Climate Data for Physical Risk Assessment in Finance

Alexis Tantet
LMD/IPSL, Ecole polytechnique, Palaiseau, France

Peter Tankov
CREST/ENSAE, Palaiseau, France

Abstract

Financial institutions are recognizing the importance of physical climate-related risks and are expected to disclose the presence of such risks in their portfolios. The assessment of physical risks is based on climate data and in particular with the advent of climate change, it will increasingly be based on the forward looking climate scenarios. Most financial actors, especially in the banking sector, presently lack the knowledge and tools required to access and use the climate datasets. The purpose of this document is to review the main sources of climate data, to provide an access point to this field for finance professionals.

A word of caution

The objective of this report is to give a brief overview of the type of climate data that may be used for physical climate risk assessment. All data sources mentioned here are prone to errors due to measuring devices precision, resolution, assumptions, etc. For a proper use of these data, we highly recommend interacting with scientists who have produced, or worked with, the data. Moreover, the degree of uncertainty associated to future climate, especially in high CO2 emission scenarios is such that precise evaluation of long-term climate-related risks is impossible, and one can at best hope to identify the most severely affected assets and the order of magnitude of impacts.

1 Introduction

According to the recent report by Hubert, Evain, and Nicol 2018, financial actors currently lack tools to assess how physical climate events could affect their assets, yet they are increasingly expected to disclose climate risk in their portfolios, in accordance with the recommendations of the Task Force on Climate-related Financial Disclosures (TCFD).

Historically, risk assessment has relied mainly on some measure of damages, rather than on the climatic origin of the damage itself. As our exposure to various kinds of risks increases and since climate change is expected to induce changes in the physical climate risk in the world (Stocker et al. 2013), physical climate risk assessments will increasingly need to rely on climate data.

The focus of this report is to give a brief overview of the climate datasets that may be useful for physical climate risk assessment studies. Economic, sociological and anthropological data on the assets and populations of interest, as well as the risk assessment methodology is not discussed.

After a quick primer on physical risk analysis in the next session, we introduce general guidelines to help select a dataset in Section 3. In Section 4, the main types of climate data-sources are presented. Then, Section 5 introduces downscaling, an often unavoidable, yet sensitive, step to translate climate data into physical risk. Finally, Section 6 warns about model biases and uncertainty. A fully worked-out example related to heat waves in the Paris area is provided in Section 7. Section 7.2.2 gives a few links to web pages that may prove useful in finding an appropriate dataset. A glossary of main climate-related terms is provided in the Appendix for reader's convenience. Additional references and data sources are given throughout the text.

2 A primer on physical risk analysis

Physical climate risks are the risks of destruction of assets and / or disruption of operations, trade routes, supply chains and markets by weather events related to climate change and climate variability. One may distinguish chronic climate risks related to slowly evolving phenomena, like sea-level rise, temperature rise, changing precipitation patterns etc., and the acute climate risks related to extreme weather phenomena (floods, extreme drought, heatwaves, wildfires, hurricanes), whose frequency and severity may increase with the onset of climate change.

The physical climate risk is a combination of three factors:

- The climate hazard: specific weather pattern or event whose frequency or severity may change;

- The exposure: presence of an asset or a system which may be affected by the climate hazard;

- The vulnerability: extent to which the asset or system may suffer from exposure to the climate hazard.

Climate data is necessary to evaluate the first factor, that is, the type, frequency and severity of expected climate hazards. Evaluating the two other factors requires access to asset level data: the geographical position and the specific characteristic of exposed assets or systems. Such data is not discussed in this report.

The evaluation of physical risks typically starts with the evaluation of climate hazards expected in the geographical area of interest. The next step is to identify the potential impact of these hazards on the concrete economic objects of interest: assets (factories, infrastructure, agriculture), transport links, supply links, markets, economic environment etc. These are then translated into estimates of balance sheets of specific companies and then possibly into the value of financial assets such as stocks and bonds.

The characterization of climate exposures requires different scales of climate information, from global to local level. Climate models provide a resolution of about 100km; they can be downscaled to lower resolutions either dynamically or statistically. Regional information on non-climatic components (relief etc.) should also be used.

Finally, one needs to keep in mind that there are large uncertainties both about the local response of the earth system to the radiative forcing associated with increased concentrations of greenhouse gases (GHGs) and about the capacity of the economic system to mitigate and adapt to changing climate. Internal variability of the climate and inter-model uncertainty are dominant in the short-term, while factors such as the level of GHG emissions are dominant at longer time scales.

3 Choosing a dataset

Recognizing that no dataset is perfect and that risk estimates should always be taken with care, several characteristics can guide the choice of a dataset: the period covered, whether historic or future; the spatial coverage; the presence of biases; the spatio-temporal scale of interest; the climate variable; the type of hazard (chronic or acute).

Past data vs. forward looking data Assessing a risk based on past data relies on the assumption that the system may be considered statistically stationary on time scales relevant for the study and that the available data allows one to sufficiently sample the underlying physical phenomena (in terms of frequency, duration and quality).

Considering that the state of the climate system varies on a continuous range of time scales, the statistical stationarity assumption may be infringed due to (i) low-frequency variability — variability on time scales longer that than the length of the time series from which risk is assessed, and sufficiently large (e.g. in terms of variance) to impact the risk — (ii) climate change — associated with changes in the forcing of the climate system, in particular, anthropogenic climate forcing through GHG emissions, land-use change, etc.

