Review of Rust et al. for HESS

This paper aims to explore relationships between periodicity in the NAO and UK river flows. It uses wavelet analysis and then examines the propagation of observed, apparent cyclicities from meteorological inputs to river flow response, and analyses the contribution of catchment characteristics (specifically BFI and a groundwater response variable). It builds well on previous work on wavelet applications in groundwater records by the team. The paper is well organised and clearly presented.

I am no specialist in wavelets but the analysis appears appropriate and well executed – although I have raised some potentially important questions. There is also a novelty to this work – there have been a few papers on cyclicities in UK river flows (that are not cited here: Sen, 2009; Franco Villoria et al. 2012) but this has not been studied extensively, nor has there been the consideration of hydrological propagation pathways. The work has suitable novelty and potential international interest for publication in HESS.

I do struggle to fully grasp the significance of the work, however. The wavelet analysis reveals intriguing patterns but these periodicities are difficult to interpret. I find this paper like many similar wavelet/frequency papers in that a lot is down to the interpretation of the plots and assumptions made on driving processes - there is a lot of faith here seems to be on 'NAO like' signals which are taken and run with but to me are far from clear in the plots, and wrt previous work. There is no actual statistical link with the NAO (cf. FrancoVilloria, etc) which should be made clear from the outset (possibly even title?).

As far as I understand the NAO literature it is somewhat debatable as to whether such patterns are physically meaningful. They do crop up in some papers but I have not seen so much that is convincing as to plausible physical interpretations (this can be said about a lot of work on cyclicities in hydroclimatic series in general). Most studies indicate that if NAO cyclicity exists it is weak, sensitive to choice of NAO index, season (DJF, DJFM) and tends to come and go over time as noted in the original Hurrell paper (these non-stationaririties being another barrier to practical applications; see also Franco-Villoria on nonstationarity of NAO-flow linkages). This should be elaborated on. See also important statistical work of Mills (2004).

There is some discussion of some of this literature in the current paper (a lack of references in the reference makes it hard to find some of the key papers cited in relation to NAO periodicity) but I find some claims quite tenuous as to what the cited literature reveals wrt the findings of Rust et al. The discussion is very wide ranging, but at times highly speculative and overreaches beyond the analysis made here, and as noted below sometimes reflects misinterpretation of existing papers, or reference to work that is not directly relevant.

Finally, I am not sure of the practical significance of such findings. The authors suggest this reveals a 'degree of forecasting' and 'critical process understanding' but I am not sure how water managers can really use such information beyond a general raised awareness. And for that, there are already operational physically based seasonal to decadal prediction systems (the Met Office DePreSys system bein one) which already give an indication of potential NAO states (and moreover, more importantly, the large-scale variables that drive the NAO, like ENSO, QBO etc which are much more predictable). How would findings really be used in reality? Perhaps the authors could comment. That said, there are wider applications of these findings in explaining general time series evolution (trends and variability) a huge area of interest in the literature.

The added interest in this paper is the catchment propagation but while this element of the paper is well executed (albeit with some questions below) I am not sure it tells us much new? We know that BFI modulates precipitation signals anyway, and there's a reasonable handle on lag times from past literature, so the results are not that surprising. Great to see the GRT as a predictor though.

In summary I think the paper is novel, well executed and potentially a valuable contribution of the literature, but its framing could be improved, as well as the coverage of the significance of the findings. There are some moderate technical and presentation issues which also need to be addressed before publication in HESSD.

Specific Comments

L18 – 'degree of forecasting' – I see this word is being used in a general sense but I don't think this is really what the paper is offering up, rather some general tendencies of multiyear river flow behaviour in time series, itself possibly useful but not really 'forecasting', although perhaps for general preparedness. I think any findings from studies like this are more useful in providing context for studies of long-term trends and variability (helping shed light on widely reported 'flood poor', 'drought rich' periods and so on) rather than providing any information on preparedness. The authors could comment on this potential application too.

