



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

**MODELAGEM DA DISTRIBUIÇÃO DA CONCENTRAÇÃO DE CLOROFILA
INTEGRADA NA COLUNA D'ÁGUA COM TÉCNICAS DE
GEOPROCESSAMENTO**

Natália de Moraes Rudorff

Trabalho final do curso de Introdução ao Geoprocessamento

INPE
São José dos Campos
2009

ÍNDICE

1 INTRODUÇÃO	6
1.1 Importância do estudo	6
2 Problema e hipótese.....	8
3 REVISÃO BIBLIOGRÁFICA	9
3.1 Distribuição da biomassa fitoplanctônica no oceano	9
3.2 Aplicações dos modelos ecológicos	9
3.3 Geoprocessamento em estudos ambientais.....	10
3.4 Modelagem de dados em Sistema de Informação Geográfica	11
4 MATERIAIS E MÉTODOS	15
4.1 Área de estudo	15
4.2 Parâmetros oceanográficos	16
4.3 Modelagem dos dados no Sistema de Informação Geográfica.....	16
4.4 Definição do Projeto e importação dos dados para o SPRING	19
4.5 Processamentos.....	20
4.6 Análise exploratória e geração das grades retangulares	20
4.7 Determinação das variáveis para o modelo de estimativa.....	20
4.8 Modelos Fuzzy	21
4.9 Mapas finais.....	23
5 RESULTADOS E DISCUSSÃO	24
5.1 Análise exploratória e geração das grades retangulares	24
5.2 Interpolação	26
5.3 Análise espacial das variáveis interpoladas.....	28
5.4 Modelo de Regressão Espacializado	30
5.5 Modelos Fuzzy	30
5.6 Mapas finais e comparação dos métodos	37
6 CONCLUSÕES.....	43
7 RECOMENDAÇÕES.....	44
8 REFERÊNCIAS BIBLIOGRÁFICAS	45
9 ANEXOS	46

ÍNDICE DE FIGURAS

Figura 1: Área de estudo com os pontos de coleta do projeto DEPROAS dos cruzeiros do verão de 2001 e 2002 na plataforma e talude da costa sudeste brasileira.	15
Figura 2: Modelo OMT-G do banco de dados para a modelagem da estimativa da concentração de clorofila integrada na coluna d'água.....	18
Figura 3: Histogramas e gráficos de curva de distribuição das variáveis.	24
Figura 4: Mapas de distribuição da clorofila integrada na coluna d'água com diferentes métodos de interpolação.	27
Figura 5: Mapas de distribuição das variáveis interpoladas com o método de média ponderada com imagem em níveis de cinza. O preto corresponde aos menores valores o branco aos maiores.....	28
Figura 6: Mapa de distribuição da clorofila integrada estimada pelo modelo de regressão linear múltiplo de Leão et al (2007).	30
Figura 7: A) Gráfico de distribuição da frequência de Brunt-Vaissala (N2) versus a concentração de clorofila integrada; B) Gráfico de distribuição de N2 em ordem crescente; C) Função fuzzy de N2; D) Mapa da distribuição de N2_fuzzy.	31
Figura 8: A) Gráfico de distribuição do inverso da concentração de nitrato integrado (INeu) versus a concentração de clorofila integrada; B) Gráfico de distribuição de INeu em ordem crescente; C) Função fuzzy de INeu; D) Mapa da distribuição de INeu_fuzzy.	32
Figura 9: A) Gráfico de distribuição da temperatura superficial do mar (TSM) versus a concentração de clorofila integrada; B) Gráfico de distribuição de TSM em ordem crescente; C) Função fuzzy de TSM; D) Mapa da distribuição de TSM_fuzzy.....	33
Figura 10: A) Gráfico de distribuição da profundidade da termoclina (zt) versus a concentração de clorofila integrada; B) Gráfico de distribuição de zt em ordem crescente; C) Função fuzzy de zt; D) Mapa da distribuição de zt_fuzzy.	33
Figura 11: A) Gráfico de distribuição da profundidade da concentração máxima de clorofila (zmcs) versus a concentração de clorofila integrada; B) Gráfico de distribuição de zmcs em ordem crescente; C) Função fuzzy de zmcs; D) Mapa da distribuição de zmcs_fuzzy.	34
Figura 12: Critérios para calculo de pesos no AHP_Fuzzy_Modelo1.	35
Figura 13: Mapa de distribuição do MNT do modelo AHP_Fuzzy_1 de estimativa da concentração de clorofila integrada.....	35
Figura 14: Critérios para calculo de pesos no AHP_Fuzzy_Modelo 2.....	36
Figura 15: Mapa de distribuição do MNT do modelo AHP_Fuzzy_2 de estimativa da concentração de clorofila integrada.....	37
Figura 16: Mapa da distribuição da concentração da clorofila integrada na coluna d'água com interpolação por média ponderada.....	39
Figura 17: Mapa da distribuição da concentração da clorofila integrada estimada pelo modelo de regressão linear múltiplo proposto por Leão et al (2007).....	40
Figura 18: Mapa da distribuição da concentração da clorofila integrada estimada pelo modelo AHP_Fuzzy_1.	41
Figura 19: Mapa da distribuição da concentração da clorofila integrada estimada pelo modelo AHP_Fuzzy_2.	41

RESUMO

O presente trabalho teve como questão central propor modelos de estimativa da distribuição da concentração de clorofila integrada na coluna d'água utilizando técnicas de interpolação, lógica fuzzy e AHP (Processo Analítico Hierárquico), em ambiente SIG. Os modelos testados tiveram como parâmetros de entrada a temperatura superficial da água (TSM), a profundidade da termoclina, o nitrato integrado na zona eufótica, a frequência de Brünt-Väisälä e a profundidade da concentração máxima de clorofila. Estas variáveis foram determinadas com base num estudo prévio de inferência da clorofila integrada por meio de um modelo de regressão múltipla, realizado por Leão et al, (2007), e análises exploratórias realizadas neste trabalho. A área de estudo para a aplicação do modelo foi na costa sudeste do Brasil, numa área que abrange a plataforma e talude, entre o Cabo de São Tomé (22°S) e a Ilha de São Sebastião (23°50'S). Os dados usados para gerar e validar os modelos foram obtidos por meio da base de dados da tese do Dr. Milton Kampel (KAMPEL, 2003). Esta tese foi parte integrante do projeto de *Dinâmica do Ecossistema de Plataforma da Região Oeste do Atlântico Sul* (Deproas), realizado em parceria dos institutos IO/USP e INPE. Os dados foram gerados por meio de análises de coletas *in situ* realizados em 2 cruzeiros oceanográficos durante o período de verão dos anos 2001 e 2002, cobrindo perfis longitudinais na plataforma e talude. A concentração da clorofila integrada na coluna d'água é um índice da biomassa fitoplanctônica distribuída ao longo da zona eufótica. O perfil vertical desta distribuição e a concentração total da clorofila integrada variam de acordo com os parâmetros de estratificação da coluna d'água, sendo, portanto, possível a sua estimação através da modelagem desses parâmetros. Os modelos baseados em lógica fuzzy são mais flexíveis permitindo uma modelagem mais representativa do gradiente de transição dos fenômenos naturais. O uso do SIG para a integração, processamento e representação dos dados, permite uma modelagem que considera a distribuição espacial das variáveis através de métodos de interpolação com inferência determinística e geoestatística. O método AHP para definição dos pesos, das variáveis de entrada no modelo, permite manipular os critérios combinando conhecimentos empíricos e teóricos do fenômeno estudado. Como resultado final foi possível aplicar estas técnicas e obter um modelo de estimativa da concentração e distribuição da clorofila integrada, com resultados razoavelmente próximos àqueles obtidos pelos dados coletados *in situ*. Apesar do modelo não ter aferido tão bem os valores, o mesmo delineou bem feições de processos ocorrentes na área durante o período estudado, como os fenômenos de ressurgência no litoral do Rio de Janeiro. O modelo mostrou-se simples e de fácil aplicação para estudos oceanográficos de distribuição da biomassa fitoplanctônica e análise das interações de variáveis hidro-físico-químicas. As ferramentas de geoprocessamento apresentaram grande potencial para o desenvolvimento deste tema de pesquisa, com o aprimoramento dos modelos propostos e a elaboração de novos, testando as diferentes técnicas disponíveis nos programas de SIG.

ABSTRACT

The present work aimed to offer models for the estimation of the chlorophyll concentration integrated in the water column, using techniques of interpolation, fuzzy logic and AHP (Analytical Hierarchical Process), on GIS environment. The proposed model had as input parameters the superficial temperature of water, thermocline depth, nitrate integrated in the euphotic zone, frequency of Brunt-Vaisala and depth of maximum concentration of chlorophyll. The study area chosen for the application of the model is situated on the southwest coast of Brazil, at an area embracing the continental platform between Cabo de São Tomé (22°S) and the São Sebastião Island (23°50'S). Data used to generate and validate the models were obtained by the data base of Milton Kampel's thesis (KAMPEL, 2003). This thesis was a part integrated of the Dynamic of the Ecosystem of Platform of the West Region of South Atlantic ("*Deproas*") project, with a partnership between IO/USP and INPE institutes. The data was generated from *in situ* analyses collected on 2 oceanographic expeditions along longitudinal profiles over the platform and talude of the study area, during 2001 and 2002. The chlorophyll concentration integrated in the water column is the total biomass distributed along the euphotic zone. The chlorophyll's vertical profile and total concentration vary according to the stratification of water parameters, thus it is possible to be estimated by modeling these parameters. Models based on fuzzy logic rules are more flexible than traditional ones and consider the smooth gradient of transitions, typical of natural phenomena. The use of GIS for integrating, processing and representing data, also allows a modeling which considers space variation of the parameters with methods of spatial analyses and geostatistical inference. The AHP method for the definition of weights also allows user to manipulated criteria combining empirical and theoretical knowledge of the phenomenon studied. As final result it was possible to apply these techniques and obtain a model of the estimation of concentration and distribution of integrated chlorophyll, with results reasonably close to the *in situ* data distribution. Although values of model weren't so close to *in situ* data, the features of main process were delineated, such as resurgence currents at the coast of Rio de Janeiro during the studied period. The model proposed showed to be simple of easy application for oceanographic studies of phytoplankton biomass distribution and its interactions with hydro-physical-chemical parameters. Geoprocessing techniques demonstrated great potential for application of the studied theme, with the improvement of models proposed and elaboration of new ones, using different techniques available in GIS programs.

1 INTRODUÇÃO

1.1 Importância do estudo

A concentração de clorofila_a é uma medida usada como índice da biomassa fitoplanctônica presente na coluna d'água. Esta medida é de extrema importância para os estudos de produtividade marinha e ciclos biogeoquímicos do carbono em escalas locais e globais (KAMPEL, 2003). A distribuição vertical da biomassa fitoplanctônica tem um perfil em formato de uma *curva gaussiana* ao longo na zona eufótica, a qual é determinada por fatores físico-químicos, como a atenuação e transmissão da irradiação solar ao longo da coluna (coeficiente vertical de atenuação difusa), a profundidade da zona eufótica, a distribuição dos nutrientes inorgânicos dissolvidos, o perfil vertical de temperatura (termoclina), etc. (PLATT e JASBY, 1978; KRIEGER et al, 2005). A concentração total da clorofila_a ao longo deste perfil vertical é o valor da biomassa integrada na coluna d'água. Quanto à sua distribuição horizontal, esta vai depender também de outros fatores como as correntes oceânicas, padrões de vento, proximidade de fontes de nutrientes, dentre outros (PEREIRA et al., 2004; KRIEGER et al, 2005).

Como a distribuição da clorofila está, portanto, estreitamente ligada a estes fatores hidro-físico-químicos, que por sua vez podem ser utilizados como parâmetros de entrada em modelos de estimativa da biomassa fitoplanctônica. Os modelos ecológicos têm auxiliado muito nos estudos de compreensão dos fenômenos naturais e nas simulações de cenários futuros (CÂMARA et al., 2001). Os modelos para estimativa de concentração de clorofila podem têm uma série de aplicações desde o monitoramento de áreas de maior produtividade para rastreamento de cardumes, monitoramento e previsão de *blooms* fitoplanctônicos, como a maré vermelha e diversos estudos da dinâmica costeira e oceânica (PEREIRA et al., 2004). Contudo, os modelos são sempre uma abstração da realidade (CÂMARA et al., 2001). Modelos simples, com poucas variáveis não conseguem englobar e representar toda a complexidade do ambiente natural, enquanto que os modelos complexos são mais representativos, mas de difícil aplicação para uso rotineiro.

Os métodos de inferência por lógica difusa têm sido cada vez mais usados nos mais diversos campos de conhecimento, vindo como uma alternativa de tornar modelos simples e robustos de forma mais flexíveis, considerando um padrão de distribuição mais de acordo com o gradiente natural (PEREIRA et al., 2004). Neste método os valores discretos são convertidos numa escala de 0 a 1 gerando intervalos de incerteza que suavizam os gradientes transicionais característicos dos fenômenos naturais. Assim a aplicação de modelos para estimativa da clorofila_a integrada na coluna d'água, utilizando métodos de lógica difusa (*fuzzy logic*) pode oferecer aproximações mais condizentes com o perfil real da distribuição da biomassa fitoplanctônica. Este método já foi aplicado à esta estimativa em outros trabalhos como (PEREIRA et al., 2004), mas que não exploraram a aplicação da distribuição espacial (horizontal).

Como os parâmetros de entrada no modelo também têm diferentes graus de influência na determinação da variável, cada qual pode ter um valor ponderado, estimado por pesos segundo os respectivos graus de importância. Estes pesos podem ser determinados subjetivamente segundo o conhecimento teórico, ou empiricamente através de relações geoestatísticas e de estatística multivariada. O método AHP para definição dos pesos, das variáveis de entrada no modelo, permite manipular os critérios combinando conhecimentos empíricos e teóricos do fenômeno estudado (FUCKS et al., 2004; GEMITZIA et al., 2007).

No ambiente SIG é possível realizar uma série de procedimentos desde a análise exploratória dos dados, a espacialização por métodos de interpolação, com inferência nas áreas não amostradas, e a elaboração de modelos de estimativa para estudo dos processos e simulações de cenários (CÂMARA et al., 2001). Assim constitui-se de uma excelente ferramenta de estudo e integração de dados distribuídos do espaço. Desta forma pode-se extrair informações dos processos de uma forma mais interativa e dinâmica do que nos procedimentos convencionais de uso de mapas cartográficos e análises não espaciais (CÂMARA et al., 2001).

