Spatiotemporal dynamics of insect-fire interactions

Heather J. Lynch

Paul R. Moorcroft

Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA.

Summary. In this paper, we extend traditional methods of spatial statistics to study causal spatiotemporal correlations between two different point processes. After developing the methodology, we apply this analysis to a particular case study of interest in ecology, the interaction between damage by a particular forest pest (western spruce budworm) and forest fires. Our analysis, which covers parts of British Columbia in the 25 year period from 1971-1995, indicates that areas affected by budworm infestation have a significantly decreased risk of forest fire for the 9 years following the infestation. Conversely, forest fires increase the risk of budworm infestation for the first year and then decrease the risk of infestation for a 2-5 year period after the fire. These temporal correlations extend over a spatial range of at least 10 km. Our study rejects the common assumption that insect infestation necessarily results in increased fire risk. This case study illustrates the utility of this approach to understanding ecosystem dynamics extending over both space and time.

Keywords: British Columbia, fire ecology, spatiotemporal interactions, Ripley's K-function, western spruce budworm

1. Introduction

Throughout history, insect outbreaks and forest fires have proven themselves powerful agents of forest disturbance and mortality. In 2002, in the United States alone, over 16.1 million acres of forest (U.S.D.A., 2003) were damaged or destroyed by insect pests and another 6.9 million acres of forest (N.I.F.C., 2003) were lost to forest fires despite widespread programs to control insect outbreaks and to manage forest fires. It is often stated in the literature that insect outbreaks and forest fires are non-independent forms of disturbance. The suggestions are that insect outbreaks may leave forests

more (Knight, 1987) or less (Despain, 1990; Kulakowski et al., 2003) vulnerable to forest fires, that forest fires may 'inoculate' forests against certain insect outbreaks (Kulakowski et al., 2003; Veblen et al., 1994; Bebi et al., 2003) while leaving trees vulnerable to others (McHugh et al., 2003), and that two-way interactions between forest fires and insect outbreaks may act synergistically to determine forest composition (Schmid and Hinds, 1974; McCullough et al., 1998). Since both the sign and the magnitude of potential correlations is still under debate and the mechanisms by which forest fires and insect outbreaks interact are complex, it is important to study the empirical spatiotemporal correlations of actual disturbances. In doing so, we can determine the net interaction between these two disturbance modes and, from this, draw conclusions about the dominant mechanisms mediating these interactions. Quantitative studies of the spatiotemporal dynamics between insect damage and forest fires have, however, been extremely limited (Fleming et al., 2002) due to a shortage of spatially-explicit, large-scale datasets and techniques with which to analyze such data.

In this paper, we develop a technique to answer these types of ecological questions by extending a traditional estimate of second-order properties, Ripley's cross-K function $K_{12}(s)$ (Cressie, 1993) to a space-time cross-K function $K_{12}(s,t)$. In addition, we modify the expressions for $K_{12}(s,t)$ and its variance $V_{12}(s,t)$ to account for the causality inherent in having two different point processes interacting in time.

2. Description of the data

The data used in this analysis was drawn from the British Columbia Natural Disturbance Database compiled by the Canadian Forest Service. The western spruce budworm (*Choristoneura occidentalis*) is a major forest pest in British Columbia and at times has been the most destructive forest defoliator in western North America (Furniss and Carolin, 1977). Its primary hosts are Douglas-fir and grand fir, although it has a number of secondary hosts such as Engelmann spruce, western larch, and subalpine fir. The western spruce budworm larvae will feed on older growth from previous seasons but prefer new growth, damaging expanding needles and developing cones (U.S.D.A., 1982). Successive years of severe defoliation can cause decreased growth, deformity and ultimately, mortality. Because the western spruce budworm feeds on specific host species, in British Columbia the western spruce budworm is primarily confined to the Interior Douglas Fir (IDF) biogeoclimatic zone. For this reason, and to

minimize heterogeneity in disturbance risk, we limited the spatial scope of our analysis to fires and *C. occidentalis* infestations falling within the IDF zone plotted in Figure 1A. Figures 1B and 1C show the 25 year dataset of western spruce budworm damage and forest fires used in this analysis. In the period examined, 1971-1995, there were 21,042 fires events and 6,525 recorded regions of budworm infestation, affecting up to 0.14% and 12.0% (respectively) of the IDF zone in any given year.

To study the interaction between insect and fire occurrence, we transformed the historical records shown in Figure 1C into point data sets, digitizing the insect outbreak and fire areas onto a 250 m by 250 m grid according to majority rule. Since only fire sizes were reported in the dataset, each fire event was assumed to have burned a circular area centered on the reported fire center with an area equal to the reported fire area. Each affected grid cell was defined as an event indexed by its position (x,y) and the year (t) in which the disturbance occurred.

