Ecological Informatics 7 (2012) 35–40

Contents lists available at ScienceDirect

Ecological Informatics

journal homepage: www.elsevier.com/locate/ecolinf

Open source evaluation of kilometric indexes of abundance

Damiano G. Preatoni [b], Clara Tattoni [a,*], Francesco Bisi [b], Elisa Masseroni [b], Davide D'Acunto [b], Stefano Lunardi [c], Ivana Grimod [c], Adriano Martinoli [b], Guido Tosi [b]

[a] *Museo delle Scienze, Sezione Zoologia dei Vertebrati, Via Calepina, 14, 38122 Trento - Italy*
[b] *Unità di Analisi e Gestione delle Risorse Ambientali, Dipartimento Ambiente-Salute-Sicurezza Università degli Studi dell'Insubria, Via J.H. Dunant 3, 21100 Varese (Italy)*
[c] *Studio associato Progetto Fauna, Fraz. Gimillan n. 180 Cogne (AO), Italy*

ARTICLE INFO

Article history:
Received 21 September 2010
Received in revised form 8 July 2011
Accepted 9 July 2011
Available online 20 July 2011

Keywords:
KAI
GIS
GRASS
Free and Open Source
Wildlife monitoring
Mountain hare

ABSTRACT

Kilometric Abundance Index (KAI) is a common measure used in wildlife studies because it allows a straightforward comparison of species abundance in different sites or at different times. KAI expresses the ratio of the total number of individuals (or of signs of presence) observed along a transect by the total transect length covered at each site. v.transect.kia is a new tool for GRASS GIS, developed for automating the evaluation of KAI, reducing the risk of manual errors especially when handling large datasets. It can also split the transects according to one environmental variable (typically habitat type) and evaluate true 3D transect length. It calculates KAI using a point map of sightings and saves the results in the attribute table, the output can be displayed in any GIS or used for further statistical analysis. The tool has been tested on field data from Northern Italy for mountain hare (*Lepus timidus*), allowing a first wide-area estimate.

1. Introduction

The Kilometric Index of Abundance (KAI or KI, Eq. (1)) is a common measure used in wildlife studies because it allows a straightforward comparison of species abundance in different sites or at different times (Buckland et al., 1993; Vincent et al., 1991). The presence of one species in a site can be confirmed by direct sightings of the animals or by the signs of their passage that can be reliability detected by a trained field biologist: footprints, tracks, droppings, feeding marks, dens, nests, hairs or feathers as well as carcasses or body parts. These indirect signs of presence are very useful when dealing with elusive species that are difficult to spot or capture. KAI expresses the ratio of the total number of individuals (or of signs of presence) observed along a transect by the total transect length covered at each site.

$$\mathrm{KAI} = \frac{\text{Number of presence sightings or signs}}{\text{Transect length covered (km)}} \quad (1)$$

Line transects surveys have the advantage of being a simple method to implement, they require limited equipment (basically, a GPS receiver) and a limited number of operators compared to other census techniques, thus resulting less expensive and offering the practicality, sensitivity and robustness, "desirable qualities" discussed in Engeman(2005).

The method has been designed and is mainly used for vertebrate species (de Thoisy et al., 2008; Engeman, 2005; Maillard et al., 2001), but it can be adapted to any species depending on its behaviour, spatial distribution and access to the study area. KAI can provide qualitative and quantitative information about the presence of a species and its population trends over time (see (Marchandeau et al., 2006)), but has obvious limits. Since it is a relative measurement, it cannot be used to infer population density, if not coupled with other field methods (Burnham et al., 1980; Seber, 1973) that can provide a density estimate.

KAI is often used in preliminary studies to identify sites where it is worth applying other more demanding techniques such as live trapping. KAI field protocols can also be used with slight modifications to apply Distance Sampling techniques (Buckland et al., 1993), thus measuring population density with slight or no change in field protocols.

