

5 Resultados e Discussão

5.1. Cromatografia dos padrões PIANO

A partir do padrão PIANO foram obtidos cromatogramas de cada uma das frações de hidrocarbonetos primeiramente em separado (parafinas, aromáticos, naftênicos e olefinas). Em seguida foi obtido o cromatograma com padrão PIANO contendo todas as frações de hidrocarbonetos. Após a correção dos tempos de retenção com cada uma das frações, montou-se a tabela de referência com um total de 139 compostos, conforme resultado predito pelo padrão PIANO.

Para testar a eficiência da tabela de referência foi feita análise de uma amostra de gasolina comercial, onde se verificou que alguns picos de área considerável não haviam sido identificados na análise PONA.

Com o objetivo de melhorar a eficiência da tabela de referência, foi feita uma análise por espectrometria de massas com cromatografia gasosa acoplada (CG-MS), sendo novos compostos identificados e incluídos na tabela 5 (tabela de referência) perfazendo um total de 172 compostos. A tabela de referência mostra os compostos com seus respectivos tempos de retenção e grupo componente.

Tabela 5 – Tabela de referência obtida por CG-MS
(A – aromático; E – etanol; N – naftênico; O – olefina; P – parafina).

Pico - nº	composto	T.R.	grupo
2	isobutano	4,39	P
4	2-metil-1propeno	4,70	O
6	n-butano	4,80	P
7	Trans-2-buteno	4,98	O
10	etanol	5,37	E
11	3-metil-1-buteno	5,47	O
12	isopentano	5,83	P
14	1-penteno	6,15	O
16	2-metil-1-buteno	6,32	O

17	n-pentano	6,44	P
18	Isoprene	6,59	O
19	Trans-2-penteno	6,64	O
20	2-metil-2-butenos	6,88	O
21	Cis-2-penteno	7,04	O
22	butilciclopropano	7,12	N
23	Trans-1,3-2-penteno	7,54	O
25	2,2-dimetilbutano	7,72	P
26	ciclopenteno	8,46	N
28	4-metil-1-penteno	8,53	O
29	3-metil-1-penteno	8,84	O
30	ciclopentano	8,90	N
31	2,3-dimetilbutano	9,01	P
35	2-metilpentano	9,11	P
38	3-metilpentano	9,90	P
39	2-metil-1-penteno	10,25	O
40	1-hexeno	10,36	O
43	n-hexano	11,02	P
44	Trans-3-hexeno	11,22	O
45	Cis-3-hexeno	11,33	O
47	Trans-2-hexeno	11,50	O
48	2-metil-2-penteno	11,68	O
49	3,3-dimetilciclobuteno	11,75	N
50	Cis-3-metil-2-butenos	11,92	O
51	4-metilciclopenteno	12,21	N
52	Cis-2-hexeno	12,55	O
53	Trans-3-metil-2-penteno	12,90	O
55	2,2-dimeteptano	13,14	P
56	metilciclopentano	13,46	N
58	2,4-dimetilpentano	13,83	P
60	2,2,3-trimetilbutano	14,32	P
64	3-metilciclopentano	15,49	N
65	benzeno	16,06	A
68	3,3-dimetilpentano	16,50	P
71	cicloexano	16,75	N
78	2-metilexano	17,88	P
79	2,3-dimetilpentano	18,21	P
80	1,1-dimetilciclopentano	18,53	N
82	3-metilexano	19,73	P
84	Cis-1,3-dimetilciclopentano	20,15	N
85	Trans-1,3-dimetilciclopentano	20,56	N
86	3-etilpentano	20,81	P
87	Trans-1,2-dimetilciclopentano	21,06	N
88	1-hepteno	21,19	O
93	Trans-3-hepteno	22,55	O
95	n-heptano	22,80	P
96	Cis-3-hepteno	22,93	O

97	2metil-2-hexeno	23,18	O
99	Trans-2-hepteno	23,49	O
101	4-metil-1,4-hexadieno	24,26	O
103	Cis-2-hepteno	25,08	O
106	metilcicloexano	25,46	N
108	2,2-dimetilexano	25,74	P
111	etilciclopentano	26,77	N
112	2,5-dimetilexano	27,00	P
113	2,2,3-trimetilpentano	27,64	P
114	2,4-dimetilexano	27,82	P
115	1-cis=2-trans-4-cis-trimetilciclopentano	28,97	N
116	1-etilciclopenteno	29,14	N
117	cis-1-trans-2- cis-3-trimetilciclopentano	29,71	N
126	tolueno	29,60	A
132	2,3-dimetilexano	30,89	P
135	2-metileptano	31,19	P
136	4-metileptano	31,46	P
139	cis-1,2,4- trimetilciclopentano	31,82	N
141	3-metileptano	32,03	P
142	3-etilexano	32,31	P
143	cis-1,2- trans-4- trimetilciclopentano	32,59	N
146	trans-1,4- dimetilcicloexano	32,60	N
147	1-etil-2-metilciclopentano	33,27	N
148	2-octeno	33,34	O
149	1-octeno	33,57	O
150	1-etil-1-metilciclopentano	33,86	N
162	Trans-1,2-dimetilcicloexano	34,50	N
163	4-octeno	34,59	O
164	n-octano	34,70	P
165	Cis-1,2,3-trimetilciclopentano	34,82	N
168	Trans-2-octeno	35,05	O
169	isopropilciclopentano	35,27	N
174	Cis-2-octeno	35,90	O
182	Cis-1,2-dimetilcicloexano	37,36	N
183	n-propilciclopentano	37,69	N
184	3,5,5-trimetil-1-hexeno	37,96	O
185	Cis-1,3,5-trimetilcicloexano	38,08	N
186	2,5-dimetileptano	38,27	P
194	3,5-dimetileptano(D)	38,39	P
195	1,1,4-trimetilcicloexano	38,37	N
196	3,5-dimetileptano(L)	38,47	P
197	etilbenzeno	39,23	A
198	Cis-1,2-trans-4-trimetilcicloexano	39,45	N
199	2,3-dimetileptano	39,85	P
201	Meta-xileno	39,93	A
202	3,4-dimetileptano(D)	40,45	P

203	3,4-dimetileptano(L)	40,54	P
204	2-metiloctano	40,82	P
205	3-metiloctano	41,35	P
206	Cis-1trans-2-cis-4-trimetilcicloexano	41,47	N
207	3,3-dietilpentano	41,77	P
208	Para-xileno	42,15	A
209	1,1,2-trimetilcicloexano	42,28	N
210	isobutilciclopentano	42,45	N
238	1-noneno	42,58	O
242	isopropilcicloexano	42,99	N
243	Trans-3-noneno	43,15	O
244	cis-3-noneno	43,40	O
245	n-nonano	43,59	P
247	Trans-2-noneno	43,84	O
251	Cis-2-noneno	44,40	O
253	Orto-xileno	44,60	A
255	2,2-dimetiloctano	45,09	P
256	n-butilciclopentano	45,24	N
258	isobutilcicloexano	46,09	N
259	3,3-dimetiloctano	46,49	P
274	isopropilbenzeno	46,83	A
276	n-propilbenzeno	47,40	A
281	3-etil-1-metilbenzeno	47,59	A
282	2,3-dimetiloctano	47,81	P
283	4-etil-1-metilbenzeno	47,95	A
285	2-metilnonano	48,53	P
287	3-etiloctano	48,75	P
288	1,3,5-trimetilbenzeno	48,92	A
290	3-metiloctano	49,01	P
292	2-etil-1-metilbenzeno	49,64	A
293	1,2,4-trimetilbenzeno	49,89	A
294	Trans-1-metil-2-n-propilcicloexano	50,22	N
295	1-deceno	50,37	O
296	tert-butilbenzeno	50,40	A
297	Trans-1-metil-2-(4MP)ciclopentano	50,89	N
298	isobutilbenzeno	50,70	A
299	n-decano	50,84	P
300	Sec-butilbenzeno	51,31	A
301	1-metil-3-isopropilbenzeno	51,76	A
302	1-metil-4-isopropilbenzeno	52,10	A
304	Indana	52,49	A
316	1-metil-2-isopropilbenzeno	52,92	A
317	1-metil-3-propilbenzeno	53,49	A
318	1-metil-4-propilbenzeno	53,75	A
333	n-butilbenzeno	54,01	A
334	5-etil-1,3-dimetilbenzeno	54,47	A
335	1,2-dietilbenzeno	54,67	A

339	1-metil-2-propilbenzeno	55,29	A
340	2-etil-1,4-dimetilbenzeno	55,41	A
345	1Mindana	55,75	A
347	4-etil-1,2-dimetilbenzeno	56,00	A
348	2-etil-1,3-dimetilbenzeno	56,36	A
356	3-etil-1,2-dimetilbenzeno	57,05	A
357	n-undecano	57,24	P
358	1,2,3,4-tetrametil-5-metilbenzeno	57,69	A
361	1,2,4,5-tetrametilbenzeno	57,91	A
366	4Mindana	59,83	A
370	2-metilbutilbenzeno	60,20	A
371	5Mindana	59,93	A
373	n-pentilbenzeno	60,28	A
374	1-tert-3,4-dimetilbenzeno	61,54	A
375	1-etil-4tert-butilbenzeno	61,84	A
396	n-dodecano	62,38	P
397	1,3,5-trietilbenzeno	63,39	A
398	1,2,4-trietilbenzeno	64,30	A
399	n-hexilbenzeno	65,27	A
422	n-tridecano	66,85	P
436	n-tetradecano	70,58	P
437	n-pentadecano	74,02	P

Com os dados da tabela de referência corrigidos, foram obtidos os cromatogramas das 120 amostras de gasolinas comerciais para obtenção dos valores de PONA de cada uma delas.

Um exemplo de cromatograma obtido para as amostras de gasolinas é apresentado na figura 38.

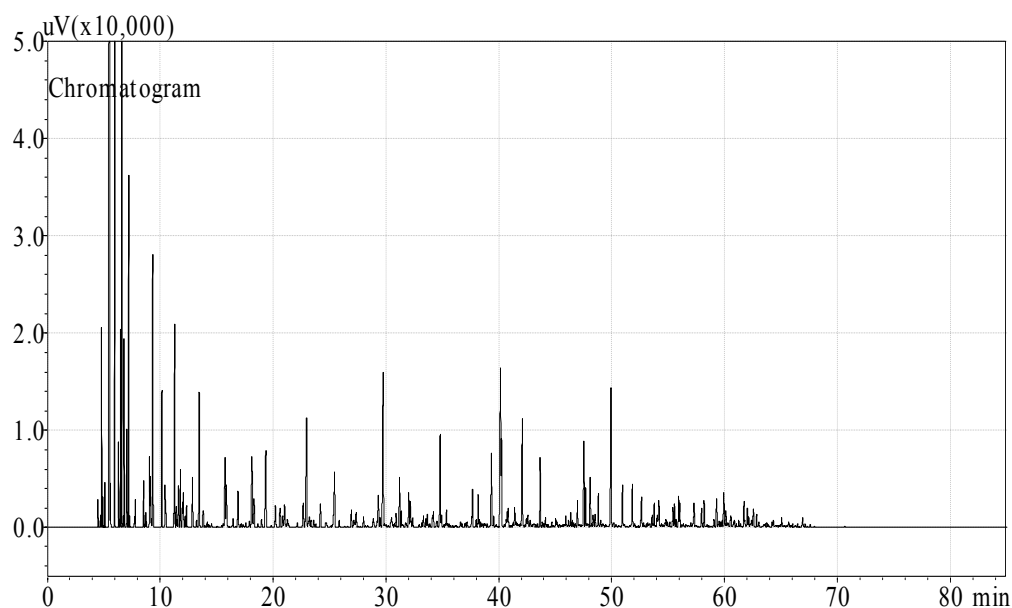


Figura 38 – Cromatograma em fase gasosa da amostra de gasolina 7514.

5.2. Fator de correção para o etanol

Para determinação do teor de etanol nas amostras de gasolinas, se fez necessário a introdução de um fator de correção de acordo com a resposta do equipamento empregado para a cromatografia gasosa durante a obtenção dos resultados pelo método PONA já que este método não prevê a presença de etanol. Para os hidrocarbonetos, os fatores de resposta já estão inseridos no programa.

