

Estatística apoiada pela tecnologia: uma proposta para identificar tendências climáticas

Júlio César Penereiro
Denise Helena Lombardo Ferreira

RESUMO

O objetivo deste trabalho é envolver e motivar um grupo de estudantes de graduação no uso de alguns conteúdos estatísticos envolvendo testes paramétricos e não paramétricos. Faz-se um estudo das variações da temperatura média e da precipitação pluviométrica para identificar possíveis tendências climáticas. Foram empregados análises de regressão e os testes de Mann-Kendall e Pettitt em dados medidos para a cidade de Campinas no período de 1910 a 2009. Os resultados indicam que não houve tendência na variável precipitação, mas foi identificado um significativo aumento da temperatura média que, possivelmente, está relacionada à influência antrópica iniciada a partir de 1955. Os estudos resultaram em várias discussões entre os estudantes e os pesquisadores envolvidos, especialmente sobre como detectar tendência a uma série de dados e decidir por qual teste estatístico empregar. Dessa maneira, constatou-se que a pesquisa atendeu algumas das expectativas dos estudantes, como vincular os testes estatísticos utilizados com a realidade do cotidiano aliados à tecnologia.

Palavras-chave: Ensino de estatística. Testes estatísticos. Educação ambiental.

Statistics supported by technology: A proposal to identify climate trends

ABSTRACT

The goal of this work is to engage and motivate a group of undergraduate students in the use of some content involving statistical parametric and nonparametric tests. A study is made of variations in average temperature and precipitation to identify possible trends in climate. It has been utilized regression analysis and the Mann-Kendall and Pettitt tests on measured data for the city of Campinas from 1910 to 2009. The results indicate that there was no trend in the variable precipitation, but was identified significant trend of increasing temperature is, possibly, related to human influence started from 1955. The studies resulted in many discussions among students and researchers involved, especially on how to detect a range of trend data and decide which statistical tests to employ. In this way, we meet some of the students' expectations, linking the statistical concepts with the reality of daily life combined with technology.

Keywords: Teaching statistics. Statistical tests. Environmental education.

Júlio César Penereiro é Doutor em Astrofísica (IAG-USP, 1993), professor da Faculdade de Matemática e eng. Ambiental do CEATEC/PUC-Campinas. Endereço para correspondência: Rod. D. Pedro I, Km 136 – Parque das Universidades, Campinas (SP) – CEP 13086-900. E-mail: jcp@puc-campinas.edu.br

Denise Helena Lombardo Ferreira é Doutora em Educação Matemática (UNESP/Rio Claro, 2000) e professora da Faculdade de Matemática do CEATEC/PUC-Campinas. Endereço para correspondência: Rod. D. Pedro I, Km 136 – Parque das Universidades, Campinas/SP – CEP 13086-900. E-mail: lombardo@puc-campinas.edu.br

Acta Scientiae	Canoas	v. 13	n.1	p.87-105	jan./jun. 2011
----------------	--------	-------	-----	----------	----------------

INTRODUÇÃO

Alguns projetos desenvolvidos com estudantes em cursos universitários às vezes associam aplicações da Modelagem Matemática como uma maneira de motivá-los a apreender e/ou empregar diferentes conteúdos matemáticos. A esse respeito, Barbosa (2006) comenta que a Modelagem Matemática é um ambiente de aprendizagem em que os estudantes são convidados a investigar situações e problemas extraídos do dia a dia, inclusive envolvendo assuntos relacionados com outras ciências. Isso pode despertar nos estudantes um interesse pelos conteúdos matemáticos, além de valorizar o senso crítico dos mesmos.

A literatura apresenta diferentes abordagens para a utilização pedagógica da Modelagem Matemática, destacando-se: a *perspectiva pragmática*, que busca habilitar os estudantes a usarem a Matemática para a solução de problemas reais, priorizando o conhecimento técnico; a *perspectiva científica-humanística*, que procura estabelecer relações com outras áreas a partir da própria Matemática, favorecendo o conhecimento matemático (KAISER; SRIRAMAN, 2006) e a *perspectiva sociocrítica*, que visa criar oportunidades para que os alunos reflitam sobre o papel da matemática na sociedade (BARBOSA, 2006).

É no emprego da perspectiva sociocrítica que a Modelagem Matemática se alavanca como uma estratégia pedagógica que deve contemplar outros interesses que extrapolam os aspectos puramente matemáticos. Atuando dessa forma, talvez seja possível sensibilizar os estudantes para algumas questões sociais, econômicas, ambientais, políticas, culturais, etc., motivando-os a pensar e agir. Isso vai ao encontro do que Skovsmose (2008) destaca como a importância de formar um estudante crítico, investigador, questionador, consciente dos problemas que afligem a sociedade e, além disso, interessado em contribuir na busca de soluções para tais problemas.

Quando se propõe um projeto de investigação com dados reais, de fontes diversas, com uso de equipamentos, dados de publicações ou da Internet, dentre outros, possibilita-se aos estudantes envolvidos apreciar a importância de se usar conteúdos estatísticos e, conseqüentemente, se interessar pela Estatística como meio de abordar problemas variados da vida real, que poderão estar relacionados às suas experiências e necessidades. Sendo assim, é necessário contribuir para o enriquecimento de práticas pedagógicas que valorizem os trabalhos em grupos, as atividades exploratórias e de investigação, além de fomentar o gosto pela resolução de problemas, bem como o uso de instrumentos e metodologias inovadoras para uma participação ativa dos estudantes, favorecendo as reflexões críticas e as discussões.

A Estatística ocupa um papel fundamental, porque, além de oferecer vários recursos matemáticos que facilitam a realização de um trabalho pedagógico ou uma pesquisa científica, também pode ser útil para o desenvolvimento de uma postura investigativa, reflexiva e crítica do pesquisador, auxiliando-o em possíveis tomadas de decisões.

Constata-se ainda que, em particular, no ensino da Estatística há uma grande relutância por parte dos professores em usar *softwares* aplicados a essa ciência, porém, o

mais grave é que em quase todas as Escolas, sejam particulares ou públicas, o ensino da Estatística é deixado em segundo plano. Nas grades curriculares da maioria dos cursos superiores na área de Ciências Exatas, é fácil perceber a pouquíssima carga horária dedicada para essa disciplina. Em muitos casos, o que se verifica é uma extensa ementa de assuntos a serem abordados pela disciplina, sem levar em conta um aprendizado crítico e contextualizado.

O estudo aqui relatado visa dar uma contribuição na área do ensino da Estatística, transmitindo uma experiência extraclasse com um grupo de cinco estudantes do 4º semestre de um curso de Engenharia Ambiental, que se mostrou interessado em trabalhar com um tema relacionado às mudanças climáticas. Apesar da experiência dos docentes envolvidos, vários desafios apareceram ao longo do estudo, pois o tema, apesar de excitante e atual, foi uma novidade na prática docente, o que resultou em várias pesquisas bibliográficas e consultas a outros pesquisadores da área em questão. No entanto, o aspecto mais importante desse trabalho foi a busca por conteúdos matemáticos, estatísticos e computacionais para estudar as variações das temperaturas e precipitações pluviais e, com isso, identificar possíveis tendências nessas duas grandezas climáticas. Como decorrência desse procedimento, foi possível motivar os estudantes e, ao mesmo tempo, envolvê-los com métodos estatísticos aplicando os dados climáticos a testes paramétricos e não-paramétricos. Certamente, alguns testes aqui abordados não são contemplados nas ementas das disciplinas de Estatística no Ensino Superior, resultando na elaboração de muitas questões e discussões sobre o emprego dos conteúdos estatísticos explorados, além dos resultados encontrados pelas interpretações relativas às questões ambientais.

