Impact assessment of a high-speed railway line on species distribution: Application to the European tree frog (Hyla arborea) in Franche-Comté
Céline Clauzel, Xavier Girardet, Jean-Christophe Foltête

To cite this version:

HAL Id: hal-00814451
https://hal.archives-ouvertes.fr/hal-00814451
Submitted on 19 May 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Impact assessment of a high-speed railway line on species distribution: Application to the European tree frog (Hyla arborea) in Franche-Comté

Céline CLAUZEL, assistant professor * (celine.clauzel@univ-fcomte.fr) - corresponding author
Xavier GIRARDET, PhD student* (xavier.girardet@univ-fcomte.fr)
Jean-Christophe FOLTÊTE, professor* (jean-christophe.foltete@univ-fcomte.fr)*
ThéMA, CNRS - University of Franche-Comté
32 rue Mégevand
F-25 030 Besançon cedex France
Tel : +33 381 66 59 54

Abstract

The aim of the present work is to assess the potential long-distance effect of a high-speed railway line on the distribution of the European tree frog (Hyla arborea) in eastern France by combining graph-based analysis and species distribution models. This combination is a way to integrate patch-level connectivity metrics on different scales into a predictive model. The approach used is put in place before the construction of the infrastructure and allows areas potentially affected by isolation to be mapped. Through a diachronic analysis, comparing species distribution before and after the construction of the infrastructure, we identify changes in the probability of species presence and we determine the maximum distance of impact. The results show that the potential impact decreases with distance from the high-speed railway line and the largest disturbances occur within the first 500 m. Between 500 m and 3500 m, the infrastructure generates a moderate decrease in the probability of presence with maximum values close to -40%. Beyond 3500 m the average disturbance is less than -10%. The spatial extent of the impact is greater than the dispersal distance of the tree frog, confirming the assumption of the long-distance effect of the infrastructure. This predictive modelling approach appears to be a useful tool for environmental impact assessment and strategic environmental assessment. The results of the species distribution assessment may provide guidance for field surveys and support for conservation decisions by identifying the areas most affected.

Keywords

Environmental assessment; Transport infrastructures; Landscape graphs; Connectivity; Species distribution model; Hyla arborea.
1. Introduction

The extension of linear infrastructures is thought to be a major threat to biodiversity, much like other anthropogenic phenomena such as more intensive farming and urban sprawl (Forman and Alexander, 1998). The construction of transport infrastructures causes a direct loss of ecological habitats and road traffic generates wildlife–vehicle collisions (Coffin, 2007; Fahrig et al., 1995; Forman and Alexander, 1998; Geneletti, 2006; Trombulak and Frissell, 2000). Apart from these direct effects, linear infrastructures cause the loss of landscape connectivity (Forman and Alexander, 1998; Geneletti, 2004), which is recognized as a key functional factor for the viability of species and their genetic diversity (Fahrig et al., 1995). Major infrastructures such as motorways or high-speed railway lines act as barriers to the movement of animals through collisions and infrastructure avoidance because of traffic noise and isolate organisms in small subpopulations which become more sensitive to the risk of extinction (Forman and Alexander, 1998). This is especially the case for populations of amphibians whose daily movements and seasonal migrations mean they regularly cross the landscape matrix (Alford and Richards, 1999; Allentoft and O’Brien, 2010; Cushman, 2006; Eigenbrod et al., 2009; Fahrig et al., 1995; Lengagne, 2008; Scherer et al., 2012).

Several case studies have contributed to identifying and quantifying the effects of linear infrastructures on species distribution in many regions of the world, using various methods. Authors have related data describing species (e.g., abundance, road mortality) to proximity of infrastructures (Brotons and Herrando, 2001; Fahrig et al., 1995; Huijser and Bergers, 2000; Kaczynski et al., 2003; Li et al., 2010) and to the degree of habitat fragmentation (Fu et al., 2010; Serrano et al., 2002; Vos and Chardon, 1998). These studies measure the real impact of the infrastructure using species data collected after its construction. However, before the construction phase, an impact prediction stage is also necessary to compare alternative infrastructure routes (Fernandes, 2000; Geneletti, 2004; Vasas et al., 2009) or to guide the mitigation measures from the beginning of the project (Mörthberg et al., 2007; Noble et al., 2011).

Reviews by Geneletti (2006) and Gontier et al. (2006) show that the effects of landscape fragmentation are more difficult to predict than the direct loss of habitat. According to these authors, current assessment methods are often restricted to protected areas or to a narrow strip on either side of the infrastructure. However, landscape fragmentation may have consequences on a far broader scale (Forman, 2000).

