

1
2 Discharge-calcium concentration relationships in streams of the Amazon and Cerrado of Brazil:
3 Soil or land use controlled.

4

5

6

7 Daniel Markewitz^{1,*}, E. Conrad Lamon III², Mercedes MC Bustamante³, Joaquin Chaves⁴, Ricardo O.
8 Figueiredo⁵, Mark S. Johnson⁶, Alex Krusche⁷, Christopher Neill⁴, and José S.O. Silva³

9

10 ¹Warnell School of Forestry and Natural Resources, The University of Georgia, Athens, GA
11 30602, USA

12 ²Statistical Ecology Associates LLC, Canyon Lake, TX 78133, USA

13 ³Department of Ecology, University of Brasilia, Brasilia, DF, 70910, Brazil

14 ⁴Ecosystems Center, Marine Biological Laboratory, Woods Hole, MA 02543, USA

15 ⁵EMBRAPA Amazônia Oriental, Belém, PA, 66095, Brazil.

16 ⁶Institute for Resources, Environment and Sustainability and Department of Earth and Ocean
17 Sciences, University of British Columbia, Vancouver, British Columbia, V6T 1Z4, Canada

18 ⁷Laboratório de Ecologia Isotópica, Centro de Energia Nuclear na Agricultura, University of São
19 Paulo, Piracicaba, SP, 1341 6, Brazil

20

21

22 *Corresponding author: dmarke@warnell.uga.edu Phone:706-542-0133 Fax:706-542-8356

23 **Abstract**

24 Stream discharge-concentration relationships are indicators of terrestrial ecosystem
25 function. Throughout the Amazon and Cerrado regions of Brazil rapid changes in land use and
26 land cover may be altering these hydrochemical relationships. The current analysis focuses on
27 factors controlling the discharge-calcium (Ca) concentration relationship since previous research
28 in these regions has demonstrated both positive and negative slopes in linear \log_{10} discharge-
29 \log_{10} Ca concentration regressions. The objective of the current study was to evaluate factors
30 controlling stream discharge-Ca concentration relationships including year, season, stream order,
31 vegetation cover, land use, and soil classification. It was hypothesized that land use and soil
32 class are the most critical attributes controlling discharge-Ca concentration relationships. A
33 multilevel, linear regression approach was utilized with data from 28 streams throughout Brazil.
34 These streams come from three distinct regions and varied broadly in watershed size (<1 to >10⁶
35 ha) and discharge (10^{-5.7} to 10^{3.2} m³ sec⁻¹). Linear regressions of \log_{10} Ca versus \log_{10} discharge in
36 13 streams have a preponderance of negative slopes with only two streams having significant
37 positive slopes. An ANOVA decomposition suggests the effect of discharge on Ca concentration
38 is large but variable. Vegetation cover, which incorporates aspects of land use, explains the
39 largest proportion of the variance in the effect of discharge on Ca followed by season and year.
40 In contrast, stream order, land use, and soil class explain most of the variation in stream Ca
41 concentration. In the current data set, soil class, which is related to lithology, has an important
42 effect on Ca concentration but land use, likely through its effect on runoff concentration and
43 hydrology, has a greater effect on discharge-concentration relationships.

44

45

46 **Introduction**

47 Streamwater discharge-concentration relationships are indicators of terrestrial ecosystem
48 function (Bond, 1979). The slope of the discharge-concentration relationship, whether positive
49 or negative, has been used to infer the sources and flowpaths of dissolved constituents to streams
50 (Saunders and Lewis, 1989). Source waters that travel long flowpaths such as groundwaters and
51 interact with primary minerals in bedrock tend to contribute high concentrations of the rock
52 derived elements (e.g., Ca^{+2} , Mg^{+2} , and Si) during low flow (Drever, 1997). In contrast, source
53 waters that are quickly transported to streams during runoff events may be dilute in the rock
54 derived elements but rich in organic carbon or nitrogen due to interaction with the soil O horizon
55 (Hornberger et al., 1994). In this case, organic C and N may have a positive discharge-
56 concentration relationship, at least during the earlier stages of storm runoff, while the rock
57 derived elements present a negative discharge-concentration relationship as groundwaters are
58 diluted by surface waters (Lewis and Grant, 1979). Empirical studies commonly observe
59 negative discharge-concentration relationships for the rock derived elements with positive
60 relationships being atypical (Meyer et al., 1988).

61 The current analysis focuses specifically on discharge-Ca concentration relationships in
62 the Amazon and Cerrado of Brazil since previous research in a watershed on highly weathered
63 soil, which is common in both regions, demonstrated a positive discharge-Ca concentration
64 relationship (Markewitz et al., 2001). Positive slopes in Ca-discharge concentration relationships
65 were reported by Meyer et al. (1988) but no mechanism was identified. In the Amazonian
66 watershed where a positive slope in Ca-discharge was observed, two competing hypotheses were
67 proposed: 1) it is possible that these positive relationships could result where soils and
68 underlying parent material have become so depleted of Ca that surface runoff concentrations

69 exceed groundwater concentrations or 2) land use conversion through slash-and-burn practices
70 can so enrich surface soils in Ca that surface runoff concentrations exceed groundwater
71 concentrations (Markewitz et al., 2001). Significant differences in stream water Ca
72 concentrations (as well as other cations) have been demonstrated to vary with lithology in the
73 Amazon Basin but effects on discharge-concentration relationships has not been thoroughly
74 investigated (Stallard and Edmond, 1983). The prevalence of positive slopes in discharge-Ca
75 concentration relationships in the Amazon and Cerrado is unknown and whether these slopes
76 result from differences in lithology and soil type or from land use conversion remains uncertain.

77 Throughout the Amazon and Cerrado regions of Brazil rapid changes in land use and land
78 cover (INPE, 2006) are altering the hydrological (Moraes et al., 2006; Williams and Melack,
79 1997) and hydrochemical (Germer et al., 2009; Neill et al., 2001) relationships in these streams
80 and possibly altering the expected discharge-concentration relationships in these water bodies.
81 As the landscape of Brazil continues to be altered in the coming decades it will be important to
82 understand regional differences in stream water chemistry (Richey et al., 1990; Stallard, 1985)
83 and differences in processes of land-water coupling (Biggs et al., 2002). Regulatory agencies in
84 Brazil will be tasked with assessing changes in water quality with continued land use conversion
85 and will need to be able to interpret concentration differences with lithology, season, or flow
86 from those changes due to human alterations.

87 The objective of the current study is to evaluate slopes (+/-) of discharge-calcium
88 concentration relationships for previously studied streams and evaluate the influence of year,
89 season, stream order, vegetation cover, land use, and soil classification on the regression
90 relationship. A multilevel linear regression approach is utilized.

91

92 **Methods**

93 Data from 28 different streams with 51 total sampling stations (i.e., >1 sampling
94 station/stream) were utilized in this analysis (Table 1). These streams are situated in eight
95 different locations and three distinct regions (Figure 1). Site descriptions and specific details of
96 stream water sampling and analysis within each watershed are available in references provided in
97 Table 1. At all sites investigators identified current land use and existing soil types. In many
98 cases stream waters were collected as grab samples on a weekly or biweekly basis, while at
99 Rancho Grande an automated ISCO sampler was utilized. A number of sites also had automated
100 stage height recorders while others recorded stage height during collections. In all cases waters
101 were filtered prior to analysis and all sites used ion chromatography for Ca analysis. Stream Ca
102 concentration data were available for all sampling stations while discharge was measured in 18
103 of the streams at 28 sampling stations. Sampling stretched over 12 yrs (1994, 1996-2007) and all
104 months of the year (i.e., season).

105 Stream order and land use were taken from site descriptions. Land use was comprised of
106 seven total categories; four within lowland moist tropical forest and three within Cerrado
107 savannah. Within these two land use classes some watersheds were nearly 100% natural
108 vegetation (broadleaf forest (Forest) or Cerrado scrub savannah (Cerrado)) while many others
109 possessed some natural vegetation (34-70% primary or secondary forest or 12-50% Cerrado)
110 mixed with pastures (19-46%) and agricultural (5-50%) land uses (Fmixed or Cmixed). Some
111 lowland forest watersheds in the Amazon had been nearly 100% converted to pasture (Pasture).
112 Finally, if forested or Cerrado watersheds in either location possessed substantial urban
113 development they were classified as Furban (1-2%) or Curban (6-27%).

114 Vegetation Cover of each watershed was characterized based on the 1988 Mapa de
115 Vegetação do Brasil at a 1:5,000,000 scale (<http://na.unep.net/datasets/datalist.php>). Soil
116 classification was similarly obtained from the 1981 Mapa de Solos do Brasil at a 1:5,000,000
117 scale. Given the available map scales each watershed and thus all the sampling stations were
118 within a single class. Furthermore, all vegetation cover and soil class designations were generally
119 consistent with site specific descriptions.

120 To analyze individual station regressions where there was sufficient data, simple linear
121 least square regression was utilized on the $\log_{10}\text{Ca}$ (in μM) - $\log_{10}\text{Q}$ (in $\text{m}^3 \text{sec}^{-1}$) relationship. To
122 analyze the data from all stations simultaneously, a multilevel modeling approach (Congdon,
123 2001) was utilized to estimate a linear model for prediction of $\log_{10}\text{Ca}$. The main predictor
124 variable was discharge or $\log_{10}\text{Q}$, which was centered by subtracting the mean of the $\log_{10}\text{Q}$ and
125 dividing by the range. If discharge was recorded as zero ($n=56$) discharge was considered a
126 missing value.

127 In the Bayesian multilevel modeling approach, which is nearly identical mathematically
128 to the classical random effect model (Clayton, 1996), adjustments to the regression relationship
129 between the dependent variable $\log_{10}\text{Ca}$ and the independent variable $\log_{10}\text{Q}$ are incorporated for
130 covariates at all levels, including observation and higher level groups (i.e., stream order, soil
131 class, etc). This approach allows for the simultaneous accounting of contextual and individual
132 variability in the outcome (Congdon, 2001). Adjustments to the linear regression parameters β_0
133 (the intercept) and β_1 (the slope) were estimated at all levels. In contrast, a multivariate
134 regression using a completely pooled regression model would use each factor as a separate
135 predictor but would have little chance of satisfactory results using data from such a large region.
136 Implicit in using a pooled model would be an assumption that a single slope and intercept could

137 describe the relationship everywhere. Since there is evidence to the contrary, the multilevel
 138 approach utilized allows for some variability in parameters, based on the chosen factors.