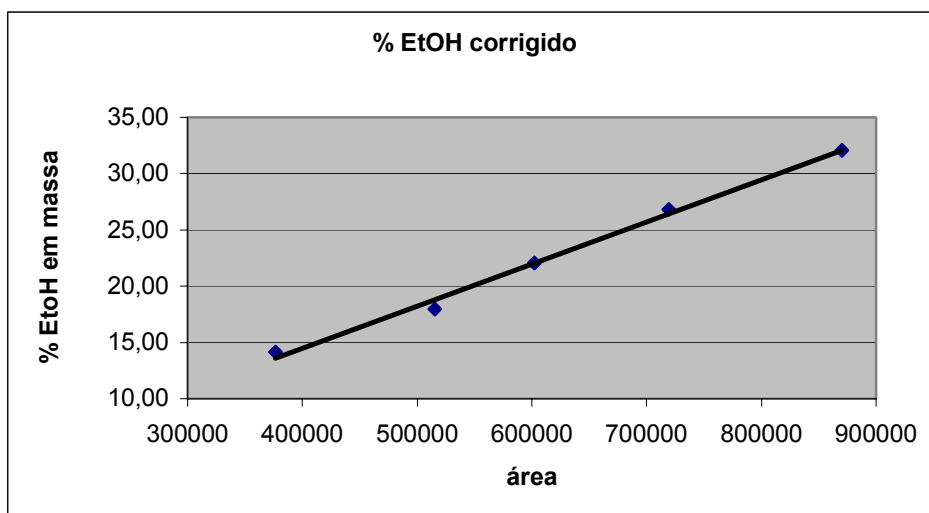
A princípio foram feitas soluções com concentrações variadas com objetivo de se determinar o tempo de retenção específico para o etanol, sendo encontrado o valor de 5,371 min.

Para se determinar o fator de correção da área par o etanol, foram utilizadas soluções de concentrações diferentes de etanol anidro com gasolina tipo A, obtendo-se as áreas características de etanol para cada concentração, calculando-se o fator de correção e o coeficiente de correlação entre a concentração de etanol versus a área do cromatograma, conforme tabela 6.

Tabela 6 – Cálculo do fator de correção para o etanol.

amostras	EtOH % teórico	área EtOH	área total	% HC	área HC	área EtOH teórico.	Fator	área EtOH cor.	área Total cor.	% EtOH cor.
1	14,08	240431	2526689	85,92	2286258	374639	1,56	376354	2662612	14,13
2	18,73	329184	2681767	81,27	2352583	542108	1,65	515282	2867865	17,97
3	21,88	384731	2513714	78,12	2128983	596324	1,55	602231	2731215	22,05
4	26,38	459543	2424534	73,62	1964991	704202	1,53	719337	2684328	26,80
5	31,72	555788	2397564	68,28	1841776	855504	1,54	869992	2711769	32,08

A curva de regressão linear mostrou uma correlação satisfatória entre a área de etanol corrigida e a concentração de etanol nas amostras de gasolinas, conforme mostra a figura 39.

Figura 39 – Teor de etanol corrigido (% m); $R^2 = 0,9945$ para fator de correção de 1,56.

O fator de correção 1,56 foi utilizado para corrigir a área de etanol anidro para todas as amostras de gasolinas, sendo portanto recalculadas as porcentagens em massa para os grupos componentes restantes. Estes resultados foram considerados os valores de referência.

Para a montagem dos modelos dos PCA e PLS foram utilizados os resultados finais de PONA obtidos por cromatografia gasosa das 120 amostras de gasolinas conforme tabela 7 e 8, respectivamente, utilizando-se 84 amostras para o conjunto calibração e 36 amostras para o conjunto teste.

Tabela 7 – Resultados de PONA para amostras de gasolinas – conjunto calibração.

Amostras	% aromáticos	% etanol	% naftênicos	% olefinas	% parafinas
7263	15,11	25,67	14,39	17,75	27,09
7271	15,36	25,71	13,84	17,70	27,40
7279	14,61	25,79	13,08	18,23	28,29
7284	14,70	25,44	14,29	17,66	27,92
7285	15,06	25,42	13,43	17,43	28,66
7289	14,54	25,45	14,39	16,86	28,76
7292	15,31	25,43	13,78	17,29	28,19
7295	15,14	25,41	13,49	17,26	28,70
7296	15,32	25,78	14,44	16,79	27,67
7300	15,57	25,19	13,39	17,61	28,23
7303	13,61	24,04	16,03	18,09	28,23
7305	15,97	27,01	12,58	19,34	25,10
7306	14,44	25,63	12,76	20,87	26,29
7311	15,84	25,66	13,22	17,45	27,84
7312	14,95	28,05	12,73	16,95	27,32
7313	16,11	25,82	13,49	17,27	27,31
7314	16,06	25,74	13,45	17,21	27,22
7321	15,29	25,49	13,40	17,80	28,02
7325	15,16	25,36	14,05	17,50	27,94
7329	15,14	25,61	13,83	17,35	28,07
7337	14,17	25,34	14,34	17,62	28,53
7341	14,85	25,49	14,29	17,46	27,91
7345	15,40	25,29	14,00	17,10	28,21
7361	15,56	25,33	13,53	17,26	28,32
7377	14,16	25,85	13,97	16,33	29,69
7393	15,22	26,09	13,95	16,95	27,78
7401	14,74	25,19	13,45	16,29	30,33
7402	15,37	25,14	14,52	18,29	26,67
7406	15,11	25,64	12,56	21,62	25,06
7502	15,49	25,44	12,61	16,84	29,62
7507	14,97	24,86	13,55	16,87	29,75
7508	16,17	25,00	12,72	17,19	28,91
7512	16,11	24,97	12,46	16,93	29,53
7514	15,94	25,80	12,10	16,20	29,97
7590	12,54	25,03	13,79	16,38	32,25
7594	15,45	24,87	13,11	17,68	28,88
7599	15,75	25,51	13,50	17,03	28,21
7600	14,63	24,91	12,73	17,20	30,54
7603	14,65	25,04	12,90	17,19	30,23
7605	13,62	25,47	13,65	16,25	31,01
7606	15,33	25,63	13,30	17,09	28,65
7610	15,53	25,38	13,18	17,30	28,60

7611	16,04	25,68	13,75	17,12	27,41
7615	12,70	25,23	14,02	16,03	32,03
7618	11,83	25,18	13,61	15,86	33,53
7627	15,53	25,40	13,44	17,51	28,12
7629	14,24	25,49	13,71	17,95	28,61
7631	14,23	25,40	13,19	17,62	29,56
7634	15,92	25,81	11,46	19,62	27,20
7635	15,94	25,76	13,09	16,84	28,36
7639	14,22	25,25	13,90	18,20	28,43
7643	14,83	24,51	13,16	17,91	29,60
7645	14,44	25,03	12,73	18,43	29,36
7649	15,19	24,96	13,84	17,77	28,24
7650	15,36	25,23	13,78	17,57	28,06
7651	14,38	25,61	13,86	17,79	28,36
7652	14,38	25,31	13,79	17,86	28,66
7654	12,06	25,21	14,08	16,33	32,31
7656	15,62	25,54	13,54	17,30	28,00
7658	16,16	24,95	11,62	17,01	30,26
7661	16,58	24,88	12,84	17,57	28,13
7662	16,04	24,88	12,55	17,51	29,02
7694	16,11	25,03	12,55	17,32	28,99
7696	16,93	25,46	11,55	16,45	29,60
7699	16,47	24,99	12,48	17,33	28,73
7700	16,10	25,66	12,80	17,33	28,11
7701	16,27	24,92	12,12	17,89	28,80
7702	15,79	25,83	11,00	16,85	30,53
7703	16,47	25,76	11,15	16,87	29,75
7706	16,29	25,02	11,45	17,36	29,88
7707	16,82	25,02	12,48	16,05	29,64
7710	16,57	24,81	12,09	17,15	29,39
7711	16,00	24,95	12,39	17,55	29,11
7717	15,35	24,84	10,77	15,27	33,78
7720	20,15	25,00	9,58	17,47	27,81
7722	19,49	26,06	7,61	21,25	25,58
7724	19,35	25,54	8,14	18,58	28,39
7726	15,52	24,99	12,19	17,44	29,87
7727	14,92	24,66	10,84	15,44	34,14
7728	17,87	25,08	9,08	16,74	31,24
7729	15,93	25,32	11,87	17,10	29,77
7730	19,13	25,34	8,91	17,60	29,02
7731	18,80	25,04	8,82	17,21	30,13
7732	16,37	25,16	11,49	17,09	29,88

Tabela 8 – Resultados de PONA para amostras de gasolinas – conjunto teste.

amostra	% aromáticos	% etanol	% naftênicos	% olefinas	% parafinas
7277	14,52	25,15	14,02	17,96	28,35
7286	14,78	25,47	14,66	16,72	28,37
7294	15,64	25,70	13,58	17,10	27,97
7298	15,01	25,35	13,84	17,65	28,15
7301	14,97	25,88	13,53	17,28	28,35
7304	15,25	25,39	13,37	17,36	28,63
7307	14,47	26,53	14,58	17,25	27,16
7308	15,39	25,29	13,66	17,64	28,02
7309	15,01	25,35	13,84	17,65	28,15
7317	13,63	27,60	12,61	16,85	29,30
7365	15,36	25,47	13,15	17,12	28,90
7369	14,04	25,49	13,65	16,81	30,01
7381	15,25	25,69	13,77	17,31	27,99
7403	15,45	25,32	14,38	17,90	26,95
7503	16,08	25,34	13,10	17,35	28,12
7510	15,83	24,78	12,76	17,31	29,32
7511	16,76	25,17	13,45	17,19	27,42
7591	14,22	25,77	13,75	16,00	30,27
7595	15,59	25,34	13,62	17,89	27,56
7598	14,76	25,56	12,81	16,55	30,31
7601	14,85	25,18	13,00	17,37	29,61
7602	15,62	24,85	13,05	17,47	29,01
7604	13,89	25,29	14,06	18,15	28,60
7607	16,37	26,35	10,95	22,17	24,15
7624	13,89	25,03	13,70	18,23	29,14
7638	14,24	25,42	13,78	18,00	28,56
7641	16,46	24,95	12,69	17,02	28,88
7647	14,41	25,60	13,01	18,80	28,18
7648	14,60	24,82	12,91	19,06	28,61
7653	16,30	24,77	13,44	17,65	27,84
7655	16,17	25,30	10,40	18,66	29,47
7657	15,24	25,53	13,81	16,14	29,28
7682	16,66	24,77	11,95	17,35	29,26
7697	15,43	25,33	12,76	17,72	28,76
7714	19,34	24,79	9,04	17,77	29,06
7725	18,86	24,90	9,15	17,49	29,60

5.3. - Análise dos componentes principais (PCA) dos Espectros FT-Raman

O conjunto de espectros FT-Raman das amostras de gasolinas obtidos (figura 40), mostra que algumas amostras (7512; 7618 e 7722) diferiram das demais nas regiões de comprimento de onda entre $110 - 2050 \text{ cm}^{-1}$ e $2300 - 3600 \text{ cm}^{-1}$. Este comportamento anômalo pode estar ligado ao fenômeno de fluorescência relatado por Hansens (29) que é possível ocorrer em espectros FT-Raman em amostras de gasolinas. Com tratamentos matemáticos pode-se corrigir esta fluorescência, podendo este espectro ser remarcado. Apesar dessas amostras em questão aparentarem serem amostras anômalas, elas não foram descartadas inicialmente, pois não havia nenhum qualquer outro motivo relevante para isto.

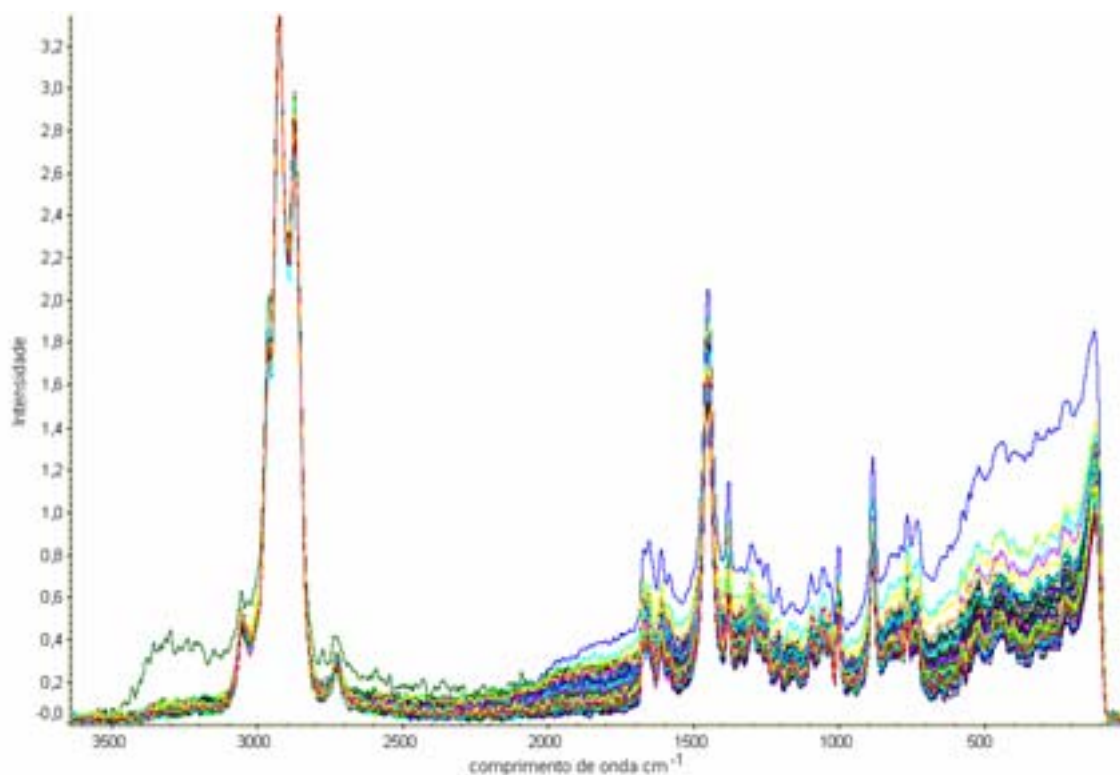


Figura 40 – Espectro FT-Raman de 120 amostras de gasolina, na região de $98 - 3600 \text{ cm}^{-1}$.

