

Tendências, Sazonalidade e Ciclos no Fluxo de Turistas no Parque Nacional do Iguaçu - PR

Wagner Moura Lamounier, D.Sc.¹

Resumo

Em termos gerais, pretendeu-se com essa pesquisa comprovar estatisticamente a existência dos padrões comportamentais específicos de tendência, ciclos, e sazonalidade no fluxo de visitantes no Parque Nacional do Iguaçu na cidade de Foz do Iguaçu / PR na década de 90. Não se identificou uma tendência expressiva de elevação no número de visitantes nesse período; o que pode indicar uma relativa perda de atratividade do parque em relação à demanda turística, que pode estar se deslocando para outros parques, ou mesmo para outros tipos de roteiros. Os resultados indicaram a incidência de um ciclo sazonal (12 meses de duração) e de três outros ciclos. Dois ciclos de curto-prazo com de 3 e 6 meses; e um ciclo de longo-prazo correspondente a um período de 60 meses (5 anos) de duração. Finalmente, em relação a sazonalidade, observou-se um forte padrão sazonal determinístico na série com destaque para o mês de janeiro como o de maior número de turistas, e o mês de junho como o de maiores quedas na demanda turística em relação à média observada para a década.

Palavras chaves: Tendência, Sazonalidade, Ciclos, Demanda Turística, Parque Nacional do Iguaçu.

Abstract

In general terms, it was intended with this research to prove the existence of the specific patterns of trend, cycles, and seasonality in the visitors' flow of tourists at the National Park of Iguaçu in the city of Iguaçu Falls / PR in the decade of 90. It wasn't identified any expressive upward trend in the number of visitors in that period. This indicate a relative loss of attractiveness of the

¹ Professor adjunto do Centro Universitário de Ciências Gerenciais - CCG/UNA.
E-mail: wlamounier@bol.com.br

park in relation to the tourist demand, that can be moving to other parks, or even for other types of routes. The results indicated the incidence of a seasonal cycle (12 months of duration) and of three other cycles. Two short-period cycles with 3 and 6 months; and a long-period cycle corresponding to a period of 60 months (5 years) of duration. Finally, a deterministic seasonal pattern was observed in the series with prominence for the month of January as the one of larger number of tourists, and the month of June as the one of larger falls in the tourist demand in relation to the average observed for the decade.

Key words: Trend; Cycles; Seasonality; Touristic Demand; National Park of Iguazu.

1. Introdução

Localizado no extremo oeste do Paraná a 640 Km. de Curitiba, o Parque Nacional do Iguazu faz fronteira com a Argentina, além de delimitar-se com diversos municípios paranaenses, abrangendo uma área total de cerca de 185.000 ha.

Criado em janeiro de 1939 e tombado em 1986 pela UNESCO como Patrimônio Natural da Humanidade, o parque constitui-se numa das maiores reservas florestais da América do Sul, bem como numa reserva de proteção dos recursos naturais renováveis do Estado do Paraná. O tipo de vegetação predominante é a mata pluvial e a mata de araucária, com a presença de palmeira, imbuía, caviúna, ervamate etc.

No Parque, a principal atração se refere ao espetáculo proporcionado pelas Cataratas do Iguazu. De forma semicircular, com 2700 metros de largura, as quedas se caracterizam pela sua rara beleza que muito apraz os visitantes. Além das exuberantes cataratas, possui em seu interior outras atrações como a fauna, bastante representativa, o Poço Preto, o Salto do Macuco, o Centro de Visitantes, a Estátua de Santos Dumont, responsável pela transformação da área das cataratas em Parque Nacional, e um hotel de lazer.

Nas trilhas do Parque, o turista encontra tranquilidade, ar puro, e muita bioenergia, proporcionadas pela mata densa, formada por árvores de grande porte como as figueiras bravas, canelas, pau marfim, jerivás e palmitos. O Parque Nacional do Iguazu tem valor incalculável como banco genético para futuros reflorestamentos. Além disso, o parque é um dos poucos ecossistemas onde ainda se pode

"Reuna, Volume 7, nº 1(18), Janeiro - Março/2002"

encontrar espécies como a onça-pintada, anta, capivara, veado, guaxinim, macaco-prego, quati e o jacaré-do-papo-amarelo. Todavia, apesar das qualidades e potencial turístico apresentado pelo parque, pode-se verificar, conforme ilustra a Figura 1, que o fluxo de visitantes no parque vem apresentando um comportamento extremamente irregular. Além dessa irregularidade, tem-se indícios de que este fluxo, quando observado ao longo das duas últimas décadas, vem diminuindo ao longo do tempo, o que indicaria uma possível perda de atratividade do parque em relação à demanda turística. Isto, em última instância, poderia implicar em uma série de perdas para agentes econômicos que de maneira direta ou indireta tem suas rendas dependentes do turismo na região, e particularmente no Parque Nacional do Iguazu.

2. Objetivos

Em termos gerais, pretende-se nesta pesquisa comprovar estatisticamente a existência padrões comportamentais específicos chamados de tendência, ciclos, e sazonalidade no fluxo de visitantes no Parque Nacional do Iguazu na cidade de Foz do Iguazu / PR na década de 90.

Especificamente, pretende-se analisar como a dinâmica desse fluxo é influenciada por esses componentes, e quais os seus possíveis impactos sobre a renda dos agentes econômicos ligados ao turismo no Parque. Pretende-se ainda apontar possíveis políticas e medidas a serem tomadas por agentes da esfera pública e privada, afim de ampliar o fluxo turístico no parque e melhor equacionar a infraestrutura do mesmo à sua demanda futura.

