Constraining the occurrence probability of compound hot and dry summers in CMIP5 models by using an observational constraint on the dependence between temperature and precipitation

Master Thesis
Faculty of Science, University of Bern

handed in by

Sidharth Sivaraj

2021

Supervisor

Dr. Jakob Zscheischler

Oeschger Centre for Climate Change Research

Climate and Environmental Physics, University of Bern

Abstract

In this project, compound hot and dry summers over a region are defined as summers where the average temperature exceed the 90^{th} percentile and at the same time the average precipitation fall below the 10^{th} percentile of the 20^{th} century summer climatology over the region. Such compound extremes pose disproportionate and uncertain future risk to both society and the environment. The dependence between the drivers of compound events, in this case the mean summer temperature and precipitation, could significantly affect the future occurrence probability of such hot and dry summers. We constrain the projections of the occurrence probability for compound hot and dry summers, from 40 different CMIP5 global climate models, by a dependence constraint between the inter-annual summer means of temperature and precipitation. Statistical comparison of independent correlations and empirical copulas are employed for constraining the models. No significant difference is observed for the mean probability projection of compound hot and dry summers between the constrained model ensemble and the all model ensemble for the 21^{st} century projections under the RCP 8.5 scenario. Some significant difference is observed for the 20^{th} century. However, the regions where the difference is significant also spatially correlate with regions of uncertain data due to lack of observation stations for this period.

Contents

1 Introduction

The 2018 growing season in Germany saw record breaking high temperatures and low precipitation (Zscheischler and Fischer (2020)). This concurrent hot and dry growing condition caused severe impacts to the country's agriculture production and its forests (Buras et al. (2020)). The disastrous event, part of many such events across northern hemisphere mid-latitudes that occurred in 2018, manifested due to a hemisphere-wide wave-7 circulation pattern (Kornhuber et al. (2019)). However, researchers have shown that it is unlikely for such events to have occurred without human-induced climate change (Vogel et al. (2019)). One may be tempted to think that such simultaneous occurrences of record breaking hot conditions and record breaking dry conditions are a low probability event. However, Manning et al. (2019) have shown that the probability of extremely long-duration dry periods co-occurring with extremely high temperatures has significantly increased for the present (1984–2013), compared to the reference period (1950–1979), for most parts of Europe. When such extremes co-occur, they often transcend the capacity of underlying systems to cope, leading to disasters (Zscheischler et al. (2020)). Therefore, it is important to understand, analyse, and accurately predict the occurrences of such concurrent extreme events, also referred to as compound events.

A weather or climate event can be categorised as a compound event, when there are multiple climate drivers (such as cyclones, cold fronts, stationary high-pressure systems etc.) and/or hazards (such as droughts, heat waves, floods etc.) contributing to the associated potential risk. Here, the potential risk could be societal, environmental, or both (Zscheischler et al. (2018)). Many studies have shown that the impacts associated with compound events are often disproportionately higher than those associated with individual climate/weather extremes (Martius et al. (2016); Zscheischler et al. (2014); Leonard et al. (2014)). For example, the co-occurrence of a heatwave and a drought over land, i.e. a compound hot and dry event, can potentially lead to increased tree mortality in the affected region (Allen et al. (2010)), can cause increased propensity to wild fires (Flannigan et al. (2009)) and adversely affect the vegetation health to a significant degree (Ciais et al. (2005)). These make the compound hot and dry event an important factor while considering the land carbon cycle. Being able to accurately project the probability of a hot and dry event occurring in the future, therefore, is of great significance for the environment, agriculture, human-habitability and the carbon budget.

Moreover, it has also been observed that in the warmest 3 months over land ("summer"), the average temperature and precipitation show a negative inter-annual correlation pattern, over many regions of the planet (See Fig.1). This suggests that the occurrence probability of compound hot and dry summers could be higher in these regions, than if temperature and precipitation were independent of each other. This directly translate to a higher future risk. Therefore, in this project, we aim to develop a constraint based on the inter-annual dependence between average summer temperature and average summer precipitation, based on the observational data. The constraint is computed over each grid point on the entire land domain (except for Antartica). This 'dependence constraint based on the observational data' is then used to constrain an ensemble of single runs from 40 'Coupled Model Intercomparison Project - Phase 5' (CMIP5) global climate models. The subset of CMIP5 models, at each grid point, containing only those models that show 'comparable' dependence between temperature and precipitation as in the observational data, are referred to as the ensemble of constrained models, or simply "good models". Subsequently, we analyse the mean likelihood of compound hot and dry summers occurring, as projected by the ensemble of all CMIP5 models and the ensemble of 'good' CMIP5 models at each grid. This comparison is carried out for the 20$^{\text{th}}$ century projections (with historical forcing) and the 21$^{\text{st}}$ century projections (under the 'Representative Concentration Pathway 8.5' or 'RCP 8.5' warming scenario). Such an analysis can provide a global insight, into the association between how well the CMIP5 models capture the observed dependence of temperature and precipitation, and their corresponding projections for the frequency of compound hot and dry summer occurrences in the future. This, as we observed, has important implications in terms of future risk reduction and planning.

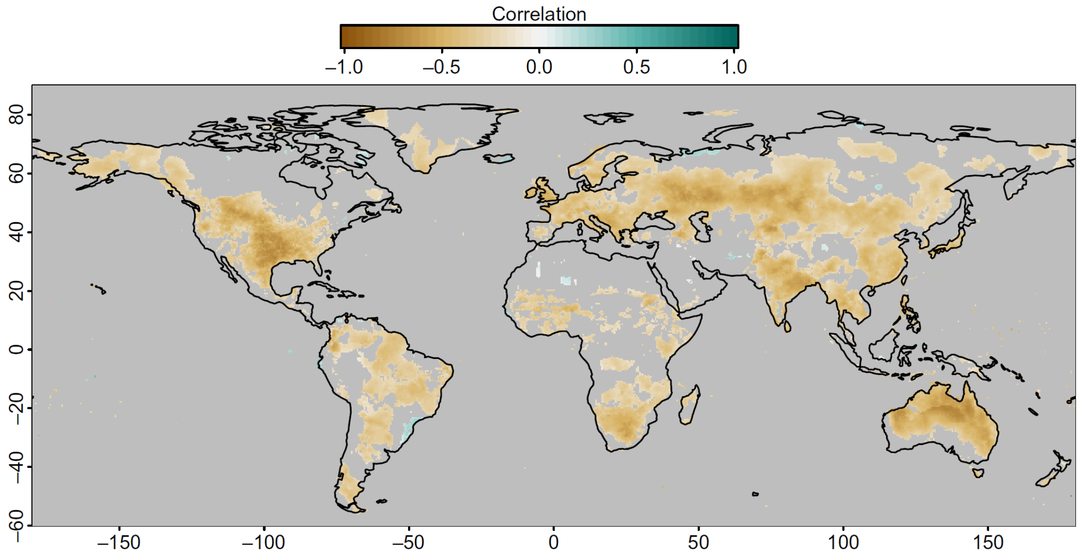


Figure 1: **Inter-annual correlation between yearly averaged temperature and precipitation over the warmest 3-month period (summer) based on the climatology of each grid point**
Figure shows average of the correlations based on the observational data sets CRU (1901–2013), Princeton (1901–2012), and Delaware (1901–2012). Regions in grey indicate Oceans and where correlations are insignificant at the 0.05 level in at least 2 data sets (Plot from Zscheischler and Seneviratne (2017))

Conventional research on hot and dry extremes have often focused on univariate statistics, i.e. either that of temperature or precipitation. For example, there have been studies that analyse the spatial/temporal patterns of dry/wet weather while completely ignoring the temperature component (Martin-Vide and Gomez (1999); Singh et al. (2014)), and vice versa (Diffenbaugh and Ashfaq (2010); Salameh et al. (2019)). However, of late, researchers have started looking into such extremes as compound events involving concurrent extremes of multiple climatic variables, and using methods of multivariate statistics to study them. There are many examples for such studies in the recent years from across the globe such as, from the USA (Mazdiyasni and AghaKouchak (2015)), from the Mediterranean (Vogel et al. (2021)), from Brazil (Geirinhas et al. (2021)) and from India (Sharma and Mujumdar (2017)). While many such researches have only used observation data (or reanalysis data) to analyse past extremes, there are also studies present, which made use of model projections into the future for computing risk associated with compound extremes (Lemus-Canovas and Lopez-Bustins (2021), p.1).

This master thesis project is partly inspired by one such study conducted by Zscheischler and Seneviratne (2017). In their study, using copula analysis on CMIP5 projections of temperature and precipitation, Zscheischler and Seneviratne (2017) showed that the occurrence frequency projections of compound hot and dry summers increased substantially, from the historical period to the 21st century, over many land regions of the world (even up to 10 folds in some cases). Moreover, using simulated temperature data along with observations and CMIP5 projections, Zscheischler and Seneviratne (2017) were also able to demonstrate that the dependence between temperature and precipitation, marked by negative correlation in most regions, contributed towards increasing the occurrence frequency of compound hot and dry summers for the historical period as well as the future. This implies that the failure to faithfully incorporate the dependence between temperature and precipitation could run the risk of underestimating the future risks associated with hot and dry compound extremes (Zscheischler and Seneviratne (2017)). Following this, in a recent study, Zscheischler and Fischer (2020) constrained CMIP5 models using an observational constraint on the dependence between temperature and precipitation over western Germany. The authors noted that the constrained CMIP5 models tend to project a slightly lower return period for a compound hot and dry growing season such as the one in 2018 (Zscheischler and Fischer (2020)).

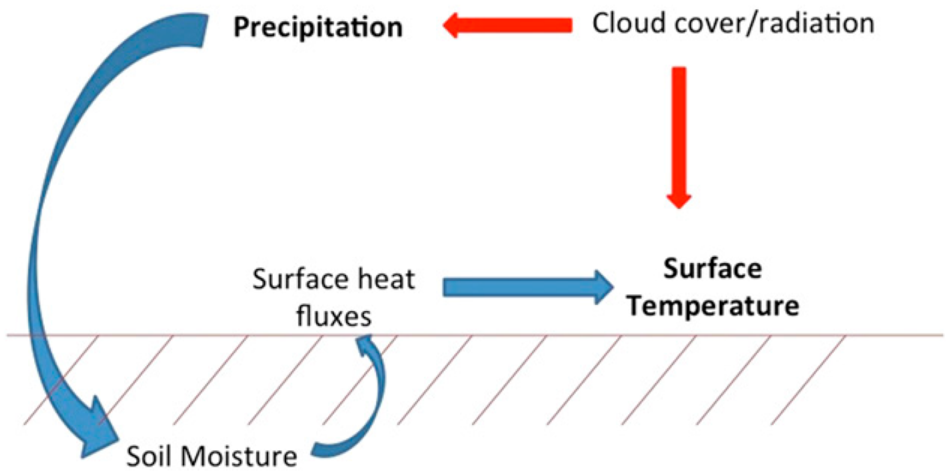


Figure 2: **Simplified representation of mechanisms determining summer correlation between temperature and precipitation over land**: cloud-atmosphere interactions are represented by red and land–atmosphere feedbacks by blue (from Berg et al. (2015))

Berg et al. (2015) have suggested that a combination of atmospheric and surface processes lead to the widely observed inter-annual anti-correlation between summer mean temperature and precipitation over land (see Fig.2). The land - atmosphere feedback happens locally, when a decrease in precipitation results in the reduction of moisture content in soil, which increases the latent heat flux. This will result in enhanced sensible heating at the surface and therefore, causing higher near surface temperatures. Due to the same reasons, high precipitation results in reduced near surface temperatures. The cloud - atmosphere feedback stems from the simple association that lack of cloud cover (or precipitation) results in increased incoming shortwave radiation, which leads to higher surface temperatures. For this phenomenon too, the converse holds true as well (Berg et al. (2015)). Based on the analysis of different models belonging to the 'Global Land–Atmosphere Coupling Experiment of CMIP5' (GLACE-CMIP5), Berg et al. (2015) observed that the CMIP5 models show considerable differences with respect to one another, in terms of incorporating the physical processes that govern the dependence between temperature and precipitation. Therefore, constraining the CMIP5 models based on the observed dependence between temperature and precipitation is a meaningful exercise, in that we may potentially also be constraining the models, at each grid point, for the underlying physical feedback mechanism/ processes that govern the dependence.

