

## **Analysis of wild boar (*Sus scrofa* L., 1758) distribution in Northeast of European Russia: A quantitative approach**

**Nickolai I. Markov, Nickolai D. Neifeld & Lyman L. McDonald**

**ABSTRACT.** Analysis of a species' distribution pattern in the marginal part of its geographic range is usually performed on a qualitative level. At the same time, quantitative estimation of area of occupancy could be of great importance for distinguishing the sites inhabited by the species from those where it is "occasionally" found. For large mammals in areas of current or recent expansion an analysis of the distribution patterns is especially complicated by high mobility of the animals, lack of data about habitat preferences and, consequently, the impossibility of using the assumptions of "closed-population models" for estimating area of occupancy. In this study we report a retrospective quantitative analysis of wild boar distribution in one of the areas of European Northeast Russia based on records of occasional detections for the period 1984–1999. Indices of "minimum site occupancy" and "constancy of species' presence" (the proportion of years, when the species has been recorded within the study area) were used for estimation of species "extent of presence" for the whole study area and distinct parts of it. The wild boar is "vagrant" or not present in most areas situated close to the Ural Mountains. In the plains areas, the species has at least "visitor" status. Thus, only about 50% of the study area could be treated as part of wild boar geographic range. The quantitative method developed allows estimation of the species' status within a study area and avoids biases in analysis of resource selection by excluding the territories where it is "occasional" or "vagrant".

**KEY WORDS:** wild boar, taiga, distribution, site occupancy, constancy of presence.

*Nickolai I. Markov [nimarkov@ipae.uran.ru], Institute of Plant and Animal Ecology, Urals Branch of Russian Academy of Sciences, 8 Marta 202, Ekaterinburg 620144, Russia; Nickolai D. Neifeld, Pechoro-Ilychsky State Biosphere Nature Reserve, Yaksha, Troitsko-Pechorskii region, Komi Republic 169436, Russia; Lyman L. McDonald, Western EcoSystems Technology, Inc., 2003 Central Avenue, Cheyenne, Wyoming, USA.*

## **Анализ распространения кабана (*Sus scrofa* L., 1758) на северо-востоке Европейской части России: количественный подход**

**Н.И. Марков, Н.В. Нейфельд, Л.Л. Макдоналд**

**РЕЗЮМЕ.** Анализ распространения вида на периферии ареала обычно проводят в качественных терминах. В то же время, количественная характеристика области распространения вида может иметь большое значение для разделения устойчиво заселенных участков и районов "отдельных встреч". Для крупных млекопитающих в районах недавней экспансии анализ закономерностей пространственного распределения осложняется высокой подвижностью животных, отсутствием данных о биотопических предпочтениях и, как следствие, невозможностью применения модели "закрытой популяции" для количественной характеристики области распространения. В нашей работе мы приводим результаты количественного анализа распространения кабана в одном из районов европейского северо-востока России, основанного на данных о случайных встречах вида в период 1984–1999 гг. Для характеристики "степени присутствия" вида в целом для района исследования и отдельных его частей были использованы индексы "минимальной доли заселенных участков" и "постоянства присутствия вида". Дикий кабан имел статус "случайного" вида или отсутствовал в большинстве участков, расположенных вблизи Уральских гор. В равнинной части района исследований вид имел статус как минимум "регулярно встречающегося". Таким образом, только около 50% района исследований можно отнести к ареалу вида. Разработанный подход позволяет проводить количественную оценку статуса вида в пределах района исследований и позволяет избежать ошибок при оценке биотопических предпочтений вида путем исключения из рассмотрения зоны "отдельных заходов".

**КЛЮЧЕВЫЕ СЛОВА:** кабан, тайга, распространение, заселение участков, постоянство присутствия вида.

## Introduction

The size of the distribution area is an important parameter describing a species on the evolutionary scale (Brown, 1984). The pattern of distribution indicates, in particular, the spectrum of habitats that could be used by the species and the adaptation to those habitats (Swihart *et al.*, 2003). However, drawing the boundary of the distribution area is difficult, mainly because in the periphery of a species' range, individuals are "rare." Researchers typically face very fragmented data on the periphery concerning the pattern of occurrence and abundance of the species. The situation is especially complicated when a strong trend for expansion of the species' range has been reported (Hanski, 1999). This is, particularly, the case of wild boar *Sus scrofa* in the taiga zone of the Northeast of European Russia.

