a crop model, we would need to use in situ data.

line 56 "non-stationary" -> "non-stationarity"

Done

line 57 This is effectively "stress-testing" and using scenarios, and has been standard practice in catastrophe analysis and emergency preparedness since well before process-based models were available or sufficient data for training empirical were available.

This is a good remark. We have added a sentence in the manuscript (lines 62-64) to reflect this: "This kind of "stress-testing" based on the use of scenarios has been standard practice in catastrophe analysis and emergency preparedness, even outside of the context of climate change (see for example De Bruijn et al. (2015))."
de Bruijn, K. M., Lips, N., Gersonius, B. and Middelkoop, H.: The storyline approach: a new way to analyse and improve flood event management, Nat. Hazards, 81(1), 99–121, doi:10.1007/s11069-015-2074-2, 2016.

lines 71-72 I am not following this. What uncertainties are we talking about here? You are setting up conditions to mimic 2016, so is the uncertainty just a measure of how well you succeeded?

Indeed, we focus on the "range of values" of the simulated events. We do not treat uncertainties per se.

line 93 "raw" sounds like not detrended, but then that contradicts the rest of the sentence.

Indeed, this sentence is misleading. We have removed it from the manuscript (line 99).

lines 79-80 Which also includes all of Belgium and bits of the UK, Germany, and Switzerland (I think we can ignore Luxembourg).

We included the comment into the manuscript (lines 83-86):
"This region also includes parts of the UK, Germany, Belgium and Switzerland and does therefore not exactly match the studied area of Ben-Ari et al. (Ben-Ari et al., 2018). The seasonal meteorological conditions we study here are related to large scale events and averaging over a larger rectangle therefore seems appropriate."

line 102 At the end of what season? Is this something that is run for a season?

As the simulations are run individually for each simulated season, "at the end of the season" means "at the end of the simulation". We thank the reviewer for pointing out that this can be confusing and made it more explicit in the revised manuscript (line 108).

lines 103-104 Well... that depends on the size of the perturbations.

Indeed, perturbations have to be small for their method for their simulations to be physically consistent. We added that comment in line 110.

line 105-106 But if you delete one simulation, and then replace it with another plus perturbation, if the number of simulations is not large with respect to the number of days of simulation, then do you not end up with the case that your multiple realisations are essentially all the same except for the end? e.g. if you are simulating December- February with 30 simulations, then in the end the realisations will all be identical in December, only diverging some time in late January and February?

We think that there has been a misunderstanding here: We do not follow the approach proposed by Ragone et al. (2017). We only summarize their approach to give an overview of the theory behind stochastic weather generators. Our simulations solely rely on observed analogues that are assembled in a certain way by stochastic weather generators. It is true that for the dynamic weather generator differences between the trajectories increase with time. Trajectories diverge fast enough to produce different weather events for single month simulations.

lines 119-120 Why Z500 for one season and SLP for the other, and different regions?

We thank the reviewer for noting that some more information is required. While SLP is a standard choice for the classification of weather patterns, Z500 could be more appropriate for the simulation of heat events as the effect of temperatures on Z500 is smaller than on SLP. As this has already been discussed in Jézéquel et al. (2018) we refrain from going into detail again here. As shown in Fig. A1, the results are not sensitive to the choice of this variable.
The anticyclonic circulation that leads to warm December temperatures can well be characterized within a region covering most of continental Europe. For wet spring seasons, transient storms coming from the Atlantic are of interest which is why we chose a different region for the analogue definition of this event.
We extended the paragraph about the analogue selection in the experimental set up section (lines 164-171).

lines 154-157 You claimed above that an advantage of your approach is that it does not need the crazy amount of simulation taken by e.g. the weather@home approach. But 68000 years (or perhaps 68000*5 months) is not any smaller. But then it is unclear still whether you are using climate model output or reanalysis output (both are implied above). But with reanalysis output I do not see how one can "start independently" 1000 time series if there are only 79 values to start with (1 Decembers during 1950-2018).

Our approach does not require supercomputers or a complex infrastructure like weather@home to generate 68000 simulations. The approach relies on combining circulation analogues in a certain way to obtain random but extraordinary events. The random selection is independent (in the statistical sense) from one trajectory to another, even with identical initial conditions, as the simulated random numbers are IID.

line 159 "chosen a region" -> "a chosen region"

Done

Figure 3 caption, lines 2 and 4 "for 3 days" -> "for three more days"

Done

lines 166-169 This will get you partly there. But I wonder if it would work even better if you restricted possible switches to when it is not raining in either the original or new trajectories? This would avoid leaving or arriving in a trajectory during the middle of a potentially heavy rainfall event. The weighting mentioned in lines 170-171 seems to oppose your desire for physical consistency, by deliberately arriving into an imminent or potentially in-progress precipitation event.

We thank the reviewer for this useful suggestion! This would indeed be an interesting alternative. We chose to switch between trajectories after fixed intervals in order to be able to evaluate the sensitivity to the number of days after which we change the trajectories. This wouldn't be straightforward to analyse if we were to switch after a variable number of days. As this is the first attempt to adapt the framework of SWG's to precipitation events we prefer to stay as close as possible to the original (and well documented) SWG.
As we think that the suggestion refinement is worth being studied we briefly discuss it in lines 319-320. If the reviewer and editor think that this refinement would improve the manuscript, we would consider to include it.

line 181 Is this a strong test?

The Pearson correlation test is indeed a simple test. However, the result that both variables are not correlated is a rather strong indication for the absence of a link between these variables. We added the p-value of the correlation test in the revised manuscript (line 188).
In addition we now argue, that from an energy point of view it is unlikely that December temperatures influence precipitation 4 months later citing Peixoto Oort, 1992, sec. 14.6.2 (see lines 188-190).

line 184 Or did the presence of only a few cold days "lead" to the abnormally warm winter?

We fully agree with the reviewer that we cannot make any statement about a causal relationship here. We changed the sentence to “The winter preceding the 2016 crop loss was abnormally warm, with only a few cold days.” (line 193)

line 188 Does detrending Z500 but not temperature complicate things?

Detrending temperatures would require to work on temperature anomalies. In general, this could be useful. Here, as we also consider the number of cold days (with temperatures below 10°C), we have to work on absolute temperature and removing a trend would add a layer of uncertainty that has advantages and disadvantages.

line 200 Is the 4-significant digit precision of "4221" supported?

This number is estimated from a beta-binomial density distribution and we do not think that we can estimate the return period with that precision. In the revised manuscript we changed the number to 4000 (line 209).

line 229 Something is wrong here. Grammar?

This sentence was indeed almost incomprehensible. We reformulated it: “As expected, the inter-annual variations are smaller in the dynamic SWG simulations than in the static SWG simulations because the dynamic SWG evolves freely with the starting conditions as only tie to the observed circulation.” (lines 237-239)

line 236 "noa"?

This is again a wrongly displayed reference to the following website: https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/na The reference is displayed correctly in the revised manuscript (line 246).

lines 241-245 Perhaps you could elaborate on why this analysis is included (as with the similar analysis for the warm December). With my understanding of midlatitude dynamics at that time of year, it is hard to imagine a wet spring in northern France that would not involve low pressure just to the northwest of northern France (as in the figures). So is this a check that the method is working?

We include this analysis for two reasons: It shows the reader how the circulation of our simulations roughly looks like and it shows how the static SWG and the dynamic SWG differ in their results. We included the following sentence to clarify the purpose of this analysis: “We next analyse what kind of large scale circulation characterizes our SWG simulations by comparing them to a few exemplary observed events.” (lines 247-248)

lines 255-259 One thing that is nagging me here is that if wet events are due to the passage of a low pressure system, then as a general rule some high pressure must pass by before another low pressure system can.(yes, it is more complicated, but just thinking of canonical mid-latitude flow). I think your method, by requiring similarity on SLP and having the 5-day segments, probably works toward ensuring this in the static case. I am less clear on the dynamic case. If this is not ensured, then can this still be considered physically-based? Anyway, some commentary on that point here or early would be useful.

The reviewer raises a valid concern about the plausibility of the simulations of our dynamic SWG. As the SWG favours analogues that bring high precipitation, it mainly selects analogues with low pressure over northern France. We choose a relatively large region for the analogue definition, therefore these analogues mostly have an upcoming high pressure system over the Atlantic that would appear over northern France if the trajectory was followed. When switching trajectories most analogues that are considered have this high pressure system westward of northern France and due to the general eastward flow it will eventually pass over our region of interest.
We spend quite some time on thinking about how we can “quantify” the plausibility of our simulations without coming to a satisfactory conclusion.
We added gif’s showing a few simulations (and the observations) to the supplementary information to demonstrate the above described behaviour.

line 266-267 But it might be dynamically? Would one expect the same large-scale synoptic systems to be responsible for rainfall in April as in August?

We agree with the reviewer that our explanation was too simplistic here. We therefore removed our hypothesis from the revised manuscript.

line 282 "the most relevant" -> "a more relevant"

Agreed

lines 286-287 Your analysis does not suggest anything about whether "the mechanism that led to the crop failure in 2016 is... understood" nor "therefore... more extreme crop losses".

This statement was indeed based on a few assumptions that we failed to explicitly state. We thank the reviewer for noting that. We now changed the sentence to:
"If crop yields respond to the number of cold days in winter and to the precipitation rate in spring as shown in Ben-Ari et al. 2018, then we have shown here that in current climate an even worse crop loss event would be possible." (lines 307-309)

lines 287-288 "Especially": Did you examine other periods than April-July?

Here, the "especially" is meant to emphasize on the precipitation part of the compound event as compared to the warm December part.

line 291 How so "similar to a summer heat wave"? In that in order to get warmth in northern France one needs to have flow from the subtropics?

The event was similar with respect to the large-scale atmospheric flow. We reworded the sentence to:
"Although the circulation pattern of the warm December 2015 was similar to a summer heat wave with an anticyclonic pattern over France, special care was required to assure that our simulated events are actually realisations of winter weather." (lines 311-313)

lines 294-295 Which "weather event" of the "series of passing storms"?

We changed "weather event" to "wet spring season" in line 316.

lines 295-296 How does not having persistence make simulation a major challenge?

Here, the challenge arises from the functioning of SWGs: Choosing an analogue with high precipitation every day, easily leads to extremely persistent weather conditions as precipitation mostly occurs during low pressure conditions. This persistence can be a desired feature when simulating persistent heat waves for example. However, for a series of precipitation bringing storms the design of the SWG had to be changed. We clarified this in the revised manuscript (lines 315-318).

lines 296-297 How were "success" and "reasonableness" defined? I do not recall this being assessed. Line 298 suggests that it was not too.

As a first attempt to simulate plausible long lasting wet periods, we propose to reassemble 5-day chunks of observed weather instead of single days. Evaluating how plausible the constructed events are is a major challenge for these kind of studies and addressing this challenge is an important task for the future.
In the revised manuscript we extended the discussion on challenges in evaluating the plausibility of our simulations (lines 319-332).

line 305 What does "considerable chance of unprecedented winter rainfall" mean?

In their conclusions, Thomson et al. write: "There is a 34% probability of an unprecedented winter monthly rainfall total in at least one month in at least one region—it is therefore likely that we will see unprecedented winter rainfall within the UK in the next few years."
Here we solely want to highlight an alternative approach to the question of how the probability of unprecedented weather events can be accessed. In our view, the actual probability that Thompson et al. found for monthly rainfall extremes in England is less relevant here. We are however happy to include more details on the study if the reviewer or the editor thinks that this would be useful.

line 310 There is some dependence here on what you mean by "crop model". Models of crop behaviour can be process-based or based on empirical observation in controlled conditions. I am not sure anything based on correlation of observed yield against observed weather, which seems to be what you are referring to here, would normally be termed a crop model.

We fully agree with the reviewer: The model we are referring to here is not a process-based model but a model based on correlations between meteorological variables and wheat in northern France. We clarified this in the revised manuscript (see lines 333-336).

Figures A1, A3 Please describe the box and whiskers. I am confused by the double boxes (red and blue) within one set of whiskers.

