Comment on essd-2022-51
Anonymous Referee #1

Referee comment on "A global marine particle size distribution dataset obtained with the Underwater Vision Profiler 5" by Rainer Kiko et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-51-RC1, 2022

The manuscript by Rainier Kiko and his cohort introduces the database that they have assembled for particle size profiles made around the world and that they placed online to make the database generally available. The data were obtained using UVP5 particle sizers after the instruments had been being calibrated in a consistent manner. The data have been stored in an easily accessible manner using a standard form. Their efforts have produced a valuable resource for the community. The UVP5 is probably the best calibrated optical particle counter used in oceanography. As with all techniques, there are quirks to the data and how they are interpreted that may not be realized by a user of this data set. I believe that it would help new users if the authors added a couple of paragraphs discussing some of the limitations of the data.

The observations are sorted into depth and size ranges, with the number of particles in each depth and diameter range divided by the associated water volume and diameter range to calculate a size spectrum. These are further processed by multiplying the number spectrum by the volume of a particle in the relevant size bin to yield a volume spectrum. The results in the database are given as these two spectra, which are probably the most useful forms for most people studying particle distributions and dynamics.

Problems with calculated spectral data arise when there is only 1 particle observed in a depth and size range. The number of particles sampled in each depth and size interval can be calculated from the sample volumes and size ranges given in the in the data tables. Would the sampled water have only one particle of the observed size in any water sample of the same volume? Or, is the observed particle the one lucky enough to be sampled when the average particle concentration is actually one tenth or one hundredth of the above estimate? Because large particles are rarer than small ones, this uncertainty is more of an issue with them. It is particularly a problem for measures that multiply by particle volume, which is greater for the larger particles.
One solution for an individual using the data is to decrease the uncertainty by setting to 0 the spectral value calculated with only one particle. Alternatively, increasing the sample volume by increasing the depth range for each sample would decrease the depth resolution of the observations but should yield more particles in each depth bin, hopefully decreasing the number of size intervals with only 1 particle. Those sampling at sea have the option of simply increasing their sample volumes within a given depth range.

Understanding this issue is important for those who plan to use the data that the authors have accumulated.

Smaller issue: the manuscript needs to be consistent in its punctuation of references in the text, particularly in the use of parentheses. For example, line 27 uses only one set of parentheses, while line 30 uses two sets. Most of the text has author names inside parentheses for citations; lines 52-55 do not.

In summary, this is an important and useful paper for oceanography. It would benefit from adding couple of paragraphs describing issues associated with using discrete counts as if they are continuous concentrations. The manuscript also needs another pass to smooth out typographic inconsistencies.