Wild boar expansion to the taiga zone has been reported since the 1940s (Fadeev, 1987). Supposedly, the main reasons for this expansion were: growth of the species' population in the broad-leaf forest zone, particularly in Byelorussia and the Baltic states (Rusakov & Timofeeva, 1984), a series of releases of wild boar in central Russia (Pavlov *et al.*, 1974), and alteration of natural communities by man (Fadeev, 1987). Wild boar started settling the Northeast of European Russia (Komi Republic) in about 1972 (Neifeld, 1998). The species has been detected in all parts of the region, however only a portion is treated as constantly inhabited by the species (Markov *et al.*, 2004). Proper description of wild boar distribution in European Northeast Russia requires search for the areas actually inhabited by the species within this region, i.e. an analysis of the species' "area of occupancy" (Gaston, 1991).

A common measure of area of occupancy is "site occupancy", the number of occupied units of a grid placed over a map of the region of interest (Gaston, 1991 and references therein). Application of this approach to the estimation of the wild boar's area of occupancy close to its northern distribution line is complicated by two factors. First, labor and material costs of data collection is high (e.g., random sampling of territory is extremely expensive) and standardization of sampling effort over seasons and years is difficult. In our case, occasional records of detection of wild boar were the only sources of information about the species' distribution and abundance within the study area. Second, there was no standardized statistical protocol for data collection. A new and popular method for estimation of site occupancy is based on the assumptions of "closed-population models" (MacKenzie *et al.*, 2002). Unfortunately, these assumptions are not applicable for many large mammal populations, especially in a recently colonized area (Hanski, 1999; Pollock *et al.*, 2002).

Thus the quantitative analysis of wild boar distribution and status in the periphery of its range required an adaptation of existing approaches for estimation of the species' area of occupancy (Gaston, 1991; Hanski, 1999) using data on occasional records of the species close to the range border. The methods we develop and apply

would be fully applicable if the occasional records were the result of sampling sessions with the effort randomized over space and seasons but not over years. Unfortunately, this assumption is never totally satisfied in a retrospective analysis of occasional records, but analysis under this assumption is probably the best that can be done with existing data. The species records obtained during such "sampling sessions" were mapped and we calculated the proportion of territorial units where animals were found. The proportion of "occupied" territorial units could not be used as a measure of species abundance because large mammals are able to move rather big distances and one individual or group could be recorded in several sites. However, the proportion could be treated as the estimated probability of detecting the species during the study period with a random search. Since in some sites animals could be present but not detected it is reasonable to treat this index as *the estimated minimum probability of recording the species during the study period with a random search*.

We examined whether using this index as a measure of species presence could provide reasonable information about the status of wild boar in the study area and distinct parts of it. We addressed the problems of determining an appropriate spatial scale for the analysis of data on occasional records, quantitative estimation of spatial and temporal patterns of species' "extent of presence" and interpretation of the data in terms of qualitative definitions of species status, particularly those, recommended by IUCN (2003).

## Study area

The area of the Troitsko-Pechorskii administrative district is approximately 40700 km<sup>2</sup>, however to obtain a rectangular study area we extracted from the map a little bigger piece of territory which is about 47000 km<sup>2</sup>. It is situated in the south-eastern part of Komi Republic, approximately between 61° and 63° N. (Fig. 1A). Topographically it is naturally divided into plains, foothills and mountainous parts. In the east it is bordered by the Ural Mountains (highest altitude in this area is 1617 m). Typical landscape of the plains and foothills zones of the western part of the region is uplands (*parma*). Annual precipitation in the mountains typically ranges from 700 to 800 mm, while in the foothills, annual precipitation decreases to about 500 mm. The average depth of snow at the end of winter is about 80 cm. Average temperature of January is typically from 18 to 20° C below zero. Lowest temperatures reach 50 to 53° C below zero. The average July temperatures are 14.5 to 17°C (Olenev, 1965). Human population density is estimated at 0.4 persons/km<sup>2</sup> and the estimated density of roads is 6 km/1000 km<sup>2</sup>. About 81% of the district is covered by forests, mainly black spruce (64% of the forested area). The agricultural areas cover only an estimated 9.7 % of the territory and are primarily dedicated to producing forage for livestock.

