



# Tvorba modelu sorpce a desorpce $^{85}\text{Sr}$ na krystalických horninách za dynamických podmínek metodou nelineární regrese

Závěrečná práce 12. licenčního studia Pythagoras

Fakulta chemicko-technologická, katedra analytické chemie,  
Univerzita Pardubice

Ján Lengyel  
27. říjen 2010

[jan.lengyel@ujv.cz](mailto:jan.lengyel@ujv.cz)



# Popis problému

- Radionuklid  $^{90}\text{Sr}$  patří mezi významné dlouhodobé štěpné nebo aktivační produkty uranu, které jsou obsaženy i v radioaktivních odpadech.
  - Vzhledem k dlouhému poločasu jaderné přeměny radionuklidu  $^{90}\text{Sr}$ ,  $T_{1/2} = 28,8$  roků, byl pro experimenty použitý krátkodobý radioizotop  $^{85}\text{Sr}$ ,  $T_{1/2} = 64,85$  dnů.
  
- Krystalické horniny pocházející ze středočeského plutonu jsou typické pro dlouhodobé hlubinné úložiště radioaktivních odpadů. Jsou to:
  - Diorit I,
  - Diorit II,
  - Gabro,
  - Granit,
  - Tonalit.



# Popis problému

- ❑ Po dlouhodobém uložení radioaktivních odpadů v řádu stovek let , existuje reálné riziko uvolnění radionuklidů do životního prostředí.
- ❑ Pro model sorpce a desorpce Sr na zmíněných krystalických horninách, za dynamických podmínek, a určení jejich parametrů byla použita statistická metoda nelineární regrese.

# Cíl práce

- Zpracovat experimentální data statistickou metodou nelineární regrese pomocí publikovaného modelu.
  - **Publikovaný model je vyjádřený rovnicí (rov.1):**

$$A_{rel} = R' \exp \left[ \left( R - n_{pv} \right) / \left( R n_{pv} / Pe \right) \left( \pi R n_{pv} / Pe \right)^{0,5} \right]$$

- Porovnat výsledky získané pomocí nelineární regrese s publikovanými výsledky.



# Cíl práce

- ❑  $R = 1 + (\rho \cdot K_d / \theta)$  je **retardační koeficient**,  $\rho$  je sytná váha a  $\theta$  je porosita pevné fáze,  $K_d$  je rozdělovací koeficient radionuklidu mezi pevnou a kapalnou fází,
- ❑  $n_{pv}$  je **počet pórových objemů**,
- ❑  $Pe = v_s \cdot L / D_d$  je **Pecletovo číslo**,  $v_s$  je axiální prosakovací rychlost kapalně fáze a  $D_d$  je hydrodynamický dispersní koeficient
- ❑ a současně  $R' = d_p / L$ , kde  $d_p$  je délka pulzu a  $L$  je délka kolony.  
V ideálním případě  **$R' = R$** .



# Použité programy

## ❑ MS Excel verze 2007

- Pro tabulkové a grafické znázornění průběhu závislosti závislé proměnné  $A_{rel}$  na nezávislé proměnné  $n_{pv}$  a pro export dat do statistického softwaru.

## ❑ QC.Expert verze 3.1

- Pro statistické zpracování experimentálních dat metodou nelineární regrese.

## ❑ Statistica verze 6

- Pro ověření výsledků z programu QC.Expert pro krystalickou horninu Diorit I.



# Nelineární regrese

- Návrh regresního modelu a odhadů parametrů.
- Odhadování parametrů
- Posouzení kvality odhadů parametrů
- Grafické posouzení vhodnosti modelu
- Základní statistické charakteristiky
- Regresní diagnostika
- Predikční schopnost modelu
- Souhlas s požadavky fyzikálního smyslu



# Návrh regresního modelu

## □ Regresní model

- Rovnici popisující sorpci a desorpci  $S_r$  na krystalických horninách, za dynamických podmínek, lze napsat do programu QC.Expert ve tvaru:

$$Arel \sim p1 * Exp(-(p2 - (npv))^2 / (4 * p2 * (npv) / p3)) / sqrt(4 * 3.14 * p2 * (npv) / p3)$$

- parametr  $p1$  představuje  $R'$  z rovnice (rov.1),
- parametr  $p2$  představuje  $R$  z rovnice (rov.1),
- a parametr  $p3$  představuje  $Pe$  (Pecletovo číslo) z rovnice (rov.1).



# Odhad parametrů

- U metody nelineární regrese je nutné zadávat první odhady parametrů.
  - V tabulce jsou uvedeny prvotní odhady parametrů  $p_1$ ,  $p_2$  a  $p_3$ , příslušná spočítaná minima reziduálních součtů čtverců a počet provedených iterací z QC.Expert a prvotní odhady pro program STATSTICA.

	QC.Expert					STATISTICA		
	p1	p2	p3	Minimum RSC	Počet iterací	b1	b2	b3
Hornina								
Diorit I	70	27	10	0,0118	15	100	30	10
Diorit II	50	20	10	0,0241	20	50	20	7
Gabro	40	15	10	0,0270	14	40	15	10
Granit	15	5	5	0,0739	1	15	5	5
Tonalit	40	12	7	0,0160	22	40	20	10

# Posouzení kvality odhadů parametrů

- Pro výpočet optimálních hodnot parametrů byla použita Gauss-Newtonová metoda, založená na klasickém derivačním algoritmu linearizace modelu v blízkosti parametrů.
  - **Kvalita odhadů parametrů je daná intervaly spolehlivosti odhadů a rozptylem, resp. směrodatnou odchylkou a korelační maticí parametrů modelu.**

Diorit I				
Parametr	Odhady	SD	Dolní mez	Horní mez
p1	81,35	0,77	79,76	82,94
p2	73,11	0,34	72,41	73,82
p3	9,44	0,19	9,05	9,83

# Posouzení kvality odhadů parametrů

- Pro srovnání v tabulce jsou výpočtené optimální hodnoty parametrů pro Diorit I, pomocí programu STATISTICA.
  - Hodnoty parametrů, v programu STATISTICA označené písmenem b, jsou prakticky shodné s hodnotami parametrů z programu QC.Expert.

