LINEAR REGRESSION ANALYSIS OF INDICES DESCRIBING LATVIAN WOOD PROCESSING INDUSTRY

Janis Krumins, **Ingus Smits**, **Salvis Dagis**, **Dagnis Dubrovskis**, **Irina Arhipova**
Latvia University of Agriculture
janis.krumins@llu.lv

Abstract
To forecast how different indices describing Latvian forest industry will develop in the future, the decision support program which is based on the research of supply chains and production processes as well as on systematic analysis of the whole industry should be developed. One of the most common types of analysis is modelling of processes. In this paper processes of forestry and wood processing in Latvia are modelled by using general approach for system modelling, and each process is described as an abstract system where only its input and output values were identified. Correlations of input and output value time series with a simple and multiple regression analysis method were analyzed. Regression analysis was created for 77 pairs of indices which theoretically could have significant correlations. Input and output factor linear regression analysis for set of processes Forestry and Wood processing shows that not always there is significant correlation between time series of chosen model factors. Some of correlations between time series of input and output data in the model of forest sector even showed controversial results. The amount of information about processes and their quantitative indicators in Latvian wood processing industry is not sufficient for development of precise simulation models. There is need to develop a list of criteria for missing information and carry out its collection process.
Key words: forest sector modelling, Latvian forest industry, linear regression analysis.

Introduction

In Latvia, timber resources have a significant economical, social and ecological value as forests cover 54% of territory and timber stocks are slightly increasing year by year. During last ten years, annual cutting volume of timber resources has been 10 to 13 million m^3 (State Forest Service, 2012). Forest industry consists of two important sectors – forestry and wood processing industry both closely working together. Together with connected sectors (transport, building, energy industries, science, etc.), forest industry employs around 14% of able-bodied population, from which most are employed in rural areas of Latvia. Total export value of goods produced by wood processing industry in 2011 and 2012 reached 1.2 billion LVL (Ministry of Agriculture ..., 2012).

To forecast how different indices describing Latvian forest industry will develop in the future, a decision support tool should be developed. This tool should be in the form of IT program where economic modelling of the Latvian wood resource utilization and further processing could be done. This kind of program should be based on the research of supply chains and production processes as well as on systematic analysis of the whole industry (Toppinen and Kuuluvainen, 2010). One of most common types of analysis is modelling of processes. Its main task is to describe general production and supply processes in sufficient degree of detailed elaboration. To fulfil this task, it is possible to use different approaches (Becker and Kahn, 2003).

Since 2012 in Latvia as a decision support tool in strategic planning for the whole forestry sector programme 'MESTRA", jointly developed by Latvia University of Agriculture and Latvian State Forest Research Institute "Silava", is used. There is no decision support tool for strategic planning of wood processing sector though.

In this paper processes concerning forestry and wood processing in Latvia are described. They are being modelled by using general approach for system modelling and each process is described as an abstract system where only its input and output values are identified (Hangos and Cameron, 2001). To study most important processes and their relationships on a conceptual level, this kind of model is used in the research. For further research of identified relationships simple and multiple regression analysis of indicators acquired from different statistical sources was used. As wood processing companies are closely linked with forestry which is the main wood supplier to wood processing industry, then looking for indicators that influence Latvian wood processing industry, indicators of the whole forest sector were observed.

Our research had two main tasks. The first task was to identify indicators that influence Latvian wood processing industry and collect their time series. The second task was to analyze correlation of collected data from at least the last ten year period and evaluate which data could be used for development of programme prototype for economic modelling of Latvian wood resource utilization and further processing.

Materials and Methods

For analysis of indicators that influence Latvian forest processing industry only those indicators were selected which had information for at least last ten years. This minimal length of time series was necessary to get as feasible data as possible from simple and multiple linear regression analysis of the indices. Most of the collected data had time series for at least twelve years. Time series data were collected from different publicly available data bases - Central Statistical Bureau of Latvia (Central Statistical Bureau ..., 2012), Ministry of Agriculture and State Forest Service. Information was gathered in the following groups: the number of employees, area of forest land, stock volume of forest stand, area of felled forest land, volume of felled timber, area of regenerated forest land, net turnover, profit and profitability of forest sector enterprises, non-financial investments in forestry enterprises, value added, value added per employee, average gross salaries of employees, sales of manufactured products, import and export of forest industry, balance of foreign trade, prices of exported production.