We now slightly shifted the boxplots such that the whiskers are clearly distinguishable.

Figure A2 caption, line 2 "the parameter we choose" -> "the value used in the analysis"

Agreed

Figures A4, A5, A6 Please describe the box and whiskers.

Done

Figure A7 Are the colours, etc. the same as in the previous plots?

Yes, we added the information in the caption.

2 Response to Referee 2: Moreno Dumont Goulart, Henrique

The paper is based on a previous study (Ben-Ari et al., 2018) that identified two climatic conditions, more specifically not enough cold days in December of the year before harvesting and too much precipitation during the spring before harvesting, to be key factors in diminishing the wheat productivity in northern France for the year 2016. The paper works around these two compounded conditions and aims at exploring how extreme each of them actually is in terms of physical plausibility, finally establishing how rare this event of 2016 was and what are the odds of them happening again.

We thank the reviewer for reading our manuscript carefully and for providing constructive feedback which helped to improve the manuscript considerably!

In the introduction section, it would be perhaps interesting to explain the underlying factors that makes France a major wheat producer with high yields, maybe climate conditions or the practices used.

France is a major wheat producer due to its intensive agriculture generating high yields. Based on FAO data we changed the first sentence of the introduction to:
"France is one of the major wheat producers and exporters in the world, thanks to yields that are roughly twice as high as the world average (FAO, 2013)." (lines 22-23)

The methodology proposed by the authors, an adaptation of stochastic weather generators (SWG), is innovative in the field and duly addresses the original research question. In addition, it is data-driven based, suggesting a more flexible and cheaper approach to simulating extreme conditions with respect to the physical climate models.

Maybe explaining the methods before the data section would make more sense in this work?

We thank the reviewer for this suggestion. We agree that as it is now the data section is not written elegantly as we cannot directly refer to the part of our methods that requires the described datasets. However, the other way around we would have the same problem. We therefore prefer to stick with the methods section after the data section.

The authors mention the paper is designed following the storyline concept, however it seems a bit shallow and too implicit the theoretical conception, in spite of the main references being rightly cited. Some minor alterations in the section presentation (lines 57 – 64, especially this passage "In this paper, we construct a climate storyline of a warm winter followed by a wet spring that is likely to lead to extremely low wheat crop yield in France" could better demonstrate the rationale behind the storyline approach used and provide a clearer description of the importance of the storylines in the current work. In my perspective, it should be more explicitly explained that the starting point of the simulations stems from the 2015/2016 season and that the counterfactuals obtained are all based on these real occurrences.

The transition between the theoretical presentation of storylines and our own approach was a bit shallow indeed. We reworded as follows:
"These storylines have the potential to be fed into impact models and to provide a tool to engage with stakeholders about their vulnerability (de Bruijn et al (2016), Symstad et al (2017))."
In this paper, we propose a methodology to compute scenarios of plausible extreme meteorological events. This is a first step towards the storyline of a compound event consisting of a warm winter followed by a wet spring that is likely to lead to extremely low wheat crop yield in France. This methodology is based on an ensemble of simulations of temperature and precipitation with a stochastic weather generator that we nudge towards extreme behaviour. The starting point of these simulations stems from 2015/2016, from which we derived counterfactual events, that could have happened instead of the observed event.

On the data section, it would be profitable to justify the choice of averaging the rectangle encompassing the northern France (line 79). Ben-Ari, 2018 decided to average the area within each department of the country and justified this by stating there was not much spatial variability within each of these departments. Perhaps a similar justification could be added so others can better understand the reason this decision was made.

We agree that a comment about this rectangular region would be helpful here. We included the following sentences: "This region also includes parts of the UK, Germany, Belgium and Switzerland and does therefore not exactly match the studied area of Ben-Ari et al. 2018. The seasonal meteorological conditions we study here are large scale events and averaging over a larger rectangle therefore seems appropriate." (lines 83-86)
Also, as we stated in a reply to the first reviewer: We agree that reanalysis data is not optimal. However, it is not very easy to get access to the Meteo France observational data and for this study, which is mostly a proof of concept we felt that using reanalysis data was sufficient. If we were to extend this study and feed our generated events into a crop model, we would need to use in situ data.

The paragraph starting at line 251, which describes the way precipitation data were grouped, could be possibly improved in a way to better explain the decision behind the 5-day selection of the data chunk length. It is understandable that 1 day would not work well and that 5 days are a good representation of a coherent time series, but what prevents the chunks from being longer or slightly shorter? According to figure A4, 4 or 7 days could work as well. Perhaps some explanation on this side to justify the parameter value selection would add some value to the work.

We fully agree that the 5-day chunks are a heuristic choice. As the reviewer points out other chunk sizes would also work. In the paragraph 266-272 we shortly discuss the sensitivity of the results to the number of days before switching trajectories. We now added a sentence clarifying that there is a range of reasonable chunk sizes and that the 5-day chunks are a heuristic choice: "Note that taking 5-day chunks is a heuristic choice and that chunk sizes between four and seven days might work similarly well."

In addition, the following paragraph starting at line 261 behaves in a similar way but this time on the amount of precipitation alpha parameter and it is not exactly evident the choices behind selecting the chosen values. For both paragraphs, it is my opinion some further explanation on the reasoning behind the parameters choices will improve the general understanding of the work.

We thank the reviewer for the suggestion and agree that some more explanation on how parameters are chosen is required. We added the following paragraph:
"As for the other free parameters of the SWG, this sensitivity test does not directly justify the choice of the α parameter. It rather gives guidance on the values that would be appropriate choices for our application. In the end the parameter is heuristically chosen considering the trade-off of creating high precipitation events and keeping as much randomness as possible in our simulations." (lines 280-283)

In line 283 some reference would be welcome so that the cold days can be duly justified.

We added a reference to Ben-Ari et al. (2018) which is the main source of information for this event.

The conclusions section is clear and concise.
The last paragraph, line 321, holds a statement that could be better contextualized. Since the extreme events are within given scenarios, it is not exactly assessing all possibilities in the world (climatic or non-climatic). It may very well be it is not the purpose of the paper to account for that, but then it would be interesting to make explicit these limitations, such as the uncertainty of the scenarios, non-climatic drivers (pests, supply chain, management, economy and so on).

We reworded the last paragraph to reflect your point. It now reads:
"This approach is rather flexible and could be adapted to simulate compound extremes using climate model outputs based on different scenarios of climate change. This could lead to a first evaluation of the impact of climate change on worst case scenarios of crop yields. This type of data has some limitations, related to the uncertainty of models and scenarios, and it fails to take into account non-climatic drivers of crop yields such as pests, supply chain, or economical concerns. We however believe it could be useful to estimate what could be plausible in terms of purely meteorological events, in a changing climate." (lines 343-351)

Some minor mistakes encountered along the text:
Line 167: “the the”;

Done

Figure 6: “te black line”;

Done

Line 304: “Thompson et al” – no date;

Done

Simulating compound weather extremes responsible for critical crop failure with stochastic weather generators

Peter Pfleiderer[1,2,3], Aglaé Jézéquel[4,5], Juliette Legrand[6], Natacha Legrix[7,8], Iason Markantonis[10], Edoardo Vignotto[9], and Pascal Yiou[6]

[1]Climate Analytics, Berlin, Germany
[2]Humboldt University, Berlin, Germany
[3]Potsdam Institute for Climate Impact Research, Potsdam, Germany
[4]LMD/IPSL, ENS, PSL Université, École Polytechnique, Institut Polytechnique de Paris, Sorbonne Université, CNRS, Paris France
[5]Ecole des Ponts, Marne-la-Vallée, France
[6]Laboratoire des Sciences du Climat et de l'Environnement, UMR8212 CEA-CNRS-UVSQ, IPSL & U Paris-Saclay, 91191 Gif-sur-Yvette, France
[7]Climate and Environmental Physics, Physics Institute, University of Bern, Bern, 3012, Switzerland
[8]Oeschger Centre for Climate Change Research, University of Bern, Bern, 3012, Switzerland
[9]Research Center for Statistics, University of Geneva, Geneva, Switzerland
[10]National Centre of Scientific Research "Demokritos", INRASTES Department, Aghia Paraskevi, Greece

Correspondence: Peter Pfleiderer (peter.pfleiderer@climateanalytics.org)

Abstract. In 2016, northern France experienced an unprecedented wheat crop loss. ~~This~~ The cause of this event is not fully understood yet and none of the most used crop forecast models were able to predict the event (Ben-Ari et al., 2018). However, this extreme event was likely due to a sequence of particular meteorological conditions, i.e. too few cold days in late autumn-winter and an abnormally high precipitation during the spring season. ~~The cause of this event is not fully understood yet and none of the most used crop forecast models were able to predict the event (Ben-Ari et al., 2018).~~ Here we focus on a compound meteorological hazard (warm winter and wet spring) that could lead to a crop loss.

This work is motivated by ~~two main questions: were the~~ the question whether the 2016 meteorological conditions were the most extreme possible conditions under current climate? and what would be the worst case meteorological scenario ~~that would lead to the worst crop loss~~with respect to warm winters followed by wet springs? To answer these questions, instead of relying on computationally intensive climate model simulations, we use an analogue-based importance sampling algorithm that was recently introduced into this field of research (Yiou and Jézéquel, 2020). This algorithm is a modification of a stochastic weather generator (SWG) that gives more weight to trajectories with more extreme meteorological conditions (here temperature and precipitation). This approach is inspired from importance sampling of complex systems (Ragone et al., 2017). This data-driven technique constructs artificial weather events by combining daily observations in a dynamically realistic manner and in a relatively fast way.

This paper explains how a SWG for extreme winter temperature and spring precipitation can be constructed in order to generate large samples of such extremes. We show that, with some adjustments, both types of weather events can be adequately simulated with SWGs, highlighting the wide applicability of the method.

We find that the number of cold days in late autumn 2015 was close to the plausible ~~maximum~~minimum. But our simulations of extreme spring precipitation show that considerably wetter springs than what was observed in 2016 are possible. Although the crop loss of 2016's relation to climate variability is not fully understood yet, these results indicate that similar events with higher impacts could be possible in present-day climate conditions.

3 Introduction

France is one of the major wheat ~~producer and exporter~~ producers and exporters in the world, thanks to ~~very high yields~~ yields that are roughly twice as high as the world average (FAO, 2013). Given the prominent role of wheat production in France, crop failures can ~~have a dramatic impact on~~ impact the national economy. When an unprecedented disastrous harvest was registered in 2016, especially in ~~the northern region~~ northern parts of France, with a loss in production of about 30% with respect to 2015 (Ben-Ari et al., 2018), France registered heavy losses in farmers incomes and a loss of approximately 2.3 billion dollars in the yearly trade balance ~~(OEC, 2020).~~ (OEC, 2020).

Interestingly, the extreme crop failure of 2016 was not predicted by any forecasting model, which all strongly overestimated yields even just before the harvesting period (Ben-Ari et al., 2018). Thus, classical crop yield forecasting models, based on a combination of expert knowledge and data-driven methods (Müller et al., 2019; MacDonald and Hall, 1980), could not anticipate this unprecedented event because it was outside their training range. To overcome these limitations Ben-Ari et al. (2018) developed a logistic model that links the meteorological conditions in the year preceding the harvest with the probability of a crop failure. ~~They identified~~

In their study, Ben-Ari et al. 2018 attribute the crop loss to a combination of two meteorological events: an insufficient number of cold days in the December preceding the harvest and an abnormally high precipitation during spring~~as the main meteorological drivers of the observed crop failure.~~

~~In the past few years, the study of compounds events, i. e. the combination of two or more meteorological extreme events has gained a lot of attention (Leonard et al., 2014). Compound events often arise from the interaction of different climatic processes, an abnormally~~. It was argued that this low wheat yield was a preconditioned event wherein a mild autumn and winter favoured the build-up of biomass and parasites, which in combination with excess precipitation in late spring resulted in favourable conditions for root asphyxiation and fungus spread (ARVALIS, 2016). There could also be a direct influence of the meteorological conditions on plant development. For both potential mechanisms it is crucial to study the meteorological conditions leading to the crop loss as a compound event as only the combination of warm winter and ~~a very wet spring in this case, leading to a significant impact , here the extreme crop yields loss (Zscheischler et al., 2018)~~wet spring had this unprecedented impact on wheat yields (Zscheischler et al., 2020).

