

**Modelo de regressão na estimativa de sólidos em suspensão por meio de imagens multiespectrais TM-Landsat 5 e CCD-CBERS 2B -
Estudo de caso: Planície de inundação do Alto Rio Paraná**

Luiz Henrique da Silva Rotta¹
Nilton Nobuhiro Imai¹
Monique Sacardo Ferreira¹
Tiago Matsuo Samizava¹
Paulo César Rocha¹
Evlyn Márcia Leão de Moraes Novo²

¹Universidade Estadual Paulista – UNESP
Faculdade de Ciências e Tecnologia
Rua Roberto Simonsen, 305 - 19060-900 - Presidente Prudente - SP, Brasil
luizhrotta@yahoo.com.br; nnimai@fct.unesp.br;
mosacardo@hotmail.com; tiagosamizava@yahoo.com.br; pcrocha@fct.unesp.br;

²Instituto de Nacional de Pesquisas Espaciais – INPE
Av. dos Astronautas, 1758– 12.227-010 – São José dos Campos, SP, Brasil
evlyn@ltd.inpe.br

Abstract. It is very important to know the spatial distribution of water component concentrations, such as suspended solids, but for that, a large number of samples spread on the water body are needed, what makes the data collection time and cost consuming. Thus the use of remote sensing to predict water component concentration has been attempted since the launch of the early satellite missions. Building on the previous efforts, this paper presents the results of using regression models to estimate the concentration of suspended solids in two sets of images: TM-Landsat 5 and CCD-CBERS 2B. The images were acquired concurrently to a ground mission during which, water samples were collected, preserved and sent to laboratory analyses for determination of suspended solid concentration. TM-Landsat image was converted into reflectance and corrected for atmospheric effects and the average ground reflectance was acquired for a 3 x 3 pixels window for both data sets. The regression of image variables against suspended solid concentration allowed to derived two models which were applied to the images. The paper compares the results showing that CCD-CBERS 2B was able to get suitable information for this estimation.

Palavras-chave: remote sensing, floating solids, regression model, sensoriamento remoto, sólidos suspensos, modelo de regressão.

1. Introdução

A atividade humana tem um papel essencial no que se refere à qualidade da água. Como a água doce é um recurso que não pode ser ampliado naturalmente e a sua demanda sofre um crescimento expressivo com o aumento da população e dos meios de produção, é de vital importância que o homem tenha consciência da necessidade de um gerenciamento racional dos recursos hídricos disponíveis (PEREIRA FILHO, 2000).

Nesse contexto é fundamental conhecer a distribuição espacial da qualidade da água, a qual está correlacionada com o desenvolvimento e manutenção da vida em ambientes aquáticos. O conhecimento sobre a distribuição espacial de componentes da água, como os sólidos suspensos, embora fundamental para a compreensão de seus processos geomorfológicos e ecológicos é de difícil obtenção.

Para a obtenção informações sobre a concentração de sólidos em suspensão são necessárias medidas de campo, as quais, dependendo do tamanho do sistema, envolve um grande número de elementos amostrais para a obtenção de previsões com acurácia adequada. Assim o uso de sensoriamento remoto na caracterização espacial de sólidos em suspensão torna-se uma boa alternativa, pois os sedimentos em suspensão possuem a propriedade de

retro-espalhar a radiação incidente em diferentes proporções ao longo do espectro eletromagnético, tornando-se possível assim correlacionar a concentração de sedimentos em suspensão presente em corpos d'água com a reflectância da água medida por sensores remotos (BARBOSA, 2005).

A água como objeto de estudo é um alvo complexo, pois na natureza ela pode conter diversos elementos em seu volume, com respostas espectrais distintas. A presença de sedimentos, Clorofila, matéria orgânica e outros elementos da natureza imprimem na água uma resposta espectral diferenciada em cada situação (PEREIRA FILHO, 2000).

A reflectância da superfície do corpo de água pode ser estimada através da imagem de satélite. Alguns fatores naturais alteram a composição físico-química da água e, conseqüentemente, sua resposta espectral. Os principais fatores que afetam a coloração e, assim, a detecção dos habitats aquáticos pelos métodos de sensoriamento remoto são: a presença de plantas aquáticas, sólidos em suspensão, nível de clorofila e a transparência da água (PEREIRA et al., 2007).

