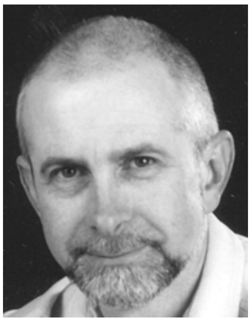
Art, Science and Informatics – Visualisation of Large, Complex Data Sets in High-Speed Measurement of the Microstructure of Wood

Robert Evans
CSIRO Forestry and Forest Products, Australia



Robert Evans was born in Sydney, Australia, in 1949 and lived for 8 years in a shop next to a movie theatre in a very small town on the New South Wales South Coast. It was at Coledale Primary School that Mrs Troy taught her young students about astronomy and the origins of the solar system. At that point Rob decided to be a scientist. After 5 more primary schools and 3 high schools in Australia and New Zealand, he attended the University of Sydney, joined a rock band, married Lynn, and specialized in the thermodynamics and quantum electrodynamics of polymer-stabilised colloidal dispersions under Don Napper.

At the end of his PhD in 1976, there were three choices: a post-doctoral position in France, a post-doctoral position in America, or the life of an itinerant musician. The first required a year at Oxford University working on quasi-elastic neutron scattering, the second was closer to Disney World and the third would have involved not being a scientist. He joined Phil Luner for an exciting 2 years at the Empire State Paper Research Institute in Syracuse, New York, working on colloidal interactions in cellulosic dispersions. He and Lynn dug their rusty old Chevy out of the snow after the great blizzard of ’77. In ’78 they returned to sunny Australia.

Rob joined Peter F. Nelson’s group at Australian Paper Manufacturers Research Division in Melbourne. That group was one of three that had independently discovered anthraquinone catalysis in pulping. After working on everything from pulping kinetics, through paper physics to effluent decolorisation, he took a position at CSIRO on April Fools’ Day, 1985, working with Adrian Wallis on the characterisation of cellulose using viscometry, x-ray diffractometry and size-exclusion chromatography.

In 1988, the Australian pulp and paper industry advised CSIRO to improve wood quality in Australian plantations. Rob had an idea for a machine to perform rapid analysis of wood microstructure. A few years later, the first prototype (SilviScan-1) was running and he was awarded the 1997 Appita L.R. Benjamin medal. Soon after completion of SilviScan-1, an Australian Government Cooperative Research Centre was established with sufficient funding to develop SilviScan-2, which came to the attention of the Marcus Wallenberg Prize Selection Committee, leading eventually to a very enjoyable presentation and dinner in Stockholm last year. Also in 2001, he and his team were awarded the CSIRO Chairman’s Gold Medal. Rob is a Fellow of the International Academy of Wood Science and is currently leading the SilviScan-3 development team.

Introduction

Informatics encompasses the generation and management of information. The underlying concepts of informatics are not new to us; they have always been central to human existence.

Effective communication is fundamental to informatics. The first recollection of my scientific education was an explanation by my primary school teacher of the formation of the Solar System according to the theory of Harold Jeffreys and James Jeans. Accompanying the lesson was a striking diagram in coloured chalk of a star passing by the sun, drawing out a long cigar-shaped gas cloud that was to condense into the planets. My only memory of that lesson is the diagram itself. Mrs Troy understood the effectiveness of scientific visualisation, and throughout my career I have relied to a large extent on the graphic arts in my presentations to colleagues and students.

I now work for the Commonwealth Scientific and Industrial Research Organisation (CSIRO), which is responsible for a large part of government research in Australia. Wood fibre research has been a central activity at CSIRO Forestry and Forest Products for several decades. Many of the questions we ask today are essentially the same as those asked fifty years ago: What are the key wood fibre properties influencing the properties of pulp, paper and wood products? What is the natural variation of each of these properties? How can we include wood fibre quality traits in tree improvement programs?

Some of these questions were asked and answered in part by CSIRO researchers such as Alan Wardrop, Alf Watson and Eric Dadswell (Wardrop, 1951; Watson et al, 1952, 1961, 1962, 1964a, 1964b). Alan Wardrop, in the 1950's, was working in my current position at CSIRO while my primary school teacher was inspiring me to become a scientist.

In those days, a large fraction of time was invested in the acquisition of the data required to study variation and to develop relationships between fibre properties and product properties. Technology has now largely removed those limitations, and we are faced with the problem of data overload.

Over the last decade, we have developed systems that automatically acquire huge quantities of fibre property data from small wood samples taken from standing trees. SilviScan combines x-ray diffractometry, x-ray densitometry and image analysis to produce several gigabytes of fibre property data each day in the form of images, tables and lists. The information is used for a wide range of purposes such as gene discovery, resource evaluation, genetic selection and product property prediction.

