Consequences of ignoring dispersal variation in network models for landscape connectivity

Lauren L. Sullivan[1,2], Matthew J Michalska-Smith[3,4], Katie P. Sperry[2,5], David A. Moeller[6], and Allison K. Shaw[2]

[1]Division of Biological Sciences, University of Missouri, Columbia, MO, U.S.A.
[2]Department of Ecology, Evolution, and Behavior, University of Minnesota, Saint Paul, MN, U.S.A.
[3]Department of Veterinary Population Medicine, University of Minnesota, Saint Paul, MN, U.S.A.
[4]Department of Plant Pathology, University of Minnesota, Saint Paul, MN, U.S.A.
[5]Marine and Environmental Sciences, Northeastern University, Boston, MA, U.S.A.
[6]Department of Plant and Microbial Biology, University of Minnesota, Saint Paul, MN, U.S.A.

1 **Abstract**

2 Habitat loss and fragmentation can negatively impact population persistence and biodiversity,

3 but these effects can be mitigated if species successfully disperse between isolated habitat

4 patches. Network models are the primary tool for quantifying landscape connectivity, yet as

5 practiced, they take an overly simplistic view of species dispersal. These models often ignore

6 individual variation in dispersal ability, assuming all individuals move the same fixed distance

7 with equal probability. Here we develop a modeling approach to address this problem. We

8 incorporate dispersal kernels into network models to determine how individual variation in

9 dispersal alters our understanding of landscape-level connectivity, and test our approach on a

10 fragmented grassland landscape in Minnesota. We show that ignoring dispersal variation

11 consistently overestimates a population's robustness to local extinctions while simultaneously

12 underestimating its robustness to local habitat loss. Furthermore, a simplified view of dispersal

13 underestimates the amount of habitat sub-structure for small populations but overestimates

14 habitat sub-structure for large populations. Our results demonstrate that considering biologically

15 realistic dispersal alters our understanding of landscape connectivity for ecological theory and

16 conservation practice.

17

18 **Introduction**

19   Loss of habitat due to land-use alteration is one of the largest anthropogenic threats to

20 Earth's planetary systems (Rockström 2009), contributing to major declines in biodiversity

21 (Newbold et al. 2016) and other ecosystem services (Haddad et al. 2015). Habitat loss

22 fundamentally alters landscapes by simultaneously decreasing the overall amount of native

23 habitat, and changing how the remaining habitat patches are arranged with respect to each other

24   through fragmentation *per se* (Fahrig 2017). The negative effects of fragmentation (Fletcher et

25   al. 2018) can be mitigated if species are still able to move between physically isolated habitat

26   patches, maintaining connectivity. However, the extent of species' movement among patches

27   remains an open question (Fahrig 2017). Thus, a complete understanding of the degree to which

28   current (Haddad et al. 2015) and ongoing (Wright & Wimberly 2013) fragmentation disrupts

29   connectivity requires accounting for potential species movement among patches.

30   Recent work examining how organisms move among fragmented patches draws on

31   network modelling (e.g. Saura & Rubio 2010; Fletcher et al. 2013; Ziółkowska et al. 2014;

32   Wimberly et al. 2018). This approach converts spatial data on habitat locations to networks (or

33   graphs), where nodes represent habitat patches and two patches are connected by an edge if

34   organisms can disperse between them (Urban & Keitt 2001). These networks can be analyzed to

35   inform conservation decisions by calculating patch-based or network-based 'connectivity'

36   metrics, and identifying sets of patches that are connected via dispersal and thus function as a

37   unit (termed components). Network modelling has enabled researchers and managers to infer

38   connectivity, identify habitat patches with high conservation value, and quantify the scale of

39   dispersal necessary to maintain connectivity (e.g. O'Brien et al. 2006; Saura & Rubio 2010;

40   Creech et al. 2014; Wimberly et al. 2018). Generally, while models that infer connectivity tend

41   to take into account detailed information about matrix quality between habitat patches and

42   potential ease-of-flow through this matrix (Moilanen & Hanski 1998; McRae et al. 2008;

43   Wimberly et al. 2018), they often make simplistic assumptions about species' movement

44   dynamics. Specifically, network models, particularly unweighted or binary networks that simply

45   consider whether or not patches are connected, tend to define species' dispersal as a single fixed

46   distance, effectively assuming that all individuals are equally able to disperse up to that distance,

47   and unable to disperse past that distance. Building network models with the same sets of

48   simplifying assumptions limits our ability to understand how a broad range of biological factors

49   (such as dispersal behavior) influence connectivity, which is especially problematic since

50   dispersal variation can have many consequences (Snell et al. 2019; Shaw 2020). Simplified

51   dispersal assumptions may over- or under-estimate the degree of connectivity, or fail to capture

52   important connectivity patterns altogether, thereby preventing accurate estimations of landscape-

53   level habitat use.

54   An alternative to viewing dispersal as fixed, is to account for variation in dispersal, thus

55   more accurately representing movement behavior. Inherent variation among individuals (e.g. sex,

56   personality, body condition), populations (e.g. density), and the environment (e.g. habitat quality,

57   habitat configuration, and season) can cause differences in dispersal ability (Snell et al. 2019;

58   Shaw 2020). This dispersal variation can be captured with a dispersal kernel, that describes the

59   proportion of individuals traveling any given distance (Shoemaker et al. 2020). Dispersal kernels

60   thus account for variation in distance traveled as well as variation in the proportion of the

61   dispersing population traveling each distance (Kot et al. 1996). In most species, the majority of

62   dispersing individuals travel short distances, remaining close to their source location and thus

63   contributing to local population dynamics (Moles & Westoby 2004). Simultaneously, few

64   individuals move longer distances and they drive processes like colonization (Soons et al.

65   2004b), range expansions (Kot et al. 1996) and range shifts (Davis & Shaw 2001). This 'long-

66   distance dispersal' is often defined by the distance travelled by the farthest 1% of individuals

67   (Nathan 2006). Finally, population size can influence dispersal; populations with more

68   dispersing individuals will more fully 'realize' the dispersal kernel – thus being more likely to

69   successfully reach farther distances. Although network models have the potential to account for

70   variation in dispersal (i.e. via weighted edges, Shanafelt et al. 2017), most either assume fixed

71   dispersal (e.g. Minor et al. 2009), or use weighted edges to describe how easily an organism can

72   move through a given matrix, ignoring dispersal variation (e.g. Ziółkowska et al. 2014;

73   Wimberly et al. 2018). Those network models that do consider dispersal as a function of distance

74   tend to do so by using simulated draws from dispersal kernels (Fletcher et al. 2011, 2013).

75   Models that more fully explore weighted networks using dispersal kernels would allow for a

76   more nuanced representation of species' movement capacity, and provide a better understanding

77   of habitat connectivity and the impacts of fragmentation. A deeper understanding of connectivity

78   would influence both the conservation of rare and threatened species that have had natural

79   movement patterns altered by fragmentation, as well as the control of invasive species that have

80   strong movement abilities that are unaltered by fragmentation (Damschen et al. 2008).

81   Here, we create a series of deterministic network models to ask the question: how does

82   variation in dispersal alter estimates of landscape connectivity? First, we generate networks

83   based on fixed dispersal distances, allowing us to compare our results to previous studies (e.g.