Uma amostra é chamada de anômala (*outlier*) quando ela apresenta-se diferente das demais não podendo ser descrita pelo modelo ou por influenciá-lo demais. A análise dos componentes principais vai descrever como essa amostra é diferente das outras, mesmo que isso seja irrelevante para modelar as propriedades mais importantes

presentes nas demais amostras. Através da análise dos componentes principais, as amostras *outliers* detectadas visualmente podem ser confirmadas através dos gráficos dos “scores” e dos valores de resíduos e “leverages”, em geral combinados em um gráfico de influência (41).

O gráfico dos “scores” mostra o perfil das amostras em relação a um, dois ou três componentes principais. Neste gráfico identifica-se uma amostra que se situe longe das demais, indicando se esta é uma amostra “*outlier*”.

Os resíduos são uma medida de quanto as amostras ajustam-se ao modelo. Uma amostra com resíduo alto é pobremente descrita pelo modelo, o qual descreve bem as demais amostras. Esse tipo de amostra é considerado como um elemento estranho em relação ao conjunto de amostras bem descritas pelo modelo, podendo ser considerada uma amostra “*outlier*”.

“*Leverage*” é uma medida de quanto um ponto de um conjunto de dados é extremo, comparado com a maioria. Em PCA e PLS, “*leverage*” pode ser interpretado como a distância entre o ponto projetado e o centro do modelo. Pontos com alto “*leverage*” têm uma alta influência no modelo e podem ou não ser amostras “*outliers*”. Uma amostra com alto “*leverage*” e alto resíduo é considerada um caso crítico e pode ser detectada usando-se o gráfico de influência.

A análise dos componentes principais, com os dados centrados, foi realizada com todas as variáveis X (comprimento de onda) disponíveis, indicando que 03 (três) componentes principais (PCs) descrevem bem o modelo construído a partir dos espectros obtidos como mostra a figura 41 (linha verde do gráfico), e confirma que algumas amostras apresentam um comportamento diferenciado das demais, podendo ser consideradas amostras “*outliers*” (amostras anômalas). Os gráficos dos “scores” e de influência são mostrados, respectivamente, nas figuras 42 e 43, para as amostras de gasolinas. Nesses gráficos, os pontos envolvidos com um círculo referem-se às amostras que apresentaram o fenômeno de fluorescência durante o processo de obtenção dos espectros FT-Raman.

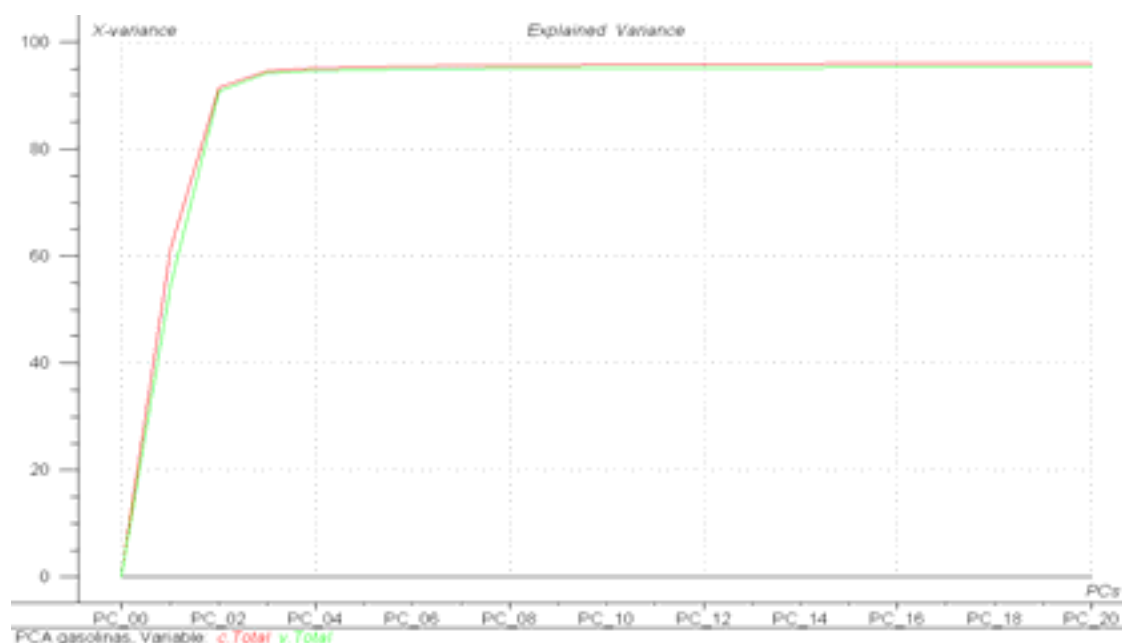


Figura 41 – Gráfico da variância, indicando 3 componentes principais para o modelo descrito a partir dos espectros FT-Raman de amostras de gasolinas.

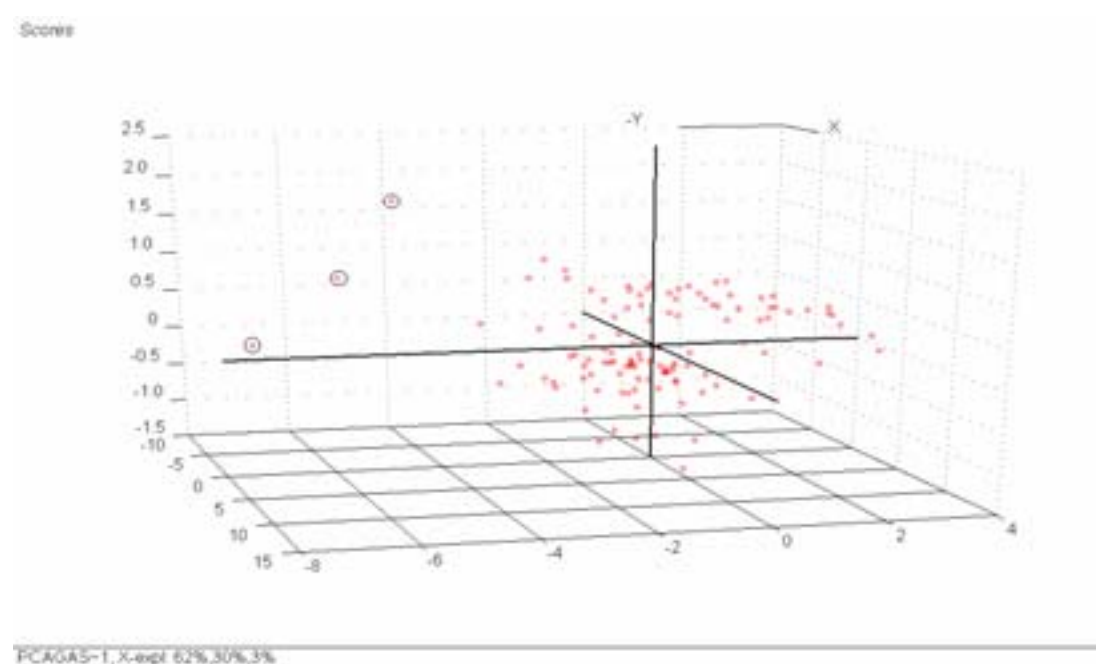


Figura 42 – Gráfico dos “scores” nos três primeiros componentes principais dos espectros FT-Raman das amostras de gasolinas da figura 40. Os pontos marcados com um círculo (○) representam amostras outliers.

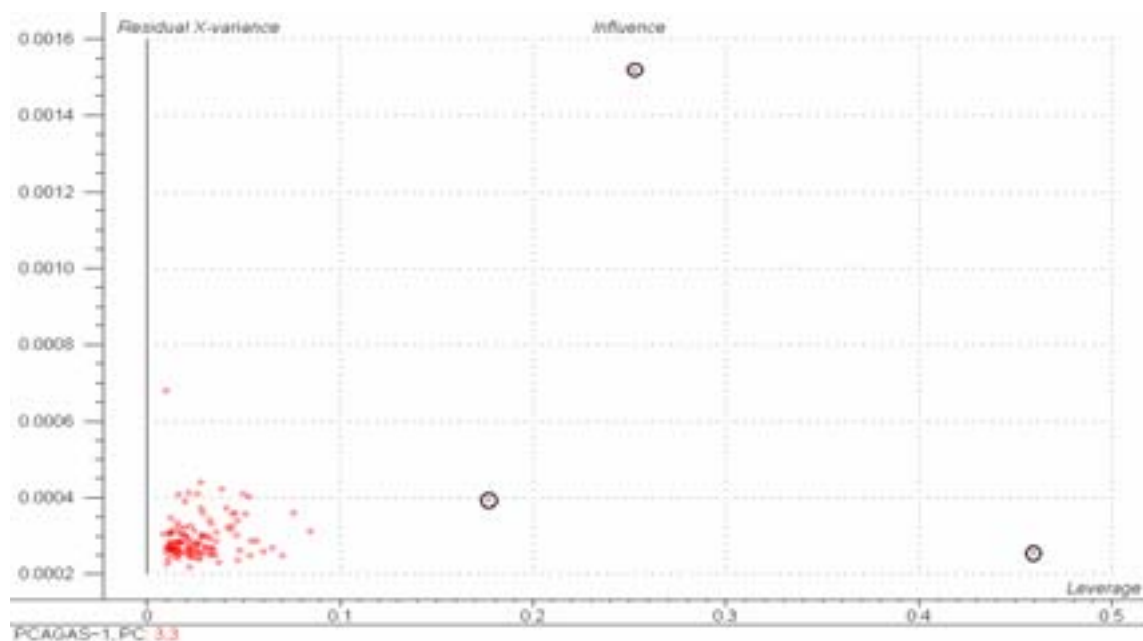


Figura 43 – Gráfico de influência dos espectros FT-Raman de amostras de gasolinas da figura 40.

Os pontos marcados com um círculo (○) representam amostras outliers.

5.4. - Seleção de variáveis

Nem sempre a utilização de todas as variáveis X (comprimentos de ondas) é o melhor caminho para se prever uma determinada propriedade, causando alguns transtornos, por exemplo, tempo computacional excessivo, tendência de superajuste (“*overfitting*”), colinearidades, alta complexidade dos modelos e outros. Para evitar esses transtornos, utilizaram-se alguns critérios de seleção das variáveis:

- utilização de todas regiões disponíveis nos espectros de $98 - 3600\text{ cm}^{-1}$;
- utilização de um conjunto de variáveis relacionados com a propriedade a ser predita, selecionado de acordo com conhecimentos químicos e análise de padrões (PIANO);
- utilização de um conjunto de variáveis selecionadas de acordo com os gráficos fornecidos pelo programa estatístico utilizado: coeficientes de regressão e cargas das variáveis X (“*X-loadings weights*”).

O gráfico dos coeficientes de regressão é utilizado para verificar a importância das diferentes variáveis X na predição de Y. Valores absolutos elevados de coeficientes de

regressão indicam variáveis significantes e valores pequenos de coeficientes de regressão indicam variáveis sem significância (41).

O gráfico das cargas das variáveis X, para um determinado componente principal, contra o número da variável também é utilizado como um dos métodos para se determinar quais variáveis X são importantes para predizer Y. Esse gráfico mostra a relação entre o componente especificado e as diferentes variáveis X. Um valor de carga elevado, positivo ou negativo, significa que a variável é importante para o componente. Variáveis com valores elevados nos primeiros componentes são aquelas que explicam a maior variação em Y (41).

