Impact of land surface parameters and large-scale warming on near-surface temperature extremes: A study using rare event sampling

A Thesis

submitted to

Indian Institute of Science Education and Research Pune

in partial fulfillment of the requirements for the

BS-MS Dual Degree Programme

by

Abel Shibu

Indian Institute of Science Education and Research Pune

Dr. Homi Bhabha Road,

Pashan, Pune 411008, INDIA.

April, 2023

Supervisor: Joy Monteiro

© Abel Shibu 2023

Certificate

This is to certify that this dissertation entitled Impact of land surface parameters and large-scale warming on near-surface temperature extremes: A study using rare event sampling towards the partial fulfilment of the BS-MS dual degree programme at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Abel Shibu at Indian Institute of Science Education and Research under the supervision of Joy Monteiro, Assistant Professor, Department of Earth and Climate Science, during the academic year 2022-2023.

Joy Monteiro

Committee:

Joy Monteiro

Argha Banerjee

This thesis is dedicated to my little sister

Declaration

I hereby declare that the matter embodied in the report entitled Impact of land surface parameters and large-scale warming on near-surface temperature extremes: A study using rare event sampling are the results of the work carried out by me at the Department of Earth and Climate Science, Indian Institute of Science Education and Research, Pune, under the supervision of Joy Monteiro and the same has not been submitted elsewhere for any other degree.

Abel Shibu

Acknowledgments

First and foremost, I wish to thank my supervisor Dr. Joy Monteiro for his excellent guidance and constant support throughout this project and for helping me grow, both at a professional as well as at a personal level. I thank him for being kind and patient with me and for gently guiding me whenever I felt stuck.

I thank Dr. Freddy Bouchet and his group at ENS de Lyon for providing me with the opportunity to visit Lyon and work with them, without which the rare event sampling algorithm could not have been implemented. I thank Corentin, Bastien, Clément, Alessandro, Sohan and Valeria for the productive discussions and insights. I specially thank Anna and Clément for their support and company during my time in Lyon.

I specially thank my family for being understanding during this time and for letting me figure things out at my own pace. I thank my labmates Hardik, Dilip and Devosmita for their constant support during this project and for the good times. I thank my friends, especially Daniel, Pranav, Adish and Arun for their valuable inputs at various points throughout this project.

I thank and acknowledge IISER Pune for providing me access to the support and resources of the PARAM Brahma Facility under the National Supercomputing Mission, Government of India, which I have used for all the climate model simulations in this project.

Last but not least, I wish to thank my friend Sarathchandran for patiently guiding me through some of the most technical parts of this project. Without his timely help, this project could not have been completed.

x

Abstract

Heatwaves are large deviations of near-surface temperatures from the climatology, over a period of around a week to sometimes as long as a month. In addition to severely impacting human health, heatwaves also affect the economy and infrastructure. In the current warming scenario, heatwaves have been predicted to become more common due to projected changes in land surface properties and the large-scale circulation due to warming.

In this project, we study the effects of land surface properties and large-scale warming on heatwaves. We do this by setting up multiple model configurations with different land surface properties and large-scale warming. To sample heatwaves in a computationally efficient manner, we utilise the GKLT algorithm, a rare event sampling algorithm. Using the GKLT algorithm, we calculate the return times of extreme heatwaves.

We find that the return times of temperature extremes change as the land surface properties and the large-scale circulation are changed. We see that the shapes of the extreme tails of the temperature distribution change between the different model configurations, leading to the changes in the return times observed.

Finally, we perform a statistical analysis on the heatwaves sampled by the algorithm for the different configurations and study how the changes in return times can be understood in terms of the intensity, duration and number of heatwaves. We find that for the land surface configurations, the return times of the time-averaged temperature anomalies decrease uniformly across anomalies as the relative humidity over land is decreased. For the warming configurations, the return times of the time-averaged temperature anomalies decrease predominantly for the larger anomaly values as the atmosphere is warmed, while the return times of the lower anomaly values remain close to that of the baseline.

Contents

List of Figures

1

4

List of Tables

Chapter 1

Introduction

Heatwaves are the deadliest class of extreme weather events, with a higher mean attributed annual mortality than any other class (Kyselý, 2002). In the absence of a formal definition, heatwaves are episodes of extreme positive deviation of temperature from the climate. The Paris 2003 (Vandentorren et al., 2006), Chicago 2005 (Karl and Knight, 1997), Russia 2010 (Wright et al., 2014) and India 2015 (Ghatak et al., 2017) heatwave incidents are some of the more infamous examples, where persistent hot temperatures had a pronounced, adverse effect on human well-being. In addition to severely affecting human health, resulting in high mortality rate (Patz et al., 2005), heatwaves also have a strong impact on the economy through a decrease in the labour capacity (Dunne et al., 2013) and increased demand for resources like electricity (Hsiang et al., 2017).

From the available observational records, the frequency and intensity of temperature extremes have been shown to be increasing through the decades (S. E. Perkins et al., 2012), with the frequency and intensity expected to continue to increase in a warming climate (Dosio et al., 2018). Given the extensive impact of heatwaves and the current trend, we wish to understand the large-scale dynamics that drive temperatures to the extremes and study how the extremes are affected by changes in large-scale dynamics.

Previous studies of heatwaves over the Indian region (Rohini et al., 2016; Sandeep and Prasad, 2018) have shown that these heatwaves are associated with quasi-stationary Rossby wave patterns over the mid-latitudes. The ensuing anomalous anticyclonic flow in the middle and upper troposphere and persistent highs cause a sinking motion, which warms the surface

by adiabatic compression. Moreover, these conditions also result in clear skies, maximising the insolation that reaches the surface. The other observed mechanism for heatwaves is atmospheric blocking (Xoplaki et al., 2012). The European, Russian and Chicago heatwaves have been attributed to atmospheric blocking episodes (Sarah E. Perkins, 2015).

In addition to the aforementioned large-scale atmospheric dynamics, land-surface interactions also need to be accounted for to better capture the variability from the climate (Sandeep and Prasad, 2018; Seneviratne, Lüthi, et al., 2006). Surface temperatures are strongly influenced by land-surface properties like soil moisture and surface fluxes. Low soil moisture can trigger positive land-atmosphere feedbacks that enhance sensible heat flux, leading to further depletion of moisture, favouring the initiation of heatwaves, and contributing to their duration, intensity and propagation (Miralles et al., 2019).

Consequently, changes in the large-scale atmospheric circulation or the land-surface properties will have an effect on the intensity and frequency of temperature extremes. Changes in equator-pole temperature gradient due to anthropogenic radiative forcing can alter the large-scale circulation and favour the occurrence of temperature extremes (Coumou et al., 2015; Mann et al., 2017). Moreover, in a global warming scenario, the land relative humidity is projected to decrease (Byrne and O'Gorman, 2016), further enhancing the effect.

As temperature extremes are more sensitive to the variance of temperature as compared to the mean temperature (and more so the rarer the event is) (Katz and Brown, 1992), the probability distribution function (PDF) of temperature around the mean is essential to study temperature extremes. With horizontal advection being the primary mechanism for temperature variability at a synoptic scale, reduction in the equator-pole temperature gradient can reduce variability and influence extremes (Schneider et al., 2015). In addition to the mean and variance, higher moments like the skewness of the distribution can also have serious implications in a warming scenario, with short warm tailed distributions being much more susceptible to extremes than analogous Gaussian distributions under a simple uniform warming across the PDF (Loikith, Neelin, et al., 2018).

In this thesis, we wish to study separately the effect of atmospheric and surface dynamics on the PDFs of near-surface temperature distribution, especially the tails. We achieve this by using an idealised Atmospheric General Circulation Model (AGCM), similar to Frierson et al., 2006. The radiation in the model is independent of air humidity, removing the radiation-water vapour feedback from the model. We study different large-scale circulation scenarios

by varying the optical depth profile of the atmosphere; To model changes in near-surface relative humidity, we vary the ratio of outgoing sensible heat flux to outgoing latent heat flux (Bowen ratio) at the surface via scaling of the relative humidity over land.

In this study, we quantify the tails of the near-surface temperature PDFs in terms of the return time associated with these extremes. As heatwaves are characterised by large anomalies over a few weeks and sometimes as long as a month (IMD, 2022), we specially focus on the return times of time averages of near surface temperature. The return time of an event is the average time between two successive occurrences of the event, and is a very useful metric, which has been utilised in other studies to quantify the risk associated with extreme events (Kumari et al., 2019; Christidis et al., 2015). To compute return times reliably, we need a good sampling of the tail.

As extreme temperature events occur rarely, sampling them directly from climate model integrations is computationally costly. Therefore, we implement and utilise a rare event algorithm (Lestang et al., 2018) that selectively samples from the tail of the time averaged temperature distribution by performing a biased sampling. This algorithm was used in a previous study on European heatwaves (Ragone, Wouters, et al., 2018) to sample extreme temperature events, and compute their return times at a fraction of the computational cost of a direct run. In addition to efficiently calculating return times, the algorithm also provides a biased ensemble of trajectories that are real solutions of the model. These trajectories provide us with a large sample of heatwaves, which can be used subsequently to analyse the dynamics of these heat waves.

Contributions of the thesis

- The Frierson boundary layer scheme was implemented and validated on CliMT.

- A baseline Atmospheric General Circulation Model (AGCM) was designed using CliMT. The baseline configuration was run for a duration of 1000 years.

- The GKLT algorithm was implemented and validated on the baseline configuration. The algorithm was validated by comparing with the 1000 year long run.

- Modifying the baseline configuration, we designed four other configurations - Drier land, Wetter land, 2 degrees warming and 4 degrees warming configurations.

- Using the algorithm, we compared the return times of near-surface temperature extremes for the different configurations.

- From the data obtained from the algorithm, a statistical analysis was performed. We examined how the heatwave properties differ between configurations.

Chapter 2

Methods

2.1 Model description

For this project, we use an idealised AGCM created using CliMT (Monteiro et al., 2018), a climate modelling toolkit. CliMT's modular design and flexibility enables us to easily and quickly work with different components and parameterisation.

The AGCM configuration is very similar to the model setup in Frierson et al., 2006; Differences in components used and parameterisations are highlighted below.

The model is an aquaplanet, with zonally symmetrical strips of land between 20° and 40° in both hemispheres. The surface is an idealised slab, with prescribed depth and heat capacity value. The area type can be either land or sea, with different depth and heat capacity values for sea and land grids. The prescribed depth value for land is 1 m, with a heat capacity value of 2000 $Jkg^{-1}K^{-1}$. For the ocean, the prescribed depth value is 2 m, with the heat capacity being the heat capacity of water, 4182 $Jkg^{-1}K^{-1}$. The temperature of the surface is controlled dynamically by the energy balance at the surface. There is no topography in our model.

For the atmosphere, the grey radiation scheme is used, with values for atmosphere opacity τ prescribed as follows.

$$\tau = \tau_0 \left(\frac{p}{p_s} \right) \tag{2.1}$$

$$\tau_0 = \tau_{0e} + (\tau_{0p} - \tau_{0e}) \sin^2(\phi) \tag{2.2}$$

where $\tau_{0p} = 1.5$ and $\tau_{0e} = 6$ are values of atmosphere opacity at the surface at the pole and equator, respectively. p and p_s are pressure and surface pressure, respectively. ϕ is the latitude.

For stability reasons, the vertical profile of optical depth with pressure in our model is purely linear, unlike in Frierson et al., 2006, where a combination of linear and quartic terms were used. The incoming solar flux values R_s are prescribed as in Frierson et al., 2006, but for May conditions, with maximum flux at 10°N, and decreasing towards the poles with functional form

$$R_s(\phi) = \begin{cases} R_{max}[1 + \Delta_s p_2(\phi - 10°)] & \text{for } \phi \geq -80° \\ R_{max}[1 + \Delta_s p_2(-90°)] & otherwise, \end{cases} \tag{2.3}$$

where

$$p_2(\theta) = \frac{1}{4}[1 - 3\sin^2\theta] \tag{2.4}$$

is the second Legendre polynomial and ϕ is the latitude. $R_{max} = 150$ Wm^{-2} and $\Delta_s = 1.4$ controls the meridional gradient of solar flux.

The model has no seasonality and no diurnal cycle. Moist convection in the model is parameterised using the Emanuel convection scheme (Emanuel and Živković-Rothman, 1999).

The surface flux and boundary layer formulation of Frierson et al., 2006 was implemented with some modification (see next subsection) as a component in CliMT for this project. The relative humidity at the surface is set to the saturation specific humidity at the surface temperature scaled by a parameter $S_L = 0.5$. This parameter models the moisture limitation

over land and controls the Bowen ratio.

These components are run along with a spectral dynamical core. The dynamical core performs the integration of the primitive equations that model the global atmospheric flow. The model grid resolution is 62 x 128 (2.76° x 2.79°), which equates to a grid length of approximately 310 km at the equator. The model has 28 height levels and an integration time step of 20 minutes. As we are interested in the large scale dynamics of rare events, this model setup is a good trade-off between the model resolution and the computational cost needed for rare event sampling.

2.1.1 Modification to the Frierson boundary layer scheme

Frierson et al., 2006 calculates the diffusion coefficients at height z as-

$$K(z) = \begin{cases} K_b(z) & \text{for } z < f_b h \\ K_b(f_b h)\frac{z}{f_b h}\left[1 - \frac{z - f_b h}{(1 - f_b)h}\right]^2 & \text{for } f_b h < z < h, \end{cases} \tag{2.5}$$

where h is the boundary layer height, $f_b = 0.1$ is the surface layer fraction, and K_b is the surface layer diffusion coefficient given by

$$K_b(z) = \begin{cases} \kappa u_a \sqrt{C} z & \text{for } Ri_a < 0 \\ \kappa u_a \sqrt{C} z\left[1 + \frac{Ri}{Ri_c}\frac{lnz/z_0}{(1 - Ri/Ri_c)}\right]^{-1} & \text{for } Ri_a > 0, \end{cases} \tag{2.6}$$

and Ri defined as

$$Ri(z) = \frac{gz[\Theta_v(z) - \Theta_v(z_a)]/\Theta_v(z_a)}{|v(z)|^2} \tag{2.7}$$

However, K_b in Eqn.(2.6) is not consistent at $Ri_a = 0$ as the value of K_b is discontinuous at $Ri_a = 0$. To solve this, we have modified the multiplier in the second case as

$$\left[1 + \frac{Ri_a}{Ri_c}\frac{lnz/z_0}{(1 - Ri_a/Ri_c)}\right]^{-1} \text{ for } Ri_a > 0 \qquad (2.8)$$
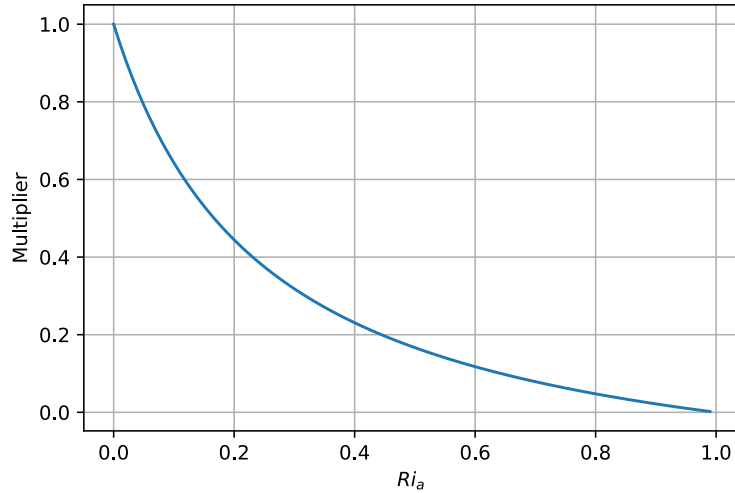
Figure 2.1: The modified multiplier with $Ri_c = 1$ and $ln(z/z_0) = 5$

The modified multiplier is consistent at $Ri_a = 0$, and decreases from 1 to 0 as Ri_a approaches Ri_c. As the surface layer diffusion coefficient Ri_a indicates stability close to the surface, this behaviour makes physical sense. The diffusion coefficient at the surface is small when the near-surface atmospheric temperature structure is stable, and is zero beyond a threshold of stability, given by Ri_c.

