Veronika Hunková, Karel Janák[1]

Data Filtration Methods of Electronic Measurement of Log Dimensions

Metode filtriranja podataka elektroničkog mjerenja dimenzija trupaca

ABSTRACT • *The article deals with the processing of log dimension data collected by electronic measurement during reception at sawmills. The subject of the work concerns the filtration of data before their use for calculation. Filtration methods were designed based on simple mathematical and statistical methods, and compared and evaluated with the use of a designed comparative methodology. As a result, five filtration methods were selected that best suit the reception requirements. At the same time, the impact of filtration on the measurement results is evaluated in relation to the calculation method of wood volume. Furthermore, the calculation by sections is recommended, as it is less affected by filtration errors.*

Key words: log yards, round wood, electronic reception, electronic measurement, data filtration

SAŽETAK • *Članak se bavi obradom podataka dobivenih elektroničkim mjerenjem dimenzija trupaca pri ulasku u pilanu. Posebna se pozornost pridaje filtriranju podataka prije upotrebe za izračun. Metode filtriranja dizajnirane su na osnovi jednostavnih matematičkih i statističkih metoda te uspoređene i vrednovane primjenom dizajnirane komparativne metodologije. Rezultat rada je pet odabranih metoda filtriranja podataka koje najbolje odgovaraju zahtjevima preuzimanja trupaca. Također, procijenjen je utjecaj metode filtriranja podataka na metodu proračuna obujma trupaca. Osim toga, preporučena je metoda za izračun obujma trupaca na koju najmanje utječu pogreške pri filtriranju podataka.*

Ključne riječi: stovarište trupaca, oblo drvo, elektronički prijem, elektroničko mjerenje, filtriranje podataka

1 INTRODUCTION
1. UVOD

In 2009, electronic receptions of logs were performed in approx 45 plants in the Czech Republic (Solař and Janák 2011). Regarding the number of sawmills, the number of installed devices is relatively small. Despite this fact, 6.8 million m³, i.e. 75 % of all sawlogs, passed through them. The majority (70 %) is measured by 3D devices, which scan the whole peripheral curves of sawlogs. These devices are usually installed at large plants whose output exceeds 100 000 m³ of processed logs per year. 2D devices are typical for sawmills with the output of 30 000 m³ to 100 000 m³. Although these devices prevail in the Czech Republic (70 %), only 30 % of the volume of electronically received material passes through them. 1D scanning for the needs of the reception is rather rare (approx 1 % of electronically received material).

The experience of suppliers and entities processing and buying logs with the results of electronic reception are different. The suppliers are mostly complaining (not always) about the lower final material volume calculated by 3D measuring. In contrast, the

---

entities processing and buying logs are satisfied with this situation. According to the experience of both sides, the reception results of 2D measurements are more or less satisfactory. The experience is always based on manual measurements of the material in the forest. The accuracy of measuring diameter and length of sawlogs by all checked devices is high and balanced. According to the results of measurements performed by an accredited calibration laboratory, the mean deviation of sawlog diameter measurements ranges around ± 2 mm.

On the basis of the conducted research at the Faculty of Forestry and Wood Technology, Mendel University in Brno (Janák *et al*., 2005), it is possible to divide the factors affecting the results of electronic reception into the following areas:

- wood properties (shape, defects, mechanical damage),
- method of dimension scanning (mainly 3D and 2D),
- data processing method (evaluation of dimensions and volume of sawlogs).

Foreign works mainly deal with the accuracy of the measuring device itself, simplification of calculation operations, etc. The work *Automated Detection of Surface Defects on Barked Hardwood Logs and Stems Using 3-D Laser Scanned Data* (Thomas, 2006) can be mentioned as an example. Similarly, possible methods of classifying logs for specific purposes on the basis of surface features, detectable with the use of 3D scanning, are dealt with by Jäppinen (2000) in his dissertation. Many works concern improved accuracy of harvester measuring devices and their calibration, respectively. The processes of log dimension and quality evaluation with different degree of automation is dealt with by e.g. Marshall (2005). Among others, he describes the factors affecting the measurement quality. He considers the type of measuring device, speed of dimension scanning, degree of automation, tree species, and the method of sawlog volume calculation as the significant factors.

The research, partly addressed by this study, is focused on processing the collected data.

The method to determine sawlog dimensions and volumes is specified in regulations and standards, according to which the electronic reception is performed. The regulation defines directions and intervals of sawlog diameter scanning, the method of data conversion from mm into cm, determination of diameter at the measurement place, bark measurement, sawlog mid diameter calculation, and the method of volume calculation. The mentioned steps can be performed by different methods, with different accuracy, and in different order. The specific regulation used for a given reception is always agreed on between the supplier of logs and customer.

The determination of volume from the measured data is performed by two methods: calculation based on mid diameter (so-called Huber method), or calculation by sections into which the sawlog is divided (e.g. of the length of 10 cm). Previous research makes it clear that

the use of Huber method is subjected to measurement errors significantly more than the calculation by sections (Šmelko, 2003). This author claims that other errors occur due to the trunk shape and position of the sawlog in the stem. The volume determined by Huber method is lower for butt logs by approx 4 %, and by approx 5 % for top parts than the real volume. In the central part of sawlogs, the error does not exceed 1 %.

The experimental verification of the difference between the above mentioned methods of determining sawlog volume (Janák and Peter, 2004) confirmed that the total deviation of the method by sections from Huber method was 1.0 %, when calculated without rounding.

Measuring systems provide just a raw source signal with a number of errors occurring at the optical scanning of sawlog surface (shape anomalies of sawlogs, surface damage, torn grains, branch rests, bark, dirt, etc). Before using the collected data for the calculation specified in the regulations, it is necessary to eliminate the errors. This is the role of filtration, which then becomes a necessary step for the processing of the measured results. 2D scanning devices currently use the filtration based on simple methods, predominantly on moving average. Apart from the elimination of errors, the filtration may distort the data on shape and dimensions of the measured logs and thus affect the resultant volume. However, the filtration is not described in the regulations at all and neither is the method of its performance.