**Model:  $A_{rel} = b_1 \cdot \text{Exp}(-b_2 \cdot npv) / \sqrt{4 \cdot b_2 \cdot npv / b_3}$**

Parametr	Odhad	SD	Dolní mez	Horní mez
<b>b1</b>	81,37	0,77	79,78	82,96
<b>b2</b>	73,11	0,34	72,41	73,82
<b>b3</b>	9,44	0,19	9,05	9,83



# Posouzení kvality odhadů parametrů

Diorit II				
Parametr	Odhady	SD	Dolní mez	Horní mez
p1	72,64	2,34	67,69	77,59
p2	61,11	0,98	59,02	63,19
p3	6,71	0,45	5,75	7,66

Gabro				
Parametr	Odhady	SD	Dolní mez	Horní mez
p1	45,21	1,70	41,54	48,87
p2	37,53	0,84	35,72	39,34
p3	6,51	0,51	5,42	7,61



# Posouzení kvality odhadů parametrů

Granit				
Parametr	Odhady	SD	Dolní mez	Horní mez
p1	24,20	1,80	20,27	28,13
p2	18,65	0,78	16,94	20,36
p3	5,04	0,67	3,59	6,49

Tonalit				
Parametr	Odhady	SD	Dolní mez	Horní mez
p1	43,02	1,22	40,39	45,64
p2	41,14	0,63	39,80	42,48
p3	9,25	0,53	8,11	10,39

