Vol 1, No 1, pp. 10–13, 2009
International Journal of
Mathematical and Computational Forestry & Natural-Resource Sciences

AVAILABLE ONLINE AT HTTP://MCFNS.COM

Submitted: Jun. 15, 2008
Accepted: Jun. 18, 2008
Published: Feb. 28, 2009
Last Correction: Mar. 25, 2009

"TOTAL-BALANCING" AN INVENTORY

A method for unbiased inventories using highly biased non-sample data at variable scales

Kim Iles

Kim Iles & Associates Ltd., 412 Valley Place, Nanaimo, BC, Canada. Ph.&FAX: 250.753.8095

Abstract. I describe here a method that can provide unbiased estimates and sampling errors with increasingly precise polygon information from non-sample sources that are often free and readily available. This method is extremely flexible, and it appears to be in line with the main trend of modern sampling – you first estimate using any information available, and then you sample to adjust those estimates. At later dates, further readjustment can be done at will, as long as the total is held constant.

Keywords: Unbiased methods, total-balancing, data adjustment, forest inventory, sampling methods

1 Background

In about 1991, the Province of British Columbia began to put together a modern multi-disciplinary forest inventory[1]. I was appointed as the sampling specialist in charge of the design. The process involved many levels of government, tedious meetings, and all of the difficulties of modern government natural resource management. Some of the debacles and techniques have been published earlier, as noted in the literature section.

The area to be inventoried was approximately 250,000,000 acres, or 100,000,000 hectares. At this scale, with any reasonable number of plot clusters, the distance between sampling positions was very large. Entire mapsheets might be unsampled. The design was intended to withstand legal scrutiny, and statistically clean results were highly desirable. At the same time, the working managers wanted good answers for individual polygons, which was quite reasonable and in line with all the major trends in forest sampling that call for detailed and useful information at a polygon level.

The design used initial estimates that were corrected by a sample of statistically chosen positions with which to calibrate the estimates. All went smoothly until a formal presentation of the process at a conference a few years later for the "little people" who actually ran the forest. Very quickly one of the managers saw that we were preparing to throw away a great deal of data that he had spent many years patiently tending, and with his installation of many plots in local areas of interest. This had been expensive, very useful in his opinion, and at a sampling density several orders of magnitude better than the one a new Provincial inventory could provide. He kindly offered it to us.

With great patience, I explained that we could not use his plots, which were selected with a great deal of bias – albeit in areas of great interest to himself, and perhaps even correctly measured with only slight differences from our standards and definitions. I explained the issues of sampling probabilities, samples of convenience, quality control, definitions and so on, in what I thought was a reasonable way. I think I had explained the issue of unbaisedness, its desirability, its legal advantages and the role of statistical controls pretty clearly. "*You can see the problem, right ?*", I remember saying to this practical, experienced and well meaning individual who had a daily interest in accurate information about his own area and who was offering massive amounts of relevant and local data that I was prepared to reject out of hand. I remember even more clearly when he leaned over the table and said *"I think that the problem is you, bubba"*.

He was right, of course. It is the inventory specialist's job to solve problems and to use data – not to reject information on the basis of simplicity (to him) and his own convenience. No matter how satisfying a set of brand new data might be, sparkling with sharp definitions, clear instructions and flawless quality control - it is not the answer. In a world of massive amounts of data, local concerns and fine resolution decisions, the old methods are just no longer appropriate.

It took less than a week to see the solution and incorporate it. It was highly desirable to have an unbiased system. Perhaps unbiased systems are mainly an emotional advantage, but certainly I was not going to court

[1] BC Ministry of Forests VIWG (1992)

Iles (*2009*) (MCFNS 1(1):10–13). ISSN 1946-7664. Manuscript editor: Chris J. Cieszewski

with methods that I would be obliged to admit in testimony were "clearly biased". Unbiasedness would be maintained. A valid estimate of the sampling error was just as necessary. The question was how to use all this other data and still preserve the advantages of unbaisedness and valid sampling errors.

