Author/s:
Uhe, P;Otto, FEL;Haustein, K;van Oldenborgh, GJ;King, AD;Wallom, DCH;Allen, MR;Cullen, H

Title:
Comparison of methods: Attributing the 2014 record European temperatures to human influences

Date:
2016-08-28

Citation:
Uhe, P., Otto, F. E. L., Haustein, K., van Oldenborgh, G. J., King, A. D., Wallom, D. C. H., Allen, M. R. & Cullen, H. (2016). Comparison of methods: Attributing the 2014 record European temperatures to human influences. Geophysical Research Letters, 43 (16), pp.8685-8693. https://doi.org/10.1002/2016GL069568.

Comparison of Methods: Attributing the 2014 record European temperatures to human influences

P. Uhe[1,6], F. E. L. Otto[1], K. Haustein[1], G. J. van Oldenborgh[2], A. King[3], D. Wallom[6], M. R. Allen[1,5] and H Cullen [4]

[1]Environmental Change Institute, University of Oxford, Oxford, UK.

[2]Royal Netherlands Meteorological Institute (KNMI), P.O. Box 201, 3730 AE De Bilt, The Netherlands

[3]ARC Centre of Excellence for Climate System Science, School of Earth Sciences, University of Melbourne, Melbourne, Australia

[4]Climate Central, One Palmer Square, Suite 330, Princeton, NJ 08542, USA

[5]Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, UK

[6]Oxford e-Research Centre, University of Oxford, Oxford, UK

Key Points:

- Human influences played a strong role in the record 2014 temperatures in Europe
- An attribution result is very sensitive to the spatio-temporal definition of the event
- The increase in risk due to climate change declines on average moving to smaller spatial scales

Abstract

The year 2014 broke the record for the warmest yearly average temperature in Europe. Attributing how much this was due to anthropogenic climate change and how much it was due to natural variability is a challenging question, but one that is important to address. In this study, we compare four event attribution methods. We look at the the Risk Ratio (RR) associated with anthropogenic climate change for this event, over the whole European region, as well as its spatial distribution. Each method shows a very strong anthropogenic influence on the event over Europe. However, the magnitude of the RR strongly depends on the definition of the event and the method used. Across Europe, attribution over larger regions tended to give greater RR values. This highlights a major source of sensitivity in attribution statements and the need to define the event to analyse on a case-by-case basis.

1 Introduction

The warming of the global climate system is unequivocal [*Stocker et al.*, 2013] and 2014 broke the record for the warmest year in Europe, in the observational record [*Photiadou et al.*, 2015]. Record-breaking daily temperatures over Europe have increased in recent decades due to climate change [*Bador et al.*, 2015], and record-breaking annual mean temperatures are also expected to increase. Europe, in 2014, saw many records of high annual temperatures broken regionally (e.g. in central England, *King et al.* [2015]) as well as the continent as a whole. A signal attributable to increasing anthropogenic greenhouse gas (GHG) emissions has been robustly detected over the last decades in large-scale temperature extremes [*Houghton et al.*, 2001; *Solomon et al.*, 2007; *Stocker et al.*, 2013] and over Europe [*Bador et al.*, 2015], however the quantification of the signal for an individual event is less straightforward.

An increase in temperatures over seasons as well as on shorter timescales puts strain on society, ecosystems, and infrastructure. Hence, it is valuable to determine to what extent, if any, anthropogenic climate change has contributed to the frequency of temperature records being broken. This quantification improves our understanding of the impacts of rising GHG emissions and helps to begin assessing the socioeconomic costs of anthropogenic climate change today. It remains under debate whether such estimates are of immediate relevance for climate change policies [*Hulme*, 2014; *Thompson and Otto*, 2015; *James et al.*, 2014]. However, whether or not anthropogenic climate change played a detectable role in the likelihood of current events is already debated outside the scientific community so the inclusion of scientific evidence can only improve this discussion.

To understand the implications of event attribution results, the design of the attribution study needs to be considered [*Otto et al.*, 2015]. It is particularly important that event attribution findings not only state the answer but also the exact question that is addressed. For example, attributing the change of likelihood of an extreme event due to the observed trend answers a slightly different question to attributing the change in likelihood between modelled climates with and without anthropogenic greenhouse gas emissions. In addition, using a coupled ocean-atmosphere model changes the attribution question slightly compared with using a model with prescribed sea-surface temperatures (SSTs).

Employing multiple methodologies allows a quantification of the uncertainties in the attributable signal, and provides a consistency check in the results between each method. This leads to greater confidence in the quantification of the attribution of human influences. In this study we contrast the results of four methodologies, applied to the question: how much did anthropogenic climate change alter the likelihood of the European record annual mean temperature of 2014? Each of these methods takes a different approach and involves slightly different assumptions regarding the modelling of anthropogenic influence on climate.

In sections 2.1–2.3, each of the four methodologies are introduced, then compared in section 2.4. Values for the Risk Ratio (RR) calculated for each of the methodologies are presented in section 3.1. Expanding on this, section 3.2 looks at the spatial distribution of the RR and section 3.3 investigates how the RR varies when different sized regions are analyzed. In conclusion (section 4) we find that while the dependency on the region size is expected, the extent to which results differ is very large. This highlights the need for attribution studies to be conducted on a case-by-case basis.

