THEME SPA.2013.1.1-05

EUCLEIA

(Grant Agreement 607085)

EUropean CLimate and weather Events: Interpretation and Attribution

Deliverable D7.1

Attribution of test heatwaves

Deliverable Title	*Attribution of test heatwaves*
Brief Description	This report investigates the 2015 heatwave as a case study for the assessment of attribution framing and methods. The event was the subject of a near-real-time attribution analysis and was subsequently reassessed in two peer-reviewed publications taking a range of different and complementary approaches to the attribution question. All above studies were supported by EUCLEIA. We compare results and conclusions of these studies, and show that for a relatively straightforward attribution question such as this one, qualitative conclusions are robust to the event definition, the method used and the time-scale over which the analysis was undertaken.
WP number	WP5
Lead Beneficiary	*Friederike Otto, University of Oxford*
Contributors	*Geert Jan van Oldenborgh, KNMI* *Peter Uhe, University of Oxford* *Myles Allen*
Creation Date Version Number Version Date	
Deliverable Due Date Actual Delivery Date	

Nature of the Deliverable	X	*R - Report*
		P - Prototype
		D - Demonstrator
		O - Other
Dissemination Level/ Audience	X	*PU - Public*
		PP - Restricted to other programme participants, including the Commission services
		RE - Restricted to a group specified by the consortium, including the Commission services
		CO - Confidential, only for members of the consortium, including the Commission services

Version	Date	Modified by	Comments
0.1	27.1.17	Fredi Otto	Initial compilation of report

Table of Contents

1. Executive Summary

The summer of 2015 in Europe was the driest on record and the third warmest for the whole of Europe and was characterised by short heat waves lasting days to weeks. The magnitude of the warming over the whole summer is comparable with previous European summer heatwaves in 2003 and 2010. Here we investigate the 2015 heatwave as a case study for the assessment of attribution framing and methods. The event was the subject of a near-real-time attribution analysis and was subsequently reassessed in two peer-reviewed publications taking a range of different and complementary approaches to the attribution question. All above studies were supported by EUCLEIA. We compare results and conclusions of these studies, add assessments of available methodologies not previously applied and show that for a relatively straightforward attribution question such as this one, qualitative conclusions are robust to the event definition, the method used and the time-scale over which the analysis was undertaken.

The increase in the probability of the event, or the probability ratio, is highly sensitive to its definition and in particular the spatial and temporal averaging interval considered, due to a change in the shape of the return-time. Focussing on the correct diagnostic is also important. For example, defining the intensity of heatwaves in terms of the wet-bulb-temperature or temperature alone has a large effect on the probability ratio. Causal analyses identifying mechanisms whereby human influence on climate or other external factors may be impacting the risk of a specific extreme weather event, while not contributing directly to the assessment of relative risk with and without human influence, can either considerably strengthen conclusions or provide a valuable note of caution.

2. Project Objectives

With this deliverable, the project has contributed to the achievement of the following objectives (DOW, Section B1.1):

No.	Objective	Yes	No
1	Derive the requirements that targeted user groups (including regional stakeholders, re-insurance Companies, general public/media) have from attribution products and demonstrate the value to these users of the attribution products developed under EUCLEIA.		✓
2	Develop experimental designs and clear ways of framing attribution studies in such a way that attribution products provide a fair reflection of current evidence on attributable risk.	✓	
3	Develop the methodology for representing the level of confidence in attribution results so that attribution products can be trusted to inform decision making.	✓	

4	Demonstrate the utility of the attribution system on a set of test cases of European weather extremes.	✓	
5	Produce traceable and consistent attribution assessments on European climate and weather extremes on a range of timescales; on a fast-track basis in the immediate aftermath of extreme events, on a seasonal basis to our stakeholder groups, and annually to the BAMS attribution supplement.	✓	

3. Detailed Report

Introduction

The summer 2015 in Europe was the driest on record and the third warmest (Dong et al., 2016; Sippel et al. 2016) for the whole of Europe, with many countries and cities setting all-time temperature records, including Germany with temperatures of 40.3^{o}C measured in Kitzingen in early July, and various records in extreme indices including the hottest day of the year, TXx (Dong et al. 2016). In contrast to the Russian heat wave of 2010, which was characterised by a persistent blocking high that lasted for all of July into mid-August (Barriopedro), the summer of 2015 saw a number of shorter heat waves lasting days to weeks, with record temperatures in the Western part of the continent early in the summer (Sippel et al. 2016) and extreme heat in Central and Eastern Europe in July and August. Thus while the magnitude of the warming over the whole summer is comparable with previous European summer heatwaves in 2003 and 2010 (Dong et al. 2016), the circulation patterns were different, with a characteristic Omega blocking event in the early summer and a stable high pressure system as a result of a northward displaced Jetstream later in the season (Sippel et al. 2016). Seasonal mean anomalies of various key temperature indicators are shown in figure 1, reproduced from Dong et al. (2016).

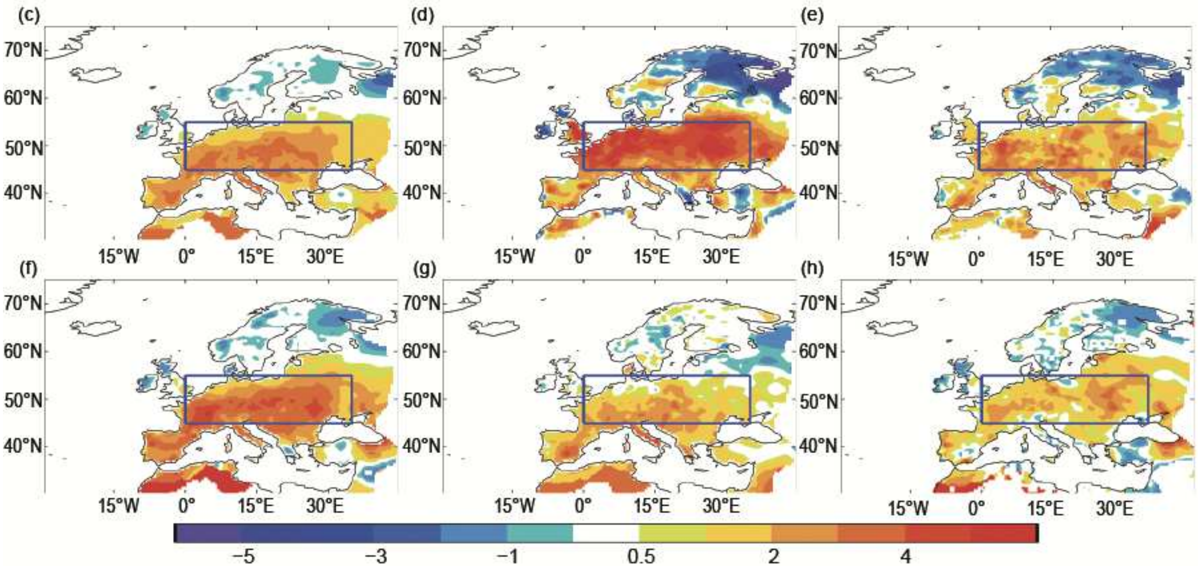


Figure 1: *The summer of 2015 in various temperature diagnostics. Anomalies relative to 1964-93 of seasonal mean (June-August, JJA) surface air temperature (SAT, top left), hottest day of the year (TXx, top middle), hottest night of the year (TNx, top right), JJA mean daily maximum temperature (Tmax, bottom left), nightly minimum temperature (Tmin, bottom middle), and diurnal temperature range (DTR, bottom right) from the E-OBS dataset*

(Haylock et al., 2008), reproduced from figure 12.1(c-h) of Dong et al. (2016). The box shows the area 45-55N, 0-35E used for scalar diagnostics in this report.

The daylong heat and in particular high nighttime temperatures are what primarily causes impacts and affects human health. However, the actual heat related mortality was much lower in 2015 than for example in the heatwave of 2003 (Mitchell et al. 2016) despite comparable temperature anomalies. This heatwave is thus a very good example of how changes in the vulnerability lead to great differences in the impact of the events even though they are comparable considering impact related meteorological variables.

