RESEARCH ARTICLE

Roads and forest edges facilitate yellow fever virus dispersion

Paula Ribeiro Prist1 | Leandro Reverberi Tambosi1,2 | Luís Filipe Mucci3 | Adriano Pinter3 | Renato Pereira de Souza4 | Renata de Lara Muylaert5 | Jonathan Roger Rhodes6 | César Henrique Comin7 | Luciano da Fontoura Costa8 | Tatiana Lang D'Agostini9 | Juliana Telles de Deus3 | Mônica Pavão10 | Márcio Port-Carvalho11,12 | Leila Del Castillo Saad9 | Maria Anice Mureb Sallum13 | Roberta Maria Fernandes Spinola9 | Jean Paul Metzger1

1Department of Ecology, Institute of Bioscience, University of São Paulo, São Paulo, Brazil; 2Center for Engineering, Modelling and Applied Social Sciences, Federal University of ABC, Santo André, Brazil; 3SUCEN—Endemics Control Superintendence, São Paulo, Brazil; 4Adolfo Lutz Institute—CLR XII. Praça Cel. Vitoriano, Taubaté, Brazil; 5Molecular Epidemiology and Public Health Laboratory, Hopkirk Research Institute, Massey University, Palmerston North, New Zealand; 6School of Earth and Environmental Sciences, The University of Queensland, Brisbane, QLD, Australia; 7Department of Computer Science, Federal University of São Carlos, São Carlos, Brazil; 8São Carlos Institute of Physics, University of São Paulo, São Carlos, Brazil; 9Center for Epidemiology Surveillance ‘Dr Alexandre Vranjac’, Coordination for Disease Control, Public Health Branch, São Paulo, Brazil; 10Geoprocessing and Spatial Analysis Core, Environment Research Institute. Infrastructure and Environment Secretariat of São Paulo, São Paulo, Brazil; 11Conservation Biodiversity Nucleus, Environmental Research Institute, Infrastructure and Environment Secretariat of São Paulo, São Paulo, SP, Brazil; 12Post Graduated Program in Biodiversity of Conservations Units, National School of Tropical Botanical—Rio de Janeiro Botanical Garden, Rio de Janeiro, Brazil and 13Epidemiology Department, Public Health Faculty, University of São Paulo, São Paulo, Brazil

Correspondence
Paula Ribeiro Prist
Email: pprist@hotmail.com

Present address
Paula Ribeiro Prist, EcoHealth Alliance, 520 Eighth Avenue, Ste. 1200, New York, NY 10018, USA

Funding information

Handling Editor: Aníbal Pauchard

Abstract
1. Landscape connectivity is important for a wide range of ecological processes, including to disease spread, once it describes the degree to which landscapes facilitate or impede vector and hosts dispersion. Understanding connectivity is extremely important to identify where pathogens can move, and at what speed, allowing the organization of vaccination campaigns or other preventive measures.

2. To better understand the effects of landscape connectivity on yellow fever virus (YFV) dispersion in Brazil, we used a network approach and modelled the movement of non-human primates’ cases, the so-called epizootic events, over time. The networks consider each epizootic event as a node and the dispersion between nodes as links. The links were established considering, respectively, the date of each epizootic event, the distance among the nodes and the permeability of the landscape between each pair of nodes.

3. Our results demonstrated that on average YFV dispersed 1.42 km/day, with the largest movement being 6.9 km/day. Dispersions were longer in summer (1.2 km/day) than in winter (0.22 km/day). Most dispersal movements occurred up to 1 km/day (71%) and within a week after the arrival of the virus in the source node (73%), except in winter, where dispersions occurred within a period of up to 20 days. The best model indicates that YFV disperses mainly through roads adjacent to forest areas, and along forest edges (within a range of 100 m) in interface with agricultural areas, water and forestry areas. Core areas of urban, agricultural and forest regions were important barriers for virus movement.
1 | INTRODUCTION

Emerging infectious diseases are a growing and significant threat to global public health (Parrish et al., 2008), with zoonotic pathogens counting for 75% of these diseases (Taylor et al., 2001). Landscape structure and connectivity (the capacity of landscapes to facilitate flows) can affect the transmission of these diseases, since they influence the dispersal of vectors, and modulate the persistence of host, reservoir and vector species (Estrada-Peña et al., 2014). Understanding the relationship between landscape connectivity and disease spread in tropical regions is particularly relevant not only to promote effective prevention actions but also because tropical areas harbour high diversity and are experiencing rapid land use and land cover dynamics (de Andreazzi et al., 2020).

In epidemiology, connectivity can be determined by the movement of infected or infectious hosts and vectors, and is influenced by multiple factors, including landscape features that can inhibit or facilitate dispersion (Russell et al., 2006). Habitat fragmentation, for example, can alter the functional connectivity and consequently the flow of organisms across a landscape (Crooks & Sanjayan, 2006), affecting the potential for disease spread (McCallum & Dobson, 2002). As such, landscape can be considered a key issue in determining the spatio-temporal dispersion of a disease.

Yellow fever virus (YFV) is a mosquito-borne flavivirus that causes a haemorrhagic fever on about 200,000 people and 30,000 deaths annually (Garske et al., 2014). In Brazil, the virus circulates in a sylvatic transmission cycle among non-human primates and diurnal mosquitoes Haemagogus spp. and Sabethes spp., with spillover occurring when mosquitoes transmit the virus to humans (Thoisy et al., 2020). A geographic expansion of the disease into the south-eastern region has been observed (Brazilian Ministry of Health, 2017), which may be linked to vector/host dispersion through the landscape from endemic areas.