Information about each group time series was collected in the most detailed level possible. For example, information about employment, net turnover, profit and profitability of forest sector enterprises, non-financial investments in forestry enterprises, value added, value added per employee and average gross salaries of employees were collected and analyzed in groups by the type of activity – forestry and logging (NACE 02), manufacture of timber products (NACE 16) and manufacture of furniture (NACE 31). In addition, information about sales of manufactured products, import and export volumes and value of forest industry was collected and analyzed by different type of production.

After all data were collected, a general model of forest industry for description of the processes and their relationships was created (Figure 1).

For construction of model the so called 'black box' principle was used which is common in modelling and analyzing different solutions of information

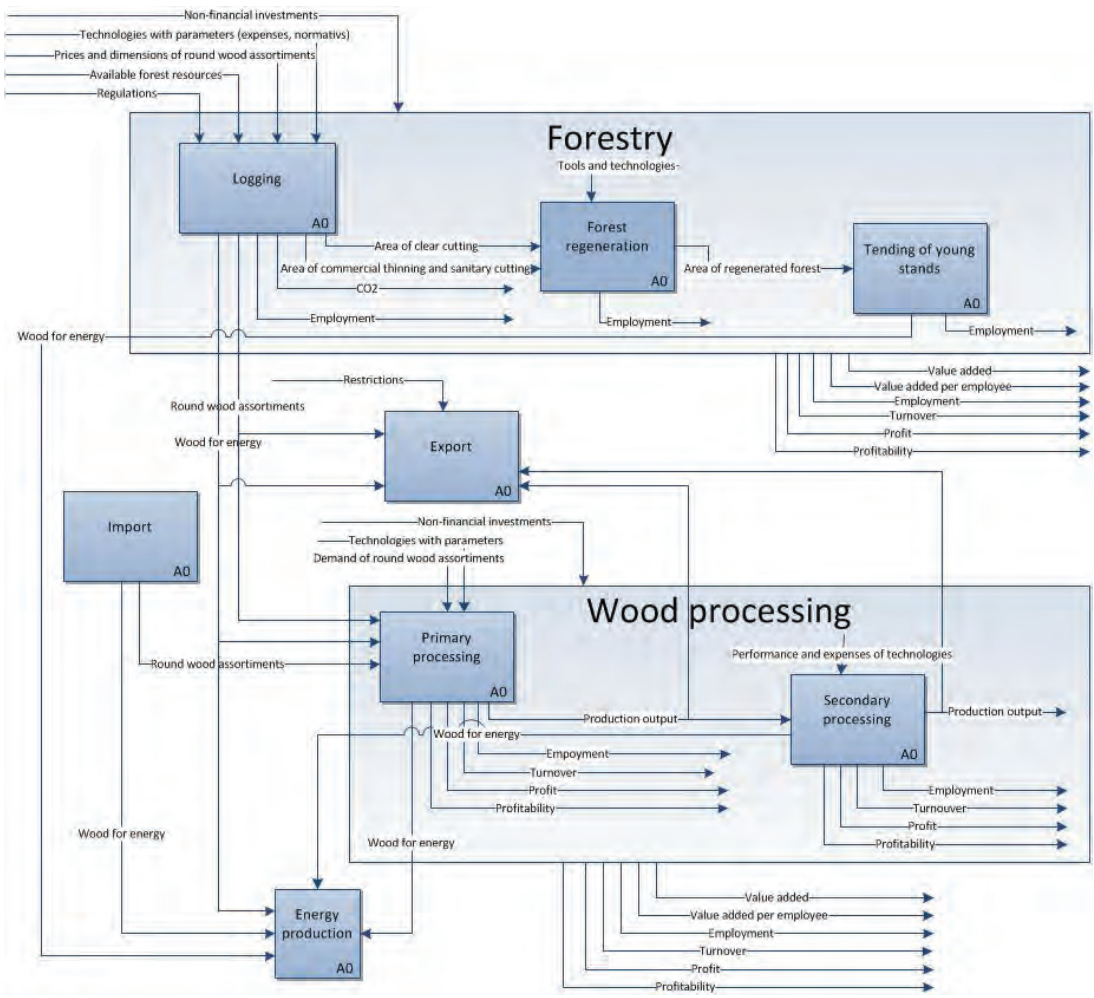

Figure 1. General model of forest industry for description of the processes and their relationships.

