

4 Metodologia

4.1 Origem das amostras

As amostras analisadas neste trabalho foram coletadas durante a realização do projeto SMS, realizado por algumas companhias, através do convênio coordenado pela PUC-Rio, envolvendo o Ministério do Meio-ambiente, o IBAMA e a Petrobras.

As amostras foram coletadas por uma compabnhia ao longo de 2002 e diferentes determinações como granulometria, HPA e metais foram realizados por outra companhia, nestas amostras. Os resultados destas determinações encontram-se em relatórios e foram utilizadas neste presente trabalho.

A malha amostral foi feita através de radiais perpendiculares (*transects*) a batimétrica local. Os *transects* ficaram separados entre si por uma distância de 20 milhas náutica (aproximadamente 37 km), abrangendo toda grande extensão batimétrica das áreas de exploração e produção de petróleo (E&P) da Bacia de Campos, sendo a coleta de sedimento realizada em 13 radiais determinadas (vide Fig. 13).

Da isobáta de 2000 m (zona profunda) até o limite interno da E&P (zona rasa) foram intercaladas estações de coleta, obtendo-se uma distância de 10 mn (milhas náuticas) entre cada estação, aumentando com isso a amostragem das áreas de produção. A cada 20 mn (milhas náuticas) foi realizada coleta em triplicata de sedimento e nas estações intercaladas entre elas não foram realizadas réplicas. Este procedimento foi para amostras de sedimento (estações e poços) que foram coletadas com Box-corer (vide Fig. 14).

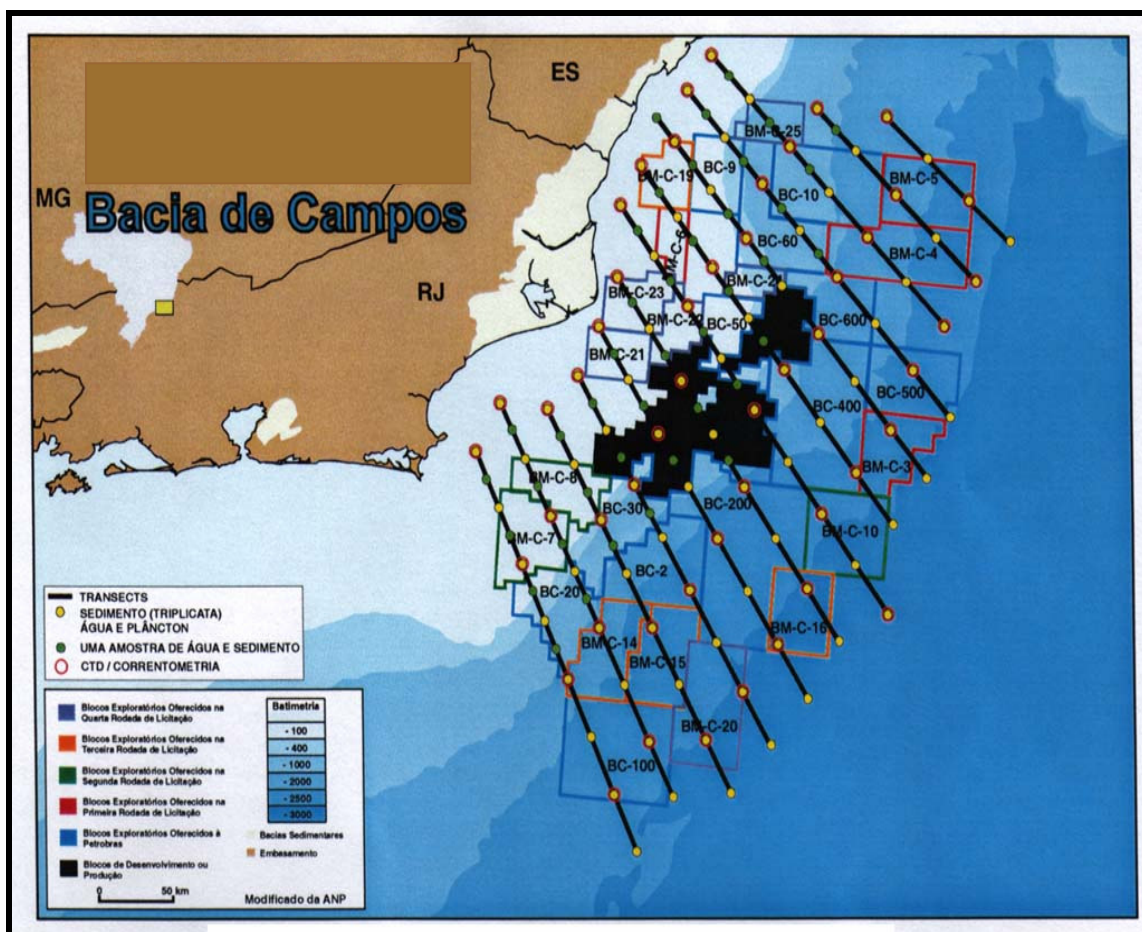


Figura 1-Pontos de amostragem da Bacia de Campos (Petrobras, 2002)