~~Crop losses in 2016 were associated with extreme winter temperatures and extreme spring precipitation.~~ The research question we want to address is: what would be a worst case meteorological scenario for this kind of crop loss event under current climate, with enhanced winter temperatures and spring precipitations? This ~~questions~~ question arises from the fact that we only lived one possible realisation of our climate. Even under unchanged climate conditions, unprecedented extreme events would occur as time goes on. Thus, to be able to put in place effective preventive measures, it is important to understand how severe an extreme event could be. ~~Unfortunately, estimating the worst case scenario of extreme crop failures of such a specific and rare event as the one observed in 2016 is challenging as we can rely on just a few decades of reliable reanalysis data.~~

To estimate how extreme a crop loss similar to the 2016 event could be, we need tools that all come with their assumptions and caveats. A standard approach would be to use large ensemble simulations based on circulation models of current climate conditions (Massey et al., 2015a). If the ensemble was large enough and physical mechanisms are adequately reproduced in the circulation model, one would find the most extreme possible version of the 2016 crop loss event and could even estimate its occurrence probability. This approach has two main drawbacks: the often huge computational cost associated with a large number of simulations and the possibly flawed representation of physical processes in climate models that could introduce a systematic uncertainty that cannot be overcome easily (Shepherd, 2019).

A second approach relies on the analysis of historical data. There are many statistical methods that could be used in this context. Specifically, copula-based techniques (Jaworski et al., 2010) can be used to study the dependence between two or more climate hazards, while models based on extreme value theory (Cooley, 2009) are suited for analyzing particularly rare events. These methods have the merit of being computationally cheap and of relying only on observed data, but dealing with ~~non-stationary~~ non-stationarity can be challenging with these methods.

As another data-driven alternative, the so-called storyline approach has emerged recently. The idea is to construct a physically plausible extreme event that one can relate to without necessarily focusing on the statistical likelihood of such an event (Hazeleger et al., 2015; Shepherd et al., 2018; Shepherd, 2019). Rather than asking what the most likely representation of the climate would be, one could ask how some extreme realisations of climate could be like. It has been argued that for adaptation planning the latter question could be more relevant (Hazeleger et al., 2015). This kind of “stress-testing” based on the use of scenarios has been standard practice in catastrophe analysis and emergency preparedness, even outside of the context of climate change (see for example de Bruijn et al. (2016)).

In this paper, we construct a climate storyline of a warm winter followed by a wet spring that is likely to lead to extremely low wheat crop yield in France. This storyline is based on an ensemble of simulations of temperature and precipitation with a stochastic weather generator that we nudge towards extreme behavior.

Here, we adapt analogue-based stochastic weather generators (SWGs) presented by Yiou (2014) and Yiou and Jézéquel (2020), which simulate spatially coherent time series of a climate variable, drawn from meteorological observations. Those SWGs were mainly tested on European surface temperatures. A version was developed to simulate extreme summer heatwaves (Yiou and Jézéquel, 2020). This paper optimizes the parameters of the SWG of Yiou and Jézéquel (2020) to simulate extreme warm winters (especially December) and extreme wet springs (especially May).

The goal is to construct a large sample of extreme climate conditions and assess the atmospheric circulation properties leading to those conditions of high temperatures and precipitation. The rationale of ensemble simulations is to determine uncertainties on the range of values that can be obtained.

Section 4 details the data that is used in this paper and explains the methodology of importance sampling with analogue simulators. Section ~~4.3~~ 5 describes the experimental results of the simulations of temperature and precipitation. Section ~~5 provides discussions on~~ 6 provides a discussion of the results. ~~The paper concludes with Section 6.~~

4 Methods

4.1 Data

We use temperature and precipitation observations from the E-OBS database (Haylock et al., 2008). The data is available on a 0.1×0.1 degree grid from 1950 to 2018. As an estimate of northern France temperature and precipitation we average these two fields over a rectangle spanning 1.5W-8.0E and 45.5N-51.5N (see Fig. 1). This region also includes parts of the UK, Germany, Belgium and Switzerland and does therefore not exactly match the studied area of Ben-Ari et al. (Ben-Ari et al., 2018). The seasonal meteorological conditions we study here are related to large scale events and averaging over a larger rectangle therefore seems appropriate.

We use the reanalysis data of the National Centers for Environmental Prediction (NCEP) (Kistler et al., 2001) for the analysis of atmospheric circulation. We consider the geopotential height at 500mb (Z500) and mean sea level pressure (SLP) over the North Atlantic region for computation of circulation analogues and *a posteriori* diagnostics. We used the daily averages between January 1st 1950 and December 31st 2018. The horizontal resolution is 2.5° in longitude and latitude. The rationale of using this reanalysis is that it covers 70 years and is regularly updated.

One of the caveats of this reanalysis dataset is the lack of homogeneity of assimilated data, in particular before the satellite era. This can lead to breaks in pressure related variables, although such breaks are mostly detected in the southern hemisphere and the Arctic regions (Sturaro, 2003), and marginally impacts the eastern north Atlantic region.

Z500 patterns are well correlated with western European temperature and precipitation, because those quantities and their extremes are related to the atmospheric circulation (Yiou and Nogaj, 2004; Cassou et al., 2005). Since Z500 values depend on temperature, we detrend the Z500 daily field by removing a seasonal average linear trend from each grid point. This preprocessing is performed to ensure that the ~~results do not depend on~~ analogue selection is not influenced by atmospheric trends. ~~All the analogue computations of this paper were performed on detrended and raw Z500 data, so that the results do not depend on Z500 trends.~~

4.2 Stochastic Weather Generators and importance sampling

The idea behind *importance sampling* is to simulate trajectories of a physical system that optimize a criterion in a computationally efficient way. Ragone et al. (2017) used such an algorithm to simulate extreme heatwaves with an intermediate complexity climate model.

The procedure of importance sampling algorithms, say to simulate extreme heatwaves with a climate model, is to start from an ensemble of S initial conditions and compute trajectories of the climate model from those initial conditions.

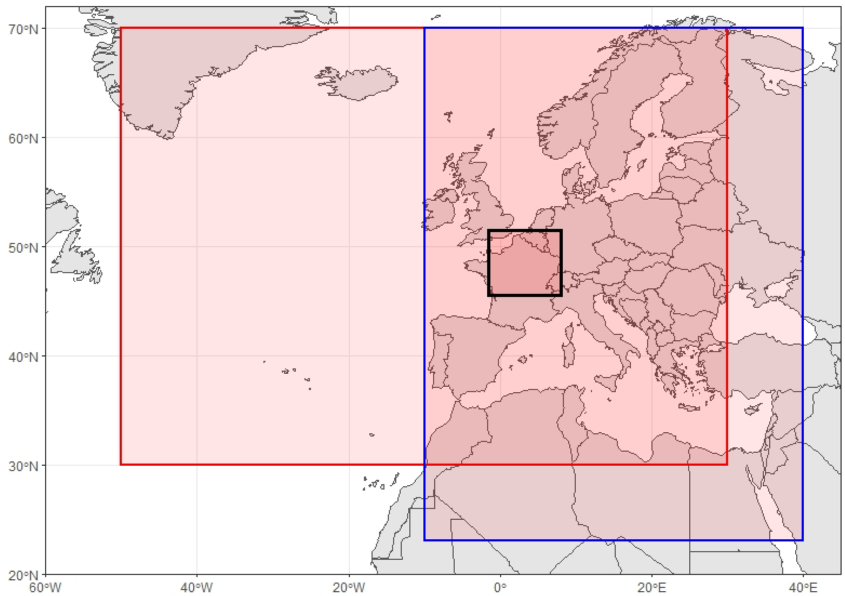

Figure 1. Regions used to identify circulation analogues for December temperatures (blue) and spring precipitation (red). The black rectangle indicates the region over which temperatures and precipitation are averaged in northern France.

An optimization *observable* is defined for the system. In this case, it can be the spatially-averaged temperature or precipitation over France. The trajectories for which the observable (e.g. daily average temperature) is lowest during the first steps of simulation are deleted, and replaced by small perturbations of remaining ones. In this way, each time increment of the simulations keeps trajectories with the highest values of the observable. At the end of ~~the season~~a simulation, one obtains S ~~simulations~~ trajectories for which the observable (here average temperature over France) has been maximized. Since those trajectories are solutions of the equations of a climate model, they are necessarily physically consistent ~~.~~ (given that the perturbations are small).

Ragone et al. (2017) argue that the probability of the simulated trajectories is controlled by a parameter that weighs the importance to the highest observable values: if 1 trajectory is deleted at each time step, the simulation of an ensemble of M-long trajectories has a probability of $(1-1/S)^M$. Hence one obtains a set of S trajectories with very low probability after M time increments, at the cost of the computation of S trajectories.

For comparison purpose, if one wants to obtain S trajectories that have a low probability (p) observable, then the number of necessary "unconstrained" simulations is of the order of M/p, so that most of those simulations are left out. Systems like weather@home (Massey et al., 2015b) that generate tens of thousands of climate simulations are just sufficient to obtain $S = 100$ centennial heatwaves, and the number of "wasted" simulations is very high. Therefore, importance sampling algorithms are very efficient ways to circumvent this difficulty. The major caveat of this approach is that one needs to know the equations that drive the system and be able to simulate them. We use an alternative method that does not require such knowledge on the system.

We use two stochastic weather generators (SWG) based circulation analogues (Yiou and Jézéquel, 2020) to simulate events of either warm temperature in December or high precipitation in Spring. These SWGs resample daily weather observations in a plausible manner to simulate new weather events (Yiou, 2014).

Circulation analogues are computed on SLP (or detrended Z500) from NCEP, between 1950 and 2018. For each day in 1950-2018, $K = 20$ best analogues are determined by minimizing a spatial Euclidean distance between SLP (or Z500~~maps. For winter temperatures, we used analogs of Z500 over a zone covering 23°N-70°N and 10°W-40°E. For spring precipitation, we used analogs of SLP over a zone covering 30°N-70°N and 50°W-30°E.~~) maps.

As explained by Yiou and Jézéquel (2020), the SWG randomly samples analogues by weighting the analogue days with a criterion that favors high temperatures or high precipitation. Hence, the importance sampling is summarized by the procedure of giving more weight to analogues that yield temperature (or precipitation) properties. There are two types of importance sampling for the analogues, which are illustrated in ~~Figure~~ figure 2.

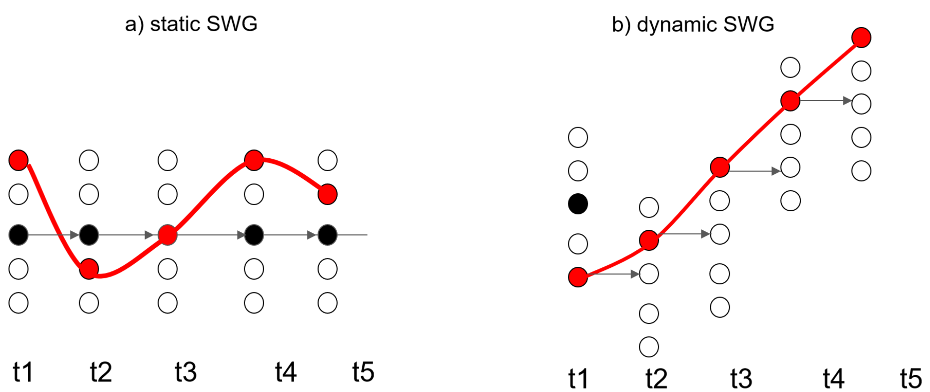

Figure 2. Illustration of the analogue-based importance sampling. (Left) The static SWG replaces each day in the observed trajectory (black dots) by one of its analogues (red dots). (Right) The dynamic SWG replaces the first day in the observations (black dot) by one of its analogues, reads the following day of this analogue and repeats the procedure until creating a new trajectory (red dots).

