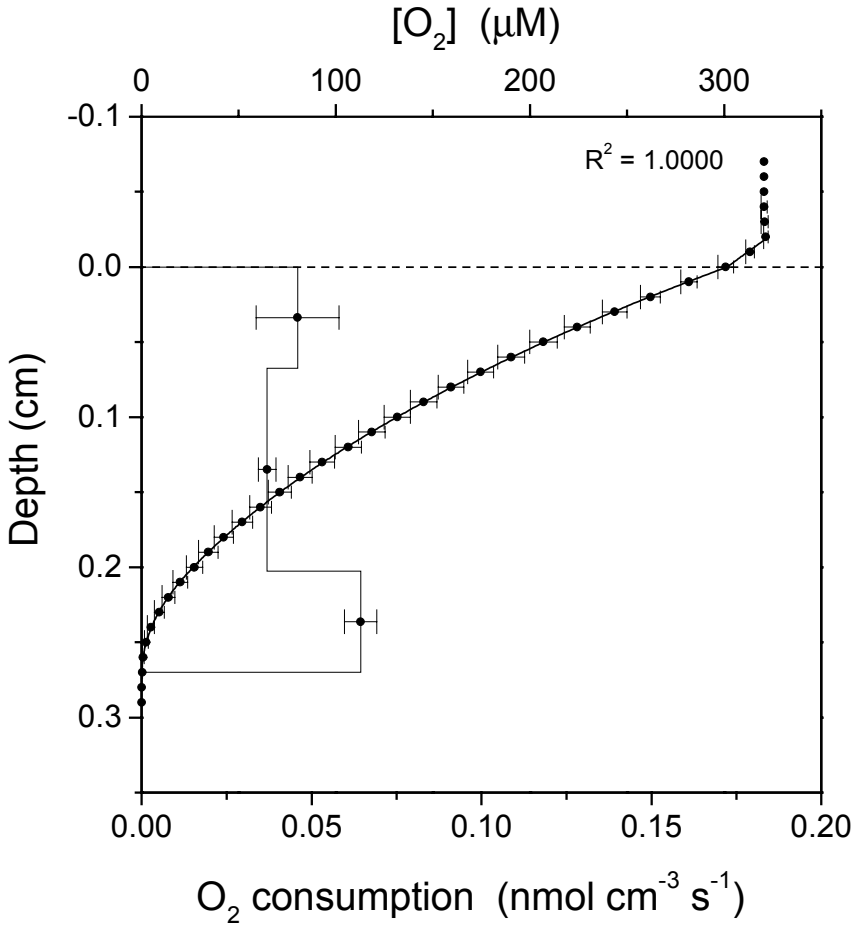
Manual for PROFILE

(version 1.0)



Content:

1. Introduction

This short manual is written as a supplement to our L&O paper (Berg, P., N. Risgaard-Petersen, and S. Rysgaard. 1998. Interpretation of measured concentration profiles in the sediment porewater. 1998, 43:1500-1510) where the profile interpretation procedure (PROFILE) was originally described. Users are recommended to read the paper as well as this manual before using PROFILE.

The main focus of this manual is on: 1) how to generate an input file using the possible options in PROFILE; 2) how to use the F statistics; and 3) how to read and use the output from PROFILE.

PROFILE is runs under Windows 95, Windows 98 and Windows NT.

PROFILE version 1.0 is given to users for free. For that reason, the user has the full responsibility for how the procedure or results are used.

2. The input file

PROFILE needs one input file which has the structure as shown in Fig. 1. Note that the line numbers in the left margin are not a part of the input file. (The listed example of input data is for the profile and its interpretation shown in Fig. 3 in our paper and also on the front page of this manual.)

The input file is a text file and can be created/changed in Notepad in Windows 95, 98 or NT, or any other editors.