2.1.2 Column test for boundary layer implementation

To test the Frierson boundary layer implementation, a column test was run, along with a slab surface, RRTMG radiation component and the Emanuel convection component.

The atmospheric temperature and potential temperature profiles after spinup are as below.
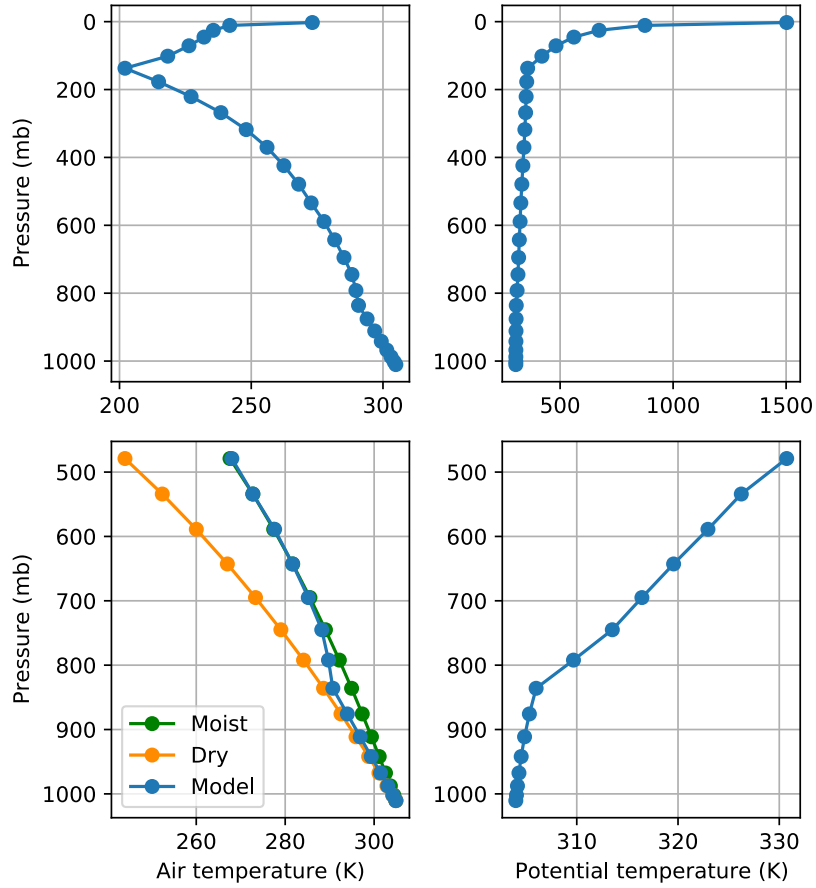
Figure 2.2: Top left - The air temperature profile. Bottom left - Temperature profile of model compared against dry and moist adiabatic lapse rate profiles. Top right - Potential temperature profile. Bottom right - Potential temperature profile close to the boundary layer.

The profiles in Fig. 2.2 highlights the characteristic features of the boundary layer. From the bottom right panel, we see that the potential temperature gradient within the boundary layer is close to zero, with the region being slightly more stable than neutral. The bottom left panel affirms that the temperature profile in the boundary layer region is close to a dry adiabatic profile.

2.2 The observable

In this study, we consider a fixed rectangular region of land between 20°N and 40°N latitude, with a longitudinal extent of 30° (shown below). As we are particularly interested in the large-scale dynamics of temperature extremes, the lengths of the region of interest are taken to be of order of the synoptic scale (more than 1000 km). The position of the land is chosen such that the land is close to the meridional location of the jet stream and the storm track. This allows us to capture and study the dynamical effects associated with the the jet stream and the storm track, along with other processes like large-scale advection (Schneider et al., 2015; Garfinkel and Harnik, 2017). As the model is zonally symmetric, the longitudinal position of the land is arbitrary. We arbitrarily set the longitudinal extent of the region of interest to be from 100°E and 130°E.
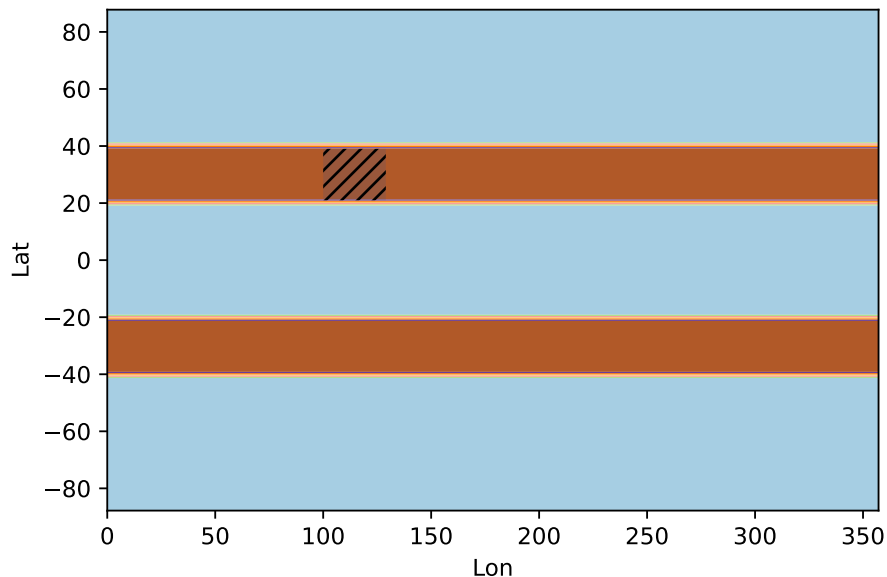
Figure 2.3: Surface configuration; Blue and brown regions are sea and land grids, respectively. The hatched region is the piece of land that is under study.

The spatial average of near-surface temperature over the piece of land $T_L(t)$ is a good indicator of large-scale temperature extremes within the region of interest. The PDF and autocorrelation of 6 hourly values of $T_L(t)$ computed from a long run (1000 years) of the model described in section 2.1 are given below. The autocorrelation plot was drawn using the python statsmodels package.
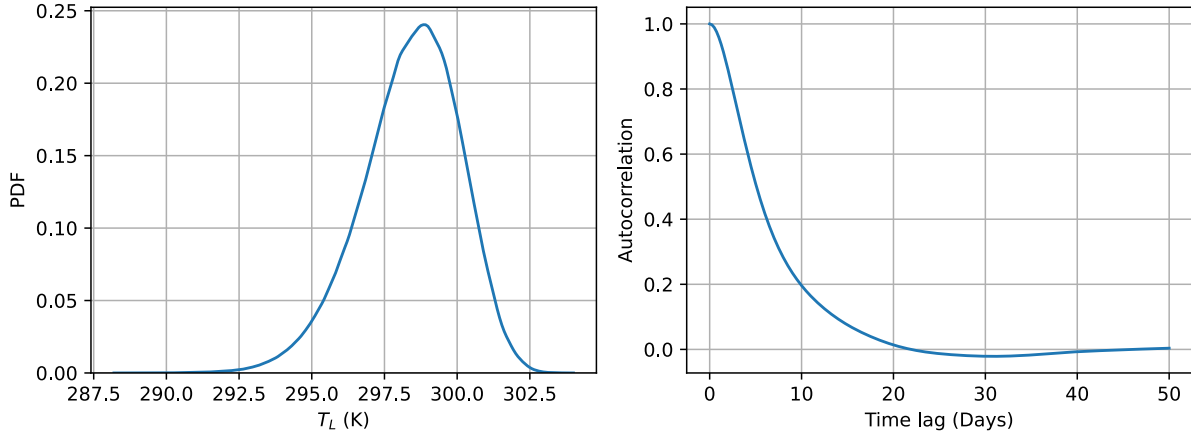
Figure 2.4: Left - PDF of T_L. Right - Autocorrelation function of T_L.

The mean of $T_L(t)$ $\mu \approx 298.38$ K and the standard deviation $\sigma \approx 1.69$ K. The PDF is negatively skewed (skewness $\approx$ -0.47), with a longer tail for values below the mean. With our region of interest being equatorward of the zonal jet stream, the negative skewness observed is consistent with previous studies that have shown that the skewness of temperature PDFs are associated with the relative location from the jet stream and the storm track, with the PDFs being positively skewed on the poleward side of the jet and negatively skewed on the equatorward side (Garfinkel and Harnik, 2017; Loikith, Waliser, et al., 2015).

The autocorrelation time τ_c of $T_L(t)$ has been estimated using the methodology detailed in Ragone and Bouchet, 2020 as $\tau_c \approx 6.7$ days. As we are studying temperature extremes, we also compute the anomaly timeseries $A_L(t)$, where $A_L(t) = T_L(t) - \mu$. As $A_L(t)$ and T_L are same up to a translation of μ, both $A_L(t)$ and T_L have been used in subsequent analysis, sometimes interchangeably.

2.3 The estimator

2.3.1 Block maximum estimator

To compute the tails and the return times of extreme events in our observable, we utilise a block maximum estimator (detailed in Lestang et al., 2018). The estimator is as follows.

17

If we have a timeseries $\{A(t)\}_{0 \leq t \leq T_d}$, we divide it into M blocks of length $\Delta T \gg \tau_c$, where τ_c is the autocorrelation time, such that $T_d = M\Delta T$.

For $1 \leq m \leq M$, we define block maximum a_m to be

$$a_m = max\{A(t)|(m-1)\Delta T \leq t \leq m\Delta T\} \tag{2.9}$$

We assume that the distribution of extreme events follow a Poisson distribution with rate $\lambda = 1/r(a)$, where $r(a)$ is the return time of events with threshold a, with the assumption that occurrences of extreme events are random and independent of each other. Hence, within the limits $\tau_c \ll \Delta T \ll r(a)$, , the probability $q_m(a)$ that a_m is larger than some threshold a is well approximated by

$$q_m(a) \simeq \Delta T/r(a) \tag{2.10}$$

$q_m(a)$ for rare event threshold a can be alternatively computed as $\frac{1}{M}\sum_{m=1}^{M} s_m(a)$, where

$$s_m(a) = \begin{cases} 1 & \text{for } a_m > a \\ 0 & \text{otherwise,} \end{cases} \tag{2.11}$$

Hence, we can estimate $r(a)$ using the block maximum estimator

$$r_B(a) = \frac{T_d}{\sum_{m=1}^{M} s_m(a)} \tag{2.12}$$

2.3.2 Modified block maximum estimator

The block maximum estimator is valid only in the limit $\Delta T \ll r(a)$. A better estimator of $r(a)$ is the modified block maximum estimator, valid even when $\Delta T/r(a)$ is of order 1. It is obtained by taking $q_m(a) = 1 - e^{-\Delta T/r(a)}$, with the estimator being

$$r'_B(a) = \frac{-\Delta T}{ln(1 - \frac{1}{M} \sum_{m=1}^{M} s_m(a))} \tag{2.13}$$

The modified block maximum estimator lets us compute return time plots very efficiently, while being equivalent to other direct methods of return time computation (Lestang et al., 2018).

Derivation of the modified block estimator

In this subsection, we show how the modified block maximum estimator (Eqn. 2.13) can be derived from our assumption that the distribution of extreme events follow a Poisson distribution. We also show that the block maximum estimator (Eqn. 2.12) is a special case of the modified block maximum estimator, with an additional constraint imposed.

Consider an ensemble of M trajectories, or equivalently, M blocks. Let E be the number of blocks that satisfy the extreme event condition $a_m > a$, where a_m is the block maximum and a is some threshold. $N = M - E$ is the number of trajectories that do not satisfy the above extreme event condition.

When the block lengths are 0, $E = 0$ and $N = M$. If we then simulate the blocks for a small time dt, assuming that that the distribution of extreme events follow a Poisson distribution with rate $\lambda = 1/r(a)$, where $r(a)$ is the return time of events with threshold a, the change in E can be expressed as

$$dE = \frac{Ndt}{r(a)} = \frac{(M - E)dt}{r(a)} \tag{2.14}$$

rearranging, we have

$$\frac{dE}{(M - E)} = \frac{dt}{r(a)} \tag{2.15}$$

integrating Eqn.(2.15), we get

19

$$-ln(M - E)\Big|_0^{E'} = \frac{t}{r(a)}\Big|_0^{T'} \tag{2.16}$$

rearranging, we get

$$\frac{E'}{M} = q(a) = 1 - e^{-T'/r(a)} \tag{2.17}$$

where T' is the total block length and E' is the number of blocks with the rare event.

Eqn.(2.17) is the modified block maximum estimator relation, which is general and valid even when $T'/r(a)$ is of order 1.

With an additional constraint $T' \ll r(a)$, Eqn.(2.17) can be simplified to

$$\frac{E'}{M} = q(a) = T'/r(a) \tag{2.18}$$

giving us the block maximum estimator relation, which is valid only in the limit $T' \ll r(a)$.

2.4 The GKLT algorithm

Climate model integration is computationally costly, making model runs longer than a 1000 years infeasible. This restricts our sampling to events with return times of around a 1000 years.

To overcome this barrier, we utilise rare event sampling algorithms. We use the GKLT algorithm in particular, which has been used in a previous study to sample temperature extremes in a climate model (Ragone, Wouters, et al., 2018). The GKLT algorithm selectively samples from the tails of the observable of interest, increasing the efficiency of extreme event sampling by a few orders of magnitude. The algorithm also provides an ensemble of biased trajectories where the extreme events are a common occurrence, which can be used for subsequent dynamical analysis.

The GKLT algorithm is specifically designed to study time averaged observables. Therefore, we focus on the time averaged observable a_L in the subsequent analysis, where

$$a_L = \frac{1}{T} \int_0^T A_L(t)\, dt \tag{2.19}$$

where $A_L(t)$ is the anomaly timeseries as described in section 2.2.

The estimator assumes that the observable a_L are independent and identically distributed (iid) variables (for an in-depth treatment, see Rohwer et al., 2015). For a correlated process, the averaging time should atleast of order $10\tau_c$ for this assumption to be reasonably valid, where τ_c is the autocorrelation time of the process. Thus, the lower limit for the averaging time T that can be used is set by the autocorrelation time of the process. For instance, in Lestang et al., 2018, an averaging time of 10 was used for the GKLT implementation on an Ornstein–Uhlenbeck process with an autocorrelation time of 1. As the autocorrelation time of our observable $T_L(t)$ is around 6.7 days, the minimum averaging time that can be used is around 60 days.

For this study, we use an averaging time $T = 90$ days, as in Ragone, Wouters, et al., 2018, which is the extent of a season. In this case, the algorithm selectively samples for large positive deviations in temperature over the duration of a season. The algorithm is insensitive to the particular value of T, provided that it is greater than 60 days.

A brief description of the algorithm is given below, more details can be found at Ragone, Wouters, et al., 2018 and Ragone and Bouchet, 2020.

2.4.1 Algorithm description

We begin by defining $P_0(\{X(t)\}_{0 \leq t \leq T_a} = \{x(t)\}_{0 \leq t \leq T_a})$ as the probability that our trajectory $X(t)$ is some $x(t)$ in our unbiased model. We simulate N model trajectories, with the starting initial conditions sampled randomly and uniformly from P_0.

The trajectories are simulated for a total duration of T_a, with resampling done at intervals of τ. Thus, resampling is done i times, with $i = 1, .., T_a/\tau$. At resampling time $t_i = i\tau$, we stop the simulation and assign to each trajectory n the weight

$$W = \frac{e^{k \int_{t_{i-1}}^{t_i} A(X_n(t))dt}}{R_i} \tag{2.20}$$

where

$$R_i = \frac{1}{N} \sum_{n=1}^{N} e^{k \int_{t_{i-1}}^{t_i} A(X_n(t))dt} \underset{N \to \infty}{\sim} E_0[e^{k \int_{t_{i-1}}^{t_i} A(X_n(t))dt}] \tag{2.21}$$

$A(X_n(t))$ is the observable of interest, the spatial average of near-surface temperature for the nth trajectory in the ensemble. k is a parameter that can be tuned to control the strength of the biasing (selection parameter).

E_0 denotes an average over P_0. The error associated with evaluating averages over observables goes as $1/\sqrt{N}$ (Ragone and Bouchet, 2020), with the relation being true asymptotically for large N.

