Efficiency in inventories of ants in a forest reserve in Central Amazonia

Jorge Luiz Pereira de Souza(1), Carlos Alberto Ribeiro de Moura(2) and Elizabeth Franklin(1)

(1) Instituto Nacional de Pesquisas da Amazônia, Departamento de Entomologia, Avenida André Araújo, nº 2936, Aleixo CEP 69060-001 Manaus, AM, Brazil. E-mail: jlpsouza@inpa.gov.br, beth@inpa.gov.br (2) Instituto Nacional de Colonização e Reforma Agrária, CEP 69306-665 Boa Vista, RR, Brazil. E-mail: moura.carlos@gmail.com

Abstract – The objective of this work was to evaluate an inventory method efficiency for ants. We used subsamples collected in 24 transects of 100 m, distributed in 6 plots of 600 ha each in primary forest, as part of a long-term project. Ten litter subsamples were extracted per transect using Winkler extractors. Ants were identified to genus level, and Crematogaster, Gnamptogenys and Pachycondyla genera to species/morphospecies level. To evaluate the consequences of reduced sampling on the retention of ecological information, we estimated the lowest number of subsamples needed to detect the effects of environmental variables. Multidimensional scaling (MDS) was used to generate dissimilarity matrices, and Mantel correlations between each reduced-sampling effort and maximum effort were used as an index of how much information was maintained and could still be used in multivariate analyses. Lower p-values was observed on the effect of soil pH in the community of genera, and on the effect of the litter volume for the community of Crematogaster. The trend was still detectable in the analysis based on reduced-sampling. The number of subsamples can be reduced, and the cost-efficiency of the protocol can be improved with little loss of information.

Index terms: Formicidae, environmental impact assessment, soil biodiversity, Winkler extractors, tropical forest, sampling protocols.

Introduction

On Amazonian dry land forest soils, the number of ants can reach about 10% (~600 individuals m$^{-2}$) of the total soil arthropod abundance (Adis & Schubart, 1984). Thus, associated with the observational scale and the taxonomic challenge, sampling and identifying ants is time-consuming and expensive. Nevertheless, many sampling programs establish an initial protocol that results in a huge amount of biological material, requiring the involvement of many people and a high number of scientist-hours to process the samples (Lawton et al., 1998; Purvis & Hector, 2000). As a consequence, the cost may be too high, limiting the work to a few years.

(Danielsen et al., 2003), or resulting in the complete exclusion of the hyperdiverse groups of invertebrates in traditional protocols.

Several strategies have been proposed, in order to reduce time and costs for hyperdiverse groups, such as the identification of taxa whose diversity is correlated to those of others (surrogates), coupled with: identification of sampling methods that produce diversity estimates, representative of more intensive sampling; the use of morphospecies inventories generated by nonspecialists (parataxonomists); selection of larger species, using presence/absence matrices; and genus richness as a surrogate for species (Longino, 1994; Andersen, 1995; Oliver & Beattie, 1996; Longino & Colwell, 1997; Andersen et al., 2002), among others.

The reduction of the sample effort applied in intensive inventories is crucial, and the biggest challenge is to pursue this reduction retaining sufficient ecological information to capture gradients associated with environmental variables, as the use of ants as bioindicators, which is more effective when supported by their relationship with the environment. We focused on the ants identified to genus, and three genera identified to species (Crematogaster, Gnamptogenys and Pachycondyla), to investigate methods for reducing the time spent to process the ants, to determine the reduction in cost, and the loss of information involved. Our objective was to conduct a pilot study, in order to identify sampling protocols that are cost-effective for large-scale biological surveys. The strategy was to reduce the effort by diminishing the number of subsamples per transect in the field. Therefore, as the aim was not merely reduction in cost, but also retaining sufficient quality of information, we evaluated the reduction of information looking at the reduction in multivariate dissimilarity between plots, and also evaluated the ability of multivariate ordination methods based on the reduced subsamples to capture gradients associated with ecological variables.

