LICENČNÍ STUDIUM

GALILEO

STATISTICKÉ ZPRACOVÁNÍ DAT

SEMESTRÁLNÍ PRÁCE

Tvorba grafů v programu ORIGIN

doc.Dr.Ing.Vladimír Pata

Univerzita Tomáše Bati ve Zlíně
Fakulta technologická
Ústav výrobních technologií
Vavrečkova 275
760 01 Zlín
Česká republika

Obsah

Tvorba grafů v programu ORIGIN

Příklad č.1 3

Příklad č.2 9

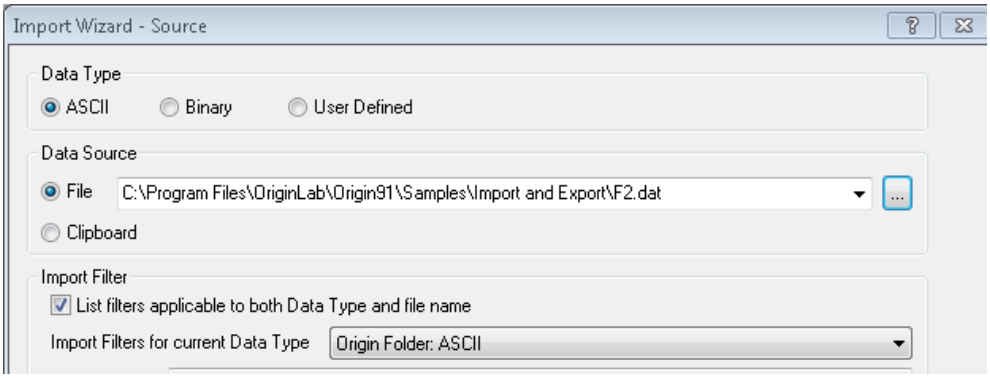
Příklad č.3 23

Použitá literatura 37

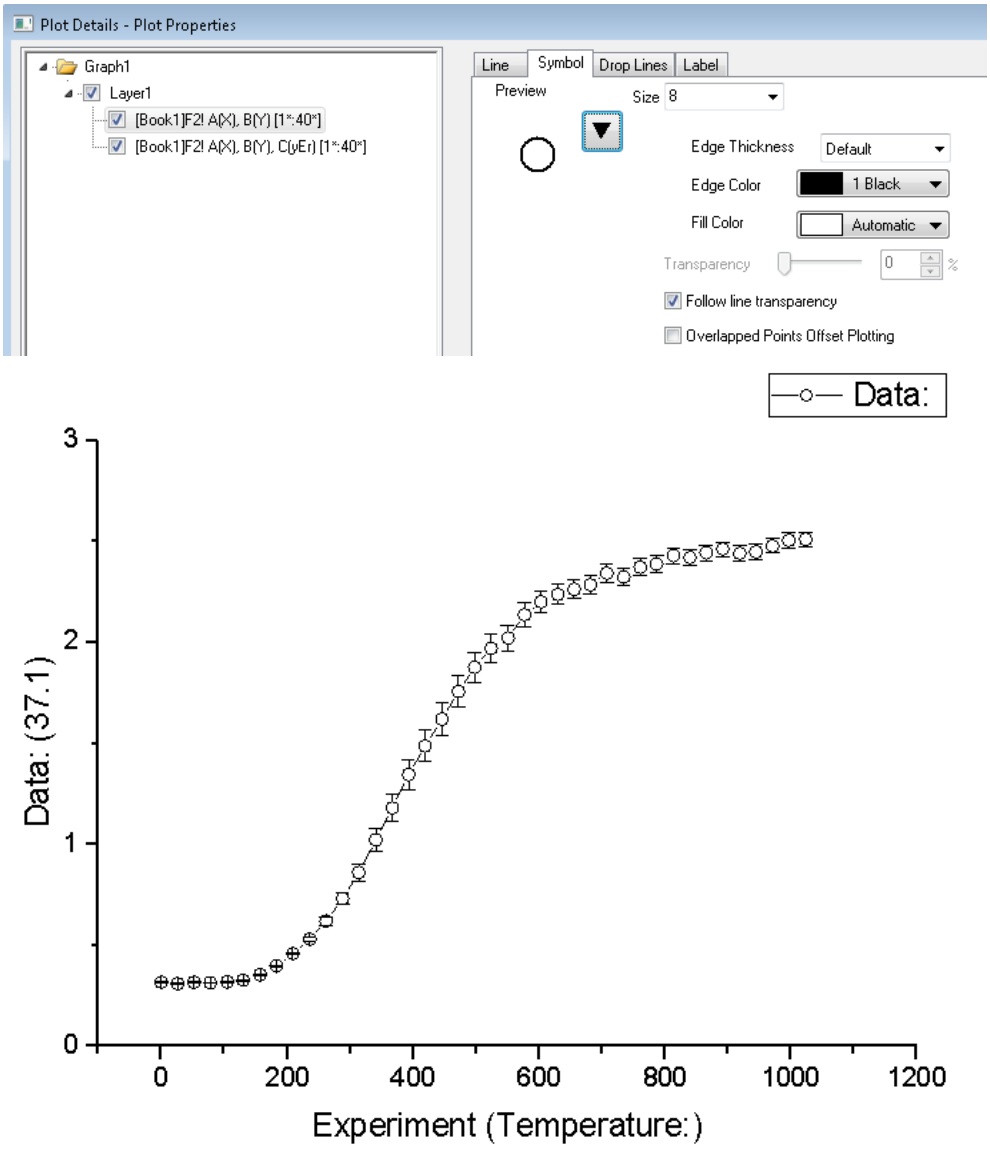
1.Příklad

Pomocí **Import Wizard** naimportujte data **Origin8/Samples/Import and Export/F2.dat**. Z načtených dat vytvořte bodový graf tak, aby měl body velikosti 8 a prázdná kolečka, chybové úsečky tvořené sloupcem C, všechny 4 osy, pouze dolní X a levá Y budou popsány, všechny záseky Tick uvnitř, popisky grafu písmem Time New Roman, černé pozadí grafu.

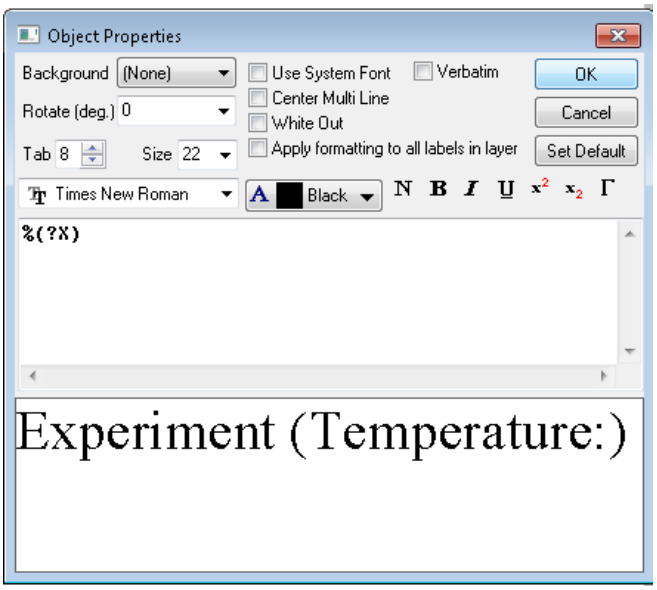
1.***Importování dat pomocí Import Wizard ze zdroje: Origin8/Samples/Import and Export/F2.dat.***

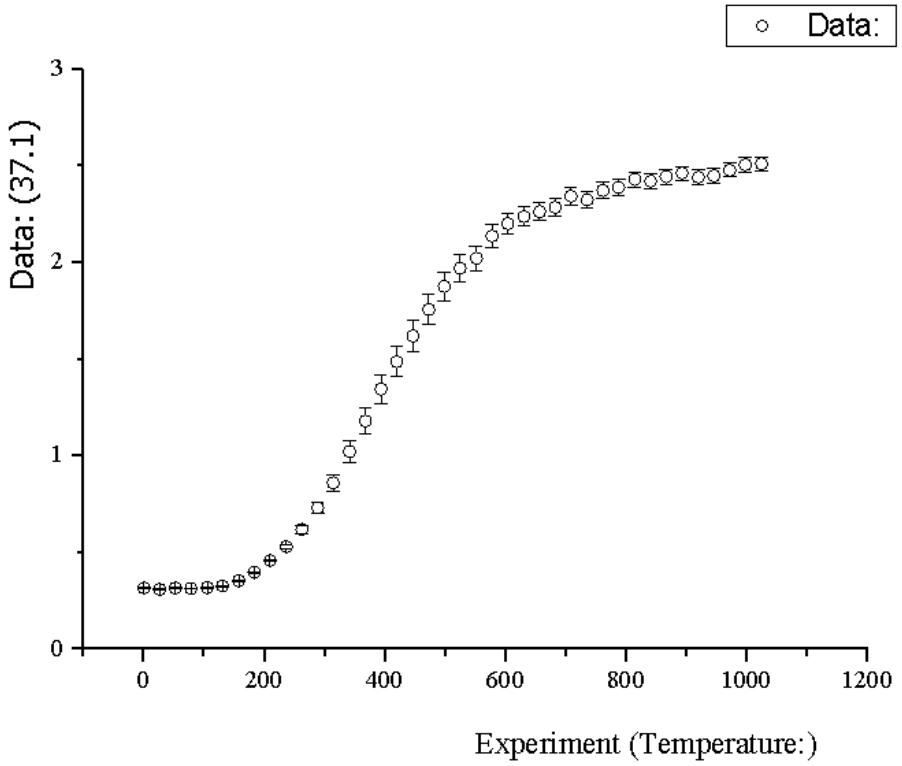


2.Vytvoření bodový graf tak, aby měl body velikosti 8 a prázdná kolečka, chybové úsečky tvořené sloupcem C

