We thank the reviewers for their valuable and constructive comments. We improved our manuscript based on the comments. The major changes to be found in the revised manuscript are as follows.

1. The gridded land air temperature dataset (CRUTEM3) was updated to CRUTEM4 (Jones et al. 2012). We recalculated all processes to find that all statistics were nearly the same as those in the previous manuscript. Therefore, an updated record was used in the revised manuscript.
2. Based on the comments, Figure 1, Figure 3 and Table 2 were improved in order to show full information appropriately. Supplementary materials (Figure S1, S2 and Table S1) were newly added on the reviewer's suggestion.
3. We added explanations in the body text and the legends so that the manuscript can be more readable.
4. The title "Climate Field Reconstruction of East Asian Spring Temperature" was changed to "Climate Field Reconstruction of Northeast Asian Spring Temperature" on the reviewer's suggestion.

Specific responses to reviewer's comments are listed below. Reviewer's comments are written in blue, while author's comments appear in black.

Response to reviewer #1
The manuscript based on five tree-ring chronologies by Ohyama et al., is a valuable contribution to understand the temperature variations for East Asia during 1794-1990. Although the majority of materials were not original, it was an interesting attempt to reconstruct regional climate. However, there are still many major problems need to be considered to improve the quality of the manuscript, and the expression of the paper need to be improved, too.

Reviewer#1 comment:
1. In the abstract, "Mean March–May temperature derived from a gridded land air temperature dataset (CRUTEM3) between 35–40°E and 125–140°N was reconstructed for the period of 1784–1990 AD", the units of Latitude and Longitude should change;

Response: We correct the units of latitude and longitude in the abstract.

Reviewer#1 comment:
2. In the abstract, "A fairly good agreement was found with cold periods as estimated from documentary records in China and Japan", China has a large geographical area, the authors should point out which part of China, for example, southeast China;

Response: We agree with the suggestion. We will use "southeast China" instead of "China".

Reviewer#1 comment:
3. For figure 1, the author should clearly explain the meaning of each symbols in the legend, such as black dots, triangle…., and the author should give exact positions of the grids from CRUTEMS or HadCRUT2 in figure 1, rather than solid lines or dotted lines;

Response: We improved Figure 1 in order to clearly show the meaning of symbols and the position of land air temperature grids employed in this study. All forest sites are indicated as black dots on the map instead of triangles, diamonds and black dots. The 5x5 degree grids of CRUTEM4 for calibration trials are indicated by dotted rectangles, while those for final calibration by blue rectangles. We believe these changes enable readers to clearly recognize the forest sites and the grids of CRUTEM4 used in this study. Please see the page below.

In addition to the improvement, we removed the description of HadCRUT3 from our

manuscript based on Reviewer #2's suggestion.

4. Actually, this manuscript only used five ring-width chronologies to reconstruct the regional spring temperature, the other two chronologies (JAPA01 and 007) mentioned in the manuscript was not included in the final reconstruction for "unsubstantial coherency with the other records", so we suggest to delete this two chronologies and corresponding description in this paper;

Response:
    We partly agree with the suggestion. As pointed out, JAPA001 and 007 (Kojo 1987) were not included in the final reconstruction. We eliminated them from Figure 1 in order to avoid readers to misunderstand the forest sites used for the final reconstruction. However, we think it is better to describe the process of calibration trial for reconstruction, because we need to show that the calibration trial was performed using all available tree-ring data in the region. Therefore, the description of JAPA001 and 007 in the calibration trial will be kept in the manuscript.

5. The author should give some detail information of the new chronology in north-eastern Japan, for example, the number of cores, the parameters, to what extent it was sensitive to the spring temperature, et al., to help readers to understand the quality of the new chronology;

Response:
    We will add the suggested information concerning the new chronology (NFI) as supplementary materials (Figure S1 and Table S1). Those will contain chronology statistics for NFI (northeastern Japan) and the results of the correlation analysis between the chronology and climatic data at the nearest meteorological station. The result shows that the NFI chronology is clearly sensitive to spring temperature. Please see the page below.