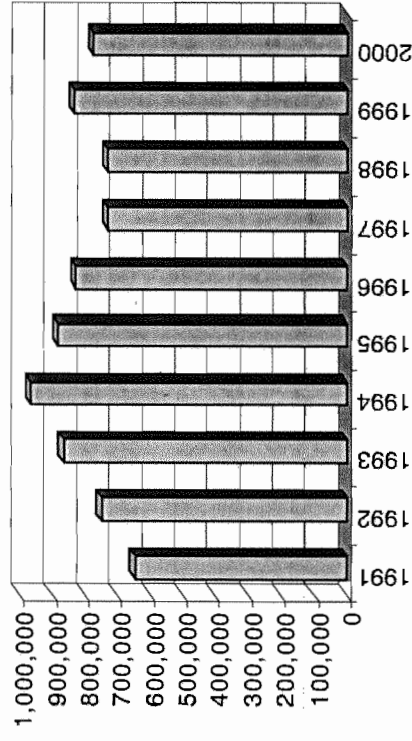


Figura 1: Fluxo de Visitantes no Parque Nacional do Iguazu (1991 / 2000)

Fonte: Secretaria de Estado do Esporte e Turismo do Paraná

3. Justificativa

A análise proposta torna-se importante à medida em que uma gestão eficiente e moderna tanto a nível governamental quanto privado se caracteriza como primordial nos dias atuais. Particularmente, no tocante à administração e exploração sustentável de recursos naturais, como é o caso do Parque Nacional do Iguaçu, uma melhor compreensão do padrão de visitas ao parque pode auxiliar estes administradores a melhor direcionar seus esforços para otimizar o potencial turístico do parque de uma maneira sustentável e não degradante dos recursos deste espaço.

Além disso, para os diversos agentes econômicos que dependem de forma direta ou indireta do turismo na região do Parque (tais como comerciantes, trabalhadores, hotéis, pousadas, agências de viagens, etc.); a busca de informações e "mecanismos" capazes de minimizar suas incertezas quanto à demanda futura por seus serviços e produtos se torna fundamental, no sentido em que estas informações podem lhe proporcionar maior capacidade de planejamento e adequação em relação ao futuro de seus negócios.

Ainda nesse sentido, SCHWAGER (1996), ressalta a importância desse tipo de análise, dentro de processos de tomada de decisões sob situação de incerteza, que incide sobre os agentes econômicos em diversos ramos de atividade.

4. Metodologia

Uma das principais formas de análise de séries temporais se dá pela busca da decomposição das mesmas em termos de seus componentes principais. Esquemáticamente, esses componentes da variação em uma série temporal podem ser caracterizados como tendências, sazonalidades e ciclos.

A Tendência em uma série temporal está ligada, conforme aponta CHATFIELD(1996), à uma mudança no nível médio da série no longo-prazo. Ou seja, a tendência irá refletir o declínio, a elevação ou a estabilidade (quando de fato não houver tendência) do valor médio da série temporal no longo-prazo; que pode-se convencionar como sendo referentes a observações relativas a períodos de tempo de 25 anos, aproximadamente.

Os Ciclos em uma série de tempo se referem a movimentos recorrentes de elevação e queda em torno do nível médio da tendência, e portanto, também são relati-

vos à dados que refletem o longo-prazo da variável em questão. De acordo com SOUZA(1989), deve-se ressaltar que os movimentos para cima e para baixo em torno da tendência podem ser estritamente periódicos ou aproximadamente periódicos, assim sendo, eles podem variar em comprimento e magnitude.

O terceiro componente que pode estar presente em uma série temporal é aquele relativo à Sazonalidade. Assim como o componente cíclico, a sazonalidade também está relacionada a movimentos para cima e para baixo em torno de um valor médio, contudo ela difere basicamente em dois aspectos do primeiro: a sazonalidade possui um comprimento constante de 12 meses, repetindo-se nesta base periódica regular; as variações sazonais podem ser observadas tendo-se por base períodos menores de tempo (médio e curto-prazo) e, assim sendo, os dados a serem analisados não deverão estar distribuídos em observações anuais, mas sim mensais ou trimestrais, para que seja possível a modelagem da sazonalidade.

4.1 Análise do Componente de Tendência

O primeiro passo na análise da tendência no fluxo de visitantes ao Parque Nacional do Itatiaia (que será aqui denotado genericamente por y_t) é dado verificando-se a estacionariedade da série em questão. Em termos práticos, a estacionariedade que é testada se refere à estacionariedade fraca. Isto é, deve-se verificar se a série temporal (y_t) terá média e variância finitas. Caso isto se confirme a mesma será considerada estacionária em covariância. Os principais testes para verificação da estacionariedade ou não da série (chamados de testes de raízes unitárias²) que serão empregados são os testes de Dickey-Fuller Expandido (ADF) e o teste de Phillips-Perron (PP).

A tendência em uma série pode ser provocada por dois tipos de fenômenos que apresentam características particulares. De acordo com a sua natureza e padrão comportamental, a tendência poderá ser caracterizada como sendo do tipo determinística ou do tipo estocástica.

A característica essencial de uma *tendência determinística* se refere ao fato de que a variação no nível médio de uma dada variável se dará, de forma previsível, como uma função do tempo. Ou seja, conforme apontam PINDYCK e RUBINFELD (1998), e BOWERMAN e O'CONNELL (1979), uma série de

² Para maiores detalhes acerca dos procedimentos destes testes ver HAMILTON(1994).

tempo Y_t , poderá evoluir em montantes absolutos em cada período de tempo de acordo com a seguinte equação:

$$Y_t = T_t + \varepsilon_t \quad (1)$$

em que: T_t é o componente de tendência determinística do modelo e ε_t é um termo de erro com média zero e variância igual a σ^2 .

O valor de Y_t poderá ser previsto de acordo com a equação (1) sendo que o componente de tendência (T_t) poderá assumir várias formas, tais como:

$$T_t = \varphi_0 \quad (2)$$

$$T_t = \varphi_0 + \varphi_1 t \quad (3)$$

$$T_t = \varphi_0 + \varphi_1 t + \varphi_2 t^2 + \dots + \varphi_k t^k \quad (4)$$

Nesse tipo de modelos, a média da série poderá evoluir de forma linear ou curvilínea ao longo do tempo, sendo que uma gama enorme de possibilidades poderá ser verificada, de acordo com os sinais dos coeficientes (φ_i) do polinômio.

A tendência determinística de uma série também pode assumir formas mais complexas, como a do *modelo de crescimento exponencial*:

$$Y_t = A \cdot e^{\delta t} \varepsilon_t \quad (5)$$

que pode ser estimado após tomar-se o logaritmo de ambos os lados de (5), que resulta:

$$\ln Y_t = \varphi_0 + \varphi_1 t + u_t \quad (6)$$

em que: $\varphi_0 = \ln A$, $\varphi_1 = \delta$, que é a *taxa de crescimento* da série, e $u_t = \ln \varepsilon_t$.

