

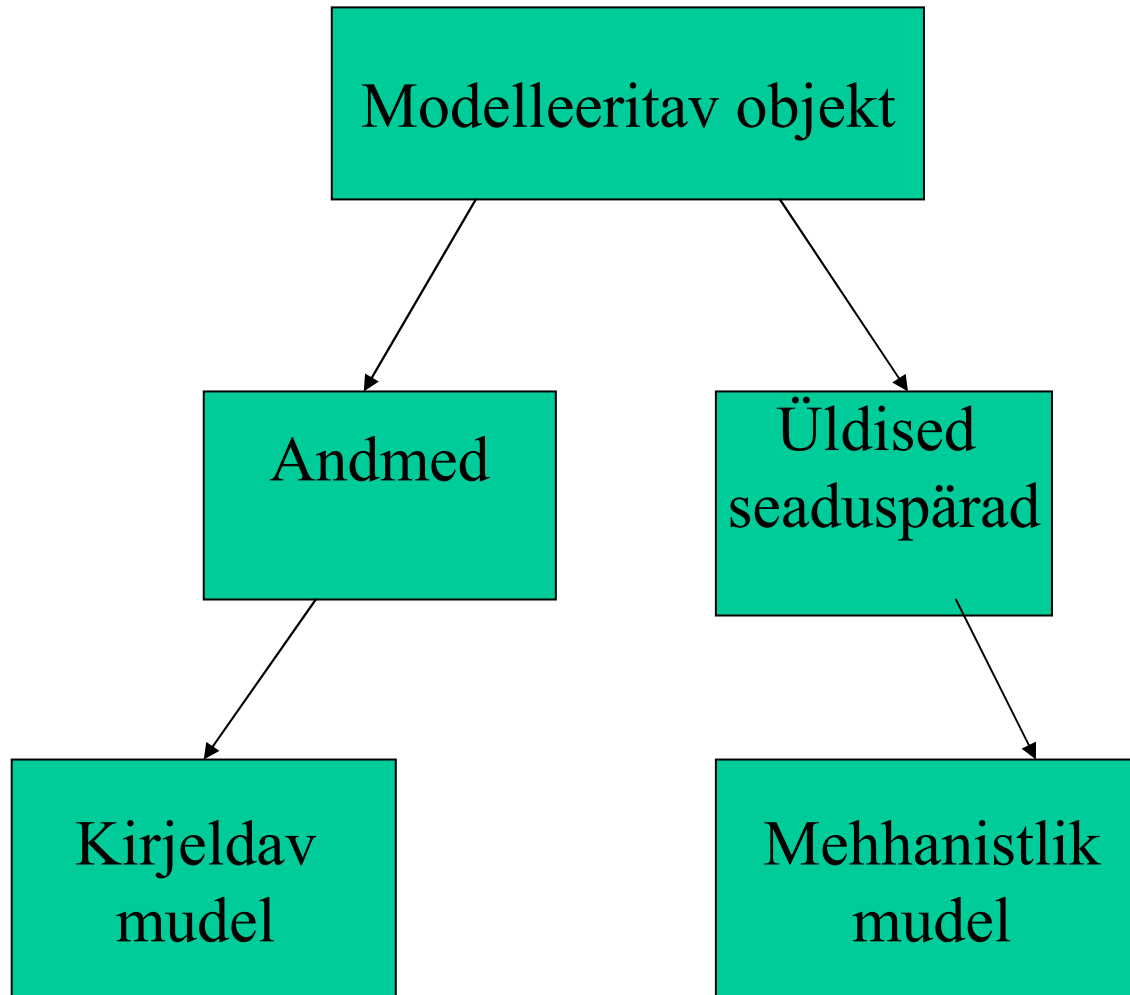
Andmed ja statistika

Margus Pihlak

Tallinna Tehnikaülikooli
matemaatikainstituut

Viitna 26.05.2011

Modelleerimise 2 strateegiat



Andmeanalüüsi 4 põhiliiki

- Regrassioonanalüüs
- Faktoranalüüs
- Dispersioonanalüüs (ingl. ANOVA)
- Diskriminantanalüüs