Non-human primates are the main hosts of YFV (Monath, 2001), with the howler monkey (Alouatta spp.) being the species most susceptible to infection (de Azevedo Fernandes et al., 2021). Its death can indicate the circulation of the virus in specific geographic locations and the immediate need for human population vaccination if the virus is confirmed through proper diagnostics. YFV notification in non-human primates is called epizootic events (i.e. a disease event in an animal population; Almeida et al., 2014), and in Brazil, it requires immediate report to the health authorities (Brazil, 2006).

The genus Haemagogus and Sabethes include arboreal species that are vectors of YFV. Their presence and abundance are positively related to the presence of food resources (Pittendrigh, 1950), high temperatures and low humidity (Pinto et al., 2009). These conditions are most common in landscapes with low amounts of forest cover (Alencar et al., 2018), and close to cattle raising areas. Females become infected either through a bite in a host during viraemia or through transovarial transmission (Mondet et al., 2002).

There is a consensus among experts that non-human primates are not responsible for the YFV spread (Possas et al., 2018). Their home ranges and daily travelling range in forests are restricted, and they normally do not use the ground and deforested areas to migrate from one place to another (Jung et al., 2015), so they are unlikely to carry the infection over long distances (Possas et al., 2018). Therefore, it is reasonable to assume that YFV expansion is being mediated by the dispersal of the infected vectors (Causey et al., 1950).

The landscape structure can affect the spread of the YFV in different ways. Roads, for example, may facilitate the spread of the disease by bringing people closer to source areas of infected vectors (i.e. forest areas; Melo et al., 2007), or by transporting infected vectors and hosts through motor vehicles (Gubler, 1989) or wind tunnels (Almeida et al., 2019), or even viraemic humans to distant regions (Vasconcelos et al., 2004). Furthermore, roads are considered agents of land use change and can facilitate the process of landscape

4. Synthesis and applications. Through landscape connectivity analyses, we provided here the first evidence that highly fragmented landscapes with a wide road network and large densities of forest edges facilitate yellow fever virus propagation, and that the maintenance of large blocks of forest can help to inhibit this spread. These results can contribute to guide forest restoration and landscape management actions in order to amplify health benefits related to restoration projects, in addition to their benefits for biodiversity conservation and climate change mitigation.

**KEYWORDS**
disease ecology, landscape ecology, landscape management, nature-based solutions, virus movement, yellow fever dispersion, zoonoses
fragmentation (Laurance et al., 2002), which can isolate the populations of vectors and hosts (Hess, 1996). In this case, virus persistence would depend on the hosts/vector’s dispersion capacity, and on the landscape’s ability to allow this dispersion, which is strongly linked to the configuration of the remaining habitat.

Usually, as fragmentation intensifies, the amount of forest decreases and the proportion of forest edge in a landscape increases (Cagnolo et al., 2009). Forest edge habitats have the potential to increase mosquito larval growth, survivorship and adult densities (Vittor et al., 2009) because they are well lit and warmer than rainforest interiors (Magnago et al., 2015). These characteristics increase the abundance of YFV vectors (Hendy et al., 2020; Santos et al., 2020), that have great capacity to adapt to secondary forests and disturbed areas (Camargo-Neves et al., 2005), and make forest edges the perfect environment for the persistence of YFV. Beyond that, forest edges allow a wider interface with humans, increasing the virus transmission risk. Furthermore, edge habitats and roads are significantly more susceptible to strong winds (Magnago et al., 2015), an aspect that may be important for the dispersal of infected mosquitoes, and consequently of the virus (Almeida et al., 2019).

A better understanding of the movement of YFV in fragmented landscapes and how these movements may be influenced by landscape structure can be used to better understand virus spatial dynamics and to inform health authorities about the potential risk areas before the outbreak, allowing mitigation measures and vaccination campaigns to be organized. Here, we modelled the dispersion of the YFV in a fragmented landscape of the São Paulo state, the most populated in Brazil. Our modelling procedure was based on empirical data, and network and circuit theory. The objectives of this study were to estimate virus dispersion in the different seasons of the year and to determine which landscape features facilitate or prevent virus connectivity. We tested different hypotheses about how YFV can disperse through the landscape. Our main predictions are that forest edge areas and roads favour the dispersion of the virus because roads lead to fragmented landscapes with high amounts of forest edges, promoting ecological conditions (i.e., high temperatures, low humidity and high abundance of YFV vectors) that favour the spread of the virus.

2 | MATERIALS AND METHODS

2.1 | Study area

The study area encompasses 30 municipalities of the state of São Paulo, located in the Atlantic Forest biome (Figure 1). These municipalities were chosen because they have a reliable YFV surveillance system, and consequently a trustworthy epizootic data collection once surveillance was systematic and intensive (with active and passive surveillance for non-human primates’ death and illness) along the 2016–2018 period outbreak (see below). More details about the study area can be found in Appendix S1.