Todavia, o problema com essa classe de modelos de análise é que nem sempre a tendência nos dados em questão é do tipo determinística. Ao contrário, ela pode estar mudando de forma aleatória ao longo do tempo se caracterizando como um componente de tendência do tipo estocástica.

A diferença principal entre estas duas formas de tendência se refere ao fato de que a tendência estocástica implica em uma variação percentual média na série em dado período de tempo, porém, ao contrário da determinística, em cada período a mudança provocada pela tendência em relação ao seu nível médio será um montante aleatório e imprevisível, ao invés de constante dado por uma determinada taxa. Ou seja, no caso da existência de uma tendência do tipo estocástica, tem-se que os coeficientes (φ_k) de uma equação para a descrição da tendência, como os apresentados nos modelos anteriores, poderão não ser constantes, mas também variar aleatoriamente ao longo do tempo.

Existem diversos modelos de séries temporais que incorporam tendências estocásticas, e que vem ganhando destaque na bibliografia recente sobre análise de séries temporais. Nessa classe de modelos, destaca-se o *modelo de passeio aleatório (random walk)* dado por:

$$Y_t = Y_{t-1} + \varepsilon_t \quad (7)$$

ou:

$$\Delta Y_t = \varepsilon_t \quad (8)$$

em que: ε_t é um erro do tipo *ruído branco*, com média 0 e variância σ^2 , e Δ é um *operador de diferenças finitas de primeira ordem*.

O modelo de passeio aleatório é claramente um caso especial de um processo auto-regressivo de ordem 1 [AR(1)], dado por:

$$Y_t = \gamma_0 + \gamma_1 Y_{t-1} + \varepsilon_t \quad (9)$$

em que: $\gamma_0 = 0$ e $\gamma_1 = 1$.

A variância da série Y_t em questão será dependente do tempo pois os choques exógenos (ε_t) terão efeitos cumulativos com o passar do tempo sobre a série. Assim sendo o processo passeio aleatório será não-estacionário, pois $t \rightarrow \infty \Rightarrow \text{Var}(Y_t) \rightarrow \infty$.

Um outro modelo que incorpora tendência estocástica, e que é bastante comum na literatura acerca de tendências estocásticas³, é uma variação do modelo de

³ Conforme pode-se observar nos trabalhos de ENDERS (1995), HAMILTON (1994) e CAMPBELL, LO e MACKINLAY (1997)

passaio aleatório que incorpora uma parte determinística ao modelo de análise das mudanças na série Y_t . O modelo de passeio aleatório com intercepto (*random walk plus drift*), que aumenta o modelo de passeio aleatório com a inclusão de um termo constante, é dado por:

$$Y_t = Y_{t-1} + \gamma_0 + \varepsilon_t \quad (10)$$

Assim sendo, dada um condição inicial Y_0 , a solução geral para a série Y_t , será dada por:

$$Y_t = Y_0 + \gamma_0 t + \sum_{i=1}^t \varepsilon_i \quad (11)$$

De acordo com ENDERS (1995), o comportamento da série Y_t , portanto, será governado por dois componentes não estacionários, quais sejam: uma tendência linear determinística ($\gamma_0 t$) e uma tendência estocástica ($\sum \varepsilon_i$). A variação determinística em cada nova observação de Y_t será dada por γ_0 e após t períodos de tempo será igual a $\gamma_0 t$; ao passo que a variação provocada pela tendência estocástica será dada pelo somatório dos diversos choques ε_i que terão efeitos permanentes na média de Y_t .

4.2 Análise do Componente Cíclico

A análise de séries temporais pode ser implementada segundo diversas abordagens, que nem sempre são concorrentes entre si. Esse é o caso da *análise no domínio da frequência* (ou *Análise Espectral*), que representa uma forma de análise de séries temporais que fornece informações complementares àquelas propiciadas pela *análise no domínio do tempo*, informações estas que geralmente estão ligadas ao estudo da ocorrência de ciclos nas séries.

A essência da análise espectral reside na representação de uma série temporal $\{Y_t\}_{j=-\infty}^{\infty}$ como uma soma ponderada de funções periódicas do tipo $\text{sen}(\omega t)$ e $\text{coseno}(\omega t)$, em que ω representa uma frequência particular; e na determinação de quais importantes são os ciclos de diferentes frequências, caso eles existam, para explicar a variância de $\{Y_t\}_{j=-\infty}^{\infty}$, e consequentemente, o comportamento dessa série.

A análise espectral para verificação de ciclos em uma série temporal $Y(t)$ tem como ponto de partida um teorema conhecido como *Teorema da Representação Espectral*⁴ segundo o qual, qualquer processo estacionário discreto $Y(t)$, com frequências pertencentes ao intervalo $[0, \pi]$, pode ser representado da seguinte maneira:

$$Y_t = \mu + \int_0^\pi [\alpha(\omega) \cdot \cos(\omega t)] d\omega + \int_0^\pi [\beta(\omega) \cdot \text{sen}(\omega t)] d\omega \quad (12)$$

que é a forma mais comum para expressar a Representação Espectral do Processo Y_t .

Da equação (12) surge, portanto, o conceito de que a análise espectral descreve o valor de uma variável Y_t como uma soma ponderada de funções periódicas da forma $\text{cos}(\omega t)$ e $\text{sen}(\omega t)$, em que w representa uma frequência particular no intervalo $[0, \pi]$ ⁵.

A frequência angular $\omega = \pi$, medida em termos de radianos por unidade de tempo, é conhecida como a *Frequência de Nyquist*, e denota o valor mais alto com que se pode obter informações sobre ciclos periódicos em um conjunto de dados. Portanto, tem-se que o ponto principal da introdução da equação (12), se refere ao fato de que ela indica que cada frequência no intervalo $[0, \pi]$, pode contribuir para a variação de uma série de tempo Y_t , qualquer.

Assim como na análise de uma série temporal no domínio do tempo o conhecimento da *Função de Autocovariância*⁶ é de valor imprescindível para descrever a estrutura de tal processo; na análise no domínio da frequência existe uma função chamada *Função de Densidade Espectral*, ou simplesmente *Espectro Populacional* do processo $\{Y_t\}_{j=-\infty}^{\infty}$, cuja integral definida em $[0, \pi]$, chamada de

⁴ Esse teorema não será demonstrado aqui pois envolve *Cálculo Integral Estocástico*, que é um tópico avançado de análise matemática, e foge do escopo do presente trabalho. Para maiores detalhes desse teorema, ver CRAMÉR e LEADBETTER (1967).