Neste contexto, o trabalho teve como objetivo envolver estudantes de um curso superior num estudo exploratório, que proporcionou a organização e análise de dados empregando conteúdos matemáticos e estatísticos para a verificação das tendências da temperatura média e da precipitação anual na cidade de Campinas, induzindo-os para a formação de uma postura crítica perante aos problemas que enfrentarão no ambiente de trabalho profissional. Para tanto, foi necessário usar uma base de dados históricos dessas grandezas climáticas existente na literatura. Nas seções subsequentes é discutida a metodologia empregada no tratamento e análise dos dados, em seguida são feitas as aplicações e análises dos testes estatísticos paramétricos e não-paramétricos e conclui-se realizando uma abordagem das perspectivas desta pesquisa.

METODOLOGIA NO TRATAMENTO DE DADOS E ANÁLISES EXPLORATÓRIAS

Para atingir os objetivos referidos consideraram-se alguns parâmetros climatológicos, como as temperaturas (mínimas, médias e máximas) e a precipitação pluviométrica (índice de chuvas) de uma determinada localidade, pois, segundo Ayoade (1996), esses são os índices mais frequentemente usados para caracterizar o clima de uma determinada região. Além da temperatura média do ar e da precipitação pluviométrica, é fato que outras informações são relevantes para caracterizar e estudar o clima de uma região,

dentre as quais a velocidade do vento, a radiação solar e a umidade relativa do ar. Porém, neste trabalho se optou por explorar conteúdos estatísticos somente nos dois parâmetros comentados anteriormente.

Na literatura, é frequentemente citado que uma série possui certa tendência quando há uma mudança sistemática e contínua em qualquer parâmetro de uma dada amostra, excluindo-se mudanças periódicas ou quase periódicas (YEVJEVICH, 1972). A tendência climática é uma mudança do clima caracterizada por um contínuo acréscimo ou decréscimo nos valores da variável observada num determinado período (GOOSSENS; BERGER, 1986).

Com o apoio dos professores envolvidos neste trabalho, o grupo de estudantes realizou um levantamento e efetuou análises dos dados exploratórios a partir de um banco de dados climáticos investigados no endereço eletrônico do Instituto Agrônomo de Campinas (IAC). Esse endereço disponibiliza a quaisquer interessados várias informações, inclusive climatológicas, em formato de tabelas, mapas e gráficos (IAC/ CIIAGRO, 2009).

Os dados de temperatura média e da precipitação total mensal e anual utilizados foram coletados no período de 1910 a 2009, sendo que todas as informações foram organizadas e trabalhadas em planilhas do programa Microsoft Excel. Essa conduta possibilitou tanto aos professores como aos estudantes se envolverem com essa importante ferramenta computacional de fácil acesso, utilizando diversos conteúdos estatísticos em atividades de ensino e pesquisa.

Para cada variável trabalhada foi realizada uma análise de regressão linear aplicada aos pontos das médias móveis. Essas médias foram calculadas usando ordem 5, por intermédio da seguinte relação (SELLO, 1999):

$$\bar{V}_n = \frac{1}{6} \left[\sum_{i=n-2}^{n+2} V_i + \frac{1}{2} (V_{n+3} + V_{n-3}) \right] \quad (1)$$

onde: (\bar{V}_n) representa o valor médio da variável climática em estudo para o i -ésimo mês (V_i). A escolha da eq. [1] foi motivada pelo fato de permitir evitar possíveis flutuações impostas pelos dados. Dessa maneira, trabalha-se com uma suavização desses dados e, assim, pode-se traçar uma linha de tendência da variável em análise com boa precisão. Para identificar a hipótese de ruptura das séries históricas, foram realizadas análises aplicando testes estatísticos não paramétricos de Mann-Kendall (SNEYERS, 1975; SNEYERS et al., 1990) e de Pettitt (PETTITT, 1979). Em particular, esses testes não paramétricos também foram empregados de forma similar por outros pesquisadores em trabalhos específicos voltados à climatologia da região sul do Brasil (BACK, 2001) e ao comportamento hidrológico de rios (PELLEGRINO et al., 2001; MORTATTI et al., 2003; GROppo et al., 2005), porém, os embasamentos expostos nesses trabalhos não foram elaborados com o intuito de despertar o interesse exploratório do ensino e aprendizagem

dos conteúdos estatísticos embutidos em cada teste e, tão pouco, foram aplicados a um grupo de estudantes do ensino superior procurando motivá-los ao desenvolvimento de uma postura investigativa, reflexiva e crítica, auxiliando-os em possíveis tomadas de decisões. Com tais finalidades, são apresentados a seguir os testes estatísticos empregados neste trabalho.

Análise de regressão

Para que o grupo de estudantes pudesse detectar uma possível tendência na base de dados históricos de uma variável climatológica, após a organização dos dados numa planilha, empregou-se o cálculo das médias móveis para fazer a aplicação da análise de regressão linear. Do ponto de vista matemático, e pelo aspecto didático com que foi trabalhado com os estudantes, definiu-se a série de dados pela variável Y (no caso, temperatura ou precipitação), no tempo t (ano). Assim, Y é uma função de t representada por $Y = f(t)$.

Nesse procedimento, como de conhecimento prévio por parte dos membros do grupo, utilizou-se a equação da reta de tendência ($y = a(x) + b$), onde a é o coeficiente angular da reta e, no caso, $x = t$. Assim, $a > 0$ significa uma tendência crescente e $a < 0$ indica uma tendência decrescente. Ao efetuar esse procedimento, o valor da estatística do R-quadrado (R^2) é devolvido como uma medida da qualidade do ajuste efetuado. Essa informação foi explorada com os estudantes, alertando-os para a importância do valor de R^2 ser o mais próximo possível da unidade, garantindo um bom ajuste e a verificação de uma menor flutuação dos valores dos dados tratados (FREUND, 2006).

Teste de Mann-Kendall

Para verificar se realmente uma série histórica de dados possui tendência, ou se a mesma está mascarada pela eventual flutuação da base de dados, foi necessário que os estudantes realizassem outros testes estatísticos mais sensíveis à identificação de tendência.

O teste não paramétrico sequencial de Mann-Kendall, proposto inicialmente por Sneyers (1975), é o método mais apropriado para analisar mudanças climáticas e, segundo Goossens e Berger (1986), ele permite detectar e localizar de forma aproximada o ponto inicial de determinada tendência. Neste teste, considera-se que, na hipótese de estabilidade de uma série de dados, a sucessão de valores ocorre de forma independente e a distribuição de probabilidade deve permanecer sempre a mesma (série aleatória simples).

