

**PRINCIPIOS BÁSICOS E IMPORTANCIA
DE LA UTILIZACIÓN DE LOS SISTEMAS DE REGRESION ALEATORIA
EN LOS PROGRAMAS DE SELECCION Y MEJORA ANIMAL**

Octubre 6 y 7, 2005
Universidad de Córdoba,
Departamento de Genética.
Córdoba, España

A. Menéndez Buxadera
Centro de Control Pecuario,
Ministerio de la Agricultura
Ciudad de La Habana, Cuba.

INDICE.

I. Introducción.....	3
II. Conclusiones generales del trabajo de selección y mejora Genética.....	4
III. Tipos de datos mas usados en genética animal.....	6
III.1. Particularidades del uso de las mediciones repetidas.....	8
III.2. Aspectos teóricos.....	8
IV. Análisis estadísticos de las medidas repetidas.....	11
IV.1. Modelos de Repetibilidad.....	11
IV.1.1. Modelo Animal con datos repetidos.....	11
IV.1.2. Modelo animal del día de control.....	12
V. Modelos multicastreres y de regresión aleatoria.....	14
V.1. Funciones de covarianza.....	14
V.2. Estimación de funciones de covarianza mediante el uso de modelos mixtos.....	17
VI. Aplicación práctica y discusión general.....	20
Anexo 1. Análisis de Componentes Principales.....	22
Anexo 2. Comparación de modelos.....	25
Anexo 3. Modelo animal uní variado, multicastreres y regresiones aleatorias.....	27
Anexo 4. Evolución de las Varianzas Genéticas y uso de Regresiones Aleatorias....	41
Anexo 5. Estudio de la curva de lactancia mediante regresión aleatoria	53
Anexo 6. Interacción genotipo-ambiente y regresión aleatoria.....	67

I. Introducción.

En la actualidad la población humana sigue creciendo y consecuentemente la demanda de alimentos, particularmente de origen animal, de manera que será necesario desarrollar métodos y procedimientos que permitan marcados incrementos en los niveles productivos tanto en regiones desarrolladas como en los países tropicales. Estas necesidades requieren que se intensifique la producción, sin embargo, esta alternativa puede producir graves consecuencias en términos de las acumuladas agresiones al clima y el efecto que puede tener sobre el bienestar de la humanidad.

Independientemente del tipo de animal o de situaciones específicas de un país o región, resulta obvio que producir con eficiencia, es un objetivo indiscutible de cualquier actividad productiva, lo que puede traducirse por mayor volumen de producto final por unidad de tiempo y recursos invertidos en la explotación de los animales. En general, esto puede cuantificarse por la relación input/output. Se conoce que el incremento en la productividad del vacuno, ovino y caprino es una función de un gran número de componentes entre los que existe una interdependencia mutua y como resultado de los cuales un alimento (supuestamente no consumible por el hombre) se transforma en un producto de alto valor biológico como la leche y la carne (Output), para lo cual es necesario invertir un importante volumen de recursos tanto materiales como humano (Input). Lo expuesto implica la existencia de un adecuado sistema de controles individuales que posibilite evaluar la marcha de todo el proceso.

Es significativo el hecho de que en la actualidad existe un alto desarrollo en la informática aplicada a las diferentes esferas de la sociedad humana, existiendo controles para todo o casi todo, esta condición prácticamente no se discute. Sin embargo, en el caso de la producción animal tal afirmación no siempre se cumplimenta existiendo una amplia diferencia entre lo que se requiere y lo que realmente se controla. En el caso particular que nos ocupa para el trabajo de selección animal el sistema de control de los datos es una herramienta imprescindible sin la cual no resulta simple de llevar a cabo.

Estos controles se pueden efectuar periódicamente (pesos vivos a diferentes edades, consumo de alimentos semanal, etc.) antes que el animal comience su fase productiva, a partir de donde existen otros registros sistemáticos (control mensual de la producción de leche, grasa o la producción de huevos etc.), con los que se estima la producción total en un tiempo dado. También existen otros datos que se repiten a diferentes intervalos en la vida de un animal, tal es el caso de la fertilidad de la hembra, condición corporal antes del parto, el peso al destete de sus progenies etc. Estos tipos de datos requieren un tratamiento específico para la evaluación genética de los animales cuyas características generales son precisamente los objetivos que trataremos en estos días.

Durante el proceso de selección y mejora, muchas veces no se toma en cuenta que los animales no solo transmiten los genes que determinan el carácter por el cual se seleccionaron sino una muestra de todo su genotipo, de manera que es muy recomendable conocer la evolución de otros rasgos durante el proceso de selección. Para demostrar estos puntos de vistas, permítasenos presentar una rápida y corta revisión de las conclusiones generales del trabajo de genética animal en animales domésticos con vistas a guiar nuestra discusión acerca del porque es necesario insistir en la necesidad de disponer de un buen

sistema de controles y en los procedimientos de análisis que nos permita estudiar y profundizar en la naturaleza de las causas de variación de los caracteres de interés económicos.

II. Conclusiones generales del trabajo de selección y mejora genética.

Independientemente de otras consideraciones vale la pena presentar brevemente una síntesis de las tendencias de los programas de mejora genética. La tabla 1 muestra el balance global del trabajo de selección artificial en animales domésticos, los cuales están altamente inspirados en las publicaciones de Rauw y col. (1998) y Hill y col. (2000).

Tabla 1. Efectos directos y colaterales de la selección artificial en porcino; pollos de ceba y ganado de leche. (adaptado de Rauw y col., 1998 y Hill y col., 2000).

	GANADO PORCINO	POLLO DE CEBAS	GANADO DE LECHE
EFFECTOS POSITIVOS	<ul style="list-style-type: none"> • Aumentó el crecimiento diario en 26%. • Decreció el consumo de alimento por kg de peso vivo en 21%. • Mejoró la composición de la canal. 	<ul style="list-style-type: none"> • Disminuyó la edad al sacrificio en 16 días. • Aumentó el peso al sacrificio en 35%. • Mejoró la conversión de alimento en 27%. • Mejoró la composición de la canal. 	<ul style="list-style-type: none"> • Incrementó la producción de leche en 75% en cada lactancia. • Consecuentemente aumentó la producción de grasa y proteína.
EFFECTOS NEGATIVOS	<ul style="list-style-type: none"> • Disminuyó el comportamiento reproductivo. • Mayores problemas de salud. • Mayores problemas óseos. • Mayor dificultad de locomoción. 	<ul style="list-style-type: none"> • Disminuyó el comportamiento reproductivo. • Se afectó el mecanismo hipotalámico del control de la alimentación. • Más problemas de salud. • Serios problemas óseos. 	<ul style="list-style-type: none"> • Disminuyó el comportamiento reproductivo. • Aumentó los problemas de enfermedades. • Disminuyó la longevidad.
CONSECUENCIAS FISIOLÓGICAS GENERALES	<ul style="list-style-type: none"> • El metabolismo se especializó en unas pocas funciones productivas las cuales emplean la mayor parte de los recursos nutritivos consumidos por el animal, dejando menos energía disponible para otros procesos fisiológicos. • En tal situación los primeros caracteres afectados fueron aquellos que conforman el 'FITNESS'. • Los animales son más productivos, pero menos capaces de enfrentar procesos estresantes tales como alteraciones violentas de efectos climáticos, enfermedades, parasitismo. • Los animales son menos longevos. 		
CONSECUENCIAS ÉTICAS Y SOCIALES	<ul style="list-style-type: none"> • Existe un consenso universal de que hay que respetar y mantener el bienestar de los animales. • Nuevas enfermedades han aparecido (vaca loca; dioxina) que han llamado la atención a todos, aumentando los controles de laboratorio sobre el estado higiénico y la calidad de los productos. • La intensificación de la producción animal ha incrementado aun más la contaminación ambiental y esto tiene efectos negativos en la percepción de la sociedad sobre estos problemas. • Los altos volúmenes de producción de leche han promovido al establecimiento de cuotas para los diferentes países, con vistas a mantener los precios. Esto resulta éticamente criticable, pues una gran parte de la humanidad tiene déficit de alimentos. 		

Es coincidente que en las tres especies se manifestaron significativos aumentos en los principales indicadores productivos, de manera que puede plantearse que un animal promedio de finales de la década del 90 del pasado siglo produce aproximadamente igual que dos animales nacidos 50 años previos. Sin embargo, también es coincidente el hecho de que se ha deprimido el comportamiento reproductivo, la capacidad de adaptación y resistencia, todo lo cual de conjunto, ha producido una reducción en el **FITNESS** de los animales. Todos estos negativos efectos colaterales se deben a que el proceso de selección artificial produjo una especialización en el metabolismo de los animales en unas pocas funciones productivas cuyas consecuencias deben asumirse. Para demostrar este punto de vista se presentara brevemente las bases fisiológicas y las implicaciones de un programa de mejora para el crecimiento.

Los objetivos de un programa de selección artificial tratan de maximizar el **OUTPUT** (generalmente un producto comercializable), pero con el menor o mas eficiente uso del **INPUT**, con vistas a obtener mayores beneficios económicos. En el caso de la producción animal, es el alimento consumido el principal componente del **INPUT**, el cual representa la mayor parte de los costos de producción (entre 60 a 70%).

Los animales utilizan la Energía Consumida (**EC**) para diversas funciones vitales y no solo para un objetivo productivo comercial (**OUTPUT**). Por otro lado el volumen de **EC** esta restringido a ciertos limites, dados por las múltiples interacciones de carácter físico y biológico entre *animal-alimento-tiempo-proceso metabólico*. En animales en la etapa de crecimiento, la **EC** se cuantifica según la:

$$\text{Ecuación 1} \quad y_{EC} = b_1PM + b_2GD + e_r$$

donde:

PM es el peso metabólico generalmente estimado como $PV^{.75}$ y **PV** es el peso vivo.

GD es la ganancia diaria de peso.

b₁ y **b₂** son los coeficientes de regresión que cuantifican el uso de **EC** para cada proceso. Estos parámetros, bajo ciertas circunstancias, son equivalentes a los denominados coeficientes de utilización de la energía para los procesos de mantenimiento (**b₁**) y ganancia de peso (**b₂**).

e_r es conocido como error residual, en el cual se incluyen todos aquellos efectos no considerados en la ecuación y que afectan a la variable dependiente **y_{EC}**

Ecuaciones como la 1 se han estimado experimentalmente en diferentes países, donde los animales se encuentran en condiciones de laboratorio con ausencia de agentes estresantes, y una dieta uniforme siempre disponible. Lo anterior resulta el método más recomendable ya que se trata de establecer un patrón comparativo. En términos prácticos existe consenso de opiniones en señalar varios puntos críticos:

- Se considera que no hay variabilidad entre animales para los coeficientes **b₁** y **b₂**.
- Se han empleado animales genéticamente diferentes a las generaciones actuales.
- Se asume que la composición corporal de la ganancia diaria de peso es la misma para todos los animales.
- No considera que en condiciones prácticas la energía consumida se utiliza para múltiples procesos tales como: *stress térmico (ST)*; *competición por el espacio* y el

alimento (C); nivel de actividad para cosechar el alimento, el cual no siempre está disponible en igual cantidad y calidad (A); movimiento diario continuo (M); resistencia a enfermedades (RE); parasitismo (P) y comportamiento reproductivo (CR), entre otros.

- Otros elementos, tal como el error de muestreo (e_m).

Según lo expuesto y empleando la simbología indicada anteriormente, es posible construir una nueva ecuación:

$$\text{Ecuación 2} \quad y_{EC} = b_1PM + b_2GD + b_3ST + b_4C + b_5A + b_6M + b_7RE + b_8P + b_9CR + e_m$$

Nótese la similitud entre las ecuaciones 1 y 2. Los dos primeros términos son los mismos, mientras que el resto de los coeficientes $b_3 \dots b_9$ corresponden con las variables explicativas mencionadas previamente que no fueron incluidas en 2 y que por definición son parte del término ' e_r ' el cual se conoce como Residual Feed Consumption **RFC**. Por simple comparación de ambas ecuaciones se tendrá que:

$$e_r = \text{RFC} = \frac{b_3ST + b_4C + b_5A + b_6M + b_7RE + b_8P + b_9CR + e_m}{\downarrow \text{CARACTERES ADAPTATIVOS}}$$

De esta forma, resulta obvio que el término e_r es una medida de **RFC** y explica la proporción de energía empleada para una serie de procesos fisiológicos ligados a lo que se denomina como **componentes funcionales adaptativos**. Es importante señalar que en **RFC** se "absorbe" las variaciones individuales en b_1 y b_2 así como las de composición de la ganancia de peso vivo, de manera que el significado biológico de **RFC** tendrá mayor aplicación práctica. En otras palabras e_r puede considerarse como un término que puede manifestar variaciones individuales cuyo origen puede conocerse mediante una ecuación de regresión aleatoria y cuya aplicación puede tener una amplia repercusión en el programa de mejora ya que esta vinculada a caracteres adaptativos. La derivación de estas fórmulas fueron adaptadas para animales en crecimiento (aves, caprinos, ovinos, porcino y vacunos) y fueron descritas en detalle por Luiting y col. (1997).

Lo expuesto pudiera interpretarse como fuera de contexto, sin embargo el objetivo es todo lo contrario. La experiencia indica que el trabajo de selección dirigido hacia unos pocos caracteres productivos producirá importantes beneficios, pero también implican un costo nada despreciable en otros rasgos de mucho interés económicos. Para guiar y corregir la estrategia del programa de mejora es necesario el análisis del comportamiento de los animales en su sentido mas amplio posible y esto solo puede llevarse a cabo mediante un adecuado sistema de controles y análisis de los datos.

III. Necesidad de un sistema de controles y tipo de datos más usados en genética animal.

Existe una confusión histórica acerca de la importancia de la aplicación y uso de los sistemas de control individual en la actividad pecuaria. La producción animal es un

complejo proceso donde intervienen la acción e interacción de un gran número de factores ligados al animal, al clima y al hombre, que los trata de modificar en su beneficio. Resulta obvio que es imposible tener conocimiento preciso del papel, naturaleza e importancia de cada uno de sus componentes, si no se dispone de un amplio Sistema Global de Control (SGC) el cual posibilite procesar los datos y producir una información imprescindible (tabla 2).

Tabla 2. Razones que justifican la aplicación de los controles.

POR QUÉ	PARA QUÉ
<ul style="list-style-type: none"> • Es la única forma de saber qué tenemos y cómo lo usamos. • Produce un dato el cual es la materia prima más importante para producir información. • Forma simple de bajo costo para incrementar la capacidad analítica del hombre. • Posibilita aumentar el conocimiento sobre el proceso productivo. 	<ul style="list-style-type: none"> • Permite tomar decisiones inteligentes en tiempo y forma • Hacer el máximo uso de los recursos disponibles. • Posibilita aplicar un trabajo analítico colectivo. • Chequear permanentemente la marcha de las decisiones.

El valor del SGC, y particularmente de la información que genera, cobra mayor relevancia en condiciones económicas difíciles, ya que el no empleo de la misma puede hacer fracasar los planes trazados para la actividad. El uso adecuado de la información permite incrementar la capacidad de pensar analíticamente, se incrementa el conocimiento sobre los procesos, lo que por su parte, garantiza el mejoramiento de la actividad productiva y su propia sostenibilidad. Aplicar un sistema de controles y analizar en forma sistemática la información que se almacena, es una forma de ser eficientes, ya que posibilita brindar o dirigir la prioridad hacia aquellos eventos, que producen una alta y rápida respuesta y esto es particularmente necesario en la actividad ganadera en su conjunto y no en una parte de ella.

En el caso de la producción vacuna el empleo de un amplio SGC posibilita:

- Al conocer los niveles productivos de cada animal, es posible aplicar un amplio programa de selección artificial.
- Se puede efectuar un uso más eficiente de los suplementos suministrados según la producción de cada animal.
- Posibilita el agrupamiento de los animales de acuerdo al volumen de producción, lo cual coadyuva a un mejor uso de los pastos y otros recursos alimenticios.
- Permite identificar cuáles rebaños tienen un adecuado comportamiento y al propio tiempo se puede conocer aquellos factores que representan un freno en otras unidades de producción que no han alcanzado iguales resultados.

- La acumulación de resultados productivos individuales permite la creación de un banco de datos con los cuales se pueden realizar diversos estudios tendientes a identificar las principales causas de variación que afecta los caracteres de mayor interés económico.

Esta amplia utilización de los resultados de SGC indica que los controles no se realizan sólo para hacer ‘genética’, sino para el manejo general de esta actividad productiva. En todo caso el mejoramiento animal puede lograrse como un subproducto de los sistemas de controles.

Es necesario apuntar que el hecho de disponer de un SGC implica el tratamiento estadístico del mismo y en tal caso, los datos disponibles presentan características especiales que deben tomarse en cuenta. De acuerdo a Fitzhugh, (1976), la mayor parte de los datos utilizados en producción animal son de tres tipos:

- **Estáticos:** son aquellos en los que uno o varios rasgos se registran una sola vez en cada individuo (peso vivo al destete, peso al año de edad etc.). En general, muchos de los análisis realizados en genética animal han empleado este tipo de datos, los cuales pueden ser analizados por modelos uni o multivariados, brindando poca información sobre el patrón de desarrollo de los animales.
- **Corte transversal:** similares a los anteriores cada individuo se representa una sola vez en un estadio de desarrollo, pero otros animales se miden en otras etapas, de manera que globalmente se puede estimar la curva de crecimiento del conjunto.
- **Longitudinales:** todos los individuos son medidos en cada estadio del desarrollo. Este tipo de datos contiene toda la información que brindan los anteriores a lo cual debe añadirse las posibilidades de medir las variaciones individuales. Como inconveniente tiene el tiempo que debe esperarse para que todos los individuos estén representados en los diferentes puntos de la trayectoria.
- **Longitudinales mixtos:** esto es una mezcla de los previamente mencionados. Es menos preciso que lo datos puramente longitudinales pero manteniendo ciertos requerimientos puede aportar importante y valiosa información acerca del patrón de crecimiento y desarrollo de la población.

Cualquiera sea el tipo de datos disponibles, el uso de toda la información en un momento dado, con vistas a incrementar la precisión y posibilidades del proceso de selección y mejora genética, se ha convertido en una estrategia de mucha envergadura, particularmente facilitada por los espectaculares avances en materia de computación. En tal sentido resulta obvio interrogarse sobre las ventajas de utilizar ‘una variable resumen (dimensión finita)’ (ejemplo producción total de leche en una lactancia o peso vivo a una edad fija) en comparación al empleo de las ‘variables originales repetidas temporalmente (dimensión infinita)’ (ejemplo pesajes mensuales de leche) a partir de donde se estimaron aquellas. Múltiples trabajos se han publicado sobre este tema, aunque le correspondieron a Ptak y Schaeffer (1993) el merito de haber presentado una primer versión coherente de la evaluación de los registros de producción láctea en el día de control, internacionalmente conocido como Test Day Model (TDM). Un año mas adelante Schaeffer y Dekkers, (1994), presentaron las bases de un procedimiento general para tratar los TDM mediante lo que se denomina Random Regresión Model (RRM). A partir de esta publicación se han generado una gran cantidad de resultados de aplicación de RRM a datos de la producción de leche (Swalve, 2000), crecimiento en el ganado vacuno (Meyer y Hill, 1997), consumo

de alimento en porcino (Schnyder y col 2001) así como crecimiento en peces (McKay y col 2002) y también en árboles maderables (Apiolaza y col 2000).

La mayor parte de los estudios realizados de genética animal hasta el presente, se basan en '*modelos de dimensiones finitas*', en los que algunos rasgos de interés económico se evalúan a una edad fija (peso vivo a 18 meses por ejemplo) o a un estadio determinado (producción de leche total en primera lactancia). Aún cuando este tipo de análisis brindan resultados satisfactorios, estos pueden mostrar parámetros con tendencias muy variables y de difícil interpretación. En el crecimiento o desarrollo de un animal es lógico asumir que su trayectoria presente una respuesta armoniosa, sin los picos que se encuentran en los análisis multi caracteres. Para resolver estos problemas se introdujo un '*enfoque dimensional infinito*', en el cual se asume que el fenotipo de un animal es una función continua del tiempo, de manera que para su mas adecuada descripción, este debe ser registrado un numero infinito de veces, (Kirpatrick y col.1990).

Estos autores señalan algunas ventajas a este punto de vista:

- Predice la evolución del crecimiento total sin asumir nada a priori, sobre el tipo de curva necesaria para ajustar los datos.
- Proporciona un método para analizar el patrón de variación genética que puede revelar cambios potenciales en la trayectoria del crecimiento.
- Reduce los sesgos en la estimación de la variación genética, pues no requiere ajustes previos a edad fija.
- Se toma en consideración el período de tiempo entre las edades en las que se registró el dato.
- Permite proyectar la evolución de la trayectoria del crecimiento, aún cuando los datos de cada animal son registrados a diferentes edades.

Tomando en cuenta lo expuesto se presentara una simple descripción de la fundamentación general de estos nuevos procedimientos y se mostraran las posibilidades de los mismos.

III.1. Particularidades del uso de las mediciones repetidas.

El uso de los registros sistemáticos de la productividad de los animales puede brindar importantes beneficios. Según van der Werf y Schaeffer, (1997), estos datos manifiestan cambios graduales en el tiempo cuya naturaleza (origen genético) puede coadyuvar a mejoras sustanciales en el programa de mejora. Por otro lado, el uso de las variables originales incrementa la precisión de los estimados y pueden disminuir los costos de los sistemas de controles o hacerlos mas eficientes.

En condiciones tropicales o en sistemas de producción extensivos donde resulta muy difícil garantizar una identificación adecuada de todos los animales., el uso de mediciones repetidas puede tener una mayor relevancia e importancia. Así por ejemplo, si se estiman los componentes genéticos y no genéticos de las mediciones repetidas, es posible contribuir al proceso de selección compensando en algo los deficientes controles de paternidad.

Para hacer efectiva estas posibilidades deben considerarse varios aspectos. En primer lugar se conoce que dos medidas en un mismo animal están más correlacionadas que entre dos individuos diferentes, de igual forma las medidas adyacentes tienen mayor grado de relación que aquellas separadas por un intervalo mas largo de tiempo. Según estos

elementos hay que tomar en cuenta la estructura de covarianza entre medidas repetidas si se desea extraer el máximo posible de respuesta de la información disponible.

III.2. Aspectos teóricos.

El fenotipo de un individuo (**P**) es una forma de medir la contribución de los efectos genéticos (**G**) y ambientales (**E**), es decir:

$$\text{Ecuación 4} \quad \mathbf{P} = \mathbf{G} + \mathbf{E}$$

En 4 se asume que **G** y **E** son independientes. En general, solo se dispone del valor de **P**, de manera que bajo ciertos principios y procedimientos, **G** puede deducirse. Para ello es necesario asumir que el valor de **P** está medido en una gran cantidad de condiciones ambientales y donde están representados una gran cantidad de genotipos. El problema de **E** es mas complejo, así por ejemplo el efecto del ambiente materno puede afectar el valor de **P** de su progenie, o es posible que el ambiente antes del parto puede tener influencia en el valor de **P** en el propio individuo. Por otro lado el genotipo y el ambiente generalmente presentan una interacción, de manera que es preferible una modificación de la ecuación anterior:

$$\begin{aligned} \text{Ecuación 5} \quad \mathbf{P} &= \mathbf{G} + \mathbf{GEp}_e + \mathbf{GEt}_e + \mathbf{Ep}_e + \mathbf{Et}_e \\ &= \mathbf{G} + \underbrace{\mathbf{GEp}_e + \mathbf{Ep}_e}_{\text{DIFERENCIAS PERMANENTES}} + \underbrace{\mathbf{GEt}_e + \mathbf{Et}_e}_{\text{DIFERENCIAS TEMPORALES}} \end{aligned}$$

donde **GE** significa la interacción genotipo ambiente, la cual se puede manifestar en un efecto ambiental permanente (**p_e**) o temporal (**t_e**). Si se considera independencia entre los efectos, entonces se puede confeccionar una ecuación muy importante en el trabajo de selección animal:

$$\begin{aligned} \text{Ecuación 6} \quad s_p^2 &= s_G^2 + \underbrace{s_{GEp_e}^2 + s_{GEt_e}^2}_{\text{DIFERENCIAS PERMANENTES}} + s_{Ep_e}^2 + s_{Et_e}^2 \\ &= s_A^2 + s_D^2 + s_I^2 + s_{GE}^2 + s_{Ep_e}^2 + s_{Et_e}^2 \end{aligned}$$

Es importante resaltar que la aplicación de la ecuación 6 permite descomponer la varianza fenotípica total (s_p^2) en sus componentes ambientales y genéticos. Este último representa las diferencias genéticas totales entre individuos, es decir aquellas debidas a efectos aditivos (s_A^2) de dominancia (s_D^2) y de epistasia (s_I^2).

Cuando un rasgo se mide varias veces en un mismo animal, entonces el término s_p^2 puede descomponerse en varianza intra y entre individuos. En el contexto de la ecuación 6, la varianza entre individuo está formada por causas de origen permanente, tanto genéticas

como de carácter ambiental ($S_G^2 + S_{Ep_e}^2$), cuya magnitud debe ser expresadas en forma relativa a la variación fenotípica total, de manera que:

$$\Gamma = \frac{S_G^2 + S_{Ep_e}^2}{S_P^2}$$

en la que Γ se denomina repetibilidad del carácter. Este estimador no es un parámetro genético, aunque puede tener utilización en los programas de selección y mejora, tal y como se expondrá más adelante. Note que Γ mide la proporción de la varianza fenotípica total cuyo origen tiene un carácter permanente y su valor equivale al coeficiente de regresión de un registro sobre el próximo.

Para una mejor representación de estas posibilidades considere que dispone de n registros del animal i con media \bar{y}_i . Asumiendo nula la interacción **GE** la varianza de las y observaciones se representa como:

$$\begin{aligned} S_{y_i}^2 &= S_G^2 + S_{Ep_e}^2 + S_{Et_e}^2 && \text{para } \mathbf{1} \text{ observación.} \\ S_{\bar{y}_i}^2 &= S_G^2 + S_{Ep_e}^2 + S_{Et_e/n}^2 && \text{para } \mathbf{n} \text{ observaciones} \end{aligned}$$

El objetivo es estimar el Valor Genético (VG) del animal \hat{a}_i , lo cual se estima como:

$$\begin{aligned} \hat{a}_i &= \mathbf{b} (y_i - \mu) && \text{para } \mathbf{1} \text{ observación.} \\ \hat{a}_i &= \mathbf{b} (\bar{y}_i - \mu) && \text{para } \mathbf{n} \text{ observaciones.} \end{aligned}$$

donde μ representa la media general de las observaciones y \mathbf{b} es el coeficiente de regresión de \bar{y}_i sobre el valor genético aditivo a_i :

$$\mathbf{b} = \frac{\text{cov}(\mathbf{a}, \bar{y})}{\text{var}(\bar{y})}$$

Ahora bien,

$$\begin{aligned} \text{cov}(\mathbf{a}, \bar{y}) &= \text{cov}(\mathbf{a}, \mathbf{g} + \mathbf{E}p_e + \mathbf{E}t_e/n) \\ &= \text{cov}(\mathbf{a}, \mathbf{g}) \\ &= S_A^2 \end{aligned}$$

Por otro lado,

$$\text{var}(\bar{y}) = S_G^2 + S_{Ep_e}^2 + S_{Et_e/n}^2$$

Incorporando el significado de Γ en la ecuación 6 se tendrá:

$$\text{var}(\bar{y}) = [\Gamma + (\mathbf{1} - \Gamma) / \mathbf{n}] S_Y^2$$

Por tanto,

$$\mathbf{b} = S_A^2 / [\Gamma + (\mathbf{1} - \Gamma) / \mathbf{n}] S_Y^2$$

$$= \frac{\mathbf{n} \mathbf{h}^2}{[\mathbf{1} + (\mathbf{n} - \mathbf{1}) \Gamma]}$$

Observe que el valor de **b** ahora es una función de la heredabilidad del rasgo (\mathbf{h}^2), del coeficiente de repetibilidad y del número de repeticiones. Por definición el valor de **b** para un solo registro es igual a \mathbf{h}^2 . De manera que resulta obvio que la precisión de la estimación del valor genético ($\hat{\mathbf{a}}_i$) aumentará con el uso de medidas repetidas, y con el valor de Γ . En términos generales para caracteres con bajo valor de \mathbf{h}^2 (rasgos reproductivos por ejemplo) se puede recomendar este tipo de alternativa.

Cuando se dispone de varias mediciones del mismo rasgo en el individuo, se introduce un elemento que aumenta el nivel de similitud entre estos registros, cuyo origen es de carácter permanente que explican por que un animal es más productivo que otro. Este es el fundamento del uso de la ‘Habilidad Productiva Mas Probable’ (HPMP) debida a Lush. La importancia de conocer Γ es simplemente para evaluar cuanto se puede ganar en precisión en las predicciones. Cuando Γ es alto poco será el incremento, por el contrario para valores bajos de Γ , mucho se podrá aumentar en precisión por el hecho de realizar mediciones repetidas. También es cierto que un valor alto de Γ permite predecir tempranamente el comportamiento futuro de un animal para el mismo rasgo. Estas ventajas se deben a que cuando existen **n** repeticiones del mismo rasgo, la varianza debida a efectos ambientales temporales se reduce en una magnitud igual a $\mathbf{S}_{Ete/n}^2$.

IV. Análisis estadísticos de las medidas repetidas.

Según van der Werf y Scheffer 1997 existen tres tipos de modelos para tratar las medidas repetidas:

- **Modelo de Repetibilidad (MR)** : de uso simple y requiere de pocos recursos computacionales. Sus resultados pueden ser sesgado cuando las correlaciones entre mediciones son significativamente inferior a 1. Se han empleado mucho en genética animal.
- **Modelo Multicaracteres (MMC)** : produce resultados más exactos ya que toma en cuenta diferentes correlaciones entre los rasgos aunque demanda de muchos recursos de computación. Cuando el número de caracteres es muy grande se presentan problemas de sobreparametrización.
- **Modelo de Regresión Aleatoria (MRA)** : son modelos de análisis en los que los se miden la evolución de los valores fenotípicos en función del tiempo (tambien del nivel de ambiente) mediante coeficientes de regresión aleatorios. Estos modelos utilizan las funciones de covarianzas propuestas por Kirpatrick y col (1990).

IV.1. Modelos de Repetibilidad.

Consciente o inconscientemente los **MR** son los que mas se han usado en los estudios de genética. Los análisis de peso al destete, algunos rasgos reproductivos y la producción de leche por lactancia son algunos de estos ejemplos. Existen varios tipos de modelos estadísticos para estudiar este tipo de datos.

IV.1.1. Modelo Animal con datos repetidos.

