UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

REFLEXÕES SOBRE O TRATAMENTO DE DADOS GEOQUÍMICOS

Arlei Benedito Macedo

Sistematização Crítica de parte da obra do candidato, para concurso para obtenção do Título de Livre-Docente junto ao Departamento de Geologia Econômica e Geofísica Aplicada do Instituto de Geociências da Universidade de São Paulo, na Área de Conhecimento: Prospecção e Avaliação de Jazidas Minerais.

SÃO PAULO
1996

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

REFLEXÕES SOBRE O TRATAMENTO DE DADOS GEOQUÍMICOS

Arlei Benedito Macedo

INSTITUTO DE GEOCIÊNCIAS
BIBLIOTECA
T2302
USP

Sistematização Crítica de parte da obra do candidato, para concurso para obtenção do Título de Livre-Docente junto ao Departamento de Geologia Econômica e Geofísica Aplicada do Instituto de Geociências da Universidade de São Paulo, na Área de Conhecimento: Prospecção e Avaliação de Jazidas Minerais.

DEDALUS - Acervo - IGC
30900005122

SÃO PAULO
1996

Agradecimentos

Este texto é resultado de trabalho feitos desde 1969, para os quais contribuíram muitas pessoas. Agradeço a todas e desculpo-me por não poder incluí-las todas neste texto. Agradeço especialmente:

Ao professor Nabor R. Rüegg, pela orientação e estímulo na primeira fase da minha permanência no Instituto de Geociências. Sua morte prematura foi uma perda muito grande para o ensino e a prática da Geologia no Brasil.

Aos professores Paulo M. B. Landim e Gilberto Amaral, pelos ensinamentos e pelo entusiasmo pelos métodos quantitativos em Geologia, que tão bem me transmitiram.

Aos colegas da Companhia de Pesquisa de Recursos Minerais.

Ao professor Aledir P. Barbour, pela orientação no doutorado e pelo apoio na volta ao Instituto de Geociências.

Ao professor Ginaldo A. C. Campanha, pelo apoio geológico aos trabalhos de Geoprocessamento.

Aos alunos estagiários de Geoprocessamento, e aos que participaram como co-autores dos trabalhos, Marcelo Braghin, Flávio A. F. Salles, Tays Ribeiro, Marcelo Barbara, Sérgio V. Liotte, Carlos C. Araújo e Alexandre C. N. Silva, especialmente aos três últimos, que muito auxiliaram na preparação de figuras para este texto.

Aos professores João B. Moreschi e Fábio Taioli, pelo auxílio na revisão do texto.

A todos os membros do Departamento de Geologia Econômica e Geofísica Aplicada do IG-USP, pelo auxílio na fase de preparação deste texto e do memorial, e por terem assumido boa parte da carga administrativa a mim atribuída no período.

À equipe da gráfica do Instituto de Geociências, por seu eficiente e atencioso trabalho de impressão e encadernação.

E aos que mais me auxiliaram, e que suportaram minha ausência durante a longa fase de redação, filhos Heitor e Stella e esposa Anna, que também muito colaboraram, principalmente na organização do memorial.

ÍNDICE

1. APRESENTAÇÃO 1
2. INTRODUÇÃO 2
2.1 Dados gerais sobre as áreas trabalhadas 4
2.1.1 Área Mina do Paqueiro 8
2.1.2 Área Mina do Perau 15
3. AQUISIÇÃO, ARMAZENAMENTO E RECUPERAÇÃO DE DADOS 19
3.1 Aquisição de dados 19
3.1.1 Generalidades 19
3.1.2 Procedimentos operacionais para coleta e conversão de dados 21
3.2 Armazenamento e recuperação 23
4 CRÍTICA DE DADOS 25
5. ANÁLISE DE DADOS UNIVARIÁVEL 31
6. ANÁLISE DE DADOS MULTIVARIÁVEL 42
6.1. Análise a duas variáveis 42
6.1.1. Introdução 42
6.1.2. Gráficos de dispersão 43
6.1.3. Coeficientes de correlação 47
6.1.3.1. Coeficiente de correlação linear produto-momento de Pearson 48
6.1.3.2. Coeficiente de correlação não-paramétrica de Spearman 48
6.1.4. Gráficos teor-distância 50
6.2. Múltiplas variáveis 53
6.2.1. Análise de agrupamentos 53
6.2.2. Análise de Componentes Principais 55
6.2.3. Análise fatorial 56
6.2.4. Regressão Múltipla 58
6.2.4.1. Regressão de teores de elementos principais a partir de elementos traços em rochas basálticas da Bacia do Paraná 58
6.2.4.2. Regressão da distância à mineralização a partir de teores de elementos principais e traços em rochas na área do Perau 65
6.2.5. Análise Discriminante 68
6.2.6 Comentários sobre as conclusões gerais 69

7. VARIAÇÃO ESPACIAL - GEOPROCESSAMENTO 79
7.1 Organização 81
7.2 Visualização 82
7.2.1 Generalidades 82
7.2.2 Visualização a duas, duas-e-meia e três dimensões 83
7.3 Consulta espacial 89
7.4 Combinação 89
7.5 Análise 89
7.5.1 Fundamentos teóricos 91
7.5.2 Análise com um mapa 94
7.5.2.1 Representação de teores e anomalias 94
7.5.2.2 Análise de superfícies de tendência 98
7.5.2.2.1 Generalidades 98
7.5.2.2.2. Aplicação ao estudo da variação dos teores do minério da mina do Perau 99
7.5.2.2.3 Aplicação para o estudo da variação dos teores de chumbo na área Mina do Paqueiro 99
7.5.2.3 Autocorrelação 102
7.5.2.4. Variogramas 102
7.5.3 Análise com múltiplos mapas 104
7.5.3.1 Combinação de anomalias 104
7.5.3.2 Análise de componentes principais 108
7.6 Predição 112
7.7 Apoio à decisão 112
7.8 Produção cartográfica 112
7.9 - Exemplo de aplicação de geoprocessamento à análise metalogenética regional: folhas Cerro Azul e Apiaí 114
8 CONCLUSÕES 141
9 BIBLIOGRAFIA CITADA 143

REFLEXÕES SOBRE O TRATAMENTO DE DADOS GEOQUÍMICOS

Arlei Benedito Macedo

"What could be cutter
Than to feed a computer
With wrong information
But naïve expectation
To obtain with precision
A Napoleonic decision ?"

Major Alexander de Seversky
apud Davis (1973)

1. APRESENTAÇÃO

Este trabalho analisa parte da obra do autor, o qual, além de trabalhar com métodos de Prospeção Mineral, também desenvolve atividades em Educação e Geologia Ambiental. Em todas essas áreas utiliza técnicas quantitativas, especialmente com computadores.

Aqui é revisada a experiência do autor no tratamento de dados geoquímicos, iniciada em 1969, quando aluno de graduação e bolsista de Iniciação Científica no Centro de Pesquisas Geocronológicas do Instituto de Geociências da USP. Desde aquela época os equipamentos e técnicas progrediram enormemente. Isto faz com que atualmente sejam necessárias uma formação especializada, sempre atualizada, e uma reflexão sobre as técnicas a empregar, para que possa ser feito o tratamento de dados com sucesso.

O autor espera contribuir com esta revisão no sentido de facilitar o caminho dos que tratam dados geoquímicos, apontando as técnicas que usou, as que tiveram sucesso e, provavelmente mais importante, os erros e

problemas. Os trabalhos realizados pelo autor e aqueles nos quais teve participação serão analisados, visando uma atualização, confrontando os resultados obtidos na época com aqueles possíveis com as técnicas atuais. Além dos trabalhos publicados em revistas ou anais de congressos, também serão citados relatórios de projetos executados no período em que o autor trabalhou na Companhia de Pesquisa de Recursos Minerais - CPRM (1978-1987), apresentados em tiragens restritas, e para cuja consulta ou cópia é necessário recorrer às bibliotecas da CPRM, do Departamento Nacional da Produção Mineral - DNPM, do Instituto de Pesquisas Tecnológicas - IPT ou, para alguns projetos, do Instituto de Geociências da USP. Embora estes trabalhos não possam ser, a rigor, considerados "obra publicada", sua citação é aqui obrigatória, sob pena de cortar o comentário de técnicas importantes.

Uma nota sobre a impressão das figuras: na maioria dos casos elas foram impressas a partir do próprio programa de análise (IDRISI, Probplot), ou com pequenas modificações introduzidas com programas de desenho matriciais (Paintbrush, Photostyler, Photoshop). Uma alternativa melhor seria passar do IDRISI para um programa de desenho vetorial (como o Corel Draw), o que permitiria desenhos melhor definidos. Isto não foi feito porque o ganho em aparência seria pequeno, comparado com a carga de trabalho, demasiada para os recursos disponíveis.

2. INTRODUÇÃO

Este trabalho visa revisar parte da obra do autor, aquela mais constante entre suas atividades, desde 1969. A revisão não será cronológica, em relação à data de publicação dos trabalhos, mas obedecerá à ordem de complexidade no tratamento dos dados, o que resulta aproximadamente numa ordem cronológica do desenvolvimento das técnicas. Embora esta maneira de comentar os trabalhos torne mais difícil para o leitor a compreensão dos trabalhos, principalmente dos de maior extensão, cujas partes são apresentadas em diferentes partes do texto, ela facilita a apresentação das técnicas e a reflexão sobre elas, meta desta revisão.

Três são os objetivos gerais deste texto: em primeiro lugar rever os trabalhos realizados pelo autor, atualizando-os; em segundo lugar, refletir sobre as técnicas de tratamento de dados geoquímicos, comparando as conclusões do autor com aquelas de outros que pensaram sobre o assunto e finalmente, e talvez o mais importante, apresentar os resultados de forma clara, que permita a utilização do texto na formação de outros técnicos e pesquisadores. Esta finalidade didática é coerente com o trabalho principal do autor, que se considera muito mais um professor do que um técnico ou pesquisador.

A revisão das técnicas não será exaustiva, concentrando-se naquelas em que o autor tem experiência direta, com pequena menção a outras disponíveis.

Serão apresentadas as técnicas para aquisição, armazenamento e recuperação de dados, sua crítica, sua análise univariável e multivariável, inicialmente sem levar em conta sua localização espacial e, finalmente, a sua apresentação e análise espacial. Será dada maior ênfase às técnicas de apresentação e análise espacial, principalmente ao Geoprocessamento, devido à importância que essas técnicas alcançaram nos últimos anos.

A divisão dos tópicos e alguns dos comentários de técnicas aqui apresentados são derivados de trabalho de revisão anterior, não publicado, escrito pelo autor como tema de exame de qualificação para doutorado (Macedo, 1985).

Um ponto a considerar desde o início é a natureza do trabalho quantitativo em Geologia, e particularmente em Geoquímica. O geólogo emprega técnicas quantitativas, algumas delas bem estabelecidas em Estatística, outras ainda não bem definidas, e que carecem de testes de significância e desenvolvimento matemático rigoroso, chamadas "quase-estatísticas" ou "proto-estatísticas" por Davis (1986). Mesmo para técnicas bem definidas, freqüentemente não é possível fazer estimação de parâmetros populacionais e testes de hipótese rigorosos porque a amostra estatística disponível não cumpre os requisitos de amostragem casual, principalmente a igual probabilidade dos pontos da população de fazer parte da amostra (devido à dificuldade de acesso a todos os pontos da população) e a

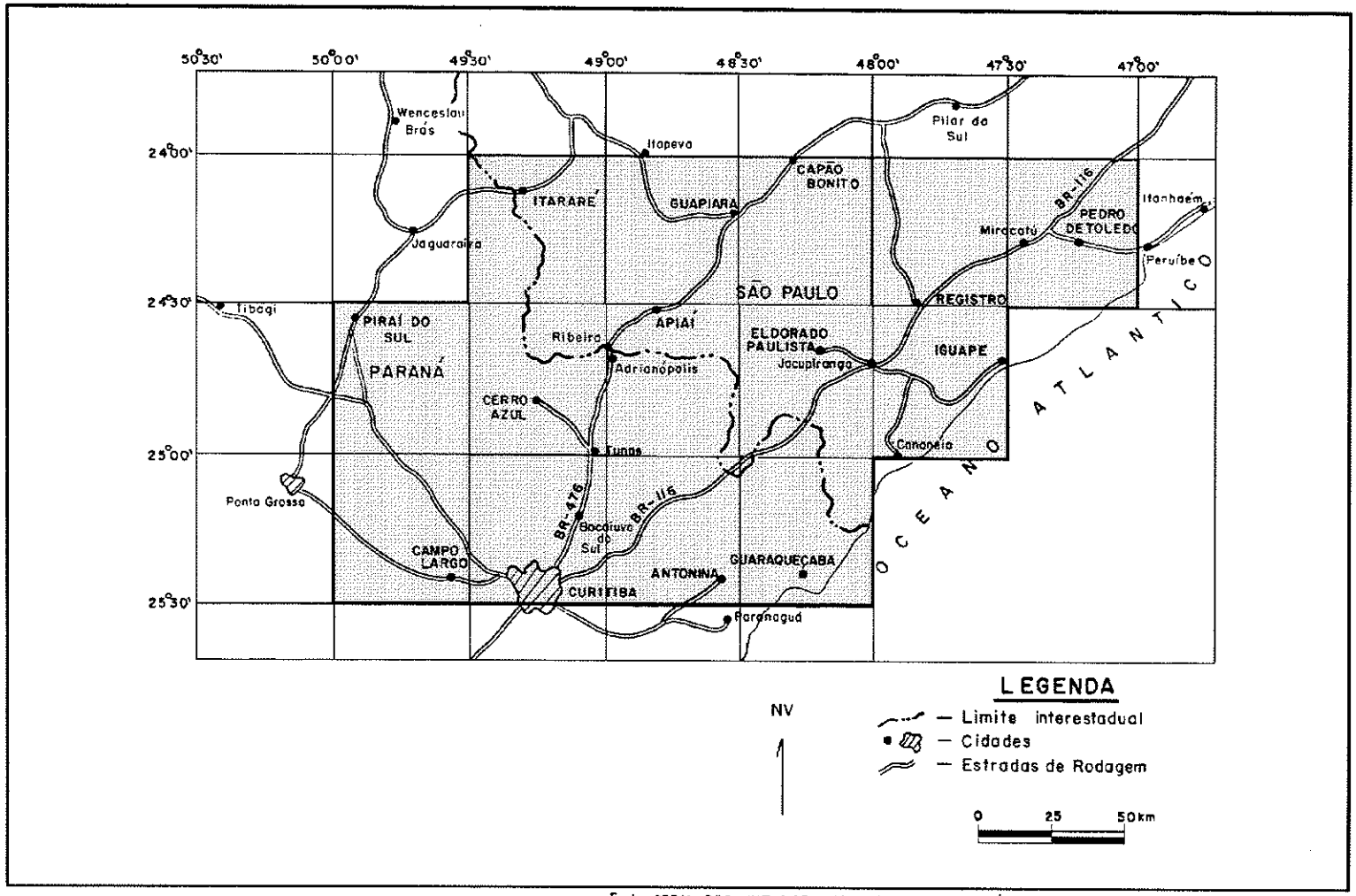
independência das amostras (muitas vezes são auto-correlacionadas). Por isto o título desta revisão fala em "tratamento de dados" e não "estudo estatístico".

2.1 Dados gerais sobre as áreas trabalhadas

Os métodos a seguir comentados foram aplicados sobre áreas amostradas para prospecção geoquímica ou para detecção de poluição. Quase todos foram aplicados a áreas do Vale do Ribeira, sendo descritos em trabalhos publicados e apresentados de 1981 a 1995. Aqui se apresenta uma descrição geral da região abrangida pela Província Mineral do Vale do Ribeira, a qual é descrita de modo mais completo em Batolla et alii (1981) e em Campanha (1991, 1995).

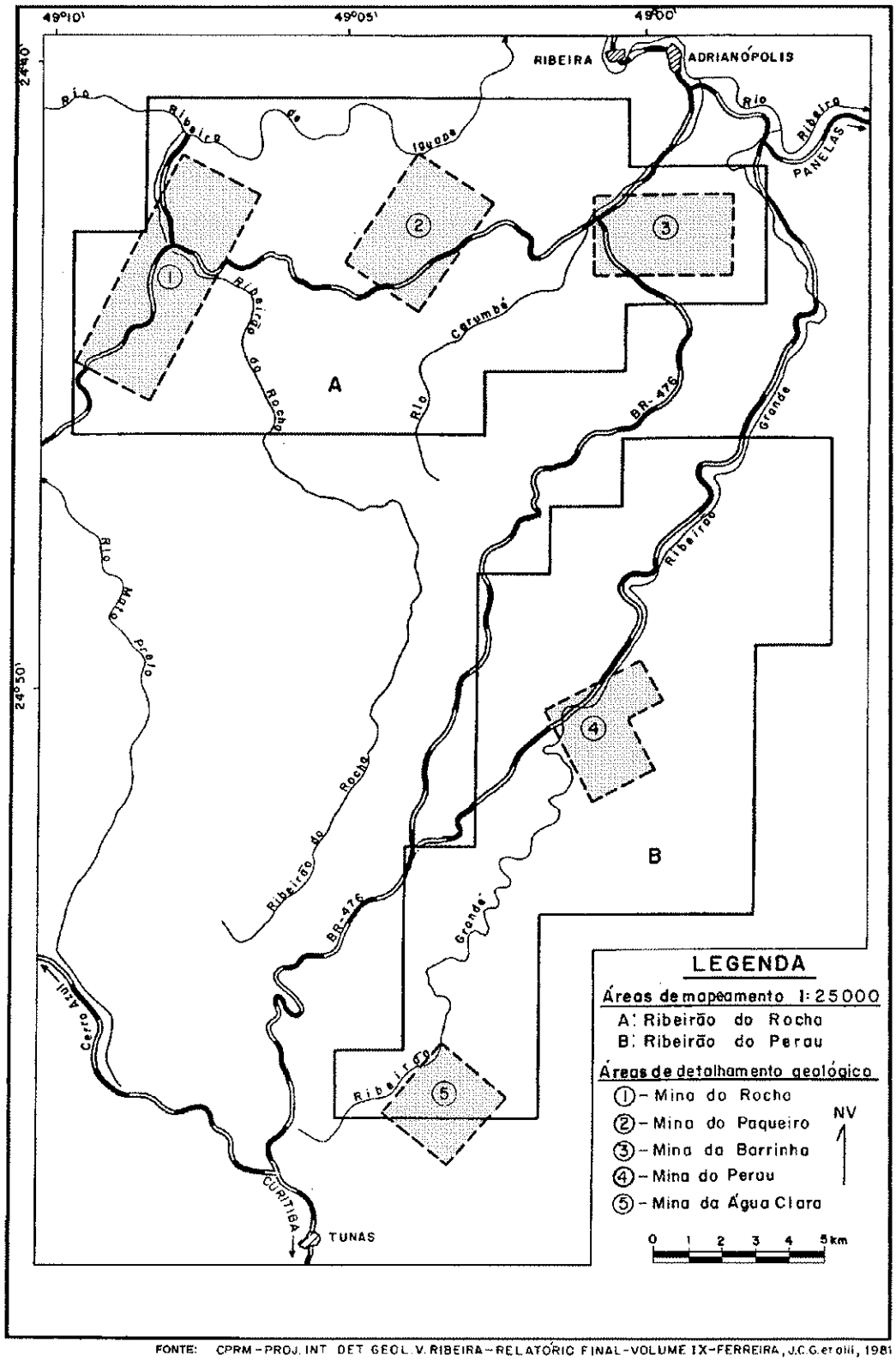
Os leitores terão uma dificuldade, que será a de acompanhar os trabalhos nas diversas áreas de forma fragmentada, pois eles serão apresentados por técnica, descrevendo-se, para cada técnica, os resultados nas diversas áreas. Esta dificuldade é incontornável. A alternativa seria apresentar os resultados por área, descrevendo-se, para cada uma, os resultados das técnicas ali aplicadas. Foi escolhida a apresentação por técnica, pois o principal objetivo desta revisão é metodológico, e não da descoberta ou apresentação de novos fatos sobre as áreas.

Entre 1978 e 1981 o autor foi encarregado dos trabalhos de Geoquímica e Geofísica do Projeto Integração e Detalhe Geológico no Vale do Ribeira - PIDGVR, cujos resultados estão expostos em Batolla et alii (1981). Nesse projeto foi feita a integração geológica a 1:100.000, na área mostrada na figura 1 e levantamento geológico de detalhe em duas áreas a 1:25.000 (área do Rocha e área do Perau). Seis áreas com mineralizações conhecidas foram detalhadas a 1:5.000, fazendo-se uma descrição e interpretação das dessas mineralizações, acompanhada de amostragem de solos, rochas e sedimentos de corrente, e levantamento magnetométrico. A localização de todas as áreas de levantamentos de detalhe é mostrada na figura 2.



Fonte: CPRM - PROJ. INT. DET. GEOL. V. RIBEIRA - RELATÓRIO FINAL - VOLUME IX - FERREIRA, JCG et alii, 1981

Figura 1 - Localização da região estudada pelo Projeto Integração e Detalhe Geológico no Vale do Ribeira, destacando a área abrangida pelas folhas onde foi executada a integração geológica a 1:100.000.



FONTE: CPRM - PROJ. INT DET GEOL. V. RIBEIRA - RELATÓRIO FINAL - VOLUME IX - FERREIRA, J.C.G. et alii, 1981

Figura 2 - Localização das áreas de mapeamento à escala de 1:25.000 e de detalhamento geológico.

Em Macedo (1986) é assim descrita, resumidamente, a geologia regional da área da Província Mineral do Vale do Ribeira, baseada em Chiodi (1984):

> As rochas arqueanas predominam na parte Sul, também ocorrendo no núcleo de antiformas (Anta Gorda, Água Clara e Perau) e janelas na região de Capão Bonito. São cobertas por metamorfitos proterozóicos de distribuição complexa, em faixas de orientação predominante NE-SW, muito afetadas por extensas falhas (Lancinha, Agudos Grandes, Figueira, Morro Agudo, Carumbé e Espírito Santo são as principais). Intrusões graníticas de grande porte também seguem a orientação NE-SW. A NW da área rochas sedimentares da Bacia do Paraná recobrem as metamórficas, que são também cortadas por diques de diabásio, derivados da reativação mesozóica.

Até este nível de descrição não há grandes polêmicas. Quando se trata de subdividir os diversos conjuntos de rochas, atribuir-lhes idades e contextos tectônicos de formação, montando a estratigrafia, há grande discordância entre os diversos autores que estudaram a área (ver Campanha, 1991 e 1995). Esta discordância não será tratada aqui. Nas áreas onde foi feita amostragem geoquímica de detalhe será adotada a nomenclatura usada no Projeto Integração e Detalhe Geológico no Vale do Ribeira. No capítulo de Geoprocessamento, onde se relatam as tentativas de análise metalogenética regional, utilizada a divisão estratigráfica de Campanha (1995).

O relevo em quase toda a área pré-cambriana é acidentado, chegando a altitudes de 1.300 m, com menores altitudes na baixada do Ribeira, a SE, e relevo menos acidentado, mas com altitudes de até 1.000 m, na área sedimentar a NW da região coberta pelo projeto mencionado.

O clima é caracterizado como mesotérmico, do tipo subtropical sem estação seca e com verão quente, com média anual das isotermas entre 16°C e 19°C, com maiores temperaturas e precipitações a SE, na Baixada do Ribeira.

A hidrografia da área estudada é dominada pela bacia do rio Ribeira de Iguape, não pertencentes a ela apenas pequenos cursos d'água próximos ao litoral, desaguando diretamente no mar, e a área da bacia sedimentar do Paraná pertencente à bacia hidrográfica do mesmo nome.

Nas áreas mais meridionais do Vale do Ribeira, e nas de maior altitude, a vegetação original era de floresta subcaducifólia subtropical com araucária, passando a floresta pluvial tropical na Baixada e nas áreas de menor altitude próximas aos rios principais. Grande parte da área foi desmatada, sendo ocupada por vegetação secundária e, em menor proporção, por pastagens e culturas. A área remanescente ainda representa a maior extensão contínua florestada dos estados de S. Paulo e Paraná, sendo sua conservação um problema ambiental mal resolvido.

Duas áreas foram tratadas em maior detalhe pelo autor nos trabalhos geoquímicos desde o PIDGVR (Macedo e Batolla, 1981; Ferreira et alii, 1981; Silva et alii, 1981; Macedo, Addas e Batolla, 1981; Macedo, Batolla e Addas, 1982) e em trabalhos posteriores (Macedo, 1986; Barbour, Macedo e Hypolito, 1988; Barbour e Macedo, 1991; Macedo, 1992, 1993*a* e *b*; Macedo, Braghin e Moreira, 1993; Macedo et alii, 1993*a* e *b*; Araújo, Macedo e Campanha, 1995; Salles, Macedo e Liotte, 1995). São as áreas denominadas Mina do Paqueiro e Mina do Perau, prolongando-se esta para a bacia do Ribeirão Grande, onde foi estudada a poluição causada pela mineração.

Para que sejam compreendidos os resultados geoquímicos aqui comentados, é apresentada uma descrição sucinta das áreas amostradas, sendo elas melhor descritas em Ferreira et alii (1981), Silva et alii (1981) e Macedo (1986). Os resultados da aplicação das diversas técnicas são apresentados nos capítulos seguintes.

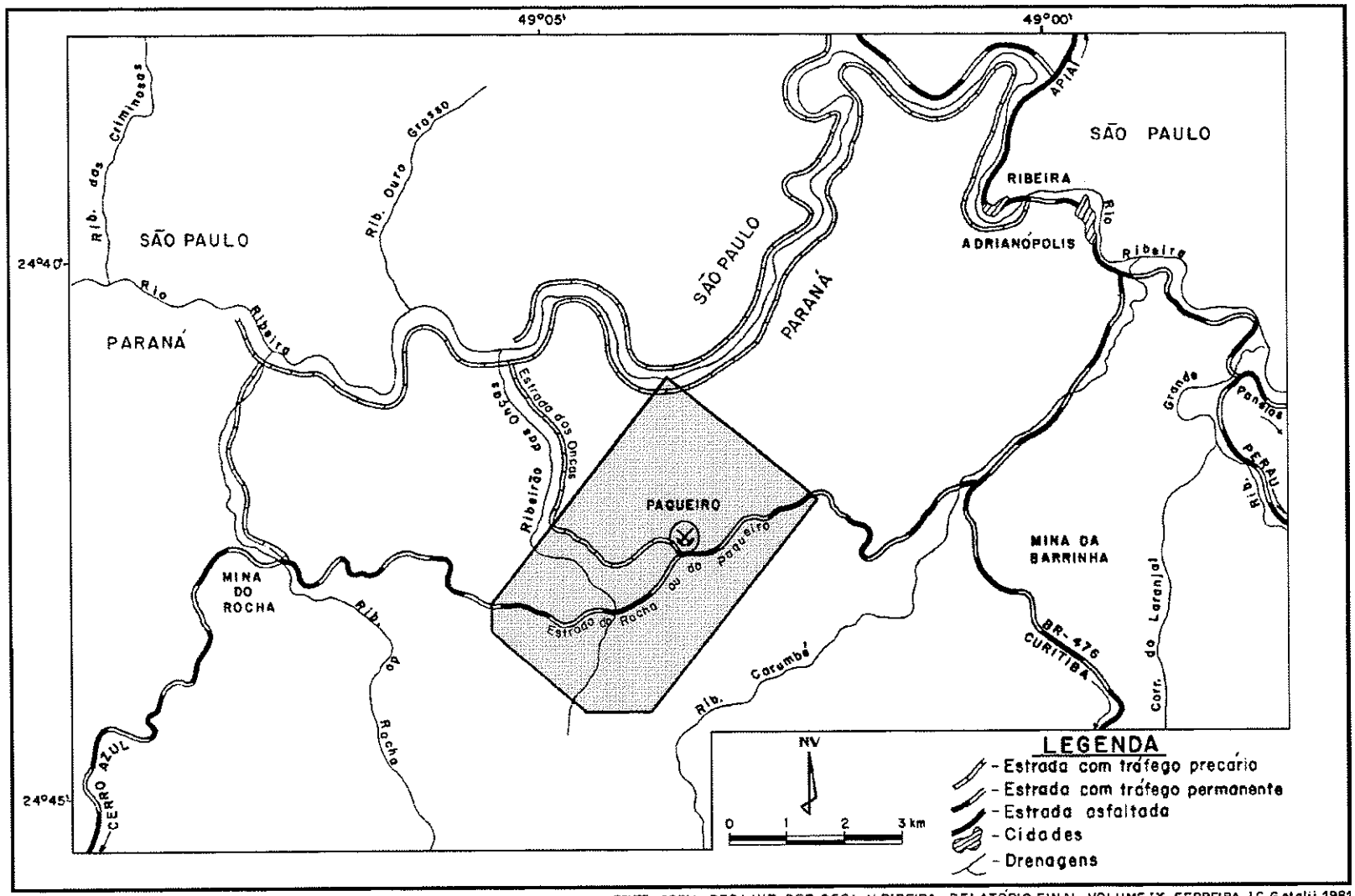
2.1.1 Área do Mina do Paqueiro

A área Mina do Paqueiro situa-se 14 km a sudoeste de Adrianópolis, PR. Nela ocorrem metassedimentos do precambriano médio a superior, terrígenos (xistos, meta-quartzo-arenitos e metaconglomerados) e carbonáticos (calcários magnesianos e quartzo-arenosos), intrudidos por diques de diabásio mesozóicos.

As rochas precambrianas estão dobradas e são cortadas por falhamentos segundo NE-SW a ENE-WSW, sendo os principais denominados Lineamento Bueno e Lineamento Paqueiro, que condicionam as mineralizações de chumbo. A figura 3 mostra a localização da área e os pontos geográficos mais importantes, e a figura 4 mostra a litologia e as estruturas principais.

O depósito foi lavrado entre 1953 e 1971, produzindo galena, que era vendida à Plumbum S/A. Segundo Ferreira et alii (1981) há uma contradição entre as informações dos relatórios anuais de lavra, que relatam uma produção total de 16.300 t de minério, com teor médio de 9,1% de Pb e reservas indicadas de 1.800 t de minério com 5% de Pb, e as de Damasceno (1967), que relata produção de 1.700 t de chumbo apenas entre 1957 e 1966 e uma reserva provada de 300 t e estimada entre 1.200 e 1.500 t de chumbo contido. O teor de Ag era admitido como de 250 g/t de Pb. Análises de Damasceno (1967) registram teores de Pb de 15,4 a 53,3 %, de Zn de 0,3 a 1,2 %, de Cu de 0,1 a 0.3% e de Ag de 240 a 960 g/t de chumbo contido. A composição mineralógica dos filões, segundo o mesmo autor, é de galena, pirita, calcopirita, esfalerita e arsenopirita, como primários; limonita, cerusita, piromorfita, covelita e malaquita, como minerais oxidados; ganga de quartzo e carbonatos. A associação foi considerada mesotermal. Análises do PIDGVR revelaram também bornita e hematita (Ferreira et alii, 1981).

Damasceno (1967) efetuou uma campanha de prospecção geoquímica, visando encontrar novos indícios de mineralização no prolongamento dos já conhecidos. Coletou 1.200 amostras de solo, em 9,5 km de perfis com direção N30W, perpendiculares à atitude média geral dos calcários, com espaçamento entre amostras de 10 m e entre perfis de 30 m, amostrando na profundidade de 0,8 m, com trado manual. As amostras foram analisadas por colorimetria, usando ditizona como indicador, tendo sido feita uma experiência com espectrografia de raios X, considerada positiva para indicação de anomalias, e vantajosa devido à maior rapidez.



FONTE: CPRM – PROJ. INT. DET. GEOL. V. RIBEIRA - RELATÓRIO FINAL - VOLUME IX - FERREIRA, J.C G. et alii, 1981

Figura 3 - Localização e vias de acesso da área Mina do Paqueiro.

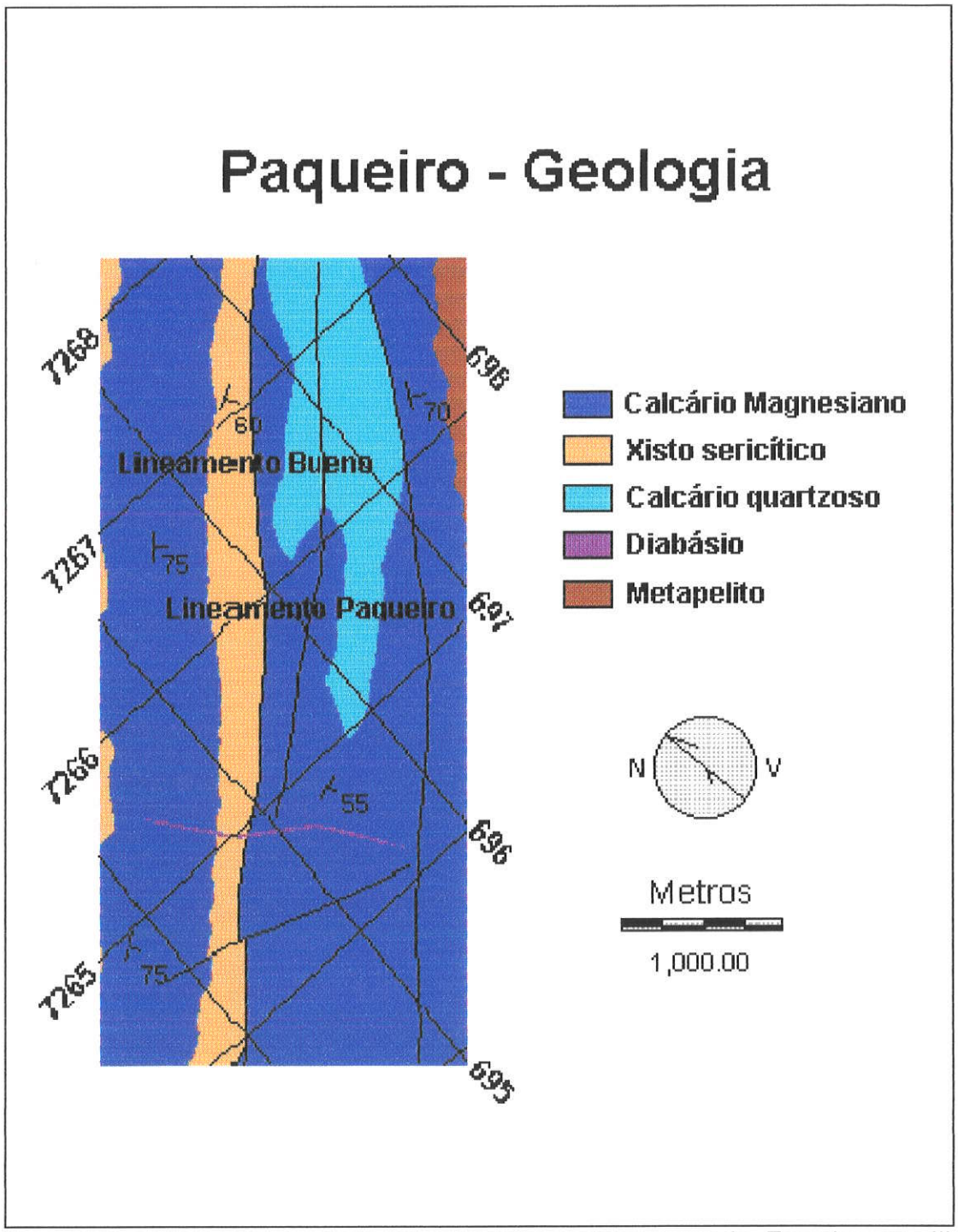


Figura 4 - Geologia da área Mina do Paqueiro, segundo Ferreira et alii (1981), simplificada.

A localização das mineralizações principais e das áreas amostradas no estudo orientativo e no levantamento de detalhe do PIDGVR é apresentada na figura 5.

Durante o Projeto IDGVR foi feito um levantamento geoquímico de solos, com a coleta de 1.009 amostras, precedida de um estudo orientativo. Os resultados do estudo orientativo são mostrados na figura 6. A mineralização se reflete claramente nos teores de chumbo, cobre e zinco das amostras coletadas nos três horizontes do solo. O horizonte A mostra efeitos da contaminação pela lavra e transporte de minério, refletida claramente nos gráficos teor-distância, que mostram aumento de teor em superfície próximo a um caminho e uma estrada, usados para transporte de minério.

A partir desse estudo, e da observação das anomalias nos perfis de Damasceno (1967), foi adotado, como valor preliminar, o limiar de anomalia de 100 ppm de Pb. Para detectar anomalias desse tipo, o espaçamento de 30 m entre amostras, em linhas traçadas a cada 100 m, foi considerado ideal. Com base na variabilidade detectada em furos amostrados de 30 em 30 cm, a amostragem foi feita na parte mais argilosa do horizonte B, que ficava geralmente entre 0,8 e 1,0 m. O horizonte A foi descartado pela possibilidade de contaminação, que, conforme o orientativo, é alta nas proximidades de estradas e caminhos utilizados para transporte de minério.

Nesta revisão foram elaborados variogramas em direções coincidentes com as linhas de amostragem e perpendiculares a elas, para testar a eficiência do espaçamento adotado no PIDGVR. Os variogramas dos teores logaritmizados dos três elementos mostram uma amplitude várias vezes superior ao espaçamento de amostragem, nas duas direções, mostrando serem os dados plenamente suficientes para o estudo do halo geoquímico em solos na região. Os gráficos e um melhor comentário são apresentados no item 7.5.

Os resultados analíticos da área Mina do Paqueiro, por serem em número elevado, e por apresentarem anomalias claramente associadas à geologia local, foram escolhidos para teste de metodologias de tratamento de dados, tanto não-espaciais quanto espaciais. Os trabalhos foram relatados em Macedo et alii

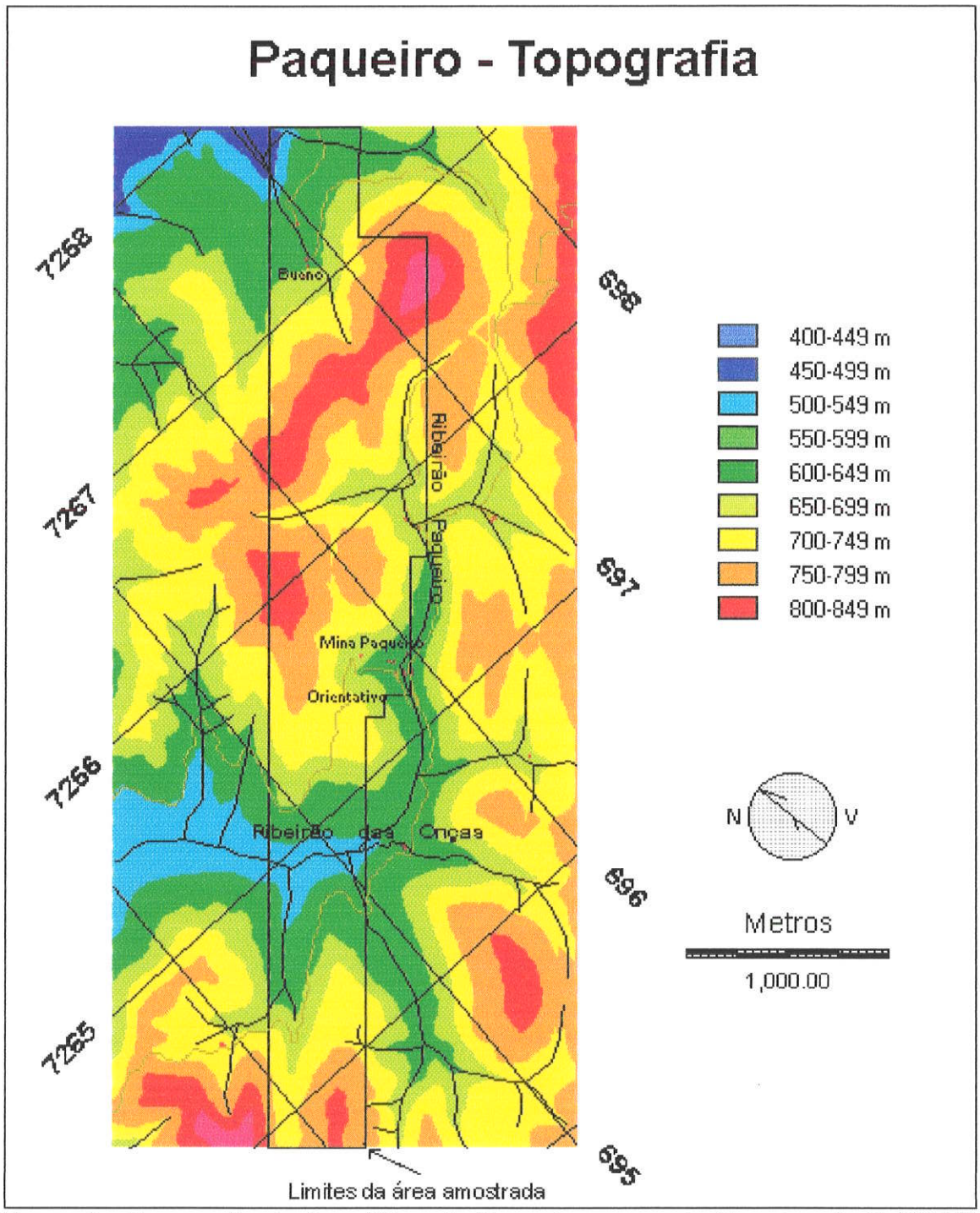
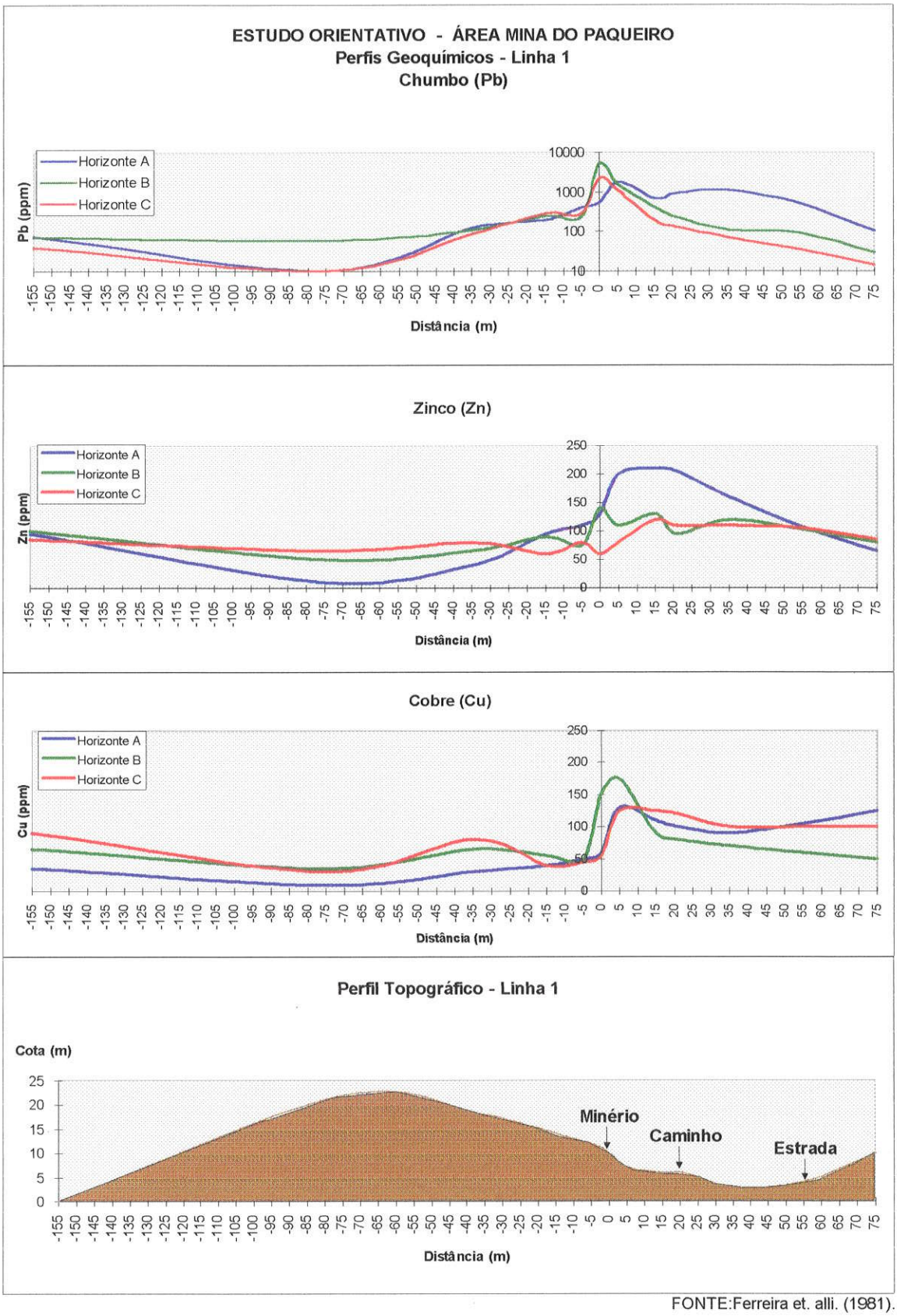


Figura 5 - Topografia da área Mina do Paqueiro, segundo Ferreira et alii (1981), simplificada.



FONTE:Ferreira et. alli. (1981).

Figura 6 - Resultados do estudo orientativo para prospecção geoquímica, área Mina do Paqueiro.

(1993*a*) e Salles et alii (1995) e seus resultados são apresentados nos ítens dedicados a cada técnica.

Juntamente com a prospecção geoquímica foi efetuada prospecção geofísica, visando verificar a aplicabilidade de medidas magnéticas para indicar mineralizações ou traços geológicos e estruturais na área. Foram efetuadas medições de intensidade do campo magnético total com magnetrômetro Geometrics G-826 e de susceptibilidade magnética, com kappameter ABEM KT-3. Os resultados de ambos os levantamentos não indicaram anomalias que pudessem ser associadas a mineralizações, níveis magnéticos ou estruturas associáveis à mineralização, sendo apenas marcante uma anomalia associada ao grande dique de diabásio que corta a parte sul da área trabalhada, ao longo do Ribeirão das Onças (Ferreira et alii, 1981).

2.1.2 Área Mina do Perau

A área Mina do Perau é importante por conter uma das maiores jazidas lavradas no vale do Ribeira, que é assim descrita em Macedo (1986):

> A jazida do Perau situa-se no município de Adrianópolis, 30 km a Sul da divisa São Paulo-Paraná. Consiste de galena, pirita, calcopirita e blenda como minerais de minério principais, com uma reserva inferida de 1,8 milhões de toneladas de minério, com 4% de Pb, 2% de Zn e 85 g/t de Ag, além de uma mineralização associada de 830.000 t de minério com 2% de cobre. O corpo de minério de chumbo está encaixado em rochas calciossilicáticas, estando estas assentadas sobre quartzitos e cobertas por xistos com intercalações de anfibolitos, compondo todo este conjunto a Formação Perau do Grupo Setuva, que se encontra sobreposta a xistos e gnaisses do embasamento pré-Setuva.