As in all event attribution studies there are different ways to characterise the event that may lead to different quantitative assessments of the role of human influence on climate on the likelihood and magnitude of the event occurring (Sippel and Otto 2014). Event definition depends on whether the effect on human health is the central outcome of the assessment, whether the focus is on explaining the records being set in routinely measured meteorological variables or on explaining the specific atmospheric circulation state associated with this event. A record in TXx may not correspond to a record in heat stress, which combines temperature and relative humidity, in particular not in a dry summer. At the same time while the temperatures might have set all time records, the circulation state might be very typical for the time of year but local processes such as land surface feedbacks have amplified the effect on maximum temperatures (e.g. Seneviratne et al., 2013, Miralles et al., 2014.). Event definition not only influences the return time of the analysed variable but also which methodologies of extreme event attribution are available and which climate models can reliably simulate the event.

The 2015 heatwave provides an excellent case study for the assessment of attribution framing and methods, having been the subject of the first ever near-real-time attribution analysis (http://www.climatecentral.org/europe-2015-heatwave-climate-change), subsequently reassessed in two peer-reviewed publications taking a range of different and complementary approaches to the attribution question, all supported by the EUCLEIA project. In this report, we compare results and conclusions of these studies, and show that for a relatively straightforward attribution question such as this one, qualitative conclusions are robust to the event definition, the method used and the time-scale over which the analysis was undertaken.

Attribution approaches

The near-real-time analysis used two methods to address the attribution question. The first was a statistical analysis of observations of maximum temperatures in station data to quantify the impact of large-scale temperature trends on the distribution of weather variables. In this method, causal impact is assessed by modelling changes in statistical moments of the distribution as a linear combination of components proportional to large-scale drivers (including in this case global temperature changes since 1900), and a stationary residual (van Oldenborgh et al, 2007). It has the advantages that only observational data and a statistical model are used, such that results are not sensitive to dynamical model biases, and very low computational cost, such that results can be readily computed in near-real-time. The caveats are that it only demonstrates an association with a large-scale trend, not a specific external driver, and the limited length of the observational record means that uncertainty estimates can be relatively large.

The second method in the near-real-time analysis focussed on comparing maximum temperatures in station data with a very large ensemble of weather@home regional climate simulations using the UK Met Office's HadAM3P with nested HadRM3P model in a distributed computing framework (Massey et al. 2015). This method examines the frequency of occurrence of an event, normally defined by a scalar variable exceeding a specific threshold, in an ensemble of atmospheric model simulations driven with "actual" observed atmospheric composition and sea surface temperatures from a seasonal forecast initialised about two months prior to the event (Haustein et al, 2016). This frequency is compared with a similarly-constructed “natural” ensemble from which a range of estimates of anthropogenic influence on climate change, including both atmospheric composition and various estimates of sea surface temperature and sea ice extent (SST and SIE) change. Causal impact is unambiguous in this design, at the expense of reliance on an atmospheric model that may be subject to biases in both mean state and response. The use of forecast (as opposed to observed) SSTs for the actual ensemble means that the impact of local, short-duration SST anomalies is not captured, but allows a near-real-time analysis.

A third, closely related, analysis (Sippel et al, 2016) revisited the ensemble used in the near-real-time analysis, using the same analysis methods to evaluate whether these findings would hold using quality-controlled station data instead of that which was instantaneously available. This study also made a thorough evaluation and bias correction of the large ensemble simulations, and extended the analysis to include an estimate of heat stress on the human body, the 3-day daily maximum wet bulb temperature. This allows an assessment of the impact of the framing of the attribution question on results (Otto et al., 2016).

A fourth analysis, also the subject of a peer-reviewed publication (Dong et al. 2016), compared the contributions to the 2015 heatwave of anomalous sea surface temperatures in 2015 (relative to the 1964-1993 climatology) with the direct impact of atmospheric composition change (due to GHG and sulphur dioxide emissions) relative to the same period. Since some contribution of anthropogenic influence is present in both the 1964-1993 climatology and 2015 SSTs, this does not provide an explicit attribution to human influence on global climate, but it does provide insight into the relative importance of atmospheric composition versus contemporaneous SST forcing for this particular temperature event.

A fifth approach, that was not used in this particular near-real-time analysis but has been used extensively in related near-real-time attribution exercises, is the evaluation of relative risk in a pre-computed coupled Atmosphere-Ocean General Circulation Model (A-OGCM) ensemble (King et al, 2014). The CMIP5 ensemble of coupled models is often used for this purpose. The advantages of this approach are that risks can be pre-computed, so it is consistent with a near-real-time analysis, and fully takes into account coupled ocean-atmosphere variability. The disadvantages are that coupled A-OGCMs are typically coarser resolution than the global atmospheres (and certainly the nested RCMs) used in attribution studies with prescribed SSTs: hence results from this approach were not used in the near-real-time analysis in 2015 which focussed on individual station data. Climatological biases in A-OGCMs can also be more substantial than in SST-driven Atmosphere GCMs. Although this approach was neither used in the near-real-time study nor in any subsequent peer-reviewed paper, results are included here for comparison with other approaches.

Finally, a sixth approach considered here is the method of analogues of Yiou et al. (2007). Like that of Dong et al (2016), this does not explicitly address the question of attribution to anthropogenic climate change in the heatwave event itself, but allows us to assess both the role of human influence on atmospheric circulation and the role of circulation in the event that occurred. Some preliminary results from the application of this approach to the 2015 heatwave are presented here, and will be further investigated in subsequent papers.

Event definition

In every attribution study, formal event definition is a crucial step that may have a substantial impact on the outcome of the study. In order to compare occurrence-probabilities straightforwardly, it is necessary to reduce the complex, high-dimensional datasets that represent the event in question into one or more critical scalar indices for which thresholds can be determined and exceedance probabilities computed. A choice must always be made regarding the diagnostic used and spatial and time averaging period considered. In general, smaller spatial scales are more relevant to most impacts, while the appropriate temporal averaging period depends on the impact considered: analysis of mortality in the 2003 heatwave (Mitchell et al. 2016) indicates that sustained high nighttime temperatures over a period of 1-2 weeks cause most harm, more so than a single very hot day. Impacts on agriculture may depend more on daytime temperatures, and depending on the crop, may be sensitive to shorter timescales.

While the ideal approach to event definition is purely in terms of impact, three further practical considerations are relevant. Both observational and model-based approaches to attribution may be made easier by focussing on a large-area, multi-day average. Unless long observational data records happen to exist for precisely the location and diagnostic of interest, it may be necessary to interpolate. Sampling noise is always a significant issue for extreme weather indices, and can be reduced by averaging over observational records. Atmospheric models also typically reproduce the statistics of large-area, long-period averages better than they reproduce observed variables on smaller and shorter scales.

A second issue is multiple testing: if an event is defined on a very small scale, then depending on the degree of coherence of the overall meteorological phenomenon, a potentially very large number of individual "local" events could be derived from a single large-scale event. Methods exist for assessing field significance in the meteorological literature (see Wilks, 2016, for a survey), but have not thus far been widely applied to the analysis of probability ratios widely used in attribution. "Numbers of stations at which weather records are broken" is sometimes used informally to indicate the magnitude of an event, but is clearly a problematic statistic unless supported by a very careful treatment of spatial covariance. If only a small number of results are presented for high-impact locations, this is less of a problem, but the issue of to what extent multiple positive attribution results for different locations and diagnostics actually support each other remains unresolved. For a spatially coherent event such as a heatwave, spatially incoherent attribution conclusions would indicate a reason for caution, but the converse is not true: spatially coherent conclusions are to be expected, and do not suggest a higher level of confidence than is indicated by any individual test.

Third, there is the choice of observational dataset to use: a compromise may have to be made between precision and coherent long-term records. For example, in the E-OBS observational dataset, the most precise available, which extends over 80 years, the return time of 2015 seasonal mean area averaged temperatures over the main region of the event, depicted in figure 1, is estimated to be 83 years, but much higher local return-times emerge in Austria and the Czech Republic (Figure 2, left panel). In the somewhat longer Berkeley Earth data set (figure 2, right panel), which goes back to 1900, the seasonal area averaged temperature has an estimated return time of 196 years. The map of local return-times in the Berkeley Earth data set is necessarily smoother because of the higher level of spatial smoothing that is used in the construction of this dataset, so this smoothness may not be representative of the real world.