Moraes et al. (1995) e Back (2001) descrevem esse método considerando uma série temporal de Y_i de N termos a ser analisada ($1 \leq i \leq N$). Segundo esses autores, esse teste consiste na soma t_n do número de termos m_i da série, relativo ao valor Y_i cujos termos precedentes ($j < i$) são inferiores ao mesmo ($Y_j < Y_i$), isto é: $t_n = \sum_{i=1}^N m_i$. Assim,

orientou-se o grupo de estudantes para que, ao trabalharem com séries envolvendo um grande número de termos (N), usassem a hipótese nula (H_0) de ausência de tendência. Dessa forma, t_n apresentará uma distribuição normal com média e variância dadas, respectivamente, por (BACK, 2001; MORTATTI et al., 2003; GROPPPO et al., 2005):

$$E(t_n) = \frac{N(N-1)}{4} \quad (2)$$

e

$$\text{var}(t_n) = \frac{N(N-1)(2N+5)}{72} \quad (3)$$

Testando a significância estatística de t_n para a hipótese nula, usando um teste bilateral, esta pode ser rejeitada para grandes valores da série estatística $U(t_n)$ por meio da equação dada por:

$$U(t_n) = \frac{[t_n - E(t_n)]}{\sqrt{\text{var}(t_n)}} \quad (4)$$

Após a determinação dos valores de $U(t_n)$, foi necessário verificar o valor da probabilidade α_1 , para que $\alpha_1 = \text{prob}(|U| > |U(t_n)|)$, calculada por meio de uma tabela da normal reduzida. Nesse momento, verificou-se a necessidade de explicar e mostrar aos estudantes a importância da determinação do nível de significância do teste estatístico empregado (OLIVO; BATANERO, 2007). Assim, admitindo que se α_0 é o nível de significância do teste aplicado ($\alpha_0 = 0,05$ e $0,10$) para significativo e levemente significativo, respectivamente), a hipótese nula será aceita se $\alpha_1 > \alpha_0$. Caso a hipótese nula seja rejeitada, implicará a existência de tendência significativa, sendo que o sinal da estatística $U(t_n)$ indica se a tendência é crescente ($U(t_n) > 0$) ou decrescente ($U(t_n) < 0$).

O ponto de início de uma mudança na série $U(t_n)$ pode ser determinado aplicando-se o mesmo princípio à série inversa (BACK, 2001). Neste caso, em sua versão sequencial, a eq. [4] é calculada no sentido direto da série, partindo do valor $i = 1$ até $i = N$, gerando a estatística $-1,96 < U(t_n) < 1,96$ (sendo que $1,96$ corresponde a $\alpha_0 = 0,05$), e, no sentido inverso da série, partindo do valor $i = N$ até $i = 1$, gerando a estatística $U^*(t_n)$. A intersecção das duas curvas $U(t_n)$ e $U^*(t_n)$ é onde se localiza o ponto aproximado de mudança de tendência, se esse ponto ocorre dentro do intervalo de confiança $-1,96 < U(t_n) < 1,96$.

Teste de Pettitt

Além do teste de Mann-Kendall, decidiu-se também aplicar o teste não paramétrico de Pettitt (PETTITT, 1979; MORAES et al., 1995). Assim como relatou Back (2001), esse

é um teste que utiliza uma versão do teste de Mann-Whitney, em que se verifica se duas amostras Y_1, Y_2, \dots, Y_t e $Y_{t+1}, Y_{t+2}, \dots, Y_T$ são ou não da mesma população (FREUND, 2006).

Aplicando o mesmo procedimento efetuado por Back (2001), recomendou-se ao grupo de estudantes que, para a estatística $U_{i,T}$, se fizesse uma contagem do número de vezes que um membro da primeira amostra é maior que um membro da segunda, determinando essa situação pela condição escrita através da relação:

$$U_{i,T} = U_{i-1,T} + \sum_{j=1}^T \text{sgn}(Y_i - Y_j); \text{ para } t = 2, \dots, T; \quad (5)$$

onde: $\text{sgn}(x) = 1$ para $x > 0$; $\text{sgn}(x) = 0$ para $x = 0$ e $\text{sgn}(x) = -1$ para $x < 0$.

A partir disso, a estatística $U_{i,T}$ foi calculada para valores de $1 \leq t \leq T$, sendo que a estatística $K(t)$ do teste fornece o máximo valor absoluto de $U_{i,T}$, por meio da equação $K_T = \max_{1 \leq t \leq T} |U_{i,T}|$ (PETTITT, 1979).

A estatística assim desenvolvida localiza o ponto em que houve uma mudança brusca na média de uma série, o que tornou tarefa fácil para um estudante do grupo verificar essa situação através da planilha de dados do Excel. Associado a isso, foi também possível para ele verificar e avaliar o nível de significância desse teste através da equação (BACK, 2001):

$$p \cong 2e \left(\frac{-6(K(t))^2}{(T^3 + T^2)} \right) \quad (6)$$

Por meio dos dados tratados no Excel, os estudantes constataram que o ponto de mudança brusca é o valor para t onde ocorre o máximo de $K(t)$. Por meio da eq. [6] também foi possível inferir os valores críticos de K através da relação:

$$K_{crit.} = \pm \sqrt{\frac{-\ln(p/2)(T^3 + T^2)}{6}} \quad (7)$$

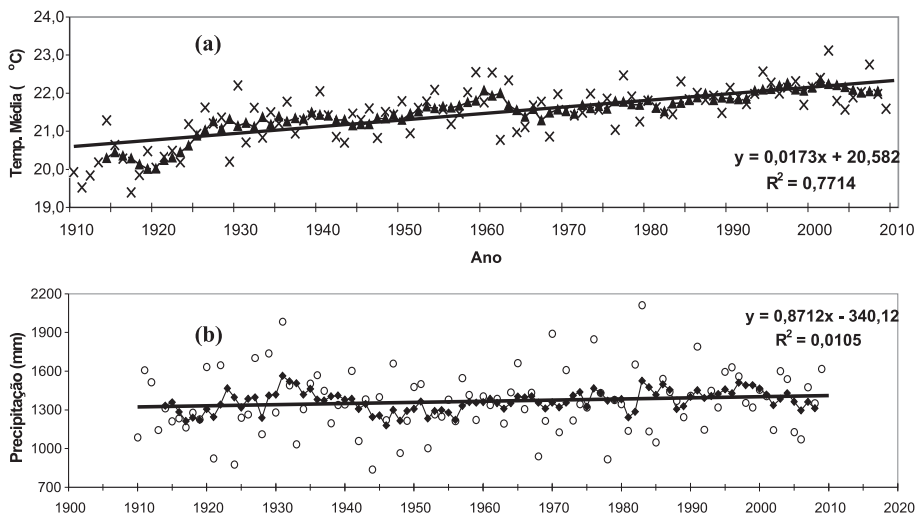
Assim, como no teste de Mann-Kendall, os níveis de significâncias do teste estatístico de Pettitt foram estimados para os intervalos de confiança em 5% e 10%, o que corresponde aos valores de desvios 1,96 e 1,65, respectivamente (FREUND, 2006).