Um aspecto importante relacionado ao sensoriamento remoto é a interferência da atmosfera nos dados registrados. A radiação refletida e espalhada pelo solo, nuvens e aerossol é devolvida parcialmente ao espaço, e captada por sensores a bordo de um satélite. Dependendo do comprimento de onda de observação, a radiação recebida é influenciada por diferentes componentes atmosféricos. (CEBALLOS, 2007).

Os objetivos deste trabalho foram: 1) desenvolver um modelo empírico baseado na técnica de regressão múltipla, para estimar a concentração de sólidos em suspensão da área de estudo por meio de imagens de satélite de média resolução espacial TM-Landsat 5 e CCD-CBERS 2B; 2) aplicar este modelo de regressão na imagem para obtenção do mapa da distribuição espacial de TSS (Total de Sólidos em Suspensão); 3) comparar os resultados obtidos pelos diferentes sensores.

2. Área de Estudo

A área de estudos está localizada no alto curso do rio Paraná, entre os estados de Mato Grosso do Sul e Paraná, na foz do rio Paranapanema e Ivinheima (Figura 1). Trata-se de uma área em que a dinâmica hidrológica vem sendo influenciada diretamente pela construção de um conjunto de barragens hidroelétricas a montante, que têm reduzido os pulsos de inundação na planície fluvial.

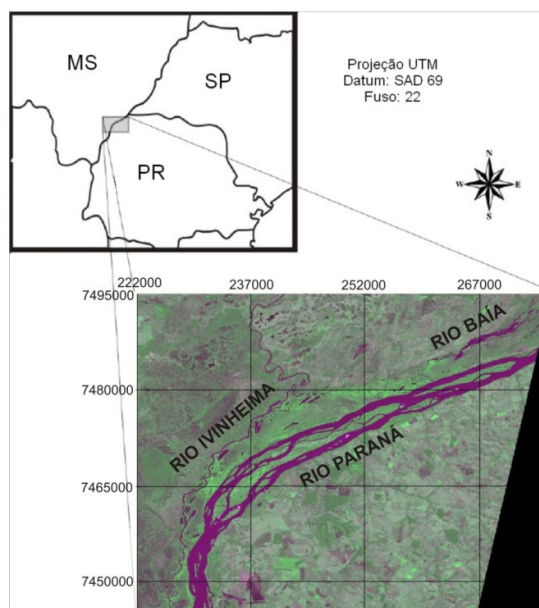


Figura 1. Área de estudo: Alto Rio Paraná.

3. Procedimentos Metodológicos

3.1. Coleta de Dados em Campo e Análise Laboratorial

É de fundamental importância para o presente trabalho a etapa de coleta de dados em campo, a qual necessita ser executada simultaneamente à tomada da imagem multiespectral com a qual os dados levantados devem ser associados. Foram utilizadas imagens TM-Landsat 5 por ser distribuída gratuitamente, pelos sensores estarem estáveis e permitirem análises de correlação, bem como pelo fato de que essas imagens TM estavam disponíveis nas datas necessárias para o desenvolvimento do presente trabalho de pesquisa. Foi utilizado também

imagens CCD-CBERS 2B tomadas poucos dias antes do levantamento de campo, de forma a viabilizar uma análise comparativa entre as inferências obtidas pelos dois produtos.

