An exercise in archaeological demography: estimating the population size of Late Neolithic settlements in the Central Balkans

Marko Porčić
Department of Archaeology, Faculty of Philosophy, University of Belgrade, RS
mporcic@f.bg.ac.rs

ABSTRACT - This paper reflects on the methodology for estimating population size from settlement data. Archaeologists are faced with a static record of houses, which is the result of dynamic processes of population growth and house use. There is no simple relationship between the total number of houses and population size. In order to produce more realistic estimates of population size, a quantitative model is presented which takes into account population dynamics and the dynamics of house accumulation. The model is used to estimate the population size of three Late Neolithic settlements in Southeast Europe: Gomolava, Divostin and Uivar.


KEY WORDS – Late Neolithic; Vinča culture; demography; population size; settlements

Introduction

Population size is considered to be an important variable in most anthropological and archaeological theories. Its role is especially prominent in approaches grounded in processual archaeology and evolutionary archaeology. In the processual perspective, its importance stems from the fact that population size (and density) is highly correlated with socio-cultural complexity and plays a key role in domains related to subsistence and cultural ecology in general (Binford 2001; Carneiro 1962; 1986; 2000; Ember 1963; Feinman 2011; Johnson and Earle 2000; Johnson 1982; Kosse 1990; 1994; Peregrine et al. 2004). In evolutionary archaeology, population size is important because cultural evolutionary theory views culture as a population phenomenon – it deals with processes at the population level (Boyd and Richerson 1985; O’Brien and Lyman 2000; Richerson and Boyd 2005; Shennan 2002). Given this great theoretical importance for archaeological research, various methods have been developed to estimate population size and changes in population size from archaeological data (e.g., Chamberlain 2006; Hassan 1978; Schacht 1980; 1981; Shennan, Edinborough 2007). One of the most frequently employed methods is to estimate settlement population size on the basis of house floor or living floor area (Brown 1987; Casseberry 1974; Dohm 1990; Kolb 1985; Kramer 1982; LeBlanc 1971; Naroll 1962; Wiesner 1974).

DOI: 10.4312/dp.38.25
However, very few attempts have been made to estimate population size of late Neolithic settlements in the central Balkans (e.g. Chapman 1981; Müller 2006; 2007; Porčić 2010). One reason for this lack of research in the domain of archaeological demography has primarily to do with the fact that the main stream of Balkan archaeology has been dominated by an outdated version of a cultural-historical approach. The other reason is that demographic reconstruction is inherently difficult, given the many factors that need to be taken into account in order to arrive at a sound population estimate. On the optimistic side, very rough estimates (e.g., order of magnitude) of population size are quite sufficient for most archaeological purposes (Drennan, Dai 2010). Apart from the methodological problems, the estimation of population size is difficult because data requirements are high. In the central Balkans, very small portions of late Neolithic sites have been excavated. Small-scale excavations produce small and biased samples, and any estimates based on such samples would be unreliable.

In recent times, geophysical surveys have become more frequent in field projects with a focus on the late Neolithic of the central Balkans and Southeast Europe in general (Crnobrnja, Simić 2008; Crnobrnja et al. 2010; Drașovean 2007; Müller 2007). These methods have enabled archaeologists to gain significant insight into the architectural contents of settlements – e.g., number of houses and their spatial distribution. In this way, data is produced at relatively low cost in comparison with large scale excavations.

The first systematic estimates of the size of late Neolithic Vinča settlement populations were made by Chapman (1981). On the basis of total site areas, Chapman estimated that populations sizes ranged from 30–300 to 1000–2500 people (Chapman 1981, 48). Müller used data from a geophysical survey to estimate late Neolithic Butmir population size in Okolište and the entire Visoko basin – the population of Okolište was estimated to be 1000 people, while the estimate for the entire Visoko basin was 3500 (Müller 2006; 2007). These estimates were made on the assumption that the entire site area was used simultaneously (Chapman), or that the houses from one building horizon were mostly contemporaneous (if I understand the procedure presented in Müller 2007, 26).