3. Spatiotemporal Model

One commonly used method of estimating whether a single stochastic process is clustered in space and time is Ripley's K-function (Cressie, 1993), defined as

$$K(s,t) = \frac{1}{\lambda}\mathrm{E}[\#\text{ of further events occurring within distance } s \text{ and time } t \text{ of an arbitrary event}] \quad (1)$$

where the number of events (N) and the area (A) may be used, under the assumptions of homogeneity, to estimate the intensity of the point process $\lambda = N/A$. Unlike the Knox test for spatial correlation, Ripley's K-function allows us to examine space-time clustering over a range of spatial and temporal scales. As detailed in Appendix A, we extended the traditional K-function to examine space-time correlations between two distinct, causally-related, stochastic point types (insect pest outbreaks (p) and forest fires (f)). Specifically, we define a space-time cross K-function as

$$K_{pf}(s,t) = \frac{1}{\lambda_f}\mathrm{E}[\#\text{ of further fire events occurring within (s,t) of a given insect pest event}] \quad (2)$$

where $\lambda_f \sim N_f/A$ represents the intensity of the fire point process. The converse function, K_{fp}(s,t) is defined analogously. Causality is inferred by considering only those point pairs which are causal. For

example, in calculating Eq.(2) we consider only those point pairs for which a fire event occurred t (years) following an insect pest infestation event. Note that in this case, $\mathrm{K}_{pf}(\mathrm{s,t}) \neq \mathrm{K}_{fp}(\mathrm{s,t})$, and therefore, you need all $f \times p$ functions to completely describe the second-order properties of the process.

Rather than analyzing the raw cross-correlations we calculated the derived function

$$D_{o,pf}(s,t) = \frac{K_{pf}(s,t) - K_{pf}(s) \cdot K_{pf}(t)}{K_{pf}(s) \cdot K_{pf}(t)} \tag{3}$$

where $\mathrm{K}_{pf}(\mathrm{s})$ and $\mathrm{K}_{pf}(\mathrm{t})$ are the cross K-functions in pure space or pure time respectively. This function has been used by a number of authors in the analysis of geographical epidemiology (see, for example, Diggle et al. (1995) and Gatrell et al. (1996)). In this context, $\mathrm{D}_{o,pf}(\mathrm{s,t})$ may be interpreted as representing the increased risk of a fire event due to an insect pest event at spatial distance s and time lag t, and represents the portion of the total space-time correlation that remains after correcting for purely spatial and purely temporal correlations.

A nice feature of the $\mathrm{D}_{o,pf}$ function (Eq. 3) is that it eliminates correlations between the two processes arising from purely spatial covariates such as elevational gradients and purely temporal covariates such as weather-related increases in fire frequency and insect outbreak frequency. In doing, we isolate the effect of only those processes that are correlated in both space and time. It is important to emphasize that our null hypothesis is that western spruce budworm outbreaks and forest fires exhibit no *space-time* interactions. Purely spatial or purely temporal correlations may exist due to a variety of ecology heterogeneities that are difficult to exclude. However, as discussed by Diggle et al. (1995), the statistical tests used in this analysis remain valid even if the underlying process is non-stationary.

The statistical significance of the $\mathrm{D}_{o,pf}(\mathrm{s,t})$ estimated was determined by calculating

$$R_{pf}(s,t) = \frac{D_{o,pf}(s,t)(K_{pf}(s) \cdot K_{pf}(t))}{\sqrt{V_{pf}(s,t)}} \tag{4}$$

where $\mathrm{R}_{pf}(\mathrm{s,t})$ is analogous to the standardized residuals of the $\mathrm{K}_{pf}(\mathrm{s,t})$ function once purely spatial and purely temporal correlations have been accounted for (Diggle et al., 1995), and $V_{pf}(\mathrm{s,t})$ is the variance of $\mathrm{D}_{o,pf}(\mathrm{s,t})$. The statistical significance of $\mathrm{D}_{o,fp}(\mathrm{s,t})$ was calculated in a similar manner. If two processes are space-time independent, $\sim 95\%$ of $\mathrm{R}_{pf}(\mathrm{s,t})$ values should lie within $\pm$ 2. $\mathrm{R}_{pf}(\mathrm{s,t})$ values whose absolute value is larger than 2 indicate $\mathrm{D}_{o,pf}(\mathrm{s,t})$ values which are statistically significant. The results

presented below are are based on calculations using a buffer edge correction method (Cressie, 1993) in which all points falling within 10 km of the edge were excluded from consideration but were used as neighbors to events lying in the interior. Estimating $D_{o,pf}(s,t)$ and $D_{o,fp}(s,t)$ using a re-weighting edge-correction scheme (Cressie, 1993) gives similar results.

$V_{pf}(s,t)$ and $V_{fp}(s,t)$ were calculated analytically (see Appendix A) and checked numerically using a Monte Carlo method in which new data sets (consistent with the initial first-order properties of the original data) were created by random permutations of the time 'labels'. The Monte Carlo estimations for the variance using N=50 simulations of the data are somewhat smaller that the theoretically calculated estimate for the variance. This is likely due to the fact that the computation time required to calculate the $D_o(s,t)$ function limited the number of simulations we could perform. In either case, however, there was consistency between the two methods as to which data points were statistically significant. The results discussed below, and represented in Figures 2 and 3, represent the more conservative (larger) estimate for the variance obtained using the analytical calculation detailed in Appendix A.

4. Results

The results of our analysis are plotted in Figures 2 and 3. Figure 2 shows $D_{o,pf}(s,t)$, the additional risk of forest fire following an *C. occidentalis* outbreak. The filled points indicate results which are significant to 5% confidence levels. We see that budworm outbreaks lead to a significantly decreased risk of forest fire. In the immediate vicinity of a spruce budworm outbreak, the risk of fire during the first year following the outbreak is reduced by approximately 40%. This reduction in fire risk decreases steadily over the next five to ten years following the budworm damage and also declines with increasing distance from the infection center but remaining significant to at least 10 km, the longest length scale examined in the study.