84   Urban & Keitt 2001; O'Brien *et al.* 2006; Wimberly *et al.* 2018). Then, we generate networks

85   based on dispersal kernels, which incorporate variation in dispersal and include the effect of

86   population size. By comparing networks created either with fixed dispersal distances or with

87   dispersal kernels, we explore how variation in dispersal alters our understanding of aspects of

88   landscape connectivity including: habitat sub-structure, robustness to habitat loss, and robustness

89   to local extinction. Our approach provides a starting point for conservation managers interested

90   in understanding how traditional methods might over- or under-estimate connectivity based on

91   simplified assumptions about dispersal. We apply these models to the fragmented grasslands in

92   Minnesota (Fig. 1), where there is renewed interest from state and local managers in considering

93   connectivity in their restoration efforts. Minnesota managers have created plans for protecting

94   existing grasslands and building future restorations to promote connectivity through the creation

95   of grassland corridors (Minnesota Prairie Plan Working Group 2018), yet these plans were

96   designed with very minimal information on species movement because little is known. Thus,

97   Minnesota grasslands are an excellent study system to demonstrate the utility of our broader

98   approach because prior knowledge of, and interest in, connectivity exists, and managers there are

99   open to considering how to incorporate more realistic information on species' dispersal into

100   future conservation plans (Minnesota Prairie Plan Working Group 2018; Wimberly et al. 2018;

101   Sperry et al. 2019).

102

103   **Materials and Methods**

104   We created deterministic network models with and without dispersal kernels in order to

105   draw conclusions about how including biologically meaningful knowledge of dispersal alters

106   predictions about connectivity as compared to ignoring dispersal variation. We assumed that

107   dispersal kernels more accurately represent species movement than the assumption of fixed

108   models where all individuals travel all distances with equal likelihood. Our models are general,

109   and thus could apply to any species of interest, including Minnesota grasslands species across a

110   range of dispersal distances: prairie coneflower (~9m; Ison et al. 2014), ground squirrels (53-

111   80m; Rongstad 1965), dickcissels (222m; Walk et al. 2004), and burrowing owls (2802m; Catlin

112   & Rosenberg 2008) – all indicating mean dispersal distances.

113

114   *Habitat Selection*

115   We developed our models for the prairie region of western and southern Minnesota (also

116   called the prairie parkland region). This region was historically grassland but has been

117   fragmented and reduced to ~1% of its original area (Minnesota Prairie Plan Working Group

118   2018). The spatial locations of the remaining grasslands are well documented, and exist in a

119   matrix of mostly agriculture. We refer to each separate grassland fragment as a 'patch'

120   throughout. To build our networks, we used a comprehensive spatial grassland habitat database

121   for the region (The Nature Conservancy 2015). This dataset combined the Minnesota

122   Department of Natural Resources' native prairie layer, the United States Fish and Wildlife

123   Services' (USFS) Habitat and Population Evaluation Team's (HAPET) 2014 reclassification

124   dataset, and the United States Department of Agriculture's Cropland Data Layer (CDL) dataset.

125   The USFS National Wetlands Inventory layer was used for corrections in classifying wetland

126   and open water areas. The resulting database consisted of all grassland types, including native

127   remnant prairie, reconstructed or disturbed grasslands, and hay/pasture fields. We included all of

128   these grassland types in our network analysis because they represent potentially suitable habitat

129   for grassland species (e.g. birds, insects, mammals, plants). Similar to Wimberley et al. (2018)

130   we used ArcGIS 10.4 to select patches that were five acres (2.023 ha) or larger. This resulted in

131   ~37,000 grasslands (*N*=37,091 patches in the network, see Appendix S1 for all parameters) to

132   use in our connectivity analysis (Fig. 1a,b).

133

134   ***Networks Using Fixed Dispersal Distances***

135   First, we generated networks for our grasslands assuming that organisms had no dispersal

136   variation (i.e. all individuals traveled a fixed dispersal distance). To do this, we calculated the

137   nearest-edge distance (i.e. the distance between the closest points) of all pairs of patches in

138   ArcGIS using the geodesic method, and a max search radius of 4000m (Fig. 1c). We chose this

139   radius, which is within the range used in similar studies (Wimberly et al. 2018), for

140   computational simplification, but note that it is smaller than the movement ability of extremely

141   vagile Minnesota grassland species (e.g. the red fox disperses on average ~31,000m; Storm et al.

142   1976). We then generated 2000 networks, one for each fixed dispersal distance (d') that we

143   considered (1,...,2000m). For each dispersal distance d, we generated a binary adjacency matrix

144   A (of size N x N) where each element described whether (1) or not (0) the distance between a

145   pair of patches was less than d (i.e., whether an individual traveling that distance could move

146   between these two patches). This generated 2000 binary adjacency matrices A (of size N x N)

147   where each element described whether (1) or not (0) two patches were connected for each

148   dispersal distance. We then used each matrix to create a non-directed network using the igraph

149   package in R (Csardi & Nepusz 2006). To guide readers through our methods, we also created a

150   'toy' network (Fig. 2a).

151

152   ***Networks Using Dispersal Kernels***

153   We also generated networks for our grasslands assuming that individual organisms varied

154   in their dispersal ability (i.e. their movement was described by a dispersal kernel, Fig. 1d).

155   Specifically, we used the exponential distribution (Fig. 3a), where the proportion of individuals

156   traveling any distance d is

157                              $$e^{-bd}$$ (1)

158   with rate parameter, b. This distribution is commonly used as a dispersal kernel since it often

159   matches empirical data (Hovestadt et al. 2011; Shaw et al. 2019). As with the fixed distance

160   networks, here we also considered 2000 dispersal distances. However instead of considering

161   these to be the exact distances traveled, we considered these to be the farthest 1% value (d^*) for

162   defining 'long-distance dispersal' for our dispersal kernels. Thus we established a dispersal

163   kernel for each d_x^* value ($x = 1,...,2000$m) as follows. We determined the proportion of

164   individual dispersing each distance d or more, given by the complementary cumulative

165   distribution function (CCDF) for the exponential kernel (Fig. 3a). We set this proportion f to be

166   0.01, plugged in each d_x^*, and solved for the corresponding b_x value (Fig. 3c), that is

167    $$b_x = \frac{-\ln(0.01)}{d_x^*} \ . \qquad\qquad (2)$$

168   This b_x value describes a dispersal kernel where only 1% of individuals dispersed a distance of

169   d_x^* or further. Next, we calculated the nearest-edge distance between all pairs of patches up to a

170   maximum distance of 4000m, resulting in a Euclidian distance matrix, $\boldsymbol{D}$ (Fig. 3b). Setting a

171   maximum distance for these calculations (rather than calculating all pairwise distances) saved

172   computational time while ensuring we calculated all relevant distances needed for our kernels

173   below. Finally, for each dispersal distance d_x^* we converted the distance matrix, $\boldsymbol{D}$, into a matrix

174    $\boldsymbol{M}_x$, describing the proportion of the modelled population that disperses between each patch (Fig.

175   3c). To do this, for each d_x^* value, we used the CCDF to calculate the proportion of individuals

176    $m(i,j)$ with dispersal kernel defined by b_x that would travel at least the distance $d(i,j)$ between

177   each patch i and j. We then used these $\boldsymbol{M}_x$ matrices to generate weighted non-directed networks,

178   where the weight of each edge corresponds to the proportion of dispersing individuals that can

179   move between the two patches the edge connects.

