HYDROLOGICAL PROCESSES
Hydrol. Process. **16**, 1353–1372 (2002)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/hyp.1003

The cation and silica chemistry of a Subandean river basin in western Amazonia

J. A. Sobieraj,[1] H. Elsenbeer[1]* and M. McClain[2]

[1] *University of Cincinnati, Department of Civil and Environmental Engineering, Cincinnati, Ohio 45221-0071, USA*
[2] *Florida International University, Department of Environmental Studies, Miami, Florida 33199, USA*

Abstract:

We sampled river water at 13 locations in the Pichis basin, a 10 500 km^2 large rainforest-covered drainage basin in Peru, to assess the influence of lithological variability and seasonality on water chemistry. The concentrations of major cations and silica show a strong seasonal dependence and a remarkable variability over short distances that is only weakly reduced in the wet season; cation concentrations in streams differ by up to 100% within a few kilometres. The lowest cation concentrations were associated with relatively cation-depleted upper Tertiary and lower Quaternary formations, whereas relatively cation-rich lower Tertiary and Jurassic formations left a clear calcium and sodium signal in the respective rivers. Cluster analysis, in conjunction with boxplots, suggests that the sampling locations can be segregated into three groups based on similarities of their geochemical signals. According to the previously defined criteria, one river is classified as a Group 2 river with $200 < TZ+ < 450$ μeq/L, whereas all other rivers fall into Group 3 with $450 < TZ+ < 3000$ μeq/L (where TZ+ refers to the total cation charge). Based on a comparison with other studies at different sections of the Amazon mainstem, the river chemistry of our study area is relatively enriched in K^+, Mg^{2+} and Ca^{2+}, and, consequently, has a higher TZ+ value, while being relatively depleted in silica. The influence of lithological variability on water chemistry must be considered in land-use change studies even at watershed areas of 26–3382 km^2. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS western Amazonia; Subandean zone; lithology; river chemistry; spatial variability; tropical rainforest

INTRODUCTION

The Amazon River drainage basin extends from about 79 °W to 46 °W, and from 5 °N to 17 °S, and covers $7{\cdot}05 \times 10^6$ km^2 (Sioli, 1984). Given the continental scale of this entirely tropical river system, a considerable lithological diversity is to be expected (Vital and Stattegger, 2000), for which the influence on fluvial geochemistry was first studied in detail by Stallard and Edmond (1983). The still popular, colour-based classification of Amazonian rivers, however, seems to be the product of early nineteenth century travellers (Reindl, 1903). Sioli (1964, 1965) provided a scientific basis for this classification and related it to the lithology of the source areas. Whitewater (high suspended sediment concentration, near-neutral pH) rivers originate in the Andes, which delimit the Amazon basin to the west; the best examples for this type of river include the Amazon itself and tributaries such as the Rio Madeira. Clearwater rivers (low suspended sediment concentration) drain the shield areas to the north and south of the Amazon trough, such as the orographically right-hand tributaries Rio Tapajós and Rio Xingú. Blackwater rivers (low suspended sediment concentration, high in dissolved organic matter, low pH) are associated mainly with extensive areas of podsols in the northern part of the Amazon basin.

* Correspondence to: H. Elsenbeer, University of Cincinnati, Department of Civil and Environmental Engineering, Cincinnati, Ohio 45221-0071, USA. E-mail: helmut.elsenbeer@uc.edu

Received 13 February 2001
Accepted 18 June 2001

Stallard and Edmond (1983) introduced quantitative categories based on total cation charge (TZ+) and related them to lithology: (Group 1) rivers with $0 < TZ+ < 200$ μeq/L drain the most intensely weathered materials, have high levels of Fe^{3+}, Al^{3+}, H^+ and colouration, and are enriched in Si relative to other major species, (Group 2) rivers with $200 < TZ+ < 450$ μeq/L drain siliceous terrains and also are enriched in silica relative to other species, (Group 3) rivers with $450 < TZ+ < 3000$ μeq/L drain marine sediments or red beds with high cation concentrations and have high levels of Ca^{2+}, Mg^{2+}, alkalinity and SO_4^{2-}, (Group 4) rivers with $TZ+ > 3000$ μeq/L drain massive evaporites and are enriched in Na^+ and Cl^- (Stallard and Edmond, 1983).

This approach, the usefulness of which has been confirmed by, for example, Edmond *et al*. (1996), offers a considerable spatial refinement over the tri-colour typology, which is not useful at any scale other than the continental. With increasing interest in the effect of land-use changes in the Amazon Basin on water quality, the focus of river chemistry studies has begun to shift to mesoscale (e.g. Mortatti *et al*., 1992) and smaller basins (Biggs *et al*., Accepted). At these scales, even more refined categories will become necessary to be able to detect hydrochemical land-use change signals against the noise of lithology related hydrochemical variability. Studies such as the ones cited invariably involve some paired-basin approach that assumes a uniform lithological signal, or at least a hydrochemical variability that is smaller than the suspected land-use signal.

Our study addresses the link between hydrochemical and lithological variability at the submesoscale level in the Pichis basin, a whitewater region of the Amazon Basin presumably in the $450 < Tz+ < 3000$ μeq/L bracket (Stallard and Edmond, 1983). Our specific objectives were (i) to discern lithological signals and (ii) to determine their season dependence in the water chemistry of tributaries and headwaters of the Pichis basin.

RESEARCH AREA

Geology

The Amazon Basin consists of four major morphostructural units: (i) the shields (Guyana Shield to the north and the Brazilian Shield to the south), (ii) the Central Amazon Trough, (iii) the Subandean Zone and (iv) the Andean Cordillera to the west (Figure 1). The study area is located in the Subandean Zone, which is part of a tectonically active region owing to low-angle subduction of the oceanic Nazca Plate (Suarez *et al*., 1983).

The Subandean Zone, which is situated between the Andes proper and the Iquitos intraforeland uplift, encompasses four retroarc intraforeland basins: the Pastaza-Maranon, Ucayali, Madre de Dios-Beni, and Acre basins (Räsanen *et al*., 1992). These intraforeland basins (<500 m a.s.l.) are delineated by emerging arches or structural highs (Räsanen *et al*., 1990), and central portions of the Pastaza-Maranon and Ucayali basins contain significant swamp areas that are characterized by widespread inundations and aggradation (Kalliola *et al*., 1991). In the Ucayali Basin, reverse faulting at the eastern edge of the Shira Mountains and uplift of the Serra do Moa intraforeland uplift caused the tilting of alluvial plains and the formation of terraces (Räsanen *et al*., 1992).

The Pichis basin is part of the Ucayali intraforeland basin. It is delineated by the San Matias anticline to the west and the Shira Monocline to the east (Figure 1). The flow of the Pichis, Azupizu and Nazarategui rivers corresponds roughly to the axis of the gently dipping syncline in between these two folded mountain ranges, whereas the orientation of tributaries, such the Apurucayali, Chivis, Esperanza, Lorenzo and Sungaro rivers is oblique to normal to the axis of the syncline (Figure 2).

The stratigraphy of the study area is derived from a geological map that covers only a portion of the study area (ONERN, 1981b); it comprises lithological units ranging in age from the Permo-Carboniferous Obentini Formation to Quaternary alluvial deposits (Table I). The latter consists of highly weathered, relatively cation-depleted, unconsolidated to slightly consolidated gravel, sand, silt and clay that occur in proximity to river channels. In the Ucayali intraforeland basin, surficial sandy beds occur at erosional banks that are carved into