Finally, a decision needs to be made whether to define an event in terms of an absolute threshold or an estimated return time. Given substantial biases in many of the models used for attribution, defining an event in terms of exceeding the relevant 100-year return-time is generally much easier than focusing on an absolute threshold, at the cost of a slightly higher level of abstraction. This is the approach used in most of the examples in this report.

In summary, event definition matters, and matters in a predictable way: Uhe et al (2016) demonstrate that the larger the spatial scale considered, the higher estimated probability ratios tend to be, because of the lower level of noise in the diagnostic. A similar effect is noted here.

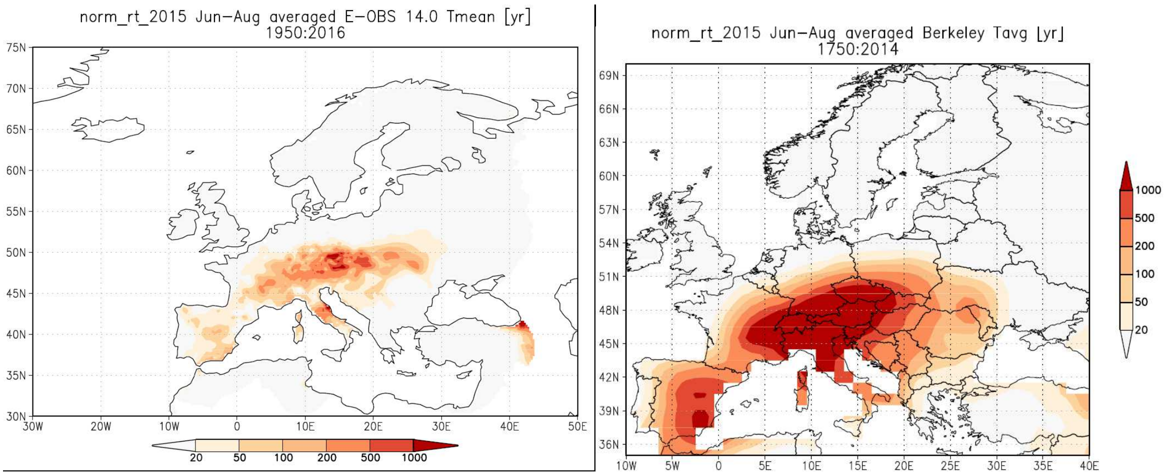


Figure 2 *Return time of the 2015 JJA mean temperatures in the gridded data sets of E-OBS (left hand side) and Berkeley Earth Surface Temperature (right hand side)*

Approach 1: statistical modelling of observations

The near-real-time analysis http://www.climatecentral.org/europe-2015-heatwave-climate-change focussed on 5 individual locations where long, quality-controlled observations are available: De Bilt in the Netherlands, Madrid Airport in Spain, Mannheim in Germany, Beauvais in France, and Zürich in Switzerland. Using the method of van Oldenborgh et al (2007), modelling the statistical moments of a Generalised Extreme Value distribution fitted to the individual records as a component proportional to global average temperature and a residual, that study found a best-estimate increase in the risk of a high temperature event, defined in terms of exceeding the observed threshold for 3-day mean temperature, of a factor of 7 (with a 5% lower bound of 4), 4 (2), 8 (4), 4 (2.4), and 8 (2.5) respectively at the different locations.

For comparison with other studies, this analysis is repeated here applied to annual TXx values (hottest day of the year) for the area-average of the European region shown in figure 1. Spatial averages of an annual index like TXx over a large area are not a particularly impact-relevant quantity, but our aim here is to provide a comparison with other approaches. Results are shown in Figure 3, based on the HadEX2 dataset. Although the temperature offset between the “actual” and “1900” return-time curves is very similar to that found at the individual stations, the probability ratio is much greater because the slope of the return-time curve is much steeper (meaning, return-times increase more rapidly with for a given rise in threshold) for the large-area-averaged quantity, owing to the suppression of variability arising from the averaging. The observed value of TXx, shown by the pink line, is estimated to have a return-time of about 100 years in the current climate in HadEX2, increasing to well over 1000 years (with a lower bound of 200 years) in a climate typical of the year 1900.

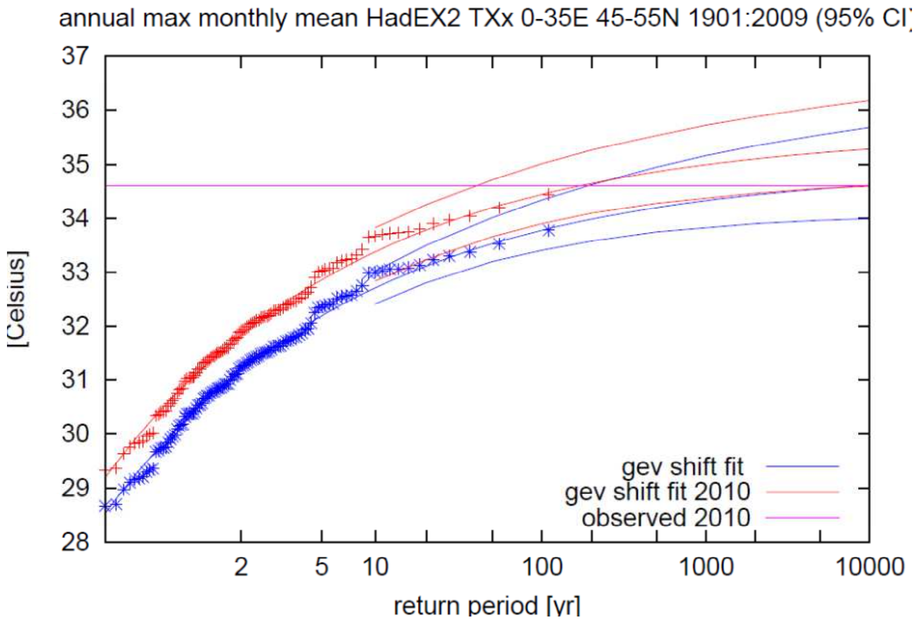


Figure 3. *Maximum daily- and area-averaged temperature TXx over heat wave region in Europe in the current (2010, the last year HadEX2 data are available) climate (red) and in the climate of 1900 (blue), fitted to a GEV distribution with location parameter assumed proportional to global mean temperature. Observed TXx value for 2015 depicted by the horizontal pink line.*

Figure 3 shows the return time of TXx in the heat wave region in the current climate in red and the climate of 1900 in blue as well as the event of 2015 (pink horizontal line) in the HadEX dataset using a GEV fit with global mean surface temperatures as the covariate following the methodology of van Oldenborgh (2007). HadEX is only available until 2010 so it does not contain the 2015 value which is added in pink from Dong et al. (2016).The return time in the current climate is about a 100 years while the same event in the climate of 1900 is so unlikely that a return time can only be given as infinity on average, with a lower bound of 500 years and a best estimate around 5000 years: for such high probability ratios, lower bounds are generally better defined that the best guess or upper bound.

Approach 2: large ensemble simulation of atmosphere/regional climate models

The second approach considered here is large-ensemble simulations with a global atmosphere model (HadAM3P) driving a regional climate model (HadRM3P), comparing simulated weather risks under near-observed conditions for 2015 (“Actual”) and a hypothetical world constructed by removing various estimates of anthropogenic influence on both atmospheric composition and the resulting warming in surface temperatures (see Sippel et al., 2016 for details). Results obtained using this methodology allow us to verify assumptions made using the observations based approach above. Results presented in the near-real-time analysis focussed on local 3-day-averaged temperatures. For comparison with other studies, such as Dong et al (2016), figure 4 repeats the analysis for area averaged maximum daily temperatures, focussing on the entire region shown in figure 1.

