

6 Metodologia Proposta para Seleção de Modelos

Neste capítulo será apresentada uma metodologia de seleção de modelos em EMP. Esta metodologia possui duas etapas e está descrita em Dal Bello & Vieira (a ser publicado). Cabe registrar que tal metodologia também pode ser utilizada para seleção de modelos em EM sem a inclusão de variáveis de processo.

Na primeira etapa da metodologia, ajusta-se um Modelo Base utilizando o método tradicional passo-a-passo (*stepwise, forward* ou *backward*). Nesta seleção, devem ser considerados todos os termos da combinação multiplicativa dos termos do modelo de Scheffé com o modelo para as variáveis de processo.

Na etapa seguinte, pode-se obter um modelo melhor, levando em consideração, além dos termos do modelo base, todos os termos equivalentes aos termos do modelo base. Nesta etapa, utiliza-se o critério de indiferença de AIC_c para a seleção dos termos do modelo, já que a utilização deste critério pode levar à escolha de um modelo com menor $PRESS$, MSE ou variância de uma futura resposta no ponto do que o obtido pelo critério AIC_c mínimo.

6.1. Primeira Etapa

Utilizando os dados do EMP apresentado no Capítulo 4, foi ajustado um Modelo Base com a utilização do método passo-a-passo. Para o caso em questão, foi considerada agora a combinação multiplicativa do modelo cúbico de Scheffé para mistura com o modelo quadrático reduzido para as variáveis de processo. A forma do modelo completo foi apresentada na Equação (4.7).

Partindo do modelo completo, ajustou-se o seguinte modelo com o método passo-a-passo (*backward*):

$$\begin{aligned} \hat{y} = & 5,99v_1 + 12,51v_2 - 38,48v_3 - 6,47v_1v_2 + 87,34v_1v_3 + \\ & + 0,94v_1z_2 + 78,55v_2v_3 + 0,57v_2z_2 + 0,70v_2z_1z_2 - \\ & - 58,94v_1v_3(v_1 - v_3) - 3,51v_1v_2z_1z_2 - 3,50v_2v_3z_1z_2 \end{aligned} \quad (6.1)$$

O Modelo (6.1) é denominado de Modelo Base. Este modelo é igual ao Modelo (4.19). Na Tabela 9 foram apresentados os resultados do teste t para o Modelo Base.

No caso em questão, o método *backward* levou à obtenção de um modelo com menor *PRESS* e *MSE* do que os dos modelos obtidos com os outros métodos passo-a-passo (*stepwise* e *forward*).

Com os gráficos de diagnóstico (Figuras 22 a 25), foi possível verificar as suposições de normalidade, independência, aditividade do modelo inerente ao modelo linear e variância constante, sendo, portanto, verificada a adequação do Modelo Base.

6.2. Segunda Etapa

Partindo do Modelo Base, pode-se obter um modelo melhor levando em consideração, além dos termos do Modelo Base, todos os termos equivalentes aos termos do Modelo Base. O termos equivalentes são obtidos com a restrição de EM apresentada na Equação (2.1). Por exemplo, o termo v_1v_2 é equivalente ao termo $v_1(1 - v_1 - v_3)$ ou ao termo $(1 - v_2 - v_3)v_2$. Na Tabela 40 são apresentados os termos equivalentes aos termos do Modelo Base.

Tabela 40 - Termos Equivalentes

Termos do Modelo Base	Termos Equivalentes
v_1v_2	$(v_1 - v_1^2 - v_1v_3); (v_2 - v_2^2 - v_2v_3)$
v_1v_3	$(v_1 - v_1^2 - v_1v_2); (v_3 - v_3^2 - v_2v_3)$
v_2v_3	$(v_2 - v_2^2 - v_1v_2); (v_3 - v_3^2 - v_1v_3)$
v_1z_2	$(z_2 - v_2z_2 - v_3z_2)$
v_2z_2	$(z_2 - v_1z_2 - v_3z_2)$
$v_2z_1z_2$	$(z_1z_2 - v_1z_1z_2 - v_3z_1z_2)$
$v_1v_3(v_1 - v_3)$	$(v_1^2v_3 - v_1v_3^2); (v_1^2 - v_1^3 - v_1^2v_2 - v_3^2 + v_3^3 + v_2v_3^2)$
$v_1v_2z_1z_2$	$(v_1z_1z_2 - v_1^2z_1z_2 - v_1v_3z_1z_2); (v_2z_1z_2 - v_2^2z_1z_2 - v_2v_3z_1z_2)$
$v_2v_3z_1z_2$	$(v_2z_1z_2 - v_2^2z_1z_2 - v_1v_2z_1z_2); (v_3z_1z_2 - v_3^2z_1z_2 - v_1v_3z_1z_2)$

Uma vez conhecidos todos os termos candidatos (termos do Modelo Base e termos equivalentes) ao modelo de EMP, utilizam-se critérios de seleção de modelos baseados na Teoria da Informação.