180   Finally, we explored the influence of population size on connectivity metrics. A dispersal

181   kernel describes the distribution of distances that would be observed across a very large number

182   of dispersal events. However, since species vary in population size and fecundity, they will also

183   vary in how well the kernel is 'realized'. These differences will appear most strongly for the low-

184   probability longer distances (the dispersal kernel tail); a species with a smaller population size or

185   lower fecundity will have fewer realized dispersal events and thus across the population there

186   will be fewer dispersal distances represented by the tail of the kernel. To mimic different

187   population sizes with our models, we thresholded the matrix M at three values, by keeping the

188   75% highest dispersal probabilities (i.e. setting to zero all dispersal probabilities less than 25%;

189   '75%-realized', representing a small population), 99% ('99%-realized', medium population), and

190   99.99% ('99.99%-realized', large population). This is equivalent to truncating the dispersal

191   kernel at three increasingly long maximum distances, but does not incorporate uncertainty and

192   represents the simplifying assumption that small populations are less likely to reach longer

193   distances than large populations. Imposing a maximum dispersal distance also kept the dispersal

194   kernel from becoming infinite (i.e. there is a very small proportion of individuals dispersing

195   infinitely far). We note that an alternative approach to examining population size is to multiply

196   all weighted edges within the network by these proportions (for large populations, 0.9999, for

197   medium populations 0.99, for small populations 0.75), and then conduct network analyses. This

198   approach leads to qualitatively similar results for patch-level metrics to those we present here

199   (Appendix S2).

200   In total, we considered 2000 different measures of 'long-distance dispersal' and 3

201   different measures of population size, generating 6000 weighted networks. As above, we

202   calculated network and patch-level metrics for each network, some of which were modified to

203   accommodate the weighted network structure.

204

205   ***Connectivity Metrics***

206   *Network-level Metrics*

207   For each network generated with fixed dispersal (non-weighted) and dispersal kernels

208   (weighted), we calculated three network-level metrics (Fig. 2b), to quantify different aspects of

209   network structure and connectivity. The first two metrics do not take into account weights and

210   thus were calculated the same way for both non-weighted (binary) and weighted networks. (1)

211   Number of components, (see Appendix S3 for igraph functions): two patches are in the same

212   component if they are connected by an edge; fully isolated patches are their own component.

213   Patches in different components are isolated from each other, thus the number of separate

214   components in a network provides a rough sense of overall fragmentation across the network

215   (Calabrese & Fagan 2004). (2) Maximum component size: the number of patches in the largest

216   component of the network provides a measure of effective network size (Urban & Keitt 2001).

217   Both the number of components and the size of the largest component represent an estimate of

218   the amount of habitat sub-structure present. (3) Average clustering coefficient: this metric

219   quantifies the extent to which a network contains well-connected clusters of patches and thus

220   provides an estimate of local landscape connectivity. In non-weighted networks, for a given

221   patch i, that is connected to k_i neighboring patches (see patch-level metrics below), there can be

222   at most

223                         $(1/2) \, k_i \, (\, k_i - 1)$ (3)

224   connections among its neighboring patches. The clustering coefficient for this patch is the

225   fraction of those possible connections that actually occur (Watts & Strogatz 1998), a metric used

226   to quantify the local connectivity for landscape networks (Wimberly et al. 2018). Average

227   clustering coefficient can be considered a measure of robustness to habitat loss, as networks with

228   higher clustering will more easily maintain their sub-structure even as habitat fragmentation

229   removes either edges or patches. For weighted networks, we used weighted distances between

230   patches based on dispersal proportion (Csardi & Nepusz 2006). This weighted clustering

231   coefficient is calculated as

232    $$\frac{1}{s_i(k_i-1)}\Sigma_{j,h}\left\{\frac{1}{2}[m(i,j)+m(i,h)]a(i,j)a(i,h)a(j,h)\right\} \quad (4)$$

233   where s_i is the strength of patch i (see below), k_i is the degree of patch i, $m(i,j)$ are the elements

234   of the weighted matrix M, and $a(i,j)$ are the elements of the adjacency matrix A. We note that

235   there was little difference between the mean and median values for clustering coefficient, except

236   for small to moderate dispersal distances in the '75%-realized' kernels where the mean value was

237   larger than the median value (Appendix S4).

238

239   *Patch-level Metrics*

240   For each patch within each network, we also calculated two patch-level metrics (Fig. 2c),

241   and summarized them by looking at the 25th, 50th, and 75th quantiles of all values for patches

242   within each network. (1) Degree centrality: for non-weighted networks, this metric is calculated

243   as the number of connected neighbors each patch has (k_i), i.e. the number of patches that an

244   individual could potentially reach via dispersal as defined by the model (Wimberly et al. 2018).

245   For weighted networks we calculated strength, the weighted version of degree centrality, as

246    $$s_i = \Sigma_j m(i,j) \qquad\qquad\qquad (5)$$

247   where $m(i,j)$ are the elements of the weighted matrix M for all connected neighbors j of patch i.

248   Degree centrality or strength quantifies the number of colonization opportunities to or from each

249   patch and represents a measure of short-term robustness to local (patch-level) extinction. Patches

250   with low degree/strength are likely to be isolated and vulnerable to reductions in species richness

251   as any local extinction would be unlikely to be recovered by recolonization from other patches.

252   (2) Closeness centrality: for non-weighted networks, this metric quantifies the importance of

253   each patch i for overall connectivity in the network as

254    $$\frac{1}{\Sigma_{i \neq j} p_{ij}} \tag{6}$$

255   where p_{ij} is the shortest path, or the number of steps (i.e. sequential dispersal events) it takes to

256   reach every other patch j in the network from the focal patch. If two patches are not connected

257   (i.e. p_{ij} is infinite), the total number of patches (N) is used instead of p_{ij} for this pair. Thus

258   closeness is a measure of the average number of sequential dispersal events required to

259   recolonize the network, and represents a measure of long-term robustness to local extinction. We

260   chose closeness as our centrality metric (rather than 'betweenness' as used by Minor and Urban

261   (2007)) because closeness more accurately represents dispersing organisms that do not always

262   take the most efficient route between patches (Borgatti 2005). For weighted networks, we

263   calculated a weighted version of closeness, as

264    $$\frac{1}{\Sigma_{i \neq j} q_{ij}} \tag{7}$$

265   where q_{ij} is the sum of inverse probabilities m^{-1} along the shortest path between patch i and patch

266   j. Since the inverse of the proportion of dispersers gives an expected number of events needed

267   (e.g., a 0.5 proportion of dispersers would take about 2 dispersal events), weighted closeness is

268   again a measure of long-term robustness to local extinction because it tallies the expected

269   number of sequential dispersal events required to recolonize the entire network. As for non-

270   weighted networks, if two patches are not connected, the total number of patches (N) is used

271   instead of q_{ij} for that pair. Note that this correction for unconnected patches (while a suitable

272   approximation for non-weighted networks) is actually an underestimate of the number of

273   sequential dispersal events for weighted networks. Since the degree to which it underestimates

274   dispersal events (and thus overestimates weighted closeness) interacts with the different

275   population sizes we considered, there is no meaningful way to compare across different

276   truncations of the dispersal kernels for this metric. Thus, we only calculated weighted clustering

277   coefficient for the networks based on the '99%-realized' dispersal kernels.

278   All analyses were run in R v3.4.4 (R Core Team 2017), data and code are available at the

279   Dryad Digital Repository xxxxx.

280

281   **Results**

282   *Network-level Metrics*

283   Networks were less fragmented (had less habitat sub-structure) for larger dispersal

284   distances, with fewer components (Fig. 4a) and larger largest components (Fig. 4b). These

285   relationships were starkest for networks created from dispersal kernels with longer realized

286   kernel tails (i.e. larger population size or higher fecundity). In other words, the '99.99%-realized'

287   kernel showed the fastest drop in the number of components, and the fastest increase in size of

288   largest component with increasing d^*, while the '75%-realized' dispersal kernel showed a

289   markedly slower decrease in the number of components and slower increase in maximum

290   component size with increasing d^*. The fixed dispersal distance produced accurate estimates for

291   populations of intermediate size (the '99%-realized' dispersal kernel, Fig. 4a-b). Intuitively, this

292   result occurs because a network from a fixed dispersal distance of d' is structurally equivalent to

293   a network with a '99%-realized' dispersal kernel with distance d^* (the same patches are

294   connected in both when considering non-weighted (or binary) network metrics like the number

295   of components and largest component size). However, fixed dispersal distance underestimated

296   habitat sub-structure for smaller populations ('75%-realized') and overestimated habitat sub-

297   structure for larger populations ('99.99%-realized').

