The Power of Crown-Indicator Variables to Detect Change

William A. Bechtold[1] KaDonna C. Randolph[2] Stanley J. Zarnoch[3]

ABSTRACT: *The goal of FHM Detection Monitoring is to identify forest ecosystems where conditions might be deteriorating in subtle ways over large areas. This study applies statistical power analysis to FHM Crown-Indicator variables to determine how many plots are necessary to detect various degrees of change at various levels of statistical power. Results show that the base Phase 3 sampling intensity provides sufficient power to support analysis at the regional level, but not necessarily at the state level.*

The objective of this study is to apply statistical power analysis to the Forest Inventory and Analysis (FIA) Phase 3 (P3) Crown Indicator to determine how many plots are necessary to detect meaningful change. The Crown Indicator includes three main variables recorded for trees at least 5.0-inches dbh on P3 plots—Crown Density (CDEN), Foliage Transparency (FTRAN), and Crown Dieback (CDBK). CDEN is the amount of crown biomass that blocks light penetration through the crown. FTRAN is the amount of skylight visible through small holes in the live portion of the crown where foliage normally occurs. CDBK is recent mortality of branches with fine twigs that begins at the terminal portion of a branch and proceeds inward toward the trunk. All three variables are recorded as percentages. More information about the Crown Indicator is available in Schomaker et al. (2007) and at the web site http://srsfia2.fs.fed.us/crowns/.

Statistical power is based on hypothesis testing. Two types of error are associated with hypothesis tests. **Type I error** (α) **is a false positive**, where the null hypothesis is rejected and the test incorrectly concludes there is some significant effect. **Type II error** (β) **is a false negative**, where the test fails to detect a true problem. **The *power* of a test is defined as** ($1-\beta$)**; it is the probability of correctly detecting a meaningful difference.** When monitoring forest health, the consequences of a false positive might be costly in terms of launching unnecessary follow-up studies, but a Type II Error (failure to recognize a problem) could be disastrous. For Detection Monitoring, it's important to know if the sampling intensity of an indicator is sufficient to support acceptable alpha and beta probability levels.

Power Analysis

We used the SAS POWER procedure to accomplish this analysis (SAS Institute Inc. 2004). The TWOSAMPLEMEANS option was applied to determine the number of plots necessary to detect a difference between two independent samples (such as two different regions or two different panels). The PAIREDMEANS option was used to determine the number of plots necessary to detect a difference using *paired* observations (such as survivor trees from remeasured plots). The required inputs are:

- the number of observations (n),
- the target alpha probability level (α) ,
- the target power level ($1-\beta$),
- the target effect size (mean difference between groups),
- the standard deviation (s)
  (or the coefficient of variation (cv) for log normal data),
- the underlying frequency distribution
  (normal or log normal), and
- a correlation coefficient (r), which is required for the
  PAIREDMEANS option only.

SAS will solve for whichever variable is not specified, so for this analysis n was set to null. The standard deviations, underlying data distributions, and correlation coefficients are estimated (usually from a pilot study) while the other input variables are simply specified by the analyst.

We used data from existing plots to estimate the ranges of means, standard deviations, correlation coefficients, and data frequency distributions expected for each crown variable. These values were then plugged into the power analysis. Table 1 shows the number of plots needed to detect changes in the crown variables for numerous combinations of input specifications. We set the power level ($1-\beta$) to alternate from 0.8 to 0.9, and α from 0.01 to 0.05. Effect sizes were set to specify differences in the crown variables that were judged to be biologically meaningful.

The last row listed for each variable in Table 1 (highlighted in yellow) shows the input specifications that best fit each of the three crown variables based on existing crown data. Fixing the power level at .9 and alpha at .05, we concluded that 63, 74, and 130 independent observations would be needed to evaluate plausible scenarios involving CDEN, FTRAN, and CDBK, respectively. For paired observations (i.e., remeasured plots) it would take at least 21, 31, and 36 plot pairs. These last three numbers were used to produce the maps described below.

