

6

LINEÁRNÍ REGRESNÍ MODEL Y

6.1 Jednorozměrné lineární regresní modely

Vzorová úloha 6.1 *Postup výstavby modelu a regresní diagnostika*

Na úloze **J6.01** *Model teplotní závislosti přechodového tlaku bismutu* ukážeme postup analýzy jednorozměrného lineárního regresního modelu. Houck studoval přechodový tlak bismutu I - II p jako funkci teploty t , str. 501 v cit⁶². Nalezněte lineární regresní model, který bude adekvátní daným datům. Vyšetřete regresní triplet a indikujte vlivné body. Vykazují klasická rezidua normalitu? Vyšetřete rankitový a kvantilový graf.

Řešení:

1. Návrh modelu: na začátku analýzy vždy zařadíme absolutní člen β_0 a navržený regresní model (přímky) bude mít tvar $y = \beta_0 + \beta_1 x$.

2. Předběžná analýza dat: poloha a proměnlivost proměnných y , x se posuzuje na základě průměru a směrodatné odchylky hodnot každé proměnné. *Pearsonův párový korelační koeficient r* ukazuje na vysokou korelaci proměnných y a x .

| Proměnná | Průměr | Směrodatná odchylna | Párový korelační koeficient | Spočtená hladina významnosti |
|----------|---------|---------------------|-----------------------------|------------------------------|
| y | 24859.0 | 427.41 | 1.0000 | ----- |
| x | 30.313 | 10.701 | -0.9983 | 0.000 |

3. Odhadování parametrů: klasickou metodou nejmenších čtverců MNČ byly nalezeny nejlepší odhady úseku β_0 a směrnice β_1 . Studentův t -test ukázal při porovnání t -kritéria s kritickou hodnotou $t_{0.95}(23-2) = 2.080$, že úsek (absolutní člen) β_0 je statisticky významný a směrnice β_1 je rovněž statisticky významná.

| Parametr | Odhad | Směrodatná odchylna | Test $H_0: b_j$ vs. $H_A: b_j \neq 0$ t-kritérium | hypotéza H_0 je | Hlad. význam. |
|----------|---------|---------------------|--|-------------------|---------------|
| b_0 | 26068.0 | 1.6169E+01 | 1612.2 | Zamítnuta | 0.000 |
| b_1 | -39.874 | 5.0419E-01 | -79.084 | Zamítnuta | 0.000 |

4. Základní statistické charakteristiky: absolutní hodnota párového korelačního koeficientu r ukazuje, že navržený lineární regresní model je statisticky významný. Vysoká

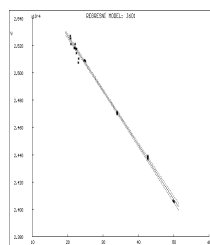
hodnota koeficientu determinace $D = 100 \% r^2$ (99.67 %), představuje procento variability, vysvětlené modelem. Predikovaný koeficient determinace R_p^2 ukazuje na predikční schopnost modelu, je však vyčíslen jinak než r^2 , místo RSC se ve vztahu užíje MEP. Střední kvadratická chyba predikce MEP a 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 | : 0.99833 |
| Koeficient determinace, D [%] | : 99.665 |
| Predikovaný koeficient determinace, R_p^2 | : 99.804 |
| Střední kvadratická chyba predikce, MEP | : 685.46 |
| Akaikovo informační kritérium, AIC | : 150.54 |

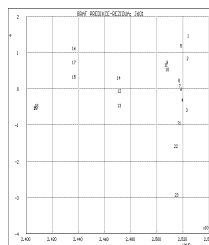
5. Regresní diagnostika: obsahuje pomůcky a postupy pro interaktivní analýzu (a) dat, (b) modelu, (c) metody, což jsou složky tzv. *regresního tripletu*.

Data: Regresní diagnostika se skládá z analýzy několika druhů grafických diagnostik (obr. 6.1-1a, b) a tabulek různých druhů reziduí.