It is fitting that Alan Wardrop, now Emeritus Professor at La Trobe University, was an advisor to the research program that produced SilviScan-2.

Throughout the development of SilviScan I was unaware of the existence of informatics as a discipline. It is ironic, therefore, that I am giving this lecture using SilviScan as an example of an application of informatics principles. The development of SilviScan is indeed a journey through the field of informatics - including theory, design, data acquisition, storage, reduction, visualization, interpretation and communication. In this presentation, I will recount highlights of the journey, which, after 14 years, we have barely begun.

Back on track

When I joined CSIRO in 1985, I left an industrial position in which there was little prospect of continuing in research. Peter Nelson, my manager at Australian Paper Manufacturers, encouraged me to make the move, and was admonished for allowing me to do so.

I had no idea what to expect at CSIRO, except a longer career path in applied research. At the time, CSIRO was fully funded by the Australian Government, although there was little money beyond salaries. I joined a Division that had very recently acquired a new Chief, Warren Hewertson, and had moved into new, modern premises. Warren came from ICI and his first actions were to make the work of the Division more relevant to industry. My experience at APM was probably an important factor in his decision to employ me.

When this work began in 1988, CSIRO Forestry and Forest Products was served by the Wood Fibre Research Advisory Group (WFRAG). The group consisted of technical leaders in the Australian pulp and paper industries, as well as CSIRO personnel. I was appointed to a WFRAG working party charged with the task of improving pulpwood quality in plantation trees. It was immediately evident that the current techniques for fibre characterisation would not be capable of dealing with the thousands of samples generated by genetic and silvicultural trials. In addition, evaluation of the fibres produced by a large range of pulping, bleaching and refining processes for an even larger range of products would have been impractical.

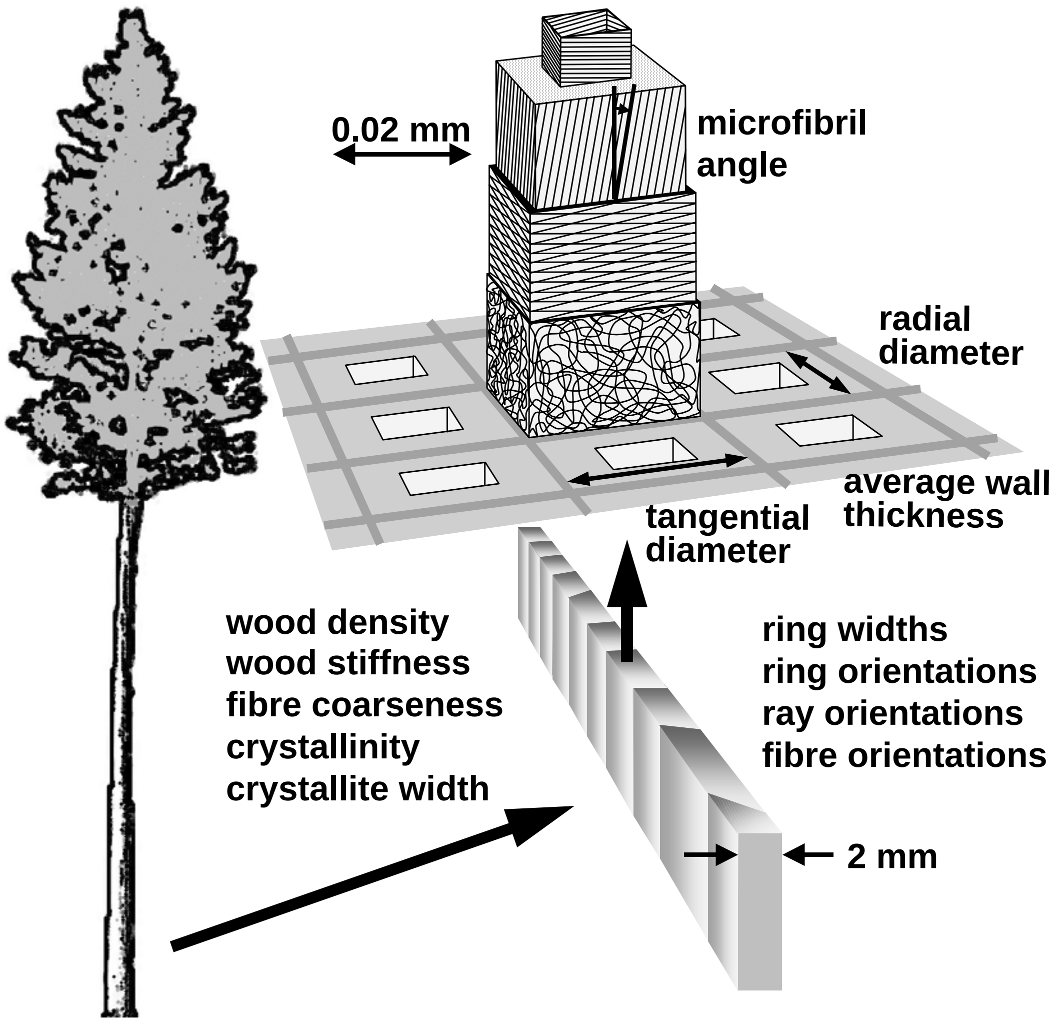


Figure 1. Sampling and wood microstructure

Why wood?