The climate change effects become material for horizons starting from 20–30 years ahead, but climate variability may manifest itself at shorter horizons (5–10 years).

Climate models vs. integrated assessment Climate models produce projections of future climate corresponding to specific scenarios of future economic activity (CO2 emissions, pollutants, land use etc.). These scenarios are created with economic reasoning, however the feedbacks of the climate system onto the economic system (such as the increase in mitigation actions after the widespread adverse consequences of climate change start to be felt) are not modelled. By contrast, integrated assessment models aim to describe jointly the evolution of the climate system and the economic system, sometimes with two-way feedbacks. Unfortunately, most integrated assessment models to date only contain a very basic climate module which may not even be based on the equations of physics, and therefore are not realistic enough to be used for physical risk analysis. These models are therefore not considered in this report, which focuses on climate models/data sets.

The Climate Data Guide[1] provides a starting point to get a quick overview of available climate datasets and their characteristics. It is not, however, exhaustive.

4 Selected climate data

In this section, we describe main types of climate data that may be used for risk assessment. We have ordered them by increasing dependence on modeling. We also focus on the following non-exhaustive list of key climate variables related to hazards: temperature, drought, precipitation, sea level and winds.

4.1 Observations

Observations may be used to assess past climate conditions and relate them to historical damages. Moreover, if the environmental sub-system of interest may be considered statistically stationary, observations may also be used to assess future risks.

Observational data being upstream of the modeling chain, they offer the advantage over reanalyses and projections to be less dependent on model assumptions and errors. In situ observations, such as from weather stations, tend to be as close as possible to the object studied, although observational devices need to be calibrated and never give a perfect representation of the measured variables. Teledetection from the ground or from space via satellites tend to offer a greater spatial (vertical and/or horizontal) coverage. However, they also tend to require more data processing and calibration than in situ observations.

Both in situ and teledetected data may be further processed to facilitate their use. For instance, NOAA's MLOST, NASA's GISTEMP, and the UK's HadCRUT datasets[2] provide a compilation of in situ observations on a regular grid. One should keep in mind that the strategy used to compile and interpolate the data impacts its quality.

[1] https://climatedataguide.ucar.edu/

[2] National Center for Atmospheric Research Staff (Eds). Last modified 25 Mar 2014. "The Climate Data Guide: Global Temperature Data Sets: Overview & Comparison Table." Retrieved from `https://climatedataguide.ucar.edu/climate-data/global-temperature-data-sets-overview-comparison-table`.

In any case, the *representativity* of the data should always be questioned. For instance, wind observations from a weather station on grass land may not be representative of the wind conditions over a nearby forest. On the other hand, the wind averaged over a 100km large grid box may not be representative of the wind conditions at a particular location within the box if the wind conditions vary greatly within this box.

4.2 Reanalyses

4.2.1 Motivation

Purely observational datasets are limited by their spatial (horizontal and vertical) and temporal coverage, and by the number of observed variables. This is particularly true in sparsely populated, economically little developed regions. On the other hand, while numerical weather predictions give a complete description of the atmosphere, they diverge from the future state of the climate system within a few days, due to combined observational errors and irreducible chaotic dynamics, and to model errors.

To alleviate this issue, one approach could be to interpolate variables of interest. Whatever the interpolation method used, there is no reason for the interpolated data to represent well the state of the climate system. In particular, the interpolated state may not be consistent with the physical laws governing the climate system. For instance, the important role of topography or of land-ocean contrasts may be taken into account.

Another approach that has proved particularly efficient for numerical weather prediction and for climate studies is to produce analyses by assimilating observational data into simulations. In brief, given present observations, data assimilation seeks to modify the past state of a climate model simulation in order for the present climate state predicted by model integration to best reproduce observations.

Data assimilation can thus be seen as a way to seamlessly interpolate observational data that is consistent with physical equations as encoded in a climate model. Yet, due to the sparsity of observations and to the strong dependence of analyses on model assumptions, data assimilation is better seen as a methodology to constrain model simulations with observations.

As a climate model, or a data assimilation method, is improved and as more observations become available, past analyses become deprecated. For that reason, new analyses are produced over the past record with the new version of the climate model and with the integration of new observational datasets. The resulting dataset is thus called a *reanalysis*. Another advantage of reanalysis datasets is that they provide a number of variables represented in, or diagnosed from, the model that may be difficult to observe.

The quality of a reanalysis strongly depends on the density and quality of observations injected in the data assimilation system. The sparser the observations at a given location, the less constrained is the model used to produce the reanalysis. This means that the data is more prone to model errors and more chaotic. This last point deserves attention. Indeed, risk assessment may require climate data that is synchronized with observed damages. However, integrating sparse observations in a reanalysis tends to constrain large scale dynamics only. This means that, while the statistical properties of small-scale processes may be well resolved by a reanalysis, chaotic dynamics may prevent considering the state of the reanalysis at a particular moment to be representative of the state of the climate system at that moment and at these scales.