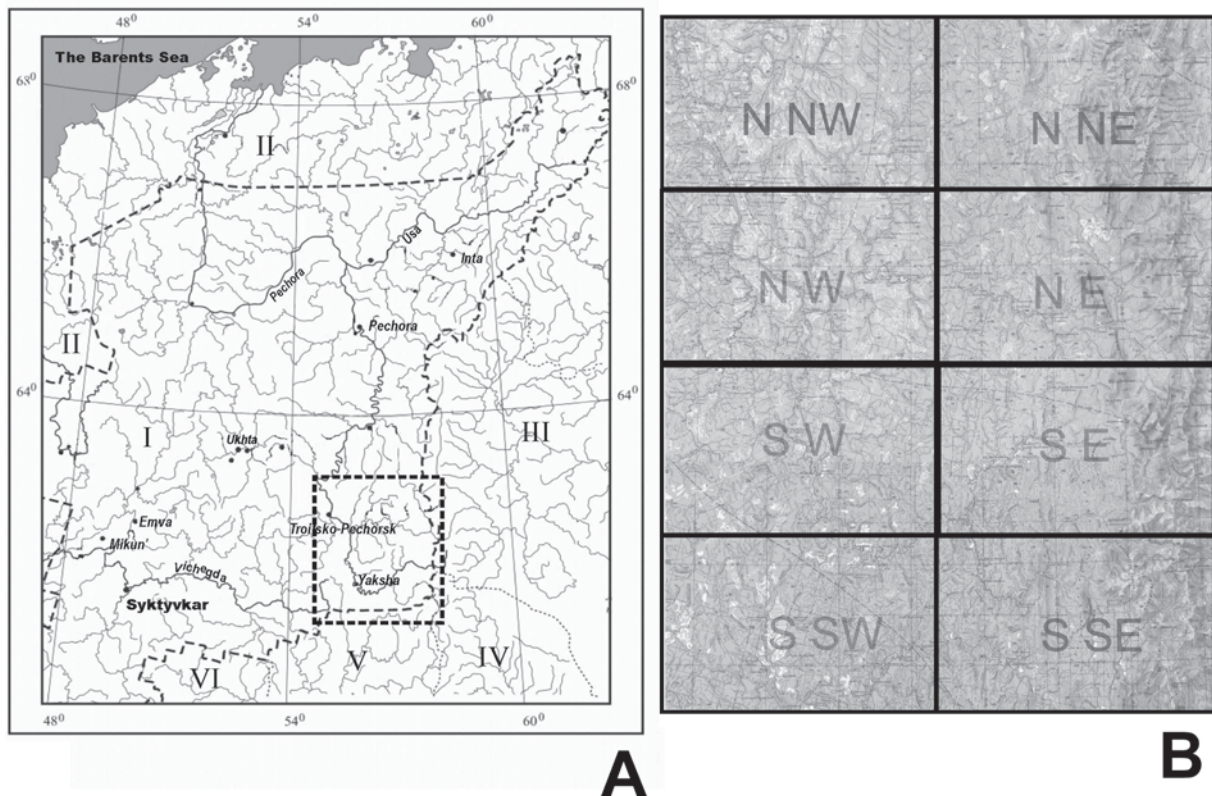


Figure 1. Geographical position and map of the study area. A. European Northeast Russia. Dotted line indicates the position of the study area within the region; I — territory of the Komi Republic, II — Arkhangelsk Province, III — Tyumen Province, IV — Sverdlovsk Province, V — Perm Province, and VI — Kirov Province. B. Topographic map and position of subareas within the study area.

## Methods

The main source of data was information on wild boar records from the database of Pechoro-Ilychsky Nature Reserve for the period from 1984 to 1999. During this time, the Troitsko-Pechorsky administrative district was monitored during all months of the year by research personnel of the nature reserve, forest service, and game management service. They collected data by themselves and used also the information presented by game management law enforcement officers. In addition to these data, an interrogation survey was carried out in 1999–2000. Hunters, foresters, rangers and other people living in rural areas were asked to answer the questionnaire. In total we obtained 78 records of wild boar within the study area for the 15 year period. Each record included the year and location of direct observation of wild boar(s) or signs of their presence (e.g., footprints or rootings). Most included the date (month) of observation and a few included sex and age of the observed individuals and habitat type where animals (or signs of their presence) were found.