139 In the current analysis, year, season, stream order, land cover, land use, and soil class
 140 were the factors, and each factor had multiple levels (e.g., season has 12 monthly levels). As
 141 such, the observation model for $\log_{10}\text{Ca}$ was

$$\log_{10}(\text{Ca conc. } \mu\text{M }_i) \sim N(\mu_i, \tau_l), \quad (1)$$

143 where τ_l is the error precision, and $\tau_l = 1/\sigma_l^2$. A uniform prior was used on σ_l (Gelman,
 144 2005b). The mean of the normal distribution for observations i (μ_i) was given by a linear
 145 regression which specifies the mean, conditional on the covariate $\log_{10}\text{Q}$ such that:

$$\mu_i = \beta_0 + \beta_1 * \log_{10}\text{Q}_i, \quad (2)$$

147 where,

$$\beta_0 = \mu_{\beta_0} + \beta_{0\text{year}_j} + \beta_{0\text{seas}_k} + \beta_{0\text{order}_l} + \beta_{0\text{cov}_m} + \beta_{0\text{use}_n} + \beta_{0\text{soil}_o} \quad (3)$$

$$\beta_1 = \mu_{\beta_1} + \beta_{1\text{year}_j} + \beta_{1\text{seas}_k} + \beta_{1\text{order}_l} + \beta_{1\text{cov}_m} + \beta_{1\text{use}_n} + \beta_{1\text{soil}_o} \quad (4)$$

150 and:

$$\mu_i = (\mu_{\beta_0} + \beta_{0\text{year}_j} + \beta_{0\text{seas}_k} + \beta_{0\text{order}_l} + \beta_{0\text{cov}_m} + \beta_{0\text{use}_n} + \beta_{0\text{soil}_o}) + \\ (\mu_{\beta_1} + \beta_{1\text{year}_j} + \beta_{1\text{seas}_k} + \beta_{1\text{order}_l} + \beta_{1\text{cov}_m} + \beta_{1\text{use}_n} + \beta_{1\text{soil}_o}) * \log_{10}\text{Q}_i \quad (5)$$

153 and $j = 1, \dots, 12$ years, $k = 1, \dots, 12$ seasons (months), $l = 1, \dots, 5$ stream orders, $m = 1, \dots, 7$
 154 vegetation covers, $n = 1, \dots, 7$ land uses, and $o = 1, \dots, 7$ soil classes. In the multilevel model of
 155 equation 5, μ_{β_0} is an overall mean intercept term, while $\beta_{0\text{year}_j}$, $\beta_{0\text{seas}_k}$, $\beta_{0\text{order}_l}$, $\beta_{0\text{cov}_m}$, $\beta_{0\text{use}_n}$,
 156 and $\beta_{0\text{soil}_o}$ are adjustments to this overall intercept due to the six factors year, season, order,
 157 cover, use, and soil, respectively. Similarly, μ_{β_1} in equation 5 is an overall mean slope for the
 158 $\log_{10}\text{Q}$ term, while $\beta_{1\text{year}_j}$, $\beta_{1\text{seas}_k}$, $\beta_{1\text{order}_l}$, $\beta_{1\text{cov}_m}$, $\beta_{1\text{use}_n}$, and $\beta_{1\text{soil}_o}$ are additive adjustments
 159 to this overall mean according to the same six factors, respectively. The sample size for each

160 level of a factor can vary and will influence the uncertainty within the parameter estimates.
161 Similarly, the matrix of all combinations of all factors may not be fully represented within the
162 observational data.

163 A non-informative, proper prior distribution was utilized for the regression coefficients,
164 such that each coefficient was assumed to have a normal distribution, with a separate mean
165 (μ) and precision ($\tau = 1/\sigma^2$). The use of a normal distribution for the regression coefficients
166 stems from the usual assumptions made regarding regression residuals. Regression coefficients
167 of a linear model are linear functions of the residuals, and if we assume the residuals are normal
168 *iid*, then so are the regression coefficients. Again, a uniform prior on each σ (in units of $\log_{10}\text{Ca}$
169 concentration) was used (Gelman, 2005a), such that $\sigma \sim U(0,100)$, and an initial value of 0 was
170 used for μ .

171 The model was estimated using a Markov Chain Monte Carlo (MCMC) simulation
172 following Lamon and Qian (2008). MCMC is a simulation technique for solving high
173 dimensional probability distribution problems. The basic idea of MCMC is to find a numeric
174 algorithm to make probabilistic inference on random variables with algebraically intractable
175 probability distributions. The Bayesian Analysis Using Gibbs Sampler (BUGS) project
176 distributes and supports flexible software for the Bayesian analysis of complex statistical models
177 using MCMC methods ([http://www.mrc-bsu.cam.ac.uk/bugs/ welcome.shtml](http://www.mrc-bsu.cam.ac.uk/bugs/welcome.shtml)), and winBUGS is
178 for use on PC platforms (Spiegelhalter et al., 2003). The model was initiated by sampling from
179 the prior distributions for each estimated coefficient and distributions were updated based on the
180 log-likelihood estimations for the observed and predicted values. As presented here, a posterior
181 distribution of all model coefficients was obtained after 100,000 iterations.

182 To evaluate parsimony, the six factor adjustments were compared to other five, four, and
 183 three factor adjustment models (e.g., without season or soil, etc.). The deviance information
 184 criterion (DIC) is a hierarchical modeling generalization of the Akaike information criterion
 185 (AIC) and the Bayesian information criterion (BIC). It is particularly useful in Bayesian model
 186 selection problems where the posterior distributions of the models have been obtained by
 187 MCMC simulation, as was done here (Spiegelhalter et al., 2002).

188 The deviance information criterion was calculated as

$$189 \quad DIC = p_D + \bar{D} \quad (6)$$

190 The deviance D is a measure of model fit analogous to a residual standard deviation. It is
 191 estimated by the log-likelihood after each iteration and is defined as

$$192 \quad D(\theta) = -2 \log(p(y|\theta)) + C \quad (7)$$

193 where y are the data, θ are the unknown parameters of the model including β , σ , and τ , and
 194 $p(y|\theta)$ is the likelihood function. C is a constant that cancels out in comparison of different
 195 models. The expectation of D

$$196 \quad \bar{D} = E^\theta [D(\theta)] \quad (8)$$

197 is an average of the log-likelihoods and is a measure of how well the model fits the data; the
 198 larger this value the worse is the fit. The effective number of parameters of the model was
 199 computed as

$$200 \quad p_D = \bar{D} - D(\bar{\theta}) \quad (9)$$

201 where $\bar{\theta}$ is the expectation of θ . This is a measure of model complexity that is particularly useful
 202 in hierarchical models where the number of independent parameters may be difficult to
 203 determine. A larger p_D implies that more parameters are being used in the model and thus the
 204 model is better able to fit the data.

205 The idea is that models with smaller DIC should be preferred to models with larger DIC.
206 Models were evaluated both by the value of D, which favors good fit, but also by model
207 complexity, as measured here by the effective number of parameters p_D . Since D will tend to
208 decrease as the number of parameters in a model increases, the p_D term compensates for this
209 effect by favoring models with a smaller number of parameters.

210

211 **Results**

212 *Data Distribution*

213 Across the dataset (n=3155) $\log_{10}\text{Ca}$ in μM ranged over two orders of magnitude with a
214 mean of 1.32 (Table 2) and discharge ($\text{m}^3 \text{sec}^{-1}$) ranged more broadly covering five orders of
215 magnitude with a mean $\log_{10}Q$ of -1.70 (Table 2). The data covered 1994 to 2007 with 1994 and
216 2007 having fewer samples and 2005 the most (Table 3). All months of the year were well
217 represented and there were five stream orders in the dataset (1, 2, 3, 5, and 6) with the majority
218 of data points from 1st or 2nd order streams (Table 3). Urupá and Ji-Paraná@Cacoal are the 5th
219 and 6th order streams, respectively.

220 There were seven vegetation covers identified from land cover maps with a majority of
221 samples from dense tropical forest with secondary forest and agricultural activities. This land
222 cover class D included all the Paragominas and Igarapé-Açu samples. Land use as identified by
223 researchers working within each site (see references in Table 1) was also comprised of seven
224 classes with forest watersheds under mixed land use being in greatest abundance, which included
225 many of the same samples identified above under dense tropical forest with secondary forest and
226 agricultural activities. Samples classified under Cerrado land uses comprised 17% of the dataset.

227 Finally, there were seven soil types classified in the watersheds with the largest number
228 of sample points represented by Latossolos amarelos distrófico which were predominant in all
229 the Paragominas streams and Juruena B1 (Table 3). Argissolos vermelho-amarelos eutróficos
230 were next most common being present in both Rancho Grande and Juruena B2. Latossolos
231 vermelho escuro represented most of the Cerrado samples. Two other soil orders were also
232 present with Cambissolos identified in two Cerrado watersheds (Pulador and Capão da Onça)
233 and Neossolos found in a single watershed in the Ji-Paraná basin (Ji-Paraná@Cacoal).
234 Latossolos, Argissolos, Cambissolos, and Neossolos are generally equivalent to Oxisols,
235 Ultisols, Inceptisols, and Entisols in US Soil Taxonomy (Soil Survey Staff, 1997).

236 From a design standpoint, it would be best to have observations for all combinations of
237 factor values. In other words, the ideal would be to have samples from every vegetation type, on
238 every soil type, under all land uses, for every stream order, month and year. This is seldom the
239 case for studies using observational data. The configuration of samples in the matrix of all
240 possible sampling combinations of the various factors (i.e., yr x month x stream order x land
241 cover x land use x soil class) is an important attribute of the analysis and can affect the
242 uncertainty in the estimated beta adjustments. For example, if there are certain months or soil
243 types or month x soil type combinations that are not represented by actual samples there is little
244 information with which to estimate adjustments and there is large uncertainty. The
245 multidimensional matrix is difficult to represent in total (i.e., 246,960 combinations from 12 yrs
246 x 12 months x 5 stream orders x 7 land covers x 7 land uses x 7 soil classes) but coplots can
247 represent three factors simultaneously (Figure 2). The coplots indicate that while every
248 combination of factors is not represented in every month, the data are far from perfect colinearity
249 among the factors. In the case of perfect colinearity, the coplots would show one and only one

250 factor value on the y axis corresponding to each factor value on the x axis. The coplots indicate,
251 however, that soil and land use are well represented in most years and months but are sparser
252 with stream order or with vegetation cover (Figure 2, coplots by year and cover not shown).

253

254 *Discharge-Concentration Regression Analysis*

255 $\log_{10}\text{Ca}$ – $\log_{10}\text{Q}$ relationships for 25 stream stations with sufficient data were analyzed
256 for each stream-station (Table 4). Within these individual station regressions for the 25 streams,
257 13 regressions had slopes significantly different from zero with a clear preponderance having
258 negative slopes (Figure 3). Ji-Paraná@Cacoal and IG54-S5 (IG54 at station 5) were the only
259 stream stations with significant positive slopes. Of the available stations that had both discharge
260 and concentration data but slopes not different from zero only the Rancho Grande Forest stream
261 had large sample size ($n=187$); all others had <13 samples.

262 Using various combinations of the available factors to analyze the $\log_{10}\text{Ca}$ – $\log_{10}\text{Q}$
263 regression relationship across all streams and stations the multilevel linear model was utilized to
264 partially pool the data. Using the available factors (i.e., year, season, order, cover, use, and soil)
265 the model search results suggest that the complete model is the best (i.e., lowest DIC) at
266 predicting Ca concentration (Table 5). A number of the five component models provide good
267 fits but each is improved by inclusion of the additional adjustment parameter. Comparison of
268 some of the 3, 4, or 5 factor models with or without land use or soil class (e.g., season veg soil vs
269 season veg use) suggest that models including land use were slightly improved.

270 To investigate the relative contribution of the various factors (i.e., year, season, order,
271 cover, use, and soil) to the overall variance in the $\log_{10}\text{Ca}$ concentration response an ANOVA
272 decomposition analysis was utilized to interpret the multilevel linear model results (Figure 4).

273 For the model containing all variables, the graphically based ANOVA decomposition indicates
274 that variance explained by the model intercept term (Int) exceeds the unexplained variance (s.y.).
275 In addition, discharge (i.e., FLOWREG) has a relatively large effect on Ca, although over this
276 broad data set, this slope term is not extremely well defined. The intercept is affected by stream
277 order, soil type, land use, and land cover. Season and year have a small but measureable effect
278 on the intercept. In contrast, land cover, season, and year have a larger effect on the $\log_{10}\text{Ca}$ -
279 $\log_{10}\text{Q}$ regression slope than do soil type, stream order, or land use (Figure 4).

