FOREST ENTOMOLOGY
Dispersal of Invasive Forest Insects via Recreational Firewood: A Quantitative Analysis
FRANK H. KOCH,1,2 DENYS YEMSHANOV,3 ROGER D. MAGAREY,4 AND WILLIAM D. SMITH1
J. Econ. Entomol. 105(2): 438–450 (2012); DOI: http://dx.doi.org/10.1603/EC11270

ABSTRACT Recreational travel is a recognized vector for the spread of invasive species in North America. However, there has been little quantitative analysis of the risks posed by such travel and the associated transport of firewood. In this study, we analyzed the risk of forest insect spread with firewood and estimated related dispersal parameters for application in geographically explicit invasion models. Our primary data source was the U.S. National Recreation Reservation Service database, which records camper reservations at >2,500 locations nationwide. For >7 million individual reservations made between 2004 and 2009 (including visits from Canada), we calculated the distance between visitor home address and campground location. We constructed an empirical dispersal kernel (i.e., the probability distribution of the travel distances) from these “origin-destination” data, and then fitted the data with various theoretical distributions. We found the data to be strongly leptokurtic (fat-tailed) and fairly well fit by the unbounded Johnson and lognormal distributions. Most campers (∼53%) traveled <100 km, but ∼10% traveled >500 km (and as far as 5,500 km). Additionally, we examined the impact of geographic region, specific destinations (major national parks), and specific origin locations (major cities) on the shape of the dispersal kernel, and found that mixture distributions (i.e., theoretical distribution functions composed of multiple univariate distributions) may fit better in some circumstances. Although only a limited amount of all transported firewood is likely to be infested by forest insects, this still represents a considerable increase in dispersal potential beyond the insects’ natural spread capabilities.

KEY WORDS biological invasion, firewood, invasive forest pest, human-mediated dispersal, long-distance dispersal

The potential for accidental, long-distance transport of invasive insects and pathogens in untreated firewood has become a topic of considerable concern in North America (Haack et al. 2010, Tobin et al. 2010). The issue has been the subject of national-scale public awareness campaigns in both the United States and Canada (e.g., see Canadian Food Inspection Agency [CFIA] 2011, The Nature Conservancy 2011), as well as similar campaigns by individual states and provinces. Recently, the USDA Animal and Plant Health Inspection Service (APHIS) released a comprehensive risk assessment of the movement of firewood in the United States (USDA-APHIS 2011b), and also requested comments on recommendations issued by a National Firewood Task Force regarding regulatory measures, voluntary activities, and outreach efforts at the state and national levels (USDA-APHIS 2010). Currently, a majority of U.S. states have imposed restrictions on firewood movement, in some cases enforced with fines and/or other penalties for violations (The Nature Conservancy 2011). In Canada, any person who moves firewood out of an area regulated for a quarantine pest without approval from the Canadian Food Inspection Agency is subject to fines and/or prosecution (CFIA 2011).

In the United States, discussion about the risks associated with firewood movement essentially began in August 1996, when the Asian longhorned beetle, Anoplophora glabripennis (Motschulsky) (Coleoptera: Cerambycidae), was discovered in the Brooklyn borough of New York. Shortly thereafter (September 1996), a second infestation was found in Amityville, NY, ∼50 km east of the first detection (Haack et al. 1997). Although the Brooklyn infestation probably originated through international trade (i.e., from infested wood packing materials associated with imported cargo), it is thought that A. glabripennis may have been introduced to Amityville when infested tree sections were transported from Brooklyn for sale as firewood. In fact, fear of the likelihood of such a scenario is what prompted surveyors to investigate relatively distant areas such as Amityville after the
initial dispersal potential of *A. glabripennis* is relatively low (Smith et al. 2001, 2004), human activities including firewood transport are also assumed to have been responsible for the insect’s spread to other parts of the New York metropolitan area, and likewise, for its spread in Chicago (IL), Worcester (MA), Toronto (ON), and Bethel (OH), a community ~50 km southeast of Cincinnati (USDA-APHIS 2011a).

The *A. glabripennis* example notwithstanding, a primary impetus for the current focus on firewood has been the invasion of eastern North America by the emerald ash borer, *Agrilus planipennis* Fairmaire (Coleoptera: Buprestidae). Native to Asia, *A. planipennis* was first detected in 2002 near Detroit (MI) and neighboring Windsor (ON) (Haack et al. 2002). It has since been found in 14 other U.S. states as well as the province of Quebec (National Agricultural Pest Information System [NAPIS] 2011); movement of infested firewood has been strongly implicated in the insect’s range expansion (Petrie and Haack 2006). Because all North American ash (*Fraxinus* sp.) tree species are susceptible, the economic impact of *A. planipennis* is expected to be substantial. For the 10-yr period 2009–2019, the anticipated treatment, removal, and replacement costs for ash trees on developed land in U.S. communities (i.e., costs not including forestland impacts) have been estimated at greater than $10 billion (Kovacs et al. 2010).

