RESEARCH ARTICLE

Predicting Breeding Values in Animals by Kalman Filter: Application to Body Condition Scores in Dairy Cattle [1]

Burak KARACAÖREN * ✎   Luc L.G. JANSS **   Haja N. KADARMIDEEN ***

  * Department of Animal Science, Faculty of Agriculture, University of Akdeniz, TR-07070 Antalya - TURKEY
 ** Department of Genetics and Biotechnology, Research Centre Foulum Blichers Alle 20, University of Aarhus, DENMARK
*** Section of Quantitative and System Genetics, Department of Animal and Veterinary Sciences, Faculty of Life Sciences, University of Copenhagen, DENMARK

Makale Kodu (Article Code): KVFD-2012-6107

Summary

The aim of this study was to investigate usefulness of Kalman Filter (KF) Random Walk methodology (KF-RW) for prediction of breeding values in animals. We used body condition score (BCS) from dairy cattle for illustrating use of KF-RW. BCS was measured by Swiss Holstein Breeding Association during May 2004-March 2005 for 7 times approximately at monthly intervals from dairy cows (n=80) stationed at the Chamau research farm of Eidgenössische Technische Hochschule (ETH), Switzerland. Benefits of KF were demonstrated using random walk models via simulations. Breeding values were predicted over days in milk for BCS by KF-RW. Variance components were predicted by Gibbs sampling. Locally weighted scatter plot smoothing (LOWESS) and KF-RW were compared under different longitudinal experimental designs, and results showed that KF-RW gave more reasonable estimates especially for lower smoother span of LOWESS. Estimates of variance components were found more accurate when the number of observations and number of subjects increased and increasing these quantities decreased standard errors. Fifty subjects with 10 observations each, started to give reasonable estimates. Posterior means for variance components were found (with standard errors) 0.03 (0.006) for animal genetic variance 0.04 (0.007) for permanent environmental variance and 0.21 (0.02) for error variance. Since KF gives online estimation of breeding values and does not need to store or invert matrices, this methodology could be useful in animal breeding industry for obtaining online estimation of breeding values over days in milk.

Keywords: *Kalman filter, Body condition score, Bayesian methods*

Damızlık Değerlerinin Kalman Süzgeci ile Tahmini: Süt Sığırlarından Toplanan Vücut Kondüsyon Puanları İçin Bir Uygulama

Özet

Bu çalışmanın ana amacı; rassal yürüyüş taslamını kalman süzgeci (RY-KS) ile kullanarak çiftlik hayvanlarında zamana dayalı damızlık değerlerinin tahmin edilebilmesinin uygunluğunun araştırılmasıdır. Veri seti Chamau enstitüsündeki süt ineklerinden, İsviçre Holştayn birliği tarafından elde edilmiştir (n=80). Yerel ağırlıklı uzanım tahmincisi (LOWESS) yöntemi ve RY-KS benzeşim yolu ile farklı deneme desenleri için karşılaştırıldı, sonuçlar benzer bulunsa da, bazı şartlar altında RY-KS daha uygun sonuçlar verdi. Birey ve gözlem sayılarının varyans unsurları ve üreteç tahminine olan etkileri incelendi ve her ikisinin arttırılmasının daha doğru tahminleri daha küçük standart hatalarla verdiği saptandı. Varyans unsurları gibbs örneklemesi ile tahmin edildi. Soncul ortalamalar genetik varyans için 0.03 (0.006), kalıcı çevre için 0.21 (0.02) ve hata varyansı için 0.21 (0.02) olarak hesaplandı. Kalman Süzgeci damızlık değerlerinin güncel tahminlerini vermesi ve matrislerin devrilmesine ihtiyaç göstermemesi nedeni ile büyük veri setlerinden zamana dayalı damızlık değerlerinin tahmin edilmesinde faydalı olabilir.

