Statistical software in forestry research

by **Liz Jupe**, of GenStat

Many readers will be familiar with the common use of computers throughout the forest industry. Forest management involves a great deal of number crunching and computers facilitate analysis of crops, modelling exercises and, more recently, visualisation and mapping. Computers control harvesting heads, timber cranes and, with GPS, are used in the field down to the individual tree scale. Then there are business and communication applications. The industry has been transformed by the use of computers, although forestry businesses generally have been rather slow to take advantage, compared with some other sectors.

Readers may be less aware of the value of computers in forest research, but everything that can be quantified has to be measured and the data analysed.

For example, field experiments on tree growth, might include measurement of ground and under-storey vegetation, soil contamination, levels of infection by insects, fungi and pathogens, as well as damage caused by squirrels or deer. All these affect not only the speed of tree growth, but the way in which the tree grows and develops.

Growth trials are, by their very nature, long-term projects; for example, a survey on the effects of formative pruning on broadleaved trees (such as European ash, cherry, European beech and English oak) meant that form and growth were assessed for up to nine years. Many factors were monitored and recorded in that time, producing a wealth of data to be analysed. This particular trial's results suggested traditional pruning methods were likely to produce an improvement in the quality of the timber.

Forestry experiments also require very careful design due to the large number of factors included in the trials, or at least to be aware of, when interpreting and analysing data. Even before any field trial is laid out, researchers have to make sure a series of other activities has taken place. For example, for a tree introduction programme, seed collections that appropriately represent species, provenances or populations are needed; researchers need to be sure that there are enough seedlings raised for the experiment to be worthwhile and a suitable trial site (or sites) needs to be found, and so the list goes on – and this is even before the trial has actually started.

In general terms, experiments usually test hypotheses, but in forestry research they can also be used for other reasons, such as, estimating wood production. Many tree improvement trials will measure timber quality (proportions of sawlogs, pulp or fuel wood), as well as pest or disease resistance, essential oil production, or the use of the foliage for animal fodder. With so many factors to take into account, careful design, planning, recording and data analysis are crucial to success.

GenStat, one of the first, and only UK-produced, statistical software packages, has largely become the tool of choice for forestry researchers worldwide. Customers include forestry research stations and universities around the globe. GenStat was originally developed by statisticians at Rothamsted Experimental Station (now Rothamsted Research), the largest agricultural research centre in the UK and one of the oldest in the world, meaning that many of the tools have been designed specifically for field trials and the complexities of agricultural data.

A research team at Moulton College in Northamptonshire has been working on a field trial to investigate how water availability affects the growth rate of young trees, and the influence on developing new woodland. The experiment was arranged in a randomised block design, which meant that given the GenStat analysis of variance facilities can analyse any generally balanced design (and certain partially balanced incomplete block designs), analysis of their trial was straightforward. Another key issue for the researchers was the ability of GenStat to handle such large data sets.

A study by Forest Research on the regeneration of oaks within oak woodland used GenStat's generalised linear model facilities to investigate relationships between location and vegetation variables and seedling data. This same experiment also used linear regression, a standard and simple to use function, to analyse the height and diameter of seedlings and of vegetation cover.

Forestry experiments can involve measurements repeated over time and that are likely to be correlated because of spatial proximity. This data can be unbalanced, however GenStat can help to overcome these issues with the use of its mixed model facilities (REML).

There are some statisticians in forestry research, but the majority are research scientists with a working knowledge of statistics, so a package that combines advanced statistics and ease of use is vital.

GenStat has made its mark in forestry research and is committed to supporting this industry worldwide. VSNi is a commercial business, it offers the developing world access free of charge to recent, but not the current version of GenStat. A major beneficiary is KEFRI (Kenya Forest Research Institute). For more information on GenStat visit www.vsni.co.uk or contact support@vsni.co.uk

VSN International Ltd
5 The Waterhouse
Waterhouse Street
Hemel Hempstead
Hertfordshire
HP1 1ES
UK

t: +44(0)1442 450230
f: +44(0)8701 215653
e: info@vsni.co.uk

www.vsni.co.uk