This index can give more insights about a species compared to the traditional formula of $\frac{\text{sights}}{\text{km}}$ (Buckland et al., 1993; Vincent et al., 1991), if calculated considering the length covered per each habitat type and the actual length covered, taking into account topography that in mountain areas can exert strong influence in abundance index calculations. When dealing with presence signs rather than sightings, it can also be of interest to weight each presence sign type according to its effectiveness as a presence indicator (for example a single dropping is less important than a pile of dung). The spread of digital

* Corresponding author. Tel.: +390461882696; fax: +390461882672.
E-mail addresses: prea@uninsubria.it (D.G. Preatoni), tattonic@ing.unitn.it (C. Tattoni).

1574-9541/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecoinf.2011.07.002

cartography, GIS and GPS tools allows a more complex and spatially-oriented approach to wildlife studies, including the possibility to evaluate KAIs for a given land cover class or elevation range, an useful technique to underline distribution differences at small scale. The aim of this work is to create a tool that can automate the geodata set-up and the evaluation of KAI. This tool is designed to process georeferenced line transects and points of presence in order to calculate:

- “classic” KAI (2D);
- “3D-corrected” KAI, with transect length compensated for terrain morphology;
- partial KAIs, according to a user-defined polygon vector;
- “weighted” KAI, according to user-defined weights per sign of presence type.

The four KAI “flavors” can be mixed in any suitable way, e.g. one can calculate a 3D-corrected, weighted KAI, breaking transect paths by land cover class.

2. Materials and methods

This work has been accomplished using Free and Open Source Software (FOSS) in a Debian GNU/Linux environment. GRASS 6.x (GRASS Development Team, 2008; Neteler and Mitasova, 2008) with SQLite (Owens, 2006) support and tcl/tk GUI provided all the tools for spatial data analysis and operations on the attribute tables as well as the possibility to automate individual steps in a script (Neteler and Mitasova, 2008). R 2.7.1 (R Development Core Team, 2005) was used for statistical analysis and graphics, QGIS 1.x (Quantum GIS Development Team, 2011) and OpenJump 1.1 (The JUMP Pilot Project, 2008) for map layout. The scripts have been written in ASCII format as UNIX (bash) shell scripts with many different FOSS text editors. The choice of using FOSS allows to exploit the huge diversity of tools available from different programs, selecting the best one for each step and guaranteeing interoperability with other software. The use of FOSS in ecological studies supports the free spread of knowledge and allows others to repeat the experiments, a fundamental principle of research (Steinigera and Hay, 2009).

2.1. GeoData set for case study

The script has been tested with data collected during 2 years of field work (2006–2008) at regional scale in Val d'Aosta (Italian Alps near the French border, 45.537–45.888 N, 6.825–7.874 E, WGS84).

Line transect surveys of mountain hare (*Lepus timidus*) were carried out within the framework of the GESTALP (Modèles de GESTion pour la valorisation de la biodiversité et du pastoralisme dans le territoire ALPin transfrontalier) project, aimed at studying management models of biodiversity and pastures in the French and Italian Alps (Decout, 2008). Mountain hare was chosen because it is a typical alpine environment species, threatened by habitat destruction and global warming, for which knowledge on distribution is still scarce. On the whole region (about 3200 km^2), about 20 transects (from 3 to 5 km long) were identified and were covered monthly by two operators walking with a GPS receiver (Garmin GPS 60) that logged the track. Operators marked a waypoint (WP) at any sign of mountain hare presence they can detect in a two meter wide strip around the track. Field (GPS) data were stored in two shapefiles, one containing 250 transects and the other about 7600 WPs.

Other geodata used in the present work include: the Global Land Cover Facility SRTM Digital Elevation Model at 90 m spatial resolution (USGS, 2006), and a detailed vegetation map derived from field surveys around each transect, accounting for 26different vegetation types (unpublished data). Coordinate reference system used is UTM, zone 32N, WGS84 (EPSG:32032).