Firewood is relevant as a vector for other high-profile forest pests, primarily wood-boring insects (especially Coleoptera species). For instance, the movement of firewood from Mexico is believed to have led to the establishment of *Agrilus prionanus* Chevrolat, a major pest of Western soapberry (*Sapindus saponaria* variety *drummondii*), in Texas (Haack 2006). Notably, long-distance transport of firewood for recreational purposes appears to be a fairly common occurrence. Haack et al. (2010) reported on a multi-year survey of drivers who surrendered firewood at Michigan’s Mackinac Bridge. (Transport of firewood across this bridge between the state’s Upper and Lower Peninsulas is prohibited under a *A. planipennis* quarantine.) According to the survey, 16% of vehicles surrendering firewood were from out of state, while 1% of vehicles were from Canada. Additionally, campground surveys in various parts of the United States indicate that 8–57% of campers bring their own firewood from home, frequently traveling distances of 160–320 km and crossing state lines (USDA-APHIS 2011b).

Whereas such findings are informative, they are insufficient to quantify the nationwide potential for long-distance dispersal of forest insects in recreational (i.e., camper-transported) firewood. In short, the combined results of these surveys do not fully describe the probability of firewood being transported over certain distances (i.e., how far, and how often, firewood-carrying campers can be expected to travel), particularly long distances. This latter limitation is significant: Long-distance dispersal events substantially increase the rate of spread of organisms, especially at larger (e.g., regional) spatial scales (Higgins et al. 2003, Nathan et al. 2003). Notably, for many invasive species, these instances of long-distance dispersal are almost exclusively associated with human transport activities (Hastings et al. 2005).

Predictive models of the spatial spread of invasive species can be informative for decision makers who must implement regulatory and control measures despite incomplete information and scarce resources (Neubert and Parker 2004). However, reliable estimates of the rate and extent of long-distance dispersal events are necessary to accurately model spread for most invaders (Suarez et al. 2001, Hastings et al. 2005, Pitt et al. 2009). Long-distance dispersal is especially important from a pest management perspective because it may lead to the formation of satellite populations, which (although they may greatly increase an invader’s overall rate of expansion) are more easily controlled than large populations within the main invasion front (Lockwood et al. 2007). Unfortunately, because long-distance dispersal events are rare, empirical data about them are difficult to obtain (Clark et al. 1995, Hovestadt et al. 2001, Brown and Hovmøller 2002, Higgins et al. 2003, Hastings et al. 2005), and in turn, model-based estimates of their likelihood tend to be highly uncertain (Nathan et al. 2003). The long-distance dispersal potential of firewood has been depicted in some spatially explicit invasion models, particularly for *A. planipennis* (BenDor and Metcalf 2006, BenDor et al. 2006, Muirhead et al. 2006, Prasad et al. 2010, Harwood et al. 2011), but only in a simplified (i.e., theoretical) fashion or based on limited empirical data. Typically, these and other invasion models incorporate the concept of stratified diffusion (Henugeveld 1989, Shigesada et al. 1995, Nathan et al. 2003, Lockwood et al. 2007), which acknowledges that species may spread by more than one process (e.g., natural as well as human-mediated dispersal). Subsequently, each process may be modeled with an individual probability distribution function (“kernel”) that depicts the likelihood of dispersal via that process over a range of distances during a specified time interval. Instead, some models use “mixed kernels” that integrate functions representing both short- and long-distance dispersal processes (see Higgins et al. 2003, Nathan et al. 2003). In either case, long-distance dispersal processes such as the movement of forest pests in firewood typically require application of leptokurtic (also called fat- or long-tailed) functions that are suited to representing the greater occurrence of rare events (Kot et al. 1996, Nathan et al. 2003). For instance, negative exponential or power law functions have been regularly applied to depict long-distance dispersal in theoretical models of biological invasions (Cannas et al. 2006), including sophisticated spatial simulations (e.g., Carrasco et al. 2010).

Given that a data-driven approach is preferable to other methods (e.g., based on expert judgment) for estimating dispersal probabilities (Quigley and Revie 2011), it would be beneficial if sufficient data were available as an empirical foundation for defining appropriate distance-dependent distribution functions to represent firewood dispersal potential in invasion.
modeling efforts (i.e., as one particular long-distance component of a stratified diffusion framework). Although data on firewood transport and utilization are lacking in this regard, an alternative is to undertake a more general approach by exploring the travel behavior of campers rather than their actual use of firewood. For this study, we had the opportunity to analyze an extensive (multi-year and national-scale) and geographically referenced database of campground visits in the United States. Our primary objective was to develop realistic, empirically based dispersal kernels for describing key patterns of camper travel under various scenarios (e.g., travel to specific popular destinations). In turn, by incorporating the limited available information about firewood usage, we also hoped to provide basic estimates of the potential contribution of recreational firewood movement to the spread of invasive forest insects in the United States.

