2008:49

LICENTIATE THESIS

On Fingerprint Traceability in the Forestry Supply Chain

Luleå University of Technology
LTU Skellefteå
Division of Wood Science and Technology

2008:49 | ISSN: 1402-1757 | ISRN: LTU-LIC -- 08/49 -- SE

Abstract

Tracing material through the forestry supply chain is a relatively untapped source for process and product improvements. Since the supply chain handles large quantities at high speeds and has a diverging flow and several different subprocesses and people involved along the way, traceability rapidly becomes very complex. The papers presented in this thesis have focused on tracing wood material by means of the fingerprint approach. The fingerprint approach rests on the foundation that each piece of wood is a unique individual with unique features and that it is possible to identify and connect individual pieces in the supply chain in the same way that human beings can be identified by the use of their fingerprints.

The results from Paper I show the importance of preserving the shape of the log and handling bark assessment at an individual level when trying to connect logs between the log-sorting station and saw intake using their 3-D outer shape.

Paper II and Paper III show very encouraging results in connecting sawn timber to the log they were sawn from by using 3-D and x-ray data for the logs and surface scanning for the sawn timber. The results show that over 95% of the sawn timber could be connected to the correct log.

Keywords: Fingerprint, Traceability, Log sorting, Green sorting, X-ray log scanner, 3-D scanning, Surface scanning

Preface

When I left upper secondary school, I was fed up, and I swore I would never to return to any sort of education, and look what happened!

This thesis was carried out at Luleå University of Technology, the Department of Wood Technology, Skellefteå Campus, under the supervision of Professor Anders Grönlund and Associate Professor Johan Oja. I would like to express my sincere gratitude to my supervisors for giving me very free reins in my work and at the same time always being there for valuable advice and insight. I would also like to express my general gratitude to coworkers at SP and the University for creating a very enjoyable atmosphere to work in.

The last and greatest thank-you goes to my wife Pia and our wonderful children Ella, Annie and our newborn Hugo. Thank you for taking my mind to things that really matter.

Skellefteå, December 2008

Jens Flodin

List of papers

I. Flodin, J. Oja, J. Grönlund, A. 2008. Fingerprint traceability of logs using the outer shape and the tracheid effect. For. Prod. J. 58:4.

II. Flodin, J. Oja, J. Grönlund, A. 2007. Fingerprint traceability of sawn products using x-ray log scanning and sawn timber surface scanning. Proceedings of COST E53-Quality control for wood and wood products. October 15-17. 2007. Warsaw. Poland.

III. Flodin, J. Oja, J. Grönlund, A. 2008. Fingerprint traceability of sawn products using industrial measurement systems for x-ray log scanning and sawn timber surface scanning. For. Prod. J. 58:11.

The collection of data in Paper II and Paper III was performed by Johan Skog and Johan Oja. All other work in the papers presented here has been performed by the author with valuable insight and guidance along the way from Johan Oja and Anders Grönlund.

Introduction

Traceability can be defined in many different ways. Töyrylä (1999) defines traceability as follows: "Traceability is the ability to preserve and access the identity and attributes of a physical supply chain's objects." The ability to identify and connect a specific manufactured object between positions in the supply chain brings an abundance of opportunities when it comes to controlling the quality of that object and the process that produced it.

An issue of growing interest for today's sawmills is the utilization of the raw material, *i.e.*, producing the most suitable product from each specific log. If this can be achieved, there is a major benefit to be gained when the production of products that don't meet quality requirements can be reduced, along with the loss in revenue that these products bring. In order to obtain knowledge about the relationship between the raw material and the sawn product, one needs individually associated data between the two.

With individually associated data, it is subsequently possible to build sorting models in which the inner and outer characteristics of the raw material can be connected to a specific quality and/or volume yield of the sawn product. Traditionally, this connected data has been the product of test sawings in which logs have been manually marked and then tracked and recorded from the harvester or the sawmill's log sorting through the green sorting or trimming station. This is, however, a time- and money-consuming task, which suggests that an automated technique for achieving the individually associated data would be well worthwhile.

Modern sawmills often have sophisticated measurement equipment, such as x-ray (Anon. 2008d) and 3-D scanning (Anon. 2008c, Anon. 2008e) for logs and surface scanning (Anon. 2008a) for sawn timber, that generates large quantities of data at an individual level. These data are collected at certain points along the production chain, but are unfortunately almost exclusively used as a means to control the production process close to the measurement point. Most of the generated data for a specific piece of wood are therefore discarded after the piece has moved past the measurement point. The challenge is therefore not to generate data, but to connect the generated data to individual pieces. If the data for each specific piece were to be collected and stored in a database, the final product could then "be considered as an information intensive product," (Uusijärvi 2003) and the database could be used for online process control. If the purpose of the traceability is offline process control, only a partial amount of the total flow needs to be tracked, which would require a less complex and expensive system compared to an online system. An online system might therefore be overkill, depending on the purpose. In order to be cost efficient, it must be proven that a system where every piece is traced gives large benefits. For strength-grading purposes, Brännström *et al.* (2007) could not find such benefits from the traceability-based data. On the other hand, it is shown that there is a large potential for process optimization in using data that can be collected by a system for tracing a partial amount of the flow (Broman *et al.* 2007).

Since sawmills have a diverging flow, and modern sawmills have high production speed, the tracing and storing of data is not well suited for manual labor. A better alternative for handling the tracing and tracking is some form of automated identification (McFarlane and Sheffi 2002). There are a number of alternative methods for practically making the connection between measurement data and the individual piece of wood. Many of these alternatives are based on some form of marking/reading technique. Two well-known methods are barcode identification (Palmer 1995) and radio frequency identification (RFID) (Finkenzeller 2003). Barcode identification is a noncontact method used in almost every supermarket checkout counter in which the bars in the code are optically read by a laser scanner. RFID is also a noncontact method, wherein an antenna picks up the RFID tag's unique identification number when it enters the antenna's reading range. For forestry traceability applications, RFID is probably better suited due to the fact that the tags can be read without an optical scan, thus making the dirt and handling involved in logging almost without influence on the reading result, as opposed to reading barcode identification under the same circumstances. The drawback is the price for the RFID tags. A sawmill that produces 150,000 m^3 of sawn wood and has an average log volume of 0.18 m^3 handles approximately 1.8 million logs annually. The price for RFID tags is approximately 1–2 € ($0.75–$1.50) per tag (Uusijärvi 2003). If every log is to be tagged, the annual cost for tags alone will then be millions of dollars.

An alternative and more cost-effective way of identifying individual pieces of wood is to use the already existing measurement data and make identification by means of the fingerprint approach (Chiorescu 2003). The fingerprint approach rests on the foundation that each piece of wood is a unique individual with unique features. These can be the piece's outer as well as inner features. If one could measure these individual features accurately enough, it would then be possible to identify individual pieces in the production chain in the same way that human beings can be identified by the use of their fingerprints.

The objective of this research is to show the so far rather hidden possibilities that lie in being able to connect individual pieces in the forestry supply chain. The three papers presented here focus on making further advancement in the development of the fingerprint approach towards a traceability concept in the forestry supply chain. All results from the three papers are derived and thus limited to Scots pine that has been harvested in northern Sweden.

Paper I is a spinoff from earlier results by Chiorescu (2003) in tracing logs between the sawmill's log-sorting station and saw intake using the log's 3-D outer shape. Chiorescu found that bark had a very negative effect on the fingerprint recognition rate. The aim of Paper I was therefore to investigate if a more individualized bark assessment based on the tracheid effect could improve the recognition rate between the log-sorting station and the saw intake.

Paper II and Paper III describe a new fingerprint-approach method in which the possibility was investigated of connecting logs from the log-sorting station to sawn center yield planks from the green-sorting station using 3-D and x-ray data for the logs and surface-scanning data for the planks. Paper II is an initial study, and Paper III is an extended study to verify and enhance the results found in Paper II.

Material & Methods

Paper I:

This study was hosted by a mid-size sawmill situated in northern Sweden with an annual production of approximately 150 000 m^3 of sawn timber. The sawmill handles both Scots pine (*Pinus sylvestris*) and Norway spruce (*Picea abies* (L.) Karst.). Scots pine was chosen in this study since the species has three different types of bark along the stem, which makes bark assessment more problematic compared to Norway spruce, which only has one type of bark.

The data collection involved three of the stations at the sawmill. These were, sequentially, the log-sorting station, a combined debarker/butt-end reducer and the saw intake. The butt-end reduction was performed by a knifed rotating ring with fixed diameter (see Figure 1). The diameter of the ring used in this study was 320 mm.

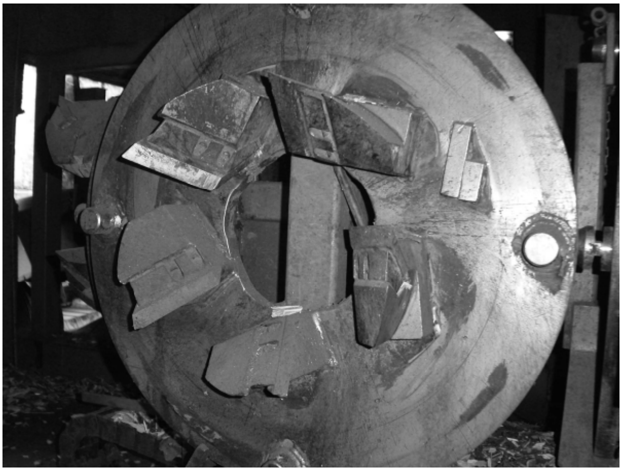

Figure 1: Knifed ring for butt-end reduction

The study involved two groups with 50 randomly chosen logs in each group. The first group consisted of small-sized logs with a top diameter in the range of 148–154 mm. The second group consisted of larger-sized logs with a top diameter in the range of 253–265 mm. The second group was chosen because the diameter class in the 253–265-mm range is the largest diameter class on which the 320-mm butt-end reducer ring is used in normal production. Most of the logs in the larger-sized group were therefore greatly affected and consequently shape-changed after they had gone through the butt-end reducer (see Figure 2).

Figure 2: The same log before and after butt-end reduction

The measurement equipment used to generate and collect the measurement data was two identical Sawco 3-D scanners (Anon. 2008e) that were installed at both the log-sorting station and the saw intake. These scanners are used in the sawmill's normal production. The log sorting station scanner was equipped with Sawco's ProBark application that uses the tracheid effect to automatically assess whether the measured surface is bark or clear wood and then uses this information to calculate bark thickness and perform bark compensation. Figure 3 shows the difference in how the laser line is spread in bark and in clear wood.

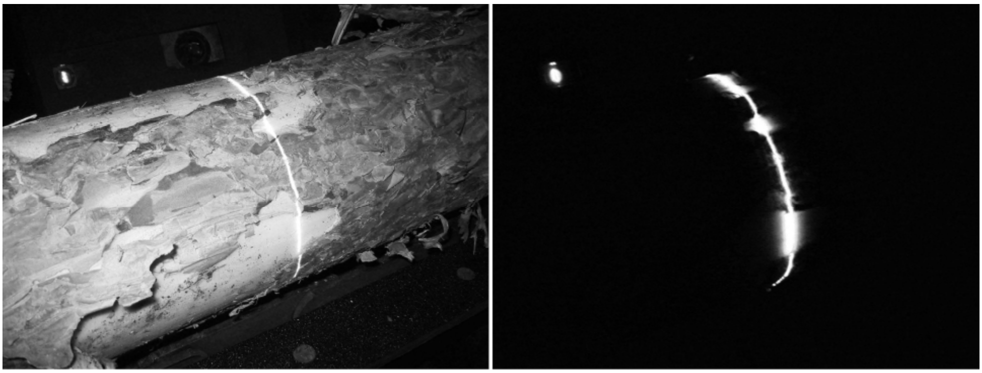

Figure: 3: Log photographed with flash (left) and without flash (right)

The recording of raw data was done with the automatic bark assessment (ABA) both active and inactive. When ABA was inactive, the raw data files were recorded with the bark present in all cross sections. When ABA was active, each cross section in the raw data files was compensated with the individual log's calculated bark thickness. With ABA active, one could think of the logs as "virtually debarked".

After the log data had been collected at the saw intake and at the log sorting station with ABA both active and inactive, the data was analyzed using MatLab (Anon. 2008b). Eleven descriptive measurements were calculated for each log (Table 1) using the raw data from both the log sorting station and the saw intake.

