

Figure 1 of the Introduction page of this chapter shows the size *ranges* commonly associated with different components of natural waters. This page discusses the theory of how to quantify the number of particles as a function of their size. The Level 3 page Creating Particle Size Distributions from Data discusses the mechanics of creating particle size distributions from measured data.

## Non-uniqueness of Particle Size

The first problem is to decide what is meant by particle “size,” especially when speaking of statistical measures like the “average size” of a collection of particles.

Suppose we have three spherical particles of radii  $r = 1, 2,$  and  $3$ . What is the “average size” of these three spheres?

If the average size is based on the radius, then the average (radius) sphere has a radius of  $(1 + 2 + 3)/3 = 2$ .

If the average is based on the surface area of the spheres,  $A = 4\pi r^2$ , then the average surface area is  $4\pi(1^2 + 2^2 + 3^2)/3$  which corresponds to a radius of  $r = 2.16$ . The same holds true for the cross-sectional area of the spheres.

If the average is based on the volume of the spheres,  $V = \frac{4}{3}\pi r^3$ , then the average volume is  $\frac{4}{3}\pi(1^3 + 2^3 + 3^3)/3$  which corresponds to a radius of  $r = 2.29$ .

So there is ambiguity in how to define the “average particle size” even for a collection of spherical particles.

The situation is ever worse for non-spherical particles, which do not have a single parameter like radius that can be used to specify the size of the particle. The “size” of a non-spherical particle is sometimes taken to be

- the largest dimension of the particle,
- the arithmetic mean of the largest dimension of the particle,  $D_L$ , and the smallest dimension,  $D_S$ :  $\frac{1}{2}(D_L + D_S)$ ,
- the geometric mean of the largest and the smallest dimensions:  $\sqrt{D_L D_S}$ ,
- the diameter of a sphere with the same surface area as the particle (an “area-equivalent” sphere), or
- the diameter of a sphere with the same volume as the particle (a “volume-equivalent” sphere),

and there are also other measures of the “size” of a non-spherical particle. Note that the arithmetic and geometric means are equal for spherical particles, but are in general not equal.

All of these measures of particle size are valid, but one measure may be optimum for a particular application. For example, a Coulter Counter measures the change in electrical conductivity when a particle passes through the sensor. This change is proportional to the amount of material in the particle, i.e. to the particle volume. Thus a Coulter Counter measures the size of a volume-equivalent sphere. Laser diffraction instruments measure diffraction of light caused by a particle in the light beam. The angular shape of the diffraction

pattern is determined by the cross-sectional (projected) area of the particle as seen by the beam. Thus laser diffraction measures the size of an area-equivalent sphere.

The diameters of area-equivalent and volume-equivalent spheres can be considerably different. Consider a cubical particle of length  $\ell$  on a side. The volume is  $V_c = \ell^3$  and the equivalent-volume sphere ( $V_s = \pi D^3/6$ ) has a diameter

$$D_{\text{vol}} = \left(\frac{6}{\pi}\right)^{1/3} \ell = 1.24\ell.$$

The surface area of the cube is  $A_c = 6\ell^2$  and of a sphere is  $A_s = \pi D^2$ , which leads to

$$D_{\text{area}} = \left(\frac{6}{\pi}\right)^{1/2} \ell = 1.38\ell.$$

If cubes of this size are seen randomly oriented in a laser beam, the average projected area is one-fourth of the surface area (Cauchy's Average Projected Area Theorem, which shows that for a convex polyhedron, the average projected area over all orientations, i.e. the average cross section, is one-fourth the surface area of the polyhedron). Thus the laser sees, on average, particles of area  $6\ell^2/4$ , and the equivalent-projected-area sphere has diameter  $\pi D^2/4$ . This again leads to  $D_{\text{area}} = \sqrt{6/\pi}\ell$ .

Thus for a cube the diameter of the equivalent-area sphere is 1.11 times the diameter of the equivalent-volume sphere. Coulter-principle and laser-diffraction instruments will therefore report somewhat different equivalent-sphere sizes for the same particle. Clearly, it is important to understand exactly what an instrument measures and what instrument was used to measure particle sizes.

Unfortunately, publications sometimes fail to state exactly what measure they are using for the size of non-spherical particles.

## Particle Size Distributions

First decide on some measure of particle size; call it the diameter  $D$ . If the particles are known to be spherical (e.g., fog droplets or small bubbles in water), then  $D$  is the diameter of the sphere. For non-spherical particles,  $D$  is most often taken to be the diameter of the volume-equivalent sphere. The next step is to quantify how many particles there are of each size  $D$ . There are two equivalent ways to do this.

The *cumulative number size distribution* ( $CSD_n$ )  $N(D)$  is usually defined (in the particle-sizing literature) as the number of particles per unit volume *larger* than size  $D$ .  $N(D)$  is usually expressed in units of  $\text{m}^{-3}$ . In principle, this function can be measured simply by counting the numbers of particles larger than a given size  $D$ , as  $D$  in principle ranges from 0 to  $\infty$ , but in practice  $D$  ranges from some minimum  $D_{\text{min}}$  to some maximum  $D_{\text{max}}$ . These minimum and maximum sizes are usually determined by the instrument used to do the counting.

The *particle number size distribution* ( $PSD_n$ )  $n(D)$  is a function defined so that  $n(D)dD$  is the number of particles per unit volume between size  $D$  and  $D + dD$  (or in the size range

$D \pm D/2$ ). The units of  $n(D)$  are usually number of particles per cubic meter per micrometer of size range, i.e.  $\text{m}^{-3} \mu\text{m}^{-1}$ .

