FLOWNORM 2.0

- a Visual Basic program for computing riverine loads of substances and extracting anthropogenic signals from time series of load data

User's Manual

2004-04-13

Anders Grimvall
Department of Mathematics
Linköping University

Introduction

Monitoring of river water quality and discharge plays a key role in many water management programmes. In particular, such monitoring can provide important information about the loads of nutrients and toxic substances that are carried by rivers to the sea. However, increases or decreases in riverine loads must be interpreted with care, because the natural variation between years can be very large. Both the leaching of dissolved substances from soil to water and the transport of particulate matter can increase significantly during periods of high precipitation. Moreover, runoff, temperature and light conditions can strongly influence processes in the river itself. This calls for efficient procedures to separate human impact from natural variation in the collected data.

The Visual Basic programme FLOWNORM has been developed to facilitate:

- **calculation of riverine loads of substances;**
- **extraction of anthropogenic signals from time series of riverine loads.**

FLOWNORM, Version 2.0, consists of five Visual Basic macros having the names and functions listed below. Two Excel worksheets containing concentration and flow data for an arbitrary number of sampling sites form the starting point for an analysis. The final result consists of time series of **normalised loads**, i.e. riverine loads that have been adjusted to remove natural fluctuations and clarify anthropogenic impacts.

Macro	**Function**
auditdailydata	Identify illegal entries in raw data
matchflowandcons	Define pairs of flow and concentration data for load calculations
computeloads	Compute monthly and annual loads from flow and concentration data ordered by date
definenormalisationmodels	Define response variables and normalisation models
flownormalise	Select normalisation models by cross-validation and compute monthly and annual normalised loads

Each macro operates on predefined worksheets for inputs and outputs. The table below shows which worksheets that are used for the different macros. Further details are given in the documentation of each macro.

Macro	Input worksheets	Output worksheets
auditdailydata	"Concentration by date" or "Flow by date"	The same as input, and "Concentration data summary" or "Flow data summary"
matchflowandcons	"Concentration by date" and "Flow by date"	"Matched pairs"
computeloads	"Matched pairs", "Flow by date", and "Concentration by date"	"Annual totals" and "Seasonal totals"
definenormalisationmodels	"Seasonal totals"	"Normalisation models"
flownormalise	"Normalisation models"	"Normalisation models", "Normalised seasonal totals", and "Normalised annual totals"

Theoretical background

Estimation of monthly and annual riverine loads

Monthly and annual riverine loads are calculated by first expanding the time series of observed concentration and flow data to complete series of daily data and then summing daily values of the product of concentration and water discharge. The expanded values are computed by connecting observed values with straight lines.

Normalisation of riverine loads

Monthly values of riverine loads are normalised by employing parametric and semiparametric regression models to remove or suppress the temporal variation that can be attributed to fluctuations in water discharge, water temperature, salinity, or other indicators of natural variation.

The semiparametric normalisation model has the general form

$$y_{ij} = \alpha_{ij} + \beta_{1,j} x_{1,ij} + ... + \beta_{p,j} x_{p,ij} + \varepsilon_{ij}, \quad i = 1, ..., n \quad j = 1, ..., m$$

where y_{ij} is the observed response for the jth month of the ith year, $x_{k,ij}$, k=1, … ,p represent contemporaneous values of p explanatory variables, and ε_{ij} is a random error term with mean zero. The slope parameters ($\beta_{k,j}$, k=1, … ,p) are permitted to vary with the season (j) under consideration, and the intercept (α_{ij}) is permitted to vary with both season (j) and year (i). However, rapid changes in the intercept are controlled by so-called roughness penalty factors (λ_1 and λ_2), and the intercept and slope parameters are estimated by minimising the expression

$$S(\alpha,\beta)=\sum_{i,j}(y_{ij}-\alpha_{ij}-\beta_{1,j}x_{1,ij}-\ldots-\beta_{p,j}x_{p,ij})^2+\lambda_1\sum_{i,j}(\alpha_{ij}-\frac{\alpha_{i+1,j}+\alpha_{i-1,j}}{2})^2$$

$$+\lambda_2\sum_{i,j}(\alpha_{ij}-\frac{\alpha_{i,j-1}+\alpha_{i,j+1}}{2})^2,$$

where the first sum ranges over all values of *i* and *j* for which both the response variable and the explanatory variables have been observed. A univariate form of this model was first used by Stålnacke and co-workers (1999), and detailed information about algorithms for parameter estimation has been published by Stålnacke and Grimvall (2001).

The parametric normalisation model is an ordinary multiple regression model of the general form

$$y_{ij}=\alpha+\beta_{1,j}\,x_{1,ij}+\ldots+\beta_{p,j}\,x_{p,ij}+\varepsilon_{ij},\quad i=1,\ldots,n\quad j=1,\ldots,m$$

where y_{ij}, $x_{k,ij}$, and ε_{ij} have the same meaning as above. The slope parameters ($\beta_{k,j}$, k=1, … ,p) are permitted to vary with the season (j) under consideration, whereas the intercept is assumed to be constant.

Selection of roughness penalty factors and assessment of the predictive ability of the tested normalisation models

The penalty factors λ_1 and λ_2 are determined by cross-validation. With this technique, the entire data set is separated into an estimation set (or training set) and a test set. The model is first fitted to the estimation set and is subsequently used to predict the observations in the test set, that is, the values that have been left out of the estimation step. If the observation period covers *m* years, we define *m* estimation sets M_i, i=1, … ,*m* by leaving out one-year-long blocks of observations, and then we compute a so-called PRESS-value (i.e., a sum of squared prediction errors):

$$S(\lambda_1,\lambda_2)=\sum_i\sum_{(i,j)\notin M_i}(y_{ij}-\hat{\alpha}_{ij}-\hat{\beta}_{1,j}x_{1,ij}-\ldots-\hat{\beta}_{p,j}x_{p,ij})^2.$$

Finally, the factors λ_1 and λ_2 are selected in such a way that $S(\lambda_1,\lambda_2)$ is minimised, and the corresponding Root Mean *PRESS* value

$$min\left\{\sqrt{\frac{1}{N}S(\lambda_1,\lambda_2)};\lambda_1>0,\lambda_2>0\right\}$$

is used as a measure of the predictive ability of the normalisation model under consideration.

Literature references

Stålnacke P, Grimvall, A., Sundblad, K., and Wilander, A.: 1999, 'Trends in nitrogen transport in Swedish rivers', *Environ. Monit. Assess*. **59**, 47-72.

Stålnacke, P., and Grimvall, A.: 2001, 'Semiparametric approaches to flow-normalisation and source apportionment of substance transport in rivers', *Environmetrics* **12**, 233-250.

Input data

The concentration and flow data that shall be analysed are pasted on the worksheets 'Concentration by date" and 'Flow by date', respectively. Figure 1 shows how the input data shall be organised for the worksheet 'Concentration by date'. For each sampling site, a date column shall be followed by an arbitrary number of columns containing the values of the monitored variables. Strings found in the row immediately above the first observations (row 3 in Figure 1) are interpreted as variable names. Sampling site names are extracted from the date columns two lines above the first observation. Missing values shall be entered as empty cells. 'Less-than-values' of the form <0.05 are permitted.