Considere un ejemplo simple de un grupo de vacas con dos o más lactancias, hijas de x sementales y z madres, algunas de las cuales pueden aparecer en el vector de datos. El pedigree es conocido para todos los animales. El modelo de análisis es el siguiente (Kennedy, 1989):

$$\text{Modelo A} \quad y = Xb + Za + Wp_e + e$$

donde :

y es un vector de observaciones.

b es un vector de efectos fijos.

a es un vector aleatorios de efectos genéticos aditivos debidos al animal.

p_e es un vector de efectos aleatorios de ambiente permanentes y efectos genéticos no aditivos.

e es un vector aleatorio de efectos residuales.

X, Z y W son matrices de incidencia que relacionan los registros con los efectos fijos, animal y efectos permanentes del ambiente respectivamente.

Se asume que los efectos de **a**, **p_e** y **e** tienen distribución independiente con media cero y varianzas igual a s_a^2 , $s_{p_e}^2$ y s_e^2 respectivamente. Luego,

$$\text{var}(p_e) = \mathbf{I} s_{p_e}^2 ; \text{var}(e) = \mathbf{I} s_e^2 = \mathbf{R}; \text{var}(a) = \mathbf{A} s_a^2 ; \text{var}(y) = \mathbf{ZAZ}' s_a^2 + \mathbf{W} \mathbf{I} s_{p_e}^2 \mathbf{W}' + \mathbf{R}$$

donde **I** ; **A** y **R** son matrices de identidad, del numerador de la relación de parentesco y de efectos residuales respectivamente.

Las soluciones de este sistema de ecuaciones BLUE de **b** y BLUP de **a** y **p_e** son:

$$\begin{bmatrix} \hat{b} \\ \hat{a} \\ \hat{p}_e \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{X} & \mathbf{X}'\mathbf{Z} & \mathbf{X}'\mathbf{W} \\ \hat{\mathbf{Z}}'\mathbf{X} & \hat{\mathbf{Z}}'\mathbf{Z} + \mathbf{A}^{-1}a_1 & \hat{\mathbf{Z}}'\mathbf{W} \\ \hat{\mathbf{W}}'\mathbf{X} & \hat{\mathbf{W}}'\mathbf{Z} & \hat{\mathbf{W}}'\mathbf{W} + \mathbf{I}a_2 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}'\mathbf{y} \\ \hat{\mathbf{Z}}'\mathbf{y} \\ \hat{\mathbf{W}}'\mathbf{y} \end{bmatrix}$$

donde $a_1 = s_e^2 / s_a^2$ y $a_2 = s_e^2 / s_{p_e}^2$.

El procesamiento de este sistema de ecuaciones brinda soluciones para los efectos fijos y aleatorios. En particular se puede conocer los valores genéticos (*a*) para todos los animales y los efectos permanentes (*p_e*) para los individuos con registros. Con tales resultados se puede estimar la HPMP = *a* + *p_e*. Al mismo tiempo si se conoce la media general (**m**) entonces se puede predecir el comportamiento futuro del animal:

$$y_i = \mathbf{m} + \text{HPMP}$$

En los **MR** se asume que las medidas repetidas corresponden a un mismo rasgo con una correlación igual a uno e igual (co)varianza ambiental entre todos los pares de registros. Estas consideraciones son difíciles de cumplimentar, no obstante esto puede analizarse también como un multicaracter que elimina tales sesgos, aunque demanda mayor esfuerzo de computación.

IV.1.2. Modelo animal del día de control.

En el caso de la producción de leche es usual estimar el valor genético (VG) de los animales (hembra y machos) para la producción a 305 días de lactancia (L305), la cual es una función de los controles mensuales efectuados. El TDM es simplemente un caso particular de MR al cual se le ha incorporado algunas covariables para tomar en cuenta la forma de curva de lactancia y su procesamiento es igual que en el modelo antes descrito.

Considere que se dispone de un conjunto de datos correspondientes a los controles mensuales de producción de leche de un conjunto de vacas, distribuidas en forma similar a cualquier otro tipo de estudio. El registro mensual puede representarse por:

$$\text{Modelo B} \quad y_{ijkl} = HTD_i + \sum_{m=1}^4 b_{mj} X_m + a_k + p_k + e_{ijkl}$$

donde :

- y_{ijkl}** es el registro del día de control.
- HTD_i** es un efecto fijo del rebaño-día de control.
- X_{mj}** es una covariable, se pueden emplear diferentes funciones para describir la curva de lactancia (CL).
- a_k** es un efecto aleatorio debido a la vaca que produce el registro.
- p_{ek}** es un efecto aleatorio permanente debido a las repeticiones de registros mensuales de la misma vaca.
- e_{ijkl}** es un efecto aleatorio del error residual.

En este modelo los efectos se estiman por un sistema de ecuaciones similar al descrito previamente. La estructura de covarianza de los residuos se consideró como una matriz identidad, es decir varianza residual homogénea para todos los días de lactancia. Se han comparado diversas covarianzas, en dependencia de estudios realizados sobre la forma de la curva, en el modelo B se asume que la forma de la CL es constante entre todos los animales, lo cual es uno de sus principales problemas, no obstante es un procedimiento simple y promete mucho.

El estimado de a_k representa el valor genético del animal para la producción de leche en un día de control, de manera que se puede multiplicar por 305 y hacer de esta forma un equivalente del a_k para L305. La literatura internacional indican una correlación superior a 0.9 entre a_k para TDM y a_k para L305.(ver Swalve, 2000).

Varias ventajas e inconvenientes tiene esta alternativa.

Ventajas:

- Se representa de una forma más exacta el nivel de ambiente donde se explota el animal (grupo de contemporáneo GC como rebaño-día de control) respecto al clásico GC que considera rebaño-año-época de parto.
- Al disponerse de más registros por vacas se incrementa la precisión del estimado.
- Se pueden incluir otros factores que afectan la producción de leche, tal como la duración de la gestación, grupo de vacas acorde a su nivel de producción, estado de salud en un mes determinado, etc. Recientemente se ha incorporado los efectos de un

índice de temperatura y humedad relativa correspondiente al mes de control, para medir la variación genética de la adaptación al calor, (Ravagnolo y Misztal, 2002).

- Se puede emplear diferentes funciones como representativa de la CL.
- Las vacas pueden evaluarse con uno o más pesajes mensuales, evitando los riesgos de los factores de ajustes.
- Pueden disminuirse los costos de los sistemas de controles puesto que permite reducir el número de registros a lo largo de la lactancia.
- Es un procedimiento simple y demanda pocos recursos computacionales.

Desventajas:

- Se requiere conservar un gran volumen de información.
- Demanda mayor tiempo de computación, y modelos matemáticos mas complejos.
- Tradicionalmente el criador está más familiarizado con L305, aunque esto puede resolverse fácilmente.
- En el caso de cruzamiento B. taurus x B. indicus existe un problema de duración de la lactancia el cual no se ha estudiado adecuadamente hasta el presente.

V. Modelos multicaracteres y de regresión aleatoria.

En el ejemplo anterior se asumió que los diferentes registros mensuales de producción de leche se trataban de un mismo carácter. Sin embargo es lógico esperar que las covarianzas entre las expresiones del denominado ‘el mismo carácter’ cambien gradualmente en función del período de tiempo entre uno y otro registro. Para tomar en cuenta este hecho se recomienda aplicar otro tipo de modelo tal como **MMC** y **MRA** que tienen mucho en común ya que ambos tratan los cambios de las covarianzas en el tiempo, por ello serán presentados de conjunto. Sin embargo, previamente es necesario unos breves comentarios sobre el concepto de funciones de covarianza.

V.1. Funciones de covarianza.

El término *funciones de covarianza (CF)* fue introducido por Kirpatrick y Heckman (1989) y se define como ‘una función continua que proporciona la varianza y covarianza de rasgos medidos en diferentes puntos de la trayectoria’. Esta **CF** se usa para describir la estructura de covarianza fenotípica, pero en principio también es apropiada para definir una **CF** para cada uno de los efectos aleatorios que explican la variación. En el contexto de la genética animal, las **CF** se usan comúnmente para describir la estructura de covarianza genética.

Para estimar *la función de covarianza genética (CFG)* se debe conocer previamente la matriz de covarianza genética aditiva $\mathbf{CG}_{(a_i, a_j)}$ por los procedimientos clásicos de la genética cuantitativa. La talla de un individuo a dos edades a_i y a_j se considera que son dos caracteres diferentes y $\mathbf{CG}_{(a_i, a_j)}$ es la covar genética aditiva entre esas dos edades. Considerando que los registros de talla se miden a n edades, se tendrá entonces una $\mathbf{CG}_{n \times n}$. Los valores que componen a $\mathbf{CG}_{(a_i, a_j)}$ proporcionan un estimado directo de la **CFG** a n^2 puntos ya que :

$$\text{Ecuación 7} \quad \mathbf{CG}_{(a_i, a_j)} = \mathbf{CFG}_{(a_i, a_j)}$$

Los valores de **CFG** entre las edades medidas se estiman por interpolación usando lo que se denominan ‘smooth curves’ en las cuales se utiliza el hecho implícitamente asumido de

que las var y covar no cambian de una manera anárquica. Una variedad de procedimientos existen para estimar **CFG** a partir de la matriz $\mathbf{CG}_{(a_i, a_j)}$; Kirpatrick y col. (1990), elige las funciones ortogonales ya que los coeficientes obtenidos son muy útiles para analizar el patrón de variación genética en la trayectoria de crecimiento. Para estimar la **CF** de una matriz de datos repetidos en el tiempo, hay que utilizar funciones normalizadas y ortogonales. Los polinomios de Legendre (**LP**) son uno de ellos. Hay muchas funciones normalizadas y polinomiales, la elección de una u otra no afecta las estimaciones de **CF**. El autor favorece los polinomios (como **LP**) porque tienen una base biológica la que permite esperar que la **CF** para el crecimiento sea relativamente ‘smooth’ en lugar de oscilatorio.

Según van der Werf y Schaeffer (1997), en términos matemáticos una **CF** para la covarianza entre los valores genéticos μ_1 y μ_2 sobre un animal para rasgos medidos a las edades X_1 y X_m se estima por:

$$\text{Ecuación 8} \quad \text{cov} (m_1, m_m) = f(x_1, x_m) = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} f_i(x_1^*) f_j(x_m^*) k_{ij}$$

donde f_i es el i^{th} ($i = 0, \dots, k-1$) polinomio de orden k , x^* es la edad estandarizada y k_{ij} son los coeficientes de la **CF**. Esa edad en forma estandarizada x_i^* con valores entre -1 a $+1$ se estima por $x_i^* = -1 + 2 \left(\frac{a_i - a_{\min}}{a_{\max} - a_{\min}} \right)$ donde a_{\max} y a_{\min} corresponden con las edades máximas y mínimas respectivamente.

La **CF** puede escribirse en términos matriciales como:

$$\mathbf{CG} = \mathbf{\Phi} \mathbf{K} \mathbf{\Phi}'$$

donde **CG** es la matriz de cov genética de orden t para los VG estimado a t edades, $\mathbf{\Phi}$ es una matriz de t por k de polinomios ortogonales, esta matriz $\mathbf{\Phi}$ puede también escribirse como $\mathbf{M}\mathbf{\Lambda}$, donde \mathbf{M} es una matriz de t por k cuyos elementos se conforman por $m_{ij} = a_i^{(j-1)}$ ($i = 1, \dots, t; j = 1, \dots, k$) y $\mathbf{\Lambda}$ es una matriz de orden k con los coeficientes polinomiales. \mathbf{K} es una matriz de los k_{ij} coeficientes asociados a **CF**.

Veamos ahora parte del procedimiento desarrollado por Kirkpatrick para estimar **CFG**, se utilizará el mismo ejemplo de su publicación. Considérese la siguiente matriz de (co)var genética (**CG**) del peso vivo en ratones medidos a 2; 3 y 4 semanas de edad:

$$\mathbf{CG} = \begin{bmatrix} 436.0 & 522.3 & 424.2 \\ 522.3 & 808.0 & 664.7 \\ 424.2 & 664.7 & 558.0 \end{bmatrix}$$

En este caso simple, la edad estandarizada es $-1, 0$ y 1 , luego la matriz \mathbf{M} tendrá $t = 3$ filas (una para cada edad) y tres columnas $k=3$ (intercepto, termino lineal y cuadrático):

$$\mathbf{M} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

Para estimar el j^{th} **LP** normalizado se estimara por :

$$P_{n+1}(x) = \frac{1}{n+1} ((2n+1)xP_n(x) - nP_{n-1}(x))$$

Estas cantidades se normalizan para dar:

$$f_n(x) = \left(\frac{2n+1}{2}\right)^5 P_n(x)$$

La matriz Λ con los primeros tres primeros **LP** es:

$$\Lambda = \begin{pmatrix} 0.7071 & 0 & -0.7906 \\ 0 & 1.2247 & 0 \\ 0 & 0 & 2.3717 \end{pmatrix}$$

Sustituyendo tendremos que $\Phi = M\Lambda$:

$$\Phi = \begin{bmatrix} 0.707 & -1.225 & 1.581 \\ 0.707 & 0 & -0.791 \\ 0.707 & 1.225 & 1.581 \end{bmatrix}$$

y los coeficientes de la matriz **K** se estiman por:

$$\mathbf{K} = \Phi^{-1} \mathbf{CG}(\Phi^{-1})^{-1}$$

Haciendo las sustituciones correspondientes se tendrá:

$$\mathbf{K} = \begin{bmatrix} 1348 & 66.5 & -112.0 \\ 66.5 & 24.3 & -14.0 \\ -112.0 & -14.0 & 14.5 \end{bmatrix}$$

Mientras que la matriz de coeficientes **C** de la **CFG** se obtiene por $\mathbf{C} = \Lambda \mathbf{K} \Lambda'$:

$$\mathbf{C} = \begin{bmatrix} 808.0 & 71.2 & -214.5 \\ 71.2 & 36.4 & -40.7 \\ -214.5 & -40.7 & 81.6 \end{bmatrix}$$

Finalmente en términos de $f(x_l, x_m) = \sum_{i=0}^{k-1} \sum_{j=0}^{k-1} \Phi_i(x_l) \Phi_j(x_m) k_{ij}$ definida previamente tendremos:

$$f(x_l, x_m) = 808 + 71.2(x_l + x_m) + 36.4x_l x_m - 40.7(x_l^2 x_m + x_l x_m^2) - 214.5(x_l^2 + x_m^2) + 81.6(x_l^2 x_m^2)$$

Observe que esto es una ecuación de regresión con la cual se puede predecir el valor de y para cualquier punto entre los edades representadas en los datos.

Utilizando esta función se puede estimar la covarianza entre combinaciones de edades representadas en **CG**, mediante interpolación también se puede estimar la covarianza entre edades que nunca fueron medidas pero que están comprendidas en el rango de edades representadas en **CG**. Por ejemplo la covarianza entre el peso a la edad 3 ($x_1 = 0$) y el peso a 3.5 semanas ($x_2 = 0.5$) será: $f(0,0.5) = 808 + 71.2(0.5) - 214.5(0.5^2) = 789.9$.

Según Kirpatrick y Heckman (1989), una de las motivaciones del uso de funciones ortogonales para estimar **CFG**, es que la matriz de coeficientes **K** puede emplearse para analizar el patrón de herencia, en particular la **K** se usa para calcular la **función propia (eigenfunction-EFU)** y el **valor propio (eigenvalue-EVA)** de **CFG**. Estos **EFU** son análogos a los **vectores propios** del análisis de componentes principales¹ de matrices de covarianza. Cada **EFU** es una función continua que representa la posible deformación en la trayectoria media de crecimiento (**TMC**) del conjunto de datos estudiados. Cualquier **TMC** puede ser vista como la suma de la actual **TMC** de la población más una combinación de **EFU** de su correspondiente **CFG** aditiva. Cada **EFU** tiene un valor asociado conocido como **EVA** el cual es proporcional a la cantidad de variación genética correspondiente a ese **EFU**. Los **EFU** con mayores **EVA** son deformaciones para las cuales existe sustancial variación genética en esa población. Por tanto la forma de la **TMC** puede evolucionar rápidamente junto a tales deformaciones si están favorecidas por la selección. De acuerdo a lo dicho, los **EFU** y sus correspondientes **EVA** contienen información de gran valor para entender el potencial evolutivo de la trayectoria del crecimiento y se estiman a partir de la matriz **K**. El primer **EFU** es una deformación que incluye un incremento o decrecimiento de la talla en todas las edades. En el anexo 1 se brindan algunas explicaciones generales sobre análisis de componentes principales y una tabla de resultados para un ejemplo de clasificación del tipo en animales vacunos.

El método descrito hasta aquí, (Kirpatrick y col., 1990) es muy engorroso pues se requiere a priori de un conocimiento de la matriz **CG**. Afortunadamente existe un procedimiento REML propuesto por Meyer y Hill (1997), que realiza simultáneamente todos los estimados y que puede usar el Likelihood Ratio Test con el cual se pueden valorar diferentes modelos de covarianza.

V.2. Estimación de funciones de covarianza mediante el uso de modelos mixtos.

Si se va a emplear el concepto de **CF** en las evaluaciones genéticas, es necesario conocer como estas **CF** pueden implementarse en el sistemas de ecuaciones de modelos mixtos. Los modelos con datos repetidos, pueden tener tres componentes aleatorios : uno genético aditivo; otro de efectos ambientales permanentes y otro de efectos ambientales temporales. Estos dos últimos generalmente se identifican como medidas del error. De acuerdo a ello, se puede escribir un modelo donde no las observaciones sino los efectos aleatorios subyacentes sean reemplazados por una **CF**. Dado que las **CF** son aditivas, al igual que sus correspondientes matrices de covarianzas, ellas pueden sumarse para estimar la **CF** fenotípica asumiendo que los efectos aleatorios son no correlacionados.

¹ Ver Anexo 1 con algunas explicaciones muy generales sobre Análisis de Componentes Principales.

Considere el siguiente modelo:

$$\text{Modelo C} \quad y_i = \mathbf{a} + \mathbf{m}_i + pe_i + \epsilon_i$$

donde \mathbf{m}_i es un vector con efectos genéticos aditivos para las observaciones medidas en el animal i , mientras que pe_i y ϵ_i son vectores de efectos ambientales permanentes y temporales respectivamente. En este caso se tendrá:

$$\text{var}(\mathbf{m}_i) = \mathbf{G}_0 \quad ; \quad \text{var}(pe_i) = \mathbf{C}_0 \quad \text{y} \quad \text{var}(\epsilon_i) = \mathbf{I}\mathbf{S}_\epsilon^2$$

Si todos los animales tienen medidas en las mismas edades, \mathbf{G}_0 y \mathbf{C}_0 son iguales para cada animal. El modelo C puede verse como un modelo del tipo MMC donde la matriz residual de covarianza es:

$$\text{var}(\mathbf{e}) = \text{var}(pe_i + \epsilon_i) = \mathbf{C}_0 + \mathbf{I}\mathbf{S}_\epsilon^2 = \mathbf{R}_0$$

Dado que los errores de mediciones son independientes entre edades, solamente se escribe la CF para los efectos aditivos y ambientales permanentes. Si se asume CF estimada con LP para cada efecto aleatorio y un mismo orden de ajuste:

$$\mathbf{G}_0 = \Phi \mathbf{K}_a \Phi' \quad \text{y} \quad \mathbf{C}_0 = \Phi \mathbf{K}_p \Phi'$$

Entonces el modelo C puede escribirse como:

$$\text{Modelo D} \quad y_i = \mathbf{m} + \Phi_i a_i + \Phi_i p_i + \epsilon_i$$

Simplemente se reemplazó \mathbf{m}_i por $\Phi_i a_i$ y pe_i por $\Phi_i p_i$. Si CF se estimó con un ajuste completo, los modelos C y D son equivalentes, y tendrán las mismas valores esperados y las mismas varianzas, luego:

$$\begin{aligned} \text{var}(\Phi_i a_i) &= \Phi_i \text{var}(a_i) \Phi_i' = \Phi_i \mathbf{K}_a \Phi_i' = \mathbf{G}_0 = \text{var}(\mathbf{m}_i) \\ \text{var}(\Phi_i p_i) &= \Phi_i \text{var}(p_i) \Phi_i' = \Phi_i \mathbf{K}_p \Phi_i' = \mathbf{C}_0 = \text{var}(pe_i) \end{aligned}$$

Estas equivalencias no serán exactas si los animales tienen registros a diferentes edades, dado que la CF se estimará con un orden de ajuste reducido. Note que este modelo CF solo requiere una matriz Φ_i diferente para cada conjunto de edades, pero los coeficientes de regresión tienen la misma estructura de covarianza (\mathbf{K}_a y \mathbf{K}_p) para cada animal, independientemente del grupo de edades a que fueron medidos. Note también que en el modelo D el número de efectos genéticos aditivos es igual al orden de ajuste de la CFG, de orden \mathbf{K}_a , de igual forma el número de efectos ambientales permanentes es igual al orden de ajuste de \mathbf{K}_p .

Debe apuntarse que el modelo D es un modelo de regresión en el cual los datos están regresados sobre LP con variables regresoras en Φ_i y coeficientes de regresión en \mathbf{a} y \mathbf{p} . Estos coeficientes no son los mismos para cada animal, sino que son representativos de una población de coeficientes de regresión, en otras palabras, los coeficientes de regresión en \mathbf{a} y \mathbf{p} son coeficientes de regresión aleatorias con $\text{var}(\mathbf{a}) = \mathbf{K}_a$ y $\text{var}(\mathbf{p}) = \mathbf{K}_p$. De hecho, se ha

conformado un modelo mixto **MMC** con una **CF** en forma de un modelo uní variado de **MRA**, en el que cada efecto aleatorio tiene **k** coeficientes de regresión aleatorios. Un modelo de este tipo para **n** observaciones de **q** animales puede escribirse de la siguiente forma:

$$\text{Modelo E} \quad y = Xb + \sum_{j=0}^{k-1} Z_j a_j + \sum_{j=0}^{k-1} Z_j p_j + \epsilon$$

Donde Z_j son matrices de **n** por **q** para los i^{th} polinomios y \mathbf{a}_j y \mathbf{p}_j son vectores de coeficientes de regresión aleatorios para todos los animales para los efectos genéticos aditivos y los efectos ambientales permanentes. La matriz Z contiene las variables de regresión, es decir aquellas donde los coeficientes son los polinomios en Φ . Si los vectores \mathbf{a}_j y \mathbf{p}_j son unidos y ordenados por animal, cada uno tendrá **k** coeficientes en **a** y **k** coeficientes en **p**. En tal caso la matriz Z se puede escribir como un bloque de una matriz diagonal de orden **n** por **k*p**, con cada bloque animal $Z_i = \Phi_i = M_i \Lambda$.

El modelo mixto puede escribirse como:

$$\text{Modelo F} \quad \mathbf{y} = \mathbf{Xb} + \mathbf{Z}^* \mathbf{a} + \mathbf{Z}^* \mathbf{p} + \boldsymbol{\epsilon}$$

con $\mathbf{a}' = \{a'_1, \dots, a'_q\}$ y $\mathbf{p}' = \{p'_1, \dots, p'_q\}$, en los que \mathbf{a}_i y \mathbf{p}_i son un conjunto de coeficientes de regresiones aleatorias para el animal **i** para los efectos genéticos aditivos y de ambiente permanente respectivamente. Si todos los animales tienen medidas en los mismos puntos de edad, todas las Z son iguales y $Z^* = I_q \otimes \Phi$. Las varianzas y covarianzas de los efectos aleatorios serán: $\text{var}(\mathbf{a}) = \mathbf{A} \otimes \mathbf{K}_a$, $\text{var}(\mathbf{p}) = \mathbf{I} \otimes \mathbf{K}_p$ y $\text{cov}(\mathbf{a}, \mathbf{p}) = 0$. En donde \mathbf{K}_a y \mathbf{K}_p son los coeficiente para la **CF** para los efectos genéticos aditivos y de ambiente permanente respectivamente.

Los modelos **MMC** y **MRA** tienen mucha similitud. Para un modelo **MMC** con **t** rasgos, tenemos:

$$\text{Modelo G} \quad \mathbf{y} = \mathbf{Xb} + \mathbf{Zu} + \mathbf{e}$$

En este caso $\text{var}(\mathbf{y}) = \mathbf{ZGZ}' + \mathbf{R}$. Si \mathbf{u} es un vector de efectos genéticos aditivos del animal entonces $\text{var}(\mathbf{u}) = \mathbf{G}$. Si \mathbf{u} así como \mathbf{y} están ordenados por animal y rasgo dentro animal, Z es una matriz diagonal de bloques, donde estos bloques diagonales Z_j pertenecen a cada animal y Z_j es una matriz de \mathbf{n}_j por **t**, en la que \mathbf{n}_j es el numero de caracteres por animal **j** y $Z_j(m,k)$ será igual a 1 si el m^{th} registro es medido para el carácter **k** y 0 en caso contrario. Si se definió que $\text{var}(\mathbf{u}) = \mathbf{G} = \mathbf{A} \otimes \mathbf{G}_0$ y $\text{var}(\mathbf{e}) = \mathbf{R} = \mathbf{I} \otimes Z_i \mathbf{R}_0 Z_i'$ con \mathbf{G}_0 y \mathbf{R}_0 matrices de orden **t** de varianzas y covarianzas residuales, entonces si no hay valores perdidos y las matrices de incidencia son iguales para cada carácter tendremos que $Z_i = I_t$ y $\mathbf{R} = \mathbf{I} \otimes \mathbf{R}_0$.

Si el modelo de **CF** confeccionado por **MRA** tiene un ajuste de orden **k** y $\mathbf{k} \leq \mathbf{t}$ y todos los animales tienen medidas en cada una de las **t** edades, en tales condiciones se pueden comparar los modelos **MMC** y **MRA**:

$$\text{var}(Z_a^*) = \mathbf{A} \otimes \Phi \mathbf{K}_a \Phi' \equiv \mathbf{A} \otimes \mathbf{G}_0 = \text{var}(\mathbf{Zm})$$

y

$$\text{var}(Z_p^* + \epsilon) = I_q \otimes \Phi K_p \Phi' + I S_e^2 \cong I_q \otimes R_0 = \text{var}(e)$$

El vector de los valores genéticos **MMC** para el animal **i**, **u_i** se puede escribir como $\Phi \mathbf{a}_i$, o sea los coeficientes de regresión aleatorios se multiplican por Φ . Cuando el orden de ajuste es completo (**k** = **t**) el **MMC** y el **MRA** son exactamente equivalentes. Cuando **k** es mucho menor que **t** la estructura de covarianza descrita por **CF** será más suave y probablemente más correcta que para **t** rasgos.

Cuando diferentes animales tienen medidas para diferentes conjuntos de puntos, el bloque diagonal en la matriz de incidencia **Z** no serán todas iguales y la varianza de **y** será:

$$\text{var}(y_i) = \Phi_i K_a \Phi_i' + \Phi_i K_p \Phi_i' + I S_e^2 = G_{0i} + R_{0i}$$

donde G_{0i} y R_{0i} son las covarianzas genéticas y ambientales para los rasgos medidos en los puntos de la *i* edad. En **MMC** con valores perdidos se obtienen las soluciones para todos los animales (con datos y sin datos) y caracteres a partir de correlaciones o de información del mismo rasgo medido sobre animales emparentados. En los modelos **CF** la información sobre cada animal para cada efecto aleatorio se conglo mer a en **k** coeficientes de regresión y las soluciones de los valores genéticos para cada punto de edad deseada puede generarse a partir de $\mathbf{u}_i = \Phi \mathbf{a}_i = \mathbf{M}_i \Lambda \mathbf{a}_i$ donde las edades están en \mathbf{M}_i .

Los resultados de **MRA** son más flexibles que un **MMC** para **t** rasgos ya que puede estimar soluciones a diferentes edades y los resultados son equivalentes a los de **MMC** en función, obviamente, de la precisión con que se ajusto la **CF** sobre la matriz de covarianza de **t** dimensiones.

VI. Aplicación práctica y discusión general.

Varios anexos se han incorporado donde se presentan ejemplos de aplicación de la mayor parte de los elementos expuestos en este documento. El objetivo no es discutir la validez de estos resultados sino de contribuir a crear una masa crítica de ideas acerca del análisis de datos de la producción animal y particularmente los procedimientos para estimar el componente genético de la variabilidad total que se manifiesta.

Finalmente se presenta unas reflexiones sobre un viejo concepto de Plasticidad y una primera aplicación práctica con datos de producción, con el objetivo de fomentar una discusión sobre el posible uso adicional de los nuevos procedimientos de regresión aleatoria.

VI-Referencias.

- Apiolaza, L.A.; Gilmour, A.R.; y Garrick, D.J., 2000. Variance modelling of longitudinal height data from a *Pinus radiata* progeny test. *Can. J. For. Res.* 30: 645-654.
- Fitzhugh, H.A., 1976. Analysis of growth curves and strategies for altering their shape. *J. Anim. Sci.* 42(4),1036-1051.

- Hill, W.G.; Bishop, S.C.; McGuirk B.; Mc Kay, J.C.; Simm, G. y Webb, A.J., 2000. The challenge of genetic change in animal production. BSAS Occasional Publication No 27, 86 pag.
- Johnson, D.L. y Thompson, R., 1995. Restricted Maximum Likelihood estimation of variance components for univariate animal models using sparse matrix techniques and average information. *J. Dairy Sci.* 78: 449-456.
- Kennedy B.W., 1989. Animal Model BLUP. Erasmus Intensive Graduate Course, Trinity College, Dublin, Ireland.
- Kirpatrick, M.; Lofsvold, D. y Bulmer, M., 1990. Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* 124: 979-993.
- Kirpatrick, M., y Heckman, N., 1989. A quantitative genetic model for growth, shape and other infinite-dimensional characters. *J. Math. Biol.* 27: 429-430.
- Luiting, P.; Vangen, O.; Rauw, W.M.; Knap, P.W. y Beilharz, R.G., 1997. Physiological consequences of selection for growth. 48th Meeting of EAAP, Vienna, Austria, 21 pag.
- McKay, L.R.; Schaeffer, L.R. y McMillan, I., 2002. Analysis of growth curves in rainbow trout using random regression. Seventh World Congress of Genetics Applied to Livestock Production, paper 241.
- Meyer, K. y Hill, W.G., 1997. Estimation of genetic and phenotypic covariance functions for longitudinal or 'repeated' records by restricted maximum likelihood. *Livest. Prod. Sci.* 47: 185-200.
- Ptak, E. y Schaeffer, L.R., 1993. Use of test day yields for genetic evaluation in dairy sires and cows. *Livest. Prod. Sci.* 34: 23.
- Rauw, W.M.; Kanis, E.; Noordhuizen-Stassen, E.N. y Grommers, F.J., 1998. Undesirable side effects of selection for high production efficiency in farm animals: a review. *Livestock Prod. Science* 56 : 15-33.
- Ravagnolo, O. y Misztal, I., 2002. Effect of heat stress on nonreturn rate in Holstein cows: genetic analyses. *J. Dairy Sci.* 85:3092-3100.
- Schaeffer, L.R. y Dekkers, J.C.M., 1994. Random regressions in animal models for test-day production in dairy cattle Proceedings of Fifth World Congress of Genetics Applied to Livestock Production, vol. XVIII. Guelph, Ont., Canada, pp. 443-446.
- Schnyder, U.; Hofer, A. y Kunzi, N., 2001. Impact of variation in length of individual testing periods on estimation of covariance components of a random regression model for feed intake of growin pigs. *J. Anim. Breed. Genet.* 118: 235-246.
- Swalve, H.H., 2000. Theoretical basis and computational methods for different test-day genetic evaluation methods. *J. Dairy Sci.* 83: 1115-1124.
- van der Werf, J. y Schaeffer, L.R., 1997. Random Regression in Animal Breeding. Course Notes . CGIL June 25-28; 58 pag.