RESULTADOS E DISCUSSÕES

Após a organização dos dados em planilhas do Excel os estudantes realizaram cálculos para determinar as médias móveis visando os ajustes das linhas de tendências e as

análises das regressões lineares. Para essas tarefas utilizaram os seguintes procedimentos do Excel: *Gráfico*, opção em *Dados de Origem...* e *Adicionar linha de tendência...*. Na sequência, obtiveram os gráficos das Figuras (1.a) e (1.b) para a temperatura média e a precipitação pluvial, respectivamente, no intervalo de tempo avaliado (1910 a 2009). Como se tornou fácil de constatar por meio desses gráficos, a Figura (1.a) mostrou uma tendência de aumento para a temperatura média anual da ordem de 1,8°C, enquanto que a Figura (1.b) indicou que a precipitação anual não sofreu tendência considerável no período. Além disso, foi possível perceber, por intermédio das análises desses dois gráficos, as flutuações (ou dispersões) dos valores anuais das variáveis climáticas trabalhadas em curtos períodos, assim como os diferentes valores de R^2 indicando as qualidades dos ajustes efetuados.

FIGURA 1 – Regressão linear aplicada aos pontos das médias móveis. (a) Tendência da temperatura média do ar em triângulos pretos e interligados, as cruzes indicam os valores anuais. (b) Tendência da precipitação pluvial em pontos pretos e interligados, os círculos são os valores anuais.



Para obterem o intervalo de confiança em 95% do valor estimado do coeficiente angular da reta de regressão, os estudantes utilizaram os seguintes passos do Excel: *Ferramentas, Análise de dados..., Regressão...* Ao abrirem a janela do programa, como mostrada na Figura 2, foram inseridos os *Intervalos do eixo Y e X*, e selecionou-se o *Nível de confiança* em 95% e o *Intervalo de Saída* do arquivo trabalhado. Posteriormente, o *software* retornou um *Resumo dos Resultados* com o intervalo de *confiança inferior e superior* (também contido na Figura 2). A Tabela 1 resume os resultados da aplicação desse procedimento para os parâmetros climáticos trabalhados pelos estudantes.

FIGURA 2 – Síntese dos procedimentos realizados pelos estudantes para encontrar o intervalo de confiança de 95% através do software Excel

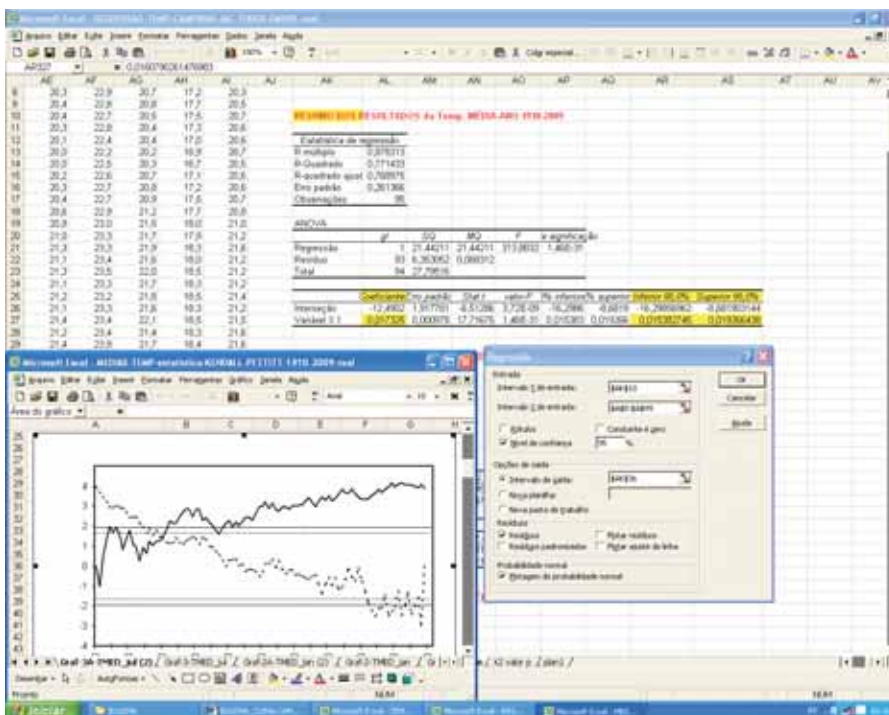


TABELA 1 – Resultados das análises de regressão linear realizada pelo grupo de estudantes.

Variável analisada	Coefficiente angular	R2	Interv. de confiança (95%)
Temperatura média anual	0,0173	0,7714	0,01538 a 0,01927
Precipitação total anual	0,8712	0,0105	0,29540 a 1,44702

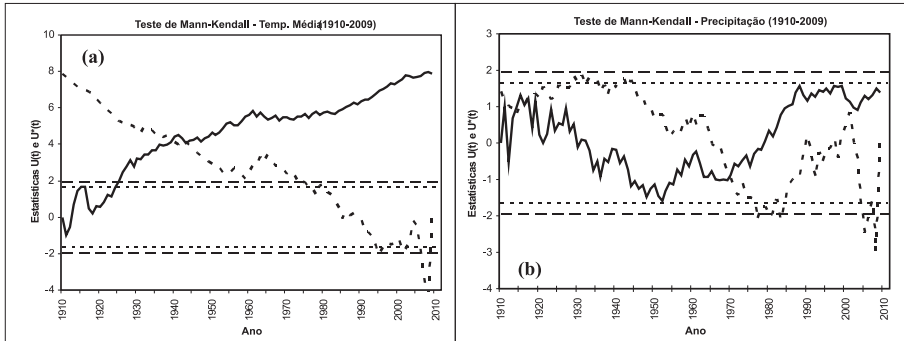
Assim como descreveram Olivo e Batanero (2007) sobre a importância dos estudantes saberem interpretar o intervalo de confiança em diferentes tipos de amostras, esses autores identificaram que os equívocos cometidos pelos estudantes estão mais relacionados à interpretação do intervalo de confiança como uma estatística descritiva, ignorando a natureza inferencial do problema. Esses autores também observaram que há uma deficiência conceitual na compreensão da maneira como se relacionam os diferentes conceitos associados com o intervalo de confiança. Muitas vezes esses equívocos parecem associados com o entendimento da utilidade dos intervalos de confiança, pois eles devem ser encarados como uma tomada de decisão sobre a hipótese definida *a priori*. Neste aspecto, o trabalho de Olivo e Batanero (2007) corrobora com a pesquisa aqui apresentada, tendo em vista que os estudantes envolvidos no trabalho aqui discutido possuíam muitas

deficiências em conceitos básicos de Estatística, assim como na compreensão a respeito do intervalo de confiança de uma amostra. Assim, algumas perguntas foram formuladas pelos próprios estudantes sobre esse tema, nos seguintes aspectos: Como, do ponto de vista matemático, calcular um intervalo de confiança para uma amostra de dados? Para que serve o intervalo de confiança calculado numa análise?