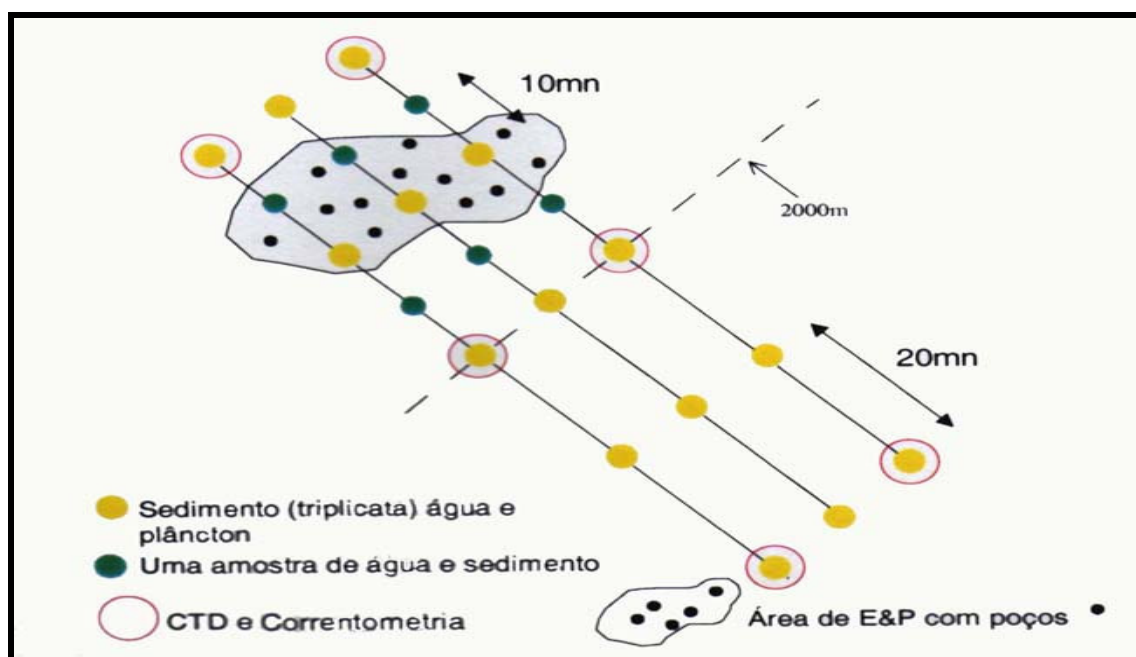


Figura 14-Esquema de desenho amostral (Petrobras, 2002)

4.2 Reagentes

Os ácidos utilizados nas digestões foram HCl 37%v/v, que foi adquirido da Merck (Darmstadt, Germany), o HNO₃ que foi destilado por sub-ebulição e o HF da Merck (Darmstadt, Germany)

As soluções padrão de: Al, Fe, Mn, Cr, Ni, V, Cu, Zn e Pb contendo 1000mg/L de cada elemento foram adquiridos da Merck (Darmstadt, Germany), sendo padrão multielementar para ICP-solução IV, certipur, exceto do elemento Vanádio que foi utilizado o padrão monoelementar da Merck (Darmstadt, Germany), Titrisol.

Antes de serem usados, os frascos de polipropileno utilizados para a determinação das concentrações dos metais das amostras, foram armazenados por vários dias em HNO₃ 10%v/v, posteriormente lavados com água ultra purificada da Gehaka (Goettingeh, Germany), secos e armazenados em potes de polietileno hermeticamente fechados.

4.3 Instrumento

As determinações das concentrações dos metais foram feitas no ICP-OES modelo Optima 4300 (Perkin-Elmer, Shelton, U.S.A.). O instrumento é equipado com amostrador automático AS-90 e controlado pelo programa PE Winlab. As condições de operação do instrumento são mostradas na Tabela 2.

As digestões foram feitas no digestor de microondas da Provecto Analítica modelo DGT 100 Plus (Brasil). O instrumento é operado numa frequência de 2458 MHz, com potência de 0-900W com ajuste de incremento de 10 W, contendo um carrossel para 12 amostras, porém não apresenta controle de temperatura. O frasco de amostra apresenta um volume de 100 mL, contudo suporta um volume máximo de 20 mL de mistura ácida, dependendo do procedimento empregado.

