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

Article

Published Version

Wouters, J. ORCID: https://orcid.org/0000-0001-5418-7657, Schiemann, R. K. H. ORCID: https://orcid.org/0000-0003-3095-9856 and Shaffrey, L. C. ORCID: https://orcid.org/0000-0003-2696-752X (2023) Rare event simulation of extreme European winter rainfall in an intermediate complexity climate model. Journal of Advances in Modeling Earth Systems, 15 (4). e2022MS003537. ISSN 1942-2466 doi: 10.1029/2022MS003537 Available at https://centaur.reading.ac.uk/111397/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1029/2022MS003537

Rare Event Simulation of Extreme European Winter Rainfall in an Intermediate Complexity Climate Model

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

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

Abstract We test the application of a rare event simulation (RES) 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 RES of seasonal precipitation extremes.

Plain Language Summary Rare events, such as winters with a lot of rainfall, can have a high impact. High rainfall can cause floods, risking lives, and damaging property. To plan for such events, we want to know what the probability is to experience a lot of rainfall in any given year. One way to make an estimate is to run climate models and look at the proportion of winters having rainfall above a high threshold. This method can be very costly and time-consuming, as climate models require powerful computers and we need to run them for a long time until we see a few winters with very high rainfall. In this article, we apply 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 saves time 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 (a) quantifying the probability of extreme event occurrence and detecting changes over time, (b) attributing such changes to natural variability as well as natural and anthropogenic forcing, including event attribution for individual observed extremes, (c) predicting occurrence probabilities at lead times of days to decades, and (d) 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 initialized 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 modeling 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 (M. D. K. Priestley et al., 2017; Schaller et al., 2016). Due to the length of the event itself many simulated years

are required to sample extreme seasons. At the same time, higher model resolution benefits the representation of storm tracks, storm structure, the water cycle and extreme precipitation, and a global coupled modeling setup allows for the representation of teleconnections with the tropics or the Arctic (M. D. Priestley & Catto, 2022; Roberts et al., 2018; Schiemann et al., 2018; Vannière et al., 2019; Zappa et al., 2013). Limited computing resource may require compromises between these different demands.

Efficient simulation strategies offer a method to resolve such dilemmas. Rare event simulation (RES) algorithms, as applied in this article, provide ways to generate many more samples of an extreme event. This larger set of samples in turn allows for more accurate estimates of the probability of occurrence of the event, or to obtain the same accuracy with less numerical effort. Originally developed in statistical physics (Kahn & Harris, 1951), RES algorithms have found applications in, for example, queuing networks (Heidelberger, 1995), biochemistry (Kuwahara & Mura, 2008), and aircraft design (Morio & Balesdent, 2015). In Earth system modeling, rare event algorithms have previously been successfully applied to atmospheric models, to over-sample heat waves (Ragone & Bouchet, 2021; Ragone et al., 2017) 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. While RES has already been successfully demonstrated on Earth system models before, different types of extremes develop differently over time. A successful RES setup for one quantity, such as temperature, may not necessarily carry over, with the same efficiency improvements, to other quantities, such as precipitation. As experiments with high resolution climate models are numerically demanding, there is a need for exploratory demonstrations of RES on intermediate complexity models. We therefore implement RES 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 precipitation far into the tail of the distribution and quantify how our RES approach has significantly reduced the numerical effort required, compared to traditional brute force approaches.

We will use an ensemble based genealogical algorithm similar to those used in Ragone and Bouchet (2021), Ragone et al. (2017), and Webber et al. (2019), which are, in turn, variants of the algorithm described in Del Moral and 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, as quantified by the relative sampling error of the estimate.

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 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 min.

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 Figure 6).

PlaSim can be considered an intermediate complexity model, which makes it suitable for explorations of novel computational methods, 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 (Boschi et al., 2013; Fraedrich, 2012; Margazoglou et al., 2021), and linear response theory (Lucarini et al., 2017; Ragone et al., 2016).

