Artificial neural networks enable discrimination of taxa using morphological characters: an example with a problematic small mammal, the water vole (genus *Arvicola*) (Rodentia, Cricetidae)

Jean-Luc Giraudel¹, Jean-Pierre Quéré²,* and François Spitz³

¹ Université de Bordeaux, UMR 5255, ISM, Rue Doyen Joseph Lajugie, 24019 Périgueux Cedex, France
² INRA, CBGP (Centre Biologie et Gestion des Populations), UMR 1062, Campus International de Baillarguet, CS 30 016, 34988 Montferrier/Lez, France, e-mail: quere@supagro.inra.fr
³ INRA, Centre de Recherches de Toulouse, CEFS, BP 52627, 31326 Castanet-Tolosan Cedex, France

*Corresponding author

Abstract

The aim of this study is to assess the reliability of different artificial neural networks as tools for identifying taxonomic or ecotypic subdivisions in European water voles (genus *Arvicola*). Self-organizing maps show that taxa at subspecies or species level are organized along morphological gradients correlated with ecotypes (from extreme forms of fossorial *Arvicola terrestris* to strictly aquatic *Arvicola sapidus*). An automatic discrimination process based on a multilayer feed-forward network, working with 17 French and one Spanish reference populations, proved able to correctly identify species in test samples from France, but sometimes misidentified specimens from other countries. Three morpho-ecotypes were identified using a self-organizing map with a single-species dataset (*Arvicola terrestris* only). However, there were no clues as to how these morpho-ecotypes might correlate with the known phylogenetics of the species. Our model can identify species on the basis of a limited range of cranial measurements which are available from vole skull fragments commonly found in owl pellets. This can have very practical risk-assessment applications as it allows easier identification of the more fossorial *Arvicola terrestris* populations which are a threat to crops.

Keywords: *Arvicola sapidus*; *Arvicola terrestris*; discrimination; neural networks; skull measurements.

Introduction

Artificial neural networks (ANNs) have seldom been used for discriminating among animal taxa (but see an example in Giraudel et al. 2000; see a review in Giraudel and Lek 2002; also an attempt for *Taterillus* in Dobigny et al. 2002).

It is generally agreed that the genus *Arvicola* Lacépède, 1799 is represented in Europe by at least two species: *Arvicola sapidus*, Miller, 1908, and *Arvicola terrestris* (Linnaeus, 1758) (Niethammer and Krapp 1982, Gromov and Polyakov 1992). However, morphologists disagree about the subdivision of *A. terrestris* into ecotypes (fossorial and semi-aquatic: Saucy 1999), subspecies (35 subspecies have been described to date: Wust-Saucy 1998), or even into species (Miller 1912 described seven species for Western Europe: amphibius, illyricus, italicus, musignani, sapidus, scherman, terrestris).

The genus has been the subject of various taxonomical and biogeographical studies in Western Europe, from Heim de Balsac and Guislain (1955), through Spitz and Morel (1972), to Ventura and Sans-Fuentes (1997). Species and subspecies have been differentiated by morphometric measurements (Engels 1975, Warmerdam 1982, Ventura and Gosálbez 1989, Escudé et al. 2008), non-metric characters (Ventura and Sans-Fuentes 1997), cytogenetics (Peshev and Belcheva 1978, Díaz de la Guardia and Pretel 1979), allozymes (Graf 1982, Saucy et al. 1994), and mitochondrial DNA polymorphism (Wust-Saucy 1998). These studies all concur on the existence of two species (*Arvicola terrestris* and *Arvicola sapidus*).

Wust-Saucy (1998) subdivided *Arvicola terrestris* into three genetic lineages (Eastern, Italian, and Iberian) and three ecotypes (aquatic, fossorial, and intermediate). Wilson and Reeder (2005) classified semi-aquatic populations as *Arvicola amphibius* (Linnaeus, 1758) and fossorial populations as *Arvicola scherman* (Shaw, 1801). Their classification was based on body mass, pelage coloration, social behavior, mating system, use of space, skull size and shape, and incisor protrusion. They did not use intermediate ecotypes.