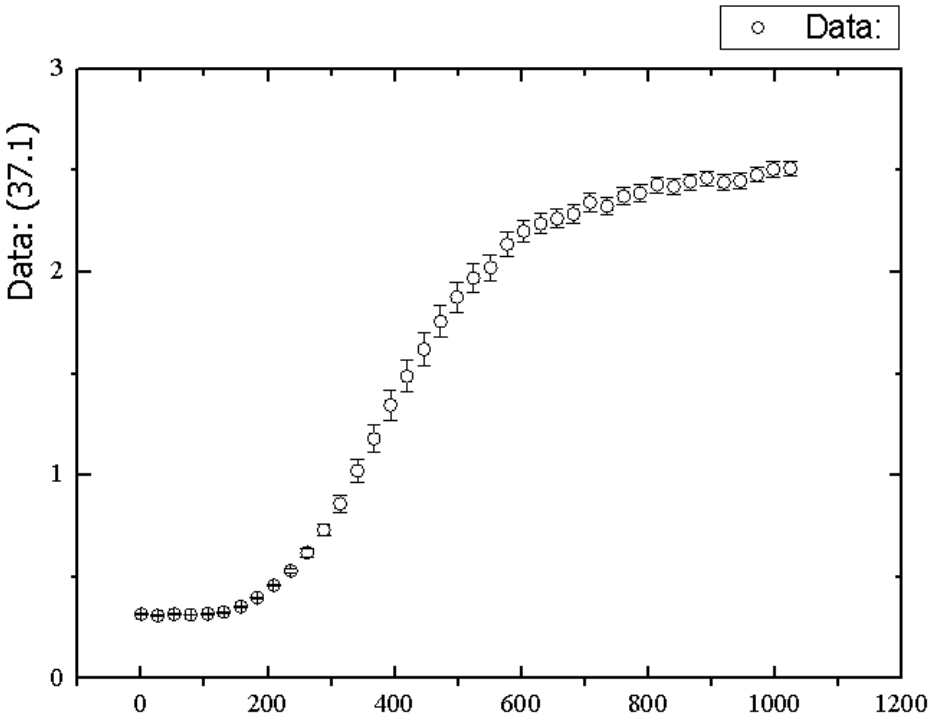


3. Tvorba popisků k osám dolní X a levá Y písmem Time New Roman

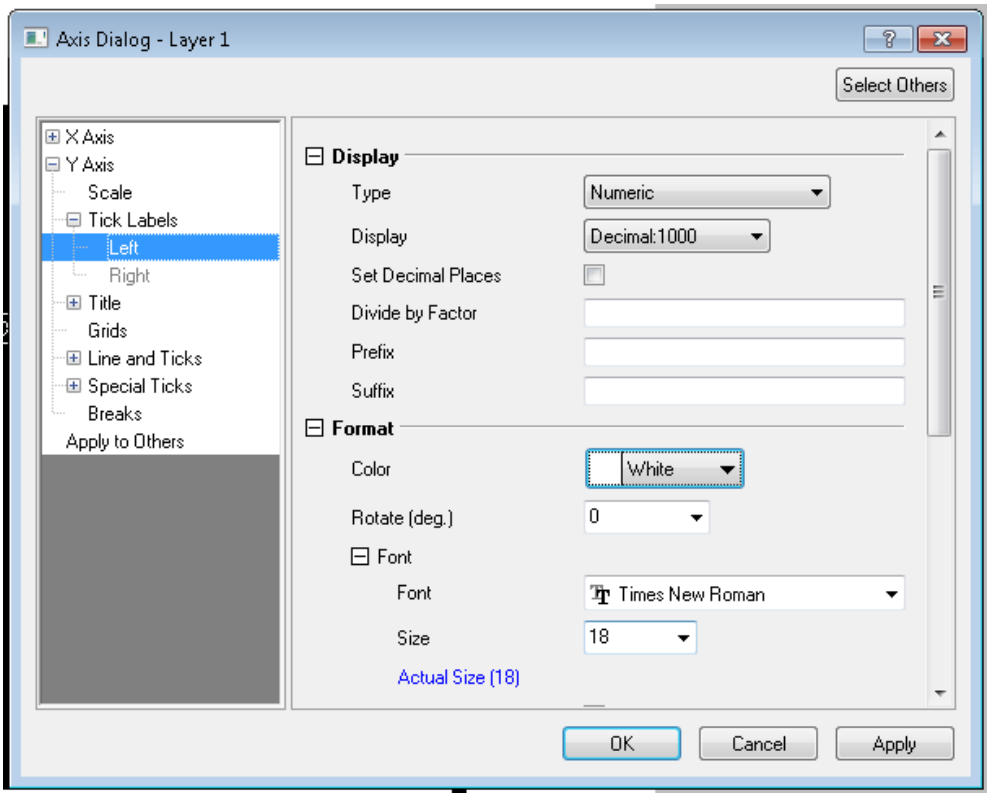


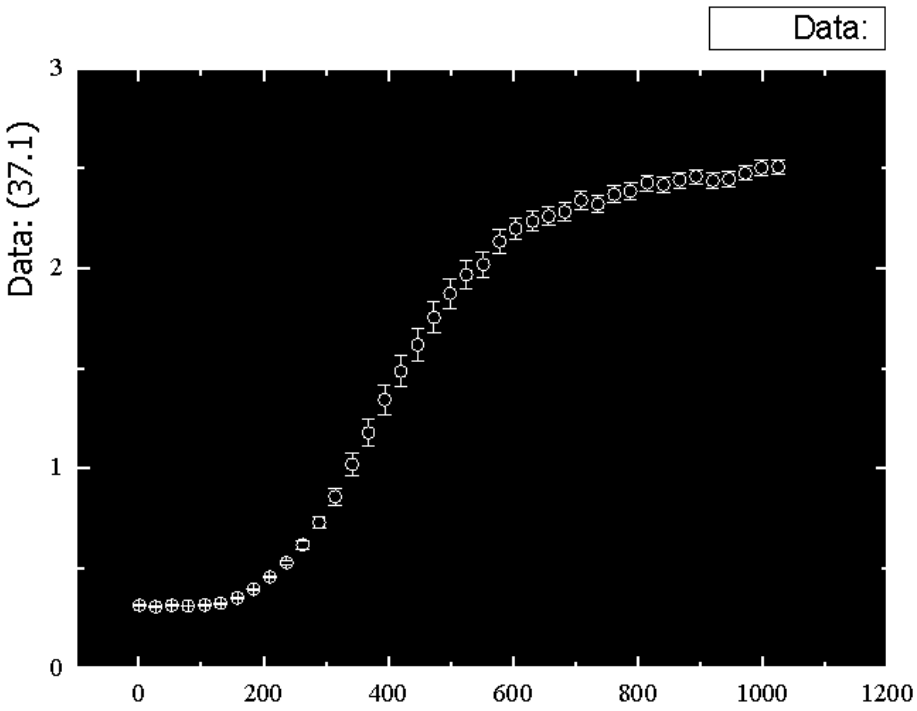


4.Tvorba všech čtyřech os se záseky Tick uvnitř



5. Nastavení černého pozadí grafu



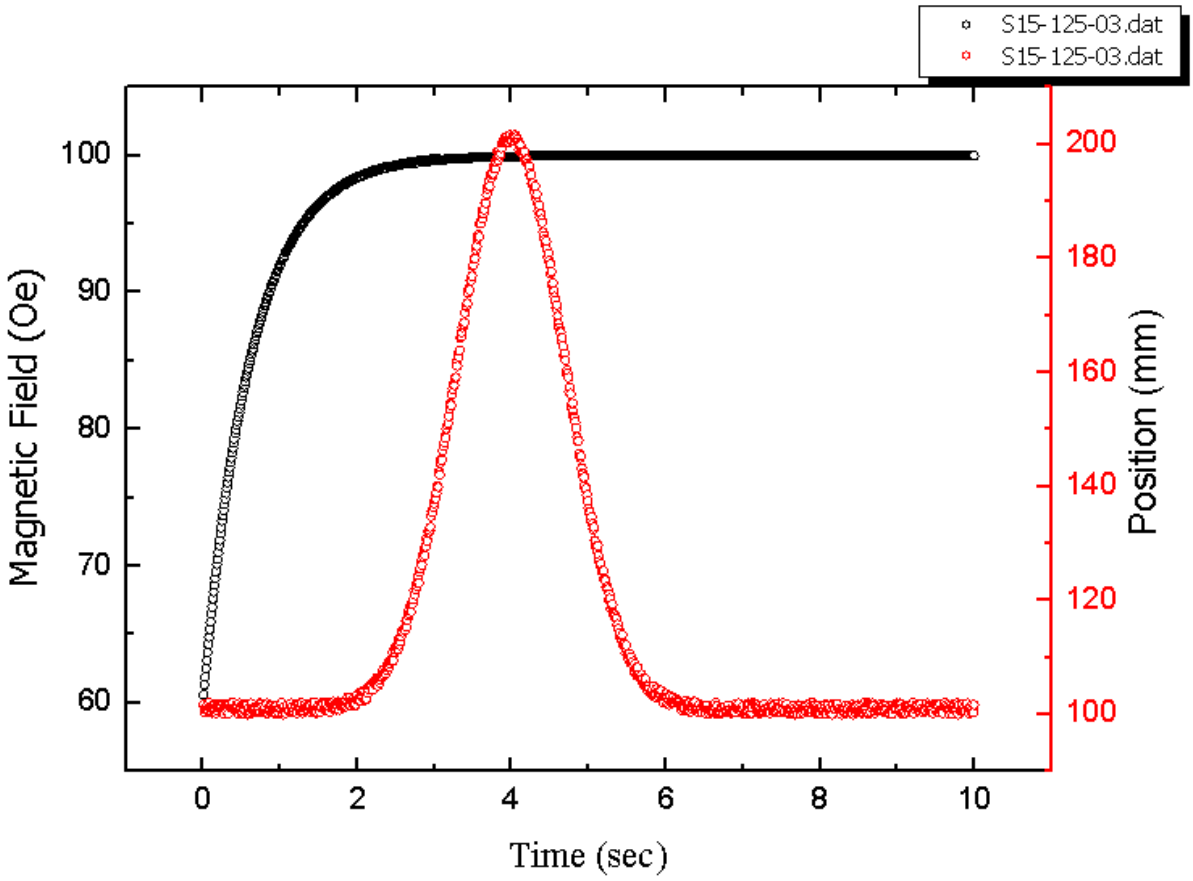


Finální graf s využitím výše uvedených nástrojů

2.Příklad

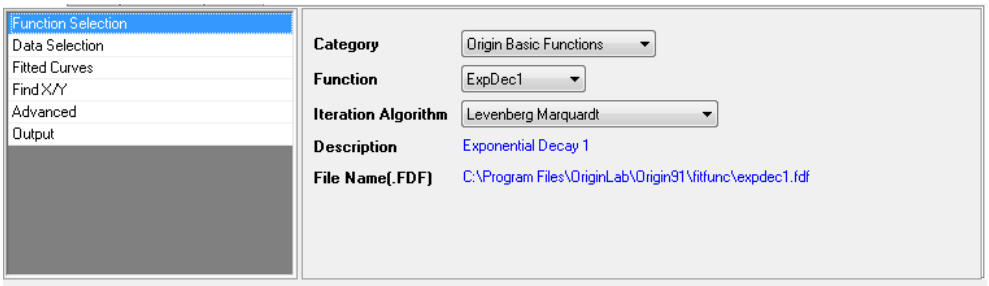
Načtěte data **Origin8/Samples/Import and Export/S15-125-03.dat** a vytvořte z nich bodový graf tak, aby na levé ose Y byla data **Magnetic Field** a na pravé ose Y data **Position**. Jednotlivé grafy od sebe barevně rozlište, body utvořte z prázdných koleček velikosti 5. Nelineárním prokládáním se pokuste nalézt alespoň dva nelineární modely pro každou křivku a uveďte, který z nich je lepší a vysvětlete také proč.

1. ***Načtěte data Origin8/Samples/Import and Export/S15-125-03.dat a vytvořte z nich bodový graf tak, aby na levé ose Y byla data Magnetic Field a na pravé ose Y data Position.***
2. ***Jednotlivé grafy od sebe barevně rozlište, body utvořte z prázdných koleček velikosti 5.***

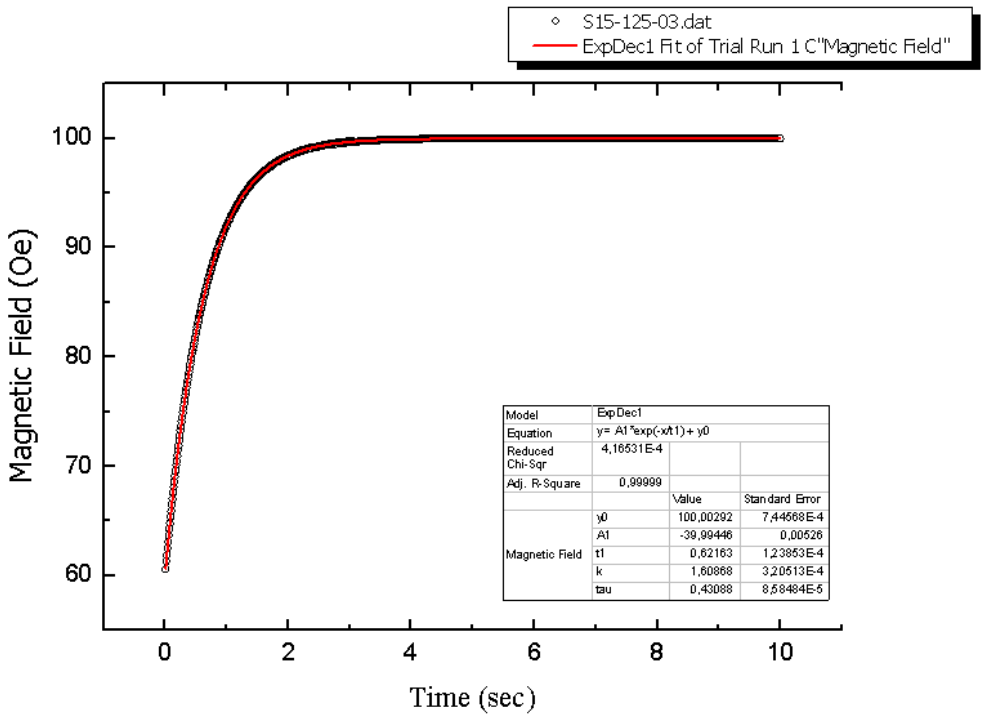


3. ***Nelineárním prokládáním se pokuste nalézt alespoň dva nelineární modely pro každou křivku a uveďte, který z nich je lepší a vysvětlete také proč.***

1.model pro data S-125-03 (Time = fnc (Magnetic field))



NO.	Param	Meaning	Fixed	Value
1	y0	offset	☐	100
1	A1	amplitude	☐	-41,66743
1	t1	decay constant	☐	0,17872



Description	Nonlinear Curve Fit
User Name	ucitel
Operation Time	3.5.2016 09:37:20
Iteration Algorithm	Levenberg Marquardt
Model	ExpDec1
Number of Parameters	3
Number of Derived Parameters	2
Number of Datasets	1
Equation	y = A1*exp(-x/t1) + y0
Report Status	New Analysis Report
Special Input Handling	
Data Filter	No

Input Data

		Dep/Indep	Data	Range	Weight Type
Magnetic Field	x	Indep	[Book1]"Trial Run 1"!A"Time"	[1*:1000*]	No Weighting
	y	Dep	[Book1]"Trial Run 1"!C"Magnetic Field"	[1*:1000*]	No Weighting

Parameters

		Value	Standard Error
Magnetic Field	y0	100,00292	7,44568E-4
	A1	-39,99446	0,00526
	t1	0,62163	1,23853E-4
	k	1,60868	3,20513E-4
	tau	0,43088	8,58484E-5

Reduced Chi-sqr = 4,1653141457E-4
COD(R^2) = 0,99999030601944
Iterations Performed = 6
Total Iterations in Session = 6
Fit converged. Chi-Sqr tolerance value of 1E-9 was reached.
k, tau are derived parameter(s).

Statistics

	Magnetic Field
Number of Points	1000
Degrees of Freedom	997
Reduced Chi-Sqr	4,16531E-4
Residual Sum of Squares	0,41528
Adj. R-Square	0,99999
Fit Status	Succeeded(100)

Fit Status Code :
100 : Fit converged. Chi-Sqr tolerance value of 1E-9 was reached.