2.2. Rare Event Simulation

2.2.1. Importance Sampling

The fundamental idea behind many RES 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 (RE), defined as $\mathrm{RE}_{\mathrm{MC}} = \sqrt{\mathbb{V}(\hat{\gamma}_A)}/\gamma_A$, the standard deviation of the estimator relative to γ_A, is approximately $1/\sqrt{N\gamma_A}$ for small γ_A (Rubinstein & Kroese, 2017). Therefore, for rare events with $\gamma_A \ll 1$ we would need a very large number of samples N to keep the RE at an acceptably low level.

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, with $\tilde{\rho}$ such that $\tilde{\rho}(x) > 0$ whenever $\mathbb{1}_A(x)\rho(x) > 0$. By generating independent samples $\tilde{X}_i$ from $\tilde{X}$ the IS 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 RE of this estimator, $\mathrm{RE}_{\mathrm{IS}} = \sqrt{\mathbb{V}(\tilde{\gamma}_A)}/\gamma_A$, can be greatly reduced in comparison to that of the brute force Monte Carlo estimator ($\mathrm{RE}_{\mathrm{MC}}$). The exact reduction depends on the probability distributions involved and can rarely be computed analytically. At least theoretically, an IS estimator with zero RE can be constructed by taking $\tilde{\rho} = (\rho/\gamma_A)\mathbb{1}_A$, although it is of no practical use, since it involves knowledge of γ_A, the very quantity we want to estimate. This optimal change of measure does illustrate that the change of measure should make the rare event more common. It should also be noted that a poor choice of distribution $\tilde{\rho}$ (one that makes the rare event even rarer), can lead to an increase in RE. For more details on IS, see, for example (Bucklew, 2004; Rubino & Tuffin, 2009; Rubinstein & Kroese, 2017).

2.2.2. Rare Event Simulation Algorithm

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 1/1000 years event. The aim of RES algorithms is to perform the change of measure that is necessary to implement IS, spending more effort simulating relevant realizations. 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.

The idea of IS can be generalized to stochastic processes. In such cases, a change of measure can be carried out by altering the dynamical equations defining the process, for example, by changing the drift of a diffusion process. The probabilities of sample paths is then altered, with the Radon-Nikodym derivative given by the Girsanov theorem. This can again be used to implement an unbiased estimator, as in the univariate case above (Vanden-Eijnden & Weare, 2012). When considering a high-dimensional climate model, this method faces several challenges. First of all, it is often not evident how the system should be altered to generate more realizations of the rare event
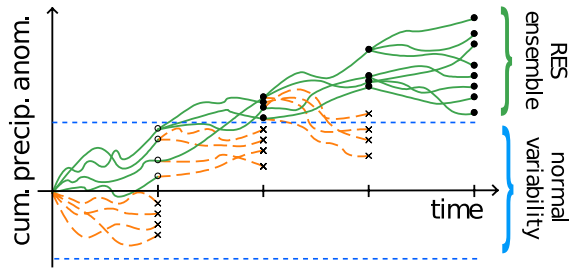
Figure 1. A graphical representation of the cumulative precipitation anomalies over time in an ensemble under the rare event simulation algorithm. Trajectories not developing an extreme cumulative precipitation event have a high probability of termination (orange dashed lines). At the final time extreme cumulative precipitation has become common in the ensemble.

in question. Second, altering a complex model to add such forcing may be technically demanding. Finally, calculating the Radon-Nikodym derivative in very high dimension is numerically challenging.

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 a 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 have a higher likelihood of being terminated. In our application this would be ensemble members with little precipitation over Europe. 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 realization 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 IS estimator 2, the change of measure needs to be taken into account in the estimation step 4. The probability of each realization has been changed at each selection step by a factor $\bar{W}_k^{(i)}$, so the overall probability has been changed by the product $L[\xi_{i,t_n}] = \left(\prod_{k=1}^{n-1} \bar{W}_k[\xi_i] \right)^{-1}$. This, therefore, needs to be divided out to make the estimation unbiased.