This paper is somewhat conservative in outlook and considers that there are just two *Arvicola* species – *A. sapidus* and *A. terrestris* – and two ecotypes (aquatic and fossorial) (Figure 1).

Our aim is two-fold: 1) to evaluate two ANN models as tools (i) for detecting relevant taxonomic subdivisions within a data set from a collection of known populations, and (ii) for classifying unknown specimens. 2) To explore possible correlations between the ANN subdivisions (3 morphotypes) and the phylogenetic tree of the genus.

Material and methods

Material

Specimens were provided by museums, research laboratories, and private collections. They came from 138
localities (117 in France, 6 in Germany, 7 in Spain, 3 in Switzerland, 2 in the United Kingdom, 2 in Poland, and 1 in Italy). A few specimens were excluded because of doubts about their precise locality of origin, or because one or more measurements could not be made. The resulting dataset totaled 563 specimens.

Species were identified initially using the descriptive characters proposed by Chaline et al. (1974), Blanco (1998), and Le Louarn and Quéré (2003). A total of 33 Arvicola sapidus specimens were identified. The fact that they were collected from regions (in western France and southern Spain) well outside the known range of A. terrestris was supporting evidence for their identification. These 33 specimens were subdivided into Mediterranean (n=11) and non-Mediterranean (n=22) “reference populations”.

For reference populations of Arvicola terrestris, a new selection criterion was added. Most were collated from a homogeneous series of individuals captured in the field (not found in owl pellets), each population at a specific locality, in conditions that excluded the possibility of them being A. sapidus. Most individuals selected in this way belonged to fossorial populations. A reference population for the semi-aquatic type was made up of specimens from the flood plain of northernmost France (Nord-Pas de Calais=NPC). Three populations [Ouroux and Prisy (ORP), Nord-Pas de Calais (NPC), and Saint Alban de Montbel (SAL)] were collected from owl pellets. In all, 17 reference populations of A. terrestris (n=304) were included in the study (Figure 2).

**Skull measurements**

Eight skull measurements were made, as described elsewhere for Micrurus (Terricola) (Spitz 1978), meaning that even partly fragmented specimens (e.g., those found in owl pellets) could be included in the study. These measurements were:

- inter-orbital distance,
- nasal bone length (measured between the narrowest posterior part and widest anterior part),
- maximum width of the anterior part of the nasal bone,
- minimum width of the posterior part of the nasal bone,
- length of the anterior palatal foramina,
- height of cranium measured between the first and second upper molar and the top of cranium,
- length of the upper molar tooth row,
- length of the maxillary diastema.

**Modeling techniques**

Skull measurements were analyzed using two categories of ANNs (Giraudel et al. 2000, Giraudel and Lek 2001, 2002). ANNs are simply computer programs organized into successive layers. The basic rule of a neural network is that every neuron in a given layer is connected with every neuron in the preceding and the following layers. Two types of ANNs were used.

The first type of network was a topological self-organizing map (Kohonen 1995), hereafter named SOM. Such
networks are made up of two layers. Neurons in the input layer are connected to input data (in the present study, eight neurons for eight measurements). The output layer consists of a rectangular grid of hexagonal cells (we used a SOM with eight columns and five lines, i.e., 40 cells). The network is designed to organize the structure of the map by itself, placing the most similar observations (an observation is a specimen characterized by its eight measurements) in the same cell or in contiguous cells. During map-building, the network is given no other information than the eight measurements, in an ‘unsupervised learning’ process. At the end of this learning phase, dissimilarity between contiguous cells can be measured. For practical purposes, dissimilarity is generally represented as a difference of elevation. Consequently, SOMs may look like topographical maps with flat regions (interpreted as homogeneous sets of observations) separated by slopes of varying steepness (generally interpreted as heterogeneous sets of observations of intermediate status). Input data values can be displayed on the SOM. This means that homogeneous regions (i.e., homogeneous groups of specimens) can be linked with particular associations of measurements (only when the range of values for a measurement in a group is clearly separate from another group’s range).