(a) *Analýza klasických reziduí* \hat{e} není příliš spolehlivá, protože klasická rezidua \hat{e} jsou korelovaná, s nekonstantním rozptylem, jeví se normálnější než náhodné chyby (*efekt supernormality*) a nemusí indikovat silně odlehle hodnoty. Grafická analýza \hat{e} vs. \hat{y}_p (obr. 6.1-1b) je schopna indikovat pouze podezřelé body, trend a nekonstantnost rozptylu, tzv. heteroskedasticitu. Míry 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 y . Odhady šikmosti a špičatosti by měly indikovat Gaussovo normální rozdělení reziduí.



Obr. 6.1-1a Graf regresního modelu, *ADSTAT*.



Obr. 6.1-1b 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 | $e_{r, i}$ |
| 1 | 2.5276E+04 | 2.5239E+04 | 7.1310E+00 | 3.7376E+01 | 1.4787E-01 |
| 2 | 2.5256E+04 | 2.5235E+04 | 7.0971E+00 | 2.1363E+01 | 8.4587E-02 |

| | | | | | |
|---|------------|------------|--------------|-------------|-------------|
| 3 | 2.5216E+04 | 2.5231E+04 | 7.0635E+00 | -1.4649E+01 | -5.8095E-02 |
| 4 | 2.5187E+04 | 2.5195E+04 | 6.7704E+00 | -7.7630E+00 | -3.0821E-02 |
| 5 | 2.5217E+04 | 2.5187E+04 | 6.7076E+00 | 3.0212E+01 | 1.1981E-01 |
| 6 | 2.5187E+04 | 2.5187E+04 | 6.7076E+00 | 2.1178E-01 | 8.4085E-04 |
| 7 | 2.5177E+04 | 2.5175E+04 | 6.6153E+00 | 2.1738E+00 | 8.6342E-03 |
| 8 | 2.5177E+04 | 2.5171E+04 | 6.5851E+00 | 6.1612E+00 | 2.4472E-02 |
| 9 | 2.5098E+04 | 2.5079E+04 | 5.9642E+00 | 1.8871E+01 | 7.5187E-02 |
| 10 | 2.5093E+04 | 2.5079E+04 | 5.9642E+00 | 1.3871E+01 | 5.5277E-02 |
| 11 | 2.5088E+04 | 2.5071E+04 | 5.9178E+00 | 1.6845E+01 | 6.7145E-02 |
| 12 | 2.4711E+04 | 2.4712E+04 | 5.5947E+00 | -1.2920E+00 | -5.2284E-03 |
| 13 | 2.4701E+04 | 2.4712E+04 | 5.5947E+00 | -1.1292E+01 | -4.5715E-02 |
| 14 | 2.4716E+04 | 2.4708E+04 | 5.6116E+00 | 7.6953E+00 | 3.1135E-02 |
| 15 | 2.4374E+04 | 2.4365E+04 | 8.1762E+00 | 8.6087E+00 | 3.5319E-02 |
| 16 | 2.4394E+04 | 2.4365E+04 | 8.1762E+00 | 2.8609E+01 | 1.1728E-01 |
| 17 | 2.4384E+04 | 2.4365E+04 | 8.1762E+00 | 1.8609E+01 | 7.6315E-02 |
| 18 | 2.4067E+04 | 2.4078E+04 | 1.1197E+01 | -1.1301E+01 | -4.6957E-02 |
| 19 | 2.4057E+04 | 2.4070E+04 | 1.1286E+01 | -1.3327E+01 | -5.5396E-02 |
| 20 | 2.4057E+04 | 2.4070E+04 | 1.1286E+01 | -1.3327E+01 | -5.5396E-02 |
| 21 | 2.5147E+04 | 2.5171E+04 | 6.5851E+00 | -2.3839E+01 | -9.4798E-02 |
| 22 | 2.5107E+04 | 2.5147E+04 | 6.4087E+00 | -3.9915E+01 | -1.5898E-01 |
| 23 | 2.5077E+04 | 2.5151E+04 | 6.4374E+00 | -7.3902E+01 | -2.9470E-01 |
| Reziduální součet čtverců, RSC | | | : 1.3449E+04 | | |
| Průměr absolutních hodnot reziduí, M_e | | | : 1.8314E+01 | | |
| Průměr relativních reziduí, $M_{e,r}$ | | | : 7.3476E-02 | | |
| Odhad reziduálního rozptylu, $s^2(e)$ | | | : 6.4043E+02 | | |
| Odhad směrodatné odchylky reziduí, $s(e)$ | | | : 2.5307E+01 | | |
| Odhad šikmosti reziduí, $g_1(e)$ | | | :-1.098 | | |
| Odhad špičatosti reziduí, $g_2(e)$ | | | : 4.689 | | |

