

Improved estimation of bovine weight trajectories using Support Vector Machine classification

Jaime Alonso^{(1)*}, Alfonso Villa⁽²⁾, Antonio Bahamonde⁽¹⁾

⁽¹⁾*Artificial Intelligence Center
University of Oviedo at Gijón, Asturias, Spain*

⁽²⁾*Association of Breeders of Asturiana de los Valles (ASEAVA)
Abarrio, N° 24, E-33424 Llanera, Asturias, Spain*

Abstract

The benefits of livestock breeders are usually closely related to the weight of their animals. In this paper we present a method to anticipate the weight of each animal provided we know the past evolution of the herd. Our approach exploits the geometrical relationships of the trajectories of weights along the time. Starting from a collection of data from a set of animals, we learn a family of parallel functions that fits the whole data set, instead of having one regression function for each individual. In this way, our method enables animals with only one or a few weights to have an accurate estimation of their future evolution. Thus, we learn a function F defined on the space of weights and time that separates the trajectories in such a way that F has constant values on each trajectory. The key point is that the specification of F can be done in terms of ordering constraints, in the same way as preference functions or ordinal regressors. Therefore, F can be obtained from a classification SVM (Support Vector Machines). To evaluate the method, we have used a collection of real world data sets of bovines of different breeds and ages. We will show that our method outperforms the separate regression of each animal when there are only a few weights available and we need medium or long term predictions.

*Corresponding author: Tel: +34 985 182 032

Email addresses: jalonso@aic.uniovi.es (Jaime Alonso⁽¹⁾), avilla@aseava.com (Alfonso Villa⁽²⁾), antonio@aic.uniovi.es (Antonio Bahamonde⁽¹⁾)

Keywords: Support Vector Machines (SVM), weight prediction, growth curves, beef cattle.

1. Introduction

The estimation of the weight of bovines, and other livestock, as a function of time is very important for breeders. The efficiency of growth determines the economic benefits not only in beef cattle, but also in dairy bovines (Arango and Van Vleck, 2002).

There are different approaches to study weight evolutions in the literature. In (West et al, 2001), the biological fundamental principles of the growth of many diverse species are emphasized. They show that it is possible to fit the same sigmoid for cows, pigs, guppies or shrimps, provided that it is performed a transformation of time and weight data to a common *dimensionless* scale. 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 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 (Stajanko et al., 2008; Tasdemir et al., 2011a,b) the authors used digital image processing procedures.

A different point of view is used when the objective is the selection of stud bulls. In this case, the bulls are housed together in central evaluation stations; they stay there for several months, depending of the management policies of breeds and beef markets requirements. During that time each bull is weighted approximately every 15 or 30 days; at the end of the testing period, bulls are ranked according to their genetic merits. The methods used to compute this ranking include the use of a pedigree tree in order to consider not only the observed rate of weight gain, but mainly the capability of bulls to transmit genetically that gain to their progeny (Meyer, 2002a,b, 2005; Schenkel et al., 2002; Freetly et al., 2011). A comparison of different genetics methods can be seen at (Jaffrézic and Pletcher, 2000).

In this paper we are concerned with finding accurate predictions of weights of beef cattle, neither for the whole specie nor for the expectations of the progeny. We want to anticipate the weight of each single animal in order to improve the incomes of breeders (Díez et al., 2003; Alonso et al., 2007, 2013), given that market prices follow a well known annual cycle.

The approach proposed in this paper exploits the geometrical relationships of the trajectories of weights along the time. Starting from a collection of data from a set of animals, we learn a family of parallel functions that fits the whole data set, instead of having one regression function for each individual. In this way, our method enables animals with only one or a few weights to have an accurate estimation of their future evolution. To learn this family of functions, in the next section, we describe how to build a classification data set. The idea is to deal with an ordering of the trajectories in the same way used to learn ordinal regression (Herbrich et al., 2000; Herbrich, 2002; Alonso et al., 2008) or to learn preferences (Joachims, 2002; Bahamonde et al., 2004); see also (Schultz and Joachims, 2003).

To evaluate the performance of the method, we have used real world data sets of bulls and cows of different breeds and ages. Of special interest are the results achieved with young bulls of *Asturiana de los Valles*, a beef breed of the North of Spain. We will show that, in all cases, our method outperforms the individual regression of each animal, in special when there are only a few weights available and we need reliable medium or long term (more than 100 days) predictions.