The  $CSD_n$ ,  $N(D)$ , as defined above decreases as  $D$  increases, so  $dN(D)/dD$  is negative. The corresponding PSD is then the negative of the derivative of the CSD, and is usually written as

$$n(D) = \left| \frac{dN(D)}{dD} \right| \quad (\text{m}^{-3} \mu\text{m}^{-1}). \quad (1)$$

Because of the large range of values for both  $N(D)$  and  $D$ , it is customary to plot these functions on log-log scales and to work with logarithmic size intervals. You then see (e.g., Junge (1953) and Liley (1992))  $n(D)$  defined in terms of a log derivative, which is related to  $dN(D)/dD$  by

$$\frac{dN(D)}{d \log D} = \frac{D}{\log e} \frac{dN(D)}{dD}. \quad (2)$$

The  $\log e$  comes from a change of base from base 10 logarithms, which are convenient for working with data, to base  $e$  natural logarithms, which are convenient for mathematics. Thus PSDs defined by logarithmic and linear derivatives will differ by a factor of  $D/\log e = 2.30D$ :

$$n'(D) = \left| \frac{dN(D)}{d \log D} \right| = \frac{D}{\log e} \left| \frac{dN(D)}{dD} \right| = \frac{D}{\log e} n(D).$$

The converse of Eq. (likesection1) is

$$N(D) = \int_D^\infty n(D) dD \quad (\text{m}^{-3}).$$

The total number of particles per unit volume is given by

$$N_t = N(0) = \int_0^\infty n(D) dD \quad (\text{m}^{-3}).$$

In the particle sizing literature, it is customary to use a diameter  $D$  as the measure of particle size. However, Mie Theory usually expresses its size parameter in terms of the particle radius  $r = D/2$ . If using a PSD( $D$ ) in Mie calculations as described on the Mie Theory page, it is necessary to use  $PSD(r) = 2PSD(D)$ .

## Area and volume size distributions

The previous definition (likesection1) of the PSD was for the *number* of particles in a unit volume, hence the notation  $PSD_n$  or  $n(D)$ , with the subscript  $n$  indicating number of particles, and the name number size distribution. We can also define size distributions for the area and volume of the same particles. The surface area of a sphere is  $\pi D^2$ , so the PSD for area (subscript  $a$ ) is

$$PSD_a(D) = \pi D^2 PSD_n(D) = \pi D^2 n(D) \quad (\mu\text{m}^2 \text{m}^{-3} \mu\text{m}^{-1}). \quad (3)$$

Note that the units of the area size distribution  $PSD_a(D)$  are particle surface area, usually in  $\mu\text{m}^2$ , per unit volume per unit size interval. The size distribution for the cross-sectional

area of the particles is  $\frac{\pi}{4}D^2PSD_n(D)$ , with the same units. The PSD for particle volume,  $\frac{\pi}{6}D^3$ , is then (subscript  $v$ )

$$PSD_v(D) = \frac{\pi}{6}D^3PSD_n(D) = \frac{\pi}{6}D^3n(D) \quad (\mu\text{m}^3 \text{ m}^{-3} \mu\text{m}^{-1}). \quad (4)$$

Now the units of the volume size distribution are particle volume, usually in units of  $\mu\text{m}^3$ , per unit volume per unit size interval. For particle volume measured in cubic micrometers,  $PSD_v(D)$  thus gives a size distribution for particle volume in parts per million, per unit size interval. Of course, these formulas are based on the assumed validity of area- and volume-equivalent spheres.

The total surface area of particles per unit volume is

$$A_t = N_a(0) = \int_0^\infty PSD_a(D)dD = \pi \int_0^\infty D^2n(D)dD \quad (\mu\text{m}^2 \text{ m}^{-3}).$$

The total volume of particles per unit volume is

$$V_t = N_v(0) = \int_0^\infty PSD_v(D)dD = \frac{\pi}{6} \int_0^\infty D^3n(D)dD \quad (\mu\text{m}^3 \text{ m}^{-3}).$$

If the quantity of interest is the number of particles (e.g., the number of phytoplankton of various sizes), then the number size distribution  $n(D)$  is the relevant distribution. However, some processes depend more on the surface area of the particles than on their number. For example, chemical reactions such as combustion (e.g., burning of coal dust in a power plant) depend strongly on the surface area, and laser diffraction is governed primarily by particle cross-sectional area. For such problems,  $PSD_a$  is the distribution of interest. For processes that depend on particle volume or mass (e.g., the mass of sediment material resuspended in a bottom boundary layer or particle buoyancy in a fluid), the volume distribution  $PSD_v$  is the distribution of interest.

Many commercial particle sizing instruments allow the user to select the output as a number, area, or volume PSD. Keep in mind that the conversions between one form of PSD and another are being made on the basis of an assumed equivalent spherical particle, even though the underlying measurement may be based on particle area (laser diffraction) or volume (Coulter counter). Thus a laser diffraction instrument may give a good measurement for the area  $PSD_a$  even for nonspherical particles, but errors can be introduced when converting the area PSD to number or volume PSDs. A Coulter counter may give a good measurement of a volume distribution  $PSD_v$  even for nonspherical particles, but errors can be introduced with converting the volume distribution to area or number distributions.

## Comments on terminology

As is so often the case, different authors sometime use different terminology. For example, Bader (1970) calls  $N(D)$  the cumulative number distribution function and  $n(D)$  the number distribution function. Jonasz (1983) calls  $N(D)$  the cumulative size distribution and  $n(D)$  the particle size distribution. I use the term cumulative number size distribution ( $CSD_n$ ) for  $N(D)$  and particle number size distribution ( $PSD_n$ ) for  $n(D)$  because that seems to be the most common in the literature (subscript  $n$  for number).