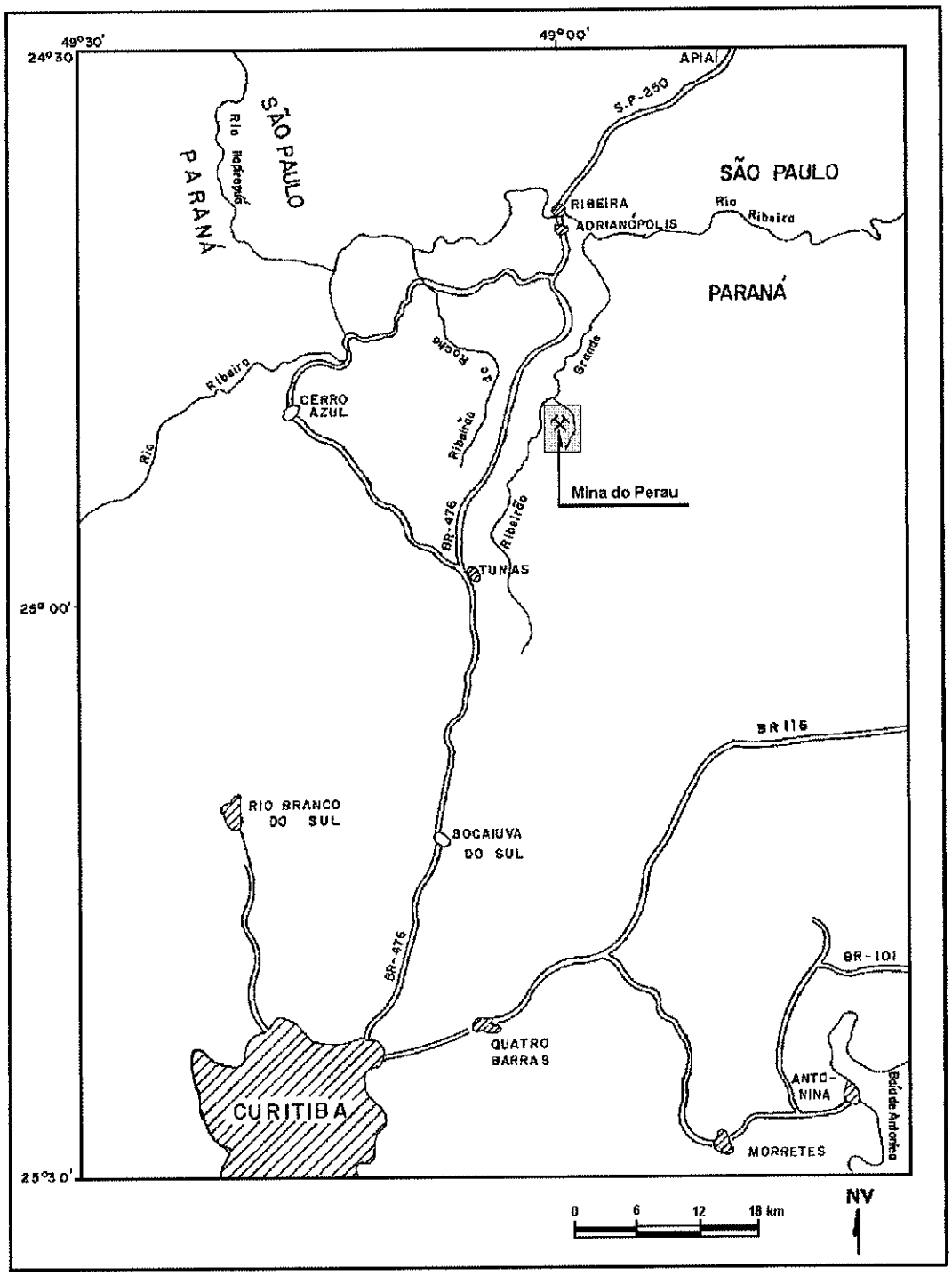
A localização da área é apresentada na figura 7, e um esboço geológico na figura 8. A área do Perau apresenta grande interesse, por sua complexidade geológica, e pela potencialidade mineral, alta para os padrões

do vale do Ribeira, e de tamanho médio para os padrões internacionais, Se lhe forem adicionadas as jazidas e ocorrências de Canoas, Água Clara, Tigre, Tirivas e Araçazeiro, próximas e geologicamente relacionadas a ela, configura-se uma província mineral que merece mais estudo.

Foram efetuados trabalhos de prospecção geoquímica de solos na área por Alves et alii (1977), de sedimentos de corrente e solos por Macedo, descritos em Silva et alii (1981), de rochas, por Macedo (1986), e de solos, pelo Projeto Anta Gorda, descritos em JICA/MMAJ (1984).

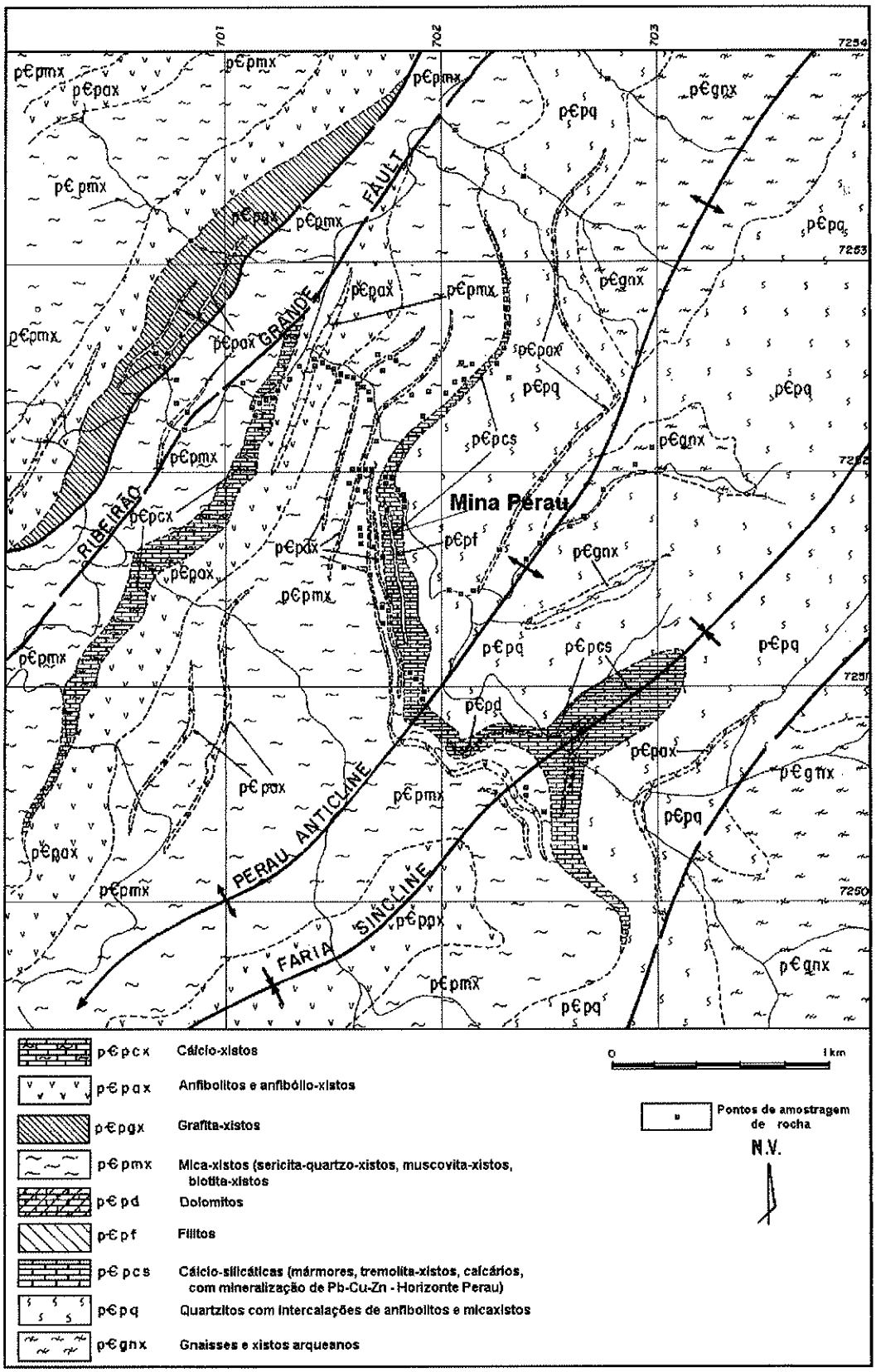
Recentemente aumentou o interesse pelos efeitos ambientais da mineração, sendo estes estudados na área tanto em relação aos impactos diretos nas proximidades da mina, empregando geoprocessamento (Macedo et alii, 1993 *a*), quanto aos impactos nos cursos d'água (Macedo, 1992 e 1993 *c*), estudados com amostragem de sedimentos de corrente.

Os métodos empregados no tratamento dos dados para essas duas áreas, tanto de cartografia geológica, quanto de prospecção e detecção de poluição serão comentados nos ítens próprios de cada fase de tratamento.



FONTE: MACEDO (1986).

Figura 7 - Localização da área Mina do Perau.



FONTE : MACEDO (1986), Prospecção Litogeoquímica na Área da Mina do Perau.

Figura 8 - Geologia da área Mina do Perau.

3. AQUISIÇÃO, ARMAZENAMENTO E RECUPERAÇÃO DE DADOS

3.1 Aquisição de dados

3.1.1 Generalidades

A aquisição de dados é uma etapa crucial no estudo geoquímico. Um erro nesta fase afeta todas as outras, e pode levar à anulação de todo o esforço. Já diziam os que trabalhavam em computadores na década de 60:

"GIGO", significando: **"Garbage In - Garbage Out"**.

Três tipos de dados devem ser adquiridos: de campo, de laboratório e secundários, podendo ser os de campo e secundários espaciais ou alfanuméricos.

O primeiro passo para aquisição de dados de campo, e em alguns casos, de laboratório e secundários, é a **amostragem**. Os dados geoquímicos quase sempre provêm de populações grandes, que cobrem extensas áreas de terreno e cujos elementos são em grande parte inacessíveis na prática, por ocorrerem em locais ínvios ou em profundidade. Isto obriga a um planejamento cuidadoso da amostragem, com escolha explícita do método e compreensão do que é realmente possível inferir em relação à população visada, considerando-se esta e a amostra estatística. Landim (1979) adverte para a necessidade de diferenciar entre a população *visada* e a população *amostrada*, de compreender como a população é constituída, os níveis de mensuração utilizados, e conseqüentemente, o que é realmente possível inferir. Krumbein e Graybill (1979) e Howarth (1983) apresentam bom tratamento do assunto.

Cabe aqui um esclarecimento quanto à nomenclatura. Há uma confusão, não resolvida, entre o uso do termo **amostra** pelos geoquímicos e geólogos, principalmente os de prospecção e aquele definido pelos estatísticos. Para os geoquímicos, **uma amostra** é uma quantidade de material coletada em campo, em um determinado meio (rocha, solo). Quando

analisadas as amostras, cada uma fornece um **teor**, expresso por um número na unidade mais conveniente (percentagem, ppm, ppb), o qual, juntado aos outros determinados no mesmo lote, ou nas amostras coletadas no mesmo meio, fornecem um **conjunto de teores**, ou **conjunto de dados**. Para um estatístico, a amostra é o conjunto de dados, chamando-se uma observação aquilo que o geólogo ou geoquímico chama de amostra. Este texto seguirá a nomenclatura usada nos trabalhos comentados, a mesma usada comumente pelos geoquímicos. Quando o termo amostra for utilizado na concepção dos estatísticos será qualificado como **amostra estatística.**

As amostras geoquímicas coletadas em quase todos os trabalhos efetuados pelo autor seguiram o esquema de amostragem sistemático, com espaçamento aritmético para levantamentos regionais e de detalhe, e espaçamento logarítmico para alguns dos estudos orientativos. Apenas em Macedo (1979) foi empregado um sistema misto, com amostragem casual aplicada a quadrículas de igual tamanho, para estabilizar a distribuição por área de uma amostra estatística pré-existente. A amostragem casual, ou uma combinação casual-sistemática, seria melhor que a sistemática simples. Não foi aplicada devido a problemas operacionais, pois exigiria um trabalho muito maior para locação dos pontos de amostragem, além do permissível com os recursos disponíveis para os trabalhos. Como conseqüência, os resultados de todos os trabalhos aqui comentados têm validade estatística limitada aos dados analisados, não podendo, a rigor, ser usados para a estimação de parâmetros populacionais. A extrapolação dos resultados além dos dados estudados (amostra estatística) é, como na maioria dos trabalhos geológicos, baseada na experiência pessoal e na de outros pesquisadores. Os resultados podem ter boa verossimilhança geológica, mas não têm validade estatística. Neste ponto é importante assinalar que quase sempre os conjuntos de dados analisados não apresentam distribuição normal ou lognormal, sendo quase sempre representativos de misturas de populações, o que obriga a tratamentos adequados, comentados adiante.

3.1.2 Procedimentos operacionais para coleta e conversão de dados

Algumas regras básicas se aplicam a toda a etapa de aquisição de dados:

- as operações devem ser padronizadas, reduzindo o efeito da variação inter-equipes;
- elas devem ser, no máximo possível, automatizadas, pois a possibilidade de erro humano é sempre muito alta;
- todo o procedimento deve ser documentado, de forma a permitir verificação posterior, por operadores diferentes dos originais;
- todos os dados devem ser testados em relação a coerência, exatidão e precisão, antes de sua incorporação definitiva ao banco de dados.

O autor tem procurado seguir estes princípios, descrevendo em Macedo e Batolla (1981) e Macedo et alii (1981, 1982) as técnicas de coleta, registro e crítica de dados empregados nos trabalhos geoquímicos do Projeto Integração e Detalhe Geológico no Vale do Ribeira, as quais, com pequena atualização, foram utilizadas também nos trabalhos posteriores.

Em resumo, os dados de campo foram registrados utilizando os formulários padronizados da CPRM, com preenchimento de todos os campos compatíveis com o meio de amostragem utilizado. A locação de pontos no campo, no caso de amostragem de solo, foi feita com levantamento plani-altimétrico expedito, a trena e bússola, controlado por perímetro levantado a teodolito. A marcação dos pontos amostrados no campo foi feita com estacas e plaquetas de alumínio, nas amostragens de solo. Para sedimentos de corrente e rochas, foi feita locação em mapas ou fotos de escala mínima 1:25.000 e marcação com tinta spray em pontos visíveis do terreno.

O custo dos levantamentos topográficos e a locação exata é justificável por terem sido todos os trabalhos com finalidade acadêmica ou de levantamento básico governamental, o que obriga a que tenham boa documentação, para que sejam utilizáveis por outros pesquisadores. Para levantamentos privados com finalidades de prospecção bastaria levantamento a passo e bússola, desde que houvesse boa marcação dos locais para retorno às possíveis áreas anômalas. A locação de pontos de coleta de sedimentos

em áreas ínvias, como partes do vale do Ribeira, tem sido um problema, pois às vezes torna-se difícil distinguir as drenagens nos mapas e fotos aéreas. Atualmente isto pode ser sanado com sistemas de posicionamento global (GPS), que permitem locação em relação à posição de satélites geoestacionários. Outra melhoria na coleta é a utilização de medidores portáteis de pH e Eh, pequenos e baratos, substituindo, com maior precisão, as fitas de papel utilizadas nos antigos levantamentos.

A aquisição de dados em laboratório deve ser automatizada, sem uma etapa de cópia manual. Esta freqüentemente causa erros. Em quase todos os trabalhos do autor, apareceram nos boletins de análise teores altos isolados, que obrigaram à reanálise daquelas amostras. Os teores altos quase sempre não foram confirmados pela reanálise, indicando erro de leitura do instrumento ou de cópia.

A aquisição de dados alfanuméricos secundários pode ser automatizada, por transferência de arquivos digitais. Isto é feito no projeto da folha Itararé, em andamento, que usa dados do SIGA (Sistema de Informações Geoquímicas, da CPRM), da Mineropar e do IDEM (Índice de Depósitos Minerais, produzido pelo IG-UNICAMp para o Pró-Minério). Também pode-se recorrer ao escaneamento de páginas e reconhecimento dos textos por programas de reconhecimento de caracteres (OCR), já utilizado pelo autor para montagem de exercícios didáticos de interpretação geoquímica. Em ambos os casos evita-se o erro humano, o qual ainda é possível no uso de OCR, que, conforme o tipo de letra e qualidade do original pode obrigar à intervenção do operador e depender de sua interpretação.

Os dados gráficos podem ser digitalizados manualmente, por mesa digitalizadora ou por mouse em tela; ou automaticamente, utilizando programas vetorizadores, que transformam imagens matriciais obtidas por escaneamento em arquivos vetorizados, com ou sem auxílio de operador. Algumas experiências do autor são narradas em Macedo (1993), Araújo, Macedo e Campanha (1995) e Liotte et alii (1995).

Os dados do projeto Itararé e dos trabalhos acima listados foram em parte digitalizados manualmente, em parte vetorizados após escaneados a

partir de overlays unitemáticos em película indeformável Mylar (técnica empregada no United States Geological Survey - USGS (Soller et alii, 1990).

Na revisão da geologia da folha Itararé (Macedo et alii, 1993) houve muitos problemas, pois foi tentada a combinação de três operações ao mesmo tempo: digitalização, combinação de mapas diferentes e revisão conceitual da geologia. Isto levou a um trabalho lento e impreciso, que necessitou sucessivas revisões.

Em resumo, juntando a experiência do autor e o que observou no USGS e na CPRM, as técnicas mais recomendáveis para digitalização são:

- preparar os originais de forma a não haver problemas de interpretação durante a digitalização. Se os originais não forem de boa qualidade, ou se quisermos juntar dados de originais diferentes, é necessário fazer um overlay em película indeformável (Mylar, Cronaflex), com caneta nanquim com traço da menor largura possível (00 ou menor). Originais de boa qualidade podem ser simplesmente marcados, indicando pontos de controle de coordenadas, numeração dos polígonos, etc.
- transferir os dados por scanner ou mesa digitalizadora. Neste último caso editar à medida em que é feita a digitalização. Para dados escaneados melhores resultados são obtidos por procedimento semi-automático, com a atuação do operador reduzindo o trabalho posterior de edição.
- editar os arquivos vetoriais, fazer pré-processamento para ajuste de coordenadas e plotar em película transparente para conferência.

Em todos estes passos, a pressa e a fadiga devem ser evitadas, e cada fase deve ser exaustivamente conferida antes de se passar à seguinte. Se a equipe não tiver prática, é muito recomendável a execução de um trabalho piloto, passando-se por todas as fases em uma parcela pequena dos dados.

3.2 Armazenamento e recuperação

O armazenamento de dados deve ser feito de maneira a manter sua integridade, evitando tanto a sua destruição por meios físicos quando por erro humano. Além disso o armazenamento deve permitir rápida recuperação e comunicação dos dados com outros programas.

Após uma verificação de exatidão da anotação dos dados, principalmente dos que passaram por uma etapa manual de cópia, e a sua crítica para verificar a adequação aos objetivos do projeto e da instituição, devem ser montados bancos de dados. Os gerenciadores de bancos de dados mais simples e mais gerais dão melhores resultados que aqueles elaborados especialmente, cujos programas e procedimentos não podem ser transferidos, e, às vezes, nem mesmo os dados podem ser aproveitados em outros projetos. Isto acontece também com alguns programas comerciais. Um bom gerenciador de banco de dados comercial (Dbase ou Access, por exemplo) permite manter e explorar os dados a custo baixo e alta eficiência. Para pequenos conjuntos uma planilha eletrônica pode ser utilizada. Os sistemas de geoprocessamento, comentados adiante, permitem combinar os bancos de dados tradicionais, de atributos, com dados espaciais, facilitando e aumentando o poder de análise dos dados. Tanto para dados espaciais quanto não-espaciais a versatilidade do banco de dados é fundamental. É imprescindível que os dados possam ser lidos e utilizados por outros programas, e não apenas por uma "caixa-preta" que usa um formato próprio incomunicável.

É importante manter um banco geral, cuja integridade dos dados seja garantida por proteção contra modificações, e só trabalhar conjuntos copiados dele, montando bancos auxiliares temporários, por projeto ou tarefa.

4 CRÍTICA DE DADOS

O conjunto de teores obtidos em uma campanha geoquímica apresenta erros provindos de duas fontes principais:

- *de amostragem*: que combinam os efeitos da variabilidade natural dentro do raio de influência dos pontos amostrados e os erros introduzidos pela técnica de amostragem;
- *de análise*: derivados do equipamento e das técnicas de preparação e análise química.

O autor comentou as técnicas de detecção e controle desses erros em Macedo e Batolla (1981), Macedo et alii (1981, 1982) e Macedo (1985, 1993*b*).

Para o controle dos erros analíticos a melhor opção é incluir em cada lote uma réplica provinda de uma amostra homogeneizada, o que correspondeu, nas amostras enviadas ao LAMIN-CPRM, a 4% das alíquotas analisadas. O coeficiente de variação para os logaritmos dos teores em ppm das réplicas de controle do Projeto Integração e Detalhe Geológico no Vale do Ribeira, analisadas por absorção atômica no LAMIN-CPRM, foi, segundo Macedo e Batolla (1981) de 2,84% para Cu, 2,82% para Pb e 4,43% para Zn, considerados aceitáveis para prospecção geoquímica (onde comumente é aceita variação de até 10%).

O controle dos erros, nos casos em que foram disponíveis réplicas e duplicatas, foi feito segundo o método de Garrett (1969, 1973).

Para esse método são comparadas duas séries de amostras por uma análise de variância modificada. No Projeto Integração e Detalhe Geológico no Vale do Ribeira (Macedo e Batolla, 1981) uma das séries era constituída de amostras coletadas durante as amostragens normais, a outra de amostras de controle coletadas no final da amostragem em cada área de amostragem, em pontos situados a um metro do ponto de coleta original. As diferenças de teores entre as amostras normais e as duplicatas são devidas conjuntamente à amostragem e à análise. Os resultados para as amostras coletadas na área Mina do Paqueiro estão apresentados na tabela 1.

Tabela 1: *Resultados de análises por absorção atômica de amostras normais e duplicatas coletadas na área Mina do Paqueiro, no projeto Integração e Detalhe Geológico no Vale do Ribeira (Macedo e Batolla, 1981).* Núm: número de campo; CuN: teores de Cu em amostras normais; CuD: teores de cobre em amostras duplicatas.

AMOSTRAS NORMAIS				AMOSTRAS DUPLICATAS			
Núm	CuN	PbN	ZnN	Núm	CuD	PbD	ZnD
591	35	45	26	1480	80	70	75
622	55	75	60	1481	65	90	80
651	50	70	50	1482	60	80	75
781	90	290	70	1483	100	180	85
701	75	55	60	1484	95	50	65
726	110	35	40	1485	120	50	50
756	70	70	45	1486	120	80	80
776	16	8	12	1487	20	16	18
802	40	16	60	1488	65	20	80
849	50	85	45	1489	50	90	50
821	27	30	45	1490	30	35	55
892	50	100	80	1491	40	90	75
933	50	60	40	1492	35	45	35
960	14	70	30	1493	20	95	50
1018	23	40	35	1494	40	30	55
1068	10	14	8	1495	12	20	10
1112	28	40	40	1496	35	50	50
1172	35	55	30	1497	55	80	55
1226	40	380	130	1498	50	460	180
1305	30	55	30	1499	35	80	50
1310	45	500	55	1500	55	400	70
1367	40	27	45	1501	45	35	65
1368	40	18	40	1502	40	20	50
1470	40	55	55	1503	40	40	65
1427	10	20	16	1504	16	26	27
1447	24	35	70	1505	26	55	60

A variância combinada de amostragem e análise é calculada pela fórmula:

$$\sigma^2_{sa} = \frac{1}{n}\sum_{i=1}^{n}\frac{(x_{1i} - x_{2i})^2}{2}, \text{ onde}$$

σ^2_{sa} = variância combinada de amostragem e análise

n = número de pares comparados

x_{1i} = logaritmo do teor para amostra de rotina

x_{2i} = logaritmo do teor para amostra duplicada

A variância total é calculada pelo método normal:

$$\sigma^2 = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2}{n-1}$$

onde:

σ^2 = variância das amostras normais

n = número de amostras

x_i = logaritmo do teor para amostra de rotina

$\bar{x}$ = média dos logaritmos das amostras de rotina

n = número de pares comparados

A comparação é feita pela divisão:

$$F = \frac{\sigma^2}{\sigma^2_{sa}}$$

No caso relatado, os valores de F são, para os seguintes elementos:

Cu = 6,63

Pb = 18,03

Zn = 4,16

O valor crítico de F a ser comparado é, segundo Garrett (1973), o tabelado para a estatística F de Fisher-Snedecor, para N - 1 e N graus de liberdade, o que dá $F_{24,\ 25,\ 0.05}$ = 1,97. Por este critério os dados podem ser considerados perfeitamente aceitáveis para prospecção geoquímica.

Outro processo empregado, para comparar réplicas de laboratório, é o método gráfico apresentado em Stanton (1966).

Por este método foram comparados dados do estudo orientativo e da amostragem regional do levantamento da poluição causada por mineração

pelas minas do Perau, Barrinha e Canoas no Ribeirão Grande (Macedo, 1993b). Plotando-se num gráfico os teores das amostras coletadas no estudo orientativo nas abscissas, e os teores das mesmas amostras, quando analisadas juntamente com as amostras coletadas na amostragem regional nas ordenadas, revela-se facilmente uma diferença sistemática entre os teores dosados nos dois lotes de amostras, o que impossibilitou seu emprego em conjunto. Apenas os teores das amostras analisadas no lote do estudo regional foram utilizadas nos tratamentos posteriores.

Um exemplo dos gráficos é apresentado na figura 9. Os dados são lançados em escala logarítmica. São traçadas linhas de referência para X = Y (onde os pontos cairiam se os teores comparados fossem idealmente iguais) e linhas de X + 25% e Y + 25%. Considera-se aceitável a série em que um número de amostras menor que o equivalente ao nível de significância (por exemplo 5%) cai fora das linhas de referência, ou seja, menos de 5% das amostras tem variação maior que 25%. Conforme o tipo de análise e a finalidade do estudo os limites de variação e o nível variam. Os aqui aplicados são geralmente utilizados para prospecção geoquímica com análise por espectrometria de absorção atômica.

Confirmando as conclusões do teste pelo método de Garret (1973), as amostras coletadas em duplicata na área do Paqueiro passam pelo teste gráfico de Stanton (1966); apenas uma das 25 amostras cai fora dos limites aceitáveis de 25%, o que pode ser observado na figura 10.

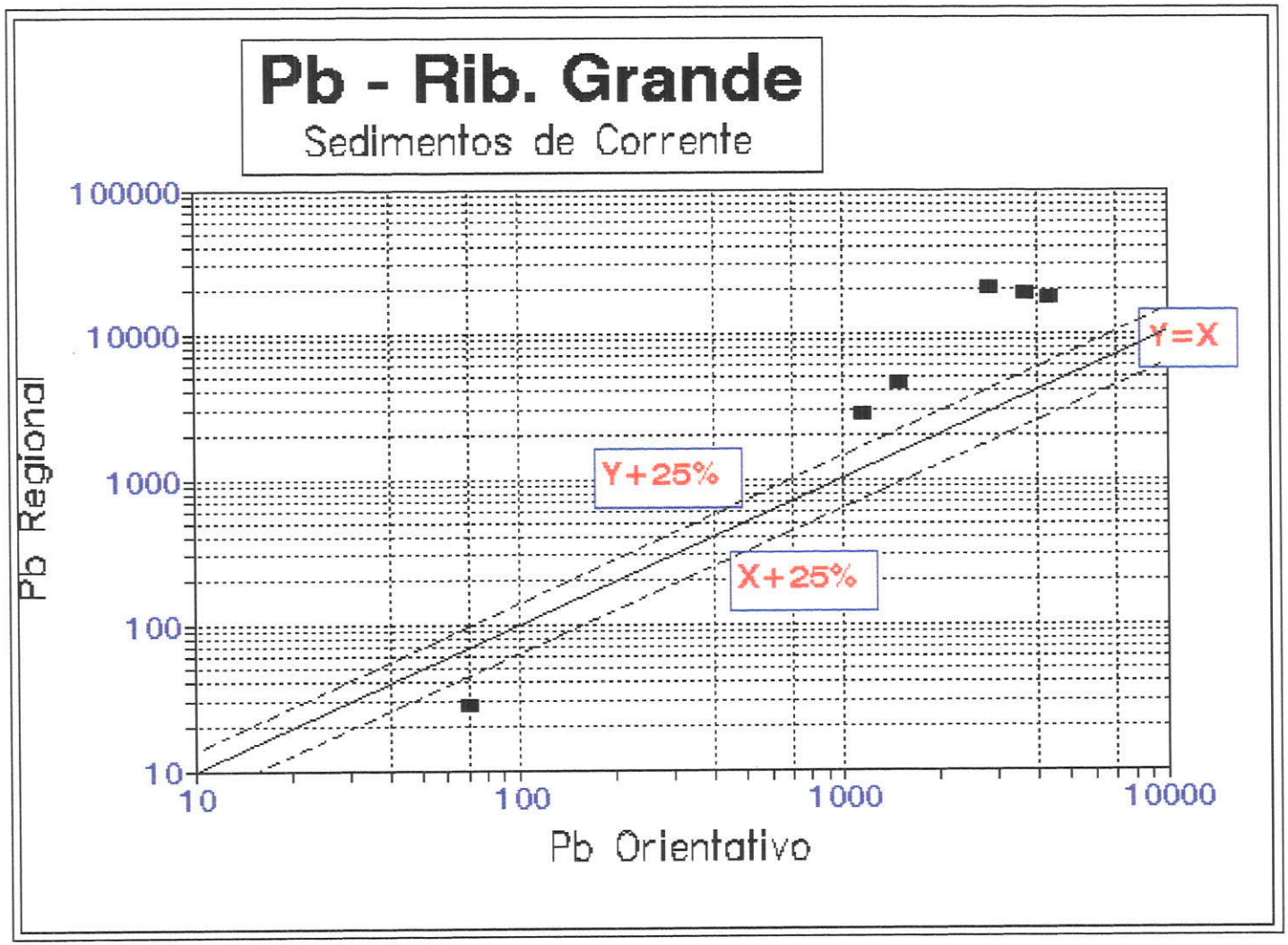


Figura 9 - Diagrama de dispersão lognormal, comparando teores de chumbo em amostras de sedimentos de corrente, coletadas no Ribeirão Grande para estudo orientativo, analisadas duas vezes, nos lotes de laboratório do orientativo e do regional.

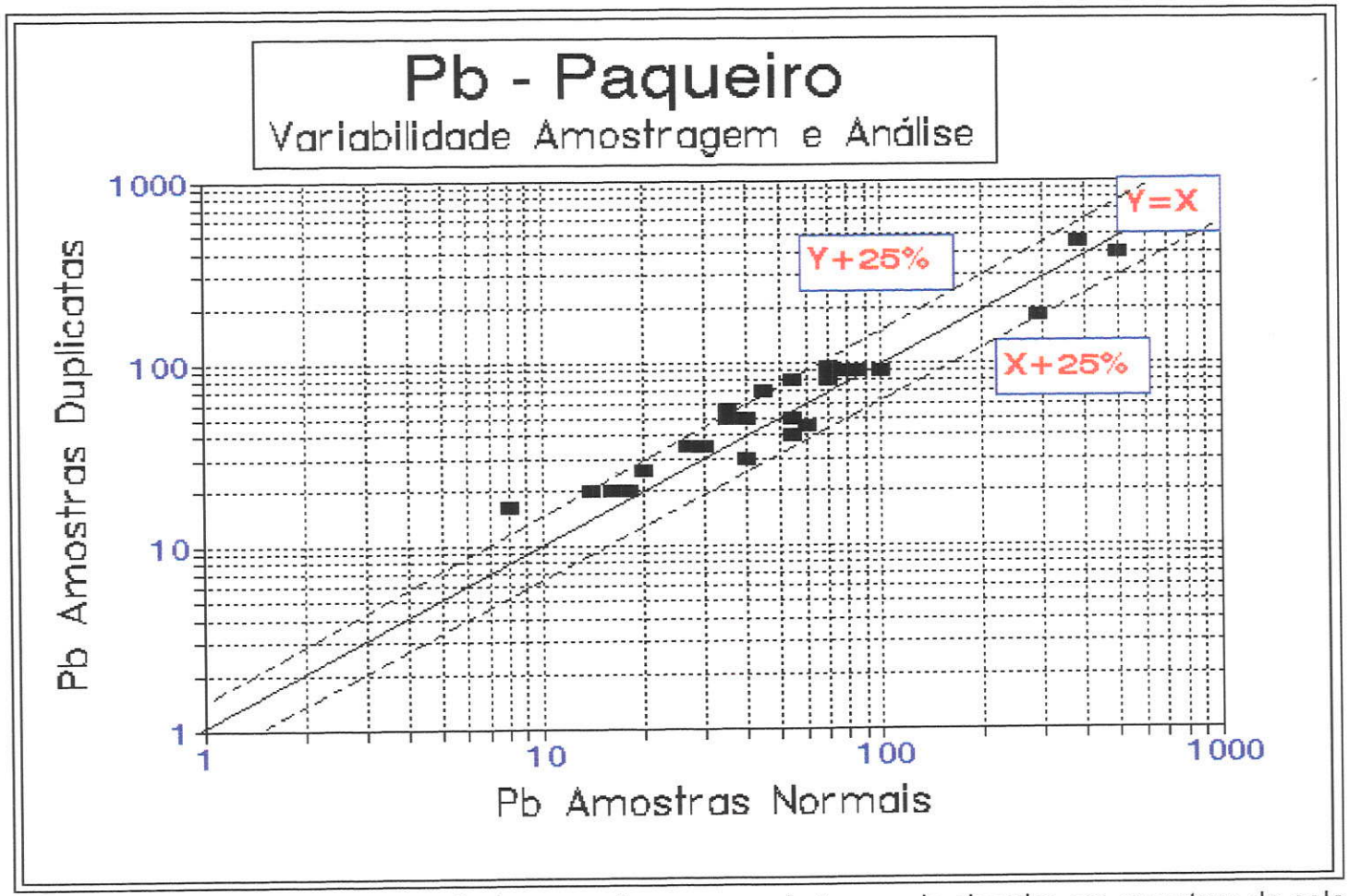


Figura 10 - Diagrama de dispersão lognormal, comparando teores de chumbo em amostras de solo normais e duplicatas, coletadas na área Mina do Paqueiro, para determinação da variância conjunta de amostragem e análise.

5. ANÁLISE DE DADOS UNIVARIÁVEL

Neste capítulo serão comentados os métodos que não levam em conta a distribuição espacial das amostras. A variação espacial será comentada no capítulo 7.

A análise univariável de dados geoquímicos visa:

- verificar sua distribuição de freqüência, determinando seus parâmetros, no caso em que esta distribuição pode ser aproximada de uma função conhecida;
- determinar a ocorrência de anomalias de distribuição, que podem ser atribuídas a anomalias geoquímicas.

Os métodos mais simples, e que devem ser usados em primeiro lugar, não assumem nenhuma hipótese de distribuição de probabilidade padrão: contagem, separação de dados por valores, determinação de máximo, mínimo e amplitude de variação; separação de dados por intervalos e sua representação, por diagrama de barras, histogramas, curvas acumulativas, diagrama de ramos e folhas (Suslick et alii, 1984).

Os testes de aderência a distribuições de freqüência conhecidas empregados são os do Qui-quadrado (que utiliza as somas das diferenças entre as freqüências esperadas, na distribuição padrão testada e as freqüências calculadas para o conjunto de dados estudado, para os mesmos intervalos) e o de Kolmogorov-Smirnof (que utiliza a máxima diferença entre freqüências acumuladas e esperadas). Infelizmente, na maioria dos casos, não são muito úteis para o trabalho em interpretação de dados geoquímicos estudados com finalidades de prospecção, a não ser como um primeiro passo, para determinar que o conjunto estudado não se encaixa em uma distribuição conhecida. Em grande parte dos casos, os conjuntos de dados estudados não dão resultados que permitam enquadrá-los nas distribuições normal ou lognormal. Por que é assim ?

Os teores de um elemento traço, determinados em uma campanha de amostragem geoquímica resultam de fenômenos múltiplos e complexos, que atuam sucessivamente no material a ser amostrado.

Para as amostras de rocha, influem inicialmente as condições de formação da rocha: cristalização e diferenciação do magma, em rochas ígneas, intemperismo, transporte e sedimentação nas sedimentares e de metamorfismo, nas metamórficas. Posteriormente as rochas podem sofrer mudanças de pressão e temperatura, introdução e remoção de elementos e outras ações de alteração, diagênese e metamorfismo. Quando uma rocha é submetida às condições da superfície, ocorre o intemperismo, e é quase impossível coletar uma amostra de rocha em superfície em clima tropical que não seja alterada.

Para as amostras de solo, além da rocha-mãe, influem os fatores de formação e alteração do solo; para sedimentos de corrente as condições de erosão, seleção, transporte e deposição dos sedimentos. Além dessa variabililidade natural, influi também aquela devida à manipulação da amostra (amostragem, preparação e análise).

Mesmo com esta enorme quantidade de fatores, aparecem regularidades na distribuição estatística dos teores. Já há muito tempo foi notado que, em rochas não mineralizadas, os elementos principais tendem a apresentar distribuição normal, os elementos traços distribuição lognormal. O mesmo acontece com sedimentos e solos, desde que tenham condições genéticas homogêneas e, principalmente, se as amostras coletadas provêm de material derivado de uma mesma rocha original.

A mineralização tende a introduzir uma distorção na distribuição. Quando outros fatores não afetam demasiadamente os teores, é possível distinguir estatisticamente, na distribuição destes, duas distribuições componentes: uma representativa do material original, sem mineralização, outra representativa do material introduzido pelos processos mineralizantes. Em outros casos é possível distinguir amostras provindas de populações diferenciadas pela litologia. Nestes casos é necessário separar as amostras em grupos coletados em áreas homogêneas, antes de outros tratamentos.

O método mais empregado para teste e separação de amostras provenientes de populações estatísticas múltiplas é o do gráfico de probabilidade, no qual são lançadas no eixo das abscissas as freqüências acumuladas e no das ordenadas os teores (para teste de normalidade) ou

seus logaritmos (para teste de lognormalidade). Se os teores seguirem a distribuição de freqüência esperada, os pontos ficarão em linha aproximadamente reta. Caso a distribuição não seja a esperada, os pontos formarão curvas, que podem ser interpretadas assumindo que resultam da mistura de teores provindos de duas ou mais populações. Este tipo de gráfico foi usado em Geoquímica pela primeira vez por Razumovskii, em 1940 (*apud* Vistelius, 1960). A separação das populações foi tratada em Williams (1967), que apresenta um processo de cálculo matematicamente rigoroso e diversos exemplos geológicos, mas, seguindo o destino das publicações do Hemisfério Sul, não é citado quase nunca. Lepeltier (1969) apresenta um processo inexato, partindo de um histograma e considerando a distribuição de pontos como retas truncadas. Até hoje é seguido, por ter publicado num periódico de prestígio, Economic Geology. Sinclair (1974,1976) plota todos os valores e freqüências acumuladas e usa limiares fixados nos percentis 1 e 99 das diversas populações. Bridges e McCammon (1980) e Stanley (1987) apresentam programas de computador que automatizam a comparação da curva dos teores com curvas obtidas pela combinação de populações teóricas.

Em Macedo (1979) foram empregados gráficos de probabilidade normal e lognormal para amostras de rochas basálticas da Bacia do Paraná, que permitiram mostrar que elas provinham de diferentes populações estatísticas,

> "que indicam variações regionais devidas à extensão e complexidade dos fenômenos magmáticos que geraram as rochas magmáticas da província".

Em Macedo e Batolla (1981) é descrita a técnica de cálculo empregada pelo autor para amostras de solo no Projeto Integração e Detalhe Geológico no Vale do Ribeira e projetos posteriores da CPRM, uma adaptação da de Williams (1967). Os dados são divididos em conjuntos provavelmente derivados da mesma litologia; os teores são plotados em papel de probabilidade, assumindo-se uma distribuição lognormal, resultando para os teores de chumbo coletados sobre calcários (709 amostras), com corte dos teores em 1000 ppm, o gráfico apresentado na figura 11. Citando esse trabalho,

"observa-se que os pontos se organizam numa linha que pode ser interpretada como a junção de duas curvas, uma delas com um ramo maior aproximadamente reto, constituindo o conjunto uma curva complexa com ponto de inflexão situado aproximadamente no local onde a linha corta a marca de 90%. Isto indicaria, segundo Williams (1967) e outros, uma mistura de duas populações, uma abrangendo 90% dos dados, com valores menores, e outra, abrangendo 10%, com valores maiores. Considerando que a curva estudada deriva de uma área sustentada por um único tipo litológico, na qual ocorrem mineralizações conhecidas de chumbo, considerou-se a população dos valores menores como representante da dispersão normal de chumbo nos solos da área (*background*) e a população dos valores maiores como indicativa do halo formado pela mineralização no solo de suas proximidades. Para determinação dos parâmetros da população de *background* foi utilizado o processo simplificado, tomando-se uma reta aproximadamente paralela ao ramo maior do gráfico, passando pelo ponto de cruzamento dos valores 49 ppm - 55%, ponto este obtido lançando o valor calculado da freqüência do valor 49 ppm na população de background isolada = 50 x 100 / 90. A partir desta reta obtivemos os valores:

DG *(desvio geométrico)* = 1,85

MG *(média geométrica)* = 46 ppm

MG.DG = 86 ppm

$MD.DG^2$ = 156 ppm

$MG.DG^3$ = 290 ppm

Estes valores resultaram num DG médio de 1,85. Empregando este valor e arredondando, chegamos aos valores definitivos:

MG = 46 ppm

DG = 1,85

MG.DG = 85 ppm

$MD.DG^2$ = 157 ppm

$MG.DG^3$ = 291 ppm.

Se tivéssemos utilizado os valores obtidos por cálculo numérico, que são representativos de mistura das duas populações, teríamos obtido valores de

MG.DG = 119 ppm; MG.DG^2 = 277,61 ppm. Isto nos levaria a fixar como limiar um valor alto demais, que não concorda com resultados obtidos durante o estudo orientativo realizado na área, nem com trabalhos anteriores (Damasceno, 1967). Desta maneira foram adotados os valores obtidos pelo processo gráfico para os limites dos intervalos do mapa de isoteores nessa área.

Para determinação dos valores da MG foi tomado o teor referente à percentagem acumulada de 50%, para MG.DG o de 84,13%, MG.DG^2 o de 97,72% e para MG.DG^3 o de 99,87%, sendo estes valores das probabilidades cumulativas (área sob a curva normal) para Z + 0, Z + 1, +2 e +3.

Para esta revisão foi feita a comparação entre métodos de cálculo estatístico de separação de populações e cálculo de anomalias, utilizando-se para tanto o conjunto total dos teores de chumbo em amostras de solo da área Mina do Paqueiro (1.009 amostras).

Calculando-se com o programa PROBPLOT (Stanley, 1987), para o total dos dados (1.009 amostras), sem divisão de populações, resulta uma média geométrica de 47,32 ppm e um desvio geométrico de 2,45, resultando num produto MG.DG = 115.93 ppm e um MG.DG^2 = 284.04 ppm.

Separando-se as populações usando o mesmo programa, com escolha visual do ponto de inflexão em 92%, obtêm-se duas populações, com os parâmetros:

	Background	**Anômala**
MG	40,05 ppm	300,26 ppm
DG	2,01	1,87

que resultam num limiar de anomalia (MG.DG^2 da população de *background*), de 161,80 ppm.

O gráfico desenhado por este programa é apresentado na Figura 12. Notar que os círculos assinalam os pontos originais teor x freqüência, as linhas retas inclinadas representam os gráficos teor x freqüência das populações lognormais componentes e a linha sinuosa o gráfico que resulta da mistura das populações componentes. A eficiência dos parâmetros escolhidos

é testada pelo ajuste entre os dados brutos e os resultados, refletido na maneira que a curva sinuosa passa pelos pontos representantes dos dados brutos.

Escolhendo a percentagem de mistura automaticamente pelo programa, usando um algoritmo de ajuste de máxima verossimilhança, ela foi calculada em 92.9 e os parâmetros ficam:

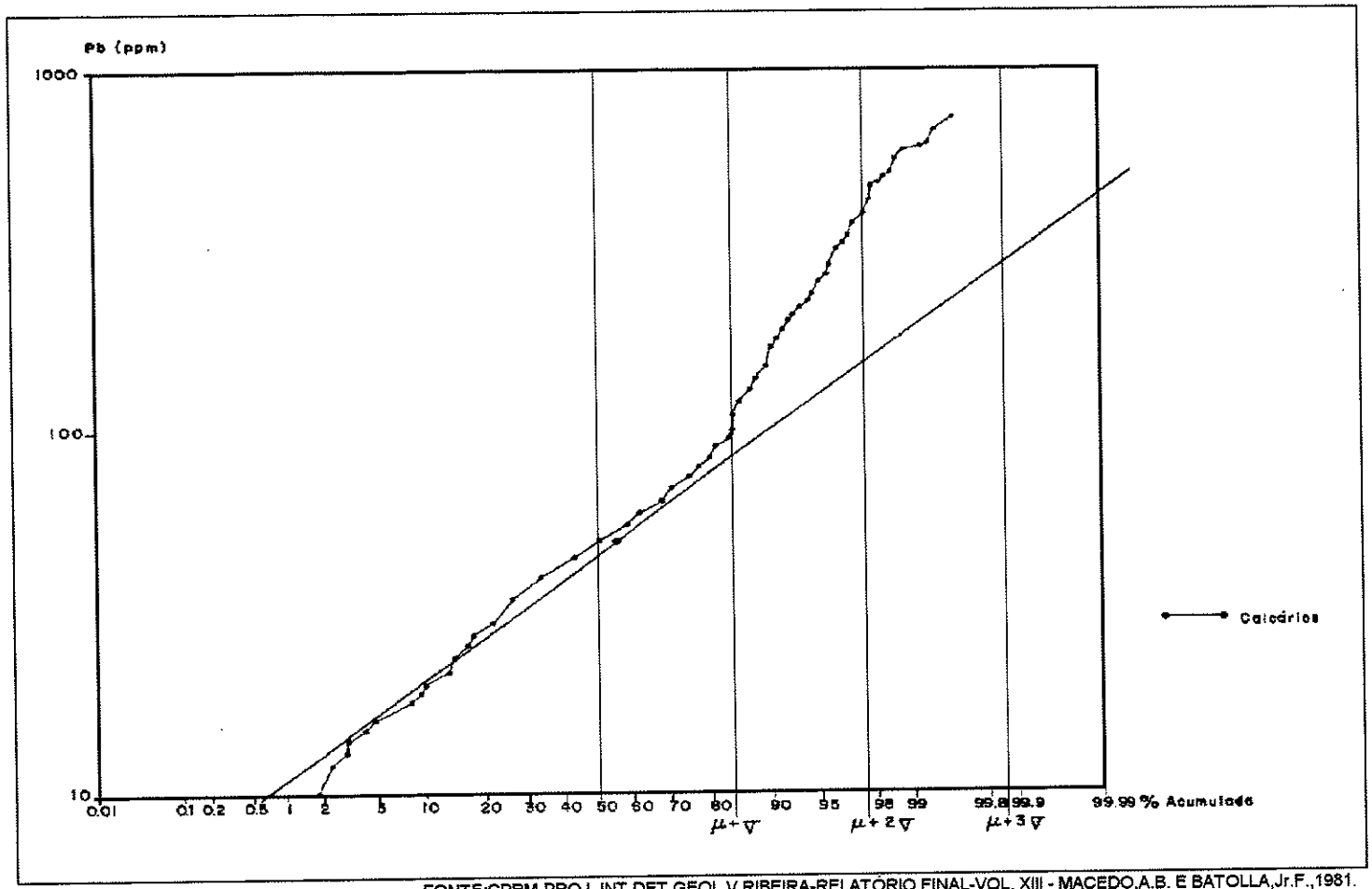
	Background	**Anômala**
MG	41,13 ppm	283,07 ppm
DG	2,09	1,98

que resultam num limiar de anomalia ($MG.DG^2$), de 179,66. Escolhendo os limiares pelo programa (valores de MG/DG^2 e $MG.DG^2$), a população dos teores menores teria limiares inferior e superior de 9,37 e 180,50 (ligeiramente diferente do anterior pelo arredondamento) e a dos teores maiores de 71,94 e 1113,78.. Neste caso, os valores entre 72 e 367 ppm (aproximadamente 20% dos dados) tanto podem pertencer a uma quanto a outra população. Estes cálculos são ilustrados na figura 13.

Neste ponto é interessante refletir sobre a finalidade desses cálculos: a determinação de anomalias de distribuição, que indicarão anomalias geoquímicas, as quais (espera-se) indicarão a localização da mineralização. Desses cálculos nota-se que o valor do limiar mais aceito ($MG.DG^2$ da população de *background*) é quase o mesmo, sendo 157 ppm pelo cálculo manual simplificado, adotando-se o ponto de inflexão de 90%; 161,80 ppm calculado pelo programa, com escolha visual do ponto de inflexão em 92%; e 180,50 ppm calculado com escolha automática do ponto de inflexão. Há grande diferença se os dados são tomados em conjunto, sem separação de populações, pois assim o limiar seria de 282,04 ppm.

Se fosse adotado o limiar apontado por Damasceno (1967), usado para determinação do espaçamento de amostragem no Projeto Integração e Detalhe Geológico no Vale do Ribeira (Ferreira et alii, 1981), este seria de 100 ppm.

Calculando-se a estatística de lacuna (*gap*) conforme Miesch (1981), com o programa incluído em Koch (1987), ela indica mudanças significativas na distribuição entre 35 e 40 ppm (ao nível de 5%) e 45 e 50 ppm (ao nível de



FONTE:CPRM-PROJ. INT.DET.GEOL.V.RIBEIRA-RELATÓRIO FINAL-VOL. XIII - MACEDO,A.B. E BATOLLA,Jr.F.,1981.

Figura 11 - Diagrama de probabilidade lognormal, para frequências percentuais acumuladas de teores de chumbo em amostras de solo, coletadas sobre calcários na área Mina do Paqueiro, com os parâmetros da população de background determinadas pelo método simplificado.