To assess the long-distance effect of linear infrastructures on species distribution, models must include connectivity metrics that take into account both structural (arrangement of habitat patches) and functional (behaviour of the organisms) aspects. With this aim in mind, the development of methods based on graph theory is promising (Dale and Fortin, 2010; Urban et al., 2009). Graph theory, which originated in mathematics and the social sciences, is being used increasingly in landscape modelling (Galpern et al., 2010). Landscape graphs provide a simplified representation of ecological networks and require relatively few species data (Urban et al., 2009). This approach is considered an interesting trade-off between information content and data requirements (Calabrese and Fagan, 2004). It is used to quantify landscape connectivity for focal species (Ricotta et al., 2000; Urban and Keitt, 2001), select corridors between habitat patches (Alagador et al., 2012) and identify the key landscape features to be conserved in order to maintain connectivity (Bodin and Saura, 2010; Saura and Pasqual-Hortal, 2007). More recently graph-based connectivity metrics have been integrated as predictors into species distribution models (Awade et al., 2011; Decout et al., 2012; Foltête et al., 2012a; Pereira et al., 2011).

In this paper, we propose to assess and to map the long-distance effect of a high-speed railway (HSR) line on the distribution of the European tree frog (Hyla arborea) in the Franche-Comté region (eastern France). This study is based on a predictive modelling approach that estimates the potential impact of the HSR but does not measure the real impact. This approach is therefore put in place before the construction of the infrastructure and allows areas potentially affected by isolation to be mapped.
The choice of the European tree frog is justified by the recent decline in populations of this species in western Europe. This decline has several causes: climate change, increased UV-radiation (Alford and Richards, 1999), predation or competition, pollution and eutrophication of ponds (Borgula, 1993), road-kill (Elzanowski et al., 2008), and in particular the destruction and fragmentation of the tree frog’s habitat (Andersen et al., 2004; Cushman, 2006; Vos and Stumpel, 1996). In the region under study, the development of the HSR and consecutive changes in connectivity may therefore impact the viability of tree frog populations. As the potential effect of the HSR is assumed to occur up to a long distance from the line, and not only in its direct vicinity, we used the methodological framework of Girardet et al. (2013) and Foltête et al. (2012a) to integrate graph-based connectivity factors in the impact assessment. This method allowed us (1) to integrate connectivity metrics in a distribution model of the European tree frog defined at an initial state before the presence of the infrastructure, (2) to extrapolate the model after the construction of the infrastructure and (3) to compute the rate of change between the two probability maps from which to make a spatial assessment of the impact of the HSR.

2. Materials and methods

2.1. Data preparation

2.1.1. Study area

The study was carried out in the Ognon valley, a zone of 4600 km² in the Franche-Comté region of eastern France (Fig. 1). In this zone, where altitude ranges from 184 to 768 m, the landscape mosaic is dominated by forests (42% of total area), arable land (27%) and meadows (20%). The Ognon valley is strategic for environmental conservation because it contains many threatened species of birds (such as the little bustard, *Tetrax tetrax*), mammals (lesser horseshoe bat, *Rhinolophus hipposideros*), reptiles (European pond turtle, *Emys orbicularis*) and amphibians (European tree frog, *Hyla arborea*) (Paul, 2011).

In December 2011, a high-speed railway line came into service in the Franche-Comté region after four years of construction. It is part of a larger project which will improve connections for eastern France with both Paris and the south of France. This infrastructure is 140 km long and 30 m wide on average and crosses the study area from west to east following the Ognon valley. The line includes a total 1300 m of viaducts and 2000 m of tunnels. In this study, this infrastructure is considered as impassable either because it forms a physical barrier or because traffic noise can disrupt the animal behaviour (Eigenbrod, 2009; Lengagne, 2008) even when the infrastructure is on a viaduct. This simplified representation of reality is used to better predict the potential impact of the infrastructure on landscape connectivity and by repercussion on tree frog populations.
2.1.2. Study species

The tree frog is widely distributed in Europe from Spain to western Russia, but its populations have declined in north-western Europe over the last 50 years (Corbett, 1989). The species is classified as endangered and is on the IUCN Regional Red List of Threatened Species for Franche-Comté where it is mainly present in the Ognon valley (Fig. 1) (Pinston et al., 2000).

The tree frog occupies both an aquatic habitat for breeding and its larval period and a terrestrial habitat after breeding and during hibernation. The breeding pond consists in shallow and sunny ponds, marshlands, gravel-pits or river pools. Pond size does not appear decisive and ranges from 1 m² to 4000 m² (Grosse and Nöllert, 1993). Although its breeding pond is essential for reproduction, the tree frog spends most of its time on land. Its terrestrial habitat is composed of dense vegetation (trees,
shrubs, bushes) which is well exposed to the sun (Stumpel, 1993). The species often prefers edge habitats in agricultural environments: banks, ditches, edges of fields or of forests (Pellet et al., 2004). The seasonal migrations between this terrestrial habitat from the breeding pond range from 250 m to 1000 m (Stumpel, 1993). Vos and Stumpel (1996) report that its presence in a pond is correlated with the amount of terrestrial habitat surrounding the pond.

Dispersal events are very important in the tree frog’s life cycle. Juveniles disperse from the breeding pond into the surrounding landscape to join other ponds. Despite the fidelity of the tree frog to its breeding ground, some adults may disperse and change ponds (Fog, 1993). These dispersal events allow individuals to colonize new ponds or to recolonize sites where the species is nearing extinction (Vos, 1999). Observed dispersal distances are generally less than 2000 m but may reach up to 4000 m (Vos and Stumpel, 1996). In this study, 2500 m was arbitrarily considered as the maximum dispersal distance.