A coleta de dados em 37 estações amostrais (vide Figura 2) foi feita nos dias 19, 20, 21 e 22 de julho de 2008, e a data de aquisição da imagem Landsat ocorreu no dia 20 de julho de 2008 e a CBERS no dia 15 de julho de 2008. Essas datas representam o período de águas baixas, em que a influência dos pulsos de inundação, característicos dessas áreas, é mínima. O delimitamento amostral foi realizado baseado em Samizava et al., 2008

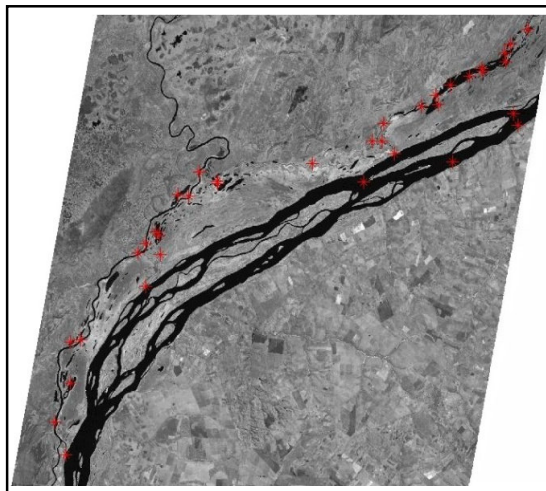


Figura 2. Distribuição espacial dos elementos amostrais

A análise dos sólidos em suspensão na água coletada em cada ponto foi realizada de acordo com o procedimento proposto pela Associação Brasileira de Normas Técnicas - ABNT/ NBR 10664/Abr1989. Dessa maneira foi possível obter valores da concentração em mg/l do TSS (Total de Sólidos em Suspensão), SSF (Sólidos Suspensos Fixos) e SSV (Sólidos Suspensos Voláteis).

3.2. Processamento dos dados de Sensoriamento Remoto

3.2.1. Imagens TM-Landsat 5

A imagem TM-Landsat datada de 20/07/08 foi georreferenciada por meio do aplicativo Spring adotando como referência a imagem GeoCover 2000 Landsat, devido a ausência de outra fonte de dados georreferenciados.

Para a realização da correção atmosférica da imagem TM-Landsat utilizou-se produtos gerados pelo sensor MODIS/Terra de 20 de julho de 2008 para a estimativa de parâmetros atmosféricos. Os dados referentes à espessura óptica, ozônio e vapor d'água foram gravados no formato tif e importados no software ENVI 4.4 para registrar parte da cena MODIS com a cena TM-Landsat. Feito isso, foi realizada uma análise estatística dos pixels dos segmentos das imagens MODIS, os quais foram processados numa planilha de cálculo e obtidos os valores de parâmetros atmosféricos para a correção radiométrica dos dados TM-Landsat. Para realizar a correção dos efeitos atmosféricos adotou-se o aplicativo SCORADIS de Zullo Jr. (1994), o qual aplica o modelo de transferência 5S (Simulação do Sinal do Satélite dentro do Espectro Solar) (TANRÉ et al., 1990).

Para executar a rotina SCORADIS são fornecidos os dados atmosféricos e, assim, calcular a imagem de reflectância real. Como a rotina está previamente configurada com os parâmetros de calibração radiométrica das bandas espectrais do sensor TM-Landsat, a rotina gera uma imagem com correção atmosférica já transformada em reflectância real de superfície.

Dessa imagem de reflectância real foram extraídos valores amostrais de reflectância nas mesmas posições onde foram coletadas as amostras utilizadas para medir a concentração de sólidos suspensos. Para isso adotou-se uma janela de 9 pixels (3x3) cujo centro possui as coordenadas de um dos elementos amostrais. O valor médio do fator de reflectância correspondente a cada elemento amostral levantado no campo foi calculado e, numa planilha de cálculo do MiniTAB, foi estimada a correlação do valor de reflectância real de cada banda (TM1, TM2, TM3 e TM4) com o valor da variável TSS.

3.2.2. Imagens CCD-CBERS 2B

Adquiriu-se a imagem do sistema sensor CCD/CBERS 2B referente ao dia 15 de julho de 2008 no website do INPE (<http://www.dgi.inpe.br/CDSR/>). A imagem foi georreferenciada como descrito para a imagem do TM-Landsat

Sabe-se da importância da calibração radiométrica de imagens multiespectrais para análise temporal. Parâmetros de calibração do sensor bem como variáveis atmosféricas podem ser usadas com o intuito de transformar os níveis de brilho registrados em reflectância real. Entretanto, optou-se por não utilizar os valores de calibração de pré-lançamento e também por não realizar a correção atmosférica da cena CCD-CBERS e, assim, utilizar os dados digitais registrados em lugar de valores de reflectância real devido à incerteza inerente à aplicação desses parâmetros de calibração do sensor, os quais não foram verificados após o sistema ter sido colocado em funcionamento. Além disso, como o objetivo deste trabalho não envolve o uso do modelo de regressão ajustado em outra cena (outra data), mas a avaliação da habilidade do sistema sensor em registrar dados adequados para inferência da concentração de TSS, não há necessidade de realizar a estimativa de valores de reflectância real, com risco de perder parte das informações de interesse.