2. Material and methods

2.1. Data

In the research reported in this paper we used several real world data sets of bovines of different breeds and ages. A data set of 351 Angus bulls from the Indiana Beef Evaluation Program (IBEP). All animals have 6 weights recorded taken at ages 180 to 509 days.

A second data set of 822 bulls of *Asturiana de los Valles* with more than 5 weights after the adjustment period. They have from 5 to 11 weights (after the adjustment period), but most of them (557 animals) have 6. The ages of the animals range from 241 to 508 days.

Finally, a third data set obtained from Meyer’s program to compute the relative genetic merits (Meyer, 2002a). The data includes the weights of a group of cows from the Wokalup Selection Experiment in Western Australia. We only considered the records of the animals (358) with at least 2 weights; the maximum number of weights is 6 recorded between 19 and 82 months.

2.2. The method of weight trajectories

Let us consider a set B of bovines. For each animal we have recorded a number of weights. The whole data available can be represented by a set of pairs (w_{ai}, d_{ai}) where w_{ai} is the i^{th} weight of the animal a taken when it was d_{ai} days old.

A first attempt to generalize these data consists in computing one regression function f_a for each animal trying to optimize the loss produced by the differences

$$w_{an} - f_a(d_{an}), \forall a \in B \tag{1}$$

Provided that for each animal $a \in B$ we want to predict its weight w_{an} at day d_{an} .

However, if we want to be able to estimate the weight of a new animal (not in B), we have two options. If we have obtained reasonably good approximations with the family $(f_a : a \in B)$, and these family of functions belong to the same class, say linear for instance, then we can try to induce a new regressor of the same class for the new animal. Unfortunately, faithful regressors need a large collection of observations, and in this case this requires time, what it is contradictory with our intention of anticipating the weight in future days. Additionally, the acquisition of data is costly and in general a risky task both

for people and for the animals that become stressed during weighting sessions (Goyache et al., 2001).

The second option consists in trying to generalize somehow the set $(f_a : a \in B)$ to obtain a kind of universal function able to estimate the weight of any animal at any time. It is not clear how to do this; and, on the other hand, it is necessary to consider individual differences: in some cases they are really important. But if we were able to devise a method to take advantage of all pairs (w_{ai}, d_{ai}) at the same time, we can wait an improved generalization performance, since the number of such pairs in practice is usually high.

The method that we are proposing in this paper emphasizes the role of the trajectories followed in the space of weights and days ($W \times D$) by the successive weights of the animals. Thus, we first look for a function $F : W \times D \rightarrow \mathfrak{R}$ able to separate the trajectories considered as subsets of $W \times D$. In other words, a function F such that

$$F(w, d) = \text{constant} \tag{2}$$

defines implicitly all reasonable trajectory for the weights of the animal (see Fig. 1).

Then, given a set of measurements of an animal a , $\{(w_{ak}, d_{ak}) : k = 1, \dots, n-1\}$, we define the constant for its trajectory as the average of F values obtained on each point. In symbols,

$$\text{constant}_a = \frac{1}{n-1} \sum_{k=1}^{n-1} F(w_{ak}, d_{ak}) \tag{3}$$

Therefore, to estimate w_{an} , the weight of the animal at day d_{an} , we only need to obtain from (2) the explicit version of the function relating w and d .

Returning to the optimization problem of (1), to express it in terms of F , we need to constraint the form of this function. So, let us assume that $F(w, d)$ is a linear combination of the first argument and a general function of d , say $g(d)$; that is,

$$F(w, d) = \alpha w + g(d), \alpha \in \mathfrak{R} \tag{4}$$

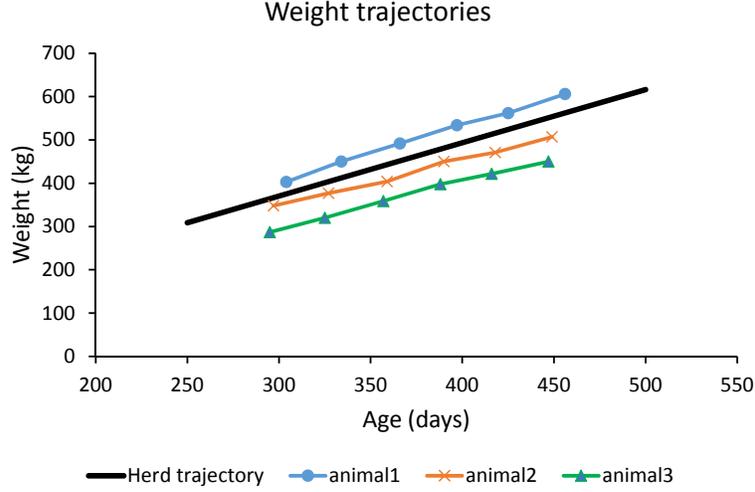


Figure 1: Herd trajectory found for *Asturiana de los Valles*; comparison with three sample animals. Trajectories of each animal can be obtained from herd trajectory and a constant value.