4.2.2 Available reanalyses

The reanalyses.org project provides an overview of most available reanalyses. Atmospheric reanalyses may be divided into lower-resolution global reanalyses and higher-resolution regional reanalyses. State-of-the-art global reanalyses include ECMWF's ERA5 (Copernicus Climate Change Service (C3S) 2017), NCEP's CFSR (Saha et al. 2010), NASA's MERRA-2 (Gelaro et al. 2017) and JMA's JRA-55 (Kobayashi et al. 2015). A detailed comparison of the design and performance of these reanalyses is given by Fujiwara et al. 2017 and (Long et al. 2017). See also Buizza et al. 2005 and Wang et al. 2011.

Let us also mention reanalyses for the entire 20th century, such as ECMWF's ERA-20C (Poli et al. 2013), although the latter should be used with great care, since the observational constraints weaken as one goes further back in time (this is particularly true before the early 80s, the beginning of the satellite era). As an example of regional reanalysis, let us mention the COSMO reanalyses over Europe (Bollmeyer et al. 2015; Wahl et al. 2017). Such regional reanalyses may improve the representation of small scales — which may in turn impact large scales — relevant for impact studies.

4.3 21st century simulations

4.3.1 Motivation

Observations and reanalyses only provide information on past climate conditions. Due to the chaotic nature of the climate system, the exact future states of the climate may not be predicted. Much can be said, however, about the future evolution of statistics of the climate system, and, in particular, about physical climate risk. Due to the impact of low-frequency variability and changes in the forcing due to anthropogenic climate change, statistics of some key climate variables have changed and are expected to change further (Taylor, Stouffer, and Meehl 2012). This is particularly true regarding surface temperature and sea level.

While low-frequency variability and anthropogenic climate change are both sources of non-stationarity, the two do not have the same implications, as far as risk assessment is concerned. The long term — say over a period of several decades — response of the statistics of the climate system, is mainly constrained by changes in the radiative forcing associated with increasing greenhouse gases. Simulating this response thus depends first on the socio-economic scenario followed by current and future emitting countries, rather than on the initial state of the climate system. Such simulations are referred to as *projections*. On the other hand, decadal predictions is an emerging field that attempts to predict the slow evolution of the climate system associated with low-frequency variability. In this case, information about the initial state of the climate system is used, although most of it is lost. Whether this information is useful to impact studies is still being investigated and is not further discussed in this report.

4.3.2 Projections and scenarios

Global 21st century projections are structured by the the WCRP Coupled[3] Model Intercomparison Project (CMIP). CMIP's 5th phase (Taylor, Stouffer, and Meehl 2012) contributed to IPCC's 5th Assessment Report (AR5), while CMIP's 6th phase (Eyring et al. 2016) is contributing to IPCC's 6th Assess-

[3] While weather models tend to represent dynamically the atmosphere only, the long-term evolution of the climate system strongly depends on the interaction of the atmosphere with other components of the Earth system. Thus, climate projections rely on general circulation models coupling the atmosphere with oceans, sea ice, land surfaces, bio-geo-chemistry, etc.

ment Report (AR6). As the resolution of models increases, together with the number of processes represented, projections are expected to improve from one phase to the next. As more data becomes available, processing this data for impact studies may, however, prove challenging.

The main source of uncertainty in projections is the scenario used to force the models. Since the 5th phase of CMIP, Radiative Concentration Pathways (RCPs) are used to impose changing greenhouse gases concentrations in climate models. It is these changing concentrations that are responsible for changes in the radiative forcing eventually resulting in global warming. In parallel, different socio-economic scenarios leading to a particular RCP are developed by economists and social scientists.

In addition to scenario-related uncertainties, model errors as well as spread due to intrinsic variability also impact the projections. Since all models have errors and since projections are not constrained by observations, it is important to (i) be aware of model biases and (ii) use multi-model ensembles. Regarding this last point, it is indeed known that aggregating information from several models allows one to improve projection skills compared to that of a particular model. Moreover, multi-model ensembles allow one to partially estimate model errors and the spread due to intrinsic variability.

CMIP data may be found on the ESGF data portal.

4.3.3 Regional projections

Computational resources limit the resolution of global climate projections. For these reasons, global projections may not be suited for risk assessment. This is why Regional Climate Models (RCMs) are sometimes used. The latter resolve a limited area, allowing to increase the resolution model and the number of resolved processes, for the same computational resource. To produce regional projections, RCMs are forced at their boundaries[4] by projections from global models.

Just as the CMIP project for global projections, the CORDEX project organizes regional climate projection efforts. It defines standard domains over multiple regions as well as standard variables and entry points. Evaluation of models participating in the CORDEX project is provided by the RCMES.

[4]Note that the boundary of the domain of RCMs may be source of spurious instabilities. Modelers attempt to design RCMs so as to dump these instabilities, but some may persist and lead to spurious results in impact studies.

For more information on evaluation and decision support, see Whitehall et al. 2012.

CORDEX data may be found on the ESGF data portal.

4.4 Indices

Some hazards may not be directly characterized by a particular climate variable. For instance the severity of a drought depends, in particular, on precipitation, temperature and on soil parameters and variables. One option is to define an index translating and/or aggregating information into a hazard intensity.

