Hydrol. Earth Syst. Sci. Discuss., 12, C7146–C7158, 2016
www.hydrol-earth-syst-sci-discuss.net/12/C7146/2016/
© Author(s) 2016. This work is distributed under
the Creative Commons Attribute 3.0 License.

Hydrology and Earth System Sciences
Open Access
Discussions

Interactive comment on "Climate change increases the probability of heavy rains like those of storm Desmond in the UK – an event attribution study in near-real time" *by* van Oldenborgh et al.

van Oldenborgh et al.

oldenborgh@knmi.nl

Received and published: 1 July 2016

Reply to Interactive comment #7.

T*he submitted paper investigates to what extent climate change has altered the odds of a rainfall extreme like the one observed during storm Desmond in the UK. The authors use three alternative approaches to assess the likelihood change of the observed event due to external factors. The authors submitted the manuscript within a remarkably short period of only 4 days after the event occurred to demonstrate that real-time attribution statements are possible in the time period during which media interest in the event*

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

is still high. Thereby they imply that such scientific near-real-time assessments are possible and meet scientific standards even of peer-reviewed publications.

Ironically my interpretation after having read the manuscript is quite the opposite. The authors made an impressive effort in doing a comprehensive analysis and submitting a manuscript in such a short period. I applaud for comparing three different methodological approaches. However, frankly speaking, I do not see much robust added value beyond a simple generic Clausius-Clapeyron argument that any climate scientist could have given to the media while the event was unfolding. The whole method intercomparison framework seems to imply a rigorous scientific assessment of an accuracy that goes far beyond such a general thermodynamic argument. But does it really do so? I am not convinced, maybe because, at most, the manuscript meets the standards of a blog article rather than the ones of a scientific paper. I recommend revisions and I am convinced that the manuscript has the potential to make a valuable contribution to the literature. But more work is needed and more generally I consider the exercise for a rigorous real-time event attribution (within days after the event) that adds substantial value beyond simple arguments on existing literature to have failed. I detail my comments below.

Indeed we realised afterwards that the paper was written too quickly and have adjusted our procedures to take more time in future studies. We took the criticisms of this paper into account when writing the next study we did using these methods, doi:10.5194/hess-2016-308. Specifically, we took more time and used more standard building blocks and check-lists to make sure no elements were left out of the presentation.

We disagree with the overall assessment of the scientific value of this study and give detailed responses to the criticisms below.

Definition of event

Despite a whole section on the event definition, it remains unclear what definition is finally used in the attribution statements. The event thresholds used seem to even differ between the three methodologies. The precipitation totals mentioned in the text dramatically vary between some 341 mm in 24h to less than 30mm/d used in Fig. 6. I understand the scale difference between in-situ measurements and gridded observations and models, nevertheless, I think a common event definition informed by observations should be used in all methods.

The presentation was indeed unclear. We added text that should make the event definitions clearer. At the end of the first paragraph we added 'To summarise: we use one-day precipitation for 0-24 datasets and when computing only statistics from 9-9 datasets, and 2-day precipitation to compare the event itself to its historical record when 0-24 data is not available. The 18:30-18:30 24-hr record is not used in his analysis.'

The remaining differences are a matter of scale and dictated by the lack of public and real-time observations in the region. This forced us to use three scales: the point observation at Eskdalemuir, the regional averages in the UK Met Office regions of NW England and S Scotland and finally the large box that we used for the main results. Observational estimates for all but the single station were not available at the time, so we relied on the ECMWF analysis (short-term forecast) instead. At revision time we added the observations for Northwest England and South Scotland en the ERA-interim estimate for the large box, clearly marked as added later.

We added the underlying assumption to the text: 'Precipitation averaged over smaller areas such as the basins of the rivers that flooded, and indeed point data at rain gauges, are assumed to have similar changes in the probability of extreme precipitation due to global warming. The extremes themselves do vary with spatial scale, but the ratios of extremes at different scales are assumed to be constant in time. For large-scale winter precipitation events such as storm Desmond we know no evidence that would contradict this assumption.'

HESSD

12, C7146–C7158, 2016

CC BY

It is unclear why one should have confidence in the precipitation totals of the ECMWF 24-hour forecast for such an exceptional event.