2.1.3. Species data

Species presence was identified by listening for the calling males during the breeding period (April to July) and recorded between 1997 and 2010 by field specialists of a wildlife association (LPO Franche-Comté). Listening was done from sunset to midnight on warm, wet nights. Surveys were conducted near ponds known to be occupied by the tree frog and near other potentially favourable ponds. Each site was visited several times during the breeding period to ensure maximum probability of detection. Field specialists considered that a pond was occupied if at least one calling male was heard. In this study, the presence points corresponded to breeding ponds that were occupied at least one year from 1997 to 2010.

These surveys identified 87 geo-located presence points in the study area. These presence points were sampled using a rule of a minimum distance of 500 m between points to remove redundant points, such as the same point at different times, or points too close together. In this way 57 points were selected in all. From the surveys, individuals were mostly observed in ponds (22), marshlands (15) and river pools (13). The remaining 7 individuals were found in grasslands and forest edges. Half of the presence points were located within a 2500 m-wide strip on either side of the HSR line.

Because absence data for the tree frog were unavailable, the study used pseudo-absences instead of real absences (Pearce and Boyce, 2006). Starting from a grid of 5000 m-sized cells corresponding to twice the dispersal distance, a set of pseudo-absence data was generated by randomly sampling one point within each cell without a presence point. In this sampling procedure, a rule of inter-point distance of 2500 m was applied to avoid having two points close together. The resulting set of 119 pseudo-absence points and 57 presence points was considered as the target variable likely to be modelled from landscape predictors.

2.1.4. Landscape data

The study required the creation of two landscape maps, the first one describing the initial state before construction of the HSR line, the second map including the linear infrastructure. Except for the HSR line, the land use was identical on both maps so only the impact of the infrastructure would be estimated. Different data sources were combined using ArcGis 10 (ESRI, 2011). The 1 m-accuracy vectorial landscape databases provided by French cartographic services (DREAL Franche-Comté, BD Topo IGN) were used to represent wetlands, hedgerows, forests, buildings, roads, railways and rivers. In agricultural areas, remotely sensed data (IRS PAN+LISS-III, 7 m spatial resolution) were used to separate grassland and bare ground. A morphological spatial pattern analysis (MSPA) (Vogt et al., 2007) was also applied to the forest category to identify forest edges. All these data elements were combined into a single raster layer with a resolution of 10 m. In this raster layer, the HSR line was represented by a 3-pixel wide line to reflect its actual size. It also avoided potential discontinuities...
induced by the conversion of linear elements in a raster format (Adriaensen et al., 2003; Gurrutxaga et al., 2011). Altogether, thirteen landscape categories were obtained (Table 1).

Table 1. Sets of resistance values for each of the 13 landscape categories.

<table>
<thead>
<tr>
<th>Landscape categories</th>
<th>Function</th>
<th>Uniform</th>
<th>Equidistant</th>
<th>Compressed</th>
<th>Contrasting</th>
<th>Highly contrasting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ponds</td>
<td>Aquatic</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Hedgerows</td>
<td>Terrestrial</td>
<td>1</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Forest edges</td>
<td>Terrestrial</td>
<td>1</td>
<td>25</td>
<td>10</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Rivers</td>
<td>Suitable</td>
<td>1</td>
<td>50</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Wetlands</td>
<td>Suitable</td>
<td>1</td>
<td>50</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Wooded grasslands</td>
<td>Suitable</td>
<td>1</td>
<td>50</td>
<td>20</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Grasslands</td>
<td>Suitable</td>
<td>1</td>
<td>75</td>
<td>30</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td>Roads</td>
<td>Unfavourable</td>
<td>1</td>
<td>75</td>
<td>30</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td>Railways</td>
<td>Unfavourable</td>
<td>1</td>
<td>75</td>
<td>30</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td>Bare ground</td>
<td>Barrier</td>
<td>1</td>
<td>100</td>
<td>40</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>Forests</td>
<td>Barrier</td>
<td>1</td>
<td>100</td>
<td>40</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>Buildings</td>
<td>Barrier</td>
<td>1</td>
<td>100</td>
<td>40</td>
<td>100</td>
<td>1000</td>
</tr>
<tr>
<td>Motorway (A36)</td>
<td>Barrier</td>
<td>1</td>
<td>100</td>
<td>40</td>
<td>100</td>
<td>1000</td>
</tr>
</tbody>
</table>

2.1.5. Definition of cost values

As the ability of the tree frog to disperse depends greatly on the surrounding matrix, it is expected that cost distances should be more relevant than Euclidean distances. The ecological literature and expert opinions were thus used to classify each landscape category according to its resistance to movement (Table 1). Radio tracking experiments (Pellet et al., 2004; Vos and Stumpel, 1996) have concluded that wooded grasslands and linear elements like hedgerows or forest edges facilitate movement and often provided the species’ terrestrial habitat. Rivers and wetlands are less favourable because their permeability depends on the density of vegetation. Conversely, grasslands, roads and railways tend to constrain movement. Finally, the cores of forest patches, bare ground, buildings and motorways are considered highly impassable and are mostly avoided by the tree frog.