2 Problema

As perguntas centrais que nortearam este trabalho foram:

- É possível elaborar modelos simples de estimativa da distribuição da concentração de clorofila integrada na coluna d'água, por meio de técnicas de geoprocessamento em ambiente SIG?
- Quais as melhores técnicas geoespaciais para elaboração dos modelos no caso particular estudado?
- Os resultados retratam bem os processos naturais e os valores medidos *in situ* no ambiente de estudo?

2.1.1 Hipóteses

Para responder a estas perguntas as hipóteses foram:

- Com o uso de variáveis determinantes na distribuição vertical e horizontal da clorofila fitoplanctônica é possível elaborar modelos simples com estimativas razoáveis;
- Os métodos de espacialização permitem uma melhor análise dos processos e inferências considerando a localização das estações;
- A lógica difusa pode gerar resultados melhores que métodos tradicionais por lógica booleana;
- O método de análise multi-critério (AHP) pode modelar modelos mais flexíveis combinando conhecimentos empíricos e teóricos do pesquisador.

3 REVISÃO BIBLIOGRÁFICA

3.1 Distribuição da biomassa fitoplanctônica no oceano

A concentração de clorofila_a é uma medida usada para determinar o índice de biomassa fitoplanctônica presente na coluna d'água. A distribuição vertical da biomassa fitoplanctônica tem um perfil em formato de uma *curva gaussiana* ao longo na zona eufótica. A zona eufótica compreende toda a coluna d'água até a profundidade máxima de penetração da luz. No perfil vertical da clorofila a concentração máxima não ocorre na primeira camada da superfície, mas numa profundidade sub-superficial por isso o formato de *curva gaussiana* (PLATT e JASBY, 1976).

Diversos fatores hidro-físicoquímicos interferem nas concentrações e no perfil de distribuição vertical da clorofila como a profundidade da zona eufótica, a distribuição dos nutrientes inorgânicos dissolvidos, o perfil vertical de temperatura (termoclina), os parâmetros de estratificação como a frequência de Brunt-Vaisala (PLATT e JASBY, 1976; PEREIRA et al 2004; KRIEGER et al, 2005; KAMPEL, 2003). A concentração total da clorofila_a ao longo deste perfil vertical é o valor da biomassa integrada na coluna d'água.

Quanto à sua distribuição horizontal, esta vai depender também de outros fatores como as correntes oceânicas, padrões de vento, proximidade de fontes de nutrientes, etc (referencia). Segundo Leão et al. (2007) as variáveis mais representativas na estimativa da concentração de clorofila integrada na zona eufótica podem ser a profundidade da termoclina, a frequência de Brunt-Vaisala, o nitrato integrado na coluna d'água e temperatura superficial do mar. Estas variáveis variam de acordo com os processos naturais da região como correntes, e da época do ano, pois possuem uma sazonalidade na área de estudo da costa sudeste do Brasil (GAETA et al., 1994).

3.2 Aplicações dos modelos ecológicos

Todo modelo é uma abstração redutiva da realidade, no entanto os modelos ecológicos auxiliam uma série de estudos. Dentre estes estão a compreensão dos processos naturais, os principais fatores que interagem na cadeia de relações e

aplicações na determinação de áreas ou épocas favoráveis de ocorrência de determinados fenômenos, simulações de cenários futuros, entre outros (GEMITZIA et al., 2007).

Mais especificamente os modelos de estimativa de concentração de clorofila são bastante aplicados para compreensão da dinâmica costeira e oceânica. Estes por sua vez podem ser aplicados para o manejo de recursos pesqueiros, estudo de áreas favoráveis a ocorrência de *blooms* algais (ploriferação de espécies), estudos de impactos ambientais e de remediação em áreas eutrofizadas (organicamente enriquecidas), dentre outros (PEPREIRA et al., 2004).

3.3 Geoprocessamento em estudos ambientais

O geoprocessamento representa a área do conhecimento que utiliza técnicas matemáticas e computacionais, fornecidas pelos Sistemas de Informação Geográfica (SIG), para tratar os processos que ocorrem no espaço geográfico (CAMARA, et al., 2001).

A natureza não é uma forma estática da realidade, mas está em constante processo de mudanças resultantes de uma série de interações de diferentes componentes. Assim para os estudos ambientais é necessário que se leve em consideração os principais componentes que interferem nas relações e os processos envolvidos. As ferramentas de geoprocessamento permitem analisar uma série de dados de diferentes variáveis distribuídas no espaço e combiná-las de forma a estudar as relações de interação e extrair informações dos processos envolvidos (referencia). Neste contexto, um *banco de dados geográfico* pode ser visto como uma forma organizada de se manter dados que possam estar correlacionados (CAMARA, et al., 2001).

No ambiente de um SIG as entidades do mundo real podem ser descritas por atributos espaciais, temporais e temáticos. Os atributos espaciais guardam informações sobre localização, topologia e geometria das entidades. Os atributos temporais referem-se à idade do objeto de estudo, à data ou à frequência de aquisição. Os atributos

temáticos referem-se a outras propriedades das entidades, que não são de localização nem temporais (CAMARA, et al., 2001).

As operações de manipulação destes atributos em análise geográfica são usadas para combinar atributos diferentes com o objetivo de encontrar alguma correlação espacial entre eles. Estas operações são denominadas de álgebra de mapas (FUCKS et al., 2004). Assim é possível manipular dados distribuídos no espaço e organizados em bancos de dados através de operações de overlays, gerando novas informações do ambiente em estudo.

3.4 Modelagem de dados em Sistema de Informação Geográfica

Para a modelagem dos dados no SIG é preciso descrever de forma matemática, simbólica ou funcional o processo de interesse a partir de medidas representativas do universo do mundo (CAMARA, et al., 2001). Para isso é preciso conhecer a natureza dos dados que podem ser definidos na representação computacional por geo-campos e geo-objetos. Os geo-campos são utilizados para dados de natureza contínua no espaço e os geo-objetos por elementos definidos num espaço restrito, que podem ter atributos não espaciais e múltiplas localizações. Estes últimos podem ser representados em linhas, pontos e polígonos.

No âmbito dos geo-campos as medidas podem ser em níveis de medidas numéricas ou por níveis de medidas temáticos (CAMARA, et al., 2001). Em termos de representação da espacialidade dessas medidas existem 3 tipos distintos:

- a) pontuais: grade regular de pontos e amostras irregulares que compõe os modelo numéricos de terreno (MNT) do tipo amostra 3D (coordenadas e atributo);
- b) superfície: que são as grades triangular (TIN) e isolinhas, que representam as variáveis por superfícies contínuas com valores variando dentro das regiões e com continuidade nos limiares.
- c) partes constantes: grade regular de células e regiões contíguas, que representam as variáveis por superfícies horizontais representados pelos MNTs de grade retangular.

3.4.1 Interpolação

Nos estudos ambientais geoespacializados, raramente é possível fazer o recobrimento de toda a área de estudo. Normalmente se faz amostragem de pontos distribuídos de forma mais homogênea possível com espaçamento regulares ou irregulares. Assim para se ter uma avaliação da distribuição contínua da variável no espaço é preciso utilizar modelos de inferência para os pontos intermediários não amostrados. Neste caso, supõe-se que a variação entre amostras pode ser descrita por funções matemáticas que representem os valores de forma quase contínua, ao longo do campo da variável, por uma grade regular de pontos. A função de interpolação usada depende da aplicação e da natureza dos dados, mas, em geral, a forma desta função nem sempre é definida claramente (FUCKS et al., 2004).

Existem dois tipos básicos de métodos de interpolação: os determinísticos e os geoestatísticos. Os métodos determinísticos consideram apenas os pesos de determinação dos pontos interpolados em função da distância ($\lambda_i = f(d_i)$). Neste caso a anisotropia comum à variação dos fenômenos naturais não é considerada e pode haver super ou subestimação de valores provocados por formação de clusters (agrupamentos). O raio de influencia das amostras também é determinado arbitrariamente. Os métodos geoestatísticos determinam os pesos a partir da correlação espacial das amostras, avaliado com base no semivariograma ($\lambda_i = f[\gamma(h)]$). Neste caso a área de influência é indicada pelo alcance (dado no semivariograma), e é possível analisar a anisotropia do fenômeno com o semivariograma de superfície. Se existir a anisotropia é possível modelar a função de acordo com os eixos de menor e maior continuidade espacial. Além disso, este método trata redundâncias provocadas pela formação de clusters na amostragem. A dificuldade do uso deste método é que as amostras devem ter uma distribuição próxima à normal para modelar o semivariograma. As etapas são constituídas pela análise exploratória, geração do semivariograma, ajuste do modelo, validação e por fim a krigeagem (FUCKS et al., 2004).

3.4.1.1 Lógica *Fuzzy*

Segundo Burrough e McDonnell (1998), a utilização de um conjunto *Fuzzy* é indicada sempre que se tiver que lidar com ambigüidade, abstração e ambivalência em modelos matemáticos ou conceituais de fenômenos empíricos. Os conjuntos *Fuzzy* é uma metodologia que permite atribuir valores de incerteza no gradiente de transição dos valores reais medidos e os estimados, assim não há limites rígidos (bordas) entre classes (FUCKS et al., 2004). Na natureza de fato os fenômenos raramente possuem limites rígidos. Os métodos baseados em lógica fuzzy são originários de métodos de inteligência artificial que se baseiam no comportamento dos processos naturais. Nele os valores discretos são convertidos numa escala de 0 a 1 gerando intervalos de incerteza que suavizam os gradientes transicionais. Desta forma os modelos baseados em funções fuzzy conseguem representar melhor os gradientes de transição que ocorrem de fato na natureza.

Existem diversos tipo de funções fuzzy para a modelagem das variáveis, mas os mais utilizados são o linear, gaussiano e senoidal. Os parâmetros de determinação dos modelos nas funções vão depender da distribuição das amostras utilizadas e do conhecimento do pesquisador a respeito dos valores limites para inferência da variável (GEMITZIA et al., 2007).

3.4.2 Técnica de suporte à decisão: Análise Multi-critério - *Analytical Hierarchical Process (AHP)*

O Processo Analítico Hierárquico (AHP) é uma técnica que possibilita juntar informações de diferentes naturezas, atribuindo a cada uma delas uma importância relativa, cujo resultado final é uma representação (pode ser na forma de mapa) que contém informação sobre o problema em questão. Os diferentes níveis de informação – os Planos de Informação (PI) – são comparados dois a dois e a importância relativa de cada um deles é definida pelo usuário de acordo com critérios por ele mesmo estabelecidos, baseados no seu conhecimento do problema (FUCKS et al., 2004). Assim este método permite combinar conhecimentos empíricos e teóricos a respeito do

fenômeno estudado, oferecendo maior flexibilidade para o pesquisador interferir na formulação do modelo (GEMITZIA, et al., 2007).

Os passos neste procedimento são:

- a) Compara os critérios dois a dois;
- b) Verificar a consistência dos dados comparando a matriz de pesos com uma matriz aleatória. Considera-se consistente se a probabilidade da matriz ser aleatória foi menor que 10%;
- c) Produzir os pesos de cada variável, cuja soma é 1 e fazer a inferência por média ponderada.

4 MATERIAIS E MÉTODOS

4.1 Área de estudo

Os dados utilizados para a realização deste trabalho foram obtidos da tese de doutorado do Dr. Milton Kampel (KAMPEL, 2003). Seu estudo estava inserido no projeto DEPROAS - *Dinâmica do Ecossistema de Plataforma da Região Oeste do Atlântico Sul*, realizado numa parceria dos institutos IO/USP e INPE.

A área de estudo se estende do Cabo de São Tomé (22°S) até a Ilha de São Sebastião (23°50'S). Os pontos de coleta foram distribuídos em perfis transversais à linha de costa, abrangendo desde profundidades mais rasas na plataforma continental (até aproximadamente 120-180m) e mais de 1000 m no talude. Os dados utilizados neste trabalho foram de coletas de 2 cruzeiros oceanográficos realizados nos verões de 2001 e 2002, totalizando 34 pontos de coleta (Figura 1).

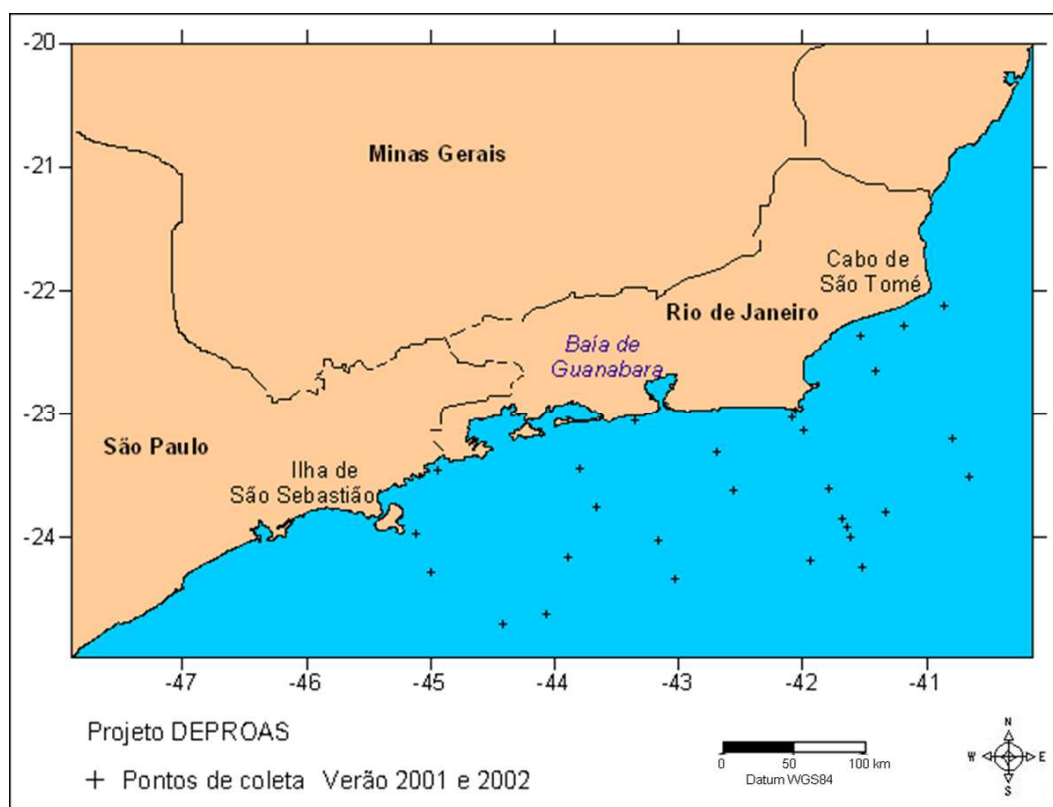


Figura 1: Área de estudo com os pontos de coleta do projeto DEPROAS dos cruzeiros do verão de 2001 e 2002 na plataforma e talude da costa sudeste brasileira.