5.5. - Determinação dos modelos de regressão PLS

Foram obtidos diferentes modelos PLS para as diversas propriedades estudadas, utilizando-se o espectro inteiro sem pré-tratamentos; o espectro inteiro com os pré-tratamentos; subconjunto de variáveis sem pré-tratamento e com os pré-tratamentos. Os pré-tratamentos utilizados foram padronização, validação cruzada e centralização dos dados em torno da média. Durante o processo de validação, um passo crítico é estabelecer o número ótimo de componentes principais que serão utilizados no modelo. Um exemplo pode ser visto na figura 44 que apresenta o gráfico RMSEP (que pode ser interpretado como o erro médio quadrático da validação) *versus* o número de componentes principais utilizado, do modelo construído para a predição do teor de etanol das amostras de gasolinas. Pode-se observar que o número ideal de componentes principais com o menor valor de RMSEP é de onze, como pode ser visto pela linha verde do gráfico mostrado na figura 44.

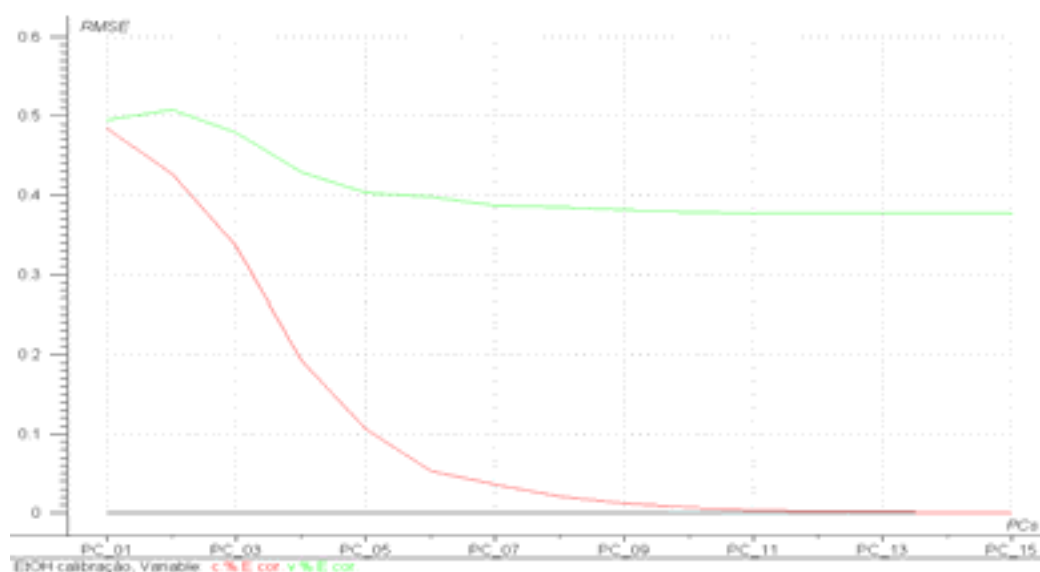


Figura 44 – Gráfico RMSEP *versus* o número de componentes principais, do modelo PLS construído para a determinação do teor de etanol nas amostras de gasolinas.

Os componentes principais utilizados nos modelos finais para as propriedades das amostras de gasolinas estão na tabela 9.

Tabela 9 – Modelo PLS selecionado para predição das propriedades das amostras de gasolinas.

Propriedade (PONA)	Banda (cm ⁻¹)	nº de componentes principais	RMSEP %			média Predição (%)
			calibração	validação	predição	
Parafinas	710 – 840 1280 – 1330 2840 – 2950	4	0,69	0,98	0,77	28,44 ± 1,20
Olefinas	1630 – 1740 2820 – 3100	4	0,45	0,60	0,51	17,46 ± 0,9
naftênicos	720 – 1000 1420 – 1560 2820 – 2960	4	0,62	0,98	0,74	13,21 ± 1,05
aromáticos	900 – 1100 3014 – 3110	5	0,37	0,92	0,79	15,27 ± 0,69
Etanol	200 – 1000	11	0,004	0,38	0,39	25,45 ± 0,3

5.5.1. - Modelo PLS para parafinas

O modelo PLS gerado para o teor de parafinas presentes na gasolina está de acordo com os estudos feitos por Cooper (23), onde ele demonstra que no espectro FT-

Raman a região entre $700 - 800 \text{ cm}^{-1}$ é característica para alcanos ramificados (isoparafinas).

O trabalho de Hansens (29), também, corrobora esta interpretação, com identificação de parafinas na região de 722 cm^{-1} , característico de estiramento simétrico de C – C envolvendo carbono terciário (isoparafinas). As regiões entre $1300 - 1450 \text{ cm}^{-1}$ e a região próximo à $\sim 2875 \text{ cm}^{-1}$ foi também detectada por ele, o que pode justificar a montagem do modelo PLS para predição de parafinas.

Para o conjunto teste das parafinas nenhuma amostra apresentou erro relativo superior a 10% em relação aos valores de referência mostrados na tabela 8. Os resultados de RMSEP do conjunto calibração obtidos foram satisfatórios, menor que 1 (RMSEP = 0,69) como mostra a figura 45. O modelo PLS apresentou também um bom resultado para o coeficiente de correlação ($R^2 = 0,81$) do conjunto calibração entre os resultados de PONA e os resultados obtidos por espectroscopia FT-Raman (figura 45). O conjunto validação apresentou um bom coeficiente de correlação ($R^2 = 0,62$) com RMSEP de 0,98 (figura 46). O conjunto teste apresentou um coeficiente de correlação ($R^2 = 0,5975$) inferior ao de validação com RMSEP de 0,75 (figura 47) o que vem a ser um bom resultado, já que o modelo de análise por multivariadas prevê um erro máximo de 1.

Levando-se em consideração os resultados da tabela 8, comprova-se uma pequena diferença entre os valores de referência e os valores preditos, já que todas as amostras se mantiveram dentro de uma mesma faixa estreita de variação, ou seja, em torno de $28,44\% \pm 1,20$. Este desvio padrão provoca um erro percentual analítico muito baixo (4,21%), causando uma certa centralização dos dados no conjunto de resultados, influenciando o coeficiente de correlação. Mas quando se observa os resultados de predição do conjunto teste (tabelas 8 e 10), constata-se que os resultados de análise obtidos estão muito próximos aos de referência, o que pode ser comprovado pelo RMSEP de 0,75 obtido pela análise PLS.

Tabela 10 – Dados de referência, predição, erro relativo do conjunto teste para parafinas.

amostra	predito	referência	erro %
7277	28,57	28,35	0,75
7286	28,38	28,37	0,04
7294	29,24	27,97	4,54
7298	27,22	28,15	-3,30
7301	29,19	28,35	2,98
7304	28,30	28,63	-1,13
7307	27,58	27,16	1,55
7308	28,30	28,02	1,01
7309	27,19	28,15	-3,40
7365	29,90	28,90	3,45
7369	29,94	30,01	-0,24
7381	27,90	27,99	-0,32
7403	27,17	26,95	0,81
7503	29,32	28,12	4,27
7510	29,54	29,32	0,76
7511	28,41	27,42	3,60
7591	30,49	30,27	0,73
7595	27,30	27,56	-0,97
7598	30,77	30,31	1,51
7601	29,00	29,61	-2,06
7602	28,78	29,01	-0,77
7604	26,92	28,60	-5,89
7607	24,76	24,15	2,53
7624	28,35	29,14	-2,70
7638	28,23	28,56	-1,16
7641	27,35	28,88	-5,30
7647	29,17	28,18	3,49
7648	29,16	28,61	1,93
7682	28,94	29,26	-1,09
7697	27,67	28,76	-3,81
7714	27,84	29,06	-4,19
7725	29,08	29,60	-1,77

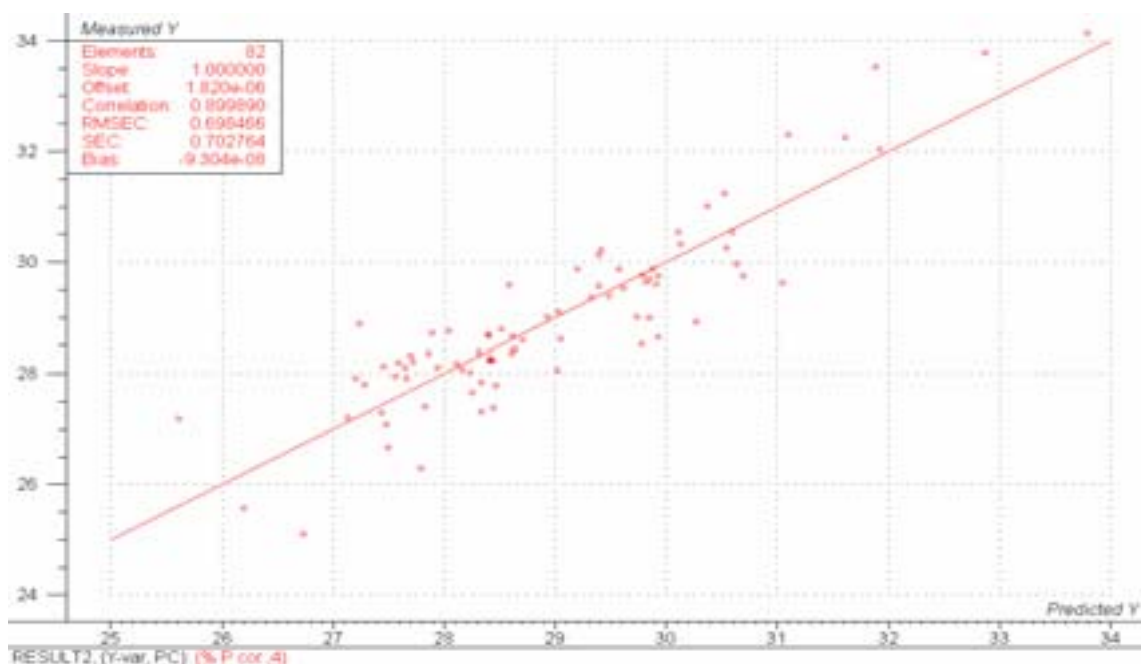


Figura 45 – Teor de parafinas nas amostras de gasolinas conjunto calibração ($R^2 = 0,8098$) ;

— teórico.

RMSEP = 0,69

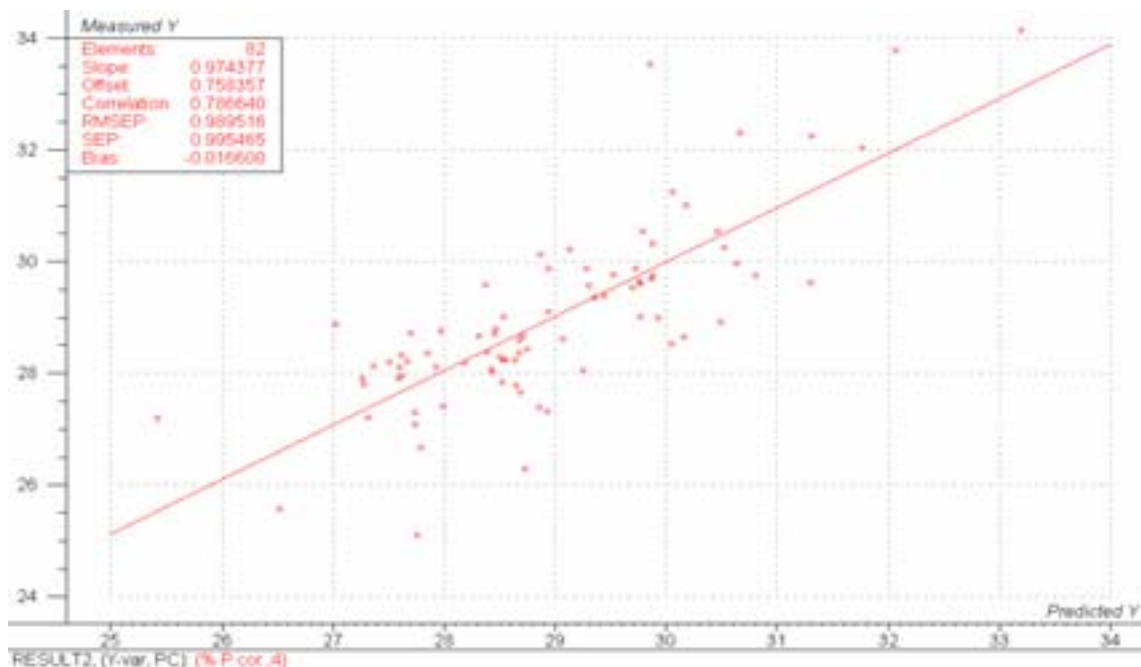


Figura 46 – Teor de parafinas nas amostras de gasolinas conjunto validação ($R^2 = 0,6188$) ;

— teórico.

