

Comment on acp-2022-792

Anonymous Referee #1

Referee comment on "Variations in global zonal wind from 18 to 100□km due to solar activity and the quasi-biennial oscillation and El Niño–Southern Oscillation during 2002–2019" by Xiao Liu et al., Atmos. Chem. Phys. Discuss.,
<https://doi.org/10.5194/acp-2022-792-RC1>, 2023

General comments:

This manuscript applies multiple linear regression (MLR) to monthly mean zonal wind data in the stratosphere, mesosphere, and lower thermosphere obtained from SABER observations, MF and meteor radar observations, and MERRA2 meteorological reanalysis to examine the effects of QBO, ENSO, and solar activity as well as seasonal changes and long-term trends. Although many similar studies based on the MLR analyses have been conducted using long-term meteorological reanalysis data, there have been few research above the stratopause due to the difficulty of observing winds. In this sense, the efforts in this manuscript are commendable. On the other hand, the method of MLR analysis and statistical significance are not well documented, and the consideration of the short data period is not sufficient. In addition, English grammar check by a native speaker is also recommended. Therefore, I think that this manuscript needs substantial revision before publication. Detailed comments are given below.

Major comments:

- Time interval of the data

It seems that 18 years are too short to fit the 11-year solar cycle. Although the authors evaluated its impacts by changing the time interval, half or more of the data periods overlap, which does not seem very meaningful. Rather, a comparison using 40 years of

MERRA2 data would be more meaningful. As the authors say, the MLS has been assimilated since 2004, but its effect appears to be strong only for the vertical structure of temperature, not so much for the meridional gradient of temperature and the distribution of zonal wind that is related to the meridional gradient of temperature.

- Method of MLR analysis

From the explanation in section 2.2, it appears that eq. (2) is applied to data for 216 months over 18 years, in which case only one regression coefficient is obtained for the entire period. On the other hand, section 3 shows that regression coefficients were obtained for each month, suggesting that eq. (2) without including the seasonal variation term was actually applied to 18 years of data for each month. In that case, I do not know how the seasonal variation was estimated. The authors need to properly explain the MLR method.

- Multicollinearity

In the MLR analysis, multicollinearity often leads to meaningless results. The authors need to evaluate and indicate whether the correlations between regressors are sufficiently small before performing the MLR analysis.

- Statistical significance

In this manuscript, the regression coefficient is considered statistically significant if it is greater than 1σ . However, there is no description of how σ is calculated. In addition, when determining whether a regression coefficient is statistically significant in the MLR analysis, it is common practice to use the p-value of each regression coefficient. Unless there is a special reason to use σ , the p-value should be used (e.g., Mitchell et al. (2015)).

- Impact of SSW

In general, if the effect of SSW is large, it should occur that the regression coefficient is not statistically significant despite its large value. The authors should first check to see if this is the case, especially in the high latitudes of the winter northern hemisphere.

Furthermore, it is questionable whether it makes sense to apply the MLR analysis to spline interpolated data. Also, it should be explicitly stated which latitude bands were replaced by spline interpolation. Looking at Fig. 10c, it appears that all winters were replaced by spline interpolation, but major SSW does not occur every year. It should be explicitly stated by what criteria SSW is defined.

Minor comments:

- L. 143-145

Is it safe to consider data from a single point observation as the same as the zonal average, even though it is a monthly average? For example, how does this compare to the data of Smith et al. (2017)?

- Fig. 2

It is hard to see the phases from the arrows. I recommend to show the amplitudes by contours and the phases by colors.

- L. 237

Please clarify how the annual mean response was calculated. Is it an annual mean of the regression coefficient for each month? Or did you apply the MLR to the data including whole months (216 months)?

- L. 275

higher southern (northern) latitudes in summer (winter) higher latitudes in the winter hemisphere

- L. 275-277

I cannot see the signal at 50S/N at $z=50-80$ km.

- L. 381-420

Trend fitting is sensitive to the values at both edge points. The authors need to mention this point.

- L. 455-456

I think that the seasonal asymmetry is explained by semiannual and terannual components to some extent.

References:

Mitchell, D.M., Gray, L.J., Fujiwara, M., Hibino, T., Anstey, J.A., Ebisuzaki, W., Harada, Y., Long, C., Misios, S., Stott, P.A. and Tan, D. (2015), Signatures of naturally induced variability in the atmosphere using multiple reanalysis datasets. *Q.J.R. Meteorol. Soc.*, 141: 2011-2031. <https://doi.org/10.1002/qj.2492>

Smith, A. K., Garcia, R. R., Moss, A. C., & Mitchell, N. J. (2017). The Semiannual Oscillation of the Tropical Zonal Wind in the Middle Atmosphere Derived from Satellite Geopotential Height Retrievals, *Journal of the Atmospheric Sciences*, 74(8), 2413-2425. Retrieved Jan 18, 2023, from <https://journals.ametsoc.org/view/journals/atsc/74/8/jas-d-17-0067.1.xml>