

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

**APLICAÇÃO DE ESTATÍSTICA DESCRITIVA E MULTIVARIADA EM ANÁLISES
QUÍMICAS E FÍSICO-QUÍMICAS DE ÁGUAS SUBTERRÂNEAS DA ÁREA DE RECARGA
DO SISTEMA AQUÍFERO GUARANI (SP)**

Monografia

TF – 16/05

Aluno: Bruno Carlos Iatallese Ferreira Pinto

Orientadora: Alexandra Vieira Suhogusoff

Co-Orientador: José Luiz Albuquerque Filho

São Paulo
Dezembro, 2016

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS

**APLICAÇÃO DE ESTATÍSTICA DESCRITIVA E MULTIVARIADA EM ANÁLISES
QUÍMICAS E FÍSICO-QUÍMICAS DE ÁGUAS SUBTERRÂNEAS DA ÁREA DE RECARGA
DO SISTEMA AQUÍFERO GUARANI (SP)**

Monografia
TF – 16/05

Aluno: Bruno Carlos Iatallese Ferreira Pinto



Orientadora: Alexandra Vieira Suhogusoff
Co-Orientador: José Luiz Albuquerque Filho

IGc-Instituto de Geociências



Aplicação de estatística descritiva e multivariada em
análises químicas e físico-químicas de
M27124

São Paulo
Dezembro, 2016

"Ora, com os diabos, fui imortal até morrer."

Vovô Jake, em "Fup" de Jim Dodge.

AGRADECIMENTOS

Eu cheguei a pensar que a USP duraria para sempre. Largar um emprego para começar uma nova faculdade, de período integral, como a Geologia – USP não foi fácil, mas definitivamente valeu a pena.

Primeiro eu gostaria de agradecer a toda minha família pelo apoio desde sempre, inclusive pelo nesta nova jornada. Eu amo muito vocês. Esta conquista é dedicada especialmente ao meu Vô Nelson e Vó Lourdes, as pessoas que eu mais amo no mundo. Obrigado por tudo. Vô, escrevi nosso nome na USP. E isso foi possível graças ao apoio da nossa família.

Não posso deixar de agradecer meus amigos daqui do ABC. Nós crescemos juntos e eu tenho vocês como irmãos para mim. Aqui vai um agradecimento especial para o Canisso Derico Careca, que me incentivou desde o começo e por ter salvo minha vida quando eu caí da árvore. Gostaria de agradecer também ao Breno Ramone, que me chamou pra trocar uma idéia sempre que via que o bicho tava pegando pro meu lado. Valeu manos.

Aos meus parceiros de banda: Lama Bomba, Carnage, Banda Rosana e outras bandas que nunca chegaram a ter um nome. Foram dias maneiros.

Gostaria de agradecer aqui à uma cabritinha chamada Beatriz, compartilhamos uma parceria como nunca imaginei. Te amo muito. Beh!

Se hoje eu celebro esta conquista, não posso deixar de agradecer às pessoas que pavimentaram o caminho e tornaram este momento possível. Gostaria de agradecer à grande maioria dos professores do IGc-USP e demais institutos onde estudei, em especial à Prof. Dr. Alexandra Suhogusoff, pela orientação neste trabalho de formatura, ao José Luis Albuquerque Filho, coorientador, Ana Maciel e ao Luís Carlos Ferrari cuja ajuda foi de grande importância para o êxito deste trabalho. Agradeço também aos demais funcionários do IGc e da USP em geral, dos motoristas ao pessoal do bandeirão.

Durante o curso tive uma oportunidade que mudou meu modo de ver o mundo, que foram os dias que passei em Baltimore, MD. Obrigado a todos os envolvidos, ao pessoal do Granny's, em especial à Cheff Annette e ao Luís Rojas.

Vários motivos me levaram à Geologia, especialmente à área Ambiental, minha meta desde o início. Um dos principais motivos foi a minha vontade de fazer com que meu trabalho não se restringisse apenas a uma fonte de renda, mas sim que tivesse um impacto positivo no planeta e na vida das pessoas. Este objetivo felizmente foi e está sendo realizado. Agradeço, portanto, ao pessoal da BTX, por tudo que aprendi e pelos grandes amigos que formei lá. São dias dos quais eu me recordo com muito carinho. Obrigado ao pessoal da Geoklock, também pela amizade, pela confiança e pelas portas que abriram para mim.

Durante o curso, eu tive a honra de, praticamente sem querer, conhecer um cara que se tornou um grande parceiro e um irmão, o PP. Valeu pelos dias em que dividimos uma rep. Foram com certeza os dias mais loucos.

No meio do curso, eu praticamente caí de turma, e o que deveria ser um motivo de preocupação se tornou um motivo de grande alegria, pois me permitiu conhecer pessoas que se tornaram grandes amigos. Tenho orgulho de dizer que faço parte do Rebanho 55! Obrigado por tudo, Café e Hugo, grandes parceiros, Júlio, Bundinha, Cauê, Lucão, Mormaço, Tênia, Manja e Constipado. Que essa amizade se estenda muito além dos dias de faculdade.

Aos amigos e colegas de outros anos, valeu pelos momentos que compartilhamos.

Para o final eu reservei os agradecimentos mais importantes que gostaria de fazer. Hoje eu olho para trás e vejo como evolui como pessoa, e como evoluímos juntos, graças à nossa amizade. Nossos dias de campo, nossos rolês, viagens, almoços no bandeirão, festas no CEPEGE, nas reps dos parceiros, BIFE e tudo o mais ficarão marcados pra sempre em minha memória, e foi o que fez toda essa jornada valer a pena. Valeu por estarem juntos nos momentos de grande alegria e nos momentos em que precisei. Foi uma grande lição de amizade. Só tenho a agradecer por este tempo juntos. Essa vai para toda a Família GEOBINHO: Azeite e Esperma, dois caras que considero como irmãos, Bolket (você sabe como é fazer Geo como segunda facul, e nossa amizade sempre me deu força pra continuar), Eunuco (vamo do tá ligado, tá ligado? "lhhhhh" estar pronto é tudo), Pseudo (tiozão), Preps (tiozinho), Kudu, Lontra, Always, Bytcha, Gardenal (Hei....tor), Kankro, Molestado, Palestina, Soka (Pei), Xupin, Marijuana (vou na paz, já fui), Orka (uma das pessoas mais queridas na minha vida).

Obrigado por tudo.

SUMÁRIO

RESUMO	6
ABSTRACT	7
1. INTRODUÇÃO	8
2. ÁREA DE ESTUDOS	9
3. METAS E OBJETIVOS	9
4. LEVANTAMENTO BIBLIOGRÁFICO	10
4.1. SISTEMA AQUÍFERO GUARANI	10
4.2. GEOLOGIA	10
4.3. HIDROGEOLOGIA	11
4.4. HIDROQUÍMICA	13
4.5. ANÁLISE DE COMPONENTES PRINCIPAIS	13
4.6. ANÁLISE DE AGRUPAMENTO HIERÁRQUICO (HCA)	16
5. MATERIAIS E MÉTODOS	18
5.1. TRATAMENTO DOS DADOS E VALIDAÇÃO DAS ANÁLISES	18
5.2. ANÁLISE ESTATÍSTICA DESCRITIVA E MULTIVARIADA	19
5.3. ELABORAÇÃO DE MAPAS, GRÁFICOS E DIAGRAMAS	20
6. RESULTADOS OBTIDOS	21
6.1. TRATAMENTO DOS DADOS E VALIDAÇÃO DAS ANÁLISES	21
6.2. ANÁLISE ESTATÍSTICA DESCRITIVA E MULTIVARIADA	22
6.3. ELABORAÇÃO DE MAPAS, GRÁFICOS E DIAGRAMAS	25
7. INTERPRETAÇÃO DOS RESULTADOS	30
8. CONCLUSÕES	38
9. REFERÊNCIAS BIBLIOGRÁFICAS	39

RESUMO

O Sistema Aquífero Guarani (SAG) trata-se da principal reserva subterrânea de água doce transfronteiriça da América do Sul. O Brasil concentra 90% da exploração atual do aquífero, sendo a maior parte no Estado de São Paulo, onde aflora como uma faixa alongada de direção NE-SW totalizando 17.376 km². Esta região constitui a sua principal zona de recarga regional, e é uma das regiões de maior vulnerabilidade à ação antrópica e à interação com as águas superficiais.

Análises físico-químicas de águas subterrâneas cedidas pelo Instituto de Pesquisas Tecnológicas (IPT), provenientes de poços de abastecimento localizados na área de recarga do SAG no Estado de São Paulo, foram submetidas ao tratamento estatístico descritivo e de análise multivariada por meio das técnicas de Análise de Componente Principal (ACP) e Análise de Agrupamento Hierárquico (HCA).

A ACP forneceu 4 Componentes Principais (CPs), que explicam 80,90% da variância total do sistema. A CP-1, interpretada como a componente da salinidade, possui forte relação com o teor de Sólidos Totais Dissolvidos, a Condutividade Elétrica e com os íons Ca^{2+} , Mg^{2+} , K^+ , HCO_3^- e SiO_2 . A CP-2 apresentou forte correlação inversa com os íons NO_3^- e Cl^- , e foi interpretada como influência da atividade antrópica. A CP-3 está inversamente relacionada com o íon Na^+ e a CP-4 com o íon SO_4^{2-} .

A HCA dividiu as análises em 2 grupos principais, o Grupo A de alta salinidade e altos teores para os principais íons, e o Grupo B de baixos teores em geral.

O Grupo A foi dividido nos subgrupos SG-1 e SG-2. No primeiro destacam-se os teores para NO_3^- e Cl^- , e no segundo os teores para SiO_2 , HCO_3^- e Ca^{2+} , o que foi interpretado como consequência da dissolução dos aluminossilicatos e da cimentação carbonática presentes nas formações Pirambóia e Botucatu.

O Grupo B, também foi dividido em dois subgrupos SG-3 e SG-4, sendo que a principal diferença entre os grupos reside na característica mais sódica e menores teores em geral do SG-4.

Palavras-chave: Análise de Componentes Principais; Análise de Agrupamento Hierárquico; Hidrogeoquímica; Sistema Aquífero Guarani

ABSTRACT

The Guarani Aquifer System (GAS) is the main groundwater reserve in South America. Brazil holds 90% of the actual exploitation of the aquifer, mostly in the State of São Paulo, where it outcrops in a NE-SW strip of 17.376 km². This area is the main regional recharge zone, and it's one of the most vulnerable zone regarding human activities and the interaction with surface water.

Physical-Chemical analysis of groundwater were given by the Instituto de Pesquisas Tecnológicas (IPT), collected from supplying wells located in the recharge area of GAS in the State of São Paulo. These analyses have been submitted to descriptive statistics and multivariate analysis, through Principal Component Analysis (PCA) and Hierarchical Clustering Analysis (HCA).

ACP provided 4 Principal Components (PCs), which explain 80,90% of the total variance of the system. PC-1, was interpreted as the "salinity" component, and holds Strong relation with Dissolved Total Solids, Electric Conductivity, and the ions Ca²⁺, Mg²⁺, K⁺, HCO³⁻ and SiO₂. PC-2 showed strong inverse correlation with the ions NO³⁻ and Cl⁻ and was interpreted as due the human activities. CP-3 is inversely related with Na⁺ and PC-4 with the ion SO₄²⁻.

The HCA divided the analysis in 2 main groups, Group A with high salinity and high concentration for the most of the ions, and Group B with lower concentrations.

Group A was divided into 2 subgroups SG-1 and SG-2. The first shows high concentrations for NO³⁻ and Cl⁻ and the second for SiO₂, HCO₃⁻ and Ca²⁺, which was interpreted as the dissolution of silicates and carbonate cement of the Pirambóia and Botucatu formations.

Group B, was also divided into two subgroups SG-3 e SG-4, where the main difference resides in the higher sodic characteristic and lower concentrations of the SG-4.

Keywords: Principal Component Analysis; Hierarchical Clustering Analysis; Hydrogeochemistry; Guarani Aquifer System.

1. INTRODUÇÃO

O Sistema Aquífero Guarani (SAG) constitui a principal reserva subterrânea de água doce transfronteiriça da América do Sul, estendendo-se pelo Brasil, Uruguai, Paraguai e Argentina. No Brasil engloba os estados de São Paulo, Goiás, Mato Grosso, Mato Grosso do Sul, Minas Gerais, Paraná, Rio Grande do Sul e Santa Catarina.

Trata-se de uma importante reserva estratégica para o abastecimento da população e o desenvolvimento das atividades econômicas.

De acordo com OEA (2009), estima-se que 92 milhões de pessoas residam sobre o SAG, entre 80 e 90% em áreas urbanas. Dessa população, 90% são de nacionalidade brasileira, o que equivale a 82,8 milhões de pessoas, ou 43% da população do país. Aproximadamente 90% da exploração atual do SAG ocorre em território brasileiro, sendo que a maior parte se concentra no Estado de São Paulo, onde 70% da água extraída é destinada ao abastecimento público, 24% para uso industrial, 5% para uso rural e 1% para recreação. Embora estes números variem entre cada estado, o principal uso para essas águas é o abastecimento público.

A porção aflorante do SAG, no Estado de São Paulo, corresponde a uma faixa alongada de direção NE-SW totalizando 17.376 km² e constitui a sua principal zona de recarga regional. Esta é uma das regiões do aquífero de maior vulnerabilidade à ação antrópica e à interação das águas subterrâneas com as águas superficiais.

De acordo com Gomes *et al.* (2006), está inserida no Planalto Médio Paulista, e pode ser dividida em três faixas: norte, central e centro-sul, nas quais predominam a cultura de cana de açúcar ao longo das três faixas, mas destacam-se também, pastagem nas faixas central e centro-sul, e arroz irrigado e citros na faixa central.

Além do uso agrícola da terra, diversos centros industriais estão localizados dentro dos limites da área aflorante do SAG, como São Carlos, Araraquara, Ribeirão Preto, entre outros.

Em face da importância do recurso hídrico subterrâneo para o abastecimento das populações e desenvolvimento agrícola e industrial e do uso e ocupação dos solos, associada à alta vulnerabilidade desta porção do aquífero, torna-se indispensável o conhecimento das características hidroquímicas do aquífero, de modo que se possa garantir sua potabilidade e orientar as ações que visam proteger a qualidade das águas subterrâneas.

Neste trabalho, amostras de água subterrânea provenientes da área de recarga do SAG no Estado de São Paulo, foram submetidas à análise estatística descritiva e multivariada com o objetivo de se identificar parâmetros físico-químicos de comportamento semelhante e agrupar as amostras de acordo com suas características hidrogeológicas.

2. ÁREA DE ESTUDOS

A Figura 1 apresenta a área de estudos, destacando a localização do SAG na América do Sul, bem como sua porção aflorante, com destaque para o Estado de São Paulo.

A área de estudos relativa aos poços analisados encontra-se inserida na porção centro-oeste do Estado de São Paulo, compreendendo uma faixa contínua de orientação SW-NE que se estende desde os municípios de Tejubá e Itai (sudoeste) até os municípios de Rifaina e Cássia dos Coqueiros (nordeste).

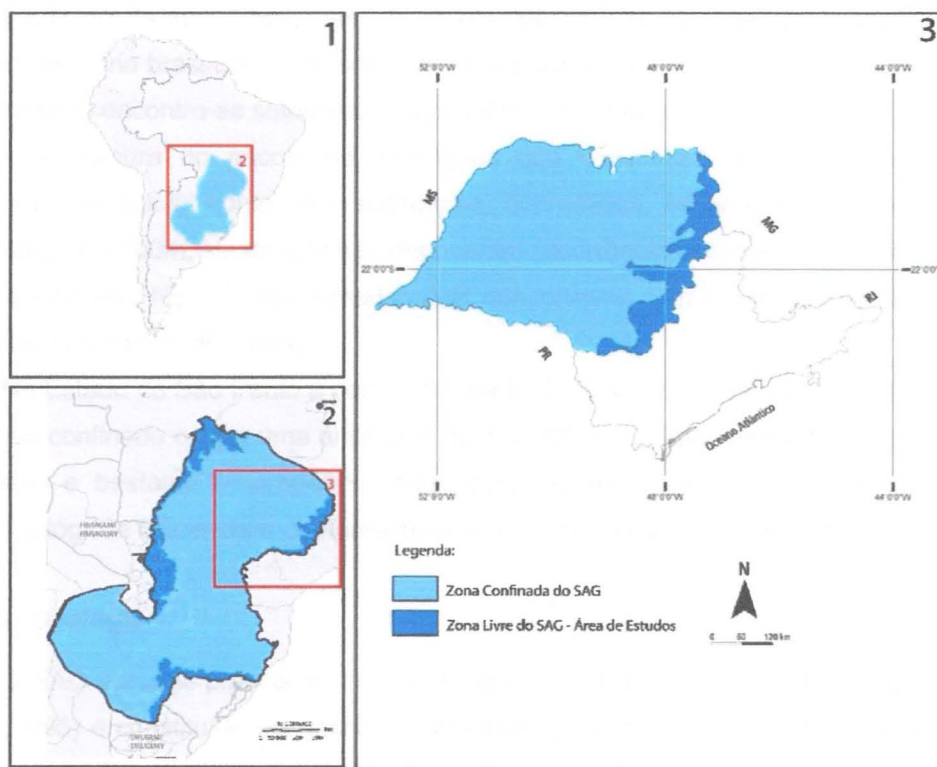


Figura 1 – Área de estudo (modificado de OEA (2009) e Iritani e Ezaki (2009)).

3. METAS E OBJETIVOS

Esse trabalho de formatura tem como principal objetivo a caracterização hidrogeoquímica do Sistema Aquífero Guarani em sua área de recarga no Estado de São Paulo. Análises físico-químicas e químicas de águas subterrâneas provenientes de poços de abastecimento público e privado foram submetidas ao tratamento estatístico descritivo e de análise multivariada por meio das técnicas de Análise de Componente Principal (ACP) a fim de identificar padrões de comportamento semelhantes entre os parâmetros analisados e Análise de Agrupamento Hierárquico (HCA – *Hierarchical Clustering Analysis* ou apenas *Clustering*.) de modo que as amostras possam ser separadas em conjuntos que possuem características hidrogeoquímicas semelhantes.