Importantly, an estimator of the variance of the GPA estimator, and hence the RE, has been derived (Del Moral & Garnier, 2005), allowing us to compare a posteriori the accuracy of the GPA estimator to brute force Monte Carlo (see Equation B2 in Appendix B).

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

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

can be used, favoring realizations with a sustained high value of the integral of V. The control parameter C determines the strength of the selection procedure, allowing the user 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 terminating and cloning of realizations, it can be applied to complex climate models.

2.2.3. The RES Configuration Used

The random variable we are interested in is the seasonally integrated total precipitation over a given area

$$X = \bar{p} := \int_{(T_0,T_1)} \int_\Omega p(x(t), r) \, dr \, dt$$

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 −20°E and 50°W and 30° and 70°N. The temporal integration is performed over December, January and February, that is, between $T_0 = 1$ December 00:00 and $T_1 = 1$ March 00:00.

The initial distribution is sampled $M = 250$ times on 1 December from a 500 years control simulation. Cloning is done every 5 days, that is, $t_k - t_{k-1} = 5$ days. This cloning interval is chosen such that clones have sufficient time

to diverge in between cloning, but don't revert to the mean. The error doubling time for PlaSim is estimated to be 2.7 days (Lyu et al., 2018). As control simulation we perform 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, that is, $V(x) = \int_\Omega p(x, r)\mathrm{d}r$ and

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

The parameter C was chosen based by fitting a Gamma distribution to a histogram of the variable of interest $\bar{p}$, total winter precipitation over Europe, 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 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. The resulting value is $C = 371.30$.

3. Results

3.1. Effectiveness of the RES Algorithm

We first verify that the RES is working effectively. Figure 2a shows the cumulative rainfall anomalies generated in the RES versus the brute force control run. We see that, as expected, over time the realizations with the highest precipitation have been selected, tilting the distribution at the final time toward extreme events. Figure 2b shows histograms of winter precipitation generated from 490 years of the 500 years control run (discarding the initial 10 years) and from RES. 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 0.220 m. Only 13 winters in the 490 years control run have winter rainfall exceeding this level.

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 RE for the brute force estimator and for the RES. The normalized RE is the RE 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 0.21–0.23 m, see Figure 2b) there is a strong reduction in the normalized RE. 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. This is the case, at least, in the range the RES is preferentially sampling. By varying the parameter C, this range can be tuned to a range of interest. As the RE of the brute force estimator is proportional to $1/\sqrt{N}$, an increase of N to $r^2 N$ would be required to reduce the RE by a factor r with brute force Monte Carlo. Said otherwise, if RES can reduce the normalized RE by a factor r, this translates into a reduction of computational effort by a factor r^2. For the rarest events (in the range (0.2275, 0.2325)), RE 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, interval probabilities $\mathbb{P}(\theta - \Delta\theta/2 < \bar{p} < \theta + \Delta\theta/2)$ are estimated. 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 still gives a positive lower bound of 0.00559. By definition the probability estimates of the brute-force MC are positive (see Equation 1). The negative lower bound of the 2σ interval is a manifestation of the increasing RE caused by a high probability that, at this sample size, an event of this rarity will not be realized, resulting
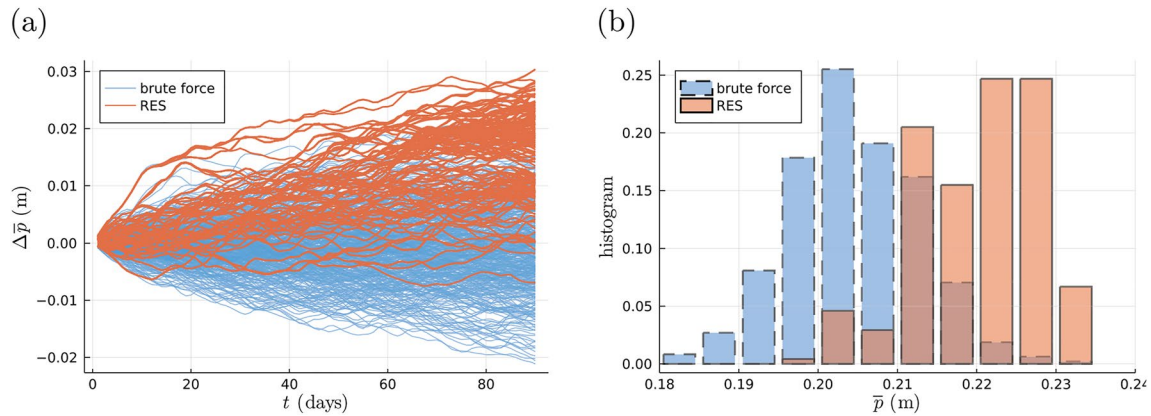
(a)