Para o caso em estudo foi elaborada uma rotina Matlab® para o cálculo e armazenamento dos valores de AIC_c e seleção dos modelos indiferentes, ou seja, aqueles que apresentaram um Δ_i entre 0 e 2. Apenas dois modelos são indiferentes pelo critério AIC_c . Segue o modelo indiferente com 9 parâmetros:

$$\begin{aligned} \hat{y} = & 6,11v_1 + 5,22v_2 + 32,11v_3 + \\ & + 0,59z_2 + 7,35v_2^2 - 2,58v_2z_1z_2 - \\ & - 110,24v_3^3 - 53,48v_1v_3(v_1 - v_3) + 3,31v_2^2z_1z_2 \end{aligned} \quad (6.2)$$

Na Tabela 41 são apresentados os resultados do teste t para o Modelo (6.2).

Tabela 41 - Teste do Modelo (6.2)

	estimativa dos coeficientes	desvio-padrão	t -valor	p -valor
v_1	6,11152	0,256418	23,834	0,0000
v_2	5,21749	1,10369	4,727	0,0001
v_3	32,1055	1,74901	18,356	0,0000
z_2	0,594357	0,117559	5,056	0,0000
v_2^2	7,35180	1,21957	6,028	0,0000
$v_2z_1z_2$	-2,58407	0,820489	-3,149	0,0037
v_3^3	-110,242	7,13286	-15,456	0,0000
$v_1v_3(v_1 - v_3)$	-53,4797	5,72710	-9,338	0,0000
$v_2^2z_1z_2$	3,30533	0,979931	3,373	0,0021

Analisando a coluna p -valor da Tabela 41, pode-se concluir que todos os termos são significativos.

As Figuras 80 a 83 apresentam os gráficos de diagnóstico utilizados na verificação da adequação do modelo ajustado.

No gráfico de probabilidade normal dos resíduos *studentized* mostrado na Figura 80, pode-se observar que não há indicação de que a suposição de normalidade não deva ser aceita, já que não há pontos muito fora do alinhamento.

Para verificar a suposição de independência, há o gráfico de resíduos *studentized* das observações na ordem em que foram realizados os experimentos (ver Figura 81). Como os resíduos do gráfico mostrado na Figura 81 estão aleatoriamente distribuídos e sem qualquer tendência evidente que os correlacione, não há razão para se suspeitar que a suposição de independência não seja válida.

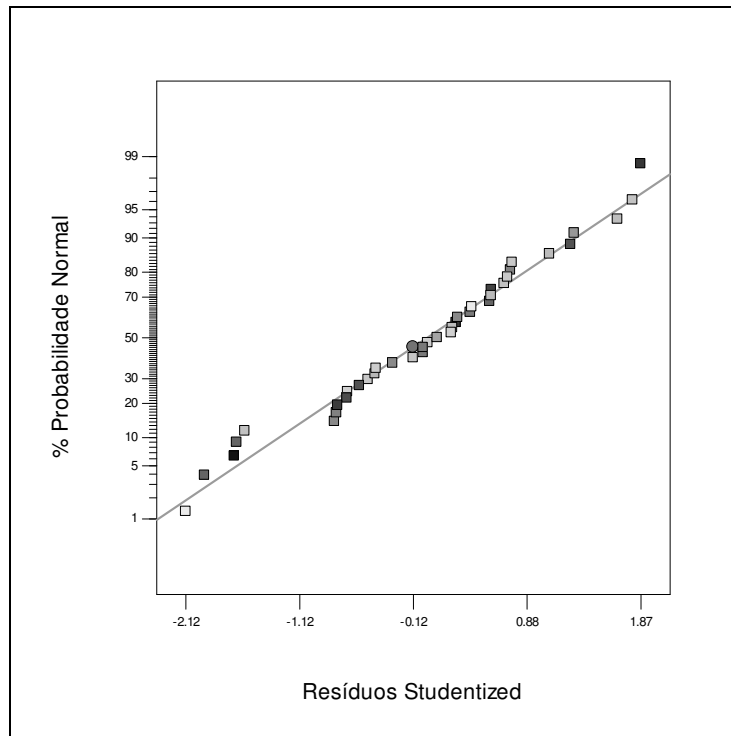


Figura 80 - Gráfico de probabilidade normal dos resíduos *studentized*

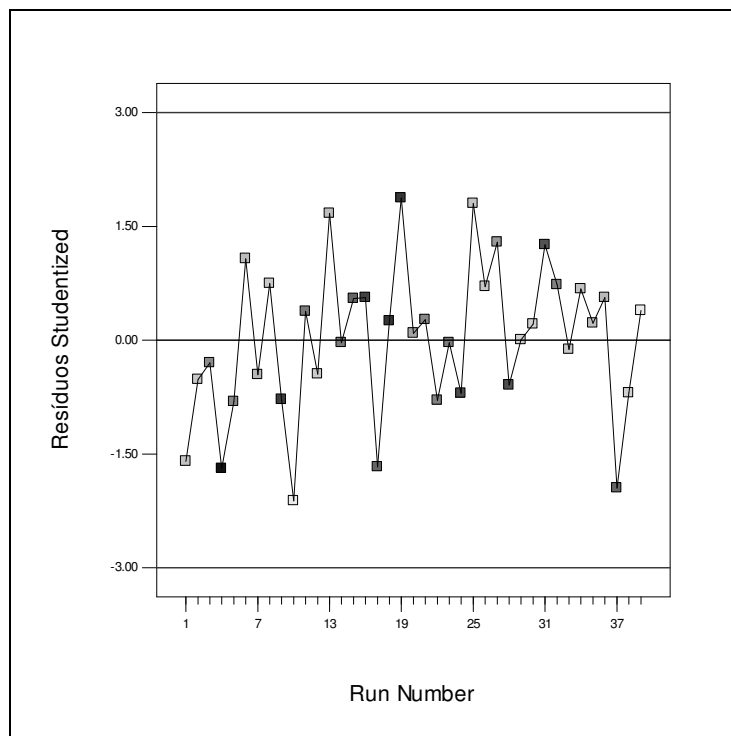


Figura 81 - Gráfico de resíduos *studentized* versus *run number*

Para verificar a aditividade do modelo, inerente ao modelo linear, há o gráfico dos resíduos *studentized* versus valores ajustados, mostrado na Figura 82.