However, in probability theory, a cumulative distribution function  $CDF(D)$  refers to the how much of a total is *less* than  $D$ , rather than greater than  $D$  as used for  $N(D)$ . Runyan et al. (2020) compute a true  $CDF(D)$  by dividing the  $PSD(D)$  by the total number of particles over all sizes,  $N_t$ , to get a true  $PDF(D)$ , and then computing a (nondimensional) CDF as  $CDF(D) = \int_0^D [n(D')/N_t] dD'$ . They also employ a carefully designed nomenclature with  $N'$ ,  $A'$ , and  $V'$  denoting the number, area, and volume PDFs, which they refer to as density functions. This is somewhat in analogy to a spectral density function, which shows how much of something there is per unit interval of distance or frequency. Their true CDFs are denoted by  $CDF_n(D)$ ,  $CDF_a(D)$ , and  $CDF_v(D)$  for the number, area, and volume CDFs computed from the respective density functions.

Any of the functions  $CSD_n(D) = N(D)$ ,  $PSD_n(D) = n(D)$ ,  $PSD_a(D)$ ,  $PSD_v(D)$ ,  $CDF_n(D)$ ,  $CDF_a(D)$ ,  $CDF_v(D)$  etc. is properly termed a particle size distribution. They are all different ways of describing how some feature of particles (number, area, volume) change with size, and one can be computed from and of the others (under certain assumptions like spherical particle shapes).

I suspect that the “backwards” definition of the cumulative size distribution as used in particle analysis traces back to measurement of size distributions of soil or other particles using a succession of wire mesh sieves. You first sieve out the large particles (gravel, for example) and count those. Then a smaller mesh collects the next smaller size (sand, for example), and so on down to the finest particles (e.g., silt). The fine meshes are very delicate, so you have to remove the large particles first. Thus as the sieving process works from large to small particles sizes, you are accumulating the number of particles larger than the current sieve size, and building up a  $CSD(D)$  that shows how many particles are *larger* than  $D$ . Given a normalized probability distribution function  $PDF(D)$  (e.g.  $PDF_n(D) = n(D)/N_t$ ), it is convenient to integrate (or sum) the PDF from small to large values of  $D$ , and thus build up a  $CDF(D)$  that shows the fraction of the total that is *less* than  $D$ .

## Models for Particle Size Distributions

Many models have been proposed for particle size distributions. These differ according to the environment (atmosphere, ocean, industrial process) and particular data set being analyzed. The classic studies by Junge (Junge (1953), Junge (1955)) were concerned with aerosol size distributions. He found that a number size distribution of the form

$$\frac{dN(D)}{d \log D} \propto D^{-3} \quad (5)$$

gave a good fit to aerosols down to about  $1 \mu\text{m}$  in diameter. Note that according to Eq. (likesection2), the  $D^{-3}$  dependence of Eq. (likesection5) corresponds to a  $D^{-4}$  dependence when written as

$$n(D) = \left| \frac{dN(D)}{dD} \right| \propto D^{-4} \quad (6)$$

Thus you usually see the statement than a Junge distribution has a slope of -4 when  $n(D)$  and  $D$  are plotted on logarithmic axes.

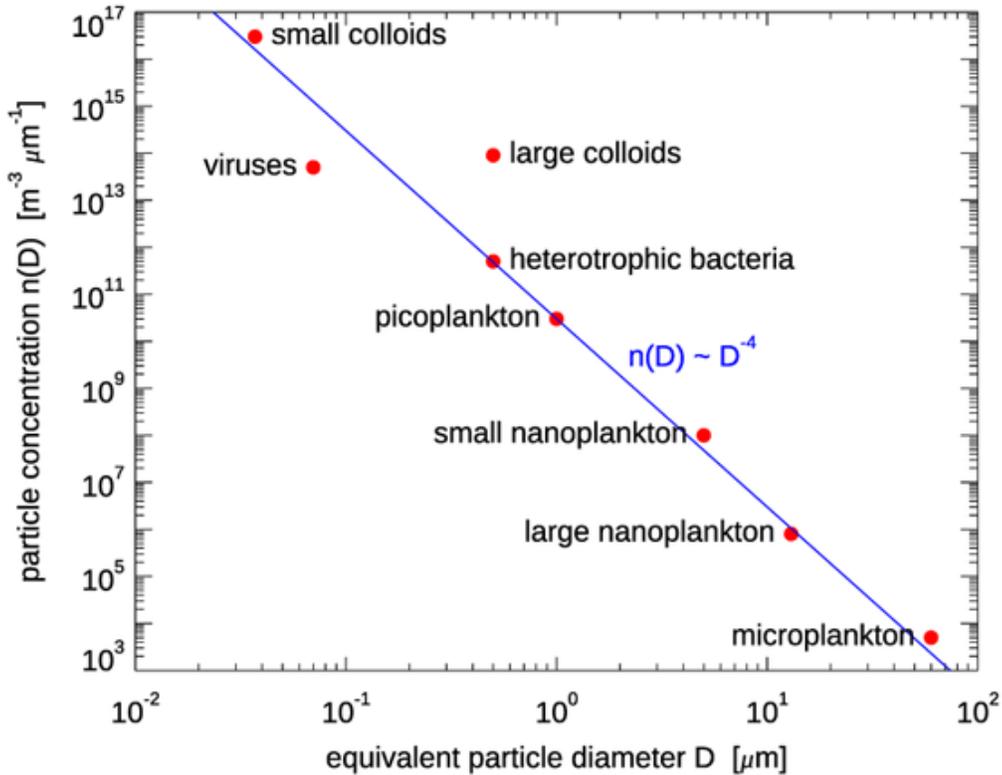


Figure 1: Approximate size distribution of biological particles in the ocean. The data points correspond to the mid-points of the size classes. The blue line is a Junge particle size distribution centered on the picoplankton point. Data courtesy of D. Stramski from data originally published in Stramski and Kiefer (1991).

Figure figure1 shows one instance of the number PSD for various size classes of marine biological particles;  $D$  is the diameter of a volume-equivalent sphere. The Junge distribution is shown as a blue line, with the magnitude fixed by the value for picoplankton.

