Response to Referee #1

General Comments:

1) This study examines the characteristics and trends across the Eurasian continent from 1966 to 2012. To do so, the authors assemble snow depth data from 1103 stations across the study area. How representative are the station (point) snow depth data of the overall regional landscapes of interest? For instance, are snow depth data in forested areas collected at airports or other open areas, that may not represent the regional snow characteristics?

Reply：Thank you for your comments and concerns. The spatial representativeness of stations is always a key and difficult problem in snow depth research or any ground-based studies at various regional scales. In fact, we did not do spatial interpolation of snow depth using these in-situ data across the study area just because the uneven distribution of stations spatially and among different landscapes. The passive microwave remote sensing snow depth products may mitigate the regional coverage problem, their low spatial resolution (25×25 km) and high uncertainties (up to 200%) provide no better help to the issue. The combination of in-situ snow depth data with satellite remote sensing snow depth will be a better approach but it is out of the scope of this study. Here for the first time, we present all data we can possibly collect from various countries over the continent and show snow depth spatial variations and temporal changes. We are fully aware the shortcomings of station distribution but this in-situ dataset and its coverage is unprecedented. We may read a lot of published literatures regarding snow cover extent in regional or hemispheric scales, but not snow depth. In this study, we present spatial and temporal changes in snow depth using available in-situ data.

2) Further to this, snow course data from the former USSR are also employed in establishing the snow depth climatology (see Section 2). Is it therefore a fair comparison to present the station (point) data with those from local (spatially averaged) data?

Reply: In our study, 440 stations have both snow course and station data. We compared the snow course averaged data and the station point data and found that they were statistically significantly correlated, and the goodness of fit reached to 94% (Fig. 1). Therefore, we are confident that it was a fair practice to combine the snow course average data and the station point data together in this study.

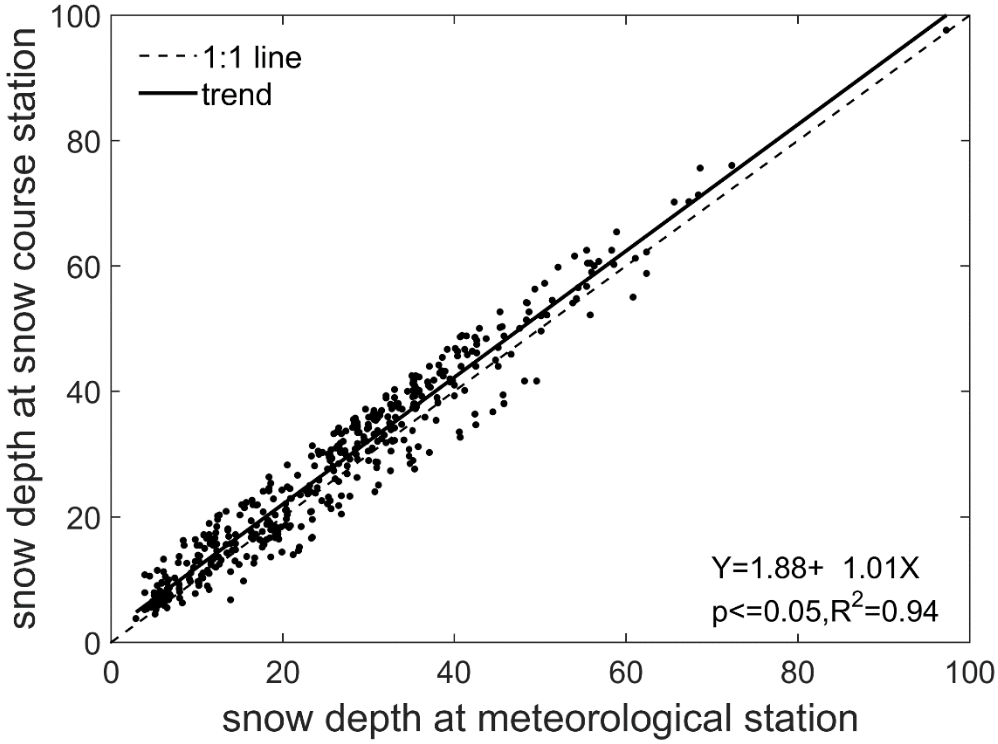

Figure1. The relationship between snow depth of meteorological station and snow course at 440 stations.