4.2 Parâmetros oceanográficos

Em cada estação foi realizado um perfil vertical para análise da distribuição das variáveis oceanográficas na coluna d'água. Dados de temperatura e salinidade foram obtidas com auxílio de um perfilador CTD (*Conductivity, Temperature, Depth*). Para análise de concentrações de clorofila a e nutrientes inorgânicos dissolvidos foram coletadas amostras em profundidades de 0, 5, 10, 25, 50, 75, 100, 150 e 200 m, com um sistema de garrafas coletoras acopladas ao CTD. Foram medidos também a profundidade da zona eufótica com disco de Secchi, e o coeficiente de atenuação difusa (K_{PNF}) com auxílio de um Perfilador de Fluorescência Natural (PNF-300, *Biospherical Instruments Inc.*). Os valores de TSM foram obtidos a partir de algoritmos *Multi-Channel Sea Surface Temperature* (MCSST) de imagens NOAA-AVHRR. Outras variáveis também foram medidas no estudo de Kampel (2003), mas apenas estas que foram explicitadas foram utilizadas no presente trabalho. Maiores detalhes dos métodos de análise de cada variável podem ser consultadas em Kampel (2003).

As variáveis neste trabalho utilizadas foram: as profundidades da termoclina (zt), a concentração máxima de clorofila_a (em metros); frequência de Brünt-Väisälä (N^2) ($s^{-1} \times 10^{-5}$); nitrato integrado na zona eufótica (Neu) (em $\mu m.m^{-2}$); temperatura da superfície do mar (TSM) ($^{\circ}C$); e a concentração de clorofila_a integrada na coluna d'água (Ce_u) ($mg.m^{-2}$). Os valores do nitrato inorgânico dissolvido e a concentração clorofila foram integrados na coluna d'água ao longo da zona eufótica utilizando as medidas de cada profundidade e o método de integração dos trapézios (KAMPEL, 2003).

4.3 Modelagem dos dados no Sistema de Informação Geográfica

A modelagem proposta neste trabalho visou utilizar as técnicas de geoprocessamento em ambiente SIG para gerar superfícies contínuas a partir de amostras pontuais e integrar as variáveis oceanográficas com diferentes métodos para a estimativa da concentração da clorofila integrada na coluna d'água. O trabalho foi realizado no programa SPRING 4.3.

O modelo do banco de dados geográficos (OMT-G) para a realização do trabalho está descrito na figura 2, como as variáveis geográficas em seus respectivos formatos e o conjunto de operações que foram realizados no processo de modelagem.

O programa SPRING, Sistema para Processamento de Informações Georeferenciadas, utilizado no desenvolvimento deste trabalho, foi desenvolvido pelo INPE (Instituto Nacional de Pesquisas Espaciais). Foi escolhido este programa devido ao acesso gratuito e possibilidade da aplicação das ferramentas propostas no trabalho, desde a construção do banco de dados até os processamentos para o resultado final. O programa possui uma interface amigável e suporta diversos tipos de dados matricias, vetoriais e não espacial (em formato de tabelas cadastrais). A Linguagem Espacial para Geoprocessamento Algébrico – LEGAL permite a realização de diversos tipos de processamentos e manipulação do banco de dados, com uma linguagem de fácil programação.

Os resultados finais dos modelos foram representados na forma de cartas elaboradas no ambiente SCARTA do SPRING 4.3.

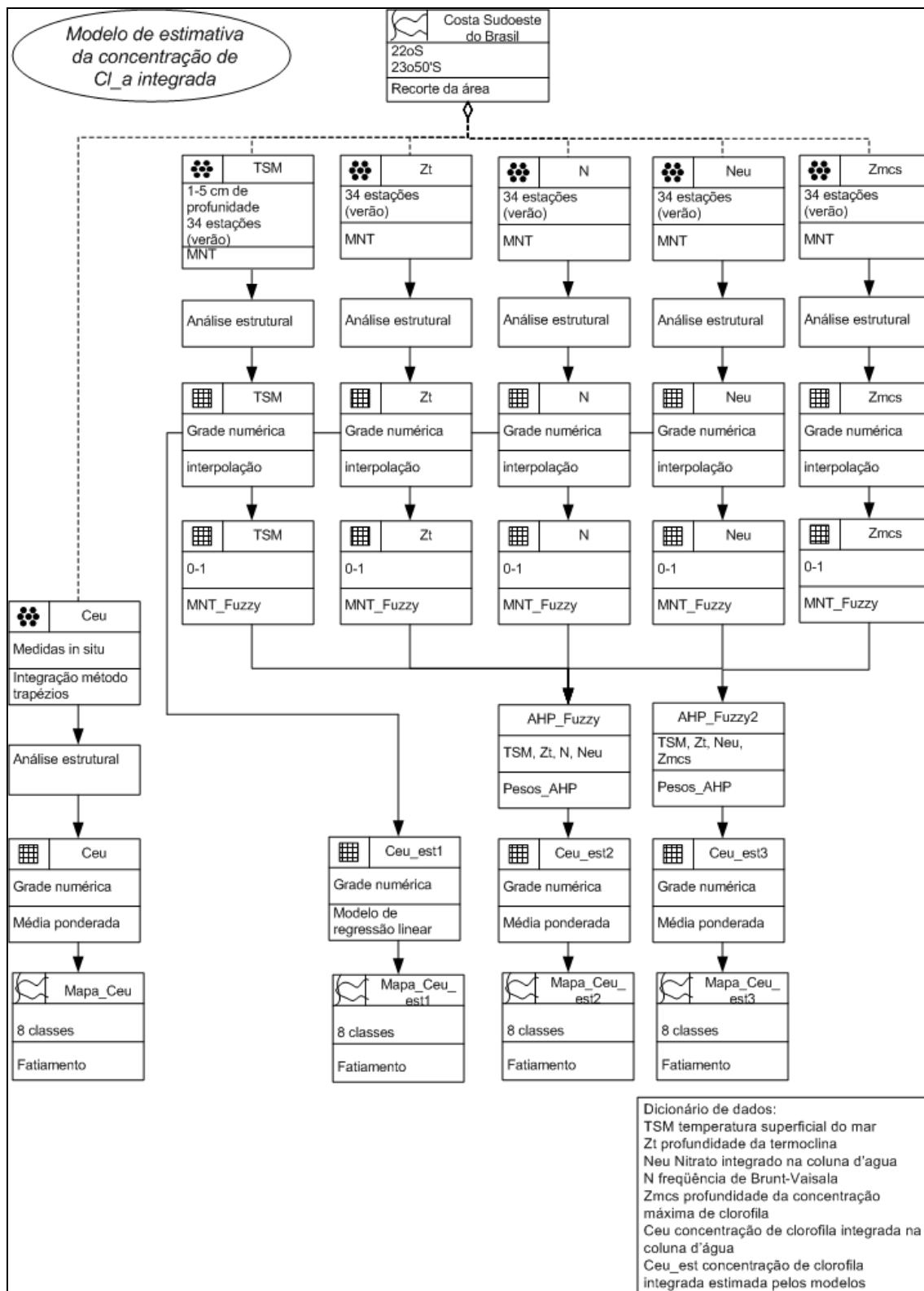


Figura 2: Modelo OMT-G do banco de dados para a modelagem da estimativa da concentração de clorofila integrada na coluna d'água.

4.4 Definição do Projeto e importação dos dados para o SPRING

Primeiro foi criado o banco de dados sob nome Clorofila e um projeto com o nome Deproas. O projeto foi definido com a projeção geográfica Cilíndrica equidistante, datum WGS 84. Esta projeção conserva a proporção entre as distâncias, em determinadas direções, na superfície representada. Os limites do retângulo envolvente foram de 47° 57' 0.00" oeste e 25° 1' 0.00" sul para os valores mínimos de longitude e latitude, e 40° 3' 0.00" oeste e 19° 54' 0.00" sul, para os valores máximos. Optou-se por utilizar esta projeção por se tratar de uma área de estudo ampla e de pequena escala (CAMARA et al., 2001).

Posteriormente foi criado o modelo do banco de dados segundo as categorias que seriam usadas para a geração de cada plano de informação, no desenvolvimento do trabalho. Ao longo do trabalho também foram incorporados novas categorias e as respectivas classes de acordo com a necessidade dos dados gerados.

Para definição da área de estudo foi utilizado um arquivo vetorial de extensão SHAPE com um recorte da linha de costa do sudeste do Brasil. Este recorte foi utilizado como máscara para espacializar os dados apenas na área do mar.

Os dados de cada variável oceanográfica foram transformadas em arquivos de extensão DBASE (versão 4) com as coordenadas x, y dos pontos amostrais e o valor z de medida da propriedade. Estes arquivos foram transformados em arquivos ASCII SPRING com formatos de pontos 3D no modelo MNT. Por fim, estes foram importados ao SPRING e associados à categoria amostras previamente criada no modelo de banco de dados. Todo os dados foram trabalhados em escala de 1:500000, com resolução de 1000 m (x,y).

4.5 Processamentos

4.6 Análise exploratória e geração das grades retangulares

Segundo Burrough et al. (1998) a análise exploratória é realizada com o objetivo de averiguar as propriedades estatísticas do conjunto de dados e identificar eventuais problemas de amostragem. Assim, primeiramente foi realizada uma análise exploratória dos dados, com os parâmetros estatísticos de média, variância e distribuição.

Foram testados diversos métodos de interpolação testados para a geração das grades regulares. O método de interpolação que foi mais adequado à distribuição das variáveis oceanográficas foi escolhido para a geração das grades retangulares para distribuição espacial das amostras.

4.7 Determinação das variáveis para o modelo de estimativa

Para a determinação das variáveis mais representativas que seriam usadas nos modelos de estimativa da concentração de clorofila integrada na coluna d'água, primeiramente foram realizadas análises do comportamento da distribuição das variáveis e suas relações com a concentração de clorofila integrada, obtida por medidas *in situ*.

A priori foram escolhidas as variáveis que foram utilizadas por de Leão et al (2007), num modelo de regressão linear múltipla para esta mesma estimativa, e que usou os mesmos dados de Kampel (2003). Em seu trabalho Leão et al (2007), separou as variáveis mais representativas, eliminando variáveis redundantes (com comportamentos semelhantes) e elaborou um modelo com R^2 de 0,61 e R^2 ajustado de 0,55.

A equação do modelo para a estimativa da clorofila integrada na plataforma e talude da área estudada e para o período de verão foi:

$$\text{Log } C_{EU} = 4,591 - 0,002N^2 - 0,033Z_T - 0,033TSM - 0,336 \frac{1}{N_{EU}};$$

Onde N^2 é a frequência de Brünt-Väisälä; Z_t a profundidade da termoclina; TSM a temperatura superficial do mar; e $1/Neu$. o inverso da concentração de nitrato integrado na coluna d'água.

Além destas variáveis, mais uma foi utilizada nos modelos geoespaciais por ter apresentado uma relação mais forte com a concentração da clorofila integrada na coluna d'água. Esta variável foi a profundidade da máxima concentração de clorofila na zona eufótica (Z_{mcs}), que teve um R^2 de 0,36, apresentando uma relação inversa com a clorofila integrada.

4.8 Modelos Fuzzy

Para a elaboração dos modelos de função fuzzy de cada variável primeiramente foi feito um estudo da relação de cada variável independente do modelo de estimativa com a concentração de clorofila integrada na coluna d'água calculada pela integração do método dos trapézios com as medidas *in situ*. Este estudo prévio foi para avaliar quais os valores mínimos e máximos no qual a variável possivelmente apresenta uma influencia na concentração de clorofila.

Numa segunda etapa foram analisados os gráficos da curva de distribuição das variáveis em ordem crescente, para auxiliar na determinação no valor médio correspondente de z (0,5) para a determinação do coeficiente da função (α). E também para avaliar qual função fuzzy se ajustava melhor à distribuição da variável.

Foram testados três tipos de funções: a linear, a senoidal e a gaussiana. A que melhor se adequou à distribuição de todas as variáveis foi a função senoidal. Esta função é descrita pela seguinte equação:

$$f(z) = \frac{1}{1 + (\alpha(z - \beta))^2}$$

onde, z é o valor referente à medida da variável, β é o valor máximo estipulado para a para a função e α o parâmetro a determina a forma da função, sendo descrito pela equação:

$$\alpha = \frac{1}{(z_{(0.5)} - \beta)^2}$$

Sendo que, $z_{(0.5)}$ corresponde ao valor de z no ponto central.

Depois da elaboração dos modelos de função fuzzy foram gerados MNTs aplicando a função fuzzy de cada variável aos MNT das interpolações utilizando programação em LEGAL. Todos os programas LEGAL estão em anexo.

4.8.1 Análise Multi-critério - *Analytical Hierarchical Process (AHP)*

Para a definição dos critérios usados no cálculo dos pesos pela análise multi-critério, foi usado como base a ordem a proporção de importância das variáveis em relação à clorofila integrada. No trabalho de Leão et al (2007), os autores descrevem estas proporções em porcentagem de importância, com os respectivos valores:

N²	5,80
TSM	7,88
Inverso N_{EU}	9,06
Z_T	77,26

Lembrando que N² é a frequência de Brünt-Väisälä; Z_t a profundidade da termoclina; TSM a temperatura superficial do mar; e 1/N_{eu}, o inverso da concentração de nitrato integrado na coluna d'água.

Os dados utilizados para o modelo AHP foram as grades retangulares dos valores normalizados pelos modelos de funções fuzzy. Foi gerado um modelo as quatro variáveis do modelo de regressão de Leão et al (2007) e um modelo substituindo a frequência de Brunt-Vaisala (N²) pela profundidade do máximo de clorofila (Z_{eu}). Esta última teve um critério qualitativo maior que as outras variáveis por apresentar o maior valor de R² (0,36) em relação à clorofila integrada.

Para cada modelo foi gerado um mapa MNT com programa em LEGAL que podem ser encontrados em anexo.

