RESEARCH ARTICLE

Using electric network theory to model the spread of oak processionary moth, *Thaumetopoea processionea,* in urban woodland patches

Daniel J. Cowley · Oliver Johnson · Michael J. O. Pocock

Abstract

Context Habitat fragmentation is increasing as a result of anthropogenic activities, especially in urban areas. Dispersal through fragmented habitats is key for species to spread, persist in metapopulations and shift range in response to climate change. However, high habitat connectivity may also hasten the spread of invasive species.

Objective To develop a model of spread in fragmented landscapes and apply it to the spread of an invasive insect in urban woodland.

Methods We applied a patch-based model, based on electric network theory, to model the current and predicted future spread of oak processionary moth (OPM: *Thaumetopoea processionea*) from its source in west London. We compared the pattern of 'effective distance' from the source (i.e. the patch 'voltage' in the model) with the observed spread of the moth from 2006 to 2012.

Results We showed that 'effective distance' fitted current spread of OPM. Patches varied considerably in their 'current' and 'power' (metrics from the model), which is an indication of their importance in the future spread of OPM.

Conclusions Patches identified as 'important' are potential 'pinch points' and regions of high 'flow', where resources for detection and management will be most cost-effectively deployed. However, data on OPM dispersal and the distribution of oak trees limited the strength of our conclusions, so should be priorities for further data collection. This application of electric network theory can be used to inform landscape-scale conservation initiatives both to reduce the spread of invasives and to facilitate large-scale species' range shifts in response to climate change.

D. J. Cowley (✉)
Bristol Centre for Complexity Sciences, University of Bristol, Queen's Building, University Walk, Bristol BS8 1TR, UK
e-mail: daniel.cowley@bristol.ac.uk

D. J. Cowley · M. J. O. Pocock
School of Biological Sciences, University of Bristol, Bristol Life Sciences Building, 24 Tyndall Avenue, Bristol BS8 1TQ, UK
e-mail: michael.pocock@ceh.ac.uk

O. Johnson
School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK
e-mail: o.johnson@bristol.ac.uk

M. J. O. Pocock
Centre for Ecology and Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK

Introduction

The study of networks, i.e. systems of interacting components, is a rapidly increasing field of research with wide applications across life, physical and social sciences (Newman 2003). Within biology, networks have been used to model many different types of interactions from genes within cells, to species in ecosystems and habitat patches in landscapes (Bascompte 2007). The term 'ecological network' has been used to refer to the interactions between species (e.g. in food webs), networks of suitable habitat patches or protected areas (Lawton et al. 2010; Galpern et al. 2011), and the combination of the two: meta-networks of interacting species in fragmented landscapes (Massol et al. 2011). Here, we consider the spatial network of habitat patches linked by species dispersal.

Spatial networks representing landscape structure have been widely studied to evaluate the effect of habitat loss and fragmentation (Fortuna et al. 2006). This has resulted in many new techniques for modelling ecological processes on spatial networks, in particular, measuring the connectivity of the landscape. Improving connectivity across large regions is widely seen as one of the most important strategies for mitigating the effects of climate change and maintaining biodiversity (Heller and Zavaleta 2009; Lawler et al. 2013). A well-connected landscape facilitates meta-population dynamics (Hanski 1999), improves gene flow (Garroway et al. 2008) and facilitates range shifts (Nuñez et al. 2013). However, well-connected landscapes can also accelerate the establishment and spread of invasive species (Kinezaki et al. 2010), which are one of the major threats to biodiversity (Mack and Simberloff 2000; Mooney and Cleland 2001) and result in substantial economic costs (Pimentel et al. 2005).

In this study, we specifically consider the spread of the oak processionary moth (OPM), *Thaumetopoea processionea* (Lepidoptera: Notodontidae), across urban green infrastructure in the United Kingdom. OPM is a univoltine species found in many countries across Europe (Groenen and Meurisse 2012) and is a major defoliator of oak trees (Wagenhoff and Veit 2011). The larvae live communally in nests and feed almost exclusively on the leaves of oak trees (*Quercus* spp.) (Forest Research 2014). An additional concern with OPM is that its larvae have setae (hairs) containing an urticating toxin (thaumetopoein) (Maier et al. 2003).

These hairs can be dispersed up to 500 m (Fenk et al. 2007) and can be harmful to humans and domestic animals, causing cutaneous reactions and potentially serious respiratory symptoms (Maier et al. 2003).

OPM was accidentally introduced to two parts of west London in 2006 on young oak trees imported from continental Europe, from which it has become established and spread (Townsend 2008, 2013). The spread from these areas is concerning, especially if it reaches semi-natural woodlands outside London, because it will then be impossible to control. The green infrastructure of woods, parks and trees in London could therefore contribute to the spread of the moth, but as yet, there is no accepted model to predict the likely spread of OPM. It is especially important to identify how the green infrastructure contributes to the spread of OPM.

In this study, we developed a patch-based method derived from electric network theory to assess species spread across highly fragmented landscapes; and applied the method to model changes in the distribution of OPM occupied woodland patches in west London. Given the patchy distribution of oak trees, there are likely to be particular 'pinch points' and stepping stones between well-connected clusters of patches that could act as foci for management interventions to restrict the further spread of the moth. We assessed how well our models match the observed spread of OPM and used the models to identify these potential 'pinch points'. This has wide relevance beyond OPM in the analysis of spatial ecological networks, particularly for informing conservation policy (to facilitate the movement of native species in response to climate change) and management (to inhibit the spread of invasive species).

Electric network theory and random walks

Electric network theory (or circuit theory) explains the flow of electricity through a network of resistors when a voltage or potential difference is placed across the network (i.e. by connecting a battery and earthing or grounding part of the circuit). Here we use the theory as an analogue of movement across a spatial network of patches. The system of equations representing the electric network model is mathematically identical to that of a random walk on the same network moving from a source until it reaches a sink. In the random walk model the resistors are 'links' between patches,

with their conductances related to the probability of the random walker taking that particular path (Doyle and Snell 1984).

The connection between random walks and electric network theory was first established by Nash-Williams (1959). This relationship has been widely discussed in the mathematical literature (Doyle and Snell 1984; Tetali 1991; Klein and Randić 1993; Chandra et al. 1996; Palacios 2001; Volchenkov 2011). Electrical network theory has been used in a variety of ecological contexts: to understand genetic isolation in heterogeneous landscapes (McRae 2006; McRae and Beier 2007; Spear et al. 2010); to model the resistance of the landscape to species movement (McRae et al. 2008; Lawler et al. 2013); and to calculate the speed of range shifts in fragmented landscapes (Hodgson et al. 2012).