# Posouzení kvality odhadů parametrů

## – Korelační matice parametrů

Diorit I	p1	p2	p3
p1	1	0,39	-0,69
p2	0,39	1	-0,22
p3	-0,69	-0,22	1

Diorit II	p1	p2	p3
p1	1	0,52	-0,73
p2	0,52	1	-0,14
p3	-0,73	-0,14	1

Gabro	p1	p2	p3
p1	1	0,43	-0,72
p2	0,43	1	-0,31
p3	-0,72	-0,31	1

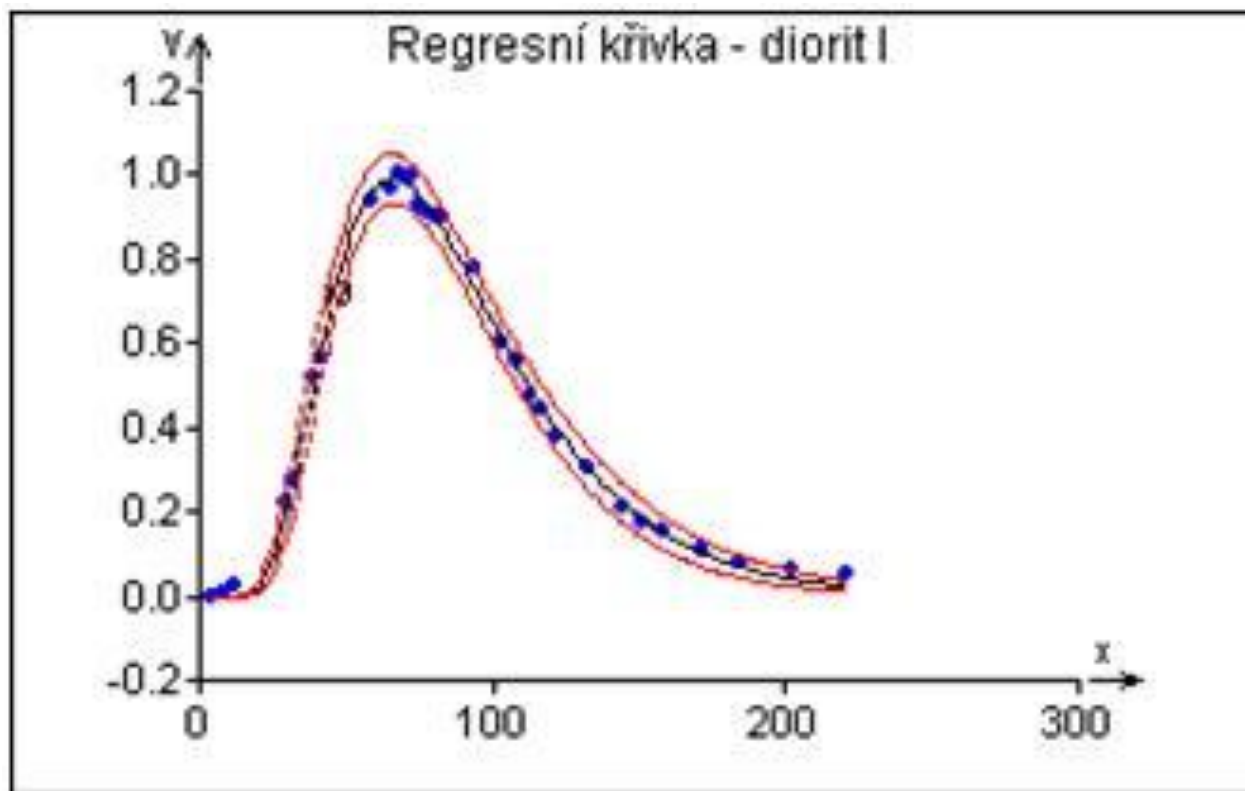
# Posouzení kvality odhadů parametrů

Granit	p1	p2	p3
p1	1	0,62	-0,41
p2	0,62	1	-0,30
p3	-0,41	-0,30	1

Tonalit	p1	p2	p3
p1	1	0,31	-0,75
p2	0,31	1	-0,17
p3	-0,75	-0,17	1

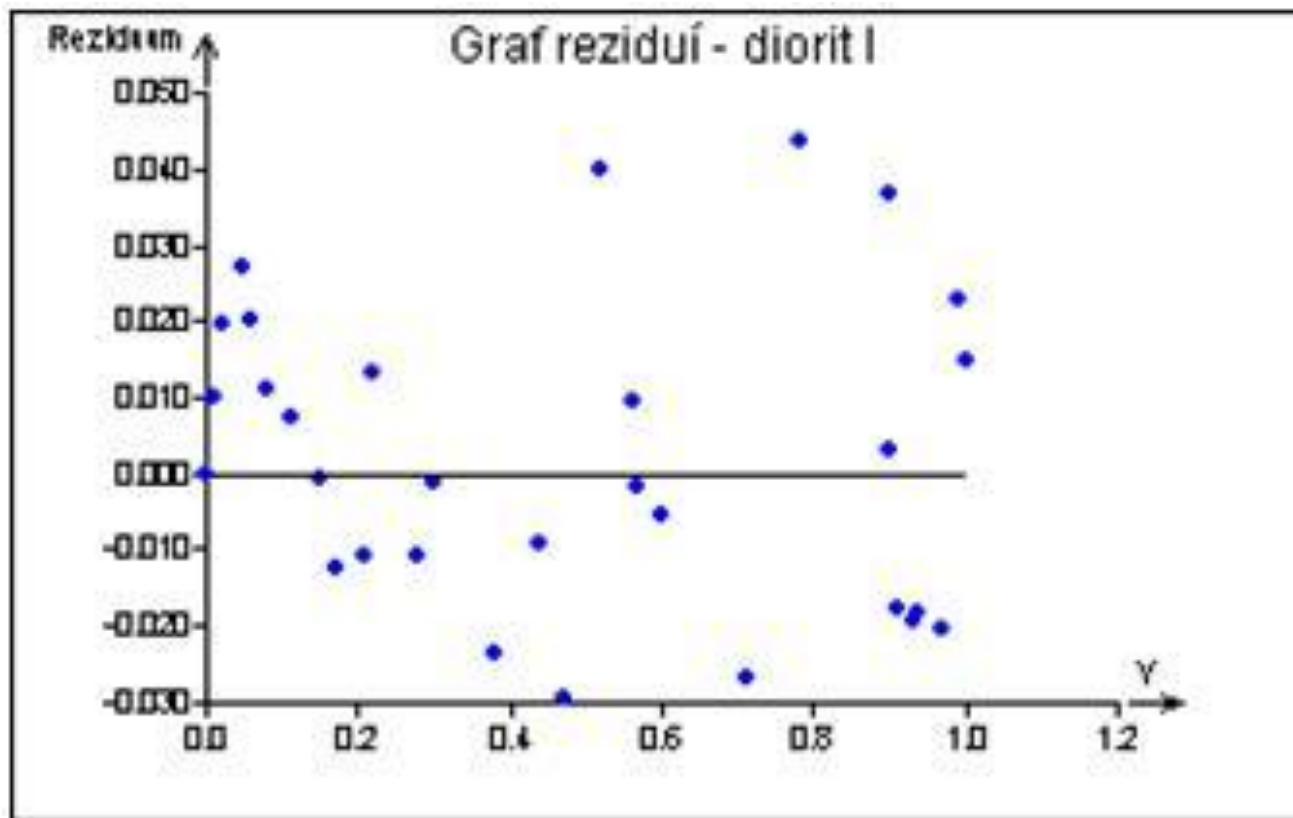
- Intervaly spolehlivosti jsou úzké a i směrodatné odchylky jsou malé.
- Hodnoty korelačních koeficientů mimo diagonálu nejsou velmi blízké +1, resp. - 1, model není třeba re-parametrizovat (zapsat v jiném algebraickém tvaru).

- Graf y regresní křivky s 95 % pásy spolehlivosti a grafy normovaných reziduí slouží k posouzení vhodnosti modelu.
  - Těsné proložení regresní křivky a úzký pás spolehlivosti
  - a v grafu normovaných reziduí body tvoří „mrak“.