It became clear that efforts should be concentrated on methods for the measurement of fibre properties in solid wood for the following reasons:

1. Wood is the raw material common to all processes and products, therefore the range of samples to be characterised is minimised.
2. Wood is the starting point for all current and future processes and therefore knowledge of wood fibre properties will always be of value.

3. Fibres in wood are arranged approximately in parallel and in a reasonably predictable pattern, so that image processing and analysis algorithms can be greatly simplified.
4. Wood cells fill the available space, allowing the indirect estimation of properties such as coarseness (mass per unit length) and wall thickness.
5. Much of the biological history of the tree is encoded in the radial patterns of wood fibre property variation. This makes possible the evaluation of effects of environment and silvicultural treatment on wood properties and the development of models to predict future resource properties. Genetic inheritance can also be estimated as a function of tree age, giving a sound statistical basis for early selection.

The most practical method for non-destructive sampling of trees has long been increment coring at breast height using a hollow drill bit. 12mm diameter increment cores were chosen to be the basic sample type for our analyses. After drying, cores are cut with a twin-blade saw to 2mm thickness in the tangential direction and 7mm in the longitudinal direction (Figure 1). Softwoods are extracted with acetone to remove resin, and then one transverse face is polished to reveal the fibre cross-sections. Note that I use 'fibre' and 'tracheid' interchangeably in this lecture.

SilviScan-1

I envisaged a machine that could rapidly and automatically measure many properties such as density, fibre diameter, wall thickness and coarseness, and their distributions, in solid wood samples cut from increment cores. Figure 1 illustrates the sample cut from a core taken from breast-height, together with an illustration of the wood microstructure and the properties that I hoped to measure on each sample. Our 1964 x-ray machine was due for replacement so I was able to obtain funding to build my measurement system. Some parts from the old x-ray machine were included to save money. My Program Manager, Geoffrey Gartside gave me considerable freedom to explore ideas. For the next 2 years we worked on the system with excellent workshop and programming support. Not many staff members knew what I was doing. The fact that I worked in a very small lab inside another lab with no handle on the outside of my door may have contributed to my seclusion. Fortunately for my ongoing employment, SilviScan-1 eventually lived up to expectation.

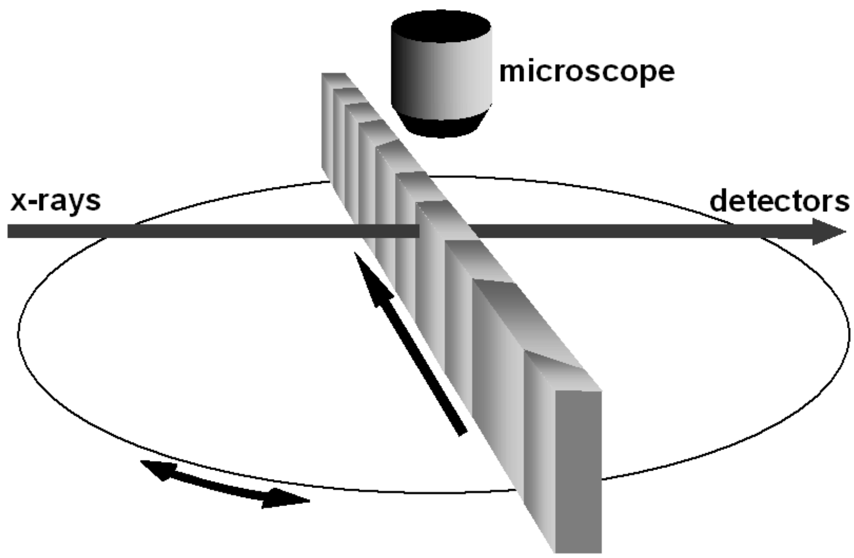


Figure 2. Schematic diagram for automated measurement of wood microstructure

By combining x-ray densitometry, x-ray diffractometry and image analysis (Figure 2), I expected to estimate more properties than was possible with each technique alone. Coincidentally, the idea was based in part on a method used by Ken Britt (1965, 1966), for whom my wife Lynn worked when we were in Syracuse over a decade earlier. Ken counted the fibres in the cut ends of small wooden dowels to estimate the average cross-sectional area of the fibres. He then multiplied this area by the density of the dowel to obtain fibre coarseness. Coarseness is the mass of wall substance in a given length of fibre and was regarded as key determinant of pulp and paper properties. Later, a similar method was used by Scallan and Green (1974, 1975) to estimate coarseness in blocks of wood. They also estimated fibre wall thickness, as did Harris (1969) in New Zealand. Those researchers demonstrated that properties like coarseness and wall thickness, which were very difficult to measure directly, could be estimated rapidly from density and fibre area, which are far more easily measured.