6. The explained variance of the reconstruction was only 19.4%, and it was strange it can pass the statistical tests. It was too low to be used to reflect regional temperature;

Response:
    Generally, explained variances of tree-ring reconstructions in mid-latitude Asia published so far are not very high, e.g. $aR^2 = 0.239$ (ring width, Yonenobu and Eckstein 2006) and $R = -0.521$($aR^2 = 0.271$)($\delta^{18}O$, Tsuji et al. 2008), indicating tree-ring records are a weak, but valid proxy in the region. It should be noted that our tree-ring reconstructions on both sides of the Sea of Japan represented similar fluctuations, which are likely to be due to the influence of climate. Historical records described in Sec. 4 (Comparison with other proxy records) in our manuscript also support our reconstructions.
    We agree that the explained variances are not high. To overcome this uncertainty, multiple statistical tests were used to demonstrate that our reconstruction is a valid temperature proxy. The benchmarks are all used in recent dendroclimatological works (e.g. Wilson et al., 2008, Cook and Pedersen, 2011).

7. Before combine the five chronologies together, the author should give a figure of comparison of these five chronologies, to show how good they cohere with each other. And this may partly explain why the explained variance of the reconstruction was so low (19.4%);

Response:
    We will add a comparison of the five chronologies as a supplementary material (Figure S2). It shows their graph during the calibration period. Please see the page below.

Reviewer#1 comment:
8. In Line 10-14 of page 3538, what is the correlation coefficient between the PCs and March-May mean temperature? And in the following sentences (Line 15-18 in the same page), "Since some divergence was observed between the actual and estimated CRUTEM3 data during the calibration trials, a frequency range higher than approximately 60 yr was removed by subtracting a high-pass filtered curve from the original data to remove the upward bias in the actual data series". The temperature record from CRUTEMS was the basis for analysis and reconstruction, the upward trend in the actual data series was a true record of temperature, why make such change?

Response:
The correlation coefficients between the principal components (PC) and March-May mean temperature are 0.423 ($p < 0.01$) for PC1 and 0.225 ($p < 0.05$) for PC2, respectively.

We tried to reconstruct spring temperature including not only higher frequency (interannual to interdecadal) but lower frequency (>60 years). A number of standardization method were examined, e.g., negative exponential, Hugershoff, and cubic spline functions, and the Regional Curve Standardization (RCS) as well, but those did not show good skill to reconstruct the lower frequency, failing to pass the validation statistics. We had to focus on reconstructing spring temperatures mainly at annual and interannual-to-interdecadal scales by applying the conservative detrending method. Finally, we were able to reconstruct the spring temperatures at the above-mentioned frequency range. Thus, our discussion is limited to variations with frequencies less than 60 years of the reconstructed CRUTEM4. We will add the reason why we removed upward bias from original CRUTEM4 in the manuscript.

Reviewer#1 comment:
9. The authors should give the regression model for the temperature reconstruction;

Response:
Although we showed the regression model in the notes of Table2, it seemed inconsiderate to readers. We will describe the regression model in the body text on this suggestion.

Reviewer#1 comment:
10. In Line 3 of page 3540, we can't get "Higher interannual fluctuation can be observed in period I (1792–1806)", contrarily, higher interannual fluctuation seemingly can be observed in the 1820-1830s (not exactly);

Response: We will describe them exactly on this suggestion.

Reviewer#1 comment:
11. The origin (references) of the figures d-f in figure 3 should be given in the legend.

Response: We added the references in the legend of Figure 3. See the page below.