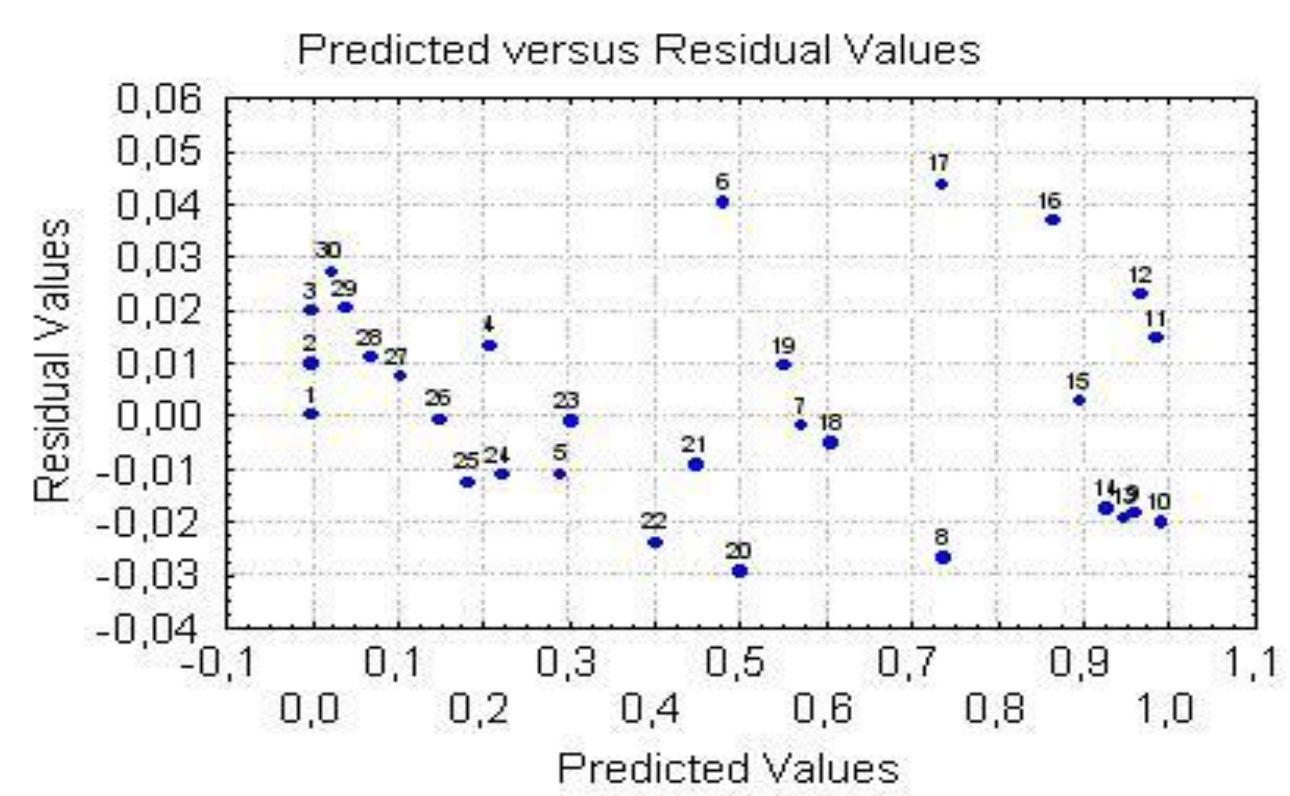


# Grafické posouzení vhodnosti modelu



- Jak uvedené grafy, tak hodnoty šikmosti  $g_1$  a špičatosti  $g_2$  (uvedené dále) svědčí o normalitě rozdělení reziduí a o vhodnosti modelu.

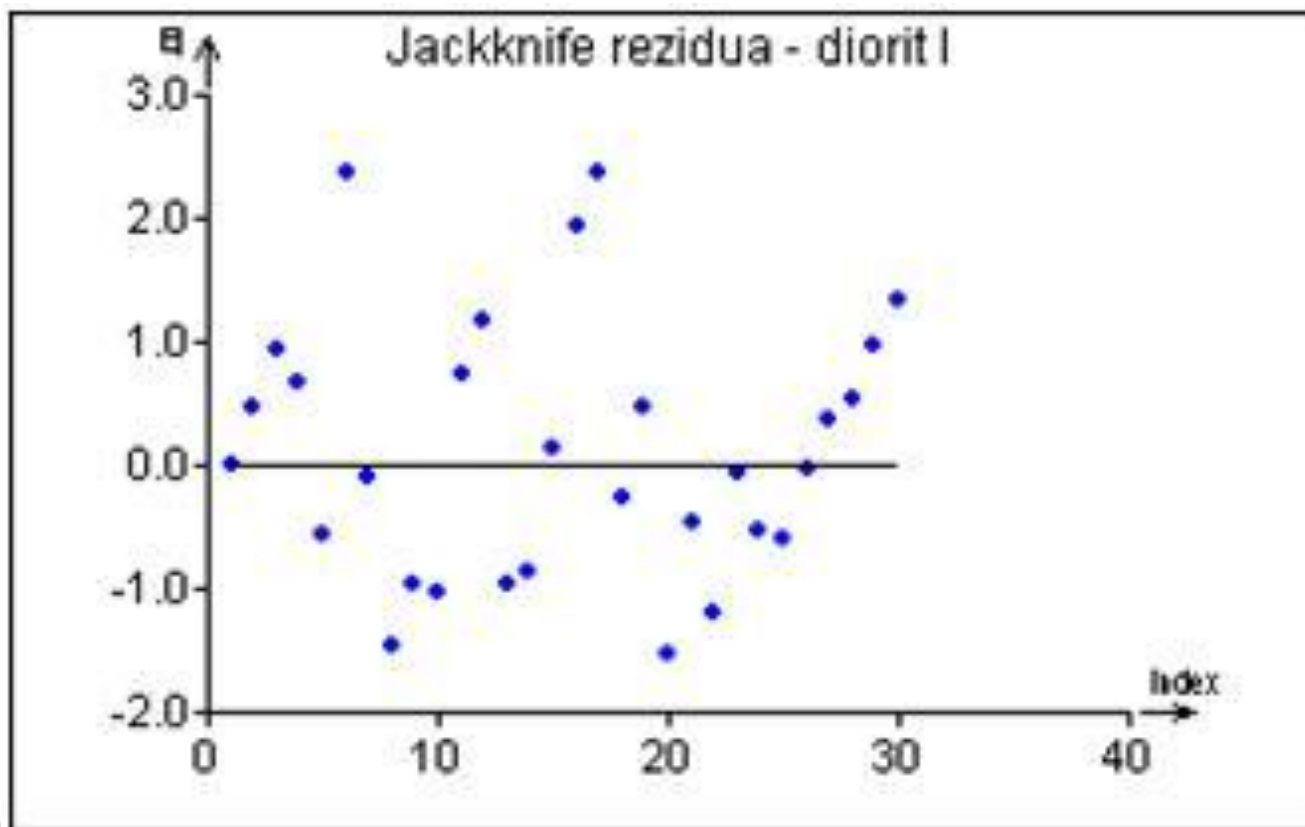
# Grafické posouzení vhodnosti modelu



- Pro srovnání je zde uvedený identický graf reziduí pro Diorit I z programu STATISTICA..

