EFFECTS OF FILTRATION ON THE RESULTING IMAGE OF A LOG AT THE ELECTRONIC SENSING ITS DIMENSIONS

Veronika Hunková – Karel Janák

Abstract

The aim of the evaluation of methods of data filtration, which were taken at the electronic reception of logs, was to propose them and to entertain the rate of their effect on original data. Methods were described from the most simple up to methods, which were the combination of several procedures. Subsequently, they were applied to data sets with marked extreme values and the way of affecting the data series by filtration was visually evaluated. From the proposed methods of filtration, a method, which calculates medians from seven successive values, appears to be the most suitable method. However, it is necessary all methods to be tested on larger sets of data. If it is possible results of methods of filtration will be compared with actual values measured right on logs.

Key words: *data filtration, electronic measurement, log dimensions, running average, median, log diameter*

INTRODUCTION

At present increasing electronic reception of wood shows several drawbacks, which have not been fully worked out yet. Filtration of data taken by a measuring device ranks among them. The aim of filtration is to remove values of measurement, which do not correspond to the real shape of a stem and resulted on the basis of sensing defects on its surface. It is carried out by comparing values of neighbouring measurements according to a certain predetermined algorithm. Thus, searching for this algorithm is the paper objective.

MATERIAL

The proposed methods were tested on data obtained from a measuring device INFRAMAT of Eltes Šumperk Ltd. It refers to 2D measurement, the basis of which consists of double sensing frames working in two directions perpendicular at each other. For the filtration, data coming from a measuring unit to a control computer were used. Values of couples of log diameters perpendicular at each other are available (thereinafter termed X and Y diameters) in mm for every 10 cm length of a log. To elucidate effects of a filtration method data were used with marked problematic values, ie values considerably deviating from the anticipated course of the log surface.

METHODS OF FILTRATION

Methods of filtration are based on simple mathematical and statistical procedures, which are subsequently variously combined. It is necessary to deal with two basic problems, viz. first, how to find an incorrect value and second, which value should be used to replace the incorrect value.

Basic procedures

a) Mutual comparison of values of diameters measured in perpendicular directions in one place with respect to the log length. A difference in X and Y values is compared with the given value. This value has to be determined in such a way values caused by flattening the log not to be considered erroneous. It can be prevented by determining a higher value than standard flattening. In such a case, of course, only values considerably differing are termed erroneous. Thus, it is possible to filter off eg, values measured at outstanding bark, torn fibres or part of a branch. Moreover, it would be possible to eliminate effects of flattening by comparing several successive values between X and Y differences.

b) Mutual comparison of values of diameters measured consecutively in one direction. There, a difference between X_a and X_{a+1} is compared with a certain value. And again, it is necessary to determine the maximum possible size of a difference on the basis of experience and practical trials. Its minimum value can be based on a gradient. However, the method does not find a defect, which will deviate in several successive measurements. It can be modified by comparing not with neighbouring values but with farther values. There, however, it would be necessary to determine changes in comparison instead of with subsequent values to compare with previous values with respect to the end of a series.

c) Comparison of the coefficient of growth of successive values. The X_a/X_{a+1} ratio is termed coefficient of growth. And again, it is compared with a certain determined value and similarly as in the previous case successive measured values need not be members of the ratio. However, even here, it would be necessary to determine changes in comparison with respect to the end of a series.

d) Calculation of a value by means of gradient. The supposed value is determined by the conversion/subtraction of the value of increase to/from a smaller/greater neighbouring or farther value. Modification would consist in the value determination as a mean from the double calculation.

e) Fitting the surface curves by means of running averages. “Running average” is a statistical method used to evaluate time series, eg in meteorology or economy. To every moment on the time series (in case of filtration, to every place of measurement) a value is assigned calculated as a mean from a series of values, the centre of which is created by the original value. In using the even number of values it is necessary the assigned value to be further calculated. Thus, for the purpose of filtration, it is simpler to use the odd number of values. A figure giving the number of members serving for the running average calculation is termed width of smoothing. If there are special requirements for data smoothing it is possible to apply the smoothing algorithm several times. A figure giving the number of repetitions of the smoothing algorithm is termed depth of smoothing. A disadvantage of this method is, however, a fact that it is not possible to calculate values for edge data. If the objective of filtration consists in the determination of mid diameter of logs for the calculation of its volume, limit values are not important. However, they are necessary for the determination of a top diameter and in using running averages, it is necessary to approximate them by another method. One of possibilities consists in fitting the curve by calculated running averages. (Fig. 1)