technologies. In this case the process is described without going in details about its structure and its ongoing activities. The main characteristic feature of the process is a set of information about input and output where output data of one process can be initial information to another process. In figure 1 eight main processes of the forest sector and their characterizing data can be seen. The processes which are included in the model are as follows: Logging, Forest regeneration, Tending of young stands, Imports, Exports, Primary wood processing, Secondary wood processing and Energy production. To simplify the initial version of the model according to the available data set, processes of wood extraction, forest regeneration, tending of young stands are viewed as a single process Forestry. Similarly also primary and secondary wood processing processes are analyzed as a single process Wood processing. The main reason for this kind of analysis is the limited set of statistical information that does not allow creating a more detailed analysis of all processes in the model.

Import and Logging can be considered as the beginning of the model (the first processes) where data of the process Import are used as information for further analysis, but relations inside the process at this stage of modelling are not covered. Consequently, as the main output information here is volume of round wood and energy wood.

Set of processes Forestry is viewed on a more detailed level. It has five general input data items, which during development process of the model are divided into more detailed information. Input and output data can be divided into several groups - economic indicators, data characterizing technological processes, regulatory information and data describing resources. In this case, the economic performance is characterized by non-financial investments in forestry enterprises and the prices of assortments from input side as well as by value added per employee, employment, turnover of forestry and logging companies, profit, profitability and value-added on output side. Information characterizing the technological processes from input side is an assortment of round wood dimensions and prices and description of technologies with parameters. The most important information that is necessary for the modelling and analysis of the process Forestry is descriptive data about the resources available, which in a given model diagram makes one common input ‘available forest resources’. Within the process Forestry resource transformation takes place, and results of transformations are characterized with output information - volume of round wood assortments and energy wood.

Next block in the model is Wood processing, and it describes processing of the received round wood assortments. This block consists of two processes - Primary wood processing and Secondary wood processing. Input data set for this block consists of a compilation of information that describes the available volumes of round wood assortments, technologies with parameters and non-financial investments. In output and internal circulation of information there are indicators that describe the results of each process, as well as the overall outcome. Main shortcoming for introduction of this block into real model is the large diversity of manufactured products as well as a relatively small set of statistical information, which is available for analysis of the internal processes.

Next step was to look for pairs of input and output time series for each of the processes out of pre-collected information of indicators that influence Latvian forest processing industry that theoretically could have significant correlations. For each pair of indices simple or multiple linear regression analysis was made. Regression analysis was performed by means of SPSS software where coefficient of determination (R^2) and input factor significance were calculated as well as model of regression was found.

After the regression analysis was performed, pairs of indices whose correspondence to real data was at least 30% ($R^2>0.3$) were selected for detailed analysis and evaluation if they could be used in programme prototype for economic modelling of Latvian wood resource utilization and further processing. In this paper a detailed description of relationship between input and output factors of the model processes was performed for those pairs of indices whose R^2 was at least 0.3 in set of processes Forestry and at least 0.7 in set of processes Wood processing.

Results and Discussion

Regression analysis was performed for 77 pairs of indices which theoretically could have significant relations. However, in most of the cases correlation was unsatisfactory.

Input and output factor linear regression analysis for the set of processes Forestry shows that not always there is a significant correlation between time series of chosen model factors (Table 1).

Regression analysis shows that profit of forestry and logging enterprises is significantly dependent ($p<0.05$) on the amount of non-financial investments in enterprises of forestry and logging. Profit of forestry and logging enterprises increases for 931 LVL if amount of non-financial investments in forestry and logging enterprises increases by one thousand LVL and model correspondence to real data is 50.5% because coefficient of determination R^2 is 0.505.

