

Papers published in *Hydrology and Earth System Sciences Discussions* are under open-access review for the journal *Hydrology and Earth System Sciences*

Statistical distribution of series of 12 monthly concentration samples for environmental classification of rivers

J. Eliasson¹ and T. Thordarson²

¹University of Iceland, Institute of Environmental and Civil Engineering, Iceland

²University of Iceland, Research Station in Hveragerdi, Iceland

Received: 20 July 2007 – Accepted: 26 July 2007 – Published: 7 August 2007

Correspondence to: J. Eliasson (jonase@hi.is)

2561

Abstract

Environmental monitoring and classification of rivers in the northern hemisphere is frequently hampered by lack of infrastructure in the scarcely populated areas of the north. Carefully designed economical methods are important. Analysis of 15 constituents in
5 14 rivers in Iceland show that monthly samples for a period of 1 year are sufficient for classification provided that the correct statistical distribution is known. Normalizing and plotting all the constituents in each river by rank shows systematic deviations from both the normal and lognormal distributions. When the constituents are pooled by river the result is one distribution for each river, all very similar. A new cumulative distribution
10 function (DoC) is formed as the average of these. It has a long tail similar to that of the lognormal distribution but below the 60% quantile, the DoC differs a lot from the lognormal so if it is to be used, an unbiased estimate of the scale and location parameters will in most cases be difficult to obtain if more than 30–40% of the highest points is used. The influence of the DoC on the classification result is very strong when the
15 90% quantile is used for classification, but fades out at the 60% quantile. It is shown that the storage effect in rivers with a lake that holds some weeks flow in storage, can have a great influence on the classification result.

1 Introduction

In Europe the classification systems for the degree of anthropogenic impact on surface
20 water is based on ecological status involving both the biology and the chemistry of the rivers as set forth in the European Union Water Framework Directive (WFD, EU 2000/60/EC).

The climate in Iceland is cold oceanic with cool summers and relatively warm winters. The mean temperature (1961–1990) of the coldest (January) and warmest (July) month
25 is respectively -3.2°C and 10.5°C in Akureyri and -0.5°C and 10.6°C in Reykjavík. The mean annual temperature in North-Iceland (Akureyri) is 3.2°C and the mean annual

2562

precipitation 490 mm. For Southwest Iceland (Reykjavik) these values are 4.3°C and 799 mm. This data and further information may be compiled from the official website of the Icelandic Meteorological Office: <http://andvari.vedur.is/english/>.

Iceland is mostly made up of Quaternary and Tertiary igneous basaltic rock that strongly influences the chemistry of rivers (Stefansson, 2001). Owing to continental drift the youngest rock is in a rift zone which for the most part extends southwest from the northeastern part of Iceland, over the central highlands to the Reykjanes peninsula (Fig. 1). The rift zone is volcanically active and in areas with postglacial lava fields surface water is uncommon owing to high permeability of the bedrock, maybe as high as $1-10^{-2}$ m/s (Eliasson, 1994). Rivers in the outskirts of the Quaternary rock formations are often spring fed and their water very pure (Sigurðsson and Ingimarsson, 1990) but the clastic sediments and hyaloclastites in this zone sometimes have $10^{-3}-10^{-6}$ m/s, which is enough to hold surface streams (Eliasson, 1994).

The concentration of natural dissolved solids is relatively stable in spring fed rivers but fluctuates in glacial and direct run-off rivers which also have higher concentration and lower discharge in winter time. The rivers acquire their natural solutes mostly from precipitation or from leaching from rocks but many rivers in Iceland receive runoff from geothermal aquifers where the concentration is very dependent on temperature. An excellent review of surface water chemistry in Iceland is given in (Stefansson, 2001) and (Stefansson et al., 2001). Generally the concentration of chemical constituents is lower in Icelandic rivers than the world average with the exception of silica (Si), sodium (Na), ammonium (NH_4) and phosphorus (P) (Stefansson et al., 2001).

Only 23 805 km² of Iceland is covered with vegetation (Statistics_Iceland, 2007), mostly the lowlands. The uppermost reaches of most of the rivers studied are in areas with only thin and patchy vegetative cover.

Agriculture is considered to be the most common source of anthropogenic pressure on surface water in Iceland followed by rural settlements and summer cabins that are for the most part sparsely populated. Urbanized areas are usually a relatively small part of drainage areas in Iceland. Both agriculture and urbanized areas are situated

2563

in the lower reaches of the drainage areas. The upper reaches of the drainage areas are generally governed by pristine conditions but used as pasture for sheep in the summertime.

The data in this paper was collected in order to classify the rivers according to water quality and degree of human impact, the basis for the pollution classification of rivers in Iceland, set forth in Regulation 796/1999, will be investigated and the methods used evaluated. Water quality monitoring, compatible with the requirements of the EU WFD, is a problem in a country like Iceland because of the large number of streams and the small population. This is a problem common for the northern regions of Europe. Consequently an economical classification method is important. In the following we describe a classification method designed to meet the requirements of the ruling of the government of Iceland and possible modifications are discussed.

2 Study area and sampling sites

The sampling sites were chosen near the estuaries of the rivers at a place of easy access. The rivers in this study are all in the Southwest and the Northeast regions of Iceland, they are listed and numbered in Table 1. Seven rivers in southwest Iceland and seven in south Iceland were sampled on a monthly basis 12–14 times. In both regions the rivers are located either in the active rift zone or outside it. The combined drainage area of the studied rivers in southwest Iceland is 1310 km² but 9877 km² in north Iceland. The total drainage area of the rivers in both regions is approximately 11% of the total surface area of Iceland. The elevation of the discharge areas extended from near sea level to approximately 900 m over sea level in southwest Iceland and to approximately 2000 m over sea levels in north Iceland.

