

O desenvolvimento de ferramentas de predição de balneabilidade baseadas em níveis de precipitação: estudo de caso da praia de Cachoeira das Emas (SP)

The development of predictive tools for recreational water quality based on rainfall levels: case study of Cachoeira das Emas beach (SP)

Fabio Muller Hirai¹, Monica Ferreira do Amaral Porto²

RESUMO

Este artigo apresenta resultados de um estudo realizado na praia Cachoeira das Emas, Rio Mogi-Guaçu, Pirassununga (SP), onde foram desenvolvidos modelos estatísticos simples de regressão linear utilizando níveis de precipitação como variáveis explanatórias, por meio da correlação positiva observada entre concentração de indicadores fecais na água e precipitação acumulada em 24, 48 e 72 horas anteriores à amostragem no rio. Os dados aplicados na calibração (período de 2009 a 2011) e validação (período de 2012) das ferramentas obtidas são oriundos das redes de monitoramento pluviométrico do Departamento de Águas e Energia Elétrica (DAEE) e de balneabilidade de água doce da Companhia Ambiental do Estado de São Paulo (Cetesb). Como essas redes possuem objetivos distintos, nem todos os eventos de precipitação que ocorreram durante o período de estudo foram relacionados com um valor de concentração de indicador fecal na água, e vice-versa. Logo, as equações de regressão obtidas ficaram baseadas em um baixo número de observações. Os resultados sugerem que há potencial de aplicação prática desse tipo de ferramenta, porém requer-se monitoramento específico e conjunto dos parâmetros do modelo a fim de obter uma quantidade significativa de dados para a calibração e a validação da ferramenta. A aplicação desse recurso de gerenciamento de balneabilidade seria ideal para um período específico do ano com maior afluência de pessoas, em praias cuja principal fonte de poluição seja difusa e consequente do escoamento de drenagem. Teoricamente, seria possível empregar outras variáveis ambientais ao modelo além dos níveis de precipitação, tais como turbidez e vazão, desde que os parâmetros fossem monitorados com a qualidade da água.

Palavras-chave: gerenciamento; qualidade da água; recreação; modelagem.

ABSTRACT

This article presents the results of a case study conducted on Cachoeira das Emas beach, at Mogi-Guaçu River, in Pirassununga (SP). Using precipitation levels as explanatory variables, simple statistical models based on linear regression were developed. Positive correlation between the concentration of faecal indicators in water and accumulated rainfall in 24, 48 and 72 hours, prior to sampling on the river, were observed. Data sets used for calibration (2009-2011 period) and validation (period of 2012) of the predictive tools were originally acquired from distinct monitoring networks: rainfall levels data by the Departamento de Águas e Energia Elétrica (DAEE) and recreational water quality data by the Companhia Ambiental do Estado de São Paulo (CETESB). Due to different goals of these monitoring networks, not every rainfall event during the study period was related to a concentration value of fecal indicator in the water, and vice versa, so the regressions obtained were based on a low quantity of observations. The results indicate that there is a potential for applying the predictive tool, but it requires monitoring the model parameters in conjunct, in order to obtain the amount of data necessary to calibration and validation. Ideally, the use of this recreational water management tool would be during a certain period of the year, when there are more people at the beach site, on locations where the main source of pollution is diffuse, resulting from runoff drainage. It would also be possible to add other environmental variables to the models, such as flow and turbidity, provided that they are monitored with water quality conditions.

Keywords: management; water quality; recreation; modeling.

¹Mestre em Engenharia Civil pela Universidade de São Paulo (USP) - São Paulo (SP), Brasil.

²Doutora em Engenharia Civil pela USP - São Paulo (SP), Brasil.

Endereço para correspondência: Fabio Muller Hirai - Rua Antonio Bastos, 265/12 - Vila Bastos - 09040-220 - Santo André (SP), Brasil - E-mail: fabiohirai@yahoo.com

Recebido em: 26/02/14 - **Aceito em:** 01/04/16 - **Reg. ABES:** 131249

INTRODUÇÃO

O uso da água para recreação em ambientes naturais pode trazer uma série de perigos àqueles em contato direto com a água (WHO, 1999), sobretudo a exposição a microrganismos patogênicos eventualmente presentes no ambiente aquático. Tradicionalmente, a avaliação das condições de balneabilidade é feita pelo monitoramento da concentração de microrganismos indicadores fecais na água. No Brasil, a regulamentação que define os critérios de avaliação e classificação de balneabilidade dos corpos hídricos é a Resolução do Conselho Nacional do Meio Ambiente (Conama) n.º 274, de 2000 (BRASIL, 2001).

Nas principais técnicas analíticas de cultura microbiológica, hoje bastante consolidadas e empregadas (membrana filtrante e tubos múltiplos), os resultados demoram cerca de 24 a 48 horas para serem apontados (USEPA, 2010). Logo, as classificações de balneabilidade obtidas ficam baseadas no histórico da qualidade microbiológica da água de semanas anteriores. Assim, praias cujos resultados retrospectos são ruins estariam mais sujeitas à poluição (situação imprópria) em comparação àquelas cujos resultados retrospectos apontam boa qualidade da água (situação própria).

As ferramentas de predição têm como objetivo estimar as condições de balneabilidade no instante presente por meio do uso de modelos que possivelmente variam muito em relação à complexidade de construção e operação. Modelos estatísticos podem ser fundamentados em regressão linear simples ou múltipla e empregam parâmetros ambientais que estão correlacionados com a qualidade da água como variáveis explanatórias, tais quais: níveis de precipitação, vazão, altura de ondas e turbidez.

Este artigo apresenta os resultados de um estudo realizado na praia de Cachoeira das Emas, no Rio Mogi-Guaçu, município de Pirassununga (SP). O estudo teve como propósito desenvolver ferramentas de predição baseadas em modelos estatísticos, utilizando como variáveis explanatórias os níveis de precipitação registrados por postos pluviométricos próximos da praia, na bacia hidrográfica do Rio Mogi-Guaçu, na Unidade Hidrográfica de Gerenciamento de Recursos Hídricos (UGRHI) 9.