3. Running the program

The program, named *v.transect.kia*, is a bash shell script that works inside an active GRASS GIS session (GRASS Development Team, 2008), and it can be run, as any GRASS program, either from the command line or using GRASS Graphical User Interface facilities, as shown in Fig. 2. The program can as well be called from another script.

The script *v.transect.kia* is designed to automate KAI evaluation according to user choices and available input data. The program calculates KAI using a point map of sightings and a line map of transect paths, and saves its results in the original transect paths vector map attribute table. The tool can also split transects according to a further, optional vector polygon map (where available), calculating “partial” KAIs per each polygon class. The script is heavily commented, describing in detail the various operational phases presented in this paper and in Fig. 1. In order to add the script to a GRASS installation, it is sufficient to issue the GRASS *g.extension* command, fetching *v.transect.kia* from the official GRASS SVN Addons repository. Alternatively, the script can be copied into the appropriate binaries directory and *v.transect.kia* will be available among the other modules.

3.1. Inputs and outputs

The required inputs for the “classic” (i.e. basic) KAI calculation are two vector maps (either in GRASS vector or in ESRI Shapefile format), one for transect paths and one for point locations (waypoints) recorded along each path, and the name for the output map. The double format choice (GRASS native or ESRI Shapefile) allows the user to prepare the input and read the output with most of the available GIS software. The transect paths map must have linear features with no closed circuits, along with any relevant attribute (transect name, date, etc.); the waypoints map must contain point features, one for each sign of presence. Both maps must contain in their attribute table a field working as primary key (e.g. a *TRANS_ID* field), containing transect unique identifiers; furthermore, the waypoints map attribute table must contain two more fields, specifying the type of presence sign recorded (*TYPE*, at least 4 characters long), and the number of signs found at each waypoint, if applicable (*N*, numeric, 4 digits long, no decimals). This data configuration is the bare essential to query both attribute tables and calculate the index.

Other optional inputs for *v.transect.kia* are:

DEM an user-defined Digital Elevation Model, in GRASS raster format, to correct transect paths for the true 3D length;

Weights, a text file, formatted after the GRASS command *r.reclass*, containing weights to be assigned to different classes of presence signs;

Class map, a polygon vector map, used to segment transect paths and calculate partial KAIs, segmentation will be made on a “polygon class” attribute (see below);

Class field the field in *classes map* attribute table that identifies land use or elevation classes. If class information is stored in a field named *class* this option is not mandatory.

The additional input maps (i.e. *DEM* and *class map*) must be already present in a GRASS mapset. The weights file is a simple text file where each line refers to a particular sign of presence type code, giving it an user-defined weight:

type 1 = value
type 2 = value
…
type n = value

4. Description of KAI computation

The first step of *v.transect.kia* is to set up a SQLite connection to work with: all tables created by the script will in fact be in SQLite,