For the Netherlands, an area of similar dimensions of O(200 km), 0-24 precipitation is publicly available from the 32 automatic weather stations. In winter, the average of these stations has a correlation of 0.95 with the ERA-interim 24-hr forecast used in the reanalysis, with an underestimation of 10% that is not that much larger than the differences between differing types of rain gauges. A scatter plot (attached) does not show larger deviations for extreme events. For a flat area these types of events are simulated very well by the ERA-interim and hence EC-Earth model. The effects of orography have been discussed above: this does not affect the central value of the risk ratio but only the uncertainty range.

Even though the change in likelihood may not be particularly sensitive, the event definition should not be arbitrary. If in this exercise the availability of observations only days after the event was the limiting factor, the whole real-time attribution effort should maybe first focus on improving the immediate availability of observations rather than investing all resources and efforts in developing complex modelling frameworks.

We agree completely, and in the European EUCLEIA project have devoted and are devoting considerable energy in compiling datasets that can be used within days of an extreme weather event.

For this event, the manuscript needs to be revised using a proper observational estimate of the event once it has become available and using the same event definition throughout all methods.

For all analyses we have added the observational estimates as they were available at revision time, clearly marked as such.

For future real-time event attribution exercises I recommend waiting for proper observational data. I simply do not see the point of implying accurate scientific event attribution *before even the observational estimates have become available. The argument that for the attribution statement it does not matter how extreme the event was, does not foster public confidence in the scientific rigour of such real-time attribution exercises.*

Indeed we have refrained from doing a few studies for lack of observational data, e.g., the severe rainfall that caused severe flooding in Malawi in early 2015. We estimated that given the nature of this event, a winter secondary low that weather models can represent well, we could trust the (re)analyses enough to compensate for the lack of observations in the area. The observations available at revision time did not change the conclusions, validating the choices we made in December a posteriori.

Looking forward, we have incorporated more real-time datasets into our system and more and more countries in Europe are following the US lead and switching to an open data policy that greatly enhances our possibilities. This allows us to do more rapid attributions in areas with enough observations.

Observational analysis and internal variability

The observational analysis is interesting and valuable but both the methodology and the findings require further documentation. You analyse southern Scotland and north-western England separately and find different observed trends in heavy precipitation. Are you suggesting that the difference relate to systematic differences in the forced precipitation response or due to internal variability?

We state in the article that the difference is entirely compatible with internal variability. This was deduced by considering whether the uncertainty of difference in trends includes zero, which it easily does. The correlation of seasonal maxima in these two regions is very low, $r = 0.16$. Case in point: the most extreme event is 15 December 2015 in South Scotland, with 215 mm/day averaged over this large region. On the same day Northwest England only had 0.7 mm/day.

I suspect it is primarily the latter and internal variability could have easily masked the heavy precipitation response even at these relatively long time scales. This should be tested based on the EC-EARTH initial condition ensemble. By how much do the trends at these time scales differ between different realizations? I suggest to use a large enough area or to pool enough stations in order to increase the signal-to-noise ratio and reduce the effects of internal variability.

Of course we checked at the time of writing that in the model sets that we had available the forced response does not change much from northern England to southern Scotland. A first quick look at the CMIP5 ensemble using the KNMI Climate Change Atlas (climexp.knmi.nl/atlas/) shows no gradients in this area (figure attached). This ensemble underestimates spatial variability due to the fact that it includes many members with low resolution, though.

As this case involved heavy precipitation both in Northwest England and South Scotland we considered the mean of the risk ratios of the two subregions a reasonable estimate of the risk ratio of the affected region. It is not clear how a forced response difference between the two sub-regions would affect the conclusions. If the reviewer deems the difference important we can do the calculation he suggests.

Attribution based on individual station series or single gridpoints that are heavily influenced by internal variability does not make sense. The choice the area size required to make robust statements may be informed by the EC-EARTH ensemble. However, simply averaging the likelihood ratio for South Scotland and Northwest England as done here seems arbitrary.

It makes sense in two ways: the differences are compatible with internal variability and the area of heavy 1-day precipitation on 5 December 2015 encompassed both regions.