298   Networks were also more connected for larger dispersal distances, with higher clustering

299   coefficients (Fig. 4c). In other words, populations with larger dispersal distances are more robust

300   to habitat loss leading to lost patches or connections. However, the fixed dispersal distance

301   consistently underestimated robustness to habitat loss compared to all three populations sizes

302   ('75%-, 99%-, 99.99%-realized' dispersal kernels, Fig. 4c), with the largest difference for the

303   largest population sizes ('99.99%-realized' kernel). The fixed network and the '75%-realized'

304   kernel produced similar results for low dispersal distances, but the clustering coefficient then

305   plateaued for the fixed distance while the '75%-realized' kernel continued to increase for larger

306   dispersal distances.

307

308   *Patch-level Metrics*

309   Patches in networks with larger dispersal distances were on average connected to more

310   neighbor patches (higher degree centrality, Fig. 5a), representing a higher short-term robustness

311   to local extinctions. Networks from fixed dispersal distances consistently overpredicted

312   robustness compared to networks from dispersal kernels, a gap that increased with dispersal

313   distance. In other words, fixed kernel networks systematically overpredict the number of

314   neighbors (and thus expected number of recolonization opportunities) each patch has, compared

315   to dispersal kernel networks. Within the dispersal kernel networks, smallest populations ('75%-

316   realized' kernel) had patches with the lowest robustness, followed by medium ('99%-realized')

317   and large ('99.99%-realized') population sizes, however these ranges overlapped substantially.

318   Similarly, patches in networks with larger dispersal distances had higher closeness values (Fig.

319   5b). Fixed dispersal distances consistently overpredicted closeness, and thus underpredicted the

320   number of sequential dispersal events needed to recolonize a network following extinction,

321   compared to the networks created with dispersal kernels.

322

323   **Discussion**

324   We built deterministic network models from fixed dispersal distances and dispersal

325   kernels, and contrasted them to more fully explore how weighted networks that use dispersal

326   kernels affect estimates of landscape connectivity. As with other simulation-based connectivity

327   models that incorporate dispersal variation (Palmer et al. 2014), we found that network models

328   based on dispersal kernels generated a markedly different understanding of population

329   connectivity than network models based on a fixed dispersal distance (Figs. 4-5, Appendix S5).

330   Specifically, using fixed dispersal consistently overestimated a population's robustness to local

331   extinctions while simultaneously underestimating robustness to habitat fragmentation. Our

332   results from fixed dispersal distances qualitatively match similar network analyses for other

333   grasslands (Wimberly *et al.* 2008) and for forests (Urban & Keitt 2001), suggesting that current

334   habitat management using fixed dispersal networks is based on inaccurate estimates of

335   population connectivity. As there is ample evidence that most organisms have substantial

336   variation in dispersal (e.g. Baguette 2003; Krkošek et al. 2007; Sullivan et al. 2018), connectivity

337   models must account for such variation by using dispersal kernels. Other network models that

338   use dispersal kernels to match empirical movement data find these methods to be a good

339   approximation of movement ability (Fletcher et al. 2011, 2013). These findings have

340   implications for managers that plan for conservation based on connectivity metrics. Some

341   species of concern may need more total habitat, while others rely on continual recolonization and

342   thus would differ in whether fixed models over- or underestimated their connectivity.

343   The magnitude of differences between fixed and dispersal kernel connectivity metrics

344   depended on how we modeled the tail of the dispersal kernel, which reflected a examining

345   different population sizes of organisms. The underestimate of robustness to habitat fragmentation

346   (clustering) was the largest for large populations ('99.99%-realized' dispersal kernel, Fig. 4c). In

347   contrast, the overestimate of robustness to local extinction (degree centrality) was similar for all

348   population sizes, but slightly larger for small populations ('75%-realized', Fig. 5a). Degree

349   centrality estimates the expected number of patches that can be colonized with a single set of

350   dispersal events. Since fixed dispersal effectively assumes 'perfect' dispersal (patches within a

351   fixed distance will always be reached), networks with fixed dispersal will always overestimate

352   colonization ability.

353   In light of our results, explicit consideration of conservation goals can help guide the

354   appropriate use of dispersal kernels for management and planning. Inherent in the use of

355   dispersal kernels is the understanding that most individuals move shorter distances, and few

356   individuals move farther distances. Therefore, the conservation goals at the heart of maintaining

357   connectivity should take population size into account when appropriate. For example, often the

358   goal of promoting connectivity between patches is to build a functioning meta-population for

359   rare species where individuals can move freely and breed between patches (Hanski 1998). As

360   rare or threatened species are often dispersal-limited due to small population sizes and low

361   fecundity (Baur 2014), considering a less realized dispersal kernel (i.e. '75%-realized') could

362   more accurately represent likely connectivity outcomes for this particular goal. Moreover, if

363   small population sizes are of serious concern, other methods might need to be incorporated,

364   including individual based models (Grimm & Railsback 2005). Another goal of maintaining

365   connectivity might be to allow for the possibility of species' response to climate change via

366   range shifts (Krosby et al. 2010). Range expansions often proceed through the dispersal of a few

367   individuals over a long distance (Davis & Shaw 2001). To successfully track climate change,

368   large populations must produce the few individuals that disperse long distances, thus a more

369   realized dispersal kernel (i.e. '99.99%-realized') would be more appropriate to include in

370   network models to achieve this goal. Consideration of these highly realized dispersal kernels is

371   also appropriate for controlling invasive species like the cane toad, which have high movement

372   ability (Perkins et al. 2013). Finally, for sessile organisms like plants, managers might be

373   interested in distinguishing between maintaining high genetic diversity to decrease the

374   probability of inbreeding depression - which requires the movement of gametes (i.e. pollen),

375   versus allowing for species recolonization to increase species diversity - which requires the

376   movement of individuals (i.e. seeds) (Elistrand 1992; Brudvig et al. 2009). In this case managers

377   should consider defining dispersal kernels that represent pollen and seed dispersal separately in

378   order to match their management goals. When looking to define dispersal kernels, managers can

379   use measurement-based (e.g. Stevens et al. 2010), trait-based (e.g. Soons et al. 2004a), or

380   genetic-based approaches (e.g. Bacles et al. 2006) to estimate kernels.

381   Grasslands are globally important, yet they are among the most threatened due to land

382   use conversion and fragmentation (Soons et al. 2005; Newbold et al. 2016). Our network models

383   help elucidate how likely species are able to move between grassland patches and maintain

384   connectivity at a broader scale. Our results are comparable to Wimberly et al. (2018), which

385   determined connectivity of the grasslands in the Prairie Coteau region of Minnesota and the

386   Dakotas, but use a fixed dispersal distance. Extrapolating their results based on our findings from

387   network models with dispersal kernels, we might expect that for species with large population

388   sizes there might be increased connectivity, with fewer, larger components that are more robust

389   to fragmentation than what Wimberly et al. (2018) found, but we might expect the opposite for

390   species with smaller populations. To aid Minnesota grassland managers, we have created a web-

391   based app to allow for the direct application of network models to existing grasslands in

392   Minnesota (Sperry et al. 2019). This approach could be easily updated to incorporate known

393   dispersal kernel information for species of interest (e.g. grassland plant species - Sullivan et al.

