Rare event simulation of extreme European winter rainfall in an intermediate complexity climate model

Jeroen Wouters[1], **Reinhard K.H. Schiemann**[2], **Len C. Shaffrey**[2]

[1]Department of Mathematics & Statistics
University of Reading
Reading, UK
[2]National Centre for Atmospheric Science
Department of Meteorology
University of Reading
Reading, UK

Key Points:

- A rare event simulation algorithm is applied to a climate model.
- Using the algorithm a large sample of model realizations with high winter rainfall over Europe is generated.
- This allows to estimate probabilities of high threshold exceedances at a significantly lower numerical effort, compared to conventional brute-force simulation.

Corresponding author: Jeroen Wouters, `j.wouters@reading.ac.uk`

Abstract

We test the application of a rare event simulation algorithm to accelerate the sampling of extreme winter rainfall over Europe in a climate model. The genealogical particle analysis algorithm, an ensemble method that interrupts the simulation at intermediate times to clone realizations in which an extreme event is developing, is applied to the intermediate complexity general circulation model PlaSim. We show that the algorithm strongly reduces the numerical effort required to estimate probabilities of extremes, demonstrating the potential of rare event simulation of seasonal precipitation extremes.

Plain Language Summary

Rare events, such as winters with a lot of rainfall, can have a high impact. For example, high rainfall can cause floods, risking lives and damaging property. To plan for such events, we would like to know what the probability is to experience a lot of rainfall in any given year. One way to make an estimate of these probabilities is to run climate models and look at the proportion of winters having rainfall above a certain threshold. This method can be very costly and time-consuming, as climate models require powerful computers and we need to keep them running for a long time until we see at least a few winters with very high rainfall. In this article, we have applied algorithms to speed up such model simulations. Instead of keeping the model running for a long time, we run a large number of copies at the same time. At regular intervals, we check which simulations are producing a lot of rain. These we copy, while others are terminated. Even though we have made high rainfall more likely than it should be, we can still estimate what the correct probabilities of high rainfall are. We show that this method can save time spent running our computer model, while still giving accurate results.

1 Introduction

The impact of anthropogenic climate change is felt most acutely during extreme weather and climate events, such as windstorms, heavy precipitation, flooding, heatwaves, and drought. Applied research in extremes addresses the questions of (i) quantifying the probability of extreme event occurrence and detecting changes over time, (ii) attributing such changes to natural variability as well as natural and anthropogenic forcing, including event attribution for individual observed extremes, (iii) predicting occurrence probabilities at lead times of days to decades, and (iv) deriving centennial and longer climate-change projections (Seneviratne et al., 2021). Designing and interpreting these kinds of studies is underpinned by, and advances, understanding of the physical processes involved.

All of these applications are also heavily reliant on computer simulation. For example, large ensembles of global climate model simulations are driven with transient historical and scenario forcings to detect forced changes in the historical period and in projected future climates (Deser et al., 2020), and systematic variation of the natural and anthropogenic forcings used in model experiments allows for the attribution of trends in extremes to these forcings (Wan et al., 2019). Similarly, event attribution contrasts the occurrence of specific historically observed extremes in ensembles of all-forcings simulations and counterfactual simulations in which changing anthropogenic forcings are disregarded (Otto et al., 2018). In forecast applications, ensembles of simulations are initialised with observation-based estimates of the state of the atmosphere, ocean, cryosphere, and land surface with one aim to quantify changes of extreme event occurrence probability in the forecast period compared to a climatological baseline (Davini et al., 2021).

Conducting these model experiments is costly in terms of human and computer time, data analysis and storage, and energy consumption. The nature and rareness of the investigated events imposes competing demands on a modelling setup that is fit-for-purpose in terms of model complexity, the length and number of simulations (ensembles members), and resolution. A case in point are seasonal precipitation extremes in the midlatitudes (Schaller et al., 2016; M. D. K. Priestley et al., 2017): Due to the length of the event itself many simulated years are required to sample extreme seasons. At the same time, model resolution benefits the representation of storm tracks, storm structure, the water cycle and extreme precipitation, and a global coupled modelling setup allows for the representation of teleconnections with the tropics or the Arctic (Zappa et al., 2013; M. D. Priestley & Catto, 2022; Vannière et al., 2019; Schiemann et al., 2018; Roberts et al., 2018). Limited computing resource may require uncomfortable compromises between these different demands.

Efficient simulation strategies are required to resolve such dilemmas. Rare event simulation (RES) algorithms, as applied in this article, provide a way to generate many more samples of an extreme event. Originally developed in statistical physics, RES algorithms have found applications in, e.g., queuing networks, biochemistry and aircraft design. In Earth system modelling, rare event algorithms have previously been successfully applied to atmospheric models, to over-sample heat waves (Ragone et al., 2017; Ragone & Bouchet, 2021) and extremely intense tropical cyclones (Webber et al., 2019).

The aim of this study is to demonstrate the potential of RES for simulating seasonal precipitation extremes. We implement rare event simulation of winter rainfall in an intermediate complexity climate model, as a test bed for further studies on higher resolution models. We will estimate probabilities of threshold exceedance of winter precipiation far into the tail of the distribution and quantify how our RES approach has significantly reduced the numerical effort required.

We will use an ensemble based genealogical algorithm similar to those used in (Ragone et al., 2017; Ragone & Bouchet, 2021; Webber et al., 2019), which are, in turn, variants of the algorithm described in (Del Moral & Garnier, 2005). In this family of algorithms, an ensemble is run until an intermediate stopping time, at which time an objective function is applied to each realization. Realizations with a higher value of the objective function are deemed to be more promising and are cloned. Realizations with low values may be terminated, hence focusing numerical effort on the extreme event.

In Section 2, we will describe the atmospheric model used, introduce rare event simulations methods in general and the setup used in this study in particular. In Section 3, we will then present the outcome of the experiment, demonstrating that the method used can generate a much larger sample of extreme events, reducing strongly the total amount of simulation time needed to obtain the same accuracy.