The reported locations of wild boar were mapped as points on a topographic map (1:500000) for each year within the period under consideration (1984–1999). Given the map scale, locations less than 0.5 cm from

each other were mapped as one point. These points were then used as indicators of species presence for a given area and time interval (year).

**Selection of spatial scale.** Estimation of spatial scale for description of wild boar distribution in the study area required selection of the size of grid for the analysis of site occupancy. We required two conditions, the grid size has to be fine enough to avoid overestimation of the occupied area and big enough to provide maximum possible information of two different locations. We required the grid size to be the one which provides not more than one location per territorial unit. We searched for this unit size by dividing the map consequently into 2, 4, 8 etc. equal parts and relating the number of records to the number of “occupied” grid cells. We stopped when this value was equal to 1 and calculated the size of grid at this stage by measuring it in centimeters and then recalculating the distance in km using the scale of map (1 cm to 5 km). We also counted the number of units of this size within the study area. This search has been performed for each year in our time series of data.

**Estimation of site occupancy.** We treated a unit as being sampled annually if a human settlement was situated within it or in one of the neighboring units (units were defined to be neighbors if they share a side) or if it



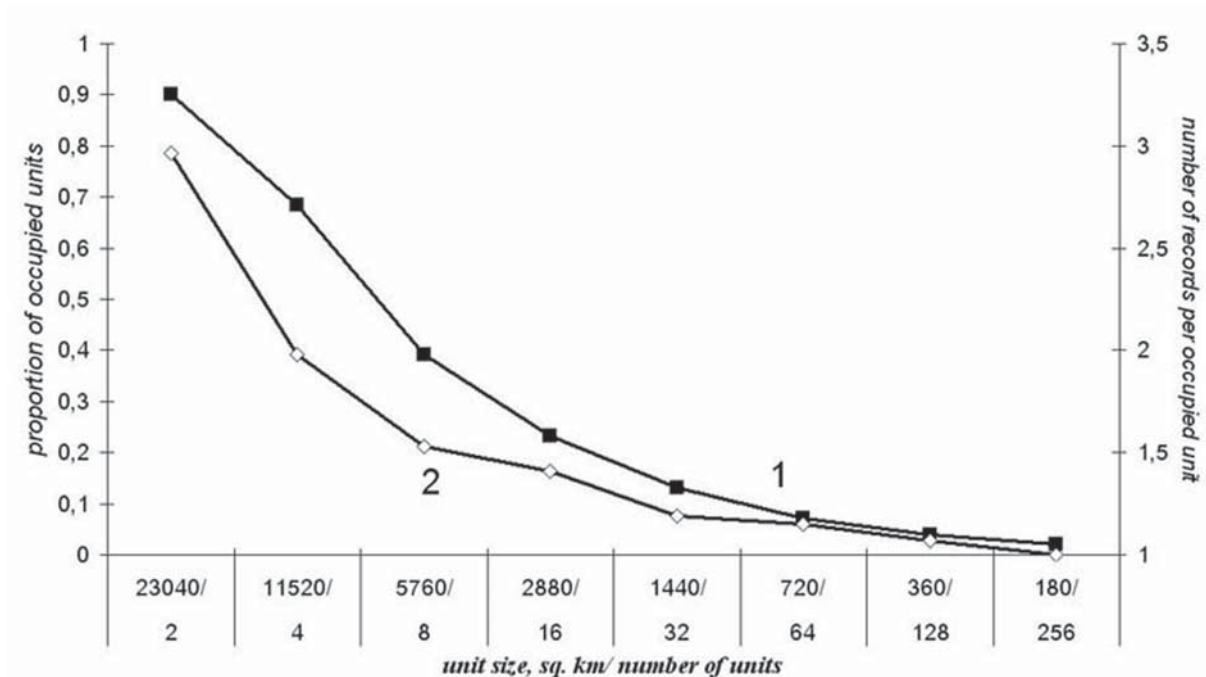


Figure 2. Dynamics of the proportion of average occupied territorial units (1) and average number of records per occupied unit (2) versus the unit size.

was visited regularly by nature reserve personnel. We calculated the index of site occupancy by dividing the number of occupied units by the total number of sampled units (Manly *et al.*, 2002). The number of occupied sites obtained from mapping the locations of the wild boar records was treated as the *minimum number of occupied units* because the animal could have been present but not detected. Thus, we computed the minimum site occupancy (MSO),

$$\text{MSO} = U_+ / U,$$

where  $U_+$  is a number of occupied sites and  $U$  is the number of sampled units within the area under consideration.