A more general number size distribution model is the *power law distribution*, which is usually written as

$$n(D) = n(D_o) \left( \frac{D}{D_o} \right)^{-S}. \quad (7)$$

The exponent  $S$  is usually called the slope parameter.  $n(D_o)$  is the value of the distribution in  $\text{m}^{-3} \mu\text{m}^{-1}$  at a reference diameter  $D_o$ . The Junge distribution (like section 6) is a special case of a power law distribution for  $S = 4$ . (Some authors call the power law distribution a Junge distribution, but others are careful to use the term Junge distribution only for  $S = 4$  because that is the only slope Junge considered in his papers.)

Figure figure2 shows 168 number size distributions  $n(D)$  measured over the equivalent-sphere size range  $0.8 \leq D \leq 120 \mu\text{m}$  in Arctic waters. When fit to the power law model of Eq. (equation 7), the best-fit values of  $-S$  have the distribution of values shown in Fig.

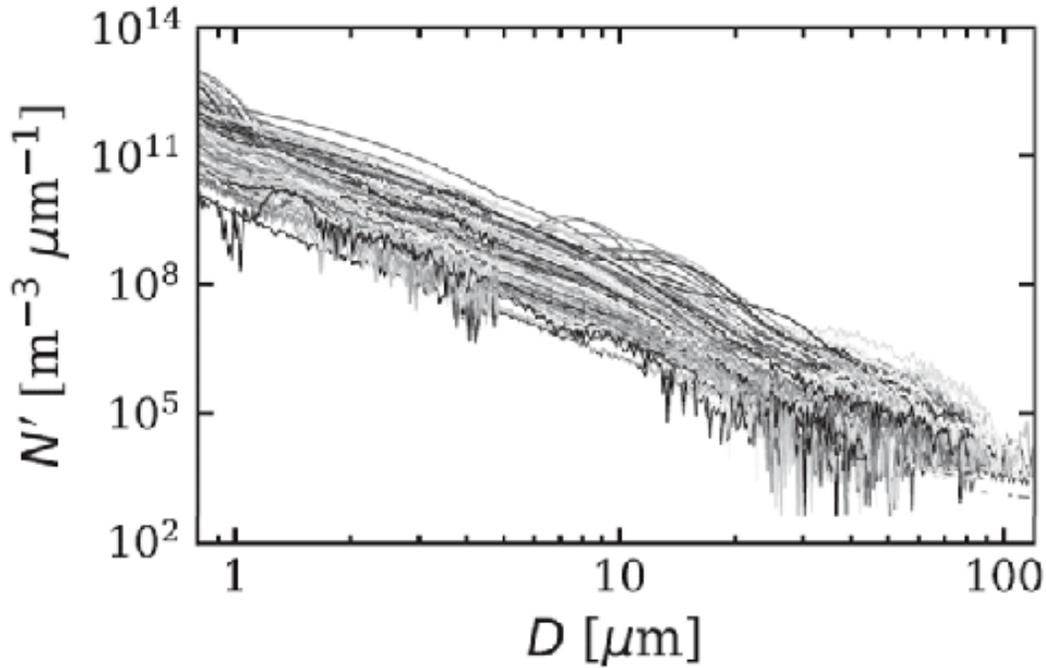


Figure 2: 168 particle number size distributions  $N' = n(D)$  taken in Arctic waters. From Fig. 2 of Runyan et al. (2020). Reproduced by permission under a Creative Commons license.

figure3. The average of the 168 slope parameters is  $S = 3.6 \pm 0.37$ .

Figure figure4 shows two of the PSDs of Fig. figure2 along with the corresponding area and volume distributions, and the cumulative distribution functions. Although the fit of the power-law model (equation7) to these number PSDs appears quite good on a log-log plot, it must be noted that there are order-of-magnitude deviations from the best-fit curves, namely near  $40\text{-}50\ \mu\text{m}$  in the upper left figure, and near  $1.5\ \mu\text{m}$  in the lower-left figure. Such deviations are common in fits such as these and can result, for example, from a bloom of a particular species of phytoplankton. It should be noted from the CDFs in the right panels that the total number of particles comes mostly from particles less than about  $3\ \mu\text{m}$ , but the total volume of particles comes mostly from the larger particles greater than about  $30\ \mu\text{m}$ . This is simply because one particle with  $D = 10\ \mu\text{m}$  has as much volume as 1000 particles of size  $D = 1\ \mu\text{m}$ , for example.

It sometimes happens that the small size and large size ends of the distribution have different slopes. A better fit is then obtained by separate fits of a power law for the two size regions, but other distributions have also been used.

It is important to recognize that a power-law distribution implies that the size range extends from  $D = 0$  to  $\infty$ . This is never the case in nature. The smallest phytoplankton have a diameter of about  $0.2\ \mu\text{m}$  and the largest are about  $200\ \mu\text{m}$ . Thus a PSD for phytoplankton, whatever its mathematical form, should be applied only for the  $D = 0.2$  to  $200\ \mu\text{m}$  size range, and perhaps for an even smaller range. Particular species or size classes of phytoplankton have much smaller size ranges. The upper panel of Fig. figure5 shows