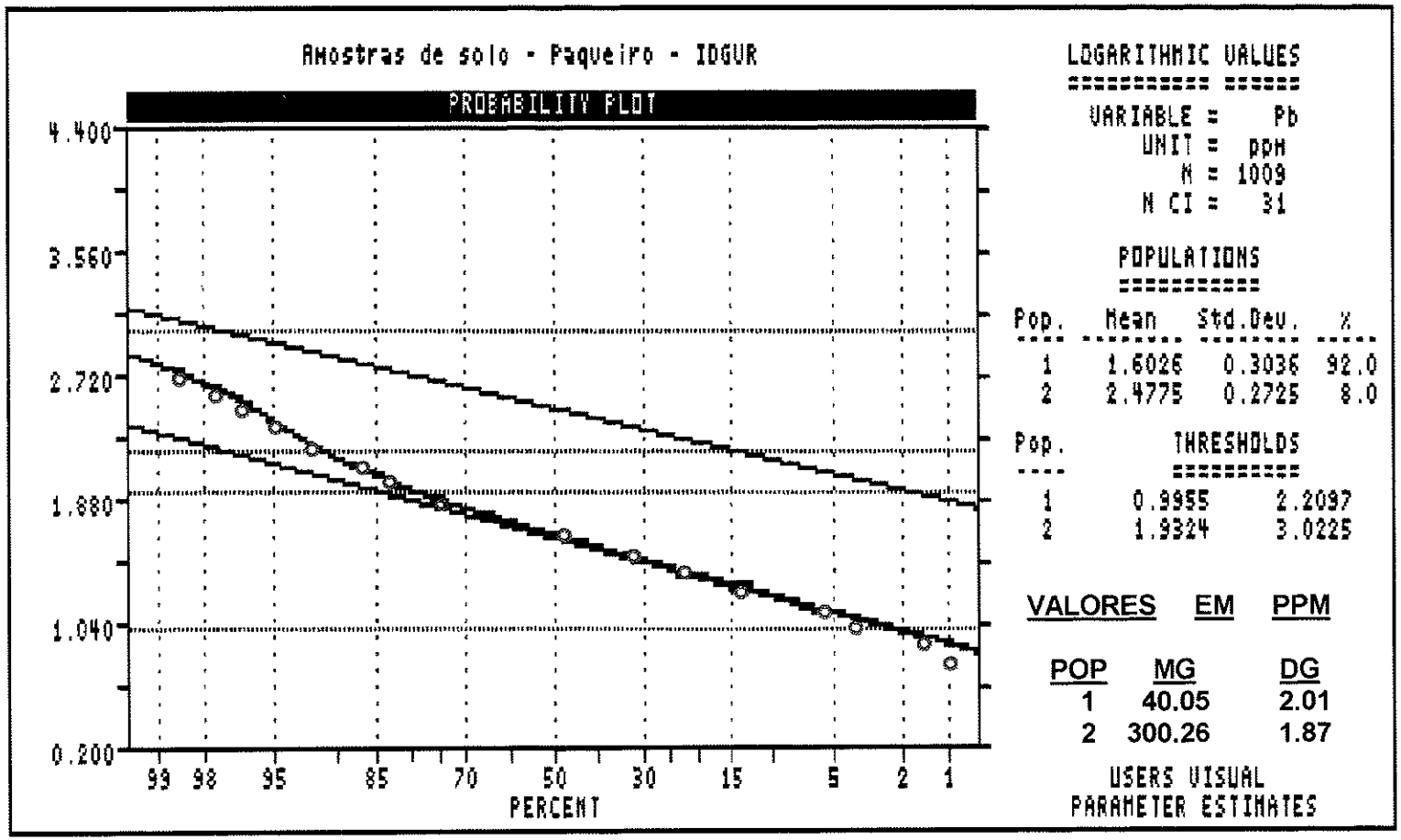


Figura 12 - Diagrama de probabilidade lognormal, para frequências percentuais acumuladas de teores de chumbo em amostras de solo, coletadas na área Mina do Paqueiro, com os parâmetros da população de background determinadas pelo programa PROBPLOT, a partir da avaliação visual do ponto de inflexão da curva.

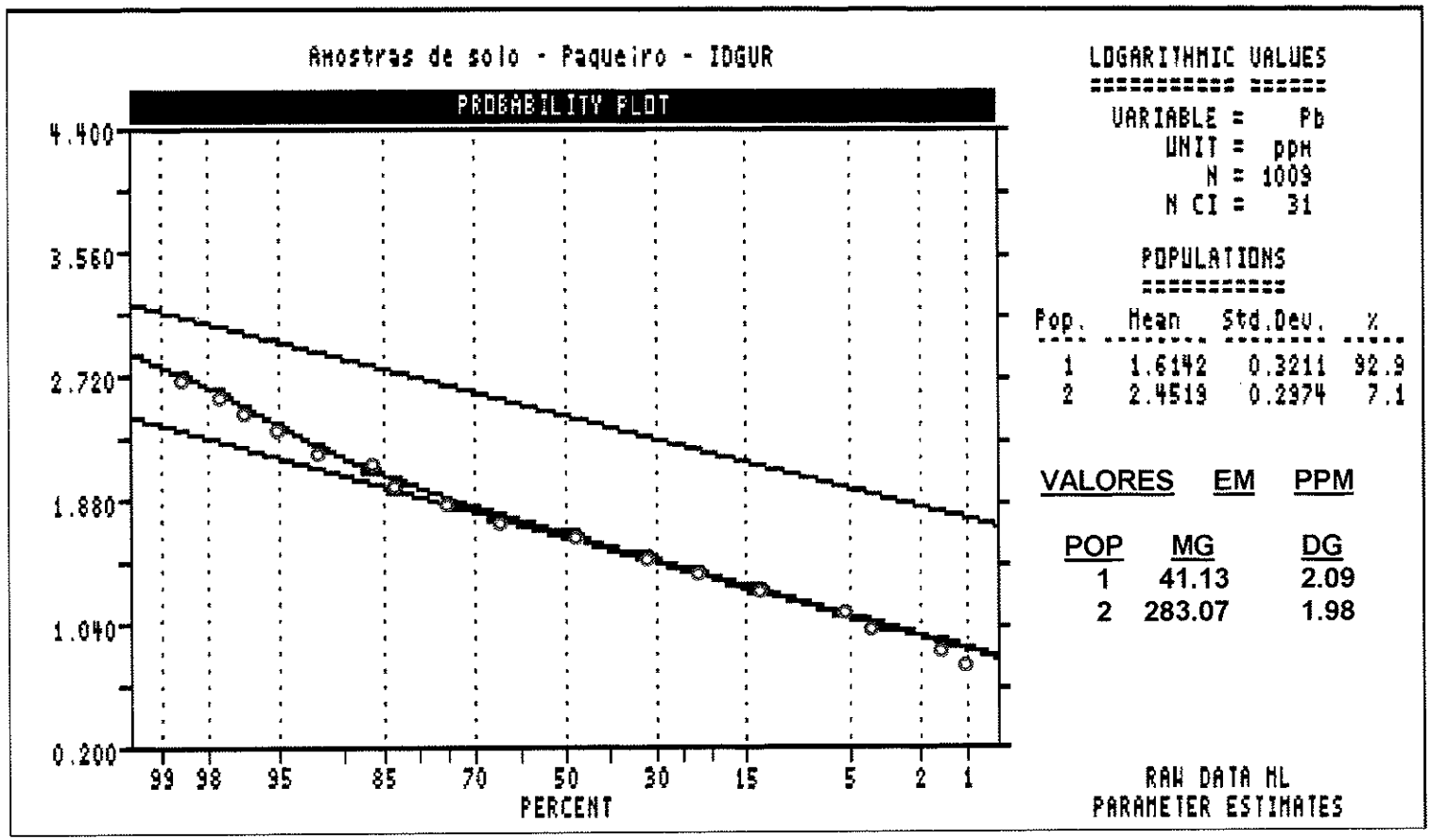


Figura 13 - Diagrama de probabilidade lognormal, para frequências percentuais acumuladas de teores de chumbo em amostras de solo, coletadas na área Mina do Paqueiro, com os parâmetros da população de background determinadas pelo programa PROBPLOT, pelo método de máxima verossimilhança.

1%), valores que podem ser sumariamente descartados como possíveis limiares de anomalias provocadas por mineralização, considerando-se a geologia da área amostrada e os resultados obtidos pelos outros métodos. Indicariam a divisão dos dados em dois grandes blocos, o que poderia ser devido a amostragem efetuada em solos derivados de litologias diferentes. Isto não é confirmado nem pela observação de campo nem pelos outros testes. Embora haja diferença significativa, entre as médias dos teores das amostras coletadas em solos derivados de rochas carbonáticas e xistos sericíticos (MG 46 e 27,5 ppm respectivamente, segundo Ferreira et alii, 1981), o pequeno número das coletadas sobre xistos (15%) não deveria afetar tanto a distribuição total. Fica aqui registrado como experiência prática da aplicação de uma técnica que parte de um princípio diferente das outras aqui empregadas. Um *gap* pode indicar um afastamento significativo da normalidade sem que necessariamente indique uma mistura de populações ou uma anomalia geoquímica.

Para verificar qual seria o efeito, na prática, de se adotar cada um dos limiares até agora calculados, foi feita uma simulação. A partir de um mapa de teores da área amostrada do Paqueiro, para cada limiar foram determinadas as superfícies das áreas anômalas (com teores acima do limiar) e de background (com teores abaixo do limiar). Num trabalho de detalhamento, a superfície da área anômala seria aquela coberta por trabalhos mais intensivos, condicionando então o custo das etapas posteriores. Os resultados estão expostos na tabela 2, seguinte:

Tabela 2: *Superfícies ocupadas por áreas anômalas e de background e percentagem da área anômala, para diversos limiares, para o total das amostras de solo da área do Paqueiro.*

LIMIAR (ppm)	Área de background (m^2)	Área anômala (m^2)	% anômala
100	3.415.275	1.030.725	30,17
157	4.073.300	372.700	9,15
161	4.107.525	338.475	8,24
180	4.225.575	220.425	5,22
277	4.377.975	68.055	1,56

Comparando-se as superfícies das áreas que seriam consideradas anômalas, nota-se pouca diferença naquelas determinadas usando o limiar de $MG.DG^2$ da população de *background*, calculado pelo método manual simplificado (157 ppm) ou pelo programa Probplot, escolhendo manualmente o ponto de inflexão (161 ppm). Uma diferença um pouco maior acontece se adotamos o ponto de inflexão escolhido pelo programa, com o limiar de 180 ppm. O limiar escolhido arbitrariamente (100 ppm) levaria a uma área anômala muito grande, e se não tivessem sido separadas as populações, resultando num limiar de 277 ppm, a área seria pequena demais. Mas qual é realmente o significado do limiar ?

O limiar não pode ser considerado absoluto, nem interpretado isoladamente. Os parâmetros da população de *background* fornecem limites convenientes para curvas de isoteor, mas não podemos tomar decisões de prospecção, tais como determinar áreas para detalhamento, apenas com cálculos sobre dados geoquímicos derivados de um elemento. É necessário juntar as informações provindas de diversos elementos químicos, os dados litológicos, estruturais, geofísicos, e todos os outros disponíveis.

Por mais bem apoiada estatisticamente que seja uma técnica de determinação de limiar, ela apenas lida com números, cabendo ao geólogo julgar em que medida esses números indicam feições e fenômenos geológicos. Particularmente importante é associar esses números à localização das amostras, e trabalhar espacialmente todas as informações (geológicas, geoquímicas e geofísicas). Isto foi feito, no projeto aqui comentado, de forma manual, sobrepondo as informações geológicas a curvas de isoteores dos três elementos analisados (Cu, Pb e Zn). Atualmente isto pode ser feito melhor por computador, assunto que será melhor comentado e exemplificado, com dados da mesma área, no item "Geoprocessamento" (capítulo 7).

6. ANÁLISE DE DADOS MULTIVARIÁVEL

6.1. Análise a duas variáveis:

6.1.1. Introdução

Em prospecção geoquímica, em muitos casos, um problema não pode ser resolvido estudando-se o comportamento dos teores de apenas um elemento. Torna-se necessário o estudo conjunto dos teores de vários elementos, ou de teores e outras variáveis representativas de parâmetros geológicos ou ambientais.

É necessário muito cuidado ao empregar técnicas multivariáveis. O menor malefício conseqüente pode ser o gasto inútil de tempo e recursos, fazendo análises multivariáveis, quando o estudo de uma variável resolveria a questão. Mais grave é a utilização de uma bateria de testes sofisticados, até que os resultados numéricos confirmem a idéia preconcebida do pesquisador. Quanto mais complicado o teste, mais possível é que assim aconteça. A partir da experiência, as regras seguintes podem ser enunciadas:

- sempre começar com os tratamentos univariáveis mais simples, progredindo para os mais complexos apenas quando os mais simples tiverem sido totalmente interpretados e compreendidos. Só usar métodos especiais, como a divisão em populações lognormais componentes de uma mistura de distribuições, quando houver muito bons indícios de que realmente ocorra o fenômeno testado;
- só passar para o estudo multivariável quando houver interpretado totalmente os testes univariáveis;
- só usar os testes multivariáveis mais complexos quando houver esgotado os mais simples. Primeiramente testes a duas variáveis, como gráficos de dispersão e correlações simples, depois os multivariáveis;
- para cada teste, interpretar os resultados associando os resultados estatísticos a outros dados geológicos, sempre colocando os resultados em mapa, antes de passar a outros testes.

O cumprimento dessas regras teria evitado que o autor tivesse mais de um metro de espessura de listagens de computador, referentes a processamentos feitos para sua tese de doutorado (Macedo, 1986), com menos mapas, do que seria de se desejar. Os equipamentos e programas atuais facilitam fazer o processamento correto, por sua natureza interativa, que permite pedir um processamento, analisar seus resultados, mudar os parâmetros, pedir outro, etc. Quando um grande número de pessoas utilizava um computador de grande porte, dos quais havia muito poucos disponíveis, era necessário especificar uma série de processamentos em cada trabalho, pois era necessário esperar horas para seus resultados.

6.1.2. Gráficos de dispersão

O tratamento a duas variáveis mais simples é lançar os valores em gráfico, com eixos de coordenadas aritméticas ou logarítmicas. Dois tipos são importantes para interpretação geoquímica: o gráfico de dispersão de teores e o gráfico teor-distância.

Os gráficos de dispersão são bem menos utilizados do que poderiam ser, especialmente considerando a facilidade de sua execução. Eles deveriam sempre ser examinados quando interpretando coeficientes de correlação, principalmente coeficientes de Pearson. Estes podem ser muito influenciados por algumas amostras, com teores altos, dando altos coeficientes de correlação que não são confirmados pela observação do gráfico ou por medidas de correlação não-paramétricas.

Um exemplo disso é a correlação entre os teores de Pb e Ag em rochas na área do Perau. Este conjunto de dados compreende 126 pares de teores em ppm, sendo 76 obtidos por espectrografia ótica de emissão e 50 por absorção atômica, em amostras de rocha total.

Na maioria das amostras um dos teores é inferior ao limite de detecção (0,5 ppm para Ag, 10 ppm para Pb). Neste caso foi adotado o critério do USGS, também usado pela CPRM, para valores qualificados: atribui-se o valor de 0,3 vezes o limite para teores assinalados N (não detectado o elemento) e 0,5 vezes quando qualificado L (detectado em teor abaixo do limite).

Um conjunto de dados como este, misturando análises diferentes, dos laboratórios da CPRM e da PB (British Petroleum Mineração, Rio de Janeiro), deve ser usado com cuidado, sendo empregado aqui apenas para ilustração da variabilidade dos coeficientes de correlação, e não para interpretação geoquímica.

Se considerados os valores de prata, apenas 30 apresentam teores superiores ao limite de detecção. Para verificar o que pode acontecer quando se calculam coeficientes com pouca crítica de dados, foram comparados os resultados para o total dos dados (126 amostras); para o total menos uma amostra com teores muito altos nos dois elementos (125 amostras); dos dados com teores de Ag acima do limite de detecção (30 amostras);, e destes, retirando-se a amostra discrepante (29 amostras).

Além do cálculo numérico, cujas fórmulas e coeficientes serão melhor explicados no item seguinte, os dados foram lançados em gráficos de dispersão, em escalas aritmética e logarítmica. Os resultados, para as amostras com teores de prata acima do limite de detecção, são ilustrados na figura 14, onde aparece claramente a amostra discrepante, com teores de 16.736 ppm de Pb e 27,8 ppm de Ag. Ela influencia demasiadamente os resultados com dados aritméticos, nos quais condiciona um coeficiente de correlação alto, que não aparece quando ela é removida. Os resultados com e sem a amostra discrepante são menos diferentes para os dados logaritmizados.

Para testar mais rigorosamente a influência da amostra discrepante, no cálculo da correlação entre Ag e Pb, nas amostras de rocha do Perau, foram calculados os coeficientes de correlação de Pearson entre os teores destes dois elementos, tanto para dados brutos quanto logaritmizados. Também foram calculados os coeficientes de Spearman, que são iguais para dados brutos e logaritmizados, pois se baseiam não nos teores mas nos postos, ou seja, na colocação dos dados numa sucessão crescente de valores. Foram empregadas as fórmulas apresentadas no item 6.3, e os coeficientes de Spearman foram calculados duas vezes, uma usando a fórmula normal, outra a fórmula corrigida para dados com muitas repetições. Os limites críticos da

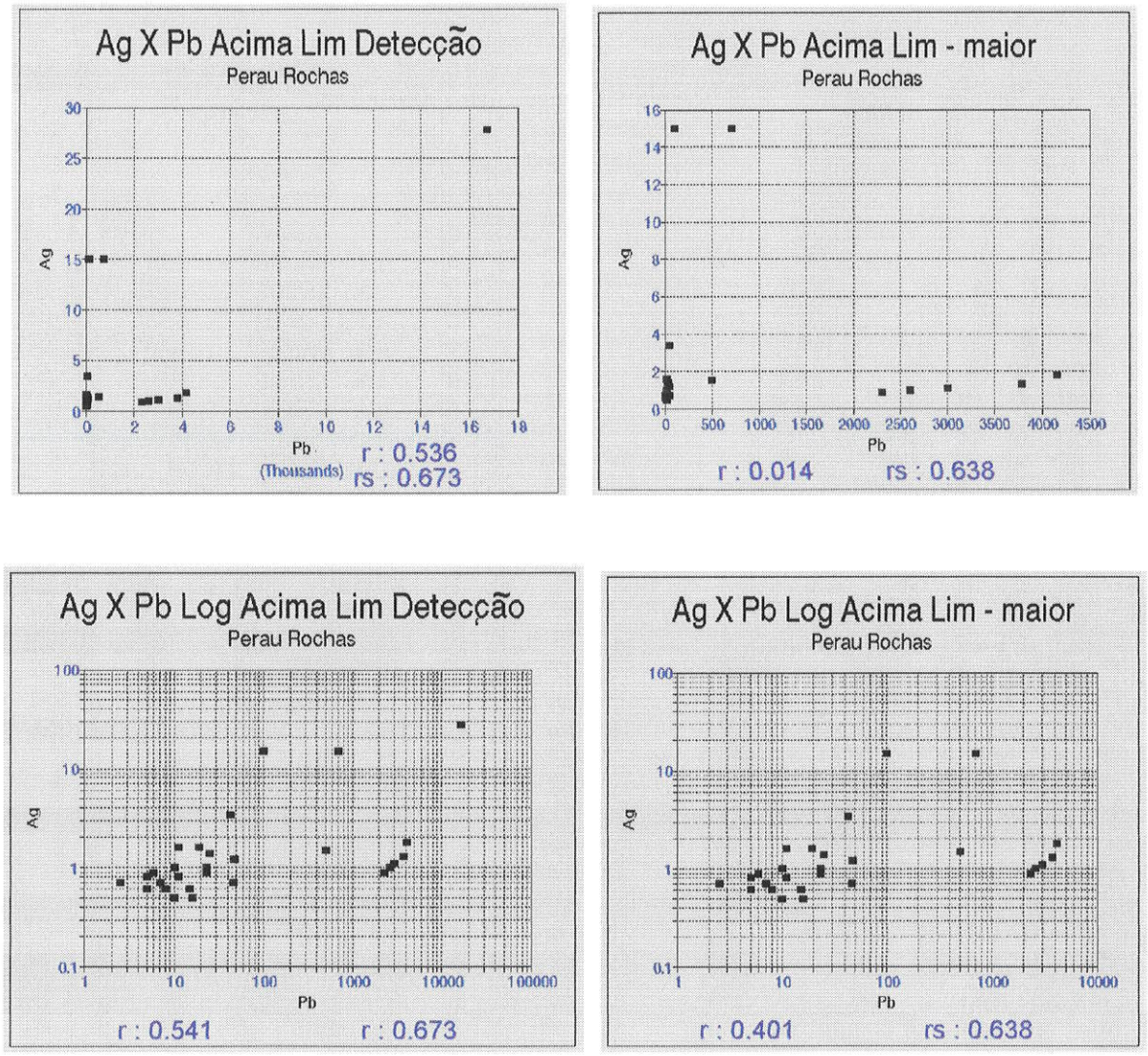


Figura 14 - Gráficos de dispersão de teores de prata x teores de chumbo, para amostras de rocha coletadas na área Mina do Perau. Os gráficos assinalados - ***maior*** foram elaborados retirando uma amostra discrepante, com teor muito maior que a média. (**r** = coeficiente de correlação de Pearson. **rs** = coeficiente de correlação de Spearman).

estatística *t* de Student foram tomados ao nível de significância de 5%, na tabela 2.11 de Davis (1986).

Tabela 3: *Coeficientes de correlação linear de Pearson e Spearman para teores de Ag e Pb de amostras de rocha da mina do Perau, e valores calculados e críticos da estatística t de Student.*

Coeficientes de correlação linear de Pearson				
Grupo	**Coeficiente**	**n-2**	**t calc**	**t crit**
rp ar 126	0,757	124	12,886	1,658
rplog 126	0,450	124	5,609	1,658
rp ar 125	0,134	123	1,542	1,658
rp log 125	0,363	123	4,316	1,658
rp ar 30	0,536	28	3,364	1,701
rplog 30	0,541	28	3,410	1,701
rp arit 29	0,014	27	0,070	1,703
rp log 29	0,401	27	2,271	1,703
Coeficientes de correlação linear não-paramétrica de Spearman				
Grupo	**Coeficiente**	**n-2**	**t calculado**	**t crítico**
r_s 126	0,382	124	6,884	1,6669
r_s repet 126	0,304	124	5,047	1,6669
r_s 125	0,367	123	6,485	1,6671
r_s 30	0,673	28	8,399	1,701
r_s repet 30	0,670	28	8,318	1,701
r_s 29	0,638	27	7,385	1,703

Se é considerado o total dos dados, o valor do coeficiente de Pearson é de 0,757 para dados aritméticos e de 0,450 para dados logaritmizados; estes valores caem para 0,134 e 0,363 respectivamente, quando se remove a amostra discrepante.

Para os valores das amostras onde Ag foi detectada acima do limite de detecção, o coeficiente de correlação de Pearson entre teores de Ag e Pb é igual a 0,536; entre dados logaritmizados é de 0,542; se tirarmos o dado discrepante, os coeficientes passam a 0,014 para teores brutos e 0,401 para dados logaritmizados.

A correlação pode ser sempre considerada significativa ao nível de 5%, se admitirmos o total dos dados. Ela não é significativa, para os coeficientes calculados com dados aritméticos, quando se retira a amostra discrepante.

Para o total dos dados, o coeficiente de Spearman é de 0,382, calculado pela fórmula normal e 0,304 pela fórmula que leva em conta o número de postos repetidos (segundo Landim, 1979); se tirarmos o último par, com os

dados discrepantes, o coeficiente se torna 0,367, uma mudança muito pequena.

Para os pares de dados de amostras com teores acima do limite de detecção para Ag os coeficientes são respectivamente 0,673, 0,670 e 0,638. Em todos os casos as correlações não-paramétricas podem ser consideradas significativas ao nível de 5%.

Esta grande variação nos resultados de correlação, para essencialmente o mesmo conjunto de dados, mostra a necessidade da análise de cada conjunto disponível, antes que seja definitivamente incorporado ao banco de dados, e antes das análises mais complexas e da interpretação. A distribuição de freqüência deve ser testada, pelo menos para as hipóteses normal e lognormal. Para teores com amplo intervalo de variação, pode ser necessário trabalhar com dados logaritmizados, e devemos tomar cuidado com amostras discrepantes, que devem ser na maioria dos casos eliminadas antes do cálculo de estatísticas agregadas. Quando forem estudadas as correlações, pelo menos um diagrama de dispersão aritmético e um logarítmico devem ser examinados.

Não é possível, porém, fixar uma regra geral sobre o tipo de coeficiente a utilizar, ou se devemos sempre usar dados logaritmizados. Isto depende da distribuição de freqüência dos dados, do tipo de teste e da finalidade do estudo. Incluir ou não dados qualificados (acima ou abaixo do limite de detecção), também é uma decisão que deve ser tomada após o exame do total dos dados.

O gráfico de dispersão também pode mostrar uma correlação não-linear, expressa por uma equação de regressão, entre as duas variáveis, com grau superior a um, ou por funções mais complexas. O coeficiente de correlação linear pode ser baixo, mas os pontos no gráfico se alinharão segundo uma curva.

6.1.3. Coeficientes de correlação

Caso os pontos no gráfico de dispersão de duas variáveis se distribuam aproximadamente numa reta, isto indica que é possível expressar a relação entre os teores dos dois elementos como uma equação de primeiro grau, da forma y = ax + b. O grau de correlação linear é calculado mais comumente por dois coeficientes:

6.1.3.1. Coeficiente de correlação linear produto-momento de Pearson:

É calculado pela fórmula:

$$r_{xy} = \frac{\mathrm{cov}_{x,y}}{\sigma_x \sigma_y} = \frac{\dfrac{\sum (x_i - \bar{x})(y_i - \bar{y})}{n-1}}{\sqrt{\dfrac{(x_i - \bar{x})^2}{n-1} \cdot \dfrac{(y_i - \bar{y})^2}{n-1}}},$$

onde $\mathbf{n}$ é o número de pares de valores para $\mathbf{x_i}$ e $\mathbf{y_i}$, variáveis com distribuição normal e $\bar{\mathbf{x}}$ e $\bar{\mathbf{y}}$ são os valores médios para $\mathbf{x_i}$ e $\mathbf{y_i}$, segundo Landim, 1979.

O teste usado para verificar se a correlação é ou não significativa é o t de Student unicaudal, com (n-2) graus de liberdade, calculando-se

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

O coeficiente de Pearson, como já foi mostrado acima, pode ser influenciado por dados com valores muito discrepantes dos da maioria. Além disso, só pode ser testado para distribuições normais ou transformadas na normal. O coeficiente não-paramétrico (que pode ser testado independente da distribuição dos dados) mais usado é o de Spearman.

6.1.3.2. Coeficiente de correlação não-paramétrica de Spearman

Ele é calculado, segundo Landim (1979), ordenando-se as variáveis x_i e y_i, de n_i pares de valores, e atribuindo-lhes seus valores de postos (localização na ordenação), x_i' e y_i'. A diferença d_i entre x_i' e y_i' é elevada ao quadrado, e calcula-se o coeficiente de correlação não paramétrica de Spearman:

$$r_s = 1 - \frac{6\sum d_i^2}{n^3 - n}$$

Para os casos em que ocorrem muitos pares com valores de posto empatados, usa-se a fórmula:

$$r_s = \frac{\sum x_e + \sum y_e - d_i^2}{2\sqrt{\sum x_e \cdot \sum y_e}}, \text{ onde}$$

$$\sum x_e = \frac{n^3 - n}{12} - \sum T_x$$

$$\sum y_e = \frac{n^3 - n}{12} - \sum T_y$$

$$T = \frac{t^3 - t}{12},$$

sendo t o número de observações repetidas num determinado posto.

O teste de significância, para mais de 10 pares de amostras, é o mesmo do teste de Pearson.

Em quase todos os trabalhos do autor foram empregados coeficientes de correlação. Em alguns trabalhos a correlação era um problema central, como na dissertação de mestrado, de título: "Estudo estatístico das relações entre elementos principais e traços em rochas basálticas da Bacia do Paraná" (Macedo, 1976), na tese de doutoramento (Macedo, 1987) e em seus trabalhos derivados (Barbour et alii, 1988; Macedo, 1989; e Barbour & Macedo, 1991). Estes trabalhos serão comentados adiante.

Quando se empregam coeficientes de correlação entre muitas variáveis a compreensão da matriz de coeficientes pode melhorar se eles forem

apresentados de forma gráfica, como na figura impressa em Macedo, 1979, e reproduzida na figura 15.

6.1.4. Gráficos teor-distância

Este tipo particular de gráfico, embora simples, é uma das ferramentas mais valiosas para a interpretação de dados em prospecção geoquímica. Um exemplo já foi mostrado, na figura 6, do estudo orientativo da mina do Paqueiro. Outro caso é o apresentado em Macedo (1992, 1993b), onde foram plotados os teores de Cu, Pb e Zn em relação à distância à mina do Perau, tomada no curso do Ribeirão Grande, reproduzido na figura 16. Pelo gráfico pode-se perceber não apenas a redução dos teores com o afastamento da mina, quanto o aumento nos pontos em que o ribeirão recebia afluentes vindos das minas de Canoas e Barrinha. Para que seja possível identificar a tendência geral dos pontos, o gráfico foi suavizado aplicando a média móvel, ficando cada ponto com o valor da média entre seu valor e dos dois pontos adjacentes.

	SiO_2	TiO_2	Al_2O_3	Fe	MnO	MgO	CaO	Na_2O	K_2O	Ba	Co	Cr	Cu	Ni	Rb	Sr	V	Zn	Zr
SiO_2				-■○	-□○	-□○	-■○		■○	○	-□	-□○		-□○	■○		-■○		
TiO_2				□○												□○	□		□○
Al_2O_3																			
Fe	-■○	□○			■○	□	□		-□		□○				-■○		■●		
MnO	-□○			■○			□		-□○						-□○		□○		
MgO	□○			□	□		■●	-□○	-■○	-□○		□○		□○	-□○		□○		-□○
CaO	-■○			□	□	■●		-□○	-■●	-□○		□		□○	-■○		□○		-□○
Na_2O						-□○	□		□○	□○				-□○	○				
K_2O	■○			-□	-□○	-■○	-■	□○		○	-□○			□○	■●		-□○		□○
Ba	○					-□○	-□	□						○	○				○
Co	-□			□○					-□○						-□○		□○		
Cr	-□○					□○	□							□○	-□				
Cu																	○		
Ni	□○					□○	□○	-□○	-□○	○		□○			-□○				-□○
Rb	■○			-■○	-□○	-□○	-■○	○	■●	○	-□○	-□		-□○			-□○		
Sr		□○																	□○
V	-■○	□○		■●	□○	□○	□○		-□○		□				-□○				
Zn																			
Zr		□				-□○	-□○		□○	○				-□		□○			

Figura 15 - Representação gráfica dos valores de coeficientes de correlação de Pearson (quadrados) e Spearman (círculos), calculados para teores de elementos traços e óxidos de elementos principais em 77 análises de rochas basálticas da bacia do Paraná (Macedo, 1979). Símbolos cheios indicam valor do coeficiente superior a 0,8; símbolos vazios valor entre 0,5 e 0,8.

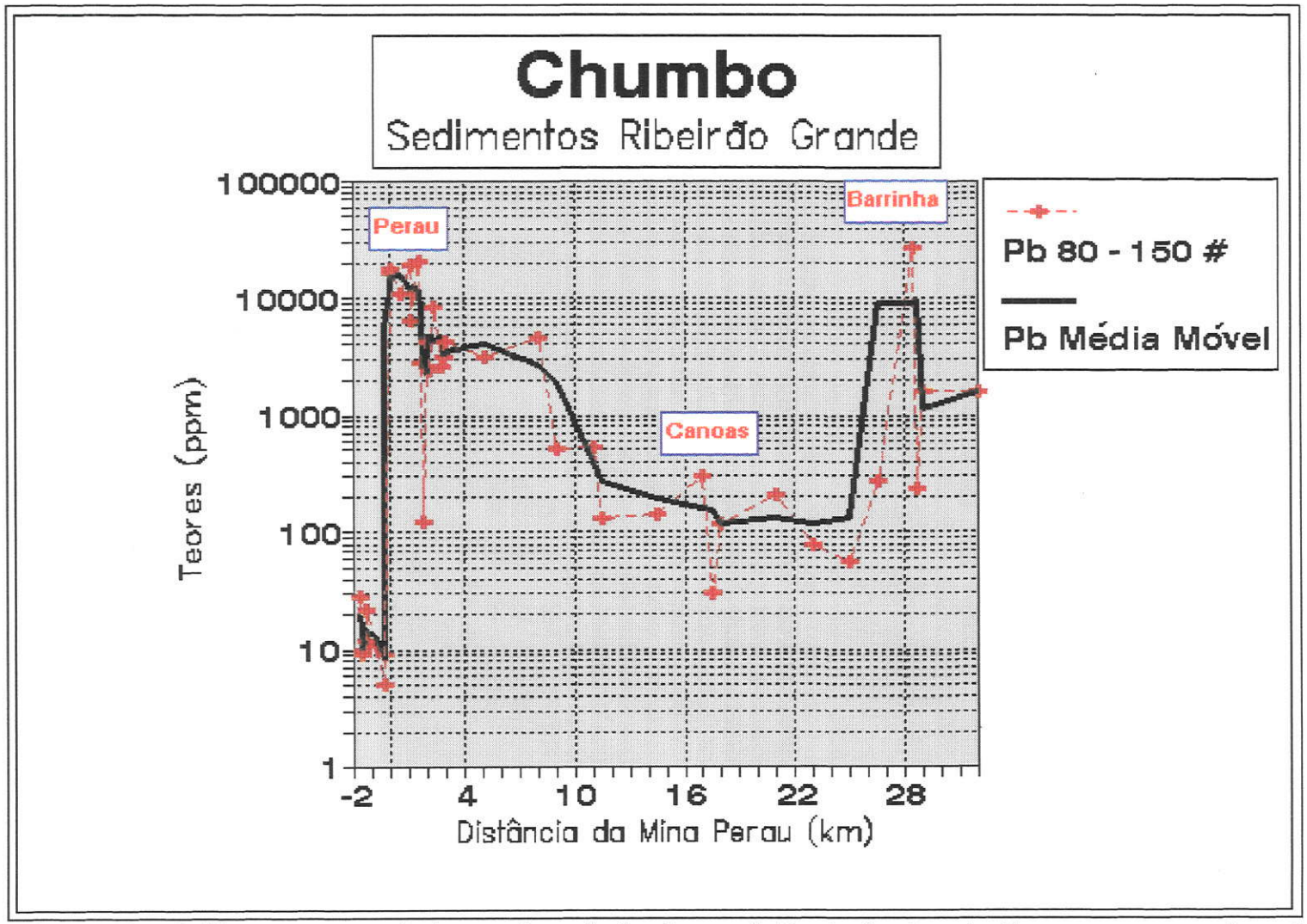


Figura 16 - Gráfico comparando teores logaritmizados e distância à mina do Perau, para amostras de sedimentos de corrente coletadas no Ribeirão Grande (Macedo, 1992, 1993*a*).

6.2. Múltiplas variáveis

6.2.1. Análise de agrupamentos

Partindo de uma matriz de coeficientes de similaridade, é possível classificar variáveis ou amostras de comportamento mais semelhante, juntando-as em agrupamentos, e representando estes agrupamentos por diagramas, dos quais o mais usado é o dendrograma. Davis (1986) apresenta diversos modos de classificar objetos, a partir de coeficientes de similaridade. Este método geralmente não está incluído nos programas estatísticos mais comuns. Um bom programa para fazer os dendrogramas é o Cluster, incluído no sistema GEOQUANT da CPRM (Branco, 1986), usado em Macedo (1986).

A análise de agrupamentos foi utilizada na obra do autor em Macedo (1979 e 1986). No primeiro trabalho a matriz de coeficientes de correlação de Pearson, de onde saíram os coeficientes representados por símbolos quadrados na figura 15, foi transformada em dendrograma pelo método de pares ponderados, pelo programa apresentado em Davis (1973).

Este método se baseia em encontrar na matriz simétrica de coeficientes de correlação, inicialmente, os coeficientes mais altos em cada coluna. Dentre estes, selecionam-se os que apresentam as correlações mutuamente mais altas entre um par de objetos estudados (no caso teores em percentagem de óxidos de elementos principais e em ppm de elementos traços de rochas basálticas da bacia do Paraná). O par é reunido e tratado como um só objeto, sendo recalculadas as correlações dos outros elementos com o par, através da média das correlações que cada elemento mantinha com cada um dos elementos do par. O processo é repetido até que reste apenas uma matriz 2 x 2, indicando a correlação entre os dois principais grupos de elementos. O resultado foi o dendrograma apresentado na figura 17, representando a divisão dos elementos entre cinco grupos, que abrangem todos eles menos o Al e o Cu. Os óxidos e elementos traços distribuem-se pelos grupos conforme sua ligação com os minerais formadores de rochas.

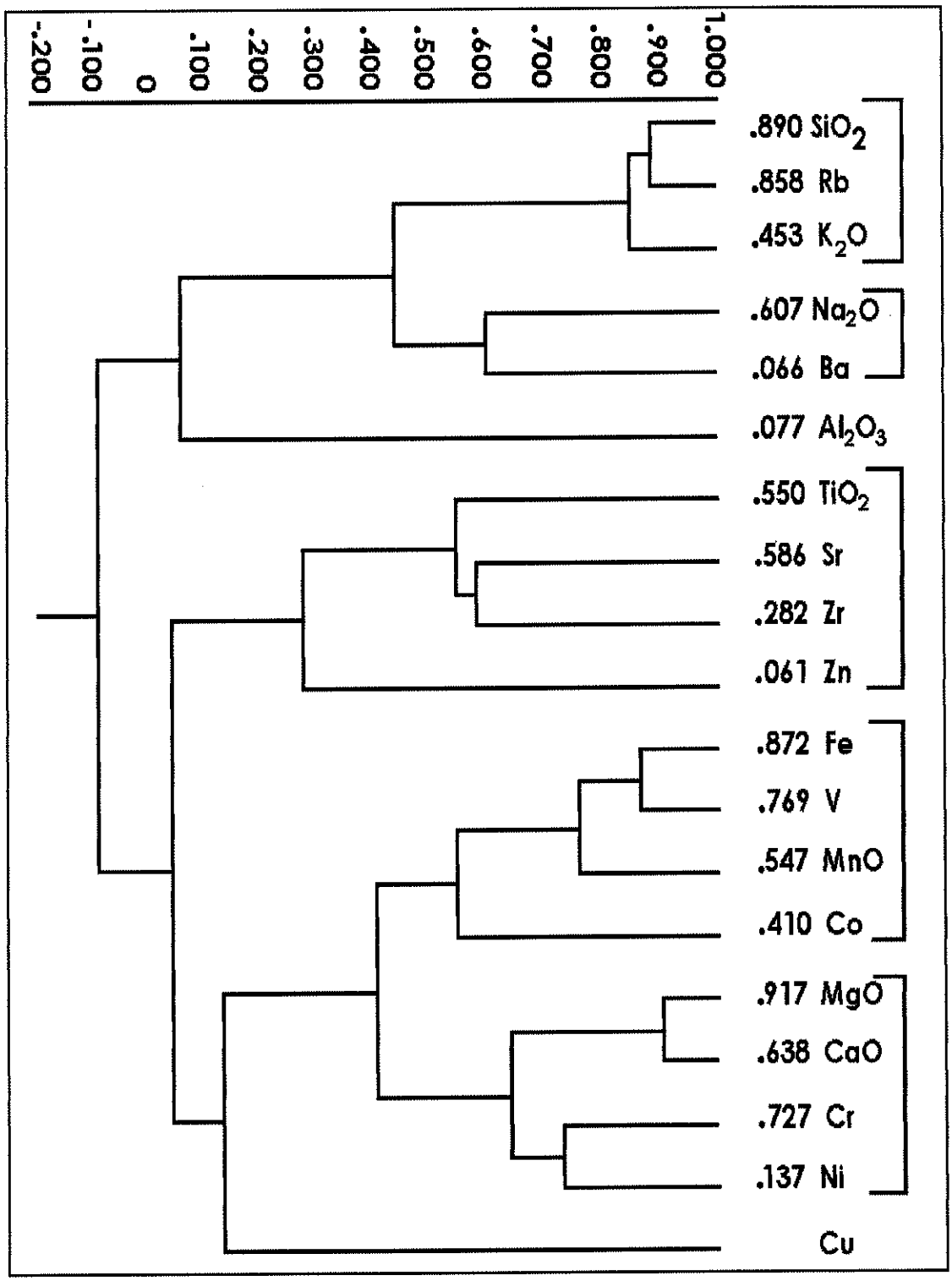


Figura 17 - Relações entre grupos de elementos principais e traços, em 77 análises de rochas basálticas da Bacia do Paraná, representadas por dendrograma calculado a partir de coeficientes de correlação de Pearson, segundo Macedo (1979).

Em Macedo (1986) a análise de agrupamentos foi também utilizada, tanto entre elementos quanto entre amostras. Entre elementos mostrou o agrupamento destes de acordo com os minerais formadores de rocha; entre amostras não apresentou agrupamentos que pudessem ser correlacionados com composição ou com localização em relação ao minério.

A análise de agrupamentos permite, de forma rápida e de fácil compreensão, examinar as relações entre as variáveis ou entre as amostras em uma matriz de dados. Porém apresenta os seguintes problemas:

- os agrupamentos variam conforme o método de cálculo usado para a redução da matriz; é necessário discernimento na sua escolha e deve-se resistir à tentação de ficar experimentando até que uma técnica de montagem de agrupamentos dê resultados semelhantes aos que o pesquisador espera.
- embora haja maneiras de calcular a distorção que o cálculo de agrupamentos introduz na matriz dos coeficientes, como os coeficientes cofenéticos usados em Macedo (1979), não existem testes estatísticos rigorosos para testar a eficiência dos agrupamentos (Landim, 1995).

6.2.2. Análise de Componentes Principais

A análise de componentes principais foi empregada pelo autor no tratamento dos dados da área Mina do Paqueiro (Salles, Macedo & Liotte, 1995). Na forma usada pelo módulo PCA do IDRISI (Eastman, 1992*a* e *b*) ela converte uma série de imagens (tais como bandas de sensoriamento remoto ou mapas de distribuição de teores em amostras de solo, como no caso estudado), em outra série, que exprime a mesma variabilidade, porém através de variáveis obtidas por cálculo (principais componentes, calculadas a partir dos autovetores e autovalores da matriz de coeficientes de similaridade entre os valores nas diversas imagens), não correlacionadas entre si. Esta análise deve permitir identificar relações entre as variáveis e amostras, não aparentes nos dados brutos e, em muitos casos, também reduzir o número de variáveis necessário para expressar a variabilidade dos dados, uma vez que geralmente os primeiros componentes são responsáveis pela maior parte da variabilidade.

Os principais componentes podem ser lançados em mapa, o que facilita a interpretação. Os resultados referentes ao Paqueiro são apresentados e discutidos no capítulo de Geoprocessamento, juntamente com outros processos para tratamento de dados georreferenciados.

6.2.3. Análise fatorial

A análise fatorial vai um passo além da análise de componentes principais, pois procura explicitamente reduzir o número de componentes, transformando-os em fatores, em número menor que as variáveis (análise modo R) ou amostras (modo Q). A análise fatorial foi empregada pelo autor em seus dois primeiros trabalhos publicados (Macedo & Rüegg, 1974 *a* e *b*), onde foi estudada a variação de elementos principais e traços de rochas basálticas da Bacia do Paraná.

No primeiro destes trabalhos, o autor escreve:

> Para efetuar a redução de grande número de variáveis, que é o caso de análises químicas de rochas, a coeficientes fatorais, podemos representar uma matriz de dados na forma do produto de duas matrizes mais simples, segundo o procedimento proposto por Manson (1967), conforme a equação abaixo:
>
> $$X_{(N,n)} = A_{(N,m)} \cdot F_{(m,n)} + E_{(N,n)},$$
>
> onde N = número de observações (análises químicas)
>
> n = número de variáveis (elementos dosados)
>
> $X_{(N,n)}$ = *matriz inicial* de dados
>
> m = número de fatores capaz de explicar a variabilidade dos dados. Este número é determinado por dois critérios: deve ser capaz de explicar a maior porcentagem possível da variabilidade e, para facilidade de interpretação e representação gráfica, deve ser o menor possível.

$A_{(N,m)}$ = *matriz de coeficientes fatoriais*, descrevendo as N análises como proporções de m membros finais (fatores).

$F_{(m,n)}$ = *matriz de pesos fatoriais*, que descreve a composição de cada um dos fatores (membros finais) como combinação das n variáveis.

$E_{(N,n)}$ = matriz da variabilidade não representada por A e F. Uma boa solução matemática deve fazer com que esta matriz tenha o valor mais próximo possível de zero. Geralmente esta condição não é atingida, porque quando reduzimos o número de fatores em relação às variáveis, perdemos em precisão uma parte do que ganhamos em simplicidade. Além disso os erros de medidas também contribuirão para formar a matriz E.

Nos dois trabalhos os autores utilizaram matrizes de pesos fatoriais, fornecidas por Manson (1967) e Prinz (1967), para comparar o comportamento de análises de rochas basálticas da Bacia do Paraná com o de rochas de outras áreas do mundo, por gráficos representando três e quatro fatores. Também calcularam independentemente matrizes de pesos fatoriais, obtendo fatores correlacionáveis com a composição química e mineralógica das rochas basálticas.

A comparação com rochas de outras áreas forneceu resultados discordantes com os dados mineralógicos. Isto pode ser interpretado de duas maneiras: uma, aquela que os autores adotaram nos trabalhos, de que as rochas basálticas da bacia do Paraná apresentam um comportamento anômalo na relação mineralogia-química; outra, também possível, e não testada, de que as matrizes de pesos fatoriais utilizadas, calculadas por Manson e Prinz, a partir de 1996 e 267 análises, respectivamente, não tivessem a generalidade necessária para bem posicionar as rochas estudadas em gráficos de classificação. Isto pode ter acontecido, considerando-se que as discrepâncias observadas foram maiores para os elementos traços, onde a matriz de pesos fatoriais, calculada a partir de 267 análises, que não incluíam basaltos da Bacia do Paraná, foi usada para plotar 101 análises dessa área.

Estes trabalhos ilustram bem os problemas da aplicação dos métodos matemáticos mais complexos:

- há quase sempre uma solução matemática, que nem sempre tem validade geológica, que vai depender da representatividade da amostra estatística utilizada;
- não se deve confiar em índices, gráficos, etc, de trabalhos feitos por outros pesquisadores, mesmo que sejam os luminares da área de conhecimento. O número de 267 análises empregado por Prinz (1967), parece realmente muito pequeno para servir como padrão para comparação.

6.2.4. Regressão Múltipla

A regressão múltipla foi empregada pelo autor em sua dissertação de mestrado (Macedo, 1979), com a finalidade de estudar as relações entre os teores de elementos principais e traços em rochas basálticas da Bacia do Paraná e em sua tese de doutorado (Macedo, 1986), para determinar a aplicabilidade da prospecção litogeoquímica em jazidas do tipo Perau.

6.2.4.1. Regressão de teores de elementos principais a partir de elementos traços em rochas basálticas da Bacia do Paraná

Segundo o primeiro trabalho, Macedo (1979):

> A regressão múltipla procura, a partir dos valores de uma variável dependente (Y) e de um número qualquer de variáveis independentes (X_1, X_2, ...X_n), estabelecer uma equação da forma
>
> $$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots\ldots + \beta_m X_m$$
>
> que permite prever o valor de Y quando conhecidos os valores das variáveis independentes.
>
> Os termos "independente" e "dependente" empregados neste contexto são arbitrários, sendo comuns nas análises de regressão múltipla a

utilização de um grupo de variáveis, ora como independentes, ora como dependentes (ver Draper & Smith, 1967 e Nie, Bent & Hull, 1970).

Os coeficientes β_j são estimados por coeficientes b_j, ficando a equação com a forma:

$$Y' = b_0 + b_1X_1 + b_2X_2 + \ldots\ldots + b_m X_m$$

Em geral o valor calculado da variável dependente Y' não é exatamente igual ao valor observado, resultando em cada caso *(teor, nos estudos em pauta)*, um resíduo:

$$e_i = Y'_i - Y_i$$

A eficiência da equação de regressão para explicação das relações entre a variável dependente e as variáveis independentes, e também para previsão dos valores da variável dependente quando conhecidos os valores da variáveis independentes, é testada por dois valores:

R múltiplo: coeficiente de correlação entre os valores observados e calculados da variável dependente. O quadrado de R, (R^2), representa a percentagem da variância da variável dependente explicada pela regressão.

F = razão entre a variância devida à regressão e a variância devida ao resíduo, que é calculada como uma análise de variância da forma:

Fonte	**Soma de quadrados**	**Graus de Liberdade**	**Quadrados médios**	**Estatística F**
Regressão	SS_R	m	MS_R	
Resíduos	SS_E	n - m - 1	MS_E	MS_R / MS_E
Total	SS_T	n - 1		

O valor de F é testado em comparação com os valores da estatística F de Fisher-Snedecor, ao nível de significância escolhido.

Analisando a tabela, notamos que haverá uma tendência à diminuição do valor de F, pelo aumento do número de graus de liberdade do numerador, e sua diminuição no denominador, toda vez que acrescentarmos uma nova variável independente à equação de regressão. Isto pode fazer com que, ao aumentarmos o número de variáveis, embora aumente o valor de R, o valor de F se torna tão baixo que a equação se torna não significante.

O controle da significância (que implica a aplicabilidade) da equação foi feito, nos dois trabalhos citados, pela técnica da regressão múltipla seqüencial, pela qual é acrescentada à equação uma variável independente de cada vez, aquela que tem a maior correlação com a variável dependente. São testados o R múltiplo e o F; caso haja um aumento no R e o F continue significante, a variável é mantida; caso contrário é excluída e passa-se a outra variável, até que não haja mais contribuição ao R mantendo F significante.

