Computers and Electronics in Agriculture 91 (2013) 116–120

Contents lists available at SciVerse ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Support Vector Regression to predict carcass weight in beef cattle in advance of the slaughter

Jaime Alonso [a,*], Ángel Rodríguez Castañón [b], Antonio Bahamonde [a]

[a] Artificial Intelligence Center, University of Oviedo at Gijón, Asturias, Spain
[b] Association of Breeders of Asturiana de los Valles (ASEAVA), Abarrio, No. 24, E-33424 Llanera, Asturias, Spain

ARTICLE INFO

Article history:
Received 22 June 2012
Received in revised form 8 August 2012
Accepted 17 August 2012

Keywords:
Support Vector Machines (SVMs)
Support Vector Regression (SVR)
Carcass weight
Beef cattle

ABSTRACT

In this paper we present a function to predict the carcass weight for beef cattle. The function uses a few zoometric measurements of the animals taken days before the slaughter. For this purpose we have used Artificial Intelligence tools based on Support Vector Machines for Regression (SVR). We report a case study done with a set of 390 measurements of 144 animals taken from 2 to 222 days in advance of the slaughter. We used animals of the breed Asturiana de los Valles, a specialized beef breed from the North of Spain. The results obtained show that it is possible to predict carcass weights 150 days before the slaughter day with an average absolute error of 4.27% of the true value. The prediction function is a polynomial of degree 3 that uses five lengths and the estimation of the round profile of the animals.

1. Introduction

In bovine beef cattle, carcass weight is the most important data to compute the price of the carcass. In this paper we present a method to estimate that weight in advance of the slaughter of the animal. The study was done with animals of the breed *Asturiana de los Valles*, a beef breed of the North of Spain. This is a specialized breed with many double-muscled individuals; their carcass have dressing percentages over 60%, with muscle content over 75%, and with a low (8%) percentage of fat (Piedrafita et al., 2003). The market target of these carcasses is made up of those consumers that prefer lean meat without any marbling (del Coz et al., 2005; Díez et al., 2005).

In the breed *Asturiana de los Valles*, the price of the carcass is computed from three factors: sex, conformation, and carcass weight. Since sex and conformation are known, the core data is carcass weight. Thus, the prediction function presented in this paper is an important tool to manage the profitability of livestock farms.

Now, the procedure typically used to estimate the carcass weight consists in computing a percentage of live weight. Thus, in (Martínez et al., 1999) suggest for males 66% for double-muscled, and 61% for single-muscled; in the case of females the percentages are 3 percentage points less in each case.

However, this method has several drawbacks. On the first hand, it requires that animals must be weighted just before the slaughter in order to achieve accurate results; in practice this is not feasible. On the other hand, the main disadvantage of the percentages is that they do not consider any individual morphological feature; in fact not all double (or single) muscled animals are equal.

The approach presented here is based on Machine Learning procedures that aim to learn a function to map carcass weights from a collection of morphological measures of the animals and the number of days until the slaughter. In this way, individual peculiarities of the animals are involved in weight estimation.

To learn the function able to predict the carcass weight we used Support Vector Machines (SVMs) (Vapnik, 1998) for regression (SVR) (Smola, 1996). We discuss the use of different options to configure this algorithm using *kernels* functions. We used this technology since SVM (SVR) are acknowledged as the most powerful learning algorithms in many application fields.

From a mathematical point of view, given a classification (or regression) learning task, a SVM (SVR) solves a *convex* optimization problem. The solution gives rise to a hypothesis able to predict unseen cases drawn with the same distribution of the initial learning task. The advantage of the convexity in this context is that it guarantees that there exists only one optimal solution. Therefore the optimizer of SVM (or SVR) will not return a local minimum instead of the best one as happens, for instance with Artificial Neural Networks.

