**Response to Anonymous Referee #2**

Air temperature downscaling is important for the mountain regions. This manuscript described a down scaling tools (maybe a software), which were validated by ground observed data in Alps and Qilian mountain.

The authors would like to thank the reviewer for the constructive feedback, and the thorough assessment of the manuscript. Below we provide a point-to-point response to each comment, reviewer comments are given in black, responses are given in blue. Additionally, we have included details of how we intend to address these changes in a revised submission.

I have several concerns:

(1) What is new? REDCAPP is a new one, or come from Fiddes and Gruber 2014; Gupta and Tarboton 2016, and Gao et. al. 2012? You should make a declaration clearly.

Though the upper-air temperatures ($T_{pl}^{f}$ *and* $T_{pl}^{c}$) are obtained by following Fiddes & Gruber (2014) (as we described in first part of Section 4.1), REDCAPP is a new method extending previous work. This is because REDCAPP disaggregates the difference of upper-air and near-surface temperatures ($\Delta T$) as a proxy of surface effects. In the reference methods, surface effects are either ignored (REF2) or treated as spatially invariant at the fine scale (REF1 and REF3). But in REDCAPP, a fine-scale DEM-based land surface correction factor (*LSCF*) is simulated to derive the fine-scale surface effects caused by cold air pooling and topography influences (e.g. hypsometric position described in Section 2.2). In the revised manuscript, we will make a clear declaration of new issues of REDCAPP by comparing with the methods used by Fiddes & Gruber (2014), Gao, Bernhardt, & Schulz (2012) and Sen Gupta & Tarboton (2016) as the first paragraph of new Section of 6.1 "Comparison with other downscaling techniques". Furthermore, the manuscript title "Parameterizing valley inversions in air temperature data downscaled from re-analyses" points to the key difference with respect to earlier work.

*"Though the upper-air temperatures ($T_{pl}^{f}$ and $T_{pl}^{c}$) are obtained following* Fiddes & Gruber (2014)*, disaggregating the difference of upper-air and near-surface temperatures as a proxy of surface effects ($\Delta T$) makes REDCAPP a new method. Additionally, the $\Delta T$ in REDCAPP is adjusted to fine scale responding to spatially heterogeneity of surface effect based on LSCF derived from DEM and observations, rather than ignored (REF2) or treated as spatial invariant (REF1 and REF3)."*

(2) Data introduction is weak. specially, the description of DEM process is too short. I can not understand what you did on the DEM data.

(a) We reformulated the Section of "Observations and quality control" (see below) to give a more detailed introduction of observations (Please also see our response to Referee #1).

*"The temperature from MeteoSwiss is observed using the Thygan instrument which has an accuracy of ± 0.01 °C, and temperatures from IMIS are measured by several different sensors (including Rotronic MP100H, Rotronic MP102H/HC2, Rotronic MP103A, Campbell Scientific CS215), with sensor accuracies ranging from ± 0.1 to ± 0.9 °C. In the Qilian Mountains, temperature sensor HMP155 with a typical accuracy of ± 0.2 °C are used. The 395 stations used cover an elevation range of ~250–4150 m as well as different topographic positions including peaks, slopes, plains and deep valleys (Figure 2a).*

*All temperature observations were filtered using a threshold (plausible values from -60 to 60 °C), and the outliers of temperature time series were removed by visually check. Time offsets between observations and ERA-Interim are avoided by conducting all analyses in UTC time. When using mean daily temperature, days with missing data were removed before further analysis. Though there are in total 395 stations used here, not all of them are available in a single year (Figure 2b). In total, there are ~2.5 × $10^6$ observations of mean daily temperature in or after 1980 used here."*

(b) About the DEM noise I simply copied part of the ASTER Global Digital Elevation Model Version 2–Summary of Validation Results (Meyer et al., 2011) here

"the addition of higher-frequency topographic signal in GDEM2 as compared to GDEM1 came at the cost of added, nearly ubiquitous, high frequency noise, as is visually apparent and as indicated by the higher standard deviation of differences from benchmark elevations (USGS) and from SRTM postings (NGA) despite the general reduction of artifacts such as pits and spikes."

In this case, we aggregated the original ASTER DEM with a spatial resolution of 1 arc-second to a spacing of 3 arc-second to avoid the noise. We added a new figure in appendix as Figure A1 (see below) to support the introduction of DEM aggregation.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 3270 | 3014 | 3465 | 3386 | 2810 | 4237 | 2210 | 2796 | 3147 |
| 2066 | 2783 | 2356 | 2699 | 3979 | 2556 | 4431 | 3736 | 3873 |
| 2844 | 3166 | 3864 | 4011 | 2517 | 3009 | 3421 | 4048 | 4206 |
| 2927 | 2104 | 2594 | 4016 | 2010 | 4054 | 3948 | 2345 | 3519 |
| 3532 | 3936 | 3555 | 2093 | 2478 | 2684 | 2877 | 2245 | 2837 |
| 2428 | 4095 | 2689 | 2655 | 4098 | 3374 | 3670 | 4224 | 2286 |
| 3894 | 4500 | 2923 | 2203 | 2782 | 2205 | 2846 | 2350 | 2912 |
| 2614 | 4451 | 2181 | 2468 | 4244 | 2954 | 4188 | 2172 | 4002 |
| 2026 | 2221 | 4449 | 4211 | 3203 | 4251 | 3716 | 3187 | 4446 |

