Developing Survey Grids To Substantiate Freedom From Exotic Pests

John W. Coulston[1], Frank H. Koch[2], William D. Smith[3], and Frank J. Sapio[4]

Abstract.—Systematic, hierarchical intensification of the Environmental Monitoring and Assessment Program hexagon for North America yields a simple procedure for developing national-scale survey grids. In this article, we describe the steps to create a national-scale survey grid using a risk map as the starting point. We illustrate the steps using an exotic pest example in which the purpose of the survey is to substantiate freedom from the pest.

Introduction

Exotic insects and pathogens pose a serious threat to forests in the United States. Damage and control costs for these pests have been estimated to exceed $4.3 billion annually (Pimentel *et al.* 2006). Moreover, these pests may cause severe ecological effects, in some cases virtually eliminating important tree species. In a recent example, the emerald ash borer (*Agrilus planipennis*), first detected in the United States in 2002, has killed more than 20 million ash (*Fraxinus* sp.) trees in Michigan, Ohio, and Indiana and continues to expand its range (USDA Forest Service *et al.* 2006). Exotic forest pests are commonly introduced to the United States through imports of raw wood products, live plants, and other commodities or in packing materials. If an introduced pest becomes established in a particular location, it may expand to new areas by a variety of human-mediated pathways (e.g., interstate shipment of contaminated nursery plants) and by its natural dispersal ability. Pests that are currently found only in limited distribution in portions of the United States but with the potential for major impact if spread elsewhere, include *Phytophthora ramorum* (which causes sudden oak death) and the sirex woodwasp (*Sirex noctilio*).

The 1999 Presidential Executive order establishing the National Invasive Species Council stated that Federal agencies have several critical duties in minimizing the impacts of exotic pests. In particular, these agencies—including the Forest Service, U.S. Department of Agriculture—are mandated to detect and rapidly respond to populations of invasive species in a cost-effective, environmentally sound manner and to perform reliable and accurate monitoring of invasive species populations (Clinton 1999). Fulfilling these mandates, especially at broad geographic scales, requires the establishment of protocols that address the detection, assessment, and monitoring phases of exotic pest surveillance. In the detection phase, we are generally interested in determining whether the exotic pest exists outside its original introduction area. If the pest exists outside the area of original infestation, then it may be important to assess the infestation's geographic extent; in particular, this assessment may determine if quarantine or other regulatory protocols are required. In areas where the pest of interest is established, monitoring to determine whether the prevalence level is increasing or decreasing is essential. Focusing specifically on the detection phase, key protocols include determining the sample size and reliability of surveys to substantiate freedom from an exotic pest and designing the exotic pest survey in a consistent manner once the sample size is known.

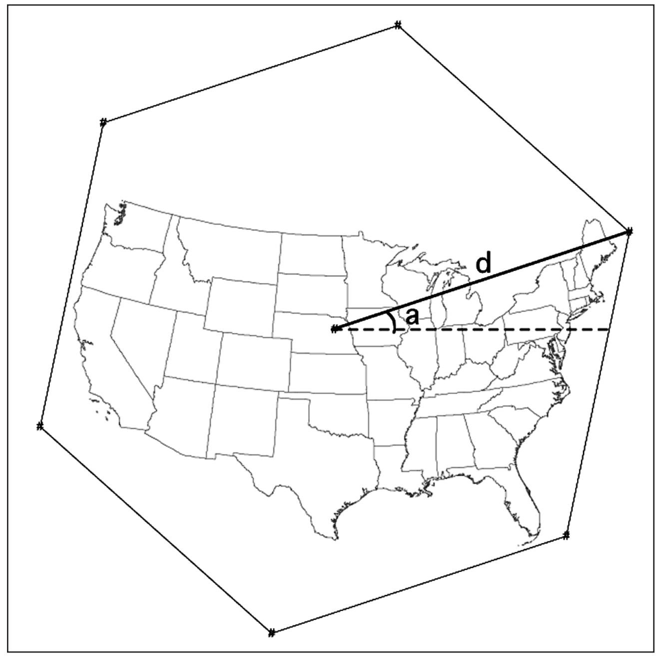
Substantiating freedom from an exotic pest requires that two conditions be defined *a priori*. First, a prevalence threshold must be set. The prevalence threshold sets the detection limit, meaning the minimum detectable level of infestation in any given landscape (e.g., 1 percent, 5 percent). If the prevalence