Using the analogy between electric networks and random walks, the flow of a species across the landscape can be thought of as either a random walk across the network of habitat patches or, equivalently, the distribution of current from a battery across an electrical network of resistors. The advantage of this model over other measures of connectivity or landscape resistance (e.g. least-cost pathway analysis, Chardon et al. 2003; Adriaensen et al. 2003) is that the inclusion of all possible pathways across the network is implicit in any network metric (e.g. voltage, current, expected commute time etc. see "Methods" section for details). This modelling framework is equally applicable whether considering movement of an individual (in which case the model represents an average trajectory of an individual, McRae et al. 2008) or spread at a population level, i.e. movement of many individuals over multiple generations (in which case the dispersal kernel models the successful dispersal events and colonisation of patches, Hodgson et al. 2012).

Patch-based and raster-based electric networks

We developed a patch-based approach to model movement across the landscape (see also Hodgson et al. 2012) where the structure of a habitat patch network is characterised by the position and size of habitat patches. Previously, electric network theory has been applied as a raster-based approach, e.g. Circuitscape (Shah and McRae 2008). In Circuitscape, nodes in the electric network are raster grid cells, each connected to their eight adjacent grid cells (McRae

et al. 2008), although other connections could be included, whereas in our method, nodes are habitat patches and are connected to all other patches. Circuitscape requires the resistance of every grid cell to be parameterised. This can provide a realistic representation of the whole landscape (including varying quality of matrix in between habitat patches), but parameterisation can be difficult and require estimation of unknown information. In a raster representation, movement from one cell is restricted to adjacent cells so dispersal events that cover multiple cells are dependent on a sequence of steps across the raster grid, constraining the form of the dispersal kernel. In contrast, in our patch-based approach, the landscape is categorised as either suitable or unsuitable. Therefore less information is required about the landscape and it is relatively straightforward to model different dispersal characteristics by using different dispersal kernels (including fat-tailed and multi-modal dispersal kernels). These can be altered to implicitly incorporate variation in the matrix habitat quality, where this is quantified. For large landscapes, a patch-based approach reduces the size of the data set, which means less computing power is required to analyse the model. Finally, the square raster grids used in most land cover maps are not rotationally invariant (Dunn 2009) whereas a patch-based approach avoids issues with rotational invariance (Etherington 2012). We believe a patch-based approach provides a valuable alternative to raster-based methods like Circuitscape, particularly for habitat that is highly fragmented or when considering the movement of habitat specialists, when a binary classification of the landscape is reasonable. Here we consider a case study meeting these criteria, where OPM (an oak specialist) is spreading in urban areas where oak trees occur in highly fragmented patches.

Methods

Distribution of oak processionary moth and woodland patches

Surveys for OPM larval nests have been undertaken in west London each year from 2006 to 2012 (N. Straw, Forest Research and A. Hoppit, Forestry Commision, unpubl. data used with permission, Fig. 1c). These surveys were undertaken by multiple contractors with
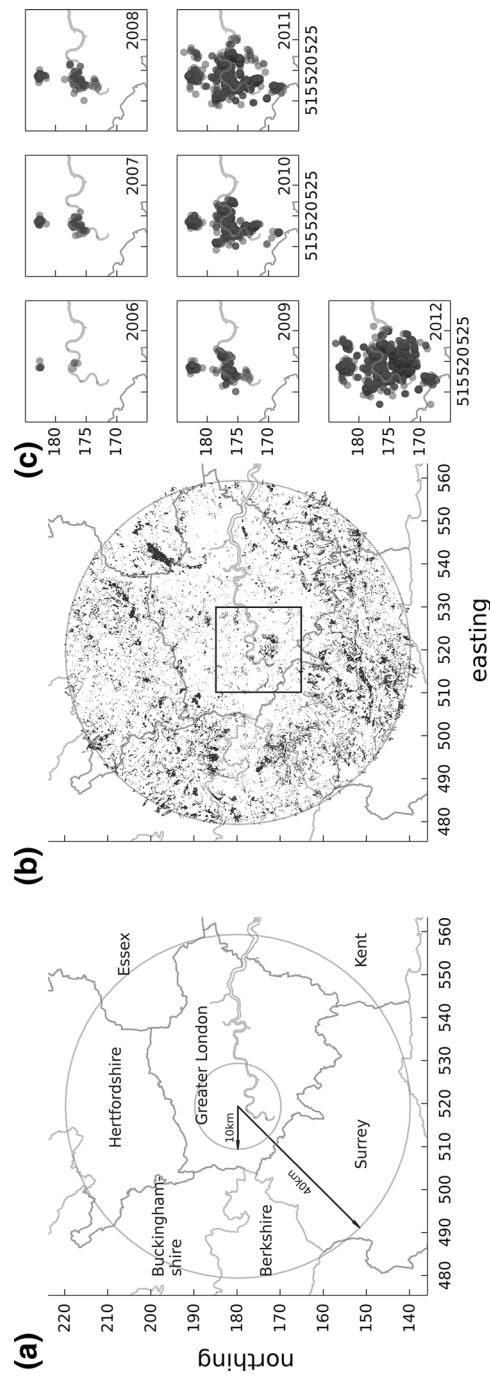
Fig. 1 Maps showing **a** the extent of the region 40 km from the initial establishment of oak processionary moth (*Thaumetopoea processionea*: OPM), **b** the deciduous woodland patches in the region and **c** locations of OPM records for each year of 2006–2012, single points are translucent so *darker grey* indicates a higher density of points. The deciduous woodland habitat patches are © Natural England copyright. Contains Ordnance Survey data © Crown copyright and database right, 2013). In **b** the *box* indicates the location of the zoomed in *square plots* in (c)

the aim of providing a complete census of the population (in order to direct management interventions). Resources were limited so, in practice, once OPM became established, surveys were focused at the edges of the known distribution so that population spread could be determined (the surveys are discussed by Townsend 2013). Therefore, although occupancy in the core of the range was probably under-estimated, we expected that population spread was accurately assessed. Follow-up sampling was not routinely conducted, so we could not assess the impact of management interventions. Given that OPM was continually spreading during our period of study (Fig. 1c), we made the assumption that once detected in a location, OPM remained present throughout the period.

The main host of OPM, i.e. oak trees (*Quercus* spp.), forms an important component of the green infrastructure in London, being a constituent of woodlands, parks and linear green features (railway embankments and road verges). However, no directory of oak trees exists for this region so, for our model, we used an inventory of woodlands from the Natural England Habitat Inventory (Natural England 2013). We selected all polygons where the main habitat type is described as "deciduous woodland" within a circle with a 40 km radius centred on the locations of first recorded instances of OPM. This circle extended to semi-natural woodland around the edge of London. This dataset therefore did not include single trees standing in other habitat types e.g. grassland, private gardens, cemeteries or parks.

