Use of multivariate approaches in biomass energy plantation harvesting: logistics advantages

Corrado Costa[1*], K. Rotsios, G. Moshonas, D. Folinas

(1. *Consiglio per la Ricerca e la sperimentazione in Agricoltura (Unità di ricerca per l'ingegneria agraria) – Monterotondo Scalo (Roma) – Italy*;

2. *Consiglio per la Ricerca e la sperimentazione in Agricoltura (Unità di ricerca per le produzioni legnose fuori foresta) – Roma – Italy*)

Abstract: Agricultural biomass supply chain consisting of multiple harvesting, storage, pre-processing, and transport operations. This network operates in space and time coordinates and produces empirical data used for many purposes, including wood-flow planning, harvesting cost calculation and work rate setting. The aim of this study was to explore and propose the use of a multivariate approach, namely, the Partial Least Squares (PLS) multivariate regression approach and compare its performance with the commonly used Ordinary Linear Regression (OLS). In particular, the study aimed at comparing the main statistical significance of indicators attributed to models calculated with OLS and PLS regressions from the same original datasets, for the purpose of quantifying the eventual improvement, obtained with the new techniques. The dataset is composed by a series of measurements (harvesting distance, load carried, plantation production, numbers of plants harvested, and tractor engine power) conducted in a harvesting yard of a poplar plantation, to forecast the demanded working times. The technical analysis was accompanied by economic scenarios, based on three hypothetical harvesting yards. The results indicated that the PLS innovative approach is better performing; model error indicators are 5%-6% lower than those estimated with the OLS method. From an economic point of view the harvesting cost *per* ton ranges among 8.69-14.59 € t^{-1}, 12.10-16.56 € t^{-1} and 13.18-16.31 € t^{-1} referring to the different load capacity of the trailers, using the PLS model. Based on these results the differences between PLS and OLS varied up to 40 € ha^{-1}. PLS modeling and more in general the advanced multivariate approach, are getting increasingly popular, because they are very robust and are particularly suitable for modeling complex systems.

Keywords: harvesting, biomass, logistics, machine costs, multivariate statistics, ordinary linear regression, partial least square

Citation: Costa, C., G. Sperandio, and S. Verani. 2014. Use of multivariate approaches in biomass energy plantation harvesting: logistics advantages. Agric Eng Int: CIGR Journal, Special issue 2014: Agri-food and biomass supply chains, 70−79.

1 Introduction

The use of renewable energy alternatives to fossil fuels, which are considered the main causes of climate change since the end of the last millennium, is increasing. Among these, the woody biomass plays an important role (Vande Walle et al., 2007). For the emerging non-food bioenergy industry to ramp up to a mature, sustainable, and commercially viable industry, one of the challenges is the determination of supporting logistics, including strategic design of a storing/distribution network, a feedstock supply, residue handling, and a tactical (year-round) operation schedule (Tembo et al., 2003; Zhu et al., 2011). Recent advances in computational tools have made it possible to build mathematical models for analysis and optimization of complex supply systems. These tools are applied successfully to manufacturing, transportation, and supply chain management of many

Received date: 2013-02-13 **Accepted date:** 2013-12-01

* **Corresponding author: Corrado Costa**, Consiglio per la Ricerca e la sperimentazione in Agricoltura (Unità di ricerca per l'ingegneria agraria), Via della Pascolare, 16, 00015, Monterotondo Scalo (Roma) – Italy. Phone: +39–0690675214 – FAX +39–0690625591. Email: corrado.costa@entecra.it.

goods and services. The agricultural biomass supply logistics consists of multiple harvesting, storage, pre-processing, and transport operations. The entire network operates in space and time coordinates. Agricultural biomass supply logistics are characterized by a wide areal distribution of biomass, time and weather-sensitive crop maturity, variable moisture content, low bulk density of biomass material and a short time window for collection with competition from concurrent harvest operations. In this context, advanced technologies and analyses could help in optimizing logistics operations. From a technological point of view, an analysis of the processes in the supply chain from forests to mills revealed that there is a potential to streamline operations and make more efficient the use of resources by implementing an RFID-based (Aguzzi et al., 2012; Costa et al., 2012) log tracking system in the chain (Timpe, 2006). On the other hand, an optimized collection, storage and transport network can ensure timely supply of biomass with minimum costs (Sokhansanj et al., 2006). These kinds of studies often produce empirical models used for many purposes, including wood-flow planning, harvesting cost calculation, and work rate setting. At a more fundamental level, performance studies also allow understanding the behavior of harvesting machines and systems under varying stand and terrain conditions (Visser and Spinelli, 2011). Empirical performance models are generally developed by collecting field data and testing the statistical significance of any relationships with regression analysis. Pure analytical approaches, where the machine operations/performances are explicitly modeled in terms of their part operations, proposed harvesting strategies to minimize costs (Sorensen, 2003; Sogaard and Sorensen, 2004). The most commonly used regression type is Ordinary Least Square (OLS) linear regression. This technique is used to "calculate" an equation capable of representing the relationship between a dependent variable (typically time consumption or productivity) and one or more independent variables. The interest in exploring alternatives to OLS, such as multivariate predictive modeling based on the recombination of principal components (Principal Component Regression – PCR) or latent variables (Partial Least Square – PLS), is increasing (Costa et al., 2012). PLS is particularly useful when predicting one or more dependent variables from a large set of independent variables, often collinear. This technique originated within the field of economics (Wold, 1966) but became popular first in computational chemistry (Geladi and Kowalski, 1986) and then in human sensory evaluation (Martens and Naes, 1989). Today PLS regression is becoming a tool of choice in the social sciences, as a multivariate technique for non-experimental and experimental data (Costa et al., 2011).

