

## Influência de Características do Solo na Variação do Nível d'água em Região de Recarga do Aquífero Guarani

Taiana Homobono Gouvêa, Edson Cesar Wendland

EESC-USP – São Carlos

*homobono@sc.usp.br; ew@sc.usp.br*

Recebido: 18/10/09 - revisado: 22/11/10 - aceito: 17/03/11

---

### RESUMO

O objetivo desta pesquisa é analisar a influência de características de solo na variação do nível d'água em área de afloramento do Aquífero Guarani. A variação do nível d'água no aquífero livre foi monitorada em 12 poços no período de setembro de 2004 a agosto de 2008, na bacia hidrográfica do Ribeirão da Onça - SP. A Estação Climatológica do CRHEA-USP, localizada no município de Itirapina-SP, forneceu os dados diários de precipitação, para o mesmo período. Para alcançar os objetivos propostos foram utilizados os seguintes métodos estatísticos: técnicas de agrupamento e análise de correlação. Os dendrogramas, gerados pelas técnicas hierárquicas, mostraram que todos os atributos (características físicas) utilizados no estudo influenciam na formação de grupos de poços de monitoramento, com ênfase para o coeficiente de evapotranspiração. Os coeficientes de correlação entre variação do nível d'água nos poços de monitoramento localizados em áreas de mesma cultura apresentaram maiores valores ( $r > 0,5$ ). Esse resultado, aliado aos grupos formados, mostra que o uso do solo (cultura) tem maior influência na variação do nível d'água em regiões de recarga. No entanto, a porosidade, também considerada neste estudo, mostrou-se como atributo de importância, uma vez que, a partir da análise dos grupos é possível perceber que sua formação também pode ser definida segundo essa característica. Assim, o presente trabalho mostra que o estudo de características de solo e seu uso são fundamentais para o projeto de rede de monitoramento de recursos hídricos subterrâneos, principalmente em áreas de recarga.

**Palavras-chave:** águas subterrâneas, chuva, métodos estatísticos, agrupamentos.

---

### INTRODUÇÃO

A quantificação da recarga das águas subterrâneas é um requisito fundamental para o manejo destes recursos. As informações obtidas pelas redes de monitoramento de recursos hídricos são essenciais para detectar os impactos de mudanças do clima e de atividades humanas na quantidade e na qualidade dos aquíferos. O processo de interação entre aquífero e ambiente pode ser observado somente por uma rede de monitoramento (Zhou, 1996). Há uma grande dificuldade na determinação da recarga de aquíferos e suas estimativas possuem elevado grau de incerteza. Nas áreas de recarga dos aquíferos, uma parte da água que infiltra no solo, retorna como escoamento básico, enquanto que outra parte pode ir para as porções profundas ou confinadas podendo se transformar em recarga profunda (Gomes, 2008).

A América do Sul abriga um dos maiores mananciais de água subterrânea do mundo: o Sis-

tema Aquífero Guarani (SAG), com extensão de 839.800 km<sup>2</sup>. No Brasil, nas regiões Centro-Oeste, Sul, e Sudeste abrangem 71% da área total do Aquífero, incluindo estados com grande importância econômica, com áreas de intensa atividade urbana, industrial e agrícola.

Faz-se indispensável um sistema de monitoramento no que diz respeito à sua disponibilidade, uma vez que a exploração não deve ser maior que a capacidade de recarga. As áreas de recarga, portanto, devem receber especial atenção de estudos hidrológicos com o intuito de garantir um bom planejamento e uso dos recursos hídricos subterrâneos, obedecendo ao balanço hídrico do SAG.

Face ao exposto, o presente trabalho busca avaliar a influência de características do solo na variação do nível da água em poços de monitoramento, em região de afloramento do Aquífero Guarani. Para tal foram utilizados os seguintes métodos estatísticos: análise de correlação e análise de agrupamento. O estudo foi realizado com base em dados hidrológicos coletados no período de outubro de

2004 a setembro de 2008 na bacia piloto do Ribeirão da Onça. A importância desse estudo se encontra na possibilidade de verificar comportamentos semelhantes e, eventualmente, permitir a redução do número de poços monitorados. Aliado a esse fato, o estudo também fornece subsídios a implantação de uma rede de monitoramento completa, capaz de definir todo o balanço hídrico da região estudada.

## RECARGA DE AQUÍFEROS E CARACTERÍSTICAS DE SOLO

A recarga das águas subterrâneas é definida como o fluxo de água que alcança o aquífero, constituindo uma adição ao reservatório de água subterrânea. Isso ocorre, em geral, a partir da percolação vertical proveniente da zona vadosa e do escoamento lateral a partir das fronteiras do sistema. A recarga subterrânea de certa área é normalmente considerada a infiltração em excesso na mesma área (Devries e Simmers, 2002).

A interação entre clima, geologia, morfologia, condições do solo e vegetação determina o processo de recarga, principalmente por ter origem nos fenômenos de precipitação. As taxas de recarga são limitadas, em grande parte, pela água disponível na superfície do terreno, controlada por fatores climáticos, como a precipitação e evapotranspiração e pelas características geomorfológicas superficiais (topografia, solo e cobertura vegetal) (Maziero, 2005). Assim, algumas características do solo merecem atenção:

### A. Porosidade

Com relação ao processo de variação de nível ou armazenamento em aquífero livre (sujeito à pressão atmosférica), a porosidade é a mais importante das propriedades físicas dos solos ou rochas a ser determinada. É expressa pela relação entre o volume de vazios (ou ocupados por ar ou água) de um material e seu volume total. A porosidade total representa a quantidade a quantidade máxima de água que um dado volume de solo pode conter

Gomes (2008) enfatiza que a porosidade acima definida é denominada porosidade primária. Além da porosidade primária, há aquela referente às fraturas e cavidades de dissolução (porosidade secundária), as quais resultam de uma gama de processos físicos e químicos.