METODOLOGIA

Empregaram-se conceitos de modelos estatísticos de regressão múltipla e da ferramenta *rain threshold levels*, descrita pela *United States Environmental Protection Agency* (USEPA, 2010). Dados da concentração de coliformes termotolerantes na praia de Cachoeira das Emas, em UFC.100 ml⁻¹, foram obtidos da rede de monitoramento semanal de balneabilidade de água doce da Companhia de Tecnologia de Saneamento Ambiental (Cetesb). As séries de dados de precipitação diária, em mm.dia⁻¹, foram obtidas da rede

de monitoramento do Departamento de Águas e Energia Elétrica (DAEE) do estado de São Paulo, para várias estações pluviométricas na UGRHI da praia de estudo, a de número 9. Em razão da sua enorme variabilidade, todos os dados microbiológicos relativos à concentração de coliformes termotolerantes na água foram transformados em valores de logaritmos na base 10 antes de qualquer análise.

Devido à disponibilidade de dados, o período de estudo ficou restrito aos anos de 2009 a 2012. Para desenvolvimento e calibração das ferramentas, consideraram-se dados de 2009 até 2011. Dados de 2012 foram usados para a validação das ferramentas.

Experiências internacionais (BOEHM *et al.*, 2007; FRANCY & DARNER, 2006) mostraram que modelos estatísticos de predição produzem melhores resultados quando sua aplicação é subdividida em temporadas seca e chuvosa, pois a dinâmica dos patogênicos é influenciada de maneira diferente. Sendo assim, foram desenvolvidas ferramentas para dois períodos do ano, separando os dados respectivos:

- período seco — de abril a setembro;
- período chuvoso — de outubro a março.

A correlação entre precipitação e balneabilidade foi analisada nesses períodos separadamente, tendo em vista os níveis de precipitação acumulados em 24, 48 e 72 horas anteriores à data de amostragem de água no rio. Para cada estação pluviométrica examinada, foram definidas as seguintes séries de parâmetros:

- P_{24} — precipitação acumulada em 24 horas antes da data de amostragem de água;
- P_{48} — precipitação acumulada em 48 horas antes da data de amostragem de água;
- P_{72} — precipitação acumulada em 72 horas antes da data de amostragem de água.

Parearam-se os dados microbiológicos com cada série de precipitação acumulada, tendo como referência a data de amostragem da água do Rio Mogi-Guaçu. Foram calculados os coeficientes de correlação linear de Pearson (R) e construídos os diagramas de dispersão dos parâmetros. Pelo fato de as ferramentas propostas serem baseadas na ocorrência de chuvas, o cálculo dessas correlações desconsiderou os pares de valores cuja variável de precipitação possuísse valor menor que 1 mm. Estações pluviométricas cujos dados não apresentassem níveis de correlação significantes entre nenhum parâmetro de precipitação acumulada e a concentração de indicadores na água não foram levadas em conta. O nível de significância adotado em todas as análises estatísticas foi de 95%. Isto é, resultados estatísticos significantes possuíram valor p menor que 0,05.

O modelo estatístico foi estabelecido conforme a Equação 1:

$$Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n + E \quad (1)$$

Em que:

Y = variável dependente, é a concentração de indicadores fecais a ser estimada na água;

b_0 = intercepto do modelo;

b_n = coeficiente de regressão;

X_n = variáveis explanatórias, são os níveis de precipitação;

E = variação randômica ou erro.

Os erros foram assumidos como independentes e igualmente distribuídos (distribuição normal), com média zero e desvio padrão.

Na construção das equações de regressão, as precipitações acumuladas significantes foram subdivididas em seus níveis diários, de forma a distinguir os dias em que ocorreu chuva, tendo em vista que podem ter diferentes influências sobre a qualidade da água. Por exemplo, um nível acumulado P_{48} igual a 10 mm pode ser igual a diversas combinações da soma da precipitação de dois dias anteriores à amostragem, como 4 e 6 mm ou 9 e 1 mm. Essa subdivisão foi feita da seguinte maneira:

- $P_{24} = P_{d-1}$ — em que P_{d-1} é o nível de precipitação no primeiro dia anterior à amostragem;
- $P_{48} = P_{d-1} + P_{d-2}$ — em que P_{d-2} é o nível de precipitação no segundo dia anterior à amostragem;
- $P_{72} = P_{d-1} + P_{d-2} + P_{d-3}$ — em que P_{d-3} é o nível de precipitação no terceiro dia anterior à amostragem.

Os parâmetros P_{d-1} , P_{d-2} e P_{d-3} foram, portanto, empregados como variáveis explanatórias na Equação 1 (dados de entrada da ferramenta), enquanto o logaritmo na base 10 da concentração de indicador fecal na água foi a variável resposta do modelo. Assim, nos casos em que P_{24} fosse significativa, apenas a variável P_{d-1} se aplicaria em regressão linear simples, enquanto nos casos em que P_{48} e P_{72} fossem significantes, puderam-se obter regressões lineares múltiplas com P_{d-1} , P_{d-2} ou P_{d-3} (esta última apenas quando P_{72} fosse significativa).

Para cada ferramenta desenvolvida, foi determinado um nível P_m , a precipitação mínima que define a aplicação da ferramenta. Pelo fato de as ferramentas se basearem na ocorrência de chuvas anteriores, elas não se aplicam quando os níveis de precipitação anteriores são nulos ou pouco significantes, isto é, abaixo de P_m . A definição dos valores P_m foi feita por meio da análise dos diagramas de dispersão e dos valores das séries de dados pareados de balneabilidade e precipitação, observando a partir de qual nível de precipitação a relação linear entre as variáveis foi constatada. Trata-se de um critério que

possui certo grau de subjetividade; idealmente, com o valor inicial adotado, o parâmetro P_m deve ser ajustado conforme o emprego da ferramenta, de maneira a obter o melhor desempenho possível ao longo da experiência adquirida.