Variables in the input file can be given in any units as long as they are consistent (see detailed description in section 5). The input file contains some text, which is not used by PROFILE, but serves only as an aid for the user. Each line in the input file is described below:

Line1: Not used by PROFILE, but serves as a description of the data in the input file. As a minimum a blank line must be given.

Line 2: The depth coordinate for the top of calculation domain. The number must be identical to one of the depths given in line 16 or below. In the example, a value of -0.02 is specified, which means that the concentrations listed in line 16 to 20 are not included in the interpretation. Never specify any depth above the diffusive boundary layer.

Line 3: The depth coordinate for the bottom of calculation domain. The number must be identical to one of the depths given below line 16 and larger than the depth coordinate given in line 2. In the example, a value of 0.27 is specified, which means that the concentrations listed in line 51 and below are not included in the interpretation.

Line 4: The maximum number of equally spaced zones that will be considered in the interpretation. If, for example, 7 is specified the zones are gradually increased from 1 to 7 in the first step of the interpretation. The maximum allowed value is 12. The specified number must always be smaller than the number of measured concentrations included in the interpretation (specified through line 2 and 3).

(PS: PROFILE runs in a DOS-window which has a certain width that can not be changed in Windows 95 or 98. This makes the screen print look odd when more than 10 zones are specified. Using Windows NT the width can be changed.)

Line 5: A number between 1 and 5 specifying the chosen combination of the two boundary conditions (BC) needed in the interpretation. Note that in some options both BC is specified at the top (option 2) or the bottom (option 3) of the calculation domain. The possible combinations are as follows:

1) first BC is a concentration at the top, second BC is a concentration at the bottom
2) first BC is a concentration at the top, second BC is a flux at the top
3) first BC is a concentration at the bottom, second BC is a flux at the bottom
4) first BC is a concentration at the top, second BC is a flux at the bottom

5) first BC is a flux at the top, second BC is a concentration at the bottom.

(PS: if irrigation is included ($\alpha \neq 0$) only combination 1 is valid.)

Line 6: First BC according to the number specified in line 5.

Line 7: Second BC according to the number specified in line 5.

Line 8: The diffusivity for the solute in free water (D).

Line 9: A number between 1 and 3 specifying the preferred expression for the sediment diffusivity (D_s) as a function of porosity (φ) and diffusivity in free water (D). The possible options are as follows:

1) $D_s = \varphi D$
2) $D_s = \varphi^2 D$
3) $D_s = D / (1 + 3(1 - \varphi))$

Line 10: The concentration of the solute in the overlying water column (C_0). The specified value is only used if irrigation is included ($\alpha \neq 0$).

Line 11: A minimum limit for the production rates that can be calculated in the interpretation. In most cases just specify a large negative value here as in the example where a value of -1.0 10^{20} is specified, which practically means that all realistic production rates are allowed in this interpretation. The option of specifying a minimum value, can be used in situations where, for example, only positive production rates (no consumption) are to be considered, in which case the value 0 is specified.

Line 12: A maximum limit for the production rates that can be calculated in the interpretation. In most cases just specify a large positive value here. In the example a value of 0 is specified, which means that only negative production rates (consumption) are considered in this interpretation.

Line 13: A stop criteria (in %) for the iterative process of finding the best fits. If a value is specified, that is too large an incorrect solution will be found. If a value is specified, that is too small the correct solution will be found, but the calculation time will be longer than necessary. The appropriate stop criteria to use can vary a lot from problem to problem, but values such as 0.001 or 0.0001 are generally good starting points. It is a good strategy when a desired solution is found to redo the interpretation with a stop criteria that is, for example, 10 times smaller and verify that exactly the same results are found.

Line14: The P value, or level of significance, used in the F test to calculate recommended choices of zones.

Line15: Not used by PROFILE but serves as a header for the values listed below. As a minimum a blank line must be given.