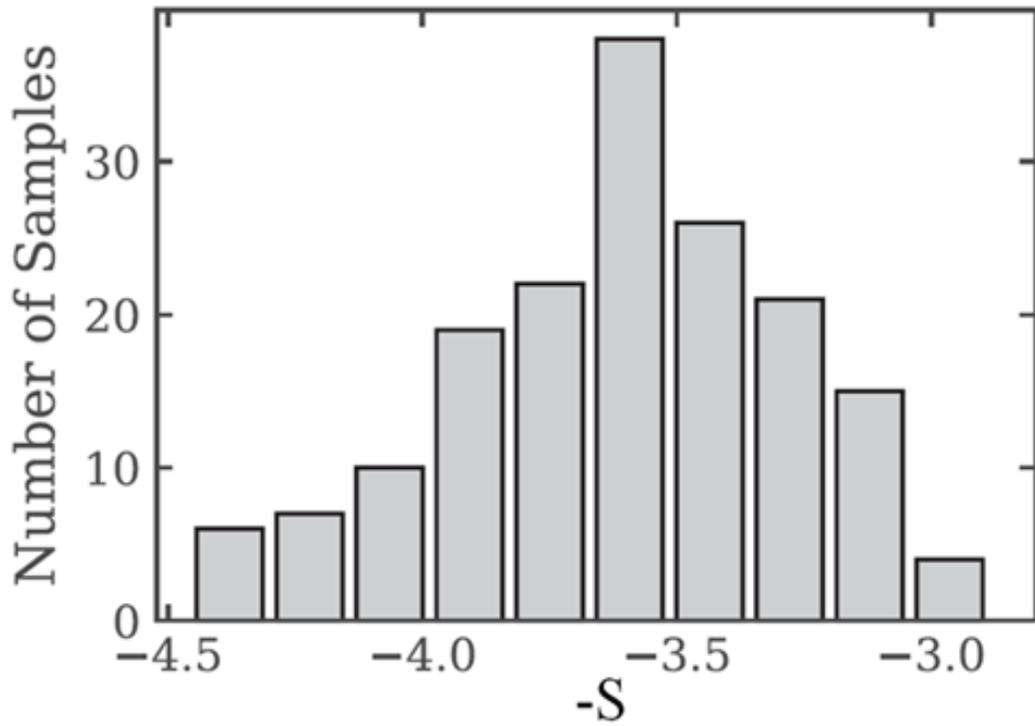


Figure 3: Distribution of the best-fit exponents of the power-law model of Eq. (equation7) when fit to the 168 PSDs of Fig. figure2. From Fig. 3 of Runyan et al. (2020). Reproduced by permission under a Creative Commons license.

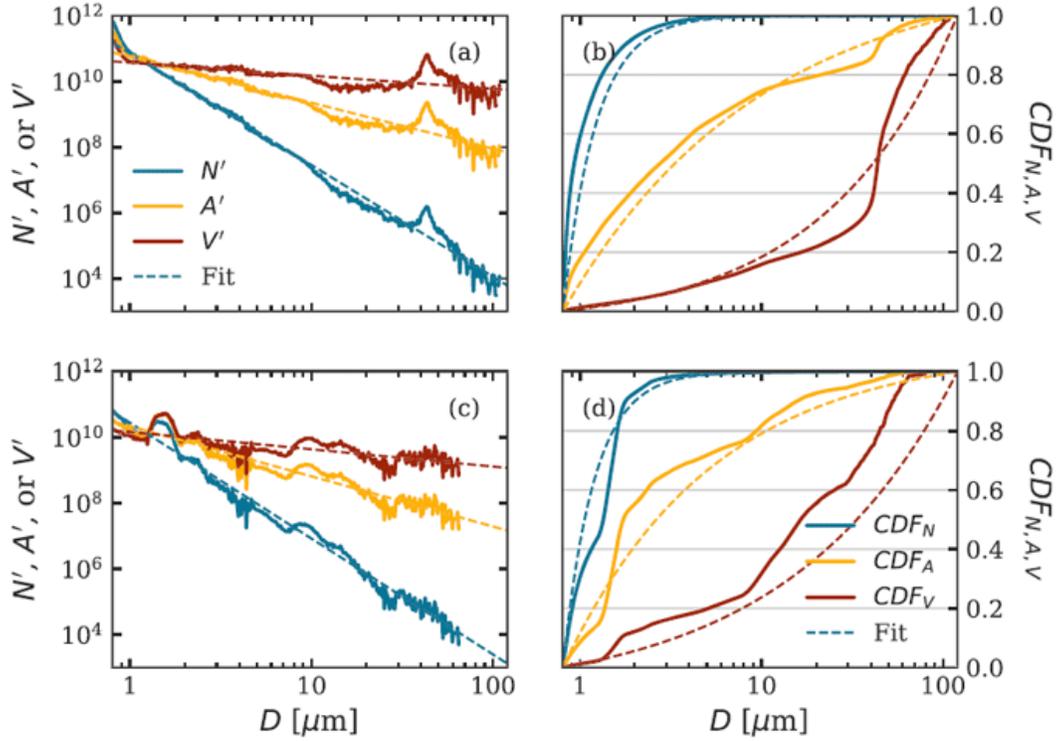


Figure 4: Two sets of PSDs from the deep chlorophyll maximum at two locations (top and bottom rows) in the Arctic. The left panels show the number ( $N' = n(D)$ ), area ( $A' = PSD_a(D)$ ), and volume ( $V' = PSD_v(D)$ ) distributions. The right panels show the corresponding cumulative distribution functions for number, area, and volume. The dashed lines show the power law distributions, based on the best-fit to the number distributions. Figure 5 of Runyan et al. (2020). Reproduced by permission under a Creative Commons license.

measured number PSDs for 18 classes and species of phytoplankton (VIRU is viruses, HBAC is heterotrophic bacteria, PROC is *prochlorococcus*, ..., MICA is *Prorocentrum micans*; see Stramski et al. (2001) for the full listing and description of these microbes). The bottom panel of the figure show the sum of the individual PSDs, with the concentrations of each class chosen within the ranges of typical values so that the total (except for viruses) obeys a Junge distribution ( $S = 4$  in Eq. (equation7)). Modeled distributions for detritus, minerals, and bubbles are also shown. These PSDs were used to model the inherent optical properties of oligotrophic waters; the total chlorophyll of the sum is  $0.18 \text{ mg Chl m}^{-3}$ .