The estimation of the weight of bovines has been studied since a long time ago; see, for instance (Enevoldsen and Kristensen, 1997). The use of Artificial Intelligence tools to predict beef cattle scores is

* Corresponding author. Tel.: +34 985 182 032.
E-mail addresses: jalonso@aic.uniovi.es (J. Alonso), arcastanon@aseava.com (Á.R. Castañón), antonio@aic.uniovi.es (A. Bahamonde).

0168-1699/$ - see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.compag.2012.08.009

not new. In (González-Velasco et al., 2011 and Alonso et al., 2007) assessment functions for beef cattle are presented. Moreover, to estimate live weights of bovines in (Stajnko et al., 2008; Tasdemir et al., 2011a,b) the authors used digital image processing procedures.

2. Material and methods

2.1. Data

In the research reported in this paper we used a dataset of 390 numerical descriptions of animals of the breed *Asturiana de los Valles* corresponding to 144 bovines of both sexes; some of them were measured several times in order to grasp the evolution of the estimations depending of the time.

The ages of the animals when they were measured range from 7 to 14 months, the ages where the commercial activities are concentrated in this breed. The measurements sessions took place with 2–222 days in advance of the slaughter, the average value is 74 days. The carcass weights of the animals average 221.6 kg ranging from 116 to 459 kg.

Table 1 summarizes the description of the dataset used in the paper.

2.2. Numerical description of animals

In (Alonso et al., 2006, 2007) we identified and described a set of morphological features relevant to characterize the animals of the breed *Asturiana de los Valles* as meat producers. Thus, to describe life animals we chose 5 lengths in centimeters (see Fig. 1) in addition to the assessment of the round profile (RP) in a scale from 1 to 5 points (Alonso et al., 2008). To this six variables we added the sum of L2 and L3 that provides a good representation of the length of the animals and it is useful for predictions independently of its components.

Table 1
Summary of the features of the dataset used in the paper.

Size	390
Animals	144
Ages (months)	[7,14]
Days before slaughter	
Range	[2,222]
Average	74
Carcass weight (kg)	
Range	[116,459]
Average	221.6

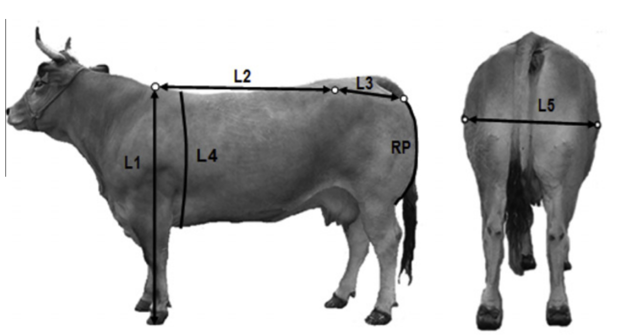

Fig. 1. Zoometric measurements used in this paper: L1 (withers height), L2 (loin length), L3 (rump length), L4 (chest girth), L5 (thighs width), and RP (round profile) assessment.

The zoometric measurements were taken by the first author of this paper using a Lydtin stick and measuring tape in different farms in Asturias. The data about carcass weight was provided by *Xata Roxa*, the company that sells the carcasses.

The assessment of the round profile is very important in animals of *Asturiana de los Valles* since it indicates the muscular developments of the animals that is correlated with the economic value of the carcasses. The assessment 5 (the best) to 1 (the worst) are given by experts by visual appreciation; see Fig. 2.

Since the aim is to predict the carcass weight in a given date, we have to include in the set of predictive variables the difference in days from the date of the measurements and the slaughter. This variable is DB (for *days before*). Finally we include the sex of the animals (codified by 1 male and 2 female) to form a set of nine predictive variables: six lengths, RP, DB, and sex.

2.3. Support Vector Regression (SVR)

The foundations of Support Vector Machines (SVMs) have been developed by Vapnik (1998) and are widely used due to many attractive features and promising empirical performance. SVMs were developed to solve classification tasks, and then they have been extended to handle regression tasks (Smola, 1996; Vapnik, 1998), in this case these algorithms are called Support Vector Regression (SVR).