Those two main types of analogue SWGs are described by Yiou (2014) and Yiou and Jézéquel (2020):

1. A *static* weather generator replaces each day by one of its K circulation analogues *or* itself. With this type of SWG, simulated trajectories are perturbations (by analogues) of an observed trajectory.

2. A so-called *dynamic* weather generator has a similar random selection rule, *but* the "next" day to be simulated follows the selected analogue, rather than the observed actual calendar day. A probability weight ω_{cal} that is inversely proportional to the distance to the calendar day is introduced:

$$\omega_{\text{cal}} = A_{\text{cal}} e^{-\alpha_{cal} R_{cal}(k)},$$

where A_{cal} is a normalizing constant, $\alpha_{cal} \geq 0$ is a weight, and $R_{cal}(k)$ is the number of days that separate the date of kth analog to the calendar day of time t. This rule is important to prevent time from going "backward". This type of SWG generates new trajectories by resampling already observed ones. They are not just perturbations of observed trajectories.

Those random selections of analogues are sequentially repeated until a lead time T.

An importance sampling is applied while selecting an analogue at each time step by weighing probabilities with the variable to be optimized (temperature or precipitation). The $K = 20$ best analogues and the day of interest are sorted by daily mean temperature or precipitation. The probability weights are determined by Yiou and Jézéquel (2020). If $R(k)$ is the rank (in terms of temperature or precipitation) of day k in decreasing order and ω_k the probability of day k to be selected, we set:

$$\omega_k = Ae^{-\alpha R(k)} \tag{1}$$

where A is a normalizing constant so that the sum of weights over k is 1. The α parameter controls the strength of this importance sampling on temperature or precipitation.

The useful property of this formulation of weights is that the values of ω_k do not depend on time t, because the rank values $R(k)$ are integers between 1 and $K+1$. The weight values do not depend on the unit of the variable either, so that this procedure is the same for temperature or precipitation. If $\alpha = 0$, this is equivalent to a stochastic weather generator described by Yiou (2014).

Combining the weights on the calendar day and on the climate variable, the probability of day k to be selected becomes:

$$\omega'_k = Ae^{-\alpha R(k)} e^{-\alpha_{cal} R_{cal}(k)}. \tag{2}$$

The generators thus give more weight to the warmest or wettest days when computing trajectories of December temperature or Spring precipitation. We thereby simulate extreme events, e.g. warm Decembers and wet Springs (May to July).

4.3 Experimental Set Up

The parameters of the SWG depend on the variables and the seasons to be simulated. We determine those parameters experimentally and detail them hereafter. Table 1 lists all parameters used for the simulation of December temperature and spring precipitation. These parameters were set after performing a number of sensitivity tests that are going to be discussed in the result section. Table 2 lists all values tested for α and α_{cal}. Most figures related to these tests can be found in the appendix.

The procedure we follow is:

- Start and end day of simulations: For each year from 1950-2018, 1000 simulations are started independently for temperature in December and precipitation in spring. The temperature simulations start on the 1st of December and end on the 31st. Precipitation simulations start on the 1st of April and end on the 31st of July. This results in 68000 independent simulations of December temperatures and spring precipitation.

parameter	choice for warm Decembers	choice wet April-July periods
start day	01/12	01/04
end day	31/12	31/07
variable for analogues	Z500	SLP
region for analogues	70N-23N 10W-40E	-50W-30E and 30N-70N
weighting on temp. or precipitation (α)	0.75	0.5
weighting on calendar day (α_{cal})	6	0.5
number of days before selecting a new analogue (n_{days})	1	5

Table 1. Parameters used for the static and dynamic SWG to simulate warm Decembers (second column) and wet April-July period (last column).

- Identification of circulation analogues: Weather analogues are identified by evaluating the similarity of weather patterns of an atmospheric variable in ~~chosen a~~ a chosen region. For December temperature, analogues are based on detrended geopotential height at 500mbar (Z500) over a region covering most of Europe (~~70N-23N 10W-40E~~70°N-23°N 10°W-40°E) (see Fig. 1~~. Indeed,~~). Jézéquel et al. (2018) showed that Z500 is better suited than SLP to simulate temperature anomalies, and that rather small domains lead to better reconstitutions. This result is supported by sensitivity tests we performed on the choice of variable ~~and region~~ for the computation of the circulation analogues used to simulate December temperature. For spring precipitation~~we use mean sea level pressure (SLP) as a proxy of atmospheric circulation in a region including western Europe and~~ , we use analogs of SLP over a zone covering 30°N-70°N and 50°W-30°E as shown in figure 1. This region includes large parts of the ~~North Atlantic (50W-30E and 30N-70N) as shown in figure 1~~northern Atlantic where rain bringing storms usually come from.

- Number of days before selecting a new analogue: For the simulation of long lasting precipitation events the ~~the~~ consistency of day to day variability is important to ensure a plausible water vapour transport. We therefore adapt the ~~the~~ stochastic weather generator (both *static* and *dynamic*): instead of choosing a new analogue every day, we stay on an observed trajectory for a number of days (n_{days}) before choosing a new analogue (see Fig. 3). For the analogue selection we weigh the analogues based on the accumulated precipitation of the analogue and the following n_{days} days giving more weight to analogues that bring more precipitation in the following n_{days} days.

- Selection of circulation analogues by the generators: The α-parameter controls the strength of the importance sampling on either temperature of precipitation while α_{cal} controls the influence of the calendar date when selecting an analogue. For temperature simulations, we use $\alpha = 0.75$ and $\alpha_{cal} = 6$. Note that we thus strongly condition on the calendar day to restrict the SWGs to winter and late autumn days. For precipitation we set both α and α_{cal} to 0.5.

experiment	tested parameter	fixed parameters	tested values	figure
December	variable for analogues	$\alpha = 0.5$, $\alpha_{cal} = 6$, $n_{days} = 1$	Z500, SLP	Fig. A1
December	α_{cal}	$\alpha = 0.5$, $n_{days} = 1$	0, 0.2, 0.5, 1, 2, 4, 6, 8, 10	Fig. A2
December	α	$\alpha_{cal} = 6$, $n_{days} = 1$	0, 0.1, 0.2, 0.5, 0.75, 1	Fig. A3
April-July	n_{days}	$\alpha = 0.5$, $\alpha_{cal} = 0.5$	1, 2, 3, 4, 5, 7, 9	Fig. A4
April-July	α	$\alpha_{cal} = 0.5$, $n_{days} = 5$	0, 0.1, 0.3, 0.5, 0.7, 0.9, 1	Fig. A5
April-July	α_{cal}	$\alpha = 0.5$, $n_{days} = 5$	0, 0.2, 0.5, 1, 2, 5, 10	Fig. A6

Table 2. Performed sensitivity tests for the parameters used to simulate warm Decembers (first 3 rows) and wet April-July periods (last 3 rows). The second column lists the parameters of which the sensitivity is assessed. The third column indicates at which levels all other parameters are fixed for the test. The fourth column lists all tested values and the last column indicates the figure where the results of the test are shown.

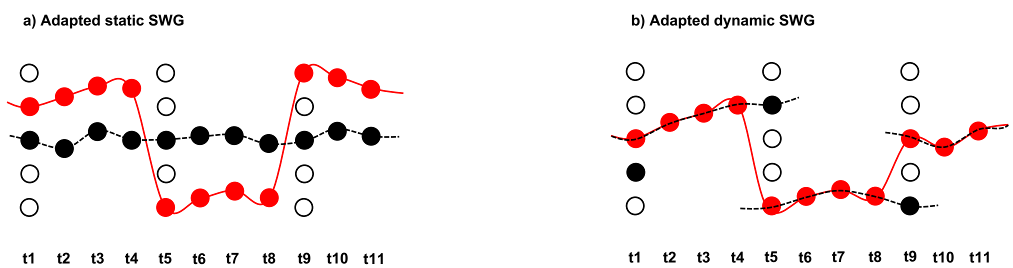

Figure 3. Adapted dynamic weather generators. a) The adapted static SWG selects a new analogue every n-th day (4th day in ~~tis~~ this illustration) and follows the observed trajectory (black dotted line) of that day for ~~3~~ three days. The resulting simulation combines observed 4-day chunks into an artificial trajectory (red line). b) The adapted dynamic SWG replaces the first day of the observations by one of its analogues and follows the observed trajectory of that analogue for ~~3~~ three days. Then a new analogue of the following day in the observed trajectory is chosen.

5 Results

A lack of cold days in December 2015 and an exceptionally wet spring caused the 2016 crop loss in northern France. Although the interplay between these two meteorological events is crucial for the resulting crop loss, the two events (warm December and wet spring) seem to have happened independently from each other~~. Indeed ,~~: the correlation between temperature in December and precipitation 4 months later is not significantly different from zero ~~.~~ and we cannot reject the hypothesis that both variables are not correlated (p-value of the Pearson correlation > 0.6). Also, from an energy point of view, the characteristic time scale of the atmosphere does not exceed 35 days (Peixoto and Oort, 1992, sec. 14.6.2). This implies that it is unlikely to find links between climate variables in December and the following May. We therefore consider that is it reasonable to simulate warm Decembers and wet springs independently.

5.1 December temperature simulations

The winter preceding the 2016 crop loss was abnormally warm~~leading to only~~ , with only a few cold days. Here, cold days are defined as days with daily maximal temperatures between 0 and 10°C. This December was the hottest in the observational record and also the December with fewest cold days.

Figure 4a shows the observed averages of daily maximal temperatures and the results from static and dynamic SWG simulations. The observed December temperatures fluctuate around 6°C with a small warming trend of 0.2°C per decade over the whole time series (p-value=0.03). Simulations from the static SWG are consistently around 3.5°C warmer and follow the year to year variability of the observations. With an average of 12°C, the dynamic SWG simulations are significantly warmer than the static SWG simulations and interannual variability is strongly reduced. This is to be expected as the dynamic SWG evolves freely from the starting day and is therefore less bound to each years circulation.

In years with higher December temperatures, the number of cold days with maximal temperatures between 0 and 10°C is reduced (see Fig. 4b). Over the period 1950–2018 no trend in the number of cold days is observed and the number of cold days fluctuates around 25 days. As the SWG simulates warmer Decembers the number of cold days is on average 8 days lower in the static SWG and 16 days lower in the dynamic SWG. Nearly half of the simulations of the dynamic SWG have thereby less cold days than what was observed in December 2015.

The 2015 December was unprecedented in terms of missing cold days and we simulate a number of warm Decembers with even fewer cold days. To estimate the probability of such extreme December, we fit a Beta-Binomial distribution (Jézéquel et al., 2018) to the observations and find that 2015 was a one in ~~4221~~ 4000 years event and that 25% of our dynamic SWG simulations are one in 1000 years events or even rarer (see A3).

As shown in figure 5, December 2015 was characterized by a persistent anticyclonic circulation with its center over the Alps. The circulation in the coldest December (1969) was opposite to 2015 with negative Z500 anomalies over Europe and positive anomalies over the Atlantic. In 2008, the December with most cold days in the observations, resembles 1969 but with less pronounced anomalies.

For all example years, the circulation in the static SWG simulations exhibits the same features as the observed circulation. The dynamic SWG always simulates high pressure anomalies over France irrespective of the starting conditions. These anomalies are however more pronounced in 2015 where the starting circulation favours the anticyclonic pattern over France.

The simulations of warm Decembers are most sensitive to the weighting on the calendar date. If this parameter is chosen too loosely, simulations would include days from other seasons which are generally warmer. As shown in figure A2, for ~~$alpha_{cat} \geq 6$~~ $\alpha_{cal} \geq 6$ over 70% of all days in the simulations are sampled from the November-February period. Increasing the weighting of the calendar day further doesn't show a significant effect.

The simulations are also sensitive to the weighting on daily maximal temperatures ~~$alpha$ (fig~~α (Fig. A3). For ~~$alpha \geq 0.75$~~ $\alpha \geq 0.75$ we simulate a large number of December's that are more extreme than 2015.