RMSEP = 0,98

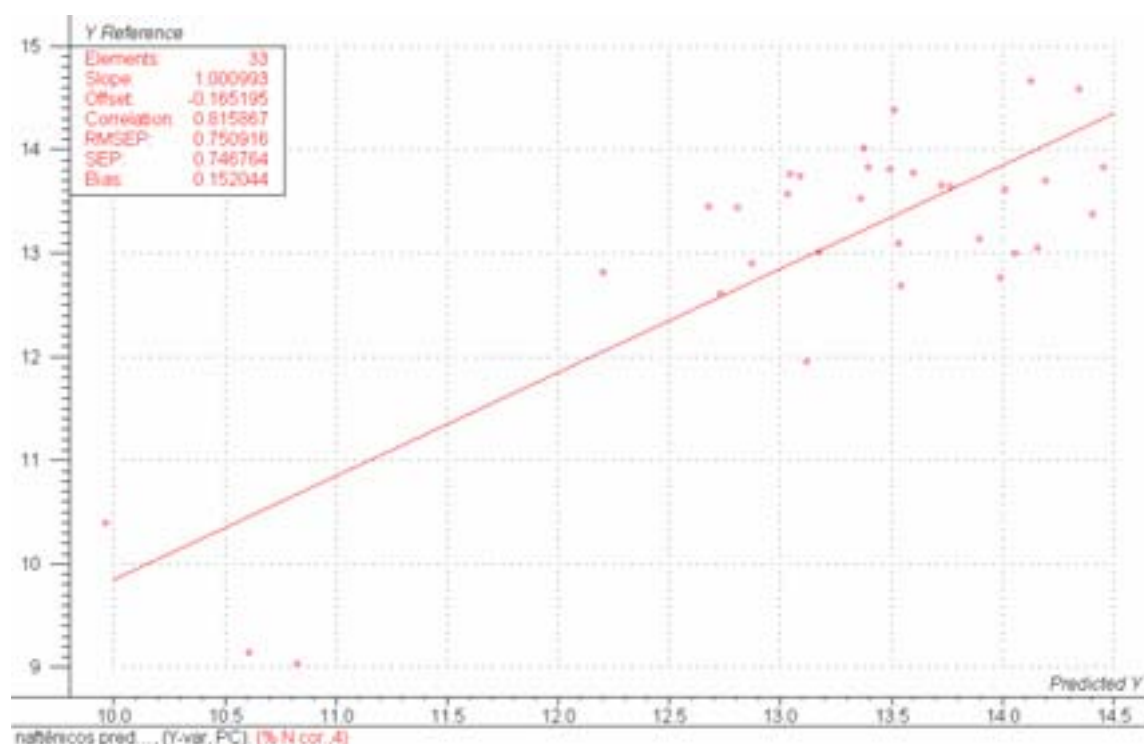


Figura 47 – Teor de parafinas nas amostras de gasolinas conjunto predição ($R^2 = 0,5975$) ;

— teórico.

RMSEP = 0,77

5.5.2. - Modelo PLS para olefinas

O modelo PLS gerado para predição do teor de olefinas em amostras de gasolinas, identificou duas regiões do espectro FT-Raman como características para a predição desta propriedade. As regiões foram entre $1630 - 1740 \text{ cm}^{-1}$, a qual foi detectada também por Cooper (23) e a região de $2820 - 3100$ detectada por Cruz (42) em seu trabalho de dissertação de Doutorado.

A propriedade teor de olefinas presentes na gasolina, apresentou um dos melhores resultados de coeficiente de correlação para os modelos PLS de calibração, validação e predição. Os dados de predição e erro relativos para o conjunto teste são mostrados na tabela 11. O erro médio relativo de predição ficou em 2,25%, ficando todas as amostras do conjunto teste com erro menor que 10%. Para o coeficiente de correlação de predição foi obtido $R^2 = 0,78$ e RMSEP de 0,51 (figura 50). O coeficiente de correlação para o conjunto calibração foi $R^2 = 0,80$ com RMSEP de 0,45 (figura 48), enquanto que o coeficiente de correlação para o conjunto validação foi de $R^2 = 0,65$ com RMSEP de 0,60

(figura 49). As amostras 7503 e 7653 não fizeram parte do conjunto teste para o modelo PLS, apesar de apresentarem valores dentro da faixa média do teor de olefinas para as amostras de gasolinas, ou seja, $17,46\% \pm 0,9$.

Um RMSEP de predição de 0,51 e os valores de predição comparados aos valores de referência na tabela 11 justificam que o modelo PLS para olefinas está bem ajustado.

Tabela 11 – Dados de referência, predição e erro relativo do conjunto teste para olefinas.

amostra	predito	referência	erro %
7277	17,01	17,96	-5,29
7286	16,65	16,72	-0,44
7294	17,21	17,10	0,63
7298	17,66	17,65	0,06
7301	17,17	17,28	-0,65
7304	17,46	17,36	0,59
7307	17,12	17,25	-0,74
7308	17,46	17,64	-1,05
7309	17,49	17,65	-0,92
7317	17,30	16,85	2,68
7365	16,66	17,12	-2,73
7369	16,51	16,81	-1,76
7381	17,42	17,31	0,64
7403	17,88	17,90	-0,13
7510	16,54	17,31	-4,48
7511	17,49	17,19	1,73
7591	16,38	16,00	2,39
7595	17,36	17,89	-2,97
7598	16,68	16,55	0,77
7601	16,58	17,37	-4,51
7602	16,96	17,47	-2,93
7604	17,59	18,15	-3,11
7607	21,57	22,17	-2,73
7624	17,57	18,23	-3,66
7638	18,06	18,00	0,33
7641	18,19	17,02	6,88
7647	17,97	18,80	-4,40
7648	18,17	19,06	-4,67
7655	18,33	18,66	-1,74
7657	16,18	16,14	0,26
7682	16,79	17,36	-3,25
7697	17,88	17,72	0,89
7714	18,52	17,77	4,27
7725	17,89	17,49	2,27

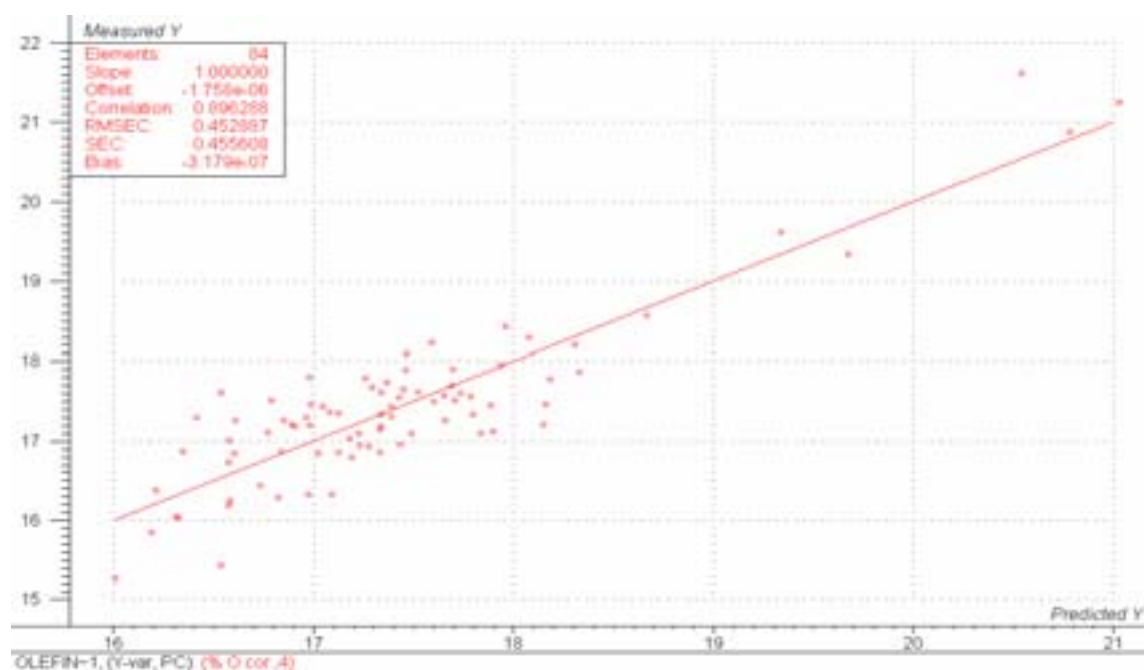


Figura 48 – Teor de olefinas nas amostras de gasolinas conjunto calibração ($R^2 = 0,80$) ;

— teórico.

RMSEP = 0,45

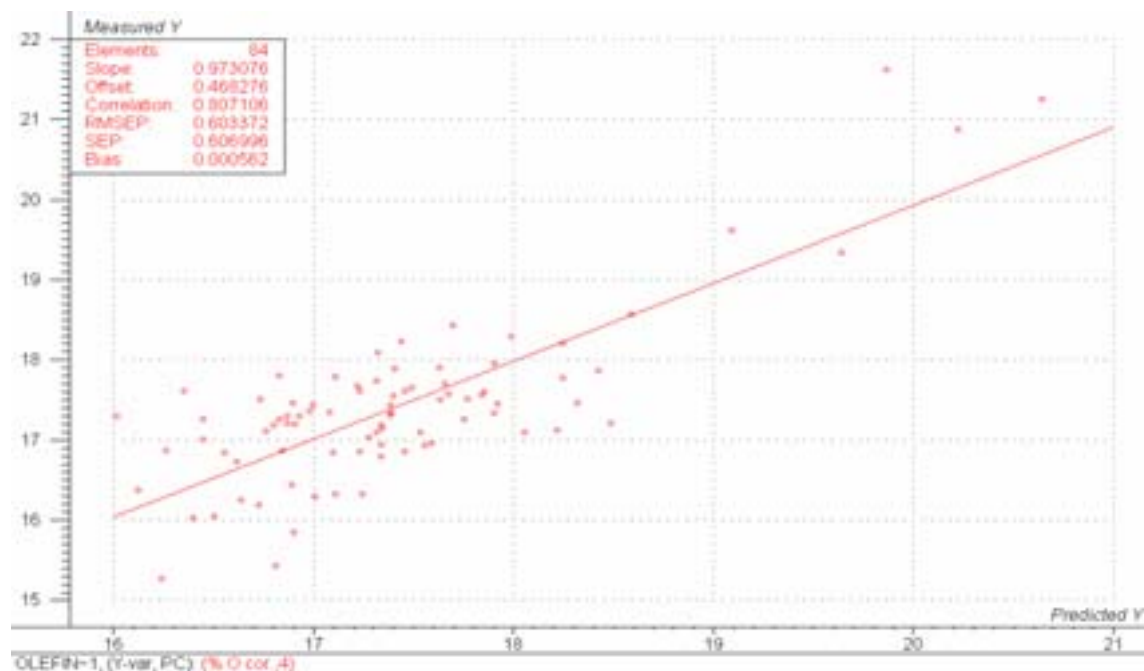


Figura 49 – Teor de olefinas nas amostras de gasolinas conjunto validação ($R^2 = 0,65$) ;

— teórico.