Given the difficulty in assigning cost values to landscape categories from the ecological literature (Rayfield et al., 2010), several resistance sets (Table 1) were tested in order to assess the sensitivity of the model to these cost values (Pereira et al., 2011; Verbeylen et al., 2003). In the uniform resistance set, all landscape categories had a resistance value of 1, meaning that the landscape matrix was considered as homogeneous and Euclidean distance alone mattered. The equidistant set included values from 1 to 100 in increments of 25. The compressed set included values from 1 to 40 in increments of 10. The contrasting set included values close to either the minimum (1) or the maximum (100) values. The highly contrasting set was similar to the contrasting set with high cost values for the ‘barrier’ category.

2.2. Landscape graph analysis

2.2.1. Landscape graph construction

The nodes of the graphs corresponding to the habitat patches were defined as being the aquatic habitat units (breeding ponds) adjacent to an area of potential terrestrial habitat. From the resistance sets (Table 1), five sets of links characterized by a complete topology were defined to connect all the patches (Galpern et al., 2010). Each link contained both the recording of its Euclidean distance and its cumulative cost distance. As least-cost distances were used, the maximum dispersal distance of 2500
m was converted to a value expressed in cost units, using a regression analysis where the Euclidean distance of all the links was considered as a linear function of the cost distance. Given the maximum dispersal distances thus obtained, five thresholded graphs were built by removing the links whose cumulative cost was higher than the maximum dispersal distance. As the HSR line on the second landscape map was considered to be an absolute barrier, its cost value was chosen so as to remove any links crossing the infrastructure when the graph was thresholded.

2.2.2. Computation of patch-level metrics

Many connectivity metrics have been developed in the literature (Galpern et al., 2010; Rayfield et al., 2011). They may be calculated at different levels of the graph structure (global, by component, local) but in this study, only the patch-based level is relevant for providing predictive variables in a distribution model. From the works of Bunn et al. (2000), Foltête et al. (2012a) and Urban and Keitt (2001), three metrics were defined representing the way in which a habitat patch can contribute to population dynamics.

(1) For a given patch i, recruitment \( R_i \) represents the demographic potential which is intrinsic to the patch and independent of the graph. This metric may be quantified by patch size or the area of available resources within a given perimeter \( a_i \) such that:

\[
R_i = a_i \quad \text{(equation 1)}
\]

In this study, the potential recruitment of a patch corresponds to the amount of terrestrial habitat within a 500 m radius around the ponds, following Vos and Stumpel (1996) and Pellet et al. (2004) who emphasize the greater importance of such habitat over pond size.

(2) The weighted dispersal flux (F) represents the capacity of a patch to disperse individuals to surrounding patches independently of its own potential recruitment. This metric depends on the patches connected to the focal patch, weighted by both their recruitment and a decreasing function of the distance that separates them from the focal patch:

\[
F_i = \sum_{j=1}^{n} R_j \exp(-\alpha d_{ij}) \quad \text{(equation 2)}
\]

where \( R_j \) is the recruitment of patch \( j \), \( d_{ij} \) is the distance between patches \( i \) and \( j \) and \( \alpha \) is a parameter representing the intensity of the distance effect.

(3) The betweenness centrality of long distances (BCl) represents the potential for a patch to be crossed by a path linking other patches, by giving more weight to the longer paths (Foltête et al., 2012a). This metric was used to access the potential role of long-distance paths for re-colonization of habitat patches after an extinction (Urban et al., 2009). Each path is weighted both by the recruitment of the two end patches and by the probability of occurrence of the path to improve its ecological significance, in the same way as the adaptation of the BC index proposed by Bodin and Saura (2010). For a patch i, this index is given by:

\[
BCli = \sum_{j=1}^{n} \sum_{k=1}^{n} a_j a_k (1 - p^*_{jk}) \quad \text{for } j, k \in P_{jk} \quad \text{(equation 3)}
\]

with \( a_j \) and \( a_k \) being the potential recruitment of the patches \( j \) and \( k \), \( p^*_{jk} \) the maximum product probability between patches \( j \) and \( k \), and \( P_{jk} \) the set of patches crossed by the least-cost path between patches \( j \) and \( k \).
F and BCI were adjusted using the same threshold distance of the graphs as a reference value corresponding to a movement probability of 0.05. In order to test the sensitivity of the model to the distance used both to threshold the graph and to adjust the metrics, four other threshold values were tested, in accordance with the movement abilities of the tree frog: 1000 m and 4000 m for the dispersal distance, 250 m and 750 m for the terrestrial habitat around ponds. Graphab 1.0 software (Foltête et al., 2012b) was used to construct the landscape graph and compute the patch-level metrics.