Finally the choice of geopotential height or mean sea level pressure to classify circulation analogues does not influence the simulations (see Fig. A1).

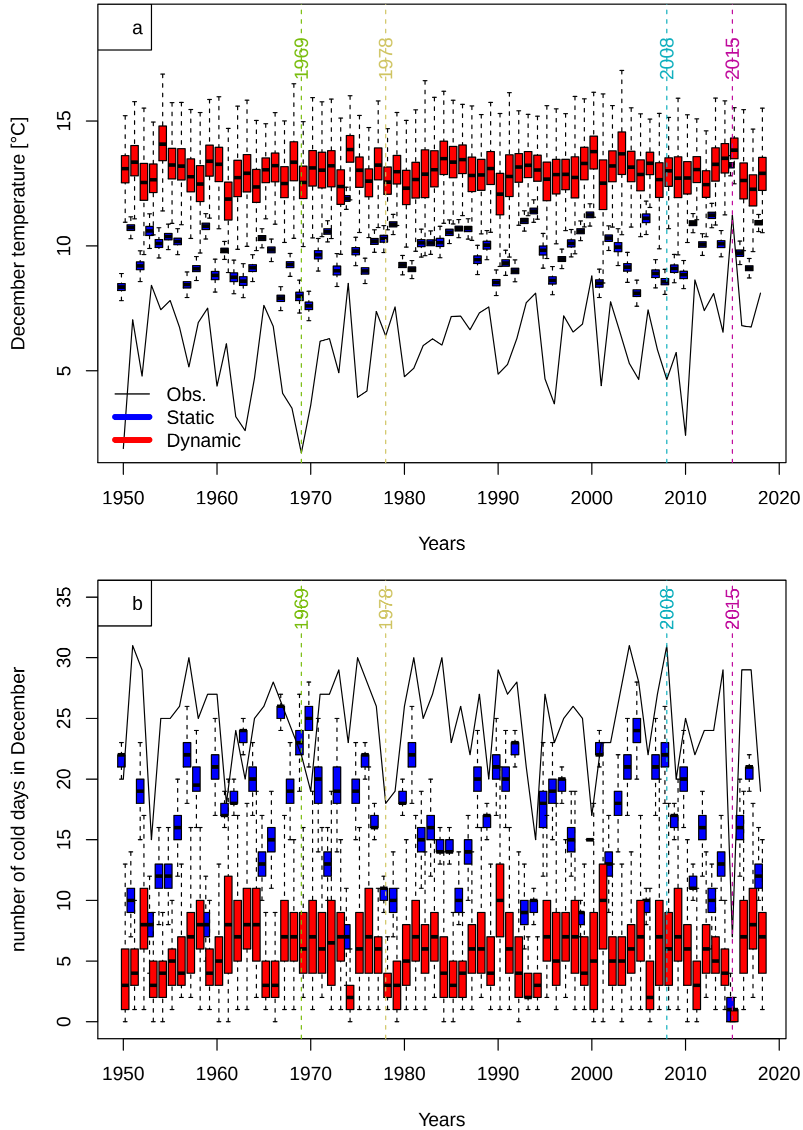

Figure 4. a) Daily maximal temperature in December from 1950 to 2018. ~~Te~~ The black line shows E-OBS observations. The boxplots represent the ensemble variability of the simulations of the static (blue) and the dynamic (red) SWG for each year. The boxes of boxplots indicate the median ($q50$), lower ($q25$) and upper ($q75$) quartiles. The upper whiskers indicate $\min[\max(T), 1.5 \times (q75 - q25)]$. The lower whisker has a symmetric formulation. The points are the simulated values that are above or below the defined whiskers. b) as a) but for the number of cold days. The vertical colored lines indicate the coldest December (green), a median December (yellow), a December with 31 cold days (cyan) and the warmest December (purple).

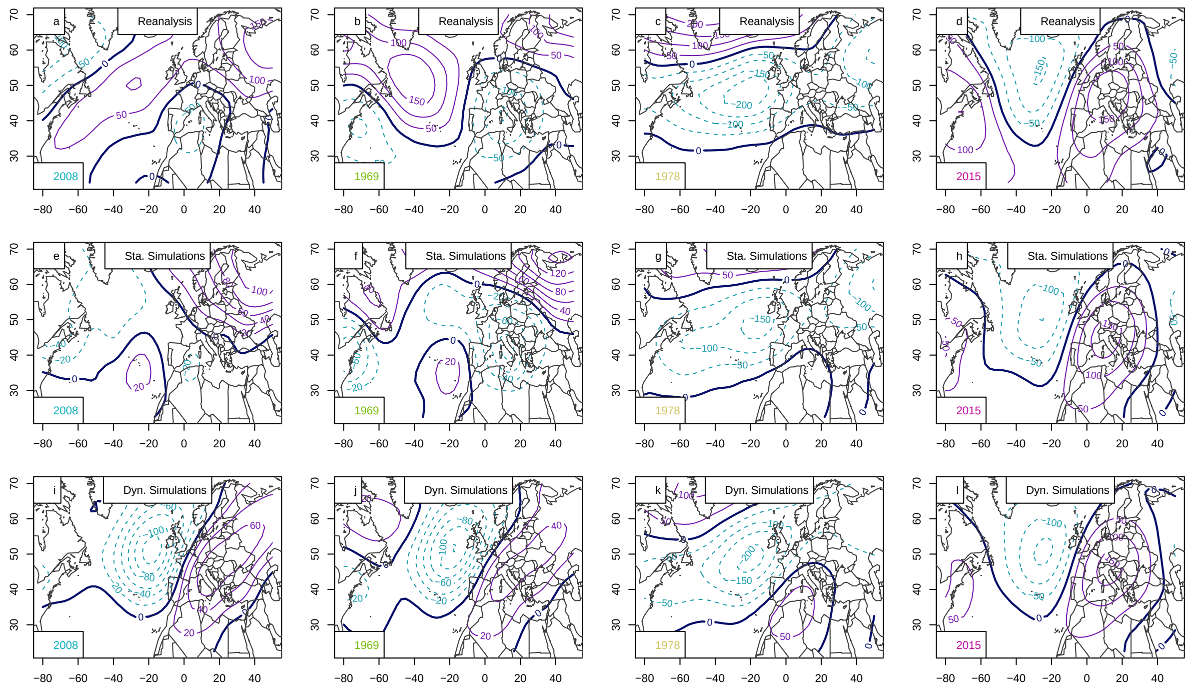

Figure 5. Geopotential height anomaly at 500 mb (Z500) composites for a year with 31 cold days (2008), the coldest (1969), median (1978) and the warmest December (2015). Upper row (panels a–c): mean ~~HGT~~ Z500 from NCEP reanalyses. Center row (panels d-f): Static SWG simulations. Bottom row (panels g-i): Dynamic SWG simulations. Isolines are shown with 100m increments. Positive Z500 anomalies are shown with purple continuous isolines; negative anomalies are shown in cyan dashed lines; the 0 anomaly is shown in thick continuous black line.

5.2 Spring Precipitation

An extremely wet period from April to July 2016 followed the warm December 2015 with an average precipitation of 2.7mm per day and 332mm for the whole period. This is more than the long-term 75th percentile but it is topped by some years including 1983, 1987 and 2012.

Figure 6 shows the daily mean precipitation for April-July periods over 1950-2018. Accumulated April-July precipitation fluctuates around 256mm with a strong year to year variability. Over the observed period no trend is detected.

Simulations from the static weather generator (blue boxplots in Fig. 6) also show a strong inter-annual variability but with significantly larger amounts of precipitation. The average seasonal precipitation for all simulations and all years is around 487mm - 190% of the observed average. Single simulations even reach daily mean precipitation of 6mm for April-July which is three times as high as the ~~maximal~~ observed precipitation in 1983.

April-July periods simulated by the dynamic SWG are even wetter than the simulations of the static SWG with an average seasonal precipitation of 590mm. As expected, the inter-annual variations are smaller ~~as~~ in the dynamic SWG simulations than

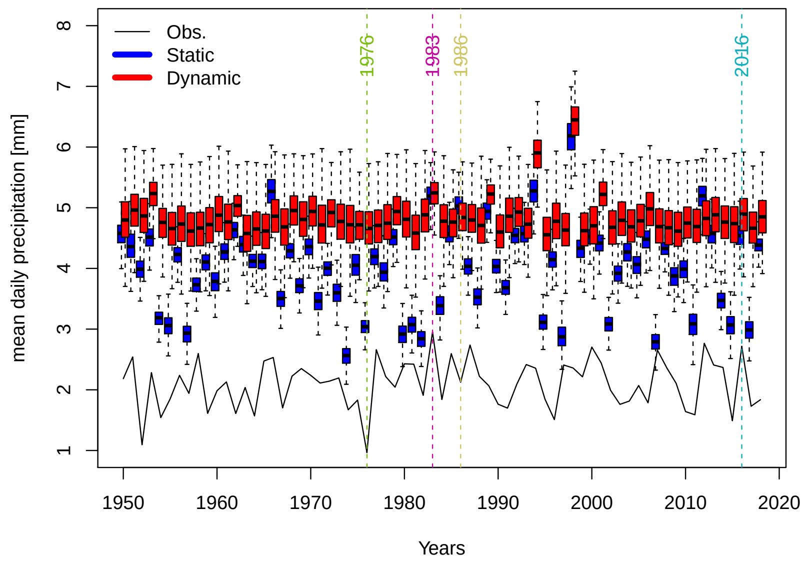

Figure 6. Daily precipitation averages for April-July from 1950 to 2018. ~~To~~ The black line shows E-OBS observations. The boxplots represent the ensemble variability of the simulations of the static (blue) and the dynamic (red) SWG for each year. The boxes of boxplots indicate the median ($q50$), lower ($q25$) and upper ($q75$) quartiles. The upper whiskers indicate $\min[\max(T), 1.5 \times (q75 - q25)]$. The lower whisker has a symmetric formulation. The points are the simulated values that are above or below the defined whiskers. The colored vertical lines indicate the driest April-July period (green), the wettest period (1983), a median period (1986) and 2016.

in the static SWG simulations because the dynamic SWG evolves freely with the starting conditions as only tie to the observed circulation.

We estimate the return periods of our simulated events by fitting a normal distribution to the observed April-July precipitation events. As we average over a quite large region and over four months a normal distribution represents the observations well (even though the analysed variable is precipitation). We find that the 2016 April-July period was a one in 17 years event while the majority of our SWGs simulations are one in ten thousand year events.

In April-July 2016, the atmospheric circulation was characterized by a moderate low pressure anomaly north of France and north of the Azores (Fig. 7a). The North Atlantic Oscillation (NAO) index switched from slightly positive to negative in May and remained negative until the end of June ~~NOAA (2020)~~(NOAA, 2020).

We next analyze the large-scale atmospheric circulation patterns that characterize our SWG simulations by comparing them to a few examples of observed events. Figure 7a-d shows the mean sea level composites of 2016, the driest (1976), a median (1986) and the wettest (1983) April-July periods. The main feature in the median event (Fig. 7c) is a low pressure anomaly north-westward of the British Isles. The wettest event (Fig. 7d) is characterized by a strong dipole over the northern Atlantic with low pressure in the east and high pressure in the west. In the driest event (Fig. 7b) this dipole is reversed and slightly shifted to the east.

For all four events, the static SWG tends to create events with stronger low pressure anomalies over northern France (Fig. 7e-h). Similarly, the simulations from the dynamic SWG all show a strong low pressure anomaly over northern France (Fig. 7i-l). For the dynamic SWG simulations even in 1976, which was the driest April-July period, a low pressure anomaly is simulated for northern France where a high pressure system had been observed. In the static SWG, the high pressure anomaly is relocated to the west leading also to a low pressure anomaly over northern France.

