Hydrol. Earth Syst. Sci. Discuss., 12, C7136–C7145, 2016
www.hydrol-earth-syst-sci-discuss.net/12/C7136/2016/

Hydrology and Earth System Sciences 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 #6.

General comments

This study looks at the likelihood of one-day and two-day precipitation events over Northern England and Southern Scotland for current day conditions and the influence that anthropogenic climate change had on these likelihoods. The methodology consists of analysing historical observed trends, coupled climate model simulations and

ensembles of regional climate simulations. The authors conclude that "the effect of climate change is positive, making precipitation events like this 40% more likely" with an uncertainty estimate ranging from 5-80% likelihood.

I see several difficulties with this study listed below. The paper is not well written (perhaps a consequence of the tight timing) in that it is confusing in several aspects, see points below, and lacks the scientific rigour expected from a contribution to HESS. As a result, I cannot recommend the manuscript for publication.

We agree that the paper was badly written and will improve his in the future attribution studies, by giving ourselves more time and having more building blocks and references available before the event. (This process will in the end lead to non-scientific operational reports for simple events, just like seasonal forecasts do not merit scientific papers.)

For this article we attempt to address the concerns below.

Major comments

1. *The manuscript aims to analyse the probability of heavy precipitation events in NW England and S Scotland similar to storm "Desmond" of 4-6 Dec 2015. As two of the three methodologies use dynamical circulation models the immediate question arises of how well these models are able to correctly simulate the event in the first place. It is well known that precipitation, and in particular heavy precipitation events are difficult to simulate for models on a range of horizontal resolutions. A major problem here is the substantial underestimation of the rainfall intensities compared to observations, even with numerical weather forecast models. The ECMWF 24-hour analysis which is used in this study also greatly underestimated the rainfall intensities for "Desmond". As discussed in the paper, this is to be expected from models that cannot resolve the local orography suffi-*

ciently. Given these inherent problems with rainfall magnitudes in models, what can be reasonably expected to be said with any confidence about extreme events that fall outside the range of the model worlds? Are these models adequate tools to quantify (relatively small) changes of observed extremes that inherently come with large uncertainties? This is a difficult question to answer positively, and unless any evidence (e.g. synoptic studies) is presented to convince ourselves that the models indeed are able to simulate structures that are reminiscent of storm "Desmond", we have to assume that the answer is No.

In this case, we argue that the answer is yes. We added a section on model evaluation to the manuscript that should have been there in the original and plan to add this to all future attribution papers.

It does not make much sense to compare local rain gauge observations to model output in hilly terrain, as all but the highest-resolution models will indeed fail to capture the local extremes (modern non-hydrostatic model are a notable exception but large enough climate run ensembles are not yet available for attribution). However, that is also not the relevant quantity. In this case, the main effect of the rainfall was flooding, which is caused by area-averaged rainfall. Models represent large-area averages much better than local rainfall. We chose to study an area with the size of the rain zone of the secondary low of Desmond (the depression itself was located near Iceland), which is O(300 km). In this case, the models we use have been derived from weather forecast models that are quite good at simulating the large-scale features of the frontal disturbance that brought the rain to northern England and southern Scotland.

Unfortunately it is very hard to compare the models against observations directly as we only have available 0-24 UTC output for the models and 9-9 UTC output for the stations. In the Netherlands, the 32 automatic weather stations also give 0-24 UTC output. The correlation between the ERA-interim output (which uses a very similar model as EC-Earth) averaged over the Netherlands O(200 km) and the

observations averaged over the 32 stations is $r = 0.95$ with an underestimation of 10%, not much larger than the differences between our different gauges. The discrepancy in magnitude will be larger in the study area due to the orography in England and Scotland. (The Netherlands are quite flat.)

We do not need the model to faithfully reproduce the magnitude of the precipitation, as long as the underestimation is the same for this event and the statistics of the past climate the result are valid. We have no evidence that for these winter storms the ratio of the model precipitation to observations is changing with time.

Finally, the Weather@Home models show that the Risk Ratio p_1/p_0 does not depend strongly on the return time, so even in this case where we could not determine the return time well the risk ratio is in fact well-defined.

We have added the evaluation of the ECMWF model over the Netherlands and mention that the synoptic structures that caused the rainfall are adequately resolved and simulated by the models that we employ.

