

6.5 Vícerozměrné lineární regresní modely

6.5.1 Vzorová úloha

Na úloze **M6.19** *Vliv tří parametrů na obsah kadmia v potravinářské pšenici* ukážeme postup analýzy vícerozměrného lineárního regresního modelu

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m,$$

kde $\beta_0, \beta_1, \beta_2, \dots, \beta_m$ jsou odhadované parametry: u vzorků potravinářské pšenice byl zjišťován obsah kadmia v zrna y v závislosti na obsahu kadmia v otrubách x_1 , ve stonku s listy x_2 a v kořenovém systému x_3 . (1) Vyšetřením regresního tripletu nalezněte nejlepší vícerozměrný regresní model. (2) Využijte k tomu regresní diagnostiku a pomocí parciálních regresních a parciálních reziduálních grafů diskutujte významnost jednotlivých parametrů v modelu stejně jako i jejich fyzikální smysl.

Data: Obsah v otrubách x_1 [mg/l], ve stonku s listy x_2 [mg/l] a v kořenovém systému x_3 [mg/l], obsah kadmia v zrna y [mg/l]:

x_1	x_2	x_3	y	x_1	x_2	x_3	y	x_1	x_2	x_3	y
1.5	1.5	1.5	1.6,	1.5	1.6	1.3	1.6,	2.0	1.9	2.2	2.1,
2.0	2.0	2.2	2.1,	6.6	7.1	7.6	8.1,	7.1	8.2	6.6	7.9,
7.8	9.1	7.1	8.4,	8.4	10.3	7.8	10.3,	8.4	9.6	8.6	9.6,
8.6	10.0	9.1	10.8,	9.0	12.3	10.5	13.1,	10.2	13.1	11.8	15.1,
1.3	1.3	1.3	1.3,	1.1	1.3	1.2	1.2,	1.3	1.6	1.3	1.5,
1.5	1.6	1.2	1.5,								

Řešení: 1. Návrh modelu: začíná se vždy od nejjednoduššího modelu, u kterého vystupují x_1, x_2, x_3 v prvních mocninách a nevyskytují se žádné interakční členy. Na začátku analýzy vždy zařadíme i absolutní člen β_0 , takže pro daná data bude navržený regresní model tvaru

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.$$

2. Předběžná analýza dat: polohu a proměnlivost proměnných y, x_1, x_2, x_3 charakterizuje průměr a směrodatná odchylka hodnot každé proměnné. Párový korelační koeficient y vs. x_1, y vs. x_2, y vs. x_3 ukazuje na vysokou korelaci, všechny tři nezávislé proměnné x_1, x_2, x_3 jsou se závisle proměnnou y spjatý silnou lineární závislostí.

Párové korelační koeficienty mezi dvojicemi vysvětlujících proměnných ukazují na silnou korelaci i mezi nezávisle proměnnými. Nejsilnější lineární vztah existuje mezi x_1 vs. x_2 a u x_1 vs. x_3 a x_2 vs. x_3 je rovněž silná korelace.

Blok INDIKACE MULTIKOLINEARITY vykazuje ve všech kritériích multikolinearitu, protože je značná korelace mezi x_1 a x_2 .

Proměnná	Průměr	Směrodatná odchylka	Párový korelační koeficient	Spočtená hladina významnosti
y	6.0125E+00	4.8734E+00	1.0000	----
x_1	4.8937E+00	3.5692E+00	0.9837	0.000
x_2	5.7813E+00	4.5296E+00	0.9935	0.000
x_3	5.0813E+00	3.8782E+00	0.9948	0.000
Párové korelační koeficienty mezi dvojicemi vysvětlujících proměnných				Spočtená hladina významnosti

x1 versus x2:	9.9344E-01	0.000		
x1 versus x3:	9.8693E-01	0.000		
x2 versus x3:	9.8847E-01	0.000		
INDIKACE MULTIKOLINEARITY:				
Č [j]	Vlastní čísla korel. matice I[j]	Čísla podmí- něnosti K[j]	Variance inflation faktor VIF[j]	Vícenás. korel. koef. pro X[j]
1	6.4568E-03	4.6141E+02	8.3272E+01	0.9940
2	1.4307E-02	2.0823E+02	9.4324E+01	0.9947
3	2.9792E+00	1.0000E+00	4.7508E+01	0.9894
Maximální číslo podmíněnosti K:		4.6141E+02		

Nápověda: $K[j]$, $K > 1000$ indikuje silnou multikolaritu, $VIF[j] > 10$ indikuje silnou multikolaritu.

3. Odhadování parametrů: klasickou metodou nejmenších čtverců (MNČ) byly nalezeny nejlepší odhady čtyř parametrů $\beta_0, \beta_1, \beta_2, \beta_3$. Studentův t -test ukázal, že absolutní člen β_0 je statisticky nevýznamný, zatímco ostatní parametry statisticky významné jsou. To je v souladu i s biologickou interpretací: β_0 se týká zbytkového obsahu kadmia v zrnu, když je obsah kadmia v otrubách ($x_1 = 0$), ve stonku ($x_2 = 0$) a v kořenovém systému ($x_3 = 0$) nulový. Je zřejmé, že když v celé rostlince bude obsah kadmia nulový, musí být nulový obsah i v zrnu a zbytkový obsah proto nemá ani biologický smysl. Absolutní člen β_0 je proto nutno ve zpřesněném modelu vynechat.