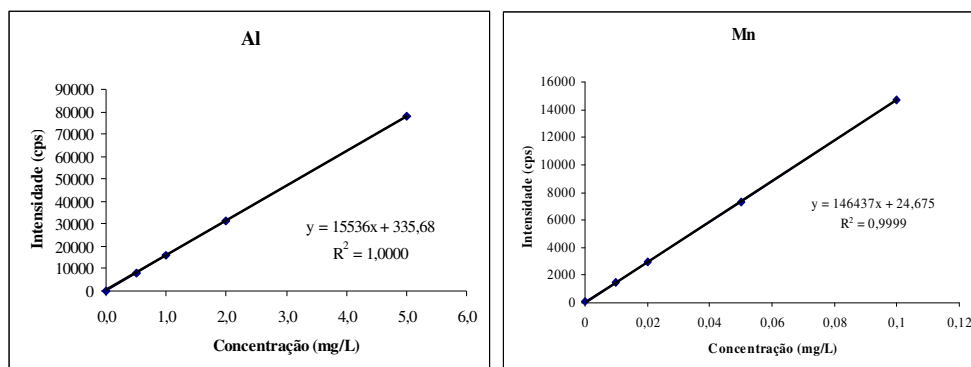
Parâmetros	Valores
Potência de radiofrequência	1500W
Vazão de argônio: plasma	15,0 L/min
auxiliar	0,2 L/min
nebulizador	0,7 L/min
Vazão de aspiração de amostra	1,5 L/min
Tipo de nebulizador	Concêntrico
Câmara de nebulização	Ciclônica
Correção de background	Automática
Tempo de espera	15 s
Tempo de rinsagem	20 s
Número de réplicas	3

4.3.1

Parâmetros característicos de desempenho da técnica de ICP-OES

A calibração do ICP-OES foi realizada com as soluções padrão estoque de 1000mg/L de cada elemento proveniente da Merck (Darmstadt, Germany), sendo diluídas com HNO₃ 0,5%v/v obtendo-se as seguintes concentrações: 0,05; 0,10; 0,20 e 0,50 mg/L.

Os parâmetros de interseção, inclinação e coeficientes de correlação das curvas dos elementos quantificados por ICP-OES estão apresentados na Tabela 3. Podem-se observar pela Figura 15 que as curvas de calibração apresentam excelentes linearidades, porque os coeficientes de correlação variaram de 0,9996 a 1,0000.