[1] Supervisory Research Forester, U.S. Department of Agriculture (USDA) Forest Service, Southern Research Station, 4700 Old Kingston Pike, Knoxville, TN 37919. E-mail: jcoulston@fs.fed.us.
[2] Research Associate, Department of Forestry and Environmental Resources, Box 8008, North Carolina State University, Raleigh, NC 27695. E-mail: fkoch@fs.fed.us.
[3] Assessment Coordinator, USDA Forest Service, Southern Research Station, 3041 Cornwallis Road, Research Triangle Park, NC 27709. E-mail: bdsmith@fs.fed.us.
[4] Forest Health Technology Enterprise Team Director, 2150 Centre Avenue, Building A, Suite 331, Fort Collins, CO 80526. E-mail: fsapio@fs.fed.us.

threshold is zero, then a complete enumeration is needed. Second, the desired confidence must be identified. The selected confidence should reflect the required level of certainty (e.g., 95 percent, 99 percent). Once these conditions are defined, "substantiating freedom from disease" has context, and an appropriate method can be developed to estimate the sample size required to, for example, substantiate that an exotic pest of interest does not exist outside its original introduction area, above a 1-percent prevalence threshold, with 99 percent confidence. Coulston *et al.* (2008) developed techniques to estimate required sample size to substantiate freedom from exotic pests based on desired confidence levels and detection thresholds. These sample size estimation techniques were based on existing epidemiological approaches (Cameron and Baldock 1998) extended to the spatial domain. Here we do not address sample size estimation; rather, we describe techniques to develop survey grids in a consistent fashion, based on a global sampling design, once the required sample size has been identified.

White *et al.* (1992) developed a global Environmental Monitoring and Assessment Program (EMAP) sampling grid which serves as the basis for the Forest Service Forest Inventory and Analysis (FIA) Phase 2 (Forest Mensuration) and Phase 3 (Forest Health). The EMAP sampling grid was developed from a truncated icosahedron made up of 20 hexagons and 12 pentagons covering the planet, with one hexagon advantageously placed to cover North America (fig. 1). A noteworthy aspect of the EMAP grid's configuration is that this hexagon can be systematically intensified, yielding a wide range of potential sample frames. In short, this process provides a straightforward framework for creating systematic survey grids. The chief objective with this article is to demonstrate intensification of the EMAP North American hexagon to create a survey grid for substantiating freedom from a hypothetical exotic pest. In support of this objective, we develop a nonlinear regression model to estimate the appropriate intensification factor, provide a table with all possible intensification factors of the EMAP North American hexagon, and describe spatial intersection techniques to implement the sampling grid based on a risk map.

Figure 1.—*The North American hexagon used to generate the Ecological Mapping and Assessment Program sampling grid and other sampling grids.*



Methods

We developed a simulated risk map of susceptibility to a hypothetical forest pest for the conterminous United States. The risk map divides the country into two strata: low risk and high risk; areas within each stratum have equal probability of exposure to a hypothetical exotic pest. For this demonstration, we assumed that 500 sample areas throughout the high-risk stratum would be adequate to substantiate freedom from our hypothetical exotic pest above a 5-percent prevalence threshold with 90 percent certainty. We then intensified the EMAP North American hexagon to develop a sample grid with the appropriate number of points. We identified a final set of sample areas for our survey by spatial intersection of the sample grid and the risk map.