2.3. Species distribution modelling and diachronic analysis

2.3.1. Integration of patch-based metrics in a species distribution model (SDM)

An SDM was developed consisting in modelling the presence/absence variable from predictors representing functional connectivity. Following Foltête et al. (2012a), the patch-based metrics computed from the graphs were generalized to any point of the study area to provide predictive factors that could be integrated into this SDM. As the data points were not necessarily located within a patch, they were connected to the nearest patches according to the maximum dispersal distance and were associated with the mean of the local values of R, F and BCI. As individuals have a decreasing probability of dispersion with distance away from the patches, the values are weighted by a decreasing function of the distance from the edge of the patch such that

\[ w = \exp(-\alpha d) \]

The statistical analysis was performed using Microsoft® Excel 2010/XLSTAT©-Pro (Version 2012.6.02, 2003, Addinsoft, Inc., Brooklyn, NY, USA). In order to reduce multicollinearity, the correlation coefficients between each pairwise combination of explanatory variables were calculated using a Spearman rank correlation at a 5% confidence level. If pairs were highly correlated (>0.70), the less significant variable in relation to presence of the tree frog was eliminated. The relationship between the presence and absence of the species and predictive variables was analysed using a generalized linear model (McCullagh and Nelder, 1989) with binomial distribution and a logit link function (Hosmer and Lemeshow, 2000). Several logistic regression models were computed by testing the dependent variable against all possible combinations of variables. The model that better explained the presence of the tree frog was selected based on the minimization of the Akaike information criterion (AIC) value. Models with AIC differences (\( \Delta \text{AIC} \)) of less than 2 with the best model (i.e. the one with the smallest AIC) were considered as competing models (Burnham and Anderson 2002). McFadden R-square values were calculated for the final model to give a measure of model fit by quantifying the amount of variability in the dependent variable explained by predictors. The assessment of the final model’s accuracy was based on the measures derived from the confusion matrix: sensitivity, i.e. the proportion of observed presences correctly predicted; specificity, i.e. the proportion of observed absences correctly predicted; Cohen’s kappa and the area under the ROC curve (AUC) (Fielding and Bell, 1997). Spatial autocorrelation of the residuals of the best-fitting model was checked by using a Moran’s I correlogram to test assumptions of constant variance. As the data set was too small to be split into separate data sets, the final model was evaluated by the bootstrap resampling technique, i.e. using a random sampling with replacement from the original dataset, following Efron and Tibshirani (1993). If the statistical model is globally validated, the SDM can be extrapolated to the entire study area with a spatial resolution of 50 m providing a continuous map of the probability of presence before the construction of the HSR line (time \( t_0 \)).

2.3.2. Species distribution assessment

The landscape map was then modified by including the infrastructure (time \( t_1 \)); a new graph was built with the same parameters as for time \( t_0 \) and the patch-based metric values were computed. From these data describing the state at \( t_1 \), the model was again extrapolated to map the probability of tree
frog presence after the construction of the HSR line. Diachronic data were thus produced describing
the species distribution models before and after the construction of the infrastructure.

The rate of change in the probability of presence was calculated for each cell from the formula:
\[ \Delta p = (p_{t1} - p_{t0})/p_{t0}. \] This rate represents the local variation in probability of tree frog presence according
to the model. The mapping of \( \Delta p \) in ArcGis 10 (ESRI, 2011) was used to visualize the impact of the
HSR line and to identify areas likely to be most affected by the infrastructure. To investigate the
relationships between the level of disturbance and the distance from the route, a regular grid with a
100 m spatial resolution was generated over the entire study area. A point was sampled at each node
of this grid and contained both the rate of change in the probability of presence and the distance to the
HSR line. Since this procedure is irrelevant if the species was absent before the construction of the
infrastructure, only points with a presence probability of more than 0.5 before the HSR line was
constructed were selected. The average and maximum rate of change was plotted against the
distance from the infrastructure grouped into class intervals of 500 m. The two curves provide
information about the spatial structure of the impact and the maximum distance of disturbance.

3. Results

On the landscape map for before the HSR line was built, 1464 habitat patches were identified varying
from 0.01 to 86 ha (mean 0.6 ha). The resistance sets proposed in Table 1 were used to generate five
graphs. For each graph, the metrics R, F and BCI were computed and analysed as predictive
variables for tree frog presence. The correlation analysis (Table 2) showed that the R and F metrics
were closely correlated. Of the two metrics, F was removed because it is the least significant variable
and the more closely correlated with BCI.

Table 2. Spearman correlation coefficients among study variables with a significance level of 5%

<table>
<thead>
<tr>
<th>Variables</th>
<th>R</th>
<th>F</th>
<th>BCI</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>0.835</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BCI</td>
<td>0.516</td>
<td>0.659</td>
<td></td>
</tr>
</tbody>
</table>

The model’s sensitivity to the cost values was analysed by comparing the global results of the logistic
regressions (Table 3). For each resistance set, the best model was always the one with R and BCI
metrics. The highly contrasting resistance set was the most relevant with the smallest AIC value
(124.68). The other four resistance sets were less relevant with AIC differences > 10.