2 Methodology and Validation

Four methodologies of event attribution, assessing to what extent anthropogenic climate change altered the likelihood of a year as warm as observed in 2014, are considered in this paper: (1) empirical methodology using historical data [*van Oldenborgh*, 2007; *van Oldenborgh et al.*, 2012], (2) methodology employing the Coupled Model Intercomparison Project Phase 5 (CMIP5; *Taylor et al.* [2012]) ensembles, using historical and historicalNat simulations [*King et al.*, 2015] and (3 and 4) methodologies using very large ensembles of simulations in the weather@home framework [*Massey et al.*, 2014]. The RR is defined as $P_{\mathrm{actual}}/P_{\mathrm{natural}}$ where P_{actual} and P_{natural} are respectively the probabilities of the event occurring in the actual world with climate change and in the world without human influence. For example, a RR of 10 means the event is 10 times more likely in the current climate compared to a world without climate change. Note that the definition of this counterfactual world differs between methods (see below for details). For each of the methodologies, gridded datasets are used and each gridbox of this data can be analysed independently, or they can be averaged over sub-regions or the whole European domain.

2.1 Empirical event attribution using historical data records

To use the observational record for event attribution, a two-step process is performed. Firstly, the observed trend since pre-industrial times is computed (excluding the event itself). Secondly, the event is related quantitatively by physical and observational arguments to a quantity for which a model-based attribution study has been performed. This can be used to estimate the extent that anthropogenic influences have contributed to the trend.

For the first step we used a long observational dataset. We chose the Berkeley Earth Surface Temperatures [*Rohde et al.*, 2013] dataset as its very large decorrelation scales match the value needed for annual mean temperature, while suppressing inhomogeneities. The area with data in 30°–76° N, 25° W–45° E was taken to be the European temperature series.

The trend in extremes was determined by fitting a Generalised Pareto Distribution (GPD) to the top 20% of the distribution above a threshold that varies linearly with the low-pass filtered global mean temperature as a measure of global warming [*King et al.*, 2015]. A non-parametric bootstrap of 1000 samples gives an estimate of the uncertainties. The inverse GPD is evaluated in 1901 and 2014 to determine the return times for the current climate and the pre-industrial climate respectively. The ratio between the return times describes the increase of likelihood of the 2014 event occurring. Details can be found in Text S1.

The second step, attribution, is done by referring to *Stott* [e.g., 2003], who find good agreement between the CMIP5 modelled trend in mean temperature over Europe and the observed trend, but not with the results of simulations using only natural forcings. This allows us to attribute the increase in likelihood of the event to anthropogenic forcings. The uncertainties of this attribution step are not included in the results.

2.2 Event attribution using the CMIP5 archive

In the second approach we compare a set of multi-model simulations representing the world as it was in the years around 2014 with simulations representing the world had humans not influenced the climate. This methodology follows that of *Lewis and Karoly* [2013] and *King et al.* [2015].

Firstly, 81 historical simulations from 17 different CMIP5 models were regridded onto a common 1.5° by 1.5° regular grid over the European region (30°–75° N, 12° W–45° E). The E-OBS dataset was interpolated to the same resolution. For both the observations and all model simulations, European average annual time series were calculated as anomalies from the 1961-1990 climatological period.

A Kolmogorov-Smirnov (KS) test was then employed to determine whether the model simulations adequately represented the European-average variability in annual temperature anomalies in the observational record over the common 1951-2005 period. To account for the high lag correlations in the observed time series (approximately 0.6 at a 1-year lag) the degrees of freedom used to determine the significance of the KS-statistic were reduced. Models were deemed to have passed this test if at least three historical simulations were available for analysis and no more than one of the historical simulations was significantly different ($p<0.05$) from the observational time series. This resulted in the selection of simulations from 15 climate models to be investigated further.

To use the HistoricalNat simulations in the RR analysis, the temperature anomalies were adjusted (from the base period of 1961-1990) by accounting for the warming prior to that period. This was done as per *King et al.* [2015], and resulted in a shift of the HistoricalNat PDFs for each model depending on the magnitude of warming seen in each model. The CMIP5 models and the number of ensemble members used are shown in Table S1.

Using 41 RCP8.5 (2006-2020) and 50 HistoricalNat (1900-2005) runs from these models, an estimate of the RR was calculated. Ten thousand estimates of the RR were calculated by bootstrap resampling 50% of the pairs of RCP8.5 and HistoricalNat simulations. Using these 10000 RR estimates, a median RR and 10th percentile RR were calculated.

2.3 Event Attribution using weather@home

The final methods use the distributed computing framework climate*prediction*.net to produce large ensembles of climate simulations. These experiments are part of the weather@home project [*Massey et al.*, 2014], and use the Met Office Hadley Centre regional climate model HadRM3P over the EURO-CORDEX domain [*Jacob et al.*, 2014], at 0.44° horizontal resolution, embedded in the atmosphere-only global circulation model HadAM3P at 1.25° x 1.875° resolution.

For this analysis, the return time is calculated from the fraction of ensemble members which have an annual mean temperature for 2014 greater than observed (employing the Berkeley data) in a particular region. The Berkeley (weather@home) data used were anomalies relative to observations (simulations) from 1961–1990. This method of determining the likelihood of an event does not depend on the length of an available temperature record to obtain significant results. However, the uncertainty depends on the size of the ensembles, and return times of extremely rare events will still be highly uncertain.

This study uses four weather@home ensembles. The first ensemble (actual2014) provides simulations forced by observed SSTs of 2014, as well as current GHG and aerosol concentrations. The second ensemble (actualClim), is used for the estimation of the return times under current climatological conditions. This combines ensembles of annual simulations of possible weather for the years 2001–2012. For the actual2014 and actualClim ensembles, the model is forced with observed SSTs and sea ice extent using the OSTIA data set [*Stark et al.*, 2007].

The third ensemble (natural2014) is a counterfactual ensemble of the year 2014 without anthropogenic climate change. As per *Schaller et al.* [2014], the anthropogenic signal was removed from the OSTIA SSTs used to force the simulations. As the uncertainty in estimating the pattern of anthropogenic warming is large, 12 different patterns from CMIP5 models were used (see Text S2). We also employ a fourth ensemble representing the 2001–2012 period (naturalClim) which has a similar set-up as the actualClim ensemble but with the anthropogenic influences removed in the same way as the natural2014 ensemble. However, it only uses a single model (HadGEM2-ES) to represent the anthropogenic change in the SSTs.