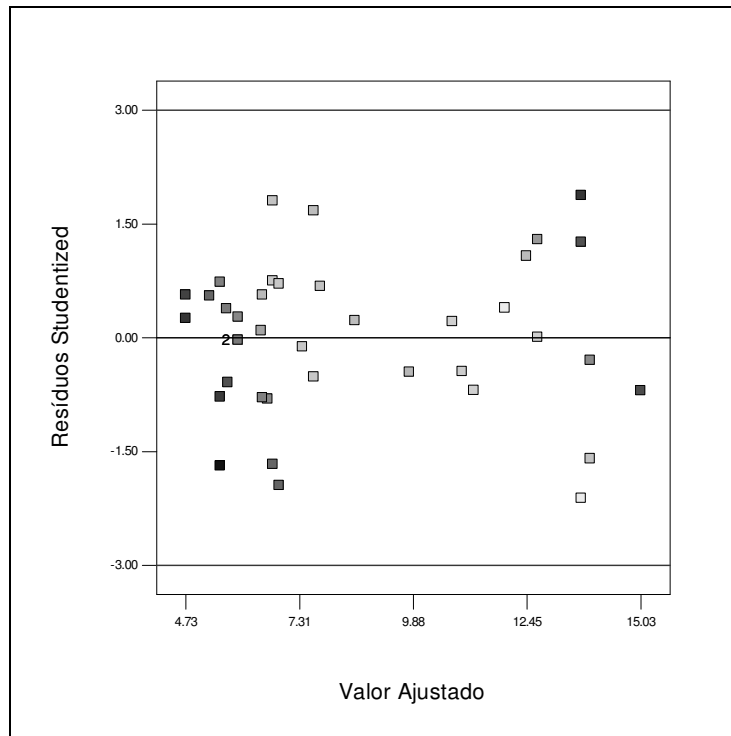


Figura 82 - Gráfico de resíduos *studentized* versus valor ajustado

Como os resíduos mostrados no gráfico da Figura 82 seguem distribuídos aleatoriamente em torno de zero, não há razão para se suspeitar que a suposição de aditividade não deva ser aceita.

Para verificar a suposição de variância constante, utiliza-se o gráfico de valores absolutos dos resíduos *studentized* versus valores ajustados, o qual é mostrado na Figura 83.

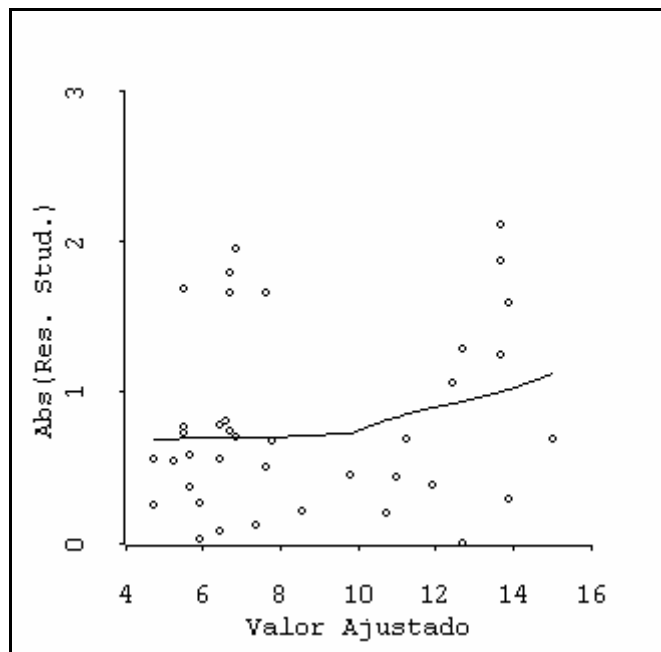


Figura 83 - Gráfico para verificação de não constância da variância

O gráfico mostrado na Figura 83 não apresenta a indicação de crescimento da variância com o aumento do valor ajustado. Ademais, utilizando a Equação (3.2), tem-se que $ET = 1,46$. Na distribuição qui-quadrado com um grau de liberdade, $ET = 1,46$ corresponde a um p -valor de 0,227. Então, não se rejeita a hipótese de que a variância não aumente quando a média aumenta.

Com isso, foi verificada a adequação do Modelo 6.2.