ANEXO 1.

Notas sobre el uso y significado biológico del Análisis de Componentes Principales.

El análisis de componentes principales (ACP) es una técnica multivariada que se emplea para examinar las relaciones entre varias variables cuantitativas. Dado un conjunto de p variables numéricas, el ACP hace transformaciones lineales de cada una y crea un nuevo conjunto de p variables que se denominan componentes principales (CP). Los CP se obtienen estimando los eigenvalue (EVA) y los eigenvector (EVE) de la matriz de covarianza o de correlaciones, aunque se prefieren estas últimas.

Dado que el ACP se lleva a cabo con las variables originales expresadas en forma estandarizada, la varianza será igual al número total de variables. Si los CP estimados se sortean por orden descendente de EVA, entonces el primer CP explica la mayor parte de la varianza de las variables. Los EVA son las varianzas de las variables PC. El mayor EVA se considera como una indicación del verdadero número de variables (esto es llamado FACTOR) que se necesita para describir la conducta del conjunto total de variables.

Por otra parte los EVE son los coeficientes de la transformación lineal de las variables estandarizadas. Estos EVE se usan para crear los valores observados de las nuevas variables CP y se utilizan para describir la estructura de tales variables. Los EVE se emplean también para computar las correlaciones entre las variables originales y las nuevas variables del FACTOR. Estas correlaciones, llamadas FACTOR LOADINGS ó FACTOR STRUCTURES, son usualmente mejores indicadores de la naturaleza de las variables PC que los propios EVE.

Los CP tienen algunas propiedades interesantes:

- § Los EVE son ortogonales, de manera que los CP representan una dirección conjunta perpendicular a través del espacio de las variables originales.
- § Los coeficientes del CP están incorrelacionados entre sí.
- § El primer CP tiene la mayor varianza de cualquier combinación lineal de las variables observadas. El j CP contiene la mayor varianza de cualquier combinación ortogonal a los $j-1$ CP. El último CP tiene la menor varianza de cualquier combinación de las variables originales.
- § El primer CP₁ mantiene la mayor correlación con las variables originales y por tanto minimiza la correlación residual.
- § Los coeficientes de CP₁ contienen la mayor varianza generalizada posible de cualquier combinación de variables originales.
- § El primer CP proporciona la solución de mínimos cuadrados al modelo:

$$Y = XB + E$$

donde:

- Y: es una matriz $n \times p$ de las variables originales estandarizadas.
- X: es una matriz $n \times j$ de los coeficientes sobre el primer PC.
- B: es una matriz $n \times p$ de EVE.
- E: es una matriz $n \times p$ y se desea minimizar la traza de $E'E$ (suma de todos los elementos cuadrados en E).

En otras palabras el primer j CP es el mejor predictor lineal de las variables originales. Cuando pocos CP explican la mayor parte de la varianza de las variables originales quiere decir que existe información redundante. A mayor valor del coeficiente significa que es la variable más importante para discriminar medidas.

En SAS existen tres tipos de procedimientos para tratar relaciones entre variables, independientemente de las definiciones sobre cuales son dependientes y cuales independientes. Estos son:

- **Principal Components (ACP).** Deriva un pequeño conjunto de combinaciones lineales (CP) de un conjunto de variables que retienen la mayor parte de la varianza de las variables originales.
- **Factor Analysis (AFC).** Se usa para explicar las correlaciones o covarianzas entre un conjunto de variables, en términos de un número limitado de variables subyacentes.
- **Canonical Analysis (ACC).** Se emplea para sumarizar las relaciones entre dos conjuntos de variables mediante un número pequeño de combinaciones lineales de cada conjunto que tiene la mayor posible correlación entre conjuntos.

Los resultados de AFC y ACP son iguales excepto que los coeficientes de AFC son normalizados y de varianza unidad (AFC produce coeficientes de ACP) mientras que PROC ACP produce coeficientes con varianza igual al correspondiente EVA. El método implícito en AFC es el APC. Este procedimiento lleva a cabo una variedad de análisis de FACTOR y PRINCIPAL COMPONENTS. Como input puede emplear una matriz de correlaciones o de covarianzas o los datos originales. Sus resultados pueden salvarse en conjunto de datos SAS. Como input puede también emplear output de otros procedimientos. Los coeficientes del GLM (manova) pueden emplearse en este procedimiento.

Para ejemplificar el uso de ACP se utilizaran los resultados publicados por Menéndez Buxadera y col (1998) en un artículo donde se estudiaron los resultados de $p = 25$ variables fenotípicas de la clasificación lineal del tipo en 1523 vacas Zebu de primer parto. Este conjunto produce una matriz de correlaciones $M = p(p-1)/2$, es decir 300 estimados lo cual es muy difícil e inapropiado para interpretar. Este conjunto de datos se estudio mediante un analisis de ACP y los principales resultados se presentan en la tabla1. De los 25 PC se retuvieron los 5 primeros ($PC_1 \dots PC_5$) cuyos EVA explicaron el 76.6% de la varianza original de los datos [$100*(14.71 + 1.67 + \dots + 0.72)/25 = 76.6$], en particular PC_1 y PC_2 son los dos componentes que mayor influencia ejercen.

Notese que todos los coeficientes de PC_1 son positivos lo cual quiere decir que su uso permite diferenciar entre animales con mayores versus menores atributos en aquellas características indicadas en la última columna de la tabla 1. Por esto es que se considera a PC_1 como una expresion general del factor '*talla*'. Para PC_2 , generalmente conocido como un componente que permite explicar cambios en la '*forma*' se pueden observar coeficientes positivos y negativos que permiten identificar animales de mejor apariencia general y mejores patas y pezuñas, sin embargo son mas pequeños y de menor capacidad corporal e inferior características carnicas. Es importante señalar que este PC_2 está

incorrelacionado con PC₁ y ambos explican el 65.5% de la varianza de las variables originales, y de conjunto brindan la misma información que aproximadamente 16 variables originales (14.71 + 1.67 = 16.38).

Tabla 1. Resultados del Análisis de Componentes Principales para rasgos de clasificación lineal del tipo en animales Zebu (tomado de Menéndez Buxadera y col 1998).

RASGO	PC1	PC2	PC3	PC4	PC5	CARACTERÍSTICAS GENERALES
Tamaño	0.6513	-0.2933	0.2220	0.4399	-0.0187	APARIENCIA GENERAL
Temperamento	0.6528	0.3627	0.4710	-0.1006	-0.1018	
Sexualidad	0.6918	0.3969	0.4389	-0.0999	-0.0453	
Ubre	0.6709	0.2795	0.2571	-0.0838	0.0193	
Cabeza	0.7190	0.2117	0.2619	-0.0525	0.1066	
Cuello	0.8241	0.0413	0.0693	-0.1734	0.0501	
Cruz	0.6869	-0.2159	-0.0822	-0.1421	0.4806	
Giba	0.7518	-0.0257	-0.0894	-0.3178	0.4002	
Pecho	0.8702	-0.1588	0.0640	0.0465	0.0429	CAPACIDAD CORPORAL
Torax	0.8090	-0.3815	0.0908	0.1425	-0.0798	
Costilla	0.8205	-0.2301	-0.0287	0.0207	0.1482	
Profundidad	0.8214	-0.3535	0.1040	0.1375	-0.9397	
Longitud	0.8307	-0.0144	0.0721	0.0395	-0.0622	GRUPA
Ancho	0.8301	-0.1523	-0.0453	-0.2362	-0.3364	
Inclinación	0.6638	0.0183	-0.3082	-0.2362	-0.3364	
Nacim. del rabo	0.6964	0.1044	-0.3083	-0.4222	-0.1816	
Vulva	0.8226	0.0098	-0.0457	-0.0032	-0.0229	
Cuartilla	0.7321	0.4252	-0.1929	0.1721	0.1116	PATAS Y PEZUÑAS
Corvejon	0.6796	0.2639	-0.3437	0.3163	0.1054	
Pezuñas	0.7075	0.4270	-0.2166	0.1961	0.1141	
Aplomos	0.7345	0.3755	-0.2048	0.3139	-0.1183	
Dorso	0.8240	-0.2406	-0.0439	-0.0725	-0.0563	CARACTER CARNICERO
Lomo	0.8739	-0.2289	-0.0349	0.0444	-0.0617	
Cuarto anterior	0.8706	-0.2260	0.0105	-0.0127	-0.0569	
Cuarto posterior	0.8472	-0.0365	-0.0827	-0.1579	-0.1625	
VECTOR PROPIO	14.71	1.67	1.10	.95	.72	
% VARIANZA ACUMULADA	58.9	65.5	69.9	73.7	76.6	

De hecho, esto representa una clara evidencia de la capacidad de ACP para reducir el número de variables originales disponibles en un pequeño conjunto de combinaciones lineales de los rasgos originales y al propio tiempo es una herramienta muy importante para brindar una explicación biológica a la estructura de varianza y covarianzas de este gran volumen de caracteres medidos en cada animal los cuales se pueden agrupar según su '*talla*' y su '*forma*'.

Aun cuando está implícito en las definiciones de ACP debe reiterarse que las variables originales tienen que expresarse en forma estandarizada, es decir $Z_i = (X_i - \bar{X}) / s$.

ANEXO 2

PROCEDIMIENTO PARA COMPARAR DIFERENTES MODELOS DE ESTIMACION DE COMPONENTES DE VARIANZA.

Muchas veces nos enfrentamos a este tipo de problema, cuando se comparan modelos con variaciones en el número de componentes de varianza a estimar. Debe tomarse en cuenta que se asume que los efectos fijos son los mismos. El ejemplo mas simple es un modelo con efectos geneticos directos y maternos y se desea conocer si la covarianza entre ambos, debe o no incorporarse en el analisis final. Según varios trabajos publicados, lo que se debe hacer es comparar los resultados de logL de dos modelos, uno con la inclusion de cov y otro excluyendo. Si la diferencia entre ambos logL es superior a un limite establecido en la tabla de chi-cuadrado entonces se asume que el modelo con el mayor logL es el adecuado. La crítica que se plantea es que en este enfoque se comparan dos modelos con diferentes número de parámetros estimados y generalmente sucede que aquel modelo con mas parametros tiene un logL superior.

De acuerdo a Foulley y Robert-Granie (2002) y copio textualmente:

“It should be strongly emphasized that information criteria provide rules of thumb to discriminate between several statistical models. They should never be used or interpreted as form of statistical test of significance”. (Verkele & Molenberghs, 2000).

Según este punto de vista, todos los procedimientos empleados solo deben interpretarse como indicativos y no como pruebas formales. La literatura mas frecuente utiliza dos metodos adicionales a logL, conocidos como Akaike Information Criteria (AIC) o el Bayesian Information Criteria (BIC) que brevemente presentamos la síntesis de ambos.

$$\begin{aligned} \text{AIC} &= -2 \log L + 2p \\ \text{BIC} &= -2 \log L + p \log(N-x) \end{aligned}$$

donde:

p es el número de componentes de (co)varianza estimados.

N es el número de observaciones.

x es el número de efectos fijos (N-x , son los grados de libertad del residuo)

En terminos generales el criterio que se sigue es: mayor valor de logL y menor valor para AIC y BIC. Para ejemplificar estas aplicaciones se usara el ejemplo de la publicación de Meyer (2001). En este artículo el autor comparó un total de 34 modelos de regresión aleatoria con datos de crecimiento predestete en dos razas de vacuno. De ellos solo se comentará aquellos seleccionados como los mejores por el autor para una de las razas.

Tabla 1. Indicadores de logL; AIC y BIC para seleccionar el mejor modelo de regresión aleatoria, según la publicación de Meyer (2001).

Modelo^a	No parámetros^b	logL	AIC	BIC
5062	43	-39.4	164.9	503.9
6064	56	-14.8	141.5	583.0
6066	67	-6.2	146.4	674.6

a)-Se refiere al orden de ajuste para los efectos genéticos aditivos directos ($d = 5$); genético aditivo materno ($m = 0$); ambiente permanente directo ($p_i = 6$) y ambiente permanente materno ($p_m = 2$).

b)-El número de parámetros se estima por el número de componentes de varianza estimados, por ejemplo para el modelo 5062, esto es 15 componentes para $d + 0$ para $m + 21$ para $p_i + 3$ para $p_m + 3$ de la varianza residual heterogénea + la media = 43. Algunos autores no incluyen la media.

De acuerdo al propio autor el modelo 6066 es el más completo y el de mayor logL, sin embargo el modelo 6064 no difiere significativamente y tiene 11 parámetros menos (67-56) a estimar y fue seleccionado como el mejor modelo según AIC, que penaliza por el número de parámetros. La tabla de chi-cuadrado para $\alpha = 0.05\%$ y 11 grados de libertad es 19.65.

El análisis de los coeficientes de regresión aleatorio del modelo 6066 indica que los términos cuárticos y quinticos para d y p_m presentaron muy poca variación (es decir los eigenvalues son casi cero y no contribuyen en forma importante a la varianza explicada) por tanto fueron eliminados y se corrió el modelo 5062 que tiene solo 43 parámetros y fue seleccionado por BIC como el mejor modelo, que hace una mayor restricción en comparación a AIC. Como se puede apreciar en este ejemplo los dos criterios informativos (AIC y BIC) pueden arrojar resultados diferentes, lo importante es la discusión de estos modelos. No siempre se alcanza este tipo de variación en las comparaciones. Así por ejemplo en el artículo de Fischer y col (2004) sobre crecimiento en ovinos, se manifestó uniformidad en la elección del mejor modelo.

Referencias.

Fischer T.M., J.H.J. van der Werf, R.G. Banks and A.J. Ball 2004. Description of lamb growth using random regression. Liv. Prod. Sci. 2004 accepted.

Foulley J. L. and C. Robert-Granie 2002. Basic tools for the statistical analysis of longitudinal data via mixed model. August 18, 2002, Montpellier, France 159 pag.

Meyer K. 2001. Estimates of direct and maternal covariance functions for growth of Australian beef calves from birth to weaning. Gen. Sel. Evol. 33:487-514.

Parámetros genéticos del peso vivo de machos Cebú en prueba de comportamiento en condiciones de pastoreo de Cuba, mediante modelo animal univariado, multicares y regresiones aleatorias¹.

A. Menéndez Buxadera^a; D. Guerra^b; T. Planas^c, F. Ramos^c y L. Fernández Chuairey^d

a- Centro de Control Pecuario. c-Dirección de Genética.

Ministerio de la Agricultura, Conill y Boyeros C. de La Habana, Cuba.

b- Centro de Investigación Mejora Animal.

Loma de Tierra, Cotorro, C. de La Habana, Cuba.

d- Universidad Agraria de La Habana "Fructuoso Rodríguez Pérez",

San José de la Lajas, La Habana Cuba.

Resumen

Los resultados individuales del peso vivo PV_i ($i = 7, 8...18$ meses de edad) de 1091 machos jóvenes (13092 estimaciones de PV_i) Cebú de raza pura fueron estudiados mediante modelo animal univariado (MAU); de multicares (MAM) y de regresiones aleatorias (MRA), con el objetivo de estimar las magnitudes de las (co)var genéticas en toda la trayectoria del periodo de tiempo representado. Estos animales eran hijos de 107 sementales y 893 madres y estaban sometidos a prueba de comportamiento en condiciones de pastoreo de Cuba con una duración de 12 meses. Las varianzas genéticas manifestaron un patrón poco homogéneo con estimados altos y bajos cuando se emplea MAU y MAM; por el contrario los resultados de MAM presentaron una tendencia armoniosa en todo el periodo de tiempo. Los estimados de heredabilidad (h^2), fueron muy semejantes en los tres procedimientos aunque se evidencio mayor coherencia en MRA, en general existe una ligera depresión en h^2 en la medida que avanza la edad, con valores promedios de $h^2 = 0.27$ para PV_7 y $h^2 = 0.23$ para PV_{18} . Las correlaciones genéticas (R_g) fueron $R_g > 0.9$ entre los PV_i a edades adyacentes y disminuyen en la medida que se incrementa la diferencia de edad entre las estimaciones. Los estimados a PV_{13} o mas meses de edad manifestaron una $R_g > 0.8$ con PV_{18} lo cual sugiere la posibilidad de una selección temprana de los animales. Como mejor MRA se selecciono uno con un orden de ajuste de 5 para la matriz de coeficientes de regresión genético (Ka) y de orden 6 para los efectos de ambiente permanentes debido a las repeticiones de la variable dependiente en cada animal. El 71.8 % y el 24.1% de la variación fue explicado los dos primeros valores propios de la matriz Ka , lo cual sugiere posibles cambios en la forma de la curva de crecimiento mediante el trabajo de selección y mejora genético. Esta posibilidad quedo ejemplificada cuando se selecciono el mejor 2% de los animales en prueba según su Valor Genético para PV_{18} , dentro de este grupo de sementales elites se pudo identificar dos subconjunto de animales, uno con un comportamiento estable en toda la trayectoria mientras que el otro manifestó una forma de curva de crecimiento totalmente diferente y dinámica. El estudio demuestra las posibilidades del uso de las regresiones aleatorias.

a) email: cencop@minag.gov.cu

Palabras claves: Crecimiento en Cebú; parámetros genéticos; modelo animal; regresiones aleatorias.

Enviado a publicación Rev. Cubana de Ciencia Agrícola 2005.

Introducción

En una publicación anterior se presentaron algunas causas de variación que afectan la prueba de comportamiento (PC) en machos jóvenes de la raza Cebú en condiciones de pastoreo de Cuba, las cuales pueden producir sesgos en los resultados de la estimación del Valor Genético (VG) de este tipo de animales, (Menéndez Buxadera y col., 2004). Al propio tiempo en este artículo se muestran evidencias de una significativa variación individual en los parámetros que definen la curva de crecimiento, lo cual puede tener una importante significación práctica y económica.

El método actual de evaluar el VG en este sistema de PC se basa en el uso de un modelo animal, el cual mejora sustancialmente la precisión de los estimados (Guerra y col., 2002). En este procedimiento solo se considera el peso (o la ganancia diaria) de cada animal al final de los 12 meses de duración de la PC. En tal sentido vale la pena cuestionarse: por que utilizar una sola estimación del peso vivo (*una variable resumen*), cuando están disponibles los registros sistemáticos y periódicos del peso vivo de cada animal (*variables originales repetidas*)?. Esta misma interrogante fue planteada en ganado de leche, existiendo consenso actual en emplear los pesajes mensuales de control en la evaluación genética de las razas lecheras (ver Schaeffer, 2004). En ganado de carne también se han aplicado estos mismos puntos de vistas (Meyer, 2004).

Aquellos caracteres que pueden medirse repetidamente a lo largo de la trayectoria de un periodo de tiempo de un animal, reciben el nombre de datos longitudinales. Desde el punto de vista genético, este tipo de datos requieren de procesamientos estadísticos algo diferentes a los que mas usualmente se han empleado, no obstante, los mismos pueden brindar importantes beneficios a los objetivos del programa de mejora, particularmente necesarios para condiciones tropicales. Según la literatura disponible, las causas de variación de tales *variables repetidas* en el tiempo, se pueden estudiar mediante modelos de repetibilidad, modelos multicares y de regresiones aleatorias (van der Werf y Schaeffer 1997).

Recientemente se ha publicado una revisión muy completa sobre los parámetros genéticos de algunos rasgos de interés económicos en ganado vacuno del tipo *B. taurus*, *B. indicus* y poblaciones cruzadas en países tropicales (Lobo y col., 2000) que pueden servir de guía para los programas de selección en estas regiones. Sin embargo, para el caso de los animales Cebuínos, las referencias disponibles pueden estar sujetas a sesgos importantes debido a problemas de métodos de estimación, cantidad de observaciones etc, los cuales habían sido justamente señalados por Mercadante y col., (1995). Estos últimos autores presentaron 95 estimados de heredabilidad (h^2) en animales del tipo Cebú con valores entre $h^2 = 0.24$ a $h^2 = 0.31$ para el peso vivo entre los 12 y 18 meses de edad. Los resultados de Cuba son muy similares a estos valores promedios (Rico y col., 1995), no obstante Guerra y col., (2002) estudiaron mediante un modelo animal completo, una muestra de cerca de 10000 animales Cebú en PC en Cuba y estimaron que $h^2 = 0.20$ para el peso vivo a 18 meses de edad, valor muy semejante a lo presentado por Lobo y col., (1994) con animales Nelore de Brasil y un similar procedimiento estadístico. Por otro lado Plasse y col., (2002), estudiando los resultados del ganado Brahman de Venezuela sometidos a un programa de selección por 30 años, presentaron estimados de $h^2 = 0.13$ para el peso vivo a 18 meses. Independientemente de estas inconsistencias, en los trabajos citados se presentan

evidencias que demuestran la existencia de una importante variabilidad de origen genético en las razas B. indicus que justifican y brindan amplias perspectivas económicas a la aplicación de programas de mejora genética en el trópico.

Los estudios sobre correlaciones genéticas entre peso vivo a diferentes edades no son tan abundantes en animales Cebú. En la revisión de Mercadante y col., (1995) solo se presentan 9 estimados entre peso vivo a 12 y 18 meses de edad, con valores que oscilan entre 0.59 a 1.11 y una estimación promedio de 0.85. Rangos muy similares fueron publicados por Rico y col., (1995) con una muestra de 1441 machos jóvenes Cebú cubanos con pesos vivos entre 6 y 18 meses de edad. Este patrón es muy semejante a la mas completa revisión realizada sobre parámetros genéticos con animales vacunos mantenidos en condiciones tropicales (Lobo y col., 2000) y sugieren la existencia de una base genética común para el crecimiento en este período de tiempo.

Las tendencias antes mencionadas se han basado en estudios que utilizaron modelos de dimensiones finitas, es decir, rasgos medidos en unas pocas edades. Mas recientemente se han desarrollado nuevos procedimientos, tal como modelo animal multi carácter y regresiones aleatorias, que han demostrado manifestar no solo menores sesgos en sus resultados sino que han posibilitado identificar la existencia de una importante variabilidad genética adicional ligada a la forma de la curva de crecimiento la cual puede proporcionar importantes beneficios en el aumento de la productividad en animales Cebú (Albuquerque y Meyer, 2001; Nobre y col., 2003).

De acuerdo a lo expuesto el objetivo de este trabajo es realizar un análisis comparativo de los parámetros genéticos del peso vivo de machos jóvenes Cebú en prueba de comportamiento en condiciones de pastoreo de Cuba, mediante la aplicación de modelo animal uni variado, multi carácter y de regresiones aleatorias.

Material y Método.

Para este estudio se utilizaron los resultados de la PC de 1091 machos de la raza Cebú, mantenidos en condiciones de pastoreo de una gran empresa ganadera de Cuba, entre los años 1981 y 1996. Todos los animales se pesaron mensualmente entre el inicio de la PC (inmediatamente después del destete) hasta los 18 meses de edad aproximadamente (12 meses de duración de la PC). Estos animales eran hijos de 107 sementales (85% de los cuales pertenecen a inseminación artificial) y 893 madres, el pedigree fue extendido por la vía padre y madre de cada macho joven con registros, para un total de 2291 animales. Mas detalles del sistema de manejo utilizado en esta PC y de la edición de los datos pueden obtenerse en Planas y Rico (1990) y Menéndez Buxadera y col., (2004) respectivamente.

Varios modelos estadísticos previos fueron desarrollados, los cuales indicaron que los efectos maternos, tanto genéticos como de ambiente permanente, no alcanzaron el nivel de significación, tampoco la edad al parto de la madre ejerció influencia sobre la variable dependiente estudiada en todo el trayecto entre los 7 y 18 meses de edad. En general, estos resultados pueden ser los esperados ya que las evidencias indican un patrón muy semejante para el crecimiento post destete en este tipo de animales (Albuquerque y Meyer, 2001; Nobre y col., 2003). El modelo de análisis final no considero estos efectos y los datos disponibles fueron estudiados mediante modelo animal de acuerdo a tres estrategias entre las cuales existe mucho en común según se presentara mas adelante. Todos los estimados

se llevaron a cabo con el software ASREML desarrollado por Gilmour y col., (2000). En forma matricial el modelo general aplicado fue el siguiente:

$$\text{Modelo 1} \quad y_{ij} = Xb + Z_1a + e_1$$

donde :

y_{ij} es un vector de $n \times ij$ correspondientes a los n registros y los j^{th} caracteres analizados simultáneamente en cada i animal. Cuando $j = 1$ se trata de un modelo animal univariado (MAU), para $j = 1, 2, \dots, t$ estimaciones de PV_{ij} entonces el modelo animal es multivariado (MAM).

b es un vector de efectos fijos debido a grupo de contemporáneos (GC) con 44 niveles, similar para los procedimientos MAU y MAM .

a es un término aleatorio debido al efecto genético aditivo del animal.

e_1 es un efecto residual aleatorio.

X y Z_1 son matrices de incidencia que conectan los efectos fijos y aleatorios con el vector de datos.

Los registros disponibles contenían el peso vivo (PV_j para $j = 7, 8 \dots 18$ meses de edad) de cada animal. Esta base de datos se utilizó en MAU y MAM. Los componentes de varianza estimados según MAU (12 corridas individuales para PV_7 a PV_{18}) fueron incorporados como parámetros iniciales en MAM. Para los estimados de (co)var se realizaron 38 corridas del MAM con varias combinaciones para $j = 2, 3$ y 4 rasgos y los valores que se presentan son los promedios en los que estaban incluidos los correspondientes PV_j .

Para la aplicación de modelo de regresión aleatoria (MRA) se requiere una preparación especial de los datos con una estructura similar al caso de un modelo animal univariado. Para ello se confeccionó un fichero longitudinal, en el cual cada animal está representado repetidamente con el PV estimado en cada una de las edades dentro de la trayectoria entre el inicio (7 meses de edad) y fin de la PC (18 meses de edad). El nuevo conjunto contenía un total de 13092 observaciones pertenecientes a los 1091 animales presentes en la prueba de comportamiento. El modelo estadístico empleado se puede representar como:

$$\text{Modelo 2} \quad y = Xb + Z_2^*a_m + Z_3^*p_m + e_2^g$$

donde :

y es un vector del peso vivo de cada animal.

b es un vector de efectos fijos con matriz de incidencia X , similar al del modelo 1, incluye una ecuación de regresión fija de orden $k_f = 5$ intra grupo contemporáneo.

a_m y p_m son conjuntos (de dimensión $n =$ numero de animales) de m coeficientes de regresión aleatorios de orden k_a y k_p respectivamente para los efectos genéticos aditivos y de ambiente permanente debido a las repeticiones de la variable dependiente en cada animal.

Z_2^* y Z_3^* corresponden a matrices de incidencia de dimensión $n \times k_i$ que como elementos contienen $Z_i^* = \Phi_i = \nabla_i(t_{ij}^*)$ donde ∇_i son los coeficientes del polinomio seleccionado y t_{ij}^* es la edad expresada en forma estandarizada entre -1 a $+1$ lo que es necesario para la utilización de los polinomios ortogonales de orden k_i .

e_2^g es un vector de efecto residual aleatorio con varianza heterogénea para $g = 9$ grupos de edades (entre 7 y 8 meses; 9 y 10 meses y 11 y 12 para los tres primeros grupos, posteriormente los intervalos de clase fueron de 1 mes de edad para los animales de 13 a 16 meses y finalmente el noveno grupo se conformo con todos los animales entre 17 y 18 meses de edad.)

El sistema de ecuaciones del modelo mixto (EMM) para obtener soluciones de b y a , según el modelo 1 será:

$$\begin{bmatrix} X'R_1^{-1}X & XR_1^{-1}Z_1 \\ Z_1'R_1^{-1}X & Z_1'R_1^{-1}Z_1 + G^{-1} \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} XR_1^{-1}y \\ Z_1R_1^{-1}y \end{bmatrix}$$

en la cual $R_1 = R_0 \otimes I_n$ y $G^{-1} = G_0^{-1} \otimes A^{-1}$ donde R_0 y G_0 son matrices de var-covar entre los residuos y de los componentes genéticos aditivos respectivamente entre los j caracteres analizados simultáneamente, A es una matriz del numerador de las relaciones de parentesco entre animales, I_n es una matriz identidad de orden n (numero de animales con datos) y \otimes es el símbolo de producto. Se considera que $y \sim N [0, (S_y^2 = Z_1(G_0^{-1} \otimes A^{-1})Z_1' + I_n S_{e1}^2)]$. Lo descrito se aplica a modelos MAM, para MAU las matrices R y G se refieren solamente a un solo rasgo.

Para el modelo 2 el sistema EMM será:

$$\begin{bmatrix} X'R_2^{-1}X & X'R_2^{-1}Z_2 & X'R_2^{-1}Z_3 \\ Z_2'R_2^{-1}X & Z_2'R_2^{-1}Z_2 + A^{-1} \otimes K_a^{-1} & Z_2'R_2^{-1}Z_3 \\ Z_3'R_2^{-1}X & Z_3'R_2^{-1}Z_2 & Z_3'R_2^{-1}Z_3 + I_n \otimes K_p^{-1} \end{bmatrix} \begin{bmatrix} b \\ a \\ p \end{bmatrix} = \begin{bmatrix} X'R_2^{-1}y \\ Z_2'R_2^{-1}y \\ Z_3'R_2^{-1}y \end{bmatrix}$$

Nótese la similitud entre los sistemas EMM del modelo 1 y 2 en el cual la var-covar genética es $G = A^{-1} \otimes K_a^{-1}$. Los términos K_a y K_p representan las matrices de var-covar de los coeficientes de regresión aleatorios para los efectos genéticos aditivos y de ambiente permanente respectivamente. En el modelo MRA se incorporó un efecto ambiental permanente debido a repeticiones de la variable dependiente en el mismo animal ($Z_3'R_2^{-1}Z_3 + I_n \otimes K_p^{-1}$). Un total de de 12 modelos del tipo MRA fueron comparados mediante la prueba denominada likelihood ratio test (logL), las diferencias se debian al orden de ajuste de los polinomios k_a y k_p ; conformándose todas las combinaciones posibles entre una ecuación lineal, orden de ajuste 1, hasta de orden 5.

