

Možnosti využití polynomu 1. stupně pro regresní vztahy mezi základními kvalitativními ukazateli hnědých uhlí

Möglichkeiten der Nutzung des Polynoms des 1. Stufens für Regressionsbeziehungen zwischen grundlegenden Qualitätsmerkmalen von Braunkohlen

Als wichtige Qualitätsmerkmale der Braunkohle gilt der Inhalt nichtverbrennbarer Stoffe (Aschestoffe), der analytisch als Aschegehalt bewertet ist, und der Wasserinhalt. Beide diese Parameter hängen mit der scheinbaren Dichte der Kohlemasse zusammen. Die Findung der gegenseitigen, genügend engen Regressionsbeziehungen zwischen diesen Kennziffern erlaubt eine Beschränkung der Laborprobeanalysen bei der Sicherstellung der notwendigen Informationen über die Qualität des bewerteten Rohstoffes sowie auch eine genauere Führung des Aufbereitungsprozesses. Eine Beurteilung der Tauglichkeit des Polynoms 1. Stufens als entsprechende Regressionsfunktion für Interpretierung der gegenseitigen Abhängigkeit zwischen o.g. Kennziffern ist in diesem Beitrag dargestellt.

Possibilities of Using a 1st Level Polynomial for Regression Relations among Basic Brown Coal Quality Indicators

The ash content (incombustible matter), analysed as ash and water contents, are significant brown coal quality indicators. Both quality parameters relate to apparent density of coal. Discovering common, significant regressive relations between these indicators would limit the sample laboratory analyses required when supplying necessary information about the quality of the assessed raw material and allow stricter control of the finishing process. Assessment of 1st level Polynomial suitability as a suitable regressive function to express the relationship between the indicators above stated is the subject of this article.

Возможности использования полинома 1-ой степени для регрессивных отношений между основными показателями качества бурых углей

К значительным показателям качества бурых углей относятся содержание негорючих веществ (зола), аналитически оцениваемое как содержание золы (зольность) и содержание воды. Оба приведенные параметры качества связаны с кажущейся плотностью угольной массы. Найдение взаимных, достаточно плотных, регрессивных отношений между этими показателями позволяет сокращение числа лабора-торных анализов проб при обеспечении требуемой информации о качестве оцениваемого сырья и более точное управление эвентуальными процессами обогащения.

Обсуждение пригодности полинома 1-ой степени в качестве соответствующей регрессивной функции для выражения взаимозависимости между приведенными выше показателями является предметом данной статьи.

Možnosti využití polynomu 1. stupně pro regresní vztahy mezi základními kvalitativními ukazateli hnědých uhlí

Významnými kvalitativními ukazateli hnědé uhlí jsou obsah nespalitelných látek (popelovin), hodnocený analyticky jako obsah popela (popelnatost) a obsah vody. Oba uvedené kvalitativní parametry souvisí se zdánlivou hustotou uhelné hmoty. Nalezení vzájemných, dostatečně těsných, regresních vztahů mezi těmito ukazateli dovoluje omezení laboratorních analýz vzorků při zajištění potřebných informací o kvalitě hodnocené suroviny i přesnější řízení event. úpravnického procesu. Posouzení vhodnosti polynomu

1. stupně, jako vyhovující regresní funkce pro vyjádření vzájemné závislosti mezi výše uvedenými ukazateli, je obsahem tohoto příspěvku.

1. Úvod

Obsah popela a vody patří mezi základní kvalitativní parametry hnědého uhlí. Na složení uhelné hmoty zkoumané částice lze pak usuzovat podle úrovně její zdánlivé hustoty. Nalezení regresních vztahů mezi výše uvedenými parametry uhelné hmoty pak při dostatečné úrovni korelačních koeficientů umožní snížení počtu laboratorních analýz k zabezpečení odpovídající úrovně poznatků o hodnocené uhelné hmotě. Těsnou funkční závislost mezi kvalitativními parametry a zdánlivou hustotou lze také efektivně využít k řízení případného separačního procesu. Uvedený postup je využíván v metodě mikrovzorků dle VÚHU pro stanovení upravitelnosti uhlí [1].