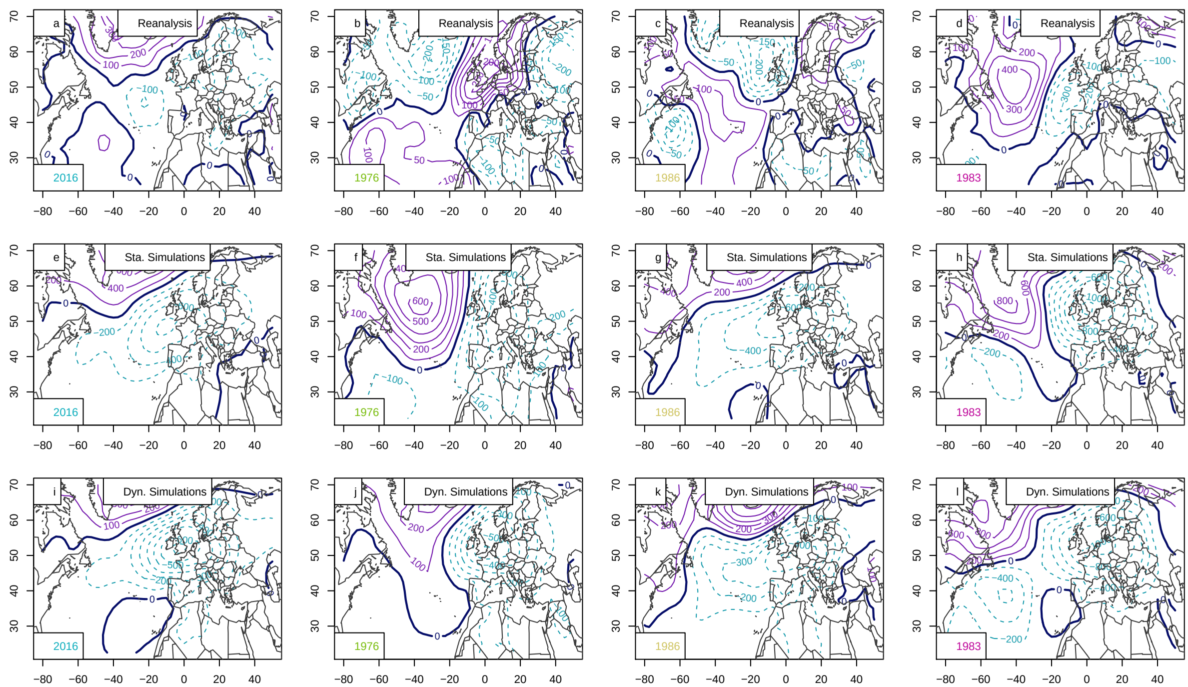

Figure 7. SLP anomaly composites (Pa) for April-July 2016, the driest period, median (1986) and the wettest period (1983). Upper row (panels a–d): mean SLP from NCEP reanalyses. Center row (panels e–h): Static SWG simulations. Bottom row (panels i–l): Dynamic SWG simulations. Isolines are shown with 100Pa increments. Positive SLP anomalies are shown with purple continuous isolines; negative anomalies are shown in cyan dashed lines; the 0 anomaly is shown in thick continuous black line.

Besides a general tendency towards low pressure anomalies over northern France, the 2016 April-July period was characterized by an increased daily pressure variability west of France (compare B1a and B1c). This indicates an enhanced storm track activity downstream of our region of interest and could explain the increased precipitation observed in 2016. In contrast to the persistent anticyclonic anomaly that led to a continuously warm December in 2015, the wet April-July period was favoured by a number of storms passing over northern France.

Our simulations of April-July periods combine 5-day chunks of observed weather into one coherent time series. By using 5-day chunks instead of combining single day observations we constrain our simulations to observed day to day variations that appear to be crucially important for precipitation events. This ensures that in our simulations storms predominantly travel eastwards and that the moisture transport in the simulations is reasonable - at least during these 5 days (see animated gif files in the supplementary files).

Indeed sensitivity tests show that simulations where a new analogue is chosen every day result in significantly higher precipitation with 7mm per day for the dynamic SWG simulations (see Fig. A4). The amount of precipitation steadily decreases with the length of the observed chunks that are assembled by the SWGs (n_{days}). This is to be expected as with longer assembled chunks and fewer analogue choices the simulated weather events resemble more and more the observations. There is an especially strong decrease in simulated precipitation from one to 3 days which suggests, that when analogues are chosen more frequently than every third day potentially unreasonable weather events are created. Note that taking five day windows is a heuristic choice and that window sizes between four and seven days give similar results.

The simulations are by definition sensitive to the weighting on the amount of precipitation α. As shown in figure A5, already with a relatively small weight of 0.1 most dynamic simulations bring more precipitation than what was observed in 2016. This could be due to the length of the simulations: it is rather unlikely that extreme weather endures over 4 months but already with a weak weighting on wet weather, simulations can result in a long lasting consistent wet periods. This increase in precipitation saturates after $\alpha \approx 0.5$ and increasing α further has no effect on the final results.

As for the other free parameters of the SWG, this sensitivity test does not directly justify the choice of the parameter α. It rather gives guidance on the values that would be appropriate choices for our application. In the end the parameter is heuristically chosen considering the trade-off between creating high precipitation events and keeping as much *randomness* as possible in our simulations.

As shown in Fig. A6 the weighting on the calendar day has limited influence on the amount of precipitation in northern France simulated by our SWG's.

For precipitation in northern France the weighting on the calendar day is less relevant as there is no pronounced seasonal cycle in precipitation (see Fig. A6).

Finally, one feature in the simulations of April-July deserves some more attention: For both static and dynamic SWG simulations precipitation is exceptionally high in 1994 and 1998. Although observed precipitation in these years was relatively high, this cannot explain the amount of precipitation in the simulations. One explanation for these outlier years could be a loop in the simulations leading to an excessive repetition of the same (wet) sequence of days. As shown in figure A7 in 1998 one date is indeed repeated 10 times in both the static and dynamic weather generator. In most other years, repetitions of single dates are rare. As our results do not rely on simulations of single years, this feature doesn't affect the overall findings of the study.

These simulations show, that there are many possible April-July periods that would be significantly wetter than what was observed in 2016 and also wetter than the ~~observational record~~ observed record preciptiation (1983).

6 Discussion

In 2016 northern France suffered an unprecedented crop loss that can be related to an abnormally warm December in 2015 and a following wet April-July period in 2016 (Ben-Ari et al., 2016). Here we investigated how extreme these meteorological precursors of the crop loss could be in current climate. Using stochastic weather generators (SWG) we simulate warm Decembers and wet April-July periods independently.

The warm December 2015 resulted in few cold days with temperatures between ~~0–10~~0 and 10°C. Our simulations show that substantially warmer Decembers would be possible. However, in terms of cold days, which is ~~the most~~ a more relevant indicator for wheat phenology in that season (Ben-Ari et al., 2018), December 2015 was already extreme and only few simulations show lower numbers of cold days.

For April-July precipitation, we find that much wetter periods than what was observed in 2016 would be plausible. The simulated events bring more than twice as much precipitation than in 2016.

~~Although the mechanism that led to the crop failure in 2016 is not fully understood yet, our analysis suggests~~ If crop yields responds to the number of cold days in winter and to the precipitation rate in spring as shown in Ben-Ari et al. 2018, then we have shown here that in current climate ~~more extreme meteorological conditions and therefore also more extreme crop losses~~ an even worse crop loss event would be possible. Especially the April-July period could be significantly wetter than what was observed in 2016.

We used stochastic weather generators to simulate extreme but plausible weather events. While the method is established for summer heat waves ~~Yiou and Jézéquel (2019)~~ (Yiou and Jézéquel, 2019) the weather events we studied here brought new challenges: Although ~~from~~ the circulation pattern of the warm December 2015 was similar to a summer heat wave with an anticyclonic pattern over France, special care was required to assure that our simulated events are actually realisations of winter weather. Here we assured for this by strongly weighting the calendar date when selecting analogues.

The wet April-July 2016 period was characterized by a series of passing storms that brought considerable amounts of precipitation. The main feature of this ~~weather event~~ wet spring season was therefore not persistence and simulating plausible day-to-day variations with SWGs was a major challenge. ~~By reassembling~~ SWGs that select a new analogue every day tend to simulate persistent rainfall events over spring, with little day-to-day variation.

As a first attempt to simulate plausible long lasting wet periods, we propose to reassemble 5-day ~~chunks~~ windows of observed weather instead of single days~~we succeeded in simulating reasonable long lasting wet periods.~~. This ensures, that low- and high pressure systems predominantly travel eastward at a speed which is tightly linked to observations. An alternative approach could be to switch trajectories on dry days instead of switching after a fixed number of days. This would additionally avoid changing trajectories during precipitation events.

Evaluating the plausibility of our simulations remains a challenge: although sensitivity tests and an analysis of the simulated circulation patterns reveal a robust and well interpretable behaviour of the SWGs, further tests would be required to assess whether all simulated events could really happen in our climate. It could for instance be interesting to analyze the simulated wet April-July periods with respect to more climate variables ~~as relative humidityin the atmosphere~~ (e.g. relative humidity) to evaluate whether the water transport is physically plausible throughout the simulated period.

To further evaluate the plausibility of our simulations one could also compare them to extreme events simulated by large ensemble climate modelling experiments. ~~A~~ In a study using a near term climate prediction model, Thompson et al. (2017) found that for England there is a considerable chance of unprecedented winter rainfall. Replicating a similar study for northern

France spring precipitation would not only provide an alternative estimate of extreme spring precipitation but would also allow to further evaluate the circulation features of our weather simulations.

Finally, ~~using crop models as the model developed by Ben-Ari et al.~~ our simulated extremes could be used ~~to estimate how large the crop losses would be for our simulated weather extremes~~as input of the regression-based yield model of Ben-Ari et al. (2018). These results should however be interpreted cautiously as our simulated weather extremes lie outside of the observed range and thereby also the range on which the ~~crop models were trained~~yield model was trained. They could also be used in process based crop models, as a worst-case meteorological scenario.

7 Conclusions

This paper is a proof of concept of importance sampling for the simulation of a compound event (warm autumn-winter and wet spring) that would have an impact on crop yield. It relies on a data-resampling approach to maximize temperature and precipitation during extended periods of time.

The simulations are based on the a priori knowledge (from expertise on ~~phenology~~crop failures in northern France) that warm autumns-winters followed by wet springs have detrimental effects on crops.

The first application of SWGs to warm winter periods and wet springs is an important advance in this research field. It also shows that with only a few adaptations SWGs can be applied to new weather phenomena, highlighting the merits of the method. Moreover, the SWG parameters can be adapted to other types of crops (with other phenological parameters and key dates).

This approach is rather flexible and could be adapted to simulate compound extremes ~~in scenario climate model simulations, in order to evaluate~~ using climate model outputs based on different scenarios of climate change. This could lead to a first evaluation of the impact of climate change on worst case scenarios of crop yields. This ~~is left to future investigations.~~ type of analysis has some limitations, related to the uncertainty of models and scenarios, and it fails to take into account non-climatic drivers of crop yields such as pests, supply chain, or economical concerns. We however believe it could be useful to estimate what could be plausible in terms of purely meteorological events, in a changing climate.

8 Acknowledgements

We acknowledge the E-OBS dataset from the EU-FP6 project UERRA (http://www.uerra.eu) and the Copernicus Cli-340mate Change Service, and the data providers in the ECAD project (https://www.ecad.eu). We also acknowledge the NCEP Reanalysis data provided by the NOAA/OAR/ESRL. P.P. acknowledges support by the German Federal Ministry of Education and Research (01LN1711A). E.V. acknowledges funding from the Swiss National Science Foundation (Doc.Mobility Grant 188229). This work was initiated during a school supported by the DAMOCLES COST action (grant No. CA17109).