### B. Rendimento Específico

O rendimento específico é um importante parâmetro na obtenção da taxa de recarga, principalmente, nos métodos que utilizam flutuações do nível d'água (WTF – *Water Table Fluctuation*; CRD – *Cumulative Rainfall Departure*). Contudo, os valores relatados para o rendimento específico na literatura são bastante variáveis, assim como os métodos para sua obtenção. Outras vezes, acontece a substituição do parâmetro *rendimento específico* por algum outro (porosidade efetiva, coeficiente de armazenamento) (Gomes, 2008).

## ANÁLISE DE AGRUPAMENTO

Análises de agrupamentos (*cluster analysis*) são técnicas estatísticas de reconhecimento de padrões de distribuição de informações. Mostram a existência de estruturas intrínsecas ou comportamento de dados amostrais sem que sejam feitas suposições prévias sobre estes. Assim, estas técnicas objetivam classificar objetos (dados amostrais, informações) em categorias ou grupos (classes) baseados nas similaridades ou proximidades entre tais objetos. Desta forma, os dados, também chamados de casos, com maiores similaridades são agrupados, a partir da aplicação de algoritmos, em uma mesma classe, de maneira tal que haja dissimilaridades máximas entre as classes (Vega *et al*, 1998).

Os resultados oriundos desta técnica auxiliam na interpretação dos dados obtidos em uma pesquisa e indicam padrões. Muñoz-Díaz e Rodrigo (2004) ressaltam que os objetos alocados em um grupo tendem a serem similares segundo atributos que podem ser apenas calculados através de algoritmos ou ainda, definidos pelo pesquisador. No entanto, esses atributos, quando utilizados a critério do pesquisador, não indicam a formação prévia de grupos, são apenas parâmetros a serem considerados no agrupamento.

As análises de agrupamentos têm sido utilizadas em diversas áreas. Nos estudos de climatologia se mostram de grande importância. Ramos (2001) utilizou as técnicas para analisar os padrões de distribuição das chuvas e suas mudanças, na região mediterrânea espanhola, ao longo dos anos. Muñoz-Díaz e Rodrigo (2004) aplicaram, em uma série de 88 (oitenta e oito) anos de dados de precipitação, as análises de agrupamento, objetivando encontrar zonas climáticas homogêneas em 32 (trinta e duas) localidades da Espanha.

Na Hidrologia, os agrupamentos também têm sendo utilizados de maneira satisfatória. Panda *et al* (2006), com o objetivo de estudar os processos antropogênicos e naturais, bem como os fatores responsáveis pelas características hidrológicas no sistema de rios Mahanadi, na Índia, utilizou de análise fatorial e análise de agrupamentos. Este autor afirma que a partir dos dendrogramas, as relações hidrológicas e de qualidade das águas ficam bem explícitas, mostrando que nas estações de coleta de dados as características dessas relações possuem similaridades. Ferraro (2006) analisou o comportamento de piezômetros localizados nas bacias do Tijucu Preto e Gregório, a partir dos métodos de técnicas de agrupamento.

Muitas investigações utilizando técnicas de agrupamentos têm como ponto de partida uma matriz  $n \times n$  com elementos que refletem medidas quantitativas de proximidade, denominadas mais comumente de *dissimilaridades*, *distância de similaridade* ou, pelo termo geral, *proximidade*. Tais proximidades podem ser determinadas de maneira direta ou indireta.

Há dois tipos de agrupamentos: os hierárquicos e os divisivos. O objetivo da técnica divisiva é separar objetos em grupos consistentes. A escolha preliminar dos objetos pode ser feita aleatoriamente. Ao final do agrupamento, cada objeto encontrará-se alocado em um grupo. Já nas técnicas hierárquicas, os objetos são progressivamente agregados até que se juntem em um único grupo. Cada objeto inicia em um grupo singular. A partir de então os grupos mais próximos se agregam formando novos grupos, até que reste apenas um grupo contendo todos os objetos (Muñoz Diaz e Rodrigo, 2004).

Segundo Muñoz-Díaz e Rodrigo (2004), as técnicas hierárquicas (de fusão) são mais implementadas. Seu propósito é o de formar grupos, minimizando as perdas de informações (sem que haja o descarte de dados). De acordo com o critério de agregação, há diferentes métodos de hierarquização. Nas técnicas de agrupamento hierárquicas, a distância entre as amostras é utilizada como uma medida de similaridade. Neste tipo de agrupamento podem ser utilizados três tipos de algoritmos: *complete linkage*; *average linkage*; *Ward's method*. (Vega *et al.*, 1998)

O modelo de agrupamento mais utilizado tem sido o método hierárquico seqüencial, onde cada dado é considerado um grupo separado. Milligan e Cooper (1987) ressaltam a possibilidade de escolha do pesquisador, o qual pode utilizar o agrupamento hierárquico completo como solução, ou

ainda, selecionar um nível que represente o número de grupos específicos do seu interesse.

O progresso e os resultados de uma análise de agrupamento são convencionalmente ilustrados a partir de *dendrogramas* ou diagramas de árvore. Trata-se de uma figura bidimensional que representa a seqüência e as distâncias métricas em que as observações são agrupadas. As distâncias entre os grupos antes de se agregarem também são indicadas nesses diagramas (Muñoz Diaz e Rodrigo, 2004).

Os métodos de agrupamento baseiam-se em medidas de distância para realizar os agrupamentos, sendo a Distância Euclidiana (*Ward's Method*) a mais utilizada. Esta medida calcula a distância entre as classes que, por sua vez, é calculada pela soma dos quadrados destas classes, incluindo todas as variáveis. A cada etapa, a soma dos quadrados é minimizada por todas as partições obtidas pela fusão de duas classes de uma etapa anterior.

A distância entre dois objetos é dada por:

$$d_{rs}^2 = \sum_{j=1}^r (X_{rj} - X_{sj})^2 \quad (1)$$

em que  $d_{rs}$  é a distância euclidiana da matriz X que contém as observações e variáveis a serem agrupadas, r e s indicam as observações agrupadas dois a dois para todas as variáveis. A raiz quadrada da distância euclidiana define a distância existente entre os grupos formados, utilizando dois objetos (Jobson, 1992).