(b)

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 realizations of the rare event simulation (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 RES. The upward shift of the histogram in panel (b) is also reflected in the positive anomalies in the RES ensemble at day 91 in panel (a).

in a zero probability estimate. It should be noted that this occurs when the inverse probability becomes of a similar magnitude to the sample size.

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 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 RES 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 RES algorithms for the study of extreme winter precipitation over Europe. We have applied a variant of the class of genealogical algorithms (Del Moral & Garnier, 2005). In our setup realizations in an ensemble simulation with sustained rainfall over Europe are cloned into multiple copies to favor the development of extreme winter rainfall.

As the results in Section 3 have shown, using the GPA RES algorithm described in Section 2.2.2 and Appendix B, we can sample much more efficiently realizations with extreme European winter precipitation. On the one hand, this allows to obtain more accurate estimates of rare event probabilities
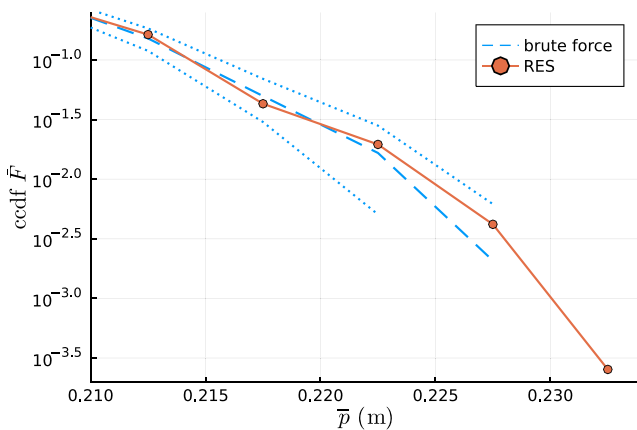

Figure 3. The complementary cumulative distribution function estimated using the brute force estimator 1 (dashed line, with dotted lines indicating the $\pm 2\sigma$ interval, where $\sigma = \sqrt{(\bar{F}(1-\bar{F}))/N}$) and using the rare event simulation estimator B1 (solid line). Wherever values for brute force estimates are not shown, they are zero or negative.
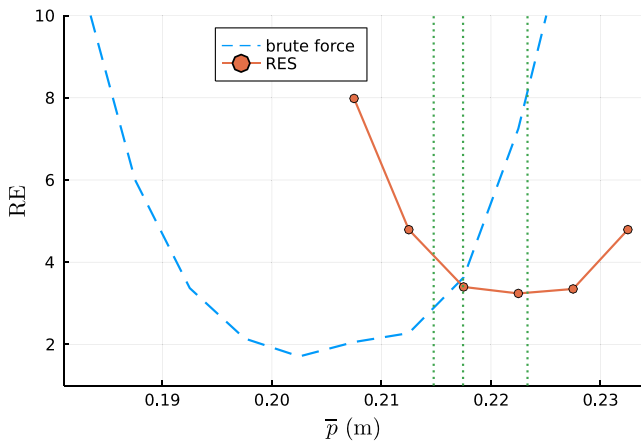
Figure 4. The normalized relative error (RE) for interval probability estimation, for the brute force simulation $\sqrt{\hat{p}_\theta(1-\hat{p}_\theta)}/\hat{p}_\theta$ (dashed blue line), where $\hat{p}_\theta$ is the probability estimate for the interval $(\theta - \Delta\theta/2, \theta + \Delta\theta/2)$ based on the brute force Monte Carlo estimator [1], and for the rare event simulation (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).

at the same numerical cost. As demonstrated in Figure 4, for winter rainfall above approximately 0.2175 m, 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 lower RE compared to the brute force estimates based on the control (see Figure 5).