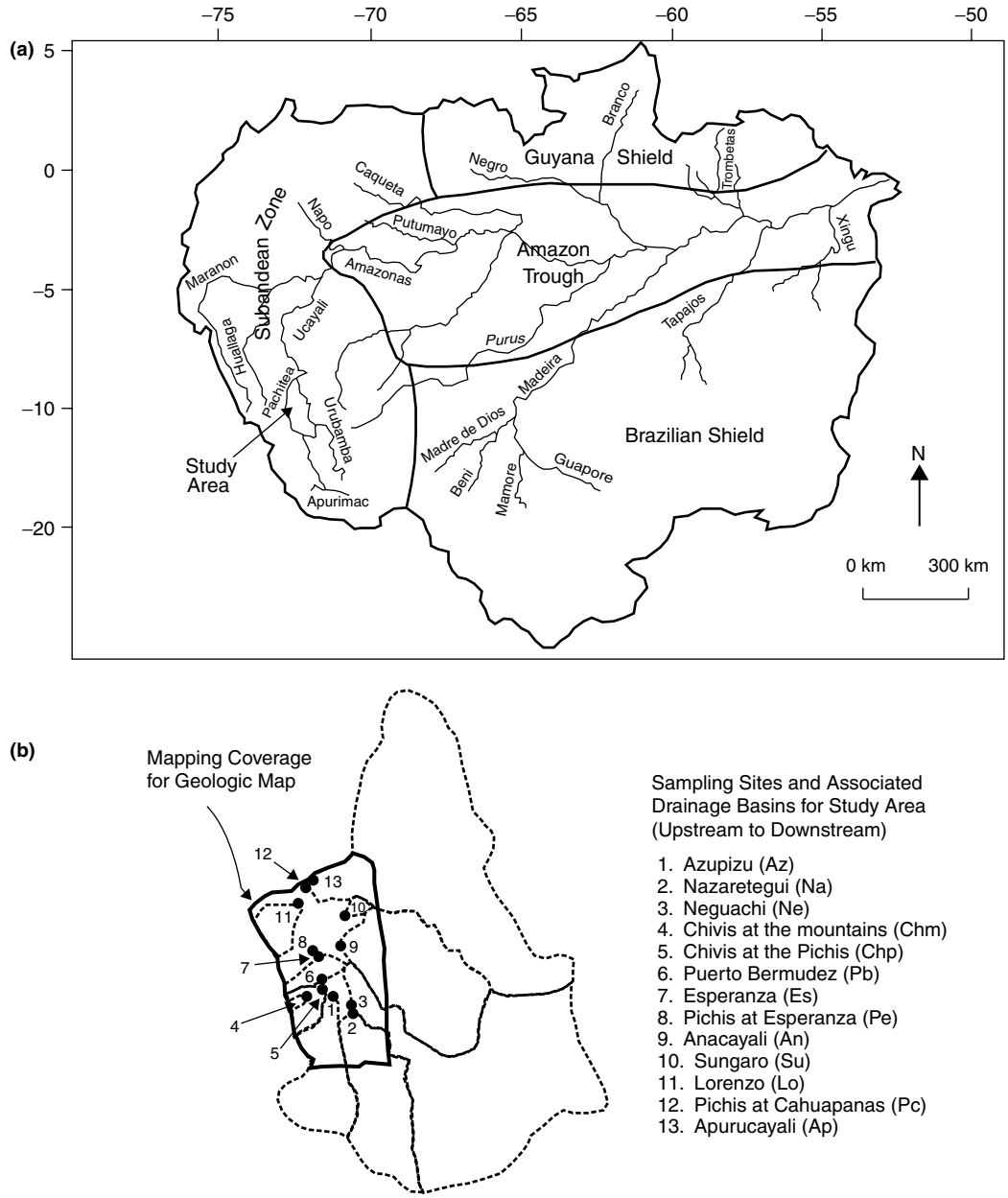

Figure 1. The location of the study area (a) within the context of major physiographic provinces, and (b) the various sub-basins comprising the Pichis basin

terraces along present-day rivers (Räsanen *et al.*, 1990). A basal horizontal gravel layer of the Quaternary Ucayali Formation unconformably overlies the Tertiary deposits of the Ipururo Formation.

Tertiary sediments are the most areally extensive deposits in the mapped sections of the study area, cropping out in the central regions of the syncline. The Ipururo Formation, Chambira Formation and

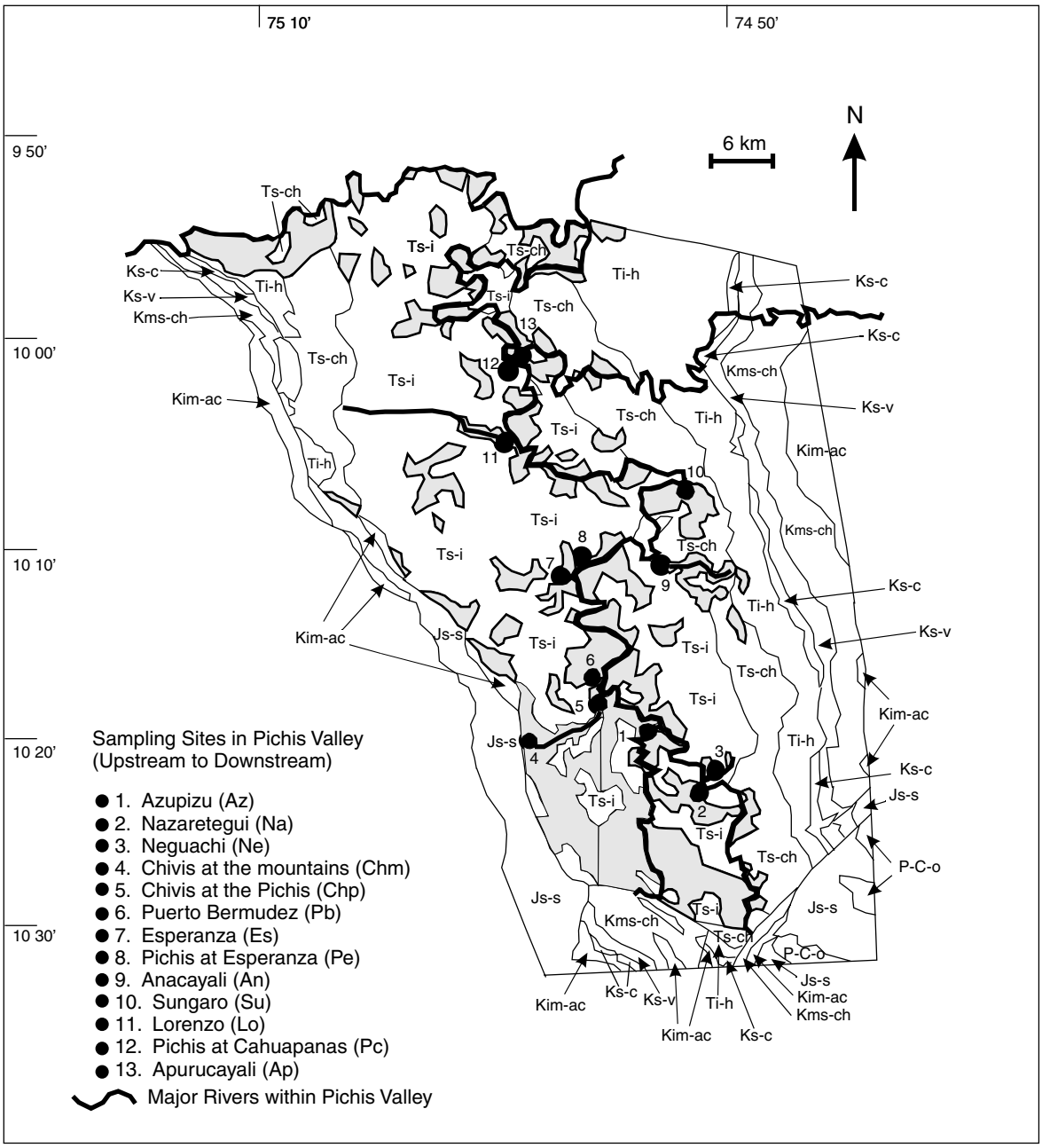

Figure 2. Geological map of the Pichis Valley (modified from ONERN, 1981a; Reproduced by permission of Instituto Nacional de Recursos Naturales (INRENA)), showing major tributaries and sample locations. Symbols for lithological units correspond to those listed in Table I. Shaded areas indicate undifferentiated Quaternary units

Huayabamba Group consist of sandstone and shale, with minor amounts of glauconite, anhydrite and limestone. An unconformity separates Tertiary deposits from Cretaceous deposits. The Cretaceous Cachiyacu, Vivian, Chonta and Agua Caliente formations crop out in relatively thin bands along the eastern and western edges of the syncline. Deposits range from cross-bedded quartz arenites, shale, mudstone, limestone and marl,