Segue agora o modelo indiferente com 10 parâmetros:

$$\begin{aligned} \hat{y} = & 6,11v_1 + 5,30v_2 + 31,83v_3 + 0,44z_2 + \\ & + 0,39v_1z_2 + 7,27v_2^2 - 2,62v_2z_1z_2 - \\ & - 109,07v_3^3 - 53,20v_1v_3(v_1 - v_3) + 3,31v_2^2z_1z_2 \end{aligned} \tag{6.3}$$

Os gráficos de diagnóstico (Figuras 84 a 87) permitiram a verificação da adequação do Modelo (6.3). Ademais, utilizando a Equação (3.2), $ET = 1,28$, que na distribuição qui-quadrado com um grau de liberdade corresponde a um p -valor de 0,258. Então, não se rejeita a hipótese de que a variância não aumente quando a média aumenta.

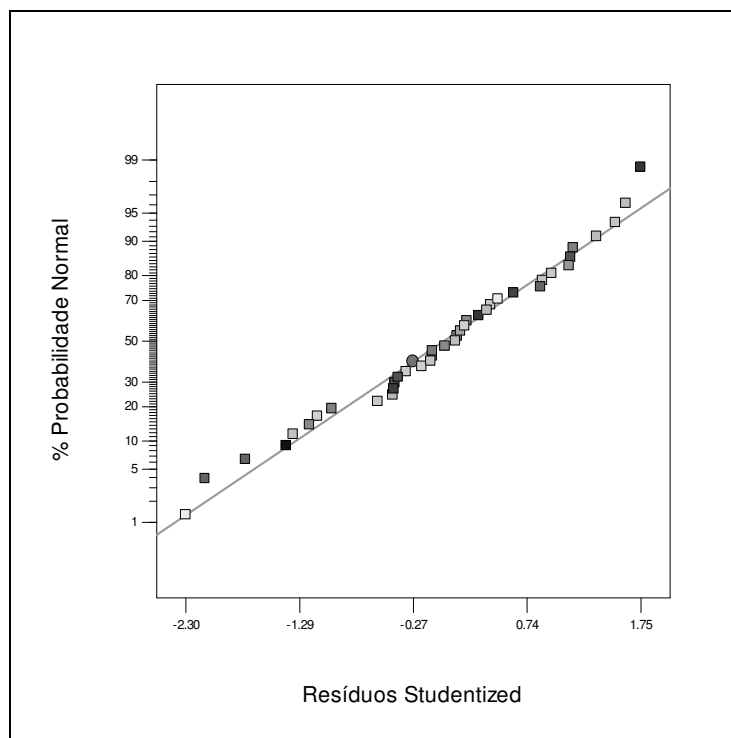


Figura 84 - Gráfico de probabilidade normal dos resíduos *studentized*

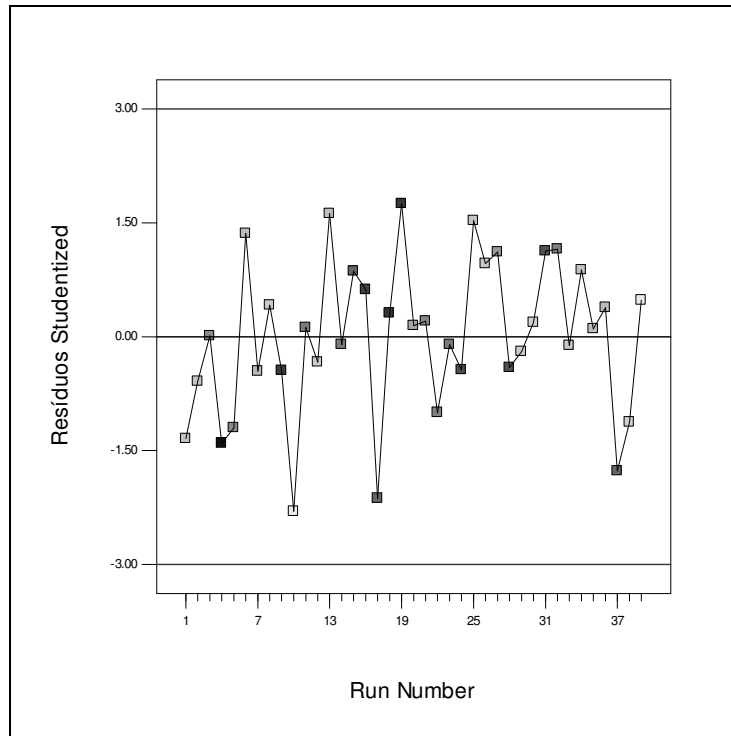


Figura 85 - Gráfico de resíduos *studentized* versus *run number*

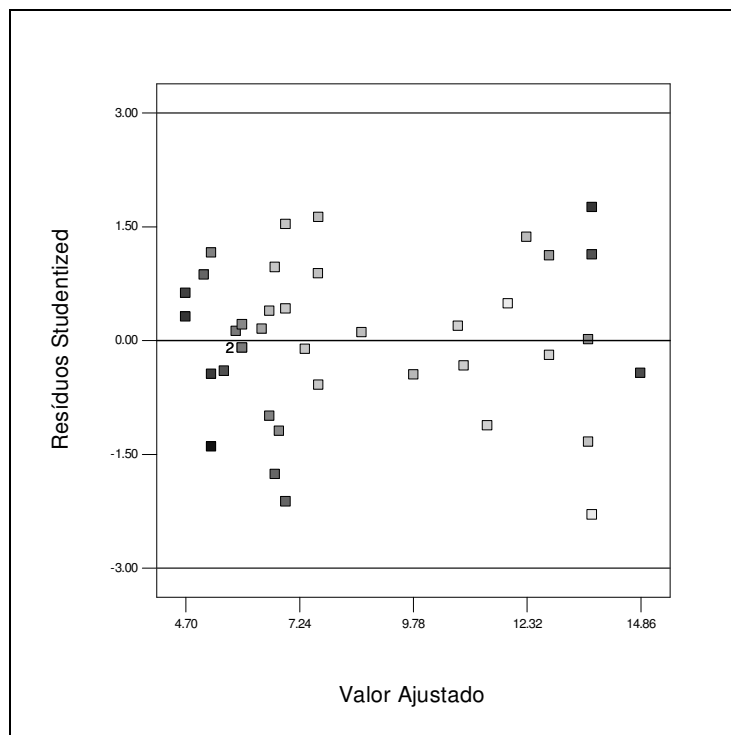


Figura 86 - Gráfico de resíduos *studentized* versus *valor ajustado*

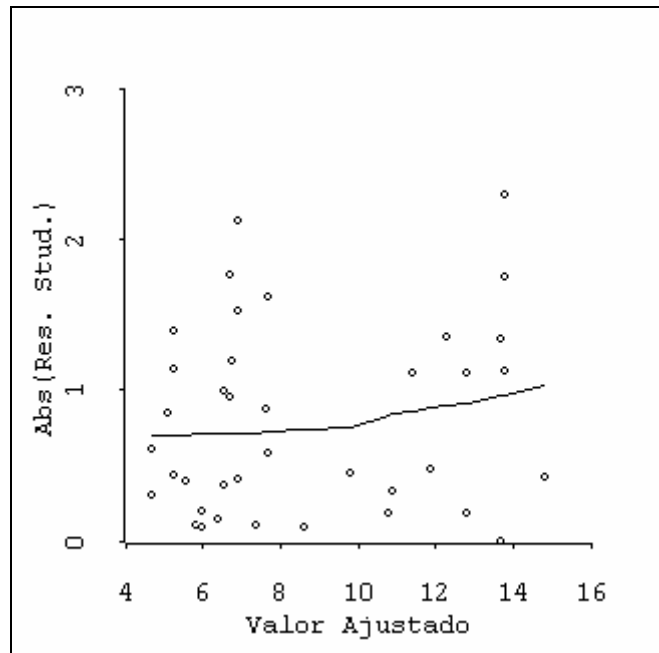


Figura 87 - Gráfico para verificação de não constância da variância

Na Tabela 42 é apresentado um comparativo dos modelos (6.1), (6.2) e (6.3), contendo as estatísticas *PRESS* e *MSE*.

Tabela 42 - Comparativo dos Modelos

Modelo	Equação	<i>PRESS</i>	<i>MSE</i>
Base	(6.1)	33,57	0,5905