Anahtar sözcükler: *Kalman süzgeci, Vücut kondüsyon puanları, Bayesçi yöntemler*

✎ **İletişim (Correspondence)**

☎ +90 242 2274560

✉ burakkaracaoren@akdeniz.edu.tr

INTRODUCTION

A filter is a device that is used for example to separate water from particles. By analogy this idea was extended to separation of signals from noise [1] in engineering context. Kalman Filter (KF) was defined with Bayesian features by Harrison and Stevens [2] as Bayesian Dynamic Linear Model[3] although West and Harrison [4] quoted as "Bayesian forecasting is Kalman Filtering" is akin to saying that statistical inference is least squares". Migon et al.[5] described some special characteristics of KF as following; i) all relevant information's are used, including history, factual or subjective experiences, and knowledge of forthcoming events ii) routine forecasting is produced by a statistical model and exceptions can be considered as an anticipation or in a retrospective base iii) prospective (what happened) and retrospective (what if) analysis are easily accommodated iv) model decomposition, a full Bayesian forecasting model may be decomposed into independent dynamic linear models, each one describing particular features of the process under analyses. More details about some of literatures about Kalman Filter could be found in [6].

KF methodology could be useful in animal breeding industry when analyzing very big time series data sets since it does not need store or invert matrices. In addition online estimation of breeding values could be useful for selection schemes over time. As was shown by Van Bebber et al.[7] KF could also be used for detecting false measurements in experimental farms.

We investigated the possible application of KF-RW in field conditions. In this context, we chose body condition score (BCS) in dairy cows [8,9]. Body weight, feed (dry matter) intake and milk production level together form an important cluster of functional traits that determines the amount of fat reserves stored in the body; this is now recorded in many countries as 'body condition score' or BCS [8,9]. A common body condition scoring system has been developed to estimate the BCS of cows in a herd. This system provides producers a relative score based on an evaluation of fat deposits in relation to skeletal features. The scoring method involves a manual assessment of the thickness of fat cover and prominence of bone at the tail head and loin area. The most widely used body condition scoring system for dairy cattle assigns scores from 1 to 9 in North America and from 1 to 5 in most European countries, with the lowest score meaning emaciated and carrying virtually no fat and the highest score meaning excessively fat. Veerkamp and Brotherstone [10] estimated variance components for BCS at calving and for average BCS over the first 26 wk of lactation; they reported heritability estimates for BCS ranging between 0.24 and 0.43. Jones et al.[11] reported for Holstein Friesian heifers moderate heritability estimates of BCS, varying with stage of lactation from 0.23 to 0.28.

Sallas and Harville [12] suggested to use KF for prediction of breeding values for consecutive lactations. Forni et al.[13] used dynamic linear model via KF to modeling of cattle growth data. The main aim of this study was to provide theoretical developments of KF-RW in the context of adapting it to analysis of animal breeding data with an emphasis on application of KF-RW in predicting breeding values over DIM for BCS. This would be the first use of KF-RW for predicting breeding values over DIM. Simulations are performed to check the validity of methods and comparison of different sub-models.

MATERIAL and METHODS

Statistical Models and Analyses

Random Walk Model: For demonstration purposes we used random walk model and it is given below.

$$y_t = \alpha_t + \varepsilon_t \, , \, \varepsilon_t \propto N\left(0, \sigma_e^2\right)$$

$$\alpha_{t+1} = \alpha_t + \eta_t \, , \, \eta_t \propto N\left(0, \sigma_n^2\right) \qquad (1)$$

In (1) the first equation is called the observation equation and the second equation is called the state equation. We assumed that observations y_t, depends on unobservable quantity α_t, and our aim was to do statistical inference on α_t (states). We assumed constant variances for ε_t and η_t as σ_e^2 and σ_n^2 respectively with independent, identically and normally distributed random variables with zero means. LOWESS [14] smoother was used to make comparisons with random walk model (1) in recursive form (KF-RW).

Parameter values assumed were $\sigma_n^2 = 100$ and $\sigma_e^2 = 25$ for simulating observations, *y*, from random walk model. A total of 1000 Gibbs cycles with first 50 cycles used as burn in period used to obtaining estimates of states, $\hat{\theta}$, and variance components σ_n^2, σ_e^2. For genetic analyses of traits following mixed model is normally used in animal breeding;

$$\mathbf{y} = \mathbf{X}\hat{\mathbf{a}} + \mathbf{Z_a}\mathbf{a} + \mathbf{Z_p}\mathbf{p} + \mathbf{e} \qquad (2)$$

where **y** is the vector of observations, **â** is the vector of fixed effects, **a** is the vector of breeding values, **p** is the vector of random permanent environmental effects, **X**, **Z_a**, **Z_p** are design matrices and **e** is the vector of random residual effects.