Table 1: Measurements calculated for each log

Variables	Description
Length (Len)	Distance from the log's top to butt end (mm)
Physical Volume (PhV)	The log's physical volume (dm3)
Top Diameter (ToD)	The log's top-end diameter (mm)
Middle Diameter (MiD)	The log's middle diameter (mm)
Butt Diameter (BuD)	The log's butt-end diameter (mm)
Total Taper (ToT)	The absolute value of the change in diameter per meter of log length from the top end to a point measured one meter from the butt end (mm/m)
Top Taper (TpT)	The absolute value of the change in diameter from the top end of the log to a point measured one meter from the top end (mm/m)
Butt Taper (BuT)	The absolute value of the change in diameter from the butt end of the log to a point measured one meter from the butt end (mm/m)
Bow Height (BoH)	The maximum distance between the log's curvature and a straight line connecting the center of the log's end surfaces (mm)
Bow Radius (BoR)	The radius of a circle fitting the log's length and bow height (m)
Bow Position (BoP)	The distance from the log's top end to the point of maximum bow height (mm)

The measurements from the log-sorting station data were calculated with no bark compensation, bark compensation with traditional bark functions (Zacco 1974) and bark compensation with ABA. The fingerprint matching was done by means of score values from multivariate principal components (Wold *et al.* 1987; Eriksson *et al.* 2001). This method is well suited for handling data that might contain some noise and has shown good results in previous studies (Chiorescu 2003).

The matching was done by taking one log at a time from the saw intake and comparing it, one by one, to all logs from the log-sorting station. The actual comparison was done by using the score values and calculating a Euclidian distance value. For the log from the log sorting station that showed the smallest Euclidian distance value for a specific log from the saw intake, a positive identification was considered to have been established between the two places.

After the first findings, the raw data files from the large-size group were virtually crosscut in the butt end in an attempt to enhance the results and handle the problem with the butt-end reducer. Matching was, after the crosscutting, performed on the remaining part of the log.

Paper II & Paper III:

These studies were hosted by a large-size sawmill situated in northern Sweden with an annual production of approximately 400 000 m^3 of sawn timber. The sawmill handles only Scots pine (*Pinus sylvestris*), which was thus the only material used in the two papers. Data for the study were collected from a total of 435 logs with top diameters in the interval 153–321 mm. All logs were sawn by 2-ex patterns into 870 pieces of center-yield planks of various dimensions (see Table 2).

Table 2: The Scots pine material used in the study

Group	Logs		Planks		
	Quantity	Top diameter (mm)	Quantity	Thickness (mm)	Width (mm)
1	70	153–187	140	50	100
2	70	174–213	140	50	125
3	70	193–229	140	50	150
4	40	208–260	80	63	150
5	75	225–277	150	63	175
6	110	253–321	220	63	200

All data used in the studies were collected by the industrial measurement systems that are used in the sawmill's normal production. The first point of data collection was the sawmill's log-sorting station where data from the logs were collected with a one-directional x-ray log scanner from Rema Control AB (Anon. 2008d) in combination with a 3-D optical scanner from MPM Engineering Ltd. (Anon. 2008c). The data extracted from these systems were the total length of the logs according to the 3-D scanner and the position and length of the whorls in the logs according to the x-ray log scanner. The second point of data collection was a cross-fed Finscan Boardmaster surface-scanning system (Anon. 2008a) situated at the sawmill's green sorting station. The total length and the positions of surface knots were recorded for each of the sawn planks.

The analysis of the collected data was performed using MatLab. Log groups one and two were used together in Paper II for analysis and construction of the fingerprint-matching algorithm. Log groups three, four, five and six were incorporated in Paper III to verify the results. Paper III also focused on further developing the matching algorithm.

The first step in the data analysis in Paper II was to investigate the correlation between the total length measurements from the log-sorting and green-sorting stations. This was done in order to create a length filter for the matching algorithm. The length filter was made to exclude all logs from the matching procedure that had a length that could not realistically belong to the plank being compared. The matching algorithm was designed to work iteratively by taking one plank at a time and comparing its summarized surface-knot positions with the positions of knot whorls for each log that had passed the length filter (see Figures 4 and 5).

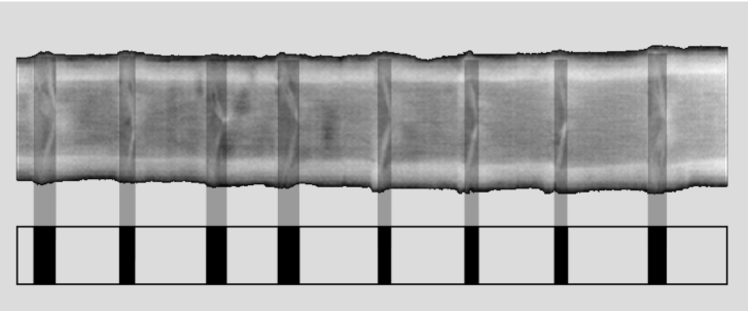

Figure 4: Lengthwise positions of knot whorls in a log.

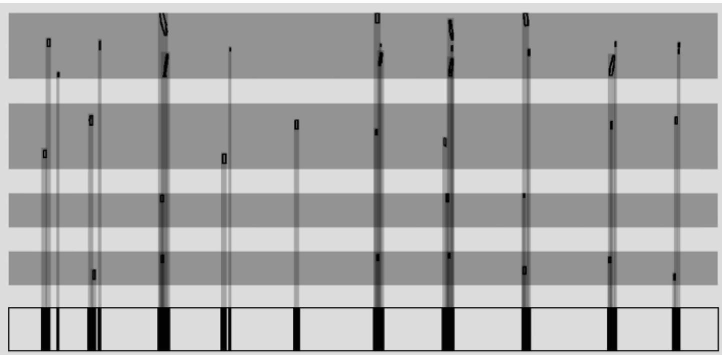

Figure 5: Lengthwise positions of surface knots on a plank. The plank's four faces are summarized (bottom).

The length-filtered log that showed the highest correlation in knot positioning to the plank was considered a positive match. In Paper III, the algorithm was improved and extended with a minimum difference value filter in order to improve the certainty in the matching procedure. The minimum-difference-value filter made it possible to exclude planks from the matching procedure if they were at risk of being matched to the incorrect log.

Results

Paper I

The results from the matching procedure are presented in Table 3. The figures presented for the large-size logs are before virtual crosscutting of the butt end (on the left) and after virtual crosscutting (on the right).

Table 3: Fingerprint recognition rate with different bark compensation methods

Bark compensation	Recognition rate (%) small-size logs	Recognition rate (%) large-size logs	
		Before virtual crosscut	After virtual crosscut
No compensation	77.6	54.0	66.0
Bark functions	83.7	62.0	70.0
ABA	88.4	63.3	76.7

The results show that the recognition rate can be improved with more sophisticated bark assessment and that the results for the large-size logs can be improved by performing a virtual crosscut.

Paper II & Paper III

The results from the matching procedure in Paper II are presented in Table 4.

Table 4: Fingerprint recognition rate, Paper II

Thickness	Width	Number of ingoing planks	Number of correctly matched planks	Percentage of correctly matched planks
50	100/125	280	259	92.5

The results from the matching procedure in Paper III are presented in Table 5.

Table 5: Fingerprint recognition rate, Paper III

Thickness	Width	Number of ingoing planks	Number of correctly matched planks	Percentage of correctly matched planks
50	100/125	280	268	95.7
50	150	140	136	97.1
63	150	80	77	96.3
63	175	150	146	97.3
63	200	220	212	96.4

Table 5 shows that the improved matching algorithm performs better than the original (Table 4) and that the results are consistent regardless the size of the logs and sawn lumber.

Figure 6 shows how the certainty in the matching procedure can be improved by using a minimum-difference-value filter.

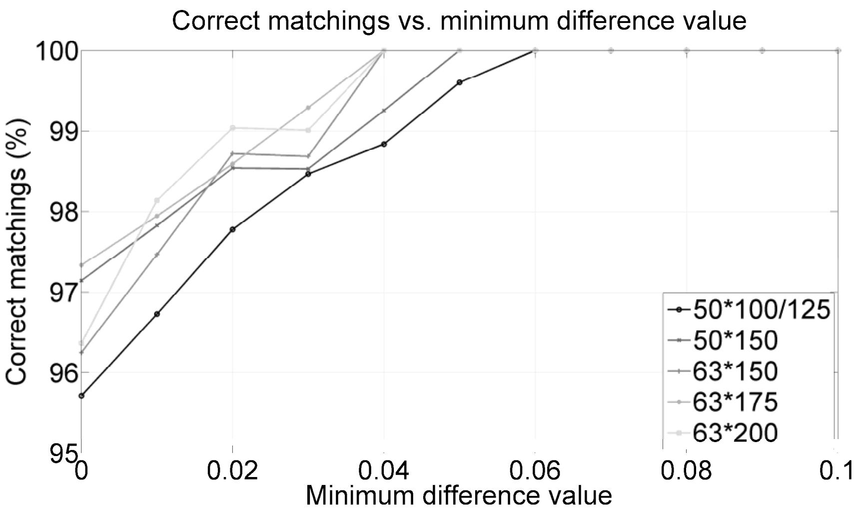

Figure 6: Illustration of how an increased number of correct matches can be achieved with the minimum difference value.

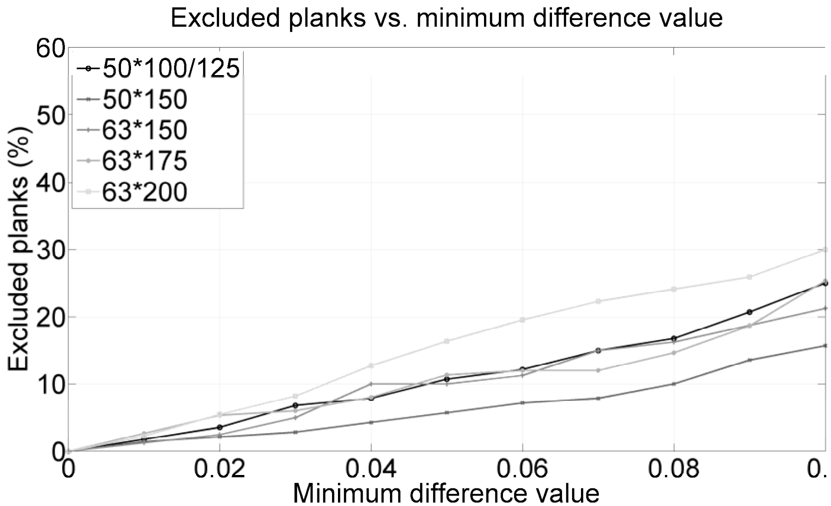

Figure 7: Illustration of how an increasing number of ingoing planks are excluded from the final matching when using a minimum difference value.

If, for example, the minimum difference value is set to 0.06, the recognition rate increases to 100% for all dimensions, while 5%–20% of the ingoing planks are excluded from the matching procedure; *i.e.*, 100% of the remaining planks are correctly matched.

Discussion

Paper I

Based on these results, ABA appears to be a better alternative than traditional bark functions for handling bark compensation on an individual level. This is also in line with what could be expected, since ABA was introduced in order to handle bark compensation on an individual level, as compared with bark functions that are sufficient for handling bark compensation on a group level, but that lack precision to handle it on an individual level.

The virtual crosscutting of the large-size logs did improve the results, but not to the levels that were found in the small-size group. This is probably due to the fact that the virtual crosscutting eliminated some of the natural variation found within the logs, but also due to the fact that log classes with large-size logs don't have the same mixture of butt, middle and top logs that can be found in log classes with smaller-size logs.

The results from this study show that the bark assessment and the integrity of the log's shape are important factors when trying to connect logs with the fingerprint method presented here. This study does, however, need to be validated with a larger test set of independent material. The figures presented here should therefore be viewed in relation to each other rather than as absolute values.

Paper II & Paper III

The results from this study are very encouraging for further development of this fingerprint-tracing method. The results indicate that this method of tracing could be a very cost-effective way to collect and connect data, as opposed to the traditional test sawings, which involve a lot of manual labor in the data collection. The connected data are essential for following up whether changes in process parameters, such as log class limits or sorting towards a certain quality, have had the desired effect, and if not, finding out which specific logs have failed to meet the requirements. The individually associated data could also be used to form the foundation from which to build statistical log-sorting models, since one gets the connection between the logs' inner and outer properties and the sawn planks' quality and volume yield.