Figure 3 shows the effect of forest fire on future risk of *C. occidentalis* damage. In the year immediately following a fire, there is a 10%-25% increase in the risk of *C. occidentalis* infestation in areas within 5 km of the fire. This elevated risk declines over the next two years, and by the third year there is up to a 12% lower risk of *C. occidentalis* infestation. This rises again to a small peak after the sixth year following fire and then levels off to very small levels of additional risk.

The result shown in Figure 2 directly contradicts the common assumption that insect damage necessarily increases the risk of forest fire (McCullough et al., 1998; Dale, 2001; O'Driscoll, 2004; Robbins, 2004). While studies suggest that this may be the case for certain insects in particular forests (e.g. jack pine budworm in jack pine forests (McCullough et al., 1998)), the overall situation is more complex. Previous studies have suggested several mechanisms by which insect damage may promote or inhibit fires via changes in the vertical distribution of fine fuels within the canopy (Knight, 1987; Despain, 1990) and changes in the light, temperature and humidity conditions within the forest (Kulakowski et al., 2003; Reid, 1989). For example, insect-related damage can lead to extensive patches of dead needles within the forest canopy. If these remain in the overstory, the resulting matrix of fine, dry fuels may increase the risk of a stand-replacing canopy fire (Knight, 1987; McCullough et al., 1998). On the other hand, if dead needles quickly fall off the trees, fuel in the canopy layer may decrease, reducing the risk of canopy fire (Despain, 1990). Similarly, insect activity can alter the occurrence of surface fires. If needle-drop occurs due to insect activity and dead needles accumulate on the forest floor, this increases the availability of dry surface fuels, promoting the occurrence of surface fires (McCullough et al., 1998). Conversely, defoliation following insect activity can lead to an opening of the canopy that promotes growth of understory herbs and other ground cover (Reid, 1989), and the resulting increase in understory moisture may lead to a decreased risk of surface fires following insect defoliation (Kulakowski et al., 2003).

5. Discussion

Analyzing the empirical data on actual fire and insect outbreaks enables us to determine the net effect of these competing factors on the probability of fire. While additional field studies will be required to study the balance between the altered risks of surface and canopy fires due to insect activity in this system, our analysis suggests that the net affect, on average, is that *C. occidentalis* outbreaks decrease forest fire risk in the 5-10 years following an outbreak.

There are several possible reasons for the initial increase in risk of *C. occidentalis* infestation in the first year following forest fires seen in Figure 3. One often cited possibility is that trees that are damaged but not killed by a forest fire may be less able to defend themselves against insect infestation and may act

as epicenters for outbreak which then spread into surrounding healthy forest (Knight, 1987; McCullough et al., 1998; Dale, 2001). However, these studies have focused on interactions between bark beetles and forest fires, and there are several reasons to suggest that weakened tree defenses are unlikely to play a major role in precipitating outbreaks of western spruce budworm. Previous work has demonstrated that the primary mechanisms of Douglas-fir resistance to *C. occidentalis* are high growth rate and late bud burst phenology as opposed to chemical defenses via secondary compounds (Chen et al., 2001, 2002) and studies of the closely related *C. fumiferana* suggest that outbreaks are not initiated by a single epicenter but occur in synchrony over large areas (Royama, 1984).

A second explanation for why fires may promote insect outbreaks is that forest fires perturb the balance between *C. occidentalis* and its predators or parasites. It is believed that predation on *C. occidentalis* larvae by ants and birds is effective in regulating budworm populations at endemic levels, but are ineffective once populations have reached outbreak levels (Torgersen and Campbell, 1982; Torgersen et al., 1990; Campbell et al., 1983). If the ecosystem disruption of a forest fire allows *C. occidentalis* to escape from its predators long enough to reach outbreak levels, this may account for the increased risk of insect outbreak in the year following a forest fire. This may occur due to post-fire changes in diet selection among birds known to feed on both seeds, which become more prevalent following a forest fire, and *C. occidentalis* larvae†. Finally, if low intensity forest fires reduce competition for light and nutrients in the forest, unusually vigorous growth in the next growing season might increase the carrying capacity for defoliators such as *C. occidentalis*, allowing the budworm populations to reach outbreak levels.

In addition to the immediate effects of fire, Figure 3 also shows evidence of a delayed increase in insect outbreak risk six years after a fire, indicating the existence of both long-term as well as short-term spatiotemporal correlations between forest fires and subsequent insect outbreaks. A potential explanation for this lagged response to fire may be insect infestation on young trees established following fire related mortality.

†Of the five bird species known to feed on *C. occidentalis* (*Hesperiphona vespertina*, *Spinus pinus*, *Spizella passerina*, *Junco hyemalis* and *Piranga ludoviciana*), four also feed on seeds (Torgersen and Campbell, 1982).

6. Conclusion

The interactions between fires and *C. occidentalis* outbreaks found in this study have important implications for the management of forests in western North America. In particular, future climate change and shifting policies regarding fire suppression are likely to have significant consequences for insect populations, forest fires, and their interactions and an improved understanding of these complex feedbacks has been cited as a research priority (Ayres and Lombardero, 2000). While our analysis has focused exclusively on the interaction between *C. occidentalis* and forest fires, our approach provides a general method for exploring for space-time interactions among different modes of disturbance wherever long-term, spatially explicit, historical records are available.

