

Comment on **essd-2021-44**

Xiangping Wang

Community comment on "A biomass equation dataset for common shrub species in China" by Yang Wang et al., Earth Syst. Sci. Data Discuss., <https://doi.org/10.5194/essd-2021-44-CC2>, 2021

This manuscript (MS) reports a large dataset for allometric equations to estimate shrub biomass, for a variety of species and sites across China. Interestingly, it seems that most equations were constructed by the present study based on sampling of 738 sites using a unified method. This greatly improved the quality and comparability of the equations, which avoid the weakness of compiling data from literatures (which were generally somehow different in sampling methods and thus led to uncertainty in data quality). Consequently, the dataset is clearly useful for improving the carbon pool/sink estimation of shrub ecosystems.

Considering the importance of this dataset, here I have some suggestions for the authors to improve the MS:

- 1) The Methods section needs to be clearer. As mentioned above, a major advantage of the dataset is that they have many equations based on their own measurements (the abstract said that they have 738 sites). However, the methods to obtain these equations were introduced together with the equations compiled from literatures. This leads the readers not very clear about the details of the methods. For instance, did they obtain one or more equations for each of the 738 sites? How many equations from the 822 equations were measured by this study? In my opinion, similar methods issues are better introduced independent of the equations collected from the literatures.
- 2) As for the equations from the literatures, these are also good data in supplementary to the measured equations. However, the methods to compile, select and validate equations from literatures clearly are different, and may be better to be introduced in another section.
- 3) In the Excel file reporting the data, I suggest to add a column in the "Equation" sheet, which clearly indicate the data source of the equation (e.g. "this study" or "the reference"). Presently, this information is reported in the "General" sheet. This is not convenient for the readers (personally I would prefer the equations measured by the present study, as explained above). Meanwhile, I also suggest the ranges of shrub height, crown, diameter etc. to be given for each equation in the "Equation" sheet. These ranges are critical for readers to determine whether an equation can be used for their estimation. However, these ranges are now only given for each species, which is not convenient for the readers.

4) The abbreviations in the "Equation" sheet seemed not well described. For instance, what does Vc, Ma, Ac, N, etc. mean? I did not find these in the "Description" sheet. Did I miss something?