Line16 to "end of file": A list with five variables per line specifying depth profiles needed in the interpretation. The five variables are as follows:

1) Depth coordinate (x) where the actual concentration was measured. If the sediment water interface is included, it must have the depth coordinate $x = 0$
2) Porosity (φ) which must be 1 in the water column
3) Biodiffusivity (D_B) which must be 0 in the water column
4) Irrigation coefficient (α) which must be 0 in the water column
5) Concentration (C)

```
1   Interpretation of O2-profile in paper (Fig. 3)
2   -0.02       Depth at top of calculation domain
3    0.27       Depth at bottom of calculation domain
4    7          Max number of equally spaced zones in interpretation (1 to 12)
5    3          Type of boundary conditions (1:t=C b=C, 2:t=C t=F, 3:b=C b=F 4:t=C b=F 5:t=F b=C)
6    0.22573    First boundary condition
7    0.0        Second boundary condition
8    11.7E-06   Diffusivity in water (D)
9    2          Expression for sediment diffusivity (Ds) (1: Ds=FI*D, 2: Ds=FI**2*D, 3: Ds=D/(1+3*(1-FI))
10   321.4123   Concentration in water column (C0)
11  -1.0E+20    Minimum for production rate
12   0.0        Maximum for production rate
13   0.001      Maximum deviation (in %) when accepting a calculated minimum
14   0.01       Level of significance in the F statistics
15        X   FI   DB  ALFA          C
16   -0.07    1    0     0   320.5841
17   -0.06    1    0     0   320.5841
18   -0.05    1    0     0   320.5841
19   -0.04    1    0     0   320.5841
20   -0.03    1    0     0   320.8101
21   -0.02    1    0     0   321.4123
22   -0.01    1    0     0   313.216
23    0.00  0.8    0     0   300.8086
24    0.01  0.8    0     0   281.7761
25    0.02  0.8    0     0   261.9157
26    0.03  0.8    0     0   243.6379
27    0.04  0.8    0     0   224.3779
28    0.05  0.8    0     0   206.9242
29    0.06  0.8    0     0   190.2216
30    0.07  0.8    0     0   174.5708
31    0.08  0.8    0     0   159.2973
32    0.09  0.8    0     0   145.2283
33    0.1   0.8    0     0   131.6854
34    0.11  0.8    0     0   118.5933
35    0.12  0.8    0     0   106.4043
36    0.13  0.8    0     0    92.86111
37    0.14  0.8    0     0    81.57462
38    0.15  0.8    0     0    71.19043
39    0.16  0.8    0     0    61.18299
40    0.17  0.8    0     0    51.77705
41    0.18  0.8    0     0    42.29603
42    0.19  0.8    0     0    34.32065
43    0.2   0.8    0     0    27.17016
44    0.21  0.8    0     0    19.87087
45    0.22  0.8    0     0    13.77456
46    0.23  0.8    0     0     9.03325
47    0.24  0.8    0     0     4.66795
48    0.25  0.8    0     0     2.40918
49    0.26  0.8    0     0     0.75301
50    0.27  0.8    0     0     0.22573
51    0.28  0.8    0     0     0
52    0.29  0.8    0     0     0
```

Fig 1. Example of an input file. Note that the line numbers in the left margin are not part of the input file. See remarks on units in section 5. The listed input data is for the profile and its interpretation shown in Fig. 3 in our paper and also shown on the front page of this manual.

3. Running PROFILE - using the F statistics

PROFILE is run by double clicking on the PROFILE.exe file (or through the Start/Run dialog box in Windows) which prompts a DOS window to pop up on the screen. Here the user is asked to specify first the name of the input file (see section 2), and second the name of one of the two output files (see section 4). Another output file, or log file, is automatically given the name PROFILE.log (see section 4) and it contains a copy of all important information printed on the screen when PROFILE is running. Fig. 2 shows an example of a log file which is generated from the input listed in Fig. 1.