Response to Reviewer #2
General comments: The paper has a potential to be published in Climate of the Past, but it has to be significantly revised. I think as particularly important following comments:

Reviewer#2 comment:
a) The authors selected five chronologies to reconstruct spring temperature by using PC regression, from the contribution percentage, accumulation contribution is 54%, which is little bit lower. The explained variance in Table 2 and on page 3538, line 27, "the explained total variance is 19.4% " are both not very high, which directly caused a big difference between estimation and actual temperature in Fig 3 a . So could you try other reconstruction approaches?

Response:
As described above (Response to Reviewer #1, 6), the explained variance is generally not high in this region. We tried to reconstruct spring temperature using all possible methods, and found the regression model is optimal.

Reviewer#2 comment:
b) Page 3539, What does it mean "PC1 (29.0% of the variance) represents the Korean chronologies, while PC2 (25.3% of the variance) represents the Japanese chronologies", how did you make a judgment?

Response:
According to the loadings of the correlation matrix in the principal component analysis, the factor loading for PC1 was high for Korean chronologies (BTR, STR and SPR), while that for PC2 was high for Japanese chronologies (NFI and O2127). We though such an expression in the text is often seen in dendroclimatic (e.g. Cook et al., 2003) and other statistical inference in related fields. This probably indicates that the geographic pattern is more significant than the difference between species.

Reviewer#2 comment:
c) Page 3537, please give more detailed meteorological data information. Why you choose two datasets of CRUTEM3 and HadCRUT3? From your data source, only tree-ring records were selected, no marine sediments or corals were used, and I suggested that land air temperature reconstruction is enough in this manuscript.

Response:
We agree with the suggestion. We will remove the description about HadCRUT3.

Reviewer#2 comment:
d) Page 3545, you mentioned the reference of Zhu et al, 2009, the reconstruction of Zhu et al is Feb.-Apr. temperature, same as your spring time, but from your comparison with other reconstructions part, I did not see that this reconstruction was shown in the Fig 3. I compared them personally, and found many differences between of them, so could you give some reasons for the differences?

Response:
The forest sites used in Zhu et al. 2009 were located between 42°12' - 43°27' N and 127°46' - 128°15' E. The sites were outside the grids for our chronologies. In the calibration trial, we tried to include the grid that contains the Zhu et al.'s sites (40-45°N, 125-130°E), but unfortunately the reconstructions were not successful. This may indicate that there is a dissimilarity of spring temperature between the area we employed in this study and the area between 40-45°N and 125-130°E, probably due to a geographical reason. Actually, a disagreement in the visual comparison of graph can be observed between Zhu's reconstruction and the reconstruction for central Japan (O2127) used in this study (Zhu et al. 2009, Figure 5).

Specific comments:
Reviewer#2 comment:
a) On the title, since this reconstruction only covered small region, which did not match with "East Asian", a large area, please change with another name.

Response:
As pointed out, "East Asia" was inappropriate for our reconstruction. We will use "northeast Asia" instead of "East Asia".

b) Caption of Figure 1 is not clear, some information has been included in the text, but here it is also important to show full information to the readers.

Response:
We will add information in the legend as follows:
1. All forest sites are indicated as black dots on the map instead of triangles, diamonds and black dots.
2. The 5x5 degree grids of CRUTEM4 for calibration trials are indicated by dotted rectangles, while those for final calibration by blue rectangles.
3. The references for the forest sites are cited.

Reviewer#2 comment:
c) More data analyses need to be done in the results, e.g. interannual - interdecadal variability.

Response:
Our discussion is mainly limited to variations with interannual-to-interdecadal scales. Analysis in this frequency range was done as much as possible in Section 4 (Comparison with other proxy records). Please also see the response to Reviewer#1's comment #8.

Reviewer#2 comment:
d) Page 3542, line 3-8, do you think the low temperature caused by volcanic eruptions did not impact the tree growth? Or the precipitation was enough, which was hidden behind the influence of low temperature.