Considerando duas classes,  $C_k$  e  $C_r$ , que se fundem para a formação de uma nova classe,  $C_m$ , a distância entre esta nova classe e outra,  $C_j$ , é:

$$d_{J,M} = \frac{((n_J + n_K)d_{JK} + (n_J + n_L)d_{JL} - n_J d_{KL})}{n_J + n_M} \quad (2)$$

em que  $n_j$ ,  $n_k$ ,  $n_l$  e  $n_m$  são os números de objetos das classes J, K, L e M, respectivamente, e  $d_{jk}$ ,  $d_{jl}$  e  $d_{kl}$  representam a distância entre as observações nas classes J e K, J e L, e K e L, respectivamente (Ramos, 2001).

Na técnica divisiva, os objetos são aleatoriamente dispostos em um número de classes previamente definidas. Eles podem ser distribuídos em uma única classe. A partir da aplicação do algoritmo "*k-mean*", os objetos são transferidos para classes com as maiores similaridades (Ramos, 2001). O algoritmo *k-mean* é desenvolvido computacionalmente, partindo de uma repartição inicial aleatória especificada por um vetor  $P$  (K,M) com M observações alocadas em K grupos. O centróide (média das vari-

áveis analisadas) de cada grupo é computado e a similaridade entre as observações nos grupos é medida por uma função-erro  $e$  definida por:

$$e[P(M, K)] = \sum_k D[L, L(I)]^2 \quad (3)$$

em que  $L(I)$  é o grupo contendo  $I$  casos (observações),  $D[L, L(I)]^2$  representa a soma dos quadrados dos desvios de cada centróide do grupo (Ramos, 2001).

Apesar de existirem outras medidas de distância, no trabalho será utilizada a Distância Euclidiana, pois segundo Vega *et al.* (1998), os agrupamentos hierárquicos gerados a partir do método *Ward*, por produzirem menores efeitos de distorção, usando maiores informações (dados amostrais) nos grupos, têm fornecido mecanismos mais poderosos de agrupamento, além de terem maior significado estatístico. Muñoz-Diaz e Rodrigo (2004) também afirmam que, mesmo os agrupamentos sendo sensíveis ao método escolhido, o método *Ward* fornece as soluções mais acuradas.

É importante ressaltar que os algoritmos de agrupamento não restringem dados de entrada a distribuições estatísticas particulares, porque as análises de agrupamento são mais um método objetivo de quantificação de características, de uma gama de observações, que uma ferramenta de inferência estatística (Muñoz Diaz e Rodrigo, 2004).

## ANÁLISE DE CORRELAÇÃO

A teoria das correlações objetiva estudar o grau de relacionamento entre duas variáveis, ou seja, busca definir uma medida de covariabilidade. Na análise de correlação não se aceita a idéia de dependência entre as variáveis, existindo, assim, uma influência mútua ou conjunta entre estas. Neste tipo de estudo, as variáveis são, em geral, consideradas aleatórias e a amostra, proveniente de uma distribuição conjunta dessas variáveis (Demetrio e Zocchi, 2003).

Para medir o grau de correlação linear simples entre duas variáveis pode-se utilizar o coeficiente de correlação de Pearson (em geral, indicado por  $r$ ), o qual varia entre -1 e +1. Sendo o valor de  $r$  correspondente a um desses valores extremos, todos os pontos da amostra devem cair sobre uma reta. O coeficiente de correlação de Pearson é dado pela seguinte fórmula:

$$r = \frac{\sum xy - \frac{1}{n} \sum x \sum y}{\sqrt{\left[ \sum x^2 - \frac{1}{n} (\sum x)^2 \right] \left[ \sum y^2 - \frac{1}{n} (\sum y)^2 \right]}} \quad (4)$$

em que  $x, y$  são os valores medidos para as duas variáveis e,  $n$  é o número de valores de cada variável.

A correlação entre duas variáveis é positiva se os valores mais altos de uma das variáveis estiverem associados aos valores mais altos da outra variável, e é negativa, se os valores mais altos de uma das variáveis estiverem associados aos valores mais baixos da outra variável. No caso da ocorrência de zero (ou valores muito próximos a ele) no valor do coeficiente, significa que não há correlação entre as variáveis. Segundo Vieira (2004), mesmo o coeficiente sendo igual a zero não significa que não exista correlação, e sim que não há correlação linear. O valor de  $r$  mede o grau de dispersão dos pontos em torno de uma reta. Quanto maior for a dispersão dos pontos ao redor da reta, menor será o valor de  $r$  (Vieira, 2004).

Se todos os pontos relativos às amostras das variáveis não estiverem próximos a uma reta, diz-se que a correlação é *não-linear*. Desta forma, equações não-lineares mostram-se mais apropriadas. Além disso, a ocorrência de valores baixos para o coeficiente de correlação, como já dito, pode demonstrar não-linearidade dos valores estudados. Uma importante alternativa é a transformação dos valores para a escala logarítmica e a posterior aplicação do cálculo de correlação (Spiegel, 1970).

## LOCAL DO ESTUDO

A Bacia-Piloto do Ribeirão da Onça, área de estudo deste trabalho, está localizada na região Centro-Oeste do Estado de São Paulo, no município de Brotas, entre os paralelos 22°10' e 22°15' de latitude Sul e os meridianos 47°55' e 48°00' de longitude Oeste.

Na região onde se localiza a bacia há predominância de uso agrário e pecuário, não sendo observadas áreas consideradas pavimentadas e nenhuma área industrial. A cobertura vegetal original dos solos era o cerrado. Atualmente, é composta por pastos, cana de açúcar, laranja, eucalipto, áreas de várzea e culturas de alta rotatividade como soja e amendoim (Gomes, 2008).

Esta bacia tem representado foco de estudos sobre o comportamento dos processos de recarga do aquífero, bem como do balanço hídrico da região (Contin Neto, 1987; Davino, 1984; Pompêo, 1990;

Cunha, 2003; Barreto, 2006; Wendland, Barreto, Gomes 2007), uma vez que apresenta características representativas das regiões de afloramento do Sistema Aquífero Guarani (SAG).