To achieve an economical method sampling frequency has to be suitably low. Norwegian rules (Norsk_Institutt_for_vannforskning, 1997) state that 1 monthly sample is minimum for big rivers, but twice a year for small rivers in pristine conditions. In this investigation it was decided to use the monthly sampling frequency until some experi-

2564

ence is gained.

3 Statistical investigation

3.1 Objectives

The main objective of the statistical investigation is to find the distribution of the constituents and compare it to the normal and lognormal distributions. Then evaluate the information we get about the average properties in Table 1, and use them in classification. In order to do so, all the constituents were normalized by subtracting the 12 months average and dividing by the standard deviation (Eliasson, 1997). By this we create a statistic with zero mean and standard deviation equal to one, for all the constituents in each river. Now there is one number for each month and each constituent in every river. The months can be ranked and the distributions can be plotted and compared. If the constituents have a common two-parameter distribution it will conveniently show in graphical display.

3.2 Normal distribution

Using the fact that the constituents do not have a strong seasonal correlation they can be compared to the normal, or Gaussian, distribution. There are only 12 numbers for each constituent in each river so the distributions plots as in Fig. 2. It shows the constituents from a river Laxa in Thingeyjasysla, no. 3 in Table 1. The river originates from the spring fed lake Myvatn, which is an eutrophic lake with high production, rich bird life and special ecology. This picture can be compared to the smaller pictures in Fig. 3.

Figure 2 is so produced that a smoothed line is drawn through all 12 values of the ranked constituents so the distribution they indicate can be compared to a red line of the normal probabilities of the same rank. The selected smoothing method exaggerates the differences between the ranks but at the same time makes it easier to identify the

2565

actual line and follow it through all 12 ranks. Three reasons that cause the distributions to deviate from the red line for normality can be identified. Practically all rank 1 values are above the red line, most rank 4–10 values are below it and rank 11 and 12 points are above it. This is a systematic error which can be seen on Fig. 3 for 8 more rivers, but statistical tests do not reject the normal distribution for most of these series. The six rivers not plotted in Fig. 3 contain no new information.

Practically the same features as above cause the logarithms of the values for the concentrations of the constituents to deviate from the normal distribution for all the streams and rivers as can be seen in the example in Fig. 4. All values on the vertical axis are in log space.

The lognormal distribution is very popular in environmental studies and many times used without any discussion about its applicability. This is not unjustified when extreme values are concerned as will be demonstrated later.

3.3 Non-parametric approach

The constituents show a clear deviation from the normal and lognormal distributions, but the deviation is not the same from river to river. Take as an example the TC line for no. 13 Fossá. It shows high values in the high and low ranks, but is relatively flat in between. In the absence of more information an observational error in the high and low values would be suspected. But because the other rivers show similar high rank and low rank values without the flat middle we know that this is not necessarily the case. Besides, the TC line is the sum of the TOC and IC lines and they do not behave in this manner as can e.g. be seen in Fig. 2.

The bold lines in Fig. 5 show the variations of the normalized data when the fourth highest (high), the median (med) and the fourth lowest (low) of all constituents of same rank in a river are plotted. The bold lines thus show the 21%, 50% and 79% average quantiles of constituents of all 14 rivers and also the standard deviations, all 14 rivers considered. This is used to produce the lines marked plus and minus which indicate one standard deviation around the respective quantile. According to the central limit

2566

theorem the distribution around the bold lines in Fig. 5 should be close to the normal distribution, as each value on that line is an average for 14 rivers. Then 68% of the probability mass should be within the slim lines of same color around the bold high, med and low lines. All the data can be plotted as points on Fig. 5 and 78% of the points would fall within the bold lines. Figure 5 thus shows that the variance between constituents (the difference between the bold lines) and the variation between rivers (difference between the slim lines and the same color bold line) is small compared to the variation within the constituents themselves (note the difference between ranks on the vertical scale). Having only 12 points for each river in each year, this strongly indicates, that the best thing to do, is to use the same distribution for all the constituents and all rivers, and in this case it is not the lognormal distribution as widely suggested (see discussion later).

3.4 The constituent distribution

If a two-parameter distribution function exists for the constituents in each river, we can now find it by pooling the results behind Fig. 5 by river (Eliasson, 1997). The result is on Fig. 6. It shows the empirical distribution of each river found by pooling 11 constituents. Of the original 15 the following are not included: pH, Conductivity and Fecal coliforms due to significant correlation to air temperature in some rivers, and TC as it is a sum of two other parameters. However, it does not make any significant change to exclude them. The resulting distributions for the rivers are very much alike so DoC (Distribution of Constituents) is now found as the average of all rivers (bold red line).

3.5 The probability density function of DoC

Discrepancies in empirical distribution functions do show better in probability density functions than in cumulative distribution functions. The DoC probability density function is shown in Fig. 7.

Figure 7 shows that even though the DoC looks nice on Fig. 6 it is by no means

2567

flawless. It is compared to the lognormal distribution we would get by sliding the distribution on the horizontal axis 2.3 to the right, probability densities calculated and slid back. This is necessary in order to avoid negative values and must be done as negative concentrations do not occur in practice.

