Review AMT - egusphere-2026-1256

Title: Configuration of climatological limits for surface radiation measurement quality control: A global assessment using a novel radiation climate classification

First author: Zhiwen Wang

Summary

This paper describes the methodological implementation of a regime-specific quality control standards for the extremely rare limit (ERL) climatological test used to filter anomalous data from BSRN data. In its current implementation the ERL is applied globally to all sites in the same manner to remove outliers outside of the range of physical expectations. The authors argue that this test is far too conservative and develop a method for deriving regime dependent limits that more closely follow the data.

Overall Feedback

I think that this paper is well written and worthy of publication with only minor revisions and response to some of my following questions. Most of the feedback below consists of recommendations about places in the manuscript to refine and be more numerically descriptive around figures and methodology.

At a pretty basic level, I need some clarification regarding why the ERL criterion on the plots are displayed as regions. The test, as I read the inequalities in eq. 1-3, appears to define a fixed negative lower bound and a SZA dependent upper bound limit. However, in figure 1 these are plotted as regions rather than a single upper-bounded functional limit. Could you please explain more directly in your response and in the text how these inequalities produce this bounding region? Does the area of the region come from the joint solar zenith and the sun-earth distance dependence, and thus there is a distribution at each SZA for different times of year (changing R)?

Another domain specific clarification – and I'm sure this is likely a conventional nomenclature thing, but I found the symbolic transition from irradiances to transmittances kind of confusing initially. Some of it follows from the irradiances: Bn becomes kb (beam), and Dh becomes kd (diffuse) – so why does Gh not become kg but instead is written as kt (total?/transmittance?)? Perhaps a brief response is all that's necessary, but it could easily be addressed with a short sentence in-text I suppose.

Could you potentially comment on why the northern "bar" and "tik" BSRN sites don't find their way into the category 2 despite sharing higher latitude seasonality of some of the

other polarward clustered sites. Are tik and bar cloudier than the other members of group 2 that are likely very dry? The ena site is very cloudy, but is it really comparable to these higher latitude group members? Is this a sign that diffuse, either from clouds or surface reflection is playing a common role for these sites? In figure 5 I can see two rows that appear to have greater distance in some features – are these bar and tik? It is my understanding that some BSRN sites have information regarding upwelling measurements as well – is there perhaps something that can be learned here from surface reflectance in addition to transmitted radiation? It seems to me that these two cases might be distinguished by whether or not diffuse radiation is traced back to surface reflection or from scattering within clouds.

Specific Feedback

1. Figure 1: It would also be might also be interesting if figure 1 indicated the number and/or percentage of screened observations. It's clearly very small by visual inspection, but it's hard to tell with small scattered points in low-density regions – I think this change would bolster the numerical point the figure is making visually.
2. Section 2.2.1: Could you further explain your selection of the harmonic order maximum of 25. In text you say that it was selected empirically – but by what metric? Is there perhaps a more solid physical motivation for the number of harmonics that are appropriate? If you divide 4 years into 25 intervals that's something on the order of two months, was that the shortest oscillation period you wanted to be able to resolve? Could this reasoning be important to document for future studies? Furthermore, can you explain exactly how the two features are extracted (e.g., the maximum – minimum across all 4 years vs. average yearly maximum minus average yearly minimum)? Does this have consequences in how you consider annual and interannual variability in radiation climatology later?
3. Line 131: Opening sentence of the paragraph at the start of section 2.2.2 is a little confusingly worded. I think it's specifically because you're listing the transmittance variables and later mentioning irradiance components and thus your use of "latter" appears to refer to a list with more than a pair of variables (as is commonly how former and latter is used. Would this change maintain your meaning? :
    a. "The density-based features are extracted from ~~kt, kb, and kd~~ the transmittance components, rather than from the irradiance components themselves, because the latter random variables are "contaminated" by yearly and diurnal cycles and are thus unable to fully reflect sky conditions."
4. Figure 3: For your beam transmittance fits with power-sn-sn mixture of distributions of the beam transmittance (kb), are you getting densities that exceed zero above the

upper maximum of the distribution? Similarly there is also a poor fit near zero – presumably these are both features driven by a linear constant being added to your distribution. Could these discrepancies exceeding maximum and near minimum values of your distribution cause issues later?

5. Line 357: change “distinct regime physics” to “distinct physical regimes”
6. Figure 5: Could you provide some quantification of the average *intragroup* distance (e.g., group 2 has greater distances within its own group, indicating greater diversity) and *intergroup* distances (e.g., how similar is group 2 to 4, or group 1 to 5?). For example it looks to me like group 1 and group 5 have a smaller WD to one another than other groups – given that the explanation for these group is different, then why? Their distributions in figure 7 also look similar (though with less low SZA sampling in group 5).
7. Figure A6: It is noteable that both GAN and GUR seem to have anomalous dense sampling of high SZA – is there some issue with the temporal sampling at these sites? There’s a relatively high sampling of observations above SZA~80 and a sharp cutoff below that doesn’t appear in the comparable sites you selected. Is the SZA sampling itself anomalous compared to other sites? What does that mean for this regime cluster? Is this just a data collection artifact? The dispersion in BNI seems like a reasonable explanation for why these groups cluster, but what causes this unique SZA dependent sampling? I didn’t notice this in figure 7, why?