

# VLIV NEORTOGONALITY PLÁNU EXPERIMENTU NA STATISTICKOU KOREKTNOST MODELU

J. Morávka<sup>1</sup>, B. Maroš<sup>2</sup>, K. Michalek<sup>3</sup>

<sup>1</sup>Třinecký inženýring, a.s. Třinec, <sup>2</sup>VUT Brno, FSI, Ústav matematiky, <sup>3</sup>VŠB – Technická univerzita Ostrava, FMMI, Katedra metalurgie

## Abstrakt

Kvalitní příprava, realizace a analýza vědecko-výzkumných experimentů z laboratorních, či (polo)provozních pokusů předpokládá osvojení si určitého souboru poznatků a systematicky vybudované metodiky známé pod zkratkou *DOE* – Design Of Experiments.

V případě použití „klasického“ *regresního* přístupu k analýze dat experimentů, a to dokonce i pro původní (netransformovaná) data z experimentu sestaveného a realizovaného podle metodiky *DOE*, může dojít k různým nesrovnalostem a numerickým, statistickým i principiálně závažným problémům, s dopadem na znehodnocení celé experimentální práce.

Příspěvek na poměrně jednoduchém technickém příkladu tvařitelnosti oceli demonstruje výskyt a analýzu těchto nečekaných problémů. Interpretační problém se vyskytl při aplikaci *lineárního regresního modelu s interakcemi*. Podstata rozporu spočívala v nesprávné indikaci *statistické nevýznamnosti* regresního koeficientu technologicky i fyzikálně vlivného faktoru (kterým byla *teplota tváření*).

V příspěvku jsou rozebírány *numerické* a následně *statistické* aspekty zmíněného nesouladu. V závěru jsou doporučeny *zásady a přístupy* umožňující vyvarování se obdobných problémů, směřující k efektivnímu návrhu a zpracování experimentů se získáním adekvátního (přiměřeného, vhodného, použitelného) matematického modelu.

## 1 Úvod

Správný návrh, provedení a následná analýza (vyhodnocení) technologicko-technických experimentů nezbytně patří (nebo rozhodně by měla patřit) k základnímu „vybavení“ studentů i absolventů (nejen) technických univerzit, výzkumných i vědeckých pracovníků.

Zanedbání určitých principů plánování a provádění experimentů může vést ke znehodnocení experimentátorské práce. Cílem příspěvku je proto stručné seznámení s pojmy, principy a metodikou *plánování experimentů*, jako i upozornění a dokumentace možných problémů na srozumitelném příkladu.

## 2 Základní pojmy a principy plánování experimentů

Častým úkolem (nejen) technologicko-technické praxe je zjištění vazeb a vztahů mezi určitými veličinami zkoumaného procesu. Jde zvláště o případy, kdy proces je velice složitý a neexistuje pro jeho popis (dostatečně) vhodný matematicko-fyzikálně-chemický model.

### 2.1 Základní pojmy DOE

Základním a nejčastějším cílem experimentu je určení, zda určité *faktory* (ovlivňující, vstupní, vysvětlující veličiny) mají vliv na sledovanou (ovlivňovanou, výstupní, vysvětlovanou) veličinu, často nazývanou *odezva* (response). Další, či následující úlohou může být nalezení takové úrovně faktorů,

aby bylo dosaženo *optimum* (maximum, minimum) sledované veličiny.

Potřebná data pro sestavení modelu lze získat buď *pozorováním* veličin procesu bez cíleného zásahu do něj (tzv. *pasivní*, či neplánovaný experiment) anebo uskutečněním *experimentu* s cílenými zásahy do procesu (tzv. *aktivní*, či *plánovaný* experiment). Avšak ani v případě aktivních experimentů nemají vždy experimentátoři dostatek znalostí o principech efektivního provedení experimentů a vůle k jejich uskutečnění. Je třeba s politováním konstatovat, že *žádnou analýzou (žádnou metodou) experimentálních dat nelze obejít špatně či nedostatečně připravený experiment* [MAROŠ & TRÁVNÍČEK 2006].

Termínem *experiment* se označuje soustava *pokusů* (také měření, pozorování), která je v případě plánovaného experimentu *vhodným způsobem uspořádána*. V podstatě lze říci, že v plánovaném experimentu jde o vytvoření takových podmínek, aby rozsah experimentu byl co nejmenší, ale objem i forma informací co nejkvalitnější. V důsledku to také znamená, že u faktorů budeme vybírat jejich vhodné *úrovně* v rámci zvolených (technologicky vhodných) intervalů.

Požadavek na efektivitu se uplatňuje ještě mnohem výrazněji u experimentu s mnoha *faktory* (nezávislými proměnnými). Jeho důsledné respektování vedlo k vytvoření samostatného odvětví aplikované statistiky, tzv. plánovaného (řízeného) experimentu. Různé návrhy uspořádání *měření* a metody jejich *vyhodnocení* se souhrnně označují jako plánování (navrhování) experimentů. V literatuře se často označují jako *Design of Experiment (DOE)*.

## 2.2 Základní principy a přístupy metody DOE

Mezi základní principy a přístupy tvorby plánů experimentů podle metody DOE patří následující:

- **replikace:** je opakování měření při stejné úrovni nebo kombinaci úrovní faktorů. Tímto způsobem lze odhadnout nepřesnost měření (náhodnou složku), zvýšit spolehlivost závěru i díky možnosti použít test adekvátnosti (přiměřenosti, vhodnosti, použitelnosti) modelu,
- **rozdělení do bloků:** pro poměrně stejné podmínky experimentu za účelem odlišení dalšího zdroje variability,
- **znáhodnění:** aby nevznikla systematická chyba (např. stejným pořadím úrovní faktorů v každém bloku), provedeme jednotlivé pokusy (měření) experimentu v náhodném pořadí.

Celé schéma komplexního experimentu pak nazýváme *znáhodněné (replikované) bloky*. Ty dovolují rozložit celkovou variabilitu na složku odpovídající efektům úrovní zkoumaného faktoru, složku nepřesnosti měření, složku odpovídající blokům a reziduální (zbytkovou, nevysvětlenou) složku, jež zahrnuje vliv ostatních (nezahrnutých) činitelů.

Z důvodu minimalizace počtu pokusů v experimentu se používají (pouze) *dvou a tříúrovňové* experimenty, které mohou být *úplné* (faktorové) a *zkrácené*.

## 2.3 Možné problémy při nepoužití metodiky DOE

Pokud budeme spoléhat jen na tzv. „selský rozum“, či „expertní posouzení“ (což zvláště „láká“ u malého počtu, tj. u 2 a 3 faktorů i zkušené experimentátory), pak může (téměř určitě) dojít k několika nepříjemným skutečnostem, např.:

- špatně provedeme *výběr bodů* (hodnot pokusů, úrovní faktorů) s důsledkem *nesprávně indikované statistické nevýznamnosti* některých faktorů (přičemž tyto jsou významné a vlivné), a tím i ke stanovení nesprávného („okleštěného“) modelu,
- sestavíme nevhodný *návrh s přebytečnými údaji* pro model za „dobrým“ účelem získání maxima (ale ne všech potřebných) informací,
- *metoda „pokus-omyl“* jenom málokdy (když se „náhodou strefíme“) poskytne užitečné informace o závislostech,
- *metoda změny pouze jednoho faktoru* při dalších faktorech konstantních (tzv. jednofaktorový experiment, plán) vede jak k příliš (zbytečně) vysokému počtu pokusů (a s tím souvisejícího

dlouhého času a mnoho peněz na takovýto experiment), tak i k zásadní skutečnosti nemožnosti podchycení (často existujících) interakcí faktorů,

- použití *nehodné metody analýzy* (při správné metodice návrhu a provedení experimentu metodou DOE) – např. lineární regresní analýzy místo analýzy přístupem DOE – na *původní data* může také (neočekávaně a překvapivě) vést k výše uvedené *nesprávné indikaci statistické nevýznamnosti* některých (ve skutečnosti významných) faktorů s důsledkem na stanovení nesprávného (příliš zjednodušeného) modelu. **Právě tato překvapující nepříjemná skutečnost je dokumentovaná na níže uvedeném příkladu.**

Plánovaný experiment (DOE) poskytuje návod jak *současně měnit všechny faktory* na několika málo jejich úrovních najednou (z toho vyplývá *minimalizace* počtu pokusů v experimentu) při získání *maxima* informací pro výstavbu vhodného a použitelného modelu zkoumaného procesu.

