ADDITIONAL VIEW IN THE MOSS TECHNIQUE BY EMPLOYING SOME SIMPLE STATISTICS

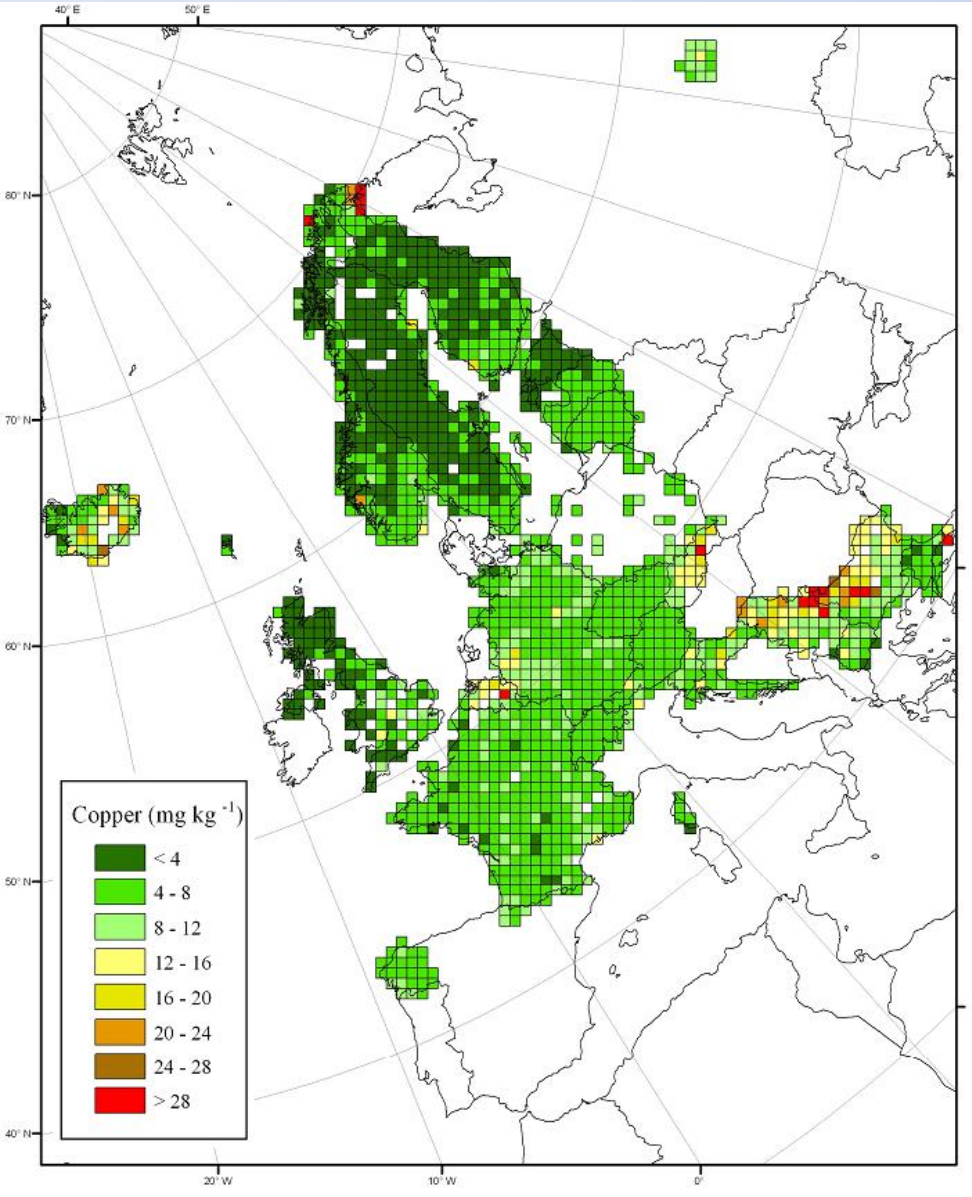
Miodrag Krmar[1], Ilija Arsenić[2], Dragan Radnović[3]

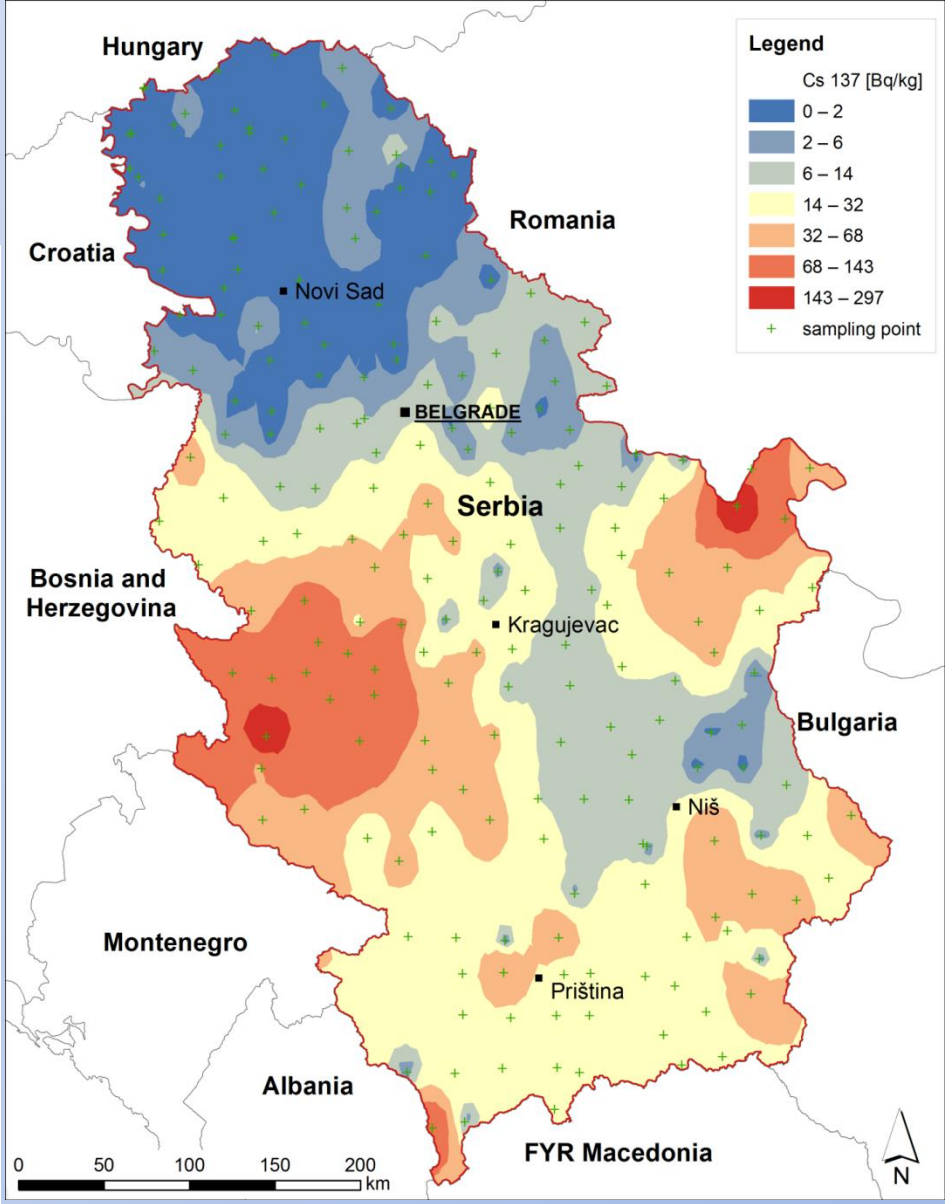
[1] Physics Department, University Novi Sad, Serbia

[2] Faculty of Agriculture, University Novi Sad, Serbia

[3]Department of Biology and Ecology, University Novi Sad, Serbia

Excellent maps shoving spatial distribution of some heavy metals or radionuclide can be obtained using biomonitors (mosses).





Draft summary data European moss survey 2005/6 (18-02-08)

Country	Parameter	As (mg/kg)	Cd (mg/kg)	Cr (mg/kg)	Cu (mg/kg)	Fe (mg/kg)	Hg (mg/kg)	Ni (mg/kg)	Pb (mg/kg)	V (mg/kg)	Zn (mg/kg)	Al (mg/kg)	Sb (mg/kg)	N (%)
Poland	Number		272		273	273		272	272		272			
Poland	Min		0.01		3.34	98		0.13	0.01		18.4			
Poland	Max		1.94		13.10	1210		4.63	17.49		124.2			
Poland	Median		0.25		6.58	300		1.64	4.17		34.1			
Poland	Mean		0.33		6.88	354		1.65	5.11		37.8			
Poland (Marina)	Number	30		30		30		30		30	30	30	30	
Poland (Marina)	Min	0.30		1.00		112		0.83		1.07	26.5	490	0.13	
Poland (Marina)	Max	3.12		10.29		3086		6.36		11.66	125.2	7406	0.68	
Poland (Marina)	Median	0.90		2.74		775		2.56		2.61	64.3	1237	0.36	
Poland (Marina)	Mean	1.03		3.42		868		2.57		2.93	66.8	1503	0.36	
Russia (all)	Number	220	74	220	76	220		220		220	220	220	220	
Russia (all)	Min	0.02	0.03	0.21	3.41	26		0.43		0.08	11.4	79	0.02	
Russia (all)	Max	2.47	1.06	48.06	22.48	23490		22.78		68.51	331.1	12865	2.47	
Russia (all)	Median	0.23	0.24	3.64	8.94	679		2.74		2.27	40.1	850	0.23	
Russia (all)	Mean	0.28	0.30	5.12	9.74	1212		3.53		3.64	47.7	1296	0.28	
Serbia	Number	193	193	193	193	193		193	193	193	193	193	193	
Serbia	Min	0.22	0.04	2.00	3.04	670		1.70	1.03	1.94	13.2	1117	0.06	
Serbia	Max	21.60	1.11	78.81	451.04	16100		23.83	248.64	32.67	258.6	31180	1.37	
Serbia	Median	1.41	0.26	6.44	11.11	2267		4.43	16.73	5.76	29.0	3946	0.24	
Serbia	Mean	2.01	0.34	9.69	17.69	3174		5.38	20.74	7.48	34.0	5572	0.29	

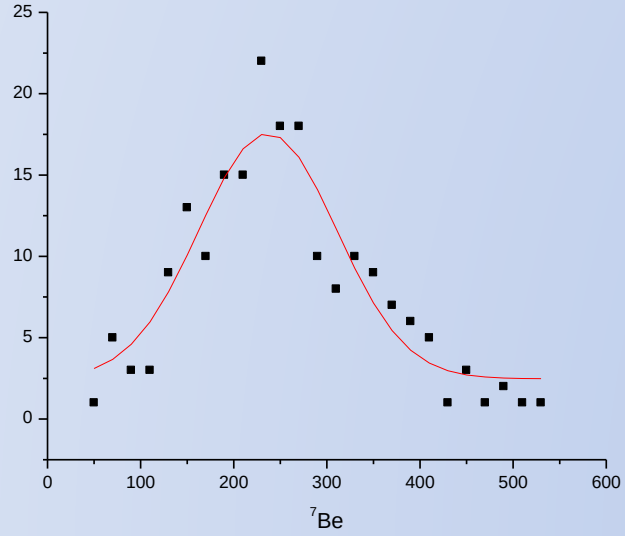
Several statistical parameters as Minimal value, Maximal value, Mean and Median are usually given.

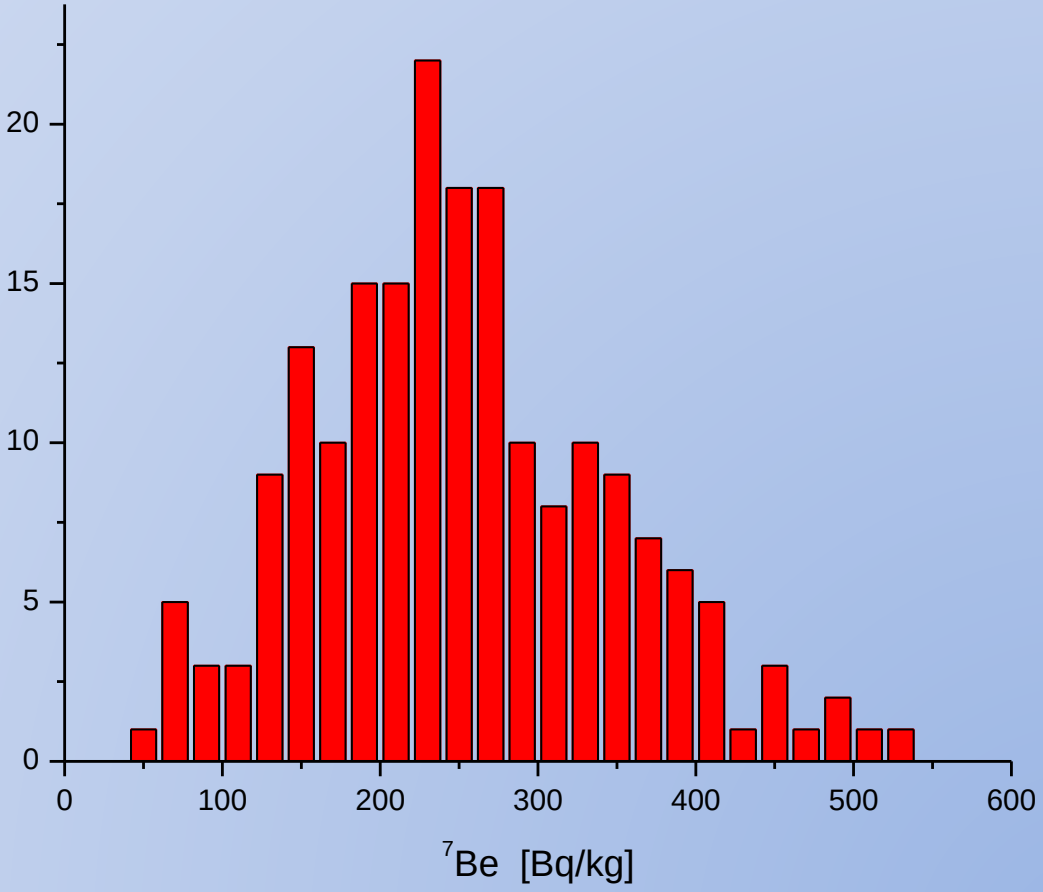
Mean is always higher than Median.
It is consequence of specific distribution of measured values.

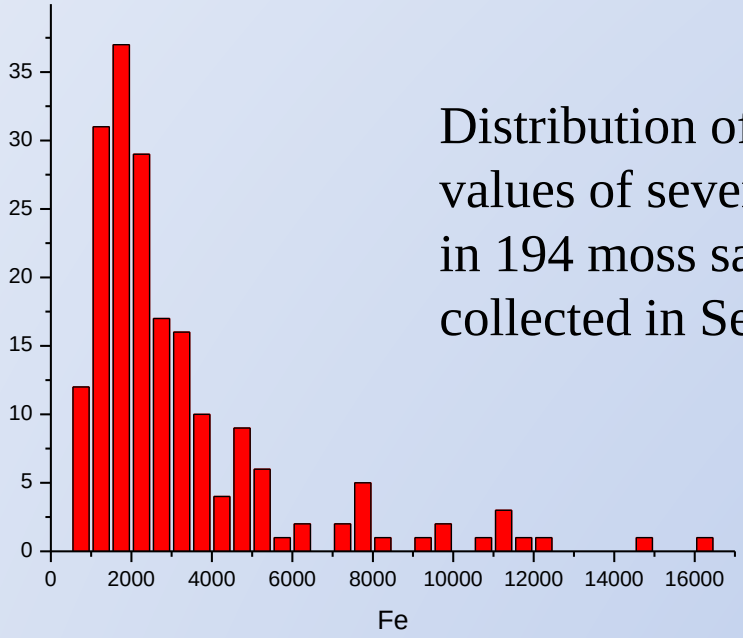
Let us see the distribution of measured values!!!!

radionuclide	range [Bq kg^{-1}]	Arithmetic mean [Bq kg^{-1}]	Median [Bq kg^{-1}]
^{7}Be	60 - 538	251(94)	244
^{40}K	128 - 2235	465(268)	400
^{137}Cs	mda - 296	23(43)	8
^{212}Pb	3 - 259	58(40)	50
^{214}Bi	mda - 793	114(116)	79

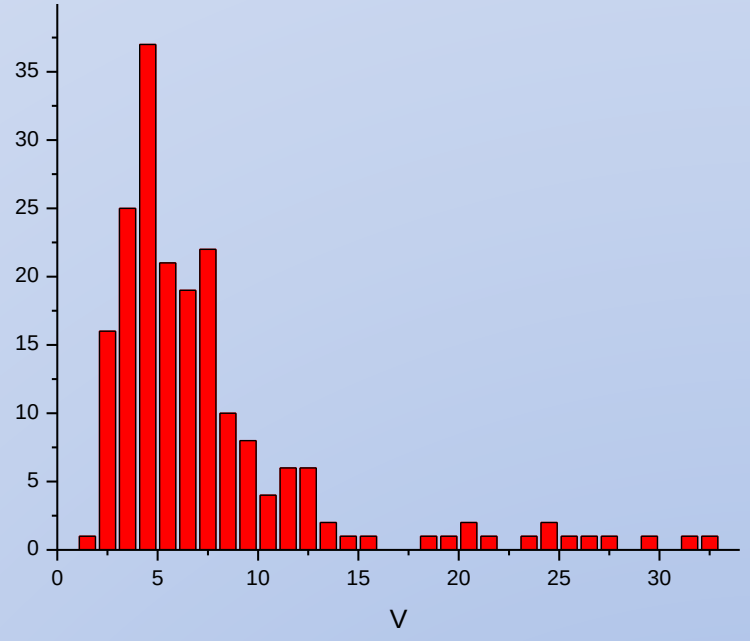
Only one symmetric distribution (Gauss - like) is obtained between all measurements of moss samples. It was in analysis of ^{7}Be activity concentrations measured in 196 moss samples collected in Serbia.