9 Code availability

All R scripts used for the analysis and the production of figures are openly available under https://doi.org/10.5281/zenodo.3859976

Appendix A: Sensitivity tests

A1 December temperature

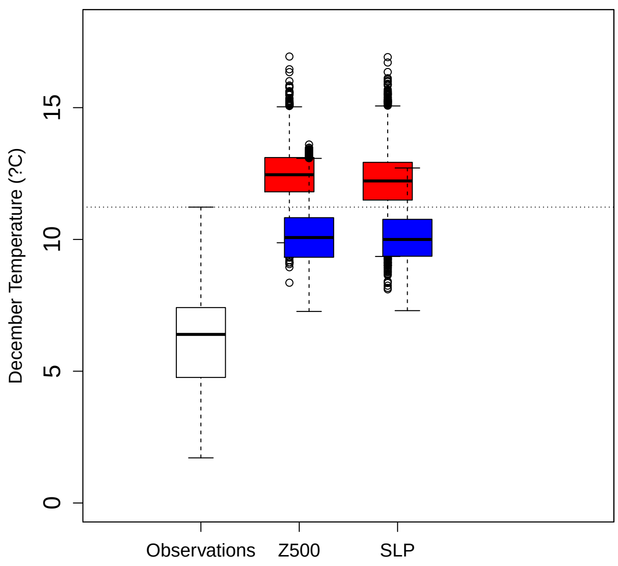

Figure A1. Distribution of the daily maximum temperature in December averaged in December in observations (white) and in simulations computed by the static (blue) and dynamic (red) generators, using circulation analogues computed using the SLP or Z500. The horizontal dotted line corresponds to the daily maximum temperature observed in December 2015. The boxes of boxplots indicate the median (q50), lower (q25) and upper (q75) quantiles. The upper whiskers indicate min[max(T),1.5×(q75-q25)]. The lower whisker has a symmetric formulation. The points are the simulated values that are above or below the defined whiskers.

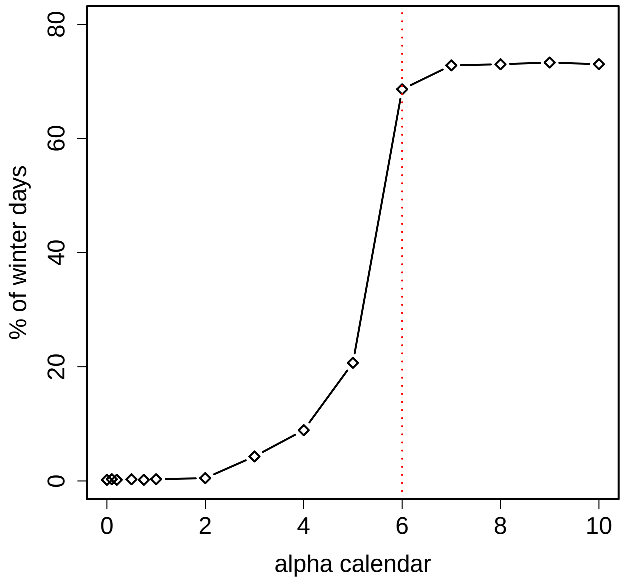

Figure A2. Percentage of days sampled between November and February by the dynamic generator when running 100 simulations of December temperatures, as a function of the parameter α_{cal}. The red dotted line is for $\alpha_{cal} = 6$ (which is the ~~parameter we choose~~value used in the analysis).

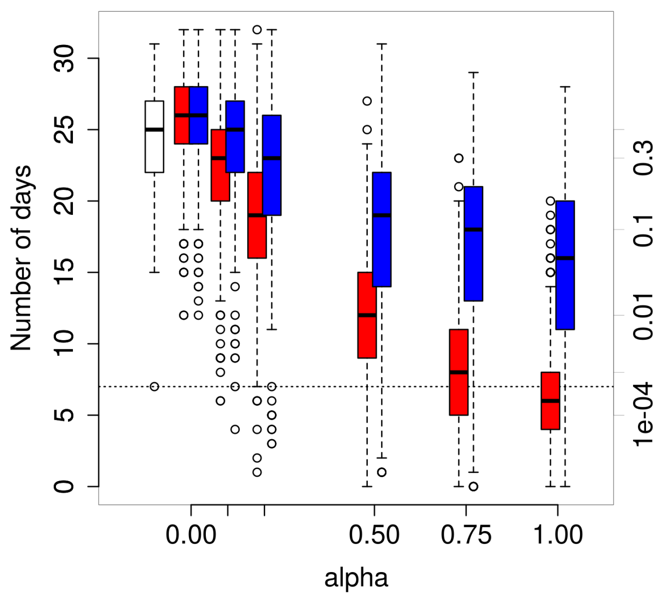

Figure A3. Distribution of the number of December days with maximal temperatures between 0 and 10°C in observations (white) and in simulations computed by the static (blue) and dynamic (red) generators as a function of α. The axis on the right indicates the probability of occurrence, assuming a Beta-Binomial distribution of the number of winter days with parameters estimated from white boxplot. The horizontal dotted line corresponds to the observed number of days in December 2015. The boxes of boxplots indicate the median (q50), lower (q25) and upper (q75) quartiles. The upper whiskers indicate min[max(T),1.5×(q75-q25)]. The lower whisker has a symmetric formulation. The points are the simulated values that are above or below the defined whiskers.

A2 Spring precipitation

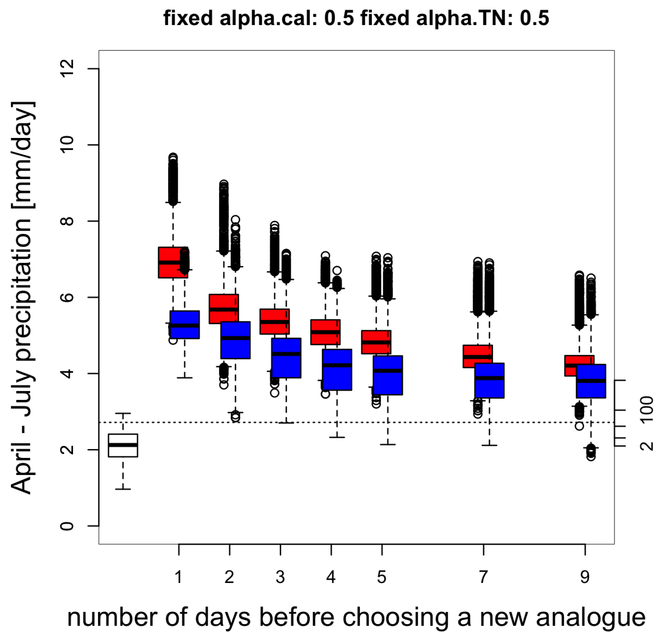

Figure A4. Distribution of April-July daily precipitation in observations (white) and in simulations computed by the static (blue) and dynamic (red) generators as a function of the number of days before selecting a new analogue n_{days}. The axis on the right indicates the probability of occurrence, assuming a normal distribution of daily precipitation with parameters estimated from white boxplot. The horizontal dotted line corresponds to the observed daily precipitation in April-July 2016. The boxes of boxplots indicate the median (q50), lower (q25) and upper (q75) quartiles. The upper whiskers indicate min[max(T),1.5×(q75-q25)]. The lower whisker has a symmetric formulation. The points are the simulated values that are above or below the defined whiskers.

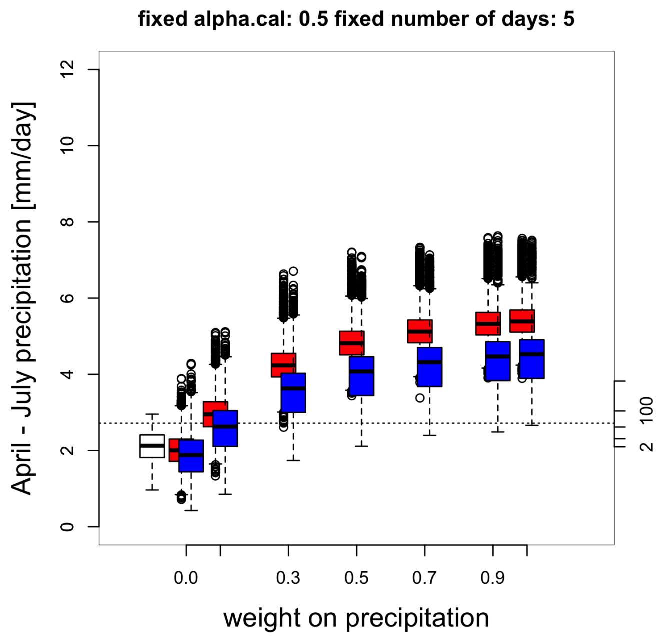

Figure A5. Distribution of April-July daily precipitation in observations (white) and in simulations computed by the static (blue) and dynamic (red) generators as a function of α. The axis on the right indicates the probability of occurrence, assuming a normal distribution of daily precipitation with parameters estimated from white boxplot. The horizontal dotted line corresponds to the observed daily precipitation in April-July 2016. The boxes of boxplots indicate the median (q50), lower (q25) and upper (q75) quartiles. The upper whiskers indicate min[max(T),1.5×(q75-q25)]. The lower whisker has a symmetric formulation. The points are the simulated values that are above or below the defined whiskers.

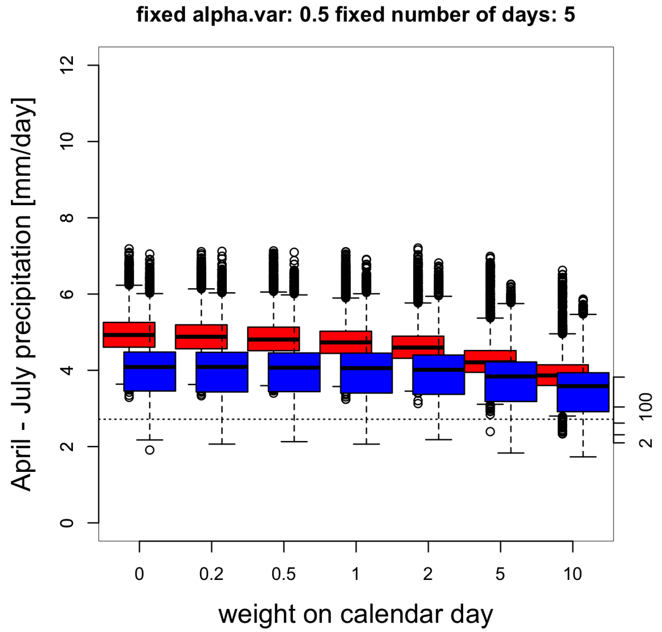

Figure A6. Distribution of April-July daily precipitation in observations (white) and in simulations computed by the static (blue) and dynamic (red) generators as a function of α_{cal}. The axis on the right indicates the probability of occurrence, assuming a normal distribution of daily precipitation with parameters estimated from white boxplot. The horizontal dotted line corresponds to the observed daily precipitation in April-July 2016. The boxes of boxplots indicate the median (q50), lower (q25) and upper (q75) quartiles. The upper whiskers indicate min[max(T),1.5×(q75-q25)]. The lower whisker has a symmetric formulation. The points are the simulated values that are above or below the defined whiskers.

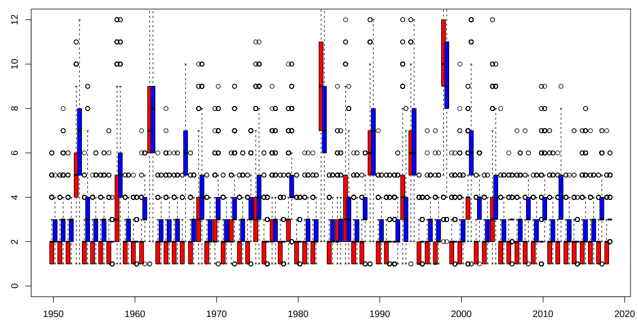

Figure A7. Maximal number of times a single date is repeated for each simulated year. The boxplots indicate the range of this maximal repetition number for the 1000 simulations for simulations of the static (blue) and dynamic (red) stochastic weather generator. The boxes of boxplots indicate the median (q50), lower (q25) and upper (q75) quartiles. The upper whiskers indicate min[max(T),1.5×(q75-q25)]. The lower whisker has a symmetric formulation. The points are the simulated values that are above or below the defined whiskers.