The second type of neural network is a multilayer network working with the back-propagation of error algorithm (Rumelhart et al. 1986), abbreviated hereafter as MNBPE. This type of neural network is devised to segregate pre-established categories, such as species or subspecies. The network comprises three layers. The input layer is made up of eight neurons (eight measurements). The output layer has a neuron for each expected category. A hidden intermediate layer contains a variable number of neurons. The hidden layer is connected to both the input and output neurons and it plays a key role in the network results. During the learning phase, the group to which each observation belongs is identified in the network (one of the 19 reference populations). The network is designed to modify the synaptic weights (i.e., the role played by each connection between two neurons of two successive layers) until the correct answer is obtained. When all (or a fixed proportion of) the observations of the learning set (i.e., reference population specimens) are correctly classified (i.e., placed in the correct population), the network is ready to operate on unclassified specimens. As each reference is represented by an output neuron, the expected answer is 1 for one neuron and 0 for all the other neurons. The network delivers a value between 0 and 1 for each output neuron, meaning that each specimen belongs “more or less” to the different populations. These values are not probabilities, but must be understood, as in fuzzy logic, as a score characterizing a degree of membership. We considered that a specimen should be included in the population for which its output value was the highest, provided that this value exceeded 0.05. In fact, a large majority of the higher values exceeded 0.90, and most specimens had only one very high output value, the other values being very low or 0. Two different discriminant-neural-network procedures were used. The usual procedure involves a learning set of all confirmed reference specimens, and a test set of all unclassified specimens. The “leave-one-out” procedure (Kohavi 1995) was used to test the reliability of identification of the reference specimens, at the species or population level. The learning set included all individuals but one, the omitted individual making up the test set. Each individual was tested in turn.

Reference specimens and newly classified specimens were then studied from geographical and ecological viewpoints.

Data processing, neural networks, and distributional operations were carried out using Dbase® (Dbase IV® Borland Software Corporation, Austin, Texas, USA), Access® (Access 97® Microsoft Corporation, Redmond, Washington, USA), Matlab® (Matlab® v.6 The Mathworks Inc., Natick, Massachusetts, USA) and MapInfo® (MapInfo® v.6 Mapinfo Corporation, New York, USA). The Kohonen map models were programmed by one of the authors (Giraudel 2001).

Our classification and discrimination process was as follows:

1. apply SOM to the entire collection (species-level discrimination),
2. validate species identification using MNBPE,
3. apply SOM to confirmed reference populations of Arvicoola terrestris: subspecies level discrimination (ecotype or geographical variant),
4. identify, using MNBPE, to which subspecies reference specimens (validation) or previously unclassified specimens belong to.

Results

Reference population distribution

We focused on the cells containing the majority of observations, or modal areas, for each reference population. Specimens of Arvicoola sapidus occupied five cells in the...
top right corner exclusively. Figure 3 shows that the modal areas of *Arvicola terrestris* populations arranged themselves depending on their dissimilarity from *A. sapidus* and were distributed along a vertical gradient. Populations from northern, eastern and central France appeared closest to *A. sapidus* and certain populations from the Pyrenees, the Alps, and the Jura were the most remote.

Our first conclusion is that the skull characteristics of *Arvicola sapidus* are much more homogeneous than those of *Arvicola terrestris*.