Assessing the suitability of the habitat data for modelling the distribution of OPM

It was important to assess the degree to which the inventoried woodland patches captured the observed distribution of oak trees occupied by OPM. Presence of OPM on oak trees outside inventoried woodland patches would indicate that our woodland patch dataset did not completely capture suitable breeding habitat for this species.

We compared the spatial extent of the observed point distribution of OPM and woodland patches positive for OPM records up to 2012 (inclusive) by calculating the overlap between the two distributions. There are several methods used in home range estimation which would be suitable to describe the
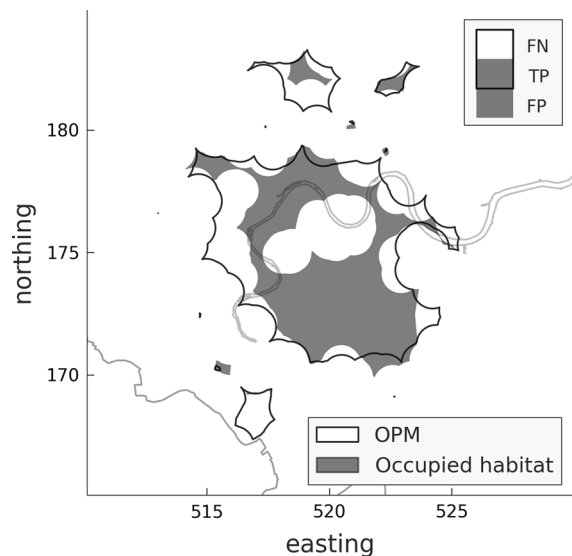
Fig. 2 The distribution of all oak processionary moth (OPM) records up to 2012 (*white with black surround*) and woodland patches with OPM records up to 2012 (*grey*) as described by negative α-hulls (α = 800 m). The binary classification categories are shown where FN is false positive, TP is true positive and FP is false positive

extent of each point distribution; the minimum convex polygon is one of the most straightforward and commonly-used (Powell 2000). However, we used negative α-hulls because they avoid some of the biases introduced by outlying points, which are associated with minimum convex polygons (Burgman and Fox 2003), while remaining simple to implement. Negative α-hulls are constructed by surrounding the distribution of points to be approximated, by all the possible circles of radius α that do not overlap any points in the distribution. The negative α-hull is the remaining area not covered by any circles (Edelsbrunner et al. 1983). The shape of the hull depends on α and because there is no method to select a 'correct' value of α we manually selected a value that adequately described the irregular outline of the distribution as simply as possible.

We wanted to assess whether woodland patches were adequate to consider the distribution of OPM (the alternative being that there were many oaks hosting OPM outside the woodland patches, or many apparently suitable woodland patches without OPM). We assessed this with 'precision' and 'sensitivity' by comparing the α-hull of the known distribution of OPM against the α-hull of woodland patches with

OPM records (Fig. 2). True positives (TP) are given by the area of intersection of the α-hulls, false positives (FP) by the area covered by the woodland patch α-hull but not the OPM record α-hull and false negatives (FN) by the area covered by the OPM record α-hull and not the woodland patch α-hull (Fig. 2). We could not use measures that include true negatives since their number depend on the extent of the landscape included in the analysis, and by increasing the size of the landscape more true negative patches far away from the outbreak are added, hence inflating assessments of fit. Precision was calculated as the proportion of the occupied woodland α-hull that intersected the OPM record α-hull: TP/(TP + FP), whereas sensitivity was calculated as the proportion of the OPM record α-hull that captured the occupied woodland α-hull: TP/(TP + FN).

Modelling framework

There were two stages to the modelling process: firstly we built a network which represents the arrangement of habitat within the landscape and the probability that a species can successfully move from one patch to another, and secondly we modelled the spread of the focal species across the network (in this case OPM spreading out of west London).

Building the network involves dividing the landscape into suitable habitat (the deciduous woodland polygons) and non-suitable habitat (everything else). Suitable habitat patches were the nodes and species dispersal was represented by links. The links were weighted using a dispersal kernel representing the probability that OPM from one patch can colonise another patch. We modelled the probability of a successful colonisation event from patch i to patch j as a function of Euclidean distance, d_{ij}, and area of each patch, A_i and A_j (further details below). We weighted the link between patches i and j by 'conductance' (the inverse of 'resistance') given by

$$C_{ij} = A_i A_j d_{ij}^{-\beta}. \tag{1}$$

Since the OPM is a habitat specialist, feeding and reproducing almost exclusively in oak trees, we assumed that the matrix is uniform and that Euclidean distance is an appropriate distance metric. If this assumption did not hold then it would have been straight-forward to implement a different distance metric e.g. a least-cost pathway (Adriaensen et al. 2003), although this could have been computationally expensive depending on the number of links in the network. We represented the dispersal kernel by a power law function of the form $d^{-\beta}$ where the exponent, β, dictates the rate of decay in dispersal probability. We chose the power law function because it has a 'fat tail' so incorporating a higher probability of rare long distance dispersal events than exponential functions (Nathan 2001). An alternative choice of dispersal kernel is easily applied in the model.

The area of the donor patch (A_i) was included since a larger patch can potentially accommodate a larger population of OPM and therefore it is more likely that a successful dispersal event from that patch will occur. We assumed that the population density is constant in occupied patches, so larger area means more organisms starting from that point. The area of the target (A_j) patch was included since a dispersing individual is more likely to encounter a large patch than a small patch and we assumed this was directly proportional to the area.

We modelled the spread of the species across the network using the analogy between electric network theory and random walks. We were interested in the spread of OPM over many generations so each step in the random walk across the network represents the dispersal event from the emergence of the moth after pupation through to reproduction at a new site. We defined a source area (where the battery is connected or where the random walk starts), specifically this was the location of the records of OPM in 2006 in Ealing and Kew, and a sink area (where the circuit is earthed), specifically the perimeter of a 40 km circle centred on the two sources. This circular arrangement modelled spread from the source to the sink, in all directions. By using undirected edges we imposed no preference on each step in the random walk. These sources and sinks gave a system of equations for the voltage at each node where voltage is the potential difference between the node and the earth, as follows.