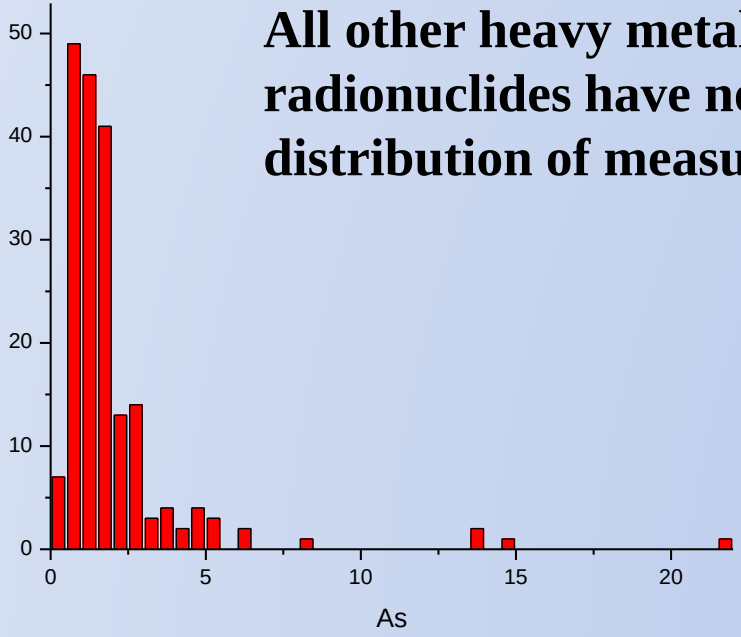




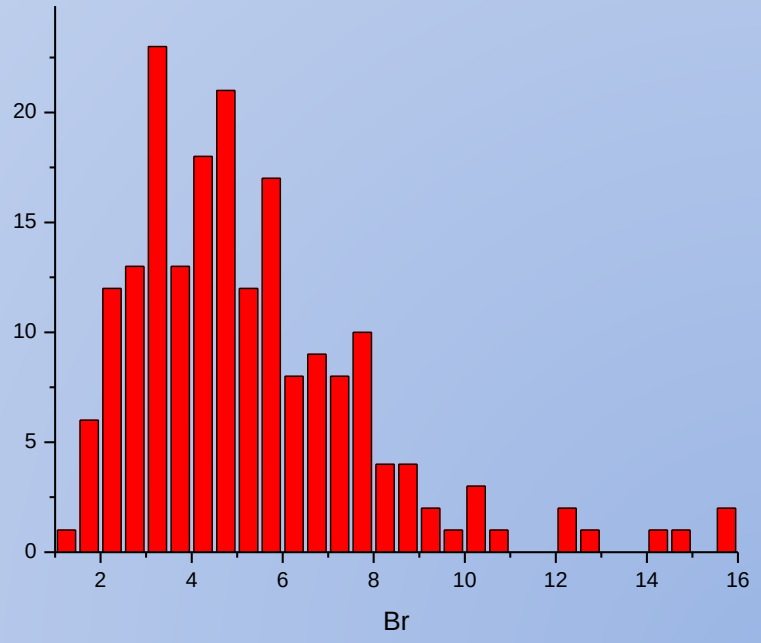


Distribution of measured values of several elements in 194 moss samples collected in Serbia.





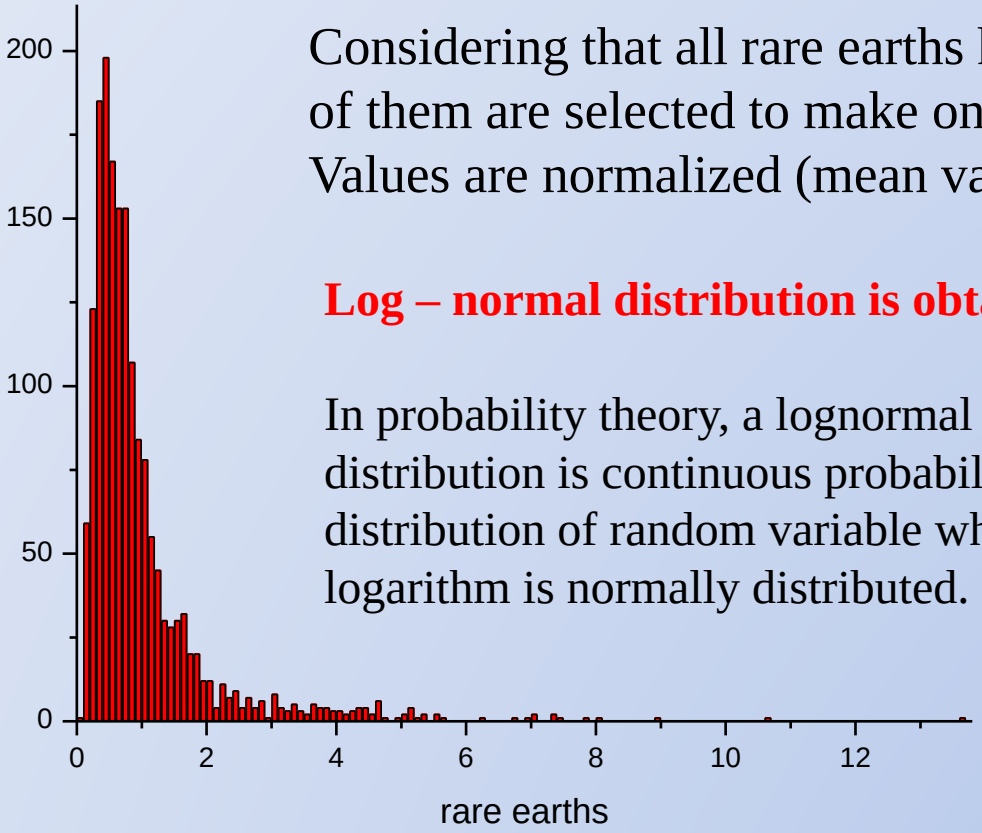
All other heavy metals or radionuclides have not symmetric distribution of measured values.

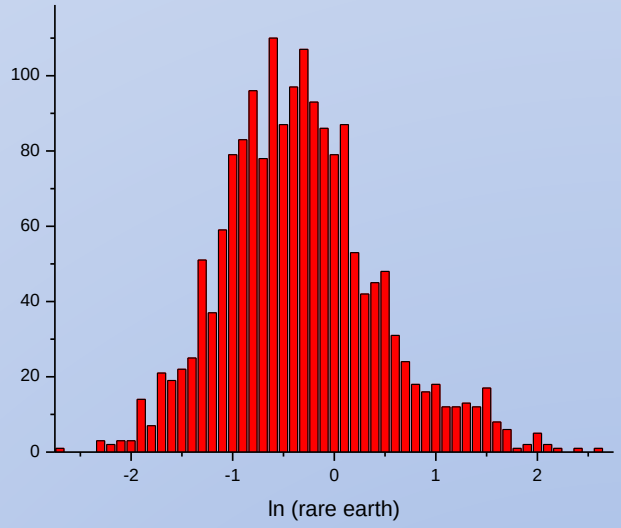


Considering that all rare earths have similar distribution, eight of them are selected to make one data set (1738 values). Values are normalized (mean value = 1)

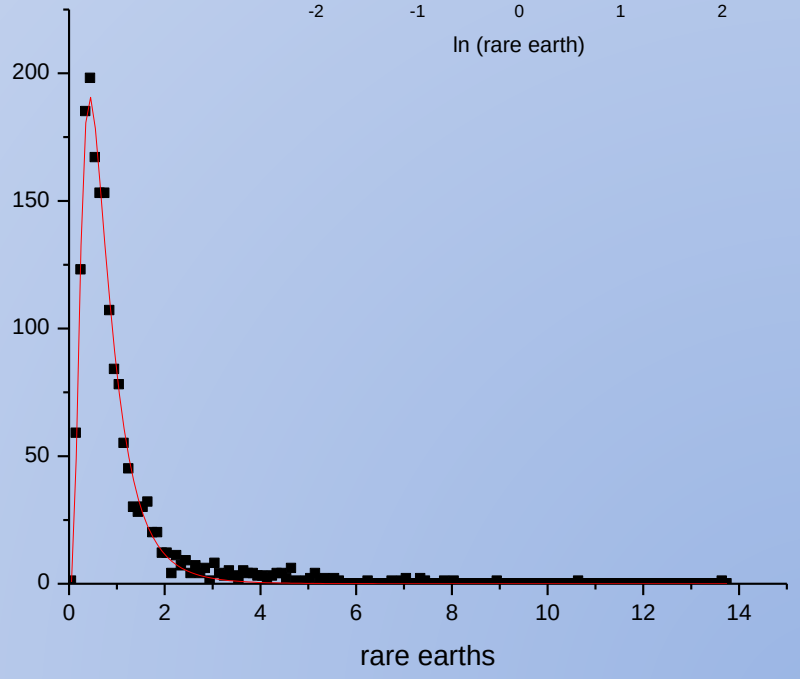
Log – normal distribution is obtained

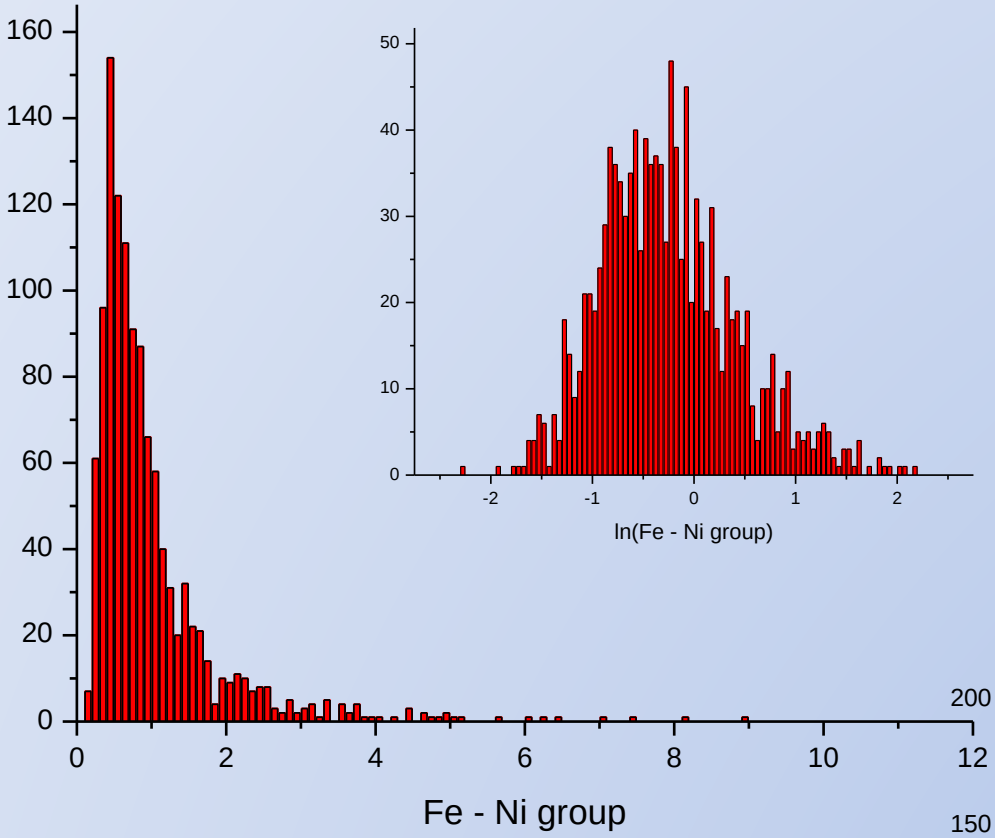
In probability theory, a lognormal distribution is continuous probability distribution of random variable whose logarithm is normally distributed.





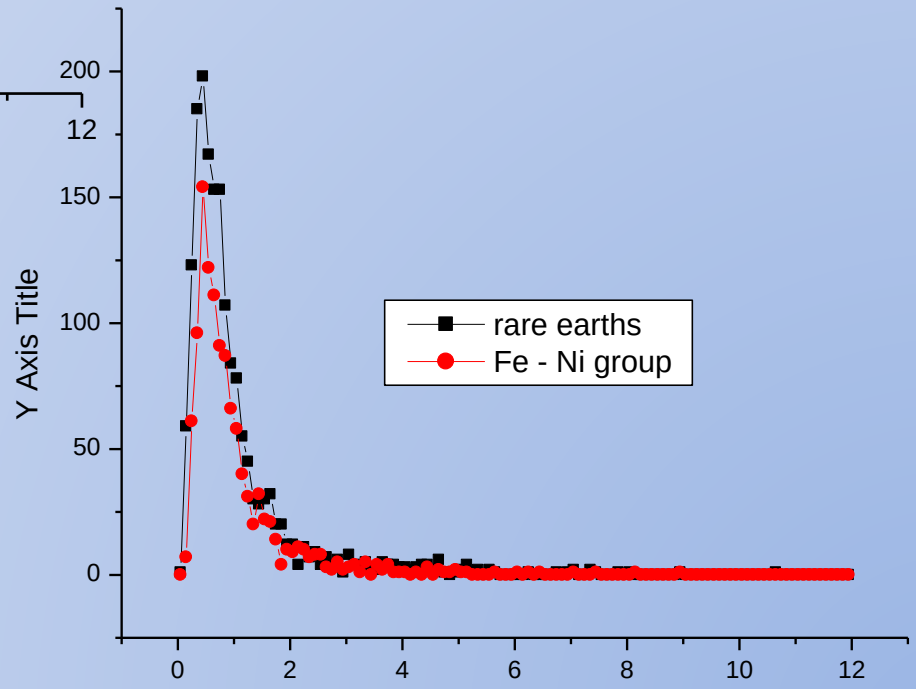
$$y = 165 \frac{exp\left(-\frac{\left(ln\frac{x}{0.657}\right)^2}{0.645}\right)}{\sqrt{2\pi\ 0.645\ x}}$$



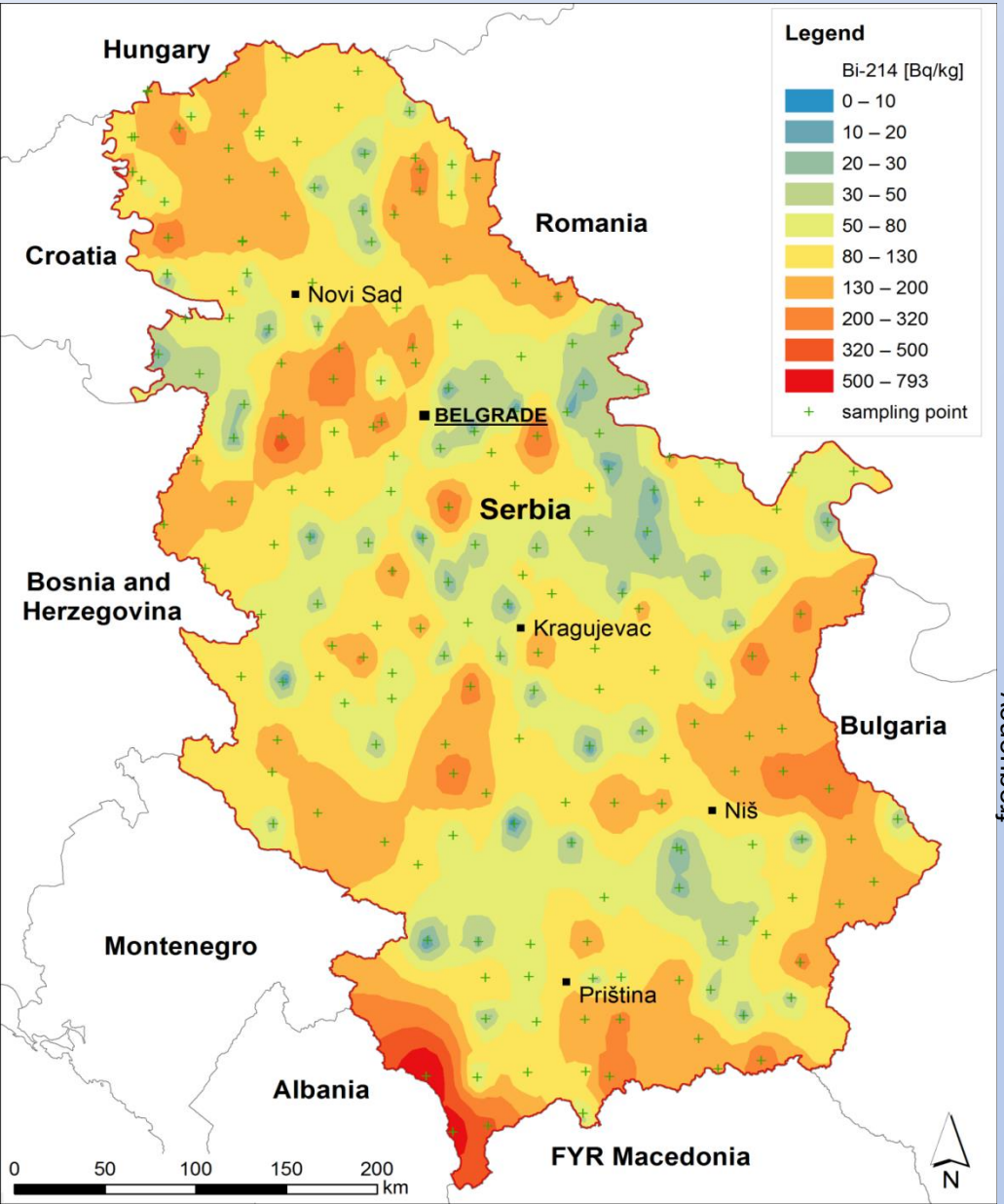


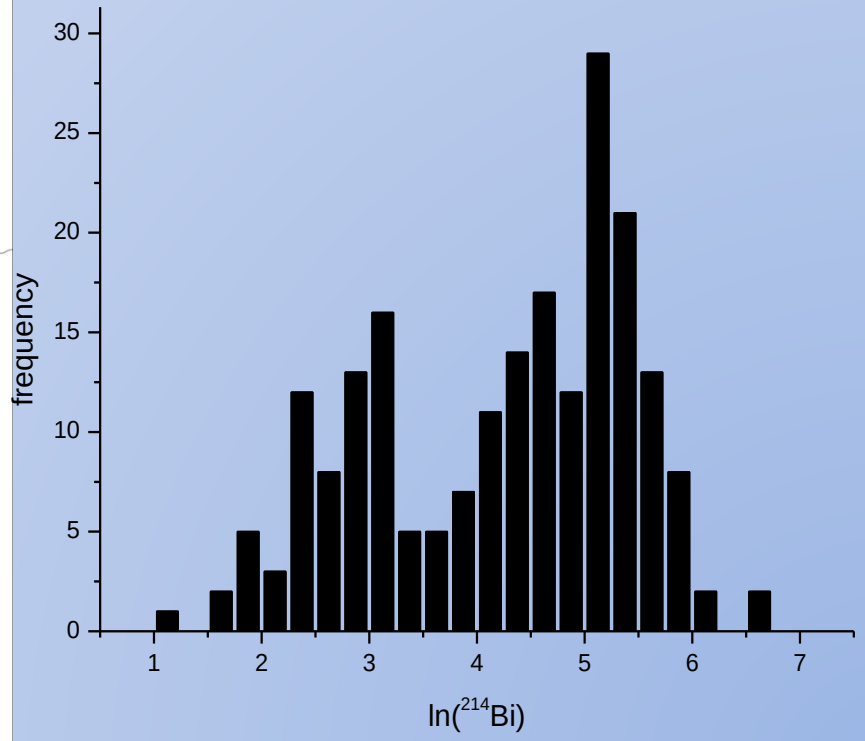
The very same procedure was repeated for sum of normalized values of Sc, Ti, Cr, Fe, Co and Ni (1159 values).
Very similar lognormal distribution was obtained

$$y = 108 \frac{exp\left(-\frac{ln\frac{x}{0.688}}{0.571}\right)}{\sqrt{2\pi\ 0.571\ x}}$$