(b) **Analýza ostatních reziduí:** Jackknife rezidua indikují odlehlé body, z diagonálních prvků H_{ii} projekční matice H a diagonálních prvků H_{mii} zobecněné projekční matice H_m pouze extrémy. Ostatní druhy reziduí a kritéria v tabulce pak obojí (značeno hvězdičkou u dotyčné hodnoty). Jackknife rezidua $e_{j,i}$ ukazují, že bod č. 23 je odlehlý, stejně tak i Cookova vzdálenost D_i , Atkinsova vzdálenost A_i ukazují na č. 23 a kritérium DF_i pak na č. 23 a věrohodnostní vzdálenosti $LD(b)_i$ a $LD(s^2)_i$ na č. 23 a $LD(b, s^2)_i$ na č. 23. Diagonální prvky H_{ii} projekční matice H ukazují na extrémy č. 18, 19, 20 a 23, diagonální prvky zobecněné H_{mii} projekční matice H_m na extrémy 23.

Indikace vlivných bodů: (* indikuje odlehlý nebo vlivný bod)

| Bod | Standardizované reziduuum | Jackknife reziduuum | Predikované reziduuum | Diagonální prvky |
|-----|---------------------------|---------------------|-----------------------|------------------|
| i | e_{Si} | $e_{j,i}$ | e_{p_i} | H_{ii} |
| 1 | 1.5393E+00 | 1.5949E+00 | 4.0600E+01 | 7.9400E-02 |
| 2 | 8.7947E-01 | 8.7453E-01 | 2.3187E+01 | 7.8649E-02 |
| 3 | -6.0282E-01 | -5.9345E-01 | -1.5887E+01 | 7.7906E-02 |
| 4 | -3.1836E-01 | -3.1144E-01 | -8.3614E+00 | 7.1573E-02 |

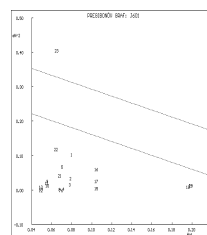
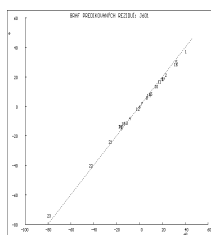
| | | | | |
|----|-------------|--------------|-------------|-------------|
| 5 | 1.2381E+00 | 1.2549E+00 | 3.2495E+01 | 7.0253E-02 |
| 6 | 8.6791E-03 | 8.4700E-03 | 2.2779E-01 | 7.0253E-02 |
| 7 | 8.8994E-02 | 8.6866E-02 | 2.3333E+00 | 6.8333E-02 |
| 8 | 2.5215E-01 | 2.4644E-01 | 6.6087E+00 | 6.7709E-02 |
| 9 | 7.6729E-01 | 7.5952E-01 | 1.9980E+01 | 5.5543E-02 |
| 10 | 5.6398E-01 | 5.5461E-01 | 1.4686E+01 | 5.5543E-02 |
| 11 | 6.8463E-01 | 6.7571E-01 | 1.7820E+01 | 5.4683E-02 |
| 12 | -5.2348E-02 | -5.1090E-02 | -1.3584E+00 | 4.8874E-02 |
| 13 | -4.5753E-01 | -4.4874E-01 | -1.1872E+01 | 4.8874E-02 |
| 14 | 3.1184E-01 | 3.0504E-01 | 8.0933E+00 | 4.9171E-02 |
| 15 | 3.5945E-01 | 3.5187E-01 | 9.6120E+00 | 1.0438E-01 |
| 16 | 1.1945E+00 | 1.2075E+00 | 3.1943E+01 | 1.0438E-01 |
| 17 | 7.7699E-01 | 7.6941E-01 | 2.0777E+01 | 1.0438E-01 |
| 18 | -4.9796E-01 | -4.8885E-01 | -1.4052E+01 | 1.9576E-01* |
| 19 | -5.8835E-01 | -5.7896E-01 | -1.6635E+01 | 1.9889E-01* |
| 20 | -5.8835E-01 | -5.7896E-01 | -1.6635E+01 | 1.9889E-01* |
| 21 | -9.7560E-01 | -9.7443E-01 | -2.5570E+01 | 6.7709E-02 |
| 22 | -1.6304E+00 | -1.7025E+00 | -4.2650E+01 | 6.4130E-02 |
| 23 | -3.0196E+00 | -3.9175E+00* | -7.9015E+01 | 6.4707E-02 |