As estatísticas usadas para avaliar e comparar os diferentes modelos de regressão obtidos foram os valores de R^2 ajustado e a estatística C_p de Mallows. Consideram-se bons modelos com maior valor de R^2 ajustado e menor valor de C_p . Foram utilizados os *softwares* Microsoft Excel e SigmaPlot 9.01 para realizar todas as análises estatísticas — este último efetua as regressões e verifica a conformidade das hipóteses de normalidade e homoscedasticidade (resíduos como variável independente, com média zero e variância constante). Logo, todos os modelos apresentados como resultados estão de acordo com essas hipóteses e são tidos como os melhores entre os diversos possíveis de ser obtidos com as séries de dados disponíveis. Também foram investigadas possíveis colinearidades entre as variáveis explanatórias empregadas na modelagem, de forma a evitar o uso de duas ou mais que apresentassem colinearidade entre si. Para tanto, calcularam-se os fatores de inflação de variância (VIF) de cada variável empregada, em cada ferramenta desenvolvida (valores não são exibidos neste artigo). Nenhum dos modelos selecionados teve colinearidade significativa entre suas variáveis explanatórias.

A metodologia *rain threshold levels* (USEPA, 2010) busca estabelecer níveis de chuva correlacionados com a ocorrência de altas concentrações de bactérias indicadoras fecais na água da praia, de modo a subsidiar a emissão de alertas de risco microbiológico para os usuários. Por sua natureza simples, as ferramentas não visam determinar a predição numérica dos níveis de concentração de bactérias na água, embora as equações de regressão resultem numa estimativa para tais níveis. A resposta da ferramenta é limitada simplesmente a uma predição de condição boa ou má de balneabilidade, baseada na estimativa obtida acima ou abaixo do valor padrão considerado.

O padrão de balneabilidade para concentração de coliformes termotolerantes em amostra única (última amostra de um conjunto) é estabelecido na Resolução Conama n.º 274, de 2000, com valor de 2.500 coliformes termotolerantes por 100 mL, cujo valor de logaritmo na base 10 é igual a 3,398. Portanto, esse é o valor padrão a ser considerado neste estudo. A predição de má qualidade é definida quando o resultado da modelagem ultrapassa esse valor padrão.

Para validação das ferramentas desenvolvidas, as predições obtidas para o ano de 2012 foram comparadas aos valores das análises microbiológicas realizadas. O desempenho das ferramentas foi avaliado de acordo com os possíveis resultados:

- Predição de má condição — a balneabilidade de fato é má (acerto);
- Predição de má condição — a balneabilidade de fato é boa (erro tipo I);

- Predição de boa condição — a balneabilidade de fato é má (erro tipo II);
- Predição de boa condição — a balneabilidade de fato é boa (acerto).

O erro tipo II é o mais crítico do ponto de vista de saúde pública, pois pode expor a população aos perigos microbiológicos existentes e não comunicados. Já o erro tipo I protegeria a população de um risco inexistente, fato que diminuiria a confiabilidade das previsões, mas não colocaria as pessoas em situação de risco.

Descrição da área de estudo

A praia de Cachoeira das Emas localiza-se à margem do Rio Mogi-Guaçu, no município de Pirassununga, inserido na UGRHI 9. As coordenadas geográficas são: Latitude 21°55'32" e Longitude 47°22'06". O Rio Mogi-Guaçu é de classe 2 e nasce no estado de Minas Gerais, no município de Bom Repouso. Sua bacia hidrográfica possui área de drenagem total aproximada em 18.938 km², e só no estado de São Paulo há 15.040 km² de área de drenagem (CBH-MOGI, 2008). O uso do solo à montante da praia é majoritariamente agrícola, sobretudo cultivo de cana-de-açúcar, pastagens e fruticulturas. Há também atividades de silvicultura, associadas às indústrias de papel e celulose, na cidade de Mogi Guaçu e proximidades (CBH-MOGI, 2008).

A série de dados de concentração de indicadores fecais na praia teve início em 26 de agosto de 2009 com intervalos irregulares (aproximadamente uma amostra por semana). O ponto de monitoramento foi denominado MOGU 02351, pela Cetesb. Empregaram-se coliformes termotolerantes como indicador fecal, e os dados microbiológicos foram originalmente expressos em unidades formadoras de colônias por 100 mL, transformados para valores de logaritmo na base 10 antes de ser examinados.

A análise de correlação incluiu dados de níveis pluviométricos diários de dez estações situadas na UGRHI 9, apresentada no Quadro 1. A Figura 1 traz a localização aproximada da praia de Cachoeira das Emas e das estações pluviométricas na UGRHI 9.

RESULTADOS

A Tabela 1 contém os valores dos coeficientes de correlação de Pearson entre a série de dados de concentração de indicadores fecais na praia de Cachoeira das Emas e cada série de precipitação acumulada (24, 48 e 72 h) de cada estação pluviométrica, nos períodos seco e chuvoso. Nenhuma série de dados oriunda das estações C4-019, D3-071 e D3-015 indicou correlação significativa. Sendo assim, essas estações foram desconsideradas do restante do estudo. O passo seguinte foi construir os diagramas de dispersão das correlações significantes verificadas, de acordo com as Figuras 2 e 3.

Os valores P_m foram representados nos diagramas pela linha vertical tracejada e adotados como o nível de precipitação acumulada a partir do qual se observou linearidade entre as variáveis, definindo a aplicação das ferramentas, isto é, a partir de qual magnitude de chuva podiam ser válidas.

Entre os diversos modelos de regressão possíveis de ser desenvolvidos com as variáveis explanatórias (P_{d-1} , P_{d-2} e P_{d-3}) dos vários postos pluviométricos, são apresentados apenas os melhores para cada nível de precipitação acumulada correlacionada significativamente com a concentração de indicadores fecais, tendo em vista a colinearidade entre as variáveis (nenhum modelo demonstra colinearidade significativa) e as medidas estatísticas de eficiência da regressão linear — C_p de Mallows e R^2 ajustado. As ferramentas obtidas são destacadas no Quadro 2, com as respectivas informações.

Quadro 1 – Estações pluviométricas consideradas no estudo.