Papers II and III have only dealt sawn center-yield planks from pine logs. It is therefore unclear how this method would handle sideboards and sawn lumber from other species such as, for example, Norway spruce. Since sideboards are cut closer to the surface of the log, this could be a potential difficulty, because the knot whorls in the log don't necessarily reach that far, especially in large butt logs. Tracing Norway spruce could also be a challenge, since a lot of minor knots grow in between the main annual whorls, thus making it more difficult to establish position and stretch of knots and knot whorls.

The algorithm can in its present state only handle the case wherein the logs and the sawn timber are of approximately equal length. The algorithm would therefore need some further development in order to handle sawn timber that has been crosscut or taper sawn.

Future research

The results, especially from Paper II and Paper III, are very encouraging for further development. An interesting idea would be to use the method in Papers II and III and add a step in the algorithm that would connect the planks that have been cut from the same log before connecting the planks to the log. By doing this, one would get sawn timber surface information from eight sides (2-ex pattern) instead of four.

It would also be interesting to investigate the possibility of using the method from Papers II and III for tracing sawn center yield from 3-ex and 4-ex patterns as well as sideboards and other species of wood.

References

Anon. 2008a. Finscan Oy, Lukupurontie 2, PL 125, 02201 Espoo, Finland. <www.finscan.fi>.

Anon. 2008b. The MathWorks, Inc, 3 Apple Hill Drive Natick, MA 01760-2098, United States. <www.mathworks.com>.

Anon. 2008c. MPM Engineering Ltd, 200-5465 192: nd Street, Surrey, British Columbia, V3S 8E5, Canada. <www.mpmeng.com>.

Anon. 2008d. RemaControl Sweden AB, Pilgatan 21, SE-721 30 Västerås, Sweden. <www.rema.se>.

Anon. 2008e. Sawco AB, Västra Kvarngatan 62, 611 32 Nyköping, Sverige, <www.sawco.se>.

Broman, O., Nyström, J. and Oja, J. 2007. *Effektiv produktion av golvprodukter med krav på estetiska egenskaper*. Technical report 2007:09. LTU Skellefteå, Department of wood technology. ISSN: 1402-1536.

Brännström, M., Oja, J. and Grönlund, A. 2007. Predicting board strength by X-ray scanning of logs: the impact of different measurement concepts. *Scandinavian Journal of Forest Research* 22(1):60–70. ISSN: 0282-7581

Chiorescu, S. 2003. *The forestry-wood chain, Simulation technique-Measurement accuracy-Traceability concept*. Doctoral thesis 2003:03. Luleå University of Technology, Skellefteå, Sweden.

Eriksson, L., Johansson E., Kettaneh-Wold, N. and Wold, S. 2001. Multi- and megavariate data analysis. Umetrics AB, Umeå, Sweden.

Finkenzeller, K. 2003. *RFID handbook: Fundamentals and Applications in Contactless Smart Cards and Identification*, Second Edition. John Wiley & Sons Ltd., Chichester, West Sussex, England.

Flodin, J. 2007. *Evaluation of transition from manual to automatic bark assessment using a 3D-scanner*. Master's thesis 2007:011. Luleå University of Technology, Skellefteå, Sweden. (In Swedish with English summary.)

McFarlane, D. and Sheffi, Y. 2002. The impact of automatic identification on supply chain operations. *International Journal of Logistics Management* 14(1):1–17.

Palmer, R.C. 1995. *The Bar Code Book*. Helmers Publishing, ISBN 0-911261-09-5

Töyrylä, I. 1999. *Realizing the potential of traceability – A case study research on usage and impacts of product traceability*. Doctoral thesis MA:97. Helsinki University of Technology, Espoo, Finland.

Uusijärvi, R. 2003. *Linking raw material characteristics with Industrial Needs for Environmentally Sustainable and Efficient Transformation processes (LINESET)*. QLRT-1999-01476 Final Report. SP Technical Research Institute of Sweden, Wood Technology (SP Trätek). Research Report no. P 0309034.

Wold, S., Esbensen, K. and Geladi, P. 1987. Principal component analysis. *Chemometrics and Intelligent Laboratory Systems* 2:37–52.

Zacco, P. 1974. *The bark thickness of saw logs*. Swedish Royal College of Forestry, Department of Forest Products. Report No. R 90, 87 pp. (In Swedish with English summary).

I

Fingerprint traceability of logs using the outer shape and the tracheid effect

Jens Flodin
Johan Oja
Anders Grönlund

Abstract

Traceability in the sawmilling industry is a concept that, among other benefits, could be used to more effectively control and pinpoint errors in the production process. The fingerprint approach is a traceability concept that in earlier studies has shown good potential for tracing logs between the log sorting station and the saw intake. In these studies, bark has been identified as a large source of measurement inaccuracy. This study was set out to investigate whether the fingerprint recognition rate could be improved when compensating for bark with traditional bark functions or a new automatic bark assessment based on the tracheid effect. The results show that the fingerprint recognition rate can be improved by using more sophisticated bark compensation. Compared to no bark compensation, improvements can be made by using the existing bark functions, and even further improvements can be made by using automatic bark assessment based on the tracheid effect. The results further show that the butt-end reducer between the log sorting station and the saw intake has a very negative effect on the fingerprint recognition rate, but that significant improvements in the recognition rate can be achieved by excluding the section of the log's butt end that is affected by the butt-end reduction.

Traceability can be defined in many different ways. Töyrylä (1999) defines traceability as follows: "Traceability is the ability to preserve and access the identity and attributes of a physical supply chain's objects." The ability to attach and access the history of a specific manufactured object brings an abundance of opportunities when it comes to controlling the quality of that object and the process that produced it. A good example is the possibility to investigate circumstances surrounding rework and costumer return of faulty products. The ability to trace a product's history makes it possible to isolate and correct errors in the manufacturing process, hence preventing the same errors from occurring again (Wall 1995, Töyrylä 1999). For the same reason, many benefits may accrue as a result of being able to trace products within the wood production industry (Kozak and Maness 2003).

A large-scale issue that is often brought up is the problem with illegal logging. This problem has a negative effect on both the environment and the economy of the affected countries (Dykstra et al. 2003). Traceability would, from this viewpoint, be a way to ensure that harvested logs and their final products originate from a certified harvest site. Since the wood production chain has a diverging flow with a number of people and companies involved in various steps of the handling (Uusijärvi 2000), traceability on a smaller scale could be viewed as a tool for the sawmilling industry to increase knowledge and understanding of factors that influence product quality and the manufacturing process.

Modern forestry and sawmilling companies often have sophisticated measurement equipment that generates large quantities of data at an individual level. These data are collected at certain points along the production chain but are unfortunately almost exclusively used as a means to control the production process close to the measurement point. Most of the generated data for a specific piece of wood are therefore discarded after the piece has moved past the measurement point. If the data for each specific piece were to be collected and stored in a database, the final product could then "be considered as an information intensive product" (Uusijärvi 2003). The challenge is therefore not to generate data, but to connect

The authors are, respectively, Research Scientist, Luleå Univ. of Technology, Dept. of Wood Technology, Skellefteå Campus, Skellefteå, Sweden (jens.flodin@ltu.se); Associate Professor, SP Technical Research Inst. of Sweden, Wood Technology, Skellefteå, Sweden (johan.oja@sp.se); and Professor, Luleå Univ. of Technology, Dept. of Wood Technology, Skellefteå Campus, Skellefteå, Sweden (anders.gronlund@ltu.se). This paper was received for publication in June 2007. Article No. 10369.

Forest Prod. J. 58(4):21-27.

the generated data to each individual piece of wood. The reconnected data would make it possible to investigate and analyze large as well as small sections of the production chain. A good example is the connection between the diameter classes for logs in the log sorting station and the volume recovery of sawn planks and boards. Without reconnection of data, one is reduced to comparing physical volume for a larger group of logs with the physical volume of their planks and boards. With traceability, i.e., reconnection of data, one is given the opportunity not only to analyze and find the individual logs in the group that yield high volume recovery, but perhaps even more important, to find the logs in the group that yield low volume recovery for a specific sawing pattern. Being able to make this distinction then makes it possible to adjust process parameters such as log class limits or sawing patterns for an overall higher volume recovery.

Since sawmills have a diverging flow, and modern sawmills have high production speed, the tracing and storing of data are not well suited for manual labor. A better alternative for handling the tracing and tracking is some form of automated identification (McFarlane and Sheffi 2002). There are a number of alternative methods for making the connection between measurement data and the individual piece of wood. Many of these alternatives are based on some form of marking/reading technique. Two well-known methods are barcode identification and radio frequency identification (RFID). Barcode identification is a noncontact method used in almost every supermarket checkout counter, where the bars in the code are optically read by a laser scanner. RFID is also a noncontact method wherein an antenna picks up the RFID tag's unique identification number when it enters the antenna's reading range (Finkenzeller 2003). For forestry traceability applications, RFID is probably better suited due to the fact that the tags can be read without an optical scan, thus making the dirt and handling involved in logging almost noninfluential on the reading result, as opposed to reading barcode identification under the same circumstances. The drawback with RFID is the price for the RFID tags. A sawmill that produces 150,000 m^3 of sawn wood and has an average log volume of 0.18 m^3 handles approximately 1.8 million logs annually. The price for RFID tags is approximately 1 to 2 € ($0.75 to $1.50) per tag (Uusijärvi 2003). If every log is to be tagged, the annual cost for tags alone will then be millions of dollars. An alternative way of identifying individual logs is to use the already existing measurement data and make identification by means of the fingerprint approach (Chiorescu 2003).

The fingerprint approach rests on the foundation that each log that enters a sawmill has unique individual features. This can be the log's outer features, such as diameter, length, taper, crook and ovality, as well as inner features, such as knot volume, distance between knot whorls, heartwood/sapwood content and so on. If one could measure these individual features accurately enough, it would then be possible to separate individual logs in the same way that human beings can be separated by the use of their fingerprints. Hence measurement accuracy is the key to being able to uniquely define and recognize a log amongst others with the fingerprint approach (Chiorescu 2003). In the research, Chiorescu also identified bark as being a factor that has a negative effect on measurement accuracy and consequently the fingerprint recognition rate. The results showed that 3-D-scanner recognition rate dropped from 89 percent to 57 percent between log sorting station and saw intake when the log sorting measurements were made on non-debarked, rather than debarked, logs. Swedish log sorting stations use bark functions to compensate for the bark thickness. The log's on-bark diameter is used in a linear regression model to calculate double bark thickness, which is then subtracted from the on-bark diameter to get the under-bark diameter (Zacco 1974). This method for bark deduction is, however, more suited for pricing and scaling purposes on groups of logs than for defining bark deduction on an individual level. Both the variation in bark thickness and the amount of missing bark will lead to errors in the diameter compensation.

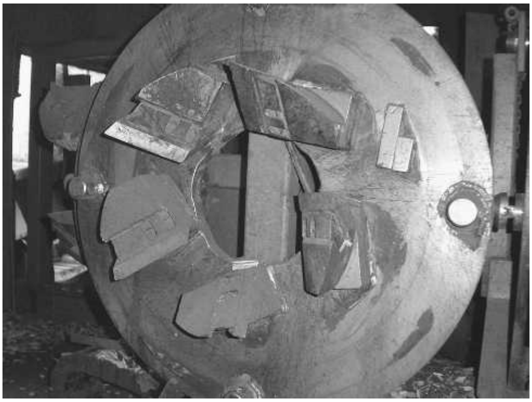

Figure 1. — Knifed ring for butt-end reduction.

A recent 3-D-scanner application that handles the bark issue on an individual level uses the tracheid effect to estimate bark thickness and missing bark. The tracheid effect is the physical phenomenon of laser light's ability to spread more along than across the wood fiber (Nyström 2002). The 3-D-scanner application for bark assessment uses the tracheid effect to determine whether the scanned surface is clearwood or bark. This is made possible by the fact that bark's ability to spread laser light is very poor compared to that of wood. By calculations based on the spread of the laser light, the application is able to virtually debark the measured log and make geometrical calculations "under bark" (Forslund 2000, Flodin 2007).