Our results highlight the potential for further applications of RES in weather and climate science. Operational 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 and/or greater sophistication, hence allowing for a more in-depth investigation of the physical drivers of extremes. 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. Extrapolations based on limit theorems, such as extreme value theory or large deviation theory, can provide probabilistic information, but not time-dependent geographical trajectories of individual realizations, as does GPA.

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. Another possible avenue 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.
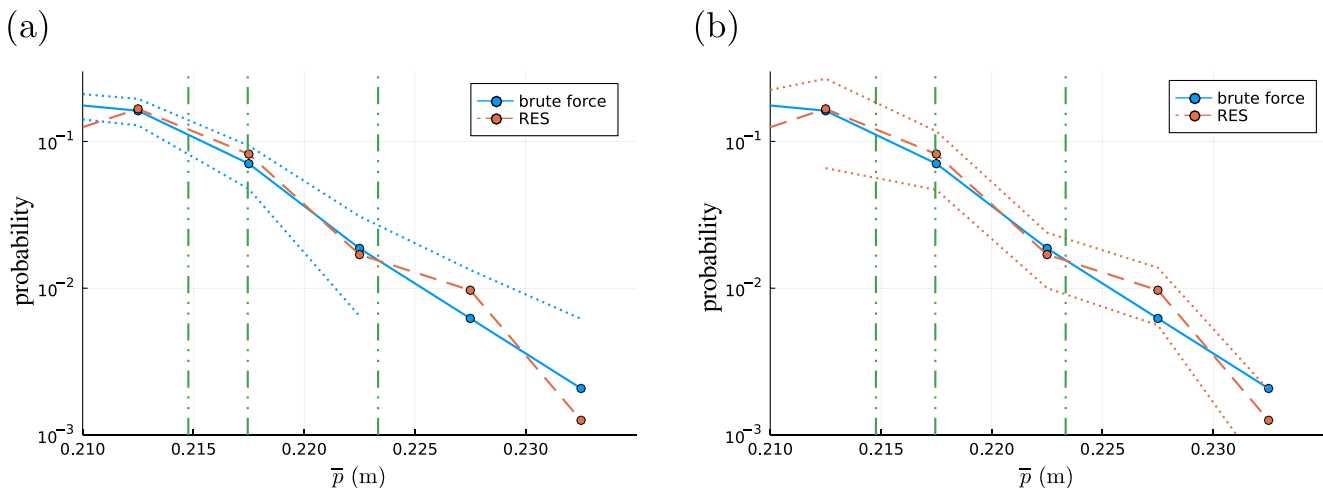