Materials and Methods

We performed all statistical analyses in the R software environment (R Development Core Team 2011), in some cases using specific developer-contributed analytical packages (individually cited below). We used ArcGIS 9.3 software (Environmental Systems Research [ESRI] 2009a) to perform all geospatial analyses.

Campground Visitation Data. Our primary data source for this study was the National Recreation Reservation Service (NRRS). The NRRS handles reservations for campgrounds and other recreational facilities operated by the U.S. Army Corps of Engineers, the Bureau of Land Management, the Bureau of Reclamation, the USDA Forest Service, and the National Park Service. Members of the public (including individuals from outside the United States) can make reservations at these facilities through the NRRS online portal (http://www.recreation.gov), by telephone, or at specific field offices. Notably, the NRRS does not manage reservations for all federal recreational facilities, as a small percentage (e.g., facilities in certain national parks) have independent reservation systems. However, NRRS-managed facilities are well distributed across the United States.

We used NRRS reservation records spanning the period from January 2004 to September 2009; these data were previously procured from the NRRS by USDA APHIS-Plant Protection and Quarantine (PPQ). Each record included the following variables: visitor origin state and ZIP code, name and state of the destination campground, reservation date (month and year), the number of reservations, and the number of nights. The latter two variables represent summations of the raw data by NRRS (at the request of USDA APHIS-PPQ) in cases where records were identical with respect to the other variables (i.e., records with the same visitor origin, destination campground, and reservation date).

Before analysis, we filtered the data by deleting records associated with campers or NRRS facilities in Alaska, Hawaii, or Puerto Rico. We reasoned that it would be unlikely for a visitor to transport firewood between these locations and the conterminous United States. For the same reason, we also deleted records from international visitors, except for visitors from Canada; in their study of firewood surrendered at Michigan’s Mackinac Bridge, Haack et al. (2010) noted multiple firewood-carrying vehicles from three different Canadian provinces (Ontario, Alberta, and Newfoundland and Labrador). Finally, we also removed any records where the number of nights was zero; these records indicated reservations for day use rather than camping. After filtering (that eliminated >500,000 reservations), the data encompassed >7.2 million individual reservations made at 2,525 campgrounds and related recreational facilities.

The NRRS supplied geographic coordinates for these facilities, some of which USDA APHIS-PPQ manually corrected after performing a spatial review in a geographic information system. We also assigned geographic coordinates for each visitor ZIP code in the data set (or postal code for the Canadian visitor records) based on the centroids of the ZIP code polygons (ESRI 2009b, Natural Resources Canada 2010). Then, for each individual visitor reservation record in the data set, we calculated the Euclidean (i.e., straight line) distance between the visitor’s origin ZIP code and the destination campground.

We extracted six subsets from the full NRRS data set for further analysis. Our main objective in analyzing these subsets was to determine how the distribution of travel distances for a particular subgroup of campground visitors might differ from the distribution for all visitors. First, we separated the full NRRS data set into four geographic subregions according to the location of the destination campground cited in each record: northeastern, southern, midwestern, and western United States (Fig. 1). Second, to evaluate potential differences in the distance distribution for visitors traveling to specific popular destinations, we created a subset of all records of visits to campgrounds within 50 km of the 20 most visited national parks in the continental United States (National Park Service [NPS] 2011). Third, to examine potential distributional differences for campground visitors from large cities, we created a subset of all NRRS records origi-
in the 20 most populous U.S. urban areas (based on U.S. Census Bureau population estimates and urban area boundary delineations).

**Kernel Density Estimation.** For the full data set and each of the subsets, we developed an empirical estimate of the probability density function of the travel distance data using kernel density estimation. The kernel density estimate provided a basic setting for visually assessing the fit of various theoretical distribution functions to the data (see next section). For this task, kernel density estimation avoids the discontinuity of histograms that arises from binning the data, but has the potential disadvantage of smoothing out some of the fine-scale variation, particularly in long-tailed distributions (Silverman 1986). A kernel’s degree of smoothing depends on a parameter, $h$, known as the bandwidth; to determine $h$, we applied the commonly used Silverman’s rule of thumb:

$$h = 0.94A^{1/5}$$

where $A$ is the minimum of the standard deviation and the interquartile range divided by 1.34, and $n$ is the sample size (Silverman 1986, Sheather 2004).

**Theoretical Distribution Fitting.** For the full NRRS data set and each of the subsets, we fitted the distribution of travel distances using a number of univariate theoretical distribution functions, many of which have been used previously to depict dispersal kernels in spatial invasion models (Kot et al. 1996, Lockwood et al. 2007, Pitt et al. 2009). The distribution functions that we evaluated were the beta, Cauchy, exponential (i.e., negative exponential), gamma, lognormal, and (two-parameter) Weibull distributions, as well as the bounded and unbounded Johnson distributions (often denoted as the Johnson SB and SU distributions, respectively); notably, the lognormal distribution is also considered part of the Johnson distribution family (denoted in this context as Johnson SN). We fit each theoretical distribution function to the data via maximum likelihood estimation using the `fitdistrplus` package in R (Delignette-Muller et al. 2010). For each tested data set, we identified the best-fitting function as the one with the lowest value of Akaike’s information criterion, a commonly used metric for comparing the goodness-of-fit of models estimated by maximum likelihood (Akaike 1973).