Indiferente 1 (9 parâmetros)	(6.2)	25,16	0,5172
Indiferente 2 (10 parâmetros)	(6.3)	27,31	0,5088

Observando a Tabela 42, pode-se concluir que o modelo indiferente com 9 parâmetros, apresentado na Equação (6.2), é o que proporciona o menor *PRESS* e o modelo indiferente com 10 parâmetros, apresentado na Equação (6.3), é o que proporciona o menor *MSE*. Conclui-se, portanto, que a implementação da segunda etapa da metodologia resultou na obtenção de modelos melhores que o Modelo Base e também melhores do que os modelos obtidos anteriormente.

6.3. Otimização da Resposta

Nesta seção é apresentada a otimização da resposta levando em consideração os Modelos (6.2) e (6.3). Nas Figuras 41 e 42 são apresentados os gráficos de contorno da previsão da resposta e do desvio-padrão da média de

futuras respostas em função dos L-pseudocomponentes e do nível das variáveis de processo ($z_1 = 1$ e $z_2 = 1$), considerando o Modelo (5.16).

Analisando os gráficos das Figuras 88 e 89, pode-se observar que, ao ser superposta a curva de contorno correspondente ao tempo de 8 segundos com as curvas de contorno de desvio-padrão, obtém-se um desvio-padrão mínimo próximo a 0,25. Entretanto, deve-se realizar um procedimento de otimização mais rigoroso.

