Image Cover Sheet

CLASSIFICATION	SYSTEM NUMBER
UNCLASSIFIED	499908

TITLE

APPROXIMATION OF THE SIZE DISTRIBUTION OF MARINE PARTICLES BY A SUM OF LOG-NORMAL FUNCTIONS

System Number:

Patron Number:

Requester:

Notes:

DSIS Use only:

Deliver to:

Limnol. Oceanogr., 41(4), 1996, 744–754

Approximation of the size distribution of marine particles by a sum of log-normal functions

Miroslaw Jonasz
M. Jonasz Consultants, 217 Cadillac St., Beaconfield, Québec H9W 2W7

Georges Fournier
Defence Research Establishment Valcartier, Courcelette, Québec G0A 1R0

Abstract

A simple algorithm is presented that decomposes the size distribution of marine particles into a sum of log-normal components of the 0th order. The algorithm was applied to 412 particle-size distributions in a particle-diameter range of ~0.5–200 μm measured by different researchers in various water bodies. A size distribution from this population may have from 1 to 6 log-normal components. The variability of the number of components reflects the variability in the shape of the size distribution and variations in the size range in which the data are available. The full-width-at-half-maximum of a component is approximately proportional to the peak diameter of the component. The maximum value of a component is approximately proportional to the inverse of the square of the peak diameter.

Two standard components of the marine particle-size distribution were identified. The peak diameter (D_{peak}), width parameter (σ), and maximum value (FD_{max}), of the first component are, respectively, 0.66 μm, 0.673, and 1.34×10^4 μm^{-1} cm^{-3}. These parameters for the second component are 10.5 μm, 0.366, and 3.3 μm^{-1} cm^{-3}.

The size distribution of particles suspended in seawater is essential information in many marine sciences ranging from optics to sedimentology. Modern particle-sizing instruments make it relatively easy to determine particle size distribution at a high particle-size resolution. Analyzing the resulting large data sets is greatly aided by a method of approximating the experimental data that reduces the large number of parameters of the original data set (each data point can be regarded as a parameter).

Besides reducing the number of parameters, a successful approximation technique has several advantages compared to the original representation of the data. Such an approximation technique can greatly simplify the classification of the size distributions (e.g. Kitchen et al. 1975) and give an insight into processes governing the formation of the distribution via linking the parameters of the approximation to physical models of such processes (e.g. Kiefer and Berwald 1992). An approximation technique that yields an analytical representation of the distribution makes possible analytical predictions of other physical properties of the particle population (e.g. the attenuation of light, Casperson 1977) as functions of the approximation parameters of the size distribution.

Several functional representations of the size distribution of marine particles have been used to date. These representations include *the power law* (Bader 1970, also called the Junge distribution after Junge 1963), *log-normal distribution* (Jonasz 1987; Lambert et al. 1981), *Weibull distribution* (Carder et al. 1971), *gamma function* (Risovic 1993; Ulloa et al. 1992), and *characteristic vectors* (Jonasz 1983; Kitchen et al. 1975, also referred to as principal components).

Even a casual visual examination of the complex size distributions characteristic of biologically active waters (e.g. Jonasz 1983) leads to the conclusion that a simple function will not accurately approximate details of such distributions. Such size distributions frequently contain characteristic intermediary maxima linked to the presence of dominating phytoplankton populations (Chisholm 1992; Hood et al. 1991; Jonasz 1983) or to the dynamics of the populations of nonliving particles (e.g. McCave 1983). Thus, a simple function, such as the power-law function, cannot realistically express such complex distributions at a fine size resolution, although the function can express a first-order approximation at a global size scale. In fact, the power-law function has been shown to be well founded in the dynamics of marine particles (Kiefer and Berwald 1992; Platt and Denman 1977).

Jonasz (1983, 1980) compared several functional representations of the size distribution of marine particles and suggested that a combination of the power-law function and a Gaussian function (or a sum of Gaussian functions) may be a versatile representation of the complex size distributions observed in biologically active surface waters of the ocean (e.g. Sheldon et al. 1972). Recently, Risovic (1993) presented an alternative composite model of a marine particle size distribution that is based on the combination of two generalized gamma functions. Van Andel (1973) and Spencer (1963) indicated that a sum of log-normal functions may describe particle size distribution in sediments, but they pointed out computational difficulties involved in decomposing a particle size distribution into a sum of log-normal components.

Fitting a combination of nonlinear functions, such as

Acknowledgments

This research was supported by contract W7701-1-4528/01-XSK of the Defence Research Establishment Valcartier.

We thank the anonymous reviewers for their comments.

those previously mentioned, to a particle size distribution leads in general to a nonlinear minimization problem, usually with many local minima. A successful application of a nonlinear fitting algorithm to such a problem requires realistic initial values of the parameters.

In this paper, we present a simple algorithm (M. Jonasz Consultants unpubl. rep.) that permits automated decomposition of a particle size distribution into a sum of log-normal components. This algorithm reduces the fitting to a global minimization problem with no dependency on the initial values of the parameters. Once the values of the parameters are derived by means of this algorithm, such values can in principle be used as initial guesses in nonlinear fitting algorithms.

We also present and analyze the results of applying this algorithm to several hundred size distributions from a database compiled by Jonasz (M. Jonasz Consultants unpubl. rep.).

Rationales for choosing the log-normal function

Of the functions considered, the log-normal function attracted our attention. Such a function was used to model light scattering by suspended particles (Casperson 1977; Kerker 1969). The log-normal distribution has three important features: a finite value for any value of the argument, a limit of 0 at the zero value of the argument, and it can assume any width at any peak diameter >0.

The log-normal function has been used extensively in geology (Kranck and Milligan 1991; Jonasz 1987) to describe the size distribution of sediment particles. The log-normal function has also been used in biological oceanography to describe the size distribution of marine bacteria (Morel and Ahn 1990). The size distributions of many marine organisms in cultures also resemble the log-normal function (e.g. Stramski and Reynolds 1993).

The log-normal function has been linked to the random fragmentation process (Middleton 1970) in which the probability of fragmentation is independent of particle size (Tenchov and Yanev 1986). If the probability of fragmentation is proportional to particle size, the Weibull distribution is more appropriate. However, as pointed out by Tenchov and Yanev (1986), the difference between a log-normal distribution and a Weibull distribution can be minimized by appropriately selecting the adjustable parameters. Such difference may be difficult to discern at the measurement precision characteristic of the techniques of particle size analysis applicable to marine particles.

The present algorithm uses the 0th-order log-normal functions to represent components of a complex size distribution. We will refer hereafter to the 0th order log-normal function as the log-normal function. The log-normal function, through logarithmic transformation, leads to a polynomial that can be fitted to a data set by a simple least-squares procedure that searches for the *global* minimum of the approximation error. The gamma function, although it can also support this approach, leads to a more complex fitting algorithm. Also, the approximation errors of the marine size distributions obtained when the gamma function is used as an approximating function are greater than those characteristic of other functions (Jonasz 1980).

In addition to bringing simplicity to the fitting algorithm, the log-normal function provides a reasonable 2nd-order approximation to modeling of the natural curvature of the particle size distributions (with the multisegment power law being the first-order approximation). Thus, there is no need to "break" the power-law approximation to reflect the changes in the slope of the distribution (e.g. Jonasz 1983). The log-normal function leads to mass and area distributions that also are log-normal (e.g. Kerker 1969). Finally, the log-normal function reduces to the power-law function when the quadratic coefficient of the log-normal function reduces to 0.