Table I. Stratigraphical descriptions for lithological units in the study area

Era	System	Epoch/Age	United stratification	Lithological description	Symbol
Cenozoic	Quaternary	Holocene	Alluvial deposits	Loose horizontal alluvium with sand, silt and clay recently rounded	Qr-a
			Colluvial deposits	Unconsolidated subangular gravel and subrounded sand and clay	Qhc
		Pleistocene	Ucayali formation	Slightly consolidated older alluvium consisting of clay, mud, silt and sand in subhorizontal layers with a basal horizontal gravel layer	Qp-u
	Tertiary	Pliocene	Ipururo formation	Red and thick grey sand exhibiting cross-bedding, sometimes bioturbated with glauconite. Intercalations of grey marl, calcareous claystone, with tuffaceous quartzose sediments	Ts-I
		Miocene	Chambira formation	Red and brown claystone with minor amounts of shale, fine whitish sandstone, limestone and marl with traces of anhydrite	Ts-ch
		Paleocene–Oligocene	Huayabamba group	Mudstone and red and purple claystone, firm and friable, intercalated with glauconitic, whitish, calcareous siltstone beds, tuffs, and red sandstone throughout most of the formation	Ti-h
Mesozoic	Cretaceous	Maastrichtian	Cachiyacu Formation	Greenish-grey shale, fissile with siltstone, brown and red mudstone. White fine sand at the top of the formation	Ks-c
			Vivian formation	White, brown and yellowish, fine-grained to coarse-grained quartz arenite with cross-beds; durable with moderate to high porosity	Ks-v
		Cenomanian-Coniacian	Chonta formation	Grey limestone, well-stratified with shale and light marl with brownish cemented siltstone, and thin white arenites and fossiliferous beds locally	Kms-ch
		Albian	Agua caliente formation	Whitish arenites, fine, friable and durable with thin beds of cemented siltstone and shale, sometimes calcareous, with plant remains and cross-bedding	Kim-ac
	Jurassic		Sarayaquillo formation	Dark yellow, red and brown, fine, calcareous sandstone, sometimes ferruginous in outcrop, with salt contents; minor amounts of cemented siltstone and calcareous, red shale	Js-s
Palaeozoic	Permo-Carboniferous		Obentini formation	Red and yellow shale, with calcareous sandstone, limestone and grey fissile shale; quartz arenite with cross-bedding at the base	P-C-o

with some units containing fossils. The Cretaceous deposits are bounded by unconformities. The Jurassic Sarayaquillo Formation is a calcareous sandstone with local salt deposits. Palaeozoic units consist of the Permo-Carboniferous Obentini Formation, which is comprised of shale, calcareous sand, limestone and quartz arenite. This formation crops out only in the southeastern part of the study area.

Soils

Soil information comes from a soil map of the area (ONERN, 1970). The mapped soil coverage included the Pichis Valley, which is part of the Pichis basin, areas to the south of the Pichis Valley, and the eastern slope of the San Matias anticline to the west. Soil information is not available for the easternmost drainage basins within the study area. Entisols and Inceptisols occur as thin bands adjacent to rivers and tributaries within the Pichis Valley, but Ultisols are the predominant soil taxonomic order within the Pichis Valley. Inceptisols are the primary soil taxonomic order in mountainous regions, although a thin band of vertisols coinciding with the spatial distribution of the Huayabamba Group occurs along the eastern part of the Pichis Valley.

Land use and land cover

ONERN (1981a, p. 227) conducted a survey of land use and land cover in a roughly rectangular area of 1286 km^2 that covered the axis of the Pichis Valley and abutted the slopes of the San Matias anticline to the west. The survey did not, however, extend near the extensive Shira monocline to the east or the headwater mountains to the south, areas with a population density even lower than the central valley. In the year of the survey (1977), agricultural land-use within the Pichis Valley comprised 5% of the area surveyed. The overall areal extent of agricultural land-use in the Pichis drainage basin as a whole therefore was much less than 5%. We conclude that the anthropogenic influence on river chemistry within the study area was negligible in 1977. There is, however, no available land-use information for the period between 1987 and 1989, the sampling period for this study. Even though the areal extent of agricultural areas has certainly increased from 1977 to 1987, we believe that only a small percentage of the area was subject to anthropogenic influences.

METHODS

Sampling design and drainage area characterization

We collected water samples at the following nine tributaries just upstream of their confluence with the Pichis River: Anacayali (An), Apurucayali (Ap), Azupizu (Az), Chivis (Chp), Esperanza (Es), Lorenzo (Lo), Nazarategui (Na), Neguachi (Ne) and Sungaro (Su). In addition, the Pichis River mainstem was sampled at Puerto Bermudez (Pb), at Cahuapanas (Pc) and at Esperanza (Pe), and the Chivis River was sampled at the location where it exits the mountains (Chm). These locations, shown in Figure 2, will hereafter be referred to by their abbreviations. Table II summarizes the sampling frequency, beginning and ending sampling dates, and total number of observations for each location; each observation represents the median of three samples.

The spatial and topographic characteristics of the study area were determined in a geographical information system (ArcInfo) using publicly available digital data and data digitized from paper maps. A digital elevation map (DEM) with 1-km-grid resolution was used to infer the topography, and river channels were hand digitized from Peruvian national maps. Drainage areas above each sampling point were delineated by following topographic highs between river segments. The calculated area of each drainage basin, as well as the minimum, maximum and mean elevations, are reported in Table III. Available lithological data were hand digitized (ONERN, 1981b) and queried to determine the areal and percentage coverage of geological units within each drainage basin. Although the geological map covers only 20% of the total area drained by our most downstream sampling points, it covers 100% of five of the 13 drainage basins investigated, all of which are western tributaries of the Pichis. Table IV highlights

Table II. Sampling frequency at the different locations

Location	Start date[a] (m/d/y)	End date (m/d/y)	Approximate Frequency[b]	Sample Total[c]
Anacayali (An)	1/20/88	4/12/89	Monthly	16
Apurucayali (Ap)	12/9/87	4/12/89	Monthly	19
Azupizu (Az)	4/27/88	4/12/89	Bimonthly	24
Chivis at the mountains (Chm)	5/18/88	2/8/89	Bimonthly	17
Chivis at the Pichis (Chp)	5/18/88	4/12/89	Bimonthly	24
Esperanza (Es)	3/16/88	7/5/89	Weekly	68
Lorenzo (Lo)	1/20/88	7/5/89	Monthly/Weekly	28
Nazarategui (Na)	4/27/88	4/12/89	Bimonthly	24
Neguachi (Ne)	4/27/88	4/12/89	Bimonthly	24
Pichis at Puerto Bermudez (Pb)	4/27/88	4/12/89	Bimonthly	24
Pichis at Cahuapanas (Pc)	7/15/87	6/21/89	Monthly/Weekly	46
Pichis at Esperanza (Pe)	1/7/87	7/5/89	Weekly	102
Sungaro (Su)	1/20/88	4/12/89	Monthly	17

[a] Indicates commencement of regular sampling.
[b] Excludes minor gaps in sampling intervals.
[c] Represents largest sampling total for constituent of concern.

Table III. Calculated drainage basin area and topographic information for each sampling location

Sample Location	Drainage basin area (km^2)	Minimum elevation (m a.s.l.)	Maximum elevation (m a.s.l.)	Mean elevation (m a.s.l.)
Anacayali (An)	1336·2	193	1155	524·9
Apurucayali (Ap)	3381·9	184	1579	715·9
Azupizu (Az)	663·0	257	938	562·3
Chivis at the mountains (Chm)	26·0	297	416	384·2
Chivis at the Pichis (Chp)	136·0	258	526	368·6
Esperanza (Es)	88·1	192	419	328·9
Lorenzo (Lo)	166·2	181	417	309·8
Nazarategui (Na)	901·1	271	1004	658·7
Neguachi (Ne)	2903·9	257	1983	998·7
Pichis at Puerto Bermudez (Pb)	4690·8	234	1983	541·6
Pichis at Cahuapanas (Pc)	6836·2	178	1983	438·2
Pichis at Esperanza (Pe)	4807·9	210	1983	509·3
Sungaro (Su)	38·9	185	406	301·2

the percentage area of relatively cation-depleted units, the percentage area of relatively cation-rich units, and the percentage area in which there was no available geological information (percentage unmapped area).