Nome do posto	Código DAEE	Município	Lat.	Lon.	Altitude (m)	Início de operação	Distância horizontal até a praia (km)
Pirassununga	C4-033	Pirassununga	2159	4.725	670	1939	9,3
Cresciumal	D4-030R	Leme	2.210	4.717	600	1936	28,1
São João da Boa Vista	C3-031	São João da Boa V.	2.157	4.648	740	1943	58,8
Fazenda Santana	D4-029	Araras	2.215	4.723	660	1936	35,9
Fazenda da Glória	D3-071	Sto. Ant. do Jardim	2.209	4.643	900	1972	72,2
Moji-Mirim	D3-008	Mogi Mirim	2.226	4.658	640	1936	70,1
Itapira	D3-015	Itapira	2.227	4.649	660	1936	80,9
Descalvado	C4-071	Descalvado	2.158	4.742	780	1956	35,0
Santa Eudoxia	C4-019	São Carlos	2.147	4.747	620	1959	45,9
Usina Capão Preto	C4-092	São Carlos	2.152	4.748	610	1960	45,7

Fonte: www.sigrh.sp.gov.br

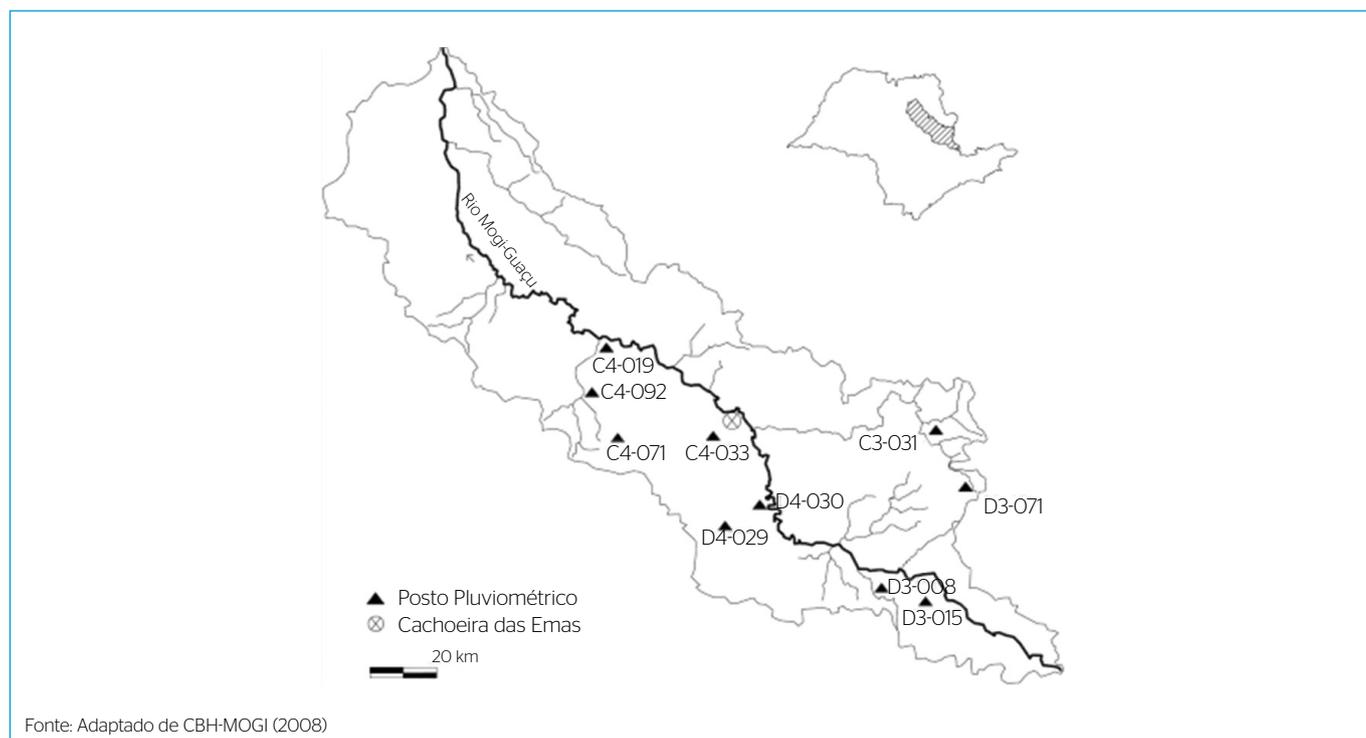


Figura 1 - Localização das estações pluviométricas na UGRHI 9.

Tabela 1 - Coeficientes de correlação linear de Pearson entre os níveis de precipitação acumulada e os valores de logaritmos na base 10 da concentração de coliformes termotolerantes (dados de 2009 a 2011).

Precipitação Acumulada	Estação C4-033			Estação D4-030			Estação C4-071			Estação D4-029			Estação D3-071		
	R	valor p	n												
Período chuvoso (out. – mar.)															
P24	0,366	0,164	16	0,057	0,827	17	0,652	0,003	19	0,604	0,008	18	0,200	0,412	19
P48	0,461	0,018	26	0,586	<0,001	30	0,393	0,039	28	0,529	0,003	30	0,148	0,427	31
P72	0,288	0,110	32	0,529	<0,001	39	0,399	0,013	38	0,425	0,012	34	0,213	0,211	36
Período seco (abr. – set.)															
P24	0,791	0,209	4	-0,595	0,405	4	0,212	0,687	6	0,718	0,108	6	0,426	0,293	8
P48	0,652	0,080	8	0,586	0,167	7	0,224	0,507	11	0,616	0,044	11	0,497	0,144	10
P72	0,571	0,085	10	0,493	0,103	12	0,261	0,390	13	0,500	0,098	12	0,420	0,175	12
Precipitação Acumulada	Estação C4-092			Estação C4-019			Estação C3-031			Estação D3-008			Estação D3-015		
	R	valor p	n												
Período chuvoso (out. – mar.)															
P24	0,429	0,052	21	0,330	0,134	22	0,085	0,707	22	0,423	0,071	19	0,276	0,283	17
P48	0,286	0,141	28	0,315	0,074	33	0,328	0,067	32	0,369	0,045	30	0,141	0,491	26
P72	0,155	0,413	30	0,285	0,087	37	0,319	0,054	37	0,332	0,055	34	0,146	0,440	30
Período seco (abr. – set.)															
P24	0,793	0,060	6	0,458	0,438	5	0,637	0,124	7	0,534	0,354	5	0,270	0,605	6
P48	0,792	0,006	10	0,540	0,086	11	0,694	0,026	10	0,505	0,113	11	0,459	0,115	13
P72	0,557	0,075	11	0,559	0,059	12	0,646	0,017	13	0,489	0,090	13	0,424	0,131	14

Os resultados da validação de cada ferramenta são apresentados a seguir e incluem os índices de ocorrência dos diferentes tipos de erros e das previsões corretas, bem como as porcentagens

de especificidade (proporção de condições boas corretamente previstas) e sensibilidade (proporção de condições más corretamente previstas).