Log-normal function

The log-normal size distribution is defined by Casperson (1977) in accordance with the historically accepted notation for the log-normal distribution of the nth order (Kerker 1969):

$$FD_n(D) = FD_{\max,n} D^n \exp[-(\ln D - \ln D_{\text{peak},n})^2/(2\sigma_n^2)]. \quad (1)$$

D is the particle diameter, n is the order of the distribution, $FD_{\max,n}$ is the maximum value of the function, and σ_n is the width parameter. (Units given in list of notation.) We will restrict ourselves to the 0th order ($n = 0$) and drop the n subscript from now on. By taking the logarithm of both sides of Eq. 1, with $n = 0$, and performing simple algebraic transformations, we have

$$\log FD(D) = B_0 + B_1 \log D + B_2 \log^2 D. \quad (2)$$

$$B_0 = \log FD_{\max} - \log^2 D_{\text{peak}} \ln 10/(2\sigma^2), \quad (3a)$$

$$B_1 = \log D_{\text{peak}} \ln 10/\sigma^2, \quad (3b)$$

and

$$B_2 = -\ln 10/(2\sigma^2). \quad (3c)$$

Notation

D	Particle diameter (μm)
FD	Particle size distribution (μm^{-1} cm^{-3})
$FD_{\max}$	Maximum value of the 0th-order log-normal function (μm^{-1} cm^{-3})
D_{peak}	Value of the argument of the 0th-order log-normal function for which the function attains its maximum value (μm)
σ	Width parameter of the 0th-order log-normal function
B_0, B_1, B_2	Log-normal coefficients, defined in Eq. 3a–c
D_1	Smallest diameter for which a log-normal component of size distribution FD assumes the value of half its maximum value (μm)
D_2	Largest diameter for which a log-normal component of size distribution FD assumes the value of half its maximum value (μm)
f	$= (2\ln 2)^{1/2}$

Equation 3a and b can be solved for FD_{max}, D_{peak}, and σ:

$$\log FD_{max} = B_0 - B_1^2/(4B_2), \quad (3d)$$

$$\log D_{peak} = -B_1/(2B_2), \quad (3e)$$

and

$$\sigma^2 = -\ln 10/(2B_2). \quad (3f)$$

The particle size distribution (FD) we model here is the *frequency* size distribution, defined by the relationship $dN = FD(D)\, dD$, where dN is the number of particles with diameters in a range of $(D, D + dD)$.

From Eq. 1, we see that the log-normal size distribution has an extremum of FD_{max} at a diameter $D = D_{peak}$. The width of the distribution is determined by the nondimensional parameter σ, which is related to the ratio of the maximum and minimum diameters of the full-width-at-half-maximum of the 0th-order log-normal function through

$$D_2/D_1 = \exp(2f\sigma). \quad (4)$$

$$f = (2 \ln 2)^{1/2} \approx 1.177, \quad (5a)$$

$$D_1 = D_{peak}\exp(-f\sigma), \quad (5b)$$

and

$$D_2 = D_{peak}\exp(f\sigma). \quad (5c)$$

D_1 and D_2 fulfill an inequality, $D_1 < D_2$, and are the diameters at which $FD(D)$ assumes half the value of the maximum $FD = FD_{max}$.

We consider only functions defined by Eq. 2 that are characterized by $B_2 < 0$. This condition ensures that the extremum of the function is a maximum.

Outline of the algorithm

The algorithm is based on an assumption, supported by biological arguments, that the number size distribution of marine particles, $FD(D)$, is composed of a cascade of log-normal components. Particle size distribution data for both natural (Jonasz 1983) and cultured marine particles (Stramski and Reynolds 1993; Morel and Ahn 1990) provide experimental evidence that each such component may represent a population of particles.

According to the underlying assumption, each log-normal component of the size distribution dominates in a particular size interval. Thus, if that size interval is somehow identified, one could determine the parameters of the respective log-normal component (e.g. by using the least-squares fitting procedure for the log-log transformed original data). In our algorithm, the size interval in question is identified by repeatedly scanning the data set with a window of variable width.

The algorithm consist of three steps. In *step 1*, the log-normal function is fitted to all data in the set. This is the average fit.

The following steps are repeated for the starting window width incremented from a minimum to a maximum value. In *step 2*, the data set is repeatedly scanned with a window. The first scan is performed for the preset starting window width. After each scan, the window width is incremented by unity. If a log-normal component is found during a scan, it is subtracted from the input data set, and the difference set becomes the new input data set. Only those differences that are each greater than a fixed fraction of the previous version of the respective data are retained. Thus, after a component has been identified, the number of data points in the updated data set usually decreases. The scanning normally ends when all window widths have been tried on all data points. The scanning is prematurely stopped when there are insufficient data left or the number of components found exceeds a preset maximum number of components.

In *step 3*, the set of components is sorted according to the order of the decrease in the approximation error of the original data caused by addition of the component. The only components retained are those that significantly reduce the total approximation error according to the Fisher test (e.g. Hudson 1964). The sorting of components is meaningful only within a framework of a particular size distribution. This algorithm does not generate a set of standard components for all examined distributions, such as those obtained with a characteristic vectors algorithm. However, the components generated by the present algorithm may indeed by standard components if the size distributions examined are originally composed of such components. We discuss this possibility later. If no components are found, the average fit is used.

In this manner, a series of component sets is generated for each size distribution, and each set is for a different value of the starting window width. The first set consists of a single component: the average fit. Each set is characterized by the error with which is approximates the original data. The selection of the best set can be manual or automatic. The criterion of the automatic selection is the minimum of the absolute value of the relative approximation error per data point. This error is determined as

$$\text{relerr} = (1/N)[\Sigma_1^N(FD_{a,i} - FD_i)^2/FD_i^2]^{1/2}. \quad (6)$$

N is the number of data points, subscript a denotes the approximation of the original data by all components, and subscript i numbers the data points. A computer program implementing the algorithm in FORTRAN is available for a nominal fee from M. Jonasz.

Examples of the application of this algorithm to noiseless and to noisy data are shown in Tables 1–3. The noisy data example (Table 2) is also illustrated in Fig. 1 for the starting window width of 11 data points. After the first approximating component (1, Fig. 1A) is found in the input data ("in," Fig. 1A), it is subtracted from the those data. Only those differences ("out," Fig. 1A) are retained that are greater than a preset fraction (10^{-10}) of the input data. Because of the low, arbitrarily set, numerical value of that fraction, virtually all positive differences are retained. The retained differences become the input data for the next round of scans ("in," Fig. 1B). After the second component is found (2, Fig. 1B), the new set of differences is created ("out," Fig. 1B). This procedure

Table 1. The data used in Fig. 1, the algorithm for retrieval of log-normal components of the size distribution. The *noiseless* data were obtained by summing two log-normal functions that resemble typical components of the marine particle size distribution. Parameters of these components are shown in Table 2. The *noisy* data were created by adding ±50% random noise to the noiseless data. The magnitude of the noise corresponds to the measurement error at a relatively low concentration of 5 cm^{-3} of particles with diameters $D > 10$ μm. Because the relative noise is approximated by the inverse square root of the particle concentration (Jonasz 1983) and the latter varies approximately as an inverse power of the D, the noise at $D < 10$ μm is overestimated and at $D > 10$ μm is underestimated. All values given in arbitrary units.

