Dear editor and two reviewers:

Thank you very much for your comments that have provided a helpful roadmap for significant improvement of this manuscript. Please thank the reviewers for their constructive comments and suggestions. We have therefore made extensive changes to the manuscript and added additional analyses to address their questions and concerns. We believe that the manuscript is now in a much improved form, but we are willing to work with you if further revisions are necessary. The following outline the major changes we made in this revision:

1) We used the latest instrumental data and CRU data (1951-2012). And therefore, all the resulting statistics were calculated for the regression model, the results were quit similar with those in the first version.

2) We added further introduction of the methods and statistical skill measures in the revision.

3) Both leave-one-out method and split-sample verification method were employed in the revision.

4) We changed the MTM spectrum analysis to Wavelet analysis according to the reviewers′ comments, and performed additional analysis on the relationship between ENSO and reconstructed temperature using moving-correlation, and,

5) We provided more discussions of relationships between tree growth and Tmin, as well as some discussions of the discrepancy between the instrumental data and estimated data in the earlier period.

Attached, please find separate answers to the reviewers'questions and comments. If you have any questions, please let us know. Thank you again for your support and we look forward to hearing from you soon.

Sincerely yours,

Yong Zhang (on behalf of all authors)

The following is the replies on the comments.

Part1: C68-replies on comments of CP

Response of Qilian juniper to climate:
The authors mention two papers that reconstruct temperature using Qilian juniper. In one study Qilian juniper ring-widths reflect December to April temperatures, in the other study ring-widths reflect previous September to April temperatures. In this study January to August temperatures are reconstructed. Qin et al. (2013) find precipitation to be limiting for growth of Qilian juniper. Could the authors explain why Qilian juniper is sensitive to so different climate variables in different places? The authors provide a short discussion in chapter 4.1. I think this manuscript would greatly benefit from further explanations.

REPLY:
The response of tree growth to climate is very much controlled by the micro growth

environment. In other words, the same species growing at different sites may respond different to various climatic factors. As mentioned in the manuscript, the reconstruction of December to April temperature (Liu et al., 2007) was developed near our study area (~ 50 km), using the ring widths and stable carbon isotope ratios (^{13}C) of Qilian juniper with 3-yr resolution, while the reconstruction of September to April (Zhu et al., 2008) was conducted in Wulan, approximately 200 km south of our sites. Both series were based on samples collected near the upper treeline. Qin et al. (2013) sampled at the elevation range of 3900-4015 m, also close to the local upper treeline, but on a very steep south-facing slope (40-65°, Table 1). According to our experience working in this region, slopes with such steep gradients in this region have thin and rocky soils that could have very low water holding capacity. Delingha is also a relatively dry site with annual precipitation of ~177 mm as compared to our site. Therefore, it is not surprising to see strong signals of precipitation or drought in the TRW series. In order to secure our samples' sensitivity to temperature signals, we screened our samples first by eliminating those from trees growing on rocks or in crevices, and then further by excluding those samples with mean sensitivity values greater than 0.45. By doing so, we have eliminated samples that potentially have strong precipitation signals and preserved those with strong temperature signals.

According to the comments, some further explanations were added in section 4.1, P14:

"The significant positive correlations between the tree-ring data and Tmean and Tmin in most months (Fig. 4) indicated that the HY chronology was temperature-sensitive. A similar climatic response has been reported for the timberline forests on the eastern and north-eastern TP (Shao and Fan, 1999; Bräuning, 2006; Liu et al., 2007; Liang, 2006, 2009). In previous studies, Tmin has been found to constrain radial growth in the TP (Gou et al., 2007; Li et al., 2011; Liang et al., 2008; Liang et al., 2010; Lv and Zhang, 2013). Even though the cambium tissues of trees are dormant in winter and early spring, the phloem sap may have freezing damages when temperatures are low during this period (Kimmins, 1987). Additionally, warmer winter Tmin protects roots from cold damages, and hence, trees may experience increased growth in the subsequent growing season (Pederson et al., 2004). On the other hand, low summer night air temperatures (Tmin) result in low soil temperatures, which may persist into the daytime, especially in the understory environments (Tranquillini 1979). The growth of roots and their function of water uptake will be limited by low summer soil temperatures at the timberline (Körner, 1999; Mayr, 2007). Tmin is known to be a critical factor affecting conifer tracheid division and enlargement by changing the timing and duration of the growing season at the timberline (Deslauriers et al. 2003, Rossi et al. 2008). Therefore, monthly minimum temperature is an important climatic and ecological indicator, not only because it is one of the directly measured meteorological variables, but also because of its potential impacts on tree growth and biological activities."

Is there a possibility to exclude changing growth limitations (seasons and from temperature to precipitation) during the reconstruction period?

REPLY: It may not be possible to entirely eliminate signals of precipitation from

the tree ring series. However, by carefully selecting the sampling sites (as high as possible in the study region) and screening our samples as described above, we have maximized the potential for temperature reconstruction, as evidenced by stronger correlations to temperature variables that those to precipitation variables (Fig. 4).

Figs 3 and 7(a): The period 1925-1940 looks very interesting. The ring-width chronology experiences a massive trend and inter-annual variability seems rather low (as far as this is visible from the figure). Is this trend as well found in other tree ring series? Why is the inter-annual variability so low? Could the authors expand on this very interesting period?

REPLY: The warming trend during 1925-1945 is visible in the Wulan reconstruction (Fig. 9c), although not as dramatic, Dunde ice core, and the Jones et al. (1998) series. The interannual variability during this period may not as low as it appears because it was dominated by an upward trend. Similar warming trends between 1925-1945 also appeared in two summer temperature reconstructions, one for southeastern Tibetan Plateau (Zhu et al. Palaeogeography, Palaeoclimatology, Palaeoecology, 2011) and another for the Asian region (Cook et al. Clim. Dyn,, 2012). The early-20th-century warming (1925-1945) was also found in the Arctic region (IPCC AR4, 2007, Bengtsson et al. 2004). We agree that this is something very interesting. However, a detailed analysis on the spatial structure and exploring the physical mechanisms of this event is outside the scope of our study and requires further work. We would like to take on this task in the future.

Bengtsson, Lennart, Vladimir A. Semenov, Ola M. Johannessen, 2004: The Early Twentieth-Century Warming in the Arctic-A Possible Mechanism. J. Climate, 17, 4045–4057. doi: http://dx.doi.org/10.1175/1520-0442(2004)017<4045:TETWIT>2.0.CO;2

Calibration issues:
To me it remained unclear what (beside the highest r value) the motivation for reconstructing over an eight month period is. In this eight month period May and March are included, two months that seem (partly) precipitation sensitive. The influence of JJA temperatures on ring-width seems almost as strong as the influence of January through August temperatures. Could the authors as well calculate partial correlations for T JJA when the influence of T JFM has been partialled out, and vice versa?