Chemical Analysis

Water samples were analysed for Ca^{2+}, K^+, Mg^{2+}, Na^+, SiO_2, electric conductivity and pH; the latter two parameters were determined immediately after sampling. One-hundred-millilitre, unfiltered subsamples with 1% (vol/vol) phenyl-mercury-acetate (to prevent microbial growth) were transported to the La Molina Agricultural Research Station in Lima within a week and filtered (0·45 μm) before further analysis. The Ca^{2+}, K^+, Mg^{2+} and Na^+ were determined by conventional atomic adsorption spectrometry (Perkin Elmer 2380),

Table IV. Percentage area of cation-poor and cation-rich lithological units, unmapped area and total area for drainage basins of each sampling location

Location	Percentage cation-poor area[a]	Percentage cation-rich area[b]	Percentage unmapped area
Anacayali (An)	1·8	8·1	79·6
Apurucayali (Ap)	0·8	2·3	92·4
Azupizu (Az)	14·9	15·7	67·8
Chivis at the mountains (Chm)	0·0	79·8	20·2
Chivis at the Pichis (Chp)	44·4	32·2	23·4
Esperanza (Es)	58·1	24·2	17·7
Lorenzo (Lo)	75·2	17·6	2·1
Nazarategui (Na)	9·6	8·9	75·5
Neguachi (Ne)	1·7	1·1	94·1
Pichis at Puerto Bermudez (Pb)	8·1	5·5	83·0
Pichis at Cahuapanas (Pc)	14·7	6·6	72·9
Pichis at Esperanza (Pe)	9·9	5·7	81·2
Sungaro (Su)	0·0	28·7	0·0

[a] Cation-poor lithological units consist of colluvial and alluvial deposits and the Ucayali and Ipururo formations.
[b] Cation-enriched formations consist of the Chonta, Agua Caliente, Sarayaquillo and Obentini formations.

and molybdate-reactive silica was determined colorimetrically at 660 nm. The relative precision for these analyses was 5% for cations and 10% for silica.

Data Analysis

Three data sets were evaluated:

1. a complete data set containing all observations (All);
2. a data set containing only dry season observations (between June and September, Dry);
3. and a data set containing only wet season observations (between October and May, Wet).

Boxplots. We constructed boxplots (Emerson and Strenio, 1983) for the three data sets (All, Dry and Wet) for all variables but show only those for Ca^{2+}, Mg^{2+}, Na^{+} and TZ+ (Figures 3–6). These boxplots were selected because they exhibit many of the features and trends found in the boxplots for all the variables, as well as show interesting results unique to only those variables.

Cluster Analysis. To detect similarities between the 13 sample locations based on their chemical signals, we performed cluster analysis, which is a technique for partitioning a set of objects (13 sample locations) into relatively homogeneous subsets based on interobject similarities. The goal of cluster analysis is to group together objects into clusters that display small within-cluster variation but large between-cluster variation (Kachigan, 1986). Two clustering algorithms, 'partitioning around medoids' and 'agglomerative nesting', were chosen to evaluate the data set. A discussion of clustering algorithms is presented in Kachigan (1986) and Mathsoft (1999a,b).

The method PAM was used for the 'partitioning around medoids' algorithm (see Chapter 2 of Kaufman and Rousseeuw (1990) for complete discussion). For this data set, there were 13 objects and seven variables (Ca^{2+}, K^{+}, Mg^{2+}, Na^{+}, SiO_2, CE and pH). Following the example of Rousseeuw (1987), a 13×7 matrix was constructed using the mean value of each variable for each data set (All, Dry and Wet) and dissimilarities were then calculated. The algorithm computes k representative objects, called medoids, which together determine a clustering. Each object is then placed to the cluster corresponding to the nearest medoid (Mathsoft,

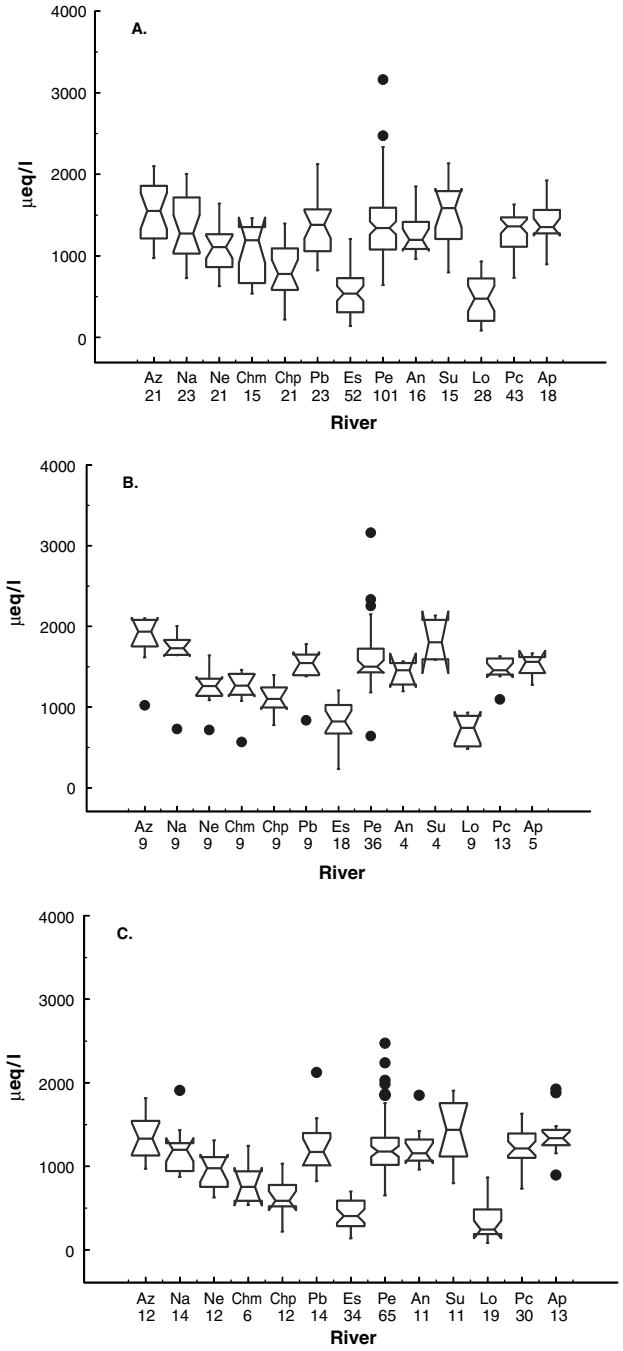

Figure 3. Boxplot comparisons of calcium (Ca^{2+}) concentrations for sample locations including: (A) All data, (B) Dry season data, and (C) Wet season data. The crossbar within the box shows the median, the length of the box reflects the fourth spread, and the fences or tails of the boxplot are marked by the extremes. In addition, the notches in the boxplots represent the 90% confidence intervals around the median. Outliers, which are shown as circles, are defined as values that are greater than 1·5 times the distance of the interquartile range away from the upper fourth or the lower fourth

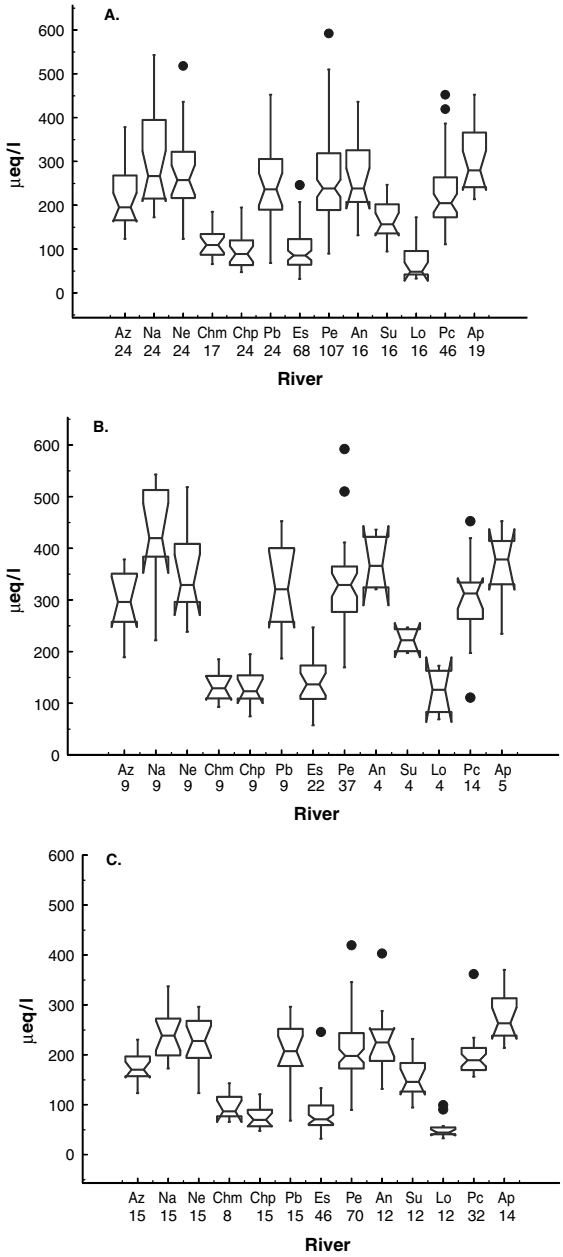

Figure 4. Boxplot comparisons of magnesium (Mg^{2+}) concentrations for sample locations including: (A) All data, (B) Dry season data and (C) Wet season data

1999a). For example, object i is placed in cluster v_i when medoid m_{vi} is nearer than any other medoid m_w:

$$d(i, m_{v_i}) \leq d(i, m_w) \text{ for all } w = 1, \ldots, k$$

where d represents distance.