Üldine lineaarne mudel

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k + \varepsilon,$$

$$\varepsilon \sim N(0, \sigma)$$

Mann-Kendalli test

$$T = \sum_{j < i} \text{sgn}(Z_i - Z_j),$$

$$E(T) = 0$$

$$\text{Var}(T) = n(n-1)(2n+5)/18$$

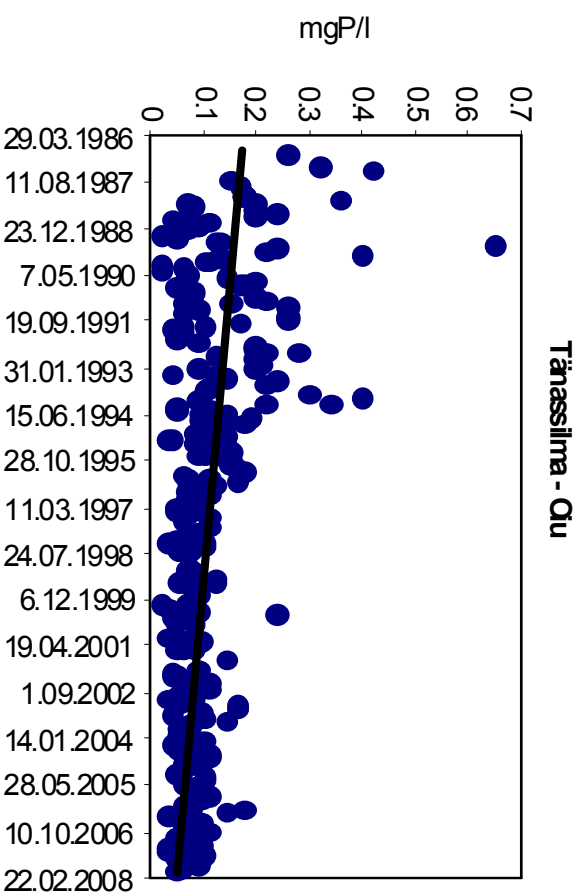
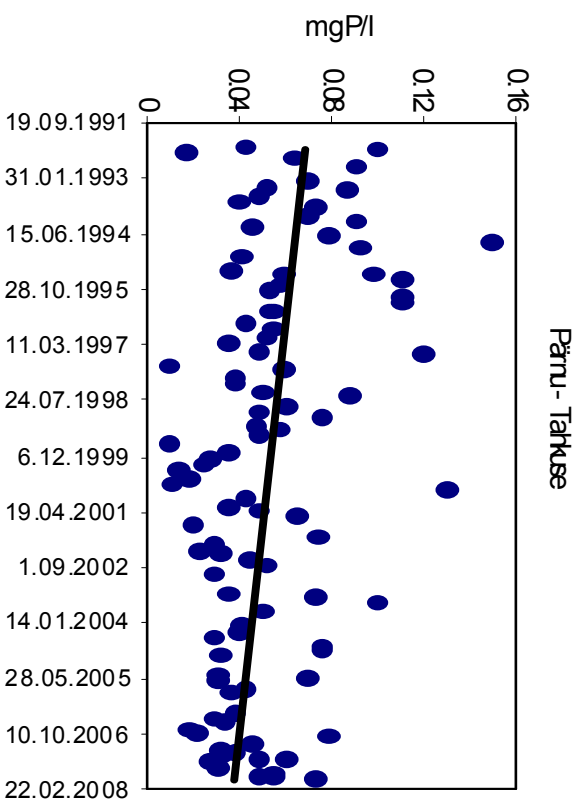
Tsentraalne piirteoreem