Ohm's law gives a relationship between current, voltage and conductance (Ohm 1827)

$$I_{ij} = C_{ij}(v_i - v_j). \tag{2}$$

I_{ij} is the current in the link joining nodes i and j, v_i is the voltage at node i, and C_{ij} is the conductance of the link between nodes i and j, given by Eq. (1).

Table 1 Mathematically equivalent relations between electric network theory and random walks, and the proposed ecological interpretations. Adapted from McRae et al. (2008) with the addition of descriptors of 'power'

Electric network descriptors	Random walk equivalents	Ecological interpretation
Voltage/potential, v_i	Escape probability	'Effective distance' across the network
Edge current, I_{ij}	Expected net number of traverses	Flow of organisms between patches
Node current, I_i	Expected net number of visits	Flow of organisms through a patch
Effective resistance	Commute time	Traversibility of the landscape
Edge power, E_{ij}	N/A	Distance-weighted flow between patches
Node power, E_i	N/A	Distance-weighted flow through a patch

Kirchhoff's law states that (Kirchhoff 1847)

$$\sum_{j \neq i} I_{ij} = 0, \tag{3}$$

so at each node (not the source or the target) the total current is equal to zero, that is incoming or positive current is equal to the outgoing or negative current. We obtain a system of equations for the voltages by applying (2) and (3) at each node. At node i, this gives an equation of the form

$$\left(\sum_{j \neq i} -C_{ij} \right) v_i + \sum_{j \neq i} C_{ij} v_j = 0. \tag{4}$$

We can think of (4) as defining the ith row of a symmetric $\mathcal{N} \times \mathcal{N}$ matrix, M, with off-diagonal entries given by C_{ij} (i.e. the 'conductance' or 'inverse resistance' between patches) and diagonal entries that make the row sums zero. Equation (4) requires us to solve for the voltages, v, with boundary conditions at the source (voltage = 0) and target (voltage = 1) nodes imposed by b

$$Mv = b. \tag{5}$$

This system is equivalent to that for the escape probability of a random walker on the same network (Doyle and Snell 1984) and is the same system of equations considered by Hodgson et al. (2012).

The random walk and electrical network modelling schemes are based on different physical interpretations of the same underlying system and their equivalence allowed us to make analogies with ecological concepts (see Table 1, adapted from McRae et al. 2008). Voltage increases as you move from the source towards the target so it can be thought of as 'effective distance' across the network (i.e.

'escape probability' in a random walk model). In a regular one-dimensional lattice (a string of nodes the same distance apart) the change in voltage is linear with distance. However in more complex networks this is not the case; the structure of the network influences the change in voltage, so voltage represents the 'effective distance' in terms of ease of crossing the network for the organisms being modelled. The electric current gives an indication of how much 'flow' is expected along an edge or through a patch. It is an assessment of how connections and patches facilitate the flow across the landscape. In the random walk model, this represents the expected net number of times the random walk will cross that edge or patch before it reaches the target area. The 'effective resistance' is a whole network metric of the overall permeability or resistance of the landscape (i.e. the commute time, which is the expected time for a random walker to traverse the network) and can be used to compare landscapes or evaluate the effect of adding or removing patches.

We also considered the power, E_{ij}, in an edge (resistor) given by

$$E_{ij} = v_{ij} I_{ij}. \tag{6}$$

This metric has no direct equivalence in random walk models, but we can explain it as 'effective distance' (voltage, v) multiplied by 'flow' (current, I), i.e. distance-weighted flow. This means that long inter-patch links have higher 'power' than shorter ones with the same flow, and that links with high flow have higher 'power' than links of the same distance but with lower flow. Ecologically, 'high power' indicates larger numbers of organisms or more biological material, travelling further. Power is a useful metric because in our unpublished related work, we have shown

mathematically that the power of a node accurately predicts the change in effective resistance of the network caused by the removal of that node, whereas the current of a node does not accurately predict this. Power can therefore be used to assess the importance of patches within a network in terms of the overall resistance of the landscape.

Calibrating the model to predict OPM arrival time

If voltage is a measure of 'effective distance' across the network then the value of 'voltage' can be taken as an indication of the time elapsed before OPM arrives at a location. In order to calibrate the relationship between voltage and 'OPM arrival time', we estimated for each year for which we had OPM records the threshold of v up to which OPM had spread. To do this, we chose for each year a range of v-thresholds up to which OPM has supposedly spread and compared the resulting presence/absence maps to the observed OPM presence/absence map of that year. For each year we selected the threshold which maximised the agreement, using a metric known as the F_1 score (van Rijsbergen 1979), given by = $2TP/(2TP + FP + FN)$, where: true positives (TP) are patches with simulated OPM presence and observed OPM presence; false positives (FP) are patches with simulated OPM presence but observed

OPM absence; and false negatives (FN) are patches with simulated OPM absence but with observed OPM presence. A scatterplot of time (year) and voltage should reveal a relationship that allows linking voltage to OPM arrival time.

Models were run in the Python programming language (Python Software Foundation 2012) and systems of linear equations were solved using the LAPACK package (Anderson et al. 1999).
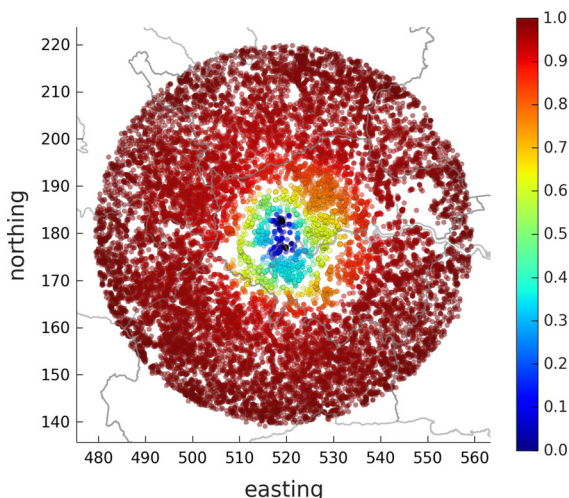
Results

Summary and data validation

There were 4439 records of OPM from west London between 2006 and 2012. In general, these data indicate that the moth is spreading (Fig. 1c). However, the spread is not even and so far appears to be fastest to the south-west (Fig. 1c). The woodland patch dataset comprised 1069 patches within 10 km of the initial sources and 16,869 patches within 40 km of the initial sources.