For a rare event, large ensembles are needed to obtain statistically significant results. The largest return times that can robustly be computed without using extreme value statistics can be an order of magnitude less than the ensemble size [*Sippel et al.*, 2015]. The ensemble sizes are given in Text S2.

2.4 Comparison of approaches

Each method described here has strengths and weaknesses, and relies on different assumptions. The empirical methodology reveals how the return time of the event in the observed temperature record is different in the current climate compared to the climatology of an earlier time period.

This methodology assumes that the climatology changes in a simple way as a function of time or global mean temperature, e.g. shifting or scaling. It assumes that a GPD is a reliable representation of the distribution. It also relies on the accuracy of the observational dataset for the whole historical record. It has the advantage of attributing the change in likelihood to the observed trend without any other complicating factors, but is limited by the amount of data in the historical record. The empirical method works best for small-scale events in space and time, as there are many data points to analyse for these events.

When using climate models, the response of the climate to anthropogenic influence is determined by the physics of each model. Models are limited by resolution and the need to parametrize small scale or complex phenomena such as land surface, clouds and convection. These parametrizations are imperfect and may result in systematic biases in the model. The return time in the world without climate change is obtained from a counterfactual (natural) ensemble without anthropogenic forcings instead of comparing with observations of the early 20th century. The use of many simulations and different models imply that the uncertainties in the response to these forcings are included.

The CMIP5 method has the advantage of using highly complex coupled models with a range of different physics. It is most sensitive to structural differences between climate models. However, because of the additional complexity of the models, they have a coarser resolution and fewer simulations compared to the weather@home ensemble.

The weather@home method is able to produce thousands of simulations to increase the statistical significance of the results when the signal-to-noise ratio is poor, however it only includes one representation of the physics of the climate system. Weather@home also answers different questions by prescribing the SSTs [*Otto et al.*, 2015]. In the third method, by prescribing the SSTs of 2014, we attribute the change in likelihood of the 2014 event given everything else being equal. In the fourth method, when using simulations from 2001–2012, we look at the change in likelihood of the 2014 event under current climatic conditions, without requiring the exact patterns of SSTs observed in 2014. This renders it more comparable to the framing in the CMIP5 methodology while employing a very different modelling approach.

3 Results

3.1 Risk ratio for 2014 temperature event

The increase in likelihood of high average temperature over Europe is robust across each of the methods presented here, even though the return times vary considerably. The one sided 90% confidence intervals of the RR were calculated, and the RR was found to be at least 500 using the empirical method. The weather@home and CMIP5 methodologies produce infinite RRs, as the observed 2014 European average temperature anomalies were greater than any of the natural simulations. There is a large uncertainty in these calculations due to the rarity of the event in the natural climate.

Using the empirical method, the 2014 event has a return time of 30 yr in the current climate, with a large uncertainty: the 95% confidence interval starts at 2 yr. A century ago the lower bound would have been 4000 yr. For the weather@home ensembles, the return times in the actual2014 and actualClim simulations were 39 and 151 yr respectively (confidence intervals [32,50] and [117,203]). We were unable to reproduce any events with a European average temperature as warm as 2014 in the natural2014, naturalClim or CMIP5 historicalNat simulations, as expected from the return time estimated from the observations and the number of ensemble members. For reference, the previous record had a natural return time of 1960 yr becoming 3.7 yr in the current climate for the CMIP5 method.

We can also explain the RR from the anomaly of the observed temperature for 2014 relative to the temperature modeled for each method, the difference in temperature between the natural and actual climate (trend) for each method, and also the standard deviation of the data in each method (Table S2). The Empirical method has the largest natural anomaly as a result of the combination of a large actual anomaly and trend. The weather@home actualClim experiments have a larger anomaly than the actual2014 experiments, due to the simulations being from the cooler period of 2001-2012. However, the trend in the naturalClim ensemble (using the HadGEM2-ES model) is smaller than the average trend in the natural2014 ensemble, resulting in similar anomalies in the natural2014 and naturalClim ensembles. Lastly we note that the natural2014 experiments have a slightly greater standard deviation than the actual2014 simulations due to the use of 12 natural SST patterns. This will make the RR value more conservative than if the variability of the actual2014 and natural2014 simulations were the same.

3.2 Spatial distribution of the risk ratio

As impacts of most events are realised on local rather than continental scales, a more meaningful way of looking at the RR is to calculate it for small regions. It is expected that the local RR for temperature will be correlated to the regional RR, and is likely to be lower [*Angélil et al.*, 2014; *Christiansen*, 2015], however, the spatial pattern will depend on the specific event.

The anomaly between the observed temperature and the average temperature modelled for the present climate is shown in Figure 1 for each of the methods. This gives an idea of how extreme the temperature was for different areas in Europe compared to what is expected by the models. The temperatures for each of the methods show good agreement in the large scale features, showing that 2014 was especially warm in central Europe and Scandinavia but cooler in Portugal, Ukraine and the western part of Russia.

The spatial distribution of RR is shown in Figure 2, calculated for individual grid-boxes using each of the methods. While the ratio is consistently greater than one, its magnitude and pattern varies for each method.

