The manuscript describes a new empirical model (called GWMT-D) for the determination of weighted-mean temperature (Tm). The main focus laid on the modelling of daily variations, the finding of the optimal period for the determination of semi-annual and annual variations from numerical weather model (NWM) data and the reduction of global bias and RMS with respect to radiosonde and NWM data, especially at higher atmospheric levels.

GWMT-D provides mean values, annual, semi-annual and daily variations on a 5 °×5 ° global grid and four distinct height levels at 0, 2, 5 and 9 km. These parameters were derived from four years of NCEP2 data. Unfortunately it remains unanswered why exactly these height layers were chosen.

Nevertheless, in contrast to other state-of-the-art empirical models based on spherical harmonics (GTm-III, GWMT-IV and GTm_N), the gridded GWMT-D model has a smaller global mean RMS on surface level and on distinct height levels up to 9 km above surface.

From Figure 2 only small Tm variations are visible during daytime. A comparison of daily mean values with daily variations is missing. Thus it is not entirely clear how the modelling of daily variations improves the performance of the model.

In consequence major potential for improvements is seen in the analysis of daily variations and the description of its impact on the model performance. Further the authors should make transparent their decision making process for the selection of the four height levels at 0, 2, 5 and 9 km. In the following I provide some further recommendations and corrections, separately for content and language.

***Response**: Many thanks for your comments. The two issues mentioned above will be specifically explained below.*

Content:

- Page 2, Line 13 "Rv is the specific gas constant for the air; [: : :]" Rv is the specific gasconstant for water vapour

***Response**: Amended.*

- Page 2, Line 14: "e is the WV pressure (in hPa); [: : :]" Water vapour pressure e does not appear in Eq. 3 but rather density v of liquid water. In order to be consistent with Eq. A1 the ideal gas equation v =e/R*T should be added or at least the relation between e and v should be explained here.

***Response**: Amended by adding "Using the ideal gas law for the water vapour, ρv can be written as $\rho_v = e\ T / R_v$, where e is the WV pressure (in hPa)" (Page 2 Line 15).*

- Page 4, Line 26: Why is the minimum number of valid levels set to 20? The values seem to be too large. In the text above only 17 standard pressure levels are defined.

***Response**: These two values refer to two different data or observations. The reanalysis data (NCEP2 in this paper) is provided on the 17 standard pressure levels. The value 20 of valid levels is used in the quality control process of radiosonde observations. In order to remove the effect of the balloon drift and other error sources, the minimum number of valid levels is set to a relatively large value (Page 4 Line 22).*

- Page 5, Line 18: Why did you select exactly these four heights layer for modeling of vertical Tm lapse rate? Please give an inside into the decision making process.

***Response**: Generally speaking, the reference heights levels in GWMT−D are determined empirically.*

The reason for choosing the 2 km height is that the atmosphere below 2 km suffers the most from the terrain effect, and the reason for neglecting 10 km height is the Tm above this level may be not-a-number (zero water vapor above this height lead to 0/0 according to the definition of the Tm, so the Tm cannot be determined for this case). The results show that multi-layer model with more than four layers cannot archieve significantly improvence. Therefore, 0, 2, 5 and 9 km height levels are selected in the GWMT−D model (Page 6 Line 14).

- Page 8, Line 8: "Section 3.1.2 shows that the piecewise linear algorithm [: : :] is better than the direct modeling of Tm [: : :]." In section 3.1.2 the piecewise linear algorithm is mentioned as new model feature. Up to now no results are shown that the piecewise linear algorithm is better suited than any other approach.

***Response**: Amended. I agree. Section 3.1.2 only shows the horizontal variation of Tm lapse rate. Nevertheless, the results in Section 4 Validation of Tm models confirm that "<u>The results also show that the piecewise linear interpolation of T_m used in GWMT−D is better than the direct modelling of T_m lapse rate in GWMT−IV or the constant-value method used in both GTm−III and GTm_N.". </u>Therefore, this sentence has been moved to the Conclusion section (Page 11 Line 7).*

- Page 18, Figure2: The daily variations seem to be rather small, how large is the improvement when 6 hour values are used and interpolated by splines in comparison to daily mean values? Is it worth to add daily variations to the model? Please provide some numbers.