2. Podstata problému a postup řešení

Dosavadní výzkumy ukázaly, že zformování dostatečně přesných regresních vztahů mezi sledovanými parametry v širokém rozmezí platnosti ztěžuje proměnlivá struktura uhelné hmoty a její různorodé petrografické složení. Vlastnosti petrografických složek, tvořících uhelnou hmotu, jsou závislé na vlastnostech původního rostlinného materiálu, ze kterého vznikly a na podmínkách a stupni jeho přeměny. Složení uhelné hmoty jednotlivých kusů a jejich vlastnosti jsou pak důsledkem různého stupně jejich promíchání. Vedle petrografických složek se na stavbě uhelné hmoty podílejí také popeloviny a voda. Rozdílné fyzikální a chemické vlastnosti petrografických složek i doprovodných minerálů jsou příčinou rozdílné kvality (popelnatosti) uhelných částic při stejné zdánlivé hustotě [5]. Proto je většinou nezbytné sestavit potřebné regresní závislosti mezi zkoumanými parametry vždy zvlášť pro různé části dobývacího prostoru, aby na jejich základě vypočtené hodnoty byly dostatečně přesné. Dosud byla analyzována řada typů regresních vztahů od lineárních až po polynomiální. Vzhledem k jednoduchosti a operativnosti při používání a výpočtech je tento příspěvek věnován analýze regresního vztahu ve formě polynomu 1. stupně.

K posouzení těsnosti vzájemných regresních vztahů mezi zdánlivou hustotou, popelnatostí a obsahem vody u hnědých uhlí bylo využito údajů 11 souborů dat získaných vyhodnocením 11 jádrových vrtů dorozvědkové etapy průzkumu vybrané hnědouhelné sloje. Těsnost vzájemných vztahů sledovaných veličin v závislosti vyjádřené polynomem 1. stupně byla ověřována samostatně na všech jedenácti souborech a současně také na souboru všech sledovaných údajů.

Ke specifikaci konstant polynomu 1. stupně (lineární vztah) pro funkční závislosti $A^d = f(d_a^r)$, $W_t^r = f(d_a^r)$ a $W_t^r = f(A^d)$ byla použita tzv. příмка odhadu v obecném tvaru

$$y = b_0 + b_1 x \quad (1)$$

Konstanty b_0 a b_1 byly vypočteny za použití metody nejmenších součtů čtverců řešením tzv. normálních rovnic ze vzorců

$$b_1 = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad (2)$$

$$b_0 = \frac{\sum y_i - b_1 \sum x_i}{n} = \bar{y} - b_1 \bar{x} \quad (3)$$

kde n počet údajů zkoumaného parametru v souboru

x_i hodnoty nezávisle proměnné veličiny

y_i hodnoty závisle proměnné veličiny

$\bar{x}$ střední hodnota nezávisle proměnné veličiny

$\bar{y}$ střední hodnota závisle proměnné veličiny

Úroveň korelace naměřených parametrů podle přímkové závislosti byla hodnocena výběrovým koeficientem korelace r_{xy} určeným pro všechny analyzované soubory ze vzorců

$$r_{xy} = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{\left[n \sum x_i^2 - (\sum x_i)^2 \right] \left[n \sum y_i^2 - (\sum y_i)^2 \right]}^{1/2} \quad (4)$$

$$r_{xy} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2 \right]^{1/2}} \quad (5)$$

Číselný údaj koeficientu r_{xy} pak charakterizuje těsnost a přesnost vypočtené lineární závislosti mezi oběma sledovanými veličinami [1], [3], [4].