⁵ O intervalo de integração é $[0, \pi]$, e não $[0, \infty]$, devido à seguinte característica das funções trigonométricas do tipo coseno:

$$\begin{aligned} \cos[(\omega + n\pi)t] &= \cos \omega t \quad \forall n \text{ e } t \text{ inteiros, com } n \text{ par} \text{ e} \\ \cos[(\omega + n\pi)t] &= \{\cos(\pi - \omega)t \quad \forall n \text{ e } t \text{ inteiros, com } n \text{ ímpar} \} \end{aligned}$$

Assim sendo, variações em frequências maiores do que π não poderão ser distinguidas de variações nas frequências correspondentes no intervalo $[0, \pi]$, pois as mesmas irão se sobrepor.

⁶ Dada por $\gamma(k) = E\{[Y(t) - \mu][Y(t+k) - \mu]\} = \text{Cov}\{Y_t, Y_{t+k}\}$

Função de Distribuição Espectral, é a representação espectral da função de autocovariância. Esse conceito também se configura num desenvolvimento teórico de grande importância para a compreensão da estrutura variacional do processo. O *Espectro Populacional* do processo $\{Y_t\}_{t=-\infty}^{\infty}$, pode ser escrito como:

$$f_Y(\omega) = \frac{1}{2\pi} \left[\gamma_0 + 2 \sum_{j=1}^{\infty} \gamma_j \cos(\omega j) \right] \quad (13)$$

Conforme aponta HAMILTON (1994), como o espectro representa uma função contínua de ω , definida no conjunto dos números reais; e γ_j representa as autocovariâncias de um processo estacionário, tem-se que $f_Y(\omega)$ terá um valor positivo para qualquer ω . Assim como ocorre na análise no domínio do tempo, em que a importância prática da função de autocorrelação não se dá com a interpretação direta dos seus coeficientes, mas sim com a análise do *correlograma* da série; na análise espectral, os coeficientes do espectro não representam o principal ponto de interesse para análise. É o *gráfico do espectro*, que também é chamado de espectro, que fornecerá as informações relevantes acerca da *variância* da série Y_t e sobre quais as frequências que contribuem para um comportamento cíclico dessa série.

Contudo, as séries temporais a serem analisadas empiricamente, geralmente são *amostras finitas* da população $\{Y_t\}_{t=-\infty}^{\infty}$, e portanto, não se pode estimar, a partir delas, o número infinito de parâmetros da equação (13). Tem-se portanto que, assim como ocorre nos modelos de análise estatística e econométrica "tradicionais"; geralmente o que se tem, para a implementação da análise espectral, são amostras finitas dos dados populacionais das séries de tempo a serem analisadas, deve-se trabalhar com um conceito análogo ao do espectro populacional, para a análise espectral de dados amostrais. Este conceito é conhecido na literatura como o *Periodograma Amostral*, ou simplesmente *Periodograma* e a forma geral dos estimadores consistentes do espectro $f_Y(\omega)$ da equação (13), apontada por JUDGE et alii. (1985), é dada pela seguinte expressão:

$$\hat{f}_Y(\omega) = \frac{1}{2\pi} \left(\lambda_0 c_0 + 2 \sum_{j=1}^{M_T} \lambda_j c_j \cos(\omega j) \right) \quad (14)$$

em que a seqüência de pesos $\lambda_j = \{\lambda_0, \lambda_1, \lambda_2, \dots, \lambda_{M_T}\}$ é chamada de *janela de defasagem* (ou de suavização), e o termo $M_T (< N)$ é chamado de *ponto de truncamento*. Uma regra sugerida por CHATFIELD (1996) para a escolha do valor para o ponto de truncamento é a de que deve-se escolher um ponto de truncamento (M_T) dado por:

$$M_T = 2 \cdot \sqrt{N} \quad (15)$$

De acordo com o sugerido por RAUSSER e CARGILL (1970), para se verificar se em um espectro estimado, existem picos estatisticamente significativos, pode-se empregar dois critérios: O primeiro se refere à *análise visual* da função de densidade espectral estimada buscando se identificar se alguns picos se destacam nesse espectro em relação aos demais. O segundo critério, mais rigoroso, é baseado no *teste da hipótese* (nula) de que o espectro estimado não é estatisticamente diferente do espectro de um *processo puramente aleatório* (ruído branco). Isto é feito comparando os valores estimados do periodograma com uma distribuição exponencial.

4.3 Análise do Componente Sazonal

Diversos métodos podem ser empregados para a análise do componente sazonal em uma série de tempo. Desde *análise visual* do gráfico da série, que se apresenta como a mais simples, porém útil, opção; passando pelos *Modelos de Regressão Linear com Variáveis Independentes Binárias* (*Variáveis Dummy*) até a utilização de métodos mais sofisticados como a propiciada pelos *Modelos de Análise Espectral* (que trata a sazonalidade como um ciclo de período fixo igual a 12 meses) e dos *Modelos de Box e Jenkins Sazonais - SARIMA*.

Para uma série y_t mensal que apresente o componente sazonal, os diversos pontos amostrais deverão apresentar algum grau de correlação com os dados correspondentes com *lags* e *leads* de 12 meses. Ou seja, espera-se a existência de correlação entre y_{t-12} e y_{t+12} com os valores de y_t . Essas correlações deverão se manifestar na função de autocorrelação amostral, dada por (16), pela presença de picos significativos para os valores estimados de $\hat{\rho}_k$, em que $k = 12, 24, 36, 48$ e assim por diante.

$$\hat{\rho}_k = \frac{\sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y})}{\sum_{t=1}^T (y_t - \bar{y})^2} \quad (16)$$

A mensuração dos efeitos, ou impactos, que os diferentes meses do ano exercem sobre a série de tempo em questão pode ser efetuada com a estimação de *índices de sazonalidade*, que de forma agregada representam o componente sazonal. ABRAHAM e LEDOLTER (1983) descrevem o componente sazonal, e os referidos índices, com a ajuda da seguinte equação:

$$S_t = \sum_{i=1}^s \delta_i IND_{it} \quad (17)$$

em que; $IND_{it} = 1$ se t se refere ao período sazonal i , e $IND_{it} = 0$ para os demais períodos, δ_i representam os valores dos índices que deverão ser adicionados ou subtraídos da série sazonal afim de se eliminar esse efeito e s representa o número de períodos de sazonalidade que poderão ocorrer. Por exemplo, s será igual a 12, se o fenômeno sazonal a ser analisado ocorrer mensalmente ao longo do ano e s será igual a 4 se o fenômeno sazonal estiver ligado às estações do ano (trimestral).