No presente trabalho, o fato de haver valores constantes multiplicados ou adicionados aos valores de reflectância real não deve perturbar as correlações entre a concentração de sólidos suspensos e a grandeza radiométrica alterada, igualmente, em todos os pixels da cena. Por outro lado, transformações inadequadas podem conduzir a perda de informações relacionadas com a reflectância volumétrica de corpos de água.

Assim, após o georreferenciamento as imagens foram importadas no ENVI para realizar a leitura do nível de cinza registrado para cada elemento amostral. Foram gravados os níveis de cinza obtidos em uma janela 3x3 pixels na qual o pixel central corresponde à coordenada do ponto. Com os nove valores de cada elemento amostral foi calculada uma média final de nível de cinza em cada uma das quatro bandas.

3.3. Análise estatística dos dados

Primeiramente foi verificada a normalidade do conjunto de dados referente à variável TSS, visto que um requisito para ajuste de modelos de regressão é a normalidade da variável resposta (Y). O teste de normalidade utilizado foi o de Anderson-Darling.

A Análise de regressão é uma coleção de ferramentas estatísticas usadas para estimar parâmetros num modelo de regressão. Então, esse modelo ajustado de regressão é tipicamente usado na previsão de observações futuras ou para estimar a resposta média em um nível particular (MONTGOMERY e RUNGER, 2003).

As variáveis que apresentaram maior correlação foram submetidas a uma análise de regressão, na qual a banda espectral foi considerada a variável independente e o TSS foi considerado variável resposta. Obteve-se assim uma equação matemática que representa o comportamento destes dados além de um coeficiente de determinação (R^2) - que representa a porcentagem da variância que o modelo consegue explicar. Foi feita uma análise de variância utilizando o teste F de Snedecor para verificar se o coeficiente angular da função ajustada é igual a zero. Também foi realizada uma análise sobre os resíduos gerados, verificando-se a normalidade dos dados e aplicando-se o teste de Durbin-Watson para verificar sua homocedasticidade (variância constante) ou grau de autocorrelação dos resíduos. A normalidade e a variância constante dos resíduos ou aleatoriedade são pressupostos requeridos na análise de regressão.

Para a validação de cada um dos modelos foram excluídos, aleatoriamente, pontos utilizados no ajuste dos mesmos e gerados modelos, neste trabalho, denominados “modelos de validação”. Aplicou-se cada modelo de validação nos pontos excluídos para geração desse modelo e comparou-se o valor inferido com o valor real. Além disso, verificou-se se o valor

inferido e o valor real encontravam-se no intervalo de predição do modelo de validação ajustado.

Para comparar os modelos de validação ao modelo de regressão obtido com toda amostra, foram analisados os valores dos coeficientes β_0 e β_1 , sendo que o β_0 é o coeficiente linear da função e β_1 é o seu coeficiente angular. Dessa forma, essa comparação dos coeficientes β_0 e β_1 de cada modelo de validação baseou-se na verificação de que esses valores encontram-se no intervalo de confiança dos coeficientes do modelo original. O intervalo de confiança adotado foi de 95%.

Finalizando, aplicou-se as equações de regressão geradas nas imagens TM-Landsat 5 de 20 de julho de 2008 e CCD-CBERS 2B de 15 de julho de 2008. Para tanto, utilizou-se o aplicativo Idrisi. A partir dos mapas da distribuição espacial da concentração de TSS analisou-se os resultados obtidos.

4. Resultados e Discussão

Com a metodologia exposta anteriormente obteve-se os seguintes resultados de sólidos suspensos na etapa laboratorial do trabalho (Tabela 1).