4. LEVANTAMENTO BIBLIOGRÁFICO

4.1. Sistema Aquífero Guarani

O Sistema Aquífero Guarani (SAG) era inicialmente denominado de Aquífero Gigante do Mercosul por Araújo *et al.* (1995). O nome "Guarani" foi uma sugestão do geólogo uruguaio Danilo Antón, como forma de homenagear os índios guaranis que originalmente habitavam esta área.

O SAG é um aquífero sedimentar formado pelos arenitos mesozóicos da Bacia Sedimentar do Paraná. Possui área de 1.194.000 km², sendo que aproximadamente 70% está em território brasileiro, 19% argentino, 6% paraguaio e 5% uruguaio. Cerca de 90% de sua extensão encontra-se sotoposta aos basaltos da Formação Serra Geral.

A espessura do pacote de formações que constituem o SAG é controlada por estruturas geológicas como altos estruturais, depressões, falhamentos e borda da bacia. Pode chegar a 800m, sendo que seu depocentro (ocorrências superiores a 500m) ocorre ao longo de um eixo NE-SW, aproximadamente concordante com o eixo deposicional da Bacia do Paraná (Araújo *et al.*, 1995).

No Estado de São Paulo a porção aflorante do SAG ocupa uma área de 17.700 km² e a porção confinada ocupa uma área total de 137.700 km². Trata-se de um sistema aquífero complexo e bastante heterogêneo, onde condicionantes geotectônicas, sedimentares e geomorfológicas influenciam diretamente suas características (Iritani e Ezaki, 2009).

4.2. Geologia

O SAG abrange praticamente toda a extensão da Bacia do Paraná e, segundo Araújo *et al.* (1995) é constituído por rochas reservatório do Triássico e do Jurássico, subjacentes aos derrames basálticos da Formação Serra Geral, datados do Cretáceo.

Assine *et al.* (2004) propõem a divisão dos arenitos sotopostos à Formação Serra Geral em Formação Pirambóia e Formação Botucatu, uma vez que ocorrem mudanças evidentes entre ambas sobretudo em sua cor e na dimensão das estratificações cruzadas. Estas formações afloram na borda nordeste da Bacia do Paraná, incluindo o Estado de São Paulo, mas são reconhecidas em subsuperfície em grande parte da bacia.

A Formação Pirambóia, de idade triássica, é constituída predominantemente por arenitos de granulação muito fina com teores de argila acima de 20% (Rocha, 1997). Possuem estratificação cruzada de médio porte, com ocorrências restritas de estratificação cruzada de grande porte. São observadas também estratificações cruzadas de baixo ângulo e plano-paralelas.

Seu ambiente deposicional é interpretado como campos de dunas e interdunas úmidas, onde se observa uma migração para um ambiente mais desértico em direção ao

topo da unidade. Destaca-se, também, a ocorrência de fácies de arenitos com seixos, por vezes conglomeráticos, de origem fluvial (Brighetti, 1994).

A Formação Botucatu, por sua vez, é descrita por Assine *et al.* (2004) como sendo composta majoritariamente por arenitos com estratificação cruzada, planar ou acanalada, de médio a grande porte. O ângulo de mergulho é alto nas camadas frontais dos estratos cruzados tomando-se tangencial na base, o que indica que o tamanho dos grãos diminui do topo para a base. Observa-se também, alternância de camadas de arenitos finos e arenitos médios. A Formação Botucatu é interpretada como um grande paleodeserto (sistema eólico seco), de idade jurássica.

A camada confinante superior do SAG é conhecida como Formação Serra Geral, que corresponde a uma série de derrames de lava de composição predominantemente básica, ocorridos no Cretáceo Inferior e ligados à abertura do Atlântico Sul. De acordo com Peate *et al.* (1992), a Formação Serra Geral é constituída por basaltos toleíticos e andesitos basálticos, com ocorrência subordinada de riolitos e riodacitos. Destacam também as atividades ígneas intrusivas, toleíticas e básicas, na forma de sills e diques associados. Os derrames basálticos, essencialmente sub-horizontais, assentam-se diretamente sobre os arenitos da Formação Botucatu, sendo que o contato entre estas duas formações é concordante e muitas vezes interdigitado (Marques e Ernesto, 2004).

Em sua base, o SAG está confinado por rochas permotriássicas de baixa permeabilidade que no Estado de São Paulo é representada pelos argilitos, siltitos, folhelhos e arenitos finos da Formação Corumbataí.

Destaca-se que os depocentros de lava da Formação Serra Geral (que provocaram a subsidência do aquífero), a ativação regional do sistema de falhamentos, a ativação dos arcos do Rio Grande e Ponta Grossa e o soerguimento das bordas da bacia (responsáveis pela delineação de um contorno estrutural concêntrico, e que permitiram a erosão das camadas superiores, possibilitando o afloramento das unidades formadoras do Aquífero Guarani) são fatores importantes na conformação geométrica atual do SAG (Araújo *et al.*, 1995).

4.3. Hidrogeologia

As feições estruturais possuem importante papel na hidrodinâmica do SAG. Lineamentos podem influenciar o fluxo no sistema formando vias preferenciais ou anisotropia de permeabilidade. Sendo assim, Araújo *et al.* (1995) dividem o SAG em dois regimes hidrogeológicos. Esses regimes são separados pelo enxame de diques do Arco de Ponta Grossa que se estende até a calha do Rio Paraná e age como uma grande barreira hidráulica.

O primeiro regime hidrogeológico, localizado ao norte do Arco de Ponta Grossa, possui fluxo preferencial de norte para sudoeste, a partir da sua porção livre, nos estados de São Paulo, Goiás, Mato Grosso e Mato Grosso do Sul.

O segundo regime, ao sul do referido Arco, distingue-se pelos altos gradientes hidráulicos e por proeminentes áreas de descarga. O sentido preferencial do fluxo, a partir da borda leste, é para sudoeste.

A Figura 2 apresenta um modelo conceitual regional para o fluxo de águas subterrâneas do SAG.

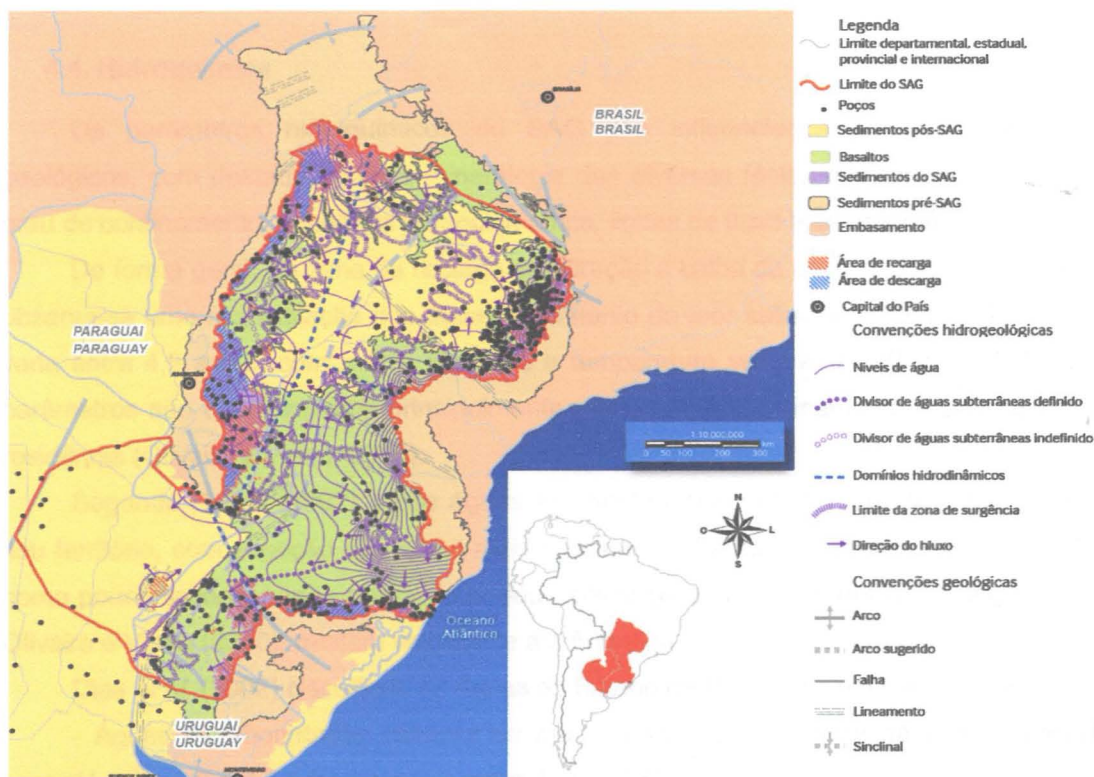


Figura 2 – Mapa potenciométrico (OEA, 2009)

A zona livre do SAG constitui sua principal zona de recarga. Esta área, no entanto, pode revelar uma dinâmica própria de fluxo local, até mesmo interagindo com as drenagens locais, em virtude de sua geomorfologia (OEA, 2009).

Barretos (2006) estima que a recarga direta do sistema, em sua zona aflorante, está em torno de 400 mm/ano e, em sua zona confinada, entre 0 a 40 mm/anuais.

Cabe destacar que a interação entre o SAG e o aquífero fraturado representado pela Formação Serra Geral vem sendo apontada por diversos autores. DAEE (1974) já indicava a ocorrência de troca de águas entre esses dois sistemas. Spiller (2005) informa que a ocorrência de água subterrânea na Formação Serra Geral está associada à presença de fraturamentos e do basalto vesicular, com suas estruturas características.

O relatório da OEA (2009) informa que os gradientes hidráulicos, na porção que compreende o Estado de São Paulo, variam de 3 a 5 m/km na porção livre, reduzindo para cerca de 0,1 m/km em sua porção confinada.

Estudos realizados pelo Departamento de Águas e Energia Elétrica (DAEE, 1974) na região de Ribeirão Preto (SP) apontam valores de transmissividade entre 65 e 240 m²/dia, condutividade hidráulica de 3,5 m/dia e coeficiente de armazenamento entre 10⁻⁵ e 10⁻³. A heterogeneidade dos valores obtidos para os parâmetros hidrodinâmicos dos poços refletem as diferentes condições geológicas e estruturais.

4.4. Hidroquímica

Os parâmetros hidroquímicos do SAG são influenciados por diversos fatores geológicos, com destaque para a mineralogia das diversas fácies que formam o aquífero, grau de confinamento, superfície potenciométrica, linhas de fluxo e estratigrafia.

De forma geral, da zona de recarga em direção à calha da bacia, no oeste do Estado, observa-se uma alcalinização das águas e aumento do teor salino e da temperatura. O pH varia entre 4,9 a 9,2, com média de 6,9 e a temperatura varia entre 20°C a 66°C. Estes parâmetros são condicionados principalmente pelo grau de confinamento e pelo tempo de residência (Oliveira e Vieira, 2010).

Segundo Rocha (1997), estas águas apresentam boa potabilidade na maior parte de seu território, com exceção das áreas mais profundas. O autor classifica as águas do SAG como pouco salinas, sendo o teor de resíduo sólido geralmente inferior a 300 mg/L, porém Oliveira e Vieira (2010), relatam valores de até 800 mg/L.

Dias *et al.* (2002) classificou as águas do Estado de São Paulo em três famílias:

- Águas bicarbonatadas cálcicas ou magnesianas que ocorrem na zona confinada, com pH entre 5,3 a 8,5 e temperatura entre 21°C a 36°C.
- Águas bicarbonatadas sódicas, que também ocorrem na zona confinada, pH maior que 8,5 e temperatura entre 23°C e 34°C.
- Águas cloro-sulfatadas cálcicas ou magnesianas, que ocorrem na zona aflorante e tem seu pH entre 4,8 a 7 e temperatura entre 20°C e 27°C.

Rocha (1997) ainda destaca a ocorrência de concentrações anômalas de fluoreto, em locais como Presidente Prudente (SP), associadas a feições estruturais.

4.5. Análise de Componentes Principais

A Análise de Componentes Principais (ACP) tem por principal objetivo a redução do número de dimensões das observações (Hardle e Simar, 2007). Por meio desta técnica de análise multidimensional, é possível diminuir o número de variáveis para evidenciar relações

ou classificar elementos de uma população, em um espaço multivariado, ou seja, quando existem diversas variáveis interagindo sobre determinado fenômeno (Andriotti, 1997).

Segundo Davis (1973), uma matriz pode ser expressa na forma de uma série de vetores em um espaço multidimensional, os quais podem ser considerados eixos de um elipsóide multidimensional. Os autovetores de uma matriz orientam a direção dos eixos principais deste elipsóide, e os autovalores indicam o tamanho desses eixos.

Componentes principais (CPs), portanto, são os autovetores de uma matriz de variância-covariância ou de uma matriz de correlação e o método conhecido como ACP busca encontrar esses eixos e medir suas magnitudes.

Uma vez que a ACP parte de uma matriz de variância-covariância simétrica, os autovetores serão mutualmente ortogonais, e as medidas no diagrama serão feitas na mesma unidade das variâncias.

Em um exemplo utilizado por este autor, a partir de um determinado conjunto de dados, foi obtida a matriz de variância-covariância s^2 abaixo, de foram calculados os autovetores I e II.

$$[s^2] = \begin{bmatrix} 20,3 & 15,6 \\ 15,6 & 24,1 \end{bmatrix}$$

$$I = \begin{bmatrix} 0,66 \\ 0,75 \end{bmatrix}$$

$$II = \begin{bmatrix} 0,75 \\ -0,66 \end{bmatrix}$$

Os autovalores são de 37,9 para I e 6,5 para II.

A variância total pode ser definida como a soma das variâncias individuais. Uma vez que as variâncias estão localizadas na diagonal da matriz de variância-covariância, é equivalente a calcular o traço da matriz. Para as variáveis originais, a variância total é de $20,3 + 24,1 = 44,4$, sendo que a primeira variável corresponde a $20,3/44,4$, ou seja, aproximadamente 46% da variância total, e a segunda variável corresponde ao restante, 54%.

Neste ponto, se for necessário reduzir o número de dimensões, a remoção de uma variável faz com que os novos dados expliquem apenas 54% da variância total, ou seja, há uma perda significativa de informação.

Uma das propriedades dos autovalores é que a soma dos mesmos é igual ao traço da matriz ($37,9 + 6,5 = 44,4$), e uma vez que eles representam a magnitude dos eixos principais, a soma deles representa a variância total do sistema, e a divisão dos autovalores pelo traço da matriz fornece a variância explicada por cada autovalor.

Desta forma, o autovetor I explica $37,9/44,4$, ou 86%, da variância total do sistema, e o autovetor II é responsável por 14%.

Se multiplicarmos a matriz contendo os dados originais [X], pela matriz contendo os autovetores [U], vamos obter a matriz [SR], cujos elementos são denominados *scores* das componentes principais. Os elementos da matriz [U] por sua vez são chamados de *loadings* das componentes principais.

$$[X].[U] = [SR]$$

Classificando e plotando-se as variáveis originais em um espaço definido pelas componentes principais obtidas, chega-se a uma nova matriz de variância-covariância. Os autovetores calculados para essa matriz são semelhantes aos autovetores calculados inicialmente:

$$I = \begin{bmatrix} 0,68 \\ 0,74 \end{bmatrix}$$

$$II = \begin{bmatrix} 0,74 \\ -0,68 \end{bmatrix}$$

No entanto, os novos autovalores são extremamente diferentes, sendo o primeiro 44,2 e segundo 0,2. Dessa forma, a variância explicada pelo primeiro autovetor é de 44,2/44,4, que representa mais de 99% da variância total do sistema.

A esta altura, portanto, é possível remover as demais componentes sem acarretar em grande perda de informação, uma vez que a primeira componente principal explica a maior parte da variância.

Este método vem sendo muito aplicado em hidrogeologia, uma vez que análises químicas e de parâmetros físico-químicos de água, quase sempre, representam um conjunto de dados multivariado. Diversos autores vêm aplicando o método com o objetivo de avaliar as condições hidrogeoquímicas da água subterrânea (Avtar *et al.*, 2013; Yidana *et al.*, 2010; Zakhem e Hafez, 2015), as interações entre a água superficial e subterrânea (Menció *et al.*, 2008), os impactos da atividade humana em água subterrânea (Kim *et al.*, 2014), entre outros.

A Figura 3 apresenta uma representação visual da aplicação do método. Neste exemplo, sete conjuntos de amostras foram analisados para 14 parâmetros inorgânicos. São apresentados os resultados obtidos para as duas primeiras CPs. Desta forma, torna-se evidente a relação entre determinados íons, de modo que foi possível distinguir dois grupos principais.

Parameters	C1	C2	C3	C4	C5	C6	C7
N:	17	15	16	21	24	23	28
Ca ²⁺	7.9	25.0	48.5	42.0	35.5	78.0	75.0
Mg ²⁺	14.0	29.0	22.0	17.2	19.8	39.0	35.0
Na ⁺	210.0	450.0	6.5	59.0	40.5	57.0	13.4
K ⁺	14.10	15.00	1.25	4.50	4.80	10.00	4.15
HCO ₃ ⁻	460.0	321.3	226.9	<u>215.4</u>	274.6	346.4	332.8
Cl ⁻	48.0	510.0	<u>8.1</u>	35.0	21.5	73.0	15.0
SO ₄ ²⁻	9.0	22.0	18.5	34.0	15.0	58.0	49.0
Fe ²⁺	0.0190	0.2800	<u>0.0060</u>	0.0290	0.1550	0.1200	0.5400
Mn ²⁺	0.0010	0.0580	<u>0.0003</u>	0.0110	0.0490	0.0150	0.0885
Br ⁻	0.130	2.300	<u>0.010</u>	0.050	0.065	0.034	<u>0.010</u>
Sr ²⁺	0.656	1.100	<u>0.095</u>	2.000	0.330	1.260	0.336
F ⁻	0.99	0.48	<u>0.11</u>	1.00	0.31	0.34	0.16
Ba ²⁺	0.220	0.120	<u>0.075</u>	0.100	0.140	0.210	0.140
HS ⁻	0.20	0.04	<u>0.02</u>	<u>0.02</u>	<u>0.02</u>	<u>0.02</u>	<u>0.02</u>
TDS ^a	368	1131	276	310	239	509	383
Total hardness ^a	72	196	211	158	164	355	336
Casing depth (m)	21	34	10	14	27	15	9

N: Number of samples; bold values: highest values; underlined values: lowest values.
^a Calculated (total hardness as CaCO₃).