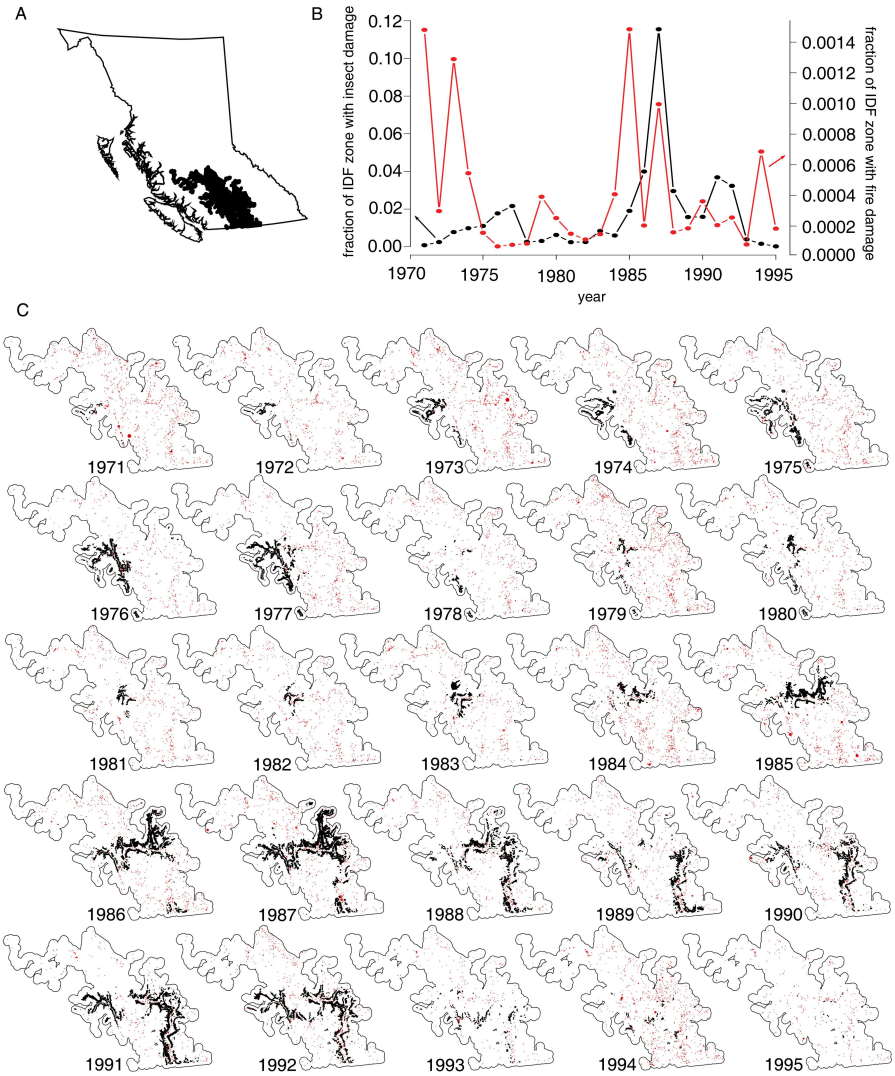

Figure 1: A. Shaded region represents the Interior Douglas Fir biogeoclimatic zone within the province of British Columbia. **B.** Time series plot of the fraction of the Interior Douglas Fir biogeoclimatic zone affected by forest fires (red) and western spruce budworm (black) during the period 1971-1995. **C.** Interior Douglas Fir Zone overlayed with the fire (red) and western spruce budworm (black) disturbance data for the period 1971-1995.

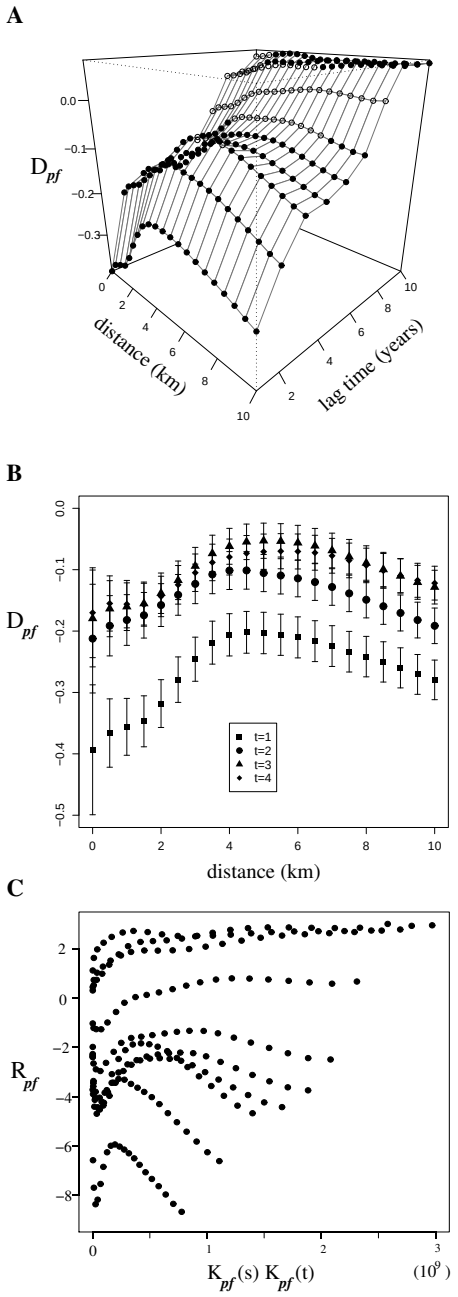

Figure 2: **A.** $D_{pf}(s,t)$ indicating increased ($D_{pf} > 0$) or decreased ($D_{pf} < 0$) risk of forest fire following western spruce budworm infestation. **B.** Cross-section of $D_{pf}(s,t)$ for the first 4 years following western spruce budworm damage. Error bars represent 1σ error as calculated analytically according to the expression derived in Appendix A. **C.** $|R_{pf}(s,t)|$ values larger than 2 indicate $D_{pf}(s,t)$ that are significant; these points are indicated by solid circles in A.