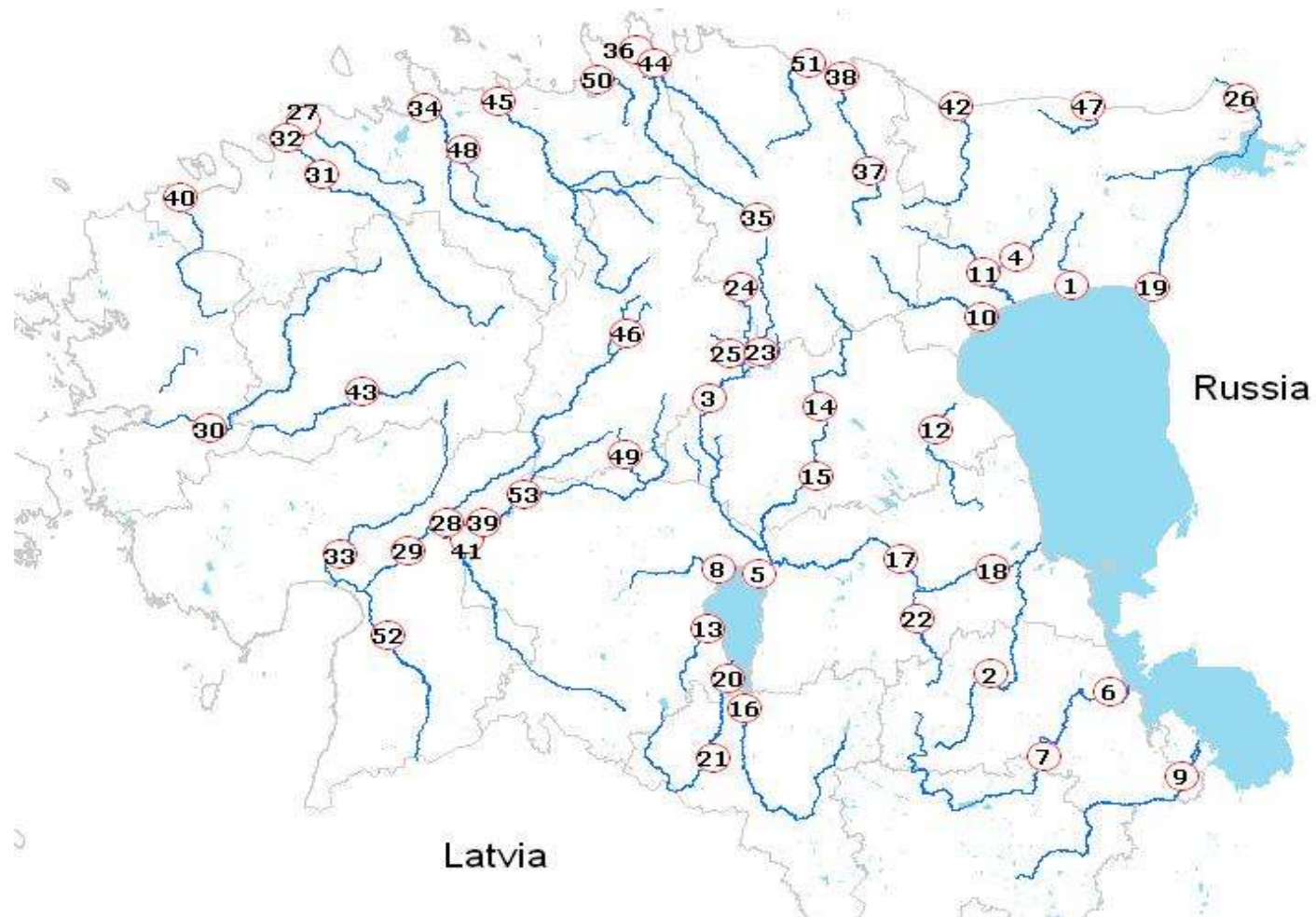
$$S_n = \frac{\sum_{i=1}^n X_i - na}{\sqrt{nb}} \sim N(0,1)$$

$$E(X) = a$$

$$\text{Var}(X) = b^2$$

Eesti jõgede fosfori koormuse trendigraafikuid





Sulfaatide faktoranalüüs, faktoreid on 2

	Faktor 1	Faktor 2
KAAPA	0.891	0.256
PIUSA	0.530	0.513
EMAJ	0.500	0.761
VEMAJ	0.711	0.619
AVIJ	0.544	
ALAJ	0.599	
RANNAP		-0.140
VOHANDU	0.488	0.749
AHJA	0.823	
NARVAN		-0.152
NARVAV	0.144	0.174
PORIJ	0.946	-0.166
PURTSE		
OHNE	0.892	

Sulfaatide faktoranalüüs, faktorite kirjeldusvõimed

	Faktor 1	Faktor 2
SS loadings	5.124	1.975
Proportion Var	0.366	0.141
Cumulative Var	0.366	0.507

Sulfaatide faktoranalüüs, faktoreid on 3

	Faktor 1	Faktor 2	Faktor 3
KAAPA	0.604	0.634	0.325
PIUSA	0.680	0.293	
EMAJ	0.898	0.182	
VEMAJ	0.861	0.353	0.176
AVIJ	0.315	0.311	0.486
ALAJ	0.336	0.312	0.645
RANNAP	-0.108	-0.108	0.472
VOHANDL	0.880	0.158	
AHJA	0.309	0.844	
NARVAN			0.183
NARVAV	0.218		
PORIJ	0.237	0.878	0.295
PURTSE			
OHNE	0.341	0.785	0.263