Then, we have that the explicit version of the function that predicts the weights of the animal a , f_a , is given by

$$f_a(d) = \frac{1}{\alpha} \left(\frac{1}{n-1} \sum_{k=1}^{n-1} F(w_{ak}, d_{ak}) - g(d) \right) \quad (5)$$

Therefore,

$$\begin{aligned} w_{an} - f_a(d_{an}) &= w_{an} - \frac{1}{\alpha} \left(\frac{1}{n-1} \sum_{k=1}^{n-1} F(w_{ak}, d_{ak}) - g(d_{an}) \right) = \\ &= \frac{1}{\alpha} \left(F(w_{an}, d_{an}) - \frac{1}{n-1} \sum_{k=1}^{n-1} F(w_{ak}, d_{ak}) \right) \end{aligned} \quad (6)$$

Hence, the deviation in weight predictions of a given animal is proportional to the deviations of F values on the weight-day points of the weight trajectory of that animal. In other words, the minimization of (1) can be achieved if we

are able to ensure the uniformity of F values on the weight trajectories of the animals. For this purpose we propose a heuristic approach. Since we do not know the right value of $F(w, d)$ for any (w, d) , we are going to ask F to fulfill a set of constraints. Let us consider two animals a and b weighted the same sequence of days ($d_i = 1, \dots, n - 1$), then we say that the *trajectory* of a is *higher* than that of b if and only if

$$w_{ai} > w_{bi}, \forall i = 1, \dots, n - 1 \quad (7)$$

In this case we ask F to fulfill the constraints:

$$F(w_{ai}, d_i) > F(w_{bj}, d_j), \forall i, j = 1, \dots, n - 1 \quad (8)$$

If the dates of weights are not exactly the same, we can approximate the values by means of interpolation, and then we can decide if the trajectories of two animals have any comparative relationship or not. Notice that interweaving trajectories will not yield to any constrain like those in equation (8).

2.3. The classification data set entailed by a set of relative comparisons

In this section we are going to see how a set of relative comparisons, as those of equation (8), give rise to a classification data set. The idea is that a classifier induced by a *SVM* provides a function F that satisfies as many constraints (expressed as relative comparisons) as possible. The approach followed is not new, see for instance (Schultz and Joachims, 2003); sometimes this method was used to learn ordinal regression (Herbrich et al., 2000) or to learn preferences (Joachims, 2002; Bahamonde et al., 2004).

For ease of reference, let us consider a finite set R of relative comparisons of F values given by

$$F(w_{r1}, d_{r1}) > F(w_{r2}, d_{r2}), \forall r \in R \quad (9)$$

If we assume that F is a linear function, there is a vector $\vec{c} = (c_1, c_2)$ such that $F(w, d) = \vec{c} \cdot (w, d) = c_1 w + c_2 d$. And then, the comparisons can be expressed as

$$F(w_{r1} - w_{r2}, d_{r1} - d_{r2}) > 0, \forall r \in R \quad (10)$$

or equivalently,

$$F(w_{r1} - w_{r2}, d_{r1} - d_{r2}) > 0, F(w_{r2} - w_{r1}, d_{r2} - d_{r1}) < 0, \forall r \in R \quad (11)$$

Therefore, from a geometrical point of view, the function F represents the distance to the hyperplane (of vectors perpendicular to \vec{c}) multiplied by the norm of \vec{c} . The search of \vec{c} is a NP-hard problem; however, it is possible to approximate the solution like in classification Support Vector Machines (Vapnik, 1998). More formally, the search for a reliable function linear F satisfying (9) can be expressed as an optimization problem for margin maximization. We must

$$\begin{aligned} \text{minimize} \quad & V(\vec{c}, \xi) = \frac{1}{2} \|\vec{c}\|^2 + C \sum_{r \in R} \xi_r \quad (12) \\ \text{subject to} \quad & \forall r \in R, \vec{c} \cdot (w_{r1} - w_{r2}, d_{r1} - d_{r2}) > 1 + \xi_r \\ & \forall r \in R, \xi_r \geq 0 \end{aligned}$$

where C is a parameter that allows trading-off margin size against training error.