4.9 Mapas finais

Por fim para a comparação qualitativa do desempenho dos modelos, os mapas de superfície contínua gerados pelos diferentes processamentos foram transformados em mapas temáticos por meio da operação de fatiamento. Cada mapa foi gerado com 8 classes de valores maiores e menores de concentração de clorofila. Os mapas estão em diferentes escala de valores devido aos diferentes processamentos aplicados, mas as classes são equivalentes. Às classes foram atribuídos valores qualitativos de concentrações de clorofila com: baixa concentração de clorofila, baixa, moderadamente baixas, média, moderadamente alta e alta. Os mapas temáticos foram editados no Scarta 4.3, um programa do Spring para a elaboração das cartas finais.

Os mapas finais foram:

- Concentração da clorofila integrada na coluna d'água obtida por medidas *in situ* e o método de integração dos trapézios;
 - Método de interpolação por média ponderada
 - Método de interpolação por krigeagem
- Concentração de clorofila integrada estimada pelo modelo de regressão linear numa escala de log.
- Concentração de clorofila integrada estimada pelo método AHP fuzzy_1
- Concentração de clorofila integrada estimada pelo método AHP fuzzy_ 2, usando uma diferente combinação de variáveis.

5 RESULTADOS E DISCUSSÃO

5.1 Análise exploratória e geração das grades retangulares

Nenhuma das variáveis oceanográficas analisadas apresentou uma distribuição próxima à normal como mostra na figura 3. Em geral todas tiveram distribuições foram assimétricas e a média diferente da moda e mediana. As variâncias também foram bastante altas.

A clorofila integrada (Ce_u) apresentou uma média de 23.63 mg.m⁻² e variância de 213.89. A frequência de Brunt-Vaisala (N₂) teve uma média de 40.98 (s⁻¹ x 10⁻⁵) e variância de 1190.41. O nitrato integrado (Ne_u) teve média de 18.75 e variância de 317.42. A temperatura superficial do mar (TSM) teve média de 24.09 °C e variância de 10.95. Este foi o parâmetro que teve menor variância em relação à média. A profundidade da termoclina teve média de 22.56 m e variância de 142.50. E a profundidade da concentração máxima de clorofila teve uma média de 54.96 e variância de 114.29.

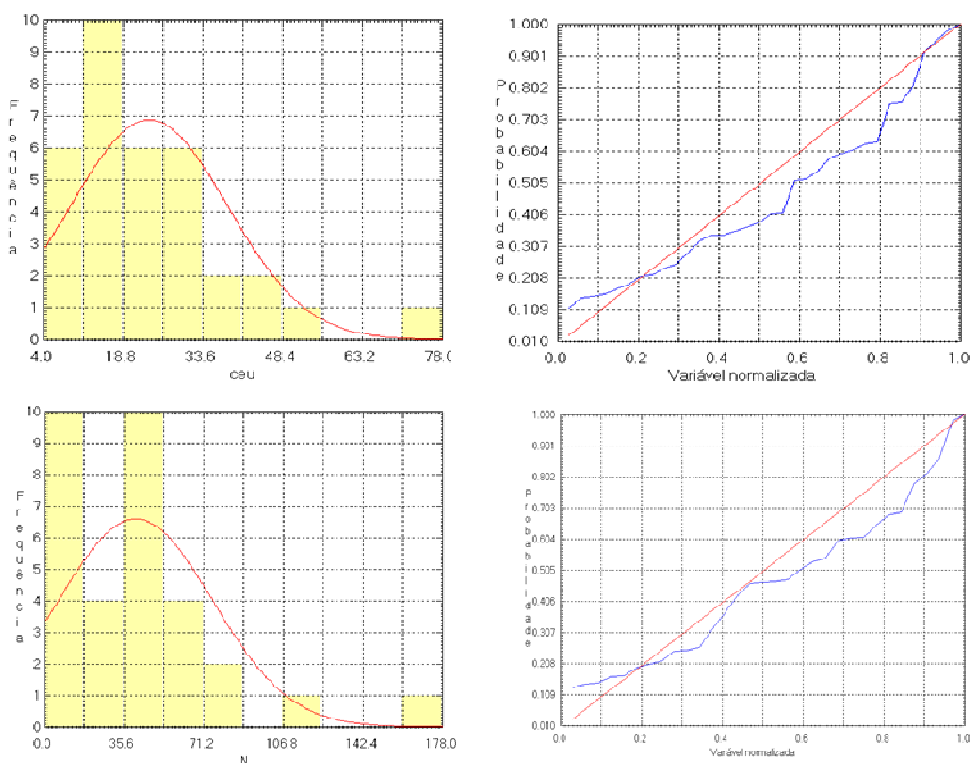


Figura 3: Histogramas e gráficos de curva de distribuição das variáveis.

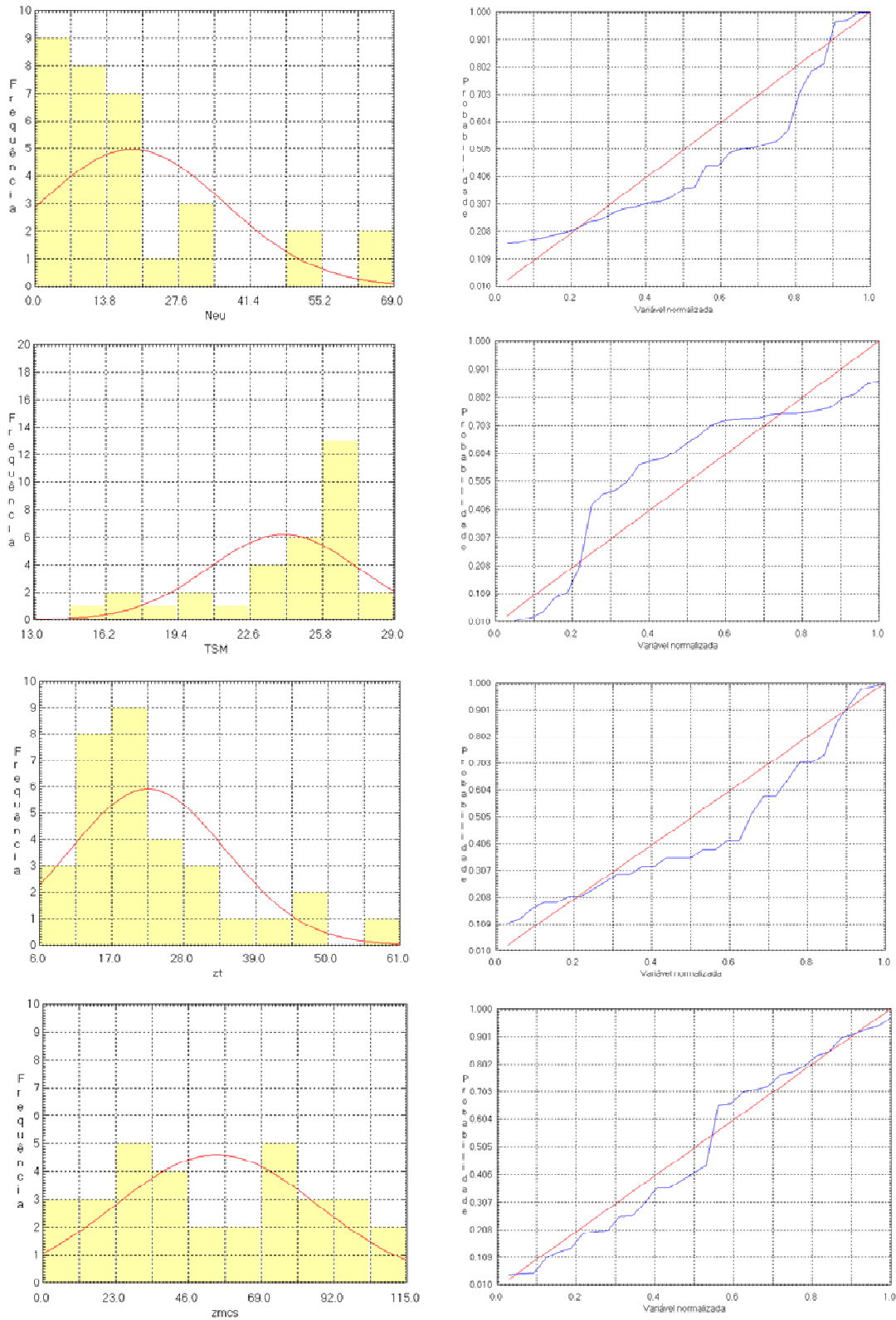


Figura 3 (continuação): Histogramas e gráficos de curva de distribuição das variáveis.

O que pode ter contribuído para a grande variabilidade das variáveis na área de estudo é o fato de abranger uma área bastante extensa e com dinâmicas distintas, com estações que vão desde bem próximas à costa até o talude. No trabalho de Kampel (2003) o autor separou as variáveis em ambiente de plataforma e talude devido às distribuições distintas que elas apresentaram nestas áreas.

5.2 Interpolação

Como as variáveis não possuíam uma distribuição normal não foi possível aplicar um modelo de semivariograma para fazer a interpolação por krigeagem ordinária. Assim, foram testados os modelos de interpolação por métodos determinísticos. O método que apresentou a melhor distribuição foi escolhido para dar seguimento ao trabalho. Os métodos foram testados para todas as variáveis, mas como todas tiveram comportamentos de distribuição semelhantes, apenas uma foi usada para ilustrar os resultados dos diferentes métodos, que foi a clorofila integrada (figura 4).

O método mais simples é o de vizinho mais próximo que atribui a todos os pontos mais próximos da estação o seu valor da variável. Como é possível notar na figura x, este método não representa o gradiente de variação dos fenômenos como ocorre na natureza. Portanto para expressar a distribuição das variáveis em estudo este método não é bem aplicado. A interpolação por média simples ficou um pouco melhor que o vizinho mais próximo, mais ainda não representou bem o gradiente de distribuição das variáveis em função da distância.

O mapa de distribuição por interpolação com média ponderada foi o que apresentou a melhor distribuição das variáveis. Representou um gradiente suave de transição ponderada pela distância das estações. Ainda assim é uma representação bastante simplista já que não considerada a correlação entre os pontos e os eixos de maior e menor variação como discutido anteriormente, mas para um método determinístico os resultados foram bastante razoáveis.

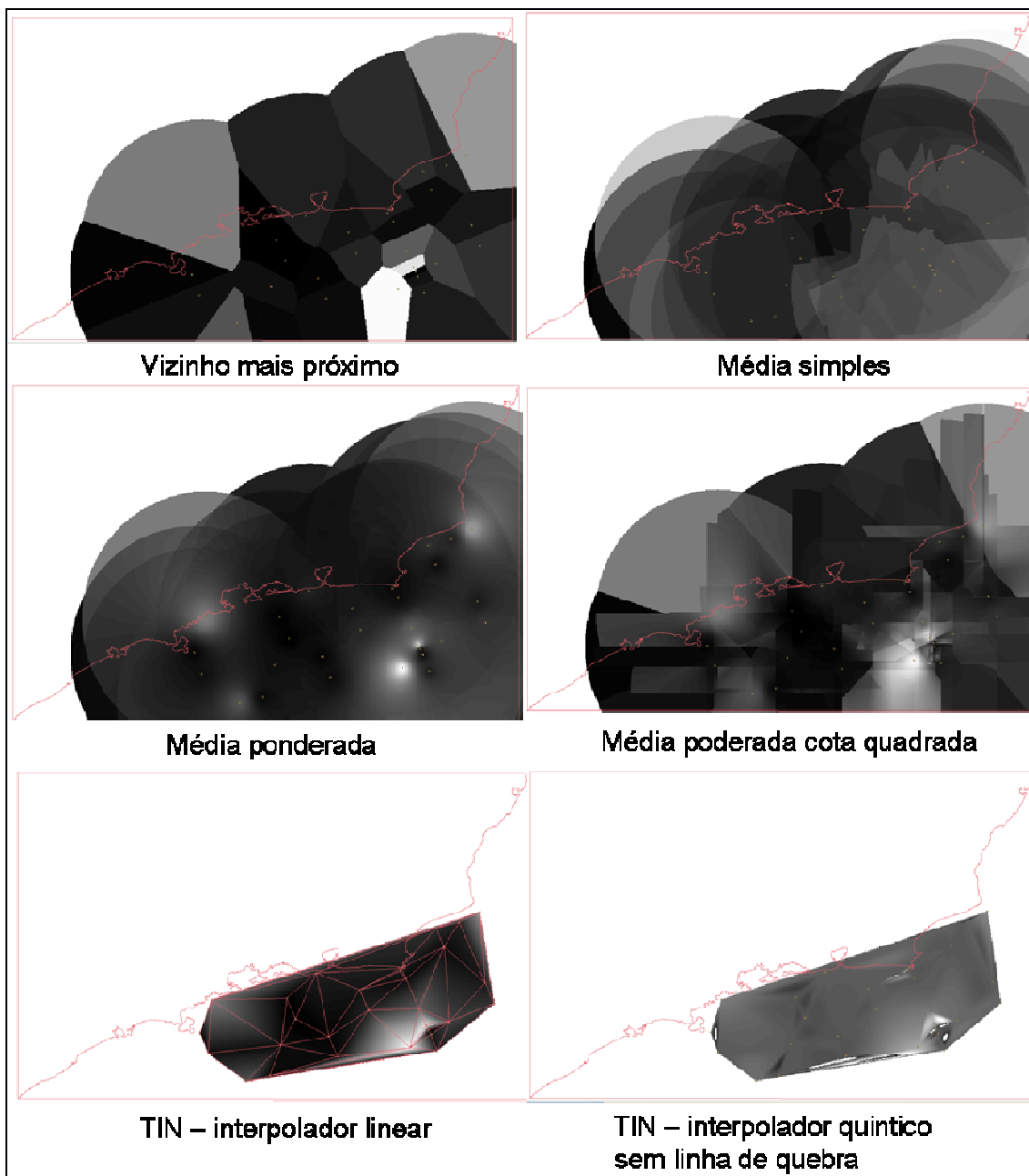


Figura 4: Mapas de distribuição da clorofila integrada na coluna d'água com diferentes métodos de interpolação.

A interpolação por média ponderada por cota quadrada não teve uma distribuição muito boa. Os interpoladores utilizados para a grade triangular para este tipo de variável também não apresentaram bons resultados. Assim, o interpolador escolhido para a modelagem das variáveis oceanográficas foi o de média ponderada.

5.3 Análise espacial das variáveis interpoladas

Na figura a seguir serão apresentados os mapas de distribuição resultado da interpolação de todas as variáveis com o método da média ponderada.