2 Methodology

2.1 The general circulation model

The model used in this study is the open-source Planet Simulator (PlaSim) model (Fraedrich et al., 2005), which uses the Portable University Model of the Atmosphere (PUMA) as its dynamical core. PlaSim gives a reasonable representation of atmospheric dynamics and of their interactions with the land surface and with the mixed layer of the ocean; it includes parameterizations of radiative transfers, diagnostic cloud cover, large-scale precipitation, convective precipitation, and dry convective adjustment.

The model is used in a T42 horizontal resolution with 10 vertical layers, fixed default model parameters and climatological SST and sea ice extent with a seasonal cycle. The time step is 20 minutes.

The model climate has been shown to represent basic features of the general circulation, including the North Atlantic storm track (Andres & Tarasov, 2019) (see also Fig. 6).

PlaSim can be considered an intermediate complexity model, which makes it suitable for explorations of novel computational methods such as ours, where computational complexity needs to be balanced with model realism. Previous applications of PlaSim include parameter tuning (Lyu et al., 2018), thermodynamic analysis (Fraedrich & Lunkeit, 2008; Lucarini et al., 2010), climate multistability (Fraedrich, 2012; Margazoglou et al., 2021; Boschi et al., 2013) and linear response theory (Ragone et al., 2016; Lucarini et al., 2017).

2.2 Rare event simulation

2.2.0.1 Importance sampling The fundamental idea behind many rare event simulation methods is that of importance sampling (IS). Assume we wish to estimate a small probability $\gamma = \mathbb{P}\{X \in A\}$ for a random variable X. The set $A \subset \mathbb{R}$ could be, for example, those values exceeding a high threshold $A = (\theta, +\infty)$, or a small interval of width $\Delta\theta$ around a target value $A = (\theta - \Delta\theta/2, \theta + \Delta\theta/2)$. The random variable X could be a physical observable, such as temperature, wind intensity or precipitation, possibly geographically and temporally averaged.

A method to estimate γ is to generate N independent samples X_i of X and use the brute force Monte Carlo estimator

$$\hat{\gamma}_A = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_A(X_i)\,, \tag{1}$$

where $\mathbb{1}_A$ is the indicator function on the set A ($\mathbb{1}_A(x) = 1$ for $x \in A$ and 0 otherwise). The problem with this approach is that the relative error scales as $1/\sqrt{N\gamma_A}$ for small γ_A. Therefore, for rare events with $\gamma_A \ll 1$ we would need a very large number of samples N to keep the relative error under control.

A solution is to sample a different random variable $\tilde{X}$ whose distribution is linked to that of X in a known relation. Let's assume that both X and $\tilde{X}$ have a continuous distribution with probability density $\rho(x)$ and $\tilde{\rho}(x)$ respectively. By generating independent samples $\tilde{X}_i$ from $\tilde{X}$ the importance sampling estimator

$$\tilde{\gamma}_A = \frac{1}{N} \sum_{i=1}^{N} \frac{\rho(\tilde{X}_i)}{\tilde{\rho}(\tilde{X}_i)} \mathbb{1}_A(\tilde{X}_i) = \frac{1}{N} \sum_{i=1}^{N} L(\tilde{X}_i) \mathbb{1}_A(\tilde{X}_i) \tag{2}$$

can be implemented, where $L(x) := \rho(x)/\tilde{\rho}(x)$ whenever $\tilde{\rho}(x) > 0$ and zero otherwise. It is important to note here that the densities ρ and $\tilde{\rho}$ do not need to be known. Only the proportionality factor $L(x)$ needs to be known, which, as we will see below, is the case in the algorithm that we use.

The IS estimator is still unbiased and by choosing a density $\tilde{\rho}$ that makes the rare event A more common, the relative error of this estimator can be greatly reduced in comparison to that of the brute force Monte Carlo estimator. For more details on importance sampling, see, for example, (Bucklew, 2004).

2.2.0.2 Rare event simulation algorithm The aim of rare event simulation algorithms is to perform the change of measure that is necessary to implement importance sampling. The random variable X is now a random process in time and the

probability distributions ρ and $\tilde{\rho}$ should be thought of as probability distributions on possible paths of this process.

When running a brute force simulation of a dynamical process, like the Earth's climate system, the model spends most of the time simulating a completely normal state of the climate, perhaps 99.9% of the time when considering a 1/1000y event. Instead, we can run an ensemble simulation with a number of parallel model runs, and at intermediate times interrupt the calculation to perform a selection procedure. This way, some paths will become more likely than others, performing the above mentioned change of probability measure. Since we can keep track of how much more likely the event has become due to the algorithm, we can find the proportionality factor L, necessary to construct an unbiased estimator from the manipulated ensemble.

In the class of algorithms known as *genealogical particle analysis* (GPA) (Del Moral & Garnier, 2005), trajectories that are not developing a rare event, in our application those with little precipitation, have a higher likelihood of being terminated. The most promising trajectories are, on the contrary, cloned, with more clones assigned to realizations with more precipitation. The algorithm is presented in Appendix B.

The central idea behind GPA is the following: most of the time the different realizations follow a free evolution (step 2a in Appendix B). At intermediate selection times t_k, the ensemble is manipulated to make the rare event of interest more common. At step 2b a weight is calculated, which determines how many clones will be created in step 2c. The number of clones can be zero, meaning this realisation will be terminated. Note that no changes to the model code are necessary to implement GPA, which is an important characteristic when working with complex numerical codes.

The selection steps 2b and 2c are where the change of measure is performed. As in the standard importance sampling estimator (2), the change of measure needs to be taken into account in the estimation step 4. The probability of each realisation has been changed at each selection step by a factor $\bar{W}_k^{(i)}$, so the overall probability has been changed by the product $L[\xi_{i,t_n}] = (\prod_{k=1}^{n-1} \bar{W}_k[\xi_i])^{-1}$. This, therefore, needs to be divided out to make the estimation unbiased.

