Response to the short comments

We would like to thank Dr. Li Ma for the short comments. These suggestions and comments are useful for improving the present version manuscript. In the following, we address all comments point-by-point according to the comments. All revisions are highlighted in the context.

Producing the high-resolution air temperature dataset in the mountain area is useful for regional climate or hydrological studies. The general target of this study is important. I have some suggestions and comments in below:

1. The method and the used lapse-rates should be more detail described in section 3.3.

-Answer: Thanks a lot for the comment. We agree that the method and the lapse rate should be presented in more detail. We added more information on the correction method in the section 3.3, especially an example on the internal lapse rate scheme (**P6 L2-11**). We also added the analysis on the temporal variability of lapse rate in Section 4.2 (**P8 L29-31, P9 L1-20**).

2. In table 1, is the elevation of the sites as same as the height from the 1 km SRTM DEM grid? it seems not the same (from P9, l18 ., I recommend to list the sites height in 1km DEM.

-Answer: Thanks a lot for the comment. Dr. Ma pointed out a very important issue. The elevations of sites are not the same as the height of 1 km DEM grids. The elevation differences between averaged 9 grids and station elevations are quite small with an average of -8 m (Table 1). Except the No. 9, the rest stations have less than 50 m elevation differences. Thus, the DEM generally matches the station elevations. The averaged 9 DEM grids height is added in the Table S1 (**supplementary Table S1**) in the revision.

Table 1 Elevation of averaged 9 DEM grids and the elevation differences with station elevation (m).

ID	averaged 9 DEM grids height	elevation Difference
1	1305	50
2	306	14
3	477	2
4	467	-26
5	764	30
6	672	-9
7	1885	-34
8	893	25
9	2004	-251
10	1101	3
11	868	5
12	940	-18
13	2462	-4
14	1057	-2
15	11	24
16	1221	8
17	978	-2
18	1066	33
19	937	-5
20	1635	3
21	433	46
22	1814	-85
23	758	-21
24	1548	20
Average		**-8**

3. Addition to comment 2. As the height in DEM grid and site point is different, in P6l13, why do you use averaged 9 grid points to evaluate the downscaled results. In my opinion, since each point in 1 km grid is downscaled according to its DEM height and ERA height, the nest grid point or the nest height grid of 9 point should be used for comparison.

-Answer: Thanks a lot for the comment. The reviewer 1 (Dr. Gerlitz) also pointed out that using the average 9 grid DEM height may lead to a systematic bias since the station elevation does not coincide with the mean elevation of the considered grid cells perfectly. Dr. Gerlitz also suggested correcting the temperatures directly to site scale. In the validation, the 3*3 grid cells surrounding each station were averaged. This approach was suggested by a referee when the authors evaluated the ERA-Interim temperature over the Tibetan Plateau (Gao et al., 2014). He/she claimed that this way can evaluate the ability of ERA-Interim on different topographies by selecting 3*3 grids with the station located in the center grid. Thus, in this study we took this suggestion. Despite, we found that the elevation differences are very tiny (smaller than 2m) among the 9 grids at 1km *1km grid resolution. Thus, we think this approach does not affect the validation very much (**P7 L12-21**).

4. The lapse rate varies in the different topographical situation and different timestamp like during the nighttime (Li et. al. 2014). The authors should discuss more about the method and results on the diurnal scale.

-Answer: Thanks a lot for the comment. Other reviewer also pointed out this issue. It is necessary to discuss the lapse rate in more detail. We added a new section in section 4.2 to analyze the temporal variability of lapse rates (**P8 L29-31, P9 L1-20**). Until now we do not have the sub-daily observations, we only compared the daily lapse rates between observation and ERA-Interim. The daily lapse rate is aggregated into monthly scale. Because the sites elevation ranges from 35 to 2458 m, thus for convenience, the ERA-Interim lapse rate was calculated using the temperature and geopotential height at 925 hPa and 700 hPa levels. The geopotential height at these two pressure levels range from around 150 m to 3000 m, which is close to the sites' elevation.

Figure 1 shows the temporal variation of monthly lapse rates. In general, the ERA-Interim has a higher temperature gradient than observation for the whole year. However, ERA-Interim captures the variability of observed lapse rate very well, especially in the warmer months (May to August). The inter-monthly variability of observed lapse rate is much higher than ERA-Interim, especially from September to January. The temperature gradient decreases significant from September, which represents the transition month from warm to cold climate regime. The temperature gradient increases significant from March, which represents the climate regime transfers from cold to warm conditions. Table 2 shows the monthly lapse rates over all sites in 1979-2013. The lapse rate differences are small (less than 0.5 °C km^{-1}) from May to August, while the differences are larger than 1 °C km^{-1} from September to December and January.