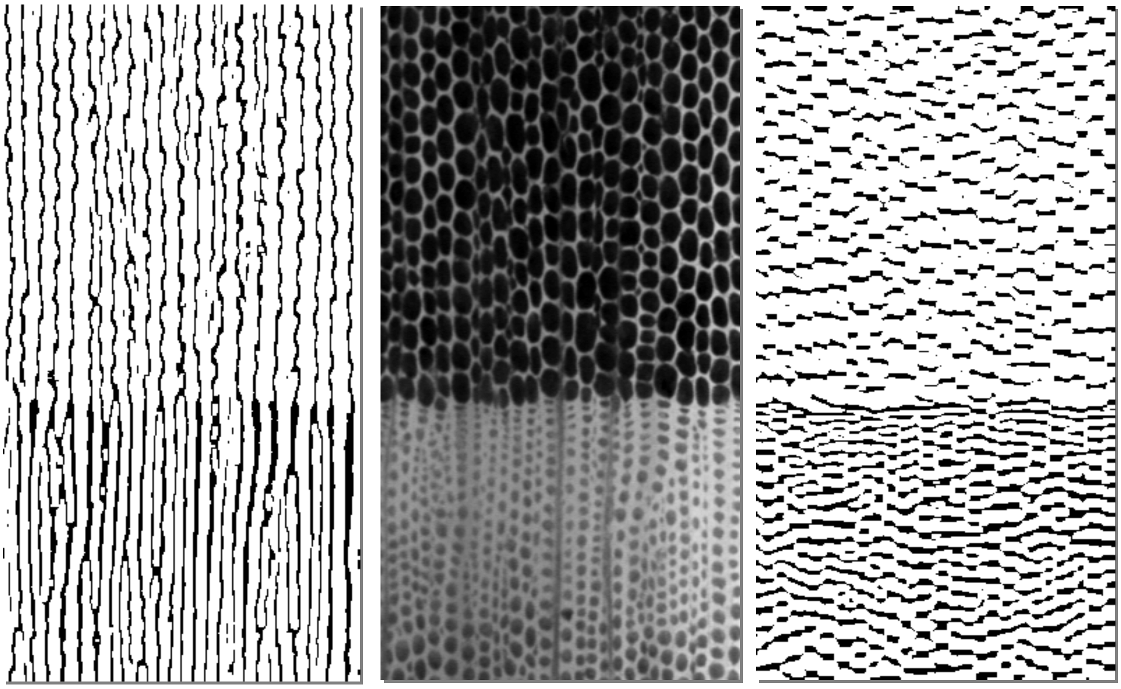


Figure 3. Transmitted light illumination of fibre cross-sections in the wood (centre), together with extraction of fibre wall images in the radial (left) and tangential (right) directions

There were many problems to overcome before an automated analysis system could be built. For example, high-contrast images of fibre cross-sections in wood are very difficult to obtain consistently. I tried many methods to improve the images and eventually discovered that the most effective technique was to use transmitted light for imaging. The light travels along the fibre walls to give the effect shown in Figure 3 (centre). The images are processed to identify the fibre walls in the radial and tangential directions in the tree. Fibre diameter and fibre area are reliably estimated from these processed images. The angles of the annual growth rings across the sample are also measured by image analysis and used to automatically align the growth rings with the x-ray beam in the subsequent densitometry scan. As mentioned above, the density and fibre cross-sectional area data are combined to give coarseness and wall thickness data. All properties are obtained at 50-micron intervals and summary reports are produced (Figure 4). The bottom three profiles in Figure 4 (fibre coarseness, wall thickness and specific surface area) are calculated from the directly measured profiles (radial diameter, tangential diameter and density). Many tables, text files and binary files are also produced for more detailed examination and further processing.

