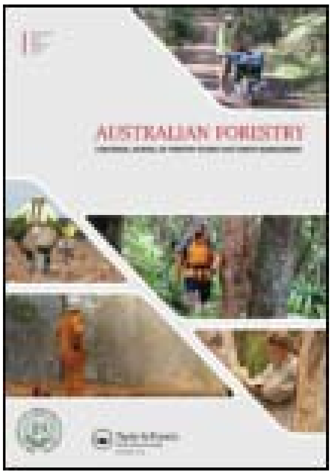

Australian Forestry
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tfor20

Not everything that can be counted counts

Cris Brack[a]

[a] Australian National University
Published online: 22 Aug 2014.

To cite this article: Cris Brack (2014) Not everything that can be counted counts, Australian Forestry, 77:2, 75-77, DOI: 10.1080/00049158.2014.908484

To link to this article: http://dx.doi.org/10.1080/00049158.2014.908484

Australian Forestry, 2014
Vol. 77, No. 2, 75–77, http://dx.doi.org/10.1080/00049158.2014.908484

GUEST EDITORIAL

Not everything that can be counted counts

... not everything that counts can be counted.

(Cameron 1963)

It seems to me that there are two substantial but not necessarily mutually exclusive groups of people. One group fears 'numbers', or at the very least distrusts them. In the words of one of the pioneers of practical expert systems, people 'do not believe in mathematics' (D'Avignon & Winkels 1986, p. 76). A typical response to such fear is to shy away from using mathematical analysis to help make decisions. Former Prime Minister John Howard represents many people in this group when he avows there is no reason to act on climate change because 'in his gut' he does not believe it will be bad. As D'Avignon and Winkels (1986) assert, many managers will rely on their 'own reasoning and intuition' and no numbers will be able to change a gut conclusion.

But those of us in the forestry profession or industry cannot sit back and claim a moral superiority over those who do not accept numbers, for such denial is rife here too. For example, many foresters don't accept the results of inventories. The simplest of these inventories belong to the class of 'design-based' sampling. Design-based approaches essentially do not need to make any assumptions about the objects being measured, except that they are selected randomly. The mathematics is very simple, and estimating the total product available just boils down to a single equation: total volume = sampled volume/probability of selecting that sample. Under a design-based approach, this total volume estimate, as simple as it is calculated, is unbiased and likely to be close to the true volume of the population. How close requires a little more arithmetic, but only an assumption that you have taken a reasonable number of sample measurements—25 or more will probably do it.

At best, however, these estimates tend to be treated as a maximum. Some formal processes, including methodologies proposed for the Commonwealth Government's Carbon Farming Initiative (CFI), mandate that only a fraction of the sample estimate can be considered to be tradable or even real, ignoring the substantially greater probability that the volume (or mass in the CFI case) will exceed this fraction. Some CFI methodologies propose using fractions of the sample-based estimates that have a 97.5% probability of being less that the true volume/mass present. Is this really just a safe and conservative estimate or an underlying fear that mathematics is not reliable?

Given the above, I suppose I should not have been surprised by an almost universal distrust of forest planning approaches based on numbers. Operations research, a field that grew out of necessity during the Second World War, promised 'optimal' management decisions. Proactive forest management agencies promoted the use of linear optimisation packages in the 1970s to estimate optimal harvesting levels and regimes for their plantations. Provided the managers knew relatively simple constraints like how much wood they were required to produce; how much that wood was worth; how much area they had; and inventory-based estimates of volume and yield, these packages were 'mathematically guaranteed' to generate an optimal harvesting level and allied regimes. Any alternative was consequently guaranteed to be sub-optimal. However, no one to the best of my knowledge ever implemented the harvesting levels or regimes as produced. Sometimes the change was attributed to risk, but more usually just to a need to generally 'be conservative'. Such conservatism has not always stood the profession in good stead—conservative estimates of the growth and yield of a forest mean you estimate a requirement for more land area to reach a predefined sustained yield, which opponents then denounce, with mathematical justification, as a land grab!