Particle diam (μm)	Noiselesss data	Noisy data
2.00	5.950E-02	6.664E-02
2.52	2.320E-01	1.628E-01
3.17	7.382E-01	5.846E-01
4.00	1.907E+00	2.175E+00
5.04	3.993E+00	3.052E+00
6.35	6.771E+00	5.039E+00
8.00	9.308E+00	1.151E+01
10.08	1.039E+01	8.940E+00
12.70	9.482E+00	1.029E+01
16.00	7.163E+00	9.505E+00
20.16	4.626E+00	2.758E+00
25.40	2.738E+00	1.671E+00
32.00	1.665E+00	2.317E+00
40.32	1.139E+00	1.529E+00
50.80	8.560E-01	5.513E-01
64.00	6.487E-01	7.588E-01
80.63	4.676E-01	6.023E-01
101.59	3.129E-01	1.990E-01
128.00	1.931E-01	2.374E-01
161.27	1.097E-01	8.523E-02
203.19	5.734E-02	8.074E-02
256.00	2.757E-02	3.655E-02
322.54	1.219E-02	1.301E-02
406.37	4.962E-03	3.472E-03
512.00	1.858E-03	1.125E-03
645.08	6.397E-04	8.437E-04
812.75	2.027E-04	1.276E-04

ends after the third component is found (3, Fig. 1C) because the number of data points left after that scan is insufficient.

Although in this example three components were created, only the first two were retained. The addition of the third component would have worsened the approximation. The first two components (Fig. 1D) passed the Fisher test and were included in the component set for the present value of the starting window width (11 points).

The approximation with two components was found to be better (Table 3) than the average fit (relerr = 0.1647). Out of the component sets for all starting window widths (ranging from 4 to 12), the starting window width of 4 points provided a slightly better approximation (relerr = 0.0594) with 5 components than the next best (relerr = 0.0627, starting window width of 11 points) with 2 components. However, given the experimental error of 0.0851 (Eq. 6), the approximation for the starting window width of 4 points is not significantly better than that for 11 points but is more likely to fit the noise in the data. At present, the experimental errors of the size distribution are not accounted for in selecting the best approximation because such errors are not specified for most of the data used in this study.

Generally, if the number of components becomes independent of the starting window width (Table 2), it is a sign that the right components have been found. In such a case, the parameters of the retrieved components tend to stabilize, especially those of the first retrieved component. However, if a component set was retrieved with a large starting width, components whose size scale is smaller than the starting window width will not be retrieved.

As seen in the noisy data example, the multiscan procedure assures that the best *overall* combination of components is found, and the possibility of spurious fits to the high-frequency noise in the data is minimized. However, such possibilities are not eliminated. Specifically, although subtraction of a component is not likely to affect data outside the immediate influence of the component (Fig. 1A), the relative noise of the difference can increase by an order of magnitude in the overlap region of two (or more) components (*see Fig. 1B*). In Fig. 1B, the second component has been identified by means of the rightmost group of data because the differential data in the overlap region of the first and the second components were too noisy. Thus, the second component was determined with a less than ideal data set and its parameters are less accurate than those of the first component.

Input data

We used the particle size distributions compiled in a computer-readable database of particle size distributions (M. Jonasz Consultants unpubl. rep.). That database contains 433 particle-size distributions measured by 12 researchers in various regions of the Atlantic, Pacific, and Indian Oceans, and in the Baltic Sea for various seasons and at various depths. Most of these size distributions (412) were obtained with Coulter counters. For consistency, only those size distributions are used in this study.

Results and discussion

The components of a particle size distribution are derived with the present algorithm purely from analyzing numerical data. Thus, the algorithm cannot attach any geological or biological interpretation to a component. However, the results of this study suggest that such components may have potentially broader meaning than do convenient mathematical constructs.

Note that the components are identified within the global pool of components of all examined size distributions only by the component parameters and not by the component order numbers used for a shorthand identification within a set of components determined for a particular size distribution. Thus, two components ob-

Table 2. Comparison of parameters of log-normal components of the noiseless data (Table 1) with parameters of components retrieved from *noiseless* and *noisy* data (Table 1) by means of the present algorithm. For the *noiseless data,* the minimum approximation error (Eq. 6) of 0.00013 occurs for three components at the starting scan window width (minpts) of 4 and 6 data points. However, in each of these two cases the third component is negligibly small and the parameters of the other two components are almost identical in all cases except minpts = 12. This latter window width is too large to properly resolve the components. The retrieved values are listed here for minpts = 10 (the next lowest approximation error of 0.00014). Parameter errors (1 SD) were calculated by the algorithm. The other forms of log-normal parameters are also given (calculated from Eq. 3a–c). For the *noisy data,* the parameters obtained for minpts = 11 (relerr = 0.0627) are shown. These parameters are also used to plot the retrieved components in Fig. 1D. All parameters are within 2 SD of their original values.

			Noiseless data		Noisy data	
Comp.	Parameter	Original value	Value	Error	Value	Error
1	B_0	−3.6051	−3.6053	0.00001	−3.9888	0.2770
	B_1	9.2103	9.2107	0.00003	9.9904	0.6626
	B_2	−4.6052	−4.6054	0.00002	−4.9806	0.3625
2	B_0	−3.9250	−3.9228	0.0006	−7.9394	2.2163
	B_1	5.3144	5.3123	0.0005	9.0570	2.0370
	B_2	−1.7989	−1.7984	0.0001	−2.6584	0.4604
1	FD_{max}	10	9.9994		10.4965	
	D_{peak}	10	9.9976		10.0677	
	σ	0.5	0.5000		0.4808	
2	FD_{max}	1	1.0005		0.5953	
	D_{peak}	30	29.9871		50.5206	
	σ	0.8	0.8001		0.6581	

tained for two different size distributions and bearing the same order number do not necessarily correspond to each other. If we refer to the component by its number, it is always within the framework of a particular size distribution, the only exception being the standard components of marine size distribution discussed later. The results discussed here apply to components selected by the program on the basis of the minimum relative approximation error of the size distribution by a combination of such components.

A typical example of the approximation of the particle size distribution by a single log-normal component is shown in Fig. 2. If the power-law function were used to approximate that distribution, at least two power-law segments would have been required to represent the curvature of the distribution.