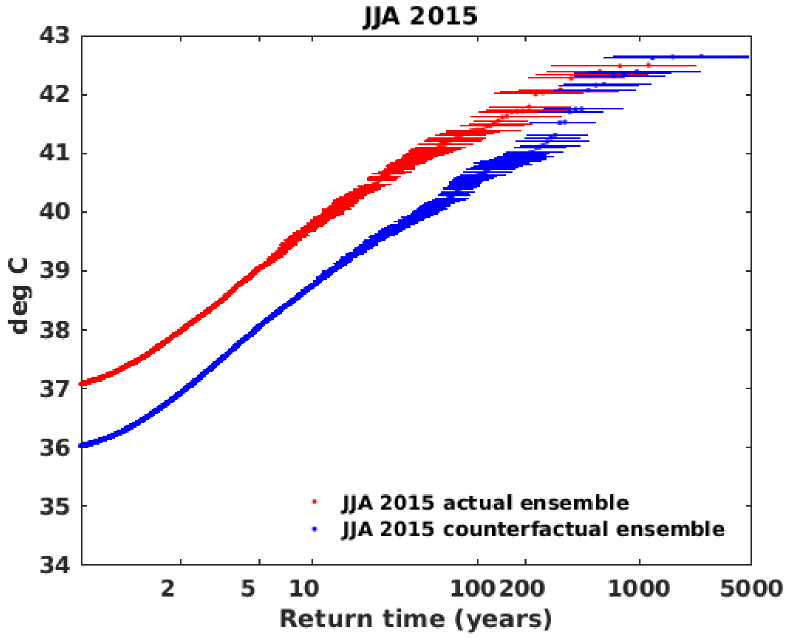


Figure 4. Estimating the return time for maximum daily temperatures in weather@home as would have been done in the real time analysis. In the real-time analysis we only looked at individual gridpoints, this is for the same region as TXx in HadEX above.

Figure 4 shows a significant bias in temperatures for June-August 2015 in weather@home, meaning that it is not possible to directly compare with the temperature of the heatwave from the observations. However, considering a return period of about 100 years in the actual 2015 simulations, this corresponds to roughly a 1 in 250 year event in the natural simulations.

Approach 3: sensitivity to diagnostic -- application to wet bulb temperature

The near-real-time analysis of http://www.climatecentral.org/europe-2015-heatwave-climate-change focussed solely on temperature, but human health impacts, specifically heat stress, depend more on the combination of temperature and humidity, which is well characterised by wet-bulb temperature. Sippel et al. (2016) extended the near-real-time results with a more sophisticated bias correction and analysis of wet bulb temperature.

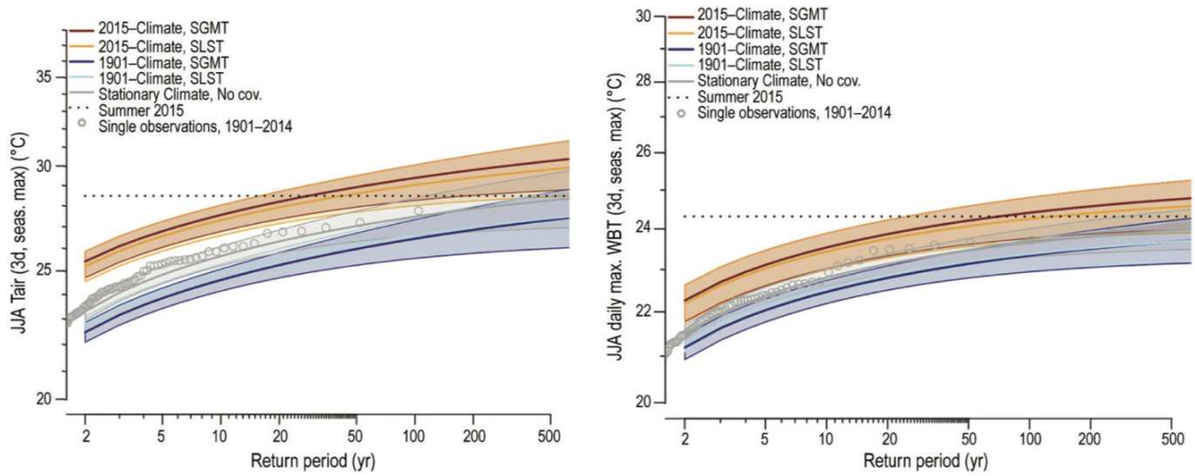


Figure 5. Return times for June-August seasonal maximum, three day mean temperature (a)

and wet bulb temperature (b), at Jena (55^o 56'N 11^{o}35'E) from Sippel et al. Figure 11.1. The data here is based on a GEV fit to the observed station data.

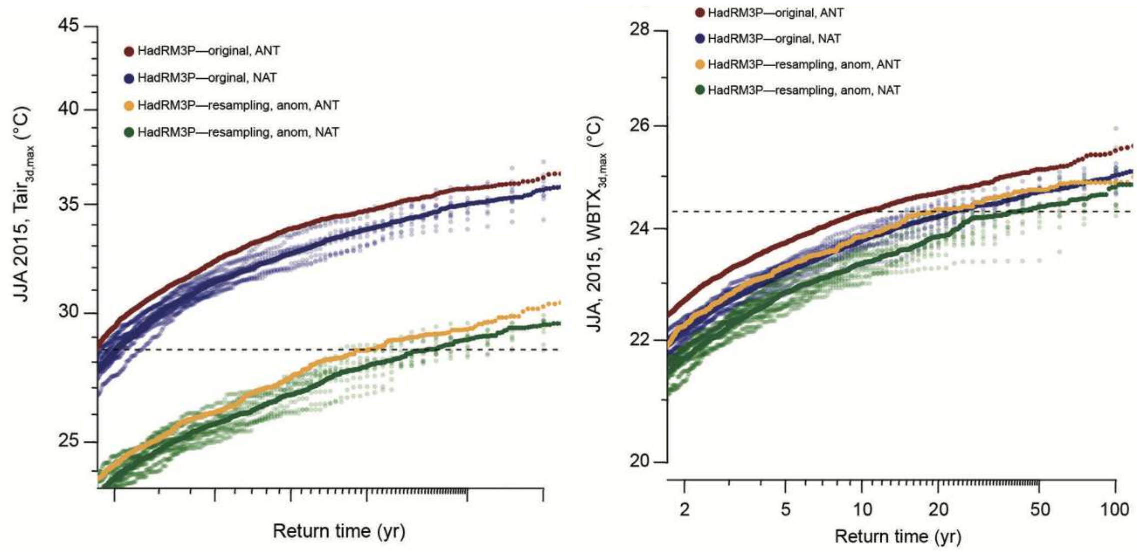


Figure 6. Return times for June-August seasonal maximum, three day mean temperature (a) and wet bulb temperature (b), at Jena (55^o 56'N 11^{o}35'E) from Sippel et al (2016), Figure 11.2. The data here shows the weather@home data for ACT and NAT cases, with and without bias correction.

Figures 5 and 6 show the change in return time for the seasonal maximum three day temperature and wet bulb temperature. The observed data show a large change in return time with the event less than 20 times as likely in 1901 compared to 2015. The weather@home results also show a clear change although smaller than the observations. The bias corrected actual 2015 weather@home simulations give a return time comparable to the observational data for the event of 2015.

These results mainly corroborated the findings that the model estimates of the probability ratio are much smaller than those based on observations partly due to the different framing of the attribution question in both cases (Otto et al., 2016) with the ability in the model to simulate the effect of anthropogenic GHG and aerosol emissions alone while the observations include other drivers of the trend. The second reason is that in a variable like 3-day maximum temperature with a clear upper bound the change in risk strongly depends on the return time hence the results are very sensitive to even small biases. The probability ratio in the wet bulb temperature was found to be very similar to maximum temperatures in all 4 cities analysed. This is in contrast to other studies comparing temperature and wet-bulb temperature attribution (Sippel and Otto 2014) and shows that in a dry heat wave, the maximum temperature actually is a good indicator of the overall change in heat risk.

Approach 4: ensemble simulation of causal drivers of extreme temperature

A second peer-reviewed publication on the 2015 summer heat wave asked the question: 'What caused these anomalous summer conditions over central Europe in 2015?' Dong et al (2015) analysed the respective contributions of anomalous sea surface temperatures compared to the 1964-1993 climatology as well as the GHG and sulphur dioxide emissions. In this study three modeling experiments were undertaken using the Met Office HadGEM2-A global climate model: a control run using 1964-1993 climatological SSTs and greenhouse gas (GHG) and aerosol forcing from the same period, an ensemble simulation of 2015 with observed SST and aerosol concentrations, and a simulation with 2015 SSTs but climatological GHG and aerosol forcings from the 1964-1993 period.