Tabela 1. Estimativa de sólidos suspensos dos elementos amostrais.

	Valor Mínimo	Valor Máximo	Média	Variância
TSS (mg/l)	0,6	26,2	6,95	41,42
SSF (mg/l)	0,0	20,0	4,86	26,85
SSV (mg/l)	0,2	6,2	2,08	1,99

Foi necessário se aplicar uma transformação logarítma na variável dependente TSS para que a mesma tivesse sua distribuição normalizada. Vale ressaltar que alguns pontos considerados outlier foram excluídos do modelo devido à influência das margens do rio no resultado da resposta espectral dos elementos amostrais. Com a variável resposta normalizada, efetuou-se a análise de correlação (Tabela 2).

Tabela 2. Correlação entre Log(TSS) e as bandas do TM-Landsat e do CCD-CBERS.

	TM1	TM2	TM3	TM4	CCD1	CCD2	CCD 3	CCD 4
Log(TSS)	0.066	0.583	0.923	0.880	-0.399	0.050	0.794	0.129

Conforme observado, obteve-se maior correlação entre o logaritmo de TSS e a banda 3 dos sensores. Essas variáveis apresentam um alto grau de relacionamento e foram submetidas à análise de regressão para geração dos modelos. A equação matemática que explica o relacionamento de tais dados bem como os resultados dos testes aos quais esses modelos foram submetidos são apresentados na Tabela 3.

Tabela 3. Análise de regressão de log(TSS) em TM3 e CCD3.

Equação	$\log TSS = -0.7321 + 24.2 \cdot TM3$	$\log TSS = -1.9838 + 0.08392 \cdot CCD3$
P valor (β_0)	0.000	0.000
P valor (β_1)	0.000	0.000
R ²	85.3%	63.1%
F	150.6	42.77
DW	1.84310	1.67733

Os coeficientes β_0 e β_1 foram bem ajustados (*p valores* = 0.000), ou seja, os coeficientes são significativos. O coeficiente de determinação elevado, seguindo o alto coeficiente de correlação, indica que o modelo de regressão estimado explica 85.3% e 63.1% dos dados utilizando as bandas TM3 e CCD3, respectivamente. A análise de variância da regressão forneceu um valor de F maior que o F_{tab} . Com relação à análise de independência dos resíduos o valor do teste de Durbin Watson encontrado foi maior que o d_u tabelado, o que indica que os

resíduos não são autocorrelacionados, determinando a aprovação do modelo em relação a esse teste. A Figura 3 apresenta o ajuste dos dados em relação aos modelos lineares obtidos.

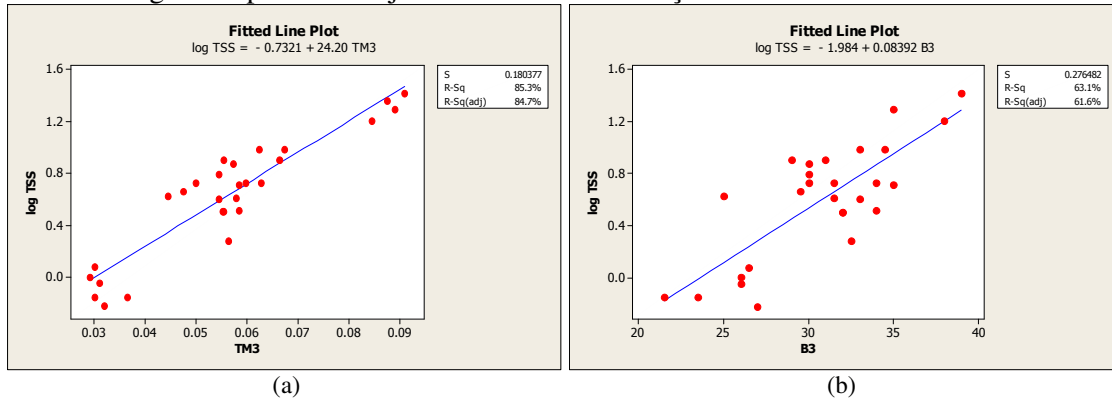


Figura 3. Ajuste do modelo de regressão para as variáveis TM3 x Log TSS (a) e CCD3 x Log TSS (b).

