Depth-Averaged Concentration

NAME

avgc – average concentration for a layer

SYNOPSIS

`avgc [filename]`

The program has two operating modes, depending on whether a filename is provided on the command line. If the program is invoked with no arguments (*i.e.*, no filename is provided), the program will provide an interactive data entry screen (Figure 2).

If the name of a suitable data file (extension ".`dat`") is provided, the program will run in "batch" mode and will create two new files (overwriting any existing files of the same name). The names of the new files are obtained by first removing any extension to obtain the "basename". The program will read data from "`basename.dat`", and will write the results on "`basename.out`". A processing log will be written on "`basename.log`".

PURPOSE

This program computes the depth-averaged (chlorophyll) biomass concentration, $\langle C \rangle$, in a layer. The chlorophyll profile is modelled by a shifted-Gaussian curve, as described by Platt *et al.* (1994). Given the specific attenuation coefficient for biomass, k_c, and the attenuation coefficient for water, k_w, the results of this program can be used to compute the average attenuation coefficient for a layer, $\langle K \rangle = k_w + k_c \langle C \rangle$.

It is sometimes necessary to distinguish between depth profiles for photosynthetically active biomass, $B(z)$, and for the biomass (including phaeopigments) which contributes to light attenuation, $C(z)$ (for the formal distinction between B and C, see Sathyendranath *et al.* (1989)). In the open ocean, the contribution of phaeopigments to $C(z)$ is negligible (Longhurst *et al.* 1995). Accordingly, we may use $B(z)$ to represent both quantities, and compute the diffuse, vertical, attenuation coefficient from $B(z)$ as described by Sathyendranath and Platt (1988). In this appendix, however, consistent with the notation of Platt *et al.* (1994), biomass will be represented using $C(z)$.

DESCRIPTION

The program reads a file containing input data: the parameters of the biomass profile and the upper and lower depths which define the layer for which the average is to be computed. The program creates two output files: one contains values for $\langle C \rangle$, the other is a log which provides a record of the calculation, including entries for any processing errors that may have occurred.
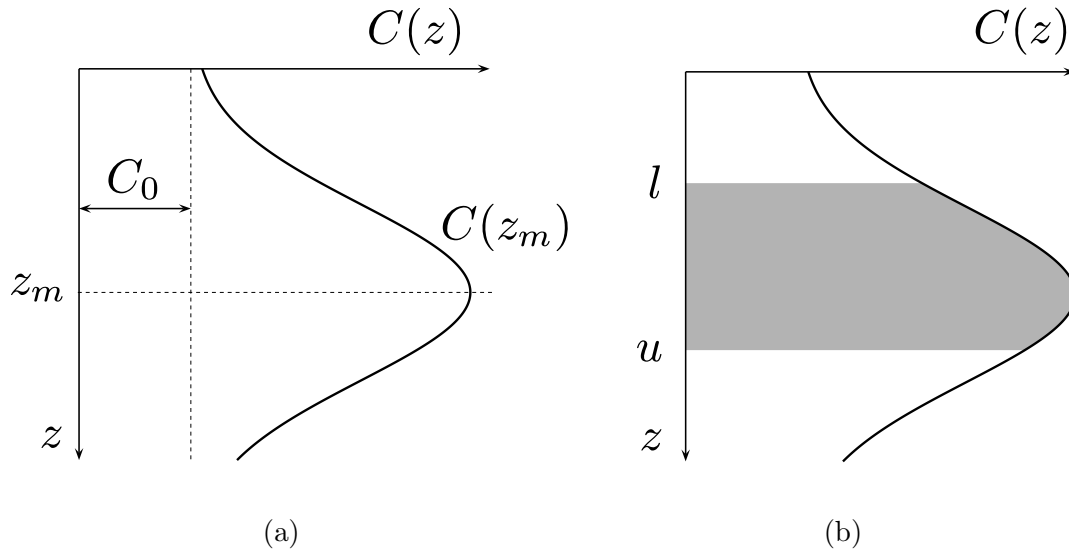
Figure 1. Shifted-Gaussian model for biomass–depth profile: a) illustrating definitions of the parameters z_m and C_0, and b) indicating the layer for which the average concentration will be computed.