1.1 State of knowledge and research questions

Though not specifically referred to as compound events, the consequences of multiple extreme events co-occurring have long been inferred by climate scientists. However, in the context of the 'Intergovernmental Panel on Climate Change' (IPCC), the chapter titled 'Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation' in SREX report (Field et al. (2012)) was the first to distinctly address compound events as having great importance in risk assessment. The definition coined by Field et al. (2012) for compound events was further generalised by Zscheischler et al. (2018) into the now widely accepted definition, introduced earlier. The research into compound events and their impacts today have grown exponentially, and exists in the intersection of many disciplines such as climate-impacts research, statistics, computer science etc. (Zscheischler et al. (2020)). This is a testament to the breadth and complexity of the umbrella subject. Although, despite being a thriving field today, many methodologies employed in compound events research, can still considered to be in their developing stages (Zscheischler et al. (2020)).

It is in such circumstances that, under the European Cooperation in Science and Technology ('COST') action 'CA17109', 'Understanding and modeling compound climate and weather events' ('COST-DAMOCLES'), was instituted in 2018. The key objective of COST-DAMOCLES is to overcome the current limitations in climate and impact modelling of compound events, which makes designing and implementation of appropriate adaptation/mitigation strategies a complicated affair (DAMOCLES-MoU (2018)). Following the aims outlined by COST-DAMOCLES in achieving the key objective, climate scientists have made huge strides in standardising the domain of compound events research. A notable mention is the 2020 paper titled 'A typology of compound weather and climate events' (Zscheischler et al. (2020)). In this paper, the authors have categorised compound events into the following four broad types;

- **Multivariate compound events :** Where multiple drivers and/or hazards co-occur within the same geographical boundaries
- **Temporally compounding events :** Where hazards occurring in succession lead to an impact
- **Preconditioned compound events :** Where existing preconditions of weather/climate worsen the impacts of a novel hazard
- **Spatially compounding events :** Where impacts are worsened by co-occurrence of hazards in multiple connected geographical regions

Following the classification, attempts are also being constantly made in exploring, standardising and introducing novel methodologies for studying each type of compounding events, as mentioned above. The workshop on 'Compound climate events and extremes in the midlatitudes' conducted online from 7–9 September 2020 (Messori et al. (2021)) is one such example among many.

In this project, we look at compound hot and dry summers as multivariate compound events, over each grid point on land. Following Zscheischler and Fischer (2020), we define a compound hot and dry summer as a summer where the average temperature exceed the 90$^{\text{th}}$ percentile and at the same time the average precipitation is below the 10$^{\text{th}}$ percentile, with respect to the 20$^{\text{th}}$ century summer climatology over the grid point. The methodologies used in this study are motivated by Zscheischler and Seneviratne (2017) and Zscheischler and Fischer (2020). The observation was made by Zscheischler and Fischer (2020) that the ensemble of CMIP5 models which show 'comparable' dependence between temperature and precipitation, to that shown by observation data in the historical period (i.e. 'good models'), tend to project lower mean return periods for a hot and dry growing season such as the 2018 one, over Germany, compared to the all model ensemble. This led us to the following research questions for this project.

- On a global scale over land, will 'good' CMIP5 models project a higher occurrence frequency for compound hot and dry summers on average, compared to an ensemble of all the CMIP5 models?
- How will the difference between mean frequency projections of 'good' CMIP5 models and all the CMIP5 models vary in the historical period without warming (1901-2000) compared to the future period under the warming scenario of RCP 8.5 (2001-2100)?
- Will the interpretations vary across geographical regions (in particular grid points), depending on whether the corresponding dependence between mean summer temperature and precipitation over the region is anti-correlation or not?

To address these questions, we generalise the observation made by Zscheischler and Fischer (2020) over Germany, to arrive at the following hypothesis.

Hypothesis : In general, over land regions where the mean summer temperature and precipitation are anti-correlated, the good models will, on average, project higher occurrence frequency for compound hot and dry summers compared to the all model ensemble. This difference will be more apparent in the 21st century, under RCP 8.5 warming, compared to the historical period with no warming (1901-2000).

If shown to be valid, this hypothesis will have significant implications for future risk assessment and planning, with respect to compound hot and dry events. Moreover, such an analysis can be replicated for many other compound events. The insights could be further incorporated into testing skills of CMIP models in predicting such events, by taking dependence between the variables into account. This is an aspect, which often gets overlooked while testing the skills of models. For e.g: (Ridder et al. (2021)).

2 Data

This project uses two primary climate variables for its analysis. One, the monthly means of daily mean surface temperatures, which is measured/projected as the air temperature from 2 metres above ground. This is referred to as '2m air temperature' or simply as 'temperature' in this thesis. Two, total hydrological precipitation (dry + wet) recorded/projected per month. We refer to this as 'precipitation'.

The variables of both temperature and precipitation are utilised from observation based dataset named the 'Climatic Research Unit gridded Time Series Version 4' ("CRU TS" or simply "CRU" [1901-2019]). The variable of precipitation is also utilised from the observation based data set named the 'Global Precipitation Climatology Centre Climatology (CLIM) V2020' ("GPCC CLIM" or "GPCC" [1901-2019]). And the time series of global mean surface temperature change is utilised from the observation based data set named the 'Goddard Institute for Space Studies Surface Temperature Analysis version 4' ("GISTEMP v4" or "GISSTEMP" [1880-Present]).

Similarly, the variables of both temperature and precipitation are also utilised from 'The European Centre for Medium-Range Weather Forecasts twentieth century reanalysis' ("ECMWF ERA 20C" or "ERA 20C"[1901-2010]) dataset and from 40 different 'Coupled Model Intercomparison Project Phase 5' ("CMIP5" [1901-2100]) global climate models. A short introduction to each of these datasets are provided below.

2.1 Observation based datasets

2.1.1 CRU TS [1901-2019]

The 'Climatic Research Unit gridded Time Series Version 4' ("CRU TS" or simply "CRU") is an observation based dataset that provide monthly data over all land domains, barring Antarctica, for the period 1901-2019. The data is available for the following 7 variables; temperature (minimum, maximum and mean), volume of precipitation, vapour pressure, number of wet days and cloud cover. The resolution of data is in 0.5° latitude by 0.5° longitude grid. CRU TS uses 7 principal sources to collect observation data and then converts them into anomalies first. Later these anomalies are transformed into 3 primary variables. The final data products are obtained through a process named angular-distance weighting (ADW). However, it is cautioned that CRU TS may not be the best dataset for exploring the global/regional climate (change) trends. This is so because CRU TS is designed primarily with the overarching objective of providing data over all land domain. This forces it to patch data in certain regions using various techniques. However CRU TS *"can be used"* for the said purpose, as extreme scientific accuracy and correction methodologies are employed in the production of the data.
Reference for this section : Harris et al. (2020)

2.1.2 GPCC [1901-2019]

The 'Global Precipitation Climatology Centre Climatology (CLIM) V2020' ("GPCC CLIM" or "GPCC") is one of the four precipitation products from the the Global Precipitation Climatology Centre (GPCC). High resolution monthly global precipitation data, for the period 1901-2019, obtained from ground based rain gauge measurements are made publicly available since 1989. The resolution of data is in 0.5° latitude by 0.5° longitude grid as well.
Reference for this section : Becker et al. (2013)

2.1.3 GISSTEMP [1880-Present]

The 'Goddard Institute for Space Studies Surface Temperature Analysis version 4' ("GISTEMP v4" or "GISSTEMP") provides anomalies in annual global mean surface temperature from 1880-Present, with respect to the 1951-1980 annual mean global surface temperature. They collect observation data from multiple sources around the planet and process them to provide multiple temperature data sets.
References for this section : GISTEMP-Team (2021), Lenssen et al. (2019)

2.2 ERA 20C reanalysis datasets [1901-2010]

'The ECMWF twentieth century reanalysis' ("ERA 20C"), is the primary outcome of the 'European Reanalysis of Global Climate Observations Project' (ERA-CLIM). ERA 20C relies on an 'Integrated Forecast System' (IFS) of the ECMWF. The IFS consists of two main components, i.e. an atmospheric general circulation model (AGCM) and a variational analysis scheme. These two components, usually employed in short range forecasts, are adapted accordingly for ERA 20C analysis spanning well over a century (1901-2000). As observational input, ERA 20C directly assimilates only surface pressure (over land and ocean) and wind velocity (only above the ocean). Also the surface temperature observations are assimilated through forcing (over land and ocean). No other observations are directly assimilated. The Final output of ERA 20C provide high resolution data for numerous variables, including temperature and precipitation, over all domains (land and ocean). The data is made publicly available on monthly as well as daily resolutions for the 20$^{\text{th}}$ century (1901-2010) (ECMWF (2021)).
Reference for this section : Poli et al. (2016)

2.3 CMIP5 global climate model datasets [1901-2100]

The 'Coupled Model Intercomparison Project Phase 5' ("CMIP5") refers to a set of coordinated experiments in climate modelling, by 20 different modelling groups from around the globe. The aims of CMIP5 was to better the understanding of climate processes and to facilitate improved climate projections for the future. And most importantly, to address the pertinent scientific queries raised by IPCC AR4. CMIP5 was meant to provide a framework for coordinating the efforts in global climate modelling from 2008-2013. There are in total 40 different global climate models (GCMs) belonging to the CMIP5 project. We utilised single runs of monthly mean temperatures and precipitation from all the 40 CMIP5 GCMs for this project. The names of all the GCMs in CMIP5 are summarised in Table.1. These models provide high resolution projections of several variables for the 20^{th} century, till 2005, with historical forcing. The future projections, from 2006-2100, takes RCP 8.5 warming scenario into account. Presently, CMIP Phase-6 (CMIP6) has succeeded CMIP5 in its role of coordinating the global climate modelling efforts.

Table 1: List of CMIP5 global climate models used in this project

Sr.No	Model Name	Sr.No	Model Name	Sr.No	Model Name	Sr.No	Model Name
1	ACCESS1-0	11	CMCC-CM	21	GISS-E2-H	31	IPSL-CM5B-LR
2	ACCESS1-3	12	CMCC-CMS	22	GISS-E2-H-CC	32	MIROC-ESM
3	bcc-csm1-1	13	CNRM-CM5	23	GISS-E2-R	33	MIROC-ESM-CHEM
4	bcc-csm1-1-m	14	CSIRO-Mk3-6-0	24	GISS-E2-R-CC	34	MIROC5
5	BNU-ESM	15	EC-EARTH	25	HadGEM2-AO	35	MPI-ESM-LR
6	CanESM2	16	FGOALS-g2	26	HadGEM2-CC	36	MPI-ESM-MR
7	CCSM4	17	FIO-ESM	27	HadGEM2-ES	37	MRI-CGCM3
8	CESM1-BGC	18	GFDL-CM3	28	inmcm4	38	MRI-ESM1
9	CESM1-CAM5	19	GFDL-ESM2G	29	IPSL-CM5A-LR	39	NorESM1-M
10	CMCC-CESM	20	GFDL-ESM2M	30	IPSL-CM5A-MR	40	NorESM1-ME

Reference for this section : Taylor (2009)

3 Methods

The analysis part of this project involved three components. Namely, developing the dependence constraint and constraining the CMIP5 models, comparing the mean frequency projections of compound hot and dry summers between the constrained ('good') model ensemble and the all model ensemble, and exploring other factors such as station number constraints, uncertainty in correlation measurements etc., which may have influenced the observed results. The three stages are explained in three subsections below.