Para validar os modelos gerados em relação ao TM3 e CCD3, excluiu-se aleatoriamente 6 elementos amostrais para cada modelo de validação gerado. Foram utilizados 5 modelos de validação. O valor real de log TSS de todos os elementos amostrais foram contemplados nos Intervalos de Predição dos modelos de validação. A partir disto, fez-se a análise do Intervalo de Confiança dos coeficientes β_0 e β_1 para finalizar a etapa de validação dos modelos. Como os valores de β_0 e β_1 dos modelos de validação estavam no intervalo de confiança dos coeficientes do modelo original, validou-se os modelos de regressão de TSS na área de estudos a partir de imagens TM3/Landsat 5 e CCD-CBERS 2B.

Mapas de classes de concentração de sólidos suspensos foram obtidos pela aplicação dos modelos de regressão validados e posterior fatiamento (Figura 4 e 5).

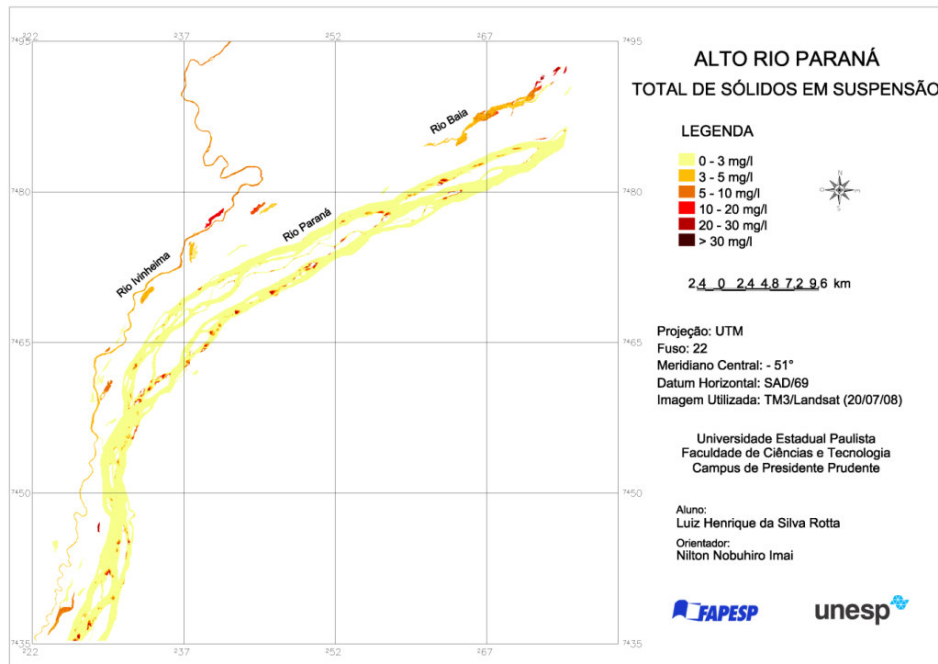


Figura 4. Mapa de concentração de TSS a partir de imagem TM/Landsat.