Among the various crops for biomass option, especially, short rotation coppice (SRC) is regarded as a strategic resource of wood products (Verani et al., 2008) and seems to best reflect the expectations of farmers who used it to short return times and generally shows little enthusiasm for traditional wood plantations harvested at 10-30 years intervals (Spinelli et al., 2009). The SRC system is an intensive cultivation. The fast-growing hardwoods at high density are employed and the average period of rotation is less than 10 years (Rockwood et al., 2004). In Italy, during the last 10 years 7,000 ha of SRC has been established with poplar (*Populus spp.* L.), especially modern hybrids mainly in the Po river valley, where biomass plants for heat generation or for heat and power cogeneration have been recently built and were the regional program for rural development includes a series of financial incentives to support the establishment the plantation (Facciotto and Bergante, 2011). In the management of the energy plantation the harvesting operation is very important, because its costs can strongly influence the economic performance of the overall supply chain. The harvesting of SRC can be performed principally with two different systems: the first one, the cut and storage system, requires that the trees are cut and moved to a storage area and chipped after storage; in the second one, the combined cut and chips system which is also the more used one, the plantation is harvested ~~or~~ with a modified forage harvester machine, whose standard header is replaced by a special cutting head (Spinelli et al., 2009; Schweier and Becker, 2012). The chips are blown into an accompanying tractor-pulled trailer, which transports the chips to a collection point (Sambra et al., 2008).

The aim of this study is to explore and propose the use of a multivariate approach, such as PLS multivariate regression approach, innovative for this kind of logistics applications, and compare its performance with the commonly used OLS. In particular, the study aimed at comparing the main statistical significance indicators attributed to models calculated with OLS and PLS regressions from the same original datasets, for the purpose of quantifying the eventual improvements obtained with the new techniques. The dataset is composed by a series of measurements (harvesting distance, the load carried, plantation production, numbers of plants harvested, and tractor engine power) conducted in a harvesting yard of a poplar plantation in order to predict the working times. The technical analysis is followed by economic scenarios based on three different hypothetical harvesting yards.

2 Materials and methods

2.1 Data collection

The study was carried out in the site “Le risaie” in the Viterbo municipality, Latium region (Central Italy), [42° 22′47″ N, 12°02′21″ E]. The plantation R2S2 (two years root and two years stem), was established with poplar clone AF2 covering an area of 15.4 ha. The cutting were planted in single rows with a spacing of 0.66 m while the distance between the rows was 2.50 m and the density of plantation was 6,060 cuttings/ha. The Claas forager Jaguar 880, with the header GBE-1 was used to harvest the plantation. The chips were blown into a trailer pulled by a tractor and transported to storage. Three trailers of different volumes, namely 25, 16, and 13 m^3, were employed during the harvesting. The load capacity of the different trailers, established as average value of three weighing for each trailer, was of 7.37, 4.74, and 3.66 t. The crew consisted of four people. The trailers were pulled by a Lamborghini 165 DT tractor (120 kW), a Fiat 115 DT (84 kW) tractor, and a Fiat 80/90 DT tractor (58 kW), respectively. For experimental data, the cycle time of machine was divided into time elements (working phases) that were considered typical of the work (Karcha et al., 2005; Puttock et al., 2005; Verani et al., 2010b; Picchio et al., 2012). Work time was recorded for every single phase, using a Minerva chronometric table equipped with three centesimal chronometers (Harstela, 1991; Acuna et al., 2012).

2.2 Modeling approaches

The OLS linear regression approach was based on the following general equation:

$$T = A + Bx_1 + Cx_2 \quad (1)$$

where, T is the gross work time for harvesting cycle (min); A, B and, C are constants to be determined and x_1 and x_2, the harvesting distance (m) and the load carried (t), respectively (Ghaffariyan et al., 2009; Gallis and Spyroglou, 2012). The harvesting distance is given by the sum of the lengths of the single rows needed to fill a trailer, or part of them. The distance was measured by the laser gauge. The OLS regression was analyzed with ANOVA test. The surface harvested for single load was calculated multiplying the rows' length by the distance between the rows (2.50 m).