REPLY: While it is true that January-August minimum temperature rendered a higher r value in correlation analysis, we consider this as the evidence of significance of minimum temperature to tree growth in this region. This parameter covers the current growing season as well the months to the prior. In section 4.1 of the revision we explained the potential ecological impact of January-August minimum temperature on tree growth, including freezing injuries in winter and early spring and slowed growth during the growing season (p.14).
We calculated partial correlations of T JJA with tree ring index when the influence of T JFM has been partialled out, the correlation coefficient of JJA changed from 0.65

to 0.5(p<0.05) when considering the influence of T JFM, and vice versa, the correlation coefficient of JFM changed from 0.58 to 0.52(p<0.05). This means that neither T JJA nor T JFM alone can fully represent the information contained in the tree ring series. No significant influence was found between T JJA and T JFM. Therefore, we concluded that the January-August minimum temperature is a valid climatic factor of tree growth.

P 351 lines 3-6: In an earlier section the authors mention the short calibration period that makes a split period approach impossible. In this section the authors are, in fact, splitting the calibration period in a period 1960 – 1984 and 1985 – 2011. For univariate (ordinary least squares) regression, regression parameters (a and b) and Pearson's r are directly related (b = r * sy / sx, and a = ymean– b * xmean). Therefore, I think the authors should decide whether a split period approach is possible or not. If they decide it is possible, they should use classical RE and CE statistics and as well give the correlation coefficient and p-value for the period 1960 – 1984.

REPLY: Many thanks to the valuable comments. In the revision, we performed the split-period verification (see section 3.3, p. 12). We used the period 1985-2012 to build the regression model and the periods 1960-1984 as the independent verification period (Table 3). The model itself showed good quality, with a R value of 0.845 and relatively high RE of 0.619. The overall sign test (ST) passed the 99% confidence level. However, for the verification period 1960-1984, the R was low (0.213) and CE was negative. Additionally, the sign test for the first-order difference (FST) did not pass the 95% confidence level, suggesting poor predicting power in the high-frequency domain. From Fig. 6b it can be seen that the reconstructed values did not match the observed values well during 1975-1984. After eliminating the data during 1975-1984, there were slightly improved verification results (r increased from 0.213 to 0.219, RE increased from 0.619 to 0.679, and CE became less negative), but the result of FST still did not reach the 95% confidence level. Considering the overall quality of the model (Table 3), we concluded that our model is reliable, especially in the low-frequency domain.

Relatively poor matches between tree ring data and meteorological data prior to the 1990s seem to be a common problem for many previous studies in China (see section 4.2.1.). Therefore, we attempted to further validate our reconstruction by comparisons with other reconstructions, including both those in the region and those for the NH (see section 4.2.3).

"It can be seen that the reconstructed values matched the observed values particularly well after 1985, while the agreement between the observed and reconstructed series was not as good before 1985, especially during 1975-1984. For this reason, we performed split verifications using data for 1960-1984 and 1960-1974, for the regression model obtained from the independent calibration period of 1985-2002 (Table 3). While the model itself had a respectful r value of 0.845, the verification period of 1960-1984 produced a relatively high reduction of error (RE) at 0.619, and the overall sign test (ST) passed the 99% confidence level. These all indicate the overall reliability of the model. However, the verification r was low (0.213) for 1960-1984 and the coefficient of efficiency (CE) was negative.

Additionally, the sign test for the first-order difference (FST) did not pass the 95% confidence level, suggesting poor predicting power of the model in the high-frequency domain. After eliminating the data during 1975-1984, there were slightly improved verification results (r increased from 0.213 to 0.219, RE increased from 0.619 to 0.679, and CE became less negative), but the result of FST still did not reach the 95% confidence level. ……Results of the split verification suggest that the reconstructed temperature series should be more reliable in reflecting low-frequency (multi-year and longer time scales) variation pattern"

P 351 lines 7-8: The authors mention a discrepancy between tree-ring data and meteorological data between 1960 and 1984. This discrepancy might indeed be caused by lower quality of meteorological data prior to 1985. A possibility of testing this hypothesis is to compare the (temperature) data of the two stations prior to 1985. Is the correlation between the two stations lower prior to 1985? Liu et al. (2007) don't find significant inhomogeneities among five stations including Zhangye. According to Liu et al. (2007) the Zhangye record starts in 1951 (with some missing values up to 1953). Is the data quality so low that the authors decided to omit these data?

REPLY: the correlation of annual Tmean between Zhangye and Yeniugou stations was high with r=0.8 for the period of 1960-2013, but the correlation coefficients for the early period (r=0.36, 1960-1984) and later period (r=0.81, 1985-2009) were quite different. Apparently, these discrepancies could have been caused by the uncertainties in the instrumental data for the earlier years, which were resulted from inconsistent observation practices, poor quality control, and location and instrumentation changes of numerous meteorological stations before 1990s. We added further discussion in section 4.2.1, page 15-16:

"Figure 6b showed discrepancies between the observed and predicted values by the regression model, especially during the period of 1975-1984. Such discrepancies seem to be a common problem in tree-ring based temperature reconstructions in China, in which the agreement between observed data and estimated data became worse before the 1990s in many regions across China (Deng et al., 2013; Fan et al., 2010; Fan et al., 2009; Gou et al., 2007; Lv and Zhang, 2013; Shi et al., 2010; Song et al., 2014; Yang et al., 2009, 2010; Zhang et al., 2013; Zhu et al., 2011). Similar issues were also found for the tree-ring based precipitation reconstructions (Chen et al., 2011; Fan et al., 2008; Liu et al., 2011; Shao et al., 2005; Yang et al., 2011, 2014; Zhang et al., 2011; Zhang et al., 2014). In addition, the correlation coefficient of annual Tmean between Zhangye and Yeniugou stations for the period of 1960-1984 is 0.36, while for the period of 1985-2009, the correlation coefficient between two stations is 0.81. Apparently, these discrepancies could have been caused by the uncertainties in the instrumental data for the earlier years, which were resulted from inconsistent observation practices, poor quality control, and location and instrumentation changes of numerous meteorological stations before 1990s (Ren et al., 2012). Data in more recent years have gone through rigorous quality assessment/quality control procedures and the uncertainties in data have been significantly reduced (Ren et al., 2007)."

We decided that the two-station averages are a better representation of the regional climatic condition for our reconstruction (p. 10). Therefore, we only used the common period of the two stations 1960-2012 in the relevant analyses.