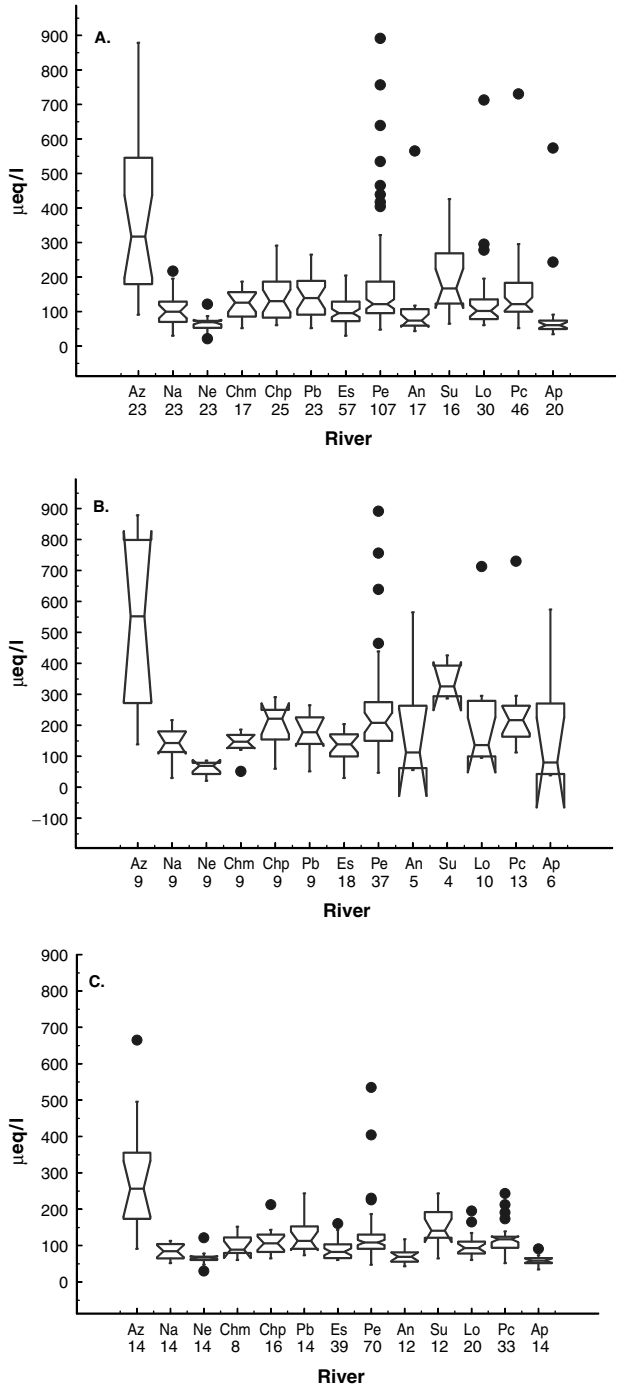

Figure 5. Boxplot comparisons of sodium (Na^+) concentrations for sample locations including: (A) All data, (B) Dry season data and (C) Wet season data

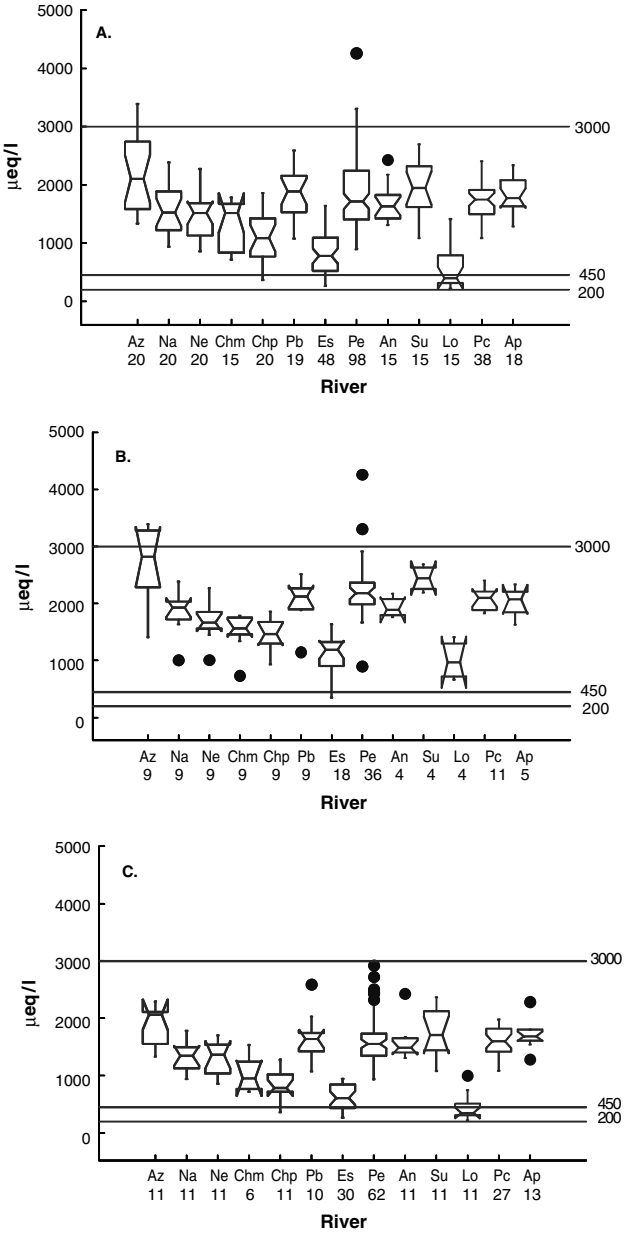

Figure 6. Boxplot comparisons of TZ+ concentrations for sample locations including: (A) All data, (B) Dry season data and (C) Wet season data

The k representative objects should minimize the sum of dissimilarities for all objects to the nearest medoid:

$$\text{objective function} = \min \sum_{i=1}^{n} d(i, m_{v_i})$$

It is difficult to determine the appropriate number of clusters for a data set. A common approach is to solve for different numbers of clusters and then select the best number based on the results of the algorithm, in

conjunction with professional judgement and common sense (Kachigan, 1986). The results of the 'partitioning around medoids' algorithm are presented in Table V. The three data sets (All, Dry and Wet) were evaluated for k values of 2, 3 and 4, which correspond to 2, 3 and 4 clusters, respectively. The 'partitioning around medoids' cluster analysis is displayed by means of a silhouette plot (Figure 7). A silhouette plot uses silhouette values for each object i, $s(i)$, which are computed and plotted as a bar of length $s(i)$ where

$s(i) \approx 1$ means the object is well classified
$s(i) \approx 0$ means the object lies between two clusters
$s(i) \approx -1$ means the object is poorly classified

The silhouette of a cluster collectively represents each object's silhouette value within that cluster, ranked in decreasing order (Figure 7). A silhouette plot shows the silhouettes of all clusters next to each other, thereby allowing a comparison of the quality of the clusters. The overall average silhouette width represents the average of the silhouette values, $s(i)$, over all objects i in the data set. Silhouette widths are a measure of how well an object lies within its cluster and a value less than 0·25 indicates that there is no substantial structure (Mathsoft, 1999a,b).

Rather than partitioning a set of objects into a given number of mutually exclusive clusters (as in the 'partitioning around medoids' technique), clusters can be formed sequentially in a hierarchical or 'nested' fashion in which smaller clusters occur within larger ones. Agglomerative hierarchical clustering is a common technique for achieving this. The function AGNES (Kaufman and Rousseeuw, 1990) was used for the 'agglomerative nesting' cluster analysis. The algorithm is based on dissimilarities only. Agglomerative nesting begins by computing a dissimilarity matrix, and the 13 objects initially each form a separate cluster. Iterations are performed that:

1. merge the two clusters with the smallest between-cluster dissimilarity;
2. compute the dissimilarity between the new cluster and all remaining clusters;
3. continue iterations until only one large cluster remains, which contains all objects.