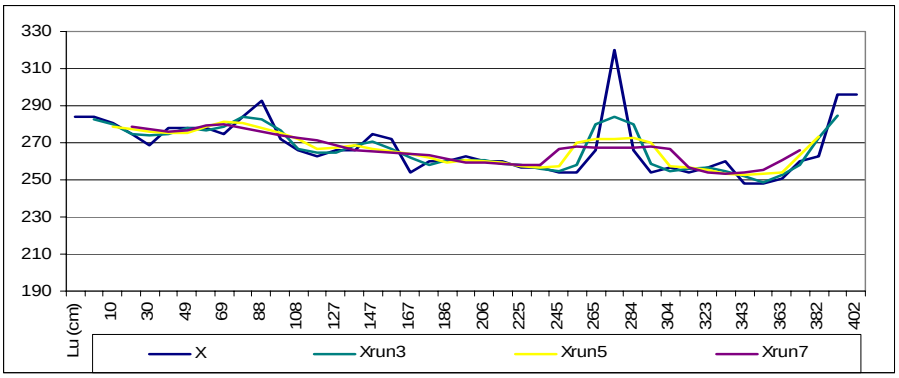

Fig. 1 Fitting the surface curve by means of running averages (width of smoothing 3, 5, 7)

f) Fitting the surface curve of a stem by means of "running" medians. The calculation procedure is the same as in the previous method and only the mean value is replaced by a mid value. (Fig. 2)

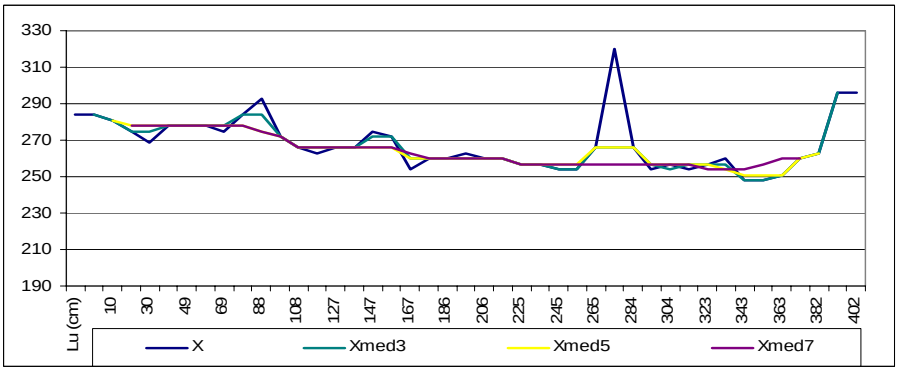

Fig. 2 Fitting the surface curve by means of "running" medians

g) Linear regression. Linear regression represents the approximation of given values by the first order polynomial (straight line) using the least squares method. This procedure can be applied right on measured values, but considerably farther values can markedly deflect the whole straight line. It is more suitable to use the procedure on the set of values, which was already modified by means of another method. (Fig. 3)