Para as 77 amostras estudadas no trabalho em pauta, o F crítico a 5% será, de uma a 10 variáveis, de 1,65 a 1,69.

A partir desta técnica foram calculadas, em Macedo (1979), equações de regressão que permitiram reproduzir os teores de óxidos de elementos principais (SiO_2, TiO_2, Al_2O_3, Fe, MnO, MgO, CaO, Na_2O, K_2O), a partir dos teores de elementos traços (Ba, Co, Cr, Cu, Ni, Rb, Sr, V, Zn e Zr), todos analisados em 77 amostras, distribuídas de forma aproximadamente homogênea em área, de rochas basálticas da Bacia do Paraná. O estudo foi feito iniciando com os valores brutos, depois acrescentando valores logaritmizados, depois elevados à 2ª, 3ª e 4ª potências. Daí resultaram as tabelas 4 e 5, reproduzidas abaixo.

Tabela 4 : *Parâmetros de controle da regressão dos valores das proporções de óxidos de elementos principais em relação aos teores de elementos traços em rochas basálticas da Bacia do Paraná, tabela modificada de Macedo (1979).*
R = coeficiente de correlação entre valores observados e calculados;
F = estatística F de Fisher-Snedecor.

	1º Grau		1º e Log		2º Grau		3º Grau		4º Grau	
	R	F	R	F	R	F	R	F	R	F
SiO_2	95	169	96	174	96	189	96	189	96	189
TiO_2	89	67	92	77	92	78	94	53	94	58
Al_2O_3	51	6	62	7	68	7	63	7	70	8
Fe	95	137	96	104	96	88	96	87	97	73
MnO	81	33	81	27	84	25	85	25	85	27
MgO	92	100	93	88	93	93	94	81	94	81
CaO	94	149	97	161	97	161	97	121	97	122
Na_2O	75	23	77	16	74	22	76	16	96	16
K_2O	96	221	96	156	97	145	97	149	97	151

Tabela 5: *Equações de previsão para percentagens de óxidos de elementos principais em relação a teores de elementos traços em ppm, obtidas através de regressão múltipla, a partir de análises químicas de rochas basálticas da Bacia do Paraná.*

$SiO_2 = + 53,232 + 0,092Rb - 0,0085Cu - 0,011V - 0,0031Sr$
$TiO_2 = - 0,289 - 0,0011Sr + 0,0037V + 0,00504Zr - 0,0046Rb$
$Al_2O_3 = + 15,225 + 0,011Cu^2 - 0,040V^2 - 1.048Rb^2 - 1,405 \log Co - 0,231Zr - 0,031Sr^2 - 2,733 \log Sr - 0,525Ni^2$
$Fe = + 10,640 + 0,012V - 0,041Rb + 0,005Zr - 0,020Ni + 0,003Cu$
$MnO = + 0,221 + 0,0001V - 0,001Rb - 0,00047Ni + 0,00005Cu$
$MgO = + 7,359 + 0,0063Cr - 0,025Rb - 0,0067Zr - 0,013Co$
$CaO = + 11,486 + 0,0046Cr - 0,035Rb - 0,0066Zr - 0,00107Ba$
$Na_2O = + 2,844 + 0,00047Ba - 0,0056Ni + 0,00045Sr - 0,00072V$
$K_2O = + 0,125 + 0,017Rb + 0,015Zr + 0,00082Sr - 0,00051Cu$

As conclusões foram:

a - em quase todos os casos, os coeficientes de correlação entre os valores observados e calculados, bem como o erro padrão e a estatística F de Fisher-Snedecor mostram-se plenamente satisfatórios desde a equação linear

simples, que utiliza apenas valores de primeiro grau, tanto nas variáveis independentes quanto nas dependentes;

b - nas equações lineares, os parâmetros estatísticos são baixos para Al_2O_3 e razoáveis para Na_2O. Aumentando o grau das equações, melhoram os resultados para Al_2O_3 e não para Na_2O;

c - a não ser no caso do Al_2O_3, as pequenas melhoras nos parâmetros de regressão não justificam a utilização de equações de graus superiores ao primeiro.

d - no caso do Al_2O_3, a melhor equação é a de segundo grau, que combina simplicidade com parâmetros estatísticos aceitáveis.

Estes cálculos ilustram os procedimentos para escolha de equações de previsão, devendo-se notar que é preferível ter melhor conhecimento anterior sobre as distribuições de freqüência dos dados do que fazer um excessivo número de cálculos, o que foi o caso no trabalho comentado.

De forma análoga foram determinadas equações de regressão para os índices elaborados por de la Roche (1968) e Leterrier e de la Roche (1972), que haviam sido empregados por Rüegg (1975*a* e *b*) para estudar a geoquímica das rochas basálticas da Bacia do Paraná. Estes índices possibilitavam uma classificação das rochas da província, a partir dos dados químicos, mais concordante com os dados mineralógicos do que outros índices tradicionalmente utilizados.

Os índices de de la Roche (1968), calculados por regressão a partir dos teores de elementos traços, apresentaram altas correlações e significância com os dados observados para Al/3 - K (R = 91, F = 176), e menores para Al/3 - Na (R = 68, F = 16). Eles foram testados, plotando-os em um gráfico, feito da mesma forma que os originais, o qual permitiu chegar às mesmas conclusões que aquele feito com os originais.

Os índices de Letelier e de la Roche (1972), que permitem uma classificação mais apurada das rochas basálticas por dados geoquímicos, também foram calculados, com os resultados seguintes:

Diagrama	Índice	Fórmula	R	F
A	X	K - Na/2 - Ca	95	344
A	Y	Si/3 - K - 3Na/3 - Ca/3	78	22
B	X'	Si/4 - K - Na - Ca	89	54
B	Y'	Fe + Ti	94	139

Também estes índices obtidos por regressão foram testados graficamente, plotando-os em gráfico e comparando a classificação obtida com aquela resultante do uso dos dados originais.

A figura 18 ilustra um desses testes. Ela representa a linha de tendência de diferenciação B, que reúne o maior número de amostras entre os cinco grupos estudados, representando as amostras com maiores teores de Fe, Ti, Ca e Na.

Conforme Macedo (1979),

> "das 39 amostras... cujos índices caem dentro da área de influência desta linha, 17 apresentam índices em que os valores calculados pelas equações de regressão caem na mesma sub-área em que caem aqueles derivados dos valores observados, 10 que caem muito próximos da mesma sub-área, 11 que caem dentro da área de influência da linha de tendência, mas em sub-áreas diferentes e um índice calculado que cai no âmbito de uma linha diferente daquela em que cai o índice observado."

Macedo (1979) considerou, nas conclusões, que

> "foi possível prever com grande precisão a composição das rochas em elementos principais, utilizando apenas elementos traços" e que "a técnica da regressão múltipla permitiu elaborar índices de diferenciação de rochas basálticas baseados em elementos traços, utilizáveis em aplicações práticas".

Estas conclusões podem ser consideradas válidas, tendo em vista os resultados dos testes numéricos e gráficos. Fica porém a dúvida de se seriam realmente representativas as 77 amostras, para elaborar índices válidos para

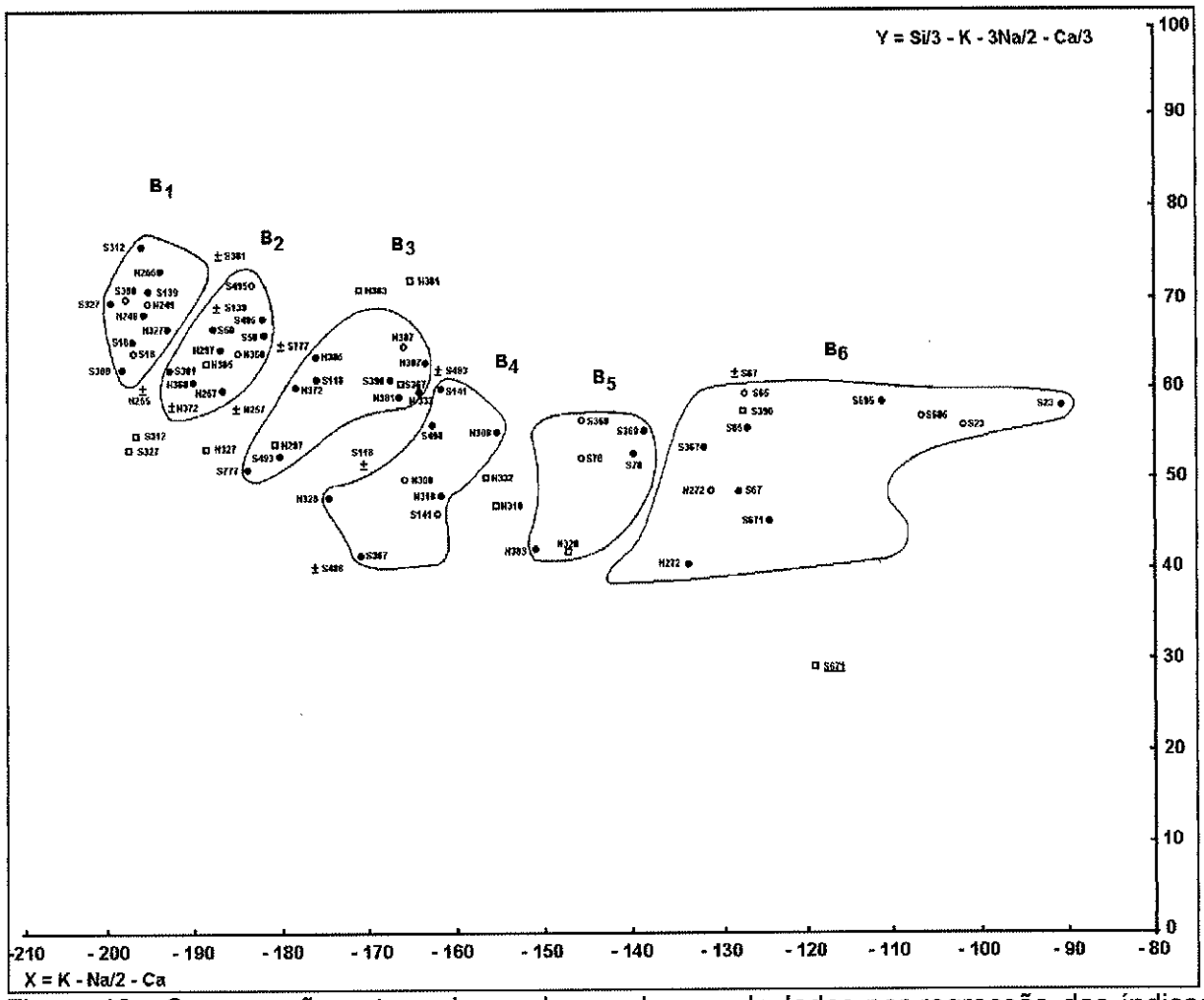


Figura 18 - Comparação entre valores observados e calculados por regressão dos índices de Leterrier e da La Roche para a linha de tendência B entre as análises de rochas basálticas da bacia do Paraná, no diagrama de Leterrier e de la Roche (1972), divididas em sub-áreas B_1 a B_6, conforme Ruegg (1975 a).Os pontos são lançados conforme a convenção:

- ● N312 - valor observado
- o N300 - valor calculado, caindo o ponto na área de influência da mesma linha e dentro da mesma sub-área que o observado.
- ± N300 - valor calculado, caindo o ponto na área de influência da mesma linha e e muito próximo da mesma sub-área que o observado, não podendo ser atribuído a outra sub-área.
- □ N300 - valor calculado, caindo o ponto na área de influência da mesma linha, porém no âmbito de uma sub-área diferente daquela em que caiu o valor observado;
- □ N300: (número sublinhado: valor calculado, caindo o ponto na área de influência de uma linha de tendência diferente daquela em que caiu o observado.

um conjunto tão grande e diferenciado quanto as rochas basálticas da província estudada. Novamente, e não pela última vez, fica a advertência da necessidade da validação geológica, tanto da lógica do tratamento de dados, quanto dos seus resultados.

6.2.4.2. Regressão da distância à mineralização a partir de teores de elementos principais e traços em rochas na área do Perau

Em Macedo (1986), relatado resumidamente também em Macedo (1989), a regressão múltipla foi apenas um dos diversos métodos usados para a interpretação de um estudo orientativo de prospecção litogeoquímica. Foram processados os teores de elementos principais (Fe, Mg, Ca, Ti) e traços (Ag, As, Au, B, Ba, Be, Bi, Cd, Co, Cr, Cu, La, Mo, Nb, Ni, Pb, Sb, Sc, Sn, Sr, V, W, Y e Zn), analisados por espectrografia ótica de emissão - EE, num conjunto de 78 amostras de rochas da área Mina do Perau (17 da lapa e 61 da capa do minério) e por espectrometria de absorção atômica - AA (Cu, Pb, Zn, Co, Ni, Sr, Rb, Al_2O_3, Fe_2O_3, FeO, TiO_2, SiO_2, MnO, CaO, MgO, Na_2O, K_2O, P_2O_5) num subconjunto de 25 amostras do Perau e 6 de Água Clara. Todas as análises foram executadas no LAMIN - Laboratório de Análises Minerais da CPRM.

Diversas vezes foi aplicada a regressão múltipla naquele trabalho:

a - Calculando os teores dos elementos analisados por absorção atômica, tomando os valores de sílica como variáveis independentes e utilizando os resíduos como teores corrigidos para outros estudos, com resultados comentados abaixo.

b - Calculando os teores dos elementos analisados por AA a partir dos analisados por espectrografia ótica. Os resultados variaram de ótimo (r^2 maior que 97%) para CaO e Pb a péssimo (r^2 menor que 20%) , para Na_2O e P_2O_5.

c - A principal aplicação da regressão múltipla no trabalho analisado foi o cálculo das distâncias à mineralização das amostras na capa e na lapa e de amostras provenientes apenas de anfibólios e de xistos da capa. Para os teores dosados por absorção atômica também foi testada a eficiência de dois processos de correção para a variação dos teores de acordo com o tipo de rocha. Um deles (assinalado com o sufixo N no símbolo do elemento na tabela

6), foi feito pela divisão dos teores brutos de um elemento analisado em um determinado tipo litológico pela média dos teores daquele elemento naquele tipo litológico, conforme a tabela de Beus (1976), para os elementos principais e a de Rose et alii (1979) para elementos traços.

O outro processo foi fazer a regressão dos teores de cada um dos elementos analisados por AA em relação aos teores de SiO_2, e empregar os resíduos da regressão em lugar dos teores brutos (assinalados com o prefixo R na tabela 6).

Em resumo, o trabalho analisado concluía que:

- as funções de regressão apresentam melhores índices de ajuste quando introduzidos na equação valores logaritmizados. Mesmo a distribuição dos elementos não sendo lognormal, a transformação reduzia a excessiva assimetria das distribuições de freqüência dos teores.
- as rochas analisadas por AA apresentaram melhores resultados do que as analisadas por EE, resultados interpretados como conseqüência conjunta da maior precisão e do fato das amostras escolhidas para análise por AA serem as menos alteradas intempericamente. Nestas, combinando os resultados por AA e espectrografia de emissão os resultados eram melhores.
- os conjuntos de amostras separadas por litologia, embora menores, deram índices melhores que o total das amostras analisadas por EE. Analogamente, os teores das amostras analisadas por AA apresentaram melhores resultados quando corrigidos por tipo litológico (N). O trabalho recomenda que a correção litológica seja sempre aplicada para interpretação de dados litogeoquímicos.
- Os resultados corrigidos por regressão de sílica (R) apresentaram resultados que variam de piores a muito pouco melhores que os obtidos com teores brutos para os mesmos conjuntos de dados. Isto indica que este tipo de correção somente pode ser bem empregado com rochas que integram uma mesma série magmática, não servindo para conjuntos de rochas de composição variada, como no caso em pauta.

Tabela 6 *- Resultados de regressões múltiplas dos valores da distância à mineralização do Perau em relação aos teores dos elementos analisados e suas transformações(L: dados logaritmizados; R: Resíduos da regressão dos valores de SiO_2; Q: análises quantitativas, por absorção atômica; N: teores corrigidos por tipo de rocha; segundo Macedo, 1986).*

Grupo (nº de amostras)	Variáveis testadas	R^2	Erro padrão	Variáveis na Equação
Lapa (17)	Fe - Zr	68,3	371	+La, -Co, +B
Lapa (17)	Fe - LZr	77,7	105	-Lsn, -Co, +LCu
Capa (61)	Fe - Zr	37,6	405	+Ti, +Sn. +Ni, -Sr
Capa (61)	Fe - LZr	47,7	375	-LPb, -LMn, +LTi, -LLa, +Sn
Anfibolitos(17)	Fe - Zr	60,8	355	-Sr, +Ti, -Fe, +Co, -B
Anfibolitos(17)	Fe - LZr	71,8	301	+LTi, -LPb, +Lzn, -Fe
Xistos (30)	Fe - Zr	66.1	256	+Be, +V, -Co, +Sn
Xistos (30)	Fe - LZr	66.1	256	+Be, -Co, +LSn
An. por AA (25)	Fe - Lzr; CuQ - Pf	73,4	212	+Be, -Ti, -Pf
An. por AA (25)	RFe - RZr	65,4	242	+Rsr, +Rbe, + Rcr, -RTi,
An. por AA (25)	RFe - Rzr; RCuQ - RPf	78,4	191	+Rsr, +Rbe, + Rcr, -RTi, +RFeQ, -RCo
An. por AA (25)	RCuQ - RPf	49,2	293	$-RTiO_2$, -RCoQ, +RZnQ, $+RAl_2O_3$
An. por AA (25)	FeN - ZrN	82,2	174	+BeN, -WN, +SnN, -TiN; +BaN
An. por AA (25)	FeN - ZrN; CuQN - P_2O_5N	86,2	152	+BeN, -WN, +SnN, -TiN; +NiQn, $-SiO_2N$
An. por AA (25)	FeN - ZrN (sem W) ; CuQN - P_2O_5N	67,4	234	+BeN, -SnN, +ZnQn, -TiN

- os resultados foram considerados suficientes para o cálculo da distância da mineralização ("erro padrão de 105 a 405 m, maioria por volta de 200 m"), advertindo o autor que "não é o objetivo deste trabalho fornecer uma equação que calcule magicamente a localização da mineralização, a qual teria sua aplicabilidade restrita aos dados com os quais foi calculada);
- o comportamento dos diversos elementos foi analisado, assinalando os que tinham seu teor aumentado (enriquecido) ou diminuído (empobrecido) com a aproximação da mineralização.

Os resultados foram combinados com os obtidos pelos outros métodos (correlações simples, análise discriminante) para escolha dos elementos a analisar para prospecção e serão comentados adiante.

6.2.5. Análise Discriminante

Segundo Davis (1986), o objetivo da análise discriminante é transformar o conjunto original de medidas em uma observação (amostra de rocha em nosso caso) em um índice discriminante único. A análise procura uma transformação que proporcione a máxima separação entre os dois grupos, ao mesmo tempo em que minimiza a variação dentro do grupo. Calcula-se a função discriminante, que fornece um índice discriminante (*discriminant score*) R, que resume a variabilidade de todas as variáveis empregadas na equação (teores, em nosso caso).

A eficiência da técnica para a separação entre os grupos de amostras próximas e distantes da mineralização foi testada usando dois indicadores: o índice lâmbda de Wilks (segundo Nie, Bent & Hull, 1970), que é inversamente proporcional à variância explicada pela separação: e a percentagem das amostras corretamente classificadas pelas funções discriminantes. Foram feitas análises usando os mesmos grupos estudados pela regressão (amostras da lapa e da capa, analisadas por EE, amostras analisadas por AA, amostras de xistos e anfibolitos separadamente), variando, em cada caso, a distância à mineralização entre 100 a 600 m, para montagem dos grupos, e em um caso, verificando a separação entre as amostras da capa e da lapa.

Os resultados foram positivos em todos os casos, com o índice lâmbda de Wilks apontando resultados significantes e percentagem de acerto na classificação igual ou próximo a 100% na maioria dos casos, com um mínimo de 80%.

As conclusões foram, em resumo:

- a distância de 200 m foi considerada ideal para separação;

- não houve ganho significativo na introdução de teores logaritmizados nas equações;
- a correção por tipo litológico foi menos proveitosa que no caso da regressão múltipla;
- os diversos elementos apresentaram comportamento individual, como participantes das equações discriminantes, semelhante ao apresentado para regressão.

6.2.6 Comentários sobre as conclusões gerais

Comparando os resultados de todos os métodos, foram apresentadas as conclusões resumidas abaixo:

- a análise de agrupamentos não se mostrou adequada para interpretação de resultados de prospecção litogeoquímica na área, com agrupamentos muito mais governados pelo tipo litológico que pela localização em relação à mineralização.
- os outros processos de tratamento dos dados (gráficos teor-distância, correlações simples entre teores e distância à mineralização, análises de regressão e discriminantes) todos permitiram distinguir os comportamentos dos grupos de elementos que sofrem enriquecimento ou empobrecimento relativo com a distância à mineralização. Os resultados são em geral concordantes. Os elementos Pb, B, Mg, Cu e K, na capa da mineralização, apresentam sempre enriquecimento e Sn, Zn, Be, Ni e Cr empobrecimento; outros elementos respondem melhor a um ou outro método e, outros ainda (Ti e Sr), apresentam resultados erráticos, mostrando enriquecimento num método e empobrecimento em outro. Embora houvesse poucas análises da lapa, La, B e Cu mostraram enriquecimento com a proximidade da mineralização nessa área.
- combinando gráficos de dispersão, os erros padrões da regressão e os resultados das equações discriminantes, foi considerada uma largura

segura do halo de dispersão da mineralização do Perau de 200 m, chegando a influência até a 400 m.

- para todos os métodos foi estudado o comportamento das amostras mais discrepantes, sendo as causas interpretadas como variação litológica (composição muito diferente dos xistos e anfibolitos predominantes) e alteração intempérica.

Estes resultados foram comparados com os de estudos semelhantes efetuados em outras áreas, apresentados na literatura, verificando que:

"os comportamentos apresentados por Cu, Pb, Ni, Co, Mn, K e Mg são semelhantes aos apresentados em jazidas semelhantes; Zn apresenta empobrecimento, ao contrário do que acontece na maioria dos casos estudados, e semelhante ao que acontece em três áreas tropicais, o que sugere que o comportamento discrepante pode ser explicado pela alteração intempérica; o Na apresenta enriquecimento, diferente da maioria e semelhante ao de jazidas distais, tipo a que provavelmente se filia a do Perau. O comportamento dos elementos, concordante em quase todos os casos com o das jazidas de caráter exalativo, reforça esta interpretação para a gênese do depósito do Perau.

A partir dessas conclusões, do estudo de superfícies de tendência, que será comentado no capítulo de análise espacial e da análise da mineralização e do halo secundário, não comentados aqui, Macedo (1986) chega às seguintes conclusões para prospecção:

a) as análises rápidas, efetuadas por espectrografia ótica de emissão, permitiram a detecção do halo geoquímico da mineralização do Perau em rocha quase tão bem como as análises de óxidos e elementos traços por processos mais precisos e muito mais caros. Uma análise de 30 elementos por EE custa *(preços de 1986)* Cz$ 89,40, uma AA Cz$ 13,70 para o primeiro elemento e Cz$ 4,70 pelos seguintes e uma análise quantitativa de 13 óxidos Cz$ 878,00, no Laboratório de Análises Minerais (LAMIN) da

CPRM. Os elementos analisados simultaneamente pelos dois processos apresentaram resultados quase idênticos, a não ser no caso do Zn, que só pode ser estudado através de análise por absorção atômica, devido ao limite inferior de detecção relativamente alto por EE (200 ppm). Os outros elementos com teores mais baixos que os limites de EE não são geralmente utilizados em pesquisas de sulfetos maciços. Dos elementos maiores que mostraram variação pela influência da mineralização apenas o potássio não é determinado pela espectrografia ótica de emissão. A correção dos teores por resíduos em relação à sílica não mostrou resultados que compensassem o aumento de custos.

b) os processos de tratamento de dados utilizados apresentaram resultados em sua maioria concordantes, a não ser no caso da análise de agrupamentos, cujos resultados não indicaram seu aproveitamento para tratamento de dados de prospecção litogeoquímica na área. As amostras com maiores resíduos na regressão da distância à mineralização são em geral as mesmas mal classificadas pela análise discriminante.

c) considerando uma largura do halo de 200 m, seguramente detectada neste trabalho, o espaçamento de amostragem deve ser de aproximadamente 100 m, para obter para cada perfil, em média, duas amostras anômalas ao se cruzar a área mineralizada.

d) a correção pelo cálculo de teores relativos, dividindo cada teor pela média de seu tipo litológico, foi a que se mostrou mais eficiente, particularmente na definição dos halos pelos gráficos teor-distância.

e) o custo da prospecção litogeoquímica, na região do Perau, considerando um espaçamento aproximado (controlado pela existência de afloramentos e variação litológica) de 100m, é aproximadamente o mesmo da prospecção geoquímica de solos, considerando-se, para um espaçamento de 50 m e o mesmo tipo de análise, uma produção diária de amostras de solo duas vezes maior que a de amostras de rocha e técnico coletor para rochas mais habilitado, no mínimo geólogo iniciante, com salário aproximadamente o dobro daquele do técnico de mineração do nível médio, coletor de amostras de solo. Haverá melhor aproveitamento do trabalho, tanto do ponto de vista econômico quanto do técnico, se a coleta for efetuada pelo mesmo geólogo

que fizer o mapeamento, e integrada com este. Não existem dados suficientes para comparar os custos de detecção por sedimentos de corrente e rocha, na área do Perau.

Daí derivavam as seguintes recomendações para prospecção:

a) efetuar coleta em perfis perpendiculares às estruturas principais, na área da Formação Perau, preferivelmente em sua base, onde ela puder ser definida por indicações de superfície.

b) o espaçamento deverá ser de aproximadamente 100 m no perfil, procurando amostrar os principais tipos litológicos encontrados. Considerando a necessidade da identificação das rochas, a coleta deve ser feita por geólogo, e integrada aos trabalhos de levantamento geológico.

c) tendo em vista a numerosa suite de elementos que pode mostrar variações com a mineralização, a análise deve ser multielementar, por processos rápidos, para o total das amostras. A precisão das análises por espectrografia ótica de emissão é suficiente para os trabalhos de rotina. Nas áreas definidas como potencialmente anômalas, as amostras deverão ser reanalisadas para os elementos-minério por processo de maior precisão (absorção atômica, raios X ou plasma).

d) o processamento dos dados deve ser feito inicialmente por gráficos de teor-distância e cálculo de estatísticas básicas de teores brutos e relativos (corrigidos por divisão pelos teores médios para cada tipo de rocha).

e) para as áreas julgadas em primeira instância como anômalas deverá ser feito processamento dos dados por análise discriminante e elaboração de gráficos teor-distância com os elementos indicadores combinados em índices compostos.

f) para detalhamento, considerando os bons resultados de ambos os meios de amostragem, deverá ser feita coleta de solo e rocha, por geólogo, em perfis com malha mínima de 50 m, e análise de Cu, Pb, Zn, B, Mg, K, Fe, Mn e Ca por processos quantitativos. Os três primeiros por sua ligação direta com o minério, os três seguintes por serem indicadores da proximidade da mineralização, Fe e Mn pelo mesmo motivo e também para controle da

absorção de elementos traços em solo ou produtos de alteração e Ca para indicação dos níveis carbonáticos, hospedeiros da mineralização.

Estas conclusões se mantêm após a revisão do trabalho ?

Os cálculos estatísticos indicam realmente dois conjuntos de elementos que aparentam comportamento diferente em relação à proximidade da mineralização. Isto é mostrado pelo comportamento das correlações, das regressões múltiplas e da análise discriminante. Não é possível fazer uma afirmação estatística rigorosa, pois os conjuntos de dados estudados não apresentam, em sua maioria, normalidade ou lognormalidade, exigidas pela maioria dos testes. No caso particular da regressão, isto não seria um problema insolúvel, pois, pelo teorema de Gauss-Markov, citado em Agterberg (1974), a regressão permite a previsão em casos de não-normalidade, desde que os resíduos sejam aleatórios e não correlacionados à variável calculada. Em Macedo (1979) os gráficos de dispersão dos resíduos foram analisados, o que não feito no trabalho de 1986, pelo grande volume de cálculos efetuados. Não existe então, para os resultados deste último, a garantia da ausência de correlação entre resíduos e a variável calculada, comprometendo a aplicabilidade da regressão.

Alguns problemas não foram adequadamente tratados e apareceram nesta revisão:

- Em alguns casos foram incluídos, nas equações de regressão e discriminantes, elementos representados por poucos teores superiores ao limite de detecção. Isto é uma conseqüência do método de análise da maioria da amostras (espectrografia ótica de emissão), o qual, para muitos elementos, tem o limite de detecção muito alto. Havia, na época do trabalho, a preocupação de aproveitar totalmente os dados, em número não muito grande (78 amostras). Teria sido preferível um conjunto menor, mas com resultados mais confiáveis.
- Nos casos de poucas amostras com teores superiores ao limite de detecção e em alguns outros, como o da prata, já comentado, valores discrepantes influíram aumentando os valores dos coeficientes de correlação de Pearson.

- Por falta de programas adequados (como o apresentado em van den Boom (1984), que faz análise discriminante não-paramétrica, usando uma variação do teste de Kolmogorov-Smirnof, que foi solicitado ao seu autor e não adquirido por problemas burocráticos para transferência de divisas, difícil na época), foram quase sempre aplicados métodos multivariáveis paramétricos, que exigiam normalidade para sua aplicação correta.

As equações de regressão realmente produzem gráficos índice-distância, montando os índices de acordo com as equações, que indicam maiores valores com a proximidade da mineralização. Esses gráficos não foram incluídos nem mesmo na tese de 1986, pois sua validade é restrita aos conjuntos de dados usados para elaborá-los.

Procurando testar a generalidade das conclusões anteriormente apresentadas foram elaborados nesta revisão gráficos somativos, montando índices a partir da diferença entre a soma dos teores normalizados (valores de **z**) dos elementos considerados como indicadores da proximidade e a soma dos teores normalizados daqueles decrescentes com a proximidade da mineralização. Daqueles reconhecidos como enriquecidos, Pb, B, Mg, Cu e K, o potássio não foi utilizado, por ser representado por poucos teores; o mesmo aconteceu com Sn e Zn entre os empobrecidos, dos quais ficaram nos gráficos Be, Ni e Cr. Todos eram considerados como exibindo esses comportamentos na capa da mineralização, enquanto B e Cu também apresentavam enriquecimento na lapa. Foram feitos gráficos com os teores brutos e com aqueles corrigidos por tipo de rocha, representados pela razão entre o teor bruto e o teor médio para cada tipo de rocha, conforme já explicado. Estas razões também foram normalizadas por **z** (sendo-lhes subtraída a média das razões e o resultado dividido pelo desvio-padrão).

Os resultados são apresentados nas figuras 19 a 21. O primeiro gráfico, da figura 19, descreve o comportamento dos dados brutos, e foi suavizado por uma média móvel entre três dados contíguos. Os picos maiores na capa são aqueles próximos à mineralização e os da região dos 600 m, na área das cálcio-silicáticas, também aparentes nos gráficos individuais de alguns elementos (Pb, Ca, B). Aparecem outros picos menores, mas há uma tendência muito clara em direção à mineralização.

Um comportamento muito semelhante apresenta o gráfico feito com teores corrigidos por rocha (figura 20). Os picos são os mesmos; as amostras mais distantes, exibem uma queda menos acentuada que nos teores brutos.

As diferenças de comportamento entre teores brutos e corrigidos por rocha são comparadas na figura 21. Confirma-se não haver grande ganho, neste caso, na correção por tipo de rocha, ao contrário do que acontece nos gráficos teor-distância dos elementos individuais (não exibidos aqui).

Este resultado é semelhante ao daquele do teste gráfico prático usado em Macedo (1979) e apresentado na figura 18. Ele mostra ser possível detectar a região mineralizada a partir do gráfico da capa da mineralização, com os elementos considerados como enriquecidos e empobrecidos em Macedo (1986). No entanto, a tendência não é tão clara quanto o que seria de esperar pelas conclusões da tese, que se referem aos diversos métodos.

A razão mais provável para um resultado não muito generalizável é que a amostra estatística utilizada (78 amostras de rocha, sendo apenas 61 da capa), fosse muito pequena para que se tirassem dela todas as conclusões extensamente reproduzidas acima. Este número é bastante diferente do de treze mil amostras (115.000 determinações), usado por Govett e seus rientandos no distrito de Bathurst, New Brunswick, Canadá, para estudos orientativos para prospecção litogeoquímica de sulfetos maciços (Govett, 1983).

Num balanço final sobre as conclusões de Macedo (1986), conclui-se que os resultados positivos na detecção da mineralização, nos diversos métodos, foram combinados na seleção dos elementos considerados como enriquecidos ou empobrecidos, sendo a eles atribuída uma generalidade maior que aquela permitida pelos dados disponíveis.

No caso de serem retomadas as pesquisas na área, a qual é abandonada e retrabalhada em ciclos de mais ou menos 20 anos, seria interessante fazer uma nova coleta, mais extensa, e submeter as amostras a análises mais precisas, o que daria resultados mais confiáveis como estudo orientativo para prospeção geoquímica.

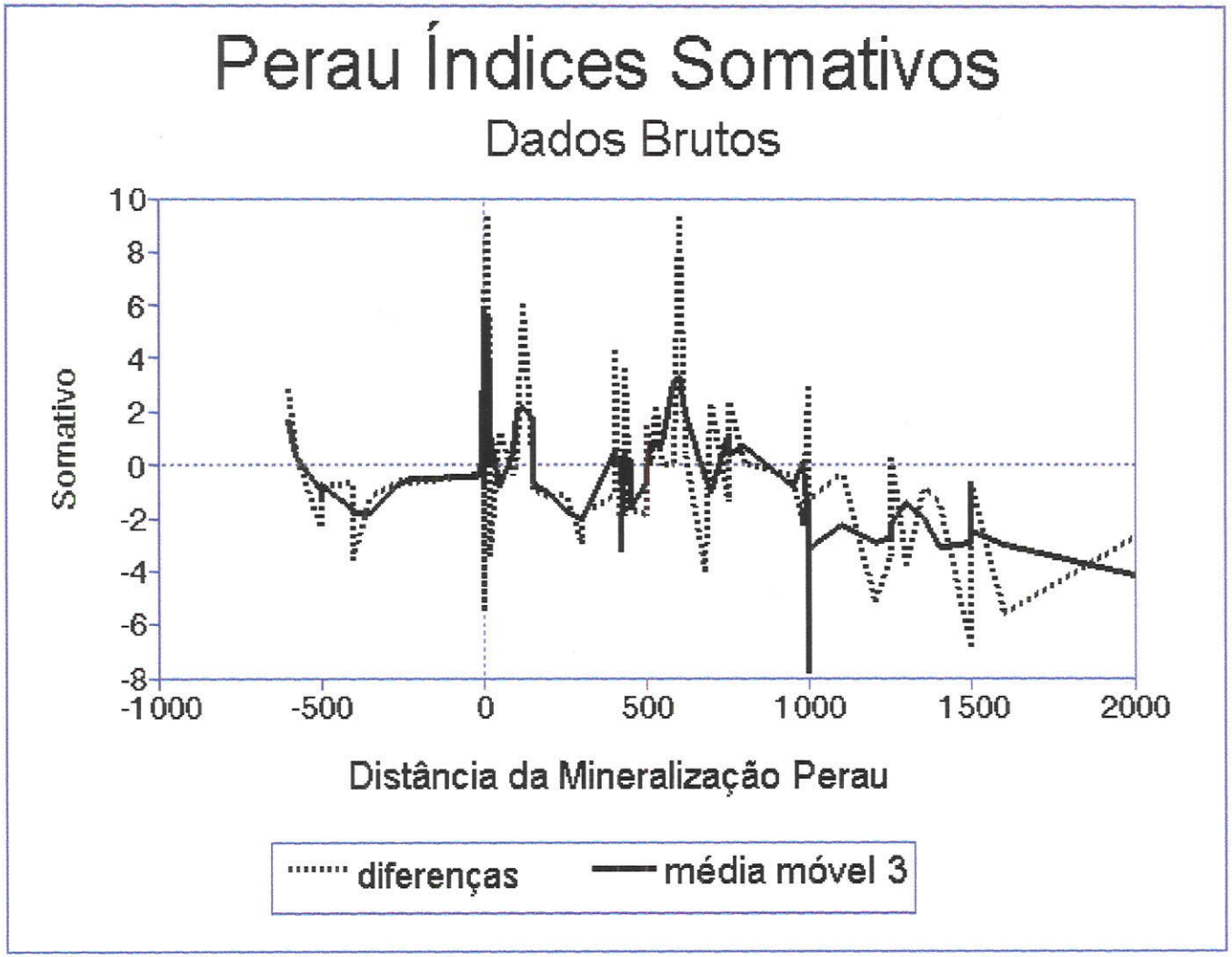


Figura 19 - Gráficos teor-distância somativos, representando as diferenças entre as somas de teores normalizados de elementos considerados como enriquecidos com a proximidade da mineralização, e as de elementos considerados como empobrecidos, para 77 amostras de rocha da área Mina do Perau. A linha pontilhada representa as diferenças; a linha cheia a suavização da anterior pela média móvel entre três teores contíguos.

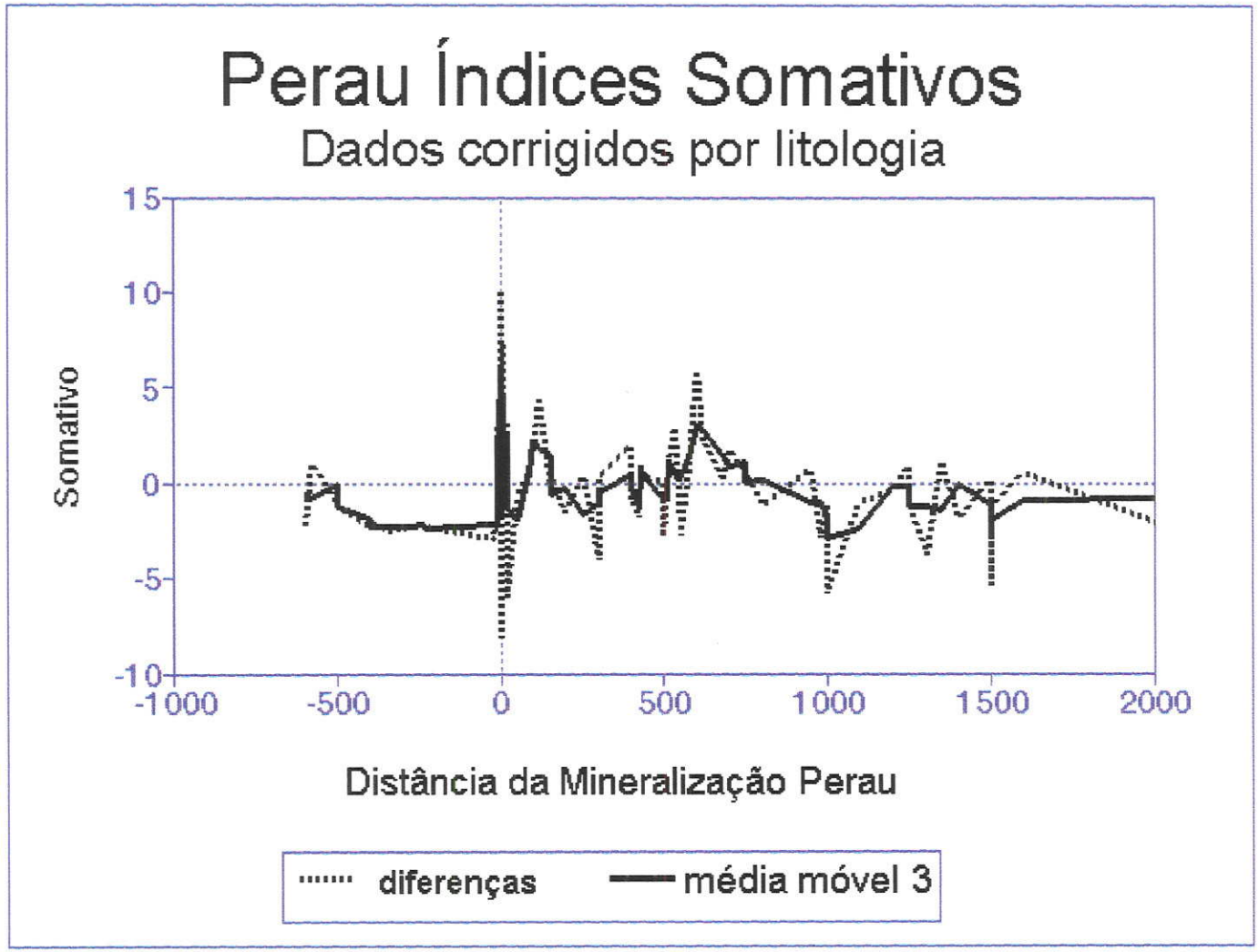


Figura 20 - Gráficos teor-distância somativos, feitos com teores corrigidos por litologia, representando as diferenças entre as somas de teores normalizados de elementos considerados como enriquecidos com a proximidade da mineralização e as de elementos considerados como empobrecidos, para 77 amostras de rocha da área Mina do Perau. A linha pontilhada representa as diferenças; a linha cheia a suavização da anterior pela média móvel entre três teores contíguos.

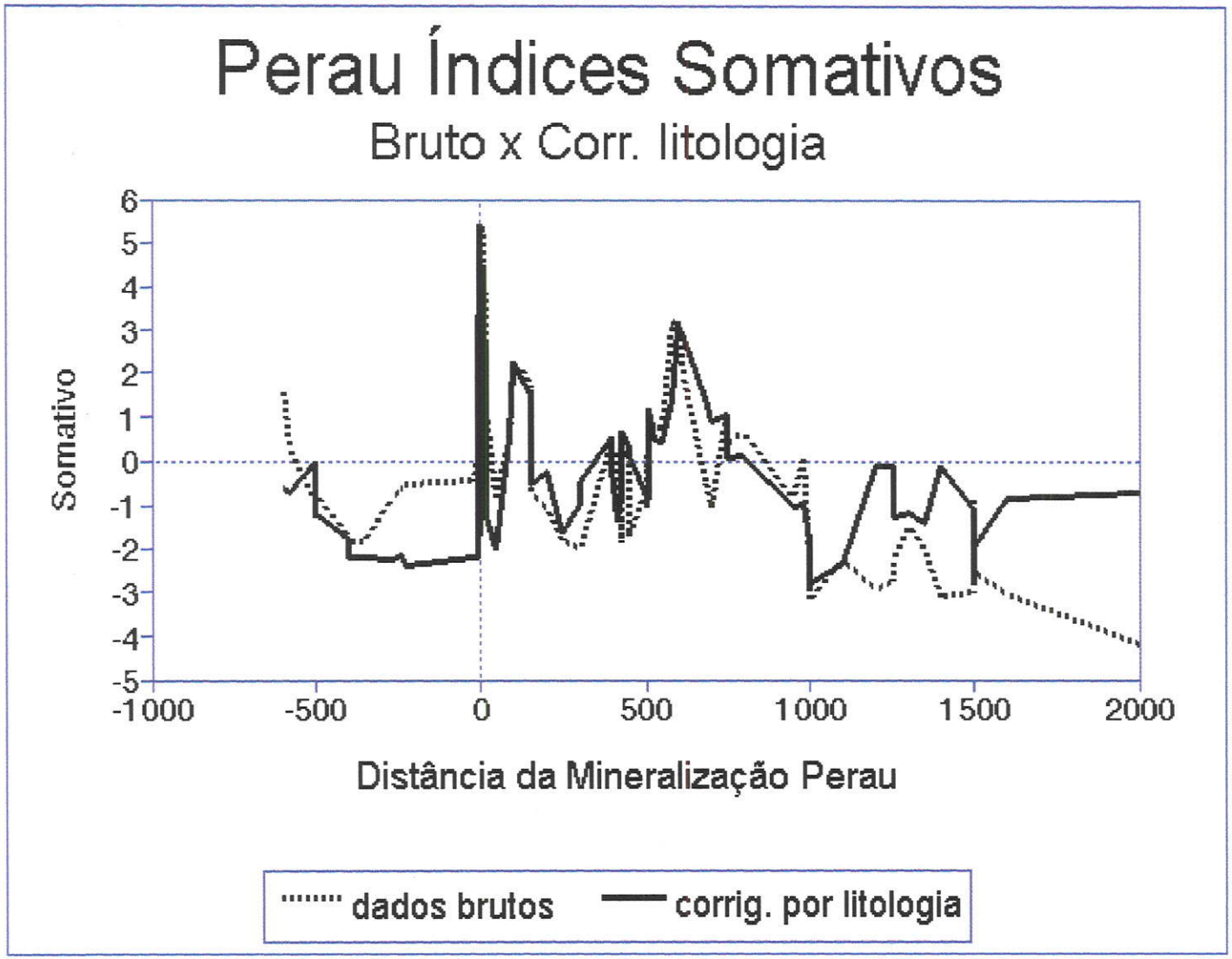


Figura 21 - Comparação entre gráficos teor-distância somativos, feitos com teores brutos e teores corrigidos por rocha, representando as diferenças entre as somas de teores normalizados de elementos considerados como enriquecidos com a proximidade da mineralização e as de elementos considerados como empobrecidos, para 77 amostras de rocha da área Mina do Perau. Ambos são suavizadas por média móvel entre três teores contíguos.

7. VARIAÇÃO ESPACIAL - GEOPROCESSAMENTO

O trabalho geológico, que inclui o geoquímico, é essencialmente espacial. Nestas áreas de conhecimento, uma informação numérica, por melhor que seja tratada, não tem sentido se não for associada a uma informação espacial. Apenas na década de 80 o aumento da capacidade dos microcomputadores e sistemas periféricos de entrada e saída de dados permitiu o tratamento de dados gráficos em computador e a apresentação dos resultados, a custos compatíveis com o de outras fases do trabalho geoquímico. Foram então desenvolvidos os sistemas de Geoprocessamento, programas e técnicas de trabalho que permitem empregar os equipamentos para o tratamento de dados espaciais.

Existem dúzias de definições de ***Geoprocessamento*** e de ***Sistema de Informações Georreferenciadas*** (geralmente expresso como **SIG**, ou como **GIS**, pelas iniciais em inglês). A confusão começa com o nome (Sistema de Informações Geográficas, Sistema de Informação Geográfica, Sistema de Informações Georreferenciadas, Sistema Geográfico de Informações?).

O assunto foi revisado por Teixeira et alii (1995), que propõem sua "melhor definição de SIG": "*Conjunto de programas, equipamentos, metodologias, dados e pessoas (usuários), perfeitamente integrados, de forma a tornar possível a coleta, o armazenamento, o processamento e a análise de dados georreferenciados, bem como a produção de informação derivada de sua aplicação*".

Talvez a definição mais prática seja a de Bedell (1994): "um *SIG é qualquer sistema que trata dados espaciais*". Ele especifica: "*é uma forma particular de sistema de informações aplicadas a dados geográficos; um sistema é um grupo de entes e atividades que interagem por um propósito comum; um sistema de informações é um conjunto de processos, executados sobre dados brutos, para produzir informações úteis para decisão; um sistema de informações geográficas usa dados referenciados geograficamente e dados não espaciais e inclui operações que permitem análise espacial*". Outro

geólogo pioneiro na área, Bonham-Carter (1994), define SIG como um "*sistema de computador para tratamento de dados espaciais*".

Em nosso grupo de pesquisa preferimos a denominação **Geoprocessamento**, conforme definida por Rodrigues (1990): "*conjunto de tecnologias de coleta e tratamento de informações espaciais e de desenvolvimento, e uso de sistemas que as utilizam*". Assim fazendo, consideramos que o nome Geoprocessamento enfatiza mais o processamento que o armazenamento de informações ou o desenvolvimento de técnicas de computação, sendo mais coerente com o nosso trabalho.

Câmara e Freitas (1994) consideram "*a característica fundamental de um sistema de Geoprocessamento a faculdade de armazenar, recuperar e analisar* ***mapas*** *num ambiente computacional*". Bonham-Carter (1994) define as funções de um SIG: organização de dados, sua visualização, consulta espacial, combinação, análise e predição. Pode-se acrescentar, acompanhando Eastman et alii (1995) e o próprio Bonham-Carter (com. pessoal), o apoio à decisão, e como fase final, a produção cartográfica. A seguir encontra-se um resumo das técnicas mais comuns, juntamente com exemplos e reflexões sobre seu uso nos trabalhos do autor.

É importante lembrar que, ao contrário dos trabalhos de tratamento de dados não-espaciais, que são individuais, ou, se em colaboração, o autor neles realizou todo o processamento, os trabalhos a seguir comentados foram quase sempre em colaboração, contando com a participação ativa de alunos e bolsistas, sem os quais seria impossível executar os enormes trabalhos de conversão e edição de dados. Nos trabalhos mais recentes, e naqueles em andamento, os membros do grupo realizam não apenas essas funções, mas também os trabalhos de análise.