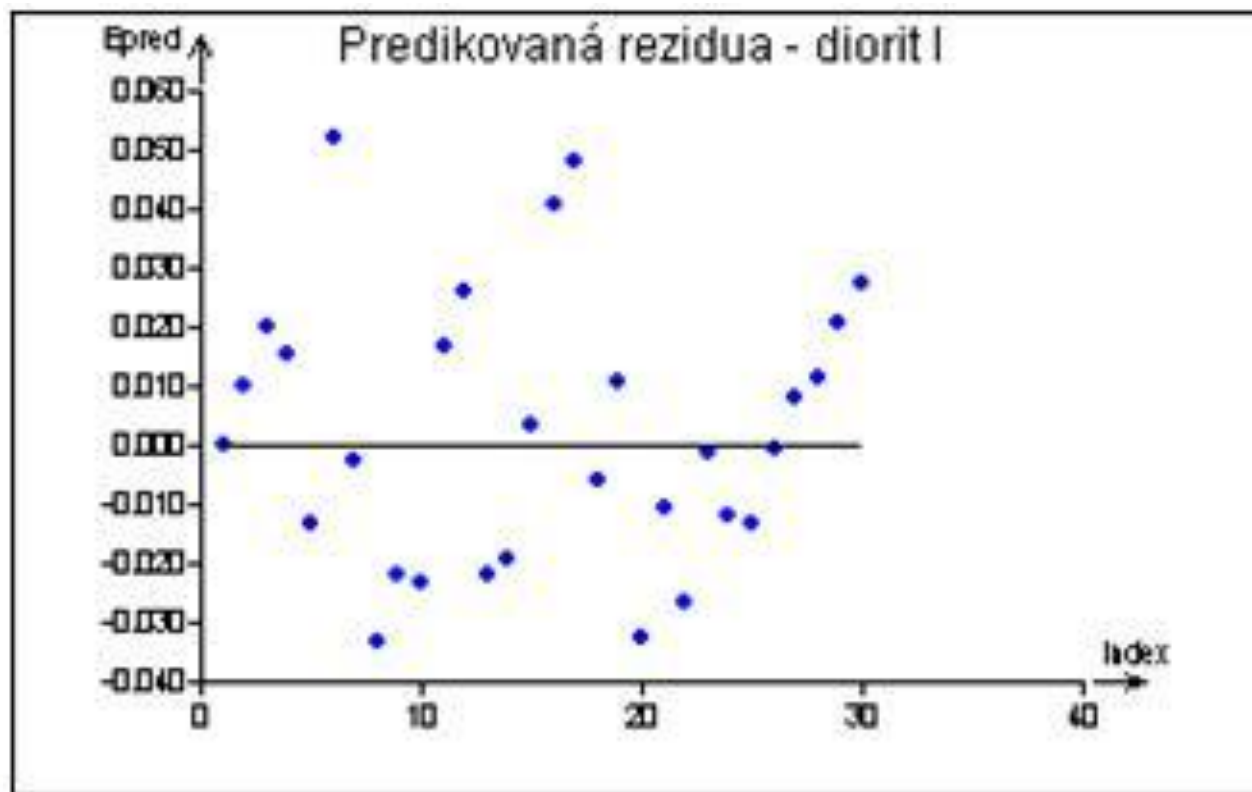
## □ Graf Jackknife reziduí

- Jackknife rezidua, se počítají tak, že pro  $i$ -tý bod je použita směrodatná odchylka rezidua získaná vynecháním  $i$ -tého bodu. Grafy Jackknife reziduí graficky indikují vybočující body.



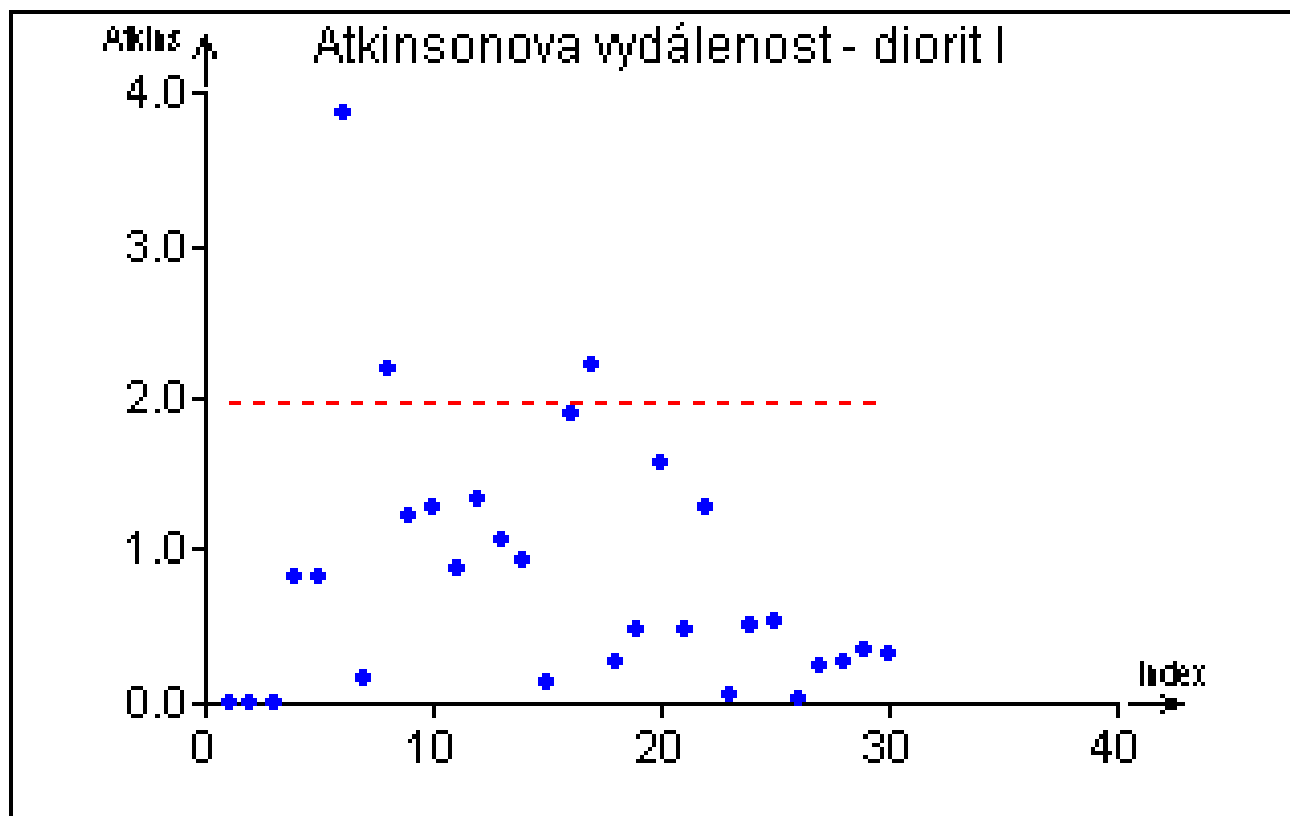
## □ Graf predikovaných reziduí

- **Predikovaná rezidua, se počítají jako rozdíl  $i$ -té hodnoty nezávisle proměnné od modelu získaného po vynechání  $i$ -tého bodu. Toto reziduum citlivěji indikuje vybočující body.**