2.2 | Epizootic data

From October 2016 to January 2018, 551 dead monkeys were reported by local authorities and had biological material collected in the studied area, of which 222 were negative (151 Callithrix, 30 Alouatta, 5 Sapajus and 36 no species identification), and 328 positives for YFV (243 Alouatta, 20 Callithrix, 9 Callicebus, 2 Sapajus and 54 no species identification). For modelling purposes, we considered only the positive events of Alouatta species (n = 243), because they are the best spatial and temporal indicator of virus circulation (de Azevedo-Fernandes et al., 2021), once they have a restricted territory and a high predictive positive value (are highly susceptible to the virus), dying a few days after exposure to virus and becoming infectious. Therefore, Alouatta events represent with high precision the presence of the virus in time and space. The other genera (Callithrix, Sapajus, Callicebus) can have some degree of resistance to infection and may develop immunity to the virus (de Azevedo Fernandes et al., 2021), with their role in the transmission cycle not yet being fully elucidated (Cunha et al., 2019). In fact, infected individuals died of other causes, and consequently, the death event does not accurately indicate time and space of virus occurrence. For the negative samples, we considered that all genera present a high negative predictive value (de Azevedo Fernandes et al., 2021; Cunha et al., 2019) and can truly represent the absence of the virus in a landscape, which allowed us to use the negative data from all species in the modelling process (n = 186). Furthermore, given the high number of positive samples of the genus Alouatta (75% of the total data), we believe that there is no bias in making this cut, as the other species represent only 9.45% of the total data. Points with unidentified species were excluded from both negative and positive datasets.

Positivity for YFV was determined by molecular assays in the Instituto Adolfo Lutz (the reference laboratory) through standard tests as determined by the yellow fever guide of the Brazilian Ministry of Health (Brazilian Ministry of Health, 2017), and the result was validated by the São Paulo State Epidemiological Surveillance Center (CVE).

2.3 | Spatio-temporal network analysis

A directed network is a widely used method to evaluate interactions between species (Godfrey, 2013) and was adopted to represent the pathways for pathogen transmission and dispersion. In that way, each primate sampled for YFV and included in our dataset was represented as a node, while a link between any two positive nodes represented pathogen transmission, and potential virus dispersion between them. For the analysis, we kept the nodes with no links (i.e. dead monkeys negative for YFV), because they can bring important information about landscape structure aspects that can restrain virus dispersion. To establish the network connections, the dates and spatial locations of each positive epizootic event were considered, given that dispersion is spatial and temporal dependent (Hill et al., 2020). We assumed that dispersion occurs from the
closest point in space and in time. Therefore, our network was established considering the closest node in time; however, if there was more than one option for route considering only time, we considered the closest node in space. If the closest node in time was too far in space to allow dispersal (>15 km/day; the maximum distance observed for a YFV vector is 11.5 km; Causey et al., 1950), we chose the closest node in both space and time at the same time. In this case, the network is established with the closest point in time that has an estimated dispersion of up to 15 km/day. It is worth mentioning that a node can act as a source for several dispersion events.

From the spatial-temporal network and considering only nodes with links, we calculated the node degree (i.e. number of contacts from the nodes), the characteristic path length (average path length between all pairs of nodes), the diameter (i.e. the largest distance in the network; Wasserman & Faust, 1994) and the average and maximum distance travelled by the virus. Additionally, we measured the average, minimal and maximum speed (km/day) of virus dispersion considering the different months and seasons of the year (Minor & Urban, 2007).

2.4 | Land use/land cover map

A land use/land cover map (LULC) with 5-m resolution, obtained from the Brazilian Sustainable Development Foundation (http://geo.fbds.org.br/) and generated using RapidEye images, from 2013 (1:10.000 scale), was used as the basis for modelling and analyses. This mapping was chosen because there was almost no change in the landscape from 2013 to the date of the yellow fever outbreak and data collection (2016–2018; IF, 2020), and because it is the only available mapping with a high resolution, which is essential for the type of analysis performed.

2.5 | Landscape resistance parameterization

Resistance provides a quantitative estimate of how environmental parameters affect animal movement (Zeller et al., 2012), recognizing that this movement is influenced by landscape features (Ferreras, 2001). Its estimation is made through environmental variables parameterization across a ‘resistance’ to movement continuum.
(i.e. resistance matrix). In this matrix, low resistance values denote ease of movement or high permeability, and high resistance values denote restricted or absolute barrier to movement (Zeller et al., 2012). These resistance matrices form the basis for modelling connectivity and reflect the influence of landscape features on the movement of YFV.

The resistance parameters used in this study were selected with the intention of capturing a wide range of hypotheses on how different land use and land cover types can potentially affect the movement of the YFV through the landscape. They were chosen through the opinion of five experts who work with YFV (i.e. an approach used in 80% of the studies that parameterized resistance surfaces; Zeller et al., 2012), and on available studies, and considered that dispersion was facilitated in forest areas, especially in forest edge areas, and in certain land uses (see Table S1). In addition, we also created matrices with different values of what was suggested by the experts and considered an opposite hypothesis (i.e. contrary hypotheses), in order to provide other options than just those provided by the experts’ opinions, once they can provide suboptimal parameterization (Seoane et al., 2005).

For this, we developed 49 different scenarios in the form of resistance matrices, based on characteristics of landscape composition, forest edge and roads effects. In each resistance matrix, the parameters were set according to a hypothesis about virus dispersion. For each pixel in the final map, a resistance value was assigned to create a contrast between forested and non-forested areas, edges and non-edge areas, roads and the six different land uses. More permeable areas had a low resistance equal to 1, while less permeable areas had a high resistance of 500. Below we described how each group (landscape composition, forest edge and road effects) of the resistance matrix was set, and in Table S1, we describe in more details all the 49 resistance matrices, the hypothesis behind each one and the parameterization values that were specified.