In the case of droughts, the *Standardized Precipitation Index* (SPI) quantifies the number of standard deviations by which the observed anomaly in precipitation deviates from the long-term mean. The *Standardized Precipitation Evapotranspiration Index* (SPEI) extends the SPI to take into account potential evapotranspiration in order to capture the main impact of increased temperatures on water demand[5]. Yet another index, the *self-calibrating Palmer Drought Severity Index* (sc-PDSI) is also sometimes used. The relevance of a particular index depends on the hazard, geographical location, local environmental conditions and index design.

Last, climdex.org collects a number of indices of extreme events computed from different global datasets. While such tools can facilitate physical climate risk assessment, care should however be taken to evaluate the local representativity of global indices.

5 Dynamical and statistical downscaling

The objects of risk assessment, and of impact studies in general, tend to depend on climatic conditions at relatively small scale, that is, the scale of a field, a city, a plant, a building, a river, and so on. On the other hand, global projections and reanalyses, with a resolution of about a degree or half a degree, tend to properly resolve climatic variations around a few degrees (i.e. about 110 km at the equator, 60 km at a latitude of 45°). To pass from

[5] Vicente-Serrano, Sergio M. & National Center for Atmospheric Research Staff (Eds). Last modified 18 Jul 2015. "The Climate Data Guide: Standardized Precipitation Evapotranspiration Index (SPEI)." Retrieved from `https://climatedataguide.ucar.edu/climate-data/standardized-precipitation-evapotranspiration-index-spei`.

the scale resolved by such global models to the scale of interest for impact studies several methodologies, known as *downscaling* are often used. These methodologies all rely on the addition of small-scale information in some way or another. Moreover, some approaches attempt to downscale climatic information into a small scale climate variable or index, while other directly downscale climatic information into some measure of impact or damage.

Downscaling may be divided in two broad categories: *dynamical* and *statistical* downscaling (for an introduction to the topic, see Benestad 2016). Dynamical downscaling relies on the resolution of earth science equations at small scales, via the integration of a regional, or limited-area, climate model. This the case of the COSMO reanalysis and of the CORDEX simulations discussed in Section 4.2.2 and 4.3.3, respectively, whereby scales of a few kilometers or tens of kilometers are reached. Even smaller scales, about a 100 m, can be reached with the use of computational fluid dynamics models. At such scales, dynamics are turbulent, so that only statistical information is relevant.

Dynamical downscaling presents the advantage of relying on fundamental equations and on integrating information at small scales about topography, land use, etc. Because of this, one expects dynamical downscaling approaches to be relatively robust to parameter or forcing changes, although some parameters of the model are still calibrated on the historical period. It is known to add value to global climate data (Hastie, Tibshirani, and Friedman 2009). However, dynamical downscaling is also a computationally intensive exercise. In addition, the information downscaled by RCMs is still climatic, so that integrated approaches are needed to translate this small-scale information into an impact.

Statistical downscaling provides an alternative that does not rely on earth science laws but rather on local information and statistical learning (Hastie, Tibshirani, and Friedman 2009). Take as an example the assessment of heat wave occurrence (a statistic) in the 21st century in the city center of Grenoble, France. We know that historical temperatures in the city center of Grenoble are not well reproduced by global reanalyses. Regional reanalyses tend to improve these results by better resolving the impact of the Alps surrounding Grenoble, but are still not satisfactory: the urban heat island is not well resolved. Thus, not only reanalyses cannot be used here to project the increase in temperatures in Grenoble associated with global warming, but global climate projections may not be expected to resolve well local temperature distributions. Assume, however, that a number of temperature sensors are scattered in Grenoble so that time series of temperatures there for the last few years are available. This information on the local climatic conditions could be combined with information about global warming provided by projections to estimate future changes in the temperature distribution in Grenoble, and in particular the occurrence of heat waves. For instance, perhaps the most basic approach would be to estimate a historical temperature distribution from the local measurements, compute the difference between the mean of this distribution and the temperature predicted by the global model during the same period, and predict future temperature distributions by shifting the local temperature distribution to follow the changes in the mean temperature projected by the global model. This particular approach is known as bias correction.

The design and choice of a particular downscaling method is an art in itself. The particular impact considered, its location, the data and resources available, need to be considered. Thus, Estrada et al. 2013 advise against applying downscaling automatically. The VALUE project (Maraun et al. 2015) provides a framework to assess the skills of different downscaling approaches which has been applied other Europe by Gutiérrez et al. 2018. In particular, downscaling studies often rely on reanalyses, the skills of which have a strong impact on the result of downscaling (Brands et al. 2011).

6 Climate data bias and uncertainty

The goal of this section is not to give an overview of all known biases in climate data, as their relevance depends on the object of the risk assessment study. Rather, we would like to warn that the availability of a large amount of climate data should not lead to over-confidence in their use for risk assessment. While global mean changes in temperature are relatively well resolved by reanalyses and projections, this is not the case of all variables and statistics:

- Some variables are more sensitive to measurement errors and modeling assumptions. For instance, precipitation directly depends on cloud microphysics which still constitutes one of the main weaknesses of earth system models (Stocker et al. 2013).
- The finer the spatial scale, the more biased the models. This is true for both reanalyses and projections.

- Statistics giving more weight to tails of distributions, such as extreme values, tend to be more biased than, for instance, averages.