De maneira geral, a implementação do cálculo de tais índices consistirá da especificação de uma equação de regressão, que será estimada por MQO, tendo como variável dependente a série y_t , livre do componente de tendência. Assim sendo, a equação a ser estimada será:

$$y_t = \beta_0 + \sum_{i=1}^{12} \beta_i D_{it} + \varepsilon_t \quad (18)$$

em que: D_{it} ($i = 1, 2, 3, \dots, 12$ e $t = 1, 2, 3, \dots, n$) são variáveis *dummy* assumindo o papel das variáveis IND na equação (17); ε_t é um termo de erro aleatório e β_i representa o efeito sazonal no i -ésimo período.

Como pode ser observado, a equação (18) possui intercepto. Esse modelo não poderá ser estimado com essa especificação, pois ocorrerá aí o problema econométrico de *multicolinearidade perfeita*⁸. Assim sendo, um dos valores de β_i deverá ser igualado à zero (β_{12} por exemplo) e portanto a interpretação dos

⁷ Lags se referem a valores defasados e leads se referem a valores futuros.

⁸ O somatório das colunas das *dummies* apresentará colinearidade perfeita com a coluna de números um que multiplicam o coeficiente de intercepto. A matriz $(X'X)^{-1}$ será singular.

índices sazonais será modificada. Cada um dos índices β_i (para todo $i \neq 12$) representará o efeito sazonal do i -ésimo mês em relação ao valor observado no 12º mês. Este tipo de modelo se adequa à modelagem da sazonalidade determinística da série. Neste caso o efeito sazonal apresenta um comportamento relativamente estável e previsível ao longo dos anos.

Contudo, conforme o discutido por VANDAELE (1983) o padrão da sazonalidade na série y_t poderá ser do tipo *estocástico*. Isso ocorrerá nos casos em que o componente sazonal possuir um comportamento instável ao longo do tempo, se modificando de forma imprevisível. Assim sendo, a estimação de um modelo como o dado pela equação (18) não irá ter nos seus resíduos uma série y_t livre do fenômeno sazonal.

Alguns modelos de séries temporais, tais como o SARIMA, desenvolvido por BOX e JENKINS (1976) incorporam a sazonalidade na forma estocástica em suas estruturas, para propiciar análises e previsões mais precisas dos valores futuros da série. Os modelos sazonais de BOX & JENKINS (1976), também conhecidos como modelos SARIMA (p,d,q) x (P,D,Q)_s, visam explorar a correlação serial existente entre diferentes valores de uma série temporal Y_t . A equação fundamental deste modelo é dada pela seguinte expressão:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)(1 - \Phi_1 B^s - \dots - \Phi_p B^{ps})(1 - B)^d (1 - B^s)^D y_t = (1 - \theta_1 B - \dots - \theta_q B^q)(1 - \Theta_1 B^s - \dots - \Theta_D B^{Ds}) \varepsilon_t \quad (19)$$

Na primeira parte desta expressão tem-se que o primeiro termo entre parêntesis⁹ corresponde aos coeficientes de um modelo auto-regressivo -AR(p); o segundo termo corresponde aos coeficientes de um modelo auto-regressivo sazonal -SAR(P); o termo d corresponde à ordem de integração da série, ou seja o número de diferenças tomadas e D corresponde ao número de eventuais diferenças sazonais tomadas.

Na segunda parte desta expressão, após o sinal de igualdade, tem-se que o primeiro termo entre parêntesis corresponde aos coeficientes de um modelo do tipo média móvel - MA (q) e o segundo termo corresponde aos coeficientes de um modelo de médias móveis sazonais - SMA (Q).

⁹ Em que B é um operador de defasagens. Por exemplo $BY_t = Y_{t-1}$ e $B^k Y_t = Y_{t-k}$.

Utilizando-se o operador de diferenças ∇ , este modelo pode ser reescrito, de maneira mais compacta, como:

$$\varphi(B)\Phi(B^s)\nabla^d\nabla_s^D y_t = \theta(B)\Theta(B^s)\varepsilon_t \quad (20)$$

Para se identificar um modelo SARIMA $(p,d,q) \times (P,D,Q)S$, deve-se inicialmente determinar os valores de d e D por meio da análise do gráfico das funções de autocorrelação e de autocorrelação parcial. Se for observado um comportamento não estacionário com persistência de elevados valores para os coeficientes de autocorrelação e de autocorrelação parciais nos pode-se inferir pela necessidade de diferenciação da série.

Da mesma forma, se valores elevados e persistentes nos lags com defasagens próximas aos múltiplos do período sazonal (próximas de 12, 24 e 36 por exemplo) pode-se inferir também pela necessidade de diferenciação sazonal da série.

4.4 Dados e Softwares Estatísticos Utilizados

Para a implementação da presente pesquisa foram utilizados dados mensais secundários do Fluxo de Visitantes ao Parque Nacional do Iguaçu no período de janeiro de 1991 a dezembro de 2000 fornecidos pela Secretaria de Turismo do Estado do Paraná.

Os principais Softwares que foram empregados no trabalho de pesquisa para o processamento dos dados e dos modelos de análise nos domínios do tempo e da frequência foram: STATISTICA 5.0 da Stat Soft Inc., o SPSS 8.0 da SPSS Inc. e o EVIEWS 3.1 da Quantitative Micro Software.

5. Resultados e Discussão

Conforme pode-se observar na Figura 2, o comportamento do fluxo de visitantes ao Parque Nacional do Iguaçu na década de 90 apresentou ao longo dos meses um comportamento extremamente volátil. Todavia alguns padrões comportamentais podem ser identificados pela simples observação do gráfico da série. Primeiramente pode verificar uma tendência geral de crescimento no número de visitantes até a metade da amostra (1995); tendência esta que se reverteu após esse período. Os ciclos na série se tornam menos nítidos no gráfico dado que são componentes de longo prazo e poderiam ser identificados mais facilmente pelo gráfico

do fluxo de visitas apenas se fosse tomada uma série com um maior número de observações. Contudo pode-se verificar que ao longo de toda a amostra se configura um ciclo aparentemente completo.