Sulfaatide faktoranalüüs, faktorite kirjeldusvõimed

	Faktor 1	Faktor 2	Faktor 3
SS loadings	3.701	2.994	1.211
Proportion Var	0.264	0.214	0.086
Cumulative Var	0.264	0.478	0.565

Sulfaatide faktoranalüüs, faktoreid on 4

	Faktor 1	Faktor 2	Faktor 3	Faktor 4
KAAPA	0.723	0.541	0.285	
PIUSA	0.382	0.656		
EMAJ	0.285	0.844	-0.106	0.217
VEMAJ	0.466	0.810	0.130	
AVIJ	0.367	0.282	0.452	
ALAJ	0.358	0.287	0.683	0.223
RANNAP			0.452	-0.146
VOHANDU	0.262	0.834		0.175
AHJA	0.869	0.194		
NARVAN		-0.103	0.180	
NARVAV				0.920
PORIJ	0.909	0.128	0.249	
PURTSE				
OHNE	0.832	0.224	0.228	0.160

Sulfaatide faktoranalüüs, faktorite kirjeldusvõimed

	Faktor 1	Faktor 2	Faktor 3	Faktor 4
SS loadings	3.594	3.083	1.141	1.045
Proportion Var	0.257	0.220	0.081	0.075
Cumulative Var	0.257	0.477	0.558	0.633

			MK-stat	P-value	MK-stat	P-value
Num.	Station	Period	Ntot	Ntot	Ptot	Ptot
1	Alajõgi-Alajõe	1984-2006	-2,74	0,0061	2,74	0,0061
2	Ahja-Kiidjärve	1985-2006	-1,41	0,1574	-0,24	0,8117
3	Põltsamaa-Rutikvere	1986-2006	-2,80	0,0050	-1,30	0,1940
4	Rannapungerja-Roostoja	1992-2006	-2,69	0,0072	2,04	0,0415
5	Emajõgi-Jõesuu	1987-2006	-0,05	0,9602	-1,23	0,2191
6	Võhandu-Räpina	1986-2006	-2,04	0,0411	-0,88	0,3769
7	Võhandu-Himmiste	1991-2006	-2,46	0,0140	-2,02	0,0434
8	Tänassilma-Oiu	1986-2006	0,11	0,9158	-2,69	0,0072
9	Piusa-Korela	1992-2006	-2,72	0,0065	0,35	0,7268
10	Avijõgi-Mulgi	1992-2006	-2,67	0,0077	0,34	0,7304
11	Tagajõgi-Tudulinna	1992-2006	-3,18	0,0014	0,82	0,4128
12	Kääpa-Kose	1986-2006	-0,66	0,5064	-0,69	0,4890
13	Tarvastu-Põdraoja	1986-2006	-3,11	0,0018	-2,73	0,0064
14	Pedja-Jõgeva	1986-2006	-0,90	0,3703	1,32	0,1875
15	Pedja-Tõrve	1986-2006	-0,05	0,9625	0,62	0,5348
16	Väike-Emajõgi-Pikasilla	1986-2006	-2,58	0,0099	-2,91	0,0036
17	Emajõgi-Tartu	1984-2006	0,37	0,7077	0,75	0,4543
18	Emajõgi-Kavastu	1986-2006	-0,55	0,5834	-1,42	0,1553
19	Narva-Vasknarva	1992-2006	-0,41	0,6835	0,46	0,6437
20	Õhne-Suislepa	1986-2006	-2,34	0,0195	1,49	0,1371
21	Õhne-Roobe	1991-2006	-2,44	0,0146	1,85	0,0642
22	Porijõgi-Reola	1992-2006	-2,04	0,0409	-1,73	0,0841

	<i>BHT7</i>	<i>Ntot</i>	<i>Ptot</i>	<i>sulfaadid</i>
<i>BHT7</i>	1			
<i>Ntot</i>	-0,10815	1		
<i>Ptot</i>	0,12772	0,204917	1	
<i>sulfaadid</i>	0,161766	0,114066	0,136304	1

Täna tähelepanu eest!