| Bod | Zobecněné diag. prvky | Cookova vzdálenost | Atkinsonova vzdálenost | Vliv na predikci |
|-----|--------------------------|-----------------------|---------------------------|---------------------|
| i | H_{mii} | D_i | A_i | DF_i |
| 1 | 1.8327E-01 | 1.0218E-01 | 1.5177E+00 | 4.6838E-01 |
| 2 | 1.1258E-01 | 3.3013E-02 | 8.2795E-01 | 2.5551E-01 |
| 3 | 9.3862E-02 | 1.5351E-02 | 5.5896E-01 | -1.7250E-01 |
| 4 | 7.6054E-02 | 3.9067E-03 | 2.8020E-01 | -8.6472E-02 |
| 5 | 1.3812E-01 | 5.7915E-02 | 1.1178E+00 | 3.4496E-01 |
| 6 | 7.0257E-02 | 2.8459E-06 | 7.5444E-03 | 2.3283E-03 |
| 7 | 6.8684E-02 | 2.9044E-04 | 7.6230E-02 | 2.3525E-02 |
| 8 | 7.0531E-02 | 2.3087E-03 | 2.1521E-01 | 6.6415E-02 |
| 9 | 8.2020E-02 | 1.7311E-02 | 5.9683E-01 | 1.8419E-01 |
| 10 | 6.9848E-02 | 9.3529E-03 | 4.3582E-01 | 1.3450E-01 |
| 11 | 7.5782E-02 | 1.3557E-02 | 5.2662E-01 | 1.6252E-01 |
| 12 | 4.8998E-02 | 7.0407E-05 | 3.7528E-02 | -1.1581E-02 |
| 13 | 5.8355E-02 | 5.3782E-03 | 3.2962E-01 | -1.0172E-01 |
| 14 | 5.3574E-02 | 2.5145E-03 | 2.2478E-01 | 6.9367E-02 |
| 15 | 1.0989E-01 | 7.5293E-03 | 3.8925E-01 | 1.2013E-01 |
| 16 | 1.6524E-01 | 8.3153E-02 | 1.3358E+00 | 4.1223E-01 |
| 17 | 1.3013E-01 | 3.5181E-02 | 8.5115E-01 | 2.6267E-01 |
| 18 | 2.0526E-01 | 3.0179E-02 | 7.8153E-01 | -2.4119E-01 |
| 19 | 2.1209E-01 | 4.2969E-02 | 9.3477E-01 | -2.8847E-01 |
| 20 | 2.1209E-01 | 4.2969E-02 | 9.3477E-01 | -2.8847E-01 |
| 21 | 1.0996E-01 | 3.4563E-02 | 8.5092E-01 | -2.6260E-01 |
| 22 | 1.8259E-01 | 9.1073E-02 | 1.4441E+00 | -4.4566E-01 |
| 23 | 4.7079E-01* | 3.1540E-01* | 3.3389E+00* | -1.0304E+00* |