A maioria dos artigos aqui comentados é produto de um programa de trabalho iniciado em 1991, destinado a introduzir a tecnologia de Geoprocessamento no IG-USP, através de trabalhos com finalidade principalmente metodológica. Partindo dos dados da folha Itararé (SG.22-X-B) do Projeto Mapas Metalogenéticos e Provisionais (DNPM/CPRM), apresentados em Hama e Algarte (1996), e depois incorporando dados e interpretações mais modernas do grupo do IPT (Rondinelli et alii, 1986;

Campanha, 1991 e 1995; Campanha et alii, 1988, 1996; Bordignon e Etchebehere, 1992), e trabalhando em conjunto com esse grupo para a digitalização dos dados, diversos trabalhos foram produzidos e apresentados, visando principalmente a análise metalogenética regional por geoprocessamento (Macedo, 1993*b*, 1994, Macedo et alii, 1993*a* e *b*, e 1995*a* e *b*; Araújo et alii, 1995; Salles et alii, 1995). Esse plano de trabalho continua em andamento, com três mestrandos, seis bolsistas e seis estagiários desenvolvendo novas aplicações de Geoprocessamento para Recursos Minerais e Meio Ambiente. Apenas duas das aplicações, que podem ser mais rigorosamente classificadas como de tratamento de dados geoquímicos, são comentadas do decorrer deste capítulo.

7.1 Organização

Nesta fase do Geoprocessamento se monta o banco de dados, fazendo-se a aquisição de dados de campo, arquivo e laboratório, sua crítica para eliminação de erros e certificação de qualidade, conversão à forma digital e montagem do banco de dados. Os princípios, procedimentos operacionais e experiências desta fase foram relatados nos capítulos 3 e 4. Alguns pontos fundamentais devem ser lembrados, mesmo à custa de repetição:

- cada conjunto de dados e cada produto de uma fase de trabalho deve ser conferido e testado exaustivamente antes da fase seguinte;
- os programas escolhidos para a organização do banco de dados devem permitir fácil comunicação com outros programas, sob pena ter de recomeçar o trabalho ao mudar o sistema;
- para os dados espaciais, devem ser mantidos os arquivos em forma vetorial, que preservam integralmente a precisão dos dados, conforme coletados, e os arquivos raster, nas resoluções espaciais mais convenientes para análise e visualização.

7.2 Visualização

7.2.1 Generalidades

A visualização dos dados e resultados é parte fundamental do geoprocessamento, tanto durante a análise quanto para apresentação das conclusões em tela ou como produtos em papel.

Nos primórdios do tratamento de dados espaciais em computadores a visualização se confundia com a produção cartográfica. Os computadores acessíveis aos pesquisadores brasileiros nas décadas 60 e 70 não tinham monitor, ou tinham apenas um monitor ligado a um "console", que controlava a operação da inacessível máquina, restando aos mortais comuns fazer a entrada de dados por cartões e fitas magnéticas e a saída exclusivamente por impressora "de linhas", que imprimia uma linha completa de cada vez, com os caracteres gravados em cilindros colocados lado a lado. O tamanho dos cilindros e a necessidade da operação rápida limitava o número de caracteres disponíveis, sendo o alfabeto apenas em maiúsculas e sem acentos. Mesmo assim já se fazia algum geoprocessamento.

Howarth (1983) apresenta os resultados de tentativas heróicas de conseguir boas representações gráficas com meios limitados. Em Macedo (1986) as análises de superfícies de tendência foram feitas pelo programa SYMAP, um antepassado dos programas de geoprocessamento atuais, fornecendo resultados bem pouco visíveis.

O processamento distribuído e os mini-computadores (como o PDP-11 instalado no CCE na década de 80) permitiam uma visualização limitada, por monitores primitivos, então conhecidos como CRT (cathode ray tubes). Com os microcomputadores, os monitores melhoraram, principalmente porque um grande mercado de entretenimento pagava caro por figuras bem definidas e coloridas. No final da década de 80 difundiram-se as GUI (graphic user interfaces), começando com os computadores Apple e passando aos do tipo IBM-PC, firmando-se a Windows como padrão de quase todo o mercado.

James Elliot, pioneiro em integração de dados geológicos do USGS, relatou, em palestra no IG-USP, que os primeiros trabalhos de integração de

sensoriamento remoto, geoquímica, geologia e geofísica tinham que ser processados no EROS Data Center, única instalação nos EUA com capacidade de processamento e visualização adequadas no início da década de 80, usando programas caríssimos, em alguns casos desenvolvidos especialmente para a função. Os mesmos trabalhos podem agora ser feitos em um microcomputador de tamanho "notebook" com tela colorida, com um programa de US $ 500 (IDRISI).

Esta facilidade de visualização, aliada à enorme capacidade de processamento dos processadores atuais (Pentium, por enquanto), permitem aos geólogos realizar seus sonhos de combinar dados, mudar escalas e projeções, testar interpretações e ver os resultados quase instantaneamente. Mesmo assim o geoprocessamento ainda não está devidamente difundido, e, mesmo quando usam computador para mapas, muitos geólogos apenas o utilizam como ajudantes de desenhista, com programas de CAD (Computer Assisted Design), inadequados mesmo para os desenhos geológicos e que não possibilitam análise. Um grande trabalho de divulgação ainda deve ser feito, para que os geólogos percebam o imenso progresso na capacidade de trabalho que lhes dá a utilização dos Sistemas de Informações Georreferenciadas.

7.2.2 Visualização a duas, duas-e-meia e três dimensões

A visualização a duas dimensões é fácil e imediata, associando-se cores aos atributos não espaciais, e atribuindo essas cores aos pixels (***pic***ture ***el***ements) das imagens, de acordo com as dimensões e a localização dos atributos espaciais. Mesmo que os dados sejam armazenados de forma vetorial, é preciso transformá-los em raster para a visualização. As figuras deste texto, tanto oriundas do IDRISI quanto de programas de desenho (Paintbrush, Adobe Photoshop), são armazenadas em raster. Na maioria dos casos os programas permitem que uma cópia da tela (*screen dump*) seja reproduzida em papel, numa forma fácil e rápida de produção cartográfica, limitada pela escala. Uma forma melhor de visualização e produção cartográfica é gerar a figura a partir de vetores, montando em raster apenas a

tela ou a imagem para impressão, a exemplo dos programas Corel Draw e Tmapper, este especializado para edição de vetores gerados no IDRISI.

Um problema mais complexo é a visualização de três dimensões. Deve-se aqui distinguir duas situações, às vezes confundidas. Uma é o caso de matrizes de dados que têm em cada ponto apenas um valor para a terceira dimensão. É o caso das superfícies topográficas, ou das superfícies do topo de corpos geológicos. Outra é a situação onde a terceira dimensão tem valores variáveis, como por exemplo os teores de um corpo de minério, que variam continuamente na vertical como nas outras dimensões. Outro exemplo é o dos teores de poluentes em aqüíferos subterrâneos. O primeiro caso é chamado em geoprocessamento de **duas e meia dimensões**, o terceiro de **três dimensões verdadeiras.**

Neste texto serão tratados apenas os métodos e experiências do autor em dados com duas e meia dimensões. Espera-se que, com os novos programas e equipamentos adquiridos pelo IG-USP, logo haverá na equipe uma competência no tratamento de dados a três e até quatro dimensões. Esta última situação ocorre também nos estudos de poluição, quando são calculados e visualizados os diversos estágios da propagação dos poluentes no espaço e no tempo.

Para que os dados a duas ou mais dimensões sejam visualizados, e também para sua utilização no cálculo de volumes e teores de reservas minerais, ou de locais poluídos, é necessário, como primeiro passo, transformar as informações brutas (por exemplo cotas, teores), com suas localizações, numa rede de pontos (*grid* ou *lattice*) igualmente espaçados, com os valores da terceira dimensão interpolados. Este modelo é freqüentemente denominado MNT (modelo numérico de terreno) ou MDT (modelo digital de terreno), pois sua aplicação inicial foi a dados topográficos. Os métodos de cálculo para a produção do *grid* e de curvas de contorno que facilitam a visualização fogem ao escopo desta revisão, e serão comentados apenas o necessário para o entendimento dos trabalhos efetuados. Boas introduções a eles e a suas aplicações podem ser encontradas em Howarth (1983), Davis (1986) e Yamamoto (1986, 1988 *a* e *b*, 1991 e 1995).

Nos primeiros trabalhos do autor a interpolação era feita por triangulação, com desenho manual das curvas; depois auxiliada por computador, mas com produção manual dos mapas finais (por exemplo em Ferreira et alii, 1981).

Em Barbour e Macedo (1991) são apresentados diagramas tridimensionais desenhados em computador, usando o programa SURFER (Golden Software, versão 3). A pouca velocidade do PC utilizado, que obrigou ao uso de *grids* com poucos elementos, e a impressão por impressora matricial, resultaram em figuras pouco visíveis.

A partir de 1991, utilizando o programa IDRISI (Eastman, 1995), e combinando-o com programas de desenho computadorizado (Corel Draw, Paintbrush), foi possível mostrar em tela e em papel os resultados dos cálculos de MNT, tanto para dados topográficos quanto para dados geoquímicos.

A figura nº 22 ilustra a forma com que o MNT é visualizado em tela, para resultados da área Mina do Paqueiro. Esta figura foi gerada, usando o módulo Ortho, a partir do MNT calculado diretamente no módulo Intercon do IDRISI, com o algoritmo de inverso do quadrado da distância, sendo os dados digitalizados por curvas de nível.

É necessário cuidado na utilização dos modelos tridimensionais (ou bi-e-meio-dimensionais). Em primeiro lugar, conforme o algoritmo utilizado e os seus parâmetros de controle, redes muito diferentes podem ser geradas com os mesmos dados. Na forma mais popular de cálculo do *grid,* inverso da potência da distância, a variação no valor da potência resulta numa representação do relevo mais ou menos acidentada, que pode facilitar ou dificultar a interpretação dos dados. O raio de procura, dentro do qual são aproveitados os pontos para o cálculo do valor em um dos nós da rede; o número de pontos utilizados no cálculo, e a possibilidade de usar a média de pontos muito agrupados num quadrante ou octante da área de procura também devem ser levados em conta e ajustados conforme as necessidades da pesquisa. Alguns programas (como o SURFER) permitem estes ajustes e a utilização de outros algoritmos (como a krigagem e a triangulação). Outros são rígidos (IDRISI, com inverso da potência sem ajustes, e ARC/INFO (ESRI, 1990), com triangulação). Na figura 22 nota-se o efeito do cálculo por um algoritmo sem opções, no caso o do inverso do

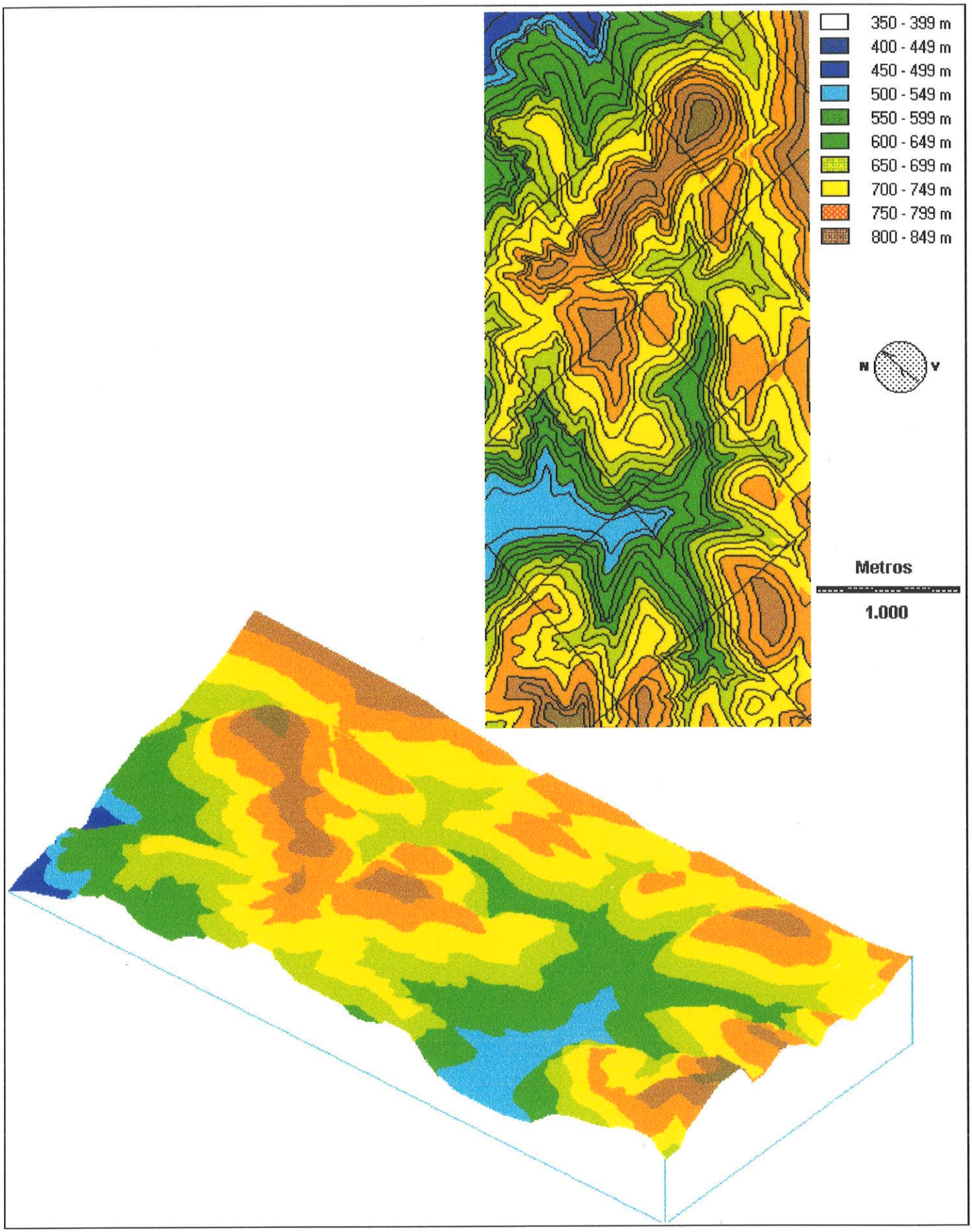


Figura 22: Modelo Numérico de Terreno, área Mina do Paqueiro, em bloco-diagrama e mapa com curvas originais sobrepostas, calculado pelo IDRISI

quadrado da distância, com procura dos seis pontos mais próximos. Ocorrem defeitos, como por exemplo, duas elevações, na parte central direita do modelo, representadas pela área de coloração marron claro, indicando 750 a 799 m, entre, as quais ocorre uma outra pequena elevação de forma arredondada, que não tem correspondente nas curvas utilizadas para o cálculo. Os autores do programa, em sua *newsletter* mais recente, prometem melhores algoritmos para a próxima versão.

Em alguns casos consegue-se melhor resultado fazendo o cálculo do MNT num programa de maiores recursos e transportando os resultados para um programa de geoprocessamento, para integração com outros dados e visualização. Isto foi feito com os dados topográficos da área Mina do Paqueiro, com os resultados apresentados em Salles, Macedo e Liotte (1995).

A figura 23 mostra um MNT calculado pelo programa SURFER, usando o mesmo algoritmo empregado na figura anterior, mas limitando o raio de procura e ajustando o efeito da concentração dos pontos por quadrantes. O relevo se apresenta suavizado, em relação ao expresso no modelo da figura 22, tanto pela forma de cálculo original, quanto porque só é possível, na presente versão do IDRISI para Windows, importar até 16.000 pontos, que devem ser reinterpolados para modelos de maiores dimensões. Neste caso houve perda de detalhe na representação do relevo.

A própria visualização apresenta variadas opções e necessita cuidado. O diagrama tridimensional pode ser apresentado com e sem sobre-elevação, o que pode levar a exagerar ou minimizar as variações do relevo topográfico ou geoquímico. Se a visualização é feita como relevo sombreado, a duas dimensões, o azimute e o ângulo de elevação da iluminação podem também alterar a interpretação.

A imagem que é sobreposta ao modelo tridimensional também dá sua contribuição ao efeito final do diagrama. Os dispositivos de visualização (monitores) e impressão (plotters e impressoras) têm limitações de resolução (número de pontos por unidade de medida) e de número de cores que podem exibir. Um MNT expresso em números reais deve ser reclassificado para uma escala de 15 ou 255 intervalos, que serão expressos em cores na figura "vestida" sobre o bloco-diagrama tridimensional. Os diferentes valores de corte que podem ser escolhidos para o "fatiamento" do modelo podem dar efeitos

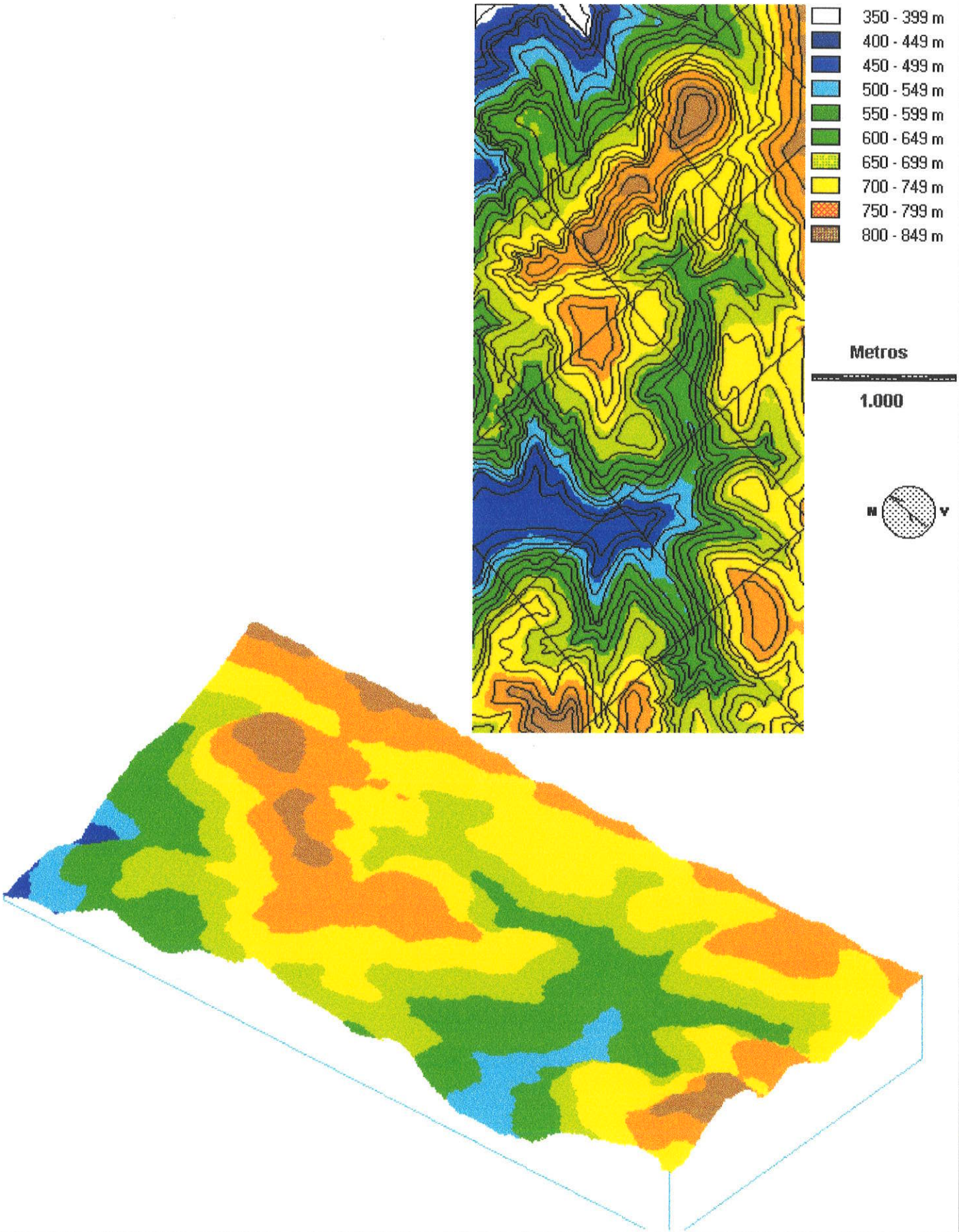


Figura 23: Modelo Numérico de Terreno, área Mina do Paqueiro, em bloco-diagrama e mapa com curvas originais sobrepostas, calculado pelo SURFER e convertido para IDRISI

muito diferentes. A escala de cores escolhida também pode acentuar diferentes intervalos do universo de valores da terceira dimensão.

Assim sendo, só é possível interpretar corretamente um diagrama de MNT quando os parâmetros de cálculo e exibição dos resultados são conhecidos.

7.3 Consulta espacial

As consultas ao banco de dados permitem obter, de forma gráfica ou tabular, informações classificadas, tais como: localização de anomalias geoquímicas, localização de ocorrências classificadas por tamanho, localização de ocorrências até uma dada distância de um lineamento, etc.

Os programas mais simples (como IDRISI para DOS) fornecem apenas respostas únicas à consulta espacial (apenas na camada que está sendo exibida na tela) e exigem o uso de um gerenciador de banco de dados independente (Dbase IV, no caso do IDRISI para dos) para consultas mais complexas.

Os programas mais avançados (por exemplo o próprio IDRISI para Windows, o ARC/INFO e muitos outros) permitem ligação dinâmica dos atributos espaciais com bancos de dados internos (do próprio programa) e externos (DBase, Access e Oracle são os mais comuns). Quando se marca um ponto na tela pode-se consultar a tabela completa dos dados sobre aquele ponto. Por exemplo, a consulta a um ponto de coleta geoquímica permite obter suas coordenadas, teores dos elementos químicos, parâmetros de coleta, etc.

A consulta espacial é parte integrante dos trabalhos mais avançados, relatados nos tópicos seguintes.

7.4 Combinação

Os dados espaciais e de atributos são combinados, permitindo juntar, por exemplo, anomalias de vários elementos numa imagem de fator geoquímico, facilitando a interpretação e os processamentos mais complexos. Este trabalho, às vezes denominado **modelamento cartográfico,** representa

uma fase preliminar da análise dos dados. Exemplos e comentários serão apresentados no tópico seguinte.

7.5 Análise

Segundo Bonham-Carter (1994) a análise é o **processo de descobrir o significado dos dados**. As análises podem ser espaciais e não espaciais. Estas últimas usam técnicas de estatística e de modelagem, por exemplo para determinar limiares de anomalias geoquímicas e ajuste de distribuições a curvas teóricas, como a lognormal, e foram comentadas nos capítulos precedentes. Com a introdução da localização dos dados como propriedade a ser analisada, passamos à análise espacial, que pode ser visual, pode ser feita sobre um ou vários mapas, pode se limitar à combinação de mapas ou incorporar técnicas estatísticas, como a análise de componentes principais com saída gráfica e a análise de superfícies de tendência.

Bonham-Carter (1994) apresenta como fase preliminar à análise a de transformação de dados, podendo pontos ser transformados em outros pontos, por interpolação, em linhas, por curvas de contorno de um MNT interpolado a partir dos pontos originais, ou em áreas, por subdivisão de áreas de influência em polígonos de Thiessen; linhas e áreas podem ser transformadas em pontos, linhas e áreas. Destas transformações, as duas primeiras, que envolvem interpolação, são as mais usadas, sendo também muito importante a geração de corredores, que são áreas situadas até uma dada distância de uma linha. Exemplos de interpolação já foram apresentados juntamente com a visualização de MNT; exemplos de corredores serão apresentados quando comentada a análise dos dados da folha Itararé.

7.5.1 Fundamentos teóricos

A este ponto é necessária uma reflexão sobre os problemas tratados pelas técnicas de análise, predição e apoio à decisão em geoprocessamento. Estes três passos são divididos por conveniência didática. O geoprocessamento é sempre utilizado para uma decisão, explícita ou implícita. Para isto deve-se prever quais partes da superfície da Terra, coberta pelos mapas analisados, atendem ao enunciado do problema estudado. As técnicas de análise permitem chegar à predição e à decisão.

O geoprocessamento tem sido utilizado em áreas de conhecimento extremamente variadas. Um congresso como o "GIS in the Rockies", ou o "Congresso Brasileiro de Geoprocessamento", cobre temas de ciências exatas, biológicas e humanas, com aplicações para áreas variadas da atividade humana, como saúde, administração municipal, localização de áreas para instalação de McDonald's, atendimento de chamadas de bombeiros e policiais, estabelecimento de fronteiras de distritos eleitorais ("electronic gerrymandering"), etc, havendo menor variedade de trabalhos no congresso nacional, ainda muito preso à Cartografia e a aplicações na administração pública.

O grupo de trabalho de Geoprocessamento do DGE-IG-USP dedica-se preferencialmente a aplicações para Recursos Minerais e Meio Ambiente. Dois exemplos são apresentados neste texto, um deles já bastante comentado, a interpretação de um levantamento geoquímico de solos na área do Paqueiro; outro será a escolha de áreas para prospecção de recursos minerais nas folhas Cerro Azul e Apiaí a 1:100.000, conforme dois modelos de mineralização.

O primeiro problema é direto e relativamente simples. Numa área pequena (2.300 x 5.100 m), parte da qual coberta por um levantamento de solo com 1.009 amostras, determina-se a localização geográfica de anomalias geoquímicas, e combinam-se essas anomalias com feições estruturais previamente julgadas como controladoras das mineralizações conhecidas (lineamentos Paqueiro e Bueno). As áreas cobertas por combinações de anomalias são consideradas aptas para detalhamento e possível descoberta

de novas mineralizações. As litologias que sustentam as áreas amostradas, com exceção de uma pequena área de dique de diabásio, são julgadas igualmente promissoras. Este problema é aqui tratado principalmente para demonstrar técnicas de tratamento espacial combinadas a técnicas não-espaciais, empregadas em trabalhos antigos e já comentadas nos capítulos anteriores. Os resultados da aplicação dos métodos de análise aos dados da área do Paqueiro serão apresentados juntamente com a exposição dos fundamentos de cada técnica.

O segundo problema é mais complexo. A área envolvida é muito maior (30 minutos por um grau geográficos), com grande variação litológica e estratigráfica, centenas de ocorrências minerais, interpretáveis por muitos modelos de mineralização. Devido a esta complexidade, os resultados da aplicação de geoprocessamento à análise metalogenética regional das folhas Apiaí e Cerro Azul serão apresentados em conjunto, no item 7.9.

A abordagem teórica da análise das possibilidades metalogenéticas de uma área deste porte já constitui por si só um problema considerável.

A análise metalogenética regional desenvolveu-se principalmente após a Segunda Guerra Mundial, contribuindo para o "boom" da descoberta de jazidas das décadas de 50 a 70. Três escolas metodológicas destacam-se para esta análise:

- a escola francesa, melhor expressa em Routhier (1980), que combina feições geológicas regionais e ocorrências minerais em "Leis da Metalogenia";
- a escola soviética, expressa em Shcheglov (1979), que combina traços geológicos, mineralizações e dados de levantamentos geoquímicos e geofísicos, para determinação de áreas favoráveis à mineralização. Os primeiros trabalhos do Projeto Mapas Metalogenéticos e Previsionais do DNPM/CPRM (tais como Hama & Algarte, 1986) seguiram uma metodologia calcada na soviética.
- a escola norte-americana, expressa em Harris (1984), e pelos trabalhos do USGS, cuja metodologia, resumida em Singer e Mosier (1981), é ainda hoje empregada, com atualizações relacionadas ao uso de métodos quantitativos, que incluem inteligência artificial (McCammon, 1987) e geoprocessamento (Elliot e Stoeser, 1994). O grupo do Geological Survey of Canada foi o que

melhor desenvolveu a síntese entre essas abordagens, o que se nota, por exemplo, em Reedy et alii (1992). Uma abordagem análoga é seguida pela escola finlandesa, sendo relatados exemplos em Gaál e Merrian (1990).

A escola norte-americana junta o componente geológico (determinação de áreas favoráveis a um determinado modelo de mineralização), com o econômico (quantificação do bem mineral localizável em uma determinada área), sempre com objetivos práticos claros e freqüentemente imediatos. Um problema a que se dedicou o USGS na última década foi a determinação do potencial mineral das áreas públicas, de forma a poder racionalmente destiná-las à produção ou preservação.

Alguns dos que desenvolvem esses estudos não utilizam geoprocessamento. Harris (1984), após analisar todos os métodos disponíveis até a época, preferia que a avaliação geológica fosse feita de forma subjetiva, em alguns casos usando técnicas estatísticas para combinar opiniões de diversos profissionais. Em Pan, Harris e Heiner (1992) já é aceita a integração estatística de informações geológicas.

A análise subjetiva na primeira fase dos estudos ainda é preferida por parte dos profissionais que fazem avaliação regional de recursos minerais no USGS, como o que acontecia no grupo do Center for Interamerican Mineral Resource Investigations, onde o autor desenvolveu parte de seu trabalho de pós-doutorado.

Outra abordagem é a divisão da área estudada em células, a determinação da presença-ausência ou de um índice numérico associado a um fator geológico ou geoquímico de mineralização e a combinação desses índices por métodos estatísticos. Este foi o processo seguido por Suslick (1986), e continua, em linhas gerais, a ser seguido em McCammon (1994).

O método original de inteligência artificial (tipo PROSPECTOR, como em McCammon, 1987), aplicado à análise metalogenética utilizava a combinação de informações da área com um banco de informações relacionadas aos modelos metalogenéticos, sendo o resultado expresso em probabilidade de ocorrência de mineralizações na área analisada. Uma atualização desta abordagem é o PROSPECTOR II, de McCammon (1993),

que incorpora mapas digitalizados para consulta ao banco de dados e exibição de resultados

Katz (1991) mostrou que é possível reproduzir o trabalho executado pelo método clássico da inteligência artificial (PROSPECTOR), com vantagens, utilizando geoprocessamento, num sistema de informações georreferenciadas. Esta abordagem é seguida por Bonham-Carter et alii (1994), que distingue dois grupos de métodos de análise metalogenética em geoprocessamento:

a - os métodos empíricos, que partem das ocorrências, ligando-as às informações geológicas e montando um modelo a partir dos dados (*data-driven*). Eles podem empregar regressão logística, pesos de evidência e redes neurais.

b - os métodos baseados em modelos de mineralização (*knowledge-driven*), que partem de regras, determinadas por peritos, que associam as informações geológicas à mineralização. Os mapas podem ser combinados de forma determinística (por mapas booleanos ou com índices somativos) ou probabilística (probabilidades simples ou condicionais, possibilidades (*fuzzy sets*), ou pelo método de Dempster-Shafer (*belief functions*).

O geoprocessamento difere do método clássico da divisão da área estudada em células porque os elementos da imagem são em número muito maior do que aquelas (cem vezes maior, nos trabalhos observados), e porque a análise é feita por mapas que representam inicialmente fatores isolados. A combinação é feita por mapa e não por célula. O trabalho clássico descrevendo este tipo de abordagem é o de Bonham-Carter et alii (1988), e uma exposição muito completa e didática é apresentada em Bonham-Carter (1994). Métodos semelhantes são usados por Eastman et alii (1993, 1995).

Um exemplo da aplicação de geoprocessamento à análise metalogenética será apresentado no item 7.9.

7.5.2 Análise com um mapa

A primeira e mais simples atividade de análise é a classificação dos dados para visualização e processamentos subseqüentes. HOWARTH (1983)

descreve os métodos de confecção de mapas geoquímicos por pontos ou suas transformações para isolinhas e mapas coloridos, para representação da distribuição de teores e anomalias.

7.5.2.1 Representação de teores e anomalias

A figura 24 apresenta os teores de Cu, Pb e Zn na área Mina do Paqueiro, divididos simplesmente em 14 intervalos logarítmicos, divididos a partir dos pontos extremos. A área amostrada é sobreposta a uma figura representando os contatos e lineamentos estruturais, considerados condicionantes de mineralização.

Uma representação como a da figura 24 tem pouca utilidade para análise. A distribuição geoquímica se torna mais compreensível quando se dividem os teores em faixas relacionadas às anomalias de distribuição. A figura 25 apresenta a área amostrada dividida em faixas correspondentes a níveis de anomalias, tomando-se como os parâmetros pelo método simplificado, apresentado em Ferreira et alii, 1981 e comentado no capítulo 5.

Os valores utilizados para traçar as figuras foram, em ppm:

Elemento	**MG**	**DG**	**Alto background**	**Anomalia possível**	**Anomalia provável**
Cu	36	1,6	58 - 91	92 - 146	> 147
Pb	45	1,9	86 - 161	162 - 308	> 309
Zn	45	1,5	68 - 100	101 - 150	> 151

Nota-se o claro condicionamento estrutural das anomalias de chumbo, menos evidente nas de cobre e zinco, tendo o cobre, ainda, suas maiores anomalias associadas ao dique de diabásio.

Outra forma de análise em um mapa é a de proximidade, que se faz medindo a distância dos pontos do mapa a feições de interesse. Esta análise é um trabalho preliminar ao traçado dos corredores (*buffers*), como os que foram traçados para os lineamentos nas folhas de Apiaí e Cerro Azul, que são apresentados e comentados à frente.

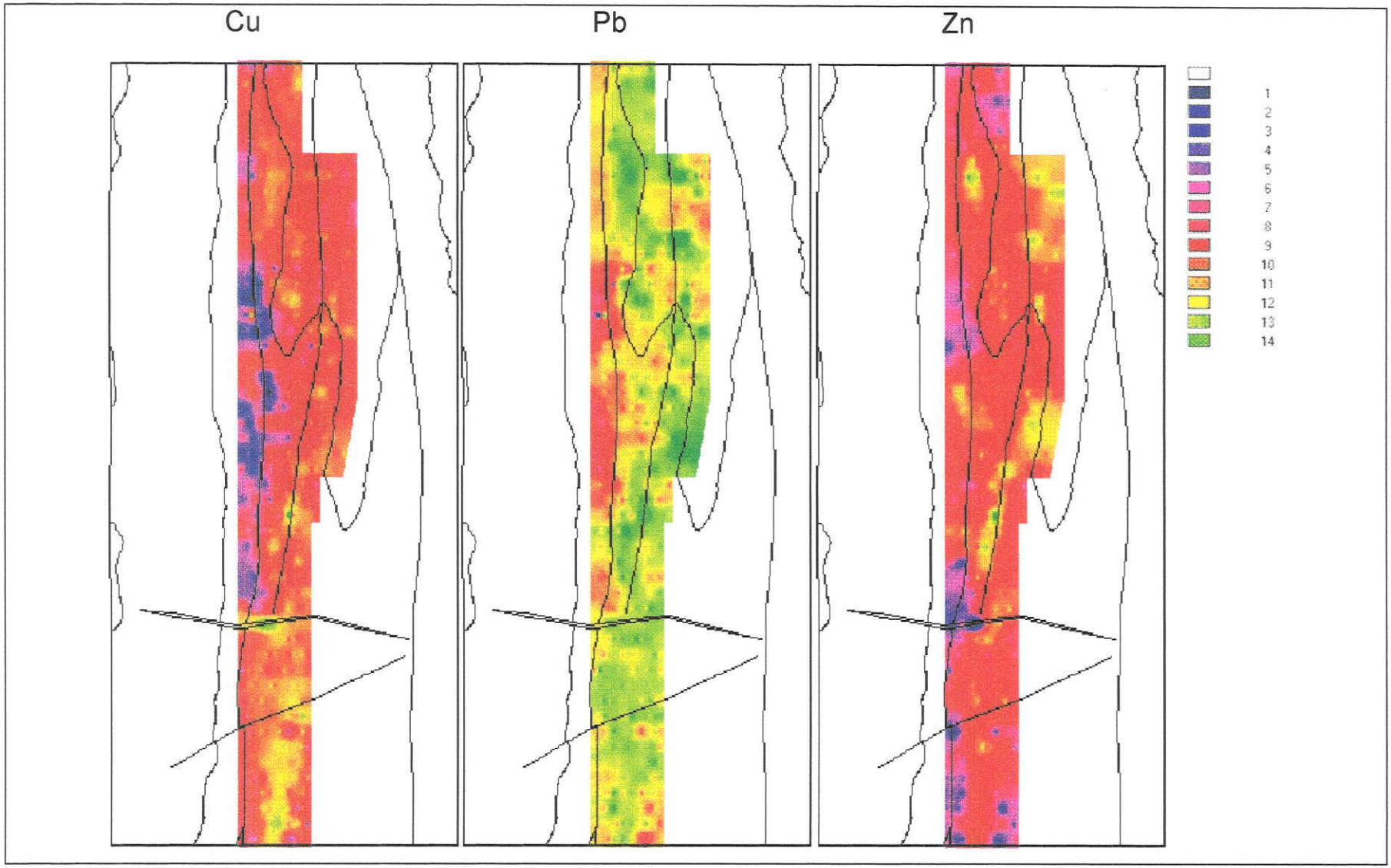


Figura 24 - Distribuição dos teores de Cobre, Chumbo e Zinco na área Mina do Paqueiro.

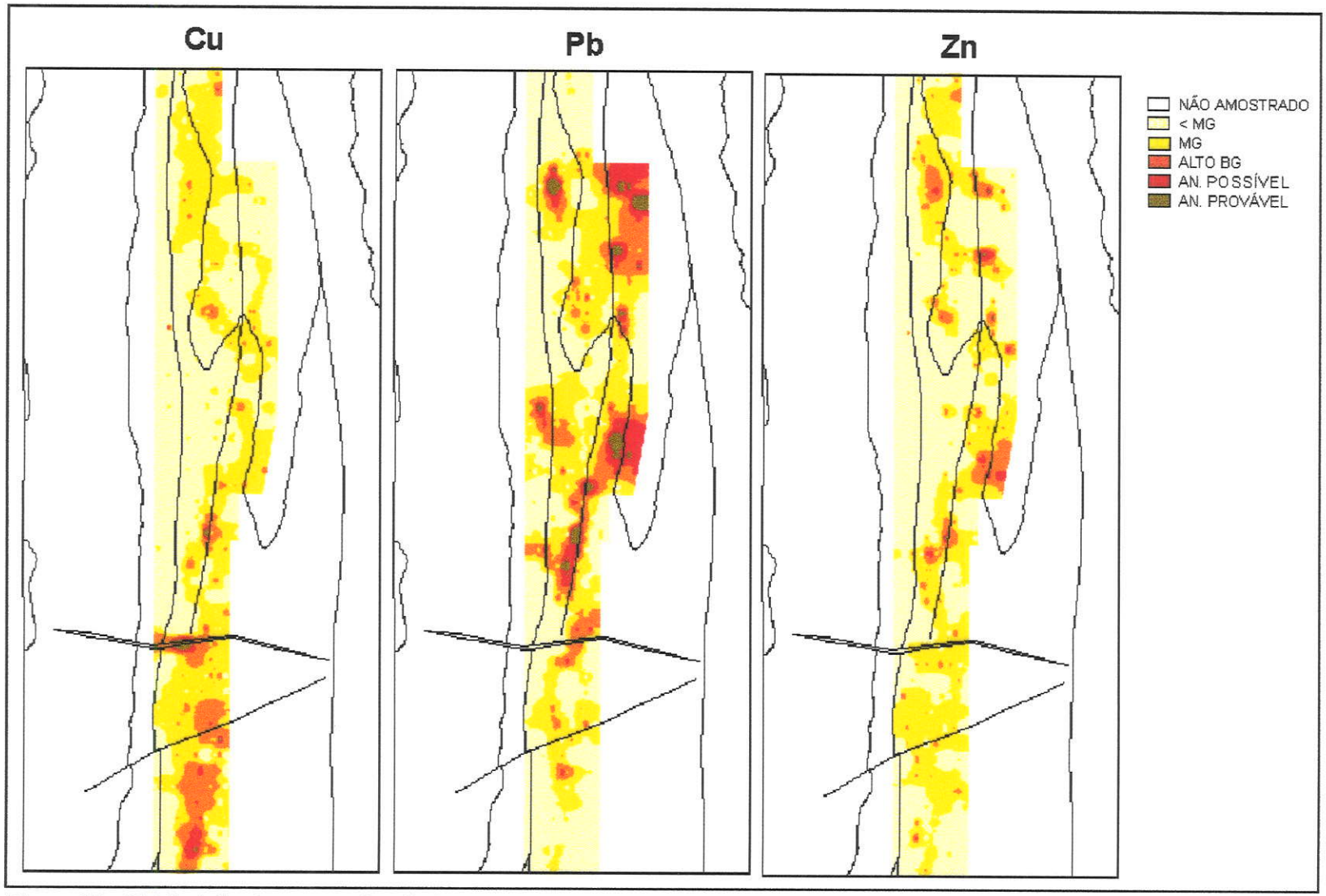


Figura 25 - Distribuição das anomalias de Cobre, Chumbo e Zinco, em amostras de solo, na área Mina do Paqueiro.

7.5.2.2 Análise de superfícies de tendência

7.5.2.2.1 Generalidades

Segundo Landim (1993), com esta técnica consegue-se separar dados mapeáveis em duas componentes, uma de natureza regional, representada pela própria superfície e outra que revela as flutuações locais, representada pelos valores residuais. Para o tratamento de dados geoquímicos, espera-se que a componente regional revele a variação geológica e os resíduos as anomalias geoquímicas.

Adotando-se a simbologia de Davis (1986), tomando-se os dados representados por suas coordenadas X_1 e X_2, representando os valores da terceira dimensão por Y, calcula-se uma superfície, expressa, quando do primeiro grau, da forma:

$$Y = b_0 + b_1 X_1 + b_2 X_2$$

Procura-se minimizar, por um método de mínimos quadrados, a diferença entre os valores de Y na superfície calculada e os valores experimentais.

Na prática é em geral impossível haver um ajuste perfeito, e a equação fica, em cada ponto, com um valor

$$Y = b_0 + b_1 X_1 + b_2 X_2 + e$$

sendo e o resíduo, ou seja a diferença entre o valor experimental de Y e o valor dessa variável na superfície calculada.

A eficiência da regressão é testada por dois parâmetros, relacionados entre si: a correlação entre os valores calculados e experimentais (sendo seu quadrado considerado o índice de ajuste - *goodness of fit* - e a razão entre a variação devida à regressão e aquela devida aos resíduos, testada pela estatística F de Fisher-Snedecor.

Muitas vezes a análise de uma superfície plana (de primeiro grau) não é suficiente para a compreensão dos dados; então usam-se equações mais complexas, da forma:

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_1^2 + b_4 X_2^2 + b_5 X_1 X_2$$

quando do segundo grau e expansões análogas para graus superiores.

7.5.2.2.2. Aplicação ao estudo da variação dos teores do minério da mina do Perau

A variação espacial dos teores dos elementos constituintes do minério da mina do Perau foi estudada em Macedo (1986) e Barbour e Macedo (1991). No primeiro trabalho foram calculadas superfícies de tendência até o sexto grau; no segundo trabalho foram desenhados em computador diagramas tridimensionais. Em ambos os casos as limitações dos dispositivos de saída fizeram com que os resultados publicados ficassem pouco visíveis.

7.5.2.2.3 Aplicação para os teores de chumbo na área Mina do Paqueiro

Foram calculadas superfícies de tendência de primeiro a terceiro graus e seus resíduos para teores logaritmizados de chumbo em amostras de solo da área Mina do Paqueiro. Foi utilizado o programa IDRISI, sendo os detalhes de cálculo apresentados em Eastman (1992*a*).

As estatísticas F para teste da significância do aumento de grau na equação foram calculadas pela fórmula:

$$F = (\% I / gl1)/((100 - \% R^2 - \% I) / gl2)$$

onde:

F = estatística F, testada pela estatística F de Fisher-Snedecor;

% I = aumento percentual do ajuste (quadrado do coeficiente de correlacão entre os dados e seus correspondentes na superfície calculada);

% R^2 = ajuste original, para a equação de grau mais baixo;

gl1 = número de graus de liberdade adicionais associados com a superfície de tendência de grau mais alto (3 para superfície quadrática sobre a linear; 4 para cúbica sobre a quadrática);

gl2 = número de graus de liberdade associados com os resíduos da superfície de tendência de grau superior igual a N menos o número de termos na equação (6 para quadrática, 10 para cúbica).

Os parâmetros de ajuste foram os seguintes:

Tabela 6: *Análise de variância para a significância das equações de regressão das superfícies de tendência de graus primeiro a terceiro, para teores logaritmizados de chumbo em amostras de solo na área Mina do Paqueiro. Nível de significância 5%.*

Grau da equação	R^2	F	F crítico	F da mudança	F crítico
Primeiro	0,1204	10899	(2,∞) 3,00		
Segundo	0,2908	13050	(5,∞) 2,21	12743	(3,∞) 2,60
Terceiro	0,3384	9042	(9,∞) 1,88	2861	(4,∞) 2,37

T2302

A alta significância do ajuste das superfícies e do aumento do grau das equações é enganosa. Eastman (1992*a*) adverte que, para a aplicação rigorosa do teste, as células usadas para o cálculo do ajuste deveriam ser amostras independentes da população. No entanto o programa utiliza o total das células, as quais têm valores com altíssima autocorrelação espacial, conforme é mostrado no item seguinte.

Uma maneira prática de usar a análise de tendência para interpretação geoquímica é representar a superfície num bloco-diagrama e sobrepor a ela um mapa cujas cores representam os valores dos resíduos, como na figura 26. Ali os resíduos foram classificados em 15 intervalos, representados da mesma forma que na figura 24. As anomalias ocupam a mesma localização daquelas determinadas pelo cálculo estatístico não-espacial e plotadas em mapa nas figuras 29 a 31.

INSTITUTO DE GEOCIÊNCIAS - USP
— BIBLIOTECA —

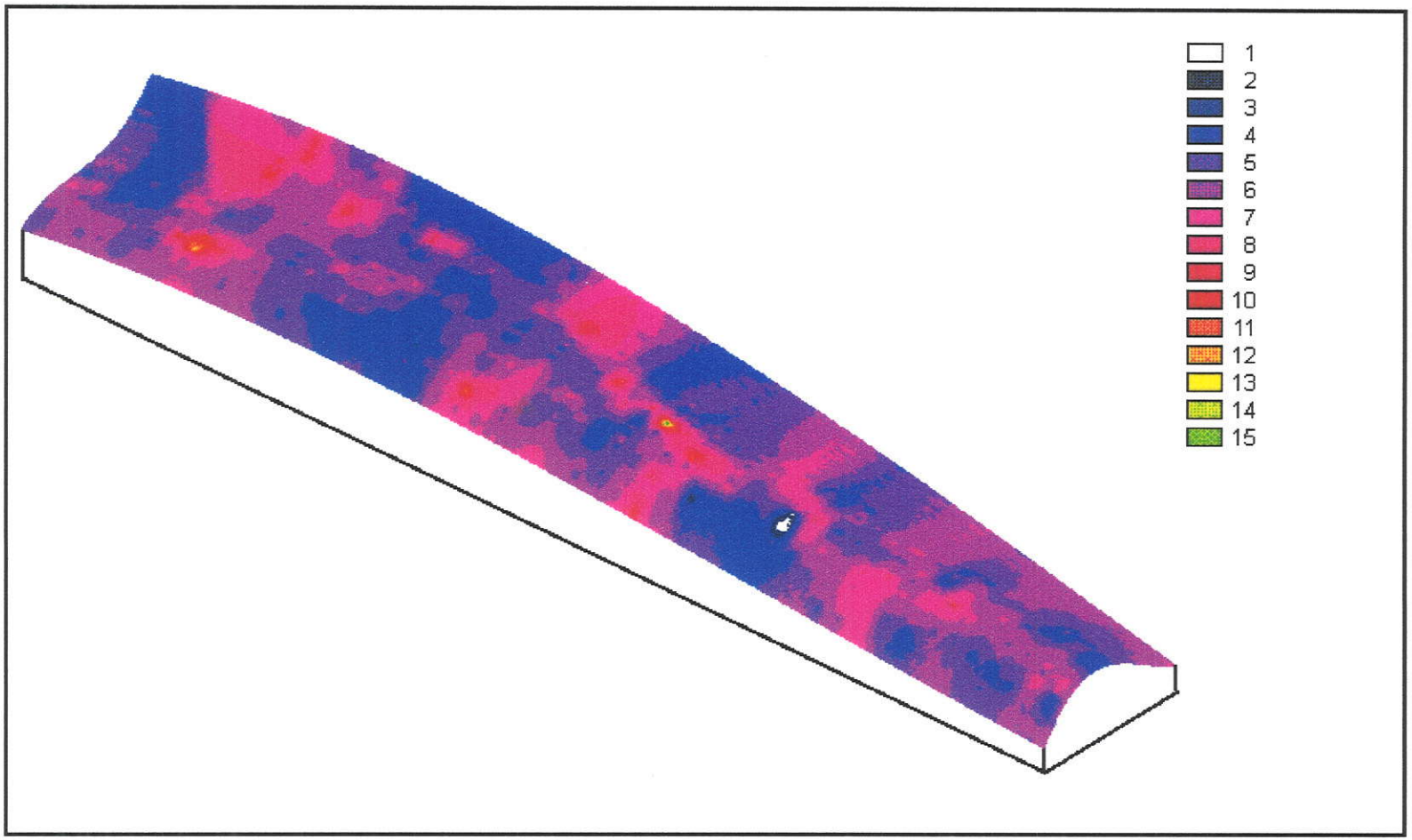


Figura 26 - Resíduos de superfície de tendência de terceiro grau, sobre a mesma superfície, para logaritmos de teores de Pb em solo na área Mina do Paqueiro, transformados em 15 intervalos para visualização.