Table 3. Comparison of the best selected models using a logistic regression under each of the five
resistance sets to predict the occurrence of the tree frog. All of the best selected models listed in the
table used both R and BCI metrics. Models were ranked using the change in the Akaike Information
Criterion (\( \Delta \text{AIC} \)). Fit statistics included the Akaike weight \( (\omega_i) \) and the area under the receiver
operating characteristic curve (AUC)

<table>
<thead>
<tr>
<th>Resistance sets</th>
<th>AIC</th>
<th>( \Delta \text{AIC} )</th>
<th>( \omega_i )</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>147.38</td>
<td>22.70</td>
<td>0.00</td>
<td>0.869</td>
</tr>
<tr>
<td>Equidistant</td>
<td>145.35</td>
<td>20.67</td>
<td>0.00</td>
<td>0.875</td>
</tr>
<tr>
<td>Compressed</td>
<td>145.63</td>
<td>20.95</td>
<td>0.00</td>
<td>0.879</td>
</tr>
<tr>
<td>Contrasting</td>
<td>135.86</td>
<td>11.18</td>
<td>0.00</td>
<td>0.911</td>
</tr>
<tr>
<td>Highly contrasting</td>
<td>124.68</td>
<td>0.00</td>
<td>1.00</td>
<td>0.941</td>
</tr>
</tbody>
</table>

From the highly contrasting resistance set, the metrics R and BCI were calculated for several
threshold distances to assess the model’s sensitivity. For each threshold distance, the best model was
still the one which included both the R and BCI metrics. For dispersal distances, the best model was
the one with the lowest values ($0 > \Delta \text{AIC} < 2$ for 1000 m and 2500 m) whatever the terrestrial habitat distance used. Conversely, models using the dispersal distance of 4000 m were less relevant with $\Delta \text{AIC}$ value $> 10$. According to the graph thresholded at 1000 m, the mapping of the links shows that the habitat network is highly fragmented in the Ognon valley (Fig. 2). Conversely, the network is more connected and joins that of the Doubs valley to the south with the graph thresholded at 4000 m. For terrestrial habitat distance, models using shorter distances (250 m or 500 m) are better whatever the dispersal distance (Table 4).

Table 4. Comparison of the best selected models using a logistic regression under each of the threshold distances to predict the occurrence of the tree frog. The dispersal distance was used to threshold the graph and to set the BCl metric. The terrestrial habitat distance was used to set the R metric. All of the best selected models listed in the table used both R and BCl metrics. Models were ranked according to $\Delta \text{AIC}$. Fit statistics included the Akaike weight ($\omega_i$) and the area under the receiver operating characteristic curve (AUC).

<table>
<thead>
<tr>
<th>Model</th>
<th>Dispersal distance (Euclidean)</th>
<th>Dispersal distance (cost unit)</th>
<th>Terrestrial habitat distance (Euclidean)</th>
<th>Terrestrial habitat distance (cost unit)</th>
<th>AIC</th>
<th>$\Delta$ AIC</th>
<th>$\omega_i$</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1000m</td>
<td>1460</td>
<td>250m</td>
<td>995</td>
<td>123.71</td>
<td>0.12</td>
<td>0.48</td>
<td>0.932</td>
</tr>
<tr>
<td>2</td>
<td>2500m</td>
<td>2392</td>
<td>250m</td>
<td>995</td>
<td>123.59</td>
<td>0</td>
<td>0.51</td>
<td>0.943</td>
</tr>
<tr>
<td>3</td>
<td>4000m</td>
<td>3323</td>
<td>250m</td>
<td>995</td>
<td>133.72</td>
<td>10.13</td>
<td>0</td>
<td>0.921</td>
</tr>
<tr>
<td>4</td>
<td>1000m</td>
<td>1460</td>
<td>500m</td>
<td>1150</td>
<td>124.8</td>
<td>0.12</td>
<td>0.48</td>
<td>0.934</td>
</tr>
<tr>
<td>5</td>
<td>2500m</td>
<td>2392</td>
<td>500m</td>
<td>1150</td>
<td>124.68</td>
<td>0</td>
<td>0.51</td>
<td>0.942</td>
</tr>
<tr>
<td>6</td>
<td>4000m</td>
<td>3323</td>
<td>500m</td>
<td>1150</td>
<td>135.03</td>
<td>10.35</td>
<td>0.00</td>
<td>0.917</td>
</tr>
<tr>
<td>7</td>
<td>1000m</td>
<td>1460</td>
<td>750m</td>
<td>1305</td>
<td>125.27</td>
<td>0.17</td>
<td>0.48</td>
<td>0.933</td>
</tr>
<tr>
<td>8</td>
<td>2500m</td>
<td>2392</td>
<td>750m</td>
<td>1305</td>
<td>125.1</td>
<td>0</td>
<td>0.52</td>
<td>0.932</td>
</tr>
<tr>
<td>9</td>
<td>4000m</td>
<td>3323</td>
<td>750m</td>
<td>1305</td>
<td>135.55</td>
<td>10.45</td>
<td>0.00</td>
<td>0.915</td>
</tr>
</tbody>
</table>

Fig. 2. Complete graphs thresholded at different dispersal distances: a. 1000 m ; b. 2500 m and c. 4000 m.