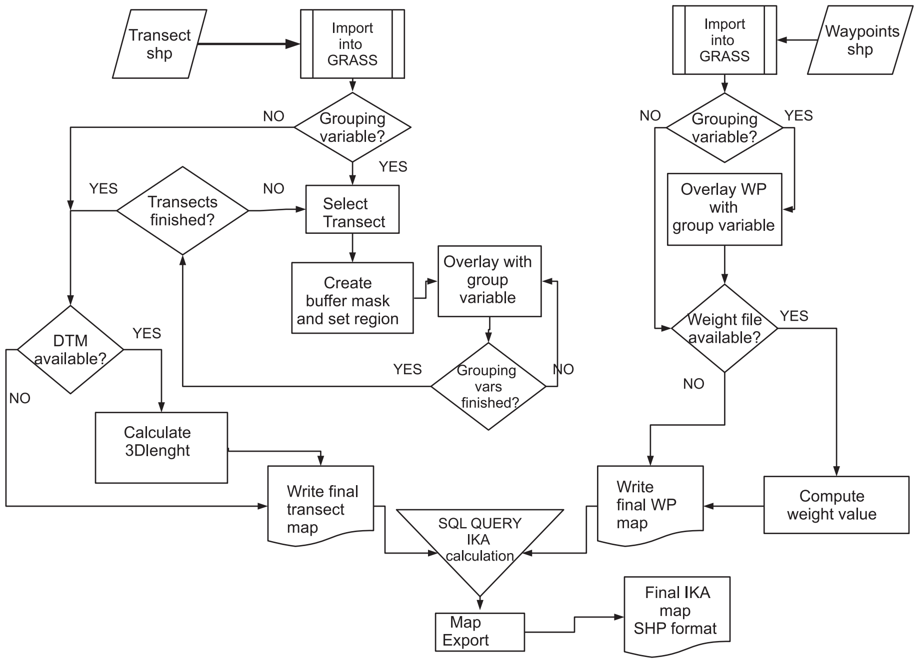

Fig. 1. Flow chart of the operations performed by the script *v.transect.kia*. See text for details.

since all the SQL manipulation capabilities needed are not present in the current GRASS 6.x *dbf* backend (Owens, 2006). With the default options, the script calculates the index in the "classic" form (Eq. (1)) using SQL to count the number of waypoints per transect, and writes the result in a new column of the output vector, preserving the original attribute table of the transect paths map (Fig. 3).

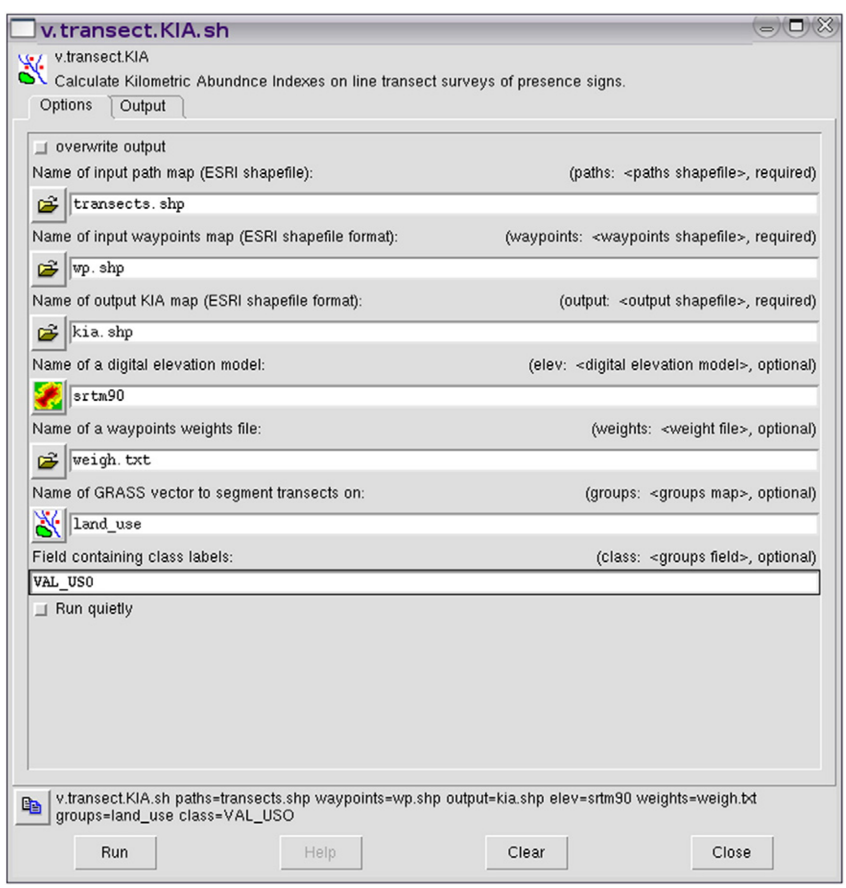

Fig. 2. Graphical User Interface for *v.transect.kia*.