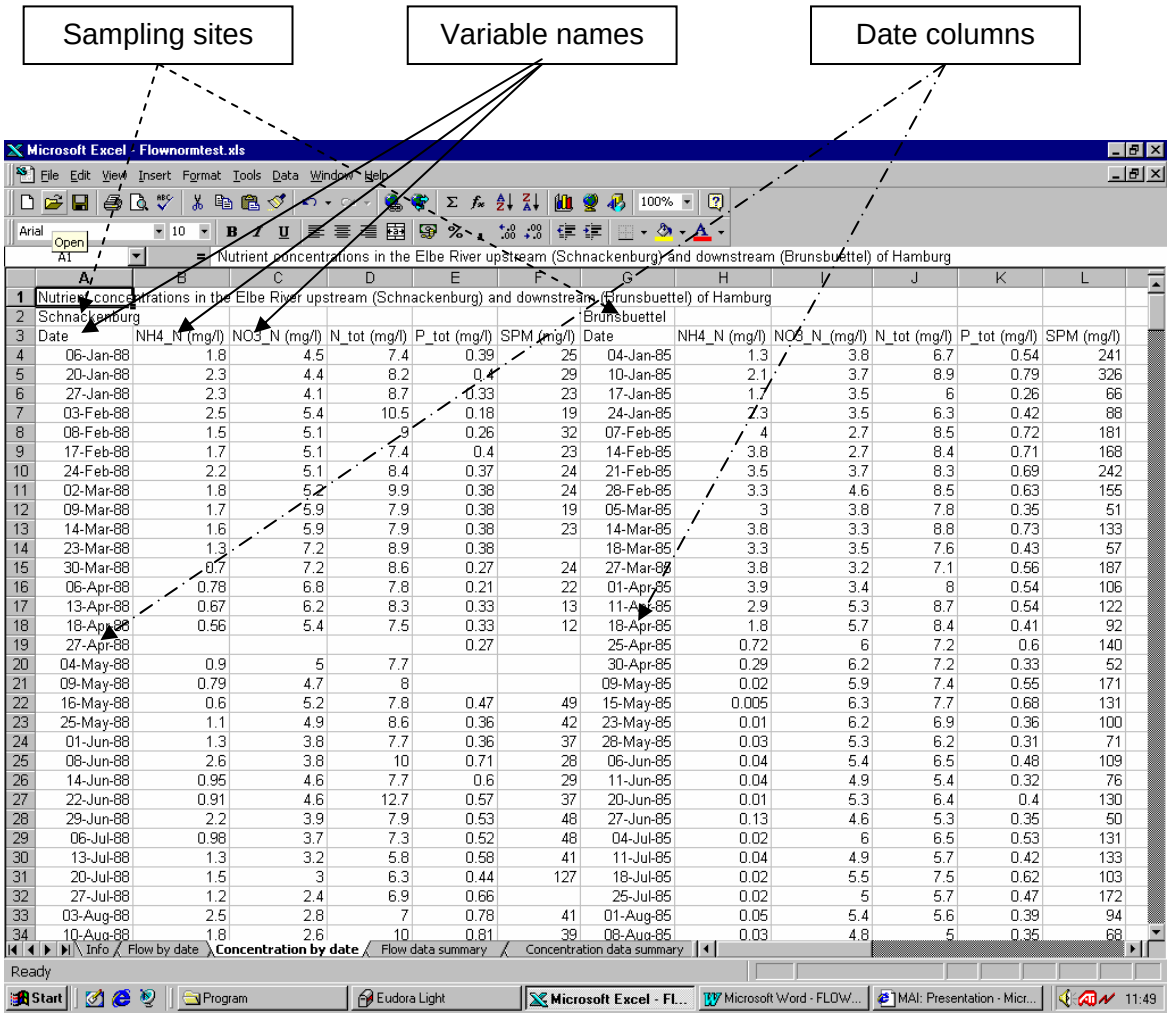


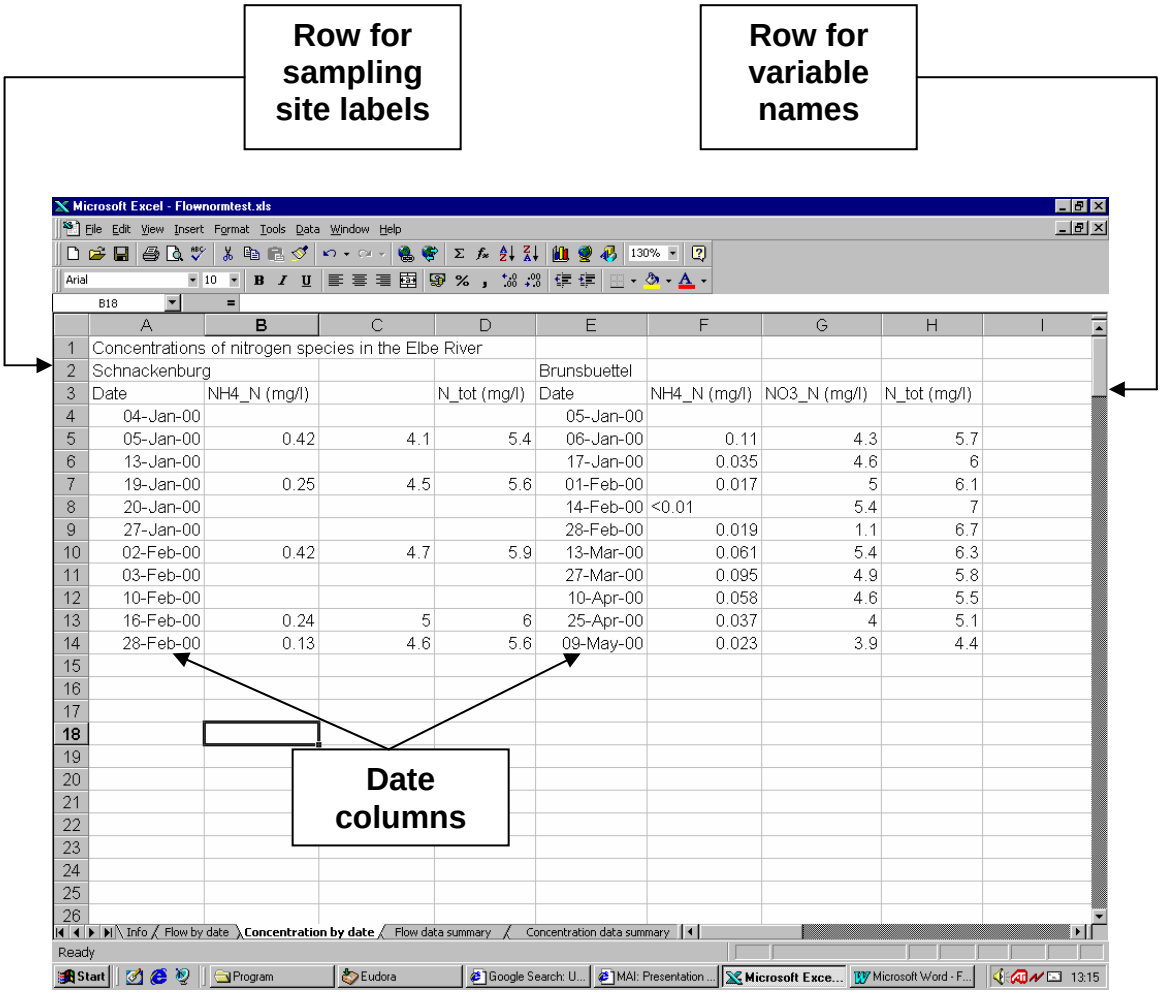
Figure 1. Worksheet 'Concentration by date' for input of concentration data.

Flow data are organised analogously. However, only one column of flow-data is permitted for each sampling site name.

Auditdailydata

This macro checks that the data pasted on the worksheets 'Concentration by date' and 'Flow by date' are of correct type and properly organised. Missing values shall be entered as empty cells. 'Less-than-values' are permitted.

When 'Auditdailydata' is run, the macro first identifies the name of active worksheet ('Concentration by date' or 'Flow by date'), and then it searches for cells containing date values. To be more precise, the macro searches for the first row containing date values and then uses the cell values in this row to determine which of the columns that shall be regarded as date columns. In figure 2, column A and E will be identified as date columns.

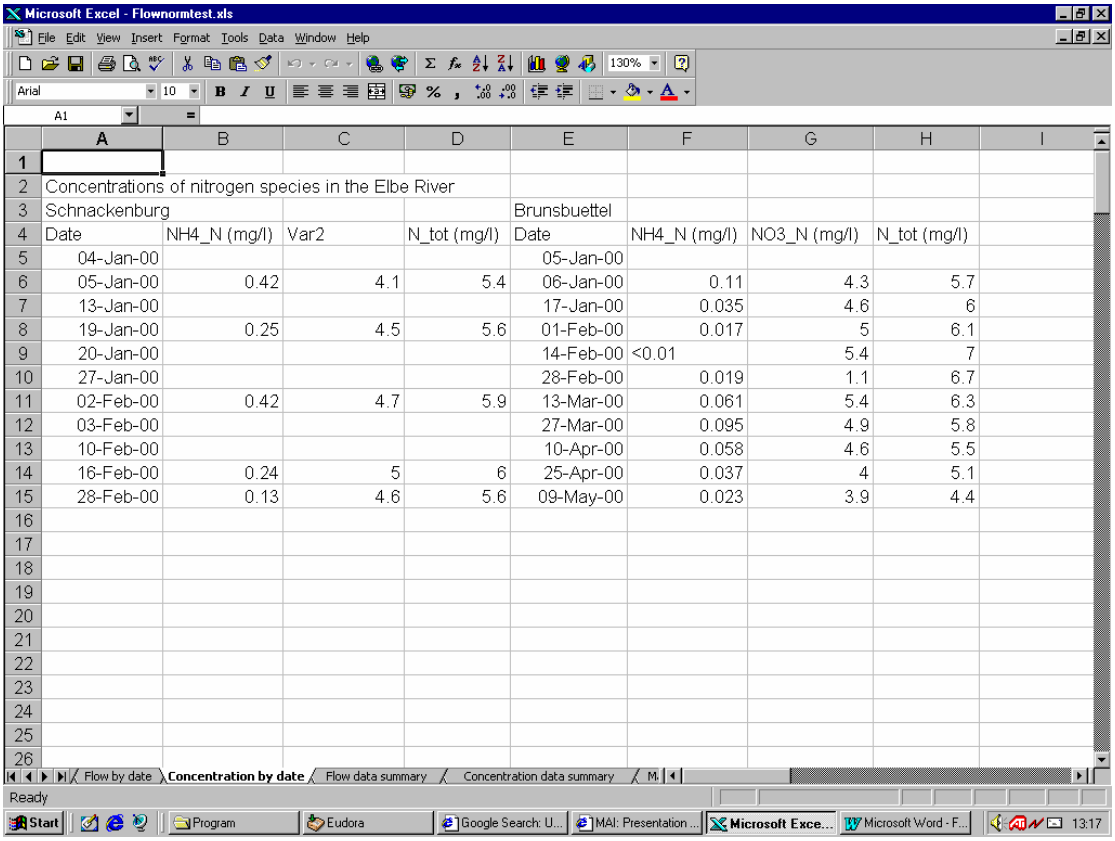


	A	B	C	D	E	F	G	H	I
1	Concentrations of nitrogen species in the Elbe River								
2	Schnackenburg				Brunsbuettel				
3	Date	NH4_N (mg/l)		N_tot (mg/l)	Date	NH4_N (mg/l)	NO3_N (mg/l)	N_tot (mg/l)	
4	04-Jan-00				05-Jan-00				
5	05-Jan-00	0.42	4.1	5.4	06-Jan-00	0.11	4.3	5.7	
6	13-Jan-00				17-Jan-00	0.035	4.6	6	
7	19-Jan-00	0.25	4.5	5.6	01-Feb-00	0.017	5	6.1	
8	20-Jan-00				14-Feb-00	<0.01	5.4	7	
9	27-Jan-00				28-Feb-00	0.019	1.1	6.7	
10	02-Feb-00	0.42	4.7	5.9	13-Mar-00	0.061	5.4	6.3	
11	03-Feb-00				27-Mar-00	0.095	4.9	5.8	
12	10-Feb-00				10-Apr-00	0.058	4.6	5.5	
13	16-Feb-00	0.24	5	6	25-Apr-00	0.037	4	5.1	
14	28-Feb-00	0.13	4.6	5.6	09-May-00	0.023	3.9	4.4	

Figure 2. Worksheet 'Concentration by date' before the macro 'Auditdailydata' is run.