Foi a partir deste ponto que os estudantes foram alertados e esclarecidos que, embora as análises de regressão mostradas anteriormente por meio de cálculos e gráficos representem as possibilidades de tendências das séries de dados estudados, os intervalos de confiança empregados para esse tipo de análise se tornam limitados para tomar a decisão se a amostra sofre ou não tendência. Por conseguinte, análises mais detalhadas e adequadas podem ser realizadas, complementarmente, empregando-se outros testes estatísticos não paramétricos, como os testes de Mann-Kendall e Pettitt. Foi esclarecido ao grupo de estudantes que essa limitação se deve, principalmente, devido a significância estatística por meio da simples regressão linear para a suposta tendência da série de dados.

Para executar os testes não paramétricos, foram desenvolvidos diversos procedimentos por meio de planilhas do Excel, empregando-se diferentes recursos de funções, de lógicas, de estatística e de gráficos desse *software*, tais como no ícone f_x : *Média*, *Desvpad*, *Cont.se*, *Procv*, *Cont.núm*, *se*, *Tipo de tendência*, *Exibir equação no gráfico*, *Exibir valor de r-quadrado no gráfico*, dentre outras. As Figuras 3 e 4 apresentam as formas gráficas obtidas para esses testes nos parâmetros temperatura média e precipitação pluviométrica, respectivamente. Nelas, os níveis de significância são mostrados nos gráficos por meio de retas horizontais, sendo de 5% com traçados contínuos e de 10% com traçados pontilhados. Assim, como lembraram os estudantes por intermédio de suas leituras na literatura, no teste de Mann-Kendall a tendência é significativa quando os valores absolutos de $U(t_n)$ são maiores que os níveis de significância e o início dessa tendência pode ser localizado pela interseção das curvas $U(t_n)$ (em traçado contínuo) e $U^*(t_n)$ (em traçado pontilhado) nas Figuras (3), porém isso deve ocorrer dentro dos valores críticos do intervalo de confiança (BACK, 2001). Por outro lado, verificou-se que no teste de Pettitt (Figura 4) o ponto de mudança brusca de $K(t)$ (em módulo) ocorre quando este for maior que o limite crítico estabelecido. No entanto, essa condição deixa de ser verdadeira quando os valores que estão em seguida ao valor crítico oscilam em intervalos próximos ao valor máximo. Neste caso, o último valor do intervalo de oscilação indicou o ponto de início da tendência (PETTITT, 1979).

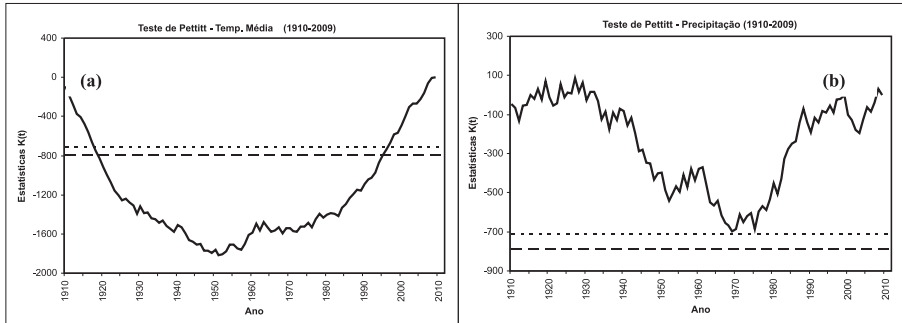
FIGURA 3 – Estatísticas $U(t_n)$ (curva contínua) e $U^*(t_n)$ (curva tracejada) do teste Mann-Kendall para (a) Temperatura média anual e (b) Precipitação pluvial anual. As linhas horizontais representam os níveis de significância de 5% (traçado contínuo) e 10% (traçado pontilhado).



Através de uma rápida inspeção nos gráficos da temperatura média, os estudantes perceberam que tanto o teste de Mann-Kendall (Figura 3.a) como o de Pettitt (Figura 4.a) aplicados a essa variável climática não apresentaram, em princípio, concordância com a análise de regressão linear. Eles constataram que, tanto pelo gráfico da regressão linear (Figura 1.a), como pelos valores do coeficiente angular $a = 0,0173$ (indicando quase nenhuma tendência da série) e o valor de $R^2 = 0,7714$ (mostrado na Tabela 1 e indicando uma boa qualidade do ajuste), houve uma leve tendência para o aumento da temperatura no período estudado, porém, eles também perceberam que o teste Mann-Kendall indica, através das curvas $U(t_n)$ e $U^*(t_n)$, que houve um aumento de tendência da série temperatura e que elas estão além dos limites de significância de 5% e 10% (Figura 3.a).

Como também foi identificado pelos estudantes através das análises dos gráficos, o teste de Pettitt apontou uma “quebra” de tendência da temperatura média em 1952 (quando essa tendência torna-se mais evidente, pois foi onde se verificou o valor crítico ($K_{crit.}$) para esse teste) e apresentou um outro pico para baixo em 1957 (Figura 4.a). Assim, o grupo de estudantes envolvidos nessa análise pode concluir que, efetivamente, a variável climática temperatura média sofreu tendência de aumento a partir de meados da década de 1950.

FIGURA 4 – Estatísticas $K(t)$ do teste Pettitt para (a) Temperatura média anual e (b) Precipitação pluvial anual. As linhas horizontais representam os níveis de significância de 5% (traçado contínuo) e 10% (traçado pontilhado).



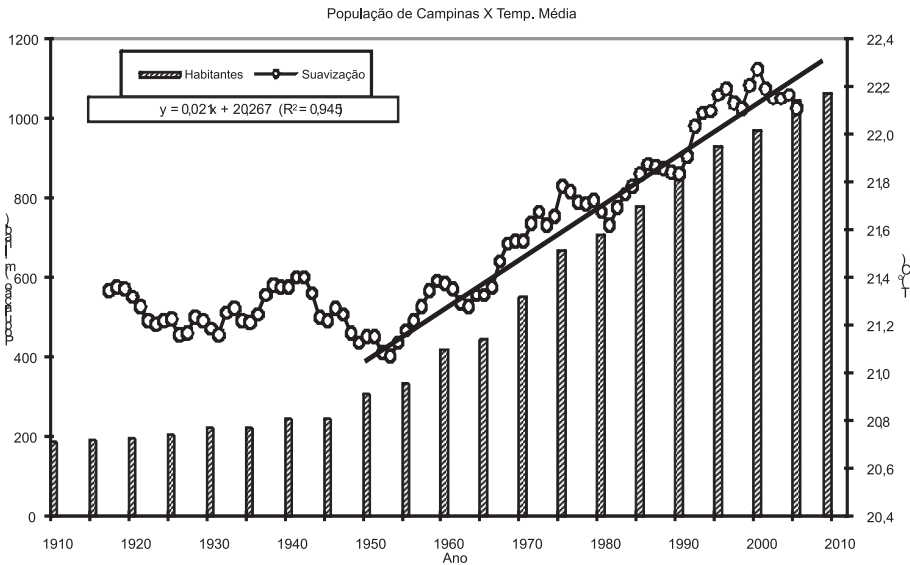
No caso dos dados de precipitação, o teste de Pettitt não apresentou concordância com a análise de regressão linear. Os estudantes notaram que, apesar da regressão apresentar um valor de coeficiente angular $a = 0,8712$, indicando uma leve tendência da série, o valor de $R^2 = 0,0105$ mostra que a qualidade do ajuste fica comprometida devido à elevada dispersão (ou flutuações) dos dados (Figura 1.b). Essa dispersão foi verificada quando se analisou o intervalo de confiança de 95% que, neste caso, ficou entre 0,29540 a 1,44702, isto é, um intervalo muito grande (Tabela 1). Quanto ao teste de Pettitt, não foram identificadas mudanças significativas para a série, visto que os valores das estatísticas $K(t)$ em nenhum momento cruzaram os intervalos de confiança (Figura 4.b). Já para o teste de Mann-Kendall, apesar das curvas estatísticas de $U(t_n)$ e $U^*(t_n)$ estarem entre os limites de significância de 5% e 10% e se cruzarem em 1972 (Figura 3.b), os estudantes perceberam que elas não apresentam tendências no período analisado.

