

# SÉRIES TEMPORAIS DAS CONDIÇÕES DE BALNEABILIDADE DA ÁGUA DA PRAIA DA PONTA DA AREIA, ILHA DO MARANHÃO, BRASIL.

**Laurinda Fernanda Saldanha SIQUEIRA (1); Jonas de Jesus Gomes da COSTA NETO (2);  
Mariano Oscar Anibal IBANEZ ROJAS (3)**

(1) Universidade Federal do Maranhão, Av. dos Portugueses s/n São Luis - MA, e-mail: [lau\\_siqueira@yahoo.com.br](mailto:lau_siqueira@yahoo.com.br)

(2) Instituto Federal de Educação Ciência e Tecnologia do Maranhão, Av. Getulio Vargas s/n Monte Castelo São Luis - MA, e-mail: [jonasneto@ifma.edu.br](mailto:jonasneto@ifma.edu.br)

(3) Instituto Federal de Educação Ciência e Tecnologia do Maranhão, Av. Getulio Vargas s/n Monte Castelo São Luis - MA, e-mail: [ibanez@ifma.edu.br](mailto:ibanez@ifma.edu.br)

## RESUMO

O presente trabalho teve como objetivo estudar as tendências e a sazonalidade em series temporais de coliformes totais (NMP/ 100 mL) da água da Praia da Ponta da Areia, Ilha do Maranhão (Brasil) como ferramenta para a gestão sustentável da orla marítima urbanizada, por meio da aplicação de testes de normalidade e transformação Box-Cox nos dados, modelos de regressão, teste de Kruskal-Wallis para verificação da significância e gráficos do tipo Box-Plot para verificação da amplitude da sazonalidade, extração da série temporal e do modelo comportamental puros. Os resultados indicam tendência decrescente nos níveis de coliformes totais e sazonalidade multiplicativa, gerando um modelo explicativo de curto prazo.

**Palavras-chave:** Análise de séries temporais, coliformes totais, Ilha do Maranhão (Brasil).

## 1 INTRODUÇÃO

O Gerenciamento Costeiro Integrado (GCI), de acordo com Cicin-Sain & Knecht (1998), é um processo contínuo e dinâmico por meio do qual decisões são tomadas visando o uso sustentável das áreas costeiras e marinhas e seus recursos. Ainda de acordo com estes autores, a atuação do GCI vincula-se às interações entre os múltiplos usos, tratando-os sistematicamente e partindo de uma escala ampla de entendimento dos problemas.

Nesta pesquisa, foi proposta a análise de séries temporais de qualidade de água para auxílio ao GCI e a sustentabilidade da orla marítima da Ilha do Maranhão (Brasil), funcionando como ferramentas de diagnóstico da evolução e do comportamento e como ferramentas de explicação e predição, preponderantes em gerar informações aos tomadores de decisões e a população. Contudo, é válido salientar, que as séries temporais e os modelos alargam as informações, mas não produzem números inquestionáveis. Tentando simplificar um sistema complexo, estes podem gerar resultados imperfeitos.

No escopo deste trabalho, teve-se o intuito de análise de tendências e sazonalidade em séries temporais de coliformes totais em água (NMP/ 100 mL) da praia da Ponta da Areia, da orla marítima urbana da Ilha do Maranhão (Brasil), buscando extrair conhecimento necessário à sustentabilidade e à gestão da área estudada, e torná-lo inteligível aos tomadores de decisão e população.

## 2 FUNDAMENTAÇÃO TEÓRICA

Séries temporais, de acordo com Morettin & Toli (2006), constituem-se como qualquer conjunto de observações ordenadas no tempo, tendo por finalidade principal a construção de modelos explicativos das mesmas segundo tanto um enfoque temporal quanto um domínio de frequências. O pressuposto básico da análise de séries temporais é de que os fatores que influenciaram o comportamento da série no passado, continuarão a fazê-lo no futuro. Assim sendo, as séries temporais geram concomitantemente modelos explicativos da variabilidade no comportamento das variáveis e da tendência ao longo do tempo,

identificando as componentes aleatórias e ou determinísticas, bem como modelos preditivos de tais características.