**Estimation of temporal patterns of species presence.** The constancy of species presence in the study area or part of it (CP) was estimated by:

$$\text{CP} = Y_+ / Y,$$

where  $Y_+$  is the number of years in which wild boars have been detected within the given sub-area and  $Y$  is the total period under consideration (1984–1999, 15 years).

**Analysis of spatial dynamics of wild boar distribution and status.** In order to analyze the possible spatial dynamics of wild boar distribution and status, the whole study area has been divided into eight equal parts (NNW, NNE, NW, NE, SW, SE, SSW and SSE) (Fig. 1B) to take into account possible latitudinal and altitudinal variation of the species occurrence. The eastern parts included mainly the territory of Pechorolychsky nature reserve, where few human settlements

exist, while the western part of the study area is more or less highly populated. This likely resulted in strong differences in sampling effort, however we found no significant correlation between the number of sampled sites and the number of records in these sub-areas ( $n=8$ ,  $R$  Spearman = 0.44,  $p=0.28$ ). Minimum site occupancy and constancy of species presence were calculated for each part of the study area as described above. We used the Mann-Whitney test (Ledermann, 1989) to examine differences in this parameter between the different parts of the study area, because the distributions departed from normality.

In order to estimate the spatial dynamics of wild boar distribution and status we performed two types of cluster analysis. In cluster analysis 1, the sub-areas were classified basing on the matrix of “raw” values of MSO estimated for each year and each subarea. In cluster analysis 2, CP indexes and the average values of MSO were used as distance measures between subareas. In both cluster analyses, UPGMA was used as an amalgamation (linkage) rule and single Euclidian distances were used as a distance measures (Sneath & Sokal, 1963).

## Results

Fig. 2 shows the change in proportion of occupied units (line 1) and number of records per occupied unit (line 2), averaged for the total period of observations and plotted versus the unit size. The number of records per unit reaches the value of 1 when the unit size is equal to 180 km<sup>2</sup> (12×15 km). Following our method,

Table 1. Minimum site occupancy (MSO, averages and their standard errors, SE) and constancy of presence (CP) of wild boar in different parts of Troitsko-Pechorsky District (Northeast of European Russia).

Subareas	Average MSO	SE	CP
NNE	0.022	0.012	0.133
NW	0.025	0.008	0.533
NE	0.014	0.007	0.267
SW	0.071	0.011	0.933
SE	0.025	0.006	0.533
SSW	0.026	0.008	0.533

we accepted units of size 180 km<sup>2</sup> to be the spatial scale for the analysis. The total area under consideration contained 256 units of this size and each of 8 subareas included 32.

Minimum site occupancy averaged 0.019 for the total study area and the period 1984–1999. Thus, a search of 1000 sites of size 12×15 km, during the study period would be expected to reveal presence of wild boar in only 19 of them. This index reached its maximum (0.057) in 1987 and minimal site occupancy (0.004; only one record) in 1993.

Estimates of average site occupancy and CP indexes for 6 subareas are presented in Table 1 (excluding

NNW and SSE subareas with no records of species). The most reliable estimates of average MSO have been obtained for NW, SW, SE and SSW subareas, where the standard error of means did not exceed 30% of the means. In the northeastern part of the study area, values of MSO varied in time with no records in several years, which determined rather high standard errors of means and less reliable estimations of the average MSO.

Pairwise comparison of MSO values for the 8 the subareas indicated that only the southwestern subarea differed significantly from all the others. The average values of minimum site occupancy in NNE, NE and SE subareas did not differ significantly from those in the NW and SSW subareas where wild boars were regularly reported, nor from NNW and SSE subareas where no wild boar records were made.

Cluster analysis 1 of the annual MSO values for the 8 subareas indicated a gradual change of this index. This result did not contribute much to estimation of species status in the study area. Cluster analysis 2 was made using two measures: the average MSO, which indicates the average minimum probability of finding wild boar during one random search using the retrospective methods, and CP, which indicates the minimum probability of finding evidence of species presence in a time series of observations. The results of this analysis indicated (Fig. 3) that the subareas could be divided into two groups.