Os solos da região originam de intemperismo da Formação Botucatu. Além dessa formação, a bacia passa, à jusante, sobre *sills* de diabásio. Estas duas unidades participam do Grupo São Bento, de idade Mesozóica.

O solo superficial denominado Sedimento Cenozóico, ocorrente na maior parte do interior do Estado de São Paulo, recobre a bacia do Ribeirão da Onça. Em geral, este solo possui espessura inferior a 10 m, freqüentemente separado da camada subjacente por uma linha de seixos (Bortolucci, 1983). Em função de sua origem, este sedimento é pouco compacto com partículas razoavelmente selecionadas, resultando em uma estrutura bastante porosa. Além disso, há a formação de duas texturas típicas: uma arenosa (predominante na bacia) e outra argilosa (Bortolucci, 1983).

Próximo ao Ribeirão da Onça está localizado o CRHEA (Centro de Recursos Hídricos e Ecologia Aplicada da Universidade de São Paulo). A Estação Climatológica do CHREA coleta e fornece todos os dados meteorológicos necessários ao acompanhamento dos fenômenos hidrológicos.

A região possui temperatura média de 20,5°C. A classificação climática de Köppen indica que o clima dessa área é subtropical úmido, com chuvas de verão, podendo apresentar variações para o clima tropical úmido com inverno seco. A precipitação média anual é da ordem de 1300-1400 mm. Pompêo (1990) apresenta uma média anual regional de longo tempo igual a 1302,4 mm (Barreto, 2006; Gomes, 2008).

## MATERIAL E MÉTODOS

A área que abrange a bacia piloto do Ribeirão da Onça possui 23 (vinte e três) poços de monitoramento, os quais foram construídos em áreas de diferentes culturas. O monitoramento do nível d'água é realizado quinzenalmente, com auxílio de um medidor de nível do tipo sensor elétrico com fita milimetrada.

Para o presente estudo foram utilizados dados de 12 dos 23 poços. O descarte de alguns poços ocorreu devido a dois fatores: quatro poços (poços 20, 21, 22 e 23) estão localizados em regiões muito próximas ao curso d'água, os tornando suscetíveis às variações do nível d'água no Ribeirão; ou-

trois sete secaram durante longos períodos (poços 1, 2, 3, 6, 7, 11 e 12). Fez-se aplicação de interpolação linear dos dados quinzenais de nível da água subterrânea, para obtenção de valores diários. Após isso, foram calculadas as variações mensais de nível da água em cada poço estudado.

Nas técnicas de agrupamento podem ser utilizados atributos para a classificação, ou seja, podem ser incluídos parâmetros ou características dos casos a serem analisados. Assim, neste trabalho, foram usados valores referentes à porosidade (Gomes, 2008) e à variação mensal da profundidade da superfície livre (nível) do aquífero ( $\Delta h$ ). Além destes, utilizou-se características do uso do solo, através do coeficiente de evapotranspiração (Kc) (Barreto, 2006), como atributo para o agrupamento. Os quadros 1 e 2 mostram os valores referentes aos atributos citados.

**Quadro 1- Coeficiente de Cultura (Barreto, 2006)**

Poços	Cultura	Coeficiente de Evapotranspiração (Kc)
9, 10	Eucalipto	1,00
4, 5, 16, 17, 18, 19	Pastagem	0,90
15, 13, 14	Cítricos	0,80
8	Cana-de-açúcar	1,04

**Quadro 2- Porosidade média nos poços da bacia do Ribeirão da Onça (Gomes, 2008)**

Poços	Porosidade (%)
4,5	15,9
8, 9, 10, 13, 14	15,1
15	8,5
16,17, 18, 19	12,3

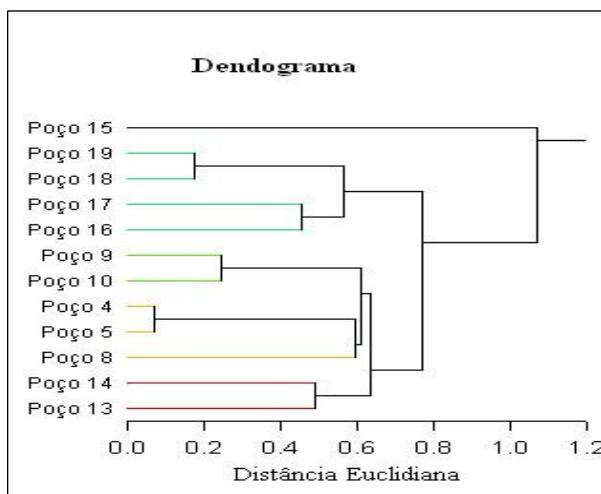
Nas técnicas divisivas são usados algoritmos *k-means* em que as medidas de similaridade entre os grupos são definidas pela Distância Euclidiana (equação 1) entre os centróide dos grupos. Através de softwares estatísticos é computada uma repartição aleatória especificada por um vetor  $P(K,M)$  em que M são observações alocadas no grupo K. A partir desse vetor, define-se o centróide inicial para o grupo, ou seja, o valor médio entre todas as variáveis das observações estudadas. A similaridade entre as observações alocadas no agrupamento são medidas pela função erro, definida pela equação 3.

As técnicas hierárquicas também utilizam a Distância Euclidiana como medida de similaridade. No entanto, utiliza-se a raiz quadrada da distância euclidiana para o cálculo da distância entre os grupos (equação 2).

O software escolhido para o desenvolvimento das técnicas de agrupamento foi o Systat 10 (SPSS Science, 2007), o qual se utiliza de algoritmos para a formação de grupos, sem que sejam definidas as quantidades de grupos a serem formados, nem o número de observações (casos) que cada um deverá conter. Para o caso dos agrupamentos divisivos o software permite um número máximo de grupos equivalente a 5, sendo o número de grupos definido pelo pesquisador.