2. *The definition of the extreme event for this study is confusing and perhaps misleading. In most parts of the paper one-day rainfall amounts are used; in some other parts it is argued that two-day rainfall amounts should be used (selection bias?). While the text mentions observed one-day rainfall amounts of 341mm at one station in Northern England and of 77mm in Southern Scotland, the event definition seems to be based on the area averaged ECMWF analysis (which is a 24-hour forecast) of 36.4mm. There is confusion here what analysis was used to derive 36.4mm – was it the operational analysis (as suggested by Fig 1 and text in Section 2) or the ERA-Interim analysis (as suggested by the text in Section 4)? The return time definition of the event is similarly confusing. A 1-in-100 year event is assumed but I cannot see much observational evidence for this assumption. From Fig 3c) and d) I'd rather think the observed return times are less than 5 years. What is the return time of the 36.4mm in the ECMWF analysis?*

The confusion derives from two sources:

- The observing stations are 9-9 stations, whereas the ECMWF analysis (and reanalysis) fields that we had were 0-24 averages. Strictly speaking these are not analyses but short-term forecast, as the 0:00 starting point is the analysis but the rainfall comes from a short forecast run. As these storms can pass through any time of day (unlike the diurnal cycle of eg thunderstorms) this does not matter for the statistics, but it makes a big difference for the description of the amount of rain associated with Storm Desmond. As we say in the paper, most rain fell in the 24 hours from midnight to midnight (this was deduced from hourly amateur stations in the region on wow.knmi.nl). The 0-24 (re)analyses (forecasts) therefore describe the event well and we took 1-day extremes. However, for the station Eskdalemuir the 2-day sum was used as half the rain fell before 9AM and the other half afterwards. The Northwest England and South Scotland series are also based on 9-9 stations, but as we estimate the observed amount from the 0-24 ECMWF forecast we again consider 1-day sums.
- The second confusion seems to stem from the spatial scales. The record precipitation is only mentioned to justify the analysis, but never used. The Eskdalemuir 2-day sum of 139 mm/day is used in Figure 4. We would have preferred to analyse more stations, but these were not available at the time of writing at the Met Office as it takes a while to collect the data from various agencies and validate them. We therefore resorted to rough estimates from the ECMWF analysis (forecast) over these regions (28 and 31 mm/day) and the large box, 36.4 mm/day. The latter value was considered too high, based on the higher resolution of the forecast model compared to ERA-interim and EC-Earth, so we did not use the return time that resulted from this value but looked at the single station to estimate a return time of O(100 yr). The ERA-interim area averaged over the box indeed turned out to be lower when

it became available three months later, 25.1 mm/day giving e return time of O(20 yr) in ERA-interim and EC-Earth. The risk ratio does not depend strongly on the return time, as can be seen from Figure 6 and from the difference between the risk ration for 100-yr return time and for a 20-yr return time that was added at revision time.

We have attempted to clarify the first point better in the text: '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.' Concerning the spatial scales, we added '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 of no evidence that would contradict this assumption.' and updated rainfall observations and return times with data that was available at revision time, clearly marked as such.

3. *I am left confused with the issue of mode biases and how the presented analysis takes them into account. The Introduction talks about the need for careful bias correction while mentioning at the same time that this was not available at the time of writing. What does this imply for the presented quantitative analysis? Section 5 mentions the dry model bias again but I cannot see how this problem has been solved or addressed adequately.*

   The main point is that the risk ratios do not depend strongly on return times. In fact, the whole GEV fitting analysis is based on the assumption that they are constant, the Weather@Home ensemble is large enough that we can verify that this assumption holds in this case.

Determining return times proved to be very difficult at the time of writing due to biases in various estimates of what happened in the real world: too few stations (one), non-availability in real time of the area averaged observations, biased reanalysis / analysis (forecast) estimates. We checked that ERA-interim and EC-Earth have similar biases: the difference in the wettest day of the October–February season is indeed zero within error margins ($4\% \pm 6\%$). However, as we indicated in the paper these have a dry bias due to the lacking orography at this low resolution, so we estimate that the ECWMF analysis value would be high compared to the ERA-interim climatology and the EC-Earth model climatology. This turned out to be the case, even more so than we thought, so the return time of precipitation in the large box has now been determined to be about 20 yr. The downgrading from 1000 yr to 100 yr in Figure 5 was therefore in the correct direction, but not strong enough.