L42 – L44 (around) maybe also worth clarifying early on that the NAO is primarily a driver of wintertime variation, noted later but should be in the intro.

L48 – [and into discussion]. I wanted to look into past research on claims of NAO cyclicities but found it a bit lacking in these papers – Tabari doesn't really look at muliannial cycles; Su found cycles but did not appear to look at the NAO. Neves relevant and useful. Kuss and Gardak not in reference list; Meinke looked at ENSO; no reference in list for Tremblay; no reference in list for Olsen, but found it and appears to be a different beast on Paleo timescales so not sure of its relevance here. It's difficult to examine the wider evidence based with these missed references. Some that are cited appear to be not entirely directly comparable – I'd recommend a careful re-reading and re-positioning of this work with the literature here (and in the discussion]. See also missed references in intro and other useful international papers e.g. Labat (2010) on various possible mechanisms of cyclicity globally.

L67, L68 – there are many more studies that look at NAO influences on UK streamflow in winter and more generally and these shed important light on regional patterns in some detail, and also catchment characteristics – would be worth referring to these. See e.g. Laize et al. 2012, Svensson et al. 2015 and references in both. Laize & Hannah really looked at this propagation question in some detail (for interannual rather than periodic behaviours but still relevant).

L136 – L139. At 705 catchments, this will inevitably be a very mixed set of catchments subject to all sorts of human influences. Ideally hydroclimate studies linking large-scale drivers to hydrological response should use relatively undisturbed catchments - human disturbances can alter the chain of propagation from signal to streamflow response. I agree however that there are few such catchments (see e.g. Harrigan et al. 2018) which would limit the range of BFI/GRTs, so using the wider set is reasonable. While I expect that the outcomes from the broad-scale national picture is unaffected, there will likely be significant heterogeneity in the catchment results. This warrants a comment.

L163. This categorisation into even BFI classes will lead to very uneven numbers in the groups, given how skewed (and slightly bimodal) the BFI distribution is. Being that there is nothing meaningful about these arbitrary thresholds why not try and make them cover the distribution better? There are very few catchments at BFI <0.25, and not so many >0.75, with the great majority in the middle two classes. (I do not have this to hand but there are NRFA BFI distributions available somewhere in the literature I am sure and it would be easy to check). How would a more even classification effect the later results?

L184. The point here is a valid one that BFI is simply an empirical property from the baseflow, but this could be generalised – rather than referring to throughflow, it can reflect any terrestrial storage in soils, lakes (lake and reservoir catchments also have high BFI) and so on.

L187 – I do not know how GRT is distributed but mirroring my BFI Comment above can the authors comment on this? It looks like it is very uneven from Fig 7 with many more in the highest class.

Fig2 caption – might need to explain why this is log GRT as you have not introduced that yet, only referred to the categories.

L204 – just a comment really for future work. This study does not consider the transition seasons, which is fine. But if the focus is really on propagation from winter NAO met signals I would have thought looking at spring would be really interesting – it may help better explain the propagation from winter.

L250. (and 2.3.4 generally). I'm not sure I fully follow the logic here so it needs some clarification – at the moment it sounds like this process is doing some heavy lifting. I failed to follow the process leading up to L250 that indicates "this produced a wavelet power for each dataset that is considered NAO like'. Why would this be considered NAO like a priori? This sounds particularly important given the 'NAO like' signal is then used to producing the residual series that is so important thereafter for capturing the 'measure of modulation of signal strength'.

L258. Following the above, it also appears to be a big assumption to produce this residual series for the summer based on winter rainfall. While in very high BFI catchments a very long lag time may be expected a priori, this is not necessarily the case in many (most?) cases (See my earlier comment about the spring season being ommitted).

I'm just concerned that two comments (while partly no doubt reflecting my lack of understanding of the process) are reflective of some major assumptions being made in this study which are (at face value) in danger of baking in some of the conclusions, somewhat – hopefully a clearer exposition will help allay these fears.