Risk Maps

We constructed the raster risk map (25 km^2 resolution) for our hypothetical pest via spatial overlay of three data layers, representing host species distribution (oak basal area), favorable recent weather conditions (mean January 2003–2005 temperatures), and potential pest spread (based on distance to significant ports). We developed the oak basal area data layer from recent FIA Phase 2 tables for each State, calculating total oak basal area (in m^2ha^{-1}) per Phase 2 plot from corresponding tree tables. We then used inverse distance weighting interpolation to generate a raster map from the (perturbed and swapped) plot points, masking nonforest areas with a forest map developed from MODIS imagery by the Forest Service Remote Sensing Applications Center. We developed the map of mean January temperatures (ºC) using weather observation station data from the National Oceanic and Atmospheric Administration's National Climatic Data Center. For each station in the conterminous United States, we calculated the mean January temperature for 2003–2005. We then created a raster map from the station data using gradient-plus-inverse distance squared interpolation, which combines multiple linear regression (with x, y, and elevation as potential explanatory variables) and distance weighting (Nalder and Wein 1998). For the potential pest spread layer, we first identified a set of "high-risk" U.S. marine ports (i.e., ports receiving raw wood products or packing materials from East Asia) based on U.S. Army Corps of Engineers foreign marine cargo data. We then constructed a national raster map depicting, for each pixel, the Euclidean distance to the closest port. By fitting a beta distribution (α=1.25, β=4) to the full range of distances in this map, we generated a second raster map of simulated introduction probabilities, with probability peaking at ~100 km from the closest port. In the risk map created from these three layers, the high-risk stratum was defined as forested areas with greater than 2.296 m^2ha^{-1} of oak basal area, mean January temperature greater than –9.44 °C, and simulated introduction probability greater than 0.2.

Estimating the Intensification Factor

The EMAP North American hexagon has an area of ~ $1.79 \cdot 10^7$ km^2 and is defined by six vertices and a center point. The hexagon can be intensified by 3, 4, and 7 or any product of these factors (see the Intensifying the North American Hexagon section for details). By intensifying in this manner, tessellations of smaller, equal-area hexagons are created and the required survey grid density can be achieved (table 1). Based on table 1, we developed a model to estimate the intensification factor required to meet the desired number of sample areas for a given survey. We used the following equation and nonlinear regression to develop the model:

Table 1.—*The associated hexagon area for each intensification factor of the North American hexagon (Spence and White 1992).*

Intensification factor	Hexagon area (km^2)
144	40,600
324	18,000
576	10,200
1,296	4,500
2,304	2,500
5,184	1,100
9,216	640
11,664	500
16,384	350

$$X = e^{\left(a - \ln\left(\frac{A}{na}\right)\right)} + \varepsilon, \tag{1}$$

where

X = the nominal intensification factor,

a = estimated parameter,

A = the total area (km^2) of the risk stratum of interest from the risk map,

na = the number of sample areas required, and

ε = error.

The value of *A/na* is the area that each intensification point represents and is analogous to hexagon area in table 1.

The idea behind equation (1) is to answer the following question: if one point represents $1.79 \cdot 10^7$ km^2 (this is the center point of the North American hexagon), then how many times should the sample be intensified so that each point represents *A/na*? By rearranging equation (1) we have simplified the relationship to the following:

$$X = e^a\left(\frac{na}{A}\right) + \varepsilon = a\left(\frac{na}{A}\right) + \varepsilon.$$

To apply this model, once the nominal intensification factor is calculated, it is compared to the intensification factors in table 2, which lists all possible factor sequences (i.e., all possible products of the factors 3, 4, and 7) up to an intensification factor of 50,176. The user then selects an 'actual' intensification factor from table 2, typically the one that is closest to the nominal intensification factor.

Table 2.—*Intensification factor and sequence used to generate sampling grids from the North American hexagon.*