3) The Introduction section is quite lengthy and could be abbreviated by focusing on past studies that report climatologies and trends in snow depth across Eurasia only and the gap being filled by the present study. Further to this, the Introduction should emphasize the novelty of this research compared to previous studies cited in the text.

Reply: We have abbreviated the introduction, and focused on the report of climatologies and trends in snow depth, the existing problems of the previous studies, as well as the characteristics of our study.

4) The authors should consider the Mann-Kendall test to assess linear trends or other non-parametric trend analysis rather than linear regressions.

Reply: Any trend analysis is an approximate and simple approach to obtain what has happened on average during the study period. Linear trend analysis provides an average rate of this change. Despite there is a nonlinearity, the linear trend analysis is also a useful approximation when a systematic low-frequency variations emerged. Meanwhile, to overcome the strong assumption in ordinary least squares (independent and normal distribution), we added a Mann-Kendall (MK) test to identify the monotonic trend in snow depth. These two test methods could provide more robust and comprehensive information of the trend analysis. We have added the method introduction in the “data and methodology” section and discussed the similarities and differences of the two kind of trend analysis results in the “results” section.

“Any trend analysis is an approximate and simple approach to get what has happened on average during the study period. Linear trend analysis provides an average rate of this change. Despite there is a nonlinearity, the linear trend analysis is also an useful approximation when a systematic low-frequency variations emerged. (Folland and Karl, 2001; Groisman et al., 2006). The linear trend coefficient of snow depth was calculated to represent the rate of change at each station. The Student T test was used to assess the statistical significant of the slope in the linear regression analysis and the partial correlation coefficients, and the confidence level above 95% was considered in our study. Meanwhile, to overcome the strong assumption in ordinary least squares (independent and normal distribution), we applied a Mann-Kendall (MK) test to identify the monotonic trend in snow depth. Confidence level above 95% was used to determine the statistically significant increase or decrease in snow depth. These two test methods could provide more robust and comprehensive information of the trend analysis.”

“The Mann-Kendall statistical curves of annual and maximum snow depth were consistent with the linear trend analysis (Fig. 5). The increasing trend of annual snow depth reached to the 0.05 confident level in the late 1980s and from the early 1990s to the mid-1990s; it reached to the 0.01 confident level in the late 1990s. The decreasing trend reached to the 0.05 confident level from the early 2000s through the mid-2000s. The intersection of the UF curve and UB curve appeared in the mid-1970s, it indicated that the rising trend was an abrupt change during this period. The abrupt change point of the maximum snow depth was in the mid-1980s, then it increased significantly ($p \leqslant 0.05$) from the early 1990s through the mid-1990s, and it reached to the 0.01 confident level from the late 1990s to the early 2010s.”