Turnover of forestry and logging enterprises is significantly dependent ($p<0.01$) on the amount of non-financial investments in enterprises of forestry

Table 1

Correlation between input and output factors in set of processes Forestry

No	Input factor X_i	Output factor Y	R^2	Regression equation
1.	X_1: Amount of non-financial investments in enterprises of forestry and logging, thousand LVL*	Y: Profit of forestry and logging enterprises, million LVL	0.505	$Y= -2.009 + 0.000931 * X_1 + e_i$
2.	X_1: Amount of non-financial investments in enterprises of forestry and logging, thousand LVL**	Y: Turnover of forestry and logging enterprises, million LVL	0.617	$Y= 107.14 + 0.003937 * X_1 + e_i$
3.	X_1: Amount of non-financial investments in enterprises of forestry and logging, thousand LVL**	Y: Profitability of forestry and logging enterprises, %	0.602	$Y= -69.97 + 0.01358 * X_1 + e_i$
4.	X1: Area of final felling, ha**	Y: Area of regenerated forest land, thousand ha	0.421	$Y= 2.094 + 0.0006 * X_1 + e_i$
5.	X1: Area of reconstructive felling, ha***	Y: Employment in forestry and logging, thousands of employees	0.639	$Y=14.40358+ 0.01327 * X_1 + e_i$
6.	X1: Area of commercial thinning, ha*	Y: Employment in forestry and logging, thousands of employees	0.320	$Y=11.624+ 0.000185* X_1 + e_i$

* - factor is significant with probability P=95% ($p<0.05$)
** - factor is significant with probability P=99% ($p<0.01$)
*** - factor is significant with probability P=99.9% ($p<0.001$)

and logging. Turnover of forestry and logging enterprises increases to 3.9 thousand LVL if the amount of non-financial investments in forestry and logging enterprises increases by one thousand LVL and model correspondence to real data is 61.72% (R^2=0.6172).

Profitability of forestry and logging enterprises is significantly dependent ($p<0.01$) on the amount of non-financial investments in enterprises of forestry and logging. Profitability of forestry and logging enterprises increases by 0.014% if amount of non-financial investments in forestry and logging enterprises increases by one thousand LVL and model correspondence to real data is 60.2% (R^2=0.602).

The area of regenerated forest land is significantly dependent ($p<0.01$) on the area of final felling (Figure 2). The area of regenerated forest land increases for 0.6 ha if the area of final felling increases by 1 ha and model correspondence to real data is 41.12% (R^2=0.421).

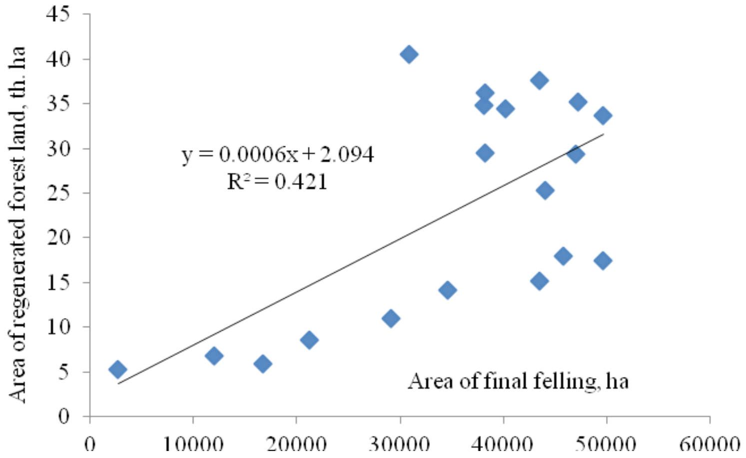

Figure 2. Correlation between area of regenerated forest land and area of final felling.