It is thus essential, in risk assessments relying on climate data, to estimate errors originating from the climate data itself. Indeed, it is often found in impact studies that a dominant source of errors comes from the climate data. To compute a first, yet incomplete, estimate of errors coming from the climate data, multiple data sources may be used. In the case of projections, multiple CMIP and CORDEX models may be used. The number of reanalyses is not as large, but several of them are available, and errors are reduced thanks to observational constraints.

When dealing with 21st century projections, the main source of uncertainty stems from the socio-economic scenario that the economic actors with the largest footprint will follow. For that reason, several typical scenarios need to be considered. As discussed in Section 4.3.2, the CMIP and CORDEX databases collect simulations from different models that are forced by different RCP scenarios. The latter may be used to prepare assessments or indicators corresponding to different socio-economic scenarios. The problem of choosing under uncertainty is still, however, pervasive.

7 Example: heatwaves in the Paris area

To illustrate the key messages of this report, we give the following example. Our objective is to compute the evolution of the number of days of heatwave over the end of the 20th century and over the 21st century at a given location and from various climate data-sources. This example is described in the next Section 7.1, while the methodology is further detailed in Section 7.2.

7.1 Description and interpretation

A heatwave is a type of extreme event increasingly impacting societies. The risk associated with a heatwave depends on the vulnerability of the population, such as its age, its access to water, etc. Here, we instead focus on the intensity of a heatwave as a climate event and follow the definition given in Ouzeau et al. 2015, Section 5.3.1:

Definition. *A heatwave is defined as an abnormally hot period lasting more than five consecutive days. Hot days are defined as days for which the maximum daily temperature is* 5 °C *larger than a climatological reference value,* *counting only days belonging to a series of more than five consecutive hot days. To obtain this reference value for each day of the year, one computes the mean annual-cycle of the daily-maximum temperature simulated for the reference period (1976–2005) [1980–2010, here], applying a running average over five days to this annual cycle.*

The SIRTA observational platform provides hourly temperature observations from 2003 to 2019 (Haeffelin et al. 2005). We choose to study heatwaves at the location of the platform, 48.718°N, 2.202°E (20 km south of Paris). This data is compared to a global reanalysis and to Regional Climate Model (RCM) runs. The MERRA-2 reanalysis (Gelaro et al. 2017) covers the 1980–2019 period at a resolution of about half a degree. The RCM runs from the Med-CORDEX exercise (Ruti et al. 2016) are concatenated to cover both the historical period (1950–2006) and the 21st century (2006–2100), using projections for the RCP 8.5 scenario (the warmest, business-as-usual) at a little less than half a degree of resolution. Two runs from the GUF and the CMCC institutes are selected. Records from the nearest grid-point of these gridded datasets to the SIRTA platform are analysed. MERRA-2's nearest grid point is at 48.500°N, 2.500°E (about 33 km from SIRTA) , while GUF's and CMCC's is at 48.524°N, 2.094°E (about 23 km from SIRTA).

Four records of daily-maximum temperature are obtained from these datasets. In addition, as a simple example of statistical downscalling, we remove from the reanalysis and the two RCM runs the bias with respect to the SIRTA data. These two sets of time series are represented in Fig. 1a and 1b. We then compute heatwave indices for each run and represent them in Fig. 1d and 1c. Thick lines correspond to smoothed versions of the yearly time-series (thin lines) using a rolling average over 20 years. Note that the rolling average uses fewer points when less than 20 years are available. For instance, at the middle of the 12-years-long SIRTA time-series, only 12 points are available and used.

We first observe that a strong bias (half a degree to several degrees) exists between the SIRTA observations (black) and both the reanalysis (gray) and the RCM runs (blue and orange). This bias may be explained by modeling errors, a poor representativity of the selected grid-box of the gridded datasets with respect to the SIRTA station, or by (less likely) observational errors in the reanalysis and the in situ observations. The RCM runs are also significantly different from each other. This shows the importance of comparing multiple data sources and validating estimations against observation over some available period.

We also note from the erratic behaviour of the thin curves compared to the thick, low-pass filtered, curves that there is a strong interannual variability in the maximum temperature. Zooming on the period covered by the SIRTA record, Fig. 2, we can see that the correlation between MERRA-2 temperatures and the SIRTA temperatures is stronger than that between the SIRTA temperatures and the RCM runs. Considering the chaotic nature of the climate system, this may be understood from the fact that, contrary to the RCM runs, the reanalysis is constrained by observations, and thus more likely to follow the observed interannual variability. On the other hand, by design, the reanalysis does not provide projections for the 21st century.

The bias correction, Fig. 1b, allows us to reduce the gap between the observations and the other records. However, we are not guaranteed that this bias is stationary over time. For instance, during the 21st century, a warmer climate may induce changes in physical processes which could result in weaker or stronger biases.

Let us now focus on the heatwave index, Fig. 1d and 1c. We can first note that the observational record (black) appears to be too short (2002–2019) to observe any trend in the number of heatwave days with this methodology. From 1980 to 2019, a small trend is visible in the reanalysis (gray) and the RCM runs (blue and orange), although its significance should be properly tested. A very clear trend is, however, visible over the 21st century for both RCM runs, although the magnitude of this trend differs significantly (by about 70 days).