Los parámetros genéticos, heredabilidad (h^2) y correlaciones genéticas (R_g), se estimaron mediante funciones lineales de los componentes correspondientes para los $j=1,2..t$ rasgos evaluados según los modelos MAU y MAM. Para el caso del modelo MRA se necesitan tomar en cuenta otras consideraciones. De acuerdo a Jamrozik and Schaeffer (1997), la varianza genética para el PV a la j edad se estima por:

$$S_{g_j}^2 = \Phi_j' K_a^{-1} \Phi_j$$

mientras que la varianza para el ambiente permanente será:

$$s_{p_j}^2 = \Phi_j' K_p^{-1} \Phi_j$$

y la covarianza genética entre las edades j y k se estimará por:

$$s_{jk} = \Phi_j' K_a^{-1} \Phi_k$$

A partir de tales componentes la h^2 para el PV a j = 12 meses de edad, (esta edad pertenece al grupo 7 de la varianza residual heterogenea) se estimará por:

$$h_{12}^2 = \frac{\Phi_{12}' K_a^{-1} \Phi_{12}}{[\Phi_{12}' K_a^{-1} \Phi_{12} + \Phi_{12}' K_p^{-1} \Phi_{12} + e_2^{g=7}]}$$

Por otro lado la correlación genética entre PV₁₂ y PV₁₅ ($R_{g_{12-15}}$) se estimará por:

$$R_{g_{12-15}} = \frac{\Phi_{12}' K_a^{-1} \Phi_{15}}{\{(\Phi_{12}' K_a^{-1} \Phi_{12}) * (\Phi_{15}' K_a^{-1} \Phi_{15})\}^{0.5}}$$

En este procedimiento de MRA se obtienen soluciones para cada animal para los m coeficientes de regresión aleatorios con los cuales se pueden estimar los Valores Genéticos (VGⁱ) para cualquier punto de la trayectoria entre los 7 y 18 meses de edad. A modo de ejemplo, para el i animal el VG₁₂ será:

$$VG_{12}^i = \sum_{m=0}^{k-1} \Phi_{12} a_i$$

donde a_i representa la solución para el animal i que viene dado por un polinomio de orden k_m . Cuando se aplica MRA aparentemente se incorpora mayor complejidad matemática a las estimaciones, no obstante son múltiples las ventajas que pueden obtenerse con los resultados. Manipulando las soluciones de a_i y los coeficientes de Φ_j se pueden estimar varios indicadores de mucha importancia, así por ejemplo la varianza genética para ganancia total de peso entre 12 y 15 meses se estimará como:

$$\Phi_{15-12} = (\Phi_{15} - \Phi_{12}) \text{ y } s_{g_{15-12}}^2 = \Phi_{15-12}' (A^{-1} \otimes K_a^{-1}) \Phi_{15-12}$$

El ASREML brinda los valores de Φ_j para todos los puntos de la trayectoria entre las edades extremas de los datos disponibles, de manera que con este procedimiento previo se pueden estimar el resto de los parámetros entre cualquier edad representado en este estudio.

Resultados y discusión.

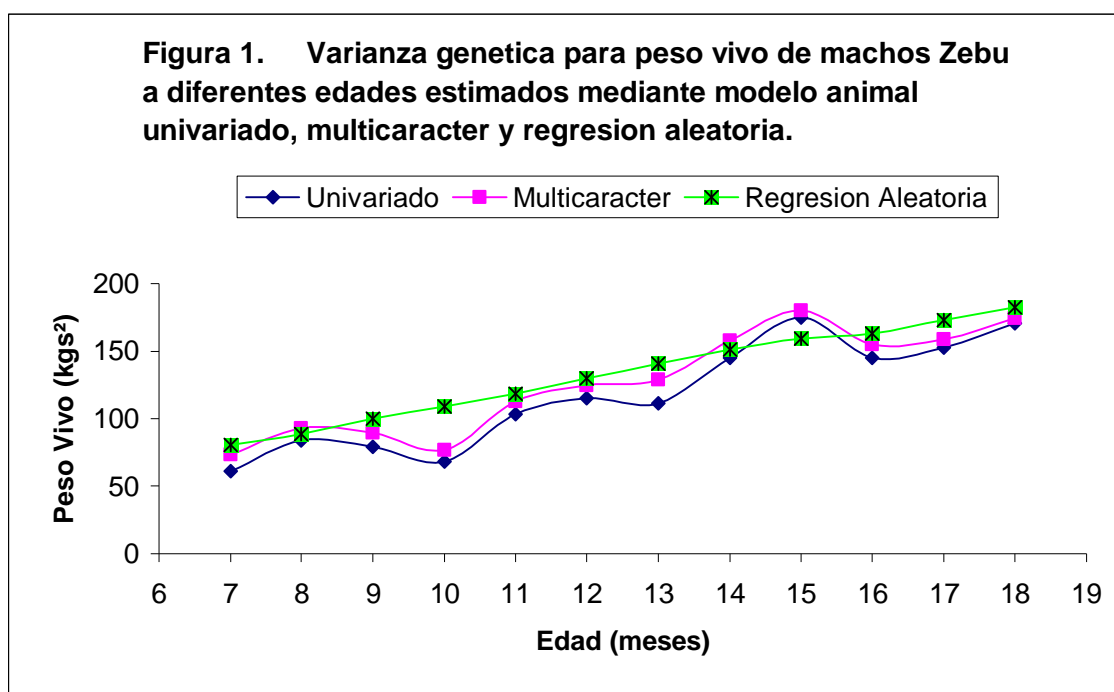
La muestra de datos analizados manifestaron una ganancia diaria de 0.552 kgs/día alcanzando 346 kgs de peso vivo a los 18 meses de edad (tabla 1), con importantes variaciones en los 12 meses de duración de la prueba de comportamiento, los coeficientes de variación aumentan en la medida que se incrementa la edad. Mas elementos descriptivos

y factores que afectan el crecimiento en esta etapa fueron presentados en un trabajo previo (Menéndez Buxadera y col., 2004).

Tabla 1. Algunos indicadores generales del crecimiento de machos Cebú en la prueba de comportamiento en condiciones de pastoreo de Cuba.

Edad (meses)	Peso Vivo (Kg.)	Coefficiente de Variación %
7	162.2	13.1
8	175.6	16.6
9	189.4	15.4
10	203.6	17.1
11	220.2	18.6
12	239.5	19.6
13	259.2	20.3
14	275.4	20.2
15	292.2	20.1
16	308.9	20.0
17	327.1	19.4
18	346.1	18.9

La evolución de las varianzas genéticas estimadas por los tres procedimientos se presenta en la figura 1. Los resultados de MAU y MAM son prácticamente semejantes y muestran un patrón irregular, con varios picos y caídas difíciles de explicar, no obstante la tendencia general muestra un positivo incremento en la medida que avanza la edad.



Esta respuesta oscilante es muy semejante a la presentada por Albuquerque y Meyer (2001) para los resultados de un modelo univariado aplicado a datos de animales Nelore (*B. indicus*) de Brasil, lo cual es una característica de estos procedimientos MAU y MAM para datos de este tipo, en los que se analizan transversalmente unos pocos puntos de esta etapa del crecimiento. En términos biológicos se debe esperar un mayor grado de coherencia en

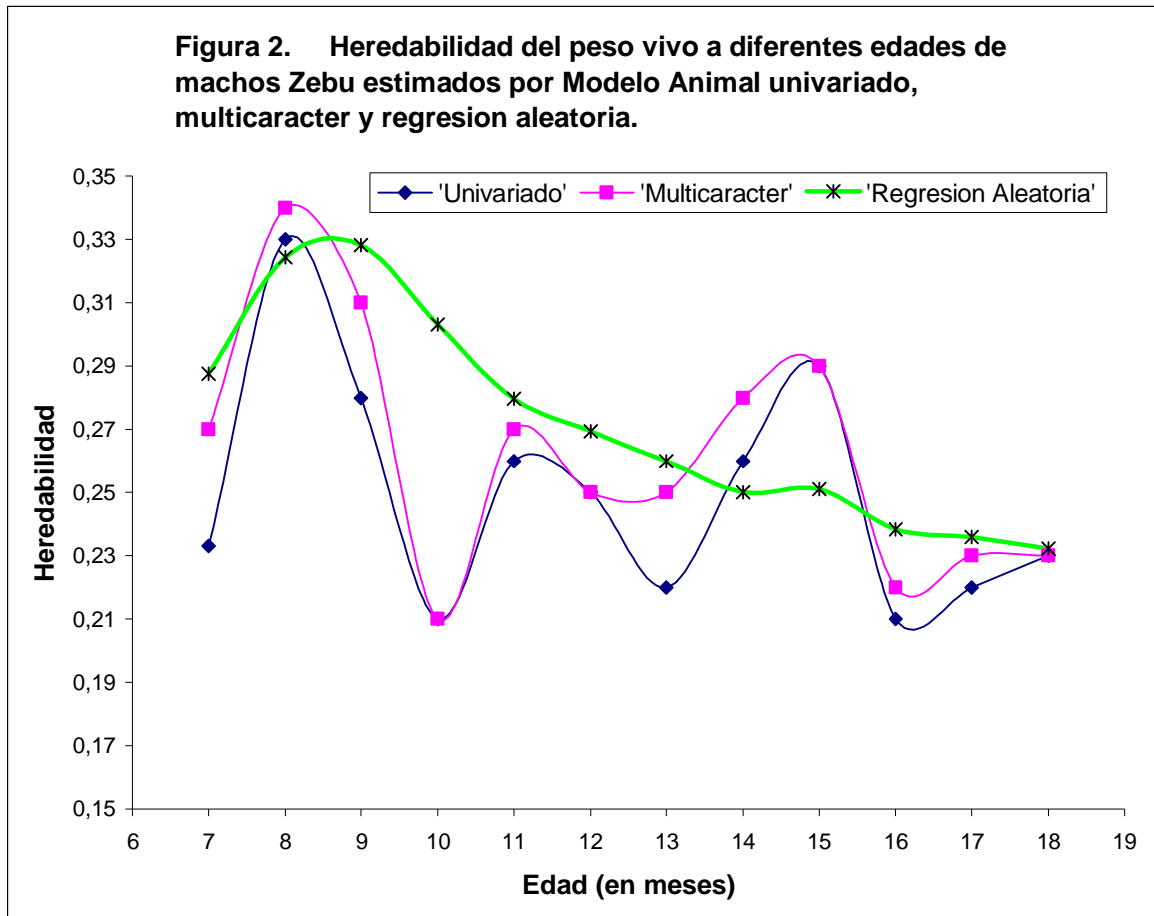
las varianzas estimadas en la trayectoria de edad representada en este estudio, lo cual es diferente a los resultados de la figura 1. Estos elementos fueron utilizados por Kirpatrick y col (1990) para argumentar y proponer el uso de modelos de dimensiones infinitas, mediante funciones de covarianza (CF), en los que la evolución de las varianzas genéticas se pueden estimar en función de la edad.

El mejor modelo MRA fue de un polinomio de orden $K_a = 5$ (ecuación cuartita) para los efectos genéticos aditivos y de orden $K_p = 6$ (ecuación quintica) para los efectos individuales de ambiente permanente. Aparentemente este orden de ajuste polinomial es el que mas se adapta a esta etapa del crecimiento en animales vacunos, no solo del tipo *B. indicus* en condiciones de pastoreo (Albuquerque y Meyer, 2001; Nobre y col., 2003), sino también en diferentes razas y cruces *B. taurus* en sistemas de alimentación con concentrado (Schenkel y col., 2002) lo cual puede servir de guía inicial a otros estudios similares. Un análisis de componentes principales realizado sobre la matriz de coeficientes de regresión aleatorios (K_a) indico que el primer vector propio explico el 71.8% de la variación, mientras que el segundo represento el 24.1%, aproximadamente igual tendencia fue indicada por Nobre y col. (2003) quienes alcanzaron el 84% de la variación en los dos primeros valores propios. Estos resultados sugieren que el proceso de selección puede ejercer una influencia nada despreciable sobre la curva de crecimiento de los animales cebuinos, lo cual es algo diferente al patrón de animales de razas europeas presentado por Schenkel y col. (2002) quienes indicaron un 97% de la variación en el primer vector propio, es decir el intercepto de la ecuación. Nuevas investigaciones deben realizarse en este aspecto.

Los correspondientes resultados de la varianza genética estimada por MRA (figura 1) presentaron un patrón muy regular y prácticamente lineal entre los 7 y los 18 meses de edad en esta muestra de animales Cebú de Cuba, lo cual esta en correspondencia con el tipo de respuesta armoniosa que debe esperarse para un modelo estadístico MRA que posee propiedades similares al procedimiento propuesto de Kirpatrick y col. (1990), aplicado a datos que varían en función del tiempo (Meyer and Hill, 1997). Los resultados y tendencias manifestados en este trabajo, están en total correspondencia con los resultados de Albuquerque y Meyer (2001) y Nobre y col. (2003) para animales cebuinos en las condiciones de Brasil. En términos generales puede indicarse que existe bastante similitud en los resultados de MAU, MAM y MRA.

Los valores de h_i^2 estimados por los tres modelos y para cada punto de la trayectoria de edad se muestran e la figura 2. En forma semejante a los resultados de las varianzas estimadas, los valores de h_i^2 para MAU y MAM, son similares, lo cual es esperado para ambos procedimientos. Para el modelo MRA se manifestó una respuesta bastante homogénea, sin esas altas y bajas de los modelos precedentes, con un incremento inicial entre 7 y 9 meses de edad y una ligera tendencia decreciente en el resto del periodo de tiempo. Los tres procedimientos muestran valores muy semejantes para la h^2 en el punto final de la trayectoria (PV_{18}).

Figura 2. Heredabilidad del peso vivo a diferentes edades de machos Zebu estimados por Modelo Animal univariado, multicaracter y regresión aleatoria.



Las características de la base de datos analizadas, en la cual los animales están representados en todos los puntos de la trayectoria, no facilita mostrar una de las ventajas de MRA no obstante, en ambas figuras se puede observar que la varianza estimada en este procedimiento fueron ligeramente superior respecto a MAU y MAM. Los parámetros estimados mediante MAU y MAM simplemente representan puntos que pueden variar a través de todo el periodo de tiempo estudiado y no toman en cuenta las diferencias de tiempo entre los PV estimados a dos edades muy distantes, por el contrario MRA elimina ese riesgo y brinda una función que representa la evolución de los componentes de (co)varianza en toda la trayectoria de edad. Los preajustes a edad fija o el uso de una covariable fija, requeridos en MAU y MAM representan una fuente de sesgo potencial ya que asumen implícitamente, una crecimiento lineal entre las dos edades que se desean ajustar. En MRA no hay que asumir nada a priori sobre la forma de crecimiento y los estimados se realizan en un solo procesamiento. Otro detalle de los resultados de este trabajo es que no se manifestó la incongruencias propias de los sistemas de regresión aleatorios en las edades extremas (ver Albuquerque y Meyer, 2001; Nobre y col. 2003), debido a las características antes mencionadas del conjunto de datos utilizados.

Los valores de h^2 de este trabajo, son ligeramente inferiores a las recopilaciones de Davies (1993); Mercadante y col, (1995) y Lobo y col. (2000), quienes presentaron las publicaciones disponibles sobre el crecimiento del vacuno en condiciones tropicales. Por el contrario, los resultados de Rico y col. (1995) con un conjunto de datos diferente de

animales de esta raza en Cuba, mostraron valores algo superiores para PV_{18} ($h^2 = 0.31$). Todas estas inconsistencias pueden estar relacionadas al procedimiento de cálculo (mínimos cuadrados) empleado en esa época. No obstante, las referencias más recientes con animales Brahman, en los que se aplicaron procedimientos estadísticos más robustos presentan valores inferiores a los estimados en este trabajo (Plasse y col., 2002; Pico y col., 2004). La similitud de resultados encontrados por los tres procedimientos aplicados en este trabajo sugieren que estos son los parámetros más probables que deben utilizarse en el programa de selección de esta raza en Cuba.

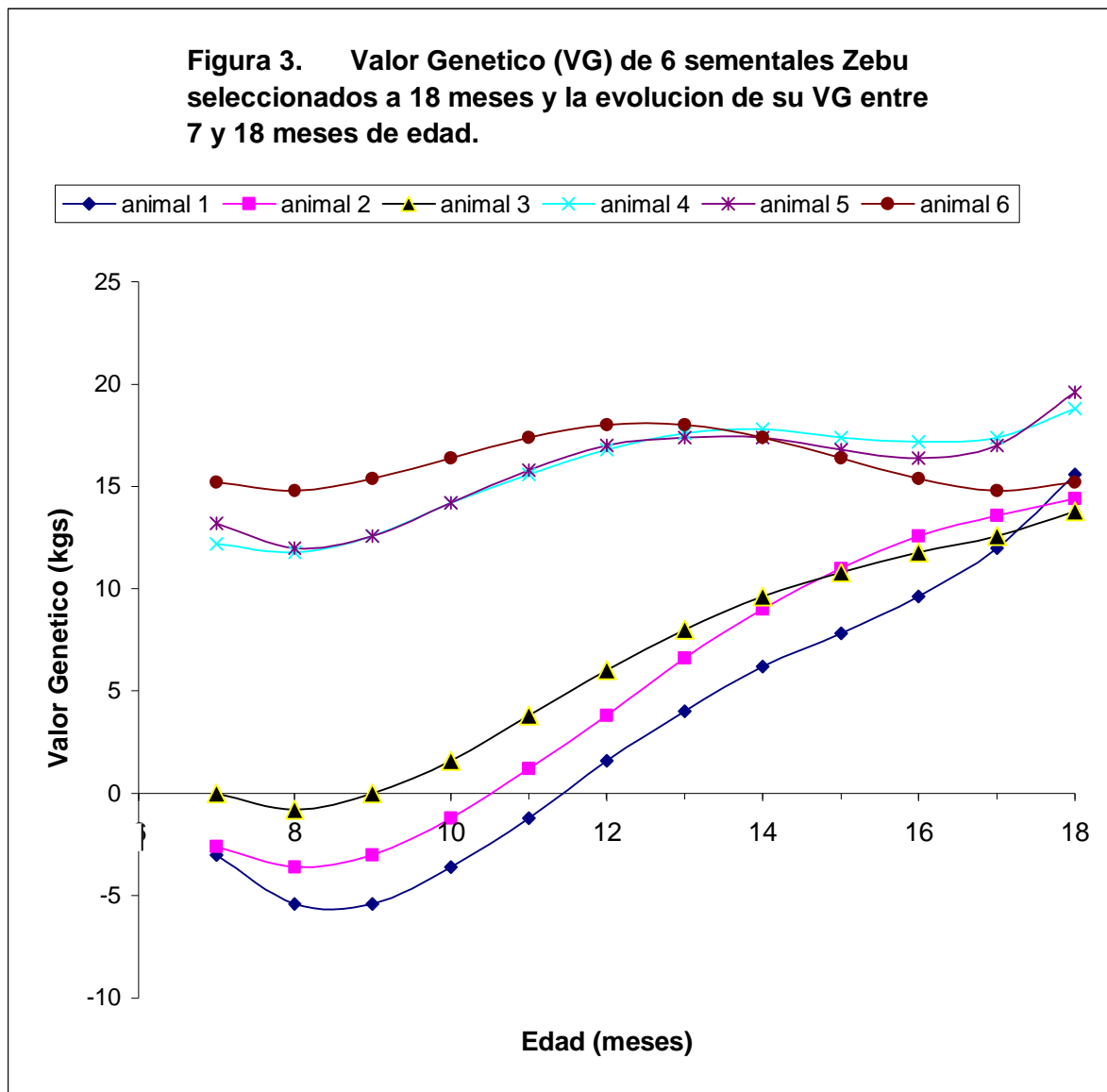
Los estimados de correlaciones genéticas entre los PV_i según MAM y MRA se presentan en la tabla 2. En términos generales, el patrón es muy semejante en ambos procedimientos aunque con mayor coherencia para MRA.

Tabla 2. Matriz de correlaciones genéticas entre peso vivo a diferentes edades en machos Cebú en prueba de comportamiento en pastoreo, estimado por modelo animal multicausal (debajo de la diagonal) y regresiones aleatorias (encima de la diagonal).

	Edad en meses											
	7	8	9	10	11	12	13	14	15	16	17	18
7		0.963	0.946	0.944	0.928	0.887	0.826	0.752	0.676	0.611	0.570	0.540
8	0.866		0.996	0.981	0.945	0.889	0.823	0.753	0.682	0.614	0.547	0.452
9	0.936	0.983		0.992	0.962	0.913	0.854	0.790	0.725	0.660	0.593	0.488
10	0.909	0.970	1.001		0.989	0.956	0.910	0.855	0.796	0.738	0.681	0.588
11	0.925	0.902	0.967	0.992		0.989	0.960	0.918	0.869	0.820	0.776	0.670
12	0.929	0.907	1.003	0.985	1.002		0.990	0.964	0.928	0.890	0.857	0.785
13	0.761	0.792	0.883	0.951	0.984	1.000		0.991	0.969	0.943	0.917	0.842
14	0.762	0.836	0.850	0.933	0.939	0.940	1.001		0.993	0.978	0.956	0.872
15	0.614	0.692	0.727	0.911	0.905	0.880	0.992	0.981		0.995	0.977	0.885
16	0.538	0.626	0.635	0.738	0.816	0.800	0.946	0.911	0.971		0.990	0.904
17	0.645	0.593	0.609	0.713	0.880	0.888	0.950	0.904	0.945	0.968		0.953
18	0.558	0.463	0.555	0.619	0.830	0.804	0.900	0.855	0.877	0.906	0.992	

Independientemente del método, todas las correlaciones son positivas, indicando que cualquier trabajo de selección individual producirá aumentos en el peso vivo en todo el trayecto entre los 7 y 18 meses de edad. Los valores de R_g son mayores de 0.9 entre edades adyacentes (unos pocos valores fuera de espacio para MAM) y tienden a decrecer en la medida que aumenta la diferencia de edad entre las estimaciones, en forma similar a lo presentado por Rico y col. (1994). Si se considera que PV_{18} constituye el objetivo de selección entonces los indicadores entre 7 y 12 meses de edad tienen un menor grado de relación ($R_g < 0.7$), sin embargo los estimados entre 13 meses y más presentan una $R_g > 0.8$ con el punto final de la prueba de comportamiento (PV_{18}), sugiriendo que es posible una selección temprana de los animales lo cual aumenta la capacidad de las instalaciones, disminuyendo costos y contribuyendo a un mayor progreso genético en esta raza. Esta sugerencia debe estudiarse más detenidamente, no obstante este punto de vista está en correspondencia con las magnitudes de los valores propios estimados a partir de la matriz K_a que se indicó anteriormente, en particular es posible identificar animales con menor, igual o superior crecimiento en los primeros 6 meses postdestete y mayor velocidad de

crecimiento posteriormente. Es esta precisamente una de las ventajas adicionales que pueden derivarse del uso de MRA si se manejan los coeficientes del primer y segundo vector propio de la matriz K_a . Para ejemplificar esta posibilidad se seleccionaron los mejores 45 animales (alrededor del 2%) según su $VG_{18} = + 16.5$ kg lo cual equivale a +2.6 desviaciones estándar genética con respecto a la media. Los resultados individuales de este grupo de alto valor genético para el peso final de la prueba de comportamiento, fueron analizados en forma pormenorizada y se identificaron dos conjuntos de 3 sementales cada uno cuya evolución en VG entre 7 y 18 meses de edad se muestran en la figura 3.



Nótese que no existen diferencias importantes entre ambos tipos de animales respecto a su VG_{18} , lo cual resulta obvio ya que pertenecen al mismo grupo de selección. Sin embargo, en términos genéticos se manifiestan importantes diferencias en la formas de la curva de crecimiento de ambos conjuntos, con características extremas que pueden representar importantes beneficios para el programa de selección de esta raza. Así por ejemplo los animales 4; 5 y 6 manifestaron un patrón ‘estable’ en su VG en toda la trayectoria entre los 7 y 18 meses de edad, mientras que por el contrario los otros 3 sementales pueden

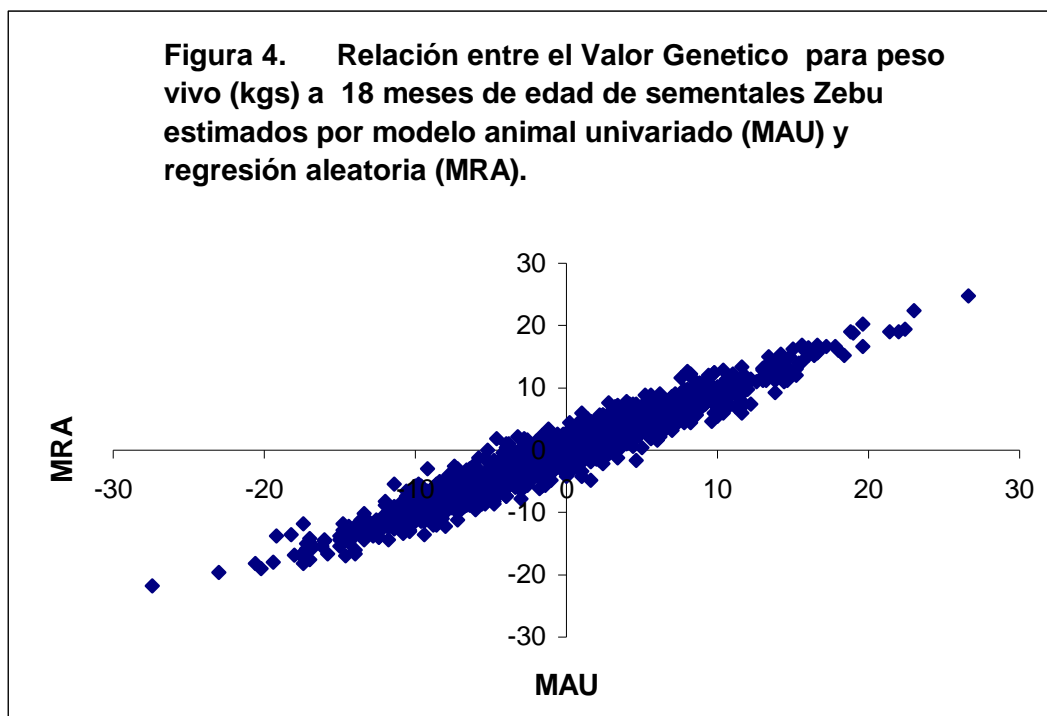
calificarse como ‘dinámicos’ con una respuesta totalmente diferente. La tabla 3 presenta los coeficientes de regresión genética promedios de ambos subconjuntos de sementales, los cuales demuestran cabalmente las diferencias existentes cuyo origen es genético, de manera que la selección final de uno u otro grupo de sementales (todos elites) pueden contribuir a cambiar la forma de la curva de crecimiento de sus progenies. Otros estudios son necesarios para discernir cual de los dos formas de crecimiento debe ser favorecida por la selección, no obstante estos efectos están presentes en la población de animales Cebú como respuesta colateral del programa de mejora en base a PV_{18} , aunque su manifestación ha estado solapada y aun es necesario cuantificar.

Tabla 3. Coeficientes de regresión genética promedio (a_i)* para los dos conjuntos de sementales Cebú elites seleccionados por su Valor Genético a 18 meses de edad (en Kg.) y representados en la figura 3.

Tipo de semental	a_0	a_1	a_2	a_3	a_4	VG_{18}
Estable	15.94	2.22	-1.53	0.023	1.33	17.90
Dinámico	4.87	9.95	0.60	-1.21	0.93	14.50

* Estos coeficientes representan las soluciones del modelo MRA y el VG para cualquier punto de la trayectoria entre los 7 y 18 meses se estima multiplicando tales coeficientes por los coeficientes del polinomio de orden 5 empleado según el modelo 2.

La importancia de esta información adicional que brinda el uso de las regresiones aleatorias solo pueden ser utilizadas en función de la relación entre los VG de los mismos sementales estimados según MAU y MRA. La figura 4 muestra estos resultados que indican una estrecha relación entre ambos procedimientos, de manera que no existe margen para cambios de orden de merito. El análisis de regresión llevado a cabo para predecir el VG según MAU de acuerdo a los resultados del VG en MRA mostró una ecuación 0.9276 ± 0.005 (coeficiente de determinación de 93.1%) sobre el total de los animales evaluados.



El conjunto de datos estudiados presentó una alta variabilidad entre los VG de los sementales, cualquiera sea el procedimiento empleado en las estimaciones. En términos prácticos se manifestó una diferencia del orden de 42 Kg. para PV_{18} entre el grupo de los peores y mejores sementales lo cual representa el 12.1% de la media general mostrada en la tabla 1. Estas diferencias entre VG sugiere que existe amplio espacio para alcanzar importantes niveles de progreso genético en la raza Cebú de Cuba.

Conclusiones.

El estudio realizado demuestra que existe una amplia variabilidad en el crecimiento de la raza Cebú de Cuba en condiciones de pastoreo. Cualquiera sea el modelo estadístico aplicado se pudo constatar que alrededor del 23 % de esa variabilidad es de origen genético aditivo, lo cual implica que existe un amplio margen para el trabajo de selección y mejora. El análisis llevado a cabo ha mostrado las grandes ventajas del modelo de regresiones aleatorias, respecto a los modelos animal univariado o de multicañerotes, para estudiar la evolución de las varianzas y covarianzas genéticas en toda la trayectoria entre los 7 y 18 meses de edad periodo en el cual se lleva a cabo la prueba de comportamiento en esta raza.

Además de la estimación del VG para todos los animales en todas las edades representadas en este conjunto de datos, el modelo de regresión aleatoria posibilitó cuantificar la existencia de importantes diferencias entre sementales en cuanto a la forma de la curva de crecimiento, la cual puede brindar beneficios económicos importantes. La matriz de correlaciones genéticas entre el peso vivo estimado a diferentes edades sugiere que es posible reducir el período de la prueba de comportamiento sin comprometer el posible progreso genético en PV_{18} . Esta variante requiere de estudios adicionales, los cuales no solo permitirán incrementar las capacidades de las instalaciones, sino también evitan riesgos de accidentes en las etapas superiores de este tipo de animales cuyo temperamento es algo particular. Se recomienda aplicar modelos de regresión aleatoria para el análisis de la prueba de comportamiento.

Referencias.

- Albuquerque, L.A. y Meyer, K., 2001. Estimates of covariance functions for growth from birth to 630 days of age in Nelore cattle. *J. Anim. Sci.* 79: 2776-2789.
- Davies, G.P., 1993. Genetic parameters for tropical beef cattle in Northern Australia: a review. *Aust. J. Agric. Res.* 44: 179-198.
- Guerra, D.; Rodríguez, M.; Ramos, F.; Planas, T.; Ortiz, J.; Torres, J.A. y Falcón, R., 2002. Evaluación genética de las razas de carne en Cuba. II Congreso Internacional Mejora Animal. Palacio de la Convenciones 18-22 Noviembre, artículo No. 561, 4 pag.
- Gilmour, A.R.; Cullis, B.R.; Welham, S.J.; y Thompson, R., 2000. ASREML Reference Manual. NSW Agric. Biom. Bull. NSW Agriculture, Locked Bag, Orange, NSW 2800, Australia.
- Jamrozik, J. y Schaeffer, L.R., 1997. Estimates of genetic parameters for a test day model with random regression for production of first lactation. *J. Dairy Sci.* 80: 762-770.