Only 36 % of woodland patches that intersected with the OPM record α-hull had OPM records, and 51 % of individual OPM records were from outside of woodland patches. However, this considered individual records, which probably under-estimated colonies in woodland, and hence under-estimated overall precision and sensitivity. In order to consider the overall pattern we approximated the recorded distribution of OPM and the distribution of woodland patches positive for OPM records with an α-hulls, with α = 800 m to produce the best approximation to the distributions (Fig. 2). Precision (proportion of occupied woodland patch α-hull that intersects with OPM record α-hull) was 93 % and sensitivity (proportion of OPM record α-hull that intersects with occupied woodland patch α-hull) was 54 %. Sensitivity rose to 70 % when considering the total extent (i.e. filling internal holes in the α-hulls). Overall, these results indicate that the range of OPM in west London can be reasonably well approximated with reference to the woodland patches alone (see "Discussion" section for more details).
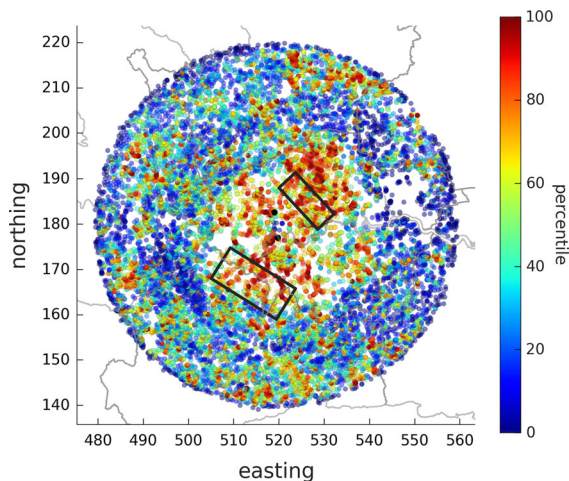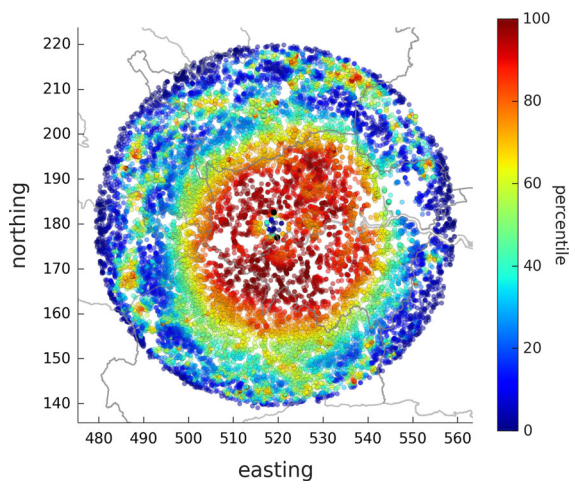

Fig. 3 The voltage of woodland patches (represented as individual points) as calculated from our electric network model. Points are coloured by the voltage ('distance' from the sources) of the patch. *Points* represent individual patches within 40 km of the two initial sources (marked in *black*) and other boundaries are described in Fig. 1

Network metrics

We ran the electric network model on the network of deciduous woodland patches and found that voltage

Fig. 4 The current of woodland patches (represented as individual points) as calculated from our electric network model. *Points* are coloured by the percentile rank of current of the patch. Points represent individual patches within 40 km of the two initial sources (marked in *black*) and other boundaries are described in Fig. 1. The *black boxes* indicate areas where monitoring effect could be focused (see "Discussion" section)

Fig. 5 The power of woodland patches (represented as individual points) as calculated from our electric network model. Points are coloured by the percentile rank of power of the patch. *Points* represent individual patches within 40 km of the two initial sources (marked in *black*) and other boundaries are described in Fig. 1

(i.e. 'effective distance') increased from the source to the edge of the network, but did not do so uniformly due to variations in the structure of the network (Fig. 3). Where voltage declines rapidly with geographic distance, this indicates regions where movement across the landscape is predicted to be hindered

(i.e. there is high resistance causing a slower predicted spread of OPM) due to a lower density of habitat (i.e. small patches or few patches; Fig. 3). Conversely, smaller changes in voltage indicate areas where the spread of OPM is predicted to be facilitated by the arrangement of patches (i.e. large and/or numerous patches).

As voltage can be taken as an indicator of 'OPM arrival time' (section 'Calibration of the model to predict OPM arrival time' presents a quantitative calibration of this relationship), results displayed in Fig. 3 suggest that OPM would initially spread south west (dark blue patches, $v < 0.1$), which was confirmed by data between 2006–2012 (Fig. 1c). The next phase of spread occurs mainly due south and to a lesser extent north-west (pale blue patches, $0.3 < v < 0.4$, Fig. 3), and the following phase of spread occurs largely to the north-east (yellow to orange patches, $0.6 < v < 0.7$, Fig. 3). In the later stages spread appears to be easier to the north east while a region with fewer patches to the south west holds up spread in that direction.

The routes through which OPM was predicted to spread were assessed by the 'current' through the patches, where high current indicates patches with a high flow (irrespective of the patch's voltage). Movement across the whole network was concentrated in a small number of high current patches and these were often arranged in pathways through the network indicating important corridors concentrating the potential spread of OPM (Fig. 4).

Considering the 'power' of individual patches (the 'effective distance'-weighted flow of species), the distribution of patches with high power (Fig. 5) is very different to the distribution of patches with high current (Fig. 4). Weighting flow by the 'effective distance' travelled highlights patches which are not just important in terms of flow per se but also with respect to the progress that flow makes across the landscape. Patches with high power are stepping stones, linking groups of well-connected patches together where you have high flow over larger distances.