Figure 2(a), using the empirical method, has RR around 10–100 over most of Europe, but greater than 1000 between Spain and France, and an area with RR less than 10 over Scandinavia. Figure 2(b), using the CMIP5 method, shows remarkably similar results in terms of the general magnitude of the RR values, albeit with differences for individual gridboxes. Like the empirical analysis, the higher RR values in the CMIP5 results are seen over southern and western Europe. Some coastal areas are undefined due to the coarser model resolutions (shown as white). Figure 2(c), calculated from the actual2014 and natural2014 ensembles, has a RR of around 10 for central Europe. There are a few areas with RR greater than 10 such as in Spain and Turkey, and RR greater than 100 over Scandinavia and Africa. Figure 2(d), calculated from the actualClim and naturalClim ensembles, is consistent with Figure 2(c), showing the attribution of this temperature event is not sensitive to the specific pattern of SSTs that occurred in 2014.

Each approach represents variability in the possible temperatures, shown by the standard deviation in Figure S1. The source of variability is different for each method. There is variability in the time series used in the empirical method, variability in the weather@home simulations exhibited across different ensemble members but damped near the coasts due to the prescribed SSTs, and variability in the CMIP5 ensemble, sampling uncertainty in the model response as well as the temporal variability. In places where the spread of values is relatively low, a small trend will change the likelihood of an event much more than in a place with higher variability (hence lower variability leads to higher RR). The trend in the temperature between the natural and current climate (Figure S2) combined with its variability (Figure S1), gives an indication of how much the likelihood of temperature events has changed, independent of any particular event. However, the RR appears to be less dependent on the magnitude of the trend. The RR is calculated for the specific conditions of 2014 (Figure 1) and shows lower values where the temperature anomaly is lower. RRs and trends for each of the SST patterns used in the natural2014 ensemble are shown in Figures S3 and S4 respectively. Which factor dominates may differ by method and by region. For example, the the high RR in Scandinavia for the weather@home model matches the pattern of higher anomaly, but there is a low RR in the same region in the empirical and CMIP5 methods, matching the pattern of higher variability.

Calculating the RR on the grid-box scale relies on the observations being known with high accuracy across the whole region. In addition the model may not perform as well in some regions compared to others, leading to unevenly distributed biases. There are also regions where the event occurs in the actual ensembles, but not in any natural simulations. This results in an infinite RR which can be interpreted as an event possible in the current climate but extremely unlikely in the past. However, there is an inherent uncertainty for events at the tail of the distribution due to the rarity of the event in the natural case. This statistical uncertainty may be reduced by spanning a wider range of possible climates, for example by conducting simulations with a wider range of possible SST forcings, using physics parameter perturbations or using different models with different physics parametrization schemes. However, it must be noted that increasing the ensemble spread (i.e. variability) will also systematically reduce the RR.

3.3 Regional dependence of the risk ratio

The previous section has shown that calculating the RR locally, then averaging over Europe, gives a very different value compared to calculating the RR from the average European temperature (compare values in Figure 2 and Table S3). It is well understood that variability changes as a function of domain size, so it is expected that the RR also exhibits some dependence on the domain size. To investigate how this occurs, this section looks at how the RR changes, calculated for average temperature over regions of varying sizes, centered about a specific location. Because of the lower resolution of the CMIP5 data, this analysis is only done for the empirical and weather@home methods.

Figure 3 shows how this behaviour depends on the size of the region and the rareness of the event in the ensemble. Figure 3(a) and (b) depict the probability of the event occurring at a specific grid-point and for different sized boxes of grid-points centered around this point, for each of the ensembles. The corresponding RR is shown in Fig. 3(d). Note that the probability of the event in the actual simulations is much more stable than the natural simulations, which decreases rapidly for larger regions. This results in the RR increasing with increasing region size. The empirical method is shown in Figure 3(c) and (e), where the same trend can be seen although with larger uncertainty.

The importance of the selection of region is clear. For example, using weather@home, Sweden experienced a very rare event where there were almost no natural ensemble members as warm as the observations across all region sizes, whereas Russia had a less rare event and the RR increased very smoothly with increasing region size. Some other regions have a very non-linear response as the region size is increased, likely due to expanding the region to include grid-boxes with a much lower RR. Hence, the choice of region for analysis is particularly important, to avoid including effects that are not relevant for a particular study.

Differences in the observed temperature thresholds used for the RR calculation may yield different assessments of the anthropogenic influence on in particular the very extreme events (see Figures S5 and S6). Here the magnitude of the probabilities of the event differ slightly while the RR is more consistent when comparing against different observational products.

When comparing the 2014 and the Clim ensembles, the magnitude of the probability of the event is different, but the RR is similar. For this particular type of event, specifying the SSTs to match the weather of 2014 compared to other recent years, does not change the result of the attribution assessment. For other parts of the world where the inter-annual variability strongly depends on the specific SST patterns this is likely to be very different (e.g.,*Otto et al.* [2013]).

4 Conclusion

Quantifying the extent to which anthropogenic climate change has contributed to record-breaking temperature extremes, such as the 2014 annual mean European temperatures, can be made by applying different methods of probabilistic event attribution [*Lewis and Karoly*, 2014; *Allen*, 2003; *van Oldenborgh*, 2007]. However, while it is relatively straightforward to obtain qualitative results for temperature extremes, the quantification and interpretation of these results is much harder. For an annual mean temperature record, the definition of the event is less ambiguous than for other extreme events where different event definitions may lead to apparently contradictory results. For example there have been very different assessments in detecting and attributing a human signal in the Russian heat wave in 2010 [*Otto et al.*, 2012] and the recent Californian drought [*Diffenbaugh et al.*, 2015; *Seager et al.*, 2014].

This study compares assessments for the same event using four methodologies involving three independent modelling approaches. The results for each of the methods are consistent in the range of the lower bound of risk ratios, although the spatial patterns vary. Comparing these methods highlights the importance of taking into account the exact framing of the attribution question that is being addressed [*Otto et al.*, 2015]. Whether the quantification of the anthropogenic signal is attempted using coupled or atmosphere-only model simulations or using observations alone will lead to different estimates of the RR. Because there is no single correct way of modelling human influence, these results should be treated more in terms of sampling estimates of the RR calculated from different scenarios and different modelling frameworks, rather than differing estimates of the same quantity.