Summary

	y0		A1		t1		k	tau	Statistics	
	Value	Standard Error	Value	Standard Error	Value	Standard Error	Value		Reduced Chi-Sqr	Adj. R-Square
Magnetic Field	100,00292	7,44568E-4	-39,99446	0,00526	0,62163	1,23853E-4	1,60868	0,43088	4,16531E-4	0,99999

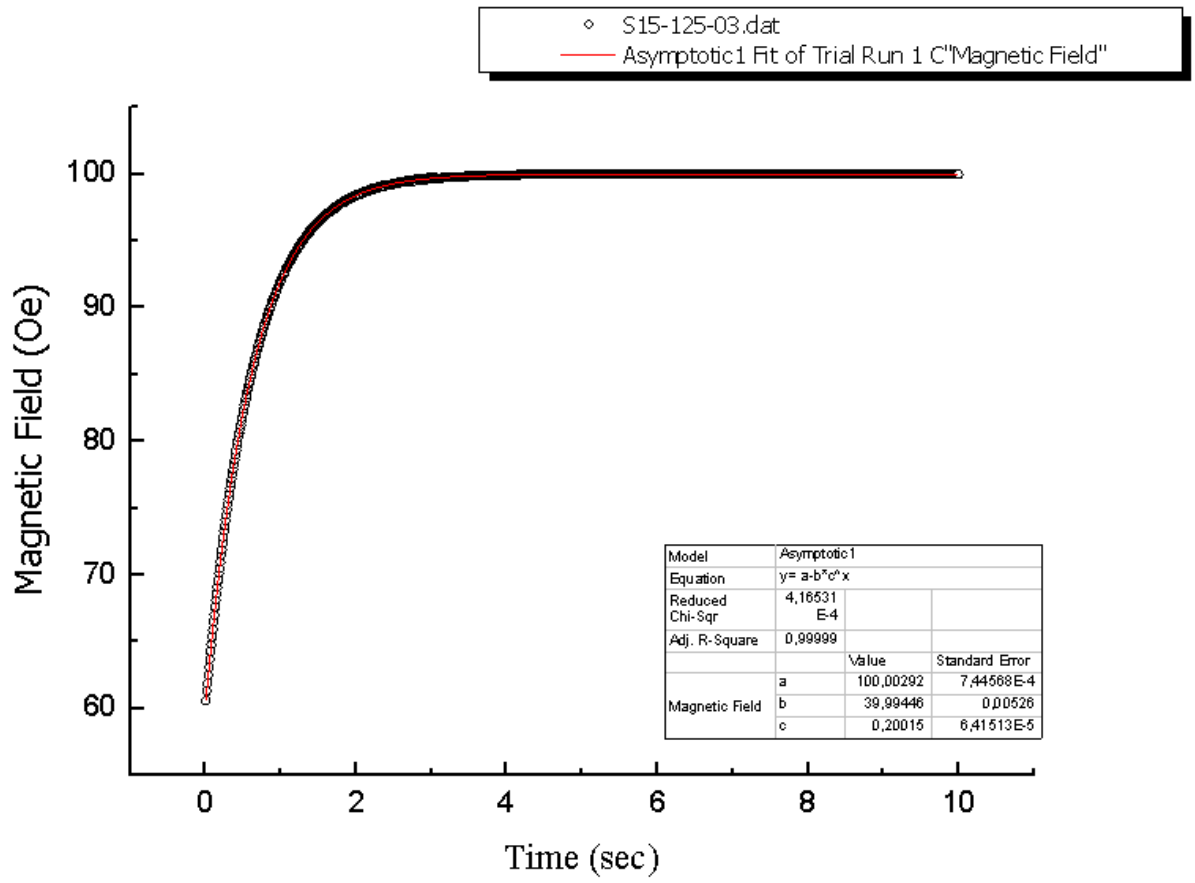
Závěr a porovnání 1.modelu (a)

Na importovaná data byly využity předdefinované modely, konkrétně:

a.) y = A1*exp(x/t1)+y0 tj. Magnetic Field = A1*exp((Time)/t1)+y0
Reziduální součet čtverců modelu: 0,41528
Adjustovaný koeficient determinace: 0,99999

Nalezené parametry 1.modelu včetně směrodatných odchylek (v programu označovaných jako Standartní chyba) – viz tab.Summary výše.

2.model pro data S-125-03 (Time = fnc (Magnetic field))



Description	Nonlinear Curve Fit
User Name	ucitel
Operation Time	9.5.2016 14:26:37
Iteration Algorithm	Levenberg Marquardt
Model	Asymptotic1
Number of Parameters	3
Number of Derived Parameters	0
Number of Datasets	1
Equation	y = a-b*c^x
Report Status	Report generated from Parameters Changed
Special Input Handling	
Data Filter	No

		Value	Standard Error
Magnetic Field	a	100,00292	7,44568E-4
	b	39,99446	0,00526
	c	0,20015	6,41513E-5

Reduced Chi-sqr = 4,1653141457E-4
COD(R^2) = 0,99999030601944
Iterations Performed = 5
Total Iterations in Session = 5
Fit converged - Chi-sqr no longer changed.

	Magnetic Field
Number of Points	1000
Degrees of Freedom	997
Reduced Chi-Sqr	4,16531E-4
Residual Sum of Squares	0,41528
Adj. R-Square	0,99999
Fit Status	Succeeded(101)

Fit Status Code :
101 : Fit converged - Chi-sqr no longer changed.

	a		b		c		Statistics	
	Value	Standard Error	Value	Standard Error	Value	Standard Error	Reduced Chi-Sqr	Adj. R-Square
Magnetic Field	100,00292	7,44568E-4	39,99446	0,00526	0,20015	6,41513E-5	4,16531E-4	0,99999

		DF	Sum of Squares	Mean Square	F Value	Prob>F
Magnetic Field	Regression	3	9,55625E6	3,18542E6	7,64748E9	0
	Residual	997	0,41528	4,16531E-4		
	Uncorrected Total	1000	9,55625E6			
	Corrected Total	999	42839,14311			

Závěr a porovnání 1.modelu (b)

Na importovaná data byly využity předdefinované modely, konkrétně:

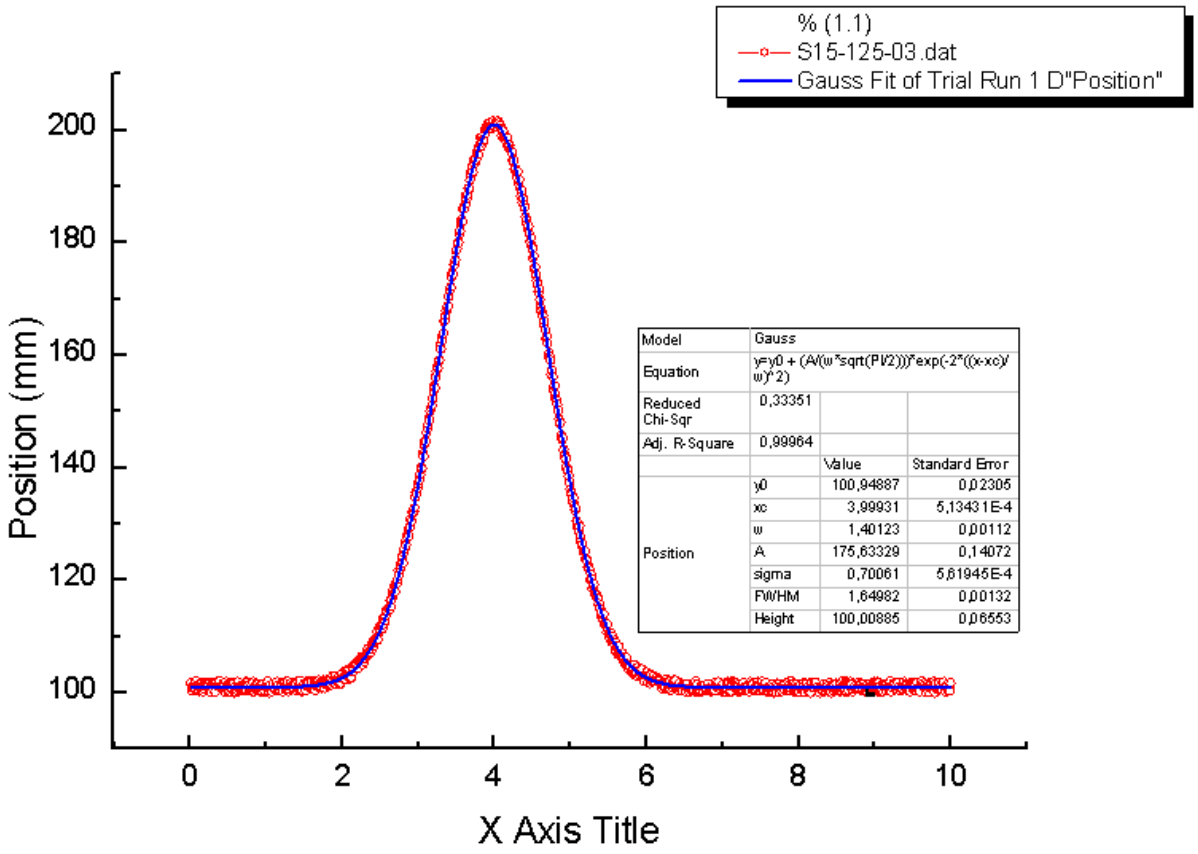
b.) y = a – b*c^x tj. Magnetic Field = a – b*c^(Time)
Reziduální součet čtverců modelu: 0,41528
Adjustovaný koeficient determinace: 0,99999

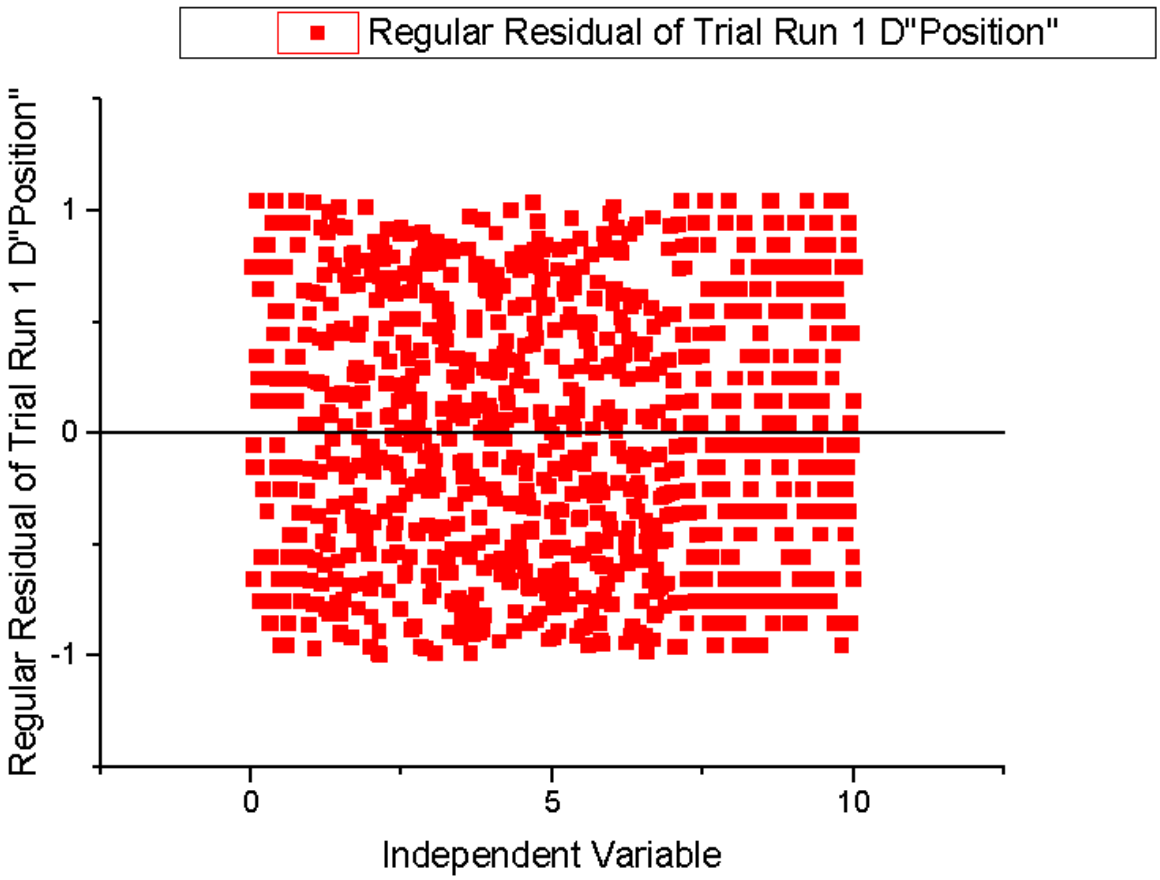
Nalezené parametry 1.modelu včetně směrodatných odchylek (v programu označovaných jako Standartní chyba) – viz tab.Summary výše.

Finální závěr k modelu č.1

Posouzení kvality modelu z důvodu absence parametru AIC a MAP je obtížné. Je možné použít pouze reziduální součet čtverců modelu a adjustovaný koeficient determinace, které v případě modelu č.1 (tedy Magnetic Field = a – b*c^(Time) a Magnetic Field = A1*exp((Time)/t1)+y0) nedokázaly rozeznat, která z modelů je pro daná data lepší. Takto se jeví, že oba modely jsou stejně kvalitní.

2.model pro data S-125-03 (Position = fnc (Time)) (a)





Description	Nonlinear Curve Fit
User Name	ucitel
Operation Time	4.5.2016 17:58:09
Iteration Algorithm	Levenberg Marquardt
Model	Gauss
Number of Parameters	4
Number of Derived Parameters	3
Number of Datasets	1
Equation	y=y0 + (A/(w*sqrt(PI/2)))*exp(-2*((x-xc)/w)^2)
Report Status	New Analysis Report
Special Input Handling	
Data Filter	No

		Dep/Indep	Data	Range	Weight Type
Position	x	Indep	[Book1]"Trial Run 1"!A"Time"	[1*:1000*]	No Weighting
	y	Dep	[Book1]"Trial Run 1"!D"Position"	[1*:1000*]	No Weighting

		Value	Standard Error
Position	y0	100,94887	0,02305
	xc	3,99931	5,13431E-4
	w	1,40123	0,00112
	A	175,63329	0,14072
	sigma	0,70061	5,61945E-4
	FWHM	1,64982	0,00132
	Height	100,00885	0,06553

Reduced Chi-sqr = 0,333508587686
COD(R^2) = 0,999644309492D1
Iterations Performed = 5
Total Iterations in Session = 5
Fit converged. Chi-Sqr tolerance value of 1E-9 was reached.
sigma, FWHM, Height are derived parameter(s).