The purpose of this study is to investigate whether it is possible to increase the fingerprint recognition rate between log sorting station and saw intake by using traditional bark functions or the tracheid effect to compensate for the previously shown negative influence of bark.

Materials and methods

The sawmill that hosted the collection of data are located in the coastal part of northern Sweden. The sawmill is a midsize mill with an annual production of approximately 150,000 m^3 of sawn timber. The mill uses fixed sawing patterns that are applied to logs that have been sorted into diameter classes. The data collection involved three of the stations at the sawmill. These were, sequentially, a log sorting station with a 3-D scanner, a combined debarker/butt-end reducer and a saw intake with a 3-D scanner. The butt-end reduction was preformed by a knifed rotating ring with fixed diameter (see **Fig. 1**). The diameter of the ring used in this study was 320 mm.

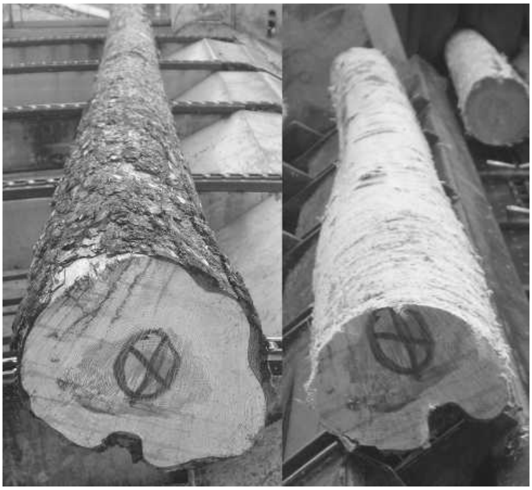

Figure 2. — The same log before and after butt-end reduction.

The study involved two groups with 50 randomly chosen Scots pine (*Pinus sylvestris*) logs in each group. The first group consisted of small-sized logs with a top diameter in the range of 148 to 154 mm. The second group consisted of larger sized logs with a top diameter in the range of 253 to 265 mm. Every log in each group was marked in both butt and top end with an ID number from 1 to 50. The second group was chosen because the diameter class in the 253 to 265-mm range is the largest diameter class to use the 320-mm butt-end reducer ring in normal production. Most of the logs in the larger sized group where therefore greatly affected and consequently shape-changed after they had gone through the butt-end reducer (see **Fig. 2**).

The measurement equipment used to generate the measurement data were two identical Sawco 3D scanners that were installed at both the log sorting station and the saw intake. The scanner has three measurement heads that use laser line triangulation to create cross sections of the log every 10 to 20 mm while it is fed through the scanner. The scanned cross sections are then stacked by the scanners software to recreate the log's outer shape (see **Fig. 3**).

The log sorting station scanner was equipped with Sawco's ProBark application that uses the tracheid effect to automatically assess whether the measured surface is bark or clearwood. **Figure 4** shows the difference in how the laser line is spread in bark and in clearwood. The recording of raw data can be done with the automatic bark assessment (ABA) active or inactive. If ABA is inactive, the raw data files are recorded with the bark present in all cross sections. If ABA is active, each cross section in the raw data files is compensated with the individual log's calculated bark thickness. With ABA active, one could think of the logs as "virtually debarked".

The log sorting station data were collected in the middle of October. The logs where then stored until the saw intake data were collected in late November. There was no influence from snow on the measurement results in either of the collections. In the October collection, no snow had yet fallen, and the snow present at the time of the November collection was removed in the debarking of the logs. The data collection with

Figure 3. — Stacked cross sections for recreation of a log's outer shape.

the log sorting station's 3-D scanner involved four runs for each group, three times with ABA active and once with ABA inactive. The data collection at the saw intake involved one run through the 3-D scanner for each group. No repeated runs of the logs were possible at the saw intake due to the machinery set-up. During each of the runs, the sequence of the logs ID numbers was noted so that each raw data file could be matched to the corresponding log.

Data analysis

MatLab® 7.0 (The MathWorks 2007) was used for calculation and analysis of the raw data files. The first step was to calculate geometry measurements that define a log. The components needed to calculate the log's geometrical measurements were found in the raw data files which hold the information about the stacked cross sections illustrated in **Figure 3.** The required components that were extracted from the files were length coordinates for the cross sections, spatial coordinates for the cross section's center of geometry, the area of the cross sections and the average diameter of the cross sections. The measurements that were calculated are shown in **Table 1.** When all 11 measurement values had been calculated, a measurement accuracy and repeatability analysis was conducted for both of the log groups using the three runs with ABA active from the log sorting station together with the run from the saw intake. The aim of the analysis was to establish and rank the reliability of the 11 variables. Each variable was evaluated by two calculated values.

The first value was for internal variation which describes the span of the measured values that the variables had assumed. Internal variation was calculated in two steps.

1. Calculate the standard deviations (SDs) for the value spans that the variables had assumed in the three log-sorting runs.
2. Set internal variation for each variable as the average value of the three SDs calculated in step one.

The second value was for the measurement difference which describes the accuracy in repeating the logs measurements between log sorting station and saw intake. Measurement difference was calculated in three steps.

1. Calculate absolute value difference between the three log sorting measurements and saw intake measurement.
2. Calculate the SDs of the absolute values obtained from step one.
3. Set measurement difference as the average value of the SDs calculated in step two.

The variables' reliability could then be determined as the internal variation divided by the measurement difference. The

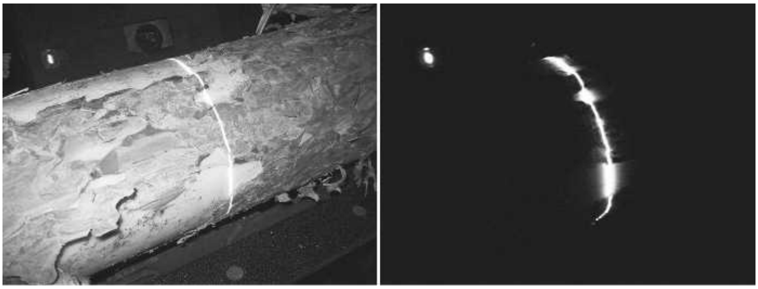

Figure 4. — Log photographed with flash (left) and without flash (right) (Flodin 2007).

Table 1. — Measurements calculated for each log.

Length (Len)	Distance from the log's top to butt end (mm)
Physical Volume (PhV)	The log's physical volume (dm^3)
Top Diameter (ToD)	The log's top-end diameter (mm)
Middle Diameter (MiD)	The log's middle diameter (mm)
Butt Diameter (BuD)	The log's butt-end diameter (mm)
Total Taper (ToT)	The absolute value of the change in diameter per meter of log length from the top end to a point measured one meter from the butt end (mm/m)
Top Taper (TpT)	The absolute value of the change in diameter from the top end of the log to a point measured one meter from the top end (mm/m)
Butt Taper (BuT)	The absolute value of the change in diameter from the butt end of the log to a point measured one meter from the butt end (mm/m)
Bow Height (BoH)	The maximum distance between the log's curvature and a straight line connecting the center of the log's end surfaces (mm)
Bow Radius (BoR)	The radius of a circle fitting the log's length and bow height (m)
Bow Position (BoP)	The distance from the log's top end to the point of maximum bow height (mm)

Table 2. — Variable reliability (a higher quotient value suggests that the related variable can be more reliably used in the multivariate matching procedure).

Small-size logs		Large-size logs	
Variable	Quotient	Variable	Quotient
Len	26.9	Len	24.9
PhV	5.4	PhV	4.4
BoH	4.1	ToD	2.2
ToT	3.3	MiD	2.1
BoR	2.8	ToT	1.7
BuT	2.6	TpT	1.4
MiD	2.6	BoR	1.2
ToD	2.3	BuD	1.2
BuD	2.2	BoH	1.2
TpT	1.7	BoP	1.1
BoP	1.6	BuT	1.0

higher the quotient value was, the more reliable the variable could be considered. **Table 2** shows the ranking of the variables according to this quotient value.

The fingerprint matching was done by means of multivariate principal components. This method had shown good results in previous studies (Chiorescu 2003). The method is a way to describe a dataset using underlying latent variables, i.e., principal components. The use of principal components is well suited to finding relationships between variables, reducing noise and allowing a dataset to be more simply described by fewer variables (Wold et al. 1987, Eriksson et al. 2001). If one imagines measured data as a point swarm in a multidimensional space where each point is an observation, principal components will align themselves orthogonally to account for as much of the variance in the point swarm as possible. The first component will account for the largest variation in the observations, the second component for the second largest variation and so on. Consequently, a linear combination of the first two or three components will not end up exactly at an original observation but will come close enough to make a good estimation. The original data matrix (X) can be projected onto the principal components to obtain the so called principal component scores (T). The score values for the observations are determined by the original data (X) and the principal component loadings (P). The relationship between original data and principal component data can be described as:

$$X = T \cdot P + E = \sum_{i=1}^{n} t_i \cdot p_i' + E$$

E is the residual matrix, i.e., the variation in the data that is not explained by the linear combination of principal components, and n is the number of components included in the model. The values in the residual matrix E will decrease for each added principal component and eventually reach zero when the number of components equals the number of original variables. Before principal components are calculated, the data are usually centered and scaled to unit variance in order to allow each variable to equally influence the observations projected score values. In this study, the variables were centered and scaled to have a mean value equal to zero and a SD equal to one.

If there is preexisting knowledge about the variables' reliability they can be further scaled after the initial unit variance scaling. One could scale up variables that are more reliable and likewise scale down variables that are less reliable in their influence on the results. If a variable is scaled up, the previously mentioned point swarm of observations will be stretched in the direction of that variable. The stretch will also affect the orientation of the principal components in the swarm, giving the stretched variable a higher loading value and subsequently more influence on the projections that gives the score values. In this study, the upscaling was done by multiplying each variable value with a corresponding scaling factor. The result will be that a unit variance variable multiplied with a scaling factor of, for example, two will get its SD changed from one into two.

The multivariate matching procedure was done with an algorithm described at the end of the data analysis section. The algorithm iterated a scaling vector containing the 11 scaling factors, i.e., SDs, to be used on the variables. The iterative

Table 3. — *Iterative range for the scaling factors used on the different variables.*

Variable	Scaling factor
Len	3 to 5
PhV	1 to 4
ToD	0 to 2
MiD	0 to 2
BuD	0 to 2
ToT	0 to 2
TpT	0 to 2
BuT	0 to 2
BoH	0 to 2
BoR	0 to 2
BoP	0 to 1

range for each scaling factor was based on the previously mentioned reliability test of the variables. **Table 3** shows the scaling factors that were used for all logs. To keep the number of iterations and computer calculation time at a reasonable level, each scaling factor in the scaling vector was given on average three different iteration values. For all the variables except length and physical volume, which had proven most reliable, the scaling factor range incorporated the value zero which when used excluded the affected variable from the matching procedure. This set-up gave the algorithm 157,464 different scaling vector combinations to work through. An additional iteration step containing three values of explained variance was also included in the matching algorithm. These values were 60, 70, and 80 percent. This step determined the number of principal components to be included in the matching procedure, i.e., the number of principal components needed to explain at least the percentage of variance in the original data given by the explained variance value. The additional iteration step for explained variance gave the algorithm a total number of 472,392 iteration steps to work through.

The actual matching for each iteration step was done using the log sorting station and saw intake score matrices (T) and then matching logs according to Euclidian distance. Euclidian distance (ED) is the shortest distance between two observations and is calculated as follows, where p and q are the score values for each principal component and n is the number of principal components:

$$ED = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2}$$

In this study, score values from the saw intake logs underwent Euclidian distance calculations one at a time with the score values of all the 50 logs from the log sorting station. Matching was then made to the nearest neighbor log with the shortest Euclidean distance within the 50 log sorting station logs. The final matching algorithm used in this study followed the sequence given below.