From a practical point of view, the classification data set is built with the positive differences (vectors of the form $(w_{r1} - w_{r2}, d_{r1} - d_{r2})$ and class +1), and the negative differences (vectors of the form $(w_{r2} - w_{r1}, d_{r2} - d_{r1})$ and class -1). Table 1 shows a sample of how to generate instances of a data set from the measurements of an animal. The hyperplane that separates these two classes defines the function F subject to the relative comparisons of (9).

3. Experimental results

We have conducted a number of experiments to evaluate our method with real world data sets. For this purpose, given a set of weight-days trajectories, we compared the predictions of our method with those found after fitting, for each animal, one regression function. We are assuming that we only knew a reduced set of weights. So, we consider only the first 2 weights of each animal, or the first 3 or 4 and finally we consider the performance in the assumption

Table 1: Example of how to build instances of a data set from pairs (weight, age) of measurements of an animal

Measurements		Instances		
Weight(Kg)	Age(days)	Weight ₁ - Weight ₂	Age ₁ - Age ₂	Class
255	213	468-337	397-304	+1
337	304	468-255	397-213	+1
468	397	337-255	304-213	+1
		337-468	304-397	-1
		255-468	213-397	-1
		255-337	213-304	-1

that we knew all the weights but the last one. Weights are taken in ascending order of the animals' ages. In all cases the target weight is the last weight recorded for each animal. The idea is to discuss the differences in performance of the proposed method and the alternative of considering a regression for each animal separately. We show that for a few measurements of the animals, the proposed method outperforms the separate regression while this approach is better when we have a lot of weights of each animal.

3.1. Experiment settings

The scores shown in Tables 2, 3 and 5 were obtained separating the available set of animals in a train set (75%) and a test set (25%). We used the test sets to assess the performance of the predictions. So if

$$S = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$$

is a testing dataset, the performance of a function f will be measured by $MAPE$ (mean absolute percentage error) and $RMSE$ (root mean square error) defined as follows:

$$MAPE(S, f) = \frac{100}{n} \sum_{i=1}^n \frac{|f(\mathbf{x}_i) - y_i|}{y_i} \quad (13)$$

Table 2: Comparison scores obtained with data of 351 Angus bulls from Performance Summaries of the Indiana Beef Evaluation Program. We compare the final weight of each animal with the weight predicted if we knew only the first weights recorded. Weights are taken in ascending order of the animals’ ages.

#weights	SVM		Linear Regression		Adv. days ^b
	MAPE(%) ^a	RMSE(Kg)	MAPE(%)	RMSE(Kg)	
1	3.9 ± 3.0	61.4	n/a	n/a	147
2	3.6 ± 2.7	57.1	8.3 ± 6.1	136.3	125
3	3.3 ± 2.5	52.9	4.6 ± 3.7	76.2	97
4	3.1 ± 2.2	49.0	3.8 ± 2.6	59.0	62
n-1 ^c	2.8 ± 1.9	43.7	2.7 ± 1.8	41.4	34

^a Mean±std (standard deviation)

^b Average number of days in advance to last weight

^c All weights but the last one

$$RMSE(S, f) = \sqrt{\frac{1}{n} \sum_{i=1}^n (f(\mathbf{x}_i) - y_i)^2} \quad (14)$$

During the experiments, we used SVM^{light} (Joachims, 1999) with the default parameters and a linear kernel. The classifying training set was formed comparing the trajectories of each animal with other trajectories randomly selected until we obtained 10 animals with higher or lower trajectories. The number 10 is not critical at all; we chose 10 to obtain manageable sizes of training set, but the results were almost the same when we changed this parameter.

In all cases discussed here, the expression for the function g of the days (see equation (4)) was linear; we tested polynomial expressions, but the results were worse. The reason is that bovines below 18 months follow a linear weight trajectory, what is the case of the animals involved in the experiments reported in Tables 2 to 4. Notice that due to beef market requirements, bulls are usually slaughtered before they reach the age of 18 months; therefore, it is difficult

Table 3: Comparison scores obtained with data of 776 bulls from the evaluation station of the Association of Breeders of *Asturiana de los Valles* (ASEAVA). We compare the final weight of each animal with the weight predicted if we knew only the first weights recorded. Weights are taken in ascending order of the animals' ages.