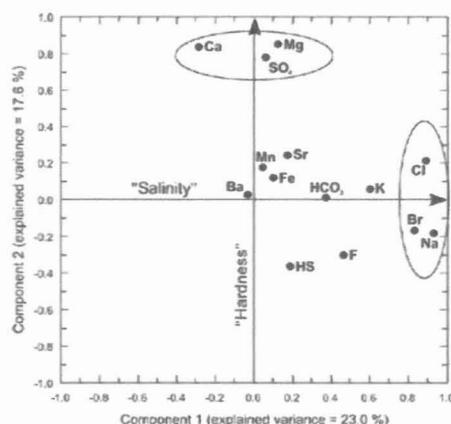


Figura 3 – Aplicação do método ACP (Cloutier *et al.* 2008).

4.6. Análise de Agrupamento Hierárquico (HCA)

A Análise de Agrupamento Hierárquico (*Hierarchical Clustering Analysis* – HCA) é uma importante ferramenta para análise de conjuntos de dados multivariados. Por meio desse método, as amostras são divididas em grupos distintos (*clusters*) e podem revelar características hidrogeológicas ou hidrogeoquímicas importantes (Guler *et al.*, 2002).

Diversos autores têm empregado este método para a determinação de grupos de amostras de água subterrânea com características semelhantes, como Yidana *et al.* (2010), Cloutier *et al.* (2008) e Marimon, *et al.* (2013).

A análise de *clusters* é realizada em duas etapas principais:

- 1) a escolha da medida de distância entre as observações,
- 2) a escolha do algoritmo utilizado para agrupar as amostras em *clusters*.

A análise parte de uma matriz de dados $n \times p$, onde n representa as observações (amostras) e p as variáveis (parâmetros analisados). A partir desta é calculada a matriz de distâncias D ($n \times n$) (Hardle e Simar, 2007).

De acordo com Bassab *et al.* (1990), nos casos em que existem parâmetros que possuem unidades de medida distintas, os dados devem ser padronizados antes de dar início à análise. A padronização é realizada por meio da Equação 1.

$$Z_i = \frac{X_i - \bar{X}_i}{S_i}$$

Equação 1

Onde, \bar{X}_i é a média dos valores e S_i representa o desvio padrão.

A medida da distância entre dois objetos em um espaço multivariado recebe o nome de coeficiente de parença. Este tipo de medida se divide em duas categorias distintas: medidas de similaridade (como o coeficiente de correlação) e medidas de dissimilaridade (por exemplo, a Distância Euclidiana).

A Distância Euclidiana é a medida mais conhecida para indicar proximidade entre dois objetos A e B quaisquer. Neste processo é calculada a Distância Euclidiana Média, a partir da Equação 1.

$$D(A,B) = \left[\sum_{i=1}^p \frac{(x_i(A) - x_i(B))^2}{p} \right]^{1/2}$$

Equação 2

onde x_i representa as coordenadas para o objeto e p o número de coordenadas envolvidas.

Após calcular a matriz de distâncias D , é empregado um algoritmo de agrupamento hierárquico. Este tipo de algoritmo parte da menor divisão possível entre as observações, onde cada amostra é considerada um *cluster*. As duas amostras mais próximas são agrupadas formando um novo *cluster*, as distâncias entre os novos grupos são recalculadas e esse passo se repete até que todas as amostras pertençam a um único grupo.

Nesta técnica é possível tratar as amostras como pontos em um espaço definido pelas variáveis (*R-mode*), ou o inverso, tratar as variáveis como pontos em um espaço definido pelas amostras (*Q-mode*).

O resultado final dessa análise é um dendograma, no qual duas amostras próximas sugerem forte correlação entre as mesmas. Em outras palavras, quanto maior a proximidade entre as medidas, maior sua similaridade (Hardle e Simar, 2007).

De acordo com Guler *et al.* (2002), a *Phenon Line* é um método subjetivo para interpretação do dendograma. Trata-se de uma linha que pode ser movida de forma a se obter um número maior ou menor de grupos de acordo com a necessidade.

A Figura 4 apresenta um exemplo de dendograma obtido por Panero e Silva (2008), onde as amostras foram divididas em dois grandes grupos, através da *phenon line* (linha pontilhada observada na figura). Um dos grupos foi interpretado como representante das águas bicarbonatadas e outro corresponde às águas não bicarbonatadas.

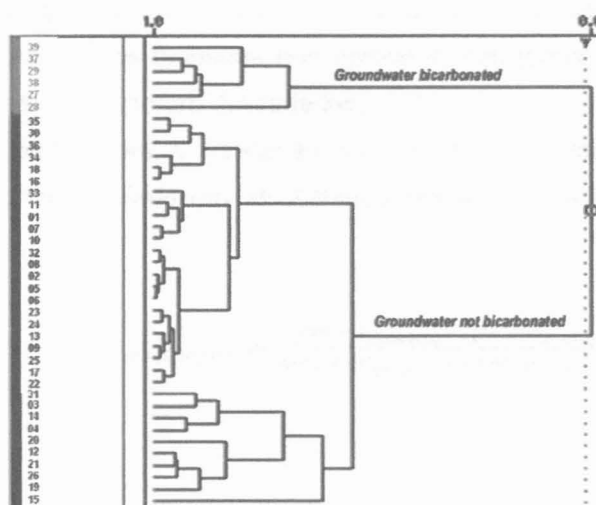


Figura 4 – Dendograma obtido por Panero e Silva (2008).

5. MATERIAIS E MÉTODOS

5.1. Tratamento dos dados e validação das análises

Para a realização do trabalho, foi cedido pelo Instituto de Pesquisas Tecnológicas (IPT) um conjunto de 251 análises químicas de água subterrânea provenientes de poços instalados na zona aflorante do SAG, no Estado de São Paulo. Parte delas é resultado da compilação de análises químicas realizadas pelo IPT junto às bases de dados do Sistema de Informação de Águas Subterrâneas – Siagas, do DAEE e de empresas de perfuração de poços. O restante das análises compreendem amostras coletadas e analisadas pelo próprio IPT.

O tratamento dos dados seguiu as seguintes etapas: avaliação dos dados; validação das análises químicas, tratamento das amostras em duplicata e identificação de *outliers*.

Foram selecionados para serem submetidos à análise estatística os compostos prioritários, expressos em mg/L:

- Cloreto (Cl^-), Sulfato (SO_4^{2-}), Bicarbonato (HCO_3^-), Nitrato (NO_3^-), Cálcio (Ca^{2+}), Magnésio (Mg^{2+}), Potássio (K^+), Sódio (Na^+) e Sílica (SiO_2).

E os parâmetros físico-químicos:

- pH, Condutividade Elétrica ($\mu\text{S}/\text{cm}$), Potencial de Óxido Redução (mV) e Sólidos Totais Dissolvidos - STD (mg/L).

Este último parâmetro foi calculado por meio da Equação 3:

$$\text{STD}_{(\text{mg/L})} = \Sigma [\text{íons}] + [\text{Sílica}]$$

Equação 3

A etapa de avaliação foi realizada a fim de averiguar se todas as análises químicas estavam completas. Apenas as amostras que apresentavam todos os parâmetros citados foram selecionadas, as demais foram descartadas.

Prosseguiu-se então com a validação das amostras restantes para análise de acurácia. Para tanto, foi calculado erro do balanço iônico para cada amostra, através da Equação 4:

$$\text{Erro}_{\text{Balanço Iônico}} = \frac{\Sigma[\text{cátions}]_{(\text{meq/L})} - \Sigma[\text{ânions}]_{(\text{meq/L})}}{\Sigma[\text{cátions}]_{(\text{meq/L})} + \Sigma[\text{ânions}]_{(\text{meq/L})}} \times 100$$

Equação 4

onde, $\sum C[\text{cátions}]$ representa a somatória das concentrações dos cátions e $\sum C[\text{ânions}]$ a somatória das concentrações dos ânions, ambas expressas na unidade de meq/L (miliequivalente por litro), que é calculada por meio da Equação 5:

$$C_{\text{meq/L}} = \frac{C_{\text{mg/L}} \times \text{Valência}}{\text{Massa Molar}}$$

Equação 5

Segundo Custódio e Llamas (1976), o erro admissível pode variar de acordo com o tipo de água e as concentrações dos compostos presentes na mesma, o que é refletido pelo parâmetro de Condutividade Elétrica. Dessa forma, águas cada vez menos salinizadas podem comportar erros de balanço iônico maiores (Tabela 1).

Tabela 1. Erro admissível (adaptado de Custódio e Llamas, 1976)

Condutividade Elétrica ($\mu\text{S/cm}$)	50	200	500	2000	>2000
Erro admissível (%)	35	10	8	4	4

As amostras consideradas válidas foram encaminhadas para a próxima etapa, onde foram identificadas as análises em duplicata, ou seja, a existência de análises provenientes do mesmo poço. Para impedir distorções na análise estatística devido ao fato de um mesmo ponto possuir mais de uma amostra, foi calculada a média dos resultados das análises químicas para cada poço onde foi identificada mais de uma amostra.

Por fim, as amostras restantes foram analisadas em busca de *outliers*, ou seja, valores discrepantes do conjunto de dados e que poderiam causar distorções no resultado final. Essa etapa foi realizada, pois, durante a execução da análise estatística multivariada notou-se que o resultado final apresentava mais influência de determinadas amostras do que do conjunto analisado, portanto foram removidas.

5.2. Análise Estatística Descritiva e Multivariada

A análise estatística descritiva foi realizada a partir dos dados selecionados e tratados na etapa anterior. Para tanto, foi utilizado o *software Excel 2010* (Microsoft, 2010) para os cálculos de média, mediana, máximo, mínimo, desvio padrão e demais cálculos que se fizeram necessários.

A análise estatística multivariada compreendeu a Análise de Componentes Principais (ACP), para os parâmetros físico-químicos e Análise de Agrupamento Hierárquico (HCA), para o conjunto de amostras selecionados na etapa anterior (*Q-mode*). Para tanto foi empregado o *software Statistica 13* (Dell Inc. 2015).

Uma vez que os parâmetros empregados na análise possuem unidades de medida diferentes entre si, os dados foram padronizados para que cada parâmetro possuísse média igual a 0 e desvio padrão igual a 1, conforme Hardle e Simar (2007).

Essa conversão foi realizada no próprio *software*, no *setup* para a realização das análises de ACP e previamente à realização da HCA.

As amostras consideradas válidas foram submetidas à HCA. Para essa etapa analítica foi empregada a Distância Euclidiana como método de medida de similaridade e o método de *Ward* como algoritmo de agrupamento, por serem os métodos que produzem grupos altamente homogêneos e mais distintos entre si. Esses métodos têm sido aplicados em diversos estudos hidrogeoquímicos (Menció *et al.*, 2008; Yidana *et al.*, 2010, Cloutier *et al.*, 2008) e, segundo, Guler *et al.* 2002, eles produzem os resultados mais satisfatórios quando se trata de análises físico-químicas de águas subterrâneas.

5.3. Elaboração de Mapas, Gráficos e Diagramas

A fim de se classificar quimicamente as águas subterrâneas da zona de afloramento do SAG, foram elaborados diagramas de Piper e de Stiff, com o auxílio do *software Aquachem* (Schlumberger, 2009).

Gráficos do tipo *box plot* foram gerados no *software Statistica 13* (Dell Inc. 2015) e os demais gráficos e tabelas foram gerados por meio do *software Excel 2010* (Microsoft, 2010).

Para a confecção do mapa de localização dos poços, foi empregado o *software ArcGis 10.4* (Esri, 2015)

6. RESULTADOS OBTIDOS

6.1. Tratamento dos dados e validação das análises

Durante a primeira etapa de tratamento dos dados, observou-se que, dentre as análises fornecidas pelo IPT, 149 amostras estavam incompletas para um ou mais dos compostos selecionados e por este motivo foram descartadas. As demais 102 amostras foram submetidas à etapa de validação das análises químicas.

Deste universo de amostras, 79 análises atenderam ao critério selecionado para o erro do balanço iônico e 23 foram descartadas por apresentarem erro superior ao estipulado.

A próxima etapa foi a de avaliação e tratamento de duplicatas e triplicatas. Observou-se que 18 pares e 2 trios de análises pertenciam cada qual aos mesmos poços coletados. Para cada ponto nesta condição foi obtida a média dos resultados das análises, de modo que ao final deste processo, para cada ponto de coleta havia apenas uma análise média, o que resultou em um total de 57 amostras.

Finalmente esses dados foram examinados em busca de *outliers*. Foram identificadas duas amostras nesta condição (#29 e #98). Na Tabela 2 é possível observar os resultados analíticos obtidos para essas amostras, comparados com as médias e os valores máximos obtidos para o restante das análises. Nota-se que para o Cl^- as concentrações das duas amostras foram cerca de 4 vezes maiores do que o valor máximo obtido para o restante das análises. Para o Na^+ , a amostra 29 apresenta valor quase 3 vezes maior do que o valor máximo obtido para as demais amostras. Assim, essas amostras também foram excluídas do banco de dados, o que culminou em um total de 55 amostras encaminhadas para análise estatística.

Tabela 2. Identificação de *outliers*

Ponto / Grupo	Cl^-	SO_4^{2-}	HCO_3^-	NO_3^-	Ca^{2+}	Mg^{2+}	K^+	Na^+	SiO_2
#29	21,00	0,00	0,00	17,27	0,81	0,41	0,68	17,00	4,50
#98	16,28	0,25	10,50	21,67	6,48	2,36	5,03	7,02	8,59
Máximo	4,50	1,80	42,00	21,26	11,00	3,40	4,90	5,90	20,00
Média	1,00	0,15	7,30	5,05	1,64	0,73	1,60	1,82	6,17

O Anexo A apresenta as análises físico-químicas para o banco de dados consolidado, constituído de 55 amostras.

6.2. Análise Estatística Descritiva e Multivariada

Os principais resultados obtidos na análise estatística descritiva, para os dados consolidados, estão apresentados na Tabela 3.

Tabela 3. Análise Estatística Descritiva

Parâmetro	Unidade	Média	Mediana	Mínimo	Máximo	D.P.
HCO ₃ ⁻	mg/L	8,41	4,50	< L.Q	42,00	10,62
NO ₃ ⁻		4,55	3,28	< L.Q	21,26	4,74
Cl ⁻		0,95	0,54	< L.Q	4,50	1,16
SO ₄ ²⁻		0,16	0,06	< L.Q	1,80	0,31
Na ⁺		1,85	1,89	< L.Q	5,90	1,23
Ca ²⁺		1,84	0,71	< L.Q	11,00	2,61
K ⁺		1,64	1,57	0,07	4,90	1,15
Mg ²⁺		0,75	0,49	0,02	3,39	0,73
SiO ₂		6,17	4,85	0,23	20,00	5,44
CE	μS/cm	30,98	18,70	2,3	121,05	28,64
pH	-	5,70	5,80	4,60	6,80	0,47
Eh	mV	466,11	468,50	309	645	77,69
STD Calculado	mg/L	26,32	19,46	5,78	79,65	19,93

L.Q. = Limite de quantificação.

D.P. = Desvio Padrão.

A quantidade de Componentes Principais (CPs) obtidas na ACP foi definida de acordo com o critério de Kaiser, segundo o qual, apenas as componentes que possuem autovalores maiores que 1 são aceitas (Kaiser, 1960).

Desta forma, foram geradas 4 CPs, que explicam 80,90% da variância total. A Tabela 4 apresenta os autovalores e variâncias obtidas para cada CP. Em negrito estão as CPs que apresentaram autovalores acima de 1.

Tabela 4. Componentes Principais

Componente	Autovalores	Variância Explicada (%)	Variância Cumulativa (%)
CP-1	6,08	46,74	46,74
CP-2	1,87	14,41	61,15
CP-3	1,53	11,80	72,95
CP-4	1,03	7,95	80,90
CP-5	0,63	4,82	85,71

Em cinza e negrito: Autovalores acima de 1 e a variância cumulativa representada pela soma das 4 CPs.

A Tabela 5 apresenta os carregamentos (*loadings*) de cada parâmetro nas Componentes Principais selecionadas. Os valores que indicaram alta correlação direta ou inversa foram destacados, ou seja, os valores acima de 0,7 ou abaixo de -0,7, respectivamente.

A CP-1, que explica a maior parte da variação do sistema, possui os principais valores de *loadings* para STD, Ca^{2+} , Mg^{2+} , CE, K^+ , HCO_3^- e SiO_2 . A CP-2 apresenta *loading* de -0,7 para Cl^- e de -0,69 para NO_3^- , sendo que este último também foi destacado devido à magnitude dos dados. Por fim, as CPs 3 e 4 apresentaram altos *loadings* para os compostos Na^+ e SO_4^{2-} , respectivamente. Destaca-se o fato de que as CP-2 e 3 revelam alta correlação inversa, uma vez que o sinal é negativo.

Tabela 5. *Loadings* dos parâmetros nas componentes principais

Parâmetro	CP-1	CP-2	CP-3	C-P4
STD	0,96	0,18	-0,11	0,02
Ca^{2+}	0,87	0,35	-0,08	0,04
Mg^{2+}	0,86	-0,27	-0,21	-0,05
CE	0,84	-0,03	-0,02	0,11
K^+	0,79	-0,07	-0,07	-0,22
HCO_3^-	0,76	0,62	-0,08	0,02
SiO_2	0,76	0,00	0,36	-0,01
Cl^-	0,46	-0,70	-0,28	-0,03
NO_3^-	0,49	-0,69	-0,34	0,02
Na^+	0,09	0,39	-0,80	0,11
SO_4^{2-}	0,36	0,05	0,24	0,82
pH	0,56	0,22	0,35	-0,52
Eh	-0,52	0,29	-0,57	-0,11

Em cinza e negrito: *Loadings* acima de 0,7 ou abaixo de -0,7.

O dendograma resultante da HCA (*Q-mode*) pode ser observado na Figura 5.

Nota-se claramente a existência de dois grupos distintos (Grupos A e B). A *Phenon Line* foi locada da forma que melhor se adequa ao dendograma, desta forma, foram obtidos 4 subgrupos (SGs), o Grupo A foi subdividido em SG-1 e SG-2 e Grupo B foi subdividido em SG-3 e SG-4.