3.1 Methods : Observation based dependence constraints

To begin with, there were global monthly data of temperature and precipitation available from observation based CRU (1901-2019), reanalysis based ERA-20C (1901-2010) and projections from 40 different CMIP5 models (1901-2100). We also had monthly data of precipitation from observation based GPCC (1901-2019)(See section.2). In the first step, we homogenised the field of all data sets by re-gridding them to a 2.5° latitude by 2.5° longitude reference grid, by using bilinear-interpolation method. The 'remapbil' operator in command line 'Climate Data Operators' (CDO) was employed to achieve this (Schulzweida et al. (2006)). CRU and GPCC only provide data over land domain, except for Antartica. Therefore, in the second step of homogenising data fields, we masked all other grid points in ERA 20C and CMIP5 model datasets of temperature and precipitation, where data was unavailable in CRU (this mainly include the Oceans and Antartica). Since our analysis is focused only on land domain, this was convenient for the study.

The next step was to identify the warmest three months (summer) for each grid point on land, where data is available, based on the 1901-2019 temperature climatology from the CRU data. Once those 3 months were identified at each grid, the average 'summer temperature' and average 'summer precipitation' were computed for each year, at each point, using that information. For example, suppose Mar-Apr-May were found to be the warmest 3 months over some grid point in Southern India, based on the 119 year CRU climatology at that grid point. We would then retain only the average value of temperature and precipitation over Mar-Apr-May, for each year from 1901-2019, over that grid point. This was performed for all datasets of temperature and precipitation (Note: Summer was determined by CRU climatology for the other datasets as well). In the end, we were left with 119 year long time series of mean summer temperature and mean summer precipitation, over each grid point, from the CRU (1901-2019), 110 year time series from the ERA 20C (1901-2010) and 200 year time series for each of the 40 CMIP5 models (1901-2100). GPCC too produced a 119 year time series for precipitation (1901-2019).

Then, we performed linear regression on both mean summer temperature and mean summer precipitation series, over all grid points, against global annual mean surface temperature time series. This was done in order to eliminate any warming trends in the data. We wanted to eliminate warming as it would influence the inter-annual dependence between the variables. The annual global mean surface temperature data was used from GISSTEMP to de-trend the CRU and the GPCC series of summer mean temperature and precipitation. For ERA 20C and each of the CMIP5 models, annual global mean temperatures were computed from the temperature data of the models themselves, as the field mean. We employed 'fldmean' operator from CDO to achieve this (Schulzweida et al. (2006)). It was impossible to determine global mean temperature as field mean from the CRU temperature data, as there were no data available over the Oceans or Antartica. After regression, the residuals were retained for all the datasets.

In the next step, to explore the dependence between mean summer temperature and precipitation, we computed correlation between the the two residual series of temperature and precipitation over each grid point. We were left with 43 sets of global data of Pearson correlation coefficients between residual mean summer temperature and residual mean summer precipitation over the land, i.e.;

a) correlation between 119 year residual mean summer temperature and precipitation data from CRU (1901-2019)

b) correlation between 119 year residual mean summer temperature from CRU and residual mean summer precipitation from GPCC (1901-2019)

c) correlation between 119 year residual mean summer temperature and precipitation from ERA 20C (1901-2010)

d) correlation between 105 year residual mean summer temperature and precipitation from each of the 40 CMIP5 models (1901-2005).

For each of the CMIP5 models, only 105 years were considered while computing the correlation for the historical period, because the CMIP5 models employ RCP 8.5 scenario in their projections from 2006 onwards until 2100. Once we obtained correlation between temperature and precipitation as the dependence parameter, our goal was then to constrain CMIP5 models based on the correlation value, at each grid point. In order to do this, we needed a statistical test that can compare two independent correlations over a grid point, and provide a measure to determine whether they are comparable or not. We employed two such tests; The Fisher-Z test and the Method of Zou.

3.1.1 Fisher-Z test

Let c_1 and c_2 be two correlation coefficients, computed from two independent discrete bi-variate distributions P_1 and P_2 of marginal sizes n_1 and n_2 respectively. We want to check whether the dependence between the two variables in P_1 is comparable to the dependence between the two variables in P_2, by comparing their corresponding correlation coefficients. Fisher-Z test achieves this by first transforming the correlation coefficients to the approximately normally distributed z coefficients.

$$z_1 = \frac{1}{2} \cdot \ln\left(\frac{1+c_1}{1-c_1}\right) \tag{1}$$

This is known as the Fisher-Z transformation. Then a significance test on the transformed coefficients z_1 and z_2 are performed in the following way. We define,

$$Z = \frac{z_1 - z_2}{\sqrt{\dfrac{1}{n_1-3} + \dfrac{1}{n_2-3}}} \tag{2}$$

The p-value of significance test is then obtained as, $2 \cdot (1-$ the probability that a standard normal variable x will be less than Z). If the p-value is less than 0.05, we can reject the null hypothesis with 95% confidence and claim that the two correlations (therefore the underlying dependencies) are not comparable. However, if the p-value is greater than 0.05, we cannot reject the null-hypothesis and therefore the correlations could be comparable (but not necessarily). It only means that there is atleast 5% chance that the correlations stem from the same true distribution.

However, in this analysis, we consider two correlation coefficients c_1 and c_2 to be comparable, if the p-value obtained from their Fisher-Z analysis, as explained in this section, is greater than 0.05.

Reference for this section : (Fisher (1921))

3.1.2 Method of Zou

Again, Let c_1 and c_2 be two correlation coefficients, computed from two independent discrete bi-variate distributions P_1 and P_2 of marginal sizes n_1 and n_2 respectively. The modified asymptotic test developed by Zou (2007) determines the confidence interval for the parameter $c_1 - c_2$, with α as the confidence level. We say that the correlations are comparable, if the confidence interval for $c_1 - c_2$ contains zero, when α is 0.05. That is to say, there is at least 5% chance that $c_1 - c_2 = 0$, i.e., $c_1 = c_2$.
Reference for this section : (Zou (2007))

Using the Fisher-Z test and the method of Zou, we compared the correlation coefficient from each of the CMIP5 models, with observation based correlation coefficient from the CRU data, over each grid point. If the correlations were found to be comparable, then the model was categorised as 'good' model at that grid point. This is what we mean by 'constraining CMIP5 model by using an observational constraint on the dependence between temperature and precipitation'.

Even though these methods were successful in constraining the models, the parameter of correlation has certain limitations. For example, correlation is not a good parameter to model the tail dependence. Therefore, we also performed a third test to constrain the CMIP5 models, using an observational constraint on the dependence. This was the comparison of empirical copulas.

3.1.3 Empirical copula test

For a bivariate discrete distribution $P \times Q$ with marginals P and Q, with sample sizes of n each, the empirical copula associated with the distribution is a discrete function $\phi : [0,1]^2 \longrightarrow [0,1]$, which summarises the dependence between P and Q. The function ϕ is defined using the order statistics of tuples $(p_i, q_i) \in P \times Q$, where i goes from 1 to n.

Rémillard and Scaillet (2009) proposed a statistical test to compare the empirical copulas between two such independent bivariate distributions. The test is implemented by the authors themselves in the R-Package 'TwoCop' (Rémillard and Scaillet (2009)). We used this test to directly check if the dependence (captured by empirical copulas) between the marginals in the bivariate distributions of residual mean summer temperature and residual mean summer precipitation obtained from each of the CMIP5 models, is comparable with the dependence between the marginals of the bivariate distribution obtained from the CRU data, at each grid. In this method, since we directly analyse dependence without defining a parameter (such as correlation), and take order statistics into account, this has the advantage of capturing tail dependence better than correlation. Again, if the dependence between the marginals of the bivariate distribution from a CMIP5 model was found to be comparable with that from CRU, the model was marked as 'good' model at that grid.

Reference for this section : (Rémillard and Scaillet (2009))

At the end of this part of the analysis, we were left with the information of which CMIP5 models were 'good' and which models were 'not good' at each gridpoint, with respect to three different tests. Namely, Fisher test of independent correlation comparison, Method of Zou for independent correlation comparison and Rémillard and Scaillet (2009)'s method of comparing empirical copulas.

3.2 Methods : Frequency projections

For this part of the analysis, following Zscheischler and Fischer (2020), we defined compound hot and dry summers over each grid point as summers where average temperature exceeded the 90^{th} percentile of average summer temperature values in the 20^{th} century temperature climatology. And at the same time, average precipitation is less than the 10^{th} percentile in the 20^{th} century precipitation climatology.

So in the first step, for each CMIP5 model over each grid point, we computed the 90^{th} percentile of average summer temperature values in the 100 years (1901-2000). Similarly, we also computed the 10^{th} percentile of average summer precipitation values in the same period. Then we simply counted the number of years where the percentile thresholds are simultaneously exceeded ($>$ for temperature and $<$ for precipitation), in the 20^{th} century (1901-2000), and in the 21^{st} century (2001-2100) separately. The same 20^{th} century percentile thresholds were used in both cases. The frequency of hot and dry summers was then simply calculated as the number of such summers divided by 100. Hundred corresponded to the number of years in both periods, i.e. (1901-2000) and (2001-2100). We chose the same number of years in both periods to give an equal footing for the analysis.

Since we had the information of 'good' models at each grid point available from constraining the models (See section.3.1), we then looked at the mean frequency of hot and dry summers projected by the ensemble of 'good' models and the ensemble of all models separately for the 20^{th} century and the 21^{st} century. We then checked to see if the difference between the mean frequencies projected by the two ensembles were significant or not. This was done by using a simple 'Student-t test' to check whether the means of two independent populations significantly differed from each other. Finally, we looked at the relative difference between the mean frequencies projected by the two ensembles, for both periods, only on the grid points where the differences were significant.

3.3 Methods : Other factors

In this stage of the analysis, we focused our attention on certain aspects, which could be influencing our results. Our first concern was the uncertainty associated with the observation data (CRU) due to the lack of stations providing data over the years. To explore this, we looked at the number of observation stations providing monthly data for CRU TS, for precipitation and temperature, averaged over the period 1901-2019 at each grid. Since precipitation seemed like the more affected variable among the two, in terms of station number constraints, we wanted to check whether this has any impact on the correlation coefficients computed from the data. To analyse this, we compared correlation values computed from CRU, against the average number of stations providing precipitation data (1901-2019) for the grid point from where the correlation value was obtained. We also wanted to check whether the deviation of observation based correlation values from the model projections of correlation values at each grid was in any way influenced by this station constraint. To see this, we also marked the points in the histogram defined above as inliers or outliers, with respect to the model projections of correlations at those corresponding grid points.

Our next concern was the fundamental uncertainty associated with the parameter of correlation itself, when the sample numbers are small, such as in our case. To explore an aspect of this problem, we conducted a numerical experiment. Firstly, we defined a theoretical bivariate normal distribution with a true correlation of -0.2 (which is comparable to the observed correlation between summer temperature and precipitation over many grid points on land). We set the marginals to be standard normal as well (mean = 0, standard deviation = 1). Then we repeatedly sampled sets of 120 data points (which is close to the 119 data points of temperature and precipitation series in CRU) from this distribution. And correlation was computed each time. Finally we plotted 10000 such correlation values from repeated sampling as a histogram.