For extremes of time-integrated observables $\int_{T_0}^{T_1} V(x(\tau))\mathrm{d}\tau$, such as the time-integrated precipitation, an exponential weight function

$$W_k[x] = \exp\left(C \int_{t_{k-1}}^{t_k} V(x(\tau))\mathrm{d}\tau\right) \tag{3}$$

can be used, favouring realizations with a sustained high value of the integral of V. The control parameter C determines the strength of the selection procedure, allowing to focus on more or less rare events.

Figure 1 illustrates the algorithm. As the algorithm doesn't require any modification of the climate model, only killing and cloning of realizations, it can be applied to complex climate models.

2.2.0.3 The RES configuration used The random variable we are interested in is

$$X = \overline{p} := \int_{(T_0,T_1)} \int_{\Omega} p(x(t), r)\, \mathrm{d}r\, \mathrm{d}t$$

where $x(t)$ is the state of the model at time t, $p(x(t), r)$ is the precipitation rate at geographical position r at time t, $\Omega \subset S^2$ is the geographical region of interest and T_0 and T_1 are the initial and final times, respectively. Here Ω is taken to be the region between -20E and 50W and 30N and 70N. The temporal integration is performed over December, January and February, i.e. between $T_0 = 1$ Dec 00:00 and $T_1 = 1$ March 00:00.

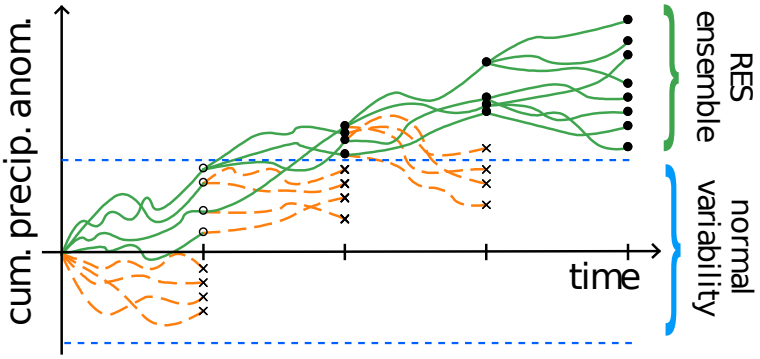

Figure 1. A graphical representation of the cumulative precipitation anomalies over time in an ensemble under the RES algorithm. Trajectories not developing an extreme cumulative precipitation event have a high probability of termination (orange dashed lines). At the final time extreme cumulative precipitation has become common in the ensemble.

Cloning is done every five days, i.e. $t_k - t_{k-1} = 5$ days. The initial distribution is sampled $M = 250$ times on 1 Dec from a 500 year control simulation. The control simulation is a stationary run (with seasonal cycle), using the default parameter settings of the PlaSim model.

For the weight function W_k we use the same observable as the one for which we are estimating rare events, i.e. $V(x) = \int_\Omega p(x,r)\mathrm{d}r$ and

$$W_k[x] \quad = \quad \exp\left(C \int_{t_{k-1}}^{t_k} \int_\Omega p(x(t),r)\mathrm{d}r\mathrm{d}\tau\right) .$$

The parameter C was chosen based by fitting a Gamma distribution to a histogram of $\int_{T_0}^{T_1} \int_\Omega p(x(t),r)\mathrm{d}r\mathrm{d}\tau$ in the control run and picking the value of C that would result in a change in the mean of the RES ensemble of 3 standard deviations (see Appendix Appendix A). The assumption of a Gamma distribution is never used in the algorithm, it is only used to obtain a reasonable value of C to target a given range of extreme events.

3 Results

3.1 Effectiveness of the RES algorithm

We first verify that the RES is working effectively. Figure 2 (a) shows the cumulative rainfall anomalies generated in the RES versus the brute force control run. We see that, as expected, over time the realisations with the highest precipitation have been selected, tilting the distribution at the final time towards extreme events. Figure 2 (b) shows histograms of winter precipitation generated from 490 years of the 500 year control run (discarding the initial 10 years) and from rare event simulation. It is evident that the RES has sampled the tail of the winter European rainfall distribution. At the final time what was a rare event in the control run has become common in the RES. The mean rainfall level in the RES ensemble lies at 0.220m. Only 13 winters in the 490 year control run have winter rainfall exceeding this level.

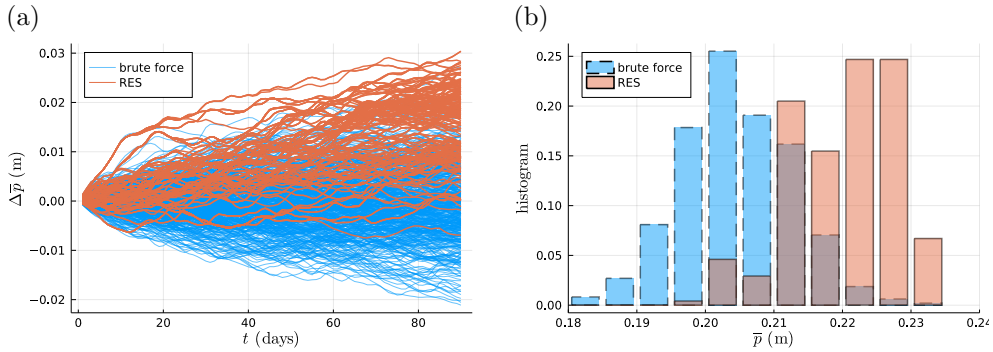

Figure 2. (a) In blue the cumulative rainfall anomaly over time generated in 490 winters of the control run. In orange the cumulative rainfall anomaly for realisations of the RES that have survived all selections until the final time (subtracting the mean of the control run). (b) A histogram of winter European rainfall based on 490 years of the control run and in the rare event simulation.