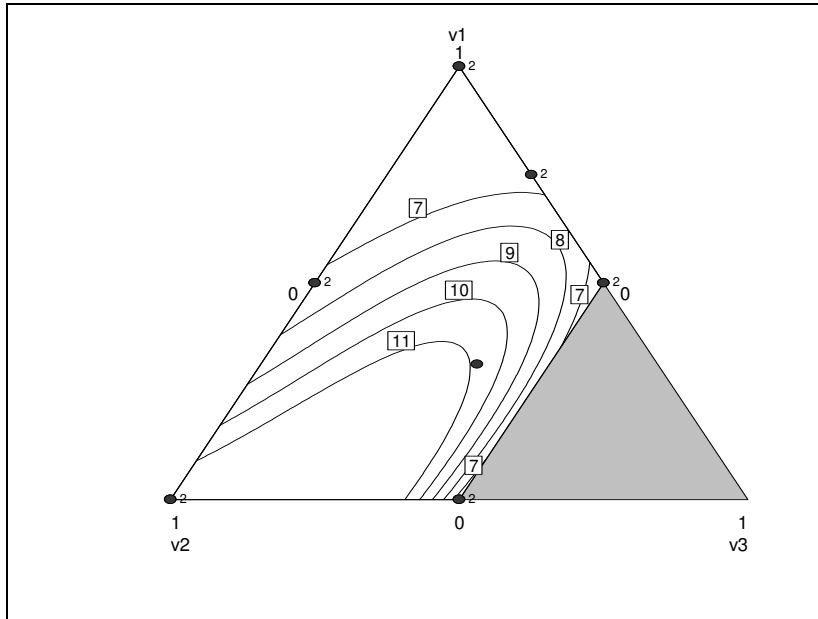


Figura 88 - Gráfico de contorno da previsão da resposta para o Modelo (6.2)

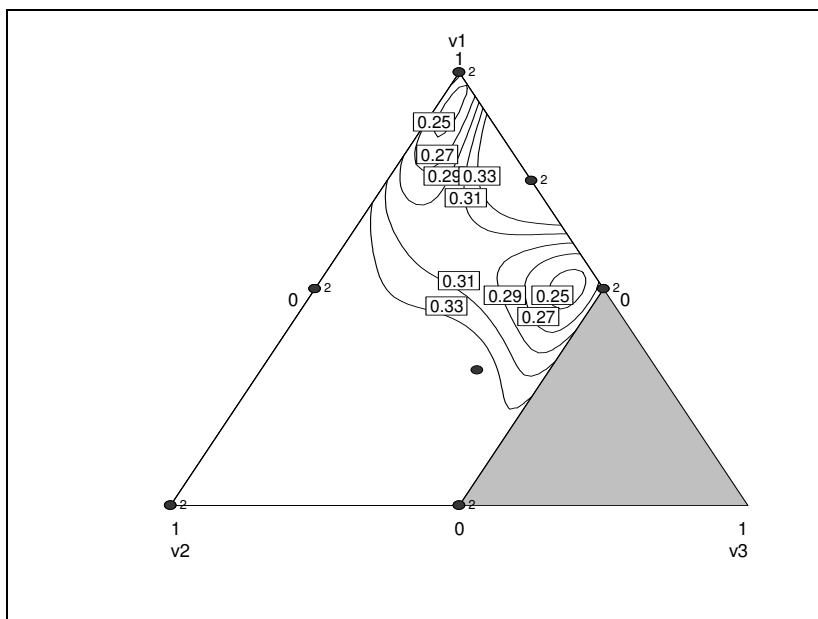


Figura 89 - Gráfico de contorno do desvio-padrão para o Modelo (6.2)

Ao minimizar a variância de uma futura resposta, foram encontradas as seguintes soluções:

Modelo (6.2):

$$v_1 = 0,4995; v_2 = 0,0652; v_3 = 0,4353; z_1 = 1; z_2 = 1;$$

$$\text{var}[\hat{y}(\mathbf{w}_0)] = 0,5757; E[\hat{y}(\mathbf{w}_0)] = 8,0000.$$

Modelo (6.3):

$$v_1 = 0,4983; v_2 = 0,0651; v_3 = 0,4366; z_1 = 1; z_2 = 1;$$

$$\text{var}[\hat{y}(\mathbf{w}_0)] = 0,5665; E[\hat{y}(\mathbf{w}_0)] = 8,0000.$$

Utilizando a Equação (2.14), tem-se a seguinte solução final em termos de componentes reais:

Modelo (6.2):

$$x_1 = 0,7900; x_2 = 0,1426; x_3 = 0,0674; z_1 = 1; z_2 = 1.$$

Modelo (6.3):

$$x_1 = 0,7899; x_2 = 0,1426; x_3 = 0,0675; z_1 = 1; z_2 = 1.$$

Analisando as soluções, pode-se concluir que os dois modelos fornecem praticamente a mesma formulação de mistura no ponto ótimo. No entanto, o Modelo (6.3) proporciona uma variância de uma nova resposta no ponto um pouco menor do que a obtida com o Modelo (6.2).

Cabe registrar que a modificação de processo proposta com relação à granulometria da mistura e a introdução do orifício no mecanismo de retardo resultam na redução da variância de uma nova resposta no ponto de interesse.

6.4. Oportunidade de Melhoria

Uma oportunidade de melhoria do método de seleção de modelos em EMP proposto consiste na diminuição do esforço computacional para o cálculo e armazenamento dos AIC_c de todos os possíveis modelos formados pelos termos do modelo base, obtido inicialmente com utilização do método de seleção passo-a-passo, e pelos termos equivalentes aos termos do modelo base. Para se ter uma idéia da quantidade de modelos possíveis, pode-se utilizar como exemplo o caso do misto de retardo apresentado no Capítulo 4. Sabe-se que, para o caso em

questão, o total de termos possíveis (termos do modelo base e termos equivalentes) de pertencer ao modelo é igual a 30. No entanto, destes 30 termos possíveis, os componentes lineares de mistura (v_1 , v_2 e v_3) devem necessariamente pertencer a todos os possíveis modelos a serem gerados. Então, pode-se calcular o número total de possibilidades de modelos em função do número de parâmetros por:

$$(\text{número de modelos}) = C_{30-3}^{p-3} \tag{6.4}$$

Com a fórmula apresentada na Equação (6.4), pode-se obter a Tabela 43, que apresenta o número de modelos possíveis em função do número de parâmetros do modelo. Para o caso em questão, devem ser consideradas mais de 130 milhões de possibilidades, o que representa um esforço computacional considerável para o cálculo e armazenamento de valores de AIC_c de todos os possíveis modelos.