Através da rápida análise exposta acima, deve-se salientar que foi mais fácil para os professores envolvidos nessa pesquisa alertar e, ao mesmo tempo, justificar aos membros do grupo de estudantes que a tendência não pode ser reconhecida e indicada apenas pelo valor do coeficiente angular (a) e pelo o gráfico da análise de regressão linear. Observou-se que existe um vínculo entre os dados trabalhados e a dispersão medida (R^2) para a série desses dados, além do comprometimento dos resultados encontrados através dos testes não paramétricos de Mann-Kendall e Pettitt. Foi fundamental advertir e explicar aos estudantes que, se a probabilidade (α_1) do teste de Mann-Kendall for menor que o nível de significância α_1 , uma tendência estatisticamente significativa existe, enquanto que se um valor de α_1 for maior que α_0 , confirma-se uma tendência insignificante para a série estudada. Vale ainda ressaltar que para a série onde não há tendência (um valor de $a \approx 0$), o valor de α_1 é próximo de α_0 . No tocante ao teste de Pettitt, para que a tendência seja reconhecida pelos próprios estudantes, os dados estatísticos de $K(t)$ tiveram, pelo menos em algum momento, que cruzar os intervalos de confiança, como ocorreu com a série analisada da Figura 4.a.

Resumindo, os estudantes puderam identificar por meio dessas análises estatísticas que não houve tendência significativa de variação na precipitação, mas, sim, tendência de aumento real da temperatura média. Essa foi uma das questões que os próprios estudantes acabaram por concluir nos gráficos gerados pelos testes estatísticos aqui apresentados. Talvez o resultado mais interessante dessa pesquisa esteja relacionado a fato que essa conclusão corrobora com a hipótese levantada por Fontana e Berlato (1997), pois reforça a ideia de que a mudança na precipitação é apenas uma oscilação de causa natural e não está necessariamente associada a uma mudança definitiva nos padrões pluviométricos globais. Quanto à temperatura, a tendência de aumento identificada nas análises acima, segundo opinião de dois estudantes do grupo, pode estar relacionada ao próprio crescimento urbano da cidade.

Discussões posteriores entre os estudantes envolvidos e os autores desta pesquisa permitiram verificar que os membros do grupo tornaram-se mais atentos às questões climatológicas vinculadas pelas mídias. Como exemplo disso, e dando sequência à pesquisa, por iniciativa própria um dos estudantes sugeriu outro desafio aos participantes: a possibilidade de averiguar se existe correlação entre o acréscimo da população da cidade com outras variáveis, tais como variação da temperatura média, aumento da mancha de ocupação do solo e o comportamento da temperatura média global da Terra. Para que fosse possível realizar essa tarefa, os próprios estudantes fizeram uma ampla pesquisa bibliográfica relacionada à população de Campinas, com dados extraídos do Instituto Brasileiro de Estatística e Geografia (IBGE, 2010). Eles também utilizaram os dados das temperaturas médias aqui discutidas, além daqueles relacionados ao desvio da temperatura média global do planeta, disponíveis no endereço eletrônico da Unidade de Pesquisas Climáticas da Universidade de East Anglia, na Inglaterra (CRU/UEA, 2010). Assim, a partir da organização das informações encontradas, foram gerados os gráficos contidos nas Figuras 5 e 6. Infelizmente, não foi possível obter nenhuma informação coerente e confiável em relação ao crescimento da área territorial ocupada pelo município com o passar do tempo.

FIGURA 5 – Evolução do número de habitantes (barras verticais) e da temperatura média (círculos abertos interligados) em Campinas no período de 1910 a 2009. Reta contínua indica a tendência da evolução da temperatura entre 1955 a 2009.

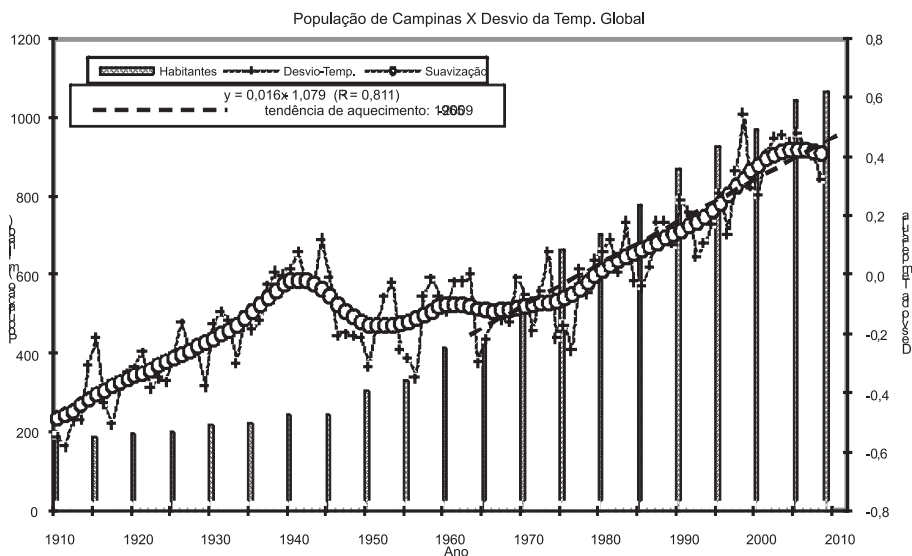


Uma análise realizada pelos estudantes, por meio das informações contidas na Figura 5, revelou que os dados do histograma em forma de barras, que representa o comportamento da população de Campinas entre 1910 a 2009 (IBGE, 2010), concordam com as medidas das temperaturas médias anuais (suavizadas através do uso da eq. [1]) no período da série histórica trabalhada (dados em círculos abertos). Além disso, verificou-se que, por meio de uma rápida inspeção do gráfico, a dificuldade em concluir que o aumento da população tenha contribuído com um aumento real da temperatura local (OKE, 1987). Entretanto, constatou-se que entre 1917 a 1950 houve uma pequena oscilação da temperatura média ($\pm 0,3^{\circ}\text{C}$), o que pode ser interpretado como ocorrência de nenhuma tendência nesse intervalo de tempo. Entretanto, após o ano 1955 essa variável climática passou a ter outro comportamento, indicando uma clara tendência de aumento da temperatura média, corroborando com a avaliação feita anteriormente através do teste de Pettitt (Figura 4.a). Por meio de um ajuste linear aplicado apenas aos dados da temperatura entre 1955 a 2009 (linha contínua da Figura 5), foi possível inferir o acréscimo de $1,13^{\circ}\text{C}$ na temperatura média do município nesse intervalo de tempo.