	Position
Number of Points	1000
Degrees of Freedom	996
Reduced Chi-Sqr	0,33351
Residual Sum of Squares	332,17455
Adj. R-Square	0,99964
Fit Status	Succeeded(100)

	y0		xc		w		A		sigma	FWHM	Height	Statistics	
	Value	Standard Error	Value	Standard Error	Value	Standard Error	Value	Standard Error	Value			Reduced Chi-Sqr	Adj. R-Square
Position	100,94887	0,02305	3,99931	5,13431E-4	1,40123	0,00112	175,63329	0,14072	0,70061	1,64982	100,00885	0,33351	0,99964

		DF	Sum of Squares	Mean Square	F Value	Prob>F
Position	Regression	4	1,49787E7	3,74467E6	1,12281E7	0
	Residual	996	332,17455	0,33351		
	Uncorrected Total	1000	1,4979E7			
	Corrected Total	999	933886,47116			

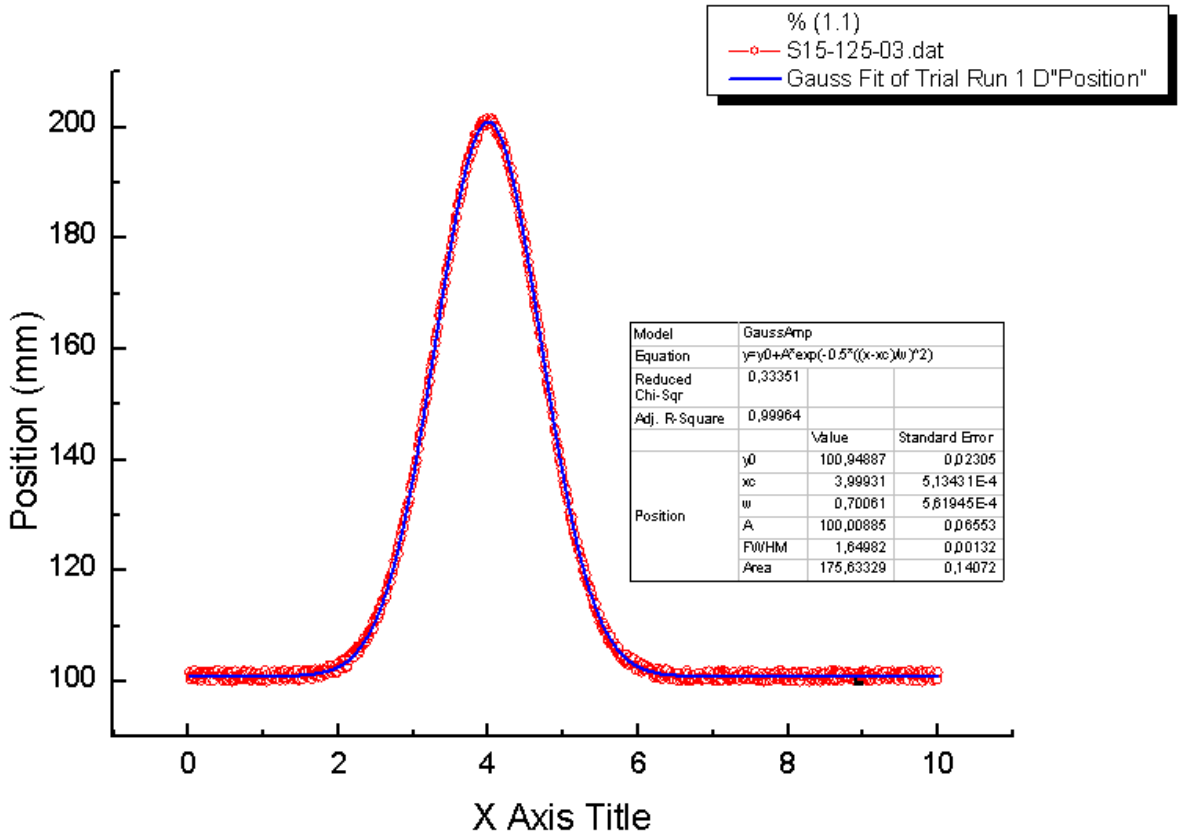
Závěr a porovnání 2.modelu (a)

Na importovaná data byly využity předdefinované modely, konkrétně:

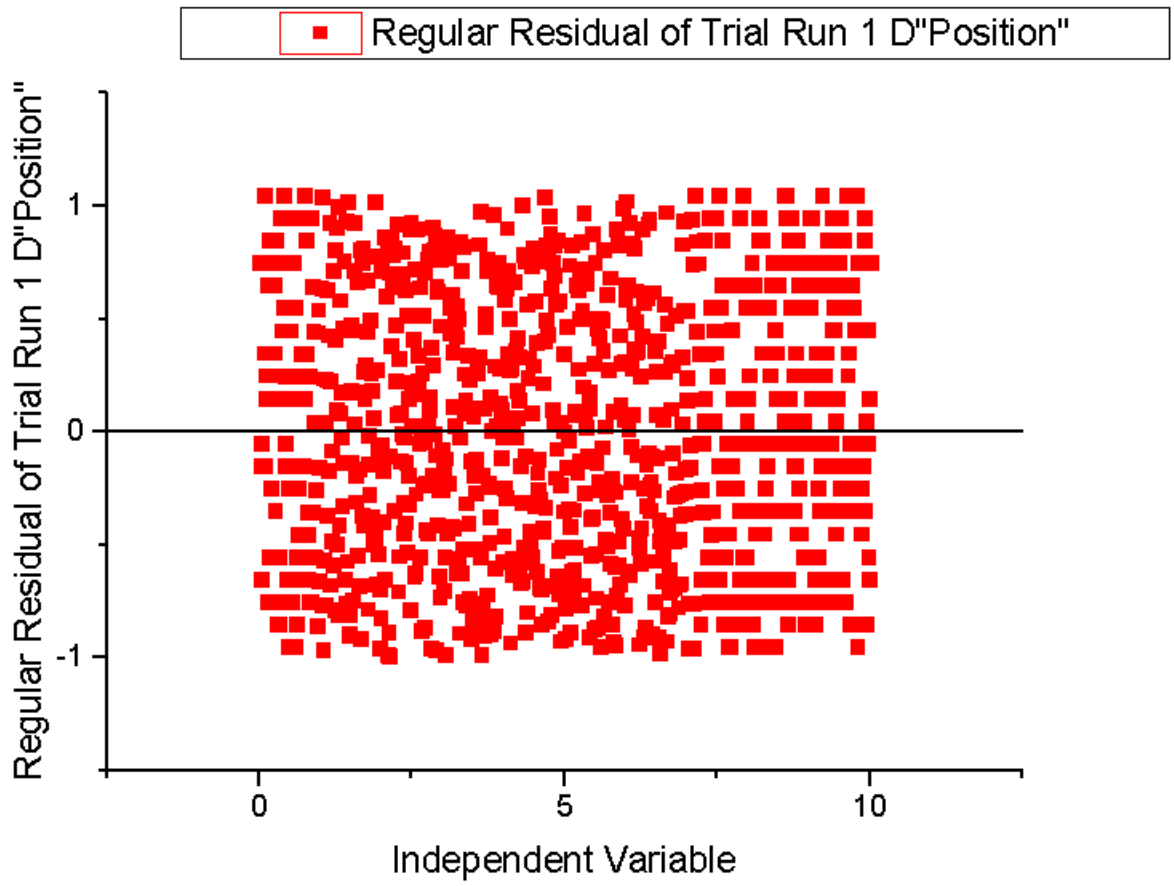
a.) y = y0+(A/(w*sqrt(PI/2)))*exp(-2*((x-xc)/w)^2)
Reziduální součet čtverců modelu: 332,17455
Adjustovaný koeficient determinace: 0,99964

Nalezené parametry 2.modelu včetně směrodatných odchylek (v programu označovaných jako Standartní chyba) – viz tab.Summary výše.

2.model pro data S-125-03 (Position = fnc (Time)) (b)



Model	GaussAmp		
Equation	y=y0+A*exp(-0.5*((x-xc)/w)^2)		
Reduced Chi-Sqr	0,33351		
Adj. R-Square	0,99964		
		Value	Standard Error
Position	y0	100,94887	0,0,2305
	xc	3,99931	5,13431E-4
	w	0,70061	5,81945E-4
	A	100,00885	0,0,6553
	FWHM	1,64982	0,0,0132
	Area	175,63329	0,1,4072



Description	Nonlinear Curve Fit
User Name	ucitel
Operation Time	9.5.2016 14:09:31
Iteration Algorithm	Levenberg Marquardt
Model	GaussAmp
Number of Parameters	4
Number of Derived Parameters	2
Number of Datasets	1
Equation	y=y0+A*exp(-0.5*((x-xc)/w)^2)
Report Status	Report generated from Parameters Changed
Special Input Handling	
Data Filter	No

		Dep/Indep	Data	Range	Weight Type
Position	x	Indep	[Book1]"Trial Run 1"!A"Time"	[1*:1000*]	No Weighting
	y	Dep	[Book1]"Trial Run 1"!D"Position"	[1*:1000*]	No Weighting

		Value	Standard Error
Position	y0	100,94887	0,02305
	xc	3,99931	5,13431E-4
	w	0,70061	5,61945E-4
	A	100,00885	0,06553
	FWHM	1,64982	0,00132
	Area	175,63329	0,14072

Reduced Chi-sqr = 0,333508587686
COD(R^2) = 0,99964430949201
Iterations Performed = 3
Total Iterations in Session = 3
Fit converged. Chi-Sqr tolerance value of 1E-9 was reached.

	Position
Number of Points	1000
Degrees of Freedom	996
Reduced Chi-Sqr	0,33351
Residual Sum of Squares	332,17455
Adj. R-Square	0,99964
Fit Status	Succeeded(100)

Fit Status Code :
100 : Fit converged. Chi-Sqr tolerance value of 1E-9 was reached.

	y0		xc		w		A		FWHM	Area	Statistics	
	Value	Standard Error	Value	Standard Error	Value	Standard Error	Value	Standard Error	Value		Reduced Chi-Sqr	Adj. R-Square
Position	100,94887	0,02305	3,99931	5,13431E-4	0,70061	5,61945E-4	100,00885	0,06553	1,64982	175,63329	0,33351	0,99964

		DF	Sum of Squares	Mean Square	F Value	Prob>F
Position	Regression	4	1,49787E7	3,74467E6	1,12281E7	0
	Residual	996	332,17455	0,33351		
	Uncorrected Total	1000	1,4979E7			
	Corrected Total	999	933886,47116			

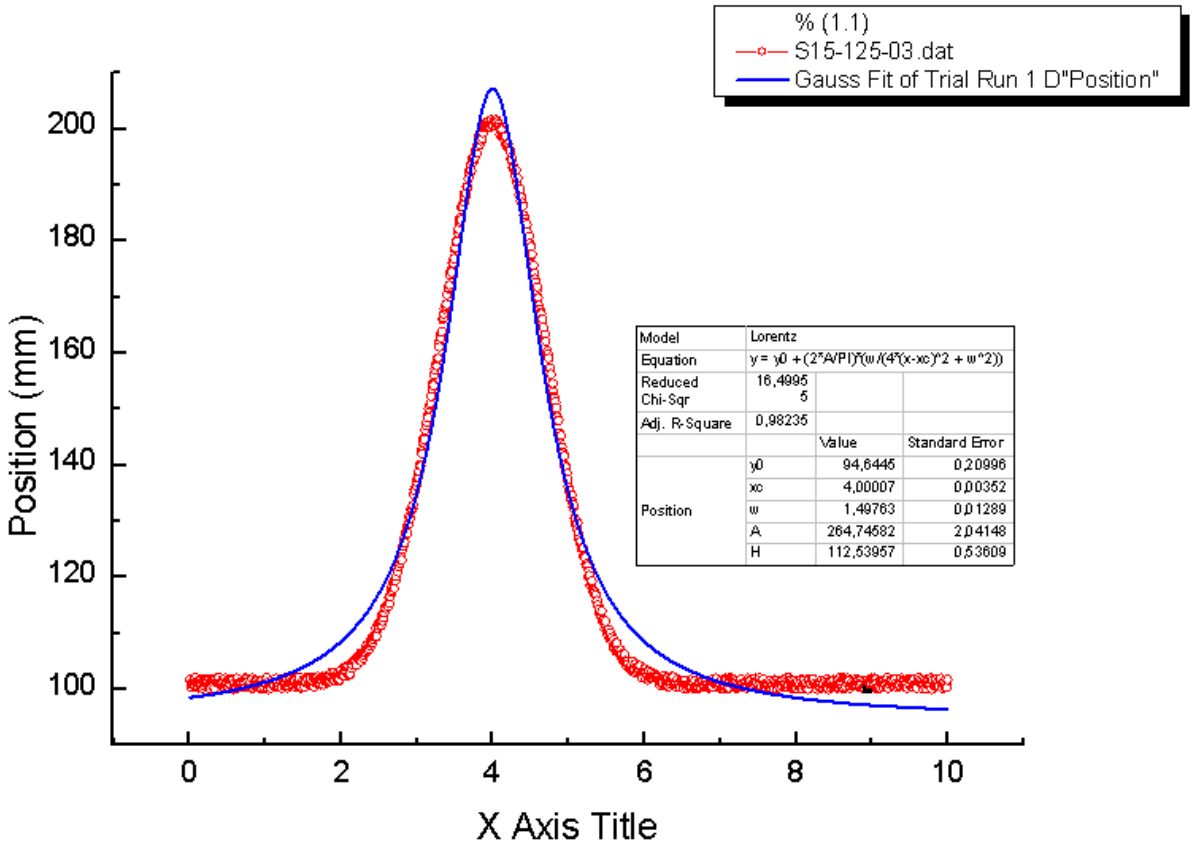
Závěr a porovnání 2.modelu (b)

