Assessing the quality of permanent sample plot databases for growth modelling in forest plantations

J.K. Vanclay, J.P. Skovsgaard & C. Pilegaard Hansen

Royal Veterinary and Agricultural University,
Danish Forest and Landscape Research Institute, and
National Forest and Nature Agency, Denmark

Abstract

Informed plantation management requires a good database, since the quality of information depends on the quality of data, growth models and other planning tools. There are several important questions concerning permanent plots: how many plots, where to put them, and how to manage them. Plot measurement procedures are also important. This paper illustrates graphical procedures to evaluate existing databases, to identify areas of weakness, and to plan remedial sampling. Two graphs, one of site index versus age, another with stocking versus tree size, may provide a good summary of the site and stand conditions represented in the database. However, it is important that these variables, especially site index, can be determined reliably. Where there is doubt about the efficacy of site index estimates, it is prudent to stratify the database according to geography, soil/geology or yield level (total basal area or volume production). Established permanent plot systems may sample a limited range of stand conditions, and clinal designs are an efficient way to supplement such data to provide a better basis for silvicultural inference. Procedures are illustrated with three data sets: teak plantations in Burma, Norway spruce in Denmark, and a clinal spacing experiment in India.

Introduction

Plantations represent a considerable investment of time and resources, and efficient management is necessary to recoup this investment. The potential value of effective plantations offers many options and incentives for silvicultural intervention. Computer simulation models and information systems can help to explore these options and their implications, but depend upon suitable data. The intensive management and financial risks characteristic of plantations demand reliable data and detailed information systems.

Fortunately, it is comparatively easy to gather data in plantations. Access is usually good, and existing records may provide a suitable basis for stratification. Plantations are normally rather uniform, so relatively small plots can be used. Since the unit cost of data collection in plantations is comparatively low and the potential benefits are considerable, it should be possible to justify a comprehensive database.

Plantation managers and planners need to know the optimal silvicultural regimes including initial spacing, intensity and timing of thinnings, rotation length and regeneration method. They also need to know the sensitivity of these operations to site quality and socio-economic assumptions. They must also forecast harvests from existing plantations, and plan the steady and sustained flow of timber (and other products and services) into the market place. Although other details are also required (e.g., prospects for weed control and fertilizing), these issues illustrate the main uses of growth models, and in turn, the data required to construct them. Thus permanent sample plots (PSPs) should sample a wide range of site and stand conditions to provide for valid inferences on these important issues.

PSPs may be classified as *experimental plots* or *passive monitoring plots*. Too often PSP systems comprise mainly passive monitoring plots in "typical" stands, and this severely restricts the ability of resulting models to provide inferences on optimal silviculture. To find out what happens when you disturb a system, you have to disturb it, not just passively observe it (Box 1966). Hence a PSP system for plantations should include experimental plots which are manipulated to provide data on a wide range of stand density, a range of thinning strategies (heavy/light, early/late, above/below), and data from stands which have been allowed to develop beyond the normal rotation age. Some plots should be established in all geographical regions in which plantations have been established or are proposed. All climatic zones and soil types should be represented in the PSP database.

These are demanding specifications, and it is important to keep them in perspective. The quality of data is paramount, so if resources are limiting, it is better to have a few well managed plots (i.e., with regular, detailed and accurate measurements), than many inadequate ones. It is critical for growth modelling applications that PSPs be optimally placed to sample the full data space, especially when resources limit the number of plots. The present study examines three data sets to illustrate strengths and weaknesses in the database highlighted by simple graphical analyses.

There are many ways to appraise the quality of a database, and the present study will consider some simple graphical approaches. The simplicity of these methods does not detract from their usefulness, and they may provide the basis for more quantitative techniques (e.g., Beetson *et al.* 1992). The beauty of the graphical approach is that it is simple, easy to use, and provides a clear picture illustrating the strengths and weaknesses of databases. The recent IUFRO conference on "Integrating Forest Information over Space and Time" called for efficient strategies to supplement deficient databases as quickly and parsimoniously as possible (Turner and Wood 1992). We hope that the methods outlined below will contribute toward this goal.

Table 1. Characteristics of Laurie and Ram (1939) data.