The trajectories are sampled proportional to their weights such that the trajectories with higher weights are kept and cloned and the trajectories with lower weights are killed off. The clones are slightly perturbed, so that they evolve differently from the original trajectory. The perturbation is introduced in the coefficients of the spherical harmonics of the logarithm of the surface pressure, with the perturbation values sampled uniformly from $[-\epsilon\sqrt{2}, \epsilon\sqrt{2}]$, where $\epsilon = 10^{-4}$.

After the first resampling, we obtain trajectories sampled according to a new PDF P_1 with

$$P_1(X(t)) = P_0(X(t)) \frac{e^{k \int_0^\tau A(X(t))dt}}{R_1} \tag{2.22}$$

Similarly, after the second resampling time, the trajectories are sampled according to PDF P_2 where

$$P_2(X(t)) = P_0(X(t)) \frac{e^{k \int_0^\tau A(X(t))dt} e^{k \int_\tau^{2\tau} A(X(t))dt}}{R_1 R_2} = P_0(X(t)) \frac{e^{k \int_0^{2\tau} A(X(t))dt}}{R_1 R_2} \tag{2.23}$$

By successively biasing the PDF at each step, after $m = T_a/\tau$ resampling steps, we have PDF P_m as

$$P_m(X(t)) = P_0(X(t)) \frac{e^{k \int_0^{T_a} A(X(t))dt}}{\prod_{i=1}^{T_a/\tau} R_i} \tag{2.24}$$

$P_m(X(t))$ for any trajectory $X(t)$ in the final ensemble is $1/N$ (one among N trajectories). We invert Eqn.(2.24) to get the probability of the final trajectories in the original distribution P_0 as

$$P_0(X(t)) = \frac{1}{N} e^{-k \int_0^{T_a} A(X(t))dt} \prod_{i=1}^{T_a/\tau} R_i \tag{2.25}$$

Thus, we obtain N trajectories, with corresponding probabilities p_n $(1 \leq n \leq N)$ in P_0. For each trajectory, we compute a_n $(1 \leq n \leq N)$, the maximum of the time averaged observable a_L with an averaging time $T = 90$ days.

$$a_n = max(\{a_L\}_{0 \leq t \leq T_a - T}) \tag{2.26}$$

The modified block maximum estimator (2.13) can be generalised for trajectories with non-equal probabilities as

$$r(a) = -\frac{T_a - T}{ln(1 - \sum_{n=1}^N p_n s_n(a))} \tag{2.27}$$

with

$$s_n(a) = \begin{cases} 1 & \text{for } a_n > a \\ 0 & \text{otherwise} \end{cases} \tag{2.28}$$

23

2.4.2 Algorithm implementation

The algorithm configuration and implementation closely follows Ragone, Wouters, et al., 2018, where specific details can be found. The implementation of the algorithm is as follows.

We simulate a total of $N = 512$ trajectories for an algorithm run. The initial conditions for the trajectories are samples from a pool of around 1000 model states obtained from a long run. The states were sampled 10 days apart, which ensures that they are statistically independent. To prepare the initial conditions, the model was run for a duration of 30 years, which is negligible compared to the return times that can be calculated from the algorithm.

The algorithm is insensitive to the the particular value of the resampling time τ, provided τ is sufficiently larger than the numerical timestep and lower than the Lyapunov time of the process. The Lyapunov time of a process the time after which the process loses memory of its initial state. We can estimate the Lyapunov time for our process by simulating processes with nearby initial conditions (see plot below). The average distance between the trajectories increases rapidly and then saturates. At saturation, the different runs are completely uncorrelated from each other. The above limit on τ ensure that the clones diverge from the original trajectory, while still being in the neighbourhood of the original trajectory.
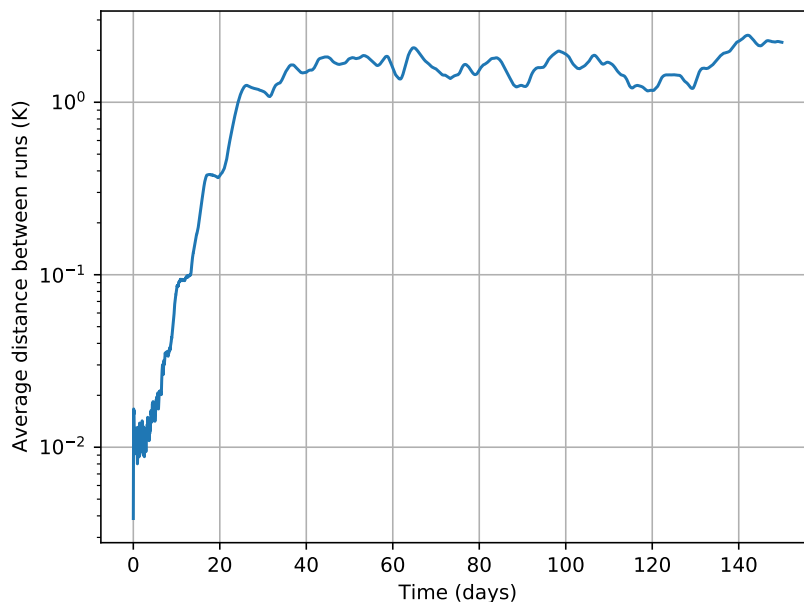

Figure 2.5: The time-evolution of average distance between the spatial average of temperature T_L from different runs (30 runs) started from nearby initial conditions.

From the plot, saturation is reached at around 30 days. Hence, we set the resampling time τ to be 8 days, close to the autocorrelation time τ_c.

The trajectory length T_a is set to 128 days, with the length of a_L, the 90-day averaged observable being 38 days. A trajectory length of 128 days ensures that the block length after time averaging is atleast a few times the autocorrelation time τ_c, while keeping the total computational cost of the algorithm low.

$A(X_n(t))$ in the weight calculation is taken to be the spatial average of near-surface temperature, as described in section 2.2. Trajectories are selected and cloned proportional to their weights. As the model is deterministic, two trajectories starting from an initial condition will evolve in the exact same way. To ensure that the trajectories diverge, we introduce a perturbation in the initial condition for the clones. The methodology for cloning and introducing perturbation in the clones is detailed in Ragone, Wouters, et al., 2018. The averaging time is taken to be $T = 90$ days.

We run the algorithm multiple times with different values of k, the selection coefficient. Return time curves are drawn as described in section 2.4.1 for $k = 10, 20, 30, 40$ and 50. For $k = 20$ and 40, the algorithm was run twice, with different initial conditions for the trajectories. The return time curves from the different runs are averaged and combined to get a best estimate return time curve. The exact methodology for combining the return time curves from different runs is detailed in Ragone, Wouters, et al., 2018.

The computational cost of a single algorithm run is around 180 years. The total computational cost across the different runs is around 1260 years, which is of the same order as that of the 1000 year long run.

2.5   Experiment details

The model configuration described in section 2.1 is run for a duration of 1000 years. This model configuration is the baseline configuration, which acts as a control for the other model configurations. We calculate the return time for temperature extremes from this long run using the modified block maximum estimator (section 2.3.2)

The GKLT algorithm is implemented for the baseline configuration. The return times

estimated from the algorithm are compared with the return times from the long run. This ensures that the algorithm is functioning as expected and agrees with the long run.

After demonstrating that the algorithm works for the baseline, we run the algorithm for modified model configurations. We are interested in the differences in return times of the modified configurations from the baseline. The modified configurations are as follows

Drier land configuration : In this configuration, only the scaling parameter that controls the relative humidity over land S_L is decreased from 0.5 to 0.2. This represents a scenario where the land is drier than in the baseline model.

Wetter land configuration : In this configuration, only the scaling parameter that controls the relative humidity over land S_L is increased from 0.5 to 0.8. This represents a scenario where the land is wetter than in the baseline model.

2 degrees warming configuration : In this configuration, the value of longwave optical depth τ_{0e} (see Eqn. 2.2) is increased from 6 to 6.6 to represent an average 2° K surface warming scenario. Other components and parameters are exactly as in the baseline.

4 degrees warming configuration : In this configuration, the value of longwave optical depth τ_{0e} is increased from 6 to 7.2 to represent an average 4° K surface warming scenario. Other components and parameters are exactly as in the baseline.

The exact version of CLiMT used for this project can be found at https://github.com/Ai33L/climt/tree/rare-event-code

The scripts for the GKLT algorithm implementation along with supporting scripts and documentation can be found at https://github.com/Ai33L/Heatwaves-GKLT

Chapter 3

Results and Discussion

3.1 Model climatology

3.1.1 Baseline configuration

The climatology of the baseline configuration is shown in Fig. 3.1. The climatology is computed as the time average of the model fields over a 30 year simulation. The spinup time of our model, the time from initialisation after which the model's climate is equilibriated, is around a year. We have spun up our models for 3 years before any analysis was performed to ensure that equilibrium was reached.

As the maximum insolation is at 10°N, the temperature maximum is in the northern hemisphere (panel a and b). The jet streams are clear in the zonally averaged u-wind plot (panel c), with a stronger jet in the winter hemisphere as expected. The jet in the northern hemisphere is at around 53°N, slighly poleward of the land, while the jet in the southern hemisphere is at around 27°S. The maximum jet speed in the northern hemisphere is around 27.6 ms^{-1}, and around 50.4 ms^{-1} the southern hemisphere. The storm track intensity, indicated by the zonal mean variance of the meridional wind (panel d) is close to the zonal jet, atleast in the northern hemisphere.

The average sensible and latent heat fluxes over land are around 19.8 Wm^{-2} and 46.5 Wm^{-2}.

Figure 3.1: Climatology of the baseline configuration. (a) Zonally averaged air temperature (K). (b) Air temperature at the lowest model level (K). (c) Zonally averaged u-winds (zonal winds) (ms^{-1}). (d) Zonal mean variance of meridional wind (m^{2}s^{-2}) - marks the location of the storm track. The zonally averaged u-wind is overlaid (black contours).

3.1.2 Drier land configuration

The climatology of the drier land configuration is given below (Fig. 3.2). The major differences from the baseline configuration are highlighted in Fig. 3.3.

The average near surface temperatures over land have increased by around 4.2 K from the baseline configuration (Fig. 3.2, panel a and b; Fig. 3.3, panel a). The northern hemisphere jet stream has shifted towards the equator by around 3° with the maximum jet speed increased by around 3.6 ms^{-1}. The maximum speed of the southern hemisphere jet has reduced by around 2.4 ms^{-1}, without any shift in latitudinal position (Fig. 3.3, panel b). The storm track intensity in the northern hemisphere has weakened by around 25 m^{2}s^{-2} and strengthened in the southern hemisphere by around 25 m^{2}s^{-2} (from Fig. 3.3, panel c).

The average sensible heat flux over land has increased by around 18.8 Wm^{-2} from the baseline, while the average latent heat flux has decreased by around 42 Wm^{-2} (not shown).
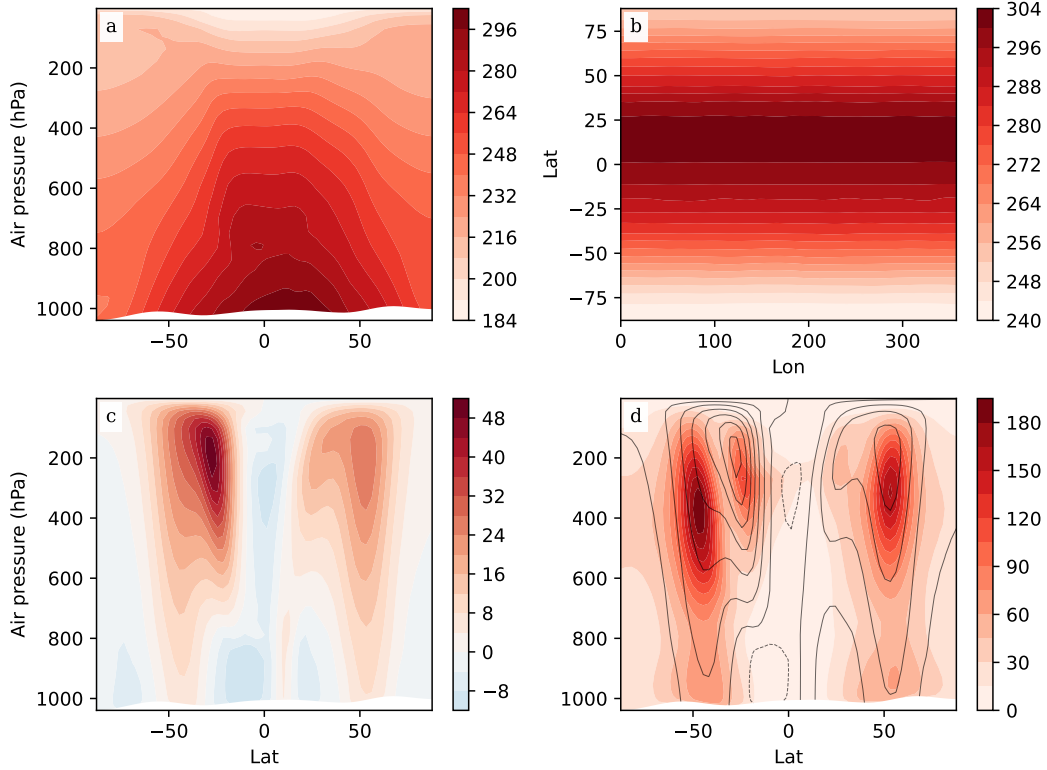

Figure 3.2: Climatology of the drier land configuration. (a) Zonally averaged air temperature (K). (b) Air temperature at the lowest model level (K). (c) Zonally averaged u-winds (zonal winds) (ms^{-1}). (d) Zonal mean variance of meridional wind (m^{2}s^{-2}) - marks the location of the storm track. The zonally averaged u-wind is overlaid (black contours).

Figure 3.3: Difference in the climatology of the drier land configuration from the baseline configuration. (a) Difference in zonally averaged temperature (K). (b) Difference in zonally averaged u-wind (ms^{-1}). (c) Difference in zonal mean variance of meridional wind (m^{2}s^{-2}). The respective baseline configuration fields are overlaid (black contours).

3.1.3 Wetter land configuration

The climatology of the wetter land configuration is given below (Fig. 3.4). The major differences from the baseline configuration are highlighted in Fig. 3.5.

The average near surface temperatures over land have decreased from the baseline configuration by around 2.1 K (clear in Fig. 3.5, panel a). The maximum jet speed in the southern hemisphere has increased by around 1 ms^{-1} from the baseline configuration, with no significant shift in its latitudinal position. (Fig. 3.5, panel b). The storm track intensity in the southern hemisphere has weakened by around 20 m^{2}s^{-2} and strengthened by around 12 m^{2}s^{-2} in the north (from Fig. 3.3, panel c).