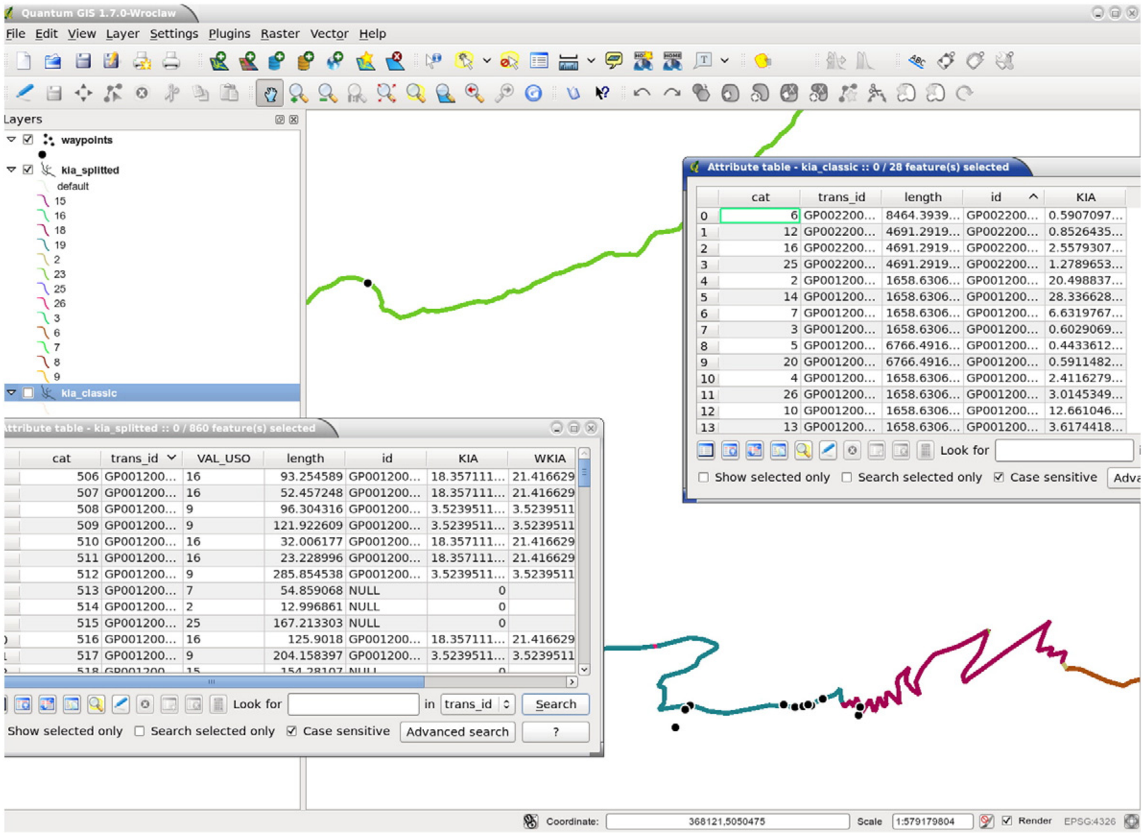

Fig. 3. Example of *v.transect.kia* output shown in a desktop GIS. The attribute table opened at lower left shows KAI values for a "classic" unweighted computation. The attribute table in the upper left shows partial weighted KAI values, split by habitat types derived from a land use map (not shown). Black dots represent GPS waypoints in a vector theme (presence signs), whereas the line represents a GPS track (transect path). Different colors indicate different habitat types.

If a DEM is available, the transect paths map is converted into a 3D vector by draping on the DEM: true length is then evaluated, added to the attribute table and used as length for KAI evaluation.

