



CIMAT

Centro de Investigación en Matemáticas A.C.

***Análisis comparativo del
crecimiento de dos tipos de planta
de agave empleando algunas
técnicas de estadística
multivariada***

T E S I S

que para obtener el grado de

Maestría en Ciencias con Especialidad en
Probabilidad y Estadística

presenta:

Esteban Bracamontes Orozco

Director de Tesis:
Dr. Fernando Ávila Murillo

Guanajuato, Gto., México, Febrero del 2006.

A mis padres:

Esteban y Martha,
*por su amor y confianza brindado
a lo largo de toda mi vida...*

A mis hermanos:

Sergio y Jessica,
*porque han sido el motor que me
motiva a seguir día a día...*

A mi tío:

José (Pepe),
*por su apoyo incondicional
brindado a cada momento...*

Agradecimientos

En este espacio quiero agradecer a todas aquellas personas e instituciones que de manera directa e indirecta hicieron posible el desarrollo y culminación de este trabajo.

Comienzo agradeciendo al CIMAT, por los recursos económicos, tecnológicos y humanos proporcionados durante la realización de mis estudios de maestría. Así mismo, agradezco al CONACyT, por la beca (con número de registro 176188) que me fue concedida para llevar a cabo dichos estudios en este mismo Centro. De igual forma, quiero agradecer al CONCyTEG, por el apoyo económico facilitado durante el desarrollo de mi trabajo de tesis. También quiero dar gracias a la empresa *Tequila Sauza*, por el apoyo económico proporcionado para la conclusión de este proyecto.

Por otro lado, quiero agradecer de manera muy especial a mi asesor y sinodales: Dr. Fernando Ávila Murillo, Dra. Eloísa Díaz-Francés Murguía, Dr. Miguel Nakamura Savoy y Dr. José Ignacio del Real Laborde, por toda su ayuda y los valiosos consejos brindados para el enriquecimiento de este trabajo.

También agradezco de manera especial a mi madre, Martha Orozco Duarte, quien siempre me ha brindado su amor y apoyo incondicional para alcanzar todas mis metas. Al igual a mi padre, Esteban Bracamontes Ramírez, quien a pesar de ya no estar con nosotros, ha sido pieza clave en mi formación profesional y siempre llevaré en mi corazón. De igual forma, quiero agradecer a mi tío José Orozco Duarte, por el apoyo que tantas veces me ha brindado en diversos aspectos, y por estar siempre al pendiente de lo que pudiese ofrecerse. Agradezco también a mis hermanos, Sergio y Jessica, quienes siempre me animaron a seguir adelante y seguirle echando ganas.

No quiero dejar de mencionar a personas que de una forma u otra pusieron su granito de arena durante el desarrollo y culminación de este proyecto, entre éstas se encuentran: Dr. Rogelio Ramos Quiroga, Dra. Graciela González Farías, Dra. Ivonne Solís Trapala, M. en C. Daniel Fernández Martínez, M. en C. María del Pilar Arcos Gamboa, M. en C. Mónica Gladis Pérez Miranda, M. en C. Claudia Nila Luévano, M. en C. José Arturo Montoya Laos, Lic. Henry Gaspar Pantí Trejo, Glenda Patricia Guevara Hernández, Gabriela Rocío Guevara Hernández, y Eduardo Antonio Solís Díaz, a quienes agradezco sinceramente su apoyo.

Quiero agradecer también a todos mis profesores, quienes con sus enseñanzas y consejos fueron parte importante en mi formación durante la maestría. Aprovecho también para agradecer a todas aquellas personas que laboran en el CIMAT, y a las que tantas veces les di un poco de lata con trámites de documentos, expedición de fotocopias, trámites de becas, etc., en especial agradezco a María Dolores Aguilera Mújica (Lolita), Berenice Segovia Puga, Jannet Vega Gutiérrez, Oscar Adolfo Sánchez Valenzuela y Sandra Jasso Ibarra; así como al personal del Staff de Cómputo y Redes (en especial Alejandra, Israel y Chuche), que siempre atendieron oportunamente a los llamados de auxilio; al personal de la Biblioteca (en especial Manuel y Alicia), quienes siempre me dieron facilidades para la adquisición de los libros necesarios para cumplir con mis deberes; agradezco también al personal que labora en la Dirección Administrativa, Eventos, Vigilancia, Limpieza, etc., que siempre tuvieron la mejor de las atenciones para conmigo.