The average sensible heat flux over land has decreased by around 12.7 Wm^{-2} from the baseline, while the average latent heat flux has increased by around 27.9 Wm^{-2} (not shown).
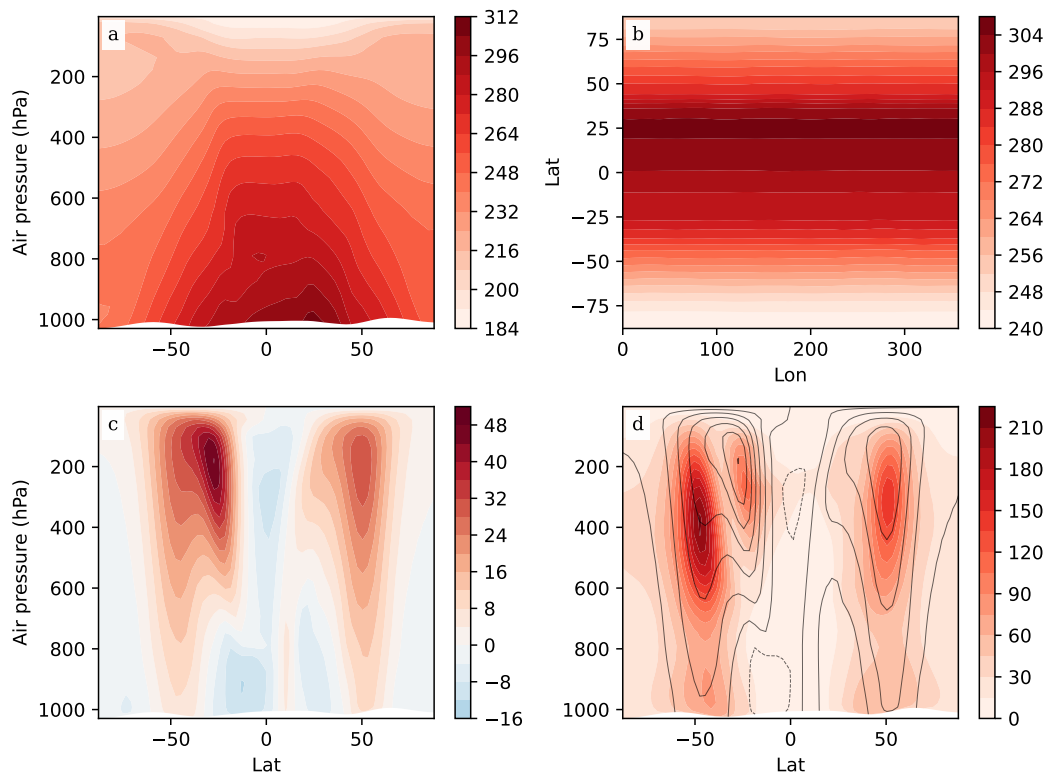
Figure 3.4: Climatology of the wetter land configuration. (a) Zonally averaged air temperature (K). (b) Air temperature at the lowest model level (K). (c) Zonally averaged u-winds (zonal winds) (ms^{-1}). (d) Zonal mean variance of meridional wind (m^{2}s^{-2}) - marks the location of the storm track. The zonally averaged u-wind is overlaid (black contours).
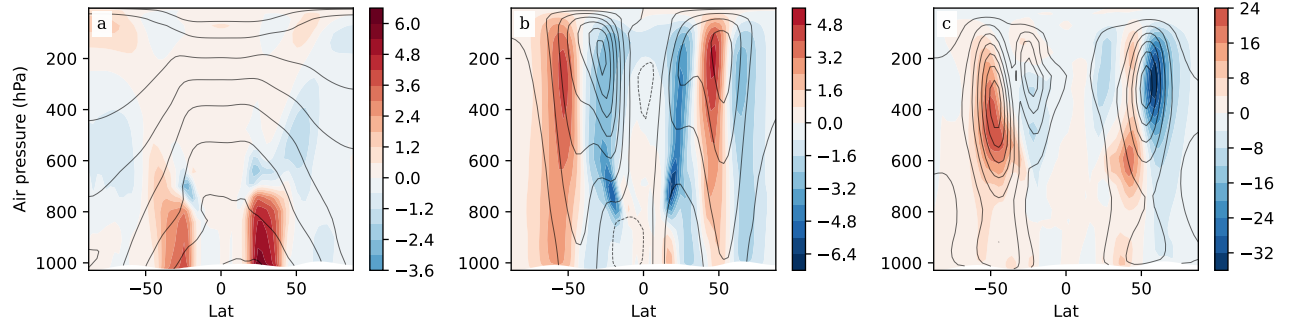

Figure 3.5: Difference in the climatology of the wetter land configuration from the baseline configuration. (a) Difference in zonally averaged temperature (K). (b) Difference in zonally averaged u-wind (ms^{-1}). (c) Difference in zonal mean variance of meridional wind (m^{2}s^{-2}). The respective baseline configuration fields are overlaid (black contours).

3.1.4   2 degrees warming configuration

The climatology of the 2 degrees warming configuration is given below (Fig. 3.6). The major differences from the baseline configuration are highlighted in Fig. 3.7.

The air temperatures have increased from baseline configuration, with the average near-surface air temperature over land increasing by around 2.5 K. Panel a of Fig. 3.7 clearly shows warming throughout the troposphere. The jet stream in the northern hemisphere has strengthened polewards, while the maximum jet speed of the southern hemisphere jet has increased by around 2.5 ms^{-1} (Fig. 3.7, panel b). The poleward flank of the storm track in the northern hemisphere has strengthened by around 12 m^{2}s^{-2} (Fig. 3.7, panel c).
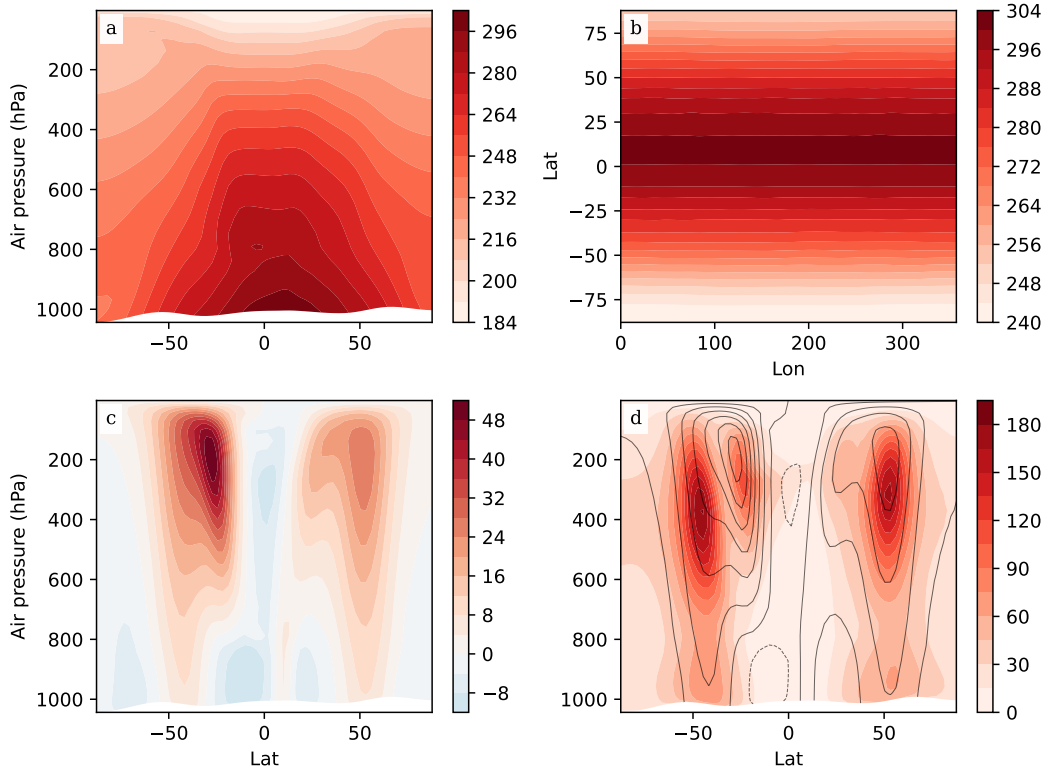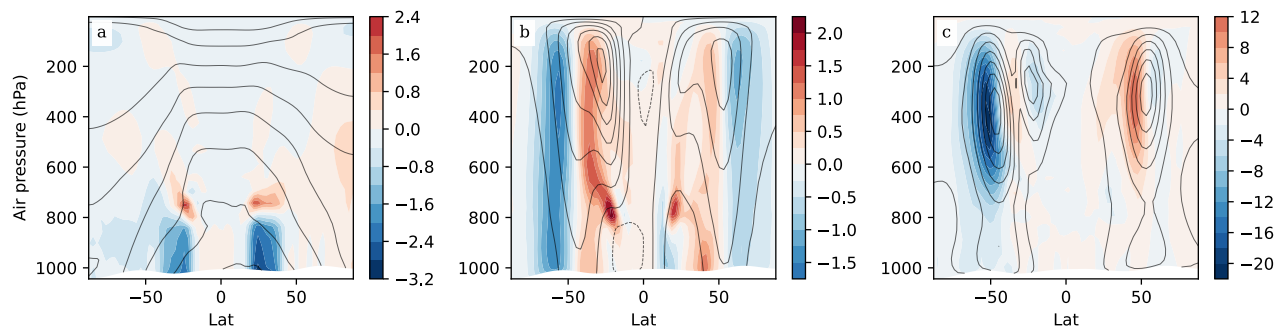

Figure 3.6: Climatology of the 2 degrees warming configuration. (a) Zonally averaged air temperature (K). (b) Air temperature at the lowest model level (K). (c) Zonally averaged u-winds (zonal winds) (ms^{-1}). (d) Zonal mean variance of meridional wind (m^{2}s^{-2}) - marks the location of the storm track. The zonally averaged u-wind is overlaid (black contours).

Figure 3.7: Difference in the climatology of the 2 degrees warming configuration from the baseline configuration. (a) Difference in zonally averaged temperature (K). (b) Difference in zonally averaged u-wind (ms^{-1}). (c) Difference in zonal mean variance of meridional wind $(m^2 s^{-2})$. The respective baseline configuration fields are overlaid (black contours).

3.1.5 4 degrees warming configuration

The climatology of the 4 degrees warming configuration is given below (Fig. 3.8). The major differences from the baseline configuration are highlighted in Fig. 3.9.

The air temperatures have increased from baseline and the 2 degrees warming configurations, with the average near-surface air temperature over land increasing by around 4.7 K from the baseline. The tropospheric warming is almost twice that in the 2 degrees warming configuration (Fig. 3.9, panel a). The northern hemisphere jet stream has strengthened polewards, while the maximum jet speed of the southern hemisphere jet has increased by around 4.7 ms^{-1} (Fig. 3.9, panel b). The southern hemisphere storm track and the poleward flank of the northern hemisphere storm track have strengthened by around 30 $m^2 s^{-2}$ and 20 $m^2 s^{-2}$ respectively (Fig. 3.9, panel c).
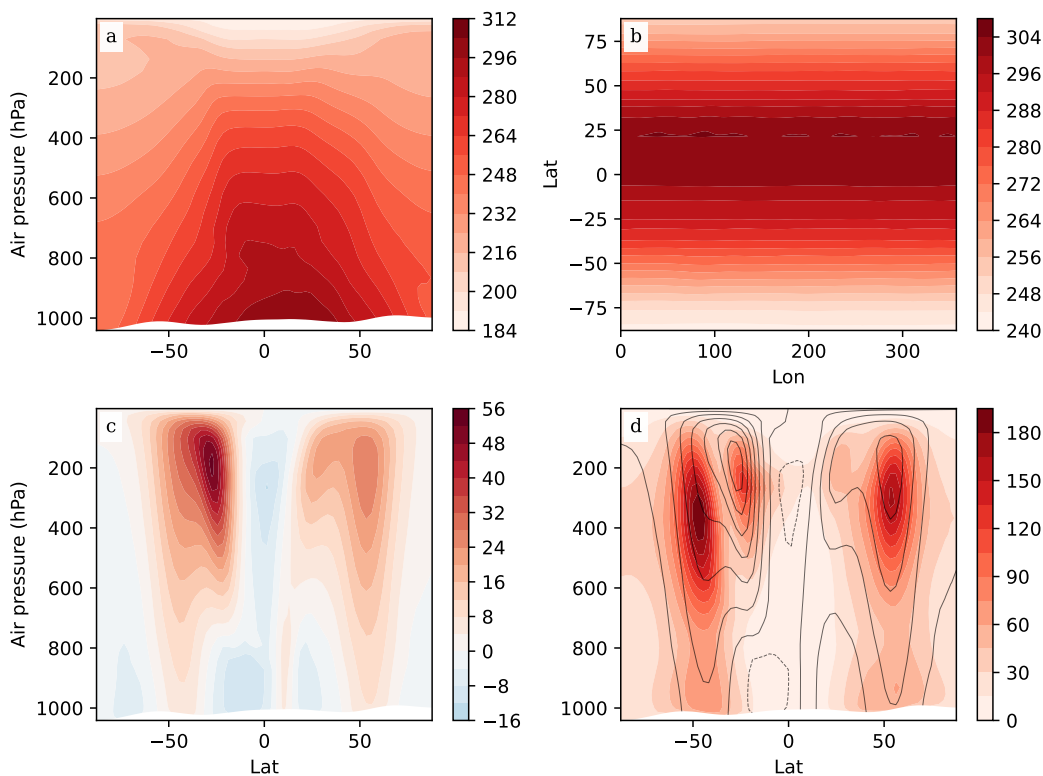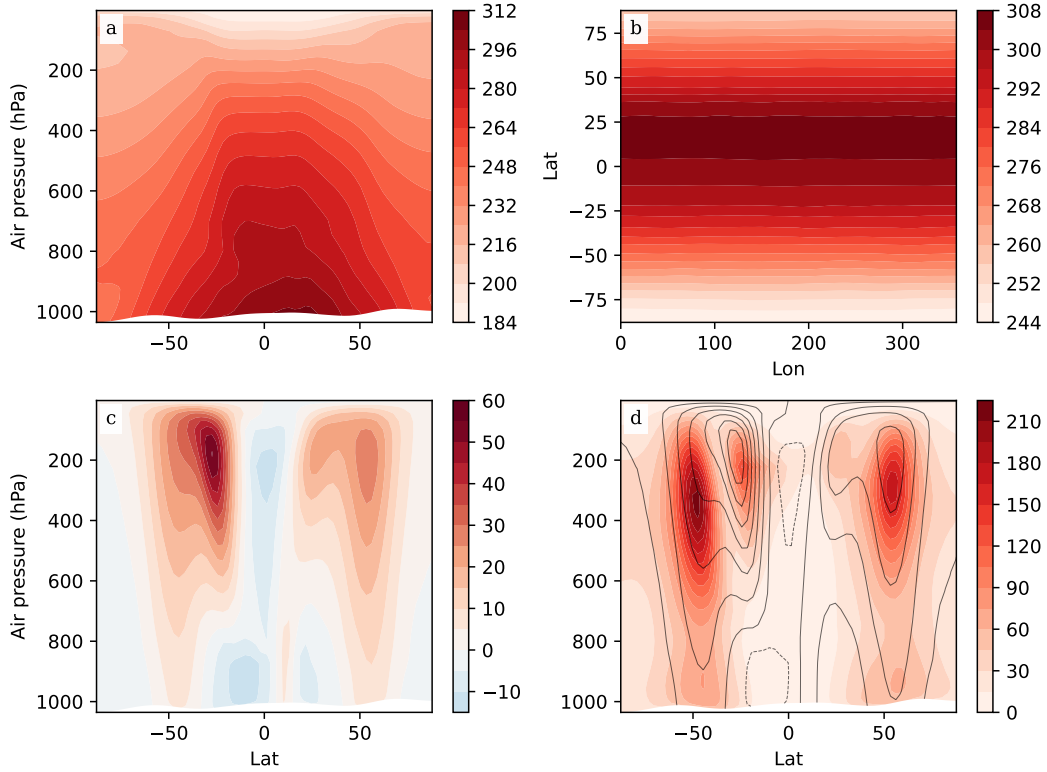
Figure 3.8: Climatology of the 4 degrees warming configuration. (a) Zonally averaged air temperature (K). (b) Air temperature at the lowest model level (K). (c) Zonally averaged u-winds (zonal winds) (ms^{-1}). (d) Zonal mean variance of meridional wind (m^{2}s^{-2}) - marks the location of the storm track. The zonally averaged u-wind is overlaid (black contours).
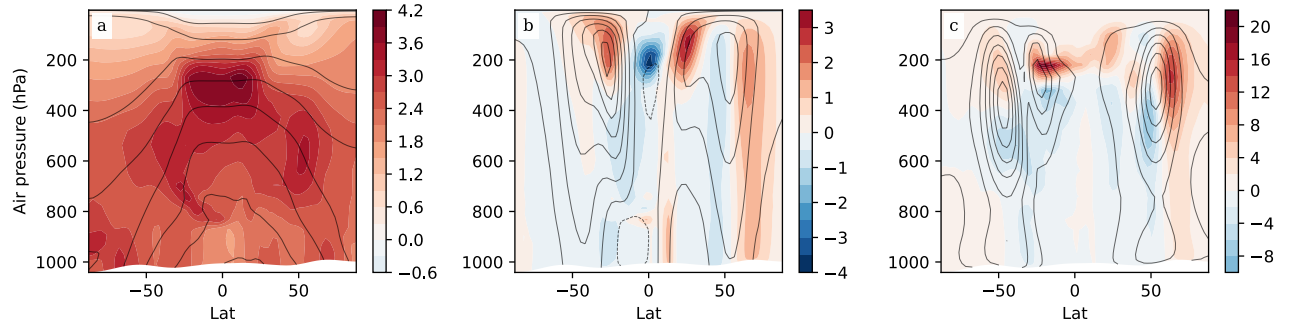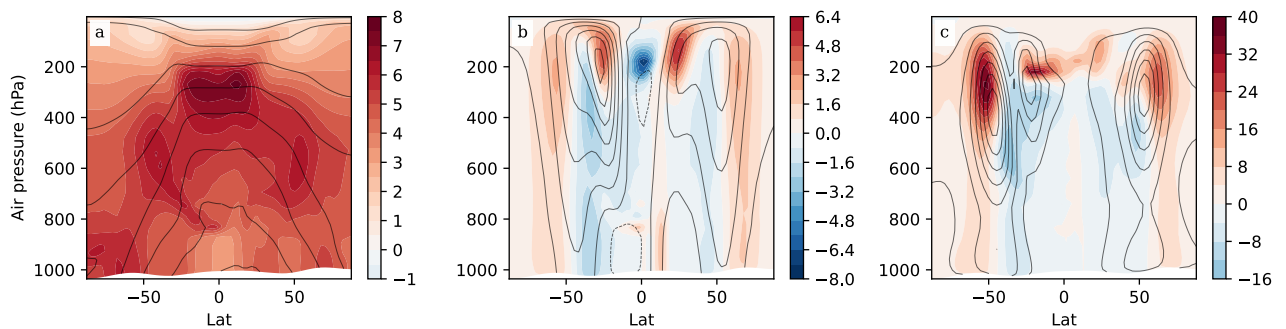

Figure 3.9: Difference in the climatology of the 4 degrees warming configuration from the baseline configuration. (a) Difference in zonally averaged temperature (K). (b) Difference in zonally averaged u-wind (ms^{-1}). (c) Difference in zonal mean variance of meridional wind (m^{2}s^{-2}). The respective baseline configuration fields are overlaid (black contours).