Table 1. Numbers of independent and paired observations required to detect a statistically significant signal, given the listed combinations of input specifications for 3 crown variables (crown density, foliage transparency, and crown dieback).

Variable	Power analysis input specifications					Plots required		
						Independent	Paired[c,d]	
	Data Distribution	Power level (1-*B*)	Alpha level (*a*)	Data variability[a]	Effect size[b]	n plots	n plots (r=.25)	n plots (r=.50)
Crown density	Normal	0.8	0.01	10	10	52	21	16
		0.8	0.01	10	15	26	12	9
		0.8	0.01	15	10	110	43	30
		0.8	0.01	15	15	52	21	16
		0.8	0.05	10	10	34	14	10
		0.8	0.05	10	15	18	8	6
		0.8	0.05	15	10	74	29	20
		0.8	0.05	15	15	34	14	10
		0.9	0.01	10	10	64	26	19
		0.9	0.01	10	15	30	14	11
		0.9	0.01	15	10	138	54	37
		0.9	0.01	15	15	64	26	19
		0.9	0.05	10	10	46	18	13
		0.9	0.05	10	15	22	10	7
		0.9	0.05	15	10	98	38	26
		0.9	0.05	15	15	46	18	13
		0.9	0.05	12	10	63	21 (r=.40)	
Foliage Transparency	Normal	0.8	0.01	5	5	52	21	16
		0.8	0.01	5	10	16	8	7
		0.8	0.01	10	5	192	74	51
		0.8	0.01	10	10	52	21	16
		0.8	0.05	5	5	34	14	10
		0.8	0.05	5	10	12	6	5
		0.8	0.05	10	5	128	50	34
		0.8	0.05	10	10	34	14	10
		0.9	0.01	5	5	64	26	19
		0.9	0.01	5	10	20	10	8
		0.9	0.01	10	5	242	93	63
		0.9	0.01	10	10	64	26	19
		0.9	0.05	5	5	46	18	13
		0.9	0.05	5	10	14	7	5
		0.9	0.05	10	5	172	65	44
		0.9	0.05	10	10	46	18	13
		0.9	0.05	6.5	5	74	31 (r=.20)	
Crown dieback	Log normal	0.8	0.01	1.5	2	118	39	25
		0.8	0.01	1.5	2.5	70	24	16
		0.8	0.01	2.0	2	160	48	29
		0.8	0.01	2.0	2.5	94	29	18
		0.8	0.05	1.5	2	80	26	16
		0.8	0.05	1.5	2.5	48	16	11
		0.8	0.05	2.0	2	108	32	19
		0.8	0.05	2.0	2.5	64	20	12
		0.9	0.01	1.5	2	150	49	30
		0.9	0.01	1.5	2.5	88	30	19
		0.9	0.01	2.0	2	204	61	36
		0.9	0.01	2.0	2.5	118	36	22
		0.9	0.05	1.5	2	106	35	21
		0.9	0.05	1.5	2.5	62	21	13
		0.9	0.05	2.0	2	144	43	25
		0.9	0.05	2.0	2.5	84	25	15
		0.9	0.05	1.8	2.0	130	36 (r=.30)	

[a] For variables with a normal distribution the measure of data variability is the standard deviation. For variables with a log normal distribution the measure of data variability is the coefficient of variation.
[b] For variables with a normal distribution the effect size is specified as the difference between two means. For variables with a log normal distribution the effect size is specified as the ratio between two means.
[c] Paired observations require the additional specification of a correlation coefficient (r). These results show the numbers of observations required when the correlation coefficient is set to 0.25 and 0.50. The last row for each variable shows correlation coefficients and other statistics that most closely match existing data.
[d] The numbers of plots listed represent pairs. Each pair represents 1 plot (i.e., grid point) with two observations.