Clausius Clapeyron

You argue that changes in precipitation extremes go beyond a simple Clausius-Clapeyron scaling and changes in atmospheric circulation may play an important role. You should test this and quantify how much of the changes in the likelihood ratio can be explained by a simple CC scaling of the precipitation extremes with the regional or large-scale warming. Do the large ensembles of weather@home and EC-EARTH really show a robust forced response in atmospheric circulation? Such large ensembles would be the ideal testbeds to address this question. As stated in the introductory paragraph I suspect that much of the changes found here are accounted for by a simple CC scaling.

In this case we do indeed find that the results are compatible with simple Clausius-Clapeyron scaling and are not influenced noticeably by forced atmospheric circulation shifts. We have added a sentence to this effect. However, we have seen too many cases in which CC-scaling did not hold to dispense with the observational and model-based analyses. Counterexamples are the rainfall and flooding in southern England in January 2014 where the change in circulation extremes increased the risk ratio by about 50% (Schaller et al, NCC, 2016), the rainfall in Central Europe that caused flooding on the Danube and Elbe in May–June 2013 that showed no trend either in observations nor in one model (Schaller et al, BAMS, 2014) and the Boulder floods of 2013 where the increase appears to be considerably lower than CC scaling (Hoerling et al, 2014; Eden et al, in preparation). See also the commentary by Otto et al that will appear soon in Nature Climate Change.

At this stage our methods to disentangle the dynamics and thermodynamics are not yet suitable for inclusion in rapid attribution. We know how to determine trends in the average circulation, but this is not necessarily the same as the trend in circulation extremes that give rise to precipitation extremes. We hope that with further research and development we will be able to add this component to rapid attribution studies in a few year's time.

HESSD

12, C7146–C7158, 2016

cc BY

Model evaluation

The authors mention but do not perform any model evaluation for this region. I understand that such an evaluation is not straightforward and bias adjustment would be particularly challenging. However, if you expect people to have confidence in these results beyond simple thermodynamic arguments some evaluation for the regional precipitation climatology of the two models is needed.

We have added some remarks on model evaluation in the revised article and have added to our templates and checklist to make sure it is included in all future studies. The EC-Earth climatology of RX1day in the larger region agrees within errors with the ERA-interim climatology ($4\% \pm 6\%$ drier), which is maybe not too surprising as both are based on similar models. ERA-interim agrees well to observations in winter extremes over the Netherlands.

For the Weather@Home model we do not use a bias correction but the return time from the observational analysis. We refer to Schaller et al, 2016 for a thorough evaluation of winter precipitation in Britain in this model.

Confidence interval

The strongest point of the paper is the comparison of three methodologies. However, it is unclear how you get from the three confidence intervals of 1.3–2.8, 1.1–1.8 and 1.05–1.8 to the 5–80% or even the best estimate of 40% highlighted in the abstract. It seems that you only used the weather@home results. Why would you have more confidence in this than the other approaches?

This should indeed have been included in the article (note that the Weather@ Home range is 1.05–1.4, this has been corrected). We have added the justification for the assessed range: 'Overall, we find a roughly 40 % increase in likelihood, with a 95 % uncertainty range of 5 to 80 % for a return time of 100 years, taken as the lower value of the lowest model result and the upper boundary of the highest model result. Given the large uncertainties of the observations we did not consider it possible to use these to reduce the range at the low end, and the high end of the observational range depends too strongly on one extreme event in 2005 in South Scotland to extend the assessed range upwards.'

Minor comments

13199: L24 *you do not introduce what f1 and f2 stands for. Is it equivalent to p1 and p2?*

Yes, fixed.

13199: L27-29 *"In the limit that the trend is completely due to anthropogenic forcings these coincide. In the UK in winter this is as far as we know a reasonable approximation." That's an odd statement. At least in the observations the role of internal variability will be substantial.* In this paragraph we meant to only discuss the forced response.

Made this more clear: 'If the trend is completely due to anthropogenic forcings the changes in return time will be the same for all methods. . . . The largest uncertainties arise from the random weather, which affects all three methods equally.'

Later, we repeated this: 'The increase in probability of these kinds of events in EC-Earth is in line with the observational one, although we expect a difference due to the different framing of the attribution question within the different methodologies. The observational analysis considers the change due to the observed trend, independent of the cause of this trend, while the coupled model shows the change due to the external anthropogenic and natural forcing prescribed in the model. The differences are mainly in the response to the aerosol and greenhouse gas forcings of the climate model used, which may differ somewhat from

HESSD

12, C7146–C7158, 2016

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

the real world. Very low frequency natural variability could also cause the results to diverge as this is averaged out more in the model ensemble than the single reality of the observations.'