Response:
We could not observe any spike-like depression just after the Laki (1783) and the Tambora (1816) eruptions both in our reconstruction and the regional reconstruction in Japan and Korea. Simply, we think it would be worth noting that in this region no abrupt climatic changes were observed from our regional and local reconstructions, although a number of paleoclimate studies have focused on such phenomena after stratospheric volcanic injections.

References

Cook, E.R., Krusic, P., and Jones, P.D. : Dendroclimatic signals in long tree-ring chronologies from the Himalayas of Nepal, International Journal of Climatology, 23, 707-732, 2003.

Cook, E.R., and Pederson, N. : Uncertainty, emergence, and statistics in dendrochronology. In M.K. Hughes, T.W. Swetnam, and H.F. Diaz (eds), Dendroclimatology: Progress and Prospects, Developments in Paleoecological Research 11: 77-112, 2011.

Kojo, Y. : A dendrochronological study of *Cryptomeria japonica* in Japan, Tree-Ring Bull, 47, 1 – 21,1987.

Tsuji, H., Nakatsuka, T., and Takagi, K. : δ18O of tree-ring cellulose in two species (spruce and oak) as proxies of precipitation amount and relative humidity in northern Japan. Chemical Geology 231, 67–76, 2006.

Yonenobu, H., and Eckstein D.: Reconstruction of early spring temperature for Central Japan from the ring widths of Hinoki cypress and its verification by other proxy records, Geophys Res Lett, 33, L10701, doi: 10.1029/2006GL026170, 2006.

Zhu, H., Fang, X., Shao, X., and Yin Z.: Tree ring-based Februry-April temperature reconstruction for Changbai Mountain in northeast China and its implication for East Asian winter monsoon, Clim Past, 5, 661–666, doi: 10.5194/cp-5-661-2009, 2009.
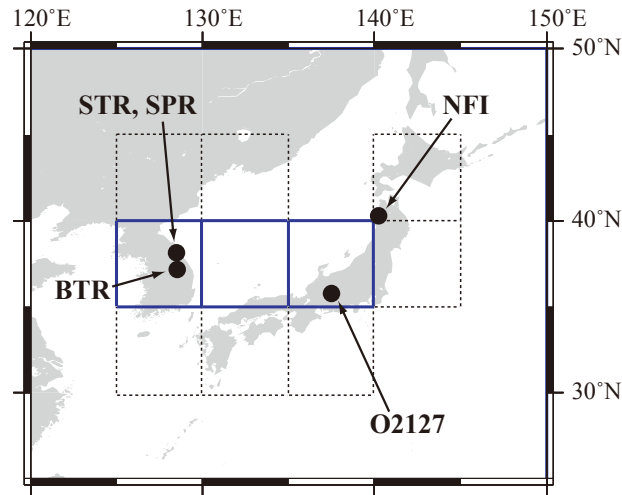
Fig. 1. Map of northeastern Asia.Black dots indicate forest sites used for the final calibration. STR, SPR, BTR (Choi et al., 1994), O2127 (Yonenobu and Eckstein, 2006) were previously developed, while NFI was newly developed in this study. The grids with dotted and blue lines show the gridded land air temperature data (CRUTEM4) (Jones et al., 2012) used for the calibration trials and the final calibration, respectively.

Table 2. Calibration and verification statistics for the northeast–Asian March–May temperature reconstruction

Calibration	1887–1990	1887–1938	1939–1990
R	0.458	0.407	0.509
aR^2	0.194	0.132	0.229
RE	0.21	0.166	0.259
ST	73+ / 31−	36+ / 16−	35+ / 17−
Verification		1939–1990	1887–1938
R		0.484	0.393
aR^2		0.234	0.155
RE		0.184	0.121
CE		0.019	− 0.040
ST		38+ / 14−	34+ / 18−

R: correlation coefficient; aR^2: explained variance after adjustment for degrees of freedom; RE: reduction of error; CE: coefficient of efficiency; ST: sign test result.