Parametr	Odhad	Směrodatná odchylka	Test $H_0: B[j] = 0$ vs. $H_A: B[j] \neq 0$		
			t-kriterium	hypoteza H_0 je	H l a d . význam.
B[0]	-7.2666E-02	1.3791E-01	-5.2692E-01	Akceptována	0.608
B[1]	-6.8505E-01	1.9165E-01	-3.5746E+00	Zamítnuta	0.004
B[2]	8.9619E-01	1.6072E-01	5.5761E+00	Zamítnuta	0.000
B[3]	8.3769E-01	1.3322E-01	6.2879E+00	Zamítnuta	0.000

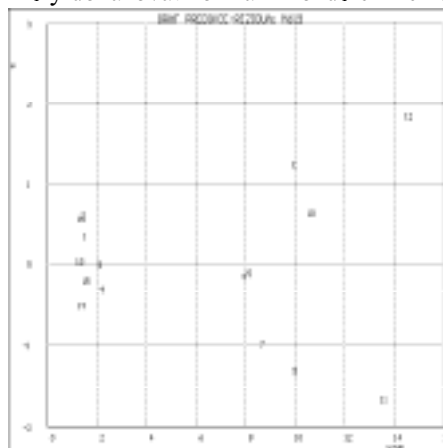
4. Základní statistické charakteristiky: vícenásobný korelační koeficient R ukazuje, že navržený lineární regresní model je statisticky významný. Vysoká hodnota koeficientu determinace $D = R^2$ ukazuje, že všechny body výtečně korespondují s modelem. Predikovaný koeficient determinace R_p^2 má podobný význam jako koeficient determinace, je však vyčíslen jinak, místo RSC se ve vztahu užije MEP . Střední kvadratická chyba predikce MEP a Akaikovo informační kritérium AIC se užívají k rozlišení mezi několika navrženými modely. Za optimální se považuje model, pro který dosahuje MEP a AIC minimální hodnotu.

Vícenásobný korelační koeficient, R	: 9.9858E-01
Koeficient determinace, D	: 9.9716E-01
Predikovaný koeficient determinace, R_p^2	: 9.9527E-01
Střední kvadratická chyba predikce, MEP	: 2.1014E-01
Akaikovo informační kritérium, AIC	: -3.6180E+01

5. Regresní diagnostika: obsahuje pomůcky a postupy pro interaktivní analýzu (a) dat, (b) modelu, (c) metody, což jsou složky tzv. *regresního tripletu*.

5.1 Data: skládá se z analýzy několika druhů grafických diagnostik a tabulek různých druhů reziduí.

(a) *Analýza klasických reziduí* není příliš spolehlivá, protože klasická rezidua jsou korelovaná, s nekonstantním rozptylem, jeví se normálnější než náhodné chyby (*efekt supernormality*) a nemusí indikovat silně odlehle hodnoty. Grafická analýza e vs. \hat{y}_p (obr. 6.5-1) je však schopna indikovat podezřelé body, trend, a nekonstantnost podmíněného rozptylu tj. heteroskedasticitu. Míry polohy a rozptýlení klasických reziduí by měly dosahovat hodnot blízkých experimentálnímu šumu. *Odhad směrodatné odchylky $s(e)$* by se měl blížit svou velikostí experimentální chybě, kterou je zatížena závisle proměnná. Odhad šikmosti a špičatosti by měly dokazovat normální rozdělení reziduí, normalitu.



Obr. 6.5-1 Analýza klasických reziduí, *ADSTAT*

Bod	Měřená hodnota	Predikovaná hodnota	Směrodatná odchylka	Klasické reziduum	Relativní reziduum
i	$y_{exp}[i]$	$y_{vyp}[i]$	$s(y_{vyp}[i])$	$e[i]$	$er[i]$
1	1.6000E+00	1.5006E+00	1.0203E-01	9.9415E-02	6.2135E+00
2	1.6000E+00	1.4227E+00	1.0587E-01	1.7733E-01	1.1083E+01
3	2.1000E+00	2.1029E+00	1.1045E-01	-2.9207E-03	-1.3908E-01
4	2.1000E+00	2.1925E+00	1.0412E-01	-9.2540E-02	-4.4067E+00
5	8.1000E+00	8.1354E+00	2.0871E-01	-3.5420E-02	-4.3729E-01
6	7.9000E+00	7.9410E+00	1.3352E-01	-4.1016E-02	-5.1919E-01
7	8.4000E+00	8.6869E+00	1.6249E-01	-2.8690E-01	-3.4155E+00
8	1.0300E+01	9.9377E+00	1.7913E-01	3.6232E-01	3.5176E+00
9	9.6000E+00	9.9805E+00	1.3555E-01	-3.8050E-01	-3.9636E+00
10	1.0800E+01	1.0621E+01	1.3061E-01	1.7918E-01	1.6591E+00
11	1.3100E+01	1.3581E+01	2.2832E-01	-4.8080E-01	-3.6703E+00
12	1.5100E+01	1.4565E+01	1.8609E-01	5.3530E-01	3.5450E+00
13	1.3000E+00	1.2908E+00	1.0514E-01	9.1821E-03	7.0632E-01
14	1.2000E+00	1.3441E+00	1.1635E-01	-1.4406E-01	-1.2005E+01
15	1.5000E+00	1.5597E+00	1.1723E-01	-5.9675E-02	-3.9784E+00
16	1.5000E+00	1.3389E+00	1.0987E-01	1.6110E-01	1.0740E+01