1. Landscape composition resistance matrices were based on the study area map, which is composed of six land use/land cover classes—water, agriculture, urban, native forest, natural non-forest and forestry areas (exotic spp. plantations). To set these resistance matrices, first we create a simple hypothesis in which virus dispersion was facilitated in forest areas, regardless of the matrix type. Therefore, the mapping was reclassified into just two classes—forest and non-forest and it was considered that dispersion was facilitated in forest areas. In a second hypothesis, all six land use/land cover classes were considered, with dispersion being dependent on the land use type. In this case, dispersion was facilitated in forest and forestry areas and hampered in water, agricultural and urban areas.

2. Forest edge resistance matrices were based on reclassification of land use/land cover maps. The reclassification processes created maps with two and six land cover classes, which were used to generate maps containing different forest edge depth effects (40, 100, 200 and 400 m). For the map created with two classes, the hypothesis postulated that virus dispersion is facilitated in forest edges and is intermediate in forest core areas. For the map created with six land cover classes, the hypothesis was that virus dispersion was facilitated in forest edges adjacent to specific land uses (agricultural and water), was intermediate in core areas of forestry and natural non-forest areas and was low in core areas of urban areas, for example. We calculated the forest edges and the core areas, defining forest edges as the interface between forest habitat and different land cover types, in both directions, into forest fragments and adjacent land cover (Harper et al., 2005). A detailed description of all matrices can be found in Supporting Information.

3. Road effect resistance matrices were created using DNIT (National Department of Transport Infrastructure—https://ide.geobases.es.gov.br/layeres/geonode:dnit_rodovias_es_31984/metadata_detail) and DERSA (Road Development S.A.) shapefiles, which contain the main paved roads of the state of São Paulo. We incorporated a 100-m buffer around each road, to capture the effects of the road degradation over the environment and hypothesized that virus dispersion was facilitated in these areas, irrespective of the landscape structure adjacent, because roads can transport yellow fever vectors through wind tunnels or through vehicles (Almeida et al., 2019; Causey et al., 1950; Moreno & Barata, 2012), facilitating virus spread. In a second step, this resistance was combined with the landscape composition resistances, and with the forest edge effects resistance to create more complex hypotheses, in which dispersion was facilitated in roads adjacent to specific land uses (i.e. forest edge areas, forest areas, forestry areas and natural non-forest areas). Table S1 contains all the 49 resistance matrices, the hypothesis behind each one and the parameterization values that were specified.

2.6 Dispersion analysis

To estimate YFV potential movement, we used a new approach borrowed from electrical circuit theory (McRae & Beier, 2007), where the ecological flow can be analogized to the electric currents and can predict the movement patterns across complex landscapes (McRae et al., 2013). This method has been used in a variety of ecological contexts and it is proven to be effective to model the resistance of the landscape to species movement (Lawler et al., 2013; McRae & Beier, 2007). For this, we used Circuitscape 4.0 (McRae et al., 2013), which describes the resistance of the virus in moving through different landscape features, and calculate connectivity (i.e. resistance values) between all pairs of nodes using the 49 resistance maps developed (Table S1) and the pairwise modelling mode.

2.7 Statistical analysis

Yellow fever virus dispersion distance and the number of days to disperse between each season of the year were compared with a t
In order to select the best resistance matrix, we compared 50 generalized linear models (GLM), with a binomial family error distribution (logit link function; 'base' package in R v. 2.1.0.1; R Core Team, 2008). For this, we used the virus presence in the arrival node as the response variable and the average resistance value for each pair of nodes (arrival and departure nodes) as the predictor variable. We also built a null model, where the virus dispersion varies at random and the resistance matrix does not influence dispersion (49 matrix resistances and one null model, forming 50 models). For model selection, we conducted a maximum likelihood model selection procedure, considering the second-order Akaike information criterion (AIC; Burnham & Anderson, 2002), choosing the model with the lower AIC.

We also calculated the Akaike information criteria weight (wAIC), which expresses the relative contribution of the model to explaining the observed pattern (Burnham & Anderson, 2002) and the evidence ratio of each model. To the best model selected, we performed a leave one out cross-validation and fitted the model using polynomial forms (first to fifth order), comparing it with linear relationships. As there was no difference between them, we kept the simplest fit (linear).

From the best supported model, we tested residuals for spatial autocorrelation by calculating Moran’s I with the inverse distance matrix and Moran’s I correlogram (Figure S1) and temporal autocorrelation with the autocorrelation function (acf). As the results showed both spatial (Moran’s I, \( p = 0.0001 \); Figure S1) and temporal autocorrelation (Figure S2), we modelled YFV dispersion probability using a spatiotemporal Bayesian model (glmmfields package) with non-informative priors, four MCMC chains, 5,000 burn-in values and 2,000 iterations. We checked model convergence through Rhat values (Table S2) and estimated the average predicted value for new observations.

### RESULTS

The spatiotemporal network for the YFV showed that for the 15 months analysed, in the selected 30 municipalities of São Paulo state, the total number of movements (number of links or virus dispersion) observed were 223. In most of the times (55.22%), one positive node resulted in only one link. In other words, from this positive node, the virus disperses to just one more node. In 28.35% of the movements, one positive node gave rise to two other positive nodes (two links), while in 8.2%, it originated three new different positive nodes; in 6.7%, four new positive nodes; and only in 1.5% of the cases generated five or more links.