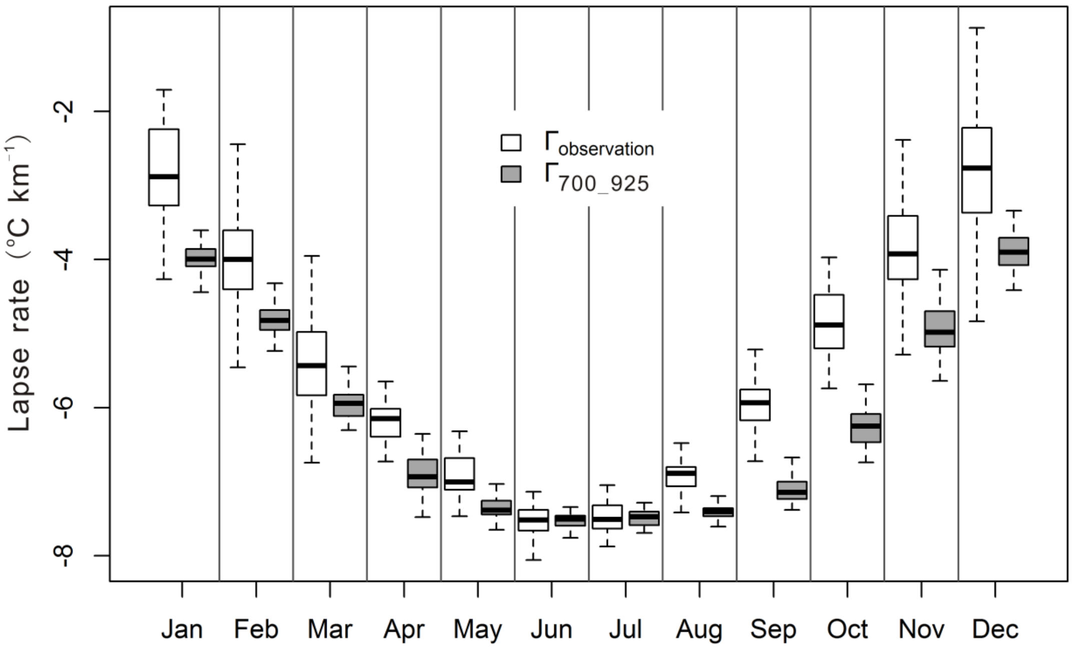

Figure 1 Boxplots of monthly lapse rates for observation and ERA-Interim (Γ_{700_925}). Thick horizontal linesin boxes show the median values. Boxes indicate the inner-quantile range (25% to 75 %) and the whiskers show the full range of the values.

Table 2 Monthly lapse rate (°C km^{-1}) over all sites in 1979-2013.

Month	observation	Γ_{700_925}
January	-2.79	-4.00
February	-4.01	-4.81
March	-5.42	-5.96

April	-6.14	-6.90
May	-6.92	-7.35
June	-7.55	-7.52
July	-7.48	-7.49
August	-6.95	-7.40
September	-5.93	-7.10
October	-4.86	-6.27
November	-3.94	-4.95
December	-2.88	-3.88

5. The authors produce the data from ～25 km to 1 km resolution. The total grid points are 818126, however, only 24 observation stations are used in validation. And in Figure 2-4, the authors only present comparison results at 4 stations, which probably have the best results. In addition, the authors only validate this 6-hourly dataset at a daily scale. To my point, the validation is somehow insufficient. It is not enough to conclude the reliability of this dataset, at least in the current level of discussion of the manuscript. I recommend the authors present some comparison analysis with ERA-Interim on the diurnal scale, and have more validation results with the station observations. Although very limited stations exist in this area, I know the diurnal max. and min. temperatures are provided in CMA station datasets, and these could be used in robust validation.

-Answer: Thanks a lot for pointing this out. It is true that 24 sites are not enough for such a large area. For Figure 2 and 4, the four stations are not the best ones but they were only selected to represent different sub-climate regimes and topography situations in the north slope of the CTM, south slope of the CTM, eastern CTM and western CTM. Prof. Dr. Guoyu Ren in National Climate Center of China and Dr. Aixia Feng in CMA told me (personal communication) that there is a high-density station data set (more than 2000 sites) but only for limited institute. We are trying to collect and apply more observations to validate the new data set. But at present, we have used the best we have. We expect other researchers to validate our product using different data resources. The validation and application are really welcome (**<u>P15 L12-18</u>**).

In order to illustrate the limitation of the new data set, we added more analysis on the maximum temperature (Tmax), minimum temperature (Tmin) and diurnal temperature range (DTR) in section 4.5. The warming trends for Tmax, Tmin and DRT also are investigated (**<u>P12, P13 L1-18</u>**). For example for DTR:

Figure 2 demonstrates the temporal variations of DTR over the 24 sites in 1979-2013. The original ERA-Interim has a more than 3 °C DTR bias compared to observations. The corrections reduce the DTR bias insignificant. The original ERA-Interim and corrections did not capture the significant decreasing trend of DTR. Table 3 shows the specific values on the trends for seasonal DTR over the 24 sites in 1979-2013. The decreasing trends are observed for annual and four seasonal DTR. Winter has the largest decreasing rate with the value of -0.384 °C $10a^{-1}$. Spring has the insignificant decreasing trend (-0.001 °C $10a^{-1}$). The original ERA-Interim and corrections capture the decreasing trends for summer and winter with smaller rates. However, they capture the opposite trends for spring and autumn, especially for spring (Table 3).