**Neural network method discrimination power**

Testing the species status of the 337 reference individuals ("leave-one-out" procedure) yielded the following results: all species identifications were confirmed, except for an isolated Spanish specimen of *Arvicola terrestris*, which was erroneously assigned to the Mediterranean set of *Arvicola sapidus*. The 227 unclassified individuals (most of them had a suspected, if not proven, species status) were then tested with the usual procedure. For most of them, there was one output value only between 0.7 and 1, confirming the robustness of the process. Nevertheless, an appreciable proportion of pre-identifications were not confirmed, showing how difficult it is to identify species from general external or skeletal characters. Species identification of French specimens of *A. sapidus* did not contradict the previously known geographic range. Of three specimens from a locality at the Somme estuary, considered as the northern limit of *A. sapidus*, one was assigned to *A. sapidus* and the other two to *A. terrestris*. In France, *A. terrestris* specimens were identified in a few additional localities of the Landes lowlands, in a few new isolated extreme points in the Seine-et-Marne and the Vienne departments, and on the Mediterranean slopes of the Cévennes in the Gard department. Isolated specimens from elsewhere in Europe yielded some unexpected results.

**Table 1** Populations belonging to each reference group of the SOM of *Arvicola terrestris*.

<table>
<thead>
<tr>
<th>Population</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>MLS</td>
</tr>
<tr>
<td>Group 2</td>
<td>LMA and PRA</td>
</tr>
<tr>
<td>Group 3</td>
<td>LFA</td>
</tr>
<tr>
<td>Group 4</td>
<td>NPC</td>
</tr>
<tr>
<td>Group 5</td>
<td>STO</td>
</tr>
<tr>
<td>Group 6</td>
<td>BY and CDH</td>
</tr>
<tr>
<td>Group 7</td>
<td>ARA, LSA and SAL</td>
</tr>
</tbody>
</table>

Region names as in Figure 2.

**Figure 4** Subdivision of SOM of *Arvicola terrestris*. Seven clusters are numbered.

**Defining seven morphometric/geographical groups of *Arvicola terrestris* and biometrical relationships in these groups**

We used the same SOM-based model to differentiate between populations of *Arvicola terrestris* only. We found that 11 of the 17 reference populations were each concentrated on a small cluster of cells, while the remaining six populations were scattered across many different regions of the SOM. Because of overlap among clusters, we considered the SOM to be subdivided into seven groups (Figure 4) constituting new reference sets for further examination. The remaining 90 specimens were put back into the set of individuals to be classified. Table 1 shows the reference population(s) belonging to each group, and the geographical regions concerned. Mapping the values of cranial measurements in the SOMs revealed associations of values for each of the seven groups. This is summarized in Table 2. For each measurement, a few groups display characteristic associations. The maximum is five groups for "minimum width

**Table 2** Cranial measurements showing definite trends in association with the seven SOM groups of *Arvicola terrestris*.

<table>
<thead>
<tr>
<th>Group</th>
<th>Observed trends</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>Posterior nasal: wide; Palatal foramen: short; Upper molar row: long</td>
</tr>
<tr>
<td>Group 2</td>
<td>Cranium height: great; Upper molar row: long</td>
</tr>
<tr>
<td>Group 3</td>
<td>Inter-orbital: narrow; Nasal: long; Cranium height: great</td>
</tr>
<tr>
<td>Group 4</td>
<td>Inter-orbital: narrow; Nasal: long; Posterior nasal: narrow; Cranium height: great</td>
</tr>
<tr>
<td>Group 5</td>
<td>Anterior nasal: wide or medium; Posterior nasal: medium; Palatal foramen: medium or long; Diastema: medium or long</td>
</tr>
<tr>
<td>Group 6</td>
<td>Anterior nasal: narrow; Posterior nasal: narrow; Upper molar row: short</td>
</tr>
<tr>
<td>Group 7</td>
<td>Anterior nasal: narrow; Palatal foramen: short; Upper molar row: long</td>
</tr>
</tbody>
</table>

Region names as in Figure 2.
of posterior part of nasal bone”, with other measurements occurring in two or three characteristic associations only. This means that no characteristics are common to groups belonging to a higher order assemblage (see below and discussion).