1. Read log sorting station and saw intake data of calculated measurements
2. Center and scale both datasets to unit variance
3. Multiply both datasets with the scaling vector
4. Calculate total number of principal components from the log sorting station data and the corresponding score and loading matrices
5. Reduce the number of principal components and also score/loading matrices to satisfy explained variance threshold value
6. Use reduced loading matrix to calculate score matrix for saw intake data
7. Calculate Euclidean distance between observations in score matrices and perform matching to nearest neighbor
8. Calculate and save the percentage of correct matching
9. Iterate explained variance threshold value and go to step 5
10. Iterate scaling vector and go to step 3

Table 4. — *Fingerprint recognition rate for small-size logs with different bark compensation methods.*

Bark compensation	Fingerprint recognition rate	Average recognition rate
	(percent)	
No compensation	77.6	77.6
Bark functions	83.7	83.7
ABA 1	85.7	
ABA 2	91.8	88.4
ABA 3	87.6	

Table 5. — *Fingerprint recognition rate for the large-size logs with different bark compensation methods.*

Bark compensation	Fingerprint recognition rate	Average recognition rate
	(percent)	
No compensation	54.0	54.0
Bark functions	62.0	62.0
ABA 1	60.0	
ABA 2	60.0	63.3
ABA 3	70.0	

Five matching runs were performed for each group between saw intake and log sorting station to observe how the bark and also the butt-end reduction for the large-size logs influenced the results. That meant one run with no compensation for bark, one run where bark had been compensated with bark functions and three runs where bark had been compensated with the ABA application.

After the first findings, two alternative approaches were tried on the ABA data to evaluate whether it was possible to improve the recognition rate for the large-size group. The first approach was to alter the raw data files from the log sorting station to mimic the effects of the butt end reducer. The second alternative tried was to virtually cross-cut the log's butt end and perform matching on the remaining part of the log. Six different length reductions were tried.

Results

During the handling between the log sorting station and the saw intake, one of the logs from the small-size group was accidentally broken in half and therefore left out of the data from the saw intake. **Tables 4, 5,** and **6** hold the best results of all the matching runs that were performed. The overall results from these tables show that the recognition rate can be improved with more sophisticated bark evaluation, but also that

Table 6. — Fingerprint recognition rate for large-size logs with different length reductions and different bark compensation methods.

Length reduction	No compensation	Bark functions	Average ABA
(mm)		(percent)	
250	54.0	58.0	63.3
500	56.0	62.0	75.3
750	66.0	70.0	76.7
1000	66.0	72.0	73.3
1250	70.0	68.0	71.3
1500	62.0	62.0	68.7

the bark issue is overshadowed by the effects of the butt-end reducer for the large-size group.

Noticeable is that there is a more than 20 percent lower recognition rate within the large-size group in **Table 5** compared to the small-size group in **Table 4**. This drop is most likely primarily caused by the shape change that occurred when the large-size logs passed through the butt-end reducer. Significant improvement in recognition rate for the large-size group could, as **Table 6** illustrates, be achieved by excluding a section of the log's butt end. The best recognition rate, 76.7 percent, was achieved using ABA and a length reduction of 750 mm. This, however, still falls behind the ABA recognition rate of 88.4 percent for the small-size group. The patterns in **Tables 4, 5,** and **6** are, on the whole, very similar. They show the lowest recognition rate for the runs where no compensation for bark was done. Compensation with traditional bark functions improves the recognition rate, which is even further improved when bark compensation is done using ABA. The attempt to alter the raw data files for the large-size group to mimic the shape change caused by the butt-end reducer did not give any significant increase in recognition rate.

Discussion

Based on these results, ABA appears to be a better alternative than traditional bark functions for handling the bark compensation on an individual level. This is also in line with what could be expected, since ABA was introduced in order to handle bark compensation on an individual level, as compared with bark functions that are sufficient for handling bark compensation on a group level but that lack precision to handle it on an individual level.

One reason why the large-size group, after length reduction, didn't reach the same recognition rate as the small-size group may be that the virtual cross cutting of the logs eliminates some of the natural variation that is used to define the logs' outer shapes. Another reason could be that the large-size group probably consisted only of butt logs, while the small-size group was composed of a random mix of butt, middle and top logs. This makes the total variation in shape more pronounced in the small-size group than in the large-size group.

The scaling factor range in the scaling vector was chosen with regard to the reliability of the variables and the number of iterations that the matching algorithm had to work through. Length was given the highest scaling factor range due to the fact that it was recognized to be far more reliable than the other variables. Bow height position, which had shown low reliability in both groups, was given a smaller range than the others. Physical volume was given the second highest values and a larger range, since it was the second most reliable variable in both groups, but at the same time sensitive to the effects of the butt-end reducer within the large-size group. Ideally, one would have wanted the study to have contained a larger amount of logs. This would have made it possible to set aside an independent test set of logs on which the matching model could have been validated. With a small amount of logs, there is a risk of over fitting the model so that it works very well on the training set, but poorly on a test set.

The number of principal components that were included in the matching procedures in any run didn't need to exceed three in order to satisfy the explained variance threshold values of 60, 70 and 80 percent. This is a good illustration of the advantage of more than 80 percent of the variance in the 11 original variables being explained by only three new latent variables, i.e., principal components. Three different threshold values were tried in order to see whether the explained variance played a large part in the recognition rate. The best results, in almost each run, came with the highest threshold value of 80 percent. It might have been interesting to try even higher threshold values, but with that comes the risk of rapidly increasing the amount of components and the modeling of noise.

The matching procedure for both the small- and large-size group was in this study done by calculating the Euclidean distance between each log measured at the saw intake to all the logs measured at the log sorting station. Consequently, a specific log could get multiple hits, i.e., be matched to several logs. An alternative approach would have been to remove matched logs one by one from the log sorting station data in order to eliminate the risk of multiple hits. However, on mismatch, this approach will automatically yield more matches that are incorrect, because the log removed from the log sorting data won't be available for a correct matching further down the line. By eliminating multiple hits, one will also eliminate the alert given that two or more logs have very similar shape. This alternative approach was therefore considered less suited for the matching procedure.

Another question raised during the study was whether each log measured at the saw intake should be matched within the group it belonged to or matched to both groups containing both small- and large-size logs. The approach chosen in this study was to match logs only within the same group. This approach would probably also be the best solution for a practical application. A sawmill of the size that hosted this study holds on average 70,000 to 80,000 logs in storage between the log sorting station and the saw intake. Each diameter class includes on average 3,000 to 4,000 logs when the class is run through the sawing procedure. If matching were to be done within a specific diameter class instead of the entire storage between log sorting station and saw intake, it would save a lot of calculation time needed to perform the matching. The drawback with this approach is that the matching becomes sensitive to mistakes; for example, if the logyard tractor by accident places timber from the wrong diameter class onto the sawing line.

All in all, the fingerprint approach offers a good potential to very cost effectively trace large amounts of logs between log sorting station and saw intake. The compromise is that the matching between logs is a probability match, rather than a secure match such as that obtained when using, for example, RFID. This indicates that the method could be well suited as a tool for process improvements where low-probability matches and multiple hits could be handled, but less suited as an origin traceability system, which requires a more secure match. An interesting idea for future work would be to investigate the extent to which "twin logs" that the matching algorithm confuses and mismatches actually differ from each other. If the purpose is process control, and these "twin logs" yield sawn timber of the same volume and quality, then a certain degree of confusion might be even more acceptable, considering the benefits that come with the fingerprint approach to tracing.

Conclusions

The fingerprint recognition rate can be improved by the use of more sophisticated bark compensation. Compared to no compensation, improvements can be made by using the traditional bark functions, and even further improvements can be made by using automatic bark assessment based on the tracheid effect. The butt-end reducer between the log sorting station and the saw intake has a very negative effect on the fingerprint recognition rate. Significant improvement in fingerprint recognition rate can be achieved by excluding the section of the log's butt end that is affected by the butt end reduction.

Literature cited

Chiorescu, S. 2003. The forestry-wood chain, simulation technique-measurement accuracy-traceability concept. Doctoral thesis 2003:03, Luleå Univ. of Tech., Skellefteå, Sweden.

Dykstra, D.P., G. Kuru, and R. Nussbaum. 2003. Tools and methodologies for independent verification and monitoring: Technologies for log tracking. Inter. Forestry Review 5(3):262-267.

Eriksson, L., E. Johansson, N. Kettaneh-Wold, and S. Wold. 2001. Multi- and Megavariate Data Analysis. Umetrics AB, Umeå, Sweden.

Finkenzeller, K. 2003. RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, 2nd ed. John Wiley and Sons Ltd., Chichester, West Sussex, England.

Flodin, J. 2007. Evaluation of transition from manual to automatic bark assessment using a 3D-scanner. Masters thesis 2007:011, Luleå Univ. of Tech., Skellefteå, Sweden. (In Swedish with English summary.)

Forslund, M. 2000. Evaluation of new technology aiming at increased accuracy when measuring the dimension of unbarked saw logs. Res. report no. P 0012041. SP Tech. Res. Inst. of Sweden, Wood Tech. (SP Trätek).

Kozak, R.A. and T.C. Maness. 2003. A system for continuous process improvement in wood products manufacturing. Holz als Roh- und Werkstoff 61:95-102.

The MathWorks. 2007. The MathWorks, Inc. Natick, Massachusetts. www.mathworks.com.

McFarlane, D. and Y. Sheffi. 2002. The impact of automatic identification on supply chain operations. Inter. J. of Logistics Management 14(1):1-17.

Nyström, J. 2002. Automatic measurement of compression wood and spiral grain for the prediction of distortion in sawn wood products. Doctoral thesis 2002:37. Luleå Univ. of Tech., Skellefteå, Sweden.

Töyrylä, I. 1999. Realizing the potential of traceability—A case study research on usage and impacts of product traceability. Doctoral thesis MA:97. Helsinki Univ. of Tech., Espoo, Finland.

Uusijärvi, R. 2000. Automatic tracking of wood-connecting properties from tree to wood product. Doctoral thesis R-99-43, Swedish Royal Univ., Stockholm. (In Swedish with English summary.)

__________. 2003. Linking raw material characteristics with industrial needs for environmentally sustainable and efficient transformation processes (LINESET). QLRT-1999-01476 Final Rept. Res. report no. P 0309034. SP Tech. Res. Inst. of Sweden, Wood Tech. (SP Trätek).

Wall, B. 1995. Materials traceability: The à la carte approach that avoids data indigestion. Ind. Manage. Data Syst. 95(1):10-11.

Wold, S., K. Esbensen, and P. Geladi. 1987. Principal component analysis. Chemometrics and Intelligent Lab. Systems 2:37-52.

Zacco, P. 1974. The bark thickness of saw logs. Rept. No. R 90. Swedish Royal College of Forestry, Dept. of Forest Products. 87 pp. (In Swedish with English summary).

II

Fingerprint traceability of sawn products using x-ray log scanning and sawn timber surface scanning

Jens Flodin[1, 2], Johan Oja[3], Anders Grönlund[4]

[1] SP Technical Research Institute of Sweden, Wood Technology, Skeria 2, 931 77 Skellefteå, Sweden, jens.flodin@sp.se

[2] Luleå University of technology, Department of wood technology, Skellefteå Campus, Skeria 3, 931 87 Skellefteå, Sweden

[3] SP Technical Research Institute of Sweden, Wood Technology, Skeria 2, 931 77 Skellefteå, Sweden, johan.oja@sp.se

[4] Luleå University of technology, Department of wood technology, Skellefteå Campus, Skeria 3, 931 87 Skellefteå, Sweden, anders.gronlund@ltu.se

Keywords: Fingerprint, Traceability, X-ray log scanner

ABSTRACT

Traceability in the sawmilling industry is a concept that could be used to more effectively control the production process and the utilization of the raw material. The fingerprint approach is a traceability concept that rests on the idea that every piece of wood is a unique individual with unique properties and hence can be identified and separated if a sufficient number of these properties are measured accurately enough. This study was hosted by a sawmill in northern Sweden and was aimed at making the fingerprint connection between logs and the center yield sawn from those logs using length and knot information. The 140 logs involved in the study were of Scots pine with top diameters spanning the range from 153 to 213 millimeters. The center yield sawn from these logs was of two dimensions. The smaller logs (153–187 mm) were sawn with a 2 ex pattern to 50 by 100 mm, and the larger logs (174–213 mm) were sawn to 50 by 125 mm with a 2 ex pattern. The data from the logs were collected at the log sorting station by an industrial one-directional x-ray log scanner in combination with a 3-D optical scanner. The data from the sawn center yield were collected by an industrial cross-fed surface scanning system situated in the sawmill's green sorting station. Both systems are used in the sawmill's normal continuous production. The results show that over 90% of all planks could be matched to the right log, which bespeaks a great potential for further development and realization of fingerprint tracing as a tool for process control and process improvement.