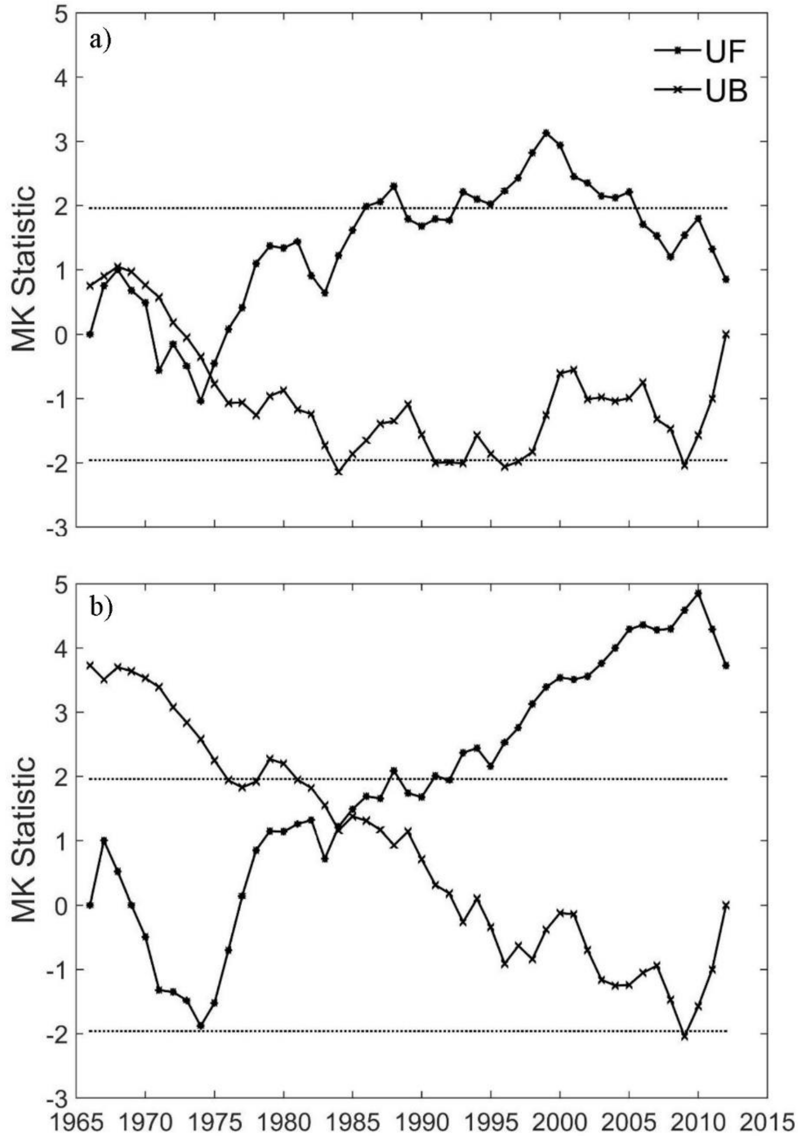

Figure 5. Mann-Kendall statistical curve of annual mean snow depth (a) and maximum snow depth (b) from 1966 through 2012 across the Eurasian continent. Straight line presents significance level at 0.05.

"In order to identify the monotonic trend in monthly snow depth, we conducted the MK test (Fig. 7). In October, snow depth represented a decreasing trend and it reached to the 0.05 confident level only after 2010. The statistically significant changes of monthly snow depth in November during the period of the late 1980s through the early 2000s, though it was not statistically significant with the linear regression. From December through March, there were increasing trends in monthly snow depth and the abrupt change point appeared in the mid-1970s. In the linear regression analysis, the variation of snow depth was not significant in December. However, the results of M-K test showed that the increasing trend of monthly snow depth reached to the 0.01 confident level during the mid-1980s through the late 1990s, and then it decreased during the 2000s. From January to March, monthly snow depth increased significantly ($p \leqslant 0.01$) from the mid-1980s to the early 2010s. In April, the statistically significant increase was found from the late 1990s to the late 2000s, and it reached to the 0.01 confident level after 2000. Consistent with the linear regression, the trend in monthly snow depth was not significant in May."