An alternative regression approach, the PLS-based was implemented. PLS is used to find the fundamental relations between two matrices (X and Y) and represents a latent variable approach to modeling the covariance structures in these two spaces. A PLS model was used to find the multidimensional direction in the X space that explains the maximum multidimensional variance direction in the Y space. A number of variants of PLS exist; in this study the SIMPLS (De Jong, 1993) algorithm was implemented. The independent variables composing the X-block consist of the following five variables: harvesting distance (m), load carried (t), plantation production (t ha^{-1}), numbers of plants harvested, and tractor engine power (kW). Both X- and Y-blocks (gross time for harvesting cycle) were transformed using the 'autoscale' procedure. PLS was computed using PLS toolbox 6.2 (Eigenvector research) for Matlab 7.1. For details on the PLS method see Costa et al. (2012). Residual error indicators, such as the Root Mean Square Errors in Calibration (RMSEC) and in Validation (RMSECV) were calculated. The predictive ability of the model was partially dependent on the number of the latent vectors used and was assessed through the following statistical indicators: Root Mean Square Error (RMSE), Standard Error of Prevision (SEP) and correlation coefficient (r). Finally, we calculated the Ratio of Percentage Deviation (RPD), which is the ratio of the standard deviation of the measured data to the RMSE (Williams, 1987). This represents the factor by which the prediction accuracy has been increased compared with using the mean of the original data. Generally, a good predictive model should exhibit high values for r and low values for RMSE and SEP. The model chosen was for the number of LV (Latent Vector) that yielded the highest r, minimum SEP for predicted and known Y-block and maximum RPD.

2.3 Economic analyses

For both OLS and PLS approaches, the production of plantation was determined extrapolating the value of load per hectare. To check which are the best trailers to be used in the harvesting, three harvesting yards employing three trailers with equal capacity have been hypothesized. The aim of the economic analysis was to identify the best (in terms of lower harvesting cost *per* ton and *per* hectare) among three hypothesized harvesting yards. The three hypothesized harvesting yards (having the same harvester Claas Jaguar 880) are composed by:

1) Hypothesis 1: three Lamborghini 165 DT tractors equipped with a 25 m^3 load capacity trailer each;

2) Hypothesis 2: three Fiat 115 DT tractors equipped with a 16 m^3 load capacity trailer each;

3) Hypothesis 3: three Fiat 80/90 DT tractors equipped with a 13 m^3 load capacity trailer each.

The hourly cost calculation of the machines and equipment for each harvesting yard is based on the analytical methods of calculation proposed by different authors (Ribaudo, 1977; Miyata, 1980). The principal elements considered in the economic analysis and the hourly costs of the single machines and equipment, are described in Table 1. Table 2 reports the hourly costs of the three harvesting yards harvesting hypotheses based on the costs reported in Table 1.

Table 1 Principal technical and economic elements used to calculate machine cost (upper part) and the relative hourly machine

costs obtained (lower part)

Description	Claas Jaguar 880	Header GBE-1	Tractor Fiat 80/90 DT	Tractor Fiat F 115 DT	Tractor Lamborghini 165 DT	Trailer 1 (25 m^3)	Trailer 2 (16 m^3)	Trailer 3 (13 m^3)
Purchase price , €	250,000	90,000	40,000	57,000	80,000	14,000	11,000	10,000
Salvage value, €	41,943	15,099	4,295	6,120	8,590	962	756	687
Life period, y	8	8	10	10	10	12	12	12
Productive machine hours, h y^{-1}	800	800	1000	1000	1000	300	300	300
Engine power, kW	350	-	58	84	120	-	-	-
Interest rate, %	6	6	6	6	6	6	6	6
Fuel consumption, l/h	44.18	-	8.27	11.73	16.43	-	-	-
Lubricant consumption, l/h	1.77	0.10	0.33	0.47	0.66	0.05	0.05	0.05
Driver cost, € h^{-1}	23	23	15	15	15	15	15	15
Fuel price, € L^{-1}	1.05	-	1.05	1.05	1.05	-	-	-
Lubricant price, € L^{-1}	9	-	9	9	9	-	-	-
Fixed costs, € h^{-1}	53.51	16.98	6.31	8.90	12.41	5.32	4.22	3.86
Variable costs, € h^{-1}	117.10	14.21	32.28	38.55	46.89	2.76	2.43	2.31
Total costs, € h^{-1}	170.61	31.19	38.59	47.45	59.30	8.08	6.65	6.17

Table 2 Harvesting yards hypotheses considered in the cost analysis of the poplar plantation harvesting (the three hypotheses used the same harvester Claas Jaguar 880)

Harvesting yard hypotheses	Fixed cost (F_c, € h^{-1})	Variable cost (V_c, € h^{-1})	Total cost (T_c, € h^{-1})
Hypothesis 1 (N.3 Lamborghini 165 DT and N.3 trailers 25 m^3)	23.68	280.26	403.94
Hypothesis 2 (N. 3 Fiat 115 DT and N.3 trailers 16 m^3)	109.86	254.23	364.08
Hypothesis 3 (N.3 Fiat 80/90 DT and N.3 trailers 13 m^3)	100.98	235.10	336.08