280 Individual adjustments for each class of each factor to the mean intercept or slope are
281 estimated and presented such that their mean is zero (Figure 5 and 6). In other words, the mean
282 intercept and slope terms from Equation 5 ($\mu_{\beta 0}$ and $\mu_{\beta 1}$, respectively) have not been added to the
283 values in Figures 5 and 6. Instead the means for $\mu_{\beta 0}$ and $\mu_{\beta 1}$ have been noted on the “zero”
284 (vertical dotted line) in these graphs. The individual adjustments for the intercept demonstrate
285 small adjustments for all months and all years (Figure 5a and b). Within the other factors a
286 number of adjustments are substantial, for example, 1st order streams, mixed forest (fmixed) land
287 use, and Cambissolos soil classes (Figure 5d, e, and f). For these three highlighted classes,
288 adjustments were negative and thus are a subtraction from the mean value. The individual
289 adjustments for each class of each factor for the slope demonstrate some different patterns with
290 effects being evident for both season and year (Figure 6a and b). May and April have the largest
291 positive adjustments and October and November the most negative. Adjustments for 1st order
292 streams, mixed forests, and Cambissolos are still evident, although positive in this case. In
293 addition, a substantial positive adjustment for open tropical forest (vegcode A) is evident.

294 The additive effects of the adjustments on the $\log_{10}\text{Ca}$ - $\log_{10}\text{Q}$ relationship predicted over
295 all years and seasons at each station (Figure 7) indicate an overall preponderance of positive

296 slopes (i.e., 29 positive, 13 negative). For locations with individual station regressions (Table 4),
297 these multilevel predictions are largely consistent except for Ji-Paraná@Cacoal, which had a
298 positive individual regression slope but is poorly defined in the multilevel model, and for
299 Taquara, which had a negative individual regression slope at $p=0.07$ (Table 4) but is predicted to
300 be positive by the model. Given the mapping scale used for each stream-station classification,
301 adjustment factors and thus slopes are similar in some cases for all stations (e.g., Capitão Poço
302 (CP 1-4)) but may differ if, for example, stream order changes downstream (e.g., Igarapé Sete
303 (IG7 1-7)).

304

305 **Discussion**

306 *Discharge-Concentration regressions*

307 This study considers many of the major controls on element supply to streams including
308 stream hydrology (discharge), stream geomorphology (order), landscape vegetation (land cover),
309 land-use practices, soil type and interannual variance (year) as they affect discharge-
310 concentration relationships. Discharge-concentration relationships are element specific but in
311 the case of rock-derived elements such as Ca there is typically a dilution of rock-derived,
312 element-enriched groundwaters by surface or stormflow runoff such that concentration decreases
313 with increasing flow (i.e. negative slope) (Drever, 1997). This pattern was observed in
314 regressions by individual station for 11 of the 13 stream datasets available (Figure 2). The two
315 streams with positive slopes (IG54-S5 and Ji-Paraná@Cacoal) were quite distinct from each
316 other in location (eastern vs western Amazon), stream order (1 vs 6), land cover (dense vs open
317 forest), and soil classification (Latosolos amarelo distrófico and Argissolos/Neossolos). In fact,
318 Ji-Paraná@Cacoal was distinct from all other streams in having Neossolos, which have a high

319 sand content. On the other hand, Ji-Paraná@Cacoal and IG54-S5 are somewhat similar in
320 having large portions of non-forest land uses (i.e., 30 and 40% pasture, respectively) in their
321 watersheds with Ji-Paraná@Cacoal possessing ~1% urban land use (Ballester et al., 2003) while
322 IG54 has ~22% row-crop agriculture (Figueiredo et al., 2010). These watersheds provide some
323 support for the proposed hypotheses regarding controlling factors of positive slopes in Ca-
324 discharge relationships (i.e., soils and underlying parent material or land use conversion) with the
325 Ji-Paraná@Cacoal watershed providing support for both alternatives and IG54-S5 providing
326 more support for the latter.

327

328 *Multilevel Analysis*

329 Rather than seeking to explain positive or negative slopes to the Ca-discharge regression
330 within individual streams based on site-specific factors, the multilevel analysis pools the
331 available data and interprets the relative effect of the various model factors on the overall
332 regression intercept and slope. The multilevel analysis clearly demonstrates an overall strong
333 effect of discharge (i.e., $\log_{10}Q$) on Ca concentration (Figure 4) with an overall mean slope that
334 is negative (Figure 6). In the intercept of the discharge concentration regression, stream order
335 explains the greatest amount of variation with 1st order streams requiring a large negative
336 adjustment (Figure 5d) indicating these streams have lower Ca concentrations. There are a
337 limited number of studies that have directly investigated the effect of stream order on stream
338 water concentration mostly focusing on N and P (Kang et al., 2008). A few studies have
339 demonstrated declining N concentration with increasing stream order while the trend for P has
340 been reversed. In the Seine River in France Ca concentrations had little variance with increasing
341 stream order (Meybeck, 1998). Data presented by Ballester et al. (2003) for the Ji-Paraná river

342 from 3rd to 7th order streams do possess increasing mean Ca concentrations. Increasing Ca
343 concentration in larger streams may reflect a greater contribution of groundwater relative to
344 surface water throughout the year.

345 Soil type and land use also affect the mean concentration of Ca. In the current analysis
346 the scale of soil maps used for classification was quite coarse but was consistent with
347 observations made within each watershed. The effect of lithology on stream chemical
348 concentrations, at least within the main tributaries of the Amazon, has been well investigated and
349 increasing Ca concentration with base-rich bedrock has been well demonstrated (Gibbs, 1967;
350 Mortatti and Probst, 2003; Richey et al., 1990; Stallard, 1985; Stallard and Edmond, 1987). At a
351 smaller scale (<13,000 km²) the effect of base –rich soil types on increasing Ca concentration in
352 the western Amazon has also been demonstrated (Biggs et al., 2002). In the present analysis,
353 Argissolos vermelho-amarelo eutrófico (ArgissolosVeAmEut) are in a eutrophic or base rich soil
354 group but do not require a positive adjustment that would reflect a higher Ca concentration. The
355 Latossolos amarelo escuro/Cambissolos association (LatossolosAmEsc/Cambissolos) and the
356 Argissolos/Neossolos association are classifications that include soils that have weak horizon
357 development and likely reflect sandy substrates. As such, these soils should be base poor with
358 potentially lower Ca concentrations. In these soils, the Cambissolos type had a negative
359 adjustment indicating a Ca concentration lower than the mean.

360 The effects of interannual variation or season on mean Ca concentration are limited for
361 explaining the variation in mean Ca concentrations across the data set. A similar pattern was
362 demonstrated for the main stem Amazon and its tributaries where inter- or intra-annual variance
363 within a river sampling station was small relative to the variance among the rivers (Mortatti and
364 Probst, 2003).

365 Interpretation of adjustment parameters on the slope of the discharge-concentration
366 relationship differs from those discussed above for the intercept term. In the case of the slope
367 adjustment, year and season explain much of the variation along with vegetation cover. Seasonal
368 adjustments in stream chemical compositions in the form of 12 monthly parameters are
369 commonly utilized to estimate changes in seasonal processes including discharge (StatSoft,
370 2010). Presently, the seasonal adjustments to slope are well defined for each month of the year
371 with the adjustment being positive in April and May (Figure 6a), which are rainy season months
372 in all locations other than the Cerrado (Markewitz et al., 2006).

373 The importance of vegetative cover to the slope adjustment as compared to land use was
374 unexpected although the vegetative cover classes do include an aspect of land use. Both the land
375 cover vegetation classes A (open tropical forest with secondary forest and agricultural activity)
376 and D (dense tropical forest with secondary forest and agricultural activity) have greater land
377 cover conversion than classes As (open tropical forest) and Ds (dense tropical forest). In fact,
378 the A and D classes both have positive slope adjustments where As and Ds are negative (Figure
379 6c). This change in adjustment is consistent with the hypothesis of land use conversion
380 increasing surface runoff concentrations. Increases in surface runoff with forest conversion to
381 pasture have been demonstrated in a number of Amazonian locations with responses being most
382 evident on watersheds $< 1 \text{ km}^2$ (Biggs et al., 2006; Germer et al., 2009; Moraes et al., 2006).
383 Only in the case of Rancho Grande have concentration-discharge relationships been
384 quantitatively evaluated with land use change (Germer et al., 2009). At this site during a number
385 of storm-event hydrographs Ca concentration increased initially with stormflow runoff in both
386 the forest and pasture watershed and remained elevated throughout the storm with Ca exports in
387 storm flow from the pasture being greater. Despite these increased Ca fluxes during the storm

388 both the forest and pasture watershed had a net Ca retention relative to inputs. In the current
389 analysis, which combined both storm-event and non-event data from Rancho Grande for
390 analysis, a similar increase in Ca concentration with increasing discharge was not evident (Figure
391 3).

392 In the land use classes some similar evidence for an effect of forest conversion is
393 apparent with the Fmixed, Curban, and Cmixed classes all requiring positive adjustment to slope
394 (Figure 6e). On the other hand, the Pasture and Furban adjustment are not positive, although
395 Furban is very poorly defined (i.e. few samples and large variance). Of course, there are many
396 studies that have demonstrated an increase in stream solute concentrations with land use
397 conversion (Likens and Bormann, 1995; Williams and Melack, 1997) but few that have
398 specifically observed changes in discharge-concentration relationships with changing land use
399 (Germer et al., 2009; Markewitz et al., 2001).

400 The predictive multilevel model indicates that the additive adjustments of all the factors
401 (year, season, stream order, land cover, land use, and soil class) on $\log_{10}\text{Ca}$, in many cases,
402 results in positive slopes for $\log_{10}\text{Ca}$ vs $\log_{10}\text{Q}$. The model, of course, reflects the data of which
403 nearly 1/3 are from IG54. This stream has a significant positive slope and shares many attributes
404 (i.e., soil, land use, land cover) with the other streams in the eastern Amazon (i.e., Region C in
405 Figure 1) and thus influences these predictions. It is uncertain how representative IG54 is for
406 this region (Davidson et al., 2010; Figueiredo et al., 2010). As such, one value of the multilevel
407 model is knowledge gained about where future sampling should occur to best learn about the
408 factors and relationships of interest. Clearly, sampling of additional streams in this rapidly
409 changing portion of the eastern Amazon would be valuable.

410

411 **Conclusion**

412 Across the Amazon and Cerrado of Brazil the hydrology of many low order streams is
413 being impacted by land use conversion as evidenced by studies demonstrating increasing surface
414 runoff, peak flows, and water yield. The factors controlling the expected responses in stream
415 concentration or concentration-discharge relationships, however, are only beginning to be
416 elucidated. In the present study the role of year, season, stream order, vegetation cover, land use,
417 and soil type were investigated for 28 streams. Ca concentrations and discharge varied across
418 three and six orders of magnitude, respectively. In 13 streams with significant concentration-
419 discharge relationships in the individual station regressions, 11 had negative slopes while two
420 had increasing concentrations with discharge. There were no readily apparent similarities
421 between these two stream watersheds and competing hypothesis of soil or land use control in
422 affecting these positive slopes were not well differentiated. Multilevel analysis of the pooled
423 data, however, indicated that soils and land use as well as stream order all explained portions of
424 the variance in mean Ca concentrations while season, year, and vegetative cover explained much
425 of the variance in the slope of the discharge-concentration regression. The utilized vegetative
426 cover classes incorporate aspects of land use and thus suggest a larger role for land use in
427 discharge-concentration slopes than soil classes.