Figure 5. Estimated rainfall probabilities for intervals of width 0.005 m between 0.21 and 0.235 m. The blue solid line indicates the brute force estimate [1], the orange dashed line indicates the rare event simulation (RES) estimate [B1]. The dotted lines indicate 2σ intervals, based on the relative error shown in Figure 4. The 2σ intervals shown in panel (a) are for the brute force estimator, whereas the 2σ intervals in panel (b) are for the RES estimator. Wherever values of estimates or interval lower bounds for brute force estimation 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).
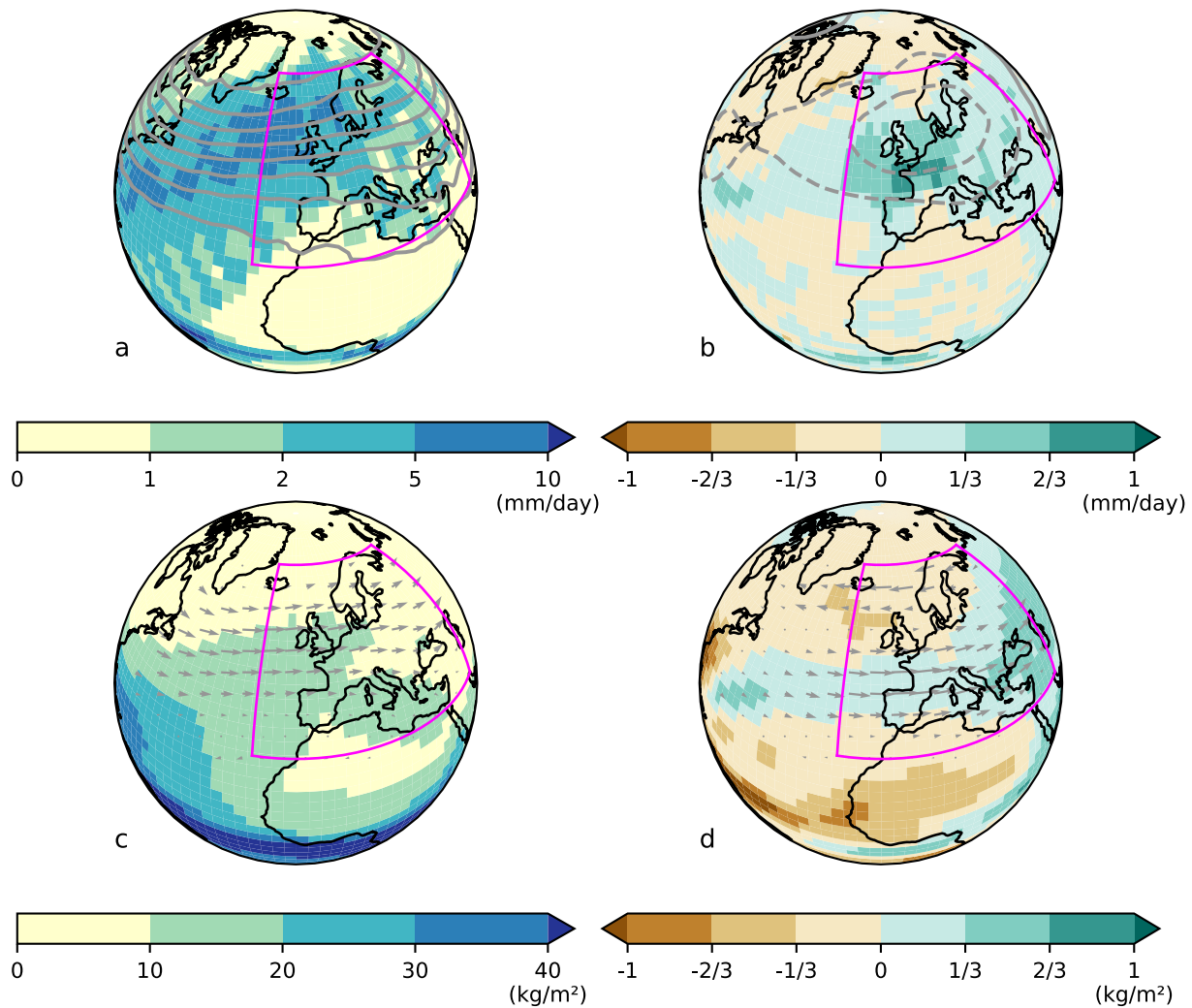
Figure 6. Mean circulation and precipitation in the CONTROL ensemble (a, c) and the difference of the mean of the rare event simulation (RES) ensemble and the mean of the CONTROL ensemble (b, d). Shading in the top row (a–b) shows precipitation (mm day^{-1}) and gray lines show 500 hPa geopotential height (m). Contour lines in panel (a) are shown from 5,050 to 5,750 m at intervals of 100 m, and lines in panel (b) from −50 to 50 m at intervals of 20 m (negative contours are dashed). Shading in the bottom row (c, d) shows precipitable water (kg/m²) and gray arrows show 800 hPa horizontal wind vectors; for reference, the wind speed over a grid box near London (0.0°E, 51.6°N) is 14.36 and 1.21 m s^{-1} in panels (c, d), respectively. The boundaries of the European RES target domain Ω are indicated by a magenta solid closed curve.