The average movement distance observed in the network was 1.42 km/day, with the largest movement being 6.9 km/day and the shortest movement being 0.1 km/day. Dispersal movements that occurred up to 1 km per day were the majority, representing 71% of the total movements. Dispersal events that involved movements between 1 and 1.5 km/day comprised 13.2% of the total, while movements between 1.5 and 2.5 km/day were 7.1% of the total. The
number of movements with distances greater than 3 km/day was rare and included only five events (~2.3%; Figure 2a).

Considering the different seasons of the year, YFV dispersion was, on average, higher in summer (N = 9 dispersion events), with longer dispersal distance values per day than the other seasons (1.2 km/day). Fall (N = 65 events) and spring (N = 140 events) had on average, similar dispersal movements, with 0.78 and 0.85 km/day, respectively, while winter (N = 16 events) had the lowest dispersal rate (0.22 km/day, p < 0.0001; Figure 2b), and the smallest confidence interval.

Additionally, most dispersal movements (~72.6%) occurred within a week after the arrival of the source node, with 10.43% of the movements dispersing after 1 day, 22.13% after 2 days and 33.8% after 3 days (Figure 3a). Dispersions after the tenth day were less frequent and represented ~17% of total movements (Figure 2c; Figure 3a). Also, dispersions in winter generally occurred after 20 days of the arrival of the source node, being longer than in the other seasons which occurred up to 1 week (p < 0.0001; Figure 2d). However, this result should be viewed with caution given the large confidence interval that exists for winter data indicating that dispersions within a few days also occur.

Nodes with flow, which represented virus movement between two points in the landscape, occurred in smaller values of resistance than those obtained for nodes without flow, which means that YFV dispersion is highly modulated by landscape permeability. Our model selection analysis presents only one best supported model (82% weight of evidence) linking YFV movement with parameters of landscape composition, 100-m forest edge with all landscape units and road effects (Figure 3b; Table 1). The spatial-temporal model shows that the average probability of YFV dispersion across the studied landscape was 0.557. This means that YFV has a 55% probability of moving around the analysed landscape. Table S2 shows the parameters derived from this model. Our cross-validation estimate for the best model showed that the test error is approximately 0.1872, while the mean absolute error (MAE) is 0.38 ($r^2 = 0.25$, RMSE = 0.43).

According to this model, virus dispersion is mediated mainly by roads adjacent to forest areas (i.e. forest roads) and secondly by roads adjacent to forestry edges and shrub areas. Forest edges, that go up to 100 m from the edge, were also important for virus movement, particularly when edges are in interface with agricultural, water, forestry and shrub areas. Core areas of urban, water, agricultural and forestry land uses offered the highest resistance values for virus movement.

**Figure 3** Yellow fever virus (YFV) spatiotemporal network with links classified by the number of days between each dispersion event in (a) the land use/land cover map and in (b) the resistance matrix selected. Only nodes with links (i.e. positive for YFV) are represented in this figure.
### TABLE 1
Results of the model selection of the best supported landscape resistance matrix to explain the yellow fever virus (YFV) dispersion in the state of São Paulo. The response variable was the presence of the YFV in the arrival node. The number of the model is equivalent as the identification number of their respective resistance matrix in Table S1.