Reziduální součet čtverců, RSC	: 1.0114E+00
Průměr absolutních hodnot reziduí, Me	: 1.9048E-01
Průměr relativních reziduí, Mer	: 4.3750E+00
Odhad reziduálního rozptylu, $s^2(e)$: 8.4282E-02
Odhad směrodatné odchylky reziduí, $s(e)$: 2.9031E-01
Odhad šikmosti reziduí, $g1(e)$: 9.2354E-02
Odhad špičatosti reziduí, $g2(e)$: 2.8755E+00

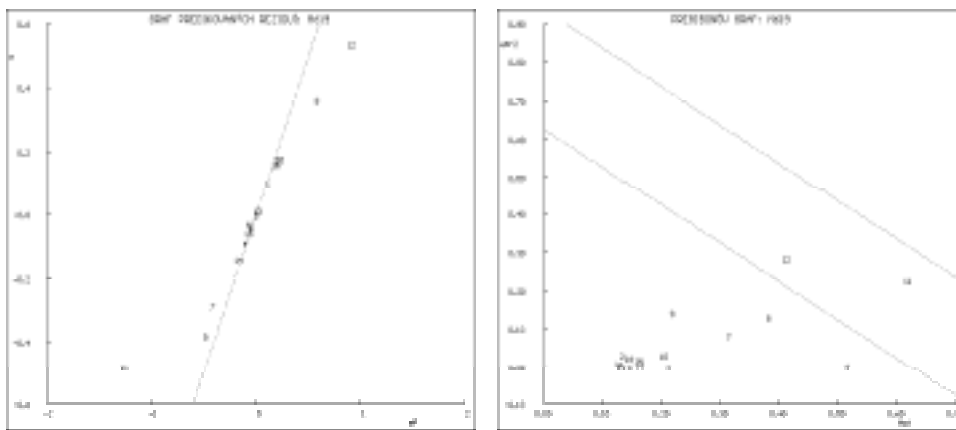
(b) **Analýza ostatních reziduí:** Jackknife rezidua indikují odlehlé body, diagonální prvky H_{ii} od projekční matice \mathbf{H} a diagonální prvky H_{mii} od zobecněné projekční matice \mathbf{H}_m pouze extrémy. Ostatní druhy reziduí a kritéria v tabulce pak obojí (značeno hvězdičkou u hodnoty). Jackknife rezidua $eJ[i]$ ukazují, že bod č. 11 a 12 je odlehlý, stejně tak i Cookova vzdálenost $D[i]$, Atkinsova vzdálenost $A[i]$ ukazují na č. 8, 11, 12, kritérium $DF[i]$ na č. 8, 11, 12 a věrohodnostní vzdálenosti $LD(b)[i]$ a $LD(s^2)[i]$ na č. 11 a $LD(b, s^2)[i]$ na č. 11 a 12. Diagonální prvky $H[i, i]$ projekční matice \mathbf{H} ukazují na extrémy č. 5 a 11, a diagonální prvky $H_m[i, i]$ zobecněné projekční matice \mathbf{H}_m pak na extrémy č. 11 a 12.

INDIKACE VLIVNÝCH BODU: (* indikuje odlehlý nebo vlivný bod)				
Bod	Standardizované reziduum	Jackknife reziduum	Predikované reziduum	Diagonální prvky
i	$eS[i]$	$eJ[i]$	$eP[i]$	$H[i,i]$
1	3.6577E-01	3.5217E-01	1.1343E-01	1.2352E-01
2	6.5601E-01	6.3966E-01	2.0453E-01	1.3298E-01
3	-1.0879E-02	-1.0415E-02	-3.4150E-03	1.4474E-01
4	-3.4147E-01	-3.2854E-01	-1.0620E-01	1.2863E-01
5	-1.7552E-01	-1.6827E-01	-7.3308E-02	5.1683E-01*
6	-1.5911E-01	-1.5249E-01	-5.2019E-02	2.1152E-01
7	-1.1925E+00	-1.2161E+00	-4.1777E-01	3.1326E-01
8	1.5859E+00	1.7078E+00	5.8504E-01	3.8069E-01
9	-1.4821E+00	-1.5700E+00	-4.8658E-01	2.1800E-01
10	6.9110E-01	6.7525E-01	2.2465E-01	2.0240E-01
11	-2.6814E+00	-4.0550E+00*	-1.2604E+00	6.1852E-01*
12	2.4023E+00	3.1923E+00*	9.0863E-01	4.1087E-01
13	3.3931E-02	3.2488E-02	1.0568E-02	1.3115E-01
14	-5.4162E-01	-5.2502E-01	-1.7163E-01	1.6062E-01
15	-2.2469E-01	-2.1557E-01	-7.1301E-02	1.6305E-01
16	5.9952E-01	5.8279E-01	1.8803E-01	1.4322E-01
Bod	Zobecněné diag. prvky	Cookova vzdálenost	Atkinsonova vzdálenost	Vliv na predikci
i	$H_m[i,i]$	$D[i]$	$A[i]$	$DF[i]$
1	1.3329E-01	4.7136E-03	2.2899E-01	1.3221E-01
2	1.6407E-01	1.6501E-02	4.3390E-01	2.5051E-01
3	1.4475E-01	5.0070E-06	7.4214E-03	-4.2848E-03
4	1.3709E-01	4.3030E-03	2.1863E-01	-1.2622E-01

