

UNIVERSIDAD AUTÓNOMA CHAPINGO

Departamento de enseñanza, investigación y Servicio en Zootecnia

**MODELOS DE FUNCIONES DE COVARIANZA EN LA ESTIMACIÓN
DE PARÁMETROS GENÉTICOS DE PRODUCCIÓN DE LECHE,
GRASA Y PROTEÍNA EN VACAS HOLSTEIN**

TESIS

**QUE COMO REQUISITO PARCIAL
PARA OBTENER EL GRADO DE:**

MAESTRO EN CIENCIAS EN PRODUCCIÓN ANIMAL

PRESENTA:

JUAN GABRIEL CANCHÉ UC



BIBLIOTECA CENTRAL U.A. CH

**Marzo de 2002
Chapingo, estado de México**

BIB 97436

**MODELOS DE FUNCIONES DE COVARIANZA EN LA ESTIMACIÓN DE
PARÁMETROS GENÉTICOS DE PRODUCCIÓN DE LECHE, GRASA Y
PROTEÍNA EN VACAS HOLSTEIN**

Tesis realizada por Juan Gabriel Canché Uc bajo la dirección del Comité Asesor
indicado, aprobada por el mismo y aceptada como requisito parcial para obtener el
grado de :

MAESTRO EN CIENCIAS EN PRODUCCIÓN ANIMAL

DIRECTOR:



M. C. CARLOS A. APODACA SARABIA

ASESOR:



DR. RODOLFO RAMÍREZ VALVERDE

ASESOR:



DR. JOSÉ ARTEMIO CADENA MENESES.

BIBLIOTECA CENTRAL U.A.C.B.

39495

**MODELOS DE FUNCIONES DE COVARIANZA EN LA ESTIMACIÓN DE
PARÁMETROS GENÉTICOS DE PRODUCCIÓN DE LECHE, GRASA Y
PROTEÍNA EN VACAS HOLSTEIN**

El jurado que revisó y aprobó el examen de grado de Juan Gabriel Canché Uc autor de la presente tesis de Maestría en Ciencias en Producción Animal estuvo constituido por:

PRESIDENTE:



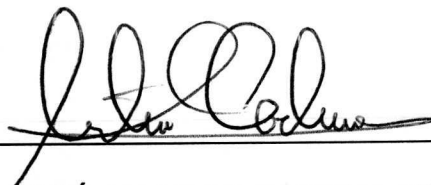
DR. RODOLFO RAMÍREZ VALVERDE

ASESOR:



M. C. CARLOS A. APODACA SARABIA

ASESOR:



DR. JOSÉ ARTEMIO CADENA MENESES

DEDICATORIA

Esta tesis va dedicada con gran cariño:

A mi familia, por infundar en mi persona la confianza que todo ser humano necesita en cualquier empresa de la vida.

A mi esposa y a mis dos hijas que son lo más valioso que tengo en la vida, por haberme brindado su comprensión durante mis estudios y porque son el motivo de mi existencia.

y

A tí señor que iluminas nuestros sendero aun a sabiendas que tus hijos se encuentran disgregados por este mundo terrenal.

AGRADECIMIENTOS:

A la Universidad Autónoma Chapingo y al Departamento de Zootecnia, por brindarme la oportunidad de realizar los estudios de Maestría en Ciencias.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT), por financiar mis estudios de Maestría.

Al M. C. Carlos Apodaca Sarabia, por su valiosa y acertada dirección y apoyo en la realización del presente trabajo.

A los Doctores Rodolfo Ramírez Valverde y José Artemio Cadena Meneses por sus invaluables sugerencias y aportaciones que enriquecieron y mejoraron el trabajo.

Al personal de la Unidad Tecnológica Lechera de la Granja Experimental, y del laboratorio de Nutrición Animal, por el apoyo recibido durante mi estancia los laboratorios.

A todos mis colegas de posgrado, por haberme brindado su amistad.

DATOS BIOGRÁFICOS

I.- Datos personales

Nombre	Juan Gabriel Canché Uc
Fecha de Nacimiento	23 de septiembre de 1976
Lugar de Nacimiento	Bécal, Calkiní, Campeche
Profesión	Ingeniero Agrónomo Zootecnista
Cédula profesional	2932776

II.- Desarrollo Académico.

1982 – 1988	Escuela Primaria Urbana Federal, Vicente Castro Bacallao, Bécal, Calkiní, Campeche.
1988 – 1991	Escuela Secundaria Técnica No 13, Vicente Guerrero, Bécal, Calkiní, Campeche.
1991 – 1994	Preparatoria Agrícola de la Universidad Autónoma Chapingo, Chapingo estado de México.
1994 – 1998	Licenciatura en el Departamento de Zootecnia de la Universidad Autónoma Chapingo, Chapingo estado de México.
2000 – 2001	Maestría en Producción Animal en el Departamento de Zootecnia de la Universidad Autónoma Chapingo, Chapingo estado de México.

Experiencia Profesional.

1998 – 1999	Asesor Técnico Pecuario del Programa de Desarrollo Agrícola y Pecuario de la Zona Norte del estado de Chiapas (PRODEAP), en el Municipio de Tumbala Chiapas.
-------------	--

ÍNDICE

CONTENIDO	PÁG.
DEDICATORIA.....	iv
AGRADECIMIENTOS.....	v
DATOS BIOGRÁFICOS.....	vi
CAPÍTULO 1. INTRODUCCIÓN	1
CAPÍTULO 2. REVISIÓN DE LITERATURA.....	6
1. Evolución de los modelos del día de registro	7
2. Ecuaciones de los Modelos Mixtos.....	14
3. Estimación de componentes de varianza	18
4. Literatura citada.....	26
CAPÍTULO 3. MODELOS DE FUNCIONES DE COVARIANZA EN LA ESTIMACIÓN DE PARÁMETROS Y VALORES GENÉTICOS PARA PRODUCCIÓN DE LECHE, GRASA Y PROTEÍNA.....	31

ÍNDICE DE CUADROS Y FIGURAS

	Página
Cuadro 1. Modelos lineales utilizados para modelar la producción de leche	11
Cuadro 2. Índices de herencia (h^2) y de ambiente permanente (EP), y varianzas genéticas (σ^2_a), de ambiente temporal (σ^2_E) y permanente (σ^2_{EP}), para producción de leche, grasa y proteína a los 305 días de lactancia.....	22
Cuadro 3. Estimación de varianzas genéticas (g) e índices de herencia (h^2) para producción diaria de grasa y proteína en leche.	23
Cuadro 4. Correlaciones genéticas estimadas (x 100) de producción de grasa (debajo de la diagonal) y proteína (arriba de la diagonal) en vacas Holstein de primer parto.	24
Cuadro 5. Varianzas, Covarianzas genéticas (bajo la diagonal) y fenotípicas (sobre la diagonal) entre las producciones diarias de leche, grasa y proteína en la lactancia...	42
Cuadro 6. Valores genéticos de producción de proteína: los días 60 (Vg_{60}) y 280 (Vg_{280}), total por lactancia (Vg_T), en el primer (Vg_{1T}), segundo (Vg_{2T}) y tercer (Vg_{3T}) tercio de la lactancia, y de persistencia ($Vg_{P1}=(Vg_{280}/Vg_{60})100$	44
Figura 1. Índice de herencia para producción de leche, grasa y proteína durante la lactancia en vacas Holstein.	43

CAPÍTULO 1

INTRODUCCIÓN

INTRODUCCIÓN

Tradicionalmente el mejoramiento genético de las características de producción de leche se realiza para producción acumulada en una lactancia de 10 meses, calculada con registros periódicos de producción por día durante la lactancia. Si la lactancia es completa una forma de calcular la producción total fue descrita por Ferris (1981) y en caso de lactancias incompletas se recurre a factores de proyección. Estas metodologías suponen que en la producción por día no son importantes los efectos no genéticos, ni las diferencias en los valores genéticos de los animales. Estas suposiciones pueden producir sesgos importantes en la selección de los animales por su valor genético.

La metodología de los modelos mixtos, y los avances en técnicas de programación y en computación han permitido: 1) contar con estimaciones y predicciones con mejores propiedades estadísticas, de máxima verosimilitud y mejor predictor lineal insesgado, 2) resolver grandes sistemas de ecuaciones simultáneas, y 3) uso de modelos más complejos (Henderson, 1974), uno de esos modelos es el denominado del día de registro (MDP). Las principales ventajas del MDP, es considerar en el modelo los efectos no genéticos sobre la observación el día en que se realizó, obtener la descripción genética de la característica en todo el periodo de su expresión, y reducir el número de parámetros a estimar (Van der Werf y Schaeffer 1997; Freeman 1998)

Las características que se observan repetidamente en un periodo de tiempo dado son susceptibles para aplicar el análisis del MDP, y se denominan longitudinales debido a que, teóricamente, se pueden registrar en un infinito número de veces. En la ganadería lechera, las características registradas en una lactancia son una ejemplo a lo antes mencionado. El

MDP es una metodología que ha evolucionado en el tiempo, donde, la idea principal es predecir los valores genéticos de los animales tomando en cuenta los efectos sistemáticos no genéticos que generan variabilidad sobre las observaciones el día en que se realizó la observación.

Las diferentes formas en que se ha usado el MDP son: 1) con un modelo fijo, ajustando la producción por día por los efectos ambientales, y posteriormente calculando la producción por lactancia y estimando los parámetros genéticos; 2) en el modelo de características repetidas, incluyendo covariables fijas con la finalidad de considerar la etapa de la lactancia y suponer que los valores genéticos y los efectos de ambiente permanente no cambian en la lactancia; y 3) considerando que los valores genéticos cambian en la lactancia, es posible: a) considerar la producción por día en diferentes etapas de la lactancia como características correlacionadas genética y fenotípicamente, y el modelo adecuado es entonces el multivariado; b) sustituir en el modelo multivariado, los valores genéticos de los animales por la ecuación de regresión que los predice, modelo denominado de regresión aleatoria; y c) sustituir las varianzas y covarianzas genéticas estimadas con el modelo multivariado por la función de covarianza que las predice, denominado modelo de funciones de covarianza (Freeman, 1998).

El modelo de mediciones repetidas es el más simple, sin embargo, la suposición de que el valor genético y de ambiente permanente del animal es el mismo en la lactancia puede no ser correcta. El análisis multivariado, requiere de la estimación de un gran número de parámetros, incrementando los requerimiento de tiempo y equipo de computo, y aumentando la posibilidad de no lograr convergencia. Los modelos de regresión aleatoria predicen los valores genéticos de los animales para cualquier día de la lactancia utilizando

ecuaciones de regresión para cada animal, y estimando las varianzas y covarianzas genéticas de los coeficientes de regresión. Los modelos de funciones de covarianza predicen tanto los valores genéticos de los animales como las varianzas y covarianzas de la característica en la lactancia (Schaeffer y Dekkers, 1994; Van der Werf y Schaeffer, 1997; Swalve, 2000). Los modelos de regresión aleatoria y de funciones de covarianza son univariados, en los que no solo se estima la varianza genética sino una matriz de varianzas y covarianzas en regresión aleatoria, y la matriz de coeficientes de la función de covarianza.