Variable names are extracted from the row immediately above the first row containing date values. If no variable name has been assigned to a variable (see cell C3 in Figure 2), the macro assigns a default name. Names of sampling sites are extracted from cells located in date columns and two rows above the first row with date values (cells A2 and E2 in figure 2). A maximum of two rows above the sampling site names can be used for comments.

Figure 3 shows the results obtained by letting ‘Auditdailydata’ operate on the worksheet in Figure 2. A default name for the variable in column C has been added, and an extra line for comments has been inserted. The number of observations for each of the identified variables is shown on the worksheet “Concentration data summary”.



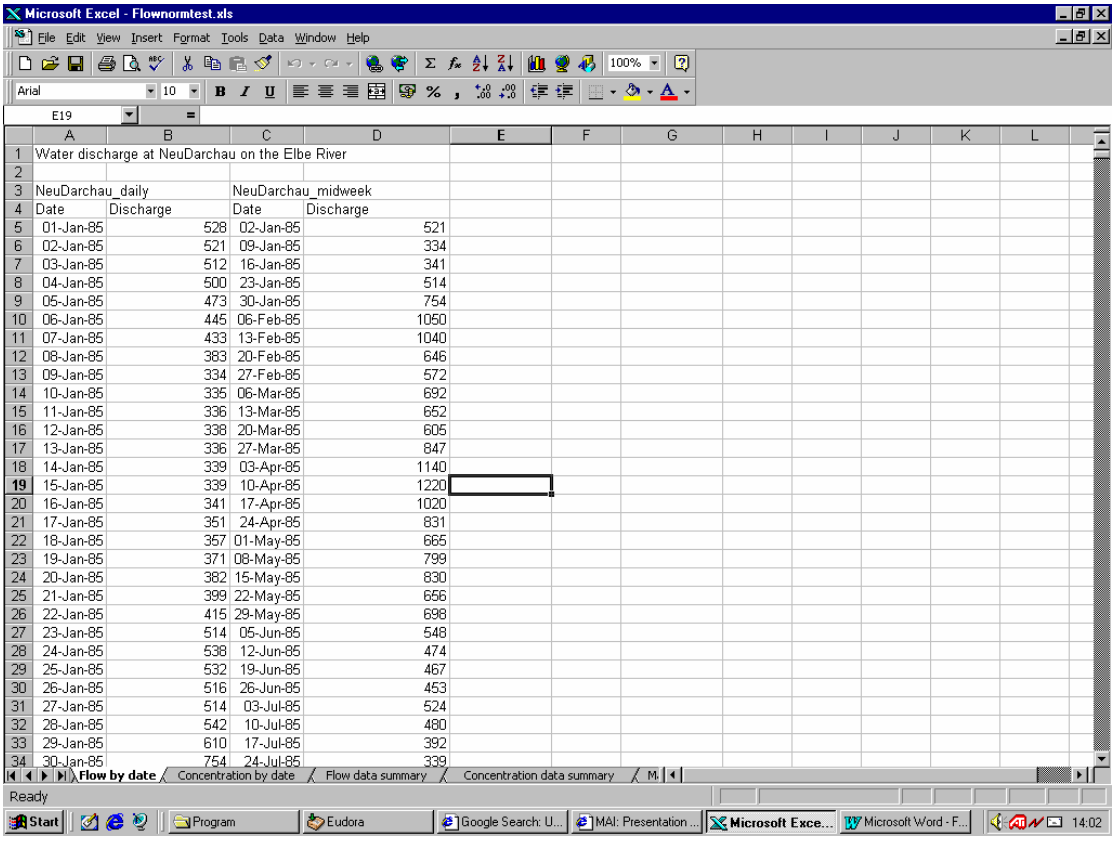
	A	B	C	D	E	F	G	H	I
1									
2	Concentrations of nitrogen species in the Elbe River								
3	Schnackenburg				Brunsbuettel				
4	Date	NH4_N (mg/l)	Var2	N_tot (mg/l)	Date	NH4_N (mg/l)	NO3_N (mg/l)	N_tot (mg/l)	
5	04-Jan-00				05-Jan-00				
6	05-Jan-00	0.42	4.1	5.4	06-Jan-00	0.11	4.3	5.7	
7	13-Jan-00				17-Jan-00	0.035	4.6	6	
8	19-Jan-00	0.25	4.5	5.6	01-Feb-00	0.017	5	6.1	
9	20-Jan-00				14-Feb-00	<0.01	5.4	7	
10	27-Jan-00				28-Feb-00	0.019	1.1	6.7	
11	02-Feb-00	0.42	4.7	5.9	13-Mar-00	0.061	5.4	6.3	
12	03-Feb-00				27-Mar-00	0.095	4.9	5.8	
13	10-Feb-00				10-Apr-00	0.058	4.6	5.5	
14	16-Feb-00	0.24	5	6	25-Apr-00	0.037	4	5.1	
15	28-Feb-00	0.13	4.6	5.6	09-May-00	0.023	3.9	4.4	
16									
17									
18									
19									
20									
21									
22									
23									
24									
25									
26									

Figure 3. Worksheet 'Concentration by date' after the macro 'Auditdailydata' has been run.

Matchflowandconc

This macro aims to facilitate the matching of flow and concentration data for load calculations. Based on the identified names of the sampling sites for flow and concentration the macro prints a preliminary list of matched pairs on the worksheet ‘Matched pairs’. This list can then be edited prior to the load calculations.

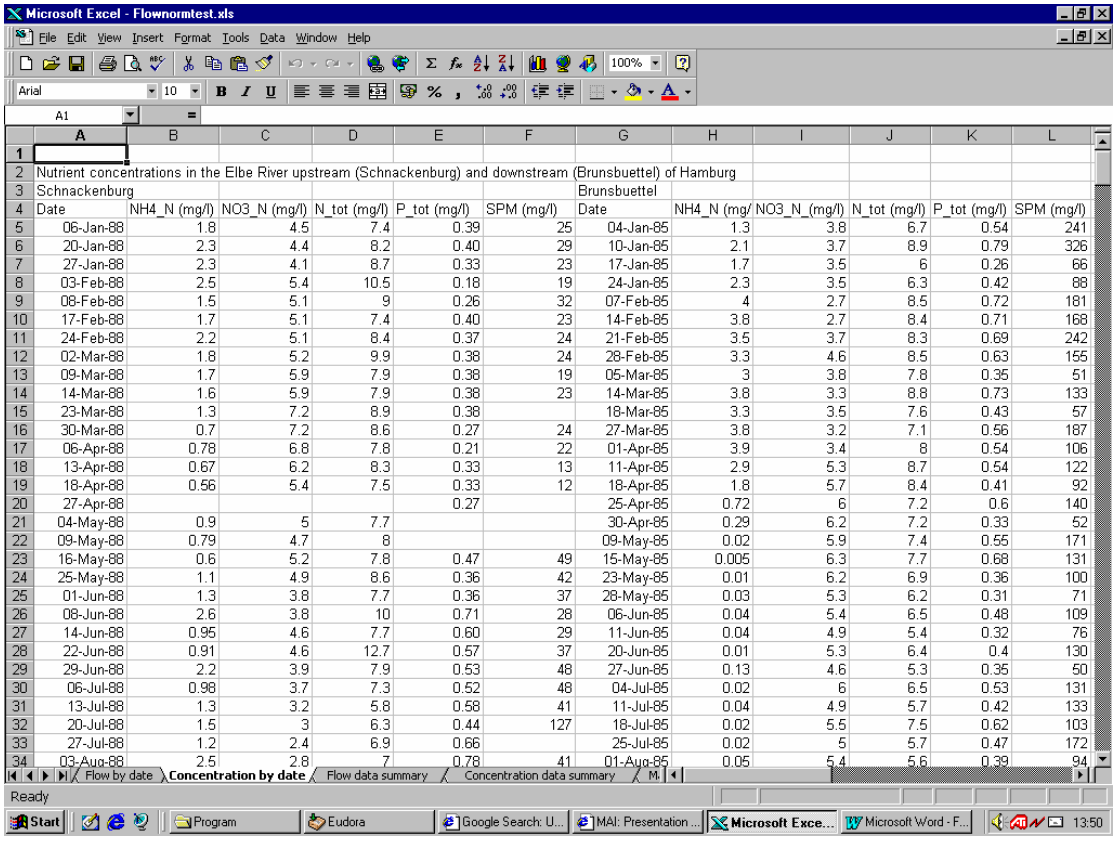
Figures 4-6 show an example of input worksheets to the macro and the output worksheet before editing. As can be seen, the macro has identified two sets of flowdata (daily data and midweek data) from NeuDarchau on the Elbe. Furthermore, concentration data has been found for two sites: Schnackenburg and Brunsbuettel. Suppose that we are only interested in calculating loads using midweek discharge data. We can then edit worksheet 'Matched pairs' as shown in Figure 7, where the pairs (NeuDarchau_midweek, Schnackenburg) and (NeuDarchau_midweek, Brunsbuettel) have been listed.