3. Výsledky analýzy

Údaje získané vyhodnocením skupiny vrtů dorozvědkové etapy průzkumu uhelné sloje a použité k sestavení polynomu 1. stupně pro závislost mezi zdánlivou hustotou (d_a^r), popelnatostí (A^d) a obsahem vody (W_t^r) v uhelné hmotě hnědé uhlí jsou pro soubory č. 1 - 8 postupně seřazeny v tabulce 1 [2]. Pro všechny soubory (1 - 11) byly vyčísleny konstanty b_0 a b_1 přímek odhadu a pomocí nich sestaveny vzorce polynomů pro všechny varianty vzájemných vztahů. Rovnice přímek odhadu pro regresní vztahy $A^d = f(d_a^r)$ jsou uvedeny v tabulce 2, sloupec 1, obdobné rovnice pro funkční závislost $W_t^r = f(d_a^r)$ ve sloupci 2 téže tabulky a pro funkční závislost

Tab. 1 Údaje použité k sestavení lineárních regresních závislostí

Soubor č.	Zdánlivá hustota d^r (kg.l^{-1})	Obsah popela A^d (%)	Obsah vody w_t^r (%)
1	1,178	6,54	45,53
	1,237	17,03	40,17
	1,341	31,64	33,56
	1,452	44,52	30,70
	1,552	53,70	27,65
	1,665	65,43	25,80
	2,002	85,74	18,48
2	1,174	6,88	48,70
	1,241	17,84	49,64
	1,343	33,77	33,03
	1,431	43,44	32,34
	1,567	55,37	26,23
	1,648	62,34	23,64
	1,765	70,18	19,94
	1,857	74,31	20,65
3	1,066	6,51	48,12
	1,169	6,86	47,75
	1,235	18,85	39,12
	1,328	33,06	33,53
	1,440	47,95	29,84
	1,546	53,08	26,00
	1,738	72,05	22,89
	1,821	73,39	20,69
4	1,075	3,64	49,87
	1,175	9,23	44,19
	1,233	20,31	37,19
	1,340	33,53	37,36
	1,436	46,15	29,22
	1,559	57,37	25,39
	1,661	65,44	26,43
	1,764	69,95	22,66
	1,856	77,56	19,52
	2,064	88,42	17,94
5	1,035	3,06	44,95
	1,077	5,52	44,21
	1,133	8,08	42,09
	1,175	9,01	40,92
	1,226	18,04	35,63
	1,271	26,40	35,55
	1,336	34,73	31,30
	1,455	48,62	26,42
	1,538	55,55	25,20

5	1,641	61,15	22,11
	1,763	70,57	21,02
	1,877	76,56	18,95
	1,974	86,13	18,08
	2,051	89,50	14,15
6	1,142	5,84	45,27
	1,178	6,50	44,74
	1,222	14,93	38,93
	1,271	23,72	37,01
	1,341	33,26	32,41
	1,457	47,53	27,41
	1,545	56,82	27,12
	1,640	61,04	24,26
	1,745	69,58	22,85
	1,856	72,17	24,13
	1,953	80,77	18,90
7	1,081	4,15	44,42
	1,140	5,79	43,21
	1,180	6,36	42,83
	1,223	15,53	38,19
	1,274	16,37	33,88
	1,336	32,56	31,79
	1,446	46,45	28,86
	1,541	55,25	27,58
	1,643	59,34	24,14
	1,740	69,72	21,68
	1,839	70,49	19,62
	1,966	78,88	17,78
	2,079	88,89	15,20
8	1,038	3,03	45,89
	1,076	4,22	44,68
	1,132	5,71	43,43
	1,177	6,01	42,83
	1,222	15,79	39,28
	1,273	24,26	34,06
	1,338	34,07	31,72
	1,443	45,67	27,58
	1,533	53,37	27,65
	1,668	64,06	22,40
	1,714	69,35	22,30
	1,838	74,19	21,41
	2,159	82,12	12,51