As séries temporais podem ter os seguintes objetivos (MORETTIN & TOLOI, 2006): (a) Descrição. Descrever propriedades e comportamentos da série (padrão de tendência, existência de variação sazonal ou cíclica, outliers, alterações estruturais); (b) Explicação. Construir modelos que permitam explicar o comportamento da série no período observado; (c) Controle. Os valores da série temporal medem a qualidade de um processo de manufatura e o objetivo é o controle do processo; e (d) Predição. Predizer valores futuros com base em valores passados, a longo e curto prazo. Aqui se assume que o futuro envolve incerteza, por isso tenta-se reduzir os erros de previsão.

Uma série temporal pode ser decomposta em um conjunto de componentes e ou movimentos (MORETTIN & TOLOI, 2006; MALETZKE, 2009). Neste aspecto, a representação gráfica dos dados sequencialmente ao longo do tempo é fundamental e pode revelar padrões de comportamento importantes. A partir do gráfico temporal, pode-se decompor a série temporal de acordo com tendências, padrões cíclicos (ou sazonalidade) ou componentes aleatórias (ou resíduos).

A tendência compreende o movimento dominante de uma série temporal, isto é, com uma mudança, em um longo intervalo de tempo, ou de longo prazo no nível médio da série temporal. (MALETZKE, 2009). A sazonalidade refere-se a movimentos similares que ocorrem regularmente em um intervalo de tempo fixo (dias, meses, anos), segundo, Chatfield (2004) e Morettin & Toloi (2006), este componente deve ser identificado, e, de acordo com o ensejo da pesquisa, ser removido ou suavizado. O componente aleatório ou resíduo são os movimentos aleatórios ou irregulares correspondentes aos deslocamentos esporádicos de uma série temporal (MALETZKE, 2009).

### 3 METODOLOGIA

A apresentação do trabalho A metodologia proposta consta de adaptações aos métodos indicados nas pesquisas de Groppo (2005) e de Cavalheiro (2003).

**a. Levantamento de dados:** Os dados desejados nesta pesquisa constam de níveis de coliformes totais (NMP/ 100 mL) de trecho da Praia da Ponta da Areia (P1 = 02° 30' 15"S; 44° 19'06.9"W), da Ilha do Maranhão (Brasil), presentes em relatórios de monitoramento da qualidade da água de praias realizados pela Secretaria de Meio Ambiente e Recursos Hídricos do Maranhão (SEMA) e Fundação Nacional de Saúde no Maranhão (FUNASA-MA). As séries utilizadas são mensais correspondentes aos anos de 1989, 2003, 2005, 2006 e 2007, mas de maneira geral os dados foram encontrados de maneira dispersa e muitas vezes incompletos.

**b.Preparação dos dados e obtenção das séries temporais:** Dados foram plotados em função do tempo ou transformados pela transformação Box-Cox, numa tentativa de modificar a série temporal não normalmente distribuída em normalmente distribuída. A opção por esta transformação foi baseada em testes de normalidade: (a) Testes de assimetria e curtose, define qual o deslocamento em relação a média; (b) Teste de Wilk-Shapiro para aderência à linha reta, quanto mais próximo de 1 maior o ajuste. Na identificação de outliers, foram adotadas etapas do Controle Estatístico de Processo (CEP) sendo as observações abaixo ou acima de três desvios-padrões (S) da média (M) ( $M \pm 3S$ ), similar ao adotado por Forester (2000).

**c. Análise de tendências:** Para identificação e caracterização de tendências, utilizou-se o Teste de Tendências Qui-quadrado, que indica tendência crescente ( $A > 0$ ) ou decrescente ( $A < 0$ ). Em seguida, traçou-se uma curva de regressão sobre a representação gráfica da série temporal e por diferença entre os dados observados e a regressão, tem-se a série temporal sem tendências (Hipel & McLeod, 1994).

**d. Análise de sazonalidade:** Foi utilizado o Teste de Kruskal-Wallis para identificação da significância da sazonalidade. Associado a gráficos do tipo Box-Plot para verificação da variabilidade dos fatores sazonais por período de análise, e por conseguinte, da amplitude da sazonalidade.