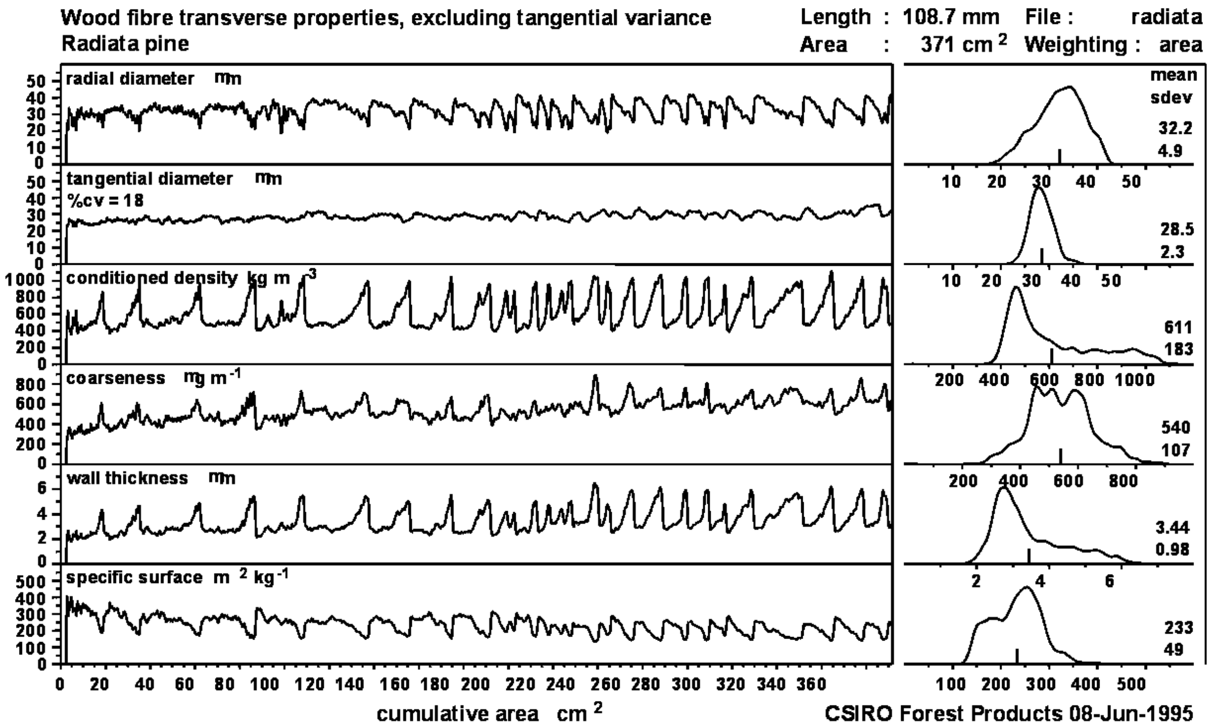


Figure 4. Summary report sheet from SilviScan-1

So far, I have described very briefly the development of SilviScan-1, which was completed by about 1992. I published the first SilviScan paper two years later (Evans, 1994). By this time, Paul Kibblewhite (Forest Research, New Zealand) and I were in a long and fruitful collaboration (eg, Evans et al, 1997, Kibblewhite et al, 1997). Computer technology was struggling to keep up with our needs at that time. In particular, memory was a limiting factor, preventing substantial improvements in the software. Nevertheless, many thousands of samples were analysed using this system.

SilviScan-2

The success of SilviScan-1 was carried into an Australian Government funded Cooperative Research Centre (CRC). The CRC for Hardwood Fibre and Paper Science included CSIRO Forestry and Forest Products, Monash University, The University of Melbourne and the Pulp and Paper Manufacturers Federation of Australia. During the 7 years from 1993-1999, sufficient funds were available to develop the second-generation system, in which I included a high-speed x-ray diffraction system for estimating properties associated with wood stiffness and strength. The development of SilviScan-2 was less straightforward than that of SilviScan-1. I had moved into a much larger laboratory, and this time everyone knew where I was. There were many visitors, many reports, many meetings and many reorganisations. Hardware problems also delayed the work. It was especially gratifying to turn on the x-ray system and watch the wisps of white smoke rising from the electronics. Once these problems were fixed, however, we found the system to be extremely reliable. SilviScan-2 has now been running almost continuously for over 42000 hours, stopping only for regular servicing every 6-8 months.

SilviScan-2 evolved in parallel with computer networking, 32-bit processing and rapidly expanding memory capacity. We took advantage of these developments and generated a system centred on a local server with a high-capacity database. Computers on the local network access the database for off-line processing of images and other SilviScan data. We produce many Gbyte of data per day, most of which must be discarded, therefore it is critical that we extract most of the information soon after acquisition. One of our clients once asked for all the raw data associated with 24 short samples. After receiving 52 CDs, they decided not ask again.