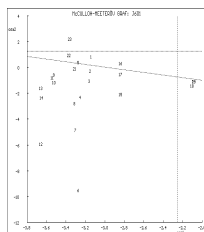
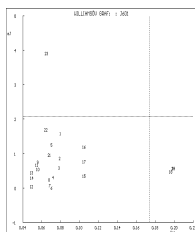
| Bod | Věrohodnostní vzdálenosti | | |
|-----|---------------------------|-------------|----------------|
| i | $LD(b)_i$ | $LD(s^2)_i$ | $LD(b, s^2)_i$ |

| | | | |
|----|-------------|-------------|-------------|
| 1 | 2.2274E-01 | 6.6813E-02 | 3.0813E-01 |
| 2 | 7.2200E-02 | 5.5015E-04 | 7.2365E-02 |
| 3 | 3.3602E-02 | 8.3059E-03 | 4.1037E-02 |
| 4 | 8.5559E-03 | 1.7810E-02 | 2.6036E-02 |
| 5 | 1.2651E-01 | 1.1418E-02 | 1.4232E-01 |
| 6 | 6.2339E-06 | 2.2387E-02 | 2.2393E-02 |
| 7 | 6.3620E-04 | 2.2015E-02 | 2.2624E-02 |
| 8 | 5.0567E-03 | 1.9460E-02 | 2.4312E-02 |
| 9 | 3.7889E-02 | 2.9345E-03 | 4.0252E-02 |
| 10 | 2.0478E-02 | 9.7036E-03 | 2.9602E-02 |
| 11 | 2.9676E-02 | 5.4651E-03 | 3.4518E-02 |
| 12 | 1.5422E-04 | 2.2260E-02 | 2.2408E-02 |
| 13 | 1.1778E-02 | 1.3480E-02 | 2.4862E-02 |
| 14 | 5.5073E-03 | 1.7986E-02 | 2.3279E-02 |
| 15 | 1.6487E-02 | 1.6639E-02 | 3.2512E-02 |
| 16 | 1.8143E-01 | 7.7907E-03 | 1.9472E-01 |
| 17 | 7.6935E-02 | 2.6720E-03 | 7.8567E-02 |
| 18 | 6.6012E-02 | 1.2071E-02 | 7.6058E-02 |
| 19 | 9.3931E-02 | 8.8254E-03 | 1.0036E-01 |
| 20 | 9.3931E-02 | 8.8254E-03 | 1.0036E-01 |
| 21 | 7.5584E-02 | 4.2910E-05 | 7.5898E-02 |
| 22 | 1.9863E-01 | 9.7932E-02 | 3.1641E-01 |
| 23 | 6.8070E-01* | 3.8060E+00* | 4.9740E+00* |

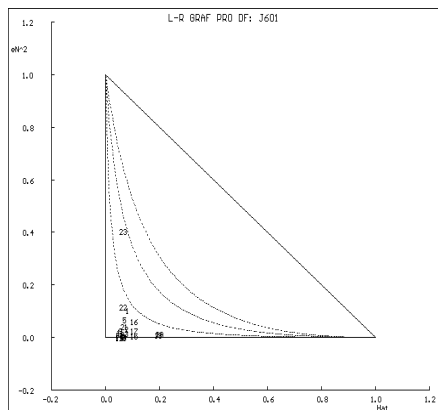
(c) **Grafy vlivných bodů** (obr. 6.1-2) jsou schopny indikovat přítomnost odlehlých hodnot a extrémů. *Graf predikovaných reziduí* ukazuje na odlehlý bod č. 23. *Pregibonův graf* ukazuje na silně vlivný bod č. 23. *Williamsův graf* indikuje č. 23 jako odlehlý bod a extrém č. 18, 19, 20. *McCullohův-Meeterův graf* dokazuje odlehlý bod č. 23, extrém č. 18, 19 a 20. Konečně *L-R graf* dokazuje odlehlý bod č. 23 a současně extrém č. 18, 19, 20. Lze uzavřít, že bod č. 23 je většinou diagnostik prokázán za odlehlý, a proto je vhodné ho dále analyzovat resp. z výběru vyloučit.