Variable	Minimum	Mean	Maximum
Age (years)	3	29	72
Site Index (m at 80 yrs)	17	21	46
Top Height (m)	7	22	36
Diameter (cm)	5	23	65
Stocking (stems/ha)	69	451	2098
Basal Area (m^2/ha)	4	12	24

Teak Plantation Data from Myanmar (Burma)

The present study was inspired by a review of data suitable for the construction of a growth model for plantation teak (*Tectona grandis*) in Myanmar (Vanclay 1992). Some of the data were collected during 1917-37 and were later published (Laurie and Ram 1939). They comprise 221 permanent plots in teak plantations in Myanmar (Burma) and 10 in the Chittagong Hills (Bangladesh). Most plots were remeasured every four or five years, and some plots were measured as many as four times. Laurie and Ram (1939) used these data to compile yield tables, but they have also been used for growth modelling (Tint and Schneider 1980). A good range of site and age was sampled (Table 1), but closer examination reveals that these data are suboptimal for growth modelling. The intent of this study is not to be critical of past data collection efforts, but to illustrate how a critical appraisal can indicate the strengths and weaknesses of a data set, and these data provide a good example to illustrate how an apparently good data set may on closer inspection, reveal deficiencies.

A graphical summary of the data may be more revealing than the common tabular approach (Table 1). Critical evaluation of a database requires careful choice of the axes used in the graphs, and a thorough evaluation may require more than one graph. Beetson *et al.* (1992) suggested two graphs illustrating the range of data sampled in natural forests: one with the important stand-level variables, basal area and site quality; the other showing the tree-level variables size (e.g., diameter) and competition (e.g., basal area in larger trees). In a mixed-species stand, the second graph would be prepared for each species.

The axes should draw on variables which (1) are fundamental to stand dynamics and growth modelling, (2) can be measured directly, and (3) have minimal measurement error. In plantations, age is an obvious candidate. Site index, stand density (e.g., basal area or stems/ha) and mean tree size (e.g.,

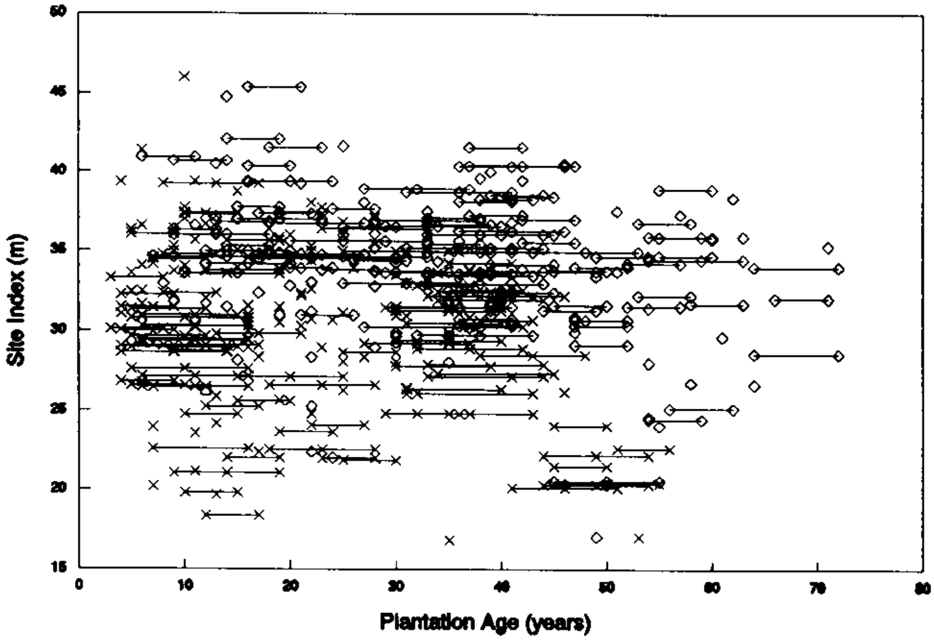

Figure 1. Range of site index and age in Burma data. Symbols indicate plots above (◇) and below (×) the mean stand basal area. Lines join remeasures on the same plot. Site index (ht at 80 years) is the average of all measures.

diameter) are also important candidates in plantations. The graph of plantation age versus site index is more revealing than for example, the height versus diameter graph, since age and site are not correlated and better illustrate the data space (Figure 1). Stand variables which can be manipulated silviculturally (e.g. stand density) provide useful guidance for the management of experimental plots. Thus Vanclay (1992) used two figures, one with site index versus age (Figure 1), and the other showing stand density versus mean tree size (Figure 2). Both figures used stand-level data, since tree-level data may clutter the graphs without providing much additional information. These figures do not illustrate the growth of the trees, but illustrate the range of site and stand conditions sampled, and thus the value of the data for defining a response surface for growth modelling. They provide a powerful way to illustrate data quality, and are interpreted more easily than tables and statistics (cf. Table 1).