P 351 lines 15- 20: 'Regardless the issues in the earlier part of the calibration period, the evaluative statistics in Table 3 indicated that our regression model was stable and reliable, and was acceptable to reconstruct the annual-to-centennial variability...' I do not think a leave-one-out approach is sufficient to substantiate this claim. Perhaps the authors could leave out 10 or even 15 consecutive years. Good performance under slightly less favourable conditions would increase the credibility of the calibration model. Unfortunately, the ring-width chronology is at its high end between 1960 and 2011. Therefore the reconstructions for most of the period before 1900 are at or beyond the end of the calibration space. As mentioned by the authors temperature variations are less nicely depicted between 1960 and 1985. Considering the entire tree-ring chronology the values for the period 1985 – 2011 are all in the fourth quartile.

Reply: In the revision, both leave-one-out and split-sample calibration-verification methods were used. Results of the split verification suggest that the reconstructed temperature series should be more reliable in reflecting low-frequency (multi-year and longer time scales) variation patterns. Comparisons with other temperature series from neighbouring regions and for the Northern Hemisphere also supported the validity of our reconstruction. We agree that our TRW data during 1960-2012 were much higher than the rest of the series, which lead to uncertainties in the lower range of the reconstructed temperature anomaly values. This also points to more work being needed in the future.

Page 350 line 5 why is the correlation calculated for 1962 – 2011 and not for 1960 – 2011?

Reply: because the SEASCORR program calculates the correlation for the previous year, and it is in default in the program.

MTM:
P 352, lines 10 -14. How is the massive trend (non-stationarity) in the 20th century affecting the MTM spectrum? Could the authors run the spectral analysis for data from AD 700 to AD 1900?
I am not an expert on MTM and significance tests but I nevertheless have a question on this topic: Are the significance levels indicated valid for one single test i.e. if the scientist is interested in the significance of the 11-year band, or are these levels accounting for multiple testing since all frequency bands are tested for significance?
The comparison to other reconstructions and forcings isn't fully satisfactory. Many frequency bands > 10-year are significant. Since solar cycles have very wide frequency bands (DeVries–Suess cycle 170–260 year), it would be surprising if none of the significant frequencies were in the solar bands. Additionally, the use of wavelet coherence analysis is possibly a more straight forward approach for comparisons

between reconstructions and forcings.

REPLY: according to the comments of reviewers, we used the wavelet analysis in the revision instead of the MTM spectrum analysis. The resulting cycles were similar with those obtained by the MTM spectrum analysis, but are more reasonable.

The following is the result of Wavelet analysis (Fig.7b in the revision, and section 3.3 in page 13-14):

"The results of wavelet analysis revealed the persistence of high-frequency patterns of approximately 2-4 years (mainly existed during AD 1000-1600) and low-frequency centennial-scale patterns of approximately 90-170 years (mainly existed during AD 1350-1700) and 40-50 years (existed during AD 900-1000) (Fig. 7b)."

We also gave a further discussion for the relationship between temperature reconstruction and solar activity in section 4.3, page 21-22.

"It is well known that solar irradiance and volcanism are the important forcing factors of global temperature variations (Crowley, 2000; Jones and Mann, 2004). The centennial cycles identified in our study were possibly associated with the frequencies of solar variations (Stuiver and Braziunas, 1989; Hoyt and Schatten, 1997; Raspopov et al., 2008). The cycles similar to the periodicities of 90-170 years have been found in other tree ring reconstructions in China (Wang et al., 2008; Gou et al., 2010; Zhang et al., 2011). These significant low-frequency cycles were prominent in the periods of AD 1350-1650 and before AD 1150, corresponding to the Sporer minimum (AD 1460-1550) and Oort minimum (AD 1040-1080) periods of low solar activities. Although the low-frequency signal was depressed since AD 1650, low temperatures during the LIA should be linked to the Maunder Minimum of solar activity (Shindell et al., 2001)."
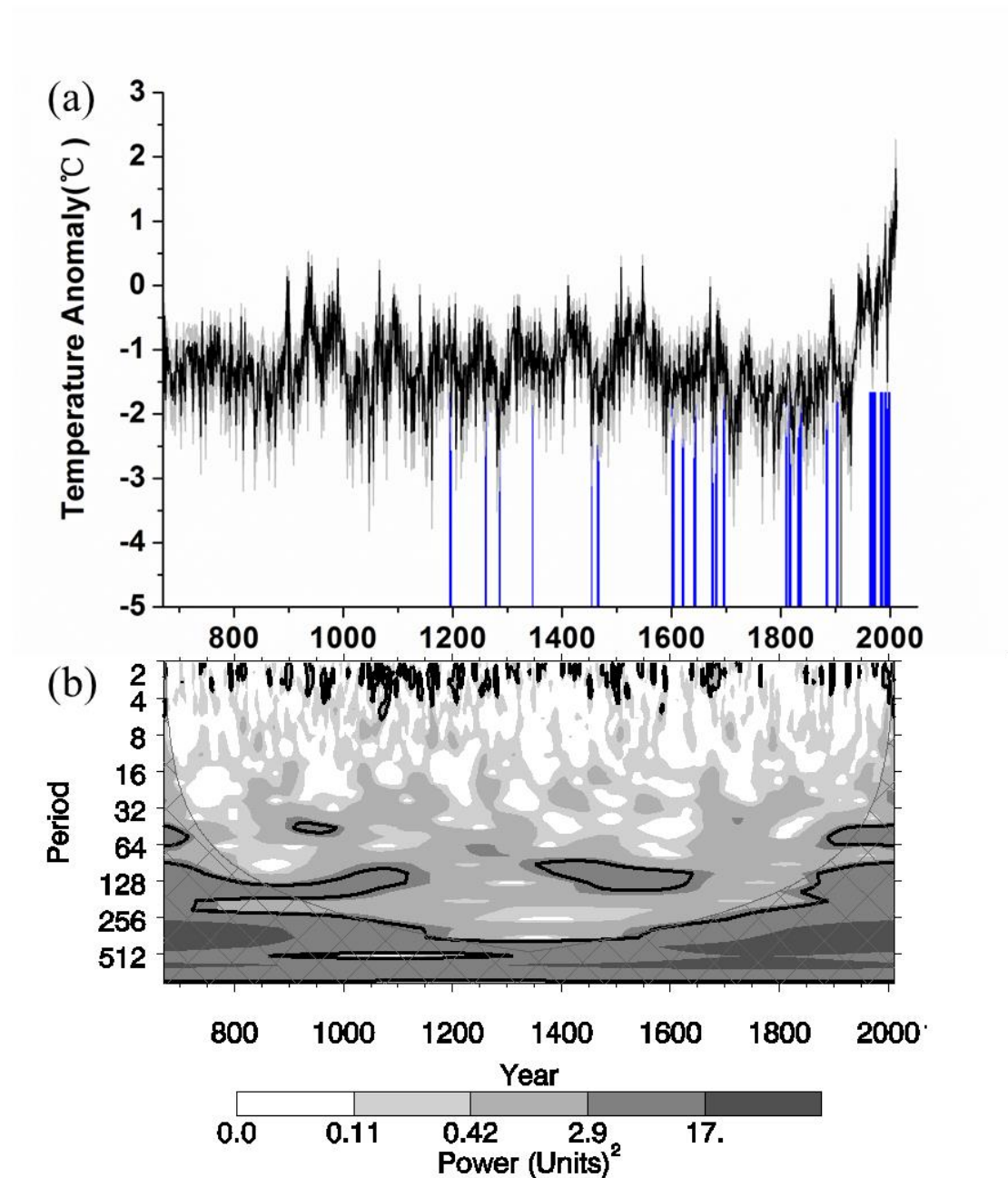
Fig. 7 (a), our reconstructed temperature in the central Qilian Mountains and 95 % confidence level (grey bars), the black dark lines indicate the 31-year running mean of our reconstruction. The vertical blue lines indicate 21 volcanic events during AD 1000-1999. (b), the wavelet power spectrum of our temperature reconstruction series. Cross-hatched regions represent the cone of influence where zero-padding of the data was used to reduce variance using a Morlet wavelet. Black contours indicate significant modes of variance with a 5% significance level using an autoregressive lag-1 red-noise background spectrum (Torrence and Compo, 1998).