Calibration of the model to predict OPM arrival time

We assessed the agreement of the observed OPM presence/absence data for each year with the simulated OPM presence/absence (patches with voltage $v < V$).
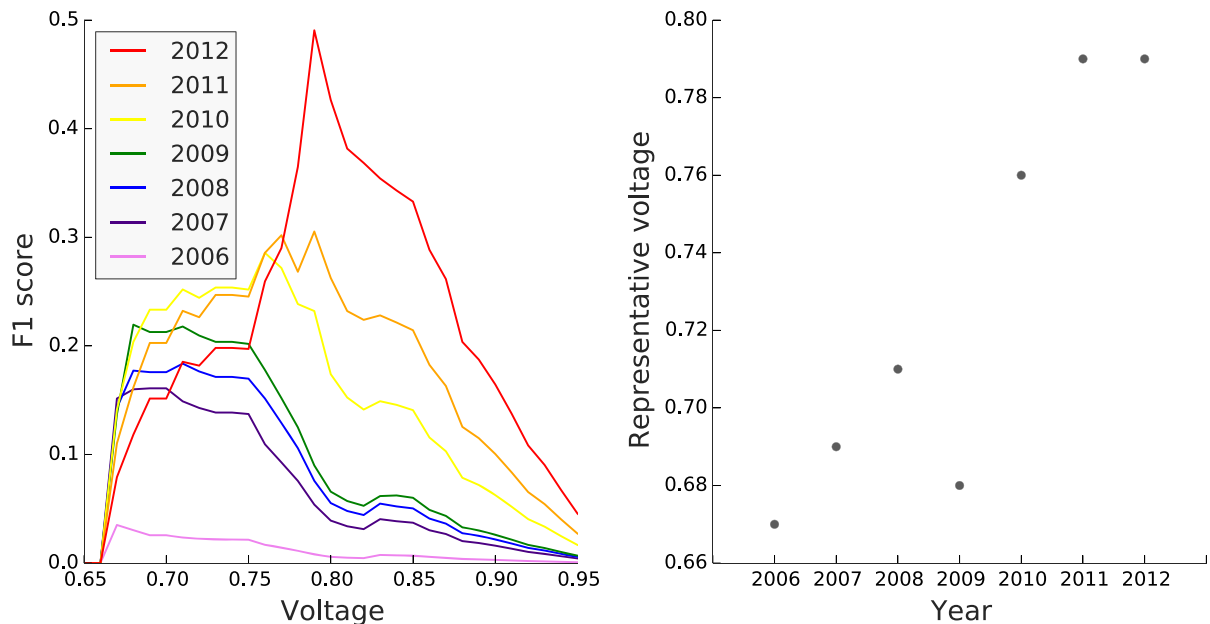
Fig. 6 a The assessment, using F_1 score (the mean of precision and sensitivity), of how well patches with a selected value of voltage (V; 'effective distance' from source) fit the distribution of oak processionary moth (OPM) each year and b the value of voltage that best fits the recorded distribution of OPM for each year, i.e. at which the F_1 score is maximised

The level of agreement (maximum F_1 score) increased for each year (Fig. 6a).

For each year, the voltage at the maximum F_1 score represented the predicted OPM presence. This voltage generally increased annually, confirming that increases in voltage approximated to arrival time, although this relationship was not linear (Fig. 6b), meaning that it is difficult to extrapolate and predict precise arrival time of OPM based on the value of patch voltage.

Discussion

We have presented a model based upon electric network theory to analyse the spread of OPM in south west London. Although the data had limitations, it served as a valuable empirical test of this model. Validating these models is particularly important when they are intended to guide landscape-scale development with expected impacts decades in the future because test data is, obviously, difficult or impossible to obtain. Here, we consider the implications (and limitations) of our case study, before considering the broader use of this electric network modelling approach and how it can be further developed.

Modelling the spread of OPM

In our case study of the spread of an insect pest (Oak Processionary Moth: OPM) in fragmented green infrastructure in urban areas, we were able to identify patches that are likely to be key stepping stones or pinch points (i.e. have high 'current' or 'power'). These areas should be the focus of management to reduce the likelihood of OPM spreading further. We showed that increasing 'effective distance' (voltage) from the points of introduction (source) was related to time of arrival, but this relationship is non-linear so it was not possible to extrapolate specific values of 'effective distance' (voltage) to specific predicted arrival times (although see Hodgson et al. 2012).

Given the limitations of the data (see below), we are cautious in making strong recommendations about the management of OPM in this region. However, our results predict initial rapid spread to the south west (as confirmed by the data), but then there is a sparsely wooded area that is relatively difficult for OPM to cross. Our results confirm that a key concern is preventing OPM from crossing this region and reaching the high density of larger woodland patches in Surrey (south and west of the existing distribution) because once there, there will be little resistance to

further spread (Fig. 3). The model also reveals that there is a high potential for OPM to spread north-east through London to reach the well-wooded area to the north-east. So, based on this model, we would recommend that surveillance and management effort is especially focused on the two areas in which there are patches with high current (Fig. 4), because flow is predicted to be concentrated through these patches and so they are of high importance in the further spread of OPM. These areas comprise (1) the patches acting as stepping stones to the Surrey woodlands to the south-west and (2) the leading edge of the block of small, but numerous and well-connected, woodlands to the north east.

Although the database of woodland patches serves as a reasonable approximation for the extent of the OPM distribution, we found that it did not well approximate the fine structure within the currently recorded OPM distribution (i.e. only half of records up to 2012, inclusive, are within woodland patches). There were four particular limitations of the habitat data, which may have impacted our work. Firstly, the monitoring of OPM was established reactively for the rapid assessment of its spread rather than to produce an accurate population distribution map, so survey effort was focused at the edge of its range. As a result, the distribution of OPM in the core of its range was less well-recorded, thus reducing the observed 'precision' of our model (Fig. 6) by inflating the number of false positives i.e. patches in the core range of OPM predicted to contain OPM but where none had been recorded. Secondly, we expected that it was easier to detect OPM on single standing trees than to detect OPM on trees within woodland because observation of the whole crown is easier from the ground level when the view is not obscured. Also some surveyors recorded OPM per 'stand' of trees rather than counting individual occupied trees. These recorder biases will have under-estimated how well the total OPM distribution can be approximated from the woodland data set. Thirdly, recording was undertaken in order to conduct local eradication procedures (although the location of eradication attempts was not known to us). Therefore, if eradication was successful, then the recorded OPM larval nest would not have been a source of dispersers. This bias would be important in mechanistic models of the spread of OPM to predict time of arrival at a location (e.g. Hodgson et al. 2012). However, detection and eradication of OPM is

difficult in practice, so it seems quite likely that OPM may have spread from locations near the eradication attempts, as confirmed by the reasonable fit of our models to the data (Fig. 4). Finally, there is no database of oak trees in our area of interest. We used a dataset of woodland patches and it is likely that most, if not all, would have contained oak trees. However, we have no knowledge of the locations of individual oak trees away from deciduous woodland e.g. in gardens, parks, cemeteries or golf courses, which would facilitate dispersal across the landscape by providing stepping-stones for population spread. This would be especially influential if the distribution of lone oaks is regionally auto-correlated.

Ultimately, better maps of green infrastructure (in this case, the distribution of oak trees) would assist in accurate modelling and contribute to effective conservation decision making. This would complement the development of urban tree maps for assessing ecosystem service provision, e.g. citizen science approaches (e.g. Opentreemap: https://www.opentreemap.org/ and Treezilla: http://www.treezilla.org/) or remote sensing by combining methods to delineate individual trees (Wulder et al. 2004; Chang et al. 2013) with those to remotely identify tree species (Xiao et al. 2002; Carleer and Wolff 2004).