Materials and Methods

As part of the ant protocol of a long-term field experimental area of the project Tropical Ecology Assessment and Monitoring (TEAM), the study was undertaken in a continuous primary forest of the Estação Científica Ferreira Penna (ECFP), an ecological station in the municipality of Melgaço (1º30’ to 1º50’S; 51º15’ to 51º45’L), at 400 km west of Belém, in the Brazilian state of Pará (Lisboa et al., 1997) (Figure 1 A). This program is an initiative of the Center for Applied Biodiversity Science (CABS) of the Conservation International (CI), and is financed by the Gordon and Betty Moore Foundation. The region is dominated by dense primary ombrophilous lowland rainforest (terra firme) (Almeida et al., 1993) (Figure 1 B). The terrain consists of Cretaceous sedimentary rocks (Alter do Chão Formation of the Amazon Basin) (Kern, 1996). Near and at the surface, the lateritic profiles have been converted into Oxisols (Yellow Latosols in the Brazilian classification), which are widespread over the region. These Oxisols have clayey to sandy texture, are deep, acidic and oligotrophic (Almeida et al., 1993). The climate of the region is classified as humid tropic of the type Am in the Köppen’s classification (Moraes et al., 1997). Annual mean precipitation is about 1,920 mm, with the peak of the rainy period registering on average 300 mm per month, from February to April. Climate is characterized by a dry season between July to December, and the driest months are October and November, with an average rainfall of 60 mm per month (Costa et al., 2003).

The ants and other invertebrates were collected by using the Winkler method, in 24 transects of 100 m, distributed in six plots of 600 ha (Figure 1 C), spread over 33,000 ha of primary forest located in the ecological station. Useful for collecting litter and soil fauna, this method is recommended for studying forested habitats (Longino & Nadkarni, 1990; Olson, 1991; Bestelmeyer et al., 2000; Vasconcelos & Delabie, 2000), and should be included in all ground-dwelling ant protocols (Delabie et al., 2000). Each plot had a trail system which divided it into 100 subplots (100x100 m), and four quadrants (500x500 m) (Figure 1 D). Four randomly placed collection transects of 100 m were installed, one per each quadrant of the plot. The subsamples were taken in the dry season, during October 26 to November 3. At 10 m intervals along each transect, a 1 m² area was used to sample litter (approximately 3–15 L). The leaf-litter was collected from the delimited area, and ants were extracted from the sieved litter in Winkler bags, through a sieve of 1 cm² mesh size (Majer et al.,...
The principle of this method is that ants and other invertebrates from within the litter sample migrate out of the litter as a behavioral response to disturbance of their habitat (Delabie et al., 2000), like humidity reduction, and fall into the pot partially filled with alcohol, suspended at the bottom of the bag. The set of Winkler subsamples was suspended for 48 hours. Before the set was suspended, the litter material was mixed to improve chances for the ants falling into the collecting pot (Parr & Chow, 2001). The sieved litter was placed into a mesh bag suspended inside of a cotton bag. The extracted invertebrates were preserved in 75% ethanol with 5% glycerin.

The genera *Crematogaster*, *Gnamptogenys* and *Pachycondyla* were identified to species, whenever possible, or to morphospecies. The species were identified using taxonomic keys made by Lattke (1990, 1995), Longino (2003) and Mackay et al. (2007). Voucher specimens were deposited in the Entomological Collection, of the Museu Paraense Emílio Goeldi (MPEG) and in a reference collection at INPA, Manaus.

To simulate reduced field collecting effort, we considered the 24 transects as units of replicates. The aim was to test if the composition of species in the taxonomic inventory would be represented in fewer subsamples per transect. To determine how much information was lost by reduced sampling, and how much could still be used in multivariate analyses, the dissimilarity between transects was calculated for qualitative data (presence-absence) by using the Sørensen dissimilarity index (Purvis & Hector, 2000). Presence-absence measures give more weight to rare species, because common species tend to occur in all transects in high numbers, and is highly recommended for description of ant distribution due to their social behavior (Longino, 2000). Dissimilarity matrices were generated for the maximum sampling effort (M; 10 subsamples/transect) and for the reduced sampling effort (M – n), for each level of reduction separately. Reduced sampling efforts ranged from only one subsample per transect (considered the minimum effort; M – 9), to 9 subsamples per transect (M – 1). Ten replicates were undertaken for each simulated effort, to evaluate