Using the ensemble generated in the RES run we can estimate probabilities using the estimator B1 and compare them to those obtained from the brute force simulation using estimator (1). Figure 3 shows the estimated complementary cumulative distribution function $\bar{F}(\theta) = \mathbb{P}(\bar{p} > \theta)$, giving the probability of exceeding a threshold θ. We see that for all estimated probabilities, the estimate for the RES is within the $\pm 2\sigma$ interval of the brute force simulation, even though the RES simulation has a total simulation time of only 61.45 years, compared to the much longer control run of 490 years.

Figure 4 shows estimates of the normalized relative error for the brute force estimator and for the rare event simulation. The normalized relative error is the relative error for sample size $N = 1$, hence comparing relative errors for equal sample sizes. Figure 4 demonstrates that in the range that the RES is preferentially sampling (roughly 0.21m to 0.23m, see Fig. 2(b)) there is a strong reduction in the normalized relative error. This reduction demonstrates that with the same numerical effort, we can get a much more accurate estimate using RES, or a similar accuracy with less numerical effort. As the relative error of the brute force estimator scales as $1/\sqrt{N}$, a reduction in normalized relative error by a factor r translates roughly into a reduction of computational effort by a factor r^2. For the rarest events (in the range $(0.2275, 0.2325)$), relative error is reduced by a factor 7.53, meaning the numerical effort could be reduced by a factor 56.84.

A complementary view to Figure 3 is presented in Figure 5. Here a histogram is estimated, showing probabilities $\mathbb{P}(\theta - \Delta\theta/2 < \bar{p} < \theta + \Delta\theta/2)$. We see that both estimates fall within each other's 2σ interval. As expected from 4, the width of the 2σ interval increases faster for the brute force estimates than for the RES estimates. Specifically, we can see from Figure 5 that, for example, the lower bound of the 2σ interval for the estimated probability for the interval $(0.225, 0.230)$ becomes negative for the brute-force simulation, while the RES gives a lower bound of 0.00559.

3.2 Characterization of extreme winter rainfall condition

An important advantage of RES methods, compared to extrapolations based on extreme value theory, is the availability of the full dynamical fields resulting from the numerical integration of the RES ensemble. Here we illustrate this property by comparing the physical fields of the RES ensemble to those of the control run.

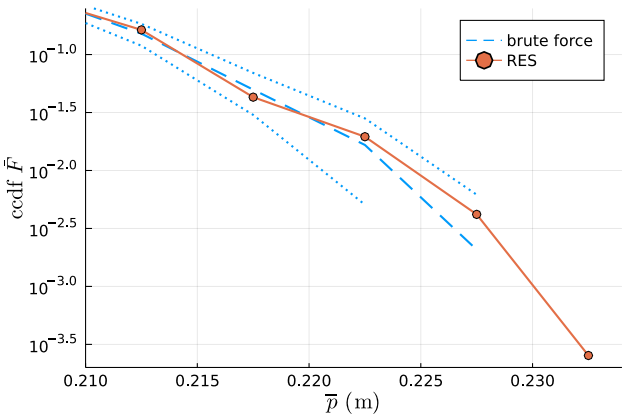

Figure 3. The complementary cumulative distribution function estimated using the brute force estimator (1) (dashed line, with dotted lines indicating the $\pm 2\sigma$ interval, where $\sigma = \sqrt{\left(\bar{F}(1-\bar{F})\right)/N}$) and using the RES estimator (B1) (solid line). Wherever values for brute force estimates are not shown, they are zero or negative.

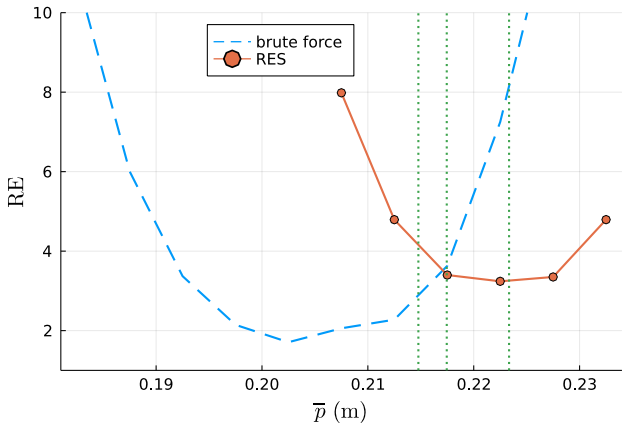

Figure 4. The normalised relative error for histogram estimation, for the brute force simulation $\sqrt{\hat{p}_\theta(1-\hat{p}_\theta)}/\hat{p}_\theta$ (dashed blue line), where $\hat{p}_\theta$ is the probability estimate for the histogram bin $(\theta - \Delta\theta/2, \theta + \Delta\theta/2)$ based on the brute force Monte Carlo estimator (1), and for the RES (solid orange line) using $\sqrt{\tilde{\gamma}_\theta^{(2)}}/\tilde{\gamma}_\theta$, where $\tilde{\gamma}_\theta$ is the probability estimate based on the RES estimator (B1) with $A = (\theta - \Delta\theta/2, \theta + \Delta\theta/2)$, and $\tilde{\gamma}_\theta^{(2)}$ is the estimate of its variance obtained using (B2). Vertical dotted green lines indicate the 0.9, 0.95 and 0.99 quantiles (from left to right).