Just as design-based sampling involves only simple mathematics, linear optimisation algorithms also simply apply a bit of algebra (admittedly it was matrix algebra) and are well within the limits of human comprehension. In fact, around about 1980, undergraduates in the BSc (Forestry) degree at the Australian National University were required to solve simple linear optimisation problems by hand as part of their forest management courses!

But at least you generally know where you stand with this group who distrust numbers—they are conservative in that they believe in less change than the numbers indicate (e.g. smaller increases, smaller quotas, less anticipated damage).

The other substantial group of people tend to accept whatever the numbers say as objective truth. I fear that many in this group accept numerical answers not so much because they understand the underpinning constructions and assumptions, but more like they are revealed truths from an unarguable authority. If they use a sufficiently powerful computer and statistical analysis package, then whatever comes out, by definition, is true. Consider the recent interest among forest conservationists in allometry, the relationship between the size of a part of an organism and the whole. I fear there are more who *believe* in allometry than there are who *understand* allometry. I appreciate how so many observations can be summarised into just two parameters in an allometric equation but I don't *believe* in allometry. The practical forester of the 1800s would have summarised the allometric relationship between say X = diameter at breast height (dbh) and Y = volume as a table and called it a Tarif. Further, given there was just one variable (dbh) that practical forester would have called this a Local Tarif and, once determined for specific local areas and species, would have confidently used it to estimate volume from a measurement of dbh. But that same forester would not have taken that Local Tarif to a different forest and used it, even though the table could be read to fractions of the unit of volume. No, despite the apparent precision of the output of a Local Tarif, the estimate would be wrong, even on average, when applied to a different area. If the user wanted to estimate volume for trees over a wider region, they would develop a Regional or Universal Tarif—one that included more than one single X or independent variable to account for changing patterns in tree shape or form associated with different environments.

The South Australian Bulletin 20 (Lewis *et al*. 1973) is probably the best-known equivalent of this sort of 'universal' predictor with independent variables incorporating dbh, height, bark thickness, and taper. Such tarifs or equations could, on average, have minimal error when applied over regions and even within sub-parts of a region. An assignment I used to make my students complete was to compare the tree volume estimates from the South Australian regional tables with their own local function for the plantations of the Australian Capital Territory (ACT). Inevitably, the average error of the South Australian predictions were no worse than that of the local function—but no one could apply their ACT local function successfully in other plantations without introducing error. It does not matter how sophisticated the model fitting technique nor whether linear, linearised or nonlinear models are used, or even if the samples taken to parameterise the model are selected systematically or randomly, restriction to a single independent variable means that the calibration is localised. A model with a single independent variable cannot adequately model changes in the shape or taper of trees, differential allocation of within-tree resources or even bark thickness resulting from environmental gradients or changing stand structure cohorts.

And yet, many scientists and policymakers or commentators accept without question the calculations from an allometric model calibrated in a different region or country, or for stands with different histories. Even if the fit statistics appear exceedingly good, a 'local allometric' will estimate the biomass or the volume of a tree of 100 cm diameter as the same regardless of whether it is a young, fast growing dominant or an old, slow growing senescent tree. Differences in shape, in internal decay and limb loss, and in bark thickness mean that the net volume and biomass must be different. And the difference will not be random—it will be systematic. Further, even if you calibrate a local allometric now and use it to estimate the biomass or volume of an 80-cm tree and then assume it will grow to 100 cm over the following few decades, it will not end up with the same biomass or volume as the tree that is currently 100 cm. The tree that is 100 cm now must have a different history to the one that is only 80 cm now (e.g. more dominant, a better-located site) and so once the 80 cm reaches 100 cm it will still be different and the original local allometric will introduce a systematic error if used to estimate the growth in volume/mass of the 80-cm tree growing to 100 cm.

