The manuscript “Estimating return intervals for extreme climate conditions related to winter disasters and livestock mortality in Mongolia” analyzed the return periods/levels of the extreme events on summer drought and winter temperature, key factors of dzud (winter weather disaster) in Mongolia. The analysis was carried out by the Extreme Value Analysis framework (i.e. fitting to two types of models). In doing so, length of the period covered by the data is attempted to extend by using tree-ling and Siberian station records, which is the author insist helpful to let the results more reliable.

Although I found potential of contribution by the current study, explanation of motivation and technical detail is not enough in the current manuscript. Before consideration for acceptance, I think substantial revision is needed.

General comments

(1) Purpose of presenting the ARIMA model and the models with climate covariants is not clear. The authors say “establishing a relationship between drought conditions and climate variables, particularly precipitation and snow, is useful in understanding the dynamics that determine dry conditions”. However, it is obvious that PDSI is dependent on precipitation, and for understanding the dynamics the conditions in the preceding some months should be analyzed.

For winter temperature, the reason why the dependence of the return level on the Siberia’s data should be considered is unclear. It may be better to first convert to the
values in Mongolia (e.g., by regression models) and then to input the converted values to GEV and GPD than to directly consider the Irkutsk’s data in GEV and GPD.

(2) Discussion on whether the reliability has been improved compared to the case where data only after 1940 is used is needed.

(3) Stationarity means there is no impact by the global warming?

(4) English should be checked by native editors.

Specific comments

Page (P) 1 Lines (Ls) 23: In the manuscript, the usages of “return level” and “return period” are confusing. (see also Ls78, 101, 103, 116, 122, 124, 286 (and Fig 5), 297)

P1 L32: “thus”: I could not understand how the previous sentence supports this sentence.

P2 Ls33-34: “Here ---insurance” These are not treated effectively by the current study. If the authors want to argue this, reorganization of the manuscript is needed.

P2 Ls 38-40: A total ---in Mongolia: present the source (literature).
P2 L49: Tachiiri et al (2008) used snow amount (SWE), not temperature, isn’t it?

P2 L63-66: I could not understand the advantages of the index-based insurance compared to the livestock-loss-based one.

P3 L69: “few studies” Do you mean there are some examples? If so it is better to briefly introduce them.

P3 L74: In Bayasgalan et al. (2009), I could understand the mechanism how drought becomes more frequent in the future, but why dzud will be more frequent is not clear.

P3 L76: For the future, models are used?

P3 L88: Does “distribution” here mean probabilistic distribution, or spatial distribution?

P3 L100 Snowfall in winter is also considered to be important, isn’t it?
“mortality assumes that the size of the population does not matter.” I do not agree with that, as in my understanding it is possible that mortality depends on the size. Also, this does not explain why analyzing climate variables is more effective to moderate the damage of dzud than analyzing mortality or loss. If the authors consider mortality is not a good indicator, how about number of livestock killed?

If livestock death depends not only on climatic factors, but also on socio-economic factors, index insurance is not reasonable, is it?

“for the future”: is the future discussed in the manuscript?

What do “many opportunities” mean here?

I could not understand what “because ...” is for.

“Kaheil and Lall” should be “Lall and Kaheil (2011)”? I could not access Lall and Kaheil (2011). Add more information on this literature in the reference list so that the readers can access that when needed.
P5 Ls151-2: “which were adjusted ....in growing seasons”: More information on how they were done is needed.

P5 L154: “to improve risk analysis of Dzud and mortality of livestock in Mongolia”: Is this consistent with Ls 89-91?

P6 Ls161-2: The Mann-Kendall test is for each time series? If so each row of Table 1 is a pair of two clusters, to show the results of the Mann-Kendall test in the right-most column of Table 1 is not appropriate?

Ps7-8 Ls190-191: how they are “scaled”?

P8 Table 2: how fine is “high resolution” of WMO and CRU data?

P9 Eq 1 (and Eq.3): What does “+” in the last term mean? Also, there is no $y_+$ in Eq 1 (may be $z$ is so?).
P9 Eq. (5): $Zt$ is equivalent to $H(x)$ in Eq. 3? (same in Eqs 6-7)

P10 L223: it is better to move Section 3.1 to Methodology section.

P10 L224: How “fit” was done?

Ps10-11 L244-254: The oscillation shown in Fig 2 is not reflected in the models? Why? Also, I want to know how good fit was obtained by the models. (for other models too)

P10 L246: “$\sigma = 0.95+0.002t$”: Do you mean $\sigma=\exp(0.95+0.002t)$?

P11 Table 3: AR(3): Why order (3) is selected as the best model? By what mechanism?

P12 L260: “could be a real feature or an artifact”: Which is more likely here?
P12 L266: Why not 1902?

P18 Table 7: Why the difference between the values for 10,50,100 years are so small? Also if PDSI of around 4 is returned every 10 years, the threshold of 1 is not too small as catastrophes?

P18 Ls346-7: “All the results show that…” What meaning does it have in light of the purpose of the study? (same for Ls419-20)

P19 L378: Why seasonality should be removed?

P19 L380: “correlation” with what should be clarified.

P20 Ls391-3: The text is not consistent with Table 8.

P21 L408: How the threshold 20 was determined should also be explained.
P22 Fig 10: Why the uncertainty is very small for East and NW?

P22 Ls425-8: “We use the GEV model because the winter minimum temperature data is a single extreme value and that the GEV model is suitable for maxima and minima of block data.” then why you also presented the result of the GPD?

P23 Fig 11: What is the cause of the variations? Siberian data?

P23 L440: what does “self-calibrated” mean here?

P23 L449: “can be used to improve the risk calculations for livestock index insurance in Mongolia” How? (also in P 24 L 463)

P23 Ls453-4: Is it the case where rainfall is included as a covariate?

P24 L462: “through early warning systems” how the lead time is considered?
P24 Ls462-3: “the estimation of extreme value distributions and return levels has the potential to improve livestock index insurance”: How?

SI P 1 L7; Whereabouts is the “inflection point” in the Figure?

SI P 1 L14: “Therefore, it is reasonable to use a threshold of 1.0.” I could not understand why you could say so from the previous sentence.

SI P1 Fig S1: What causes high ACF for x > 20 (particularly for East). Also, add explanation of the dotted line near y=0.1.

SI P2 Fig S3: Add a bit more explanation on what the figure shows. What is reparameterization? Why the sign was changed by that?

SI Ps3-4; Add more explanation on Figs S5 and S6 (on the red marks, red curves, numbers at the center of the boxes etc).
SI P6 Table S4: Add information on the statistical significance.

Typos etc

P3 L84: extraordinarily -> extraordinary?

P7 L179: ARIMA is not spelled out. Also, a brief description of \((p, d, q)\) (like \((3,0,0)\)) would be helpful.

P9 L210: explanation after “where” is confusing.

P16: Figure 6. Add unit for the x-axis.

P16 L320: chosen -> chose

P17 L328: PDSI -> for PDSI?
P23 L448: Add “for GPD” after “the upper-bound Beta distribution”.

SI P1 L5: delete one “Figure S.4”

(end)