4 Results

The results from this project are presented in three sub sections. In the first subsection, titled 'observational constraints', the dependence between temperature and precipitation (as correlation) and the comparison of dependence (correlation and copula) between observation based data, reanalysis data and the CMIP5 model data are presented. In the second subsection titled 'frequency projections', the mean frequency projections for hot and dry compound summer occurrence from the ensemble consisting of all the CMIP5 models and the ensemble consisting of only the 'good' CMIP5 models are separately presented. The results for the 20$^{\text{th}}$ century data as well as the 21$^{\text{st}}$ century data are shown separately for these two ensembles. In the final subsection, titled 'other factors', we introduce certain related results, which could potentially explain some of the main results that are obtained from the study (See section 3).

4.1 Results : Observation based dependence constraints

The warmest 3 months at each grid point (summer) based on the 1901-2019 climatology of CRU observational data is portrayed in Fig.3. Fig.3 shows the month in the middle of the summer for each grid point. It can be seen that there is a circular smooth transition of summer months from the North to South. And the summer months are nearly the same across same latitude, which is an obvious result.

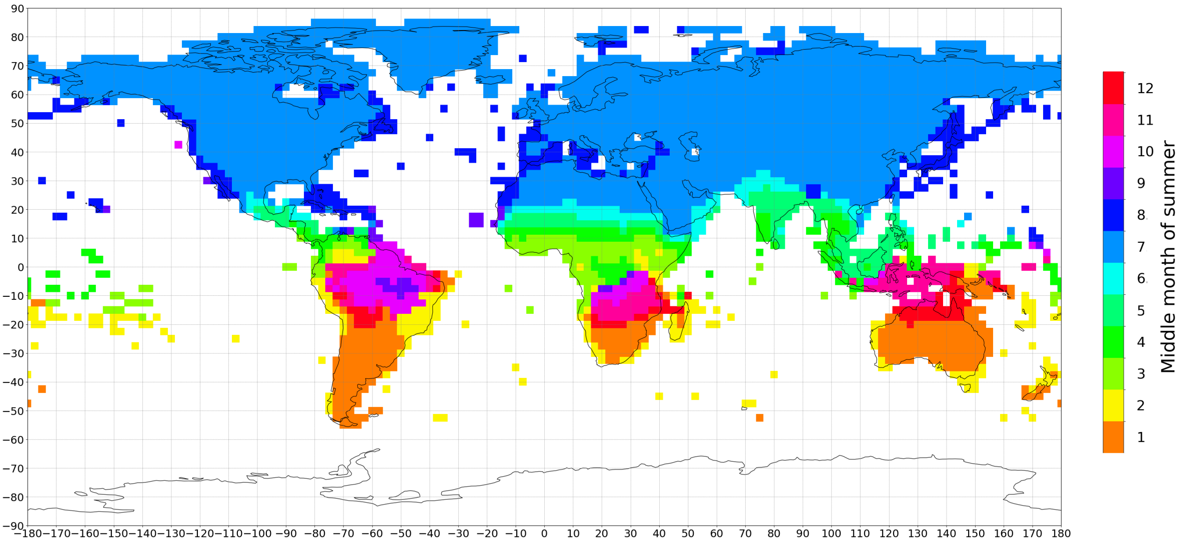
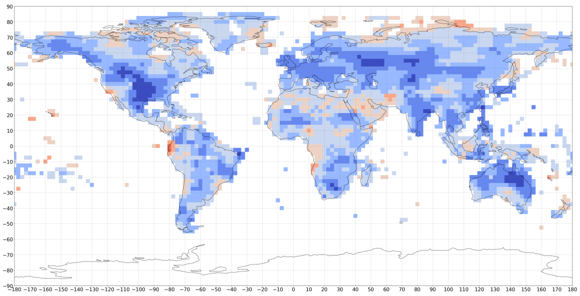
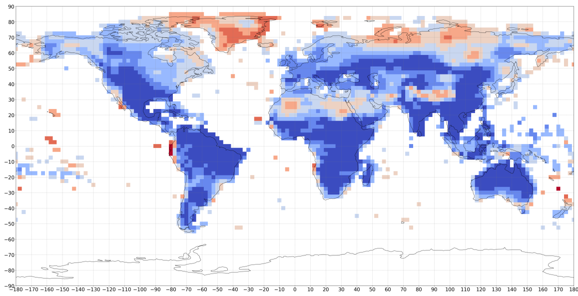


Figure 3: The middle month of the warmest 3 months (summer) at each grid point based on the CRU climatology (1901-2019). Note : 1 corresponds to January and so on.

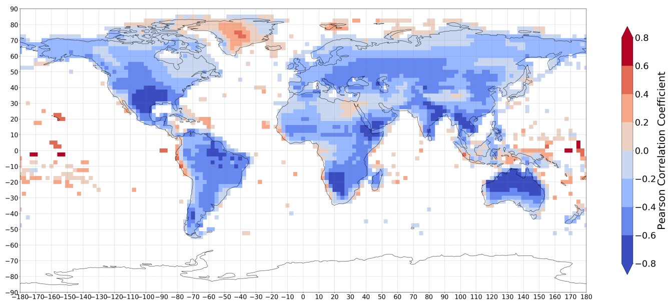
The correlation coefficients computed between yearly summer residuals of temperature and precipitation for the 20th century are presented in Fig.4. The coefficients are calculated separeately by using the observational dataset (CRU TS), the re-analysis dataset (ERA 20C) and each of the 40 CMIP5 model datasets. The results for correlation coefficients were not significantly affected by regression with global mean temperature for any datasets. However, the residuals were still used to compute correlation in order to remove any warming signals, especially towards the end of the century. It can be observed from Fig.4 that the mean summer temperature and precipitation are strongly negatively correlated over most land regions. However, there are also minimal exceptions such as some parts of Egypt and the Arabian Peninsula for example. It is interesting to observe that the correlation computed from the observation (Fig.4a) is generally less negative than the mean correlation from 40 CMIP5 models (Fig.4c), i.e. models tend to project a stronger anti-correlation compared to the observation. This was also noticed by Zscheischler and Seneviratne (2017). However, comparing correlation from ERA 20C reanalysis (Fig.4c) with the other two, shows that ERA 20C tend to project the strongest anti-correlation among the three, for most regions. In particular, the Southern hemisphere.



(a)



(b)



(c)

Figure 4: **Correlation between inter-annual summer mean residual temperature and precipitation for the historical period**
(a) Correlation computed from the CRU observational datasets of temperature and precipitation (1901-2019), **(b)** Correlation computed from the ERA20C reanalysis datasets of temperature and precipitation (1901-2010), **(c)** Average of correlation values computed from the 40 CMIP5 GCM datasets of temperature and precipitation (1901-2005)

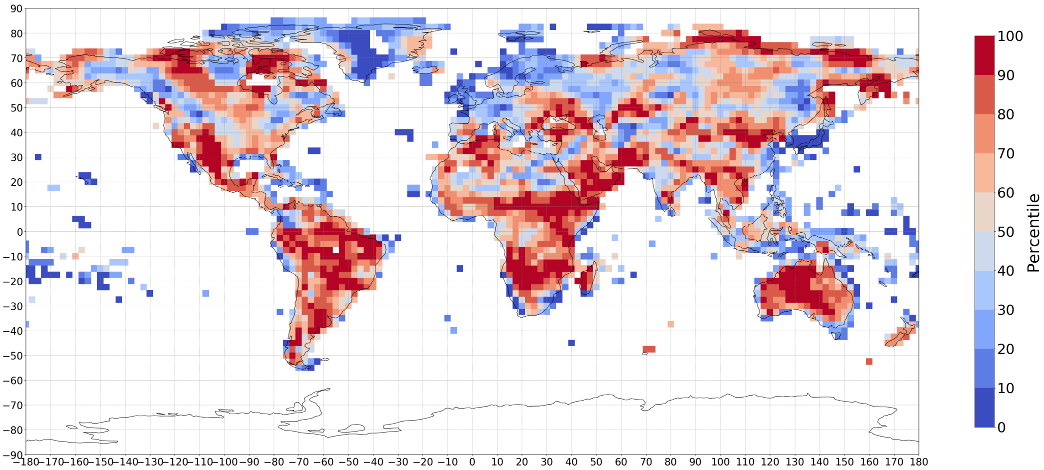
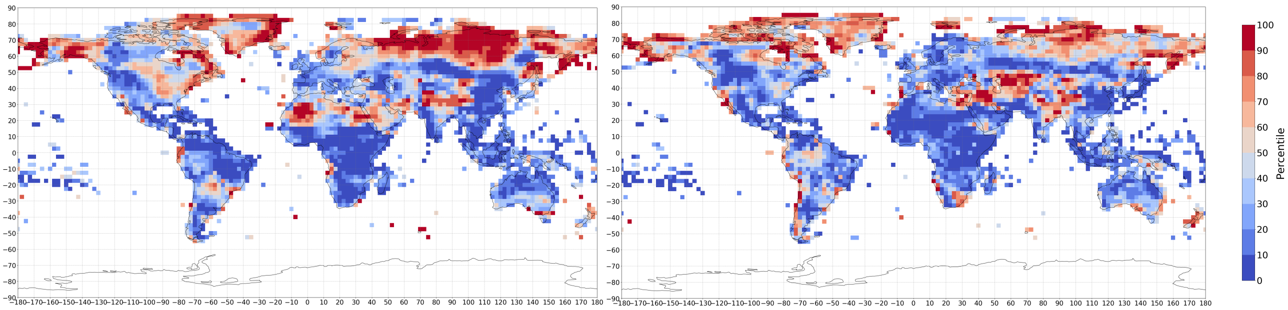


Figure 5: The percentile to which the correlation computed from CRU data (1901-2019) correspond to at each grid, with respect to a distribution of 40 correlation coefficients computed from 40 CMIP5 models



(a) (b)

Figure 6: **(a)** The percentile to which the correlation computed from ERA 20C data (1901-2010) correspond to at each grid, with respect to a distribution of 40 correlation coefficients computed from 40 CMIP5 models, **(b)** Same as in (a) for correlation coefficients computed from EC-Earth model data (1901-2005)

Fig.5 shows how the correlation coefficients computed from the CRU observational data compare with the distribution of correlation coefficients computed from the 40 member CMIP5 ensemble at each grid. The same is shown for ERA 20C data in Fig.6a. The plots show which percentile the correlation coefficients from CRU and ERA20C, respectively, correspond to, with respect to the distribution of 40 correlation coefficients from the CMIP5 ensemble. As observed from the correlation plots, we can see that the correlation from CRU tend to be in the higher 70+ percentiles of CMIP5 distributions for many regions, i.e. less negative. This is especially true for the global south. On the other hand, for ERA 20C, the coefficients are in the lower (< 30) percentiles in the global south. However, There are also regions in the global north where ERA 20C correlation is in the higher 90 percentiles. Especially around northern Russia Fig.6a. By statistically comparing the correlation coefficients obtained from ERA 20C with correlation coefficients from each of the CMIP5 models (using Fisher-Z test), we found that the 'EC-Earth' model shows comparable correlation on most grid points than any other model, i.e. 79% of total grid points where data is present (see Appendix. Table.2). Fig.6b shows percentiles that the correlation coefficient computed from EC-Earth correspond to at each grid point. It can be seen that they follow similar patterns as ERA 20C. However, there are strong exceptions too. For example, the coefficients are in the higher percentiles for ERA 20C around longitude 0° to 10° region in North Africa (Fig.6a). For EC-Earth however, the percentiles are in the lower range for this region (Fig.6b).