Detectable Impact Area

Once the target number of plots is known, the following formula can determine how much area must be impacted by a forest health problem for the plot network to detect it:

$$IA = n(E)\frac{P_t}{P_n}$$ Equation (1)

where

- IA = the minimum size of an impact area detectable by the plot network in the area of interest,
- n = the minimum number of plots needed to detect an impact (from the power analysis),
- E = the plot expansion factor in the area of interest (6,000 for P2 or 96,000 for P3),
- P_t = the total number of panels per measurement cycle in the area of interest, and
- P_n = the number of available panels in the area of interest.

Suppose we want to know if the FIA P3 plot grid is sufficient to detect a change in CDBK from remeasured (paired) plots, given a 5-panel system with 5 panels of available data. The most plausible scenario for CDBK indicates that 36 paired plots are necessary to detect the doubling of CDBK specified in the last row of Table 1, so:

$n = 36$ $E = 96{,}000$ $P_t = 5$ $P_n = 5.$

Given this input, Equation 1 indicates sufficient power to detect an impact affecting 3.5 million acres of forest. The 21 and 31 paired plots needed to identify meaningful changes in CDEN and FTRAN result in detectable impact areas of 2 and 3 million acres, respectively.

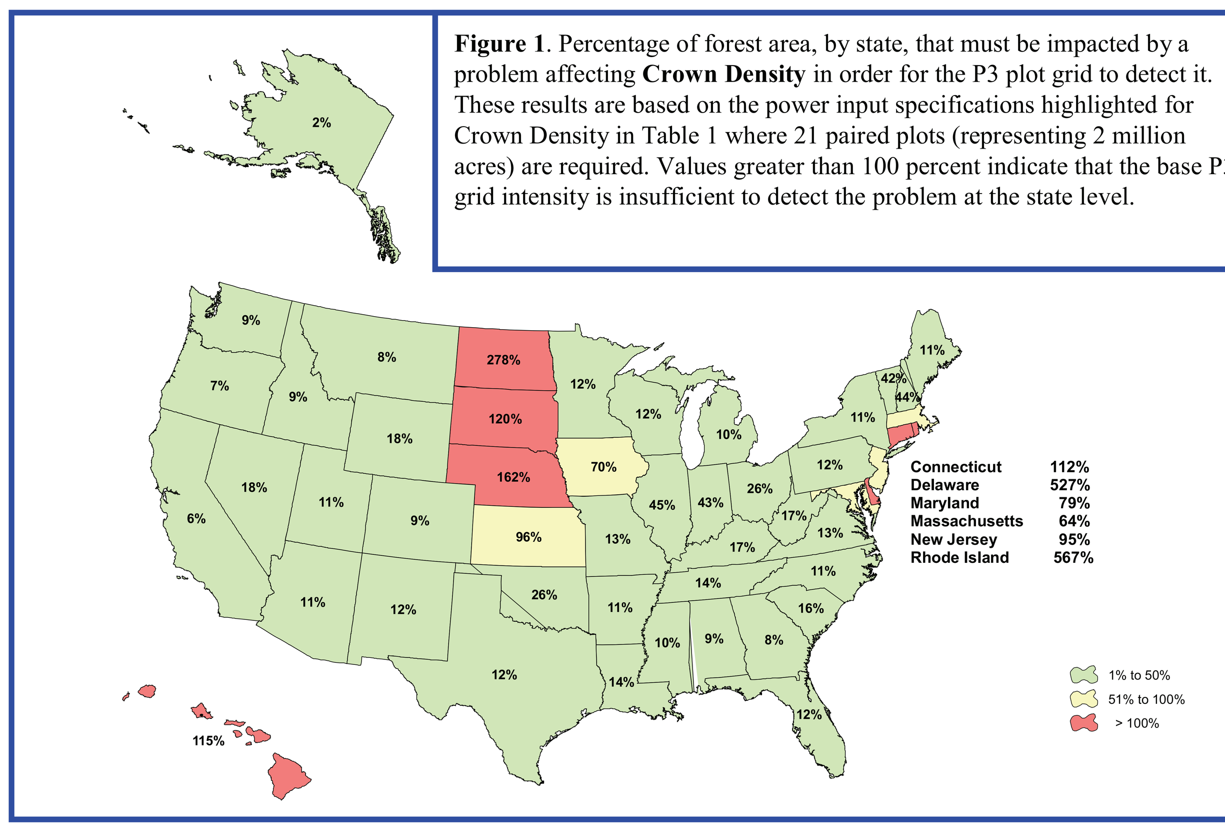

Figure 1. Percentage of forest area, by state, that must be impacted by a problem affecting **Crown Density** in order for the P3 plot grid to detect it. These results are based on the power input specifications highlighted for Crown Density in Table 1 where 21 paired plots (representing 2 million acres) are required. Values greater than 100 percent indicate that the base P3 grid intensity is insufficient to detect the problem at the state level.

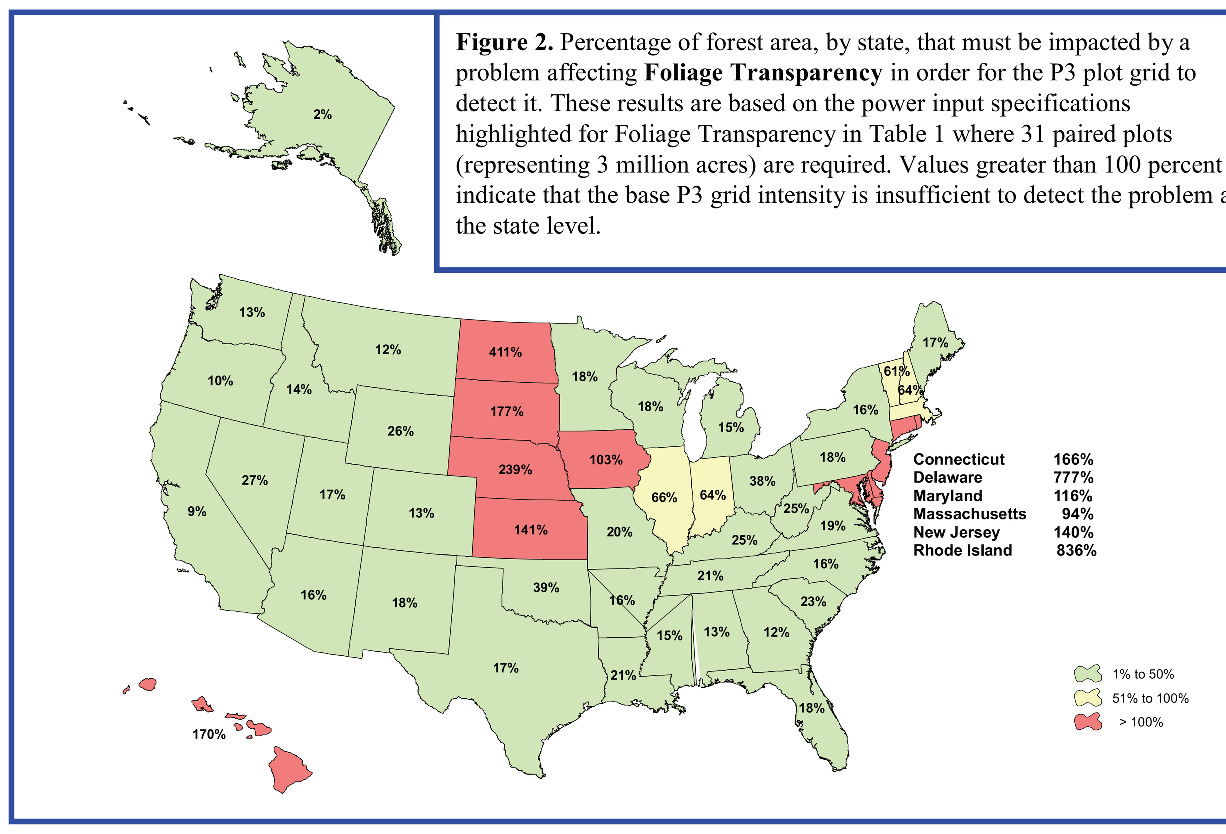

Figure 2. Percentage of forest area, by state, that must be impacted by a problem affecting **Foliage Transparency** in order for the P3 plot grid to detect it. These results are based on the power input specifications highlighted for Foliage Transparency in Table 1 where 31 paired plots (representing 3 million acres) are required. Values greater than 100 percent indicate that the base P3 grid intensity is insufficient to detect the problem at the state level.