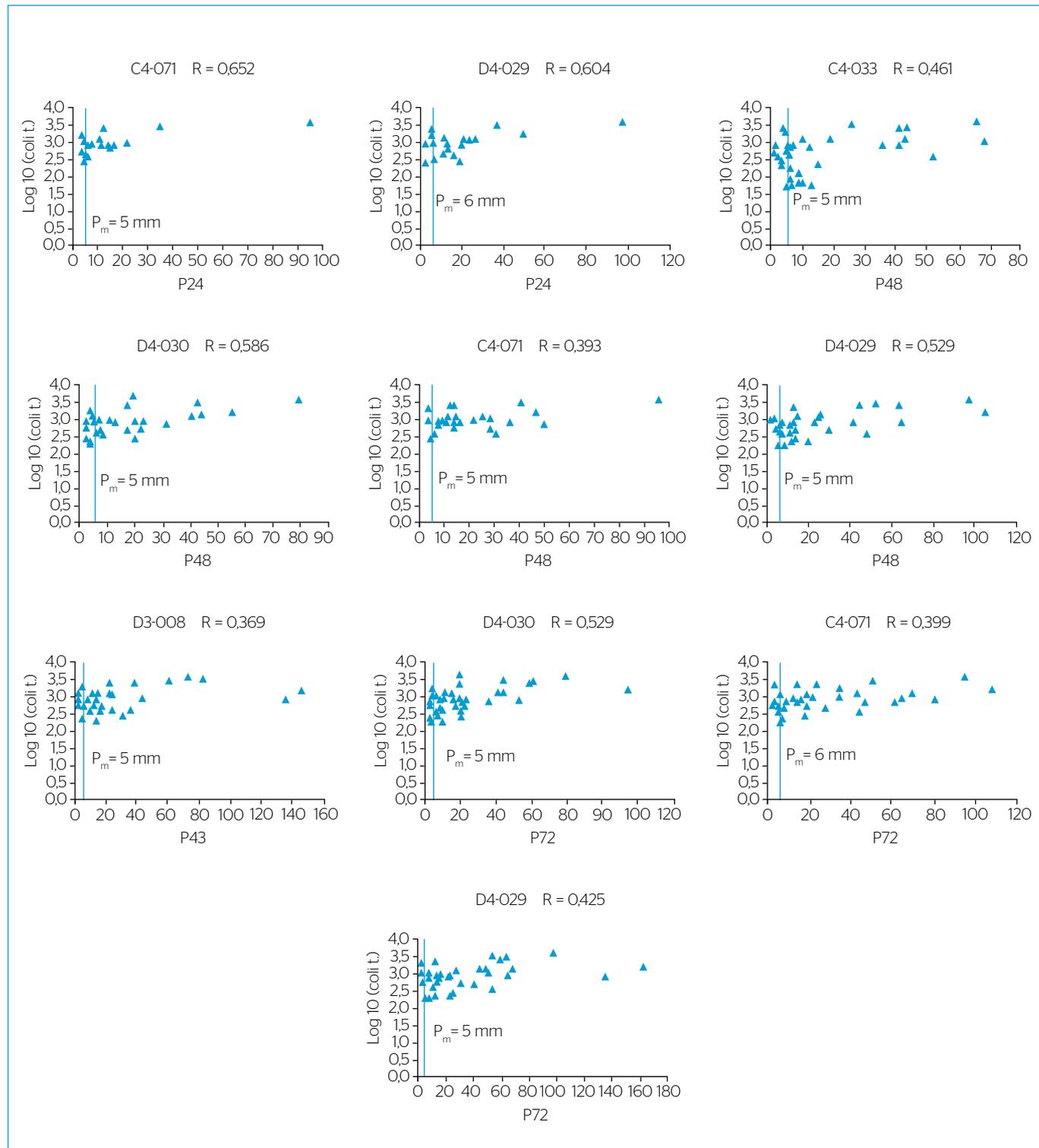


Figura 2 - Diagramas de dispersão entre precipitação acumulada e concentração de coliformes termotolerantes correlacionadas significativamente no período chuvoso.