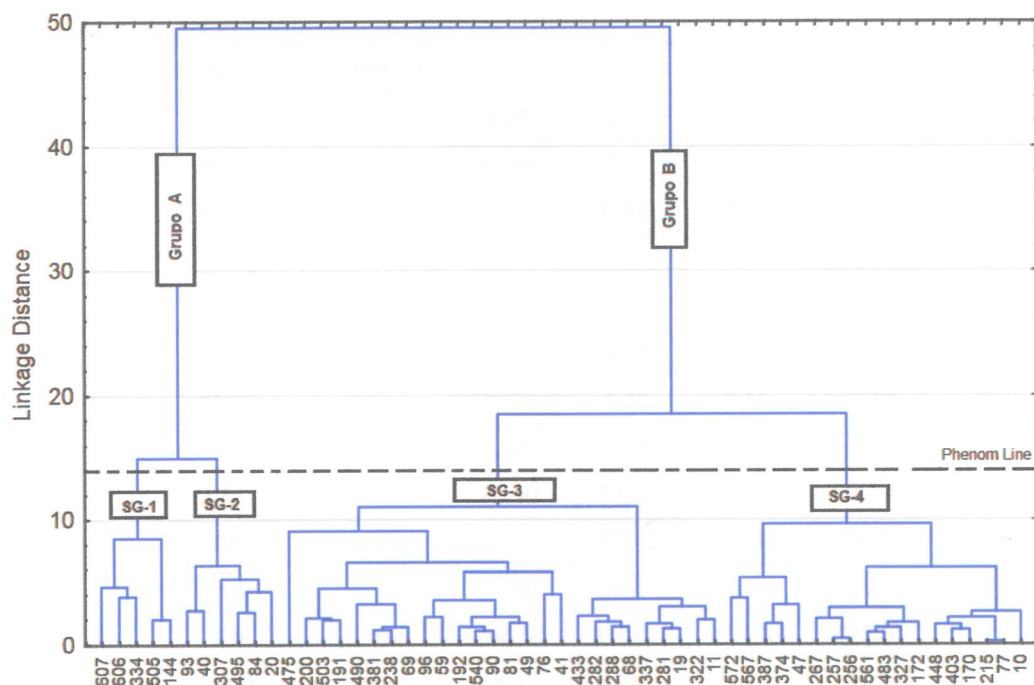


Figura 5. Análise de Agrupamento Hierárquico

Os valores médios para os parâmetros analisados, separados por grupos e subgrupos podem ser observados nas Tabelas 6 e 7 respectivamente. Na Tabela 7 os valores mais altos para cada parâmetro foram destacados em cinza e negrito e os valores mais baixos foram destacados em cinza e itálico.

Tabela 6. Médias dos parâmetros para os Grupos A e B

Parâmetro	Unidade	Grupo A	Grupo B
Cl ⁻	mg/L	2,29	0,67
NO ₃ ⁻		8,95	3,58
SO ₄ ²⁻		0,32	0,11
HCO ₃ ⁻		21,91	4,47
Ca ²⁺		5,89	0,74
Na ⁺		2,07	1,88
Mg ²⁺		1,97	0,44
K ⁺		3,21	1,16
SiO ₂		13,47	3,97
CE	µS/cm	69,38	20,09
pH	-	6,07	5,55
Eh	mV	397,42	489,23
STD	mg/L	60,07	17,02

Em cinza e negrito: maiores valores.

Tabela 7. Médias dos parâmetros divididos por subgrupos

Parâmetro	Unidade	SG-1	SG-2	SG-3	SG-4
Cl ⁻	mg/L	3,64	0,95	<i>0,51</i>	0,83
NO ₃ ⁻		13,62	4,27	<i>3,47</i>	3,68
SO ₄ ²⁻		0,21	0,43	0,15	<i>0,07</i>
HCO ₃ ⁻		9,15	34,67	5,98	2,96
Ca ²⁺		3,37	8,40	0,90	<i>0,58</i>
Na ⁺		1,68	2,47	<i>1,36</i>	2,41
Mg ²⁺		2,48	1,45	0,55	<i>0,33</i>
K ⁺		3,48	2,93	1,72	<i>0,59</i>
SiO ₂		10,60	16,33	5,66	2,29
CE	μS/cm	61,01	77,75	26,15	<i>14,03</i>
pH		5,98	6,17	5,86	<i>5,23</i>
Eh	mV	<i>384,50</i>	410,33	454,88	523,58
STD	mg/L	48,23	71,90	20,30	<i>13,74</i>

Em cinza escuro e negrito: maiores valores.

Em cinza claro e itálico: menores valores.

6.3. Elaboração de Mapas, Gráficos e Diagramas

Parte dos gráficos e diagramas gerados neste trabalho foi elaborada para auxiliar na interpretação dos resultados obtidos principalmente na Análise Estatística Multivariada, para evidenciar relações ou responder a eventuais questões surgidas na interpretação desses dados. Estes gráficos são apresentados no Item 7 – Interpretação dos Resultados.

O diagrama de Piper para o conjunto de análises validadas pode ser observado na Figura 6. As Figuras 7 e 8 apresentam os diagramas de Piper gerados para cada grupo obtido na etapa de *clustering*, de modo a facilitar a visualização e evitar a sobreposição de dados. Nestas duas figuras, os dados estão classificados de acordo com o subgrupo ao qual pertencem.

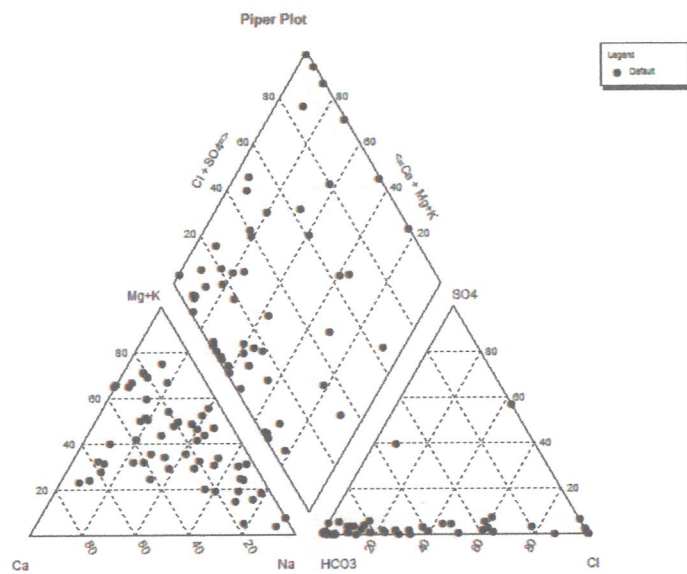


Figura 6 – Diagrama de Piper (para as 55 amostras)

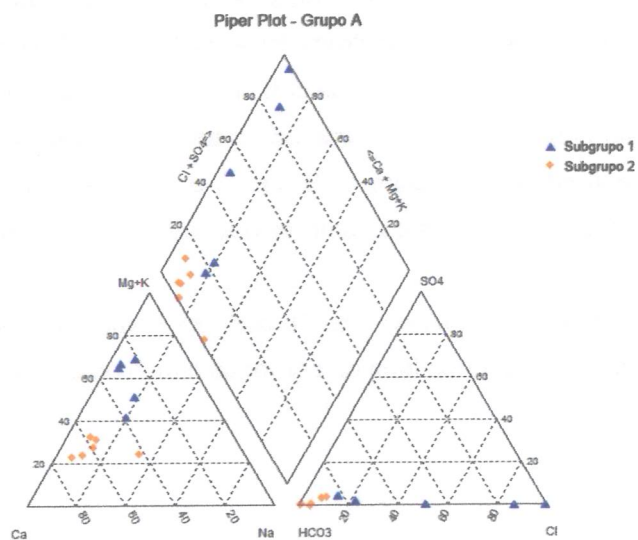


Figura 7 – Diagrama de Piper - Grupo A

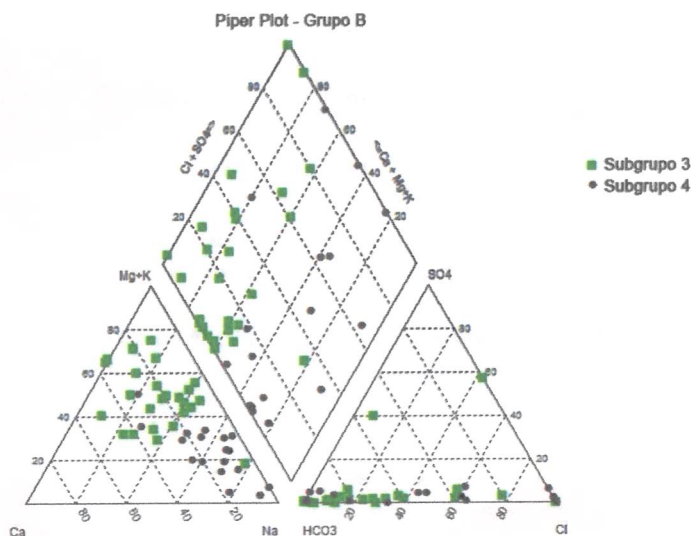


Figura 8 – Diagrama de Piper - Grupo B

Os gráficos do tipo *box plot* para todos os parâmetros, separados por subgrupos estão apresentados no Anexo B. Para facilitar a interpretação dos resultados, alguns gráficos deste tipo são reapresentados no decorrer do trabalho.

A partir das médias calculadas para os parâmetros, foram obtidos diagramas de Stiff para os 2 grupos principais (Grupo A e Grupo B) e para os 4 subgrupos (SG-1 a SG-4) delimitados na análise de *clustering*. Estes diagramas estão apresentados nas Figuras 9 a 14.

O mapa geológico do Estado de São Paulo, contendo a localização dos poços amostrados, separados de acordo com o subgrupo ao qual pertencem consta na Figura 15.

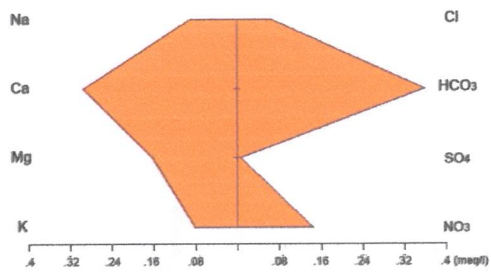


Figura 9 – Diagrama de Stiff- Grupo A

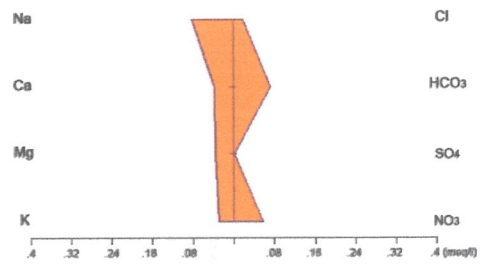


Figura 10 – Diagrama de Stiff- Grupo B

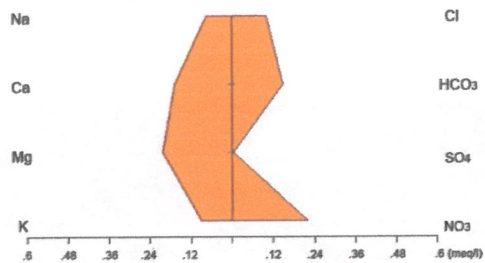


Figura 11 – Diagrama de Stiff- SG-1

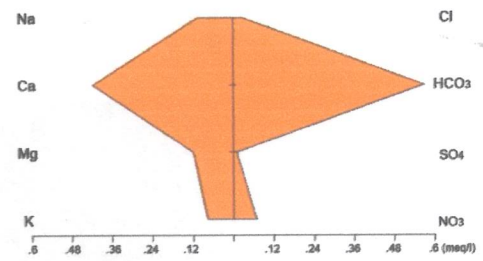


Figura 12 – Diagrama de Stiff- SG-2

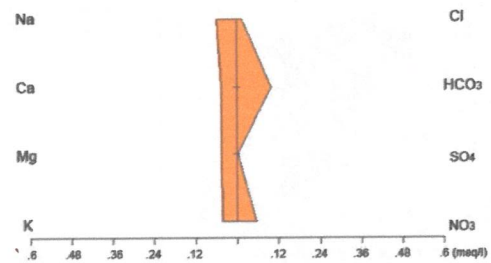


Figura 13– Diagrama de Stiff- SG-3

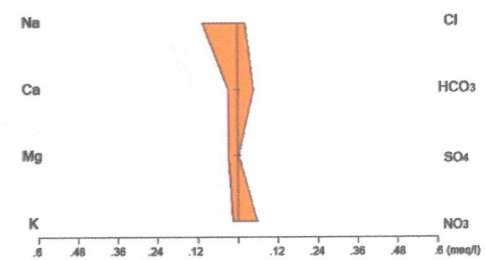


Figura 14 – Diagrama de Stiff- SG-4

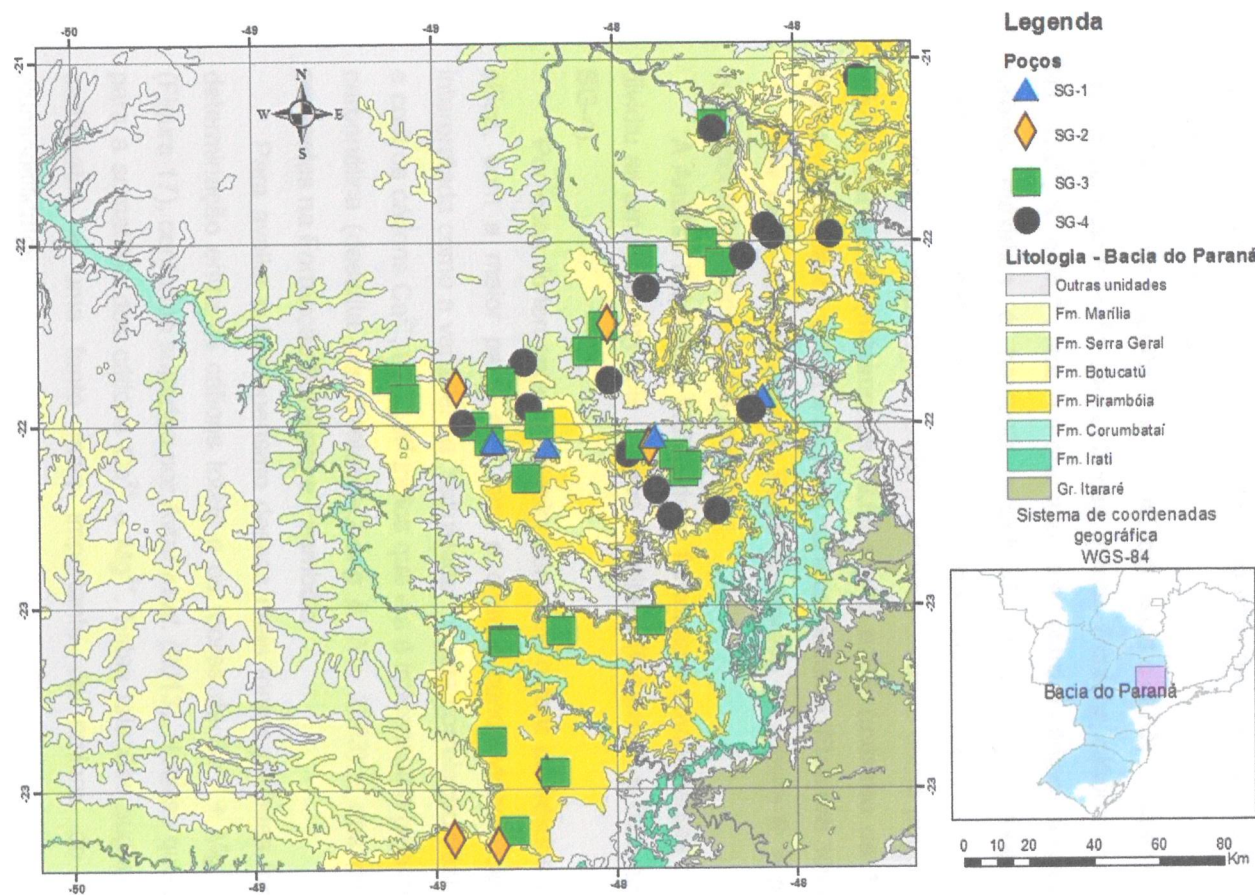


Figura 15 – Mapa de Localização dos poços de acordo com o subgrupo.

7. INTERPRETAÇÃO DOS RESULTADOS

Conforme se observa na Tabela 3, o HCO_3^- é o ânion que se destaca por apresentar os maiores teores e o maior valor médio para sua concentração. Logo em seguida destaca-se a SiO_2 como um dos compostos predominantes em termos de composição média. Comparando-se estes dois compostos nota-se que a SiO_2 possui uma distribuição muito mais homogênea que o HCO_3^- como sugerem os valores máximos e mínimos e o desvio padrão para as amostras.

No que diz respeito aos ânions, logo após o HCO_3^- , seguem-se o NO_3^- , Cl^- e SO_4^{2-} em ordem decrescente de concentração média. Para os cátions, essa ordenação resulta em Na^+ , Ca^{2+} , K^+ e Mg^{2+} . Nota-se que enquanto entre os ânions a diferença entre as médias e valores máximos é bastante acentuada, entre os cátions as concentrações são relativamente próximas, especialmente para o Na^+ e o Ca^{2+} .

Em geral estas águas podem ser consideradas como pouco salinas, dado os baixos valores para STD (máximo de 79,65 mg/L) e CE (máximo de 121,5 $\mu\text{S}/\text{cm}$), levemente ácidas (pH entre 4,6 e 6,8) e de forte caráter oxidante (Eh entre 309 e 645 mV).

Por meio do diagrama de Piper observado na Figura 6, é possível classificar estas águas como bicarbonatadas mistas.

A Análise Estatística Multivariada forneceu 4 componentes principais (CP-1 a CP-4) e dividiu as amostras em dois Grupos principais (Grupos A e B) e quatro Subgrupos (SG-1 a SG-4).