INTRODUCTION

Modern forestry and sawmilling companies often have sophisticated measurement equipment that generates large quantities of data at an individual level. These data are collected at certain points along the production chain, but are unfortunately almost exclusively used as a means to control the production process close to the measurement point. Most of the generated data for a specific piece of wood is therefore discarded as soon as the piece has moved past the measurement point. If the data for each specific piece were to be collected and stored in a database, the final product could then "be considered as an information intensive product" (Uusijärvi 2003). The challenge is therefore not to generate data, but to connect the generated data to each individual piece of wood. The reconnected data would make it possible to investigate and analyze both large and small sections of the production chain. A good example is the connection between the diameter classes for logs in the log sorting station and the volume recovery of sawn planks and boards. Without reconnection of data, one is reduced to comparing physical properties for a larger group of logs with the physical properties of their planks and boards. With traceability, i.e., reconnection of data, one is given the opportunity not only to analyze and find the individual logs in the group that yield high recovery, but perhaps even more importantly, to find the logs in the group that yield low recovery for a specific sawing pattern. Being able to make this distinction then makes it possible to adjust process parameters such as log class limits or sawing patterns for an overall higher recovery.

Since sawmills have a diverging flow, and modern sawmills have high production speed, the tracing and storing of data is not well suited for manual labor. A better alternative for handling the tracing and tracking is some form of automated identification (McFarlane and Sheffi 2002). One way of identifying individual pieces of wood is to use the already existing measurement data and make identification by means of the fingerprint approach (Chiorescu 2003). The fingerprint approach rests on the foundation that each piece of wood has unique individual features. These can be both outer and inner features. If these individual features could be measured accurately enough, it would then be possible to identify individual pieces in the production chain in the same way that human beings can be identified by the use of their fingerprints. The purpose of this study is to investigate whether the important individual connection between log and sawn center yield can be made by using the fingerprint approach based on length and x-ray information from the log sorting station and on length and surface scanning information from the green sorting station.

MATERIALS AND METHODS

The sawmill that hosted this study is a large-size mill situated in northern Sweden with an annual production of approximately 400,000 m^3 of sawn timber. The sawmill handles only Scots pine (*Pinus sylvestris*), which was also the only species included in this study. The 140 logs that were involved in the study were randomly chosen and had top diameters spanning the range from 153 to 213 millimeters. These logs were all sawn with a 2 ex pattern into two different center yield dimensions. The smaller logs (153–187 mm) were sawn to 50 by 100 mm, and the larger logs (174–213 mm) were sawn to 50 by 125 mm, making a total amount of 280 sawn center yield pieces. The sideboards produced were not included in the study.

The data used in this study were gathered at two points in the production chain from systems that are used in the sawmill's daily production. The first point was the sawmill's log sorting station where data from the logs were gathered with a one-directional x-ray log scanner from Rema Control AB in combination with a 3-D optical scanner from MPM Engineering Ltd. The data extracted from these systems were the log's total length according to the 3-D scanner and the position and length of the log's knot whorls according to the x-ray log scanner. The second point of data gathering was a Finscan Boardmaster surface scanning system situated at the sawmill's green sorting station. The total length and the positions of surface knots were recorded for each of the sawn planks. The order in which the logs and planks passed the measurement systems was written down manually from the end surfaces, which had been stamped with identification information (Skog and Oja 2007).

The analysis of the gathered data was performed using MatLab 7.3 (The MathWorks 2007). The first step in the analysis was to investigate the correlation between the total length measurements from the log sorting and green sorting stations. This was done by calculating the mean and standard deviation values for all the logs' lengths minus their corresponding planks' lengths. Once the length correlation was known, an algorithm was constructed to perform fingerprint matching between logs and planks. The algorithm was designed to work in a three-step sequence. The first and second steps in the sequence read the data into two matrices, first from the logs and then from the planks. Each row in the log matrix contained the identification, the total length and the starting position and length of all knot whorls found in the specific log. The information in the plank matrix was setup in the same way, with the difference that it contained the lengthwise starting point and length of all surface knots found on the planks. Due to edge effects in the filter, the x-ray log scanner needs a short distance before it starts registering information. Therefore, knots that were situated within 200 mm of the top and butt ends of the planks were disregarded.

The third and final step of the sequence was the actual matching procedure. The algorithm worked iteratively by taking one plank at a time and comparing its surface knot positions with the positions of knot whorls in all logs that had passed a length filtering. The length filter only allowed logs with a total length within a span based on the length correlation mean and standard deviation. The final matching was then made between the length-filtered log and the actual plank that showed the highest agreement in knot positioning. When all planks had been iterated, the total percentage of correct matches was calculated.

In order to find well-working settings for the algorithm, different values were tested for the log length filter as well as for the distance over which knots were disregarded at the plank ends. These values were between 3 and 10 cm for the length filter and between 100 and 400 mm for plank end disregarding.

RESULTS

Figure 1 and figure 2 show how the agreement in the planks' surface knots and the logs' knot whorls can be used to pair together a certain plank with a certain log.

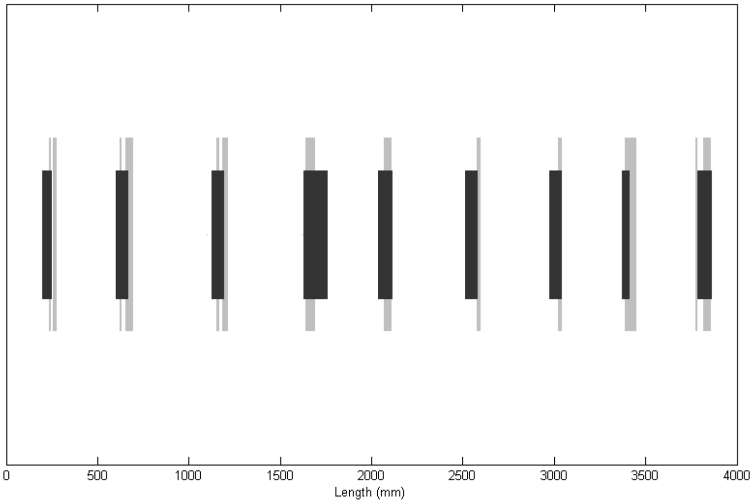

Figure 1. A correct match shows good agreement between the plank's surface knot positions (light gray) and the log's knot whorl positions (dark gray).

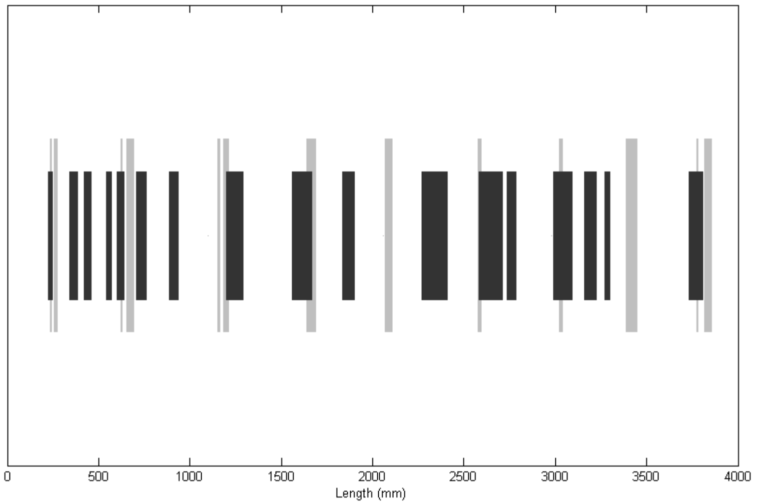

Figure 2. An incorrect match shows poor agreement between the plank's surface knot positions (light gray) and the log's knot whorl positions (dark gray).

The results from the fingerprint matching are positive, with a total correct matching percentage of approximately 90%. The length correlation between log sorting and green sorting station gave a mean value of -1.2 cm and a standard deviation of 1.6 cm. This means that the planks are in general measured slightly more than one centimeter longer than their corresponding logs and that the correct log, with very high certainty, is to be found within +/- 5 cm from the planks length minus 1.2 cm. The testing of different values for the log length filter and plank end disregarding didn't give any drastic results on the total percentage of correct matches. The percentage hovered slightly over and under 90% with the different values.

When basing the settings for the log length filter on the length correlation to +/- 5 cm (with mean correction) and using the original setting of 200 mm for plank end disregarding, a total of 92.5 % of the 280 planks were matched to the correct log.

DISCUSSION

With promising results like this, one can look forward to what might lie ahead for this method of tracing. One interesting spinoff is the ability to follow up if changes in process parameters such as, for example, log class limits have had the desired impact on the sawn product. Another idea is to use the fingerprint connected data to develop sorting models for the log sorting station, i.e., finding the outer and/or inner characteristics of the logs that yield a specific quality and/or volume recovery.

This study was conducted on Scots pine only. It's therefore hard to say how the fingerprint tracing approach would work on Norway spruce (*Picea abies*) which is the other main species of wood sawn in Sweden. Initially, it is thought that it will probably be more difficult, since Norway spruce doesn't have as clearly defined knot whorls as Scots pine because branches also grow in between the main knot whorls in the living tree. It would, however, be very interesting to investigate the possibilities of tracing Norway spruce with this method.

CONCLUSIONS

The results show a high potential for further development and realization of fingerprint tracing between log sorting and green sorting into a practical application for process control and process improvement.

ACKNOWLEDGEMENTS

The work was financially supported by the SkeWood program through TräCentrum Norr and through the Swedish Agency for Innovation Systems (VINNOVA).

REFERENCES

1. Chiorescu, S. (2003): *The forestry-wood chain, Simulation technique-Measurement accuracy-Traceability concept*. Doctoral thesis 2003:03. Luleå University of Technology, Skellefteå, Sweden.
2. The MathWorks (2007): The MathWorks, Inc., 3 Apple Hill Drive, Natick, MA 01760-2098, United States. <www.mathworks.com>.
3. McFarlane, D. and Sheffi, Y. (2002): The impact of automatic identification on supply chain operations. *International Journal of Logistics Management*. 14(1):1–17.
4. Skog, J. and Oja J. (2007): Improved log sorting combining X-ray and 3D scanning –a preliminary study. Proceedings of the COST E53 Conference on Quality Control for Wood and Wood Products, Warsaw, 15–17 Oct. 2007.
5. Uusijärvi, R. (2003): *Linking raw material characteristics with Industrial Needs for Environmentally Sustainable and Efficient Transformation processes (LINESET)*. QLRT-1999-01476 Final Report. SP Technical Research Institute of Sweden, Wood Technology, (SP Trätek). Research Report no. P 0309034.

III

Fingerprint traceability of sawn products using industrial measurement systems for x-ray log scanning and sawn timber surface scanning

Jens Flodin
Johan Oja
Anders Grönlund

Abstract

Traceability in the sawmilling industry is a concept that, for example, could be used to more effectively control the production process and the utilization of raw material. The fingerprint approach is a traceability concept that rests on the principle that every piece of wood is a unique individual with unique properties and therefore can be identified and separated if a sufficient number of these properties are measured accurately enough. This study was made with the aim of making the fingerprint connection between logs and the center yield sawn from those logs using length and knot information. The material used was Scots pine logs from six different diameter groups sawn with a two-ex sawing pattern into six different dimensions of center-yield planks. The data from the logs were collected at the log sorting station by an industrial one-directional x-ray log scanner in combination with a 3-D optical scanner. The data from the sawn center yield were collected by an industrial cross-fed surface scanning system situated in the sawmill's green sorting station. The results show that over 95 percent of all planks could be matched to the right log. This gives a high potential for further development and realization of fingerprint tracing between the log sorting and the green sorting station into a practical application for process control and process improvement.

Traceability can be defined in many different ways. Töyrylä (1999) defines traceability as follows: "Traceability is the ability to preserve and access the identity and attributes of a physical supply chain's objects." The ability to attach and access the history of a specific manufactured object brings an abundance of opportunities when it comes to controlling the quality of that object and the process that produced it. One example is the ability to ensure that harvested logs and their final products originate from a certified harvest site (Dykstra et al. 2003). Another good example is the possibility to investigate circumstances surrounding rework and customer return of faulty products. The ability to trace a product's history makes it possible to isolate and correct errors in the manufacturing process, hence preventing the same errors from occurring again (Wall 1995, Töyrylä 1999). For the same reason, many benefits may result from being able to trace products within the wood production industry (Kozak and Maness 2003).