Output from the agglomerative nesting technique is displayed in a 'clustering tree' format for the three data sets (All, Dry, Wet). A clustering tree is constructed so that the leaves of the tree (horizontal lines) represent objects, and the branches of the tree (vertical line) equal the dissimilarity between the corresponding clusters (Mathsoft, 1999a,b). For example, in the clustering tree for All data (Figure 8a), the dissimilarity between clusters Chm and Chp is between 1 and 2, whereas a greater dissimilarity (>6) exists between cluster Az and the cluster containing the remaining objects.

RESULTS AND DISCUSSION

Cation and silica chemistry

The boxplots (Figures 3–6) show a general dilution effect of the wet season that affects all solutes studied, but they also show a trend towards homogeneity: rivers with distinct hydrochemical signals in the dry season lose their individuality in the wet season. For example, the difference in the Ca^{2+} concentration in Na and Ne (Figure 3), significant in the dry season, becomes non-significant in the wet season; likewise the difference in TZ+ in An and Su (Figure 6), significant in the dry season, becomes non-significant in the wet season. Despite this trend to homogeneity in the wet season, the spatial variability of water chemistry is remarkable. Within a 10-km radius, Ca^{2+} (An and Lo, or Es and An, Figure 3B) and Mg^{2+} (Ne and Chp, or An and Su, Figure 4B) concentrations differ by 100%. In the mainstem of the Pichis River, however, this variability is not apparent: the chemical composition remains virtually constant along its channel (sampling locations Pb, Pe, Pc in Figure 3–6).

Some patterns within this variability, however, emerge from the grouped boxplots. The samples collected at Es and Lo are statistically similar (overlap of 90% confidence intervals) for most variables in the three data

Table V. Results of 'Partitioning around Medoid' cluster analysis. Members of each cluster are listed in descending order of silhouette width within that respective cluster

Partioning around medoid for $k = 2$

Data Set	Cluster 1	Cluster 1 silhouette width	Cluster 2	Cluster 2 silhouette width	Average silhouette width of total data set
All	Ap, Na, Pb, An, Pe, Su, Pc, Ne Az	0·47	Lo, Chp, Esp, Chm	0·54	0·49
Dry	Ap, An, Na, Pb, Pe, Su, Pc, Az, Ne	0·41	Lo, Chp, Esp, Chm	0·63	0·48
Wet	Na, Pb, Ap, Pe, An, Su, Pc, Ne, Az	0·48	Chp, Lo, Esp, Chm	0·58	0·51

Partioning around medoid for $k = 3$

Data Set	Cluster 1	Cluster 1 silhouette width	Cluster 2	Cluster 2 silhouette width	Cluster 3	Cluster 3 silhouette width	Average silhouette width of total data set
All	Ap, Na, Pb, An, Pe, Pc, Ne, Su	0·58	Lo, Chp, Esp, Chm	0·51	Az	0	0·51
Dry	An, Pb, Na, Ap, Pe, Pc, Ne, Su	0·44	Lo, Chp, Esp, Chm	0·61	Az	0	0·46
Wet	Na, Ap, Pe, Pb, An, Pc, Ne, Su	0·58	Chp, Lo, Esp, Chm	0·57	Az	0	0·53

Partioning around medoid for $k = 4$

Data Set	Cluster 1	Cluster 1 silhouette width	Cluster 2	Cluster 2 silhouette width	Cluster 3	Cluster 3 silhouette width	Cluster 4	Cluster 4 silhouette width	Average silhouette width of total data set
All	Ap, Na, Pb, An, Pe, Pc, Ne, Su	0·54	Chm, Chp	0·53	Az	0	Lo, Esp	0·43	0·48
Dry	Pb, Na, Ne, Pe, Pc, An, Ap	0·39	Lo, Chp, Esp, Chm	0·6	Az	0	Su	0	0·4
Wet	Ap, Ne, An, Na	0·29	Chp, Lo, Esp, Chm	0·55	Az	0	Pe, Su, Pb, Pc	0·25	0·33

sets (All, Wet and Dry). The strongest correlation between these two rivers occurred in the Dry data set, in which the Lo and Es samples were statistically similar for all variables. The dry season contains the smallest number of samples, however. Boxplots also show that samples collected at Es, Lo, Chp and Chm had the four lowest concentrations/values for Ca^{2+}, CE, K^{+}, Mg^{2+}, pH and TZ+ and had the four highest concentrations of SiO_2. Samples from Az had the highest concentration/value of Ca^{2+}, CE, Na^{+}, pH and TZ+.

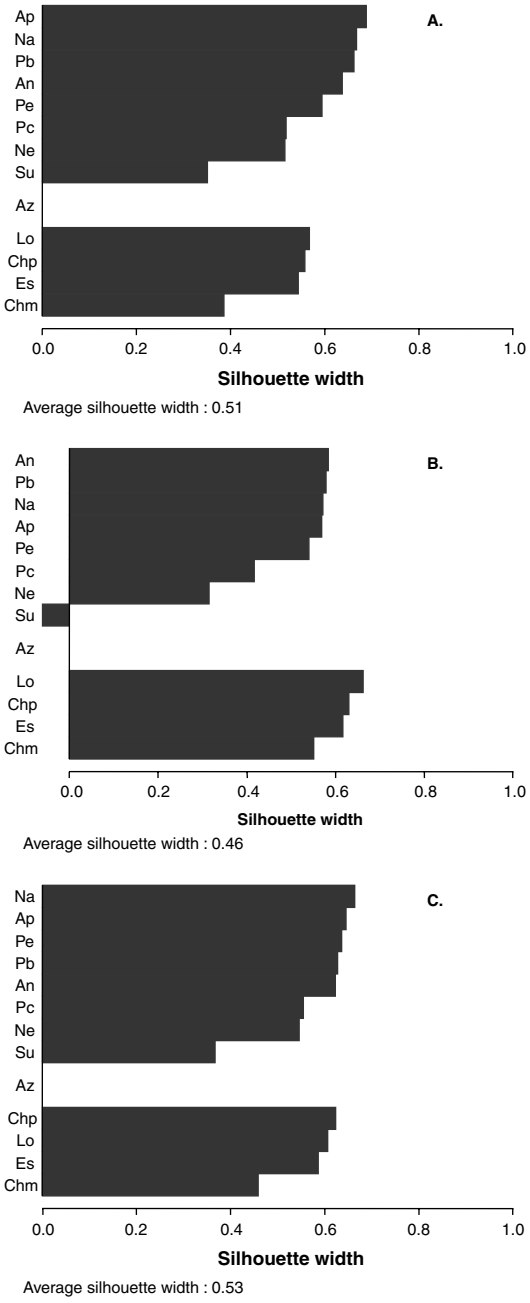

Figure 7. Silhouette plots with $k = 3$ for (A) All data, (B) Dry season data and (C) Wet season data

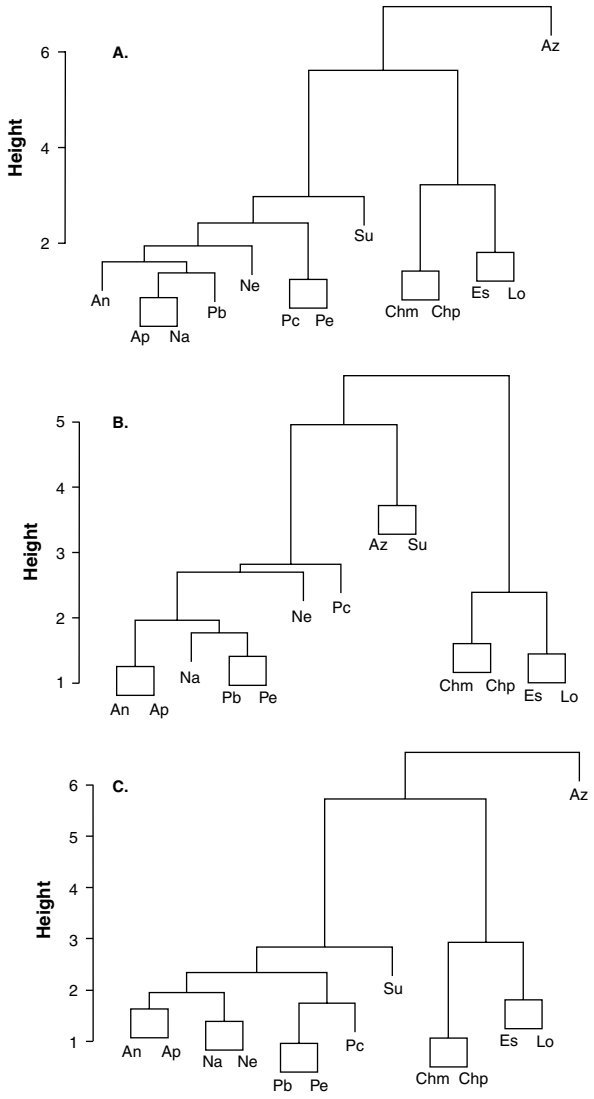

Figure 8. Agglomerative nesting cluster analysis plots for (A) All data, (B) Dry season data and (C) Wet season data