3.6 The probability factor of the DoC

The DoC is a cumulative probability distribution function, Φ_{DoC} , stating:

$$P\left(\frac{X - \bar{X}}{s_x} < x\right) = \Phi_{\text{DoC}}(x) \quad (1)$$

or

$$\frac{X - \bar{X}}{s_x} = P^{-1}(\Phi_{\text{DoC}}(x)) \Rightarrow X = \bar{X} + K s_x \quad (2)$$

P = probability

K = probability factor defined in Eq. (2)

s_x = standard deviation

X = observation value

\bar{X} = average value

K can be taken from Table 2, compiled from the data behind Fig. 6

4 Concentration classes using the DoC

4.1 Concentration limits

Having established the parameters of the distribution to be expected (Table 2) it is possible to classify the rivers in Table 1 according to the ruling of Regulation no. 796/1999

(The_Environment_and_Food_Agency_of_Iceland, 1999). In this regulation a 90% concentration class limit is prescribed, so the 90% K in Table 2. has to be used.

$$X(90\%) = \bar{X} + 1.48s_x = \bar{X}(1 + 1.48C_v) \quad (3)$$

The coefficient of variation C_v , is to be taken from Table 1 or a similar source.

- 5 Similarly, if a mean value is to be within its class on the 90% safety level, still taking only one measurement in each month, Eq. (1) must be inverted.

$$\bar{X} \leq X(90\%)/(1 + 1.48C_v) \quad (4)$$

This would be the wanted mean value to keep in order to ensure that 90% of the observations will be within the class limits when $X(90\%)$ is the higher limit for the class.

- 10 A different interpretation of Regulation no. 796/1999 is possible. That is to say that for monthly observations one year is not enough basis for classification. This is a view shared by many researchers studying censored data (Johnes et al., 1996) and (Johnes, 2007) recommends daily samples, (Guo et al., 2002) investigate nitrate-N load estimates when monitoring durations were 1, 2, 3, and 6 years, and the sampling frequencies ranged from weekly to bimonthly. They base their estimates on rating curve, ratio estimator and flow-weighted average estimator for each sampling scenario, and then use two bias correction techniques, minimum variance unbiased estimator (MVUE) and smearing estimator. They state that a desired accuracy of the estimates could be achieved either by sampling more frequently or by monitoring the site longer. This result then means that a monthly sampling frequency can be made accurate enough by using an observation period long enough, but does not really say that one year is sufficient. It is thus possible to increase the classification accuracy of monthly samples by measuring more than one year. If the average concentration of each year is now called one observation, the 90% ruling can be used on these. Doing so, two facts have to be remembered. First, instead of the C_v value in Table 1, a new coefficient of variation with the expected value $C_v^* = C_v/\sqrt{12}$ emerges. Secondly the DoC will not be valid any more, if we add up a distribution 12 times, the central limit

2569

theorem will have turned up the Normal (Gaussian) distribution, irrespective of what the mother distribution is. Now the formula changes to,

$$\bar{X} = X(90\%)/(1 + 1.16C_v^*) \quad (5)$$

- Equation (5) produces a considerably higher value than Eq. (4) if the same $X(90)$ is used .

4.2 Using the 90% rule or averages for the class definition

- As the Figs. 2–6 clearly show, statistical significance of the 90% limit is very low. In Fig. 2, 90% correspond to the rank 2 value, this fluctuates from 0,5 to more than 2, or over 1.5 standard deviations. Around the mean value there is much greater stability, so if we base the classification on 12 monthly values only, to make use of the mean value and the DoC to estimate the 90% value, is a better choice than using the rank 2 values. In this way Table 1 can easily be turned into a class determining table by using Eq. (1) to calculate the 90% values that must not be exceeded in a series of 12 monthly values and compare them to the class limits. All together there are five classes, and the classes for each constituent in the 14 rivers is shown under the 90% heading in Table 3.

- 15 In Table 3 under the 60% heading we also see, the resulting classes if the rivers are classified according to mean value. C_{max} is the highest class, signifying that the environmental state of the river as a whole is this class or better, all constituents considered. In Table 3 there are also the classes that would result from a 60% rule or classification by average.

5 A note on the current practice of classification

As mentioned above reg. 796/199 stipulates that the environmental classification of surface waters is to be done at the 90% significance level. Due to lack of data, the

2570

rivers classified so far in Iceland have been classified based on average values from 10–14 samples over approximately one year, in order to achieve realistic classification within manageable financial restraints, this corresponds to the 60% significance level according to the DoC. In order to achieve the 90% confidence interval using only the mean or the geometric mean, and not having established the validity of the DoC, a much larger number of samples would be needed. This is discussed in more detail later.

6 Discussion

The DoC gives the best fit to the data, it can be used in monitoring involving monthly samples, but the studied data is too small to recommend it to be used for other data sets. It is still a question when the DoC has to be used or if the logN may be used instead, Fig. 7 shows that the high tail of DoC follows the logN tail very closely, so investigators only concerned about the high values can use logN and achieve the same result. The problem is the estimation of the scale parameter and the location parameter, whose estimates will be biased because how the DoC deviates from the lognormal everywhere below the average. In this range the probability weights of the lognormal differ from the DoC by a factor of 2 to either side. This may be helped by taking the parameters from the tail where the distributions are almost identical and get an unbiased estimate that way. E.g. the significant value (average of the highest third of the data) may be used. Other parameters produced by maximum likelihood estimates using only the highest one third of the data may also be used.