Las funciones de covarianza, inicialmente fueron descritas por Kirkpatrick *et al.* (1990), usando polinomios ortogonales de Legendre, que requieren de la estimación previa de las varianzas y covarianzas genéticas entre algunos periodos de registro de la característica. Así mismo, Meyer y Hill (1997) y Meyer (1998) reparametrizaron el modelo de regresión aleatoria para producir la función de covarianza descrita por Kirkpatrick, generando el modelo de funciones de covarianzas, con las ventajas de no requerir la estimación previa de las varianzas y covarianzas genéticas, y de usar máxima verosimilitud restringida en lugar de cuadrados mínimos.

Una de las grandes ventajas de los modelos de regresión aleatoria y de las funciones de covarianzas es que con las ecuaciones de predicción de los valores genéticos, podemos obtener no solo el valor genético en cada día de la lactancia, sino también para producción: por lactancia, máxima por día, acumulada en cualquier periodo deseado y para persistencia. Lo anterior sin ser consideradas en el vector de observaciones, es decir con un análisis univariado.

En el presente trabajo, primeramente hace una revisión de la teoría de los modelos del día de prueba en la predicción de los valores genéticos para producción por día en la lactancia, y en la segunda parte se presenta un artículo mostrando algunos resultados, que por el reducido número de observaciones utilizado debe ser tomado más como un ensayo.

CAPÍTULO 2

REVISIÓN DE LITERATURA

1. Evolución de los modelos del día de registro

De las formas de uso del MDP referidas por Freeman (1998), los modelos multivariado, de regresión aleatoria y funciones de covarianza son los que describen genéticamente la característica en toda la lactancia; sin embargo, con fines didácticos es recomendable iniciar la evolución del MDP, con el de medidas repetidas hasta llegar al de funciones de covarianza.

Modelo de repetibilidad

Al considerar a la producción de leche diaria como una característica que se repite en una lactancia, se obtiene un modelo con un reducido número de parámetros a estimar. Sin embargo, es necesario suponer que el valor genético y de ambiente permanente del animal no cambian en la lactancia, así el modelo mixto correspondiente será:

$$y_{ijk} = HTD_i + \sum_{m=0}^n b_m X_m + u_j + p_j + e_{ijk} \quad (1)$$

donde y_{ijk} es la producción por día, HTD_i es el efecto fijo del i -ésimo Hato-día de prueba, b_m son los coeficientes de regresión parcial para cuantificar el efecto promedio de etapa de lactancia, u_j es el valor genético aleatorio del animal j , p_j es el efecto ambiental permanente; y, e_{ijk} es el efecto residual. En (1) se tiene: $E(y_{ijk}) = HTD + \sum_{m=0}^k b_m X_m$, $E(u_j) = E(p_j) = E(e_{ijk}) = 0$, $var(u_j) = \sigma_a^2$, $var(p_j) = I\sigma_p^2$, $var(e_{ijk}) = I\sigma_e^2$ y $cov(u_j, p_j) = 0$. Debido a la fisiología animal se espera que los genes que determinan la producción por día cambie durante la lactancia, así el modelo de repetibilidad puede producir sesgos, al suponer la persistencia en la producción por día es solo causada por factores no genéticos. Una

forma de incluir los cambios genéticos producidos en la lactancia es el modelo multivariado.

Modelo multivariado

Dividiendo la lactancia en diferentes etapas y suponiendo: a) la producción por día dentro de cada etapa como la misma variable, y b) la producción por día en etapas diferentes como variables correlacionadas genética y fenotípicamente, un modelo adecuado propuesto es el multivariado con tantas variables como etapas se tengan. El modelo para la i -ésima etapa es: $y_{ijk} = HTD_i + b_1 DL + u_k + e_{ijk}$, donde y_{ijk} es la producción diaria del animal k -ésimo en la i -ésima etapa, b_1 es el coeficiente de regresión para considerar el efecto lineal del día de lactancia, u_k es el valor genético del k -ésimo animal en la i -ésima etapa, y e_{ijk} es el residual, suponiendo $E(y_{ijk}) = HTD_i + b_1 DL$, $E(u) = E(e_{ijk}) = 0$, $var(u) = \sigma_{ai}^2$, y $var(e) = \sigma_{ei}^2$. El modelo para todas la etapas, en forma matricial es:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e} \quad (2)$$

En (2), $\mathbf{y}$ es el vector de observaciones compuesto por los vectores de observaciones de cada etapa, $\mathbf{X}$ la matriz de incidencia que relaciona cada efecto fijo con las observaciones, $\mathbf{b}$ el vector de efectos fijos, $\mathbf{Z}$ la matriz de incidencia que relaciona cada animal con las observaciones, $\mathbf{u}$ el vector de efectos genéticos, y $\mathbf{e}$ el vector de residuales. El modelo (2) desglosado para t etapas se escribe:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \cdots & 0 \\ 0 & X_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & X_t \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_t \end{bmatrix} + \begin{bmatrix} Z_1 & 0 & \cdots & 0 \\ 0 & Z_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & Z_t \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_t \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_t \end{bmatrix}$$

De esta forma, $E(\mathbf{y}) = \mathbf{Xb}$, $\text{Var}(\mathbf{u}) = \mathbf{G}_0$, la matriz de varianzas y covarianzas genéticas entre las producciones diaria en las t etapas, y $\text{var}(\mathbf{e}) = \mathbf{R}$, la matriz de covarianzas de los residuales que puede tomar diferentes formas, de la cual la más simple es $\mathbf{I}\sigma_e^2$, la varianza del residual multiplicada por una matriz idéntica de orden t que corresponde al orden del número de etapas consideradas en el modelo. El modelo multivariado produce mejores predicciones de los valores genéticos al usar la información de cada una de las etapas en su información propia y la aportada por las otras en la covarianza entre ellas (Núñez y Rodríguez, 1995; Van der Werf y Schaeffer, 1997).

Regresión aleatoria. Si el número de etapas es grande el modelo multivariado no es práctico, este es el caso de producción por día en una lactancia de 10 meses, en la que se busca describir los cambios graduales en los valores genéticos de los animales en cada día de la lactancia, es decir se deben estimar 300 valores genéticos para cada animal en la población. Una forma de solucionar este problema es mediante el uso de regresión aleatoria, así el valor genético del animal para cada día de lactancia se obtendrá mediante la ecuación respectiva (Snedecor y Cochran, 1979; Henderson, 1982).

El modelo de regresión aleatoria es una modificación al multivariado, de forma tal que el vector de los t valores genéticos de cada animal se escribe por la ecuación que los predice, buscando que el número de coeficientes de regresión sea menor al de etapas en

el modelo multivariado o que el del días en lactancia. El modelo puede ser descrito como:

$y_{itj} = HTD_i + \sum_{k=0}^n \beta_k X + \sum_{k=0}^n \alpha_{kj} X + e_{itj}$, donde y_{itj} es la producción por día el t -ésimo día de lactancia, HTD_i el i -ésimo efecto fijo, β y α los coeficientes de regresión que ajustan por los días en lactancia en general y los coeficientes de regresión aleatoria del j -ésimo animal (Schaeffer y Deckkers, 1994). Una relación de las ecuaciones de predicción usadas como fijas o aleatorias se muestran en el Cuadro 1. El modelo de regresión aleatoria considerando la ecuación de Guo y Swalve (1995) es: $y_{itj} = HTD_i + \beta_0 + \beta_1 t^{0.5} + \beta_2 \ln t + \alpha_0 + \alpha_{1j} t^{0.5} + \alpha_{2j} \ln t + e_{itj}$.

Este modelo puede representarse en la forma general de un modelo lineal mixto; $\mathbf{y} = \mathbf{Xb} + \mathbf{Z}\alpha + \mathbf{e}$, en donde $\mathbf{y}$ es el vector de observaciones correspondiente a la producción de leche por día en el día t de la lactancia (más de una para cada animal con registros), $\mathbf{X}$ es la matriz de incidencia que relaciona las observaciones con los efectos fijos, con ceros y unos para el efecto de clasificación y valores de la covariable, $t^{0.5}$ y $\ln t$, para los coeficientes de regresión (β), $\mathbf{b}$ es el vector de los efectos fijos cuyos elementos son los efectos fijos y los coeficientes β_k , $\mathbf{Z}$ es la matriz de incidencia que relaciona los valores de la covariable señaladas en cada uno de las observaciones para cada animal y α es el vector de los coeficientes de regresión aleatoria (α_{ij}) que predecirán los valores genéticos de cada animal y en cada día de la lactancia (Schaeffer y Dekkers, 1994), $\mathbf{e}$ es el residual, y $\text{var}(\alpha) = \mathbf{G}$ es la matriz de covarianzas aditivas de los coeficientes de regresión.

Cuadro 1. Modelos lineales utilizados para modelar la producción de leche .

Modelo ¹	Autor
$Y_t = a + bt + ct^2$	Gaskins y Anderson, 1980
$Y_t = a + bt - ct^2$	Dave (1971)
$Y = at^b e^{-ct}$	Wood, 1967
$Y_t = \exp(a - bn_1 - cn_1^2 - d/n)$	Morant y Gnanasakthy, 1989
$Y_t = ae^{-bt} - ae^{-ct}$	Brody <i>et al.</i> (1924)
$Y_t = b_0 + b_1 t + e^{(-dt)} Z(t)$	White <i>et al.</i> (1999)
$Y^{-1} = a + bt^{-1} + ct$	Nelder, 1966
$Y^{-1} = a + bt^{-1} + ct + dt^2$	Badner y Anderson, 1985
$Y_t = t/(a + bt + ct^2)$	Nelder, 1966
$Y_t = a + bt + c \exp(-0.5(\log(t)-1)/0.6)^2 t^{-1}$	Guo y Swalve, 1995
$Y_t = a + bt^{0.5} + c \ln t$	Guo y Swalve, 1995
$Y_t = a + bt + c \exp(-0.05t)$	Wilmink, 1987
$Y_t = a + bt^{0.5} + c \ln t + dt^4$	Guo y Swalve, 1995
$Y_t = a + b(t/305) + c(t/305)^2 + d \ln(305/t) + f \ln^2(305/t)$	Ali y Schaeffer, 1987
$Y_t = a + bt + c \sin(x)t^2 + d \sin(x)t^3 + f \exp(-0.055t)$	Guo y Swalve, 1995
$Y_t = a + bt + ct^2 + dt^3 + f \ln(t)$	Guo y Swalve, 1995
$Y_t = a + bt + ct^2 + dt^3 + ft^4$	Guo y Swalve, 1995
$Y_t = a + bt + ct^2 + dt^3 + ft^4 + gt^5 + ht^6$	Guo y Swalve, 1995

Fuente: Jamrozik *et al.*, 1997; Meraz, 2001.

¹ Términos del modelo: a, b, c, d, f, g, h = parámetros a ser estimados en el modelo, t = días en lactancia.

Funciones de covarianza

Las funciones de covarianza tratan de predecir las varianzas y covarianzas genéticas entre una característica que puede ser medida teóricamente un infinito número de veces, denominada longitudinal. En producción animal existen algunas características que pueden ser medidas más de una vez. En ganado lechero se encuentran todas aquellas

relacionadas con la lactancia; como producción de leche, grasa y proteína, y conteo de células somáticas.