**e. Obtenção da série temporal e modelo comportamental:** Para eliminação da tendência e, posteriormente, da elevada sazonalidade Diagnosticada, fez-se uso da relação (Equação 1) proposta por Morettin & Toloi (2006) e Cavalheiro (2003). Onde,  $Y_t$  = série temporal sem tendência ( $St + Et$ );  $Z_t$  = série de dados original;  $T_t$  = tendência estimada,  $St$  = sazonalidade estimada,  $Et$  = resíduo ou erro estimado:

$$Y_t = Z_t - T_t \quad [\text{Equação 1}]$$

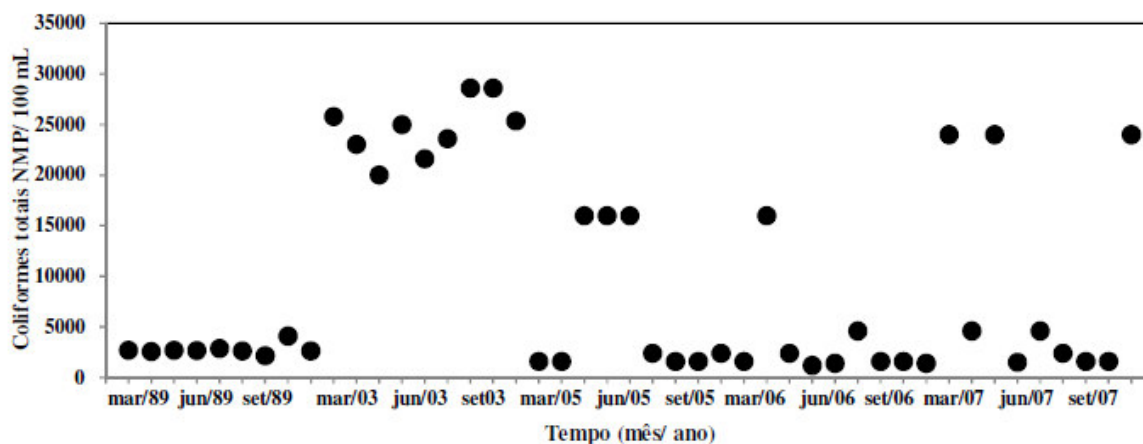
O resultado da equação também indica a sazonalidade mais o resíduo (erro). Com o conhecimento da amplitude da sazonalidade, obteve-se a sazonalidade pura necessária ao cálculo da série temporal pura. Uma vez com a sazonalidade pura mais resíduo, juntamente com a tendência, pode-se enfim encontrar a série temporal pura sem tendência, sazonalidade e resíduo pela Equação 2. A série temporal pura só é encontrada quando eliminados ou pelo menos suavizados a tendência, a sazonalidade e o resíduo (erro ou componente aleatório). Onde,  $Z_o$  = série temporal pura;  $Z_t$  = série de dados original;  $T_t$  = tendência estimada;  $St$  = sazonalidade estimada;  $Et$  = resíduo ou erro estimado.

$$Z_o = Z_t - T_t - St - Et \quad [\text{Equação 2}]$$

## 4 RESULTADOS, ANÁLISE E INTERPRETAÇÃO DOS DADOS

### a. Preparação dos dados de entrada e construção das séries temporais:

A Figura 1 traz um gráfico de dispersão que mostra a série temporal de dados brutos de coliformes totais dos anos de 1989, 2003, 2005, 2006 e 2007, de um trecho da praia da Ponta da Areia ( $P1 = 02^\circ 30' 15'' S$ ;  $44^\circ 19' 06.9'' W$ ). Vale salientar que em função da não-continuidade dos dados optou-se por não ligar com segmentos de retas os pontos plotados. Os dados representados foram obtidos junto a SEMA.