This long tail produces many problems in classification. The most outstanding is the significance of the storage effect. When a river runs through a lake containing a few weeks run-off in storage then the concentrations in the outflow vary much less than in the inflow. This may be seen in Table 3. Elliðaa (No. 4) is the outflowing river from a lake, Elliðavatn, that is a hydropower storage reservoir. Elliðaa is classified in a better class than the inflowing rivers, Holmsa and Sudura when the 90% rule is used.

2571

Changing to a 60% rule makes inflow and outflow classes equal. This difference is caused by the heavy weight of the term $1,48 C_v$ in Eqs. (1) and (2) and the fact that the C_v values are greater for the inflowing rivers because of the storage effect.

7 Conclusions

The normalized concentrations of the constituents in each river investigated define a distribution function when pooled together. It can be any distribution, not necessarily the lognormal. The pooling presumes that the constituents do belong to a common two-parameter distribution, a presumption already in use in most environmental monitoring practice.

Averaging the distributions of 14 rivers produces the DoC with surprising accuracy as the 14 distributions are very similar, using the DoC is within 0.2 everywhere in Fig. 6, so class determination will be within 0.2σ .

When high concentration values are the only concern, the lognormal distribution produces the same result as the DoC. But parameters have to be estimated with unbiased estimates using 40%, or less, of the highest values.

C_v 's can be rather high so the 90% rule of the Regulation no. 796-1999 can create a large fluctuation in class determination due to high influence of the C_v and the unavoidable fluctuations of this parameter when classification is based on only 12 monthly samples. This classification problem can be circumvented by using C_v values averaged over 6–12 years and Eq. (3), thus avoiding the influence of the long tail of the DoC.

Variance reducing storage effects can lead to the unexpected result that rivers in pristine condition can fall in a worse class than the same river in lowland areas under higher environmental pressure. Classification according to 60% rule instead of 90% avoids this flaw.

References

- Eliasson, J.: Northern Hydrology in Iceland, (in Northern Hydrology International Perspectives, edited by: Prowse, T. D., Omanney, C. S. L., and Watson, L. E.), NHRI Science Report No. 3, 41–79, Fig. 4.9, 1994.
- 5 Eliasson, J.: A statistical model for extreme precipitation, *Water Resour. Res.*, 33(3), 449–455, 1997.
- Guo, Y., Markus, M., and Demissie, M.: Uncertainty of nitrate-N load computations for agricultural watersheds, *Water Resour. Res.*, 38(10), 1185–1199, 2002.
- ISO: ISO 9308-1:2000 Water quality – Detection and enumeration of *Escherichia coli* and coliform bacteria – Part 1: Membrane filtration method, 2000.
- 10 Johnes, P., Moss, B., and Phillips, G.: The determination of total nitrogen and total phosphorus concentrations in freshwaters from land use, stock headage and population data: Testing of a model for use in conservation and water quality management, *Freshwater Biol.*, 36(2), 451–473, 1996.
- 15 Johnes, P. J.: Uncertainties in annual riverine phosphorus load estimation: Impact of load estimation methodology, sampling frequency, baseflow index and catchment population density, *J. Hydrol.*, 332(1–2), 241–258, 2007.
- Norsk_Institutt_for_vannforskning: Classification of Environmental Quality in freshwater (Klassifisering av miljøkvalitet i ferskvann; in Norwegian), Veiledning 97:04, 1997.
- 20 Sigurðsson, F. and Ingimarsson, J.: Lekt íslenskra jarðefna, in: *Vatnið og landið*, edited by: Sigbjarnarson, G., Orkustofnun, Reykjavík, 121–128, 1990.
- State of New York Department of Health: Environmental laboratory approval program certification manual, Item no. 198.3, Procedural Modification for Drinking Water Total Coliform, 3 pp., 2005.
- 25 Statistics_Iceland: Geography and environment, <http://www.statice.is/>, 2007.
- Stefansson, A.: Dissolution of primary minerals of basalt in natural waters – I. Calculation of mineral solubilities from 0 degrees C to 350 degrees C, *Chem. Geol.*, 172(3–4), 225–250, 2001.
- Stefansson, A., Gislason, S. R., and Arnorsson, S.: Dissolution of primary minerals in natural waters – II. Mineral saturation state, *Chem. Geol.*, 172(3–4), 251–276, 2001.
- 30

The_Environment_and_Food_Agency_of_Iceland:

<http://www.reglugerd.is/interpro/dkm/WebGuard.nsf/b7fd33650490f8cf00256a07003476bb/8740f17f27ea86cd00256a0800330269?OpenDocument> (in Icelandic), 1999.

Table 1. Average physiochemical properties of the rivers studied (Ave), coefficient of variation (C_{var}) and correlation coefficient (r).