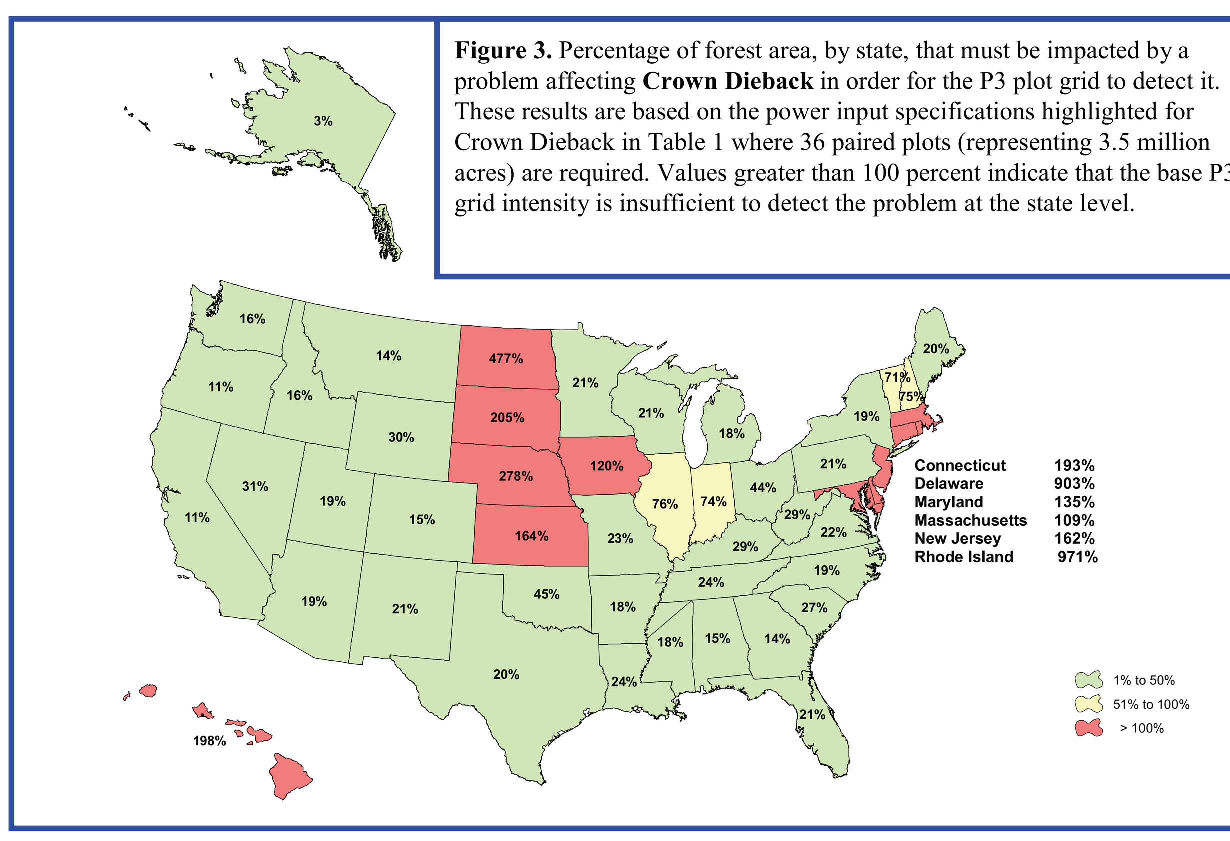

Figure 3. Percentage of forest area, by state, that must be impacted by a problem affecting **Crown Dieback** in order for the P3 plot grid to detect it. These results are based on the power input specifications highlighted for Crown Dieback in Table 1 where 36 paired plots (representing 3.5 million acres) are required. Values greater than 100 percent indicate that the base P3 grid intensity is insufficient to detect the problem at the state level.