**Figura 1 – Série temporal de coliformes totais em NMP/ 100mL), em trecho da Praia da Ponta da Areia ( $P1 = 02^\circ 30' 15'' S$ ;  $44^\circ 19' 06.9'' W$ ), Ilha do Maranhão (Brasil), 1989-2007.**

A priori, considerando-se todas as séries temporais de dados de coliformes totais da Praia da Ponta da Areia, identificou-se por meio do Teste de Normalidade de Shapiro-Wilk (W) e de medidas de assimetria e curtose

(Tabela 1), que os dados estudados apresentaram distribuição próxima da normal. Optou-se por transformar as séries de dados não normalmente distribuída em série de dados normalmente distribuídas, por meio de transformações Box-Cox, como realizado por Groppo (2005) associadas à estatística não-paramétrica na análise, como feito por Moraes et al (1995). As transformações de dados fazem com sejam distribuídos mais simetricamente havendo, portanto, duas vantagens: (1) alguns métodos não-paramétricos (e mesmo os paramétricos) assumem que os dados são distribuídos simetricamente; e (2) sendo distribuídos simetricamente se torna mais fácil a localização do centro da distribuição (GROPPO, 2005).

**Tabela 1 – Testes de normalidade e análise descritiva para as séries de dados de coliformes totais (NMP/ 100 mL), em trecho da Praia da Ponta da Areia (P1 = 02° 30' 15" S; 44° 19'06.9"W), Ilha do Maranhão (Brasil), 1989-2007.**

Testes	Dados brutos	Transformação logarítmica
	P1	Log P1
W	0.7237	0.7996
N.S. <sup>1</sup>	0.0078	0.0086
Assimetria (G1)	0.7937	0.4514
Curtose (G2)	-1.1841	-1.5897
Variância	10079e+04	0.2558
Média	9465.0222	3.6953
Relação variância/ média	10648.6807	0,0711
Desvio padrão	10039.3077	0.5058
Erro padrão (Resíduo)	1496.5716	0.0754

<sup>1</sup>Nota: N.S. = nível de significância.

As transformações Box-Cox estimaram a melhor transformação para se obter a normalidade dos dados como sendo a logarítmica, confirmada após comparação entre os valores de W do Teste de Shapiro-Wilk, assimetria e curtose (Tabela 1). As transformações também implicaram na redução do desvio das observações em relação à média, repercutindo diretamente nas medidas de assimetria e curtose. Além disso, reduziram o resíduo (erro padrão) das séries temporais e a variância entre os dados. A redução da relação entre variância e a média após as transformações só vem a confirmar a escolha adequada pelas mesmas. A série temporal dos dados do trecho P1, após a transformação, é mostrada na Figura 2.

Para a identificação de outliers, bem como de padrões comportamentais (sazonalidade), como referido na metodologia da pesquisa, foram utilizadas etapas do Controle Estatístico de Processo (CEP), similar ao adotado por Forester (2000), considerando as observações abaixo ou acima de três desvios-padrões (S) da média (M) ( $M \pm 3S$ ), sendo respectivamente o limite inferior e superior dos dados. Na Figura 3, que traz a fixação dos limites superior (LS) e inferior (LI) da análise, é possível perceber que não existiram valores discrepantes (outliers), fato que indica que a escolha da transformação logarítmica foi coerente.