The PSDs of individual species or classes of particles as seen in the upper panel of Fig. figure5 are often modeled by a log-normal distribution. In this distribution, the numbers are normally distributed when  $D$  is plotted on a logarithmic scale as in the preceding figures, or equivalently, using a normal distribution in  $\ln D$  rather than in  $D$ . That is,  $\ln D$  is distributed as a Gaussian probability distribution function:

$$pdf(\ln D) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln D - \mu)^2}{2\sigma^2}\right], \quad (8)$$

where  $\mu$  and  $\sigma$  are the mean and standard deviation of  $\ln D$ , respectively. Note that this  $pdf(\ln D)$  satisfies

$$\int_{-\infty}^{\infty} pdf(\ln D) d \ln D = 1,$$

as is required of any probability distribution function. Noting that  $pdf(D)dD = pdf(\ln D)d \ln D$  represents the same probabilities gives

$$pdf(D) = pdf(\ln D) \frac{d \ln D}{dD} = \frac{1}{D} pdf(\ln D),$$

which is the form seen in Eq. (A1) of Campbell (1995), and is equivalent to Eq. (2) of Ahmad et al. (2010). To express the log-normal distribution in terms of base 10 logarithms, the same procedure plus the observation that  $d \log D / dD = \log e / D$  gives

$$pdf(D) = \frac{\log e}{D} pdf(\log D) = \frac{1}{(\ln 10)D} pdf(\log D),$$

which is the form seen in Eq. (1) of Shettle and Fenn (1979). Now  $\mu$  and  $\sigma$  are the mean and variance of  $\log D$ . (Note that  $pdf(D)$  is defined for  $0 < D < \infty$ , which corresponds to  $-\infty < \ln D < \infty$  in Eq. (equation8).) Multiplying these probability distribution functions by magnitudes (e.g., the total number of particles per cubic meter) scales them for use as PSDs.

Log-normal distributions are commonly used to describe atmospheric aerosol size distributions (e.g., Shettle and Fenn (1979)). The current NASA atmospheric correction algorithm models aerosols as a sum of two log-normal distributions: one for small, continental (dust) aerosols and one for large, marine (sea salt) aerosols (Ahmad et al. (2010)):

$$n(r) = \frac{1}{r} \frac{dN(r)}{d \ln r} = \sum_{i=1}^2 \frac{N_{oi}}{r\sigma_i\sqrt{2\pi}} \exp\left[-\frac{(\ln r - \mu_i)^2}{2\sigma_i^2}\right]. \quad (9)$$

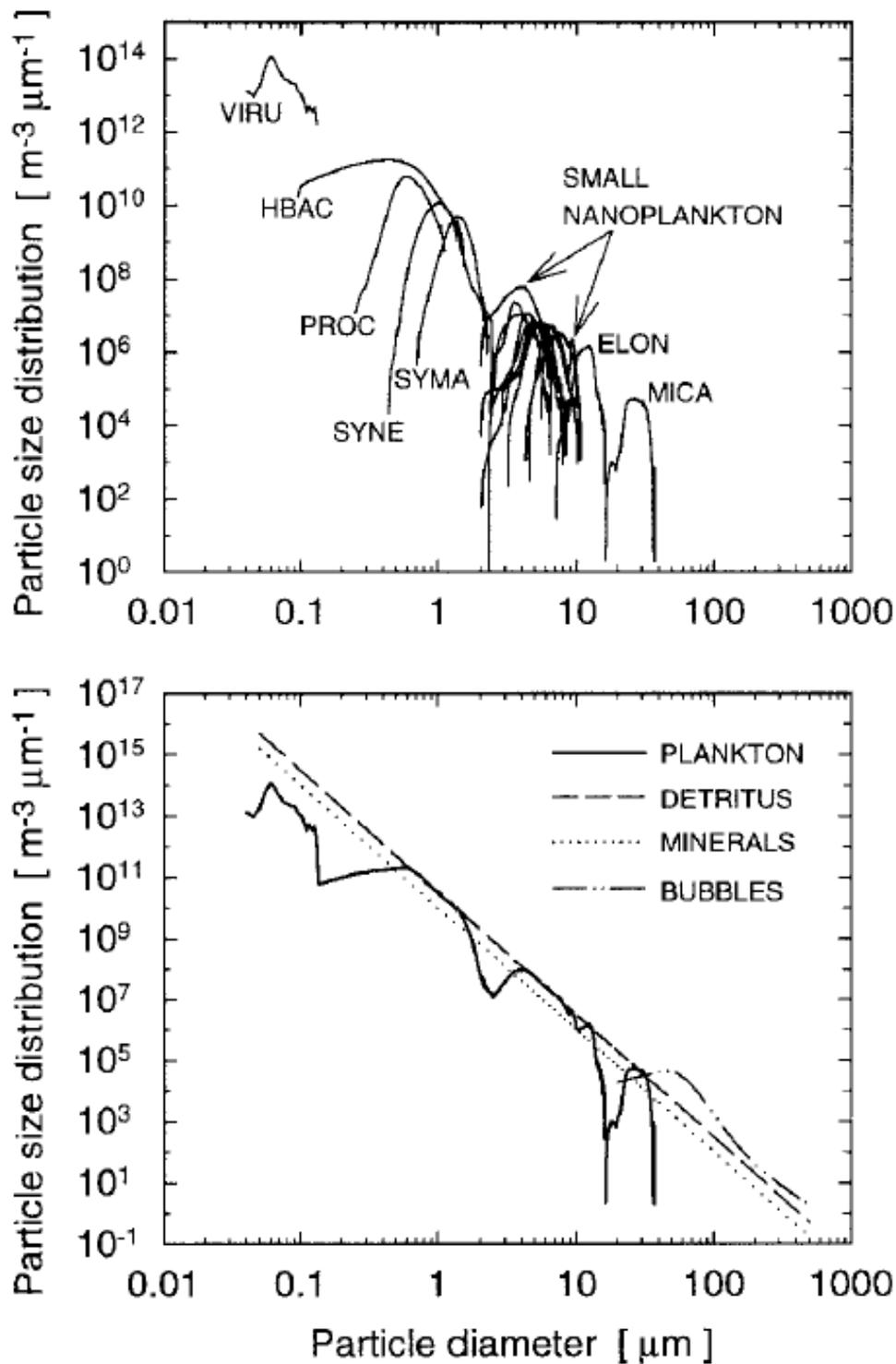


Figure 5: Top panel: measured number size distributions for 18 classes or species of microbes. Bottom panel: the sums of the microbial PSDs (solid line), and modeled PSDs for detritus, minerals, and bubbles. Reproduced from Fig. 9 of Stramski et al. (2001) by permission of the Optical Society of America under their Fair Use policy.