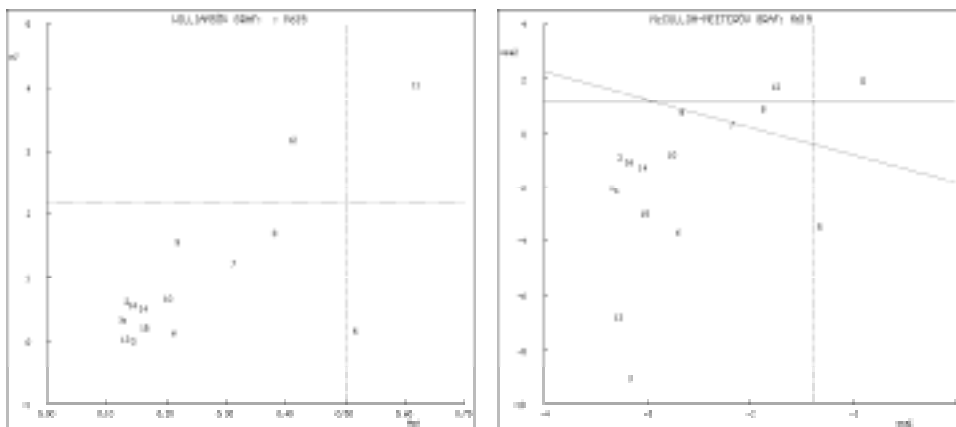
5	5.1807E-01	8.2386E-03	3.0143E-01	-1.7403E-01
6	2.1318E-01	1.6978E-03	1.3680E-01	-7.8983E-02
7	3.9465E-01	1.6218E-01	1.4226E+00	-8.2134E-01
8	5.1049E-01	3.8650E-01*	2.3192E+00*	1.3390E+00*
9	3.6116E-01	1.5310E-01	1.4358E+00	-8.2895E-01
10	2.3414E-01	3.0299E-02	5.8916E-01	3.4015E-01
11	8.4709E-01*	2.9145E+00*	8.9433E+00*	-5.1634E+00*
12	6.9419E-01*	1.0062E+00*	4.6176E+00*	2.6660E+00*
13	1.3124E-01	4.3449E-05	2.1863E-02	1.2623E-02
14	1.8114E-01	1.4034E-02	3.9779E-01	-2.2966E-01
15	1.6657E-01	2.4587E-03	1.6480E-01	-9.5150E-02
16	1.6888E-01	1.5020E-02	4.1270E-01	2.3827E-01
Bod	Věrohodnostní vzdálenosti			
i	LD(b)[i]	LD(s ²)[i]	LD(b,s ²)[i]	
1	2.5120E-02	2.2351E-02	4.6185E-02	
2	8.7766E-02	6.2215E-03	9.1797E-02	
3	2.6704E-05	3.2606E-02	3.2632E-02	
4	2.2933E-02	2.3569E-02	4.5296E-02	
5	4.3879E-02	3.0095E-02	7.1394E-02	
6	9.0523E-03	3.0538E-02	3.9045E-02	
7	8.4238E-01	3.0992E-02	9.5089E-01	
8	1.9389E+00	2.4682E-01	2.6917E+00	
9	7.9637E-01	1.5878E-01	1.0958E+00	
10	1.6079E-01	4.5425E-03	1.6232E-01	
11	1.0861E+01*	7.8276E+00*	4.4183E+01*	
12	4.6276E+00	3.4386E+00	1.3130E+01*	
13	2.3173E-04	3.2520E-02	3.2738E-02	
14	7.4671E-02	1.2499E-02	8.4425E-02	
15	1.3108E-02	2.8534E-02	4.0880E-02	
16	7.9906E-02	9.2132E-03	8.6631E-02	

(c) *Grafy vlivných bodů* (obr. 6.5-2) jsou schopny indikovat a současně i testovat, dokazovat přítomnost odlehlých hodnot a extrémů. *Graf predikovaných reziduí* ukazuje na odlehlé body č. 8, 11, 12 a částečně na extrémy č. 11 a 12. *Pregibonův graf* ukazuje na středně vlivné body č. 11 a 12. *Williamsův graf* indikuje č. 11 a 12 jako odlehlé body a extrémy č. 11 a 12. *McCulloh-Meeterův graf* dokazuje odlehlé body č. 11 a 12, extrémy č. 8 a 11. Konečně *L-R graf* dokazuje odlehlé body č. 8, 11 a 12 a extrémy č. 11, 12 a 9.