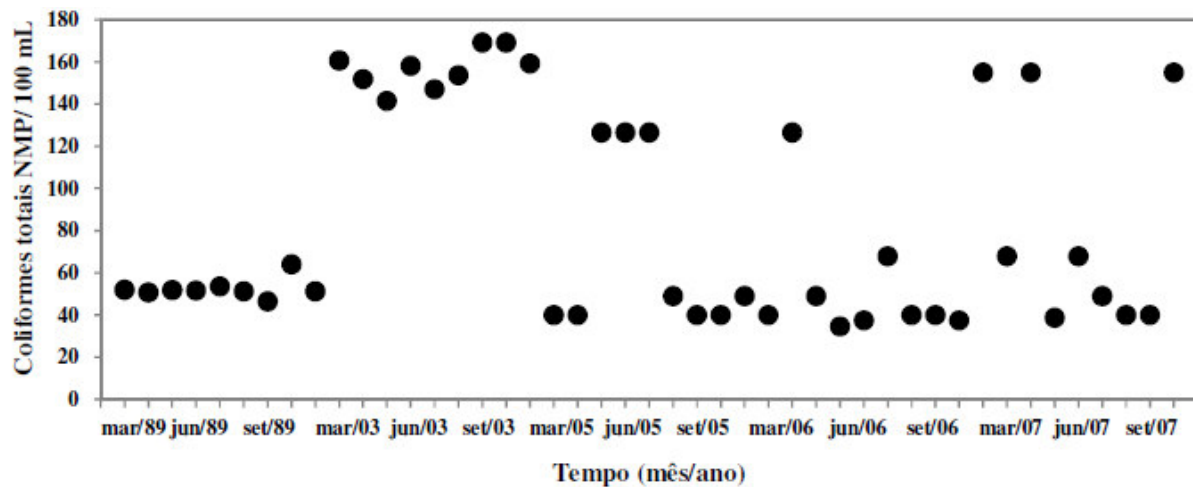


Figura 2 – Série temporal de coliformes totais em NMP/ 100 mL, em trecho da Praia da Ponta da Areia (P1 = 02° 30' 15" S; 44° 19'06.9"W), após transformação logarítmica, Ilha do Maranhão (Brasil), 1989-2007.

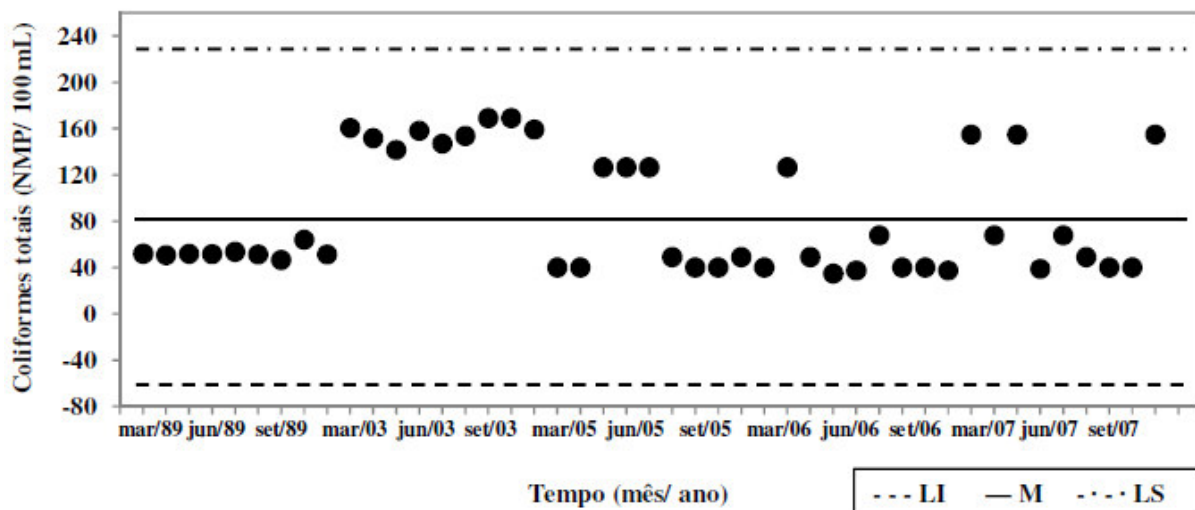


Figura 3 – Série temporal de coliformes totais em NMP/ 100 mL, em trecho da Praia da Ponta da Areia (P1 = 02° 30' 15"S; 44° 19'06.9"W), após transformação logarítmica, com o limite superior (LS), limite inferior (LI) e a média (M), Ilha do Maranhão (Brasil), 1989-2007.

#### b. Análise de tendências

Na análise das tendências na série temporal, foi escolhido o modelo de regressão (linear, exponencial, de potência, logaritmo, polinomial) que melhor se ajustou aos dados, em função do maior coeficiente de explicação ( $R^2$ ) e da maior simplicidade do modelo, dando prioridade aos modelos lineares em detrimento aos polinomiais de alta ordem (Tabela 1). O modelo que melhor representou a tendência do trecho da Praia da Ponta da Areia (P1 = 02° 30' 15" S; 44° 19'06.9"W) foi o modelo logarítmico de ordem 6 (Figura 4). A aplicação de um teste de tendência baseado nas associações entre as séries de dados caracterizou a tendência neste trecho da Praia da Ponta da Areia como decrescente ( $A0 < 0$ ), porém com baixa significância.