For weighted KAI calculation, *v.transect.kia* executes a query on the waypoints map attribute table, adding a new *weight* column containing the product of the number of signs by the relative weight, as specified in the weight table file. When this option is selected, two fields will be added to the output map, containing respectively both KAI and weighted KAI (*WKAI*) values (Fig. 3).

When a polygon map is selected for partial KAI evaluation by habitat type (or any other environmental variable, represented as polygons, even a coropleth map), the script first splits each transect into segments, overlaying each path with the polygon map and assigning to each transect segment the appropriate class code. The script then initialises two variables and counts both number of transects and habitat classes present in the maps.

Splitted segments that overlay each habitat class are stored in a temporary vector map that is patched at each step of the processing loop with the previously split transects. When all the transects have been split, the resulting attribute table is pruned of all unnecessary columns and only the columns deriving from the original transect map are kept. If requested so, *v.transect.kia* adds to the output attribute table a field containing habitat class codes, as well as a field containing the true 3D-corrected segment length.

In the case of partial KAI calculations, i.e. when the vector polygon habitat map option is selected, also the waypoint map is processed as described above, assigning each waypoint to the class of the polygon which falls into. The script then adds a new field to the waypoint attribute table, named after the habitat class field in the habitat class map, and populates it with the result of a *v.what.vect* query against the habitat map. The script then continues according to the other input options.

The waypoints and transect paths map attribute tables are at this point ready to be joined, according to the key formed by *TRANS_ID* (and *class*, if present) field present on both tables. Using tables instead of spatial relationships to assign each point to the correct transect allows the management of replicates. In fact, in abundance index-based surveys is not uncommon to have time replicates of the same transect, and plain spatial adjacence would not guarantee the correct pairing of a waypoint set with a transect path. Furthermore, waypoints almost never lie close to the path, due to GPS dilution of precision and to the survey strip width: therefore a selection based on some geometry-based function (e.g. distance) is not appropriate for all cases. Using attribute tables allows to work with two maps containing time replicates of the same transect.

The SQL query sums the number of waypoints according to the corresponding transect and habitat class pair, divides the result by the segment length (3D-corrected, if present) storing the result in a new column. If weighting is selected, the script multiplies each WP by its weight when calculating KAI.

The output of *v.transect.kia* is a new map (either in GRASS or Shapefile format) where KAI values are added to the original attribute table as new fields: the output is ready to be displayed in any GIS and the attribute table can be opened with a spreadsheet or a statistics tool.

5. Application to a case study

KAI and weighted KAI have been evaluated for all the 250 transects inside study area, 3D-correcting lengths, and results were used in R for further analysis. A first interesting result was that there was no significant advantage in giving different weights to all possible signs of presence (three-way ANOVA, interaction among transect, month and KAI type: $F_{(167,1224)} = 0.90$, $p = 0.98$; interaction between transect and KAI type: $F_{(20,1224)} = 0.23$, $p = 0.99$), at least for mountain hare, meaning that it is not necessary to record sign type when working in the field. An analysis of variance (ANOVA) (Faraway, 2002) showed a significant difference in KAI values for the different transects in time (two-way ANOVA, transect effect: $F_{(20,612)}=301.90$, $p<0.001$; month effect: $F_{(11,612)}=197.00$, $p<0.001$; interaction: $F_{(167,612)}=27.70$, $p<0.001$), meaning that the species was unevenly distributed in the study area: the spatial distribution of the average KAI per transect in the study area is presented in Fig. 4.

The output of the splitting routine is presented in Fig. 3, applied to a single transect. Spatially-partitioned partial KAIs were useful to understand the distribution of a species according to some environmental gradient. In this case study, partial KAI resulted significantly higher in some vegetation types than in others: dwarf pine (*Pinus mugo uncinata*) bush was the habitat with the highest value of signs per km, and in general semi-open environments presented higher numbers of mountain hare signs.