For the Weather@Home model we did not do a bias correction, but based the risk ratio on the return time from the other analyses. For 20 yr it is almost the same as for the 100 yr that we used at the time, 1.01 to 1.35. This has all been added to the text, again clearly marked as added at revision time.

4. *The authors argue in the Introduction that internal low-frequency variability plays a minor role. What is the basis for this statement? The discussion of Section 4 saying that "very low frequency natural variability could also cause the results to diverge" seem to suggest differently.*

   We have added the underlying computations to the text: 'Low-frequency natural variations also play a minor role here: precipitation extremes are not significantly correlated to the Atlantic Multidecadal Oscillation (AMO) or Pacific Decadal Oscillation (PDO) at $p<0.1$ over 80 years of observations.' In the conclusions we added the caveat in case somebody finds another very low frequency mode that affects rainfall extremes in this area.

5. *It is found that the regions of NW England and S Scotland show a very different behaviour in terms of precipitation trends even though they are geographically very close and one would expect similar large-scale dynamic influences. How does the discrepancy between these two neighbouring regions impact the findings of this study? Observations of NW England show no trend whereas those in S Scotland indicate a positive trend. I don't understand the sentence in Line 24, p13201: "the trends in the two regions are compatible with each other. . . ". I also don't understand the comments on the natural variability in this context.*

   The trends are not very different from each other as long as we take the uncertainties into account. These are pretty large, and both trends are compatible within these uncertainties with the average for the two trends. The uncertainties are so large because the natural variability in extreme precipitation is large. The secondary lows causing these extremes have sizes that are about the same as these areas, so sometimes South Scotland happens to get a very strong winter precipitation extreme while Northwest England does not, as on 15 December 2005. Other times it is the other way around. The uncertainty around the trend estimates is due to the random nature of these extremes. The hypothesis that the underlying trend is equal to the average of the South Scotland and Northwest England trends with the accompanying uncertainty fits all the data. In Scotland it is larger, partly due to the downpours in 2005. In Northwest England it is lower because such strong extremes happened to be absent before 2015.

6. *Why is the framing of the attribution question different in the coupled climate model EC-Earth and the observations? As far as I understood a very similar methodology based on trends and GEVs has been applied to the coupled model data, no matter what forcings were used in the model runs.*

   In the usual framing of the framing question they are indeed the same. In practice they can differ due to shortcomings in the modelled response to greenhouse gases (transient climate sensitivities among CMIP5 models differ by more than a factor two) and aerosols (even more uncertain). The second difference is that we use large numbers of ensemble members for the models, averaging out natural variability, whereas the single Earth that the observations come from can have more low-frequency natural variability.

7. *The abstract mentions a change in return times of 1.4 with a confidence interval of 1.05-1.8. How is this derived when the three individual methods give values of 1.3-2.8 for the observations, 1.1-1.8 for EC-Earth and 1.05-1.8 for the regional model?*

   This should have been documented. We took the lower bound of the lowest model result and the upper bound of the highest model result to take the model spread into account. 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 points

- *Fig 2 is not very informative – what is the motivation for showing it? Why are there two seasonal cycles shown in the plot?* We needed it to justify the season over which we consider maximum precipitation. We assumed not many readers would know that precipitation extremes in this area occur in autumn and early winter. The double seasonal cycle is a standard feature of the KNMI Climate Explorer that was used to make this plot.
- *The subplots of Fig 3 are very small (too small)* These sill be larger in the published version.
- *Fig 4 is not discussed in the text.* We have added a reference to the figure in the paragraph that discussed this figure.

- *Several places in the text where reference to wrong figures are made.* We have fixed these references and made sure that the procedure for writing articles during a rapid attribution makes it less likely that these kind of errors are made in the future.
- *Why have only the two recent winters instead of a more representative climatology been included in the analysis of the regional model?* At the time of writing, large ensembles were only available for these two winters.
- *The discussion of floods in the last paragraph of Section 6 is misleading as this study is about rainfall.* We wanted to put this one aspect of the flood into perspective and explicitly call out the limitations of the study. The roles of exposure and vulnerability are all too often neglected compared to the very modest role that climate change has in changing the odds of floods. We judged it better to say this explicitly than to restrict ourselves completely to precipitation, giving the impression that climate change is the dominant problem.

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