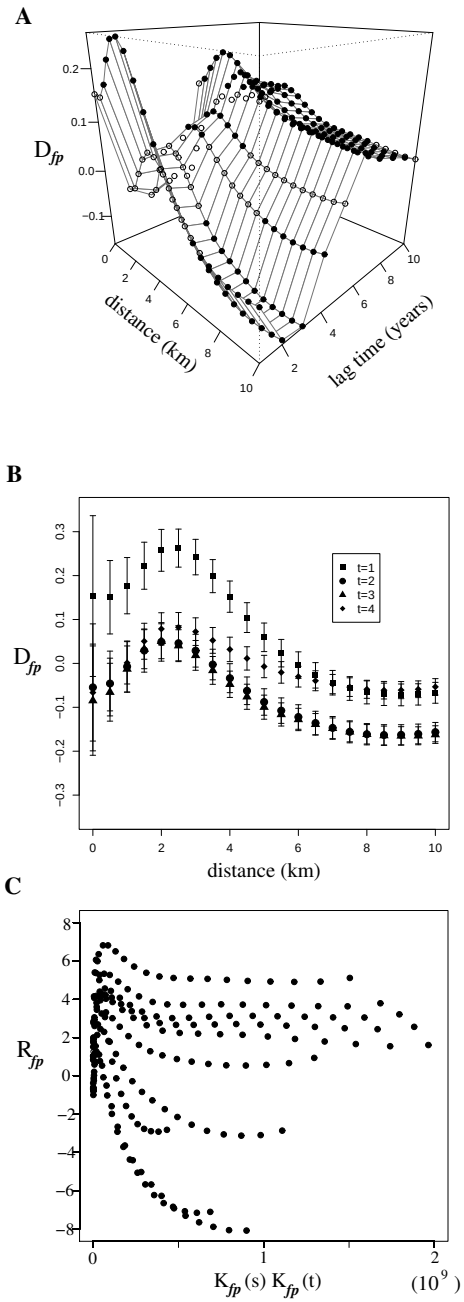

Figure 3: **A.** D_{fp}(s,t) indicating increased ($D_{fp} > 0$) or decreased ($D_{fp} < 0$) risk of western spruce budworm outbreak following a forest fire. **B.** Cross-section of D_{fp}(s,t) for the first 4 years following a forest fire. Error bars represent 1σ error as calculated analytically according to the expression derived in Appendix A. **C.** $|R_{fp}(s,t)|$ values larger than 2 indicate D_{fp}(s,t) that are significant; these points are indicated by solid circles in A.

References

Ayres, M. and M. Lombardero (2000). Assessing the consequences of global change for forest disturbance from herbivores and pathogens. *Sci. Total Environ. 262*, 263–286.

Bebi, P., D. Kulakowski, and T. Veblen (2003). Interactions between fire and spruce beetles in a subalpine rocky mountain forest landscape. *Ecology 84*(2), 362–371.

Campbell, R., T. Torgersen, and N. Srivastava (1983). A suggested role for predaceous birds and ants in the population dynamics of the western spruce budworm. *For. Sci. 29*, 779–790.

Chen, Z., T. Kolb, and K. Clancy (2001). Mechanisms of douglas-fir resistance to western spruce budworm defoliation: bud burst phenology, photosynthetic compensation and growth rate. *Tree Physiol. 21*, 1159–1169.

Chen, Z., T. Kolb, and K. Clancy (2002). The role of monoterpenes in resistance of douglas fir to western spruce budworm defoliation. *J. Chem. Ecol. 28*(5), 897–920.

Cressie, N. (1993). *Statistics for Spatial Data*. New York: John Wiley and Sons.

Dale, V. (2001). Climate change and forest disturbances. *BioScience 51*, 723–734.

Despain, D. (1990). *Yellowstone Vegetation: Consequences of Environment and History in a Natural Setting*. Boulder, CO: Roberts Rinehart Publishers.

Diggle, P., A. Chetwynd, R. Haggkvist, and S. Morris (1995). Second-order analysis of space-time clustering. *Stat. Methods Med. Res. 4*, 124–136.

Fleming, R., J. Candau, and R. McAlpine (2002). Landscape-scale analysis of interactions between insect defoliation and forest fire in central canada. *Clim. Change 55*, 251–272.

Furniss, R. and V. Carolin (Eds.) (1977). *Western Forest Insects*. United States Department of Agriculture Forest Service.

Gatrell, A., T. Bailey, P. Diggle, and B. Rowlingson (1996). Spatial point pattern analysis and its application in geographical epidemiology. *Trans. Inst. Br. Geogr. NS 21*, 256–274.