- Kirkpatrick, M.; Lofsvold, D. y Bulmer, M., 1990. Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* 124: 979-993.
- Lobo, R.B.; de los Reyes, A.; Ferraz, J.B.S.; Bezerra, L.A.F.; Mercadante, M.E.Z. y Duarte, F.A.M., 1994. Bivariate animal model analysis of growth weights and scrotal circumference of Nellore cattle in Brasil. 5th World Cong. Genetics Applied to Liv. Prod., Vol 17: 199-201. Guelph, Canada.
- Lobo, R.N.B.; Madalena, F.E. y Viera, A.R., 2000. Average estimates of genetic parameters for beef and dairy cattle in tropical regions. *Anim Breed. Abst.* 68:433-462.
- Menéndez Buxadera, A.; Guerra, D.; Planas, T. y Ramos, F., 2004. Factores que afectan el crecimiento de machos jóvenes de la raza Cebú en prueba de comportamiento en condiciones de pastoreo de Cuba. Enviado a publicacion *Rev. Cubana de C. Agric.*
- Mercadante, Maria E.; Lobo, R.B. y de los Reyes, A., 1995. Parámetros genéticos para características de crecimiento en Cebuinos de carne. *Archivos Latinoam. Prod. Anim.* 3:45-89.
- Meyer, K. and Hill, W. G., 1997. Estimation of genetic and phenotypic covariance for longitudinal or repeated records by restricted maximum likelihood. *Liv. Prod. Sci.* 47: 185-200.
- Meyer, K., 2004. Scope for a random regression model in genetic evaluation of beef cattle growth. *Liv. Prod. Sci.* 86: 69-83.
- Nobre, P.R.C.; Misztal, I.; Tsuruta, S.; Bertrand, J.K.; da Silva, L.O.C. y Lopes, P.S., 2003. Analysis of growth curves of Nellore cattle by multiple-trait and random regression models. *J. Anim Sci.* 81:918-926.
- Pico, B.A.; Naser, F.W.C. y van Wyck, J.B; 2004. Genetic parameters for growth traits in South African Brahman cattle. *South African J. Anim. Sci.* 34:44-46.
- Planas, T. y Ramos, F., 1990. Cebú Cubano. Una incursión en la promoción de sementales de Cebú cubano. *Rev. ACPA* 12:11-18.
- Plasse, D.; Verde, O.; Fossi, H; Romero, R; Hoogesteijn, R; Bastidas P. y Bastardo, J., 2002. (Co)variance components, genetic parameters and annual trends for calf weight in a pedigree Brahman herd under selection for three decades. *J. Anim; Breed. Genet.* 119: 141-153.
- Rico, C.; Planas T.; Menchaca, I. y Garcia, F., 1994. Parámetros genéticos del crecimiento a diferentes edades en ganado Cebu. *Archivos Latinoamer. Prod. Anim.* 2:1-8.
- Schaeffer, L.R., 2004. Application of random regression models in animal breeding. *Liv. Prod. Sci.* 86: 35-45.
- Schenkel, F.S.; Miller, S.P.; Jamrozik, J. y Wilton, J.W., 2002. Two-step and random regression analysis of weight gain of station tested beef bulls. *J. Anim. Sci.* 80: 1497-1507.
- van der Werf, J. y Schaeffer, L.R., 1997. Random Regression in Animal Breeding. Course Notes . CGIL June 25-28 ; 58 pag.

ANEXO 4

Evolución de las Varianzas Genéticas y el Valor Genético de Sementales Cebu para la clasificación del tipo a diferentes edades de sus hijas mediante el uso de Regresiones Aleatorias.

A. Menéndez Buxadera*¹; L. Torres*; M. Larios*; R. Peynado*, E. Brito* y L. Fernández Chuairey**.

* Centro de Control Pecuario, Ministerio de la Agricultura, Conill y Boyeros, Ciudad de la Habana, Cuba

** Universidad Agraria de la Habana “Fructuoso Rodríguez Pérez”, San José de las Lajas, La Habana, Cuba

Resumen

Los resultados de la clasificación del tipo de 23405 hembras Cebú llevadas a cabo entre 1982 y 1999 fueron utilizados para estudiar la evolución de las varianzas genéticas y la heredabilidad de 6 rasgos del tipo en toda la trayectoria entre 2 y 12 años de edad. En los datos estaban representados 928 sementales de esta raza y se estudiaron mediante un modelo semental de regresión aleatoria (RRM). Los mejores resultados se obtuvieron para un RRM con una covariable de orden 2 para los efectos aleatorios y de orden 5 para los efectos fijos de grupos contemporáneos y de orden 3 para los efectos del intervalo parto clasificación. El uso de RRM para analizar este tipo de rasgos permitió describir el patrón de herencia de estos caracteres en función de la edad, manifestándose los resultados superiores antes de los 48 meses de edad y después una fase estable. Los valores de heredabilidad (h_{ij}^2) para todos los rasgos, fueron en promedio de $h_{ij}^2 = 0.30$ (0.22 a 0.45) cuando se evalúan antes de los 48 meses de edad, mientras que disminuye y se estabilizan a $h_{ij}^2 = 0.15$ a 0.22 entre los 4 y 12 años de edad. Las correlaciones genéticas fueron superiores a 0.8 entre edades adyacentes y decrecen hasta 0.67 entre edades tempranas y estados adultos. Los coeficientes de regresión entre el Valor Genético (VG) de los sementales estimados a 36 meses de edad presentan una alta capacidad predictiva del VG a edades adultas, de manera que es recomendable su uso en condiciones prácticas. El trabajo demuestra las posibilidades de este procedimiento de evaluación del VG de los sementales mediante la aplicación de un procedimiento de regresiones aleatorias.

1) email: cencop@minag.gov.cu

Palabras claves: Clasificación del Tipo; ganado Cebú; regresiones aleatorias; parámetros genéticos.

Enviado a publicación Rev. Cubana de Ciencia Agrícola 2005.

Introducción.

La clasificación del tipo en el ganado vacuno ha sido un procedimiento tradicionalmente empleado, en el cual se lleva a cabo una comparación relativa de ciertas características morfológicas de cada individuo, respecto a un animal '*ideal*' que reúne todos los requisitos de una raza determinada. El concepto de '*ideal*' está establecido subjetivamente en los técnicos que realizan este trabajo, sobre la base de la vasta experiencia de cada uno. En este procedimiento se evalúan varias características generales que se denominan '*detalles de parte*', otorgándole una clasificación a cada uno, que en términos numéricos pueden variar, entre 60 (muy malos) y 100 puntos (excelentes), posteriormente todos los resultados se combinan para brindar una clasificación final.

Este tipo de actividad muchas veces ha sido mal interpretada, en unos casos se consideraba a priori una relación positiva con caracteres productivos, mientras que en otras se asociaba a eficiencia funcional. En su bien conocido libro Brody (1945), presenta una amplia discusión sobre este tema, indicando que en todas las especies los individuos cambian de talla y por tanto de forma (la forma es la base de todo el trabajo de clasificación del tipo) en la medida que envejecen y que existe una estrecha relación entre forma y función. Este mismo autor señala que aunque ambos atributos son dos aspectos de la misma cosa, no se puede predecir con seguridad las funciones en base a la forma, lo cual se debe a limitaciones propias de las observaciones.

Para alcanzar un mayor grado de objetividad en los resultados todo el trabajo de clasificación del tipo ha evolucionado en el tiempo. Desde hace aproximadamente dos décadas se introdujo un sistema denominado '*clasificación lineal del tipo*', en el cual a cada '*detalles de parte*' se le han incorporado diversos '*rasgos descriptivos*' a cada uno de los cuales se le brindan valores que contrastan características biológicamente extremas (desde los muy malos o indeseables hasta los excelentes o deseables), aunque en algunos casos los valores intermedios son los más indicados. De esta forma se le brinda una cuantificación más precisa a cada característica, con vistas a ser consecuente con la base de sustentación subyacente mencionada previamente. Este sistema ha recibido una gran atención en ganado lechero (ver Mark, 2004) desde que fue propuesto por Thompson y col. (1983); en porcino se realizó un primer intento (Van Steenberg, 1989) aunque no se han encontrado otras referencias. Mas recientemente se han presentado los primeros resultados en ovino (Janssens y col., 2004) y en ganado de carne (Gutiérrez y Goyache, 2002; Forabosco y col 2004).

En Cuba, la clasificación del tipo en ganado de carne se lleva a cabo de una manera sistemática, particularmente en la raza Cebú que es la más numerosa. En tal sentido, el departamento técnico del Centro de Control Pecuario (CENCOP) ha introducido recientemente un sistema de clasificación lineal, el cual mantiene los '*detalles de parte*' los que se consideran reúnen globalmente las propiedades cuyas manifestaciones se registran mediante la evaluación de 27 rasgos descriptivos que se controlan en este nuevo procedimiento, cada uno de los cuales se cuantifican con valores entre 1 a 6, siendo los extremos de esta escala para aquellos individuos que presentan funciones biológicas contrapuestas, no se utilizan valores intermedios.

Como consecuencia de dificultades prácticas para la realización del trabajo, resulta obvio que existirán importantes variaciones en la edad de los animales cuando se realizan las

visitas periódicas de los técnicos oficiales designados por CENCOP para la clasificación del tipo de cada hembra. Este problema se puede resolver incorporando una covariable en el modelo de análisis del valor genético de cada animal, sin embargo esto puede sesgar los resultados ya que se asume una ecuación de regresión fija la cual puede sobrevalorar los animales mas jóvenes y lo contrario para los mas viejos, según se puede inferir de un trabajo previo realizado con la base de datos de esta raza (Menéndez Buxadera y col., 1998a). Resulta obvio que esto requiere ser analizado.

De acuerdo a lo expuesto se decidió realizar este análisis con el objetivo de estimar la evolución de las (co)varianzas genéticas entre los resultados de la clasificación del tipo de los *'detalles de parte'* en hembras Cebú de raza pura entre 2 y 12 años de edad como paso previo a una etapa posterior donde se incorporen los resultados recientes de la clasificación lineal cuya base de datos es menos numerosa pero con más variables controladas respecto a la utilizada en este estudio.

Material y Método

Para este estudio se procedió de la misma forma que fue descrita en una publicación previa (Menéndez Buxadera y col., 1998) incorporando solo unas pocas restricciones adicionales. La tabla 1 muestra algunos indicadores generales de cada uno de los 6 *'detalles de parte'* estudiados.

Tabla 1. Indicadores generales de la clasificación del tipo en hembras Cebú de Cuba.

Detalle de Parte	Símbolo	Amplitud
Apariencia General	AG	62 a 98
Capacidad Corporal	CCO	62 a 95
Capacidad Carnicera	CCA	62 a 95
Grupa	GRU	62 a 96
Patas	PAT	62 a 92
Clasificación Final	CF	60 a 95

Datos. Para este trabajo se recibió copia de los ficheros disponibles en CENCOP los cuales fueron sometidos a un pormenorizado proceso de validación. El conjunto de datos contiene los resultados individuales de 6 caracteres controlados, que describen los *'detalles de parte'* de 24720 hembras clasificadas entre 1982 y 1999 y que eran hijas de 1073 sementales Cebú. De este total se eliminaron aquellos datos de grupo contemporáneo (rebaño-año-época, GC) con menos de 10 observaciones. Se creó la variable intervalo entre el parto y la clasificación con clases de 1 mes de amplitud, excepto en la primera que fue de dos meses. Finalmente quedaron los datos de 23405 hembras hijas de 504 sementales y 15981 vacas, de las cuales 8382 estaban representadas en el vector de los datos. El pedigree por la vía de los sementales y de los padres de las madres fue extendido en la medida de lo posible, de acuerdo a la información contenida en la base de datos nacional de CENCOP, disponiéndose de un total de 928 sementales

Argumentación y procedimiento estadístico. Los datos más comúnmente analizados en genética animal son del tipo denominado como de *dimensiones finitas* en los que solo están representados una sección o punto en el desarrollo de un animal, por ejemplo, el peso vivo a una edad fija (o a varias). Este tipo de caracteres se pueden estudiar por modelos

univariados o en otros casos multivariados cuando el rasgo se cuantifica en unas pocas edades o en diversos estadios (ejemplo diferentes lactancias). En muchos casos las matrices de (co)varianza que resultan de estos análisis muestran patrones erráticos de difícil interpretación, ya que en términos biológicos se debe esperar que se manifieste una tendencia más o menos armónica en toda la trayectoria del periodo de tiempo entre las cuales se controla el carácter.

A partir de las propuestas presentadas por Kirkpatrick y col. (1990) se han publicado numerosos trabajos en los cuales se estiman la evolución de las (co)varianzas genéticas de diferentes caracteres a través del tiempo, es decir para rasgos de *dimensiones infinitas*, mediante lo que se denomina funciones de covarianza (FC). En términos prácticos fue Schaeffer y Dekkers (1994) quienes presentaron una primera versión coherente del uso de FC al caso de los registros mensuales de producción de leche, internacionalmente conocido como Test Day Model (TDM), mediante un sistema de regresiones aleatorias (RA), previamente planteadas por Henderson Jr (1982). Inicialmente el uso de RA mediante FC resultaba muy engorroso, sin embargo, quedó demostrado que existía equivalencia entre estos procedimientos y los algoritmos del tipo REML (Meyer and Hill, 1997) lo cual facilita la utilización de estos métodos a una enorme cantidad de aplicaciones en diferentes escenarios de la mejora animal en diferentes especies (ver Schaeffer, 2004).

En Cuba, muy recientemente se han iniciado estudios de este tipo sobre TDM en ganado de leche (Fernández Chuairy, 2004) y en el análisis del crecimiento de machos Cebú en prueba de comportamiento (Menéndez Buxadera y col., 2005). En ambas referencias estaba claro que existían registros repetidos del mismo rasgo en cada individuo durante cierto periodo de tiempo, sin embargo en el caso que nos ocupa no resulta tan evidente. Debido a múltiples limitaciones y cuestiones prácticas objetivas, se manifiesta una importante variación de edad en los resultados de la clasificación del tipo, por otro lado existe el criterio de que los caracteres morfológicos que conforman los diferentes rasgos del tipo, manifiestan una evolución en función del tiempo y de los diferentes estados fisiológicos (partos, lactancia etc) a través de la vida productiva del animal. Incluso los técnicos que realizan este trabajo, por demás con alta experiencia en esta actividad, han indicado que se deben realizar dos o tres evaluaciones a cada animal para la más adecuada evaluación morfológica de los mismos. Esta alternativa sería realmente óptima, pero muy costosa tanto en recursos como en tiempo, sin embargo la propia distribución de los datos puede posibilitar un primer examen de esta interrogante.

Los datos disponibles para este estudio son denominados de corte transversal longitudinal (Fitzhugh, 1976), en el que cada animal tiene registro en una sola edad, pero de conjunto se puede representar la curva de desarrollo de la población de animales Cebú. La información que se puede obtener de este tipo de datos puede ser muy útil en términos descriptivos, sin embargo, la variabilidad individual no puede cuantificarse adecuadamente. Considere que estos mismos datos conforman n grupos de progenies, de manera que exista una representación del genotipo de los n sementales a lo largo de la trayectoria de edad, si al mismo tiempo se toma en cuenta las relaciones de parentesco entre todos los animales así como sus antecesores, se podría conectar toda la información disponible. En tales condiciones los datos pueden analizarse mediante un modelo lineal mixto incorporando la matriz de parentesco, lo cual permite estimar la evolución de las (co)varianzas genéticas de cada '*detalles de parte*' en todo el periodo de edad representada en la base de datos. Esta fue precisamente la estrategia seguida en este análisis.

En términos estadísticos se aplicó un modelo semental univariado de regresiones aleatorias (RRM) según la siguiente representación:

$$y_{ij} = \sum_{m=0}^5 b_m f_m(t_{ij}^*) + \sum_{m=0}^3 a_m f_m(t_{ij}) + \sum_{m=0}^2 S_m f_m(t_{ij}^*) + \epsilon_{ij}$$

donde:

y_{ij} es la estimación de la clasificación del tipo de las progenies del i esimo semental realizada a la j edad.

b_m y a_m representan vectores de efectos fijos de GC (b) con 205 niveles y de intervalos entre parto y la clasificación (a) de 16 niveles.

S_m son conjuntos (de dimensión s = numero de sementales) de m coeficientes de regresión aleatorios de orden k_i para los efectos genéticos aditivos del Semental (S_m).

ϵ_{ij} es un vector de efecto residual aleatorio con varianza heterogénea para 5 grupos de edades (menor de 36; de 36 a 54 ; de 55 a 65; de 66 a 72 y más de 73 meses de edad).

Los elementos $f_m(t_{ij}^*) = \Phi_i$ corresponden con las matrices de incidencia de los modelos lineales clásicos (X y Z) y como elementos contienen los m coeficientes de regresión de orden k_i y t_{ij}^* es la j edad en la que se registro cada variable dependiente expresada en forma estandarizada entre -1 a + 1. Para los efectos fijos, los m coeficientes fueron los de un polinomio de orden 5 para b y de orden 3 para a . Para el efecto aleatorio se compararon varios modelos con $k = 1$ y 2. El símbolo \sum quiere decir que el efecto total de cada factor se estima como una sumatoria del polinomio asumido y los m coeficientes estimados, es decir, m_0 (intercepto) + m_1 (coeficiente lineal) + m_2 (coeficiente cuadrático) + etc.

El sistema de ecuaciones del modelo mixto para obtener soluciones de los efectos incluidos será:

$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1} \otimes K_m^{-1} \end{bmatrix} \begin{bmatrix} b \\ s \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

donde b representa los estimados de los efectos fijos, A es una matriz del numerador de las relaciones de parentesco entre animales, \otimes es el símbolo de producto y K_m es la matriz de los coeficientes de regresión aleatorios del efecto semental, en este caso la matriz de varianzas y covarianzas genéticas será $G_s = \frac{1}{4}(A^{-1} \otimes K_m^{-1})$. Expandiendo los elementos de K_m la estructura global de (co)var para $k=2$ será:

$$V = \begin{bmatrix} K_m = \begin{pmatrix} AS_{s_0}^2 & AS_{s_{10}} & AS_{s_{20}} \\ AS_{s_{01}} & AS_{s_1}^2 & AS_{s_{21}} \\ AS_{s_{02}} & AS_{s_{12}} & AS_{s_2}^2 \end{pmatrix} & 0 \\ 0 & \sum_{c=1}^5 r_c \end{bmatrix}$$

Los componentes de (co)varianza para cada variable dependiente fue estimada por dos modelos diferentes, en total 12 corridas. Las diferencias se debían al nivel de ajuste del polinomio utilizado para los efectos aleatorios de semental, en este caso se compararon para $k=1$ y $k=2$, los resultados de logL se utilizaron para seleccionar el mejor modelo. El termino $\sum_{c=1}^5 r_c$ representa los efectos de la varianza residual heterogénea.

Todo el estudio se realizó utilizando el software ASREML (Gilmour y col., 2000) el cual brinda no solo las soluciones para todos los efectos incluidos en el modelo y los coeficientes de los polinomios (Φ_i) sino también para todos los componentes de varianza descritos previamente. Sin embargo, para estimar los parámetros genéticos, se debe aplicar un procedimiento adicional siguiendo la propuesta de Jamrozik and Schaeffer (1997), según las cuales la varianza genética para efecto genético aditivo debido al semental (s_s^2) de cada *detalle de parte* a la j edad se estimará por:

$$s_{sj}^2 = 4 [\Phi_j' (A^{-1} \otimes K_m^{-1}) \Phi_j]$$

Mientras que las covarianzas genéticas entre las edades j y k (s_{sjk}) será:

$$s_{sjk} = 4 [\Phi_{sj}' (A^{-1} \otimes K_m^{-1}) \Phi_{sk}]$$

La varianza fenotípica total para la j edad que se desee, se estima con el correspondiente componente de varianza genética, debiéndose adicionar la varianza residual a que pertenece el grupo de edad en cuestión. Finalmente los parámetros genéticos (heredabilidad y correlaciones genéticas) se estiman por procedimientos clásicos.

En este procedimiento de RRM se obtienen soluciones para los efectos de semental (S_i) para los k_m coeficientes de regresión aleatorios con los cuales se pueden estimar los Valores Genéticos (VG^i) para cualquier punto de la trayectoria entre los 30 y 120 meses de edad. A modo de ejemplo, el valor genético para el semental i a los 36 meses de edad de sus hijas VG_{36}^i será:

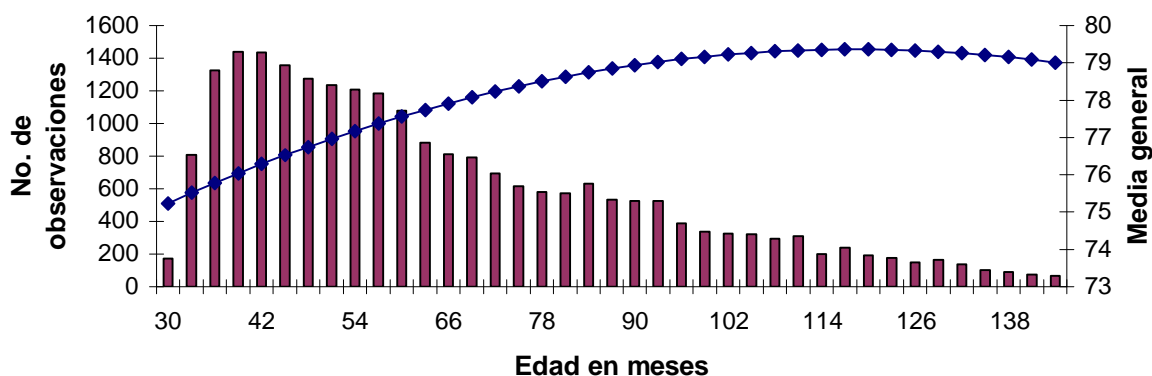
$$VG_{36}^i = \sum_{k=0}^{k-1} \Phi_{36} S_i$$

Resultados y discusión.

Un análisis de frecuencia simple de los datos utilizados en este estudio se muestra en la figura 1, donde también se incorporo la media de AG como representativa de los '*detalles de parte*' estudiados. Un análisis de regresión llevado a cabo, tanto con los datos individuales así como con las medias de cada clase de edad ajustadas para efectos de GC, arrojó que una ecuación cuadrática es la que mejor describe la evolución de estos rasgos en función de la edad, de manera que se pueden considerar como representativos de la

trayectoria del desarrollo de los rasgos estudiados entre 2 y 16 años de edad en la raza Cebú de Cuba.

Figura 1. Cantidad de observaciones y media de la Apariencia General en hembras Cebú



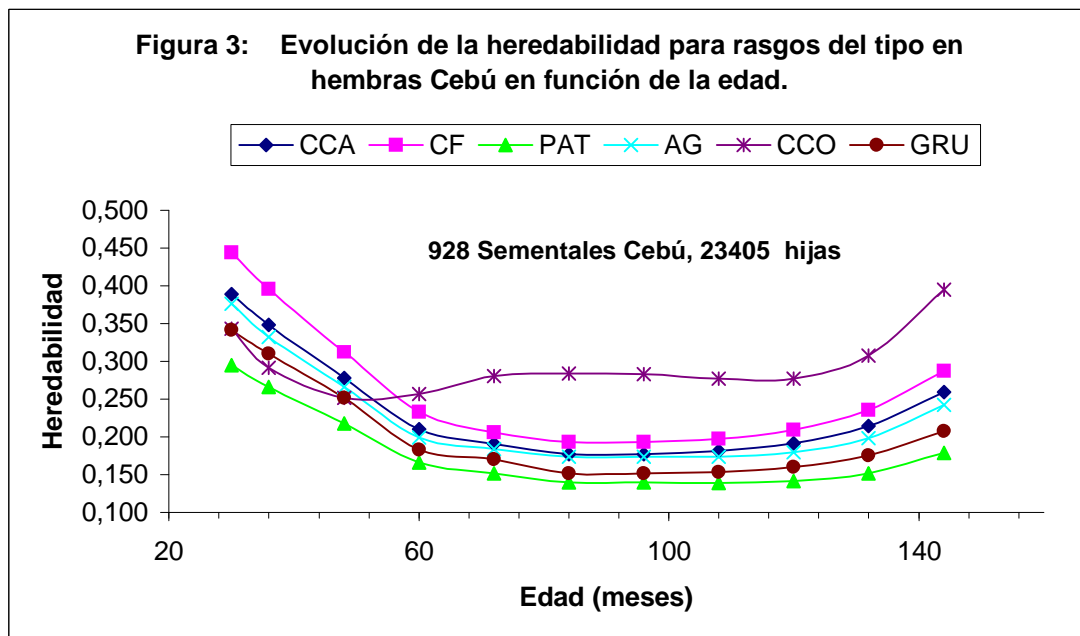
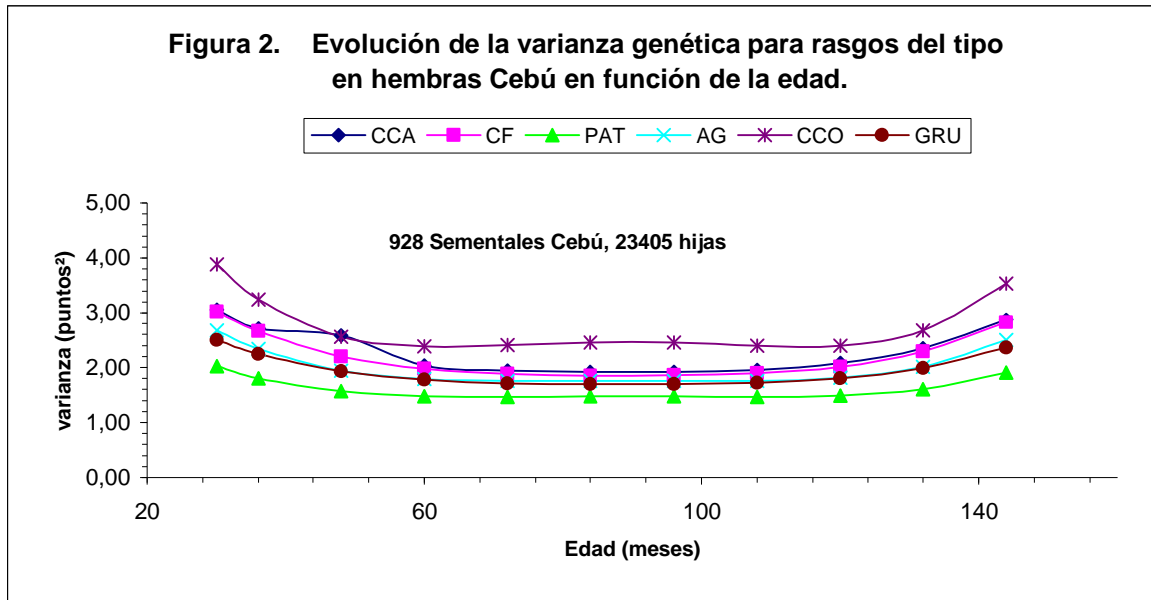
Los resultados de RRM convergieron rápidamente para $k = 1$; por el contrario cuando $k = 2$ los parámetros estuvieron fuera de espacio, sin embargo cuando se consideró nula la covarianza entre el intercepto (s_0) y los coeficientes lineal (s_1) y cuadrático (s_2) del efecto semental, se alcanzó la convergencia en las primeras 7 iteraciones. La comparación directa entre ambos logL es válida ya que el número de parámetros es el mismo y son los mismos efectos fijos incluidos en ambos modelos. En todos los casos se alcanzó mejor ajuste (tabla 2) cuando se empleó un modelo con $k = 2$, de manera que serán estos los resultados que se mostrarán.

Tabla 2. Comparación de log L de los modelos de regresión aleatoria aplicado a datos de clasificación del tipo en animales Cebú.

Detalle de Parte	k = 1	k = 2	Diferencia
Apariencia General	-35012	-34890	+ 22
Capacidad Corporal	-36149	-36115	+ 34
Capacidad Carnicera	-35928	-35827	+101
Grupa	-35798	-35694	+104
Patas	-35142	-35048	+96
Clasificación Final	-34426	-34307	+119

La evolución de las varianzas genéticas (s_{ij}^2) y heredabilidad (h_{ij}^2) para los i^{th} rasgos estudiados entre las j edades ($j = 30$ a 144 meses de edad) se muestran en las figuras 2 y 3 respectivamente. Excepto para los resultados de CCO, el patrón general fue muy semejante para todos los componentes del tipo. Durante los primeros 48 meses de edad se estimaron las mayores valores de s_{ij}^2 y h_{ij}^2 indicando que las clasificaciones más tempranas deben favorecerse en el trabajo de clasificación que lleva a cabo CENCOP, ya que no solo es imprescindible para disminuir el intervalo entre generaciones sino también coadyuvan adicionalmente a una evaluación genética más precisa. A partir de los 4 años de edad se estabilizan los parámetros, excepto en la edad extrema de los 132 meses de edad.

Este patrón de cambios en las s_{ij}^2 y h_{ij}^2 en la trayectoria de edad representada en esta base de datos, no se podía cuantificar tan fácilmente con los resultados del modelo clásico con una covariable fija para eliminar las variaciones en edad, lo cual representa de hecho una de las ventajas del uso de RRM. En términos de estimados promedios los valores de h_{ij}^2 de este estudio son muy similares a los previamente publicados en esta raza (Menéndez Buxadera y col., 1998b) y están dentro del rango de las publicaciones sobre caracteres del tipo en ganado de leche (Mark, 2004) y del ganado de carne (Gutiérrez y Goyache 2002).



Las estimaciones de correlaciones genéticas (r_g) entre el mismo carácter evaluado entre 30 y 84 meses de edad mostraron un patrón bastante similar en los 6 caracteres del tipo estudiados (tablas 3a; 3b y 3c). Cuando el rasgo se evalúa antes de los 48 meses todas las r_g fueron superiores a 0.85 dentro de ese intervalo de edad, sin embargo a partir de ese punto se manifiesta un decrecimiento importante y las r_g disminuyen hasta 0.75 entre los resultados a 30 y 84 meses de edad. No obstante cuando se utilizan los resultados a 36 o 48 meses de edad todas las r_g fueron superiores a 0.82 con edades entre 60 y 72 meses de edad, excepto para CCO entre 36 y 72 meses con $r_g = 0.773$ (Tabla 3b debajo de la diagonal) lo cual sugiere la existencia de una base genética muy semejantes en toda la evolución de estos caracteres en la trayectoria de edad representada. Estos valores de r_g permiten prever pocos cambios en el orden de merito de los sementales evaluados a edades tan distantes como los 3 y 7 años de edad de sus hijas.

Tabla 3a. Correlaciones genéticas* entre rasgos del tipo en animales Zebu evaluados a diferentes edades.

	Edad a la Clasificación (en meses)					
	30	36	48	60	72	84
30		0.993	0.940	0.850	0.752	0.672
36	0.995		0.973	0.904	0.822	0.749
48	0.951	0.978		0.978	0.930	0.878
60	0.873	0.918	0.980		0.986	0.955
72	0.780	0.838	0.932	0.985		0.991
84	0.690	0.756	0.871	0.949	0.989	

* Por encima de la diagonal los resultados de Apariencia General. Por debajo de la diagonal los resultados para Clasificación Final.