In the next step, we constrained each of the 40 CMIP5 models by using three tests, i.e. Fisher-Z test for independent correlation comparison, method by Zou for independent correlation comparison and Rémillard and Scaillet (2009)'s method for comparing two empirical copulas (See section 3). Fig.7 shows the number of 'good' models at each grid, based on the Fisher-Z test, which compared the correlation coefficients from CRU TS with that from CMIP5 models. Fig.8a and Fig.8b show the same results for Zou-method and Rémillard and Scaillet (2009)'s method, respectively. We observed that the Fisher-Z method and method of Zou provided identical results in constraining the models. No significant difference was observed between the number of constrained models, on any grid, between the Fisher method and the method of Zou. This was also further verified by comparing the mean frequency of hot and dry summers projected by the models constrained by the two methods separately, on each grid. This also showed no significant difference. Similarly copula method too gave nearly identical results to the other two methods. The copula method seemed to constrain slightly more number of models in general. Although the difference between the three methods were found to be insignificant. Moreover, copula method also had small patches were the test could not be performed, i.e. around the Gobi desert and Sahara (Fig.8b). This was due to the presence of gaps in the time series of temperature or precipitation at those points, which the test developed by Rémillard and Scaillet (2009) could not accommodate.

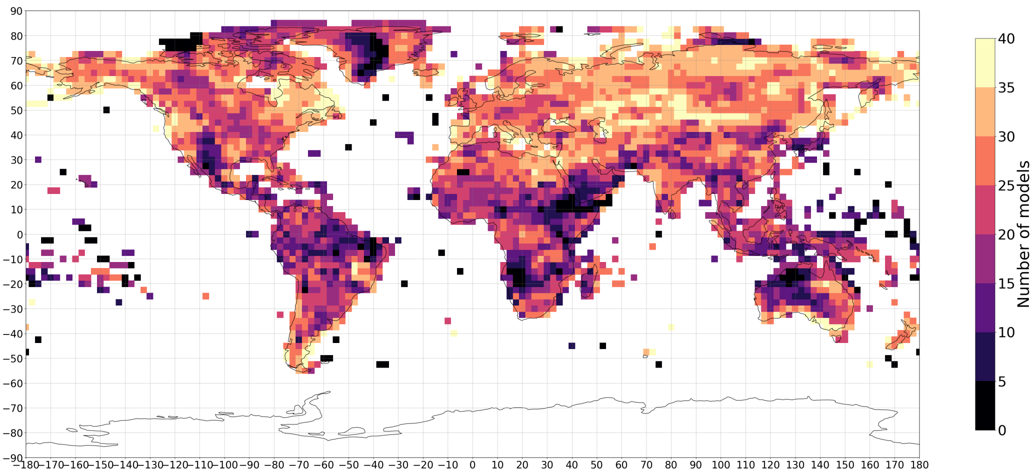
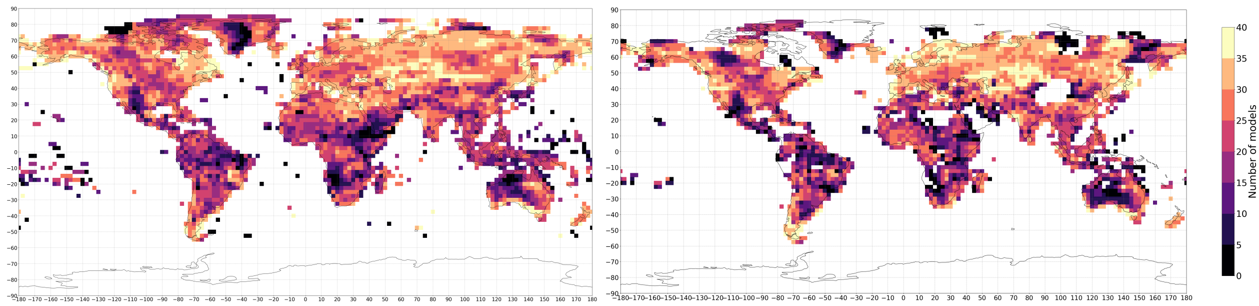


Figure 7: Number of CMIP5 models that show comparable correlation values to the correlation computed from CRU TS (i.e. number of 'good models'). Comparability is statistically determined by Fisher - Z test for comparing independent correlations



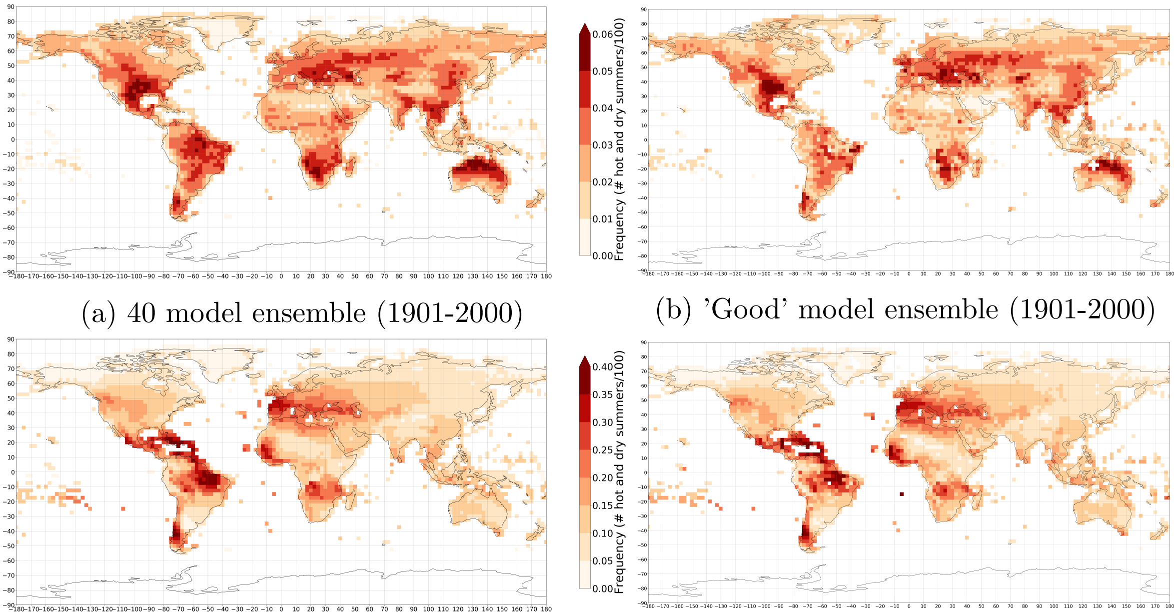
(a) (b)

Figure 8: **(a)** Number of CMIP5 models that show comparable correlation values to the correlation computed from CRU data (i.e. number of 'good models'). Comparability is statistically determined by Zou method for comparing independent correlations, **(b)** Number of CMIP5 models that have comparable dependence between temperature and precipitation to the dependence exhibited by temperature and precipitation from the CRU data (i.e. number of 'good models'). Comparability is statistically determined comparing the empirical copulas using the method developed by Rémillard and Scaillet (2009). The missing values over land are regions where the test could not be performed.

However, since all three methods performed similarly, and the copula test was at least 12 times more computationally expensive, we based rest of our analysis on the results of Fisher-Z test for comparison of correlations.

4.2 Results : Frequency projections

Here, we present the results of comparing mean frequency of hot and dry compound summer occurrences projected by the 40 model ensemble (containing all the CMIP5 models) and the 'good' model ensemble (containing only the 'good' CMIP5 models). The 'good' models referred to in this result are determined by statistical comparison of correlation values with those computed from CRU data, by using the Fisher-Z method for comparing independent correlations. The results are summarised separately for the future warming scenario of RCP 8.5 (2001-2100) and the historical scenario (1901-2000) in Fig.9.



(a) 40 model ensemble (1901-2000)

(b) 'Good' model ensemble (1901-2000)

(c) 40 model ensemble (2001-2100)

(d) 'Good' model ensemble (2001-2100)

Figure 9: **Mean frequency of compound hot and dry summer occurrences projected by two model ensembles for the 20th century (1901-2000) and the 21st century (2000-2100)**

(a) Mean frequency projected by model ensemble consisting of all the 40 CMIP5 models for the 20th century (1901-2000) **(b)** Mean frequency projected by model ensemble consisting of only the 'good' CMIP5 models for the 20th century (1901-2000) **(c)** Mean frequency projected by model ensemble consisting of all the 40 CMIP5 models for the 21st century (2001-2100) **(d)** Mean frequency projected by model ensemble consisting of only the 'good' CMIP5 models for the 21st century (2001-2100)

Note : Scale for (a) and (b) run from 0 to 0.06 while the scale for (c) and (d) run from 0 to 0.40. This happens because the frequency projections for compound hot and dry summer occurrences increase significantly from the 20th century to the 21st century under RCP 8.5

As can be seen from Fig.9a and Fig.9c, the projected mean occurrence frequency of compound hot and dry summers show considerable increase over all land domains between the two centuries. The scale of frequencies increased from (0 to 0.06) in the 20th century to (0 to 0.40) in the 21st century. In Fig.10 we show by how many folds this frequency increased at each grid point. While upper northern hemisphere generally shows slightly less increase, Western Europe shows significant projected increase. This was also observed by Zscheischler and Seneviratne (2017)) and is a cause for concern.

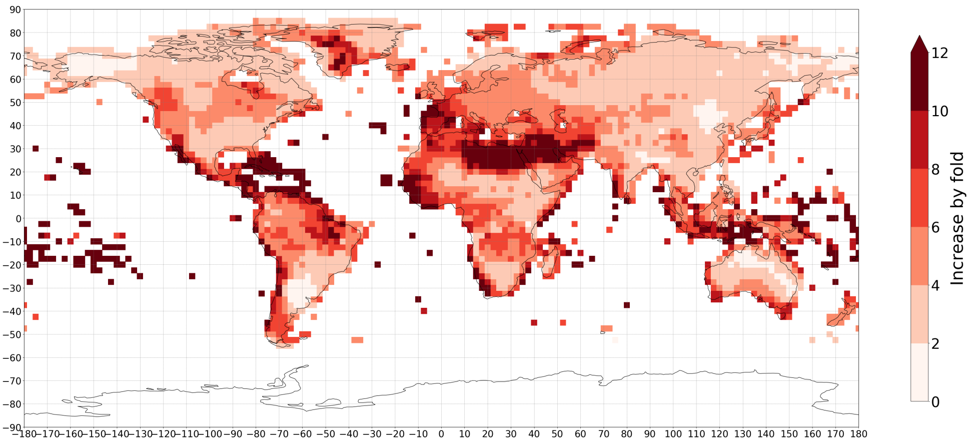


Figure 10: By how many folds the projected mean occurrence frequency of compound hot and dry summers increase between the 20th and the 21st centuries. Mean frequency is computed from ensemble consisting of all the CMIP5 models.

The main goal of this project was to compare the difference between the mean frequency of compound hot and dry summers projected by the all model ensemble and the 'good' model ensemble on a global scale. Fig.11 shows the relative difference in the projected frequency (i.e. [Mean frequency projected by 'good' models - Mean frequency projected by all models] / Mean frequency projected by all models) at each grid point, for the 20th century (1901-2000).