Knight, D. (1987). Parasites, lightening, and the vegetation mosaic in wilderness landscapes. In M. Turner (Ed.), *Landscape Heterogeneity and Disturbance*, Chapter 4, pp. 59–83. Springer-Verlag.

Kulakowski, D., T. Veblen, and P. Bebi (2003). Effects of fire and spruce beetle outbreak legacies on the disturbance regime of a subalpine forest in colorado. *J. Biogeogr. 30*, 1445–1456.

McCullough, D., R. Werner, and D. Neumann (1998). Fire and insects in northern and boreal forest ecosystems of north america. *Annu. Rev. Entomol. 43*, 107–127.

McHugh, C., T. Kolb, and J. Wilson (2003). Bark beetle attacks on ponderosa pine following fire in northern arizona. *Environ. Entomol. 32*(3), 510–522.

N.I.F.C. (2003). Wildland fire season 2002 at a glance. Technical report, Available online at http://www.nifc.gov/fireinfo/2002/index.html.

O'Driscoll, P. (2004, January 19). Fire managers fear drought's effect. *USA Today*.

Reid, M. (1989). Ph. D. thesis, University of Colorado, Boulder.

Robbins, J. (2004, July 13). Beetles take a devastating toll on western forests. *New York Times*.

Royama, T. (1984). Populations dynamics of the spruce budworm choristoneura fumiferana. *Ecol. Monogr. 54*, 429–462.

Schmid, J. and T. Hinds (1974). Development of spruce-fir stands following spruce beetle outbreaks. Technical report, USDA Forest Service Research Paper RM-131.

Torgersen, T., R. Mason, and R. Campbell (1990). Predation by birds and ants on two forest insect pests in the pacific northwest. *Studies in Avian Biol. 13*, 14–19.

Torgersen, T. R. and R. W. Campbell (1982). Some effects of avian predators on the western spruce budworm in north central washington. *Environ. Entomol. 11*, 429–431.

U.S.D.A. (1982). Forest service, forest insect and disease leaflet 53. Technical report.

U.S.D.A. (2003). Forest insect and disease conditions in the united states 2002. Technical report.

Veblen, T., K. Hadley, E. Nel, T. Kitzberger, M. Reid, and R. Villalba (1994). Disturbance regime and disturbance interactions in a rocky mountain subalpine forest. *J. Ecol. 82*(1), 125–135.

Appendix A

1. Calculation of $K_{12}(s,t)$ and $K_{21}(s,t)$ for causal point processes

Using Cressie's notation (Cressie, 1993), we extend the cross-K function to include both space and time as follows:

$$\hat{K}^{12}(s) = \frac{1}{\hat{\lambda}_1\hat{\lambda}_2\nu(A)}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} w(s_i^{(1)}, s_j^{(2)})^{-1} I(\|s_i^{(1)} - s_j^{(2)}\| \leq s)$$

$$\hat{K}^{12}(t) = \frac{1}{\hat{\lambda}_1\hat{\lambda}_2\nu(T)}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} v(t_i^{(1)}, t_j^{(2)})^{-1} I(\|t_i^{(1)} - t_j^{(2)}\| \leq t) \tag{5}$$

$$\hat{K}^{12}(s,t) = \frac{1}{\hat{\lambda}_1\hat{\lambda}_2\nu(A)\nu(T)}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} w(s_i^{(1)}, s_j^{(2)})^{-1} v(t_i^{(1)}, t_j^{(2)})^{-1} I(\|s_i^{(1)} - s_j^{(2)}\| \leq s) I(\|t_i^{(1)} - t_j^{(2)}\| \leq t) \tag{6}$$

where $\nu(X \subset R^n)$ is the n-dimensional volume of X. For homogeneous point processes, we can approximate the intensity $\hat{\lambda}_i$ as $n_i/\nu(\bullet)$. We can, therefore, simplify the above equations to get

$$\hat{K}^{12}(s) = \frac{|A|}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} w(s_i^{(1)}, s_j^{(2)})^{-1} I(\|s_i^{(1)} - s_j^{(2)}\| \leq s) \tag{7}$$

$$\hat{K}^{12}(t) = \frac{T}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} v(t_i^{(1)}, t_j^{(2)})^{-1} I(\|t_i^{(1)} - t_j^{(2)}\| \leq t) \tag{8}$$

$$\hat{K}^{12}(s,t) = \frac{|A|T}{n_1 n_2}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} w(s_i^{(1)}, s_j^{(2)})^{-1} v(t_i^{(1)}, t_j^{(2)})^{-1} I(\|s_i^{(1)} - s_j^{(2)}\| \leq s) I(\|t_i^{(1)} - t_j^{(2)}\| \leq t) \tag{9}$$

To consider only cross-correlations between two *causal* stationary point processes, we need to account for the fact that there are no longer $n_1 n_2$ pairs being considered, but only some fraction $f n_1 n_2$ which contribute to the K-function. This fraction f depends explicitly on the data sets being considered and must be calculated in advance. Note that if the two point patterns were completely independently homogenous over the space-time interval, f would equal 1/2.