The problem with these procedures is the assumption that the accumulated houses or site space in general were contemporaneous. This is a more or less justified assumption in cases where there is independent evidence that horizon durations were short (e.g., 14C dates), but it is clearly erroneous in cases where settlements lasted for a longer period, not to mention the problem of horizontal stratigraphy. Even with short-lived sites, the question of how short this period needs to be for an approximation of contemporaneity to hold remains. The answer may seem simple: it needs to be less or equal to the average use-life of a house. But the problem is that the average use-life of the house is unknown.

Porčić tried to resolve the contemporaneity issue by applying Schiffer’s discard equation (Schiffer 1972; 1976; 1987) to Vinča houses (Porčić 2010). However, this approach is also flawed. It rests on the assumption that the number of houses (and population) was more or less constant during the life of the settlement, and an attempt was made to estimate this average number of houses used contemporaneously. Again, this approach may make sense if the site duration was relatively short and growth rates were very low, but the greatest drawback of this procedure is that the average use-life of a house needs to be estimated. Unfortunately, the estimate of the average use-life of houses is little better than an educated guess in Porčić’s case.

To summarise, the methods used so far have not been entirely appropriate for estimating population size because they failed to make a systematic distinction between the systemic and archaeological ‘assemblages’ of houses (Schiffer 1972; 1976; 1987). Even when the distinction was made, as in the case of Porčić’s study, the population dynamics model was not realistic.

This paper attempts to contribute to the issue of estimating population size in Late Neolithic settlements in the central Balkans by directly addressing the issue of population dynamics and house accumulation dynamics. It will seek answer to this particular question: how can archaeologists estimate the final population size on the basis of the number of houses present in the archaeological record? The methodology for this problem will be formulated and applied to the following sites: Gomolava, Divostin and Uivar. Even though the quality of the data from these sites is far from ideal and the estimates will be approximate, the methodological insights gained from this study and its potential relevance for future work should justify the effort.
Modeling population and house accumulation dynamics

The first step in trying to develop a method for estimating population size on the basis of house remains is to understand population dynamics. In other words, the question is: how does population size change through time? In the case of the late Neolithic in the central Balkans, we can be fairly certain that the population size grew at both regional and settlement levels (Kaiser, Voytek 1983; Porčić 2010; Tringham and Krstić 1990). If the reasonable premise of growing population is accepted (at least, for this particular case), the next question is: how did population grow? This means that we have to determine the form of the population growth model and its parameters. The first part is less difficult, since most human populations obey a logistic growth model in the long run (Chamberlain 2006.21–23; Schacht 1980). The logistic model of population growth has the following form (Schacht 1980.786):

$$P_t = \frac{K}{1 + \left(\frac{K - P_0}{P_0}\right)e^{-rt}} \quad (1)$$

where $P_t$ is the population size at time $t$, $P_0$ is the initial population size, $r$ is the maximum growth rate and $K$ is the maximum population size possible (carrying capacity).

Once the model for population dynamics is formulated, two things remain: 1) to model the link between the current population size and the number of houses, and 2) to model the accumulation of houses in the archaeological record during a certain period.

The first is relatively easy to model. It is assumed that the number of houses depends on household size. Therefore, the equation which describes the systemic number of houses at any moment in time is simply:

$$H_t = \frac{K}{1 + \left(\frac{K - P_0}{P_0}\right)e^{-rt}} \quad (2)$$

Where $H_t$ is the systemic number of houses in time $t$, and $m$ is the reciprocal of the household size. The number of houses entering the archaeological record at any time $t$ depends primarily on the house use-life and is equal to:

$$\frac{dH_t}{dt} = H_t \frac{1}{L} \quad (3)$$

Where $Ha$ is the number of houses in the archaeological record at time $t$ and $L$ is the average use-life of a house. This equation provides a direct link between population growth and house accumulation. The total number of houses ($H_{total}$) accumulated on a site from time $t_0$ to time $t$ can be calculated as follows:

$$H_{total} = \int_{t_0}^t H_t \frac{1}{L} dt + H_t = \left[ \frac{K}{1 + \left(\frac{K - P_0}{P_0}\right)e^{-rt}} \right]_{t_0}^t \quad (4)$$

The first term of the formula is simply the sum of all the increments of house accumulation, while the second term represents the final systemic number of houses which instantly enter the archaeological record after the settlement is abandoned. It is assumed that the settlement is abandoned rapidly.