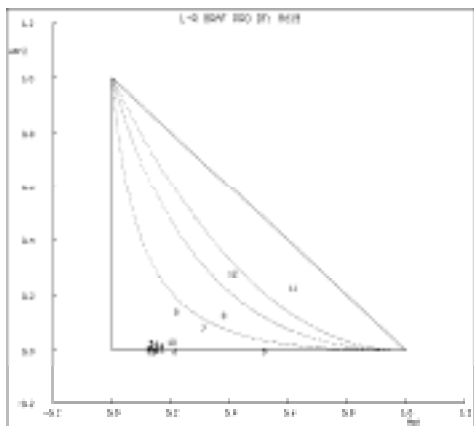
Lze uzavřít, že body č. 11 a 12 jsou většinou diagnostik prokázány za odlehlé, a proto je třeba je dále prověřit nebo z výběru vyloučit.



Obr. 6.5-2 Grafy vlivných bodů: (a) graf predikovaných reziduí a (b) Pregibonův graf, *ADSTAT*

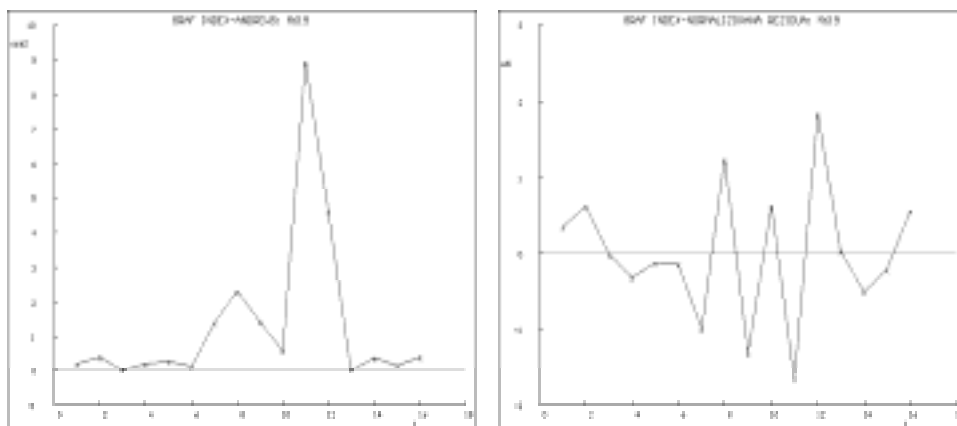


Obr. 6.5-2 Grafy vlivných bodů: (c) Williamsův graf a (d) McCulloh-Meeterův graf, *ADSTAT*

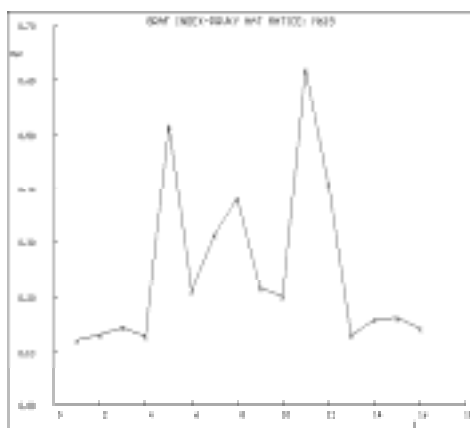


Obr. 6.5-2 Grafy vlivných bodů: (e) L-R graf, *ADSTAT*

(d) **Indexové grafy** (obr. 6.5-3) upozorňují na podezřelé body. *Andrewsův indexový graf* a *graf normovaných reziduí* ukazují na podezřelé body č. 5, 8, 11, 12. *Indexový graf prvků projekční matice H* pak na podezřelé extrémní body č. 5, 8, 11.

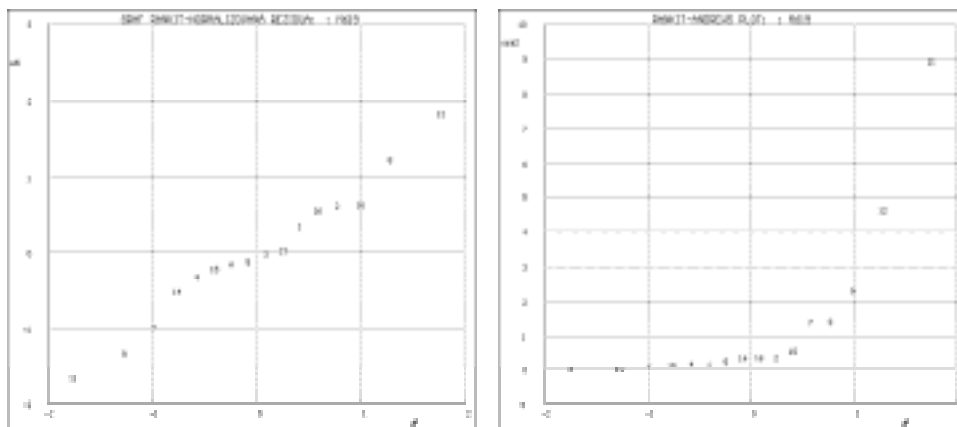


Obr. 6.5-3 Indexové grafy: (a) Andrewsův graf a (b) graf normovaných reziduí, *ADSTAT*