Part 2: C68- replies on comments of CP

My main general concern is the focus on the minimum temperature. It is mentioned in a few places that correlations between tree-ring width (TRW) and mean temperature are nearly as high as correlations between TRWand minimum temperature. Is there a tree physiological reason why growth should be related to a monthly temperature minimum that possibly covers just a very short period of the entire month? Or are minimum and mean temperature just very highly correlated and minimum temperature shows coincidently slightly higher correlations?

Reply: In most meteorological observations of temperature, daily minimum and maximum temperature are directly measured. Daily mean temperature is calculated either by averaging the maximum and minimum values or multiple measurements (4-6 or more) during the day. Then the monthly mean values are obtained by averaging the daily values. The daily minimum temperatures therefore represents the coldest time of the day and have significant ecological meanings regarding tree growth, rather than a single incidence of the lowest temperature during a month. In section 4.1 of the revision (p. 14, lines 6-11) we discussed the significance and previous cases of reconstructions of minimum temperatures in different regions. In contrary, monthly mean temperature is a combination of both maximum and minimum temperature, and because of this combined effect, it is also correlated with the TRW chronology, but not as high as the minimum temperature. The underlying assumption of using correlation analysis and response function to search for the limiting factor of tree growth is that tree growth at an optimal dendroclimatic sampling site should respond to the variation of a single climatic factor whose effect is reflected by the ring widths and, therefore, the climatic factor that has the strongest influence on tree growth should have the strongest correlation. This is considered a standard analysis method in dendroclimatological studies (Fritts, 1976; Blasing et al., 1984). The challenge, of course, is that in some cases multiple climatic factors may be at work influencing tree growth at a given dendroclimatic study site. That is why we had to consider multiple factors and using correlation analysis as a screening tool. In our case, the higher correlation of mean minimum temperature with the TRW series also helped establishment of a better regression model in reconstruction.
(Ref.: Blasing TJ, Solomon AM, Duvick DN. 1984. Response functions revisited. Tree-Ring Bulletin 44: 1-15.
Fritts, H. C.: Tree Rings and Climate, London, Academic Press, 1976.)

Also, why should temperatures from January to April influence growth? Maybe due to extended snow cover? The choice made in this manuscript seems to be just a statistical hunt for highest correlation and I miss a sound scientific argumentation for the choices that have been made. It is mentioned only once in section 3.3 that the authors reconstruct "monthly minimum temperature anomalies, averaged over the months January to August". I found the mostly used term "mean minimum temperature" not very clear.

Reply: Because of the long seasonal coverage of our temperature reconstruction, we

considered two impacts of minimum temperature on tree growth. One is freezing injuries during cold winter and spring which produces a potential influence for coming tree growth, while the other is the slowed biological activities during a cool growing season. We describe these possible causes in section 4.1 (p.14, lines 11-20) of the revision. We also described more clearly why and how the anomalies of the climatic variables (both temperature and precipitation) were used to calculate regional anomaly series on p. 7 (lines 14-20). The definitions of monthly mean minimum and maximum temperatures are now explicitly defined on p. 7.

Section 4.1., p14: "A similar climatic response has been reported for the timberline forests on the eastern and north-eastern TP (Shao and Fan, 1999; Bräuning, 2006; Liu et al., 2007; Liang, 2006, 2009). In previous studies, Tmin has been found to constrain radial growth in the TP (Gou et al., 2007; Li et al., 2011; Liang et al., 2008; Liang et al., 2010; Lv and Zhang, 2013). Even though the cambium tissues of trees are dormant in winter and early spring, the phloem sap may have freezing damages when temperatures are low during this period (Kimmins, 1987). Additionally, warmer winter Tmin protects roots from cold damages, and hence, trees may experience increased growth in the subsequent growing season (Pederson et al., 2004). On the other hand, low summer night air temperatures (Tmin) result in low soil temperatures, which may persist into the daytime, especially in the understory environments (Tranquillini 1979). The growth of roots and their function of water uptake will be limited by low summer soil temperatures at the timberline (Körner, 1999; Mayr, 2007). Tmin is known to be a critical factor affecting conifer tracheid division and enlargement by changing the timing and duration of the growing season at the timberline (Deslauriers et al. 2003, Rossi et al. 2008). Therefore, monthly minimum temperature is an important climatic and ecological indicator, not only because it is one of the directly measured meteorological variables, but also because of its potential impacts on tree growth and biological activities."

Finally in the comparison with other records, the minimum temperature reconstruction is discussed as if it would be exactly the same as a mean temperature reconstruction. Therefore, I would suggest to either focus on the mean temperature reconstruction instead of the minimum or at least show more results for how minimum and mean temperatures are related and then explain a possible mechanism why trees should respond to one temporally short monthly minimum or an average of a few monthly minima rather than average temperatures during the growing season.

Reply: As we stated above, monthly mean minimum temperature is an important parameter of climate with significant ecological meanings for tree growth. It would have been optimal if we could compare our reconstruction with similar constructions of minimum temperature in different regions. However, there are only a few reconstructions of minimum temperature (e.g., Shao and Fan, 1999; Liang et al. 2008; Lv and Zhang, 2013) and none of them covered the millennial timespan. Because of the high correlation between the minimum and mean temperatures within the range of 0.7-0.97 for the 12 months of the year, we proceeded to compare our reconstruction to

other mean temperature reconstructions in the region and of hemispheric scales. Our results suggest that our reconstruction is valid as indicated by similarities with the others, while it also offers additional information by its differences with the other reconstructions (section 4.2.3, p16-19).