Obr. 6.1-2 Grafy vlivných bodů, vlevo graf predikovaných reziduí, a vpravo Pregibonův graf, *ADSTAT*.

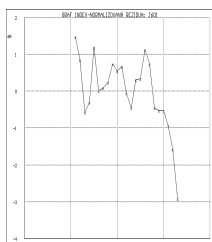
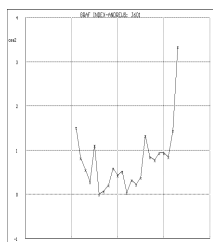


Obr. 6.1-2 Grafy vlivných bodů, vlevo Williamsův graf, a vpravo McCullohův-Meeterův graf, *ADSTAT*.

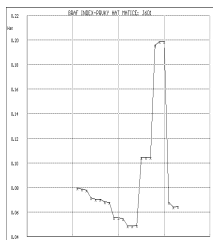


Obr. 6.1-2 Grafy vlivných bodů, L-R graf, *ADSTAT*.

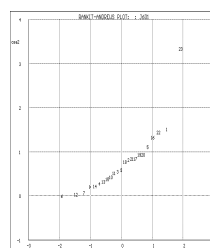
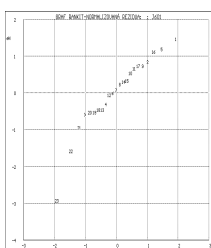
(d) *Indexové grafy* (obr. 6.1-3) upozorňují na *podezřelé body*, a kvantifikují velikost vlivu. *Andrewsův indexový graf* a *graf normovaných reziduí* ukazují na podezřelé body č. 1 a 23. *Indexový graf prvků H projekční matice* pak na extrémní č. 18, 19, 20.

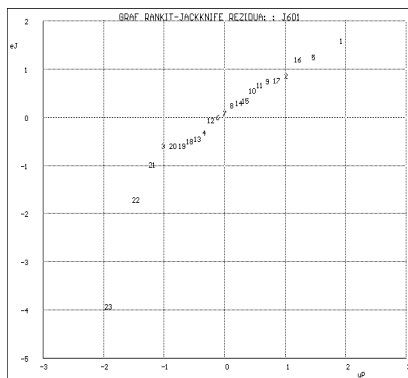
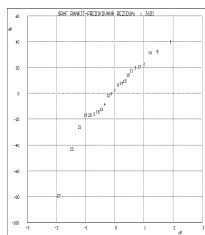


Obr. 6.1-3 Indexové grafy, vlevo, Andrewsův graf, a vpravo graf normovaných reziduí, *ADSTAT*.

Obr. 6.1-3 Graf prvků H-projekční matice, *ADSTAT*.

(e) *Rankitové grafy* (obr. 6.1-4) ukazují vedle normality rozdělení dotyčných reziduí i na vlivné (zde odlehlé) body. *Graf normovaných reziduí* ukazuje na začátku č. 23 a na konci č. 1 jako odlehlé body. *Andrewsův graf* ukazuje bod č. 23 jako odlehlý. *Graf Jackknife reziduí* ukazuje č. 23 jako odlehlý.

Obr. 6.1-4 Rankitové grafy, vlevo, graf normovaných reziduí, a vpravo, Andrewsův graf, *ADSTAT*.



Obr. 6.1-4 Rankitové grafy, vlevo, graf predikovaných reziduí, a vpravo, graf Jackknife reziduí, *ADSTAT*.