An important angle

The x-ray diffractometer on SilviScan-2 has proven to be the most significant addition to the technology. Long-chain cellulose molecules, grouped in bundles called microfibrils, are arranged in a helical fashion around the fibre axis. The angle of this helix is called the microfibril angle (MFA). When the MFA is small (close to the fibre axis) the fibre is stiff and strong. Traditionally, measurement of MFA is slow and labour-intensive. In 1997 I found a way to estimate very rapidly the average MFA in large numbers of wood fibres using x-ray diffraction (Evans, 1999). Acquisition and analysis of each diffraction pattern takes as little as 10 seconds. A sample of length 200mm may be analysed at 200-micron intervals (1000 MFA measurements) in only 3 hours. The method is particularly suitable for estimation of average MFA from an average diffraction pattern, thereby allowing all fibres (up to 3 million) in a sample cross-section to be scanned in several minutes and to be represented in an average MFA for the whole sample.

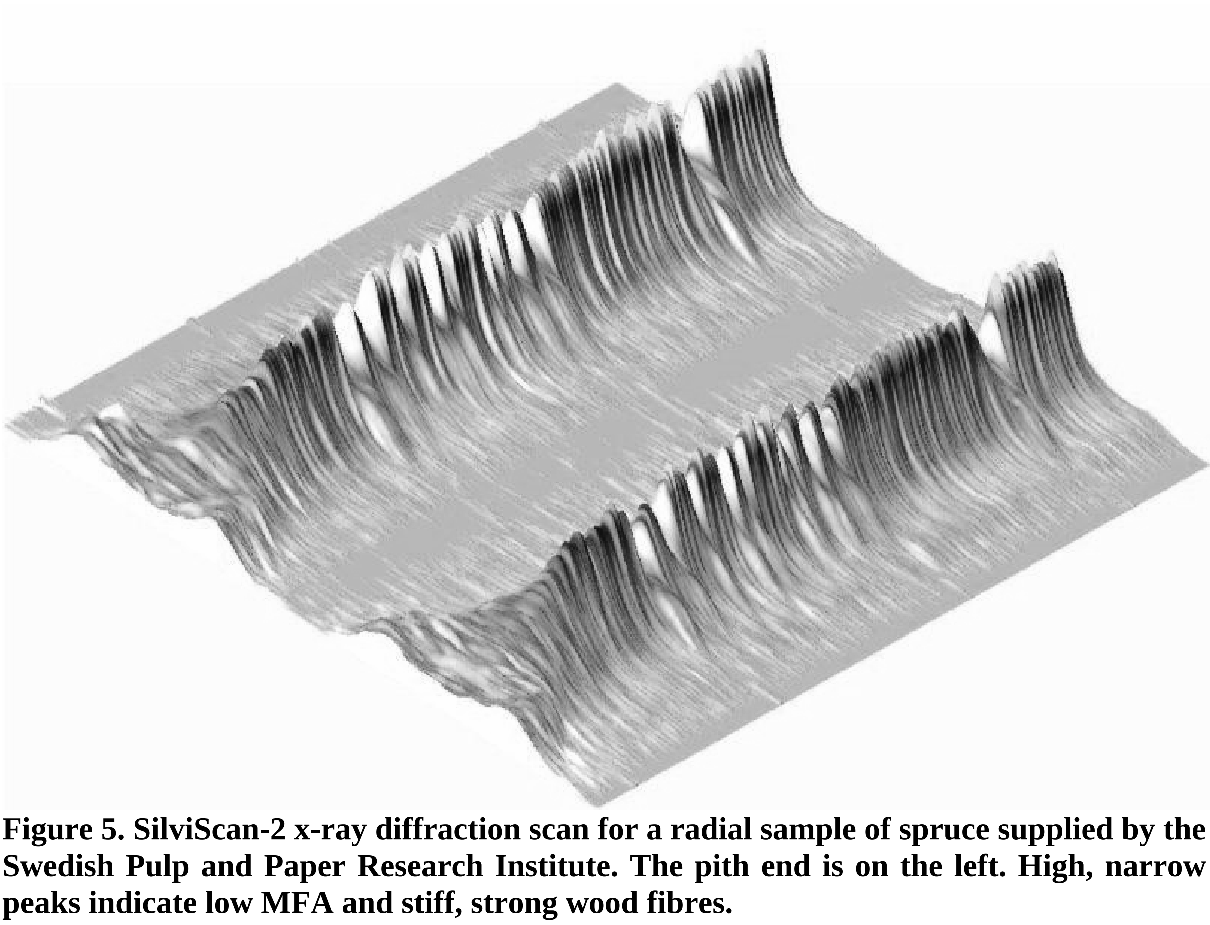


Figure 5. SilviScan-2 x-ray diffraction scan for a radial sample of spruce supplied by the Swedish Pulp and Paper Research Institute. The pith end is on the left. High, narrow peaks indicate low MFA and stiff, strong wood fibres.