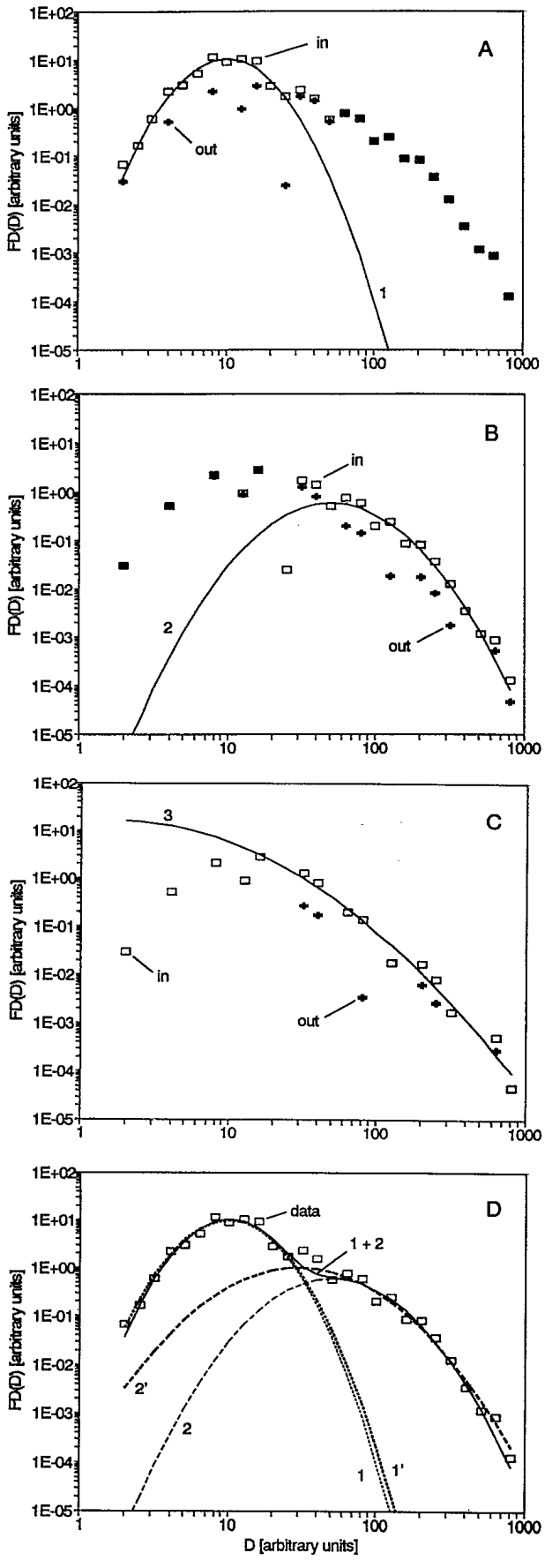
For a complex size distribution, several log-normal components are necessary to approximate the size distribution (Fig. 3). In fact, if a particle size distribution can be represented by only one component, it is sometimes the result of a limited range of particle size. Such a limited size range of the size distribution may also lead to spurious approximations.

The first component shown in Fig. 3 is almost identical in shape and magnitude to the size distribution shown in Fig. 2. The component with the smallest peak diameter in Fig. 3 could correspond to the size distribution of submicron particles measured by Longhurst et al. (1992) in Atlantic waters off Halifax, Nova Scotia, although the width of the size distribution of Longhurst et al. is a fourth that of the component from Fig. 3. The similarities between the components of the size distributions measured by different researchers in different waters and different seasons suggest that the present algorithm is not merely a convenient approximation procedure, but could provide a deeper insight into processes that create the size distribution of marine particles. We discuss this possibility later.

The present algorithm was applied to 412 particle-size distributions and generated a total of 902 components. A small fraction (~5%) of these components were so close

Table 3. The relative errors (Eq. 6) of approximation of the noiseless and noisy data by sets of log-normal components for different starting widths (minpts) of the scan window. The number of retrieved components—cpts.

	Noiseless		Noisy	
minpts	data	cpts	data	cpts
4	0.000133	3	0.059	5
5	0.000145	2	0.100	3
6	0.000133	3	0.102	2
7	0.000144	2	0.086	2
8	0.000144	2	0.086	2
9	0.000141	2	0.082	2
10	0.000140	2	0.095	2
11	0.000146	2	0.063	2
12	0.000763	2	0.142	1



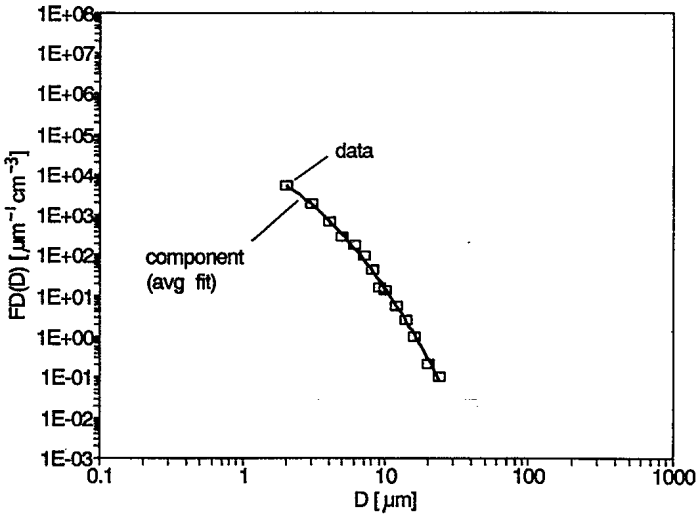


Fig. 2. A particle size distribution represented by a single log-normal component (the average fit). The data were obtained by M. Jonasz at 31.5°N, 15.8°W (mid-Atlantic), 6-m depth, on 13 April 1978 (file JONATL78.P21, col 3). Parameters of the log-normal component: $B_0 = 4.217$; $B_1 = -1.057$; $B_2 = -2.016$; relerr = 0.049.

in shape to the hyperbolic function that their maximum values (FD_{max}) were extremely large. We neglected the hyperbolic components in examining the relationships between the parameters of the components. In neglecting a component, we used the criterion

$$|B_1/B_2| > 10. \quad (7)$$

If this criterion is fulfilled, the error of approximation of a log-normal component with a hyperbola is on the order of the error of approximation of the typical particle size distribution by the log-normal function. After this criterion was applied, we had 853 components to consider. Because only a small fraction of all components could be classified as hyperbolic components, the hyperbolic shape of the size distribution of marine particles seems to be an exception rather than a rule.

By substituting criterion 7 (Eq. 7) in the definition of the component's peak diameter (Eq. 3e), a value of 0.01 nm is obtained, which is smaller than the size of the water

←

Fig. 1. Illustration of the principles of the log-normal decomposition algorithm as applied to a noisy data set from Table 1. This composite figure shows the component retrieval process for a starting scan window width of 11 data points. In panels A–C, "in" data are those used to retrieve a component (marked by a number), and "out" data are the difference data set created by subtracting the component shown in the panel from the in data. The out data set of one panel becomes the in data set of the next panel. In panel D, the components retained by the algorithm for the final representation of the size distribution are shown (1 + 2) along with the original noisy data set (data) and the original noiseless components (1′ and 2′). The components retrived from the noiseless data set are essentially identical with the original components. Refer to Tables 2 and 3 for parameter values and other details.