Uma tendência polinomial de alta ordem, embora capaz de ajustar-se bem aos dados observados, extrapolações futuras podem não ser tão boas (MORRETTIN & TOLOI, 2006). Podendo apresentar

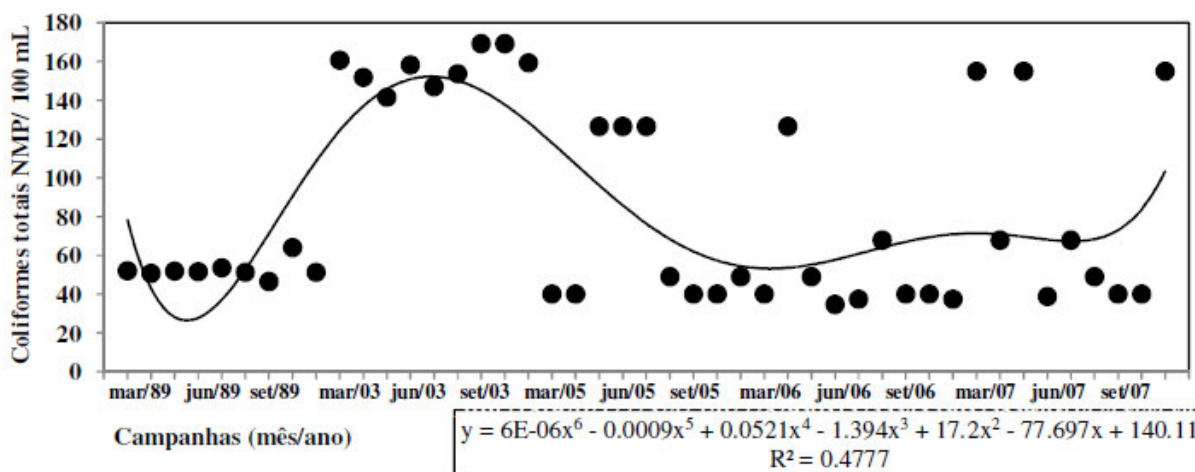
comportamento inexplicável do ponto de vista físico e ter difícil aplicação às regressões com poucos dados (BOOLMANN, 2003).

### c. Análise de sazonalidade

Quanto à sazonalidade no comportamento dos dados, o CEP permite identificação visual de padrões (Figura 3), no entanto foi considerada prudente uma confirmação estatística por meio do teste de sazonalidade não-paramétrico Kruskal-Wallis (H). Com base neste teste, foi verificado que há sazonalidade altamente significativa ( $H = 22.2409$ ;  $p = 0.0002$ ) nas séries temporais de dados de coliformes totais na Praia da Ponta da Areia, sendo que as séries dos anos de 1989, 2005 e 2006 apresentaram maior sazonalidade. A verificação de sazonalidade também foi consubstanciada pela construção de gráfico do tipo “Box-Plot” que mostra a variabilidade dos níveis de coliformes totais (NMP/ 100 mL) ao longo do tempo (Figura 5).