The following economic parameters were calculated from the results obtained by the two regression approaches (*i.e.*, OLS and PLS):

a) Biomass production:

$$Biomass_{ha} = \frac{x_2}{dist}\ 4000 \tag{2}$$

where, $Biomass_{ha}$ is the estimated biomass plantation production (t ha^{-1}); x_2 is the load carried (t), using the different hypotheses and *dist* is the length of the harvested rows for the same load carried (m), and 4000 is a constant indicating the harvesting linear distance for each hectare.

b) Harvesting productivity:

$$Pr = \frac{x_2}{T} \tag{3}$$

where, Pr is the harvesting productivity (t h^{-1}); x_2 is the load carried (t) using the different hypotheses, and T is the harvesting gross work time (h) calculated with the two regression approaches.

c) Harvesting cost (mass based)

$$HC_t = \frac{T_c}{Pr} \tag{4}$$

where, HC_t is the harvesting cost *per* ton (€ t^{-1}); T_c is the total hourly cost for each hypothesis (€ h^{-1}; Table 2) and Pr is the harvesting productivity (t h^{-1}) expressed in the Equation (3).

d) Harvesting cost (area based)

$$HC_{ha} = HC_t\ \ Biomass_{ha} \tag{5}$$

where, HC_{ha} is the harvesting cost *per* hectare (€ ha^{-1}); HC_t and $Biomass_{ha}$ are expressed in the Equations (4) and (2) respectively.

3 Results and discussion

3.1 Technical results

The average harvesting distances of the R2F2 plantation was 1,254 m (ranged from 480 m to 2,552 m) and the average of load per trailer was 5.32 t. The average gross time per trip was 11.67 min, and the average gross productivity was 28.16 t h^{-1}.

The average load carried was 3.66, 4.74, and 7.37 t, for the trailers with 13, 16, and 25 m^3, respectively. The percentages of the harvesting time are reported in Figure 1. The operating time of harvesting (composed by harvesting and reversing times) was high (89.95%) followed by the time due to mechanical and personal delay (8.61%) and waiting time (1.44%). The operating speed of the machine was equal to 6.43 ± 1.28 km h^{-1}.

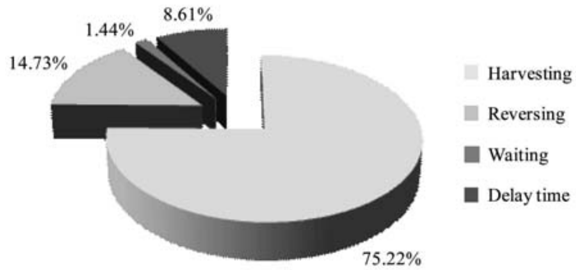

Figure 1 Percentages of working times of the plantation harvesting

Harvesting dedicated plantations performance for the Claas Jaguar with the header GBE-1 showed results as demonstrated by the high percentage of harvesting time with respect to the total working time (75.22%).

The OLS method showed that the gross time for harvesting cycle was expressed by the equation:

$$T = 0.48 + 0.0043x_1 + 1.077x_2 \tag{6}$$

where, T is the gross time for harvesting cycle (min); and x_1, x_2, the harvesting distance (m) and load carried (t), respectively. ANOVA reported a p value lower than 0.0001.

Table 3 shows the main indicators for the OLS and PLS regression models.

Table 3 Main goodness-of-fit indicators for the regression models OLS and PLS

Dataset	R2F2	
Regression analysis	OLS	PLS
Observations (n)	49	49
X Variables (n)	2	5
Latent Vectors (n)	-	3
% Cumulated Variance X-block	-	99.25
% Cumulated Variance Y-block	-	85.97
RMSEC	-	1.50
RMSECV	-	1.69
r	0.919	0.927
r^2	0.844	0.859
RMSE	1.60	1.52
SEP	1.60	1.52
RPD	2.53	2.67

The model error indicators (SEP and RMSE) were

5%-6% lower than the PLS regression model compared with the OLS one. Moreover, the r value was higher for PLS model, with an increment of 0.8% over OLS model. RPD was higher for the PLS regression model. Based on the RPD classification, PLS regression allows the possibility of increasing the predictive power of ordinary regression models. Models are both considered as “excellent”. Maybe increasing the number of observations and variability of the samples the models could reach better RPD scores. PLS is a better performing model than the OLS model as also demonstrated by Costa et al. (2012).

Table 4 shows the relative contribution (loadings) of individual X-variables to each of the latent vectors of both PLS models.

Table 4 PLS Model: X variable loadings for Latent Vectors (LVs)

	LV1	LV2	LV3
Distance	0.50	-0.19	-0.39
Load	0.46	0.41	-0.09
Production	-0.26	0.80	-0.55
Plant density	0.50	-0.19	-0.39
Tractor power	0.47	0.34	0.62

All the variables, except Production, contributed highly to the first LV. Production (0.080) has the highest effect on LV2. Tractor power and Production gave the highest contribution to the third LV (Table 4). The observed *vs* predicted independent Y variable (TL, min) for the OLS and PLS models are reported in Figure 2.