Si tenemos una estimación de la matriz de covarianzas genéticas para producción diaria de leche medida cada 30 días, $\hat{\mathbf{G}}$, es posible se requiera de covarianzas entre producción de días que no fueron medidas, para ello las funciones de covarianza son una alternativa (Kirkpatrick *et al.*, 1990; Kirkpatrick *et al.*, 1994). La función de covarianza es una forma de modelar las varianzas y covarianzas genéticas entre los registros repetidos de una característica, $G(t_j, t_{j^*}) = \sum_{i=1} \sum_{j=1} [\mathbf{C}]_{ij} \phi_i(a^*_1) \phi_j(a^*_2)$. Para ello se puede expresar; $\hat{\mathbf{G}} = \boldsymbol{\phi} \hat{\mathbf{C}} \boldsymbol{\phi}'$,

donde $\boldsymbol{\phi}$ es una matriz cuyos elementos son los coeficiente de los polinomios ortogonales estandarizados de Legendre, y $\mathbf{C}$ es la matriz de coeficientes de las estimaciones de la función de covarianza, $\hat{\mathbf{C}} = \boldsymbol{\phi}^{-1} \hat{\mathbf{G}} \boldsymbol{\phi}^T$.

La elección del grado del polinomio es de gran importancia, ya que determina el tamaño de la Matriz $\mathbf{C}$, es decir, el número de coeficientes a estimar. Kirkpatrick *et al.* (1994) propusieron estimar primero el polinomio de grado igual al tamaño de $\hat{\mathbf{G}}$, y posteriormente reducir el orden hasta encontrar el que genere covarianzas estadísticamente iguales a las de $\hat{\mathbf{G}}$, con el procedimiento recomendado de cuadrados mínimos. Van der Werf y Schaeffer (1997) propusieron iniciar con un polinomio de menor orden y aumentar este en uno, hasta en encontrar el polinomio que genere una estimación de $\mathbf{G}$ con el menor error. La prueba estadística usada en el primer caso es χ^2 y F en el segundo. En ambos casos se busca que la suma de las diferencias al cuadrado entre los componentes de varianza estimados no sea importante.

La desventaja de este procedimiento radica en que se tiene que realizar una estimación previa de $\mathbf{G}$, para lo cual debe usarse un análisis multivariado, y luego estimar las funciones de covarianza, implicando altos requerimiento de cómputo. Una forma diferente, basada en el modelo de regresión aleatoria, fue propuesta por Meyer (1998). Tomando el modelo $y_{itj} = F_i + \sum_{m=1}^{k-1} \beta_m \phi_m(t_j) + e_{itj}$, en el que y_{itj} es la producción de leche por día en i -ésimo efecto fijo, registrada el día t de lactancia en el j -ésimo animal, F_i es el efecto fijo, β_m son los coeficientes de regresión sobre el polinomio de Legendre, $\phi_m(t_j)$, de grado k , y e_{itj} , es el efecto residual. La covarianza entre dos observaciones es:

$$\begin{aligned} \text{cov}(y_{itj}, y_{i^*t^*j^*}) &= \sum_{m=1}^{k-1} \sum_{m^*=1}^{k-1} \text{cov}(\beta_m, \beta_{m^*}) \phi_m(t_j) \phi_{m^*}(t_{j^*}) + \text{cov}(e_{itj}, e_{i^*t^*j^*}) \\ &= G(t_j, t_{j^*}) + \text{cov}(e_{itj}, e_{i^*t^*j^*}). \end{aligned}$$

En la ecuación anterior se produce la función de covarianza referido por Kirkpatrick *et al.* (1994), con las ventajas de no requerir de la estimación previa de $\mathbf{G}$, de que aunque en el ejemplo se usaron polinomios de Legendre, se puede usar cualquier modelo de regresión aleatoria, y de usar el procedimiento Máxima Verosimilitud Restringida (REML). Lo anterior es de importancia por la reducción producida en los requerimientos de cómputo, y el uso de REML en lugar de cuadrados mínimos como el propuesto por Kirkpatrick *et al.* (1994).

La construcción de las matrices $\mathbf{X}$ y $\mathbf{Z}$ para el modelo de función de covarianza es igual a la construcción para regresión aleatoria, es decir, la matriz $\mathbf{Z}$ es de covarianza anidada para cada animal y $\mathbf{Z}'\mathbf{Z}$ es diagonal en bloques para producir las ecuaciones de predicción de los valores genéticos de cada animal. La diferencia es que en regresión aleatoria los elemento de la matriz $\mathbf{G}$, como fue antes mencionado, son las varianzas y covarianzas

genéticas entre los coeficientes de regresión y en funciones de covarianza $\mathbf{G}$ es reemplazada por la matriz de coeficientes de la función de covarianza, $\mathbf{C}$.

Como los efectos ambientales varían durante la lactancia, se supuso que los valores genéticos al igual que los efectos permanentes presentan un comportamiento similar, así Schaeffer y Dekkers (1994) propusieron incluir el uso de covariables en la parte aleatoria del modelo, dando lugar a los modelos de regresión aleatoria. En estos modelos se utiliza tanto el modelo de medidas repetidas, al considerar los efectos permanentes, como el multivariado en donde se toma como diferentes características a cada uno de los coeficientes de regresión de la ecuación de predicción. Lo anterior trae ventajas debido a que se obtienen los valores genéticos de los animales en toda la lactancia con un reducido número de ecuaciones.

Paralelamente al modelo de regresión aleatoria, Kirkpatrick *et al.* (1990) propuso el modelo de funciones de covarianza que permite realizar una mejor descripción del carácter en el tiempo de medición, permitiendo obtener las varianzas y covarianzas genéticas para cada día de lactancia. El modelo de funciones de covarianza puede generar las mismas varianzas y covarianzas que la regresión aleatoria, y además derivarse de las varianzas de la ecuación de predicción para la parte genética y de ambiente permanente del modelo.

2. Ecuaciones de los Modelos Mixtos

Las ecuaciones de los modelos mixtos propuestas por Henderson (1949), consideran los efectos no genéticos, fijos, y los valores genéticos de los animales, aleatorios, dentro del modelo. Las soluciones obtenidas son las mejores estimaciones (BLUE) y predicciones

(BLUP) linealmente insesgados de los efectos fijos y aleatorios respectivamente (Núñez y Rodríguez, 1995; Barreras, 1999).

Para una característica con n observaciones, p efectos fijos, q valores genéticos cuyos primer y segundo momento sean: $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$, $E(\mathbf{u}) = E(\mathbf{e}) = 0$, $\text{var}(\mathbf{e}) = \mathbf{R} = \mathbf{I}\sigma_e^2$, $\text{var}(\mathbf{u}) = \mathbf{G} = \mathbf{A}\sigma_a^2$, $\text{var}(\mathbf{y}) = \mathbf{V} = \mathbf{ZG}_0\mathbf{Z}' + \mathbf{R}$, $\text{cov}(\mathbf{y}, \mathbf{e}) = 0$. Donde $\mathbf{A}$ corresponde a la matriz de relaciones genéticas aditivas. El modelo lineal mixto, es representado como:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}$$

donde:

$\mathbf{y}$ = Vector de observaciones $n \times 1$.

$\mathbf{X}$ = Matriz que relaciona las observaciones con los efectos fijos $n \times p$.

$\boldsymbol{\beta}$ = Vector de efectos fijos $p \times 1$.

$\mathbf{Z}$ = Matriz que relaciona las observaciones con el valor genético $n \times q$.

$\mathbf{u}$ = Vector de los efectos genéticos $q \times 1$.

$\mathbf{e}$ = Vector de residuales $n \times 1$.

La forma general de las ecuaciones de modelos mixtos es la siguiente:

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}^0 \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \quad (3)$$

Generalmente $\mathbf{R}^{-1} = \mathbf{I}(1/\sigma_e^2)$ y $\mathbf{G}^{-1} = \mathbf{A}^{-1}(1/\sigma_a^2)$, entonces las ecuaciones de los modelos mixtos pueden ser multiplicadas por σ_e^2 , quedando de la siguiente forma:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\lambda \end{bmatrix} \begin{bmatrix} \beta^0 \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix} \quad (4)$$

En donde λ es el coeficiente entre la varianza residual y la varianza genética. Este sistema de ecuaciones puede ser representado por $\mathbf{Cs} = \mathbf{r}$, donde $\mathbf{C}$ se denomina matriz de coeficientes, y $\mathbf{s}$ y $\mathbf{r}$ son los vectores de soluciones y de elementos del lado derecho de las ecuaciones, respectivamente. La solución para valores conocidos de σ_e^2 y σ_a^2 , es $\mathbf{s} = \mathbf{C}^{-1}\mathbf{r}$.

En ganado lechero, es posible tener algunas modificaciones a las ecuaciones de los modelos mixtos, éstas dependen de cómo se considere a la producción de leche por día durante una lactancia. Las posibles alternativas son: a) una variable que se puede medir un gran número de veces, b) variables fenotípica y genéticamente correlacionadas, y c) una variable cuyo fenotipo y valor genético cambian durante el tiempo de su registro. De esta manera, el modelo correspondiente será: de repetibilidad, multivariado, de regresión aleatoria (RRM) y de funciones de covarianza (MFC).

En las formas anteriores de considerar la producción de leche, el modelo de repetibilidad supone que el valor genético de los animales para producción de leche es el mismo durante toda la lactancia, suposición que puede no ser cierta. Las restantes suponen la existencia de cambios graduales en el valor genético de los animales durante la lactancia cambiando la forma en que se modela. El análisis multivariado requiere de la estimación de demasiados parámetros genéticos, imposibilitando su uso rutinario. El modelo de regresión aleatoria predice los valores genéticos de los animales para producción de leche por día dentro del periodo de lactancia. Las funciones de covarianza predicen tanto los

valores genéticos de los animales, como las varianzas y covarianzas genéticas entre producción de leche por día en la lactancia (Van der Werf y Schaeffer, 1997).

Utilizando el modelo de repetibilidad se hace uso de la ecuación (4), incorporando la columna de los efectos permanentes, en donde se estima $\text{var}(u_j) = \sigma_a^2$, $\text{var}(p_j) = \mathbf{I}\sigma_p^2$, $\text{var}(e_{ijk}) = \mathbf{I}\sigma_e^2$. Para un modelo multivariado, se emplea la ecuación (3) en donde $\text{Var}(\mathbf{u}) = \mathbf{G}_0$, la matriz de varianzas y covarianzas genéticas entre las producciones diaria en las t etapas y $\text{var}(\mathbf{e}) = \mathbf{R}$, la matriz de covarianzas de los residuales que puede tomar diferentes formas, donde la más simple es $\mathbf{I}\sigma_e^2$, la varianza residual multiplicada por una idéntica de orden t que corresponde al orden del número de etapas consideradas en el modelo. En el modelo de regresión aleatoria de igual manera se utiliza la ecuación (3) con ciertas diferencias ya que la matriz $\mathbf{X}$ es una matriz de ceros y unos que adicionalmente contienen las columnas correspondientes a las covariables de los coeficientes de regresión de la ecuación de predicción de la curva de lactancia, en tanto para la matriz $\mathbf{Z}$, se incluyen las mismas columnas para las covariables pero para cada animal, de tal manera que los resultados obtenidos son los coeficientes de regresión con la cual se predicen los valores genéticos en toda la lactancia para cada uno de los animales. La $\text{Var}(\alpha) = \mathbf{G}$, es la matriz de varianzas y covarianzas genéticas de los coeficientes de regresión de la ecuación de predicción, y la $\text{var}(\mathbf{e}) = \mathbf{R}$, es la matriz de las varianzas y covarianzas de los residuales que generalmente es $\mathbf{I}\sigma_e^2$.