The formal presentation of SVR starts with a dataset

$$S = \{(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_n, y_n)\}$$

consisting of instances described by pairs $(\boldsymbol{x}_i, y_i)$, where $\boldsymbol{x}_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$. Each y_i is the desired target or output value for the input vector $\boldsymbol{x}_i$. A regression model is learned from these patterns and used to predict the target values of unseen input vectors.

Among the various types of SVR, the most commonly used is ε-SVR (Smola, 1996; Vapnik, 1998). The goal is to find a function $f(\boldsymbol{x})$ that has at most ε deviation from the actually obtained targets y_i for all the training data, and at the same time is as flat as possible. In other words, we do not care about errors as long as they are inside the ε-insensitive band (ε-tube). See Fig. 3.

Moreover, to make the learning method more robust, the image of the input data does not need to lie strictly on or inside the ε-tube. Instead, the images which lie outside the ε-tube are penalized and slack variables are introduced to take into account for these situations (analogously to the *soft margin* in SVM for classification). The objective function and constraints are typically given as follows.

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\langle \boldsymbol{w}, \boldsymbol{w}\rangle + C\sum_{i=1}^{n}(\xi_i + \xi_i^*), \\ \text{subject to} \quad & (\langle \boldsymbol{w}, \phi(\boldsymbol{x}_i)\rangle + b) - \boldsymbol{y}_i \leqslant \varepsilon + \xi_i, \\ & \boldsymbol{y}_i - (\langle \boldsymbol{w}, \phi(\boldsymbol{x}_i)\rangle + b) \leqslant \varepsilon + \xi_i^*, \\ & \xi_i, \xi_i^* \geqslant 0,\ i = 1, \ldots, n, \end{aligned} \tag{1}$$

In these equations, C is a parameter which gives a tradeoff between model complexity and training error, ξ_i and ξ_i^* are slack variables for exceeding the target value by more than ε and for being below the target value by more than ε, respectively. As was mentioned before, this corresponds to dealing with a so-called ε-insensitive loss function $-\xi-_\varepsilon$ described by

$$|\xi|_\varepsilon = \begin{cases} 0 & \text{if} |\xi| \leqslant \varepsilon \\ |\xi| - \varepsilon & \text{otherwise.} \end{cases} \tag{2}$$

Note that

$$\phi : \mathbb{R}^d \to \mathcal{F},$$

can be a nonlinear function from the input space $\mathbb{R}^d$ to a feature space $\mathcal{F}$. The regression hyperplane to be derived is

Fig. 2. Examples of round profile (RP) in beef cattle of *Asturiana de los Valles*. The leftmost cow is a paradigm of the animals which have rank 2, while the following are representative examples of ranks 3, 4 and 5 respectively; rank 1 is extremely unfrequent.

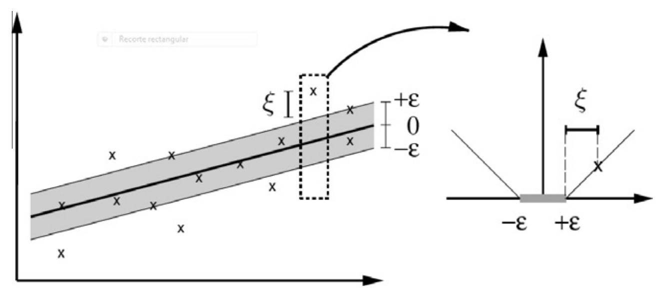

Fig. 3. Geometrical interpretation of the ε-tube and the ε-insensitive loss function in linear SVR.

$$f(\boldsymbol{x}) = \langle \boldsymbol{w}, \phi(\boldsymbol{x}) \rangle + b, \tag{3}$$

where $\boldsymbol{w}$ is the weight vector of inputs, and b is the offset.