My second general concern is in regard to the spectral analysis. The strongly autocorrelated TRW is expect to have a red spectrum. Much more details about the applied method are needed, for instance with regard to how the red spectrum is taken into account in the decision which frequencies show significant peaks. I would be very careful trying the discuss a 500-yr cycle. Data treatment such as detrending might have influenced the low-frequency variability of the reconstruction, especially if chronologies have a length of around 500 years. You should also check how the 20th century temperature trend influences the spectral characteristics.

Reply: Considering the reviewers' comments, we employed the wavelet analysis instead of MTM analysis. The 500-yr cycle is less certain because of the boundary effect, so we didn't discuss it in the revision. Our methods to remove the growth trend should preserve low frequency signals at the decadal to centennial timescales. Previous studies on variations at the decadal to centennial timescale variations have used the same methods as we did (Scuderi, 1993; Briffa et al., 1995). In the first version, we used the mtm10c_XP.exe program downloaded from Lamount Doherty Tree-Ring Laboratory, which is widely used in tree ring research (Li et al., 2011; Fang et al., 2010; Gray et al., 2004). In this program, we used the red-noise spectrum with F-test. The result of wavelet analysis is presented at the end of section 3.3 (p. 13) and further discussed in section 4.3(p.20-21).

"The results of wavelet analysis revealed the persistence of high-frequency patterns of approximately 2-4 years (mainly existed during AD 1000-1600) and low-frequency centennial-scale patterns of approximately 90-170 years (mainly existed during AD 1350-1700) and 40-50 years (existed during AD 900-1000) (Fig. 7b)."

"It is well known that solar irradiance and volcanism are the important forcing factors of global temperature variations (Crowley, 2000; Jones and Mann, 2004). The centennial cycles identified in our study were possibly associated with the frequencies of solar variations (Stuiver and Braziunas, 1989; Hoyt and Schatten, 1997; Raspopov et al., 2008). The cycles similar to the periodicities of 90-170 years have been found in other tree ring reconstructions in China (Wang et al., 2008; Gou et al., 2010; Zhang et al., 2011). These significant low-frequency cycles were prominent in the periods of AD 1350-1650 and before AD 1150, corresponding to the Sporer minimum (AD 1460-1550) and Oort minimum (AD 1040-1080) periods of low solar activities. Although the low-frequency signal was depressed since AD 1650, low temperatures during the LIA should be linked to the Maunder Minimum of solar activity (Shindell et al., 2001)."
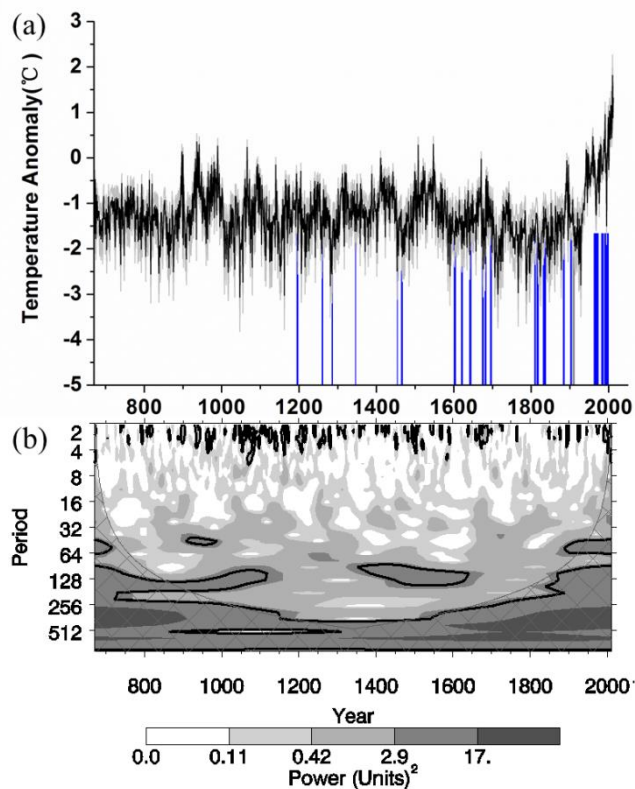
Fig. 7 (a), our reconstructed temperature in the central Qilian Mountains and 95 % confidence level (grey bars), the black dark lines indicate the 31-year running mean of our reconstruction. The vertical blue lines indicate 21 volcanic events during AD 1000-1999. (b), the wavelet power spectrum of our temperature reconstruction series. Cross-hatched regions represent the cone of influence where zero-padding of the data was used to reduce variance using a Morlet wavelet. Black contours indicate significant modes of variance with a 5% significance level using an autoregressive lag-1 red-noise background spectrum (Torrence and Compo, 1998).

(Ref: Briffa, K. R., Jones, P. D., Schweingruber, F. H., Shiyatov, S. G., and Cook, E. R.: Unusual twentieth-century summer warmth in a 1,000-year temperature record from Siberia, Nature, 376, 156-159, 1995.

Li, J., Xie, S.-P., Cook, E. R., Huang, G., D'Arrigo, R., Liu, F., Ma, J., and Zheng, X.-T.: Interdecadal modulation of El Niño amplitude during the past millennium, Nature Climate Change, 1, 114-118, 2011.

Fang, K., Gou, X., Chen, F., Li, J., D'Arrigo, R., Cook, E., Yang, T., and Davi, N.: Reconstructed droughts for the southeastern Tibetan Plateau over the past 568 years and its linkages to the Pacific and Atlantic Ocean climate variability, Climate Dynamics, 35, 577-585, 2010.

Gray, S. T., Graumlich, L. J., Betancourt, J. L., and Pederson, G. T.: A tree-ring based reconstruction of the Atlantic Multidecadal Oscillation since 1567 AD, Geophysical Research Letters, 31, -, 2004.

Scuderi, L. A.: A 2000-year tree ring record of annual temperatures in the Sierra Nevada Mountains, Science, 259, 1433-1435, 1993.)

Finally, not all readers of "Climate of the Past" will be experts in tree-ring research. Thus, I strongly suggest to explain the methods and statistical skill measures in more detail, e.g.what is a "response function", "subsample signal strength", "Rbar", "EPS", etc. and how are these interpreted.

Reply: thank you for your advice, the explanations, such as "subsample signal strength", "Rbar", "EPS", "SSS", have been added in the revision (see section 2.2, page 6-7). but not all of them, for example the "EPS", because response function analysis cannot express completely with a few words, we therefore cited a reference (Fritts, 1976) in the revision. However, the following is the detailed information of this method:

Response Function analysis (Fritts et al., 1971; 1976) is a multiple regression technique using the principal components (PCs) of monthly climatic data to estimate indexed values of ring-width growth. The regression coefficients are then multiplied by the PCs of climate to obtain a new set of regression coefficients related to the original monthly climatic data variables. The technique has become widely used to estimate how ring -width growth (portrayed by a particular standardized chronology) responds to variations in monthly climatic conditions over the span of a year or more.