THEORY

Let the vertical distribution of chlorophyll concentration be $C(z)$, with z positive down-wards. Consider an arbitrary layer extending between depths l and u ($u > l$) and let $\langle K_c \rangle$ be the average attenuation coefficient for the layer, due to chlorophyll. Then we have

$$I(u) = I(l)e^{-\langle K_c \rangle(u-l)}, \tag{1}$$

where

$$\langle K_c \rangle = k_c \langle C \rangle. \tag{2}$$

It has been found that the shifted-Gaussian model provides a suitable description of the vertical distribution of chlorophyll at a wide range of locations in the ocean (Platt *et al.* 1988). Explicitly,

$$C(z) = C_0 + \frac{h}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(z-z_m)^2}{2\sigma^2}\right] \tag{3}$$

where a background biomass C_0 has superimposed on it a Gaussian curve (Figure 1(a)) with parameters z_m, the depth of the chlorophyll maximum; σ, a measure of the thickness or vertical scale of the peak; and h, the total biomass in the peak such that $h/(\sigma\sqrt{2\pi})$ is a measure of the height of the peak above background.

For a multi-component medium, Platt *et al.* (1994) have shown that the average (total) attenuation coefficient for a layer is the sum of the average attenuation coefficients for the individual components. As a result, the background biomass can be treated separately from

```
                   Select variable, press [Enter] to edit:

** Exit (accept current values)
Record identifier: example
Depth of chlorophyll maximum: 27.000 (m)
Background component of chlorophyll-depth profile: 0.310 (mg Chl/m^3)
Gaussian height parameter: 43.400 (mg Chl/m^2)
Gaussian scale parameter, sigma: 9.000 (m)
Depth to top of layer: 0.000 (m)
Depth to bottom of layer: 46.550 (m)


                                results
Status code: 0 (0=OK, -9=not computed)
Average concentration for the layer: 1.230 (mg Chl/m^3)
```

Figure 2. Interactive data entry screen for average concentration calculations.

---

the superimposed Gaussian component and from the attenuation due to seawater itself or any other component, dissolved or suspended.

The average concentration, $\langle C \rangle$, is a function of the top, l, and bottom, u, depths (Figure 1(b)), where l is a numerically smaller value than u (a consequence of the convention by which depth is measured) and of the parameters of the biomass profile model.

FILES

In addition to the binary executable program, one input data file is required. Two new files will be created, an output data file and a log file containing a record of the processing (previously existing files having the same names will be destroyed without warning). Each file is identified by its extension (the three letters following the "." character):

1) the program (executable) itself (`.exe` extension);
2) input data (ASCII text, `.dat` extension);
3) processing log (ASCII text, `.log` extension); and
4) output data (ASCII text, `.out` extension).

All the names are determined from the command line at run time (*i.e.*, the program does not rely on any "hard-coded" file names). When no file name is given on the command line, the program creates the file "`avgc.dat`" using values entered interactively by the user. In this case, the output files will be "`avgc.log`" and "`avgc.out`".

INPUT

The program can be used interactively, in which case the user will be allowed to enter values for the input variables, or can read data from a file. When the program is run interactively, a data file called `avgc.dat` will be created together with the output file `avgc.out` and a log file `avgc.log`. These are simple ASCII text files that can be viewed with any editor. Each time the program is run these files will be overwritten, so the files must be copied to another location or renamed if the results are to be saved for future use.

The data file must provide the parameters for the shifted-Gaussian curve, z_m, C_0, h, and σ, and the depths, l and u, that define the layer. All depths are measured with the origin at the surface and positive in the downwards direction. Note that in rare instances the fitting procedure may produce a negative value for the "depth" of the chlorophyll maximum. This can occur when the maximum concentration is near the surface[1].

The input data file has two header records followed by one or more data records. The first field of each input data record is an identifying string which is simply passed on unchanged to the output record (this can be useful when relating the output to other files). The fields are shown below (the indicated ranges are those used for interactive data entry, consistent with an `f8.3` format):

1) an arbitrary string (format `an`) used to identify the record for subsequent processing;
2) "depth" of the chlorophyll maximum, $-999.999 \le z_m \le 9999.999$ (m), format `fw.p`;
3) the background component, $0 \le C_0 \le 9999.999$ (mg Chl m^{-3}), format `fw.p`;
4) height parameter for the Gaussian component, $0 \le h \le 9999.999$ (mg Chl m^{-2}), format `fw.p`;
5) Gaussian scale parameter, $0.001 \le \sigma \le 9999.999$ (m), format `fw.p`;
6) depth, $0 \le l \le 9999.999$ (m), of the top of the layer, format `fw.p`; and
7) depth, $0 \le u \le 9999.999$ (m), of the bottom of the layer, format `fw.p`.

Each input data record must have 6 numeric fields. When $h = 0$, the values of σ and z_m are ignored, but should be coded in the input file using valid numbers to avoid a run-time error while processing the input record.