Using the multi-model approach, we can say with high confidence that the annual mean European temperatures have been made at least 500 times more likely. This is however not true for individual regions. For example, for regions of around 1000km in scale in Figure 3 (also see Table S3), the lower bound of the RR was about 9 using the weather@home simulations and 5 using the empirical method. At the gridbox level, in the centre of the same regions, these lower bounds are reduced to 6 using weather@home and CMIP5 methods, and 4 using the empirical method.

This shows that for an event defined by mean temperature, the exact choice of region can change the quantification of the risk ratio by an order of magnitude. An influence of region averaging is not unexpected but has never been quantified before. To avoid this averaging effect, it may be preferable to calculate RR by pooling data over the region of interest rather than averaging. Furthermore, when conducting event attribution, the event definition should be carefully chosen on a case-by-case basis to match the impacts. It should be tested for sensitivity to choice of region, and defined consistently to avoid misleading results when comparing different attribution assessments.

Acknowledgments

We would like to thank all of the participants who computed simulations for climateprediction.net. We would like to thank our colleagues at the Oxford eResearch Centre: A. Bowery, M. Rashid and S. Sparrow for their technical expertise. We would like to thank the Met Office Hadley Centre PRECIS team for their technical and scientific support for the development and application of weather@home. Andrew King was supported by the ARC Centre of Excellence for Climate System Science (grant CE 110001028) and the NCI National Facility in Australia. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. EUCLEIA project funded by the European Unions Seventh Framework Programme [FP7/20072013] under grant agreement no. 607085. The data for this paper are available by request from the corresponding author. We acknowledge the E-OBS dataset from the EU-FP6 project ENSEMBLES (http://ensembles-eu.metoffice.com) and the data providers in the ECA&D project (http://www.ecad.eu).

References

Allen, M. (2003), Liability for climate change, *Nature*, *421*, 891–892, doi: 10.1038/421891a.

Angélil, O., D. A. Stone, M. Tadross, F. Tummon, M. Wehner, and R. Knutti (2014), Attribution of extreme weather to anthropogenic greenhouse gas emissions: Sensitivity to spatial and temporal scales, *Geophysical Research Letters*, *41*(6), 2150–2155, doi:10.1002/2014GL059234.

Bador, M., L. Terray, and J. Boé (2015), Detection of anthropogenic influence on the evolution of record-breaking temperatures over Europe, *Clim. Dynam.*, pp. 1–19, doi: 10.1007/s00382-015-2725-8.

Christiansen, B. (2015), The role of the selection problem and non-Gaussianity in attribution of single events to climate change, *Journal of Climate*, *28*(24), 9873–9891.

Diffenbaugh, N. S., D. L. Swain, and D. Touma (2015), Anthropogenic warming has increased drought risk in California, *P. Natl. Acad. Sci. USA*, *112*(13), 3931–3936, doi: 10.1073/pnas.1422385112.

Houghton, J., Y. Ding, D. Griggs, M. Noguer, P. van der Linden, X. Dai, K. Maskell, and C. Johnson (eds.) (2001), Climate Change 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change.

Hulme, M. (2014), Attributing weather extremes to climate change: A review, *Prog. Phys. Geogr.*, doi:10.1177/0309133314538644.

Jacob, D., J. Petersen, B. Eggert, A. Alias, O. Christensen, L. Bouwer, A. Braun, A. Colette, M. Déqué, G. Georgievski, E. Georgopoulou, A. Gobiet, L. Menut, G. Nikulin, A. Haensler, N. Hempelmann, C. Jones, K. Keuler, S. Kovats, N. Kröner, S. Kotlarski, A. Kriegsmann, E. Martin, E. van Meijgaard, C. Moseley, S. Pfeifer, S. Preuschmann, C. Radermacher, K. Radtke, D. Rechid, M. Rounsevell, P. Samuelsson, S. Somot, J.-F. Soussana, C. Teichmann, R. Valentini, R. Vautard, B. Weber, and P. Yiou (2014), Euro-cordex: new high-resolution climate change projections for european impact research, *Regional Environmental Change*, *14*(2), 563–578, doi:10.1007/s10113-013-0499-2.

James, R., F. Otto, H. Parker, E. Boyd, R. Cornforth, D. Mitchell, and M. Allen (2014), Characterizing loss and damage from climate change, *Nature Clim. Change*, *4*, 938–939, doi:10.1038/nclimate2411.

King, A. D., G. J. van Oldenborgh, D. J. Karoly, S. C. Lewis, and H. Cullen (2015), Attribution of the record high Central England temperature of 2014 to anthropogenic influences, *Environ. Res. Lett.*, *10*(5), 054,002, doi:10.1088/1748-9326/10/5/054002.

Lewis, S., and D. Karoly (2014), Are estimates of anthropogenic and natural influences on Australias extreme 2010–2012 rainfall model-dependent?, *Clim. Dynam.*, pp. 1–17, doi:10.1007/s00382-014-2283-5.

Lewis, S. C., and D. J. Karoly (2013), Anthropogenic contributions to Australia's record summer temperatures of 2013, *Geophys. Res. Lett.*, *40*(14), 3705–3709.

Massey, N., R. Jones, F. E. L. Otto, T. Aina, S. Wilson, J. M. Murphy, D. Hassell, Y. H. Yamazaki, and M. R. Allen (2014), weather@home–development and validation of a very large ensemble modelling system for probabilistic event attribution, *Q. J. Roy.*

Meteor. Soc., pp. n/a–n/a, doi:10.1002/qj.2455.

Otto, F., N. Massey, G. Oldenborgh, R. Jones, and M. Allen (2012), Reconciling two approaches to attribution of the 2010 Russian heat wave, *Geophys. Res. Lett.*, *39*(4).