Page 342
Line 6: "pre- and current growing season" You just analyze the previous year starting
in October when the growing season was probably already finished!

Reply: maybe there are some misunderstandings, this sentence is now changed to "prior to and during the growing season" in Page 1
.
Line 6: "with minimum temperature" do you mean daily or monthly or yearly or : : :minima?

Reply: it changed to "with monthly minimum temperature" (page 1)

Line 18: "28.8-66.2, 113.6-169.5" wide ranges of significant peaks, see also general comments on spectral analysis above.

Reply: because the content has been modified largely, these sentences have been deleted.

Line 21: Last sentence is very vague

Reply: this sentence has been change to "The Comparison between the reconstructed temperature and the index of tropical volcanic radiative forcing indicated that some cold events recorded by tree ring may be due to the impact of tropical volcanic eruptions."(page 1)

Page 344
Line 5: "different with those" -> "different from" or "in comparison to"

Reply: it was changed to "different from". (page 3)

Line 11" Why was it controversial if samples where temperature or precipitation sensitive?

Reply: because these archaeological samples were obtained from tombs, there is less information about the sources of these ancient woods. Therefore, some researchers think they are temperature sensitive (Liu et al. 2009), while others found they are moisture sensitive (Shao et al. 2010).

Line 14 ff: "Whether : : :" TRW reconstructions, which have to be detrended and otherwise statistically treated, have been shown to often not properly capture the full spectrum of climate variability. Will they be able to answer that question here?

Reply: a modified negative exponential curve or linear regression with negative or zero slope has proven to be efficient to preserve the inter-annual variations (Cook et al., 1990). Therefore our reconstruction should at least capture the inter-annual variations. Additionally, the mean segment length of the chronology we developed is 516 yrs, which suggests that our reconstruction is able to capture decadal- to at least partially centennial-scale variations.

  (Ref.: Cook, E., and L. Kairiukstis (1990), Methods of Dendrochronology, Springer, New York.)

Line 27: "We hope" I would use another expression

Reply: it changed to "This reconstruction should improve our understanding of temperature variability in the north-eastern TP for the past millennium."

Page 345
Line 14: -1.5 to – 0.7
Line 23: "healthy" -> "living" and maybe say a word about the age distribution

Reply: these words have been modified. Age distribution of the sample cores used in the chronology can be inferred from Fig. 3b.

Page 346
here you say the mean correlation between all cores would be 0.6. In Table 2 you write 0.312. What is correct?

  Reply: the first correlation of 0.6 is the result using the Cofecha program with 32-yrs spline for filtering (excluding low-frequency signals in raw measurements), where all the samples were included for crossdating. The second correlation of 0.312 is the result using Arstan program with different detrended curves (including low-frequency signals), where some higher sensitive series have been excluded for developing the chronology. This information is now added in the revised manuscript (p. 5).

"After measuring each ring to the nearest 0.01 mm, we statistically verified the cross-dating accuracy using the program COFECHA (Holmes, 1983). This program firstly removes the low-frequency variations from the raw data of ring width by fitting cubic smooth spline, and then provides the correlation test for succeeding 50-year segments with 25-year overlapping periods." And "the mean correlation of all cores in COFECHA was 0.6".

Line 20: How can a growth trend be corrected with a horizontal line? In general say why many different function had to be used to remove the growth trend and how you have chosen the functions.