The aim of this study is to design different simple and complex filtration methods, evaluate their influence on the evaluated shape and volume of stem sawlogs, and recommend suitable filtration procedures for reception. A subsequent aim is to evaluate the effect of filtration methods on the evaluated volume of sawlogs.

2 MATERIAL AND METHODS
2. MATERIJAL I METODE

The designed filtration methods will be used for the correction of data collected within a common measurement for the needs of reception and processing of sawlogs. Therefore, they need to be based on the standard and currently used structure of scanned data.

Therefore, the data used for the research come from 2D measurement of sawlog diameter (Fig. 1). The measurements were performed vertically and horizontally. The values of pairs of mutually perpendicular sawlog diameter (hereinafter marked X and Y) are available in millimetres for every 10 cm of sawlog length.

The data were scanned by Inframat 760.2, made by Eltes Šumperk, and long-term stored on discs of computers in the plants of Pila Tetčice, a.s. and JILOS Horka, s.r.o. in 2008 and 2009. In total, data on 250 000 measured logs were stored.

For comparisons of simple filtration methods, data on 189 sawlogs were selected at random, and for comparisons of complex filtration methods, data on 901 sawlogs were selected.

In order to perform tests of filtration methods for end values, the pieces with defects close to the fronts
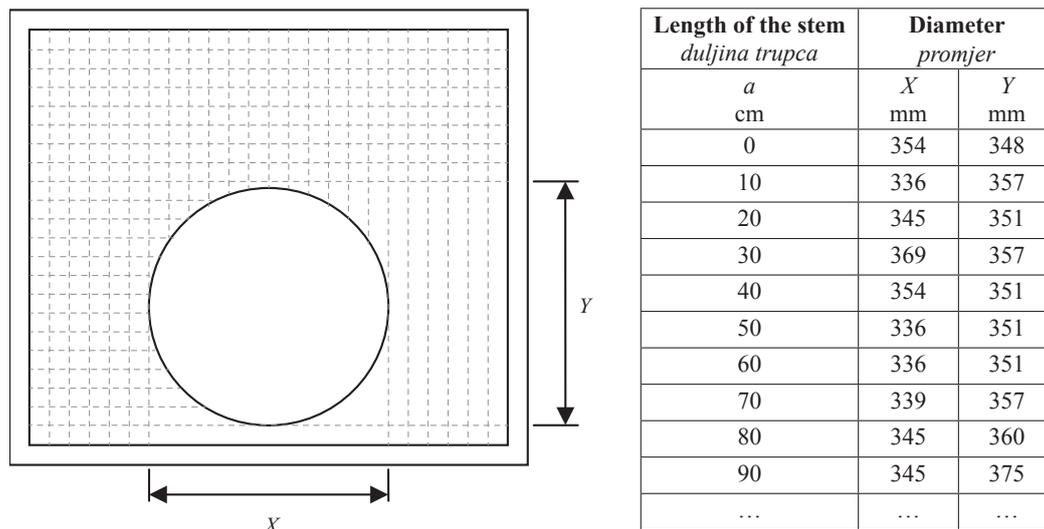
Length of the stem *duljina trupca*	Diameter *promjer*	
a cm	*X* mm	*Y* mm
0	354	348
10	336	357
20	345	351
30	369	357
40	354	351
50	336	351
60	336	351
70	339	357
80	345	360
90	345	375
…	…	…

Figure 1 2D measurement scheme and example of measured values
Slika 1. Shema 2D mjerenja i primjer izmjerenih vrijednosti

were selected. The selection was performed on the basis of a visual assessment of a graphic display of sawlog surface curve. The data set contained 27 sawlogs.

All filtration method tests were performed on sawlogs of 4 m length.

2.1 Used filtration methods
2.1. Primijenjene metode filtriranja

Filtration methods were designed in consecutive steps.

In the first phase, simple methods were designed. They were evaluated according to the effectiveness when removing values significantly differing from standard values. These methods were reliable regarding the sawlogs without shape anomalies. However, regarding more complicated shapes (buttresses, undercuts), they considerably distorted the total evaluated sawlog shape. Therefore, filtration methods focused only on these problematic sawlogs, and usually end parts of sawlogs were designed and evaluated. Based on the findings, more complex methods were designed that combine the advantages of individual simple methods and methods for end values. These methods were marked by letter *f* and a sequence number.

In total, 69 methods were designed for the filtration of measured data. They can be divided according to their purpose into:
- methods focused on filtration of values ± even course of sawlog shapes (mainly in the middle of the log),
- methods focused on identification and filtration of values of more complicated sawlog parts in terms of shape (mainly in their end parts),
- combined methods taking into account both types (both parts) of their courses.

Simple methods *(f1-f7)* use basic procedures for data processing and for the determination of general characteristics of individual files. They include moving averages and linear regressions. A method was derived for the filtration of data for electronic log reception. This method is based on the same principle as moving average, but it replaces the average with a median.

The moving average is calculated as an arithmetic average from a row of consecutive measurements and replaces its value with a value of this measurement which is in the middle of the row. When using the even number of values, it is necessary to further calculate the assigned value. Therefore, it is easier to use an odd number of values for the filtration purposes. The moving median works similarly, but assigns a median instead of arithmetic average. Thus, it is not affected by extreme values. The number of elements from which the moving average (median) is calculated is the smoothing width. Neither moving average nor median are able to calculate the values for data at the edges of data sets.

Simple filtration methods used in this study consist of moving average with the smoothing width of 3, 5 and 7 (*f1–f3*), moving median with the smoothing width of 3, 5 and 7 (*f4–f6*) and linear regression (*f7*).