Given the results of the logistic regression (Table 3), four models (1, 2, 4 and 5) were considered as competitors for predicting the presence of the tree frog. As all of these models used two variables, the final model was selected based on the highest AUC value. So, the analysis used model 2 with the dispersal distance of 2500 m and the terrestrial habitat distance of 250 m.

The two variables in this model explained 44% of the variance of the response variable (Table 5). The addition of altitude to the metrics R and BCI increased the model’s explanatory power, which provided...
an AIC of 107.56 and explained 52% of the variance. Results of the regression analysis indicate that all variables are significant (p-value<0.01) and that BCl, followed by R, are positively correlated with the presence of the tree frog. Conversely, altitude has a negative effect on the species. This model has good discriminating power with an AUC value of 0.943 and a high correct classification rate with a Cohen’s kappa of 0.69. However, absence is predicted better than presence (93% correctly classified against 74%). No significant spatial autocorrelation of the residuals was detected.

Table 5. Results of logistic regression modelling for the presence of the tree frog using a dispersal distance of 2500 m and a terrestrial habitat distance of 250 m (model 2). The limits of the 95% interval resulted from 10 000 bootstrap samples.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Standardized coefficients</th>
<th>SD</th>
<th>Wald $\chi^2$</th>
<th>Lower bound of 95% confidence</th>
<th>Upper bound of 95% confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>1.395</td>
<td>0.342</td>
<td>16.597***</td>
<td>0.588</td>
<td>2.404</td>
</tr>
<tr>
<td>BCl</td>
<td>1.501</td>
<td>0.695</td>
<td>4.658*</td>
<td>0.138</td>
<td>2.864</td>
</tr>
<tr>
<td>Altitude</td>
<td>-2.15</td>
<td>0.771</td>
<td>7.780**</td>
<td>-7.210</td>
<td>-0.373</td>
</tr>
</tbody>
</table>

*** p-value < 0.0001, ** p < 0.001, * p < 0.01.

Model with R and BCl (model 2): McFadden $R^2 = 0.44$; AIC value = 123.59
Model 2 + altitude: McFadden $R^2 = 0.52$; AIC value = 107.56

From the two landscape maps, extrapolation of this model to all cells in the study area provided two probability maps of tree frog occurrence, before and after construction of the infrastructure (Fig. 3). The highest values of presence probabilities are located in the Saone valley in the north, the Ognon valley in the centre and the Doubs valley in the south.

Fig. 3. Probability of tree frog occurrence before (a) and after (b) the construction of the HSR line. These probabilities result from the SDM using model 2 including R, BCl and altitude as predictors.
Fig. 4. Estimated impact of the construction of the HSR line on the probability of tree frog presence, according to the species distribution assessment.

The diachronic analysis shows that the impact of the HSR line is not uniform across the study area (Fig. 4). In the east, the probability values decrease sharply (from -60% to -99%) over a distance of several kilometres from the infrastructure. In the west, the impact is more focused on its route with values ranging from -20% to -70%. In the central part, the probabilities of presence seem to remain stable near the infrastructure but decrease slightly between 3 and 5 km. The superimposition of these results on the landscape map shows that the areas impacted by the HSR line correspond to the favourable landscape categories for the tree frog.

The curves of the average and maximum rate of change (Fig. 5) confirm that the impact decreases with distance from the HSR line. The largest disturbances (from -60% to -80%) occur within the first 500 m. Between 500 m and 3500 m, the HSR line generates a moderate disturbance with 90% of the points having a value of less than -50%. Beyond 3500 m the average disturbance is less than -10%. The 9th decile curve shows an increase of disturbance around 7 km with maximum values close to -25%. These points correspond to the impact estimated by the model in the north-east of the study area on the south side of the HSR line (Fig. 4). This area is occupied by several wetlands connected by a narrow corridor to the Ognon valley and could be potentially affected by the infrastructure despite its distance.
4. Discussion

This study was designed to prospectively assess the impact of the construction of an HSR line on the distribution of the tree frog in a study area located in the region of Franche-Comté. The combination of graphs and species distribution modelling is relevant in this study as it improves the ecological significance of models and makes a diachronic analysis feasible.

In terms of modelling of the species distribution, the final model is significant with only two graph-based metrics and altitude as predictors. These results confirm the role of the pond network (BCI) and the area of terrestrial habitat (R) as mentioned in Pellet et al. (2004) and Vos and Stumpel (1996). Given that BCI proves to be the most significant predictor, the results of the modelling suggest that long-distance migrations play an important role in the distribution of the tree frog. Such findings are consistent with earlier studies showing that this species is characterized by a metapopulation structure (Arens et al., 2006; Pellet et al., 2004; Vos, 1999). However, this result could be partly related to the extent of the study area and to the landscape configuration. Linear-shaped valleys result in elongated subgraphs which increase potential patch crossings by a path linking other patches. As these valleys contain many presence points, the role of long-distance paths may explain the significant contribution of the metric BCI.