Microsoft Excel - Flownormtest.xls

	A	B	C	D
1	Water discharge at NeuDarchau on the Elbe River			
2				
3	NeuDarchau_daily		NeuDarchau_midweek	
4	Date	Discharge	Date	Discharge
5	01-Jan-85	528	02-Jan-85	521
6	02-Jan-85	521	09-Jan-85	334
7	03-Jan-85	512	16-Jan-85	341
8	04-Jan-85	500	23-Jan-85	514
9	05-Jan-85	473	30-Jan-85	754
10	06-Jan-85	445	06-Feb-85	1050
11	07-Jan-85	433	13-Feb-85	1040
12	08-Jan-85	383	20-Feb-85	646
13	09-Jan-85	334	27-Feb-85	572
14	10-Jan-85	335	06-Mar-85	692
15	11-Jan-85	336	13-Mar-85	652
16	12-Jan-85	338	20-Mar-85	605
17	13-Jan-85	336	27-Mar-85	847
18	14-Jan-85	339	03-Apr-85	1140
19	15-Jan-85	339	10-Apr-85	1220
20	16-Jan-85	341	17-Apr-85	1020
21	17-Jan-85	351	24-Apr-85	831
22	18-Jan-85	357	01-May-85	665
23	19-Jan-85	371	08-May-85	799
24	20-Jan-85	382	15-May-85	830
25	21-Jan-85	399	22-May-85	656
26	22-Jan-85	415	29-May-85	698
27	23-Jan-85	514	05-Jun-85	548
28	24-Jan-85	538	12-Jun-85	474
29	25-Jan-85	532	19-Jun-85	467
30	26-Jan-85	516	26-Jun-85	453
31	27-Jan-85	514	03-Jul-85	524
32	28-Jan-85	542	10-Jul-85	480
33	29-Jan-85	610	17-Jul-85	392
34	30-Jan-85	754	24-Jul-85	339

Figure 4. Flow data on worksheet 'Flow by date'



Microsoft Excel - Flownormtest.xls

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2	Nutrient concentrations in the Elbe River upstream (Schnackenburg) and downstream (Brunsbuettel) of Hamburg											
3	Schnackenburg						Brunsbuettel					
4	Date	NH4_N (mg/l)	NO3_N (mg/l)	N_tot (mg/l)	P_tot (mg/l)	SPM (mg/l)	Date	NH4_N (mg/	NO3_N_(mg/l)	N_tot (mg/l)	P_tot (mg/l)	SPM (mg/l)
5	06-Jan-88	1.8	4.5	7.4	0.39	25	04-Jan-85	1.3	3.8	6.7	0.54	241
6	20-Jan-88	2.3	4.4	8.2	0.40	29	10-Jan-85	2.1	3.7	8.9	0.79	326
7	27-Jan-88	2.3	4.1	8.7	0.33	23	17-Jan-85	1.7	3.5	6	0.26	66
8	03-Feb-88	2.5	5.4	10.5	0.18	19	24-Jan-85	2.3	3.5	6.3	0.42	88
9	08-Feb-88	1.5	5.1	9	0.26	32	07-Feb-85	4	2.7	8.5	0.72	181
10	17-Feb-88	1.7	5.1	7.4	0.40	23	14-Feb-85	3.8	2.7	8.4	0.71	168
11	24-Feb-88	2.2	5.1	8.4	0.37	24	21-Feb-85	3.5	3.7	8.3	0.69	242
12	02-Mar-88	1.8	5.2	9.9	0.38	24	28-Feb-85	3.3	4.6	8.5	0.63	155
13	09-Mar-88	1.7	5.9	7.9	0.38	19	05-Mar-85	3	3.8	7.8	0.35	51
14	14-Mar-88	1.6	5.9	7.9	0.38	23	14-Mar-85	3.8	3.3	8.8	0.73	133
15	23-Mar-88	1.3	7.2	8.9	0.38		18-Mar-85	3.3	3.5	7.6	0.43	57
16	30-Mar-88	0.7	7.2	8.6	0.27	24	27-Mar-85	3.8	3.2	7.1	0.56	187
17	06-Apr-88	0.78	6.8	7.8	0.21	22	01-Apr-85	3.9	3.4	8	0.54	106
18	13-Apr-88	0.67	6.2	8.3	0.33	13	11-Apr-85	2.9	5.3	8.7	0.54	122
19	18-Apr-88	0.56	5.4	7.5	0.33	12	18-Apr-85	1.8	5.7	8.4	0.41	92
20	27-Apr-88				0.27		25-Apr-85	0.72	6	7.2	0.6	140
21	04-May-88	0.9	5	7.7			30-Apr-85	0.29	6.2	7.2	0.33	52
22	09-May-88	0.79	4.7	8			09-May-85	0.02	5.9	7.4	0.55	171
23	16-May-88	0.6	5.2	7.8	0.47	49	15-May-85	0.005	6.3	7.7	0.68	131
24	25-May-88	1.1	4.9	8.6	0.36	42	23-May-85	0.01	6.2	6.9	0.36	100
25	01-Jun-88	1.3	3.8	7.7	0.36	37	28-May-85	0.03	5.3	6.2	0.31	71
26	08-Jun-88	2.6	3.8	10	0.71	28	06-Jun-85	0.04	5.4	6.5	0.48	109
27	14-Jun-88	0.95	4.6	7.7	0.60	29	11-Jun-85	0.04	4.9	5.4	0.32	76
28	22-Jun-88	0.91	4.6	12.7	0.57	37	20-Jun-85	0.01	5.3	6.4	0.4	130
29	29-Jun-88	2.2	3.9	7.9	0.53	48	27-Jun-85	0.13	4.6	5.3	0.35	50
30	06-Jul-88	0.98	3.7	7.3	0.52	48	04-Jul-85	0.02	6	6.5	0.53	131
31	13-Jul-88	1.3	3.2	5.8	0.58	41	11-Jul-85	0.04	4.9	5.7	0.42	133
32	20-Jul-88	1.5	3	6.3	0.44	127	18-Jul-85	0.02	5.5	7.5	0.62	103
33	27-Jul-88	1.2	2.4	6.9	0.66		25-Jul-85	0.02	5	5.7	0.47	172
34	03-Aug-88	2.5	2.8	7	0.78	41	01-Aug-85	0.05	5.4	5.6	0.39	94

Figure 5. Concentration data on worksheet 'Concentration by date'

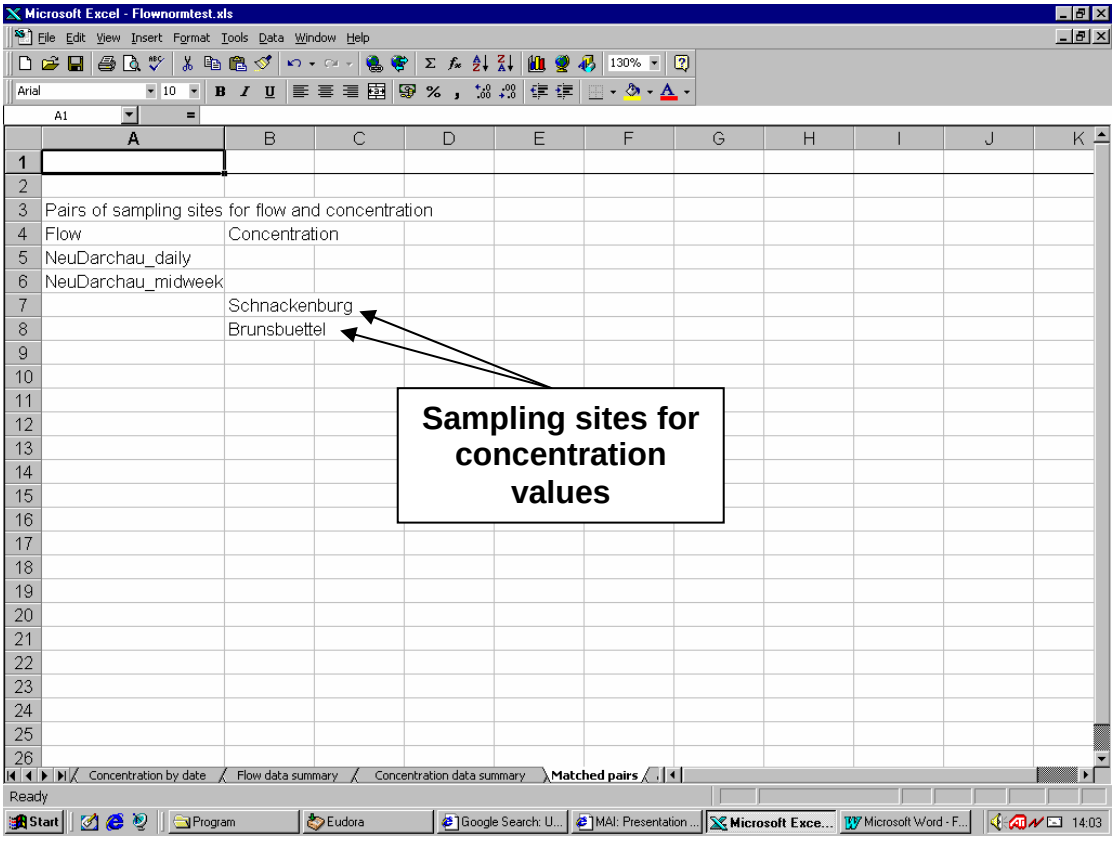


Figure 6. Identified sampling sites for flow and concentration.