Por estar fortemente correlacionada com o teor de sólidos totais dissolvidos (STD), sílica, com a maior parte dos íons e com a condutividade elétrica, a CP-1 pode ser interpretada como a variável da "salinidade". A alta correlação desta CP com a SiO_2 , HCO_3^- e com os cátions Ca^{2+} , Mg^{2+} , K^+ indica que ela é influenciada pela dissolução da cimentação carbonática (descrita por Caetano-Chang & Wu, 2003) e dos minerais aluminosilicáticos presentes na Formação Botucatu e Pirambóia.

Para avaliar o intemperismo dessas duas fontes distintas foram gerados gráficos de determinação entre os cátions totais e os teores de $\text{Ca}^{2+} + \text{Mg}^{2+}$ (Figura 16) e $\text{Na}^+ + \text{K}^+$ (Figura 17), conforme sugerido por Avtar *et al.* (2013). Nota-se que a maior correlação se dá para a somatória dos cátions Ca^{2+} e Mg^{2+} . Estas relações sugerem que a dissolução dos carbonatos é uma fonte mais importante de cátions do que a dissolução dos aluminossilicatos.

Essa CP aparentemente está correlacionada com o Grupo A obtido no *Clustering*. Conforme se observa na Tabela 6, este grupo apresenta o maior teor de STD, maior CE e os maiores teores para todos os íons analisados. Essas características sugerem que este grupo provavelmente corresponde às águas com maior tempo de residência no aquífero e

esta observação é compatível com a conclusão de que a qualidade das águas subterrâneas deste grupo é resultado de interações entre água e rocha.

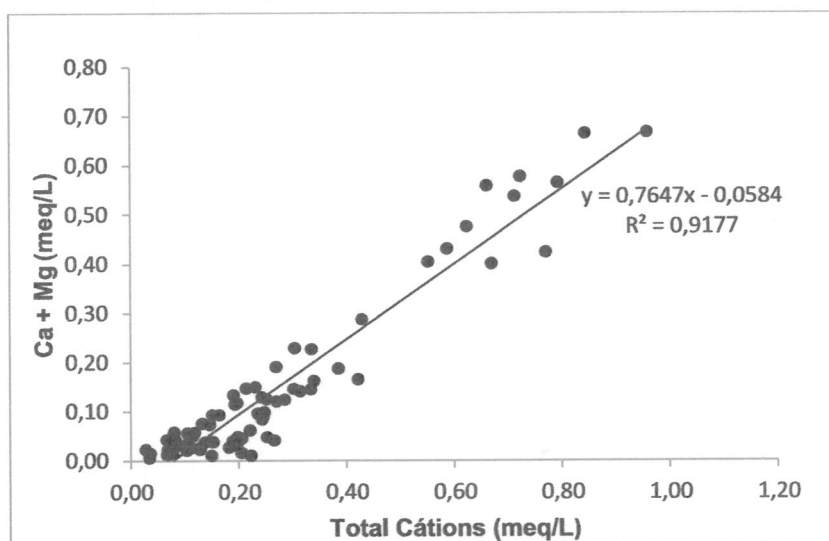


Figura 16. Ca+Mg x Cátions Totais

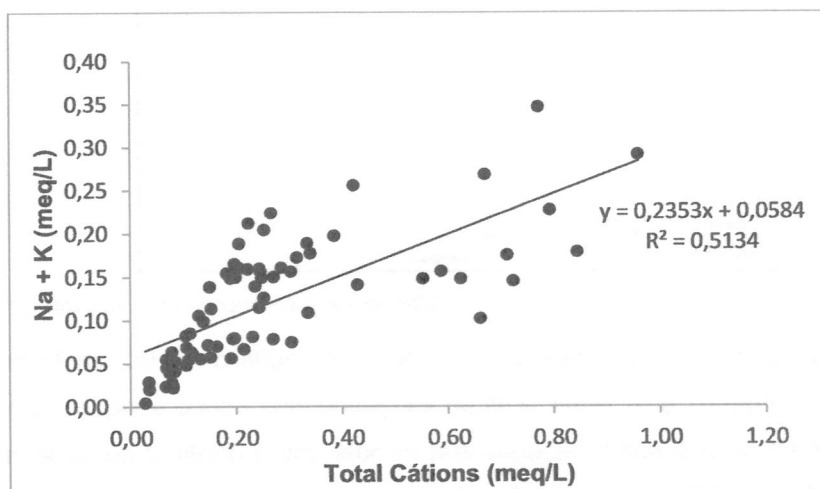


Figura 17. Na+K x Cátions Totais

A CP-2 apresenta forte correlação inversa com o Cl^- e o NO_3^- e foi interpretada como a variável que representa a atividade antrópica. Ela está evidentemente correlacionada com o SG-1, conforme nota-se nas Figuras 18 e 19. Como estes compostos apresentam sinal negativo de *loading* com a CP-2 e o SG-1 apresenta os maiores teores de Cl^- e NO_3^- , é possível afirmar que esta é a componente que indica a separação desse conjunto de poços, dos demais.

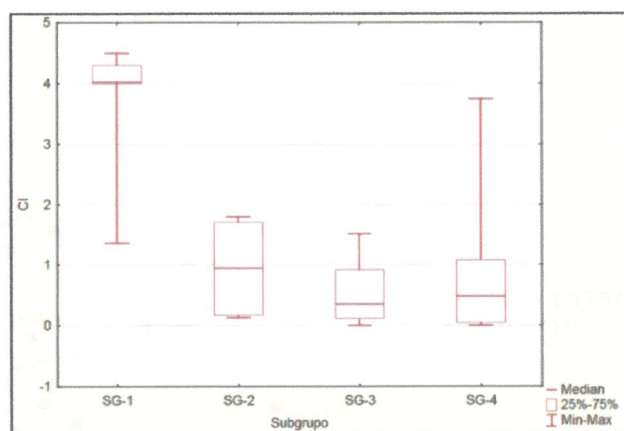


Figura 18. Box Plot para as concentrações de Cl^- .

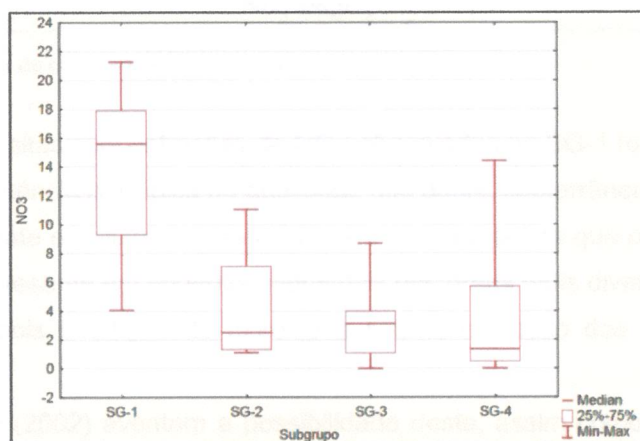


Figura 19. Box Plot para as concentrações de NO_3^- .

Segundo Fitts (2015), o efluente proveniente de sistemas sépticos para descarte de resíduos humanos, pode conter concentrações de NO_3^- entre 20 a 70 mg/L. Na maior parte dos sistemas sépticos o nitrato é um fator de preocupação. Esse tipo de sistema é comum em áreas rurais (como a maior parte da zona de afloramento do SAG no Estado de São Paulo), não atendidas por tubulações e sistemas de tratamento de esgotos.

Segundo Avtar *et al.* (2013), o NO_3^- proveniente do uso de fertilizantes deve apresentar forte correlação com os teores de K^+ , caso o emprego de fertilizantes seja a principal fonte de contaminação por nitrato. Para verificar essa possibilidade, foi traçado um gráfico de dispersão para esses parâmetros (Figura 20), obtendo-se um baixo coeficiente de determinação ($R^2=0,1791$). Portanto, apesar da zona de afloramento do SAG no Estado de São Paulo estar localizada em uma região de atividade agrícola predominante, a contaminação pelo uso de fertilizantes não é expressiva.

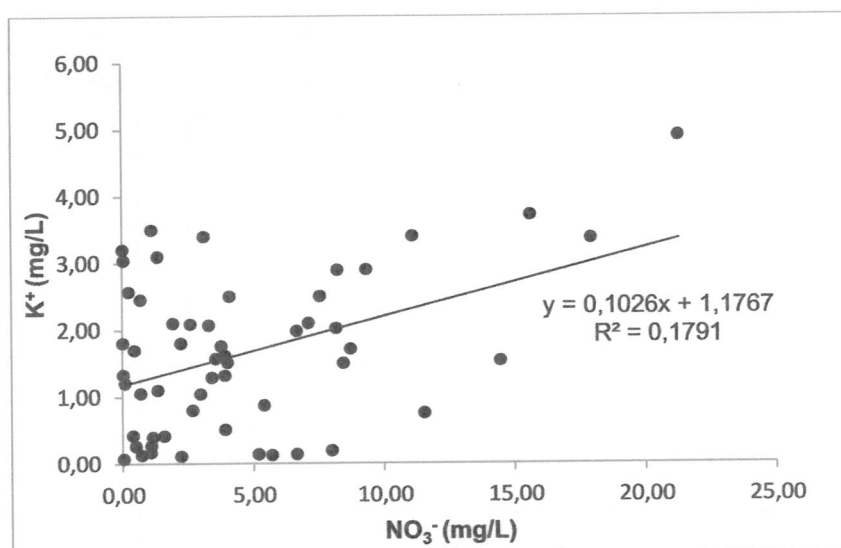


Figura 20. Diagrama de dispersão entre NO_3^- e K^+ .

Desta forma, as altas concentrações de NO_3^- observadas no SG-1 foram interpretadas como resultado da influência antrópica na qualidade das águas subterrâneas locais.

O Cl^- está presente em efluentes sanitários em concentrações que podem ultrapassar 15 mg/L. Pode estar presente em efluentes industriais das áreas mais diversas e até mesmo em águas tratadas, pois a adição de cloro leva a uma elevação dos níveis de cloreto (CETESB, 2009).

Sracek & Hirata (2002) aventam a possibilidade deste, assim como o sulfato, serem provenientes da dissolução de minerais evaporíticos presentes nas formações pouco permeáveis subjacentes à Formação Pirambóia, e estarem presentes no SAG por meio de mistura de águas.

Geisicki (2007), no entanto, cita que a infiltração de águas proveniente das formações subjacentes ao SAG é restrita à região oeste do Estado, em condições de grande confinamento do aquífero, o que não é o caso para este estudo, e também destaca que minerais evaporíticos, tais quais, halita, gipsita e mirabilita, não foram descritos nos arenitos das formações Pirambóia e Botucatu.

Durante o processo de tratamento dos dados, foram removidas duas amostras consideradas *outliers* (Tabela 2). Estas amostras apresentavam teores de Cl^- muito acima do restante das amostras. Isso pode se dar por contaminação das amostras após a coleta, mas a hipótese de serem amostras provenientes de água subterrânea contaminada por atividade antrópica não deve ser descartada.

A CP-3 destaca-se pela forte correlação inversa com o parâmetro Na^+ . O sódio, portanto, é o único cátion que não está correlacionado com a CP-1, o que sugere que este composto sofre processos distintos dos demais. De fato, comparando-se o *box plot* obtido

para o Na^+ (Figura 21) com os demais compostos, nota-se que ele não possui forte relação com o Grupo 1, estando distribuído de forma bastante homogênea entre os 4 subgrupos.

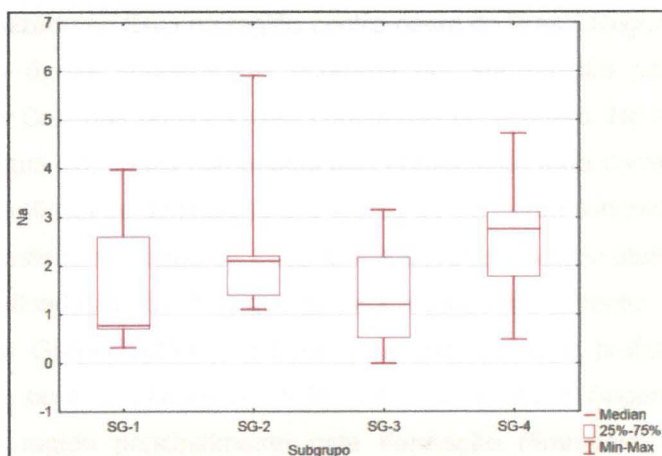


Figura 21. Box Plot para o composto Na^+ .

Uma hipótese para a obtenção de sódio seria a dissolução de feldspatos, porém, como revela a Figura 22, este cátion apresenta correlação muito baixa com a sílica ($R^2=0,0414$), portanto, a dissolução de feldspatos não poderia ser o principal mecanismo de obtenção de Na^+ .

Sracek e Hirata (2002), por meio de cálculos de balanço de massa afirmam que o incremento no teor de sódio não pode ser atribuído apenas a uma fonte interna de íons, como a dissolução de aluminossilicatos, e que uma fonte externa, como a infiltração de águas provenientes do Grupo Passa Dois, pode contribuir com o aumento na concentração do mesmo. Meng e Maynard (2001) afirmam que a fonte de Na^+ nas regiões mais profundas do aquífero provavelmente seja devido à infiltração de águas mais salinas provenientes das formações inferiores ao SAG. Cabe ressaltar novamente que este fenômeno estaria restrito às zonas de alto confinamento.

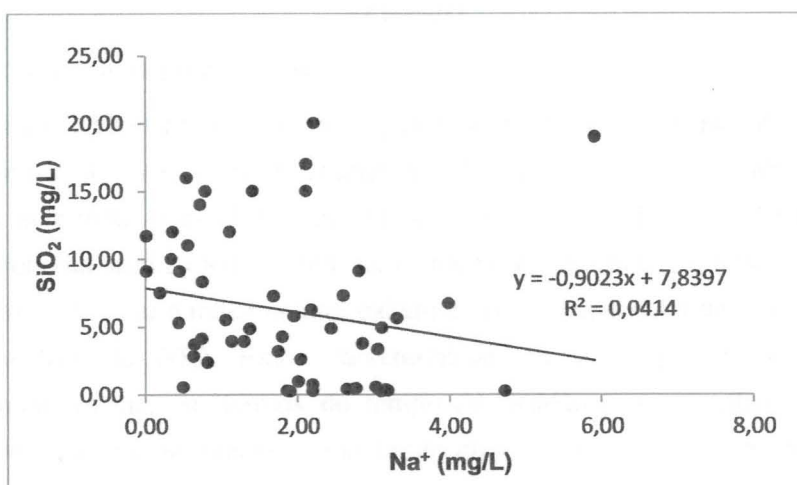


Figura 22. Diagrama de dispersão entre SiO_2 e Na^+ .

A ocorrência de troca catiônica entre o Na^+ e o Ca^{2+} ao longo das linhas de fluxo é uma hipótese já levantada por Sracek e Hirata (2002) e também observada por Gastmans (2007) em estudo realizado no SAG na região centro-oeste do Brasil. Segundo os autores, a perda de cálcio nas águas subterrâneas ocorreria por substituição pelo sódio, o que justificaria a perda de Ca^{2+} nas porções mais confinadas no aquífero. No entanto, como se pode observar na Figura 23, esses compostos não apresentam forte correlação, e nem ao menos observa-se tendência de diminuição dos teores de Ca^{2+} com o incremento de Na^+ , ou vice versa. Portanto este mecanismo de troca iônica deve apresentar atuação mais efetiva nas porções mais confinadas do SAG do que em sua zona de afloramento.

De acordo com Geisicki (2007), a troca catiônica acontece preferencialmente por adsorção à superfície, ou nos espaços interfoliares dos agilominerais. Segundo a autora este processo químico é regido principalmente pela Formação Pirambóia, devido às suas características faciológicas, com estratos argilosos característicos de formações interdunares e infiltração de argilas.

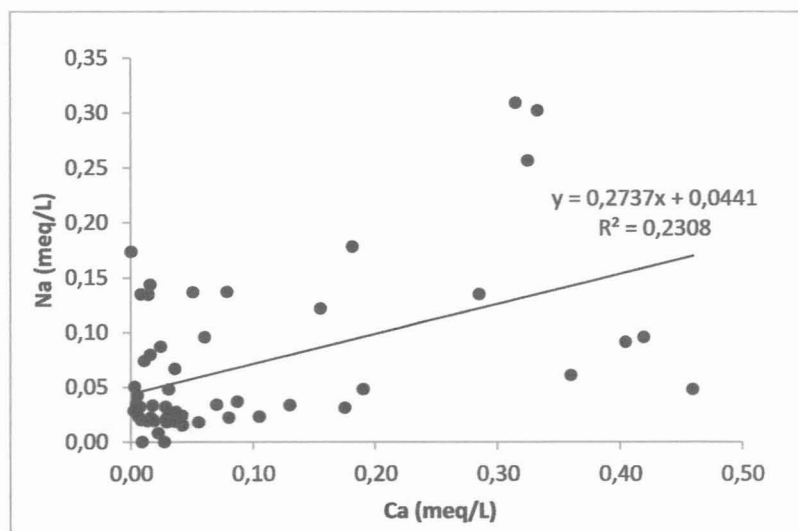


Figura 23. Diagrama de dispersão entre Na^+ e Ca^{2+} .

Apesar desta distribuição relativamente igualitária observada na Figura 21, nota-se por meio dos diagramas de Piper para os Grupos A e B (Figuras 7 e 8) que, relativamente, o SG-4 é o que apresenta o caráter mais sódico dentre os 4 subgrupos. Este subgrupo apresenta os menores teores médios para quase todos os parâmetros químicos (exceto Cl^- e NO_3^-), além de apresentar o maior caráter oxidante, menor condutividade elétrica, pH mais ácido e menor teor de STD. Estas características sugerem que estas águas são relativamente mais jovens, em termos de tempo de residência no aquífero, do que as demais. Uma vez que não se observa uma significativa variação nos teores de Na^+ , nem uma evolução destes teores do SG-4 para os demais subgrupos, é possível sugerir que o incremento de sódio nas águas da zona de recarga do SAG ocorra a nível bastante

superficial, seja pela precipitação atmosférica ou influência das águas superficiais, ou até mesmo atividade antrópica.

A existência de dissolução de feldspatos e a captura de sódio por meio de troca iônica com o cálcio são fenômenos que não podem ser descartados, e um equilíbrio entre estes dois fenômenos poderia ser a causa da heterogeneidade na distribuição dos teores de Na^+ ao longo de todos os subgrupos.