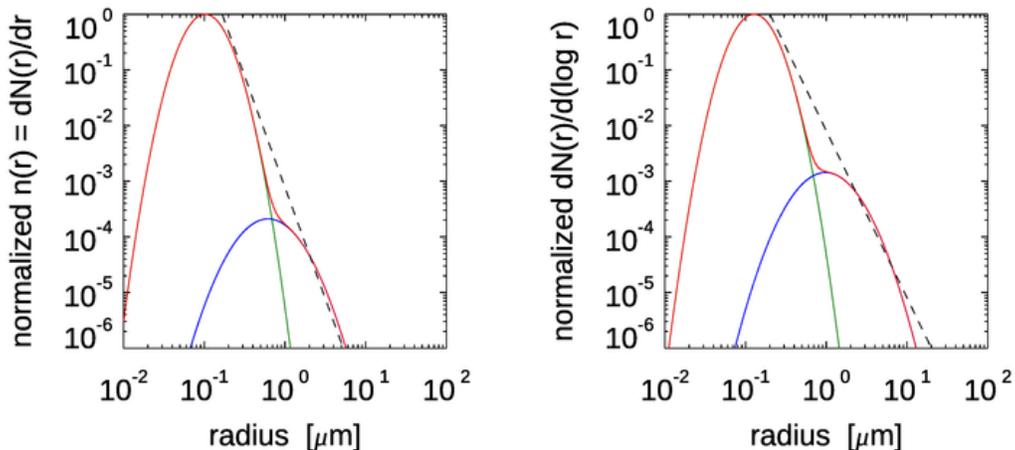


Figure 6: Typical open-ocean aerosol distributions plotted as  $dN(r)/dr$  in the left panel and as  $dN(r)/d\log r$  in the right panel. The green curve is the contribution by continental aerosols and the blue curve is the marine aerosols; the red curve is the total. The black dashed lines show the slope of a Junge distribution.

The parameters  $N_{oi}$ ,  $\mu_i$ , and  $\sigma_i$ ,  $i = 1, 2$  are adjusted to give the relative amounts of continental and marine aerosols, and the parameters also depend on the relative humidity. Figure illustrates these aerosol distributions for a typical open-ocean mix of aerosol types and for a relative humidity of 50%. The left panel plots the number PSD  $n(r) = dN(r)/dr$  and the right panel plots the same information as  $dN(r)/d\log r$ . In the left panel, the dashed line shows the  $-4$  slope of a Junge distribution (Eq. likesection6). In the right panel, the Junge distribution is represented by a slope of  $-3$  on the  $\log r$  abscissa sale (Eq. likesection5). The aerosol model of Eq. (equation9) uses particle radius  $r$  rather than diameter  $D$  because these particle size distributions are used as input to Mie calculations of aerosol optical properties.

The log-normal distribution finds many applications beyond particle size distributions. Campbell (1995) shows examples of log-normal distributions of chlorophyll concentrations, normalized water-leaving radiances, the diffuse attenuation coefficient  $K_d(490)$ , normalized photosynthetic yields, and several other biological variables. In fields other than oceanography, the log-normal distribution has been found to describe phenomena as diverse as the distribution of income (excluding billionaires), daily rainfall amounts, the populations of cities, and much more. Qualitative understanding of these observations can be obtained as follows. If a random variable is the *sum* of other random variables, then the distribution of the sum will be normal, regardless of the distribution of the individual variables. This is known as the Central Limit Theorem. (See the page on Error Estimation in Monte Carlo calculations for numerical examples.) If the total results from a *product* of random processes, then the product will obey a log-normal distribution. This is because the logarithm of a product is the sum of the logarithms, and the sum of the individual logarithms is then normally distributed. Thus, for example, if the total chlorophyll concentration is the result of a repeated daily fractional increase (daily growth rate) (e.g., the total Chl is the value at

day 1 times 1.1 to get the value at day 2, time 1.1 again to get the value at day 3, ...), the total chlorophyll concentration will be log-normally distributed.

## Comments on Units

Suppose you have some data  $Y(x)$  that you wish to model with a normal or Gaussian distribution.  $Y$  might be the number, area, or volume of particles, and  $x$  is some measure of their size. The mean and standard deviation of the data are  $m$  and  $s$ , respectively, and  $Y_t$  is the total “amount” of  $Y$  for all sizes (e.g., the total number of particles or their total volume). Then the Gaussian model of the data would be

$$Y(x) = \frac{Y_t}{\sqrt{2\pi}s} \exp \left[ -\frac{(x-m)^2}{2s^2} \right]. \quad (10)$$

If  $x$  has units, say  $\mu\text{m}$ , then  $m$  and  $s$  have the same units. The argument of the exponential is nondimensional, and

$$\int_{-\infty}^{\infty} Y(x) dx = Y_t$$

since a Gaussian distribution integrates to 1.