### 3 Příklad

V literatuře [MAROŠ & TRÁVNÍČEK 2006] je uveden velice pěkný, jednoduchý, technicky zaměřený a srozumitelně rozpracovaný motivační *příklad* pro dokumentaci použití metody DOE k návrhu a vyhodnocení experimentů.

*Příklad:* Při tváření oceli za studena nebo za „poloohřevu“ klade materiál odpor, jenž se mění jak s velikostí deformace, tak s nastavením tvářecí teploty. Úkolem je nalézt vhodný model pro ocel 14240.3, jenž bude popisovat chování tohoto *přetvárného odporu* v závislosti na *stupni deformace* a *teplotě*.

Jeden faktor je tedy stupeň deformace  $\varphi$  [-], druhým faktorem je teplota  $t$  [°C]. Sledovanou veličinou bude přetvárný odpor  $\sigma$  [MPa]. Oblast zkoumání bude v prostoru, který je vymezen oběma faktory:  $\varphi \in \langle 0.0, 1.4 \rangle$ ,  $t \in \langle 20, 750 \rangle$ .

Chceme nalézt co nejjednodušší vztah  $\sigma = f(\varphi, t)$ , ale takový, aby statistická analýza potvrdila významnost tohoto vztahu.

V uvedené literatuře je ukázán *postup* výstavby modelu (v součinnosti s upravovanou či doplňovanou maticí plánu) od jednoduchého *lineárního bez interakcí*, přes *lineární model s interakcemi* až po *úplný kvadratický model*.

Cílem příspěvku je ukázat na možné numericko-statistické problémy právě *lineárního modelu s interakcemi*, který byl v daném případě uvažován ve tvaru (pro původní proměnné):

$$\sigma = b_0 + b_1 \cdot \varphi + b_2 \cdot t + b_3 \cdot \varphi \cdot t, \quad (1)$$

kde jsou  $b_j$  - regresní koeficienty,  $j = 0, 1 \dots p$ , kde  $p = k+1 = 3$  je počet prediktorů (regresorů, tj.  $\varphi, t, \varphi \cdot t$ ),  $k = 2$  je počet faktorů ( $\varphi, t$ ),  $b_0$  je tzv. absolutní (konstantní) člen.

Tento *absolutní* (polohový) *člen* je, v případě plánovaného experimentu podle metodiky DOE, průměrem naměřených výsledků odezvy (v daném případě přetvárného odporu  $\sigma$ ). Obecně je vhodné s ním vždy počítat (zařazovat jej do modelu), protože podchycuje vliv měřítka (fyzikálních hodnot) odezvy, současně i vliv případné nezařazené vysvětlující proměnné, či existující nelinearity a umožňuje korektní stanovení statistických ukazatelů (jakými jsou např. koeficient determinace modelu, Durbinova-Watsonova statistika apod.).

#### 3.1 Plán experimentu podle DOE

S ohledem na *minimalizaci* počtu pokusů (a tím i času a peněz na jejich provedení), lze uvažovat, že každý faktor má pouze *dvě* (krajní) úrovně. Ty lze pomocí *lineární transformace* převést na *DOE normované hodnoty* s úrovněmi  $-1$  a  $+1$ . Normované proměnné budou dále označovány symboly  $x_1, x_2, \dots, x_k$ .

*Plán experimentu* je matice  $\mathbf{X}$  tvořená ze sloupců normovaných faktorů, přičemž první sloupec odpovídá fiktivní proměnné  $x_0$ , pro niž platí  $x_0 \equiv 1$  a příslušný odpovídající regresní koeficient je  $b_0$ .

Žádný sloupec nesmí být tvořen pouze z nul. Řádky matice plánu představují (normované) hodnoty úrovní jednotlivých faktorů. Počet řádků matice plánu je roven počtu všech pokusů (měření), včetně opakovaných.

Pro možnost podchycení vlivu interakce budeme uvažovat **úplný faktorový plán** se dvěma úrovněmi ( $m = 2$ ), u kterého je počet pokusů závislý na počtu hlavních (základních) faktorů (bez interakce)  $k$  podle vztahu  $n = m^k = 2^k = 2^2 = 4$  (v analyzovaném případě máme pouze 2 základní faktory – stupeň deformace  $\varphi$  a teplota tváření  $t$ ).

Z důvodu podchycení *chyb měření* a umožnění *testování adekvátnosti modelu* mohou být přidány další **opakované pokusy** (v *krajních* bodech nebo ve středu plánu čili ve *středovém, centrálním* bodu, ve kterém mají všechny normované faktory hodnotu 0).

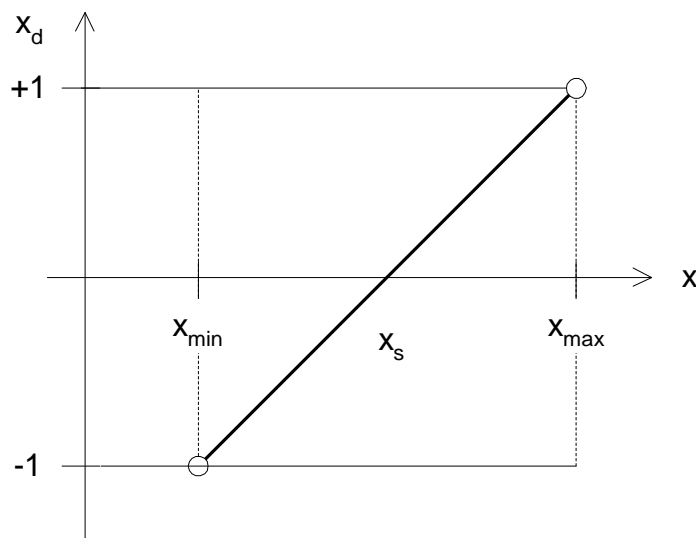
**Ortogonalní plán** je takový plán experimentu  $\mathbf{X}$ , ve kterém jsou všechny sloupce matice na sebe *kolmé* a jsou přitom *nenulové*.