#weights	SVM		Linear Regression		Adv. days ^b
	MAPE(%) ^a	RMSE(Kg)	MAPE(%)	RMSE(Kg)	
1	5.3 ± 4.4	37.0	n/a	n/a	141
2	4.9 ± 4.1	34.0	9.2 ± 9.6	37.7	115
3	4.5 ± 3.7	30.1	5.2 ± 4.8	24.9	89
4	4.0 ± 3.3	26.0	3.6 ± 3.0	18.2	61
n-1 ^c	3.4 ± 2.7	21.9	2.0 ± 1.7	13.0	18

^a Mean±std (standard deviation)

^b Average number of days in advance to last weight

^c All weights but the last one

to have data about older bulls. Thus, to deal with older bovines, they must be cows, as in the data set used to compute Table 5 ; however in this case we again obtained a linear relation between days and weights, although with a quite different slope. The ages of cows range from 19 to 82 months, but we only have available one weight per year; that is, a maximum of 6 measurements for animal; this make difficult the possibility of fitting other kind of nonlinear curves.

3.2. Results and discussion

In Table 2 we report the scores achieved with a data set of Angus bulls (351 animals) from the Indiana Beef Evaluation Program. The family of linear functions found by SVM on training set has a common slope of 3.73, while the average gain, computed for the same set, was 3.71 pounds (1.68 Kg) per day.

The scores obtained with *Asturiana de los Valles* have being reported in Tables 3 and 4. We have separated two sets of data. In the first one we included

Table 4: Comparison scores obtained with a control set of 46 bulls from the evaluation station of the Association of Breeders of *Asturiana de los Valles* (ASEAVA). We compare the final weight of each animal with the weight predicted if we knew only the first weights recorded. Weights are taken in ascending order of the animals' ages.

#weights	SVM		Linear Regression		Adv.
	MAPE(%) ^a	RMSE(Kg)	MAPE(%)	RMSE(Kg)	days ^b
1	5.0 ± 3.7	32.1	n/a	n/a	128
2	4.9 ± 3.8	32.5	11.8 ± 7.6	71.2	117
3	4.9 ± 4.0	33.0	12.6 ± 11.0	79.5	104
4	4.9 ± 3.9	32.4	7.9 ± 6.7	49.2	89
5	4.7 ± 3.8	31.5	6.4 ± 5.8	40.8	76
6	4.5 ± 3.6	30.0	5.0 ± 4.1	30.9	60
7	4.2 ± 3.4	28.2	3.6 ± 3.1	23.3	47
n-1 ^c	3.9 ± 3.3	26.8	3.3 ± 3.1	22.1	15

^a Mean±std (standard deviation)

^b Average number of days in advance to last weight

^c All weights but the last one

bulls from the initial date until November of last year; this sums 776 animals. They have from 5 to 11 weights (after the adjustment period), but most of them (557 animals) have 6. Table 3 collects the scores achieved with these data separating a train set of 75% and a test set of 25%. The second set, with 46 bulls born after November, was used as control set to test if what was learned with bulls of previous generations can be extended to future generations. The results are reported in Table 4, and exhibit the same kind of scores of Tables 2 ad 3. The slope of the linear relation of weights and days found by our method in the train set was 1.27, while the average weight gain was 1.23 kg per day.

Finally, Table 5 shows the scores obtained with the data set provided with Meyer's program to compute the relative genetic merits (Meyer, 2002a). We

Table 5: Comparison scores obtained with 358 weights of beef cow from the Wokalup Selection Experiment in Western Australia. We compare the final weight of each animal with the weight predicted if we knew only the first weights recorded. Weights are taken in ascending order of the animals' ages.