In many problems the relation between outputs and input components is nonlinear and then kernel functions are needed. The idea of the kernel function is to enable operations to be performed in the input space rather than in the potentially high dimensional feature space. An inner product in the feature space has an equivalent kernel in the input space,

$$K(\boldsymbol{x}, \hat{\boldsymbol{x}}) = \langle \phi(\boldsymbol{x}), \phi(\hat{\boldsymbol{x}}) \rangle.$$

One of the most widely adopted kernel function is the *radial basis function* (*RBF*) which is defined as

$$K_R(\boldsymbol{x}, \hat{\boldsymbol{x}}) = e^{(-\gamma \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2)}, \ \gamma > 0, \tag{4}$$

where γ is the *width* parameter of the kernel.

We also use the *linear* kernel because the solutions provided by the system in this case are easier to explain to the end user. The weights of the attributes of the examples are a good indication of the relevance of the morphological feature behind each attribute. This kernel corresponds to the identity transformation from input to feature spaces, and then it is given by the inner product in the Euclidean space $\mathbb{R}^d$.

$$K_L(\boldsymbol{x}, \hat{\boldsymbol{x}}) = \langle \boldsymbol{x}, \hat{\boldsymbol{x}} \rangle = \sum_{i=1}^{d} \boldsymbol{x}_i \hat{\boldsymbol{x}}_i. \tag{5}$$

2.4. Learning carcass weights

To learn carcass weights using SVR we tested linear and RBF kernels. In the following, we call the regressors so obtained SVR_L y SVR_R, respectively. In both cases, we used the implementation *libsvm* (Chang and Lin, 2011).

To assess the performance of the predictions, we used absolute differences. So, if

$$S' = \{(\boldsymbol{x}'_1, y'_1), \ldots, (\boldsymbol{x}'_m, y'_m)\},$$

is a testing dataset, the performance of a regressor f will be measured by *MAE* (mean absolute error) and *MAPE* (mean absolute percentage error) defined as follows.

$$MAE(S', f) = \frac{1}{m} \sum_{i=1}^{m} |f(\boldsymbol{x}'_i) - y'_i|, \tag{6}$$

$$MAPE(S', f) = \frac{100}{m} \sum_{i=1}^{m} \left| \frac{f(\boldsymbol{x}'_i) - y'_i}{y'_i} \right|.$$

Moreover, assuming a normal distribution of errors, if μ is the mean and σ is its standard deviation, we have that for an error e,

$$\Pr(\mu - 1,64 * \sigma \leqslant e \leqslant \mu + 1.64 * \sigma) \approx 90\%.$$

Thus, 95% of the errors are below the upper bound of this confidence interval (95th percentile):

$$\Pr(e \leqslant \mu + 1.64 * \sigma) \approx 95\%.$$

Each animal is described by nine numerical attributes explained in the last section. However, in previous papers like (Alonso et al., 2006), we found that these description of the animals can be considerably improved if we add the powers of lengths untill degree 3. For this reason, with the linear kernel we used these powers in addition to the original lengths (L1, L2, L3, L2+L3, L4 and L5). Therefore, we have 18 attributes plus the RP, sex and DB. That is we may use 21 attributes for each animal.

The algorithm SVR need to adjust some parameters to achieve optimal performance. We did the adjust without using in any way the data reserved to test the performance with MAE or MAPE (Eq. (6)). We used a procedure called *internal grid search* for the best parameters with a 2-fold cross validation repeated five times. When using the liner kernel, the only value to adjust is the regularization parameter C (Eq. (1)). We searched $C \in \{10^i: i = -3, -2, -1, 0, 1, 2, 3\}$.

On the other hand, an RBF kernel has two parameters to adjust: C and γ (Eq. (4)). In this case, we tested $C \in \{10^i: i = -3, -2, -1, 0, 1, 2, 3\}$ and $\gamma \in \{10^j: j = -3, -2, -1, 0, 1, 2\}$.