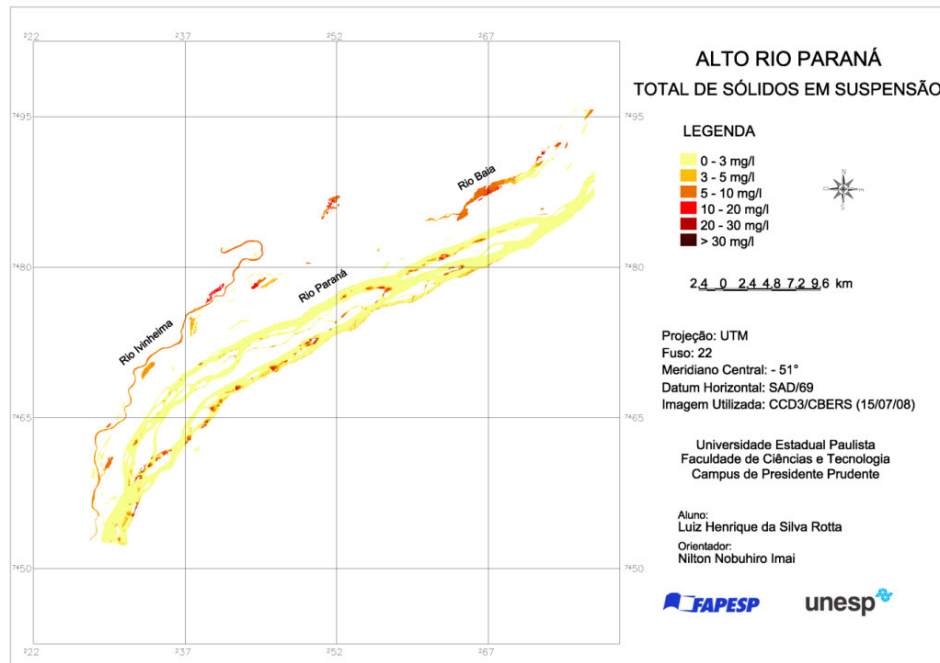


Figura 5. Mapa de concentração de TSS a partir de imagem CCD/CBERS 2B.

Confrontado-se as duas representações, percebe-se que o rio Paraná apresenta as menores concentrações de TSS em ambos os casos, entretanto nota-se a presença de áreas que alcançam concentrações entre 5 e 10 mg/l. Esse fato está associado à áreas de baixa profundidade do rio, ocasionadas pela presença de bancos de areia, que interferem na radiação captada pelo sistema sensor. A presença de tais bancos de areia foi verificada na campanha de campo.

As regiões com maior concentração de TSS verificado na planície através das imagens multiespectrais, foi a lagoa Finado Raimundo (pertencente ao Rio Ivinheima), e em algumas regiões do Rio Baía, chegando a atingir valores na faixa de 20-30 mg/l.

O ambiente que obteve a maior diferenciação na concentração de TSS entre as cartas obtidas por meio dos sensores TM e CCD foi o Rio Baía. Observando-se a imagem obtida pelo TM, nota-se que o Rio Baía apresenta pouca variação de classes, com um aspecto mais suavizado, sendo sua maioria pertencente às classes “3-5 mg/l” e “5-10 mg/l”. Na imagem obtida a partir do sensor CCD, observa-se uma maior heterogeneidade de classes, contemplando em sua maioria classes de “0-3 mg/l” e “5-10 mg/l”.

5. Conclusões

Ante o exposto, constatou-se o potencial dos sensores de média resolução espacial CCD-CBERS 2B e TM-Landsat 5 em registrar variações da concentração do Total de Sólidos em Suspensão. Estas imagens mostraram-se adequadas para geração de um modelo de inferência da concentração do TSS presentes no sistema fluvial do alto rio Paraná. Espera-se que esse modelo possa ser aplicado em cenários de diferentes datas, visto que o método empregado para o ajuste e validação do modelo mostrou-se eficiente na inferência da variável limnológica na área de estudo, considerando que os períodos apresentem características hidrológicas e climáticas semelhantes aos da data de amostragem

Embora não tenha sido realizada calibração radiométrica da cena CCD/CBERS 2B adverte-se que a mesma é de fundamental importância para que o modelo possa ser aplicado em outras datas. A necessidade da correção atmosférica decorre do fato de que os dados adquiridos por sensores remotos orbitais tratam da radiação solar que interage com a atmosfera até atingir o alvo e que alcança o sensor interagindo novamente com a atmosfera.

Quando eliminados os efeitos da atmosfera, as imagens apresentam-se num mesmo referencial, reflectância de superfície, sendo possível a realização de análises temporais. Apesar dos parâmetros estatísticos indicarem que o modelo ajustado com a imagem TM-Landsat 5 é melhor do que aquele obtido com os dados CCD-CBERS 2B, os resultados de TSS estimados por meio do sensor CCD-CBERS foram considerados aceitáveis. Uma análise dos mapas de TSS resultantes de ambos os sensores mostra que a distribuição espacial da variável limnológica em questão foi bastante similar.