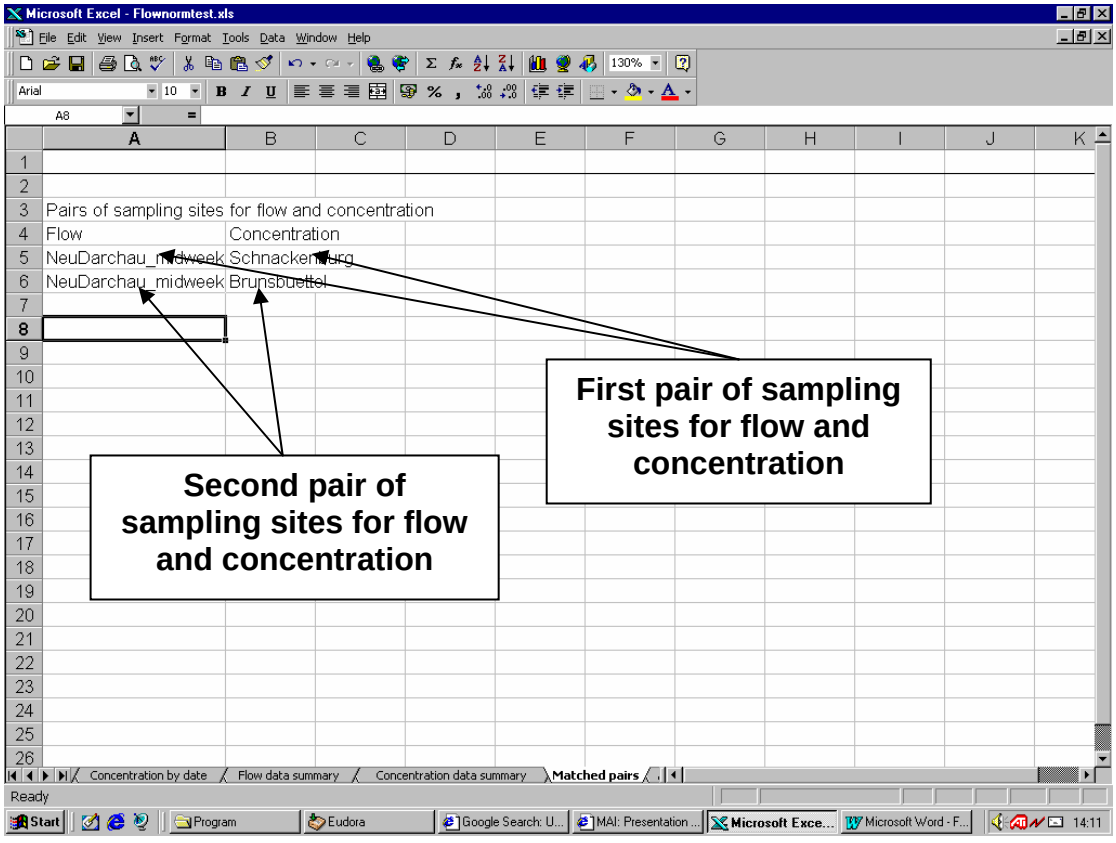


Figure 7. Worksheet 'Matched pairs' after editing

Computeloads

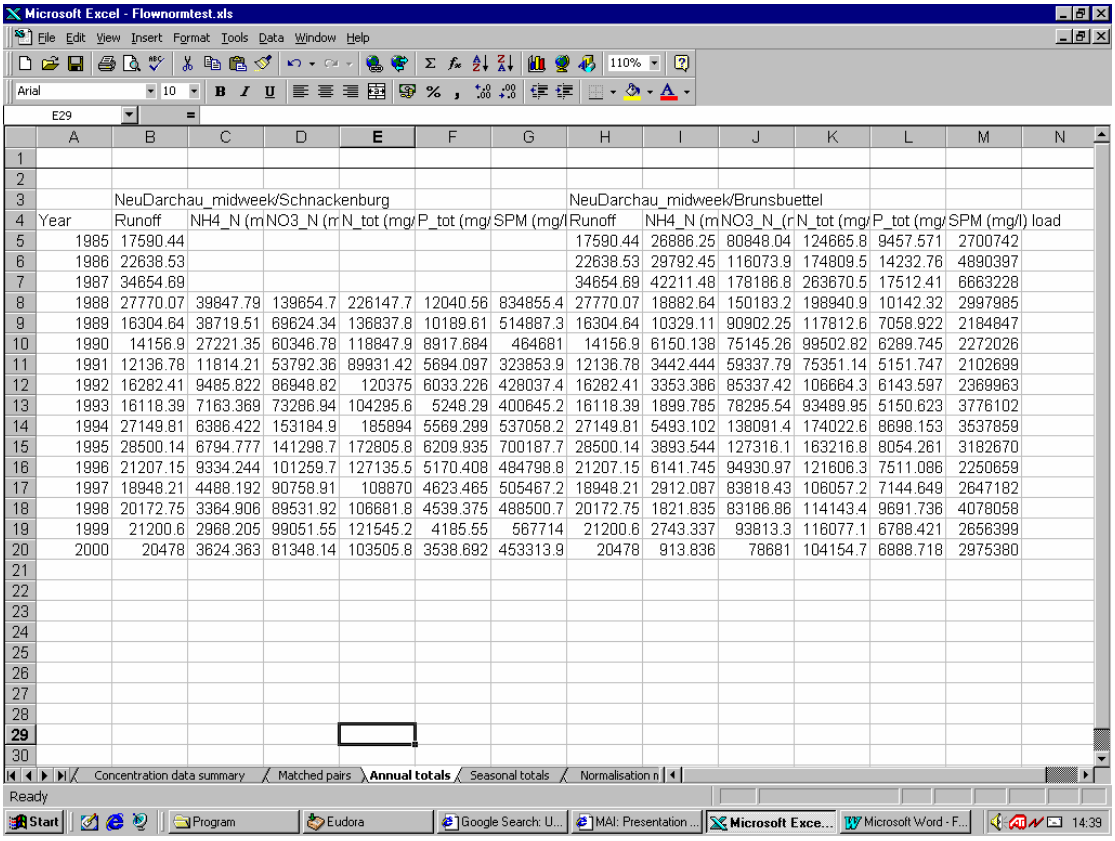
This macro operates on the worksheets 'Flow by date', 'Concentration by date' and 'Matched pairs'. The output worksheets are 'Annual totals' and 'Monthly totals'. Figure 8 and 9 illustrate the output obtained for the input shown in Figures 4, 5 and 7.

Because no concentration data were available for Schnackenburg in 1985-87, no substance loads were calculated for that period (see Figure 8). To avoid misleading extrapolations and interpolations of observed data, the following general rules are applied:

- the calculation of substance loads starts the first month for which both flow and concentration data are available and ends the last month for which such data exist;
- gaps in flow data are filled by interpolated values provided that the gap does not contain a full month without any observations;
- gaps in concentration data are filled by interpolated values provided that the gap contains a maximum of two full months without any observations;
- annual loads are only given for years with a complete set of monthly loads.

'Less-than-values' are replaced by a fixed percentage of the detection limit. The user is asked to enter the desired percentage when the macro is run.

If the water discharge is expressed in m^3/s and the concentrations in mg/l, the riverine loads will be expressed in ton/month or ton/year. The monthly and annual totals for the water discharge are expressed in 10^9 m^3 per month and year, respectively.