Without going into detail, there are two major peaks in a diffraction pattern from wood. To a first approximation, MFA is proportional to the width of these peaks. Figure 5 shows a diffraction scan for a sample of spruce. The image consists of over 1000 individual diffraction profiles, stacked to form a 3-dimensional surface. The raw data upon which this image is based occupied over 2Gbyte. On the left of the image is the pith end of the sample. Notice that the peaks are very broad, indicating that MFA is large and that the fibres are weak and very low in stiffness. As we move towards the right (bark end), we see the peaks become narrow, indicating an increase in strength and stiffness. Periodically, however, there are sections with large MFA, possibly caused by bending stresses on the tree stem resulting in the formation of compression wood.

Stiffness

In 2000, I discovered a robust method for predicting wood stiffness from SilviScan's diffraction and density information. We now use the method routinely for all wood species, although I would like to do confirmatory experiments on more hardwoods. Some day I will finish writing the paper on this method.

Back to the trees

If samples are taken from a range of heights within a tree, the spatial variation of properties can be determined by interpolation. Using this method on radiata pine, strong relationships have been found between the interpolated whole tree average properties and the corresponding whole tree kraft pulp fibre properties. More importantly, wood properties in breast-height cores correlate well with whole tree properties. Figure 6 shows examples of whole tree maps of several properties estimated by SilviScan. These maps represent averages for 25 13-year-old radiata pine trees grown in New Zealand and contain information from almost 30 million fibres.

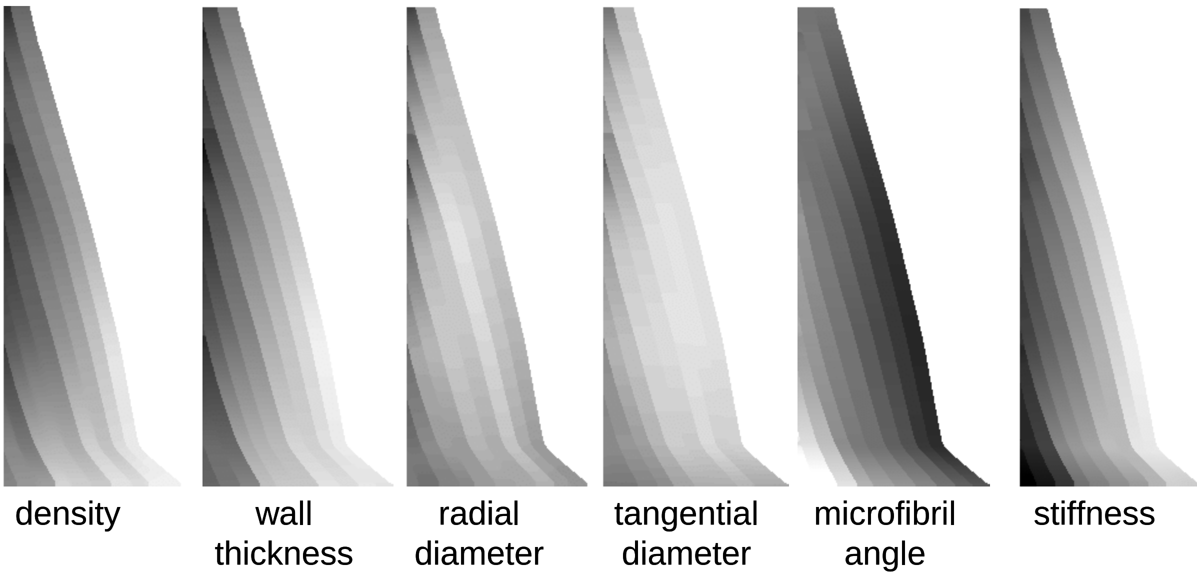


Figure 6. Maps of wood properties averaged over 25 13-year-old radiata pine trees. Lighter shades indicate higher values.

SilviScan-3

SilviScan is intended to be a development platform as well as a system for routine analysis. The next generation of instruments is currently under development and these will be installed in Sweden next year. As new techniques are discovered for analysing solid wood, they will be added to SilviScan. Currently, we are incorporating NIR spectroscopy. In time, I hope to see a network of laboratories improving the technology and exchanging basic information on wood microstructure. I am an optimist.

Scientific visualisation

The images used in this lecture were designed to attract attention, illustrate concepts and reinforce memory. The oral presentation also uses colour and contains many more examples than shown in this document. Choices of orientation, lighting, relative proportions, etc, should not be made for their own sake because the message may be obscured. If possible, artistic and scientific concepts should be synergistic to maximise the effectiveness of the message. Some of these concepts are represented in the Figures:

Figure 6 portrays the spatial variation in several wood properties averaged over 25 trees. Here, I have chosen to use an area-based mapping scheme in which horizontal distance in the image is proportional area of wood produced. This makes the whole volume of the tree visible in a 2-dimensional image. Alternative renderings can be based on a scale of mass or distance (section through the tree). A colour or grey scale may be added for quantitative work. If colour images are used in print, the colour map should be chosen such that the meaning is not lost when converted to grey scale. For example, a red-blue scale will not convert or photocopy properly because red and blue convert to the same grey level.