De modo análogo, quando foram comparados os dados do aumento populacional com o desvio (ou anomalia) da temperatura média global do planeta (dados suavizados pela eq. [1]) no período da série histórica aqui trabalhada, os estudantes verificaram que a partir de 1965 iniciou-se uma correlação entre essas duas informações (Figura 6). Entretanto, o desvio absoluto inferido por eles para a temperatura média global entre 1965

a 2009 foi da ordem de 0,56°C. Portanto, inferior ao obtido por meio da temperatura média local suavizada com os dados do IAC (Figura 5), em que pese os diferentes intervalos de tempo trabalhados nesses dois casos.

FIGURA 6 – Evolução do número de habitantes (barras verticais) e da temperatura média global da Terra (cruzes interligadas), entre 1910 a 2009. Os círculos abertos indicam a suavização dos dados e a reta pontilhada a tendência do aumento do desvio da temperatura média global entre 1965 a 2009.



Fonte: Adaptado de IBGE, 2010 e CRU/UEA, 2010.

Um fato marcante, e que chamou atenção do grupo de estudantes, ocorreu ao avaliar com mais detalhes a Figura 6. Como se pode observar, o comportamento da suavização da temperatura média global entre os anos 1925 a 1950 mostrou-se questionável, uma vez que apareceu uma forte tendência de aumento e depois diminuição desse parâmetro climático. Tanto o grupo de estudantes envolvido nessa pesquisa como os autores responsáveis pela mesma, não encontraram uma justificativa plausível para o comportamento estranho no desvio da temperatura média global do planeta, deixando essa questão em aberto.

Particularmente, as duas averiguações (envolvendo o acréscimo da população com a temperatura média local e acréscimo da população com o desvio da temperatura média global) chamaram a atenção de todos os pesquisadores envolvidos. Os mesmos concluíram que entre as décadas de 50 e 60 do século passado, época em que se iniciou de forma efetiva uma escalada da indústria e da ocupação urbana na Região Metropolitana de Campinas (RMC), pôde-se inferir o indício do aumento da temperatura média local, juntamente com um incremento do desvio da temperatura média global do planeta. Além disso, eles observaram que essas duas grandezas climáticas acompanharam o crescimento da cidade no período em que foi feita essa análise. Segundo os próprios estudantes concluíram,

o fato de ter sido identificado um aumento na temperatura média de Campinas, muito provavelmente foi também devido às causas da variabilidade natural do clima, além da própria ação do Homem ao longo da história do município.

Conclusões fechadas sobre o assunto não são possíveis de serem firmadas, mesmo porque muitas questões estão em aberto e são pautas de discussões na comunidade científica preocupada com as mudanças climáticas. Entretanto, por intermédio de reflexões sobre o propalado aquecimento global, ficou claro para os estudantes envolvidos que é uma evidência comprovada que o planeta possui sua variabilidade natural do clima e que isso se deve a própria interação da Terra com os fenômenos que ocorrem no Universo (IPCC, 2007; NOBRE, 2008). O fato de ter sido identificado um pequeno aquecimento local devido ao aumento da população de Campinas não é um aval para os moradores dessa, ou de qualquer outra cidade, continuarem a degradar o meio ambiente.

Por conseguinte, considerando que é inevitável o aumento populacional do futuro próximo, o bom senso sugere a adoção de políticas de conservação ambiental melhor elaborada, além de mudanças nos hábitos de consumo e na ocupação do solo urbano. É também nesse sentido que a Matemática e/ou a Estatística podem contribuir para tomadas de decisões técnicas, de forma a ajudar a sociedade a decidir seus rumos e estando em conformidade com os pensamentos exarados por Skovsmose (2008).

CONSIDERAÇÕES FINAIS E PERSPECTIVAS

Identificar alterações nos registros meteorológicos é de grande importância para os estudos nas engenharias que utilizam séries de dados, pois tanto simulações como aplicações de teorias de probabilidades são realizadas com a hipótese de que as séries são homogêneas (isto é, os dados da série vêm da mesma população), ou seja, não apresentam tendências. Além disso, dada a relevância de um tema tão atual, aliado às inúmeras discussões veiculadas pelas mídias, o envolvimento desses assuntos com estudantes de outros cursos, como Matemática, Estatística, Geociências, torna-se importante na medida em que se consegue explorar conteúdos estatísticos.

O presente estudo envolveu a participação de um grupo de estudantes do curso de Engenharia Ambiental, interessados em desenvolver voluntariamente um procedimento computacional para aplicar métodos estatísticos que possibilitassem identificar tendências em séries históricas baseadas em dados de temperatura média e precipitação pluvial. Os dados trabalhados e analisados foram os de medições levantadas pela estação meteorológica do IAC, localizada em Campinas (SP). Foram utilizados os dados mensais e anuais no período de quase um século de medições.

Foi possível desenvolver com esses estudantes, ao longo de vários encontros num período de aproximadamente oito meses, métodos estatísticos que despertaram o interesse por parte dos mesmos, o que permitiu inferir a existência ou não de tendências climáticas nas séries de dados. Neste aspecto, observou-se que os estudantes passaram a ter um comportamento mais crítico perante a pesquisa realizada. Essa atitude foi evidenciada pela

questão por eles levantada, ao compararem o aumento populacional com a variação da temperatura média na cidade de Campinas. Isso vai ao encontro das questões exploradas por Barbosa (2006) e Skovsmose (2008), visto que permeia os aspectos inclusos nas perspectivas abordadas por esses autores e levantadas nesse trabalho.

Arigor, foi possível identificar e mostrar aos membros do grupo, por meio de gráficos e procedimentos computacionais em atividades expositivas extraclasse, a importância e o uso do programa Excel e suas ferramentas. Por meio dessa facilidade tecnológica, foram aplicadas análises de regressão por meio de médias móveis e testes não-paramétricos de Mann-Kendall e de Pettitt (esses, especialmente desenvolvidos no Excel). Para todas as análises foram inferidos os intervalos de confiança e suas interpretações no contexto desta pesquisa, como explorado por Olivo e Batanero (2007). Por conseguinte, ficou evidenciado pelos estudantes que o *software* Excel é uma poderosa ferramenta no tratamento de dados, possuindo muitas funções da Matemática e da Estatística que invariavelmente são pouco exploradas nos cursos superiores das universidades brasileiras (BORBA; VILLAREAL, 2005).