Finalmente, agradezco a Dios, por darme la salud y fortaleza necesaria para la culminación satisfactoria de esta etapa de mi vida.

Índice

Agradecimientos	vii
1. Introducción	1
2. Generalidades sobre el agave	5
2.1 <i>Producción del tequila</i>	5
2.2 <i>Tipos de agave micropropagado y convencional</i>	7
2.3 <i>Problema de interés para la empresa</i>	7
3. Análisis de crecimiento para plantas de agave	9
3.1 <i>Descripción del crecimiento de una planta</i>	9
3.1.1 <i>Análisis de crecimiento para una planta en general</i>	9
3.1.2 <i>Análisis de crecimiento para una planta de agave</i>	11
3.2 <i>El índice morfológico de una planta de agave</i>	12
3.2.1 <i>Análisis de Componentes Principales</i>	14
3.2.1.1 <i>Introducción</i>	14
3.2.1.2 <i>Obtención analítica de los componentes principales</i>	14
3.2.1.3 <i>Componentes principales empleando matrices de correlación</i>	16
3.2.2 <i>Construcción del índice morfológico de una planta de agave</i>	18
3.3 <i>Modelos de crecimiento</i>	19
3.3.1 <i>Modelo Exponencial</i>	20
3.3.2 <i>Modelo Logístico</i>	21
3.3.3 <i>Modelo Gompertz</i>	22
3.4 <i>Algunas metodologías propuestas para la comparación del crecimiento entre dos poblaciones de plantas de agave</i>	23
3.4.1 <i>Comparación empleando la tasa de crecimiento</i>	24
3.4.1.1 <i>Tasa de crecimiento temporal</i>	24
3.4.1.2 <i>Tasa de crecimiento relativa promedio</i>	27
3.4.2 <i>Comparación empleando curvas de crecimiento</i>	29
3.4.3 <i>Modelos más generales</i>	32

4. Análisis exploratorio de los datos disponibles	35
4.1 Descripción de los datos.....	35
4.2 Análisis exploratorio (Pihuamo, Tepatitlán; Periodo: Enero-Agosto del 2004)	37
4.3 Análisis exploratorio (Pihuamo; Periodo: Enero 2004-Septiembre 2005).....	46
4.4 Diagnóstico de los datos.....	51
4.4.1 Conclusiones.....	51
4.4.2 Sugerencias y recomendaciones.....	53
4.5 Algunas conclusiones preliminares derivadas a partir de un estudio piloto de crecimiento con base en los datos disponibles.....	53
5. Comparación del crecimiento de dos tipos de planta empleado una simulación de datos	61
5.1 Introducción.....	61
5.2 Obtención de los datos simulados.....	62
5.3 Escenario 1: Poblaciones con crecimientos similares.....	66
5.3.1 Análisis exploratorio de los datos.....	66
5.3.2 Análisis de crecimiento.....	70
5.3.2.1 Comparación empleando la tasa de crecimiento.....	85
5.3.2.1.1 Tasa de crecimiento temporal.....	85
5.3.2.1.2 Tasa de crecimiento relativa promedio.....	88
5.3.2.2 Comparación empleando curvas de crecimiento.....	91
5.3.3 Conclusiones.....	97
5.4 Escenario 2: Poblaciones con crecimientos diferentes.....	98
5.4.1 Análisis exploratorio de los datos.....	98
5.4.2 Análisis de crecimiento.....	105
5.4.2.1 Comparación empleando la tasa de crecimiento.....	114
5.4.2.1.1 Tasa de crecimiento temporal.....	114
5.4.2.1.2 Tasa de crecimiento relativa promedio.....	117
5.4.2.2 Comparación empleando curvas de crecimiento.....	120
5.4.3 Conclusiones.....	126
6. Conclusiones y sugerencias	129
6.1 Resumen del trabajo.....	129
6.2 Sugerencias y recomendaciones.....	130
6.3 Directrices para nuevos proyectos.....	131
6.4 Estudios futuros.....	132