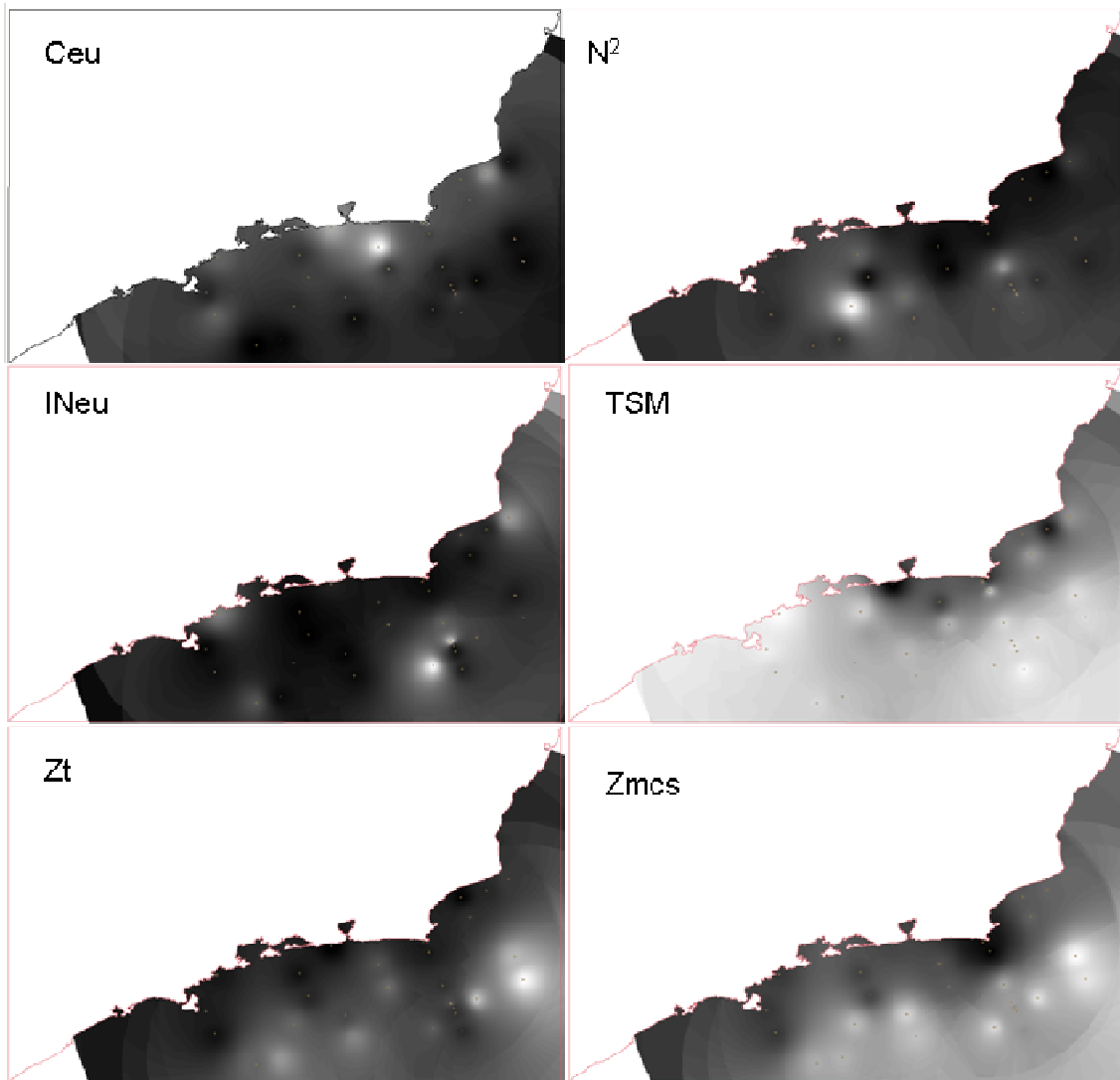


Figura 5: Mapas de distribuição das variáveis interpoladas com o método de média ponderada com imagem em níveis de cinza. O preto corresponde aos menores valores e o branco aos maiores.

A interpolação por média ponderada usa uma medida de tendência que suaviza a estimativa dos valores interpolados. Como foi observado na análise exploratória as

variáveis apresentam valores máximos e mínimos bem distantes da média (alta variância). Estes valores não são representados na interpolação por média ponderada. Assim esta interpolação gera apenas os valores de tendência.

Ainda assim é possível notar algumas tendências de distribuição em decorrência dos processos naturais que ocorrem nessa área de estudo. A clorofila integrada e o nitrato tiveram os maiores valores na plataforma, próximos à costa. Notar que no caso do nitrato foram utilizados valores do inverso da concentração (INeu), por isso os valores mais escuros correspondem a maiores valores de nitrato integrado na coluna d'água. Foi utilizada esta transformação pois foi a mesma usada para o modelo de regressão linear no trabalho de Leão et al (2007). Essa distribuição condiz com o que se espera da Água Costeira (AC), mais ricas em nutrientes inorgânicos e conseqüentemente com maiores de clorofila devido à influencia da drenagem continental (GAETA, 1994; KAMPEL, 2003)

A frequência de Brunt-Vaisala (N^2), a temperatura superficial do mar (TSM), e as profundidades da termoclina (zt) e da concentração máxima de clorofila (zmcs), tiveram uma distribuição inversa da clorofila e do nitrato, com os maiores valores mais distantes da costa em direção ao talude. Essa distribuição com menores valores de TSM próximo à costa principalmente no litoral do Rio de Janeiro pode ser explicada pelas correntes que regem os processos costeiros nesta área durante o período do verão. Analisando o mapa (figura x) podemos ver que nesta área da costa do Rio de Janeiro as águas são mais frias, com maiores concentrações de nitrato e clorofila, ou seja, são águas frias organicamente enriquecidas. Isso caracteriza águas típicas de fenômenos de ressurgência costeiras que trazem águas frias e organicamente ricas do fundo da plataforma para a superfície na costa. De fato durante o período do verão estes fenômenos são bastante ocorrentes no litoral do Rio de Janeiro. O que ocorre devido à intrusão da corrente da Água Central do Atlântico Sul (ACAS) na quebra da plataforma em direção à costa, atingindo a parte mais interna da plataforma continental (KAMPEL, 2003).

5.4 Modelo de Regressão Espacializado

Para gerar um mapa de distribuição do modelo de regressão linear múltiplo proposto por Leão et al (2007), foram utilizados dos modelos numéricos de terreno das variáveis oceanográfica e realizada a operação de algebra de mapas pela seguinte fórmula:

$$\text{Log } C_{EU} = 4,591 - 0,002N^2 - 0,033Z_T - 0,033TSM - 0,336 \frac{1}{N_{EU}}; \text{ (LEÃO et al., 2007)}$$

Onde N^2 é a frequência de Brünt-Väisälä; Z_t a profundidade da termoclina; TSM a temperatura superficial do mar; e $1/N_{EU}$ o inverso da concentração de nitrato integrado na coluna d'água. Os resultados deste modelo são dados em escala logarítmica da concentração de clorofila integrada (LogCeU).

O mapa de distribuição gerado está exposto na figura 6. A priori o modelo estimou maiores concentrações da clorofila integrada na zona próxima à costa e menores na região do talude, como era espera. Mas os resultados deste modelo serão mais detalhados no item final de comparação dos modelos com os mapas temáticos gerados pelo fatiamento.

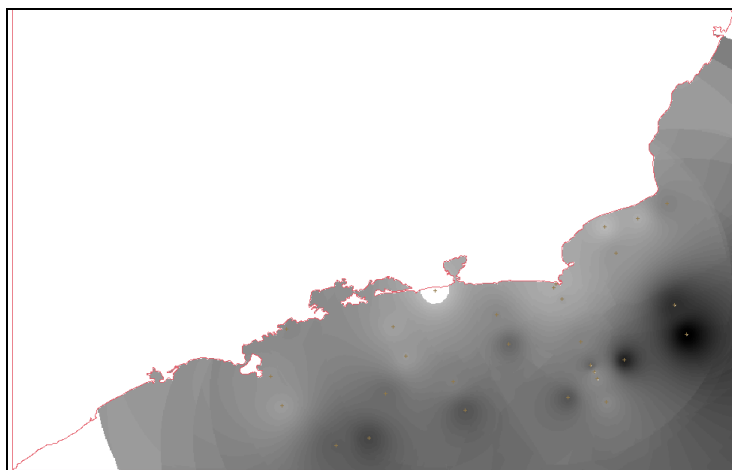


Figura 6: Mapa de distribuição da clorofila integrada estimada pelo modelo de regressão linear múltiplo de Leão et al (2007).

5.5 Modelos Fuzzy

Como já comentado na metodologia, foram testados 3 diferentes funções fuzzy para as variáveis: o modelo linear, gaussiano e senoidal. E o que se ajustou melhor em todos

os casos foi o modelo senoidal. Em geral todas as variáveis tiveram valores de R^2 baixos com a concentração de clorofila integrada, no entanto, como já apresentado na introdução e metodologia, no modelo de regressão múltipla estas variáveis combinadas tiveram um valor de R^2 elevado (0,61).

Os mapas de distribuição gerados pelos modelos fuzzy diminuíram os limites de mínimos e máximos ressaltando valores desses limites e suavizaram o gradiente de transição nos valores interpolados.

A frequência de Brunt-Vaisala teve um R^2 de 0,12 com a concentração de clorofila integrada. O valor mínimo estipulado para a função fuzzy senoidal foi de $7,56 \text{ s}^{-1} \times 10^{-5}$, o valor máximo de $109,88 \text{ s}^{-1} \times 10^{-5}$ e o α calculado foi de 0,000214 (figura 7).

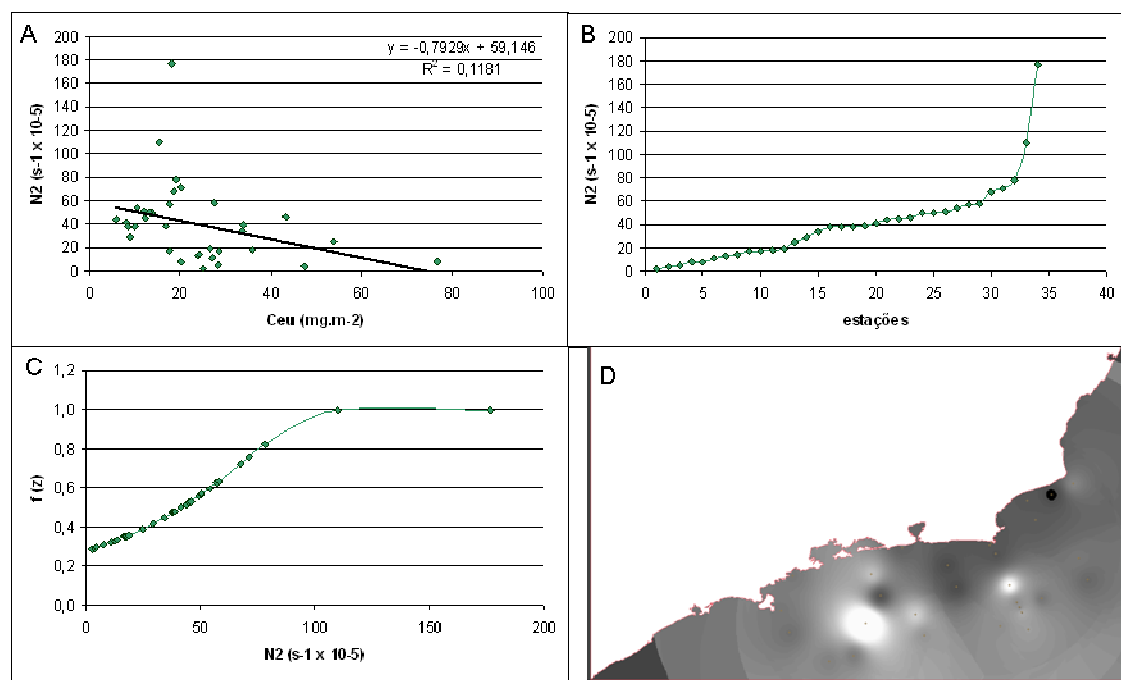


Figura 7: A) Gráfico de distribuição da frequência de Brunt-Vaissala (N2) versus a concentração de clorofila integrada; B) Gráfico de distribuição de N2 em ordem crescente; C) Função fuzzy de N2; D) Mapa da distribuição de N2_fuzzy.

O inverso da concentração de nitrato integrado na coluna d'água teve um R^2 de 0,20 com a concentração de clorofila integrada. O valor mínimo estipulado para a função fuzzy senoidal foi de 0,0092, o valor máximo de 0,37 e o α calculado foi de 73,60 (figura 8).

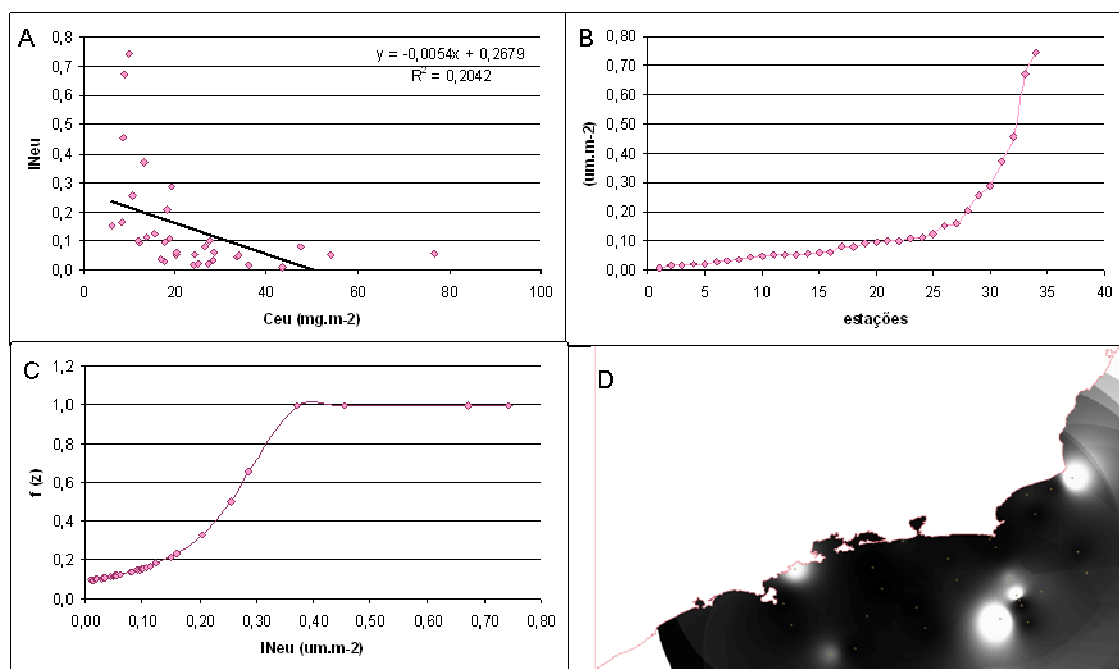


Figura 8: A) Gráfico de distribuição do inverso da concentração de nitrato integrado (INeu) versus a concentração de clorofila integrada; B) Gráfico de distribuição de INeu em ordem crescente; C) Função fuzzy de INeu; D) Mapa da distribuição de INeu_fuzzy.