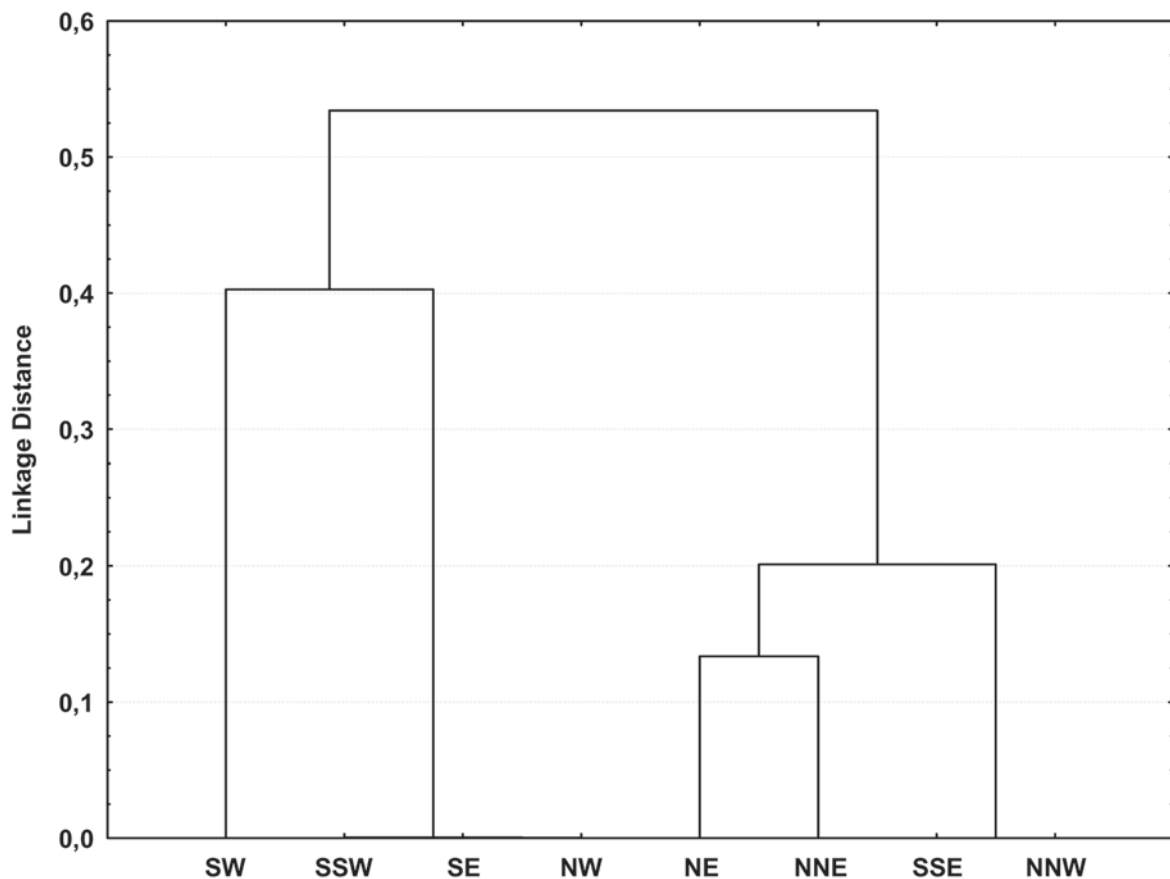


Figure 3. Dendrogram of subareas based on average minimum site occupancy and constancy of presence of the wild boar.



Figure 4. Spatial distribution of average minimum site occupancy and constancy of presence for the wild boar in Troitsko-Pechorsky District. Values are shown for each of the considered subareas.



The first group consisted of the SW subarea (represented as a separate unit) joined with the sub-cluster consisting of the NW, SSW and SE subareas. The sub-cluster is adjacent to the southwestern subarea in space. The second group also included two sub-clusters; one of them linked the NE and NNE subareas together, while the other consisted of the NNW and SSE subareas.

Fig. 4 gives the spatial position of each cluster and respective values of constancy of species presence and average MSO. Generally, both indexes appear to be lower in the eastern part of the study area. A clear east to west gradient exists in MSO, while the constancy of species presence changes gradually from the central part of the area to its periphery.