How does this model help us to estimate the population size and the number of contemporary houses in archaeological situations? If all the parameters which appear on the right side of Equation 1 were known, there would be no need for the model presented by Equation 4, since we could simply calculate the population size directly using Equation 1. However, the situation is such that the only variable which can be directly observed in the archaeological record is the number of accumulated houses ($H_{total}$), while other parameters can only be estimated with more or less certainty. This paper proposes to use different combinations of reasonable parameter estimates in order to project different values for the number of accumulated houses ($H_{total}$). The parameter combinations which produce the closest fit to the actual $H_{total}$ can be used to estimate population size.

Data and methods

Data from the Late Neolithic Vinča culture sites of Gomolava, Divostin and Uivar will be used as an empirical basis (Fig. 1). It should be emphasised that the archaeological culture label is merely that – a technical label – the anthropological reality which stands behind this label should by no means be automatically equated with a single social, political, linguistic or ethnic unit (for general information on Vinča culture see Chapman 1981; Garašanin 1979; 1982).

The relevant temporal, spatial and archaeological data for each of these sites are summarised in Table
Average household size was estimated by dividing the average house floor area with an average value of floor area per person, which was estimated at 7m²/person (see Chapter 4 in Porčić 2010). The total number of accumulated houses for Uivar was estimated by the original investigator on the basis of the geophysical survey, and this estimate was used in this paper.

The total number of accumulated houses for Divostin and Gomolava had to be estimated by proportional projection. The total area of the Divostin site was estimated to be 15ha, 2480m² (1.65%) of which was excavated (McPherron and Srejović 1988). Two Vinča horizons were defined – Divostin IIa and Divostin IIb. These two settlements lasted altogether for about 300 years, from 4900–4650 calBC (Borić 2009). A total of 5 Divostin IIa and 12 Divostin IIb houses was uncovered completely or partially (McPherron and Srejović 1988). In order to estimate the total number of accumulated houses, a proportional projection was made. If both settlements covered the same area, the estimated total number of accumulated houses is 1028. However, it is unlikely that both settlements covered the same area, so the figure of 1028 should be regarded as the maximum number of accumulated houses. If we assume that the older settlement, Divostin IIa, covered only one fourth of the total site area, then the estimated number of accumulated houses is 801. This should be regarded as the minimum number of accumulated houses.

The total area of Gomolava tell was estimated to be 18400m² (van Zeist 2002), of which 5000m² (27.17%) was excavated (Brukner 1988). There were three Vinča culture horizons: Gomolava Ia, Gomolava Iab, and Gomolava Ib, spanning a period of c. 350 years, from around 5000 to 4650 calBC (Borić 2009). A total of 31 houses were uncovered at Gomolava. Since Gomolava is a relatively small tell site, and tell sites are usually confined to a specific area, there was no need to assume that older settlements were much smaller in area (although they were certainly smaller in terms of architectural density) than the latest phase. A proportional projection was made, and it was estimated that the total number of accumulated houses should be around 114. Uivar is a Vinča culture tell site situated in Banat in western Romania (Drasovean 2007; Schier 2006; 2008). It covers an area of 3ha, which have been surveyed with magnetometer in their entirety. As a result, a complete settlement plan of the last phase

**Table 1. Temporal, spatial, and archaeological information for sites mentioned in the text. Notes: a Based on measurements of house contours from the site plan published in Drasovean (2007).**

<table>
<thead>
<tr>
<th>Site</th>
<th>Location</th>
<th>Chronology</th>
<th>Area (ha)</th>
<th>Average household size</th>
<th>Estimated number of accumulated houses</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gomolava</td>
<td>Srem, Serbia</td>
<td>5000–4650 calBC</td>
<td>1.84</td>
<td>5</td>
<td>114±20</td>
<td>Borić 2009; Brukner 1988; van Zeist 2002</td>
</tr>
<tr>
<td>Uivar</td>
<td>Banat, Romania</td>
<td>4940–4800 calBC</td>
<td>9.5</td>
<td>6a</td>
<td>70–80</td>
<td>Drasovean 2007; Schier 2006; 2008</td>
</tr>
</tbody>
</table>
(dated from 4940–4800 calBC) is available for study. Schier counted more than 70 burnt houses on the plan (Schier 2008). In Schier’s opinion, this count underestimates the total number of houses, because many did not burn at all. Schier assumes that “most of the sediment forming the settlement mound consists of transformed building material”, and uses the total volume of the tell to calculate the total number of houses for all phases of Uivar (Schier 2008:56–57). The result of this calculation is the figure of 3500–4000 houses. However, in this paper, only the last phase of the settlement along with the count of burnt structures will be taken into consideration. The main reason for this decision is that the count of burnt houses is based on solid empirical data, while Schier’s estimate is based on the premise that most of tell’s volume comes from transformed building material. In any case, the Uivar population estimate calculated in this paper may be considered a minimum estimate.