**Mixture Distributions.** An alternative to using a univariate theoretical distribution function is to use a “mixed-kernel” approach (Nathan et al. 2003), which combine two or more functions that depict separate aspects of dispersal. This sort of approach is a logical extension of the aforementioned concept of stratified diffusion (i.e., where dispersal involves multiple processes); for instance, a mixture distribution applied as a dispersal kernel in a spatial model of an organism’s spread might integrate one theoretical distribution function to represent short-distance dispersal via one vector (e.g., the organism’s own movement capabilities) with a second function to depict rare long-distance dispersal via another vector, such as movement of the organism by wind or as a “hitchhiker” on other organisms (Higgins and Richardson 1999, Higgins et al. 2003, Nathan et al. 2003). In actuality, even when an organism’s dispersal can be reasonably characterized as single-vector, it is still likely to involve multiple individual processes, each associated with a particular subpopulation or subgroup of the organism (Higgins and Richardson 1999). We explored this premise with the NRRS data, positing that, at least in some cases, camper movement (representing one particular vector for dispersal of forest insects) might best be characterized by a mixture of component distributions corresponding to the behavior of certain social groups of campers (e.g., weekend campers vs. vacationers taking extended trips). Having identified the best-fitting univariate theoretical distribution functions (see previous section), we fit mixtures of these theoretical distributions to the national parks and urban areas subsets of the NRRS data. We did this using the `mixdist` package in R (Macdonald 2010), which contains tools for fitting mixture distribution models to grouped (i.e., histogram) data. Although a lack of data can make it difficult to fit mixtures with more than two components, given the large sample size ($N > 1.1$ million) of both the national parks and urban areas subsets, we felt justified in fitting them with three- and four-component mixture distributions. For each subset, we judged the success of the tested mixtures based on the chi-squared approximation to the likelihood ratio test as well as the standard errors of the component distributions’ parameters (Macdonald 2010).

**Results**

**Exploratory Geospatial Data Analysis.** Figure 2 shows a map of the links between visitor origin ZIP codes and destination campgrounds for the full NRRS data set. The map suggests some basic trends in camper travel behavior. First, the highest-volume links (i.e., links with >1,000 reservations, depicted in red) are all <250 km in length, and commonly much shorter (<100 km). In fact, most links with a moderate (>50) to high volume of reservations are similarly short in length. Although there are a few moderate-volume links that exceed 1,000 km, the longest links generally tend to be those associated with few (<25) reservations. Regardless, instances of cross-country travel by campground visitors are evident in the map, as are a small proportion of visits originating in Canada (most prominently from locations in Alberta, Ontario, and Quebec).

Figure 2 also depicts a number of distinct clusters of camper travel activity. For instance, CA contains prominent clusters of travel from its major coastal urban areas (i.e., San Diego–Los Angeles and San Francisco–Oakland) to campgrounds in the San Bernardino or Sierra Nevada Mountains. Similar clusters can be seen in other parts of the western United States (e.g., from the Denver, CO, area into the Rocky Mountains), but the region with the highest overall level of camper travel activity appears to be the southern United States, particularly in the area extending north from eastern Texas to Missouri. The activity level in
this region may be somewhat visually exaggerated because the destination campgrounds are more widely dispersed geographically than in the western United States. Nevertheless, close examination of the map reveals a consistent pattern of high-volume travel in this region, from both rural and urban locations to destination campgrounds situated near major lakes (or reservoirs) or National Forests.

**Summary Statistics.** For the continental United States (i.e., the full NRRS data set), as well as the southern, northeastern, and midwestern U.S. subregions, the median travel distance was <100 km (Table 1). The median travel distance in the western United States was somewhat higher (nearly 150 km). This may be because destination campgrounds generally tend to be further away from populated places in the West, although this result was probably also shaped by characteristics of the geospatial data; essentially, some ZIP code areas in the western United States are quite large, and because visitor travel distances are calculated from the centroids of ZIP code areas, these distances are likely longer on average for this region than for other parts of the United States. The median travel distances for the national parks and urban areas subsets (Table 1) were even higher than for the western U.S. subregion, suggesting that particular target destinations and places of origin do have at least a marginal effect on camper travel behavior.