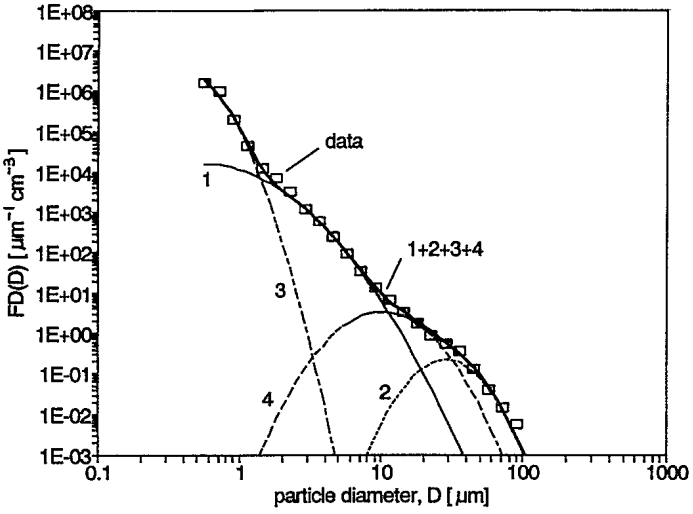


Fig. 3. A particle size distribution represented by four log-normal components. These components were determined with the starting scan window width of 6 points (*see text*). The data were obtained by K. Kranck and T. Milligan at 48.80°N, 64.58°W (northwestern Atlantic), 1-m depth, on 2 May 1986 (file KRAATL86.P01). Component 1 is almost identical in shape and magnitude to the size distribution shown in Fig. 2. Parameters of the log-normal components (the second subscript denotes component number): $B_{0,1} = 4.054$; $B_{1,1} = -1.164$; $B_{2,1} = -2.116$; $B_{0,2} = -17.039$; $B_{1,2} = 22.523$; $B_{2,2} = -7.745$; $B_{0,3} = 4.965$; $B_{1,3} = -7.042$; $B_{2,3} = -7.299$; $B_{0,4} = -4.267$; $B_{1,4} = 9.599$; $B_{2,4} = -4.807$; relerr = 0.038.

molecule (0.3 nm). Thus, criterion 7 identifies log-normal components representing, in totality, essentially all particles present in a volume of seawater.

The histograms of the parameters of the components are shown in Figs. 4–7. The histogram of a component's width (Fig. 6) is represented by the ratio $D_2 : D_1$ of the maximum-to-minimum limit of the full-width-at-half-maximum of a component. The $D_2 : D_1$ axis is stretched exponentially as compared to a σ axis (Eq. 4), but the single-modal character of the histogram is not affected.

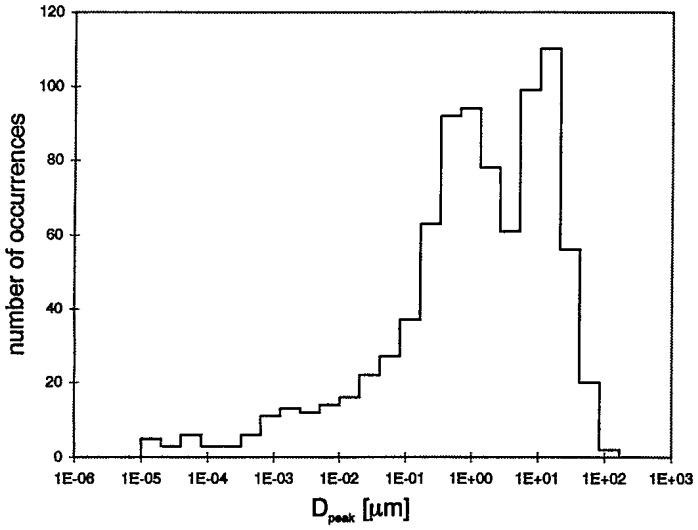


Fig. 4. Histogram of peak diameter, D_{peak}, of a component.

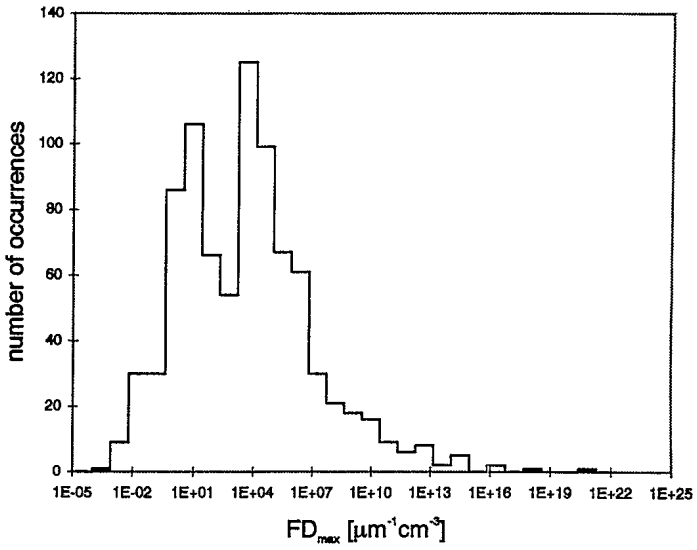


Fig. 5. Histogram of the maximum value, FD_{max}, of a component. Only the most significant range of FD_{max} is shown.

Both the histogram of the peak diameter (Fig. 4) and the histogram of the maximum value (Fig. 5) have bimodal structures. Each peak in one histogram has a counterpart in the other, as seen in the histogram (Fig. 7) of the pair (D_{peak}, FD_{max}). The counterparts permits us to identify two *standard* components (Table 4) of the marine particle size distribution. Components resembling these standard components can be found (Figs. 8, 9) in the two size distributions previously discussed (Figs. 2, 3). Out of 853 log-normal components, 141 components have a peak diameter of between 0.328 and 0.927 µm, a range

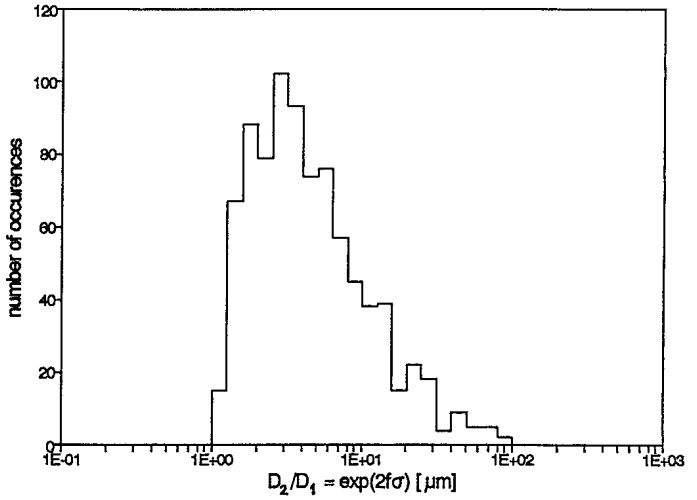


Fig. 6. Histogram of the logarithmic width of a component, D_2/D_1, where D_1 and D_2 are the particle diameters at which the size distribution of the component attains the value equal to half of the maximum value. The constant f equals $(2\ln 2)^{1/2}$.