**Figure 1.** Map of the Estação Científica Ferreira Penna (A); view of the primary ombrophilous forest (B); distribution of the six 600-ha plots of the TEAM project, in the ecological station (adapted from Martins et al. 2005) (C); and grid dividing each plot in 100 subplots, and the position of the transect (100 m) in the four quadrants (500x500 m) (D).
the mean and variability introduced by subsampling. The dissimilarity matrix of the maximum effort was compared with the matrices of each simulated effort using Mantel correlation. The value of the Mantel correlation indicates the loss of information with the reduced effort and ranges from -1 to +1 (Chust et al., 2003). The Mantel correlations were plotted against sampling effort, and Lowess regressions (tension = 0.5) were used to produce a smooth curve. In the symmetrical dot displays graphics of the values for correlations, the position of the dots of each row in relation to the y values showed the loss of information obtained with Mantel. We assumed that Mantel correlation higher than 0.7 represent a dissimilarity value among transects that indicates a satisfying reduced sampling in the inventory, and more than 30% loss of information was not considered.

To determine the effect of information loss, multivariate ordinations based on genera or species composition data were generated. Ordination axes are, by definition, linear and meet the assumptions of dependent variables for multivariate regression. Project researchers have evaluated some environmental variables like soil physics and chemistry, volume of litter, canopy opening, and structure of the vegetation (Table 1). We tested the localization of samples along the ordination axes (dependent variables) against some of chosen, independent variables (volume of litter, pH of the soil and percentage of sand) because initial analyses indicated different degrees of effect on the species composition. As these independent variables were pre-selected in our study, they were not used for tests of specific hypotheses. They were used merely to show whether an effect detected in one analysis could be detected in a subsequent analysis based on reduced-effort sampling. An evaluation of the environmental variables affecting the ant community will be presented elsewhere. Multidimensional scaling (MDS) analyses were conducted in package PATN (Belbin, 2004). We made a priori decision to reduce the dimensionality to only one axis.

The inventory was done in 8 days, in a rate of 3 transects/day to collect 240 subsamples, not counting transportation to the field station. To analyze the loss of information due to reduction in number of subsamples, we used the 42 genera of ants, and then the species communities of Crematogaster, Gnamptogenys and Pachycondyla (Souza et al., 2007) for a complete list of genera and species or morphospecies.

Cost estimates were based on laboratory-processing time (person-hours) required for processing the samples. Therefore, the reduction in processing time was evaluated with the data of 10 or less samples per transect. For the estimations of effort reduction, we considered only the results with Mantel correlation lower than 0.7. The material was sorted at US$ 297 per month, during 2.5 months with 160 hours/month.

**Results and Discussion**

Considering the 24 transects and the community of genera, the values for Mantel correlation ranged from 0.38 to 0.94. The values of the correlation remained high, with six (average of 0.73) or more subsamples per transect (Figure 2). The consistency of the sampling set considering lower number of subsamples was satisfactory, and the composition of the community of genera, in the matrix of taxonomic inventory, could be represented in fewer subsamples.

The values for Mantel correlation for the three communities of species ranged from 0.3 to 0.85 for Crematogaster, 0.06 to 0.98 for Pachycondyla, and 0.05 to 0.93 for Gnamptogenys. The composition of species of Crematogaster is well represented with six or more subsamples per transect. Gnamptogenys and Pachycondyla communities are well represented with eight or more subsamples (Figure 3).

The probability values associated with three independent variables (sand percentage, litter volume and

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Range(1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Air temperature (°C)</td>
<td>28.3</td>
<td>26.40–30.60</td>
</tr>
<tr>
<td>Air humidity (%)</td>
<td>90</td>
<td>79.0–96.0</td>
</tr>
<tr>
<td>Soil temperature (°C)</td>
<td>25.93</td>
<td>25.0–26.40</td>
</tr>
<tr>
<td>Soil pH</td>
<td>4</td>
<td>3.40–5.30</td>
</tr>
<tr>
<td>Altitude (m a.s.l.)</td>
<td>76</td>
<td>39.4–109.0</td>
</tr>
<tr>
<td>Inclination</td>
<td>9.9</td>
<td>0.7–27.0</td>
</tr>
<tr>
<td>Clay (g kg⁻¹ of soil)</td>
<td>348.8</td>
<td>0.0–860.0</td>
</tr>
<tr>
<td>Sand (g kg⁻¹ of soil)</td>
<td>162.9</td>
<td>10.0–340.0</td>
</tr>
<tr>
<td>Litter volume (L)</td>
<td>8.99</td>
<td>3.00–15.50</td>
</tr>
<tr>
<td>Canopy opening (%)</td>
<td>19.5</td>
<td>3.64–26.52</td>
</tr>
</tbody>
</table>