Tabla 3b. Correlaciones genéticas* entre rasgos del tipo en animales Zebu evaluados a diferentes edades.

	Edad a la Clasificación (en meses)					
	30	36	48	60	72	84
30		0.995	0.953	0.880	0.790	0.704
36	0.991		0.979	0.922	0.847	0.769
48	0.911	0.959		0.982	0.936	0.878
60	0.776	0.855	0.967		0.986	0.951
72	0.645	0.773	0.901	0.981		0.989
84	0.555	0.660	0.842	0.949	0.991	

* Por encima de la diagonal los resultados de Capacidad Carnica. Por debajo de la diagonal los resultados para Capacidad Corporal.

Tabla 3c. Correlaciones genéticas* entre rasgos del tipo en animales Zebu evaluados a diferentes edades.

	Edad a la Clasificación (en meses)					
	30	36	48	60	72	84
30		0.995	0.956	0.890	0.814	0.743
36	0.994		0.981	0.931	0.867	0.803
48	0.946	0.976		0.984	0.946	0.899
60	0.870	0.919	0.982		0.988	0.961
72	0.791	0.854	0.946	0.989		0.991
84	0.731	0.800	0.908	0.968	0.994	

* Por encima de la diagonal los resultados de Grupa. Por debajo de la diagonal los resultados para Patas.

Para ejemplificar la importancia de las correlaciones estimadas se preparo la tabla 4 donde se muestran los coeficientes de regresión del VG de los sementales evaluados tempranamente (30, 36 o 48 meses de edad de las hijas), sobre los VG de los mismos sementales a diferentes edades de sus hijas (30 a 84 meses). En la parte inferior de la tabla 4 se muestran los resultados de la predicción de VG_{CF} a edades adultas, en función de los resultados del propio rasgo a edades tempranas. Todos los valores de b fueron superiores a 0.8 (excepto para la predicción de VG_{CF84} a partir de VG_{CF30}). Cuando la evaluación se realiza según VG_{CF36} se alcanzan altas precisiones y un coeficiente de determinación $R^2 = 90\%$ de manera que resulta ventajoso la evaluación de CF a edades tempranas.

Tabla 4. Coeficientes de regresión (b) del Valor Genético (VG) para Clasificación Final del tipo sobre el VG para otros componentes del tipo estimados entre 30 y 84 meses de edad.

Rasgos del Tipo	Edad a la Evaluación	Coef. de b para predecir el VG a diferentes edades*					Error Estándar	$R^2\%$
		30	36	48	60	84		
Capacidad Corporal	30	0.9934	0.9733	0.9355	0.9007	0.8407	0.010	73.3
	36	1.0116	0.9983	0.9722	0.9466	0.8972	a	a
	48	1.0245	1.0251	1.0230	1.0164	0.9898	0.019	89.9
Capacidad Carnicera	30	0.9343	0.9143	0.8751	0.8369	0.7638	0.010	71.7 a
	36	0.9567	0.9402	0.9072	0.8743	0.8087	a	91.7
	48	0.9789	0.9697	0.9501	0.9287	0.8801	0.015	
Apariencia General	30	0.8727	0.8552	0.8217	0.7897	0.7306	0.060	72.2
	36	0.8935	0.8799	0.8533	0.8269	0.7749	a	a
	48	0.9140	0.9087	0.8965	0.8819	0.8458	0.014	94.5
Grupa	30	0.8210	0.8063	0.7776	0.7497	0.6963	0.011	68.7
	36	0.8449	0.8329	0.8090	0.7851	0.7373	a	a
	48	0.8728	0.8666	0.8529	0.8375	0.8023	0.015	87.1
Patas	30	0.6806	0.6729	0.6580	0.6433	0.6148	0.012	62.5
	36	0.7020	0.6870	0.6865	0.6753	0.6513	a	a
	48	0.7288	0.7290	0.7277	0.7241	0.7097	0.014	79.8
Clasificación Final	30		0.9748	0.9262	0.8790	0.7924	0.003	76.8
	36	1.0212		0.9584	0.9176	0.8287	a	a
	48	1.0392	1.0264		0.9721	0.9120	0.0141	99.8

* Los resultados para 72 meses de edad no se muestran por problemas de espacio, pero fueron intermedios entre los valores de 60 y 84 meses. (928 sementales Cebú con 23405 hijas).

Los resultados de CF fueron empleados como variables independientes para predecir el VG del resto de los rasgos estudiados. Los resultados de la tabla 4 se organizaron en forma ascendente según el nivel de *b*. El patrón de respuesta fue muy similar para todos los rasgos. En general los niveles de *b* aumentan en la medida que se incrementa la edad a que se evaluó el VG_{CF} , lo mismo sucede con los estimados de R^2 . Las predicciones son más precisas entre edades adyacentes.

Conclusiones.

Todos los estimados de este trabajo apuntan en el sentido de que es posible predecir los VG de rasgos del tipo en animales adultos en función de sus evaluaciones a edades tempranas, particularmente a los tres años de edad. Según estos resultados la creencia de que hay que esperar a que los animales sean adultos para evaluar adecuadamente los cambios morfológicos de estos caracteres no se justifican y pueden ser reemplazados por una política de evaluación temprana, que puede contribuir a un mayor beneficio en el programa de mejora genética de esta raza.

El estudio realizado demuestra que es factible aplicar RRM a los datos de clasificación del tipo, sin embargo aparentemente se incorpora mayor complejidad matemática a las estimaciones, no obstante son múltiples las ventajas que pueden obtenerse con los resultados. Se pueden estimar los valores genéticos de todos los animales a todo lo largo de la trayectoria de edad representada. Esta posibilidad puede aportar beneficios en modo alguno despreciables, ya que facilita la identificación de aquellos animales con una desviación genética importante de la talla y forma durante su desarrollo hasta el estado adulto.

Referencias.

- Brody, S., 1945. Bioenergetic y Growth. Reinhold Pub. Corp. New York, USA. Chapter 17: pag 575- 671.
- Fernández Chuairey, L., 2004. Modelos Estadísticos-Matemáticos en el análisis de la curva de lactancia y factores que la afectan en el genotipo Siboney de Cuba. Tesis presentada en opción al grado científico de doctor en Ciencias Veterinarias. La Habana, 100 p.
- Fitzhugh, H.A., 1976. Analysis of growth curves and strategies for altering their shape. J. Anim. Sci. 42(4),1058-1072.
- Forabosco, F.; Groen, A.F.; Bozzi, R.; van Arendonk, J.A.M.; Filippini, F.; Boettcher, P. y Bijma, P., 2004. Phenotypic relationship between longevity, type traits and production in Chianina beef cattle. J. Anim. Sci. 82: 1572-1580.
- Gilmour, A. R.; Cullis, B.R.; Welham, S.J. y Thompson, R., 2000. ASREML Reference Manual. NSW Agric. Biom. Bull. NSW Agriculture, Locked Bag, Orange, NSW 2800, Australia.
- Gutierrez, J.P. y Goyache, F., 2002. Estimation of genetic parameters of type traits in Asturiana de los Valles beef cattle breed. J. Anim. Brred. Genet. 119: 93-100.

- Henderson, C.R.Jr., 1982. Analysis of covariance in the mixed model: higher-level, non homogeneous and random regressions. *Biom.* 38: 623-640.
- Jamrozik, J. y Schaeffer, L.R., 1997. Estimates of genetic parameters for a test day model with random regression for production of first lactation. *J. Dairy Sci.* 80: 762-770.
- Janssens, S.; Winandy, D.; Tylleman, A.; Delmotte, Ch.; Van Moeseke, W. y Vandepite, W., 2004. The linear assesment scheme for sheep in Belgium: breed averages and assesor quality. *Small Rum. Res.*, 51 (1) 85-95.
- Kirkpatrick, M.D.; Lofsvold, D. y Bulmer, M., 1990. Analysis of the inheritance, selection and evolution of growth trajectories. *Genetics* 124: 979-993.
- Mark, T., 2004. Applied genetic evaluations for production and functional traits in dairy cattle. *J. Dairy Sci.* 87: 2641-2652.
- Menéndez Buxadera, A.; Torres, L.; Jiménez, M. y Larios, M., 1998a. Clasificación del tipo en ganado Cebú. 1. Resultados generales. *Rev. Cubana Ciencias Agric.* 32(2) 121-133.
- Menéndez Buxadera, A.; Torres L.; Jiménez M. y Larios, M., 1998b. Clasificación del tipo en ganado Cebú. 2. Parámetros genéticos. *Rev Cubana Ciencias Agric* 32(2) 134-140.
- Menéndez Buxadera, A.; Guerra, D.; Planas, T.; Ramos, F. y Fernández Chuairey, L., 2004. Parámetros genéticos del peso vivo de machos Cebú en prueba de comportamiento en condiciones de pastoreo de Cuba, mediante modelo animal univariado, multicaacteres y regresiones aleatorias. En preparación.
- Meyer, K. y Hill, W. G., 1997. Estimation of genetic and phenotypic covariance for longitudinal or repeated records by restricted maximum likelihood. *Liv. Prod. Sci.* 47: 185-200.
- Schaeffer, L.R. y Dekkers, J.C.M., 1996. Random Regressions in animal models for test day production in dairy cattle. *Proc. 5th World Congress Genet. Applied to Liv. Prod. Vol XVIII Ont. Guelph, Canada* pp 443-446.
- Schaeffer, L.R., 2004. Application of random regression models in animal breeding. *Liv. Prod. Sci.* 86: 35-45.
- Thompson, J.R.; Lee, K.L. y Freeman, A.E., 1983. Evaluation of a linear type appraisal system for Holstein cattle. *J. Dairy Sci.* 66: 325-331.
- Van Steenbergen, E.J., 1989. Description and evaluation of a linear scoring system for exterior traits in pigs. *Liv. Prod. Sci.*, 23 (1-2) : 163-181.

Evaluación de la producción de leche en vacas bajo un sistema de control lácteo alternativo mediante regresión aleatoria.

A. Menéndez Buxadera* y J. Caunedo**

* Centro de Control Pecuario, Ministerio de la Agricultura
Conill y Boyeros, Ciudad de La Habana, Cuba.

**Empresa Pecuaria “Niña Bonita”, Cangrejera, La Habana, Cuba.

La aplicación de un sistema de control individual de la producción de leche (CPL) es un método comúnmente empleado en cualquier ganadería desarrollada. Cuba no es una excepción, y desde hace casi tres décadas se lleva a cabo este procedimiento. Sin embargo en los últimos años sólo una proporción, relativamente pequeña, de los rebaños lecheros están bajo algunas formas de CPL, debiéndose incrementar aceleradamente la aplicación de este sistema de control a la mayor parte de la masa ganadera dedicada a esta actividad productiva y no sólo a una parte de ella, denominada como ‘genética’.

Varios son los procedimientos aplicados en el mundo con vistas a la implementación de CPL. En general, estos pueden agruparse según la periodicidad del control individual, entre aquellos que utilizan el control diario o semanal a través de la lactancia de cada vaca u otros que emplean el pesaje mensual o bimensual. Desde el punto de vista de precisión, es evidente que los resultados del control diario brindan los estimados más exactos, no obstante, las experiencias internacionales demuestran que no existen grandes diferencias entre este y los estimados según el control individual mensual. Los interesados pueden encontrar una amplia revisión y discusión actualizada sobre el tema en McDaniel (1965); Schaeffer y col. (2000) y Berry y col., (2005). Más recientemente están disponibles las normas internacionales para esta actividad (ICAR, 2004).

Cualquiera sea la variante de CPL es necesario realizar dos visitas a cada rebaño en el horario del ordeño de la tarde (generalmente reconocido como el primer ordeño) y en la madrugada. Evidentemente si se aspira a desarrollar masivamente el empleo de CPL se necesitarán grandes recursos para ambas visitas. Para facilitar esta tarea se ha propuesto un sistema alternativo que se basa en el pesaje mensual pero sólo efectúa el control individual de un solo ordeño y el otro es estimado por el uso de factores de ajustes.

Este procedimiento permite un sustancial ahorro en los recursos materiales necesarios para el control, mostrando además resultados muy satisfactorios, sin embargo, en su ejecución requiere del uso de un control automático del intervalo de tiempo entre dos ordeños (Schaeffer y col., 2000 y Berry y col., 2005). Una importante adecuación fue desarrollada por Wiggans (1981), en el cual se sustituye el control del tiempo, por el volumen total de leche producido en el rebaño, pero proporcionalmente en cada ordeño. Esta nueva alternativa fue comprobada su efectividad en las condiciones de Cuba (Menéndez Buxadera y col., 1986), sin embargo por diversas razones, esto no fue aplicado en condiciones prácticas.

En este sentido precisamente se desarrolló este trabajo, que tiene como objetivo principal la evaluación de una nueva alternativa para estimar la producción diaria de leche de las vacas en control individual.

MATERIAL Y METODO:

Para este trabajo se recibió una copia de los controles de pesaje individual llevados a cabo durante los años 2001 al 2004 en 40 vaquerías de una empresa pecuaria situada en la provincia Habana. En total se disponía de 45394 registros totales del ordeño de la tarde (PM) y de la madrugada (AM) correspondientes a 3433 vacas entre el 1 al 12 parto. Una pormenorizada edición de los datos fue aplicada, eliminándose aquellas observaciones pertenecientes a vacas enfermas el día de pesaje, carga inicial de los datos donde se registró solamente el total de producción y no lo correspondiente a cada ordeño; animales secos o eliminados del control, animales con tres ordeños, así como unos pocos datos del año 2001, etc. Finalmente quedaron disponibles 33982 registros de 3107 vacas entre el primero y el séptimo o mas partos distribuidas en 31 vaquerías. Estas hembras eran progenies de sementales Holstein; Siboney de Cuba; Cebú Lechero; Caribe de Cuba y un grupo de padres desconocidos pero cuyas vacas fueron clasificadas como del tipo lecheras, según los requerimientos establecidos en CENCOP. Este último conjunto de datos (A) fue la base para todo el trabajo desarrollado.

Varios estudios fueron desarrollados cuyas características generales se presentan a continuación.

Forma de la curva de lactancia. Los controles efectuados en los ordeños AM; PM y el TOT=AM+PM, contenidos en el conjunto de datos A fueron analizados de acuerdo a un modelo lineal generalizado en el que se incluyeron los efectos de unidad de pesaje (31 niveles); fecha del pesaje (40 niveles); número de lactancia (7 niveles); número de controles (12 niveles) y los días de lactancia en este caso como clases de 7 días con 44 niveles (de la semana 2 a la 45). Por el momento sólo se está interesado en representar la forma de la curva de lactancia en los controles AM; PM y TOT ya que la mejor ecuación formal se estudiará mas adelante.

Importancia de los controles parciales. Los mismos datos anteriores fueron empleados para analizar la importancia relativa de la predicción de la producción TOT en función del control de un solo ordeño. En términos estadísticos se aplicó un modelo muy similar al descrito previamente pero incorporando una covariable intra mes de pesaje (12 niveles, de esta forma se representaba de una mejor manera este importante efecto). En este caso se estudiaron los resultados según la producción AM; PM y un sistema alterno (ALT) que consiste en un mes se controla el pesaje AM y en el próximo PM. Las correlaciones entre los resultados de cada ordeño y la producción total del día se estudio aplicando otro modelo lineal mixto con los mismos efectos descritos para la forma de la curva de lactancia pero incorporando al padre de la vaca como efecto aleatorio. Esto último se llevó a cabo para las razas Holstein; Siboney de Cuba, Caribe de Cuba y Cebú Lechero.

Parámetros genéticos de los controles parciales. Este análisis se realizó sólo en los datos de la raza Holstein debido a una mayor cantidad de sementales representado. Sin embargo se decidió aplicar una estrategia que posibilitara la mayor representatividad posible de la estructura de los datos, para ello el vector de los efectos aleatorios debido al semental padre de las vacas (S_v), estuvo conformado por dos vectores, $S_v = S_h + S_o$, donde S_h son los sementales Holstein que se denominaron '*informativos*', en términos de los componentes de varianza estimados y los parámetros genéticos de este genotipo, mientras que el resto de

los sementales (So) se incorporan en los componentes de varianza totales pero serán '**no informativos**' a los efectos de los parámetros genéticos. El pedigree de los sementales '**informativos**' fue extendido por la vía padre, resultando finalmente un total de 345 sementales de esta raza.

En términos estadísticos los 21399 registros mensuales de producción individual de leche descritos en la primera parte de este informe fueron estudiados por un modelo semental uní variado de regresiones aleatorias según la siguiente representación:

$$y = \sum_{m=0}^{k-1} b_m f_m(t_{ij}^*) + \sum_{m=0}^{k-1} S_m f_m(t_{ij}^*) + \sum_{m=0}^{k-1} I_m f_m(t_{ij}) + \sum_{m=0}^{k-1} a_m f_m(t_{ij}) + \epsilon_{ij}$$

donde:

- y es un vector de los registros de producción de las progenies del i ésimo semental.
- b_m son vectores de efectos fijos debidos a fecha de control (40 niveles), unidad de control (28 niveles) y número de lactancia (7 niveles).
- S_m y I_m son conjuntos (de dimensión $s =$ número de sementales) de m coeficientes de regresión aleatorios de orden k_i para los efectos genéticos aditivos del Semental Holstein que se denominó '**informativo**' (S_m) y de los otros sementales (I_m) cuyos efectos solo se incluyen en la estimación de la varianza fenotípica total..
- a_m es un efecto aleatorio de ambiente permanente (P_i) debido a las repeticiones de la variable dependiente en cada progenie (1270 en total) de un semental '**informativo**'.
- ϵ_{ij} es un vector de efecto residual aleatorio con varianza heterogénea para 6 clases de intervalos de duración de la lactancia –agrupadas en decenas–(entre 1 y 6; de 7 a 12; de 13 a 20; de 21 a 24; de 25 a 28 y mas de 29).

Los elementos $f_m(t_{ij}^*) = \Phi_i$ corresponden con las matrices de incidencia de los modelos lineales clásicos (X y Z) y como elementos contienen los m coeficientes de regresión de orden k_i . El término t_{ij}^* es el j día de lactancia en la que se registró cada variable dependiente expresada en forma estandarizada entre -1 a + 1, lo que es necesario para la utilización de los polinomios ortogonales de orden k_i . Para los efectos fijos, los m coeficientes fueron los de un polinomio de orden $k_i = 5$, excepto para el numero de lactancia que fue de orden $k_i = 2$ (ecuación cúbica). Para los efectos aleatorios a , I_m y S_m se compararon varios modelos con $k_i = 1, 2, 3$ y 4. El símbolo \sum quiere decir que el efecto total de cada factor se estima como una sumatoria del polinomio asumido y los m coeficientes estimados, es decir, m_0 (intercepto) + m_1 (coeficiente lineal) + m_2 (coeficiente cuadrático) + etc.

Para el modelo descrito el sistema de ecuaciones será:

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z_1 & X'R^{-1}Z_2 & X'R^{-1}Z_3 \\ Z_1'R^{-1}X & Z_1'R^{-1}Z_1 + A^{-1} \otimes K_{Si}^{-1} & Z_1'R^{-1}Z_2 & Z_1'R^{-1}Z_3 \\ Z_2'R^{-1}X & Z_2'R^{-1}Z_1 & Z_2'R^{-1}Z_2 + I_{So} \otimes K_{So}^{-1} & Z_2'R^{-1}Z_3 \\ Z_3'R^{-1}X & Z_3'R^{-1}Z_1 & Z_3'R^{-1}Z_2 & Z_3'R^{-1}Z_3 + I_{Pi} \otimes K_{Pi}^{-1} \end{bmatrix} \begin{bmatrix} b \\ S_i \\ S_o \\ P_i \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z_1'R^{-1}y \\ Z_2'R^{-1}y \\ Z_3'R^{-1}y \end{bmatrix}$$

donde b representa los estimados de los efectos fijos indicados previamente, A es una matriz del numerador de las relaciones de parentesco entre los sementales '*informativos*', \otimes es el símbolo de producto Kronecker y K_{Si} es la matriz de los coeficientes de regresión aleatorios del efecto semental *informativo*, en este caso la matriz de varianzas y covarianzas genéticas será $G_s = \frac{1}{4}(A^{-1} \otimes K_{Si}^{-1})$. Finalmente I_{So} e I_{Pi} representan matrices de identidad de orden igual al número de sementales *no informativos* y número de vacas con registros repetidos respectivamente, mientras que K_{So} y K_{Pi} son las matrices de varianzas y covarianzas para los correspondientes coeficientes de regresión aleatorios.

Los componentes de (co)varianza para cada variable dependiente fueron estimados por diferentes modelos. Las diferencias se debían al nivel de ajuste del polinomio utilizado para los efectos aleatorios de semental *informativo*, en este caso para $k_i = 1$ y 2 . Para los efectos S_o y P_i , se compararon modelos $k_i = 1, 2$ y 3 . En total se realizaron 12 corridas. Los resultados de logL se utilizaron para seleccionar el mejor modelo.

Todo el estudio se realizó utilizando el software ASREML (Gilmour y col., 2000) el cual brinda no solo las soluciones para todos los efectos incluidos en el modelo y los coeficientes de los polinomios (Φ_i) sino también para todos los componentes de varianza descritos previamente. Sin embargo, para estimar los parámetros genéticos, se debe aplicar un procedimiento adicional siguiendo las propuestas de Jamrozik y Schaeffer (1997), según las cuales la varianza genética para efecto genético aditivo debido al semental (s_s^2) para el j día de lactancia se estima por:

$$s_{s_j}^2 = 4 [\Phi_j' K_{Si}^{-1} \Phi_j]$$

mientras que las covarianzas genéticas entre los días j y k ($s_{s_{jk}}$) será:

$$s_{s_{jk}} = \Phi_{s_j}' K_{Si}^{-1} \Phi_{s_k}$$

Por otro lado las varianzas para los efectos P_i se estiman por $s_{pi}^2 = \Phi_j' K_{Pi}^{-1} \Phi_j$ y para los efectos de los sementales no informativos será $s_{so}^2 = \Phi_j' K_{So}^{-1} \Phi_j$. Finalmente la heredabilidad de cada variable dependiente en cada día de lactancia que se desee, se estimarán por:

$$h_j^2 = \frac{s_{s_j}^2}{s_{s_j}^2 + s_{so_j}^2 + s_{pi_j}^2 + \epsilon_{ij}}$$

Por otro lado la correlación genética entre la producción de leche en el día de lactancia j y la producción en el día k será:

$$R_{g_{jk}} = \frac{S_{s_{jk}}}{\{ S_{s_j}^2 * S_{s_k}^2 \}^{0.5}}$$

Estos parámetros se refieren a un día de lactancia específico, no obstante con este procedimiento de MRA se pueden obtener también los mismos indicadores para las producciones acumuladas. El elemento adicional es la sumatoria de los coeficientes polinomiales (Φ_j) que corresponden a la etapa que se desea. De igual manera con las soluciones para cada semental (a_i) para los m coeficientes de regresión aleatorios se pueden estimar los Valores Genéticos (VG^i) para cualquier punto de la trayectoria del periodo de lactancia o para la producción acumulada. A modo de ejemplo, para el i animal el VG para la producción en el día 100 de la lactancia (VG_{100}^i) será:

$$VG_{100}^i = \Phi_{100} a_i$$

mientras que para el mismo semental se puede estimar su VG para la producción acumulada en los primeros 100 días de lactancia (\overline{VG}_{100}^i) será:

$$C = \sum_{i=5}^{100} \Phi \quad \text{y} \quad \overline{VG}_{100}^i = C * a_i$$

donde a_i representa la solución para el animal i que viene dado por un polinomio de orden k_m utilizado en el modelo. Cuando se aplica MRA aparentemente se incorpora mayor complejidad matemática a las estimaciones, no obstante son múltiples las ventajas que pueden obtenerse con los resultados. Manipulando las soluciones de a_i y los coeficientes de Φ_j se pueden estimar varios indicadores de mucha importancia, así por ejemplo el VG para la producción acumulada a 200 días de lactancia será:

$$C = \sum_{i=100}^{200} \Phi \quad \text{y} \quad \overline{VG}_{200}^i = C * a_i$$

Con tales resultados se puede estimar el VG para los segundos 100 días de lactancia así como la persistencia (VG_P) de la producción de leche de las hijas del semental. En este caso:

$$VG_P = \frac{VG_{200} - VG_{100}}{VG_{100}}$$

Toda esta labor se facilita con el uso del ASREML que brinda los valores de Φ_j para todos los puntos de la trayectoria de la duración de la lactancia.

RESULTADOS.

Indicadores generales. El comportamiento de los genotipos representados en este estudio (tabla 2) puede catalogarse de modesto. Aún cuando el propósito de este informe no es la comparación entre estas razas representadas, varios elementos deben resaltarse. En primer término, la duración de la lactancia promedio es muy baja, lo cual puede estar relacionado a insuficientes niveles de alimentación según puede inferirse del bajo peso vivo antes del parto (momento del secado) de estos animales. En segundo lugar, se manifiesta una altísima variación del orden de alrededor al 50% (desviación estándar/media del rasgo) tanto para los pesajes mensuales como los indicadores clásicos de producción a un tiempo

fijo. En general los resultados de la muestra estudiada están dentro del contexto de los niveles productivos del vacuno en las condiciones actuales de Cuba (ver Borroto, 2002), por tanto se consideran como representativas de la población ganadera del país.

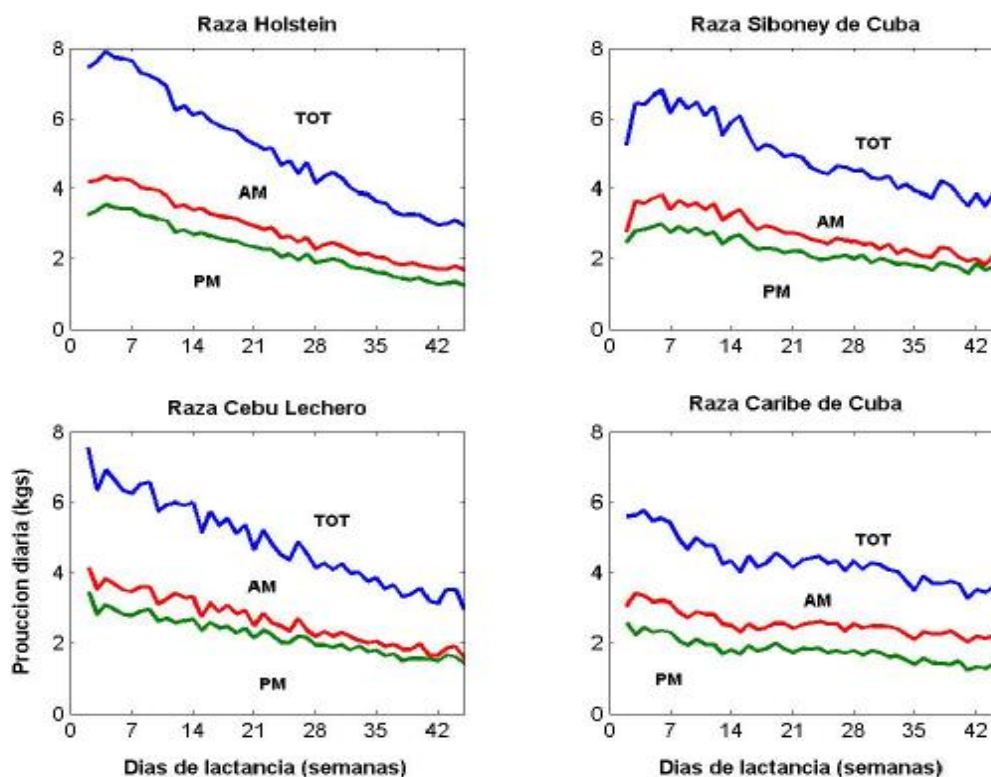
Tabla 2. Indicadores productivos (kgs) de vacas lecheras*.

	Holstein	Siboney	Caribe	C. Lechero	Varias
No. Sementales	240	44	22	25	--
No. Vacas	1270	341	240	376	880
Controles mensuales	13684	3740	3281	3794	9483
Controles por vaca	5.1(2.8)	4.9(2.7)	4.9(2.7)	4.9(2.7)	6.1(2.8)
Producción AM	3.4(1.7)	3.2(1.4)	2.7(1.4)	3.1(1.5)	2.9(1.5)
Producción PM	2.8(1.5)	2.4(1.1)	2.3(1.2)	2.5(1.3)	2.4(1.2)
Producción Total	6.3(3.2)	5.6(2.5)	4.9(2.5)	5.7(2.7)	5.3(2.7)
Peso Vivo	434(47)	409(66)	429(39)	427(69)	437(55)
Leche en 100 días	732(312)	619(266)	593(249)	640(280)	541(283)
Leche en 200 días	1161(578)	964(488)	913(448)	1018(518)	818(521)
Leche en 244 días	1269(668)	1044(560)	991(517)	1110(597)	883(595)
Leche en 305 días	1335(730)	1095(616)	1035(563)	1162(653)	920(645)
Duración Lactancia	215(78)	188(86)	203(77)	200(83)	172(85)

* Entre paréntesis la desviación estándar de cada rasgo.

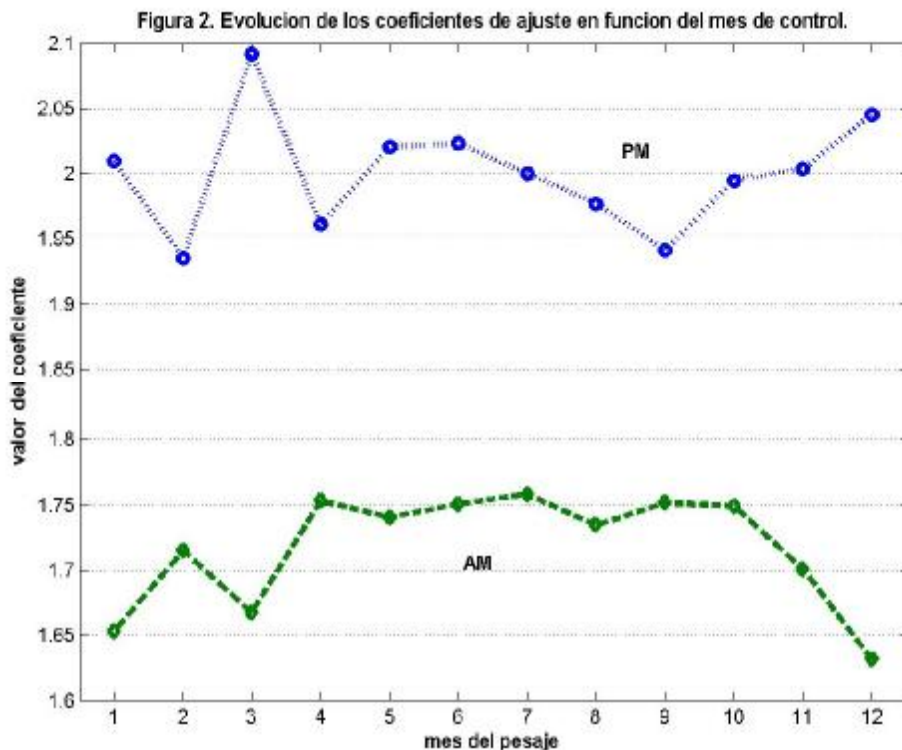
Forma de la curva de lactancia. Todos las causas de variación incluidas en los modelos de análisis de los pesajes mensuales (AM; PM y TOT) ejercieron efectos altamente significativos. Las constantes para duración de la lactancia (semanas 2 a la 45) se muestran en la figura 1.