Figure 1 illustrates that a good range of sites and ages have been sampled. Lines join remeasures on the same plot, and show that although the data span 70 years, no single plot has been monitored for more than 13 years. The different symbols indicate plots above and below the mean basal area for the database, providing a simple illustration of the range in stand density with site and age. It is not surprising that young stands have more plots below the mean basal area, but

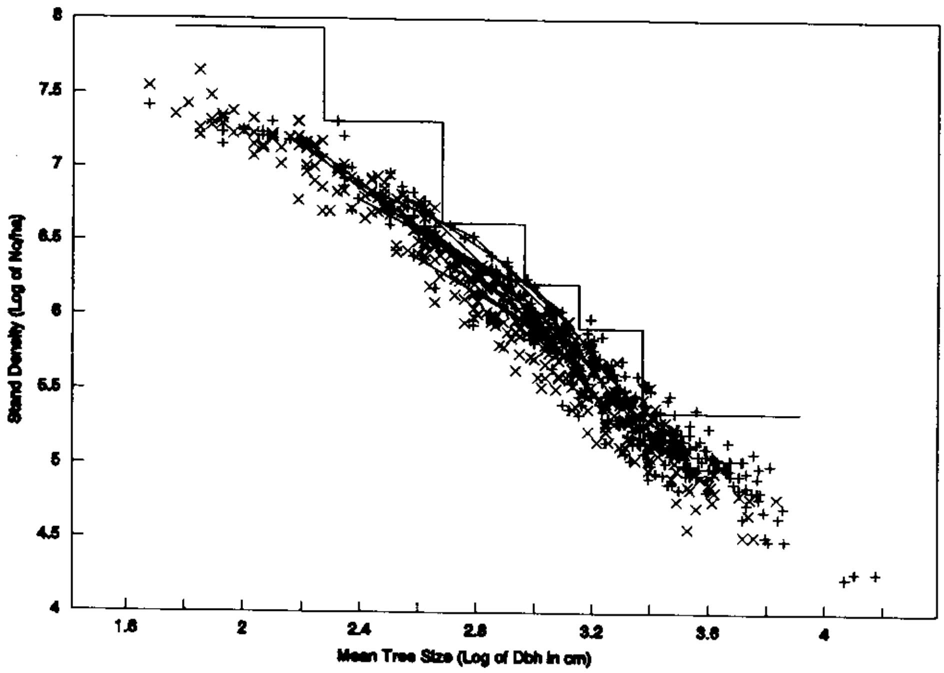

Figure 2. Mean tree size and stand density in Burma data. Lines join remeasures on the 15 plots measured 4 times. The stepped line is a silvicultural prescription (Streets 1962). Symbols indicate above (+) and below (×) average site index.

this does reflect on the quality of the data: the ideal data set for modelling should have some plots sampling extremes of stand density at all ages. It does not matter that these stand conditions do not occur in routine management; they remain necessary to define properly the response surface (Vanclay 1991).

Figure 1 also indicates that many of the plots sample plantations aged about 40 years with site index 35 metres. If cost-cutting measures demanded that the number of plots should be reduced, then it would be preferable to abandon plots in this part of the data space, than to loose plots at the extremes of the site and age.

Figure 2 illustrates the range of stocking (stems/ha) and mean diameter sampled. Another alternative would be to use stocking versus tree height. Figure 2 is more revealing than Figure 1, and indicates a deficiency in these data: all the data are clustered in a narrow belt sampling a limited range of stocking at any stage of development. A logarithmic transformation has been used to compress the data onto a useful scale, but means that a single unit represents almost a three-fold change (using natural logarithms to base e). The logarithmic scale on both axes provides a stand density diagram which may provide a useful basis for management prescriptions (Reineke 1933). It is apparent that these data are drawn from passive monitoring plots and do not include plots in open-grown stands with less between-tree competition. The stepped line in Figure 2 indicates a silvicultural prescription for teak (Streets 1962, Lamprecht 1989) which suggests that the Laurie and Ram (1939) data sample managed stands of relatively low density in which there is little natural mortality. The stocking reductions evident in Figure 2 must be due to thinning, which should show as a stepped line (e.g., the silvicultural prescription) if measurements are made before and after thinning.

These data do not provide a good basis for making inferences about silvicultural alternatives because of the narrow range of stand conditions sampled and the failure to discriminate deaths and harvests. They are well suited for preparing standard yield tables and site index curves, but are less well suited to the construction of dynamic growth and yield models. Such passive monitoring plots should be supplemented by thinning and spacing experiments to provide data at silvicultural extremes, to define response surfaces more fully and allow better predictions from the model. Supplementary data should include both heavy thinnings (or wide initial spacing) and some unthinned plots to allow natural mortality through self-thinning. Without additional experimental plots, the database cannot provide a good basis for modelling silvicultural options.