The heatwave index is less sensitive to biases, as it is defined as a difference with respect to a reference annual cycle computed from the same dataset (except for the SIRTA heatwave index, which is computed using the bias-corrected MERRA-2 annual cycle).

Finally, note that we have focused here on the estimation of the number of heatwave days from different climate data to illustrate basic methodologies and their pitfalls. It remains then to estimate the actual damages associated with these climate events.

7.2 Methodology

We now explain in more details the methodology used to obtain the results of Section 7.1.

7.2.1 Downloading the data

The in-situ observations, the reanalysis and the projections are obtained from the SIRTA observatory, NASA's MERRA-2 and the Med-CORDEX initiative.

SIRTA The SIRTA database is accessible free of cost for public research and teaching applications[6]. We download the SIRTA (Haeffelin et al. 2005) temperature record via File Transfer Protocol (FTP) from SIRTA's data-download page[7], selecting *atmospheric state/surface meteorology/Air temperature/Meteorological variables with LMD station in zone 2* for the 2003/01/03–2019/01/01 period. The data is sampled every minute.

MERRA-2 NASA promotes the full and open sharing of all data with the research and applications communities, private industry, academia, and the general public. The MERRA-2 data is accessible from NASA's GES DISC[8], after registration. To get the daily maximum temperature, the single-level daily-statistics dataset[9] should be selected.

Med-CORDEX Data from the Med-CORDEX project available on this server is provided without charge and may be used for research and education only. Commercial use of the data is not permitted. The Med-CORDEX data is accessible through the Med-CORDEX FTP[10]. For instance, the daily-maximum temperatures from the GUF historical run and RCP8.5 scenario were respectively downloaded from `ftp://www.medcordex.eu/MED-44/GUF/MPI-ESM-LR/historical/r1i1p1/GUF-CCLM4-8-18/v1/day/tasmax/` and `ftp://www.medcordex.eu/MED-44/GUF/MPI-ESM-LR/rcp85/r1i1p1/GUF-CCLM4-8-18/v1/day/tasmax/`.

7.2.2 Bias correction

The bias correction adjusts for differences in the mean between the SIRTA maximum-temperature record and that of the reanalysis and RCM records.

[6] See SIRTA's Data Policy.

[7] `https://sirta.ipsl.fr/data_download.html`

[8] `https://disc.gsfc.nasa.gov/`

[9] `https://disc.gsfc.nasa.gov/datasets/M2SDNXSLV_5.12.4/summary?keywords=MERRA-2`

[10] `ftp://www.medcordex.eu/`

To do so, we first select the period over which both records overlap. The mean over this period of the record to be corrected is removed from it and we add instead the mean of the SIRTA record over the same period. As a result, the mean of the corrected record is the same as the mean of the SIRTA record over the overlapping period.

To go further

- The Climate Data Guide describes different sources of climate data.
- reanalyses.org focuses on reanalyses.
- https://s-rip.ees.hokudai.ac.jp/pubs/index.html provides data from different reanalyses on common grids.
- Data used for downscaling from the VALUE project `http://www.value-cost.eu/data`.
- The S-RIP report (Fujiwara et al. 2017) gives an overview of the different reanalysis systems.
- Chapter 12 of the Working Group I on *The Physical Science Basis*[11] contribution to the 6th Assessment Report of the IPCC and Chapter 16 of the Working Group II on *Impacts, Adaptation and Vulnerability*[12] will cover climate risk assessment.

Acknowledgement

The authors would like to acknowledge SIRTA for providing the temperature data used in this study.

[11] `https://www.ipcc.ch/report/sixth-assessment-report-working-group-i/`

[12] `https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/`

A Glossary

Climate model	A physical model for the evolution of the Earth climate in the long term (100+ years). Climate models are used to make climate projections.
Climate variability	Natural variability of the Earth climate at the scales longer than one year.
CMIP	Coupled Model Intercomparison Project, an exercise led by IPCC wherein different climate models are run under the same conditions to produce multiple climate scenarios which may be compared to evaluate the uncertainty of future climate.
Downscaling	A procedure to refine the scale of climate models down to that of local phenomena of interest.
ECMWF	European Centre for Medium-Range Weather Forecasts
Integrated assessment model	A model which aims to describe the joint evolution of the economy and the climate system.
Physical climate risk	Risk of destruction of assets / disruption of systems by weather events related to climate change and climate variability.
Reanalysis	In a reanalysis exercise, a climatological model is run over a long period of time in the past, assimilating all available observations as model constraints. This allows to interpolate available observations over regular time and space domain. The standard practice in climatology, which we folliw in this paper, is to replace historical data with reanalysis data.
RCP	Representative Concentration Pathway, a scenario of greenhouse gas emissions used to parameterize climate models

References

[Ben16] Rasmus Benestad. “Downscaling Climate Information”. en. In: *Oxford Research Encyclopedia of Climate Science* (July 2016). DOI: `10/gfvn8q`.

[Bol+15] C. Bollmeyer et al. “Towards a High-Resolution Regional Reanalysis for the European CORDEX Domain”. In: *Quarterly Journal of the Royal Meteorological Society* 141.686 (Jan. 2015), pp. 1–15. ISSN: 0035-9009. DOI: `10/f6z65b`.