Model: *Parciální regresní grafy* a *parciální reziduální grafy* jsou určeny pro vícerozměrné lineární regresní modely a nemají proto u jednorozměrného regresního modelu smysl. Vhodnost modelu se posuzuje přímo v grafu, obsahujícím data a průběh modelové funkce. Je patrné, že v tomto případě je přímka akceptovatelná a data nevykazují nelineární průběh.

Metoda: do této části patří vyšetření splnění základních předpokladů metody nejmenších čtverců MŇČ čili statistické testy předpokladů, za kterých by měla metoda vést k nejlepšímu nestrannému lineárnímu odhadům regresních parametrů:

Fisherův-Snedecorův test významnosti regrese potvrdil, že navržený model je přijat jako významný, jinými slovy: závisle proměnná y a nezávisle proměnná x jsou silně lineárně závislé.

Scottovo kritérium multikolinearity nemá u jednorozměrného regresního modelu smysl.

Cookův-Weisbergův test heteroskedasticity ukazuje, že rezidua vykazují heteroskedasticitu (nekonstantnost rozptylu).

Jarqueův-Berraův test normality reziduí ukazuje, že klasická rezidua nevykazují Gaussovo, normální rozdělení.

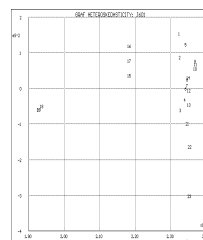
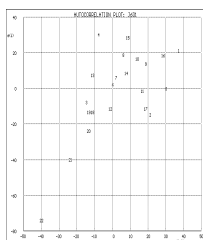
Waldův test autokorelace ukazuje, že klasická rezidua jsou autokorelována. To je vážné upozornění ke zhodnocení provedeného experimentu a dokazuje, že došlo k narušení podmínek, např. objevil se vliv nežádoucího konkurenčního děje (zanedbané proměnné), nevhodných podmínek měření na téže soustavě atd.

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

| TESTOVÁNÍ REGRESNÍHO TRIPLETU (DATA + MODEL + METODA): | |
|---|----------------|
| Fisherův-Snedocorův test významnosti regrese, F_{exp} | : 6.2543E+03 |
| Tabulkový kvantil, $F_{1-\alpha}(m-1, n-m)$ | : 4.3248E+00 |
| Závěr: Navržený model je přijat jako významný. | |
| Spočtená hladina významnosti | : 0.000 |
| Scottovo kritérium multikolinearity, M | : -4.07117E-15 |
| Závěr: Navržený model je korektní. | |
| Cookův-Weisbergův test heteroskedasticity, S_f | : 1.1084E+02 |
| Tabulkový kvantil, $\chi^2_{1-\alpha}(1)$ | : 3.8415E+00 |
| Závěr: Rezidua vykazují heteroskedasticitu. | |
| Spočtená hladina významnosti | : 0.000 |
| Jarqueův-Berraův test normality reziduí, $L(e)$ | : 7.3537E+00 |
| Tabulkový kvantil, $\chi^2_{1-\alpha}(2)$ | : 5.9915E+00 |
| Závěr: Normalita není přijata. | |
| Spočtená hladina významnosti | : 0.025 |
| Waldův test autokorelace, W_a | : 3.1791E+01 |
| Tabulkový kvantil, $\chi^2_{1-\alpha}(1)$ | : 3.8415E+00 |
| Závěr: Rezidua jsou autokorelována. | |
| Spočtená hladina významnosti | : 0.000 |
| Znaměkový test, D_t | : -2.5225E+00 |
| Tabulkový kvantil, $N_{1-\alpha/2}$ | : 1.6449E+00 |
| Závěr: Rezidua nevykazují trend. | |
| Spočtená hladina významnosti | : 0.006 |

Graf autokorelace (obr. 6.1-5) vykazuje přibližně mrak bodů bez výrazné orientace.

Graf heteroskedasticity (obr. 6.1-5) vykazuje klín, což odpovídá heteroskedasticitě, nekonzantnosti podmíněného rozptylu.