For all of the tested data sets, the average travel distance was two to three times higher than the median distance, while the maximum travel distance was >3,000 km (and >5,500 km for the full data set and some subsets). These high average and maximum travel distance values relative to the median indicate

<table>
<thead>
<tr>
<th>Region/city category</th>
<th>N</th>
<th>Avg. distance (SD) (km)</th>
<th>Median distance (km)</th>
<th>Max. distance (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continental United States</td>
<td>7,220,563</td>
<td>235.8 (463.2)</td>
<td>92.6</td>
<td>5,565.9</td>
</tr>
<tr>
<td>Southern United States</td>
<td>3,140,537</td>
<td>175.0 (340.1)</td>
<td>67.7</td>
<td>5,565.9</td>
</tr>
<tr>
<td>Northeastern United States</td>
<td>358,584</td>
<td>180.4 (364.4)</td>
<td>85.3</td>
<td>4,531.6</td>
</tr>
<tr>
<td>Midwestern United States</td>
<td>1,338,091</td>
<td>170.4 (312.9)</td>
<td>78.4</td>
<td>3,279.5</td>
</tr>
<tr>
<td>Western United States</td>
<td>2,383,351</td>
<td>360.9 (631.7)</td>
<td>148.3</td>
<td>5,435.9</td>
</tr>
<tr>
<td>20 most visited national parks</td>
<td>1,144,999</td>
<td>487.7 (756.9)</td>
<td>213.4</td>
<td>5,435.9</td>
</tr>
<tr>
<td>20 most populous urban areas</td>
<td>1,117,898</td>
<td>370.7 (605.5)</td>
<td>171.6</td>
<td>4,538.7</td>
</tr>
</tbody>
</table>
that the tested data sets are positively skewed, as illustrated by the histogram for the full data set (Fig. 3). The histogram also shows the strongly leptokurtic nature of the data. For the full NRRS data set, ≈10% of visitors traveled >500 km, and ≈5% traveled >1,000 km.

**Theoretical Distribution Fitting.** The inset plot in Fig. 3, which is a close-up up of the full data histogram at distances <1,000 km, reveals something else significant: the peak density (i.e., the mode) of the distribution of travel distances for the full data set is at a distance of ≈50 km. This implies that theoretical distribution functions such as the (negative) exponential or inverse power law, where the probability density declines from a maximum at zero, are unlikely to provide a good fit to the data. Indeed, for the full data set and all of the subsets, the best-fitting theoretical distribution function, based on minimum Akaike's information criterion, was the unbounded Johnson (SU) distribution, followed closely by the lognormal distribution (i.e., another member of the Johnson distribution family). For the full data set, a plot of density as estimated by the Johnson SU distribution versus the kernel density estimate (Fig. 4A) suggests a good fit across the entire range of travel distances; the only apparent flaw is under-prediction by the Johnson SU at distances between 100 and 200 km (discernable in the inset plot). However, a plot of the logarithm of density versus distance (Fig. 4B), which more clearly depicts differences between the kernel density estimate and the Johnson SU at longer distances, reveals over-prediction at distances between 500 km and 1,000 km and increasing under-prediction at distances >1,000 km. (Note that the plot in Fig. 4B is truncated.

![Fig. 3. Histogram of visitor travel distances based on the full NRRS data set. Inset plot shows the histogram for distances up to 1,000 km.](image)

![Fig. 4. (A) Plot of density versus travel distance for the unbounded Johnson (SU) distribution and the kernel density estimate for the full NRRS data set. Inset plot shows distances up to 1,000 km; (B) plot of the logarithm of density versus travel distance for the Johnson SU distribution and the kernel density estimate for the full NRRS data set.](image)
at distances >4,000 km; beyond this distance, the number of observations was insufficient to derive a smooth logarithmic estimate of the kernel density.

This pattern of fit generally holds true for each of the tested geographic subregions (Figs. 5 and 6) and for the national parks and urban areas data subsets (not shown). The kernel density estimates for the northeastern United States and western U.S. subregions (Fig. 5A, D, respectively) display a degree of multimodality (i.e., secondary peaks at distances ≈200 km) that is not well captured by the univariate Johnson $S_U$ distribution, suggesting that mixture distributions might be more appropriate for these subsets. Figure 6 shows a plot of the logarithm of density versus distance for each of the geographic subregions. (As with the full NRRS data set, the logarithmic estimates of the kernel density for each subregion become less reliable as they approach the maximum recorded distance, which differed region-to-region; see Table 1.) The subregions exhibit under-prediction at longer dis-

![Fig. 5. Plots of density versus travel distance (distances up to 1,000 km) for the unbounded Johnson ($S_U$) distribution and the kernel density estimate, based on data for the four analytical subregions: (A) northeastern United States, (B) southern United States, (C) midwestern United States, (D) western United States.](image1)

![Fig. 6. Plots of the logarithm of density versus travel distance for the unbounded Johnson ($S_U$) distribution and the kernel density estimate, based on data for the four analytical subregions: (A) northeastern United States, (B) southern United States, (C) midwestern United States, (D) western United States. Distances >4,000 km have been omitted.](image2)
stances in a manner similar to the full data set, although in the southern United States (Fig. 6B), under-prediction begins at a comparatively shorter distance (~750 km), and in the western United States (Fig. 6D), at a comparatively longer distance (~2,000 km).