Employment in forestry and logging is significantly dependent ($p<0.001$) on the area of reconstructive felling. The number of employees in forestry and logging increases by 13.27 if the area of reconstructive felling increases by 1 ha and model correspondence to real data is 63.9% (R^2=0.639). This result of statistical analysis is controversial as from logical point of view the area of reconstructive felling could not have significant influence on employment as reconstructive felling in Latvia is done only in the area of a few hundred hectares annually.

Employment in forestry and logging is significantly dependent ($p<0.05$) on the area of commercial thinning. The number of employees in forestry and logging increases by 0.2 if the area of reconstructive felling increases by 1 ha and model correspondence to real data is 32% (R^2=0.320).

Input and output factor linear regression analysis for the set of processes Wood processing shows that not always there is a significant correlation between time series of chosen model factors (Table 2).

Value added (manufacture of timber products, manufacture of furniture) is significantly dependent on import of round wood ($p<0.1$) and type of activity ($p<0.01$) (Figure 3). The interaction effect of two factors is not significant. The

Table 2

Correlation between input and output factors in set of processes Wood processing

No	Input factor X_i	Output factor Y	R^2	Regression equation
1.	X_1: Import of round wood in forest sector, th. m^3 # X_2: Type of activity (manufacture of timber products, manufacture of furniture)** $X_1 * X_2$: Interaction effect of two factors	Y: Value added by type of activity (manufacture of timber products, manufacture of furniture), th. LVL.	0.716	Y=39863.24 + 139981.7236 (if X_2=manufacture of timber products) + 21.8288 * X_1 + 49.1889 * X_1 (if X_2=manufacture of timber products) + e_i
2.	X_1: Import of round wood, th. m^3*** X_2: Type of activity (manufacture of timber products, manufacture of furniture)* $X_1 * X_2$: Interaction effect of two factors	Y: Employment, (manufacture of timber products, manufacture of furniture), thousands of employees	0.920	Y=8.0256 + 17.9089 (if X_2=manufacture of timber products) + 0.0017 * X_1 + 0.00278 * X_1 (if X_2= manufacture of timber products) + e_i
3.	X_1: Import of round wood, th. m^3*** X_2: Type of activity (manufacture of timber products, manufacture of furniture)*** $X_1 * X_2$: Interaction effect of two factors	Y: Turnover of enterprises by type of activity (manufacture of timber products, manufacture of furniture), million LVL	0.927	Y= 80.95 + 419.0279 (if X_2=manufacture of timber products) + 0.060742 * X_1 + 0.29559 * X_1 (if X_2=manufacture of timber products) + e_i
4.	X_1: Import of firewood, th. tons X_2: Type of activity (manufacture of timber products, manufacture of furniture)*** $X_1 * X_2$: Interaction effect of two factors	Y: Employment, (manufacture of timber products, manufacture of furniture), thousands of employees	0.899	Y=9.0026 + 18.41 (if X_2=manufacture of timber products) + 0.02829 * X_1 + 0.3425 * X_1 (if X_2=manufacture of timber products) + e_i
5.	X_1: Import of firewood, th. tons X_2: Type of activity (manufacture of timber products, manufacture of furniture)*** $X_1 * X_2$: Interaction effect of two factors	Y: Turnover of enterprises by type of activity (manufacture of timber products, manufacture of furniture), million LVL	0.834	Y=99.799+ 515.618 (if X_2=manufacture of timber products) + 7.0162 * X_1 + 32.4655 * X_1 (if X_2=manufacture of timber products) + e_i
6.	X_1: Import of firewood, th. tons# X_2: Type of activity (manufacture of timber products, manufacture of furniture)** $X_1 * X_2$: Interaction effect of two factors	Y: Value added by type of activity (manufacture of timber products, manufacture of furniture), th. LVL.	0.740	Y=43398.6746+125035.7013 (if X_2 = manufacture of timber products)+3245.15*X_1+14605.476 * X_1 (if X_2=manufacture of timber products) + e_i
7.	X_1: Amount of non-financial investments in enterprises of manufacture of furniture, thousand LVL***	Y: Turnover of enterprises in manufacture of furniture, million LVL	0.794	Y= 82.7379 + 0.002707 * X_1

* - factor is significant with probability P=95% ($p<0.05$)