7.5.2.3 Autocorrelação

A autocorrelação entre os teores de Pb em amostras de solo da área Mina do Paqueiro foi testada calculando-se o índice I de Moran, usando o procedimento Autocorr do IDRISI. Pelo método da torre de xadrez (testando nas direções E-W e N-S), o coeficiente foi 0,9957; pelo método do rei (testando E-W, N-S, NW-SE e NE-SW), o coeficiente foi 0,9932, ambos significantes a nível inferior a 0,01 %, conforme Eastman (1992*a*).

Com dados deste tipo, os resultados das análises de variância para determinação da significância das superfícies de tendência têm valor apenas indicativo, não servindo como teste estatístico.

7.5.2.4. Variogramas

Para determinar se o espaçamento utilizado para a amostragem na área Mina do Paqueiro foi suficiente para expressar a variabilidade dos teores estudados, foram calculados, para os teores logaritmizados de Cu, Pb e Zn, analisados em amostras de solo na área Mina do Paqueiro, variogramas nas direções paralela (direção 0 nos gráficos) e perpendicular (direção 90) às linhas de amostragem, o que corresponde aproximadamente a direções perpendicular e paralela aos lineamentos, que condicionam as mineralizações no local. O espaçamento de cálculo (*lag*) foi o mesmo de amostragem, sendo de 30 m nas linhas (direção 0) e 100 m entre elas (direção 90).

Ajustando-se os variogramas experimentais ao modelo esférico, observou-se que os variogramas traçados na direção paralela às linhas de amostragem apresentam uma amplitude de 172,5 a 284 m, mostrando ser perfeitamente adequado o espaçamento de amostragem. A amplitude na direção perpendicular às linhas apresentou valores de 299 a 920 m, mostrando que o espaçamento entre linhas foi também adequado.

Os variogramas experimentais e as curvas e eles ajustadas segundo o modelo esférico são apresentados na figura 27.

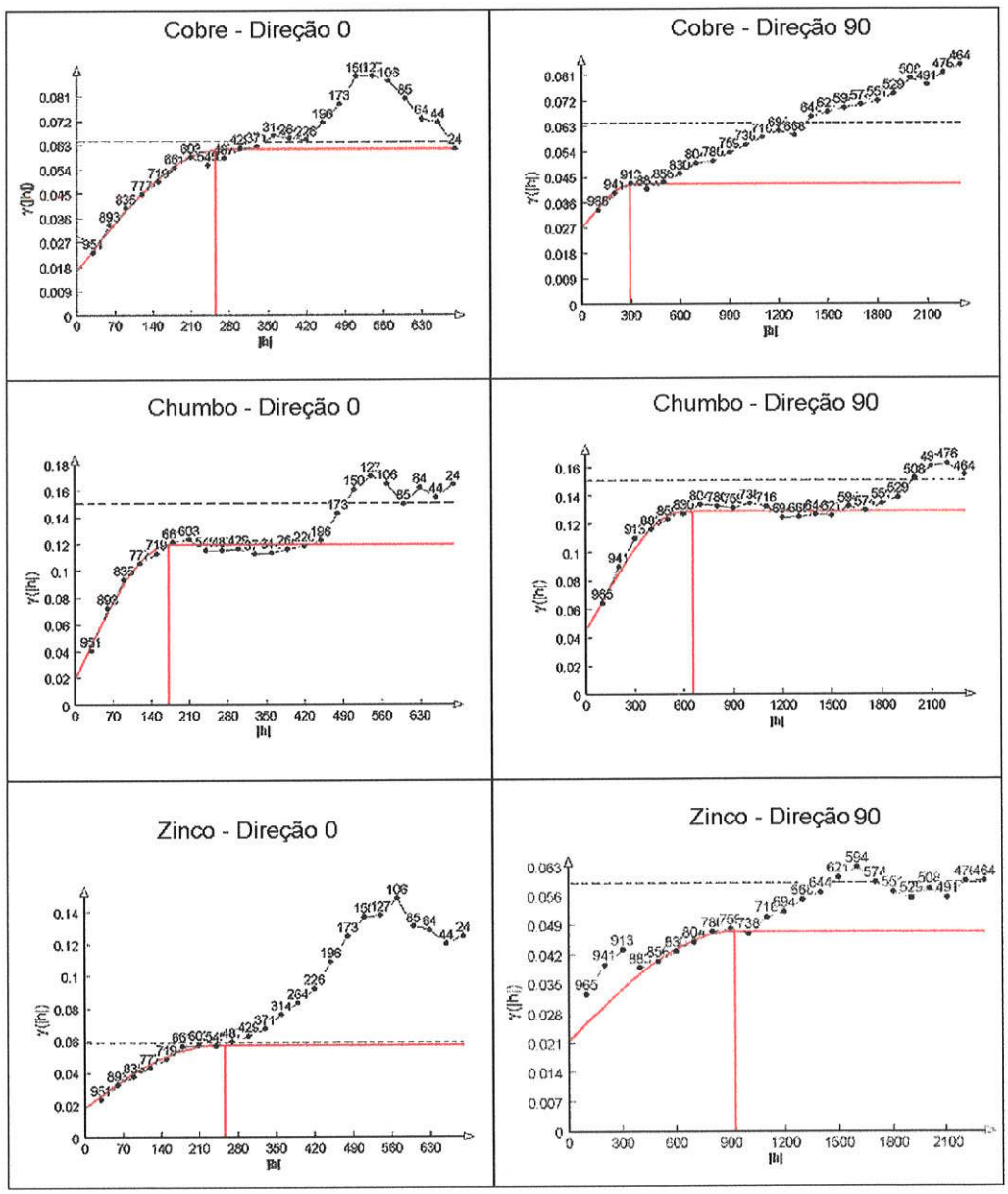


Figura 27 - Variogramas experimentais e ajustados ao modelo esférico, para teores logaritmizados de Cu, Pb e Zn, em amostras de solo da área Mina do Paqueiro, nas direções 0 (paralela às linhas de amostragem, perpendicular aos lineamentos) e 90 (perpendicular às linhas de amostragem, paralela aos lineamentos).

7.5.3 Análise com múltiplos mapas

7.5.3.1 Combinação de anomalias

Os mapas podem ser combinados, tanto dois a dois quanto em maior número. A sobreposição (*overlay*) é uma das operações de análise mais utilizadas. Nesta, os valores dos pixels correspondentes (com idênticas localizações) dos diversos mapas são combinados em operações matemáticas, tais como soma, subtração, divisão, ou são substituídos conforme regras como, por exemplo: "substituir o valor do pixel do mapa 1 pelo valor do pixel do mapa 2, se este não for zero".

Um exemplo de sobreposição é a combinação de mapas de anomalias de Cu, Pb e Zn da área Mina do Paqueiro. Ela pode ser feita por simples soma de pontos (considerando-se um ponto para área com valores inferiores à média, dois para a faixa da média, três para alto background, quatro para anomalia possível e cinco para anomalia provável). O resultado é apresentado na figura 28. As anomalias associadas ao lineamento Paqueiro aparecem claramente, bem como a associada ao dique de diabásio. Este tipo de combinação tem um problema: não é possível saber a composição das anomalias combinadas em relação aos elementos constituintes.

Outra maneira, melhor, é reclassificar os mapas para duas faixas: 0 para área sem anomalia, 1 para área com valores de anomalia possível e maiores, e depois atribuir pesos variando geometricamente aos diversos elementos, conforme recomendado por Elliot e Stoeser (1994). Neste caso foram atribuídos os pesos: 1 a Cu, 2 a Pb e 4 a Zn. Quando se somam os mapas, os valores são únicos, de modo, que, por exemplo, um valor 3 indica sempre a soma de Cu e Pb, 4 a presença de Zn isolado, etc. Ajustando-se as cores a esses valores é possível um mapa muito ilustrativo, que indica a localização, ao mesmo tempo, das anomalias por elementos isolados e as combinações. Um exemplo, com os mesmos dados da área Mina do Paqueiro, é apresentado na figura 29. Nota-se então a predominância de Pb nas anomalias associadas ao lineamento Paqueiro e ao contato entre os tipos de calcário; o Cu predomina nas anomalias a Sul da área amostrada.

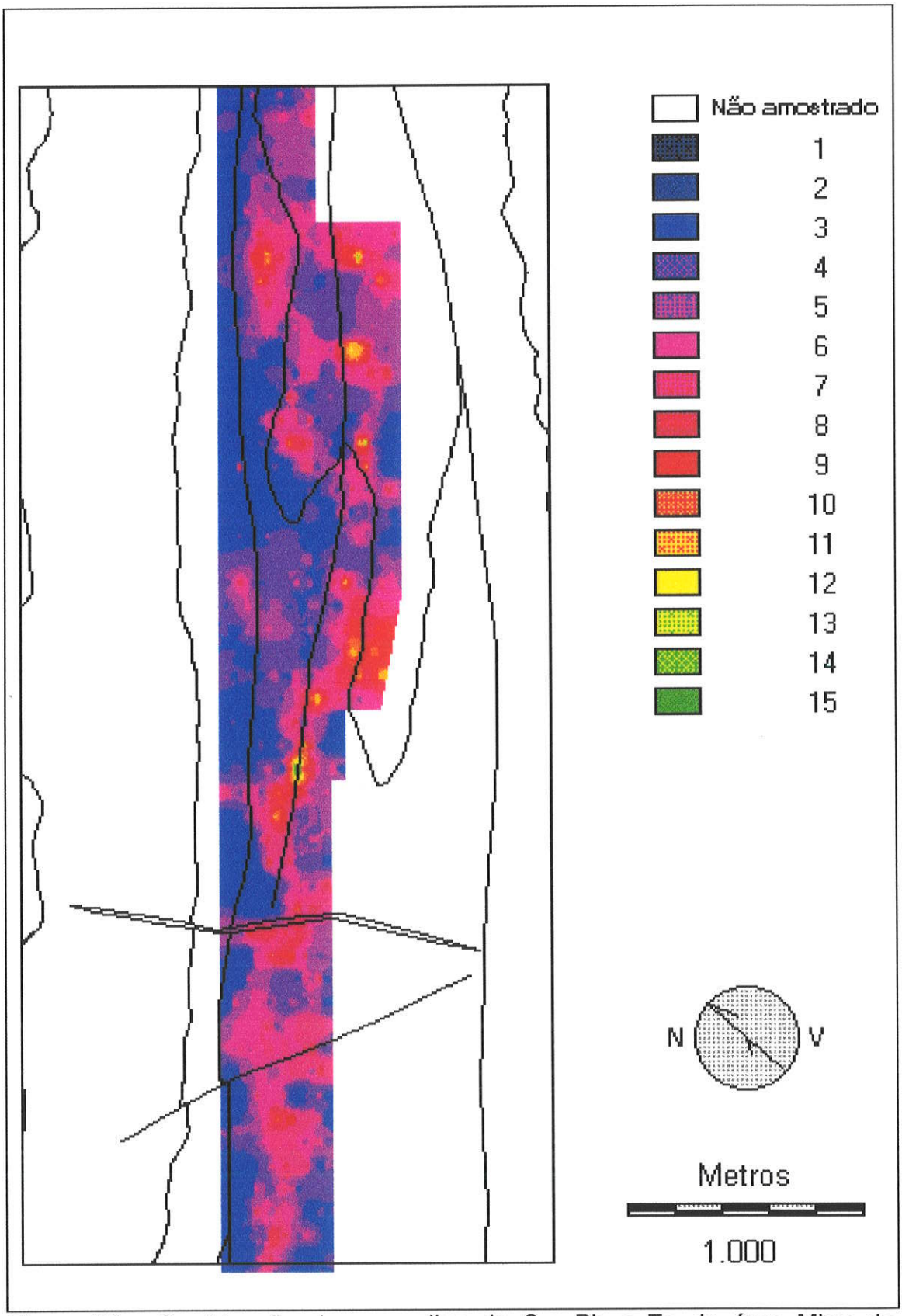


Figura 28 - Combinação de anomalias de Cu, Pb e Zn da área Mina do Paqueiro por soma simples de pontos.

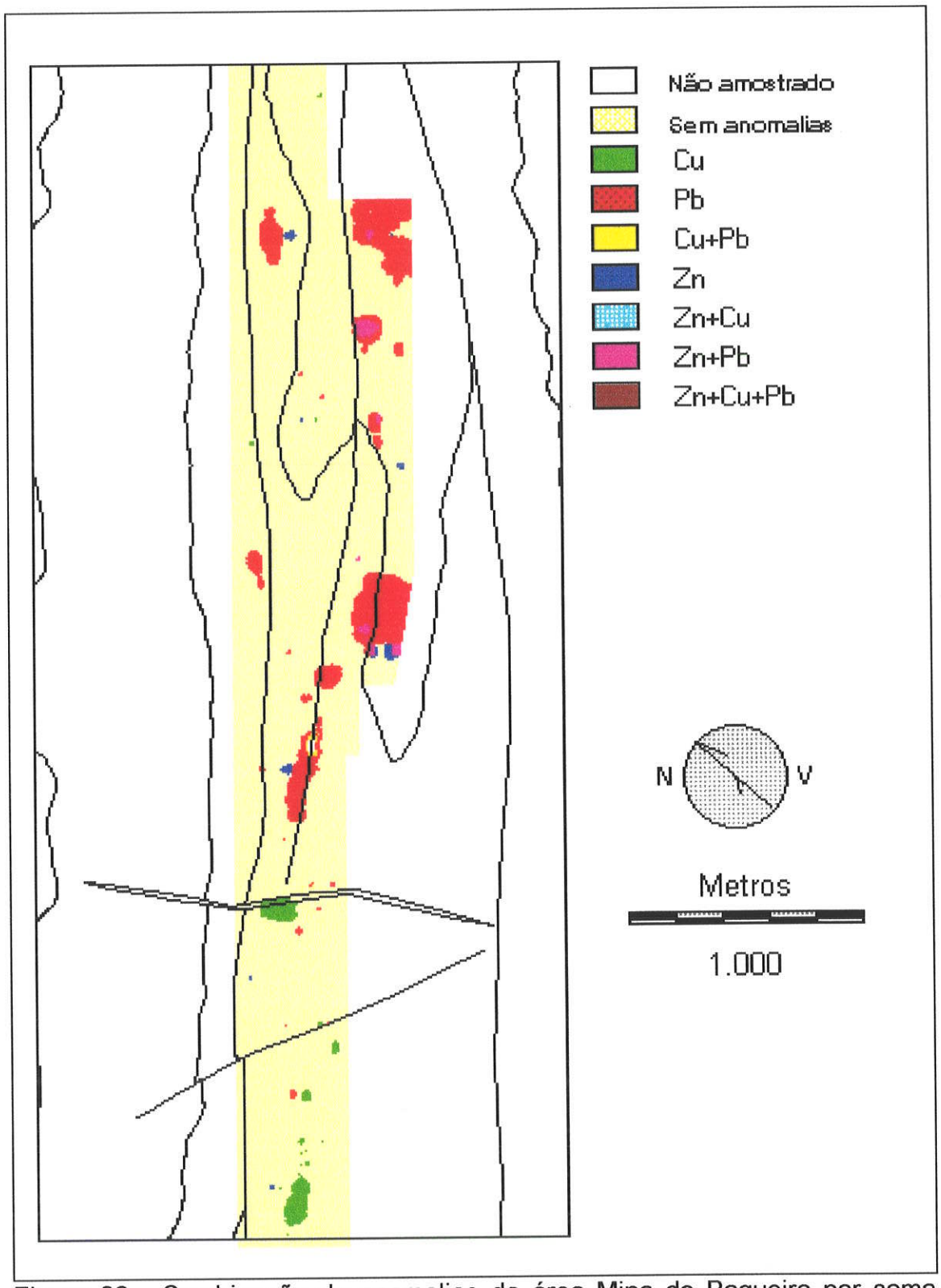


Figura 29 - Combinação de anomalias da área Mina do Paqueiro por soma com pontos variando geometricamente.

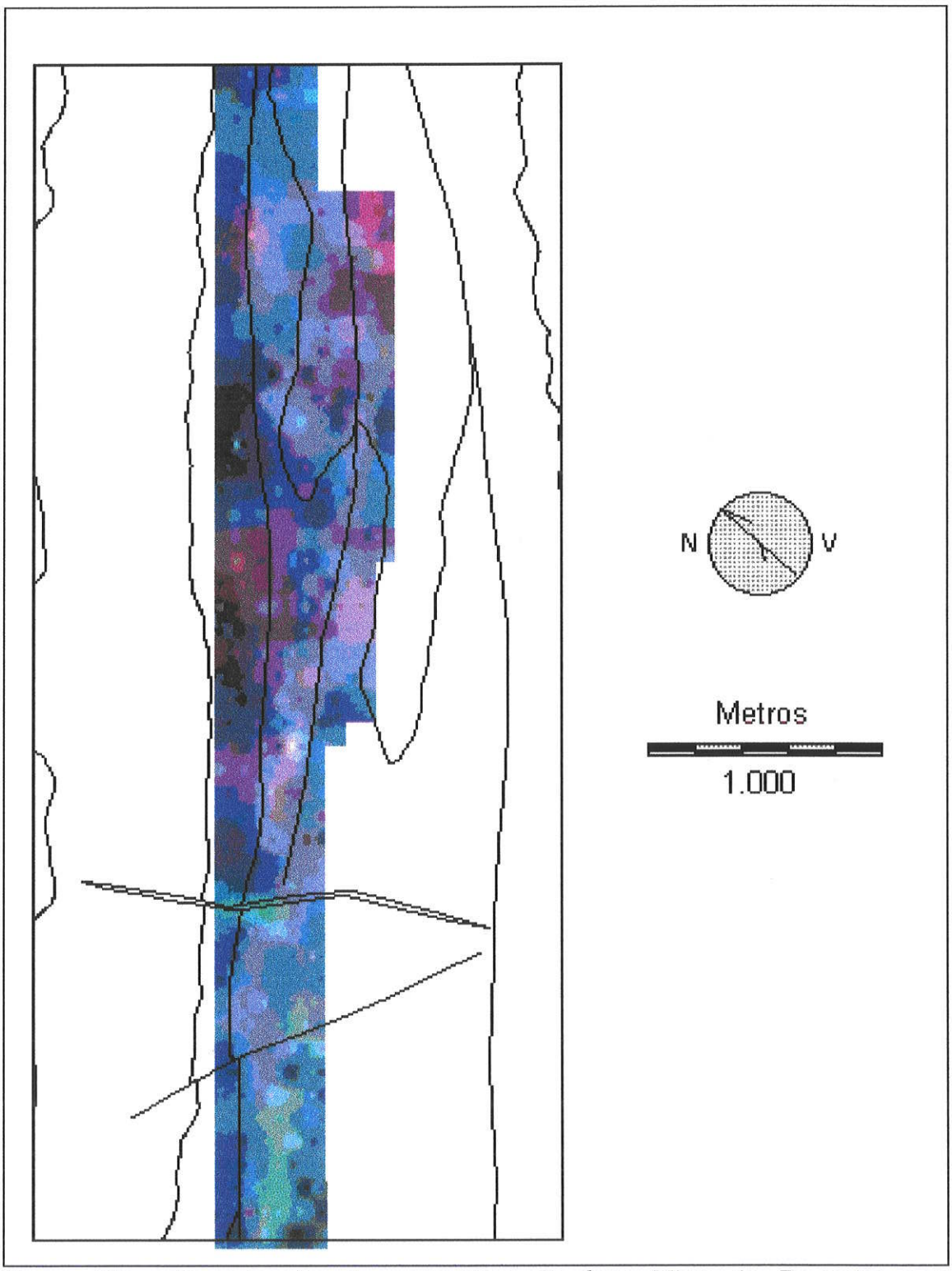


Figura 30 - Combinação de anomalias da área Mina do Paqueiro por composição colorida: Pb:vermelho; Cu: verde; Zn: azul.

Ainda outra maneira é utilizar o módulo de composição colorida dos programas de tratamento de imagens, como o IDRISI, e atribuir cores aos elementos. No caso, atribuímos vermelho ao Pb, verde ao cobre e azul ao zinco, resultando no mapa apresentado na figura 30, bem menos útil para interpretação que o da figura 29, embora também neste caso seja possível determinar o elemento predominante nas anomalias.

7.5.3.2 Análise de componentes principais

Um processo mais elaborado de combinação de informações geoquímicas e outras informações geológicas é a transformação dos dados em componentes principais, calculando conforme as fórmulas do item 6.2.2.

Para esta combinação foram utilizados os teores de Cu, Pb e Zn em amostras de solo da área Mina do Paqueiro e também as anomalias magnéticas, determinadas na área a partir de medidas terrestres da intensidade do campo magnético total, com magnetrômetro Geometrics G-826, efetuadas nos mesmos pontos de coleta de amostras de solo e tendo sido os dados processados pelo geofísico Francisco J.F. Ferreira.

O cálculo, usando o módulo PCA do IDRISI, primeiramente obriga à reclassificação dos valores em cada mapa em 255 intervalos (expressos em números de um byte), que depois são combinados. O mapa da figura 31 ilustra as anomalias magnéticas, após essa reclassificação, e sofrendo ainda uma condensação de intervalos para ser impresso em 14 cores.

Os resultados da análise de componentes principais para os logaritmos de teores de Cu, Pb e Zn, combinados com as anomalias magnéticas na área Mina do Paqueiro (todos eles inicialmente reclassificados para 255 intervalos) foram os seguintes:

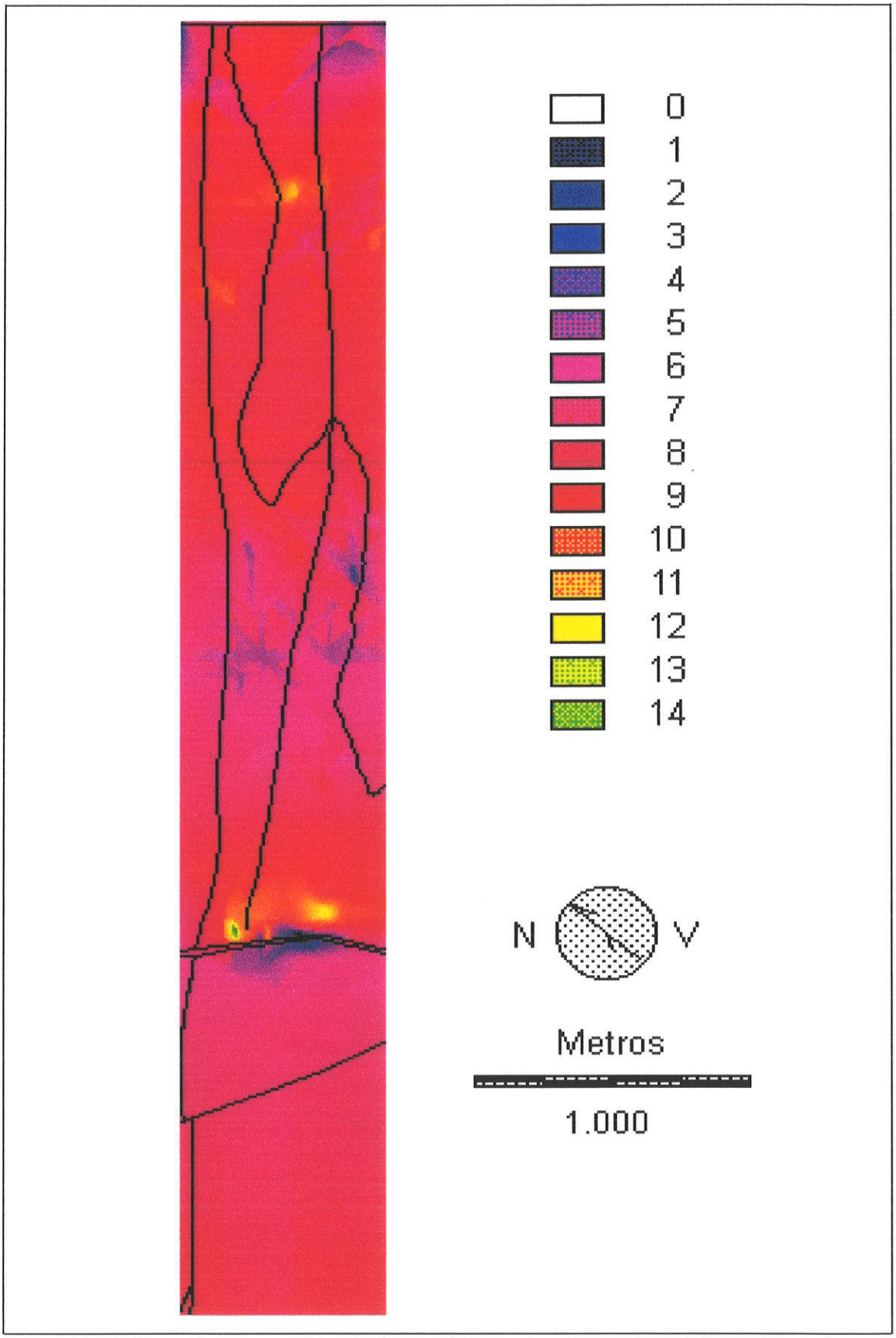


Figura 31 - Mapa de anomalias de intensidade do campo magnético total, área Mina do Paqueiro, reclassificado.

Tabela 7: *Resultados da análise de componentes principais de dados geoquímicos e geofísicos da área Mina do Paqueiro:*

Correlação	Cu	Pb	Zn	Anom. Mag.
Cu	1,000	0,156	0,773	- 0,118
Pb	0,156	1,000	0,369	0,001
Zn	0,773	0,369	1.000	- 0,070
Anom. Mag	- 0,118	0,001	-0,070	1,000

Componente	Comp. 1	Comp. 2	Comp. 3	Comp. 4
% variância	62,17	25,99	6,47	5,37
autovalor	1158,39	484,31	120,58	100,11
autovetor 1	0,605	-0,393	0,629	0,288
autovetor 2	0,344	0,913	0,211	0,061
autovetor 3	0,717	-0,105	-0,648	- 0235
autovetor 4	- 0,029	0,035	- 0,373	0,926

Cargas	Comp 1	Comp. 2	Comp. 3	Comp. 4
Cu	0,875	-0,367	0,293	0,122
Pb	0,501	0,859	0,099	0,026
Zn	0,952	-0,091	-0,276	- 0,092
Anom. mag.	- 0,096	0,077	- 0,402	0,908

Estes componentes principais estão expressos na figura 31. Os componentes podem, em primeira aproximação, ser interpretados como representando a mineralização (componentes 1 e 2), e o fator litológico, com o dique bem assinalado (componente 3), e fator geofísico, sem associação com mineralização e apenas indicando o grande dique de diabásio do Ribeirão das Onças. O componente 1 indica muito bem as anomalias associadas aos dois lineamentos.

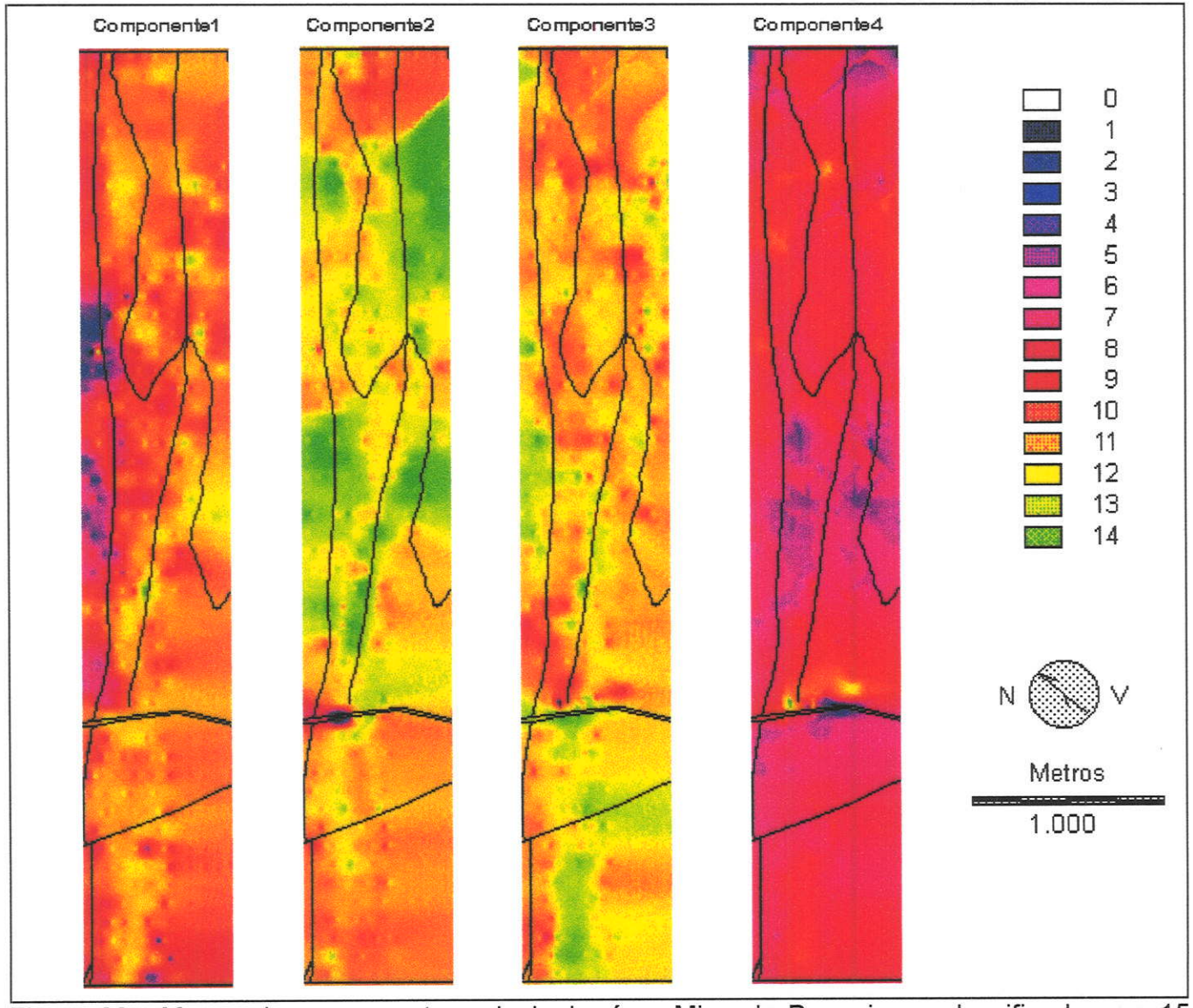


Figura 32 - Mapas de componentes principais, área Mina do Paqueiro, reclassificados em 15 intervalos para visualização.

7.6 Predição

O passo seguinte é a aplicação da análise espacial para predição, em nosso caso, a determinação de áreas potenciais para pesquisa de recursos minerais e para alocação de áreas para diversos usos. Para isso, as técnicas de análise espacial mais utilizadas são as de sobreposição de mapas, com combinação de atributos pelas regras booleanas e por índices somativos e multiplicativos, os métodos de *fuzzy logic* e estatística condicional (bayesiana), conforme já utilizados pelo grupo (Macedo, 1993; Araújo, Macedo e Campanha, 1995) e cuja metodologia é detalhada em Eastman et alii (1993, 1995) e Bonham-Carter (1988, 1994).

7.7 Apoio à decisão

Finalmente, a alocação de áreas pode ser apoiada por técnicas de comparação entre os possíveis usos, usando técnicas de Multi-objetivos/Multi-critérios, conforme a denominação de Eastman et alii (1993, 1995). Um problema ocorrente na área, cuja resolução pode ser muito auxiliada por essas técnicas, é a decisão entre a alocação de uma determinada área para mineração, preservação ambiental ou agricultura.

Este problema está sendo tratado em dois projetos, em início de execução, por um mestrando (Marcos de Souza Campos) e um bolsista de iniciação científica (Alexandre Carnier Nunes da Silva), que analisarão a folha Iporanga a 1:50.000, verificando a favorabilidade da alocação de áreas para mineração ou preservação ambiental.

7.8 Produção cartográfica

Os sistemas de Geoprocessamento permitem o ajuste de escalas, a produção de legendas e padrões gráficos e a impressão de mapas básicos ou derivados. Isto é feito com resultados variados, conforme a ênfase dos autores do sistema. Assim, o ARC/INFO e o MGE da Intergraph apresentam resultados

de alta qualidade, usados para impressão de mapas em grandes edições. O IDRISI (Eastman, 1992 *a* e *b*, 1995), embora um bom sistema para análise espacial, deixa a desejar na produção cartográfica. Isto levou Jeffrey Jones, já autor do módulo de digitalização Tosca (Jones, 1994), a fundar uma companhia e produzir um complemento, TMAPPER (Tsoft Co, 1994), para imprimir resultados do IDRISI. Esta produção de complementos (*add-on programs*) é comum em Geoprocessamento, havendo dezenas de programas deste tipo para adaptar o AutoCAD, destinado a auxiliar projetos de engenharia, para trabalhos em Geologia e Geoprocessamento.

Os trabalhos de Geoprocessamento, principalmente os de Geologia e Meio Ambiente, desenvolvidos no IG-USP, exigem para sua edição equipamentos raster (eletrostáticos, a laser ou a jato de tinta), pois grandes superfícies devem ser cobertas de tinta, o que não é feito eficientemente por plotters de pena. A grande eficiência e o baixo custo de aquisição e operação dos novos equipamentos, principalmente de jato de tinta, tornaram definitivamente obsoletos os processos manuais de edição de mapas, levando para o mesmo caminho os de composição fotográfica, a não ser para a edição de grande número de cópias de um mesmo original.

Devido a dificuldades de adaptação de equipamentos (o plotter de jato de tinta do IG-USP foi adquirido recentemente) e ao custo, nesta revisão apenas são apresentadas composições cartográficas de uma página, produzidas em impressoras a laser e jato de tinta.

7.9 - Exemplo de aplicação de geoprocessamento à análise metalogenética regional: folhas Cerro Azul e Apiaí

A área estudada, com a finalidade de desenvolvimento metodológico, é a coberta pelas folhas Cerro Azul e Apiaí a 1:100.000 (SG22-X-B-IV e V), sendo parte da folha Itararé a 1:250.000 (SG.22-X-B). A geologia da área é descrita em Campanha et alii (1996)

A geologia da folha Itararé, segundo esses autores, caracteriza-se por rochas supracrustais de grau metamórfico fraco a médio. O embasamento é constituído por rochas gnáissico-migmatíticas com núcleos charnockíticos e intercalações de metassedimentos. São intrudidas por granitos do Proterozóico Superior ao Cambro-Ordoviciano e por alcalinas básicas mesozóicas. As rochas supracrustais são predominantemente do Proterozóico Médio, enquanto seu embasamento pertence ao Proterozóico Inferior e ao Arqueano.

As figuras 33 a 38 apresentam traços geológicos das folhas Apiaí e Cerro Azul a 1:100.000, decompostos em litologia, estratigrafia e idades das unidades estratigráficas.

As feições estruturais adotadas para a análise foram todas as consideradas como falhamentos e contatos por falha, nas áreas analisadas, segundo o mapa de Campanha et alii (1996). Não foram incorporadas as direções de dobramento nem feita distinção entre tipos de falhamentos.

Para a geoquímica foram digitalizadas as anomalias em teores de sedimentos de corrente e concentrados de batéia, apresentadas nos mapas do Projeto Mapas Metalogenéticos e Previsionais, folha Itararé, compiladas pelo autor e sua equipe da CPRM (Hama e Algarte, 1986). Foi montado um banco de dados, que compreende as bacias de drenagem, armazenadas em formato vetorial e matricial, e uma tabela padrão Access com teores e informações de coleta.

Esta escolha de dados tem os defeitos seguintes:

- utiliza as anomalias já determinadas nos projetos dos quais foram compiladas para o Mapa Metalogenético, em detrimento dos dados brutos. Isto foi feito, em primeiro lugar, por não serem em muitos casos disponíveis os dados brutos. Nos casos em que eram disponíveis, havia dificuldade em processar

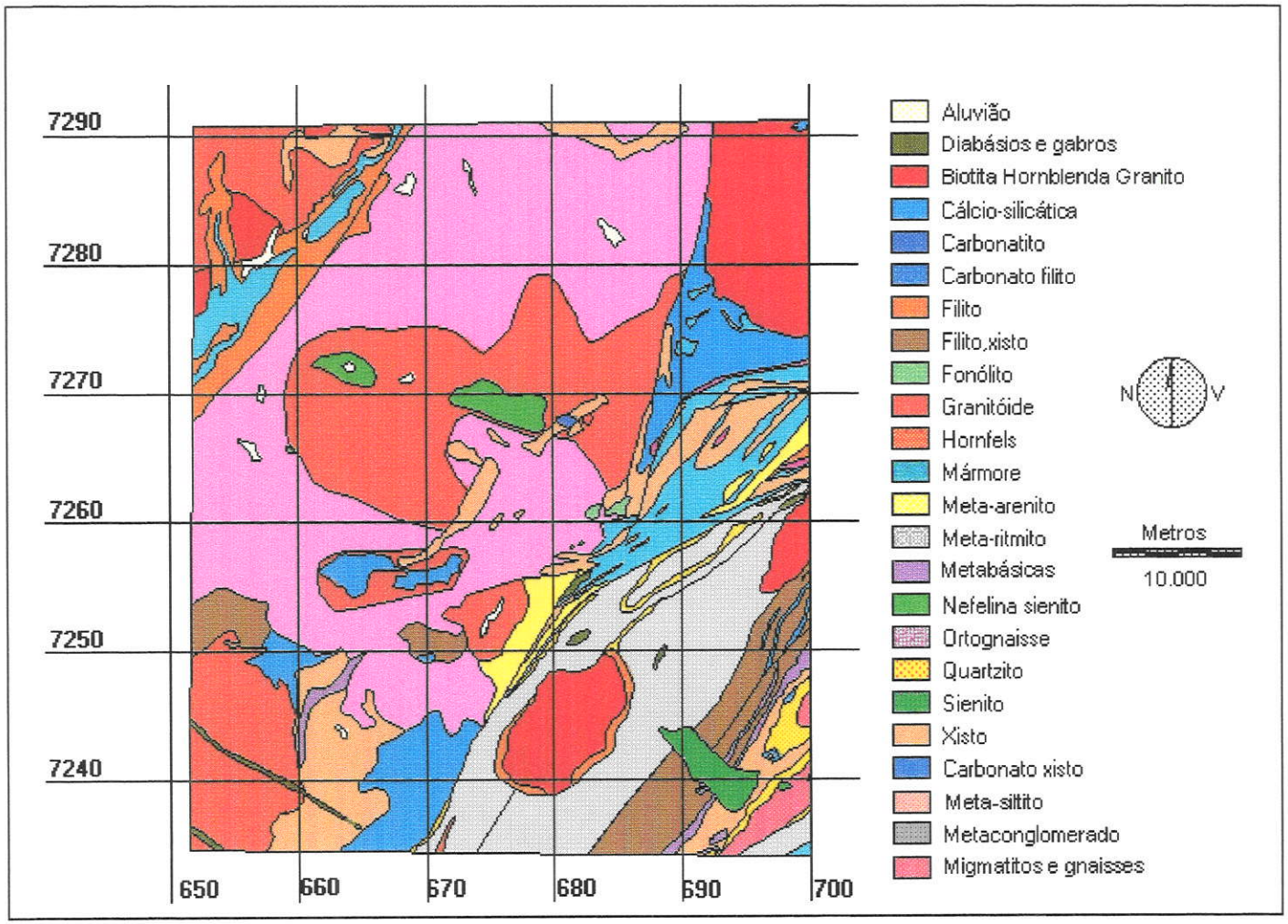


Figura 33: Litologia - Quadrícula Cerro Azul

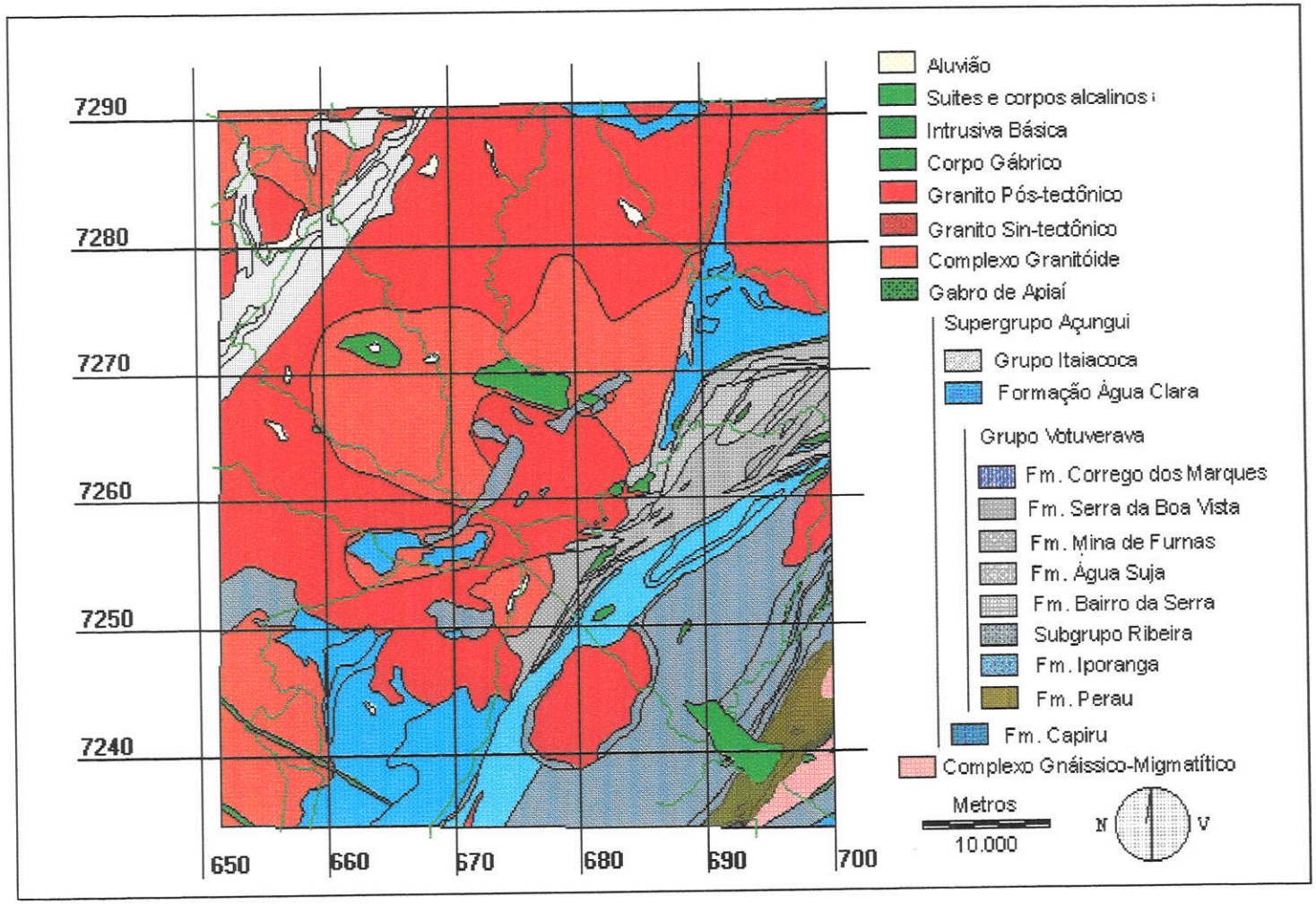


Figura 34- Estratigrafia - Quadrícula Cerro Azul

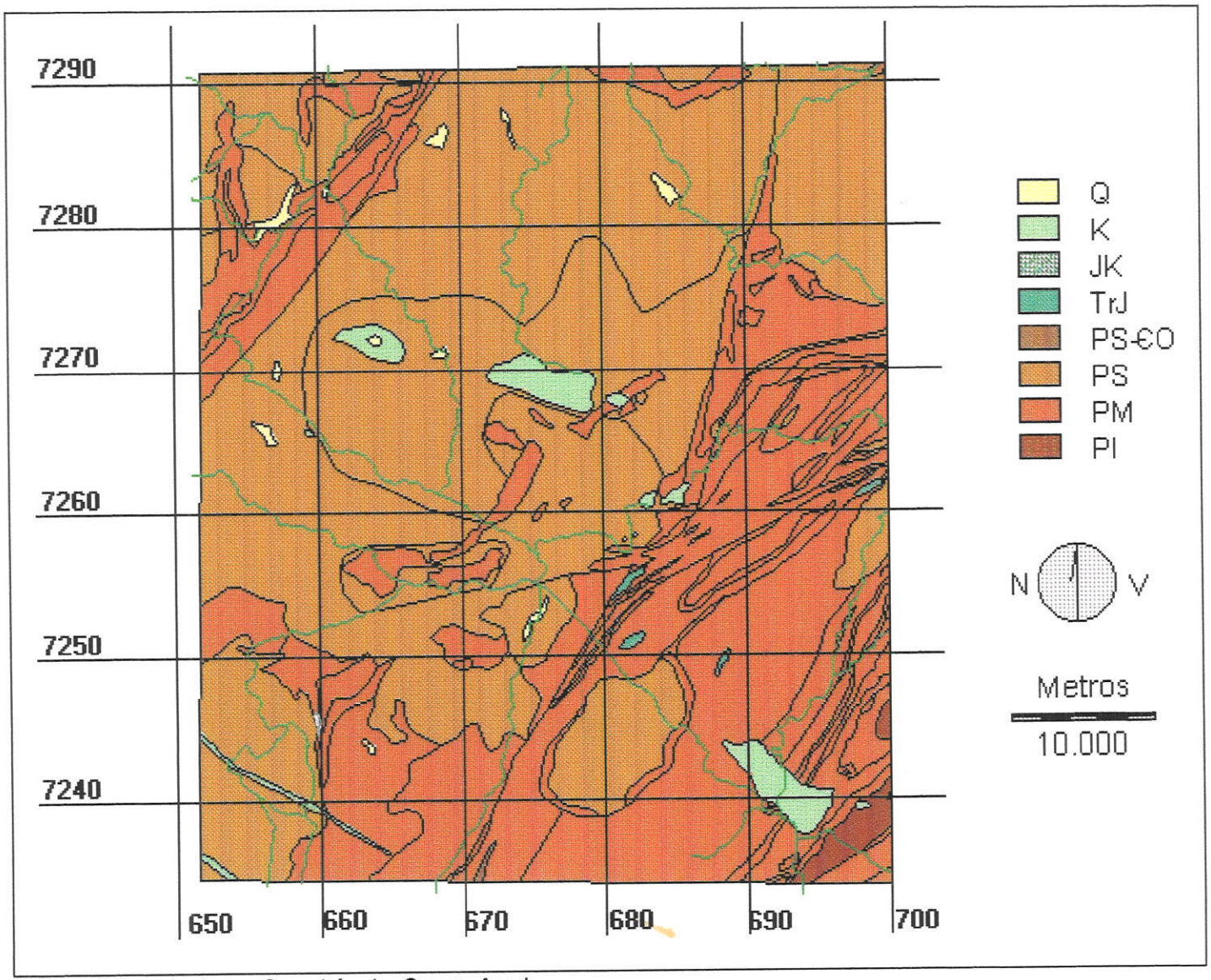


Figura 35 - Idades - Quadrícula Cerro Azul.