Apéndice	135
<i>A.1 Tablas</i>	135
<i>A.2 Gráficas adicionales</i>	146
<i>A.3 Programas</i>	157
<i>A.4 Análisis Discriminante</i>	159
 Bibliografía y Referencias	 163

Capítulo 1

Introducción

El *tequila* es la bebida mexicana mejor conocida a nivel internacional gozando actualmente de muy alta estima por su sabor y calidad. Es producida por docenas de empresas, algunas con una gran trayectoria a lo largo de muchos años.

El tequila es elaborado a partir de los azúcares extraídos de la planta de *agave*, Tequilana Weber en su variedad azul, que de acuerdo a la Norma Mexicana, es el único autorizado para su elaboración.

A diferencia de otras plantas como la vid (planta de la cual se utiliza su fruto, la uva, para la elaboración de vinos), la planta de agave aún no ha sido estudiada a fondo para conocer en detalle como es su crecimiento a lo largo del tiempo, sobretodo en sus primeros años de vida. La empresa *Tequila Sauza* está desarrollando, con la colaboración del *CIMAT*, una serie de proyectos de investigación relacionados con el desarrollo del agave, con el objetivo de un mejor aprovechamiento de los recursos involucrados en los procesos de producción de agave en el campo.

Un aspecto que en la actualidad es de interés para dicha empresa es comparar el desarrollo entre los dos tipos de plantas de agave que son utilizados para elaborar sus productos. Los dos tipos son referidos como *micropropagado* y *convencional*, donde su principal diferencia radica en la manera en que se obtienen las plantas de cada tipo. Para estos dos tipos de planta se ha observado, de manera informal, que su desarrollo en el campo no es el mismo, pero, sin embargo, a la fecha no se han realizado estudios comparativos sustentados científicamente.

Como es costumbre en la Estadística, podemos empezar con el planteamiento de una hipótesis (*nula*) sobre la igualdad en el desarrollo de dos poblaciones de plantas (las del tipo micropropagado y las del tipo convencional). El objetivo central del presente trabajo es precisamente encontrar evidencia a favor o en contra de dicha hipótesis, a partir de la información contenida en mediciones tomadas periódicamente en una muestra seleccionada de plantas de los dos tipos.

Para describir el desarrollo en las plantas nos apoyaremos, como una de las hipótesis de trabajo, del hecho de que el *crecimiento* de éstas es precisamente un indicador de su desarrollo, por lo que será necesario analizar variables que cuantifiquen tal crecimiento a lo largo del tiempo. Las variables consideradas durante el presente trabajo serán: *Altura de la planta*, *Número de hojas desprendidas* y *Longitud de hoja*, las cuales representan características físicas de las plantas en lo que respecta a su morfología y pueden considerarse como indicadoras de crecimiento según lo descrito por C. L. Beadle en Coombs et al., (1985). Sobre esto se hablará un poco más a detalle en el capítulo 3.

Entre las herramientas estadísticas que se emplearán para los análisis presentados en este trabajo destacan: análisis discriminante, análisis de componentes principales, pruebas de hipótesis para tasas de crecimiento, ajustes de modelos no-lineales y simulación.

El trabajo se encuentra estructurado de la siguiente manera:

El capítulo 2 presenta algunos conceptos generales sobre las plantas de agave, iniciando con una sección donde se dará una descripción general sobre el proceso que siguen éstas hasta la obtención del tequila. Así mismo, en este capítulo se cuenta con una sección donde se habla sobre los dos tipos de agave, *micropropagado* y *convencional*, utilizados en la actualidad por la empresa *Tequila Sauza* para la elaboración de su tequila, y sobre los cuales interesa contrastar su crecimiento como función del tiempo. Finalmente en la última sección se describe con más detalle el problema de interés planteado por la empresa, el cual como ya se ha mencionado consiste a grandes rasgos en comparar el desarrollo entre dos poblaciones de plantas de agave.