A temperatura superficial do mar (TSM) teve um R^2 de 0,30 com a concentração de clorofila integrada. O valor mínimo estipulado para a função fuzzy senoidal foi de 17,53 °C, o valor máximo de 27,04 °C e o α calculado foi de 0,12 (figura 9).

A profundidade da termoclina (zt) teve um R^2 de 0,21 com a concentração de clorofila integrada. O valor mínimo estipulado para a função fuzzy senoidal foi de 8 m, o valor máximo de 35 m, e o α calculado foi de 0,0039 (figura 10).

A profundidade da concentração máxima de clorofila (zmcs) teve um R^2 de 0,36 com a concentração de clorofila integrada. O valor mínimo estipulado para a função fuzzy senoidal foi de 1 m, o valor máximo de 97,92, e o α calculado foi de 0,000432 (figura 11).

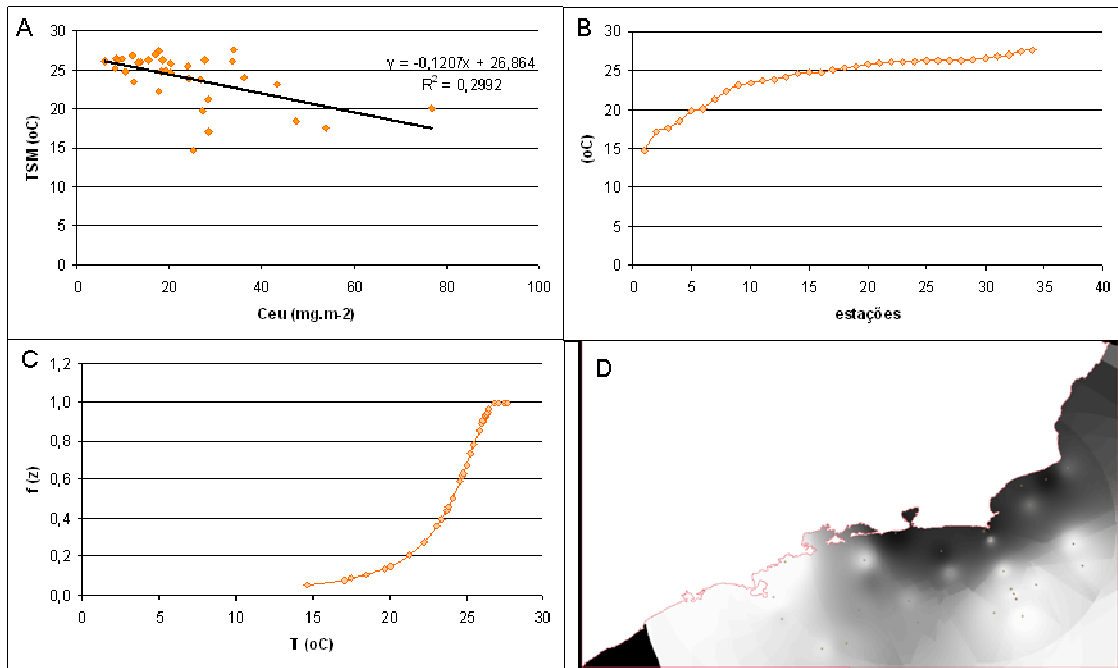


Figura 9: A) Gráfico de distribuição da temperatura superficial do mar (TSM) versus a concentração de clorofila integrada; B) Gráfico de distribuição de TSM em ordem crescente; C) Função fuzzy de TSM; D) Mapa da distribuição de TSM_fuzzy.

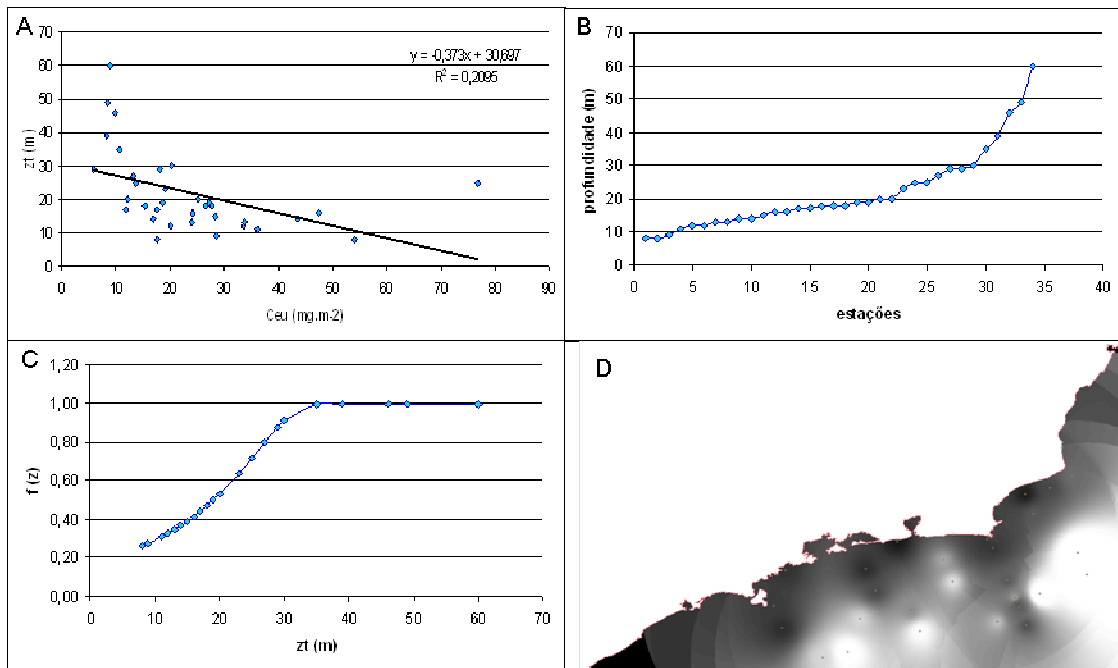


Figura 10: A) Gráfico de distribuição da profundidade da termoclina (zt) versus a concentração de clorofila integrada; B) Gráfico de distribuição de zt em ordem crescente; C) Função fuzzy de zt; D) Mapa da distribuição de zt_fuzzy.

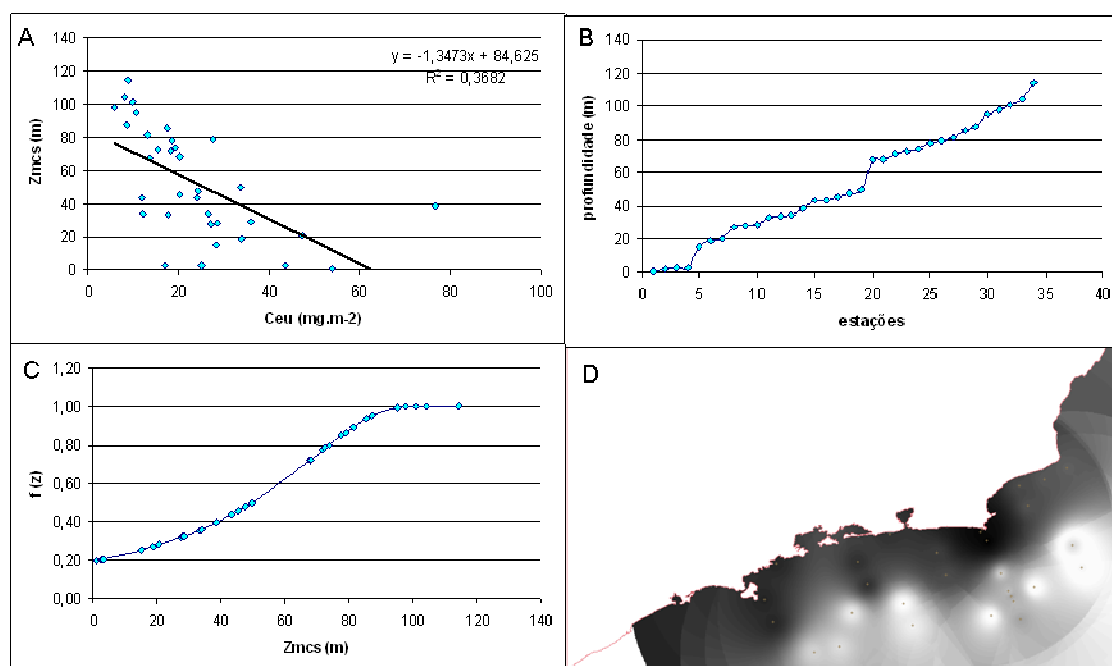


Figura 11: A) Gráfico de distribuição da profundidade da concentração máxima de clorofila (zmcs) versus a concentração de clorofila integrada; B) Gráfico de distribuição de zmcs em ordem crescente; C) Função fuzzy de zmcs; D) Mapa da distribuição de zmcs_fuzzy.

5.5.1 Análise Multi-critério - *Analytical Hierarchical Process (AHP)*

5.4.1.1. Modelo AHP_Fuzzy_1

No primeiro modelo de análise multicritério para a estimativa da concentração da clorofila integrada na coluna d'água foram utilizados os MNTs dos modelos fuzzy das variáveis: frequência de Brünt-Väisälä; profundidade da termoclina; temperatura superficial do mar; o inverso da concentração de nitrato integrado na coluna d'água. As mesmas utilizadas no modelo de regressão linear proposto por Leão et al., (2007). Como já comentado os critérios para cálculo dos pesos foram determinados com base na porcentagem de importância de cada variável em relação à concentração de clorofila integrada apresentado em Leão et al., (2007).

O inverso da concentração de nitrato (INeu_Fuzzy) (9,06% de importância) foi considerado algo melhor que a frequência de Brunt-Vaisala (N_Fuzzy) (5,08% de

importância) e um pouco melhor que a temperatura superficial do mar (TSM_Fuzzy) (7,88% de importância). A TSM_Fuzzy foi considerada um pouco melhor que o N_Fuzzy. A profundidade da termoclina (zt_Fuzzy) (77,26% de importância) foi bem melhor que o INeu_Fuzzy, muito melhor que o TSM_Fuzzy e criticamente melhor que o N_Fuzzy (Figura 12). A razão de consistência calculada para este modelo foi de 0,031.

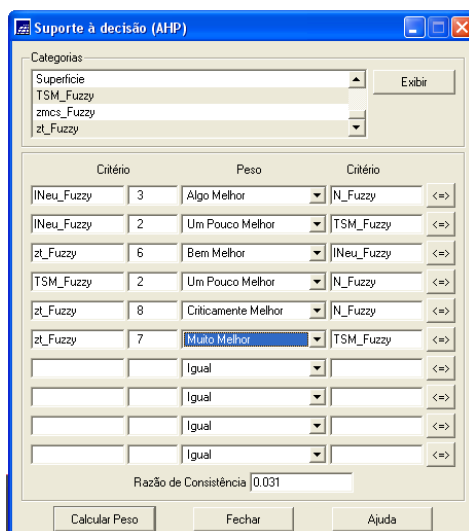


Figura 12: Critérios para cálculo de pesos no AHP_Fuzzy_Modelo1.

A equação gerada pelo modelo foi:

$$Ce_u_{est_1} = 0.158 * INeu_Fuzzy + 0.061 * N_Fuzzy + 0.097 * TSM_Fuzzy + 0.684 * zt_Fuzzy,$$

onde $ce_u_{est_1}$ é a concentração de clorofila integrada com o modelo AHP_Fuzzy_1.

O mapa de distribuição em níveis de cinza gerado pelo MNT do modelo AHP_Fuzzy_1 está na figura 13.

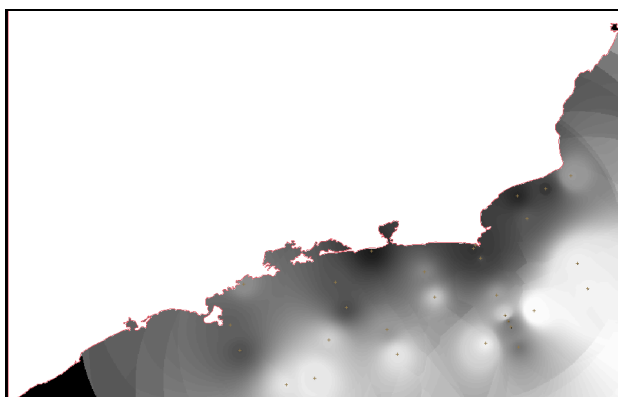


Figura 13: Mapa de distribuição do MNT do modelo AHP_Fuzzy_1 de estimativa da concentração de clorofila integrada.

5.4.1.2. Modelo AHP_Fuzzy_2

No segundo modelo de análise multicritério para a estimativa da concentração da clorofila integrada na coluna d'água a diferença foi ao invés de utilizar a variável frequência de Brünt-Väisälä, foi utilizada a profundidade da concentração máxima de clorofila, pois esta apresentou um valor de R^2 com a concentração de clorofila integrada mais elevado (0,36).

Os critérios para cálculo dos pesos do processo AHP foram os mesmos do modelo anterior, exceto neste modelo a nova variável utilizada da profundidade da concentração máxima de clorofila teve um peso maior que todos por ter apresentado um R^2 maior que todos na análise de regressão. Assim os pesos foram: O inverso da concentração de nitrato (INeu_Fuzzy) um pouco melhor que a temperatura superficial do mar (TSM_Fuzzy). A profundidade da termoclina (zt_Fuzzy) bem melhor que o INeu_Fuzzy e muito melhor que o TSM_Fuzzy. A profundidade da concentração máxima de clorofila (zmcs_Fuzzy) um pouco melhor que o zt_Fuzzy, muito melhor que o INeu_Fuzzy, e criticamente melhor que a TSM_Fuzzy. A razão de consistência calculada para este modelo foi de 0,028 (Figura 14).