Appendix B: Circulation details

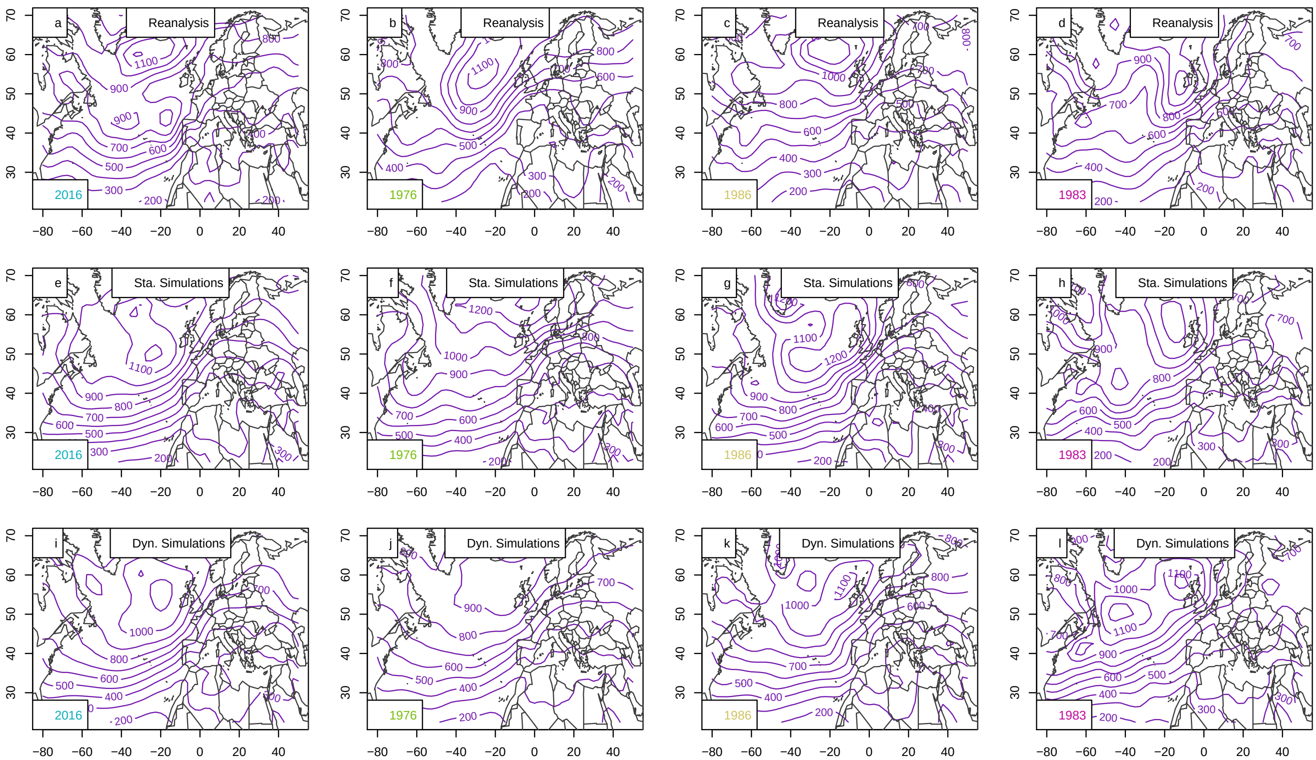

Figure B1. Standard deviation of daily SLP anomalies (Pa) for April-July 2016, the driest period, median (1986) and 2018. Upper row (panels a–d): SLP from NCEP reanalyses. Center row (panels e–h): Static SWG simulations. Bottom row (panels i–l): Dynamic SWG simulations. For the SWG simulations the average of all 1000 runs for the given year are presented.

Author contributions. P.Y. and A.J. conceived the study. N.LG., J.L, I.M., E.V. and P.P. did the analysis and created all figures. P.P. wrote the manuscript with contributions from all authors.

Competing interests. The authors declare no competing interests.

References

ARVALIS: Rendements catastrophiques du blé en 2016 : la pluie, seule responsable?, https://www.semencesdefrance.com/actualite-semences-de-france/rendements-catastrophiques-ble-2016-pluie-seule-responsable/, accessed: 2020-10-23, 2016.

Ben-Ari, T., Adrian, J., Klein, T., Calanca, P., Van der Velde, M., and Makowski, D.: Identifying indicators for extreme wheat and maize yield losses, Agricultural and Forest Meteorology, 220, 130–140, 2016.

Ben-Ari, T., Boé, J., Ciais, P., Lecerf, R., Van der Velde, M., and Makowski, D.: Causes and implications of the unforeseen 2016 extreme yield loss in the breadbasket of France, Nature Communications, 9, 1627, https://doi.org/10.1038/s41467-018-04087-x, http://dx.doi.org/10.1038/s41467-018-04087-xhttp://www.nature.com/articles/s41467-018-04087-x, 2018.

Cassou, C., Terray, L., and Phillips, A. S.: Tropical Atlantic influence on European heat waves, Journal of Climate, 18, 2805–2811, 2005.

Cooley, D.: Extreme value analysis and the study of climate change, Climatic change, 97, 77, 2009.

de Bruijn, K. M., Lips, N., Gersonius, B., and Middelkoop, H.: The storyline approach: a new way to analyse and improve flood event management, Natural Hazards, 81, 99–121, https://doi.org/10.1007/s11069-015-2074-2, http://link.springer.com/10.1007/s11069-015-2074-2, 2016.

FAO, F.: Agricultural statistics database, Rome: Wold Agricultural. Information Center. Disponível em< http://faostat. fao. org/site/567/DesktopDefault. aspx, 2013.

Haylock, M. R., Hofstra, N., Tank, A. M. G. K., Klok, E. J., Jones, P. D., and New, M.: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006, J. Geophys. Res. - Atmospheres, 113, doi:10.1029/2008JD010 201, <GotoISI>://000260598000009, 2008.

Hazeleger, W., Van Den Hurk, B. J., Min, E., Van Oldenborgh, G. J., Petersen, A. C., Stainforth, D. A., Vasileiadou, E., and Smith, L. A.: Tales of future weather, Nature Climate Change, 5, 107–113, https://doi.org/10.1038/nclimate2450, 2015.

Jaworski, P., Durante, F., Hardle, W. K., and Rychlik, T.: Copula theory and its applications, vol. 198, Springer, 2010.

Jézéquel, A., Cattiaux, J., Naveau, P., Radanovics, S., Ribes, A., Vautard, R., Vrac, M., and Yiou, P.: Trends of atmospheric circulation during singular hot days in Europe, Environmental Research Letters, 13, 054 007, https://iopscience.iop.org/article/10.1088/1748-9326/aab5da/pdf, 2018.

Kistler, R., Kalnay, E., Collins, W., Saha, S., White, G., Woollen, J., Chelliah, M., Ebisuzaki, W., Kanamitsu, M., Kousky, V., van den Dool, H., Jenne, R., and Fiorino, M.: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation, Bulletin of the American Meteorological Society, 82, 247–267, <GotoISI>://000166742900003, 2001.

Leonard, M., Westra, S., Phatak, A., Lambert, M., Hurk, B. v. d., McInnes, K., Risbey, J., Schuster, S., Jakob, D., and Stafford-Smith, M.: A compound event framework for understanding extreme impacts, Wiley Interdisciplinary Reviews: Climate Change, 5, 113–128, https://doi.org/10.1002/wcc.252, 2014.

MacDonald, R. B. and Hall, F. G.: Global crop forecasting, Science, 208, 670–679, 1980.

Massey, N., Jones, R., Otto, F., Aina, T., Wilson, S., Murphy, J., Hassell, D., Yamazaki, Y., and Allen, M.: weather@ home—development and validation of a very large ensemble modelling system for probabilistic event attribution, Quarterly Journal of the Royal Meteorological Society, 141, 1528–1545, 2015a.

Massey, N., Jones, R., Otto, F. E. L., Aina, T., Wilson, S., Murphy, J. M., Hassell, D., Yamazaki, Y. H., and Allen, M. R.: weather@home—development and validation of a very large ensemble modelling system for probabilistic event attribution, Quat. J. Roy. Met. Soc., 141, 1528–1545, https://doi.org/10.1002/qj.2455, 2015b.

Müller, C., Elliott, J., Kelly, D., Arneth, A., Balkovic, J., Ciais, P., Deryng, D., Folberth, C., Hoek, S., Izaurralde, R. C., et al.: The Global Gridded Crop Model Intercomparison phase 1 simulation dataset, Scientific data, 6, 1–22, 2019.

NOAA: North Atlantic Oscillation, https://www.cpc.ncep.noaa.gov/products/precip/CWlink/pna/nao.shtml, accessed: 2020-01-10, 2020.

OEC: The Observatory of Economic Complexity, https://oec.world/en/, accessed: 2020-04-23, 2020.

Peixoto, J. P. and Oort, A. H.: Physics of climate, 1992.

Ragone, F., Wouters, J., and Bouchet, F.: Computation of extreme heat waves in climate models using a large deviation algorithm, Proceedings of the National Academy of Sciences, p. 201712645, https://www.pnas.org/content/pnas/115/1/24.full.pdf, 2017.

Shepherd, T. G.: Storyline approach to the construction of regional climate change information, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 475, https://doi.org/10.1098/rspa.2019.0013, 2019.

Shepherd, T. G., Boyd, E., Calel, R. A., Chapman, S. C., Dessai, S., Dima-West, I. M., Fowler, H. J., James, R., Maraun, D., Martius, O., Senior, C. A., Sobel, A. H., Stainforth, D. A., Tett, S. F., Trenberth, K. E., van den Hurk, B. J., Watkins, N. W., Wilby, R. L., and Zenghelis, D. A.: Storylines: an alternative approach to representing uncertainty in physical aspects of climate change, Climatic Change, 151, 555–571, https://doi.org/10.1007/s10584-018-2317-9, 2018.

Sturaro, G.: A closer look at the climatological discontinuities present in the NCEP/NCAR reanalysis temperature due to the introduction of satellite data, Climate dynamics, 21, 309–316, https://doi.org/10.1007/s00382-003-0334-4, 2003.

Thompson, V., Dunstone, N. J., Scaife, A. A., Smith, D. M., Slingo, J. M., Brown, S., and Belcher, S. E.: High risk of unprecedented UK rainfall in the current climate, Nature Communications, 8, 1–6, https://doi.org/10.1038/s41467-017-00275-3, https://www.nature.com/articles/s41467-017-00275-3, number: 1 Publisher: Nature Publishing Group, 2017.

Yiou, P.: AnaWEGE: a weather generator based on analogues of atmospheric circulation, Geoscientific Model Development, 7, https://doi.org/10.5194/gmd-7-531-2014, 2014.

Yiou, P. and Jézéquel, A.: Simulation of Extreme Heatwaves with Empirical Importance Sampling, Geosci. Model Dev. Discuss., 2019, 1–26, https://doi.org/10.5194/gmd-2019-164, https://www.geosci-model-dev-discuss.net/gmd-2019-164/, 2019.

Yiou, P. and Jézéquel, A.: Simulation of extreme heat waves with empirical importance sampling, Geosci. Model Dev., 13, 763–781, https://doi.org/10.5194/gmd-13-763-2020, https://www.geosci-model-dev.net/13/763/2020/, 2020.

Yiou, P. and Nogaj, M.: Extreme climatic events and weather regimes over the North Atlantic: When and where?, Geophys. Res. Lett., 31, https://doi.org/10.1029/2003GL019119, 2004.

Zscheischler, J., Westra, S., Hurk, B., Seneviratne, S., Ward, P., Pitman, A., AghaKouchak, A., Bresch, D., Leonard, M., Wahl, T., and Zhang, X.: Future climate risk from compound events, Nature Climate Change, https://doi.org/10.1038/s41558-018-0156-3, 2018.

Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R. M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., Mahecha, M. D., Maraun, D., Ramos, A. M., Ridder, N. N., Thiery, W., and Vignotto, E.: A typology of compound weather and climate events, Nature Reviews Earth & Environment, 1, 333–347, https://doi.org/10.1038/s43017-020-0060-z, http://www.nature.com/articles/s43017-020-0060-z, 2020.