L271 - Given this pairwise testing between groups, is my earlier comment about the irregular distribution of BFI a potential issue (would this look different with different more representative groups?)

L285 I am no expert on wavelets. But when I look at Figs 3 and 4 I wouldn't say a c.7 year cycle leaps out at me – rather, higher powers at a range of years <c.8 years. Especially for winter rainfall. That is, I see there are peaks in significance but are you really that confident in there being a (even approx.) 7 year cycle in these? IN fact I don't really see the 'two discrete bands of periodicity' (l279). I guess this is all down to interpretation but this causes some modest concern if this is the basis of the identification of 'NAO like' signals – please elaborate on this. It's also really difficult to see the variability in the wider cloud of catchments shaded grey, but it looks like there is a range, especially for rainfall – some comment of this would be useful.

L290. 'Wavelet p values indicate the detected wavelet powers are not the result of external forcing'. Is this strictly true, I thought this just indicates it is not AR1 generated – I assume it does not rule out that it is internal variability, which could be driven by all sorts of long-term persistence processes (see the extensive work on the Hurst phenomenon and many papers of Demetrios Koutsoyiannis), as opposed to external forcing. (corollary to this, re: my point in the general intro about physical significance, when I backtrack into the literature on NAO periodicities, back to the Hurrell papers cited, it seems far from clear as to being settled whether NAO periodicities, as they are, are externally forced or internal variability).

L294 – should this say 'river flow records' instead of groundwater?

L340 onwards and Fig 6 – note whether previous question of BFI distribution has any impact on these findings?

L416 – I would not have said this strong conclusion on the difference between the periodicities between winter ('present') and summer ('absent') rainfall really emerges from Figs 2 and 3 as noted earlier. I may be missing something but this seems quite an open interpretation of those data. Important as the seasonal differences are majored on.

L420 – L425. Following on from this, this discussion hinges on there being multiple periodicities at different peaks between the seasons, but my reading of those plots makes it hard to really pick out any of these as 'peaks'.

L425 – I also think this 16 – 32 periodicity is very difficult to see, let alone link to the EA pattern – and I could find no work on this in the Rust et al. 2019 paper cited? An important general point though is that while the NAO is the leading mode of variability there is a whole laundry list of (interacting) influences (Scandinavia pattern, AO, etc) as well as the lower frequency SST drivers (AMO, ENSO) that are not considered here given focus on NAO-like, even though other work suggests they could also manifest themselves on similar timescales (see e.g. Labat, Villoria). Worth comment in intro & discussion?

L453. Haarsma not in the reference list, But on looking at this paper I don't see this SE England outcome on any of their results maps? Please clarify. In general, I find this whole gulf stream section really speculative. I looked (admittedly quickly) but could not really find much in these papers to support this e.g. concurrent but lagged correlations in Watelet rather than periodicities; little specific mention of GS in Hurrel and Deser. A lot is made of the GS as a mechanism for the key NAO-like behaviour central to this study, so this reference to other work is important and could be checked and strengthened.

L576 – L582. Related to my points in the intro, good to see potential applications but this is quite a long way off from what is discovered in this paper so some of these claims could be moderated.

References

Franco-Viloria et al. 2012. http://www.jenvstat.org/v03/i06/paper

Sen: https://onlinelibrary.wiley.com/doi/abs/10.1002/hyp.7224

Svensson et al. 2015. https://iopscience.iop.org/article/10.1088/1748-9326/10/6/064006

Harrigan et al. 2018. https://iwaponline.com/hr/article/49/2/552/37838/Designation-and-trend-analysis-of-the-updated-UK

Labat, 2010. https://www.sciencedirect.com/science/article/pii/S0022169410001071

Laize & Hannah 2010. https://www.sciencedirect.com/science/article/abs/pii/S0022169410003331?via%3Dihub

Mills: https://rmets.onlinelibrary.wiley.com/doi/epdf/10.1002/joc.1003