Review - Yoann Robin 2025

24 November 2020 12:08

This paper extends previously published methods to estimate the changing likelihood from the past to the future of a specified extremes metric due to historical forcings and future emission scenarios. In addition the likelihoods from a counterfactual world, one without anthropogenic emissions is estimated. The new extension to the methodology is enabling a single counterfactual world to be estimated from multiple different scenario runs. The authors also revise the Bayesian approach to reflect advances in this field. For the analysis the authors have developed a software package that is publicly available to facilitate other researchers wishing to perform similar tasks.

My request for major revision is not due to significant concerns of the analysis they have performed but because of the lack of placing their work in the context of the field more generally and the belief that in its present form the text is only accessible to a very small audience. I found what the authors have exactly done and why quite difficult to ascertain and feel that I have only got the general gist rather than a good understanding.

1) Context of other work

There is now quite a body of work addressing the production of posteriors of climate variables from the combination of observations with climate models more generally and extreme specific. It would be helpful if they placed their work in the context of these other pieces of work.

General Bayesian climate projection

Making climate projections conditional on historical observations.
Ribes, A., Qasmi, S., & Gillett, N.P. (2021). Science Advances, 7(4), DOI: 10.1126/sciadv.abc067

Energy Budget Constraints on the Time History of Aerosol Forcing and Climate Sensitivity
C.J. Smith, G.R. Harris, M.D. Palmer, N. Bellouin, W. Collins, G. Myhre, M. Schulz, J.-C. Golaz, M. Ringer, T. Storelvmo and P.M. Forster (2021) J. Geophys. Res. Atmos. doi: 10.1029/2020JD033622

Towards consistent observational constraints in climate predictions and projections
Gabriele C Hegerl, Andrew P Ballinger, Ben Booth, Leonard F Borchert, Lukas Brunner, Markus Donat, Francisco Doblas-Reyes, Glen Harris, Jason Lowe, Rashed Mahmood, Juliette Mignot, James Murphy, Didier Swingedouw, Antje Weisheimer (2021) Frontiers in Climate. doi: 10.3389/fclim.2021.678109

Comparing Methods to Constrain Future European Climate Projections Using a Consistent Framework.
Lukas Brunner, Carol McSweeney, Andrew P. Ballinger, Daniel J. Befort, Marianna Benassi, Ben Booth, Erika Coppola, Hylke de Vries, Glen Harris, Gabriele C. Hegerl, Reto Knutti, Geert Lenderink, Jason Lowe, Rita Nogherotto, Chris O'Reilly, Said Qasmi, Aurelien Ribes, Paolo Stocchi, Sabine Undorf
J. Climate (2020) 33 (20): 8671-8692. https://doi.org/10.1175/JCLI-D-19-0953

Quantifying uncertainty in European climate projections using combined performance-independence weighting
Lukas Brunner, Ruth Lorenz, Marius Zumwald and Reto Knutti
Environmental Research Letters, Volume 14, Number 12
Citation Lukas Brunner et al 2019 Environ. Res. Lett. 14 124010
DOI 10.1088/1748-9326/ab492f

UKCP18 Land Projections: Science report
Murphy JM, Harris GR, Sexton DMH, Kendon EJ, Bett PE, Clark RT, Eagle KE, Fosser G, Fung F, Lowe JA, McDonald RE, McInnes RN, McSweeney CF, Mitchell JFB, Rostron JW, Thornton HE, Tucker S, Yamazaki K (2018). https://www.metoffice.gov.uk/pub/data/weather/uk/ukcp18/science-reports/UKCP18-Land-

report.pdf

A climate model projection weighting scheme accounting for performance and interdependence
Knutti R, Sedláček J, Sanderson B M, Lorenz R, Fischer E M and Eyring V 2017
Geophys.Res. Lett. 44 1909–18 https://doi.org/10.1002/2016GL072012

A representative democracy to reduce interdependency in a multimodel ensemble
Sanderson B M, Knutti R and Caldwell P (2015), J. Clim. 28 5171–94

Probabilistic Projections of Transient Climate Change.
Glen R. Harris, David M. H. Sexton, Ben B. B. Booth, Mat Collins, James M. Murphy,
Climate Dynamics (2013) doi:10.1007/s00382-012-1647-y. Supplementary Material.