Norway Spruce in Denmark

Another way to appraise the quality of a PSP system is to contrast the extent of the database with that of the forest estate. Ideally, the PSPs should sample the full range of conditions encountered, as a subset may compromise the ability of planning systems to make useful inferences for forest management. Comparisons can be made by graphing PSP data with other resource data (e.g., temporary inventory plots, management records). We illustrate with an appraisal of Norway spruce (*Picea abies*) data in Denmark. Norway spruce is the most important timber species in Denmark, with established plantations totalling 130,000 ha, equivalent to 32% of the forested area. It is not indigenous, and all stands are planted.

This study contrasts data from 22 PSPs in most parts of the country with 12,467 inventory observations in national forests. The PSPs are thinning experiments with 2-8 different thinning treatments, and represent a total of 180 sub-plots. All but 4 PSPs contain unthinned control plots. The inventory observations sample some 30,000 ha, and are extracted from the forest management database (FMDB) at the National Forest and Nature Agency. The FMDB is updated regularly, and contains data recorded over a 15 year period and projected 0-15 years to a common status at 1 January 1993. The projection is based on computerized versions of standard growth and yield tables for Norway spruce in Denmark (Møller 1933, West-Nielsen 1950, Magnussen 1983). As practically all stands in national forests are inventoried, this database is presumably a faithful representation of the national forest estate.

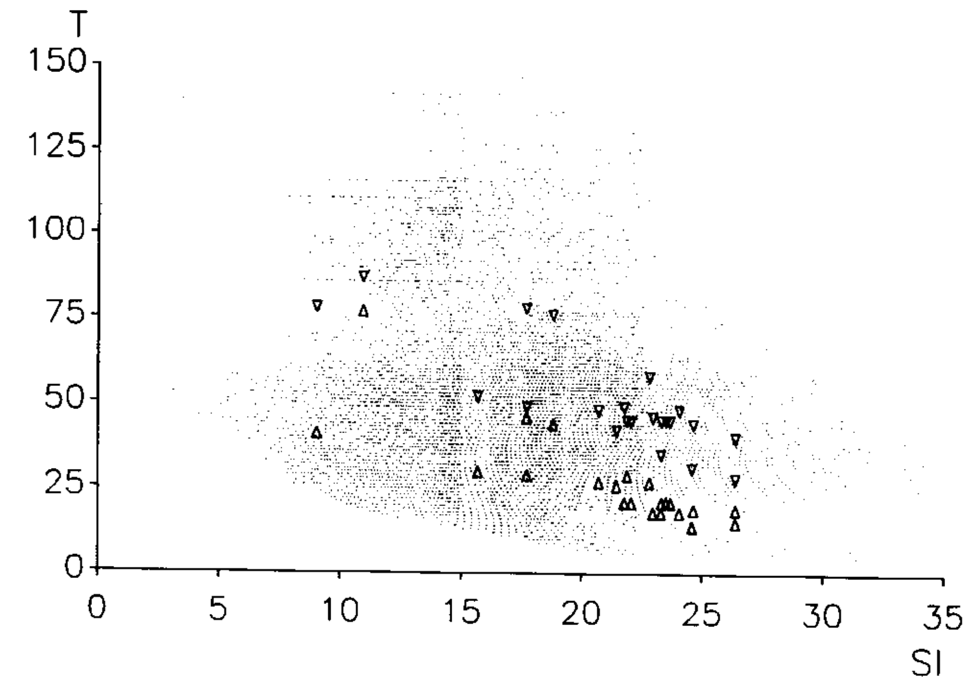

Figure 3. Stand age (T, yrs) vs site index (ht, age 50). Stands are classified on site index using a computerized version of Møller (1933) to allow comparisons. Symbols indicate the FMDB (·) and the age range represented on PSPs (▵-▿).

Figure 3 contrasts the range of age and site index sampled by the PSPs with the FMDB. Although the PSPs cover a good range of sites and ages, they sample less than half the data space represented in the FMDB. Site indices 18-27 appear well sampled, but remedial sampling would be desirable at the extremes and for site indices 10-15. None of the PSPs have reached great ages, so supplementary sampling in old stands would also be beneficial. Stem analyses of selected trees could supplement the data available for modelling the height-age relationship.

Figure 3 illustrates some weaknesses in the PSP database, but may overstate its strengths as it takes no account of regional and soil differences. First impressions suggest that site indices 20-25 may be over-represented, but site index (height at age 50) may not be an adequate indicator of site productivity, and stratification by other factors (e.g., geography, topography, soils, etc.) may reveal that the data are not as comprehensive as they first appeared. Further, the FMDB illustrates only the national forests, and may not give a reliable indication for the country as a whole. A true national appraisal may give a different emphasis to remedial sampling.