Mixture Distributions. Figure 7 shows a three-component lognormal mixture distribution applied to the data subset for the 20 most visited U.S. national parks. Table 2 shows the parameter estimates for each of the component distributions. We used the lognormal because the Johnson SU is not implemented in the mixdist package (and the two performed similarly during univariate distribution fitting). The standard errors of the parameter estimates were fairly low, suggesting the estimates were precise, and the fit of the mixture, based on the chi-squared approximation of the likelihood ratio, was highly significant ($P < 0.0001$). However, these findings were similarly true of other three- and four-component mixture distributions (as well as single-component distributions) that we tested. In short, there were a few reasonable candidate mixture distributions for the national parks subset; we selected the one that, in our judgment, appeared to provide the best overall fit to the data histogram. In any case, a

Fig. 7. Plot of a three-component lognormal mixture distribution (green) fitted to the histogram (blue) for the data subset associated with the 20 most visited U.S. national parks. Inset plot shows distances up to 1,000 km. The individual component distributions are depicted by red lines, while the red triangles indicate their mean ($\mu$) values.

Fig. 8. Plot of a four-component lognormal mixture distribution (green) fitted to the histogram (blue) for the data subset associated with the 20 most populous U.S. urban areas. Inset plot shows distances up to 1,000 km. The individual component distributions are depicted by red lines, while the red triangles indicate their mean ($\mu$) values.
The top 20 most populous U.S. urban areas

Distributions of a lognormal mixture applied to the data subset for the top 20 most visited U.S. national parks

Distributions of a lognormal mixture applied to the data subset for the urban areas subset. Similar to component 1 in the size that this resulted from the inherent structure of distributions (i.e., components 1 and 2). We hypothesized that the data reflect the behavior of distinct subpopulations of travelers. Thus, the three component distributions may be interpreted as corresponding to three hypothetical groups of campgrounds: the large majority of overnight campers, whose travel times are typically less than a few hours (component 1); campers visiting specific destination campgrounds (i.e., specific national parks) who are thus willing to travel farther (component 2); and campers with longer-term itineraries, perhaps involving the midday use of recreational vehicles and visits to multiple destinations (component 3).

Figure 8 shows a four-component lognormal mixture distribution applied to the data subset for the 20 most populous U.S. urban areas, while Table 3 shows the parameter estimates for each of the component distributions. As with the mixture distribution for the national parks data subset, the standard errors of the parameter estimates were low, and the fit was found to be highly significant ($P < 0.0001$). Also like the national parks subset, other (three- and four-component) mixture distributions that we tested performed similarly in terms of these metrics, so we selected the one that we judged to offer the best overall fit to the data histogram. One interesting aspect of the selected mixture (and of the other candidate mixtures we tested) is the presence of two overlapping component distributions (i.e., components 1 and 2). We hypothesize that this resulted from the inherent structure of the urban areas subset. Similar to component 1 in the national parks subset, these components (that together capture most of the short distances recorded in the subset) likely correspond to overnight campers. However, we suspect there is something of a split in this subpopulation; basically, overnight campers from certain densely populated, large cities (e.g., New York) likely have to travel farther to get beyond developed areas than campers from cities surrounded by a wide “transition zone” where wildlands and developed areas are highly intermixed. If we downplay this difference and assume that components 1 and 2 represent slightly different portions of the same subpopulation, then component 3 in the urban areas mixture distribution can be interpreted similarly to component 2 in the national parks mixture (i.e., representing campers visiting specific destinations and thus willing to travel farther), while component 4 can be interpreted similarly to component 3 in the national parks mixture (i.e., representing campers with longer-term itineraries).

Discussion

The analyses presented in this study lead to two major generalizations. The first generalization is that a majority of campground visits involve relatively short travel distances ($<100$ km, which is usually a $<2$-h drive from home). This is not surprising, because most campground visits probably last only a few days (e.g., a weekend), and under such circumstances, campers are likely to minimize the travel time to their destinations. Nevertheless, given that some percentage of campers are likely to carry infested firewood, these short-distance trips are highly relevant in an invasion context, because they have the potential to transport forest insect pests well beyond their natural dispersal range. For example, in flight mill studies, the median flight distance of $A.\ planipennis$ mated females was $>3$ km in 24 h (with only 1% flying $>20$ km), while mass mark-recapture studies of $A.\ glabripennis$ suggest a natural dispersal potential of 2–3 km per year (Smith et al. 2001, 2004; Taylor et al. 2010). Furthermore, based on the NRMS data, the percentage of campground visits involving longer travel distances (i.e., $>500$ km) appears to be substantial (10% in the full data set).