[Bra+11] S. Brands et al. “On the Use of Reanalysis Data for Downscaling”. In: *Journal of Climate* 25.7 (Dec. 2011), pp. 2517–2526. ISSN: 0894-8755. DOI: `10/fxp825`.

[Bui+05] Roberto Buizza et al. “A Comparison of the ECMWF, MSC, and NCEP Global Ensemble Prediction Systems”. In: *Mon. Weather Rev.* 133.5 (2005), pp. 1076–1097. ISSN: 0027-0644. DOI: `10/bcjtng`.

[Cop17] Copernicus Climate Change Service (C3S). “ERA5: Fifth Generation of ECMWF Atmospheric Reanalyses of the Global Climate”. In: *Copernicus Climate Change Service Climate Data Store (CDS)* (2017).

[Est+13] Francisco Estrada et al. “A Cautionary Note on Automated Statistical Downscaling Methods for Climate Change”. en. In: *Climatic Change* 120.1 (Sept. 2013), pp. 263–276. ISSN: 1573-1480. DOI: `10/f468pg`.

[Eyr+16] Veronika Eyring et al. “Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) Experimental Design and Organization”. In: *Geosci. Model Dev.* 9.5 (2016), pp. 1937–1958. ISSN: 19919603. DOI: `10/gbpwpt`.

[Fuj+17] Masatomo Fujiwara et al. “Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and Overview of the Reanalysis Systems”. English. In: *Atmospheric Chemistry and Physics* 17.2 (Jan. 2017), pp. 1417–1452. ISSN: 1680-7316. DOI: `10/f9t7r4`.

[Gel+17] Ronald Gelaro et al. "The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)". In: *Journal of Climate* 30.14 (May 2017), pp. 5419–5454. ISSN: 0894-8755. DOI: 10/gbk3sp.

[Gut+18] J. M. Gutiérrez et al. "An Intercomparison of a Large Ensemble of Statistical Downscaling Methods over Europe: Results from the VALUE Perfect Predictor Cross-Validation Experiment". en. In: *International Journal of Climatology* (2018), pp. 1–36. ISSN: 1097-0088. DOI: 10/gf3t6n.

[Hae+05] M. Haeffelin et al. "SIRTA, a Ground-Based Atmospheric Observatory for Cloud and Aerosol Research". English. In: *Annales Geophysicae* 23.2 (Feb. 2005), pp. 253–275. ISSN: 0992-7689. DOI: 10/fd8mbb.

[HEN18] Romain Hubert, Julie Evain, and Morgane Nicol. *Getting Started on Physical Climate Risk Analysis in Finance - Available Approaches and the Way Forward*. Tech. rep. I4CE, 2018.

[HTF09] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. 2nd ed. New York: Springer, 2009. ISBN: 978-0-387-84858-7. DOI: 10.1007/978-0-387-84858-7.

[Kob+15] Shinya Kobayashi et al. "The JRA-55 Reanalysis: General Specifications and Basic Characteristics". en. In: *Journal of the Meteorological Society of Japan. Ser. II* 93.1 (2015), pp. 5–48. ISSN: 0026-1165, 2186-9057. DOI: 10/f65v7p.

[Lon+17] Craig S. Long et al. "Climatology and Interannual Variability of Dynamic Variables in Multiple Reanalyses Evaluated by the SPARC Reanalysis Intercomparison Project (S-RIP)". English. In: *Atmospheric Chemistry and Physics* 17.23 (Dec. 2017), pp. 14593–14629. ISSN: 1680-7316. DOI: 10/gcqnrs.

[Mar+15] Douglas Maraun et al. "VALUE: A Framework to Validate Downscaling Approaches for Climate Change Studies". en. In: *Earth's Future* 3.1 (Jan. 2015), pp. 1–14. ISSN: 2328-4277. DOI: 10/gfkbvk.

[Ouz+15] G. Ouzeau et al. *Le Climat de La France Au XXIe Siècle: Scénarios Régionalisés : Édition 2014 Pour La Métropole et Les Régions d'outre-Mer*. Vol. 4. Ministère de l'Écologie, du Développement Durable et de l'Énergie, 2015.

[Pol+13] P Poli et al. *The Data Assimilation System and Initial Performance Evaluation of the ECMWF Pilot Reanalysis of the 20th-Century Assimilating Surface Observations Only (ERA-20C)*. Tech. rep. Shinfield Park, Reading: ECMWF, Sept. 2013, p. 59.

[Rut+16] Paolo M. Ruti et al. "Med-CORDEX Initiative for Mediterranean Climate Studies". In: *Bull. Am. Meteorol. Soc.* 97.7 (2016), pp. 1187–1208. ISSN: 00030007. DOI: 10.1175/BAMS-D-14-00176.1.

[Sah+10] Suranjana Saha et al. "The NCEP Climate Forecast System Reanalysis". In: *Bulletin of the American Meteorological Society* 91.8 (Apr. 2010), pp. 1015–1058. ISSN: 0003-0007. DOI: 10/cmpc8x.

[Sto+13] T.F. Stocker et al., eds. *IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 2013. DOI: 10.1017/CBO9781107415324.

[TSM12] Karl E. KE Taylor, RJ Ronald J. Stouffer, and GA Gerald A. Meehl. "An Overview of CMIP5 and the Experiment Design". In: *Bull. Am. Meteorol. Soc.* 93.4 (Apr. 2012), pp. 485–498. ISSN: 00030007. DOI: 10/b7xgw2.