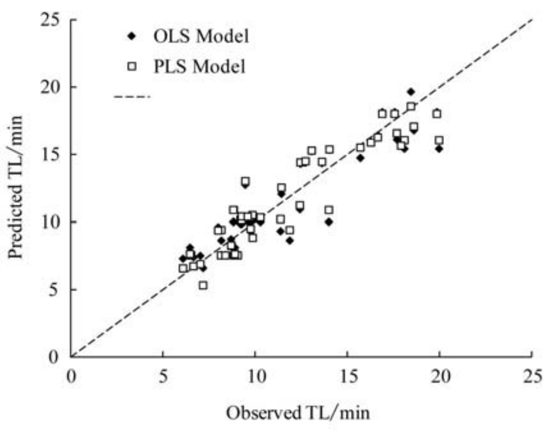

Figure 2 Observed *vs* predicted Y (TL, min) for the OLS and PLS models

The gross work time *T*, and the harvesting productivity *Pr* (Equation (3)) *per* type of trailer load, calculated using OLS and PLS models, are shown in Figure 3. In particular, the gross harvesting time (calculated with both PLS and PLS models) is expressed as a function of the harvesting distance *per* single load carried (Figure 3A, while, the harvesting productivity was obtained using the Equation (3), calculated with both OLS and PLS models, and obtained as a function of the biomass *per* hectare (Figure 3B). Figure 3A shows a similar trend for all the three hypotheses. But only hypothesis 1 (higher load capacity) showed that the curves mostly overlap; in general the trends of the PLS models are less linear as compared to the OLS ones. With lower load capacity (hypothesis 3) the OLS model overestimates the gross harvesting time; on the contrary, at medium load capacity (hypothesis 2) the OLS model underestimates the gross harvesting time. These trends are reflected and amplified in the estimated values in Figure 3B, where, in particular, the lower load capacity (hypothesis 3) tends to diverge at biomass production *per* hectare higher than 20 t ha^{-1}.

The harvesting operation could be considered as fast as demonstrated by the gross productivity obtained by the machine ranging from 22 to 39 t h^{-1}. The difference between the two approaches in estimating the gross productivity, ranged up to 6% at higher/intermediate load capacity and up to 13% at lower load capacity.

3.2 Economic results

In Figure 4 both harvesting costs *per* ton (Equation 4; Figure 4A) and *per* hectare (Equation (5); Figure 4B) scenarios for each hypotheses and regression model (*i.e.*, OLS and PLS) are presented. Harvesting costs *per* ton and hectare are decreasing and increasing, respectively, both in relation to the increase of the estimated biomass *per* hectare. Considering the estimated costs *per* ton (Figure 4A) OLS model overestimated at lower load capacity (13 m^3), while at intermediate load capacity (16 m^3) the trend was opposed with respect to the PLS models; for higher load capacity (25 m^3) the two approaches have different slopes intersecting approximately in correspondence with 16 t ha^{-1} of biomass. Observing the estimated costs *per* hectare

(Figure 4B) it is possible to observe the opposed trend with respect to the costs *per* ton graph (Figure 4A). In addition it is possible to observe a sharper curve shape of the PLS model at lower load capacity (13 m^3). In both harvesting costs *per* ton and *per* hectare at higher load capacity (25 m^3) the OLS model overestimates and underestimates the costs below and above 16 t ha^{-1} of biomass, respectively. Both provisional models (Figure 3 and Figure 4) show how the use of trailers with lower load capacity (13 m^3) minimizing the harvesting costs of the R2F2 poplar plantation increases the logistics advantages.

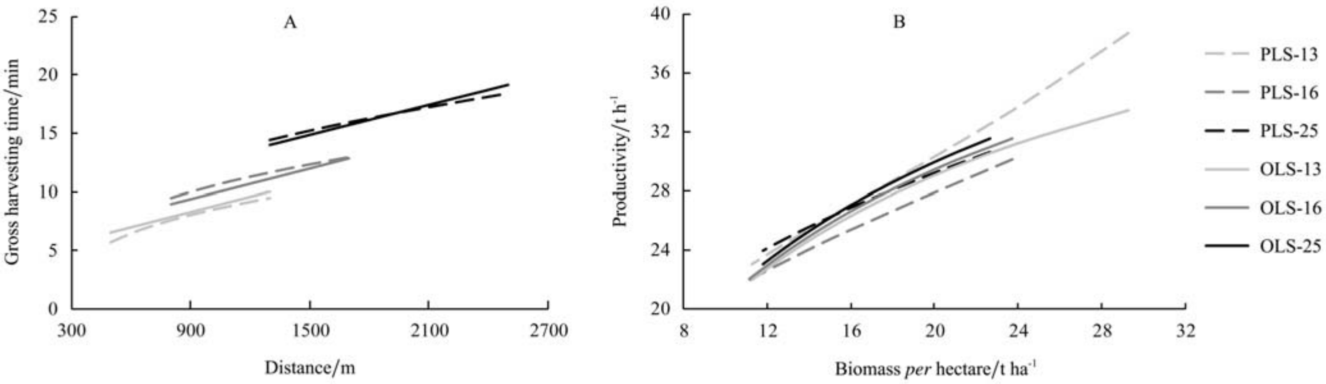

Figure 3A Gross harvesting time predicted by PLS and OLS for three yard sites in relationship with the distance *per* load.
Figure 3B Work productivity calculated with PLS and OLS model as a function of the variation of biomass produced for the three hypothetical work sites