Para las funciones de covarianza se sigue un procedimiento similar al de la ecuación (3), con la diferencia de que en $\mathbf{G}$ se sustituyen las varianzas y covarianzas de los coeficientes de regresión por una ecuación que nos permite predecir las varianzas y covarianzas de

todos los días en lactancia. Así, $\hat{G} = \phi \hat{C} \phi'$, donde ϕ es una matriz cuyos elementos son los coeficiente de la ecuación de predicción, y C es la matriz de coeficientes de las estimaciones de las funciones de covarianza, mismas que son: $\hat{C} = \phi^{-1} \hat{G} \phi^T$.

3. Estimación de componentes de varianza

Uno de los elementos de las ecuaciones de modelos mixtos son las varianzas de los efectos aleatorios considerados en el modelo. El desarrollo teórico de la metodología de los modelos mixtos supone que esas varianzas son conocidas, ó al menos las proporciones; sin embargo, en la práctica, su valor es desconocido. Generalmente, se estiman los componentes de varianza, y se usan las estimaciones como si fueran los valores reales. Kackar y Harville (1981) mostraron que cuando las varianzas son estimadas con un método que cumpla con ciertas propiedades y esas estimaciones son sustituidas en las ecuaciones de los modelos mixtos, las estimaciones de los efectos fijos y los predictores de los efectos aleatorios son insesgados. Las propiedades que se requieren para este tipo de estudio son:

- a).- Las observaciones se distribuyan simétricamente, por ejemplo, $f(y) = f(-y)$.
- b).- Los estimadores de las varianzas sean de traslación invariante.
- c).- Los estimadores de las varianzas sean funciones de las observaciones.

Los procedimientos para estimar los componentes de varianzas han ido cambiando con el objetivo de optimizar las propiedades de los estimadores que producen, y de hacerlos más eficientes y versátiles desde el punto de vista computacional.

Los principales métodos para la estimación de los componentes de varianzas son los de Análisis de Varianza (ANOVA), los de Máxima Verosimilitud (ML) y los de Máxima Verosimilitud Restringida (REML).

Los métodos tipo ANOVA, consisten en igualar la suma de cuadrados del análisis de varianza con sus esperanzas. Originalmente fueron ocupados para datos balanceados (Cadena y Castillo, 2000), sin embargo, Henderson (1953), presentó tres diferentes formas de adaptarlos para ser utilizados en datos desbalanceados. Dichas adaptaciones se conocen como los tres métodos de Henderson. El método I se usa para modelos con únicamente efectos aleatorios. Con el método II se ajustan primero los datos por los efectos fijos en el modelo y posteriormente se usa el método I para estimar los componentes de varianza. El método III se usa para modelos mixtos y se basa en las sumas de cuadrados obtenidas al ajustar un modelo lineal y sus submodelos (suma de cuadrados reducido).

Los métodos de ANOVA suponen que la información es una muestra aleatoria y consideran únicamente una relación de parentesco, comúnmente, medios hermanos paternos (MHP), ignorando las relaciones aditivas entre los progenitores. Este primer supuesto generalmente no se cumple debido a que la información proviene de hatos sujetos a selección, además de que el procedimiento puede producir estimaciones de varianza negativas. Finalmente, carecen de propiedades de distribución bien definidas cuando los datos son desbalanceados (Núñez y Rodríguez, 1995).

Otros métodos que no presentan estas limitantes son los de Máxima Verosimilitud (ML) y Máxima verosimilitud Restringida (REML), y son los métodos más comúnmente utilizados en mejoramiento genético animal (Núñez y Rodríguez, 1995; Cadena y Castillo, 2000).

El procedimiento de ML fue propuesto por Fisher (1925), y aplicado por primera vez al modelo mixto general por Hartley y Rao (1967). Este método supone que los términos del error y los efectos aleatorios tienen una distribución normal.

De manera simple: $y_i = \mu + e_i$, donde y_i es la i -ésima observación, μ es una constante (la media) común en todas las observaciones, y e_i es la desviación aleatoria de la i -ésima observación con respecto a la constante μ , donde $e_i \sim \text{NID}(0, \sigma^2)$; $E(y_i) = \mu$. El procedimiento de ML requiere de maximizar la función de verosimilitud bajo una distribución normal multivariada. Este estimador produce un estimador sesgado de la varianza al estar dividido entre el número total de observaciones, es decir, no toma en cuenta los grados de libertad al estimar los efectos fijos (Meyer, 1989).

Con la finalidad de evitar este sesgo y producir un estimador de la varianza libre de los efectos fijos, se realiza el análisis de los datos sobre la transformación de las observaciones a $\mathbf{k}'\mathbf{y}$ en lugar de $\mathbf{y}$, donde $\mathbf{E}(\mathbf{k}'\mathbf{y}) = \mathbf{0}$, denominado análisis de residuales o contraste de error. Al derivar el logaritmo de la función de verosimilitud aplicada sobre $\mathbf{k}'\mathbf{y}$ se produce un estimador de la varianza insesgado al considerar los grados de libertad perdidos por la estimación de los efectos fijos, $n-1$, para este modelo en particular (Núñez y Rodríguez, 1995; Meyer y Hill, 1997).

Las matrices de coeficientes de las funciones de covarianza para los valores genéticos se estiman con REML maximizando $\log L = -\frac{1}{2}(\text{const} + \log|\mathbf{R}| + k\log|\mathbf{G}| + \log|\mathbf{C}|$

+ $y'Py$) donde $\sigma_e^2 = y'Py / (N-r(X))$. Es posible considerar heterogeneidad de varianzas, siendo R una diagonal con diferentes elementos (Meyer, 1989; 1991; 1998).

Efectos ambientales

Una de las ventajas del modelo del día de prueba (MDP) es la corrección que se realiza para los efectos fijos, principalmente para aquellos cuyo impacto cambia en el tiempo, adicionalmente se evita el uso de lactancias incompletas (Swalve, 2000).

En los modelos de evaluación genética, es común el uso del grupo contemporáneo hato-año-estación de parto (HYS) como una forma de corregir los efectos ambientales que inciden durante el parto. Así el grupo contemporáneo clasifica a los animales que parieron en el mismo hato, el mismo año y en la misma estación. Sin embargo, no se consideran los factores que inciden directamente durante el día de registro. El MDP sustituye el grupo HYS por el HTD (Hato día de prueba), que relaciona el efecto del hato y el día en el cual se está llevando a cabo el registro de la característica, ya que esta es más importante que la fecha de parto (Pander *et al.*, 1992; Stanton *et al.*, 1992), reduciendo la varianza residual hasta en un 24 % (Meyer *et al.*, 1989; Ptak y Shaeffer, 1993).

Parámetros genéticos

La formación de grupos contemporáneos considerando el día de registro de producción, permite considerar en el modelo la variabilidad existente entre y dentro de animales (Swalve, 1998; 2000). Al reducir la varianza residual se incrementan los valores de los parámetros genéticos para producción total y para los diferentes periodos durante la lactancia como puede observarse en los cuadros 2 y 3 (Jamrozik *et al.* 1997; Jamrozik y Schaeffer, 1997). Así mismo, permite estimar las correlaciones genéticas entre cada día

de la lactancia, las cuales son mayores entre periodos adyacentes y disminuyen conforme se incrementa el intervalo entre ellos, (Cuadro 4).

Los parámetros genéticos son de gran importancia porque permiten identificar el momento óptimo para seleccionar animales como padres de las siguientes generaciones (Jamrozik y Schaeffer, 1997). Aunado a estas ventajas el MDP permite incluir en los esquemas de selección algunas características de la curva de producción con importancia económica, como es el caso de la persistencia (Dekkers *et al.*, 1998).

Cuadro 2. Índices de herencia (h^2) y de ambiente permanente (EP), y varianzas genéticas (σ^2_a), de ambiente temporal (σ^2_E) y permanente (σ^2_{EP}), para producción de leche, grasa y proteína a los 305 días de lactancia.

	σ^2_a	σ^2_E	σ^2_{EP}	h^2	EP
Lactancia total (305 días)					
Producción de leche	2.7700	5.6900	3.2600	0.24	0.28
Producción de grasa	0.0038	0.0140	0.0059	0.16	0.25
Producción de proteína	0.0019	0.0073	0.0032	0.16	0.26
Días de prueba (8 registros de producción)	2.8800	4.7900	3.6700	0.25	0.32
Producción de leche	0.0039	0.0130	0.0065	0.16	0.27
Producción de grasa	0.0019	0.0064	0.0035	0.16	0.30
Producción de proteína					

Fuente: Swalve, 1995.

Persistencia

En el uso de lactancias extendidas no se consideran las características genéticas de los animales durante su periodo de medición, como es el caso de la persistencia (Jamrozik y Schaeffer, 1997), trayendo consigo predicciones sesgadas de los valores genéticos para producción total.

Una de las ventajas de la persistencia se encuentra en la reducción de los costos de producción que es debida al uso de una menor cantidad de concentrados utilizada en la alimentación animal (Sölkner y Fuchs, 1987; Dekkers *et al.*, 1998). Adicionalmente, la mayor persistencia también presenta una reducción en los problemas reproductivos y sanitarios que se deben a la reducción del desbalance energético negativo que caracteriza a los animales en los primeros meses después del parto, lo que permite una recuperación de la actividad reproductiva y de la condición corporal de manera oportuna. Esta situación puede ser aprovechada para incrementar la rentabilidad de la actividad lechera (Dekker *et al.*, 1998).

Cuadro 3. Estimación de varianzas genéticas (g) e índices de herencia (h^2) para producción diaria de grasa y proteína en leche.

Días en lactancia	Índices de herencia					
	Leche		Grasa		Proteína	
	g	h^2	g	h^2	g	h^2
5	16.11	0.59	4.96	0.68	2.49	0.69
10	7.66	0.40	2.47	0.51	0.87	0.44
15	7.54	0.40	1.94	0.45	0.61	0.35
20	8.04	0.41	1.73	0.42	0.54	0.33
25	8.41	0.44	1.60	0.42	0.52	0.34
45	8.55	0.46	1.26	0.40	0.48	0.34
85	7.72	0.45	0.90	0.37	0.49	0.35
125	7.23	0.45	0.77	0.34	0.56	0.39
165	6.67	0.43	0.73	0.35	0.62	0.42
205	6.01	0.42	0.72	0.36	0.64	0.44
245	5.85	0.41	0.79	0.40	0.68	0.46
285	7.36	0.43	1.05	0.41	0.87	0.47
305	9.24	0.48	1.29	0.47	1.07	0.52

Fuente: Fuente: Jamroziky Schaeffer, 1997.