Filtration methods for end values (*f8–f17*) focus on areas near sawlog ends. There are often outstanding values caused by buttresses, undercuts, or oblique cuts. During a common filtration process, these values may influence the course of the whole evaluated sawlog shape. An undercut or an oblique cut are defects whose presence reduces the usable length of a sawlog. After its identification, it is necessary to reduce (filter) the sawlog length by the length of such defect. A buttress is a stem butt swelling, which, after detection, only needs to be smoothed by filtration, not removed. Each of the methods designed for the filtration of end values thus consists of a part for the identification of both defect types and subsequently for their filtration.

In order to detect an undercut or an oblique cut, the methods are based on the comparisons of sawlog diameter values measured in both mutually perpendicular directions. The difference between these values is compared with the selected limit value (methods *f8, f14* – value 150; *f9, f15* – value 100) or with one fifth (*f10*), or one quarter (*f11*) of the median, respectively, of nine values corresponding with the position measured 130 cm from the end of the sawlog. The last two methods

compare the difference of a measured value on a given position and the median of all measured values in one direction with the value of 100 (*f16*) and one fifth of the median of all values (*f17*). In case the difference is greater than the compared value, the measured piece has most likely a large undercut or an oblique cut.

The second part of the methods for end values focuses on the detection of buttresses. In the Recommended Guidelines for measuring and classification of wood in the Czech Republic (composite authors 2007), an instruction is given to remove buttresses, so that their height over the round area does not exceed 3 cm. Based on this regulation, boundary values are determined as 73 (*f8*), 53 (*f9*), one sixth of the median at the position of 130 cm (*f10*, *f12*), one seventh of the same median (*f11*, *f13*), with which the difference between the measured value and the median on the position of 130 cm is compared.

Other methods determine the presence of the buttress with the use of a median from all values and the following values are considered the limit difference: 80 (*f14*), 60 (*f15*, *f16*) and one sixth of the median of all values (*f17*).

The above limit values (150, 100; 73, 80, 60) were derived from typical or permitted shape deviations of sawlogs and subsequently adjusted according to the filtration results.

Complex filtration methods (*f18 to f69*) consist of simple methods and end value methods for the evaluation of their advantages and drawbacks.

They can be divided into two groups. The first group includes the methods working on the principle of moving medians of different degrees of smoothing (3 to 7), which have the end values complemented with moved medians or with minimum values. The second group contains methods that compare the measured values with medians in order to detect the extreme. In case their difference or quotient exceeds the limit value (15, 20, 25 mm or 1.05 and 1.08), the initial measured value is replaced with a new one. In case it is not ex-ceeded, the initial value is still valid. New values are based on the calculation of a median of a different degree of smoothing or on the detection of the minimum.

2.2 Evaluation of filtration methods
2.2. Evaluacija metoda filtriranja podataka

2.2.1 Visual evaluation of filtration methods for end values
2.2.1. Vizualna evaluacija metoda filtriranja za vrijednosti na kraju trupca

When testing the filtration methods for end values, the ability of the methods to identify individual defects is determined first (undercut or oblique cut, buttress). The values collected at the place of these defects are thus not replaced with new values, but only with numbers (*U* for undercuts, *B* for buttresses, *I* invalid value). Therefore, the values corresponding with the standard sawlog shape stay unmarked. That is the reason why the identification results are very comprehensible, but they can be evaluated only visually. In the next step, the unsatisfactory values behind the beginning of undercut are replaced with zero (as if the undercut would not continue), and the values at the areas of buttresses are replaced with the values calculated from the median. This completes the filtration.

The first evaluating criterion is the detection and replacement of buttresses, respectively. In Fig. 2, the scanned values corresponding with buttresses are marked with the area *B*, the values after the filtration are below them on the curve *F*.

The second criterion is an invalid insertion of a zero value in the place where an extreme value was measured. Fig. 2 shows a replacement of the upper extreme (*UE*) with zero value during the filtration (marked by number *0*). The method erroneously identifies an extreme as an undercut and filters it out.

The third criterion is the failure to identify an undercut or an oblique cut. The values identifying this defect are shown in Fig. 2 in the area *U*. The defect is evaluated and filtered out correctly there.
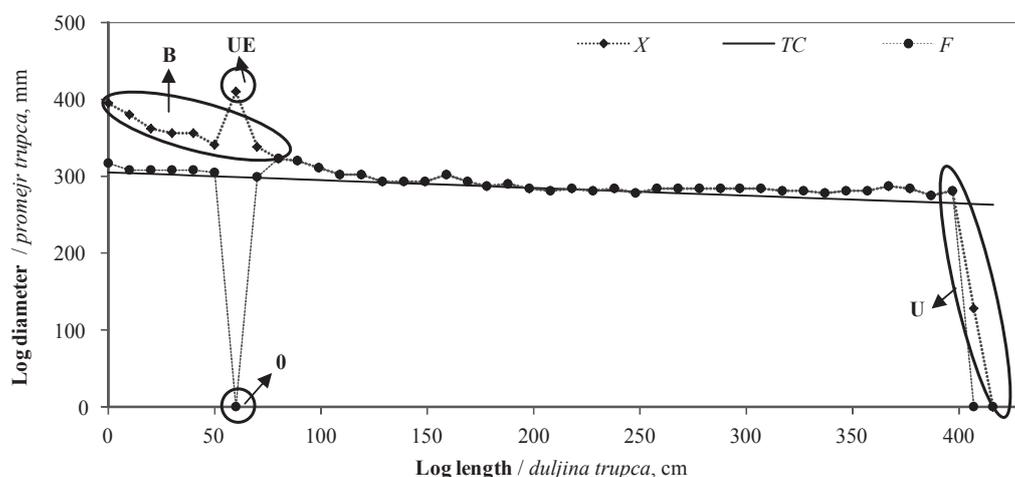