2 The Solution

The answer was simple. It applies not just to forest inventory projects, but to any sampling situation. I have never been able to find it discussed in the literature, but would certainly want to know if anyone knows of a specific reference to it so that I can appropriately give credit where it is due. The key was to recognize exactly what "unbiased" means, and what the sampling error really refers to. I think this is not well understood by most people in forest inventory, or in the sampling establishment in the larger sense. Indeed, when trolling through the statistical literature it is not clear to me that the definition is actually very clearly defined or well agreed upon. The mathematical statistics books seem to do the best job with definitions of this type, because of their need for mathematical clarity; but even here unbiasedness is not clearly and unanimously defined. As I understand it, unbaisedness refers to totals (and to averages by extension). So do sampling errors.

Any process that comes to the same total as an unbiased system with each repetition *also becomes an unbiased system with the same sampling error* — no matter what its procedures might be. Therefore, if we simply adjust the total of any set of polygons to the same total as any unbiased sample, that adjusted total becomes clearly unbiased and the problem is solved. In addition, such polygons are infinitely adjustable on a local scale. You can subsample areas of interest, use expert judgment, or do anything else you find locally useful. The sampling error does not change either, no matter how any outside information is used. The process of assigning polygon estimates might be extremely complicated and very effective, but the total used might be a simple average of plots on the area, which is easy for the public to understand.

The inventory design remained essentially the same, but was now flexible enough to accommodate any information that was provided, no matter what the source[2]. The manager's data could all be used. We did, of course, make accommodation for some outside sources to be ranked as more accurate, and therefore to be changed less when the total-balancing took place. The total for the Provincial Inventory was to be based strictly on the sample plots managed by the inventory team, and honed vary sharply in terms of definitions, testing and quality control. This process extended to the point of falling statistically chosen trees for the direct measurement of actual net volumes.

Several things happened very quickly. The first was the realization that anyone could contribute to improving the inventory process by any method they chose. Previously, most groups could only sit and watch while a single organization did all of the data gathering and controlled the entire outcome. In this case, only the total was under strict statistical control. The individual polygon values they cared about were highly flexible, and could be much improved by local knowledge, previous work and growth models.

Most of these groups had resented being kept at arms length on previous projects, but nobody seemed to resent the fact that we offered them the chance to change polygon values any way they wanted (this included allowing them to put plots into "typical" parts of stands). We would, however, check the results and there would be a full *and public* disclosure as to how much their contributions helped (or hurt) the inventory quality. In most cases they decamped at that point, but it was their choice and having that choice made a great deal of difference to them.

3 Maintaining the tree characteristics

In the end, the solution was just about as simple and flexible as it could be. We would use any information at all to change polygon values, before or after the project, as long as the polygons added up to the same total. In our case, the design was to total-balance by species. Even categorical variables could be adjusted, and we developed methods for each of the values in the inventory. To maintain the flexibility to adjust or evaluate the inventory in the future all that was needed was to hold back the exact location of the actual field measurements.

Of particular interest was the ability to change species totals in stands without changing the essential characteristics of individual trees. The solution to this was equally simple. Each tree was assigned a fixed plot size or a Variable Plot Basal Area Factor. To vary the volume of a stand by 10%, simply change these plot "sizes" by 10% and recompile. This adjusts the number of trees without changing their individual characteristics. To do the adjustment by species, different correction amounts could be applied by species or species group. Any number of adjustments could be done, as long as one final adjustment was done to maintain the overall total. This also maintained any sampling errors by category. These adjustments can use simple ratios, regressions, or any other adjustment method.

[2] BC Ministry of Forests VIWG (1995)

4 Other Applications

A few years ago I applied this method to an industrial inventory of about 200,000 acres, or 80,000 hectares, on the West Coast. In this case, the company wanted individual stand information in the form of a tree list for every polygon. The same overall process was carried out, but in this case plots from stands that were judged "similar" in tree characteristics were simply gathered and duplicated into each of the vast majority of unsampled stands. Overall species volumes were further adjusted using the BAF of the prism used for the data, as suggested earlier. At the end of the process, the total was adjusted to match the unbiased total of a sparse but correctly executed probability sample for the company (in this case, one was already available for use, although it did not provide adequate polygon information). In one sense, this is similar to the "nearest neighbor" approach, but the key difference is the use of local information, changing polygons at will, and balancing the total to insure unbiasedness and valid statistics.