**DOE normování (kódování)** základních faktorů vychází ze vztahu:

$$x_d(i) = \frac{x(i) - \frac{x_{\max} + x_{\min}}{2}}{\frac{x_{\max} - x_{\min}}{2}} \quad (2)$$

- kde je
- $x(i)$  - původní základní proměnné,  $i = 1..k$ ,  $k$  – počet základních faktorů,
  - $x_d(i)$  - normované proměnné podle metody DOE [-],
  - $x_{\max}$  - maximální hodnota původní proměnné  $x(i)$  [fyzikální jednotka],
  - $x_{\min}$  - minimální hodnota původní proměnné  $x(i)$  [fyzikální jednotka],

který představuje *lineární transformaci* (podle rovnice přímky (2)) hodnot původní proměnné z intervalu  $\langle x_{\min}, x_{\max} \rangle$  do intervalu  $\langle -1, +1 \rangle$  – viz *obr.1* (kde je symbolem  $x_s$  označena průměrná, střední hodnota):



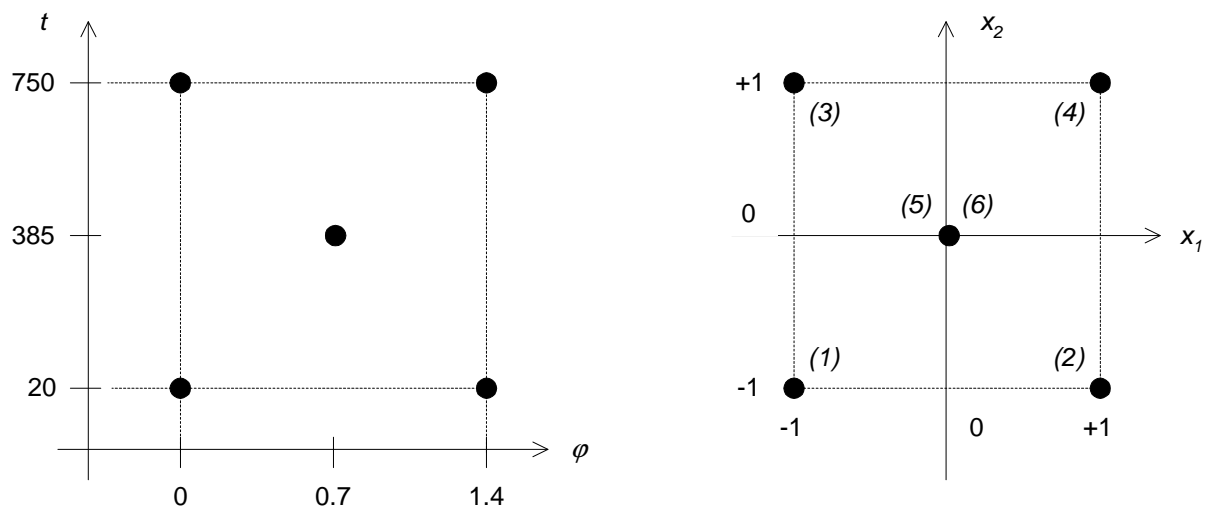
Obr. 1 : Lineární DOE normování

Ze vztahu (2) je zřejmé, že DOE normováním dochází také k převodu z původních fyzikálních jednotek faktorů na *bezrozměrový* tvar. Pro daný rozsah uvažovaných veličin vyjdou ze vztahu (2) normované základní veličiny (index  $d$  je kvůli přehlednosti vynechán):

$$x_1 = \frac{\varphi - \frac{\varphi_{\max} + \varphi_{\min}}{2}}{\frac{\varphi_{\max} - \varphi_{\min}}{2}} = \frac{\varphi - \frac{1.4 + 0}{2}}{\frac{1.4 - 0}{2}} = \frac{\varphi - 0.7}{0.7} \quad (3a)$$

$$x_2 = \frac{t - \frac{t_{\max} + t_{\min}}{2}}{\frac{t_{\max} - t_{\min}}{2}} = \frac{t - \frac{750^\circ\text{C} + 20^\circ\text{C}}{2}}{\frac{750^\circ\text{C} - 20^\circ\text{C}}{2}} = \frac{t[^\circ\text{C}] - 385^\circ\text{C}}{365^\circ\text{C}} = \frac{t - 385}{365} \quad [-] \quad (3b)$$

Z důvodu možnosti *testovat adekvátnost* (přiměřenost, vhodnost tvaru) *modelu* byly přidány ještě *opakovaná* dvě měření, a to ve *středovém* bodu [0.7, 385], což po transformaci (3a,b) přejde na bod [0, 0]. Původní a transformované hodnoty proměnných jsou názorně viditelné na *obr.2*:



Obr. 2 : Původní a transformované hodnoty plánu

Měření uskutečníme v bodech (DOE normovaných) faktorů v pořadí (1), (2) ... (6), tj. v bodech [-1,-1], [1,-1], [-1,1], [1,1], [0,0], [0,0].

Matice plánu  $\mathbf{X}$  s normovanými hodnotami úrovní faktorů v pokusech experimentu pro testovaný model tvaru (1) pak bude mít následující tvar (první sloupec odpovídá absolutnímu členu a poslední čtvrtý sloupec, obsahující *interakci 2. řádu*, vznikl násobením hodnot 2. a 3. sloupce, čili obou faktorů, tj.  $x_1 \cdot x_2$ ):

$$\mathbf{X} = \begin{pmatrix} +1 & -1 & -1 & +1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & +1 & +1 \\ +1 & 0 & 0 & 0 \\ +1 & 0 & 0 & 0 \end{pmatrix} \quad (4)$$

Pomocí skalárních součinů (viz vysvětlení níže) lze jednoduše ukázat, že tento *plán* je *ortogonální*, tzn., že sloupcové vektory jsou na sebe kolmé (pomocí tzv. *skalárního součinu* dvou vektorů, v daném případě kombinace dvojic sloupců, kdy je tento součin roven nule).

Zjištěné hodnoty  $\sigma$  z provedeného experimentálního měření jsou: 382, 900, 329, 308 a v centrálních bodech 520, 530 [MPa]. Ve vektorovém zápisu to lze zapsat pomocí řádkového vektoru naměřených hodnot  $\mathbf{y}$  následovně:

$$\mathbf{y} = (382 \ 900 \ 329 \ 308 \ 520 \ 530)^T \quad (5)$$

### 3.2 Analýza experimentu

Pro úplný faktoriální (faktorový) dvou-úrovňový plán (ale i pro další typy experimentů) je podrobný *postup* (algoritmus, popis) návrhu i *analýzy* uskutečněného experimentu *metodou DOE* uveden např. v literatuře [TOŠENOVSKÝ & NOSKIEVIČOVÁ 2000]. V této literatuře definovaným názorným postupem je možné „ručně“, „na papíře“ za pomoci kalkulačky, či elegantněji a obecněji

prostřednictvím např. tabulkového procesoru Excel, uskutečnit analýzu DOE a získat tvar i koeficienty modelu.

Je logické, že v dnešní době, je lepší (komplexnější, spolehlivější, rychlejší) použít nějaký vhodný SW, obsahující tuto metodu. Existuje speciální SW pro metodiku DOE známý pod anglickou zkratkou (akronymem) CADEX/DOE (Computer Aided Design and Analysis of Experiments / Design of Experiments), mezi něž patří např. programy *Design Expert* či *MODDE*.

Avšak asi nejčastěji mají uživatelé (studenti, výzkumníci a vědci) k dispozici obecné statistické (např. *Minitab*, *Statgraphics*, *JMP*, *Statistica*, *NCSS*), či matematické (*Matlab*) balíky programů, které lze pro dané účely použít.

Další vhodnou a přitom bezpečnou alternativou, kterou lze použít (spolehlivě jenom) pro *normované* hodnoty proměnných, je klasická **lineární regrese**, která je běžně dostupná v různých programech.

*Model* (lineární 1. řádu s interakcí 2. řádu) s DOE transformovanými proměnnými budeme potom uvažovat v analogickém tvaru jako model (1), tj.:

$$\tilde{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1 x_2, \quad (6)$$

kde je  $\tilde{y}$  - vypočtená (predikovaná) hodnota  $\sigma$  z modelu,  
 $x_1$  - transformovaná (normovaná) proměnná  $\varphi$ ,  
 $x_2$  - transformovaná (normovaná) proměnná  $t$ .