**Discrimination among groups within *Arvicola terrestris***

Again using MNBPE, we first tested the stability of the reference specimens of *Arvicola terrestris* in relation to the group they belong to. The “leave-one-out” procedure yielded 96.7% of correct results: five individuals were unclassified (all output values lower than 0.05) and one individual from group 7 was classified in group 2. Next, by the usual procedure, we assigned each of the unclassified *A. terrestris* specimens (including the six specimens excluded by the leave-one-out test) to one of the seven groups. Of the 257 specimens, 34 (13.2%) were not classified (all output values lower than 0.05). The remaining 223 were unevenly distributed among the groups. Table 3 shows that groups 6 and 7 were represented far more often than groups 3 or 4, but this raw result is meaningless in itself. A first interpretative approach was to examine the concordance between the geographical locality of the reference sets and the representation of groups in the regions, using all reference and non-reference specimens. Figure 5 shows that all

**Table 3**  Average proportions of the groups in the seven regions.

<table>
<thead>
<tr>
<th>Region</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
<th>Group 6</th>
<th>Group 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>0.10</td>
<td>0.00</td>
<td>0.05</td>
<td>0.65</td>
<td>0.10</td>
<td>0.03</td>
<td>0.08</td>
</tr>
<tr>
<td>North-East</td>
<td>0.13</td>
<td>0.13</td>
<td>0.00</td>
<td>0.01</td>
<td>0.28</td>
<td>0.23</td>
<td>0.05</td>
</tr>
<tr>
<td>Alsace-Jura</td>
<td>0.07</td>
<td>0.09</td>
<td>0.00</td>
<td>0.01</td>
<td>0.20</td>
<td>0.40</td>
<td>0.41</td>
</tr>
<tr>
<td>Massif Central</td>
<td>0.16</td>
<td>0.08</td>
<td>0.01</td>
<td>0.02</td>
<td>0.01</td>
<td>0.21</td>
<td>0.42</td>
</tr>
<tr>
<td>Landes</td>
<td><strong>0.50</strong></td>
<td>0.00</td>
<td>0.10</td>
<td>0.00</td>
<td>0.00</td>
<td>0.40</td>
<td>0.35</td>
</tr>
<tr>
<td>Pyrenees</td>
<td>0.05</td>
<td>0.00</td>
<td><strong>0.27</strong></td>
<td>0.00</td>
<td><strong>0.23</strong></td>
<td>0.07</td>
<td>0.35</td>
</tr>
<tr>
<td>Alps</td>
<td>0.08</td>
<td><strong>0.32</strong></td>
<td>0.02</td>
<td>0.00</td>
<td>0.00</td>
<td>0.08</td>
<td><strong>0.40</strong></td>
</tr>
</tbody>
</table>

Data in bold underlined font: group which is dominant in the region. Data in bold font: subdominant groups, provided that their proportion exceeds 0.2.

**Figure 5**  Distribution map of reference places of the seven groups (bold numbers) and boundaries of the geographical subdivisions of the range of *Arvicola terrestris* in France. The frequency of each group is indicated as a percentage, all identified specimens included.
regions except the Landes contain one or two reference localities, but the dominant group in a region is not always that of the nearest reference locality. Group 7 is dominant in all mountain regions: Alsace-Jura, Massif Central, Alps, and Pyrenees. Groups 1, 4, and 5 are dominant in one region each. Group 6 is subdominant in four regions outside the Alps. The Pyrenees include two subdominant groups: group 5 (dominant in the North-East) and group 3 (rare everywhere else).