(1) Range of the observed values between September and October 2003.
soil pH) were obtained for the communities of ant genera and *Crematogaster* taxa, to verify how well the reduced effort of 6 subsamples reproduced the results obtained for the maximum effort (10 subsamples). A consistent trend of lower values was detected in capturing the effect of soil pH for the community of genera, and litter volume for the community of *Crematogaster*. The trend was still detectable in the subsequent analysis based on reduced sampling (Table 2).

Our results suggest that the decision on how many subsamples to collect, in order to maintain enough information to detect the effects of environmental variables, is also dependent of the taxa being studied. The number of species of one taxon is not the unique condition to reduce the number of subsamples. There was a relationship between the frequency of individuals of each genera and the number of subsamples that could be diminished from the initial protocol of the long-term experimental area. In our study, *Gnamptogenys* (10 taxa) and *Pachycondyla* (6 taxa) were less frequent, compared to *Crematogaster* (6 taxa). Considering the species sampled with pitfall traps in the same area and period, the total of *Crematogaster, Gnamptogenys*

![Figure 2](image1.png)

**Figure 2.** Values for Mantel correlations between subsamples per transect, for the composition of Formicidae genera sampled with Winkler extractors. Continuous lines are Lowess regressions (tension = 0.5).

![Figure 3](image2.png)

**Figure 3.** Values for Mantel correlations between subsamples per transect, for the community of species of *Crematogaster, Gnamptogenys* and *Pachycondyla* sampled with Winkler extractors. Continuous lines are Lowess regressions (tension = 0.5).
and *Pachycondyla* taxa rises to 7, 14, and 15 (Souza et al., 2007), confirming the relatively lower diversity of *Crematogaster* in the area. We suppose that the sampling results were influenced by size of colony (Lattke, 2003) and the predatory behavior of *Gnamptogenys* and *Pachycondyla* (Andersen, 1997). Also, individuals of these both genera normally forage alone. *Crematogaster* have a massive foraging activity and has a higher frequency and number of individuals (Longino, 2003).

A lower number of sub-samples was satisfactory to represent the assemblage of genera of ants, and also the assemblages of *Crematogaster*, *Gnamptogenys* and *Pachycondyla*, keeping more than 70% of the information. As a result, processing costs and time could be reduced (Figure 4). We estimate the reduction of time and costs in the laboratory as a direct consequence of the reduction in the number of samples in the field. In general, eight samples per transect were satisfactory for all taxa analyzed, resulting in 20% sorting effort reduction (Figure 3), and the processing hours of laboratory work would be reduced from 800 to 640. For the communities of ant genera and *Crematogaster* species, the processing time would be reduced in 40%, considering six samples per transect, representing only 480 work hours. The estimated economy in time and expenses would be around 10%, because reduction in the number of subsamples is likely to have only a small effect on field expenses.

Ecologists and taxonomists differ in their units used to quantify biodiversity, but data have to be collected and processed in such a way that the sample can be validly compared statistically with subsamples collected in the same way at different places or times (Chust et al., 2003). Therefore, many replicates will result in a huge amount of material to be sorted. We collected around 2,188 L of litter material that passed through the sieves of the Winkler extractors, resulting in 144 L of finer litter material used to extract the invertebrates. The “ants methodology protocol” was planned to have four sampling periods a year. Considering 0.6 L of finer litter material, inside each cotton enclosure of the Winkler extractor, and having 240 sample units sampled four times a year, in 10 years it would result in 5,760 L of litter material, collected in 600 ha (0.96 L ha\(^{-1}\) per year). As a comparison, 246 L of litter was collected in 1,500 ha (0.11 L ha\(^{-1}\) per year), during 14 years, in “La Selva Ecological Station” (Longino et al., 2002).