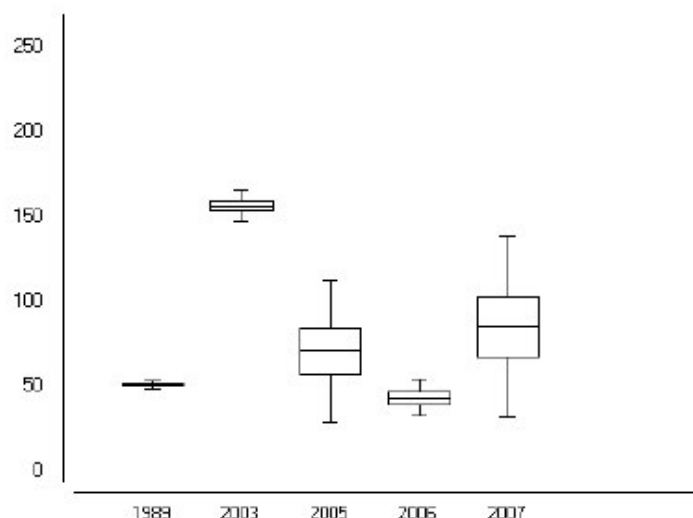
**Tabela 2 – Coeficientes de explicação para modelos aplicados às séries de dados de níveis de coliformes totais (NMP/ 100 mL) em trechos da Praia da Ponta da Areia (P1 = 02° 30’ 15’’ S; 44° 19’06.9’’W), Ilha do Maranhão (Brasil), 1989-2007.**

Modelos	Coeficientes de explicação ( $R^2$ )
Linear	0.0167
Logarítmico	0.0006
Exponencial	0.0251
Polinomial	<u>0.4777</u>
	(ordem 6)
De potência	0.0002



**Figura 4 – Análise de tendência para a série temporal de coliformes totais em m NMP/ 100 mL, em trecho da Praia da Ponta d da Areia (P1 = 02° 30’ 15’’ S; 44° 19’06.9’’W), Ilha do Maranhão(Brasil), 1989-2007.**

A obtenção de tais modelos é importante para a análise da sazonalidade e dos resíduos, e posteriormente na construção de modelos de previsão.



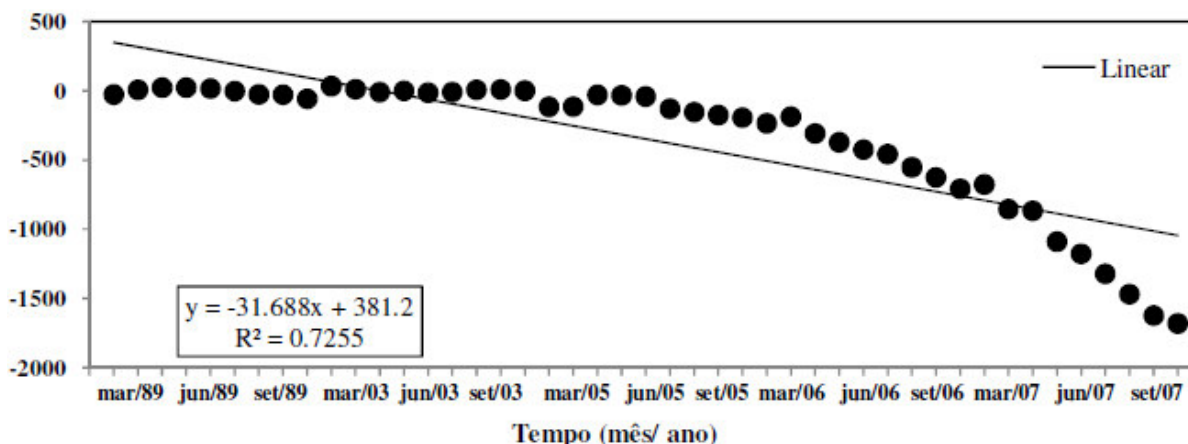
**Figura 5 – Variabilidade nos níveis de coliformes totais em NMP/ 100 mL em m trecho da Praia da Ponta da Areia (P1= 02° 30' 15"S; 44° 19'06.9"W), Ilha do Maranhão (Brasil), 1989-2007.**

Uma vez definida a significância da sazonalidade para as séries temporais, e com base na tendência identificada, entende-se como necessária a suavização ou eliminação da tendência e da sazonalidade para que se tenha um modelo de previsão que represente com mais fidedignidade o comportamento dos dados em função do tempo.