Soubor č.	Regresní lineární závislosti pro vztah		
	$A^d = f(d^r)$	$W_t^r = f(d^r)$	$W_t^r = f(A^d)$
1.	$A^d = -89,990 + 88,710 d^r$	$W_t^r = 75,990 - 29,570 d^r$	$W_t^r = 45,841 - 0,325 A^d$
2.	$A^d = -70,350 + 75,910 d^r$	$W_t^r = 88,580 - 38,520 d^r$	$W_t^r = 49,171 - 0,407 A^d$
3.	$A^d = -81,20 + 82,440 d^r$	$W_t^r = 87,220 - 37,889 d^r$	$W_t^r = 48,651 - 0,389 A^d$
4.	$A^d = -89,723 + 90,274 d^r$	$W_t^r = 79,365 - 31,903 d^r$	$W_t^r = 47,883 - 0,358 A^d$
5.	$A^d = -90,025 + 90,175 d^r$	$W_t^r = 73,930 - 29,897 d^r$	$W_t^r = 44,187 - 0,334 A^d$
6.	$A^d = -97,872 + 94,725 d^r$	$W_t^r = 76,152 - 30,253 d^r$	$W_t^r = 45,221 - 0,327 A^d$
7.	$A^d = -92,894 + 90,179 d^r$	$W_t^r = 74,463 - 29,703 d^r$	$W_t^r = 43,766 - 0,327 A^d$
8.	$A^d = -83,933 + 84,496 d^r$	$W_t^r = 76,409 - 31,034 d^r$	$W_t^r = 45,324 - 0,360 A^d$
9.	$A^d = -98,747 + 98,231 d^r$	$W_t^r = 78,719 - 32,487 d^r$	$W_t^r = 46,389 - 0,339 A^d$
10.	$A^d = -86,766 + 87,862 d^r$	$W_t^r = 75,841 - 29,655 d^r$	$W_t^r = 46,011 - 0,326 A^d$
11.	$A^d = -108,051 + 103,759 d^r$	$W_t^r = 69,786 - 25,511 d^r$	$W_t^r = 52,316 - 0,464 A^d$
1 – 11	$A^d = -86,766 + 87,862 d^r$		

Tab. 2 Rovnice polynomů 1. stupeň pro soubory 1 – 11

$W_t^r = f(A^d)$ pak ve sloupci 3 tabulky 2. V posledním řádku tabulky jsou uvedeny tvary polynomu pro všechny zkoumané údaje (soubory 1 - 11). Hodnoty výběrových koeficientů korelace, vypočtené podle vztahů (4) a (5), jsou seřazeny v tabulce 3. Graficky jsou regresní závislosti znázorněny přímkami s nejvyšší a nejnižší hodnotou koeficientu korelace a přímkou vyjadřující tentýž vztah pro všechny soubory na obr. 1 - 3. Těsnost nalezených polynomů s naměřenými hodnotami byla ověřována rovněž pomocí součtu hodnot odchylek mezi naměřenými a vypočtenými údaji závisle proměnné veličiny a jejich středních hodnot. Příslušné údaje jsou pro všechny regresní vztahy uvedeny v tabulce 4. V rámci analýzy údajů získaných v souboru bylo provedeno rovněž ověření těsnosti lineární regrese mezi zdánlivou hustotou (d_a^r) a obsahem popela (A^d) pomocí zvolených středních hustot pro po sobě následující rozmezí hustot 0,1 kg.l-1 a jím příslušných středních hodnot obsahu popela. Hodnoty středních hustot a jím příslušných údajů o obsahu popela jsou uvedeny v tabulce 5. Vyčíslením konstant b_0 a b_1 byl získán polynom 1. stupně ve tvaru