The input file is ASCII text with one record per line. The first line of the file is not used, and can be used to provide a list of the input variables or to otherwise identify the data set. The second input record must provide a FORTRAN format string that can be used to read the data records. This format must use only `an`, `fw.p`, `in`, or `x` specifiers. The input data file corresponding to Figure 2 is shown below:

```
  id,     z_m,       C_0,      h,     sigma,     l,        u
(a8,x,f8.3,x,f8.3,x,f8.3,x,f8.3,x,f8.3,x,f8.3)
example    27.000    0.310    43.400    9.000    0.000    46.550
```

---

[1] The fitting procedure yielded a negative estimate for z_m in only two of the 28 cases shown in Table 3 of Platt *et al.* (1991), representing just three profiles in a data set of more than 600 profiles.

OUTPUT

The program produces a single output file. This file contains a two-line header followed by one or more output data records, one for each input data record. Each output data record contains the record identifier, a status code, the depths u and l, and the computed value for $\langle C \rangle$. The first two records are a list of the output variables and a FORTRAN format suitable for reading the file. The remaining records contain data. The first (record identifier) and depth fields are the same as the input data file. The second field is a status code (0="no error", see DIAGNOSTICS for the interpretation of non-zero status codes). The last field is the total value of $\langle C \rangle$ including contributions from both the background and the Gaussian components. The output file corresponding to Figure 2 is:

```
id,  status,  l,     u,     < C >
(a8,x,i2,x,f6.2,x,f6.2,x,f6.2)
example   0   0.00  46.55   1.23
```

The corresponding log file is:

```
Depth-averaged concentration

Program: .\BIN\AVGC.EXE:

Files:
 Output file: avgc.out
 Log file:    avgc.log
 Input file:  avgc.dat

Limits:
   tolerance, epsilon, for (u-l):   0.1000E-03

Processing errors:

Summary:
     3 records processed (including two header records)
No processing errors.
```

Note that if an error had occured, a log entry would have been created. The corresponding output record would have a non-zero status code and a missing value code would appear in place of a value for $\langle C \rangle$. The DIAGNOSTICS section explains the error messages and resulting status codes.

Processing

The calculations are performed for a single input data record at a time. If a processing error occurs an entry will be made in the log file identifying the line at which the problem occurred and a non-zero status value is written to the output record.

The additive property allows the computations to be done separately for the background and Gaussian components and then summed to obtain the total. If $h = 0$, the Gaussian portion of the calculation is omitted and the parameters z_m and σ are ignored.

The computation of the Gaussian component follows eqn. (13) of Platt *et al.* (1994), from which it follows that:

$$\langle C \rangle = C_0 + \frac{h}{2(u-l)}(\Phi(x_0) - \Phi(x_1)),$$

where $\Phi(x)$ is the *error function* $(\Phi(x) = (2/\sqrt{\pi}) \int_0^x e^{-t^2}\, dt)$, $x_0 = (l - z_m)/\sigma\sqrt{2}$, and $x_1 = (u - z_m)/\sigma\sqrt{2}$.

Several complications due to machine arithmetic require careful handling in the numerical calculations to minimize the occurrence of run-time errors. Note that this section of code is used only in the case $h > 0$ and $l < u$.

The divisions cannot be performed when $\sigma = 0$. When $\sigma = 0$ but $h > 0$, no meaningful value can be computed. It is, however, difficult to protect against overflow in case where σ has a small non-zero value and the numerator is large.

When $u = l$ the calculation can be simplified. Since values for u and sometimes l are obtained from the input file, the smallest meaningful difference is approximately the smallest difference that can be represented in the format used to read the data. The smallest real difference between u and l that can occur depends on the scales for which the calculations make sense, and will typically be much larger. Internally, a parameter ϵ (see Limits) is used in a test of the form $|u - l| < \epsilon$. The value for ϵ can be any value large enough so the division is unlikely to fail but smaller than the physically significant difference in depths.

Special cases

In the case $u = l$, we have $x_0 = x_1 = x$, and the contribution of the Gaussian component to $\langle C \rangle$ is just the value of the Gaussian function at the indicated depth:

$$\langle C \rangle = C_0 + \frac{h}{\sigma\sqrt{2\pi}} exp\left(-x^2\right).$$

REQUIREMENTS

The numerical calculations are not demanding. An effort has been made to ensure that the results will remain consistent across a range of hardware platforms. It is assumed that double precision variables conform to the IEEE floating point arithmetic standard. This is the most efficient data type for floating point computations on modern microprocessors with hardware floating point support.

To run the program under MS-DOS, approximately 400k of free low DOS memory and 1024k of extended memory is required. The `ansi.sys` device driver must be loaded in `config.sys`.

BUGS

It is difficult to protect against all possible combinations of input parameters that generate run-time errors due to floating point underflow or overflow (see DIAGNOSTICS).

LIMITS

The parameter ϵ used in the test $|u - l| < \epsilon$ is set at compile time. The current value is shown in the the log file.