To estimate the prediction errors we performed a 10-fold cross validation. The errors reported in tables are average values throughout the folds.

3. Results and discussion

In this section we report the results of a set of experiments designed to evaluate the approach proposed in this paper. The

Table 2
Cross validation estimation of prediction errors achieved with different kernels and number of attributes. We report the average, standard deviation, and 95th percentile.

System	#Attributes	MAE (kg)		MAPE (%)	
		Mean ± sd	P95th	Mean ± sd	P95th
SVR_R	9	9.31 ± 8.00	22.43	4.12 ± 3.13	9.26
SVR_L	9	10.98 ± 11.74	30.23	4.91 ± 4.68	12.59
SVR_L	21	9.61 ± 7.90	22.56	4.31 ± 3.21	9.58

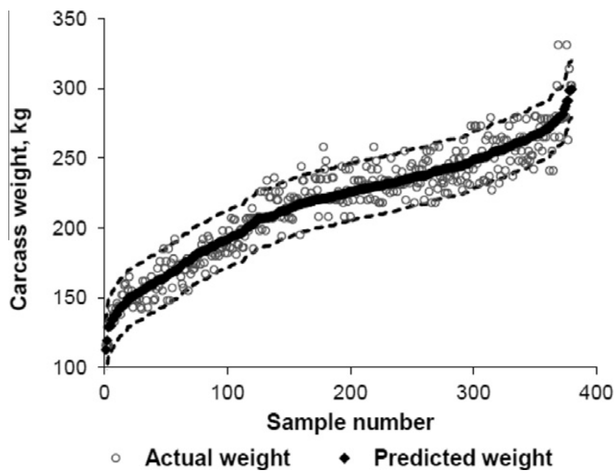

Fig. 4. Carcass weight predictions using SVR with linear kernel (SVR_L) and 21 attributes including the number of days of difference from measurements and slaughter. The horizontal axis represents the sample number; samples were ordered according to their predictions. In solid diamonds are the predictions and in empty circles are true values. Additionally, in dash lines are represented the confidence intervals at 90%.

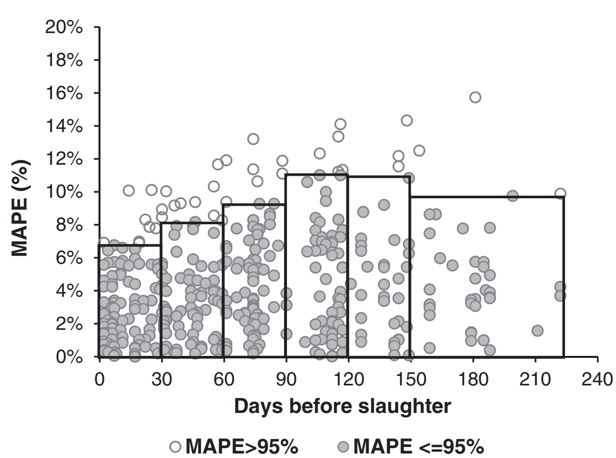

Fig. 5. MAPE errors, obtained by SVR_L with 21 attributes, depending on the number of days from measurements to slaughter. We separated months until 150 days. The rectangles include 95% of predictions.

main objective is to check the accuracy of the predictions of carcass weights using a SVR algorithm.

Table 2 reports the errors estimated with 10-fold cross validation using different options for the regressor. The results are quite similar for kernels RBF and linear, in this case when 21 attributes describe animals. We tested that there are no significant differences in these cases using a Wilcoxon signed-rank test with $p < 0.01$. In both cases the relative error in percentage (MAPE) is near 4%; thus predictions are quite accurate. See Fig. 4. Notice that the 95th percentile is just above 22 kg of error; that is near 9% of the carcass weight.

Table 3
MAPE errors obtained by SVR_L with 21 attributes including days from measurements to slaughter. We report averages, standard deviation and 95th percentiles.