#weights	SVM		Linear Regression		Adv.
	MAPE(%) ^a	RMSE(Kg)	MAPE(%)	RMSE(Kg)	mo. ^b
1	9.3 ± 6.7	72.4	n/a	n/a	54.1
2	9.5 ± 7.2	70.6	42.5 ± 32.9	336.1	43.4
3	8.6 ± 6.0	62.1	16.7 ± 13.6	136.1	32
n-1 ^c	8.3 ± 5.9	60.4	10.2 ± 8.5	83.2	12.8

^a Mean±std (standard deviation)

^b Average number of months in advance to last weight

^c All weights but the last one

only considered the records of the animals (358) with at least 2 weights; the maximum number of weights is 6. As was mentioned above, the slope found by SVM for weights and months (in this case time is measured in months) is 3.83, while the average gain was 4.70 kg per month. In this case, the difference of these values is higher than those found for young Angus or *Asturiana de los Valles*. Notice also the difference of slopes due to the older age of Wokalup cows.

3.3. Anticipation

To give a measure of the relevance of weight predictions, we included in all Tables (2 to 5) a column to display the difference in days (or months in the case of Table 5) from the last weight used to make the predictions and the date of the target weight. Thus, we observe that our approach outperforms linear regressions when the problem was to anticipate weights more than 100 days; the differences become lower until we have about 60 days to the target weight is farther; below this number of days, it is preferable to use linear regression, although in this case, the prediction problem is uninteresting: you can simply

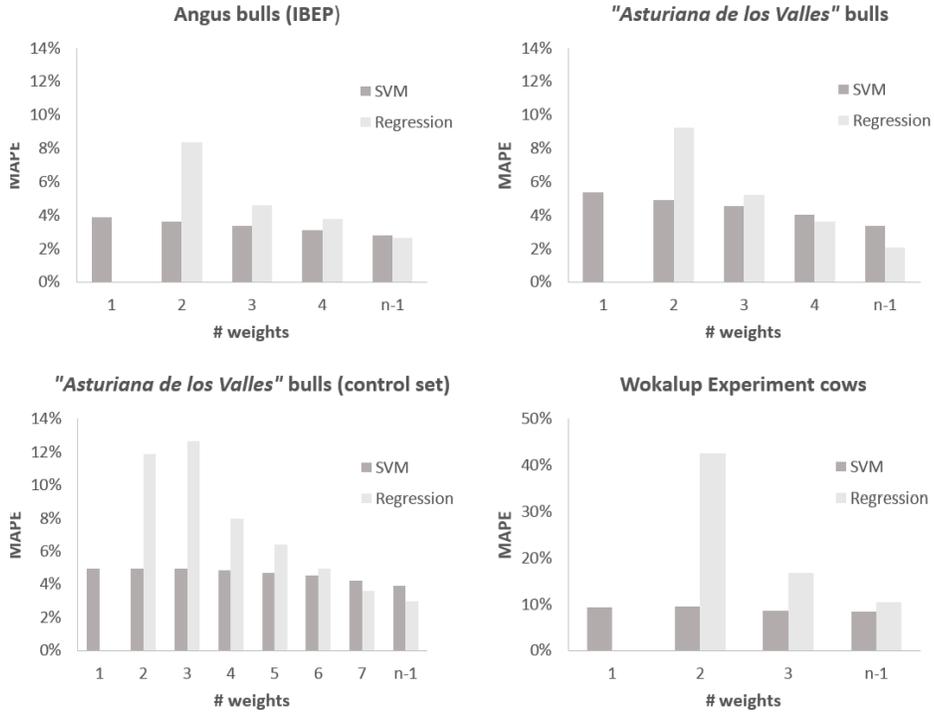


Figure 2: Comparative error charts showing the MAPE scores reported in tables 2 to 5 to compare the predictions of our method (based on SVM) with those found using linear regression.

wait and see the final weight. The case of older cows reported in Table 5 is a bit different; here our approach has lower average differences in all cases, although the margin with linear regression becomes narrower when we have to predict the weight one year before knowing the evolution of the previous years.

4. Conclusions

We have presented a method to anticipate the weights of livestock alternative to regression. Starting from a collection of weight trajectories along the time of a set of similar animals, we described a method to build a training set for a classification SVM in such a way that the hyperplane that separates the

two classes defines implicitly the family of functions that estimate the weight of each animal having available only a small number of previous weights. To evaluate the method we have carried out several experiments with real world data sets of weight trajectories of bovines of different breeds and different geographical environments. These experiments reported that our method has a higher accuracy than the individual regression of each animal, especially when there are only a few weights available and we need reliable medium or long term predictions.

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 (ASEAVA) during the acquisition of data stage. Also, the authors would like to thank: the Indiana Beef Evaluation Program (IBEP) and Karin Meyer, for making available through the web the data used; T. Joachims for his SVM^{light}.