Considering that only one of the four sampling periods a year was evaluated in the long-term field experimental area of the Project TEAM, in the Brazilian state of Pará, the laboratory work for sorting material sampled over a relatively large area of 600 ha was expensive. The costs involved (higher than 1,300 US$) processing 240 subsamples can be reduced without losing much information, as shown by this study. The majority of this money was used for person/hours in the laboratory. Costs for such projects are cumulative and, as a consequence, according to Evans & Viengkham (2001), the time cost of these surveys will outweigh any economic benefits that they may offer.

Our results showed reduction in the costs and time involved in mounting, labeling, and identifying ant specimens in the laboratory. Unlike other taxa Table 2. Probabilities estimated by multivariate multiple regression relating three variables to the first two multidimensional scaling axes for ordination, based on ten and six subsamples per transect, for the communities of genera and of species of *Crematogaster*.

<table>
<thead>
<tr>
<th>Community</th>
<th>Number of subsamples</th>
<th>Predictor Variables</th>
<th>R²</th>
<th>P</th>
<th>F</th>
<th>T</th>
<th>p (partial)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genera</td>
<td>10</td>
<td>Litter volume</td>
<td>0.164</td>
<td>0.301</td>
<td>1.304</td>
<td>-1.442</td>
<td>0.165</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>pH of the soil</td>
<td>0.168</td>
<td>0.288</td>
<td>1.346</td>
<td>-0.694</td>
<td>0.496</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Sand percentage</td>
<td>0.298</td>
<td>0.065</td>
<td>2.826</td>
<td>-2.135</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Litter volume</td>
<td>0.308</td>
<td>0.057</td>
<td>2.961</td>
<td>-2.407</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>pH of the soil</td>
<td>0.308</td>
<td>0.057</td>
<td>2.961</td>
<td>-2.407</td>
<td>0.026</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>Sand percentage</td>
<td>0.308</td>
<td>0.057</td>
<td>2.961</td>
<td>-2.407</td>
<td>0.026</td>
</tr>
</tbody>
</table>

as pteridophytes in a primary forest in Central Amazonia, whose amostral effort reduction resulted in a significant saving of time in the field (Zuquim et al., 2007), sampling ants in the field using Winkler method is relatively quick and with fixed costs, such as transportation, meaning that the costs are unlikely to be much reduced. The same pattern observed for ants was also observed for oribatid mite assemblages, in an Amazonian savanna, by Santos et al. (2008). Although these authors had detected a substantial reduction in time and costs in the laboratory, the costs in the field did not reduce much. Otherwise, economy in the field, that we estimated to be around 10%, can be substantial in a large-scale sampling program. Added to the number of temporal replicates to be taken a year by the project, such economy would reduce significantly the effort demanded by the intensive ant sampling program.

Therefore, we must consider adaptive procedures in the laboratorial work. The desirable percentage of ant species captured by a certain amount of subsamples in an inventory must be good enough to show the relationship of this assemblage to the environmental variables, and if so, we must investigate if an effect detected in one analysis can still be detected in a subsequent analysis of reduced effort.

Also, questions need to be asked separately for species groupings and different geographic regions. Other experiments are necessary and should encompass replicates of several sampling periods, and comparisons or combinations of other sampling methods, like pitfall traps and baits, to verify how successfully we would reproduce the same tendencies here observed.

Conclusions

1. It is possible to reduce the number of samples in traditional inventories and to optimize the initial protocol of the long-term project to increase field and laboratory cost-efficiency, which can allow sampling a higher number of plots, additional replicates, and hence the number of species encountered per unit effort, increasing the statistical power and the generality of the results.

2. The sampling effort reduction does not affect the capacity to detect the effect of soil pH for the
community of genera, and of litter volume of the *Crematogaster* community.

**Acknowledgements**

To the Tropical Eadogy Assessment and Monitoring Initiative and the program for Biodiversity Research, for financial support; to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Apoio à Pesquisa do Estado de Amazonas (FAPEAM), for scholarships; to William Magnusson and Flávia Costa, for assistance in statistical analyses; to Heraldo Vasconcelos, for helpful comments; and to Ana Harada, for help with the identification of ants. Financial support came from project TEAM. The work was partly funded by Program for Biodiversity Research, and scholarships from CNPq and FAPEAM.

**References**