DIAGNOSTICS

The following messages may occur:

```
** error in get_files **
```

An error occurred in the `get_files` subroutine. This message will be preceded by a message indicating the type of error that occurred.

```
** error ** file I/O
```

An error occurred while reading or writing a file. This could indicate a missing or corrupted file, a disk problem such as lack of space, or a program which uses more files than the operating system configuration allows (many systems limit the number of files a program can use; in some cases the user may be able to increase this number via a configuration option).

```
** error ** getarg
```

The system function used to obtain the command line parameters returned an error. This may indicate lack of memory, an incompatible command processor, or a command line that is too long.

```
** error ** name too long:  ...
```

A program or file name was too long.

```
** error ** limit exceeded:  too many files to read
```

This is an internal program error which should not occur. The list of input file extensions passed to the subroutine has more entries than the number of files requested.

```
** error ** limit exceeded:  too many files to write
```

This is an internal program error which should not occur. The list of output file extensions passed to the subroutine has more entries than the number of files requested.

```
** error ** opening file ...
```

The indicated file could not be opened. The file name may have been entered incorrectly or the file may have a hidden, read-only, or system attribute.

```
** I/O error **
```

An error occurred while reading from or writing to a file or the console. This could indicate a missing or corrupted file, a disk problem such as lack of space, a buffer overflow, or a control character inadvertently entered from the keyboard.

```
-- End of file --
```

This is not always an error, but may indicate a file that has been truncated or damaged.

```
** error:  input file(s) **
```

A problem occurred with an input file. The file name passed to the program may be incorrect, or the file may have a hidden or system attribute.

```
** error:  output file(s) **
```

A problem occurred with an output file. The file name passed to the program may be incorrect or the file may have a hidden, system, or read-only attribute.

```
** error:  input record format **
```

The input record format (obtained from the second line of the input file) did not have the required number and types of fields.

```
** error:  division by zero **
```

The value for σ was zero. This could result from a data entry error: either a zero value was entered or the data format is incorrect. A missing value code will be printed and the value of the status field set to -1.

The status codes that may appear in the output file are:

  0 – no processing error occurred.

 -1 – the input data include $\sigma = 0$ for $h > 0$; the Gaussian component is undefined.

REFERENCES

Longhurst, A., S. Sathyendranath, T. Platt, and C. Caverhill (1995), 'An estimate of global primary production in the ocean from satellite radiometer data', *J. Plankton Res.* **17**, 1245–1271.

Platt, T., C. Caverhill, and S. Sathyendranath (1991), Basin-scale estimates of oceanic primary production by remote sensing: The North Atlantic. *J. Geophys. Res.*, **96**, 15147–15159.

Platt, T., S. Sathyendranath, C. M. Caverhill, and M. R. Lewis (1988), Ocean primary production and available light: Further algorithms for remote sensing. *Deep-Sea Research*, **35**, 855–879.

Platt, T., S. Sathyendranath, G.N. White III, and P. Ravidran (1994), Attenuation of visible light by phytoplankton in a vertically structured ocean. *J. Plankton Res.*, **16**, 1461–1487.

Sathyendranath, S., and T. Platt (1988), 'The spectral irradiance field at the surface and in the interior of the ocean: A model for applications in oceanography and remote sensing', *J. Geophys. Res.* **93**, 9270–9280.

Sathyendranath, S., T. Platt, C. M. Caverhill, R. E. Warnock, and M. R. Lewis (1989), 'Remote sensing of oceanic primary production: Computations using a spectral model', *Deep-Sea Res. I* **36**, 431–453.

NOTATION

$C(z)$ chlorophyll concentration as a function of depth, mg Chl m^{-3}.

C_0 background concentration in shifted-Gaussian model, mg Chl m^{-3}.

$\langle C \rangle$ average concentration for a layer, mg Chl m^{-3}.

K_C attenuation coefficient due to chlorophyll, m^{-1}.

$\langle K_C \rangle$ average attenuation coefficient due to chlorophyll for a layer, m^{-1}.

h parameter for the amplitude of the chlorophyll maximum in shifted-Gaussian model, mg Chl m^{-2}.

$I(z)$ total irradiance as a function of depth, W m^{-2}.

k_c specific attenuation coefficient due to chlorophyll biomass, m^{-1}(mg Chl m^{-3})$^{-1}$.

k_w attenuation coefficient due to water, m^{-1}.

l depth (origin at the surface, positive downwards, $l < u$) to top of layer, m.

σ scale parameter for width of peak in shifted-Gaussian model, m.

u depth (origin at surface, positive downwards, $l < u$) to bottom of layer, m.

z depth (origin at surface, positive downwards), m.

z_m "depth" (in rare cases, the value of z_m may be negative) of chlorophyll maximum in shifted-Gaussian model, m.