Otto, F., E. Boyd, R. Jones, R. Cornforth, R. James, H. Parker, and M. Allen (2015), Attribution of extreme weather events in Africa: a preliminary exploration of the science and policy implications, *Climatic Change*, pp. 1–13, doi:10.1007/s10584-015-1432-0.

Otto, F. E., R. G. Jones, K. Halladay, and M. R. Allen (2013), Attribution of changes in precipitation patterns in African rainforests, *Philos. T. R. Soc. B*, *368*(1625), 20120,299.

Photiadou, C., G. van der Schrier, G. J. van Oldenborgh, G. Verver, A. K. Tank, M. Plieger, H. Mänchel, P. Bissoli, and S. Rössner (2015), 2014 warmest year on record in Europe — EURO4M Climate Indicator Bulletin, [Online; accessed 19-June-2015].

Rohde, R., R. Muller, R. Jacobsen, S. Perlmutter, A. Rosenfeld, J. Wurtele, J. Curry, C. Wickham, and S. Mosher (2013), Berkeley Earth Temperature Averaging Process, *Geoinfor. Geostat.: An Overview*, *1*(2), 1–13.

Roth, M., T. A. Buishand, G. Jongbloed, A. M. G. K. Tank, and J. H. van Zanten (2014), Projections of precipitation extremes based on a regional, non-stationary peaks-over-threshold approach: A case study for the Netherlands and north-western Germany, *Weather and Climate Extremes*, *4*, 1 – 10, doi: http://dx.doi.org/10.1016/j.wace.2014.01.001.

Schaller, N., F. E. L. Otto, G. J. van Oldenborgh, N. R. Massey, S. Sparrow, and M. R. Allen (2014), The heavy precipitation event of May–June 2013 in the upper Danube and Elbe basins [in "Explaining Extremes of 2013 from a Climate Perspective"], *Bull. Amer. Meteor. Soc.*, *95*(9), S69–S72.

Seager, R., M. Hoerling, S. Schubert, H. Wang, B. Lyon, A. Kumar, J. Nakamura, and N. Henderson (2014), Causes and Predictability of the 2011 to 2014 California Drought, *Tech. rep.*, NOAA Drought Task Force, Modeling, Analysis, Predictions and Projections (MAPP) Program of OAR/Climate Program Office.

Sippel, S., D. Mitchell, M. T. Black, A. J. Dittus, L. Harrington, N. Schaller, and F. E. Otto (2015), Combining large model ensembles with extreme value statistics to improve attribution statements of rare events, *Weather and Climate Extremes*, *9*, 25–35.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller (eds.) (2007), *Climate Change 2007: The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC*, Cambridge University

Press, Cambridge, United Kingdom and New York, NY, USA.

Stark, J., C. Donlon, M. Martin, and M. McCulloch (2007), OSTIA : An operational, high resolution, real time, global sea surface temperature analysis system, in *OCEANS 2007 - Europe*, pp. 1–4, doi:10.1109/OCEANSE.2007.4302251.

Stocker, T., D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley (eds.) (2013), 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, p. 1535.

Stott, P. A. (2003), Attribution of regional-scale temperature changes to anthropogenic and natural causes, *Geophys. Res. Lett.*, *30*(14), n/a–n/a, doi:10.1029/2003GL017324, 1728.

Taylor, K., R. L. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Amer. Meteor. Soc.*, *93*, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Thompson, A., and F. E. L. Otto (2015), Ethical and normative implications of weather event attribution for policy discussions concerning loss and damage, *Climatic Change*, *accepted*.

van Oldenborgh, G. J. (2007), How unusual was autumn 2006 in Europe?, *Clim. Past*, *3*, 659–668.

van Oldenborgh, G. J., A. van Urk, and M. R. Allen (2012), The absence of a role of climate change in the 2011 Thailand floods, *Bull. Amer. Met. Soc.*, *93*, 10471049, doi: 10.1175/BAMS-D-12-00021.1.