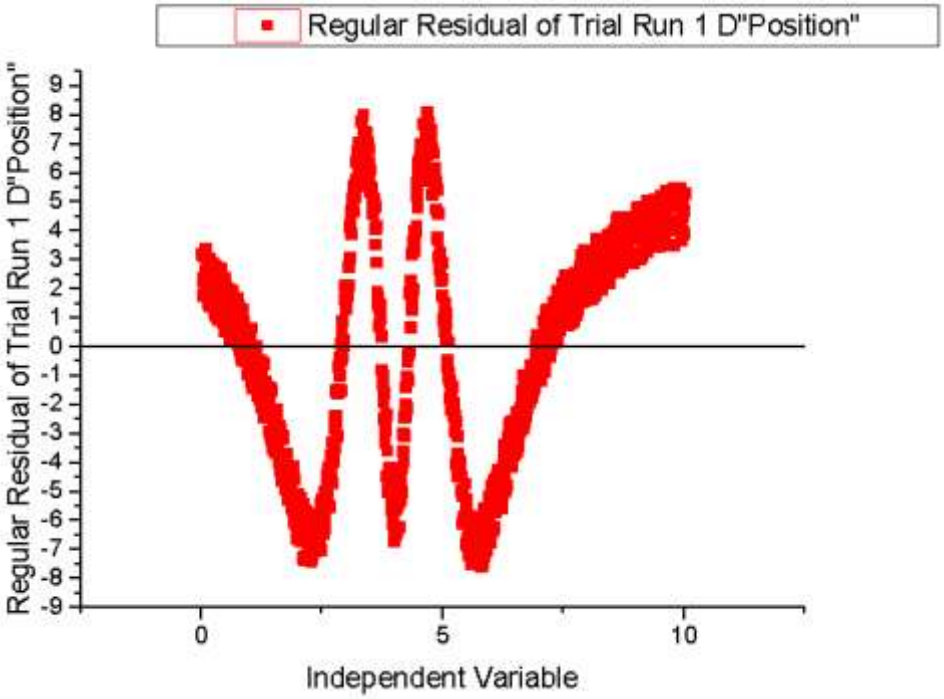
Na importovaná data byly využity předdefinované modely, konkrétně:

b.) y = y0+A*exp(-0,5*((x-xc)/w)^2)
Reziduální součet čtverců modelu: 332,17455
Adjustovaný koeficient determinace: 0,99964

Nalezené parametry 2.modelu včetně směrodatných odchylek (v programu označovaných jako Standartní chyba) – viz tab.Summary výše.

2.model pro data S-125-03 (Position = fnc (Time)) (c)





Description	Nonlinear Curve Fit
User Name	ucitel
Operation Time	9.5.2016 14:18:18
Iteration Algorithm	Levenberg Marquardt
Model	Lorentz
Number of Parameters	4
Number of Derived Parameters	1
Number of Datasets	1
Equation	y = y0 + (2*A/PI)*(w/(4*(x-xc)^2 + w^2))
Report Status	Report generated from Parameters Changed
Special Input Handling	
Data Filter	No

		Dep/Indep	Data	Range	Weight Type
Position	x	Indep	[Book1]"Trial Run 1"!A"Time"	[1*:1000*]	No Weighting
	y	Dep	[Book1]"Trial Run 1"!D"Position"	[1*:1000*]	No Weighting

		Value	Standard Error
Position	y0	94,6445	0,20996
	xc	4,00007	0,00352
	w	1,49763	0,01289
	A	264,74582	2,04148
	H	112,53957	0,53609

Reduced Chi-sqr = 16,4995503034
COD(R^2) = 0,98240305153821
Iterations Performed = 8
Total Iterations in Session = 8
Fit converged. Chi-Sqr tolerance value of 1E-9 was reached.

	Position
Number of Points	1000
Degrees of Freedom	996
Reduced Chi-Sqr	16,49955
Residual Sum of Squares	16433,5521
Adj. R-Square	0,98235
Fit Status	Succeeded(100)

Fit Status Code :
100 : Fit converged. Chi-Sqr tolerance value of 1E-9 was reached.

	y0		xc		w		A		H	Statistics	
	Value	Standard Error	Value	Standard Error	Value	Standard Error	Value	Standard Error	Value	Reduced Chi-Sqr	Adj. R-Square
Position	94,6445	0,20996	4,00007	0,00352	1,49763	0,01289	264,74582	2,04148	112,53957	16,49955	0,98235

		DF	Sum of Squares	Mean Square	F Value	Prob>F
Position	Regression	4	1,49626E7	3,74065E6	226712,15568	0
	Residual	996	16433,5521	16,49955		
	Uncorrected Total	1000	1,4979E7			
	Corrected Total	999	933886,47116			

Závěr a porovnání 2.modelu (c)

Na importovaná data byly využity předdefinované modely, konkrétně:

c.) y = y0+(2*A/PI)*(w/(4*(x-xc)^2+w^2))
Reziduální součet čtverců modelu: 16433,5521
Adjustovaný koeficient determinace: 0,98235

Nalezené parametry 2.modelu včetně směrodatných odchylek (v programu označovaných jako Standartní chyba) – viz tab.Summary výše.

Finální závěr k modelu č.2

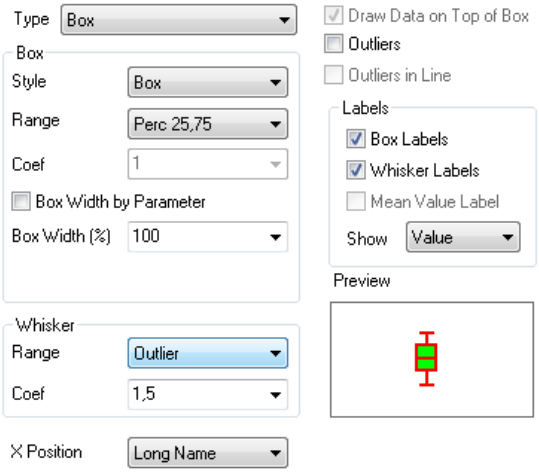
Posouzení kvality modelu z důvodu absence parametru AIC a MAP je obtížné. Je možné použít pouze reziduální součet čtverců modelu a adjustovaný koeficient determinace, které v případě modelu č.2 a, b (y = y0+(A/(w*sqrt(PI/2)))*exp(-2*((x-xc)/w)^2) a dále y = y0+A*exp(-0,5*((x-xc)/w)^2)) nedokázaly rozeznat, která z modelů je pro daná data lepší. Takto se jeví, že oba modely jsou stejně kvalitní.

Rozdílně se však jeví model c (y = y0+(2*A/PI)*(w/(4*(x-xc)^2+w^2))), který na rozdíl od a a b má vyšší hodnotu reziduálního součtu čtverců i nižší hodnotu adjustovaného koeficientu determinace. Takto je možné konstatovat, že model c je podstatně horší než modely a a b, což je vidět i z hladkosti proložení a grafů reziduí.

3.Příklad

Užitím příkladu **E4.20** z Kompendia vytvořte 6-ti vrstvý graf závislostí jednotlivých znaků cigaret. Pomocí lineární regrese pak rozhodněte, které znaky spolu korelují a proč. Nacházejí se v datech nějaké odlehlé hodnoty? Jak bude vypadat závislost po jejich odstranění (do stejného grafu)? Vysvětlete obdržené výsledky.

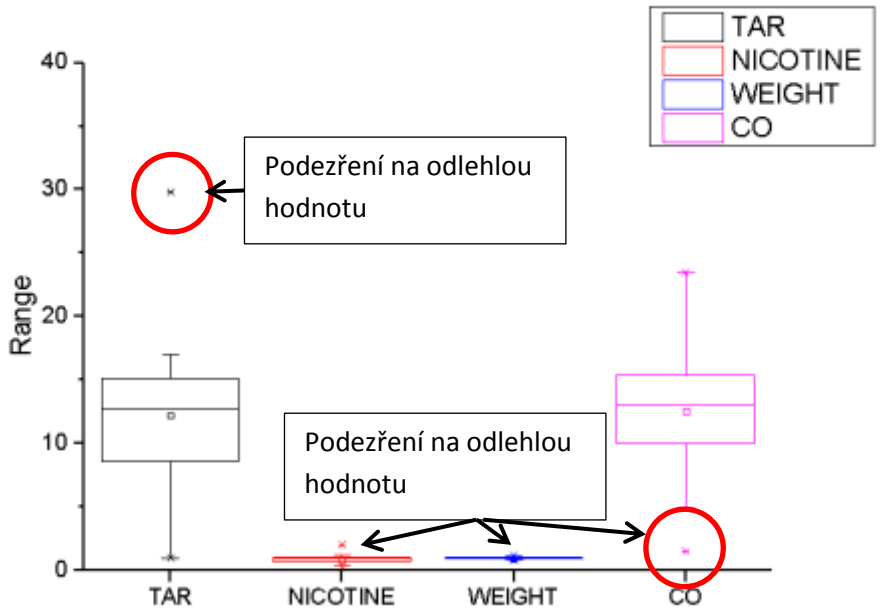
Tvorba Box_plotového grafu pro jednotlivé znaky v Layerech



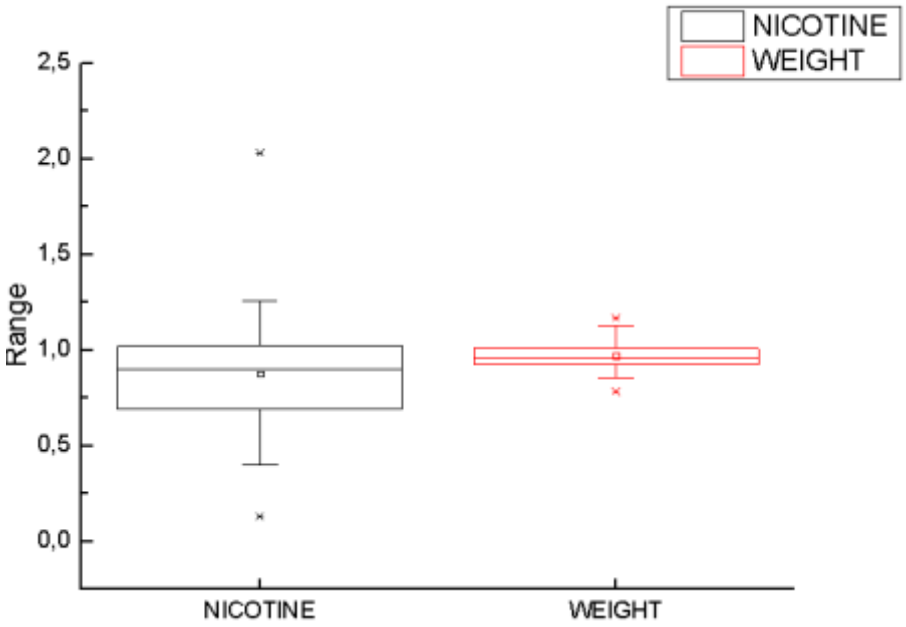
Základní statistika v layerech

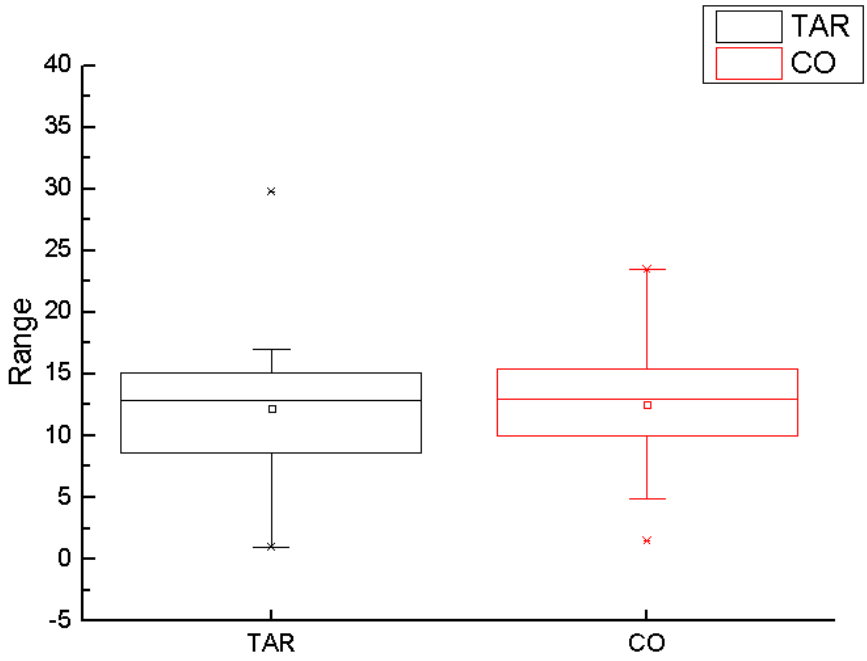
	N	Mean	SD	Sum	Min	Max
"TAR"	25	12,216	5,66581	305,4	1	29,8
"NICOTINE"	25	0,8764	0,35406	21,91	0,13	2,03
"WEIGHT"	25	0,97028	0,08772	24,2571	0,7851	1,165
"CO"	25	12,528	4,73968	313,2	1,5	23,5

Základní statistika jednoznačně vede k tvrzení, že v layerech TAR, NICOTINE, WEIGHT a CO bude nutné využít pro nalezení odlehlých hodnot rozdílných grafů, přesněji řečeno grafů pro TAR a CO a dále NICOTINE a WEIGHT, protože jejich hodnoty se liší o jeden řád.