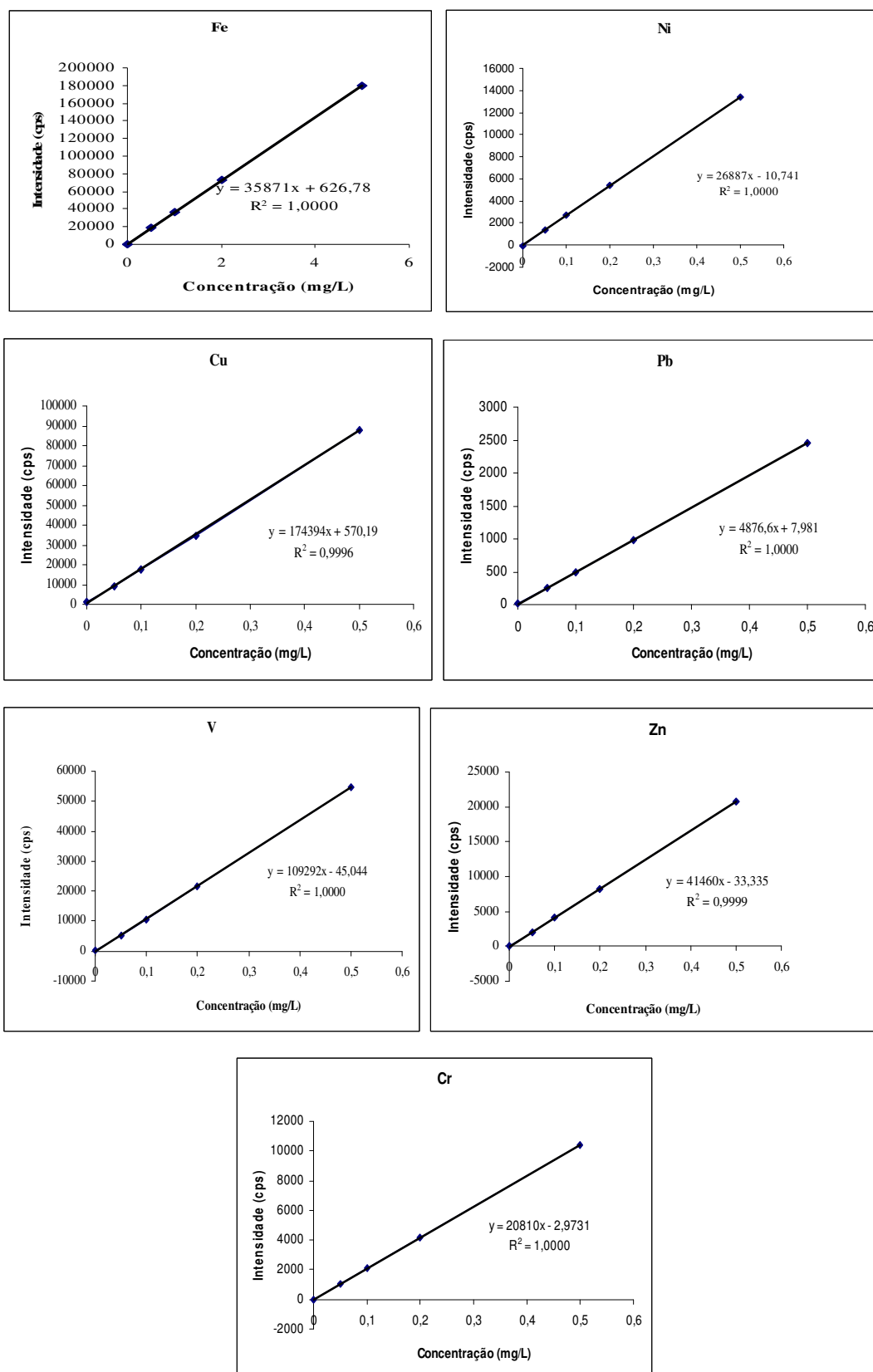


Figura 15 - Curvas de calibração dos elementos: Al, Fe, Mn, Ni, Pb, V, Zn e Cr em meio nítrico utilizadas para análise de amostras de sedimento por ICP-OES

O limite de detecção (baseado em 3σ da solução branco) e o limite de quantificação (baseado em 10σ da solução branco) foram calculados para cada elemento determinado, sendo estes valores mostrados na Tabela 4. Foi usado como padrão interno o Sc na concentração de 0,5 mg/L, para compensar as diferenças de viscosidade e salinidade entre os padrões e as amostras. A correção interelemento foi feita com o programa da Perkin Optima 4300, para compensar qualquer interferência espectral encontrada no comprimento de onda usado para os diferentes metais determinados.

Os materiais de referências certificados utilizados na digestão, no microondas, para poder avaliar os parâmetros estudados foram: MESS-3 sedimento marinho do National Research Council of Canada (NRCC) e SRM 1645 sedimento de rio do National Institute of Standards and Technology (NIST).

Tabela 3 - Parâmetros obtidos pelas curvas de calibração, em meio nítrico, utilizadas para a análise dos elementos em amostras de sedimento por ICP-OES

Analito	Linha Analítica	Sensibilidade(cps.L/mg)	Interseção com o eixo y	Coefficiente do correlação
Al	396,153	15536	336	1,0000
Fe	238,204	35871	327	1,0000
Mn	231,604	146437	25	0,9999
Ni	231,604	26887	11	1,0000
Cu	324,752	174394	570	0,9996
Pb	220,353	4877	8	1,0000
V	292,402	109292	45	1,0000
Zn	202,548	41460	33	0,9999
Cr	205,56	20810	3	1,0000

Tabela 4-Limites de detecção e quantificação ($\mu\text{g/g}$) obtidos utilizando o instrumento ICP-OES Perkin-Elmer Optima 4300

Elemento	Comprimento de onda (nm)	Limite de detecção (3σ)	Limite de quantificação (10σ)
Al	396,153	18	59
Fe	259,939	5	16
Mn	257,610	0,4	1,2
V	292,402	0,04	0,2
Ni	231,604	0,06	0,2
Cr	205,560	0,06	0,2
Cu	324,752	3,8	12,4
Zn	202,548	5,8	19,6
Pb	220,353	0,02	0,06

4.4 Procedimento

4.4.1 Digestão no microondas

Antes de cada digestão no microondas, os frascos eram limpos com 19 mL de HNO_3 10%v/v, aquecido por 30 minutos a 400W. Os frascos eram de teflon quimicamente modificado, suportando uma potência máxima de 900W de operação. O digestor de microondas não possui controle de temperatura, tem um carrossel para 12 amostras, mas foram utilizados apenas seis em cada digestão.

As etapas do programa utilizado para o aquecimento de digestão no microondas são mostradas na Tabela 5.

Tabela 5-Programa para digestão utilizando o microondas DGT 100 plus da Provecto Analítica

Etapa	Potência (W)	Tempo (min)
Programa		(Total = 40 min)
1	400	5
2	600	20
3	400	5
4	0	10

Diferentes massas de amostras têm sido utilizadas para a digestão de amostras de sedimentos com microondas, na literatura, variando de 50 mg (Bertinelli *et al*, 2000) a 1000 mg (Kodot *et al*, 1992). Neste estudo, optou-se por 250 mg que é a quantidade mínima recomendada pelos materiais de referência certificados utilizados neste trabalho.