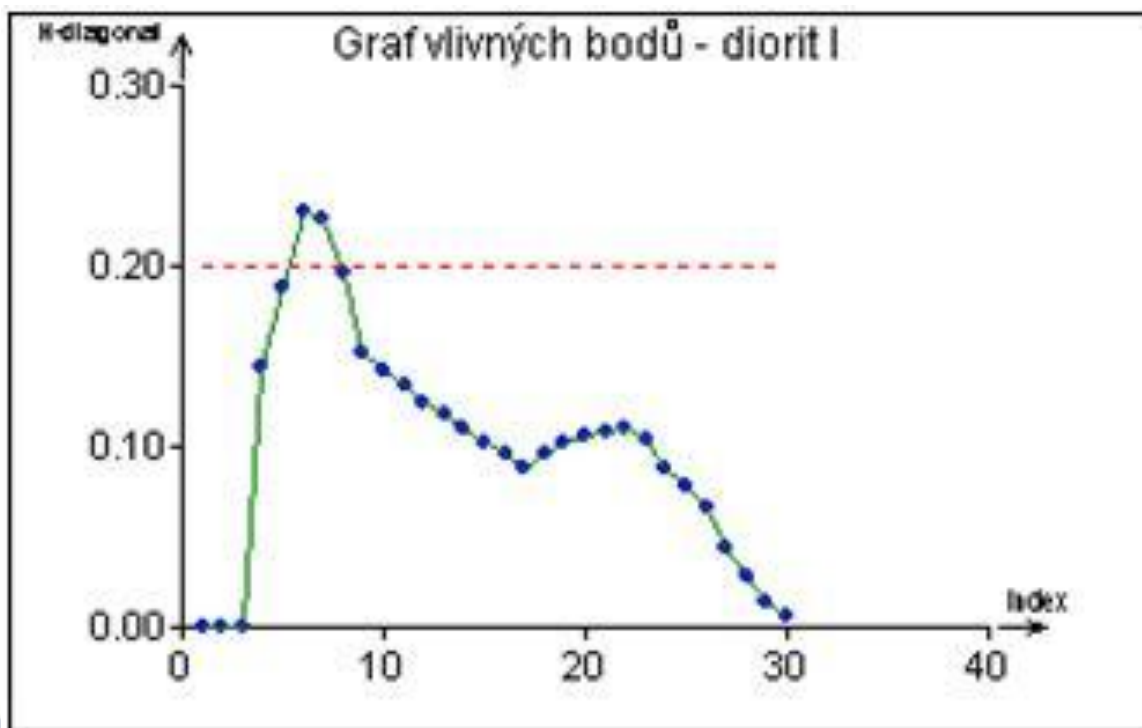
## □ Graf Atkinsonových vzdáleností

- **Graf Atkinsonových vzdáleností vyjadřuje vliv dat na parametry modelu. Data nad vodorovnou přímkou indikují silně vlivná data.**



## □ Graf vlivných bodů

- **Graf vlivných bodů - diagonální prvky projekční matice  $H=X(X^T X)^{-1} X^T$ , které vyjadřují míru vlivu jednotlivých dat na regresi (X je matice prvních parciálních derivací modelu podle jednotlivých parametrů v jednotlivých hodnotách nezávisle proměnné). Body nad vodorovnou přímkou se považují za silně vlivné.**





# Grafické posouzení vhodnosti modelu

- Grafická analýza reziduí vykazuje přítomnost vlivných bodů v maximu regresních závislostí a jeho okolí pro sledované krystalické horniny.
  - Vzhledem k malému počtu experimentálních bodů kolem maxim regresních křivek nelze je vyloučit ze statistického hodnocení.

- V tabulce jsou uvedeny základní statistické charakteristiky, z kterých stonásobek koeficientu determinace – regresní rabat, šikmost a špičatost, střední kvadratická chyba predikce – MEP, a Akaikeho informační kritérium – AIC, určují dobrou podmíněnost parametrů navrženým modelem.

Statistické charakteristiky regrese	Diorit I	Diorit II	Gabro	Granit	Tonalit
Vícenásobný korel. koef. R:	0,9984	0,9949	0,9915	0,9631	0,9960
Koeficient determinace R <sup>2</sup> :	0,9968	0,9898	0,9830	0,9276	0,9920
Predikovaný korel. koef. Rp :	0,9958	0,9803	0,9624	0,7872	0,9778
Stř. kvadratická chyba predikce MEP:	0,0005	0,0024	0,0037	0,0145	0,0026
Akaikeho informační kritérium:	-229,2477	-120,7621	-96,1302	-73,6889	-112,4116
Reziduální součet čtverců:	0,0118	0,0241	0,0270	0,0739	0,0160
Průměr absolutních reziduí:	0,4876	0,4692	0,4699	0,8037	0,4289
Reziduální směr. odchylka:	0,0209	0,0388	0,0456	0,0785	0,0339
Reziduální rozptyl:	0,0004	0,0015	0,0021	0,0062	0,0011
Šikmost reziduí:	0,4883	1,1411	1,6316	0,5158	1,1184
Špičatost reziduí:	2,4742	4,1959	4,2597	2,9025	3,5548



- V tabulce jsou statistické charakteristiky heteroskedasticity a normality reziduí.

Cook-Weisbergův test heteroskedasticity	Diorit I	Diorit II	Gabro	Granit	Tonalit
Hodnota kritéria CW:	1,2221	3,6151	3,3042	3,6485	3,4979
Kvantil $\chi^2(1-\alpha,1)$ :	3,8415	3,8415	3,8415	3,8415	3,8415
Pravděpodobnost:	0,2689	0,0573	0,0691	0,0561	0,0614
<b>Závěr:</b>	<b>Rezidua vykazují homoskedasticitu</b>	<b>Rezidua vykazují homoskedasticitu</b>	<b>Rezidua vykazují homoskedasticitu</b>	<b>Rezidua vykazují homoskedasticitu</b>	<b>Rezidua vykazují homoskedasticitu</b>
Jarque-Berrův test normality	Diorit I	Diorit II	Gabro	Granit	Tonalit
Hodnota kritéria JB:	2,7870	4,7456	5,4090	1,2955	3,3869
Kvantil $\chi^2(1-\alpha,2)$ :	5,9915	5,9915	5,9915	5,9915	5,9915
Pravděpodobnost:	0,2482	0,0932	0,0669	0,5232	0,1839
<b>Závěr :</b>	<b>Rezidua mají normální rozdělení</b>	<b>Rezidua mají normální rozdělení</b>	<b>Rezidua mají normální rozdělení</b>	<b>Rezidua mají normální rozdělení</b>	<b>Rezidua mají normální rozdělení</b>

- V tabulce jsou statistické charakteristiky autokorelace a znaménkového testu reziduí.