394   2018), patch prioritization, or matrix quality between patches (Castillo et al. 2016), which would

395   afford a more targeted understanding of which species can maintain connectivity, and which

396   might require assistance moving between patches.

397

398   **Conclusions**

399   In order to understand if and where connectivity is maintained between isolated habitat

400   fragments, we must account for how organisms move in a biologically meaningful way. We take

401   steps toward this goal by considering variability in dispersal in network models by incorporating

402   fully explored dispersal kernels, to determine how this alters our view of network-based

403   connectivity as compared to standard methods that use a fixed dispersal distance. Since

404   interspecific dispersal variation is also common, future work should examine how dispersal

405   varies across species (e.g., when different species have different dispersal kernel shapes), and

406   when there is directionality in dispersal to understand more fully how interspecific variation

407   affects connectivity. We find that models ignoring dispersal variation simultaneously

408   overestimate robustness to local extinctions while underestimating robustness to habitat loss,

409   compared to models that account for dispersal variation. The magnitude of these differences

410   depends both biological traits of the species of interest, particularly population size, and dispersal

411   distance.

412   **References**

413

414   Bacles CFE, Lowe AJ, Ennos RA. 2006. Effective seed dispersal across a fragmented landscape.

415   Science **311**:628.

416   Baguette M. 2003. Long distance dispersal and landscape occupancy in a metapopulation of the

417   Cranberry Fritillary butterfly. Ecography **26**:153–160.

418   Baur B. 2014. Dispersal-limited species - A challenge for ecological restoration. Basic and

419   Applied Ecology **15**:559–564.

420   Borgatti SP. 2005. Centrality and network flow. Social Networks **27**:55–71.

421   Brudvig LA, Damschen EI, Tewksbury JJ, Haddad NM, Levey DJ. 2009. Landscape

422   connectivity promotes plant biodiversity spillover into non-target habitats. Proceedings of

423   the National Academy of Sciences of the United States of America **106**:9328–32.

424   Calabrese JM, Fagan WF. 2004. A comparison-shopper' s guide to connectivity metrics.

425   Frontiers in Ecology and the Environment **2**:529–536.

426   Castillo JA, Epps CW, Jeffress MR, Ray C, Rodhouse TJ, Schwalm D. 2016. Replicated

427   Landscape genetic and network analyses reveal wide variation in functional connectivity for

428   American pikas. Ecological Applications **26**:1660–1676.

429   Catlin DH, Rosenberg DK. 2008. Breeding dispersal and nesting behavior of burrowing owls

430   following experimental nest predation. The American Midland Naturalist **159**:7.

431   Creech TG, Epps CW, Monello RJ, Wehausen JD. 2014. Using network theory to prioritize

432   management in a desert bighorn sheep metapopulation. Landscape Ecology **29**:605–619.

433   Csardi G, Nepusz T. 2006. The igraph software package for complex network research.

434   InterJournal **Complex Sy**:1695.

435   Damschen EI, Brudvig LA, Haddad NM, Levey DJ, Orrock JL, Tewksbury JJ. 2008. The

436 movement ecology and dynamics of plant communities in fragmented landscapes.

437 Proceedings of the National Academy of Sciences **105**:19078–83.

438 Davis MB, Shaw RG. 2001. Range shifts and adaptive responses to Quaternary climate change.

439 Science **292**:673–9.

440 Elistrand NC. 1992. Gene flow by pollen: Implications for plant conservation genetics. Oikos

441 **63**:77–86.

442 Fahrig L. 2017. Ecological responses to habitat fragmentation per se. Annual Review of

443 Ecology, Evolution, and Systematics **48**:1–23.

444 Fletcher RJ et al. 2018. Is habitat fragmentation good for biodiversity? Biological Conservation

445 **226**:9–15.

446 Fletcher RJ, Acevedo MA, Reichert BE, Pias KE, Kitchens WM. 2011. Social network models

447 predict movement and connectivity in ecological landscapes. Proceedings of the National

448 Academy of Sciences **108**:19282–19287.

449 Fletcher RJ, Revell A, Reichert BE, Kitchens WM, Dixon JD, Austin JD. 2013. Network

450 modularity reveals critical scales for connectivity in ecology and evolution. Nature

451 Communications **4**:1–7.

452 Grimm V, Railsback SF. 2005. Individual-based modeling and ecology. Princeton University

453 Press.

454 Haddad NM et al. 2015. Habitat fragmentation and its lasting impact on Earth's ecosystems.

455 Science Advances **1**:e1500052–e1500052.

456 Hanski I. 1998. Metapopulation dynamics. Nature **396**:41–49.

457 Hovestadt T, Binzenhöfer B, Nowicki P, Settele J. 2011. Do all inter-patch movements represent

458 dispersal? A mixed kernel study of butterfly mobility in fragmented landscapes. Journal of

459   Animal Ecology **80**:1070–1077.

460   Ison JL, Wagenius S, Reitz D, Ashley M V. 2014. Mating between Echinacea angustifolia

461   (Asteraceae) individuals increases with their flowering synchrony and spatial proximity.

462   American Journal of Botany **101**:180–189.

463   Kot M, Lewis MA, van den Driessche P. 1996. Dispersal data and the spread of invading

464   organisms. Ecology **77**:2027–2042.

465   Krkošek M, Lauzon-Guay J-S, Lewis MA. 2007. Relating dispersal and range expansion of

466   California sea otters. Theoretical Population Biology **71**:401–407.

467   Krosby M, Tewksbury J, Haddad NM, Hoekstra J. 2010. Ecological connectivity for a changing

468   climate. Conservation Biology **24**:1686–1689.

469   McRae BH, Dickson BG, Keitt TH, Shah VB. 2008. Using circuit theory to model connectivity

470   in ecology, evolution, and conservation. Ecology **89**:2712–2724.

471   Minnesota Prairie Plan Working Group. 2018. Minnesota Prairie Conservation Plan, Second

472   Edition.

473   Minor ES, Tessel SM, Engelhardt K a M, Lookingbill TR. 2009. The role of landscape

474   connectivity in assembling exotic plant communities: a network analysis. Ecology **90**:1802–

475   1809.

476   Minor ES, Urban DL. 2007. Graph theory as a proxy for spatially explicit population models in

477   conservation planning. Ecological Applications **17**:1771–1782.

478   Moilanen A, Hanski I. 1998. Metapopulation dynamics: Effects of habitat quality and landscape

479   structure. Ecology **79**:2503–2515.

480   Moles AT, Westoby M. 2004. Seedling survival and seed size: a synthesis of the literature.

481   Journal of Ecology **92**:372–383.

482   Nathan R. 2006. Long-distance dispersal of plants. Science **313**:786–788.

483   Newbold T et al. 2016. Has land use pushed terrestrial biodiversity beyond the planetary

484   boundary? A global assessment. Science **354**:288–291.

485   O'Brien D, Manseau M, Fall A, Fortin MJ. 2006. Testing the importance of spatial configuration

486   of winter habitat for woodland caribou: An application of graph theory. Biological

487   Conservation **130**:70–83.

488   Palmer SCF, Coulon A, Travis JMJ. 2014. Inter-individual variability in dispersal behaviours

489   impacts connectivity estimates. Oikos **123**:923–932.

490   Perkins TA, Phillips BL, Baskett ML, Hastings A. 2013. Evolution of dispersal and life history

491   interact to drive accelerating spread of an invasive species. Ecology Letters **16**:1079–87.