Results are summarised in figure 7, reproduced from Dong et al (2016), showing the increase in surface air temperature (SAT) and various other diagnostics that was attributable to the sea surface temperature and sea-ice extent anomalies (green bars) is generally smaller than the increase attributable to changing composition. The study concluded that the combination of SSTs and atmospheric forcings can explain about 2/3 of the magnitude of the heat wave using several measures averaged over the central European domain, with the remainder due to unpredictable atmospheric noise. They also concluded that atmospheric composition plays a dominant role compared to SSTs and sea ice for most diagnostics of extreme summer temperature, with the exception of night-time temperatures where the roles of SSTs and composition were comparable.

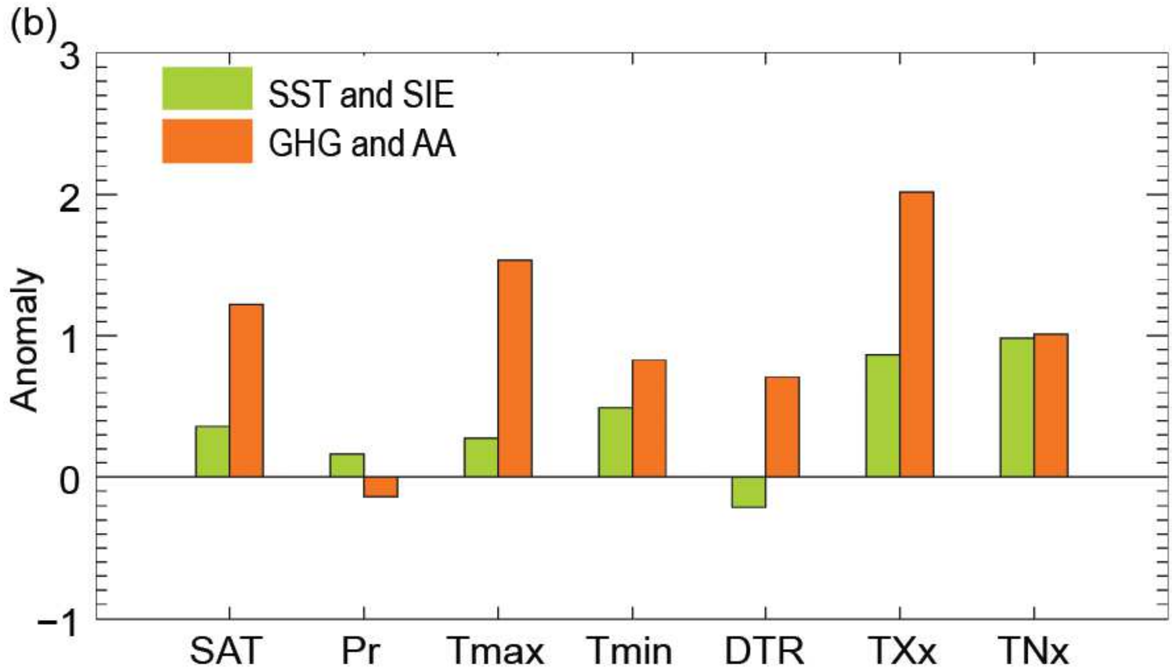


Figure 7: Increase in seasonal-mean (JJA) area-mean surface air temperature (SAT), pressure (Pr), and other diagnostics defined in figure 1 in response to SST and SIE anomalies in 2015 relative to the period 1964-93 (green bars) and in response to atmospheric composition change between 1964-93 and 2015 with no change in SST and SIE (orange bars).

Approach 5: estimating probability ratios from coupled atmosphere-ocean models

A complementary approach to the atmosphere and regional climate model used in approaches 1 and 2 is to compute probability ratios from coupled atmosphere-ocean models, following the method of King et al (2015). Here we employ climate models from the fifth phase of the Coupled Model Intercomparison Project (CMIP5; Taylor et al. 2012) to examine the event with multiple models, analysed in the same way as observations in figure 3 to estimate the change in risk of TXx exceeding a threshold defined by a 1-in-100 return-time event in the climate of the individual models. Attribution is more straightforward here, since counterfactual simulations have been performed with these models confirming that, in each case, almost all of the warming since 1900 is attributable to the model-simulated impact of human influence on climate.

A key question here is the fitness-for-purpose of the coupled models. In total there are 23 different models available that have historical simulations from at least 1900 up to 2005 under observed climate conditions and the last 10 years using RCP 4.5. For this very recent period the exact RCP scenario is relatively unimportant as the temperature response lags the emissions. In order to provide the same analysis applied for HadEX in the coupled model experiments, the models need to be evaluated with respect to their ability to simulate maximum temperatures over the region of interest. It is not possible to use the whole CMIP5 archive as one very large ensemble as data from different models is clearly drawn from different distributions, some of which may be entirely unrealistic. We therefore evaluate the GEV distributions fitted to the TXx over the same region as shown in figure 1 for 23 CMIP5 models running RCP 4.5. Figure 8 shows the scale and shape parameters of these GEV fits.

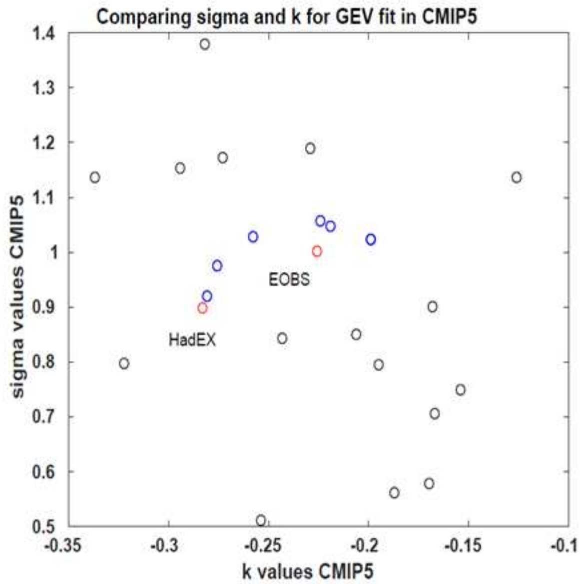
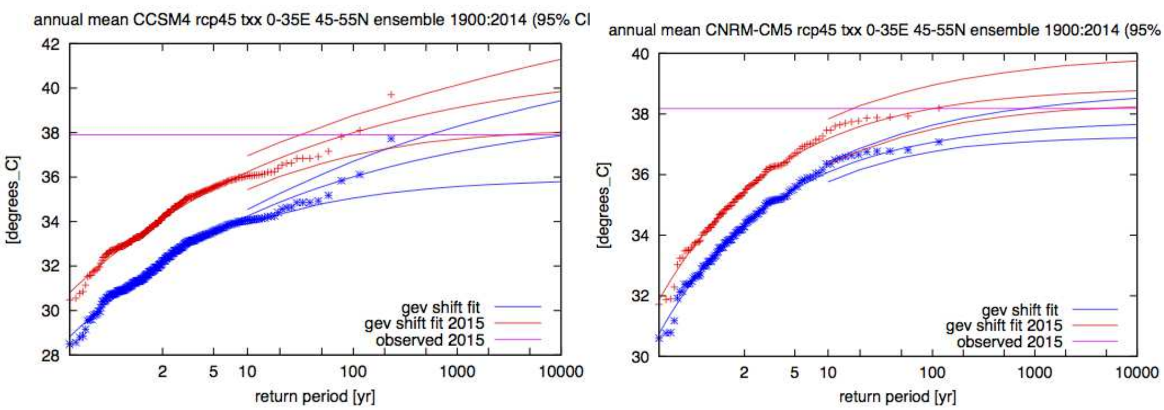


Figure 8 *Scatter plot of scale (sigma) and shape (k) parameter from GEV fits on TXx over the heat wave region indicated in figure 1 in CMIP5 (black circles) and two observational estimates (HadEX and E-OBS, red circles) for the common 1900-2015 period. The blue circles show the models closest to observations. The multimodel mean is not shown as it is not possible to fit a GEV distribution.*

Following the assessment in figure 8 it is now possible to define a maximum distance from the GEV parameters fitted to observed data to subsequently examine return time plots for the “good” CMIP5 GCMs for TXx in 2015 and 1900, applying the same methodology as in figure 3. Using a relatively arbitrary distance from the observed values we have identified 7 models for further analysis, marked in blue in figure 8. Return-time plots for these models are shown in figure 9.