428

429 **Acknowledgements**

430 This research was supported by grant #'s NCC5-686 and NNG06GE88A of NASA's
431 Terrestrial Ecology Program as part of the Large-scale Biosphere-Atmosphere Experiment in
432 Amazonia (LBA-ECO) project. We thank Luke Worsham and Paul Lefebvre for help drawing
433 the Amazon Basin inset map. We also thank others in the LBA Land Water Coupling Synthesis
434 Group including Jeff Richey, John Melack, and Eric Davidson.

435

436 **References**

- 437 Ballester, M.V., D.d.C. Victoria, A.V. Krusche, R. Coburn, R.L. Victoria, J.E. Richey, M.G.
438 Logsdon, E. Mayorga, and E. Matricardi. 2003. A remote sensing/GIS-based physical
439 template to understand the biogeochemistry of the Ji-Parana river basin (Western
440 Amazonia). *Remote Sensing of Environment* 87:429-445.
- 441 Biggs, T.W., T. Dunne, and T. Muraoka. 2006. Transport of water, solutes and nutrients from a
442 pasture hillslope, southwestern Brazilian Amazon, pp. 2527-2547, Vol. 20. John Wiley &
443 Sons, Ltd.
- 444 Biggs, T.W., T. Dunne, T.F. Domingues, and L.A. Martinelli. 2002. The relative influence of
445 natural watershed properties and human disturbance on stream solute concentrations in
446 the southwestern Brazilian Amazon basin. *Water Resources Research* 38:1150.
- 447 Bond, H.W. 1979. Nutrient Concentration Patterns in a Stream Draining a Montane Ecosystem
448 in Utah. *Ecology* 60:1184-1196.
- 449 Chaves, J., C. Neill, S. Germer, S.G. Neto, A. Krusche, and H. Elsenbeer. 2008. Land
450 management impacts on runoff sources in small Amazon watersheds. *Hydrological
451 Processes* 22:1766-1775.

452 Clayton, D.G. 1996. Generalized Linear Mixed Models, p. 275-301, *In* W. R. Gilks, et al., eds.
453 Markov Chain Monte Carlo in Practice. Chapman and Hall, London.

454 Congdon, P. 2001. Bayesian Statistical Modelling. John Wiley and Sons, LTD, New York.

455 Davidson, E.A., R.O. Figueiredo, D. Markewitz, and A.K. Aufdenkampe. 2010. Dissolved CO₂
456 in small catchment streams of eastern Amazonia: A minor pathway of terrestrial carbon
457 loss. *J. Geophys. Res.* 115:G04005.

458 Drever, J.I. 1997. The geochemistry of natural waters: Surface and groundwater environments. 3
459 ed. Prentice Hall.

460 Figueiredo, R.O., D. Markewitz, E.A. Davidson, A.E. Schuler, O. dos S. Watrin, and P. de Souza
461 Silva. 2010. Land-use effects on the chemical attributes of low-order streams in the
462 eastern Amazon. *J. Geophys. Res.* 115:G04004.

463 Gelman, A. 2005a. Analysis of variance: Why it is more important than ever (with discussion).
464 *Annals of Statistics* 35:1-53.

465 Gelman, A. 2005b. Prior distributions for variance parameters in hierarchical models. *Bayesian*
466 *Analysis* 1:1-19.

467 Germer, S., C. Neill, T. Vetter, J. Chaves, A.V. Krusche, and H. Elsenbeer. 2009. Implications of
468 long-term land-use change for the hydrology and solute budgets of small catchments in
469 Amazonia. *Journal of Hydrology* 364:349-363.

470 Gibbs, R.J. 1967. The geochemistry of the Amazon river system, I: The factors that control the
471 salinity and the composition and concentration of suspended solids. *Geological Society of*
472 *America Bulletin* 78:1203-1232.

473 Hornberger, G.M., K.E. Bencala, and D.M. McKnight. 1994. Hydrological Controls on
474 Dissolved Organic Carbon during Snowmelt in the Snake River near Montezuma,
475 Colorado. *Biogeochemistry* 25:147-165.

476 INPE. 2006. Monitoramento da floresta Amazônica Brasileira por satélite:Projeto PRODES. .

477 Johnson, M., J. Lehmann, E. Couto, J. Filho, and S. Riha. 2006. DOC and DIC in Flowpaths of
478 Amazonian Headwater Catchments with Hydrologically Contrasting Soils.
479 *Biogeochemistry* 81:45-57.

480 Kang, S., H. Lin, W.J. Gburek, G.J. Folmar, and B. Lowery. 2008. Baseflow Nitrate in Relation
481 to Stream Order and Agricultural Land Use. *J Environ Qual* 37:808-816.

482 Lamon, E.C.I., and S.S. Qian. 2008. Regional scale stressor-response models in aquatic
483 ecosystems. *Journal of the American Water Resources Association* 44:771-781.

484 Lewis, W.M.J., and M.C. Grant. 1979. Relationships Between Stream Discharge and Yield of
485 Dissolved Substances From A Colorado Mountain Watershed. *Soil Science* 128:353-363.

486 Likens, G.E., and F.H. Bormann. 1995. *Biogeochemistry of a forested ecosystem*. 2 ed. Springer-
487 Verlag, New York.

488 Markewitz, D., E.A. Davidson, R.d.O. Figueiredo, R.L. Victoria, and A.V. Krusche. 2001.
489 Control of cation concentrations in stream waters by surface soil processes in an
490 Amazonian watershed. *Nature* 410:802-805.

491 Markewitz, D., J.C.F. Resende, L.M. Parron, M.M.C. Bustamante, C.A. Klink, and E.A.
492 Davidson. 2006. Dissolved rainfall inputs and streamwater outputs in an undisturbed
493 watershed on highly weathered soils in the Brazilian Cerrado. *Hydrological Processes*.

494 Meybeck, M. 1998. Man and river interface: multiple impacts on water and particulates
495 chemistry illustrated in the Seine river basin. *Hydrobiologia* 373:1-20.

496 Meyer, J.L., W.H. McDowell, T.L. Bott, J.W. Elwood, C. Ishizaki, J.M. Melack, B.L. Peckarsky,
497 B.J. Peterson, and P.A. Rublee. 1988. Elemental Dynamics in Streams. *Journal of the*
498 *North American Benthological Society* 7:410-432.

499 Moraes, J.M.d., A.E. Schuler, T. Dunne, R.d.O. Figueiredo, and R.L. Victoria. 2006. Water
500 storage and runoff processes in plinthic soils under forest and pasture in Eastern
501 Amazonia. *Hydrological Processes* 20:2509-2526.

502 Mortatti, J., and J.-L. Probst. 2003. Silicate rock weathering and atmospheric/soil CO₂ uptake in
503 the Amazon basin estimated from river water geochemistry: seasonal and spatial
504 variations. *Chemical Geology* 197:177-196.

505 Neill, C., L.A. Deegan, S.M. Thomas, and C.C. Cerri. 2001. Deforestation for pasture alters
506 nitrogen and phosphorus in soil solution and stream water of small Amazonian
507 watersheds. *Ecological Applications* 11:1817-1828.

508 Richey, J.E., R.L. Victoria, E. Salati, and B.R. Forsberg. 1990. Biogeochemistry of a major river
509 system: the Amazon case study., p. 57-74, *In* E. T. Degens, ed. *Biogeochemistry of major*
510 *world rivers*, Vol. 42. Wiley, New York.

511 Saunders, J.F., III, and W.M. Lewis, Jr. 1989. Transport of Major Solutes and the Relationship
512 between Solute Concentrations and Discharge in the Apure River, Venezuela.
513 *Biogeochemistry* 8:101-113.

514 Silva, J.S.O., M.M.C. Bustamante, D. Markewitz, A.V. Krusche, and L. Ferreira. In press.
515 Effects of land cover on chemical characteristics of streams in the Cerrado region of
516 Brazil. *Biogeochemistry*.

517 Spiegelhalter, D.J., N.G. Best, B.P. Carlin, and A. van der Linde. 2002. Bayesian measures of
518 model complexity and fit. *Journal of the Royal Statistical Society Series B* 64:583-639.

519 Stallard, R.F. 1985. River chemistry, geology, geomorphology, and soils in the Amazon and
520 Orinoco basins, p. 293-316, *In* J. I. Drever, ed. The chemistry of weathering, Vol. 149,
521 Series C ed. D Reidel Publishing Company, Boston.

522 Stallard, R.F., and J.M. Edmond. 1983. Geochemistry of the Amazon. The Influence of Geology
523 and Weathering Environment on the Dissolved Load. *Journal of Geophysical Research*
524 88:9671-9688.

525 Stallard, R.F., and J.M. Edmond. 1987. Geochemistry of the Amazon. Weathering chemistry
526 and limits to dissolved inputs. *Journal of Geophysical Research* 92:8293-8302.

527 StatSoft, I. 2010. *Electronic Statistics Textbook*, Tulsa.

528 Williams, M.R., and J.M. Melack. 1997. Solute export from forested and partially deforested
529 catchments in the central Amazon. *Biogeochemistry* 38:67-102.

530

531

Table 1. Brazilian streams utilized for multilevel analysis of discharge-Ca concentration relationships. Sta-is number of stations on each stream.

| Location, State | Stream/River | Latitude | Longitude | Yr | Sta | Order | Ppt | Basin area | Land Cover | Land use | Soil | Ref |
|------------------------------|------------------|----------|-----------|-------|-----|-------|-----|------------|------------|----------|-------|-----|
| | | | | | | | cm | ha | | | | |
| Ji-Paraná, Rondonia | Urupá | 11°40' S | 61°30' W | 99/00 | 1 | 5 | 241 | 420900 | A | Furban | AVAE | 1,2 |
| | Ji-Paraná@Cacoal | 10°80' S | 61°80' W | “ | 1 | 6 | “ | 1755900 | A | Furban | Ag/Ne | “ |
| Juruena, Mato Grosso | B1 | 10°28' S | 58°28' W | 03/06 | 1 | 1 | 258 | 2 | As | Forest | LAD | 3 |
| | B2 | 10°25' S | 58°46' W | “ | 1 | 1 | “ | 2 | As | Forest | AVAD | “ |
| Faz. Rancho Grande, Rondônia | Forest | 10°18' S | 62°52' W | 04/05 | 1 | 1 | 230 | 1.4 | Ds | Forest | AVAE | 4 |
| | Pasture | “ | “ | “ | 1 | 1 | “ | 0.7 | Ds | Pasture | AVAE | “ |
| Fazenda Nova Vida, Rondônia | Forest1 | 10°30' S | 62°30' W | 94/01 | 1 | 2 | 220 | 1740 | A | Forest | AVAE | 5 |
| | Pasture1 | “ | “ | “ | 1 | 2 | “ | 720 | A | Pasture | AVAE | “ |
| | Forest2 | “ | “ | “ | 1 | 2 | “ | 250 | A | Forest | AVAE | “ |
| | Pasture2 | “ | “ | “ | 1 | 1 | “ | 130 | A | Pasture | AVAE | “ |
| Paragominas, Pará | IG54 | 2°59' S | 47°31' W | 96/05 | 5 | 2 | 180 | 14000 | D | FMixed | LAD | 6,7 |
| | Sete | 3°16' S | 47°23' W | 03/05 | 7 | 3 | “ | 16143 | D | FMixed | LAD | 7 |
| | Pajeú | 3°10' S | 47°17' W | “ | 3 | 2 | “ | 3246 | D | FMixed | LAD | “ |
| Capitão Poço, Pará | CP1 | 2°10' S | 47°15' W | “ | 2 | 1 | 260 | 20 | D | Forest | LAD | “ |
| | CP2 | “ | “ | “ | 2 | 1 | “ | 20 | D | Forest | LAD | “ |
| Igarapé-Açu, Pará | Cumaru | 1°11' S | 47°34' W | 06/07 | 4 | 2 | 251 | 1850 | D | FMixed | AAD | 8 |
| | Pachibá | 1°10' S | 47°37' W | “ | 2 | 1 | “ | 323 | D | FMixed | AAD | “ |