Figure 2 Graphic display of filtration methods for end values (*X* – measured values in direction *X*; *TC* – comparison curve; *F* – values determined by filtration; *B* – buttresses; 0 –invalid zero value; U – undercut or oblique cut; UE – upper extreme)
Slika 2. Grafički prikaz metoda filtriranja za vrijednosti na kraju trupca (*X* – izmjerena vrijednost u smjeru *X*; *TC* – krivulja za usporedbu; *F* – vrijednosti utvrđene filtriranjem; *B* – dno debla; 0 – pogrešna nula vrijednost; *U* – kosi rez; *UE* – gornji ekstrem)
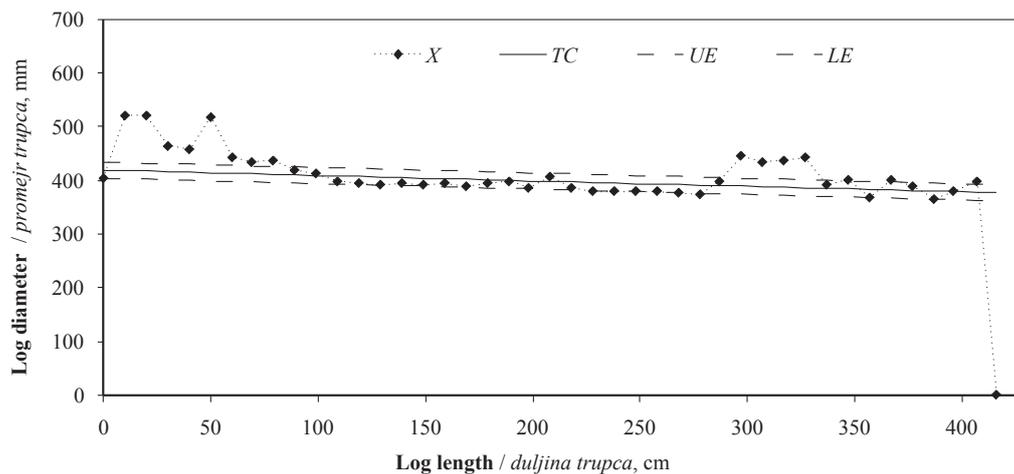
Figure 3 Graphic display of convergence line *TC* and limits for the upper (*UE*) and lower (*LE*) extremes. *X* curve represents the measured values in the horizontal direction
Slika 3. Grafički prikaz krivulje konvergencije *TC* i limita za gornje (*UE*) i donje ekstreme (*LE*); krivulja *X* predočuje vrijednosti izmjerene u horizontalnom smjeru

The errors are determined on the basis of a visual evaluation of curves and incorrectly identified values are summed.

2.2.2 Evaluation with the use of comparison curve
2.2.2. Procjena primjenom krivulje za usporedbu

In order to evaluate the filtration result without a visual check of every sawlog, a procedure is designed on the basis of a comparison of the measured values with the assumed values (a line corresponding with the normal rising gradient) and with the values calculated with the application of the filtration method.

The rising gradient line is calculated according to the rising gradient common for spruce in the Czech Republic, i.e. diameter increment of 1 cm per 1 cm of sawlog length. The line is interspersed with the median of all measured values in a given direction (*X* or *Y*). In the next step, a difference is calculated between the value determined by the given filtration method and the value according to the comparison line.

The values of the above mentioned differences of measured values, comparison values, and filtered values are compared, and it is evaluated whether:
- the value determined by filtration is identical to the measured value (not applicable to extreme values),
- the value determined by filtration stays in the field of extreme (measured value was marked as extreme),
- the measured extreme value above (below) the comparison line is replaced with a permissible value by the filtration,
- the measured value (despite not being extreme) is replaced with a new one – the value calculated by filtration,
- the calculated value, affected by an extreme, is at another position (an extreme is not removed by the filtration, it is just transferred).

With this method of evaluation of filtration procedures, the extreme values are those for which the filtration deviation of the calculated or initial value and comparison value is higher than 15 mm. An example of measured data in direction *X*, inserted comparison line *TC* and limits for marking values for the upper (*UE*) or lower extreme (*LE*) is shown in Fig. 3.

The classified values are summed in both the filtered data files and the initial measured data files. The difference is expressed in percentages. After the evaluation (in the next step), the numbers of extreme values are determined.

All criteria of effectiveness evaluation of filtration methods for individual sawlogs are gathered in data files that are statistically processed. The basic descriptive statistical indicators are calculated. The files do not have a normal classification, given by the method of effectiveness evaluation. The analysis of variance (ANOVA), which requires normality, cannot be used for the determination of the statistical significance of the difference between data files.

2.2.3 Evaluation of filtration selected methods by volume
2.2.3. Procjena odabranih metoda filtriranja primjenom obujma

Filtration methods are compared by their impact on the total sawlog volume. Two alternatives are selected for the calculation: from the sawlog mid diameter and by 10 cm long sections.

Diameter values (mid or of individual section) for the volume calculation are used in millimetres, without cm conversion, which excludes the impact of the conversion on the evaluation of filtration results.

2.2.4 Overall evaluation of filtration methods
2.2.4. Ukupna procjena metode filtriranja

Statistically processed evaluating criteria are ranked from the best to the worst regarding the required properties:
- upper extreme filtering rate (as many as possible should be filtered out),
- lower extreme filtering rate (as few as possible),
- rate of keeping initial values (as many as possible),
- rate of transfer of extremes (as few as possible),

- value of the sum of upper extremes (as low as possible),
- impact on the resultant sawlog volumes (as low as possible).

The methods are ranked with the use of robust characteristics of the descriptive statistics, i.e. median, modus, and the frequency of modus while taking into account outliers and extreme values.