RMSEP = 0,60

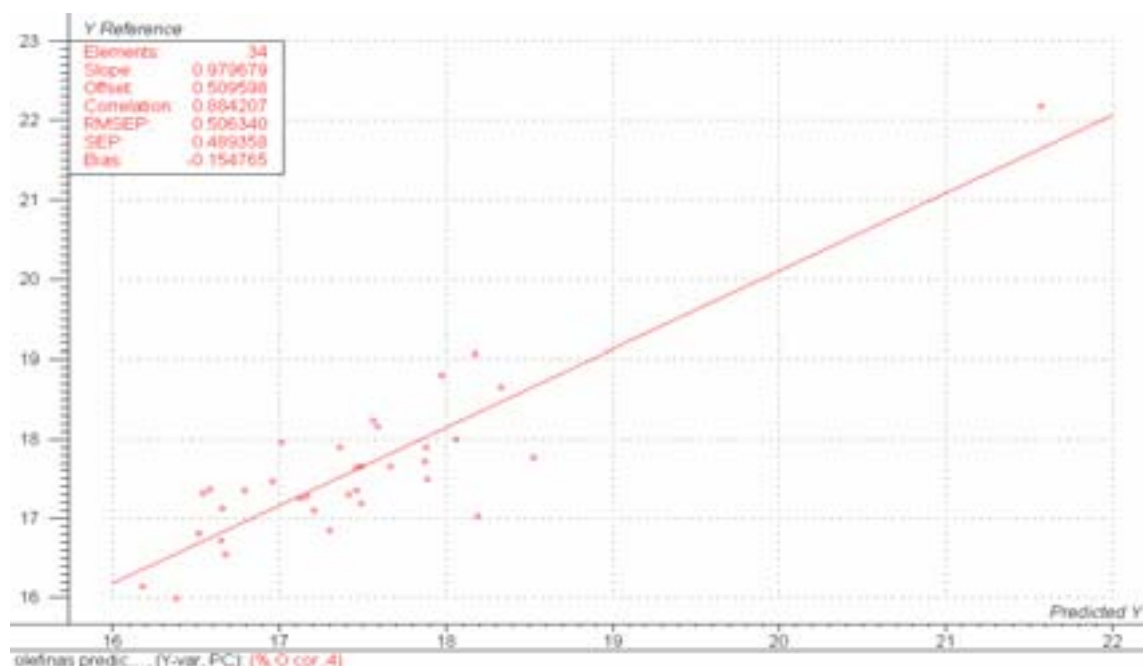


Figura 50 – Teor de olefinas nas amostras de gasolinas conjunto predição ($R^2 = 0,78$) ;

— teórico.

RMSEP = 0,51

5.5.2. - Modelo PLS para naftênicos

Para a propriedade teor de naftênicos em gasolinas, o modelo PLS desenvolvido indicou três regiões do espectro FT-Raman para a predição desta propriedade, sendo a região entre $720 - 1000 \text{ cm}^{-1}$, também detectada pelo modelo PLS desenvolvido por Cruz (42); a região entre $1420 - 1560 \text{ cm}^{-1}$ e a região entre $2820 - 2960 \text{ cm}^{-1}$, sendo esta última também detectada por Cruz (42).

O espectro FT-Raman da amostra do cicloexano comparado com o espectro FT-Raman da amostra da gasolina (figura 51) mostram uma banda forte na região de 1445 cm^{-1} do espectro tanto para a amostra de gasolina quanto para a amostra de cicloexano. A inclusão destas variáveis independentes mostraram uma melhor resposta para o modelo PLS construído para a determinação do teor de naftênicos.

O teor de naftênicos foi a propriedade que apresentou os maiores índices de erro relativo com uma média percentual de 4,77% com duas amostras apresentando um erro relativo acima de 10% como pode ser visto assinalado na tabela 12 em vermelho. Estas amostras (7714 e 7725) apesar deste comportamento, não foram consideradas amostras “outliers” conforme mostra o gráfico dos “scores” (figura 42), nos três primeiros

componentes principais dos espectros FT-Raman das amostras de gasolinas. A tabela 8 mostra para estas amostras uma variação no teor de aromáticos e naftênicos seguindo o mesmo tipo de comportamento o que sugere que estas amostras possuem composição química diferenciada das demais. Foram realizados testes para a predição sem as amostras 7714 e 7725, mas como sua ausência influenciava o modelo PLS de forma negativa, foi resolvido então mantê-las conforme pode ser observado na tabela 12.

O conjunto calibração para naftênicos apresentou um bom coeficiente de correlação ($R^2 = 0,82$) com mostra seu RMSEP de 0,62 (figura 52). Para o conjunto validação o coeficiente de correlação apresentou $R^2 = 0,55$, que está relacionado com a pouca variação do intervalo que compreende os resultados de referência e validação, apresentando este último um RMSEP de 0,98 (figura 53). O conjunto teste apresentou também um bom coeficiente de correlação ($R^2 = 0,68$) como comprova seu RMSEP de 0,74 (figura 54), bem melhor que o apresentado para validação, sendo considerado satisfatório o modelo PLS desenvolvido para naftênicos. Todas amostras apresentaram teor de naftênicos em torno da média ($13,21\% \pm 1,05$) com exceção das amostras 7714 e 7725 como mostra a tabela 12.

As amostras 7304, 7510, 7604, e 7607 não fizeram parte do conjunto predição por influenciarem de forma negativa o modelo PLS, valendo observar que a amostra 7607 para teor de naftênicos encontra-se fora da faixa média percentual encontrada para as demais amostras, como mostra a tabela 8.

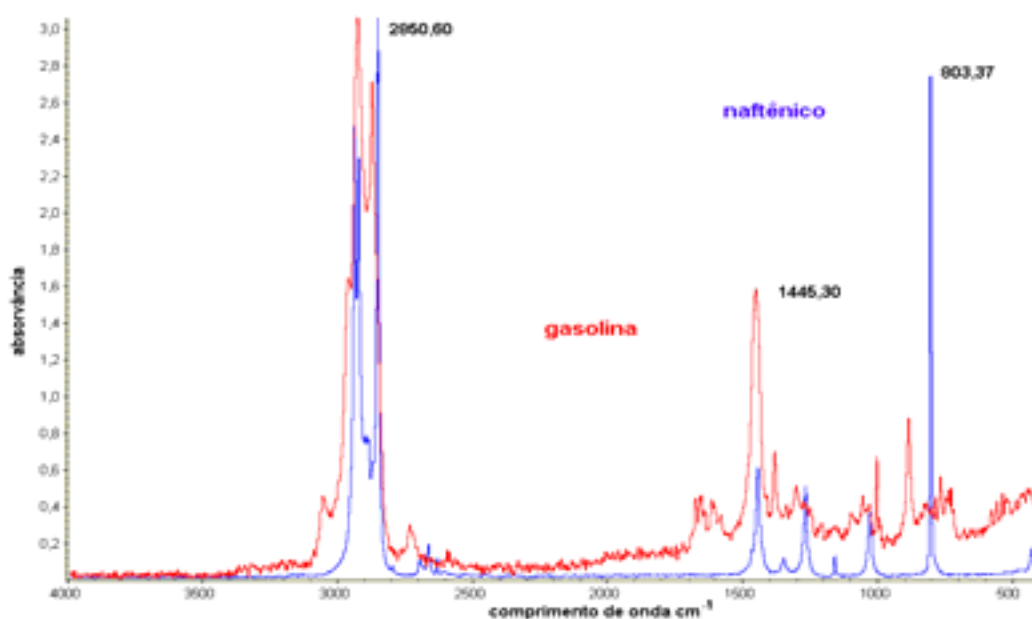


Figura 51 - Espectro FT-Raman amostra de gasolina e cicloexano.

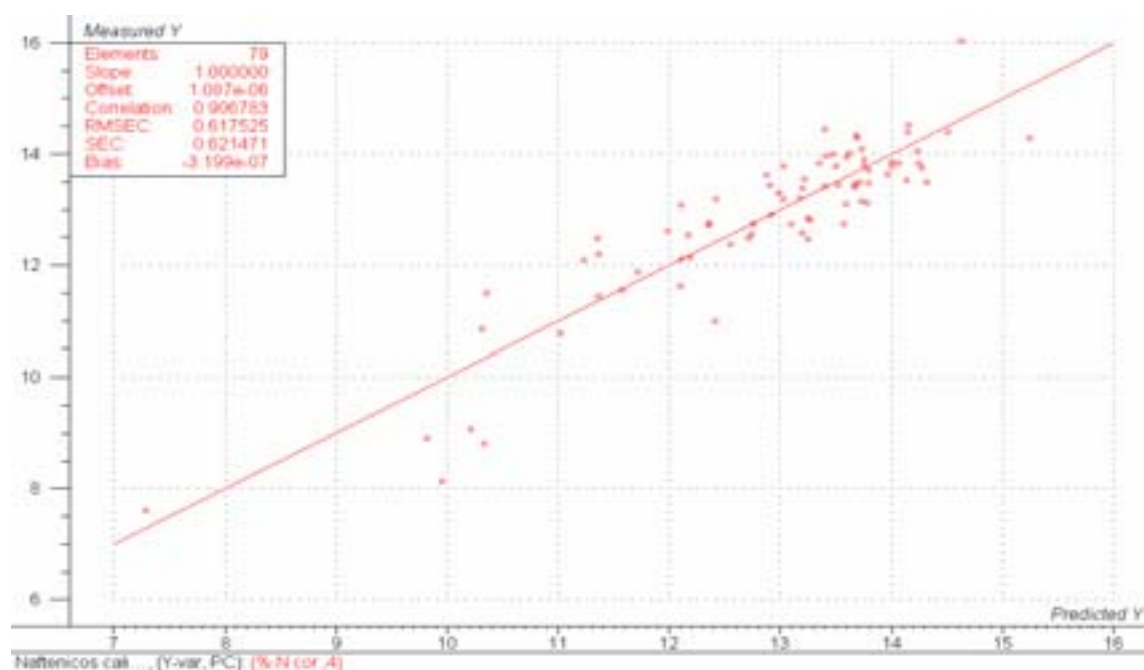


Figura 52 – Teor de naftênicos nas amostras de gasolinas conjunto calibração ($R^2 = 0,82$) ;

— teórico.

RMSEP = 0,62

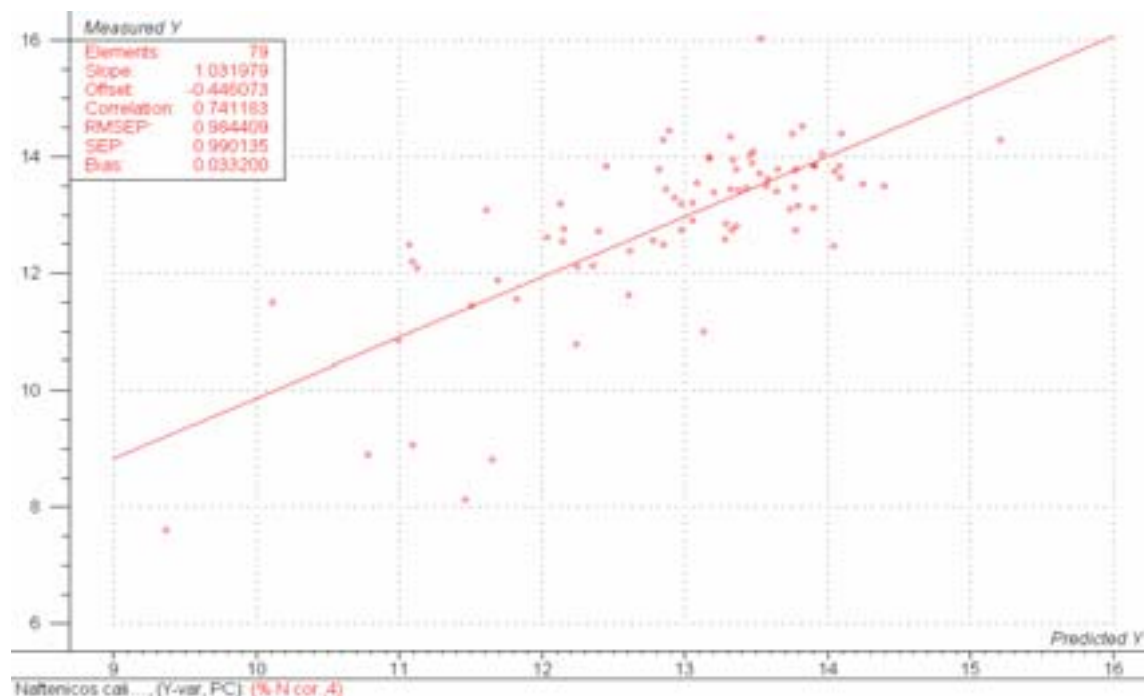


Figura 53 – Teor de naftênicos nas amostras de gasolinas conjunto validação ($R^2 = 0,55$) ;

— teórico.