En el capítulo 3 se presenta una descripción sobre las metodologías propuestas para la modelación y comparación del crecimiento de las plantas de agave, iniciando en la primera sección con una breve descripción sobre los principios básicos considerados para un análisis de crecimiento de plantas

desde el contexto biológico, así mismo se presenta una descripción sobre los aspectos a considerar para el análisis de crecimiento de plantas de agave. En la segunda sección se habla sobre el *índice morfológico* de una planta de agave y su construcción, ya que dicho índice será una de las herramientas principales para cuantificar el crecimiento de las plantas de agave, con el cual se busca resumir la información presente en las variables consideradas en este trabajo a medirse en cada una de las plantas de la muestra elegida. La tercera sección presenta algunos de los modelos de crecimiento más comunes, donde uno de éstos (el *modelo exponencial*) será empleado para el ajuste de las curvas de crecimiento que serán utilizadas para contrastar el crecimiento de ambos tipos de planta. Finalmente, en la última sección se presentan las metodologías que son propuestas en este trabajo para comparar el crecimiento de ambos tipos de planta.

En el capítulo 4 se presentan algunos análisis de carácter exploratorio, realizados con base en un conjunto de datos proporcionado por la empresa. Con dichos análisis se busca juzgar la calidad informativa de los datos referidos, y evaluar así si éstos son útiles para responder a la interrogante de interés para la empresa. A raíz de lo observado en estos análisis se proporciona un diagnóstico de los datos, dando en principio algunas conclusiones, y posteriormente algunas sugerencias y recomendaciones orientadas en el diseño de nuevos experimentos donde se busque continuar con el estudio del crecimiento de plantas de agave. Una conclusión básica a la que se llega de estos análisis es que no es recomendable emplear los datos disponibles para evaluar de manera contundente la hipótesis nula, ya que existen algunos factores que no fueron controlados en el diseño original que repercutieron significativamente en el comportamiento de los datos obtenidos. No obstante, con base en lo observado de los análisis, y tomando en consideración la presencia de estos factores, se proporcionan algunas conclusiones preliminares sobre el crecimiento de ambos tipos de planta. Así mismo, son presentados algunos de los resultados obtenidos al considerar a los datos disponibles en un estudio piloto de crecimiento, a partir del cual sea posible establecer posibles directrices encaminadas al planteamiento de nuevos proyectos donde se continúe con el análisis del crecimiento de estos dos tipos de planta.

Con el objetivo de ilustrar la utilidad de las metodologías propuestas en el capítulo 4, en el capítulo 5 se presenta un estudio de simulación considerando dos diferentes escenarios: uno en donde se tiene igualdad de crecimiento entre ambos tipos de planta, y otro donde el crecimiento es diferente. Las

simulaciones de estos conjuntos de datos se llevaron a cabo tomando como base características iniciales presentes en las plantas de agave de Pihuamo y Tepatitlán, buscando de esta forma que dichos datos se encuentren apegados, en medida de lo posible, a lo que ocurre en la realidad con las plantas de agave en el campo.

Para finalizar, en el capítulo 6 se presenta en primer lugar una recapitulación sobre lo realizado a lo largo de este trabajo. Así mismo, se dan algunas sugerencias y recomendaciones generales orientadas al diseño de futuros experimentos donde se continúe analizando el crecimiento de las plantas de agave. Finalmente, se presentan un par de secciones más donde se discute por un lado sobre posibles directrices encaminadas al planteamiento de futuros proyectos, y por otro sobre algunas posibilidades para futuros estudios, en ambos casos encaminados también en la continuación del análisis de crecimiento de plantas de agave.

Capítulo 2

Generalidades sobre el agave

En este capítulo se presentan algunos conceptos generales involucrados en los procesos de producción de agave. La primera sección comienza dando un panorama general sobre la manera en que se obtiene el tequila, así como el papel que la planta de agave juega en dicho proceso. Por su parte, la segunda sección habla sobre los dos tipos de agave, *micropropagado* y *convencional*, que son utilizados por la empresa *Tequila Sauza* para la elaboración de dicha bebida y de los cuales interesa contrastar su crecimiento en este trabajo. Finalmente, en la última sección se describe el problema de interés para la empresa, el cual es la principal motivación de los análisis y metodologías presentados en capítulos posteriores.