Days before slaughter	MAPE (%)	
	Mean ± sd	P95th
30	3.33 ± 2.54	7.50
60	3.53 ± 2.69	7.95
90	3.91 ± 2.90	8.67
120	4.17 ± 3.13	9.31
150	4.27 ± 3.21	9.53
+150	4.31 ± 3.21	9.58

Table 4
Prediction function obtained by SVR_L with 21 attributes and days from measurements to slaughter, sex and round profile. It is a linear function with 21 variables and the intercept term reported in the last row.

Var.	Coef.	Var.	Coef.	Var.	Coef.
L1	−1.003E+00	$(L1)^2$	8.649E−04	$(L1)^3$	4.185E−05
L2	6.687E−02	$(L2)^2$	−1.074E−03	$(L2)^3$	9.802E−06
L3	−9.574E−01	$(L3)^2$	3.270E−03	$(L3)^3$	4.473E−04
L2+L3	−8.299E−02	$(L2+L3)^2$	8.177E−05	$(L2+L3)^3$	1.127E−05
L4	−4.360E−01	$(L4)^2$	1.711E−03	$(L4)^3$	2.379E−05
L5	−1.068E+00	$(L5)^2$	5.970E−03	$(L5)^3$	4.787E−04
Sex	−1.843E+01	DB	7.339E−01	RP	9.359E+00
Intercept	1.528E+02				

The results are worse when the linear kernel uses only the nine original attributes. The average absolute error is almost 11 kg (MAE) and the relative error is near 5% (MAPE).

The scores shown in Table 2 confirm that the carcass weight is a nonlinear function of the input measurements and variables included in the original set detailed is Section 2.2; see Fig. 1. Recall that a linear function over the powers up to degree 3 is, in fact, a polynomial function.

Given that we estimate carcass weight with measures taken a long time prior to the slaughter, we think that the errors achieved could be related with this value. In order to check this, we gathered the predictions obtained in cross validation for the whole dataset using SVR_L and 21 attributes. We represented the errors and the number of days prior to slaughter from the measurements in Fig. 5. We split the data by months until 150 days, and then the rest of data. The rectangles in the figure represent the 95th percentiles. Thus, we left out the rectangles the 5% of the predictions with highest absolute errors.

To detail the information reported in Fig. 5, we show Table 3.

We observe in Fig. 5 that if the slaughter is no later than 60 days from the measurements, the errors in predictions are below 10% not only in 95% of the cases, but in all cases. Only a few cases give rise an error above 10%, and always are predictions taken with more than 2 months in advance.

The preceding scores are estimations about the performance on unseen cases of a prediction function computed using all the data. Notice that cross validation is a procedure to estimate these scores that uses different splits of the whole dataset in training and testing sets. The prediction function obtained in this case by SVR_L with 21 attributes is a linear function of 21 variables. The coefficients of each variable and an intercept term are reported in Table 4.

4. Conclusions

We have presented a method to estimate the weight of carcass from zoometric measurements taken months before the slaughter date. For this purpose we used an Artificial Intelligence tool, the Support Vector Machines for Regression (SVR). We discussed the use of different options for configuring these learning algorithms.

The paper includes a case study carried with animals of a specialized beef breed, *Asturiana de los Valles*. The results show that using a nonlinear function it is possible to achieve accurate predictions a long time prior the slaughter. However, the errors are smaller if the measures of the animals are taken near the slaughter.

Acknowledgements

The research reported in this paper is supported in part under the Grant TIN2011-23558 from the MICINN (Ministerio de Ciencia e Innovación, Spain). The authors would like to acknowledge the collaboration of the experts of the Association of Breeders (ASEAV-A) during the acquisition of data stage; worthy of special mention was the help supplied by Amable Fernández González, breeder from Sobrescobio, Asturias.

References

Alonso, J., Bahamonde, A., Villa, A., Alonso, L., 2006. Estimación del peso corporal en terneros de la raza asturiana de los valles usando medidas zoométricas. FEAGAS 29, 49–54.