Microsoft Excel - Flownormtest.xls

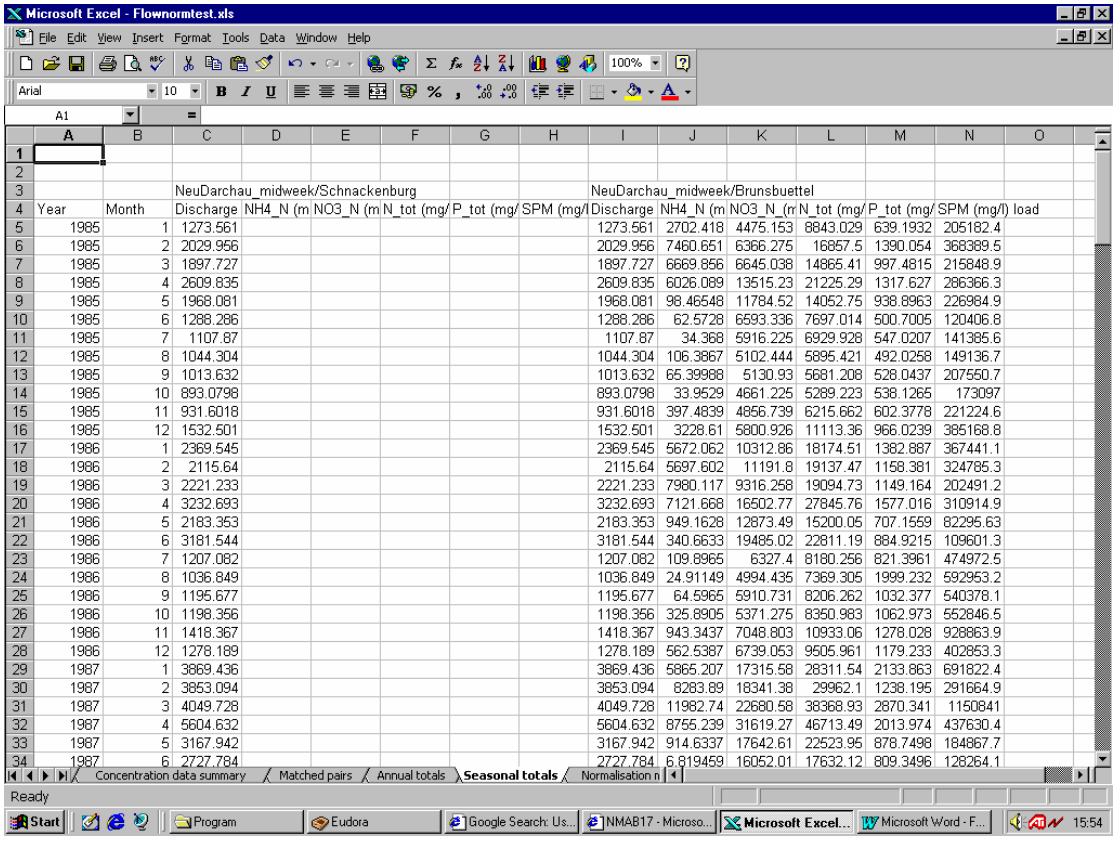
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1														
2														
3		NeuDarchau_midweek/Schnackenburg						NeuDarchau_midweek/Brunsbuettel						
4	Year	Runoff	NH4_N (m	NO3_N (m	N_tot (mg/	P_tot (mg/	SPM (mg/l	Runoff	NH4_N (m	NO3_N_(r	N_tot (mg/	P_tot (mg/	SPM (mg/l)	load
5	1985	17590.44						17590.44	26886.25	80848.04	124665.8	9457.571	2700742	
6	1986	22638.53						22638.53	29792.45	116073.9	174809.5	14232.76	4890397	
7	1987	34654.69						34654.69	42211.48	178186.8	263670.5	17512.41	6663228	
8	1988	27770.07	39847.79	139654.7	226147.7	12040.56	834855.4	27770.07	18882.64	150183.2	198940.9	10142.32	2997985	
9	1989	16304.64	38719.51	69624.34	136837.8	10189.61	514887.3	16304.64	10329.11	90902.25	117812.6	7058.922	2184847	
10	1990	14156.9	27221.35	60346.78	118847.9	8917.684	464681	14156.9	6150.138	75145.26	99502.82	6289.745	2272026	
11	1991	12136.78	11814.21	53792.36	89931.42	5894.097	323853.9	12136.78	3442.444	59337.79	75351.14	5151.747	2102699	
12	1992	16282.41	9485.822	86948.82	120375	6033.226	428037.4	16282.41	3353.386	85337.42	106664.3	6143.597	2369963	
13	1993	16118.39	7163.369	73286.94	104295.6	5248.29	400645.2	16118.39	1899.785	78295.54	93489.95	5150.623	3776102	
14	1994	27149.81	6386.422	153184.9	185894	5569.299	537058.2	27149.81	5493.102	138091.4	174022.6	8698.153	3537859	
15	1995	28500.14	6794.777	141298.7	172805.8	6209.935	700187.7	28500.14	3893.544	127316.1	163216.8	8054.261	3182670	
16	1996	21207.15	9334.244	101259.7	127135.5	5170.408	484798.8	21207.15	6141.745	94930.97	121606.3	7511.086	2250659	
17	1997	18948.21	4488.192	90758.91	108870	4623.465	505467.2	18948.21	2912.087	83818.43	106057.2	7144.649	2647182	
18	1998	20172.75	3364.906	89531.92	106681.8	4539.375	488500.7	20172.75	1821.835	83186.86	114143.4	9691.736	4078058	
19	1999	21200.6	2968.205	99051.55	121545.2	4185.55	567714	21200.6	2743.337	93813.3	116077.1	6788.421	2656399	
20	2000	20478	3624.363	81348.14	103505.8	3538.692	453313.9	20478	913.836	78681	104154.7	6888.718	2975380	
21														
22														
23														
24														
25														
26														
27														
28														
29														
30														

Concentration data summary | Matched pairs | Annual totals | Seasonal totals | Normalisation n

Ready

Start | Program | Eudora | Google Search: U... | MAI: Presentation ... | Microsoft Exce... | Microsoft Word - F... | 14:39

Figure 8. Annual loads of the substances monitored at Schnackenburg and Brunsbuettel.



Microsoft Excel - Flownormtest.xls

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1															
2															
3			NeuDarchau_midweek/Schnackenburg						NeuDarchau_midweek/Brunsbuettel						
4	Year	Month	Discharge	NH4_N (m	NO3_N (m	N_tot (mg/	P_tot (mg/	SPM (mg/l	Discharge	NH4_N (m	NO3_N_(m	N_tot (mg/	P_tot (mg/	SPM (mg/l)	load
5	1985	1	1273.561						1273.561	2702.418	4475.153	8843.029	639.1932	205182.4	
6	1985	2	2029.956						2029.956	7460.651	6366.275	16857.5	1390.054	368389.5	
7	1985	3	1897.727						1897.727	6669.856	6645.038	14865.41	997.4815	215848.9	
8	1985	4	2609.835						2609.835	6026.089	13515.23	21225.29	1317.627	286366.3	
9	1985	5	1968.081						1968.081	98.46548	11784.52	14052.75	938.8963	226984.9	
10	1985	6	1288.286						1288.286	62.5728	6593.336	7697.014	500.7005	120406.8	
11	1985	7	1107.87						1107.87	34.368	5916.225	6929.928	547.0207	141385.6	
12	1985	8	1044.304						1044.304	106.3867	5102.444	5895.421	492.0258	149136.7	
13	1985	9	1013.632						1013.632	65.39988	5130.93	5681.208	528.0437	207550.7	
14	1985	10	893.0798						893.0798	33.9529	4661.225	5289.223	538.1265	173097	
15	1985	11	931.6018						931.6018	397.4839	4856.739	6215.662	602.3778	221224.6	
16	1985	12	1532.501						1532.501	3228.61	5800.926	11113.36	966.0239	385168.8	
17	1986	1	2369.545						2369.545	5672.062	10312.86	18174.51	1382.887	367441.1	
18	1986	2	2115.64						2115.64	5697.602	11191.8	19137.47	1158.381	324785.3	
19	1986	3	2221.233						2221.233	7980.117	9316.258	19094.73	1149.164	202491.2	
20	1986	4	3232.693						3232.693	7121.668	16502.77	27845.76	1577.016	310914.9	
21	1986	5	2183.353						2183.353	949.1628	12873.49	15200.05	707.1559	82295.63	
22	1986	6	3181.544						3181.544	340.6633	19485.02	22811.19	884.9215	109601.3	
23	1986	7	1207.082						1207.082	109.8965	6327.4	8180.256	821.3961	474972.5	
24	1986	8	1036.849						1036.849	24.91149	4994.435	7369.305	1999.232	592953.2	
25	1986	9	1195.677						1195.677	64.5965	5910.731	8206.262	1032.377	540378.1	
26	1986	10	1198.356						1198.356	325.8905	5371.275	8350.983	1062.973	552846.5	
27	1986	11	1418.367						1418.367	943.3437	7048.803	10933.06	1278.028	928863.9	
28	1986	12	1278.189						1278.189	562.5387	6739.053	9505.961	1179.233	402853.3	
29	1987	1	3869.436						3869.436	5865.207	17315.58	28311.54	2133.863	691822.4	
30	1987	2	3853.094						3853.094	8283.89	18341.38	29962.1	1238.195	291664.9	
31	1987	3	4049.728						4049.728	11982.74	22680.58	38368.93	2870.341	1150841	
32	1987	4	5604.632						5604.632	8755.239	31619.27	46713.49	2013.974	437630.4	
33	1987	5	3167.942						3167.942	914.6337	17642.61	22523.95	878.7498	184867.7	
34	1987	6	2727.784						2727.784	6.819459	16052.01	17632.12	809.3496	128264.1	

Concentration data summary | Matched pairs | Annual totals | Seasonal totals | Normalisation n

Ready

Start | Program | Eudora | Google Search: Us... | NMAB17 - Microso... | Microsoft Excel... | Microsoft Word - F... | 15:54

Figure 9. Monthly loads of the substances monitored at Schnackenburg and Brunsbuettel.