Gradients on the SOM are influenced by morphological changes with age. This influence is limited to the vertical axis in the SOM of *Arvicola terrestris*. Taking this into account, we obtained a relatively simple classification of the *A. terrestris* specimens into three morphotypes only: the semi-aquatic type, restricted to group 4 (reference population NPC), a first fossorial type, or “fossorial 1”, with groups 1 and 5 (referenced by populations MLS and STO), and a second fossorial type, or “fossorial 2”, including groups 2, 3, 6, and 7 (referenced by populations LMA, PRA, LFA, BY, CDH, APA, LSA, and SAL), Figure 6 provides a synthetic view of the distribution of these morphotypes. Table 4 presents the relative occurrence of the three morphotypes by region. The map in Figure 6 shows that the semi-aquatic morphotype is present in several regions outside the range previously admitted, but it is extremely rare outside the two northern lowland regions. Nevertheless, its range noticeably overlaps the range of *Arvicola sapidus*. “Fossorial 1” and “fossorial 2” are sympatric everywhere, including the Landes and the Pyrenees. The balance between these two morphotypes definitely seems to be associated with the prevalence of mountain habitats. Nevertheless, we should also consider proximity to the area of dispersion of the subspecies *A. terrestris scherman*, which could correspond to this morphotype (90% of specimens in both easternmost regions are “fossorial 2”).

**Table 4** Distributional aspects of the classification of *Arvicola terrestris* into three morphotypes.

<table>
<thead>
<tr>
<th>Region</th>
<th>Semi-aquatic</th>
<th>Fossorial 1</th>
<th>Fossorial 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>North</td>
<td>0.65</td>
<td>0.20</td>
<td>0.16</td>
</tr>
<tr>
<td>North-East</td>
<td>0.10</td>
<td>0.41</td>
<td>0.41</td>
</tr>
<tr>
<td>Alsace-Jura</td>
<td>0.01</td>
<td>0.09</td>
<td><strong>0.90</strong></td>
</tr>
<tr>
<td>Massif Central</td>
<td>0.02</td>
<td>0.17</td>
<td>0.72</td>
</tr>
<tr>
<td>Landes</td>
<td>0.00</td>
<td><strong>0.50</strong></td>
<td>0.50</td>
</tr>
<tr>
<td>Pyrenees</td>
<td>0.00</td>
<td>0.28</td>
<td><strong>0.72</strong></td>
</tr>
<tr>
<td>Rhône-Alps</td>
<td>0.00</td>
<td>0.08</td>
<td><strong>0.90</strong></td>
</tr>
</tbody>
</table>

Data in bold underlined font: highest frequency of each morphotype in the study area as a whole. Data in underlined font: highest frequency of morphotype(s) by region.

**Discussion**

In the introduction, we asked: (i) can ANN ordination models (SOM) detect relevant subdivisions within a collection of unclassified individuals? (ii) Can ANN discriminant models (MNBPE) correctly recognize pre-established species or subspecies subdivisions within a collection of reference individuals? (iii) Once validated,
can these models determine to which species or sub-

species unidentified specimens belong? Additionally, we set out to determine whether the ANN subdivisions identified in our collection of \textit{Arvicola terrestris} specimens were correlated with the known phylogenetics and taxonomy of the species.

With regard to the first question, SOM produced satisfactory orderings which clearly isolate \textit{Arvicola sapidus} as a homogeneous subset close to the heterogeneous subset of \textit{Arvicola terrestris}.

With regard to the second question, we obtained 100\% of correct results at species level. This tool can therefore accurately identify \textit{Arvicola} skulls found in owl pellets, i.e., skulls damaged in varying degrees. For our reference populations, the SOM subdivisions within \textit{Arvicola terrestris} were validated by the discriminant model with more than 95\% confidence level.

Concerning the third question, most unidentified individuals could be assigned to one species with a score exceeding 0.7, sometimes in contradiction with their provisional pre-identification. At the subspecies level, within \textit{Arvicola terrestris}, more than 85\% of unclassified specimens were assigned to one of the SOM groups.