The analysis of sensitivity to cost values suggests that highly contrasting values between favourable and unfavourable landscape categories are especially relevant. Increasing the maximum cost value in the highly contrasting set significantly improved the model. It is essential therefore to take into account the effects of the landscape matrix, in particular barriers, when predicting the occurrence of the tree frog. Variability in the relevance of the model confirms the need to test several resistance sets as mentioned by Sawyer et al. (2011) and Spear et al. (2010). The model was also tested with several threshold distances because of the many dispersal distances of the tree frog mentioned in the literature (Fog, 1993; Pellet et al., 2004; Vos, 1999). The shorter dispersal distances (1000 m and 2500 m) are the more relevant, which is consistent with the analysis conducted over the entire region of Franche-Comté (Foltête et al., 2012a). For the terrestrial habitat, the shorter distance (250 m or 500 m) is also more relevant, which suggests that the presence of the tree frog is related to the proximity of this habitat around ponds.

The diachronic analysis of the presence probability values shows that the potential impact of the HSR line declines with distance from this infrastructure. The impact ranges from a few metres to several
kilometres, the largest decreases in the presence probability being close to the HSR line. However, in some sections, the impact is not necessarily located along the route and may occur 3 km from the HSR line. This variability is related to the landscape configuration and the initial state of the connectivity of habitat patches in the study area. Indeed, the extent of disturbance depends on the size of the subgraphs in the network, with a large subgraph increasing the distance of the impact as in the north-east of the study area. The maximum rate of change plotted against the distance from the infrastructure shows that the disturbance occurs at a distance greater than the maximum dispersal distance of the tree frog, thus confirming the assumption of the long-distance effect of the infrastructure.

This graph-based approach provides an approximation of the potential impact based on the model but not a hard and fast measurement of the true impact. The extent of the perturbation may be over-estimated here by the species distribution sampling points. In order to validate our approach, these findings could be confirmed by field observations in the coming years to test whether the real impact of the infrastructure is similar to that predicted by the model. These surveys will also identify the causes of extinction of breeding populations. Several environmental factors may lie behind the extinction process and compound to the long-distance effect of the infrastructure.

The methodological approach used appears to be a handy tool for planners in assessing the impact of linear infrastructures on different spatial scales including the regional level, which is recognized as a gap in current methods (Fernandes, 2000; Geneletti, 2006; Mörtberg et al., 2007).

The variation in the presence of the species can help to optimize the location of new protection areas or mitigation measures such as wildlife crossing structures, by identifying the areas most affected by the infrastructure. The results also provide information about the maximum distance of the impact, which is often difficult to assess. In the case of the HSR line in the Franche-Comté region, the environmental assessment studies focused only on a strip of 800 m on either side of the HSR line, which allowed the creation of new ponds to replace those destroyed by the construction of the infrastructure.

This study takes a species-oriented approach, which provides for a more realistic ecological analysis by considering the behaviour of a focal species. The species distribution model can also be validated using presence/absence data. This approach is particularly relevant to the conservation of threatened species. Nevertheless, given the diversity of wildlife, the impact assessment of an infrastructure cannot be limited to a single species. In this case, the habitat-oriented approach may seem more relevant by focussing on certain landscape categories such as forests or wetlands (Gurrutxaga et al., 2011; Mancebo Quintana et al., 2010; Tannier et al., 2012). But these habitat models raise the difficulty of the validation needed in an impact assessment process. Furthermore, the same habitat category may contain several species that have different behaviours and dispersal abilities and will not be affected in the same way by infrastructure. The habitat-oriented approach does not allow us to assess these different impacts. For a generic approach when species presence data are missing, the virtual species approach developed by Hirzel et al. (2001) to model the distribution of several species might be one way to quantify the impact in habitat patches and in their vicinities (Girardet et al., 2013).

Conclusion

The combination of landscape graphs and SDM seems a relevant way to improve methods for assessing the impact of linear infrastructure on the distribution of a given species. The species distribution assessment provides spatial indicators about the magnitude of the impact. The analysis carried out in the Franche-Comté region shows that probabilities of tree frog presence decrease by -90% in some areas. The extent of the impact is often greater than the maximum dispersal distance of the species. Used in a predictive approach, this method can assess several scenarios of linear
infrastructure routes in order to select the one that generates the least impact on species distribution. After the construction of the infrastructure, the results may guide land planners in identifying the areas most affected with a view to conservation.

Acknowledgements

The research has been funded by the French Ministry of Ecology, Energy, Sustainable Development and the Sea (ITTECOP Program). Field data about Hyla arborea were provided by LPO Franche-Comté. The graph analysis was conducted as part of the Graphab project managed by the USR 3124 MSHE 744 Ledoux. Computations were performed on the supercomputer facilities of the MSHE Ledoux.

References


Tannier, C., Foltête, J.-C., Girardet, X., 2012. Assessing the capacity of different urban forms to preserve the connectivity of ecological habitats. Landscape and Urban Planning. 105. 128-139.