The next step in the analysis is to choose the range of parameter values for Equation 4. There has been plenty of research attempting to estimate the growth rate of Neolithic populations (Bocquet-Appel 2002; Carneiro, Hilde 1966; Galeta, Bruzek 2009), but most of these estimates were based on the exponential model of growth rather than the logistic model. The main difference between these two models is that the growth rate in the exponential model is constant, while in the logistic model the effective growth rate is a function of the current population size (Mooney and Swift 1999). Therefore, the parameter in the logistic model should be interpreted only as the maximum possible growth rate. Galeta and Bruzek review some of the estimated growth rates from the literature and provide their own estimates for the exponential model. The estimated rates range from 0.001–0.03 (Galeta, Bruzek 2009). Given that the model used in this paper is logistic, it is reasonable to take into consideration even higher growth rates. Therefore, the range of possible growth rates which will be used in this analysis would be 0.0025–0.07 in increments of 0.001.

The next parameter is carrying capacity. Carrying capacity for Vinča culture sites in central Serbia was estimated at around 1400 people for the 3km catchment zone (Bankoff, Greenfield 1984). Therefore, the range for the parameter will be between 1000 and 4000 people in increments of 200.

The remaining parameter for the model is house use-life. Estimates of the average use-life for Central European Neolithic houses range from 20–50 years (Gerritsen 2008; Whittle 2003:140–141). Perhaps the best solution is to tie house use-life to the human generation length. The average human generation length is 28 years (Fenner 2005). Therefore, 3 discrete values will be used for the parameter: 28, 56 and 86 years. The initial population parameter will be in the following range: 10–500 people in increments of 20.

All combinations of parameters are used to calculate the expected totals of accumulated houses via Equation 4. For each site, 81 600 different parameter combinations will be used to project the expected number of accumulated houses. Models that have the best fit of the projected to the observed number of accumulated houses will be used to estimate the population size. The criterion for the best fit depends on the method used to estimate the total number of accumulated houses on the site. It may be assumed that these estimates are the most precise for sites where a geophysical survey was conducted (such as Uivar). Therefore, the best fit interval for Uivar is 70–80 houses. For Gomolava, the proportional projection is based on almost 30% of the site; therefore, the interval around the estimate should not be great, around ±20 houses. For Divostin, a very large interval has to be used – between 801 and 1028 houses.

Results

For Gomolava, there are 148 combinations of parameters which predict the observed number of accumulated houses (within the best fit criterion limits). Each of these models is based on a combination of the parameters presented in the previous section. Final population size estimates range from 69–285, with a mean of 153 people, standard deviation of 52.41. 95% of estimated population sizes are between 70 and 258 people. The distribution of estimates is shown in Figure 2.

There are 8496 models which predict the accumulated number of houses to be within the best fit interval for Divostin. The final population size estimates generated by these parameter combinations range from 633–3713. The mean and standard deviation of population estimate values are 1740 and 495.52, respectively. 95% of population estimates are between 868 and 2842 people (Fig. 3).

There are 333 models out of 81 600 which predict the accumulated number of houses within the specified interval for Uivar. Estimates range from 90–
325 people (Fig. 4). The mean is 194; standard deviation is 55.68; and 95% of estimates are between 92 and 309 people.