2. Calculation of V_{12}(s,t) and V_{21}(s,t)

To calculate the mean and variance of

$$\hat{K}(s,t) \equiv \frac{|A|T}{fn_1n_2}\sum_{i,j} w_{ij}v_{ij}I(d_{ij} \leq s)I(u_{ij} \leq t)I(t_i - t_j > 0) \tag{10}$$

it suffices to consider the mean and variance of the rescaled function

$$Q(s,t) = \sum_{i,j} w_{ij}v_{ij}I(d_{ij} \leq s)I(u_{ij} \leq t)I(t_i - t_j > 0). \tag{11}$$

Note that $d_{ij} = \|s_i^{(1)} - s_j^{(2)}\|$ and $u_{ij} = \|t_i^{(1)} - t_j^{(2)}\|$. The mean of Q(s,t) may be written as

$$E[Q] = \frac{1}{fn_1!n_2!}\sum_{\pi}\sum_{i}^{n_1}\sum_{j}^{n_2} W_{ij}V_{\pi^{-1}(i)\pi^{-1}(j)} \tag{12}$$

where $W_{ij} = w_{ij}I(d_{ij} \leq s)$, $V_{ij} = v_{ij}I(u_{ij} \leq t)$, and π represents all possible permutations of the time index and f represents the fraction of those permutations that obey causality. With i (point type 1) fixed and j (point type 2) fixed, each permutation occurs f(n_1-1)!(n_2-1)! times. This may be expressed as

$$\frac{1}{fn_1!n_2!}\sum_{\pi}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij}V_{\pi^{-1}(i)\pi^{-1}(j)} = \frac{f(n_1-1)!(n_2-1)!}{fn_1!n_2!}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2} V_{kl} \tag{13}$$

Using this and the new variables W_1 and V_1 as defined below, we can write E[Q] as

$$E[Q] = \frac{f(n_1-1)!(n_2-1)!}{fn_1!n_2!}W_1V_1 \tag{14}$$

$$= \frac{1}{n_1n_2}W_1V_1 \tag{15}$$

To calculate the variance of Q(s,t) we will calculate the full covariance matrix $E[QQ'] = E[Q(s,t)Q(s',t')]$ and simplify at the end setting s′=s and t′=t. Using the basic definitions of covariance

$$Cov(Q,Q') = E[(Q-\bar{Q})(Q'-\bar{Q'})] \quad (16)$$

$$= E[QQ'] - \frac{1}{n_1^2 n_2^2} W_1 V_1 W_1' V_1'. \quad (17)$$

To calculate E[QQ′], we consider the average of QQ′ over all possible permutations of the time index, leaving the spatial index fixed. With π defined to be all possible permutations of the times, this becomes

$$E[QQ'] = \frac{1}{f n_1! n_2!} \sum_{\pi} ((\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij} V_{\pi^{-1}(i)\pi^{-1}(j)})(\sum_{k=1}^{n_1}\sum_{l=1}^{n_2} W_{kl} V_{\pi^{-1}(k)\pi^{-1}(l)})). \quad (18)$$

The most general term in this expression may be written $W_{ij}V_{xy}W'_{kl}V'_{rs}$ where $\pi(x) = i$, $\pi(y) = j$, $\pi(r) = k$, $\pi(s) = l$. That is, that the same permutation π takes $x \to i$, $y \to j$ etc. Throughout, we assume that i and k indicate points of type 1, and that j and l indicate points of type 2. In general, in considering the space-time interaction between two different point types, we would want to consider clustering of type 1 'around' type 2, and clustering of type 2 'around' type 1. In what follows we will consider clustering of type 2 'around' type 1 in order to calculate the covariance structure appropriate to normalize the function

$$\hat{D}_{12}(s,t) = \hat{K}_{12}(s,t) - K_S(s) \cdot K_T(t). \quad (19)$$

There are, in Eq. (17), four important indices: i, j, k, and l. These four indices can either be all four distinct, have three distinct members or only two distinct members. (There need to be at least two distinct indices since there are two different point types to be summed over.) Most generally, if γ_1 indices for point type 1 are fixed and γ_2 indices for point type 2 are fixed, there are f(n_1-γ_1)!(n_2-γ_2)! possible permutations of the remaining 'free' indices where, again, f is the fraction of all point 1/point 2 combinations which are causal. (Because we only permute the existing time indices and we do not change the number of points in each year bin, this fraction is the same for all permutations.)

Lets start with the case of 4 distinct indices i, j, k, and l. This is the case when we are considering pairs between type 1 and type 2 points and make the restriction that i≠k and j≠l. There is only one way to generate such a set, written as

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1,k\neq i}^{n_1}\sum_{l=1,l\neq j}^{n_2} W_{ij}V_{\pi^{-1}(i)\pi^{-1}(j)}W'_{kl}V'_{\pi^{-1}(k)\pi^{-1}(l)} \tag{20}$$

and it occurs f(n_1-2)!(n_2-2)! times.

We can have three distinct indices in two different ways. Either the same type 1 point can be paired with two different type 2 points, or the same type 2 point can be paired with two different type 1 points. In the first case, only one type 1 index is fixed and two type 2 indices are fixed, so this occurs f(n_1-1)!(n_2-2)! times in the permutation of all time indices. Analogously, the second case happens f(n_1-2)!(n_2-1)! times. These two terms may be written

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{l=1,l\neq j}^{n_2} W_{ij}V_{\pi^{-1}(i)\pi^{-1}(j)}W'_{il}V'_{\pi^{-1}(i)\pi^{-1}(l)} \tag{21}$$

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1,k\neq i}^{n_1} W_{ij}V_{\pi^{-1}(i)\pi^{-1}(j)}W'_{kj}V'_{\pi^{-1}(k)\pi^{-1}(j)} \tag{22}$$

Finally, if there are only two distinct values among i, j, k, and l, this is associated with two identical pairings between a point in type 1 and a point in type 2. These occur f(n_1-1)!(n_2-1)! times.