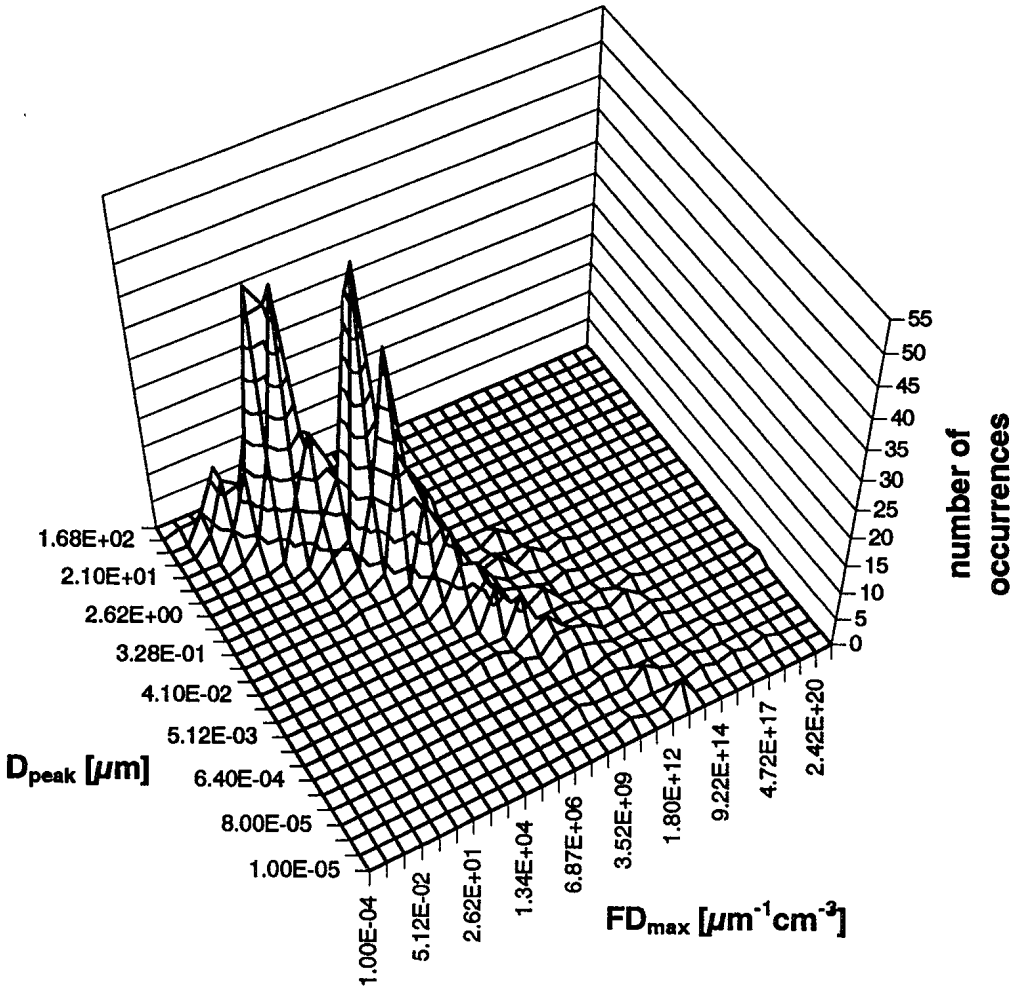


Fig. 7. Histogram of the pair D_{peak} and FD_{max} showing that two peaks in each of the histograms of D_{peak} and FD_{max} correspond to each other, and identifying two standard components of the marine particle size distribution. The histogram is plotted as a surface, S, such that $S(D_{peak,i}, FD_{max,i})$ represents the number of occurrences of the pair with values in a rectangular region defined by the corner values $(D_{peak,i-1}, FD_{max,i-1})$ and $(D_{peak,i}, FD_{max,i})$.

that includes the peak diameter of standard component 1. A total of 152 components have a peak diameter of between 5.24 and 14.8 μm, a range that includes the peak diameter of standard component 2.

Can these standard components be artifacts of the measurements of the size distribution with the Coulter counter? The particle size measurements are performed in the Coulter counter, and in other particle counters based on

Table 4. The parameters of two standard components of the marine size-particle distribution. The maximum value, FD_{max}, and peak diameter, D_{peak}, have been estimated from the histogram; the width parameter, σ, has been calculated from the regression $\sigma = \sigma(\ln D_{peak})$.

Parameter	Component 1	Component 2
FD_{max}	13,400	3.28
D_{peak}	0.65	10.5
σ	0.673	0.366

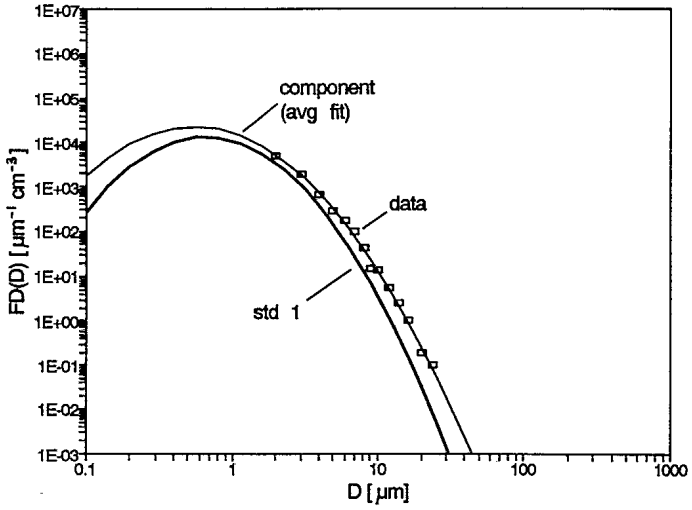


Fig. 8. Comparison of standard component 1 (Table 4) of the marine particle size distribution with the size distribution (data) and its component from Fig. 2.

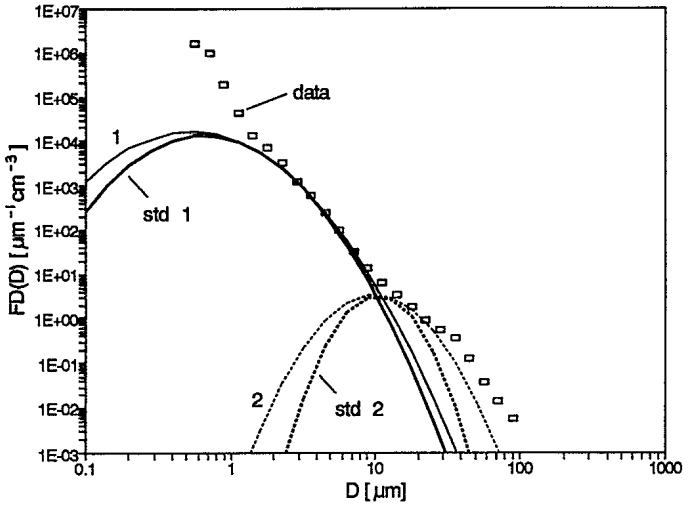


Fig. 9. Comparison of standard components 1 and 2 (Table 4) of the marine particle size distribution with the size distribution from Fig. 3 (data) and its retrieved components (1 and 2).

the electro-resistive principle (e.g. Milligan and Kranck 1991), by passing particles one after another through a small orifice in a nonconductive membrane that separates two chambers containing electrolyte. Constant electric current flows through the orifice. When a particle passes through the orifice, the resistance of the current path through the orifice increases momentarily, generating a voltage pulse. The amplitude of this pulse, at a fixed amplification, current, and orifice diameter, is proportional to the particle volume.