Cuadro 4. Correlaciones genéticas estimadas (x 100) de producción de grasa (debajo de la diagonal) y proteína (arriba de la diagonal) en vacas Holstein de primer parto.

	5	25	45	85	125	165	205	265	285	305
5		53	35	25	23	22	19	9	5	1
25	26		96	80	58	39	23	7	4	2
45	3	92		92	74	54	37	15	9	4
85	3	61	85		94	81	65	34	23	14
125	11	33	60	93		96	85	53	40	28
165	17	12	38	79	96		96	71	59	46
205	19	-4	19	62	85	96		87	78	67
265	14	-21	-10	23	51	69	86		98	94
285	10	-24	-17	8	38	54	75	98		99
305	5	-26	-23	-5	22	38	61	93	98	

Fuente: Jamroziky Schaeffer, 1997.

Experiencias con MFC y RRM

El MDP es utilizado en diferentes países como Australia, Nueva Zelanda, Alemania, España y recientemente en Chile, siendo Canadá el país que está usando oficialmente para realizar las evaluaciones genéticas de ganado lechero (Schaeffer *et al.*, 2000; Swalve, 2000; Uribe, 2001).

El MDP ofrece la oportunidad de incrementar la ganancia genética al predecir los valores genéticos en toda la lactancia, y de realizar evaluaciones a una edad más temprana sin la necesidad del uso de lactancias extendidas que pueden sesgar el valor genético de la producción. La limitante es conocer la ecuación de predicción que más se ajuste a los registros observados. Para producción de leche diferentes autores mencionan que la ecuación propuesta por Ali y Schaeffer (1987), y la de Wilmink (1987), han sido las más adecuadas.

Esta metodología puede ser utilizada en cualquier característica que sea medida en forma longitudinal en el tiempo, como es el caso de las características productivas en los animales (ganancia diaria de peso, grasa dorsal, producción de leche, consumo de alimento, y producción de huevo, entre otras).

4. Literatura citada

- Ali, T. E., And R. Schaeffer. 1987. Accounting for covariances among test day milk yields in dairy cows. *Can. J. Anim. Sci.* 67:637-644.
- Barreras, S. A. 1999. Aspectos genéticos y fenotípicos de la forma de la curva de lactancia en cabras. Tesis de Maestría en Ciencias. Producción Animal. Chapingo México. p 25-28.
- Cadena, M. J. A., y A. Castillo M. 2000. Una comparación de la estimación de componentes de varianza mediante simulación. *Agrociencia.* 34:343-352.
- Dekkers, J. C. M., J. H. Ten Hag, and A. Weersink. 1998. Economic aspects of persistency of lactation in dairy cattle. *Livest. Prod. Sci.* 53:237-252.
- Ferris, A. T. 1981. Selecting for lactation curve sharp and milk yield in dairy cattle a dissertation submitted to Michigan state University. Doctor of philosophy department of dairy sci. Michigan State university.
- Fisher, R. A. 1925. Statistical methods for research workers. Oliver and Boyd. Edimburgh. UK. 362 p.
- Freeman, A. E. 1998. Dairy cattle Breeding. *Proc. 6th World Congr. Genetic Appl. Livest. Prod.* Armidale, New South Wales, Australia. 23:292-294
- Guo, Z., and H. H. Swalve. 1995. Modelling of the lactation curve as sub-model in the evaluation of test day records. *Procedings of the open Session of the Interbull Annual Meeting.* Prague, Czech Republic. September 7-8. Bulletin No 11.1-4

- Hartley , H. O., and N. K. Rao. 1967. Maximun likelihood estimation for the mixed analysis of variance model. *Biométrieka*. 54:93-108.
- Henderson, C. R. 1949. Estimates of Changes in herd environment. *J. Dairy Sci*. 32:706.
- Henderson, C. R. 1953. Estimation of variance and covariance components. *Biometrics*. 9:226-252.
- Henderson, C. R. 1974. General flexibility of linear techniques for sire evaluation. *J. Dairy Sci*. 57:963-972.
- Henderson, C. R. Jr. 1982. Analysis of covariance in the mixed model: higher-level, nonhomogeneous and random regression. *Biometrics*. 28:623-640.
- Jamrozik, J., and L. R. Schaeffer. 1997. Estimates of genetics parameters for test day model with random regression for yield traits of first lactations in Holsteins. *J. Dairy Sci*. 80:762-770.
- Jamrozik, J., L. R. Schaeffer, and J. C. M. Dekkers. 1997. Genetic evaluation of dairy cattle using test day yields and random regression model. *J. Dairy Sci*. 80:1217-1226.
- Kackar, R. N., and A. Harville. 1981. Unbiasness of two-stage estimation and prediction procedures for mixed linear models. *Commun. Stat. A:Theory & Methods*. 10:1249.
- Kirkpatrick, M., D. Lofsvold, and M. Bulmer, 1990. Analysis of the inheritance selection and evolution of growth trayectories. *Genetics*. 124:979-993.

- Kirkpatrick, M., W. G. Hill, and R. Thompson. 1994. Estimating the covariance structure of traits during growth and ageing, illustred with lactating dairy cattle. *Gen. Res.* 64:57-69.
- Meyer, K. 1989. Estimation of genetic parameters. Reviews on molecular and quantitative approaches in honour of A. Robertson. W. G. Hill and T. F. M. Mckay, eds., pp. 161-167-199.
- Meyer, K. 1991. Estimating variance and covariance for multivariate animal models by restricted maximun likelihood. *Genet. Sel. Evol.* 23:67-83.
- Meyer, K. 1998. Derivate Free Restricted Maximun Likelihood (DFREML). User notes. 28 p.
- Meyer, K., and W. G. Hill. 1997. Estimation of genetic and phenotypic covariance functions for longitudinal or repeated records by restricted maximun likelihood. *Livest. Prod. Sci.* 47:185-200.
- Meráz, A., M. R., 2001. Comparación de ecuaciones para estimar curvas de lactancia con diferentes frecuencia de medición en bovinos Angus, Suizo y sus cruza. Tesis de Maestría en Producción Animal. Departamento de Zootecnia. Universidad Autónoma Chapingo. Chapingo, México. P:3-9.
- Núñez, D. R., y F. A. Rodríguez A. 1995. Componentes de Varianza. Memorias del curso intensivo "Introducción a la metodología de los modelos mixtos (Aplicaciones al Mejoramiento Animal)". Universidad Autónoma de Chihuahua. Facultad de zootecnia. Chihuahua, Méx. III:1-54.

- Pander, B. L., W. G. Hill, and R. Thompson. 1992. Genetic parameters of test day records of British Holstein-Friesian heifers. *Anim. Prod.* 55:11-21.
- Ptak, E., and L. R. Schaeffer. 1993. Use of test day yields for genetics evaluation in dairy sires and cows. *Livest. Prod. Sci.* 34:23-34.
- Schaeffer, L. R., and J. C. M. Dekkers. 1994. Random regression in animal models for test day production in dairy cattle. *Proc. 5th World Congre. Genet. Appl. Livest. Prod.* Guelph, ON, Canada. 18:443-446.
- Schaeffer, L. R., J. Jamrozik, and G. J. Kistemaker. 2000. Experience with a test-day model. *J. Dairy Sci.* 83:1135-1144.
- Snedecor, W., y W.G Cochran. 1979. *Métodos Estadísticos*. Edit. CECSA. p 191-197.
- Sölkner, J., and W. Fuchs. 1987. A comparison of different measures of persistency with special respect to variation of test-day milk yields. *Livest. Prod. Sci.* 16:305-319.
- Stanton, T. L., L. R. Jones, R. W. Everett and S. D. Kachman. 1992. Estimating milk, fat and protein lactation curves with a test day model. *J. Dairy Sci.* 75:1691-1700.
- Swalve, H. H. 1998. Use of test day records in genetics evaluation. *Proc. 6th World Congr. Genet. Appl. Livest. Prod.* Armidale, New South Wales, Australia 23:295-302.
- Swalve, H. H. 2000. Theoretical basis and computational methods for different test day genetic evaluation methods. *J. Dairy Sci.* 83:1115-1124.
- Uribe, H. A. 2001. Modelando el día de control; nueva técnica estadística en evaluaciones genéticas de ganado bovino lechero. *Agricultura Técnica.* 61:500-511.

Van Der Werf, J., and L. Schaeffer. 1997. Random regression in animal breeding. Course Notes p:1-59

Wilmink, J. B. M. 1987. Adjustment of test day milk, fat, and protein yields for age, season and stage of latation. Livest. Prod. Sci. 16:335.

CAPÍTULO 3

MODELOS DE FUNCIONES DE COVARIANZA EN LA ESTIMACIÓN DE PARÁMETROS Y VALORES GENÉTICOS PARA PRODUCCIÓN DE LECHE, GRASA Y PROTEÍNA¹

¹ Artículo de Canche-Uc J. G, C.A Apodaca-Sarabia.; R. Ramirez-Valverde, J. A. Cadena-Meneses, para enviarse a la revista Veterinaria México.

Resumen

El modelo de funciones de covarianza (MFC) es una de las formas del modelo del día de registro. Sus ventajas aplicadas a características de rendimiento por día en la lactancia permiten la descripción genética del carácter en el trayecto de su expresión con menor número de parámetros a estimar que los requeridos en el modelo multivariado. La reducción se logra al usar ecuaciones de regresión para predecir las covarianzas genéticas del carácter y los valores genéticos de los animales en la lactancia. La ecuación de regresión por animal permite predecir los valores genéticos de los animales para: a) el comportamiento por día en cualquier día de la lactancia, b) la producción acumulada por lactancia o en cualquier periodo de la misma, y c) cualquier definición de persistencia. Además los modelos del día de registro permiten con relativa facilidad la inclusión de efectos no genéticos sistemáticos sobre parte de lactancia y evita la proyección de lactancias incompletas, mejorando así la exactitud del valor genético predicho. Los objetivos del presente trabajo fueron discutir el MFC usando un número limitado de registros de rendimiento diario de leche, grasa y proteína. Las estimaciones de heredabilidad de las tres características presenta la posibilidad de lograr mayores respuesta a la selección al inicio y final de la lactancia así como la gran facilidad de detectar animales con diferencias genéticas importantes durante la lactancia.

Palabras clave: día de pesada, ganado lechero, parámetros genéticos, valores genéticos.

ABSTRACT

The covariance function model (MFC) is one of the forms of the test day model. The MFC has advantages if it is applied to characteristics of yield per day in the lactation, allowing the genetic description of the trait in the passage of its expression with a smaller number of parameters to estimate than the required ones in the multivariate model. Reduction is accomplished when using regression equations to predict the genetic covariances of the trait and the breeding values of the animals in the lactation. The regression equation by animal allows to predict breeding values of the animals for: a) the performance per day in any lactation day, b) the accumulated production by lactation or in any period of the same one, and c) any definition of persistence. In addition, the test day models allow the easy inclusion of systematic nongenetic effects on part of the lactation, and avoid the projection of incomplete lactation, improving the accuracy of the predicted breeding values. The objectives of this study were to discuss the possible application of the MFC, using a limited number of daily yield milk, fat and protein records of Holstein cows. Estimated heritabilities of the three traits display the possibility for accomplish larger responses to the selection at the beginning and at the end of the lactation, as well as the effortless to detect animals with important genetic differences during the lactation.