Figure 4 draws on the same data to illustrate the stand height-diameter relationship. Our introductory remarks suggested that other less-correlated

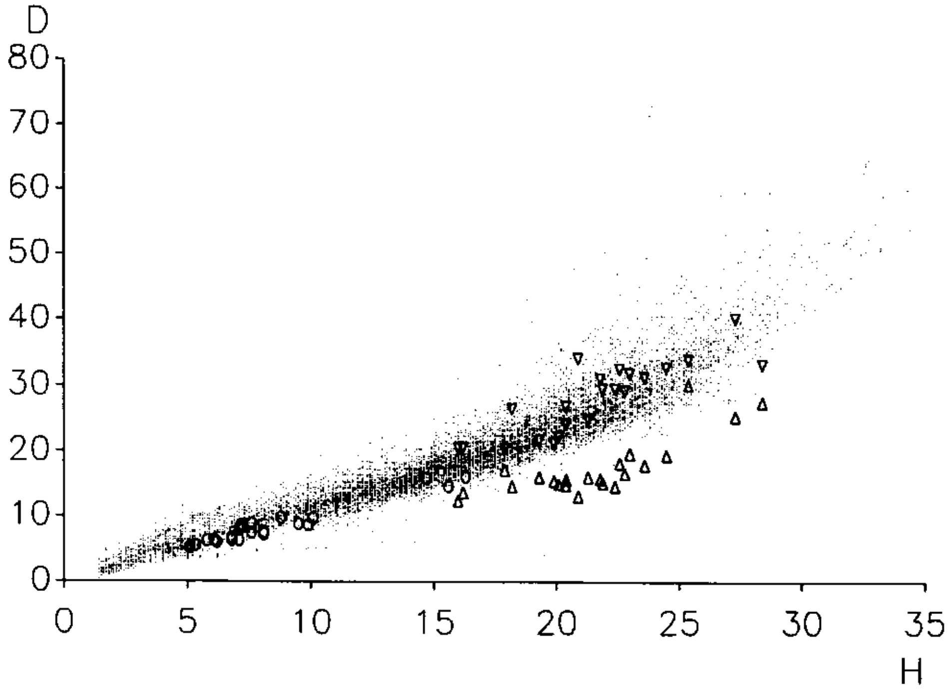

Figure 4. Quadratic mean diameter (D, cm) by stand height (H, m). Symbols indicate the FMDB (·), the initial PSP (pre-treatment) (○), and the range of diameters at last PSP observation (▵-▿).

variables were better suitable for comparing databases, but this figure may be an exception in illustrating a particular weakness. The purpose of this figure is to compare the range of thinning treatments represented by the PSPs with current practice. It would be preferable to use stand basal area as an indicator of thinning grade, as diameter is not the best indicator, but basal area data are presently available only for a few stands.

The PSPs cover a range of thinning treatments from unthinned control plots to heavily thinned shelter belts. However, Figure 4 illustrates that even heavier thinnings are practised in the national forests. These heavier thinnings may originate in overmature stands, or stands being regenerated by the shelterwood system. Figure 4 illustrates that thinning experiments should include extremely heavy thinnings, regeneration cuttings, and thinnings late in the life of the stand. Differences in stand structure prior to experiment establishment (5-10 m height) indicates that PSPs sample stands established at closer spacings than current practice. Thus the PSP database could be supplemented with experiments including wide spacings with early, heavy thinning. When planning such experiments, it is important to realize that the concept of "extreme" changes over time.

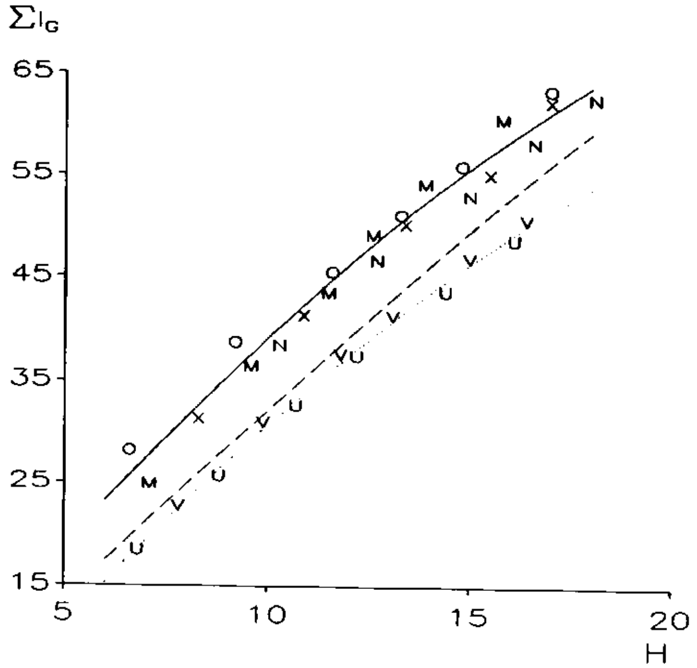

Figure 5. Total basal area production (ΣI_G, m^2/ha) vs stand height (H, m) for unthinned plots in the Rold (KM, KN, KO) and Grib (KU, KV, KX) areas. Regression lines are for Rold (solid), Grib (dashed), and KU & KV plots (dotted).