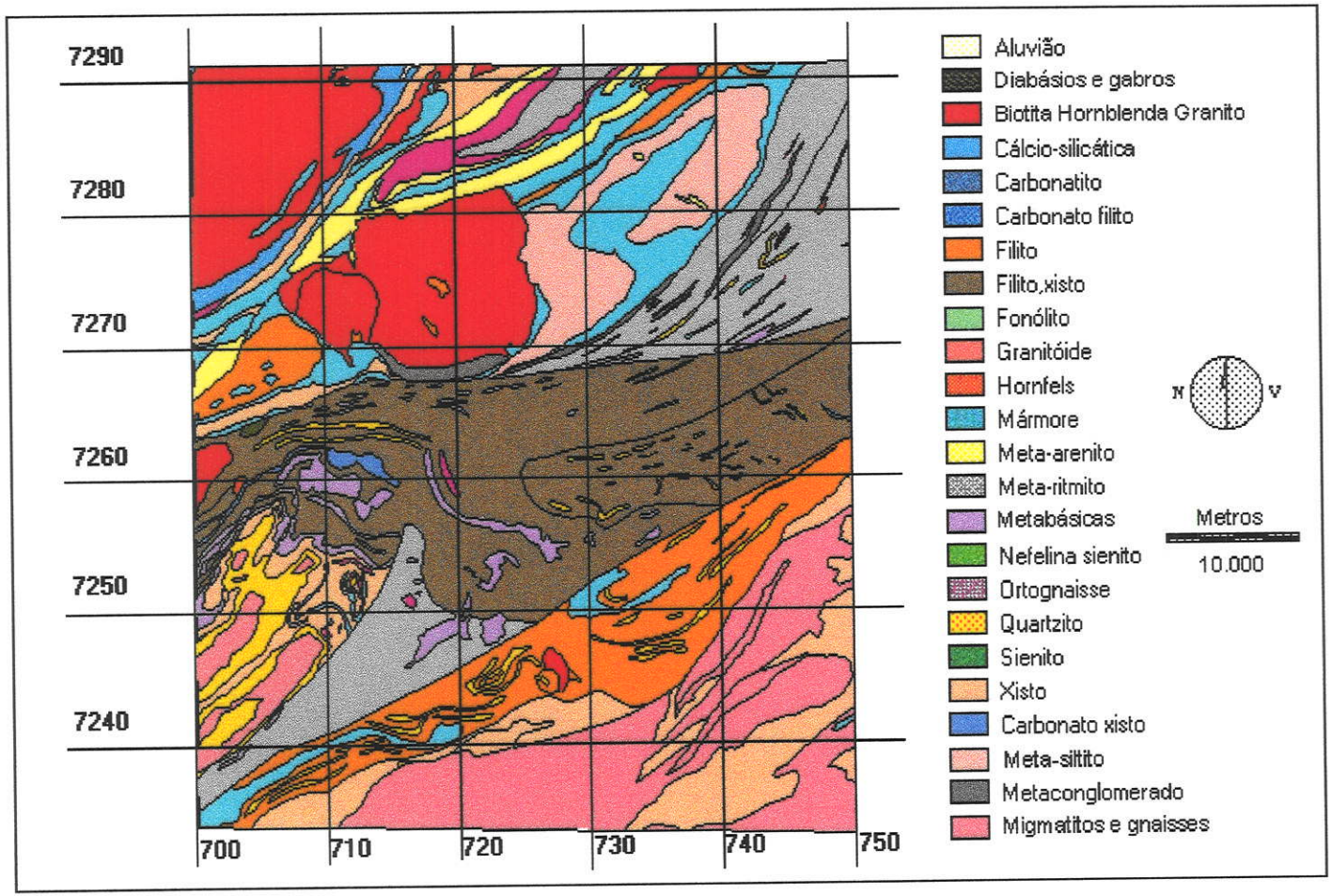


Figura 36: Litologia - Quadrícula Apiaí

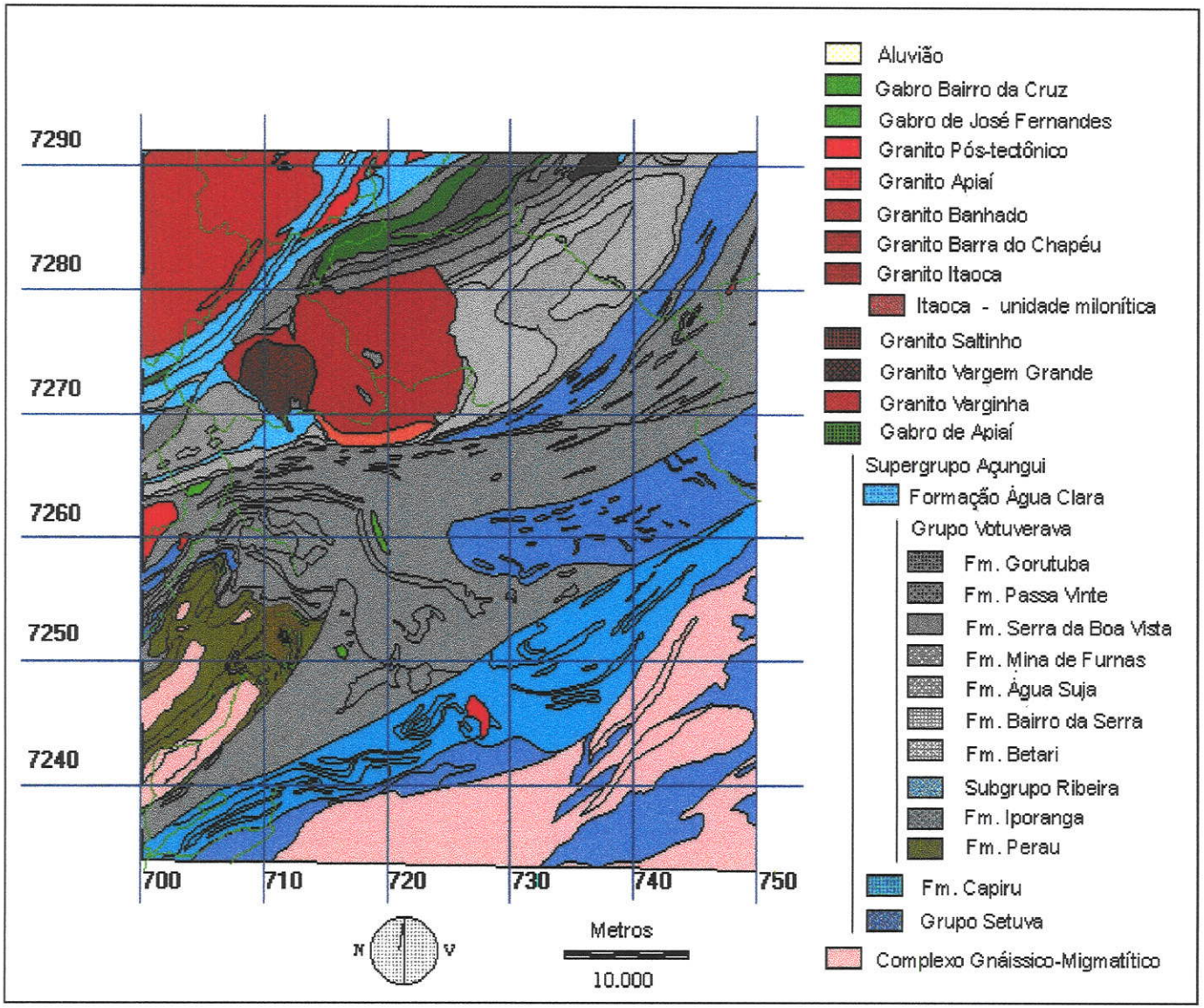


Figura 37: Estratigrafia - Quadrícula Apiaí

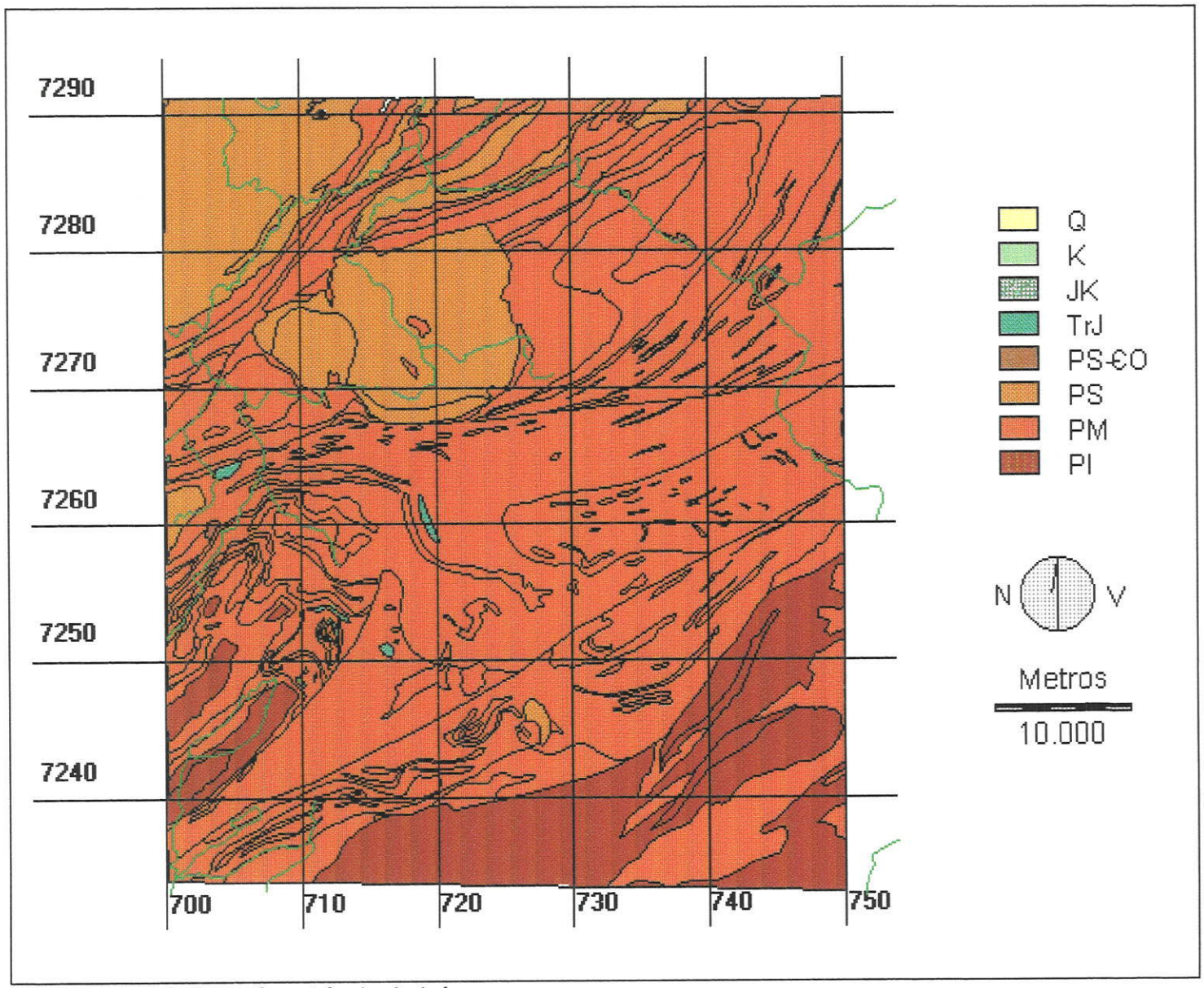


Figura 38 - Idades - Quadrícula Apiaí.

conjuntamente dados coletados e analisados por técnicas diferentes, em laboratórios diferentes, em alguns casos sem controle de qualidade.

- algumas bacias de drenagem, nas quais ocorrem depósitos conhecidos, não foram amostradas pelo projeto Geoquímica no Vale do Ribeira (Morgental et alii, 1978), principal fonte dos dados regionais. Isto foi feito para evitar a coleta de sedimentos com teores altos, derivados de anomalias naturais ou da contaminação por trabalhos de pesquisa ou lavra. Esta medida, que seria justificável num projeto puramente de prospecção, é prejudicial aos trabalhos que necessitam uma cobertura total da área e padrões completos de distribuição de teores.

- não incorpora dados mais recentes, da Mineropar e do Pro-Minério. Neste caso são disponíveis os teores brutos, mas, pelo menos para os do Pro-Minério, não são disponíveis as anomalias. Até agora não foi possível juntar estes dados aos anteriores.

Além desses problemas com os dados, a análise ressentiu-se da falta de dados geofísicos e de melhor exploração dos dados estruturais e tectônicos.

Tendo em vista os problemas apontados, os resultados devem ser considerados parciais, e deve-se levar em conta o objetivo principal dessas análises, que foi o da incorporação de tecnologia e treinamento da equipe.

Na fase atual dos trabalhos, com alguns alunos desenvolvendo projetos detalhados em áreas contidas na Folha Itararé, passamos a tornar rigorosa a coleta de dados (incorporando todos os disponíveis até a data) e a análise (com melhor exploração das interpretações dos diversos fatores). Esses trabalhos estão em andamento e constituirão um salto qualitativo em relação aos aqui apresentados.

Conforme discutido no item 7.5.1, o emprego de geoprocessamento para análise metalogenética pode ser feito por duas linhas de abordagem (Bonham-Carter, 1994):

- **baseada nos dados *(data-driven)*:** que parte das mineralizações conhecidas, analisa por geoprocessamento os padrões de associação delas com fatores geológicos e monta, a partir desses padrões, um modelo empírico-

espacial de potencialidade mineral, que será aplicado a áreas em que ocorrem padrões semelhantes àqueles de onde ocorrem mineralizações;

- **baseada no conhecimento *(knowledge-driven)*** : que utiliza modelos de mineralização conhecidos, empíricos ou genéticos, e monta modelos espaciais a partir dos fatores desses modelos.

Nos trabalhos comentados (Macedo, 1993; Araújo, Macedo e Campanha, 1995), apenas foi adotada a abordagem baseada no conhecimento, utilizando modelos adequados às mineralizações que se esperava encontrar nas áreas trabalhadas. Em trabalhos em andamento está sendo iniciada a análise baseada nas ocorrências. Isto não foi feito nos primeiros trabalhos porque é necessária a conferência minuciosa das ocorrências, para empregá-las na análise, principalmente no que concerne a sua localização geográfica e geológica, o que exige comparação dos arquivos com os trabalhos originais, e até trabalho de campo, devido ao grande número de erros encontrados nas bases de dados dados digitais (IDEM e SIGA).

Dois modelos de depósitos minerais foram considerados: o de sulfetos maciços de origem exalativa (vulcanogênica ou sedimentar-exalativa), denominado regionalmente de modelo Perau, e o de sulfetos filonianos de origem hidrotermal, com a denominação regional de modelo Rocha (JICA/MMAJ, 1984).

Na busca de mineralizações segundo o modelo de sulfetos maciços (Perau) são considerados fatores importantes para a busca de mineralizações os seguintes:

- fator litológico: litologias associadas a seqüências vulcânicas e sedimentares marinhas a elas associadas (metavulcânicas ácidas e básicas e rochas cálcio-silicáticas);
- fator geoquímico: anomalias de Cu, Pb e Zn;
- fator estrutural: proximidade de falhamentos.

Para o modelo de depósitos filonianos de sulfetos de origem hidrotermal (Rocha) foram considerados os fatores:

- litológicos: mármores, granitos e granitóides (na escala trabalhada não estão representados enclaves e restos de teto em rochas graníticas, que podem conter mineralizações, como na área de Itaoca);

- fator geoquímico: anomalias de Cu, Pb e Zn;
- fator estrutural: proximidade de falhamentos.
- fator térmico e químico: proximidade dos contatos de granitos.

Uma vez que a influência dos lineamentos não se limita a uma linha, mas abrange uma área que pode ser atingida por fluidos que se deslocam pelas falhas principais e por falhas derivadas, transversais a elas, o fator estrutural é expresso por um corredor (*buffer*) de 1000 m de largura.

Também foi feito um corredor para os granitos, igualmente de 1000 m, distância na qual considera-se que pode haver influência térmica ou química, derivada de fluidos provenientes das intrusões.

O mapa de fator geoquímico é obtido pela sobreposição (*overlay*) dos mapas de anomalias para cada um dos elementos.

Os métodos de combinação dos mapas individuais, para montar os mapas de fatores (como o de fator geoquímico), e dos mapas de fatores para deles derivar a favorabilidade a um determinado modelo de mineralização (mapa de potencial mineral), diferem na maneira que atribuem pesos aos componentes.

Eastman et alii (1993, 1995) distinguem entre fatores e restrições (*constraints*). Os fatores exprimem graus diversos de favorabilidade, e podem ser combinados por operações de soma booleana. As restrições exprimem a impossibilidade da ocorrência do fenômeno procurado, obrigando à utilização da multiplicação booleana. Assim, são feitos mapas em que os fatores são expressos por valores positivos, e as restrições são expressas por zeros. A multiplicação booleana faz com que qualquer parte do mapa final em que ocorra um zero nos mapas componentes seja também zero. Para a combinação dos fatores, o método booleano atribui pesos unitários aos diversos fatores, combinando-os por soma simples.

Um processo mais flexível que o booleano é o ponderado, também denominado de índices somativos, que consiste em atribuir pesos aos diversos fatores, conforme sejam considerados mais ou menos favoráveis ao fenômeno procurado. Os pesos podem ser atribuídos pelo pesquisador, ou, mais apropriadamente, por especialistas nos modelos considerados. Neste caso os

pesos atribuídos pelos diversos pesquisadores são combinados, resultando um mapa composto para cada fator.

Maiores refinamentos podem ser incluídos juntando-se aos pesos probabilidades, ou intervalos de variação, que exprimem a incerteza na atribuição dos pesos. Outra abordagem é anexar possibilidades, que exprimem a incerteza na definição dos fatores *(fuzzy method)*.

Para o exemplo a seguir apresentado foi empregada a soma ponderada simples, com pesos atribuídos arbitrariamente.

Os diversos mapas originais (figuras 33 a 38), e os mapas de anomalias geoquímicas para teores individuais resultaram nos mapas de fatores:

- mapas de fatores litológicos para o modelo Perau, folhas Cerro Azul e Apiaí, destacando as unidades consideradas como metavulcânicas básicas, rochas cálcio-silicáticas e os xistos, considerados como prováveis metavulcânicas ácidas, apresentados nas figuras 39 e 40.
- mapas de fatores litológicos para o modelo Rocha, folhas Cerro Azul e Apiaí, destacando os granitos, granitóides e mármores, apresentados nas figuras 41 e 42.
- Mapa de somas de anomalias em sedimentos de corrente para Cu, Pb e Zn, atribuídas às bacias de captação, apresentado na figura 43. Este mapa foi reclassificado para obter um mapa de fator geoquímico, booleano (em que os pixels das áreas em que há anomalias têm valor um e os das sem anomalia têm valor zero), não mostrado no texto.
- Mapa de fator estrutural: apresenta um corredor (*buffer)* de 1.000 m ao redor de falhamentos, para as duas folhas, na figura 44.
- Mapas de fator térmico e químico: com um *buffer* de 1.000 m ao redor de contatos de granitos, para as duas folhas, na figura 45.

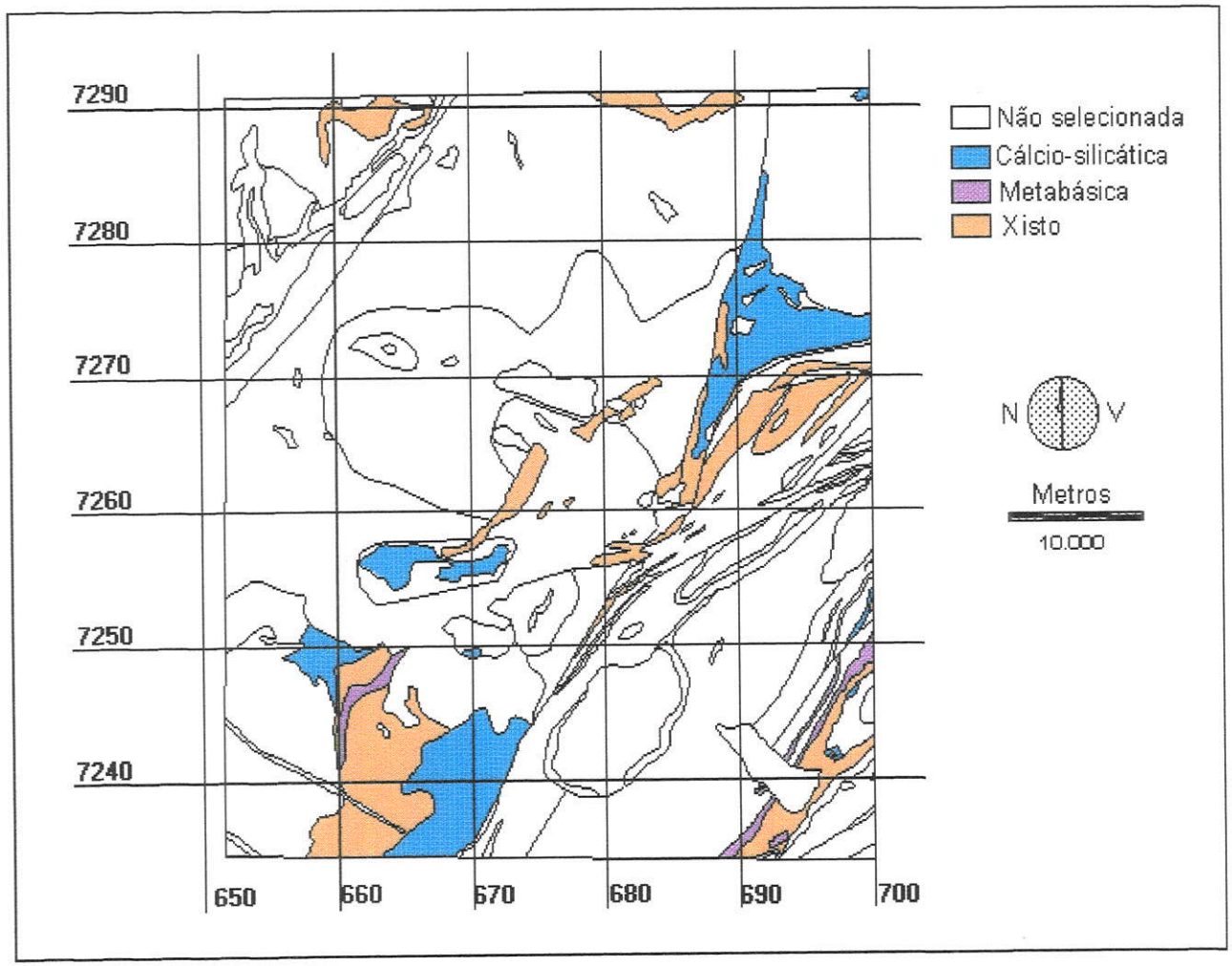


Figura 39: Mapa de fator litológico para o modelo Perau, quadrícula Cerro Azul

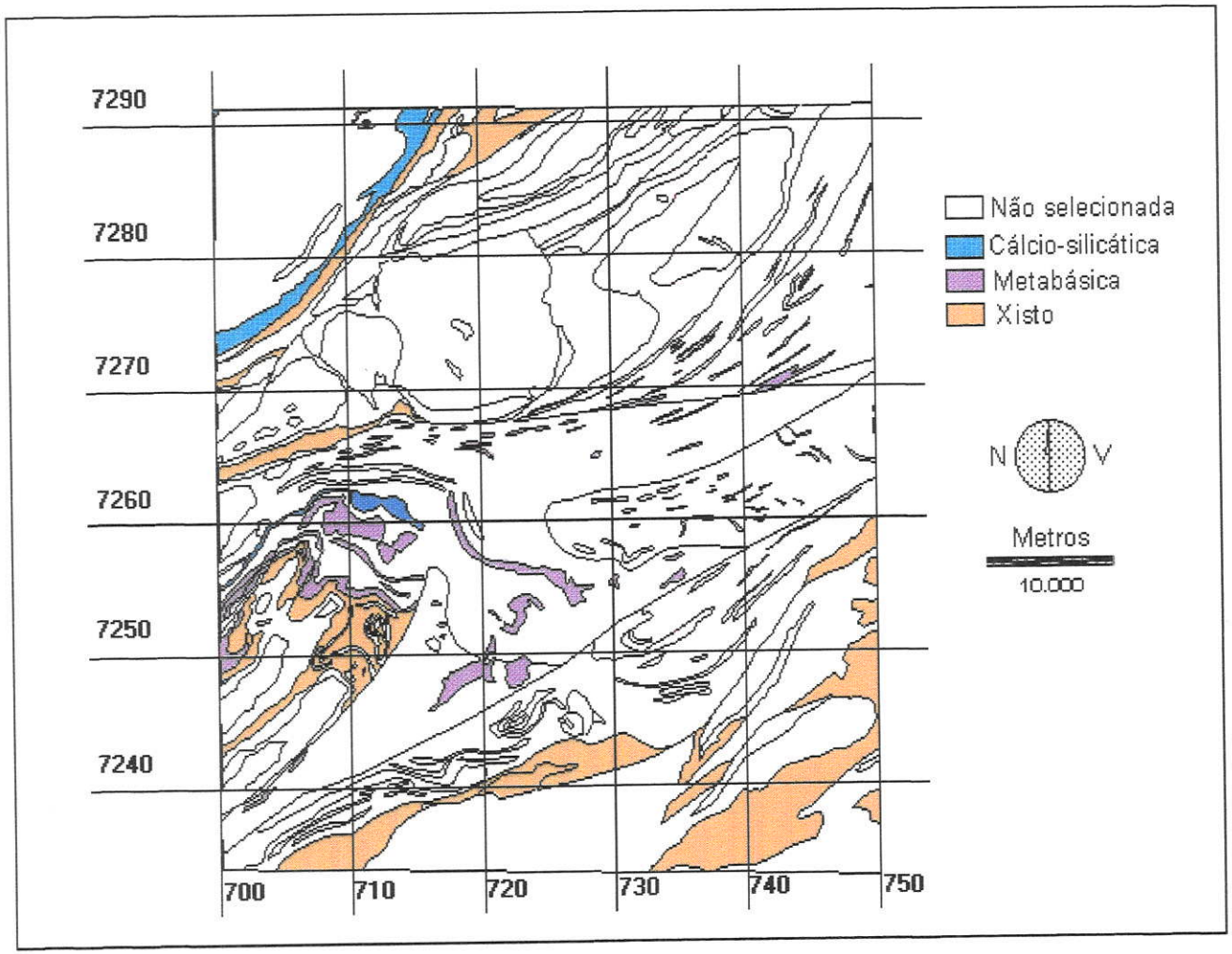


Figura 40: Mapa de fator litológico para o modelo Perau, quadrícula Apiaí

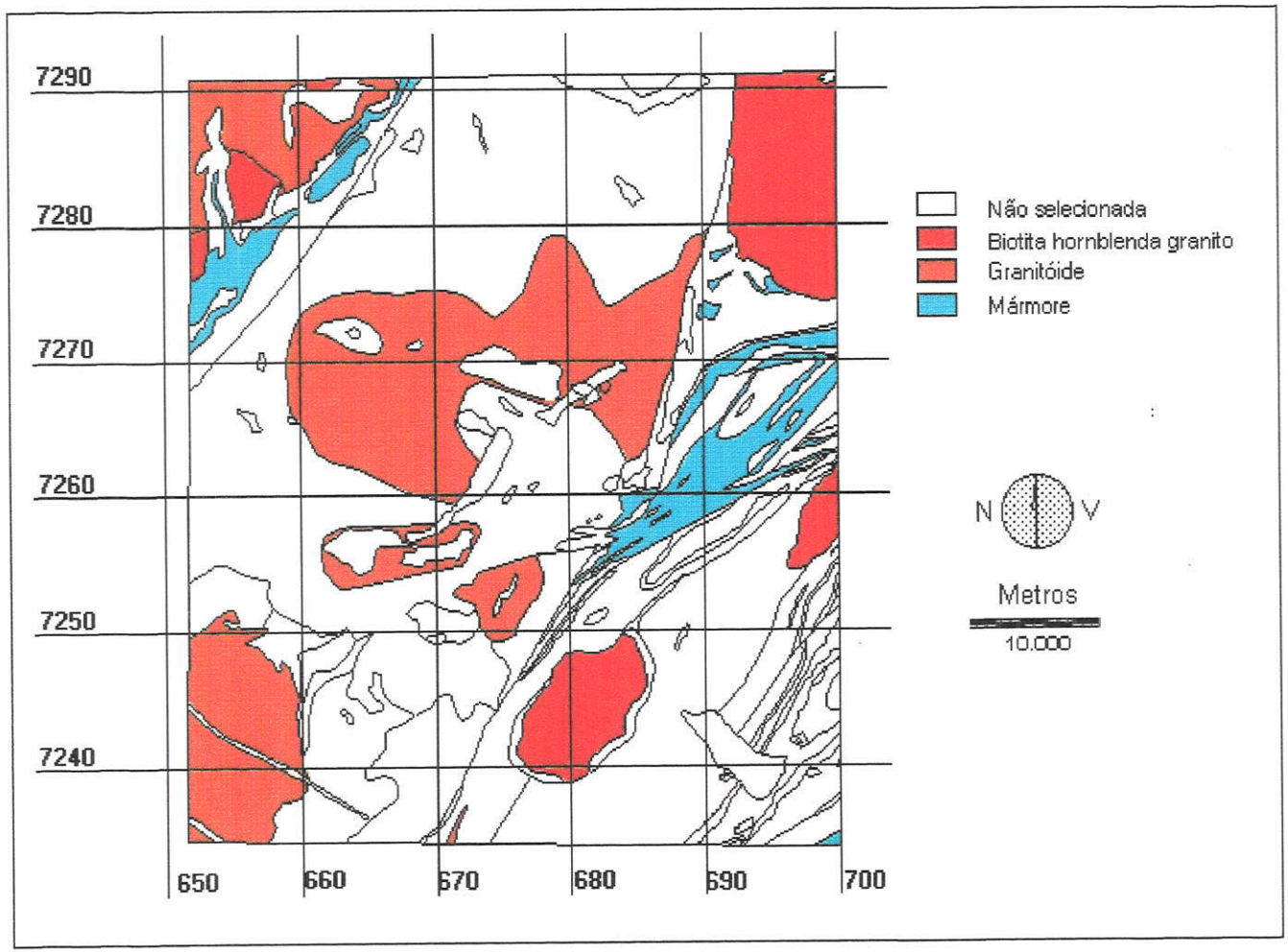


Figura 41 - Mapa de fator litológico para o modelo Rocha, quadrícula Cerro Azul.

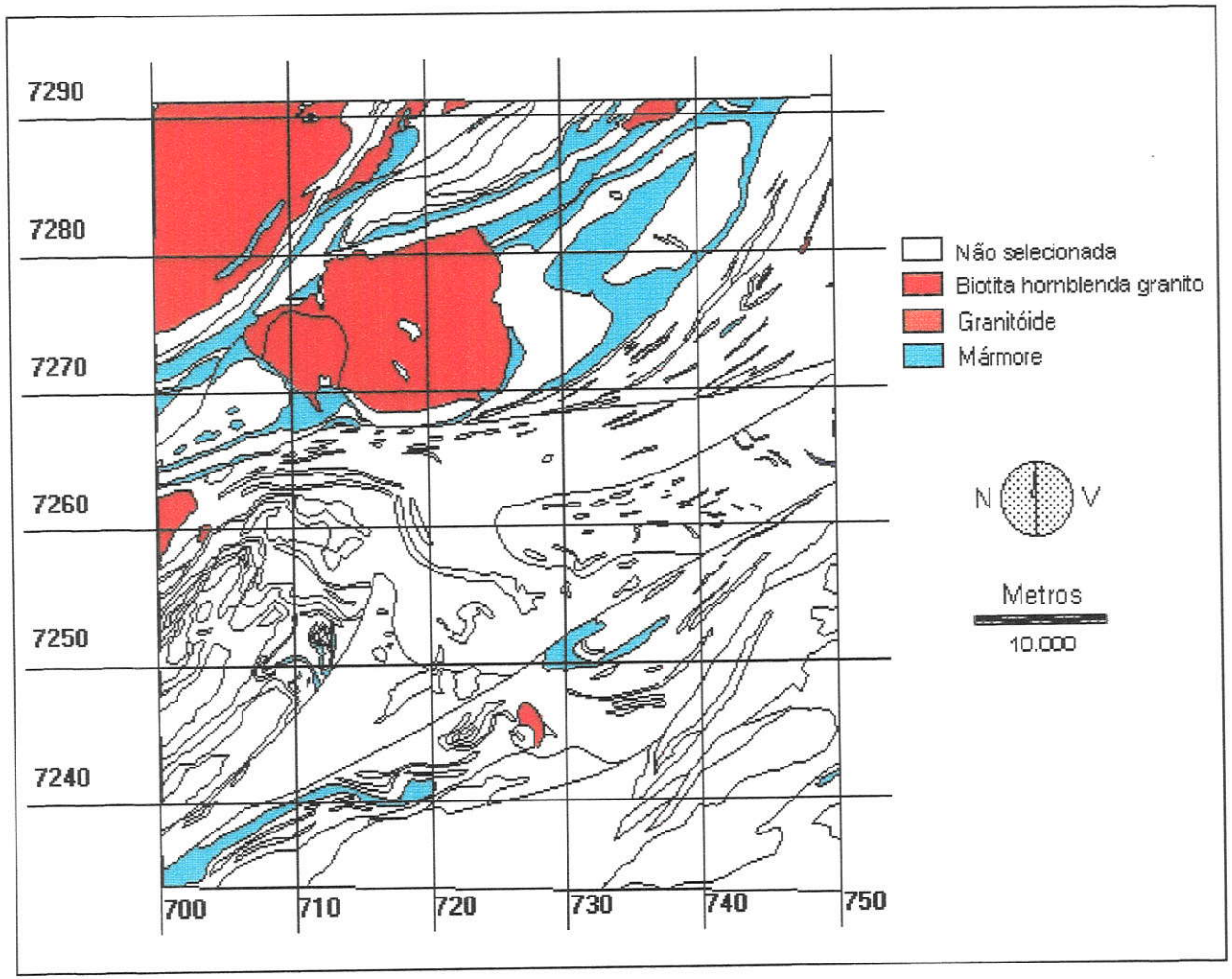


Figura 42 - Mapa de fator litológico para o modelo Rocha, quadrícula Apiaí.

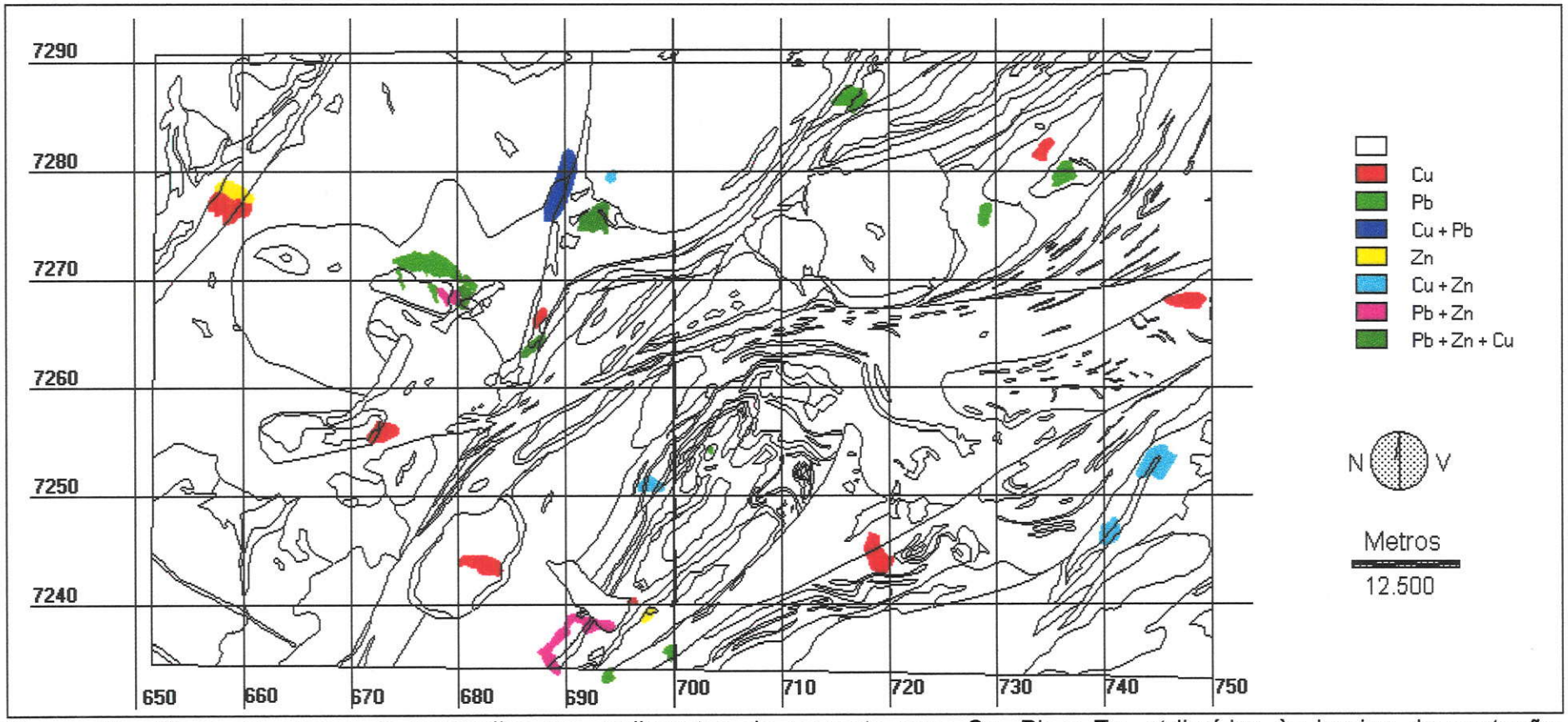


Figura 43 - Mapa de somas de anomalias em sedimentos de corrente para Cu, Pb e Zn, atribuídas às bacias de captação, quadrículas Cerro Azul e Apiaí.

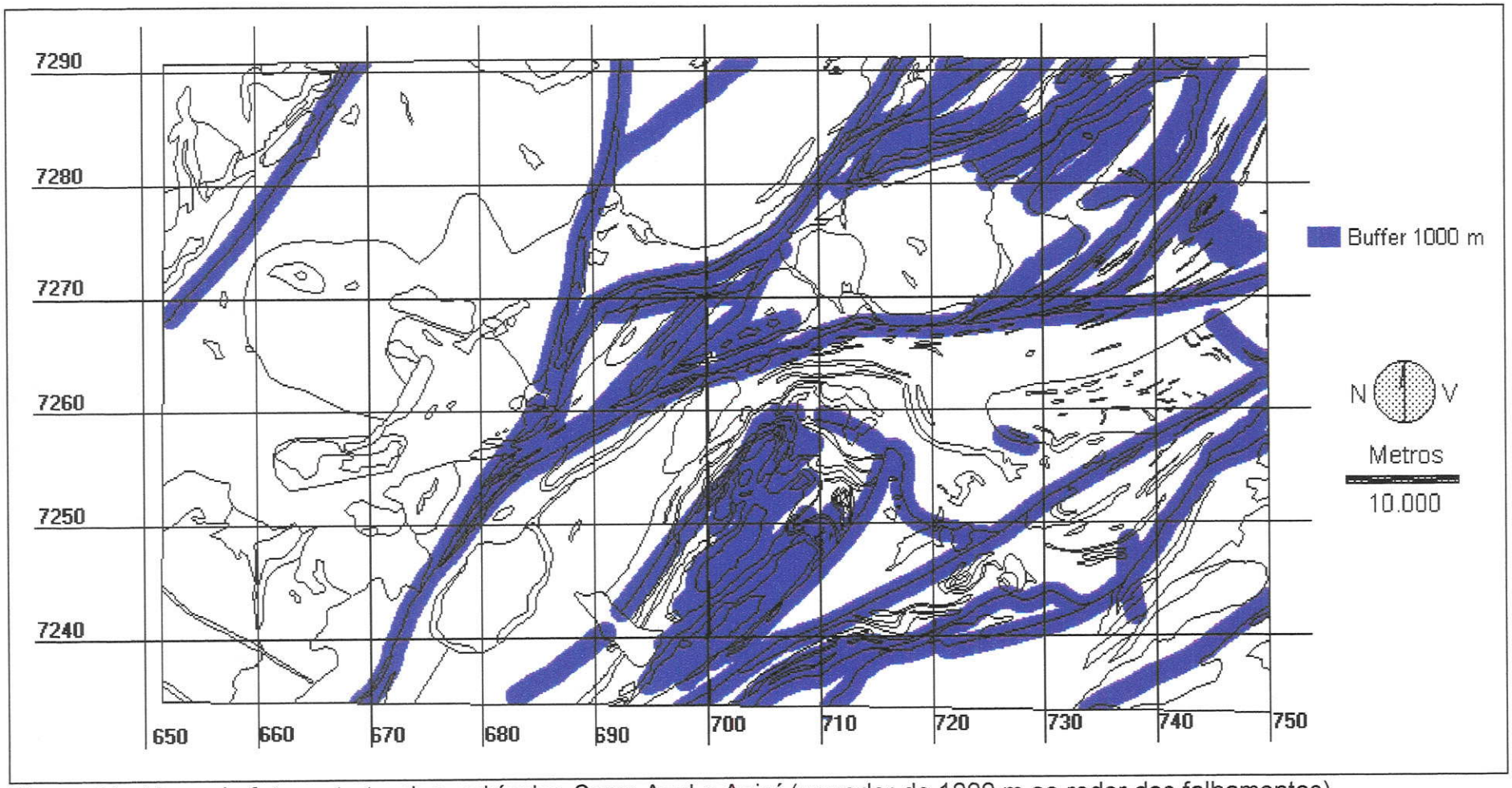


Figura 44 - Mapa de fator estrutural, quadrículas Cerro Azul e Apiaí (corredor de 1000 m ao redor dos falhamentos).

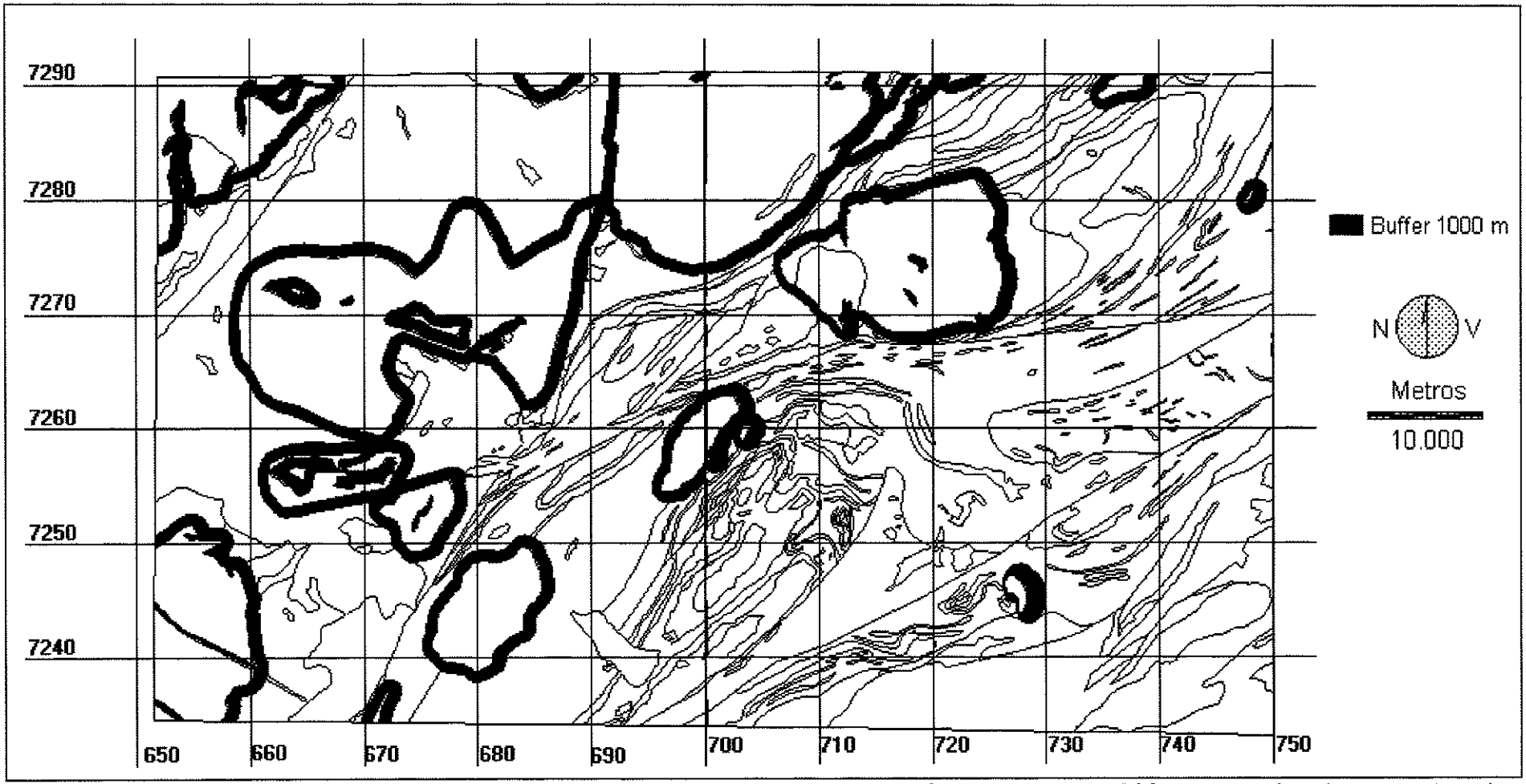


Figura 45 - Mapa de fator térmico e químico, quadrículas Cerro Azul e Apiaí (corredor de 1000 m ao redor dos contatos dos granitos).

Esses mapas foram combinados, somando os pontos para favorabilidade à ocorrência das mineralizações, com a ponderação seguinte:

- foi atribuído peso dois aos fatores litológico e geoquímico e peso um ao fator estrutural, para o modelo Perau, resultando nos mapas de favorabilidade apresentados nas figuras 46 e 47, para as folhas Cerro Azul e Apiaí, com índices de favorabilidade variando de zero a cinco.
- foi atribuído peso dois aos fatores litológico e geoquímico e peso um aos fatores estrutural e térmico-químico, para o modelo Rocha, resultando nos mapas de favorabilidade apresentados nas figuras 48 e 49, para as folhas Cerro Azul e Apiaí, com índices de favorabilidade variando de zero a seis.

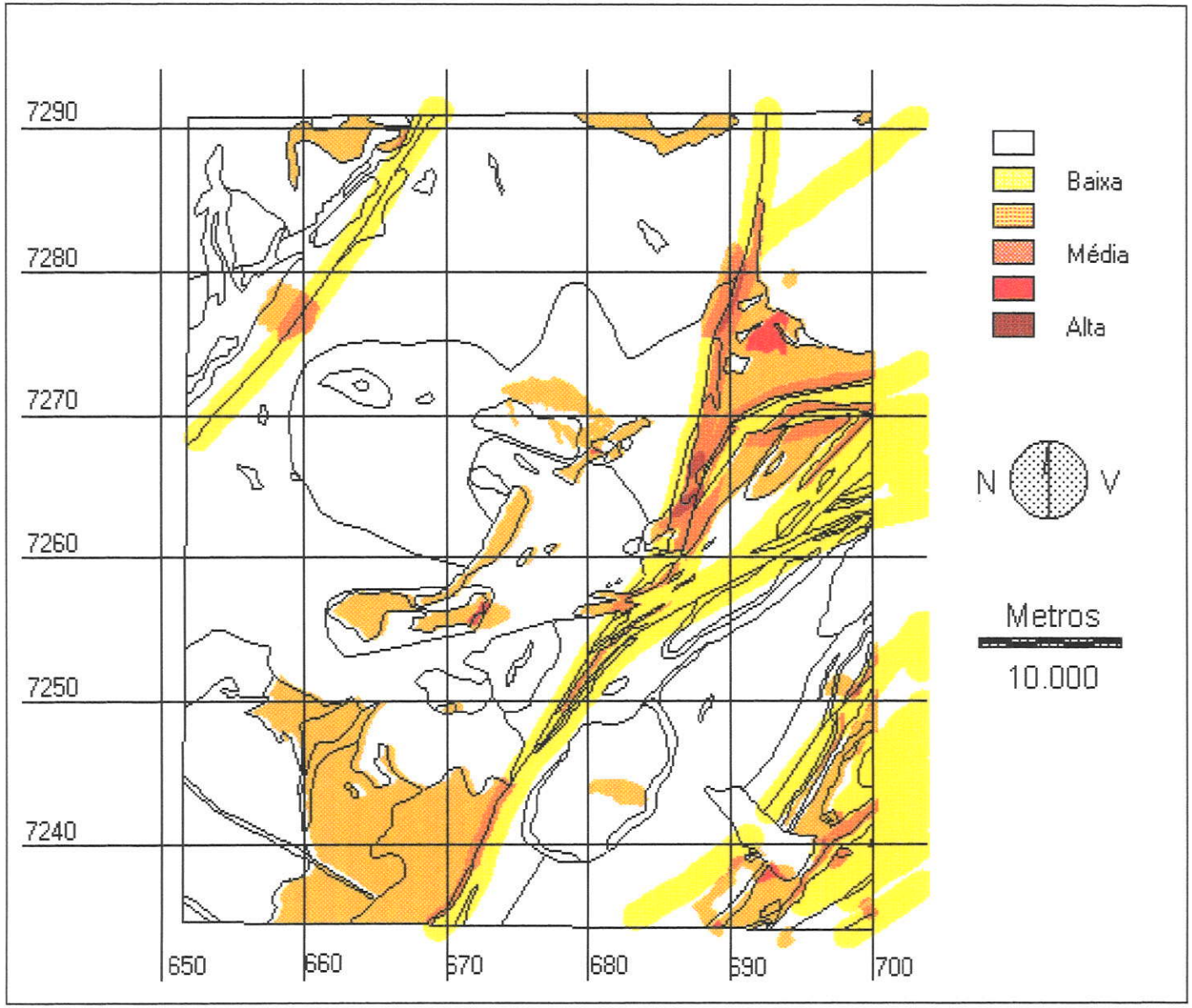


Figura 46 - Mapa de favorabilidade para depósitos minerais do tipo Perau, quadrícula Cerro Azul.