#### **d. Obtenção da série temporal pura e do modelo comportamental**

A grande dificuldade encontrada está no fato da existência da tendência polinomial de ordem elevada nas séries temporais. Este fato prejudica a previsão a longo prazo (Figura 4). Na Figura 5, observa-se a amplitude da sazonalidade é variável para cada ano (período). Uma amplitude de sazonalidade variável implica em uma sazonalidade multiplicativa, isto é, influenciada por fatores multiplicativos, calculados em função da média da sazonalidade do período (9 meses/ano) multiplicada por 2, pois a sazonalidade é aproximadamente simétrica. A sazonalidade mais resíduo foi então dividida por cada fator multiplicativo do período obtendo-se a sazonalidade pura mais resíduo.

Uma vez com a sazonalidade pura mais resíduo, juntamente com a tendência, pode-se enfim encontrar a série temporal pura sem tendência, sazonalidade e resíduo pela Equação 2. Obtendo-se, finalmente, por regressão, o modelo mais simples capaz de explicar o comportamento da variável coliformes totais com 72.55% de certeza (Figura 6).



**Figura 6 – Série temporal pura para dados de coliformes totais em NMP/ 100 mL, em trecho da Praia da Ponta da Areia (P1= 02° 30' 15" S; 44° 19'06.9"W), Ilha do Maranhão (Brasil), 1989-2007.**



## **5 DISCUSSÕES E CONSIDERAÇÕES FINAIS**

As tendências identificadas indicam diminuição na quantidade de coliformes totais na água da Praia da Ponta da Areia, Ilha do Maranhão (Brasil). As tendências decrescentes podem implicar em diminuição dos pontos de lançamento de esgoto in natura, contudo nada conclusivo pois na cidade as poucas estações de tratamento de esgoto estão desativadas. A sazonalidade variável pode indicar oscilações na concepção de que no verão seco (meses junho, julho, agosto, setembro, outubro e novembro) há redução na quantidade de coliformes e que no verão chuvoso (meses dezembro, janeiro, fevereiro, março, abril e maio) há aumento, podendo ocorrer processo oposto em alguns períodos. A grande variabilidade nos dados imposta por um modelo polinomial pode prejudicar a explicação do comportamento e da possibilidade de previsão com base neste. O modelo linear, obtido após a retirada da tendência e da sazonalidade, pode ser utilizado para este propósito contudo com ressalvas, sendo indicado para curto prazo.

## **REFERÊNCIAS BIBLIOGRÁFICAS**

BOLLMANN, H.A. Relação da densidade populacional sobre variáveis de qualidade físico-química das águas superficiais em microbacias hidrográficas urbanas sem cobertura sanitária em Porto Alegre-RS. Tese (Doutorado em Engenharia de Recursos Hídricos e Saneamento Ambiental), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2003.

CAVALHEIRO, D. Método de previsão de demanda aplicada ao planejamento da produção de indústrias de alimentos. Dissertação (Mestrado em Engenharia Mecânica), UFSC, Florianópolis, 2003.130f.

CHATFIELD, C. The analysis of time series: an introduction. 6 ed. Chapman & Hall/ CRC, New York, US, 2004.

CICIN-SAIN, B. & KNECHT, R.W. Integrated Coastal and Ocean Management: concepts and practices. Washington DC: Island Press, 1998.

FORESTER, D. L. Water quality in the credit river 1964-1998. Canadá Institute for Environmental Studies, University of Toronto, 2000.

GROPPO, J.D. Estudo de tendências nas séries temporais de qualidade de água de rios do estado de São Paulo com diferentes graus de intervenção antrópica. Dissertação (Mestrado em Ecologia de Agrossistemas). Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2005.

HIPEL, K.W. & McLEOD, A.D. Times series modeling of water resources and environmental systems. Development in water science, v.45, 1994.

MALETZKE, A. G. Uma metodologia para a extração de conhecimento e informações em séries temporais por meio da identificação de motifs e da extração de características. Dissertação (Mestrado em Ciências), Universidade de São Paulo, São Carlos, 2009.

MORAES, J.M. et al. Estudo preliminar da evolução temporal dos componentes do ciclo hidrológico da bacia hidrológica da Bacia do Rio Piracicaba. In: Anais do Simpósio de Recursos Hídricos, 11. ABRH, 1995.

MORETTIN, P. A. & TOLOI, C. M. C. Análise de séries temporais. 2 ed. São Paulo: Edgard Blucher, 2006.