34

Summary of differences

| Quantity	Baseline	Drier land	Wetter land	2 degrees warming	4 degrees warming
Table of climatological quantities					
Avg. northern hemisphere land surface temperature	301.41 K	308.82 K	297.23 K	303.69 K	305.72 K
Avg. global ocean surface temperature	283.89 K	283.92 K	283.81 K	286.0 K	287.94 K
Avg. northern hemisphere near-surface temperature	298.38 K	303.50 K	295.91 K	300.80 K	302.96 K
Avg. global ocean near-surface temperature	284.16 K	284.38 K	283.98 K	286.45 K	288.55 K
North jet position and max. wind speed	53.3 N, 27.9 ms^{-1}	50.4 N, 31.6 ms^{-1}	52.0 N, 27.7 ms^{-1}	53.3 N, 27.8 ms^{-1}	53.3 N, 28.3 ms^{-1}
South jet position and max. wind speed	27.3 S, 50.8 ms^{-1}	27.3 S, 48.2 ms^{-1}	27.3 S, 51.5 ms^{-1}	27.3 S, 53.3 ms^{-1}	27.3 S, 55.4 ms^{-1}
North storm track position and max. intensity	53.3 N, 168 m^{2}s^{-2}	50.4 N, 148 m^{2}s^{-2}	53.2 N, 169 m^{2}s^{-2}	53.2 N, 168 m^{2}s^{-2}	53.2 N, 174 m^{2}s^{-2}
South storm track position and max. intensity	47.5 S, 193 m^{2}s^{-2}	47.5 S, 215 m^{2}s^{-2}	44.6 S, 181 m^{2}s^{-2}	47.5 S, 195 m^{2}s^{-2}	47.5 S, 216 m^{2}s^{-2}
Avg. sensible flux over northern hemisphere land	18.33 Wm^{-2}	39.25 Wm^{-2}	5.32 Wm^{-2}	17.16 Wm^{-2}	16.03 Wm^{-2}
Avg. latent flux over northern hemisphere land	61.92 Wm^{-2}	11.25 Wm^{-2}	93.54 Wm^{-2}	68.06 Wm^{-2}	73.66 Wm^{-2}
Avg. sensible flux over global ocean	-6.53 Wm^{-2}	-8.52 Wm^{-2}	-5.54 Wm^{-2}	-8.15 Wm^{-2}	-9.57 Wm^{-2}
Avg. latent flux over global ocean	80.31 Wm^{-2}	82.91 Wm^{-2}	79.14 Wm^{-2}	86.37 Wm^{-2}	91.59 Wm^{-2}

Table 3.1: Summary of the various climatological values for the different configurations.

The average northern hemisphere land surface temperatures decrease as the relative humidity over land is increased and increases with warming. The near-surface temperature over the northern hemisphere land follows the same pattern, with a smaller magnitude of differ-

ences. The global ocean surface and near-surface temperatures do not change significantly across configurations.

For the warming configurations, the northern hemisphere jet stream strengthens polewards as the warming is increased (Fig. 3.7 and Fig. 3.9). However, the maximum jet speed and its location does not change. The northern hemisphere jet shifts equatorwards for both the drier and wetter land configurations, with a maximum jet speed increase in the drier land configuration. The large change in the northern hemisphere jet speed and location for the drier land configuration was unintentional and is a result of the particular land configuration we have implemented, with the meridional position of the jet being at the land-ocean boundary.

3.2 Return times

The results of the experiments detailed in section 2.5 are presented in the following subsections. We use the baseline configuration to validate and explain the algorithm procedure. For the other configurations, the results are directly presented as the algorithm procedure is essentially kept the same between configurations.

Note that as we have computed return times for the anomalies from the climatology for each configuration, the effects due to changes in the mean temperatures are removed. The changes in return times captured are only due to changes in higher moments of the 90-day averaged temperature anomaly distributions.

3.2.1 Baseline configuration

The return time curves and the corresponding cumulative probability distributions for the baseline configuration are given below.

For the baseline configuration, a 1000 year long run was used as a control. The algorithm was run 7 times, with different selection coefficient values (see section 2.4.2).
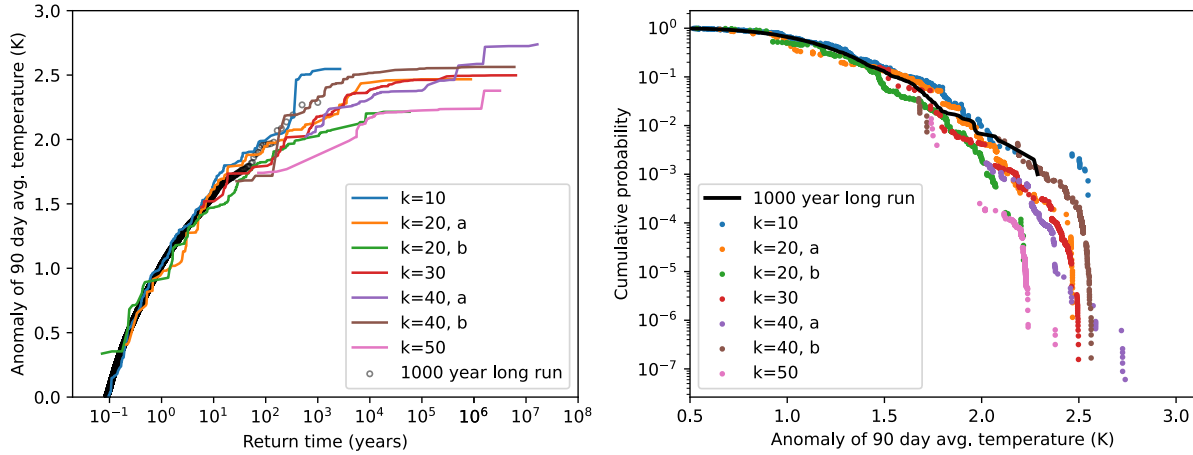
Figure 3.10: Left - Return times for the 90-day averaged temperature anomalies. Right - Cumulative probabilities for the 90-day averaged temperature anomalies. The time window taken for probability calculation is 1 year.

The return times curves computed from the different algorithm runs are trimmed and averaged as described in Ragone, Wouters, et al., 2018 to get the estimated return time curve.
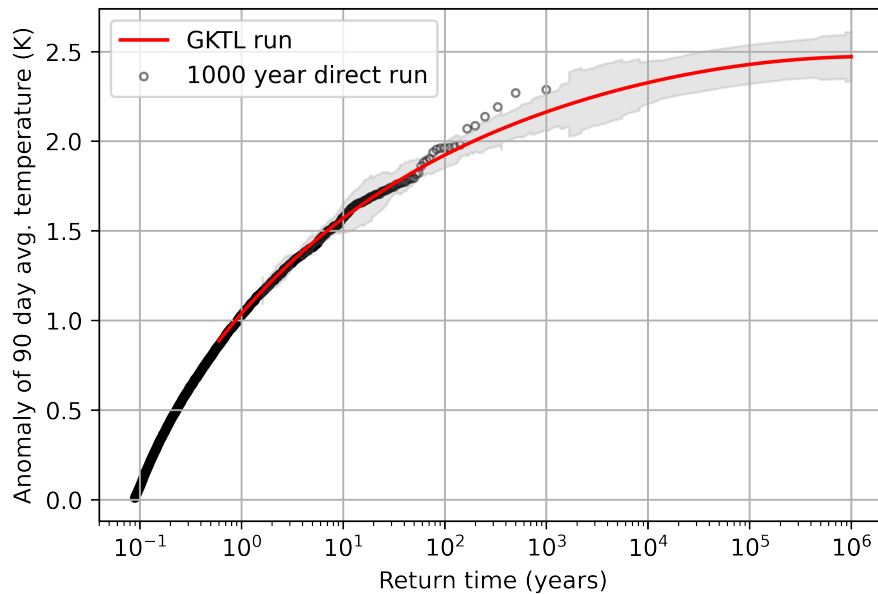

Figure 3.11: The estimated return time curve from the GKLT algorithm (red) and the return times computed from the 1000 year run (black). The grey area is the error bar, calculated as 1 standard deviation (SD) of the curves averaged at each point.

37

The algorithm correctly computes the return times for the 90-day averaged temperature anomalies with return times in the range of 10^0 to 10^3. Moreover, the algorithm lets us compute return times upto order 10^6, which was not possible with the 1000 year run. As the computational power used by the algorithm is of the order of the 1000 year direct run, the algorithm has a sampling efficiency of order 10^3.

Fig. 3.12 illustrates the basic working of the algorithm. The algorithm selectively samples around some value at the tail of the 90 day averaged temperature distribution. As the selection coefficient k is increased, the value at the tail around which the sampling is done is higher. As algorithm runs with different selection coefficients sample around different values at the distribution tail, combining multiple algorithm runs gives us a good sampling across a range of values at the tail of the distribution.
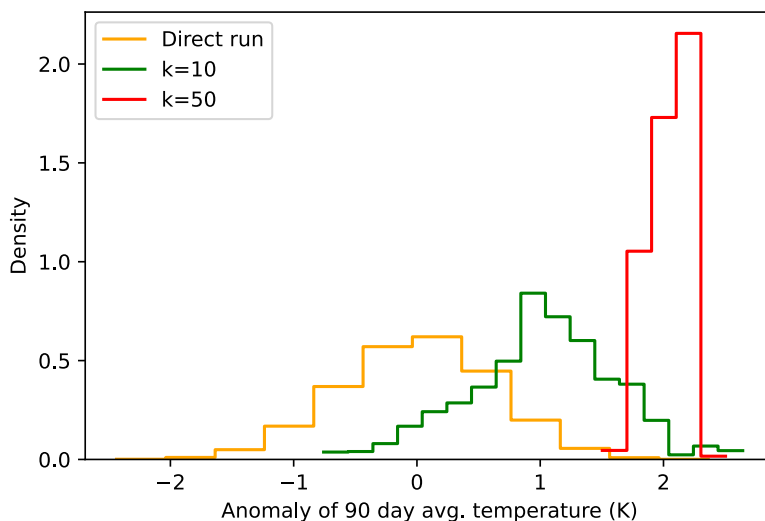

Figure 3.12: PDF of 90 day temperature anomalies computed from the direct run, compared with the PDFs computed from selected algorithm runs.
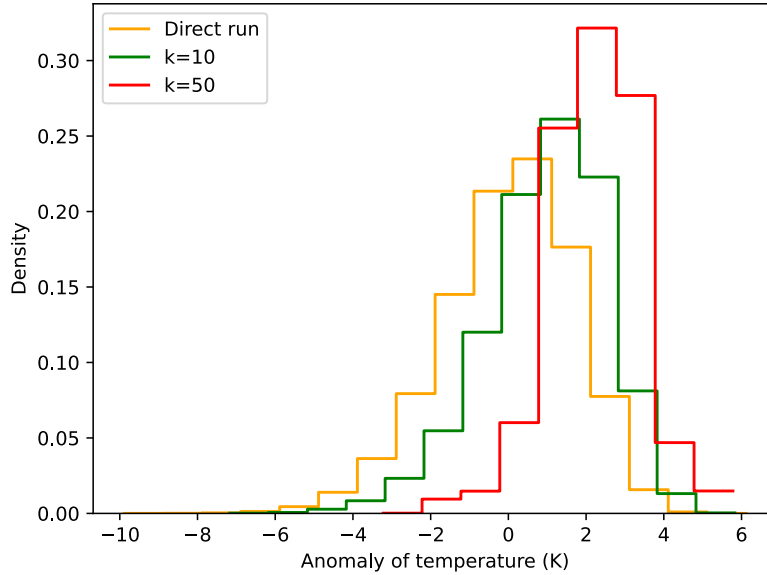
Figure 3.13: PDF of temperature anomalies (no time averaging) computed from the direct run, compared with the PDFs computed from selected algorithm runs.

However, from Fig. 3.13 we can see that even though the algorithm samples higher instantaneous temperature anomaly values, the instantaneous temperature values sampled by the algorithm and the direct run have a large overlap. From Fig. 3.12 and Fig. 3.13, it is clear that the algorithm is designed to sample a class of extreme events, with large deviations from the mean over a long period of time, 90 days in this case.

3.2.2 Drier land configuration

The return time curves and the corresponding cumulative probability distributions for the drier land configuration are given below. Note that to calculate the anomalies in this case, the climatology of drier land configuration was used.

Figure 3.14: For the drier land configuration. Left - Return times for the 90-day averaged temperature anomalies. Right - Cumulative probability of the 90-day averaged temperature anomalies.

The estimated return time curve for the 90 day averaged temperature anomalies in the drier land configuration is given below.

Figure 3.15: The estimated return time curve from the GKLT algorithm for the drier land configuration (red) and the baseline configuration (black). The grey area is the error bar, calculated as 1 standard deviation (SD) of the curves averaged at each point.

40

From the plot above, the return times of 90 day averaged temperature anomalies are lower in the drier land configuration than in the baseline configuration. This means that the probability of occurrence of these events is larger in the drier land configuration.