** - factor is significant with probability P=99% ($p<0.01$)

*** - factor is significant with probability P=99.9% ($p<0.001$)

\# - factor is significant with probability P=90% ($p<0.1$)

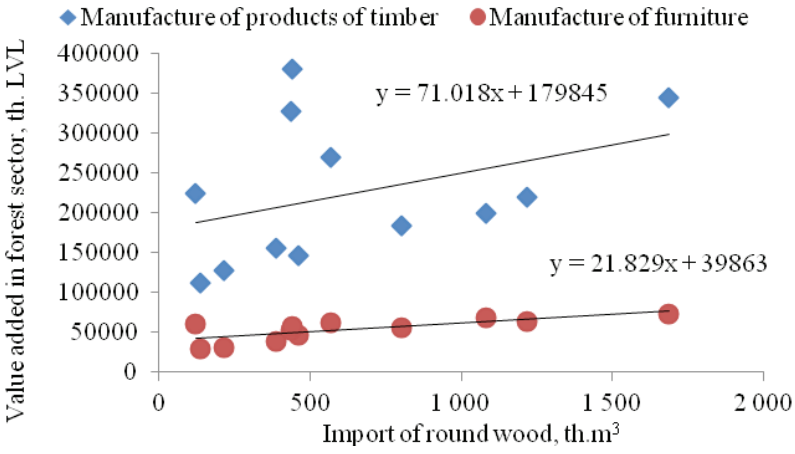

Figure 3. Correlation of value added in forest industry, import of round wood and type of activity (manufacture of timber products, manufacture of furniture).

correspondence of the model to real data was 71.6% (R^2=0.716).

General description of the model shown in Figure 3 is the following: Y = 39863.24 + 139981.72 (if X_2=manufacture of timber products) + 21.82 * X_1 + 49.19 * X_1 (if X_2= manufacture of timber products) + e_i

where

- Y is value added in forest industry (manufacture of timber products, manufacture of furniture), million LVL;
- X_1 is quantitative factor Import of round wood in forest sector, thousand m^3;
- X_2 is qualitative factor Type of activity with 2 values: manufacture of timber products, manufacture of furniture;
- e_i is random error.

When making interpretation of model direction coefficients, the model is transformed depending on the type of activity: manufacture of timber products, manufacture of furniture. As the result, following equations are obtained:

- if Type of activity is manufacture of timber products, then
  Y = 39863.24 + 139981.7236 (if X_2 = manufacture of timber products) + 21.8289* X_1 + 49.1889*X_1 (if X_2 = manufacture of timber products) + e_i or Y = 179845 + 71.018*X_1
- if Type of activity is manufacture of furniture, then
  Y = 39863.24 + 21.8289*X_1 + e_i

It means that

- if Import of round wood in forest sector (factor X_1) increases by 1000 m^3, then Value added in manufacture of timber products increases by 71.02 thousand LVL;
- if Import of round wood in forest sector (factor X_1) increases by 1000 m^3, then Value added in manufacture of furniture increases by 21.83 thousand LVL;
- taking into account that interaction effect of two factors X1 * X2 is not significant, it means that monetary increase of value added in the forest sector depending on the round wood import is not significantly different in manufacture of timber products or furniture. Graphically this means that two regression lines are almost parallel, as their increment rates are not significantly different.

Further on all other relations presented in Table 3 are described briefly.

The number of employed persons in forest sector (manufacture of timber products, manufacture of furniture) is significantly dependent on the import of round wood ($p<0.001$) and the type of activity ($p<0.05$). Interaction effect of two factors is not significant. The correspondence of the model to real data was 92.0% (R^2=0.920).

Turnover of forest sector enterprises by the type of activity (manufacture of timber products, manufacture of furniture) is significantly dependent on import of round wood ($p<0.001$), type of activity ($p<0.001$) and interaction effect of two factors ($p<0.001$). The correspondence of the model to real data was 92.7% (R^2=0.927).

The number of employed persons in forest sector (manufacture of timber products, manufacture of furniture) is significantly dependent on the type of activity ($p<0.001$), but it is not significantly dependent on the amount of imported firewood and interaction effect of two factors. The correspondence of the model to real data was 89.9% (R^2=0.899).