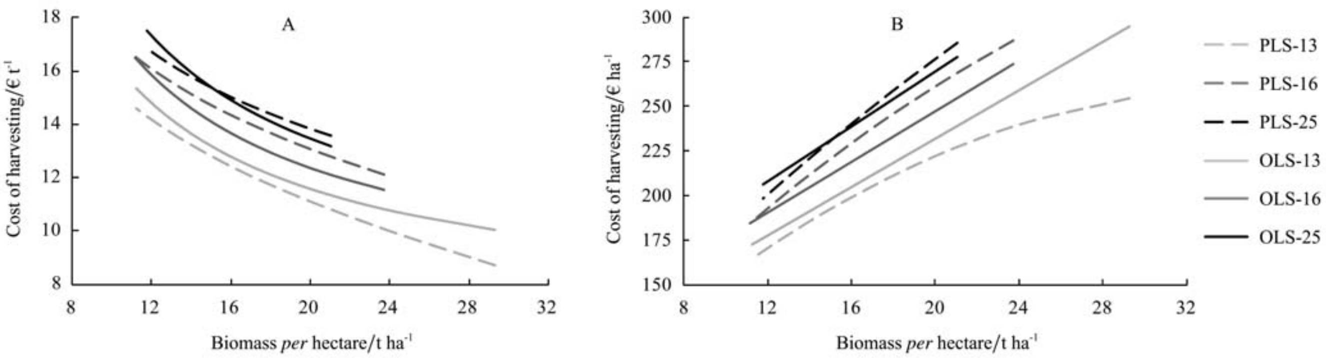

Figure 4 Comparison between OLS and PLS approaches on the cost analysis of the harvesting in relation to production of biomass *per* hectare of the plantation for three hypothesis of harvesting yard: A - Cost *per* ton (€ t^{-1}); B - Cost *per* hectare (€ ha^{-1})

Figure 5 represents the percentage differences between the two approaches at different load capacities. It is possible to observe that at lower load capacity (13 m^3) OLS tends always to overestimate ranging from 4% to 16%. The intermediate load capacity (16 m^3) OLS trend, instead, underestimates ranging from -4% to -0.3%. The higher load capacity (25 m^3) showed similar estimation with respect to PLS model ranging from -2.9% to 4%.

By using the PLS model, the harvesting cost *per* ton ranged 8.69-14.59 € t^{-1}, 12.10-16.56 € t^{-1} and 13.18-16.31 € t^{-1}; while the costs *per* hectare ranged 164-254 € ha^{-1}, 185-287 € ha^{-1} and 199-286 € ha^{-1}, and these values referring to the different trailers load capacities, 13, 16 and 25 m^3 respectively. At higher load capacity, the harvesting productivity increased, but insufficiently to balance the higher hourly cost of the harvesting yard. The differences between the results produced by the two approaches are higher especially at lower load capacities. Basing on these results, the differences between PLS and OLS varied up to 40 € ha^{-1}. These results remain valid under the same experimental conditions, where the different load capacity of the trailer did not influence the waiting time, because the distance to unload the chips at the landing was not that large and the three tractors were always ready to interchange. Moreover, when the distance from the unloading site become longer, the logistics advantage of the harvesting yard with greater load capacity will, of course, increase caused by the reduced waiting time.

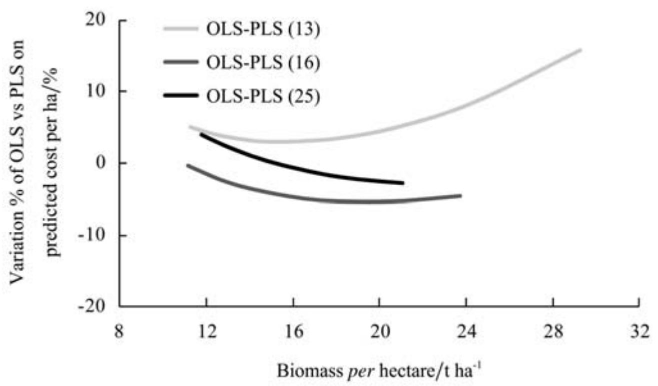

Figure 5 Comparison between OLS and PLS approach on the cost analysis of the harvesting in relation to production of biomass per hectare of the plantation for three hypotheses of harvesting yards