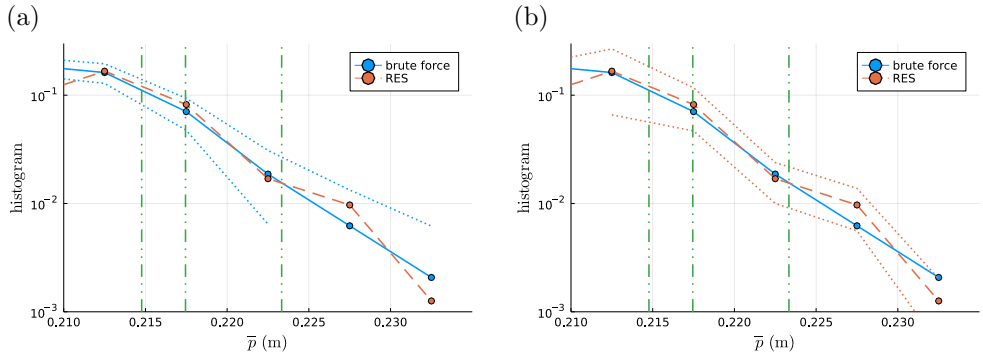

Figure 5. Estimated histograms of rainfall probabilities for bins of width 0.005m between 0.21m and 0.235m. The blue solid line indicates the brute force estimate (1), the orange dashed line indicates the RES estimate (B1). The dotted lines indicate 2σ intervals, based on the relative error shown in Figure 4. The interval shown in (a) is for the brute force estimator, whereas the interval in (b) is for the RES estimator. Wherever values for brute force estimates are not shown, they are zero or negative. Vertical dash-dotted green lines indicate the 0.9, 0.95 and 0.99 quantiles (from left to right).

Fig. 6 shows the mean circulation and precipitation over the North Atlantic and Europe as represented by the PlaSim CONTROL and RES experiments. Both experiments capture basic properties of the regional climate, like the southwest-northeast tilt in the North Atlantic storm track and the precipitation distribution across Europe. The westerlies and atmospheric moisture transport are stronger in RES than in CONTROL. Moreover, the tilt of the storm track in RES is reduced a little. These differences between the experiments result in increased precipitation in RES throughout most of the target domain, and especially over western Europe, to the south of the climatological maximum precipitation near the British Isles.

These differences appear physically plausible; we forego a more detailed analysis of the associated meteorological processes because of the comparative simplicity and low resolution of PlaSim. The results do illustrate how rare event simulation could be used in a more realistic model to investigate the physical processes driving extreme winter rainfall.

4 Conclusion

In this article we have demonstrated the applicability of rare event simulation algorithms for the study of extreme winter precipitation over Europe. We have applied a variant of the class of genealogical algorithms introduced by Del Moral and Garnier. In our setup realizations in an ensemble simulation with sustained rainfall over Europe are cloned into multiple copies to favour the development of extreme winter rainfall.

As the results in Section 3 have shown, using the Del Moral-Garnier algorithm, we can sample much more efficiently realizations with extreme European winter precipitation. On the one hand, this allows to obtain more accurate estimates at the same numerical cost. As demonstrated in Figure 4, for winter rainfall above approximately 0.2175, the RES becomes more efficient. On the other hand, estimates with the same accuracy can be obtained with less numerical effort. Our application demonstrates a combination of both, a relatively short RES of 61.45 years (compared to the 490 years of the control simulation) allows us to obtain probability estimates of rare events, with