***Response**: Amended. Figure 2 has been redrawn only for the first example at latitude 30 °N, longitude 5 °W, height 2 km since the results of the other two examples are similar. Figures 2(a) and 2(b) were kept the same as before. Figure 2(c) shows the range (max – min) of daily T_m. Figure 2(d) shows power spectrum density of T_m residuals derived from the GTm–III (using daily-mean method) and GWMT–D (using spline interpolation). Figure 2(d) indicates that GWMT–D effectively captures diurnal variations in T_m but GTm−III does not. Although the post-fitting standard deviations of these two methods are very close (~ 3 K) at all the reference times (0, 6, 12, 18 UTC), the spline interpolation used in GWMT–D can significantly remove the diurnal variation (indirect evidence can be seen from Figure 2(d)). However, this can hardly be validated in this study because all the data used are at the reference times. Other data sets (e.g., COSMIC data) may be used for the future work (Page 6 Line 1).*

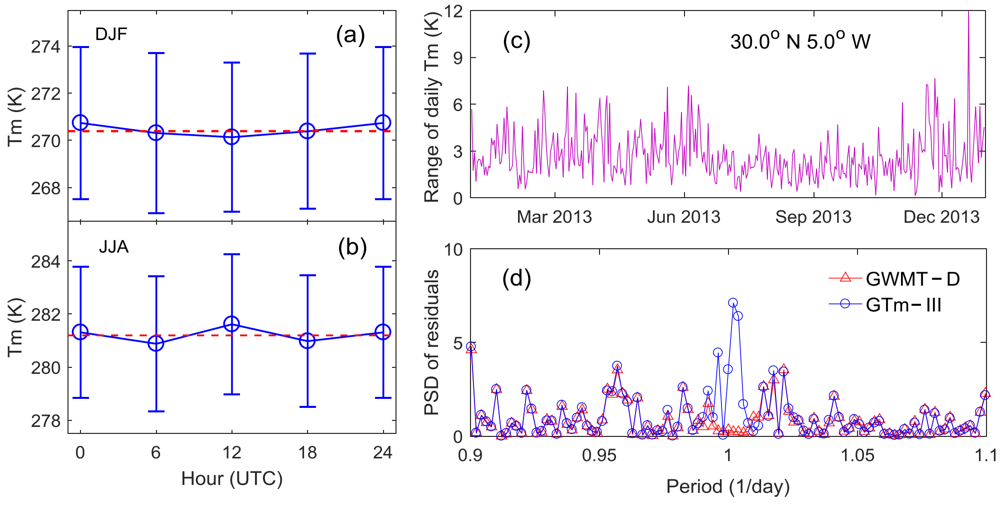

Figure 2. Statistical results of diurnal T_m (mean ± standard deviation) at 2 km height and four

reference times during (a) Dec-Jan-Feb (DJF) and (b) Jun-Jul-Aug (JJA) in 2013, (c) the range (max − min) of daily T_m and (d) power spectrum density (PSD) of T_m residuals.

- Page 8, Line 19: Table 2 shows the results only for pressure levels from 1000 hPa to 600 hPa and not below as mentioned in the text.

***Response**: Amended. In fact, the reasons were explained Section 4.1. In order to make it consistent, this part in Section 4.1 has been moved from Section 4.1 to Section 3.1.3 and rephrased as 'As a result, the Bias and RMS of all global grid points are given in Table 3 on the pressure levels of 925 hPa (~0.6 km) and 600 hPa (~5 km).' (Page 6, Line 29).*

- Page 12, Eq. A2-A4: Equations not found in given reference Aparicio et al., 2009.

***Response**: Amended by replacing with "(Ge, 2006)".*

- Page 12, Eq. A3: For the determination of gravity usually a height dependent term like (1(1+h/RE))^2 is added. Please explain why this was not used here.

***Response**: Amended. g(φ) is in fact the gravity on the geoid which has been added in the descriptions of Equation (A3). Traditionally, the geopotential height is defined based on a reference height which is usually the The dependence of height in gravity has been considered in the derivation of Equation (A2). More details can refer to Ge (2006).*

- Page 12 Eq. A5: relative humidity is abbreviated with 'f' but in the text 'RH' is used, please be consistent.