In this case study KAI values per habitat type showed a wide range of variation, highlighting the differences between vegetation types. In fact, when applying the index to a short portion of a transect that is rich in signs of presence, KAI values increase dramatically: in this case KAI values were log-transformed before carrying out a statistical analysis.

A complete discussion about the findings on mountain hare presence in Val d'Aosta goes beyond the scope of this paper: a complete analysis is found in the GESTALP project final report (Decout, 2008). The *v.transect.kia* script proved to be a useful tool in this case study, because leveraged on data automation, essential with large survey datasets, reducing the risk of manual errors when handling data. Further surveys were easily included in the data set while field monitoring proceeded, and new KAI maps were produced in minutes.

Analysis outcomes have already been used to create guidelines for future monitoring of the mountain hare, allowing a first wide-area semiquantitative estimate of an elusive species, otherwise difficult to monitor efficiently at a regional scale.

The script has been tested on the mountain hare field survey on a machine with an Intel® Core™ 2 Duo CPU with 4 GiB RAM running Debian GNU/Linux 2.6.24 testing, and the heaviest duty of processing partial KAIs for the whole dataset took about 20 minutes.

6. Conclusions

The program presented in this paper, *v.transect.kia*, a new module for the GRASS GIS, is designed to be as general as possible, allowing the user to run it from the command line or GUI, and to choose different input and processing options. Data set requirements are easy to meet, the only constraints are on field names and types for three columns in the input data attribute tables. The *v.transect.kia* script proved to be a useful tool in this case study, allowing for a high level of data processing automation, essential with large survey datasets, reducing the risk of manual errors when handling data. The automation of KAI evaluation also supported the data preparation process for further analyses, since the production of a correct KIA map guaranteed data coherency and saved time for further analysis. We hope that *v.transect.kia* will be useful to evaluate KAI or similar metrics in different contests or for other similar wide-area wildlife monitoring schemes.

The script and its manual page can be downloaded from the GRASS official add-ons page

https://svn.osgeo.org/grass/grass-addons/vector/v.transect.kia/

for the installation instructions please refer to the GRASS wiki page

http://grass.osgeo.org/wiki/GRASS_AddOns

The script is released under the terms of the GNU General Public License as published by the Free Software Foundation, the copyright belongs to GRASS Development Team, Damiano G. Preatoni and Clara Tattoni. Comments and feedback are welcome by the authors.

Acknowledgements

Authors are grateful to Emanuela Boggio and Luca Perlasca for support during fieldwork. Fundings: the module was developed during the GESTALP Project, EU-funded programme N 194 INTERREG III A. Provincia di Sondrio and Regione Autonoma Val d'Aosta supported fieldwork and logistics.