According to the evaluated properties of the tested methods, methods for the use in practice are designed that best meet practical requirements under specific conditions. They are characterized by:
- prevalent dimensions or quality of measured logs,
- requirements for accuracy of the determination of individual log dimensions,
- speed of data processing on a given device,
- share of an operator in the course of reception,
- type of a scanning device and performance of the control system.

3 RESULTS AND DISCUSSION
3. REZULTATI I RASPRAVA

3.1 Simple filtration methods
3.1. Jednostavne metode filtriranja

Simple filtration methods were applied to values scanned from 189 sawlogs. The evaluation of their effectiveness was performed with the use of a comparison curve.

The results of comparisons of simple filtration methods are as follows:
- Regression analysis is unsuitable for data filtration, since it considerably simplifies the stem shape and is too affected by extreme values. The diversion of the regression line from the initial direction may cause a significant error in the estimate of end values, which are important for the classification of logs (an example shown in Fig. 4). The method can be used for finding the end values only if they are used for data which have already been smoothed by moving averages or medians, i.e. if extremes have been removed.

- Regarding moving averages, the wider the smoothing, the smoother the curve, but also the more distorted values by a single extreme value. An example of an affected result by a single extreme value is shown in Fig. 4. The transfer of an extreme to adjacent values is an important negative property of this method.
- The issue of end values also applies to median methods. It is necessary to complement them with another method (e.g. regression, use of minimum values or initial values, calculation of a median from another section). The advantage is that there is no impact on adjacent values by extremes (Fig. 4). Median methods reliably filter high solitary extremes, which may be represented by branch stubs and hanging bark. If more extreme values are present next to each other, which may be caused by e.g. larger tree burls, it is more probable that the information on such fault is not lost.

3.2 Filtration methods for end values
3.2. Metode filtriranja za vrijednosti na kraju trupca

Filtration methods for end values were tested on 27 samples, which had buttresses, undercuts, and oblique cuts. The evaluation is divided by evaluation criteria specified by the methodology.

3.2.1 Visual evaluation of filtration methods for end values
3.2.1. Vizualna procjena metoda filtriranja za vrijednosti na kraju trupca

a) Failure to detect buttresses

Regarding the filtering out of buttresses, the most effective seems to be the method *f11* (it compares the difference between sawlog diameter in both mutually perpendicular directions with one fourth of the median of nine values at the position of 130 cm) and *f13* (it compares the difference between the measured value and the median at the position of 130 cm with one seventh of the same median). When using these methods, the highest smoothing of the curve occurs in the area of buttresses with the lowest error rate (the failure to detect buttresses only occurred in six cases and it con-
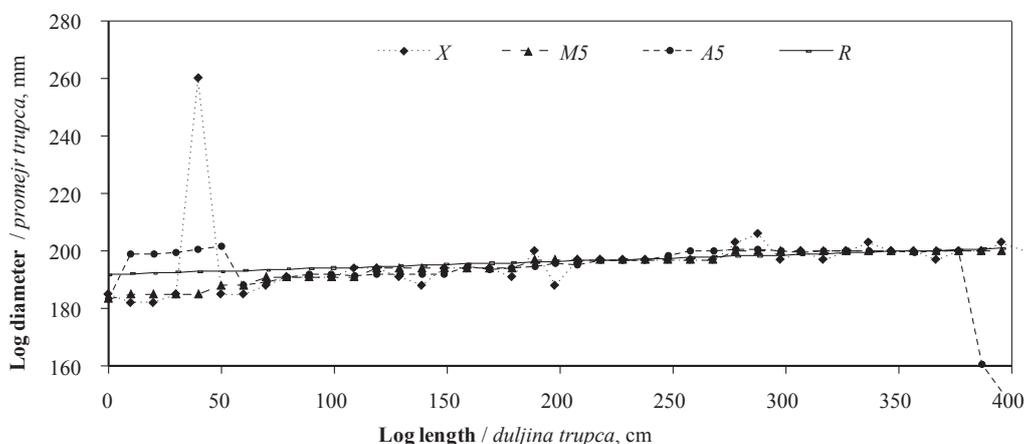

Figure 4 Impact on moving averages by a single extreme of sawlog No. 112 (*X* – measured values; *M5* – median out of five values; *A5* – moving average with the smoothing width 5; *R* – linear regression)
Slika 4. Utjecaj jednog ekstrema na trupcu broj 112 na pomične prosjeke (*X* – izmjerene vrijednosti; *M5* – medijan od ukupno pet vrijednosti, *A5* – pomični prosjek sa širinom poravnanja 5; *R* – linearna regresija)

cerned just one value, in one case two values). The methods *f8* and *f14* (they compare the difference of sawlog diameter in both mutually perpendicular directions with a fixed value) filter insufficiently, buttresses stayed undetected with 17 sawlogs. Most commonly, three values marked as buttresses stayed unfiltered, but in some cases there were five extremes.

The difference between the results of filtration of buttresses by individual methods is given by a determined comparison criterion. The methods *f8* and *f14* have the highest set fixed value for comparison. Another group consists of methods *f9*, *f15* and *f16*, where this value is lower. The third group consists of methods with a relative comparison criterion derived from the diameter of the measured sawlog (*f17*). Regarding sawlogs of smaller diameter, this approach reacts to small deviations. Regarding very thin pieces, high curve smoothing occurs. In contrast, even large deviations are tolerated in the cases of large sawlogs.

The search for extreme values with the help of a quotient of sawlog diameters brings a new calculation to the filtration process, which means higher demands on information technology and longer evaluation times.

b) Invalid insertion of zero value

Invalid insertion of zero value is the most important error in terms of the calculation of final sawlog volume. The visual checking makes it obvious that the error occurs due to an incorrect interpretation of a large difference between *X* and *Y* (sawlog diameter values scanned in both mutually perpendicular directions), such as an undercut or oblique cut. The lower the value with which the difference is compared, the more likely the difference between *X* and *Y* marked as an extreme.