3.3 General considerations

PLS regression analysis does offer some benefits over ordinary regression analysis (Lipp, 1996). The substantial improvement of all goodness-of-fit indicators is probably the most visible benefit. Moreover, other benefits of the PLS regression technique are not merely the increase of a coefficient, but the capacity of detecting significant variables otherwise discarded with ordinary regression techniques (Costa et al., 2009). This is the advantage of Latent Vectors, which are capable of integrating the effect of more independent variables. A further advantage of PLS regression over multiple linear regression lies in the definition of the new variables, whose definition takes into account not only the values assumed by the X but also their correlation with the dependent variables (Kresta, 1992). In this respect, it is most interesting to compare the X-variables included in the ordinary and PLS regression models obtained from the same datasets. Another advantage in using PLS regression is that this method could handle many collinear variables. Ordinary regression would pick one or the other, but the use of latent vectors in PLS regression makes it possible to select more than one attribute for the same characteristic, after weighing their contribution through pre-processing. The larger number of X-variables included in the PLS regression model also guarantees a more accurate description of complex processes such as biomass supply logistics, where different and often unpredictable factors influence the variable to be estimated. On the other hand, this approach requires a larger effort when gathering input data (Costa et al., 2012).

4 Conclusions

The agricultural biomass supply logistics consists of multiple harvesting, storage, pre-processing, transport operations, and networking in space and time coordinates, producing empirical models used for many purposes, including wood-flow planning, harvesting cost calculation and work rate setting. The interest in exploring alternatives to ordinary linear regression, such as multivariate predictive modeling based on latent variables, is increasing. We demonstrated how PLS regression analyses allow producing models that better fit the original data, compared to OLS. Additionally, PLS regression analyses allow handling collinear variables, facilitating the extraction of sound models from large amounts of field data obtained from biomass logistics operations. This could lead to more robust models in terms of both variable oscillations and higher repeatability. The models themselves are somewhat less applicable than standard regression equations. Nevertheless, PLS modeling, and generally the advanced multivariate approach, is getting increasingly popular because it is very robust and particularly suitable for modeling complex systems.

Acknowledgements

This publication benefits from participation in the COST actions FP-0902 “Development and harmonization of new operational research and assessment procedures for sustainable forest biomass supply” and FP-1104 “New possibilities for print media and packaging - combining print with digital”.

References

Acuna, M., M. Bigot, S. Guerra, B. Hartsough, C. Kanzian, K. Kärhä, O. Lindroos, N. Magagnotti, S. Roux, R. Spinelli, B. Talbot, E. Tolosana, and F. Zormaier. 2012. Good practice guidelines for biomass production studies. Editors: Magagnotti N., Spinelli R. COST Action FP-0902 ISBN 978-88-901660-4-4.

Aguzzi, J., J. B. Company, C. Costa, M. Matabos, E. Azzurro, A. Mànuel, P. Menesatti, F. Sardà, M. Canals, E. Delory, D. Cline, P. Favali, S. K. Juniper, Y. Furushima, Y. Fujiwara, J. J. Chiesa, L. Marotta, and I. M. Priede. 2012. Challenges to assessment of benthic populations and biodiversity as a result of rhythmic behaviour: video solutions from cabled observatories. *Oceanography and Marine Biology,* 50: 235-286.

Costa, C., C. Angelini, M. Scardi, P. Menesatti, and C. Uteri. 2009. Using image analysis on the ventral colour pattern in *Salamandrina perspicillata* (Savi, 1821) (Amphibia, Salamandridae) to discriminate among populations. *Biological Journal of the Linnean Society,* 96(1): 35-43.

Costa, C., F. Antonucci, F. Pallottino, J. Aguzzi, D. W. Sun, and P. Menesatti. 2011. Shape analysis of agricultural products: a review of recent research advances and potential application to computer vision. *Food and Bioprocess Technology,* 4(5): 673-692.

Costa, C., P. Menesatti, and R. Spinelli. 2012. Performance modelling in forest operations through partial least square regression. *Silva Fennica,* 46(2): 241-252.

Costa, C., F. Antonucci, F. Pallottino, J. Aguzzi, D. Sarriá, P. Menesatti. 2013. A review on agri-food supply chain traceability by means of RFID technology. *Food and Bioprocess Technology,* 6(2): 353-366.

De Jong, S. 1993. SIMPLS: an alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems,* 18(3):251-263.

Facciotto, G., and S. Bergante. 2011. Nine years of measurements in Italian SRC trial with 14 poplar and 6 willow clones. *19th European Biomass Conference and Exhibition*. 6-10 June Berlin, Germany.

Gallis, C., and G. Spyroglou. 2012. Productivity linear regression models of tree-length harvesting system in natural coastal Aleppo pine (*Pinus halepensis* l.) forests in the Chalkidiki area of Greece. *Croatian Journal of Forestry Engineering,* 33(1): 115-123.

Geladi, P., and B. Kowalski. 1986. Partial least square regression: a tutorial. *Analytica Chemica Acta,* 185: 1-17.

Ghaffariyan, M. R., K. Stampfer, and J. Sessions. 2009. Production equations for tower yarders in Austria. *International Journal of Forestry Engineering* ,20(1): 17-21.

Harstela, P. 1991. Work study in forestry. *Silva Carelica,* 18: 1-41.