492   R Core Team. 2017. R: A Language and Environment for Statistical Computing. Vienna,

493   Austria. Available from http://www.r-project.org/.

494   Rockström J. 2009. A safe operating space for humanity. Nature **461**:472–475.

495   Rongstad OJ. 1965. A life history study of thirteen-lined ground squirrels in southern Wisconsin.

496   Journal of Mammalogy **46**:76–87.

497   Saura S, Rubio L. 2010. A common currency for the different ways in which patches and links

498   can contribute to habitat availability and connectivity in the landscape. Ecography **33**:523–

499   537.

500   Shanafelt DW, Salau KR, Baggio JA. 2017. Do-it-yourself networks: A novel method of

501   generating weighted networks. Royal Society Open Science **4**:171227.

502   Shaw AK. 2020. Causes and consequences of individual variation in animal movement.

503   Movement Ecology **8**.

504   Shaw AK, D'Aloia CC, Buston PM. 2019. The evolution of marine larval dispersal kernels in

505   spatially structured habitats: Analytical models, individual-based simulations, and

506   comparisons with empirical estimates. The American Naturalist **3**:424–435.

507   Shoemaker LG et al. 2020. Integrating the underlying structure of stochasticity into community

508   ecology. Ecology **101**:e02922.

509   Snell RS et al. 2019. Consequences of intraspecific variationin seed dispersal for plant

510   demography, communities, evolution and global change. AoB Plants **11**:plz016.

511   Soons MB, Heil GW, Nathan R, Katul GG. 2004a. Determinants of long-distance seed dispersal

512   by wind in grasslands. Ecology **85**:3056–3068.

513   Soons MB, Messelink JH, Jongejans E, Heil GW. 2005. Habitat fragmentation reduces grassland

514   connectivity for both short-distance and long-distance wind-dispersed forbs. Journal of

515   Ecology **93**:1214–1225.

516   Soons MB, Nathan R, Katul GG. 2004b. Human effects on long-distance wind dispersal and

517   colonization by grassland plants. Ecology **85**:3069–3079.

518   Sperry KP, Shaw AK, Sullivan LL. 2019. Apps can help bridge restoration science and

519   restoration practice. Restoration Ecology:3–6.

520   Stevens VM, Turlure C, Baguette M. 2010. A meta-analysis of dispersal in butterflies. Biological

521   Reviews **85**:625–642.

522   Storm GL, Andrews RD, Phillips RL, Bishop RA. 1976. Morphology, reproduction, dispersal,

523   and mortality of Midwestern Red Fox populations. Wildlife Monographs **49**:3–53.

524   Sullivan LL, Clark AT, Tilman D, Shaw AK. 2018. Mechanistically derived dispersal kernels

525   explain species-level patterns of recruitment and succession. Ecology **99**:2415–2420.

526   The Nature Conservancy. 2015. 2015 Prairie Plan Land Cover Analysis. Minneapolis, MN.

527   Urban D, Keitt T. 2001. Landscape connectivity: A graph-theoretic perspective. Ecology

528   **82**:1205–1218.

529   Walk JW, Wentworth K, Kershner EL, Bollinger EK, Warner RE. 2004. Renesting decisions and

530   annual fecundity of female Dickcissels (Spiza americana) in Illinois. The Auk **121**:1250–

531   1261.

532   Watts DJ, Strogatz SH. 1998. Collective dynamics of "small-world" networks. Nature **393**:440–

533   442.

534   Wimberly MC, Narem DM, Bauman PJ, Carlson BT, Ahlering MA. 2018. Grassland

535   connectivity in fragmented agricultural landscapes of the north-central United States.

536   Biological Conservation **217**:121–130.

537   Wright CK, Wimberly MC. 2013. Recent land use change in the Western Corn Belt threatens

538   grasslands and wetlands. Proceedings of the National Academy of Science **110**:4134–4139.

539   Ziółkowska E, Ostapowicz K, Radeloff VC, Kuemmerle T. 2014. Effects of different matrix

540   representations and connectivity measures on habitat network assessments. Landscape

541   Ecology **29**:1551–1570.

542

543

544

545   **Figure Legends**

546   Figure 1. (a) Map of the location of the ~37,000 grassland patches across Minnesota used in our

547   network analysis, as well as (b) the histogram showing distances between patches up to 4000m

548   (the max our models examined). Examples of a subset of the network (from Clay County, MN)

549   showing patches and connections under (c) the fixed distance (with $d' = 2000$) and (d) dispersal

550   kernel (with $d^* = 2000$ and '99%-realized'). In panel (d) thicker lines correspond to a higher

551   proportion of dispersers between patches.

552

553   Figure 2. (a) Toy network with nine patches (points labeled A-I) and two components, where

554   numbers along edges indicate the probability that two patches are connected via dispersal – used

555   for calculating weighted metrics. When calculating non-weighted metrics, the probability values

556   along connections become 1. (b) Network-level metrics calculated for this network. (Since

557   clustering coefficient only counts nodes with degree>1, the only nodes used for these

558   calculations are B, G, H and I). (c) Patch-level metrics as calculated for two example patches (A

559   and B, within component #1).

560

561   Figure 3. Methods schematic for building networks from dispersal kernels. (a) Use the dispersal

562   kernel (the proportion of individuals traveling a distance d) to calculate the complementary

563   cumulative density function (CCDF, f, the proportion of individuals traveling a distance d or

564   more). Very few (1%) individuals travel a distance d^* or more, considered "long-distance

565   dispersal". (b) Use the landscape of N patches to calculate a distance matrix (D, the physical

566   distance between all pairs of patches i and j). (c) Use the CCDF to map from each "long"

567   distance (d_x^*) to corresponding dispersal kernel parameter (b_x) by setting f_x to 0.01, then use both

25

568   the CCDF and the distance matrix to create a corresponding weighted matrix ($\mathbf{M_x}$, the proportion

569   of individuals dispersing between all pairs of patches *i* and *j*), where x is the dispersal distance

570   index x=1…*n*, (*n*=2000). See Appendix S1 for full definitions of parameters.

571

572   Figure 4. Network-level metric values for networks with fixed dispersal distances (red) and using

573   the exponential dispersal kernel at various tail truncations '75%-realized' (dark gray), '99%-

574   realized' (medium gray) and '99.99%-realized' (light gray), which represent increasing abilities

575   for long-distance dispersal. Panels show the (a) number of components, (b) size of the largest

576   component, which represent measures of habitat sub-structure and (c) the clustering coefficient,

577   which represents robustness to habitat loss.