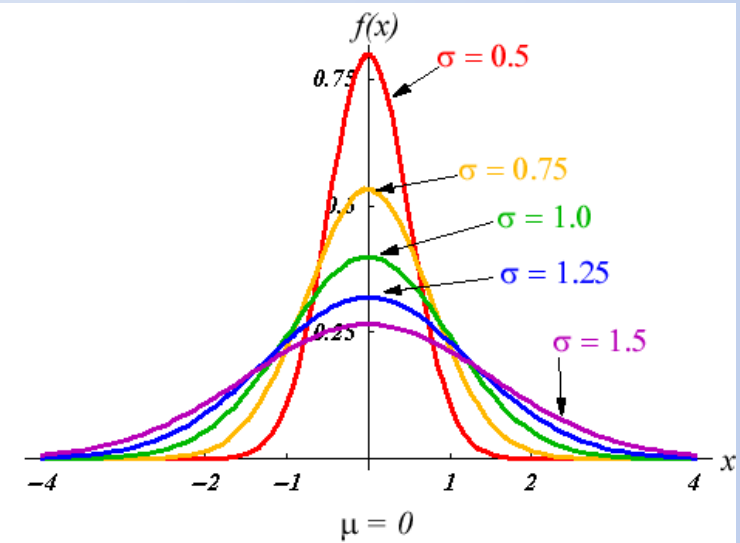
Distribution of values of ^{214}Bi (^{238}U) measured by gamma spectroscopy in 217 moss samples from Serbia.

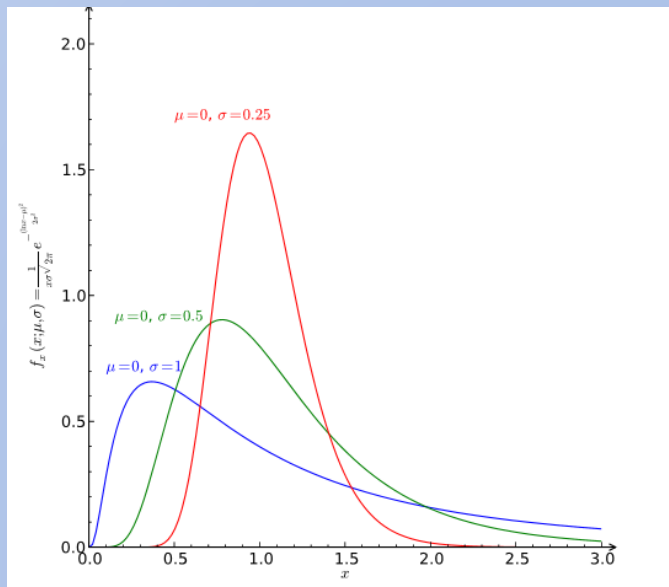


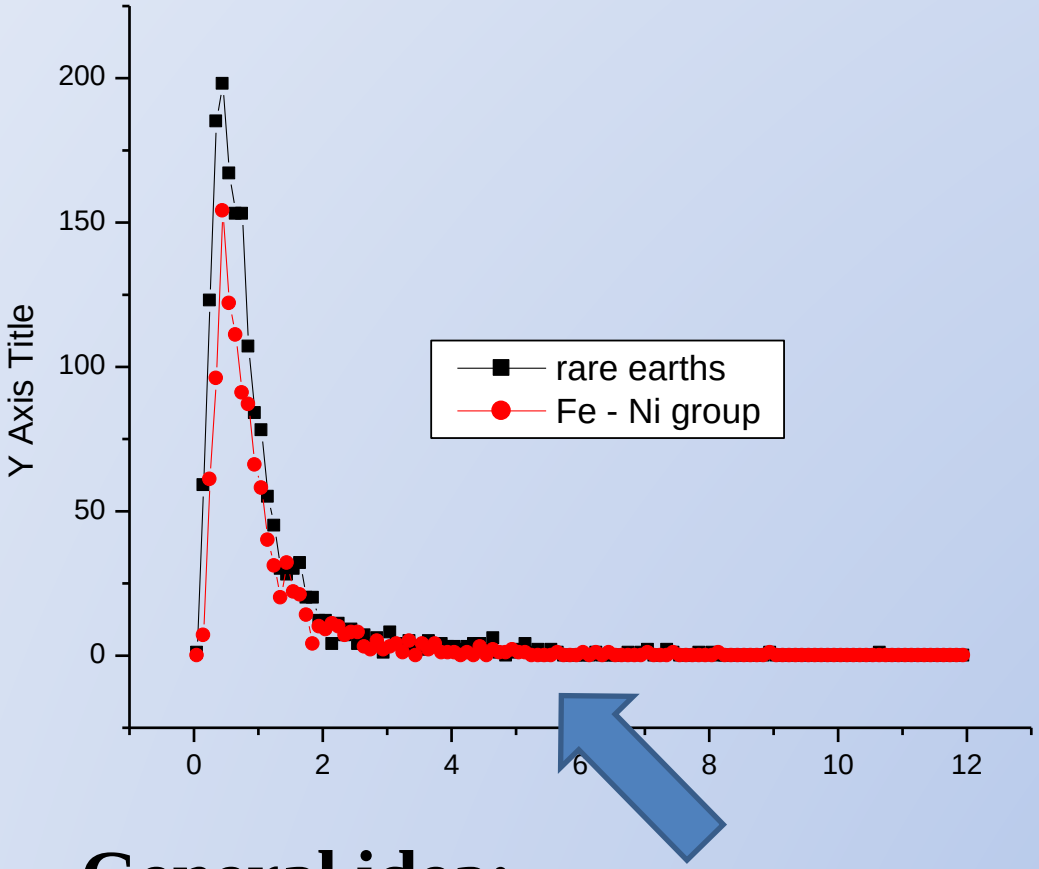


A lognormal process is statistical realization of the multiplicative product of many independent random variables. Many natural processes are driven by the accumulation of many small (differential) changes. These become additive on a log scale and transformation onto the linear scale makes the distribution lognormal.

Physical quantities that are expected to be the sum of many independent processes have normal distributions.

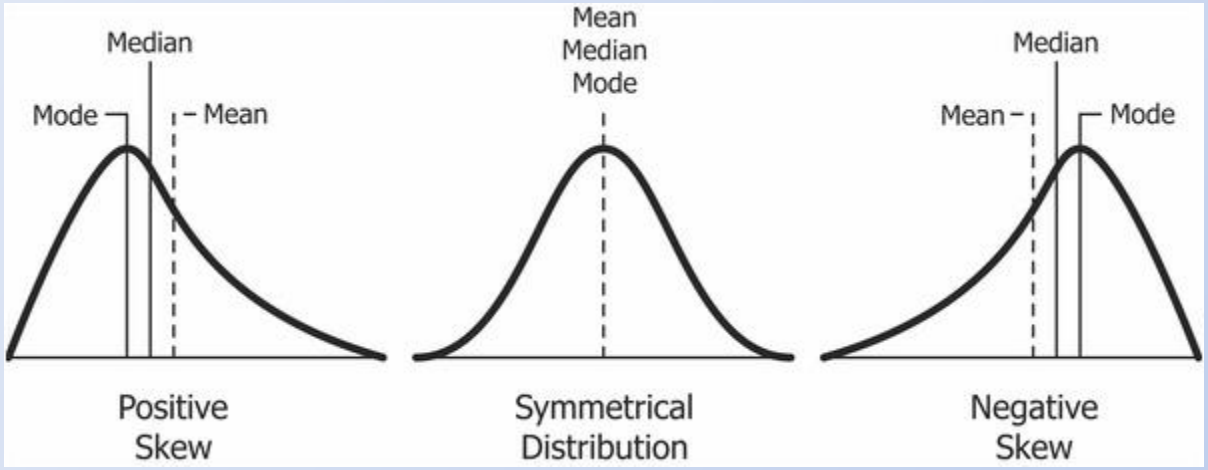






General idea:

If the lognormal distribution is characteristic distribution of measured values of heavy elements or radionuclides in moss samples, possible pollution and detection of unusually elevated values could be observed at the tail of distribution.

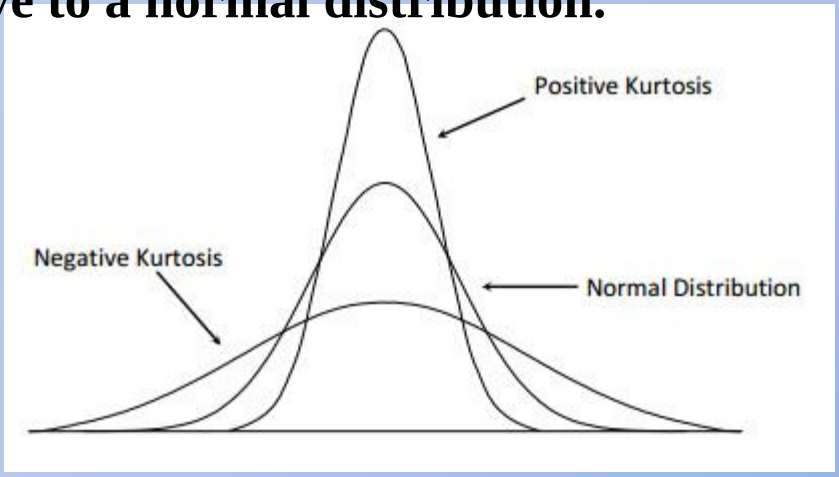


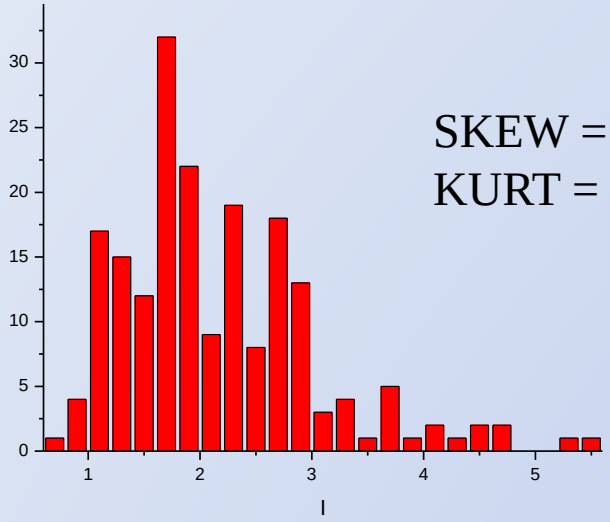
$$S = \frac{1}{n} \frac{\sum_{i=1}^{n} (X_i - X_{avg})^3}{\sigma^3}$$

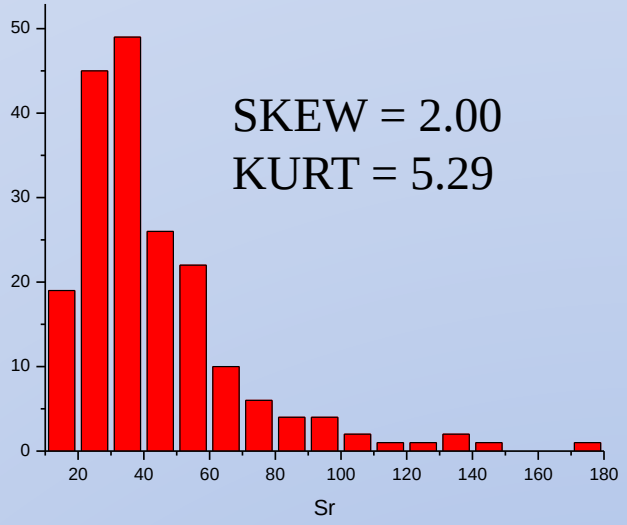
Skewness is a measure of the asymmetry of some distribution of a real valued variable about its mean.

Kurtosis is a measure of whether the data are heavy-tailed or light-tailed relative to a normal distribution.

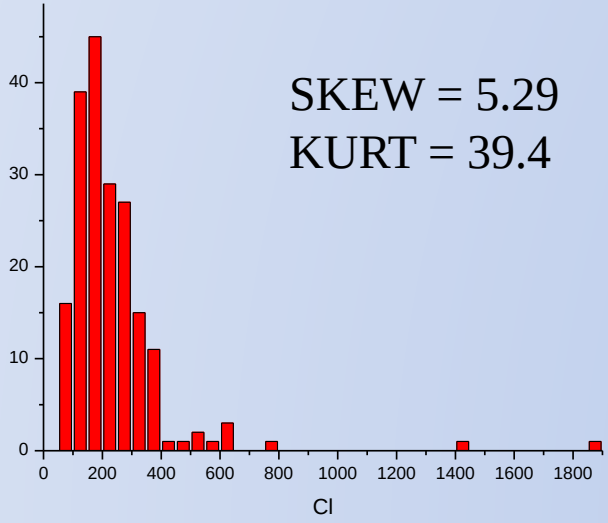
$$K = n \frac{\sum_{i=1}^{n} (X_i - X_{avg})^4}{(\sum_{i=1}^{n} (X_i - X_{avg})^2)^2} - 3$$

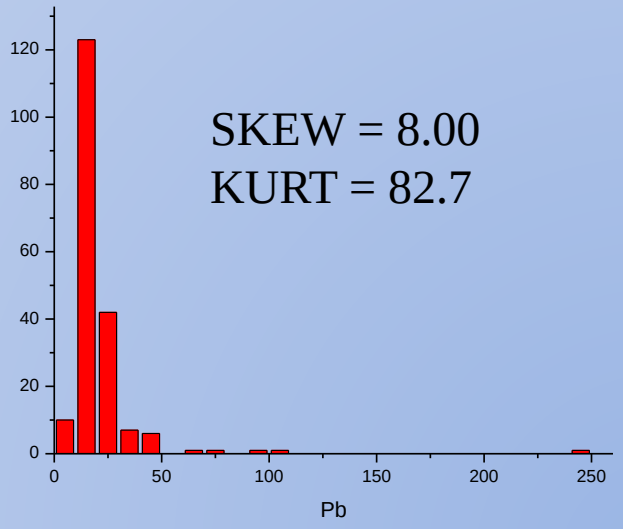


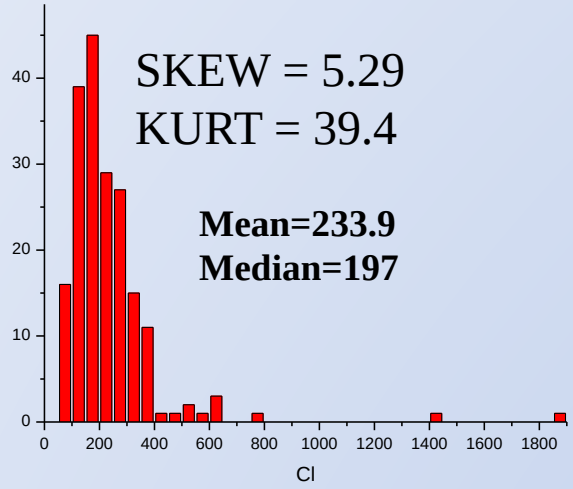


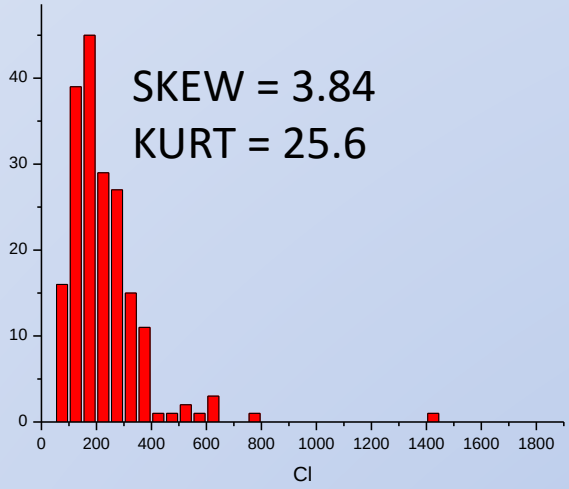


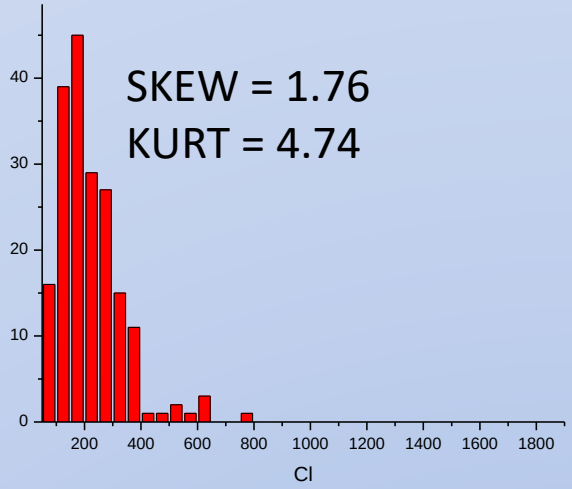
Skewness, kurtosis and distributions of several elements measured in moss samples collected 2005 in Serbia. There were 194 moss samples.