Com relação ao padrão sazonal pode-se observar um nítido comportamento em que alguns meses do ano se mostram como períodos de maior visitação ao parque, enquanto outros se caracterizam por uma acentuada redução no fluxo.

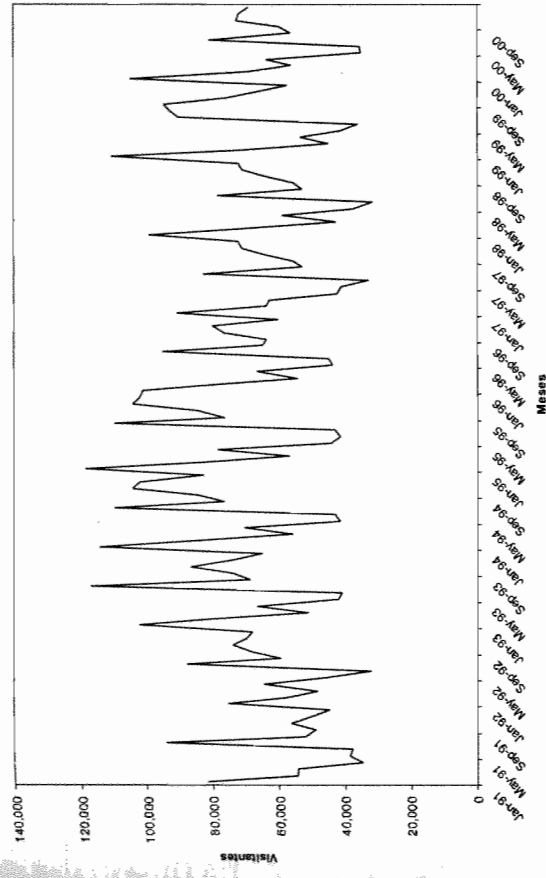


Figura 2. Fluxo mensal de visitantes ao Parque Nacional do Iguaçu - Jan. 1991 / Dez. 2000.

Fonte: Elaboração do autor a partir de dados fornecidos pela Secretaria de Estado do Esporte e Turismo do Paraná

Apesar da análise "visual" do gráfico da série se constituir em um importante instrumento na identificação de padrões comportamentais em séries temporais, deve-se proceder com testes estatísticos mais rigorosos e completos nesse processo. Assim sendo, primeiramente foram implementados testes em relação ao componente de tendência no fluxo de visitas ao parque.

Os testes de raízes unitárias descritos anteriormente se configuram como testes para identificar uma tendência estocástica na série, caso a série possua raiz unitária. No presente caso da série do fluxo de visitantes ao Parque Nacional do Iguaçu, os resultados dos testes estão sintetizados no Quadro 1.

Quadro 1 - Resultados do teste de Dickey-Fuller expandido

Estadísticas do teste de Dickey-Fuller expandido	1 defasagem	2 defasagens	3 defasagens	4 defasagens
Modelo com intercepto e tendência determinística	-6.233583	-4.576386	-5.637386	-5.815010

Fonte: Resultados da pesquisa.

Obs.: Os valores críticos de MacKinnon para a rejeição da hipótese nula de presença de raízes unitárias em níveis de significância de 1% e 5% foram, respectivamente, de - 4,0387 e -3,4481.

Conforme pode-se verificar os resultados encontrados não confirmam a hipótese de raiz unitária para a série, o que implica que a tendência na mesma, caso exista não é da forma estocástica, mas sim de natureza determinística. Nesse sentido, foram estimados os modelos de tendência dados pelas equações (1) a (6). Verificou-se que o modelo que apresentou o melhor ajustamento foi o modelo de tendência quadrática para o fluxo de visitantes no parque, cujos principais resultados são dados pela equação (21).

$$\text{Visitantes} = 56.350,76 + 536,5135 \cdot T - 4,3923 \cdot T^2 \quad (21)$$

$$\text{Valores -P} \quad (0,0000) \quad (0,0166) \quad (0,0159)$$

$$\bar{R}^2 = 0,0329 \quad DW - 1,5223 \quad F - 3,0245$$

Apesar do fato do coeficientes estimados serem todos significativos ao nível de 10% adotado na presente pesquisa, pode-se verificar que a significância geral do modelo de tendência determinística é baixa, com um coeficiente de determinação ajustado de apenas 3,29% e um valor para a estatística F igual a 3,02. Isso implica que a tendência, apesar de presente na série, não é um fator de grande relevância para explicar o comportamento do fluxo de visitas ao Parque Nacional do Iguaçu.

Os valores estimados para os coeficientes de tendência implicam que com o passar do tempo, o fluxo de turistas tendeu a aumentar, pela dominância do efeito do coeficiente linear à taxa média de aproximadamente 537 novos turistas por mês. Todavia após a metade da amostra o efeito do termo quadrático da tendência (-4,39) passa a dominar, invertendo a tendência do fluxo de visitação, projetando uma queda no número de turistas para os anos vindouros. Este comportamento pode ser melhor visualizado com a ajuda da Figura 3.

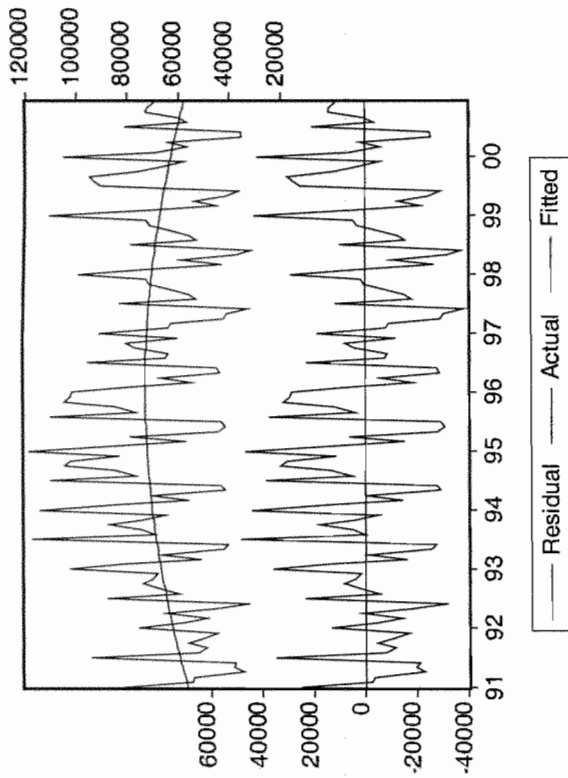


Figura 3 - Gráfico da tendência do fluxo de visitas ao Parque Nacional do Iguaçu.
Fonte: Resultados da pesquisa.