<table>
<thead>
<tr>
<th>Model no.</th>
<th>Predictor variable</th>
<th>AIC</th>
<th>∆AIC</th>
<th>df</th>
<th>Weight</th>
<th>Evidence ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>M39</td>
<td>Roads and LULC 100 m edges expert</td>
<td>120.2</td>
<td>0</td>
<td>2</td>
<td>0.8285</td>
<td>1</td>
</tr>
<tr>
<td>M38</td>
<td>Roads and LULC 40 m edges expert</td>
<td>125</td>
<td>4.8</td>
<td>2</td>
<td>0.0749</td>
<td>11.06</td>
</tr>
<tr>
<td>M1</td>
<td>Forest/non-forest</td>
<td>125.3</td>
<td>5.1</td>
<td>2</td>
<td>0.0643</td>
<td>12.88</td>
</tr>
<tr>
<td>M14</td>
<td>Land use/land cover</td>
<td>128</td>
<td>7.8</td>
<td>2</td>
<td>0.0166</td>
<td>49.9</td>
</tr>
<tr>
<td>M10</td>
<td>Forest/non-forest and 40 m edges inverse</td>
<td>130.6</td>
<td>10.4</td>
<td>2</td>
<td>0.0045</td>
<td>184.11</td>
</tr>
<tr>
<td>M29</td>
<td>Roads and Forest/non-forest 40 m edges expert</td>
<td>131.1</td>
<td>10.9</td>
<td>2</td>
<td>0.0035</td>
<td>236.71</td>
</tr>
<tr>
<td>M23</td>
<td>LULC 40 m edges with inverse values</td>
<td>131.3</td>
<td>11.1</td>
<td>2</td>
<td>0.0031</td>
<td>267.25</td>
</tr>
<tr>
<td>M37</td>
<td>Roads and land use/land cover map</td>
<td>132.1</td>
<td>11.9</td>
<td>2</td>
<td>0.0022</td>
<td>376.59</td>
</tr>
<tr>
<td>M16</td>
<td>LULC 100 m edges expert</td>
<td>133</td>
<td>12.8</td>
<td>2</td>
<td>0.0014</td>
<td>591.78</td>
</tr>
<tr>
<td>M42</td>
<td>Roads and LULC 40 m edges fixed values</td>
<td>136.6</td>
<td>16.4</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M46</td>
<td>Roads and LULC with 40 m edges inverse</td>
<td>137.2</td>
<td>17</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M12</td>
<td>Forest/non-forest and 200 m edges inverse</td>
<td>138.2</td>
<td>18</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M6</td>
<td>Forest/non-forest and 40 m edges different values</td>
<td>138.4</td>
<td>18.2</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M2</td>
<td>Forest/non-forest and 40 m edges expert</td>
<td>138.4</td>
<td>18.3</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M24</td>
<td>LULC 100 m edges with inverse values</td>
<td>138.9</td>
<td>18.7</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M15</td>
<td>LULC 40 m edges expert</td>
<td>139.1</td>
<td>19</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M11</td>
<td>Forest/non-forest and 100 m edges inverse</td>
<td>139.5</td>
<td>19.4</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M19</td>
<td>LULC 40 m edges with same value</td>
<td>139.9</td>
<td>19.7</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M28</td>
<td>Roads and forest/non-forest map</td>
<td>141.1</td>
<td>20.9</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M41</td>
<td>Roads and LULC 400 m edges expert</td>
<td>143.1</td>
<td>22.9</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M33</td>
<td>Roads and forest/non-forest 40 m edges inverted</td>
<td>144</td>
<td>23.8</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M30</td>
<td>Roads and forest/non-forest 100 m edges expert</td>
<td>145</td>
<td>24.9</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M40</td>
<td>Roads and LULC 200 m edges expert</td>
<td>145.5</td>
<td>25.3</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M31</td>
<td>Roads and forest/non-forest 200 m edges expert</td>
<td>145.6</td>
<td>25.5</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M43</td>
<td>Roads and LULC 100 m edges fixed values</td>
<td>148.1</td>
<td>27.9</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M34</td>
<td>Roads and forest/non-forest 100 m edges inverted</td>
<td>151.9</td>
<td>31.7</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M35</td>
<td>Roads and forest/non-forest 200 m edges inverted</td>
<td>152.1</td>
<td>32</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M18</td>
<td>LULC 400 m edges expert</td>
<td>152.7</td>
<td>32.6</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M8</td>
<td>Forest/non-forest and 200 m edges different values</td>
<td>153.1</td>
<td>32.9</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M4</td>
<td>Forest/non-forest and 200 m edges expert</td>
<td>153.1</td>
<td>32.9</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M7</td>
<td>Forest/non-forest and 100 m edges different values</td>
<td>153.4</td>
<td>33.2</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M3</td>
<td>Forest/non-forest and 100 m edges expert</td>
<td>153.4</td>
<td>33.2</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M20</td>
<td>LULC 100 m edges with same value</td>
<td>153.5</td>
<td>33.4</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M21</td>
<td>LULC 200 m edges with same value</td>
<td>156.2</td>
<td>36</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M47</td>
<td>Roads and land use/land cover with 100 m edges inverse</td>
<td>157.1</td>
<td>37</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M17</td>
<td>LULC 200 m edges expert</td>
<td>159.1</td>
<td>38.9</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M25</td>
<td>LULC 200 m edges with inverse values</td>
<td>160.2</td>
<td>40</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M9</td>
<td>Forest/non-forest and 400 m edges different values</td>
<td>162.1</td>
<td>41.9</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M5</td>
<td>Forest/non-forest and 400 m edges expert</td>
<td>162.1</td>
<td>41.9</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M22</td>
<td>LULC 400 m edges with same value</td>
<td>162.1</td>
<td>41.9</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M13</td>
<td>Forest/non-forest and 400 m edges inverse</td>
<td>163</td>
<td>42.8</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M26</td>
<td>LULC 400 m edges with inverse values</td>
<td>163</td>
<td>42.8</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
<tr>
<td>M49</td>
<td>Roads and land use/land cover with 400 m edges inverse</td>
<td>164.1</td>
<td>44</td>
<td>2</td>
<td>&lt;0.001</td>
<td>--</td>
</tr>
</tbody>
</table>

(Continues)
movement, while forest core areas had intermediate resistance, bringing also some resistance to virus movement.

Roads adjacent to urban and water environments had intermediate resistance to virus movement, while forest edges in proximity to urban areas also present resistance to movement, demonstrating that more important than the presence of a certain landscape feature, is the context in which it is inserted.

4 | DISCUSSION

We provide here novel insights relating landscape connectivity and yellow fever virus dispersion in the neotropics. Our results identified that the virus is moving on average 1.42 km/day, with the greatest movement observed being 6.9 km/day during summertime, but only 0.22 km/day during the wintertime. Most of the movements occurred in distances up to 1 km and within a period of 7 days of the arrival of the source node. A key finding of our study is that roads adjacent to forest areas and 100-m forest edges in interface with agricultural areas are facilitating virus movement, while the presence of urban, agriculture, water (i.e. rivers and lakes), forestry and forest core areas represent a barrier to virus dispersal. These results can be used to guide restoration and landscape management actions to ensure or amplify their health benefits.