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij}V_{\pi^{-1}(i)\pi^{-1}(j)}W'_{ij}V'_{\pi^{-1}(i)\pi^{-1}(j)} \tag{23}$$

Considering only the W terms for now (the V terms fall out in exactly the same way), we can simplify the $\sum_\pi$ of equation Eq. (18) by taking each term (where each term represents one of the four possible permutations considered above) and simplifying them as follows.

Equation (20) may be simplified as

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1,k\neq i}^{n_1}\sum_{l=1,l\neq j}^{n_2} W_{ij}W'_{kl} = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2} W_{ij}W'_{kl} - \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1,k\neq i}^{n_1} W_{ij}W'_{kj} \tag{24}$$

$$-\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{l=1,l\neq j}^{n_2} W_{ij}W'_{il} - \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij}W'_{ij} \tag{25}$$

where the second term may be rewritten as

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1,k\neq i}^{n_1} W_{ij}W'_{kj} = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_1} W_{ij}W'_{kj} - \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij}W'_{ij} \tag{26}$$

and the third term may be rewritten analogously. In the end, we can write Eq. (20) as

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_1}\sum_{l=1}^{n_2} W_{ij}W'_{kl} - \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_1} W_{ij}W'_{kj} - \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{l=1}^{n_2} W_{ij}W'_{il} + \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij}W'_{ij} \tag{27}$$

Equation (21) may be simplified as

$$\begin{aligned}\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{l=1,l\neq j}^{n_2} W_{ij}W'_{il} &= \sum_{i=1}^{n_1}\left(\sum_{j=1}^{n_2}\sum_{l=1}^{n_2} W_{ij}W'_{il}\right) - \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij}W'_{ij} &(28)\\ &= \sum_{i=1}^{n_1}\left(\sum_{j=1}^{n_2} W_{ij}\right)\left(\sum_{l=1}^{n_2} W'_{il}\right) - \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij}W'_{ij} &(29)\\ & &(30)\end{aligned}$$

Likewise, Equation (22) may be simplified as

$$\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1,k\neq i}^{n_1} W_{ij}W'_{kj} = \sum_{j=1}^{n_2}\left(\sum_{i=1}^{n_1} W_{ij}\right)\left(\sum_{k=1}^{n_1} W'_{kj}\right) - \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij}W'_{ij} \tag{31}$$

We can greatly simplify the problem by defining a few new variables:

$$W_1 \equiv \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij} \tag{32}$$

$$W'_1 \equiv \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W'_{ij} \tag{33}$$

$$W_2 \equiv \sum_{i=1}^{n_1}((\sum_{j=1}^{n_2} W_{ij})(\sum_{l=1}^{n_2} W'_{il})) \tag{34}$$

$$\tilde{W}_2 \equiv \sum_{j=1}^{n_2}((\sum_{i=1}^{n_1} W_{ij})(\sum_{k=1}^{n_1} W'_{kj})) \tag{35}$$

$$W_3 \equiv \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} W_{ij}W'_{ij} \tag{36}$$

with corresponding definitions for V_1,V'_1,V_2,$\tilde{V}_2$, and V_3.

With these new variables in hand, we can write the following expression for the covariance matrix for Q.

$$\begin{aligned} E[QQ'] &= \frac{1}{fn_1!n_2!}[f(n_1-2)!(n_2-2)![(W_1W'_1 - W_2 - \tilde{W}_2 + W_3)(V_1V'_1 - V_2 - \tilde{V}_2 + V_3)] \\ &\quad +f(n_1-1)!(n_2-2)![(W_2 - W_3)(V_2 - V_3)] \\ &\quad +f(n_1-2)!(n_2-1)![(\tilde{W}_2 - W_3)(\tilde{V}_2 - V_3)] \\ &\quad +f(n_1-1)!(n_2-1)!W_3V_3]. \end{aligned}$$

Note the analogous structure as derived for the single point type analysis in Diggle et al. (1995). This may be plugged back into Eq. (17) to give the final expression

$$\begin{aligned} Cov(\hat{K}(s,t),\hat{K}(s',t')) &= \left(\frac{A|T|}{fn_1n_2}\right)^2 \frac{1}{n_1!n_2!}[(n_1-2)!(n_2-2)![(W_1W'_1 - W_2 - \tilde{W}_2 + W_3) * \\ &\quad (V_1V'_1 - V_2 - \tilde{V}_2 + V_3)] + (n_1-1)!(n_2-2)![(W_2 - W_3)(V_2 - V_3)] \\ &\quad +(n_1-2)!(n_2-1)![(\tilde{W}_2 - W_3)(\tilde{V}_2 - V_3)] + (n_1-1)!(n_2-1)!W_3V_3] \\ &\quad -\frac{1}{n_1^2n_2^2}W_1V_1W'_1V'_1. \end{aligned} \tag{37}$$

Note that the fraction f, which accounts for the restriction to causal point pairs, cancels everywhere except for in the overall scale factor in the front. The variance can be calculated directly from the covariance in Equation (37).