For the random effects it was assumed that

$$Var\begin{pmatrix} \mathbf{a}_0 \\ \mathbf{p}_0 \\ \mathbf{e}_0 \end{pmatrix} \sim N\left[\mathbf{0}; \begin{pmatrix} \mathbf{A}\sigma_a^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\sigma_p^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2 \end{pmatrix}\right]$$

$$Var\begin{pmatrix} \mathbf{a}_t \\ \mathbf{p}_t \\ \mathbf{e}_t \end{pmatrix} \sim N\left[\begin{pmatrix} \mathbf{a}_{t-1} \\ \mathbf{p}_{t-1} \\ \mathbf{e}_{t-1} \end{pmatrix}; \begin{pmatrix} \mathbf{A}\sigma_a^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}\sigma_p^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}\sigma_e^2 \end{pmatrix}\right]$$

where σ_a^2, σ_p^2, and σ_e^2; are genetic, permanent environment and error variances. **A** is the additive genetic relationship matrix for the animals; **I** is an identity matrix.

In the following, we show general assumptions used in KF-RW method, based on Bayesian principles. Proportional joint posterior distribution without constant terms given in (3) using (2) based on following recursive relationship (4);

$$p\left(\boldsymbol{\theta}_\mathbf{t}\middle|\boldsymbol{\theta}_{(-\mathbf{t})},Y_n\right) \propto p\left(\boldsymbol{\theta}_\mathbf{t}\middle|\boldsymbol{\theta}_{\mathbf{t-1}},Y_{t-1}\right)p\left(\boldsymbol{\theta}_{\mathbf{t+1}}\middle|\boldsymbol{\theta}_\mathbf{t},Y_{t-1}\right)p\left(Y_\mathbf{t}\middle|\boldsymbol{\theta}_\mathbf{t},Y_{t-1}\right)$$

$$f\left(y\middle|b,a,p,\sigma_a^2,\sigma_p^2,\sigma_e^2\right) \propto \left(\sigma_e^2\right)^{-\frac{1}{2}N} \exp\left(-\frac{1}{2}\left(\mathbf{y}-\mathbf{Xb}-\mathbf{Z_a a}-\mathbf{Z_p p}\right)'\left(\mathbf{y}-\mathbf{Xb}-\mathbf{Z_a a}-\mathbf{Z_p p}\right)/\sigma_e^2\right)$$

$$\mathrm{x}\left(\sigma_a^2\right)^{-\frac{1}{2}N_a} \exp\left(-\frac{1}{2}\mathbf{a_1}'\mathbf{A}^{-1}\mathbf{a_1}/\sigma_a^2\right)\left[\prod_{t=2}^{T}\left(\sigma_a^2\right)^{-\frac{1}{2}N_a}\exp\left(-\frac{1}{2}\left(\mathbf{a_t}-\mathbf{a_{t-1}}\right)'\mathbf{A}^{-1}\left(\mathbf{a_t}-\mathbf{a_{t-1}}\right)/\sigma_a^2\right)\right]$$

$$\mathrm{x}\left(\sigma_p^2\right)^{-\frac{1}{2}N_p} \exp\left(-\frac{1}{2}\mathbf{p_1}'\mathbf{p_1}/\sigma_e^2\right)\left[\prod_{t=2}^{T}\left(\sigma_p^2\right)^{-\frac{1}{2}N_p}\exp\left(-\frac{1}{2}\left(\mathbf{p_t}-\mathbf{p_{t-1}}\right)'\mathbf{I}^{-1}\left(\mathbf{p_t}-\mathbf{p_{t-1}}\right)/\sigma_p^2\right)\right] \tag{3}$$

$$\mathrm{x}\left(\sigma_e^2\right)^{-\left(\frac{v_e+1}{2}\right)} \exp\left[-\frac{v_e S_e}{2\sigma_e^2}\right]\left(\sigma_a^2\right)^{-\left(\frac{v_a+1}{2}\right)}\exp\left[-\frac{v_a S_a}{2\sigma_a^2}\right]\left(\sigma_p^2\right)^{-\left(\frac{v_p+1}{2}\right)}\exp\left[-\frac{v_p S_p}{2\sigma_p^2}\right]$$