Objetivando confirmar os grupos gerados a partir das técnicas de agrupamento realizou-se o cálculo do coeficiente de correlação entre as variações mensais do nível d'água nos poços monitorados. Utilizou-se o coeficiente de correlação de Pearson (equação 4) considerando como variáveis pares de poços. Isso significa que os poços foram comparados 2 a 2, até que se obtivesse os coeficientes de correlação entre todos os poços.

## RESULTADOS E DISCUSSÕES



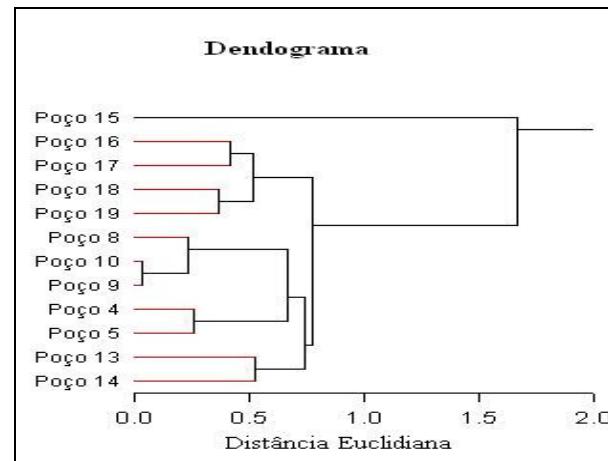
**Figura 1- Dendrograma de agrupamento para os poços monitorados, para o mês de novembro de 2004.**

O objetivo do trabalho foi detectar a formação de classes de poços de monitoramento a partir de atributos comuns (características físicas da região estudada) e variação do nível d'água.

Foram gerados dendogramas para cada mês do período estudado. Na figura 1 é apresentado o dendrograma representativo do período de setembro de 2004 a julho de 2006, uma vez que houve poucas mudanças nos agrupamentos ao longo dos meses.

Analizando o dendrograma gerado através da técnica de agrupamento hierárquico, é notado que não há grandes saltos entre as distâncias euclidianas obtidas. Segundo a revisão bibliográfica o critério para definição dos agrupamentos, nas técnicas hierárquicas, é o salto do valor da distância euclidiana. Dessa maneira, para análise dos resultados definiu-se a formação de cinco grupos, os quais foram gerados com uma distância euclidiana no valor de 0,6.

Os agrupamentos divisivos, realizados a partir do desenvolvimento de algoritmos *k-means* (para tal, fez-se uso do software estatístico Systat), por outro lado, mostraram resultados menos significativos, uma vez que o número de grupos a ser formados é decidido de maneira aleatória, pelo pesquisador. Da mesma maneira que para as técnicas hierárquicas, realizou-se o agrupamento mensal dos poços monitorados. A tabela 1 mostra o agrupamento divisivo.



**Figura 2- Dendrograma de agrupamento para os poços monitorados, para o mês de novembro de 2007.**

Em agosto de 2006 houve uma mudança de cultura em duas áreas de plantação da região estudada. O poço 8 encontrava-se até então localizado em região de cultivo de cítricos. A partir do período supracitado, passou-se a cultivar cana-de-açúcar no local. O inverso ocorre no local onde está localizado o poço 15. Passou a cultivar cítricos onde antes havia cultivo de cana-de-açúcar. Mediante essa informação

**Tabela 1-Agrupamento dos poços monitorados, a partir de técnicas divisivas, no período de setembro de 2004 a julho de 2006.**

Grupo	Distância								
Poço 5	0,61	Poço 4	0,4	Poço 15	0	Poço 16	0,26	Poço 18	0,18
Poço 8	0,18	Poço 9	0,21			Poço 17	0,26	Poço 19	0,18
Poço 13	0,32	Poço 10	0,23						
Poço 14	0,15								

**Tabela 2-Agrupamento dos poços monitorados, a partir de técnicas divisivas, para o período de agosto de 2006 a setembro de 2008.**

Grupo	Distância								
Poço 4	0,43	Poço 13	0,26	Poço 15	0,0	Poço 16	0,21	Poço 18	0,18
Poço 5	0,59	Poço 14	0,26			Poço 17	0,21	Poço 19	0,18
Poço 8	0,49								
Poço 9	0,25								
Poço 10	0,27								

os dados do período correspondente a agosto de 2006 até setembro de 2008, foram submetidos novamente a análise de agrupamento, com o propósito de detectar possíveis mudanças na formação de grupos. O dendograma da figura 2 e a tabela 2 apresentam os grupos formados pelas técnicas hierárquicas e divisivas, respectivamente.

Ambas as técnicas de agrupamento mostraram resultados consistentes. No entanto, considerase, nessa pesquisa, que os resultados dos agrupamentos hierárquicos foram mais consistentes, pois, como já citado, o número de grupos formados pelas técnicas divisivas é definido aleatoriamente. Além disso, as técnicas hierárquicas refletem a realidade dos agrupamentos, uma vez que diminuem as perdas de informações, como já explanado na revisão bibliográfica desse trabalho.

A definição dos cinco grupos foi determinada pelos valores de distância euclidiana obtidos, como já citado. Analisando os dendrogramas, nota-se a formação de 5 grupos, em que a distância euclidiana varia entre 0,5 e 0,6. Os próximos agrupamentos gerados encontram-se próximos a 1,0 (distância euclidiana). No dendrograma correspondente ao período de agosto de 2006 a agosto de 2008, essa discrepância entre as distâncias é mais perceptível.