DEM of 1 arc-second

| | | |
|---|---|---|
| 2980 | 3244 | 3540 |
| 3095 | 3051 | 3105 |
| 3251 | 3169 | 3313 |

DEM of 3 arc-second

Figure A1 Schematic illustration of DEM aggregation from a grid spacing of 1 arc-second to 3 arc-seconds by averaging. Numbers in the pixels are elevations in meter. In hypsometric simulation, the DEM with a grid spacing of 15 arc-second is derived using the same method.

(3) Discussion is weak.

(a) Section 6.1 Comparison with other downscaling techniques

We highlighted the new things of REDCAPP by comparing the reference methods used in this study in a new Section 6.1 "Comparison with the other downscaling techniques" (see below). Please also see the responses to comment (1). As mentioned by the Referee #1, a detailed discussion of comparisons of REDCAPP and exiting other methods, such as the Parameter-elevation Regression on Independent Slopes Model (PRISM) (Daly et al., 2000; Daly et al., 2002), the Daily Surface Weather and Climatological Summary (DAYMET) (e.g., Thornton et al., 1997), and the techniques used in Hijmans et al. (2005) is needed to highlight the advantages of REDCAPP. In this case, we added a detailed discussion (see below) in the Section 6.1.

6.1 Comparison with other downscaling techniques

*Though the upper-air temperatures ($T_{pl}^{f}$ and $T_{pl}^{c}$) is obtained by following* Fiddes & Gruber (2014), *disaggregating the difference of upper-air and near-surface temperatures as a proxy of surface effects*

*(ΔT) makes REDCAPP a new method. Additionally, the ΔT in REDCAPP is adjusted to fine scale responding to spatially heterogeneity of surface effect based on LSCF derived from DEM, rather than ignored (REF2) or treated as spatial invariant (REF1 and REF3).*

*Besides the lapse rate correction methods referenced in this study, many existing downscaling approaches for mountainous terrain focus on deriving fine-scale T through interpolation (e.g. truncated Gaussian weighting filter, Inverse Distance Weighting, or Kriging) of surrounding observations, and adjustments are then made based on fine-scale topography. PRISM (Parameter-elevation Regressions on Independent Slopes Model) (Daly et al., 2000; Daly, Gibson, Taylor, Johnson, & Pasteris, 2002), for example, derives a weighing function to represent the relationship of T with geographic (e.g. slopes, coastal) and meteorological (e.g. atmosphere boundary-layer) factors. Similarly, the approach by Thornton et al. (1997) calculates interpolation weights for the stations nearby, and corrected the downscaled results based on an empirical relationship of T to elevation, and Hijmans et al. (2005) conducted a second-order spline interpolation using latitude, longitude and elevation as independent variables. As observations are usually sparse in mountains, especially at higher elevation, these methods are expected to often have significant uncertainty caused by inadequately sampling of elevation and hence lapse rate. In comparison, REDCAPP relies on reanalysis data for air temperature and uses station data only for calibration of the LSCF related to CAP. REDCAPP derives lapse rates from multiple layers of upper air temperature encompassing the entire elevation range of study area. Thus, REDCAPP results are expected to be robust because both the Tsa and Tpl from reanalysis are used.*

(b) Section 6.3 Transferability

The discussion of model transferability is reinforced by clarifying the differences of transferring parameters of *LSCF* and applying REDCAPP in other mountains. Details please see our response to comment (4).

(4) In general, REDCAPP can not be transferred to different regions. It is a big shortage of this method. So, you should give more work on this issue. otherwise, others can not use your version 1.0.

In section 6.3 Transferability, we give a discussion of transferability of parameter values (α, β, γ in Eq. 6 and 7) of establishing land surface correction factor (LSCF). The parameters of land surface correction factor (LSCF) is different between Alps and Qilian Mountains, and hence hard to be directly transferred from the tested two mountains to other regions. But for REDCAPP, it could be applied in other mountains. This is because the establishment of LSCF is derived from fine-scale

DEM and observations, this fundamental concept is physically sensible and could be used in different mountains.

We reformulated the sentence of

*"The difference in estimated parameter values limits the transferability of REDCAPP as it requires tests new mountain regions to investigate suitability."*

to

*"The difference in estimated parameter values of LSCF limits the directly transferability of REDCAPP parameters from the mountain regions tested here to others as it requires new calibration in other mountain regions."* to clarify

Additionally, we added

*"REDCAPP can be applied to other mountains once the parameters (α, β and γ in Eq. 6 and 7) of LSCF are derived based on observations and a fine-scale DEM."*

at the end of Section 6.3 Transferability to clarify the difference of parameter transferability and applying models in other regions.