The second major point is that we were able to attain a good fit with theoretical distribution functions, despite the data being strongly leptokurtic. A primary reason for this success was that the NRMS database and its subsets had large sample sizes and were temporally robust (i.e., spanning a 5.75-yr period), thus providing adequate data to reasonably characterize the likelihood of long-distance dispersal events. Nevertheless, the observed fits with the Johnson $S_{B}$ distribution, which was the best-performing univariate function in all cases, were not perfect. In particular, the Johnson $S_{B}$ distribution consistently underpredicted dispersal likelihood at very long (>1,000 km) travel distances. This is not necessarily problematic for modeling dispersal of forest insect pests in firewood, because the region of under-prediction generally encompassed only the rarest long-
April 2012  KOCH ET AL.: DISPERSAL OF INVASIVE FOREST INSECTS IN FIREWOOD 447

distance events (<1–2% of the observed distances), and campers who travel such long distances may be less likely to bring firewood with them. Regardless, the superiority of the Johnson $S_U$ distribution (and log-normal distribution) to the other theoretical distribution functions we tested suggests that invasion modelers should carefully consider their assumptions regarding what distribution will provide the best kernel for modeling a long-distance dispersal process (i.e., commonly used distributions like the negative exponential may not always be appropriate).

The kernel density estimates for two of our geographic subregions displayed some degree of multimodality. Based on our fitting of mixture distributions to the national parks and urban areas subsets, it seems likely that we could also successfully fit mixtures to these subregional data sets. Mixture distributions have the additional advantage of dividing the data into different subpopulations, which, for example, could each be assigned a unique probability of transporting forest insects. The notion of subpopulations of campers with distinctive travel patterns also has important implications for developing public awareness campaigns and outreach programs. For example, each subpopulation could be targeted with an individual suite of policies, which might prove more efficient than applying a “blanket” approach to all campers.

Two important limitations of our analysis should be mentioned. First, because of privacy concerns, all personally identifiable information was removed from the NRRS data records before their release to us. As a result, we were unable to document cases where a camper made reservations at several different campgrounds on consecutive or near-consecutive dates (i.e., a multi-stop itinerary). In truth, some of the longer-distance campground visits that we portrayed as “nonstop” trips in our analysis likely involved one or more intermediate stops along the way. It seems unlikely that the campers on these multi-stop trips who began traveling with firewood brought from home were still carrying it when they reached their final destinations. However, it is also quite plausible that they kept some firewood from a previous stop, possibly hundreds of kilometers away. In effect, these campers still presented a substantial risk of introducing forest insect pests to novel locations. Ultimately, we made the simplifying assumption that the potential effect of multi-stop camper itineraries on the dispersal distance estimates was mitigated by a presumably much larger proportion of single-stop itineraries in the NRRS data, as well as by scenarios such as the one just described.

A second limitation relates to our use of Euclidean distance as the measure of the travel distance between each visitor origin ZIP code and destination campground. A more realistic distance measure would be the length of the most likely road route connecting the two locations. Unfortunately, determining the most likely route for each origin-destination combination is a difficult computational problem, made even more challenging by the large volume of such combinations in the NRRS data. Foremost, the most likely road route is not necessarily the shortest-length route; for example, various categories of roads (e.g., interstate highways vs. local access roads) may permit dramatically different travel speeds, thus influencing travelers’ route choices. Although the Euclidean distance between two geographic locations underestimates the route-based travel distance by some unknown amount, the sophisticated geospatial modeling efforts necessary to determine the latter for each unique origin-destination combination in the NRRS data were beyond the scope of this analysis. However, this is a key area for future work.

Acknowledging these limitations, our study has nevertheless presented a straightforward quantitative approach for developing and parameterizing dispersal kernels to characterize the spread of forest insects via recreational travel. Additionally, we have shown that the kernels, whether based on univariate or mixture distributions, can be customized for certain invasion scenarios and geographic subsets (e.g., analyzing the expansion of an insect pest in a particular region of interest, or focusing on campgrounds with environmental conditions especially suitable for an invader). We envision one of these kernels (or a similarly derived kernel) being used to simulate recreational (i.e., camper-related) dispersal in spatial invasion models, as a complement to algorithms depicting biological dispersal or other modes of human-mediated dispersal. To do so would require making assumptions regarding the propagule pressure (i.e., number of dispersing individuals; Lockwood et al. 2005) associated with the kernel; regrettably, with respect to recreationally transported firewood, the existing data are insufficient in quantity and geographic scope to make anything but coarse propagule pressure estimates for individual forest pests of interest. (In fact, a lack of empirical data about propagule pressure is a universal problem for invasion modeling; see Lockwood et al. 2007.)