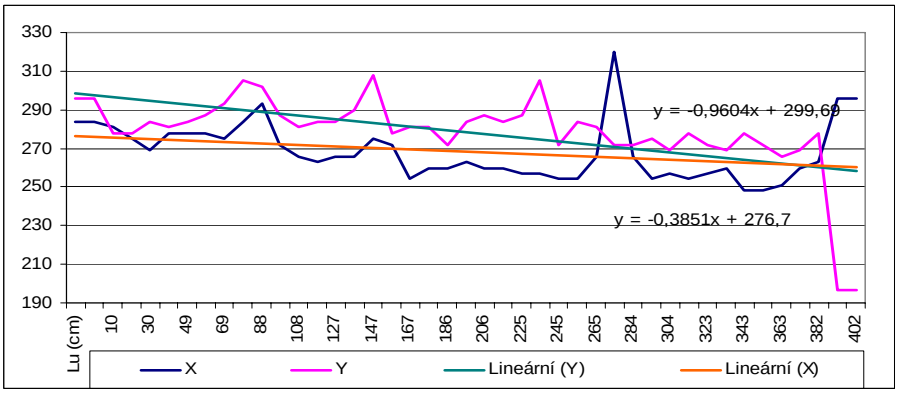

Fig. 3 Linear regression

It possible to carry out the inexhaustible amount of combinations on the basis of simple methods described above. It is given not only by their various sequences but also by possibilities, which are provided by each of the methods. Some of them will be described and completed by diagrams.

Combined procedures

h)

- Calculation of running averages (width of smoothing 5 or 7)
- Elimination of extreme values by means of comparing X and Y diameters and their replacement by values from running averages (the value of an acceptable maximum difference has to be determined)
- Smoothing the curve by running averages (width of smoothing 3, 5 or 7) by calculation from original values reduced by extremes

  Possible follow-up procedures
- Linear regression calculation – for the total fitting the curve or calculation of limit values
- Extrapolation of limit values

i)

- Calculation of running averages (width of smoothing 5 or 7)
- Comparison of running averages with the original value; too deviated values replaced by values of running averages (the value of an acceptable maximum of difference has to be determined)
- Smoothing the curve by running averages (width of smoothing 3, 5 or 7) (Fig. 4)

  Calculation of linear regression and extrapolation of extreme values can follow.

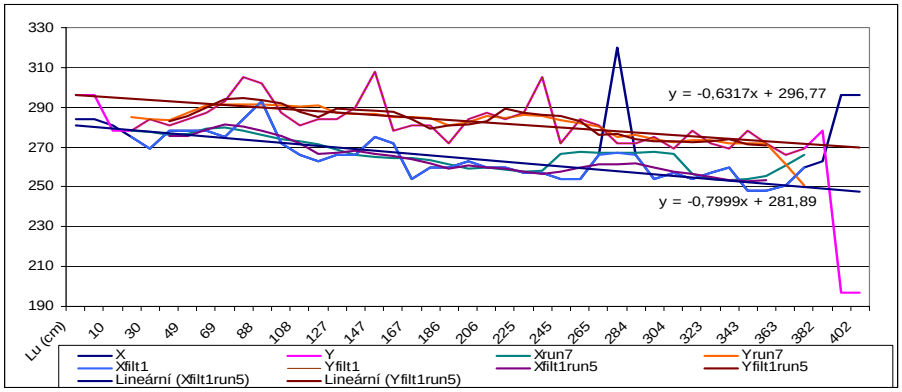

Fig. 4 Method i)

j)

- Calculation of running averages (width of smoothing 5 or 7)
- Comparison of the original value with a running average corresponding to a higher value: if the original value is higher than average it is replaced by corresponding running average
- Smoothing the curve by running averages (width of smoothing 3, 5 or 7)

  Calculation of linear regression and extrapolation of extreme values can follow.

Within the method h), it is possible to carry out the second step, ie elimination of extreme values by other simple methods, such as comparing values of diameters measured successively in one direction or comparing the coefficient of growth of successive values.

Many other combinations can occur replacing the average by median either for both calculations of running values or only for one of them. Further, to modify any of previous procedures by omitting the first and the last two or three values. The aim of the procedure is to remove effects of root swelling and obliquely cut butt ends on the resulting curve if linear regression is used. (Fig. 5)