***Response**: Amended. 'f' in the Eq. A5 is an enhancement factor defined as the ratio of the saturation vapour pressure of moist air to that of pure water vapour (WMO, 2000), so f is a coefficient calibrating the results for ideal water vapour to moist air. The 'f' in Equation (A5)−(A8) have been replaced by 'f(P)' for clarification, e.g.,* $e = f(P)exp\left(\frac{17.62\,t}{243.12+t}\right)$.

- Page 23, Figure 12: Is the RMS of PWV given in (K) or rather in (mm)?

***Response**: Amended. The unit in Figure 12(a) is mm.*

Understanding and language:

Please review language and writing. A selection of not meaningful or incorrect phrases is given in the following:

- Page 1, Line 8: "One of the most critical variables in PWV remote sensing using GPS technique is the zenith tropospheric delay (ZTD)." Not a good introduction for a paper about mean temperature.

***Response**: Amended as 'In the GPS-based PWV remote sensing, the atmospheric-weighted-mean temperature (Tm) is a crucial parameter for the conversion from zenith tropospheric delay (ZTD) to PWV over the GPS station' (Page 1 Line 8).*

- Page 1, Line 12: "using global reanalysis data from 2010 to 2014 provided by NCEP-DOE Reanalysis 2 data (NCEP2)." Please correct, e.g. in the following way "using global reanalysis data 2 provided by the National Centers for Environmental Prediction (NCEP2)."

***Response**: Amended.*

- Page 2, Line 1: "using GPS-PWV" can be eliminated

***Response**: Amended.*

- Page 2, Line 3: "over the site of the station (: : :)." Please clarify

***Response**: Amended as 'The GPS-PWV have been used to study the temporal variation of PWV, such as seasonal and diurnal variation patterns'.*

- Page 2, Line 4: "over the region covered by the stations (: : :)." Try to be more precise.

***Response**: Amended as 'It also has been used to investigate the spatial variation in PWV over the GPS network'.*

- Page 4, Line 2: The first sentence of Section 2.1 is not meaningful, reanalysis data cannot have a main aim, please correct.

***Response**: The numerical weather prediction/analysis (NWP) system project usually aims at improving NWP models, or providing atmospheric dataset for studies including, e.g., climate change and monitoring and numerical seasonal prediction. There are some papers which introduce the motivation of reanalysis data (e.g., NCEP2, ECMWF ERA-Interim, and JRA-55).*

(1) *Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S. K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP-DOE AMIP-II reanalysis (R-2), Bulletin of the American Meteorological Society, 83, 1631-1643, 10.1175/Bams-83-11-1631, 2002.*

(2) *Ebita, A., Kobayashi, S., Ota, Y., Moriya, M., Kumabe, R., Onogi, K., & Kamahori, H.: The Japanese 55-year Reanalysis" JRA-55": an interim report.Sola,7, 149-152, 2011.*

In order to avoid further misunderstanding here, this sentence has been rephrased as 'The studies of climate change and climate monitoring benefit from the National Centers for Environmental Prediction/National Center for Atmospheric Research (NCEP/NCAR) reanalysis data' (Page 4 Line 2).

- Page 4, Line 13 "Radiosonde profile data from [: : :] stations over the globe in 2014 (: : :) : : : ."

***Response**: Amended as 'Radiosonde profiles from 585 global Integrated Global Radiosonde Archive (IGRA) stations (Figure 1) are selected to validate the new GWMT−D model'.*

- Page 5, Line 5: "Due to the fact that : : : the surface Tm from GGOS data is also used : : : ." clarify

***Response**: Amended as 'The GGOS data set has been applied in the development of GTm−III and will be used in the performance assessment of this paper'.*

- Page 7, Line 4: Assuming Tm at the target location [: : :] is Tm [: : :]" clarify

***Response**: Amended as 'Assuming T_m (φ, λ, h, DOY, HOD) is a function of target location (φ, λ, h), day of year (DOY) and UTC hour (HOD),…'.*