Reply: For those cores that stared far away from the piths, there are no clearly defined declining growth trends. This is especially true for Qilian junipers older than 700 years. Therefore, both negative exponential curve and declining linear trend do not work well to represent the "growth trend". For these cores, we used a horizontal line passing the mean to represent the growth trend in obtaining the ring index with mean of 1.0. This information is clarified on line 12-16, p. 6 of the revised manuscript. Below are examples of growth trend removal using a) declining linear function, b) negative exponential curve, and c) a horizontal line passing the mean, with the width measurements before growth trend removal and the width indices afterward.。
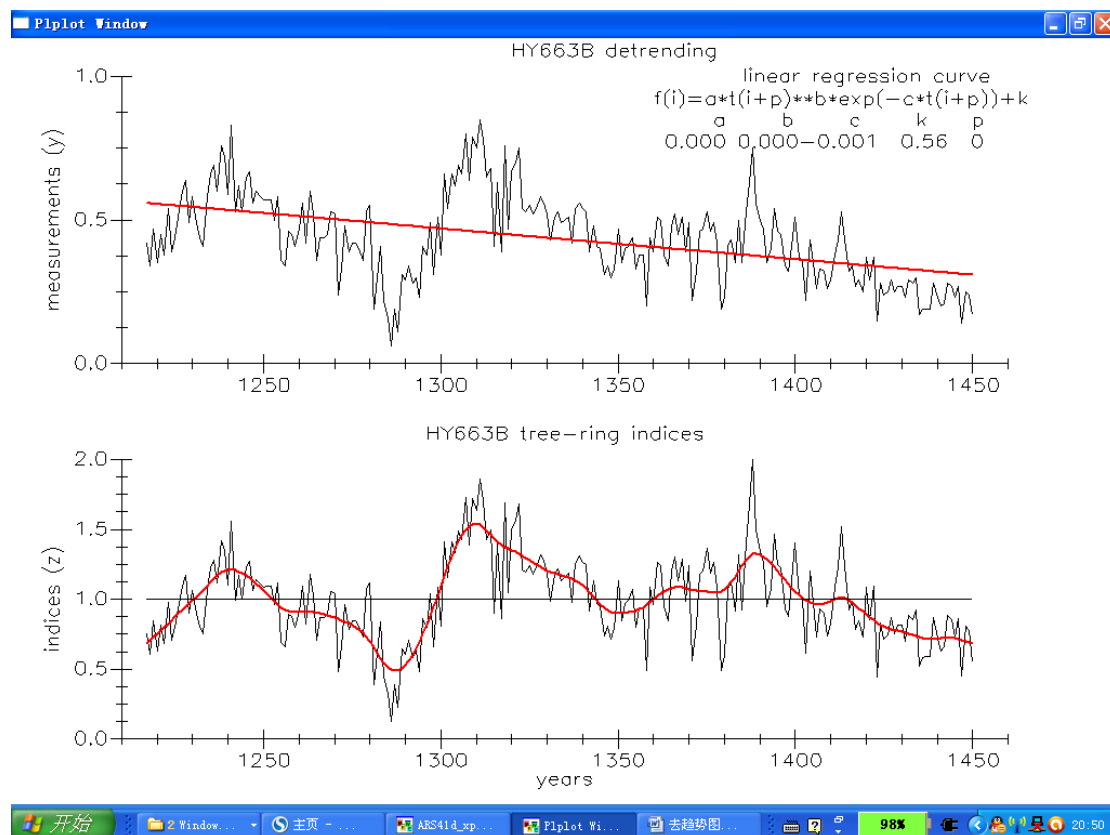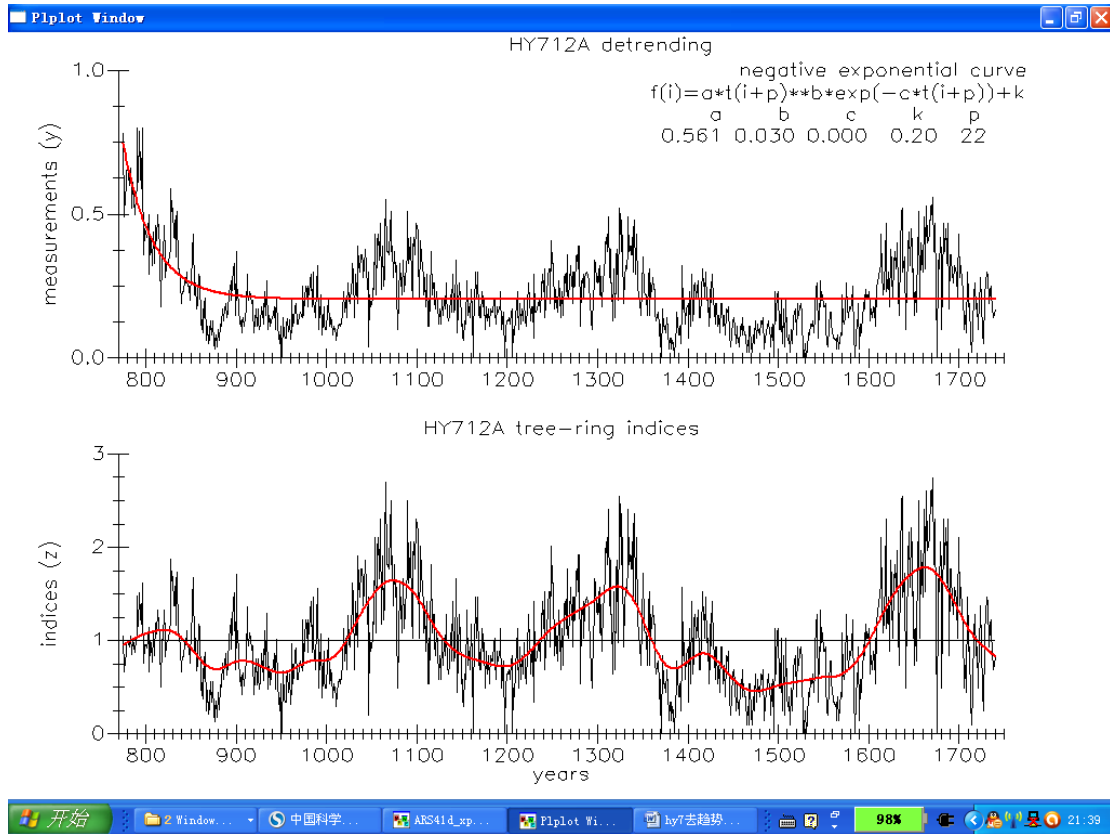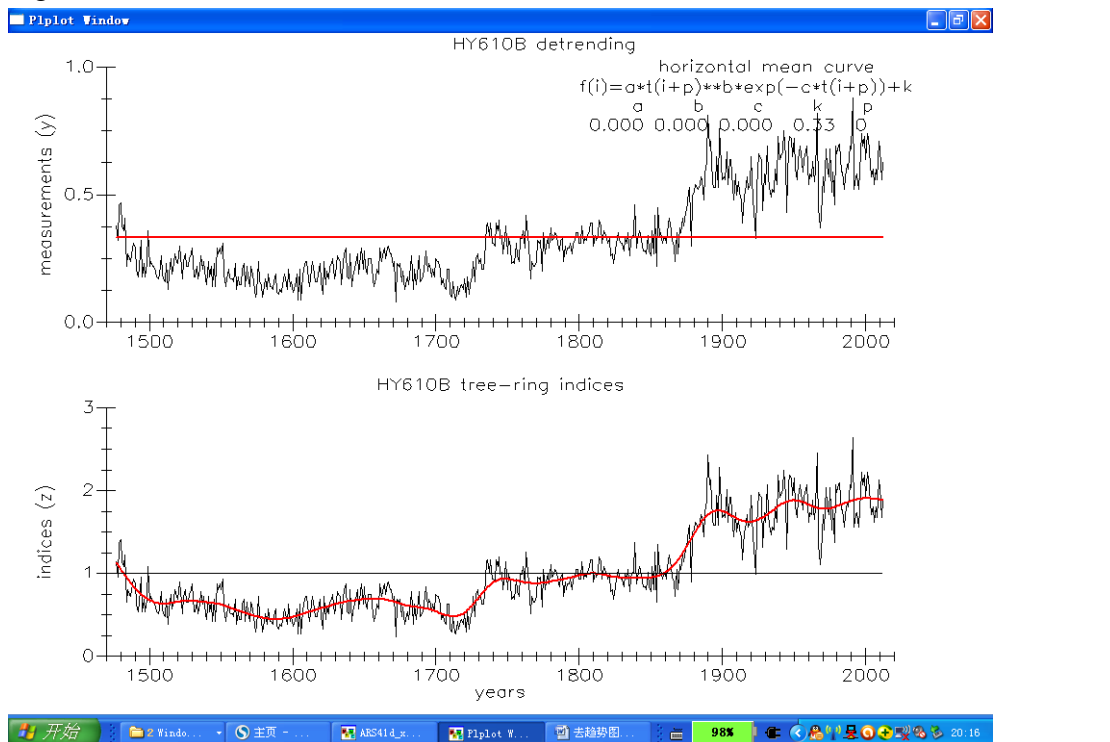

Fig a

Fig. b

Fig. c
(top is measurement, bottom is detrended curve)

Page: 347
Line 6: Maybe add a note on how these stations on very different elevations compare and why you also choose the station at so much lower elevation than the tree-ring site.

Reply: In the revision, we add more detailed information and explanation in the data and method section, see section 2.3, p. 7.
"Compared to the climate variables at a single station, the 2-station means of the climate variable have better spatial representation. Therefore we also calculated the average values of the monthly climate variable between the two stations for the common period of 1960-2012. Because of the distinct difference in elevation between Zhangye and Yeniugou stations, the 2-station means of monthly temperature and precipitation were calculated using the monthly temperature anomalies and precipitation anomalies, respectively. In other words, departures from the 1960-2012 means were first calculated for each station, and then averaged for the two stations to obtain the anomaly series."

Line 11: Is there no way to obtain data from the closed station if there is one existing?