The diffraction image in Figure 5 was designed in 3-dimensions to convey an impression of strength where the peaks are high and narrow, and weakness where the peaks appear collapsed near the pith. Information on the three-dimensional orientation of the fibres in the wood is present in the same data but a different (2-dimensional) image must be generated to make this apparent.

The complex report sheet in Figure 4 is primarily used in printed form, as it summarises most of the information available from SilviScan-1. In hindsight, I should have rotated the distributions on the right hand side to make their scales match those of the profiles themselves. Colour may be used to highlight specific areas of such a complex diagram.

Figure 3 shows a section of an image with the results of processing on either side. It is easy for the observer to appreciate simultaneously the effects of two processing paths. The high contrast in the original image (resulting from transmitted light illumination) is also evident.

Figure 2 renders SilviScan down to its essential elements and retains a semblance of reality in the structure of the wood sample. I have done this to emphasise the importance of the wood itself and to remind the observer that the annual rings are not necessarily perpendicular to the sample axis. This is the reason for the rotation stage.

Finally, Figure 1 was designed for a poster and therefore conveys many pieces of information for use in an extended discussion. The same poster additionally contained Figure 2.

I thank Professor Paul McFarlane for the honour of presenting the 2002 Burgess-Lane Lecture. I hope I have been able to demonstrate the value of artistic concepts in communicating complex data and ideas in wood science. Scientific visualisation is more than the use of default graphs, charts and fonts. With a small amount of effort, you can make considerable improvements. If your images are pleasing to the observer, and clearly illustrate the science, your message is more likely to get through. It might even be remembered for as long as Mrs Troy's 1956 astronomy lesson.

References

Britt, K.W. 1965. Determination of fiber coarseness in wood samples. Tappi *48*, 7-11.

Britt, K.W. 1966. Fiber coarseness in wood. Tappi *49*, 202-206.

Evans, R. 1994. Rapid measurement of the transverse dimensions of tracheids in radial wood sections from *Pinus radiata. Holzforschung* **48** (2):168.

Evans, R. 1999. A variance approach to the x-ray diffractometric estimation of microfibril angle in wood. *Appita J*. **52** (4):283.

Evans, R., Kibblewhite, R.P. and Stringer, S. 1997. Kraft pulp fibre property prediction from wood properties in eleven radiata pine clones. *Appita J.* **50** (1):25.

Harris, J.M. 1969. Use of beta rays in determining wood properties. Part 4 - Estimating the thickness of tracheid cell walls. New Zealand J. Forest Sci. *12*, 432-442.

Kellogg, R.M., C.B.R. Sastry, and R.W. Wellwood. 1975. Relationships between cell-wall composition and cell wall density. Wood and Fibre *7*, 170-177.

Kibblewhite, R.P., Evans, R. and Riddell, M.J.C. 1997. Handsheet property prediction from kraft fibre and wood tracheid properties in eleven radiata pine clones. *Appita J.* **50** (2):131.

Scallan, A.M. and H.V. Green. 1974. A technique for determining the transverse dimensions of the fibers in wood. Wood and Fibre *5*, 323-333.

Scallan, A.M. and H.V. Green. 1975. The effect of pulping upon the dimensions of wood tracheids. Wood and Fibre *7*, 226-233.

Wardrop, A.B. 1951. Cell wall organisation and the properties of the xylem. I. Cell wall organisation and the variation of breaking load in tension of the xylem in conifer stems. Aust. J. Sci. Res. B-4: 391-414.

Watson, A.J. and H.E. Dadswell. 1961. Influence on fibre morphology on paper properties. Part 1. Fibre length. Appita *14*, 168-176.

Watson, A.J. and H.E. Dadswell. 1962. Influence on fibre morphology on paper properties. Part 2. Early wood and latewood. Appita *15*: 116-127.

Watson, A.J. and H.E. Dadswell. 1964a. Influence on fibre morphology on paper properties. Part 3. Length:diameter (L/D) ratio. Appita *17,* 146-150.

Watson, A.J. and H.E. Dadswell. 1964b. Influence on fibre morphology on paper properties. Part 4. Micellar spiral angle. Appita *17,* 151-156.

Watson, A.J., A.B. Wardrop, H.E. Dadswell and W.E. Cohen. 1952. Influence of fibre structure on pulp and paper properties. Appita Proceedings *6*, 243-266.