The statistical evaluation of a visual check of the invalid insertion of zero value shows that the differences between the tested filtration methods are minimal. The occurrence of invalid values lies within the boundaries of statistical insignificance for all methods.

c) Undercut or oblique cut detection

The result of the evaluation of an undercut or oblique cut has only two alternatives: it was or it was not detected. The first four methods, *f8 – f11,* are very effective for the detection based on comparisons of the measured values in both directions (*X* and *Y*). The degree of effectiveness of other methods is about half of the above value. Despite this fact, the methods cannot be recommended due to their high error rate when inserting the zero value.

The mutual comparison of values *X* and *Y* can also be used for detecting an undercut or oblique cut and for the subsequent determination of sawlog length in compliance with Recommended Guidelines (composite authors, 2007). This operation can be performed only for end values as the first step before the filtration. When the difference between the end values *X* and *Y* is lower than a given criterion, the values can be considered as the new initial values, and with their use, it is possible to determine the sawlog length and the beginning of filtration.

3.2.2 Evaluation of filtration methods for end values according to a comparison line
3.2.2. Procjena metoda filtriranja za vrijednosti na kraju trupca primjenom usporedne krivulje

Similarly to the visual evaluation, when performing the evaluation according to a comparison line, it is impossible to claim with statistical significance that the methods differ from each other in a larger degree. However, it is obvious that some of them are stronger and some are weaker in terms of filtration.

The results of the evaluation of undetected buttresses are confirmed by the sum of the remaining upper extremes expressed as percentage. Regarding the filtration, buttresses are considered extreme values. Therefore, the sum of their deviations is relatively high. The statistical evaluation shows that filtration methods *f11* and *f13* are the strongest ones. Comparable results are reached by the method *f17*, which compares the difference of sawlog diameter values collected at a given place in both mutually perpendicular directions with one fifth of the median of all values. The highest sums of deviations were detected with methods *f8* and *f14*, comparing the same difference with a constant value of 150 (the highest use value).

When evaluating the sum of lower extremes, no difference was detected between the majority of methods, with the exception of methods *f8* and *f14,* which contained more errors. Nevertheless, the selected set of sawlogs contained very few extreme values under the comparison line (they occur exceptionally). Thus, the evaluation is based on a small number of observations.

A similar result can be seen when comparing the sums of deviations above and below the comparison line. The highest smoothing of the initial curve occurs with the use of filtration methods *f11* and *f13*, and the lowest with methods *f8* and *f14*.

Regarding the preservation of the initial values, which were not marked as extreme, it can be asserted that there is almost no difference between the filtration methods, and that the testing sawlogs mostly kept 78–79 % of the initial values. No transfer of extremes occurred in the used data file.

The evaluation of all criteria shows that the most acceptable method for the filtration of end values is the method *f13* and *f17,* respectively.

The method *f13* compares a measured value with the median of nine values at the position of 130 cm from the end of the sawlog. In case their difference is higher than a quarter of this median, the value is considered an undercut and replaced with zero. The calculation of an opposite difference with the aim to detect buttresses follows. The difference is compared with one seventh of the same median as in the previous step. In case the difference is higher, it is replaced with a value increased by 15 mm. The method is only applied to 10 end values.

The method *f17* works on a similar principle with the exception that a value of a median of all values in a given direction is used for the comparisons and calculations. The limit value for undercuts is one fifth of the median and one sixth for buttresses. A value marked as

a buttress is, in this case, replaced with a median increased by 25 mm.

Both these methods are relatively strong at the filtration of buttresses and extreme values, and do not show errors of invalid zero insertion, which is an important factor in practice. In order to search for extreme values, both methods use criteria based on a quotient from sawlog diameter. Therefore, the degree of curve smoothing is dependent on diameter. This is not in contrast with the real variability of log surface, which is lower with thinner pieces. In contrast, higher unevenness, occurring with thicker pieces, is shown in the sawlog value even after the filtration.

3.3 Complex filtration methods
3.3. Složene metode filtriranja

The evaluation of individual criteria leads to the following conclusions:

Simpler complex methods are very effective in removing the upper extremes, but rank among the weakest in other criteria. Their only advantage is their simplicity and speed of data processing.

Contrasting results are shown by the methods where the difference between the original measured value and the median of three or six elements is compared with the limit value of 25 mm. They are less effective in removing the upper extremes, but keep the initial values and do not transfer extremes. Thus, the lowest differences occur between sawlog volumes, calculated from the measured and filtered values.

Regarding the number of positively evaluated properties, the previous methods are ranked before the method *f55*. It is a double-step method. In the first step, it filters end values by the above mentioned method *f13*. Afterwards, all the calculated values are compared with the median of three values. In case the difference is higher than 25 mm, the value is replaced with this median. However, it keeps fewer initial values.

In the majority of criteria, even the methods which replace the detected extremes by the minimum are considered favourable. However, they are weaker in removing extremes. They also show an increase in values below the comparison curve. In addition, *f49* shows weaker filtration of extremes, even though other characteristics rank among the better evaluated ones. This method compares the ratio of the initial value and the median of three elements with the value of 1.08 (i.e. 8% tolerance). It is influenced by the measured log thickness.

Double-step methods, which provide more intense ("stronger") log curve smoothing, are more reliable for the filtration of the upper extremes. However, the consequence is a decrease in the resultant sawlog volumes. In addition, higher effectiveness methods lead to more complicated calculations.

3.4 Recommendations
3.4. Preporuke

Based on the analysis of properties, advantages and drawbacks of individual methods, the following methods are recommended for different conditions:

f21 – moving median of three values, for end values a median of the last two values is used

Advantages: simplicity, fast calculation, strong filtration of the upper extremes.

Drawbacks: removal of lower extremes, higher curve smoothing, lower percentage of preservation of initial values, transfer of extremes.