Definenormalisationmodels

This macro, which aims to facilitate the formulation of normalisation models, operates on the worksheet 'Seasonal totals'. Before the macro is run, response variables and explanatory variables shall be defined by typing 'y' or 'x' two rows above the variable names in this worksheet (see Figure 10).

Figure 11 shows a list of possible normalisation models. This list can be edited prior to the normalisation. For example we can decide to use the list in Figure 12 instead.

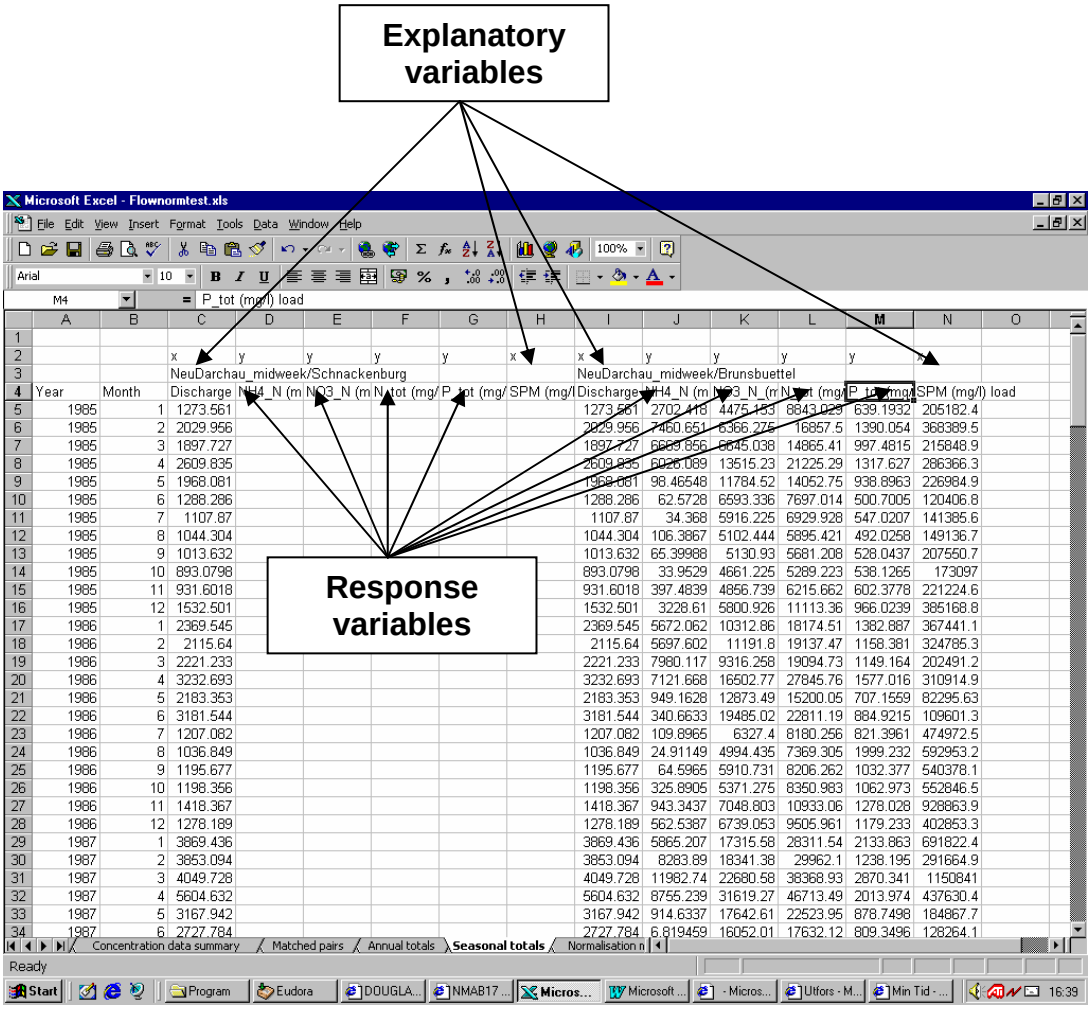


Figure 10. Response variables and explanatory variables indicated by 'y' and 'x', respectively, two rows above the variable names.

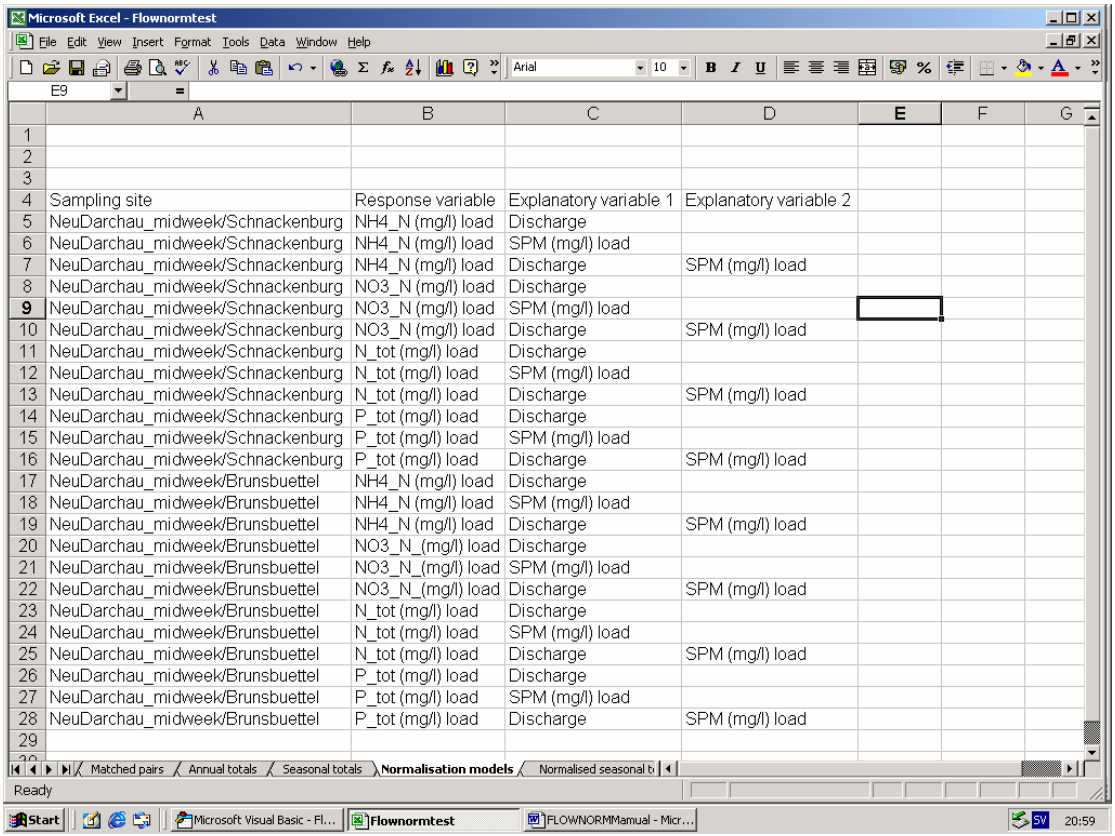


Figure 11. List of normalisation models generated by the macro ‘Definenormalisationmodels’.

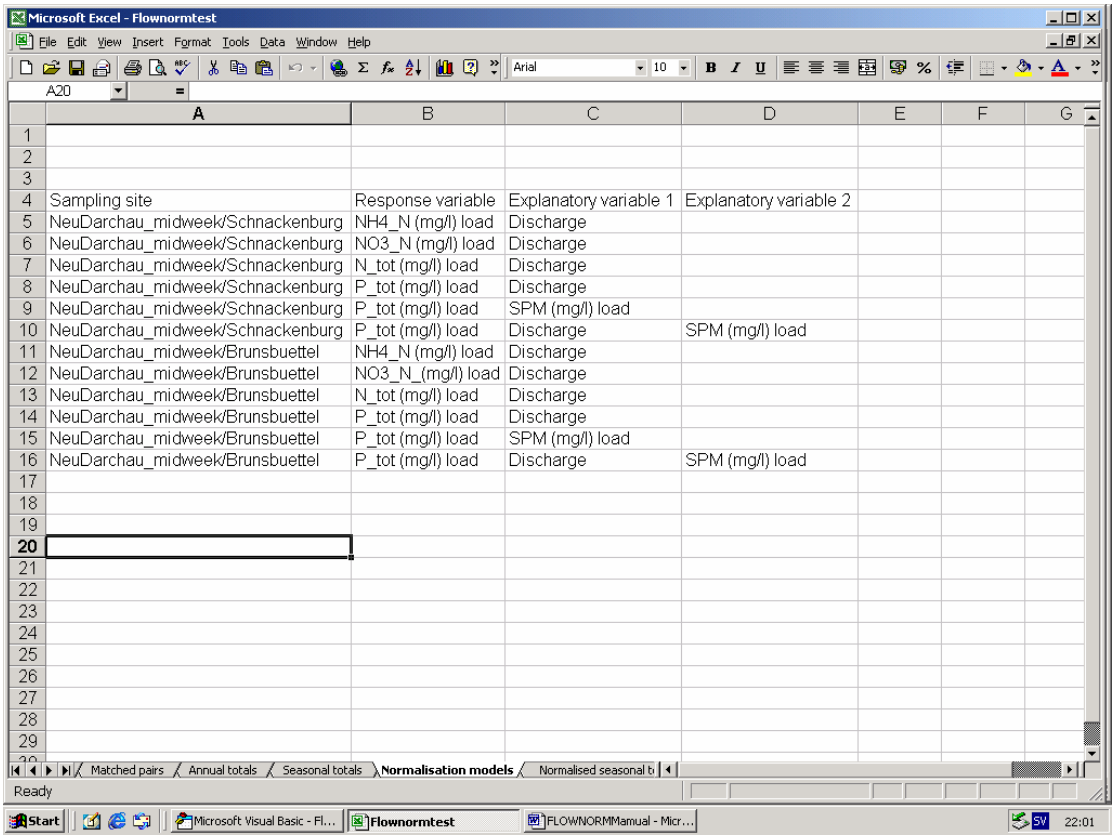


Figure 12. Edited list of normalisation models.