Os agrupamentos divisivos mostraram algumas diferenças quanto a formação de grupos de poços. No entanto, para análise e discussão dos

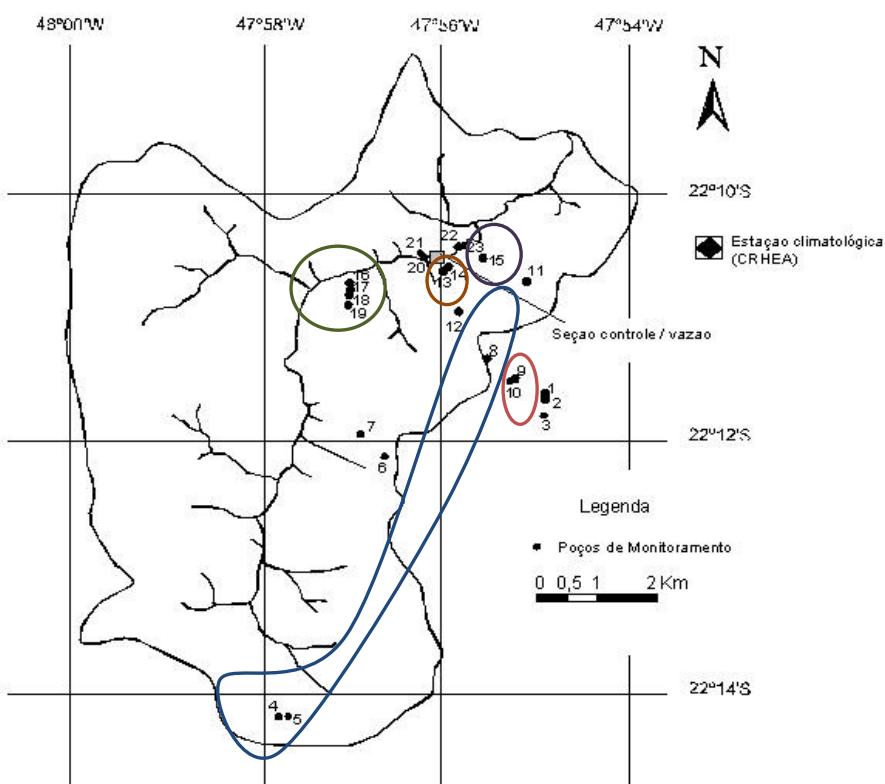
resultados fez-se uso dos dendrogramas resultantes das técnicas hierárquicas. Os gráficos a seguir mostram os grupos de poços, com as variações mensais na profundidade do nível d'água, ao longo dos quatro anos estudados.

A partir da análise dos resultados obtidos, foram formados cinco grupos de poços. O primeiro é formado pelos poços 4 e 5, que se encontram em áreas de pastagem e pelo poço 8, o qual após o ano de 2006, encontra-se localizado em área de cultivo de cana-de açúcar. No período anterior a agosto de 2006, na área em que o poço está localizado cultivava-se cítricos. O segundo grupo foi formado por quatro poços (16, 17, 18 e 19), instalados em área de mesma cultura (pastagem). O poço 15, em área de cultivo de cítricos, mas que até o ano de 2006 encontrava-se em área de cultivo de cana-de-açúcar, não formou grupo com nenhum outro poço. Os poços 13 e 14 encontram-se em áreas de cultivo de cítricos e estão agrupados. E, finalmente, o quinto grupo foi formado pelos poços 9 e 10, que estão localizados em regiões onde se cultiva eucalipto.

A análise dos grupos formados indica a variação do nível d'água como característica física de maior influência na formação de grupos de poços. No entanto o coeficiente de evapotranspiração, o qual está diretamente ligado à cultura da região, também se mostra como importante fator na definição dos agrupamentos.

**Tabela 3 - Coeficiente de correlação (Pearson) entre variações diárias do nível d'água nos poços monitorados.**

	P5	P8	P9	P10	P13	P14	P15	P16	P17	P18	P19
P4	<b>0,87</b>	0,20	0,19	-0,17	0,58	0,53	0,01	0,59	<b>0,61</b>	0,49	<b>0,79</b>
P5		0,04	0,05	-0,24	<b>0,64</b>	<b>0,64</b>	0,01	<b>0,69</b>	<b>0,64</b>	0,49	<b>0,73</b>
P8			0,37	0,21	-0,14	-0,12	-0,01	-0,09	-0,08	0,01	0,16
P9				<b>0,80</b>	-0,08	-0,08	-0,13	-0,05	-0,07	-0,02	0,08
P10					-0,30	-0,27	-0,03	-0,23	-0,32	-0,24	-0,28
P13						<b>0,93</b>	0,02	<b>0,71</b>	<b>0,82</b>	0,56	<b>0,63</b>
P14							0,01	<b>0,81</b>	<b>0,70</b>	0,42	0,47
P15								0,00	0,03	0,03	0,03
P16									<b>0,68</b>	0,38	0,46
P17										<b>0,89</b>	<b>0,76</b>
P18											<b>0,69</b>



**Figura 3 – Mapa de localização dos grupos de poços de monitoramento na bacia-piloto do Ribeirão da Onça, de acordo com as técnicas de agrupamento.**

Os agrupamentos mostram ainda que a porosidade exerce importante papel no processo de variação do nível d'água nos poços monitorados. Os agrupamentos que contém os poços 4, 5, 8, 9, 10, 13 e 14 estão localizados em regiões de alta porosidade (15,9% e 15,1%). Já os grupos formados pelos poços

16 a 19 localizam-se em áreas cuja porosidade é de 12,3%. E o poço 15 é o único localizado em área com porosidade de 8,5%, o que denota a importância de tal característica física na região. Com a mudança de coeficiente de evapotranspiração era esperado que o poço 15 mudasse o

agrupamento. O fato de não ter ocorrido a mudança ratifica a importância da porosidade como atributo de agrupamento.

Os resultados obtidos mostram a importância das variáveis utilizadas para o agrupamento, confirmando a necessidade de estudos mais aprofundados quanto aos demais fenômenos hidrológicos para o melhor entendimento do funcionamento da recarga na bacia do Ribeirão da Onça. Não é possível definir qual atributo é mais importante para a formação de grupos. Contudo, a variação do nível d'água nos poços foi definida como característica determinante por dois motivos: é a única característica que sofre variações nos valores ao longo dos anos estudados, e está diretamente ligada às demais características utilizadas, uma vez que a porosidade e o coeficiente de evapotranspiração influenciam o processo de infiltração da água no solo.

Após os agrupamentos, foi calculada a correlação entre os poços monitorados, utilizando as variações diárias do nível d'água. O objetivo foi confirmar a existência de forte relação entre variação do nível d'água e uso do solo (cultura). A tabela 3 mostra os valores obtidos.