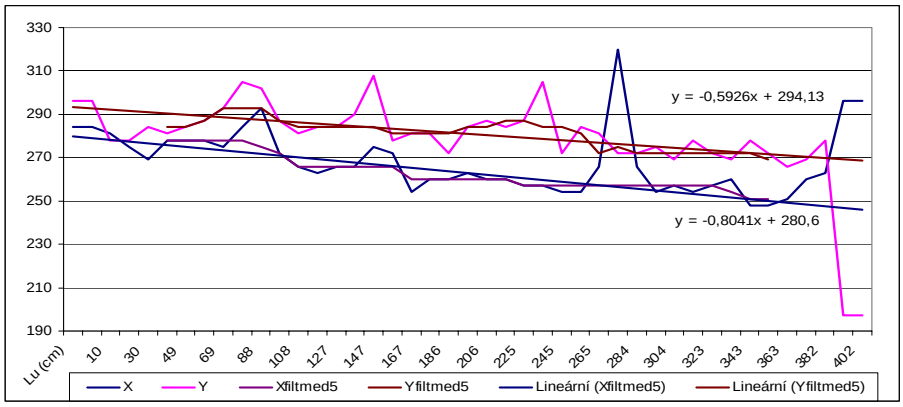

Fig. 5 Method i), the average replaced by median

All methods can be further simplified in such a way that first, the mean value is calculated from X and Y diameters in one place of measurement, eg according to Recommended rules 2008, and then particular operations of filtration are carried out.

DISCUSSION

At the experimental application of proposed methods it appeared that the simplest comparative methods did not fulfil our expectations.

Errors in their use only shift to another place. The calculated value is affected by an error and thus an inaccessible difference occurs at its comparison with the original value. Thus, comparison of original values with values calculated by means of running averages or medians appeared to be more suitable. It is valid for running averages that the larger the width of smoothing the smoother the curve. However, there are also more biased values due to one extreme value. Using the larger with of smoothing (7 values) for the calculation of comparative values at the first filtration appears to be useful. Thus, it is possible to filter off advantageously extreme values. In next step, we can use running averages from 15 values, where original values with extremes replaced by running averages from the first calculation serve as input values. Thanks to the smaller width of smoothing the character of surface will be preserved (eg knots or burls). Even better results were achieved using medians instead of running averages. Medians are not affected by extreme values. For the number of used values the same rule is valid as for averages.

As mentioned in the enumeration of methods linear regression is not suitable for application right on the set of original values. Extremes included there can deviate considerably the curve. Nevertheless, the method is very suitable if it is applied on data smoothed by means of running averages or medians. On the basis of a determined straight line it is possible to calculate missing limit values. At present, the Huber method is used for the log volume calculation most frequently. The method is based on the mid diameter of a log. In such a case, limit values are not necessary. However, if the calculation of volume is

carried out as the sum of volumes of particular sections of measurement this additional calculation is necessary. It finds its use also at grading logs according to butt end diameter.

This requirement is also a reason for the filtration of data of particular series of log diameters separately. A mean mid diameter necessary for the calculation of volume is determined from filter values of diameters.

For the present, using medians from seven successive values appears to be the simplest and at the same time sufficiently precise method. It not only sufficiently smoothes the curve but also removes effects of extreme risk values (eg measured root swelling).

CONCLUSION

From proposed and mainly visually assessed methods of data filtration taken at the electronic reception of logs a method, which calculates medians from seven successive values appears to be most suitable.

Nevertheless, it is necessary all methods to be tested on larger sets of data. If it is possible results of methods of filtration will be compared with actual values measured right on logs.

REFERENCES

[1] KOLEKTIV, Doporučená pravidla pro měření a třídění dříví v ČR 2008. Nakladatelství a vydavatelství Lesnická práce, s.r.o., Brno, 2008. ISBN: 978-80-87154-01-4
(Recommended rules for measuring and sorting the timber in the Czech Republic 2008)

[2] JANÁK, K., ONDRÁČEK, K. Elektronická přejímka dříví. MZLU v Brně, 2006. ISBN 80-7157-942-4
(Electronic Reception of Wood)

[3] ŽIŽKA, L. Vliv filtrace dat na výsledný obraz tvaru a rozměrů výřezu při elektronickém snímání. Diplomová práce. MZLU v Brně, 2008.
(Efects of filtration on the resulting image of a log at the electronic sensing its dimensions)

[4] Materiály fy Pila Tetčice, s.r.o. a fy Katres Stará Ves u Rýmařova, s.r.o.
(Data source)

[5] ÖNorm L 1021: Vermessung von Rundholz, BEV 2006, Rakousko
(Austrian standard L 1021 measuring of Round Wood)