3.2.3   Wetter land configuration

The return time curves and the corresponding cumulative probability distributions for the wetter land configuration are given below.
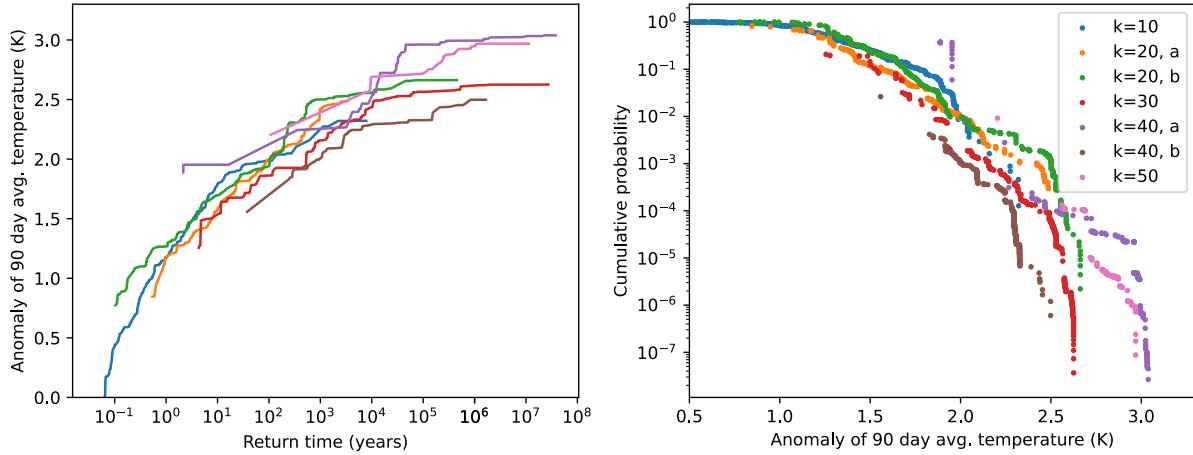

Figure 3.16: For the wetter land configuration. Left - Return times for the 90-day averaged temperature anomalies. Right - Cumulative probability of the 90-day averaged temperature anomalies.

The estimated return time curve for the 90 day averaged temperature anomalies in the wetter land configuration is given below.
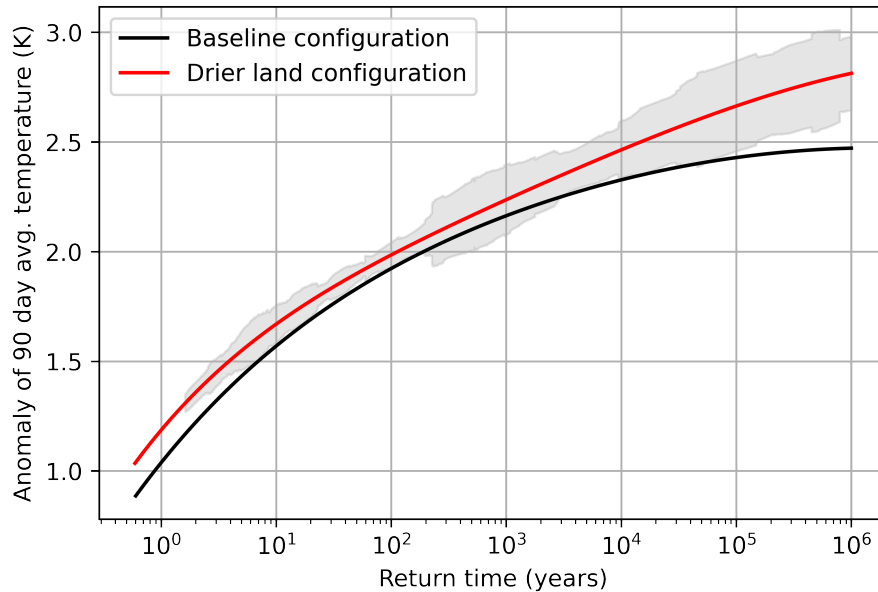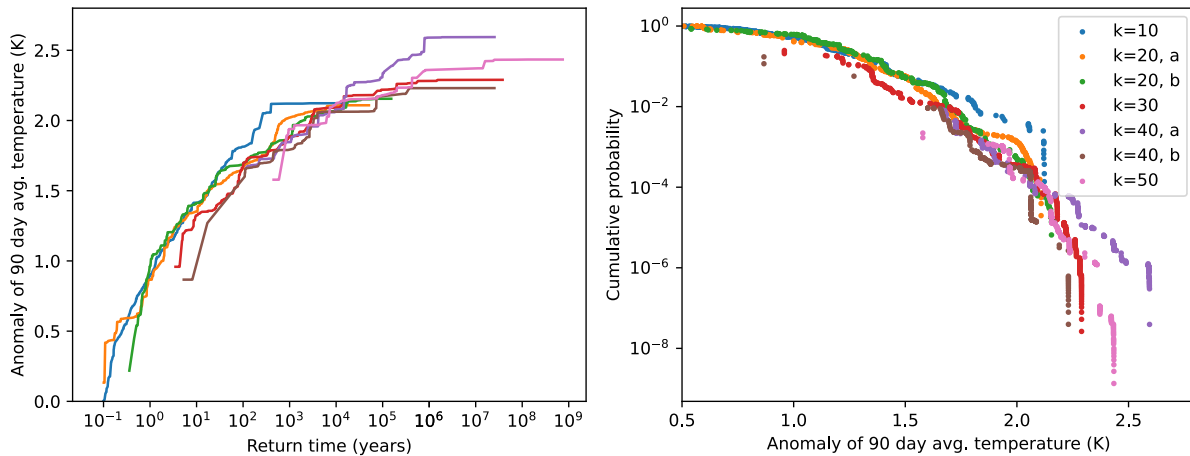
Figure 3.17: The estimated return time curve from the GKLT algorithm for the wetter land configuration (red) and the baseline configuration (black). The grey area is the error bar, calculated as 1 standard deviation (SD) of the curves averaged at each point.

The return times of 90 day averaged temperature anomalies are higher in the wetter land configuration than in the baseline configuration. This means that the probability of occurrence of these events is lower in the wetter land configuration.

3.2.4   2 degrees warming configuration

The return time curves and the corresponding cumulative probability distributions for the 2 degrees configuration are given below.

Figure 3.18: For the 2 degrees warming configuration. Left - Return times for the 90-day averaged temperature anomalies. Right - Cumulative probability of the 90-day averaged temperature anomalies.

The estimated return time curve for the 90 day averaged temperature anomalies in the 2 degrees warming configuration is given below.
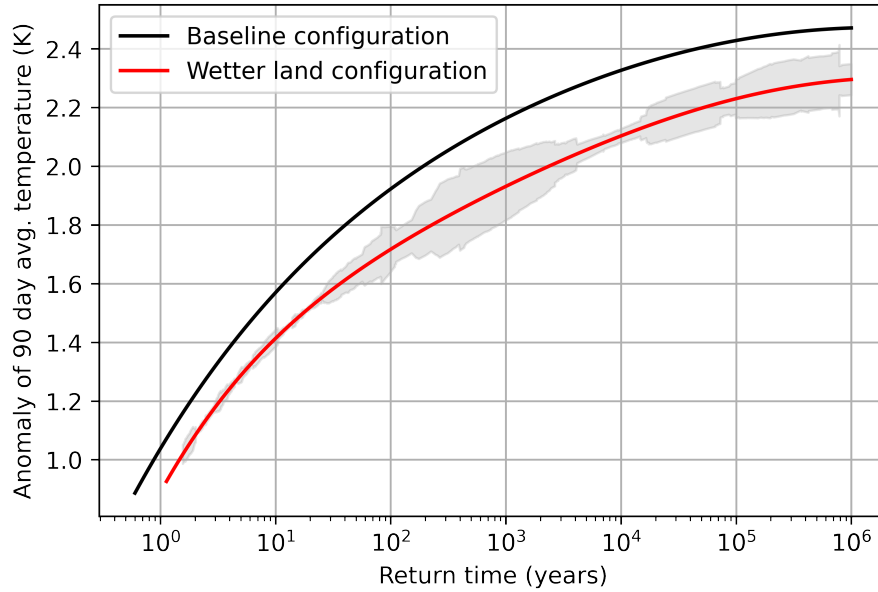

Figure 3.19: The estimated return time curve from the GKLT algorithm for the 2 degrees warming configuration (red) and the baseline configuration (black). The grey area is the error bar, calculated as 1 standard deviation (SD) of the curves averaged at each point.

From the plot above, the return times for the 90 day average temperature anomalies in the 2 degrees warming configuration are less than in the baseline configuration in the range of 10^4 to 10^6 years, and more than in the baseline configuration in the range of around 1 to 100 years. However, the baseline curve is within the error bar of the other curve, except in the range of 1 to 100 years.

3.2.5   4 degrees warming configuration

The return time curves and the corresponding cumulative probability distributions for the 4 degrees configuration are given below.
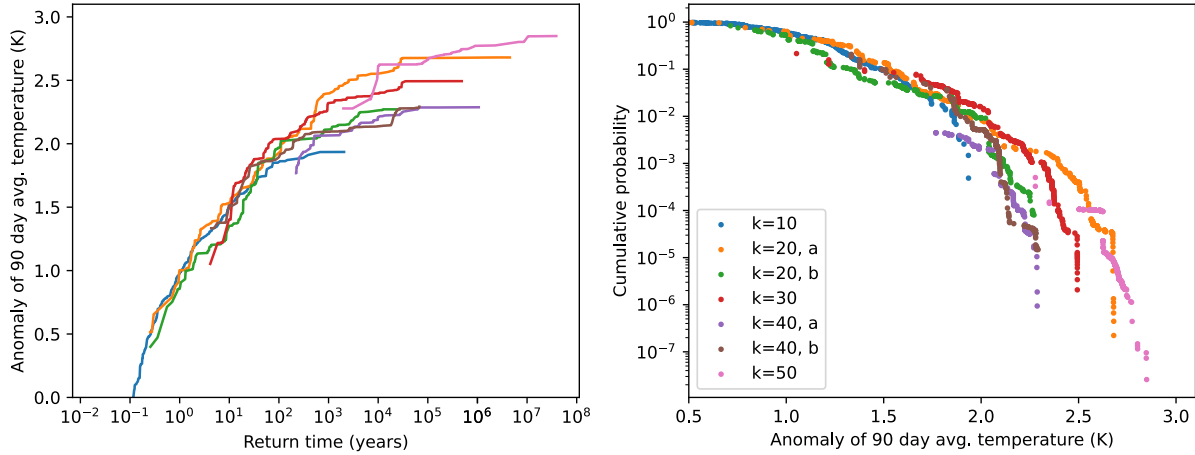

Figure 3.20: For the 4 degrees warming configuration. Left - Return times for the 90-day averaged temperature anomalies. Right - Cumulative probability of the 90-day averaged temperature anomalies.

The estimated return time curve for the 90 day averaged temperature anomalies in the 4 degrees warming configuration is given below.
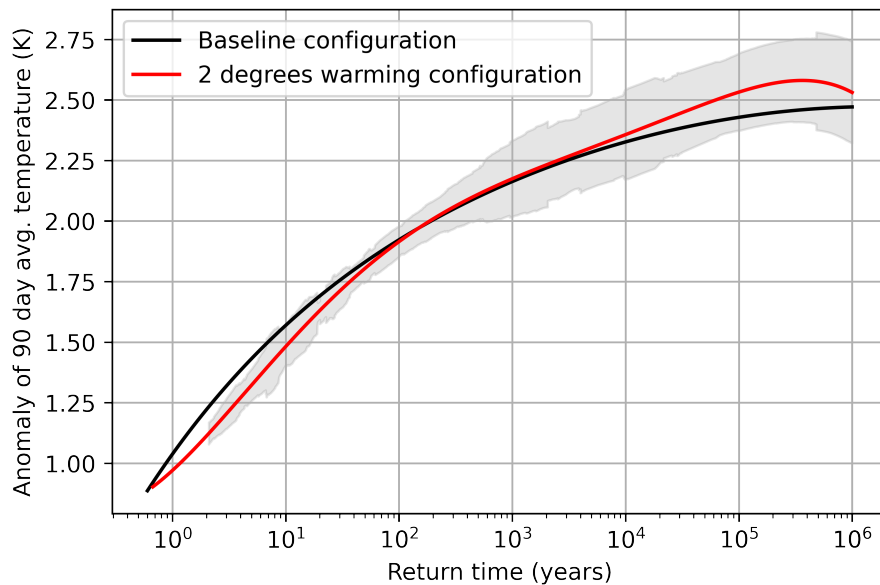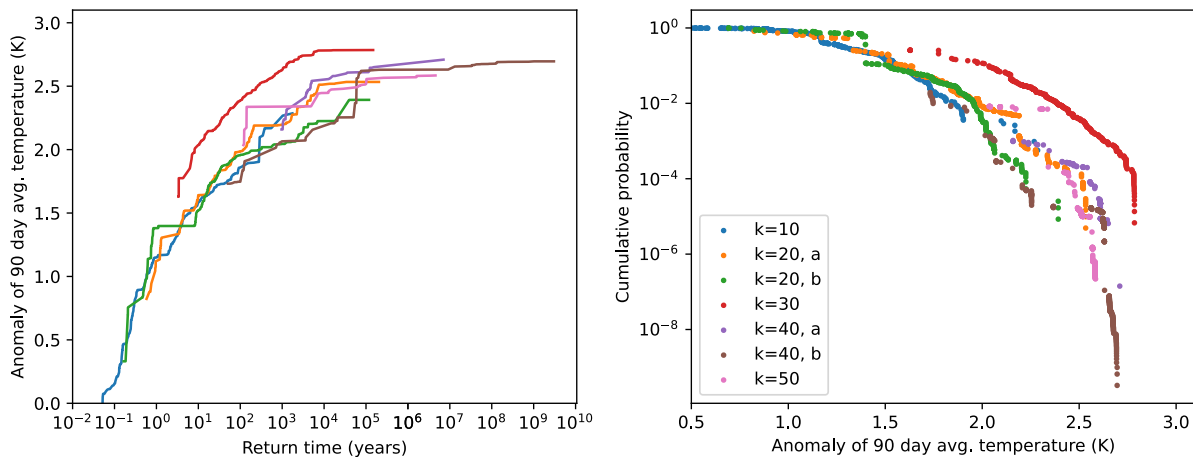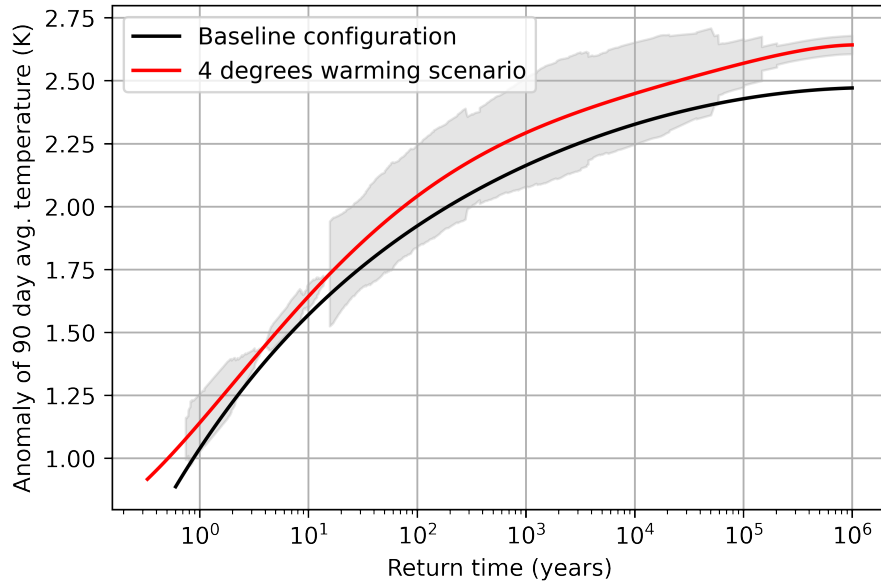
Figure 3.21: The estimated return time curve from the GKLT algorithm for the 4 degrees warming configuration (red) and the baseline configuration (black). The grey area is the error bar, calculated as 1 standard deviation (SD) of the curves averaged at each point.

The return times for the 90 day average temperature anomalies in the 4 degrees warming configuration are lower than those in the baseline configuration, especially in the range of return times higher than 100 years. Hence, the probability of occurrence of these extreme events is larger in the 4 degrees warming configuration.

A predominant difference between the land surface configurations and the warming configurations is as follows - For the land surface configurations, the return times of the 90-day averaged temperature anomalies decrease uniformly across anomalies as the relative humidity over land is decreased. In contrast, for the warming configurations, only the the return times of the larger 90-day averaged temperature anomalies decrease as the atmosphere is warmed, while the return times of the lower anomaly values remain close to that of the baseline.

3.3 Statistical analysis

From the return time plots in section 3.2, we see that the return times of the 90 day averaged temperature anomalies change in a non-trivial way from the baseline. We wish to understand what has changed between configuration to induce this effect.