Neznámé koeficienty:

$$\mathbf{b} = (b_0 \quad b_1 \quad b_2 \quad b_3)^T, \quad (7)$$

tj. jejich odhad, lze určit ze známého vztahu, vycházejícího z *metody nejmenších čtverců* (MNČ) odchylek mezi původními a modelem stanovenými hodnotami odezvy:

$$\mathbf{b} = (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{X}^T \cdot \mathbf{y} = \mathbf{M}^{-1} \cdot \mathbf{X}^T \cdot \mathbf{y} = \mathbf{V} \cdot \mathbf{X}^T \cdot \mathbf{y}, \quad (8)$$

kde je  $\mathbf{M}$  - tzv. momentová matice,  
 $\mathbf{V}$  - tzv. varianční matice, inverze momentové matice  $\mathbf{M}$ .

Prakticky lze uvedený vektor regresních koeficientů  $\mathbf{b}$  velice elegantně a názorně určit (podle vztahu (8)) v systému *Matlab*, odkud dostaneme následující vektor odhadů regresních koeficientů:

$$\hat{\mathbf{b}} = \begin{pmatrix} +494.8333 \\ +124.2500 \\ -161.2500 \\ -134.7500 \end{pmatrix}. \quad (9)$$

Nicméně, pro ověření statistické významnosti modelu, jako i významnosti jeho regresních koeficientů, je nutné stanovit další hodnoty různých statistik a uskutečnit potřebné testy.

## 4 Regresní analýza příkladu

Pokud uskutečníme statistickou vícenásobnou lineární regresní analýzu pro *DOE transformovaná a původní* (netransformovaná) data uvedeného příkladu v systému *Matlab* (ale obecně i v libovolném jiném matematickém, statistickém, či ekonometrickém programu), pak sice dostaneme (po přepočtu, po zpětné transformaci) **numericky** stejné hodnoty regresních koeficientů, avšak jejich **statistická** významnost (statistické vlastnosti) se budou u obou modelů (pro oba druhy dat) lišit.

### 4.1 Statistické „pozadí“ hodnocení regrese

Před hodnocením statistických výstupů regrese je vhodné definovat vztahy vedoucí k jejich určení.

Skutečnost, že vypočtený regresní koeficient modelu má nenulovou hodnotu ještě neznamená, že je statisticky významně odlišný od nuly (větší nebo menší než nula). To lze posoudit až po určení jeho nepřesnosti (charakterizované směrodatnou odchylkou).

Rozptyl regresního koeficientu  $b_j, j = 0 \dots p$ , kde  $p$  je počet regresorů, se určuje ze vztahu:

$$s^2(b_j) = s^2(e) \cdot \mathbf{V}_{jj}, \quad (10)$$

kde je  $\mathbf{V}_{jj}$  - hlavní diagonála varianční matice  $\mathbf{V}$ , tj.  $\text{diag}(\mathbf{V})$ ,  
 $s^2(e)$  - celkový reziduální (zbytkový, nevysvětlený) rozptyl.

Tento rozptyl je možné vypočítat z reziduálního součtu čtverců (RSC), tj. ze součtu čtverců reziduí (zbytků, chyb modelu)  $e$ :

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}, \quad (11)$$

$$RSC = \mathbf{e}^T \cdot \mathbf{e}, \quad (12)$$

$$s^2(e) = \frac{RSC}{n - (p + 1)} = \frac{\mathbf{e}^T \cdot \mathbf{e}}{n - p - 1}, \quad (13)$$

kde  $n$  je počet naměřených hodnot (a  $p + 1$  je počet regresorů modelu včetně absolutního členu).  
*Směrodatná odchylka* regresního koeficientu je druhou odmocninou jeho rozptylu:

$$s(b_j) = \sqrt{s^2(b_j)}, \quad (14)$$

odkud pro testovou *t-statistiku* statistické významnosti příslušného regresního koeficientu platí poměr (podíl):

$$t(b_j) = \frac{b_j}{s(b_j)}. \quad (15)$$

Pokud platí nerovnost:

$$|t(b_j)| \geq t(1 - \alpha/2, n - p - 1), \quad (16)$$

kde  $t(1 - \alpha/2, n - p - 1)$  je kvantil Studentova *t-rozdělení*, pak zamítáme hypotézu o statistické nevýznamnosti („nulovosti“) regresního koeficientu.

Další možností hodnocení statistické významnosti regresního koeficientu je určit tzv. *p-hodnotu* (p-value), čili dosaženou hladinu pravděpodobnosti (významnosti) *t*-statistiky tohoto regresního koeficientu a porovnat ji se zvolenou „kritickou“ hladinou významnosti testování  $\alpha$  (většinou se volí hodnota 0.05).

## 4.2 DOE transformovaná data

S využitím výše definovaných vztahů byla v programu *Matlab* uskutečněna lineární regresní analýza s diagnostikou statistické významnosti regresních koeficientů nejprve pro DOE transformovaná (DOE normovaná) data.

Výstup vytvořeného m-skriptu *TCP08\_d.m* má obvyklý tvar používaný ve výstupech statistických programů – *tab.1*:

Tab. 1: VÝSLEDKY REGRESE PRO DOE NORMOVANÁ DATA

regresor	koef	sb	tb	p-hodnota
const	494.8333	15.2208	32.5103	0.0009
x1	124.2500	18.6416	6.6652	0.0218
x2	-161.2500	18.6416	-8.6500	0.0131
x1.x2	-134.7500	18.6416	-7.2284	0.0186

Pomocí funkce *regress* ze Statistics Toolboxu je možné ještě určit i hodnoty koeficientu determinace ( $R^2$ ), Fisherovy statistiky modelu jako celku (*F-stat*) a její *p-hodnoty* (*p-F*). Pro uvedená normovaná data platilo:  $R^2 = 0.9885$  (tj. 98.85 % rozptylu regresandu bylo vysvětleno modelem), Fisherovo *F-stat* = 57.1659 a *p-F* = 0.0172.

Z uvedených výstupů je zřejmé, že všechny regresní koeficienty (včetně absolutního členu), jako i hodnota Fisherovy statistiky modelu jsou statisticky významné na běžně volené hladině významnosti  $\alpha = 0.05$ , která je ve všech případech větší než *p-hodnoty*, tj. dosažené hladiny významnosti. Koeficient determinace  $R^2$  ukazuje, že daný model až na 98.85 % popisuje variabilitu odezvy (regesandu).

### 4.3 Původní data

Výstup vytvořeného m-skriptu *TCP08.m* pro původní (netransformovaná) data má v tomto případě obdobný tvar – *tab.2*:

Tab. 2: VÝSLEDKY REGRESE PRO PŮVODNÍ DATA

regresor	koef	sb	tb	p-hodnota
const	398.5354	36.7758	10.8369	0.0084
fi	380.5479	38.7074	9.8314	0.0102
t	-0.0726	0.0722	-1.0052	0.4207
fi.t	-0.5274	0.0730	-7.2284	0.0186

A opět pomocí funkce *regress* ze Statistics Toolboxu je možné určit i hodnoty koeficientu determinace ( $R^2$ ), Fisherovy statistiky modelu jako celku (*F-stat*) a její *p-hodnoty* (*p-F*). I pro původní data byly získány stejné hodnoty těchto statistických ukazatelů:  $R^2 = 0.9885$  (tj. 98.85 % rozptylu regresandu bylo vysvětleno modelem), Fisherovo *F-stat* = 57.1659 a jeho *p-F* = 0.0172.