578

579   Figure 5. Patch-level metric values for networks with fixed dispersal distances (red), and

580   networks with dispersal kernels, '75%-realized' (dark gray), '99%-realized' (medium gray) and

581   '99.99%-realized' (light gray). Panels show the 25th, 50th and 75th quantiles for (a) patch degree

582   centrality, and (b) patch closeness which represent measures of robustness to local extinction. In

583   panel (a), the '99%-realized' (medium gray) and '99.99%-realized' (light gray) are nearly

584   overlapping with the '99.99%-realized' kernel having a slightly higher degree. For clarity,

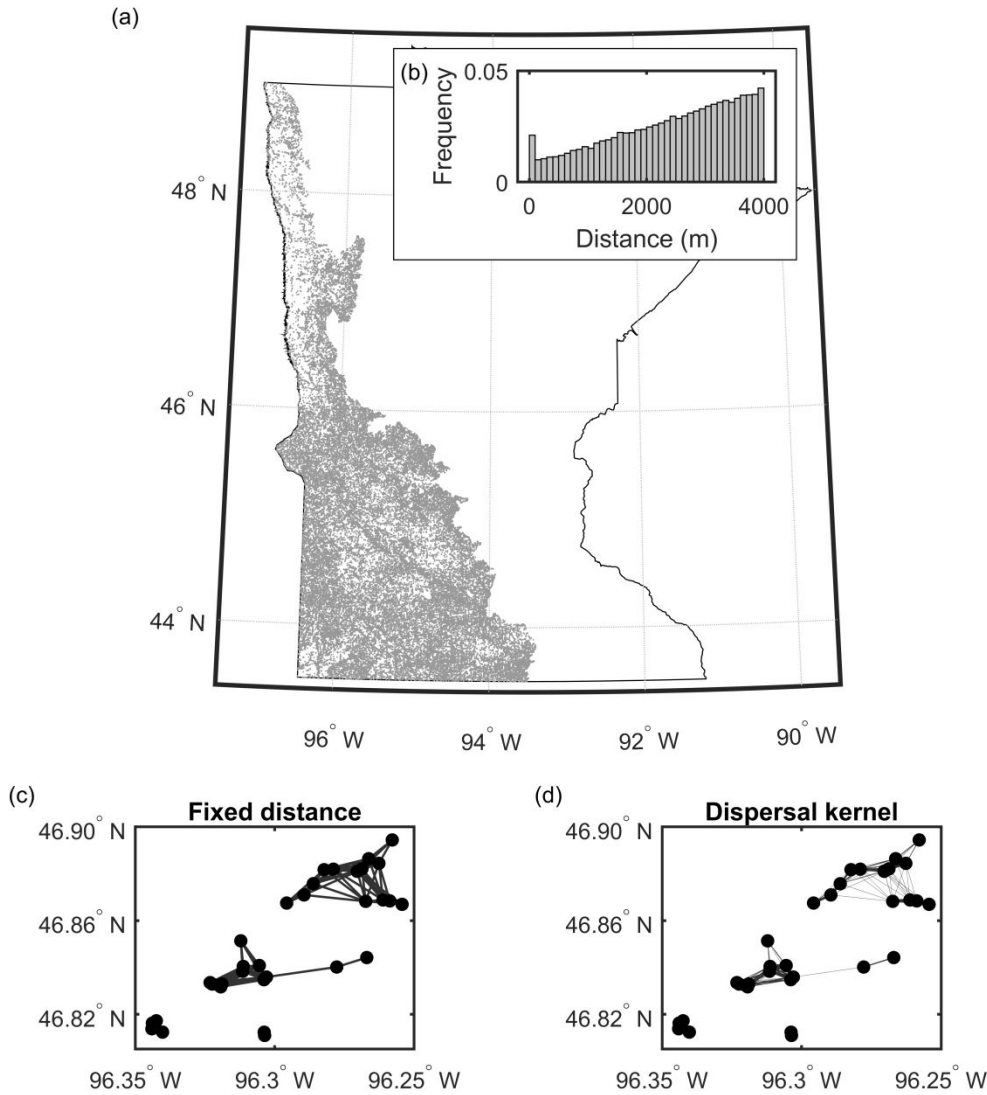585   asterisks represent the 50th quantile for each kernel.

586

587

588

589

Figure 1. (a) Map of the location of the ~37,000 grassland patches across Minnesota used in our network analysis, as well as (b) the histogram showing distances between patches up to 4000m (the max our models examined). Examples of a subset of the network (from Clay County, MN) showing patches and connections under (c) the fixed distance (with $d' = 2000$) and (d) dispersal kernel (with $d^* = 2000$ and '99%-realized'). In panel (d) thicker lines correspond to a higher proportion of dispersers between patches.
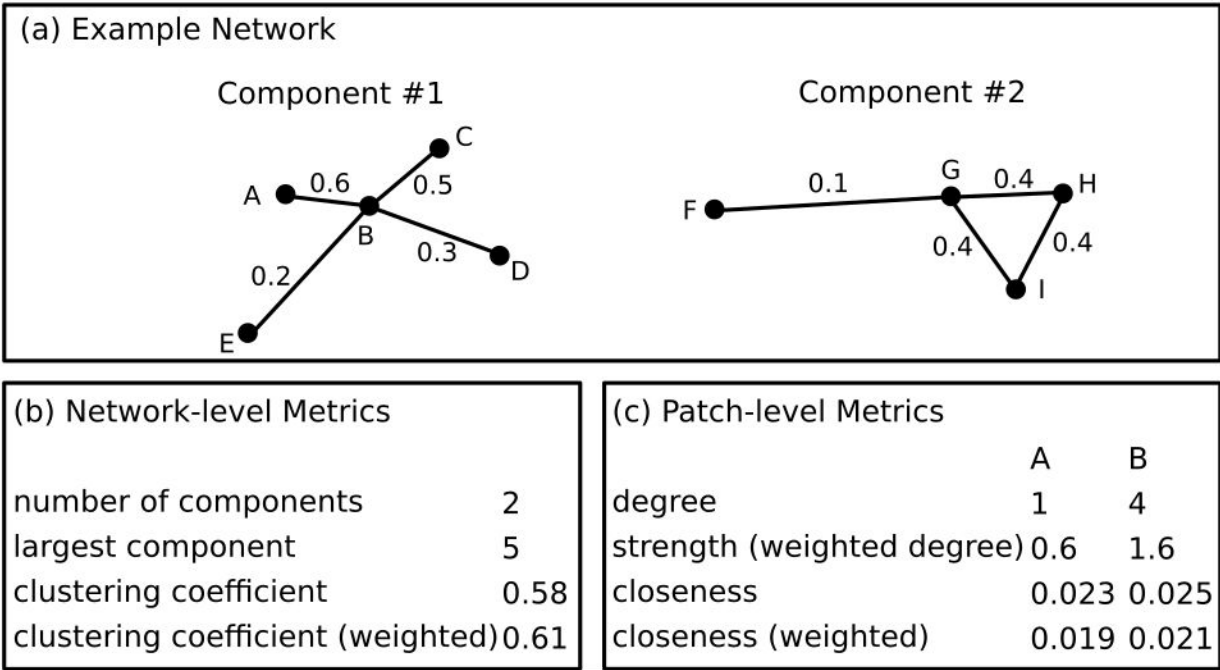
Figure 2. (a) Toy network with nine patches (points labeled A-I) and two components, where numbers along edges indicate the probability that two patches are connected via dispersal – used for calculating weighted metrics. When calculating non-weighted metrics, the probability values along connections become 1. (b) Network-level metrics calculated for this network. (Since clustering coefficient only counts nodes with degree>1, the only nodes used for these calculations are B, G, H and I). (c) Patch-level metrics as calculated for two example patches (A and B, within component #1).
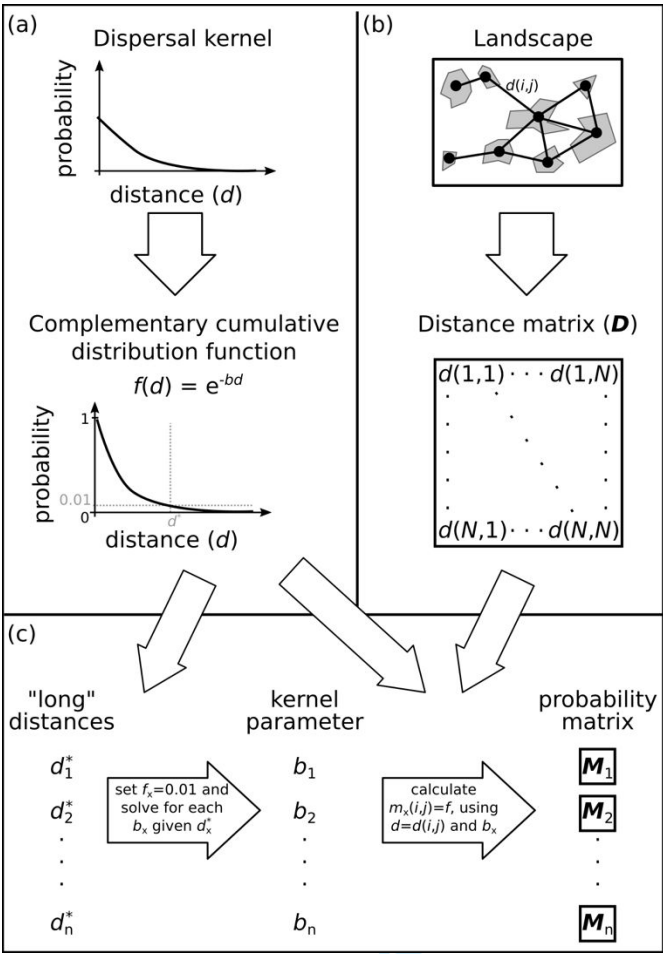
607