RMSEP = 0,98

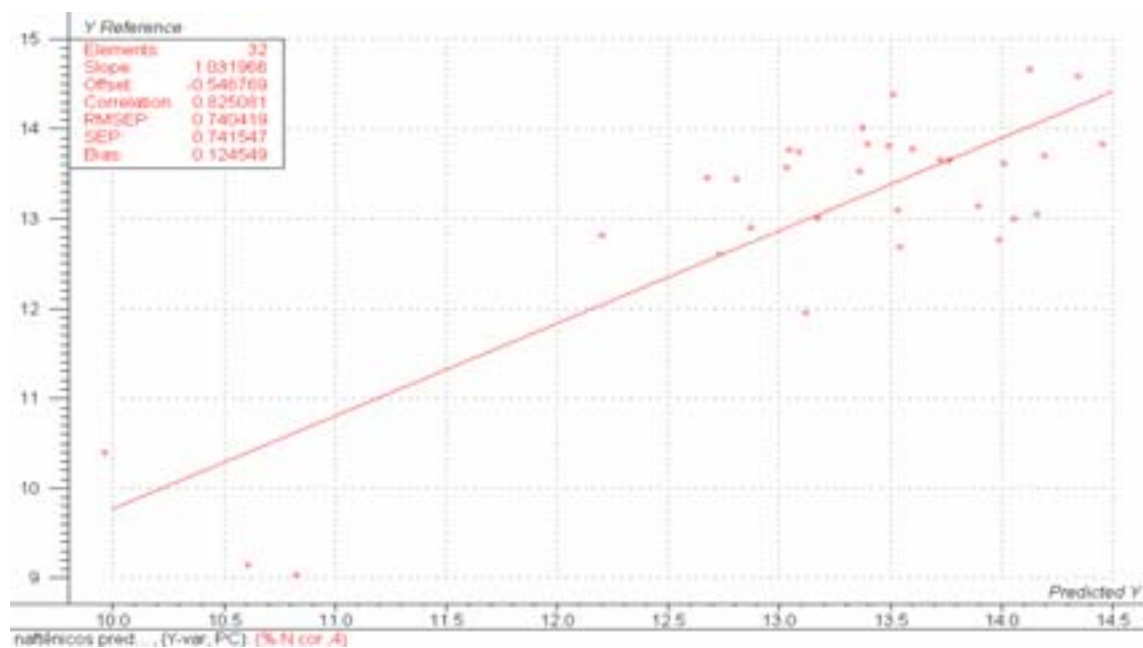


Figura 54 – Teor de naftênicos nas amostras de gasolinas conjunto predição ($R^2 = 0,68$) ;

— teórico.

RMSEP = 0,74

Tabela 12 – Dados de referência, predição e erro relativo do conjunto teste para naftênicos.

amostra	predito	referência	erro %
7277	13,37	14,02	4,80
7286	14,13	14,66	3,79
7294	13,03	13,58	4,17
7298	13,40	13,84	3,29
7301	13,36	13,53	1,27
7307	14,34	14,58	1,67
7308	13,73	13,66	-0,50
7309	14,46	13,84	-4,28
7317	12,73	12,61	-0,94
7365	13,90	13,15	-5,41
7369	13,76	13,65	-0,84
7381	13,04	13,77	5,57
7403	13,51	14,38	6,43
7503	13,53	13,10	-3,18
7511	12,68	13,45	6,12
7591	13,09	13,75	5,04
7595	14,01	13,62	-2,83
7598	12,20	12,81	5,03
7601	14,05	13,00	-7,52
7602	14,16	13,05	-7,79
7624	14,19	13,70	-3,46
7638	13,60	13,78	1,34
7641	13,54	12,69	-6,27
7647	13,18	13,01	-1,28
7648	12,87	12,91	0,25
7653	12,81	13,44	4,97
7655	9,96	10,40	4,43
7657	13,49	13,81	2,36
7682	13,12	11,95	-8,91
7697	13,99	12,76	-8,77
7714	10,83	9,04	-16,48
7725	10,61	9,15	-13,77

5.5.3. - Modelo PLS para aromáticos

Em seu trabalho, Cooper (23) demonstra que uma banda na região em torno de 1000 cm^{-1} é característica para anel aromático, o que é confirmado por Hansens (29), que em seu trabalho com gasolinas da Nova Zelândia, detectou por espectroscopia FT-Raman, que na região em torno de 3060 cm^{-1} ocorre uma banda característica de estiramento de C – H aromático. As regiões aqui citadas por Cooper (23) e Hansens (29), foram previstas pelo modelo PLS, sendo um bom indício para predição do teor de aromáticos.

O conjunto calibração para a propriedade teor de aromáticos, apresentou um coeficiente de correlação muito bom ($R^2 = 0,93$) com mostra a figura 55, comprovado pelo RMSEP de 0,37 mesmo tendo as amostras 7714 e 7725 apresentado um valor de referência elevado para aromáticos em relação às demais amostras conforme mostra a tabela 8. Todas as amostras apresentaram um erro relativo menor que 10% com exceção da amostra 7714, cujo valor encontra-se em vermelho na tabela 13. Todas as amostras se mantiveram dentro de uma faixa média de teor de aromáticos em $15,27\% \pm 0,69$, valores estes bem próximos dos de referência. A amostra 7714 foi mantida no modelo PLS desenvolvido, pois sua ausência apesar do seu comportamento diferenciado em relação as demais, produzia no modelo um efeito negativo.

O coeficiente de correlação para o conjunto validação ($R^2 = 0,58$) foi considerado satisfatório (figura 56), com um RMSEP de 0,92 que é considerado ainda um erro aceitável dentro de análise multivariada. Os resultados de predição para o conjunto teste apresentaram um desempenho melhor em relação ao conjunto calibração como pode ser observado na figura 57, o qual apresenta coeficiente de correlação com $R^2 = 0,7$ e um RMSEP de 0,79, o que para o tipo de formatação dos dados de predição, são bastante satisfatórios. As amostras 7317; 7595; 7598; 7655 e 7657 não fizeram parte do conjunto calibração pois contribuíam de forma negativa para o modelo de predição PLS, valendo observar que as amostras 7317 e 7655 não estão dentro da faixa média para teor de aromáticos.

Tabela 13 – Dados de referência, predição e erro relativo do conjunto teste para aromáticos.

Amostra	predito	referência	erro %
7277	15,12	14,52	4,17
7286	14,79	14,78	0,05
7294	15,24	15,64	-2,61
7298	15,37	15,01	2,38
7301	15,33	14,97	2,43
7304	14,51	15,25	-4,84
7307	15,13	14,47	4,57
7308	15,56	15,39	1,08
7309	15,41	15,01	2,67
7365	14,63	15,36	-4,73
7369	14,10	14,04	0,43
7381	15,68	15,25	2,82
7403	15,27	15,45	-1,22
7503	15,30	16,08	-4,89
7510	14,57	15,83	-7,93
7511	16,07	16,76	-4,17
7591	14,46	14,22	1,70
7601	15,01	14,85	1,08
7602	15,98	15,62	2,28
7604	14,76	13,89	6,23
7607	15,79	16,38	-3,55
7624	14,44	13,89	3,93
7638	14,96	14,24	5,05
7641	15,58	16,46	-5,35
7647	15,00	14,41	4,08
7648	14,96	14,60	2,46
7653	15,51	16,30	-4,87
7682	15,53	16,66	-6,77
7697	15,14	15,43	-1,87
7714	16,63	19,34	-14,02
7725	17,62	18,86	-6,58

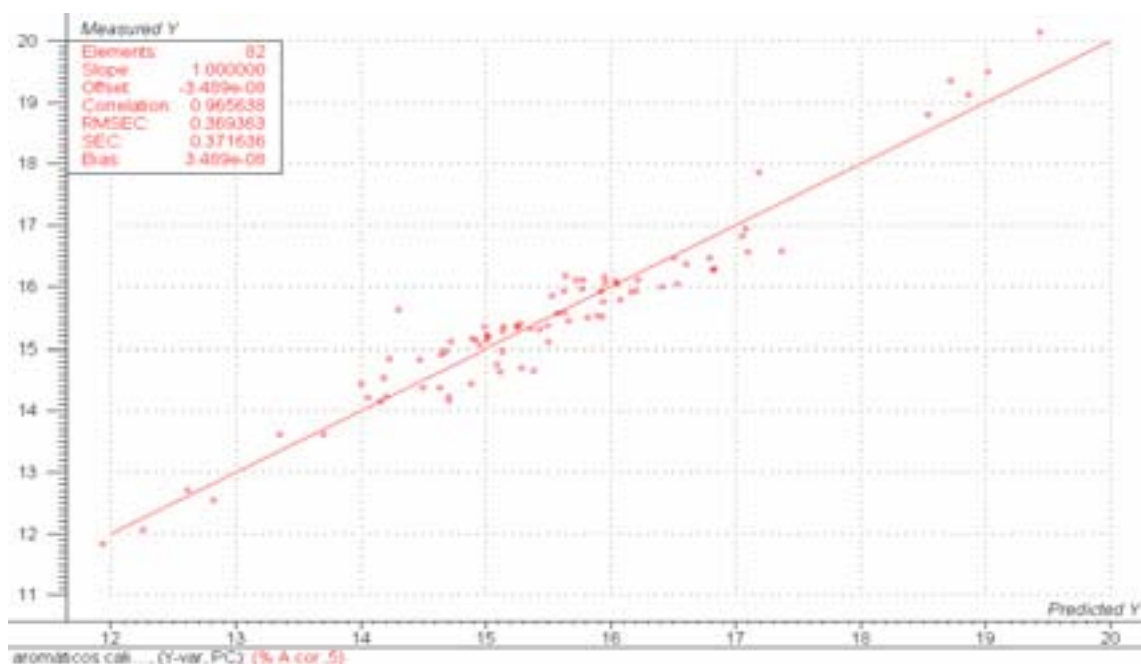


Figura 55 – Teor de aromáticos nas amostras de gasolinas conjunto calibração ($R^2 = 0,93$) ;

— teórico.

RMSEP = 0,37

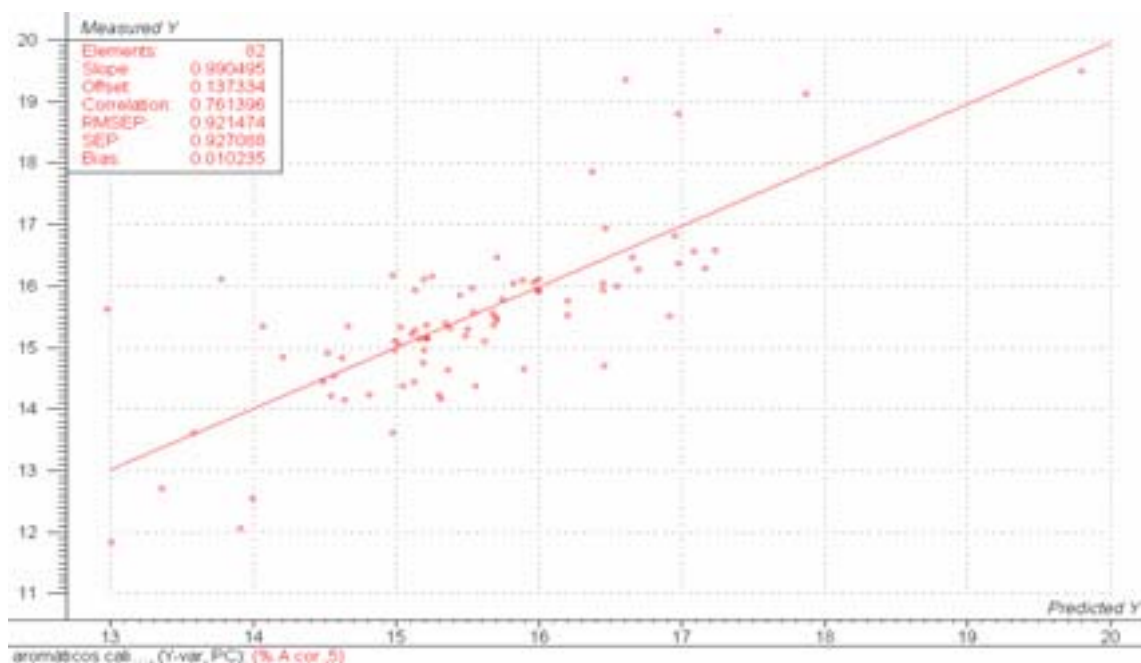


Figura 56 – Teor de aromáticos nas amostras de gasolinas conjunto validação ($R^2 = 0,58$) ;

— teórico.