| | | | | | | | | | | | | |
|--------------------|------------------|----------|----------|-------|---|---|-----|------|----|---------|-------|----|
| | São João | 1°10' S | 47°30' W | “ | 2 | 1 | “ | 570 | D | FMixed | AAD | “ |
| Brasília, Distrito | Roncador | 15°56' S | 47°53' W | 98/00 | 1 | 3 | 147 | 2000 | Sa | Cerrado | LVE | 9 |
| Federal | Pitoco | 15°55' S | 47°52' W | 05/06 | 2 | 1 | 138 | 80 | Sa | Cerrado | LVE | 10 |
| | Taquara | 15°57' S | 47°53' W | “ | 2 | 1 | “ | 150 | Sa | Cerrado | LVE | “ |
| | Vereda Grande | 15°32' S | 47°34' W | “ | 1 | 1 | “ | 3850 | Sa | Cerrado | LVE | “ |
| | Estanislau | 15°47' S | 47°37' W | “ | 2 | 1 | “ | 390 | S | Cmixed | LVE | “ |
| | Barreiro do Mato | 15°48' S | 47°36' W | “ | 1 | 1 | “ | 250 | S | Cmixed | LVE | “ |
| | Capão da Onça | 15°38' S | 48°10' W | “ | 1 | 1 | “ | 720 | S | Cmixed | LVE/C | “ |
| | Pulador | 15°40' S | 48°1' W | “ | 1 | 1 | “ | 170 | S | Curban | LVE/C | “ |
| | Mestre D'Armas | 15°36' S | 47°40' W | “ | 1 | 1 | “ | 5740 | Sa | Curban | LVE | “ |
| | Atoleiro | 15°37' S | 47°38' W | “ | 1 | 1 | “ | 2030 | Sa | Curban | LVE | “ |

Furban – forest watershed intermixed with urban areas

Fmixed – forest watershed intermixed with pasture and agricultural areas

Curban – cerrado watershed intermixed with urban areas

Cmixed-cerrado watershed intermixed with pasture and agricultural areas

A- Floresta ombrofila aberta (Floresta de transição) – Vegetação secundária e Atividades agrícolas (Open tropical rainforest (transition forest) – secondary vegetation and agricultural activities).

As- Floresta ombrófila aberta (Floresta de transição) – Submontana (Open tropical rainforest (transition forest) – sub-mountain).

ON-Áreas de tensão ecológica (contatos entre tipos de vegetação)-Floresta Ombrófila-Floresta Estacional (Ecotone {contact between two vegetation types}-tropical rainforest-seasonal forest.

Ds-Floresta ombrófila densa-submontana (Dense tropical forest – submountain).

D- Floresta ombrófila densa- Vegetação secundária e Atividades agrícolas (Dense tropical forest –secondary vegetation and agricultural activities).

Sa- Savana-Arbórea Aberta (Savannah-open woodlands).

S- Savana- Atividades agrícolas (Savannah –agricultural activities).

LAD – Latossolos amarelo distrófico (dystrophic yellow latosol)

LVE- Latossolos vermelho escuro (dark red latosol)

LVE/C – Latossolos vermelho escuro/Cambissolos (dark red latosol/cambisol)

AVAE –Argissolos vermelho-amarelo eutrófico (eutrophic red yellow argisol)

AAD-Argissolos amarelo distrófico (dystrophic yellow argisol)

AVAD - Argissolos vermelho-amarelo distrófico (dystrophic red yellow argisol)

Ag/Ne – Argissolos/Neossolos (argisol/neosol)

1 (Krusche – unpublished data); 2 (Ballester et al., 2003); 3 (Johnson et al., 2006); 4 (Chaves et al., 2008); 5 (Neill et al., 2001); 6 (Markewitz et al., 2001); 7(Figueiredo et al., 2010) ; 8 (Figueiredo - unpublished data); 9 (Markewitz et al., 2006); 10 (Silva et al., In press)

Table 2. Descriptive statistics for Log_{10}Ca concentration and Log_{10}Q for 28 streams in Brazil sampled between 1994 and 2007. Total sample size is 3155.

| Statistic | $\log_{10} \text{Ca}$ | $\log_{10}\text{Q}$ |
|--------------------------------|---|--|
| | μM | $\text{m}^3 \text{sec}^{-1}$ |
| n | 2734 | 2062 |
| Minimum | -0.432 | -6.00 |
| 1st Quartile | 1.08 | -3.243 |
| Median | 1.38 | -1.200 |
| Mean | 1.32 | -1.707 |
| 3rd Quartile | 1.64 | -0.072 |
| Max | 2.43 | 3.238 |
| Missing values | 421 | 1093 |

Table 3. Sample size available for multilevel analysis from 28 streams in Brazil sampled between 1994 and 2007.

| Year | | Month | | Stream Order | | Land Use | | Land Cover | | Soil Class | |
|------|-----|-------|-----|--------------|------|----------|------|------------|------|------------|------|
| ID | N | ID | N | ID | N | ID | N | ID | N | ID | N |
| 1994 | 21 | 1 | 291 | 1 | 1502 | Forest | 712 | A | 276 | LAD | 1336 |
| 1996 | 124 | 2 | 450 | 2 | 1407 | Fmixed | 1224 | As | 83 | LVE | 489 |
| 1997 | 271 | 3 | 366 | 3 | 198 | Furban | 48 | D | 1389 | LVE/C | 42 |
| 1998 | 340 | 4 | 206 | 5 | 24 | Pasture | 640 | Ds | 792 | AVAE | 1044 |
| 1999 | 217 | 5 | 181 | 6 | 24 | Cerrado | 350 | ON | 84 | AAD | 136 |
| 2000 | 148 | 6 | 213 | | | Cmixed | 105 | S | 126 | AVAD | 84 |
| 2001 | 73 | 7 | 172 | | | Curban | 76 | Sa | 405 | Ag/Ne | 24 |
| 2003 | 171 | 8 | 203 | | | | | | | | |
| 2004 | 589 | 9 | 273 | | | | | | | | |
| 2005 | 820 | 10 | 241 | | | | | | | | |
| 2006 | 305 | 11 | 385 | | | | | | | | |
| 2007 | 40 | 12 | 172 | | | | | | | | |

Furban – forest watershed intermixed with urban areas

Fmixed – forest watershed intermixed with pasture and agricultural areas

Curban – cerrado watershed intermixed with urban areas

Cmixed-cerrado watershed intermixed with pasture and agricultural areas

A- Floresta ombrófila aberta (Floresta de transição) – Vegetação secundária e Atividades agrícolas (Open tropical rainforest (transition forest) – secondary vegetation and agricultural activities).

As- Floresta ombrófila aberta (Floresta de transição) – Submontana (Open tropical rainforest (transition forest) – submountain).

ON-Áreas de tensão ecológica (contatos entre tipos de vegetação)-Floresta Ombrófila-Floresta Estacional (Ecotone {contact between two vegetation types})-tropical rainforest-seasonal forest.

Ds-Floresta ombrófila densa-submontana (Dense tropical forest – submountain).

D- Floresta ombrófila densa- Vegetação secundária e Atividades agrícolas (Dense tropical forest –secondary vegetation and agricultural activities).

Sa- Savana-Arbórea Aberta (Savannah-open woodlands).

S- Savana- Atividades agrícolas (Savannah –agricultural activities).

LAD – Latossolos amarelo distrófico (dystrophic yellow latosol)

LVE- Latossolos vermelho escuro (dark red latosol)

LVE/C – Latossolos vermelho escuro/Cambissolos (dark red latosol/cambisol)

AVAE –Argissolos vermelho-amarelo eutrófico (eutrophic red yellow argisol)

AAD-Argissolos amarelo distrófico (dystrophic yellow argisol)

AVAD - Argissolos vermelho-amarelo distrófico (dystrophic red yellow argisol)

Ag/Ne – Argissolos/Neossolos (argisol/neosol)

Table 4. Linear regression statistics for $\text{Log}_{10}Q$ ($\text{m}^3 \text{sec}^{-1}$) vs Log_{10}Ca (μM) for individual stream stations. Statistics include adjusted R^2 , y intercept (y_0) and standard error (SE_{y_0}), slope and standard error (SE_{slope}), p-values for tests of y-intercept (P_{y_0}) and slope (P_{slope}) different from zero.

| ID | Adj. R^2 | y_0 | SE_{y_0} | Slope | SE_{slope} | P_{y_0} | P_{slope} |
|-------------|------------|--------|-------------------|--------|----------------------------|-----------|--------------------|
| Urupá | 0.50 | 2.550 | 0.105 | -0.186 | 0.038 | 0.0001 | 0.0001 |
| Ji-ParanáCa | 0.80 | 1.019 | 0.067 | 0.262 | 0.027 | 0.0001 | 0.0001 |
| JuruenaB1 | 0.59 | -1.043 | 0.251 | -0.713 | 0.072 | 0.0001 | 0.0001 |
| JuruenaB2 | 0.79 | -0.253 | 0.096 | -0.719 | 0.030 | 0.0104 | 0.0001 |
| RGForest | 0.00 | 1.446 | 0.076 | -0.008 | 0.017 | 0.0001 | 0.6621 |
| RGPasture | 0.02 | 1.302 | 0.024 | -0.027 | 0.007 | 0.0001 | 0.0002 |
| FazNVFor1 | 0.62 | 1.779 | 0.015 | -0.122 | 0.014 | 0.0001 | 0.0001 |
| FazNVPas1 | 0.23 | 1.696 | 0.046 | -0.152 | 0.040 | 0.0001 | 0.0004 |
| FazNVFor2 | 0.38 | 1.880 | 0.030 | -0.076 | 0.019 | 0.0001 | 0.0004 |
| FazNVPas2 | 0.18 | 1.714 | 0.080 | -0.158 | 0.059 | 0.0001 | 0.0119 |
| IG54-S5 | 0.20 | 1.282 | 0.010 | 0.996 | 0.086 | 0.0001 | 0.0001 |
| IG54-S3 | 0.00 | 1.466 | 0.100 | 0.296 | 0.279 | 0.0001 | 0.3139 |
| Sete-S2 | 0.08 | 0.691 | 0.340 | -1.684 | 1.073 | 0.0691 | 0.1477 |
| Sete-S4 | 0.11 | 2.465 | 0.721 | -2.915 | 1.711 | 0.0066 | 0.1192 |
| Sete-S5 | 0.03 | 0.676 | 0.419 | 1.883 | 1.553 | 0.1381 | 0.2533 |
| Sete-S6 | 0.00 | 1.673 | 0.717 | -1.204 | 1.585 | 0.0445 | 0.4670 |
| Pajeú-S2 | 0.00 | 1.065 | 0.304 | 0.183 | 0.392 | 0.0057 | 0.6504 |
| CumaruA | 0.00 | 0.549 | 0.690 | -0.024 | 0.127 | 0.4170 | 0.8512 |
| CumaruB | 0.00 | 0.833 | 0.700 | 0.086 | 0.132 | 0.2593 | 0.5290 |
| CumaruC | 0.00 | 1.160 | 0.129 | -0.037 | 0.038 | 0.0001 | 0.3579 |
| CumaruD | 0.67 | 0.169 | 0.230 | -0.337 | 0.073 | 0.4799 | 0.0013 |
| Roncador | 0.22 | 1.426 | 0.044 | -0.342 | 0.043 | 0.0001 | 0.0001 |
| Taquara | 0.16 | -3.850 | 2.164 | -2.908 | 1.483 | 0.0970 | 0.0700 |
| Pachibá | 0.00 | 0.818 | 0.367 | 0.037 | 0.093 | 0.0546 | 0.6947 |
| São João | 0.00 | 0.893 | 0.163 | 0.015 | 0.043 | 0.0028 | 0.7501 |