Intensification factor	Sequence	Intensification factor	Sequence	Intensification factor	Sequence	Intensification factor	Sequence
3	–	448	4•4•4•7	3,969	3•3•3•3•7•7	16,384	4•4•4•4•4•4•4
4	–	567	3•3•3•3•7	4,032	3•3•4•4•4•7	16,464	3•4•4•7•7•7
7	–	576	3•3•4•4•4	4,096	4•4•4•4•4•4	16,807	7•7•7•7•7
9	3•3	588	3•4•7•7	4,116	3•4•7•7•7	19,683	3•3•3•3•3•3•3•3•3
12	3•4	729	3•3•3•3•3•3	5,103	3•3•3•3•3•3•7	20,412	3•3•3•3•3•3•4•7
16	4•4	756	3•3•3•4•7	5,184	3•3•3•3•4•4•4	20,736	3•3•3•3•4•4•4•4
21	3•7	768	3•4•4•4•4	5,292	3•3•3•4•7•7	21,168	3•3•3•4•4•7•7
27	3•3•3	784	4•4•7•7	5,376	3•4•4•4•4•7	21,504	3•4•4•4•4•4•7
28	4•7	972	3•3•3•3•3•4	5,488	4•4•7•7•7	21,609	3•3•7•7•7•7
36	3•3•4	1,008	3•3•4•4•7	6,561	3•3•3•3•3•3•3•3	21,952	4•4•4•7•7•7
48	3•4•4	1,024	4•4•4•4•4	6,804	3•3•3•3•3•4•7	26,244	3•3•3•3•3•3•3•3•4
49	7•7	1,029	3•7•7•7	6,912	3•3•3•4•4•4•4	27,216	3•3•3•3•3•4•4•7
63	3•3•7	1,296	3•3•3•3•4•4	7,056	3•3•4•4•7•7	27,648	3•3•3•4•4•4•4•4
64	4•4•4	1,323	3•3•3•7•7	7,168	4•4•4•4•4•7	27,783	3•3•3•3•7•7•7
81	3•3•3•3	1,344	3•4•4•4•7	7,203	3•7•7•7•7	28,224	3•3•4•4•4•7•7
84	3•4•7	1,372	4•7•7•7	8,748	3•3•3•3•3•3•3•4	28,672	4•4•4•4•4•4•7
108	3•3•3•4	1,701	3•3•3•3•3•7	9,072	3•3•3•3•4•4•7	28,812	3•4•7•7•7•7
112	4•4•7	1,728	3•3•3•4•4•4	9,216	3•3•4•4•4•4•4	34,992	3•3•3•3•3•3•3•4•4
144	3•3•4•4	1,764	3•3•4•7•7	9,261	3•3•3•7•7•7	35,721	3•3•3•3•3•3•7•7
147	3•7•7	1,792	4•4•4•4•7	9,408	3•4•4•4•7•7	36,288	3•3•3•3•4•4•4•7
189	3•3•3•7	2,187	3•3•3•3•3•3•3	9,604	4•7•7•7•7	36,864	3•3•4•4•4•4•4•4
192	3•4•4•4	2,268	3•3•3•3•4•7	11,664	3•3•3•3•3•3•4•4	37,044	3•3•3•4•7•7•7
196	4•7•7	2,304	3•3•4•4•4•4	11,907	3•3•3•3•3•7•7	37,632	3•4•4•4•4•7•7
243	3•3•3•3•3	2,352	3•4•4•7•7	12,096	3•3•3•4•4•4•7	38,416	4•4•7•7•7•7
252	3•3•4•7	2,401	7•7•7•7	12,288	3•4•4•4•4•4•4	45,927	3•3•3•3•3•3•3•3•7
256	4•4•4•4	2,916	3•3•3•3•3•3•4	12,348	3•3•4•7•7•7	46,656	3•3•3•3•3•3•4•4•4
324	3•3•3•3•4	3,024	3•3•3•4•4•7	12,544	4•4•4•4•7•7	47,628	3•3•3•3•3•4•7•7
336	3•4•4•7	3,072	3•4•4•4•4•4	15,309	3•3•3•3•3•3•3•7	48,384	3•3•3•4•4•4•4•7
343	7•7•7	3,087	3•3•7•7•7	15,552	3•3•3•3•3•4•4•4	49,152	3•4•4•4•4•4•4•4
432	3•3•3•4•4	3,136	4•4•4•7•7	15,876	3•3•3•3•4•7•7	49,392	3•3•4•4•7•7•7
441	3•3•7•7	3,888	3•3•3•3•3•4•4	16,128	3•3•4•4•4•4•7	50,176	4•4•4•4•4•7•7

Intensifying the North American Hexagon

To develop a method for systematically constructing survey plot networks, we first generated a triangular grid of points from the original seven points defining the North American hexagon (fig. 1). The geometric properties of the triangular grid allow for intensifications of 3, 4, 7, or any product of these factors. For implementation purposes, the North American hexagon can also be directly intensified by a factor of 9, which is a special case of the 3x3 intensification. The intensification grids are created by spawning additional points at regularly spaced distances from points in the initial grid. Conceptually, if an initial point is located at (0, 0) on the unit circle in Cartesian space, then additional points are spawned as follows:

1. 3X: add 2 points at locations (0, –1/√3) and (–1/2, –1/2√3).
2. 4X: add 3 points at locations (1/4, –√3/4), (–1/4, –√3/4), and (–1/2, 0).
3. 7X: add 6 points at locations (5/14, √3/14), (2/7, –√3/7), (–1/14, –3√3/14), (–5/14, –√3/14), (–2/7, √3/7), and (–2/7, √3/7).
4. 9X: add 8 points at locations (1/6, –1/2√3), (1/3, –1/√3), (0, –1/√3), (–1/3, –1/√3), (–1/2, –1/2√3), (–2/3, 0), (–1/3, 0), and (–1/6, –1/2√3).

