General comments: This paper presents interesting and novel results about the step-like development of many climate variables. There methods that are used are appropriate and it is good that a variety of tests and methods are applied. I still cannot recommend publication in its current form because the discussion of the hypotheses suffers from a lack of clarity (see specific comments). Thus I would recommend the authors to revise and resubmit the paper.

Specific comments: POINT 1: My main concern is that it is not clear enough what exactly the hypotheses are that are tested. Sometimes the authors say that what is tested is whether (i) internally and externally forced components of the climate system are independent or not (page 3, page 29); sometimes they say that what is tested is whether
(ii) the development of climate variables follows a trend or is step-like (abstract, Section 5). It is not the case that independence of internally and externally forced components of the climate system implies that there is a trend; this is possible, but there could also be independence and at the same time step-like behaviour. Also, it is not the case that dependence of internally and externally forced components implies that there is necessarily step-wise behavior. This could be the case, but there could also be dependence and a trend at the same time. As a result, it remains unclear what exactly is tested: (A) (Only) whether the internally and externally forced components are independent. (B) (Only) whether the climate variables follow a trend or not. (C) Whether the internally and externally forced components are independent AND whether there is a trend. (D) Whether the internally and externally forced components are dependent AND there is step-like behavior of the climate variables. Throughout the paper, the authors need to be clearer what exactly is tested.

POINT 2: Related to this, if what is tested is (C) and (D) (as is often suggested; cf. in particular the hypotheses on page 5), then it is important to see that (C) and (D) are not exhaustive (because there are also the possibilities that there is independence and a step-wise development; or that there is dependence and a trend). The authors want to test an exhaustive set of hypotheses, but (C) and (D) are not exhaustive.

POINT 3: Throughout the paper the assumption seems to be that “trend-like” and gradual as opposed to step-wise and non-gradual means that there is a linear relationship (e.g. on page 7). It is unclear why gradual implies that there is a linear relationship. There can be gradual behavior with various kinds of relationships (a quadratic relationship etc).

POINT 4: On page 7 the hypothesis states that there is a “(probably monotonic)” trend. The brackets are confusing. Is it now tested that the trend is monotonic or is it allowed that the trend is not monotonic?

POINT 5: on page 5 six tests are described. It should be clearly stated which tests test
which hypotheses (becomes clear later, but should be stated clearly early on).

POINT 6: The beginning of Section 2: here it is argued that the gradualist thesis is derived from induction. Yet, as the paper later argues, the data actually do not support the gradualist thesis and the gradualist thesis rather seems to be often adopted for no empirical reasons (convenience, simplicity). Hence it seems that the gradualist thesis is not justified by induction after all.

POINT 7: The beginning of Section 2: “The application of linear trend analysis to atmospheric warming is invariably justified as inference to the best explanation”. I am puzzled by this sentence. Why is there suddenly a reference to the inference to the best explanation (previously the matter of concern was induction).

Technical corrections: Page 12, line 28 and 29: “otherwise are p>0.05” should be inside the brackets. Page 25, line 32: spaces are missing between the papers that are cited.