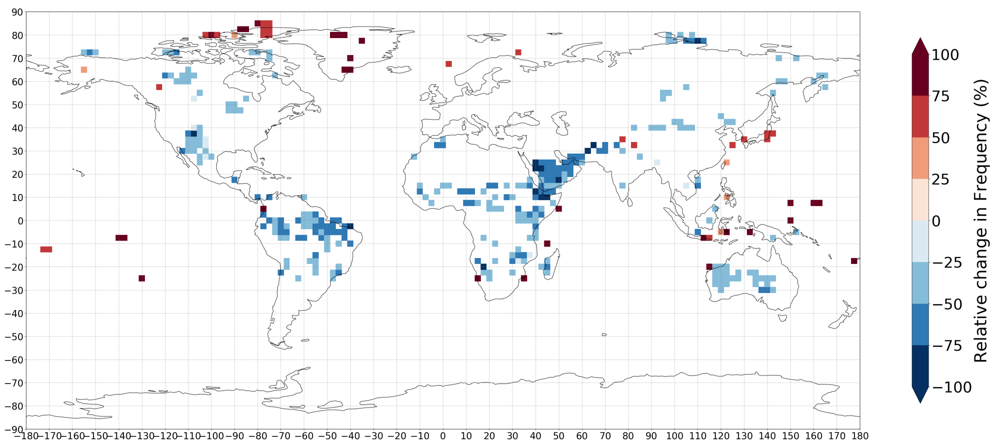


Figure 11: Relative difference between the mean frequency of compound hot and dry summers projected by the 'good' model ensemble and the all model ensemble for 20th century. Grid points where data is not available or the difference is not statistically significant are left as blank.
Relative difference is computed as (Mean frequency projected by 'good' models - Mean frequency projected by all models) / Mean frequency projected by all models, and expressed in percentages

For the 20th century, we only looked at points where the difference between the projections were statistically significant. In the majority of these points, the relative difference is negative, i.e. the 'good models' tend to project less occurrence frequency for compound hot and dry summers compared to all models. For the 21st century (2001-2100), the difference between the two ensembles were found to be field insignificant, i.e. less than 2% of grid points (with values) showed a significant difference. Therefore, no meaningful observations could be made.

4.3 Results : Other factors

In order to further explain some of the main results, we looked at possible limitations that may have been present in the data sets and the analysis that we employed. Fig.12 addresses the station number constraint present in the procurement of CRU observation data.

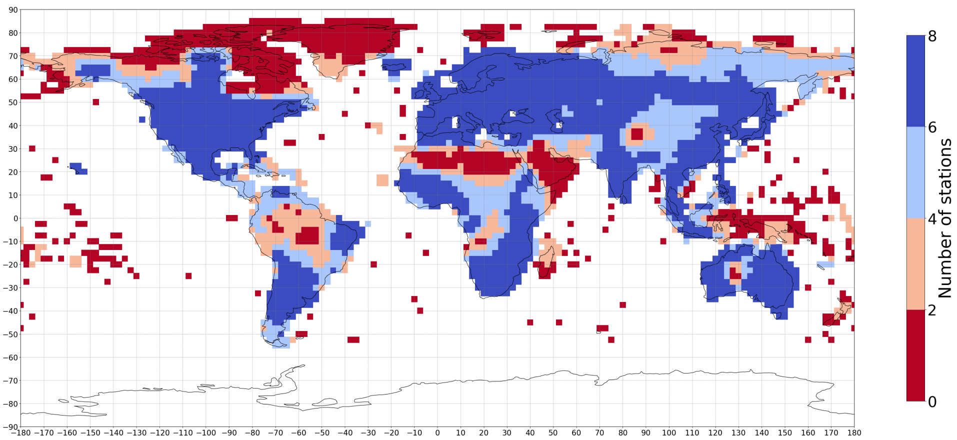
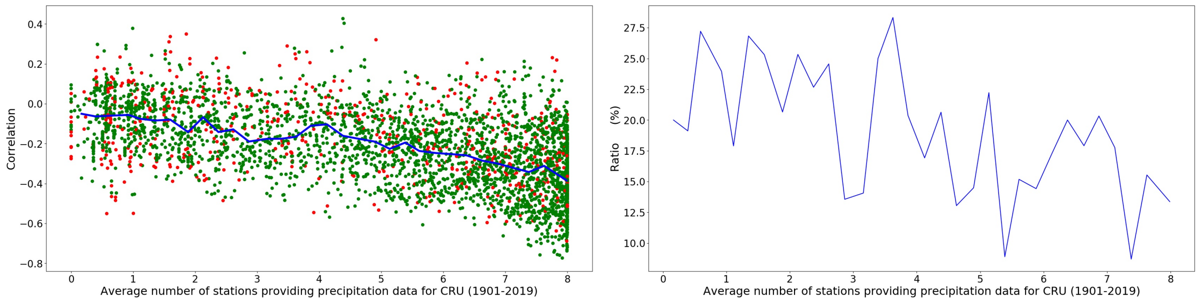


Figure 12: Number of observation stations providing monthly precipitation data for CRU TS, averaged over the period 1901-2019, at each grid

The lack of stations in collecting data was more apparent in the case of precipitation than temperature for CRU. However, both temperature and precipitation had, on average, maximum 8 stations and minimum 0 stations providing observations depending upon the region. Even though we tried to overcome this limitation by using GPCC data for precipitation, there was no significant difference in resulting correlation coefficients computed (correlation from CRU temperature and GPCC precipitation), when compared with the correlation computed from CRU alone. Even though GPCC had, on average, many more stations providing data, the spatial distribution of stations and their time evolution through the century follow the same sparse pattern as in Fig.12 (Sun et al. (2018), Fig.2). Therefore it was not anymore useful than the CRU data.

To see whether the lack of stations providing precipitation data has any significant impact on the dependence computed from the CRU data, we looked at Fig.13. In Fig.13a, the correlation values at all grid points, computed from the CRU observation data, are plotted against the average number of stations providing precipitation data for CRU at those corresponding points. The median line is also plotted. Fig.13b, also explores how many of the correlation values computed from CRU are outliers with respect to the 0.1 and 0.9 quantiles of correlation value distribution from the 40 CMIP5 models. The ratio of outliers to total number of points is plotted against the average number of stations providing the precipitation data. It can be observed that there is a clear downward trend in the median line for Fig.13a. This means that the correlations between residual mean summer temperature and precipitation computed from CRU, at grid points which do not have a severe station number constraint, tend to be slightly more negative than correlations computed at points with a station constraint. However, it is also to be noted that there are significantly more number of points in the regions where the station constraints are not severe, compared to points with few stations providing precipitation data on average. From Fig.13b we can also spot a reasonable downward trend. This indicate that the dependence computed from CRU observations tend to disagree with the general projections of CMIP5 models at those grid points more, where there is station number constraints for CRU, i.e. grid points, where the reliability of observation data could be in question.

Finally, we also explored the uncertainty associated with correlation computed from a bivariate normal distribution, due to small sampling size. Fig.14 shows the result of repeatedly sampling sets of 120 data points from the theoretical distribution with a true correlation of -0.2, and computing correlations. The results from 10000 samplings is portrayed. It is evident that the sample correlations shows a wide range from -0.5 to 0.1, even though the true correlation of the bivariate population was -0.2.



(a) (b)

Figure 13: **(a)**Pearson correlation coefficients computed between yearly mean summer temperature and precipitation (residual) over 119 years (1901-2019) from observational (CRU TS) data plotted against number of observation stations providing monthly precipitation data to CRU TS averaged over 119 years (1901-2019) at each grid. The outlier values of CRU based correlation coefficients are shown in red colour. The outliers are computed with respect to 0.1 and 0.9 quantiles of correlation coefficients from 40 distinct CMIP5 model data (1901-2005) at each grid. The corresponding inliers are represented in green. The median line is shown in blue. **(b)** The ratio of outliers (red) with total number of points (red+green) contained in bins of sizes of 0.25 stations for (a).

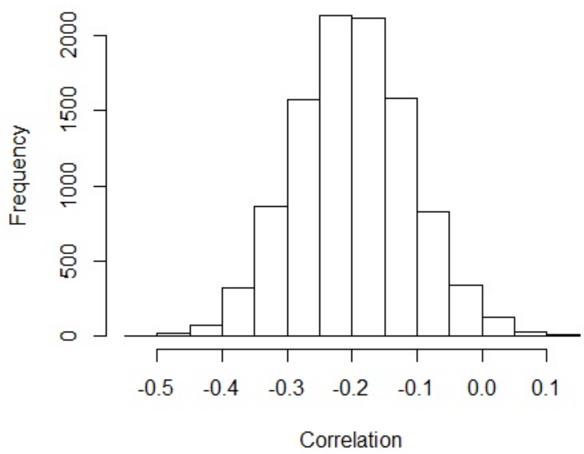


Figure 14: Histogram of correlation values computed from 10000 samplings of sets of 120 data points each. The samples are taken from a theoretical bivariate normal distribution with a true correlation of -0.2. Both the marginals in the theoretical distribution are considered to be standard normal, i.e. mean = 0 and standard deviation = 1

5 Discussion

Through this project, we were able to reassert many results observed by past researchers in the realm of compound hot and dry events. Moreover, we were able to test our own hypothesis regarding whether constraining CMIP5 global climate models for dependence between temperature and precipitation have any impact on their projections for compound hot and dry summer occurrence on a global scale.

The correlation between summer mean temperature and precipitation computed from CRU observation was generally observed to be less strongly negative than corresponding model projections (Fig.4,5). Zscheischler and Seneviratne (2017) had suggested that this could be due to the sparse distribution of stations providing data for CRU, leading to the correlations being less well constrained. We explored this in Fig.13a,13b, and observed that the assertion could be true to some extent. The trend in Fig.13b in particular is indicative of the fact that CRU correlations disagree more with model projections in regions with station number limitations. However, both these results are contingent upon the fact that there are far too many points representing higher number of stations compared to lower number of stations. Therefore, this asymmetry in distribution could also be contributing to the trend. Further exploration is required to establish the association.

The correlation computed from ERA 20C shows extreme diverging patterns in the south and in the north, towards the lower percentiles and the higher percentiles, respectively (Fig.4,6a). The EC-Earth global climate model was observed to project comparable correlation coefficients, to that projected by ERA 20C in general (Fig.6, Appendix Table.2). Berg et al. (2015) attributed this dependence pattern projected by the EC-Earth model to a combination of land-atmosphere and cloud-atmosphere feedback incorporated by the model (See section.1). Since both EC-Earth and ERA 20C are products of ECMWF, one could assume that they share some core model genealogy (See section.2.2). This could explain the observed similarity. However, there were also significant departures from similarity in some regions. This could be attributed to the fact that ERA 20C is shown to have performed poorly in terms of reproducing observations in some regions, especially outside Europe (Poli et al. (2016)). Therefore, utilising ERA 20C to constrain models for dependence is not recommended.

While comparing the correlation coefficients to constrain the models, the three statistical tests used were found to have given quite similar outputs (Fig. 7,8). This could be explained by the fact that the p-value constraints we imposed on each test is, by design, highly in favour of inclusion to 'good' models than exclusion as 'not good'. Therefore, the tests tend to include similar number of models at each grid. Moreover, the advantage the copula test have over the correlation comparisons is that the former can capture the tail dependence better than the latter. However, since we only have 119 data points at best for the bivariate distribution, the tail of the distribution is rather short. This could be a reason why the copula test fails to provide significantly different results from the other two.