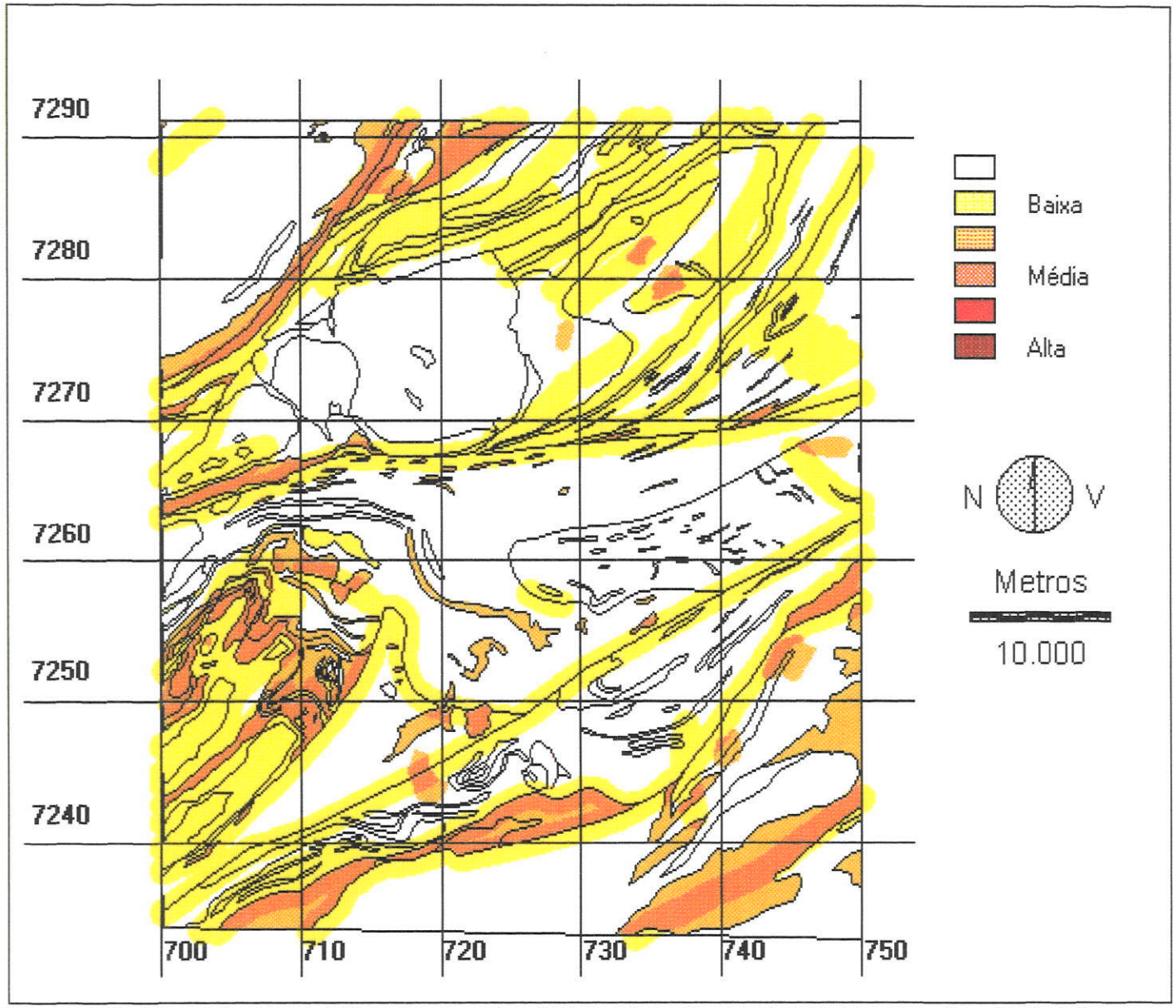


Figura 47 - Mapa de favorabilidade para depósitos minerais do tipo Perau, quadrícula Apiaí.

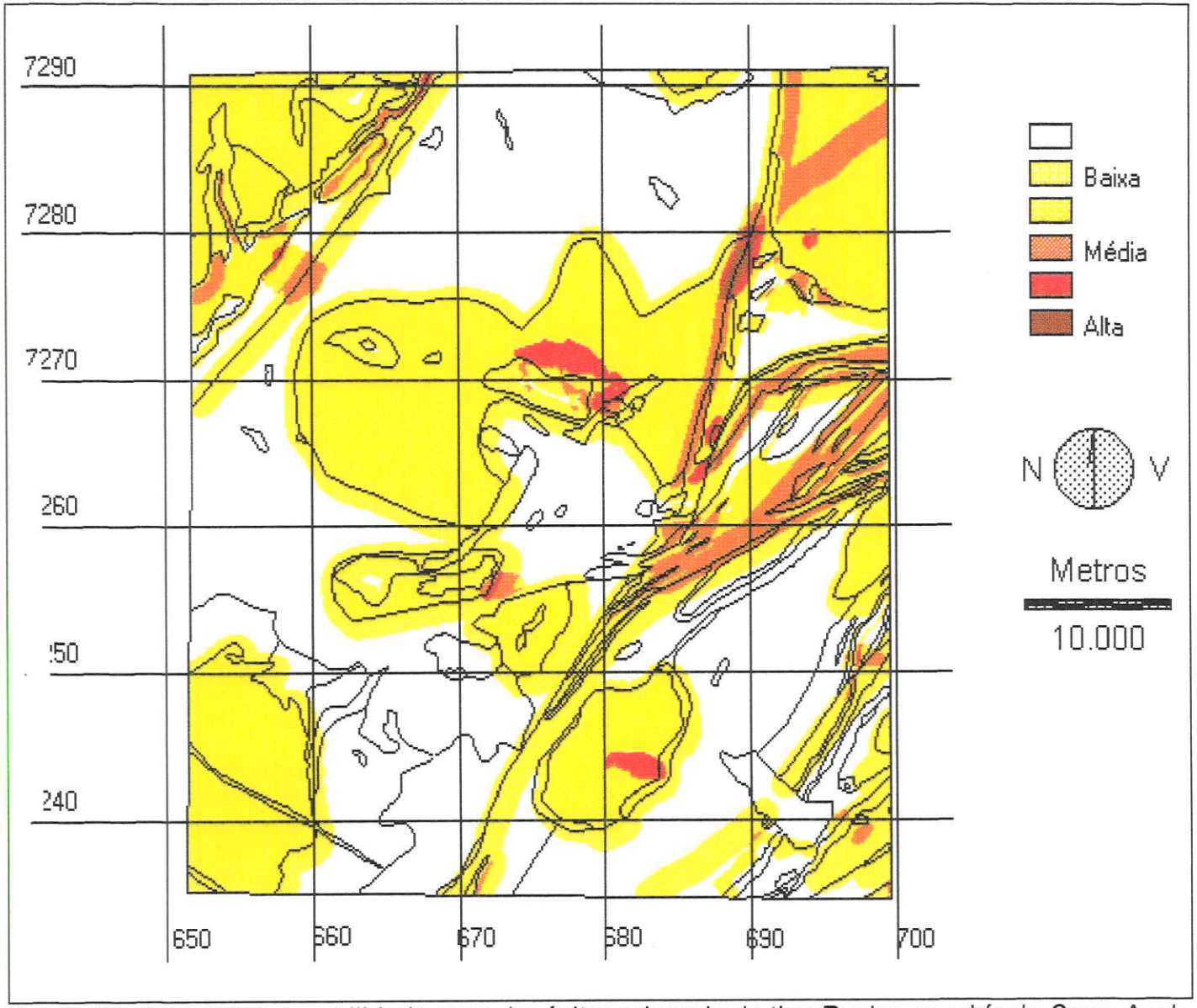


Figura 48 - Mapa de favorabilidade para depósitos minerais do tipo Rocha, quadrícula Cerro Azul.

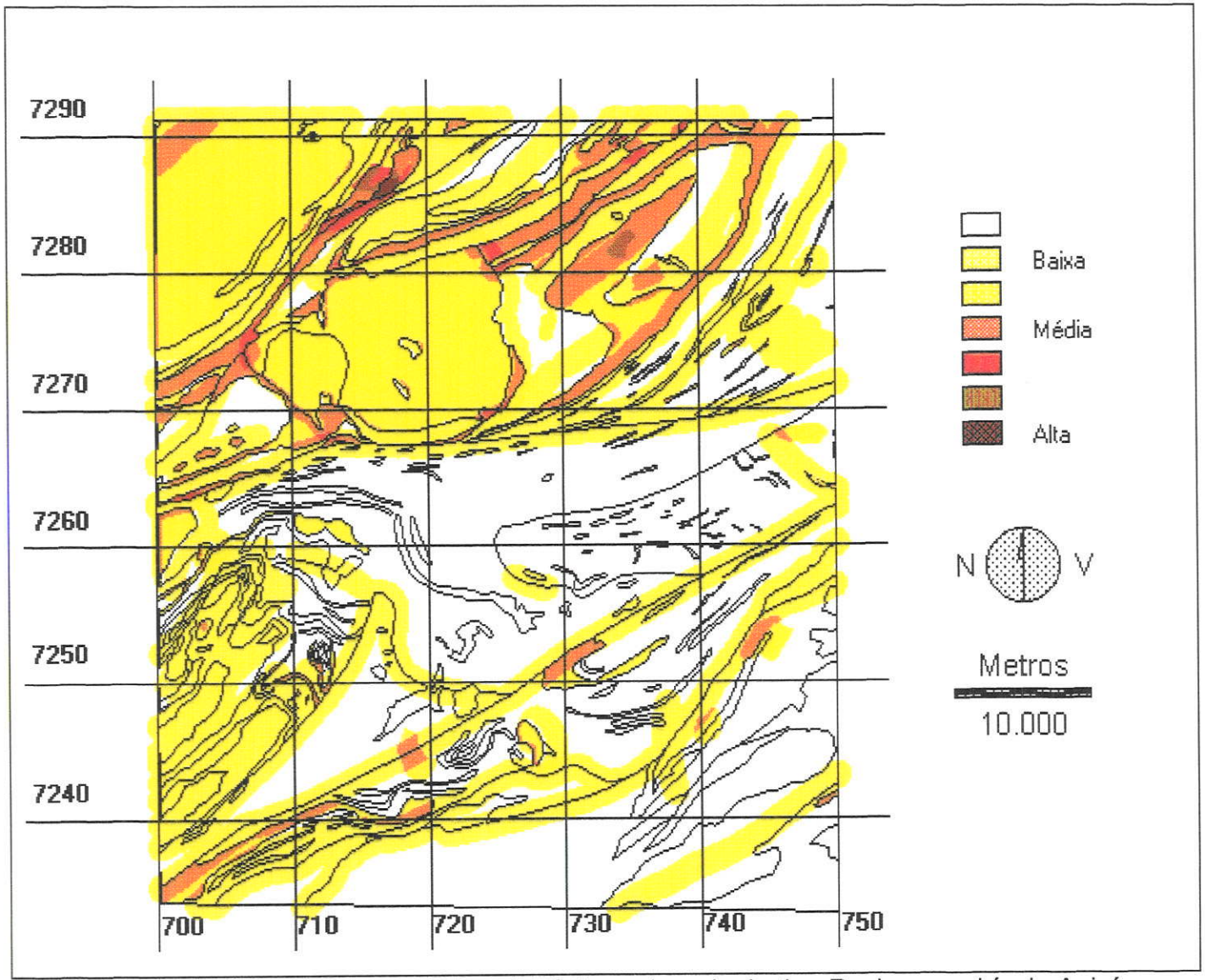


Figura 49 - Mapa de favorabilidade para depósitos minerais do tipo Rocha, quadrícula Apiaí.

Os resultados obtidos são semelhantes aos obtidos pelo processo manual (Hama e Algarte, 1986), com algumas diferenças:

- em alguns casos, como o da indicação da totalidade dos granitos como favoráveis a mineralizações do tipo Rocha, as limitações dos dados levam a excessiva generalização, com grandes áreas consideradas como de potencialidade baixa, mas positiva;
- o mesmo resultado ocorre com os buffers traçados ao redor de todos os falhamentos e granitos, sem uma análise prévia, também aumentando as áreas de favorabilidade baixa;
- também aumentam as áreas favoráveis por serem considerados apenas os fatores litológicos e não os estratigráficos, misturando rochas de idades e ambientes de formação diferentes.
- em Hama e Algarte (1986) foi atribuído peso máximo à ocorrência de mineralizações, não consideradas neste estudo.

Para um teste prático foram comparadas as ocorrências consideradas pelo projeto Mapas Metalogenéticos (Hama e Algarte, 1986) como dos tipos estratiforme e filoniano (aproximadamente correspondentes aos tipos Perau e Rocha) com as áreas favoráveis, no total das folhas Cerro Azul e Apiaí. Duas técnicas foram empregadas:

- gráfica, fazendo os mapas das figuras 50 e 51. Nestes os pontos correspondentes às ocorrências parecem associados às áreas favoráveis.
- numérica, fazendo uma tabulação cruzada espacial, pelo subprograma Crosstab do IDRISI, que determina as correspondências entre classes de duas imagens. Os resultados estão expressos na tabela seguinte:

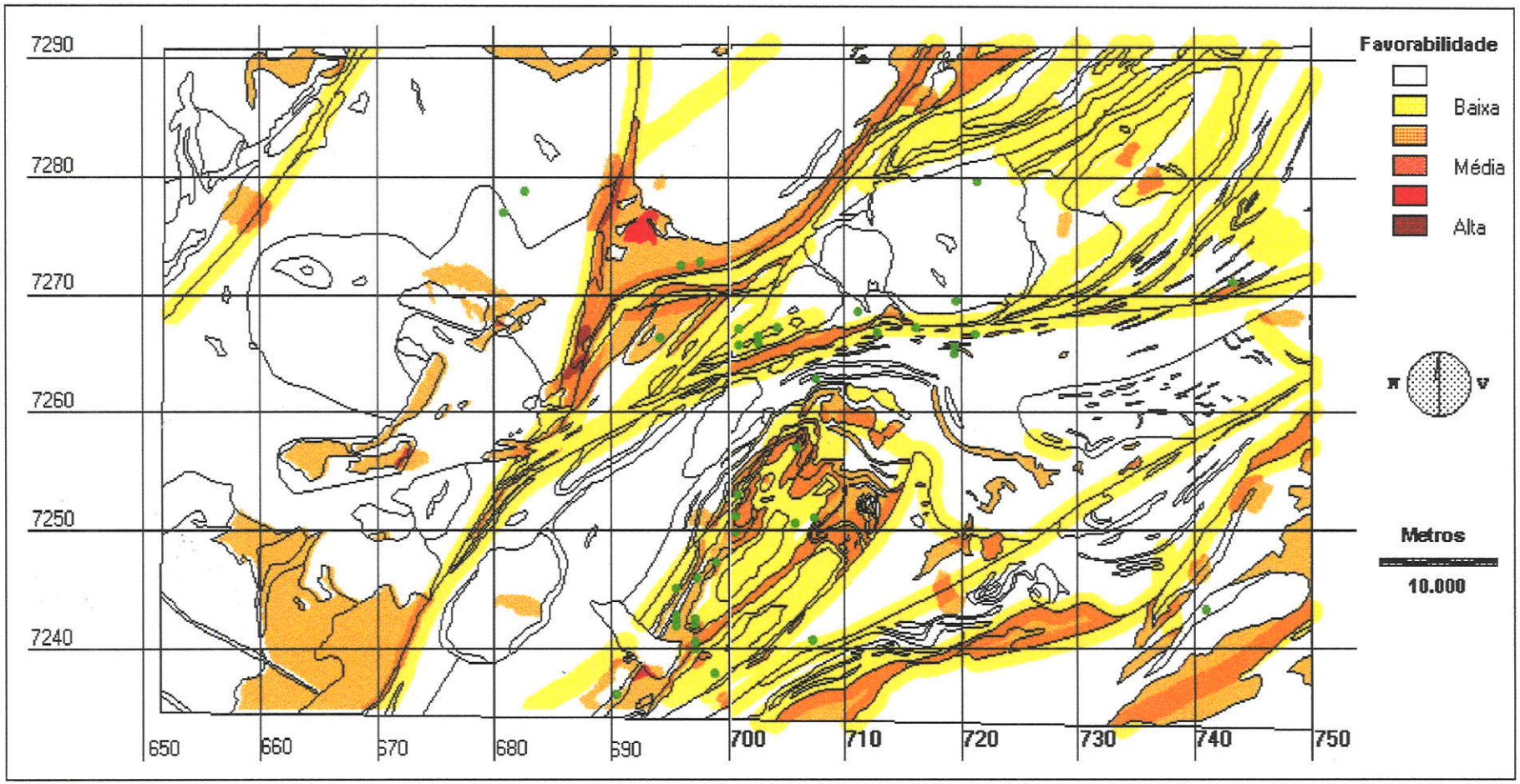


Figura 50: Comparação entre áreas consideradas favoráveis a mineralização tipo Perau e localização de mineralizações, quadrículas Cerro Azul e Apiaí.

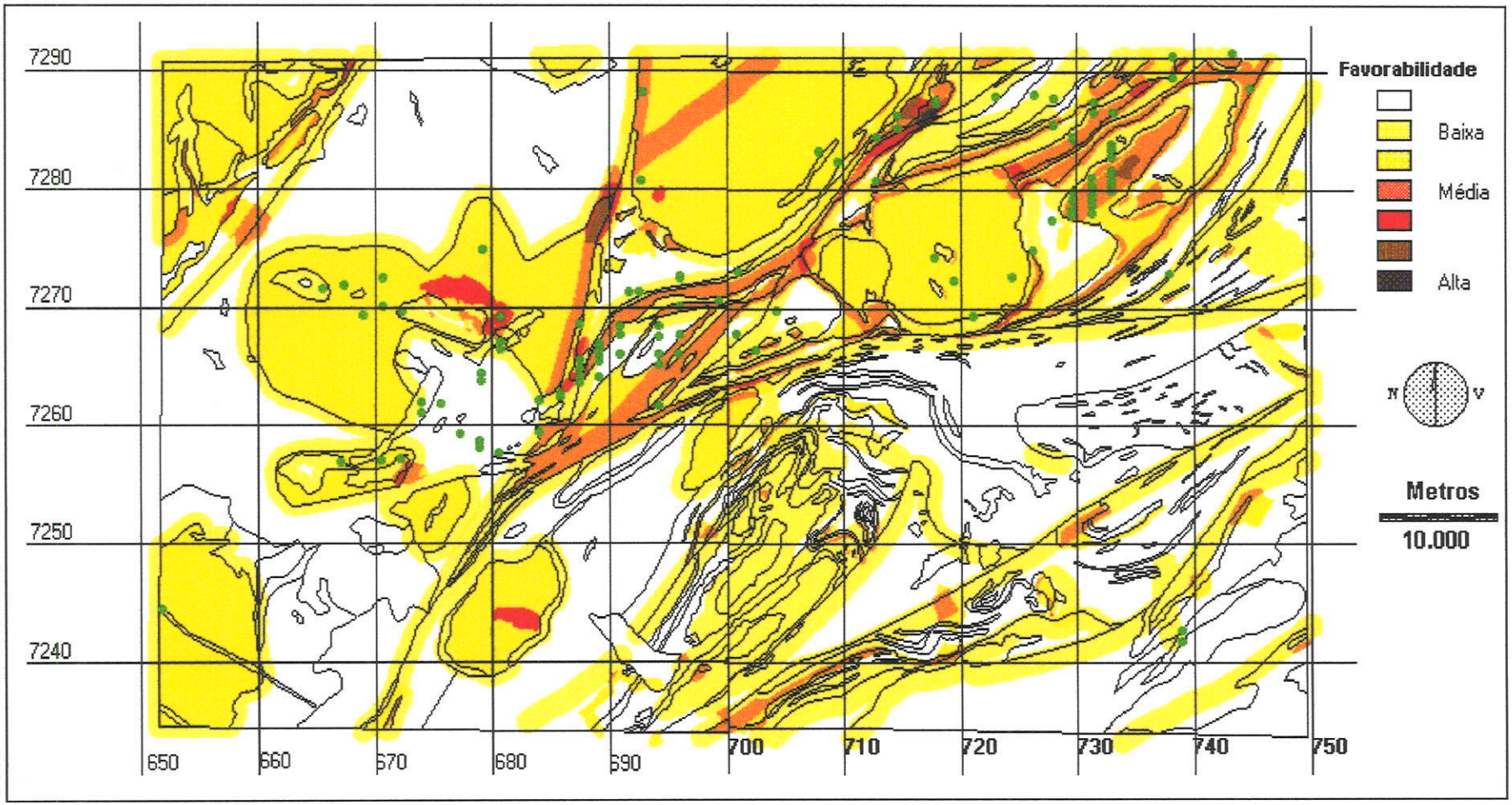


Figura 51: Comparação entre as áreas consideradas como favoráveis a mineralização tipo Rocha e localização das mineralizações, quadrículas Cerro Azul e Apiaí.

Tabela 8: *Correspondência entre ocorrências dos tipos estratiforme e filoniano e as áreas de favorabilidade nas folhas Cerro Azul e Apiaí*

Soma de Pesos de Favorabilidade	Ocorrências Estratiformes	Ocorrências Filonianas
0	11	23
1	16	30
2	7	20
3	6	41
4	0	3
5	0	1
6	-	0
Total de ocorrências	40	118

:

Um quinto das ocorrências filonianas e um quarto das estratiformes cai em áreas consideradas de nenhuma favorabilidade (soma de pontos igual a zero); as outras se distribuem entre as áreas favoráveis. Poucas se colocam nas áreas de maior favorabilidade, provavelmente porque estas sempre correspondem às de anomalias geoquímicas. Conforme já explicado, não houve amostragem de sedimentos de corrente pelo projeto Geoquímica no Vale do Ribeira nas áreas próximas a minas e ocorrências conhecidas.

Uma parte da não conformidade das ocorrências com as áreas favoráveis pode ser atribuída a deslocamento, por falta de precisão da anotação das coordenadas, feita no projeto Mapas Metalogenéticos. Foram utilizados arquivos não corrigidos, também como forma de teste do desempenho dos arquivos existentes. Para os trabalhos de análise metalogenética em andamento os arquivos serão conferidos com os trabalhos originais, em escala maior e, na folha Iporanga, com trabalho de campo.

A partir dos resultados deste item pode-se concluir que a qualidade da predição de áreas favoráveis para a pesquisa de recursos minerais depende:

- da qualidade dos modelos de mineralização empregados. Quanto mais específicos, melhor o desempenho para aquele tipo de modelo de mineralização.
- da qualidade dos dados e da sua conversão à forma digital. Quanto maior a escala original de compilação dos dados, e quanto mais próximos os

documentos das fontes (campo e trabalhos originais), maior a precisão e exatidão.

- dos processos de análise empregados. A combinação booleana é excessivamente rígida; a combinação por pesos somativos é mais flexível; é ainda melhor introduzir probabilidades e possibilidades, aumentando a precisão.

8 CONCLUSÕES FINAIS

Algumas conclusões surgem, no final desta revisão, aplicáveis a todos os capítulos:

a - os processos de tratamento quantitativo de dados permitem ao geólogo uma visão *ao* mesmo tempo muito detalhada e espacialmente abrangente das relações entre os teores de elementos químicos nos meios geológicos e as feições geológicas.

b - os equipamentos e programas atuais permitem a realização de cálculos e a execução de mapas em tempos impensáveis há poucos anos.

c - este novo poder traz ainda mais responsabilidades. É agora mais fácil que nunca aplicar métodos inúteis ou torcer os resultados até que sejam aqueles esperados de antemão.

d - a utilização de métodos de inteligência artificial e de geoprocessamento para a análise metalogenética, embora não consiga superar um geólogo experiente na exatidão dos resultados, permite a execução automática das tarefas enfadonhas e trabalhosas, mas necessárias, facilitando e acelerando a interpretação dos dados.

e - a exibição e edição computadorizada dos resultados das análises facilita a comunicação entre que executam as interpretações geológicas e entre estes e os usuários dos resultados.

Assim sendo, a conclusão final é de que a capacidade de tratamento computadorizado de dados geoquímicos e geológicos é agora uma habilidade imprescindível para os geólogos. A ela deveria ser dedicado um esforço muito maior que o atual, no treinamento dos novos geólogos e na reciclagem dos profissionais formados quando os recursos atuais não eram ainda disponíveis.

9. BIBLIOGRAFIA CITADA

AGTERBERG, F. P., 1974 - **Geomathematics.** Amsterdam, Elsevier, 596 p.

ALGARTE, J.P. et alii, 1972 - **Projeto Sudeste do Estado de São Paulo (escala 1:250.000)** - São Paulo, DNPM/CPRM, 6 v.

ALVES, M.J., CAMARGO, M.V., FERREIRA, J.C., RAABE, G.J. & THEODOROVICZ, A., 1977 - **Mapeamento geológico e levantamento geoquímico da área de cobre e chumbo, Barra do Perau.** Adrianópolis, CEGEP-UFP-Eletro São Marco, 40 p. (rel. inédito).

ARAÚJO, C.C., MACEDO, A.B. & CAMPANHA, G.A.C., 1995 - *Análise de dados geológicos por geoprocessamento visando prospecção regional de recursos minerais - Folha Cerro Azul a 1:100.000 (SG22-X-B-IV).* **VI Simpósio de Quantificação em Geociências. Boletim de Resumos Expandidos**. Rio Claro, UNESP, p. 7-9.

ARAÚJO, C.C., MACEDO, A.B. & CAMPANHA, G.A.C. , 1995 - *Geoprocessamento visando prospecção regional de recursos minerais - Folhas Cerro Azul e Apiaí a 1:100.000 (SG22-X-B-IV e V).* **3° Simpósio de Iniciação Científica da USP.** São Carlos, CNPq-USP, p. 68.

BARBOUR, A.P., BRITO NEVES, B.B. & MEDEIROS, R.A., 1990 - *Algumas implicações tectônicas na gênese das mineralizações sulfetadas do tipo Panelas no vale do Ribeira, SP e PR.* **Rev. Bras. Geociências,** 20(1-4), 46-54.

BARBOUR, A.P. & MACEDO, A.B. , 1991 - *Contribuição ao conhecimento da distribuição espacial dos elementos Cu, Pb, Zn e Ba no distrito do Perau, PR.* **Boletim do Instituto de Geociências - USP.** São Paulo, IG-USP, 22:29-50.

BARBOUR, A.P., MACEDO, A.B. & HYPOLITO, R. , 1988 - *Correlação dos elementos Ag, Pb, Zn, Cu e Fe com Ba em algumas jazidas sulfetadas do vale do Ribeira, estados de SP e Pr.* **Boletim do Instituto de Geociências - USP.** São Paulo, IG-USP, 19:1-21.

BATOLLA JR., F. et alii, 1981 - **Projeto Integração e Detalhe Geológico no Vale do Ribeira.** São Paulo, DNPM/CPRM, 15 v.

BEDELL, R.L , 1994 - **GIS for the Geosciencies - Shortcourse.** Sparks, NV, Homestake Mining Co., 193 p.

BEUS, A.S., 1976 - **Geochemistry of the Lithosphere**. Moscou, Mir 366 p.

BONHAM-CARTER, G.F., 1994 - **Geographic Information Systems for Geoscientists: Modelling with GIS.** Ottawa, Pergamon, 398 p.

BONHAM-CARTER, G.F., AGTERBERG, F.P. & WRIGHT, D.F., 1988 - *Integration of Geological Datasets for Gold Exploration in Nova Scotia.* **Photogrametric Engineering and Remote Sensing,** 54(11):1585-1592.

BONHAM-CARTER, G.F., REDDY, R.K.T. & GALLEY, A.G., 1993 - *Knowledge-driven modeling of volcanogenic massive sulphide potential with a Geographic Information System.* In. KIRKHAM, R.V., SINCLAIR, W.D., THORPE, R.I. & DUKE, J.M. - **Mineral Deposit Modeling.** St. John's, Canada, Geological Association of Canada, p. 735-751.

BOOM, G. Van den, REHDER, S. & KOTTRUP, G., 1980 - *Geochemical patterns in the rocks surrounding the Woodlawn volcanogenic ore deposit, New South Wales, Australia.* **Geologisches Jahrbuch**, D 40:169-200.

BORDIGNON, J.L. & ETCHEBEHERE, M.L.C, 1992 - *Avaliação do potencial metalogenético da porção oriental da folha Itararé (SG.22-X-B), Vale do Ribeira, SP.* **Geociências,** São Paulo, 11(2):41-80.

BRANCO, P.C.M., 1986 - **GEOQUANT.** Rio de Janeiro, CPRM (programa e manual em arquivos de computador, inéditos).

BRANCO, P.C.M., CÂMARA, M.M. & VASCONCELLOS, R.M., 1992 - *Análise espacial de dados aplicada a Prospecção Mineral - Médio Vale do Ribeira - SP.* **37o Congresso Brasileiro de Geologia**. São Paulo, v. 2, p. 218.

BRIDGES, N.J. & McCAMMON, R.B., 1980 - *DISCRIM: a computer program using an interactive approach to dissect a mixture of normal or lognormal distributions.* **Computers & Geosciences,** 6:361-396.

CÂMARA, G. & FREITAS, U.M., 1994 - *Perspectivas em Sistemas de Informação Geográfica.* Palestra apresentada no **V Simpósio de Quantificação em Geociências.** Rio Claro, UNESP, s.p. (mimeo.).

CAMPANHA, G.A.C., 1991 - **Tectônica Proterozóica no Alto e Médio Vale do Ribeira - Estado de São Paulo e Paraná (1:50.000)**, São Paulo, IG-USP, 296 p. (Tese de Doutoramento, inédita).

CAMPANHA, G.A.C, 1995 - *Geologia da Folha Itararé em 1:250.000.* **IV Simpósio de Geologia do Sudeste. Boletim de Resumos,**. Águas de São Pedro, UNESP, p. 111.

CAMPANHA, G.A.C, GIMENEZ, A. & BISTRICHI, C.A., 1996 - **Geologia da folha Itararé (SG.22-X-B).** São Paulo, IPT, s.p. (minuta de relatório, inédita).

CAMPANHA, G.A.C. et alii, 1988 - **Avaliação preliminar da geologia das folhas (em 1:50.000) Taquaral, Mina do Espírito Santo, Ribeirão**

Itacolomi, Serra do Aboboral, Jacupiranga, Guaraú, rio Turvo - Vale do Ribeira, SP. São Paulo, IPT, 334 p. (Relatório 2805-04-7).

CHIODI F°, C., 1984 - *Contribuição à geologia das regiões Sul Paulista e Leste Paranaense. Grupos Açungui e Setuva.* **Anais do XXXIII Congresso Brasileiro de Geologia**, Rio de Janeiro, SBG, 5:2394-2406

DAMASCENO, E.C., 1967 - **Geologia da Mina do Paqueiro.** São Paulo, FFCL-USP, 69 p. (Tese de Doutoramento, inédita).

DAVIS, J.C., 1973 - **Statistics and Data Analysis in Geology.** New York, Wiley, 550 p (1a. edição).

DAVIS, J.C., 1986 - **Statistics and Data Analysis in Geology.** New York, Wiley, 646 p (2a. edição).

de la ROCHE, H., 1968 - *Comportement geochimique différentiel de Na, K et Al dans les formations vulcaniques et sedimentaires: un guide pour l'étude des formations métamorphiques et plutoniques.* **C. R. Acad. Sci. Paris. Série D.** 267:39-42.

DERBY, O.A., 1889 - *Retrospecto histórico dos trabalhos geográficos e geológicos efetuados na Província de São Paulo.* São Paulo, **Comissão Geográfica e Geológica**. 26 p. (Boletim, 1)

DRAPER, N.R. & SMITH, H., 1966 - **Applied Regression Analysis.** New York, Wiley, 407 p.

EASTMAN, J.R. 1992*a* - **IDRISI - Technical Reference.** Worcester, MA, Clark Un., 213 p.

EASTMAN, J.R., 1992*b* - **IDRISI - User's Guide.** Worcester, MA, Clark Un., 178 p.

EASTMAN, J.R., 1995 - **IDRISI for Windows - User's Guide, v. 1.0.** Worcester, MA, Clark Un., 220 p.

EASTMAN, J.R., JIN, W., KYEM, P.A.K. & TOLEDANO, J., 1995 - *Raster procedures for Multi-Criteria/Multi-Objective Decisions.* **Photogrammetric Engineering & Remote Sensing,** 61(5):539-547.

EASTMAN, J.R., KYEM, P.A.K., TOLEDANO, J. & JIN, W., 1993 - **GIS and Decision Making.** Geneva, UNITAR, 110 p.

ELLIOTT, J. & STOESER, D., 1994 - **Application of Geographic Information Systems in Mineral Deposit Modelling and Resource Assessment.** Rio de Janeiro, CPRM\USGS, s.p. (notas de curso, inéditas).

ESRI, 1990 - **ARC/INFO.** Redlands, CA, Environmental Systems Research Institute. Programa v.4.

FERREIRA, J.C.G, CHIODI, C., MACEDO, A.B. & BATOLLA, F., 1981 - **Projeto Integração e Detalhe Geológico no Vale do Ribeira. Relatório Final - Área Mina do Paqueiro.** São Paulo, DNPM/CPRM, v. IX, 117 p.

FLETCHER, W.K., 1981 - **Analytical Methods in Geochemical Prospecting.** Amsterdam, Elsevier, 255 p.

GAÁL, G. & MERRIAN, D.F., 1990 - **Computer Applications in Resource Estimation.** Oxford, Pergamon, 455 p.

GARRETT, R.G., 1969 - *The determination of sampling and analytical errors in exploration geochemistry.* **Economic Geology**, 64(5):568-569.

GARRETT, R.G., 1973 - *The determination of sampling and analytical errors in exploration geochemistry - a reply.* **Economic Geology**, 68(2):282-283.

GOLDEN SOFTWARE, 1994 - **SURFER, v. 5.1.** Golden, CO, Golden Software, programa e manual em arquivo de computador.

GOVETT, G.J.S, 1983 - **Rock Geochemistry in Mineral Exploration.** Amsterdam, Elsevier, 461 p.

HAMA, M. & ALGARTE, J.P., 1986 - **Projeto Mapas Metalogenéticos e de Previsão de Recursos Minerais 1:250.000) - Folha Itararé - SG.22-X-B.** São Paulo, CPRM, 2 v.

HARRIS, D. P., 1984 - **Mineral Resources Appraisal.** Oxford, Clarendon, 445 p.

HOWARTH, R.J., 1983 - **Statistics and data analysís in geochemical prospecting.** Amsterdam. Elsevier. 437 p. (Handbook of Exploration Geochemistry, v.2).

JICA-MMAJ, 1984 - **Report on geological survey of Anta Gorda - Brazil.** s.l., JICA, 63 p.

JONES, J.R., 1994 - **Tosca - Reference Guide, version 2.1.** Worcester, MA, Clark University, 86 p.

KATZ, S.S., 1991 - *Emulating the Prospector expert system with a raster GIS.* **Computers & Geosciences,** 17(7):1033-1050.

KOCH, G.S., 1987 - **Exploration-Geochemical Data Analysis with the IBM-PC.** New York, Van Nostrand Reinhold, 179 p.

KRUMBEIN, W.C. & GRAYBILL, W.A., 1965 - **An introduction to statistical models in Geology.** New York, Mc.Graw-Hill, 475 p.

LANDIM, P.M.B., 1979 - **Análise estatística de dados geológicos.** - Rio Claro, CEGEO-UNESP, 102 p.

LANDIM, P.M.B., 1993 - **Análise estatística de dados geológicos**. - Rio Claro, CEGEO-UNESP, 116 p. (Laboratório de Geomatemática - Publicação didática nº 4).

LANDIM, P.M.B., 1995 - *Modelos Multivariantes em Geociências.* **Sessão Regular da Academia Brasileira de Ciências,** São Paulo, IG-USP, 1/12/1995 (inédito).

LEPELTIER, C., 1969 - *A simplified statistical treatment of geochemical data by graphical representation.* **Economic Geology,** 64:538-550.

LETERRIER, J., & de La ROCHE, H., 1972 - *Extension aux basalts d'une tipologie chimique des roches ignées acides et intermediaires.* **C. R. Acad. Sci., Paris,** 274:1772-1775.

LIOTTE, S.V., MACEDO, A.B., SOUZA, A.P., SALLES, F.A.F. & BRUGNERA, G.R., 1995 - *Geoprocessamento como auxílio à administração municipal: Bom Sucesso de Itararé.* **VI Simpósio de Quantificação em Geociências. Boletim de Resumos Expandidos.** Rio Claro, UNESP, p. 97-99.

MACEDO, A.B., 1979 - **Estudo estatístico das relações entre elementos principais e traços em rochas basálticas da bacia do Paraná.** São Paulo, IG-USP, 171 p. (dissertação de mestrado, inédita).

MACEDO, A.B., 1985 - *Tratamento de dados geoquímicos*, in Macedo, A.B. - **Exame de Qualificação**, São Paulo, IG-USP, p. 103-134 (inédito).

MACEDO, A.B., 1986 - **Prospecção litogeoquímica na mina do Perau, Paraná.** São Paulo, IGc-USP, 2 v. (Tese de doutoramento, inédita).

MACEDO, A.B., 1989 - *Lithogeochemical exploration of massive sulfides at Perau base metal mine, Paraná, Brasil.* In: INTERNATIONAL GEOCHEMICAL EXPLORATION SYMPOSIUM, 13. BRASILIAN GEOCHEMICAL CONGRESS, 2. Rio de Janeiro. **Abstracts**. Rio de Janeiro, Sociedade Brasileira de Geoquímica, CPRM/DNPM, p. 108.

MACEDO, A.B., 1992 - *Mining pollution in the Ribeira river valley, Southern Brazil.* **Abstracts for the Conference on Hazardous Waste Research.** Boulder, Colorado, p. 19-20.

MACEDO, A.B., 1993*a* - *Poluição por mineração na bacia do Ribeirão Grande, vale do Ribeira, Paraná.* **V Simpósio Sul-Brasileiro de Geologia - Boletim de Resumos e Programa.** Curitiba, SBG, p. 70-71.

MACEDO, A.B., 1993*b* - *Avaliação regional de recursos minerais por Geoprocessamento - Fl. Itararé (SG22-X-B-1:250.000): Problemas conceituais e práticos.* **V Simpósio Sul-Brasileiro de Geologia - Boletim de Resumos e Programa**. Curitiba, SBG, p. 84-85.

MACEDO, A.B., 1994 - *Avaliação do programa IDRISI para trabalhos de Geologia e Mineração.* **V Simpósio de Quantificação em Geociências - Boletim de Resumos Expandidos.** Rio Claro, 58-60.

MACEDO, A.B., ADDAS, W. & BATOLLA Jr, F. , 1981 - *Metodologia de prospecção geoquímica em minas do vale do Ribeira, PR.* **Atas do 3. Simpósio Regional de Geologia**, Curitiba, SBG, 1:71-81.

MACEDO, A.B., ARAÚJO, C.C. & CAMPANHA, G.A.C., 1995 - *Prospecção regional de sulfetos exalativos e hidrotermais no vale do Ribeira por Geoprocessamento.* **Sessão Regular da Academia Brasileira de Ciências,** São Paulo, IG-USP (trabalho apresentado, aguardando publicação nos Anais da Academia Brasileira de Ciências).

MACEDO, A.B. & BATOLLA, F., 1981 - **Projeto Integração e Detalhe Geológico no Vale do Ribeira. Relatório Final - Área Mina do Paqueiro.** São Paulo, DNPM-CPRM, v. XIII, 90 p.

MACEDO, A.B., BATOLLA JR., F. & ADDAS, W., 1982 - *Prospecção Geoquímica e Geofísica de Detalhe em Minas do Vale do Ribeira.* **XXXII Congresso Brasileiro de Geologia, Salvador, SBG, Resumos, Boletim n. 2**, p. 36.

MACEDO, A.B., BRAGHIN, M.A. & MOREIRA, F.R.S., 1993*a* - *Teste de métodos de interpretação geoquímica por computador na mina do Paqueiro, Adrianópolis, Paraná.* **I Workshop de Informática do IG-USP, Boletim IG-USP, Publ. Especial,** São Paulo, IG-USP, 16:23-24.

MACEDO, A.B., CAMPANHA, G.A.C., BRAGHIN, M.A. & ARAÚJO, C.C., 1993*b* - *Produção de mapas digitalizados e banco de dados para geoprocessamento. Fl. Itararé (SG22-X-B - 1:250.000).* I Workshop de Informática do IG-USP, **Boletim IG-USP, Publ. Especial**, São Paulo, IG-USP, 16:27-28.

MACEDO, A.B. & RÜEGG, N..R., 1974*a* - *A variação de elementos traços nas rochas basálticas da bacia do Paraná estudada por meio da análise fatorial.* **Anais do XXVIII Congresso Brasileiro de Geologia.** Porto Alegre, p.49-56

MACEDO, A.B. & RÜEGG, N.R., 1974*b* - *Aplicação da análise fatorial ao estudo de elementos principais nas rochas basálticas da Bacia do Paraná.* **Anais do XXVIII Congresso Brasileiro de Geologia.** Porto Alegre, p.143-151.

MACEDO, A.B., SALLES, F.A.F., LIOTTE, S.V. & GUIMARÃES, M, 1993*c* - *Geoprocessamento aplicado ao estudo do impacto ambiental da mineração e planejamento de recuperação na mina do Perau, Adrianópolis, PR.* **I Workshop de Informática do IG-USP, Boletim IG-USP, Publ. Especial,** São Paulo, IG-USP, 16:29-30.

MANSON, V., 1967 - *Geochemistry of basaltic rocks: major elements.* **In** HESS, H.H. - **Basalts, the Poldervaart treatise on rocks of basaltic composition.** New York, Interscience, p. 215-269.

McCAMMON, R.B., 1987 - **The muPROSPECTOR Mineral Consultant System.** Denver, USGS, 18 p. (USGS Bull., 1697).

McCAMMON, R.B., 1993 - *Prospector II - An expert system for mineral deposit models.* In. KIRKHAM, R.V., SINCLAIR, W.D., THORPE, R.I. & DUKE, J.M. - **Mineral Deposit Modeling.** St. John's, Canada, Geological Association of Canada, p. 679-684.

McCAMMON, R.B., FINCH, W.I., KORK, J. & BRIDGES, N.J., 1994 - *An integrated data-directed numerical method for estimating the undiscovered mineral endowment in a region.* **Nonrenewable Resources,** 3(1):109-122.

MIESCH, A.T., 1981 - *Estimation of the geochemical threshold and its statistical significance.* **Journal of Geochemical Exploration,** Amsterdam, 16:49-76.

MORGENTAL, A. et alii, 1978 - **Projeto Geoquímica no Vale do Ribeira (escala 1:100.000)**, São Paulo, DNPM/CPRM, 8 v.

PAN, G., HARRIS, D.P. & HEINER, T., 1992 - *Fundamental issues in quantitative estimation of mineral resources.* **Nonrenewable Resources,** 1(4):281-292.

NIE, N., BENT, D. & HULL, C., 1970 - **SPSS - Statistical Package for the Social Sciences.** New York, McGraw-Hill, 343 p.

PERROTTA, M.M. & CÂMARA, M.M., 1994 - *Processamento digital de imagens de satélite e aerogamaespectométricas na região entre Guapiara-Iporanga-Eldorado Paulista-Serra dos Agudos Grandes, Médio Vale do Ribeira, SP.* **38° Congresso Brasileiro de Geologia.** Camboriú - SC, vol. IV, p. 157.

PERROTTA, M.M., 1995 - *Modelamento dos Depósitos Auríferos do Vale do Ribeira.* **IV Simpósio de Geologia do Sudeste. Boletim de Resumos.** Águas de São Pedro, UNESP, p. 103.

PRINZ, M, 1967 - *Geochemistry of basaltic rocks: trace elements.* **In** HESS, H.H. - **Basalts, the Poldervaart treatise on rocks of basaltic composition.** New York, Interscience, p. 271-323.

REDDY, R.K.T., BONHAM-CARTER, G. & GALLEY, A.G., 1992 - *Developing a Geographic Expert System for regional mapping of volcanogenic massive sulfide (VMS) deposit potential.* **Nonrenewable Resources,** 1(2)112-124.

RODRIGUES, M.,1990 - *Geoprocessamento* - **Simpósio Brasileiro de Geoprocessamento**, São Paulo, EPUSP, p. 01 - 27.

RONDINELLI, D., et alii, 1986 - Avaliação preliminar da geoquímica regional das folhas (1:50.000) ... , São Paulo, Pró-Minerio/IPT, 3 v. (Relatório IPT 28.631).

ROSE, A.W., HAWKES, H.E. & WEBB, J.S., 1979 - **Geochemistry in Mineral Exploration.** London, Academic, 657 p.

ROUTHIER, P., 1980 - **Où sont les métaux pour l'avenir ? Les provinces metaliques - Essai de métallogénie globale.** Orléans, BRGM, 407 p. (Mémoire du BRGM nº 105).

RÜEGG, N.R., 1970 - *A composição química das rochas basálticas da bacia do Paraná (América do Sul) e de outras províncias gonduânicas equivalentes.* **Notas e Memórias, Museu Min. Geol. Univ. Coimbra,** 70:26-85.

RÜEGG, N.R., 1975*a* - **Modelos de variação química na província basáltica do Brasil Meridional.** São Paulo, Inst. Geociências, 2 v., 309 p. (Tese de Livre-Docência, inédita).

RÜEGG, N.R., 1975*b* - *Natureza serial do magmatismo basáltico da Bacia do Paraná.* **Naturalia,** S. José do Rio Preto, FFCL, 1:141-152.

SALLES, F.A.F., MACEDO, A.B. & LIOTTE, 1995 - *Aplicação de métodos estatísticos e de geoprocessamento para interpretação geoquímica na mina do Paqueiro (Adrianópolis, PR).* **VI Simpósio de Quantificação em Geociências. Boletim de Resumos Expandidos.** Rio Claro, UNESP, p. p. 151-152.

SHCHEGLOV, A.D., 1979 - **Fundamentals of Metallogenic Analysis.** Moscou, Mir, 335 p.

SINCLAIR, A.J., 1974 - *Selection of thresholds in geochemical data using probability graphs.* **Journal of Geochemical Exploration,** 3(2):129-140.

SINCLAIR, A.J., 1976 - **Applications of Probability Graphs in Mineral Exploration.** Richmond, BC, Canada, Assoc. Expl. Geochemists, 95 p. (Spec. Vol. 4).

SILVA, C.R., CHIODI Fº, C., MACEDO, A.B., TAKAHASHI & BATOLLA, Fº., 1981 - **Projeto Integração e Detalhe Geológico do Vale do Ribeira. Relatório Final Área Mina do Perau.** São Paulo, DNPM/CPRM v.11, 162 p.

SINGER, D.A. & MOSIER, D.L., 1981 - *A review of regional mineral resource assessment methods.* **Economic Geology,** 76:1006-1015.

SOLLER, D.R., STETTNER, W.R., LANFEAR, K.J. & AITKEN, D.S., 1990 - **A user's manual for a method of map scanning and digital editing for**

thematic map production and data-base construction. Denver, USGS, 38 p. (USGS Circular 1054).

STANLEY, C.R., 1987 - **Instruction manual for PROBPLOT.** Ontario, Ass. Expl. Geochemists, 40 p. (AEG Special Volume # 14).

STANTON, R.E. , 1966 - **Rapid Methods of Trace Analysis for Geochemical Exploration.** London, Edward Arnold, 96 p.

SUSLICK, S.B., 1996 - **Quantificação da potencialidade mineral com base em dados geoquímicos: um estudo aplicado às folhas Itaquaquecetuba, Piracaia, Igaratá, Camanducaia e Monteiro Lobato.** São Paulo, IG-USP, 304 p. (Tese de doutoramento, inédita).

SUSLICK, S.B., QUNTANILHA, J.A. & LEWIS, Jr., R.W., 1984 - *Métodos exploratórios de análise de dados geoquímicos.* **Anais do XXXIII Congresso Brasileiro de Geologia,** Rio de Janeiro, SBG, p. 4864-4876.

TSOFT Co., 1994 - **TMAPPER! Interactive Mapping Software.** Worcester, MA, Tsoft Co, 74 p.

VISTELIUS, A.B., 1960 - *The skew frequency distributions and the fundamental law of the geochemical processes.* **Jour. Geol.,** 68:1-22.

WILLIAMS, X.K., 1967 - *Statistics in the interpretation of geochemical data.* **New Zealand J. Geol. Geophys.,** 10:771-797.

YAMAMOTO, J.K., 1986 - **Representações gráficas espaciais em Geologia: aplicações no Complexo Alcalino de Anitápolis, SC.** São Paulo, IG-USP, 167 p. (Dissertação de mestrado, inédita).

YAMAMOTO, J.K., 1988*a* - *Representações gráficas espaciais em Geociências auxiliadas por computador.* **Rev. Bras. Geociências,** São Paulo, SBG, 18(1):3.26.

YAMAMOTO, J.K., 1988*b* - *Representações gráficas de modelos digitais de terreno.* ***Anais do 21º Cong. Bras.* Informática,** Rio de Janeiro, p. 924-929.

YAMAMOTO, J.K., 1991 - **Comparação de métodos computacionais para avaliação de reservas: um estudo de caso na Jazida de cobre de Chapada, GO.** São Paulo, IG-USP, 175 p. (Tese de doutoramento, inédita).

YAMAMOTO, J.K., 1995 - **Desenvolvimento de métodos de interpolação para avaliação de jazidas.** São Paulo, IG-USP, 117 p. (Sistematização crítica para concurso de Livre-Docência, inédita).