Os valores obtidos mostram que há uma boa correlação entre as culturas existentes na área em que os poços estão localizados. Os valores relativos a variação do nível d'água no poço 4 geraram bons coeficientes de correlação quando relacionados aos valores de variações dos níveis d'água dos poços 5, 16, 17, 18 e 19, os quais, assim como o poço 4, estão localizados em área de pastagem. Essa relação se repete para o poço 5, também localizado em área de pastagem, o qual gerou coeficientes de correlação de altos valores quando relacionados aos poços supracitados.

Os coeficientes de correlação do poço 8 não resultaram em valores significativos com os demais poços, nem mesmo com os poços 13 e 14, localizados em área de cultivo de citrinos (assim como o poço 8). Os valores de nível d'água dos poços 13 e 14, quando correlacionados, resultaram em elevado coeficiente ( $r = 0,93$ ). É importante ressaltar que o poço 8 (22,9 m) possui uma profundidade média maior que os poços 13 e 14 (9,8 m e 6,6 m, respectivamente), o que pode explicar os diferentes valores de correlação.

No segundo semestre do ano de 2006 a cultura da área em que está localizado o poço 8 foi modificada; passou-se a cultivar cana-de-açúcar. Essa mudança pode provocar alterações no processo de infiltração da água, fazendo com que, consequentemente, a variação do nível d'água do

poço 8 também seja alterada, o que explica os coeficientes de correlação deste poço serem menores se comparados com os poços onde há o cultivo de citrinos.

A análise de correlação entre a variação do nível d'água no poço 15 e essa variação nos demais poços não resultou em altos coeficientes de correlação, o que pode ser explicado pelo fato de ser o único poço localizado em região de cultivo de cana-de-açúcar. Na região em que o poço está localizado também ocorreu mudança de cultura. Em 2006, passou-se a cultivar cítricos.

Os valores de nível d'água dos poços 9 e 10, localizados em culturas de eucalipto, geraram coeficiente de correlação elevado entre si. Em relação aos demais poços, o coeficiente de correlação das variações de nível d'água gerados são baixos, indicando que não há correlação entre os poços localizados em região de cultura de eucalipto e os poços localizados em regiões de outras culturas.

A análise dos valores dos níveis d'água nos poços 13, 14, 16 e 17 gerou altos coeficientes de correlação. Considerando que os poços 13 e 14 estão localizados em área de cítricos, enquanto que os poços 16 e 17 estão localizados em região de pastagem, conclui-se que a profundidade de nível d'água nesses poços é fator influenciador da boa correlação, uma vez que a profundidade dos quatro poços varia entre 5,1 m (poço 16) e 11,1 m (poço 17).

De maneira geral, os melhores coeficientes de correlação foram obtidos entre poços localizados em áreas de mesma cultura. Além disso, a profundidade do poço também influencia na correlação, conforme citado, no caso dos poços 13, 14, 16 e 17.

Quando comparados aos grupos formados a partir das técnicas de agrupamento, nota-se que os resultados são consistentes, já que os poços alocados em um mesmo grupo estão localizados em área de mesma cultura. Esse resultado é de grande importância pois permite determinar que o coeficiente de evapotranspiração, que está diretamente ligado a cultura da região, tem grande influência na variação do nível d'água nos poços monitorados.

Assim, é possível definir no mapa da bacia-piloto do Ribeirão da Onça a formação dos grupos, definidas pelos métodos estatísticos, conforme mostra a figura 3.

## CONCLUSÕES

Este trabalho teve como objetivo definir a influência de características do solo na variação do nível d'água em aquífero livre, apoiado em dados de monitoramento hidrogeológico. Foram monitorados por quatro anos, dados de precipitação e profundidade do nível d'água em poços localizados na Bacia Piloto do Ribeirão da Onça, área de recarga do Sistema Aquífero Guarani.

Os agrupamentos gerados pelas técnicas hierárquicas mostraram a existência de cinco grupos de poços. Todas as características físicas escolhidas como atributo (porosidade, coeficiente de evapotranspiração e variação do nível d'água nos poços) foram de grande importância. As diferentes culturas cultivadas na região podem ser o fator que define o comportamento do nível d'água nos poços, visto o fato destes terem se agrupado de acordo com a cultura cultivada na área em que se localizam.

A porosidade também é um fator determinante no agrupamento dos poços. O poço 15 é o único localizado em área de porosidade de 8,5%, e em todos os agrupamentos realizados não formou grupo com nenhum outro poço.

Os coeficientes de correlação mais elevados foram encontrados a partir da relação entre as variações diárias do nível d'água de poços localizados em área de mesma cultura. Isso significa que o coeficiente de evapotranspiração é fator fundamental para a variação do nível d'água subterrânea.

As técnicas estatísticas utilizadas se mostraram como importantes ferramentas para a análise de dados hidrológicos e, consequentemente, geram como resultado conhecimentos mais aprofundados no que concerne ao monitoramento dos processos de recarga de aquíferos, em especial o Sistema Aquífero Guarani, bem como para o planejamento do uso dos seus recursos hídricos. Ressalta-se, contudo, que os resultados obtidos nesse trabalho não definem relações diretas entre as variáveis estudadas. Apenas mostram padrões de comportamento, deixando clara a necessidade de estudos mais aprofundados sobre os fatores que interferem no processo de recarga de aquíferos.

## REFERÊNCIAS

BARRETO, C.E.A.G. Balanço hídrico em zona de afloramento do sistema aquífero guarani a partir de monitoramen-

to hidrogeológico em bacia representativa. Dissertação (Mestrado em Hidráulica e Saneamento). EESC-USP, São Carlos, SP.2006.

BORTOLUCCI, A. A. Caracterização Geológico-Geotécnica da região urbana de São Carlos-SP, a partir de sondagens de simples reconhecimento. Dissertação (Mestrado). Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 1983.

CONTIN NETO, D. Balanço hídrico em Bacia Hidrográfica situada em Região de Recarga do Aquífero Guarani. Tese (Doutorado em Hidráulica e Saneamento). EESC-USP, São Carlos, SP. 1987.