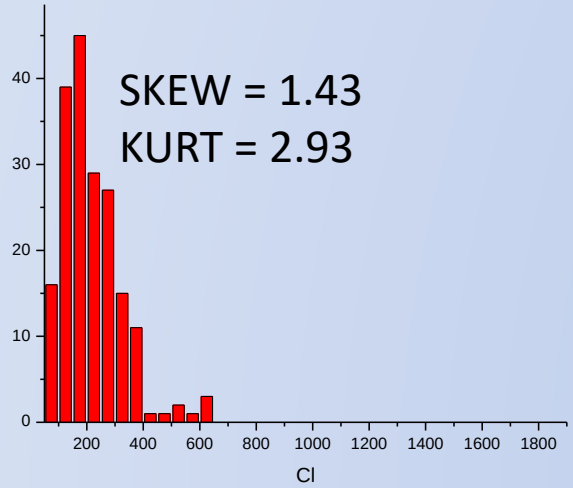


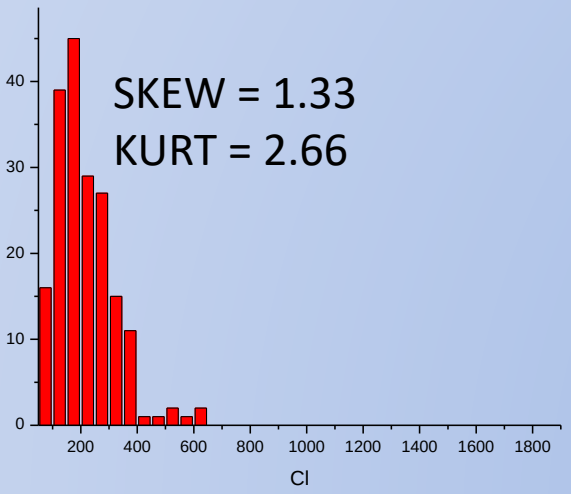


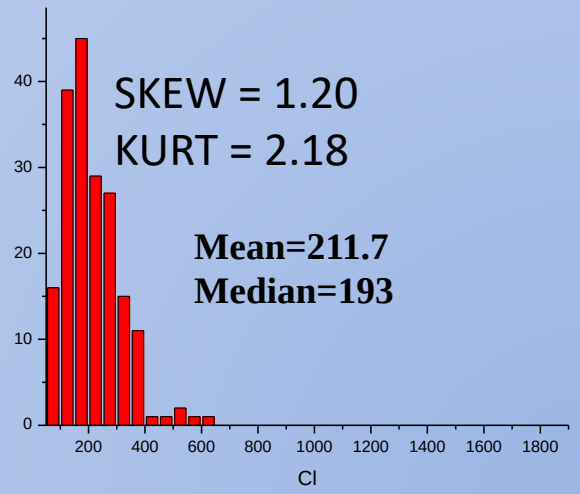


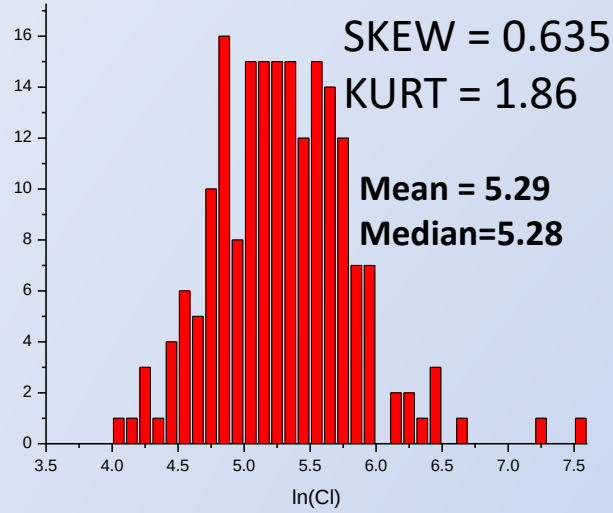


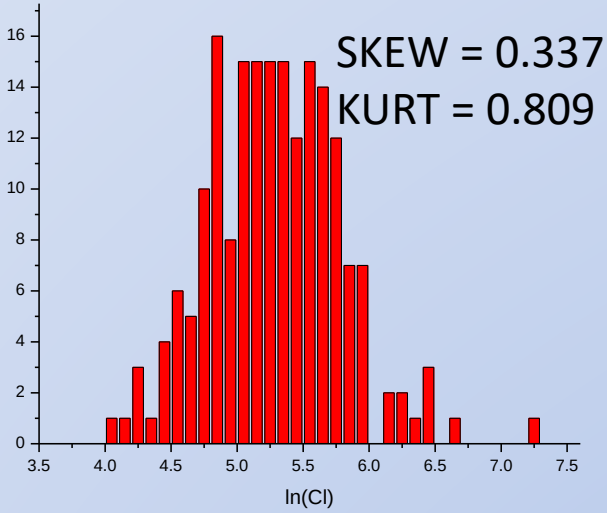
Skewness of the distribution of measured values of Cl in 194 moss samples collected in 2005 in Serbia. The highest measured values are successively deleted.

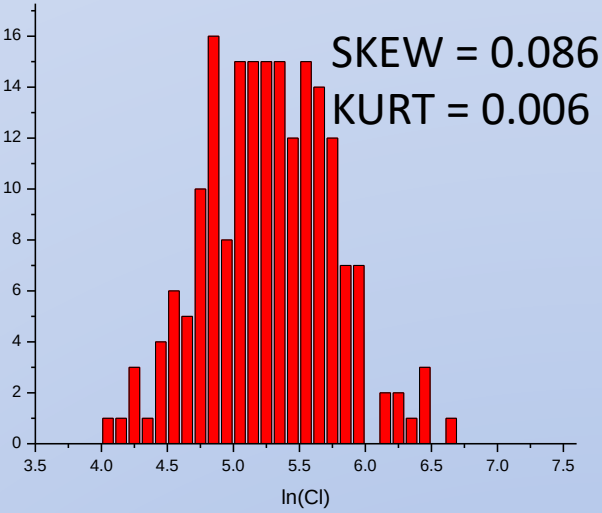




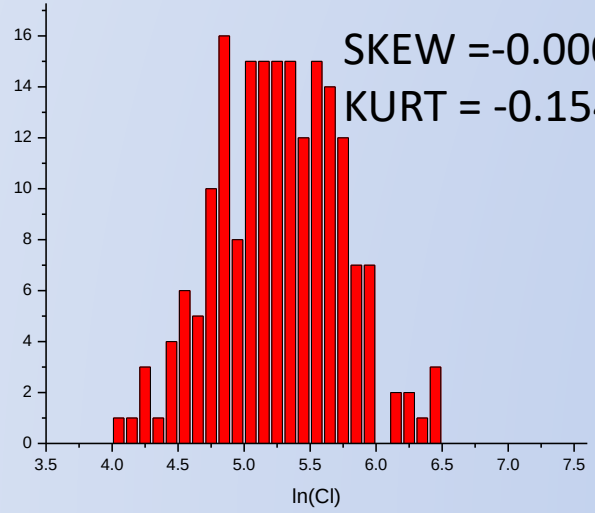


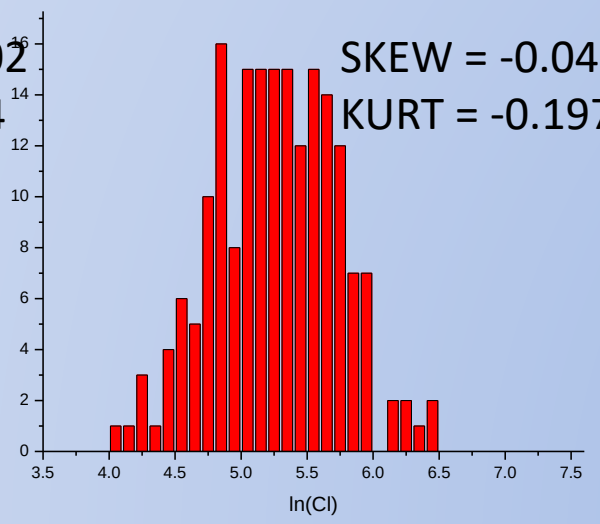


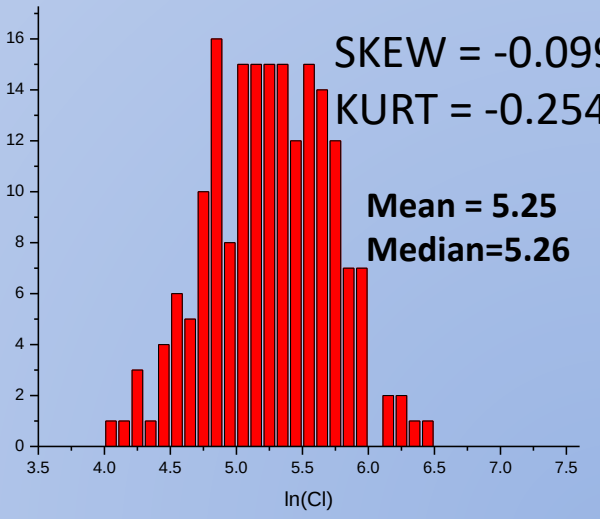




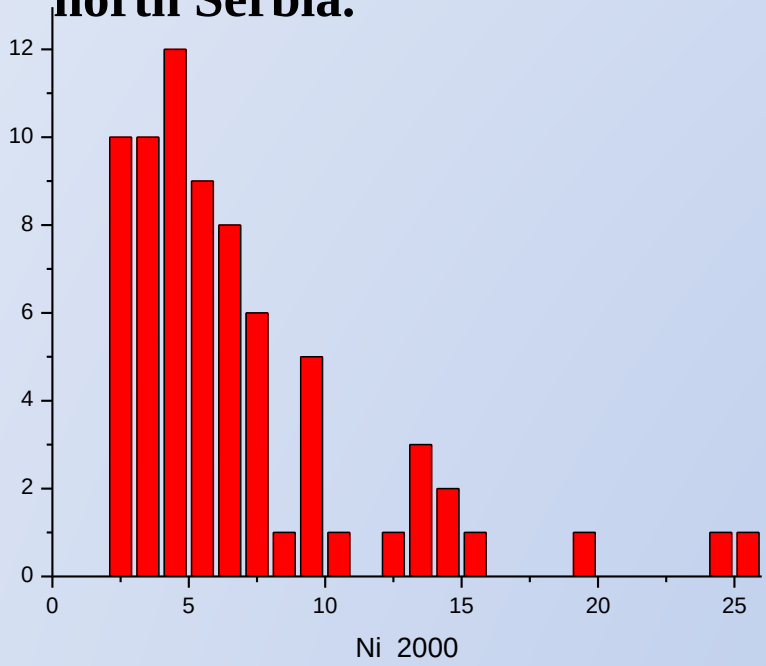
Very same calculation was repeated using logarithm values of measured concentrations of Cl in 194 moss samples.



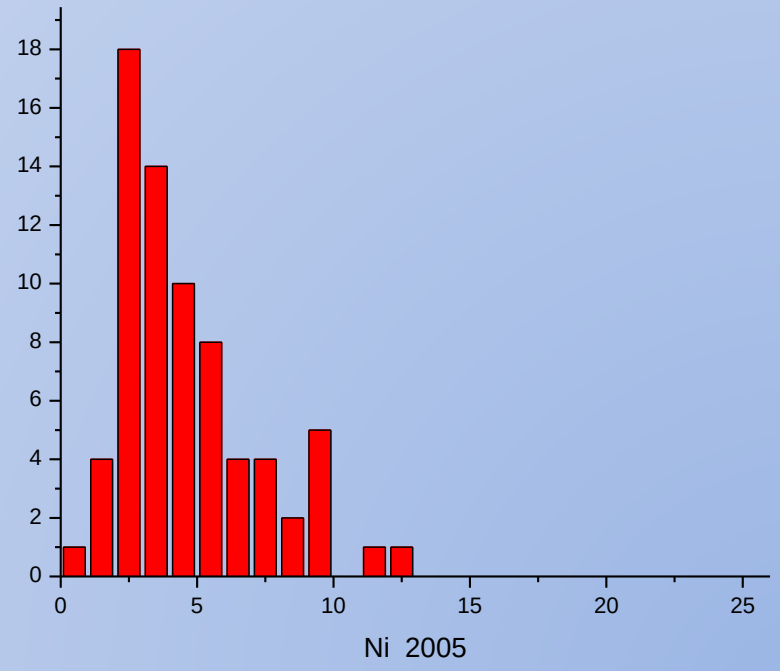




Let us compare two sets of data: moss survey 2000 and moss survey 2005. There were 75 moss samples taken from very same location in the area of north Serbia.

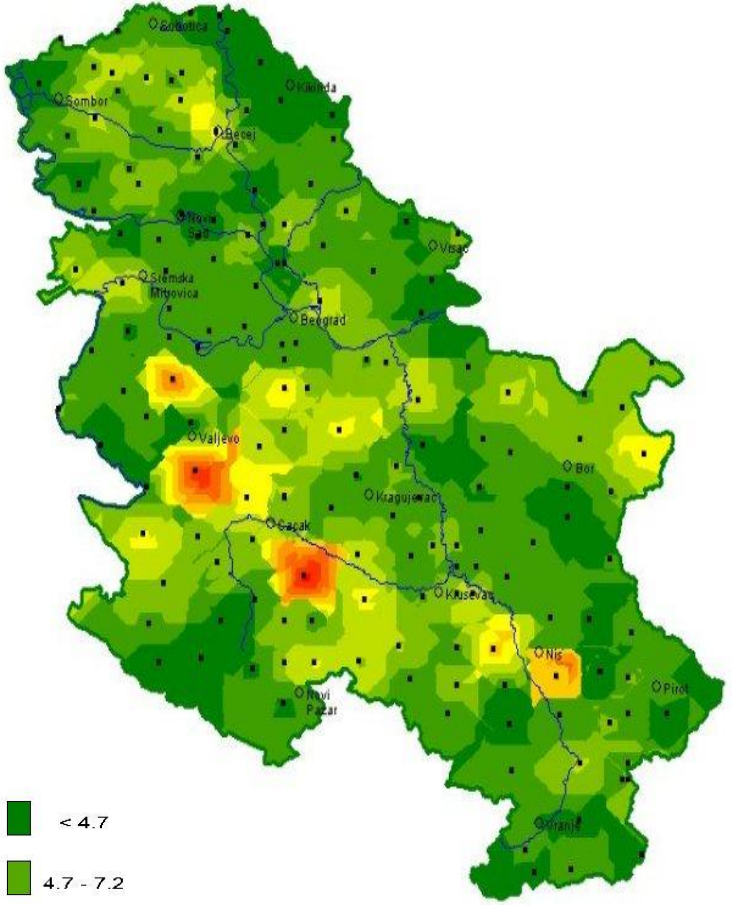


	Ni 2000	Ni 2005
Min	2.09	0.915
Max	25.7	12.2
Mean	6.87	4.61
Median	5.60	3.81
Skew	2.10	1.08
Kurt	5.10	0.736