Reply: As noted in the manuscript, there is a closer station at Sunan (p. 7-8). However, this station is not among the 700+ stations in the database we used (China Surface Climate Data (Monthly) Dataset, http://cdc.cma.gov.cn/dataSetDetailed.do?changeFlag=detai&titleName=&dsId=SURF_CLI_CHN_MUL_MON&dsid=&TITLE=&DSID=SURF_CLI_CHN_MUL_MON&LINK=&LB=sjfljs&tpcat=SURF&keyWord=SURF_CLI_CHN_MUL_MON).
Also, its data ended in 2010 and we did not have access to its most updated data.

Page 348
Line 2: "Compare : : :" sentence is not clear.
Line 4: "subset of 2-station means" do you mean a subset of certain months?

Reply: sorry for the unclear sentences, we have changed them in the revision (see section 2.3, p. 7).

Line 15: "is derived"
Line 19: "covers"
Line 22: "mean" before you said it would be the "median"
Page 349
Line 16 to 20: delete, just repetition of previous paragraph (REPLY: sorry, this paragraph presented the results of response function analysis for single station, which was different with previous paragraph presenting the results of correlation analysis, therefore, we didn't change it.)
Line 21: "correlation"
Line 21: rephrase, not clear that you generate a mean of the two stations for all temperature
variables. (Reply: change to "The results of correlation and response-function

analyses of the 2-station means of temperature variables with tree-ring index were very similar with those in single station")
Line 23: What is "mean Tmin"
Line 26: Why?

Reply: these words above have been modified in the revision. "mean Tmin" changed to "two-station average Tmin" when applicable.

Page 350
Line 2: Is Tmean also significant?

Reply: Tmean is also significant in different months and multi-month periods; however, Tmean can't pass the test for both leave-one-out and split-period verifications.

Page 351
Line 8: Do you have any indication for issues with early instrumental data?

Reply: Because of inconsistent observation practices, poor quality control, and location and instrumentation changes of numerous meteorological stations before 1990s, the quality of the earlier instrumental data was poor comparing to the later instrumental data. In the revision, we add more discussion on the issues with the earlier instrumental data. See section 4.2.1, page 15-16.
"4.2.1 Comparison between instrumental and estimated data
Figure 6b showed discrepancies between the observed and predicted values by the regression model, especially during the period of 1975-1984. Such discrepancies seem to be a common problem in tree-ring based temperature reconstructions in China, in which the agreement between observed data and estimated data became worse before the 1990s in many regions across China (Deng et al., 2013; Fan et al., 2010; Fan et al., 2009; Gou et al., 2007; Lv and Zhang, 2013; Shi et al., 2010; Song et al., 2014; Yang et al., 2009, 2010; Zhang et al., 2013; Zhu et al., 2011). Similar issues were also found for the tree-ring based precipitation reconstructions (Chen et al., 2011; Fan et al., 2008; Liu et al., 2011; Shao et al., 2005; Yang et al., 2011, 2014; Zhang et al., 2011; Zhang et al., 2014). In addition, the correlation coefficient of annual Tmean between Zhangye and Yeniugou stations for the period of 1960-1984 is 0.36, while for the period of 1985-2009, the correlation coefficient between two stations is 0.81. Apparently, these discrepancies could have been caused by the uncertainties in the instrumental data for the earlier years, which were resulted from inconsistent observation practices, poor quality control, and location and instrumentation changes of numerous meteorological stations before 1990s (Ren et al., 2012). Data in more recent years have gone through rigorous quality assessment/quality control procedures and the uncertainties in data have been significantly reduced (Ren et al., 2007)."

Page 354
Line 8: Has Qilian data been degraded to 3-yr averages before calculating correlations?

Reply: Yes. We first calculated the corresponding 3-year averages of our reconstruction and then correlated it with the Sidalong reconstruction.

Line 27: "inner-year" -> "intra-annual" ?

Reply: yes, it is modified.

Page 356
Line 12 ff: Why do you discuss the ENSO cycle extensively if you do not find any correlation between ENSO reconstructions and your temperature reconstructions but you do not discuss cycles in thermal contrasts, westerlies, etc. although you briefly mention them as possible causes of the observed frequencies?
Line 26: or it means that ENSO is not responsible for the 2-yr frequency!

Reply: In the first version, the results of correlation analysis didn't show a significant relationship between tree ring index and ENSO, but with further analysis we found that there might be a linkage that varied over time. In the revision, we found the relationship between the ENSO and temperature varied with time using 50-yr moving correlation, some significant correlations were found during several periods, especially when the 2-4 yr cycles were more prominent (please see section 4.3, page 19-20). As to the intensities of westerly circulation and land-ocean thermal contrast, the greatest difficulty is the lack of millennial-length high-resolution reconstructions.
"while significant correlations between ENSO index and reconstructed temperature were found in certain periods using 50-year moving-correlation. A series of significant positive correlations ($p<0.01$) were found in approximately AD 1340-1410, 1553-1631, and 1798-1869, with the highest correlation coefficient of 0.32 in AD 1355-1404, and meanwhile, continuous significant negative correlations ($p<0.01$) were found in AD 1060-1130, 1591-1656, and 1840-1927, with the strongest correlation coefficient of -0.33 during AD 1849-1898."

Line 16: " has been" -> "can be identified"
Page 357
Line 15: Do you mean the 11-yr sun spot cycle when you talk about the 22-yr cycle?

Reply: It has been changed to "low temperatures during the LIA should be linked to the Maunder Minimum of solar activity (Shindell et al., 2001)" (p. 21).

Page 358
Line 10: delete "to us"
Maybe mark the 21 volcanic eruptions in the reconstruction plot
REPLY: yes, we have added 21 volcanic eruptions in fig. 7a.

Page 359
Line 19: "relationsships"
Figures

Check label sizes, very different sizes make small text hard to read in the current printer-friendly pdf version.
Figure 1
Strange odd numbers on elevation scale

REPLY: these numbers indicate the lowest and highest elevations in our figure. We modified the fonts in the figure so that the labels are more readable when the figure is printed in normal size.

Figure 3
Rbar has not been explained in text

REPLY: the explanation of Rbar has been added in the revision (in section 2.2, p. 6-7).

Figure 4
Maybe just show the average of the two stations as all there parts of the plot are very similar

REPLY: The purpose to present the results for both stations is to show the consistency as a support to use the two-station averages in the following analysis. Thus, we didn't change this figure.

Figure 6
Is it mean or minimum temperature?

REPLY: sorry, it is Tmin

Figure 7
Labels and titles could be improved and duplicate scale removed

REPLY: The figure has been changed to show the results of wavelet analysis.

Figure 8
Is the "temperature anomaly" (top) again the "minimum"?

Reply: yes, it has been changed to "Tmin Anomaly"