Após o término do processo de digestão, o frasco de reação era lavado várias vezes com água ultra purificada da Gehaka (Brasil), para a transferência ser quantitativa, a amostra era centrifugada por 5 minutos com uma velocidade de 3000 rotações por minuto, na centrífuga da Fanem (Brasil), depois a parte líquida era transferida para um frasco volumétrico e o volume da amostra era completado até 50 mL, com água ultra purificada. Para cada digestão foi feito um reagente branco, para posteriormente determinação da concentração dos metais por ICP-OES.

4.4.2 Mistura ácida utilizada

Os ácidos nítrico e o clorídrico ou uma combinação destes são os mais utilizados para a digestão de amostras de sedimento de acordo com a literatura (Nadkarni, 1984; Mahan *et al*, 1987; Bertinelli *et al*, 1987; Lamothe *et al*, 1988; Rantala e Loring, 1989;

Kodot *et al*, 1992; Elwaer e Belzile, 1994; Berttinelli *et al*, 2000; Sandroni e Smith, 2002; Cid *et al*, 2002). Quatro combinações foram feitas para determinar as melhores condições de digestão para cada material de referência certificado conforme pode ser visto na Tabela 6.

Tabela 6-Ácidos utilizados na digestão com suas respectivas condições

Ácidos/ Misturas	Volume (mL)			Condição
	HNO ₃	HCl	HF	
HNO ₃	8	0	0	100% de ácido nítrico e 0% de ácido clorídrico
HNO ₃ + HCl	4	4	0	50% de ácido nítrico e 50% de ácido clorídrico
HNO ₃ + HCl	2	6	0	25% de ácido nítrico e 75% de ácido clorídrico
HNO ₃ + HCl + HF	2	6	1	22% de ácido nítrico, 67% de ácido clorídrico e 11% de ácido fluorídrico

4.4.3

Experimentos multivariados e processamento de dados

A otimização das condições de digestão foi realizada utilizando três planejamentos fatoriais 2³, com três variáveis e 2 níveis (Neto *et al*, 2001), em cada um e com duas réplicas para cada experimento de cada planejamento, realizando análise de variância (ANOVA), para se determinar se os valores eram compatíveis para a determinação da média, os cálculos foram feitos com o software Statistica 6.0. Os planejamentos fatoriais apresentavam como fator 1 mistura de ácidos em 2 níveis, o fator 2 em dois níveis e o fator 3 em dois níveis (Tabela 7) A recuperação de cada elemento foi calculada como:

$$Recuperação, \% = \frac{\text{concentração do elemento no experimento}}{\text{concentração do elemento no CRM}} \times 100$$

A resposta de cada experimento foi definida como a média das recuperações de todos os elementos nos dois materiais de referência certificados usados, como:

$$Resposta, \% = 1/18 [R_{Al\ Mess} + R_{Fe\ Mess} + R_{Mn\ Mess} + R_{V\ Mess} + R_{Ni\ Mess} + R_{Cu\ Mess} + R_{Cr\ Mess} + R_{Zn\ Mess} + R_{Pb\ Mess} + R_{Al\ NIST} + R_{Fe\ NIST} + R_{Mn\ NIST} + R_{V\ NIST} + R_{Ni\ NIST} + R_{Cu\ NIST} + R_{Cr\ NIST} + R_{Zn\ NIST} + R_{Pb\ NIST}]$$

Os cálculos para otimização dos experimentos foram feitos com o software Statistica 6.0, através do planejamento fatorial 2^k, da análise de variância (ANOVA) e dos gráficos Pareto.

Tabela 7 – Planejamentos 1, 2 e 3

Tabela 7a- Planejamento 1

Fator	Nível (-)	Nível (+)
1-Ácidos/Misturas	8 mL HNO ₃	4 mL HNO ₃ + 4 mL HCl
2-Potência (W)	600	890
3-Tempo (min)	30	40

Tabela 7b- Planejamento 2

Fator	Nível (-)	Nível (+)
1-Ácidos/Misturas	4 mL HNO ₃ + 4 mL HCl	2 mL HNO ₃ + 6 mL HCl
2-Potência (W)	600	890
3-Tempo (min)	30	40

Tabela 7c- Planejamento 3

Fator	Nível (-)	Nível (+)
1-Ácidos/Misturas	2 mL HNO ₃ + 6 mL HCl	2 mL HNO ₃ + 6 mL HCl +1mL HF
2-Potência (W)	600	890
3-Tempo (min)	30	40

4.4.4

Análise estatística univariada

Com as concentrações dos metais: Al, Fe, Mn, Cr, Ni, V, Cu, Zn e Pb das amostras de sedimento da Bacia de Campos realizaram-se análises estatísticas univariadas e multivariadas.