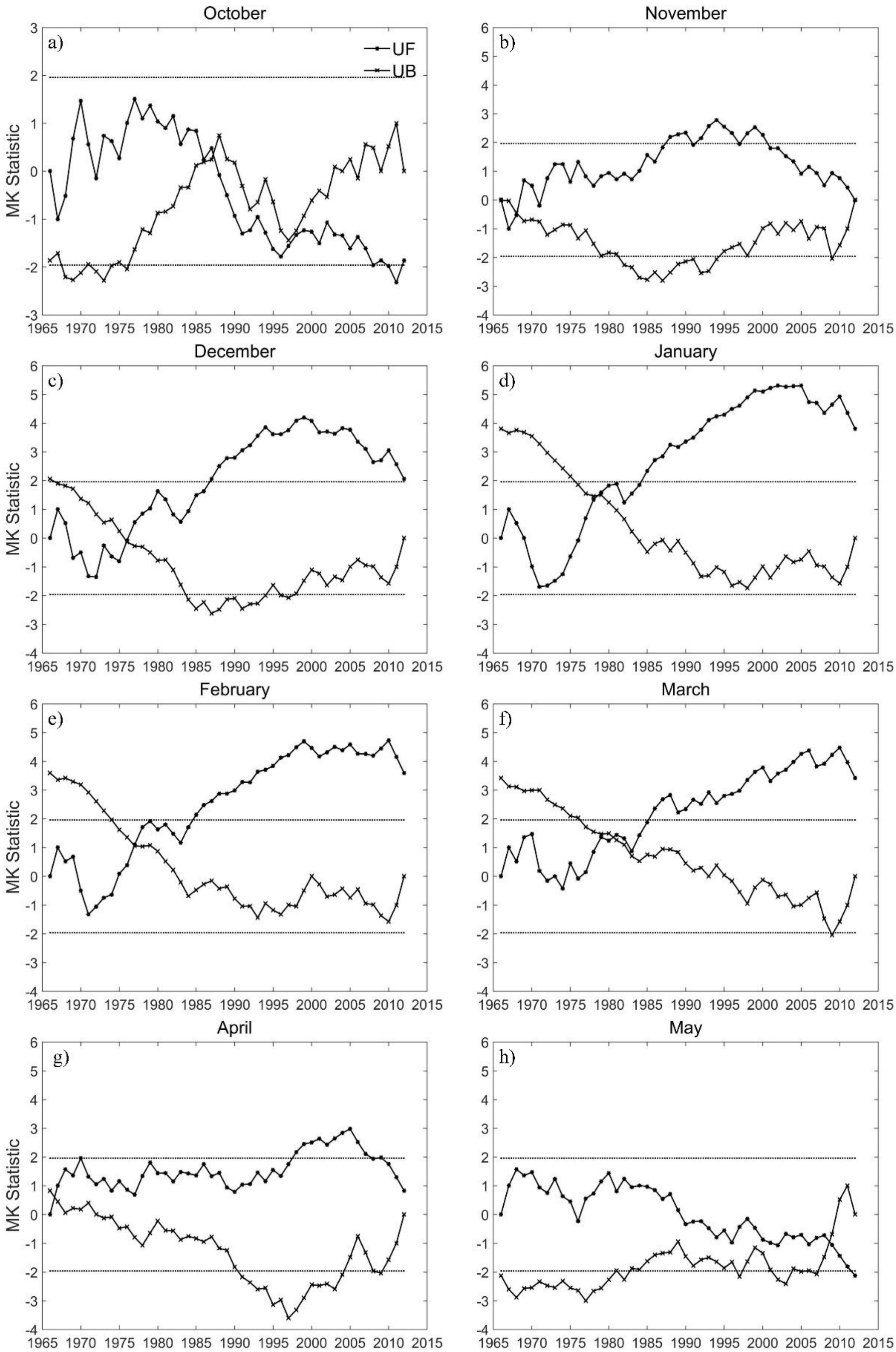

Figure 7. Mann-Kendall statistical curve of monthly mean snow depth (from October to May) from 1966 through 2012 across the Eurasian continent. (a) October, (b) November, (c) December, (d) January, (e) February, (f) March,

(g) April, (h) May. Straight line presents significance level at 0.05.

5) Do the linear trends reported in Section 3.2 exceed the variability in the snow depth data? In other words, are there "detectable" trends in snow depth, i.e. with the signal greater than the noise in the system?