The sharp edge evident at the lower limit of FMDB data in Figure 4 suggests strict adherence to the standard prescription of early and heavy thinning. The unthinned control plots (▲) may seem redundant, but these control plots provide much important information, and may be critical for interpreting specific experiments, for understanding stand dynamics, and for constructing growth and yield models.

Figure 5 shows site productivity expressed as total basal area production in two selected regions, the Rold forest area in Northern Jutland and the greater Grib forest area (including adjacent forests) in Northern Zeeland. Both forest areas are on sandy, hilly moraine, with apparently identical site indices (based on height-age curves: Rold plots KM, KN and KO have site index 23.7, 21.9, and 21.8 m respectively, and Grib plots KU, KV, KX are 23.5, 22.1 and 23.3 m). The figure is drawn from six unthinned control plots with similar spacing (3 in each region). Regional differences indicate weaknesses in the height site index method (i.e., solid vs dashed lines in Figure 5). Clearly, height site index is not an adequate indicator of site productivity for Norway spruce in Denmark, as these two regions with apparently similar growth conditions have distinctly different yield levels (the yield level is characterized by the total basal area or total volume production at a given height; Assmann 1970:161; Kramer 1988:86). The yield level also may vary considerably within small areas which are apparently homogeneous, e.g., the Grib forest area where one sample plot (KX) with a better water supply (peat soil) shows a significantly higher yield level than the two other plots (KU & KV; dotted line in Figure 5). Large variations have also been reported for Sitka spruce, both at the national level (Skovsgaard in prep.) and for single sample plots (Skovsgaard 1993).

Care is needed in selecting criteria with which to appraise database quality, especially if the appraisal seeks to cull rather than to supplement the database. In this instance, site index based on height-age curves may be a deficient criterion for assessing the PSP database, at it falls short of our previously stated objectives for variables measured directly and without error.

Graphical appraisals may reveal much about a database, but are not foolproof. Care is needed in selecting suitable scales (e.g., scaling the axes differently can change the emphasis), and transformations should be used with care (e.g., logarithmic transformations tend to exaggerate small values and understate large values).

Supplementary Sampling

Supplementary data to expand the range of stand conditions in both databases reviewed above should include thinning and spacing trials. Plot based trials such as correlated curve trend experiments (Bredenkamp 1984) can be expensive of space and resources, but suitable data can also be obtained efficiently from single tree methods such as clinal plots (Dawkins 1960), Nelder fans (Nelder 1962) and scotch plaid experiments (Zavitkovski *et al.* 1982).

We illustrate with data from clinal plots of *Eucalyptus tereticornis* in India. A systematic clinal design based on individual trees (Figure 6) enabled all permutations of 16 spacings ranging from ½ × ½ m through ½ × 5 to 5 × 5 m, to be evaluated in a compact trial comprising only 2000 trees (Adlard *et al.* 1992). These spacings represent a range in stocking from 375 to 25,000 stems/ha, compared with common plantation densities of 2,000-2,500 stems/ha.

Figure 7 illustrates the data space sampled by these clinal plots. Plotting these data as single tree data would exaggerate the range of data relative to other Figures in this paper, so data from the eight replications has been averaged prior to reporting (For rectangular spacings, points represent 16 replications, as data from e.g. ½ × 5 and 5 × ½ have been pooled). The heteroscedacity evident in Figure 7 (variance increases as stocking increases) is due to the many configurations possible at higher densities (e.g., ½ × 5 = 1.6 × 1.6 = 4000/ha) and the inability of this crown-shy species to effectively utilize the extreme configurations (e.g., ½ × 5). Clinal plots may be more fragile than traditional block designs, and care is needed during establishment and maintenance. The data in Figure 7 have not been adjusted for missing trees; some data represent as few as

6 replications, but the nominal stand density has not been adjusted.

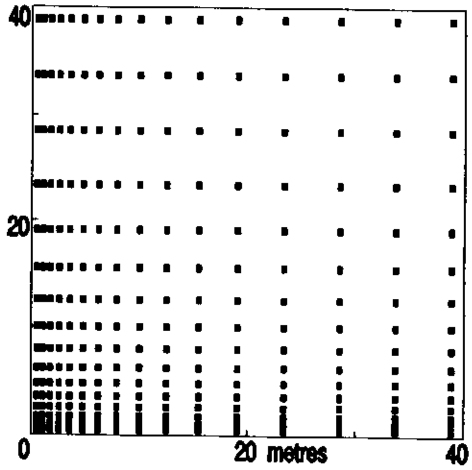

Figure 6. Planting positions used in each block of the Adlard *et al.* (1992) clinal spacing experiment.