## Discussion

In our study we used data on occasional records of wild boar to obtain quantitative estimates of its status in the study area. The first important part of our analysis was selection of grid size for estimation of site occupancy as a measure of species presence. We failed to find formal methods in the literature for selection of an appropriate spatial scale for the site occupancy analysis. The importance of this topic has been discussed by Gaston (1991), but in most studies the authors use a variety of grid sizes without adapting those to the exact species under consideration (e. g. Freitag & van Jaarsveld, 1997; Silvo *et al.*, 2002). After data analysis and writing of this paper, we discovered that the steps in our method for selection of unit size, i.e., the spatial scale at which to study site occupancy, follow an “optional” method for efficient kernel estimation of a utilization distribution (UD), the two dimensional probability density function of locations of animals based a random sample of locations (Kern *et al.*, 2003). Estimation of a utilization distribution is an alternative analysis for study of site occupancy. We plan to estimate a utilization distribution based on our retrospective analysis and contrast the results to those in this paper.

Our results indicated marked dependence of site occupancy upon the unit size. Our judgment is that the selected grid size was, in fact, the finest scale on which distribution of wild boar should be analyzed, because it provided not more than 1 record per territorial unit. For the units of smaller size, species’ detection probability could be affected by specific habitat structure, while the territory of 12x15 km was assumed to include at least one habitat type where animals could be detected if they are present. It would also be possible to use units of 30x24 km (360 km<sup>2</sup>) in our analysis, because the average number of locations per occupied unit for this grid size was only slightly bigger than 1. Operating with units of larger size and, thus extrapolation of a single location of species presence to territory bigger than 360 km<sup>2</sup> seemed hardly possible, because such large units could include both areas inhabited by the wild boar and those that never were used by the species.

The estimation of minimum site occupancy for the whole study area indicated that the wild boar is very

rare in this part of the European Northwest. The estimated average probability of recording the animals, or signs of their presence, during the time period under consideration was 0.019. However, the analysis revealed clear differences in minimum site occupancy and constancy of species presence between eight subareas within the study area. In the following discussion, we give possible qualitative interpretations of the quantitative estimates of MSO and CP, and comparison of the subareas in terms of wild boar status.

No records of wild boar were made in two sections of the study area (NNW and SSE) during the 15-year study period, thus we treat these territories as not inhabited by the species. The SW subarea had the highest MSO and CP values. Wild boar has been reported in these territories during the whole period under consideration, except one year. The values of MSO and CP for northwestern, south-southwestern and southeastern subareas are intermediate between the highest in SW and extremely low in NE and NNE. However, cluster analysis clearly linked these areas with SW, but not with the northeastern ones. In fact, this analysis separated the areas where the wild boar was “present at least from time to time” from those where it was “not present or present irregularly in small numbers”. These definitions coincide, respectively, with the terms “visitor” and “vagrant”, used by IUCN (2003) for description of species’ populations on the regional level. Thus, in the northeastern part of the study area, the species could be regarded as “vagrant”, while in the northwestern, southwestern, southeastern and south-southwestern parts of the study area the wild boar had at least “visitor” status. This allows us to treat these territories as a part of the species’ range. Generally, the probability of species presence increased from the east to the west side of the study area, however there was no clear gradient in the north-south direction. The west-east gradient of species status might be explained by strong ecological differences (e.g., depth of snow and habitat structure) between the plains, foothills, and mountainous parts of the study area. On the other hand this gradient could reflect the general pattern of species expansion over the whole Northeast of European Russia (Markov *et al.*, 2004). The differences in species “extent of presence” within the western part of the study area could probably be explained by specific of habitat structure, however, we do not address this problem in this article.

Description of wild boar distribution in the study area by outlining locations of occasional wild boar records puts almost the whole territory into the species’ range. Quantitative estimation of species status based on calculation of the MSO and CP indices demonstrated that not more than 50% of it should be treated as inhabited by the wild boar. Thus, we believe quantitative estimation of species status in the study area to be an important part of the analysis of its area of occupancy. Including areas where the species is vagrant in resource selection analyses could lead to bias in estimation of habitat preferences and errors in patterns of site occupancy. We think that further studies of factors that

affect “presence/absence” patterns and wild boar distribution in areas surrounding Pechoro-Ilychsky Nature Reserve, should include only the areas where presence of species cannot be treated as an artifact.