A CP-4, por fim, está correlacionada com o SO_4^{2-} . Nota-se por meio da Tabela 3 que o SO_4^{2-} apresenta os menores teores dentre os compostos analisados, evidenciando o caráter pouco sulfatado destas águas. A análise do *box plot* (Figura 24) para este parâmetro mostra que a maior parte das concentrações de SO_4^{2-} está abaixo de 1,2 mg/L, à exceção da amostra 475, que apresenta teor de sulfato de 1,8 mg/L (conforme consta no Anexo B). Desconsiderando-se esta amostra, o Grupo A concentra os maiores teores de sulfato.

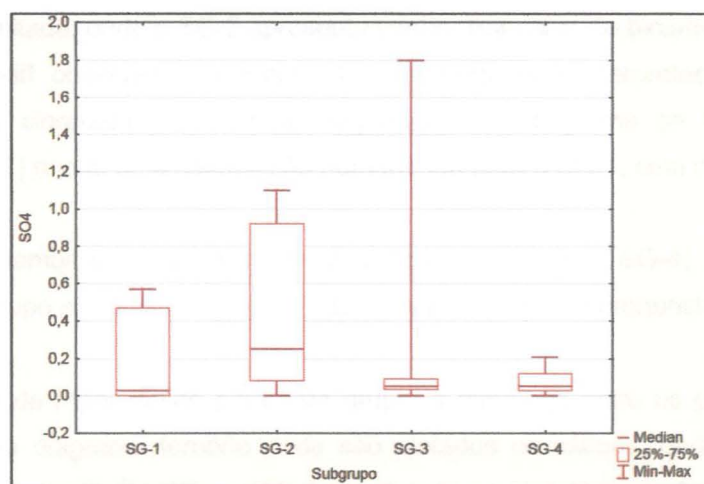


Figura 24. *Box plot* para o parâmetro SO_4^{2-} .

Segundo Custódio e Llamas (1976), o sulfato pode ser proveniente de atividades urbanas, industriais e agrícolas, ou de fontes naturais como a água da chuva, dissolução de minerais presentes em rochas ígneas, sedimentares e metamórficas, ou em terrenos formados em condições de grande aridez ou ambiente marinho. No entanto, qualquer que seja a fonte de sulfato, esta possui pouca influência na geoquímica das águas do SAG em sua zona de afloramento.

A análise do *Clustering* (Figura 5) permitiu dividir os poços em 2 grupos principais. As principais características desses conjuntos podem ser avaliadas por meio da interpretação das Tabelas 6 e 7, os *box plots* apresentados no Anexo B, dos diagramas de Piper observados nas Figuras 6 e 7 e dos diagramas de Stiff que constam nas Figuras 9 a 14.

O Grupo 1 apresenta os maiores teores para todos os compostos analisados, maior condutividade elétrica, maior quantidade de sólidos totais dissolvidos, além de pH mais neutro e caráter menos oxidante.

O Grupo 2 por sua vez possui os menores teores para os elementos químicos, pH levemente mais ácido e oxidante.

A diferença entre os dois grupos fica evidente ao se observar os *box plots* (Anexo B) e os diagramas de Stiff das Figuras 9 e 10, em que é possível verificar que o Grupo A fornece um diagrama mais robusto que o Grupo B, indicando maiores concentrações. Nota-se também que para os ânions ambos os grupos apresentam configuração semelhante, porém para os cátions o Grupo A apresenta caráter mais cálcico e o Grupo B caráter mais sódico.

O Grupo A foi dividido em 2 subgrupos, SG-1 (que apresenta os maiores teores para Cl^- , NO_3^- , K^+ e Mg^{2+}) e SG-2 (que apresentam os maiores teores para HCO_3^- , SO_4^{2-} , Ca^{2+} e Na^+).

A separação entre estes dois subgrupos é perceptível ao se analisar os *Box Plots*, especificamente os parâmetros Cl^- , NO_3^- , e Mg^{2+} para o SG-1 e HCO_3^- , Ca^{2+} e SiO_2 para o SG-2. No diagrama de Piper apresentado na Figura 7, a separação entre estes dois grupos pode ser bem delimitada, onde o SG-2 apresenta caráter francamente bicarbonatado cálcico (o diagrama de Stiff observado na Figura 12 corrobora esta interpretação) e o SG-1 classifica-se como cloretado potássico a magnésiano. O diagrama de Stiff para este subgrupo (Figura 11) revela uma distribuição homogênea para o SG-1, com destaque para o Mg^{2+} e o NO_3^- .

O Grupo B também foi dividido em 2 subgrupos (SG-3 e SG-4), no entanto, a separação deste grupo em dois subgrupos, não se revela de forma pronunciada ao analisar os *box plots*.

No diagrama de Piper obtido para este grupo, a diferença entre os grupos se torna mais perceptível no diagrama ternário onde são plotados os cátions, onde o SG-3 está deslocado no sentido do Mg^{2+} e K^+ e o SG-4 em direção ao Na^+ . Para os ânions, ambos os subgrupos se sobrepõem. Por meio dos diagramas de Stiff (Figuras 13 e 14) observa-se que ambos os subgrupos apresentam baixas concentrações, onde se destaca o HCO_3^- no SG-3 e, novamente, o Na^+ no SG-4.

8. CONCLUSÕES

A aplicação das técnicas de Análise Estatística Multivariada empregadas (HCA e PCA) se mostrou satisfatória para identificar conjuntos de poços com características hidrogeoquímicas semelhantes, bem como identificar parâmetros físico-químicos que atuam de modo semelhante sobre a qualidade das águas subterrâneas da Zona de Afloramento do SAG no Estado de São Paulo. Desta forma pode-se considerar que os objetivos propostos para o presente trabalho foram atingidos.

As águas subterrâneas presentes na porção livre do SAG no Estado de São Paulo foram divididas por meio da HCA em dois grupos principais de acordo com a salinidade. Esta divisão foi corroborada pelo resultado da PCA, onde a CP-1 está fortemente correlacionada com o teor de STD, CE e a maior parte dos íons analisados (Ca^{2+} , Mg^{2+} , K^+ , HCO_3^- e SiO_2).

O Grupo A caracteriza-se pela alta salinidade, maiores valores de CE, pH e caráter comparativamente menos oxidante. Foi dividido em dois subgrupos SG-1 e SG-2. O primeiro apresenta altos teores para NO_3^- e Cl^- , claramente correlacionado com a CP-2, foi interpretado como sido influenciado pela atividade antrópica. O segundo subgrupo apresentou altos teores para HCO_3^- e Ca^{2+} , interpretados como consequência da dissolução de carbonatos presentes nas formações Pirambóia e Botucatu.

O Grupo B, corresponde às águas de baixa salinidade, também foi dividido em dois subgrupos SG-3 e SG-4, sendo que o primeiro apresenta distribuição relativamente homogênea para todos os cátions, o segundo apresenta evidente característica sódica.

O sódio foi o único cátion analisado que não apresentou forte correlação com a CP-1, mas sim com a CP-3. Isso revela que este íon apresenta distribuição diferente dos demais, o que é confirmado pela análise do *box plot* para este parâmetro, e provavelmente está submetido a processos distintos, dentre os quais foram levantados a troca catiônica entre Ca^{2+} e Na^+ e dissolução de feldspatos. Destaca-se também, que o SG-4, que aparentemente se trata das águas mais jovens dentre os subgrupos, foi classificado como o mais sódico dentre os mesmos, por meio do Diagrama de Piper. O que indica que o incremento de sódio é oriundo de processos anteriores à infiltração da água no aquífero e sua concentração se mantém constante.

Por fim, a CP-4 está correlacionada com os teores de sulfato, porém uma análise dos dados levantados revela que estas águas apresentam em geral baixos teores para este composto.

9. REFERÊNCIAS BIBLIOGRÁFICAS

- ANDRIOTTI, J. L. S. Análise de Componentes Principais: Fundamentos de uma técnica de análise de dados multivariada aplicável a dados geológicos. *Acta Geológica Leopoldensia*, v. 20, n. 44, p. 27 - 50.1997.
- ARAUJO, L.M.; FRANÇA, A.B.; POTER, P.E. *Aquífero Gigante do Mercosul no Brasil, Argentina, Uruguai e Paraguai; Mapas Hidrogeológicos das Formações Botucatu, Pirambóia, Rosário do Sul, Buena Vista, Misiones e Tacaurémbo*. Mapas, Curitiba: UFPR-Petrobras, 1995. 4 mapas, colorido. Escala 1:5.000.000. Texto Explicativo, 16p.
- ASSINE, M. L.; PIRANHA, J.M.; CARNEIRO, C. D. R. Os Paleodesertos Pirambóia e Botucatu. In: NETO, V. M.; BARTORELLI, A.; CARNEIRO, C.D.R.; NEVES, B. B.de B. *Geologia do Continente Sul-Americano: Evolução da Obra de Fernando Flávio Marques de Almeida*. São Paulo: Beca, 2004. Cap. V.
- AVTAR, R.; KUMAR, P.; SINGH, C. K.; SAHU, N.; VERMA, R.L.; THAKUR, J. K.; MUKHERJEE, S. Hydrogeochemical Assessment of Groundwater Quality of Bundelkhand, India Using Statistical Approach . *Water Qual Expo Health*. p. 105 -115. 2013.
- BARRETOS, E. A. G. B. *Balanço hídrico em zona de afloramento do Sistema Aquífero Guarani a partir de monitoramento hidrogeológico em bacia representativa*. 2006. 249 p. Dissertação de mestrado - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos-SP.
- BASSAB, W.O.; MIAZAKI, E.S.; ANDRADE, D.F. Introdução à Análise de Agrupamentos. 9º *Simpósio Brasileiro de Probabilidade e Estatística*. São Paulo, Brasil. 1990.
- BRIGHETTI, J. M. P. *Faciologia dos sedimentos da Formação Pirambóia na Região de Rio Claro (SP)*. 1994. 124 p. Dissertação de mestrado – Instituto de Geociências, Universidade Estadual Paulista, Rio Claro-SP.
- CAETANO-CHANG, M. R.; WU, F. T. Diagênese de arenitos da Formação Pirambóia no Centro-Leste Paulista. *Geociências – Ed. Especial*, v. 22, p.33 – 39, 2003.
- CETESB COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO. *Qualidade das Águas Interiores do Estado de São Paulo*. Série Relatórios. 2009.
- CLOUTIER, V.; LEFEBVRE, R. THERRIEN, R.; SAVARD, M. M. Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *Journal of Hydrology*. n. 353, p 294 – 313. 2008.
- CUSTODIO, E.; LLAMAS, M. R. *Hidrología Subterránea*. Primeira Edição. Barcelona; Ediciones Omega, S.A. 1976.

- DAVIS, J.C. (1973). *Statistics and Data Analysis in Geology*. Segunda edição. John Wiley & Sons, Inc.
- DEPARTAMENTO DE ÁGUAS E ENERGIA ELÉTRICA DO ESTADO DE SÃO PAULO. *Estudo de Águas Subterrâneas – Região Administrativa 6 Ribeirão Preto*. Revista DAE, São Paulo, v. 2, p. 37 – 87, 1974.
- DIAS, C.L.; SILVA, M. de F. B.; CASARINI, D. C. P.; LEMOS, M. M. G. Caracterização hidrogeoquímica e qualidade do Sistema Aquífero Guarani em poços tubulares para abastecimento público no Estado de São Paulo. *Suplemento -XII Congresso Brasileiro de Águas Subterrâneas*, São Paulo, 2002.
- DOMENICO, P.A.; SCHWARTS, F.W.; (1998) *Physical and Chemical Hydrogeology*. Segunda Edição. John Wiley & Sons, Inc.
- FITTS, C. R. (2015). *Águas Subterrâneas*. Segunda edição. Elsevier Editora Ltda.
- GASTMANS, D. *Hidrogeologia e Hidroquímica do Sistema Aquífero Guarani na Porção Ocidental da Bacia Sedimentar do Paraná*. 2007. 238 p. Tese de Doutorado – Instituto de Geociências e Ciências Exatas. Universidade Estadual Paulista, Rio Claro – SP.
- GEISICKI, A. L. *Evolução diagenética das formações Pirambóia e Botucatu (Sistema Aquífero Guarani) no Estado de São Paulo*. 2007. 175 p. Tese de Doutorado – Instituto de Geociências. Universidade de São Paulo, São Paulo – SP.
- GOMES, M. A. F.; FILIZOLA, H.F.; SPADOTTO, C.A. Classificação das áreas de recarga do Sistema Aquífero Guarani no Brasil em domínios pedomorfoagroclicmáticos – subsídio aos estudos de avaliação de risco de contaminação das águas subterrâneas. *Revista do Departamento de Geografia*, São Paulo, n. 18, p. 67 – 74, 2006.
- GULER, C.; THYNE, G. D.; MCCRAY, J. E.; TURNER, A. K. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeology Journal*, p. 455 – 474, 2002.
- HARDLE, W., SIMAR, L. (2007). *Applied multivariate Statistical Analysis*. Springer.
- IRITANI, M.A., EZAKI, S. *As Águas Subterrâneas do Estado de São Paulo*, São Paulo. Secretaria do Estado do Meio Ambiente. 2009.
- KAISER, H.F. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20. p. 141- 151. 1960
- KIM, K. H.; YUN, S. T.; MAYER, B.; LEE, J. H.; KIM, T. S.; KIM, H. K. Quantification of nitrate sources in groundwater using hydrochemical and dual isotopic data combined with a Bayesian mixing model. *Agriculture, Ecosystems and Environment*. N. 199, p. 369 – 381. 2015

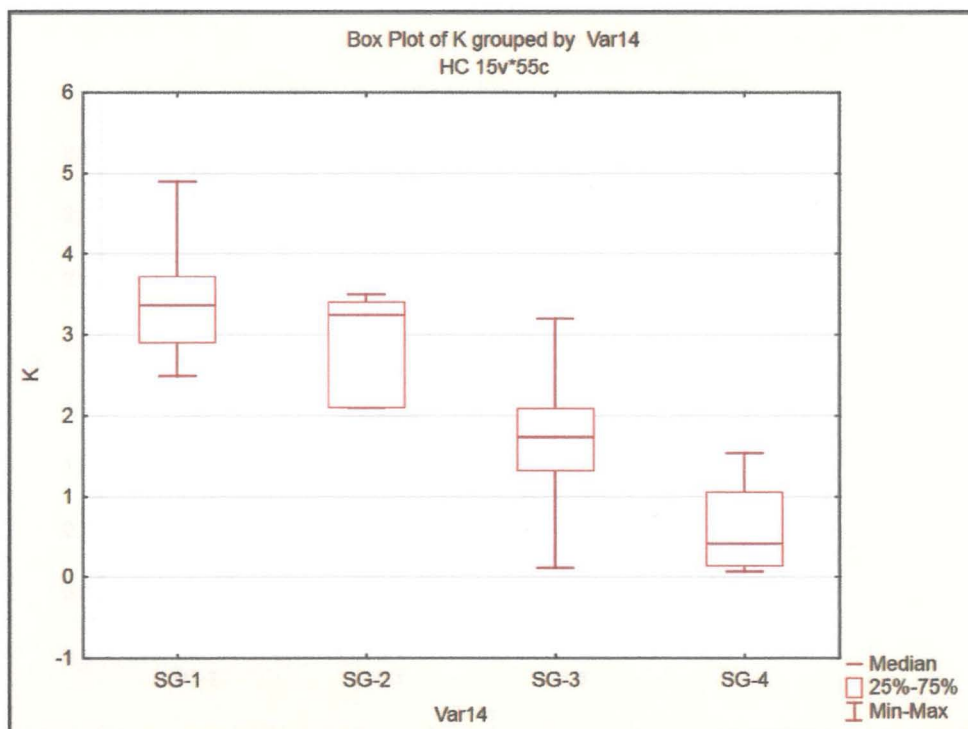
- MARIMON, M.P.C.; ROISENBERG, A.; SUHOGUSOFF, A.V.; VIERO, A.P.; Hydrogeochemistry and statistical analysis applied to understand fluoride provenance in the Guarani Aquifer System, Southern Brazil. *Environ Geochem Health*. P 391 – 403, 2013.
- MARQUES, L.S.; ERNESTO, M. O magmatismo toleítico da Bacia do Paraná. In: NETO, V. M.; BARTORELLI, A.; CARNEIRO, C.D.R.; NEVES, B. B.de B, *Geologia do Continente Sul-Americano: Evolução da Obra de Fernando Flávio Marques de Almeida*. São Paulo: Beca, 2004. Cap. XV.
- MENCIÓ, A.; MAS-PLA, J. Assessment by multivariate analysis of groundwater–surface water interactions in urbanized Mediterranean streams. *Journal of Hydrology*. n. 352, p. 355 – 366. 2008.
- MENG, S. X.; MAYNARD, J. B. Use of statistical analysis to formulate conceptual models of geochemical behavior: water chemical data from Botucatu aquifer in São Paulo state, Brazil. *Journal of Hydrology*, n.250. p 78 – 97, 2001
- OLIVEIRA, L.A.; VIEIRA, A.S. Estado da Arte do Sistema Aquífero Guarani – SAG. *Caminhos de Geografia*. v.11, n.34, p174 -189. 2010
- ORGANIZAÇÃO DOS ESTADOS AMERICANOS (OEA). *Aquífero Guarani: Programa Estratégico de Ação*. 2009.
- PANERO, F.S.; SILVA, H.E.B.; Application of exploratory data analysis for the Characterization of tubular wells of the North of Brazil. *Microchemical Journal*. N. 88, v. 194 – 200. 2008.
- PEATE, D. W.; HAWKESWORTH, C. J.; MANTOVANI, M. S. M.; Chemical stratigraphy of the Paraná lavas (South America): classification of magma types and their spatial distribution. *Bulletin of Vulcanology*, 55: 119 -139, 1992.
- ROCHA, G.A. O Grande Manancial do Cone Sul. USP, Estudos Avançados, São Paulo, v.11 n. 30, p. 191 -212, 1997.
- SPILLER, L. E.; Estruturas armazenadoras e qualidade das águas subterrâneas do Estado de São Paulo. *II Simpósio de Geologia do Sudeste*. São Paulo, Brasil. 2005.
- SRACEK O.; HIRATA, R. Geochemical and stable isotopic evolution of the Guarani Aquifer System in the State of São Paulo, Brazil. *Hydrogeology Journal*, p. 643 – 655. 2002.
- YIDANA, S. M.; YAKUBO, B. B.; AKABZAA, T. M. Analysis of groundwater quality using Multivariate and spatial analyses in the Keta basin, Ghana. *Journal of African Earth Sciences*. n. 58, p. 220 – 234. 2010.
- ZAKHEM, B. A.; HAFEZ, R. Hydrochemical, isotopic and statistical characteristics of groundwater nitrate pollution in Damascus Oasis (Syria). *Environ Earth Sci*. n. 74, p. 2781 -2797. 2015.