Table 5: Results of the model search within the ANOVA models using year, season, stream order, vegetation cover, land use, and soil type. DIC is an estimate of expected predictive error (lower including more negative deviance is better). Dbar is a Bayesian measure of fit, while pD ($pD = Dbar - Dhat$) is the estimated number of independent parameters (complexity) of the multilevel model. C is an indicator for convergence; M is an indicator that Markov chains have mixed during simulation.

| Model | Dbar | Dhat | pD | DIC | C | M |
|--------------------------------|-------------|-------------|-----------|------------|----------|----------|
| Season Veg Soil | 1581.99 | 1445.78 | 136.21 | 1718.20 | 1 | 0 |
| Season Veg Use | 1473.75 | 1269.96 | 203.79 | 1677.53 | 1 | 1 |
| Year Season Order | 726.960 | 619.82 | 107.140 | 834.100 | 0 | 0 |
| Year Season Soil | 386.899 | 189.674 | 197.225 | 584.123 | 1 | 1 |
| Season Veg Use Soil | 1431.81 | 1294.59 | 137.22 | 1569.03 | 1 | 0 |
| Year Season Use Soil | -372.663 | -722.334 | 349.671 | -22.992 | 1 | 1 |
| Year Season Order Soil | -280.310 | -609.287 | 328.976 | 48.666 | 1 | 0 |
| Year Season Order Use | 470.518 | 215.02 | 255.497 | 726.015 | 1 | 1 |
| Year Season Order Veg | -607.721 | -1090.06 | 482.338 | -125.383 | 1 | 1 |
| Year Season Veg Soil | 3.994 | -260.95 | 264.942 | 268.936 | 1 | 1 |
| Year Season Veg Use | -581.756 | -893.92 | 312.166 | -269.589 | 0 | 0 |
| Year Season Order Veg Soil | -688.375 | -1157.03 | 468.654 | -219.721 | 1 | 0 |
| Year Season Order Veg Use | -835.041 | -1296.67 | 461.634 | -373.407 | 1 | 1 |
| Year Season Order Use Soil | -560.757 | -610.80 | 50.042 | -510.714 | 0 | 0 |
| Year Season Veg Use Soil | -746.277 | -962.58 | 216.302 | -529.976 | 1 | 1 |
| Year Season Order Veg Use Soil | -991.330 | -1237.39 | 246.062 | -745.268 | 1 | 1 |

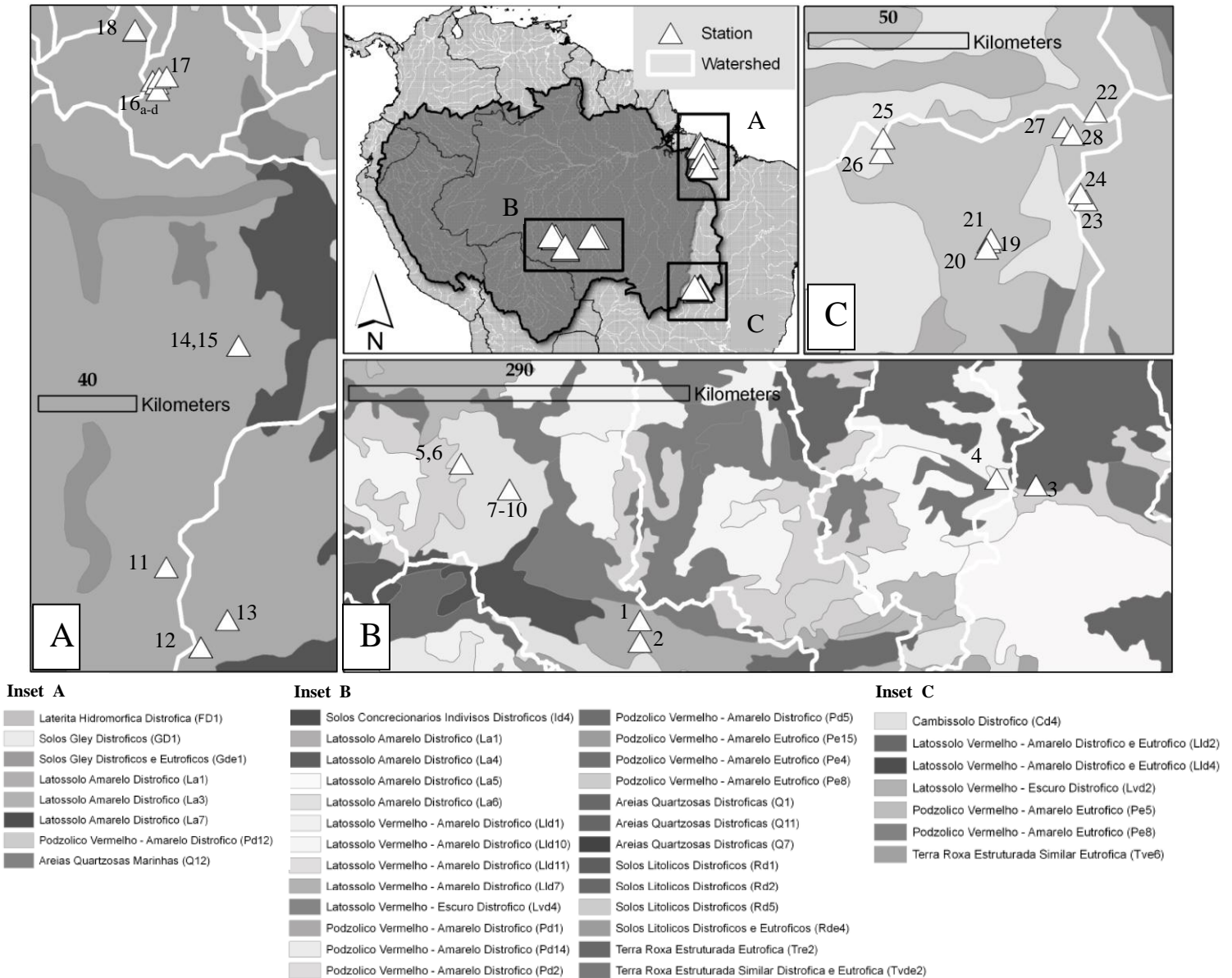


Figure 1. Locations of streams in the Amazon and Cerrado of Brazil. Underlying map is RADAM soil classifications. 1-Urupá; 2-Ji-Paraná@Cacoal ; 3-B1; 4-B2; 5,6-Rancho Grande; 7-10 Nova Vida; 11-IG54; 12-Sete; 13-Pajeú; 14,15 Capitão Poço; 16-Cumarú; 17-Pachibá; 18-São João; 19-Roncador; 20-Pitoco; 21-Taquara; 22-Vereda Grande; 23-Estanislau; 24-Barreiro do Mato; 25-Capão do Onça; 26-Pulador; 27-Mestre D’Armas; 28-Atoleiro.

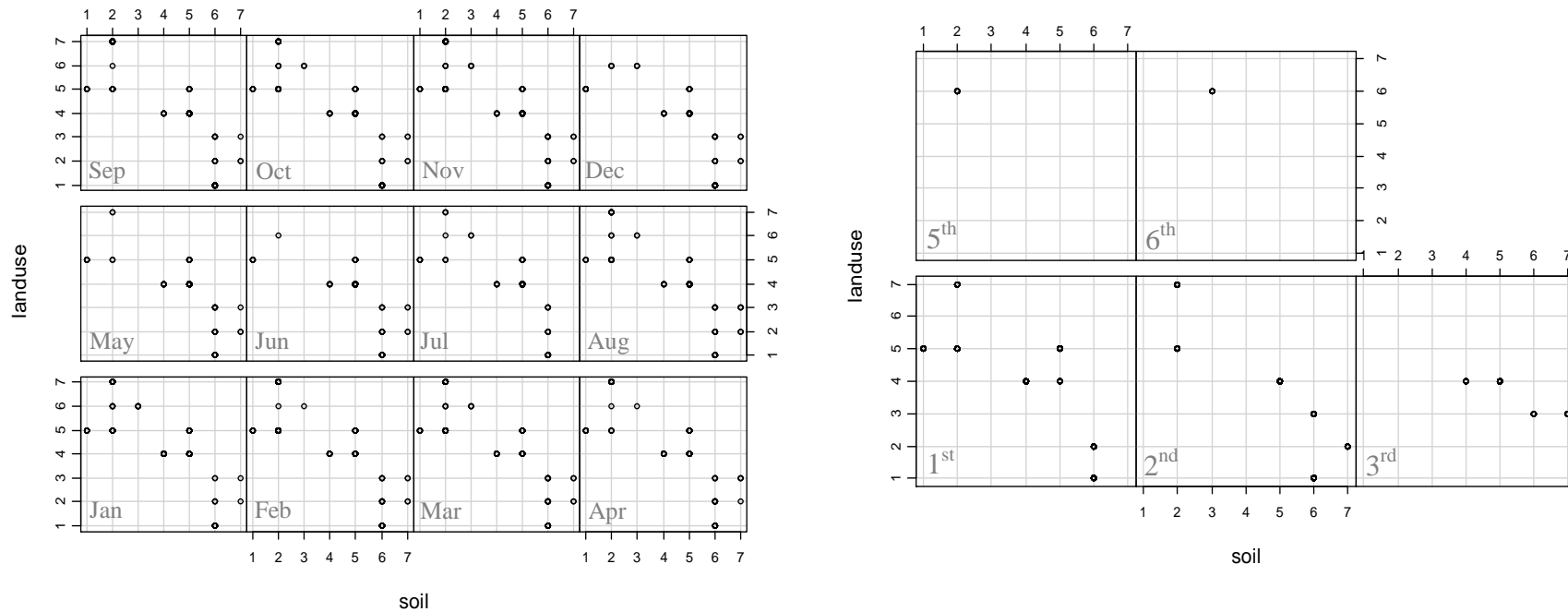


Figure 2. Coplots for landuse and soil given A) season (i.e., month) or B) stream order. Circles indicate presence of data corresponding to each soil type and land use, for each level of the marginal variable (i.e., season or order). Ideal would be a representative of each soil type in each land use for each season or stream order. Each month in season is well represented although July is missing soil type 7 (made up of land uses 2 and 3 in other months) and land use number 1 (Cerrado) is all in soil type 6, for all months. Stream order 1, 2, and 3 are well represented but order 5 and 6 are single soil and land use combinations.