## References

- Brown J.H. 1984. On the relationship between abundance and distribution of species // *American Naturalist*. Vol.124. P.255–279.
- Fadeev E.V. 1987. [Wild boar] // *Okhota i Okhotnich'e Khozyaistvo*. No.2. P.10–13 [in Russian].
- Freitag S. & van Jaarsveld A.S. 1997. Relative occupancy, endemism, taxonomic distinctiveness and vulnerability: prioritizing regional conservation actions // *Biodiversity and Conservation*. Vol.6. No.2. P.211–232.
- Gaston K.J. 1991. How large is a species' geographic range? // *Oikos*. Vol.61. P.434–438.
- Hanski I. 1999. *Metapopulation Ecology*. Oxford: Oxford University Press. 313 p.
- IUCN. 2003. *Guidelines for Application of IUCN Red List Criteria at Regional Levels: Version 3.0*. Gland, Switzerland and Cambridge, UK: IUCN Species Survival Commission, IUCN. 26 p.
- Kern J.W., McDonald T.L., Amstrup S.C., Durner G.M. & Erickson W.P. 2003. Using the bootstrap and fast Fourier transform to estimate confidence intervals of 2D kernel densities // *Environmental and Ecological Statistics*. Vol.10. P.405–418.
- Ledermann W. (ed.). 1989. [*Handbook of Applicable Mathematics*]. T.1. Moskva: *Finansy i statistika*. 271 p. [in Russian].
- MacKenzie D.I., Nichols J.D., Lachman G.B., Droege S., Royle J.A. & Langtimm C.A. 2002. Estimating site occupancy rates when detection probabilities are less than one // *Ecology*. Vol.83. No.8. P.2248–2255.
- Manly B.F.J., McDonald L.L., Thomas D.L., McDonald T.L. & Erickson W.P. 2002. *Resource Selection by Animals. Statistical Design and Analysis for Field Studies*. Second edition. Dordrecht: Kluwer Academic Publishers. 221 p.
- Markov N.I., Neifeld N.D. & Estaf'ev A.A. 2004. Ecological aspects of dispersal of the wild boar, *Sus scrofa* L., 1758, in the Northeast of European Russia // *Russian Journal of Ecology*. Vol.35. No.2. P.131–135.
- Neifeld N.D. 1998. [*Sus scrofa*, the Wild Boar or Wild Pig] // [Fauna of Northeast of European Russia. Mammals. Tom 2. Part 2]. Sankt Peterburg: Nauka. P.243–246 [in Russian].
- Olenev A.M. 1965. [The Urals and Novaya Zemlya]. Moskva: Mysl'. 215 p. [in Russian].
- Pavlov M.P., Korsakova I.B. & Lavrov N.P. 1974. [The wild boar] // [Acclimation of Game Animals and Birds in the Soviet Union]. Kirov: VNIIOZ. P. 370–436 [in Russian].
- Pollock K.H., Nichols J.D., Simons T.R., Farnsworth G.L., Bailey L.L. & Sauer J.R. 2002. Large scale wildlife monitoring studies: statistical methods for design and analysis // *Environmetrics*. Vol.13. No.2. P.105–119.
- Rusakov O.S. & Timofeeva E.K. 1984. [The Wild Boar: Ecology, Resources, and Economic Significance in the Northwestern Soviet Union]. Leningrad: Izdatel'stvo Leningradskogo Gosudarstvennogo Universiteta. 207 p. [in Russian].
- Silvo T., Reino L. M. & Borralho R. 2002. A model for range expansion of an introduced species: the common waxbill *Estrilda astrild* in Portugal // *Diversity and Distributions*. Vol.8. No.6. P.319–326.
- Sneath P.H.A. & Sokal R.R. 1963. *Numerical Taxonomy*. San Francisco: W.H. Freeman & Co. 369 p.
- Swihart R.K., Gehring T.M., Kolozsvary M.B. & Nupp T.E. 2003. Responses of 'resistant' vertebrates to habitat loss and fragmentation: the importance of niche breadth and range boundaries // *Diversity and Distributions*. Vol.9. No.1. P.1–18.