Subspecies subdivisions in \textit{Arvicola terrestris} (sensu latu) have generally been considered to follow an aquatic-fossorial gradient. More recently, geneticists have produced evidence that only two or three phyla exist (Saucy et al. 1994, Taberlet et al. 1998, Wust-Saucy 1998), with some local anatomical adaptations. The tail/neck-body ratio is smaller in fossorial populations (lineage F of Wust-Saucy) than in populations living in aquatic environments (lineage A2 of Wust-Saucy) (Wust-Saucy 1998). The same type of macroscopic differentiation was expected in skull morphology which was confirmed through our multivariate unsupervised approach, but no univariate relation was found between any of the measurements and any of our within-species subdivisions. In this context, it is not surprising that our somewhat simplistic approach led to rather simple results:

1. The skull morphology of \textit{Arvicola sapidus} is much less variable than that of \textit{Arvicola terrestris}. This result is in agreement with the literature (Cubo et al. 2006). But morphological homogeneity needs to be confirmed for \textit{A. sapidus} due to the small sample size (n=33). Nevertheless, the results from these specimens, which cover a broad geographical range, make a case for this hypothesis.

2. \textit{Arvicola sapidus} morphology is closer than expected to that of some mature specimens from certain \textit{Arvicola terrestris} populations.

3. Good species discrimination using cranial measurements requires a suitable set of reference individuals for each geographical region. Our ANN models should yield information about other local populations, in particular from Eastern and Southern Europe.

4. The apparent high variability of \textit{Arvicola terrestris} could be age-biased. This leads us to propose a parsimonious subdivision into three morphotypes only: one semi-aquatic and the other two fossorial.

5. At this point of the study, we cannot identify any single morphological trait (a particular range of values of one of our measurements) as characterizing any of the three morphotypes. Larger reference samples are required.

6. The distribution area of the semi-aquatic morphotype in France seems wider than previously supposed.

7. Both fossorial morphotypes seem to be sympatric everywhere, with a gradient of their respective frequencies ranging from mountain areas (\textit{loci typici} of \textit{Arvicola scherman}, mostly “fossorial 2”) towards eastern and southern lowlands (“fossorial 1” and “fossorial 2” at similar frequencies) and northern lowlands (mostly “fossorial 1”).

These results suggest a relatively straightforward correlation between morphology (our three morphotypes) and environment (altitude and climate). This is a qualitative association, and membership of a given group may (or may not) be associated with a majority of greater or smaller values of a particular measurement (Table 2). These results also suggest that fossorial adaptations might be acquired at varying rates depending on local conditions (Cubo et al. 2006).

Our results are in agreement with a number of studies on geographical correlates of morphotypes of \textit{Arvicola}. Engels (1975) segregates Alpine and Pyrenean fossorial populations of \textit{Arvicola terrestris}, Ventura and Gosámez (1989) differentiate Pyrenean and Asturian populations as well as populations from the Central Massif and from the Alps. Warmerdam (1982), in the Netherlands, uses principal component analysis and discriminant analysis to separate fossorial and semi-aquatic forms. However, no studies since Spitz and Morel (1972) have demonstrated the presence of a morphological gradient between \textit{Arvicola sapidus} and aquatic than fossorial forms of \textit{A. terrestris}.

Our results are also consistent with the phylogeny proposed by Wust-Saucy (1998), in which all fossorial \textit{Arvico}ola in France belong to a single lineage from a single refuge in the Iberian Peninsula. The phylogenetic position of our semi-aquatic morphotype is still unresolved, as our reference sample was small, and because no specimen from northern France was included in the molecular genetics studies: this morphotype could represent either a local adaptation of the “Iberian lineage” (lineage F of Wust-Saucy) or the westernmost extension of the northern aquatic lineage (lineage A2 of Wust-Saucy) from the Carpathian refugium. The existence of these refuges is well supported by the fossil record (Kowalski 2001, Nadachowski 2003). But these scenarios must be confirmed. Refugia are currently the subject of much discussion and it is suggested that many mammal species found permanent refuges in northern Europe during the Quaternary (Bhagwat and Willis 2008).