The main result of this project showed that there was no significant difference between the projected mean frequency for compound hot and dry summers in future (2001-2100), under RCP 8.5, between the 'good' model ensemble and the all model ensemble. There were some points where this difference was significant for the 20^{th} century. In those points, the 'good models' tend to project lesser frequency for hot and dry summers on average compared to all models. Both these observations were in contradiction to our initial hypothesis. In the case of the 20^{th} century result, it can be observed that there exists a spatial correlation between the significant points of difference (Fig. 11) and points where there are severe station number limitations in CRU (Fig.12). This could mean that the only significant difference we observed in this case could be a consequence of the poorly constrained correlation values in observation, owing to the lack of quality observations. This makes it impossible to further associate significance to the observed result.

As for the lack of significant difference in projections of hot and dry event frequencies between the constrained and the non constrained models in general, we provide couple of explanations. One, it could be argued that the limited number of data points (119 at best and 105 at worst) available while computing correlation could mean that the observed correlation is not a reflection of the true correlation of the underlying population. This was demonstrated through a simple numerical experiment in Fig.14. This could mean that the constraints have certain limitations to this extent and it might be reflecting in the final result. Secondly, and most importantly, Bevacqua et. al.(2021) [*submitted*] have shown that precipitation is the significant component in determining the projected frequency of hot and dry compound summer occurrence for the 21st century under RCP 8.5 warming. This follows from the fact that the mean temperature is projected to rise over all land domains for the future, especially in the second half of the 21st century, essentially pushing most of the future summers to the hot extreme compared to the 20th century percentile thresholds. Therefore, the probability of occurrence for multivariate hot and dry compound summers will largely depend on the univariate statistics of precipitation. This could be a reason why we fail to observe a significantly high frequency of occurrence for the multivariate extreme projected by the 'good' model ensemble, which is constrained only for dependence between temperature and precipitation, and not for precipitation statistics.

6 Conclusion

6.1 Summary

In this project, we constrained an ensemble of 40 CMIP5 global climate models over land, with a constraint on the dependence between summer temperature and precipitation over each grid point, based on the 20^{th} century observation climatology (from CRU observation data). We used correlation as the dependence parameter between temperature and precipitation in the first stage of the project and empirical copula in the second stage. Fisher-Z test was employed to statistically compare the observed correlation with the correlation obtained from each of the CMIP5 model data sets. Bi-variate empirical copulas obtained from observation data of temperature and precipitation and that from each of the CMIP5 models were directly compared using the R package TwoCop. The results of comparing correlation and copulas were found to be nearly equivalent. Therefore, we based the rest of our analysis on the comparison of correlation values with Fisher-Z method and not the copula test. The empirical copula comparison test was also computationally much more expensive than the Fisher-z test. Subsequently, the mean occurrence frequency of compound hot and dry summers as projected by the constrained models ('good models') and all the models at each grid point was compared for the 20^{th} century and the 21st century (under RCP 8.5 warming) projections. We hypothesised that the 'good' models will, on average, project higher occurrence frequency for compound hot and dry summers, wherever temperature and precipitation were anti-correlated. The hypothesis was found not to be true, for results from the period 1901-2000 as well as the future period 2001-2100, under RCP 8.5 warming. The projected mean frequency difference between the ensembles were found to be insignificant for more than 98% of grid points in the future period. For the historical period, the good models were found to be projecting a lower frequency on average, at most of the significant points. We argue that the results, which are in contradiction to our initial hypothesis, could be a product of multiple factors. One being the lack of stations providing data for the observation, which leaves the observed correlation less well constrained. Moreover, researchers have recently shown that the future occurrence frequency of compound hot and dry summers may only depend on the univariate statistics of precipitation, as temperatures will be pushed to the hot extreme constantly under the high warming scenario of RCP 8.5. This essentially renders the constraint for dependence between temperature and precipitation ineffective, as only the variability in precipitation matters for the hot and dry event occurrence.

6.2 Limitations

This project suffers from a few limitations. One of them being the low number of samples used to compute dependence between temperature and precipitation. CRU data with 119 data points (1901-2019), for temperature and precipitation, is the longest time series that we have employed in this study in this regard. But we observed from Fig.14 that there is large uncertainty associated with the correlation coefficient, when only 120 points are sampled. This could suggest that the dependence parameter that we used to constrain the models may not be reliable. We tried to overcome this limitation by using empirical copulas to assess the dependence. However, the copula comparison test too provided nearly identical results. This, again could be a consequence of only having close to 100 data points, which results in a short tail and therefore, making it nearly impossible for copulas to detect any tail dependencies that the correlation cannot. Increasing the number of data points by including individual monthly means from each year was also not advisable. Because, that will result in intra-annual correlation dominating the dependence, while our focus is on constraining the models for their inter-annual dependence between the variables. Using other data sets, such as GPCC for observational data and ERA 20C as reanalysis data, and comparing the dependence patterns among the data sets was also not an ideal solution to the problem. Because we observed that, the dependence patterns computed from these data sets too were equally or even more unreliable than those computed from the CRU data. The ideal solution for this is to have reliable observation data, across regions, that span a longer time-frame. For this, however, we may have to wait for a few more decades to obtain observational data from the future as well.

Another limitation stems from what could also be considered as a philosophical issue associated with climate modelling. Our study involved constraining global climate models, which incorporate complex, coupled physical processes to arrive at climate projections for the historical period as well as the future. We used observational dependence from CRU data and projections of dependence from each of the models, in the historical period, to constrain the models and looked at their future projections of compound extremes. One implicit assumption we made here is that, the models that are better equipped at projecting the dependence well today, will also manage to do so in the future. This may not be true as climate regime shifts in the future, owing to anthropogenic climate change, might mean that the processes governing the dependence between variables are completely different then, compared to today. Therefore, even if the models, which are considered 'good' today for projecting the dependence, projected an increased frequency for occurrences of compound extremes in the future, we can never be assured that the result amounts to absolute physical certainty. There are also pitfalls in hoping to capture the spectrum of physical processes that may govern the dependence in future, by using an ensemble of CMIP global climate models. For one, many of these models share the same climate modelling ancestry, which reduces the spectrum of underlying processes significantly (Knutti et al. (2013)). And for two, due to their extremely complex nature, we do not have much clarity regarding the physical processes that govern these models (Baumberger et al. (2017)). The solution to this problem is the better understanding of physical processes that govern the climate systems and 'demystifying' the global climate models using novel methods such as interpretable AI. This is why, fundamental research into the physical science basis of climate change is extremely crucial, even today.

6.3 Outlook

The years 2020 and 2021 have been a challenging period for humanity across the planet, owing largely to the novel corona virus outbreak. The global nature and the associated complexity and unprecedented social responses to the pandemic has prompted scholars to look for potential parallels between tackling the climate crisis and the corona crisis (Hulme et al. (2020)). Additionally, in 2021, Europe in particular has experienced concurrent contrasting weather patterns of extreme precipitation and flooding in the west (Binnie and Abnett (2021)) and severe heatwaves with record high temperatures in the south-east, with Greece bearing the brunt of the heat (Kampouris (2021)). Eminent climate scientists have made public statements regarding the potential association with climate change for such extremely divergent weather patterns that are becoming more prevalent in Europe (Wanner (2021)). Evidences suggest that we are entering a time period where lives across borders are starting to face first hand experiences in the compounding nature of impacts that climate change can bring about. In this context, the 'Assessment Report 6' from IPCC, with 'Working Group 1 - The Physical Science Basis for Climate Change' scheduled to be published in August 2021 and the rest to follow within the next year, becomes extremely crucial in communicating the scientific facts and projections related to climate change. Well-known climate scientists such as Prof. Thomas Stocker have always believed in the role of science and scientists in shaping policy (Stocker and Blülle (2021)).

Under such circumstances, a project such as this becomes relevant in communicating the potential risks into the future, if societies continue to opt for a non-mitigating pathway such as RCP 8.5. As the present study demonstrated, the projected frequency of compound hot and dry summers will be significantly higher for the 21st century compared to the 20th century over all land domains. Even up to 10 folds in certain regions. This directly translate to a disproportionate and significantly higher risk, both societal and environmental. The project also was an attempt towards constraining global climate models in order to have projections of future compound event occurrences, which take the dependence between variables into account. This could potentially improve our predictions and therefore aid in future risk assessments and their reduction. Even though the project failed to establish a significant difference in frequency projections by the constrained model ensemble compared to the all model ensemble for the future, the work has produced both valuable insights and paved way for research directions that can reduce gaps in our understanding with respect to such compound extremes.

I would like to conclude this thesis by listing out the potential future research ideas and further developments that have come out of this study.

- Replicating the analysis using CMIP - Phase 6 (CMIP6) models and comparing the results with that of CMIP5 data.
- Analysing the role of individual variables in determining the occurrence probability of hot and dry summers for the future.
- Exploring the geneology of ERA 20C and identifying whether the physical processes that govern the model and the EC-EARTH model have many similarities.
- Studying the impacts that station number constraints have on observed dependence, by making use of other observation datasets of temperature and precipitation as well as other climate variables.
- Testing whether the empirical copula test is better suited, compared to Fisher-Z test or Zou test, to analyse dependence between marginals that have longer tails. This could be explored by performing all tests on actual climate data along with on distributions that are artificially modified to have a longer tail.

A Appendix

Table 2: List of CMIP5 models in decreasing order of corresponding 'good' grid points (in percentage) with respect to ERA 20C reanalysis data. The number of 'good' grid point are those which show comparable correlation coefficient to correlation coefficient computed from ERA20C at each grid. Correlations are statistically compared using Fisher-Z test.

Sr.No	Model Name	Good grids(%)	Sr.No	Model Name	Good grids (%)
1	EC-EARTH	71	21	CanESM2	54
2	FGOALS-g2	63	22	CSIRO-Mk3-6-0	54
3	GFDL-CM3	63	23	bcc-csm1-1-m	53
4	ACCESS1-3	61	24	GFDL-ESM2G	53
5	GISS-E2-H-CC	58	25	GFDL-ESM2M	52
6	HadGEM2-AO	58	26	CESM1-CAM5	52
7	HadGEM2-ES	58	27	MPI-ESM-MR	52
8	GISS-E2-R-CC	58	28	CNRM-CM5	52
9	HadGEM2-CC	57	29	IPSL-CM5B-LR	52
10	GISS-E2-H	57	30	IPSL-CM5A-MR	51
11	MRI-CGCM3	57	31	MPI-ESM-LR	51
12	CMCC-CESM	57	32	FIO-ESM	49
13	GISS-E2-R	57	33	CESM1-BGC	49
14	MRI-ESM1	56	34	IPSL-CM5A-LR	49
15	ACCESS1-0	56	35	CCSM4	49
16	BNU-ESM	56	36	CMCC-CM	49
17	bcc-csm1-1	55	37	CMCC-CMS	48
18	MIROC5	55	38	NorESM1-M	48
19	MIROC-ESM	54	39	NorESM1-ME	46
20	MIROC-ESM-CHEM	54	40	inmcm4	42

List of Figures

List of Tables

References

Allen, C. D., Macalady, A. K., Chenchouni, H., Bachelet, D., McDowell, N., Vennetier, M., Kitzberger, T., Rigling, A., Breshears, D. D., Hogg, E. T., et al. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest ecology and management*, 259(4):660–684.

Baumberger, C., Knutti, R., and Hirsch Hadorn, G. (2017). Building confidence in climate model projections: an analysis of inferences from fit. *Wiley Interdisciplinary Reviews: Climate Change*, 8(3):e454.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M. (2013). A description of the global land-surface precipitation data products of the global precipitation climatology centre with sample applications including centennial (trend) analysis from 1901–present. *Earth System Science Data*, 5(1):71–99.