Last line of (3) are product of density of scaled inverted chi-square distributions assumed prior for variance parameters. After algebraic manipulations conditional distributions could be written as following,

$$b_t\middle|\sigma_a^2,\sigma_p^2,\sigma_e^2,a_t,p_t,y_t \sim N\left(\left(\mathbf{X'X}\right)^{-1}\mathbf{X'}\left(\mathbf{y}-\mathbf{Z_a a}-\mathbf{Z_p p_t}\right),\left(\mathbf{X'X}\right)^{-1}\sigma_e^2\right)$$

$$a_t\middle|\sigma_a^2,\sigma_p^2,\sigma_e^2,b_t,p_t,y_t \sim$$

$$N\left(\left(\frac{1}{\sigma_e^2}\mathbf{Z'_{a_t}}\mathbf{Z_{a_t}}+\frac{2}{\sigma_a^2}\mathbf{A}^{-1}\right)^{-1}\left(\frac{1}{\sigma_e^2}\mathbf{Z'}\left(\mathbf{y_t}-\mathbf{X_a b}_t-\mathbf{Z_{p_t} p}_t\right)+\frac{1}{\sigma_a^2}\mathbf{A}^{-1}a_{t+1}\right),\left(\frac{1}{\sigma_e^2}\mathbf{Z'_{a_t}}\mathbf{Z_{a_t}}+\frac{2}{\sigma_a^2}\mathbf{A}^{-1}\right)^{-1}\right)$$

$$p_t\middle|\sigma_a^2,\sigma_p^2,\sigma_e^2,a_t,b_t,p_t,y_t \sim$$

$$N\left(\left(\frac{1}{\sigma_e^2}\mathbf{Z'_{p_t}}\mathbf{Z_{p_t}}+\frac{2}{\sigma_p^2}\mathbf{I}^{-1}\right)^{-1}\left(\frac{1}{\sigma_e^2}\mathbf{Z'}\left(\mathbf{y_t}-\mathbf{X_a b}_t-\mathbf{Z_{p_{tt}} a}_t\right)+\frac{1}{\sigma_p^2}\mathbf{I}^{-1}a_{t+1}\right),\left(\frac{1}{\sigma_e^2}\mathbf{Z'_{p_t}}\mathbf{Z_{p_t}}+\frac{2}{\sigma_p^2}\mathbf{I}^{-1}\right)^{-1}\right)$$

$$\sigma_a^2\middle|\sigma_p^2,\sigma_e^2,p_t,a_t,b_t,p_t,y_t \sim \frac{\left(Q_a+v_a S_a\right)}{\chi_{DF}^2}$$

$$\sigma_p^2\middle|\sigma_a^2,\sigma_e^2,p_t,a_t,b_t,p_t,y_t \sim \frac{\left(Q_p+v_p S_p\right)}{\chi_{DF}^2}$$

$$\sigma_e^2\middle|\sigma_a^2,\sigma_p^2,p_t,a_t,b_t,p_t,y_t \sim \frac{\left(Q_e+v_e S_e\right)}{\chi_{DF}^2}$$

where in the last lines Q stands for quadratic form of the respective error terms and *DF* degrees of freedoms.

LOWESS: The LOWESS model [12] was used to capture the local variability by weighted least square regression using different smoother spans.

Implementation

All computations were made in R [14], using the package called MASS [15] implemented in R, for sampling from multivariate Normal distributions. We compared KF-RW and LOWESS approach. We investigated effect of different number of subjects and different number of observations per subject on estimation of both states and variance components. Finally we applied the theory to animal breeding data to predict breeding values over DIM.

Application of KF-RW Method to BCS Data

BCS was measured by the Swiss Holstein Breeding Association as described by Trimberger [16] using 1 to 5 scale, during May 2004-March 2005 for 7 times approximately at monthly intervals from multiparous dairy cows (*n*=80) stationed at the Chamau research farm of the Swiss Federal Institute of Technology, Switzerland. The experimental procedures of the farm followed the Swiss Law on Animal Protection and were approved by the Committee for the Permission of Animal Experiments of the Canton of Zug,

Zug, Switzerland. Results of 7 BCS sessions split into 4 periods over DIM to reduce number of missing values. Missing values filled in with subject specific averages. A summary of the dataset used for BCS analysis is given in *Table 1*. The pedigree file included 637 animals.

For genetic analyses of BCS following mixed model was used;

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Z_a a} + \mathbf{Z_p p} + \mathbf{e} \qquad (4)$$

where **y** is the vector of observations included 320 (80 animals, with 4 repeated measurements each) observations, β is the vector of fixed effects including breed ($n=80$; b_1=Holstein-Friesian, n_1=42, b_2= Brown Swiss, n_2=38) age at calving, and year-season interaction, **a** is the vector of breeding values, **p** is the vector of random permanent environmental effects, **X**, **Z_a**, **Z_p** are design matrices and **e** is the vector of random residual effects.