Cluster analysis suggests that the sampling locations can be segregated into three groups based on their similarities. The largest silhouette widths occur for a k value of 3 for two of the data sets (All and Wet) and the k value of 3 for the Dry data set is only slightly less than for a k value of 2 (Table V). The PAM algorithm generated three clusters with the following members: (Cluster 1) An, Ap, Na, Ne, Pb, Pc, Pe and Su, (Cluster 2) Chm, Chp, Esp and Lo, and (Cluster 3) Az (Figure 7). The formation of a separate cluster for Az and a cluster containing Chm, Chp, Es and Lo is consistent with our observations of the boxplots. Therefore, we believe that a k value of 3 is appropriate for the 'partitioning around medoids' cluster analysis.

For agglomerative nesting cluster analysis, clustering trees for all three data sets show the clustering of Chm, Chp, Esp and Lo with each other before they are merged with the remaining objects, indicating that dissimilarities are relatively small between these objects. A second trend for the All and Wet data sets shows that object Az merges with the remaining objects in the final step, indicating maximum dissimilarity between

this object and the other objects. The Dry data set, however, shows a cluster between Az and Su that merges with a larger cluster containing seven locations and finally merges with the cluster containing Chm, Chp, Esp and Lo (Figure 8). The remaining objects are grouped together in various clusters for the three data sets.

River chemistry as a function of lithology

Lithology has long been known to exert considerable influence on river chemistry at a variety of scales (see e.g. Walling and Webb, 1986). Within the Pichis basin, the river chemistry of Lo is characterized by a relatively low concentration of base cations and this was the only site that had an acidic median pH value for two of the data sets (All and Wet). The relatively low concentrations of cations and consequent low pH can be attributed to the largest percentage area of cation-poor Tertiary and Quaternary formations within its drainage basin (Table IV). The similar river chemistry between Lo, Es and Chp probably is a reflection of the largest percentage of cation-poor areas (75·2%, 58·1% and 44·4%, respectively) within their drainage basins (Table 4). The relatively cation-poor river chemistry at Chm is enigmatic. This location occurs at the base of the mountains and reflects drainage of the calcareous and locally sodium-rich Jurassic Sarayaquillo Formation. It is difficult to explain why there is not a strong Ca^{2+} or Na^{+} geochemical signal reflected in the samples collected from this location, especially because 79·8% of the drainage basin consists of cation-rich lithological units. It is interesting to note that the sample locations comprising Cluster 1 (Lo, Es, Chp and Chm) contain four of the five smallest drainage basins (Table III) and these basins are spatially distributed on the western side of the Pichis Valley.

The Azupizu River (Az) is characterized by high concentrations of Ca^{2+}, Na^{+} and by a high electrical conductivity, which we attribute to the calcareous Cretaceous Chonta and Agua Caliente formations and the calcareous and sodium-rich Jurassic Sarayaquillo Formation. Although these formations account for only 15·7% of its drainage area (Table IV) owing to the spatial limitations of the geological survey, structural considerations suggest that they extend considerably into the headwater of the Azupizu River. Because the Az sampling location is greater than 15 km downstream of the Ca^{2+}- and Na^{+}-rich lithological units within its sub-basin (the presumed source area for the strong signal), an even stronger geochemical signal most probably occurs further upstream (Figure 1).

We attribute the strong Ca^{2+} and Na^{+} signal in the Sungaro River (Figures 3 and 5) to the drainage from the Ca^{2+}- and Na^{+}-rich Tertiary (Chambira and Huayabamba formations) and Cretaceous (Cachiyacu and Chonta formations) lithologies, which collectively comprise 28·7% of its drainage area. The geochemical signals are not as strong as those at Az, despite the distance of the presumed source area lithologies being much shorter (6–12 km) for the Su location than those for the Az location (Figure 2). The chemistry of the remaining eight sample locations does not exhibit a characteristic geochemical signal that allows for differentiation between locations.

Regional context

Stallard and Edmond (1983) developed a TZ+-based continental-scale classification scheme for the Amazon River system, which we applied, to the mesoscale Pichis basin; the respective boxplots are shown in Figure 6. The TZ+ boxplot for the All and Wet data sets show that the median and 90% confidence interval of every location except Lo fall within the boundaries of Group 3, as defined by Stallard and Edmond (1983). The Lo locality has a median and lower confidence limit within the Group 2 boundaries, but the upper confidence limit occurs within Group 3. The TZ+ boxplot for the Dry data set shows that the median values and confidence intervals for all sample locations except Az occur within the Group 3 boundaries. Although the median and lower confidence limit occur in Group 2, the upper 90% confidence limit of Az extends into the Group 4 region.

Based on a review of the boxplots for the three data sets and the geology of the area, we believe that Lo satisfies the criteria to be placed in Group 2. In addition to the TZ+values, the Lo location also is relatively enriched in silica compared with the other locations (silica boxplots not shown). The other sample locations satisfy the conditions for Group 3. It is possible that the river chemistry further upstream of Az satisfies the

criteria for a Group 4 classification, particularly because the Jurassic Sarayaquillo Formation is known for local salt deposits.

Table VI compares the river chemistry from different areas of the Amazon with that of the study area. The river chemistry of the Pichis at Cahuapanas (Pc) was chosen as being representative for the subject area, because it is the most downstream sample location (Figure 2). Table VI shows that the river chemistry of our study area is relatively enriched in K^+, Mg^{2+} and Ca^{2+} and, consequently, has a higher TZ+ value, although being relatively depleted in silica. The concentrations of K^+ and Mg^{2+} were very similar to concentrations collected from the Upper Madeira basin in Bolivia (Seyler *et al.*, 1998). We believe the regional variability of river chemistry is a function of the variable geology between these locations.

Implications for deforestation studies at the mesoscale

The Pichis basin exhibits considerable variability in lithological signals between its tributaries at the sub-basin scale, a variability that occurs even in the general absence of anthropogenic influences such as deforestation. For watersheds and river basins in Rondonia, Brazil, Biggs *et al.* (Accepted) found that average solute concentrations in areas of deforestation were higher than predicted forested background levels. Table VII lists the means for the concentration partitions of major cations for two deforestation classes studied by Biggs

Table VI. Comparison of average composition of dissolved materials from study area with values from the Amazon mainstem and estimated world averages

Element	World average[a]	Upper Amazon[b]	Lower Amazon[c]	Upper Madeira basin, Bolivia[d]	Study Area[e]	Units
Sodium	223	278	63·4	87	153·8	µmol/L
Potassium	33	28·9	20·8	34	34·9	µmol/L
Magnesium	130	96·7	39·8	119	113·6	µmol/L
Calcium	335	477	129	182	655	µmol/L
Silica	173	185	120	NM	76·6	µmol/L
TZ+	1186	1454·3	421·8	723	1725·9	µeq/L

[a] Table 2 of Stallard and Edmond (1983).
[b] Stallard and Edmond (1983, S250, Upper Amazon at Iquitos).
[c] Stallard and Edmond (1983, S209, Lower Amazon at Obidos).
[d] Seyler *et al.* (1998, location 26, Madeira River, Bolivia).
[e] This study, mean values from Pichis River at Cahuapanas.