An issue of growing interest for today's sawmills is the utilization of the raw material, i.e., producing the most suitable product from each specific log. If this can be achieved, there is a major benefit to be gained when the production of products that don't meet quality requirements can be reduced, along with the loss in revenue that these products bring. In order to obtain knowledge about the suitability between logs and sawn products, one needs individually associated data between the two. With individually associated data, it is subsequently possible to build log-sorting models in which the inner and outer characteristics of the logs can be connected to a specific quality and/or volume yield of the sawn product. Traditionally,

The authors are, respectively, PhD Student and Associate Professor, SP Technical Research Inst. of Sweden, Wood Technology, Skellefteå, Sweden (jens.flodin@sp.se; johan.oja@sp.se); and Professor, Luleå Univ. of Technology, Dept. of Wood Technology, Skellefteå Campus, Skellefteå, Sweden (anders.gronlund@ltu.se).
The work was financially supported by the SkeWood program through TräCentrum Norr and through the Swedish Agency for Innovation Systems (VINNOVA). This paper was received for publication in November 2007. Article No. 10430.

Forest Prod. J. 58(11):100–105.

these connected data have been the product of test sawings where logs have been manually marked and then tracked and recorded from the sawmill's log sorting through green sorting station. This is, however, a time- and money-consuming task, which suggests that an automated technique for achieving the individually associated data would be well appreciated.

Modern sawmills often have sophisticated measurement equipment that generates large quantities of data at an individual level. These data are collected at certain points along the production chain, but are unfortunately almost exclusively used as a means to control the production process close to the measurement point. Most of the generated data for a specific piece of wood are therefore discarded after the piece has moved past the measurement point. If the data for each specific piece were to be collected and stored in a database, the final product could then "be considered as an information intensive product" (Uusijärvi 2003). The challenge is therefore not to generate data, but to connect the generated data to each individual piece of wood.

Since sawmills have a diverging flow, and modern sawmills have high production speed, the tracing and storing of data are not well suited for manual labor. A better alternative for handling the tracing and tracking is some form of automated identification (McFarlane and Sheffi 2002). There are a number of alternative methods for practically making the connection between measurement data and the individual piece of wood. Many of these alternatives are based on some form of marking/reading technique. Two well-known methods are barcode identification and radio frequency identification (RFID). Barcode identification is a noncontact method used in almost every supermarket checkout counter in which the bars in the code are optically read by a laser scanner. RFID is also a noncontact method wherein an antenna picks up the RFID tag's unique identification number when it enters the antenna's reading range (Finkenzeller 2003). For forestry traceability applications, RFID is probably better suited due to the fact that the tags can be read without an optical scan, thus making the dirt and handling involved in logging almost noninfluential on the reading result, as opposed to reading barcode identification under the same circumstances. The drawback is the price for the RFID tags. A sawmill that produces 150,000 m^3 of sawn wood and has an average log volume of 0.18 m^3 handles approximately 1.8 million logs annually. The price for RFID tags is approximately 1 to 2 € (U.S. $0.75 to $1.50) per tag (Uusijärvi 2003). If every log is to be tagged, the annual cost for tags alone will then be millions of dollars.

An alternative and more cost-effective way of identifying individual pieces of wood is to use the already existing measurement data and make identification by means of the fingerprint approach (Chiorescu 2003, Flodin et al. 2007). The fingerprint approach rests on the principle that each piece of wood is a unique individual with unique features. These can be the piece's outer as well as inner features. If one could measure these individual features accurately enough, it would then be possible to identify individual pieces in the production chain in the same way that human beings can be identified by the use of their fingerprints. Microwaves have shown potential in fingerprint tracing of sawn wood (Fuentealba et al. 2004). This method might, however, be more suited for tracing wood that has been dried and kept in a constant climate rather that tracing through the sawmill process, since the wood's dielectric properties change when going from frozen to thawed and from green to dried conditions (Lundgren et al. 2005).

Table 1. — The Scots pine material used in the study.

Group	Logs Quantity	Logs Top diameter	Planks Quantity	Planks Thickness	Planks Width
		(mm)		-------(mm)------	
1	70	153 to 187	140	50	100
2	70	174 to 213	140	50	125
3	70	193 to 229	140	50	150
4	40	208 to 260	80	63	150
5	75	225 to 277	150	63	175
6	110	253 to 321	220	63	200

If one wants to make a fingerprint connection between logs at the log sorting station and sawn center yield products at the green sorting station, there are, among others, two properties that remain unchanged between the two locations if one applies a typical Scandinavian sawing pattern: the total length of the pieces and the lengthwise positioning of knots in the pieces. The purpose of this study is to investigate if the important individual connection between log and sawn product can be made by using the fingerprint approach along with length and x-ray information from the log sorting station combined with length and surface scanning information from the green sorting station.

Materials and methods

The sawmill that hosted this study was a large size mill situated in northern Sweden with an annual production of approximately 400,000 m^3 of sawn timber. The sawmill handles only Scots pine (*Pinus sylvestris*) which also was the only species included in this study. Scots pine is commonly sawn in Scandinavia and has well-defined knot whorls with no knots in between the main whorls. The logs involved in the study were randomly chosen from six different top diameter groups. All logs were sawn with a two-ex sawing pattern into center yield planks of six different dimensions. The sawing patterns referred to in this study are typical Scandinavian patterns applied on local raw material where the length of the sawn center yield planks is equal to the length of the log they are sawn from. A two-ex pattern means that each log is broken down into two center yield planks with surrounding sideboards. No sideboards were however included in the study. **Table 1** shows the data for the logs in the study.

The data that were used in this study were collected at two points in the production chain from systems that are used in the sawmill's daily production. The first point was the sawmill's log sorting station where data from the logs were collected with a one-directional x-ray log scanner from Rema Control AB (RemaControl 2007) in combination with a 3-D optical scanner from MPM Engineering Ltd. (MPM 2007). **Figure 1** shows the measurement equipment used in the study, and **Figure 2** shows an x-ray attenuation image of a Scots pine log. The data extracted from these systems were the total length of the logs according to the 3-D scanner and the position and length of the whorls in the logs according to the x-ray log scanner. The second point of data collection was a cross-fed Finscan Boardmaster surface-scanning system (Finscan 2007) situated at the sawmill's green sorting station. The total length and the positions of surface knots were recorded

Figure 1. — Industrial measurement equipment used to collect log data. 3-D optical scanner (left) and x-ray log scanner (right).

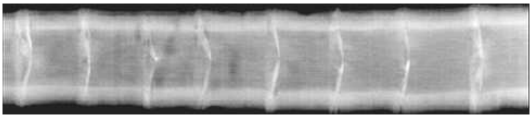

Figure 2. — X-ray attenuation image of a Scots pine log.

for each of the sawn planks. The order in which the logs and planks passed the measurement systems were written down manually from the end surfaces, which had been stamped with identification information (Skog and Oja 2007).

The analyses of the collected data were performed using MatLab 7.3 (The MathWorks Inc. 2007). Log groups 1 and 2 were used together for analysis and construction of the fingerprint-matching algorithm, while log groups 3, 4, 5, and 6 were used to verify the results. The first step in the analysis, before working on the matching algorithm, was to investigate the correlation between the total length measurements from the log sorting and green sorting stations. The unique identification allowed the sawn planks' length measurements to be individually compared with the length measurements from the logs. This was done by subtracting each plank's total length from the total length of its corresponding log. The mean and SD values were then calculated for the variation in difference between the two measurement points.

Once the length correlation was known, an algorithm was constructed to perform fingerprint matching between logs and planks. The algorithm was designed to work in a three-step sequence. The first and second steps in the sequence read the data into two matrices, first from the logs and then from the planks. Each row in the log matrix contained the identification, the total length (cm), and the starting position and length (mm) of all knot whorls found in that specific log measured from the top end, see **Figure 3**. The information in the plank matrix was set up in the same way, with the difference that it contained the lengthwise starting point and length (mm) of all surface knots found on all four sides of the planks measured from the top end, see **Figure 4**. Due to a filter in the x-ray scanner's software, the scanner needs a short distance before

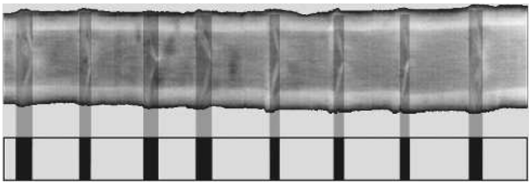

Figure 3. — Lengthwise positions of knot whorls in a log.

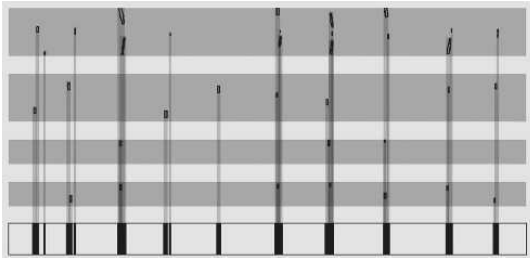

Figure 4. — Lengthwise positions of surface knots on a plank. The planks four faces are summarized (bottom).

it starts registering information. Therefore, knots that were situated within 200 mm of the top and butt ends of the planks were disregarded.

The third and final step of the sequence was the actual matching procedure. The algorithm worked iteratively by taking one plank at a time and comparing its surface knot positions with the positions of knot whorls for each log that had passed a length filtering. The length filter was based on the length correlation mean and SD and was used to screen through all the ingoing logs in order to exclude all logs that had a length that could not realistically belong to the actual plank being compared. The comparison between plank and log was made by creating two zero vectors, one for the plank and one for the log, with the same number of elements as the actual planks length in millimeters. These vectors were then filled with ones in the elements corresponding to positions of surface knots on planks and positions of knot whorls in logs. The MatLab autocorrelation function "xcorr" was used to calculate the correlation between the vectors, i.e., the correlation in knot positions between plank and log. The result from the function was normalized so that a total agreement would give a resulting value of 1.0, and a total disagreement would give a resulting value of zero. Matching between actual plank and the length-filtered logs was then done to the log that showed the highest normalized value. When all planks had been compared, the total number and percentage of correctly matched planks was calculated. In order to find well-functioning settings for the algorithm, different values were tested for the logs' length filter as well as for the distance over which knots were disregarded in the plank ends. The values tested were between 3 and 10 cm for the length filter and between 100 and 400 mm for the disregarding of end knots.

To increase the confidence in the knot agreement matching between actual plank and length filtered logs, the requirement for a certain minimum difference value between the highest normalized agreement value and second highest, was incorporated into the algorithm. If this minimum difference value

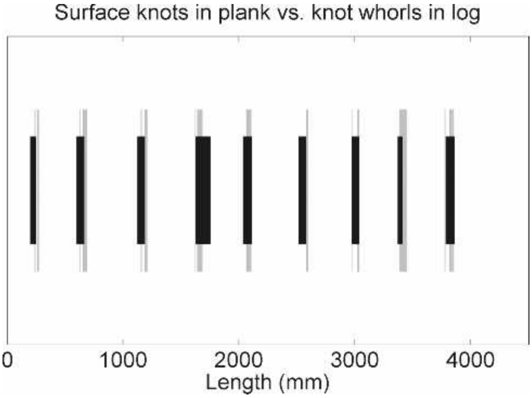

Figure 5. — A correct matching shows good agreement between the plank's surface knot positions (light gray) and the log's knot whorl positions (black).

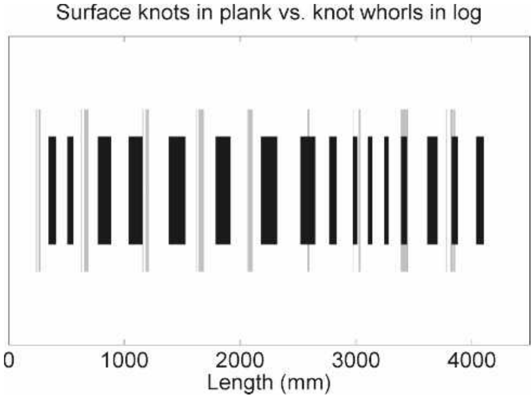

Figure 6. — An incorrect matching shows poor agreement between the plank's surface knot positions (light gray) and the log's knot whorl positions (black).

wasn't met, the actual plank was considered to be of too great a risk to be matched to the wrong log and subsequently not included in the final matching percentage. Different minimum difference values were tried, and the final matching percentage along with the number of planks left out was recorded. An analysis was also carried out on the left-out planks and their corresponding logs to investigate if they showed some sort of common characteristics, such as number or size of knots. The final stage of the study was to verify the results on groups 3, 4, 5, and 6. This verification would also show if the physical size of the logs and planks had an impact on the results.