In three years (1989-1992), the clinal plots sampled a much larger part of the data space than the teak data illustrated in Figure 2, and in time, will provide a very good database for growth modelling. These plots should be maintained, beyond the normal rotation age (8 years for pulpwood) to provide data on self thinning and the onset of competition. It may be useful to thin some plots to examine the response to thinning, but some plots should remain unthinned. Clearly, clinal plots may be very efficient in supplementing existing data to extend the range of size-density sampled and so enable more reliable inferences on stand dynamics and silvicultural options. This example draws on young plots in short-rotation plantations, but clinal plots also provide an efficient way to gather data on the long rotations common in temperate forests. Estimating the point density is complex in mature clinal plots after significant mortality, but such data may provide valuable insights into stand dynamics and greatly assist the calibration of growth models.

The logarithmic scale used in Figures 2 and 7 tends to devalue the large trees and dense stands, and care should be taken to sample these extremes as they may be critical for modelling diameter increment and mortality. Despite this weakness, the size-density diagram serves to emphasize the need to obtain a good representation of a wide range of stand conditions.

Discussion

This study has focused on forest trees in a plantation environment, but analogues of these graphs may be used to evaluate data from other situations. For example, agroforestry data may need to draw on individual tree characteristics. The important thing is to identify critical variables for tree growth, crop dynamics and growth modelling, and to appraise data in such a way to best illustrate the range of values sampled. It is not just the overall range that is important (e.g., Table 1), but also the range of each variable for each level of other variables. This helps to avoid multicollinearity which may lead to problems in regression analyses. Thus in Figure 1, we seek a good range of age for each site index.

Graphical analyses will not make it easier to collect data, but can help you to invest your measurement effort more efficiently. However, there is no

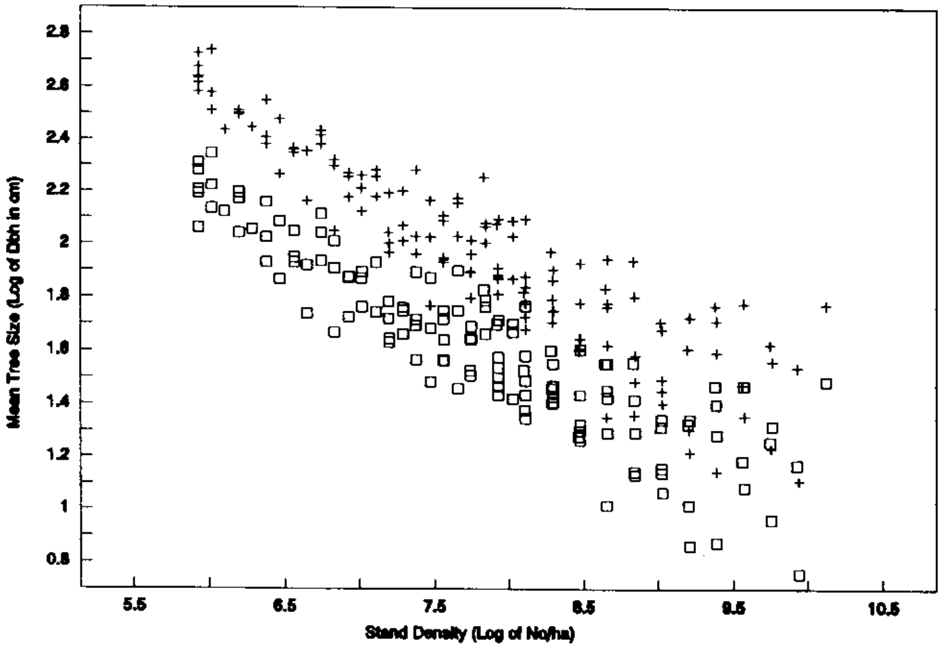

Figure 7. Size-density diagram for clinal plots of Eucalyptus in India at 3 (□) and 6 (+) years of age. Each point is the mean of up to 16 replications. Notice that a good range of density and size is sampled, even after only 3 years.

substitute for an intimate knowledge of the field conditions. In our figures, all plots appear equal, but those familiar with these data have reminded us that some plots are inferior and warrant replication. Do not be blinded by the comfortable images from your computer, but become familiar with your plots in the field.