Essa pesquisa possibilita abrir novas perspectivas no ensino e aplicação da Estatística, pois se pretende dar continuidade ao trabalho envolvendo estudantes interessados em averiguar o comportamento de outras variáveis climatológicas, inclusive envolvendo a análise de outros bancos de dados de outros municípios do Estado de São Paulo. Pretende-se, se possível, aplicar outros procedimentos estatísticos, além dos aqui apresentados, como teste paramétrico Z (disponível no Excel) e os testes não paramétricos de Sen (SEN, 1968; THEIL, 2000), Run (THOM, 1966) e o teste de Wilcoxon-Mann-Whitney (SIEGEL; CASTELLAN JR, 2006), que podem melhorar substancialmente as análises desta proposta. Ao mesmo tempo a pesquisa pode ser ampliada e aplicada para uma quantidade maior de dados e de diferentes tipos de informações, como a umidade relativa do ar, a pressão atmosférica e a radiação solar. Além disso, é possível dar continuidade ao trabalho em várias outras situações, como na averiguação do comportamento das vazões de bacias hidrográficas, no zoneamento de áreas agrícolas e agropecuárias, no reflorestamento em pequena e média escala, dentre outros.

Por fim, vale ressaltar ainda que, apesar da tecnologia disponível no momento, não existem modelos matemáticos totalmente confiáveis que abordem as questões climatológicas, daí a importância desse trabalho como fator motivador para esses estudos.

REFERÊNCIAS

- AYOADE, J. O. *Introdução à Climatologia para os Trópicos*. Tradução de M. J. Z. dos Santos. 4.ed. Rio de Janeiro: Bertrand Brasil. 1996.
- BARBOSA, J. C. Mathematical Modelling in classroom: a critical and discursive perspective. *Zentralblatt für Didaktik der Mathematik*, v.38, n.3, p.293-301, 2006.
- BACK, A. J. A percepção de análise estatística para identificação de tendências climáticas. *Pesquisa Agropecuária Brasileira* (Brasília), v.36, n.5, p.717-726, 2001.

BORBA, M. C.; VILLARREAL, M. E. *Humans-with-Media and Reorganization of Mathematical Thinking: Information and Communication Technologies, Modeling, Visualization and Experimentation*. New York: Springer Science+Business Media, Inc., 2005.

CRU/UEA. Climatic Research Unit/University of East Anglia, 2010. Disponível em: <<http://www.cru.uea.ac.uk/>>. Acesso em: 10 mar. 2010.

FONTANA, D. C.; BERLATO, M. A. A influência do El Niño: oscilação sul sobre a precipitação pluvial do Estado do Rio Grande do Sul. *Revista Brasileira de Agrometeorologia*, Santa Maria, v. 5, p. 127-132, 1997.

FREUND, J. E. *Estatística Aplicada: Economia, Administração e Contabilidade*, 11.ed., Porto Alegre: Editora Bookman, 2006.

GOOSSENS, C.; BERGER, A. Annual and seasonal climatic variations over the northern hemisphere and Europe during the last century. *Annales Geophysicae*, Berlin, v.4, n.B4, 1986.

GROPPO, J. D. et al. A. Análise de séries temporais de vazão e precipitação em algumas bacias do Estado de São Paulo com diferentes graus de intervenções antrópicas. *Geociências*, São Paulo, v. 24, n.2, p.181-193, 2005.

IAC/CIIAGRO. Instituto Agrônomo de Campinas – Monitoramento climático, 2009. Disponível em: <<http://www.ciiagro.sp.gov.br/ciiagroonline/>>. Acesso em: 22 jun. 2010.

IBGE. Instituto Brasileiro de Geografia e Estatística, 2010. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 08 fev. 2010.

IPCC Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. SOLOMON, S. et al. (Ed.). Cambridge University Press: United Kingdom and New York, 996 pp., 2007.

KAISER, G; SRIRAMAN, B. A global survey of international perspectives on modelling mathematics education. *Zentralblatt für Didaktik der Mathematik*, v.38, n.3, p.301-310, 2006.

MORAES, J. M. et al. Estudo preliminar da evolução temporal dos componentes do ciclo hidrológico da bacia do Rio Piracicaba. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, 11.; SIMPÓSIO DE HIDRÁULICA E RECURSOS HÍDRICOS DOS PAÍSES DE LÍNGUA OFICIAL PORTUGUESA, 2, Recife. *Anais...* Recife: Associação Brasileira de recursos Hídricos, 1995. p.27-32.

MORTATTI, J. et al. Hidrologia dos rios Tietê e Piracicaba: séries temporais de vazão e hidrogramas de cheia. *Revista de Ciência & Tecnologia*, v.12, n.23, p.55-67, 2003.

NOBRE, C. Mudanças Climáticas e o Brasil – Contextualização. *Parcerias estratégicas*, v.27, p.7-17, 2008.

OKE, T. R. *Boundary Layer Climates*. 2nd edition. Methuen: New York, 435pp., 1987.

OLIVO, E.; BATANERO, C. Un estudio exploratorio de dificultades de comprensión del intervalo de confianza. *UNIÓN – Revista Iberoamericana de Educación Matemática*, n.12, p.37-51, 2007.

PELLEGRINO, G. Q. et al. A. Análise espaço-temporal de componentes hidroclimáticos na bacia do rio Piracicaba, SP. *Revista Brasileira de Agrometeorologia*. Santa Maria, v.9, n.1, p.125-135, 2001.

PETTITT, A. N. A non-parametric approach to the change-point problem. *Applied Statistics*, London, v.28, n.2, p.126-135, 1979.

SELLO, S. Time series forecasting: a nonlinear dynamics approach, 1999. Disponível em: <<http://arxiv.org/abs/physics/9906035>>. Acesso em: 21 jun. 2010.

SEN, P. K. Estimates of the regression coefficient based on Kendall's tau. *Journal of American Statistics Association*, v.63, p.1379-1389, 1968.

SIEGEL, S.; CASTELLAN JR, N. J. *Estatística não-paramétrica para ciências do comportamento*. 2.ed. Porto Alegre: Artmed, 2006.

SKOVSMOSE, O. *Desafios da reflexão em Educação Matemática Crítica*. Campinas: Papirus, 2008.

SNEYERS, R. *Sur l'analyse statistique des séries d'observations*. Genève: Organisation Meteorologique Mondial, 192p. (OMM Note Technique, 143), 1975.

SNEYERS, R. et al. Climatic changes in Belgium as appearing from the homogenized series of observations made in Brussels – Uccle (1933-1988). In: SCHIETECAT, G. D. (Ed.). *Contributions à l'étude des changements de climat*. Bruxelles: Institut Royal Meteorologique de Belgique, Publications Série 124, 1990. p.17-20.

THEIL, H. A rank-invariant method of Linear and polynomial regression analysis, I, II, III. *Nederl Akad Wetensch, Proc.* 53, p.386-392, p.512-525, p.1397-1412, 2000.

THOM, H. C. S. *Some methods of climatological analysis*. Genève: World Meteorological Organizational (WMO Technical Note 81), 1966.

YEVJEVICH, V. *Probability and statistics in hydrology*. Fort Collins: Water Resources Publication, 1972.

Recebido em: out. 2010

Aceito em: mar. 2011