Still, we can use the available information on firewood to make some general statements regarding the risk of spreading forest insects in the United States. Based on data compiled from a small number of firewood inspections and usage surveys (Haack et al. 2010, USDA-APHIS 2011b), we estimate that 30–40% of campers carry firewood from home (or other distant locations). If we assume that ~20% of the firewood is infested with live wood borers (roughly equal to the percentage reported by Haack et al. 2010), then ~6–10% of campground visits involve the movement of firewood infested with viable forest insects. Accounting for factors such as the burning of firewood before any pests can escape, or the fact that some transported insect species will already be present at their eventual destinations, it seems reasonable to estimate that 3–5% of campground visits pose a potential risk of firewood-mediated dispersal of forest pests. The NRRS data used in this study recorded an average of ~1.2 million campground visits per year. Furthermore, these data represent only a subset of all campground visits in the United States; for instance, there are >9,000 private campgrounds distributed across the country, compared with 2,525 for the NRRS (USDA-APHIS 2011b).
This implies that, even if only 1% of recreational firewood contained viable insects, potentially tens of thousands of camping trips occur each year in the United States during which infested firewood is transported to a distant location. Hence, current concerns about the risk of forest pest spread in camper-transported firewood appear to be well justified.

In fact, a better question might be why we have not seen a greater proliferation of forest insect pests because of recreational firewood transport. We believe there are a few feasible explanations for this. Most likely, the environmental conditions or the reproductive capacity of the introduced insects, especially given the potential for Allee effects (Taylor and Hastings 2005), are inadequate for establishment in the vast majority of firewood-related introductions. The insects’ specific life history attributes may also affect establishment likelihood. For instance, Haack et al. (2011) noted substantial variation in the rates at which establishment likelihood. For instance, Haack et al. (2011) noted substantial variation in the rates at which different forest insect families infest host material. They estimated that an infested log \( \approx 1 \) m in length and \( \approx 10 \) cm in diameter could be expected to contain 100–250 individual bark beetles, 20–30 buprestids (e.g., \( A. \) planipennis), or 5–10 cerambycids (e.g., \( A. \) glabripennis). These estimates suggest that a bundle of infested firewood is more likely to contain a sufficient population of bark beetles for establishment than a sufficient population of larger borers (such as buprestids or cerambycids). Alongside such population-based explanations, it is possible that the existing regulations and public awareness campaigns are indeed having a positive effect on human behavior with respect to firewood use and transport. Alternatively, it may be the case that many existing infestations initiated by firewood have simply not been detected yet. The correct answer is probably some combination of these possibilities.

The spread of invasive pests via recreational travel (and associated firewood transport) is conceptually similar to other modes of human-mediated dispersal, such as the movement of invasive organisms with tradable commodities (Hulme et al. 2008). However, certain aspects of recreational travel are distinctive, so the dispersal patterns (i.e., dispersal kernels) observed in this study may not translate directly to other human-mediated dispersal processes. For example, with respect to domestic trade, commodities are generally moved along pathways linking areas of high industrial or agricultural activity to populated places (i.e., cities and towns). In contrast, recreational travel follows pathways that link populated places to destinations that are often located in sparsely populated regions. Thus, the destinations and underlying objectives of these two potential human-mediated dispersal modes are quite different. Notably, residential firewood usage, which we did not consider in this study, involves aspects of both modes: Many homeowners obtain their firewood directly from private (e.g., individually owned woodlots) or public lands (e.g., through fuelwood harvesting permits on National Forests), while others purchase it from large retail or wholesale distributors (USDA-APHIS 2011b). Given that firewood is used regularly in \( \approx 30 \) million U.S. homes (Houck et al. 1998), this topic probably deserves further analytical attention from a biological invasions perspective.

A final, technical point pertains to the appropriate application setting for the dispersal kernels presented in this study. The kernels are effectively omni-directional; in other words, the likelihood of dispersal from a given point of origin to a given destination is simply a matter of the distance, and not the orientation, between the two points. While omni-directional kernels may be satisfactory for many spatial modeling applications, it may be more realistic to use the NRRS data in a network-based setting, where dispersal occurs via specified routes (e.g., road corridors) and is restricted to the set of potential destinations (e.g., campgrounds) defined for the network. Additionally, the networked data could also serve as the basis for a pathway model, which could provide, for example, probabilistic estimates of the most likely pathways and destinations for a forest pest introduced at a given origin node. Network-based implementation of the kernels derived in this study will be a focus of future research.

In summary, our analyses appear to validate current regulatory and public outreach efforts regarding firewood transport and the potential spread of invasive forest pests. Although most campers travel relatively short distances, and even though only a small proportion of these campers are likely to be carrying infested firewood, this still translates into a sizeable increase in dispersal potential beyond the natural spread capabilities of most forest insects. This is especially true given the huge number of camping trips that occur each year. While many aspects of the firewood issue remain unclear, we have provided some preliminary answers that should prove useful to decision makers and other researchers. Additional work is needed to relate our findings to the unique circumstances of individual species of interest.

Acknowledgments

We thank Judith Pasek, David Kowalski, and Daniel Borchert (USDA-APHIS) for providing the data used in this study. We also thank Kurt Ritters and John Coulston (USDA Forest Service), as well as two anonymous reviewers, for their helpful comments. Frank Koch’s initial work on the study was supported by Research Joint Venture Agreement 10-JV-1130146-064 between the U.S. Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC, and North Carolina State University.

References Cited


Received 12 August 2011; accepted 7 December 2011.