Table 2. — Verification of results.

Thickness	Width	Number of ingoing planks	Number of correctly matched planks	Percentage of correctly matched planks
50	100/125	280	268	95.7
50	150	140	136	97.1
63	150	80	77	96.3
63	175	150	146	97.3
63	200	220	212	96.4

Results

Figures 5 and **6** show how the agreement in the planks' surface knots and the logs' knot whorls can be used to pair together a certain plank with a certain log. The results from the total length correlations between logs and planks gave a mean value of −1.2 cm and a SD of 1.6 cm, thus revealing that the planks are generally measured as a little longer than their corresponding logs. This result was used to initially set the length filter to ±5 cm of the actual plank's length (with mean correction). The length filter gave in itself two mismatches, due to a difference in measured length of more than 10 cm between log sorting and green sorting station.

The results of the first matching run were that 268 of the 280 planks could be matched to the correct log, yielding a success rate of 95.7 percent. After trying different values, the initial settings with length filter span at ±5 cm and end knot disregarding at 200 mm proved to be the best settings for the matching algorithm. Different settings showed no significant impact on the matching result. Similar results were found using planks from groups 3, 4, 5, and 6, as shown in **Table 2.** The confidence for all groups could also be increased by incorporating the previously mentioned minimum difference value between the first and second log with the highest normalized knot agreement. **Figure 7** shows that the percentage of correct matchings increases with increased minimum difference value, and **Figure 8** shows how the percentage of planks that were excluded from the matching procedure also increases when failing to fulfill the minimum difference value.

In order to find out if the excluded planks and their corresponding logs had any common characteristics, four histograms were plotted that compare mismatched and correctly

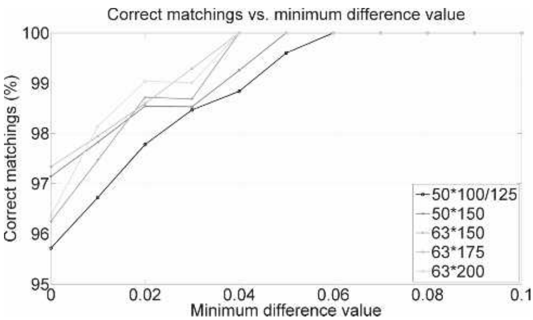

Figure 7. — Illustration of how more correct matchings can be achieved with the minimum difference value.

matched planks by 2 knot characteristics found in both the planks and the logs. The knot characteristics plotted were the amount and lengthwise size of surface knots for the planks and the amount and lengthwise size of knot whorls for the logs. **Figures 9** and **10** show the results for the planks, and **Figures 11** and **12** show the results for the logs. As **Figures 9** through **12** illustrate, no obvious grouping of the mismatched planks and logs could be found.

Discussion

The results from this study are very encouraging for further development of this fingerprint tracing method. The method can, as **Figures 7** and **8** show, be strengthened by applying a

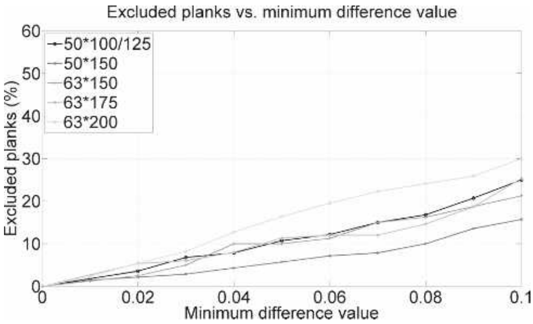

Figure 8. — Illustration of how the number of excluded planks increases with the minimum difference value.

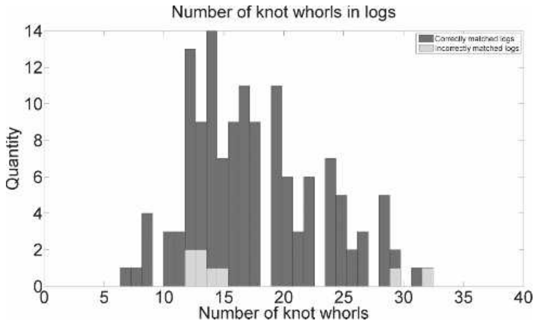

Figure 11. — The number of knot whorls in correctly and incorrectly matched logs.

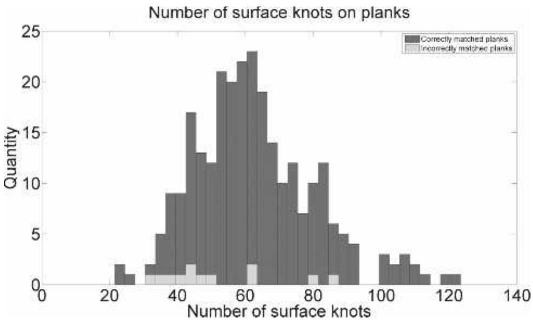

Figure 9. — The number of surface knots in correctly and incorrectly matched planks.

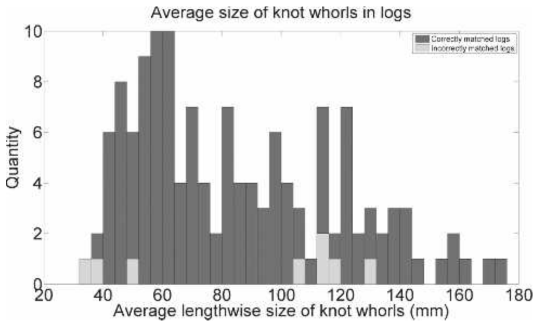

Figure 12. — The average lengthwise size of knot whorls in correctly and incorrectly matched logs.

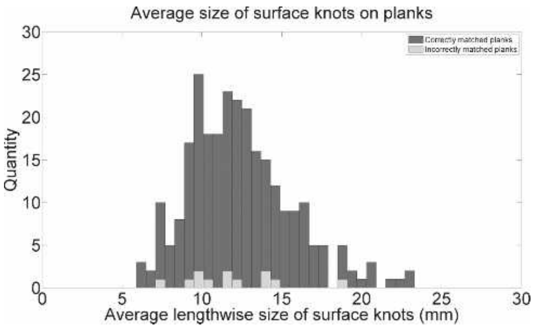

Figure 10. — The average lengthwise size of surface knots in correctly and incorrectly matched planks.

minimum difference value at the expense of throwing out some of the ingoing data. As **Figures 9** through **12** show, the exclusion of planks does not seem to take away any of the total variation in the ingoing material, which is very positive. With the results shown in **Table 2**, one might argue that the need for a minimum difference value is overkill if the object of the tracing is to develop statistical probability models for process control and process improvements (Maness 1993).

In this study, the occurrence of multiple hits, i.e., the same log being matched to more than two planks, was not given any special treatment. In order to further increase the confidence level of the matching, the algorithm could be extended to exclude logs that have received multiple hits. Another interesting approach in attempting to enhance the matching algorithm would be to start by matching together the planks that have been sawn from the same log and then use the combined knot information from these planks in order to find their corresponding log.

The results indicate that fingerprint tracing could be a very cost-effective way to collect and connect data, as opposed to the traditional test sawings which involves a lot of manual labor in the data collection. This connected data are essential for following up whether changes in process parameters such as, for example, log class limits, have had the desired effect. The individually associated data could also be used to form the foundation on which to build log-sorting models, since one gets the connection between the logs' inner and outer properties and the sawn planks' quality and volume yield. A large scale practical application would need to include a database and some form of breakpoints to indicate when batches are moved to different steps in the production. The breakpoints would make it possible to check off logs from the correct batch in the database when a suitable match is found in the batch of sawn planks from the green sorting station scanning.

This study was conducted on Scots pine only. It is therefore difficult to say how the fingerprint tracing approach would work on Norway spruce (*Picea abies*), which is the other main species of wood sawn in Sweden. The initial view is that it will probably be more difficult, since Norway spruce doesn't have as clearly defined knot whorls as Scots pine, due to the fact that branches also grow in between the main knot whorls in the living tree. The species are however normally kept separate at the sawmill and sawn one species at a time. Another interesting investigation would be to try fingerprint tracing on sideboards. Again, the initial view is that it will probably be more difficult, since the occurrence of surface knots decreases with increased distance from the logs center. The greater challenge would therefore be to find sideboards from large sawing patterns that have been applied on butt logs, but it would be very worthwhile to investigate the possibilities of tracing both Norway spruce and sideboards with this method.

The matching algorithm that was developed in this study relies on the logs and the sawn lumber to be of equal length. In order to handle sawn lumber that has been cross cut or taper sawn before surface scanning the present algorithm would need some further development.

Conclusions

The results show a high potential for further development and realization of fingerprint tracing between log sorting and green sorting station into a practical application for process control and process improvement. The results of the matching procedure can be strengthened and secured without systematically losing any of the natural variation in the ingoing material.

Literature cited

Chiorescu, S. 2003. The forestry-wood chain: Simulation technique, measurement accuracy, traceability concept. Doctoral thesis 2003:03, Luleå Univ. of Tech., Skellefteå, Sweden.

Dykstra, D.P., G. Kuru, and R. Nussbaum. 2003. Tools and methodologies for independent verification and monitoring, Technologies for log tracking. Inter. Forestry Review 5(3):262–267.

Finkenzeller, K. 2003. RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification, 2Nd ed. John Wiley and Sons Ltd., Chichester, West Sussex, England.

Finscan. 2007. Finscan Oy. www.finscan.fi. Finscan, Espoo, Finland.

Flodin, J., J. Oja, and A. Grönlund. 2007. Fingerprint traceability of logs using the outer shape and the tracheid effect. Forest Prod. J. 58(4): 21–27.

Fuentealba, C., C. Simon, D. Choffel, P. Charpentier, and D. Masson. 2004. The k-nearest neighbor method for automatic identification of wood products. *In*: Proc. of the 14th international conference on electronics, communications and computers (CONIELECOMP'04), 0-7695-2074-X/04.

Kozak, R.A. and T.C. Maness. 2003. A system for continuous process improvement in wood products manufacturing. Holz als Roh- und Werkstoff 61:95–102.

Lundgren, N., L. Hansson, O. Hagman, and A.L. Antti. 2005. FEM simulation of interactions between microwaves and wood during thawing. Presented at the 2nd Conf. on Mathematical Modelling of Wave Phenomena, Aug. 14–19, 2005, Växjö, Sweden.

The MathWorks Inc. 2007. MatLab 7.3. www.mathworks.com. The MathWorks Inc., Natick, Massachusetts.

Maness, T.C. 1993. Real time quality control system for automated lumbermills. Forest Prod. J. 43(7/8):17–22.

McFarlane, D. and Y. Sheffi. 2002. The impact of automatic identification on supply chain operations. Inter. J. of Logistics Management 14(1):1–17.

MPM Engineering Ltd. 2007. MPM. www.mpmeng.com. MPM Engineering Ltd., Surrey, British Columbia, Canada.

RemaControl. 2007. www.rema.se. RemaControl Sweden AB, Västerås, Sweden.

Skog, J. and J. Oja. 2007. Improved log sorting combining X-ray and 3D scanning. *In*: Proc. of the COST E53 Conf. on Quality Control for Wood and Wood Products, Warsaw, Oct. 15–17, 2007.

Töyrylä, I. 1999. Realizing the potential of traceability—A case study research on usage and impacts of product traceability. Doctoral thesis MA:97, Helsinki Univ. of Tech., Espoo, Finland.

Uusijärvi, R. 2003. Linking raw material characteristics with industrial needs for environmentally sustainable and efficient transformation processes (LINESET). QLRT-1999-01476 Final Rept. Res. Rept. No. P 0309034. SP Tech. Res. Inst. of Sweden, Wood Tech. (SP Trätek).

Wall, B. 1995. Materials traceability: The à la carte approach that avoids data indigestion. Ind. Manage. Data Syst. 95(1):10–11.

LULEÅ UNIVERSITY OF TECHNOLOGY

Universitetstryckeriet, Luleå