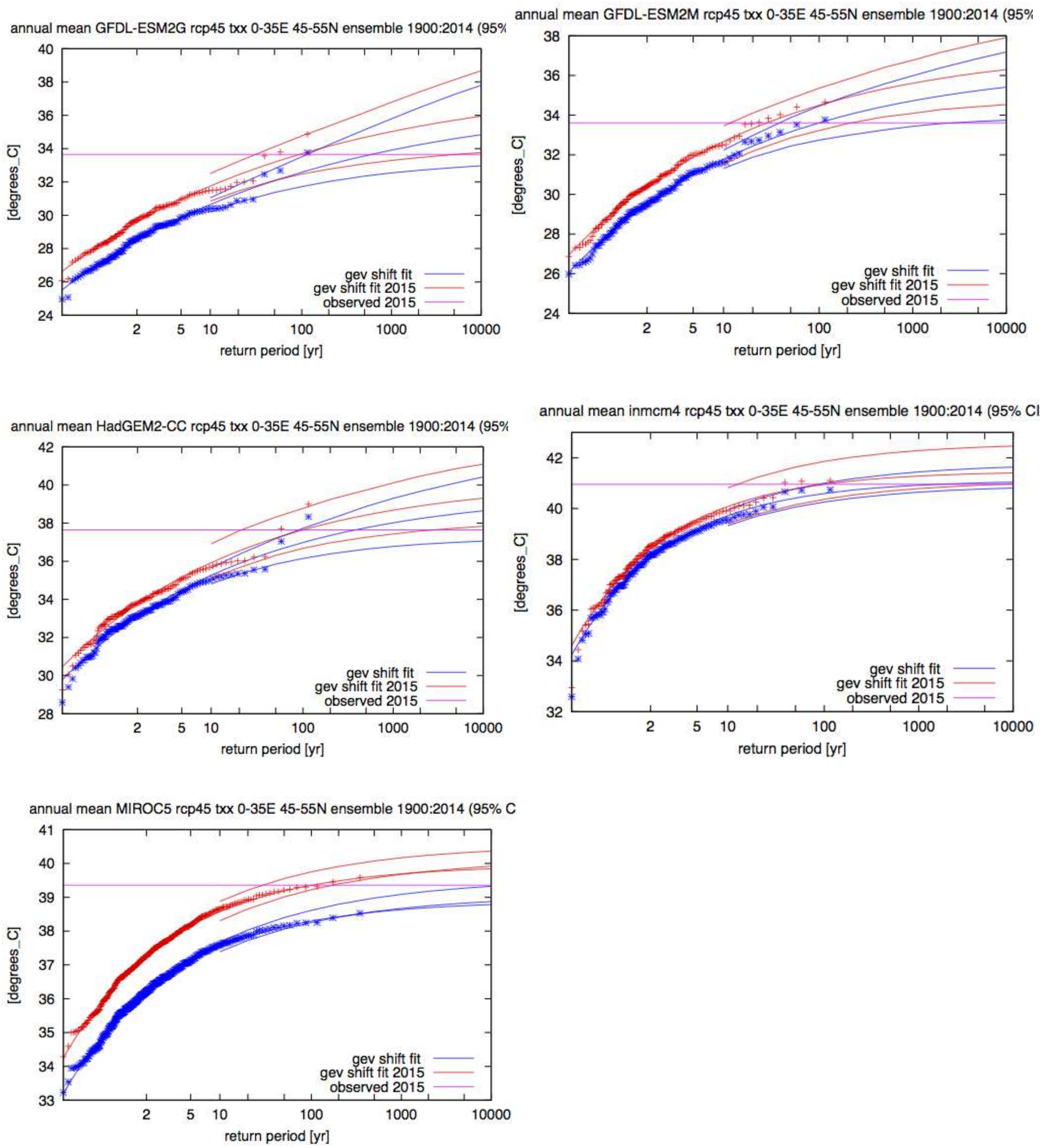


Figure 8. *Return times calculated using the same method as in figure 3 but for the 7 different CMIP5 models. The vertical pink line always indicates the 1 in 100 year event.*

Applying the same methodology to these six models the probability ratios (shown in figure 9) can be obtained. While the shape and scale parameters in these models are similar, the location parameters are not. Hence, instead of correcting for these offsets the 1 in a 100 year event is analysed. Figure 9 shows in blue the average change in risk and in red the lower bound of the overall change in risk for a 1 in 100 year event in the current climate. While the average probability ratios range from infinity to a five-fold increase in risk the lower bounds are better constrained, ranging from just positive to a twenty-fold increase in the risk of an extreme event. This very large range of probability ratios highlights that while a qualitative attribution of heat waves to anthropogenic climate change is relatively straight forward, the quantitative assessment is not.

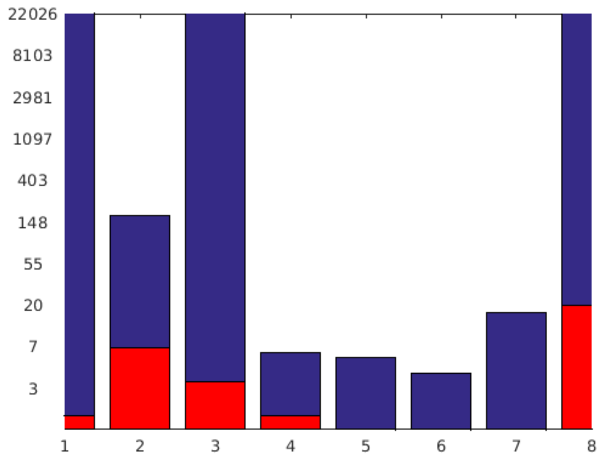


Figure 9 Mean (blue) and lower bound (red) of probability ratios for the 1 in 100 year event calculated from return time graphs as in figure 2 for HadEX (1) and the 7 closest CMIP5 models with respect to shape & scale parameter (2 = CCSM4, 3 = CNRM-CM5, 4 = GFDL-ESM2G, 5 = GFDL-ESM2M, 6= HadGEM2-CC, 7=inmcm4, 8=Miroc5). The mean probability ratio in HadEX (1), CNRM-CM (3) and Miroc5 (8) is infinity as a TXx value as high as the 1 in 100 year event in the current climate is not observed in the pre-Industrial estimates.

Approach 6: role of circulation estimated using atmospheric analogues

From thermodynamic considerations alone, we expect heatwaves like the one observed in 2015 in central Europe to become more likely in a warming climate. Anthropogenic climate change may, however, also affecting the atmospheric circulation and through this could either amplify or counteract the increase in occurrence frequency of the event from the thermodynamics (Otto et al. 2016).

One way of estimating the role of the circulation in the likelihood of the event occurring was introduced by Yiou et al (2007) and Cattiaux et al (2010). In this “flow-analogues” method consists of selecting thirty flow-analogues among the 1949 to 2014 summers in a 30-day window centered on the 5th of July 2015, a day in the middle of the first heatwave. The selection is made on the basis of minimizing the Euclidean distance of daily MSLP maps over central Europe and the north Atlantic as shown in figure 10. Using linear correlation for the ranking or Spearman correlation coefficient and using 100 instead of 30 analogues does not change our results significantly.

Applying this method, analogues of the current circulation state are defined and similar analogues are analysed in NCEP reanalysis data. We find that the number of analogues that shows similar sea level pressure patterns and magnitudes to the observed 2015 circulation are much more frequent in recent decades than in the 50s 60s or 70s (Figure 11). This is particularly apparent when only looking at the small number of analogues and thus very close analogues. With the distance function of maximal 500 Pa over the region averaged. If the number of analogues per year is higher the difference in the number of analogous per decade is still there but not as pronounced as with the smaller number however on this larger number of analogues the maximum distance from the observed analogues is 601 Pa. A caveat here is that the NCEP data is not detrended so a simple increase in pressure in a warming climate will play a certain role. This effect alone does however not explain the increase in the number of analogues in more recent decades and suggest that circulation may be playing a role in increasing the likelihood of Central European heatwaves as observed in 2015.