RESULTS

Random Walk Model for Simulated Data Sets

We compared predictions of simulated observations from random walk model with different smoother spans for LOWESS and KF-RW *(Table 2)*. Predictions were found to be both reasonable by KF-RW models and LOWESS approach.

We also investigated effect of different sampling sizes

to state and variance components estimates *(Table 3)* from KF-RW. Lowest number of observations (n=5) with lowest number of subjects (n=50) gives reasonable predictions of the states but incearesing both sampling sizes also decrease the standard errors of the estimates of the states. However prediction of variance components was found far from their parameters values with the lowest number of observations (n=5) and this effect gradually decreased by increasing the number of subject holding the number of observations as fixed. Standard errors of predicted variance components were found decreased by increasing the both number of observations and number of subjects.

Prediction of Breeding Values for Body Condition Scores by Kalman Filter

We analyzed BCS data for permanent environmental variance without genetic effects in order to decide number of Gibbs cycle and burn in period. Repeated runs of the same analyses for the same priors showed that monte carlo errors were small. Two thousand Gibbs cycles with 200 burn-in period were found reasonable. However, probably, since each cow consisted of small number of repeated measurements (n=4) different prior values gave different estimates, especially for smaller values of scaling factors, s, *(Table 4)*.

Increasing the scaling factor, s, gave more stable estimates, hence v=5 with s=0.5 chosen for parameters of scaled inverted chi-square prior distribution for genetic analyses of the BCS dataset. Scatter plots of gibbs samples

Table 1. *Number of records, means, and standard errors for body condition scores for some selected days of first lactation*
Tablo 1. *Vücut puanları için bazı günlere ait gözlem sayıları, ortalamalar ve bunlara ait standart hatalar*

DIM	Body Condition Score		
	n	X	SE
1	80	3.24	0.006
75	80	3.16	0.005
150	80	3.22	0.005
305	80	3.21	0.006

Table 2. *Means of predictions of observations simulated from random walk model by different smoother spans for LOWESS and Kalman Filter random walk methodologies*
Tablo 2. *Kalman Süzgeci rassal yürüyüş taslamı ve LOWESS (farklı üreteçler için) yöntemlerinin rassal yürüyüş benzeşimini tahmin ortalamaları*

Smoother Span	Observations	LOWESS	Kalman Filter
f=0.1	201.19 (0.12)	202.96 (0.02)	202.78 (0.08)
f=0.5	199.40 (0.11)	201.32 (0.06)	199.71 (0.07)
f=0.7	261.40 (0.25)	264.42 (0.21)	263.57 (0.23)
f=0.9	157.08 (0.17)	162.35 (0.12)	159.15 (0.15)
f=0.01	203.62 (0.13)	202.93 (0.06)	204.71 (0.09)
f=0.05	160.75 (0.15)	161.41 (0.11)	161.51 (0.13)
f=0.07	184.61 (0.16)	182.30 (0.13)	185.08 (0.14)
f=0.09	116.91 (0.18)	119.25 (0.15)	117.42 (0.17)

were visually investigated (figures are not shown), since no patterns were observed for all variance components; it was decided that sample size was reasonable. Posterior means for variance components found 0.21 (0.02) for error variance, 0.03 (0.006) for animal genetic variance and 0.04 (0.007) for permanent environmental variance. In this data set variation was found quite small within and between animal.

DISCUSSION

In the simulation study predictions of observations were found to be both reasonable by KF-RW models and LOWESS approach. However mostly KF-RW estimates were found better than those of LOWESS estimates *(Table 2)*. It is difficult to compare two methodologies directly, since each of them has their own specifications, but KF-RW model does not need any tuning parameters. However predicting variance components needs additional computing power in KF-RW.

We also investigated effect of different sampling sizes to state and variance components estimates *(Table 3)* from KF-RW. Estimates of states were getting closer to observations as number of observations and number of subjects increased. Also it was observed that increasing both quantities decreased associated standard errors. Variance components estimates were found to be more accurate when number of observations and number of subjects increased. Again increasing these quantities decreased standard errors. Fifty subjects with 10 observations each, started to gave reasonable estimates.