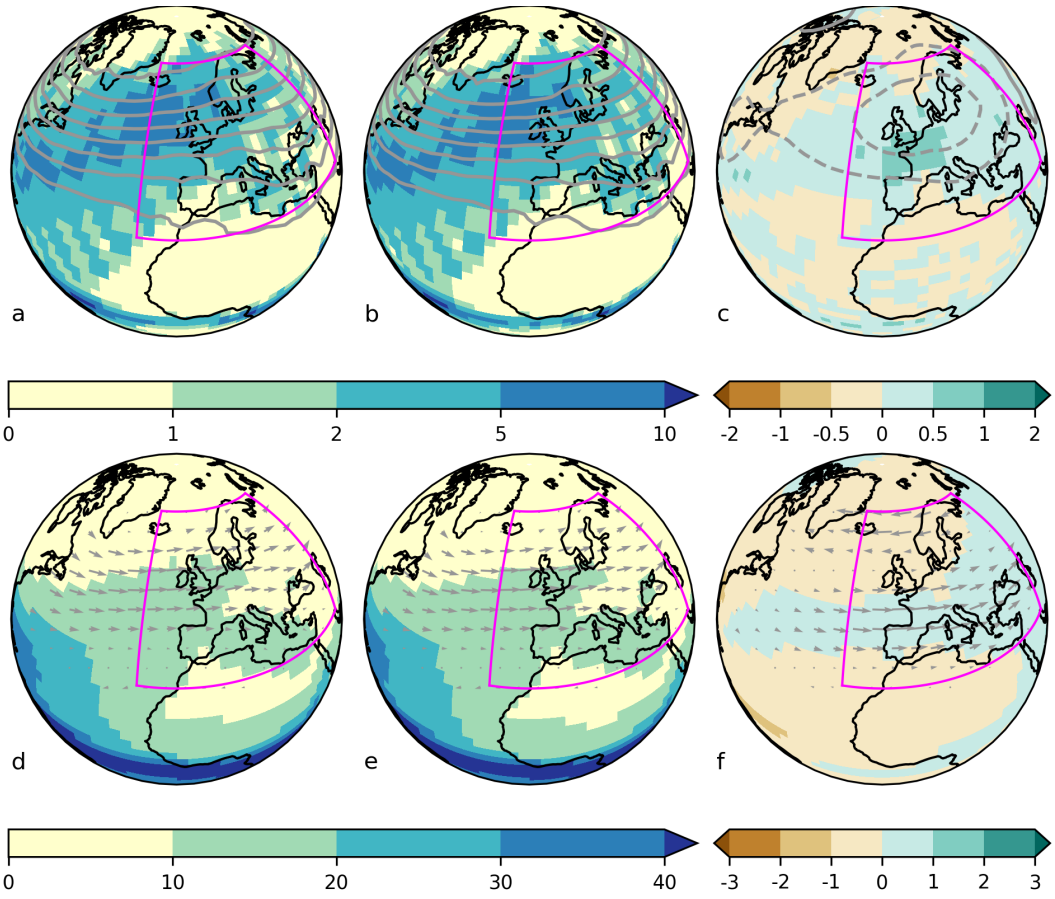

Figure 6. Mean circulation and precipitation in the CONTROL experiment (a, d), the RES experiment (b, e) and the difference RES - CONTROL (c, f). Shading in the top row (a–c) shows precipitation (mm day^{-1}) and grey lines show 500hPa geopotential height (m). Contour lines in (a) and (b) are shown from 5050 m to 5750 m at intervals of 100 m, and lines in (c) from -50 m to 50 m at intervals of 20 m (negative contours are dashed). Shading in the bottom row (d–f) shows precipitable water (kg/m^2) and grey arrows show 800hPa horizontal wind vectors; for reference, the wind speed over a grid box near London (0.0°E, 51.6°N) is 14.36 m s^{-1}, 15.57 m s^{-1}, and 1.21 m s^{-1} in panels d, e, and f, respectively. The boundaries of the European RES target domain Ω are indicated by a magenta solid closed curve.

lower relative error compared to the brute force estimates based on the control (see Figure 5).

Our results highlight the potential for further applications of rare event simulation in weather and climate science. Operational or regular activities such as seasonal forecasting or event attribution are based on model experiments that are conceptually similar to the proof of concept presented in this study, albeit using much more sophisticated, and hence expensive, models. RES would permit these applications at lower cost or greater sophistication. Moreover, RES permits to simulate events that are so rare/extreme that they are simply not accessible within a brute-force approach while still being physically plausible within the chosen model and experiment.

It should be stressed that the methodology used here is not specific to the general circulation model used. In the future, we plan to apply the method to a state-of-the-art, high-resolution coupled climate model. We also expect the method to work well for other geographical regions. Other possible avenues for future research include the testing of other types of weighting functions, taking into account the physical processes leading to rare events to sample them even more efficiently.

Acknowledgements

RKHS and LCS are funded by the National Centre for Atmospheric Science (NERC contract number R8/H12/83).

This document has been drafted using GNU TEXMACS (van der Hoeven et al., n.d.). JW would like to thank Joris van der Hoeven and the team behind GNU TEXMACS for creating their useful software.

Appendix A Exponential tilting of the Gamma distribution

Let Y be Gamma(α, θ) distributed with pdf

$$\rho_Y(y) = \frac{y^{\alpha-1}e^{-y/\theta}}{\Gamma(\alpha)\theta^\alpha}, y > 0.$$

We exponentially tilt the distribution to $\tilde{Y}$ with pdf

$$\tilde{\rho}_{\tilde{Y}}(y) = \frac{y^{\alpha-1}e^{-y/\theta}}{\Gamma(\alpha)\theta^\alpha}\frac{e^{Cy}}{\mathcal{Z}} = \frac{1}{\mathcal{Z}}\frac{y^{\alpha-1}e^{-\left(\frac{1-C\theta}{\theta}\right)y}}{\Gamma(\alpha)\theta^\alpha}$$

where $\mathcal{Z}$ is a normalising factor to ensure that the pdf integrates to 1.

We obtain again a Gamma distribution, with new parameters $(\tilde{\alpha}, \tilde{\theta}) = \left(\alpha, \frac{\theta}{1-C\theta}\right)$. Notice that $\tilde{\theta}$ diverges for $C = \frac{1}{\theta}$. The mean and standard deviation of $\tilde{Y}$ change to

$$\begin{aligned}\tilde{\mu} = \tilde{\alpha}\tilde{\theta} &= \frac{\theta}{1-C\theta}\alpha \\ \tilde{\sigma}^2 = \tilde{\theta}^2\tilde{\alpha} &= \left(\frac{\theta}{1-C\theta}\right)^2\alpha\end{aligned}$$

As $C \to \frac{1}{\theta}$, both $\tilde{\mu}$ and $\tilde{\sigma}^2$ diverge.

If we wish to pick C so as to shift the mean $\tilde{\mu}$ to k standard deviations away from μ, we wish to have $\tilde{\mu} = \mu + k\sigma$. Solving for C gives

$$\begin{aligned}\frac{\theta}{1-C\theta}\alpha &= \theta\alpha + k\theta\sqrt{\alpha} \\ \frac{1}{1-C\theta}\sqrt{\alpha} &= k + \sqrt{\alpha} \\ C &= \frac{1}{\theta}\frac{k}{k+\sqrt{\alpha}}\end{aligned}$$

Appendix B The genealogical particle analysis algorithm

Algorithm 1
Genealogical particle analysis (Del Moral & Garnier, 2005)
With interaction times $T_0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T_1$. The number of particles on the time interval (t_{k-1}, t_k) is denoted by N_k.

1. Initiate M realizations according to the initial distribution, $\xi_0^{(i)}(t_0) \sim \rho_0$ for $1 \leqslant i \leqslant N_1 = M$
2. For $k \in \{1, \ldots, n-1\}$
   (a) Time-evolve $\xi_{k-1}^{(i)}$ for $i \in \{1, \ldots, N_k\}$ under the model dynamics from t_{k-1} to t_k, resulting in $\zeta_k^{(i)}(t)$ with $t \in (t_0, t_k)$.
   (b) Calculate the cloning factor for realization i as
   $$\begin{aligned} \bar{W}_k^{(i)} &= \frac{W_k[\zeta_k^{(i)}]}{Z_k} \\ Z_k &= \frac{1}{N_k} \sum_{i=1}^{N_k} W_k[\zeta_k^{(i)}] \,, \end{aligned}$$
   storing the value of the normalizing factor Z_k.
   (c) Generate a new particle distribution $\xi_k^{(j)}$ consisting of $N_k^{(i)}$ copies of $\zeta_k^{(i)}$where $N_k^{(i)}$ is a random integer such that $\mathbb{E}(N_k^{(i)}) = \bar{W}_k^{(i)}$. We have $N_{k+1} = \sum_{i=1}^{N_k} N_k^{(i)}$.
   (d) If the model is not stochastic apply a random perturbation to make clones diverge over time.
3. Time-evolve $\xi_{n-1}^{(i)}$ for $i \in \{1, \ldots, N_{n-1}\}$ under the model dynamics from t_{n-1} to $t_n = T_1$, resulting in $\zeta_n^{(i)}(t)$ with $t \in (T_0, T_1)$.
4. Finally calculate $\frac{1}{M} \sum_{i=1}^{N_{t_n}} L[\xi_{i,t_n}] F[\xi_{i,t_n}]$ to estimate $\mathbb{E}(F[X])$, where $L[\xi_{i,t_n}] = (\prod_{k=1}^{n-1} \bar{W}_k[\xi_i])^{-1}$ with $\prod_{k=1}^{n-1} \bar{W}_k[\xi_i]$ is the product of weights given to the realisation ξ in step 2b, at each selection step along its history.