Conclusion

\textit{Arvicola sapidus} is monotypic, while \textit{Arvicola terrestris} is polytypic and thus displays great morphological variability. We confirmed that a convergence could exist between the skull morphology of extreme individuals (or populations) of the aquatic morphotype of \textit{A. terrestris} and that of \textit{A. sapidus}. We also demonstrated that a morphological analysis based on a limited number of meas-
urements (using fragmentary specimens from owl pellets) could produce results that are consistent with more complete studies. In order to cover at least the three main lineages identified by Wust-Saucy (1998), a further test of our approach should include skull samples from a much larger area of Europe, covering Italy, Germany, Scandinavia, and the Balkan Peninsula at least. ANNs proved to be efficacious at segregating populations by morpho-ecotypes and at classifying unknown specimens. The Kohonen maps are a particularly effective tool for classifying any collection of objects without prior assumptions.

Acknowledgements

We are indebted to A. Arrizabalaga (Museum of Granollers – SP), J. Cuisin (National History Museum of Paris – F), F. Poitevin (EPHE, Montpellier – F), D. Siruge (Regional Parc of Morvan – F) and L. Viroit (University of Poitiers – F) for loans of material. Thanks also to P. Brunet-Lecomte and L. Granjon who kindly reviewed an early draft of the manuscript. We are grateful to Mark Hewison and Christopher Sutcliffe for improving the English text.

References


Giraudel, J.L., D. Aurelle and S. Lek. 2000. Application of the

Giraudel, J.L., D. Aurelle and S. Lek. 2000. Application of the

self-organizing mapping and fuzzy clustering microsatellite
data: how to detect genetic structure in brown trout (Salmo

data: how to detect genetic structure in brown trout (Salmo

tutta) populations. In: (S. Lek and J.F. Guégan, eds.) Artificial


Graf, J.D. 1982. Génétique biochimique, zoogéographie et taxono-

Graf, J.D. 1982. Génétique biochimique, zoogéographie et taxono-
mie des Arvicolidae (Mammalia, Rodentia). Rev. Suisse

mie des Arvicolidae (Mammalia, Rodentia). Rev. Suisse

Zool. 89: 749–787.


Heim de Balsac, H. and R. Guislain. 1953. Evolution et spécia-

Heim de Balsac, H. and R. Guislain. 1953. Evolution et spécia-
tion des campagnols du genre Arvicola en territoire français. 


Kohavi, R. 1995. A study of cross-validation and bootstrap for 
estimation and model selection. In: Proceedings of the 14th

International Joint Conference on Artificial Intelligence. Mor-


1143.


in Information Sciences. Springer Verlag, Berlin/Heidelberg.

362 pp.


72: 1–389.

Le Louarn, H. and J.P. Quéré. 2003. Les rongeurs de France:

faunistique et biologie. 2ème édition revue et augmentée.


Miller, G.S. 1912. Catalogue of the mammals of Western Europe

(Europe exclusive of Russia) in the collection of the British


Nadachowski, A. 2003. The Carpathians as a probable late

Pleistocene refugium for small mammals: a fossil record. In:

4th European Congress of Mammalogy, Institute of Verte-

brate Biology, Academy of Sciences of the Czech Republic, 

Brno. 175 pp.


Europas. Vol. 2/1 (Rodentia). Akademische Verlagsgesell-


vole (Arvicola terrestris L., Rodentia, Mammalia) chromo-


representations by back-propagating errors. Nature 323:

533–536.

Saucy, F. 1999. Arvicola terrestris (Linnaeus, 1758). In: (A. Mitch-

ell-Jones, G. Amori, W. Bogdanowicz, B. Krystufek, P. Reijn-

ders, F. Siptzenberger, M. Stubbe, J. Thissen, V. Vohralik and 

J. Zima, eds.) The atlas of European mammals. Academic


polymorphism and genetic variability in aquatic and fossorial

populations of the water vole, Arvicola terrestris, Rodentia,


227–242.

Spitz, J. 1978. Etude craniométrique du genre Phymys. Mammal-

ia 42: 268–304.


des grands campagnols (Arvicola Lacépède) par l’analyse


D 275: 771–774.


Comparative phylogeography and postglacial colonization 


terrestris (Linnaeus, 1758) (Rodentia, Arvicolidae) in the 