Unfortunately this systematic error was ignored in a recent paper in *Nature*, which used changes in dbh to demonstrate that large trees continue to grow substantially in biomass (Stephenson *et al*. 2014). Stephenson *et al*. (2014) based their conclusions on two underpinning sets of 'numbers'—repeated observations of dbh and estimates of mass (and consequently changes in mass) derived from allometry using those dbh values. Estimates of dbh are based on a model or assumption of a circular or possibly an elliptical cross-section of a tree at breast height. While untrue, this model allows callipers (at one or two points) or girth tapes to measure something that is related to cross-sectional area or bole diameter. For young, healthy, and undamaged trees, the estimated value is related to cross-sectional area through the simple formula $A = \pi D^2/4$ (where D = dbh). As trees age, however, and particularly as they approach the large trees described in the article, the relationship between the measurements and cross-sectional area systematically deteriorates. Eucalypts in particular almost certainly will develop rot or insect damage and butt hollows as they age and so the correlation between measurements of dbh and mass, even on the same tree, becomes increasingly biased. The errors in extrapolating dbh measurement, in addition to the above concerns about the use of local allometrics, suggest that the conclusions of the paper by Stephenson *et al*. (2014) are based on invalid assumptions and faith in an immutable and simple allometric relationship between dbh and mass.

Major policy decisions are made on the basis of such powerful and reliable-looking numbers. People making or arguing their policies on the basis of these numbers passionately believe in the power of mathematics to justify the truth of their position. But such numbers may easily be wrong. Such numbers are likely to be biased.

So, if fear and faith in mathematics are both problematical, how do we proceed? We can start with the conclusion of that famous statistician, G.E.P. Box (1976, p. 791), that 'all models are wrong…' and realise that numbers that result from sophisticated allometrics—linear and nonlinear regressions—are the result of models that are wrong. Even 'measuring' is really a form of modelling, and therefore measurements are wrong too. But, as Box went on to say '…some are useful', or as I say, 'some models are wrong in useful ways'. We know that sample-based inventories produce quantitative estimates, but always and only accepting the lower confidence bound is like concluding that 2 + 2 = 3. While that statement is obviously false, 2 + 2 > 3 is obviously not false and may be a useful and 100% true conclusion. Similarly, an allometric equation fitted on measurements of healthy, relatively young trees of a given species and site will be wrong when used on those same trees in the future. But, a useful and 100% true conclusion may be that the mass of those trees in the future will be less than the estimates extrapolated by the local allometric. Measurements and models are vitally important tools in the forest manager's arsenal, but only when we realise where they are 'wrong' will we be able to use them to make true and useful statements and decisions. Foresters need to reclaim their role as being the world experts in forest mensuration from those who don't believe numbers or just use rote formulas and faith.

Cris Brack
Australian National University
Email: cris.brack@anu.edu.au

References

Box, G.E.P. (1976) Science and statistics. *Journal of the American Statistical Association* **71**(356), 791–799. doi:10.1080/01621459.1976.10480949

D'Avignon, G.R. and Winkels, H.M. (1986) Multicriteria decision aid for the management of uncertainty. *In:* Wilkin, L. (ed.) *The Management of Uncertainty: Approaches, Methods and Applications*. Martinus Nijhoff Publishers, Dordrecht, pp. 75–135.

Cameron, W.B. (1963) *Informal Sociology: A Casual Introduction to Sociological Thinking*. Random House, New York, 70 pp. [Sometimes attributed to Albert Einstein].

Lewis, N.B., McIntyre, G.A. and Leech, J.W. (1973) *Regional Volume Table for* Pinus radiata *in South Australia*, metric edn. Bulletin No. 20, Woods and Forests Department, Adelaide, South Australia, 64 pp.

Stephenson, N.L., Das, A.J., Condit, R., Russo, S.E., Baker, P.J., Beckman, N.G., Coomes, D.A., Lines, E.R., Morris, W.K., Rüger, N., Alvarez, E., Blundo, C., Bunyaveichewin, S., Chuvong, G., Davies, S.J., Duque, A., Ewango, C.N., Flores, O., Feanklin, J.F., Grau, H., Hao, Z., Harmon, M.E., Hubbell, S.P., Kenfack, D. and Lin, Y. (2014) Rate of tree carbon accumulation increases continuously with tree size. *Nature* **507**, 90–93. doi:10.1038/nature12914