To estimate γ_A with the choice of W_k as in (3) we get that $\prod_{i=1}^{n-1} W_k[x] = \exp(C \int_{t_0}^{t_{n-1}} V(x(\tau))\mathrm{d}\tau)$ so we can take $F(x) = \mathbb{1}_A[x]$ and the estimator becomes

$$\tilde{\gamma}_A = \frac{1}{M} \sum_{i=1}^{N_{t_n}} \left(\exp\left(-C \int_{t_0}^{t_{n-1}} V(x_i(\tau))\mathrm{d}\tau \right) \mathbb{1}_A[x_i] \right) \left(\prod_{k=1}^{n-1} Z_k \right) . \tag{B1}$$

In (Del Moral & Garnier, 2005), a central limit theorem result for the estimator result is derived for estimator (B1). This result is important, because asymptotic confidence intervals can be derived from it. In the case of estimation of histogram probabilities $\mathbb{P}\{X \in (\theta - \Delta\theta/2, \theta + \Delta\theta/2)\}$ with $\Delta\theta \to 0$, the leading order term in the variance can be estimated as

$$\tilde{\gamma}_\theta^{(2)} = \frac{1}{M} \sum_{i=1}^{N_{t_n}} \left(\exp\left(-2C \int_{t_0}^{t_{n-1}} V(x(\tau))\mathrm{d}\tau \right) \mathbb{1}_{(\theta-\Delta\theta/2, \theta+\Delta\theta/2)}[x] \right) \left(\prod_{k=1}^{n-1} Z_k \right)^2 . \tag{B2}$$

References

Andres, H. J., & Tarasov, L. (2019, August). Towards understanding po-

tential atmospheric contributions to abrupt climate changes: characterizing changes to the North Atlantic eddy-driven jet over the last deglaciation. *Climate of the Past*, *15*(4), 1621–1646. Retrieved 2022-02-15, from https://cp.copernicus.org/articles/15/1621/2019/ (Publisher: Copernicus GmbH) doi: 10.5194/cp-15-1621-2019

Boschi, R., Lucarini, V., & Pascale, S. (2013, November). Bistability of the climate around the habitable zone: A thermodynamic investigation. *Icarus*, *226*(2), 1724–1742. Retrieved 2022-02-15, from https://www.sciencedirect.com/science/article/pii/S0019103513001309 doi: 10.1016/j.icarus.2013.03.017

Bucklew, J. A. (2004). *An introduction to rare event simulation.* New York: Springer.

Davini, P., Weisheimer, A., Balmaseda, M., Johnson, S. J., Molteni, F., Roberts, C. D., ... Stockdale, T. N. (2021, 1). The representation of winter northern hemisphere atmospheric blocking in ecmwf seasonal prediction systems. *Quarterly Journal of the Royal Meteorological Society*, *147*, 1344-1363. doi: 10.1002/qj.3974

Del Moral, P., & Garnier, J. (2005). Genealogical particle analysis of rare events. *The Annals of Applied Probability*, *15*(4), 2496–2534. doi: 10.1214/105051605000000566

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., ... Ting, M. (2020, 4). Insights from earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, *10*, 277-286. Retrieved from http://www.nature.com/articles/s41558-020-0731-2 doi: 10.1038/s41558-020-0731-2

Fraedrich, K. (2012, May). A suite of user-friendly global climate models: Hysteresis experiments. *The European Physical Journal Plus*, *127*(5), 53. Retrieved 2022-02-15, from https://doi.org/10.1140/epjp/i2012-12053-7 doi: 10.1140/epjp/i2012-12053-7

Fraedrich, K., Jansen, H., Kirk, E., Luksch, U., & Lunkeit, F. (2005). The Planet Simulator: Towards a user friendly model. *Meteorologische Zeitschrift*, *14*(3), 299–304. doi: 10.1127/0941-2948/2005/0043

Fraedrich, K., & Lunkeit, F. (2008, January). Diagnosing the entropy budget of a climate model. *Tellus A: Dynamic Meteorology and Oceanography*, *60*(5), 921–931. Retrieved 2022-02-15, from https://doi.org/10.1111/j.1600-0870.2008.00338.x (Publisher: Taylor & Francis _eprint: https://doi.org/10.1111/j.1600-0870.2008.00338.x) doi: 10.1111/j.1600-0870.2008.00338.x

Lucarini, V., Fraedrich, K., & Lunkeit, F. (2010, October). Thermodynamics of climate change: generalized sensitivities. *Atmospheric Chemistry and Physics*, *10*(20), 9729–9737. Retrieved 2022-02-15, from https://acp.copernicus.org/articles/10/9729/2010/ (Publisher: Copernicus GmbH) doi: 10.5194/acp-10-9729-2010

Lucarini, V., Ragone, F., & Lunkeit, F. (2017, February). Predicting Climate Change Using Response Theory: Global Averages and Spatial Patterns. *Journal of Statistical Physics*, *166*(3), 1036–1064. Retrieved 2022-02-15, from https://doi.org/10.1007/s10955-016-1506-z doi: 10.1007/s10955-016-1506-z