$$A^d = -83,225 + 85,370 d_a^r \quad (6)$$

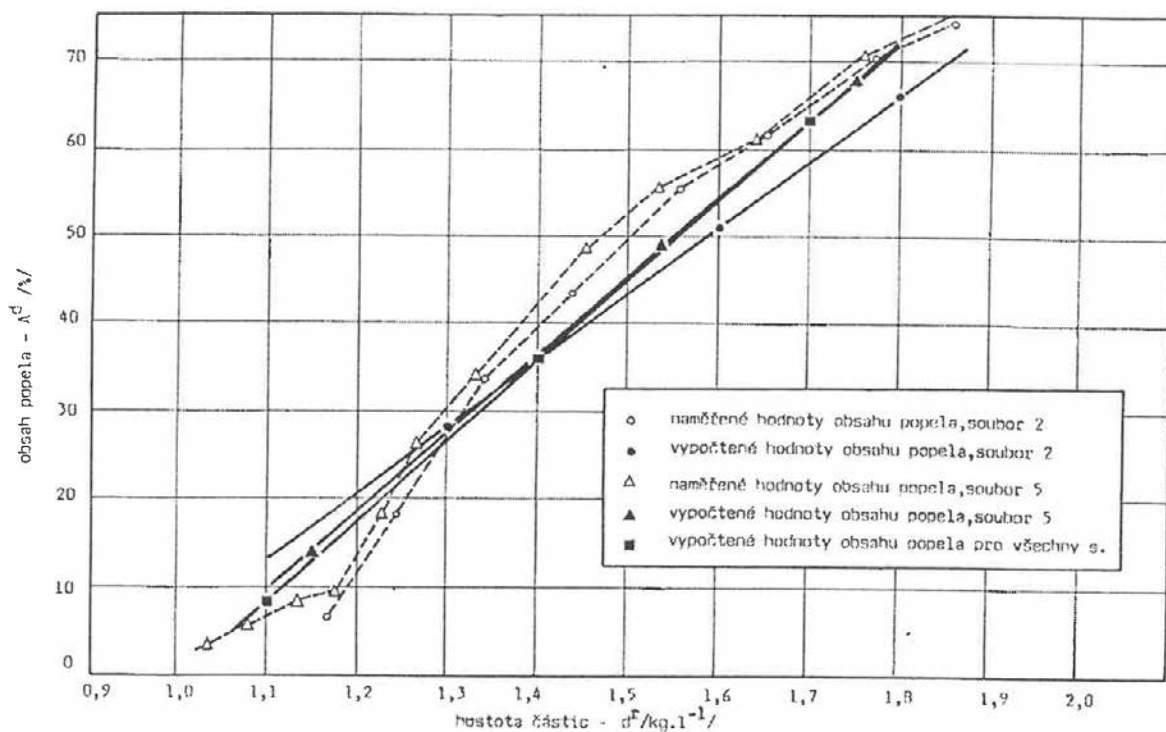
a korelační koeficient $r_{xy} = 0,983$. Součet odchylek hodnot obsahu popela vypočtených podle (6) a naměřených ve vrtech (soubory 1-11) má hodnotu $\Sigma\Delta = -0,194$, střední odchylku určenou z lineárních hodnot v úrovni $\Delta = \pm 5,157$ a z kvadratických hodnot odchylek ve výši $\Delta = \pm 5,991$.

Tab. 3 Hodnoty korelačních koeficientů pro regresní vztahy $A^d = f(d^r)$, $W_t^r = f(d^r)$, $W_t^r = f(A^d)$

Soubor č.	Hodnoty korelačních koeficientů pro regresní vztahy $A^d = f(d^r)$, $W_t^r = f(d^r)$, $W_t^r = f(A^d)$		
	$A^d = f(d^r)$	$W_t^r = f(d^r)$	$W_t^r = f(A^d)$
1.	0,980	-0,970	-0,990
2.	0,948	-0,954	-0,988
3.	0,950	-0,961	-0,988
4.	0,985	-0,954	-0,981
5.	0,990	-0,974	-0,990
6.	0,978	-0,941	-0,985
7.	0,977	-0,986	-0,985
8.	0,964	-0,975	-0,986
9.	0,977	-0,951	-0,976
10.	0,975	-0,982	-0,971
11.	0,983	-0,963	-0,902
1 - 11	0,978	-0,959	-0,975