Turnover of enterprises by the type of activity (manufacture of timber products, manufacture of furniture) is significantly dependent on the type of activity ($p<0.001$), but it is not significantly dependent on the amount of imported firewood and interaction effect of two factors. The correspondence of the model to real data was 83.4% ($R^2=0.834$).

Value added in forest sector (manufacture of timber products, manufacture of furniture) is significantly dependent on the type of activity ($p<0.01$) and import of firewood ($p<0.1$), but it is not dependent on interaction effect of two factors. The correspondence of the model to real data was 74.0% ($R^2=0.740$).

Turnover of enterprises in the type of activity-manufacture of furniture is significantly dependent on amount of non-financial investments in enterprises of manufacture of furniture ($p<0.001$), and correspondence of the model to real data was 79.4% ($R^2=0.794$).

For development of comprehensive programme prototype for economic modelling of Latvian wood resource utilization and further processing, the forest sector model should be supplemented with information about demand of production between local and export markets, influence of used wood processing technologies on different output indices, employment dependency on volume and assortment of products manufactured.

Conclusions

1. The processes which are included in the model of Latvian forest industry are the following: Logging, Forest regeneration, Tending of young stands, Imports, Exports, Primary wood processing, Secondary wood processing and Energy production. Available data set of statistical information for detailed analysis of all processes of the model is limited. Therefore, regression analysis for most of the processes is done on a less detailed level – for set of processes Forestry and Wood processing.
2. In the model of forest industry every process is described with input and output information where output data of one process can be initial information to another process. Correlations of input and output value time series were analyzed with a simple and multiple regression analysis method.
3. Regression analysis was performed for 77 pairs of indices which theoretically could have significant relations, but in most of the cases correlation was unsatisfactory. In set of processes Forestry only for 7 correlations and in set of processes Wood processing for 12 correlations correspondence to real data was at least 30% ($R^2>0.3$).
4. Some correlations between time series of input and output data in the model of forest sector showed controversial results. For example, annual increase of area of reconstructive felling by 1ha leads to annual increase of number of employees by 13 which cannot be true.
5. From analyzed pairs of indicators, in programme prototype for economic modelling of Latvian wood resource utilization and further processing, should be used those regression equations, whose correspondence to real data is at least 70% ($R^2>0.7$).
6. The amount of information about processes and their quantitative indicators in Latvian wood processing industry is not sufficient for development of precise simulation models. There is need to develop a list of criteria for missing information and carry out its collection process.

Acknowledgements

This research was carried out within the frame of the project 'Substantiation of deciduous trees cultivation and rational utilisation, new products and technologies' in Latvia University of Agriculture (contract number: 2011. 10-4/VPP-5) and preparation of manuscript: by European Social Fund's project Support for the implementation of the Latvian Agricultural University doctoral study (contract number: 04.4-08/EF2.D2.45).

References

1. Becker J., Kahn D. (2003) The Process in Focus. In: Becker J., Kugeler M., Rosemann M. (eds.) *Process Management. A Guide for the Design of Business Processes*. Springer, Berlin, pp. 1-12.
2. Central Statistical Bureau of Latvia (2012) Statistics Database. Available at: http://csb.gov.lv/en/dati/statistics-database-30501.html, 22 October 2012.
3. Hangos K., Cameron I. (2001) *Process Modelling and Model Analysis*. Academic Press, London, 543 p.
4. Ministry of Agriculture of the Republic of Latvia (2012) Meža nozares produkcijas ārējā tirdzniecība (External trade of forest industry). Available at: http://www.zm.gov.lv/?sadala=289, 12 November 2012. (in Latvian).
5. State Forest Service (2012) Meža apsaimniekošana (Forest management). Available at: http://vmd.gov.lv/?sadala=2, 05 February 2013. (in Latvian).
6. Toppinen A., Kuuluvainen J. (2010) Forest sector modelling in Europe - the state of the art and future research directions. *Forest Policy and Economics*, 12, pp. 2-8.