Fig1.png

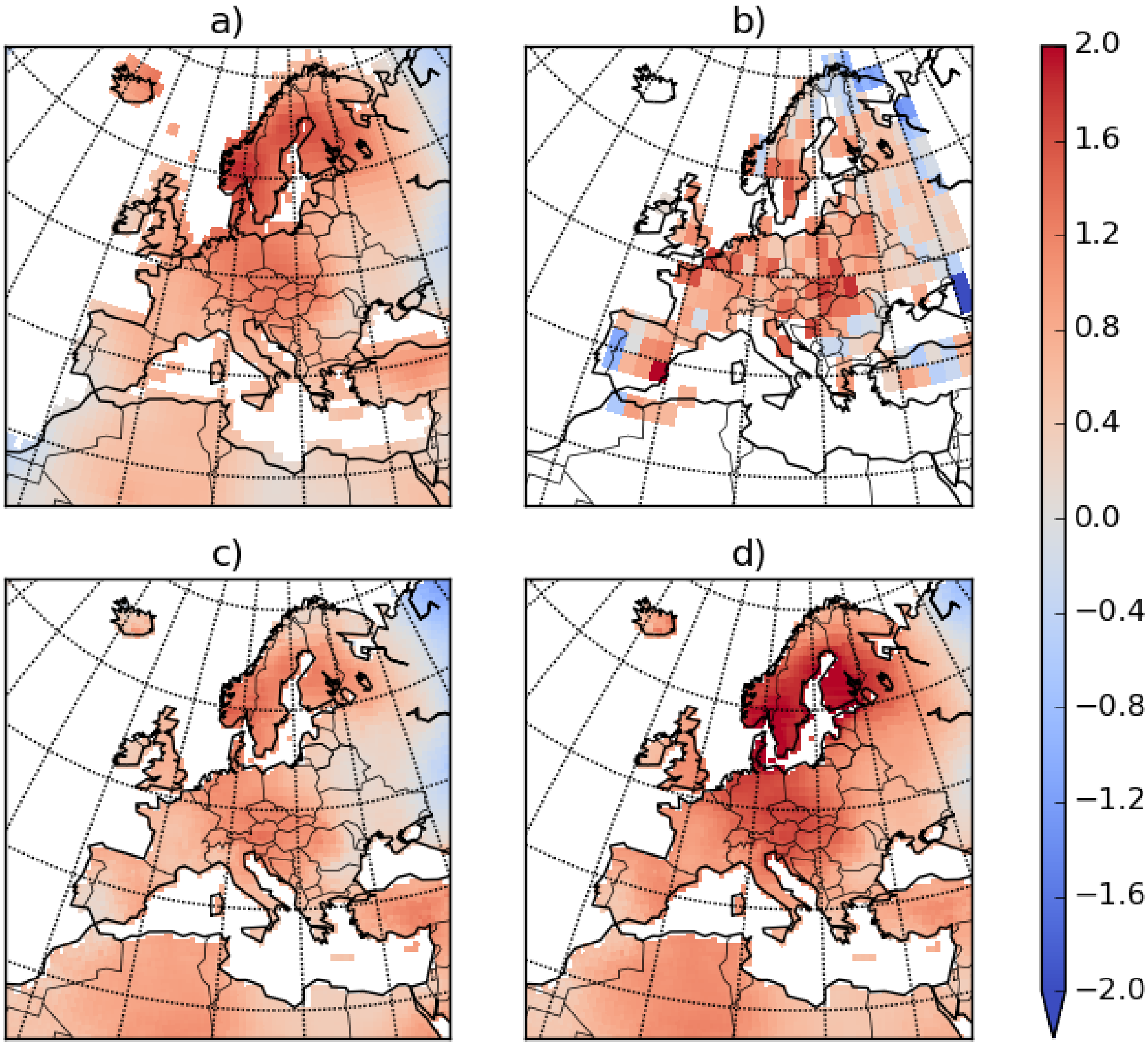
Figure 1. The spatial distribution of differences between the observed 2014 temperatures and the average temperature in the current climate modelled using different methods. a) historical data shifted to 2014 levels using the empirical method, b) CMIP5 simulations between 2006-2020, c) actual2014 simulations, and d) actualClim simulations. a), c) and d) were calculated relative to Berkeley data and b) was calculated relative to E-OBS

Fig2.png

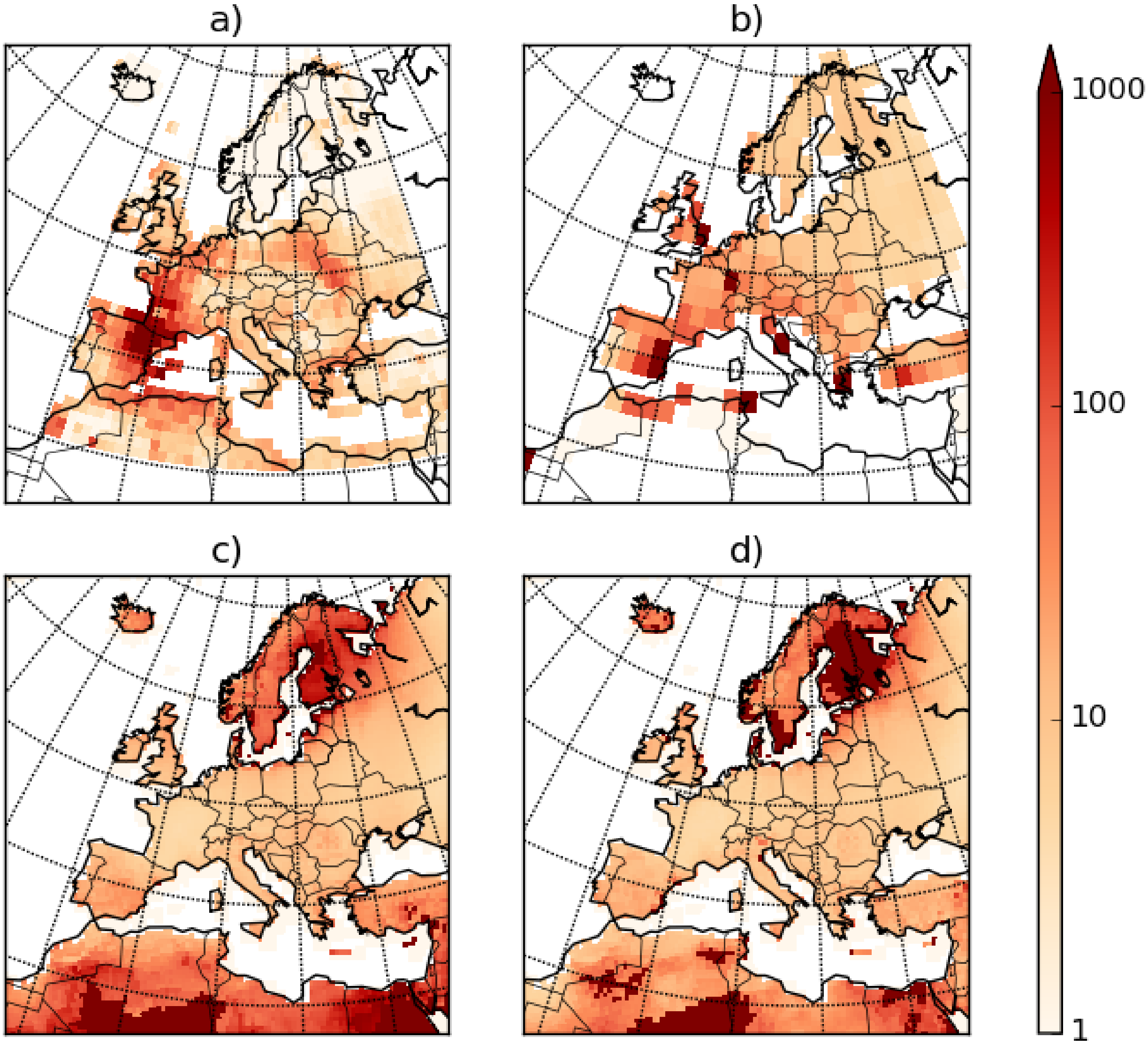
Figure 2. The spatial distribution of the lower bound of the 90% confidence interval of the RR, calculated for individual grid-boxes for different methods. a) is using the empirical method, b) using the CMIP5 method, c) is from comparing the actual2014 and natural2014 ensembles, and d) is from comparing the actualClim and naturalClim ensembles. a), c) and d) were calculated relative to Berkeley data and b) was calculated relative to E-OBS.