4.4.4.1

Determinação dos valores anômalos (*outliers*) e normalidade dos resultados

Algumas considerações devem ser tomadas antes de realizar os métodos estatísticos, como: os erros das medidas devem ser aleatórios; a variação das concentrações é decorrente das variações dos sedimentos e não a incerteza das medidas e o número de amostras devem ser maior que o número de variáveis (Johnson e Wichern, 1999).

Primeiramente foi realizada a regressão simples para identificar valores anômalos (*outliers*), que são valores medidos, mas únicos que não representam a base de dados como um todo, ou podem ser uma contaminação na medida. Isto pode influenciar na análise estatística de forma decisiva, por não estarem representando a variabilidade média da base de dados.

Depois foi aplicado o teste de Kolmogorov-Smirnov ($p > 0,20$ K-S) para verificar se os resultados encontravam-se distribuídos dentro de uma normalidade.

As análises de valores anômalos (*outliers*) e verificação da distribuição da normalidade foram realizadas com o software Statistica 6.0.

4.4.4.2

Estatística descritiva, determinação do índice de geoacumulação e fator de enriquecimento

Com os resultados obtidos foram determinados a concentração mínima, a concentração máxima, a média das concentrações, a concentração de *background*, o fator de enriquecimento (EF) e o índice de geoacumulação (IG).

A concentração de *background* pode ser determinada de várias maneiras diferentes. Alguns autores utilizam concentrações já determinadas na região de estudo e tabeladas na literatura (Yongming *et al*, 2005), mas que no trabalho em questão ainda não se encontram disponíveis para serem utilizadas. Outros autores lançam mão da concentração mínima determinada no estudo (Loska e Wiechula, 2003) e outros autores usaram a média da concentração de oito amostras da região de estudo que não se encontravam contaminadas (Yu *et al*, 2001). No trabalho em questão optou-se em utilizar a média da concentração de dez amostras, que apresentaram os resultados mais baixos, para cada elemento.

O índice de geoacumulação (IG) e o fator de enriquecimento (EF) podem ser utilizados para diferenciar entre a origem dos metais com processos naturais das atividades humanas, indicando assim o grau de influência da ação antropogênica no meio.

O índice de geoacumulação (IG) foi determinado pela primeira vez através da equação de Müller (1979):

$$I_{geo} = \log_2 \frac{C_n}{1,5B_n}$$

onde: C_n é a concentração do metal (mg/kg) no sedimento da região e B_n é a concentração de *background* (mg/kg) do metal e o fator 1,5 é utilizado para compensar possíveis variações de *background* devido a efeitos litogênicos. O índice de geo-acumulação (IG) possui sete graus de intensidade de contaminação, onde IG próximo a zero significa sedimento não-contaminado e deficiente do metal indo até IG = 6 que significa que o sedimento se encontra fortemente contaminado. O fator de enriquecimento pode ser calculado de algumas maneiras diferentes, Zonta *et al* (1994) utilizaram a seguinte equação:

$$EF(\%) = \frac{C - C_{\min}}{C_{\max} - C_{\min}} \times 100\%$$

Onde: C -média da concentração do elemento X (unidade em $\mu\text{g/g}$);

C_{min} e C_{max} - concentrações mínima e máxima determinadas no período de investigação do trabalho (unidades em $\mu\text{g/g}$).

Kemp *et al* (1976) introduziu pela primeira vez, a razão a seguir, usando como elemento de referência o Al para a normalização:

$$EF = [X_i] / [X_0] : [Ref.i] / [Ref.0]$$

Onde: $[X_i]$ e $[Ref.i]$ – concentrações do elemento X e do elemento de referência na amostragem i;

$[X_0]$ e $[Ref.0]$ – concentração de *background* do elemento X e do elemento de referência.

Vários autores têm utilizado alguns elementos de referências diferentes, que geralmente, são elementos conservativos como: Fe, Mg, Mn, Sc, Ti etc (Quevauviller *et al*, 1989; Isakson *et al*, 1997; Lee *et al*, 1998; Reimann e Caritat, 2000; Bergamaschi *et al*, 2002; Hernandez *et al*, 2003; Conrad e Chisholm-Brause, 2004; Mishra *et al*, 2004). O fator de enriquecimento foi organizado em 5 categorias por Loska e Wiechula (2003), como pode ser visto na Tabela 8. Neste trabalho optou-se pela equação de Kemp *et al* (1976) e como elemento de referência o Fe, porque na metodologia empregada para a digestão das amostras, este elemento apresentou a maior recuperação no material de referência certificado (Tab 12).