The screenshot shows the 'Suporte à decisão (AHP)' window. It features a list of categories on the left: INeu_Fuzzy, Imagem, N_Fuzzy, and Param_Fuzzy. The main area contains a comparison matrix with columns for 'Critério', 'Peso', and 'Critério'. The matrix includes comparisons between INeu_Fuzzy, zmcs_Fuzzy, zt_Fuzzy, and TSM_Fuzzy. The 'Razão de Consistência' is displayed as 0.028. Buttons for 'Calcular Peso', 'Fechar', and 'Ajuda' are at the bottom.

Critério	Peso	Critério	
INeu_Fuzzy	2	Um Pouco Melhor	TSM_Fuzzy
zmcs_Fuzzy	7	Muito Melhor	INeu_Fuzzy
zt_Fuzzy	6	Bem Melhor	INeu_Fuzzy
zmcs_Fuzzy	8	Criticamente Melhor	TSM_Fuzzy
zt_Fuzzy	7	Muito Melhor	TSM_Fuzzy
zmcs_Fuzzy	2	Um Pouco Melhor	zt_Fuzzy
		Igual	
		Igual	
		Igual	
		Igual	

Figura 14: Critérios para cálculo de pesos no AHP_Fuzzy_Modelo 2.

A equação gerada pelo modelo foi:

$Ceu_est_2=0.076*INeu_Fuzzy+0.05*TSM_Fuzzy+0.528*zmcS_Fuzzy+0.347*zt_Fuzy$
onde ceu_est_2 é a concentração de clorofila integrada com o modelo AHP_Fuzzy_2

O mapa de distribuição em níveis de cinza gerado pelo MNT do modelo AHP_Fuzzy_2 está na figura 15.



Figura 15: Mapa de distribuição do MNT do modelo AHP_Fuzzy_2 de estimativa da concentração de clorofila integrada.

5.6 Mapas finais e comparação dos métodos

Para a geração dos mapas finais de distribuição da concentração de clorofila integrada na coluna d'água obtida por medidas *in situ*, pelo modelo de regressão múltipla e pelos modelos fuzzy com análise multi-critério, os MNTs foram divididos em 8 classes “equivalentes” associando os intervalos numéricos de cada MNT. Como cada MNT tinha um valor numérico de saída diferente devido aos processos de modelagem, as classes são correspondentes qualitativamente, mas possuem intervalos numéricos diferenciados.

Na tabela abaixo estão descritos os intervalos numéricos correspondentes das classes geradas nos mapas temáticos (Tabela 1). Os valores de céu ($mg.m^{-2}$) são os valores reais medidos *in situ* e integrados com o método dos trapézios. O valores LogCeus são valores em escala logarítmica do modelo de regressão múltipla e o Fuzzy são os valores dos mapas dos modelos fuzzy que vão de 0 a 1.

Tabela 1: Intervalos numéricos correspondentes às classes dos mapas de distribuição da concentração de clorofila integrada.

Classes	Ceu (mg.m^{-2})	LogCeu	Fuzzy (0-1)
Muito baixa	<14	<1,74	<0.125
Baixa	14-23	1.74-2.02	0.125-0.25
Moderadamente baixa	23-32	2.02-2.29	0.25-0.375
Média baixa	32-41	2.29-2.57	0.375-0.5
Média	41-50	2.57-2.85	0.5-0.625
Média alta	50-59	2.85-3.13	0.625-0.75
Moderadamente alta	59-67	3.13-3.40	0.75-0.875
Alta	>67	>3.40	>0.875

Como já discutido, na interpolação por média ponderada não é levado em consideração a correlação espacial das estações, nem a variação anisotrópica das variáveis estudadas. Assim, algumas feições da variação natural não foram bem representadas principalmente nas áreas do talude onde era de se esperar uma maior continuidade espacial paralela à linha de costa. Mas foi possível observar algumas feições claras na plataforma, em áreas mais próximas à costa. Observando o mapa da distribuição da concentração de clorofila integrada (figura 16) é possível observar bem às áreas de maior concentração próximo à costa do Rio de Janeiro, que podem estar associadas não somente às Águas Costeiras, com maior aporte de nutrientes da drenagem continental, mas também às correntes de ressurgência. Estas são provocadas pela intrusão da ACAS (Água Central do Atlântico Sul), que trás águas mais frias do fundo e organicamente enriquecidas à superfície do mar nas áreas próximas à costa. Este fenômeno é comum nesta região e época do ano, de verão e provoca um aumento na produtividade do ambiente (GAETA, 1994; KAMPEL, 2003).

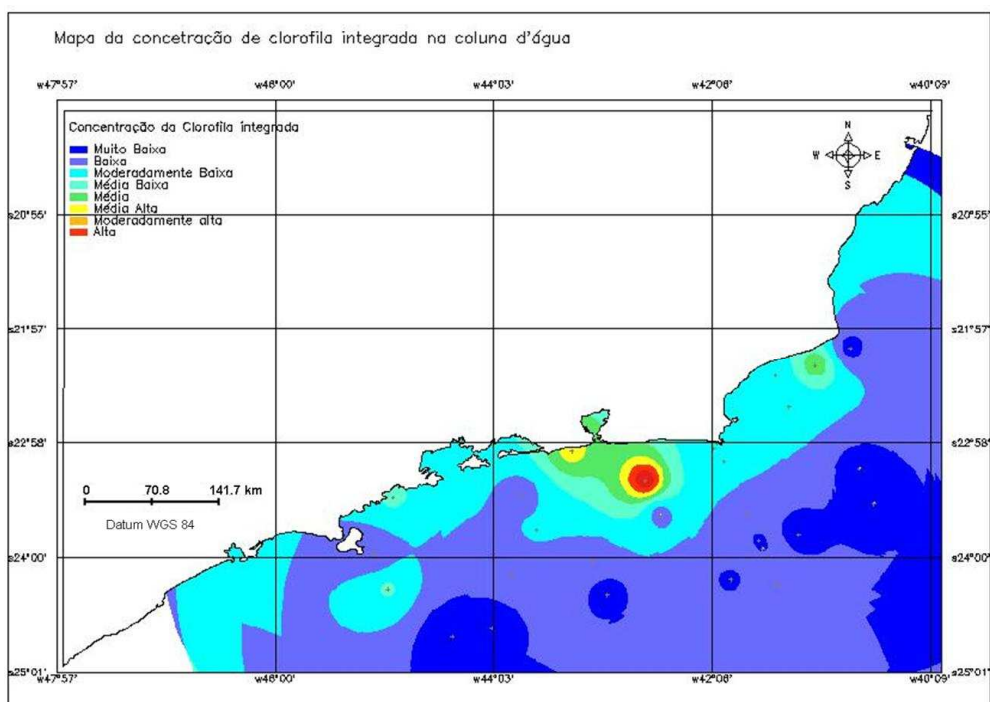


Figura 16: Mapa da distribuição da concentração da clorofila integrada na coluna d'água com interpolação por média ponderada.

O mapa de distribuição da clorofila integrada estimada gerado pelo modelo de regressão linear por álgebra de mapas está na figura 17. Este modelo apresentou padrões de distribuição semelhantes à distribuição da clorofila medida *in situ*, no entanto com valores bastante superestimados em quase toda área. Apenas em um ponto foi subestimado, que é o ponto à nordeste da Baía de Guanabara (RJ).

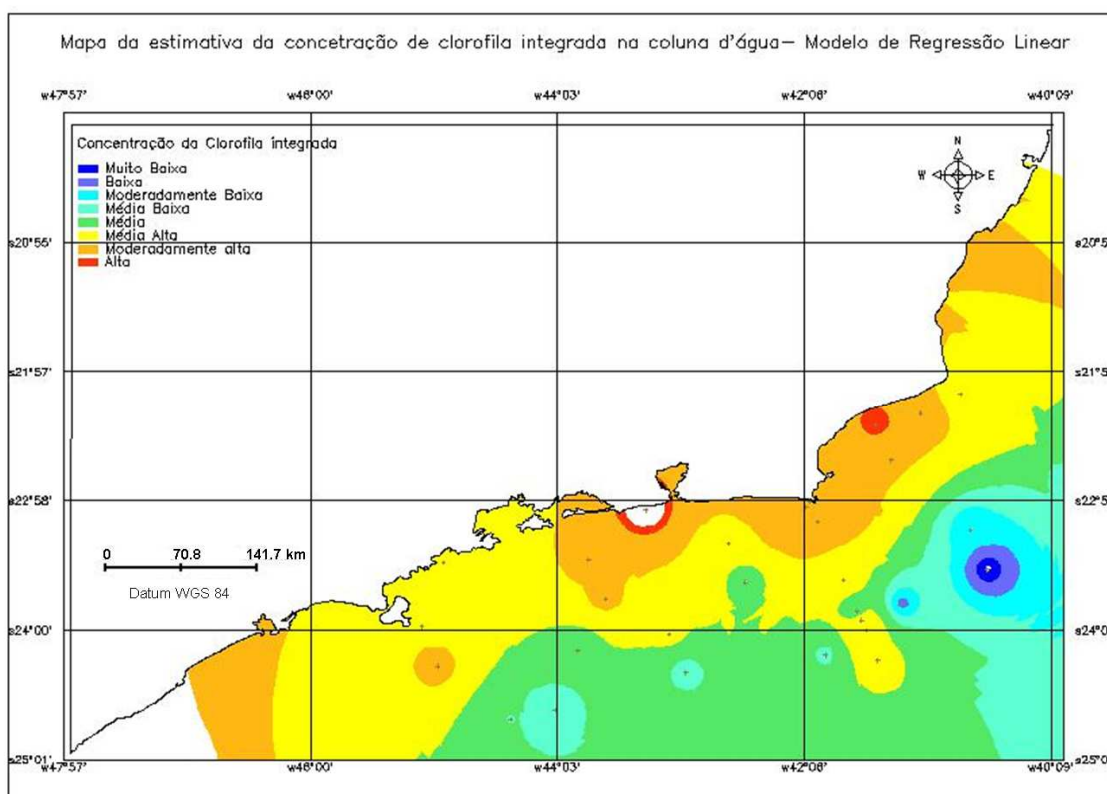


Figura 17: Mapa da distribuição da concentração da clorofila integrada estimada pelo modelo de regressão linear múltiplo proposto por Leão et al (2007).

O mapa de distribuição da clorofila integrada estimada gerado pelo modelo APH_Fuzzy_1 (figura 18) também apresentou valores superestimados da clorofila, mas dessa vez bem mais próximos à distribuição das medidas *in situ* (figura 16).

O mapa de distribuição da clorofila integrada estimada gerado pelo modelo APH_Fuzzy_2 (figura 19) teve uma distribuição bastante semelhante ao modelo AHP_Fuzzy_1. A maior diferença foi que na região do talude a estimativa foi mais condizente com a distribuição das medidas *in situ* (figura 16).

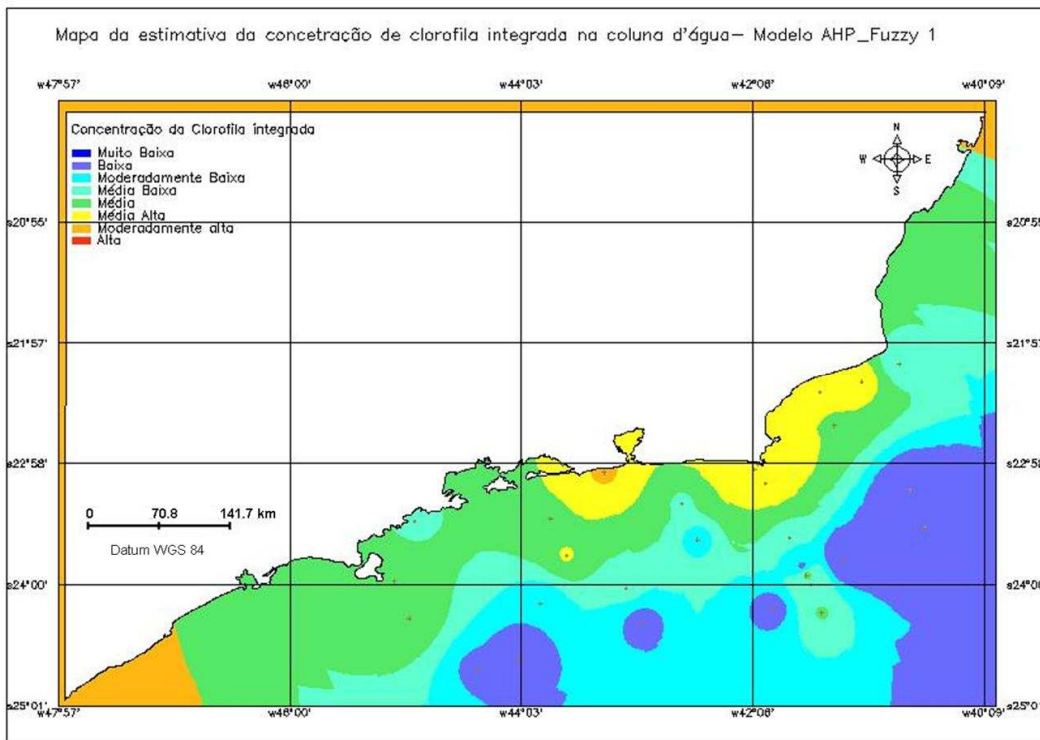


Figura 18: Mapa da distribuição da concentração da clorofila integrada estimada pelo modelo AHP_Fuzzy_1.

