KUMARAVELU, G., S. JAGDEES, R. S. V. RAI and B. SAMPSON (1995): Provenances of *Eucalyptus camaldulensis* Dehnh and *E. tereticornis* Sm suitable to South Indian conditions – results of an IUFRO trial. Annals of Forestry **3**, 129–133.

LAL, P., H. D. KULKARNI and K. SRINIVAS (1993): Eucalyptus improvement programme of ITC Bhadrachalam Paper boards Ltd. *In:* VIVEKANANDAN, K., SUBRAMANIAN, K. N., ZABALA, N. Q. and GURUMURTHI, K. (ed.) Proc. Workshop on Production of Improved Planting Material for Afforestation Programmes, 18–25 June, Coimbatore, India.

MIDGLEY, S. J., J. W. TURNBULL and K. PINYOPUSARERK (2002): Industrial Acacias in Asia: Small brother or big competitor? *In:* WEI, R. P. and XU, D. (Eds.) Proc. International Symposium on Eucalyptus Plantations, September 1–6, 2002, Guangdong, China, pp 19–36.

OSORIO, L. F., T. L. WHITE and D. A. HUNBER (2003): Age-age and trait-trait correlations for *Eucalyptus grandis* Hill ex Maiden and their implications for optimal selection age and design of clonal trials. Theoretical and Applied Genetics **106**, 735–743.

PINYOPUSARERK, K., J. C. DORAN, E. R. WILLIAMS and P. WUSAWANICH (1996): Variation in growth of *Eucalyptus camaldulensis* provenances in Thailand. Forest Ecology and Management **87**, 63–73.

POTTS, B. M. and H. S. DUNGEY (2004): Interspecific hybridization of *Eucalyptus*: key issues for breeders and geneticists. New Forests **27**, 115–138.

RAO, D. V. (1984): Provenance trials of *Eucalyptus.* Indian Forester **110**, 28–34.

VARGHESE, M., A. NICODEMUS, B. NAGARAJAN and K. SUBRAMANIAN (2000): Hybrid breakdown in Mysore gum and need for genetic improvement of *Eucalyptus camaldulensis* and *E. tereticornis*. Proceedings,QFRI/CRC-SPF Symposium on Hybrid breeding and Genetics of Forest Trees. April 9–14, Noosa, Australia. Queensland Forestry Research Institute, Gympie, pp 519–525.

VARGHESE, M., N. RAVI and R. HEGDE (2001): Seedling seed orchards of *Eucalyptus tereticornis* for improved productivity in plantations of Tamil Nadu, India. International Conference on Forestry and Forest Products Research (CFFPR), 2001, Kuala Lumpur, Malaysia, 1–3 October, 2001, pp 520–522.

WEI, X. and N. M. BORALLHO (1993): Genetic control of growth traits of *Eucalyptus urophylla* S.T. Blake in South East China. Silvae Genetica **47**, 158–165.

WILLIAMS, E. R., A. C. MATHESON and C. E. HARWOOD (2002: Experimental Design and Analysis for Tree Improvement, Second Edition. CSIRO, Melbourne.

WILLIAMS, E. R. and M. TALBOT (1993): ALPHA+ : Experimental design for variety trials. Design User Manual. CSIRO, Canberra and SASS, Edinburgh.

Short Note: Coefficients of Variation in Variables with Bounded Scales

By R. D. BURDON*)

Senior Scientist, Genetics, Scion (New Zealand Forest Research Institute Ltd),
Private Bag 3020, Rotorua, New Zealand

(Received 23th October 2007)

Abstract

With a variable that is recorded on a scale with fixed bounds, it can be appropriate to use for the denominator of the coefficient of variation the square root of the (sign-independent) product of the differences between the mean and the two bounds of the scale. A simple illustrative example is given.

Key words: coefficient of variation, subjective rating, subjective score, binary variables.

The coefficient of variation (CV), namely the standard deviation divided by the mean, is often very useful information for the breeder. In particular, the product of the phenotypic CV and the heritability is a direct measure of the potential for genetic improvement of a given trait. For an ordinary metric trait (e.g. a growth variable such as height) a CV is fully valid, although an extremely high CV may be an indicator of severe positive skewness. However, in tree improvement it is often necessary to address traits for which valid measures of CVs are not straightforward. Such traits include ones that in practice need to be assessed according to bounded scales (e.g. subjective straightness scores), or else have binary expression (e.g. survival). In such cases, the choice of a relevant mean for the denominator of a CV is problematic.

As an example, consider straightness, assessed on a 1–9 scale (1 = very crooked to 9 = very straight). With a mean score of 3, 3 would superficially be the denominator for the CV, although it seems more realistic to use 2 as the denominator, this being the difference 3–1. But if the scale is to be used as a measure of crookedness, one might use 6 (namely 9–3) as the denominator.