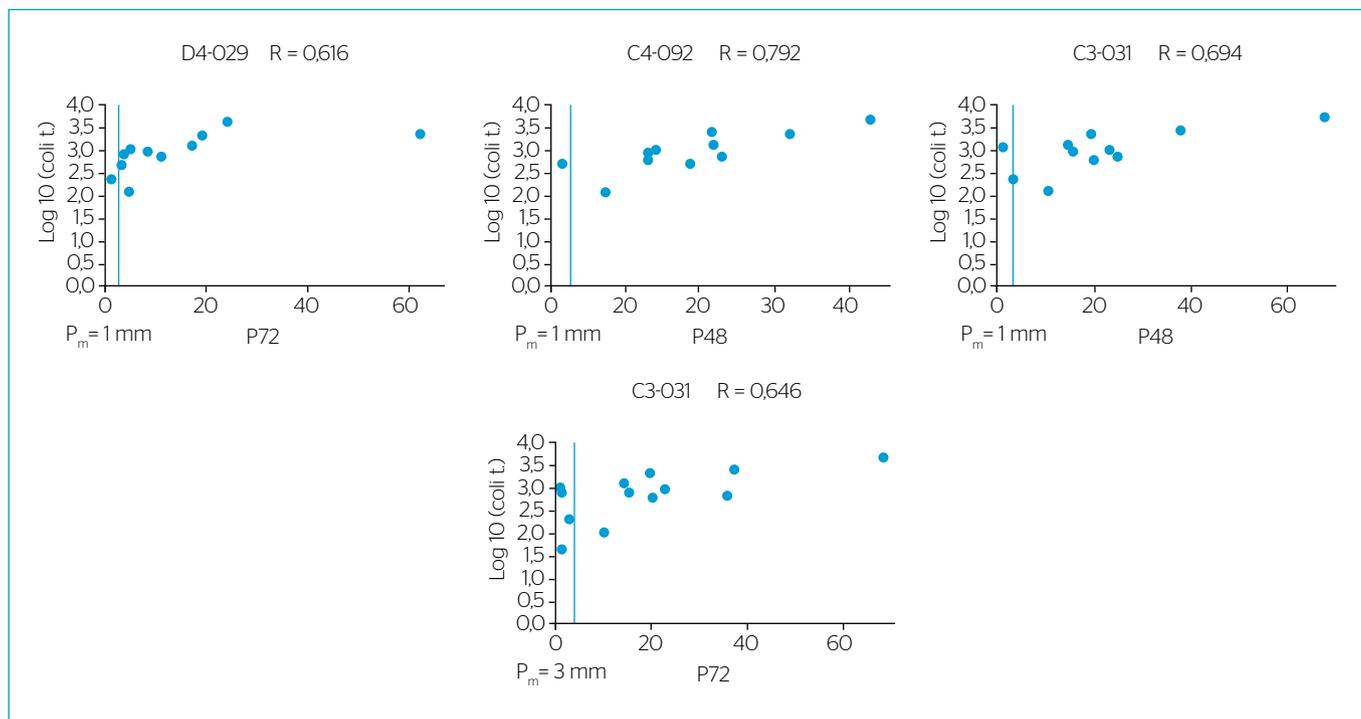


Figura 3 - Diagramas de dispersão entre variáveis de precipitação acumulada e a concentração de coliformes termotolerantes correlacionadas significativamente no período seco.

Quadro 2 - Lista das melhores ferramentas de predição de balneabilidade desenvolvidas.

Ferramenta	Precipitação acumulada significativa	Estação pluvial	Tipo de regressão linear	Variáveis explanatórias do modelo de regressão	n	Pm	R ² ajustado	Cp de Mallow
Período chuvoso								
A	P24	C4-071	simples	Pd-1	14	5	0,458	2,000
B	P48	D4-030	múltipla	Pd-1, Pd-2	12	5	0,836	5,000
		C4-071		Pd-2				
C	P72	D3-008	múltipla	Pd-2	21	5	0,451	4,000
		D4-030		Pd-2, Pd-3		6		
Período seco								
D	P48	C4-092	múltipla	Pd-1, Pd-2	10	1	0,525	3,000
E	P72	C3-031	múltipla	Pd-1, Pd-2, Pd-3	11	3	0,740	4,000

n: número de elementos considerados na calibração do modelo de regressão; Pm: nível de precipitação mínimo aplicável (mm); R: coeficiente de correlação linear de Pearson.

Tabela 2 - Resultados da validação das ferramentas de predição desenvolvidas.

Ferramenta	n	Previsões corretas (%)	Erro tipo I (%)	Erro tipo II (%)	Especificidade (%)	Sensitividade (%)
A	5	60,0	0,0	40,0	100,0	0,0
B	5	60,0	20,0	20,0	75,0	0,0
C	7	71,5	0,0	28,5	100,0	0,0
D	8	87,5	0,0	12,5	100,0	0,0
E	7	85,7	14,3	0,0	83,3	100,0

n: número de ocasiões em que a ferramenta pôde ser validada no ano de 2012.

DISCUSSÃO

De todas as estações pluviométricas consideradas neste estudo, somente algumas apresentaram dados correlacionados com a balneabilidade. As precipitações registradas pelos postos mais distantes da praia não demonstraram correlação significativa com a variação da condição de balneabilidade na praia de Cachoeira das Emas. Todas as correlações significantes ($p < 0,05$) entre os níveis de precipitação acumulada e os valores de logaritmo na base 10 da concentração de coliformes termotolerantes na praia foram correlações positivas, tanto no período seco como no chuvoso, subsidiando a hipótese de que a ocorrência de altos níveis de precipitação em dias anteriores à amostragem de água aumenta a probabilidade de uma amostra ter altos níveis de concentração de indicadores fecais.

Nos diagramas de dispersão foi possível observar, em certos casos, evidente comportamento linear entre os parâmetros modelados, com uma quantidade mínima precipitada (P_m). Os valores de P_m mostraram ser maiores no período chuvoso, sugerindo que, durante esse período, é necessária maior quantidade de chuva para se observar alguma linearidade entre os parâmetros e, portanto, as ferramentas só seriam aplicáveis com chuvas mais intensas, em comparação com o período seco. Quanto maior o nível P_m de uma ferramenta, mais restrita será sua aplicação, pois eventos de precipitação de maior magnitude possuem menor frequência de ocorrência.

O ponto fraco da predição é que ela se obtém com base em dados, os quais alguns consolidam enquanto outros prejudicam a correlação. A análise de regressão requer quantidade suficiente de dados para estabelecer relações robustas entre as variáveis explanatórias e a variável resposta, de forma a definir a confiabilidade das previsões alcançadas pelo modelo (USEPA, 2010). Entretanto as presentes ferramentas foram desenvolvidas e validadas com base em observações limitadas, conforme a disponibilidade dos dados preexistentes. Ambas as redes de monitoramento envolvidas com este estudo possuem objetivos distintos: a rede de monitoramento da Cetesb foi desenvolvida para avaliar a condição de qualidade semanal das praias de acordo com a Resolução Conama n.º 274, de 2000, enquanto a rede de monitoramento do DAEE registra os níveis de precipitação em incrementos diários. Sendo assim, nem todos os dados microbiológicos disponíveis estão pareados com níveis de precipitação anteriores à amostragem, pois em muitas situações simplesmente não aconteceu chuva nas 24, 48 ou 72 horas antes de analisar as concentrações de indicadores fecais na água da praia. Por outro lado, também houve situações em que, após a ocorrência de chuva, não existiu amostragem de água para análise da concentração de indicadores fecais.

Mediante uma extensa análise exploratória, foi possível verificar que a distribuição das séries de dados de precipitação diária de todas as estações pluviométricas mostrou ser assimétrica, com

a maioria de valores nulos e baixos, representando a característica típica para o clima do estado de São Paulo: chuvoso no verão e seco no inverno (esses dados não são apresentados no artigo). Dados microbiológicos pareados com valores de precipitação acumulada nulos, ou abaixo de P_m , não foram incluídos na calibração, pois as ferramentas são baseadas exclusivamente na ocorrência de chuvas; isso também limitou o tamanho do conjunto de dados empregados na modelagem.