Extreme specific probabilsitic projection

UKCP Additional Land Products: Probabilistic Projections of Climate Extremes
J.M. Murphy, S.J. Brown and G.R. Harris (2020)
https://www.metoffice.gov.uk/binaries/content/assets/metofficegovuk/pdf/research/ukcp/ukcp-probabilistic-extremes-report-september-2020.pdf

Climate projections of future extreme events accounting for modelling uncertainties and historical simulation biases
Simon J. Brown, James M. Murphy, David M. H. Sexton and Glen R. Harris
2014 Climate Dynamics https://link.springer.com/article/10.1007/s00382-014-2080-1

2) Accessibility

I do not underestimate how difficult this challenge is and it is difficult to know how to advise on this. Apart from the hurdle of the Bayesian terminology I think the reader will struggle to put all the pieces together. This could be alleviated with some text at the beginning of the Methods section giving an outline of the whole procedure, how the different parts fit together, the assumptions of the approach, how climate model deficiencies are accounted for and how biases with observations are dealt with.

For those readers who will probably never be that comfortable with the more statistical aspects of the paper I think they will be greatly helped if there was more emphasis and care in the physical interpretation of the method and discussion of the example analysis. Also, a comparison with other posterior work who's emphasis is more towards the physical modelling uncertainty of future climate
For example such a discussion would include the impact of this being an ensemble of opportunity, how carbon cycle uncertainty and aerosol modelling uncertainty is samples and their consequences for the results presented. Also a discussion on the relative importance of X_R & X_G, what different do they bring to the analysis.

Other general comments

1) I know section 5 is a really important part of the paper but it really does interrupt the flow. Might the authors consider moving it to an appendix?
2) Are return periods and their changes particularly helpful due to their extremely nonlinear behaviour? The issue with looking at return periods in the present day is that the return period will be dominated by uncertainty in the shape. The actual increase in extreme temperatures, however, will be dominated by the change in location. Present day factual and counterfactual comparisons if looking at return periods will be dominated by shape uncertainty whereas present to future comparisons of the 100y return level will be dominated by location change and its uncertainty.
3) The adaptation sphere is very focussed on high resolution modelling of the future climate (perhaps too much). As this paper is also concerned with determining the likelihoods of future extremes it would be useful if the authors could comment on how their method might accommodate regional climate modelling.

4) X_R and X_G - these will be highly correlated. Can you demonstrate that both are required? If you only had one will not the parameters not just get adjusted to compensate? Or put it another way, does the small bit of extra info when using both lead to a significantly better outcome?

Minor comments

Abstract and Intro

1. Abstract general: - could be improved to better describe the solution it is providing in more general terms (currently only 81 words)
2. Abstract general: does not mention the improved Bayesian sampling mentioned in the introduction. This seems a key point of the paper.
3. 3: I think the abstract needs to mention that your "observations" are ERA5
4. 5: "Counterfactual world" - this is jumping in very deep very quickly into DA jargon
5. 29 compbining
6. 24-28 I would be helpful to have some non DA focussed literature on extremes in climate
7. 29-34 very limited review of other literature attempting Bayesian approaches to present/future climate, see above
8. 34-35: Flow of text is a bit confusing. It is not clear the next paragraph is addressing these two issues because of the way it starts. The "Here," suggest the paragraph is setting off on a new topic.
9. 46-47: Could you please compare like with like. Currently it is CPU time vs wall time. Surely wall time is primarily dependent on how much compute resource you have at your disposal.
10. 53: "observed" - not true observations should be something like "as represented in ERA5"
11. 59,60: "classical attribution…specific definition" Not sure this will make sense to the reader

2 Data

1. 64: refer to refers
2. 69: being pedantic 0h to 23h misses out one hour but I know what you mean, perhaps 0:00 to 23:59?
3. 74: Presumably three is some urban warming in the Paris observations. Please could you comment on how this affects the results somewhere, perhaps in section 6?

3 Method

1. 106: I am confused by the phrase "We add to the model". Are you suggesting (X_R + X_G)? If so, put it in the definition of eq 1. But if not perhaps "In addition, we can replace X_R with X_G.." or something similar would be clearer.
2. 111: English has gone a bit wrong here
3. 123: The notation for the scenarios is grim. Could we not have X_R,A,S_i and define S_i elsewhere?
4. 123: It would be good to acknowledge the unavoidable assumption that the climate system responds linearly to forcing. e.g Mu_t and sigma_t are constant wrt different forcing scenarios.
5. 124-130: This assumes all GCMs are equal. It would be good to acknowledge this and that other approaches have seen it necessary not to make this assumption (references below)
6. 131-132: this has been said earlier.
7. 143,4: it would be good to say that the energy balance model is forced with natural forcings only and the radiative forcings are natural only.
8. 151 theta_R and theta_R - second should be theta_G?
9. 171: Fig S3 seems to indicate sigma_t is very very small. Worth commenting on the physical significance I think.
10. 171,200: Fig S3 & Fig 2c-f and their discussion. I think it would help the reader if you reminded them that these plots are site specific using your Paris observations
11. 177: I think you can only say *impossible* if the probability of the shape being >=0 is zero. I don't think you have yet shown this.
12. 189: m is used as a superscript and a subscript in this line. Is this what you mean?
13. 189: This is the first mention of bias between the climate models and truth which is very late in the paper. Too late. There should at least be a pointer earlier in the text that climate model bias is

addressed later in the method description. Also I don't think you can deal with the issue of climate model bias in just ~12 words! What sources of bias is it accounting for? In The GEV parameters? In the covariates? A description of how and how well is surely needed.