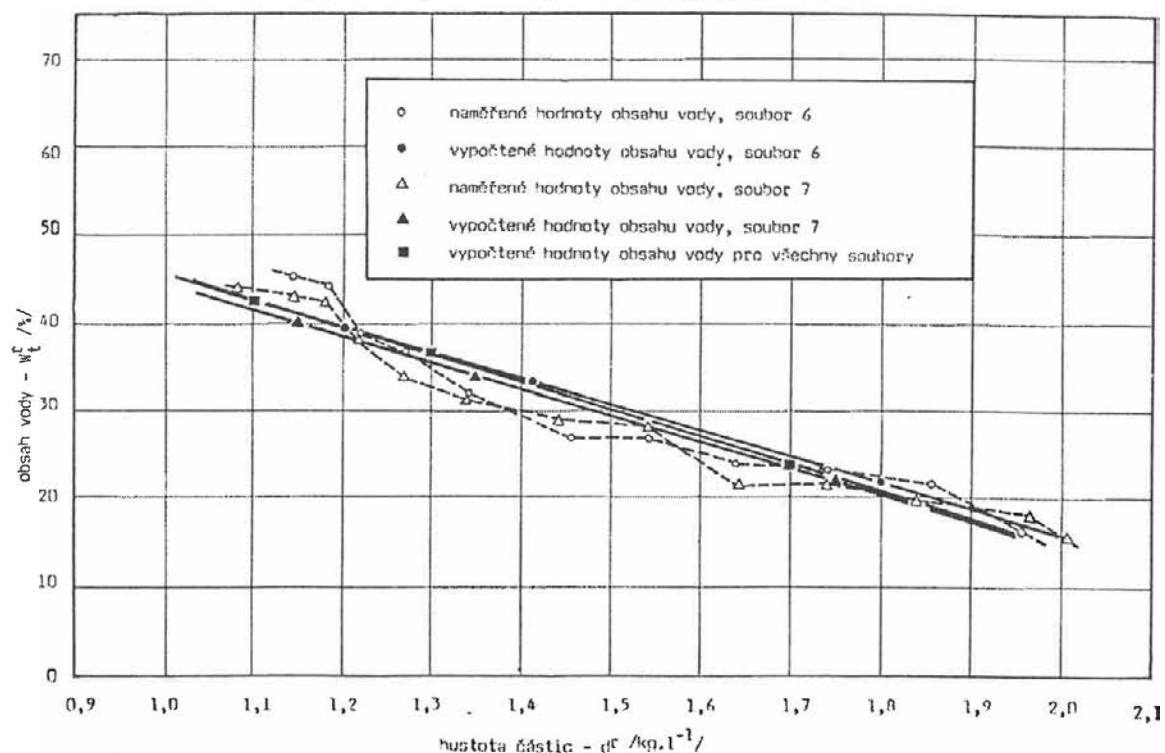
Obr. č. 1

Znázornění regresní závislosti mezi hustotou částic a obsahem popela pro hodnoty všech zkoumaných souborů a souborů s nejvyšším a nejnižším koeficientem korelace/souborů 5 a 2/