2.1 Producción del tequila.

La planta de agave es una de las materias primas principales para la elaboración del tequila. Pertenece a la familia de las amarilidáceas y entre sus características físicas destacan sus hojas largas y puntiagudas de un color verde azulado.

Existen varias especies de agave cuyo jugo puede ser fermentado y destilado para la producción de bebidas alcohólicas, sin embargo, sólo el *agave tequilana Weber* en su variedad azul es el autorizado para la elaboración del tequila.

El cultivo de esta planta requiere de la combinación de numerosos factores como son: altitud del terreno, condiciones de tierra, pluviosidad, temperatura constante, clima semiseco, etc. Los métodos más comunes para su cultivo básicamente son dos: uno por medio de *semilla* (o bulbillo), y el otro por

rizoma. El cultivo por medio de semilla es el método tradicional en la agricultura, el cual consiste como su nombre lo indica en sembrar semillas de plantas de agave una vez que la tierra es debidamente preparada para dicho proceso. Este método en la actualidad no es utilizado por los productores de agave debido al largo periodo de tiempo requerido para la obtención de las semillas necesarias para la generación de nuevas plantas.

El método de rizoma consiste en transplantar los hijuelos que brotan de rizomas de la planta (una planta usualmente proporciona de uno a dos hijuelos por año a partir de su segundo/tercer año de plantación y a dicha planta se le denomina *planta madre*), al alcanzar el hijuelo una altura aproximada de 50 cm y cuando el corazón de éste es del tamaño de una toronja se desprende de la planta madre cortándolo por medio de un barretón. La edad óptima del agave para reproducirse mediante hijuelos es a los 3 años, y la plantación de los hijuelos en el campo se lleva a cabo antes de la temporada de lluvias, ya que la planta debe ser asentada y enterrada en 75% de su volumen apisonando la tierra para asegurar la planta.

Así empieza el largo proceso de evolución del agave, tardando de 7 a 10 años hasta terminar su ciclo de vida, que es cuando la planta florea (emisión del quiote). Durante el crecimiento de la planta se van realizando algunas labores que ayudan a ésta a producir y conservar el máximo nivel de almidones en el corazón. Una vez que el agave ha llegado al final de su ciclo de vida se lleva a cabo la cosecha mediante lo que es conocido como la *jima*, ya que en la elaboración del tequila se utiliza únicamente la parte central (corazón, piña o cabeza) de la planta, lugar donde se concentra la mayor cantidad de azúcares.

Finalmente, las piñas son transportadas a la fábrica, donde éstas se someten a un proceso industrial bajo estrictos controles de calidad para la obtención del tequila. Dicho proceso comienza por la selección de las piñas a ser procesadas, luego, éstas son partidas e introducidas en hornos para llevar a cabo su cocción. Después de un perfecto cocimiento se descargan los hornos y las piñas cocidas se trasladan a un área de molienda para extraer los azúcares que se encuentran en las fibras de la planta. Con el jugo obtenido se formula el *mosto* utilizado para la fermentación y de esta manera producir el alcohol junto con otros compuestos que conforman al tequila. Luego se entra en la fase de destilación, en la cual son eliminados algunos residuos existentes en el jugo, y a su vez es enriquecido el contenido alcohólico y refinado el producto. Una vez destilado, el producto final se concentra en tinas de paso donde se

diluye para pasarse a pipones o barricas de roble blanco, donde se añejará dependiendo del tipo de tequila que se desee obtener. Por último, antes de ser embotellado el tequila es filtrado para eliminar así algunos sólidos conferidos por la madera.

2.2 Tipos de agave micropropagado y convencional.

El agave que es utilizado por la empresa *Tequila Sauza* para la elaboración de su tequila es clasificado en dos tipos: *agave convencional* y *agave micropropagado*. Dicha clasificación está dada en función del origen de los hijuelos que serán trasplantados en campo.

Las plantas de agave cultivadas a partir de hijuelos obtenidos por el método de rizoma mencionado en la sección anterior son a las que se les denomina *agave convencional*.