Karcha, K., A. Jouhiaho, A. Mutikainen, and S. Mattila. 2005. Mechanized energy wood harvesting from early thinnings. *International Journal of Forestry Engineering,* 16(1): 15-26.

Kresta, J. 1992. The application of Partial Least Squares to problems in chemical engineering. Ph. D. thesis at McMaster University, Canada: 215.

Lipp, M. 1996. Comparison of PLS, PCR and MLR for the quantitative determination of foreign oils and fats in butter fats of several European countries by their triglyceride composition. *Zeitschrift für Lebensmittel Untersuchung und Forschung A,* 202(3): 193-198.

Martens, H., and T. Naes. 1989. *Multivariate calibration*. London: Wiley.

Miyata, E. S. 1980. Determining fixed and operating costs of logging equipment. North Central Forest Experiment Station, Forest Service, USDA, St. Paul, MN: 14.

Picchio, R., A. Sirna, G. Sperandio, R. Spina, and S. Verani. 2012. Mechanized harvesting of eucalypt coppice for biomass production using a high mechanization level. *Croatian Journal of Forestry Engineering,* 33(1): 15-24.

Puttock, D., R. Spinelli, and B. R. Hartsough. 2005. Operational trials of cut-to-length harvesting of poplar in a mixed wood stand. *International Journal of Forestry Engineering,* 16(1): 39-49.

Ribaudo, F. 1977. Il costo di esercizio delle macchine agricole. *Macchine e Motori Agricoli* 11: 101-116.

Rockwood, D. L., C. V. Naidu, D. R. Carter, M. Rahmani, T. A. Spriggs, C. Lin, G. R. Alker, J. G. Isebrands, and S. A. Segrest. 2004. Short-rotation woody crops and phytoremediation: opportunities for agroforestry? *Agroforestry Systems,* 61(1-3): 51-63.

Sambra, A., C. A. G. Sorensen, and E. F. Kristensen. 2008. Optimized harvested and logistics for biomass supply chain. *Proceedings of European Biomass Conference and Exhibition*. Valencia, Spagna.

Sokhansanj, S., A. Kumarc, and A. F. Turhollow. 2006. Development and implementation of integrated biomass supply analysis and logistics model (IBSAL). *Biomass and Bioenergy,* 30(10): 838-847.

Sørensen, C. G. 2003. Workability and machinery sizing for combine harvesting. *The CIGR Journal of AE Scientific Research and Development* 5. (Manuscript PM 03 003) ISSN 1682-1130.

Søgaard, H. T., and C. G. Sørensen. 2004. A model for optimal selection of machinery sizes within the farm machinery system. *Biosystems Engineering,* 89(1): 13-28.

Spinelli, R., C. Nati, and N. Magagnotti. 2009. Using modified foragers to harvest short-rotation poplar plantation. *Biomass and Bioenergy,* 33(5): 817-821.

Tembo, G., F. M. Epplin, and R. L. Huhnke. 2003. Integrative investment appraisal of a lignocellulosic biomass-to-ethanol industry. *Journal of Agricultural Research and Economics,* 28(3): 611-633.

Timpe, D. 2006. RFID in forestry: prospects of an RFID-based log tracking system as an alternative to stamping. Fibre Science and Communications Network Report, Mittuniversitetet, Sundvall, Sweden, 53pp. ISSN 1650-5387 2006: 39.

Vande Walle, I., N. Van Camp, L. Van de Casteale, K. Verheyen, and R. Lemeur. 2007. Short-rotation forestry of birch, maple, poplar and willow in Flanders (Belgium) I Biomass production after 4 years of tree growth. *Biomass and Bioenergy,* 31(5): 267-275.

Verani, S., G. Sperandio, R. Picchio, R. Spinelli, and G. Picchi. 2008. Poplar harvesting. International Poplar Commission Working Paper IPC/8. FAO, Rome. 60 pp.

Verani, S., G. Sperandio, and R. Picchio. 2010b. First thinning in a coniferous plantation for biomass production: productivity and costs. *Proceedings of the Forest Engineering: Meeting the Needs of the Society and the Environment.* July 11-14. Padova. Italy.

Visser, R., and R. Spinelli. 2011. Determining the shape of the productivity function for mechanised felling and felling-processing. *Journal of Forest Research,* 17(5): 397-402.

Williams, P. 1987. Variables affecting near-infrared reflectance spectroscopic analysis. In: Williams, P. & Norris, K. (eds.). Near-infrared technology in the agricultural and food industries. American Association of Cereal Chemists, St Paul, Minnesota, USA. p. 143-166.

Wold, H. 1966. Estimation of principal components and related models by iterative least squares. In: Krishnaiaah, P.R. (ed.). Multivariate analysis. New York: Academic Press. p. 391–420.

Zhu, X., X. Li, Q. Yao, and Y. Chen. 2011. Challenges and models in supporting logistics system design for dedicated-biomass-based bioenergy industry. *Bioresource Technology,* 102(2): 1344-1351.