After entering the names of the input and the output files, the first step of the procedure starts where the number of equally spaced zones is gradually increased from 1 up to the number specified in the input file (Fig. 1 line 4). After this first step, a table is printed on the screen as shown in Fig. 2 (line 7 to 15). The table shows the sum of squared deviations (SSE) and R^2 for the best obtainable fits with different numbers of equally-spaced zones, as well as P values for the comparisons (F test) of the different fits. The table is similar to Table 1 in our paper.

The tabulated P values are used to determine how many zones to use in the second step of the procedure. The P values in Fig. 2 show that four zones is a reasonable choice at a significance level of 0.010 in this particular example. Details of how to read the table are given in our paper for another and slightly more complicated example. Since the interpretation of the P values can be a little complicated PROFILE automatically generates a suggested number of zones, which is written on the screen right below the table (see Fig. 2 line 17)

When the number of zones for the further calculations is entered by the user, the second step of the procedure starts where adjacent zones are lumped. The result of this lumping is listed in a table as shown in Fig. 2 (line 21 to 26). The table shows as before the sum of squared deviations (SSE) and R^2 for the best obtainable fits, as well as P values for the comparisons of the different fits. This table is similar to Table 2 in our paper.

Based on the tabulated P values the number of zones in the final result is chosen. Details of how to read the table are given in our paper. As before, PROFILE automatically generates a suggested number of zones which is written on the screen right below the table (see Fig. 2 line 28)

When the final number of zones is entered by the user, a list of calculated results is printed on the screen as shown in Fig. 2 (line 34 to 46). For further details of how to read these calculated results and their units see section 5.

The F statistics sometimes can fail to suggest a number of zones to start the lumping from in the second step of the procedure. Most often this is because too few zones are specified as an upper limit for the variation of equally spaced zones in the input file (Fig. 1 line 4), and it is easy to specify a higher number and redo the calculation. It is important that a high enough number for this upper limit is specified, so that any further

increase will not give fits that are significantly better. On the other hand, specifying the maximum allowed value of 12 in all interpretations will give a fairly long calculation time in some cases.

It is also important to keep of the right perspective on the F statistics as a helpful tool in finding the appropriate number of zones to start out with when zones are lumped. In the example shown in Fig. 2, one could argue that an R^2 value of 0.9999 that is obtained at three zones (line 11) is very high, and that an increase from three to four zones as suggested by the P values will be to take the measurements too far, or to over interpret the measured data. If this calculation is continued with three zones instead of four very similar results would be calculated. The biggest difference would be a moderate change in the extent of the top and bottom zones with elevated consumption rates (see Fig 3 in our paper or the front page on this manual), while the same depth integrated production, for example, will be calculated within 0.4%.

```
Name of input file  : Example.inp
Name of output file: Example.out

Input read from file, - the calculation begins.


Zones        SSE      R**2   Zones:    2      3      4      5      6      7

    1   .1359E+04    .9958 ----------  .000  .000  .000  .000  .000  .000
    2   .1991E+03    .9994 ----------------   .000  .000  .000  .000  .000
    3   .1774E+02    .9999 -------------------------- .000  .000  .000  .000
    4   .8216E+01   1.0000 ---------------------------------- .271  .051  .028
    5   .7820E+01   1.0000 ---------------------------------------- .030  .020
    6   .6406E+01   1.0000 ------------------------------------------------ .076
    7   .5569E+01   1.0000

The F statistics suggest  4 zones. Choose the number of zones for further
calculations:  4


Zones        SSE      R**2   Zones:    3      2      1

    4   .8216E+01   1.0000 ----------  .434
    3   .8416E+01   1.0000 ----------------  .000
    2   .1544E+02   1.0000 --------------------------  .000
    1   .1359E+04    .9958

The F statistics suggest  3 zones. Choose the number of zones in
final result:  3

The calculation is done, -    151638 steady state profiles are tested.


Calculated concentration at top      :  .3227E+03
Calculated concentration at bottom:  .2257E+00
Calculated flux at top               :  .1242E-01
Calculated flux at bottom            :  .2410E-15
Depth integration of production      : -.1242E-01
Depth integration of irrigation      :  .0000E+00

Depth integration of production and irrigation in each zone:

                          Zones     Production      Irrigation
                            1      -.3080E-02     .0000E+00
                            2      -.4996E-02     .0000E+00
                            3      -.4343E-02     .0000E+00
```