13200: L22 *rephrase sentence*

'There are two relatively long area-averaged time daily series available from the Met Office in areas with severe precipitation: Northwest England and Southern Scotland precipitation (for data and maps see http://hadobs.metoffice.gov.uk/hadukp/). We analysed these to investigate trends in daily and two-daily sums of precipitation. '

13201: L24-25 *As stated above I agree that spatial pooling or aggregation is essential here but the way the averaging of two regions is motivated and done here is confusing.*

We hope it is clearer now.

13203: L3 *4-5% is about what CC scaling would give you, right?*

Yes, given the upstream North Atlantic warming of roughly half a degree. We added a sentence to this effect to the discussion.

13204: L13-15 *Comparing only two years is a very poor test of the influence of SST variability. Even if I assume that the conclusion still holds for other years it should be tested with SST variability from about a decade or so.*

We now have the full decade of models results to be able to do this, see the next paper.

13206: L12-13 *I understand what you mean but it is confusing to argue that the anthropogenic forcing matters but the SST patterns does not if in weather@home you ultimately reflect the anthropogenic influence by an SST masking. I think what*

HESSD
12, C7146–C7158, 2016

you mean is that intra-decadal SST variability does not matter but multi-decadal SST changes do matter.

Agreed, we have reformulated he sentence to ‘This corroborates the assumption that this increase is indeed mainly due to anthropogenic climate forcings made in the observational analysis, and that the influence of other factors such as SST differences resulting from natural variability is small.’

References *It gives a bit an odd impression if more than two third of the references are from the authors or their teams. Particularly if the paper is about a subject like heavy precipitation where there is a vast body of literature.*

Agreed. We are working on a comprehensive bibliography file to be used in these rapid attribution studies, but this was not yet available at the time of this study. The references included now are necessary to document what we have done. The decision to take more time for the scientific paper in future studies should alleviate this problem as well. In the next paper the proportion was lowered to one third.

Interactive comment on Hydrol. Earth Syst. Sci. Discuss., 12, 13197, 2015.