No		Air temp °C	Water temp °C	pH	Conductivity uS/cm	Fec. colif. 100 ml	t-P (mg/l)	PO ₄ -P (mg/l)	t-N (mg/l)	NH ₄ -N (mg/l)	TOC (mg/l)	IC (mg/l)	TC (mg/l)	Cu (ug/l)	Zn (ug/l)	Pb (ug/l)	Cr (ug/l)	Ni (ug/l)
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
Fñjóska	1 Ave	6.08	3.43	7.57	52.64	4.50	0.02	0.01	0.21	0.01	1.45	4.00	5.45	0.61	2.00	0.03	1.13	0.26
	C_{var}	0.80	0.93	0.02	0.10	0.71	0.43	0.32	0.21	0.23	0.24	0.26	0.25	1.46	0.20	0.99	0.25	1.10
	Correl	1.00	0.84	0.59	-0.39	-0.16	-0.28	-0.19	-0.31	-0.24	0.13	0.33	0.28	0.12	-0.32	0.14	0.22	0.07
Skjalfanda tjót	2 Ave	5.58	4.16	7.55	97.70	13.17	0.03	0.02	0.23	0.01	1.84	7.12	8.96	1.58	2.83	0.05	1.56	0.76
	C_{var}	0.91	1.00	0.03	0.12	1.59	0.38	0.33	0.46	0.81	0.14	0.23	0.19	1.57	0.58	1.20	0.76	1.39
	Correl	1.00	0.84	0.47	-0.20	0.11	-0.21	-0.30	-0.52	-0.56	-0.02	-0.05	-0.05	0.06	0.10	0.33	0.01	0.06
Laxá	3 Ave	5.05	4.63	8.21	142.45	20.42	0.03	0.02	0.39	0.02	3.27	10.77	14.04	0.69	2.09	0.02	1.42	0.39
	C_{var}	1.10	1.02	0.07	0.17	1.43	0.31	0.26	0.31	0.80	0.27	0.27	0.26	0.54	0.22	1.01	0.24	0.57
	Correl	1.00	0.85	0.65	-0.15	-0.06	-0.55	-0.15	0.37	-0.54	0.58	0.15	0.26	-0.32	-0.09	0.13	-0.32	-0.21
Elldaár	4 Ave	6.73	6.45	8.46	90.42	12.67	0.01	0.01	0.10	0.01	1.68	4.16	5.83	0.45	0.84	0.03	0.96	0.27
	C_{var}	0.98	0.88	0.09	0.05	1.08	0.37	0.10	0.38	0.59	0.30	0.10	0.13	0.41	0.49	0.61	0.16	0.26
	Correl	1.00	0.90	0.80	-0.12	0.04	-0.26	0.04	0.55	0.36	0.76	0.20	0.60	-0.07	-0.24	0.18	-0.44	-0.21
Holmsá	5 Ave	6.09	5.16	8.12	88.08	36.67	0.01	0.01	0.10	0.01	1.28	4.11	5.39	0.32	0.77	0.02	0.98	0.29
	C_{var}	1.15	0.90	0.05	0.07	1.50	0.40	0.22	0.48	0.39	0.17	0.11	0.09	0.56	0.59	0.68	0.11	0.39
	Correl	1.00	0.93	0.81	-0.47	0.50	-0.28	-0.04	-0.28	-0.31	0.48	0.25	0.44	-0.30	-0.21	-0.41	-0.38	-0.23
Sudurá	6 Ave	6.01	5.02	8.12	90.17	18.00	0.01	0.01	0.44	0.01	1.13	3.65	4.78	0.41	0.87	0.02	0.95	0.31
	C_{var}	1.24	0.72	0.05	0.06	2.27	0.36	0.19	2.67	0.26	0.24	0.30	0.26	0.52	0.52	0.66	0.12	0.45
	Correl	1.00	0.95	0.55	-0.45	0.32	-0.26	-0.43	0.41	-0.03	-0.31	-0.46	-0.47	0.08	-0.35	-0.29	-0.10	0.03
Glerá	7 Ave	6.36	3.21	7.69	64.75	135.17	0.03	0.01	0.15	0.04	1.25	2.87	4.12	2.36	3.18	0.13	0.93	0.65
	C_{var}	1.01	0.57	0.04	0.31	1.50	0.59	0.39	0.84	0.77	0.39	0.33	0.33	0.91	0.81	1.02	0.53	0.80
	Correl	1.00	0.82	0.29	-0.27	-0.06	0.27	-0.25	-0.25	-0.23	-0.16	-0.09	-0.12	0.43	0.33	0.38	0.30	0.41
Svarfaðar dalsá	8 Ave	4.71	3.12	7.51	61.75	103.83	0.03	0.00	0.23	0.01	1.49	3.41	4.90	1.73	3.38	0.09	0.31	0.62
	C_{var}	1.07	0.84	0.05	0.16	2.61	0.71	0.76	0.83	0.57	0.38	0.22	0.23	0.70	1.33	0.95	0.37	1.40
	Correl	1.00	0.64	0.50	-0.21	0.28	0.33	0.03	0.30	-0.66	0.38	0.06	0.23	0.39	0.14	0.22	0.37	0.13
Hörgá	9 Ave	6.14	3.41	7.61	52.58	22.92	0.02	0.01	0.25	0.01	0.98	3.36	4.34	1.17	2.28	0.03	0.23	0.33
	C_{var}	0.99	0.76	0.04	0.23	1.73	0.45	0.28	1.71	0.96	0.22	0.34	0.30	0.62	0.79	0.84	0.36	0.41
	Correl	1.00	0.65	0.33	-0.46	0.20	-0.06	-0.43	0.34	-0.23	-0.13	-0.45	-0.41	0.23	0.29	0.00	0.07	-0.20
Eyja fjarðará	10 Ave	6.40	4.38	7.57	65.45	37.75	0.02	0.01	0.14	0.01	1.51	3.65	5.16	1.28	1.47	0.02	0.50	0.30
	C_{var}	1.05	0.74	0.03	0.11	0.87	0.42	0.29	0.52	0.85	0.25	0.28	0.23	1.13	0.67	0.78	0.40	0.51
	Correl	1.00	0.74	0.59	-0.14	-0.11	0.01	-0.57	-0.18	-0.75	-0.27	-0.18	-0.24	-0.23	0.08	0.22	0.20	0.05
Botnsá	11 Ave	4.25	4.28	8.34	58.42	32.50	0.01	0.01	0.55	0.01	1.04	2.86	3.89	0.26	84.01	0.04	0.73	0.37
	C_{var}	1.31	0.87	0.10	0.11	1.58	0.32	0.17	2.99	0.41	0.44	0.35	0.33	0.49	2.57	0.87	0.48	0.57
	Correl	1.00	0.95	0.20	-0.41	0.31	0.10	0.06	-0.13	-0.31	-0.17	0.05	-0.02	-0.20	0.41	-0.67	0.11	0.58
Brynju dalsá	12 Ave	4.33	4.28	7.97	54.42	9.42	0.01	0.01	0.09	0.01	0.94	2.58	3.51	0.65	44.48	0.07	0.79	0.40
	C_{var}	1.36	0.89	0.05	0.16	1.15	0.75	0.17	0.87	0.40	0.39	0.19	0.18	1.13	1.57	0.92	1.11	0.74
	Correl	1.00	0.98	0.06	-0.62	0.55	0.05	0.14	-0.02	-0.49	-0.43	-0.44	-0.62	-0.22	-0.33	-0.36	0.09	-0.19
Fossá	13 Ave	4.50	4.10	7.76	59.00	2.50	0.01	0.01	0.08	0.01	0.99	2.71	3.69	0.31	98.42	0.04	0.57	0.32
	C_{var}	1.28	0.97	0.04	0.17	1.11	0.28	0.23	1.33	0.53	0.48	0.43	0.40	0.75	1.38	0.74	0.68	0.62
	Correl	1.00	0.97	-0.22	-0.50	0.42	0.33	-0.01	-0.31	-0.13	-0.14	0.25	0.15	-0.15	-0.21	-0.43	-0.42	0.01
Kibafellsá	14 Ave	4.75	4.48	7.79	92.83	30.42	0.02	0.01	0.35	0.01	1.94	5.56	7.34	0.58	157.78	0.04	0.67	0.71
	C_{var}	1.19	0.83	0.02	0.18	1.91	0.25	0.40	1.47	0.84	0.45	0.36	0.34	0.78	1.62	0.42	0.55	0.49
	Correl	1.00	0.97	0.06	-0.24	0.11	-0.39	-0.18	-0.12	-0.24	-0.08	-0.14	-0.14	0.15	-0.09	-0.39	-0.42	0.12