References

- 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: *Proceedings of the 2008 European Conference on Machine Learning and Knowledge Discovery in Databases - Part I (ECML PKDD '08)*, Walter Daelemans, Bart Goethals, and Katharina Morik (Eds.). LNAI, Springer-Verlag, Berlin, Heidelberg, pp. 39–54.

- Alonso, J., Castañón, Á. R., Bahamonde, A., 2013. Support Vector Regression to predict carcass weight in beef cattle in advance of the slaughter. *Computers and Electronics in Agriculture* 91, 116–120.
- Arango, J. A., Van Vleck, D., 2002. Size of beef cows: early ideas, new developments. *Genetics and Molecular Research* 1 (1), 51–63.
- Bahamonde, A., Bayón, G. F., Díez, J., Quevedo, J. R., Luaces, O., del Coz, J. J., Alonso, J., Goyache, F., 2004. Feature subset selection for learning preferences: a case study. *Proceedings of the 21st International Conference on Machine Learning, ICML 2004, Banff, Canada*.
- Díez, J., Bahamonde, A., Alonso, J., López, S., del Coz, J. J., Quevedo, J., Ranilla, J., Luaces, O., Alvarez, I., Royo, L., Goyache, F. (2003). Artificial intelligence techniques point out differences in classification performance between light and standard bovine carcasses. *Meat Science*, 64, 249–258.
- Enevoldsen, C. and 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.
- Freetly, H. C., Kuehn, L. A. and Cundiff, L. V., 2011. Growth curves of crossbred cows sired by Hereford, Angus, Belgian Blue, Brahman, Boran, and Tuli bulls, and the fraction of mature body weight and height at puberty. *J ANIM SCI* 89 (8) 2373–2379.
- 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.
- Goyache, F., del Coz, J., Quevedo, J., López, S., Alonso, J., Ranilla, J., Luaces, O., Alvarez, I., Bahamonde, A., 2001. Using artificial intelligence to design and implement a morphological assessment system in beef cattle. *Animal Science*, 73, 49–60.

- Herbrich, R., 2002. Learning Kernel Classifiers: Theory and Algorithms. MIT Press.
- Herbrich, R., Obermayer, K., Graepel, T, 2000. Large margin rank boundaries for ordinal regression. In A. J. Smola, P. L. Barlett, B. Schölkopf, D. Schuurmans (ed), Advances in Large Margin Classifiers. MIT Press.
- IBEP: Indiana Beef Evaluation Program. <http://www.ansc.purdue.edu/ibep/index.htm>.
- Jaffrézic, F., Pletcher, S.D., 2000. Statistical Models for Estimating the Genetic Basis of Repeated Measures and Other Function-Valued Traits. *Genetics* 156, 913–922.
- Joachims, T., 1999. Making large-Scale SVM Learning Practical. In B. Schölkopf and C. Burges and A. Smola (ed.), *Advances in Kernel Methods - Support Vector Learning*. MIT-Press.
- Joachims, T., 2002. Optimizing search engines using clickthrough data. *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD)*.
- Meyer, K., 2002. Home page of RRGIBBS: A program to estimate variance components for simple random regression models using Gibbs sampling. Animal Genetics and Breeding Unit (AGBU), University of New England, Armidale, Australia. <http://www.agbu.une.edu.au/kmeyer/rrgibbs.html>.
- Meyer, K., 2002. RRGIBBS- A program for simple random regression analyses via Gibbs sampling. *Proceeding of the 7th World Congress on Genetics Applied to Livestock Production*. Montpellier, France.
- Meyer, K., 2005. Estimates of genetic covariance functions for growth of Angus cattle. *Journal of Animal Breeding and Genetics*, 122 (2), 73–85.
- Schenkel F. S., Miller, S. P., Jamrozik, J., Wilton, J. W., 2002. Two-step and random regression analyses of weight gain of station-tested beef bulls. *J. Animal Sciences* 80, 1497–1507.

- Schultz, M., Joachims, T., 2003. Learning a Distance Metric from Relative Comparisons, Proceedings of the Conference on Advance in Neural Information Processing Systems (NIPS).
- Stajniko, 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., 1998. *Statistical learning theory*. John Wiley.
- West, G. B., Brown, J. H., Enquist, B. J., 2001. A general model for ontogenetic growth. *Nature*, Vol. 413, 628–631.