608   Figure 3. Methods schematic for building networks from dispersal kernels. (a) Use the dispersal

609   kernel (the proportion of the population traveling a distance *d*) to calculate the complementary

610   cumulative density function (CCDF, *f*, the proportion of individuals traveling a distance *d* or

611   more). Very few (1%) individuals travel a distance *d** or more, considered "long-distance

612   dispersal". (b) Use the landscape of *N* patches to calculate a distance matrix (**D**, the physical

613   distance between all pairs of patches *i* and *j*). (c) Use the CCDF to deterministically map from

614   each "long" distance (d_x*) to corresponding dispersal kernel parameter (b_x) by setting f_x to 0.01,

615   then use both the CCDF and the distance matrix to create a corresponding weighted matrix (**M$_x$**,

616   the proportion of individuals dispersing between all pairs of patches *i* and *j*), where x is the

617   dispersal distance index x=1…*n*, (*n*=2000). See Appendix S1 for full definitions of parameters.
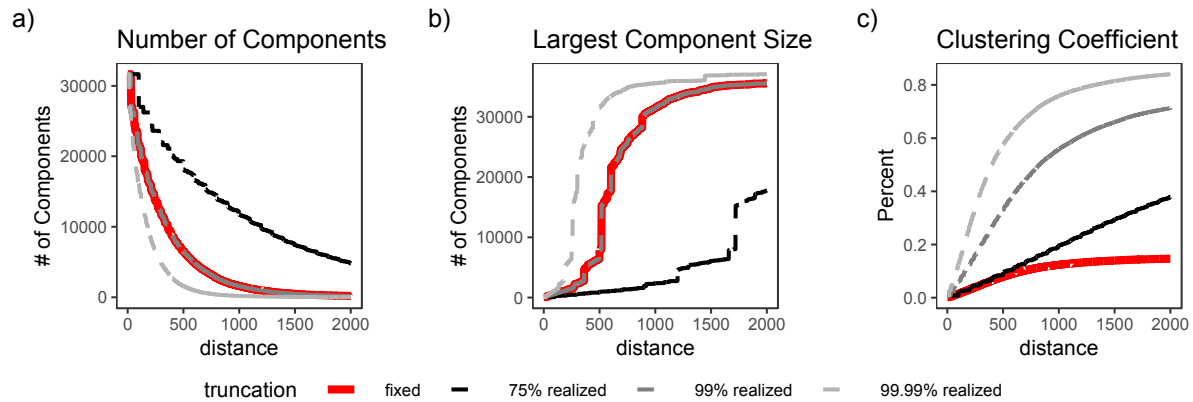
Figure 4. Network-level metric values for networks with fixed dispersal distances (red) and using the exponential dispersal kernel at various tail truncations '75%-realized' (dark gray), '99%-realized' (medium gray) and '99.99%-realized' (light gray), which represent increasing abilities for long-distance dispersal. Panels show the (a) number of components, (b) size of the largest component, which represent measures of habitat sub-structure and (c) the clustering coefficient, which represents robustness to habitat loss.
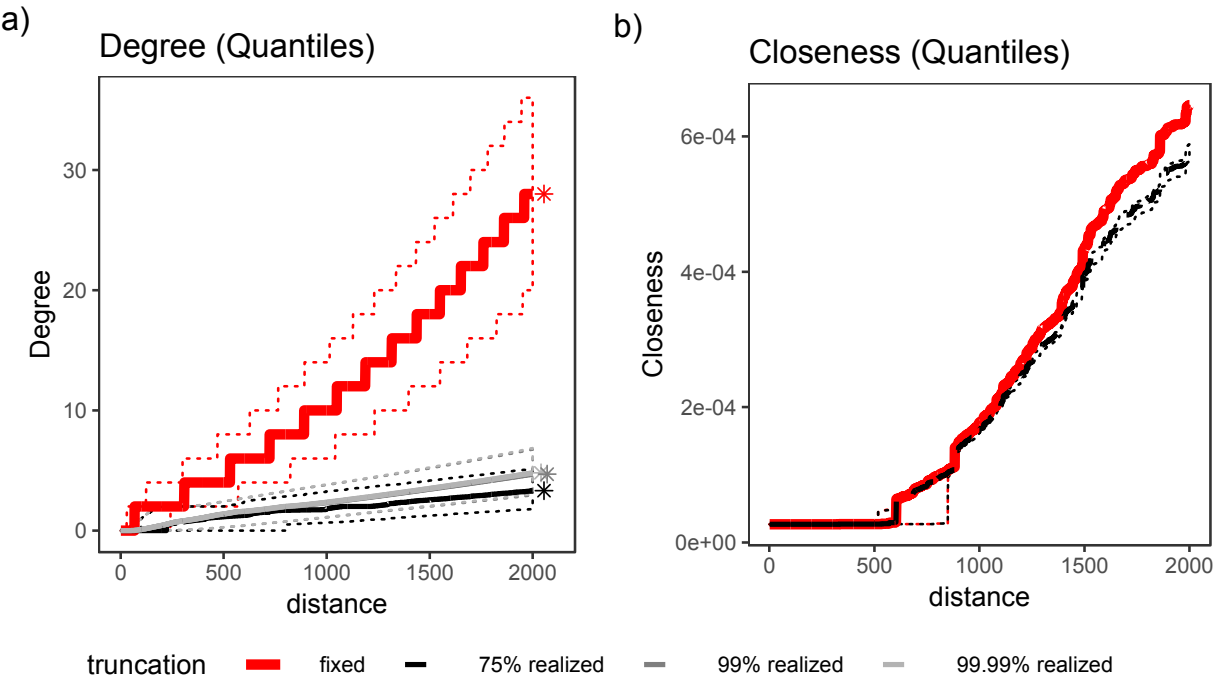
a) Degree (Quantiles)

b) Closeness (Quantiles)

truncation — fixed — 75% realized — 99% realized — 99.99% realized

629

Figure 5. Patch-level metric values for networks with fixed dispersal distances (red), and

networks with dispersal kernels, '75%-realized' (dark gray), '99%-realized' (medium gray) and

'99.99%-realized' (light gray). Panels show the 25th, 50th and 75th quantiles for (a) patch degree

centrality, and (b) patch closeness which represent measures of robustness to local extinction. In

panel (a), the '99%-realized' (medium gray) and '99.99%-realized' (light gray) are nearly

overlapping with the '99.99%-realized' kernel having a slightly higher degree. For clarity,

asterisks represent the 50th quantile for each kernel.

637

638

639