Reply: The Student T test was used to assess the statistical significant of the slope in the linear regression analysis, and the confidence level above 95% was considered in our study. We have tested the results of the linear trends in Section 3.2, and the results show that all of the "detectable" trends in snow depth were greater than the noise in the system.

6) All figures are rather small and difficult to interpret when printed on paper.

Reply: Thank you for pointing this out. We have expanded all figures.

Specific Comments:

1) P. 1, line 21: Insert "a" before "snow depth". Then insert "its" before "spatiotemporal".

Reply: Has been done.

2) P. 1, line 27: Consider a word other than "dramatically" here. Are these statistically significant trends?

Reply: We replaced it with "significantly". In our study, the trends with the confidence level above 95% were only considered.

3) P. 3, lines 10-20: Note that the tense for verbs changes throughout this paragraph.

Reply: We replaced "is" with "was" in line 15, and replaced "promotes" with "prompted" in line 19.

4) P. 3, lines 22-24: Are the soil thermal conditions reported here for winter only?

Reply: Yes, the soil thermal conditions are in winter.

5) P. 4, line 8: Delete "the" before "ecological".

Reply: Has been deleted.

6) P. 5, line 8: Delete "In order" and begin the sentence with "To obtain…"

Reply: Thank you for your suggestions. We revised it.

7) P. 6, lines 4-8: Delete “Using data from ground-based measurements” as this repeats text from the previous sentence. Also, please rephrase the statement “a detailed description of snow depth”, as this suggests the paper goes at length in describing how snow depth is defined, which is not the case. This entire sentence is awkward and quite long, so should be rephrased and perhaps divided into two sentences.

Reply: Thank you for your suggestions. We rephrased the sentence: “The objective of this study is to investigate the climatology and variability of snow depth, and analyze snow depth relationships with the topography and climate factors over the Eurasian continent from 1966 to 2012.”

8) P. 6, line 14: Snow depth data from 17 countries are apparently used in the present study; yet Table 1 lists only three countries (former USSR, Mongolia and China) as sources for the snow depth data.

Reply: Seventeen countries includes China, Mongolia and 15 countries previously belonged to the former USSR. In order to avoid the readers’ confusion, we deleted “in 17 countries”.

9) P. 6, line 15: Insert “a” before “daily”.

Reply: We inserted it.

10) P. 6, line 18: Replace “5” with “five”.

Reply: We revised it.

11) P. 6, line 22: “SWE” has not yet been defined.

Reply: We have defined SWE at P. 3, line 7.

12) P. 8, line 2: Delete “In order” and start the sentence with “To reflect…”

Reply: We revised it.

13) P. 9, line 8: What is a “scale gram”?

Reply: We deleted “gram”.

14) P. 9, line 15: Delete extra spaces before “from”.

Reply: We deleted it.

15) P. 10, line 11: “TP” is not defined.

Reply: We have defined TP at P. 5, line 7.

16) P. 11, line 4: Delete extra space before “northern”.

Reply: We deleted it.

17) P. 12, line 2: Insert “it” before “fluctuated”.

Reply: Thank you very much for your suggestion. We inserted it.

18) P. 12, line 7: Change to “decreasing trend”.

Reply: We revised it.

19) P. 12, line 25: Rephrase “fluctuant trend”.

Reply: We inserted “increasing” before “trend”.

20) P. 13, line 7: Delete “variability” before “trends”.

Reply: We deleted it.

21) P. 13, line 25: Delete the space in “95%”.

Reply: We deleted it.

22) P. 14, line 23: Delete the extra space before “snow”.

Reply: We deleted it.

23) P. 15, line 2: Variations in hydrometeorological quantities such as snow depth are due to climate variability, not climate change.

Reply: Thank you very much for your suggestion. We replaced “climate change” with “climate variability”.

24) P. 15, line 7: Here reports of significant declines in snow depth are provided, while the abstract (line 27) suggests the opposite pattern is being observed – which is correct?