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 normalizing 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

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

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

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 $\xi_k^{(i)}$ 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 $L[\xi_{i,t_n}] = \left(\prod_{k=1}^{n-1} \bar{W}_k[\xi_i]\right)^{-1}$ with $\prod_{k=1}^{n-1} \bar{W}_k[\xi_i]$ is the product of weights given to the realization 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 $\prod_{i=1}^{n-1} W_k[x] = \exp\left(C \int_{t_0}^{t_{n-1}} V(x(\tau)) d\tau \right)$ so we can take $F(x) = \mathbb{1}_A[x]$ and the estimator becomes

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

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 interval 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

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

Data Availability Statement

Jupyter notebooks and Julia scripts to run the rare event simulation described in this manuscript can be found at https://gitlab.act.reading.ac.uk/j.wouters/plasimgpa.jl. The Planet Simulator climate model can be downloaded from https://github.com/HartmutBorth/PLASIM (The version of PlaSim used in this study is Git commit ea4d2d8, dated 28/06/2017). This document has been drafted using GNU TEXMACS (van der Hoeven et al., N.D).

References

Andres, H. J., & Tarasov, L. (2019). Towards understanding potential 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. https://doi.org/10.5194/cp-15-1621-2019

Boschi, R., Lucarini, V., & Pascale, S. (2013). Bistability of the climate around the habitable zone: A thermodynamic investigation. *Icarus*, *226*(2), 1724–1742. https://doi.org/10.1016/j.icarus.2013.03.017

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

Davini, P., Weisheimer, A., Balmaseda, M., Johnson, S. J., Molteni, F., Roberts, C. D., et al. (2021). The representation of winter northern hemisphere atmospheric blocking in ECMWF seasonal prediction systems. *Quarterly Journal of the Royal Meteorological Society*, *147*(735), 1344–1363. https://doi.org/10.1002/qj.3974

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

Deser, C., Lehner, F., Rodgers, K. B., Ault, T., Delworth, T. L., DiNezio, P. N., et al. (2020). Insights from Earth system model initial-condition large ensembles and future prospects. *Nature Climate Change*, *10*(4), 277–286. https://doi.org/10.1038/s41558-020-0731-2

Fraedrich, K. (2012). A suite of user-friendly global climate models: Hysteresis experiments. *The European Physical Journal Plus*, *127*(5), 53. https://doi.org/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. https://doi.org/10.1127/0941-2948/2005/0043

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

Heidelberger, P. (1995). Fast simulation of rare events in queueing and reliability models. *ACM Transactions on Modeling and Computer Simulation*, *5*(1), 43–85. https://doi.org/10.1145/203091.203094

Kahn, H., & Harris, T. E. (1951). Estimation of particle transmission by random sampling. *National Bureau of Standards Applied Mathematics Series*, *12*, 27–30.

Kuwahara, H., & Mura, I. (2008). An efficient and exact stochastic simulation method to analyze rare events in biochemical systems. *The Journal of Chemical Physics*, *129*(16), 165101. https://doi.org/10.1063/1.2987701

Lucarini, V., Fraedrich, K., & Lunkeit, F. (2010). Thermodynamics of climate change: Generalized sensitivities. *Atmospheric Chemistry and Physics*, *10*(20), 9729–9737. https://doi.org/10.5194/acp-10-9729-2010

Lucarini, V., Ragone, F., & Lunkeit, F. (2017). Predicting climate change using response theory: Global averages and spatial patterns. *Journal of Statistical Physics*, *166*(3), 1036–1064. https://doi.org/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. https://doi.org/10.1002/2017MS001194

Margazoglou, G., Grafke, T., Laio, A., & Lucarini, V. (2021). Dynamical landscape and multistability of a climate model. *Proceedings of the Royal Society A: Mathematical, Physical & Engineering Sciences*, *477*(2250), 20210019. https://doi.org/10.1098/rspa.2021.0019

Morio, J., & Balesdent, M. (2015). *Estimation of rare event probabilities in complex aerospace and other systems: A practical approach*. Elsevier Science & Technology.