CORRÊA, U. M. P. Estudo das Águas Subterrâneas das Bacias Hidrográficas dos Rios Jacaré-Guaçu e Jacaré-Pepira no Estado de São Paulo. Dissertação (Mestrado em Hidráulica e Saneamento). EESC-USP, São Carlos, SP. 1995.

CUNHA, A T. Estimativa experimental da taxa de recarga na zona de afloramento do Aquífero Guarani, para a região de São Carlos – SP. Dissertação (Mestrado em Hidráulica e Saneamento). EESC-USP, São Carlos, SP. 2003.

DAVINO, A. Considerações Hidrogeológicas Preliminares sobre a Bacia Hidrográfica do Ribeirão da Onça, São Carlos, SP. Faculdade de Filosofia Ciências e Letras de Ribeirão Preto. Relatório apresentado a FAPESP. Página irregular. Ribeirão Preto. 1984.

DEMÉTRIO, C.G.B., ZOCCHI, S. S. Modelos de regressión aplicados a estudios en Edafología. 1.ed. Guatemala. 2003.

DE VRIES, J. J.; SIMMERS, I. Groundwater recharge: an overview of processes and challenges. Hydrogeology Journal. 10:5-17. DOI 10.1007/s10040-001- 0171-7.2002.

EVERITT, B.S., LANDAU, S., LEESE, M. Cluster Analysis. 4.ed. – London: Ed. Arnold. 2001.

FERRARO, F. Artigo de intercambio apresentado à Escola de Engenharia de São Carlos. 2006.

GIAMPÁ, C. E. Q.; SOUZA, J. C. Potencial Aquífero dos Basaltos da Formação Serra Geral no Estado de São Paulo. 2o CABAS, Salvador – BA. 1982.

GOMES, L.H. Determinação da recarga profunda na bacia piloto do ribeirão da onça em zona de afloramento do aquífero guarani. Dissertação (Mestrado em Hidráulica e Saneamento). EESC-USP, São Carlos, SP. 2008.

JOBSON, J.D. Applied Multivariate Data Analysis, vol. 2:Categorical and Multivariate Methods. Springer-Verlag. 1992.

MATTOS, I. S. Aplicação de um Modelo de Balanço Hídrico na Bacia do Rio Jacaré-Guaçu. Dissertação (Mestrado em Hidráulica e Saneamento). EESC-USP, São Carlos, SP. 1987.

- MAYER, T.D., CONGDON, R.D. Evaluating Climate Variability and Pumping Effects in Statistical Analyses. *Ground Water* 0(0).1-16. 2007.
- MAZIERO, T.A. Monitoramento de água subterrânea em área urbana: aspectos quantitativos. Dissertação (Mestrado em Hidráulica e Saneamento). EESC-USP, São Carlos, SP. 2005.
- MILLIGAN, G.W.; COOPER, M.C. Methodology Review: Clustering Methods. *Applied Psychological Measurement*. v.11. 1987.
- MUÑOZ-DIAZ, D.; RODRIGO, F.S. Spatio-temporal patterns of seasonal rainfall in Spain (1912-2000) using cluster and principal component analysis: comparison. *Anales Geophysicae*, 22. 1435-1448. 2004.
- PANDA, U. C.; SUNDARAY, S.K.; RATH, P.; NAYAK, B.B.; BHATTA, D. Application of factor and cluster analysis for characterization of river and estuarine water systems – A case study: Mahanadi River (India). *Journal of Hydrology*. 331. 434-445. 2006.
- POMPÊO, C. A. Balanço Hídrico da Zona Não-Saturada do Solo na Bacia do Ribeirão da Onça (SP). Tese (Doutorado em Hidráulica e Saneamento). EESC-USP, São Carlos, SP. 1990.
- RAMOS, M. C. Divisive and hierarchical clustering techniques to analyze variability of rainfall distribution patterns in a Mediterranean region. *Journal of Hydrology*, New York, v: 57. 123-138. 2001.
- SPIELGEL, M. R. Estatística. Coleção Schaum, Editora McGraw-Hill do Brasil, Rio de Janeiro, 580 p. 1970.
- SYSTAT. Manual do Usuário. Versão 10. SPSS Science, USA. 2007.
- VEGA, M.; PARDO, R.; BARRADO, E.; DEBAN L. Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. *Water Research*: v.32.3581-3592. 1998.
- VIEIRA, S. Bioestatística: tópicos avançados. 2.ed. Rio de Janeiro: Elsevier. 2004.
- WENDLAND, E.; BARRETO, C.; GOMES, L. H. Water balance in the Guarani Aquifer outcrop zone based on hydrogeologic monitoring. *Journal of Hydrology* .v. 342. 261-269. 2007.
- ZHOU, Y. Sampling Frequency for Monitoring the Actual State of Groundwater Systems. *Journal of Hydrology*, v. 180. 301-318. 1996.

*Influence of Soil Characteristics on Water Level Fluctuations in the Recharge Area of the Guarani Aquifer System*

**ABSTRACT**

The main objective of this research is to analyze the influence of soil characteristics on water level fluctuations in the recharge area of the Guarani Aquifer System. Water level fluctuations were monitored in 12 wells, during the period from September 2004 to August 2008, in the Ribeirão da Onça hydrographic basin. The Climatology Station of CRHEA-USP provided the data on daily precipitation for the same period. Two statistical methods were used to achieve the proposed objectives: clustering techniques and correlation analysis. The clustering techniques show that all the soil characteristics used in this study are very important for water level fluctuations, but the dendograms suggest that the most important one is the evapotranspiration coefficient. The correlation between monthly variations of the well water levels generated high coefficients ( $r>0.5$ ). This result, analyzed together with clustering, shows that soil use (crops) is highly important for water level fluctuations, emphasizing the need for more specific studies of soil characteristics; however, porosity was also shown to be a major characteristic. This research shows that studies of soil characteristics and soil use are essential to the project of monitoring groundwater resources, especially in the recharge areas.

**Key-words:** groundwater, rainfall, statistical methods, clustering