As was shown for the first time in Vautard et al., (2016) and using a slightly different methodology in Yiou et al., (2017) for extreme rainfall in the UK, it is possible to apply the “flow-analogue” methodology to climate model simulations of possible weather in the current climate as well as in counterfactual simulations as used in the large-ensemble modelling approach. In doing so, it will be possible to quantify what percentage of the overall change in risk is due to thermodynamic effects compared to the change in risk due to changes in the atmospheric circulation. While further research is needed for a quantitative assessment, these preliminary results suggest that circulation is unlikely to be acting strongly against the thermodynamic impact of a warming climate. This is important, because modelling the impact of increased greenhouse gases on atmospheric circulation remains a challenge, and hence if there were evidence of a strong role of circulation changes reducing the risk of flow patterns similar to that observed in 2015 (which is not observed in models), this would substantially reduce our confidence in our model-based results. In fact, our preliminary results suggest the opposite: if anything, the trend is towards an increased prevalence of this summer blocked flow. These considerations hold under the assumption that the reanalysis data are a good approximation of the real world. Whether this assumption holds for NCEP is however questionable as different reanalysis data sets show very different trends in summer circulation patterns.probability ratio

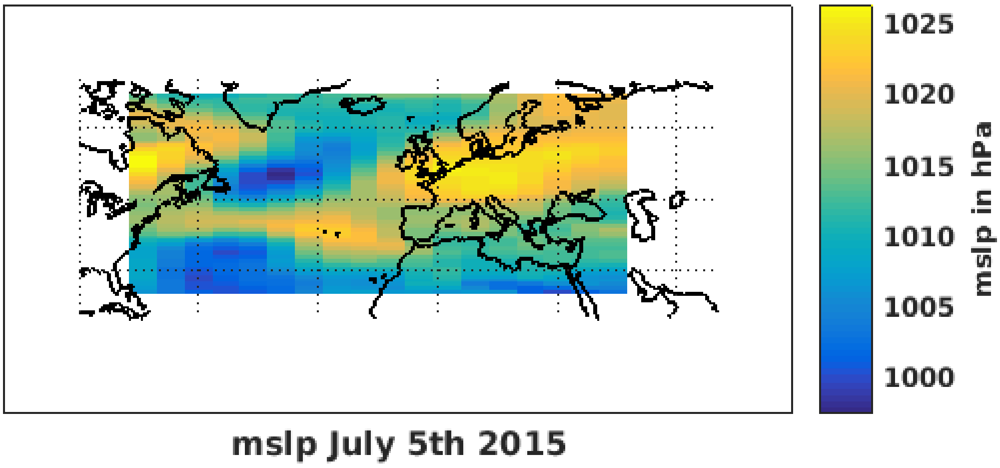


Figure 10. *MSLP over Central Europe and the North Atlantic on 5th of July in NCEP reanalysis. This is just an example of the flow, defined as mslp over europe and the north atlantic which is used to calculate the analogues. On all days during the early heatwave in Western Europe the flow looks very similar.*

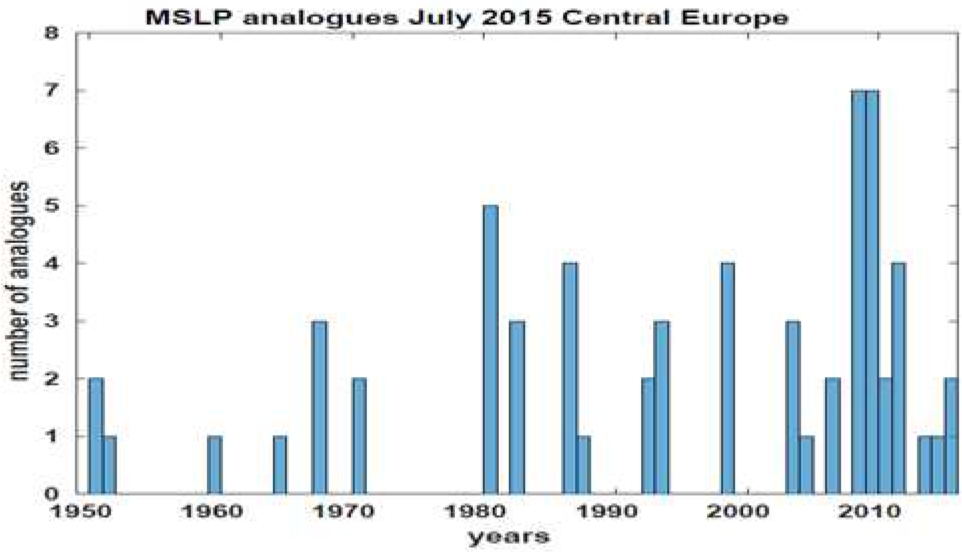


Figure 11 Number of analogues to the observed July 2015 MSLP per year. These are the number of analogues per year using MSLP over the region depicted in figure 4 as analogue and calculating only 30 analogues.

Discussion

The application of a range of attribution approaches to the same weather event, the 2015 European heatwave, provides an excellent opportunity to compare results and draw some general methodological conclusions. The most obvious is that, for this particular event, the role of human influence on climate is very clear, independent of the method of attribution or the formal definition of the event: increased greenhouse gas concentrations and consequent large-scale anthropogenic warming substantially increased the probability of occurrence of a heatwave in 2015.

That said, the amount by which the probability is increased, or the probability ratio, is highly sensitive to the definition of the event, in particular the spatial and temporal averaging interval considered. The main reason for this is that averaging tends to reduce the noise more than the signal if the averaging is over intervals larger and longer than the decorrelation scales but smaller than the event scale. Near-real-time analysis performed focussing on high 3-day-mean temperatures in individual station records and grid-box-scale output from a 25km resolution regional climate model indicated probability ratios in the region of 2-8. A repeated analysis of observational data focussing on the hottest day for average temperatures over a large region indicates a probability ratio of 50 or more (figure 3), but the corresponding analysis of the regional climate model, also averaging over a large area, indicated a similar probability ratio to the local results: a factor of 2.5. The reasons for this discrepancy between observationally-based and model-based results remains under investigation, and may indicate a role of cancellation of errors in the good agreement observed between the two approaches in the near-real-time analysis. Analysis of coupled climate models (figures 8 and 9) indicate a broad range of probability ratios, with best-estimates from different models spanning these two extremes. This illustrates, however, that it is important, if probability ratios are quoted with respect to a particular impact, to focus on a spatial and temporal scale relevant to that impact.

Focussing on the correct diagnostic is also important, as illustrated by figures 5 and 6, which show that when the intensity of heatwaves is measured in terms of wet-bulb-temperature, an indicator most relevant to heat stress and health impacts, the risk ratio for the contribution of human influence to the 2015 summer heatwave was substantially smaller than if intensity is measured in terms of temperature alone. Hence it is important, where this can be identified, to focus on the correct impact-relevant diagnostic if an attribution study is prompted by a high-impact weather event

Causal analyses identifying mechanisms whereby human influence on climate or other external factors may be impacting the risk of a specific extreme weather event, while not contributing directly to the assessment of relative risk with and without human influence, can either considerably strengthen conclusions or provide a valuable note of caution. The analysis of Dong et al (2016) demonstrated that most of the change in expected summer temperatures and temperature extreme indices in 2015, relative to the period 1963-94, arises from changing atmospheric composition, with the change in sea surface temperatures and sea ice playing an important but smaller role. This significantly strengthens confidence in attribution of the increase in risk of this event to greenhouse gas emissions, because the causal link between emissions and changing composition is much less contentious than the link between emissions and large-scale temperatures.

A preliminary analysis using the method of weather analogues (Cattiaux et al, 2010) indicates that the probability of occurrence of the pattern of surface pressure that characterised the onset phase, at least, of the 2015 heatwave appears to have increased substantially over recent decades. Attribution of this change to human influence on climate is, as yet, inconclusive, but this provides some reassurance that changing atmospheric circulation is not working against the thermodynamic impact of a warming atmosphere, which provides greater confidence in model-based attribution results, given than the response of circulation to human influence on climate remains a matter of contention.