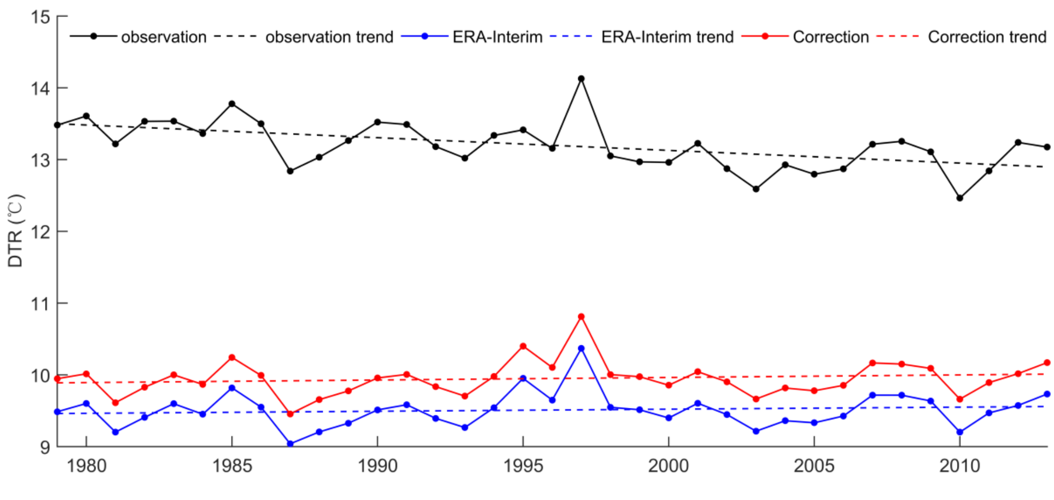

Figure 2 Temporal variations of DTR from observation, original ERA-Interim

and correction temperatures over the 24 sites in 1979-2013.

Table 3 Trends (°C $10a^{-1}$) of annual and seasonal DTR over the 24 sites in 1979-2013.

	Annual	Spring	Summer	Autumn	Winter
observation	-0.177	-0.001	-0.181	-0.132	-0.384
ERA-Interim	0.029	0.262	-0.052	0.069	-0.178
Correction	0.036	0.274	-0.044	0.068	-0.168

6. As the dataset provide temperature in 38 years, the authors should show some validation in annual and seasonal scale for this long period. In the spatial scale, I recommend to validate the dataset in each sub regions based on hydrology basin or climate zone or different elevation ranges. Figure 5 and figure 6 can not give much information of any performance skill of this dataset.

-Answer: Thanks a lot for the comments. As we answered for question 5, we are now applying the high-density observations from CMA for further validation. Meanwhile, we discussed with some researchers, for example Dr. Haijun Deng in Fujian Normal University, who did a lot of work in Kaidu river basin and Urumqi river basin. We initially agreed to validate the new data set using some automatic weather observations in a high-temporal resolution (**P13 L14-18**). Figure 5-7 showed the general climatology of the CTM based on the new data set. However, some researchers interested more in maximum and minimum temperatures rather than 1% and 99% quantiles. Thus, the figures for spatial distribution of maximum and minimum temperatures over the CTM are also provided in supplementary (**P11 L26-29, supplementary Figures S1 and S2**).

7. The dataset is very not friendly to use to me. As NetCDF format, I recommend to provide each file for the whole area at each timestamp (or each day, each month, each year) like most Grid datasets did (APHRODITE, TRMM, China meteorological forcing data from CAS. et. al.). It will be much easier for

regional climate or hydrology studies. Or save as the GRIB format like the reanalysis dataset.

-Answer: Thanks a lot for the suggestions. All reviewers pointed this problem out. We know that the data set is not very friendly at present. We have tried many ways to make it easier for end-user. For example, we put all points together for a single year in a signal NetCDF file, but it was more than 5 GB. A normal desktop cannot read it. The Matlab (we process the data in Matlab) always says out of memory. If we divided into monthly or daily, the numbers of files will be huge. Thus, we prefer to provide the smaller parts with limited points and time series. The advantage is that the potential users can download the data points according to the coordinates of study area. It is not necessary to download the whole data points. We are working on the version 2.0, which is friendlier to users. The accessibility of data set (including data format) also will be improved in the version 2.0(**P15 L12-18**).

Reference:

Li, Y., Z. Zeng, L. Zhao, and S. Piao (2015), Spatial patterns of climatological temperature lapse rate in mainland China: A multi–time scale investigation, J. Geophys. Res. Atmos., 120, 2661–2675, doi:10.1002/2014JD022978.