As the factor sequences in table 2 suggest, grids for larger intensification factors are created by first spawning points from the original seven-point grid at one of the basic factor levels (i.e., 3X, 4X, or 7X), then intensifying the resulting point grid again at one of these factor levels, and so on.

Operationally survey plot networks should only be developed using projected coordinates. Spence and White (1992) suggested using Lambert's azimuthal equal-area projection when working with the EMAP grid since it minimizes scale distortion. For our hypothetical pest example, we assumed Lambert's azimuthal equal-area projection with the following parameters: North American Datum of 1983, GRS80 spheroid, distance units in meters, and 6,378,137 m sphere radius; the center of this projection is –96° 00' 00" longitude and 37° 30' 00" latitude.

Two additional pieces of information are necessary for intensification: the angle α at which the grid deviates from the East-West direction and the distance δ between grid points. For the seven-point grid that defines the North American hexagon, given the aforementioned projection parameters, $\alpha \approx 18.8817°$ and $\delta \approx 2628774.8$ m (fig. 1).

In the following example of 3X intensification, steps that are ultimately applied to every point in an initial grid are illustrated in terms of one point. The point (x_0, y_0) is first centered on the unit circle in Cartesian space at (x_c, y_c) by $x_c = x_0 – x_0$ and $y_c = y_0 – y_0$. Two new points, $(x_{c1'}, y_{c1'})$ and $(x_{c2'}, y_{c2'})$, are then added at $\delta\cdot(0, –1/\sqrt{3})$ and $\delta\cdot(–1/2, –1/2\sqrt{3})$, respectively. The point $(x_{c1'}, y_{c1'})$ is rotated to match α by adjusting the coordinates as follows: $x_{c1} = x_{c1'}\cos(\alpha) - y_{c1'}\sin(\alpha)$ and $y_{c1} = x_{c1'}\sin(\alpha) + y_{c1'}\cos(\alpha)$. The point $(x_{c2'}, y_{c2'})$ is rotated in the same manner. The points are then shifted back to geographic space by $x_{01} = x_{c1} + x_0$ and $y_{01} = y_{c1} + y_0$ and $x_{02} = x_{c2} + x_0$ and $y_{02} = y_{c2} + y_0$. (See White *et al.* 1992 for additional information.) Once the grid has been intensified to the target density, hexagons are created by Thiessen expansion (i.e., points in the grid serve as hexagon centroids).

Spatial Intersection

After determining the intensification factor and creating an intensified point grid corresponding to the high-risk stratum for our hypothetical pest, we used spatial intersection to extract only those grid points that coincided with the high-risk stratum in our raster risk map. We accomplished this using the “SAMPLE” command in ArcInfo®, which extracts the value of a raster map pixel for each sample point. If a sample point from the high-risk grid fell within a high-risk pixel in the raster map, it was kept; otherwise, it was dropped from the final set of points. The final sampling hexagons were then the set of hexagons derived from Thiessen expansion of the retained sample points.

Results

Equation (1) was fit to the data in table 1 using PROC NLIN (SAS Institute, Inc. 2004). The final form of the model was

$$X = 5783883\left(\frac{na}{A}\right) \qquad (2)$$

The only estimated parameter was *a*, which was 5,783,883 with a standard error of 23,123 and a 95-percent confidence interval of 5,730,560 to 5,837,206. Figure 2 displays the model and original data points; notably, this model can be applied to estimate the nominal intensification factor for any future survey grid design efforts.

The final risk map had 920,975 km^2 of high-risk area and 6,857,225 km^2 of low-risk area. As previously noted, our goal was approximately 500 plots in the high-risk stratum. By applying equation (2), the nominal intensification factor for the high-risk stratum was 3,140. The actual intensification factor was 3,136 (table 2), which was implemented by intensifying $4^3\cdot7^2$. This intensification resulted in a systematic triangular grid of 4,093 points in the North American hexagon (fig. 3a). After a risk map value was extracted for each point and all points that did not fall in high-risk areas were deleted, 493 grid points remained (fig. 3b). The stage one sampling hexagons were the Thiessen polygons representing the 493 grid points (fig. 3c).