The electric network model in ecology

We found that species spread is dramatically influenced by landscape characteristics. We used electric network theory to model species spread across a fragmented habitat network. This approach models complex fragmented habitat in a relatively straightforward way and includes the effects of multiple pathways across the landscape. It provides a suite of metrics for both individual patches and whole landscapes, which have useful and intuitive ecological interpretations. This model can be used more generally in the analysis of fragmented habitat networks in light of changing global conditions to inform conservation management in landscape-scale conservation initiatives.

Our method for identifying the most important patches is particularly useful when conservation resources are limited and effective action is urgently required. This applies when seeking to impede the spread of invasive species (as we have demonstrated) but also when facilitating the movement of native species in the face of anthropogenic climate change,

e.g. from the distribution of currently occupied sites to the distribution of sites predicted to be suitable at some time in the future (Parmesan 2006; Lawler et al. 2013). Indeed, these two competing priorities may need to be addressed simultaneously in any one landscape. It is a big challenge to balance the efforts to protect native species against the requirement to restrict local dispersal of invasive species. One approach would be to model target species (e.g. from current range to predicted future range for native species, and from introduction point to vulnerable habitat for invasive species). Then monitoring and management efforts could be focused on the 'pinch points' identified from the models (i.e. patches with particularly high current) in order to seek to facilitate or impede (as appropriate) the species under consideration.

Our approach could be developed in order to improve its application. One of the key assumptions underlying our simple application of this model is that the matrix is homogeneous but in reality the traversability of the matrix will vary, often in a way that is spatially auto-correlated. This could be incorporated into the model by varying the dispersal kernel per inter-patch link, by weighting each link by the relative traversability of the matrix or using the least cost pathway distance (Watts et al. 2010). Least cost pathways are likely to become prohibitively time-consuming as the number of patches increases because the number of links scales by the square of the number of patches (in our woodland network with 1.6×10^3 patches there were 1.28×10^8 links), but with the increasing computational power of computers, this may be feasible in the future.

Using 'current' as a metric to identify the patches with high flow could be applied to other types of weighted networks as an alternative to more commonly used metrics of node importance, for example centrality metrics (Barthélemy 2011), and need not be restricted to spatial networks (Barrat et al. 2004). The model applies to flow across any type of network, e.g. flow of information in social networks or flow of resources in food webs, especially where there are defined 'sources' and 'sinks'.

It would also be informative to compare our modelling method with other similar models of species movement, for example Circuitscape (McRae et al. 2008) or Rangeshifter (Bocedi et al. 2014). Each of these different models makes different assumptions about the dispersing organism. They differ in terms of their flexibility, simplicity, usability and computational

efficiency but qualitative comparisons between the results are lacking. Although a simulation comparison of these models would be instructive, it would be far more informative to test the real world applicability of these models with high quality empirical data of the population spread for a species across a fragmented landscape.

The networks considered here are all undirected, which means the dispersal probabilities between pairs of patches are symmetric, but for case studies into the dispersal of other species these probabilities could be anisotropic, e.g. seeds dispersed by wind (Howe and Smallwood 1982; Nathan 2001) or marine larvae dispersed by prevailing water currents (Cowen and Sponaugle 2009). Anisotropic dispersal could be represented by a directed network with two directed links between each pair of patches. Then the modelling framework (Eqs. 1–4) would need to be altered so that 'dispersal in' equals 'dispersal out' for each patch, preserving the conservation of current in Ohm's law (Eq. 2). The matrix (M; Eq. 5) derived from the inter-patch conductance matrix would then not be symmetric, reflecting the asymmetry of the dispersal probability.

In conclusion, this study acts as a case study demonstrating the value of the electric network model. Models have an important role in developing our understanding of how habitat networks affect species movement, especially in a rapidly changing world. However, obtaining accurate data with which to validate and parameterise these models will be vital to ensure conservation policies can be informed by sound theoretical work.

References

Adriaensen F, Chardon JP, De Blust G, Swinnen E, Villalba S, Gulinck H, Matthysen E (2003) The application of "least-cost" modelling as a functional landscape model. Landsc Urban Plan 64:233–247

Anderson E, Bai Z, Bischof C, Blackford S, Demmel J, Dongarra J, Du Croz J, Greenbaum A, Hammarling S, McKenney A, Sorensen D (1999) LAPACK users' guide, 3rd edn. Society for Industrial and Applied Mathematics, Philadelphia

Barrat A, Barthelemy M, Pastor-Satorras R, Vespignani A (2004) The architecture of complex weighted networks. Proc Natl Acad Sci USA 101:3747–3752

Barthélemy M (2011) Spatial networks. Phys Rep 499:1–101

Bascompte J (2007) Networks in ecology. Basic Appl Ecol 8:485–490

Bocedi G, Palmer SCF, Pe'er G, Heikkinen RK, Matsinos YG, Watts K, Travis JMJ (2014) RangeShifter: a platform for modelling spatial eco-evolutionary dynamics and species' responses to environmental changes (R. Freckleton, ed). Methods Ecol Evol 5:388–396

Burgman MA, Fox JC (2003) Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. Anim Conserv 6:19–28

Carleer A, Wolff E (2004) Exploitation of very high resolution satellite data for tree species identification. Photogramm Eng Remote Sens 70:135–140

Chandra AK, Raghavan P, Ruzzo WL, Smolensky R, Tiwari P (1996) The electrical resistance of a graph captures its commute and cover times. Comput Complex 6:312–340

Chang A, Eo Y, Kim Y, Kim Y (2013) Identification of individual tree crowns from LiDAR data using a circle fitting algorithm with local maxima and minima filtering. Remote Sens Lett 4:29–37

Chardon J, Adriaensen F, Matthysen E (2003) Incorporating landscape elements into a connectivity measure: a case study for the Speckled wood butterfly (Pararge aegeria L.). Landscape Ecol 18:561–573

Cowen RK, Sponaugle S (2009) Larval dispersal and marine population connectivity. Ann Rev Mar Sci 1:443–466

Doyle PG, Snell JL (1984) Random walks and electric networks. Mathematical Association of America, Washington, DC

Dunn AG (2009) Grid-induced biases in connectivity metric implementations that use regular grids. Ecography 33:627–631

Edelsbrunner H, Kirkpatrick D, Seidel R (1983) On the shape of a set of points in the plane. IEEE Trans Inf Theory 29:551–559

Etherington TR (2012) Least-cost modelling on irregular landscape graphs. Landscape Ecol 27:957–968

Fenk L, Vogel B, Horvath H (2007) Dispersion of the bioaerosol produced by the oak processionary moth. Aerobiologia 23:79–87