4.1 | Landscape connectivity effects on yellow fever virus dispersion

A fundamental prediction in disease ecology is that pathogen spread depends on the dispersal success of infected hosts or vectors (Altizer et al., 2006). Here, we are considering that virus spread is being mediated by the movement of infected vectors, with our results showing a dispersion of 1.42 km/day. This is consistent with *Haemagogus leucocelaenus* and *Haemagogus janthinomys* dispersal, once they can move up to 3 km per day (Causey et al., 1950; Goméz et al., 2018) and supports previous studies focused on the virus phylogeographic analysis, that showed that virus lineages can spread at a mean rate of 1 km/day (Hill et al., 2020). In addition, the greatest dispersal movement observed in this study (6.9 km/day) is consistent with the findings of other study that reported 6–11.5 km dispersion distance after 10 days for *Haemagogus* spp. mosquitoes (Causey et al., 1950). Because the authors were not able to identify the dispersal distance per day, the movement registered may have occurred in long dispersions over a short period of time, facilitated by wind. Here, we hypothesize that wind may be an important factor for infected mosquito dispersion. There are several records of mosquitoes being carried considerable distances, as much as 177 km, by prevailing winds (Huestis et al., 2019; Sanogo et al., 2021).

Yellow fever virus dispersion is on average higher in summer, with a similar pattern in the fall and spring. Winter presented the lowest dispersal rate observed. However, despite moving over shorter distances in this season (with a narrow confidence interval), the virus can take as few as many days to disperse (as the large confidence interval shows in Figure 2d). This behaviour was expected and was observed before (Lapointe, 2008), once summer and rainy seasons increase the abundance of mosquitoes (Linthicum et al., 1999; Muir, 1988), a factor that can influence their dispersal (Watts et al., 2018). Moreover, as poikilotherms, the insect internal temperature varies and depends on the temperature of its surrounding environment. With this, each single task performed (i.e. flying, feeding, etc.) will only happen within a range of temperatures, out of which they become torpid and unable to move (Huëy & Stevenson, 1979), explaining lower dispersal rates in the wintertime.

Evidence of landscape-mediating disease spread has been shown previously for other diseases (Smith et al., 2002; Watts, 2018). According to our results, the YFV dispersion is mediated primarily by roads adjacent to forest areas, and secondarily along 100-m forest edges in interface with agricultural areas, water, forestry and shrub areas. Road effects were associated before to a higher risk of contracting bovine and human rabies (De Andrade et al., 2016), increased prevalence of hantavirus (Medeiros et al., 2010), malaria (Bauch et al., 2015) and the occurrence of YFV (Almeida et al., 2019), giving support to our results.

Roads were associated with YFV movement only when adjacent to forest areas. This can be explained by the vector’s requirements, once they are arboreal (Neville et al., 1988), and need forest areas to survive. In addition, forest edges have microclimates that increase *Haemagogus* sp. abundance, once they have higher temperatures and lower humidity (Couto-Lima et al., 2020; Pinto et al., 2009). With
this, YFV vectors are found in larger abundances in forest edge habitats (Wilk-da-Silva et al., 2020). Roads can also taper and increase winds, forming currents, which increases mosquito’s dispersal ability (Almeida et al., 2019; Causey et al., 1950; Moreno & Barata, 2012). In that way, roads may serve as dispersal corridors for YFV, facilitating disease spread, when adjacent to habitats that are favourable to mosquito vectors. Therefore, the creation of roads that cross forest areas, especially in protected areas, can contribute to the dissemination of zoonoses, besides all the environment degradation and other negative impacts (Laurance et al., 2009). When roads are adjacent to land uses that are not favourable to vector’s survival, as urban areas and water reservoirs, this effect does not happen once vectors are not present to disperse.

Forest edge areas, composed by 100-m width, in interface with agricultural areas, water, forestry and shrub areas, were the second most important attribute through which YFV virus is dispersing in terms of movement costs. Forest edges proved to be a favourable environment for other disease vectors before (Kantso et al., 2010; Richman et al., 2018), where they were found in higher abundances when compared to other landscape features. However, when forest edges are immersed in urban areas, they represent a barrier for wild yellow fever virus movement. This happens because urban landscapes have high temperatures (Santamouris, 2007), absence of forest canopy and air pollution, which can negatively impact wild YFV vector species. Therefore, wild YFV vector mosquitoes, which have forest habitats, do not find an adequate environment to survive. Furthermore, when immersed in open habitats, as agricultural areas, forest edge habitats are susceptible to strong winds (Magnago et al., 2015), leading to an increased vector’s chance of dispersal. This point must be taken into consideration in landscape management programmes and especially in forest restoration projects. In rural areas immersed in agricultural matrices, restoration projects should need to consider this information, weighing up the advantages and disadvantages of creating corridors thinner than 100 m. In some situations, the expansion of existing forest areas should be preferred to the creation of new patches or corridors formed entirely of edges, especially if they are near roads.

Our results showed that forest core areas can restrain virus movement, in a similar result of what was found for the spread of malaria (Bauch et al., 2015). This may happen because forest core areas have more diverse communities of both mosquitoes and hosts. With this, vectors are probably less abundant and end up feeding on hosts less competent to the pathogen, diluting the effect of the most susceptible host, in a dilution effect (Keesing et al., 2006). Moreover, forest core areas are protected from wind by the shelter of dense vegetation and provide a much more abundance of sites for mosquito oviposition, what may decrease the need of long-distance flights. Therefore, mosquitoes present in these areas probably present little dispersion (Service, 1980). Therefore, large tracts of forests can hinder virus dispersion and decrease the transmission risk for this disease. Protected areas then can play a key role in stopping or reducing the spread of diseases and should be seen not only as elements for biodiversity conservation but also as key components of healthier landscapes.