Figura 1. Forma de la Curva de Lactancia en hembras lecheras de Cuba



Los picos máximos de producción se manifestaron alrededor de la tercera a cuarta semana después del parto en Holstein y Siboney. Similares resultados fueron obtenidos por Fernández Chuarey (2004) para animales Siboney y aproximadamente los mismos niveles de producción. El patrón del Cebú Lechero y el Caribe de Cuba fue bastante errático con variaciones mensuales durante todo el período de lactancia. Todos los genotipos presentan una tasa decreciente muy marcada en el resto de la lactancia, lo cual también es un reflejo de los bajos niveles de alimentación. En todos los casos, la producción de leche en el ordeño AM fue superior a la PM, sin variaciones en la forma entre ambos y obviamente en TOT.

Controles mensuales. Los coeficientes de determinación ($R^2\%$) de los modelos para predecir los resultados de TOT en función del control AM; PM y ALT fueron entre $R^2 = 94$ a 96% para las tres alternativas, lo cual implica una alta precisión en los estimados muy similares a los publicados por Schaeffer y col. (2000) y más recientemente Berry y col (2005). Es necesario destacar que el uso de la covariable intra mes de pesaje, manifestó importantes diferencias según el control utilizado (AM o PM), dado que estos serán los factores de ajustes a utilizar para predecir TOT según el control AM o PM, se presentan en la figura 2. El uso de estos factores es muy simple, se multiplican los resultados del ordeño controlado (AM o PM) por el valor del coeficiente que le corresponde y el resultados es una predicción de la producción total del día.



En general, las diferencias no solo se manifestaron en la forma de la curva de respuesta, sino también en la magnitud de los coeficientes. Los del ordeño AM fueron alrededor de 17% inferiores en comparación a los de PM. Independientemente de tales diferencias, el alto nivel de ajuste de los modelos aplicados debe posibilitar una precisión muy alta en las predicciones. En tal sentido deben interpretarse los valores de las correlaciones entre los

efectos del semental, de acuerdo a los niveles productivos de sus hijas en el ordeño AM, PM o ALT y los resultados del día de control (tabla 3).

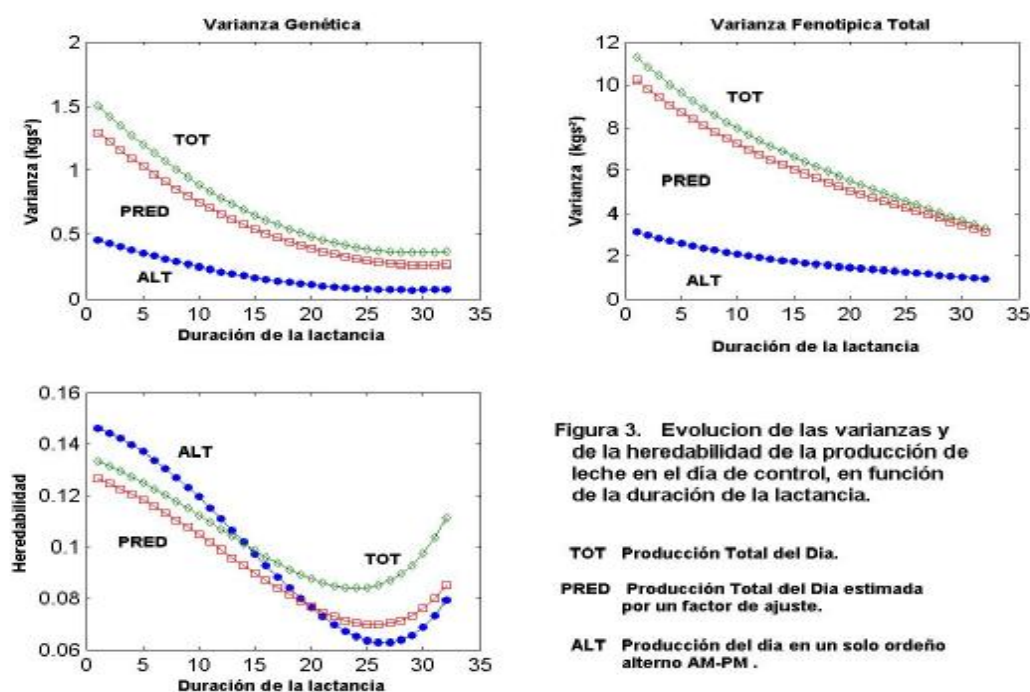
Tabla 3. Correlaciones entre el efecto Semental* para el control mensual total y los sistemas alternativos, AM, PM o ALT.

Raza	N	AM	PM	ALT	Rango (kgs)
Holstein	149	0.988	0.984	0.995	3.0 a 8.0
Siboney de Cuba	29	0.974	0.949	0.977	4.2 a 6.8
Cebu Lechero	21	0.987	0.982	0.994	4.7 a 6.4
Caribe de Cuba	17	0.993	0.991	0.988	2.7 a 4.3

- Sementales con al menos 20 pesajes mensuales, por ello las diferencias con los valores de la tabla 1.

De acuerdo a estos resultados, se debe esperar que el orden de mérito de los sementales es el mismo, cualquiera sea el control lácteo alternativo utilizado, respecto al clásico total diario (TOT=AM + PM) empleado hasta el presente. Con la excepción de los resultados del Caribe de Cuba, todas las correlaciones entre el sistema ALT fueron superiores a las de AM o PM. A pesar de los bajos niveles productivos alcanzados en estas unidades, nuevamente debe resaltarse las variaciones entre sementales, lo cual puede indicar un posible espacio nada despreciable, para el trabajo de selección.

Parámetros genéticos. La comparación de los resultados según el procedimiento estadístico descrito previamente, indicó que el modelo de mejor ajuste fue aquel que incluyó una ecuación lineal para el efecto de semental, tanto ‘*informativo*’ como ‘*no informativo*’, mientras que fue un polinomio de orden 3 (ecuación cúbica) para el efecto de ambiente permanente debido a repeticiones del mismo rasgo en cada vaca. A continuación se presentarán los parámetros estimados según este modelo en toda la trayectoria de la lactancia (figura 3).



Los resultados manifiestan una misma tendencia para las tres variables estudiadas. Las varianzas tanto genética como fenotípicas, se deprimen marcadamente en la medida que avanza la lactancia. Es necesario resaltar la semejanza de los resultados cuando se usa un solo ordeño y el factor de corrección para predecir (PRED) la producción total del día. La magnitud absoluta de las varianzas del sistema ALT es inferior a los estimados de TOT y PRED lo cual es lógico ya que se trata de un solo ordeño, sin embargo debe tomarse en cuenta que los resultados de esta alternativa de control presenta un valor de heredabilidad (h^2) ligeramente superior en los primeros 150 días de lactancia. Siguiendo la descripción del procedimiento estadístico, se estimaron los valores de h^2 para la producción de leche acumulada en diferentes períodos de lactancia y los resultados se presentan en la tabla 4.

Tablas 4. Estimados de heredabilidad de la producción de leche acumulada en diferentes períodos de lactancia y por diferentes sistemas de control*.

Carácter	AM+PM Real	ALTERNO	PREDICCIÓN
Leche a 100 días	0.124	0.135	0.117
Leche a 200 días	0.114	0.109	0.106
Leche a 305 días	0.108	0.107	0.106
Leche entre 100 y 200 días	0.116	0.119	0.110
Leche entre 200 y 305 días	0.111	0.108	0.101

*Los sistemas se describen en el texto.

Los valores de h^2 para la producción de leche acumulada manifestaron una depresiva respuesta en la medida que avanza la lactancia, en forma similar a los valores de los pesajes mensuales. En todos los casos los estimados de PRED son ligeramente inferiores. De acuerdo a estos resultados puede indicarse que alrededor del 12 % de las diferencias productivas entre los animales tienen un origen genético aditivo. Estos parámetros son similares, ya sean expresados en forma de control mensual real (AM+PM); un solo ordeño (ALTERNO) o predecidos (PRED) según un solo control (AM o PM) y el uso de un factor de ajuste, o para la forma clásica de producción acumulada en diferentes períodos de lactancia. En términos generales los resultados de este estudio son muy similares a los resultados recientes de Cuba con una muestra de datos muy numerosa y para la producción de leche acumulada en 305 días de lactancia (Menéndez Buxadera, 2005) con $h^2 = 0.151$ en animales Holstein, mientras que Guerra y col., (2002) estimaron valores de $h^2 = 0.15$ en leche acumulada en 244 días de lactancia en animales Siboney de Cuba. Finalmente, el único estimado disponible mediante un modelo de MRA en animales Siboney de Cuba para la producción de leche en el día de control (Guerra, comunicación personal 2005) arrojó valores de $h^2 = 0.10$ muy similar a los presentados en la figura 4. Los autores de este informe no han encontrado referencias disponibles, tanto internamente como en la literatura internacional donde se estudien los parámetros genéticos de la producción de leche en un solo ordeño, aunque el sistema AM+PM es reconocido oficialmente (ICAR 2004).

De las diferentes variables estudiadas, se prefiere recomendar los parámetros estimados a 200 días de lactancia, ya sea para los controles mensuales reales o predecidos así como para la producción de leche acumulada. Por un lado los valores de h^2 son muy similares al resto de los estimados y por otro manifiestan un alto grado de correlación genética (r_g) con el resto de los días de lactancia (tabla 5).

Tabla 5. Correlaciones genéticas entre la producción mensual de leche a diferentes días de lactancia.

Días de lactancia	50	100	150	200	250	305
50	1.000	0.994	0.967	0.897	0.752	0.518
100		1.000	0.989	0.940	0.895	0.607
150			1.000	0.980	0.965	0.718
200				1.000	0.966	0.842
250					1.000	0.953
305						1.000

Los valores de r_g disminuyen en la medida que aumenta la distancia (en días) entre los períodos de lactancia. Este es el patrón característico de los estudios publicados sobre parámetros genéticos de los controles lecheros mensuales (ver Swalve, 2000). Los valores a 200 días se consideran como bastante adecuados.

El hecho de que los valores de h^2 son relativamente bajos pudiera interpretarse como que no existe espacio para el trabajo de selección y mejora. Sin embargo la realidad es todo lo contrario, aunque efectivamente el proceso será mas complejo y de respuestas mas lentas, pero en modo alguno despreciables. En tal sentido, es oportuno revisar la magnitud de las variaciones existente en el Valor Genético de los sementales. Para cumplimentar este punto de vista, se utilizaron las soluciones para el efecto de semental para cada una de las variables dependientes empleadas, así como las posibles nuevas alternativas que se plantearon previamente. El MRA aplicado brinda todos los parámetros necesarios para llevar a cabo estas estimaciones y constituye de hecho, una de las ventajas del uso de este procedimiento. En la figura 4 se presentan globalmente una síntesis de estos nuevos resultados.

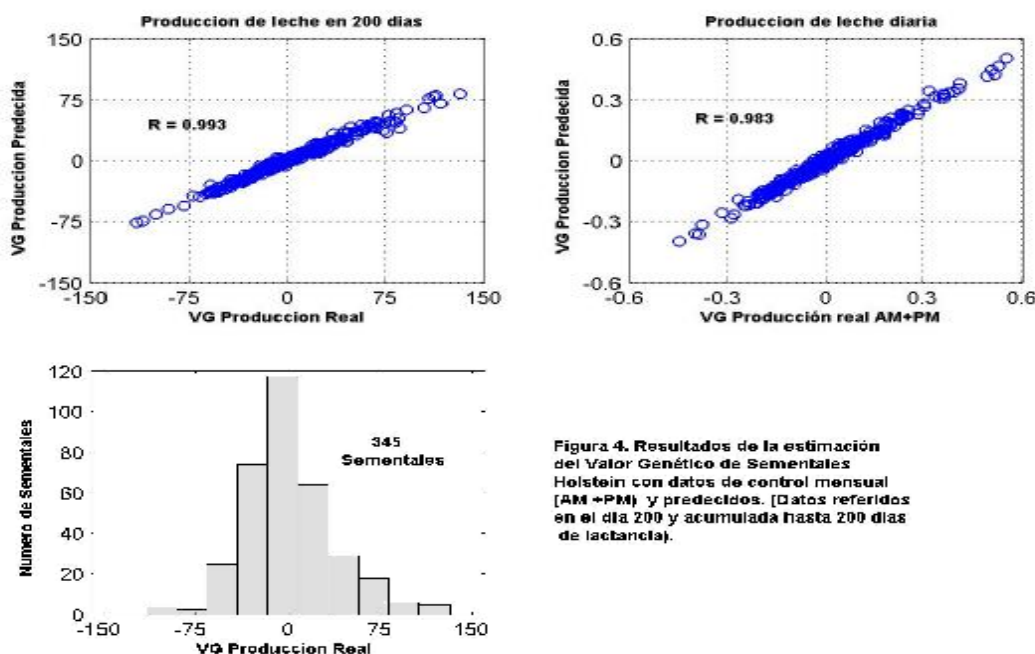
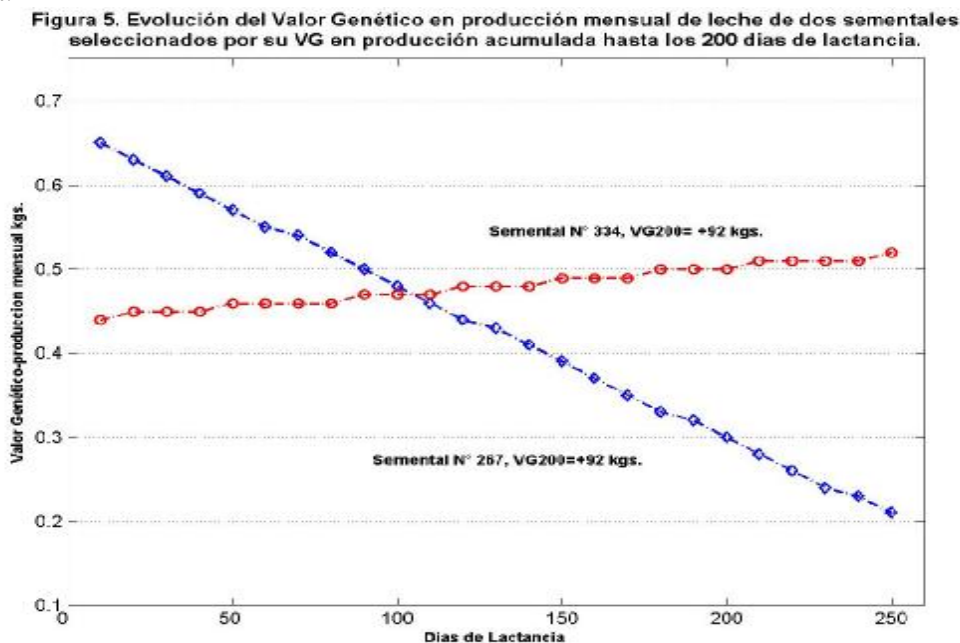


Figura 4. Resultados de la estimación del Valor Genético de Sementales Holstein con datos de control mensual (AM +PM) y prediccidos. [Datos referidos en el día 200 y acumulada hasta 200 días de lactancia].

El primer elemento a resaltar es el alto nivel de correlación entre los VG estimados con los datos reales y los predecidos (parte superior de la figura 4), sean estos la producción en el día de control o la acumulada hasta un punto determinado de la trayectoria de la lactancia, en este caso se eligió los 200 días por las razones ya mencionadas. Estos resultados implican que el método empleado en este estudio puede aplicarse con alto grado de confiabilidad en condiciones prácticas. Con respecto a las variaciones, los resultados demuestran que en esta muestra de datos, se manifiesta un total de 263 kgs de diferencias entre los VG de los mejores y peores sementales, lo cual es una evidencia que aún en estos bajos niveles de producción existe margen para el trabajo de mejora.

Los resultados de la figura 4 muestran algunas de las ventajas de los modelos de regresión aleatoria, no obstante, existen otros elementos que deben resaltarse. Debe recordarse que en los procedimientos, digamos clásicos, y aplicados hasta el presente, se asumía que las diferencias genéticas se debían a los niveles de producción total ya que se consideraba que no existían diferencias en cuanto a la forma de la curva de la lactancia. Dado que los resultados del MRA utilizado permiten estudiar este punto de vista, se decidió utilizar la metodología presentada en la sección de material y método. Se estimó el VG de los sementales para la producción de leche acumulada a 200 días de lactancia, cuyos resultados se presentaron en la figura 4. De este total se seleccionaron los 35 mejores sementales (10%) y se examinaron en forma minuciosa, los VG de los mismos sementales para la producción real mensual. Se pudo identificar importantes diferencias respecto al mérito genético de tales sementales cuando son evaluados en toda la trayectoria de la curva de lactancia. A modo ilustrativo se preparó la figura 5 con dos sementales de este grupo elite que manifiestan importantes diferencias genéticas respecto a la forma de la curva de lactancia.

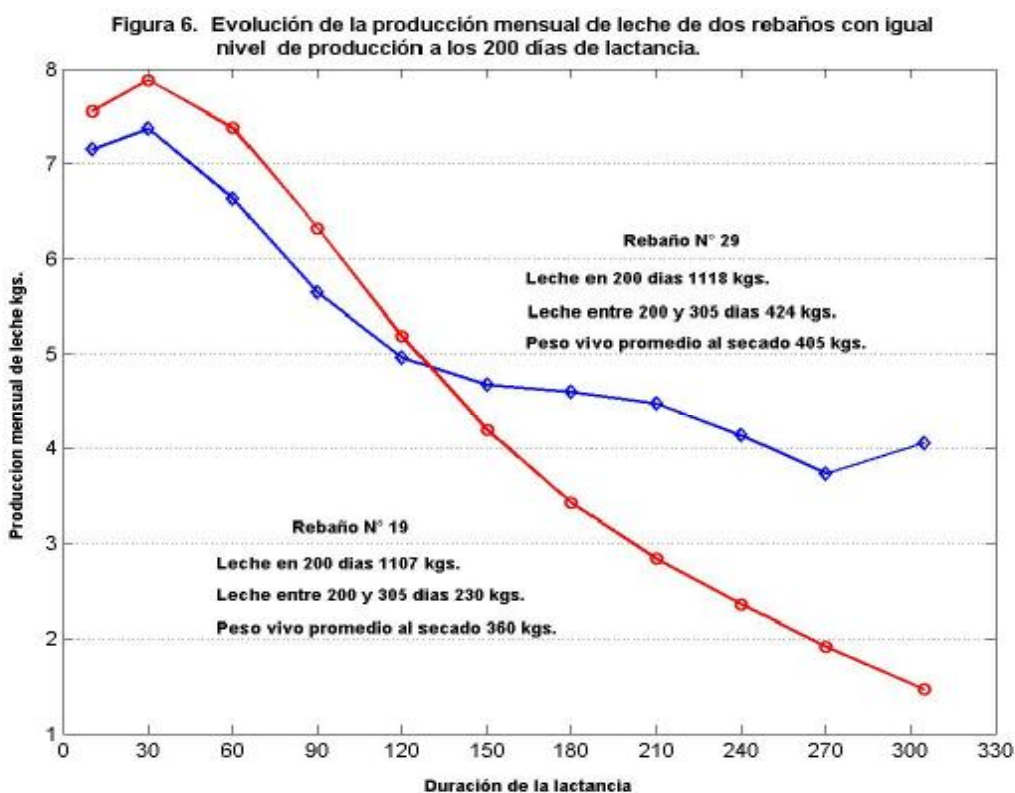


Ambos sementales presentan un $VG_{200} = +92$ kgs, sin embargo el semental N° 267 posee un componente genético según el cual la mayor parte de su potencial se manifiesta en la primera parte de la lactancia, por el contrario el semental N° 334 presenta un

comportamiento más estable en toda la lactancia. La respuesta de cuál es el mejor semental o el más recomendable requiere de más información y no está en el contexto de este estudio, aunque debe investigarse cuanto antes. En tal sentido, lo importante a resaltar es que este tipo de interrogante sólo puede plantearse cuando se utiliza un procedimiento estadístico de regresión aleatoria.

Finalmente, resulta obvio que los objetivos básicos de una empresa no sólo son conocer o contribuir a la mejor comprensión del conocimiento de la base genética de la producción de leche, sino incrementar los niveles productivos de sus rebaños. Para ello es determinante identificar a tiempo, aquellos factores que representan una restricción al incremento sostenido de la producción. Es este punto de vista precisamente, el que justifica la aplicación de un sistema de controles.

Para ser parcialmente consecuente con el planteamiento inicial del uso de la base de datos para el manejo general del rebaño, y no sólo para el trabajo de selección y mejora, se utilizaron los ficheros de salida de los modelos utilizados para estimar los parámetros genéticos. A partir de las medias ajustadas para los efectos de unidad de control, según los diferentes factores fijos y aleatorios incluidos en el modelo descrito, se identificaron dos rebaños (los números 19 y 29) cuyos niveles de producción de leche a 200 días de lactancia son muy similares. En la figura 6 se muestra la evolución de la producción mensual de ambas vaquerías.



Antes de comentar estos resultados es oportuno señalar que los valores en la figura 6, representan los niveles de producción de dichos rebaños, libres de cualquier diferencia genética, época del año o número de lactancia, ya que tales efectos se incluyeron en el modelo de análisis, de manera que se trata de problemas ligados a diferencias en el manejo

entre las dos unidades. Ambos rebaños manifiestan una curva de lactancia muy semejante en la primera etapa, sin embargo la unidad 19 presenta una depresiva tendencia en la segunda mitad de la lactancia mucho más marcada que la vaquería 29. En los primeros 100 días de lactancia el rebaño 19 produjo 691 kgs mientras que el 29 alcanzó 631 kgs (aproximadamente medio litro más por día para un rebaño de 100 vacas en ordeño), no obstante, esta superioridad fue compensada por la mayor persistencia de la vaquería 29. Aunque a los 200 días de lactancia ambas unidades produjeron un volumen muy similar, para una lactancia promedio de 305 días, la vaquería 29 logró producir 13.5% más de leche que la 19. ¿Qué factor de manejo puede explicar tales diferencias?

Aún cuando la base de datos empleada, ha sido la más completa de las disponibles en el país, no contiene toda la información necesaria para una respuesta adecuada a la interrogante anterior (ejemplo no está controlado los cambios de peso vivo durante la lactancia), no obstante es significativo el hecho de que la vaquería 29 tiene un peso vivo al momento del secado 12% superior a la unidad 19.

CONCLUSIONES.

De acuerdo a los resultados alcanzados se pueden hacer las siguientes conclusiones generales.

- El método de control lácteo mediante el registro del volumen de producción en un solo ordeño, pero de forma alterna, ha demostrado ser muy útil y preciso, por tanto puede emplearse en condiciones prácticas. Existe otro procedimiento aún más sencillo del cual se poseen evidencias de su efectividad y debe examinarse más adelante.
- En términos estadísticos, es posible indicar algún sesgo en los resultados ya que los factores de ajuste fueron obtenidos con la misma base de datos, no obstante, el alto nivel de precisión que se manifestó en las correlaciones entre los Valores Genéticos de los sementales, permite esperar que sus efectos serán mínimos.
- Otro elemento a destacar es que sólo se ha estudiado el volumen de leche, debiéndose llevar a cabo un trabajo similar para la composición de la leche.
- El procedimiento estadístico de Modelos de Regresión Aleatoria (MRA) ha demostrado sus ventajas, tanto en los estudios genéticos como para evaluar rebaños, de manera que se debe recomendar su uso. Es necesario mayor divulgación sobre este tipo de análisis que se usa internacionalmente como modelo de evaluación genética.
- El uso de MRA ha posibilitado demostrar que las diferencias genéticas no sólo se deben a los niveles de producción de los animales, sino también en la forma de la trayectoria de la curva de lactancia. Este nuevo componente puede producir beneficios en modo alguno despreciables.
- El estudio realizado no deja margen de dudas sobre la existencia de una importante variabilidad de origen genético, por tanto hay espacio para el trabajo de selección y mejora aún cuando los niveles productivos no son tan altos en la muestra analizada.

BIBLIOGRAFIA.

- Berry D.P., Olori V.E., Cromie A.R., Veerkamp R.F., Rath M. and P. Dillon, 2005.** Accuracy of predicting milk yield from alternative milk recording schemes. *Animal Science* 80: 53-60.
- Borroto, O. , 2002.** Informe de la Ganadería Vacuna. Ministerio de la Agricultura
- Guerra, D; Evora, J; De Bien, R; Prada, N., 2002.** Evaluaciones genéticas de la raza Siboney de Cuba. ISBN 959-7164-32-9. Disponible en: Web.<http://www.softcal.cu>.
- Fernández Chuarey, Lucía, 2004.** Modelos Estadísticos–Matemáticos en el análisis de la curva de lactancia y factores que la afectan en el genotipo Siboney de Cuba. Tesis presentada en opción al grado científico de doctor en Ciencias Veterinarias. La Habana, 99p.
- Gilmour, A.R.; Cullis, B.R.; Welham, S.J.; y Thompson, R., 2000.** ASREML Reference Manual. NSW Agric. Biom. Bull. NSW Agriculture, Locked Bag, Orange, NSW 2800, Australia.
- ICAR, 2004.** International Agreement of Recording Practices. Approved by General Assembly held in Sousse, Tunisia, June 2004. 457 pag.
- Jamrozik, J. y Schaeffer, L.R., 1997.** Estimates of genetic parameters for a test day model with random regression for production of first lactation. *J. Dairy Sci.* 80: 762-770.
- McDaniel B.T., 1969.** Accuracy of sampling procedures for estimating lactation yields: a review. *J. Dairy Sci* 52 (11) 1742 – 1761
- Menéndez Buxadera A., 1986.** Fundora Cordovés H. y Tarrero López R. Nuevo sistema para estimar la producción diaria de leche en hembras bajo control individual. Concurso ACPA, Nov., 1986, 15 pag.
- Menéndez Buxadera A., 2005.** Definición e importancia de la plasticidad fenotípica para la genética animal en el trópico. -ejemplo de aplicación en el ganado lechero de Cuba-; III Congreso Internacional de Mejora Animal, CIMA, Nov., 2005, 4 pag.
- Schaeffer, L.R., Jamrozik J., Van Dorp R., Kelton D.F. and D.W.Lazenby, 2000.** Estimating daily milk yield of cows from different milking schemes. *Liv. Prod. Sci.* 65: 219-227.
- Swalve, H.H., 2000.** Theoretical basis and computational methods for different test-day genetic evaluation methods. *J. Dairy Sci.* 83: 1115-1124.
- Wiggans G.R., 1981.** Methods to estimate milk or fat yield from AM and PM plans. *J. Dairy Sci* 64 (12) 1621–1624.

DEFINICIÓN E IMPORTANCIA DE LA PLASTICIDAD FENOTÍPICA PARA LA GENÉTICA ANIMAL .- APLICACIÓN EN EL GANADO LECHERO DE CUBA-

A. Menéndez Buxadera*
Centro de Control Pecuario
Ministerio de la Agricultura
Conill y Boyeros, C. de La Habana, Cuba

RESUMEN.

Un genotipo puede manifestar diferentes fenotipos en función de su capacidad de adaptación a las condiciones ambientales imperantes. La magnitud de esos cambios se denomina como la **'plasticidad'** o capacidad adaptativa de ese organismo para ese carácter. Tales cambios del valor fenotípico en función del ambiente representa una expresión de interacción genotipo-ambiente (I_{GA}) y es por eso precisamente que la **Plast** puede brindar una solución a la medición cuantitativa de la adaptación a ambientes heterógeneos. Según tales resultados se pueden clasificar los animales como **'estables o robustos'** si no manifiestan cambios en su comportamiento en toda la gama de condiciones ambientales, es decir animales que manifiestan una adecuada capacidad adaptativa, y como **'no estables o plásticos'** en caso contrario. Para examinar estos puntos de vista se realizó un pormenorizado análisis de las causas de variación de 226402 lactancias de vacas Holstein hijas de 519 sementales de esta raza que habian parido en las condiciones de toda Cuba. Este conjunto de datos se examinó mediante un modelo de regresión aleatoria, equivalente a un Modelo de Norma de Reacción (**RNM**) que permitió estimar la evolución de los parámetros genéticos y el valor genético (**VG**) de los sementales en todo el gradiente de condiciones ambientales representados en este estudio. Solo el 33 % de los mejores sementales seleccionados en ambiente alto (nivel de producción de leche del rebaño de 4800 kg) fueron los mejores en ambiente bajo (nivel de producción de leche del rebaño de 2400 kg) es decir existen importantes efectos de I_{GA} . El análisis en detalle del 5% de los mejores sementales permitieron identificar que 1 de cada 4.5 sementales elites pueden calificarse como **'estables o robustos'** con un patrón muy de alta capacidad adaptativa en toda la trayectoria de ambiente manifestando un **VG** = +243kg (nivel de producción del rebaño de 4800 kg) y **VG** = + 222 kg (nivel de producción del rebaño de 2400 kg), mientras que otros sementales clasificados como **'plásticos'** presentaron un **VG** = + 14 kg en ambiente bajo y **VG** = +477 kg en ambiente alto. Se discute la importancia de la **Plast** para la mejora genética animal en el trópico y se recomienda el uso de los modelos **RNM** como los de mayor aplicación práctica en condiciones ambientales de Cuba.

Palabras claves. Plasticidad fenotípica, interacción genotipo ambiente, modelo de norma de reacción, producción de leche, trópico.

***mail:** cencop@minag.gov.cu

Texto de la conferencia que será ofrecida en el III Congreso Internacional de Mejora Genética en el Trópico. Noviembre 5-6 del 2005, Ciudad Habana, Cuba.

INTRODUCCIÓN.

Todos estamos familiarizados con el hecho de que los organismos, ya sean vegetal como animal, manifiesten un adecuado nivel fenotípico cuando las condiciones ambientales son favorables, esto es **ambiente alto**. Por el contrario ese potencial productivo se reduce considerablemente en caso contrario, es decir en **ambiente bajo**. En cierta medida esto ha dado lugar a ciertas tendencias según las cuales el proceso de selección debe llevarse a cabo en **ambiente alto** donde las diferencias genéticas pueden manifestarse más ampliamente. Este punto de vista considera que es **'el genotipo'** el factor determinante, lo cual es contrario al principio básico de que es **'el ambiente'** el elemento que permite identificar los mejores y más adaptados animales. Al propio tiempo se obvia la correspondencia necesaria que debe existir entre **'el genotipo'** y **'el ambiente'**, en otras palabras se asume implícitamente que la Interacción Genotipo-Ambiente (I_{GA}) no es importante. Asumir que I_{GA} no sea importante no implica necesariamente que sus efectos no estén presentes y puedan sesgar los resultados esperados.