Fig 2. Example of an log file (automatically named PROFILE.log). Note that the line numbers in the left margin are not part of the file. The file was calculated from the input data shown in Fig. 1. See remarks on units in section 5.

4. The output files

Two output files are generated when running PROFILE. The units of the listed values in the output files depends entirely on the units used in the input file. For further details see section 5.

The first output file is a log file, which simply contains a copy of all important information printed on the screen when PROFILE is running. This file is automatically named PROFILE.log (see comments in section 3 and example in Fig 2).

The second output file is given a name by the user when PROFILE is running. As the example in Fig. 3 shows this output file contains a list of depth dependent variables (profiles) calculated in the interpretation. Note that only the top and the bottom of the file is shown. The listed values are calculated from the input shown in Fig. 1. The first line is a text line, describing the seven listed variables which are as follows:

1) Depth coordinate (x)
2) Porosity (φ). The values are found by linear interpolation between the values given in the input file
3) Biodiffusivity (D_B). The values are found by linear interpolation between the values given in the input file
4) Irrigation coefficient (α). The values are found by linear interpolation between the values given in the input file
5) Concentration (C) for the best fitting profile
6) Irrigation rate ($\varphi\alpha(C_0 - C)$) at the given depth
7) Production rate (R) at the given depth (or consumption rate if negative)

Variables in this output file can be read easily by most graphic programs for visual inspection and final presentation. Visual inspection of calculated results is recommended in all interpretations.