More recently, the University of Georgia, in connection with the Forest Service FIA group, has used the unbiased total for the well maintained existing grid of FIA plots, then produced polygon level information with non-sample and biased information for much more precise decisions and localized inventory information[3].

5 Summary

Using this method can provide unbiased estimates and sampling errors with increasingly precise polygon information from non-sample sources that are often free and readily available. It is extremely flexible, and it appears to be in line with the main trend of modern sampling – you first estimate using any information available, and then you sample to adjust those estimates. At later dates, further readjustment can be done at will, as long as the total is maintained.

The situation with forest inventory is very similar to mapmaking. For many years the only acceptable method of improving a map was to start over with a fresh sheet of paper and do the entire job again with great fidelity to current map accuracy standards. Those days are over. The same is true of forest inventory. A better concept is "let's just change the parts that are not good enough". The other parts often change so little as not to be noticed.

Stratified cruises with standard descriptions for multiple stands and repeating the process every 20 years while ignoring the existing inventory are old and outdated processes. In an age where information pours down upon us from every direction, it is time we started to use it effectively. It is so easy to assure statistical unbiasedness, good polygon estimates, and valid sampling errors by the process described here that it is hard to imagine why anyone would strike the old forest inventory off the records and independently do it again from a standing start.

[3] Cieszewski et al. (2005)

End Note

The original undedited versions of Iles papers[4] related to this subject are available at http://www.island.net/~kiles/jf_ar.pdf and at http://www.island.net/~kiles/p-boise.pdf, while his forest inventory book[5] Iles (2003) (especially page 248 or page 313) covers much of the related background material.

Acknowledgements

The author wishes to thank an anonymous manager, who was kind enough to teach him some humility, statistics, and practical psychology; and who contributed greatly the design of the BC Vegetation Inventory. As well, thanks are due to two anonymous reviewers who provided constructive comments.

References

BC Ministry of Forests VIWG. 1992. Report of the timber inventory task force on the current timber inventory with recommendations for the future. Pp. 111. http://ilmbwww.gov.bc.ca/risc/o_docs/-teveg/pni/viwg_rpt-08.htm

BC Ministry of Forests VIWG. 1995. Final Report from the Vegetation Inventory Working Group on a Proposed New Inventory. Pp. 101. http://ilmbwww.gov.bc.ca/risc/o_docs/-teveg/pni/viwg_rpt-08.htm

Cieszewski, C.J., K. Iles, R.C. Lowe, and M.J. Zasada. 2005. Proof of concept for an approach to a finer resolution inventory. Page 69 – 74. In R. E. McRoberts, G. A. Reams, P. C. Van Deusen, and W. H. McWilliams. Proceedings of the Fifth Annual Forest Inventory and Analysis Symposium. November 18–20, 2003, New Orleans, LA. Gen. Tech. Rep. WO-69. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station.

Iles, K. 1994. Some Directions for Forest Inventory. Journal of Forestry. Invited lead article for the special edition on Forest Inventory.

[4] Iles (1994, 1998)

Iles, K. 1998. Some Practical Aspects of Designing a Large Inventory. International meeting on Forest Inventory, Boise, Idaho.

Iles, K. 2003. A Sampler of Inventory Topics. Second print, ISBN 097321980-7.

Biographical Note

Kim Iles is a specialist in sampling and forest inventory, and inventor of many new statistical techniques. He received his B.S. in Forest Management and an M.Sc. in Forest Biometrics from Oregon State University, with a Ph.D. Forest Biometrics from the University of British Columbia in 1979. He was the Biometrician and Head of the Growth and Yield Department of Macmillan Bloedel Ltd., Nanaimo, B.C. for 12 years, and since 1991 and has been a consultant. He is the author of a textbook on Forest Inventory[5] (now in second printing), and he was the principle inventory design consultant for the British Columbia Provincial Inventory design (this inventory covered an area of 250,000,000 acres (100,000,000 hectares). Kim Iles specialties of Variable Plot Sampling and 3P sampling are the dominant inventory systems used in North America. He has taught inventory techniques to several thousand professional cruisers on 4 continents, and is the major contributor to the Cruising and Inventory Newsletter published by John Bell & Associates. He has been a member of several committees of national standing, and has developed and introduced a number of innovations for cruising systems in North America.