**Discussion and conclusion**

The methodology used in this paper makes it possible to reduce the number of potential population growth models and to arrive at a set of population size estimates by introducing the constraint that the candidate model must produce the observed number of accumulated houses. The precision of the estimate depends directly on the ability to determine precisely the total number of houses in the archaeological record. Moreover, estimates depend directly on the range of input values which are used as parameters. The ability to constrain the range of possible values is directly related to the precision of the estimate – more constrained ranges will produce more precise estimates.

Many of the parameter values used in this analysis may be regarded as *a priori* unlikely. For example, given that this is a deterministic model, a house use-life value of 84 years is very improbable. Moreover, initial population is not likely to be 100, 300 or 500 people for sites such as Divostin and Gomolava. For example, if we constrain the initial population at Divostin to between 10 and 50 people, assume that the most probable use-life value of houses is 56 years, and use a single estimate of carrying capacity of 2000 people – the result would be a narrower interval estimate (95% of estimates would be between 1926 and 1997 people, Fig. 5). Likewise, initial population is not likely to be 10 people for the latest Neolithic settlement at Uivar, where there is clear evidence of earlier occupation. Therefore, population estimates based on models which assume higher initial population size are more probable for Uivar. For example, if only the initial population size for Uivar is constrained between 100–200 people (which is a more realistic estimate), the resulting population estimates would range between 136 and 221 people. Carrying capacity can be estimated with greater rigor if there is good environmental data. Growth rates can be estimated independently from skeletal data (*Bocquet-Appel* 2002). All this information can be included to enhance the precision of the estimate. One way to do this would be to build a stochastic model – a numerical simulation where parameter values would be randomly generated from pre-specified probability distributions.

How different are the estimates from this study from estimates based on other procedures? It is clear that
they are much lower than estimates based on the assumption that most of the houses from one horizon were contemporaneous. For example, it is estimated here that in its final phase, Neolithic Gomolava was inhabited by approximately 150 people, 250 people at most. If the proportional projection was made on the basis of Gomolava Ib record, this estimate would be about 350–400 people (24 houses were excavated in the Gomolava Ib, and the excavated area is approximately one third of the entire site area). The error would even be greater in the case of Divostin. On the assumption that all or most Divostin Ib houses were used contemporaneously, one would reach estimates that would range between 5000–6000 people. On the other hand, estimates made using the model from this paper are usually higher than estimates calculated by Porčić (e.g., final population sizes for Gomolava Ib and Divostin Ib were estimated at 50 and 1000 people respectively, see Porčić 2010. 342, Tab. 7.6). This is because Porčić estimated the average number of inhabitants during a certain period, while this procedure estimates the final (maximum) population size. The latter figure makes more sense than average population size. The average population size estimates would make sense only in cases where growth rates are very low, or when the population is not monotonically increasing – e.g., when the end of the settlement was not abrupt.

This leads to the final and most important question: are there ways to test the population growth model? The most straightforward way to test the model would be to have independent data on demographic trends (e.g., from skeletal data) or to have a great number of dates from a single site (e.g., from a single horizon). If the growth model is correct, the number of samples from appropriate time periods should follow the logistic curve. The problem that may arise is that 14C dates are difficult to distinguish on time scales at which the settlement dynamic operates (e.g., 50–100 years for a single building horizon). One possible solution to this problem would be to use seriation (e.g., seriate house or pit assemblages), perhaps in combination with 14C dates, in order to achieve a fine-grained relative chronology.

The most important lesson to be learned from this exercise in archaeological demography is that population size estimation from settlement data is far from straightforward. The situation faced by archaeologists looking at a deceptively clear settlement plan produced by a large scale excavation or magnetometer survey is actually a textbook case of equifinality (see Rogers 2000). The same number of houses could indicate very different population sizes, depending on the demographic scenario which actually took place. Fortunately, not all scenarios are possible or equally likely; therefore, some can be eliminated. The remaining scenarios can be used as a basis for making population estimates. The need for high quality data is a constant imperative that needs no special elaboration. In this context, it is more important to stress the need for better models and better parameter estimates.

**Fig. 5. Distribution of population estimated for Divostin when the range of parameters is constrained to more realistic values.**
REFERENCES


An exercise in archaeological demography: estimating the population size of Late Neolithic settlements in the Central Balkans

nauka i umjetnosti Bosne i Hercegovine. Centar za balkanoška ispitivanja, Sarajevo: 79–212.