2575

Table 2. K values for DoC.

$P\%$	5	10	25	50	60	75	90	95	97.5	99
K	-1.25	-0.96	-0.64	0.27	0	0.51	1.48	2.17	2.65	2.93

2576

Table 3. Determined classes for the rivers in Table 1 by 90% values and averages (60%).

	Region	Fec. colif.		t-P		PO4-P		t-N		TOC		Cu		Zn		Pb		Cr		Ni		C _{max}	
		90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%
Confidence level		90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%	90%	60%
Fniðská	N	1	2	2	1	3	2	2	1	2	2	2	2	1	1	1	1	2	2	1	1	3	2
Skjalafanda	N	2	1	3	2	3	2	2	1	2	2	3	2	2	1	1	2	2	2	2	2	3	2
Laxá	N	2	2	3	2	3	3	2	2	3	3	2	2	1	1	1	1	2	2	2	1	3	3
Elldas E1	S	2	1	1	1	2	1	1	1	2	2	2	2	1	1	1	1	2	2	1	1	2	2
Holmsá	S	3	2	1	1	1	1	1	1	2	1	2	2	1	1	1	1	2	2	1	1	3	2
Sudurá	S	2	2	1	1	2	1	4	1	2	1	2	2	1	1	1	1	2	2	1	1	4	2
Glerá	N	4	3	3	2	2	1	2	1	2	1	3	2	2	1	2	1	2	2	2	1	4	2
Svarfabardalsá	N	4	3	3	2	2	1	2	1	2	1	3	2	2	1	2	1	2	2	2	1	4	3
Hörgá	N	2	1	3	2	2	1	3	1	2	1	2	2	2	1	1	1	2	1	1	1	3	2
Eyjafjarðará	N	2	2	2	1	2	1	1	1	2	2	3	2	1	1	1	1	2	2	1	1	3	2
Bolnsá	S	2	2	1	1	1	1	4	1	2	1	2	1	5	2	1	1	2	2	2	1	5	2
Brynjufalsá	S	2	1	2	1	2	1	1	1	2	1	2	2	4	3	1	1	2	2	2	1	4	3
Fossá	S	1	1	2	1	2	1	1	1	2	1	2	1	4	4	1	1	2	2	1	1	4	4
Kibafellsá	S	3	2	2	1	3	2	3	1	3	2	2	2	5	4	1	1	2	2	2	2	5	4
Mean																						3.6	2.4