In real biological dataset, different from simulated dataset, number of observations per animals was low (*n*=4) and this was in effect for predicting variance components *(Table 4)*. Different prior values gave very different variance component predictions especially for smaller values of scaling factors, however it was stabilized by increasement of both degree of belief and scaling factors. Based on various runs we decided to use *v*=5, *s*=0.5 for genetic analyses of the BCS dataset.

We used 2000 Gibbs cycles with 200 burn in period, based on results of different runs (results not shown). Since dataset were obtained under homogenized, controlled experimental farm; small amount of variances predicted for both within and between animals. Visual inspection showed that predictions of observations were agreed with actual observations.

We provided a general theoretical framework for use of KF-RW in the analyses of animal breeding data, with an emphasis on application of KF-RW in predicting breeding values for BCS measured as a longitudinal trait. Simulations

Table 3a. *Means of estimates of states* $\hat{\theta}$ *, observations* y *. Standard errors given in bracets*

Tablo 3a. *Durumların* $\hat{\theta}$ *ve gözlemlerin* y *tahminlerinin ortalamaları. Standart hatalar parantez içinde verilmiştir*

Number of Subjects	Number of Observations							
	5		10		20		50	
	$\hat{\theta}$	y	$\hat{\theta}$	y	$\hat{\theta}$	y	$\hat{\theta}$	y
50	198.80 (0.04)	198.85 (0.05)	201.08 (0.02)	201.02 (0.02)	200.67 (0.02)	200.67 (0.01)	196.57 (0.009)	196.52 (0.001)
100	200.77 (0.02)	200.87 (0.03)	200.19 (0.01)	200.24 (0.01)	200.02 (0.008)	199.92 (0.009)	197.75 (0.005)	197.55 (0.005)
200	199.58 (0.01)	200.38 (0.02)	199.33 (0.006)	199.58 (0.07)	200.00 (0.004)	200.04 (0.004)	200.72 (0.002)	200.68 (0.002)
500	199.46 (0.004)	199.54 (0.005)	199.66 (0.003)	199.63 (0.002)	199.42 (0.001)	199.35 (0.002)	200.89 (0.0009)	200.98 (0.001)
1.000	199.63 (0.002)	199.86 (0.002)	199.95 (0.001)	200.08 (0.001)	199.48 (0.0009)	199.53 (0.0007)	200.05 (0.0005)	200.04 (0.0005)

Table 3b. *Means of predictions of variance components (* $\hat{\sigma}_e^2$ $\hat{\sigma}_n^2$ *). Standard errors given in brackets*

Tablo 3b. *Varyans unsurlarının tahminlerinin ortalamaları* $\hat{\sigma}_e^2$ $\hat{\sigma}_n^2$ *. Standart hatalar parantez içinde verilmiştir*

Number of Subjects	Number of Observations							
	5		10		20		50	
	$\widehat{\sigma}_n^2$	$\hat{\sigma}_e^2$	$\widehat{\sigma}_n^2$	$\hat{\sigma}_e^2$	$\widehat{\sigma}_n^2$	$\hat{\sigma}_e^2$	$\widehat{\sigma}_n^2$	$\hat{\sigma}_e^2$
50	71.00 (0.13)	40.52 (0.12)	26.01 (0.03)	102.58 (0.07)	29.43 (0.02)	99.97 (0.01)	26.07 (0.004)	103.99 (0.01)
100	95.58 (0.10)	35.71 (0.11)	25.42 (0.01)	101.56 (0.04)	31.34 (0.01)	93.88 (0.02)	26.39 (0.003)	92.62 (0.007)
200	97.38 (0.05)	33.92 (0.05)	27.91 (0.01)	94.67 (0.02)	23.44 (0.003)	98.43 (0.009)	26.92 (0.001)	102.15 (0.003)
500	100.31 (0.02)	36.02 (0.02)	25.47 (0.003)	99.64 (0.009)	23.18 (0.001)	102.09 (0.001)	25.10 (0.0005)	99.32 (0.001)
1.000	97.75 (0.007)	20.94 (0.01)	23.16 (0.001)	101.98 (0.004)	23.71 (0.0007)	100.93 (0.002)	24.47 (0.0003)	100.59 (0.0008)

$\hat{\sigma}_e^2$ *: Estimates of error variance* $\widehat{\sigma}_n^2$ *: Estimates of states variance*

Predicting Breeding Values...