Com relação à incidência de ciclos no comportamento temporal do fluxo de turistas no parque, a análise espectral dos dados confirmou a incidência de um forte ciclo sazonal (12 meses de duração) conforme indicam os valores no Quadro 2.

Quadro 2 - Maiores valores estimados para as densidades espectrais¹⁰

Frequência (w)	Período (P)	Coeficientes		Valores do periodograma	Densidades espectrais
		(Coseno)	(Seno)		
.083333	12	8653.57	-12339	213628472E3	92733206E2
.333333	3	13560.67	4236.5	12110398E3	81491860E2
.166667	6	9395.11	5288.5	69742096E2	46821400E2
.075000	13.33	-2290.00	2722.7	759436667	28527346E2
.091667	10.91	2254.55	1013.4	366595599	26043271E2
.325000	3.08	1258.38	-2390.6	437901826	23980008E2
.016667	60	1061.40	-6914.5	29362357E2	22007846E2
.500000	2	7192.16	0.0	31036298E2	21488699E2
.341667	2.93	427.67	278.4	15624916.	20647376E2
.158333	6.32	1244.04	1258.6	187904832.	13854476E2

Fonte: Resultados da pesquisa.

¹⁰M = 5, de acordo com a regra de Parzen que estabelece como um ponto de truncamento (M) um número próximo a 5% do número de observações.

Todavia três outros ciclos puderam ser identificados na análise espectral da série estudada. Dois ciclos de curto-prazo, correspondentes às frequências 0,3333 e 0,1666, que em termos de períodos equivalem a ciclos com duração respectiva de 3 e 6 meses. Um ciclo de longo-prazo também foi identificado. Este ciclo corresponde à frequência 0,0166 que implica em um período de 60 meses (5 anos) de duração.

O espectro estimado para a série (Figura 4) confirma a incidência dos ciclos mencionados anteriormente. Este espectro foi estimado utilizando-se a Janela de Parzen com defasagem de truncamento $M = 5$, após a retirada do componente de tendência na série afim de se evitar o problema conhecido como *leakage*. Conforme pode-se verificar existem picos significativamente destacados próximos às frequências 0,833; 0,333; 0,166 e 0,012 que correspondem aos ciclos apontados. Para as demais frequências o que se observa é um comportamento suave e sem picos significativos.

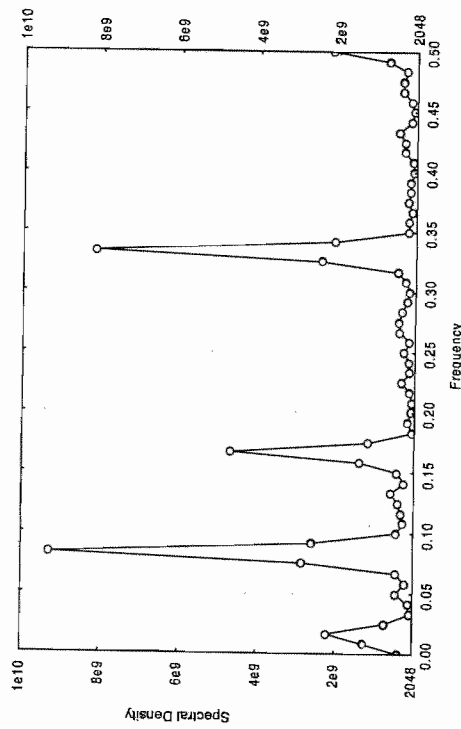


Figura 3 - Espectro estimado para o fluxo de visitantes no Parque Nacional do Iguaçu (em função de ω).

Fonte: Resultados da pesquisa.

Finalmente, em relação à análise da sazonalidade no fluxo de turistas, observou-se um forte padrão sazonal no mesmo, conforme apontado pelo espectro estimado e pelos testes e modelos específicos para a sazonalidade estimados. Primeiramente, foi estimada a função de autocorrelação amostral dada pela equação (16), cujos resultados estão sintetizados no Quadro 3. Os coeficientes de autocorrelação com defasagens iguais a 12, 24 e 36, foram todos significativos, apontando a importância do fator sazonal na série analisada.

Quadro 3 - Resultados para a função de autocorrelação amostral

Autocorrelações	Correlações parciais	Def.	Coefficiente de autocorrelação	Coefficiente de autocorrelação parcial	Estatística Q	Valor P
**	-.**	1	0.275	0.275	9.2817	0.002
*	-.*	2	0.104	0.104	10.612	0.005
*	.*	3	0.177	0.152	14.513	0.002
*	**	4	-0.163	-0.277	17.851	0.001
**	.*	5	-0.229	-0.150	24.542	0.000
**	.*	6	0.202	0.358	29.804	0.000
*	**	7	-0.163	-0.298	33.263	0.000
*	.*	8	-0.177	-0.087	37.335	0.000
*	.*	9	0.141	0.156	39.970	0.000
*	.*	10	0.036	0.135	40.142	0.000
*	**	11	0.174	0.295	44.203	0.000
****	.*	12	0.629	0.370	97.847	0.000
**	.*	13	0.206	-0.105	103.65	0.000
..	..	14	0.060	-0.011	104.15	0.000
..	..	15	0.116	-0.082	106.03	0.000
**	.*	16	-0.229	-0.131	113.38	0.000
**	.*	17	-0.271	-0.062	123.83	0.000
*	..	18	0.085	-0.016	124.86	0.000
**	.*	19	-0.208	-0.084	131.14	0.000
**	.*	20	-0.260	-0.179	141.03	0.000
*	..	21	0.079	0.041	141.95	0.000
..	..	22	-0.039	0.016	142.18	0.000
*	..	23	0.089	0.064	143.39	0.000
****	..	24	0.462	0.031	175.95	0.000

Fonte: Resultados da pesquisa.