Conclusion

Graphical analyses provide a simple but effective way to appraise database quality. Two graphs may help to appraise plantation database characteristics, one with site index versus age, and another with stocking versus tree size. These graphs provide a good summary of the site and stand conditions represented in the database. However, it is important that these variables, especially site index, can be determined reliably. Where there is doubt about the efficacy of site index estimates, it is prudent to stratify the database by other factors such as geography, soil/geology or yield level, i.e., the productive capacity of the site as measured by the total basal area or volume production. Reliable inferences require that good range of site and stand conditions are sampled. One of the more efficient ways to sample extreme stand conditions is to use clinal plots. These provide a good basis for calibrating response surfaces for growth models.

Acknowledgements

Philip Adlard kindly provided the data for Figures 6 and 7.

References

Adlard, P.G., G.S. Kariyappa and N.V. Srinivasalu, 1992. Spacing at planting of short-rotation *Eucalyptus* in Karnataka. In: I.R. Calder, R.L. Hall and P.G. Adlard (eds), Growth and Water Use of Forest Plantations. Wiley, N.Y., pp. 103-127.

Assmann, E., 1970. The Principles of Forest Yield Study: Studies in the organic production, structure, increment and yield of forest stands. Pergamon Press, Oxford, 506 pp.

Beetson, T., Nester, M. and Vanclay, J.K., 1992. Enhancing a permanent sample plot system in natural forests. The Statistician 41:525-538.

Box, G.E.P., 1966. Use and abuse of regression. Technometrics 8:625-629.

Bredenkamp, B., 1984. The C.C.T. concept in spacing research - a review. In: D.C. Grey, A.P.G. Schönau and C.J. Schutz (eds), Proceedings of the IUFRO Symposium on Site and Productivity of Fast-growing Plantations, 30 April-11 May 1984, Pretoria and Pietermaritzburg, South Africa. South African Forest Research Institute, Pretoria, 1:313-332.

Dawkins, H.C., 1960. The Pudden clinal plot; thinning experiments without surrounds. Empire Forest Review 39:168-171.

Kramer, H., 1988. Waldwachstumslehre. Ökologische und anthropogene Einflüsse auf das Wachstum des Waldes, seine Massenund Wertleistung und die Bestandessicherheit. Verlag Paul Parey, Hamburg & Berlin, 374 pp.

Lamprecht, H., 1989. Silviculture in the Tropics. GTZ, Eschborn.

Laurie, M.V. and Bakhshi Sant Ram, 1939. Yield and stand tables for teak in India and Burma. Indian Forest Record - Silviculture, New Series, Vol. IV-A, No. 1.

Magnussen, S., 1983. En tilvækstoversigt for rødgran på østersønære lerede moræneiorder. Dansk Skovforenings Tidsskrift 68:215-246.

Møller, C.M., 1933. Boniteringstabeller og bonitetsvise Tilvækstoversigter for Bøg, Eg og Rødgran i Danmark. Dansk Skovforenings Tidsskrift 18:457-513, 537-623.

Nelder, J.A., 1962. New kinds of systematic designs for spacing experiments. Biometrics 18:283-307.

Reineke, L.H., 1933. Perfecting a stand density index for even-aged stands. Journal of Agricultural Research 46:627-638.

Skovsgaard, J.P., 1993. Correcting for local variation in site productivity when assessing the effect of thinning on the volume production of Sitka spruce (Picea sitchensis (Bong.) Carr.). Forest Ecology and Management, in press.

Skovsgaard, J.P., in prep. Structure and development of unthinned stands of Sitka spruce (Picea sitchensis (Bong.) Carr.) in Denmark. To be published in Forskningsserien.

Streets, R.J., 1962. Exotic forest trees in the British Commonwealth. Oxford.

Tint, K. and Schneider, T.W., 1980. Dynamic growth and yield models for Burma teak. Mitteilungen der Bundesforschungsanstalt für Forst- und Holzwirtschaft in Hamburg-Reinbek, Nr. 129.

Turner, B., and Wood, G., 1992. Preface. In: B. Turner, and G. Wood (eds), Integrating Forest Information over Space and Time, IUFRO Conference, 13-17 January 1992, Canberra, Australia. ANUTECH, Canberra, pp.15-16.

Vanclay, J.K., 1991. Data requirements for developing growth models for tropical moist forests. Commonwealth Forestry Review 70:248-271.

Vanclay, J.K., 1992. Follow-up mission report on growth and yield modelling. FO: MYA/85/003 Field Document No. 18. FAO, Yangon, Myanmar.

West-Nielsen, G., 1950. Rødgranens produktionsforhold på den midtjyske hede. Hedeselskabets Tidsskrift 71:118-135.

Zavitkovski, J., Lundgren, A.L. and Strong, T.F., 1982. Biomass production of 4- to 9-year-old intensively cultured *Larix eurolepis* grown in "Scotch plaid" plots in Wisconsin. USDA Forest Service Research Paper NC-231.