Otto, F. E., Philip, S., Kew, S., Li, S., King, A., & Cullen, H. (2018). Attributing high-impact extreme events across timescales—A case study of four different types of events. *Climatic Change*, *149*(3–4), 399–412. https://doi.org/10.1007/s10584-018-2258-3

Priestley, M. D., & Catto, J. L. (2022). Improved representation of extratropical cyclone structure in HighResMIP models. *Geophysical Research Letters*, *49*(5), e2021GL096708. https://doi.org/10.1029/2021GL096708

Priestley, M. D. K., Pinto, J. G., Dacre, H. F., & Shaffrey, L. C. (2017). The role of cyclone clustering during the stormy winter of 2013/2014. *Weather*, *72*(7), 187–192. https://doi.org/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. https://doi.org/10.1029/2020GL091197

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

Ragone, F., Wouters, J., & Bouchet, F. (2017). Computation of extreme heat waves in climate models using a large deviation algorithm. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(1), 24–29. https://doi.org/10.1073/pnas.1712645115

Roberts, M. J., Vidale, P. L., Senior, C., Hewitt, H. T., Bates, C., Berthou, S., et al. (2018). 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*(11), 2341–2359. https://doi.org/10.1175/BAMS-D-15-00320.1

Rubino, G. & Tuffin, B. (Eds.) (2009). *Rare event simulation using Monte Carlo methods*. Wiley.

Rubinstein, R. Y., & Kroese, D. P. (2017). Simulation and the Monte Carlo method/Reuven Y. Rubinstein, Dirk P. Kroese (3rd ed.).

Schaller, N., Kay, A. L., Lamb, R., Massey, N. R., Oldenborgh, G. J. V., Otto, F. E. L., et al. (2016). Human influence on climate in the 2014 southern England winter floods and their impacts. *Nature Climate Change*, *6*(6), 627–634. https://doi.org/10.1038/nclimate2927

Schiemann, R., Vidale, P. L., Shaffrey, L. C., Johnson, S. J., Roberts, M. J., Demory, M.-E., et al. (2018). Mean and extreme precipitation over European river basins better simulated in a 25 km AGCM. *Hydrology and Earth System Sciences*, *22*(7), 3933–3950. https://doi.org/10.5194/hess-22-3933-2018

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

Vanden-Eijnden, E., & Weare, J. (2012). Rare event simulation of small noise diffusions. *Communications on Pure and Applied Mathematics*, *65*(12), 1770–1803. https://doi.org/10.1002/cpa.21428

van der Hoeven, J. (N.D.). GNU TeXmacs. Retrieved from https://www.texmacs.org

Vannière, B., Demory, M. E., Vidale, P. L., Schiemann, R., Roberts, M. J., Roberts, C. D., et al. (2019). Multi-model evaluation of the sensitivity of the global energy budget and hydrological cycle to resolution. *Climate Dynamics*, *52*(11), 6817–6846. https://doi.org/10.1007/s00382-018-4547-y

Wan, H., Zhang, X., & Zwiers, F. (2019). Human influence on Canadian temperatures. *Climate Dynamics*, *52*(1–2), 479–494. https://doi.org/10.1007/s00382-018-4145-z

Webber, R. J., Plotkin, D. A., O'Neill, M. E., Abbot, D. S., & Weare, J. (2019). Practical rare event sampling for extreme mesoscale weather. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *29*(5), 053109. https://doi.org/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. https://doi.org/10.1175/JCLI-D-12-00501.1