Z výše uvedeného výstupu v *tab.2* je však překvapivě viditelné, že jako *statisticky nevýznamný* se jeví faktor *teplota tváření* (*t*), protože jeho dosažená hladina významnosti = 0.4207  $\gg$  0.050 =  $\alpha$ . Tato skutečnost odporuje výsledkům získaným pro DOE transformovaná data (i získaná vyhodnocovací metodou DOE), tak fyzikální podstatě a dlouholeté i praktické (technologické) zkušenosti.

Na druhé straně je jednoduše ověřitelné, že samotné hodnoty regresních koeficientů jsou číselně shodné s koeficienty získanými transformací koeficientů z modelu DOE.

Závěrem lze v tomto případě konstatovat, že získaný matematický model je pro daná (původní) data jako celek *významný*, ovšem s *nesprávně stanoveným nevýznamným vlivem teploty tváření*.

### 4.4 Shrnutí

*Porovnáním výsledků* obou přístupů (pro oba druhy dat) je jasné, že i když oba modely *stejně kvalitně* vysvětlují rozptyl odezvy (vykazují stejné hodnoty koeficientu determinace), *liší se* v hodnocení statistické (ne)významnosti vlivu (efektu, koeficientu) *teploty tváření* na přetvárný odpor oceli.

Otázka tedy zní, čím (jakou skutečností) je tato disproporce způsobena, což je obsahem následující kapitoly.

## 5 Rozbor statistického problému

Po zjištění příčin vzniklého problému (podle informací v odborné literatuře) byla definována používaná indikační kritéria.



## 5.1 Příčiny problému

Na základě informací z literatury, např. [MELOUN & MILITKÝ 1994], [HEBÁK & HUSTOPECKÝ 1987], lze zjistit, že *nesprávná indikace statistické nevýznamnosti regresoru* (vysvětlující proměnné) je důsledkem tzv. **multikolinearity** (téměř souběžnosti) několika **regresorů** (jejich vzájemné korelovanosti, souvislosti), či jinak řečeno, jejich **neortogonalita** (nekolmosti). Tato skutečnost způsobuje (kromě jiného) *velké rozptyly* jednotlivých odhadů regresních koeficientů, díky kterým výsledky *t-testů* indikují nesprávně statistickou nevýznamnost těchto koeficientů.

Je překvapivé, že i v tomto případě vychází *koeficient determinace* vysoký a regresní *model* může dobře *popisovat* (avšak už ne dobře a správně *vysvětlovat*) experimentální data.

Nesprávně určena *statistická* nevýznamnost vychází z *numerického* hlediska tzv. *špatné podmíněnosti* tzv. momentové matice  $\mathbf{M} = \mathbf{X}^T \cdot \mathbf{X}$  (kde  $\mathbf{X}$  je matice plánu).

## 5.2 Indikační kritéria neortogonalita

Pro testování uvedeného jevu existuje a je doporučeno několik kritérií, a to jak numerických, tak i statistických, přičemž *kritéria*:

- **numerická** vycházející ze spektrálního rozkladu momentové a (její inverzní) varianční matice se stanovením tzv. *vlastních čísel*, *stopy* a *determinantu* matic,
- **statistická** vycházejí z *korelační* matice, používá se u nich statistických pojmů, či pro hodnocení se využívá kvantilů pravděpodobnostních rozdělení.

Další obdobné členění těchto kritérií může být podle skutečnosti, z jaké *matice* regresorů se vychází, tj. zda se vychází z matice:

- *plánu*  $\mathbf{X}$ ,
- *momentové*  $\mathbf{M} = \mathbf{X}^T \cdot \mathbf{X}$ ,
- *variační*  $\mathbf{V}$ , která je inverzní maticí k matici momentové:  $\mathbf{V} = \mathbf{M}^{-1} = (\mathbf{X}^T \cdot \mathbf{X})^{-1}$ ,
- anebo z normované verze varianční matice  $\mathbf{R}$ , která je formálně shodná s *korelační maticí* vysvětlujících proměnných (regresorů).

Seznam vybraných a na datech příkladu použitých kritérií vychází z literatury [MELOUN & MILITKÝ 1994], [HEBÁK & HUSTOPECKÝ 1987], [GARAJ & ŠUJAN 1980].

U níže uvedeného přehledu jsou u kritérií definované *kritické hodnoty*, při jejichž překročení lze uvažovat o existenci *multikolinearity* (*neortogonalita*) regresorů:

### A. Kritéria numerická

- *skalární součiny* dvojic vysvětlujících proměnných (z matice plánu  $\mathbf{X}$ ): pokud nejsou všechny *nulové*, pak plán není ortogonální,
- *číslo podmíněnosti* momentové matice  $\mathbf{M}$ : číslo podmíněnosti = podíl největšího a nejmenšího vlastního čísla matice větší než  $10^6$ , přičemž pro každou nesignulární čtvercovou matici  $\mathbf{A}$  typu  $m \times m$  nebo  $(m, m)$  platí (kde  $\lambda_j, j = 1..m$ , jsou tzv. vlastní či charakteristická čísla matice):

$$|\mathbf{A}| = \det(\mathbf{A}) = \lambda_1 \cdot \lambda_2 \cdot \dots \cdot \lambda_m, \quad (17)$$

$$K = \frac{\lambda_{\max}}{\lambda_{\min}}, \quad (18)$$

- *stopa matice*  $\mathbf{M}$ : stopa (součet diagonálních prvků) větší než  $10^6$ ,
- *determinant* varianční matice  $\mathbf{V}$ : hodnota menší než  $10^{-6}$ ,
- *stopa matice*  $\mathbf{V}$ : hodnota větší než 1.0,
- *korelační matice*  $\mathbf{R}$ : pokud nejsou všechny mimodiagonální prvky (Pearsonovy párové korelační koeficienty) rovné nule, pak plán není nekorelovaný (není nezávislý),
- *číslo podmíněnosti*  $K$  matice  $\mathbf{R}$ : větší než  $10^3$ ,

- *determinant* matice  $\mathbf{R}$ : blízký nule.

### B. Kritéria statistická

- **VIF-faktory** (Variance Inflation Factor - inflační faktory rozptylu regresorů, viz např. [Hebák & Hustopecký 1987], [ZVÁRA 2008]): rovné jedné (prediktory nejsou korelované, plán je nekorelovaný a je *ortogonální*), větší než 1 ale menší než 5 (indikace mírné korelovanosti a neortogonalita plánu), větší jak 5 ale menší než 10 (významná korelovanost a neortogonalita), větší než 10 (multikolinearita, neortogonalita). V případě  $VIF > 1$  jsou odhady regresních koeficientů (číselně) správné, ale nejsou správné jejich *p-hodnoty*, které jsou stanoveny jako diagonální prvky inverzní korelační matice ( $j = 1 \dots p$ ):

$$VIF_j = D_{jj} = \text{diag}(\mathbf{D}) = \text{diag}(\mathbf{R}^{-1}), \quad (19)$$

- *testy diagonality* výběrové korelační matice  $\mathbf{R}$ , či jinak řečeno, nulové hypotézy, že tato matice je jednotková [Hebák & Hustopecký 1987]: větší než kritická hodnota kvantilu chí-kvadrát rozdělení s počtem stupňů volnosti odpovídajících počtu párových korelací pro  $p$  proměnných (symbol  $\ln$  označuje přirozený logaritmus,  $|\mathbf{R}|$  znamená determinant korelační matice):

$$W = -n \cdot \ln|\mathbf{R}| \sim \chi_{\alpha}^2(p \cdot (p-1)/2), \quad (20)$$

- častěji je však používána *korigovaná* (zpřesněná) hodnota *testu diagonality* na počet faktorů (regresorů) podle *Farrara a Glaubera* [Garaj & Šujan 1980]:

$$W_{FG} = -\left(n-1-\frac{2p+5}{6}\right) \cdot \ln|\mathbf{R}| = -n \cdot \left(1-\frac{2p+11}{6n}\right) \cdot \ln|\mathbf{R}| \sim \chi_{\alpha}^2(p \cdot (p-1)/2), \quad (21)$$

- *F-test* regresorů způsobujících multikolinearitu: pokud je větší než kritická hodnota kvantilu Fisherova F-rozdělení s  $(p-1)$  a  $(n-p)$  stupni volnosti:

$$FD_j = \frac{n-p}{p-1}(VIF_j - 1) \sim F_{\alpha}(p-1, n-p). \quad (22)$$

## 6 Způsoby řešení problému

Podle informací z literatury byly dále sestaveny návrhy způsobů řešení problému. Pro ověření navrhovaných opatření byla uskutečněna aplikace kritérií a způsobů řešení na data příkladu s hodnocením jejich detekční schopnosti a účinnosti.