Tabela 7-Categorias de contaminação baseada no efeito de enriquecimento (EF)

EF	Enriquecimento
< 2	Deficiente
2-5	Moderado
5-20	Significativo
20-40	Muito alto
>40	Extremamente alto

4.4.4.3

Regressão simples

O objetivo da análise de regressão simples é prever uma única variável dependente a partir do conhecimento de uma variável independente. Neste trabalho foi escolhido o elemento Fe (variável independente), por ser um elemento natural, que se apresenta em

concentrações altas e constantes, não sofrendo influências antropogênicas (Förstner e Wittmann, 1979).

Os parâmetros mais importantes para serem analisados na regressão simples são: intercepto (b_0), coeficiente de regressão (b_1), o coeficiente de determinação (R^2) e o resíduo.

O intercepto (b_0) significa o valor da constante do eixo Y, que é o eixo da variável dependente, onde a reta é definida pela equação de regressão: $Y = b_0 + b_1X$. O coeficiente de regressão (b_1) que representa o montante da variação na variável dependente em relação a uma variação na variável independente. O coeficiente de determinação (R^2) que mede a proporção da variância da variável dependente em torno de sua média, sendo explicada pela variável independente. Este coeficiente varia de 0-1 e quanto maior o valor, maior será o poder de explicação da equação de regressão, sendo, portanto, melhor a previsão da variável dependente. O resíduo é o erro na previsão dos dados da amostra, raramente as previsões serão perfeitas, considera-se que o erro ocorra aleatoriamente, sendo atribuído para o erro na população uma média de 0 a uma variância constante (Hair, 2005). A análise de regressão simples foi realizada com o software Statistica 6.0.

4.4.5

Estatística multivariada

A análise multivariada é uma ferramenta de grande valia, porque possibilita a interpretação de um grande número de dados por métodos estatísticos, levando a resultados que, principalmente: possibilitam a redução de dados sem comprometer o valor da informação inicial; identifica grupos de variáveis semelhantes; identifica as relações entre as variáveis e determina como as variáveis se correlacionam.

A base de dados foi preparada, fazendo-se um escalonamento, através da padronização, para obter média zero e desvio padrão um, permitindo assim o uso de variáveis com unidades diferentes.

Por fim foram aplicados os modelos estatísticos multivariados como análise de regressão múltipla (os dados utilizados não foram padronizados, para que o intercepto- b_0 pudesse ser avaliado e não assumisse o valor zero), análise do componente principal (PCA) e análise de agrupamentos (CA) forneceram resultados que permitiram uma ampla compreensão das informações dos dados. Todas as análises de regressão múltipla foram realizadas com o software Statistica 6.0.

4.4.5.1

Regressão múltipla

A análise de regressão múltipla é semelhante à análise de regressão simples, diferindo no número de variáveis independentes, que tem que ser mais de uma, portanto, é uma técnica estatística que analisa a relação entre uma única variável dependente e várias variáveis independentes. Neste trabalho optou-se pelas variáveis independentes como sendo o Fe, o Al e o Mn, que são elementos que apresentam propriedades conservativas, elementos naturais, que se apresentam em concentrações altas e constantes e não sofrem influências antropogênicas (Förstner e Wittmann, 1993).

O grupo de variáveis independentes ponderadas origina a variável estatística de regressão, que também pode ser denominada de equação de regressão ou modelo de regressão, sendo uma combinação linear das variáveis independentes que melhor prevê a variável dependente.

Três parâmetros, o coeficiente de determinação (R^2), os coeficientes de correlações parciais (r) e o resíduo, são os mais importantes para serem analisados na regressão múltipla. O coeficiente de determinação (R^2) e o resíduo já foram explicados na regressão simples. Os coeficientes de correlações parciais (r) medem a relação entre a variável dependente e uma única variável independente, mantendo-se constante as outras variáveis independentes. Estes coeficientes podem variar de -1 a +1, onde +1 indica uma perfeita relação positiva, 0 indica que não existe relação nenhuma e -1 indica uma perfeita relação reversa, ou seja, quando uma aumenta a outra diminui (Hair, 2005).

4.4.5.2

Análise dos componentes principais (PCA)

A análise do componente principal tem como objetivo, principalmente, explicar a variância ou covariância de uma base de dados muito extensa, reduzindo a base de dados para um número menor de componentes principais que contenha o máximo de informações da base original, facilitando assim a interpretação (Johnson e Wichern, 1999).

Os componentes principais são novas variáveis, resultado da correlação entre as variáveis originais, cada componente principal agrupa variáveis que apresentam maior correlação. A matriz de correlação é determinada a partir da matriz padronizada cujo objetivo é a determinação dos autovalores e autovetores. Cada autovalor representa a variância dos componentes principais, indicando o número de variáveis que o componente explica, a soma das variâncias de todos os componentes resulta no número total de

variáveis. Para cada autovalor existe um autovetor associado (*factor loading*) na matriz de correlação, resultando na matriz de autovetores formada por vetores correspondentes a cada componente principal.