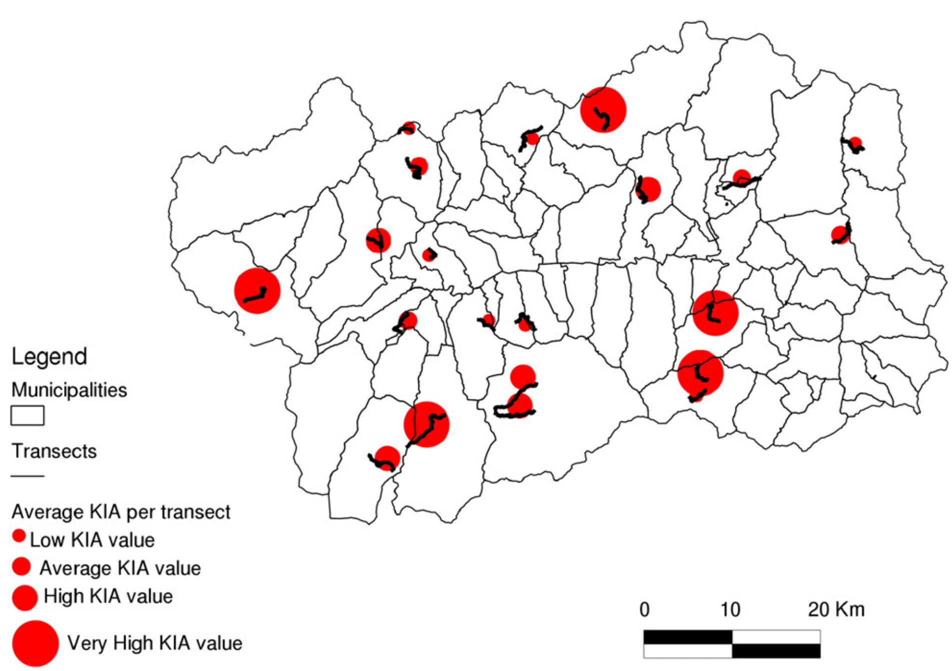

Fig. 4. Final KAI map for mountain hare study area. Red circles diameter is proportional to the average KAI value per transect replicate.

This is paper number 3 of the MoHaRe Project.

References

Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J., 1993. Distance Sampling: Estimating Abundance of Biological Populations. Chapman & Hall, London.

Burnham, K.P., Anderson, D.R., Laake, J.L., 1980. Estimation of density from line transect sampling of biological populations. Wildlife Monographs 72, 1–102.

de Thoisy, B., Brosse, S., Dubois, M., 2008. Assessment of large-vertebrate species richness and relative abundance in neotropical forest using line-transect censuses: what is the minimal effort required? Biodiversity and Conservation 17 (11), 2627–2644.

Decout, S., 2008. GESTALP. Modèles de gestion pour la valorisation de la biodiversitè et du pastoralisme dans le territoire alpin transfrontalier. Technical report.

Engeman, R.M., 2005. Indexing principles and a widely applicable paradigm for indexing animal populations. Wildlife Research 32, 203–210.

Faraway, J., 2002. Practical Regression and Anova using R.

GRASS Development Team, 2008. Geographic resources analysis support system (grass gis) software.

Maillard, D., Calenge, C., Jacobs, T., Gaillard, J.M., Merlot, L., 2001. The Kilometric Index as a monitoring tool for populations of large terrestrial animals: a feasibility test in Zakouma National Park, Chad. African Journal of Ecology 39 (3), 306–309.

Marchandeau, S., Aubineau, J., Berger, F., Gaudin, J.C., Roobrouck, A., Corda, E., Reitz, F., 2006. Abundance indices: reliability testing is crucial—a field case of wild rabbit *Oryctolagus cuniculus*. Wildlife Biology 12 (1), 19–27.

Neteler, M., Mitasova, H., 2008. Open Source GIS: A GRASS GIS Approach, volume 773 of *The International Series in Engineering and Computer Science*Third Edition. Springer, New York.

Owens, M. (2006). *The Definitive Guide to SQLite.*, apress edition. ISBN 978-1-59059-673-9 (Print) 978-1-4302-0172-4 (Online).

Quantum GIS Development Team, 2011. Quantum GIS Geographic Information System. Open Source Geospatial Foundation http://qgis.osgeo.org.

R Development Core Team, 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria3-900051-07-0.

Seber, G.A.F., 1973. The Estimation of Animal Abundance. Hafner Publishing Company, New York.

Steinigera, S., Hay, G.J., 2009. Free and open source geographic information tools for landscape ecology. Ecological Informatics 4 (4), 183–195.

The JUMP Pilot Project, 2008. OpenJUMP GIS—the free, java-based open source GIS. http://www.openjump.org.

USGS, 2006. Shuttle Radar Topography Mission, 1 Arc Second scene SRTM. Global Land Cover Facility, University of Maryland, College Park, Maryland www.landcover.org.

Vincent, J.P., Gaillard, J.M., Bideau, E., 1991. Kilometric index as biological indicator for monitoring forest roe deer populations. Acta Theriologica 36, 315–328.