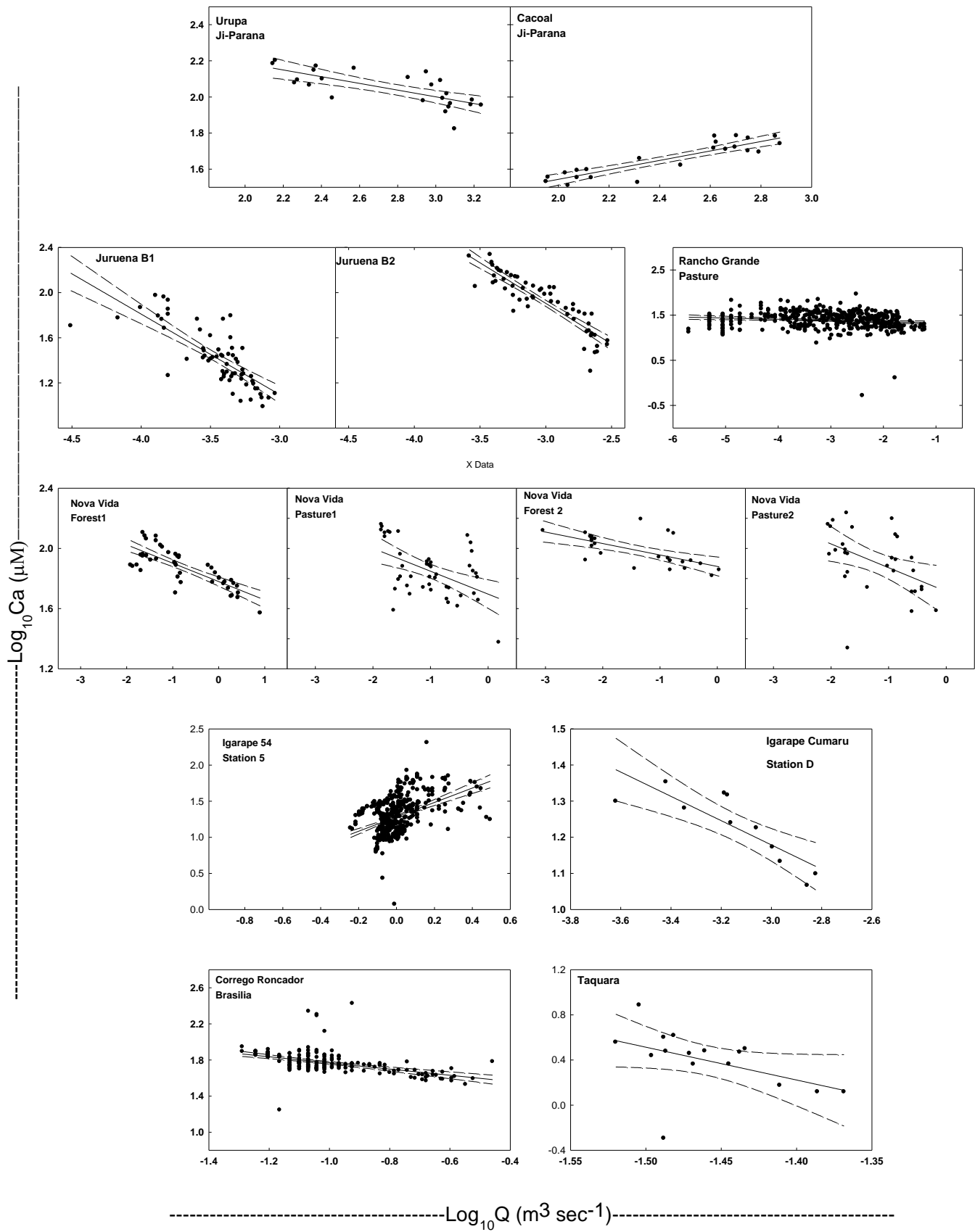


Figure 3. Log_{10}Ca (μM) vs Log_{10}Q ($\text{m}^3 \text{sec}^{-1}$) relationship for 13 streams in Brazil. Solid lines are least square linear regressions and dashed lines are upper and lower 95% confidence intervals. Data were collected between 1996 and 2005.

ANOVA TABLE

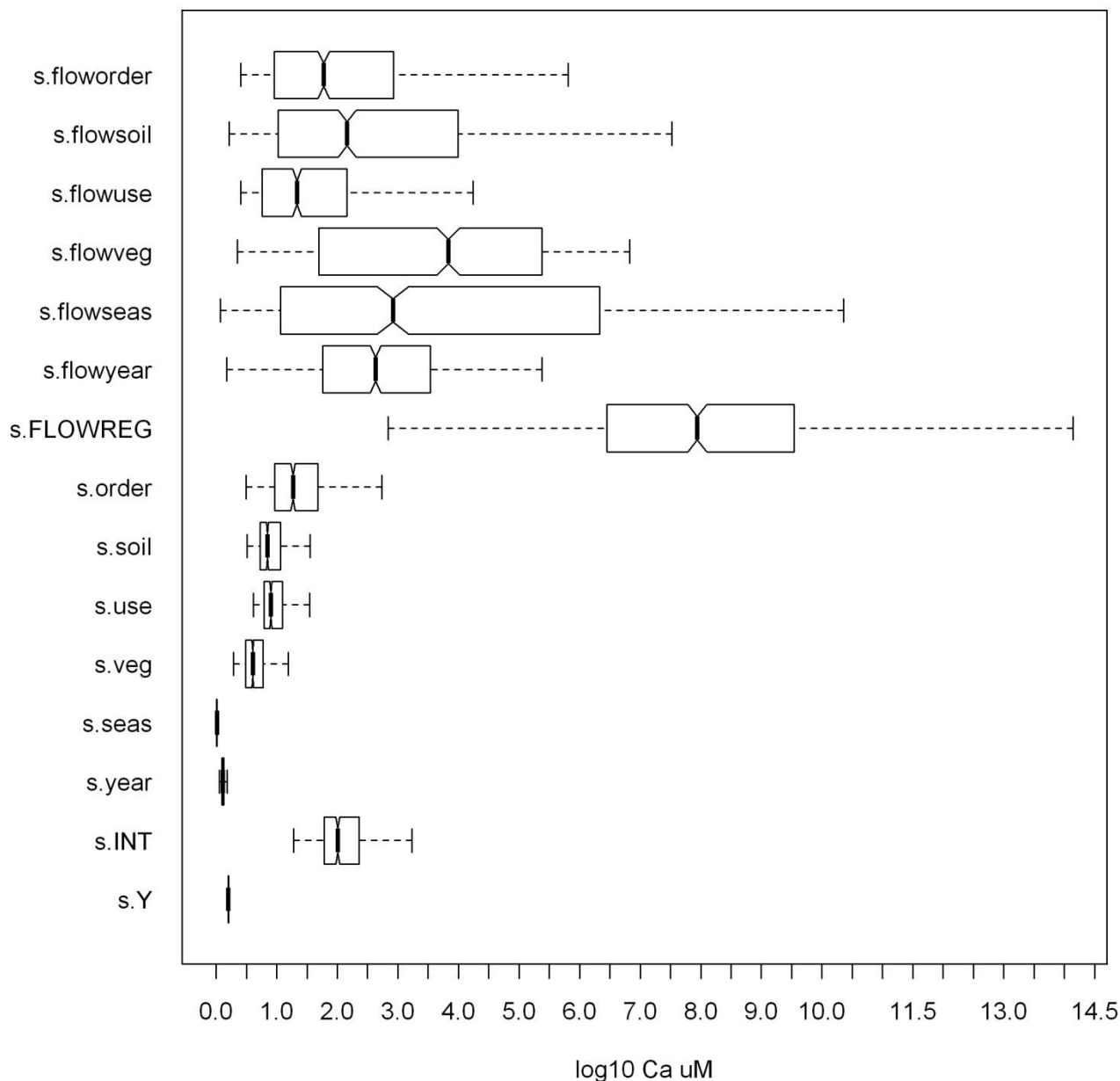


Figure 4 – ANOVA analysis for main effects on Log_{10}Ca concentration ($n=3155$). The mean of the box plots represents the proportion of the standard deviation explained by each component and the distribution represents how well the effect is determined. The upper boxes (s.flow(factor)) represent the decomposition of the variance explained by the slope of the discharge-Ca regression slope (s.FLOWREG) and the lower boxes (s.(factors)) represent the decomposition of the variance in the intercept term (s.INT). The s.y. component identifies the unexplained variance.