Finally, an important question this example allows us to address is the potential for near-real-time attribution. Re-assessment of near-real-time results in peer-reviewed publications found little change in quantitative conclusions resulting from a more detailed model evaluation and refinement of analysis methods (Sippel et al, 2016). Of much greater significance than the near-real-time versus post-hoc-analysis distinction is the choice of impact variable and spatio-temporal averaging used to define the event.

This is potentially significant for the design of an attribution service, because it illustrates the importance of close collaboration between the meteorological research community and the impact modelling and disaster vulnerability and exposure research communities to ensure that the correct impact-relevant diagnostics are used to characterise a high-impact event. This is particularly important for events such as drought-induced crop failure or the human health impacts of heat stress where multiple meteorological variables contribute to the impact. There is also an important social science aspect to this question, since what stakeholders or the general public perceive to be the most important aspects of an extreme weather event may not be immediately apparent. Close integration and collaboration, informed by social science, between event attribution services and potential users of attribution information will be important going forward.

References

Blade, I., B. Liebmann, D. Fortuny, and G. van Oldenborgh, Observed and simulated impacts of the summer NAO in Europe: implications for projected drying in the Mediterranean region, Climate Dynamics, 39 (3-4), 709{727, doi:10.1007/s00382-011- 1195-x, 2012.

Cattiaux, J., R. Vautard, C. Cassou, P. Yiou, V. Masson-Delmotte, and F. Codron (2010), Winter 2010 in Europe: A cold extreme in a warming climate, Geophys. Res. Lett., 37, L20704, doi:10.1029/2010GL044613.

Donat, M. G., et al. (2013), Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, J. Geophys. Res. Atmos., 118, 2098–2118, doi:10.1002/jgrd.50150

Dong, B., and Coauthors, “The 2015 European Heat Wave” [in “Explaining Extremes of 2015 from a Climate Perspective”]. Bull. Amer. Meteor. Soc., 97 (12), S14–S18, doi:10.1175/BAMS-D-16-0149.

Haarsma, R. J., F. M. Selten, B. J. J. M. van den Hurk, W. Hazeleger, and X.Wang, Dry mediterranean soils due to greenhouse warming bring easterly winds over summertime central europe, Geophys. Res. Lett., 36, L04,705, doi:10.1029/2008GL036617, 2009.

RJ Haarsma, FM Selten, SS Drijfhout, Decelerating Atlantic meridional overturning circulation main cause of future west European summer atmospheric circulation changes Environmental Research Letters 10 (9), Environ. Res. Lett. 10(2015) 094007 doi:10.1088/1748-9326/10/9/094007

Haylock, M. R., et al. "A European daily high□resolution gridded data set of surface temperature and precipitation for 1950–2006." Journal of Geophysical Research: Atmospheres 113.D20 (2008).

King, Andrew D., et al. "Attribution of the record high Central England temperature of 2014 to anthropogenic influences." Environmental Research Letters 10.5 (2015): 054002.

Miralles, D. G., A. J. Teuling, C. C. van Heerwaarden, and J. Vila-Guerau de Arellano, Megaheatwave temperatures due to combined soil desiccation and atmospheric heat accumulation, Nature Geosci, 7 (5), 345{349, doi:10.1038/ngeo2141, 2014. 2014.

Mitchell, Daniel, et al. "Attributing human mortality during extreme heat waves to anthropogenic climate change." Environmental Research Letters 11.7 (2016): 074006.

Otto, F. E. L., 2016: Extreme events: the art of attribution. Nat. Climate Change, 6, 342–343, doi:10.1038/nclimate2971.

Polade, S. D., D. W. Pierce, D. R. Cayan, A. Gershunov, and M. D. Dettinger, The key role of dry days in changing regional climate and precipitation regimes, Sci. Rep., 4, doi:10.1038/srep04364, 2014.

Rowell, D. P., and R. G. Jones, Causes and uncertainty of future summer drying over Europe, Clim. Dyn., 27, 281-299, 2006.

Seneviratne, S. I., et al., Impact of soil moisture-climate feedbacks on cmip5 projections: First results from the glace-cmip5 experiment, Geophysical Research Letters, 40 (19), 5212-5217, doi:10.1002/grl.50956, 2013.

Sippel, S., and Coauthors, “The Role of Anthropogenic Warming in 2015 Central European Heat Waves” [in “Explaining Extremes of 2015 from a Climate Perspective”]. Bull. Amer. Meteor. Soc., 97 (12), S14–S18, doi:10.1175/BAMS-D-16-0149.

Stott, Peter A., et al. "Attribution of extreme weather and climate□related events." Wiley Interdisciplinary Reviews: Climate Change 7.1 (2016): 23-41.

Taylor, K. E., R. J. Stou er, and G. A. Meehl, 2012: An overview of CMIP5 and the experimental design. Bull. Amer. Meteor. Soc., 93, 485–498, doi:10.1175 /BAMS-D-00094.1.

Teuling, A. J., et al., Evapotranspiration ampli_es European summer drought, Geophys. Res. Lett., 40 (10), 2071-2075, doi:10.1002/grl.50495, 2013.

Uhe, P., F. E. L. Otto, K. Haustein, G. J. van Oldenborgh, A. D. King, D. C. H. Wallom, M. R. Allen, and H. Cullen (2016), Comparison of methods: Attributing the 2014 record European temperatures to human influences, Geophys. Res. Lett., 43, 8685–8693, doi:10.1002/2016GL069568.

Ronald van Haren , Reindert J Haarsma, Hylke de Vries, Geert Jan van Oldenborgh and Wilco Hazeleger. Resolution dependence of circulation forced future central European summer drying. Environ. Res. Lett.10 (2015) 055002, doi:10.1088/1748-9326/10/5/055002, 2015.

Vautard, R., P. Yiou, F. D’Andrea, N. de Noblet, N. Viovy, C. Cassou, J. Polcher, P. Ciais, M. Kageyama, and Y. Fan (2007), Summertime European heat and drought waves induced by wintertime Mediterranean rainfall deficit, Geophys. Res. Lett., 34, L07711, doi:10.1029/ 2006GL028001.

Vautard et al. The simulation of European heat waves from an ensemble of regional climate models within the EURO-CORDEX project. Clim Dyn (2013) 41:2555-2575 DOI 10.1007/s00382-013-1714-z, 2013.

Vautard, R., Yiou, P., Otto, F., Stott, P., Christidis, N., van Oldenborgh, G.J. and Schaller, N. (2016) Attribution of human-induced dynamical and thermodynamical contributions in extreme weather events. Environmental Research Letters, 11(11). 114009.
Vidale, P., D. L uthi, R. Wegmann, and C. Sch ar, European summer climate variability in a heterogeneous multi-model ensemble, Climatic Change, 81 (1), 209{232, doi: 10.1007/s10584-006-9218-z, 2007.

Yiou, P., R. Vautard, P. Naveau, and C. Cassou (2007), Inconsistency between atmospheric dynamics and temperatures during the exceptional 2006/2007 fall/winter and recent warming in Europe, Geophys. Res. Lett., 34, L21808, doi:10.1029/2007GL031981.

Yiou, P., A., Jezequel, P. Naveau, F.E.L. Otto, R. Vautard, M.Vrac(2017) A statistical framework for conditional extreme event attribution. under review.

Wilks, D., 2016: "The Stippling Shows Statistically Significant Grid Points": How Research Results are Routinely Overstated and Overinterpreted, and What to Do about It. *Bull. Amer. Meteor. Soc.,* **97**, 2263–2273, doi: 10.1175/BAMS-D-15-00267.1.

4. Lessons Learnt

The choice of the impact variable and the spatio-temporal averaging used to define the event are essential in event attribution studies. Qualitative assessments of the role of anthropogenic climate change are robust to the event definition, time-scale of analysis and attribution methodology employed.

Close collaboration between the meteorological research community and the impact modelling and disaster vulnerability and exposure research communities is important, in order to ensure that the correct impact-relevant diagnostics are used to characterise a high-impact event.

5. Links Built

This work required collaboration with WP6 to ensure the models used in the analyses are fit for purpose. We also employed fast-track methodologies developed by WP8 and contributed the case study of a heatwave event to WP7.