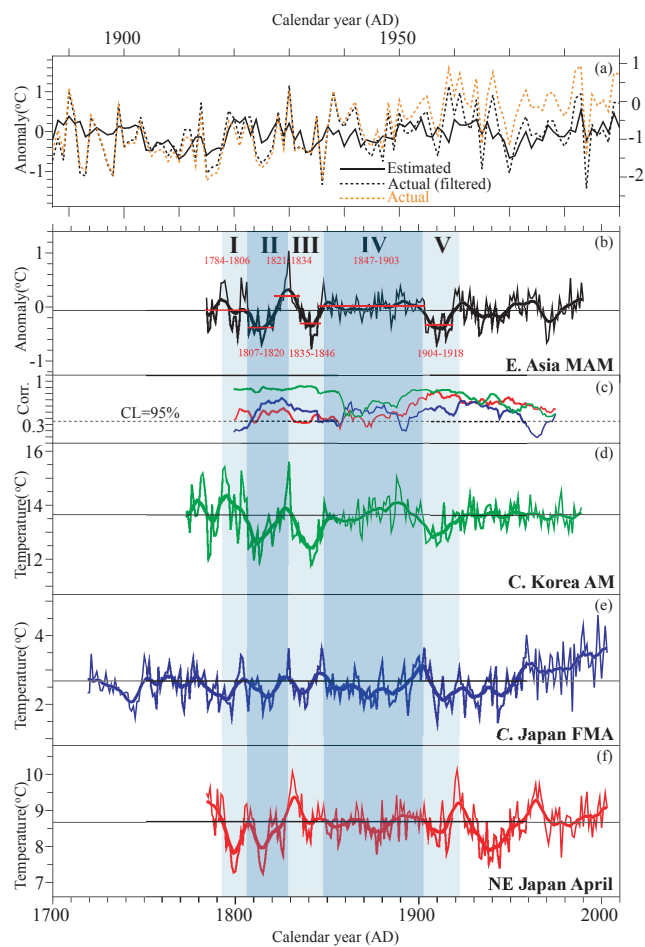Final calibration: $T_t^{MAM} = 0.12PC1_t + 0.26PC2_t − 0.06$

Fig. 3. (a)Estimated (solid line), and actual (dotted line) Mar–May (MAM) temperature anomalies. The actual record (orange line) is prewhitened to fit the low-frequency variability to the estimated series (see black dotted line). The left axis is for filtered actual and estimated records; the right axis for the actual record. (b) Reconstructed spring (MAM) temperature for East Asia. The horizontal red lines show regime shifts detected using intervention analysis(Rodionov, 2004). (c) 31-year running correlations between REAST and the regional reconstructions (d–f): (d) C.Korea, April–May (Choi et al., 1994), (e) C. Japan, February–April (Yonenobu and Eckstein, 2006) (f) NE Japan, April temperatures (this study).

Supplementary Material

Table S1: Statistics of NFI chronology

No. of trees(radii)	Time span	Mean	SD	MS	$\bar{r}_{bt}$	AC(1)	EPS
37(74)	1784-1990	1.005	0.329	0.197	0.306	0.598	0.869

MS, Mean sensitivity; SD, standard deviation; $\bar{r}_{bt}$, mean correlation between trees;