Keywords; test day, dairy cattle, genetic parameters, breeding values.

Introducción

Las evaluaciones genéticas en producción de leche, tradicionalmente se han realizado para producción acumulada en una lactancia de 305 días, usando registros en días específicos de la lactancia (Misztal *et al.*, 2000), sin considerar los factores que modifican la producción por día del registro y suponiendo la misma curva de lactancia para todos los animales cuando se proyectan lactancias. Sin embargo, si los animales tienen diferentes curvas de lactancia, la evaluación genética podría ser sesgada (Jamrozik y Schaeffer, 1997). El modelo del día de registro (MDP) se caracteriza por que la variable de estudio es la producción por día, siendo entonces posible modelar la producción en toda la lactancia y por lo tanto la inclusión de factores que influyeron sobre el comportamiento animal el día de su registro.

Uno de los primeros trabajos en considerar los efectos genéticos y no genéticos sobre el comportamiento animal el día de su registro fue el de Ali y Schaeffer (1987), quienes al incluyeron regresiones individuales y las covarianzas entre las producciones por día buscando mejorar la predicción. Meyer (1989) propuso el modelo de repetibilidad suponiendo que los valores genéticos y de ambiente permanente de los animales no cambian durante la lactancia, incluyendo además covariables, con la finalidad de ajustar por la etapa de lactancia. Trus y Buttazzoni (1990) usaron las producciones por día con la finalidad de incluir efectos como el avance de la preñez y de la lactancia, y la edad de la vaca, y así el cálculo de la producción por lactancia y registros faltantes sea realizado en un solo análisis considerando los efectos antes mencionados sobre la producción por día. Freeman (1998) mostró las diferentes formas en que se ha usado el MDP, entre ellas, la

regresión aleatoria y las funciones de covarianza, que tienen la capacidad de describir genéticamente al carácter en el periodo de su expresión.

Asumiendo que la producción de leche por día se observa repetidamente en la lactancia, una observación registrada el i -ésimo día de pesada, en la j -ésima vaca, el día t de la lactancia (y_{ijt}) podría representarse en el siguiente modelo mixto: $y_{ijt} = DP_i + \sum_{l=0}^k \beta_l t_l + u_j + p_j + e_{ijt}$, donde los efectos del día de prueba (DP) y las covariables (β_l) son considerados fijos, y corresponden a los efectos de las diferentes fechas de medición y etapas de lactancia, y u_j , p_j y e_{ijt} son los valores genéticos, de ambiente permanente y los residuales, respectivamente. En este modelo $E(y_{ijt}) = DP_i + \sum_{l=0}^k \beta_l t_l$, $E(u_j) = E(p_j) = E(e_{ijk}) = 0$, $\text{var}(u_j) = \sigma_a^2$, $\text{var}(p_j) = I\sigma_p^2$, $\text{var}(e_{ijk}) = I\sigma_e^2$ y $\text{cov}(u_j, p_j) = 0$ (Núñez y Rodríguez, 1995). El modelo anterior corresponde a un modelo de repetibilidad, asumiendo el mismo valor genético y de ambiente permanente de los animales para cada día en la lactancia (Meyer, 1989), por lo que los cambios observados en la producción por día durante la lactancia son causados por los efectos temporales, como la etapa de lactancia por ejemplo. Lo anterior puede no cumplirse si los genes responsable de la producción de leche no son los mismos durante la lactancia, es decir, que el valor genético y el efecto de ambiente permanente del j -ésimo animal cambien durante la lactancia.

Para considerar lo anterior, los valores genéticos y de ambiente permanente se pueden incluir en la ecuación de regresión que los predice, $\sum_{l=0}^k \alpha_l t_l$ y $\sum_{l=0}^k \gamma_l t_l$, respectivamente, resultando en el modelo conocido como de regresión aleatoria: $y_{ijt} = DP_i + \sum_{l=0}^k \beta_l t_l + \sum_{l=0}^k \alpha_l t_l + \sum_{l=0}^k \gamma_l t_l + e_{ijk}$, donde $\text{var}(\sum_{l=0}^k \alpha_l t_l) = \mathbf{G}_0$, $\text{var}(\sum_{l=0}^k \gamma_l t_l) = \mathbf{P}_0$, y

las restantes son iguales al modelo anterior. Los elementos de $\mathbf{G}_0$ y $\mathbf{P}_0$ son las covarianzas genéticas y de ambiente permanente entre los coeficientes de regresión respectivos (Schaeffer y Dekkers, 1994; Jamrozik *et al.* 1997; Jamrozik y Scheffer, 1997).

En características que pueden ser medidas repetidamente en el tiempo, el cambio genético en la trayectoria de expresión, se puede definir por dos funciones continuas, la de covarianzas genéticas y la del gradiente de selección: $\Delta_G = \int G(a,x) \phi(x) dx$. La función de covarianzas describe genéticamente la característica en el periodo de medición y fue descrita por Kirkpatrick *et al.* (1990) usando polinomios ortogonales de Legendre $\phi(a_i^*)$, definidos en las variables estandarizadas a_i^* , y elementos de $\mathbf{C}_G$: $G(a,x) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} [\mathbf{C}_G]_{ij} \phi(a_1^*) \phi(a_2^*)$.

Para estimar $\mathbf{C}_G$ es necesario expresar la matriz de covarianzas genéticas entre las observaciones en diferentes tiempo de registro, $\mathbf{G}_0$, en función de $\mathbf{C}_G$ y de los polinomios de Legendre, es decir, $\mathbf{G}_0 = \phi \mathbf{C}_G \phi'$, así $\mathbf{C}_G = \phi^{-1} \mathbf{G}_0 \phi^{-T}$, donde ϕ es la matriz de los coeficientes de los polinomios ortogonales de Legendre. Esta metodología usa el procedimiento de cuadrados mínimos, que tiene la desventaja principal de requerir la estimación previa de $\mathbf{G}_0$ (Kirkpatrick, *et al.*, 1990; Meyer y Hill, 1997). Una forma de evitar la estimación previa de $\mathbf{G}_0$, es usar el procedimiento de máxima verosimilitud restringida (REML), modificando el modelo de regresión aleatoria y usando como covariables fijas y aleatorias los polinomios de Legendre, lo que puede ser representado en el siguiente modelo: $y_{ijt} = DP_i + \sum_{l=0}^k \beta_{tl} + \sum_{l=0}^k \alpha_{il} a_i^* + \sum_{j=0}^k \gamma_{ij} p_j + e_{ij}$. Esto es equivalente a un modelo multivariado en donde la sumatoria es representada por ϕ , así expresando en matrices para evitar la sumatoria, $\text{var}(\Phi_i a_i) = \Phi \text{var}(a_i) \Phi' = \Phi \mathbf{C}_G \Phi' = \mathbf{G}_0$ y $\text{var}(\Phi p_j) = \Phi \mathbf{C}_P \Phi' =$

P_0 , donde C_G y C_P son matrices de coeficientes de las funciones de covarianza genética y de ambiente permanente, respectivamente. La covarianza entre dos registros es entonces

$$\text{Cov}(y_{ij}, y_{ij'}) = \sum_{i=0}^{k-1} \sum_{i'=0}^{k-1} \phi_i(a^*_{ij}) \phi_{i'}(a^*_{ij'}) \text{cov}(\alpha_{ij}, \alpha_{ij'}) + \sum_{i=0}^{k-1} \sum_{i'=0}^{k-1} \phi_i(a^*_{ij}) \phi_{i'}(a^*_{ij'}) \text{cov}(\gamma_{ij}, \gamma_{ij'}) + \text{cov}(e_{ij}, e_{ij'}) = G(a^*_{ij}, a^*_{ij'}) + P(a^*_{ij}, a^*_{ij'}) + R(e_{ij}, e_{ij'}),$$

las funciones de covarianza de los componentes de varianza.

Si se supone los e_{ij} son idénticamente distribuidos, con varianza σ_e^2 , la covarianza entre los errores es entonces la varianza del error si $j=j'$, y es igual a cero para el resto. Las matrices de coeficientes de las funciones de covarianza para los valores genéticos y de ambiente permanente se estiman con REML, maximizando $\log L = -\frac{1}{2}(nt(\log|C_G| + k\log|A| + nr(\log|C_P| + \log|C| + \log|R| + y'Py))$, donde $\sigma_e^2 = y'Py / (N-r(X))$. Sin embargo, es posible considerar heterogeneidad de varianzas y R es entonces diagonal con diferentes elementos (Meyer, 1989; 1991; 1998).

Lo anterior permite el uso de cualquier ecuación para predecir las curvas de lactancia (Swalve, 2000). Ali y Schaeffer (1987), Jamrozik *et al.* (1997) y Jamrozik y Schaeffer (1997) compararon diferentes ecuaciones, concluyendo que las mejores ecuaciones que generan menores valores en los residuales son la propuesta por Ali y Schaeffer (1987), en donde $y_t = a_0 + a_1(t/305) + a_2(t/305)^2 + a_3 \ln(305/t) + a_4(\ln(305/t))^2$ con cinco coeficientes de regresión, y la propuesta por Wilmink (1987) con tres coeficientes de regresión, $y_t = a_0 + a_1 t + a_2 e^{-0.05t}$. Lo anterior también fue señalado por Guo y Swalve (1995), quien comparó 16 diferentes ecuaciones, no encontrando diferencias importantes entre estas dos ecuaciones de predicción.

Los modelos de regresión aleatoria y funciones de covarianza proporcionan, para cada animal, una ecuación de predicción para calcular el valor genético en cualquier día de la lactancia. Esta puede ser usada para obtener el valor genético para producción total en la lactancia, para producción en cualquier parte de la lactancia y para diferentes definiciones de persistencia. Jamrozik *et al.* (1997) obtuvieron el valor genético para producción en tres etapas de 100 días cada una, y para persistencia definida como 110 veces la diferencia entre los valores genéticos del día 280 y el día 60, $P = (Vg_{280} - Vg_{60})110$. La importancia biológica y económica de la persistencia fue discutida por Sölkner y Fuchs (1987) y Dekkers *et al.* (1998).

No obstante las ventajas demostradas por la metodología del MDP, en México su conocimiento y utilización es limitado. Por lo anterior, el objetivo del presente trabajo fue mostrar la factibilidad de su utilización en la estimación de parámetros genéticos y la predicción de valores genéticos de la producción de leche, grasa y proteína para diferentes días en lactancia, para producción total por lactancia y para persistencia utilizando un modelo de funciones de covarianza con un pequeño número de registros en vacas Holstein.