This proportionality is preserved as long as the particle diameter is within ~2–40% of the orifice diameter. Thus, several orifices with different diameters are required to cover a wide particle diameter range. The standard orifice sizes are 30, 70, 100, 140, 200, 280 μm, etc. These orifices can be used to size particles in diameter ranges of 0.6–12, 1.4–28, 2–40, 2.8–56, 4–80, and 5.6–112 μm. Some mismatch may occur when the orifice is changed (e.g. Milligan and Kranck 1991), in part because the locations and size ranges of the particle counting bins usually change.

However, many components with peak diameters of ~0.6 μm (the diameter of standard component 1) have been retrieved from data obtained with a *single* 100-μm orifice. The data in question are those of M. Jonasz from the Atlantic Ocean and the Baltic Sea (database files JONATL*.* and JONBAL*.*, where * denotes a wildcard). Such an orifice (particle diameter range of 2–40 μm) also covers the diameter range of standard component 2 whose peak diameter is 10.5 μm. Size distributions with components having peak diameters of ~10 μm about the peak diameter of standard component 2) also have been measured with a single orifice (280 μm; Hood et al. 1991; files HOOPAC91.* in the database). Thus, we found no evidence that the standard components of the size distribution are artifacts of the measurement process.

Establishing the existence of standard components of marine size distribution that are identifiable by their peak diameters, and thus can potentially be associated with populations of specific marine particles, is an important result of this study. Although standard components of the marine size distribution can be generated by other techniques (e.g. by the method of characteristic vectors, Jonasz 1983), such components lack the physical and biological interpretation available for components generated with the present algorithm.

The statistics of and correlations between the parameters of 853 log-normal components of the 412 particle-size distributions are shown in Tables 5 and 6. The correlation matrix for the B coefficients representing all 902

Table 6. Correlations, expressed with r^2, between the parameters of log-normal components of 412 size distributions. The correlations between the coefficients B_0, B_1, and B_2 are greater than those between the parameters FD_{max}, D_{peak}, and σ of the log-normal distributions because of the smoothing effect of the logarithmic transformation.

	$\ln FD_{max}$	$\ln D_{peak}$	σ
$\ln FD_{max}$	1.000	0.873	0.478
$\ln D_{peak}$	0.873	1.000	0.760
σ	0.478	0.760	1.000
	B_0	B_1	B_2
B_0	1.000	0.963	0.840
B_1	0.963	1.000	0.942
B_2	0.840	0.942	1.000

Table 5. Statistics of the parameters of 853 log-normal components of 412 size distributions. $D_2/D_1 = \exp(2f\sigma)$ is the ratio of the maximum to the minimum limit of the full-width-at-half-maximum of the component, where $f = (2\ln 2)^{1/2}$.

Parameter	Avg	SD	Min	Max
B_0	−13.86	56.05	−725.11	6.347
B_1	25.95	87.58	−19.19	966.9
B_2	−13.67	36.53	−398.8	−0.301
FD_{max}	4.8×10^{17}	1.34×10^{19}	3.1×10^{-4}	3.73×10^{20}
D_{peak}	7.06	12.46	1.54×10^{-5}	103.6
σ	0.6385	0.3799	0.0537	1.9561
D_2/D_1	7.349	10.45	1.135	100.1

components is virtually the same as that shown in Table 6 for 853 log-normal components. The average values of coefficients B_0, B_1, and B_2 represent a geometrically averaged component, $FD_{avg}(D)$, such that $FD_{avg}(D) = [FD_1(D) \times FD_2(D) \times \ldots \times FD_N(D)]^{1/N}$. The average values of FD_{max}, D_{peak}, and σ do not have a simple meaning and are given here for completeness only. The two sets of averages do not yield the same function of the diameter D because they are related via nonlinear functions (Eq. 3a–c). The average error of approximation, relerr (Eq. 6), of the size distribution with the sum of log-normal components was 0.057 ± 0.030. The number of components per size distribution ranged from 1 to 6, with an average of 2.18 ± 1.22. (1 SD is shown following the ± sign).

Significant correlations exist between the D_{peak} and FD_{max}, as well as between D_{peak} and σ of the component, as can be seen in Figs. 10 and 11. The equations of the approximating lines (*see also the correlation coefficients in Table 6*) shown in these two figures are given in the figure legends.

The relationships between the B coefficients for the 853 log-normal components are

$$B_1 = (4.684 \pm 16.205) - (1.533 \pm 0.010)B_0, \quad (8)$$

$$B_2 = (-5.392 \pm 14.053) + (0.598 \pm 0.009)B_0, \quad (9)$$

$$B_2 = (-3.193 \pm 8.451) - (0.406 \pm 0.003)B_1. \quad (10)$$

1 SD of the respective parameter follows each ± sign. Given the errors of the intercepts, the latter can all be set to 0. Relationships among the B coefficients for all 902 components are virtually identical with those shown in Eq. 8–10.

Correlation between width, $w = D_2 - D_1$, and D_{peak} ($r^2 = 0.991$, $\log w$ vs. $\log D_{peak}$) is even greater than between σ and $\ln D_{peak}$ ($r^2 = 0.760$). The relationship between width w (μm), and peak diameter D_{peak} (μm) is

$$w = 1.368 D_{peak}^{0.9}. \quad (11)$$

The relationships expressed in the equations in Figs. 10 and 11 and Eq. 11 are consequences of the constraint that the envelope of all components of marine particle size distribution be described by a power law:

$$FD(D) = kD^{-m}.$$

k and m are parameters, with m assuming a value of ~4 (Jonasz 1983). Thus, the smaller D_{peak}, the larger must be FD_{max} of the component, and the narrower (in the linear scale) must be the component.

We have attempted to find correlations between the total surface and total volume of particles represented by a component and the peak diameter of that component. However, this work has so far been inconclusive. We believe that this is due to the extremely long "tail" of the 0th order log-normal distribution function. With this function, the most significant contributions to the higher moments (such as the total area and volume) come from a diameter range that is many component widths away from the peak. Higher order log-normal distributions would alleviate this problem, but would require a nonlinear fitting algorithm.

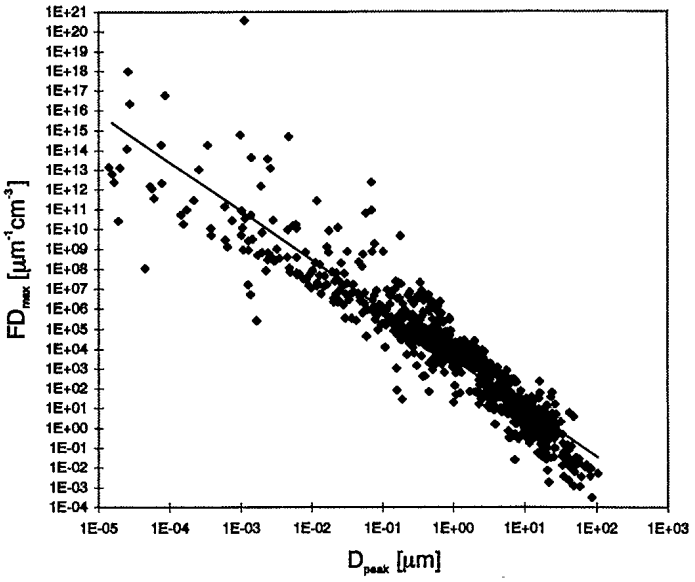


Fig. 10. Relationship between D_{peak} and FD_{max} for 853 log-normal components of 412 particle-size distributions. The range of FD_{max} values shown is limited to keep the number of decades on the FD_{max}-axis manageable. The data points not shown conform to the general trend. Approximation line equation: $\ln FD_{max} = (8.070 \pm 2.799) - (2.446 \pm 0.032)\ln D_{peak}$ (±1 SD). Correlation coefficient $r^2 = 0.873$.

