

1. First file: **mayg_2022avg.txt**. Note that the file **mayg_2022avg.txt** has columns separated by single space, making a total of 14 columns. The headings are not included in the **mayg_2022avg.txt** file, but, in order, they are: Column 1=Year, 2=Doy (Day of Year), 3=ZTD, 4=ZWD, 5=SigZTD, 6=PWV, 7=SigPWV, **8=P_s**, **9=T_s**, **10=ZHD**, 11=GradNS, 12=SigNS, 13=GradEW, 14=SigEW. Please note that the columns of interest in the **mayg_2022avg.txt** file are 8th, 9th and 10th columns!
2. Second file: **ZTD_mayg.dat**. The **ZTD_mayg.dat** file is columns that are tab-separated data. If my English makes sense. Making a total of 8 columns. Column headings are also not included in the file, but, in order, they are: Column 1= Year, 2=Doy, 3=Hour (starting from 0 – meaning 12 am of Day 1), 4=Measure, **5=ZTD**, 6=ZTDex, 7=ZTDerror, 8=Sats. If you investigate the file, it has data for the whole year but then there are days missing from the file meaning that there are #hours that are also skipped. Say for example Day 1 has 0-23 hours (12am to 11pm – just roll with it. We would like to do the average of columns 4,5,6,7,8 for each day. So, we do average 0-23 hours, considering column 3 (Hour). Then Day 2 average is from 24-47 (12 am to 11pm of the second Day). Note that you should write a script where you consider the Day and Hour, But we don't automatically assume that the number of rows of data for a random Day is 24. NO! Just average the data available. I hope average means the sum divided by the number of data. Say a day has 12 rows of data – you do the sum of each column of interest separately and divide by 12 separately. Not 24! Then output the daily averages to **ZTD_daily_mayg.dat**. NAMING IS IMPORTANT for me! I have so many files, deciding on the names brings some order in this world. Haha. You can laugh.

ZTD_daily_mayg.dat file will have new columns as follows: Column 1= Year, 2=Doy, 3=Measure, **4=ZTD**, 5=ZTDex, 6=ZTDerror, 7=Sats. NB: Hour column in the original file is not important here.

3. Third file: **ZTD_daily_mayg.dat** file. This one you generate yourself.
4. Compute (a) – (c): Write a single script to compute ZWD, T_m and PWV.
 - a) ZWD: Equation $ZWD = (ZTD * 1000) - ZHD$. ZWD is equal to ZTD*1000 minus ZHD.
 - ZHD source: Column **10=ZHD** of **mayg_2022avg.txt** file.

- ZTD source: Column **4=ZTD** of **ZTD_daily_mayg.dat**. NB: DON'T CONFUSE THE FILES!

YOUR task here is to write another script where you check if the Doy is the same, then you do the math $ZWD = ZTD \times 1000 - ZHD$. We multiply by 1000 because ZTD is in meters while ZHD is in millimeters. Therefore, the resulting ZWD is in mm.

- b) Weighted mean temperature, T_m , is computed from surface temperature T_s using equation $T_m = 70.2 + 0.72T_s$.

Source of T_s is column **9** of **mayg_2022avg.txt** file.

- c) PWV (in mm).

$$PWV = \frac{10^5}{\left(\frac{k_3}{T_m} + k'_2\right) R_v} \times ZWD$$

Where $k_3=3.776 \times 10^5$, $k'_2=16.52$, $R_v=461.495$. The only variables are T_m and ZWD, and of course the PWV. Do not mind the units. I am expecting PWV to be in the range 10-100 mm.

SUMMARY of (a) – (c): Write the data to a new file **PWV_daily_mayg.dat** file. It will be like adding new columns to **ZTD_daily_mayg.dat** file, where the first 7 columns are: Column 1=Year, 2=Doy, 3=Measure, **4=ZTD**, 5=ZTDex, 6=ZTDerror, 7=Sats. NB: You only write data for Doy that are both in the original files, as you did while computing for ZWD – you checked if a specific Doy was in both files, then you did $ZWD=ZTD-ZHD$. Next columns that you write to **PWV_daily_mayg.dat** file are: Column 8= ZWD, 9= T_m , 10=PWV.