The method can be used for lower demands on accuracy, for measuring logs with bark (errors generated by the measurement with bark are usually higher than the errors caused by filtration). This corresponds with the practices of log conversion depots or small sawmills, where no high accuracy is required, as the classification is performed slowly and with cheaper machines, and an operator can significantly influence the process.

f32 – the measured value is compared with a moving median of three values; in case the difference is higher than 20 mm, the value is replaced with this median

Advantages: no complicated calculation algorithm, identical for all values, preservation of lower extremes, higher percentage of preservation of initial values.

Its drawback is a weaker filtration of the upper extremes.

This method is suitable for medium demands on speed, medium demands on accuracy, for logs of lower quality (keeping information on more complicated shapes of logs). This corresponds with the practices of small and meddle-size sawmills, where the classification is performed with cheaper machines, and where an operator can significantly influence the process.

f49 – quotient of the measured value and moving median of three values is compared with the value of 1.08; in case the quotient is higher, the value is replaced with this median

Advantages: the same (medium complex) calculation algorithm for all values, preservation of lower extremes, higher percentage of preserved initial values.

Drawbacks include a different degree of filtration for individual thickness classes, weaker filtration of the upper extremes.

The filtration method is suitable for medium demands on speed and accuracy, logs of lower quality (keeping information on more complicated shapes of logs), for logs with a wide range of thicknesses (more initial values are preserved with high thickness), for measurements with conveyors (effective filtration of extremes – driving dogs). This is suitable for small and middle-size operations, where medium accuracy is required, the classification is slower and performed with cheaper machines, and where an operator can significantly influence the process, logs of different thicknesses and quality are supplied, or a conveyor with driving dogs passes through a measurement frame.

f56 – the method works in two steps: coarse filtration is performed by the method *f13* (it compares the difference between the measured value and the median

at the position of 130 cm with one seventh of the same median) and these values are subsequently filtered with the use of a moving median of three values. The other step is a comparison of the values from the first and the second filtration phase – the limit value for the difference is 20 mm; in case the value is exceeded, a value from the median from the second filtration phase is inserted.

Advantages: preservation of the lower extremes, higher percentage of preservation of initial values, stronger filtration of the upper extremes.

Drawbacks: more complicated calculation algorithm, different algorithms for end values, higher curve smoothing.

The method is suitable for higher demands on accuracy, for logs of higher quality (extremes are often caused by surrounding effects, their filtration does not lead to the loss of information on the curve shape), for logs with buttresses, oblique cuts and undercuts (the filtration includes algorithms only applicable to end values). This corresponds with middle-size sawmills, where higher speed and accuracy is required, and where higher performance control systems are available.

f59 (the method works in two steps: coarse filtration is performed by the method *f13* and these "filtered" values are subsequently filtered with the use of a median of five values; then follows a comparison of values from the first and the second filtration phases – the limit value for the difference is 20 mm; in case the value is exceeded, a value from the median from the second filtration phase is inserted)

Advantages: preservation of the lower extremes, higher percentage of preservation of initial values, stronger filtration of the upper extremes, low error rate.

Drawbacks: complicated calculation algorithm, different algorithms for end values, higher curve smoothing.

The method is suitable for high demands on accuracy, for logs of higher quality (extremes are more often caused by surrounding effects, their filtration does not lead to the loss of information on the curve shape), for logs with buttresses, oblique cuts and undercuts (the filtration includes algorithms only applicable to end values). This corresponds with middle-size and large sawmill operations with higher quality equipment, where an operator rarely influences the measurement.

3.5 Impact of data filtration on electronic reception of logs
3.5. Utjecaj filtriranja podataka na elektronički prijem trupaca u pilani

Wood volume is essential when selling and buying wood with the electronic reception of logs. Huber method is usually used to determine the volume. When using Huber method, the volume of the whole sawlog is dependent on the mid diameter. With the electronic reception of logs the diameter is determined by four measurements near the sawlog centre and their average value is calculated.

The filtration should remove shape anomalies and prevent a distortion of the sawlog volume calculated by an extreme value. However, the filtration of an extreme value collected near the measured area in the sawlog centre may cause the replacement of a correct value for an incorrect one (extreme transfer). The result of filtration may be contradictory to the initial intention.

Simple filtration methods do not allow determining the limit criterion for an extreme, which is to be filtered or which is not be filtered. Complex methods allow taking this limit into account.

In order to evaluate the properties of individual filtration methods, the differences in volumes determined from two values near the sawlog centre from the initial and the filtered vales were also calculated. Regarding simple methods, the difference of volume from the initial and filtered values can be both positive and negative. The difference is positive only with complex methods. Positive differences indicate lower volume after the filtration, i.e. the diameter was reduced by the filtration. Therefore, the positive extreme was filtered out. The filtration of negative extremes is undesirable.

Although the percentage deviations of extremes may reach the limit values of over 30 %, majority of extremes ranges in the area of 10 %. It is necessary to point out that the deviations in volumes only occurred at approx 10 % of all sawlogs. The total deviation, which may occur during the measurement, would then range in the area of around 1 %.

Negative deviations were also indicated with simple methods, i.e. after the filtration, the volume is higher than before the filtration. This phenomenon is caused by an impact on initially normal value by a potentially extreme adjacent value. Subsequently, the diameter increase occurs.

Therefore, it is more suitable to apply more complex methods.

In order to evaluate the effect of filtration methods on the resultant sawlog volume, it is more suitable to determine sawlog volume by sections, since it takes into account the filtration of the whole curve. The mentioned effect of filtration is expressed as the volume of a cylinder of the diameter that is equal to the log mid diameter (Huber method). When determining the sawlog volume by sections, the effect is much lower, since it only has an impact on the section affected by an error, not the whole sawlog. Therefore, it is more accurate to use the method by sections, in order to determine the effect of filtration methods on the calculated sawlog volumes.