**ANEXO A – DADOS CONSOLIDADOS SUBMETIDOS À ANÁLISE
ESTATÍSTICA MULTIVARIADA**

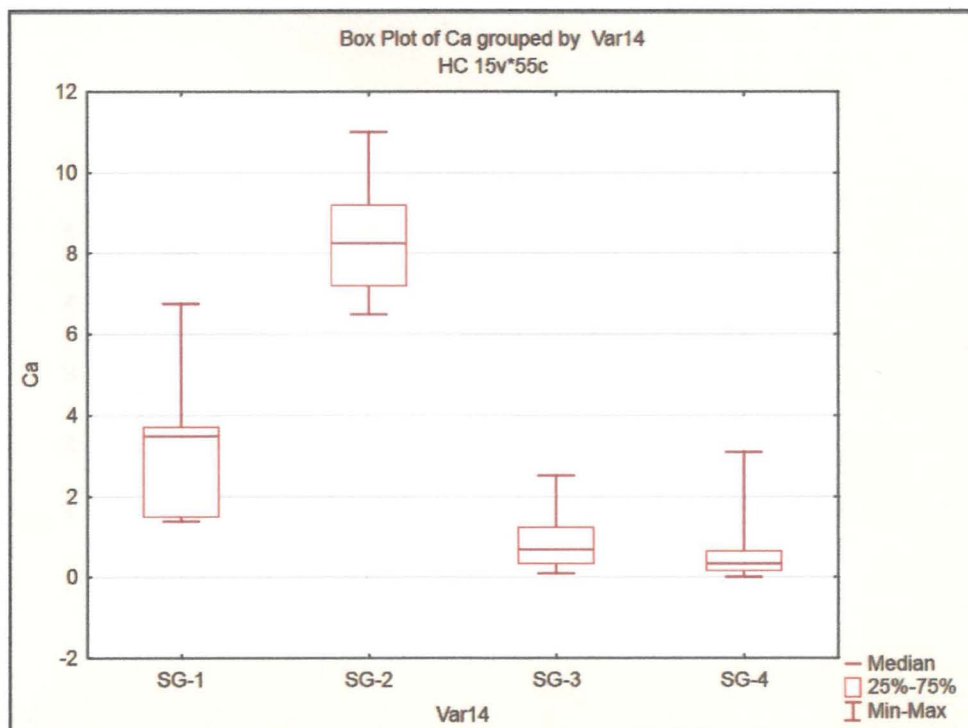
Poço Amostrado	Cl ⁻	SO ₄ ⁻	HCO ₃ ⁻	NO ₃ ⁻	Ca ²⁺	Mg ²⁺	K ⁺	Na ⁺	SiO ₂	CE	pH	Eh	STD Calculado
	(mg/L)									μS/cm		(mV)	(mg/L)
10	0,00	0,14	3,50	0,02	0,35	0,04	0,07	1,85	0,26	2,30	5,09	645,00	6,23
11	0,22	0,06	19,00	3,89	1,99	0,56	1,62	2,64	0,39	13,00	5,67	543,00	30,37
19	0,00	0,09	16,00	0,00	1,92	0,80	3,20	2,19	0,72	24,00	5,94	550,00	24,92
20	0,13	0,31	42,00	1,11	6,50	1,20	3,50	5,90	19,00	89,60	5,70	416,00	79,65
40	1,70	0,92	27,00	7,09	9,20	1,20	2,10	1,10	12,00	88,40	6,30	414,00	62,31
41	1,22	0,14	3,50	8,20	0,68	1,03	2,89	1,05	5,48	121,05	5,81	391,50	24,18
47	2,61	0,04	2,50	0,49	0,00	0,12	0,26	4,72	0,25	11,00	5,14	503,00	10,99
49	0,57	0,05	4,50	8,70	0,93	0,98	1,71	1,37	4,84	23,30	5,75	446,00	23,63
59	0,56	0,04	2,00	3,89	0,11	0,19	0,51	2,85	3,74	53,15	6,06	531,50	13,88
68	0,36	0,04	6,00	2,58	0,29	0,52	2,09	2,17	6,24	13,90	6,04	460,50	20,28
69	0,11	0,00	5,00	1,33	0,54	0,26	1,10	0,00	9,10	9,60	6,20	443,00	17,44
76	0,34	0,61	0,00	8,41	0,37	0,90	1,50	0,44	9,10	42,20	4,90	384,00	21,67
77	0,01	0,04	6,00	0,07	0,28	0,20	1,20	3,09	0,23	9,00	5,37	515,00	11,12
81	0,97	0,09	4,50	3,75	0,80	0,71	1,76	2,44	4,85	16,70	5,66	375,50	19,85
84	1,10	0,19	38,00	3,10	11,00	1,40	3,40	2,10	15,00	95,90	6,20	470,00	75,29
90	1,52	0,08	4,00	3,28	1,24	0,49	2,07	0,74	4,13	19,45	5,81	448,50	17,51
93	1,80	1,10	36,00	1,90	8,10	2,10	2,10	2,10	17,00	91,30	6,40	309,00	72,20
96	0,93	0,12	1,00	2,96	0,34	0,21	1,04	0,64	3,71	59,40	6,03	448,00	10,93
144	4,50	0,00	0,00	4,07	1,50	1,40	2,50	0,33	10,00	34,20	5,80	380,00	24,30
170	0,41	0,21	6,00	0,39	1,18	0,47	0,42	3,18	0,26	11,00	5,37	595,00	12,52
172	0,08	0,16	3,50	0,71	0,42	0,20	0,13	2,19	0,25	5,10	5,03	511,00	7,64
191	0,29	0,00	11,00	0,44	2,10	0,51	1,70	0,53	16,00	25,80	5,90	468,00	32,57
192	1,33	0,05	1,50	3,89	0,33	0,28	1,32	1,13	3,92	18,70	6,12	484,50	13,74
200	0,88	0,06	5,00	7,53	0,84	0,91	2,50	0,35	12,00	24,40	5,80	477,00	30,07
215	0,05	0,02	5,50	0,68	0,33	0,26	1,05	3,05	0,38	6,50	5,32	511,00	11,32
238	0,17	0,00	2,00	2,66	0,14	0,27	0,80	0,74	8,30	11,90	5,80	395,00	15,08

Poço Amostrado	Cl ⁻	SO ₄ ⁻	HCO ₃ ⁻	NO ₃ ⁻	Ca ²⁺	Mg ²⁺	K ⁺	Na ⁺	SiO ₂	CE	pH	Eh	STD Calculado
	(mg/L)									µS/cm		(mV)	(mg/L)
256	0,54	0,05	0,00	1,13	0,12	0,08	0,40	0,63	2,83	9,80	5,16	477,00	5,78
257	0,54	0,09	0,50	1,57	0,07	0,07	0,42	0,81	2,39	8,20	5,35	468,50	6,43
267	1,08	0,14	1,99	3,39	0,65	0,26	1,29	0,49	0,55	15,00	5,47	555,00	9,84
281	0,00	0,07	10,00	0,03	0,19	0,43	3,04	1,89	0,23	14,00	5,80	561,00	15,88
282	0,03	0,05	11,00	0,00	1,03	0,82	1,81	1,67	7,26	19,20	6,06	548,50	23,65
288	0,12	0,02	12,00	3,99	2,51	0,72	1,52	1,95	5,76	28,25	6,14	438,00	28,57
307	0,79	0,00	27,00	11,07	7,20	1,40	3,40	1,40	15,00	16,20	6,10	532,00	67,26
322	0,31	0,05	5,00	8,15	1,57	0,80	2,02	3,15	0,35	20,00	5,39	490,00	21,40
327	0,28	0,02	2,25	5,40	0,64	0,37	0,87	1,79	4,26	11,30	5,31	469,50	15,87
334	1,37	0,57	14,00	15,59	3,71	2,64	3,72	2,60	7,29	58,20	5,97	379,00	51,47
337	0,03	0,05	12,00	0,22	2,45	0,78	2,57	3,02	0,55	18,00	5,93	484,00	21,67
374	0,87	0,00	3,00	6,64	0,16	0,02	0,14	3,10	4,90	21,50	5,20	362,00	18,83
381	0,02	0,00	0,00	1,06	0,18	0,16	0,17	0,00	11,70	3,25	5,80	423,00	13,29
387	1,80	0,00	0,00	7,97	0,31	0,29	0,19	3,30	5,60	25,60	5,10	465,00	19,46
403	0,01	0,06	7,00	0,01	0,09	0,27	1,33	2,76	0,44	8,60	5,36	592,00	11,97
433	0,92	0,22	9,99	0,68	0,48	0,29	2,45	2,00	0,96	21,00	6,47	486,00	17,99
448	0,03	0,10	10,00	1,07	1,18	0,80	0,27	2,74	0,42	17,00	5,72	555,00	16,61
475	0,28	1,80	3,00	3,54	0,44	0,41	1,57	0,18	7,50	15,80	5,00	388,00	18,72
483	0,49	0,03	0,50	5,69	0,60	0,30	0,12	2,03	2,61	14,80	5,01	488,00	12,36
490	0,47	0,00	2,00	2,21	0,71	0,27	0,11	0,43	5,30	11,20	6,80	320,00	11,50
495	0,17	0,08	38,00	1,33	8,40	1,42	3,10	2,20	20,00	85,10	6,30	321,00	74,70
503	0,55	0,29	5,00	2,21	0,83	0,62	1,80	0,55	11,00	35,30	6,00	386,00	22,85
505	4,00	0,00	1,00	9,30	1,40	1,90	2,90	0,78	15,00	45,30	5,40	378,00	36,28
540	1,14	0,06	0,50	6,63	0,44	0,28	1,98	1,29	3,93	17,25	5,60	456,50	16,23
561	0,48	0,07	1,00	5,18	0,26	0,21	0,14	1,74	3,19	16,15	5,13	551,50	12,26
567	3,75	0,12	0,00	14,42	0,73	0,43	1,54	3,05	3,34	45,65	4,60	548,00	27,36
572	1,90	0,03	0,00	11,51	3,10	1,60	0,75	2,80	9,10	14,10	5,50	613,00	30,79
606	4,03	0,47	23,75	17,90	6,75	3,39	3,37	3,98	6,70	87,15	6,63	461,50	70,32
607	4,30	0,03	7,00	21,26	3,50	3,10	4,90	0,71	14,00	80,20	6,10	324,00	58,80