[Wah+17] Sabrina Wahl et al. "A Novel Convective-Scale Regional Reanalysis COSMO-REA2: Improving the Representation of Precipitation". en. In: *Meteorologische Zeitschrift* (Oct. 2017), pp. 345–361. ISSN: , DOI: 10/gcnr2m.

[Wan+11] Wanqiu Wang et al. "An Assessment of the Surface Climate in the NCEP Climate Forecast System Reanalysis". en. In: *Climate Dynamics* 37.7 (Oct. 2011), pp. 1601–1620. ISSN: 1432-0894. DOI: 10/bgrm4z.

[Whi+12] Kim Whitehall et al. “Building Model Evaluation and Decision Support Capacity for CORDEX”. In: *WMO* 61 (2012).

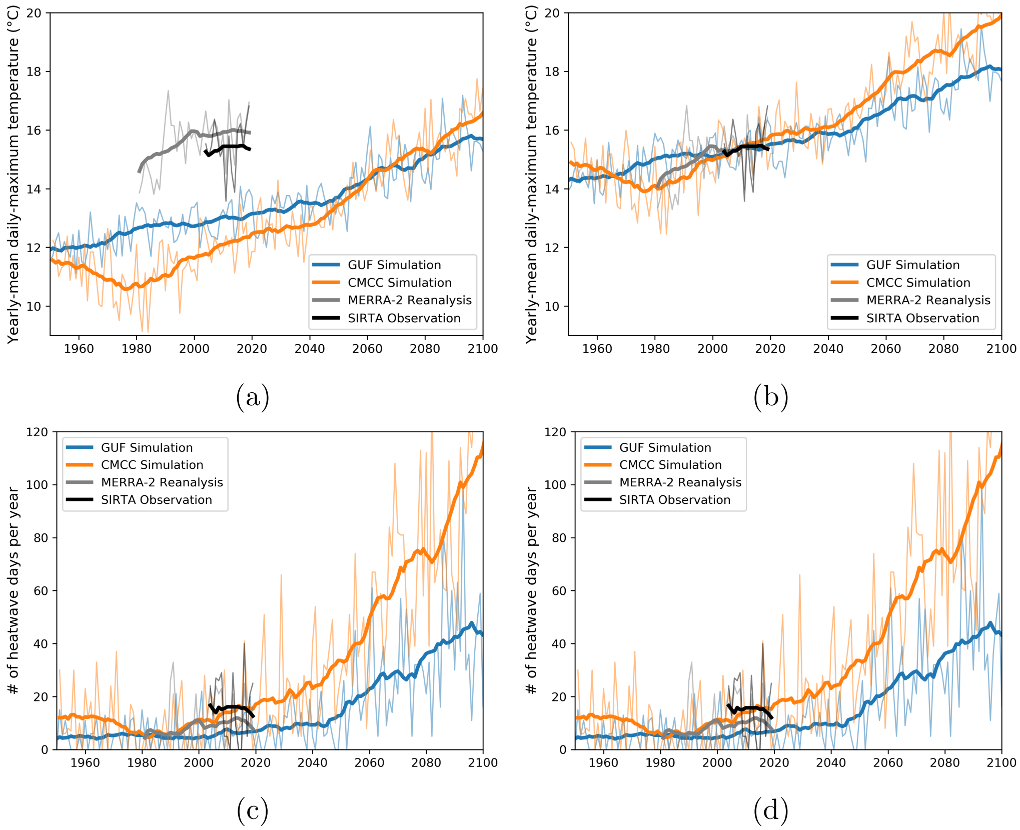

Figure 1: Yearly-mean daily-maximum temperature (top) and number of heatwave days (bottom) without (left) and with (right) bias-correction. Thick lines correspond to smoothed versions of the thin lines using a rolling average over 20 years. Black lines are obtained from the SIRTA observational data from 2003 to 2019. Gray lines are obtained from the MERRA-2 reanalysis data from 1980 to 2019. Blue lines are obtained from the concatenation of the historical and RCP 8.5 CORDEX runs by the GUF institute. Orange lines are obtained from the concatenation of the historical and RCP 8.5 CORDEX runs by the CMCC institute. The number of heatwave days are computed following the methodology described in Ouzeau et al. 2015, Section 5.3.1, with 1980–2010 as reference period. Since the SIRTA record is too short, we use the bias-corrected MERRA-2 data as reference to compute the number of heatwave days in the SIRTA record. The bias correction is simply performed by removing from all non-SIRTA records their mean over the SIRTA period and adding instead the mean of the SIRTA record.

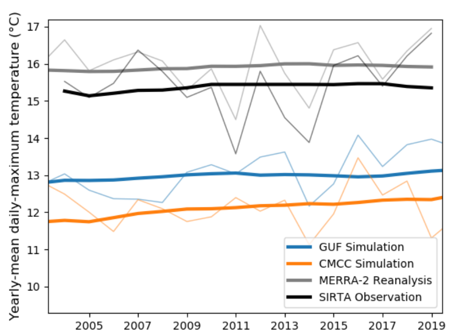

Figure 2: Zoom of Fig. 1a.