Reply: Here the result showed the relationship between snow depth and air temperature. There was a negative correlation between them. Increasing air temperature result in the snow depth decreased. However, in the abstract, the increasing trend represented the interannual variation in snow depth. They are different.

25) P. 15, line 18: Change to “increased”.

Reply: We revised it.

26) P. 15, line 27: Insert “is” before “not”.

Reply: We inserted it.

27) P. 16, lines 7-8: “differences” is used twice in succession.

Reply: We replaced the first “differences” with “discrepancies”

28) P. 17, line 5: Delete the extra space before “is the”.

Reply: We deleted it.

29) P. 22, line 20: This should read “Liston”.

Reply: Thank you for your suggestions. We revised it.

30) P. 26, Figure 1: The colors highlighting three regions (Sakhalin, Kamchatka Peninsula, and northern Xinjiang Autonomous region) are nearly indistinguishable. Please consider using colors of greater contrast. Why are these regions highlighted in the first place? A number of abbreviations are used on the map that are not defined in the figure caption (this is an issue in other figures as well).

Reply: We highlighted the three regions due to snow depths were greater in these areas. We wanted to indicate the accurately locations for readers who are not familiar with the geography of Eurasia. We have canceled the highlight because it may cause confusion for the reader. The country abbreviations were used because the space is limited and cannot be spelled out. We have spelled out the names of countries as an annex.

Abbreviation Description

Country	Abbreviation
Kazakhstan	KAZ
Ukraine	UKR
Turkmenistan	TKM
Uzbekistan	UZB
Tajikistan	TJK
Belarus	BLR
Estonia	EST
Georgia	GEO
Latvia	LVA
Lithuania	LTU

Azerbaijan	AZE
Kyrgyzstan	KGZ
Moldova	MDA

31) P. 27, Figure 2: Given the high number of sites with high average snow depth values in the northern reaches of the Eurasian continent, would the results be better depicted using contour lines instead? Consider adding the latitudinal averages of the snow data as secondary diagrams to these figures.

Reply: We tried to use the contour lines instead of the point values. But we found that there was a problem of the accuracy of interpolation with Kriging interpolation in AcrGIS, in which there was snow cover in some no snow areas. Therefore, the results cannot be depicted using contour lines instead. Snow depth distributions are affected by the topographic factors over the Eurasia. Snow depth is also affected by elevation, slope, aspect in the same latitude. The latitudinal average of snow data cannot fully reflect the snow depth distribution.

32) P. 30, Figure 4: Does the number of stations used in the composite snow depth anomalies vary over time? The statistical significance of the trends should use the symbol “⩽” rather than “<=”. Why does the last sentence in the figure caption mention “simulation” of snow depth?

Reply: Thank you for your suggestions. The number of stations used in the composite snow depth anomalies vary over time. First, we calculated the snow depth anomaly at each site, and then took the average of the anomalies as the general anomaly. We replaced “<=” with “⩽”. In the figure, X means the value of the linear regression trend, which is calculated by the snow depth anomaly with linear regression. Therefore, it is a calculation of snow depth.

33) P. 31, Figure 5: See comments for Figure 4.

Reply: We replaced “<=” with “⩽”.

34) P. 34, Figure 8: The statistical significance of the linear regressions should use the symbol “⩽” rather than “<=”. Are any of the stations below sea level? If not, panels (b) and (c) should have their x-axes begin at 0 m in elevation. The caption should also state that these are relationships between snow depth and latitude and elevation, not changes.

Reply: Thank you for your suggestions. We replaced “<=” with “⩽”. There are 9 stations below sea level in the former USSR. We revised the caption: “The relationships among annual mean snow depth, air temperature and snowfall for 386 stations from November through March during 1966-2009 over the USSR. The thick line is a linear regression trend.”

35) P. 35, Figure 9: See comments for Figure 8.

Reply: We replaced “<=” with “⩽”.

36) P. 36, Figure 10: See comments for Figure 8.

Reply: We replaced “<=” with “≤”.