Table VII. Comparison of Pichis basin lithology signals with estimated deforestation signals

	15–40% Deforestation class				40%+ Deforestation class			
	Lower limit[a]	Exceedance (%)[b]	Mean[c]	Exceedance (%)[b]	Lower Limit[a]	Exceedance (%)[b]	Mean[c]	Exceedance (%)[b]
Ca^{2+}	0·8	65	1·2	38	1·0	50	2·9	0
K^+	0·4	97	1·4	42	1·0	49	2·5	0
Mg^{2+}	0·9	54	1·7	24	2·2	19	4·5	0
Na^+	0·8	60	1·1	44	0·9	56	2·5	17

[a] Represents lower 95% confidence limit for deforestation class defined by Biggs *et al.* (Accepted), computed as ratio of observed concentration to predicted background concentration.
[b] Percentage of ratios (mean concentration of a solute in our sub-basin over that of another sub-basin), out of a possible 72 comparisons, that exceed the lower 95% confidence level and mean, respectively, in a deforestation class.
[c] Represents mean value for deforestation class defined by Biggs *et al.* (Accepted), computed as ratio of observed concentration to predicted background concentration.

et al. (Accepted), as well as the lower 95% confidence limit. For the watersheds and river basins in Rondonia, the observed dry season calcium signal in the 15–40% deforestation class is 1·2 times higher than the predicted background concentration, although the lower 95% confidence limit is only 0·8 (Table VII). For the 40%+ deforestation class, the observed dry season calcium signal is 2·9 times greater than the predicted background signal, although the lower 95% confidence limit is 1·0.

In Table VII, we also compare the lithology-derived variability of river chemistry ('lithology signal' from hereon) in the Pichis basin with the deforestation signals estimated for some basins in Rondonia (Biggs *et al*., Accepted) to see if these deforestation signals exceed, or are masked by, the lithology signal. We define this lithology signal by all possible ratios of the mean concentration of a given solute in one sub-basin over that of another sub-basin (example: mean Ca^{2+} concentration in Sungaro over mean Ca^{2+} concentration in Esperanza). The reason for this approach is that any comparison of 'deforested versus undisturbed' usually involves two watersheds, i.e. one impacted and the other unimpacted. In our hypothetical example, therefore, 72 such comparisons are possible for the nine tributary sub-basins (Az, Na, Ne, Chp, Es, An, Su, Lo and Ap, see Figure 1 for location).

Table VII shows that this 'lithology signal' within the Pichis basin exceeds the lower 95% limit of some deforestation signals, even in the >40% category. The entry for K^+, for example, tells us that in the 15–40% category the deforestation signal is not significantly different from the 'lithology signal' because virtually all (97%) sub-basin ratios as defined above exceed the lower 95% confidence limit for the mean deforestation signal. Even in the 40%+ category, the 'lithology signal' masks the deforestation signal in nearly half (49%) the cases, i.e. the 'lithology signal' as defined above is not significantly different from the deforestation signal.

In view of the recent interest in relating land-use change in the Amazon basin to changes in streamwater chemistry beyond the small-catchment scale, our results suggest that lithological variability should receive due consideration even at the submesoscale (<10 000 km^2) in any comparison of more and less deforested watersheds. Uniformity of background water chemistry should not simply be assumed even in adjacent basins, and the magnitude of the deforestation signals observed by Biggs *et al*. (Accepted) is similar to the lithology-derived spatial variability of river chemistry observed in this study.

CONCLUSIONS

The tributaries of the Pichis River in the Subandean Zone of Peru show distinctive chemical signals. The study area represents a drainage basin that was relatively unaffected by human influence. Therefore, the variable river chemistry is primarily a function of the variable lithology within the Pichis basin drainage. We conclude that mesoscale river chemistry studies in the Amazon Basin, aimed at detecting the effects of deforestation, should take into account the hydrochemical variability resulting from lithological diversity. This lithology-controlled variability is likely to mask deforestation signals derived from the comparison of river basins under the assumption of similar lithology.

ACKNOWLEDGEMENTS

J. Sobieraj was in part supported by NSF Grant No. 9554527. We would like to thank Roxanna Ayllon for her assistance with the GIS portion of the project.

REFERENCES

Biggs TW, Dunne T, Domingues TF, Martinelli LA. Accepted. The relative influence of natural watershed properties and human disturbance on stream solute concentrations in the southwestern Brazilian Amazon. *Water Resources Research*.

Edmond JM, Palmer MR, Measures CI, Brown ET, Huh J. 1996. Fluvial geochemistry of the eastern slope of the northeastern Andes and its foredeep in the drainage of the Orinoco in Colombia and Venezuela. *Geochimica Cosmochim Acta* **60**: 2949–2976.

Emerson JD, Strenio J. 1983. Boxplots and batch analysis. In *Understanding Robust and Exploratory Data Analysis*. Hoaglin DC, Mosteller F, Tukey JW (eds). Wiley: New York; 58–96.

Kachigan SK. 1986. *Statistical Analysis: an Interdisciplinary Introduction to Univariate and Multivariate Methods*. Radius Press: New York; 589.

Kalliola R, Puhakka M, Jukka S, Hanna T, Kalle R. 1991. The dynamics, distribution and classification of swamp vegetation in Peruvian Amazonia. *Annales Botaniii Fennici* **28**: 225–239.

Kaufman L, Rousseeuw PJ. 1990. *Finding Groups in Data: an Introduction to Cluster Analysis*. Wiley: New York; 349.

Mathsoft. 1999a. *S-Plus 2000 Guide to Statistic*, vol. 2. Data Analysis Products Division: Seattle, WA.

Mathsoft. 1999b. *S-Plus 2000 Modern Statistics and Advanced Graphics: User's Guide*. Data Analysis Products Division: Seattle, WA.

Mortatti J, Probst DL, Ferreira JR. 1992. Hydrological and geochemical characteristics of the Jamari and Jiparana river basins (Rondônia, Brazil). *GeoJournal* **26**: 287–296.

ONERN. 1970. Mapa de Grandes Grupos de Suelos, 1 : 400,000. *In Inventorio y Evaluacion Semidetallada de los Recursos Naturales de la Zona del Rio Pichis (Proyecto Pichis-Palcazu).* (October 1981). Oficina Nacional de Evaluacion de Recursos Naturales; Republica del Peru.

ONERN. 1981a. *Inventario y Evaluacion Semidetallada de los Recursos Naturales de la Zona del Rio Pichis (Proyecto Pichis-Palcazu).* Oficina Nacional de Evaluacion de Recursos Naturales; Republica del Peru; 1–360.

ONERN. 1981b. Mapa Geologico, 1 : 120,000. *In Inventario y Evaluacion Semidetallada de los Recursos Naturales de la Zona del Rio Pichis (Proyecto Pichis-Palcazu).* Oficina Nacional de Evaluacion de Recursos Naturales; Republica del Peru.

Räsanen ME, Salo JS, Jungnert H, Pittman RL. 1990. Evolution of the Western Amazon lowland relief: impact of Andean foreland dynamics. *Terra Nova* **2**: 320–332.

Räsanen M, Neller R, Salo J, Jungnert H. 1992. Recent and ancient fluvial deposition systems in the Amazonian foreland basin, Peru. *Geology Magazine* **129**: 293–306.

Reindl J. 1903. *Die schwarzen Flüsse Südamerikas*. Theodor Ackermann: München; 138 pp.

Rousseeuw PJ. 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics* **20**: 53–65.

Seyler P, Guyot JL, Maurice-Bourgon L, Sondag F, Elbaz-Poulichet F, Etcheber H, Quintanilla J. 1998. Origin of trace elements in the Bolivian Amazonian drainage basin. In *Hydrology in the Humid Tropic Environment*, Fernández-Jaúregui CA (ed.). IAHS 253, International Association of Hydrological Sciences: Wallingford; 47–58.

Sioli H. 1964. General features of the limnology of Amazonia. *Verhandlungen der Internationalen Vereinigung fuer Theoretische und Angewandte Limnologie* **15**: 1053–1058.

Sioli H. 1965. Bemerkungen zur Typologie amazonischer Flüsse. *Amazoniana* **1**: 74–83.

Sioli H. 1984. The Amazon and its main affluents: hydrography, morphology of the river courses, and river types. In *The Amazon. Limnology and Landscape Ecology of a Mighty Tropical River and its Basin*, Sioli H (ed.). W. Junk: Dordrecht; 126–165.

Stallard RF, Edmond JM. 1983. Geochemistry of the Amazon: the influence of geology and weathering environment on the dissolved load. *Journal of Geophysical Research* **88**: 9671–9688.

Suarez G, Molnar P, Burchfield BC. 1983. Seismicity, fault plane solutions, depth of faulting, and active tectonics of the Andes of Peru, Ecuador, and southern Colombia. *Journal of Geophysical Research* **88**: 403–428.

Vital H, Stattegger K. 2000. Major and trace elements of stream sediments from the lowermost Amazon River. *Chemical Geology* **168**: 151–168.

Walling DE, Webb BW. 1986. Solutes in river systems. In *Solute Processes*, Trudgill ST (ed.). Wiley: New York; 251–316.