To understand this, we need to look at the trajectories themselves. For each model configuration, we have 3584 biased trajectories of length 128 days each, where heatwaves are very common. As the algorithm is kept exactly the same (only the configuration changes), we can compare between the trajectories of different configurations to understand the changes statistically.

We first define heatwaves as any event where the anomaly of the spatially averaged temperature over our area of interest (see section 2.2) exceeds a threshold. We set the threshold to 2.7 K, the 95th percentile of daily temperature maximum values, computed from the 1000 year long run for the baseline configuration. The anomalies for different configurations are computed using the climatology of the particular configuration. The threshold for the heatwave is kept the same between configurations. The following results are not sensitive to the particular value of the threshold used, as similar results were obtained even when the threshold was set to 3.1 K, the 98th percentile of daily temperature maximum values.

Using this definition, we decompose the trajectories as shown in Fig. 3.22. We decompose the trajectory into segments above and below the threshold. For each segment, we calculate the duration and intensity, where the duration is the length of the segment and the intensity is the mean value of the anomaly over the segment. This procedure ensures that the 90-day averaged anomaly calculated from the decomposed trajectories remain the same.
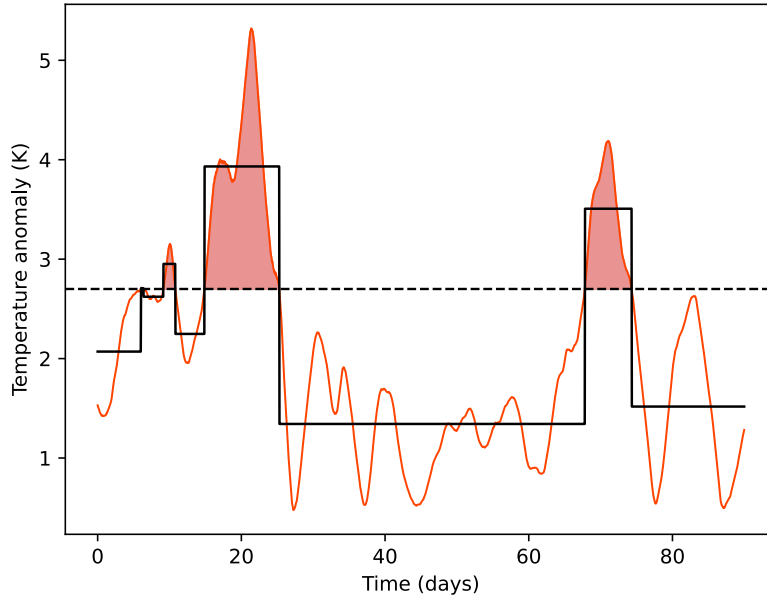
Figure 3.22: (Red curve) Original trajectory - Heatwave episodes shaded in red. (Black curve) Decomposed trajectory - mean value of temperature anomaly computed for different segments above and below the threshold.

However, from our analysis (not shown), we found that the 90-day averaged temperature anomalies could be accounted for by the heatwave episodes alone (Fig. 3.22, shown in red). Therefore, we attempt to explain the changes observed in the return times of the 90 day averaged temperature anomalies (section 3.2) in terms of the intensity, duration and the number of heatwaves in the decomposed trajectories of the different configurations. For our convenience, we consider the different land configurations and the warming configurations separately.

3.3.1 Land configurations

We look at the heatwaves in the trajectories of the different land relative humidity scaling configurations. We filter out trajectories associated with negative 90 day averaged temperature anomalies as they are of no interest to us. The results are the same irrespective of this filtering.

Fig. 3.23 shows cumulative plots for number of heatwaves above some duration/intensity

47

value. The cumulative number of heatwaves is highest in the drier land configuration and lowest in the wetter land configuration. This pattern is consistent across duration/ intensity values. Note that the duration/ intensity values are for individual heatwaves in the trajectories.

Figure 3.23: For the land configuration experiments. Left - number of heatwaves with duration larger than some value. Right - number of heatwaves with intensity larger than some value.

Fig. 3.24 shows a histogram plot of the number of heatwaves for different durations and intensities. A clear trend in intensity and duration of heatwaves between configurations is seen, with both intensity and duration decreasing from the drier land to the baseline to the wetter land configuration. From the drier land to the baseline configuration, the predominant change is the reduction in intensity, while from the baseline to the wetter land configuration, the primary change is the shortening of the duration.
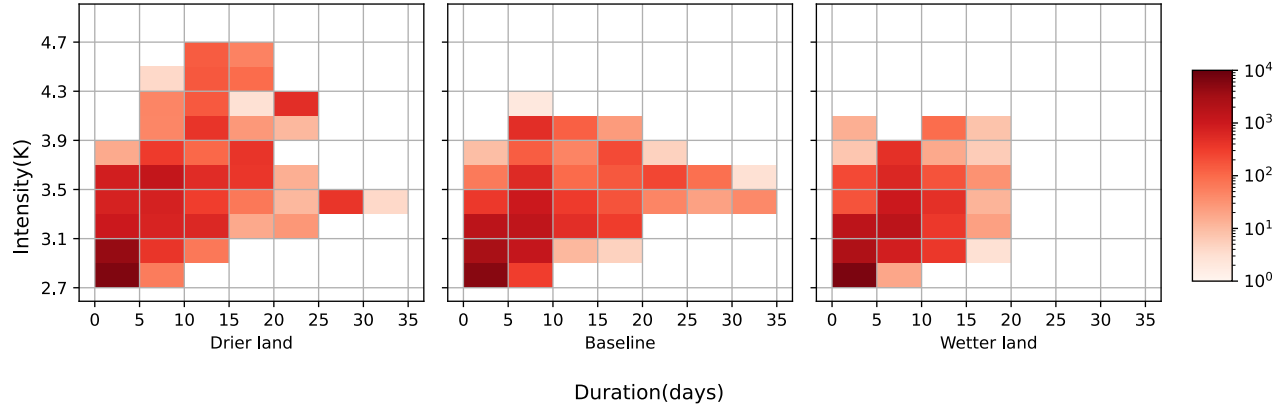
Figure 3.24: For the land configuration experiments - histogram plots for the number of heatwaves of different duration and intensity values.

Hence, the return time differences seen in section 3.2 for the land configurations can be attributed to an increase in the number of heatwaves, as the scaling of the relative humidity over land is decreased, with an increase in both duration and intensity of the heatwaves.

3.3.2 Warming configurations

We look at the heatwaves in the trajectories of the different warming configurations. We filter out trajectories associated with negative 90 day averaged temperature anomalies.

Fig. 3.25 shows cumulative plots for number of heatwaves above some duration/intensity value. The cumulative number of heatwaves is in general lowest in the baseline configuration and highest in the 4 degrees warming configuration. The clearest difference between the configuration is seen in the duration plot (left panel), with medium duration heatwaves being more frequent in the 4 degrees warming scenario than in the other warming configurations. Also, there seems to be fewer high duration events in the 4 degrees warming configuration.

Fig. 3.26 shows a histogram plot of the number of heatwaves for different durations and intensities. We see no clear trend in the intensity and duration distribution of the heatwaves between configurations. The total number of heatwaves seems to be the major change between configurations.
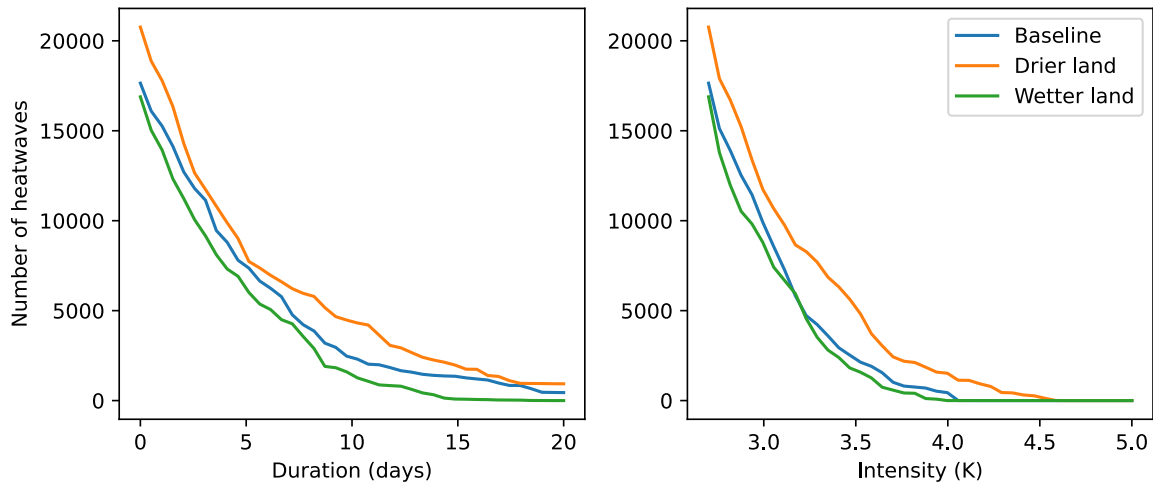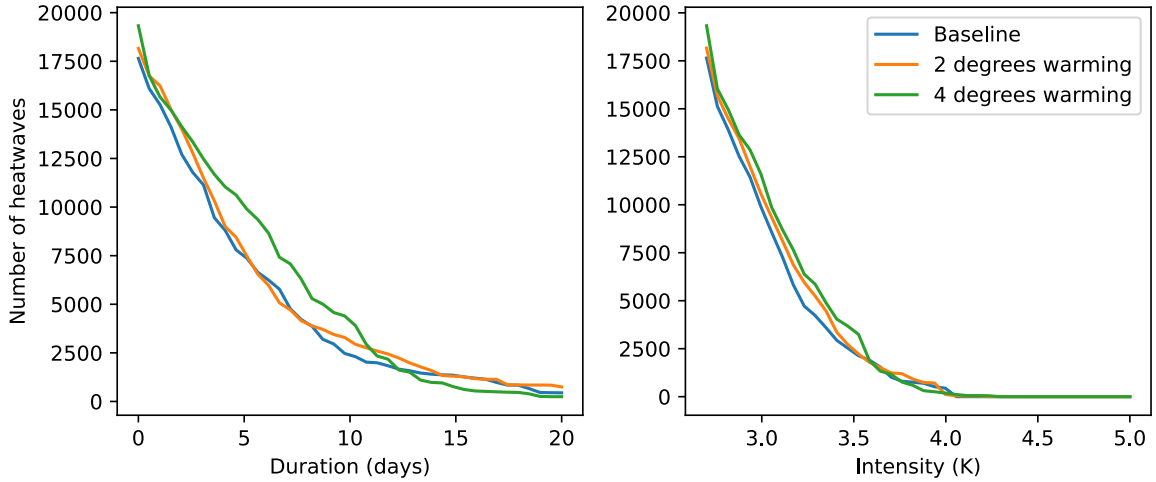
Figure 3.25: For the warming experiments. Left - number of heatwaves with duration larger than some value. Right - number of heatwaves with intensity larger than some value.
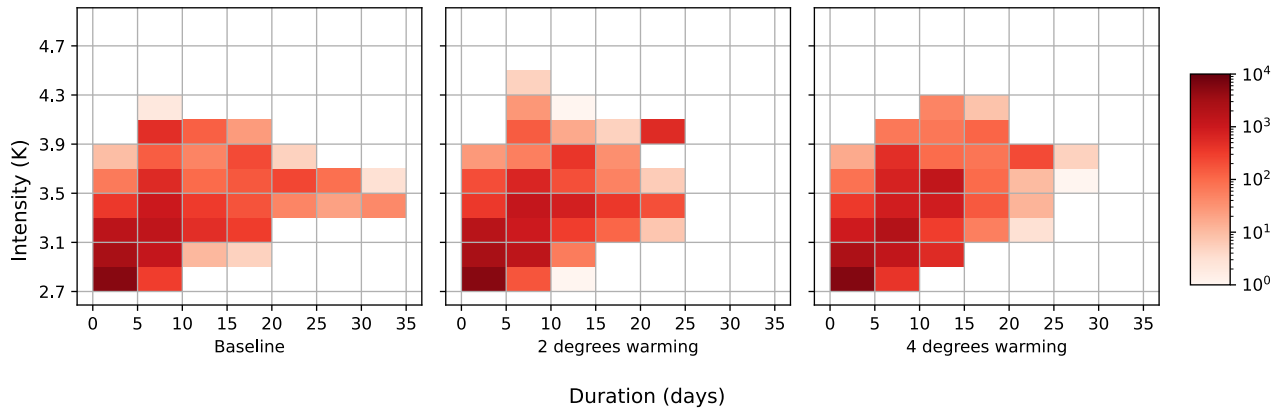
Figure 3.26: For the warming experiments - histogram plots for the number of heatwaves of different duration and intensity values.

To verify that the number of heatwaves is the dominant cause of the differences seen in the return time plots for the warming experiments (section 3.2), we perform an additional analysis, where we sample heatwaves for trajectories with an associated 90 day averaged temperature anomaly above 2.4 K. This value is chosen because the return time plots for the warming experiments (especially in Fig. 3.19) are different beyond an anomaly of 2.4 K.

Fig. 3.27 shows cumulative plots for number of heatwaves above some duration/intensity value for the filtered trajectories. The cumulative number of heatwaves is highest in the 4

degrees warming scenario and lowest in the baseline configuration.
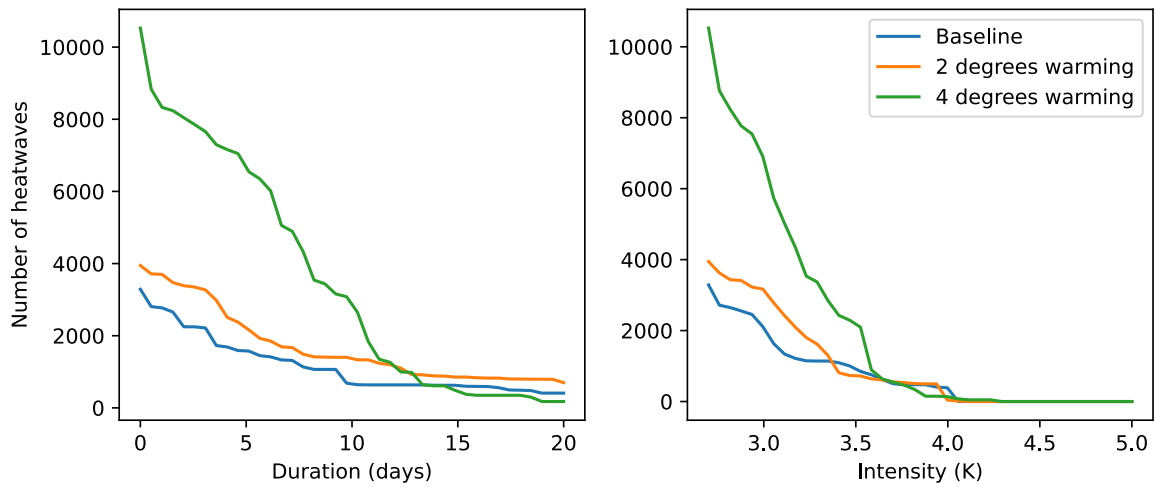
Figure 3.27: For the warming experiments and for filtered trajectories. Left - number of heatwaves with duration larger than some value. Right - number of heatwaves with intensity larger than some value.

From the histogram plot (Fig. 3.28), the clearest pattern that we observe is the decrease in duration and increase in intensity of heatwaves with warming. However, the change in the number of heatwaves seems to be the dominant difference between the warming configurations.