Obr. 6.1-5 Vlevo: graf autokorelace, a vpravo: graf heteroskedasticity, *ADSTAT*.

6. Konstrukce zpřesněného modelu:

(a) Po odstranění bodů č. 23 (*kritika dat*) byly nalezeny nové odhady parametrů zpřesněného modelu, když $t_{0,95}(22-2) = 2.086$.

| Parametr | Odhad | Směrodatná odchylka | Test $H_0: b_j = 0$ vs. $H_A: b_j \neq 0$ | | |
|----------|---------|---------------------|---|-------------------|---------------|
| | | | t-kritérium | hypotéza H_0 je | Hlad. význam. |
| b_0 | 26078.0 | 12.742 | 2046.7 | Zamítnuta | 0.000 |
| b_1 | -40.103 | 0.3930 | -102.04 | Zamítnuta | 0.000 |

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

$$y = 26\,078 (13) - 40.1 (0.4) x$$

je doložen statistickými charakteristikami: *Pearsonův párový korelační koeficient* r , *koeficient determinace* D a *predikovaný korelační koeficient* R_p dosáhly vesměs vysokých hodnot. *Střední kvadratická chyba predikce* MEP a *Akaikovo informační kritérium* AIC dosáhly nižších hodnot než u předešlého modelu, což dokazuje, že zpřesněný model je lepší.

| | |
|---|----------|
| Vícenásobný korelační koeficient, r | : 0.9990 |
| Koeficient determinace, 100 % D [%] | : 99.808 |
| Predikovaný koeficient determinace, R_p^2 | : 0.9989 |
| Střední kvadratická chyba predikce, MEP | : 414.22 |
| Akaikovo informační kritérium, AIC | : 132.62 |

Rezidua nyní vykazují normální rozdělení a nevykazují trend, stále však vykazují heteroskedasticitu, a proto lze doporučit použití metody vážených nejmenších čtverců.

(b) Užitím statistické váhy ($w_i = 1/y_i^2$) kompenzujeme heteroskedasticitu v datech. Obdržíme nové odhady parametrů, když $t_{0,95}(22-2) = 2.086$.

| Parametr | Odhad | Směrodatná odchylka | Test $H_0: b_j = 0$ vs. $H_A: b_j \neq 0$ | | |
|----------|---------|---------------------|---|-------------------|---------------|
| | | | t-kritérium | hypotéza H_0 je | Hlad. význam. |
| b_0 | 26079.0 | 12.666 | 2059.0 | Zamítnuta | 0.000 |
| b_1 | -40.110 | 0.3860 | -103.92 | Zamítnuta | 0.000 |

Opravený model má tvar (v závorce je odhad směrodatné odchylky parametru)

$$y = 26\,079 (13) - 40.1 (0.4) x$$

Jelikož došlo ke snížení rozhodujících kritérií, tj. střední kvadratické chyby predikce MEP a Akaikova informačního kritéria AIC , lze považovat tyto odhady za lepší než předešlé. *Pearsonův korelační koeficient* r , a tím pádem i koeficient determinace D vychází nepatrně lepší nebo stejný než u předešlého odhadu bez statistické váhy.

| | |
|---|----------|
| Vícenásobný korelační koeficient, r | : 0.9991 |
| Koeficient determinace, 100 % D [%] | : 99.815 |
| Predikovaný koeficient determinace, R_p^2 | : 0.9989 |
| Střední kvadratická chyba predikce, MEP | : 410.29 |
| Akaikovo informační kritérium, AIC | : 132.39 |

7. Zhodnocení kvality modelu: porovnáním hodnot regresní diagnostiky lze snadno provést zhodnocení *regresního tripletu* získaného lineárního regresního modelu pro upravená data, zbařená odlehých hodnot, metodou vážených nejmenších čtverců. Nalezený a prokázaný model teplotní závislosti přechodového tlaku bismutu má tvar,

(v závorce je vždy uveden odhad směrodatné odchylky parametru)

$$y = 26\,079 (13) - 40.1 (0.4) x$$