Tabela 43 - Número de modelos em função do número de parâmetros

p	Nº Modelos	p	Nº Modelos	p	Nº Modelos
3	1	12	4.686.825	22	2.220.075
4	27	13	8.436.285	23	888.030
5	351	14	13.037.895	24	296.010
6	2.925	15	17.383.860	25	80.730
7	17.550	16	20.058.300	26	17.550
8	80.730	17	20.058.300	27	2.925
9	296.010	18	17.383.860	28	351
10	888.030	19	13.037.895	29	27
11	2.220.075	20	8.436.285	30	1

Pode-se também obter um gráfico de evolução do AIC_c mínimo em função do número de parâmetros, o qual está apresentado na Figura 43.

Observando a Figura 90, pode-se concluir que a evolução do AIC_c mínimo em função do número de parâmetros consiste de um ramo descendente, atingindo um ponto de mínimo, e um outro ramo ascendente.

Sabe-se que, para o método de seleção de modelos de EMP proposto, só interessam os modelos indiferentes ao modelo que possui o menor valor de AIC_c . Com isso, para o caso em estudo, não seria necessário analisar os modelos com mais de 10 parâmetros, o que reduziria consideravelmente o esforço computacional. No entanto, não há, *a priori*, como saber exatamente como será o real comportamento do valor de AIC_c mínimo em função do número de

parâmetros. O desafio deste problema seria encontrar uma solução analítica do número de parâmetro correspondente ao menor valor de AIC_c , o que seria uma solução exata.

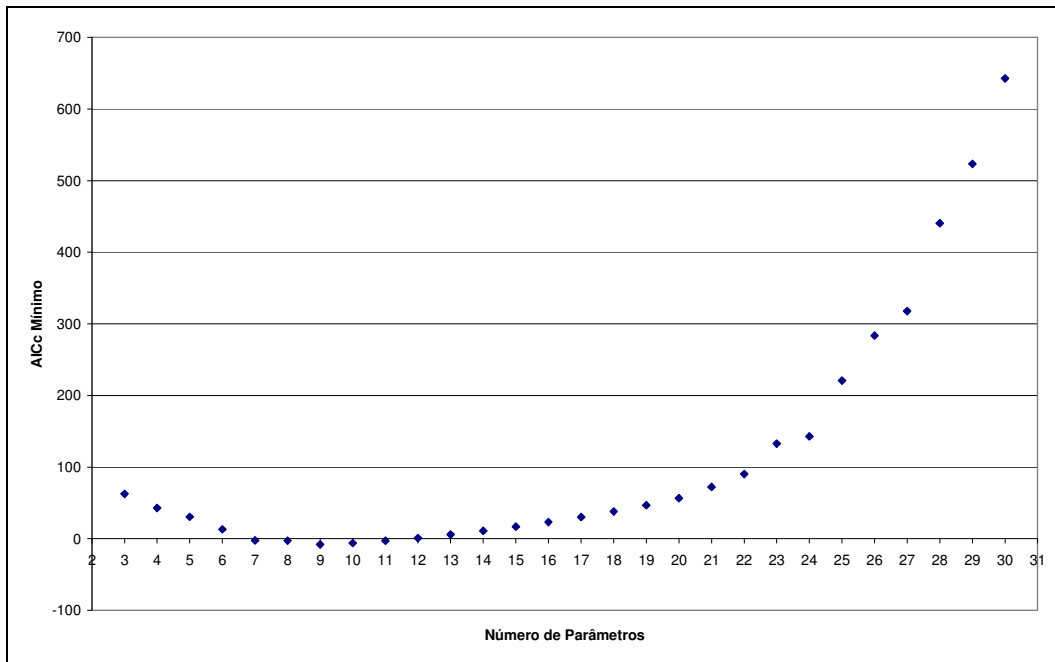


Figura 90 - AIC_c mínimo em função do número de parâmetros

Conforme apresentado, o AIC_c é composto por $n \ln(\hat{\sigma}_p^2)$, onde $\hat{\sigma}_p^2$ é o estimador de máxima verossimilhança da variância do erro, que depende do modelo considerado, e por uma função de penalidade, que é função apenas do número de parâmetros e do número de observações.

Para o caso do misto de retardo, o comportamento do primeiro termo do AIC_c mínimo, $n \ln(\hat{\sigma}_p^2)$, em função do número de parâmetros pode ser observado na Figura 91. Pode-se concluir que, para o caso em questão, $n \ln(\hat{\sigma}_p^2)$ apresenta um ramo descendente, atingindo um mínimo, e um outro ramo ascendente, a medida que o número de parâmetros aumenta.

O comportamento da função de penalidade em função do número de parâmetros pode ser observado na Figura 92. Com o gráfico mostrado na Figura 92 pode-se concluir que a função de penalidade sempre aumenta conforme aumenta o número de parâmetros, independentemente do modelo considerado.