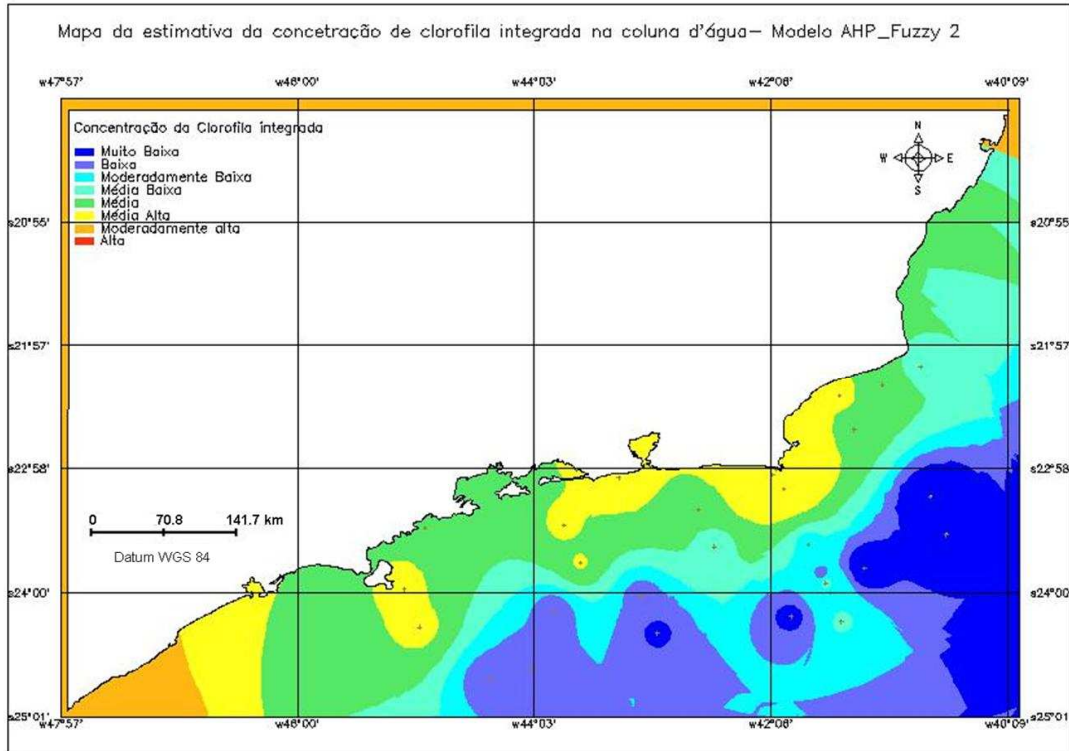


Figura 19: Mapa da distribuição da concentração da clorofila integrada estimada pelo modelo AHP_Fuzzy_2.

Ainda que os modelos não tenham conseguido valores mais próximos dos reais medidos *in situ* como observado pelos mapas, superestimando a maior parte da área, eles apresentaram um padrão de distribuição bastante semelhante. Este padrão demarcou bem a área da costa do Rio de Janeiro sob influência da AC (Água Costeira) e pelas correntes de ressurgência promovida pela ACAS (Água Central do Atlântico Sul), ambas organicamente enriquecidas e com maiores valores de concentração de clorofila integrada.

6 CONCLUSÕES

O método de interpolação utilizado neste trabalho foi mais simplista de média ponderada, pois os dados não permitiam o uso da krigeagem ordinária. Mas como o maior foco no trabalho foi de testar diferentes modelos de estimativa usando lógica fuzzy e booleana, a interpolação por média ponderada pode ser razoavelmente bem aplicada.

A distribuição do modelo de regressão múltipla proposto por Leão et al. (2007), utilizando álgebra de mapas, superestimou bastante a concentração de clorofila integrada em quase toda área. Os modelos AHP_Fuzzy apresentaram distribuições mais próximas daquela obtida por medidas *in situ*, mas ainda foram superestimadas. O modelo AHP_Fuzzy_2 apresentou uma distribuição melhor para a área do talude. Este modelo foi o que teve melhor desempenho da estimativa da concentração de clorofila integrada com as comparações qualitativas dos mapas temáticos.

Os modelos de distribuição das variáveis oceanográficas e da concentração de clorofila integrada na coluna d'água propostos neste trabalho representaram razoavelmente bem alguns processos que ocorrem na área de estudo durante o período analisado, como a ocorrência das correntes de ressurgência no litoral do Rio de Janeiro. Os modelos de estimativa por lógica fuzzy e processo analítico apresentaram um desempenho bom para análise desses processos, com base nas comparações qualitativas dos mapas temáticos.

As múltiplas funcionalidades do programa SPRING que permitem desde a análise exploratória dos dados, análise da variação espacial à modelagem, foram bem aplicadas para o desenvolvimento deste trabalho. Foi possível conhecer a natureza das variáveis com as amostras utilizadas, e seu comportamento espacial na área de estudo. Estes procedimentos auxiliaram na escolha dos interpoladores e na interpretação dos processos analisados. O geoprocessamento mostrou ser uma ferramenta de grande potencial para o desenvolvimento de modelos de estimativa que podem ser muito melhor trabalhos e elaborados em trabalhos futuros.

7 RECOMENDAÇÕES

Para trabalhos futuros de desenvolvimento deste tema de estudo recomenda-se:

- aplicar modelos de interpolação por krigagem paramétrica e considerar a anisotropia das variáveis na área de estudo;
- aplicar a correlação espacial no modelo de regressão linear múltiplo;
- testar diferentes modelos separando as áreas da plataforma e do talude e inserindo outras variáveis oceanográficas;
- elaborar um modelo de estimativa para o período de inverno;
- avaliar quantitativamente o desempenho dos modelos com os dados estimados.

8 REFERÊNCIAS BIBLIOGRÁFICAS

CÂMARA, G.; DAVIS, C.; MONTEIRO, A. M.; D'ALGE, J. C. **Introdução à ciência da geoinformação**. São José dos Campos, INPE, 2001 (2ª edição, revista ampliada, disponível em www.dpi.inpe.br/gilberto/livro).

BURROUGH, P.A.; MCDONELL, R.; **Principles of Geographical Information Systems**. Oxford, Oxford University Press, 1998.

FUCKS, S.; CARVALHO, M.S.; CÂMARA, G.; MONTEIRO, A.M. **Análise Espacial de Dados Geográficos**. Brasília, Embrapa, 2004 (ISBN: 85-7383-260-6). (Edição em papel: EMBRAPA, Brasília, 2004, disponível em www.dpi.inpe.br/gilberto/livro)

GAETA, A.S. **Oceanografia Biológica: Plâncton - Estado de São Paulo**. In: J. L. Valentin e A. S. Gaeta (ed). Diagnóstico Ambiental Oceânico e Costeiro das Regiões Sul e Sudeste do Brasil. São Paulo, p.46-119, 1994.

GEMITZI A., TSIHRINTZIS V. A., VOUDRIAS E., PÉTALAS, C., STRAVODIMOS, G. Combining geographic information system, multicriteria evaluation techniques and fuzzy logic in siting MSW landfills. **Environmental Geology**. 51: 797–811, 2007.

KAMPEL, M., **Estimativa da produção primária e biomassa fitoplanctônica através de sensoriamento remoto da cor do oceano e dados in situ na costa sudeste brasileira**. Tese de doutorado. Instituto Oceanográfico da Universidade de São Paulo. 2003.

KRIEGER, S. **Análise de dados do projeto Coroas – larga escala**. Ed. Universidade de São Paulo. 2005.

LEÃO, C.; KAMPEL, M.; LORENZZETTI, J.A.; GAETA, S.A. Modelos de regressão linear múltipla para estimar a concentração de Clorofila-a integrada na zona eufótica na costa sudeste brasileira. Anais **COLACMAR**. 2007.

PEREIRA G. C., EVSUKOFF, A. G., COUTINHO R, EBECKEN N. F. F. Coastal Environmental Management by Fuzzy System Modeling. **FUZZ_IEEE**. Budapest Hungary, 2004.

PLATT, T.; JASBY, A.D. The relationship between photosynthesis and light for natural assemblages of coastal marine phytoplankton. **Journal of Phycology**, v.12, p. 412-430, 1976.

9 ANEXOS

Programas das operações desenvolvidas em LEGAL

Modelo Regressão linear para estimativa da concentração de clorofila integrada na coluna água

```
{
// Modelo Regressão linear para estimativa da
//clorofila integrada na coluna água

//Declaração
Numerico n, tsm, ineu, zt ("Amostras");
Numerico ceureg ("CEU");

//Instanciação
n = Recuperar ( Nome= "N_mp" );
tsm = Recuperar ( Nome= "TSM_mp" );
ineu = Recuperar ( Nome= "INeu_mp" );
zt = Recuperar ( Nome= "zt_mp" );

ceureg = Novo ( Nome = "Ceureg" , ResX=1000, ResY=1000, Escala=500000, Min =
0, Max = 1 );

//Operação
ceureg = 4.591 - 0.002*n - 0.033*zt - 0.033*tsm - 0.336*ineu ;
}
```

Modelo de função fuzzy para a frequência de Brunt-Vaisalla (N^2)

```
{
// Fuzzy N (ponto de N max 109.8779 e alfa 0.000214) modelo senoidal

//Declaração
Numerico n ("Amostras");
Numerico nfuzzy ("N_Fuzzy");

//Instanciação
n = Recuperar ( Nome= "N_mp" );
nfuzzy = Novo (Nome = "N_Fuzzy", ResX=1000, ResY=1000, Escala=500000, Min=0,
Max=1);
```

```
//operação
nfuzzy = (n < 7.5563)? Numerico (0) : (n > 109.8779)? Numerico (1) : ( 1/ (1+
(0.000214 * ((n - 109.8779)^2) ) ) );
}
```

Modelo de função fuzzy para a temperatura superficial do mar (TSM)

```
{
// Fuzzy TSM (ponto de TSM max 27.0421 e alfa 0.116166) modelo senoidal

//Declaração
Numerico tsm ("Amostras");
Numerico tsmfuzzy ("TSM_Fuzzy");

//Instanciação
tsm = Recuperar ( Nome= "TSM_mp" );
tsmfuzzy = Novo (Nome = "TSM_Fuzzy", ResX=1000, ResY=1000, Escala=500000,
Min=0, Max=1);

//operação
tsmfuzzy = (tsm < 17.5266)? Numerico (0) : (tsm > 27.0421)? Numerico (1) : ( 1/ (1+
(0.116166 * ((tsm - 27.0421)^2) ) ) );

}
```

Modelo de função fuzzy para o inverso da concentração de nitrato integrado na coluna d'água (INeu)

```
{
// Fuzzy INeu (ponto de INeu max 0.3714 e alfa 73.60386) modelo senoidal

//Declaração
Numerico ineu ("Amostras");
Numerico ineufuzzy ("INeu_Fuzzy");

//Instanciação
ineu = Recuperar ( Nome= "INeu_mp" );
ineufuzzy = Novo (Nome = "INeu_Fuzzy", ResX=1000, ResY=1000, Escala=500000,
Min=0, Max=1);

//operação
ineufuzzy = (ineu < 0.0092)? Numerico (0) : (ineu > 0.3714)? Numerico (1) : ( 1/ (1+
(73.60386 * ((ineu - 0.3714)^2) ) ) );

}
```

```
}
```

Modelo de função fuzzy para a profundidade da termoclina (Zt)

```
{
```

```
// Fuzzy zt (ponto de zmcs max 35 e alfa 0.003906) modelo senoidal
```

```
//Declaração
```

```
Numerico zt ("Amostras");
```

```
Numerico zfuzzy ("zt_Fuzzy");
```

```
//Instanciação
```

```
zt = Recuperare ( Nome= "zt_mp" );
```

```
ztfuzzy = Novo (Nome = "zt_Fuzzy", ResX=1000, ResY=1000, Escala=500000,  
Min=0, Max=1);
```

```
//operação
```

```
ztfuzzy = (zt < 8)? Numerico (0) : (zt > 35)? Numerico (1) : ( 1/ (1+ (0.003906 * ((zt -  
35)^2) ) ) );
```

```
}
```

Modelo de função fuzzy para a profundidade da concentração máxima de clorofila na zona eufótica (Zmcs)

```
{
```

```
// Fuzzy zmcs (ponto de zmcs max 97.92 e alfa 0.000432) modelo senoidal
```

```
//Declaração
```

```
Numerico zmcs ("Amostras");
```

```
Numerico zmcsfuzzy ("zmcs_Fuzzy");
```

```
//Instanciação
```

```
zmcs = Recuperare ( Nome= "zmcs_mp" );
```

```
ztfuzzy = Novo (Nome = "zt_Fuzzy", ResX=1000, ResY=1000, Escala=500000,  
Min=0, Max=1);
```

```
//operação
```

```
zmcsfuzzy = (zmcs < 1)? Numerico (0) : (zmcs > 97.92)? Numerico (1) : ( 1/ (1+  
(0.000432 * ((zmcs - 97.92)^2) ) ) );
```

```
}
```


Modelo da estimativa da concentração de clorofila integrada por média ponderada com os pesos calculados pelo método AHP - modelo 1

```
// Pesos a serem aplicados
// INeu_Fuzzy = 0.158
// N_Fuzzy = 0.061
// TSM_Fuzzy = 0.097
// zt_Fuzzy = 0.684

// Razao de consistência
// CR = 0.031

// Definicao dos dados de entrada
Numerico var1 ("INeu_Fuzzy");
Numerico var2 ("N_Fuzzy");
Numerico var3 ("TSM_Fuzzy");
Numerico var4 ("zt_Fuzzy");

// Definicao do dado de saida
Numerico var5 ("AHP_Fuzzy");

// Recuperacao dos dados de entrada
var1 = Recuperare (Nome="INeu_Fuzzy");
var2 = Recuperare (Nome="N_Fuzzy");
var3 = Recuperare (Nome="TSM_Fuzzy");
var4 = Recuperare (Nome="zt_Fuzzy");

// Criacao do dado de saida
var5 = Novo (Nome="AHP_Fuzzy", ResX=1000, ResY=1000, Escala=500000,
Min=0, Max=1);

// Geracao da media ponderada
var5 = 0.158*var1 + 0.061*var2+ 0.097*var3+ 0.684*var4;
}
```

Modelo da estimativa da concentração de clorofila integrada por média ponderada com os pesos calculados pelo método AHP - modelo 2

```
{
// Pesos a serem aplicados
// INeu = 0.076
// TSM = 0.050
// zmcs = 0.528
// zt = 0.347

// Razao de consistência
```

```

// CR = 0.028

// Definicao dos dados de entrada
Numerico var1 ("INeu_Fuzzy");
Numerico var2 ("TSM_Fuzzy");
Numerico var3 ("zmcs_Fuzzy");
Numerico var4 ("zt_Fuzzy");

// Definicao do dado de saida
Numerico var5 ("<AHP_Fuzzy_2>");

// Recuperacao dos dados de entrada
var1 = Recupere (Nome="INeu_Fuzzy");
var2 = Recupere (Nome="TSM_Fuzzy");
var3 = Recupere (Nome="zmcs_Fuzzy");
var4 = Recupere (Nome="zt_Fuzzy");

// Criacao do dado de saida
var5 = Novo (Nome=" AHP_Fuzzy_2", ResX=1000, ResY=1000, escala=500000,
Min=0, Max=1);

// Geracao da media ponderada
var5 = 0.076*var1 + 0.050*var2+ 0.528*var3+ 0.347*var4;
}

```