O nível das correlações significantes observadas foi maior para o período seco. Por outro lado, no período chuvoso, todas as ferramentas puderam ser calibradas e validadas com maior número de observações, pois se têm mais chuvas nessa época do ano.

O resultado da validação das ferramentas mostrou níveis de previsão correta acima de 60%, porém apenas a ferramenta E apresentou sensibilidade. Ou seja, somente ela conseguiu prever de maneira correta a ocorrência de más condições de balneabilidade, o objetivo primário desse tipo de recurso de gerenciamento. No geral, as ferramentas demonstraram níveis de erro tipo II maiores do que erro tipo I. Especificamente, na praia de Cachoeira das Emas concentrações de indicadores fecais na água acima dos padrões são valores extremos na série de dados de balneabilidade, com baixa frequência de ocorrência. Verificou-se que, ocasionalmente, essas altas concentrações são precedidas por níveis de precipitação acumulada nulos ou abaixo de P_m , e esses dados não são incorporados na calibração das ferramentas. Portanto, fica evidente que outros fatores, além da precipitação e da consequente poluição difusa, podem influenciar as condições de balneabilidade no local.

O projeto de um programa de monitoramento é um componente crucial para um estudo focado numa ferramenta que sirva para a emissão de alertas (USEPA, 2002). Inoportunamente, os dados disponíveis dos parâmetros contemplados não foram coletados da forma sistemática requerida. Assim, as ferramentas obtidas não possuem calibração com grande número de observações. Como a praia também apresentou altas concentrações de indicadores sob condições secas, as ferramentas de predição exclusivamente baseadas em níveis de precipitação talvez não sejam as mais apropriadas para esse caso. Para melhor desenvolver e calibrar ferramentas de predição por meio da metodologia utilizada, o ideal seria coletar e analisar amostras de água após todos os eventos de precipitação durante um determinado período do ano (o qual se pretenda operar a ferramenta, quando houver maior afluência de pessoas à praia), de modo a correlacionar as concentrações microbiológicas com a quantidade precipitada registrada. Segundo Francy *et al.* (2013), é importante coletar dados por anos, de maneira a considerar as variações do clima e das condições de qualidade da água que ocorrem anualmente.

Nesse sentido, haveria maior frequência de monitoramento no período chuvoso do ano, o que significaria maior número de observações para calibrar e validar as ferramentas. Por outro lado, durante o período seco é possível que exista correlação mais forte entre balneabilidade e precipitação, conforme foi observado nas análises dos coeficientes de Pearson. Afinal, com a menor ocorrência de chuvas, tem-se mais tempo para os poluentes se acumularem sob a bacia hidrográfica, caracterizando o escoamento superficial inicial de drenagem com carga mais concentrada de poluição na eventual ocorrência de precipitação (NOVOTNY, 2003). O período chuvoso coincide com o verão e a temporada de férias escolares, quando a tendência é aumentar o número de banhistas nas praias. Além de se obter mais dados, elevar a frequência de monitoramento nesse período pode ser considerada uma ação de gestão a favor da segurança da população usuária.

O monitoramento de precipitação em incrementos horários poderia descrever os eventos de precipitação com maior nível de detalhe, aumentando a precisão e a confiabilidade das ferramentas. Ademais, seria possível utilizar a variável de intensidade da chuva no modelo (USEPA, 2010). De acordo com Ge e Frick (2007), a duração dos eventos também pode ser um fator-chave para explicar as relações entre qualidade da água e precipitação. Nesse âmbito, a estação pluviométrica deveria ser automatizada e a quantidade de dados a ser gerenciada aumentaria significativamente, acarretando maior complexidade em todas as análises.

As informações dos dados bacteriológicos disponíveis são como fotografias instantâneas das concentrações de bactérias na água, o que leva a incertezas no tocante às avaliações de balneabilidade. De acordo com Eleria e Vogel (2005), sabe-se que as tradicionais metodologias analíticas para quantificação de bactérias em amostras de água possuem considerável incerteza intrínseca.

Para melhor explicar a variabilidade das condições de balneabilidade e aumentar a capacidade de predição das ferramentas, podem-se levar em conta outros parâmetros nos modelos de regressão linear múltipla. Em praias localizadas em margens de rios, outras variáveis com potencial de serem empregadas são a turbidez, a vazão e o nível do rio, pois são parâmetros de fácil obtenção (os dados podem ser coletados na ocasião da amostragem) e estão relacionados com cargas de poluição difusa. Como exemplo dessa aplicação, em dez praias da Escócia, ferramentas de predição foram empregadas utilizando dados de precipitação e vazão para prever níveis de densidade de bactérias (MCPHAIL & STIDSON, 2009). Segundo Ge e Frick (2007), a turbidez pode estar associada com a suspensão de sedimentos na água, que podem servir como abrigo para a

sobrevivência de bactérias indicadoras fecais. A vazão dos rios é tipicamente associada aos eventos de chuva e escoamento superficial, podendo ser também um indicador das cargas de poluição. De acordo com USEPA (2010), altas vazões em geral são correlacionadas com altos níveis de bactérias e, de forma análoga, o nível do rio também está associado com a chuva e é com frequência correlacionado com densidades de bactérias ou com a probabilidade de exceção dos padrões de balneabilidade.

O monitoramento de parâmetros adicionais de natureza hidrológica, meteorológica ou de qualidade da água requer maiores investimentos. É preciso analisar os benefícios da inclusão de parâmetros nos modelos de regressão em comparação com o aporte financeiro necessário, de maneira a verificar a viabilidade econômica da ferramenta de predição.

CONCLUSÕES

A investigação da relação entre níveis pluviométricos diários acumulados e a concentração de indicadores fecais na praia de Cachoeira das Emas mostrou existir uma correlação linear positiva, que, em certos casos, pode ser significativa a partir de uma determinada quantidade de precipitação, conforme observado nos diagramas de dispersão. Os resultados deste estudo são ferramentas simples que demonstraram capacidade de predição, ainda que limitadas pela disponibilidade de dados para sua construção. O ponto fraco da predição é que ela é obtida com base em dados, os quais alguns consolidam enquanto outros prejudicam a correlação. O resultado da validação das ferramentas desenvolvidas indicou níveis de previsão correta acima de 60%, porém apenas uma (ferramenta E) apresentou sensibilidade, ou seja, conseguiu prever corretamente a ocorrência de más condições de balneabilidade.