Berg, A., Lintner, B. R., Findell, K., Seneviratne, S. I., van den Hurk, B., Ducharne, A., Chéruy, F., Hagemann, S., Lawrence, D. M., Malyshev, S., et al. (2015). Interannual coupling between summertime surface temperature and precipitation over land: Processes and implications for climate change. *Journal of Climate*, 28(3):1308–1328.

Binnie, I. and Abnett, K. (2021). As floods hit western europe, scientists say climate change hikes heavy rain. *Reuters*.

Buras, A., Rammig, A., and Zang, C. S. (2020). Quantifying impacts of the 2018 drought on european ecosystems in comparison to 2003. *Biogeosciences*, 17(6):1655–1672.

Ciais, P., Reichstein, M., Viovy, N., Granier, A., Ogée, J., Allard, V., Aubinet, M., Buchmann, N., Bernhofer, C., Carrara, A., et al. (2005). Europe-wide reduction in primary productivity caused by the heat and drought in 2003. *Nature*, 437(7058):529–533.

DAMOCLES-MoU (2018). Understanding and modeling compound climate and weather events (cost-damocles). *COST Action-CA17109*.

Diffenbaugh, N. S. and Ashfaq, M. (2010). Intensification of hot extremes in the united states. *Geophysical Research Letters*, 37(15).

ECMWF (2021). Era20c-access. *ERA 20C,https://apps.ecmwf.int/datasets/data/era20c-moda/levtype=sfc/type=an/, Dataset accessed 2021-02-15 at 10.00.*

Field, C. B., Barros, V., Stocker, T. F., and Dahe, Q. (2012). *Managing the risks of extreme events and disasters to advance climate change adaptation: special report of the intergovernmental panel on climate change.* Cambridge University Press.

Fisher, R. A. (1921). On the'probable error'of a coefficient of correlation deduced from a small sample. *Metron*, 1:1–32.

Flannigan, M. D., Krawchuk, M. A., de Groot, W. J., Wotton, B. M., and Gowman, L. M. (2009). Implications of changing climate for global wildland fire. *International journal of wildland fire*, 18(5):483–507.

Geirinhas, J. L., Russo, A., Libonati, R., Sousa, P. M., Miralles, D. G., and Trigo, R. M. (2021). Recent increasing frequency of compound summer drought and heatwaves in southeast brazil. *Environmental Research Letters*, 16(3):034036.

GISTEMP-Team (2021). Giss surface temperature analysis (gistemp) version 4. *NASA Goddard Institute for Space Studies,https://data.giss.nasa.gov/gistemp/,Dataset accessed at 2020-12-15.*

Harris, I., Osborn, T. J., Jones, P., and Lister, D. (2020). Version 4 of the cru ts monthly high-resolution gridded multivariate climate dataset. *Scientific data*, 7(1):1–18.

Hulme, M., Lidskog, R., White, J. M., and Standring, A. (2020). Social scientific knowledge in times of crisis: What climate change can learn from coronavirus (and vice versa). *Wiley interdisciplinary reviews. Climate change.*

Kampouris, N. (2021). Greece records highest temperature in its history in unprecedented heatwave. *Greek Reporter.*

Knutti, R., Masson, D., and Gettelman, A. (2013). Climate model genealogy: Generation cmip5 and how we got there. *Geophysical Research Letters*, 40(6):1194–1199.

Kornhuber, K., Osprey, S., Coumou, D., Petri, S., Petoukhov, V., Rahmstorf, S., and Gray, L. (2019). Extreme weather events in early summer 2018 connected by a recurrent hemispheric wave-7 pattern. *Environmental Research Letters*, 14(5):054002.

Lemus-Canovas, M. and Lopez-Bustins, J. A. (2021). Assessing internal changes in the future structure of dry–hot compound events: the case of the pyrenees. *Natural Hazards and Earth System Sciences*, 21(6):1721–1738.

Lenssen, N. J., Schmidt, G. A., Hansen, J. E., Menne, M. J., Persin, A., Ruedy, R., and Zyss, D. (2019). Improvements in the gistemp uncertainty model. *Journal of Geophysical Research: Atmospheres*, 124(12):6307–6326.

Leonard, M., Westra, S., Phatak, A., Lambert, M., van den Hurk, B., McInnes, K., Risbey, J., Schuster, S., Jakob, D., and Stafford-Smith, M. (2014). A compound event framework for understanding extreme impacts. *Wiley Interdisciplinary Reviews: Climate Change*, 5(1):113–128.

Manning, C., Widmann, M., Bevacqua, E., Van Loon, A. F., Maraun, D., and Vrac, M. (2019). Increased probability of compound long-duration dry and hot events in europe during summer (1950–2013). *Environmental Research Letters*, 14(9):094006.

Martin-Vide, J. and Gomez, L. (1999). Regionalization of peninsular spain based on the length of dry spells. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 19(5):537–555.

Martius, O., Pfahl, S., and Chevalier, C. (2016). A global quantification of compound precipitation and wind extremes. *Geophysical Research Letters*, 43(14):7709–7717.

Mazdiyasni, O. and AghaKouchak, A. (2015). Substantial increase in concurrent droughts and heatwaves in the united states. *Proceedings of the National Academy of Sciences*, 112(37):11484–11489.

Messori, G., Bevacqua, E., Caballero, R., Coumou, D., De Luca, P., Faranda, D., Kornhuber, K., Martius, O., Pons, F., Raymond, C., et al. (2021). Compound climate events and extremes in the midlatitudes: Dynamics, simulation, and statistical characterization. *Bulletin of the American Meteorological Society*, 102(4):E774–E781.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., Laloyaux, P., Tan, D. G., Peubey, C., Thépaut, J.-N., et al. (2016). Era-20c: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, 29(11):4083–4097.

Rémillard, B. and Scaillet, O. (2009). Testing for equality between two copulas. *Journal of Multivariate Analysis*, 100(3):377–386.

Ridder, N. N., Pitman, A. J., and Ukkola, A. M. (2021). Do cmip6 climate models simulate global or regional compound events skillfully? *Geophysical Research Letters*, 48(2):e2020GL091152.

Salameh, A. A., Gámiz-Fortis, S. R., Castro-Díez, Y., Abu Hammad, A., and Esteban-Parra, M. J. (2019). Spatio-temporal analysis for extreme temperature indices over the levant region. *International Journal of Climatology*, 39(15):5556–5582.

Schulzweida, U., Kornblueh, L., and Quast, R. (2006). Cdo user's guide. *Climate data operators, Version*, 1(6):205–209.

Sharma, S. and Mujumdar, P. (2017). Increasing frequency and spatial extent of concurrent meteorological droughts and heatwaves in india. *Scientific reports*, 7(1):1–9.

Singh, D., Tsiang, M., Rajaratnam, B., and Diffenbaugh, N. S. (2014). Observed changes in extreme wet and dry spells during the south asian summer monsoon season. *Nature Climate Change*, 4(6):456–461.

Stocker, T. and Blülle, E. (01-Jun-2021). In der schweiz herrscht immer noch ein gewisser grundegoismus. *Republik*.

Sun, Q., Miao, C., Duan, Q., Ashouri, H., Sorooshian, S., and Hsu, K.-L. (2018). A review of global precipitation data sets: Data sources, estimation, and intercomparisons. *Reviews of Geophysics*, 56(1):79–107.

Taylor, K. E. (2009). A summary of the cmip5 experiment design. *http://cmip-pcmdi.llnl. gov/cmip5/docs/Taylor_CMIP5_design. pdf.*

Vogel, J., Paton, E., Aich, V., and Bronstert, A. (2021). Increasing compound warm spells and droughts in the mediterranean basin. *Weather and Climate Extremes*, 32:100312.

Vogel, M. M., Zscheischler, J., Wartenburger, R., Dee, D., and Seneviratne, S. I. (2019). Concurrent 2018 hot extremes across northern hemisphere due to human-induced climate change. *Earth's future*, 7(7):692–703.

Wanner, H. (25-Jul-2021). Die zwei probleme des europäischen klimawandels. *Der Bund*.

Zou, G. Y. (2007). Toward using confidence intervals to compare correlations. *Psychological methods*, 12(4):399.

Zscheischler, J. and Fischer, E. M. (2020). The record-breaking compound hot and dry 2018 growing season in germany. *Weather and climate extremes*, 29:100270.

Zscheischler, J., Martius, O., Westra, S., Bevacqua, E., Raymond, C., Horton, R. M., van den Hurk, B., AghaKouchak, A., Jézéquel, A., Mahecha, M. D., et al. (2020). A typology of compound weather and climate events. *Nature reviews earth & environment*, 1(7):333–347.

Zscheischler, J., Michalak, A. M., Schwalm, C., Mahecha, M. D., Huntzinger, D. N., Reichstein, M., Berthier, G., Ciais, P., Cook, R. B., El-Masri, B., et al. (2014). Impact of large-scale climate extremes on biospheric carbon fluxes: An intercomparison based on mstmip data. *Global Biogeochemical Cycles*, 28(6):585–600.

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

Zscheischler, J., Westra, S., Van Den Hurk, B. J., Seneviratne, S. I., Ward, P. J., Pitman, A., AghaKouchak, A., Bresch, D. N., Leonard, M., Wahl, T., et al. (2018). Future climate risk from compound events. *Nature Climate Change*, 8(6):469–477.

Acknowledgement

I am extremely grateful to numerous people, without whom this master thesis project would not have been possible. Foremost, I would like to extend my deepest gratitude to Dr. Jakob Zscheischler, whose supervision and guidance was everything that I could have hoped for. The weekly meetings and constant interactions with him always helped me in numerous ways. I also thank Prof. Olivia R. Martius for her guidance at the important stages of the project. Undoubtedly, those interactions have substantially improved the quality of this work. I am also thankful to Prof. Fortunat Joos for his kind participation in the examination and enlightening questions.

I would also like to acknowledge the help and support of the members of CompoundX group, for their participation in many presentations that I gave over the course of the year, and for providing me with suggestions and feedback on both style and content. The weekly meetings with Andreia, Aris, Elisabeth, Emanuele, Lily, Mohit and Natacha have been a wonderful and invaluable educational experience for me to say the least. I also thank Lucas for the all the help during the limited time that we shared the office space at KUP. Dr. Gunnar Jansen's technical help and support was instrumental, without which I could not have completed this study. I would like to acknowledge each and every member of OCCR and KUP at the University of Bern who provided me with this opportunity.

Finally, I thank my family and friends in Switzerland as well as in India, who supported me wholeheartedly during the past two years. I am indebted to each and everyone of you.

Declaration of consent

on the basis of Article 30 of the RSL Phil.-nat. 18

Name/First Name: Sivaraj Sidharth

Registration Number: 19-100-114

Study program: M Sc. in Climate Sciences

Bachelor ☐ Master ☑ Dissertation ☐

Title of the thesis: Constraining the occurrence probability of compound hot and dry summers in CMIP5 models by using an observational constraint on the dependence between temperature and precipitation

Supervisor: Dr. Jakob Zscheischler

I declare herewith that this thesis is my own work and that I have not used any sources other than those stated. I have indicated the adoption of quotations as well as thoughts taken from other authors as such in the thesis. I am aware that the Senate pursuant to Article 36 paragraph 1 litera r of the University Act of 5 September, 1996 is authorized to revoke the title awarded on the basis of this thesis.

For the purposes of evaluation and verification of compliance with the declaration of originality and the regulations governing plagiarism, I hereby grant the University of Bern the right to process my personal data and to perform the acts of use this requires, in particular, to reproduce the written thesis and to store it permanently in a database, and to use said database, or to make said database available, to enable comparison with future theses submitted by others.

Bern, 06/07/2021

Place/Date

Signature