Agricultural, forestry and urban areas can also restrain virus movement. Agricultural core areas are formed by open habitats and do not have the requirements that vectors need to survive. Forestry areas, despite being formed by trees, are not suitable habitats for vectors (i.e. do not have hollows—essential for mosquito breeding). Thus, it is an environment that can be used for shelter when close to other suitable uses (as forest edges) but can restrain virus movement when dominant in the landscape. However, some land uses, especially urban or residential areas, are ideal environments for other vectors, such as the Aedes aegypti, and the increase in these areas can increase the spread of the diseases transmitted by it, including allowing the resurgence of urban yellow fever.

Our results on how YFV disperses across the landscape of São Paulo state were made with the support of expert opinion. Although we tried to deal with possible biases of this approach by creating resistance matrices with different values and hypotheses contrary to what was proposed by them, it is essential that studies, especially with vectors, be carried out in the field for a validation of the results. Furthermore, the use of different number of species for the positive and negative points may bring some bias to our results, underestimating the presence and overestimating the absence of the virus. This is a point that should be viewed with caution. However, little is known about YFV and the pathogenesis in neotropical non-human primates (Silva et al., 2020), once they have different degrees of sensitivity to YFV (Holzmann et al., 2010). Further investigation of non-human primate sensitivity to YFV, the duration of the immune response and yellow fever pathogenesis, can contribute to elucidate the potential for vector infection and help to improve the YFV surveillance and control programmes.

4.2 | Conservation, landscape management and health implications

Our results allowed us not only to identify which attribute of the landscape facilitates virus dispersion, what can help to design preventive efforts to mitigate human risk, but also to outline landscape management strategies to prevent the virus from spreading and guarantee multiple benefits. It is important to decrease the landscape connectivity for YFV vectors towards diminishing YFV dispersion. In general, landscapes with high amounts of forest edges formed by several forest fragments in contact with agricultural, forestry and shrub areas and with the presence of forest roads increase virus spread, while less fragmented landscapes formed by large blocks of forests with less contact with other land uses decrease virus spread (Figure 4).

In that way, landscape management actions should focus in avoiding the creation of new forest edges. It is also important to avoid the allocation of roads close or in the middle of forest areas, especially cutting protected areas. With this respect, restoration practices should be performed to increase already existing forest fragments, or to create more aggregated forest fragments, and to avoid the creation of forest edges smaller than 100 m of width. If
the creation of forest corridors cannot be avoided, they should have a width greater than 150 m, and have, preferably, a few lines of forestry around them. These strategies may prevent the spread of YFV. Thinking only about controlling YFV spread, agricultural and forestry activities should not be interspersed with forested areas, to decrease the amount of forest edge in interface with these land uses. It is also important to avoid the creation of small, isolated fragments in the middle of agricultural areas, which can act as stepping stones for virus movement.

However, these landscape designs need to be balanced considering that the proximity of forests to agriculture can favour the provision of other ecosystem services. A landscape that can be considered ‘ideal’ to promote the regulation of YFV, can end up being negative for ecosystem services that require greater interface between agricultural and forest areas, as pollination (González-Chaves et al., 2020) or pest control (Aristizábal & Metzger, 2019). This shows that there are trade-offs, and that a landscape considered ideal for providing one ecosystem service can be bad for another service. However, once forest core areas proved to be important to stop the virus movement, protected areas, including private reserves, can act as important barriers to the spread of the virus. In that way, these areas can be important not only to biodiversity conservation (Metzger et al., 2019) but also to create healthier landscapes for humans (Prist et al., 2021). Finally, understanding the virus movement through the landscape can enhance our ability to deal with the trade-offs and help monitoring and preventing outbreaks, especially in landscape configuration that may facilitate virus YFV dispersion but are designed for other essential ecosystem services (water provision, pollination, pest control).

**ACKNOWLEDGEMENTS**

This study was developed within the ‘Interface project’, supported by the São Paulo research foundation FAPESP (n. 2013/23457-6). PP was funded by FAPESP grants 2017/11666-0 and 2018/23364-1. CHC was funded by the FAPESP grant 2018/09125-4. LFC was supported by CNPq (grant no. 307333/2013-2), FAPESP (grant no. 2015/22308-2) and NAP-PRP-USP. JP was funded by the National Council for Scientific and Technological Development (CNPq, 307934/2011-0). MAMS was funded by FAPESP (grant no. 2018/14389-0).

**CONFLICT OF INTEREST**

The authors declare that no conflict of interest.

**AUTHORS’ CONTRIBUTIONS**

P.R.P., L.R.T., A.P., L.F.M., R.P.d.S., J.R.R., M.A.M.S. and J.P.M. conceived and designed the research; M.P.-C., M.P., J.T.d.D., L.D.C.S., T.L.D., R.M.F.S. and R.P.d.S. collected non-primate data and performed virus prospection analyses; P.R.P., R.L.M., C.H.C. and L.d.F.C. performed data analysis; P.R.P. wrote the manuscript; all authors contributed critically to the drafts and gave final approval for publication.

**DATA AVAILABILITY STATEMENT**


**ORCID**

Paula Ribeiro Prist https://orcid.org/0000-0003-2809-0434
Leandro Reverberi Tambosi https://orcid.org/0000-0001-5486-7310
Renata de Lara Muylaert https://orcid.org/0000-0002-6466-6210
Jonathan Roger Rhodes https://orcid.org/0000-0001-6746-7412
Juliana Telles de Deus https://orcid.org/0000-0002-1617-7340
Jean Paul Metzger https://orcid.org/0000-0002-0087-5240


**SUPPORTING INFORMATION**

Additional supporting information may be found in the online version of the article at the publisher’s website.