Waldův test autokorelace	Diorit I	Diorit II	Gabro	Granit	Tonalit
Hodnota kritéria WA:	2,1498	0,0524	0,1131	0,5113	0,1123
Kvantil $\chi^2(1-\alpha, 1)$ :	3,8415	3,8415	3,8415	3,8415	3,8415
Pravděpodobnost:	0,2689	0,0573	0,0691	0,0561	0,0614
<b>Závěr:</b>	<b>Autokorelace je nevýznamná</b>	<b>Autokorelace je nevýznamná</b>	<b>Autokorelace je nevýznamná</b>	<b>Autokorelace je nevýznamná</b>	<b>Autokorelace je nevýznamná</b>
Znaménkový test reziduí	Diorit I	Diorit II	Gabro	Granit	Tonalit
Hodnota kritéria Sg:	1,6551	1,7028	1,7765	0,4435	1,8018
Kvantil $N(1-\alpha/2)$ :	1,9600	1,9600	1,9600	1,9600	1,9600
Pravděpodobnost:	0,0979	0,0886	0,0757	0,6574	0,0716
<b>Závěr:</b>	<b>V reziduích není trend.</b>	<b>V reziduích není trend.</b>	<b>V reziduích není trend.</b>	<b>V reziduích není trend.</b>	<b>V reziduích není trend.</b>

# Regresní diagnostika

## □ Analýza klasických reziduí

- Směrodatná odchylka reziduí se rovná velikosti odhadu náhodných chyb závislé proměnné  $A_{rel}$ .
- Rezidua vykazují homoskedasticitu.
  - Homoskedasticita svědčí o konstantnosti rozptylu.
- Rezidua odpovídají normálnímu rozdělení.
- Autokorelace je nevýznamná.
  - Hodnoty pravděpodobnosti jsou pro diorit II, gabro, granit a tonalit mírně větší než hladina významnosti  $\alpha = 0,05$ , ale pro diorit I je tato hodnota významně větší vůči hodnotě 0,05, což svědčí o tom, že model není silně statisticky významný.
- V reziduích není trend.
  - To znamená, že nejsou přítomné závislosti, které nejsou postihnuty navrženým modelem. Hodnoty pravděpodobnosti ale opět ukazují na to, že model není silně statisticky významný.



# Regresní diagnostika

## □ Analýza vlivných bodů

- Grafy reziduí a v tabulkách z protokolu programu QC.Expert , ve kterých jsou uvedeny standardní, Jackknife a predikovaná rezidua, diagonální prvky projekční matice –  $\text{Diag}(H_{ii})$  a Atkinsonové vzdálenosti indikují vlivné body – v maximech regresních křivek a jejich okolí.
  - Vzhledem k malému počtu experimentálních bodů nelze je vynechat ze statistického hodnocení.

## □ Tabulka reziduí pro Gabro – pro ilustraci.

Index	Standardní	Jackknife	Predikované	Diag(Hii)	Atkins. vzdál.
1	0,3433	0,3314	0,0157	0,0018	0,0294
2	1,5071	1,5939	0,0702	0,0427	0,7010
3	0,2831	0,2729	0,0173	0,4399	0,5034
4	-1,6935	-1,8430	-0,1208	0,5911	4,6125
5	3,0212	5,3181	0,1710	0,3506	8,1348
6	-0,6321	-0,6169	-0,0337	0,2664	0,7738
7	-1,4749	-1,5530	-0,0757	0,2106	1,6699
8	-0,2672	-0,2575	-0,0136	0,1917	0,2610
9	-0,0346	-0,0333	-0,0018	0,2086	0,0356
10	-0,0058	-0,0056	-0,0003	0,2583	0,0069
11	-0,0606	-0,0583	-0,0031	0,2129	0,0631
12	0,4134	0,3999	0,0205	0,1542	0,3554
13	0,4657	0,4512	0,0217	0,0413	0,1950
14	0,3733	0,3606	0,0172	0,0200	0,1074
15	0,8677	0,8589	0,0397	0,0063	0,1423
16	0,5051	0,4901	0,0231	0,0033	0,0591

- Predikční schopnost modelu  $K$  se vypočítá z rovnice (rov.2):

$$K = RSC^2 / ((RSC(M1))^2 + (RSC(M2))^2)$$

- **M1 M2 představují rozdělené intervaly experimentálních bodů pro jednotlivé horniny:**
  - diorit I M1: body 1-15 a M2: body 16-30,
  - diorit II M1: body 1-9 a M2: body 10-19,
  - gabro M1: body 1-8 a M2: body 9-16,
  - granit M1: body 1-7 a M2: body 8-15,
  - a tonalit M1: body 1-8 a M2: body 9-17.

# Predikční schopnost modelu

- Hodnoty predikční schopnosti modelu jsou pro sledované krystalické horniny uvedeny v tabulce.