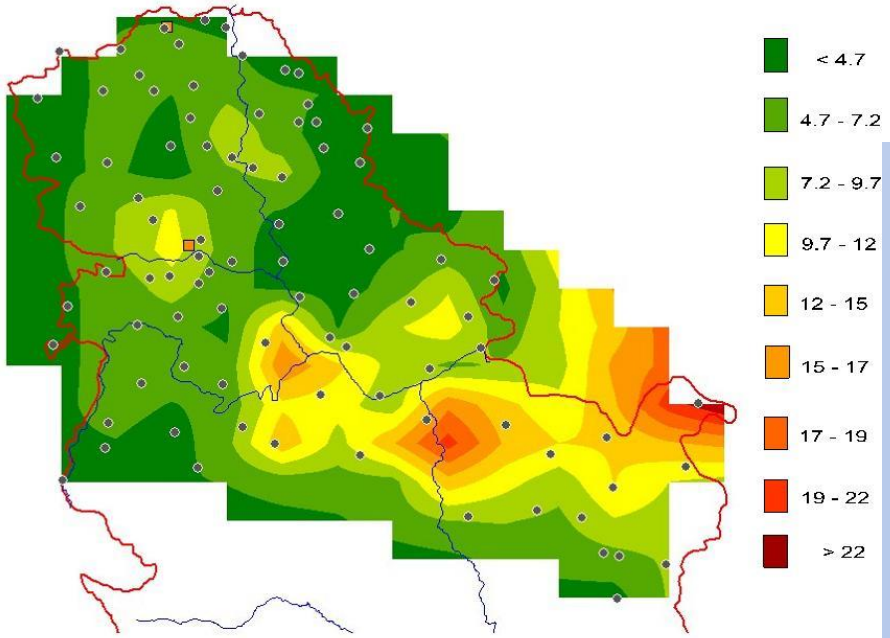
NI

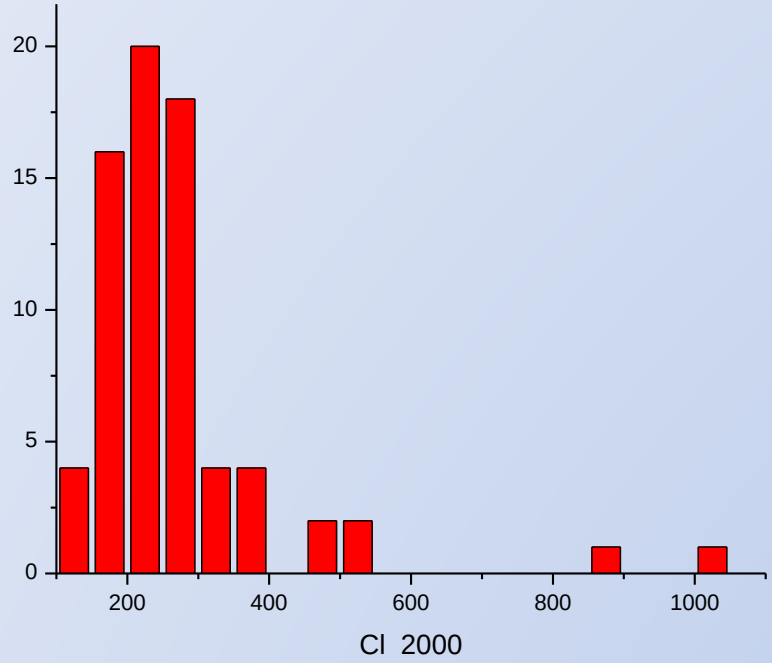
2005



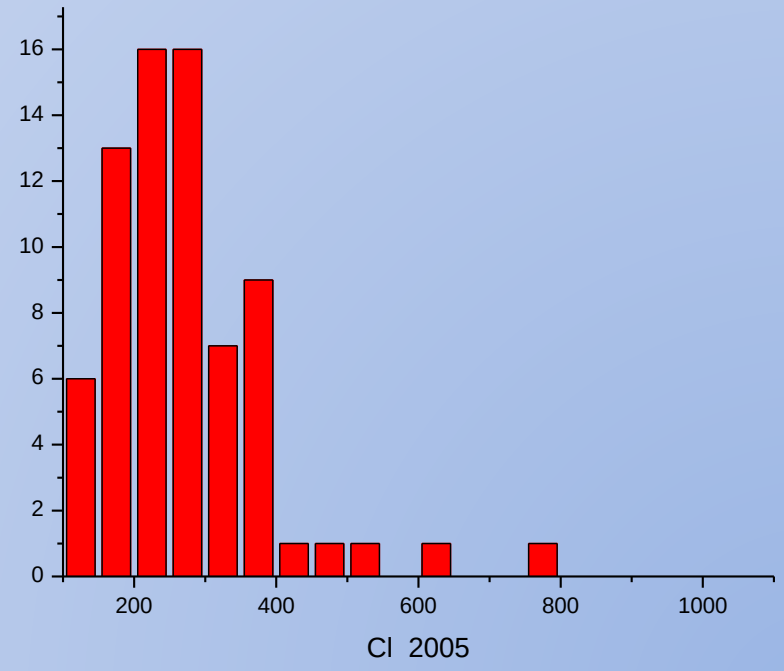
2000

Ni



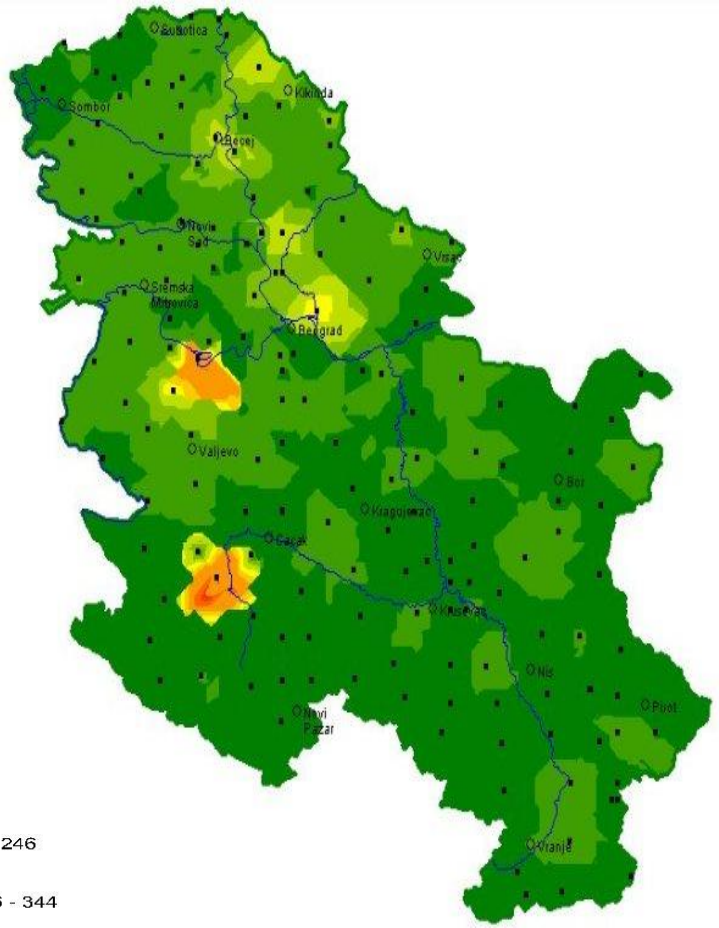


	Cl 2000	Cl 2005
Min	**131.3**	**104.1**
Max	**1032**	**793.2**
Mean	**269.6**	**269**
Median	**235.7**	**251**
Skew	**3.32**	**1.92**
Kurt	**13.8**	**6.09**



CL

2005

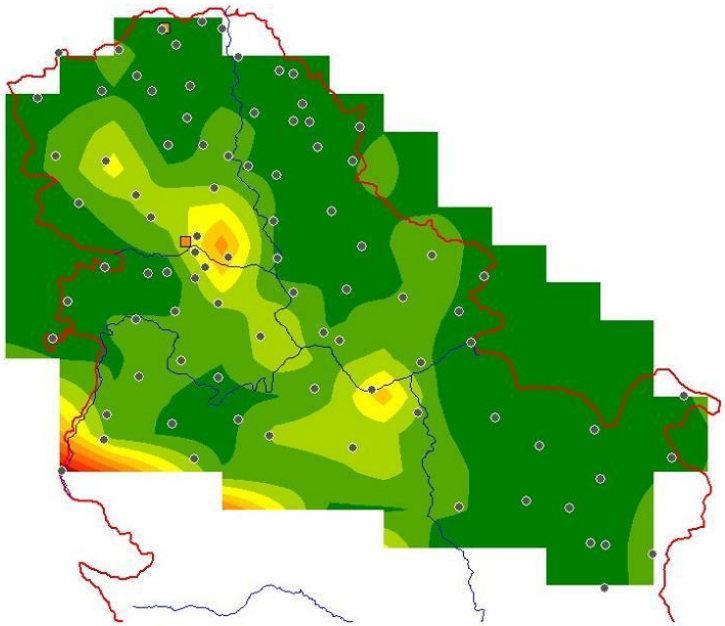


Concentration [ppm]

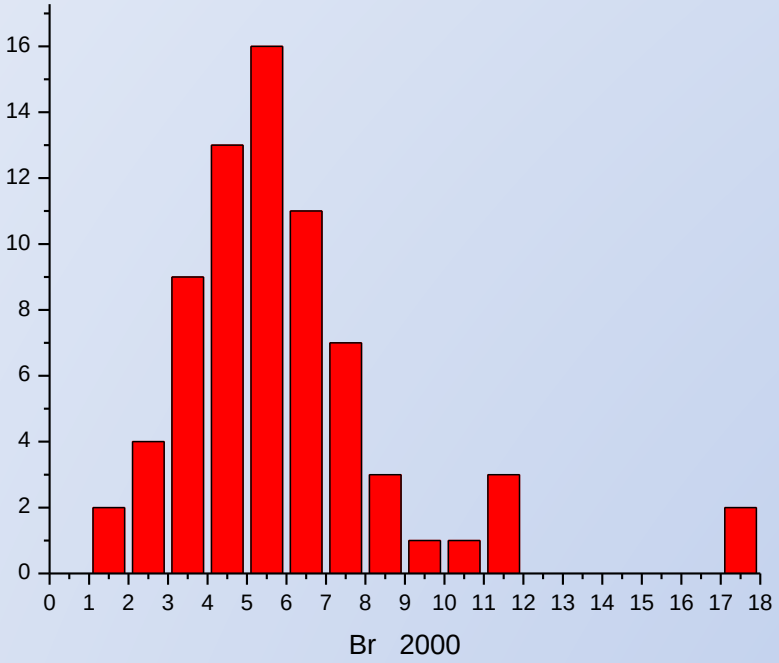
- <210
- 211-361
- 362-511
- 512-661
- 662-811
- 812-961
- 962-1111
- 1112-1261
- 1262-1411
- >1412

2000

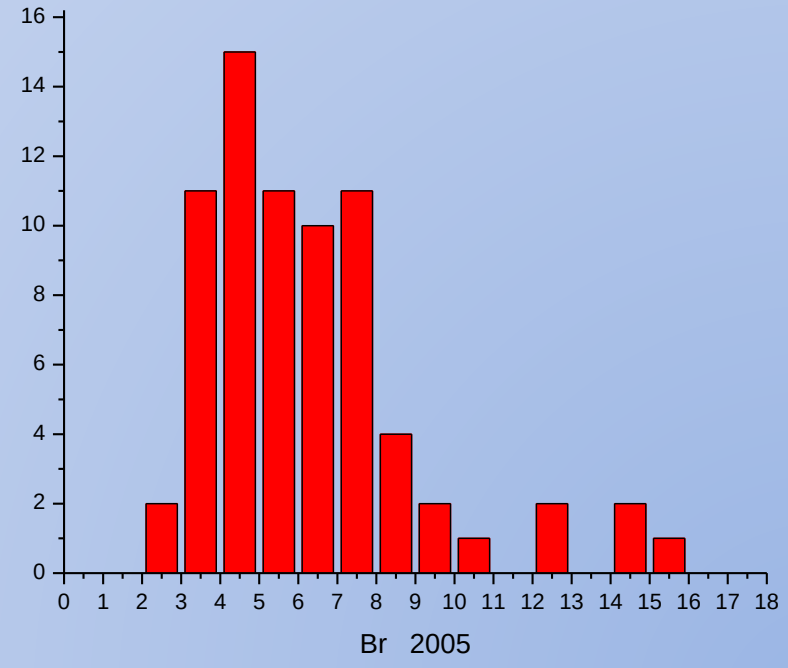
Cl

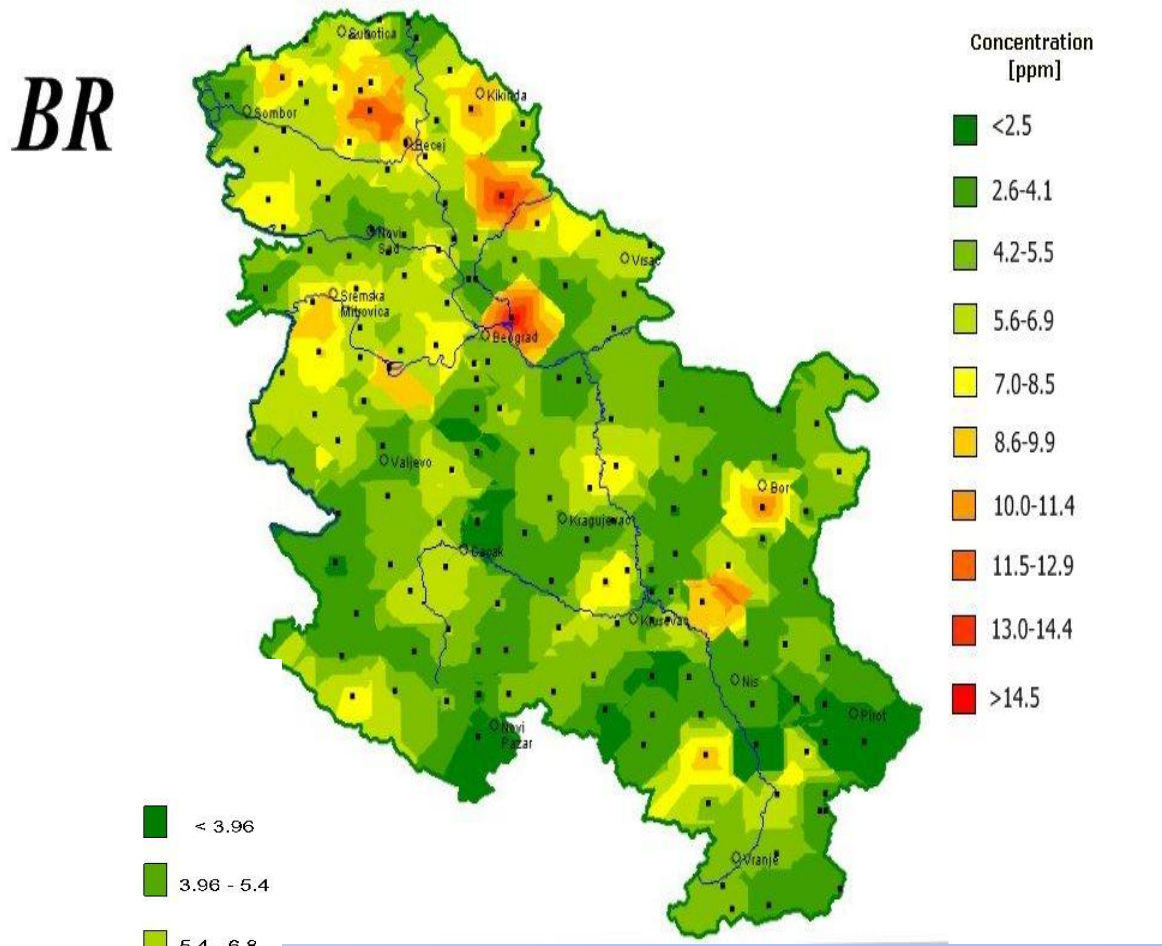


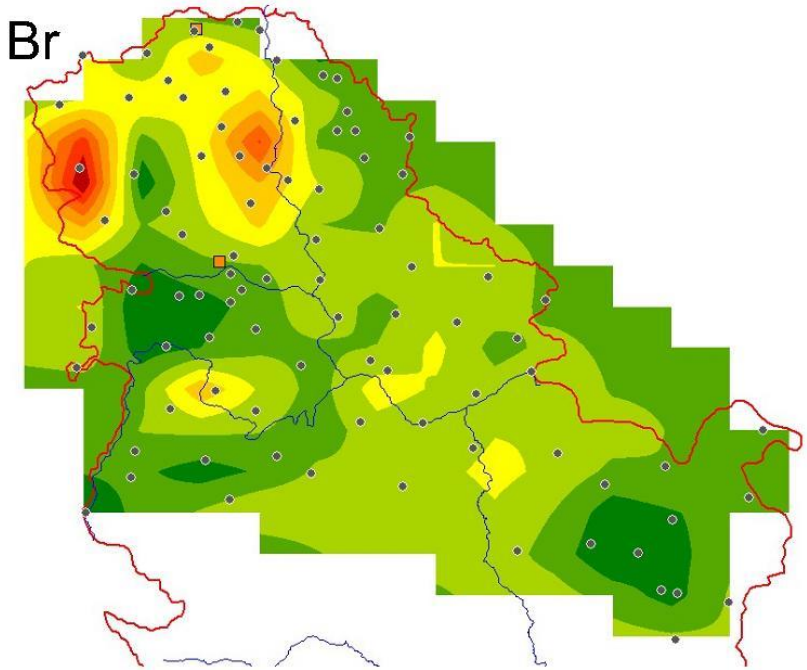
- < 246
- 246 - 344
- 344 - 441
- 441 - 539
- 539 - 637
- 637 - 735
- 735 - 832
- 832 - 930
- > 930

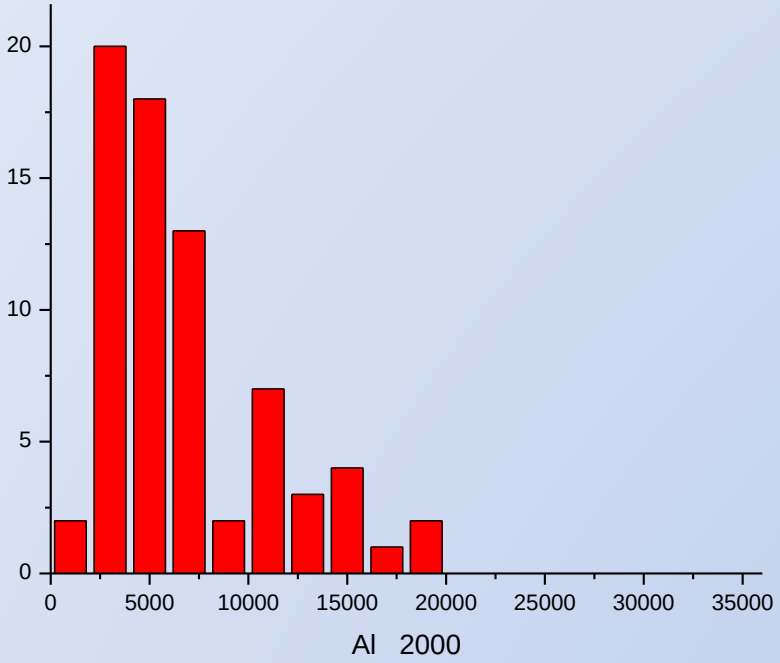


	Br 2000	Br 2005
Min	**1.83**	**2.32**
Max	**17.6**	**15.9**
Mean	**3.56**	**6.31**
Median	**5.47**	**5.77**
Skew	**1.97**	**1.51**
Kurt	**5.69**	**2.68**



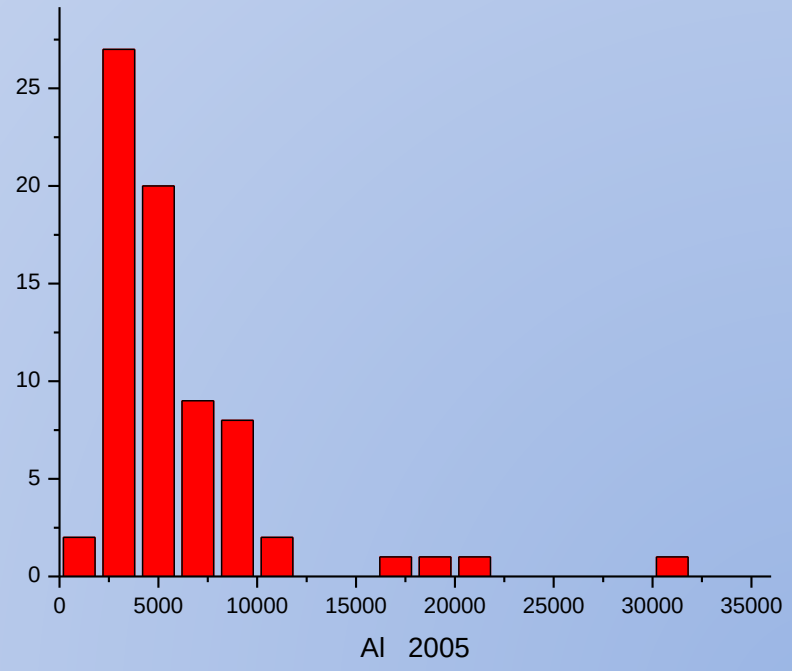

BR
Concentration [ppm]
<2.5
2.6-4.1
4.2-5.5
5.6-6.9
7.0-8.5
8.6-9.9
10.0-11.4
11.5-12.9
13.0-14.4
>14.5
Subotica
Kikinda
Sombor
Becej
Novi Sad
Visac
Sremska Mitrovica
Beograd
Valjevo
Bor
Kragujevac
Gacak
Kursumlija
Nis
Pirot
Novi Pazar
Vranje