RMSEP = 0,92

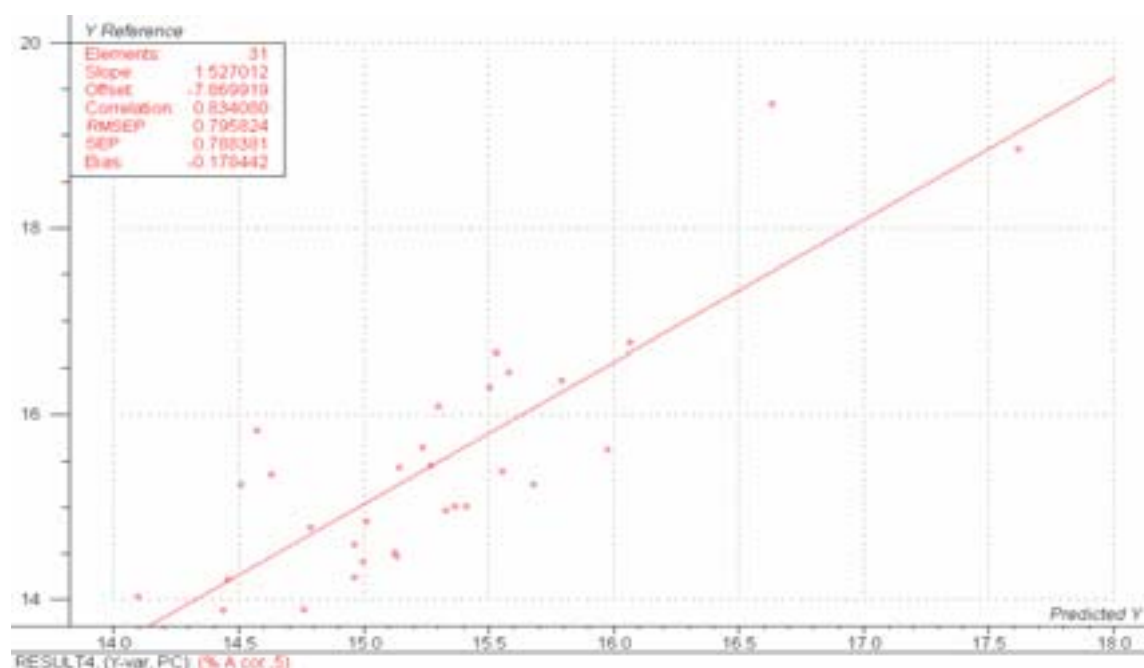


Figura 57 – Teor de aromáticos nas amostras de gasolinas conjunto predição ($R^2 = 0,7$) ;

— teórico.

RMSEP = 0,79

5.5.5. - Modelo PLS para etanol

A região do espectro prevista pelo modelo PLS desenvolvido para a predição do teor de etanol está situada entre 200 – 1000 cm^{-1} no espectro FT-Raman, região esta, que já havia sido detectada pelo modelo PLS gerado por Cruz (42), que ainda identificou outras duas regiões do espectro FT-Raman entre 830 – 1529 cm^{-1} e 2769 – 3085 cm^{-1} . Estas duas últimas regiões foram testadas pelo modelo PLS aqui gerado, para predição do teor de etanol em amostras de gasolinas, mas como os resultados não foram satisfatórios, estas regiões do espectro foram abandonadas.

O modelo PLS para etanol apresentou um excelente coeficiente de correlação para o conjunto calibração ($R^2 = 0,9999$), com um RMSEP de 0,004, conforme mostra a figura 58. Já os conjuntos validação e predição (figuras 59 e 60 respectivamente) apresentaram valores para o coeficiente de correlação baixos, $R^2 = 0,40$ e $R^2 = 0,26$ respectivamente. Estes valores baixos estão relacionados ao fato da faixa de variação para os teores de etanol ser muito estreita o que não contribui para o coeficiente de correlação. Vale observar que os valores de RMSEP de 0,38 para o conjunto validação e

RMSEP de 0,39 para o conjunto teste (predição) são baixos, demonstrando que o modelo PLS com onze PCs descreve bem a propriedade.

Observando-se a tabela 14, pode-se comprovar que os teores de etanol nas amostras de gasolinas variam dentro de uma faixa muito estreita e o erro relativo médio é de 1,35% entre o valor de referência e o valor predito, sendo esta variação muito pequena em relação a composição percentual do etanol nas amostras de gasolinas. Aliando-se estas observações ao fato do conjunto predição ter apresentado RMSEP de 0,39 conclui-se que o modelo PLS para predição dos teores de etanol foi bem desenvolvido apresentando um bom resultado de predição.

Tabela 14 – Dados de referência, predição e erro relativo do conjunto teste para etanol.

amostra	referência	predito	erro %
7277	25,15	25,24	0,36
7286	25,47	25,53	0,23
7294	25,7	25,8	0,35
7298	25,35	25,81	1,81
7301	25,88	25,68	-0,78
7304	25,39	25,76	1,44
7307	26,53	25,91	-2,36
7308	25,81	25,29	-2,01
7309	25,35	25,61	1,01
7365	25,47	25,54	0,28
7369	25,5	25,17	-1,26
7381	25,69	25,67	-0,09
7403	25,32	25,74	1,66
7503	25,34	25,29	-0,2
7510	24,78	25,28	2,02
7511	25,17	24,96	-0,85
7591	25,77	25,44	-1,28
7595	25,34	25,49	0,58
7601	25,18	25,55	1,48
7602	24,85	25,47	2,46
7604	25,91	25,29	-2,37
7607	26,35	25,84	-1,93
7624	25,03	25,68	2,57
7638	25,42	25,31	-0,41
7641	24,95	25,29	1,36
7647	25,06	25,6	2,16
7648	24,82	25,25	1,75
7653	24,77	25,39	2,52
7682	24,77	25,39	2,49
7697	25,33	25,52	0,75
7714	24,79	25,1	1,23
7725	24,9	24,63	-1,1

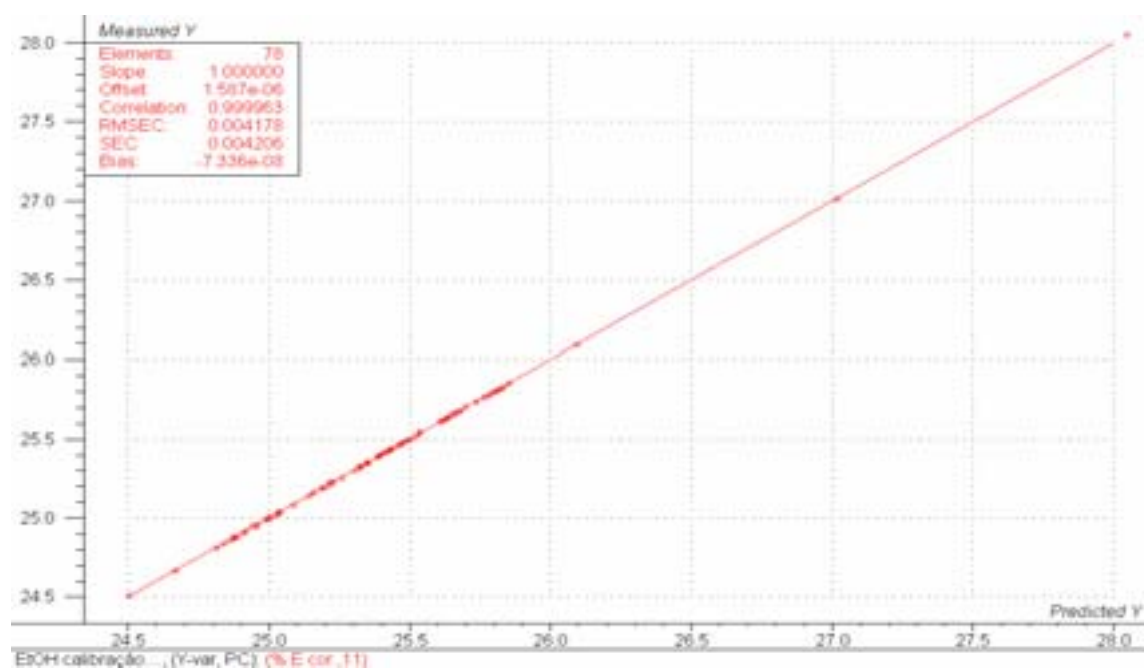


Figura 58 – Teor de etanol nas amostras de gasolinas conjunto calibração ($R^2 = 0,9999$) ;

— teórico.

RMSEP = 0,004

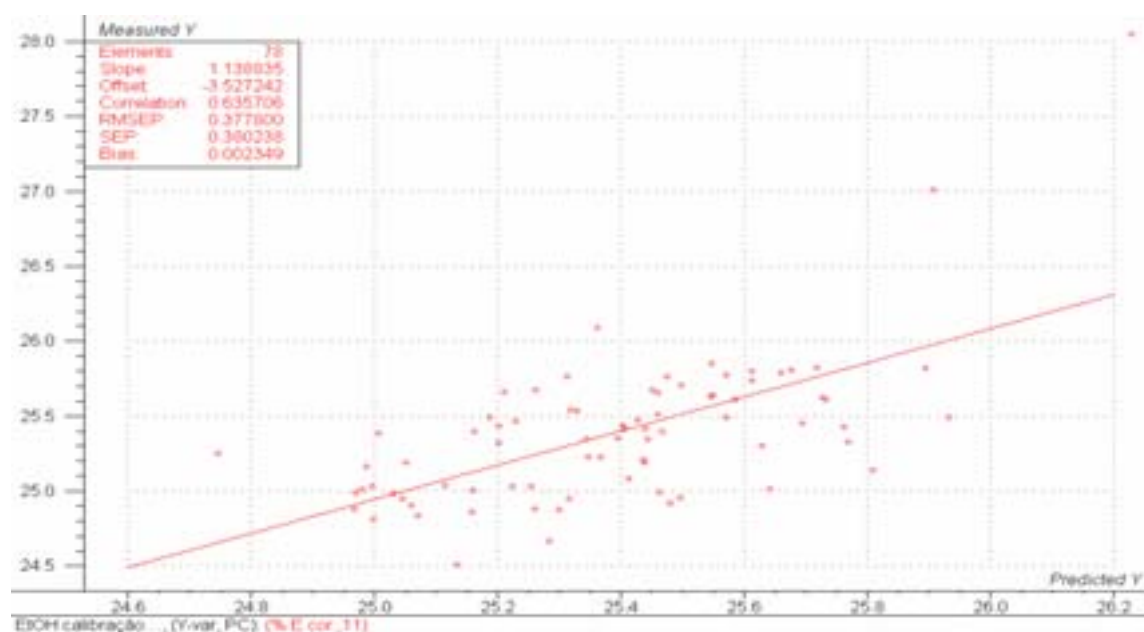


Figura 59 – Teor de etanol nas amostras de gasolinas conjunto validação ($R^2 = 0,40$) ;

— teórico.

RMSEP = 0,38

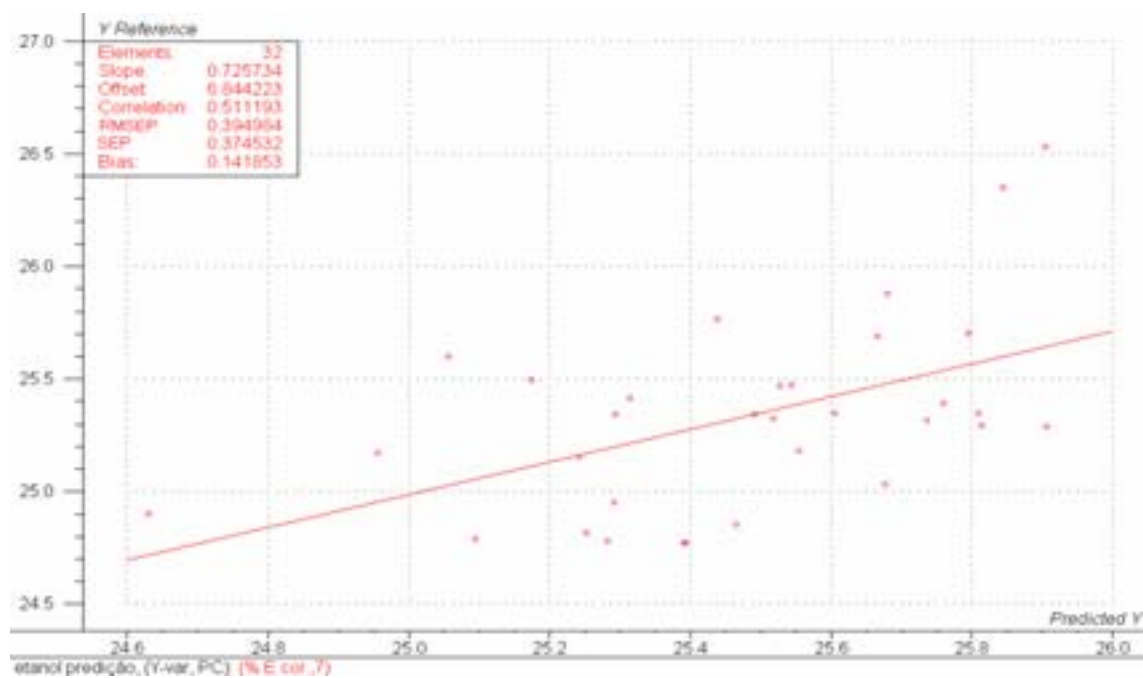


Figura 60 – Teor de etanol nas amostras de gasolinas conjunto predição ($R^2 = 0,26$) ;

— teórico.

RMSEP = 0,39