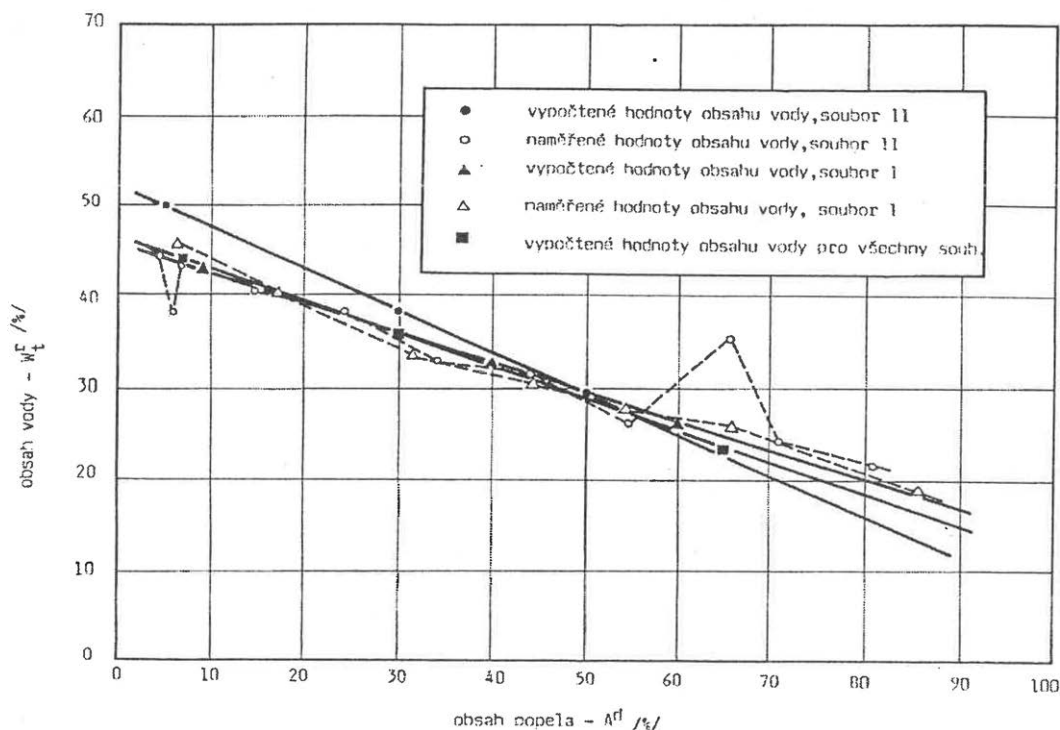


Obr. č. 2

Znázornění regresní závislosti mezi hustotou částic a obsahem vody pro hodnoty všech zkoumaných souborů a souborů s nejvyšším a nejnižším koeficientem korelace/souborů 7 a 6/



Znárodnění regresní závislosti mezi obsahem popela a obsahem vody částic pro hodnoty všech zkoumaných souborů a soubory s nejvyšším a nejnižším koeficientem korelace/soubory 11 a 1/



4. Posouzení výsledků analýzy

Údaje seřazené v tabulce 1 naznačují přímou závislost obsahu popela na zdánlivé hustotě a nepřímou závislost obsahu vody na tomtož parametru hnědého uhlí. Z regresních vztahů, vyjádřených polynomem 1. stupně a seřazených v tabulce 2, vyplývá variabilnost této závislosti pro různé části uhelné sloje a typy uhlí. Hodnoty konstanty b_0 pro funkční vztah $\Lambda^d = f(d_a^r)$ se pohybují v rozmezí hodnot $-70,35 \div 108,051$ (rozsah 37,701) a pro konstantu b_1 v rozmezí $75,910 \div 103,759$ (rozsah 27,849). U polynomu definujícího závislosti $W_t^F = f(d_a^r)$ mají tyto konstanty hodnoty v rozmezí $69,786 \div 88,580$ (rozsah 18,794) resp. $25,511 - 38,520$ (rozsah 13,009) a pro závislost $W_t^F = f(\Lambda^d)$ pak hodnoty v rozmezí $43,766 \div 52,316$ (rozsah 8,55) a $-0,325 - 0,464$ (rozsah 0,139). Všechny analyzované funkční závislosti vykazují velmi vysoký stupeň korelace vyjádřený hodnotou korelačních koeficientů, které se u závislosti $\Lambda^d = f(d_a^r)$ pohybují v rozmezí $0,948 \div 0,990$, u závislosti $W_t^F = f(d_a^r)$ v intervalu $0,941 \div 0,986$ a u vztahu $W_t^F = f(\Lambda^d)$ v rozmezí $0,902 \div 0,990$. Pro polynomy sestavené pro všechny analyzované hodnoty jsou koeficienty korelace relativně vysoké a jejich hodnoty se pohybují v rozmezí 0,959 až 0,978 (poslední řádek tabulky 3).

Přesnost lineární závislosti mezi analyzovanými parametry je zřejmá rovněž z obrázků 1 - 3, kde jsou vždy na jednom obrázku vyneseny naměřené hodnoty pro studovaný funkční vztah u souboru s nejvyšším a nejnižším koeficientem korelace a přímky znázorňující průběh příslušného polynomu. Do obrázku je vždy zakreslena také přímka, znázorňující tentýž funkční vztah pro hodnoty všech sledovaných souborů.