Br
< 3.96
3.96 - 5.4
5.4 - 6.8
6.8 - 8.2
8.2 - 9.6
9.6 - 11
11 - 12.5
12.5 - 13.9
> 13.9



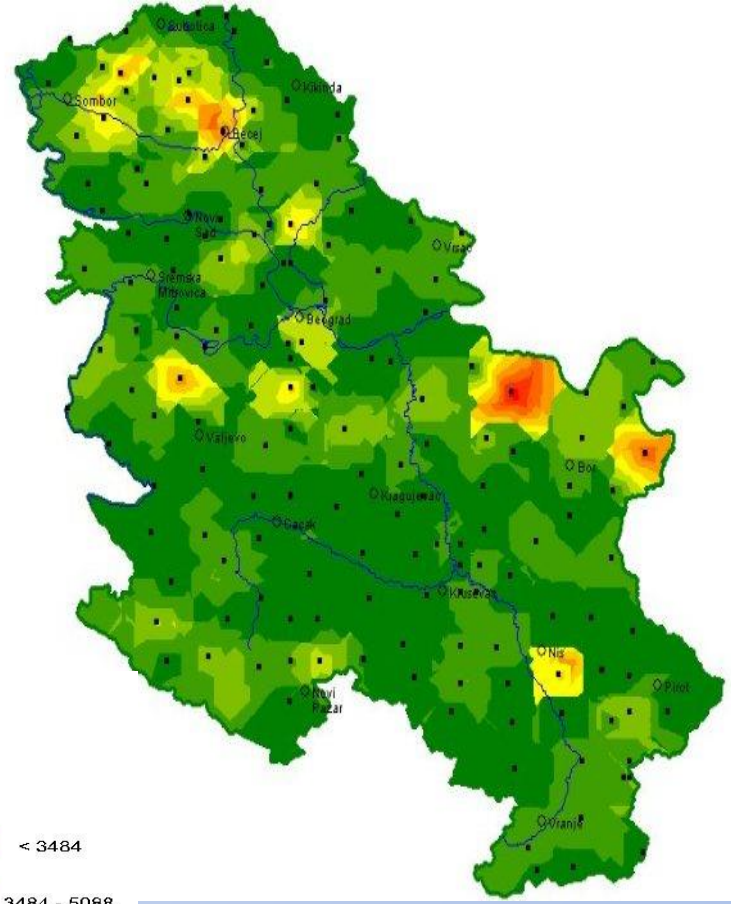
Couple examples when Skewness and Kurtosis are higher in 2005 than in 2000 moss samples.

	Al 2000	Al 2005
Min	1275	1938
Max	18660	31180
Mean	6841	5905
Median	5521	4628
Skew	1.18	3.23
Kurt	0.615	13.1



AL

2005

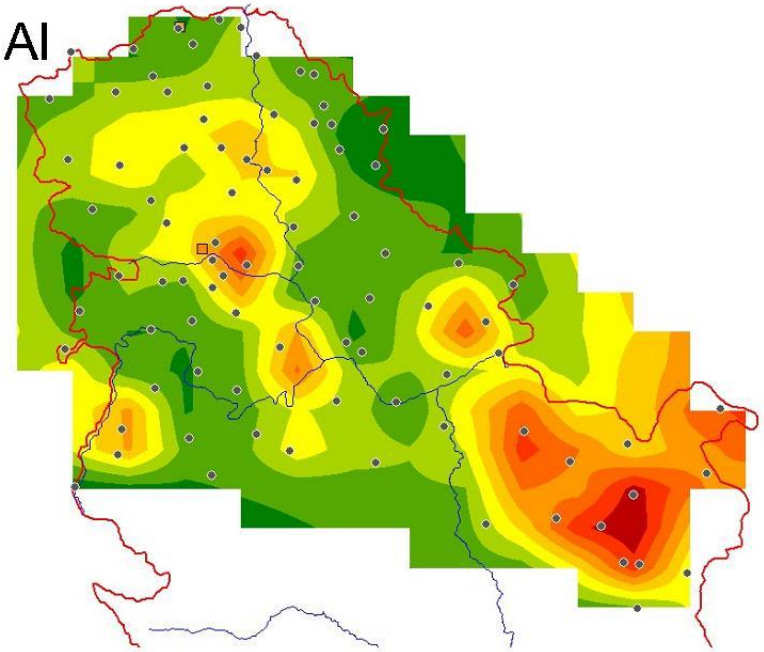


Concentration [ppm]

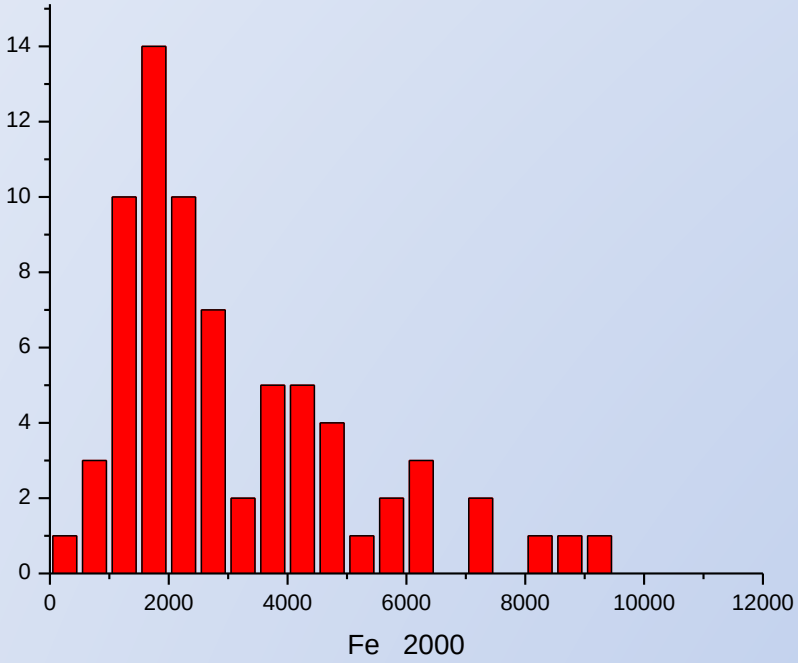
- <4190
- 4191-7189
- 7190-10187
- 10188-13186
- 13187-16185
- 16186-19183
- 19184-22182
- 22183-25181
- 25182-28179
- >28180

2000

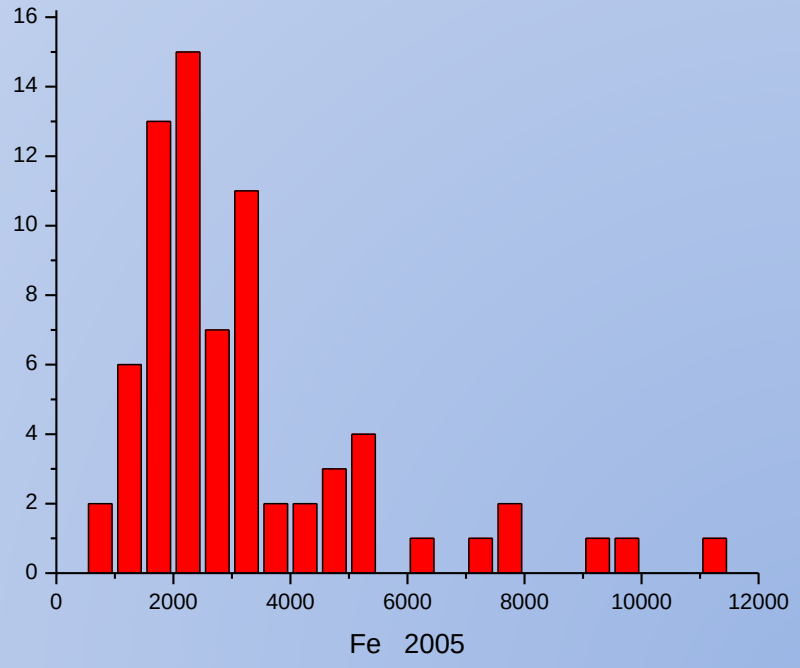
Al



- < 3484
- 3484 - 5088
- 5088 - 6697
- 6697 - 8297
- 8297 - 9901
- 9901 - 11505
- 11505 - 13109
- 13109 - 14713
- > 14713

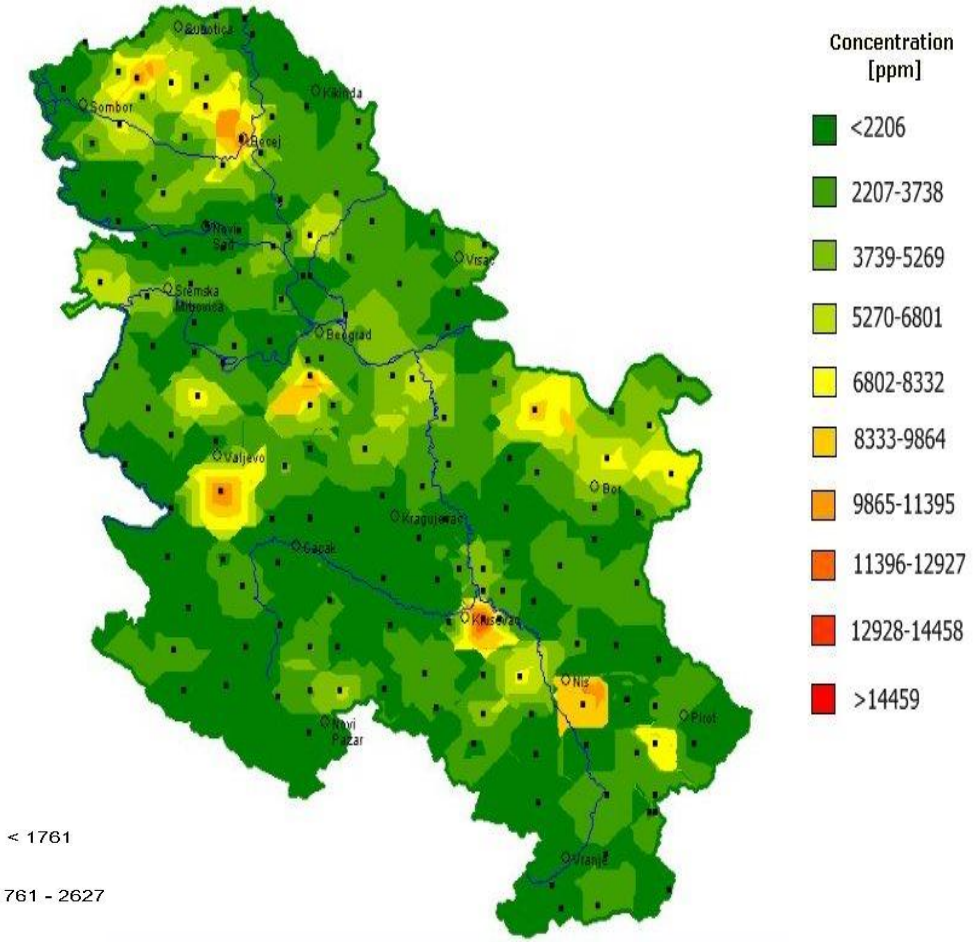


	Fe 2000	Fe 2005
Min	**337**	**675**
Max	**9229**	**11320**
Mean	**3086**	**3192**
Median	**2298**	**2540**
Skew	**1.21**	**1.88**
Kurt	**0.876**	**3.75**



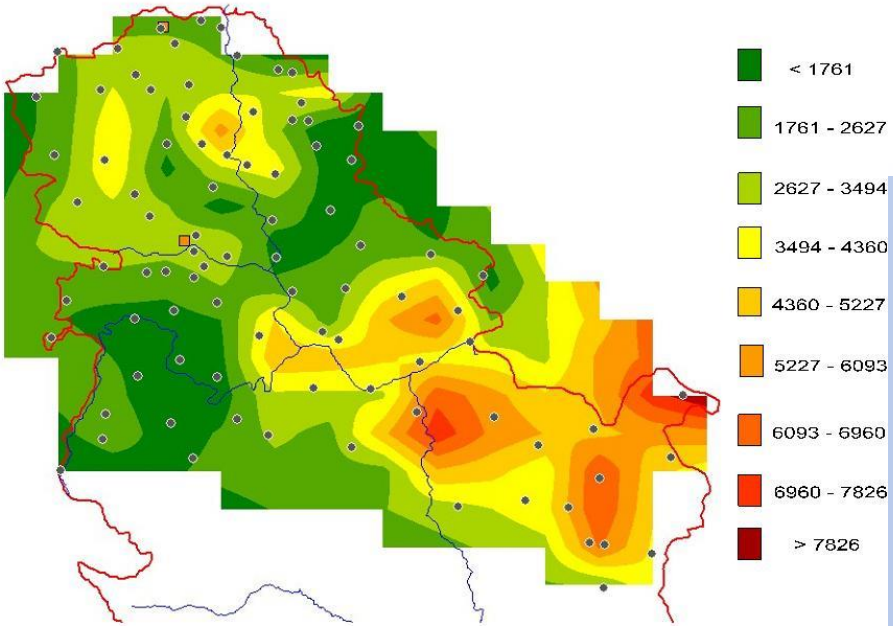
FE

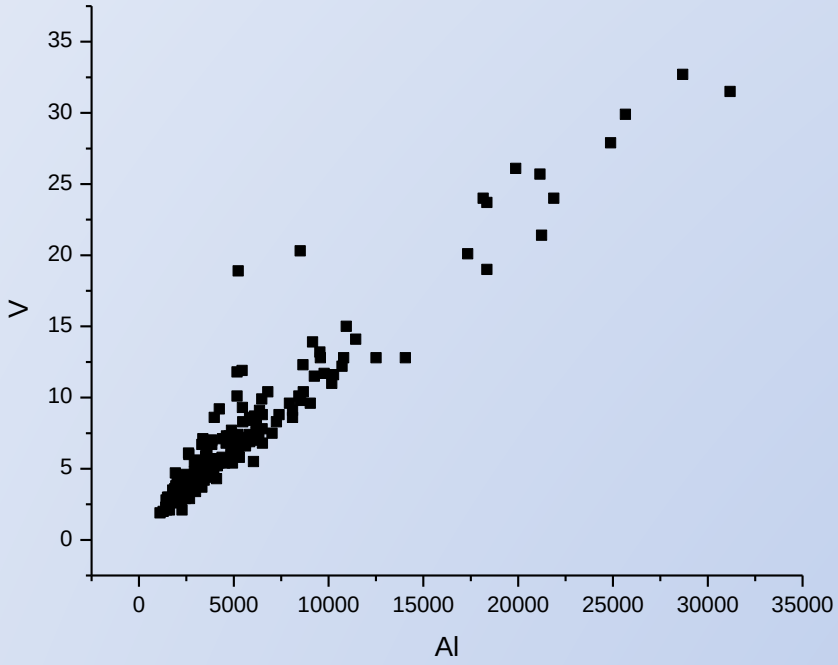
2005



2000

Fe



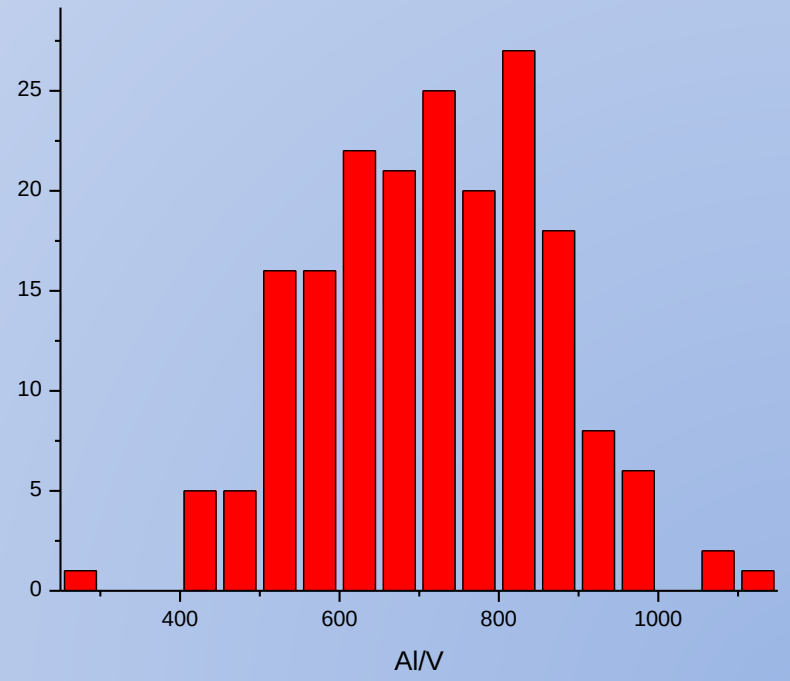


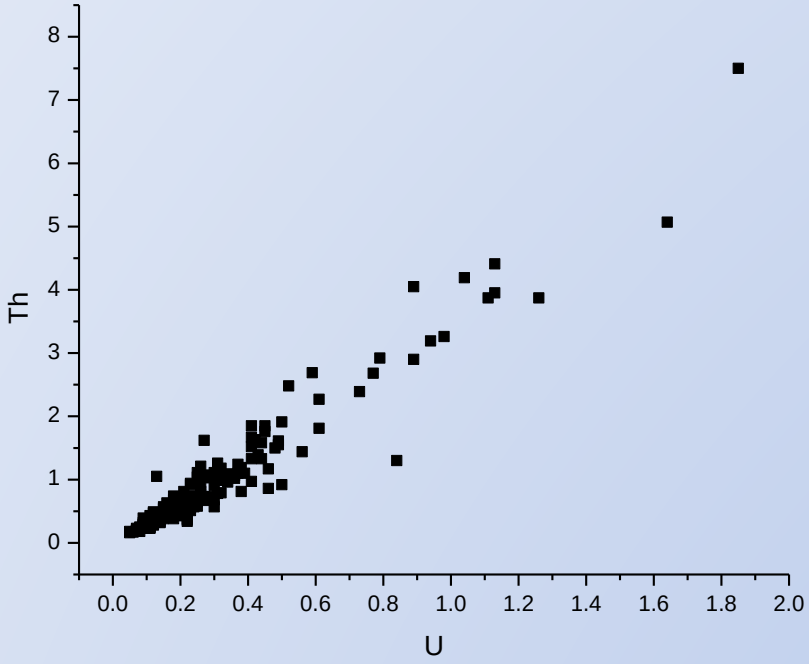
What about correlation between two elements measured in mosses?