Materiales y Métodos

Por la baja disponibilidad de información sobre producción de leche, grasa y proteína, se usaron los registros de rendimiento diario (colectados de septiembre del 2000 a agosto del 2001) de 173 animales Holstein de la Granja Experimental de la Universidad Autónoma Chapingo, localizada en una región de clima templado en Chapingo, México, bajo un régimen de estabulación total. La producción por día fue la suma de las producciones de mañana y tarde. El porcentaje de grasa se determinó por el método

Gerber y para proteína se usó microkiejhdal. Los rendimientos de grasa y proteína se obtuvieron multiplicando la densidad promedio (estimado de 60 vacas) por el porcentaje y producción de leche respectivo.

El modelo estadístico utilizado fue:

$$y_{ijk} = AENT_i + \sum_{l=0}^4 \beta_l X_{il} + \sum_{l=0}^4 \alpha_l X_{il} + \sum_{l=0}^4 \gamma_l X_{il} + e_{ijk}$$

donde:

y_{ijk} = Producción por día de leche, grasa o proteína en la i-ésima clase año-época-lactancia-día de registro de la j-ésima vaca en el k-ésimo día de lactancia.

$AENT_i$ = Efecto de la i-ésima clase año-época-lactancia-día de registro, $i=1,2,\dots,616$.

β , α , y γ = Coeficientes de regresión parcial que cuantifican el efecto promedio del día de lactancia y predicen los valores genéticos de los animales y los efectos de ambiente permanente de las vacas con registros.

X_j = Los valores de las covariables, $X_0 = 1$, $X_1 = t/305$, $X_2 = X_1^2$, $X_3 = \ln(305/t)$, $X_4 = X_3^2$, t = el día de lactancia.

e_{ijk} = El efecto residual.

Las suposiciones para este modelo son: $E(y_{ijk}) = AENT_i + \sum_{l=0}^4 \beta_l X_{il}$, $\text{var}(\alpha_l X_{il}) = \mathbf{G}_0$, $\text{var}(\gamma_l X_{il}) = \mathbf{P}_0$, $\text{var}(e_{ijk}) = I\sigma_e^2 = \mathbf{R}$, las covarianzas entre los residuales son cero, y $\mathbf{G}_0$ y $\mathbf{P}_0$ son las matrices de varianzas y covarianzas genéticas y de ambiente permanente entre las producciones diarias en la lactancia predichas con las respectivas funciones de

covarianzas. El análisis se realizó con el programa de máxima verosimilitud restringida libre de derivadas (DFREML) propuesto por Meyer (1998).

Utilizando las ecuaciones de predicción para los valores genéticos de producción por día en la lactancia se obtuvieron los valores genéticos para producción de proteína el día 60 (Vg_{60}) y 280 (Vg_{280}), la acumulada en el primero (Vg_{1T}), segundo (Vg_{2T}) y tercer tercio (Vg_{3T}) de la lactancia, total por lactancia (Vg_T), y la persistencia definida como: 1) $Vg_{P1} = (Vg_{280}/Vg_{60})100$ (Shanks *et al.* 1981; Bar-Anan *et al.* 1985; Tsuruta *et al.* 2000). Los valores genéticos antes mencionados (una vez obtenidos los coeficientes de regresión respectivos) para el animal j -ésimo son:

$$Vg_{ij} = \alpha_{0j} + \alpha_{1j}(i/305) + \alpha_{2j}(i/305)^2 + \alpha_{3j}\ln(305/i) + \alpha_{4j}(\ln(305/i))^2, i = 60, 280.$$

$$Vg_{1Tj} = \sum_{i=1}^{100} Vg_{ij} = \sum_{i=1}^{100} (\alpha_{0j} + \alpha_{1j}(i/305) + \alpha_{2j}(i/305)^2 + \alpha_{3j}\ln(305/i) + \alpha_{4j}(\ln(305/i))^2) = 100\alpha_{0j} + 16.557377\alpha_{1j} + 3.6371943\alpha_{2j} + 208.291802\alpha_{3j} + 519.121737\alpha_{4j}.$$

$$Vg_{2Tj} = \sum_{i=101}^{200} Vg_{ij} = \sum_{i=101}^{200} (\alpha_{0j} + \alpha_{1j}(i/305) + \alpha_{2j}(i/305)^2 + \alpha_{3j}\ln(305/i) + \alpha_{4j}(\ln(305/i))^2) = 100\alpha_{0j} + 49.3442623\alpha_{1j} + 25.2442892\alpha_{2j} + 72.538566\alpha_{3j} + 56.5000018\alpha_{4j}.$$

$$Vg_{3Tj} = \sum_{i=201}^{305} Vg_{ij} = \sum_{i=201}^{305} (\alpha_{0j} + \alpha_{1j}(i/305) + \alpha_{2j}(i/305)^2 + \alpha_{3j}\ln(305/i) + \alpha_{4j}(\ln(305/i))^2) = 100\alpha_{0j} + 87.0983607\alpha_{1j} + 73.2857296\alpha_{2j} + 20.3902642\alpha_{3j} + 5.49769157\alpha_{4j}.$$

$$Vg_T = \sum_{i=1}^{305} Vg_{ij} = \sum_{i=1}^{305} (\alpha_{0j} + \alpha_{1j}(i/305) + \alpha_{2j}(i/305)^2 + \alpha_{3j}\ln(305/i) + \alpha_{4j}(\ln(305/i))^2) = 305\alpha_{0j} + 153\alpha_{1j} + 102.167213\alpha_{2j} + 301.220632\alpha_{3j} + 581.11943\alpha_{4j}.$$

Para explicar las diferencias en valor genético de los animales durante la lactancia se consideraron tres animales con Vg_T lo más parecido posible y diferentes valores genéticos para persistencia. Es importante considerar que las predicciones y estimaciones de los valores genéticos y componentes de varianza, por el reducido número de registros

deben tener grandes errores estándar. Sin embargo, son un buen ejemplo del uso potencial de la metodología.

Resultados y discusión

Las varianzas y covarianzas genéticas y fenotípicas estimadas para producción de leche, grasa y proteína se muestran en el cuadro 5.

El modelo de funciones de covarianza permite la predicción de las varianzas y covarianzas genéticas y de ambiente permanente entre los días de lactancia, así, usando el modelo con cinco coeficientes de regresión (Ali y Schaeffer, 1987) es posible obtener la matriz $\mathbf{G}_0$ de orden 305, con una matriz, $\mathbf{C}_G$ de orden cinco. Lo anterior significa un ahorro importante en los requerimientos de tiempo y memoria de computo y un aumento en la posibilidad de lograr convergencia, sobre todo en los casos en que las correlaciones genéticas entre los registros en el tiempo tiende a uno (Swalve, 2000).

Las covarianzas estimadas fueron diferentes a las encontradas en la literatura. Las covarianzas genéticas y fenotípicas negativas entre diferentes días de lactancia, pueden deberse a la gran cantidad de parámetros estimados con un reducido número de registros. Resultados similares han sido reportadas por Jamrozik y Schaeffer (1997), quienes obtuvieron covarianzas genéticas negativas entre el primer tercio de la lactancia (25, 45 y 85 días), y el último (205, 265 y 305 días), mientras que otros autores (Pander *et al.*, 1992; Swalve, 1995; Swalve, 1998), reportan covarianzas y correlaciones positivas en todo el periodo de medición, con altas correlaciones (>0.90) entre periodos adyacentes, que disminuyen conforme el intervalo entre ellos se incrementa, lo que es similar a los resultados derivados del presente trabajo.

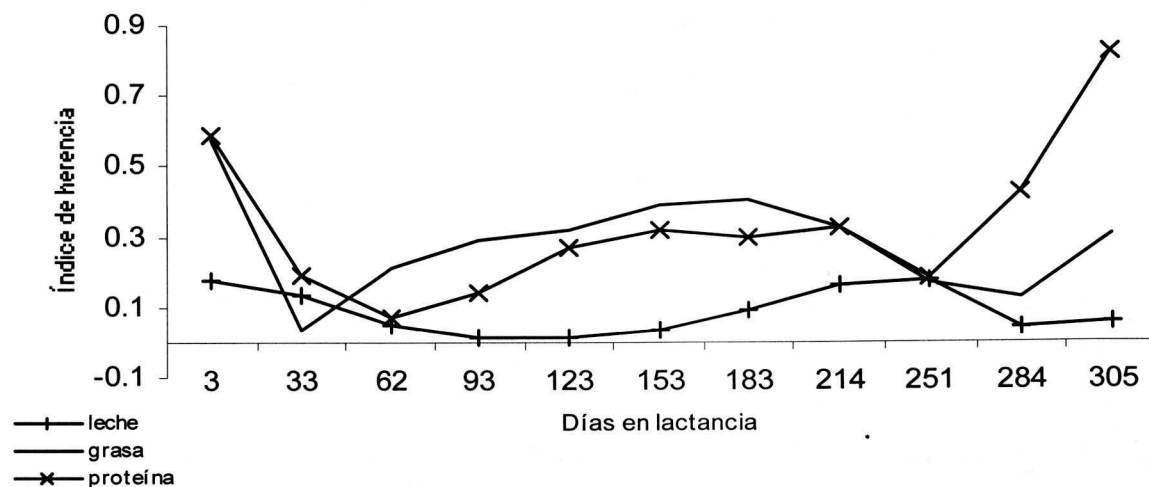
Cuadro 5. Varianzas, Covarianzas genéticas (bajo la diagonal) y fenotípicas (sobre la diagonal) entre las producciones diarias de leche, grasa y proteína en la lactancia.

Producción de leche								
DL	3	63	123	183	243	303	Varianza	
							Genética	Fenotipia
3		37.474	35.065	35.472	22.470	40.384	22.834	127.028
63	6.740		30.267	22.664	16.392	20.988	2.040	42.454
123	0.839	0.112		24.623	19.162	19.667	0.396	36.195
183	-5.100	-1.819	0.657		22.340	19.009	3.094	33.667
243	-8.758	-2.903	0.534	3.938		15.402	5.368	28.894
303	4.728	1.337	0.333	-0.686	-1.439		1.049	26.584
Producción de grasa								
3		-0.0178	-0.0150	0.0109	0.0108	0.0007	0.0840	0.1470
64	-0.0308		0.0173	0.0051	0.0019	-0.0163	0.0119	0.0533
123	-0.0211	0.0075		0.0186	0.0097	-0.0037	0.0162	0.0504
183	0.0000	-0.0012	0.0147		0.0166	0.0024	0.0211	0.0523
243	-0.0021	-0.0002	0.0094	0.0130		0.3788	0.0081	0.0407
303	0.0147	-0.0038	0.0048	0.0088	0.0042		0.0152	0.0537
Producción de proteína								
3		0.0170	0.0312	0.0208	-0.0059	0.1145	0.0858	0.1451
64	0.0155		0.0222	0.0119	0.0091	0.0496	0.0031	0.0427
123	0.0320	0.0055		0.0161	0.0081	0.0639	0.0123	0.0458
183	0.0154	0.0013	0.0075		0.0133	0.0189	0.0108	0.0363
243	-0.0101	-0.0033	-0.0020	0.0064		-0.0157	0.0095	0.0368
303	0.1243	0.0252	0.0431	0.0074	-0.0297		0.2075	0.2598

Los índices de herencia para producción por día de leche, grasa y proteína durante la lactancia mostraron variabilidad, siendo de mayor magnitud al inicio y al final de la lactancia (figura 1). comportamientos similares fueron reportados por Jamrozik y Schaeffer (1997), quienes obtuvieron altos valores al inicio y al final de la lactancia. Sin embargo,

Pander *et al.* (1992) y Swalve (1995) obtuvieron resultados diferentes al obtener bajos índices de herencia al inicio y al final de la lactancia.