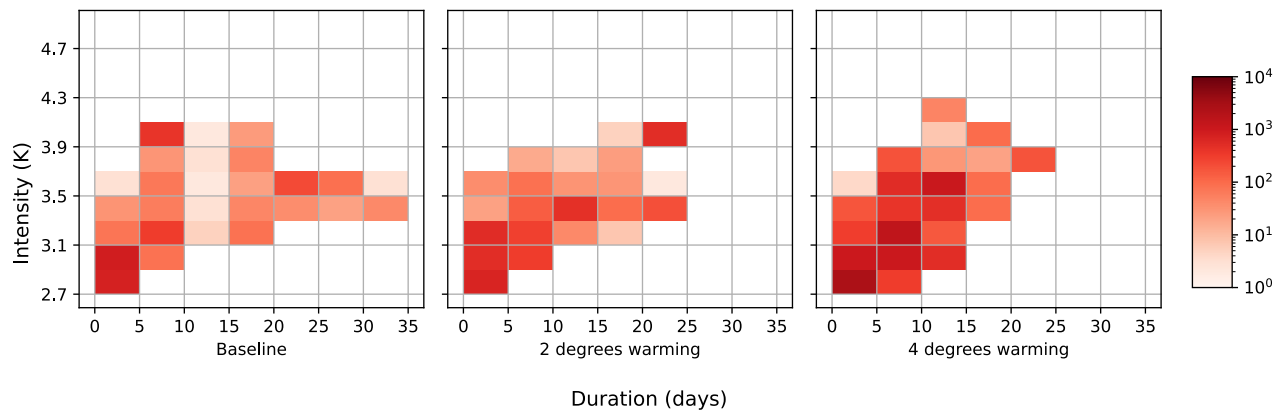Figure 3.28: For the warming experiments and for filtered trajectories - histogram plots for the number of heatwaves of different duration and intensity values.

Hence, the return time differences seen in section 3.2 for the warming configurations,

especially at the larger values of the 90 day averaged anomalies can be attributed to just the increase in the number of heatwaves as the warming is increased.

Chapter 4

Conclusion

In this project, an idealised climate model was set up using CliMT. We used an intermediate complexity model, similar to that in Frierson et al., 2006, with a complexity between simple energy balance models and full-scale GCMs with all feedbacks included. The GKLT rare event sampling algorithm was successfully implemented on this model, giving us the capability to efficiently sample extreme heatwaves and accurately compute their return times. Note that in this study, we have defined heatwaves using temperature anomalies instead of absolute temperatures. This is because we are interested in the changes in the return times due to changes in the higher moments of the distribution, associated with changes in the circulation, and not due to changes in the mean, governed by the large-scale thermodynamics and energy balances.

We find non-trivial changes in the return times of extreme heatwaves as we change the land surface relative humidity scaling and introduce large-scale warming. For the land surface configurations, the return times of the 90-day averaged temperature anomalies decrease uniformly across anomalies as the relative humidity over land is decreased. For the warming configurations, only the return times of larger 90-day averaged temperature anomalies decrease as the atmosphere is warmed. We observe a decrease in the duration and intensity of heatwaves as the land relative humidity scaling is increased, while the number of heatwaves increases predominantly with large-scale warming.

Studies like Seneviratne and Hauser, 2020 and Hirsch et al., 2021 and have been done to examine heatwaves using observational and CMIP6 data, both on regional and global

scales. Such studies are important as they estimate the likelihood of occurrence of temperature extremes at a regional scale, and how they might change in warming scenario. These estimations guide policy makers to plan better and design adaptive measures. In contrast, we approach heatwaves from a conceptual perspective. We start with a simplified model, where the processes that govern heatwaves are easier to understand, while being sufficiently realistic. Features like realistic radiation and land-atmosphere feedbacks due to soil and vegetation, which contribute to heatwaves are left out. It is essential to understand the processes and dynamics in the simplified model thoroughly before adding additional components like topography and realistic land-ocean configurations.

Incorporating the GKLT algorithm on our simplified model allowed us to study the true impact of the changes in circulation on the temperature distribution tails. For instance, the changes in return times for the warming configurations at large anomaly values could only be computed from the robust statistics obtained by biased sampling. The heatwaves sampled from the algorithm can be studied using an intensity-duration-frequency framework, where we look at the intensity, duration and frequency of individual heatwaves in the trajectories obtained from the algorithm. This framework provides a consistent way to combine the physics and the statistics of heatwaves, as the physical processes that govern the intensity, duration and frequency of heatwaves can be studied separately.

In this project, we have only focused on the statistical aspect of heatwaves, in terms on their intensity, duration and frequency. In a future study, we intend to look at how the physical processes of interest in our simplified model - the large-scale horizontal advection, subsidence and surface fluxes, quantitatively contribute to the intensity, duration and frequency of heatwaves. Also, in our current model configuration, we observe that the jet is unintentionally strengthened when the relative humidity over land is decreased. This is a limitation of the land configuration we have implemented. We intend to change the land configuration in a future work to address this issue. After gaining a thorough understanding of the dynamics at play in the simplified model, more realistic components like topography can be added to understand the more complex processes that emerge, in a systematic manner.

Bibliography

Katz, Richard W. and Barbara G. Brown (July 1992). "Extreme events in a changing climate: Variability is more important than averages". en. In: *Climatic Change* 21.3, pp. 289–302. ISSN: 1573-1480. DOI: 10.1007/BF00139728.

Karl, Thomas R. and Richard W. Knight (June 1997). "The 1995 Chicago Heat Wave: How Likely Is a Recurrence?" en. In: *Bulletin of the American Meteorological Society* 78.6. Publisher: American Meteorological Society Section: Bulletin of the American Meteorological Society, pp. 1107–1120. ISSN: 0003-0007, 1520-0477. DOI: 10.1175/1520-0477(1997)078<1107:TCHWHL>2.0.CO;2.

Emanuel, Kerry A. and Marina Živković-Rothman (June 1999). "Development and Evaluation of a Convection Scheme for Use in Climate Models". EN. In: *Journal of the Atmospheric Sciences* 56.11. Publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences, pp. 1766–1782. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/1520-0469(1999)056<1766:DAEOAC>2.0.CO;2.

Kyselý, Jan (2002). "Temporal fluctuations in heat waves at Prague–Klementinum, the Czech Republic, from 1901–97, and their relationships to atmospheric circulation". en. In: *International Journal of Climatology* 22.1. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.720, pp. 33–50. ISSN: 1097-0088. DOI: 10.1002/joc.720.

Patz, Jonathan A. et al. (Nov. 2005). "Impact of regional climate change on human health". en. In: *Nature* 438.7066. Number: 7066 Publisher: Nature Publishing Group, pp. 310–317. ISSN: 1476-4687. DOI: 10.1038/nature04188.

Frierson, Dargan M. W., Isaac M. Held, and Pablo Zurita-Gotor (Oct. 2006). "A Gray-Radiation Aquaplanet Moist GCM. Part I: Static Stability and Eddy Scale". EN. In: *Journal of the Atmospheric Sciences* 63.10. Publisher: American Meteorological Society Section: Journal of the Atmospheric Sciences, pp. 2548–2566. ISSN: 0022-4928, 1520-0469. DOI: 10.1175/JAS3753.1.

Seneviratne, Sonia I., Daniel Lüthi, et al. (Sept. 2006). "Land–atmosphere coupling and climate change in Europe". en. In: *Nature* 443.7108. Number: 7108 Publisher: Nature Publishing Group, pp. 205–209. ISSN: 1476-4687. DOI: 10.1038/nature05095.

Vandentorren, S. et al. (Dec. 2006). "August 2003 Heat Wave in France: Risk Factors for Death of Elderly People Living at Home". In: *European Journal of Public Health* 16.6, pp. 583–591. ISSN: 1101-1262. DOI: 10.1093/eurpub/ckl063.

Perkins, S. E., L. V. Alexander, and J. R. Nairn (2012). "Increasing frequency, intensity and duration of observed global heatwaves and warm spells". en. In: *Geophysical Research Letters* 39.20. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2012GL053361. ISSN: 1944-8007. DOI: 10.1029/2012GL053361.

Xoplaki, Elena et al. (Jan. 2012). "6 - Large-Scale Atmospheric Circulation Driving Extreme Climate Events in the Mediterranean and its Related Impacts". en. In: *The Climate of the Mediterranean Region*. Ed. by Piero Lionello. Oxford: Elsevier, pp. 347–417. ISBN: 978-0-12-416042-2. DOI: 10.1016/B978-0-12-416042-2.00006-9.

Dunne, John P., Ronald J. Stouffer, and Jasmin G. John (June 2013). "Reductions in labour capacity from heat stress under climate warming". en. In: *Nature Climate Change* 3.6. Number: 6 Publisher: Nature Publishing Group, pp. 563–566. ISSN: 1758-6798. DOI: 10.1038/nclimate1827.

Wright, Christopher K., Kirsten M. de Beurs, and Geoffrey M. Henebry (Dec. 2014). "Land surface anomalies preceding the 2010 Russian heat wave and a link to the North Atlantic oscillation". en. In: *Environmental Research Letters* 9.12. Publisher: IOP Publishing, p. 124015. ISSN: 1748-9326. DOI: 10.1088/1748-9326/9/12/124015.

Christidis, Nikolaos, Gareth S. Jones, and Peter A. Stott (Jan. 2015). "Dramatically increasing chance of extremely hot summers since the 2003 European heatwave". en. In: *Nature Climate Change* 5.1. Number: 1 Publisher: Nature Publishing Group, pp. 46–50. ISSN: 1758-6798. DOI: 10.1038/nclimate2468.

Coumou, Dim, Jascha Lehmann, and Johanna Beckmann (Apr. 2015). "Climate change. The weakening summer circulation in the Northern Hemisphere mid-latitudes". eng. In: *Science (New York, N.Y.)* 348.6232, pp. 324–327. ISSN: 1095-9203. DOI: 10.1126/science.1261768.

Loikith, Paul C., Duane E. Waliser, et al. (Feb. 2015). "Surface Temperature Probability Distributions in the NARCCAP Hindcast Experiment: Evaluation Methodology, Metrics, and Results". EN. In: *Journal of Climate* 28.3. Publisher: American Meteorologi-

cal Society Section: Journal of Climate, pp. 978–997. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-13-00457.1.

Perkins, Sarah E. (Oct. 2015). "A review on the scientific understanding of heatwaves—Their measurement, driving mechanisms, and changes at the global scale". en. In: *Atmospheric Research* 164-165, pp. 242–267. ISSN: 0169-8095. DOI: 10.1016/j.atmosres.2015.05.014.

Rohwer, Christian M., Florian Angeletti, and Hugo Touchette (Nov. 2015). "Convergence of large deviation estimators". In: *Physical Review E* 92.5. arXiv:1409.8531 [cond-mat], p. 052104. ISSN: 1539-3755, 1550-2376. DOI: 10.1103/PhysRevE.92.052104.

Schneider, Tapio, Tobias Bischoff, and Hanna Płotka (Mar. 2015). "Physics of Changes in Synoptic Midlatitude Temperature Variability". EN. In: *Journal of Climate* 28.6. Publisher: American Meteorological Society Section: Journal of Climate, pp. 2312–2331. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-14-00632.1.

Byrne, Michael P. and Paul A. O'Gorman (Dec. 2016). "Understanding Decreases in Land Relative Humidity with Global Warming: Conceptual Model and GCM Simulations". EN. In: *Journal of Climate* 29.24. Publisher: American Meteorological Society Section: Journal of Climate, pp. 9045–9061. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-16-0351.1.

Rohini, P., M. Rajeevan, and A. K. Srivastava (May 2016). "On the Variability and Increasing Trends of Heat Waves over India". en. In: *Scientific Reports* 6.1. Number: 1 Publisher: Nature Publishing Group, p. 26153. ISSN: 2045-2322. DOI: 10.1038/srep26153.

Garfinkel, Chaim I. and Nili Harnik (Jan. 2017). "The Non-Gaussianity and Spatial Asymmetry of Temperature Extremes Relative to the Storm Track: The Role of Horizontal Advection". EN. In: *Journal of Climate* 30.2. Publisher: American Meteorological Society Section: Journal of Climate, pp. 445–464. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-15-0806.1.

Ghatak, Debjani et al. (Aug. 2017). "The role of local heating in the 2015 Indian Heat Wave". en. In: *Scientific Reports* 7.1. Number: 1 Publisher: Nature Publishing Group, p. 7707. ISSN: 2045-2322. DOI: 10.1038/s41598-017-07956-5.

Hsiang, Solomon et al. (June 2017). "Estimating economic damage from climate change in the United States". In: *Science* 356.6345. Publisher: American Association for the Advancement of Science, pp. 1362–1369. DOI: 10.1126/science.aal4369.

Mann, Michael E. et al. (Mar. 2017). "Influence of Anthropogenic Climate Change on Planetary Wave Resonance and Extreme Weather Events". en. In: *Scientific Reports* 7.1.

Number: 1 Publisher: Nature Publishing Group, p. 45242. ISSN: 2045-2322. DOI: 10. 1038/srep45242.

Dosio, Alessandro et al. (Apr. 2018). "Extreme heat waves under 1.5 °C and 2 °C global warming". en. In: *Environmental Research Letters* 13.5. Publisher: IOP Publishing, p. 054006. ISSN: 1748-9326. DOI: 10.1088/1748-9326/aab827.

Lestang, Thibault et al. (Apr. 2018). "Computing return times or return periods with rare event algorithms". en. In: *Journal of Statistical Mechanics: Theory and Experiment* 2018.4. Publisher: IOP Publishing and SISSA, p. 043213. ISSN: 1742-5468. DOI: 10.1088/1742-5468/aab856.

Loikith, Paul C., J. David Neelin, et al. (Dec. 2018). "Short Warm-Side Temperature Distribution Tails Drive Hot Spots of Warm Temperature Extreme Increases under Near-Future Warming". EN. In: *Journal of Climate* 31.23. Publisher: American Meteorological Society Section: Journal of Climate, pp. 9469–9487. ISSN: 0894-8755, 1520-0442. DOI: 10.1175/JCLI-D-17-0878.1.

Monteiro, Joy Merwin, Jeremy McGibbon, and Rodrigo Caballero (Sept. 2018). "sympl (v. 0.4.0) and climt (v. 0.15.3) – towards a flexible framework for building model hierarchies in Python". English. In: *Geoscientific Model Development* 11.9. Publisher: Copernicus GmbH, pp. 3781–3794. ISSN: 1991-959X. DOI: 10.5194/gmd-11-3781-2018.

Ragone, Francesco, Jeroen Wouters, and Freddy Bouchet (Jan. 2018). "Computation of extreme heat waves in climate models using a large deviation algorithm". In: *Proceedings of the National Academy of Sciences* 115.1. Publisher: Proceedings of the National Academy of Sciences, pp. 24–29. DOI: 10.1073/pnas.1712645115.

Sandeep, A. and V. S. Prasad (2018). "Intra-annual variability of heat wave episodes over the east coast of India". en. In: *International Journal of Climatology* 38.S1. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/joc.5395, e617–e628. ISSN: 1097-0088. DOI: 10.1002/joc.5395.

Kumari, Savitri et al. (Apr. 2019). "Return period of extreme rainfall substantially decreases under 1.5 °C and 2.0 °C warming: a case study for Uttarakhand, India". en. In: *Environmental Research Letters* 14.4. Publisher: IOP Publishing, p. 044033. ISSN: 1748-9326. DOI: 10.1088/1748-9326/ab0bce.

Miralles, Diego G. et al. (2019). "Land–atmospheric feedbacks during droughts and heatwaves: state of the science and current challenges". en. In: *Annals of the New York Academy of Sciences* 1436.1. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/nyas.13912, pp. 19–35. ISSN: 1749-6632. DOI: 10.1111/nyas.13912.

Ragone, Francesco and Freddy Bouchet (June 2020). "Computation of Extreme Values of Time Averaged Observables in Climate Models with Large Deviation Techniques". en. In: *Journal of Statistical Physics* 179.5, pp. 1637–1665. ISSN: 1572-9613. DOI: 10.1007/s10955-019-02429-7.

Seneviratne, Sonia I. and Mathias Hauser (2020). "Regional Climate Sensitivity of Climate Extremes in CMIP6 Versus CMIP5 Multimodel Ensembles". en. In: *Earth's Future* 8.9. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1029/2019EF001474, e2019EF001474. ISSN: 2328-4277. DOI: 10.1029/2019EF001474.

Hirsch, Annette L. et al. (2021). "CMIP6 MultiModel Evaluation of Present-Day Heatwave Attributes". en. In: *Geophysical Research Letters* 48.22. _eprint: https://onlinelibrary.wiley.com/doi/pdf e2021GL095161. ISSN: 1944-8007. DOI: 10.1029/2021GL095161.

IMD (Apr. 2022). *Monthly Weather and Climate Summery for the month of March 2022.*