If we divide the detectable impact area by the forest acreage in an area of interest, we can see what percentage of the forest in that region must be impacted to detect the specified change. Figures 1-3 show the results by state. Wherever the listed number is more than 100 percent, the plot network is not sufficient to detect an impact. For example, the 971% shown for Rhode Island (Figure 3) means that the base plot grid there would have to be increased by about 10x to detect a problem involving dieback. Overall, results from the power analysis confirm that P3 sampling intensity is sufficient to detect changes in tree crown health at the regional level, but there are approximately a dozen states with too few forest plots to support analysis of the Crown Indicator at the state level.

This analysis concentrates on area by state, but Equation 1 can be used to check the plot network for a wide variety of scenarios. For example, one could also look at area by forest type, area by elevation, area by ecoregion, etc.

Strategies to Increase Statistical Power

Increasing sample size by grid intensification is the most obvious way to gain statistical power, if economically feasible. Sample size can also be increased by combining areas, but this must be done carefully. Adding adjacent states will increase power if those areas have the same problem as the original location. If not, the analysis could be compromised because adding unaffected areas will dilute any effect that may have been present at the original location.

Stratification has the potential to increase statistical power. Forest health impacts that have "clumped" distributions display higher variances than impacts that are evenly dispersed. Statistical power is adversely affected by increased variance. Stratification can help by reducing variation within strata. If a clumped distribution is suspected, post-stratify the data to isolate the clumps if possible.

Species groups are another important aspect of stratification. Grouping species together has the potential to boost power by increasing the number of plots available for analysis, but a forest health problem that impacts only a few species in a group may go undetected. If a particular species is known to be sensitive to a prospective threat, that species should be isolated. Care must always be taken to avoid different species mixes between the two groups of observations being tested, because effect size could be an artifact of the difference in species mix.

Power can also be gained by using one-tailed tests, where the alternate hypothesis (H_1) is formulated to check for change in only one direction. When one-tailed tests are substituted for the two-tailed tests used to produce the results in Table 1, the required numbers of plots drop by 15-20 percent.

Conclusions

All things considered, the Crown Indicator is performing as originally expected. About 100 independent observations, or 50 paired observations, would be enough to support most analyses. At the base P3 grid intensity, this is sufficient power for regional analysis, but not for many individual states.

Power analysis requires only three bits of information from actual data: (1) the standard deviation, (2) the underlying data frequency distribution, and (3) a correlation coefficient for paired observations.

Once the target sample size is known, situations where sampling intensity is insufficient can be identified, and options available to correct such situations can be properly quantified.

As new indicators are developed, power analysis should be part of the vetting process.

Additional details related to this analysis are available in Bechtold et al. (in press).

Literature Cited

Bechtold, William A.; Randolph, KaDonna C.; Zarnoch, Stanley J. In press. The power of FIA Phase 3 crown-indicator variables to detect change. In: McWilliams, Will; Moisen, Gretchen; Czaplewski, Ray, comps. 2008. 2008 Forest Inventory and Analysis (FIA) Symposium; October 1-23, 2008. Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 1 CD. pp xx-xx.

SAS Institute Inc. 2004. SAS OnlineDoc® 9.1.3. Cary, NC: SAS Institute Inc.

Schomaker, M.E.; Zarnoch, S.J.; Bechtold, W.A.; Latelle, D.J.; Burkman, W.G.; Cox, S. M. 2007. Crown-condition classification: a guide to data collection and analysis. Gen. Tech. Rep. SRS-102. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 78 p.

1. Research Forester, USDA Forest Service, Southern Research Station, Asheville NC.
2. Mathematical Statistician, USDA Forest Service, Southern Research Station, Knoxville TN.
3. Mathematical Statistician, USDA Forest Service, Southern Research Station, Asheville NC.