HESSD

12, C7146–C7158, 2016

Interactive Comment

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Discussion Paper

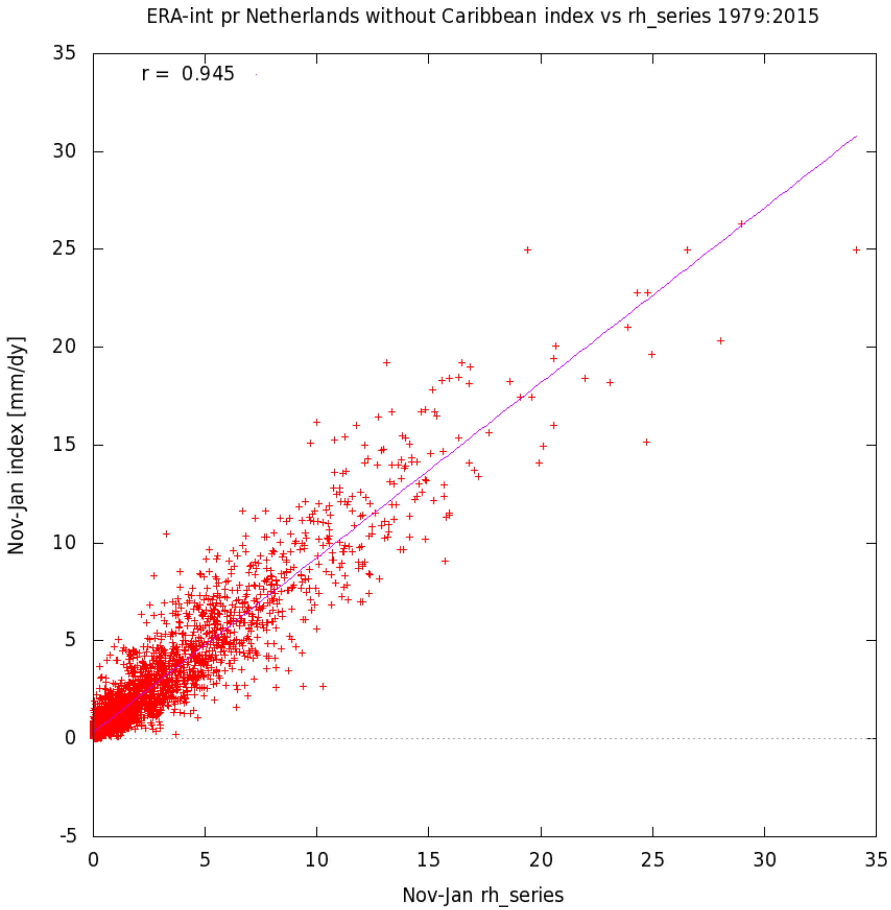

Fig. 1. Scatterplot of ERA-interim precipitation averaged over the Netherlands against the average of 32 0-24 stations

CC BY

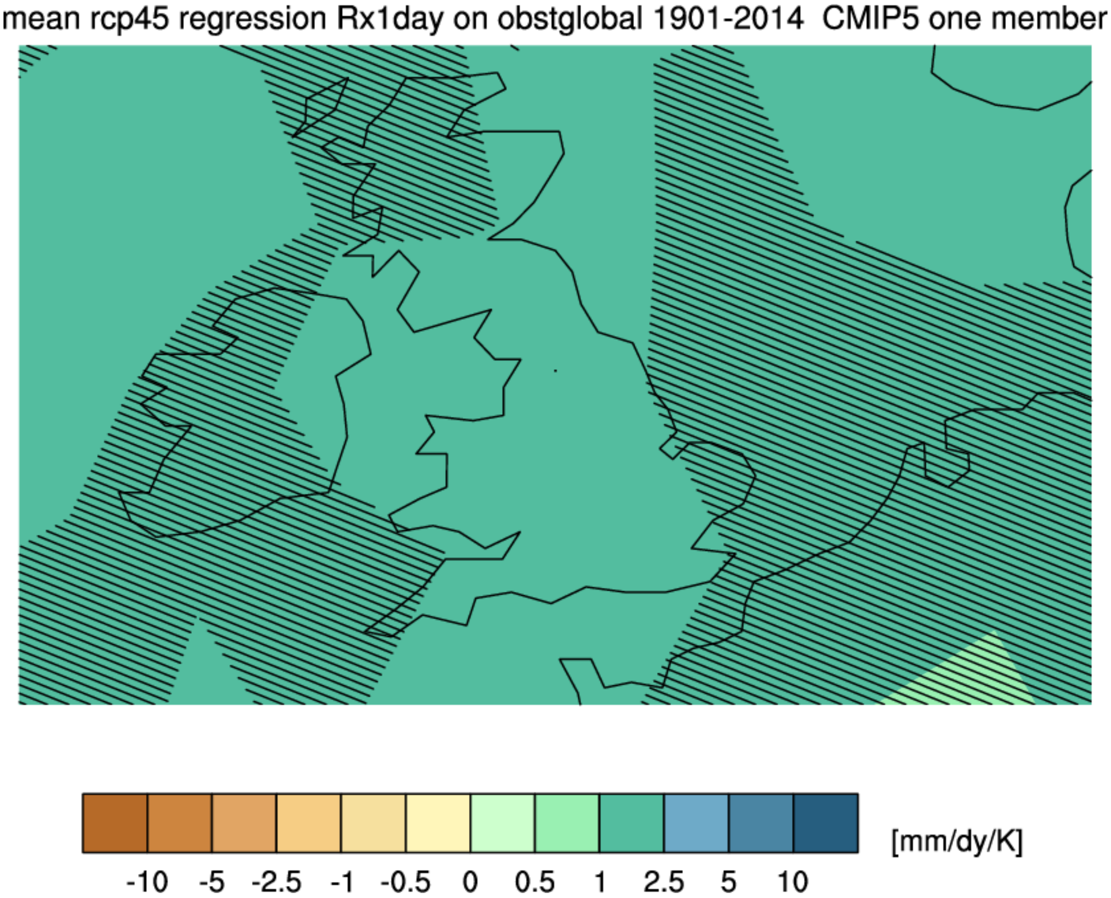

Fig. 2. Mean of the CMIP5 historical/RCP4.5 trends in wettest day of the year 1901-2014

HESSD
12, C7146–C7158, 2016