Fig3.png

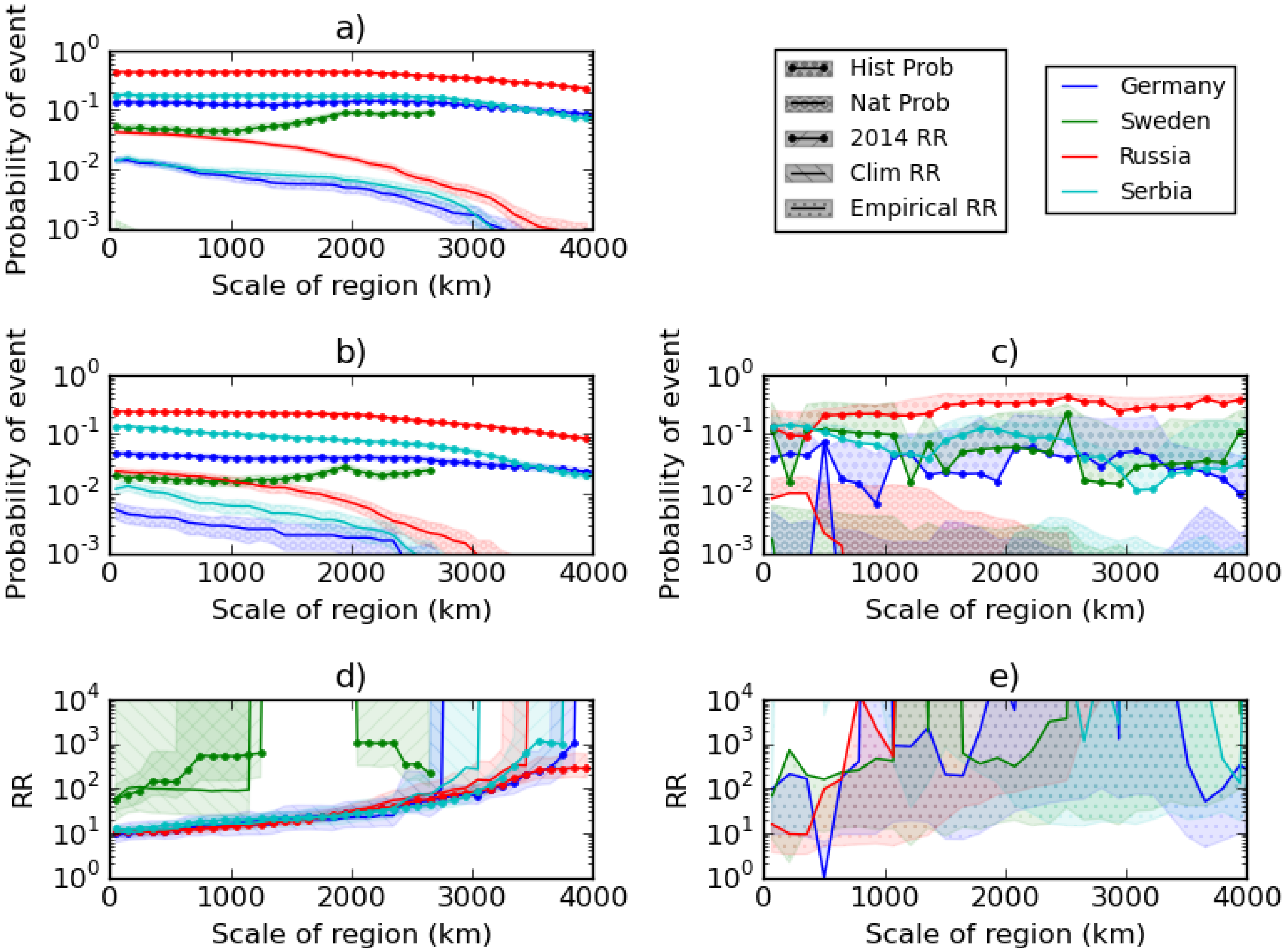
Figure 3. Plots showing the event probability and RR, for regions centred around different locations, expanding in size. Uncertainty here is defined using the 10%–90% confidence interval. a) Shows the probability of the event occurring in the actual2014 and natural2014 simulations, b) shows the probability of the event occurring in the actualClim and histClim simulations, c) shows the probability of the event occurring for the empirical data corresponding to 1901 (Nat) and 2014 (Hist), d) shows the RR calculated from the actual2014 and natural2014 (2014 RR) and actualClim and naturalClim (Clim RR), e) shows the RR calculated from the empirical data. Coordinates for the central points of the regions are given in table S4.



2016GL069568-f01-z-.png



2016GL069568-f02-z-.png



2016GL069568-f03-z-.png