14. 196: "grid point containing" earlier you were fitting to station obs for the plots (l137) - has it changed?
15. 196 Fig2b: 1940-1960 All GCMs are cool wrt Obs. Please comment.
16. 201: black ellipses missing
17. 241: covariate FOR Europe

4 Comparison

1. Section 4 I'd suggest renaming to "Comparison with the independent scenarios method".
2. 275: What are the consequences of mu_1 being so different for obs vs gcm? Some would argue that if the climate model is so biased can we trust its physical representation of real world extremes.
3. 280: "Based on the estimates of the laws" - not sure this will mean very much to most people.
4. 302: "average energy" - I wonder if this is the right term in a geophysical journal? Joules?
5. 304: I do wonder how much the average reader will get out of Fig 3 and I note that there is not that much discussion in the text for it. Perhaps just have the last column in a 2x2 format?
6. 314: "does a good job" - perhaps a bit too colloquial?
7. Is there a low bias in ERA5 for some regions? eg UK had 40C in 2022 although this was only a single day. Kay et al. (2025) has much lower return periods for single day events and one would have thought that to first order changes in return period will be somewhat similar for different metrics of extreme hot temperatures
8. I think maps of GEV parameters would be very interesting to most readers, say for 2000, 2024, 2080? In the supplementary info if needs be.

6 Example

1. Fig 4 is an odd beast. It seems like it is implicitly assuming that the climate has been stationary between 1940 and 2024. For example two points near-ish to each other might have seen the same max temperature at very different times, say 1940 and 2024 for arguments sake. The probability of those two events are very different as the 2024 climate is much hotter. Yet the calculation of 4c assumes they have the same probability of occurring. At least could we have a complimentary plot of the 100 year return level (or whatever) with and without human influence and the difference please? Also for the caption I found "mean of all scenarios" a bit confusing as during the observed period the forcings are the same? I think it would be clearer to say with and without human influence.
2. 373: given the lack of spatial dependence perhaps a warning that these numbers cannot be used to calculate the likelihood of a hot event occurring in a given region or country without a correction to account for extremal dependence.
3. 383-387: I find this spatial variation in return periods across quite small distances alarming. Some of this will be due to the issue in my point above but it does not seem physical. You could check how the observed exceedance rates compare between different locations. e.g.
    a. Those areas where 4b shows very high return periods. What are the empirically observed exceedance rates? Are hot events occurring more frequently than predicted here?
    b. Also for the regions in 4b with very frequent return rates, N Africa, E Turky are we seeing their frequent occurrence in the observations?
    c. And eastern France and western Germany (approx. Nancy & Stuttgart) to see if the different return expectations from the plots in these two places are supported by the data.
    d. Kay et al 2025 found the 2022 UK record event of exceeding 40C (admittedly a single day maximum, but one would expect different averaging periods to be somewhat in step) to be 1 in 24 years which seems rather at odds with your plot (4b) if >500y for the region where these temperatures occurred. Comment?
4. 385 I stumbled a bit over western and southern Asia. Perhaps "southern Caucasus" ?
5. 413: Don't think you need to repeat the results of the counterfactual world here as there is no reason for them to be different to 2040?
6. 404: these holes really bother me and I think they are an artifact of the methodology. Please can you diagnose their cause. Are they coming through the GEV terms or the X* terms? Maps of all of these

parameters at 2040 and 2100 (Mu, sigma, psi, and all the Xs) would be helpful in this regard.

7 Conclusions

1. Section 7 seems to be an afterthought or the authors have run out of steam (not surprising). It would be good to have:
    a. A candid discussion of the limitations of the approach
    b. A more thorough comparison of results with other studies
    c. One of the difficulties this paper will have will be convincing people that the pain of understanding it and using the code is worth the effort. At present the case for is not very clear, at least to me. What might help is a companion plot of Fig S3 For Paris for a given return period say 1 in 100 years with uncertainty (ideally including a profile likelihood approach, see Coles 2001 Fig 2.3) , for the method presented here vs a non stationary univariate GEV maximum likelihood approach with Xt's calculated directly from the GCMs used.
2. 434: From my perspective very little verification has been undertaken with respect to current literature. Please could you explain why you think this statement is justified?
3. 456: section 7.2 is rather niche and that there are far larger issues that would benefit having a discussion on. Such as how this work compares with other work that attempt to produce future posterior distributions of climate variables. I include references to such studies above.