Lyu, G., Köhl, A., Matei, I., & Stammer, D. (2018). Adjoint-Based Climate Model Tuning: Application to the Planet Simulator. *Journal of Advances in Modeling Earth Systems*, *10*(1), 207–222. Retrieved 2022-02-15, from https://onlinelibrary.wiley.com/doi/abs/10.1002/2017MS001194 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2017MS001194) doi: 10.1002/2017MS001194

Margazoglou, G., Grafke, T., Laio, A., & Lucarini, V. (2021, June). Dynamical land-

scape and multistability of a climate model. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *477*(2250), 20210019. Retrieved 2022-02-15, from https://royalsocietypublishing.org/doi/full/10.1098/rspa.2021.0019 (Publisher: Royal Society) doi: 10.1098/rspa.2021.0019

Otto, F. E., Philip, S., Kew, S., Li, S., King, A., & Cullen, H. (2018, 8). Attributing high-impact extreme events across timescales—a case study of four different types of events. *Climatic Change*, *149*, 399-412. doi: 10.1007/s10584-018-2258-3

Priestley, M. D., & Catto, J. L. (2022, 3). Improved representation of extratropical cyclone structure in highresmip models. *Geophysical Research Letters*, *49*. doi: 10.1029/2021GL096708

Priestley, M. D. K., Pinto, J. G., Dacre, H. F., & Shaffrey, L. C. (2017, 7). The role of cyclone clustering during the stormy winter of 2013/2014. *Weather*, *72*, 187-192. Retrieved from https://onlinelibrary.wiley.com/doi/10.1002/wea.3025 doi: 10.1002/wea.3025

Ragone, F., & Bouchet, F. (2021). Rare Event Algorithm Study of Extreme Warm Summers and Heatwaves Over Europe. *Geophysical Research Letters*, *48*(12), e2020GL091197. Retrieved 2022-02-03, from https://onlinelibrary.wiley.com/doi/abs/10.1029/2020GL091197 (_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020GL091197) doi: 10.1029/2020GL091197

Ragone, F., Lucarini, V., & Lunkeit, F. (2016, March). A new framework for climate sensitivity and prediction: a modelling perspective. *Climate Dynamics*, *46*(5), 1459–1471. Retrieved 2022-02-15, from https://doi.org/10.1007/s00382-015-2657-3 doi: 10.1007/s00382-015-2657-3

Ragone, F., Wouters, J., & Bouchet, F. (2017, December). Computation of extreme heat waves in climate models using a large deviation algorithm. *Proceedings of the National Academy of Sciences*, 201712645. Retrieved from http://www.pnas.org/content/early/2017/12/18/1712645115 doi: 10.1073/pnas.1712645115

Roberts, M. J., Vidale, P. L., Senior, C., Hewitt, H. T., Bates, C., Berthou, S., ... Wehner, M. F. (2018, 11). The benefits of global high resolution for climate simulation: Process understanding and the enabling of stakeholder decisions at the regional scale. *Bulletin of the American Meteorological Society*, *99*, 2341-2359. Retrieved from https://journals.ametsoc.org/bams/article/99/11/2341/70243/The-Benefits-of-Global-High-Resolution-for-Climate doi: 10.1175/BAMS-D-15-00320.1

Schaller, N., Kay, A. L., Lamb, R., Massey, N. R., Oldenborgh, G. J. v., Otto, F. E. L., ... Allen, M. R. (2016). Human influence on climate in the 2014 southern England winter floods and their impacts. *Nature Climate Change*, *6*(6), 627–634. Retrieved 2020-05-07, from https://www.nature.com/articles/nclimate2927 doi: 10.1038/nclimate2927

Schiemann, R., Vidale, P. L., Shaffrey, L. C., Johnson, S. J., Roberts, M. J., Demory, M.-E., ... Strachan, J. (2018, 7). Mean and extreme precipitation over european river basins better simulated in a 25 km agcm. *Hydrology and Earth System Sciences*, *22*, 3933-3950. Retrieved from https://www.hydrol-earth-syst-sci.net/22/3933/2018/https://hess.copernicus.org/articles/22/3933/2018/ doi: 10.5194/hess-22-3933-2018

Seneviratne, S. I., Zhang, X., Adnan, M., Badi, W., Dereczynski, C., Luca, A. D., ... Wehner, M. (2021). Weather and climate extreme events in a changing climate. In V. Masson-Delmotte et al. (Eds.), (p. 1513-1766). Cambridge University Press. Retrieved from https://www.ipcc.ch/report/sixth-assessment-report-working-group-i

van der Hoeven et al., J. (n.d.). *Gnu texmacs.* https://www.texmacs.org.

Vannière, B., Demory, M. E., Vidale, P. L., Schiemann, R., Roberts, M. J., Roberts, C. D., ... Senan, R. (2019, 6). Multi-model evaluation of the sensitivity of the global energy budget and hydrological cycle to resolution. *Climate Dynamics*, *52*, 6817-6846. doi: 10.1007/s00382-018-4547-y

Wan, H., Zhang, X., & Zwiers, F. (2019, 1). Human influence on canadian temperatures. *Climate Dynamics*, *52*, 479-494. doi: 10.1007/s00382-018-4145-z

Webber, R. J., Plotkin, D. A., O'Neill, M. E., Abbot, D. S., & Weare, J. (2019, May). Practical rare event sampling for extreme mesoscale weather. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *29*(5), 053109. Retrieved 2019-05-24, from `https://aip.scitation.org/doi/full/10.1063/1.5081461` doi: 10.1063/1.5081461

Zappa, G., Shaffrey, L. C., & Hodges, K. I. (2013). The Ability of CMIP5 Models to Simulate North Atlantic Extratropical Cyclones*. *Journal of Climate*, *26*(15), 5379–5396. Retrieved from `http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00501.1` doi: 10.1175/JCLI-D-12-00501.1