En términos teóricos, pocas veces reflexionamos en la naturaleza de este problema. Un genotipo puede manifestar diferentes fenotipos en función de su capacidad de adaptación a las condiciones ambientales imperantes. La magnitud de esos cambios se denomina como la **'plasticidad'** o capacidad adaptativa de ese organismo para ese carácter. El concepto de **'plasticidad'** (**Plast**) ha tenido mucha aplicación en el campo de la ecología y en la biología evolucionista (Via y col., 1995) y solo recientemente se comienza a introducir en la literatura de genética animal (de Jong y Bijma 2002). Tales cambios del valor fenotípico en función del ambiente representa una expresión de I_{GA} y es por eso precisamente que la **Plast** puede brindar una solución a la medición cuantitativa de la adaptación a ambientes heterogéneos. Según tales resultados se pueden clasificar los animales como **'estables o robustos'** si no manifiestan cambios en su comportamiento en toda la gama de condiciones ambientales, es decir animales que manifiestan una adecuada capacidad adaptativa, y como **'no estables o plásticos'** en caso contrario.

Hasta el presente estos problemas de I_{GA} se han interpretado en base al cambio de orden de mérito de los animales en diferentes ambientes. Este enfoque puede representar solamente una parte del problema, por tanto debe ser reemplazado por un procedimiento de mayor amplitud, mas robusto y menos sesgado que brevemente se muestra a continuación.

1. PROCEDIMIENTO CLÁSICO UTILIZADO HASTA EL PRESENTE.

Todo el trabajo de selección y mejora se basa en el análisis de las variaciones del valor fenotípico (**P**) de uno o varios caracteres de mayor interés económico, los cuales se consideran que están altamente relacionados con los objetivos de selección. Se asume que los valores de **P** son la expresión individual de una gran cantidad de genotipos (**G**) manifestados y registrados en una gran amplia gama de condiciones ambientales (**E**). Esta simple y esquemática descripción ha posibilitado confeccionar una ecuación general sobre el cual se fundamenta la selección artificial:

$$(A) \quad P = \mu + G + E$$

donde μ es la media general y los demás símbolos ya fueron definidos.

A partir de A se puede confeccionar un modelo estadístico de trabajo para datos de producción en condiciones prácticas:

$$(A1) \quad y_{ijk} = HYS_i + G_j + e_{ijk}$$

donde :

HYS es una combinación del rebaño-año-época donde se manifestó el valor fenotípico registrado (y_{ijk}) y cuya media general es equivalente a μ .

- G_j es el efecto debido al genotipo del animal.
- e_{ijk} es un efecto común a todas las observaciones o error residual aleatorio.

En A1 se considera **HYS** como efecto fijo (aunque también puede ser aleatorio), mientras que G_i y e_{ijk} son aleatorios. Este modelo permite obtener una predicción del Valor Genético (**VG**) de los animales que son la guía para la selección artificial. Al propio tiempo, se puede estimar la importancia relativa de la varianza genética-aditiva con respecto a la variabilidad total, es decir la heredabilidad (h^2) del carácter. Muchos trabajos existen donde se discuten las propiedades y mejores algoritmos para aplicar a sistemas de ecuaciones del tipo A1 a un gran volumen de datos y ha sido el modelo que más se ha empleado en los estudios de genética animal hasta el presente, con resultados satisfactorios. Sin embargo, la ciencia avanza muy rápidamente y existen nuevas interrogantes y resultados que deben tomarse en cuenta para una mayor eficiencia en el trabajo de selección.

De acuerdo a la teoría de la genética cuantitativa, el parámetro h^2 representa la magnitud de las diferencias entre los valores fenotípicos cuyo origen es genético-aditivo y solo es aplicable a la media de la población, periodo de tiempo y condiciones ambientales medias donde fueron estimados, mientras que el **VG** representa la desviación de **G** la cual está ponderada por h^2 . Sin embargo, aceptar un valor fijo para h^2 implica que tanto las variaciones genéticas como el estimado de **VG** no varían en toda la heterogeneidad de ambientes representados en la trayectoria de **HYS**, es decir se asume que es el mismo carácter en todos los niveles de ambientes representados en los datos.

Según lo expuesto, resultaría adecuado profundizar en esta falta de coherencia particularmente en los componentes del modelo A1, con vistas a estimar valores de h^2 y **VG** más adaptados a los programas de mejora genética. En tal sentido se han propuesto nuevos modelos estadísticos de trabajo, en los que tanto la h^2 como el **VG** de los animales se pueden estimar como función del ambiente, lo cual está en correspondencia con el origen y naturaleza de la variabilidad genética de los caracteres de interés económico. Resulta evidente que estos nuevos puntos de vista, serán bien recibidos por la comunidad científica de los países tropicales, donde pueden brindar resultados espectaculares. Estas regiones requieren urgentemente de mucho más evidencias sobre las bases y procedimientos para la selección de animales más adaptados a variaciones extremas de un clima caliente, húmedo y muy adverso y donde la **I_{GA}** es la regla y no la excepción.

2. NUEVOS PROCEDIMIENTOS PARA EVALUAR LA CAPACIDAD ADAPTATIVA DE LOS ANIMALES.

Existen dos modelos estadísticos para estudiar la variabilidad de **Plast**: character state model (**CSM**) y reaction norm model (**RNM**) -se ha mantenido su denominación original según la literatura internacional-.

Character State Model. El **CSM** se utiliza cuando los datos de un determinado carácter, se expresa en cada categoría de ambiente, cada uno de los cuales se denomina Character State (**CS**). La variabilidad en CS_i , donde i representa cierta clasificación de ambiente, tal como nivel de producción del rebaño o del país, puede estudiarse mediante **CSM** de la misma forma que un análisis multicaracteres, en el que se considera CS_i y CS_j como dos rasgos diferentes y la correlación (r_g) entre ambos es una primera estimación de **I_{GA}**. Si r_g es significativamente diferente de 1 quiere decir que el comportamiento del genotipo no es el mismo en diferentes ambientes, por lo tanto la **I_{GA}** es importante lo cual implica que hay variabilidad genética en **Plast**. Este es el principio clásico utilizado hasta el presente y ha sido la base subyacente de los trabajos sobre esta temática. En tal contexto y adaptado a dos niveles de ambientes, es posible interrogarse por la variabilidad de un nuevo carácter $CS_k = CS_j - CS_i$ y su varianza ($VarCS_k = VarCS_j + VarCS_i - 2CovCS_{ij}$) brindará una primera estimación de la magnitud de estas diferencias como manifestación de **Plast**.

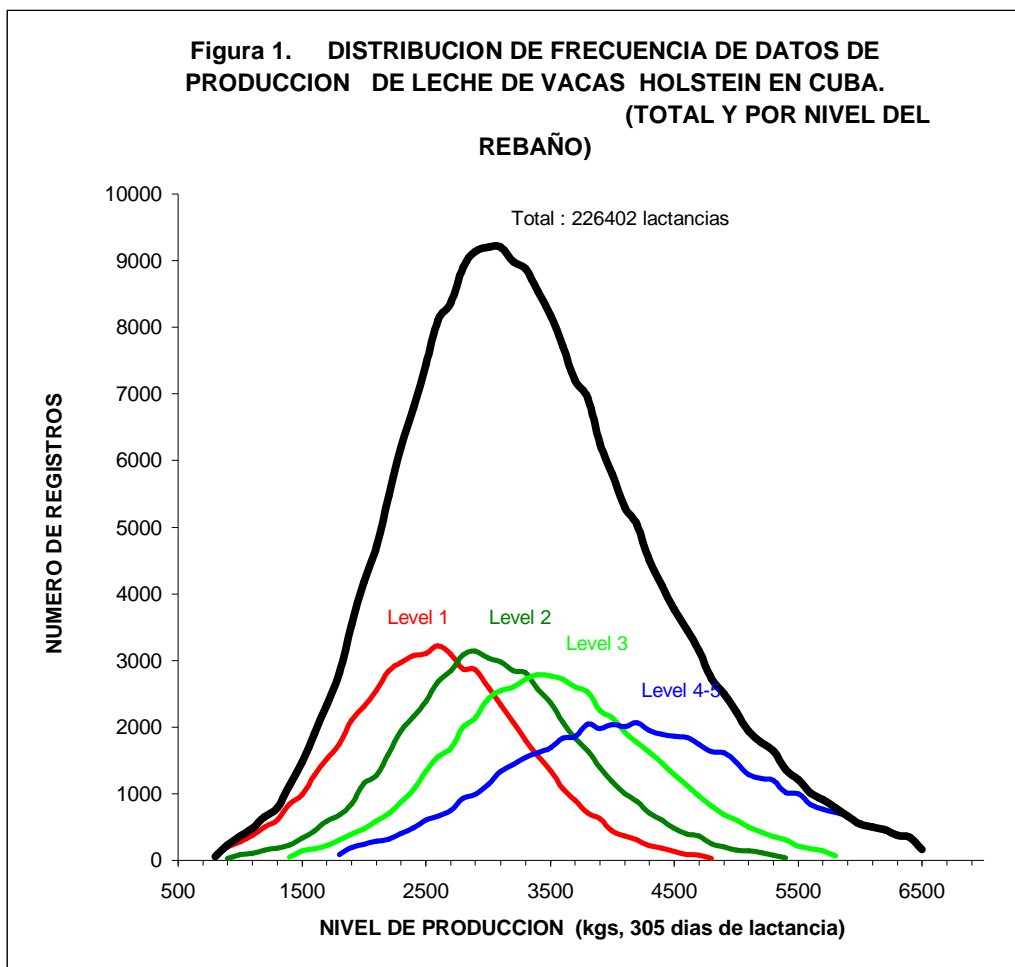
Reaction Norm Model. La norma de reacción (**RN**) de un genotipo se define como la capacidad de un organismo a manifestar un patrón propio de expresión fenotípica de acuerdo a los cambios sistemáticos de las condiciones ambientales y su magnitud puede cuantificarse por la regresión del fenotipo sobre el ambiente, una función estimable mediante **RNM**.

Existe mucha controversia respecto a la aplicación de **CSM** o **RNM**. El **CS** es la expresión fenotípica de un genotipo en cada ambiente (ambiente discreto), mientras que la **RN** significa el patrón de expresión total del mismo genotipo en todo el gradiente de ambientes (ambiente continuo) consecuentemente el **CS** es solo un punto en toda la curva de respuesta del genotipo. En términos generales se puede plantear que el **CS** es un caso particular de **RN** en el que el rasgo se ha expresado para unas condiciones ambientales dadas, de manera que bajo determinadas circunstancias los resultados de **CSM** y **RNM** pueden ser equivalentes.

Los parámetros genéticos así como el **VG** estimado por **CSM**, son aplicables a la población y nivel de ambiente en la cual fueron estimados, y su significación y grado de utilización es el mismo que para los resultados obtenidos por un modelo univariado o multivariado que se comentaron previamente. Por el contrario, el uso de **RNM** puede brindar los mismos estimados, pero en términos de una función con la cual puede estimarse la evolución de la variabilidad genética (y del **VG**) del carácter en toda la trayectoria de ambiente. Para cada semental los resultados de **RNM** proporciona el valor de una función: $f_g = \alpha + \beta * p(env)$, donde α es el intercepto o valor genético medio; β es la pendiente o valor genético para la norma de reacción o su equivalente de la I_{GA} , y depende también del coeficiente p de un polinomio, de orden k , para cada nivel de ambiente. De acuerdo a este punto de vista el procedimiento **RNM** permite estimar un **VG** para cada semental en función de las condiciones ambientales. A diferencia de **CSM**, el **RNM** posibilita hacer una discriminación más amplia y precisa de los genotipos según su capacidad de adaptación a diferentes condiciones ambientales. Es en este sentido precisamente que **RNM** pueden ser una herramienta de mucho potencial para los programas de mejora genética, particularmente para los sistemas de producción orgánicos o de bajos insumos que actualmente se llevan a cabo en países desarrollados, pero también y muy especialmente para las regiones tropicales.

3. APLICACIÓN PRÁCTICA CON DATOS DE CUBA.

Para examinar los puntos de vistas expuestos se decidió llevar a cabo un detallado análisis de las causas de variación de la producción de leche (305 días de lactancia) de 105953 vacas Holstein (226402 lactancias en total) hijas de 519 sementales de esta raza que habían parido en las condiciones de toda Cuba entre los años 1975 a 1989. La variabilidad fenotípica total se presenta en la figura 1 y este conjunto de datos fueron estudiada por un clásico modelo semental (**SSM**) del tipo A1. En esta propia figura se muestran la distribución de frecuencia de los mismos datos pero agrupados según 5 niveles de producción de leche del rebaño (los niveles 4 y 5 se presentan unidos). Esta división se basó en la media mínimo cuadrática del efecto rebaño (**LSh**) estimado por un modelo lineal mixto en el que se incluyeron además los efectos de la edad al parto de la vaca y se eliminaron las posibles diferencias genéticas al incorporar el efecto semental al modelo de análisis.



En términos generales es aceptado que el nivel de producción del rebaño es el reflejo del manejo, alimentación y atención que recibieron todos los animales del rebaño, consecuentemente el valor de LSh representa ‘el nivel de ambiente’ que recibieron esos animales. Excepto los dos extremos, se establecieron 600 kgs de diferencias (2400 kg a 4800 kg en 305 días de lactancia) entre cada nivel de producción y los valores fenotípicos de cada animal (vaca) fueron considerado como la expresión del carácter dentro del rango de las diferencias existentes intra cada nivel, es decir una manifestación de **CS**. Estos datos fueron estudiados por un modelo semental del tipo Character State Model (**CSM**). Finalmente los valores individuales de LSh se incorporaron como una covariable fija y aleatoria y se aplicó un modelo de regression aleatoria equivalente a Reaction Norm Model (**RNM**). Los efectos ambientales de carácter permanente debido a repeticiones de registros (más de una lactancia por vaca) del mismo animal fue incluido en los tres modelos. También la combinación de rebaño-año-epoca de parto, la edad al parto de cada vaca y la matriz de parentesco entre todos los sementales (607 sementales en total) se incorporaron a los tres modelos. En **CSM** y **RNM** se adicionó los efectos de varianza residual heterogénea. El software ASREML (Gilmour y col., 2000) fue utilizado en todos estos análisis.

Los tres modelos brindaron resultados muy similares (tabla 1) lo cual es una demostración de la consistencia y propiedades de estos procedimientos. Los valores puntuales de h^2 en este estudio son ligeramente inferiores a una revisión bibliográfica publicada recientemente, en 52 artículos sobre este rasgo en países tropicales (Lobo y col., 2000). Debe señalarse que en esa propia revisión se plantea que la mayor parte de estos trabajos pueden estar sujetos a crítica debido a poco número de observaciones y el uso de procedimientos estadísticos poco robustos. Dado que estos elementos no

son aplicables a este estudio, estos resultados pueden catalogarse como aceptables y representativos de la h^2 de la producción de leche en las condiciones de Cuba.

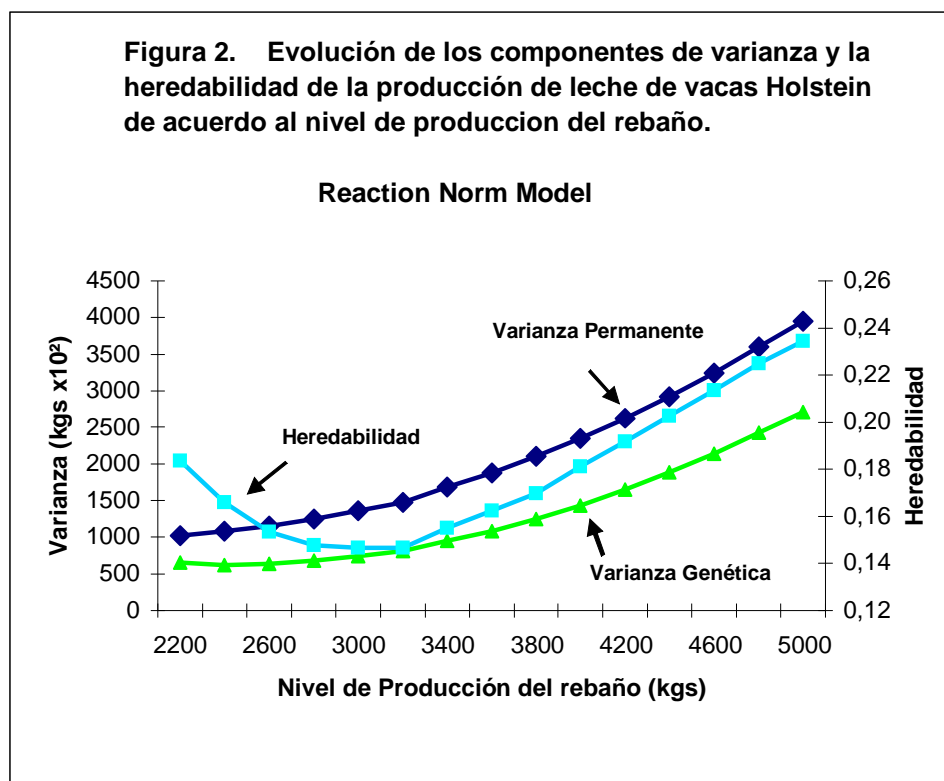
Tabla 1. Componentes de varianza y heredabilidad de la producción de leche en vacas Holstein en las condiciones tropicales de Cuba estimada por tres modelos^a.

Modelos	Varianza Genética	Varianza efectos Permanentes	Varianza Residual	Heredabilidad
SSM	90890	204900	388155	0.153 \pm 0.01
CSM	94270	207400	353816	0.161 \pm 0.02
RNM	94660	191745	419263	0.155 \pm 0.01

a. Se utilizaron un total de 226402 registros de 105953 vacas Holstein hijas de 519 sementales.

Los parámetros mostrados en la tabla 1 presentan el valor estimado de h^2 de un mismo problema biológico aunque con el significado y aplicación práctica que se señaló previamente. Observe la similitud entre los valores de h^2 en SSM y RNM ya que ambos valores son aplicables a la población total representada en la figura 1 (modelo SSM) mientras que los valores de RNM se refieren específicamente a un nivel de ambiente (es decir nivel de producción del rebaño) similar a la media de esa población. La ligera diferencia para los resultados de CSM se deben a que en este caso se refiere a una clase de ambiente (3000 a 3600 kg) dentro del rango de la cual esta situada la media de la población.

Independientemente de lo planteado, esta forma de expresión no ofrece la respuesta acerca de la posible evolución de la varianza genética y el Valor Genético en función del ambiente, lo cual fue llevado a cabo con los resultados de RNM y se presentan en la figura 2. Los resultados muestran una curva de respuesta homogénea en la cual la magnitud de la varianza genética y la heredabilidad del carácter aumentan, en forma absoluta y relativa, en todo el gradiente de ambiente representado en este estudio por el nivel de producción del rebaño.



Es importante resaltar los valores de $h^2 = 0.225_{\pm 0.02}$ cuando los animales se encuentran en un nivel de ambiente alto (4800 kgs.) lo cual es aproximadamente 50 % superiores a los estimados para la media de la población. En términos prácticos esto quiere decir que la selección puede resultar mucho más efectiva y precisa cuando se lleva a cabo en condiciones ambientales favorables o muy favorables. Este criterio asume implícitamente que los mejores sementales en ‘ambiente alto’ continúan siendo los mejores en ‘ambiente bajo’, en otras palabras que la Interacción Genotipo-Ambiente carece de importancia.

El procedimiento de **RNM** permite estimar las correlaciones genéticas (r_g) entre el denominado el mismo carácter en los diferentes niveles de producción y estos resultados se presentan en la tabla 2. El análisis de esta tabla indica que todas las r_g fueron positivas, esto quiere decir que cualquiera sea el nivel de producción donde la selección se lleve a cabo habrá una respuesta positiva en todos los ambientes. Entre niveles de ambientes adyacentes o muy próximos todas las r_g fueron mayores de 0,96, no obstante este alto grado de relación decrece en la medida que aumentan las diferencias entre niveles de producción. Cuando tales diferencias de ambiente son del orden de 1200 kg o más (dos desviaciones estándar en este estudio) los valores de r_g fueron inferiores a 0,85, de lo cual se infiere que existiran cambios en el orden de merito de los sementales seleccionados, en otras palabras los mejores sementales en ‘ambiente alto’ no son los mejores en ‘ambiente bajo’. Estos resultados significan que la importancia de I_{GA} solo puede cuantificarse cuando las diferencias entre los niveles de ambiente son muy contrastantes.

Tabla 3. Correlación genética de la producción de leche en función del nivel de producción del rebaño.

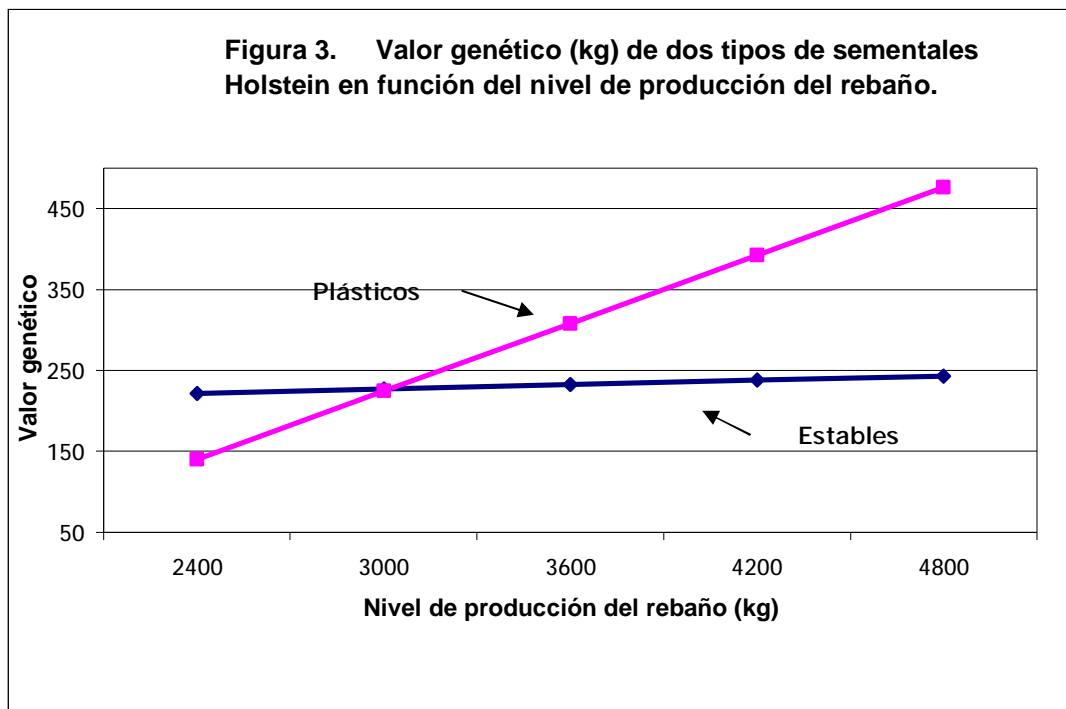
	NIVEL DE PRODUCCION DEL REBAÑO -kg-											
	2600	2800	3000	3200	3400	3600	3800	4000	4200	4400	4600	4800
2400	0,990	0,962	0,923	0,885	0,823	0,772	0,722	0,676	0,634	0,594	0,560	0,527
2600		0,991	0,968	0,942	0,895	0,854	0,813	0,773	0,736	0,702	0,671	0,642
2800			0,993	0,978	0,946	0,915	0,883	0,851	0,820	0,790	0,764	0,738
3000				0,996	0,979	0,957	0,933	0,908	0,883	0,859	0,836	0,814
3200					0,993	0,979	0,961	0,941	0,921	0,900	0,881	0,862
3400						0,996	0,987	0,975	0,961	0,946	0,931	0,917
3600							0,997	0,990	0,981	0,970	0,959	0,947
3800								0,998	0,993	0,986	0,977	0,969
4000									0,998	0,994	0,989	0,983
4200										0,999	0,996	0,992
4400											0,999	0,997
4600												0,999

Para ejemplificar las perspectivas del procedimiento **RNM** en términos de capacidad de discriminación entre los genotipos respecto a su capacidad de adaptación a la heterogeneidad de ambientes, se examinaron los resultados del **VG** de los 607 sementales representados en este estudio. Se decidió aplicar una intensidad de selección del orden del 5%, correspondiente a la categoría de sementales élites. Dos estrategias se compararon:

- Seleccionar los mejores 30 sementales según su **VG** en ‘ambiente alto’ (nivel de producción del rebaño de 4800 kgs).
- Seleccionar los mejores 30 sementales según su **VG** en ‘ambiente bajo’ (nivel de producción del rebaño de 2400 kgs).

Cuando se unieron estos dos conjuntos solamente 12 sementales se mantenían como los mejores en ambos ambientes, es decir el 33 %, lo cual es una clara evidencia de la importancia de la I_{GA} . El análisis pormenorizado del **VG** de cada semental de ambos grupos, en toda la trayectoria de ambientes permitió identificar dos patrones de comportamiento muy diferentes, que sirven para

ilustrar la importancia del concepto Plasticidad para la mejora genética de la producción lechera en las condiciones tropicales de Cuba. Estos resultados se presentan esquemáticamente en la figura 3.



Del conjunto total de sementales seleccionados se pudo identificar 5 sementales que manifiestan una gran sensibilidad a los cambios en las condiciones ambientales representadas por el nivel de producción del rebaño, estos se corresponden con la categoría de genotipos '**plásticos**'. Por el contrario, se encontraron otro grupo de sementales con un patrón de respuesta prácticamente lineal, es decir un comportamiento en el que no se presentan variaciones en toda la heterogeneidad de ambientes y estos genotipos se denominan '**estables o robustos**'. Hay que reiterar que la figura 3 se confeccionó a partir de los **VG** de los sementales evaluados en este estudio, esto quiere decir estimados obtenidos de acuerdo a un modelo estadístico que incluyen todas las causas de variación que afectan el carácter y la matriz de parentesco, en otras palabras, se trata de estimados insesgados según la denominación más usual para estos estudios. Para ejemplificar aún más las diferencias entre ambos tipos de sementales tomese en cuenta que los valores promedios de *beta* fueron de +199 kg y +13 kg para los sementales **plásticos** y **robustos** respectivamente. Tal y como se señaló previamente, este indicador *beta* es la cuantificación de las diferencias entre estos genotipos en la norma de reacción de cada uno en toda la trayectoria de cambios en las condiciones ambientales donde se explotaban sus hijas. Tales diferencias son de origen genético y están relacionadas con la capacidad de adaptación a ambientes tan heterogeneos como los que están presentes en la base de datos utilizadas en este estudio. De acuerdo a estos resultados es posible interrogarse... cual es el tipo de semental que debe recomendarse?

Para responder a esta pregunta hay que tomar en cuenta las características propias del rasgo así como del sistema de producción animal imperante.

1. La producción láctea es un carácter ligado al sexo y se manifiesta un período de tiempo posterior a la decisión de la selección de los sementales y la realización de las inseminaciones correspondientes. En buenas condiciones de manejo y alimentación, este puede ser de alrededor de 3 años, aunque en la actualidad de la ganadería del trópico se empleará no menos de 5 años. Dicho en otras palabras se trata de una inversión realizada en el año **n** y cuyos efectos se materializará **n+5** años posteriores.

2. Si el sistema de producción permite planificar los recursos necesarios para garantizar los insumos que se requieren para alcanzar altos niveles de producción, entonces los sementales que deben seleccionarse deben ser los **'plásticos'**, ya que manifiestan una tendencia de respuesta muy positiva. Esto es la práctica normal que se lleva a cabo en países desarrollados.
3. En condiciones tropicales en general, esto es muy difícil de garantizar ya que la situación económica es muy variable y dependen de circunstancias de un mercado internacional cada vez más competitivo y agresivo. En estas condiciones los sementales que deben seleccionarse son los **'estables o robustos'**, cuyo nivel de riesgo es muy bajo cualquiera sean los recursos disponibles.

Nótese la envergadura que puede implicar la variante de seleccionar sementales atendiendo solamente por sus resultados en **'ambiente alto'** es decir genotipos **'plásticos'** ya que si fallan las premisas (lo que se denominó como paralelismo entre **'el genotipo'** y **'el ambiente'**) los resultados serán totalmente contrarios cuando sus progenies se exploten en **'ambiente bajo'**.

La forma en que se han representado estos resultados puede transmitir la impresión de que este tipo de sementales **'estables o robustos'** son muy poco frecuentes, no obstante la realidad es otra. De los dos grupos de 30 sementales seleccionados en **'ambiente alto'** y **'ambiente bajo'** en realidad se seleccionaron 36 sementales en lugar de 60 ya que 12 fueron mejoradores en ambas condiciones. De este conjunto se pudieron identificar 8 sementales (22.2%, es decir aproximadamente 1 de cada 4.5 sementales seleccionados) que se califican como **'estables o robustos'** lo cual quiere decir que existe margen suficiente para llevar a cabo un proceso de selección más acorde a las condiciones existentes.

En términos de procedimiento este modelo **RNM** puede resultar algo más engorroso que el método clásico actualmente en uso (tipo modelo A1) que brinda un valor fijo para el **VG** de cada semental. Son incuestionables las ventajas que puede representar el uso de **RNM** ya que permite una mayor grado de discriminación entre los sementales evaluados disponibles y permite seleccionar aquellos genotipos más adaptados a las diversas condiciones ambientales. Resulta evidente que este debe ser el modelo recomendado para condiciones ambientales del trópico. Debe añadirse que en esta oportunidad se ha elegido 'el nivel de producción del rebaño' como indicador de 'ambiente', no obstante se pueden emplear otro tipo de variables (temperatura ambiente, alimentación disponible etc) aunque el significado será el mismo.

4 . REFERENCIAS

- de Jong, G., & Bijma, P. Selection and phenotypic plasticity in evolutionary biology and animal breeding. *Livestock Production Science* 78:3 (2002) 195-214.
- Gilmour, A.R.; Cullis, B.R.; Welham, S.J.; y Thompson, R., 2000. *ASREML Reference Manual*. NSW Agric. Biom. Bull. NSW Agriculture, Locked Bag, Orange, NSW 2800, Australia.
- Lobo, R.N.B., Madalena, F.E. & Viera, A.R. Average estimates of genetic parameters for beef and dairy cattle in tropical regions. *Animal BreedIng Abstracts* 68:6 (2000) 433-462.
- Via, S., Gomulkiewicz, R., de Jong, G., Scheiner, S.M., Schlichting, C.D. & van Tienderen P.H. Adaptive phenotypic plasticity: consensus and controversy. *Trends in Ecology Evolutive* 10: (1995) 212-217.