As a mean score for such a variable approaches a bound, the standard deviation will approach zero. Yet the alternative CVs will diverge widely according to

*) E-Mail: burdon@scionresearch.com

DOI:10.1515/sg-2008-0027
edited by Thünen Institute of Forest Genetics

Table 1. – Alternative expressions of phenotypic coefficients of variation (%) for branch habit scores (1–9) in *Pinus radiata* at two sites (from C.J.A. Shelbourne & C.B. Low unpubl. data).

Site	Mean score	Phenotypic variance	Denominator used			
			X_{mean}	X_{mean} - 1	$9 - X_{mean}$	$[(X_{mean} - 1) \times (9 - X_{mean})]^{1/2}$
A	7.56	1.591	16.7	19.2	87.6	41.0
B	5.39	2.351	28.4	34.9	42.5	38.5

whether one chooses as the denominator the difference between the mean and one or the other bound. It seems very desirable to have an expression for the CV that is independent of what may be an arbitrary choice of denominator. To this end, it is proposed that the CV denominator (D) be expressed as

$$[(X_{mean} - X_{min.})(X_{max.} - X_{mean})]^{1/2} \quad (1)$$

where X_{mean} is the mean and $X_{min.}$ and $X_{max.}$ are the lower and upper bounds of the scale respectively (cf BURDON, 1998). This expression is sign-independent. In the example considered, D would be $\sqrt{[(3-1)(9-3)]}$, which equals 3.464.

Table 1 gives a simple actual-data example involving branch-habit scores for a progeny trial of *Pinus radiata* replicated on two sites, one of which gave scores clustered near the upper bound. Using Eq. 1 eliminated an apparent wide difference between sites in CV.

The approach can be applied not only to phenotypic CVs but also to CVs for component sources of variation, e.g. genotypic and environmental effects. Estimates of such CVs are obtainable by using as numerators the square roots of the respective variance component estimates, with the same denominator.

The phenotypic variances (and hence standard deviations) for integer-value scores might be adjusted by using Sheppard's correction, which entails subtracting one-twelfth from the variance (SNEDECOR and COCHRAN, 1966). In principle, it removes variation that is an artifact of recording in discrete categories. However, experience (BURDON, unpubl.) is that this correction is only satisfactory when the scores are symmetrically distributed, which is likely to require that a mean be close to mid-range.

Since the scales for subjective ratings are inevitably somewhat arbitrary, caution will be needed in interpreting CV statistics even when using Equation 1. For instance, the implied linearity of the scale is never guaranteed, and if it exists, there would be no guarantee that the economic-worth function would be linear. Normalising transformations of data, while they may improve data properties for purposes of statistical analysis, would make expression of CVs very problematic. Moreover, any effort made to spread the scores away from mid-scale, in an attempt to achieve greater resolution of group differences, will make CVs less meaningful biologically. A similar problem arises with data such as disease ratings which, for biologically plausible reasons, may be expressed on logarithmic-type scales (e.g. GILMOUR and NOODERHAVEN, 1973) that give greater weight to differences near the zero and the 100% bounds.

Note that for a binary trait, showing 0 or 1 expression (Bernoulli distribution), with a proportion p of individuals showing one expression or the other, the phenotypic variance equals $p(1-p)$. On the basis of Equation 1, the CV is automatically 1. The components of variation will all be subject to an upper bound of 1 for their corresponding CVs.

Acknowledgments

Thanks are due to A. D. YANCHUK, M. O. KIMBERLEY and T. J. MULLIN for helpful comments on drafts, and to an anonymous referee for minor corrections.

References

BURDON, R. D. (1998): Relative performance values in genetic tests: Alternatives and their properties. Silvae Genet. **47**: 1–5.

GILMOUR, J. W. and A. NOODERHAVEN (1973): Control of *Dothistroma* needle blight by low volume aerial application of fungicides. New Zeal. J. For. Sci. **3**: 120–136.

SNEDECOR, G. W. and W. G. COCHRAN (1967): Statistical Methods, 6th ed. Iowa State Univ. Press, Ames.

Herausgeber: Johann Heinrich von Thünen-Institut. Bundesforschungsinstitut für Ländliche Räume, Wald und Fischerei.
Schriftleitung: Institut für Forstgenetik, Sieker Landstraße 2, D-22927 Großhansdorf
Verlag: J. D. Sauerländer's Verlag, Finkenhofstraße 21, D-60322 Frankfurt a. M.
Anzeigenverwaltung: J. D. Sauerländer's Verlag, Frankfurt am Main.
Satz und Druck: ADN Offsetdruck, Battenberg — Printed in Germany.

ISSN 0037-5349
DOI:10.1515/sg-2008-0027
edited by Thünen Institute of Forest Genetics