Assim, a imagem TM-Landsat 5, apesar do longo tempo de funcionamento desse sistema, e a imagem CCD-CBERS 2B mostraram-se adequadas para o monitoramento ambiental dos ambientes aquáticos da planície de inundação do alto rio Paraná. Esse tipo de aplicação possibilita redução de custos de monitoramento e pode contribuir para o entendimento da hidrodinâmica do sistema, uma vez que é possível gerar uma representação computacional da distribuição do TSS nos diversos ambientes da planície (lagoas, rios, canais, ressacos).

6. Agradecimentos

À Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP pela concessão de bolsa de estudos. Ao projeto CT - Hidro/CNPq “Hidrodinâmica, ritmo climático e zoneamento de áreas alagáveis na planície fluvial do alto Rio Paraná, Centro-Sul do Brasil” e ao Programa Nacional de Cooperação Acadêmica PROCAD/CAPES por viabilizarem a coleta de dados em campo. Aos alunos da disciplina Comportamento Espectral de Alvos do Programa de Pós Graduação em Sensoriamento Remoto do INPE, pela cooperação no trabalho de campo. Agradeço também à Prof^a Dr^a Vilma Mayumi Tachibana pela ajuda referente aos assuntos estatísticos abordados neste trabalho.

7. Referências Bibliográficas

- ABNT - Associação Brasileira de Normas Técnicas: NBR 10664. 1989. Águas-Determinação de Resíduos Sólidos- método gravimétrico. Rio de Janeiro. 14p.
- BARBOSA, C. C. F. **Sensoriamento remoto da dinâmica da circulação da água do sistema planície de Curuai/rio Amazonas.** 2005. 258p. Tese (Doutorado em Sensoriamento Remoto) – Instituto Nacional de Pesquisas Espaciais – São José dos Campos.
- CEBALLOS, J. C. Produtos para atmosfera – Modis Atmosphere, INPE. **O sensor Modis e suas aplicações ambientais no Brasil/** Organizadores: RUDORFF, B.F.T.; SHIMABUKURO, Y. E.; CEBALLOS, J.C. São José dos Campos: A. Silva Vieira Ed., 2007.
- MONTGOMERY, D. C.; RUNGER, G. C. **Estatística aplicada e probabilidade para engenheiros.** Rio de Janeiro: Livros Técnicos e Científicos Editora S.A., 2003.
- PEREIRA FILHO, W. **Influência dos diferentes tipos de uso da terra em bacias hidrográficas sobre sistemas aquáticos da margem esquerda do reservatório de Tucuruí – Pará.** 2000. 136p. Tese (Doutorado em Geografia Humana) – Universidade de São Paulo, São Paulo.
- PEREIRA, R.; SCHÄFER, A. E.; WÜRDIG, N. L. Aplicabilidade de métodos de sensoriamento remoto na avaliação e monitoramento do estado trófico de lagoas costeiras do Rio Grande do Sul (RS). In: XIII Simpósio Brasileiro de Sensoriamento Remoto, 2007, Florianópolis, p. 3513-3520. **Anais.**
- SAMIZAVA, T. M.; IMAI, N. N.; ROTTA, L. H. S; FERREIRA, M. S.; GALO, M. L. B. T; ROCHA, R. R. A.; ENNES, R. Proposta de delineamento amostral para levantamento de medidas de variáveis limnológicas e de dados espectrorradiométricos em planície de inundação. In: Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação, II, 2008, Recife. **Anais.**
- TANRÉ, D.; DEROO, C.; DUHAUT, P.; HERMAN, M.; MORCRETTE, J.J. Description of a computer code to simulate the satellite signal in the solar spectrum: the 5S code. **International Journal of Remote Sensing.** v. 11, n. 4, p. 659- 668, 1990.
- ZULLO JR., J. **Correção atmosférica de imagens de satélite e aplicações.** 1994. 204p. Tese (Doutorado em Engenharia Elétrica). Universidade Estadual de Campinas, Campinas.