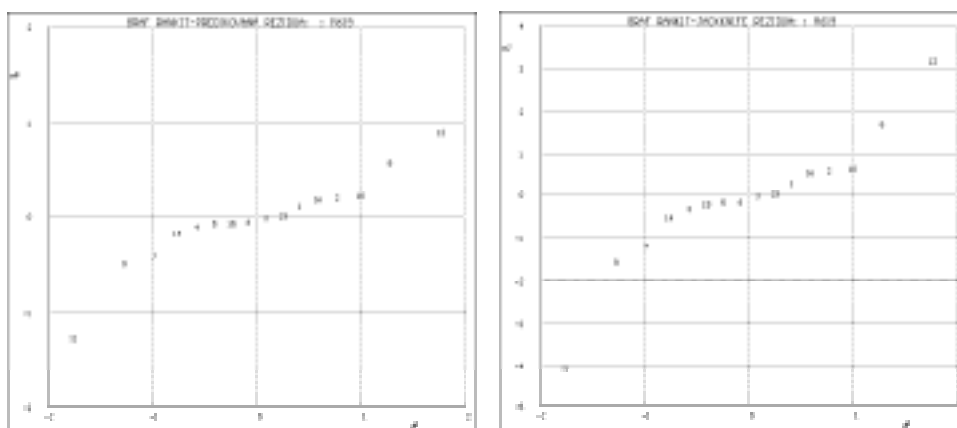


Obr. 6.5-3 (c) graf prvků H -projekční matice, *ADSTAT*

(e) **Rankitové grafy** (obr. 6.5-4) ukazují vedle normality rozdělení dotyčných reziduí i na vlivné (zde odlehlé) body. *Graf normovaných reziduí* ukazuje na začátku č. 11 a na konci č. 12 jako odlehlé body. *Andrewsův graf* ukazuje na č. 8, 11 a 12 jako odlehlé. *Graf Jackknife reziduí* č. 11 a 12 jako odlehlé. Po odstranění dvou odlehlých bodů č. 11 a 12 lze konstatovat, že zbytek dat nevykazuje odchylky od normality.

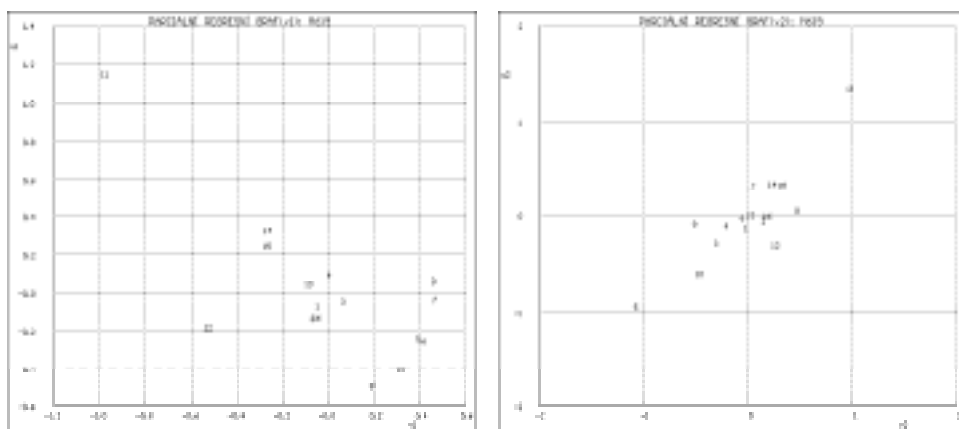


Obr. 6.5-4 Rankitové grafy: (a) graf normovaných reziduí, a (b) Andrewsův graf, *ADSTAT*

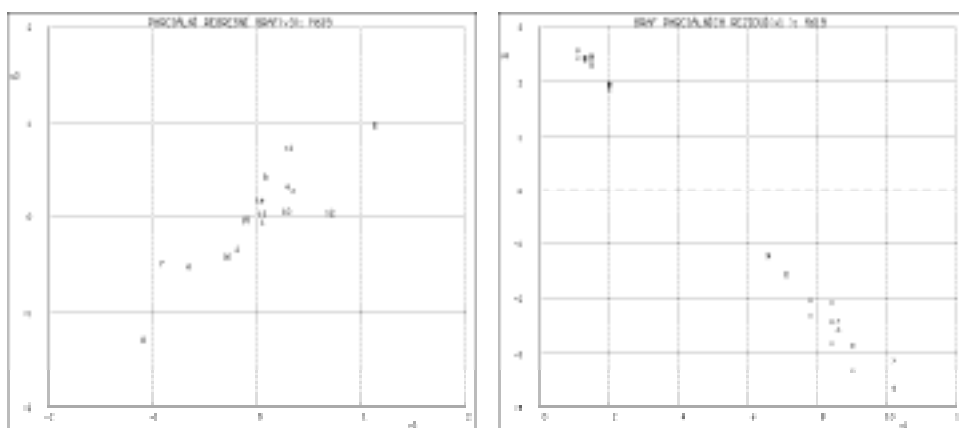


Obr. 6.5-4 Rankitové grafy: (c) graf predikovaných reziduí, a (d) graf Jackknifé reziduí, *ADSTAT*

5.2 Model: *Parciální regresní grafy* (obr. 6.5-5) ale především *parciální reziduální grafy* (obr. 6.5-6) ukazují na čisté lineární závislosti jednotlivých nezávisle proměnných. Vedle posouzení závislosti navrženého regresního modelu umožňují také indikovat vlivné body, a to č. 5, 8, 11 a 12. Navržený model se jeví stran členů $\beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$ správný, pouze β_0 je nadbytečné.