Para se obter uma melhor interpretação, maximiza-se a variância dos quadrados dos *loadings* para cada componente, isto é, rotaciona-se as componentes de maneira a maximizar os *loadings* dos autovetores para valores mais próximos de um, este processo denomina-se de Varimax (Hopke, 1992).

4.4.5.3

Análise de agrupamento (Cluster Analysis-CA)

Análise de agrupamento hierárquico tem como objetivo principal identificar grupos com homogeneidade e heterogeneidade entre os grupos, através de um conjunto de dados bastante extenso.

É representado por um dendograma que é um gráfico bidimensional, resultante da análise de agrupamentos, que identifica a semelhança ou não entre eles, que se permita julgar comportamentos como semelhantes, dentro da população de grupos que são expressos através de suas variáveis.

A escolha do coeficiente de semelhança é necessária para que se possa quantificar o quanto as variáveis são parecidas, ou é necessário medidas de similaridade, que indica que quanto maior o valor, mais similares eles são, ou de dissimilaridade, ou seja, quanto menor ele for. Esses coeficientes ou medidas podem ser: a distância Euclidiana que é a distância geométrica em um espaço multidimensional, seu cálculo é realizado a partir dos dados sem a padronização; a distância Euclidiana Quadrática que é a distância Euclidiana ao quadrado; distância de Manhattan que é a soma das diferenças através das dimensões; distância de Chebyshev que, geralmente, se utiliza esta medida quando se quer determinar se as variáveis são muito diferentes entre si num espaço multidimensional; distância de Potência que se utiliza quando se quer aumentar ou diminuir o peso progressivo que é provocado pela dimensão nas quais as variáveis são muito diferentes; porcentagem de desagregamento e Pearson estas duas últimas são utilizadas quando escolha-se a opção de matriz. O coeficiente escolhido neste trabalho foi à distância de Chebyshev, porque demonstrou resultado compatível com a análise do componente principal.

A escolha da regra de *linkage*, também é necessária para se determinar a melhor estratégia de agrupamento: *Single linkage* sendo a distância entre os grupos, é a distância entre os elementos mais próximos, ou seja, é a menor distância entre os grupos; *Complete*

linkage é definida como sendo a distância entre os indivíduos mais distantes dos grupos, ou seja, é a maior distância; *Unweighted pair-group average* é a distância determinada como a média aritmética das distâncias entre todos os pares de valores de um grupo a outro; *Weighted pair-group average* é semelhante ao anterior, sendo usado quando se suspeita que os grupos sejam extremamente desiguais; *Unweighted pair-group centroid* a distância entre os grupos é calculada pela diferença entre os centros de “gravidade” dos grupos que é o ponto médio em um espaço multidimensional definido por suas dimensões; *Weighted pair-group centroid (median)* é semelhante ao anterior, mas quando existem diferenças consideráveis entre os grupos é melhor este método porque um peso é introduzido na análise computacional que leva em consideração as diferenças dos tamanhos dos grupos e o método de Ward que utiliza a análise de variância (ANOVA) para avaliar a distância entre os grupos. A regra escolhida neste trabalho foi o método de Ward, porque foi o que mostrou resultado compatível com a análise do componente principal.

Depois de determinar o número de grupos através da análise de agrupamento pode-se fazer a análise de *K-means clustering* que examina a média para cada grupo em cada dimensão para avaliar como os grupos são distintos. A magnitude do valor de F calculado a partir da análise de variância (ANOVA) realizada em cada dimensão é uma outra maneira de indicar de quanto à dimensão é significativa.

4.4.5.4

Análise e comparação dos dados do relatório do laboratório contratado pela Petrobras com os dados da PUC-Rio

Com os resultados do relatório do laboratório contratado pela Petrobras das amostras de sedimento da Bacia de Campos foram selecionados os seguintes parâmetros: Al, Ba, Cr, Cu, Fe, Mn, Ni, Pb, V, Zn, % areias, % silte, % argilas, % carbonatos, carbono orgânico total (COT) e profundidade, a relação Níquel/Vanádio, índice de geoacumulação (IG), fator de enriquecimento (EF), a distribuição log-normal, estatística descritiva, análise de correlação, análise de componente principal (PCA) e análise de agrupamento (CA) destes parâmetros, com o auxílio do programa SPSS 9.0.

Foi feito à comparação dos dados do relatório do laboratório contratado pela Petrobras com os resultados da PUC-Rio e a esses resultados foram acrescentados alguns parâmetros do relatório e realizado análise de correlação, análise de componente principal (PCA) e análise de agrupamento (CA) com esses dados, através do programa SPSS 9.0