O aporte de bactérias nas praias e a variação de sua concentração ao longo do tempo podem depender de múltiplos fatores. Ferramentas baseadas exclusivamente em níveis de precipitação podem ter desempenho satisfatório apenas em casos particulares, em que a chuva e o escoamento superficial resultante são reconhecidamente os principais fatores que influenciam as condições de balneabilidade, por desencadarem a poluição difusa nos corpos hídricos da bacia.

O desenvolvimento e a aplicação desse tipo de ferramenta demandam recursos e investimentos adicionais para o monitoramento conjunto dos parâmetros modelados, pois, mesmo depois de desenvolvida, a ferramenta requer eventuais calibrações periódicas. É prudente avaliar a viabilidade técnica e econômica da aplicação dessas ferramentas antes de decidir utilizá-las.

REFERÊNCIAS

- BOEHM, A.B.; WHITMAN, R.L.; NEVERS, M.B.; HOU, D.; WEISBERG, S.B. (2007) Nowcasting recreational water quality. In: WYMER, L. J. [Ed.] *Statistical framework for recreational water quality criteria and monitoring*. Wiley-Interscience, Chichester, West Sussex, England.
- BRASIL. Ministério do Meio Ambiente. Conselho Nacional do Meio Ambiente. (2001) Resolução n.º 274, de 29 de novembro de 2000. Dispõe sobre os critérios de balneabilidade em águas. *Diário Oficial da União*: República Federativa do Brasil, Poder Executivo, Brasília, DF, n. 18, Seção 1, p. 70-71, 25 jan. 2001. Disponível em: <<http://www.mma.gov.br/port/conama/legiano.cfm?codlegitipo=3>>. Acesso em: 3 jan. 2011.
- COMITÊ DA BACIA HIDROGRÁFICA DO RIO MOJIGUAÇU - CBH-Mogi. (2008) *Plano da Bacia Hidrográfica 2008-2011*. Faculdade Municipal Professor Franco Montoro, Geosyntec Planejamento e Consultoria, FEHIDRO Contrato n. 092/2007. Mogi Guaçu. 169 p.
- ELERIA, A.; VOGEL, R.M. (2005) Predicting fecal coliform bacteria levels in the Charles River, Massachusetts, USA. *Journal of the American Water Resources Association*, v. 41, n. 5, p. 1195-1209.
- FRANCY, D.S.; DARNER, R.A. (2006) *Procedures for developing models to predict exceedances of recreational water-quality standards at coastal beaches*. 34 p. United States Geological Survey Techniques and Methods 6-B5. Disponível em: <<http://pubs.usgs.gov/tm/2006/tm6b5/>>. Acesso em: 18 out. 2011.
- FRANCY, D.S.; STELZER, E.A.; DURIS, J.W.; BRADY, A.M.G.; HARRISON, J.H.; JOHNSON, H.E.; WARE, M.W. (2013) Predictive models for *Escherichia coli* concentrations at inland lake beaches and relationship of model variables to pathogen detection. *Applied Environmental Microbiology*, v. 79, n. 5, p. 1676-1688.
- GE, Z.; FRICK, W.E. (2007) Some statistical issues related to multiple linear regression modeling of beach bacteria concentrations. *Environmental Research*, v. 103, n. 3, p. 358-364.
- MCPHAIL, C.D.; STIDSON, R.T. (2009) Bathing water signage and predictive water quality models in Scotland. *Aquatic Ecosystem Health & Management*, v. 12, n. 2, p. 183-186.
- NOVOTNY, V. (2003) *Water quality: diffuse pollution and watershed management*. 2. ed. Nova York: John Wiley. 864 p.
- UNITED STATES ENVIRONMENTAL PROTECTION AGENCY - USEPA. (2002) *Time-relevant beach and recreational water quality monitoring and reporting*. Office of Research and Development, National Risk Management Research Laboratory, Cincinnati, Ohio. 49 p. Disponível em: <<http://nepis.epa.gov/Exe/ZyNET.exe/P1007KG7.TXT?ZyActionD=ZyDocument&Client=EPA&Index=2000+Thru+2005&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C00thru05%5Ctxt%5C00000024%5CP1007KG7.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocuments=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=p%7Cf&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL>>. Acesso em: 5 jul. 2011.
- UNITED STATES ENVIRONMENTAL PROTECTION AGENCY - USEPA. (2010) *Predictive tools for beach notification*. Office of Water. Office of Science and Technology. 2 v. Disponível em: <<http://nepis.epa.gov/Exe/ZyNET.exe/P1009IZT.TXT?ZyActionD=ZyDocument&Client=EPA&Index=2006+Thru+2010&Docs=&Query=&Time=&EndTime=&SearchMethod=1&TocRestrict=n&Toc=&TocEntry=&QField=&QFieldYear=&QFieldMonth=&QFieldDay=&IntQFieldOp=0&ExtQFieldOp=0&XmlQuery=&File=D%3A%5Czyfiles%5CIndex%20Data%5C06thru10%5Ctxt%5C00000023%5CP1009IZT.txt&User=ANONYMOUS&Password=anonymous&SortMethod=h%7C-&MaximumDocument=1&FuzzyDegree=0&ImageQuality=r75g8/r75g8/x150y150g16/i425&Display=p%7Cf&DefSeekPage=x&SearchBack=ZyActionL&Back=ZyActionS&BackDesc=Results%20page&MaximumPages=1&ZyEntry=1&SeekPage=x&ZyPURL>>. Acesso em: 22 jul. 2011.
- WORLD HEALTH ORGANIZATION - WHO. (1999) *Health-based monitoring of recreational waters: the feasibility of a new approach (the "Annapolis Protocol")*. Geneva. 50 p. (Protection of the Human Environment: Water, Sanitation and Health Series). Disponível em: <http://www.who.int/water_sanitation_health/bathing/wsh9901/en/>. Acesso em: 13 ago. 2011.