Figura 1. Índice de herencia para producción de leche, grasa y proteína durante la lactancia en vacas Holstein.



En general, los altos valores al inicio y al final de la lactancia para todas las variables estudiadas pueden explicarse parcialmente por el bajo efecto no genético en el comportamiento animal durante esos periodos y a la tendencia general de la metodología de regresión (Draper y Smith, 1980) para producir mayor sesgo en los extremos del periodo de registro. La variabilidad mostrada durante la lactancia, confirma la posible ventaja de utilizar la metodología de funciones de covarianza y el considerar la suposición de varianzas heterogeneas de los residuales durante la lactancia (Jamrozik *et al.*, 1997; Swalve, 2000).

Un ejemplo de las diferencias en valores genéticos durante la lactancia, se muestran en los Vg_{60} , Vg_{280} , Vg_{1T} , Vg_{2T} , Vg_{3T} y Vg_{P1} de producción de proteína de tres animales (Cuadro 6). Normalmente las decisiones sobre la selección de animales se basan sobre la

producción total por lactancia, sin tomar en cuenta el cambio del valor genético de los animales durante la lactancia (Jamrozik y Schaeffer, 1997). En los resultados obtenidos, si consideramos que el criterio para seleccionar animales es la producción total por lactancia se seleccionaría al animal A, por presentar el valor genético más alto para producción total (Cuadro 6). Sin embargo, este animal no siempre es el mejor su jerarquización al considerar su valor genético en diferentes partes de la lactancia, de esta manera, para el tercer tercio de la lactancia y para el Vg_{280} ocupa el segundo lugar.

Cuadro 6. Valores genéticos de producción de proteína: los días 60 (Vg_{60}) y 280 (Vg_{280}), total por lactancia (Vg_T), en el primer (Vg_{1T}), segundo (Vg_{2T}) y tercer (Vg_{3T}) tercio de la lactancia, y de persistencia ($Vg_{P1}=(Vg_{280}/Vg_{60})100$).

Animal	Valores genéticos						Persistencia
	Kilogramos						%
	Vg_{60}	Vg_{280}	Vg_T	Vg_{1T}	Vg_{2T}	Vg_{3T}	Vg_{P1}
A	0.2237	0.0981	57.25	35.12	11.83	10.29	43.84
B	0.1900	0.1308	55.77	28.80	13.36	13.61	68.86
C	0.2189	0.0773	53.21	34.77	10.28	8.17	35.34

Por su parte, el animal B si se considera el Vg_T , ocupa el segundo lugar; Sin embargo, el primero para Vg_{280} , Vg_{2T} y Vg_{3T} y el tercero para su Vg_{60} y Vg_{1T} , y para el animal C este tendría el tercer lugar para Vg_T , Vg_{280} , Vg_{2T} y Vg_{3T} y el segundo para Vg_{60} y Vg_{2T} .

Estos resultados son de gran importancia debido a que brinda la oportunidad de seleccionar animales para un punto específico o periodo de la lactancia, y uno de las características de gran importancia económica es el considerar el grado de persistencia de los animales (Sölkner y Fuchs, 1987; Dekkers *et al.*, 1998; Jamrozik y Schaeffer, 1997). En

este ejemplo, pudiéramos elegir al animal A como el mejor semental si el criterio es producción total, sin embargo, al considerar la persistencia el mejor animal sería el animal B. El animal A presenta mayor producción total que pudiera tener altos niveles de producción al inicio de la lactancia con una fuerte caída, en comparación con el B que resultó tener una producción más sostenida, es decir, que a pesar de no presentar altos niveles de producción al inicio de la lactancia, mantiene su producción durante un mayor tiempo. Comparativamente estos dos animales presentan la misma producción total, sin embargo, la forma en que se distribuye la producción en la lactancia pudiera ser la más importante. Así el mejor animal sería aquel que cubra con las necesidades de un sistema de producción específico dado que los objetivos de producción pueden variar para cada uno de los diferentes sistemas, esta metodología podría permitir que el productor seleccione con base a sus necesidades específicas.

Una de las ventajas de considerar la persistencia en las decisiones de la selección de los animales es la oportunidad de favorecer la recuperación oportuna de la capacidad de ingestión de alimento, recuperación de la actividad reproductiva, reducción de los problemas sanitarios y por consiguiente una reducción en los costos de producción (Sölkner y Fuchs, 1987; Dekkers *et al.*, 1997).

El uso de modelos del día de registro, específicamente regresión aleatoria - funciones de covarianza, son de gran importancia en la industria lechera, pues con un modelo equivalente al multivariado se cuenta con información adicional sin incrementar sustancialmente los requerimientos de cómputo, permitiendo el desarrollo de estrategias de mejoramiento no solo sobre la producción total, sino también sobre algunos componentes de la curva de lactancia, como producción máxima y persistencia. De igual

forma, es posible predecir el valor genético de los animales con lactancias incompletas sin necesidad de su extensión, ello puede conducir a una mayor respuesta a la selección como reflejo de la reducción del intervalo generacional (Jamrozik y Schaeffer, 1997; Freeman, 1998; Swalve, 2000).

Conclusiones

Es posible usar modelos de regresión aleatoria-funciones de covarianza para predecir los valores genéticos de los animales en cualquier parte de la curva de lactancia.

La selección de animales por valores genéticos obtenidos con regresión aleatoria – funciones de covarianza, permite la elección de animales de acuerdo a las necesidades específicas del productor de acuerdo a sus objetivos de producción.

Una reducción importante en el número de parámetros necesarios a ser estimados para describir genéticamente las características que se expresan en un rango dado de la vida del animal, se puede lograr mediante el modelos de regresión aleatoria - funciones de covarianza.

El modelo de regresión aleatoria - funciones de covarianza produce como subproducto los valores genéticos de producción total, parcial y persistencia. Así, no es necesaria su inclusión en el vector de observaciones y por lo tanto se requiere de un modelo univariado, reduciendo los requerimientos de tiempo y equipo de cómputo.

Literatura citada

1. Ali TE, Schaeffer LR. Accounting for covariances among test day milk yields en dairy cows. Can. J. Anim. Sci. 1987; 67: 637-644.
2. Bar-Anan R, Ron M, Wiggans GR. Associations among milk yield, yield persistency, conception, and cullings of israeli Holstein dairy cattle. J. Dairy Sci. 1985; 68: 382-386.
3. Dekkers J, Ten Hag JH, Weersink A. Economic aspects of persistency of lactation in dairy cattle. Livestock Production Science 1998; 53: 237-252.
4. Draper NR, Smith H. Applied Regression Analysis. Second edition. Edit. John Wiley & Son. 1980; 45-54.
5. Freeman AE. Dairy cattle Breeding. Proc. 6th Worl Congr. Genetic Appl. Livest. Prod., Armidale, New South Wales, Australia. 1998; 23;292-294
6. Guo Z, Swalve H. H. 1995. Modelling of the lactation curve as sub-model in the evaluation of test day records. Proceedings of the open Session of the Interbull Annual Meeting. Prague, Czech Republic. September 7-8. 1995; 11: 1-4
7. Jamrozik J, Kistemaker GJ, Dekkers JCM, Schaeffer LR. Comparison of posible covariates for use in a random regression model for analyses of test day yields. J. Dairy Sci. 1997; 80: 2550-2556
8. Jamrozik J, Schaeffer LR. Estimates of genetics parameters for test day model with random regression for yield traits of first lactations Holsteins. J. Dairy Sci. 1997; 80:762-770.

9. Kirkpatrick M, Lofsvold D, Bulmer M. Analysis of the inheritance selection and evolution of growth trayectories. *Genetics* 1990; 124: 979-993.
10. Meyer K, Hill WG. Estimation of genetic and phenotypic covariance functions for longitudinal or repeated records by restricted maximun likelihood. *Livest. Prod. Sci.* 1997; 47:185-200.
11. Meyer K. Derivate Free Restricted Maximun Likelihood (DFREML). User notes. 1998; 1-28 p .
12. Meyer K. Estimating variance and covariance for multivariate animal models by restricted maximun likelihood. *Genet. Sel. Evol.* 1991; 23: 67-83.
13. Meyer K. Restricted maximun likelihood to estimate variance components for animal models with several random effects using derivate-free algorithm. *Genet. Sel. Evol.* 1989; 21: 317-340.
14. Misztal I, Strabel T, Jamrozik J, Mäntysaari EA, Meuwissens THE. Strategies for estimating the parameters needed for different test day models. *J. Dairy Sci.* 2000; 83: 1125-1134.
15. Núñez DR, Rodríguez FAA. 1995. Componentes de Varianza. Memorias del curso intensivo "Introducción a la metodología de los modelos mixtos (Aplicaciones al Mejoramiento Animal)". Facultad de Zootecnia. Universidad Autónoma de Chihuahua. Chihuahua, Méx. 1995; III: 1-54 p.
16. Pander BL, Hill WG, Thompson R. Genetic parameters of test day records BritishHolstein-Friesian heifers. *Anim. Prod.* 1992; 55: 11 - 21.

17. Schaeffer LR, Dekkers JCM. Random regression in animal models for test day production in dairy cattle. Proc. 5th World Congre. Genet. Appl. Livest. Prod. Guelph, ON, Canada. 1994; 18: 443-446.
18. Shanks RD, Berger PJ, Freeman AE, Dickinson FN. Genetic spect of latations curves. J. Dairy Sci. 1981; 64: 1852-1860.
19. Sölkner. J, Fuchs W. A comparison of different measures of persistency with special respect to variation of test-day milk yields. Livest. Prod. Sci. 1987; 16: 305-319.
20. Swalve HH. The effect of test day models on the estimation of genetic parameters and breeding values for dairy yields traits. J. Dairy Sci. 1995; 78: 929-938
21. Swalve HH. Use of test day records in genetics evaluation. Proc. 6th Worl Congr. Genetic Appl. Livest. Prod., Armidale, New South Wales, Australia 1998; 23: 295-302
22. Swalve. HH. Theoretical basis and computational methods for different test day genetic evaluation methods. J. Dairy Sci. 2000; 83: 1115-1124.
23. Trus D, Buttazzoni LG. A multiple trait approach to modelling the lactation curve. Proc. 4th wold congress on genetic applied to livestock production. Edimburgo, Uk, XIV, 1990; 492-495.
24. Tsuruta S, Keown JF, Van Vleck LD, Miztal I. Bias in genetic evaluations by records of cows treated with bovine somatotropin. J. Dairy Sci. 2000; 83:2650-2656
25. Wilmink JBM. Adjustment of test day milk, fat and protein yields for age, season and stage of lactation. Livest. Prod. Sci. 1987; 16: 335.

BIBLIOTECA CENTRAL U. A. CH.