This study determined the log mid diameter as an average value of four measurements near the sawlog centre. The same method is used for the determination of a section diameter, when determining the volume by sections. The compared sawlog volumes were determined from diameter values in mm, i.e. without conversion into cm. The difference of the volumes of sawlogs determined by the calculation from the mid diameter and by sections was only set in this work for illustrative purposes and substantially simplified. The finding that the sawlog volume determined by sections is on average by 1.0 % higher than the volume determined from the mid diameter is in line with the previous works (Janák and Peter, 2004).

4 CONCLUSION
4. ZAKLJUČAK

Data filtration is a necessary step for processing data collected by the electronic measurement of sawlog dimensions. It is also a factor affecting the evaluated shape and volume of sawlogs. Therefore, the filtration method has an impact on the results of the whole electronic reception and, up to a certain extent, on supplier-customer relations of wood trading.

69 methods were designed and evaluated for the analysis and recommendations of suitable data filtration methods. The methods were applied to sets of values collected from sawlogs in sawmills (not laboratories). The methods differ from each other in the calculation method (moving average, median, smoothing width) and in the type of application (simple or combined, different processing of individual sawlog parts). The impact of individual filtration methods on the resulting sawlog values was evaluated visually as well as by a comparison procedure based on the taper.

The evaluation shows that simple methods deform mainly the parts of sawlogs with buttresses and are unreliable concerning oblique cuts and bounced cuts at sawlog ends. The combined methods are more reliable but more demanding in the performance of the control system. Five filtration methods are recommended, which differ in their accuracy and calculation demands. Therefore, they are suitable for different areas of use (volume and predominant dimensions or measured log quality, requirements as to the accuracy of sawlog dimensions, degree of participation of an operator during electronic reception, type of scanning device and equipment).

The impact of filtration on the value of the calculated volume during electronic reception is mainly given by the method and effectiveness of filtration; the method of log volume calculation also plays an important role.

When calculating sawlog volume by Huber method (from the mid diameter, typical for Central Europe), the total volume of a given piece is given by four diameter values scanned near the sawlog centre. An initially correct value may be affected by an adjacent extreme value and changed when applying simple filtration methods based on moving the most common averages or medians. An error, which thus occurs, has a significant impact on the volume value of the whole piece. Therefore, it is recommended to apply more complex methods (searching for extreme values while using set criteria).

When calculating sawlog volume by sections, only one section is affected by a given error. Therefore, the calculation of sawlog volume by sections significantly reduces the impact of filtration and scanning errors, and is hence more suitable.

The research was performed in cooperation with the representatives of suppliers and customers of saw timber. Our aim was to satisfy all requirements that were identified during the research. The research results shall lead to complementing and improving the regulations for electronic log reception in the Czech Republic, and, as we believe, to improving the log reception itself.

5 REFERENCES
5. LITERATURA

1. Janák, K.; Peter, B., 2004: Elektronická přejímka suroviny. Rozbor a kvantifikace vlivů různých způsobů stanovení objemu kulatiny na výslednou hodnotu. (*Electronic reception of wood. Analysis and quantification of influence of different ways of round wood volume calculation on the resulted value*) Faculty of Forestry and Wood Technology, Mendel University Brno.
2. Janák, K. *et al.*, 2005: Elektronická přejímka dříví: Přehled tuzemských metod měření dříví. Kvantifikace skutečných odchylek a rozbor jejich příčin. (*Electronic reception of wood. Overview of inland round wood measuring methods. Quantification of real deviations and analysis of their causes*) Faculty of Forestry and Wood Technology, Mendel University Brno.
3. Janák, K. *et al.*, 2006: Elektronická přejímka dříví: Rozbor zahraničních předpisů. Návrh průběhu elektronické přejímky dříví u nás. (*Electronic reception of wood. Analysis of abroad prescriptions. Proposition of electronic reception of round wood in our country*) Faculty of Forestry and Wood Technology, Mendel University Brno.
4. Jäppinen, A., 2000: Automatic Sorting of Sawlogs by Grade. Swedish University of Agricultural Sciencies, Uppsala. Available: http://pub.epsilon.slu.se/7/1/91-576-5873-0.fulltext.pdf
5. Marshall, H. D., 2005: An Investigation of Factors Affecting the Optimal Output Log Distribution from Mechanical Harvesting and Processing Systeme. Dissertation. Oregon State University. Available: http://hdl.handle.net/1957/9570
6. Solař, A.; Janák, K., 2011: Electronic Measurements and Timber Reception. In Woodworking Techniques 2011. Czech University of Life Sciences Prague, University of Zagreb. Praha, pp. 292-299. ISBN 978-80-213-2182-3.
7. Šmelko, Š. *et al.*, 2003: Meranie lesa a dreva. (*Measuring of forest and wood*) ÚVVP LVH SR, Zvolen. ISBN 80-89100-14-7.
8. Thomas, L., 2006: Automated Detection of Surface Defects on Barked Hardwood Logs and Stems Using 3-D Laser Scanned Data. Dissertation submitted to the Faculty of the Virginia Polytechnic Institute and State University Blacksburg, Virginia. Available: http://scholar.lib.vt.edu/theses/available/etd-09202006-145847/unrestricted/dissertation-lithomas-06.pdf
9. Žižka L., 2008: Vliv filtrace dat na výsledný obraz tvaru a rozměrů výřezu při elektronickém snímání. (*Influence of data filtration on the resulted picture of log dimension and shape at electronic scanning*) Diploma work. MZLU Brno.
10. ***2007: Doporučená pravidla pro měření a třídění dříví v České republice 2008. (*Recommended rules for measuring and sorting of wood in Czech Republic 2008*) Lesnická práce, Praha.

Corresponding address:

Ing. KAREL JANÁK, CSc.

Faculty of Forestry and Wood Technology
Mendel University
Zemědělská 3
613 00 Brno, CZECH REPUBLIC
e-mail: karel.janak@mendelu.cz