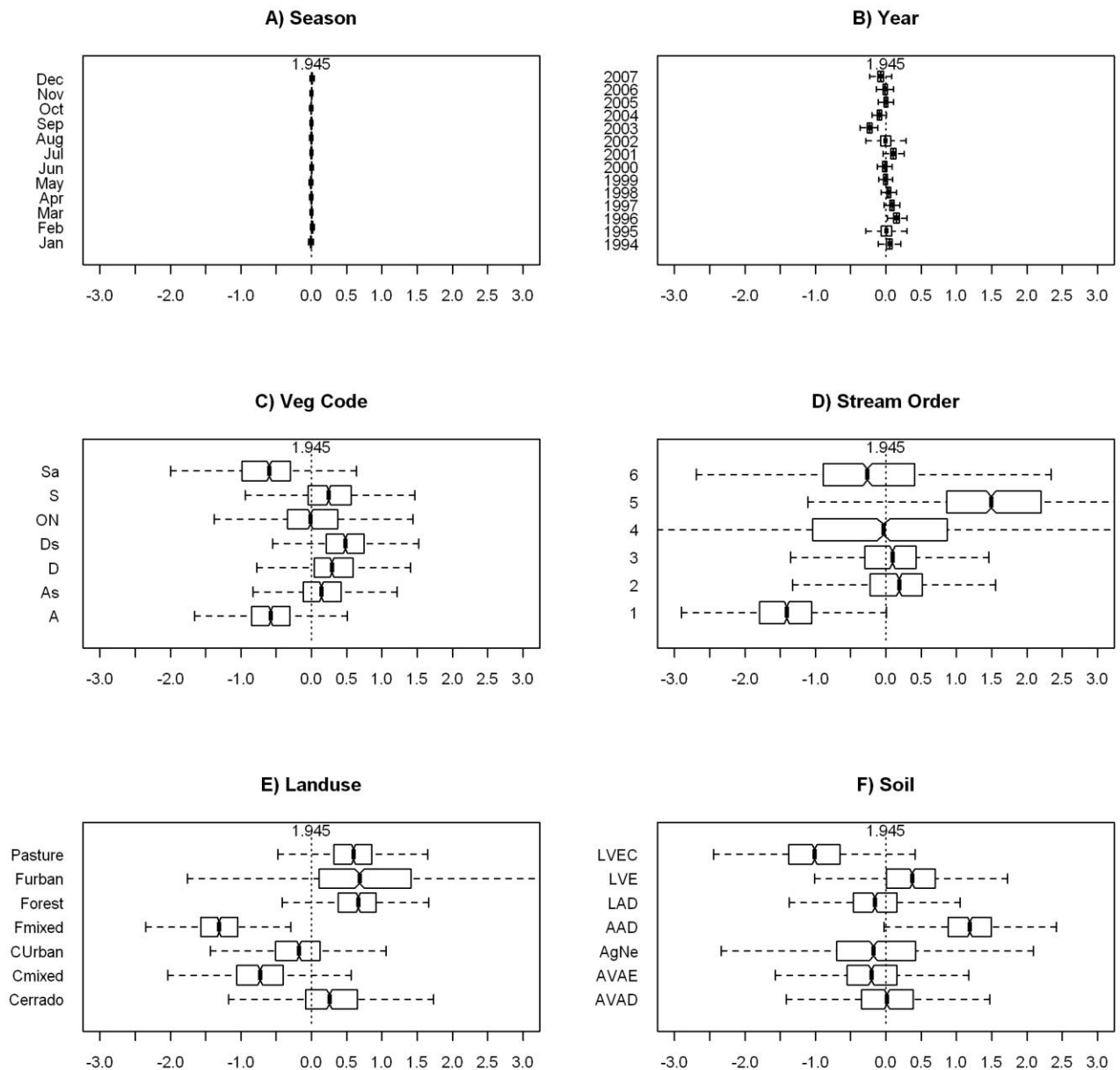


Figure 5 – Intercept adjustments associated with the $\log_{10}Q$ regression. The overall mean intercept is identified by the dotted line in each panel. A-Open tropical rainforest with secondary vegetation and agricultural activities); As-Open tropical rainforest – sub-mountain; ON-Ecotone tropical rainforest-seasonal forest; Ds-Dense tropical forest – submountain; D-Dense tropical forest –secondary vegetation and agricultural activities; Sa-Savannah-open woodlands; S-Savannah –agricultural activities). LAD – Latossolos amarelo distrófico; LVE- Latossolos vermelho escuro; LVE/C – Latossolos vermelho escuro/Cambissolos; AVAE –Argissolos vermelho-amarelo eutrófico; AAD-Argissolos amarelo distrófico; AVAD - Argissolos vermelho-amarelo distrófico; Ag/Ne – Argissolos/Neossolos.

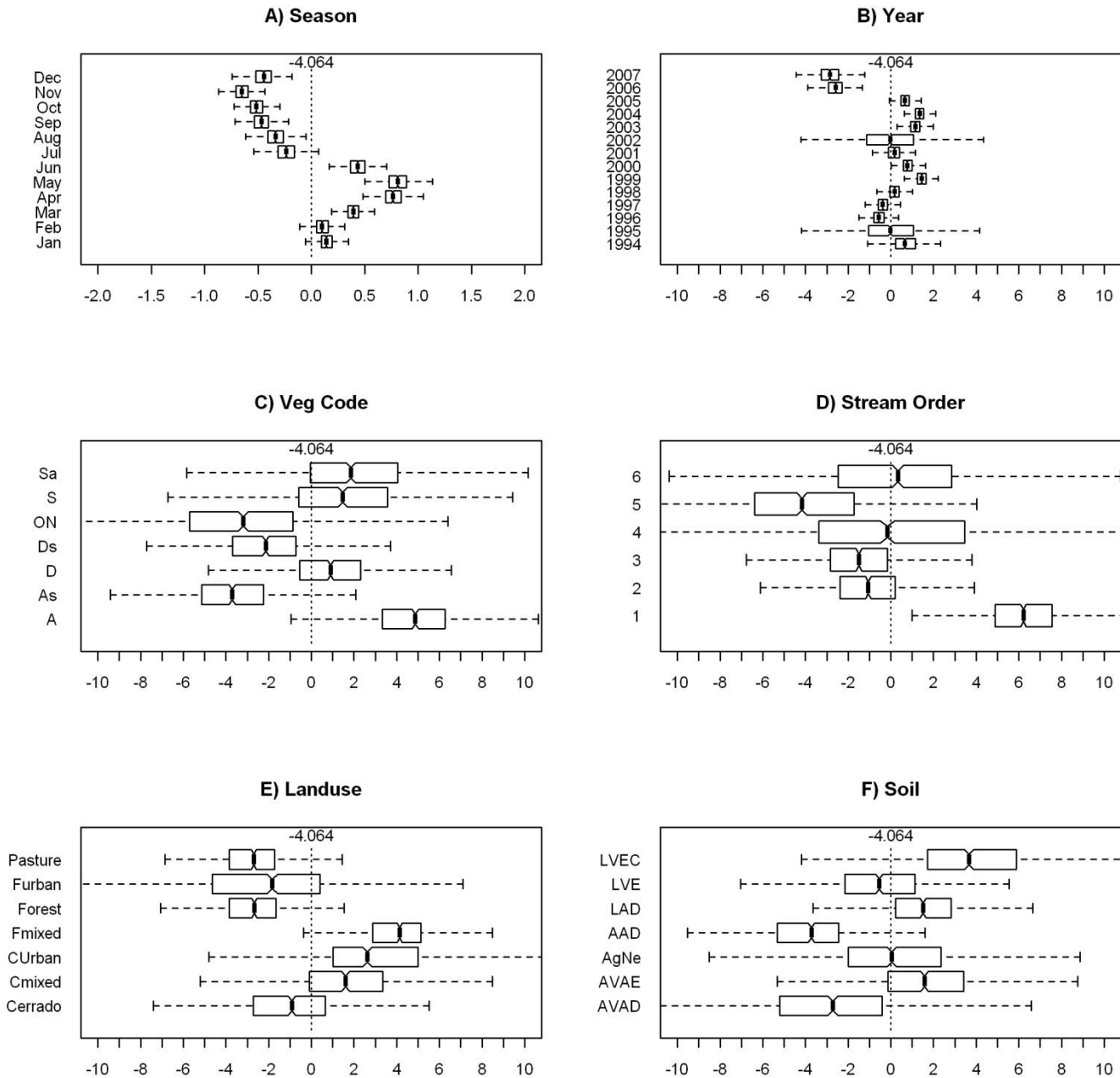


Figure 6 - Slope adjustments associated with the $\log_{10}\text{Flow}$ regression. The overall mean Slope is identified by the dotted line in each panel. A-Open tropical rainforest with secondary vegetation and agricultural activities); As-Open tropical rainforest – sub-mountain; ON-Ecotone tropical rainforest-seasonal forest; Ds-Dense tropical forest – submountain; D-Dense tropical forest –secondary vegetation and agricultural activities; Sa-Savannah-open woodlands; S-Savannah –agricultural activities). LAD – Latossolos amarelo distrófico; LVE- Latossolos vermelho escuro; LVE/C – Latossols vermelho escuro/Cambissolos; AVAE –Argissolos vermelho-amarelo eutrófico; AAD-Argissolos amarelo distrófico; AVAD - Argissolos vermelho-amarelo distrófico; Ag/Ne – Argissolos/Neossolos.

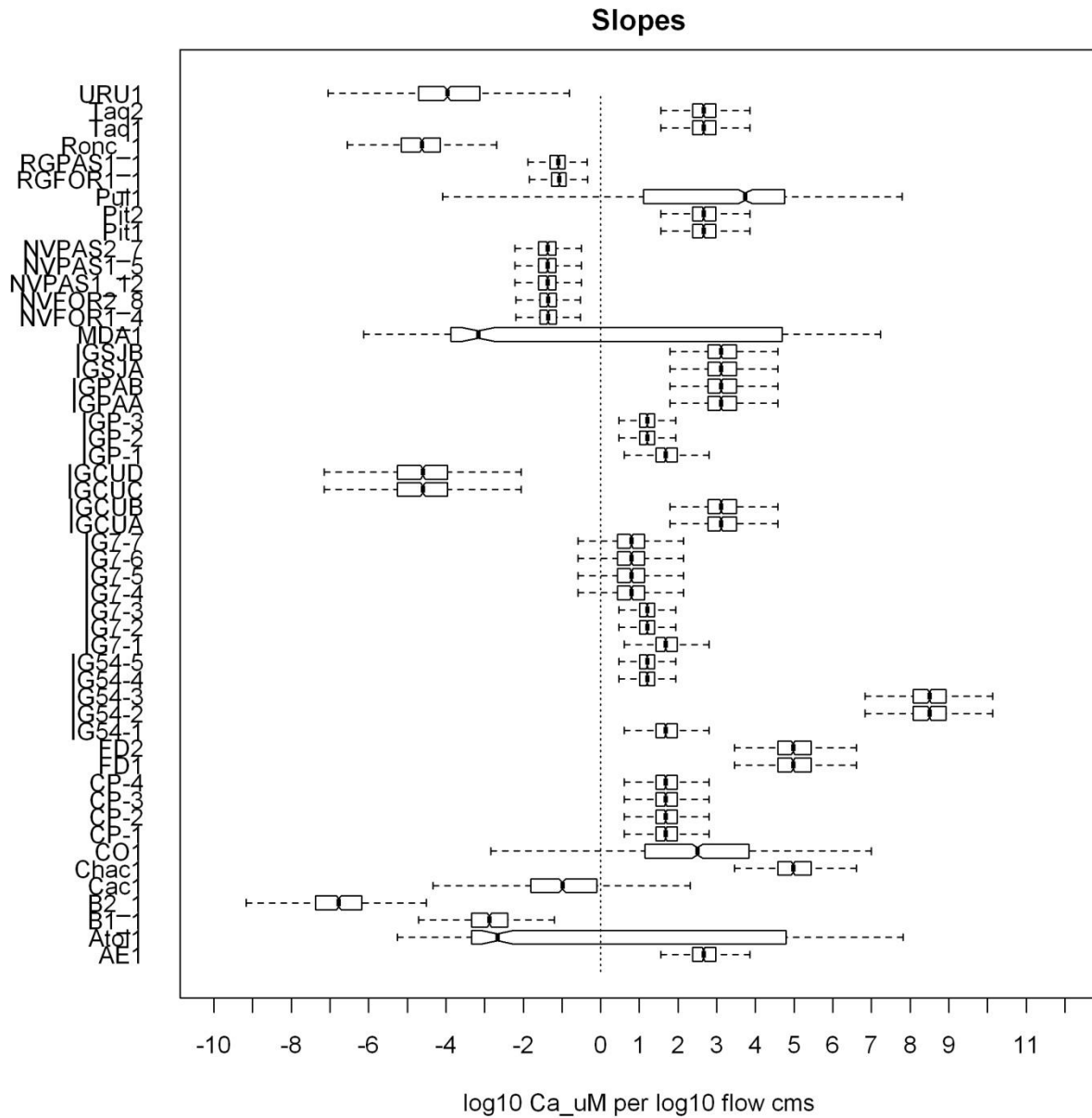


Figure 7. The log₁₀Ca vs log₁₀Q slopes for all stream stations predicted over all years from a multilevel model including adjustments for year, season, stream order, land cover, land use, and soil class. URU1-Urupa; Taq-Taquara; Ronc-Roncador; RGPAS-Rancho Grande Pasture; RGFOR-Rancho Grande Forest; Pul-Pulador; Pit-Pitoco; NVPAS-Nova Vida Pasture; NVFOR-Nova Vida Forest; MDA-Mestre D' Armas; IGSJ-São João; IGPA-Pachiba; IGP-Pajeu; IGCU-Camaru; IG7-Sete; IG54-Cinquenta e quarto; FD-Fazenda Dimas; CP-Capitão Poço; CO-Capão de Onça; Chac-Chacara; Cac-Ji-Paraná@Cacoal; B-Juruena; Atol-Atoliero; AE-Aguas Emendadas. Letters or numbers after abbreviations indicate stations within the stream.