AC(1), first-order autocorrelation; EPS, expressed population signal

All values are calculated for the common interval of 1784–1990 using the program ARSTAN (Cook 1985).
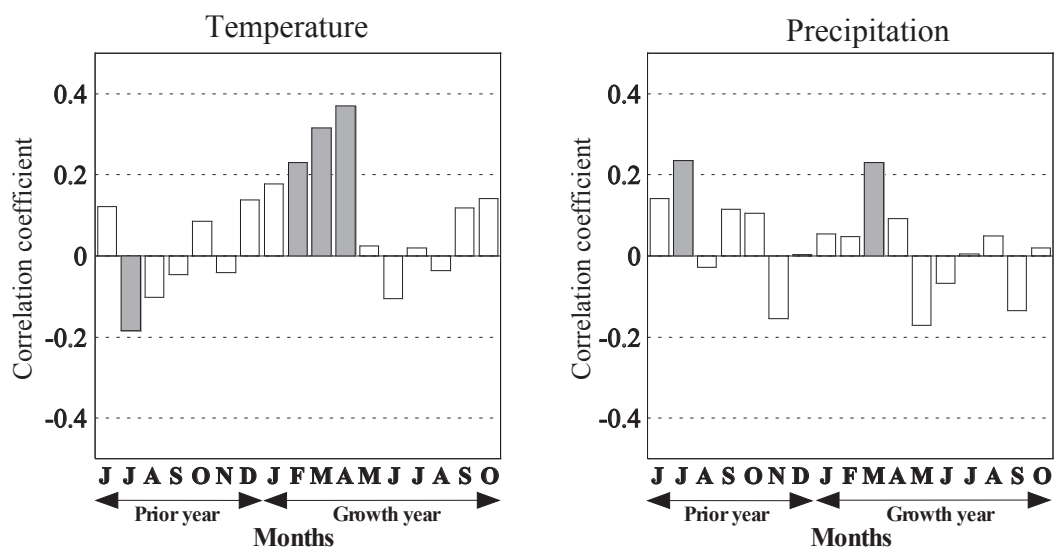
Figure S1: Simple correlation coefficient between NFI chronology and climatic records (monthly temperature and total precipitation at the Akita meteorological station, nearest to the forest site) Bars show correlation coefficients. Gray bars mean significant at 95% confidence level.
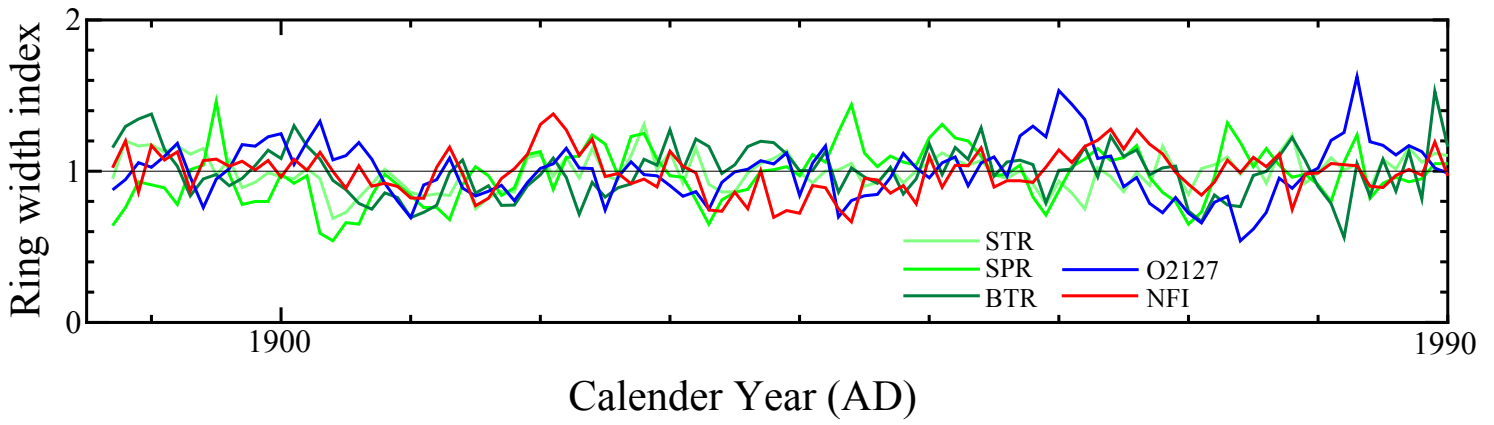
Figure S2: Comparison of five tree-ring chronologies used in this study over calibration period (1887-1990)