V literatuře [MELOUN & MILITKÝ 1994], [HEBÁK & HUSTOPECKÝ 1987] a [TOŠENOVSKÝ & NOSKIEVIČOVÁ 2000] je doporučeno v případě výskytu multikolinearity použít **lineární transformaci** dat několika následujícími způsoby (kde index, pořadí pokusů  $i = 1, 2, \dots, n$ , kde  $n$  je celkový počet pokusů):

- **centrování** hodnot proměnných:

$$x_c(i) = x(i) - \bar{x}, \quad (23)$$

kde je  $x_c$  - centrovaná proměnná,  
 $x$  - původní (netransformovaná) proměnná,  
 $\bar{x}$  - střední hodnota (aritmetický průměr) proměnné  $x$ ,

- **normování** (má význam pro *kladnou* střední hodnotu proměnné):

$$x_n(i) = \frac{x(i) - \bar{x}}{\bar{x}}, \quad (24)$$

kde je  $x_n$  - normovaná proměnná,

- **DOE normování** (kódování) - viz předchozí vztah (2),
- **standardizace** (má význam pro nekonstantní proměnnou s *nenulovou, kladnou* směrodatnou odchylkou proměnné):

$$x_s(i) = \frac{x(i) - \bar{x}}{s_x}, \quad (25)$$

kde je  $x_s$  - standardizovaná proměnná,  
 $s_x$  - (výběrová nebo populační) směrodatná odchylka původní proměnné.

Všechny uvedené způsoby lineární transformace dat vedou k *nulové střední hodnotě* a *symetrii* transformovaných dat s příznivým dopadem na zajištění *ortogonalitu* a *nekorelovanosti* plánu s takovýmito proměnnými.

Nicméně, zde je třeba obecně konstatovat, že multikolinearitu regresorů (faktorů) nelze vždy uvedenými způsoby odstranit. Je to tehdy, když jsou faktory mezi sebou z principu závislé (např. délka a objem tělesa), při používání vyšších mocnin (které jsou navzájem závislé) proměnných, apod.

Pro analyzovaná data se rozmezí jejich hodnot pro původní i transformované proměnné pohybovalo v následujících intervalech – viz *tab.3*:

Tab. 3: ROZMEZÍ (MIN/MAX), INTERVALY HODNOT PROMĚNNÝCH

Data	$\phi(x_1)$	$t(x_2)$	$\phi \cdot t(x_1 \cdot x_2)$	Poznámka
Původní	$\langle 0, +1.4 \rangle$	$\langle +20, +750 \rangle$	$\langle 0, +1050 \rangle$	
Centrovaná	$\langle -0.7, +0.7 \rangle$	$\langle -365, +365 \rangle$	$\langle -255, +255 \rangle$	
Normovaná	$\langle -1, +1 \rangle$	$\langle -0.948, +0.948 \rangle$	$\langle -0.948, +0.948 \rangle$	uvnitř $\langle -2, +2 \rangle$
DOE normovaná	$\langle -1, +1 \rangle$	$\langle -1, +1 \rangle$	$\langle -1, +1 \rangle$	-, -
Standardizovaná	$\langle -1.225, +1.225 \rangle$	$\langle -1.225, +1.225 \rangle$	$\langle -1.5, +1.5 \rangle$	-, -

Prakticky to znamená, že *normování*, *DOE normování* a *standardizace* „stlačují“ hodnoty do intervalu  $\langle -2, +2 \rangle$ , což je doporučené rozmezí s ohledem na zaručení numerické stability výpočtů a omezení kumulačních chyb. Uvedené rozmezí poskytuje nejmenší odchylky od správného řešení, které bychom obdrželi, kdyby počítač uměl zapsat každé číslo na nekonečně dlouhou mantisu.

## 6.1 Ověření kritérií a způsobů řešení

Po transformaci původních dat a výpočtu kritérií v prostředí *Matlab* lze dostat výsledky, které jsou přehledně uspořádány v *tab.4a,b* (kde jsou v *bledě modrých* políčkách uvedeny „podezřelé“, od dalších odchylné hodnoty a navíc *tlustým písmem* jsou pak označeny hodnoty kritérií, které překročily kritické hodnoty a signalizují narušení *ortogonalitu* a *nekorelovanosti* plánu):

Tab. 4a: HODNOTY KRITÉRIÍ MULTIKOLINEARITY A KORELOVANOSTI DAT

Data	det(M)	det(V)	tr(M)	tr(V)	cond(M)	Poznámka
původní	1.636E+12	6.111E-13	<b>2 671 000</b>	2.051	<b>4 159 900</b>	Narušení
centrovaná	1.636E+12	6.111E-13	794 000	0.677	271 900	Lze
normovaná	310	0.003224	17.2	0.973	1.67	OK
DOE normovaná	384	0.002604	18.0	0.917	1.50	-, -
standardizovaná	1944	0.000514	27.0	0.611	1.50	-, -

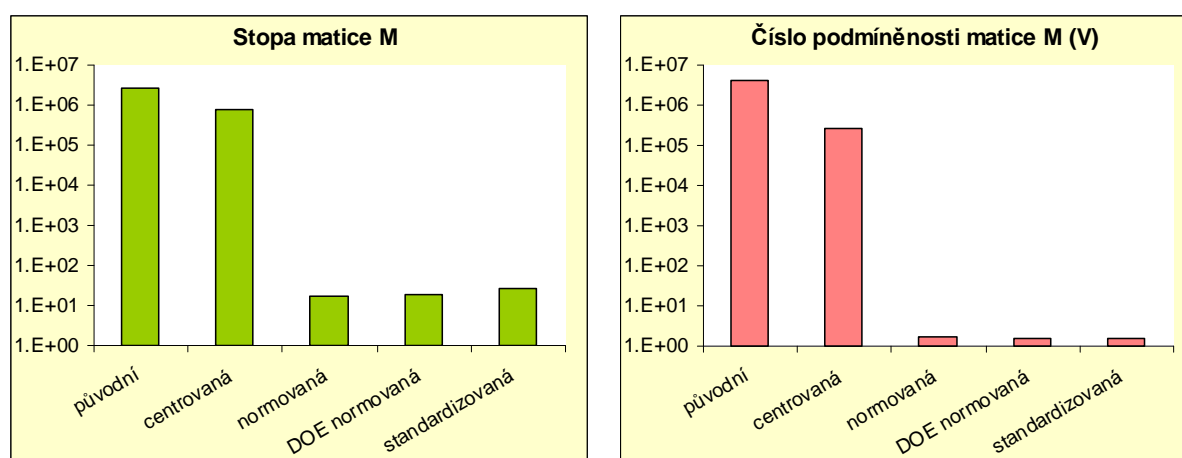
*Legenda:* *det()* ... determinant matice, *tr()* ... stopa (*trace*) matice, *cond* ... číslo podmíněnosti (*conditional number*) matice: platí, že  $cond(M) = cond(V)$ , a to díky lineární transformaci operace inverze, která sice změní hodnoty všech vlastních čísel původní matice, avšak jejich poměry (podíly) zůstanou zachovány.