Com isso, o desafio consiste em explorar e prever o comportamento do primeiro termo do AIC_c mínimo, $n \ln(\hat{\sigma}_p^2)$, que depende do modelo considerado,

com o intuito de saber previamente o p correspondente ao menor valor de AIC_c dentre todos os modelos possíveis.

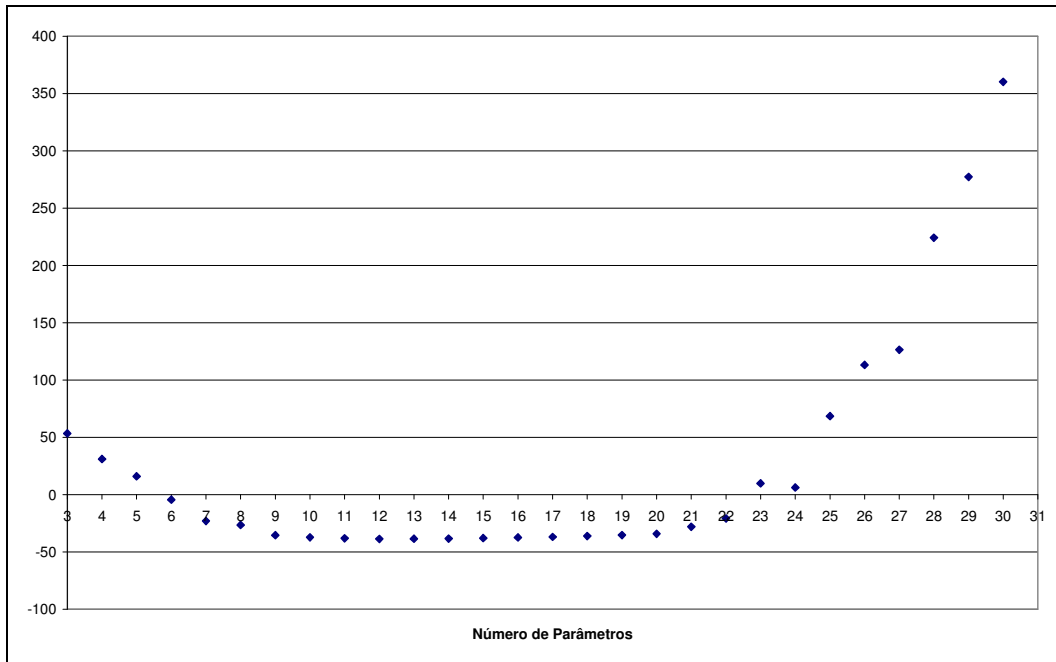


Figura 91 - $n \ln(\hat{\sigma}_p^2)$ versus número de parâmetros

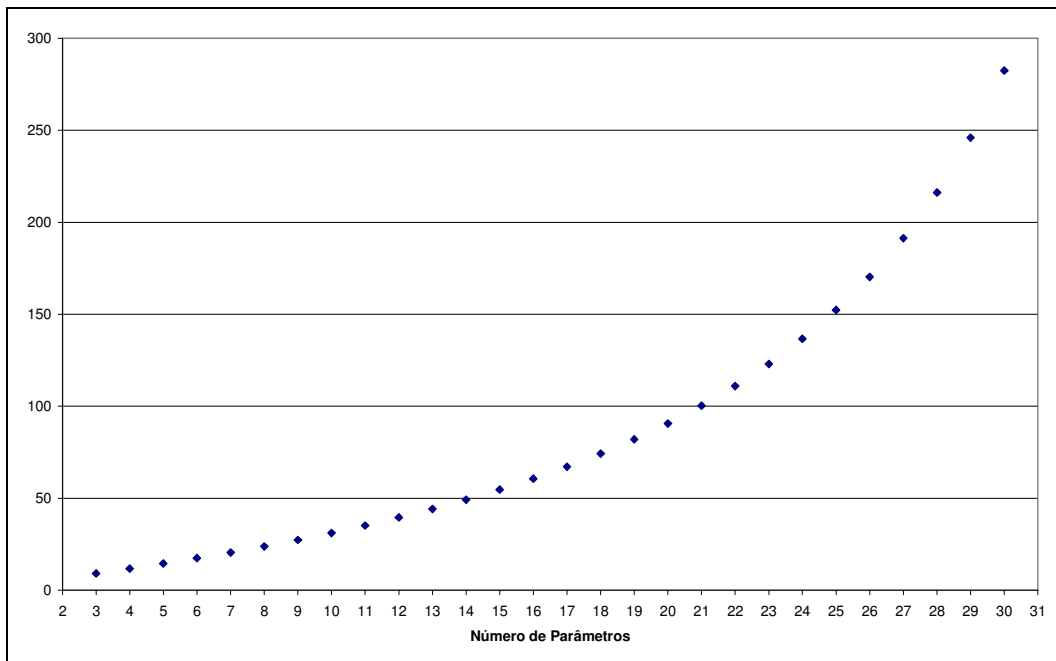


Figura 92 - Função de penalidade em função do número de parâmetros

Uma outra abordagem consiste na tentativa de obtenção de uma solução aproximada, que consiste em encontrar um padrão do comportamento do primeiro termo do AIC_c mínimo em função do número de parâmetros do modelo por meio de ajustes de funções polinomiais ou outras que podem apresentar melhores resultados.