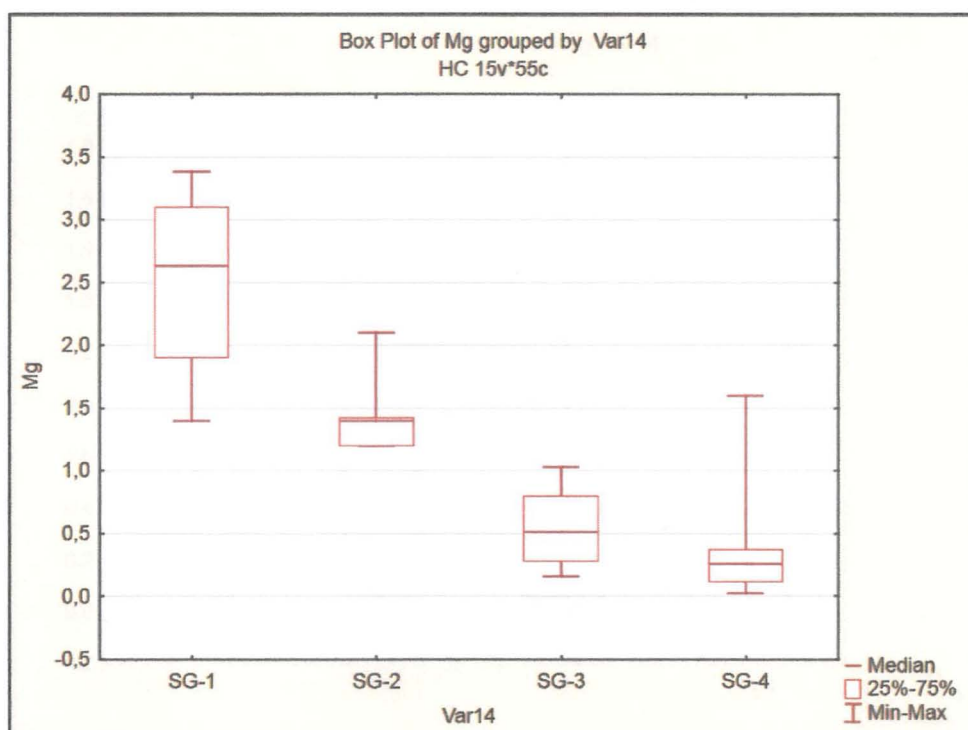
ANEXO B – BOX PLOTS



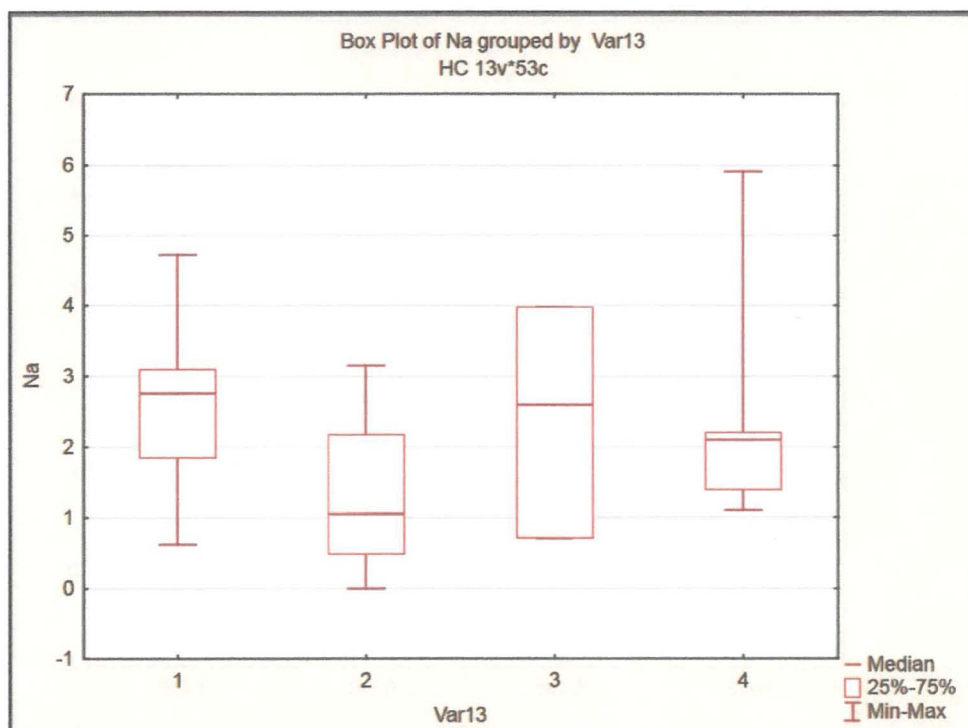
Box Plot 1 - Potássio



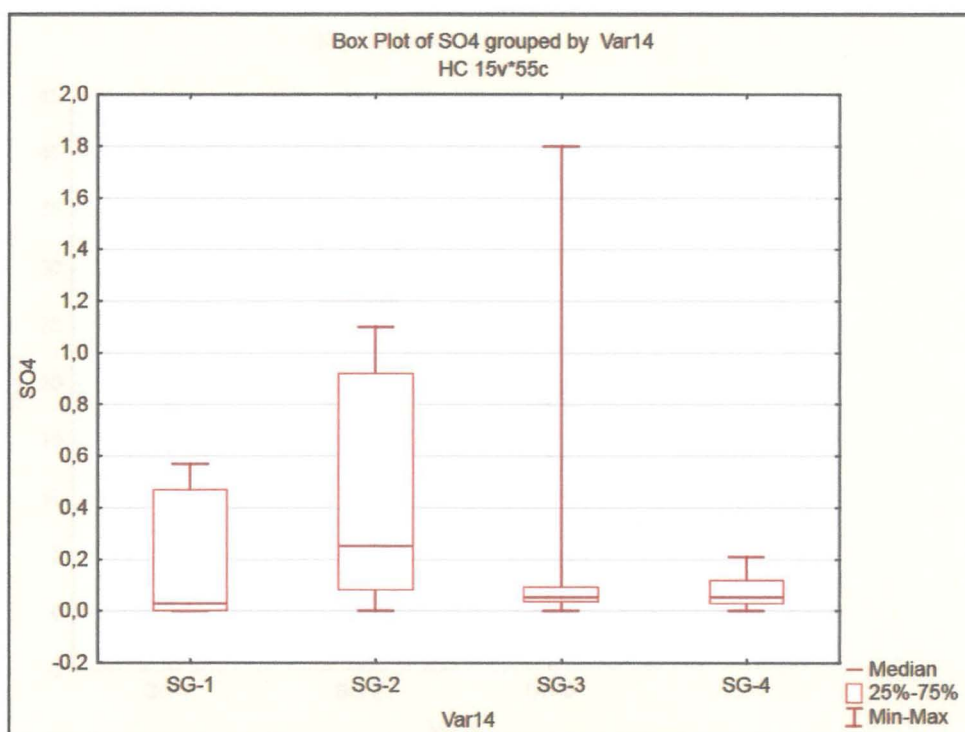
Box Plot 2 - Cálcio



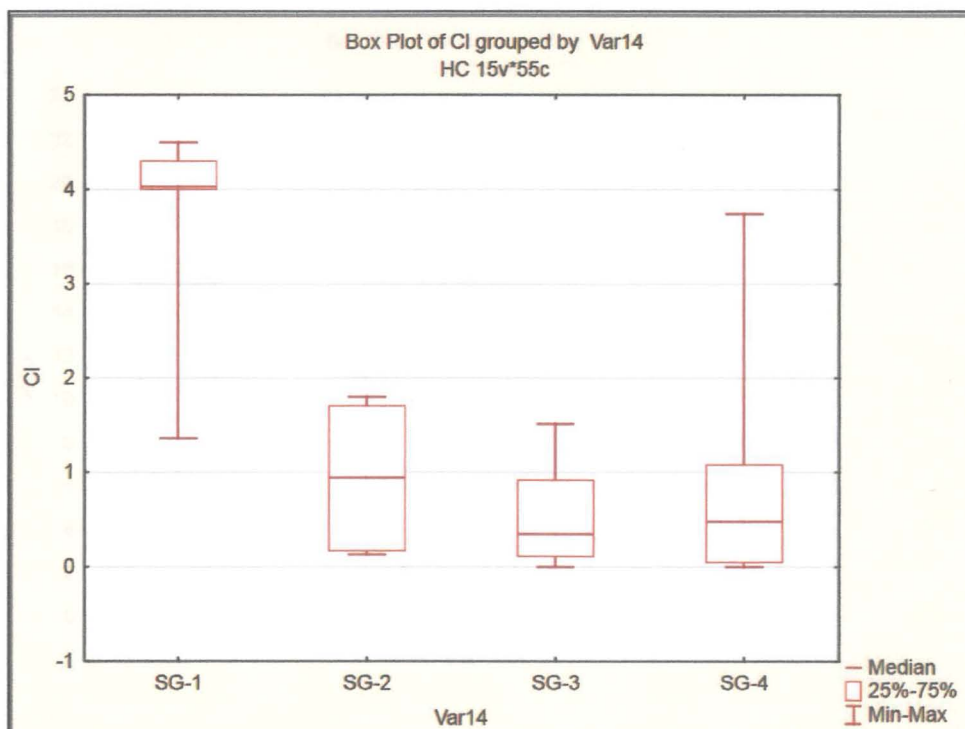
Box Plot 3 - Magnésio



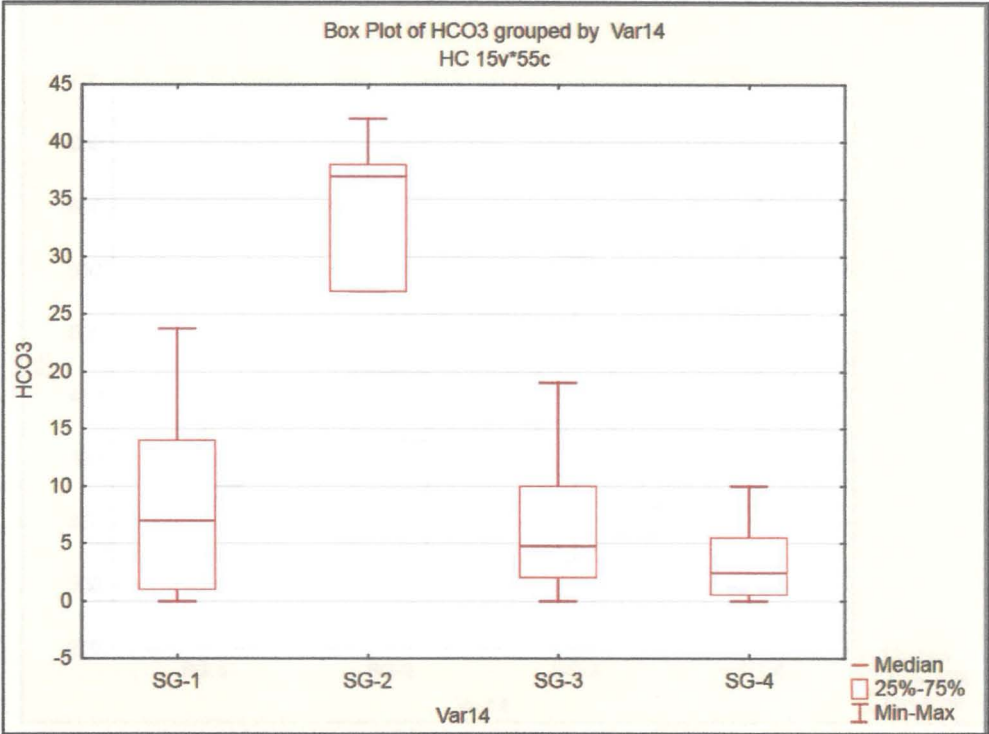
Box Plot 4 - Sódio



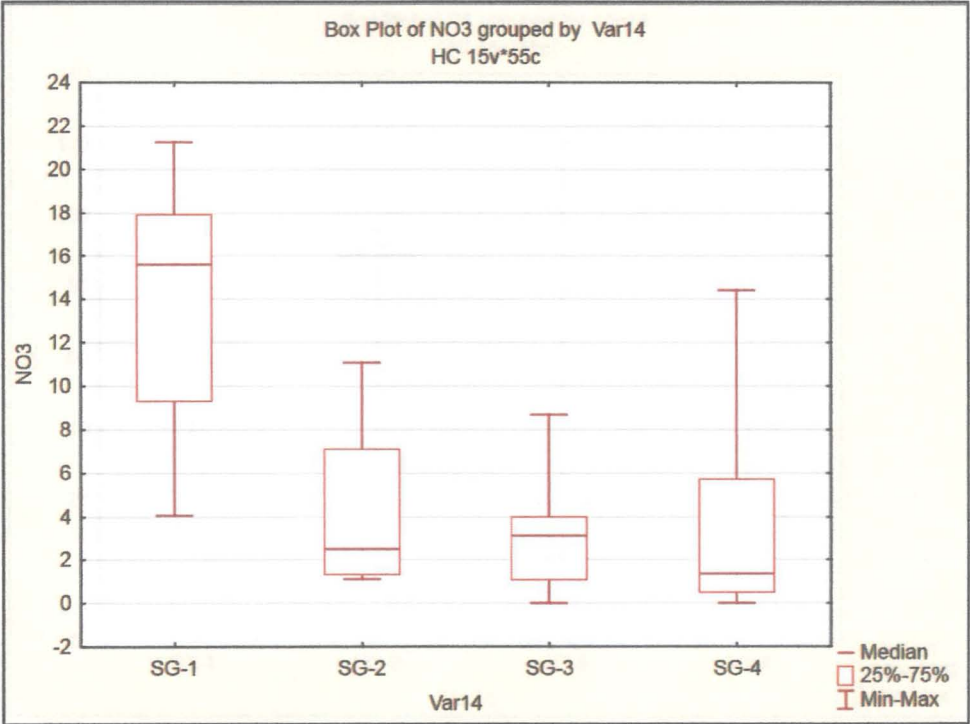
Box Plot 5 – Sulfato



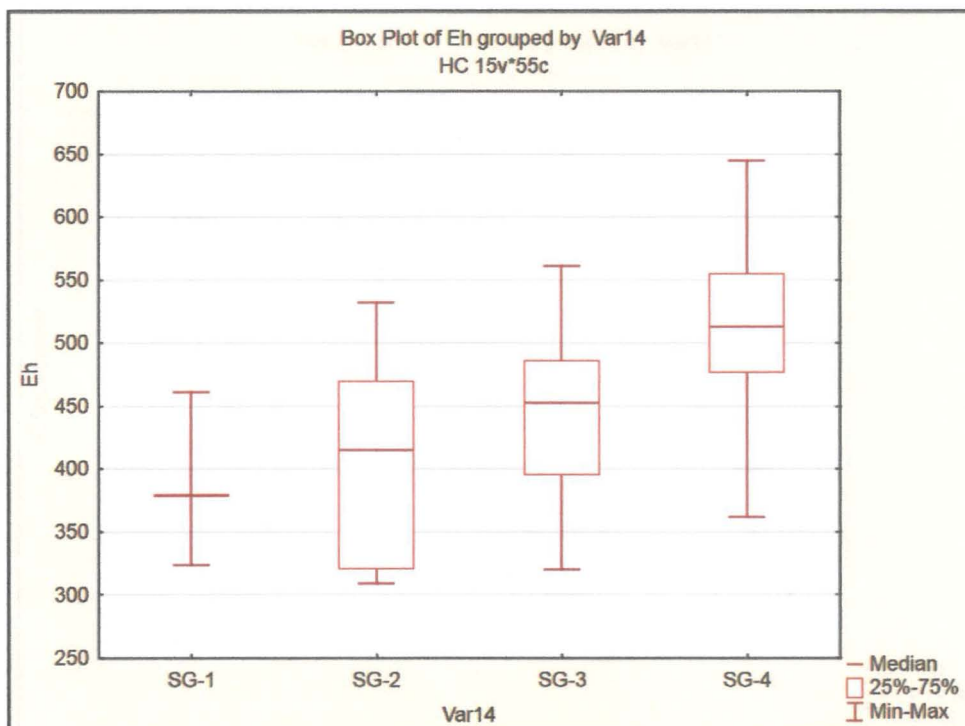
Box Plot 6 Cloreto



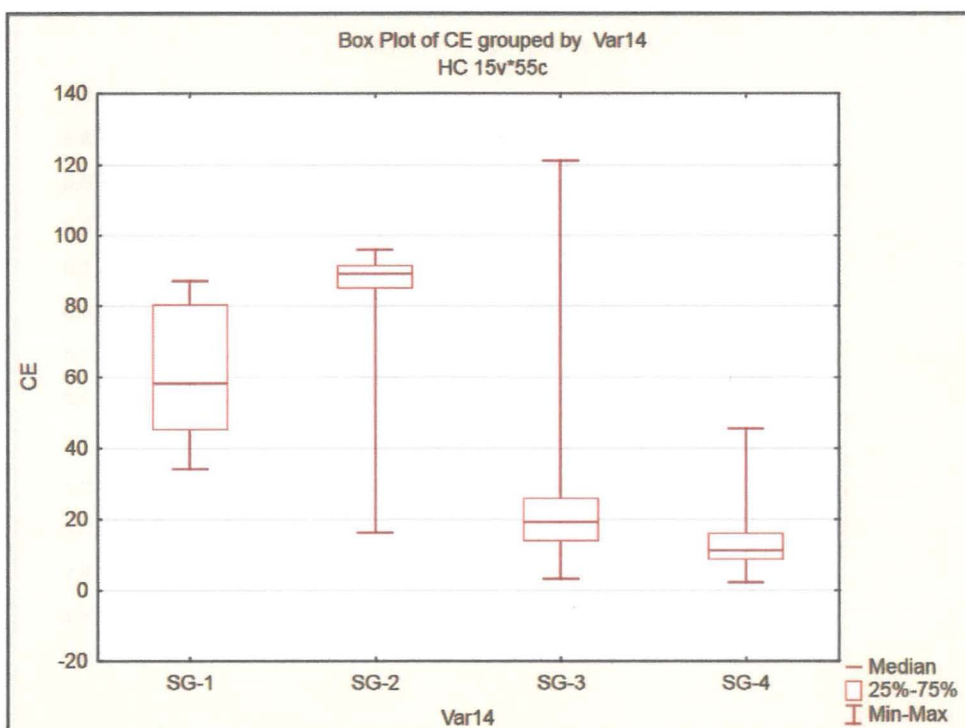
Box Plot 7 – Bicarbonato



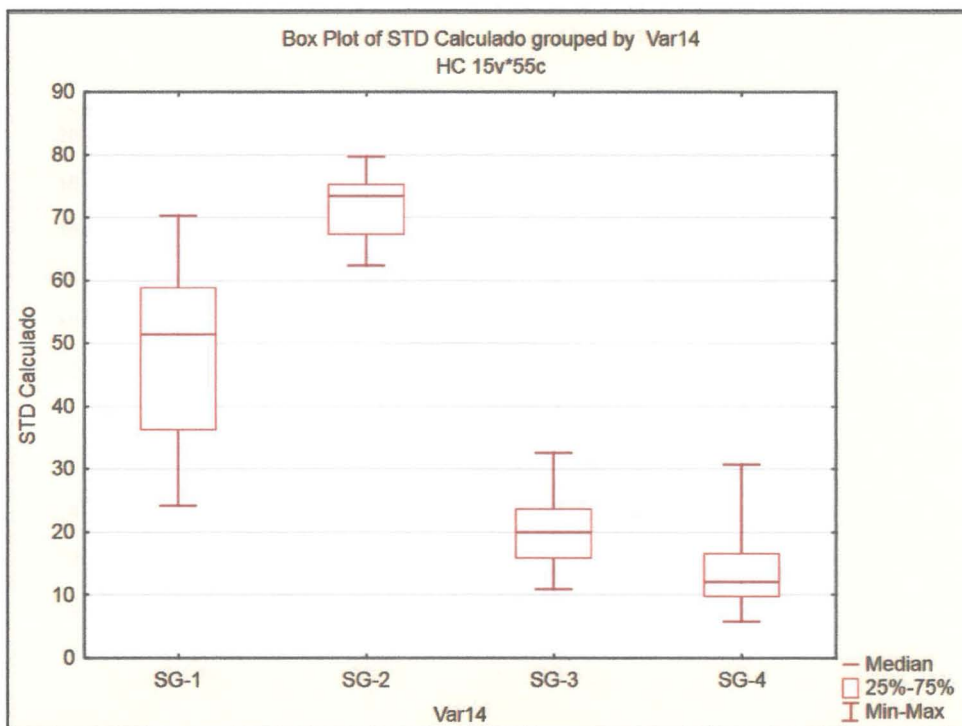
Box Plot 8 – Nitrato



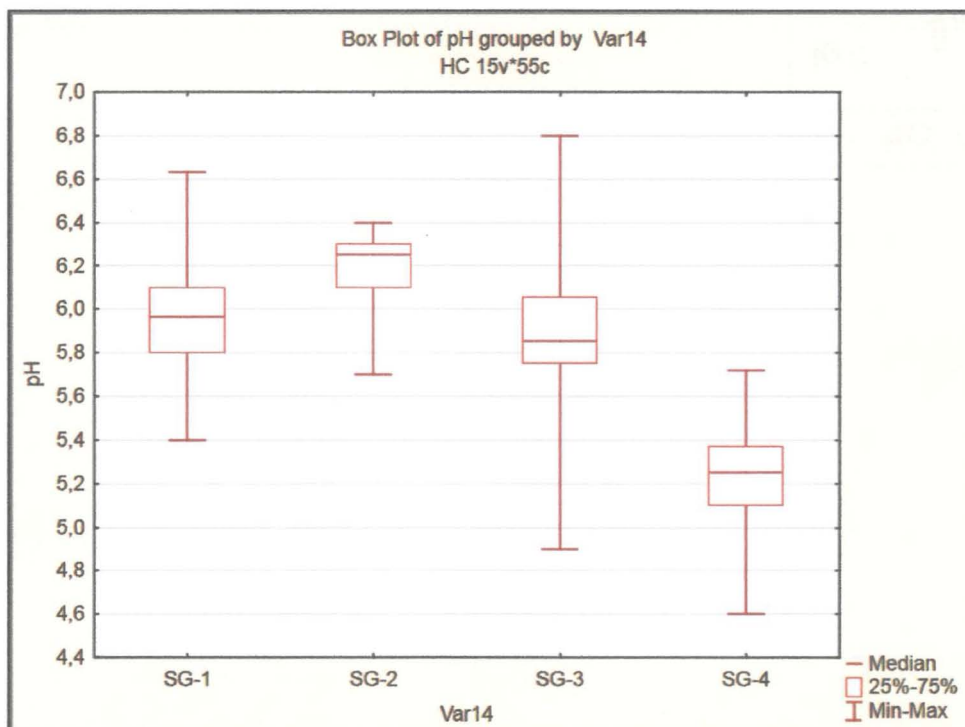
Box Plot 9 –Potencial de Oxi-redução



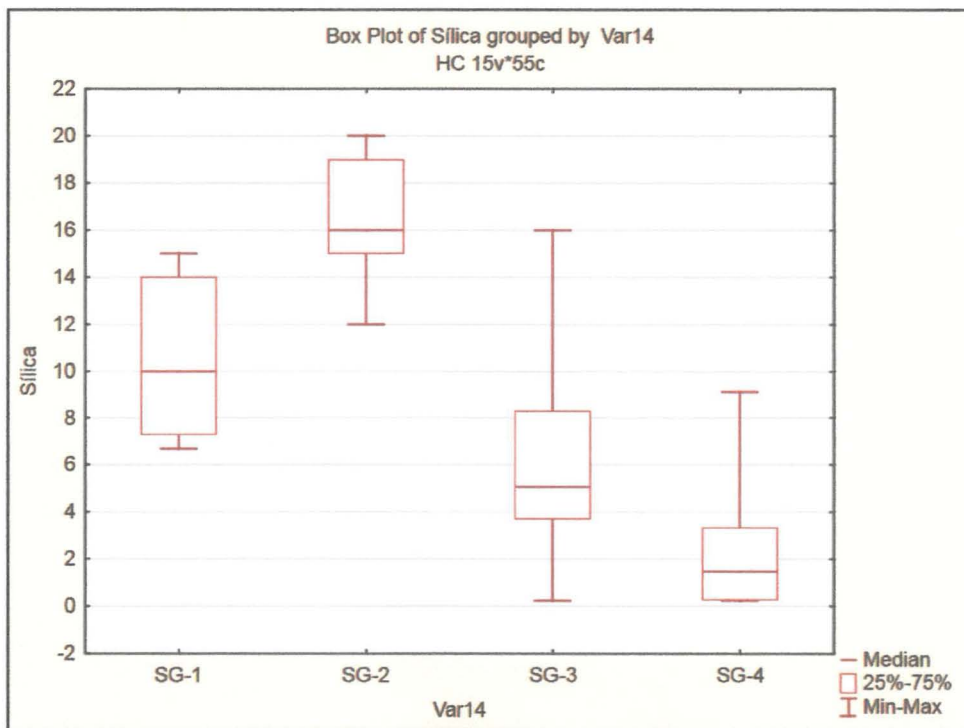
Box Plot 10 – Condutividade Elétrica



Box Plot 11 – Sólitos Totais Dissolvidos



Box Plot 12 - pH



Box Plot 12 – Sílica

DOAÇÃO Graduação
IGC

Data: 20 / 03 / 18