X	FI	DB	ALFA	C	IRRIGATION	PRODUCTION
-.2000E-01	.1000E+01	.0000E+00	.0000E+00	.3227E+03	.0000E+00	.0000E+00
-.2000E-01	.1000E+01	.0000E+00	.0000E+00	.3227E+03	.0000E+00	.0000E+00
-.1857E-01	.1000E+01	.0000E+00	.0000E+00	.3212E+03	.0000E+00	.0000E+00
-.1571E-01	.1000E+01	.0000E+00	.0000E+00	.3181E+03	.0000E+00	.0000E+00
-.1286E-01	.1000E+01	.0000E+00	.0000E+00	.3151E+03	.0000E+00	.0000E+00
-.1000E-01	.1000E+01	.0000E+00	.0000E+00	.3121E+03	.0000E+00	.0000E+00
-.7143E-02	.1000E+01	.0000E+00	.0000E+00	.3090E+03	.0000E+00	.0000E+00
-.4286E-02	.1000E+01	.0000E+00	.0000E+00	.3060E+03	.0000E+00	.0000E+00
-.1429E-02	.1000E+01	.0000E+00	.0000E+00	.3030E+03	.0000E+00	.0000E+00
-.2819E-17	.1000E+01	.0000E+00	.0000E+00	.3014E+03	.0000E+00	.0000E+00
-.2819E-17	.1000E+01	.0000E+00	.0000E+00	.3014E+03	.0000E+00	-.4563E-01
.1467E-02	.8000E+00	.0000E+00	.0000E+00	.2984E+03	.0000E+00	-.4563E-01
.4402E-02	.8000E+00	.0000E+00	.0000E+00	.2924E+03	.0000E+00	-.4563E-01
.7337E-02	.8000E+00	.0000E+00	.0000E+00	.2864E+03	.0000E+00	-.4563E-01
.1027E-01	.8000E+00	.0000E+00	.0000E+00	.2805E+03	.0000E+00	-.4563E-01
.1321E-01	.8000E+00	.0000E+00	.0000E+00	.2747E+03	.0000E+00	-.4563E-01
.1614E-01	.8000E+00	.0000E+00	.0000E+00	.2690E+03	.0000E+00	-.4563E-01
.1908E-01	.8000E+00	.0000E+00	.0000E+00	.2633E+03	.0000E+00	-.4563E-01
.2201E-01	.8000E+00	.0000E+00	.0000E+00	.2576E+03	.0000E+00	-.4563E-01
.2495E-01	.8000E+00	.0000E+00	.0000E+00	.2521E+03	.0000E+00	-.4563E-01
.2788E-01	.8000E+00	.0000E+00	.0000E+00	.2466E+03	.0000E+00	-.4563E-01
.3082E-01	.8000E+00	.0000E+00	.0000E+00	.2412E+03	.0000E+00	-.4563E-01
.3375E-01	.8000E+00	.0000E+00	.0000E+00	.2358E+03	.0000E+00	-.4563E-01
.3668E-01	.8000E+00	.0000E+00	.0000E+00	.2305E+03	.0000E+00	-.4563E-01
.2539E+00	.8000E+00	.0000E+00	.0000E+00	.1613E+01	.0000E+00	-.6434E-01
.2568E+00	.8000E+00	.0000E+00	.0000E+00	.1151E+01	.0000E+00	-.6434E-01
.2597E+00	.8000E+00	.0000E+00	.0000E+00	.7808E+00	.0000E+00	-.6434E-01
.2627E+00	.8000E+00	.0000E+00	.0000E+00	.5032E+00	.0000E+00	-.6434E-01
.2656E+00	.8000E+00	.0000E+00	.0000E+00	.3182E+00	.0000E+00	-.6434E-01
.2685E+00	.8000E+00	.0000E+00	.0000E+00	.2257E+00	.0000E+00	-.6434E-01
.2700E+00	.8000E+00	.0000E+00	.0000E+00	.2257E+00	.0000E+00	-.6434E-01
.2700E+00	.8000E+00	.0000E+00	.0000E+00	.2257E+00	.0000E+00	.0000E+00

Fig 3. Example of the second output file. Note that a large section of the file is cut out and only the top and the bottom of the file is shown. The file was calculated from the input file shown in Fig. 1. See remarks on units in section 5.

5. Units

Units can be chosen freely when specifying variables in the input file, as long as they are consistent. The units in the two output files are entirely depended on the units used in the input file. An example of consistent units is given below:

Variable	Unit
Input file (Named by user):	
Depth coordinate (x)	cm
Concentration (C)	nmol cm^{-3}
Flux (optional boundary condition)	nmol cm^{-2} s^{-1}
Diffusivity (D)	cm^2 s^{-1}
Porosity (φ)	-
Biodiffusivity (D_B)	cm^2 s^{-1}
Irrigation coefficient (α)	s^{-1}
First output file (PROFILE.log):	
Concentration (C)	nmol cm^{-3}
Flux	nmol cm^{-2} s^{-1}
Depth integrated production	nmol cm^{-2} s^{-1}
Depth integrated irrigation	nmol cm^{-2} s^{-1}
Depth integrated production in each zone	nmol cm^{-2} s^{-1}
Depth integrated irrigation in each zone	nmol cm^{-2} s^{-1}
Second output file (named by user):	
Depth coordinate (x)	cm
Porosity (φ)	-
Irrigation coefficient (α)	s^{-1}
Concentration (C)	nmol cm^{-3}
Irrigation rate ($\varphi\alpha(C_0 - C)$)	nmol cm^{-3} s^{-1}
Production rate (R)	nmol cm^{-3} s^{-1}

PS: Note that concentrations (C) are given per volume of porewater, while production rates (R) and irrigation rates ($\varphi\alpha(C_0 - C)$) are given per volume of sediment.