2577

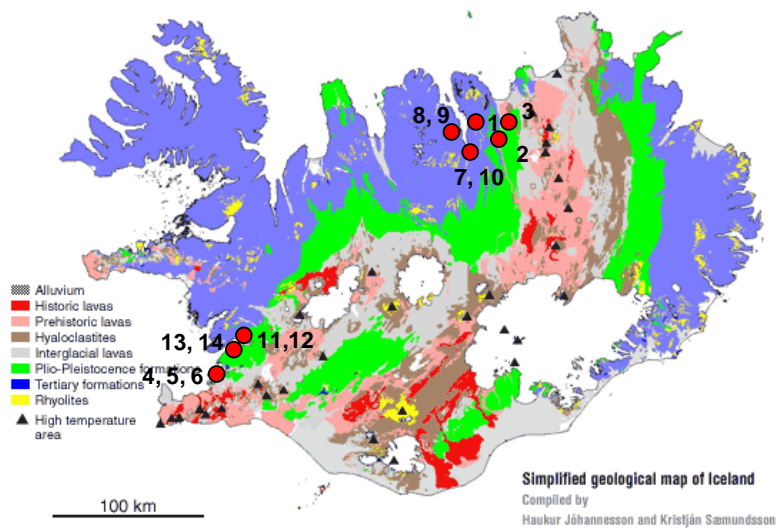


Fig. 1. Approximate location of the 14 rivers plotted on a geological map of Iceland. No. 4, 5 and 6 are in the capital area, the others in scarcely populated areas with some farming, cattle gassing and summer cabins in the catchments' area.

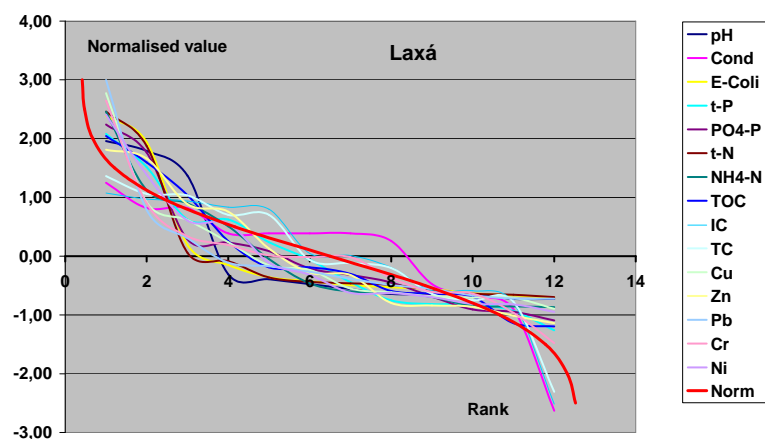


Fig. 2. Distributions of the ranked constituents in Laxá. The deviations of Conductivity and pH are unusually large. The thick line is the normal distribution

2579

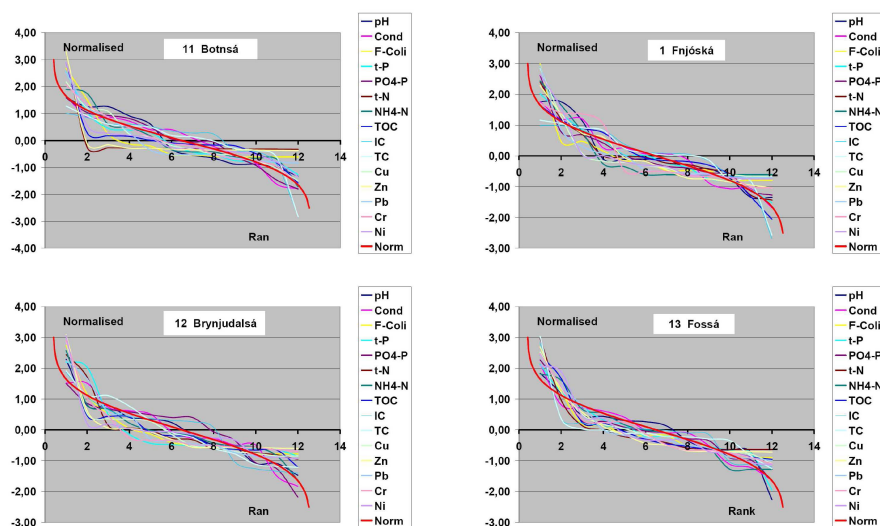


Fig. 3. Plots of the normalized constituents in 8 selected rivers compared to the normal distribution (red line). Note the rank (probability measure) is on the horizontal axis opposite to tradition in statistics. Each constituent has the same color in all 8 figures.

2580

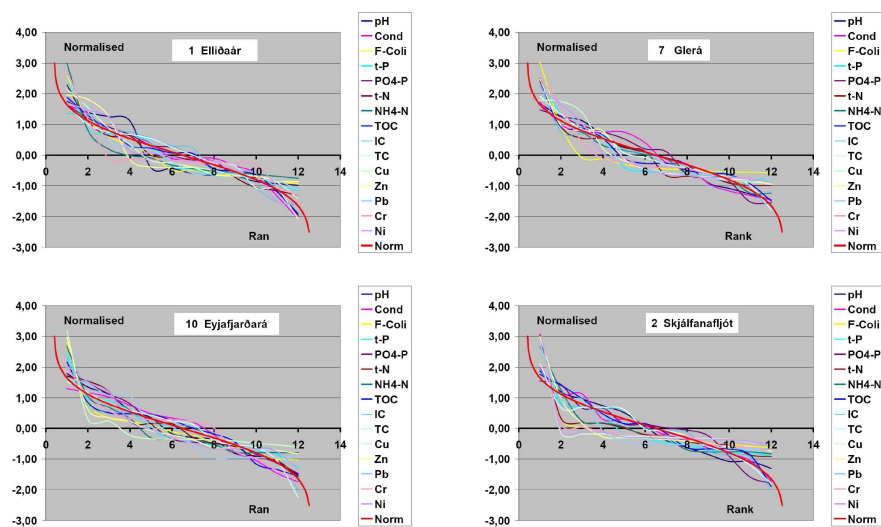


Fig. 3. Continued.