Flownormalise

This macro aims to remove or suppress the natural variation in monthly riverine loads. The response variable and the explanatory variables in each of the tested normalisation models are listed on the worksheet 'Normalisation models'. Input data are read from the worksheet 'Seasonal totals', and the outputs of the macro are printed on the worksheets 'Normalised annual totals' (see Figure 13) and 'Normalised seasonal totals' (not shown).

For each combination of response variable and explanatory variables, normalised values are computed by employing:

- a semiparametric regression model in which the intercept is permitted to vary slowly with season and year;
- an ordinary regression model with constant intercept.

The smoothing parameters in the semiparametric model can be selected by cross-validation or assigned user-defined values. When cross-validation is employed, the number of test sets is equal to the number of years in the observation period, and each test set consists of all observations made the same year.

The worksheet 'Seasonal totals' can be edited before the riverine loads are normalised. For example, it is possible to change the definition of seasons. However, the label 'Year' and the years below that label must be left unchanged, and the season numbers must be typed in the column to the right of the years.

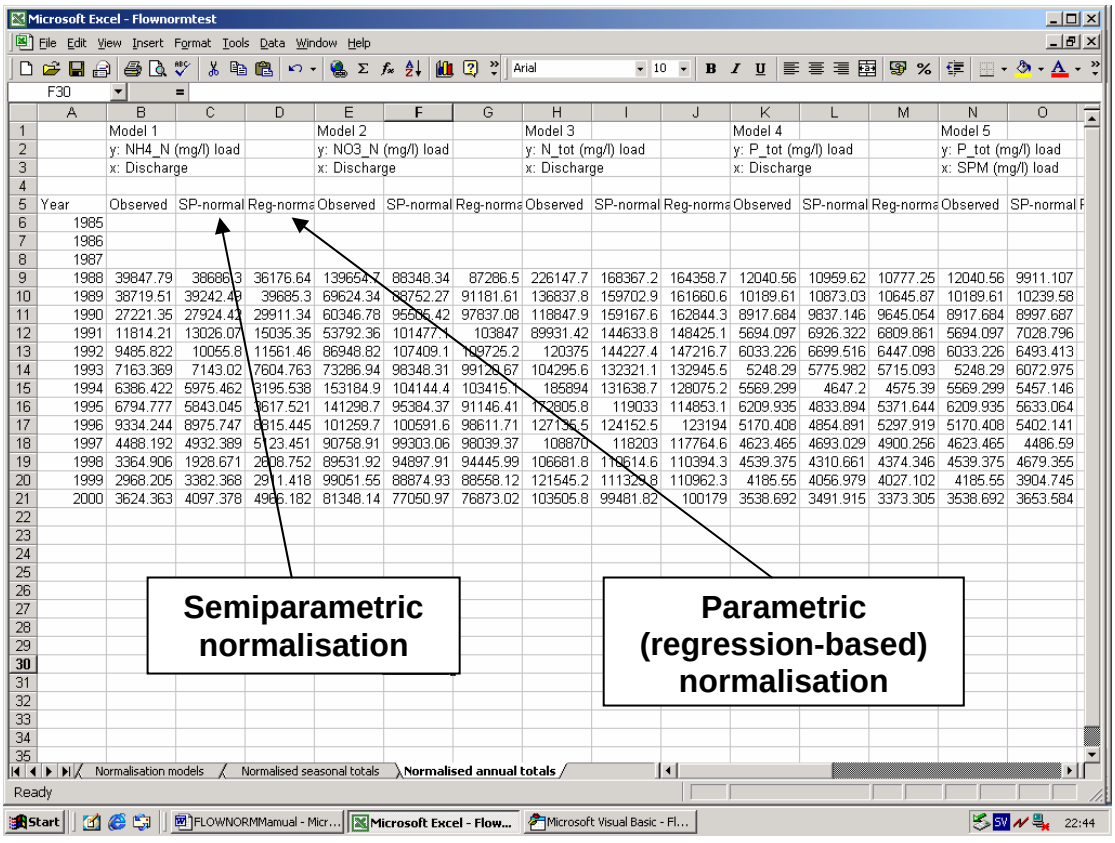


Figure 13. Normalised annual loads

The predictive ability of the different normalisation models can be assessed by computing the mean square prediction error, that shall be as low as possible. The MSPE-values in Figure 14 show that, for the tested data sets, semiparametric models perform better than the simpler parametric models. In addition, the results indicate that the load of total-P at Schnackenburg is most efficiently normalised by using a semiparametric model with load of suspended particulate matter (SPM) as explanatory variable.

The goodness-of-fit is determined by computing the mean square residual. The results obtained for total-P at Schnackenburg (Figure 14) show that the predictive ability of a normalisation model can decrease, even if the goodness-of-fit is improved when a new explanatory variable is included in the model.

The lambda-values in Figure 14 show the values of the smoothing parameters that were selected by cross-validation. To reduce the computational burden, the search for optimal smoothing parameters is restricted to a grid-set of values.

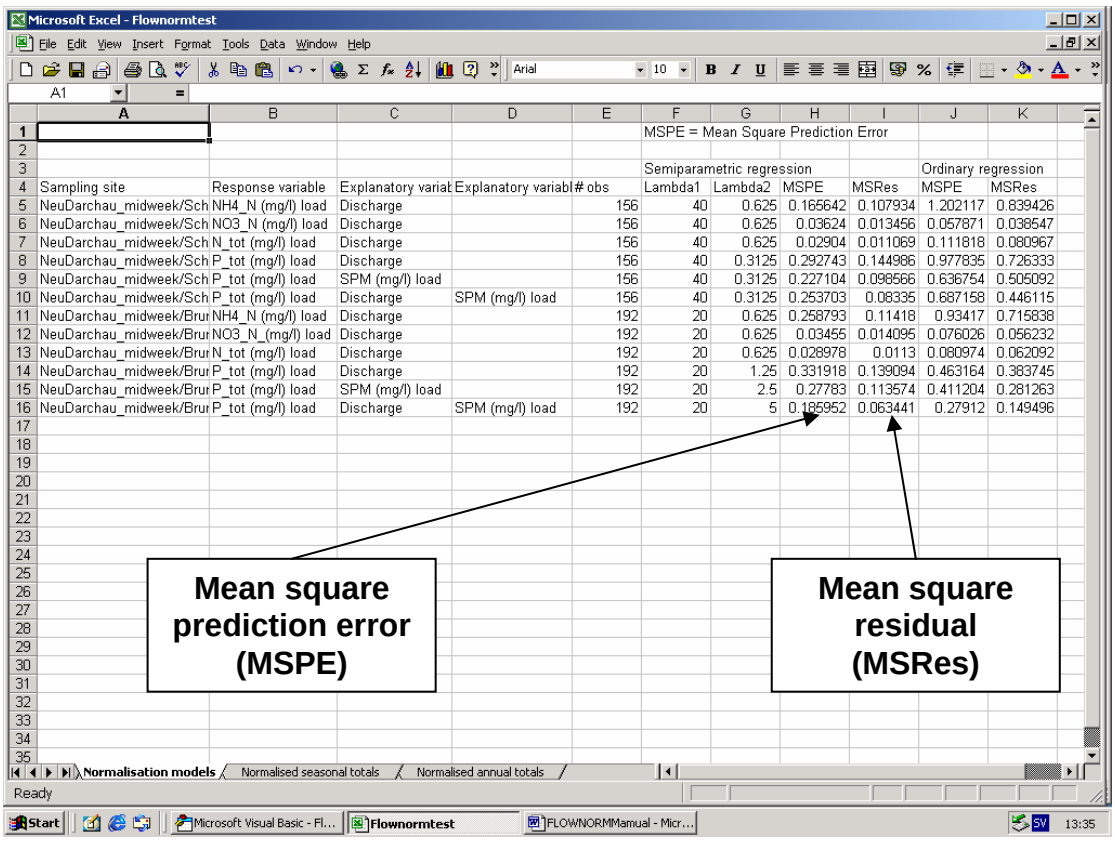


Figure 14. Predictive ability (mean square prediction error) and goodness-of-fit (mean square residual) for the tested normalisation models.