Our results lead to two important conclusions. First, by using an automated algorithm we were able to identify two standard components of the marine particle size distribution, with peak diameters of ~0.66 and 10.5 μm in

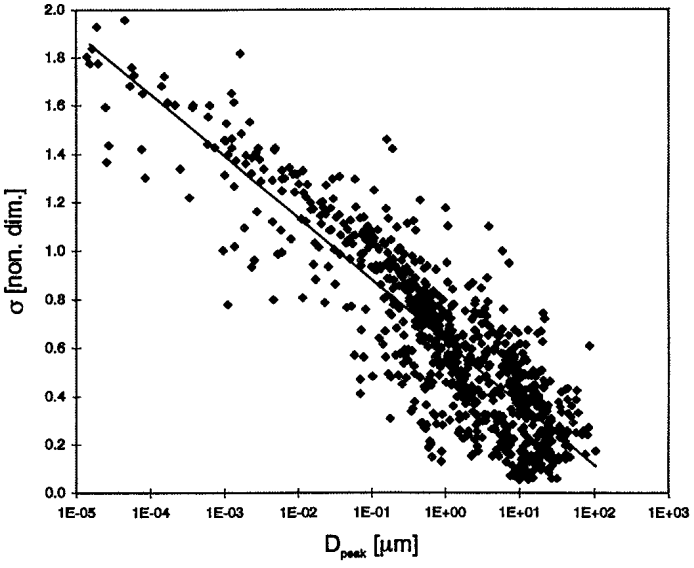


Fig. 11. Relationship between D_{peak} and σ for 853 log-normal components of 412 particle-size distributions. Approximation line equation: $\sigma = (0.626 \pm 0.186) - (0.111 \pm 0.002)\ln D_{peak}$ (±1 SD). Correlation coefficient $r^2 = 0.760$.

a population of 412 particle-size distributions obtained by different researchers in different seawater bodies and in different seasons. The existence of such components suggests that these components may be size distributions of specific biological or geological material and indicates the need to search for other possible standard components of the marine size distribution.

Second, we found that very few of the components are hyperbolic. This conclusion does rule out the first-order representation of the particle size distribution of marine particles by a hyperbola in a large range of particle diameters. Such a hyperbola would simply be an envelope of the sum of log-normal components.

References

BADER, H. 1970. The hyperbolic distribution of particle sizes. J. Geophys. Res. **75:** 2822–2830.

CARDER, K. L., G. F. BEARDSLEY, JR., AND H. PAK. 1971. Particle size distributions in the eastern equatorial Pacific. J. Geophys. Res. **76:** 5070–5077.

CASPERSON, L. W. 1977. Light extinction in polydisperse particulate systems. Appl. Opt. **16:** 3183–3189.

CHISHOLM, S. W. 1992. Phytoplankton size, p. 213–238 *In* P. G. Falkowski and A. D. Woodhead [eds.], Primary productivity and biogeochemical cycles in the sea. Plenum.

HOOD, R. R., M. R. ABBOT, AND A. HUYER. 1991. Phytoplankton and photosynthetic response in the coastal transition zone off northern California in June 1987. J. Geophys. Res. **96:** 14,769–14,780.

HUDSON, D. 1964. Statistics. CERN.

JONASZ, M. 1980. The characterization of physical properties of marine particles using light scattering [in Polish]. Ph.D. thesis, Institute of Oceanology, Sopot. 213 p.

———. 1983. Particle size distribution in the Baltic. Tellus **35B:** 346–358.

———. 1987. Nonspherical sediment particles: Comparison of size and volume distributions obtained with an optical and a resistive particle counter. Mar. Geol. **78:** 137–142.

JUNGE, C. E. 1963. Air chemistry and radioactivity. Academic.

KERKER, M. 1969. The scattering of light and other electromagnetic radiation. Academic.

KIEFER, D. A., AND J. BERWALD. 1992. A random encounter model for the microbial planktonic community. Limnol. Oceanogr. **37:** 457–467.

KRANCK, K., AND T. MILLIGAN. 1991. Grain size in oceanography, p. 332–345. *In* J. P. M. Syvitsky [ed.], Principles, methods, and application of particle size analysis. Cambridge.

LAMBERT, C. E., C. JEHANNO, N. SILVERBERG, J. C. BRUN-COTTAN, AND R. CHESSELET. 1981. Log-normal distributions of suspended particles in the open ocean. J. Mar. Res. **39:** 77–98.

LONGHURST, A. R., AND OTHERS. 1992. Sub-micron particles in northwest Atlantic shelf water. Deep-Sea Res. **39:** 1–7.

MCCAVE, I. N. 1983. Particulate size spectra, behavior, and origin of nepheloid layers over the Nova Scotia continental rise. J. Geophys. Res. **12:** 7647–7666.

MIDDLETON, G. V. 1970. The generation of log-normal size frequency distribution in sediments, p. 34–42. *In* M. A. Romanova and V. Sarmand [eds.], Topics in mathematical geology. Consultants Bur.

MILLIGAN, T. G., AND K. KRANCK. 1991. Electroresistance particle size analyzers, p. 109–118. *In* J. P. M. Syvitsky [ed.], Principles, methods, and application of particle size analysis. Cambridge.

MOREL, A., AND Y.-H. AHN. 1990. Optical efficiency factors of free-living marine bacteria: Influence of bacterioplankton upon the optical properties and particulate organic carbon in oceanic waters. J. Mar. Res. **48:** 145–175.

PLATT, T., AND K. DENMAN. 1977. Organisation of pelagic ecosystem. Helgol. Wiss. Meeresunters. **30:** 575–581.

RISOVIC, D. 1993. Two-component model of sea particle size distribution. Deep-Sea Res. **40:** 1459–1473.

SHELDON, R. W., A. PRAKASH, AND W. H. SUTCLIFFE, JR. 1972. The size distribution of particles in the ocean. Limnol. Oceanogr. **17:** 327–340.

SPENCER, D. W. 1963. The interpretation of grain size distribution curves of sediments. J. Sediment. Petrol. **33:** 180–190.

STRAMSKI, D., AND R. A. REYNOLDS. 1993. Diel variations in the optical properties of a marine diatom. Limnol. Oceanogr. **38:** 1347–1364.

TENCHOV, B. G., AND T. K. YANEV. 1986. Weibull distribution of particle sizes obtained by random fragmentation. J. Colloid Interface Sci. **111:** 1–7.

ULLOA, O., S. SATHYENDRANATH, T. PLATT, AND R. A. QUINONES. 1992. Light scattering by marine heterotrophic bacteria. J. Geophys. Res. **6:** 9619–9629.

VAN ANDEL, T. H. 1973. Texture and dispersal of sediments in the Panama Basin. J. Geol. **81:** 434–457.

Submitted: 30 November 1994
Accepted: 7 August 1995
Amended: 25 October 1995