2581

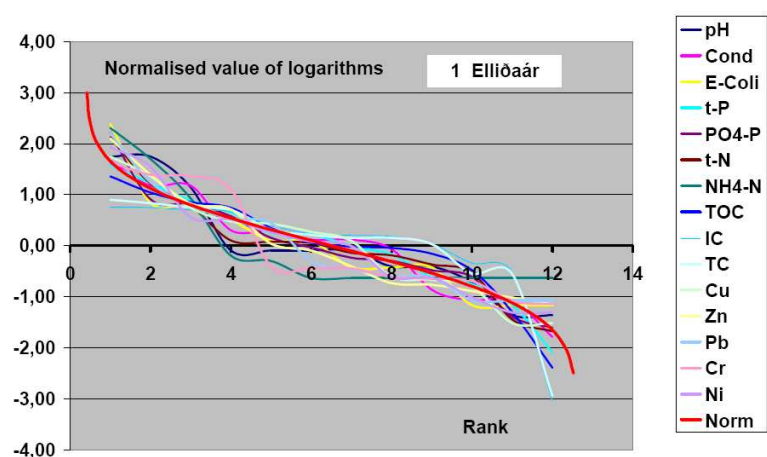


Fig. 4. Plots of the normalized constituents of Elliðaár river compared to the lognormal distribution (red line). Note that the rank (probability measure) is on the horizontal axis opposite to tradition in statistics.

2582

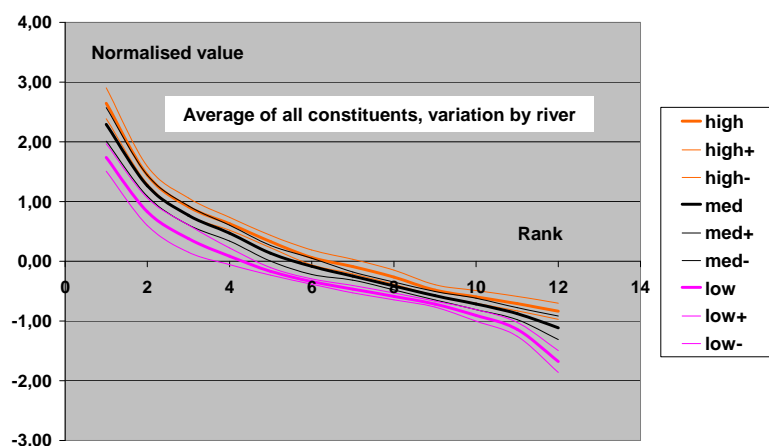


Fig. 5. A plot of constituents of all ranks in the rivers. The bold lines show the average for all rivers of the fourth highest (high), the median (med) and the fourth lowest (low) of all the constituents in the respective river. The plus and minus lines indicate one standard deviation around that estimate.

2583

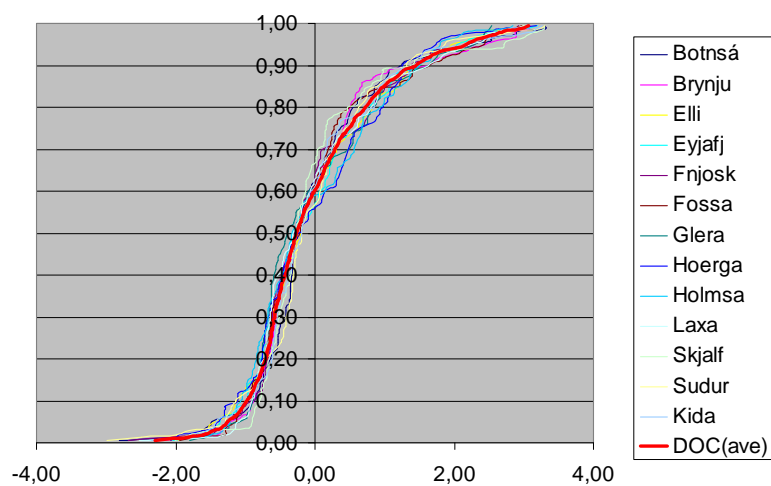


Fig. 6. Distribution of constituents (DoC) found as the average value of all the empirical distribution found by pooling each individual river.

2584

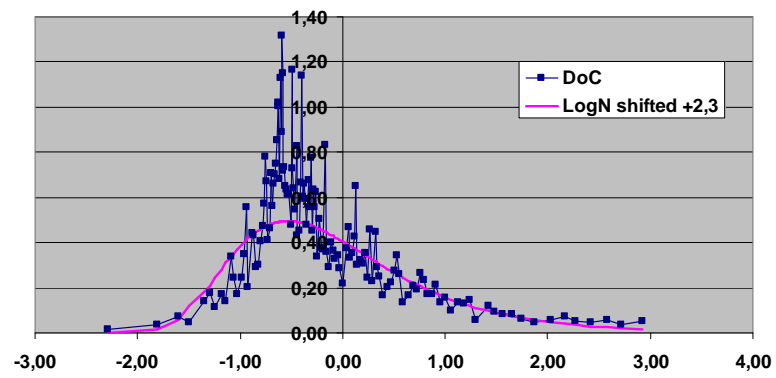


Fig. 7. Probability density function of DoC.