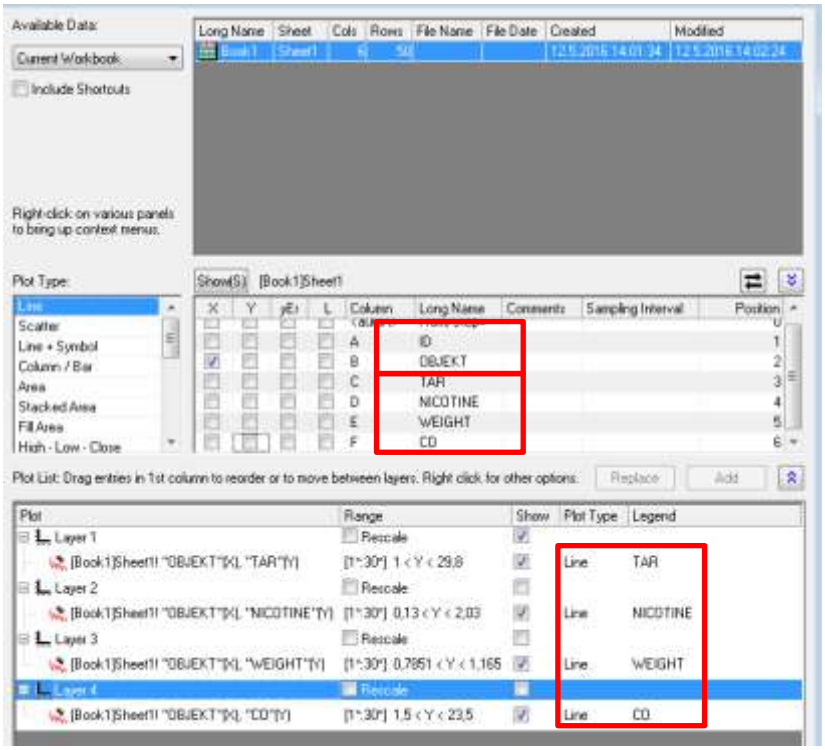
Na výše uvedeném grafu je demonstrován případ, že všechny parametry byly znázorněny do jednoho grafu. Z důvodu rozdílnosti hodnot parametrů (o jeden řád) odlehlé hodnoty parametru NICOTINE a WEIGHT zcela zanikají.

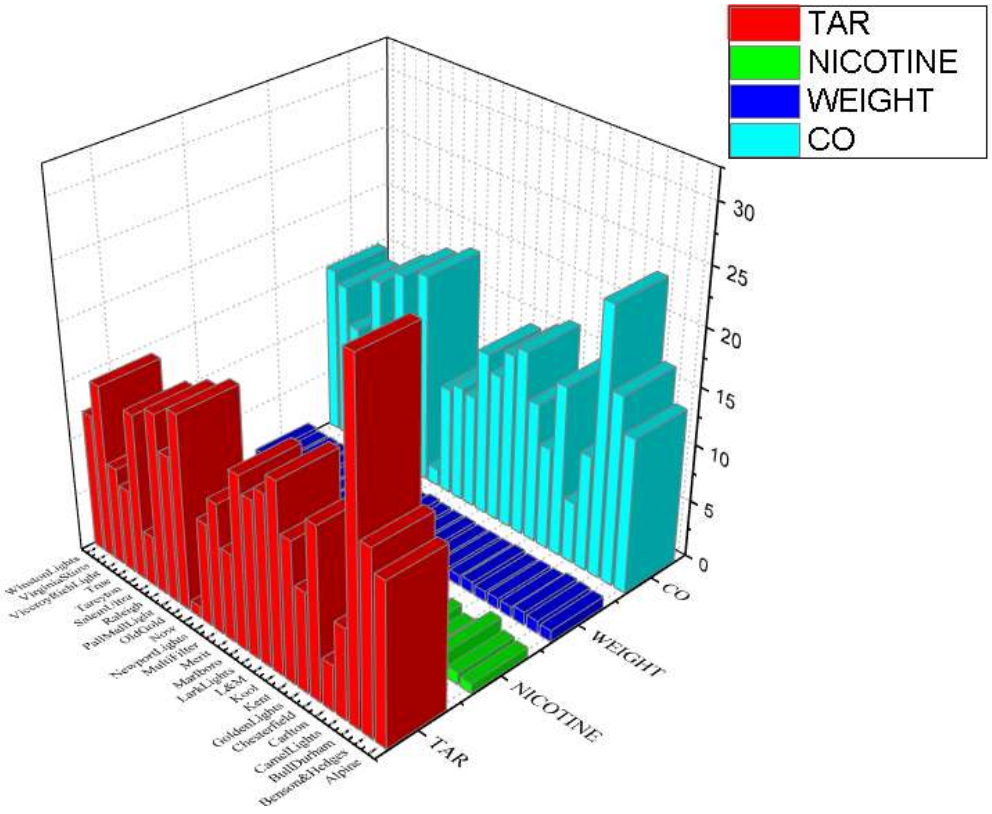




Ve výše uvedených dvojicích Box Plotových grafů je zřejné, že v datech se nacházejí odlehlé hodnoty, které jsou v Box Plotových grafech znázorněny hvězdičkou (*). Leží tedy mimo horní a dolní hradbu daného Box Plotového grafu.

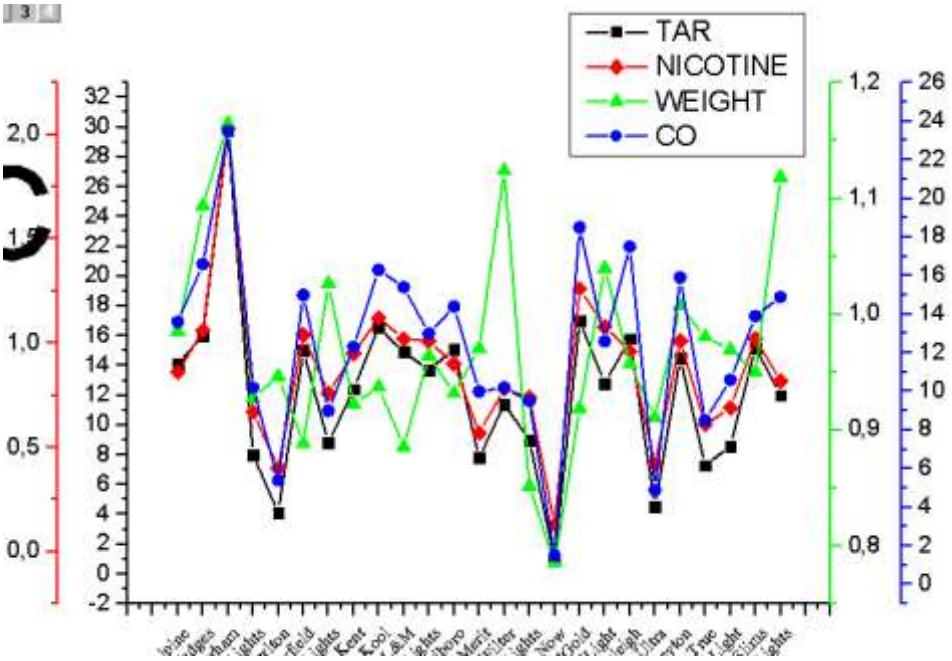
Vícevrstvý graf závislostí jednotlivých znaků cigaret.





V případě tvorby 3D grafu z jednotlivých layerů opět můžeme konstatovat, že parametry NICOTINE a WEIGHT jsou řádově menší než parametry TAR a CO. I přes grafickou „atraktivnost“ tohoto typu grafu, jeho využití v tomto případě je sporné.

Čtyřúrovňový graf jednotlivých složek s vyznačením trendových složek

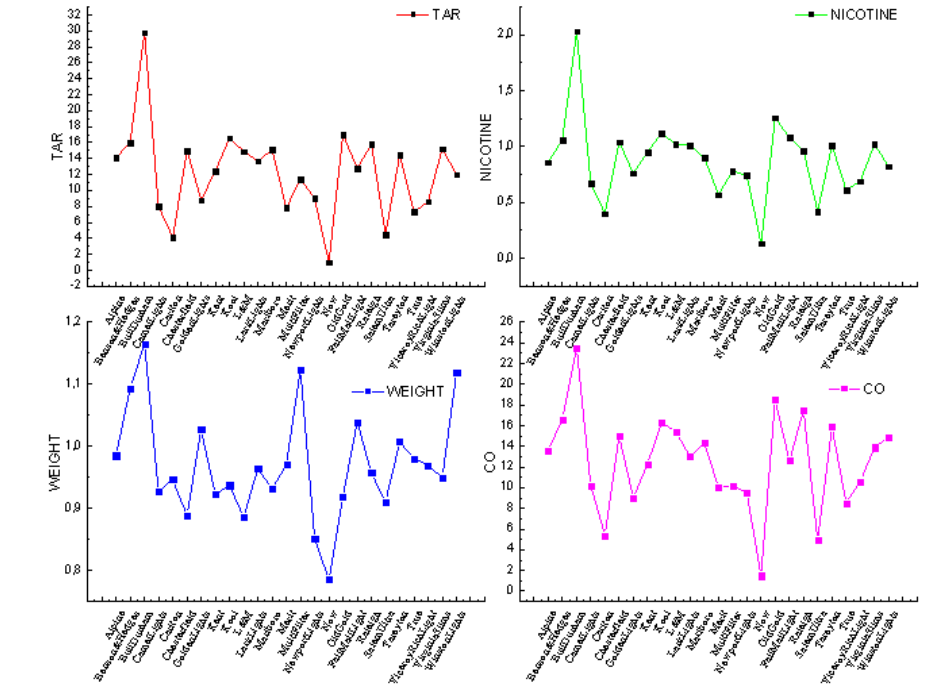


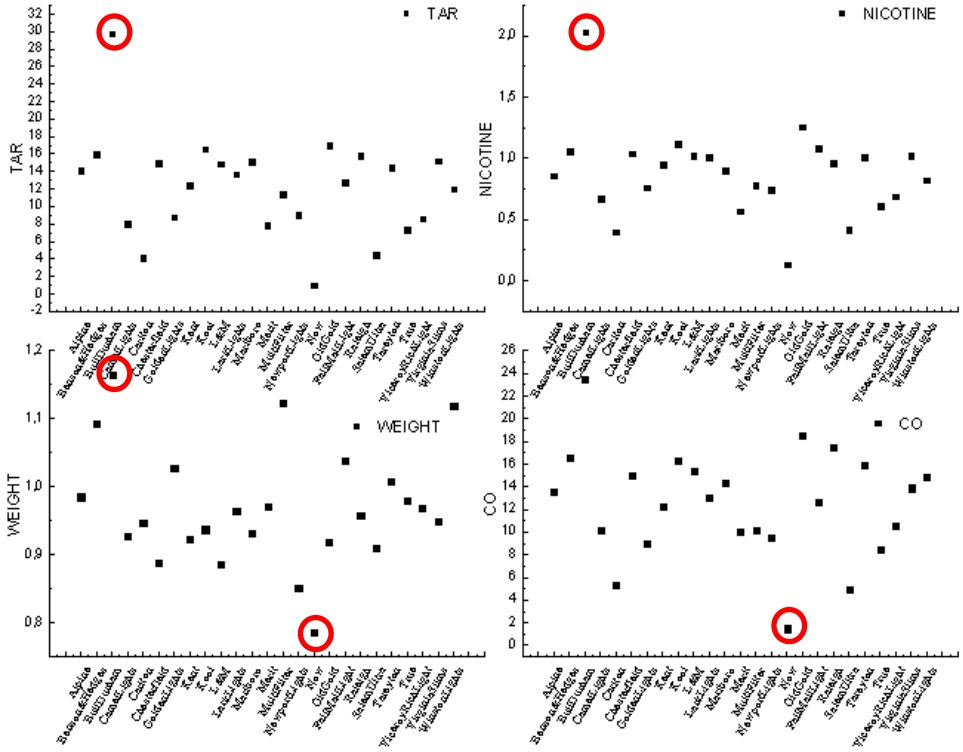
Graf zobrazuje veškeré parametry, tj. TAR, NICOTINE, WEIGHT a CO, přičemž každý z jednotlivých parametrů druhu cigaretového výrobce má vlastní Y novou osu s vlastním měřítkem.

Tento způsob však klade vysoké předpoklady na „uživatele“ grafu, konkrétně na jeho schopnost a zkušenost práce s těmito typy grafů.

Pro „nezkušeného“ uživatele je dle mého soudu tento graf spíše matoucí.

Čtyřúrovňový graf jednotlivých složek typu spojnicový a bodový graf





Tyto typy grafu již znázorňují mnohem lépe odlehlé hodnoty jednolivých parametrů uložených v layerech. Odlehlé hodnoty, které jsou ve shodě s box plotovými grafy (značené *) jsou označeny červeným kroužkem.