Tab. 4b: HODNOTY KRITÉRIÍ MULTIKOLINEARITY A KORELOVANOSTI DAT

Data	det(R)	cond(R)	WFG	VIF			FD		
				x <sub>1</sub>	x <sub>2</sub>	x <sub>1</sub> ·x <sub>2</sub>	x <sub>1</sub>	x <sub>2</sub>	x <sub>1</sub> ·x <sub>2</sub>
původní	0.321	10.4	3.596	2.113	2.000	3.113	1.67	1.50	3.17
centrovaná	1	1	0	1	1	1	0	0	0
normovaná	1	1	0	1	1	1	0	0	0
DOE normovaná	1	1	0	1	1	1	0	0	0
standardizovaná	1	1	0	1	1	1	0	0	0

Legenda: WFG ... hodnota kritéria Farrara-Glaubera porovnávaná s kritickou hodnotou kvantilu chí-kvadrát rozdělení  $\chi^2_{\alpha(p \cdot (p-1)/2)} = \chi^2_{0.05(3)} = 7.815$ , FD ... hodnota Fisherova F-testu regresorů způsobujících multikolaritu porovnávaná s kvantilem Fisherova rozdělení  $F_{\alpha(p-1, n-p)} = F_{0.05(2,3)} = 9.552$ .

Hodnoty a porovnání ukazatelů *stopa matice M*, jako i číslo podmíněnosti matic *M* i *V* (jejichž podmíněnost je stejná) jsou pro původní a transformovaná data graficky znázorněny na obr.3:



Obr. 3 : Stopa matice *M* a číslo podmíněnosti matic *M* i *V*

Z obr.3 je zřejmé, že *centrování* částečně (a snad i dostatečně) snižuje hodnoty stopy matice *M* a číslo podmíněnosti obou matic. Výrazné (téměř o 5 řádů) je však toto snížení u dalších způsobů transformace, tj. u *normování*, *DOE normování* a *standardizace* proměnných.

## 6.2 Hodnocení kritérií a způsobů řešení

Z předchozích výsledků je jasné, že *neortogonalitu* (nekolmost) a *korelovanost* (*multikolaritu* a *špatnou podmíněnost*) matice plánu s *původními* proměnnými:

- signalizovaly (pouze) tři kritéria: *inflační faktory VIF* (větší než 1), *stopa matice M*, tj.  $\text{tr}(M)$  a (stejně) *číslo podmíněnosti matic M i V*, tj.  $\text{cond}(M) = \text{cond}(V)$ , které překročily kritickou mez (obvykle o hodnotě)  $10^6$ . Je však třeba říci, že *malé* číslo podmíněnosti sice zaručuje správný numerický výsledek, avšak *velké* číslo pouze *může* (ale nemusí) způsobit numerické problémy – tzn. malé číslo podmíněnosti je pouze *postačující* (PP), ale nikoliv *nutná* podmínka (NP) k numerické stabilitě a statistické správnosti výsledku výpočtu,
- nejvýrazněji (pro techniky a technology) signalizovala *nesprávná* a *nevěrohodná statistická nevýznamnost* vlivu *teploty tváření* na přetvárný odpor.

Jako *způsoby řešení* uvedeného problému se rýsují následující:

- použití metody *DOE*, která zajišťuje *ortogonalitu* a dobrou *numerickou podmíněnost* (malá čísla podmíněnosti díky transformovaným, normovaným, tzv. kódovaným datům) plánu experimentu s dopadem na *statistickou korektnost* a *spolehlivost závěrů*,

- při použití **regresní analýzy** (např. díky nemožnosti použít metodu DOE s příslušným SW) pro objektivní a spolehlivé rozpoznání uvedených nechtěných vlivů, a tím i možnosti výskytu numericko-statistické nekorektnosti indikace nevýznamnosti (některých) regresních koeficientů, je nutné a vhodné *porovnat* tuto významnost (pomocí t-statistik či p-hodnot) pro *modely s původními a transformovanými* daty. Výraznější nesouhlas (významnost versus nevýznamnost) pak signalizuje špatnou podmíněnost plánu. Hodnocení statistické významnosti koeficientů modelu je korektní pouze u modelu s *transformovanými* proměnnými, přičemž číselné hodnoty samotných regresních koeficientů jsou správné i pro model s původními proměnnými.

Obecně je zde viditelné, že doporučené způsoby *lineární transformace* původních dat vedou k odstranění multikolinearity díky ortogonalitě a nekorelovanosti plánů s *transformovanými* proměnnými a také k numericky i statisticky věrohodným výsledkům.

## 7 Statistický důsledek neortogonalit

Jak už bylo konstatováno, matice plánu s původními hodnotami regresorů je *neortogonální, korelovaná* a hlavně tzv. *špatně podmíněna*. Poslední jmenovaná skutečnost znamená velký podíl a tím rozdíl v hodnotách tzv. *vlastních čísel* matice plánu.

Pro původní proměnné a momentovou matici  $\mathbf{M}$  je minimální vlastní číslo  $\lambda_{\min}(\mathbf{M}) = 0.561$  a maximální vlastní číslo  $\lambda_{\max}(\mathbf{M}) = 2.335 \cdot 10^6$ . Pro její inverzní (čili varianční) matici  $\mathbf{V}$  je pak minimální vlastní číslo  $\lambda_{\min}(\mathbf{V}) = 0.0000004283$  (asi  $4 \cdot 10^{-7}$ ) a maximální vlastní číslo  $\lambda_{\max}(\mathbf{V}) = 1.782$ .

Pro „technickou“ názornost a srozumitelnost lze dodat, že pokud by matice plánu byla tzv. *maticí dynamického systému*, pak uvedený velký rozdíl (podíl) vlastních čísel charakterizuje tzv. *tuhý (stiff) dynamický systém*, který je problematicky říditelný a numericky (počítačově) simulovatelný, což předpokládá užití speciálních numerických metod řešení tzv. integrace soustavy diferenciálních rovnic. Fyzikální význam vlastních čísel u dynamických systémů spočívá ve skutečnosti, že to jsou *převrácené hodnoty časových konstant* (u kmitavého systému *period kmitů*) systému [ZÍTEK 1990]. Pokud tedy uvažujeme přímo vlastní čísla, které mají charakter vlastních kmitů (frekvencí), pak je logické, že stanovení vlastních čísel matice (systému) se nazývá jejím *spektrálním rozkladem*.

Při operacích se špatně podmíněnými maticemi rostou numerické chyby díky nutnému zaokrouhlení („ořezání“) čísel při omezené délce slova počítače. To se zvláště projeví při *inverzi momentové matice*  $\mathbf{V} = \mathbf{M}^{-1}$ , ze které následně vychází výpočet rozptylů i směrodatných odchylek regresních koeficientů a příslušné t-testy jejich významnosti.

Z výše uvedeného vztahu (10) pro výpočet rozptylu regresních koeficientů je vidět *přímý dopad nepřesně určené inverzní matice* (díky špatné podmíněnosti momentové matice vycházející z matice plánu) na následně nepřesně stanovený (většinou nadhodnocený) *rozptyl regresního koeficientu*. Špatná podmíněnost momentové matice *může* být způsobena až o 2,5 řádu *rozdílnými hodnotami* (úrovněmi) základních faktorů v matici plánu (střední hodnota  $\varphi = 0.7$ , střední hodnota  $t = 385$  °C).