Now suppose you wish to model the same data using a log-normal distribution. The idea is that

$$Y(\log x) = \frac{Y_t}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(\log x - \mu)^2}{2\sigma^2} \right], \quad (11)$$

where now  $\mu = \log m$  and  $\sigma = \log s$  are the parameters of the distribution. However, *this equation is not correct because you cannot take the logarithm of a dimensional quantity.* (Nor can you compute  $e^x$  or  $\sin(x)$  or any other such function unless  $x$  is a nondimensional number.) However, you can compute  $\log(ax)$  if  $a$  has units of 1 over the units of  $x$ , so that  $ax$  is a nondimensional number. Thus, before using a log-normal distribution, you must non-dimensionalize the size  $x$ . If  $x$  is in  $\mu\text{m}$ , then this can be done by dividing all values of  $x$  by  $1 \mu\text{m}$ , i.e. setting  $a = 1/1 \mu\text{m}$ . However, you could just as well normalize the  $x$  values by dividing by  $17.3 \mu\text{m}$ . Note that converting the  $x$  data values from, say, micrometers to millimeters, or even to inches, also amounts to multiplying all  $x$  values by a scale factor  $a$ . If  $x \rightarrow ax$ , then  $m \rightarrow am$  and  $s \rightarrow as$ , and the exponential of the log-normal distribution does not change. The log-normal distribution is then properly written as

$$Y[\log(ax)] = \frac{Y_t}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(\log(ax) - \mu)^2}{2\sigma^2} \right], \quad (12)$$

where now  $\mu = \log(am)$  and  $\sigma = \log(as)$ . Of course, if  $x$  is measured in micrometers and  $a = 1/1 \mu\text{m}$ , then nothing changes numerically in Eq. (likesubsection11) vs. (likesubsection12), but it should be kept in mind when using equations like (equation8) or (equation9) that the size measure  $D$  or  $r$  must be non-dimensional. This subtlety seems never to be mentioned in the particle sizing literature, although the mathematicians have commented on this; see Matta et al. (2011) and Finney (1977).

## Single-parameter measures of particle size

A PSD contains the full information about the sizes of particles in a sample. However, a PSD is a function of size, and there is a human tendency to want to have a single number to describe the particle sizes. Using a single number to describe a distribution may sometimes be useful, but it often can be misleading or downright dangerous. As shown in the introduction, it is not even possible to define a unique “size” for more than one sphere, let alone a distribution of non-spherical particles. A power-law distribution diverges as the size goes to 0, so such a distribution can be integrated only over a finite size range  $D_1$  to  $D_2$ . Numbers like the mean or median particle size can be computed for a distribution, but the “mean” size will be different when based on a number distribution versus a volume distribution, for example.

Perhaps the most commonly used single-number size parameter is the Sauter mean diameter (SMD), which is defined by

$$SMD = \frac{\int_{D_1}^{D_2} D^3 n(D) dD}{\int_{D_1}^{D_2} D^2 n(D) dD}.$$

Recalling Eqs. likesubsection3 and likesubsection4, the SMD can be written as  $SMD = 6V_t/A_t$ , where  $V_t$  and  $A_t$  are the total volume and area of the particles in the distribution. For a single spherical particle of volume  $\pi D^3/6$  and area  $\pi D^2$ , the SMD reduces to the diameter of the sphere. Thus in words, the SMD is the diameter of a sphere that has the same total volume to total area ratio as the particles in the distribution.

Figure figure7 shows the differences in the mode, median, mean, and Sauter mean diameter for a log-normal probability distribution function,

$$pdf(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right],$$

with parameters  $\mu = 1.0$  and  $\sigma = 0.5$ . This illustrates that these measures of particle “size” are in general all different. The Sauter mean diameter is larger than the others and is determined most strongly by the largest particles of a distribution. If you are interested in total volume or mass, then the SMD is a useful quantity. This is why the SMD is used in sediment transport studies. On the other hand, if you are interested in the smallest particles, then the mode or median is probably a better statistic to use. Other single-number measures of a distribution have been developed for specific problems. Clearly, which single measure of a size distribution might be most useful depends on the problem at hand.

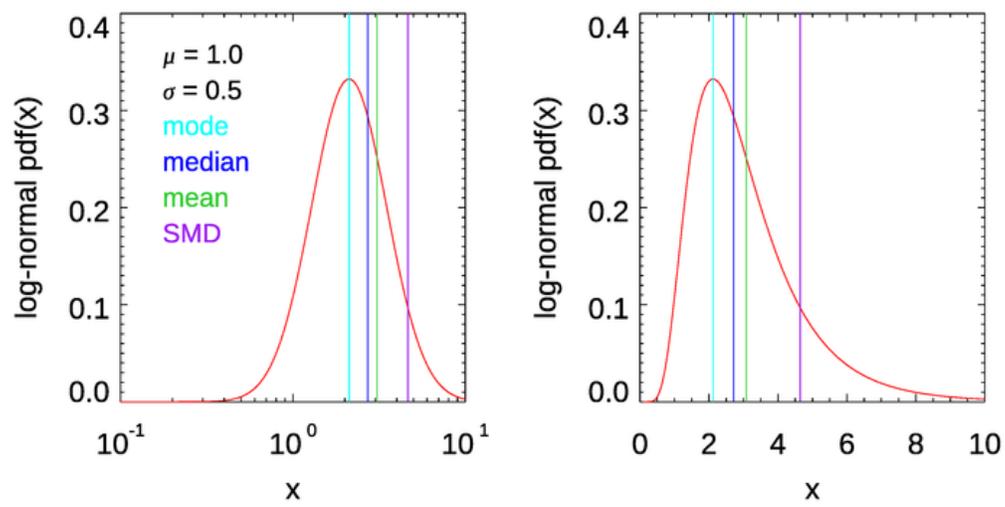


Figure 7: Differences in the mode, median, mean, and Sauter mean diameter for a log-normal distribution with parameters  $\mu = 1.0$  and  $\sigma = 0.5$ .