Hornina	RSČ	RSČ (M1)	RSČ (M2)	K
Diorit I	0,0118	0,0053	0,0065	1,98
Diorit II	0,0241	0,0185	0,0056	1,55
Gabro	0,0270	0,0239	0,0031	1,26
Granit	0,0739	0,0594	0,0145	1,46
Tonalit	0,0160	0,0104	0,0056	1,83

- **Hodnoty K se blíží k jedničce. Což znamená, že predikční schopnost modelu, i vzhledem k nízkým hodnotám střední kvadratické chyby predikce (MEP), je dobrá.**

# Souhlas s požadavky fyzikálního smyslu

- V tabulkách jsou porovnány hodnoty parametrů odhadnutých navrženým modelem a metodou nelineární regrese s hodnotami uvedených v literatuře.
  - Porovnání shody  $R' = R$  pro případ platnosti lineární sorpční izotermy (v modelu parametr p1 s parametrem p2)

Hornina	p1=R', model	p2=R, model	R, lit.	p1/p2, model	p1/R <sub>lit.</sub>
Diorit I	81,35	73,11	69	1,11	1,18
Diorit II	72,64	61,11	53	1,19	1,37
Gabro	45,21	37,53	35	1,2	1,29
Granit	24,2	18,65	13	1,3	1,86
Tonalit	43,02	41,14	36	1,05	1,2



# Souhlas s požadavky fyzikálního smyslu

- Porovnání hodnot  $R$  (v modelu parametr  $p_2$ ) s literárním údajem.

Hornina	$p_2=R$ , model	$R$ , lit.
Diorit I	73,11	69
Diorit II	61,11	53
Gabro	37,53	35
Granit	18,65	13
Tonalit	41,14	36

- Z těchto tabulek je vidět jistou oprávněnost, že lze předpokládat lineární sorpční izotermu. Hodnoty odhadnuté navrženým modelem jsou vůči literárním údajům vyšší, ale vzhledem k chybějícím statistickým údajům v literatuře, nelze provést přesnější vyhodnocení.



# Souhlas s požadavky fyzikálního smyslu

- **Porovnání hodnot Pecletova čísla (v modelu parametr p3) s literárním údajem.**

Hornina	p3, model	Pe, lit.
Diorit I	9,44	10
Diorit II	6,71	7
Gabro	6,51	9
Granit	5,04	5
Tonalit	9,25	8

- **Vzhledem k chybějícím statistickým údajům v literatuře nelze provést přesnější statistické vyhodnocení. Hodnoty Pecletova čísla, uvedené v literatuře 4, leží v intervalech spolehlivosti odhadů parametrů navrženého modelu pro diorit II a granit.**

# Závěr

- ❑ Odhad parametrů sorpčního a desorpčního chování radionuklidu  $^{85}\text{Sr}$  na krystalických horninách diorit I, diorit II, gabro, granit a tonalit lze velice dobře provést metodou nelineární regrese pomocí modelu uvedeného v literatuře.
- ❑ Statistickými charakteristikami a analýzou reziduí byla provedena kritika modelu, dat a metody.
- ❑ Získané odhady parametrů odpovídají hodnotami fyzikálnímu smyslu a lze z nich odvodit konstanty popisující sorpční a desorpční chování radionuklidu  $^{85}\text{Sr}$  na uvedených krystalických horninách.



# Dotazy a připomínky

□ V práci je uvedena řada statistických charakteristik pro posouzení kvality regresního modelu. Uveďte tři z nich, které považujete za nejdůležitější:

- **1. Regresní rabat v % – stonásobek koeficientu determinace,  $D^2$ , informuje o přiblížení navrženého modelu experimentálními body.**

**Koeficient determinace :**

$$D = 1 - U(b) / \sum_{i=1}^n (y_i - \bar{y})^2$$

- **2. Střední kvadratická chyba predikce, MEP, informuje o predikční schopnosti modelu. Čím je hodnota nižší, tím je model pravděpodobněji a má lepší predikční schopnost.**

$$MEP = \frac{1}{n} \sum_{i=1}^n (y_i - f(x_i, \beta_i))^2$$



# Dotazy a připomínky

- **3. Akaikovo informační kritérium, AIC, slouží k posouzení kvality různých modelů. Čím je hodnota AIC nižší, tím je model optimálnější.**

$$AIC = n \cdot \ln \left[ \frac{U(b)}{n} \right] + 2m$$

- **Účelová funkce  $U(b)$ , vyjadřuje minimum sumy čtverců reziduí.**



# Dotazy a připomínky

- **Korelaci mezi parametry lze odstranit reparametrizací (zápisem v jiném algebraickém tvaru). Platí to obecně?**
  - **1. Je to jedna z možností.**
  - **2. Posouzení, zda model není přeurčný – velký počet parametrů.**
  - **3. Posouzení, zda není přeurčenost parametrů vzhledem k experimentálním bodům.**
  - **4. Rozšířením rozsahu experimentálních dat.**

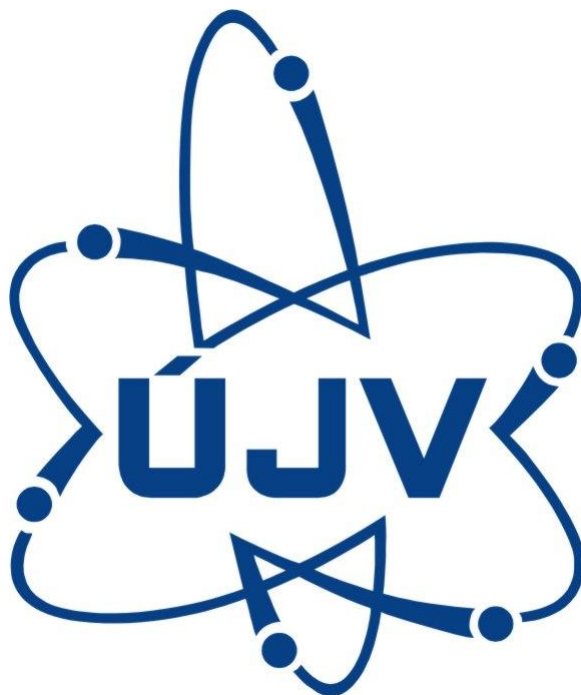


# Dotazy a připomínky

- Příčinou velkých směrodatných odchylek parametrů může být špatně navržený regresní model. Mohou to být ale i jiné příčiny. Jaké?
  - 1. Při použití metody vážených nejmenších čtverců k odstranění trendů v reziduích a heteroskedasticity může dojít ke zvětšení směrodatných odchylek odhadů parametrů.
  - 2. Malý rozsah – počet, experimentálních dat.



# Děkuji Vám za pozornost



[www.ujv.cz](http://www.ujv.cz)