Aplikace přímkové regrese na jednotlivé složky v jednotlivých layerech

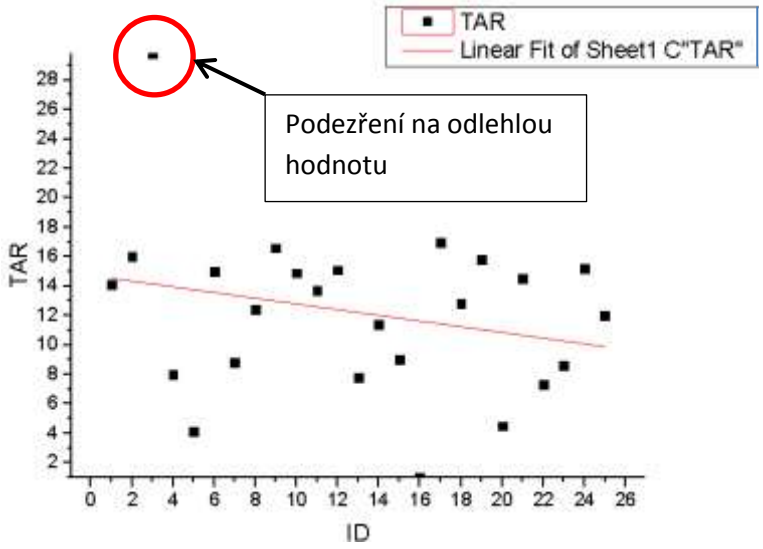
Description	Perform Linear Fitting
User Name	ucitel
Operation Time	16.5.2016 13:16:58
Equation	y = a + b*x
Report Status	New Analysis Report
Weight	No Weighting
Multi-Data Fit Mode	Independent Fit - Consolidated Report
Special Input Handling	
Data Filter	No

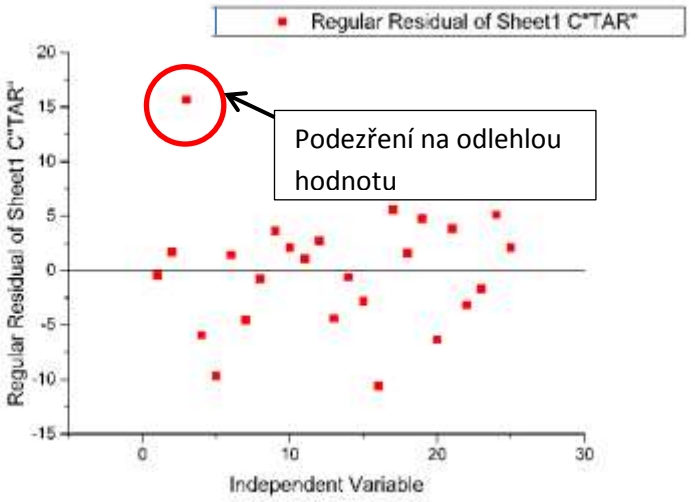
	Input X Data Source	Input Y Data Source	Range
	[Book1]Sheet1!A"ID"	[Book1]Sheet1!C"TAR"	[1*:30*]
	[Book1]Sheet1!A"ID"	[Book1]Sheet1!D"NICOTINE"	[1*:30*]
	[Book1]Sheet1!A"ID"	[Book1]Sheet1!E"WEIGHT"	[1*:30*]
	[Book1]Sheet1!A"ID"	[Book1]Sheet1!F"CO"	[1*:30*]

		Value	Standard Error
TAR	Intercept	14,734	2,30955
	Slope	-0,19369	0,15536
NICOTINE	Intercept	1,0288	0,14463
	Slope	-0,01172	0,00973
WEIGHT	Intercept	0,98101	0,03686
	Slope	-8,24769E-4	0,00248
CO	Intercept	13,895	1,96946
	Slope	-0,10515	0,13248

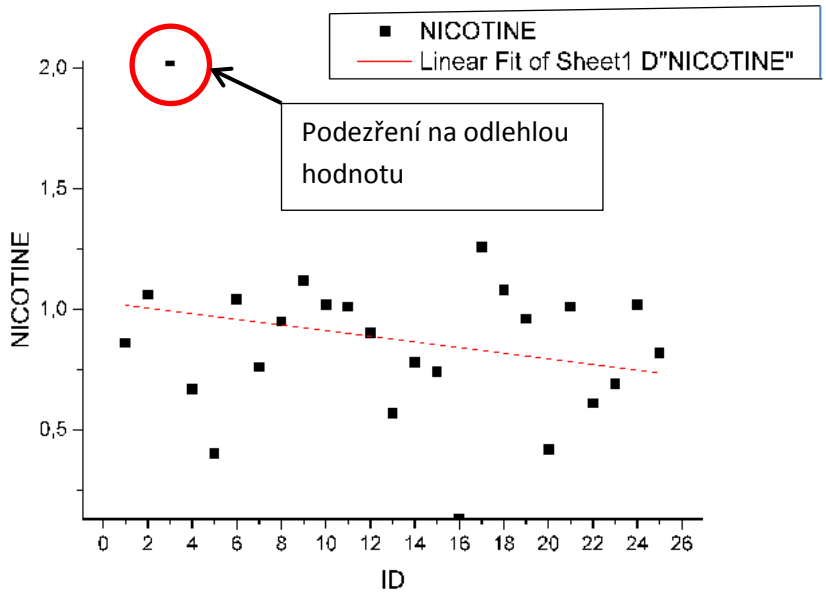
Ve výše uvedené tabulce je možné vysledovat hodnoty Intercept a Slope (tedy úsek na ose y a směrnici regresní rovnice přímky) pro parametry TAR a CO, NICOTINE a WEIGHT.

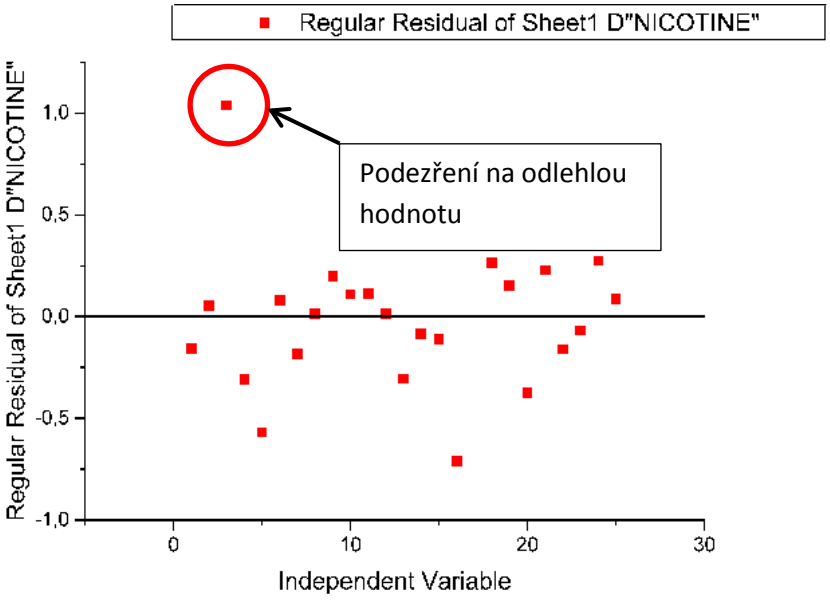
Aplikace přímkové regrese na složku TAR



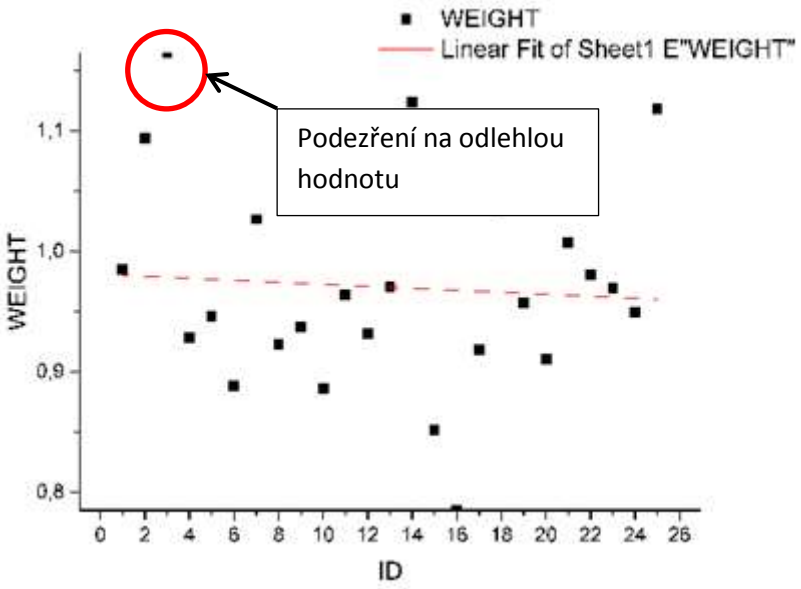


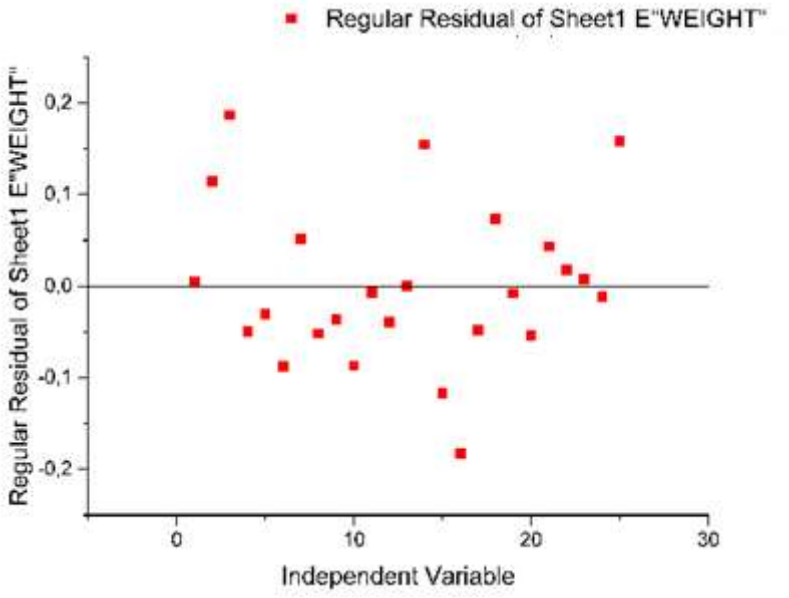
Aplikace přímkové regrese na složku Nicotine



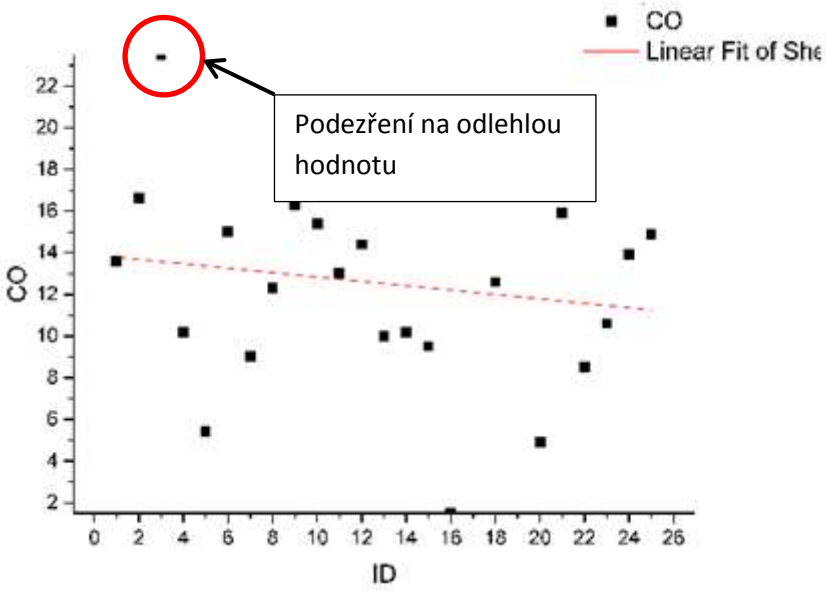


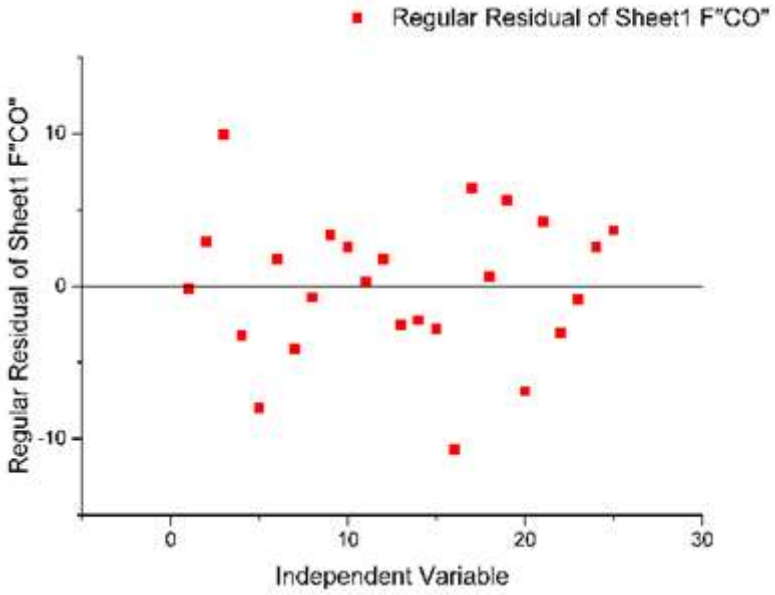
Aplikace přímkové regrese na složku Weight (hmotnost cigarety v g)