Ratio of two elements was calculated for all moss samples and distribution of obtained ratios was analyzed.

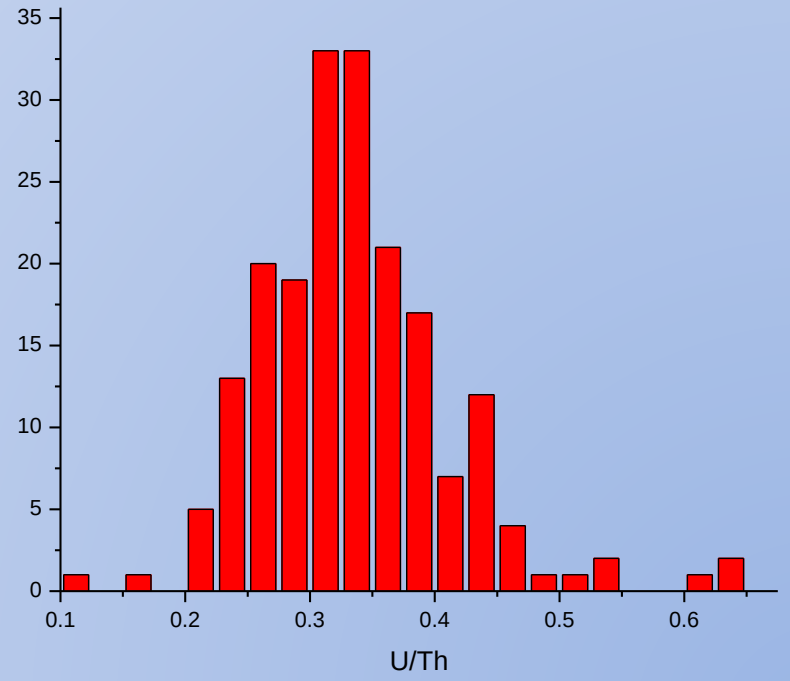
Skewness = - 0.02527
Kurtosis = - 0.1755

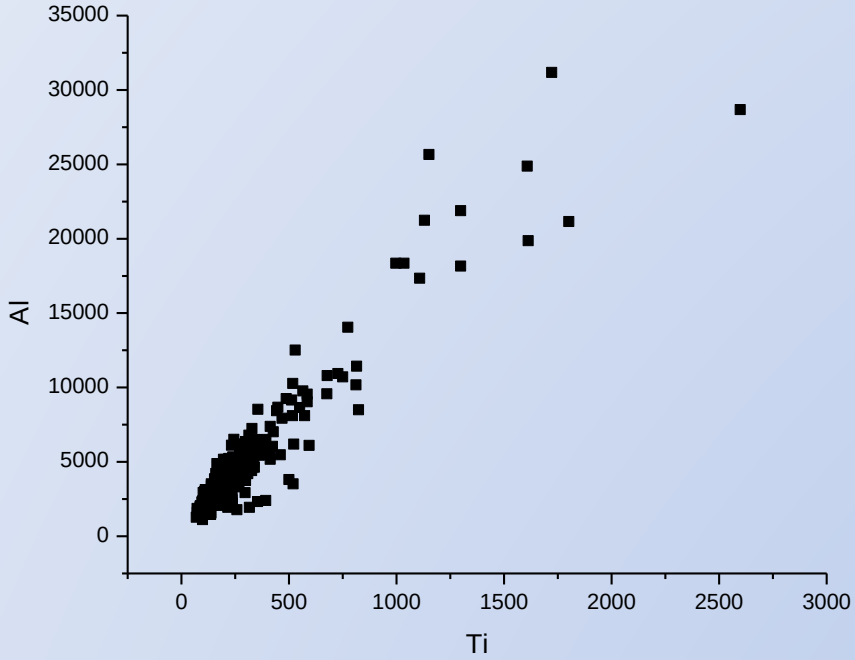
Good correlation – almost Gauss shape of distribution of calculated ratio.



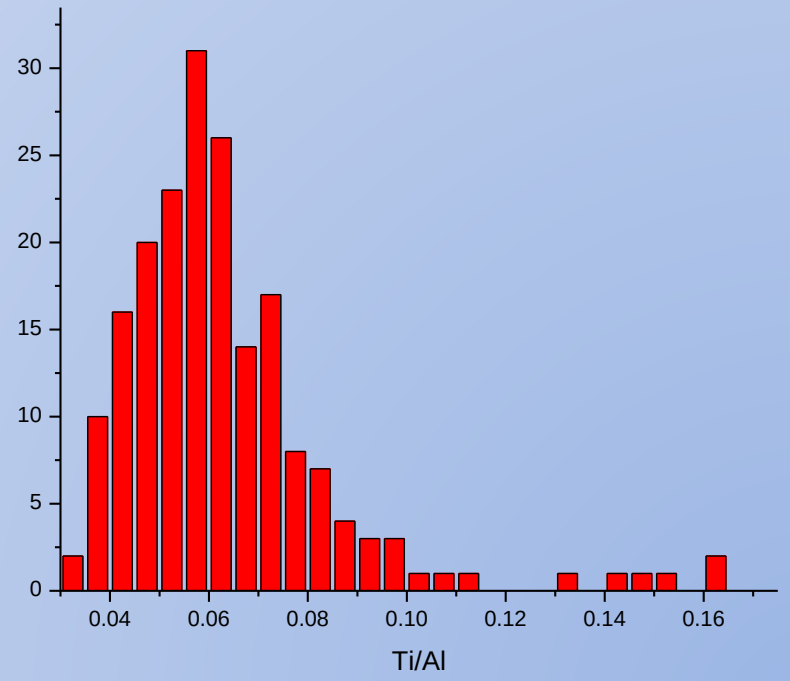


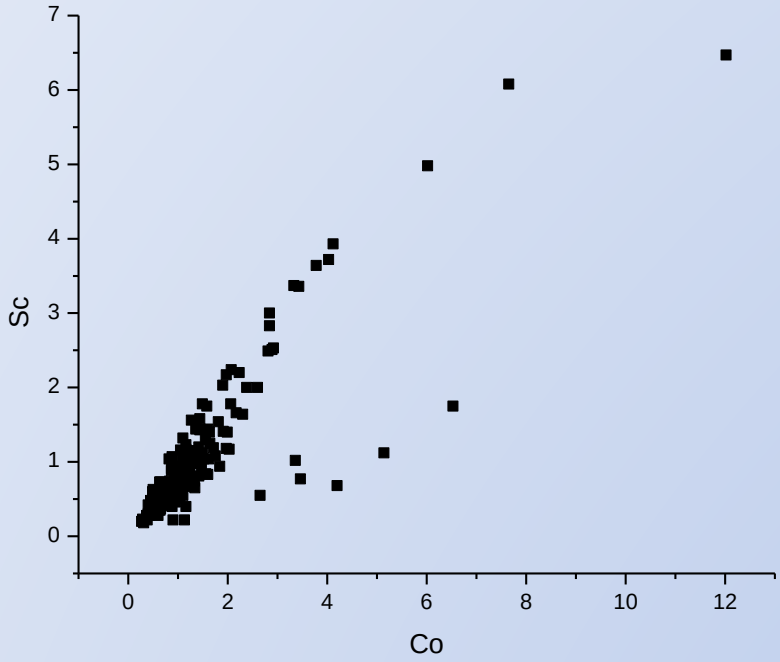
Skewness = 1.028
Kurtosis = 2.886



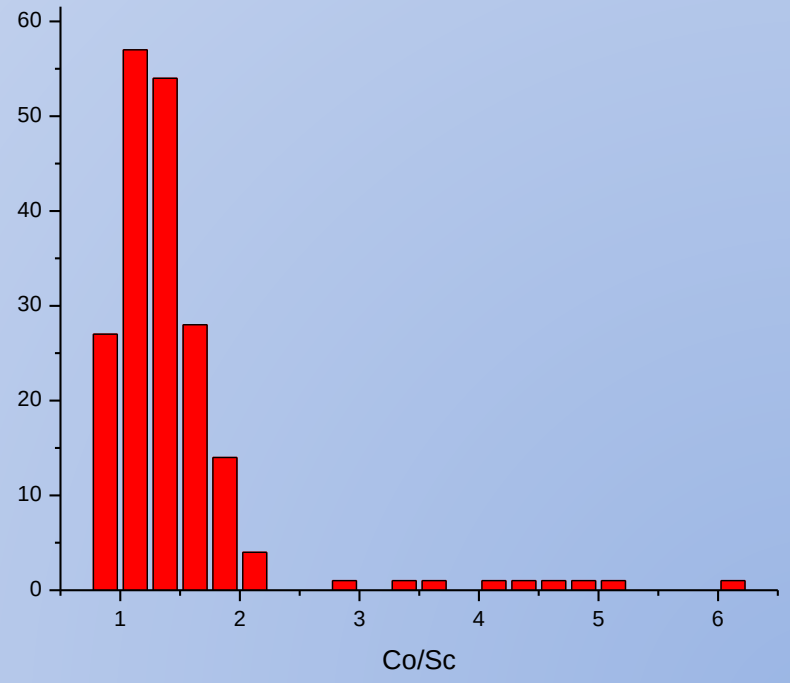


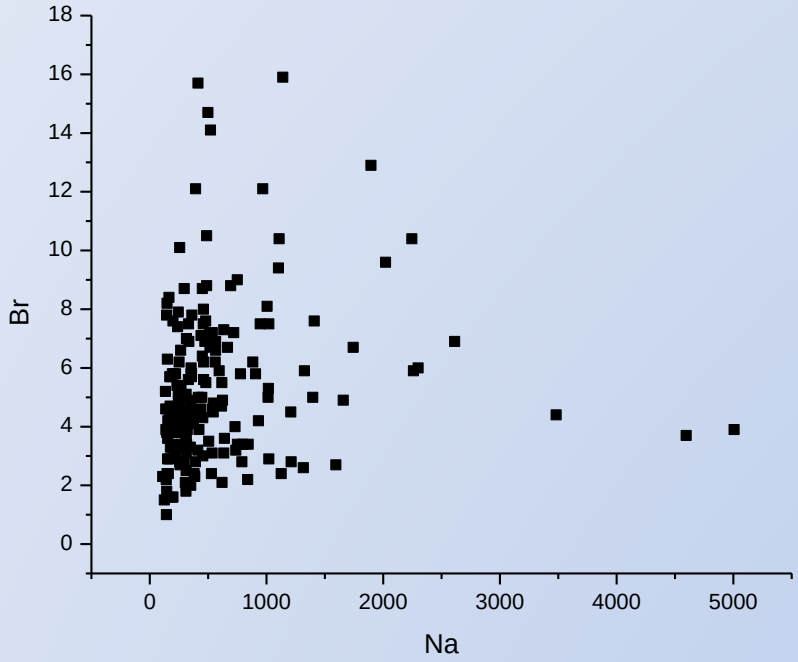
Skewness = 2.214
Kurtosis = 6.927



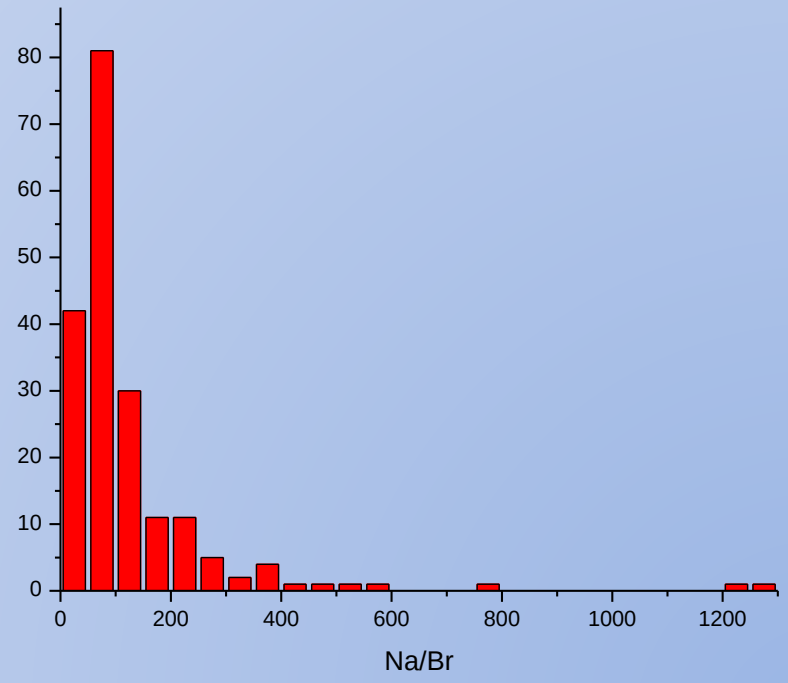


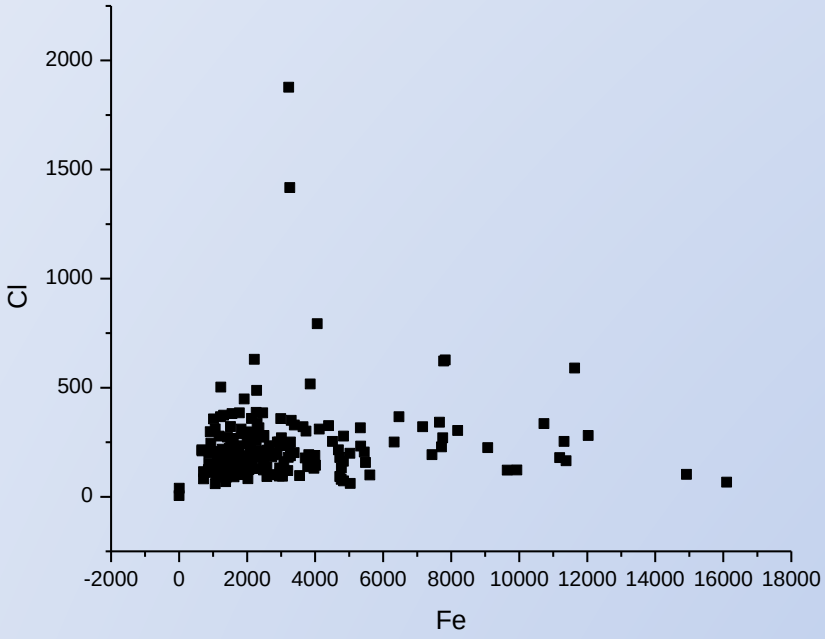
Skewness = 3.771
Kurtosis = 16.83





Skewness = 4.705
Kurtosis = 28.477

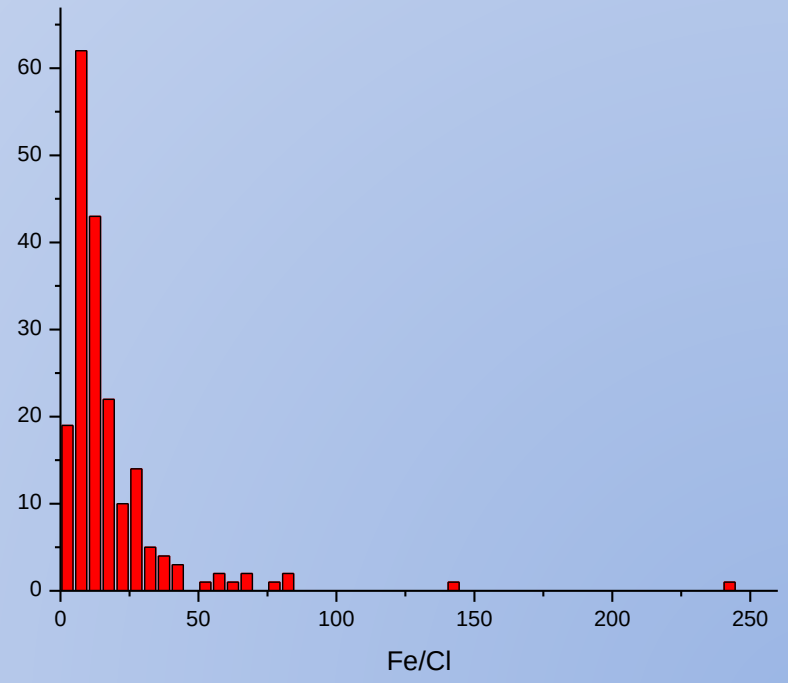


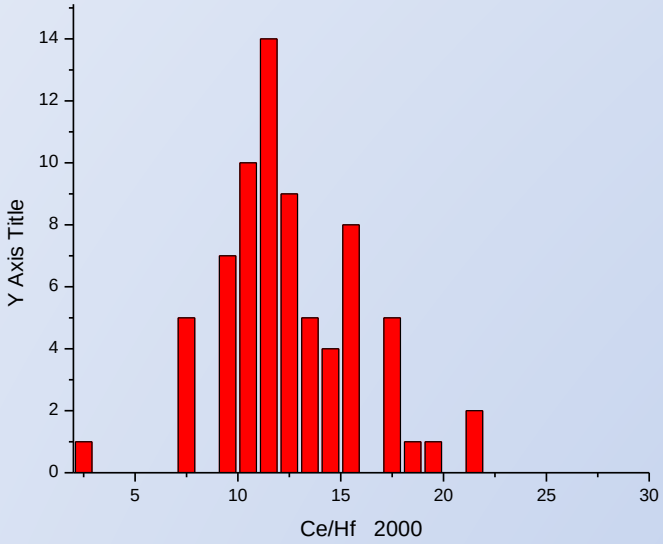


Very bad correlation – lognormal distribution of obtained ratios showing high values of skewness and kurtosis.