Forest Research. (2014). Oak processionary moth—Thaumetopoea processionea. Tree Health. http://www.forestry.gov.uk/fr/INFD-6URJCF. Accessed 10 Jan 2014

Fortuna MA, Gómez-Rodríguez C, Bascompte J (2006) Spatial network structure and amphibian persistence in stochastic environments. Proc Biol Sci R Soc 273:1429–1434

Galpern P, Manseau M, Fall A (2011) Patch-based graphs of landscape connectivity: a guide to construction, analysis and application for conservation. Biol Conserv 144:44–55

Garroway CJ, Bowman J, Carr D, Wilson PJ (2008) Applications of graph theory to landscape genetics. Evol Appl 1:620–630

Groenen F, Meurisse N (2012) Historical distribution of the oak processionary moth Thaumetopoea processionea in Europe suggests recolonization instead of expansion. Agric For Entomol 14:147–155

Hanski I (1999) Habitat connectivity, habitat continuity, and metapopulations in dynamic landscapes. Oikos 87:209

Heller NE, Zavaleta ES (2009) Biodiversity management in the face of climate change: a review of 22 years of recommendations. Biol Conserv 142:14–32

Hodgson JA, Thomas CD, Dytham C, Travis JMJ, Cornell SJ (2012) The speed of range shifts in fragmented landscapes. PLoS One 7:e47141

Howe HF, Smallwood J (1982) Ecology of seed dispersal. Annu Rev Ecol Syst 13:201–228

Kinezaki N, Kawasaki K, Shigesada N (2010) The effect of the spatial configuration of habitat fragmentation on invasive spread. Theor Popul Biol 78:298–308

Kirchhoff G (1847) Ueber die Auflösung der Gleichungen, auf welche man bei der Untersuchung der linearen Vertheilung galvanischer Ströme geführt wird. Ann Phys 148:497–508

Klein DJ, Randić M (1993) Resistance distance. J Math Chem 12:81–95

Lawler JJ, Ruesch AS, Olden JD, McRae BH (2013) Projected climate-driven faunal movement routes. Ecol Lett 16:1014–1022

Lawton JH, Brotherton PNM, Brown VK, Elphick C, Fitter AH, Forshaw J, Haddow RW, Hilborne S, Leafe RN, Mace GM, Southgate MP, Sutherland WJ, Tew TE, Varley J, Wynne GR (2010) Making space for nature: a review of England's wildlife sites and ecological network. Report to DEFRA

Mack R, Simberloff D (2000) Biotic invasions: causes, epidemiology, global consequenses, and control. Ecol Appl 10:689–710

Maier H, Spiegel W, Kinaciyan T, Kerhan H, Cabaj A, Schopf A, Honigsmann H (2003) The oak processionary caterpillar as the cause of an epidemic airborne disease : survey and analysis. Br J Dermatol 149:990–997

Massol F, Gravel D, Mouquet N, Cadotte MW, Fukami T, Leibold Ma (2011) Linking community and ecosystem dynamics through spatial ecology. Ecol Lett 14:313–323

McRae BH (2006) Isolation by resistance. Evolution 60:1551–1561

McRae BH, Beier P (2007) Circuit theory predicts gene flow in plant and animal populations. Proc Natl Acad Sci USA 104:19885–19890

McRae BH, Dickson BG, Keitt TH, Shah VB (2008) Using circuit theory to model connectivity in ecology, evolution, and conservation. Ecology 89:2712–2724

Mooney HA, Cleland EE (2001) The evolutionary impact of invasive species. Proc Natl Acad Sci USA 98:5446–5451

Nash-Williams CSJA (1959) Random walk and electric currents in networks. Proc Camb Philos Soc 55:181–194

Nathan R (2001) The challenges of studying dispersal. Trends Ecol Evol 16:481–483

Natural England (2013) GIS digital boundary datasets. http://www.naturalengland.org.uk/. Accessed 15 Nov 2013

Newman MEJ (2003) The structure and function of complex networks. SIAM Rev 15:247–262

Nuñez TA, Lawler JJ, McRae BH, Pierce DJ, Krosby MB, Kavanagh DM, Singleton PH, Tewksbury JJ (2013)

Connectivity planning to address climate change. Conserv Biol 27:407–416

Ohm GS (1827) Die Galvanische Kette, mathematisch bearbeitet. (tr., The Galvanic Circuit Investigated Mathematically). Riemann, Berlin

Palacios JL (2001) Resistance distance in graphs and random walks. Int J Quantum Chem 81:29–33

Parmesan C (2006) Ecological and evolutionary responses to recent climate change. Annu Rev Ecol Evol Syst 37:637–669

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecol Econ 52:273–288

Powell RA (2000) Animal home ranges and territories and home range estimators. In: Boitani L, Fuller TK (eds) Research techniques in animal ecology: controversies and consequences. Columbia University Press, New York, pp 65–110

Python Software Foundation (2012) Python Language Reference, version 2.7.3

Shah V, McRae B (2008) Circuitscape: a tool for landscape ecology. In: Proceedings of the 7th Python in science conference, vol 7, pp 62–68

Spear SF, Balkenhol N, Fortin M-J, McRae BH, Scribner K (2010) Use of resistance surfaces for landscape genetic studies: considerations for parameterization and analysis. Mol Ecol 19:3576–3591

Tetali P (1991) Random walks and the effective resistance of networks. J Theor Probab 4:101–109

Townsend M (2008) Oak processionary moth *Thaumetopoea processionea* (L.) (Le.:Thaumetopoeidae) in west London. Entomol Rec J Var 120:149–150

Townsend M (2013) Oak processionary moth in the United Kingdom. Outlooks Pest Manag 24:32–38

Van Rijsbergen CJ (1979) Information retrieval. Butterworths, London

Volchenkov D (2011) Random walks and flights over connected graphs and complex networks. Commun Nonlinear Sci Numer Simul 16:21–55

Wagenhoff E, Veit H (2011) Five years of continuous *Thaumetopoea processionea* Monitoring: tracing population dynamics in an arable landscape of south-western Germany. Gesunde Pflanzen 63:51–61

Watts K, Eycott AE, Handley P, Ray D, Humphrey JW, Quine CP (2010) Targeting and evaluating biodiversity conservation action within fragmented landscapes: an approach based on generic focal species and least-cost networks. Landscape Ecol 25:1305–1318

Wulder MA, White JC, Niemann KO, Nelson T (2004) Comparison of airborne and satellite high spatial resolution data for the identification of individual trees with local maxima filtering. Int J Remote Sens 25:2225–2232

Xiao X, Boles S, Liu J, Zhuang D, Liu M (2002) Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. Remote Sens Environ 82:335–348