Parametr	$A^d = f(d^r)$			$W_t^r = f(d^r)$			$W_t^r = f(A^d)$		
	nejnižší ceny	nejvyšší ceny	všechny soubory	nejnižší r_{xy}	nejvyšší r_{xy}	všechny soubory	nejnižší r_{xy}	nejvyšší r_{xy}	všechny soubory
Součet odchylek Δ od naměř. hodnot	14,037	0,006	-6,924	-0,005	0,014	-	-11,858	-0,003	-9,508
Střední odchylka $\pm \Delta$ z lineár.hodnot	$\pm 6,23$	$\pm 3,605$	$\pm 4,671$	$\pm 2,519$	$\pm 1,899$	$\pm 2,270$	$\pm 5,664$	$\pm 1,012$	$\pm 1,774$
Střední odchylka z kvadratických hodnot Δ^2	$\pm 6,79$	$\pm 4,829$	$\pm 5,507$	$\pm 2,935$	$\pm 2,151$	$\pm 2,886$	$\pm 6,93$	1,198	$\pm 2,447$

Tab. 4 Hodnoty středních odchylek mezi naměřenými a vypočtenými hodnotami závisle přeměnné veličiny

Pro vztah $A^d = f(d_a^r)$ jsou tyto body a přímky vyneseny do obr. 1, pro závislost $W_t^r = f(A^d)$ na obr. 2. Přímky hodnot všech souborů jsou ohraničeny rovnoběžkami, označujícími střední hodnoty odchylek naměřených hodnot od hodnot vypočtených z rovnice polynomu. Z průběhu těchto rovnoběžek i z údajů tabulky 4 vyplývá, že nejvyšší střední hodnota odchylek naměřených a vypočtených hodnot byla zjištěna u vztahu $A^d = f(d_a^r)$, nejnižší pak u funkce $W_t^r = f(d_a^r)$. Vzdálenost rovnoběžek střední hodnoty odchylky od hodnot polynomu vyjadřuje také přesnost dané regrese.

Tab. 5 Údaje o středních hustotách a popelnatostech použité ke stanovení regresního vztahu

Rozmezí hustoty	Střední hodnota hustoty kg.l ⁻¹	Počet hodnot obsahů popela	Rozmezí obsahu popela pro rozmezí hustot	Hodnota rozmezí obsahu popela (%)	Střední hodnota obsahu popela (%)
-1,1	1,05	8	3,03 – 6,51	3,48	4,279
1,1 – 1,2	1,15	18	5,35 – 9,23	3,88	7,032
1,2 – 1,3	1,25	18	14,50 – 27,37	12,87	19,804
1,3 – 1,4	1,35	11	31,64 – 36,01	4,37	33,831
1,4 – 1,5	1,45	11	43,44 – 48,62	5,18	46,240
1,5 – 1,6	1,55	11	53,08 – 57,43	4,35	55,433
1,6 – 1,7	1,65	10	59,34 – 68,66	9,32	63,800
1,7 – 1,8	1,75	10	69,58 – 73,73	4,15	70,808
1,8 – 1,9	1,86	10	70,49 – 80,37	9,88	75,284
1,9 – 2,0	1,95	4	77,24 – 86,13	8,89	80,755
+ 2,0	2,10	6	82,12 – 90,36	8,24	87,505

Závěr

Z analýzy souborů naměřených hodnot obsahů popela, obsahů vody a zdánlivé hustoty a příslušných hodnot závislé proměnné veličiny ve zkoumaných funkčních vztazích vypočtených podle rovnic polynomu 1. stupně lze vyvodit následující závěry.

- pro funkční vztahy $A^d = f(d_a^r)$, $W_t^r = f(d_a^r)$ a $W_t^r = f(A^d)$ byla nalezena těsná lineární regrese vyjádřitelná polynomem 1. stupně
- těsnost regrese, vyjádřena koeficientem korelace, není konstantní a kolísá pro různé soubory analyzovaných hodnot, což svědčí o proměnlivých vlastnostech uhelné hmoty
- přesnost údajů získaných využitím regresního vztahu při případné náhradě některých laboratorních analýz je závislá na úrovni střední odchylky naměřených a vypočtených hodnot a s růstem její hodnoty klesá.