Figure 2.—*Results from the nonlinear regression. The squares represent data points and the solid line represents equation (2).*

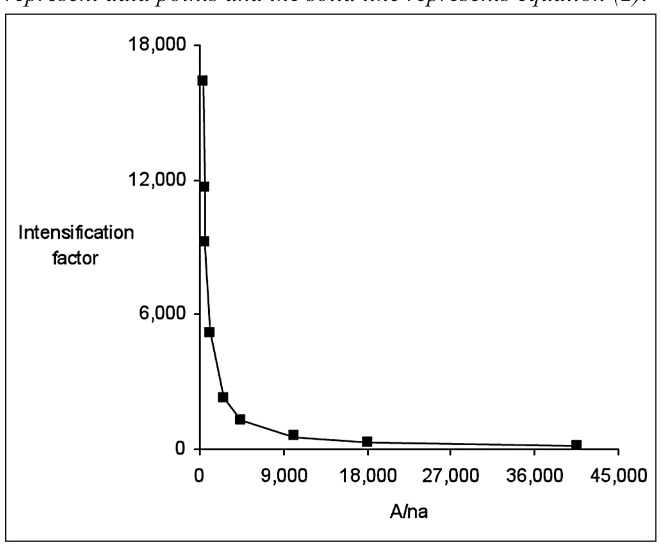
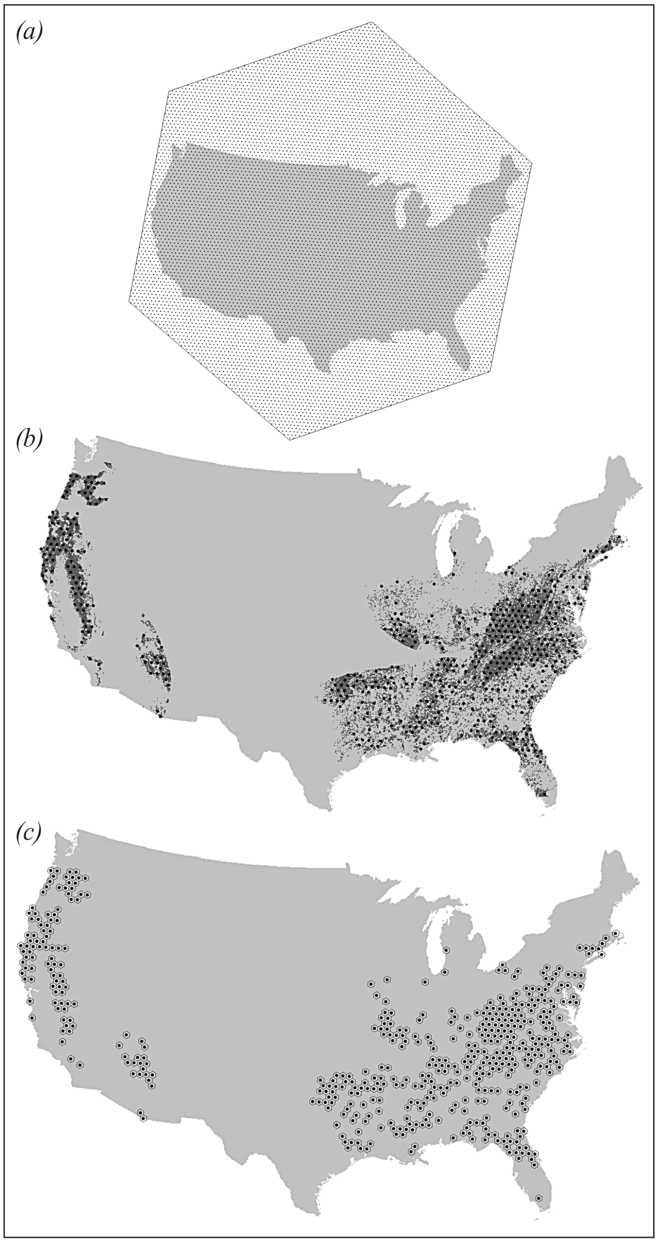


Figure 3.—*(a) The North American hexagon intensified 3,136 X; (b) the intersection of the high-risk stratum and hexagon centers resulting from the intensification; (c) final stage one sampling hexagons for the example survey.*



Discussion

The purpose of the survey grid methodology described here is to substantiate freedom from exotic pests. In general, a systematic grid will be adequate for determining whether an exotic pest of interest exists at specified levels outside the area of original detection. Certain situations exist, however, where a systematic grid may not be adequate. For example, when an exotic pest is expected to be associated with riparian areas in the conterminous United States, the approach presented here may not yield an appropriate distribution of sample areas. In situations such as this, we recommend following the guidelines of Stevens and Olsen (2004) to develop stage one sample areas based on restricted randomization.