Alonso, J., Bahamonde, A., Villa, A., Castañón, Á.R., 2007. Morphological assessment of beef cattle according to carcass value. Livestock Science 107, 265–273.

Alonso, J., del Coz, J.J., Díez, J., Luaces, O., Bahamonde, A., 2008. Learning to predict one or more ranks in ordinal regression tasks. In: Walter Daelemans, Bart Goethals, Morik, Katharina (Eds.), Proceedings of the 2008 European Conference on Machine Learning and Knowledge Discovery in Databases-Part I (ECML PKDD '08), LNAI. Springer-Verlag, Berlin, Heidelberg, pp. 39–54.

Chang, C.-C., Lin, C.-J., 2011. LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2, 27:1–27:27.

del Coz, J.J., Bayón, G.F., Díez, J., Luaces, O., Bahamonde, A., Sañudo, C., 2005. Trait selection for assessing beef meat quality using non-linear SVM. In: Saul, Lawrence, Weiss, Yair, Bottou, Léon (Eds.), Proceedings of the Neural Information Processing Systems Conference (NIPS'05), Advances in Neural Information Processing Systems, vol. 17. MIT Press, Cambridge, MA, pp. 321–328.

Díez, J., del Coz, J.J., Sañudo, C., Albertí, P., Bahamonde, A., 2005. A kernel based method for discovering market segments in beef meat. In: Jorge, Alípio Mário, Torgo, Luís, Camacho, Rui, Brazdil, Pavel, Gama, João (Eds.), Proceedings of the 9th European Conference on Principles and Practice of Knowledge Discovery in Databases (PKDD'05), LNAI. Springer-Verlag, Berlin, Heidelberg, pp. 462–469.

Enevoldsen, C., Kristensen, T., 1997. Estimation of body weight from body size measurements and body condition scores in dairy cows. Journal of Dairy Science 80 (9), 1988–1995.

González-Velasco, H., García-Orellana, C., Macías-Macías, M., Gallardo-Caballero, R., García-Manso, A., 2011. A morphological assessment system for show quality bovine livestock based on image analysis. Computers and Electronics in Agriculture 78 (1), 80–87.

Martínez, A., García, J., Noval, G., Alonso, L., Castro, P., Osoro, K., 1999. Efecto del grado de presencia del gen de la hipertrofia muscular en los rendimientos productivos del cebo intensivo y las características de la canal-culón. ITEA 20 (1), 89–91.

Piedrafita, J., Quintanilla, R., Sañudo, C., Olleta, J., Campo, M., Panea, B., Renand, G., Turin, F., Jabet, S., Osoro, K., Oliván, M.C., Noval, G., García, P., García, M., Cruz-Sagredo, R., Oliver, M., Gispert, M., Serra, X., Espejo, M., García, S., López, M., Izquierdo, M., 2003. Carcass quality of 10 beef cattle breeds of the Southwest of Europe in their typical production systems. Livestock Production Science 82 (1), 1–13.

Smola, A.J., 1996. Regression estimation with support vector learning machines. Master's thesis, Technische Universität München.

Stajnko, D., Brus, M., Hocevar, M., 2008. Estimation of bull live weight through thermographically measured body dimensions. Computers and Electronics in Agriculture 61 (2), 233–240.

Tasdemir, S., Urkmez, A., Inal, S., 2011a. Determination of body measurements on the holstein cows using digital image analysis and estimation of live weight with regression analysis. Computers and Electronics in Agriculture 76, 189–197.

Tasdemir, S., Urkmez, A., Inal, S., 2011b. A fuzzy rule-based system for predicting the live weight of holstein cows whose body dimensions were determined by image analysis. Turkish Journal of Electrical Engineering and Computer Sciences 19 (4), 689–703.

Vapnik, V.N., 1998. Statistical Learning Theory. Wiley-Interscience, New York.