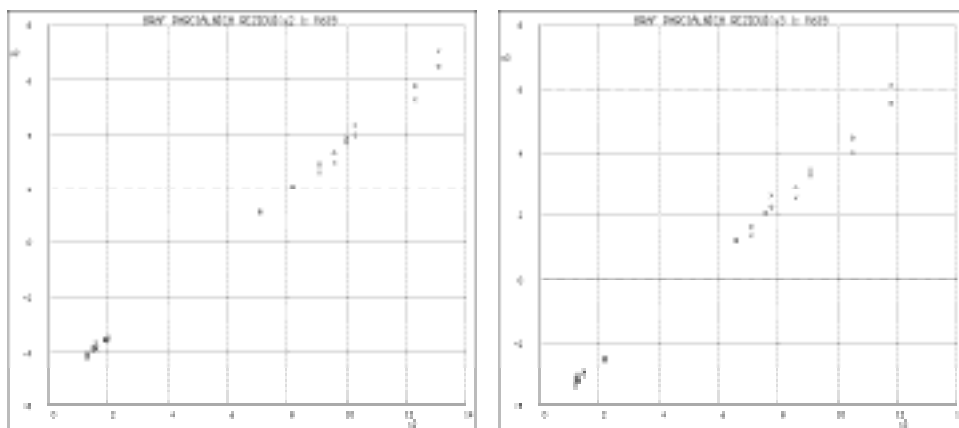


Obr. 6.5-5 Parciální regresní grafy: (a) pro proměnnou x_1 , (b) pro proměnnou x_2 , *ADSTAT*



Obr. 6.5-5 (c) pro proměnnou x_3 ,

Obr. 6.5-6 (a) Parciální reziduální graf pro proměnnou x_4



Obr. 6.5-6 Parciální reziduální grafy (b) pro proměnnou x_2 , (c) pro proměnnou x_3 , *ADSTAT*

5.3 Metoda: do této části patří vyšetření splnění základních předpokladů metody nejmenších čtverců (MNC), za kterých by měla vést k nejlepším nestranným lineárním odhadům regresních parametrů:

Fisher-Snedecorův test významnosti regrese potvrdil, že navržený model je přijat jako významný, jinými slovy: závisle proměnná y a nezávisle proměnné x_1, x_2, x_3 jsou v lineární závislosti.

Scottovo kritérium multikolinearity ukazuje, že navržený model není korektní s ohledem na vazby mezi proměnnými.

Cook-Weisbergův test heteroskedasticity dokazuje, že rezidua vykazují heteroskedasticitu (nekonstantnost rozptylu).

Jarque-Berraův test normality reziduí ukazuje, že klasická rezidua vykazují Gaussovo rozdělení.

Waldův test autokorelace ukazuje, že klasická rezidua jsou autokorelována.

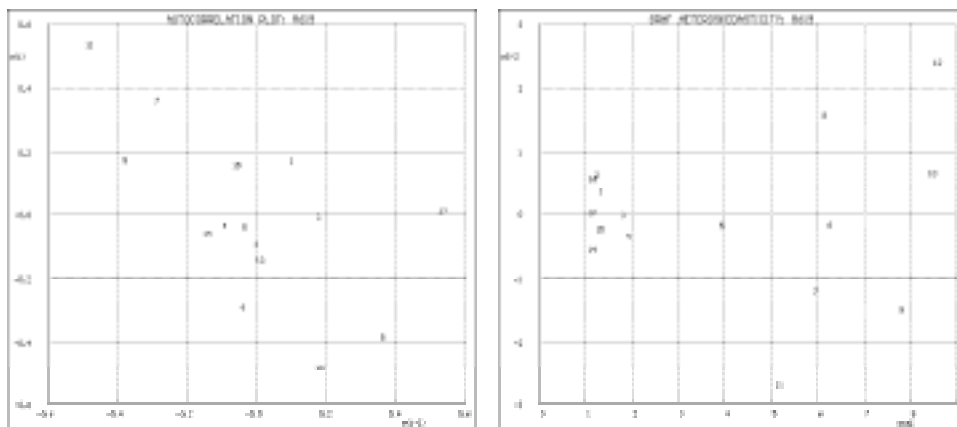
Znaménkový test prokazuje, že znaménko klasických reziduí se dostatečně střídá, a proto rezidua nevykazují žádný trend.

TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA):

Fisher-Snedecorův test významnosti regrese, F	: 1.4050E+03
Tabulkový kvantil, $F(1-\alpha, m-1, n-m)$: 3.4903E+00
Závěr: Navržený model je přijat jako významný.	
Spočtená hladina významnosti	: 0.000
Scottovo kritérium multikolinearity, M	: 9.6119E-01
Závěr: Navržený model není korektní.	
Cook-Weisbergův test heteroskedasticity, Sf	: 1.9926E+01
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua vykazují heteroskedasticitu.	
Spočtená hladina významnosti	: 0.000
Jarque-Berraův test normality reziduí, L(e)	: 3.3085E-02
Tabulkový kvantil, $\chi^2(1-\alpha, 2)$: 5.9915E+00
Závěr: Normalita je přijata.	
Spočtená hladina významnosti	: 0.984
Waldův test autokorelace, Wa	: 1.0320E+01
Tabulkový kvantil, $\chi^2(1-\alpha, 1)$: 3.8415E+00
Závěr: Rezidua jsou autokorelována.	
Spočtená hladina významnosti	: 0.001
Znaménkový test, Dt	: -1.9739E-01
Tabulkový kvantil, $N(1-\alpha/2)$: 1.6449E+00
Závěr: Rezidua nevykazují trend.	
Spočtená hladina významnosti	: 0.422

Graf autokorelace (obr. 6.5-7a) vykazuje přibližně mrak bodů reziduí.

Graf heteroskedasticity (obr. 6.5-7b) vykazuje klín, a proto rezidua vykazují heteroskedasticitu, nekonstantnost rozptylu.

Obr. 6.5-7 (a) Graf autokorelace, a (b) graf heteroskedasticity, *ADSTAT*

6. Konstrukce zpřesněného modelu: (a) Po odstranění bodů č. 11 a 12 a absolutního členu β_0 byly nalezeny nové odhady parametrů zpřesněného modelu.

Parametr	Odhad	Směrodatná odchylka	Test H0: B[j] = 0 vs. HA: B[j] \neq 0 t-kriterium	hypoteza H0 je	Hlad. význam.
B[0]	0.0000E+00	-----	-----	-----	-----
B[1]	-1.1808E+00	3.8271E-01	-3.0854E+00	Zamítnuta	0.010
B[2]	1.2454E+00	2.3610E-01	5.2751E+00	Zamítnuta	0.000
B[3]	9.1666E-01	1.5049E-01	6.0910E+00	Zamítnuta	0.000

Zpřesněný model (v závorce je uveden odhad směrodatné odchylky parametru)

$$y = -1.18 (0.38) x_1 + 1.25 (0.24) x_2 + 0.92 (0.15) x_3$$

je doložen statistickými charakteristikami: vícenásobný korelační koeficient R , koeficient determinace D a predikovaný korelační koeficient R^2_p dosáhly vesměs vysokých hodnot. Střední kvadratická chyba predikce MEP a Akaikeho informační kritérium AIC dosáhly nižších hodnot, což dokazuje lepší model než předešlý.

Vícenásobný korelační koeficient, R	: 9.9900E-01
Koeficient determinace, D	: 9.9800E-01
Predikovaný koeficient determinace, R_p^2	: 9.9792E-01
Střední kvadratická chyba predikce, MEP	: 6.0784E-02
Akaikeho informační kritérium, AIC	: -4.3472E+01

(b) Užitím statistické váhy ($w_i = 1/y_i^2$) kompenzujeme heteroskedasticitu v datech. Obdržíme nové odhady parametrů, v nichž však parametr β_1 vychází jako statisticky nevýznamný.

Parametr	Odhad	Směrodatná odchylka	Test H0: B[j] = 0 vs. HA: B[j] \neq 0 t-kriterium	hypoteza H0 je	H l a d . význam.
B[0]	0.0000E+00	-----	-----	-----	-----
B[1]	5.6441E-02	2.6023E-01	2.1689E-01	Akceptována	0.832
B[2]	6.2215E-01	1.6877E-01	3.6863E+00	Zamítnuta	0.004
B[3]	3.6328E-01	1.5078E-01	2.4094E+00	Zamítnuta	0.035

Opravený model má tvar, (v závorce je vždy uveden odhad směrodatné odchylky parametru):

$$y = 0.62 (0.17) x_2 + 0.36 (0.15) x_3$$

Jelikož došlo k významnému snížení rozhodujících kritérií, *střední kvadratické chyby predikce MEP* a *Akaikeho informačního kritéria AIC*, lze považovat tyto odhady za lepší než předešlé. *Korelační koeficient R*, a tím pádem i *koeficient determinace D* vychází nepatrně horší než u předešlého odhadu bez statistické váhy.

Vícenásobný korelační koeficient, R	: 9.9689E-01
Koeficient determinace, D	: 9.9379E-01
Predikovaný koeficient determinace, Rp ²	: 9.9505E-01
Střední kvadratická chyba predikce, MEP	: 1.2300E-02
Akaikeho informační kritérium, AIC	: -6.2092E+01

7. Zhodnocení kvality modelu: porovnáním hodnot regresní diagnostiky lze snadno provést zhodnocení *regresního tripletu* dosaženého lineárního regresního modelu pro upravená data, zbavená odlehlých hodnot a upravený regresní model bez absolutního členu a metodou vážených nejmenších čtverců. Nalezený model má tvar (v závorce je vždy uveden odhad směrodatné odchylky parametru)

$$y = 0.62 (0.17) x_2 + 0.36 (0.15) x_3$$

čili obsah kadmia v zrnu potravinářské pšenice je funkcí pouze obsahu kadmia ve stonku a v kořenovém systému a není funkcí obsahu kadmia v otrubách a dále nemá smysl uvádět ani zbytkový obsah kadmia v zrnu při nulovém obsahu kadmia ve zbytku rostlinky.