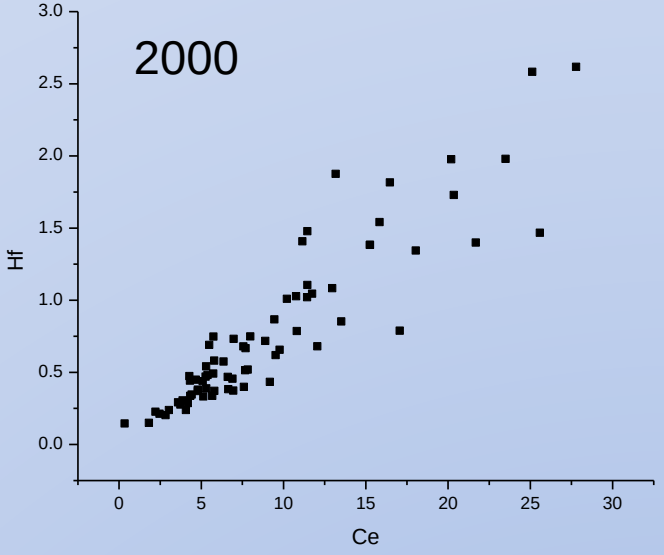
Skewness = 5.752
Kurtosis = 45.81

Skewness and kurtosis of ratio of two element concentration can be confident index of correlation.

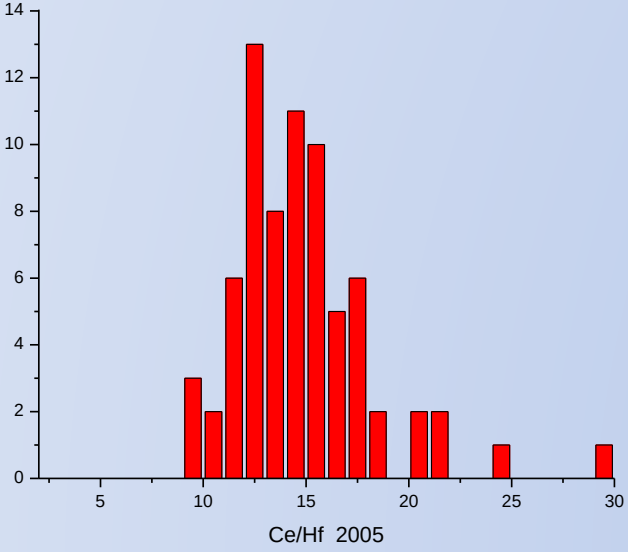




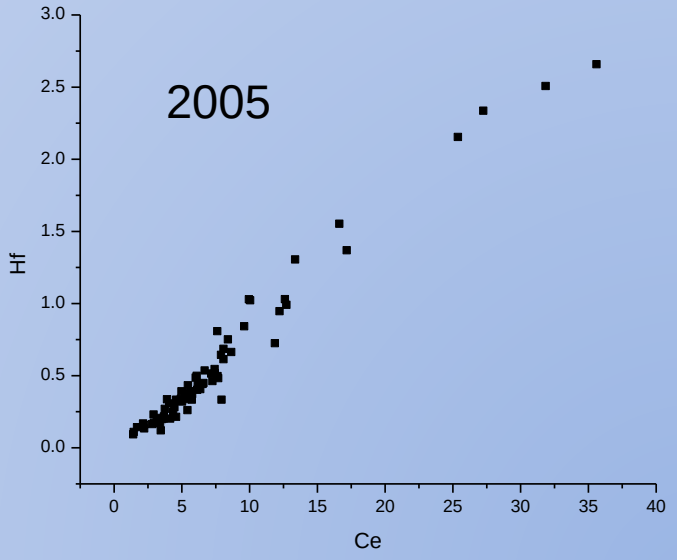
Mean = 12.6
Median = 19
Skew = 0.337
Kurt = 0.958

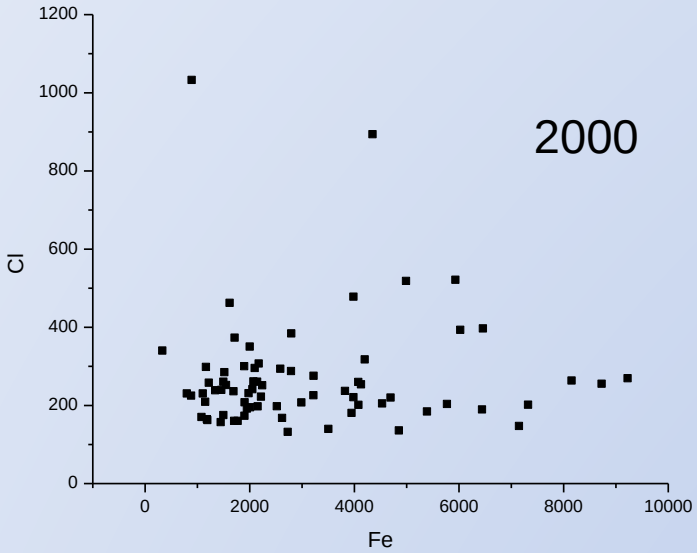


Skewness and kurtosis of ratios of two good correlated elements in two successive sampling.

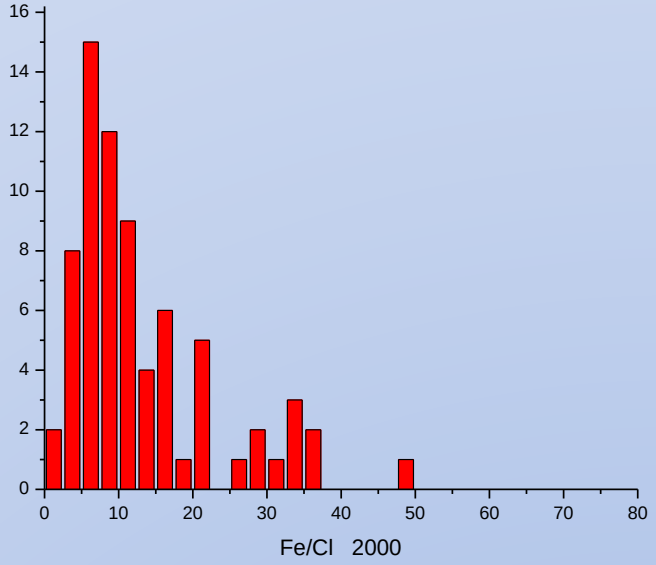


Mean = 14.8
Median = 14.3
Skew = 1.59
Kurt = 4.52

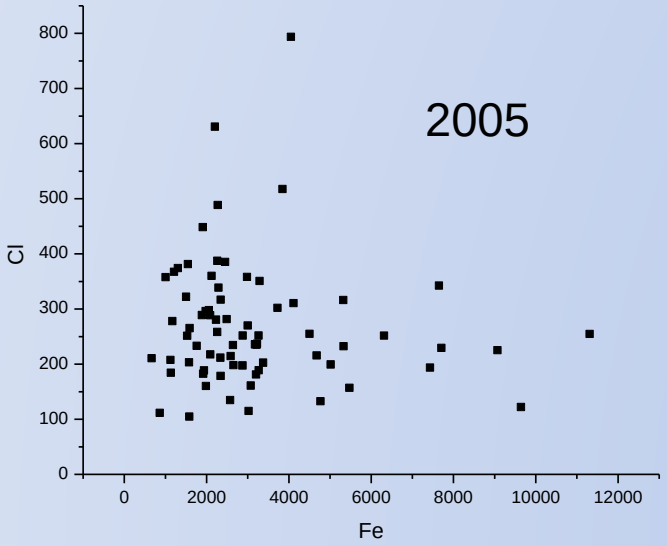




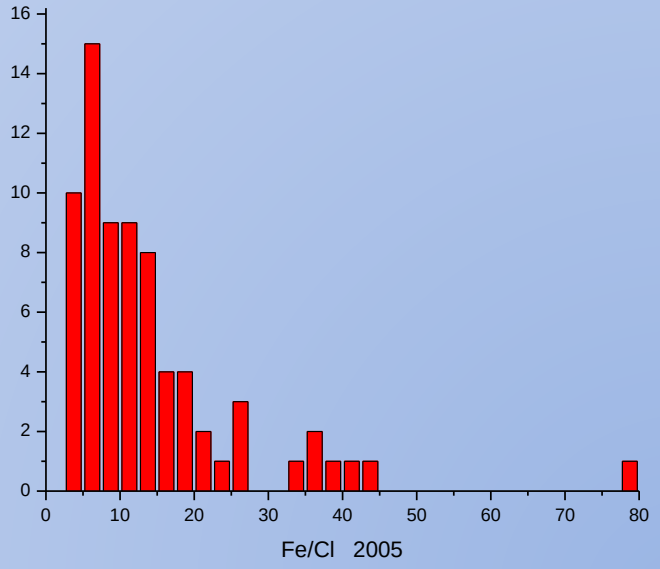
Mean = 13.2
Median = 9.70
Skew = 1.50
Kurt = 1.94

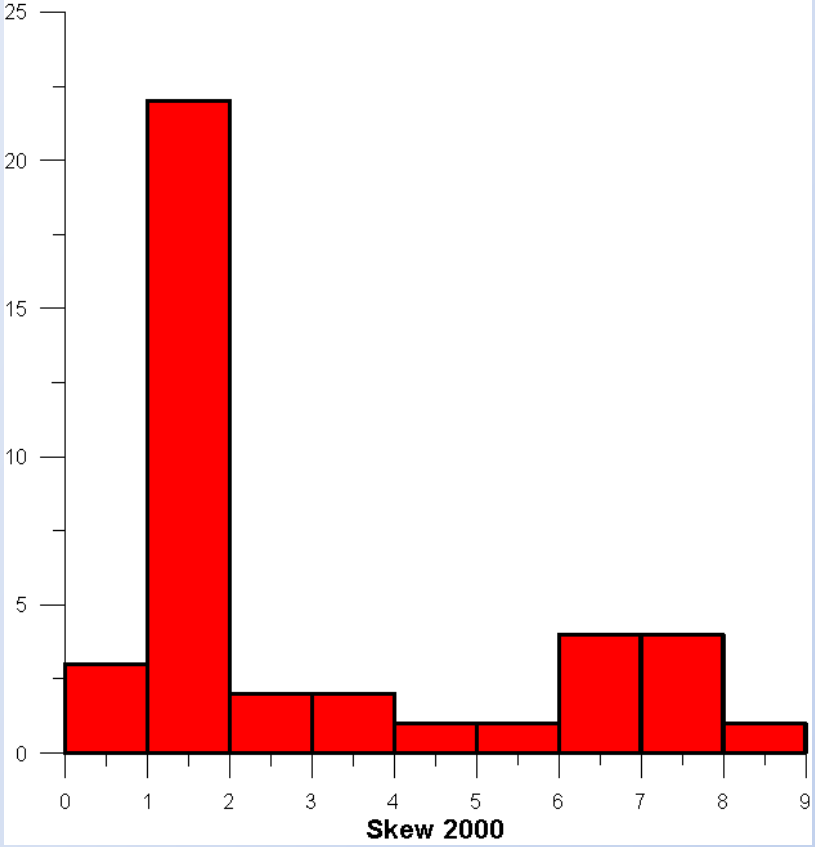


Skewness and kurtosis of ratios of two bad correlated elements in two successive sampling.



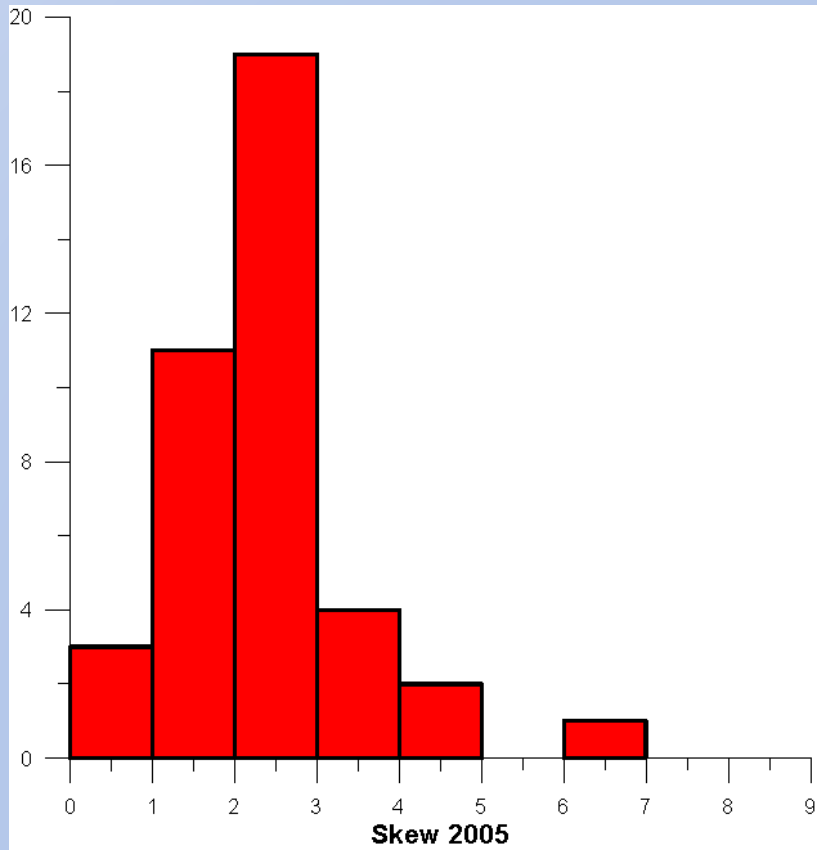
Mean = 14.1
Median = 10.8
Skew = 2.73
Kurt = 10.5

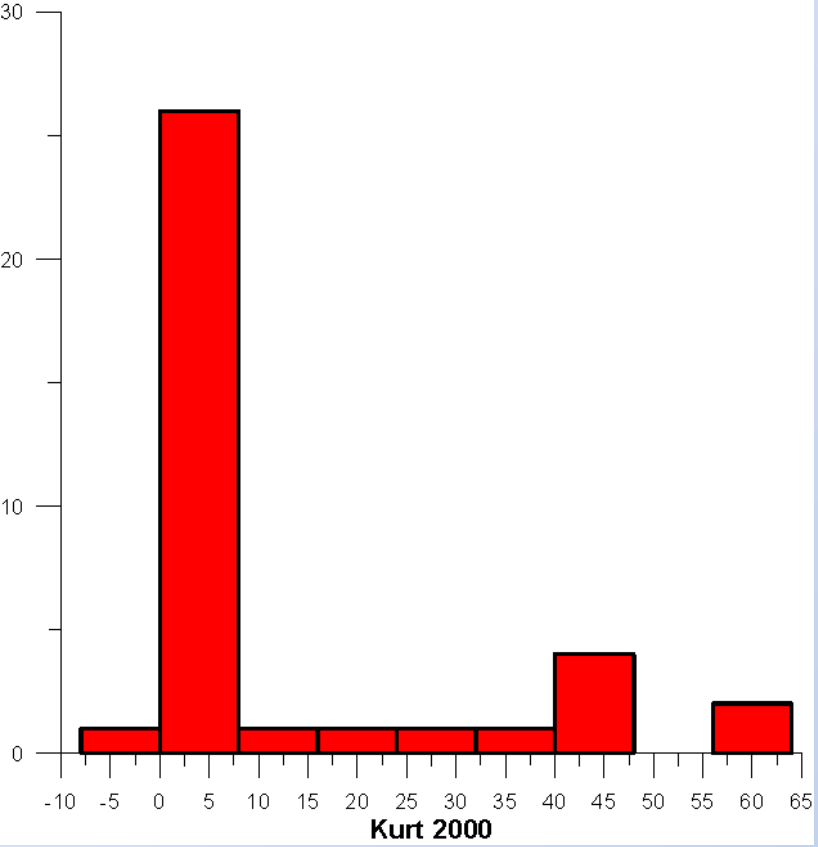




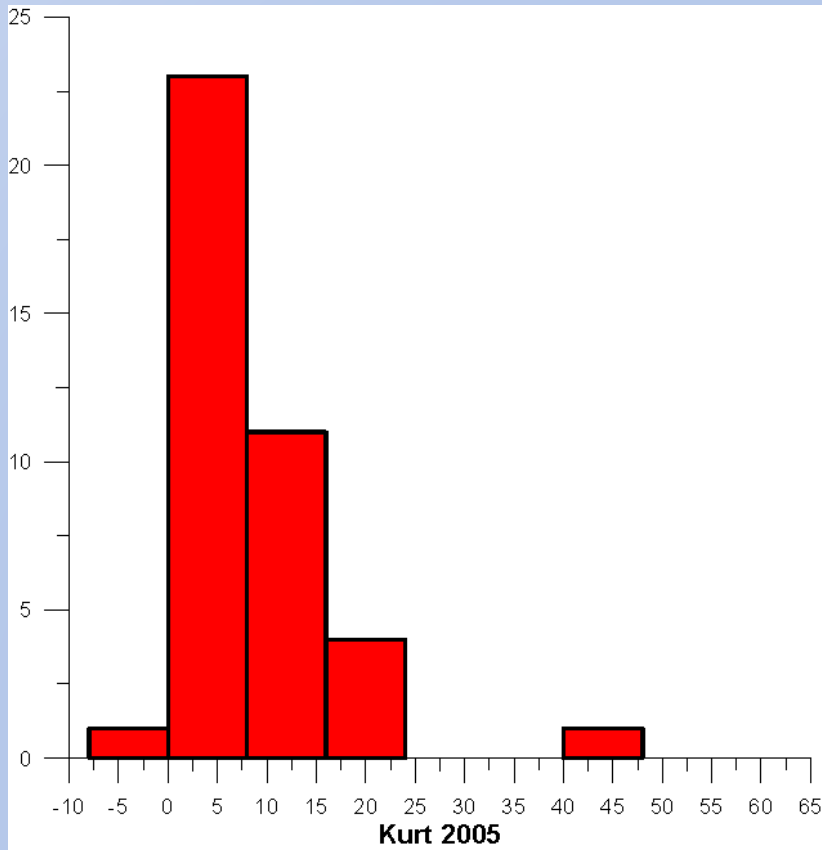
Distribution of skewness values obtained in two sets of data.

	2000	2005
Min	0.776	0.656
Max	8.10	6.42
Mean	2.96	2.44
Median	1.66	2.46





	2000	2005
Min	- 0.355	- 0.733
Max	67.5	47.7
Mean	16.34	8.39
Median	3.35	7.51



Distribution of kurtosis values obtained in two sets of data.

It is just suggestion to think together about it.

Thank you!!!!