A mensuração dos efeitos, ou impactos, que os diferentes meses do ano exercem sobre o fluxo de visitantes ao Parque Nacional do Iguaçu, pôde ser efetuado por meio da estimativa da equação (18). Todavia ela foi estimada com onze variáveis dummy afim de evitar o problema de multicolinearidade. Isso porque o termo de intercepto do modelo se refere ao valor médio do impacto provocado pelo mês de janeiro (que será o mês de referência para a análise) no fluxo médio de turistas que foi de aproximadamente 67.453 turistas / mês ao longo da década analisada. Os resultados encontrados podem ser verificados no Quadro 4.

Os resultados encontrados para os índices sazonais indicam que o mês de janeiro é o mês que, em média, verifica-se o maior acréscimo no número de turistas no Parque Nacional do Iguaçu. De acordo com o coeficiente estimado neste mês o

número de turistas tenderá a aumentar, em média, 35.581 em relação à sua média histórica. Fevereiro também é um mês em que o volume médio de visitas é maior do que a média histórica anual em 3.961 visitantes. No mês de março tem início um período de quedas sistemáticas no fluxo de visitação ao parque com uma redução de 14.634 visitantes em relação a média, seguido por uma redução em função do mês de abril da ordem de 7.371 visitantes por mês. Em maio e junho tem-se os meses de menor volume de visitas ao parque de acordo com o modelo estimado, com quedas respectivas de 26.307 e 29.863 em relação à média na década.

Quadro 4 - Resultados dos testes de sazonalidade determinística mensal

Período/ Variável	Índice Sazonal	Valor de Probabilidade
Janeiro	32581.42	0.0000
Fevereiro	-28620.85	0.0000
Março	-47215.21	0.0000
Abril	-39952.30	0.0000
Maió	-58888.89	0.0000
Junho	-62445.50	0.0000
Julho	-12059.13	0.0341
Agosto	-31401.77	0.0000
Setembro	-31889.13	0.0000
Outubro	-24071.60	0.0000
Novembro	-23859.69	0.0000
Dezembro	-30572.99	0.0000
Estatística-F	20.1885	0.0000
R2 ajustado	0.6395	

Fonte: Resultados da pesquisa.

Por outro lado, em julho tem-se o mês com o segundo maior volume de visitas com uma elevação de 20.522 no número de turistas. Uma certa estabilidade no número de turistas em relação à sua média anual ocorre nos dois meses posteriores, agosto e setembro. Todavia nos três últimos meses do ano os índices sazonais apontam um comportamento de elevação no fluxo de visitantes ao parque da ordem de 8.510, 8.772 e 2.009 respectivamente.

6. Conclusões

Apesar das qualidades e potencial turístico apresentado pelo Parque Nacional do Iguaçu, pode-se verificar com a pesquisa realizada, que o fluxo de visitantes no parque apresentou um comportamento irregular ao longo da década de 90, em que não se identificou uma tendência expressiva de elevação da no número de visitantes nesse mesmo período. Isto pode indicar uma possível perda de atratividade do parque em relação à demanda turística, que pode estar se deslocando para outros parques, ou mesmo para outros tipos de roteiros. Isto, em última instância, poderia implicar em uma série de perdas para agentes econômicos que de maneira direta ou indireta tem suas rendas dependentes do turismo na região, e particularmente no Parque Nacional do Iguaçu.

Com relação ao componente cíclico no comportamento do fluxo de turistas no parque, os resultados da pesquisa confirmaram a incidência de um ciclo sazonal (12 meses de duração) e de três outros ciclos. Dois ciclos de curto-prazo, correspondentes a períodos com duração respectiva de 3 e 6 meses e um ciclo de longo-prazo correspondente a um período de 60 meses (5 anos) de duração.

Finalmente, em relação à análise da sazonalidade no fluxo de turistas, observou-se um forte padrão sazonal no mesmo. Os resultados encontrados indicam que o mês de janeiro é o mês que, em média, verifica-se o maior acréscimo no número de turistas no Parque Nacional do Iguaçu. Além de janeiro, os meses de fevereiro, julho, agosto, setembro, outubro, novembro e dezembro também se configuraram como meses de comportamento sazonal de elevação em relação à média de visitas. Sendo que nos demais meses o comportamento sazonal é de redução em relação à média.

7. Bibliografia

- ABRAHAM, Bovas.; LEDOLTER, Johannes. Statistical Methods for Forecasting. USA. John Wiley & Sons, Inc. 1983.
- BOWERMAN, Bruce L.; O'CONNEL, Richard T.; Time Series and Forecasting. California Duxbury Press. USA, 1979.
- BOX, George E.P.; JENKINS, Gwilym M.; Time Series Analysis, Forecasting and Control, Revised edition, USA, 1976.

- CAMPBELL, John Y.; LO, Andrew W.; MACKINLAY, A. Craig. *The Econometrics of Financial Markets*. New Jersey, Princeton University Press: USA, 1997.
- CHATFIELD, C. *The Analysis of Time Series - An Introduction*. 5a ed. London: Chapman & Hall. UK 1996.
- CRAMER, Harald; LEADBETTER, M.R. *Stationary and Related Stochastic Processes*. New York, Wiley. USA, 1967.
- ENDERS, Walter. *Applied Econometric Time Series*, 1a ed. John Wiley & Sons: Inc. USA, 1995.
- HAMILTON, J. D., *Time series analysis*. New Jersey: Princeton University Press Princeton, USA 1994.
- JUDGE, G.G. et al. *Introduction to the Theory and Practice of Econometrics*. 2ª ed, New York, Wiley, USA 1988.
- PINDYCK, Robert S.; RUBINFELD, Daniel L., *Econometric Models and Economic Forecasts*. 4ª ed. Irwin/McGraw-Hill, USA. 1998.
- RAUSSER, Gordon C.; CARGILL, Thomas F.; *The Existence of Broiler Cycle: An Application of Spectral Analysis*. *American Journal of Agricultural Economics*, 52 (1): 109-21, USA, 1970.
- SCHWAGER, Jack D., *Fundamental Analysis*. John Wiley & Sons. USA, 1994.
- SOUZA, Reinaldo Castro, *Modelos Estruturais Para Previsão de Séries Temporais: Abordagens Clássica e Bayesiana*. 17o Colóq. Bras. de Matemática, IMPA. RJ. 1989.
- VANDAELE, Walter. *Applied Time Series and Box-Jenkins Models*. Academic Press USA, 1983.