Table 4. Estimates of state ($\bar{\sigma}_n^2$) and error variances ($\hat{\sigma}_e^2$) for different parameter values of scaled inverted chi-square prior distributions (v degree of belief, s scaling factors)

Tablo 4. Durumların ($\bar{\sigma}_n^2$) ve hata varyanslarının ($\hat{\sigma}_e^2$) ölçeklenmiş tersinir kikare öncül dağılışı ile tahmini (v inanç derecesi, s ölçekleme faktörü)

s	v=0.01		v=1		v=2		v=3		v=4		v=5	
	$\hat{\sigma}_n^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_n^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_n^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_n^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_n^2$	$\hat{\sigma}_e^2$	$\hat{\sigma}_n^2$	$\hat{\sigma}_e^2$
s=1	0.0005	0.18	0.02	0.18	0.04	0.19	0.04	0.20	0.05	0.21	0.06	0.21
s=0.1	0.00009	0.18	0.04	0.18	0.009	0.17	0.01	0.17	0.02	0.17	0.02	0.18
s=0.5	0.0003	0.18	0.01	0.18	0.03	0.19	0.03	0.19	0.04	0.19	0.04	0.19

Standard errors varied between 0.000001-0.14

were also performed to check the validity of methods and comparison of different sub-models. Although we used random walk model; model choice depends on the variability and prior information about dataset, different models could be more realistic in different applications. Since random walk model is not stationary it may not be suitable for animal breeding data under certain conditions. We assumed constant variances for state and error components over DIM, however this assumptions could be extended for time dependent variance components models, and it could be claimed that it would give more realistic results. Since KF gives online estimation of breeding values and does not need to store or invert matrices, this methodology could be useful in animal breeding industry for obtaining online estimation of breeding values over DIM.

ACKNOWLEDGEMENTS

REFERENCES

1. Grewal MS, Andrews AP: Kalman Filtering: Theory and practice using MATLAB. JohnWiley&Sons, 2001.

2. Harrison PJ, Stevens C: A Bayesian approach to short term forecasting. *Oper Res Quart*, 22, 341-362, 1971.

3. Meinhold RJ, Singpurwalla ND: Understanding the Kalman Filter. *Am. Stat,* 37, 123-127, 1983.

4. West M, Harrison PJ: Bayesian forecasting and dynamic models. Springer. 1997.

5. Migon H, Gamerman D, Lopes H, Ferreira M: Dynamic models. In Dey D, Rao C (Eds): Handbook of Statistics. p. 25. Elsevier, 2005.

6. Karacaören B: Phenotypic and genetic analysis of functional traits in dairy cattle under experimental management system. *PhD Thesis.* Swiss Federal Institute of Technology Zurich, 2006.

7. Van Bebber J, Reinsch N, Junge W, Kalm E: Monitoring daily milk yields with a recursive test day repeatability model (Kalman Filter). *J Dairy Sci*, 82, 2421-2429, 1999.

8. Kadarmideen HN, Wegmann S: Genetic parameters for body condition score and its relationship with type and production traits in Swiss Holsteins. *J Dairy Sci*, 86, 3685-3693, *2003*.

9. Kadarmideen HN: Genetic correlations among body condition score, somatic cell count, production, reproduction, and conformation traits in Swiss Holsteins. *Anim Sci*, 79, 191-201, 2004.

10. Veerkamp RF, Brotherstone S: Genetic correlations between linear type traits, food intake, live weight and condition score in Holstein Friesian dairy cattle. *Anim Sci*, 64, 385-392 1997.

11. Jones HE, White IMS, Brotherstone S: Genetic evaluation of holstein friesian sires for daughter condition-score changes using random regression model. *Anim Sci*, 68, 467-475, 1999.

12. Sallas WM, Harville DA: Best linear recursive estimation for mixed linear models. *J Amer Statist Assoc*, 76, 860-869, 1981.

13. Forni S, Gianola D, Rosa GJM, de los Campos G: A dynamic linear model for genetic analysis of longitudinal traits. *J Anim Sci*, 87, 3845-3853, 2009.

14. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2004

15. Venables WN, Ripley BD: Modern Applied Statistics with S. Fourth ed,. Springer, New York, 2002.

16. Trimberger GW: Dairy Cattle Judging Techniques. Prentice-Hall, Prospect Heights, Waveland Press, 1977.