Aplikace přímkové regrese na složku CO

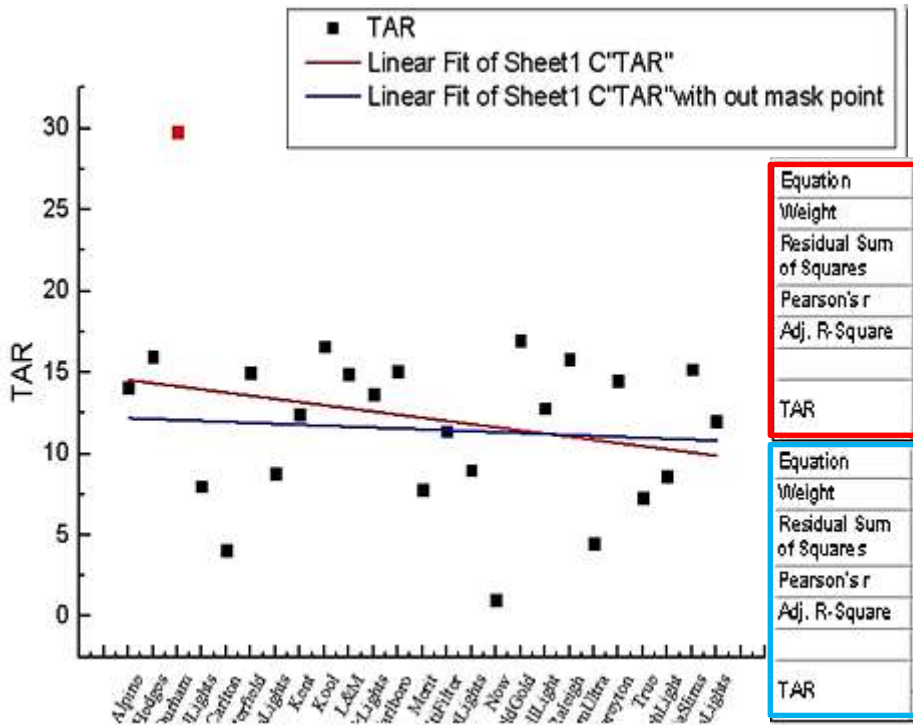




Subzávěry:

Ve všech čtyřech případech parametrů TAR, NICOTINE, WEIGHT a CO je možné konstatovat, že v důsledku odlehlých hodnot je regresní přímka jimi ovlivněna. Toto plně koresponduje j Box plotovými grafy, grafy bodovými a grafy trendovými.

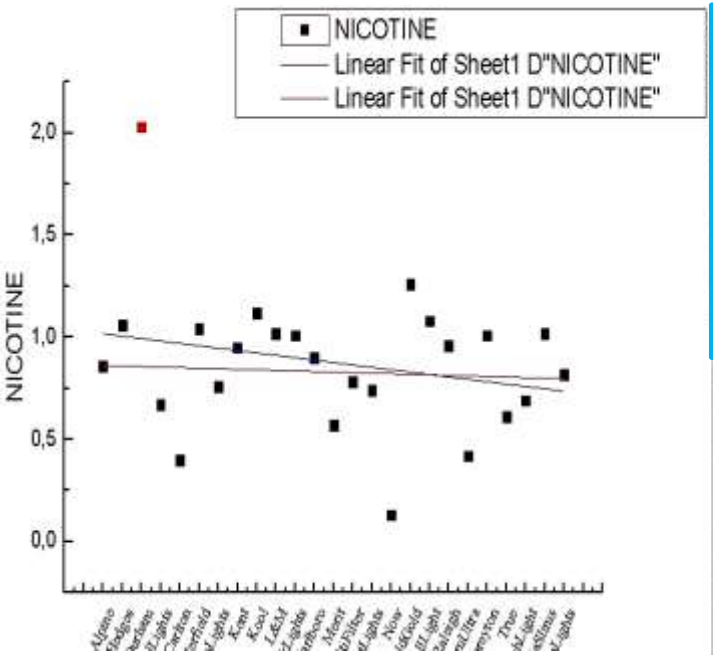
Přímková regrese pro parametr TAR s uvažováním a vypuštěním podezřelého bodu



Equation	y = a + b*x		
Weight	No Weighting		
Residual Sum of Squares	721,66188		
Pearson's r	-0,2516		
Adj. R-Square	0,02258		
		Value	Standard Error
TAR	Intercept	14,734	2,30955
	Slope	-0,19369	0,15536

Equation	y = a + b*x		
Weight	No Weighting		
Residual Sum of Squares	444,41421		
Pearson's r	-0,09373		
Adj. R-Square	-0,03627		
		Value	Standard Error
TAR	Intercept	12,25337	1,9704
	Slope	-0,05739	0,12997

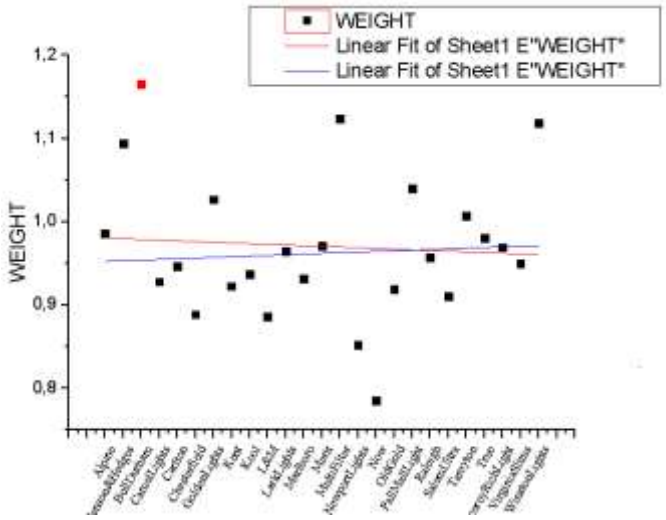
Přímková regrese pro parametr NICOTINE s uvažováním a vypuštěním podezřelého bodu



Equation	y = a + b*x		
Weight	No Weighting		
Residual Sum of Squares	2,82992		
Pearson's r	-0,24369		
Adj. R-Square	0,01849		
		Value	Standard Error
NICOTINE	Intercept	1,0288	0,14463
	Slope	-0,01172	0,00973

Equation	y = a + b*x		
Weight	No Weighting		
Residual Sum of Squares	1,81364		
Pearson's r	-0,07318		
Adj. R-Square	-0,03986		
		Value	Standard Error
NICOTINE	Intercept	0,8646	0,11873
	Slope	-0,0027	0,00783

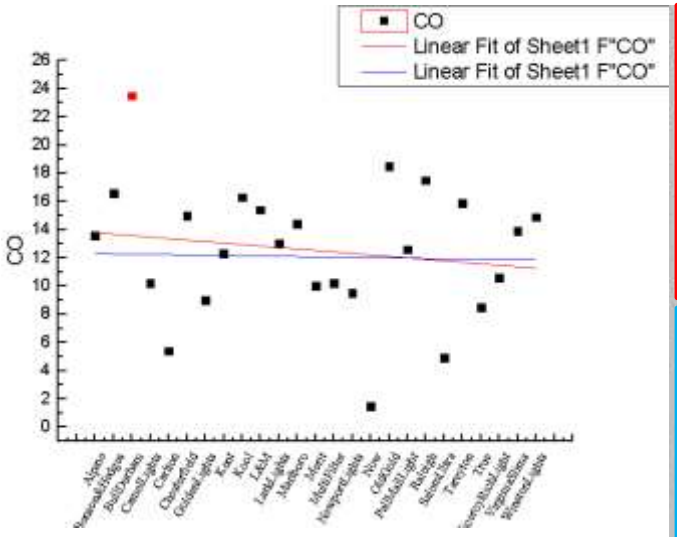
Přímková regrese pro parametr WEIGHT s uvažováním a vypuštěním podezřelého bodu



Equation	y = a + b*x		
Weight	No Weighting		
Residual Sum of Squares	0,1838		
Pearson's r	-0,0692		
Adj. R-Square	-0,03848		
		Value	Standard Error
WEIGHT	Intercept	0,98101	0,03686
	Slope	-8,24769E-4	0,00248

Equation	y = a + b*x		
Weight	No Weighting		
Residual Sum of Squares	0,14442		
Pearson's r	0,07256		
Adj. R-Square	-0,03995		
		Value	Standard Error
WEIGHT	Intercept	0,95144	0,03552
	Slope	7,99519E-4	0,00234

Přímková regrese pro parametr CO s uvažováním a vypuštěním podezřelého bodu



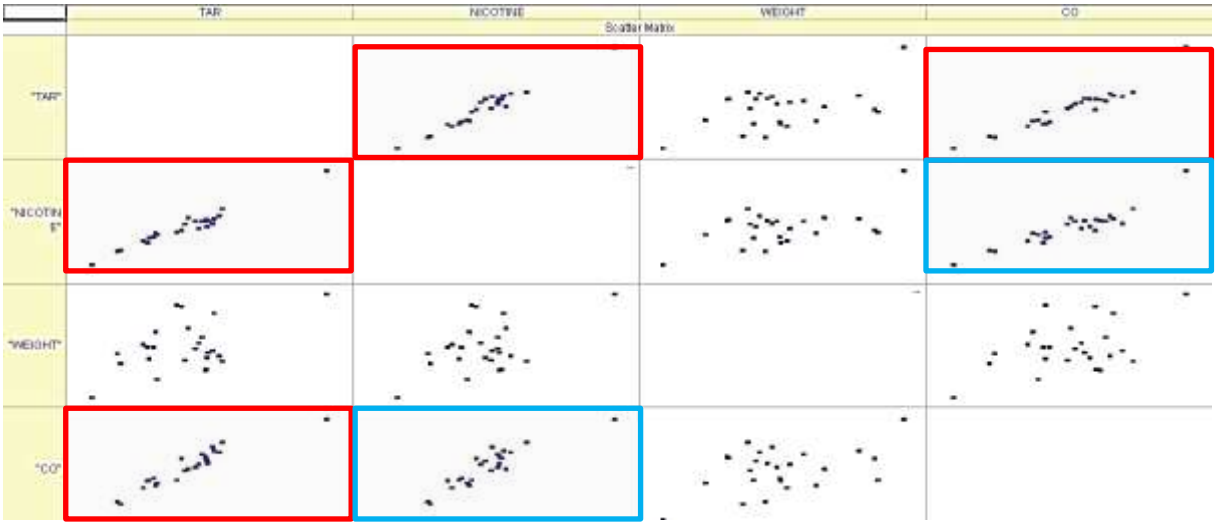
Equation	y = a + b*x		
Weight	No Weighting		
Residual Sum of Squares	524,77587		
Pearson's r	-0,16328		
Adj. R-Square	-0,01566		
		Value	Standard Error
CO	Intercept	13,895	1,96946
	Slope	-0,10515	0,13248

Equation	y = a + b*x		
Weight	No Weighting		
Residual Sum of Squares	413,32968		
Pearson's r	-0,03186		
Adj. R-Square	-0,04439		
		Value	Standard Error
CO	Intercept	12,32224	1,90024
	Slope	-0,01874	0,12534

Při hodnocení regresních přímek parametrů TAR, NICOTINE, WEIGHT a CO je možné konstatovat, že po vypuštění výše popisovaných odlehlých hodnot dochází ke změně parametrů regresní přímky – viz tab. horní (bez vypuštění) a dolní (po vypuštění) příslušného regresního grafu.

Vypuštěný bod (vypuštění je provedeno maskováním) je označen červenou barvou a plně koresponduje s výše uvedeným.

Korelace graficky a numericky



Výpočet Pearsonova korelačního koeficientu pro jednotlivé složky

Pearson Correlations

		TAR	NICOTINE	WEIGHT	CO
"TAR"	Pearson Corr.	1	0,97661	0,49077	0,95749
	Sig.	--	0	0,01274	6,55032E-14
"NICOTINE"	Pearson Corr	0,97661	1	0,50018	0,92595
	Sig.	0	--	0,01089	3,31173E-11
"WEIGHT"	Pearson Corr.	0,49077	0,50018	1	0,46396
	Sig.	0,01274	0,01089	--	0,01948
"CO"	Pearson Corr	0,95749	0,92595	0,46396	1
	Sig.	6,55032E-14	3,31173E-11	0,01948	--

2-tailed test of significance is used

Závěr:

Z výše uvedených hodnot Pearsonova korelačního koeficientu i grafického znázornění korelace parametrů TAR, NICOTINE, WEIGHT a CO je možné konstatovat, že vysokých hodnot nabývají:

Korelující parametry	Hodnota Pearsonova korelačního koeficientu
TAR cor NICOTINE	0,97661
TAR cor CO	0,95749
NICOTINE cor CO	0,92595

Použitá literatura

1. MELOUN, Milan a Jiří MILITKÝ. *Statistická analýza experimentálních dat*. Vyd. 2. uprav. rozš. Praha: ACADEMIA, 2004, 953 s. ISBN 80-200-1254-0.
2. MELOUN, Milan. *Kompendium statistického zpracování dat: metody a řešené úlohy*. Vyd. 2., přeprac. a rozš. Praha: Academia, 2006, 982 s. ISBN 80-200-1396-2.
3. MELOUN, Milan. *Statistické zpracování experimentálních dat*. 2. vyd. Praha: Ars magna, 1998, 839 s. ISBN 80-721-9003-2