In this article, we focused on the development of stage one sample areas for a single stratum survey. The outlined techniques are easily adaptable, however, to multistage samples where two or more risk strata exist. Suppose that our goal was to develop a survey grid that had 500 sample areas in the high-risk stratum and 50 sample areas in the low-risk stratum. To develop the multistage survey, we would estimate the intensification factor for each stratum using equation (2) and then select the actual intensification factor from table 2. A grid would be created for each of the two strata and intersected with the risk map. The final survey grid would then be the compilation of points from each grid that intersected the corresponding stratum in the risk map.

The spatial structure of risk maps plays a vital role in the development of useful survey grids. Risk maps with little within-strata spatial autocorrelation have a "salt-and-pepper" appearance. This autocorrelation will result in a survey grid with a similar salt-and-pepper appearance. Risk maps with high within-strata spatial autocorrelation, however, will have a smooth appearance where the likelihood that areas from the same stratum are adjacent is high. The appearance of both the risk map and the survey grid is partially related to the spatial resolution (i.e., pixel size) of the risk map and the data layers that were used to develop it. For the purposes of designing a national scale survey to substantiate freedom from an exotic pest, we suggest a pixel size of approximately 25 km^2 or larger to limit the influence of pixel size on survey grid design.

As global trade and the influx of foreign goods to the United States escalate, the probability that new exotic pests will be accidentally introduced also increases. If unchecked, some of these exotic pests are likely to become established and, in turn, impact the Nation's forest resources. The techniques described here offer a relatively simple way to monitor such emerging forest pest threats. Significantly, the foundation for these techniques is a global sampling design that is already being used by FIA; indeed, its capacity for systematic, hierarchical intensification may already be familiar to many scientists. Furthermore, adoption of standardized approaches like the one outlined here enables us to rapidly and more efficiently respond to exotic pest threats and, thus, better fulfill our current mandate in this regard.

Literature Cited

Cameron, A.R.; Baldock, F.C. 1998. Two-stage sampling in surveys to substantiate freedom from disease. Preventive Veterinary Medicine. 34: 19-30.

Clinton, W.J. 1999. Executive Order 13112 of February 3, 1999: Invasive Species. Federal Register. 64(25): 6183-6186.

Coulston, J.W.; Koch, F.H.; Smith, W.D.; Sapio, F.J. 2008. Exotic pest surveillance: survey development and reliability. Canadian Journal of Forest Research: 2422-2433.

Nalder, I.A.; Wein, R.W. 1998. Spatial interpolation of climatic normals: test of a new method in the Canadian boreal forest. Agricultural and Forest Meteorology. 92: 211-225.

Pimentel, D.; Zuniga, R.; Morrison, D. 2006. Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics. 52: 273-288.

SAS Institute, Inc. 2004. SAS/STAT® 9.1 User's Guide. Cary, NC: SAS Institute, Inc.

Spence, M.H.; White, D. 1992. EMAP sampling grid technical report. Corvallis, OR: ManTech Environmental Technology, Inc.; U.S. Environmental Protection Agency, Environmental Research Laboratory. 64 p.

Stevens, D.L.; Olsen, A.R. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association. 99(465): 262-278.

U.S. Department of Agriculture (USDA) Forest Service; Michigan Department of Agriculture; Michigan Department of Natural Resources; USDA Animal and Plant Health Inspection Service; Michigan State University; Purdue University; Ohio State University. 2006. Emerald ash borer. http://www.emeraldashborer.info/. (07 November).

White, D.; Kimerling, J.A.; Overton, W.S. 1992. Cartographic and geometric components of a global sampling design for environmental monitoring. Cartography and Geographic Information Systems. 19(1): 5-22.