Právě díky *nadhodnocení* odhadu směrodatné odchylky u regresního koeficientu teploty tváření  $t$  došlo k *podhodnocení* příslušné t-statistiky a nepřekročení kritické hodnoty s dopadem na nesprávnou signalizaci statistické nevýznamnosti tohoto vlivu (faktoru).

## 8 Závěry a doporučení

Na základě matematicko-statistického rozboru jednoduchého „technologického“ příkladu s lineárním modelem s dvojnou interakcí faktorů, kde se pro *původní* data nečekaně vyskytl numericko-statistický problém nesprávné indikace *statistické nevýznamnosti* jednoho regresního koeficientu technologicky i fyzikálně vlivného faktoru, a to *teploty tváření*, lze vyslovit *závěry a doporučení*. Tyto jsou v přehledné a shrnující tabulkové formě uvedeny v *tab.5* a pod ní popsány i slovně:

Tab. 5: PŘEHLED TYPŮ EXPERIMENTŮ A JEJICH VYHODNOCENÍ

Typ experimentu	Podtyp	ORT	MK	hrk	snvf	čpp	snkmv	Pozn.
Aktivní	DOE	ano	ne	OK	ne	m v	DOE Rtp	Nejlepší je DOE
	jiný	ne	ano	OK	ano	m/v	(RA – n)	
Pasivní (provozní)	sledování	ne	ano	???	ano	v	(RA – ?)	Nejhorší !

*Legenda:*

ORT ... ortogonalita (vzájemná kolmost) faktorů,  
 MK ... multikolinearita (souběžnost) faktorů,  
 hrk ... hodnoty regresních koeficientů,  
 snvf ... statistická nevýznamnost vlivných faktorů,  
 čpp ... číslo podmíněnosti plánu: m = malé, v = velké,  
 snkmv ... statisticky a numericky korektní metoda vyhodnocení experimentu,  
 Rtp ... regrese s transformovanými hodnotami proměnných (faktorů),  
 RA ... regresní analýza,  
 RA–n ... (pouze) numericky (ale ne statisticky) korektní,  
 RA–? ... obecně může být numericky i statisticky nekorektní (nesprávná).

Pro výběr typu experimentů a jejich způsobu vyhodnocování lze konstatovat následující skutečnosti:

- nedobře připravený (*aktivní*, ale ne *DOE* !) experiment (v laboratoři, v technologickém procesu) je *ztráta* času, peněz a důvěry ve vědu a výzkum,
- *žádnou analýzou* experimentálních dat nelze obejít špatně či nedostatečně připravený (aktivní) experiment !,
- pokud je *aktivní* experiment naplánován podle metody *DOE*, pak je vhodné a nutné z důvodu spolehlivosti výsledků jej také *touto metodou* (pro *DOE normované* proměnné) *analyzovat*. Díky ortogonalitě (a malému číslu podmíněnosti) plánu je možné *v jednom kroku* (jednorázově) správně (korektně) stanovit statisticky významné a nevýznamné faktory. Užití *regresní analýzy* pro *původní* (netransformované) proměnné může (díky *neortogonalitě* plánu) poskytnout statisticky i fyzikálně, či technologicky *nesprávné* hodnocení nevýznamnosti (některých) regresních koeficientů (v analyzovaném příkladu to byl vliv, tj. regresní koeficient *teploty tváření*), přičemž však jejich hodnoty jsou numericky správné. Proto pouze regresní analýza s *transformovanými* proměnnými poskytne i statisticky *správné* výsledky,
- pokud je experiment sice *aktivní*, plánovaný, ale *ne podle metodiky DOE*, pak jeho vyhodnocení pomocí *regresní analýzy* poskytne správné hodnoty regresních koeficientů faktorů, čísla podmíněnosti jsou (většinou) malá, ale díky *neortogonalitě* plánu a určité multikolinearitě faktorů dochází ke statistické nekorektnosti jejich hodnocení, tj. k nesprávné indikaci statistické nevýznamnosti fyzikálně či technologicky vlivných faktorů, a to jak pro jejich *původní*, tak i pro *transformované* hodnoty. Tato skutečnost znamená, že není k dispozici žádná statisticky korektní metoda vyhodnocení experimentu, ale pouze metoda numericky korektní, kterou je *regresní analýza* (analýzu metodou *DOE* nelze v tomto případě použít),
- tzv. *pasivní* experiment je nejhorší variantou pro analýzu: plán je (zásadně) *neortogonální*, faktory jsou kolineární, podmíněnost je velká, a tak lze (asi výrazněji a častěji pro větší rozsah dat a větší počet faktorů) očekávat při analýze pomocí *regresní analýzy* jak numerickou nestabilitu (díky numerickým chybám a jejich kumulaci), tj. *numericky nepřesné* a až *nesprávné* hodnoty (některých) regresních koeficientů, tak i jejich *nesprávně* indikovanou *statistickou nevýznamnost*, což znamená statistickou nekorektnost výsledků,
- metoda *DOE* poskytuje *spolehlivé* a *efektivní* (optimální) přístupy k návrhu a analýze přínosných experimentů, jako i *numericky* a *statisticky korektní* výsledky. Proto je velice potřebné, aby vědci, výzkumníci, technici i technologové, ale i studenti tuto metodiku znali a zásadně ji používali.

## Literatura

- [1] MAROŠ, B. & TRÁVNÍČEK, T. 2006. *Plánování experimentu*. In Sborník 5th International Conference *Aplimat 2006*, Bratislava : Ústav matematiky FSI STU Bratislava. 20 s.
- [2] TOŠENOVSKÝ, J. & NOSKIEVIČOVÁ, D. 2000. *Statistické metody pro zlepšování jakosti*. Ostrava: Montanex, 2000. 362 s. ISBN 80-7225-040-X.
- [3] MELOUN, M. & MILITKÝ, J. 1994. *Statistické zpracování experimentálních dat*. Praha : PLUS, 1994. 839 s. ISBN 80-85297-56-6.
- [4] HEBÁK, P. & HUSTOPECKÝ, J. 1987. *Vícerozměrné statistické metody s aplikacemi*. Praha : SNTL/ALFA, 1987. 456 s. ISBN 80-01-01076-7.
- [5] GARAJ, V. & ŠUJAN, I. 1980. *Ekonometria*. Bratislava : ALFA/SNTL, 1980. 288 s.
- [6] ZÍTEK, P. 1990. *Simulace dynamických systémů*. 1. vyd. Praha : SNTL, 1990. 420 s. ISBN 80-03-00330-X.
- [7] ZVÁRA, K. 2008. *Regrese*. 1.vyd. Praha : Matfyzpress MFF UK Praha, 2008. 253 s. ISBN 978-80-7378-041-8.

---

### **Ing. Jan Morávka, Ph.D.**

739 61 Třinec – Staré město, Frýdecká 126, e-mail: [jan.moravka@tzi.trz.cz](mailto:jan.moravka@tzi.trz.cz), tel.: 558 53 2192

### **Doc. RNDr. Bohumil Maroš, CSc.**

616 69 Brno, Technická 2896/2, e-mail: [maros@fme.vutbr.cz](mailto:maros@fme.vutbr.cz)

### **Prof. Ing. Karel Michalek, CSc.,**

708 33 Ostrava – Poruba, 17. listopadu 15, e-mail: [karel.michalek@vsb.cz](mailto:karel.michalek@vsb.cz), tel.: 59 732 5213