Hasta hace algunos años, en *Tequila Sauza*, al igual que en toda la industria, la obtención del agave necesario para la elaboración de sus productos era mediante el método de rizoma. Sin embargo, en virtud de los avances tecnológicos que se tienen hoy en día esta empresa ha involucrado algunos de dichos avances para llevar a cabo una producción de agave más controlada, con el objetivo de mantener día a día la mejor calidad en las plantas utilizadas para la elaboración de su tequila.

De esta forma, para tener un mejor control sobre las características de los hijuelos que serán trasplantados en el campo, esta empresa ha incorporado un proceso de laboratorio para su producción, el cual consiste, a grandes rasgos, en tomar una muestra proveniente de un punto de crecimiento en una planta madre seleccionada, y con la información genética contenida en dicha muestra generar hijuelos mediante la utilización de técnicas biotecnológicas. Una vez que dichos hijuelos están en condiciones óptimas, son trasplantados al campo para que comiencen así su crecimiento. A las plantas de agave generadas mediante este procedimiento son a las que se les denomina *agave micropropagado*.

2.3 Problema de interés para la empresa.

Uno de los problemas de interés para la empresa *Tequila Sauza*, el cual se aborda en este trabajo, surge de lo observado en el campo en lo que respecta

al desarrollo de los dos diferentes tipos de plantas de agave manejados por dicha empresa. Los dos tipos de agave son referidos como el agave micropropagado y el agave convencional. Del manejo de estos dos tipos de planta se ha observado en el campo que aparentemente ambos se desarrollan de una manera distinta a lo largo del tiempo, por lo cual el interés principal radica en realizar un estudio comparativo de su desarrollo, el cual cuente con un sustento científico.

Dicho estudio, desde un punto de vista estadístico, puede pensarse en el sentido de una prueba de hipótesis, donde de lo observado en el campo en lo respecta al desarrollo de las plantas la hipótesis de interés a probar sería la diferencia en el desarrollo de éstas, teniéndose así, como hipótesis (*nula*), la igualdad de desarrollo entre ambas poblaciones de plantas (micropropagado y convencional). Para dar respuesta a esta interrogante se establecerán metodologías de carácter estadístico que permitan encontrar tal evidencia en favor o en contra de la hipótesis nula antedicha.

Así, una tarea específica será, por ejemplo, utilizar curvas de crecimiento que modelen el desarrollo de las plantas, y, plantear las hipótesis sobre el comportamiento de los dos tipos de agave en términos de los parámetros de dichas curvas. En el siguiente capítulo se hablará con más detalle sobre cada una de las metodologías propuestas, las cuales están encaminadas en ayudar a encontrar la evidencia necesaria para la hipótesis previamente mencionada.

Capítulo 3

Análisis de crecimiento para plantas de agave

A lo largo de este capítulo se proponen algunas metodologías orientadas al análisis del crecimiento de las plantas de agave. Para ello, se comienza dando, en la primera sección, una breve descripción sobre el crecimiento de plantas desde un contexto biológico. Luego, en la segunda sección, se describe el procedimiento para la construcción de lo que se denominará en este trabajo el *índice morfológico* de una planta de agave, el cual usaremos como una de las herramientas principales de modelación. En la tercera sección se presentan algunos de los modelos matemáticos más comunes que han sido empleados en la literatura biológica para modelar el crecimiento de diversos organismos, donde; uno de éstos modelos se empleará para modelar el crecimiento de las plantas de agave. Finalmente, en la última sección se describen algunas metodologías propuestas para llevar a cabo la comparación del crecimiento entre las dos poblaciones de plantas de agave, el cual es el objetivo principal de este trabajo.

3.1 Descripción del crecimiento de una planta.

3.1.1 *Análisis de crecimiento para una planta en general.*

En las ciencias biológicas es de gran interés y utilidad conocer la manera en que algún organismo crece a lo largo del tiempo. La forma en que se mide dicho crecimiento depende de la característica de interés en el estudio.

En el estudio de crecimiento de plantas, Beadle (en Coombs et al., 1985, Sección 2.2) menciona que como principio básico para un análisis de crecimiento es necesario considerar dos tipos de mediciones:

1. El *peso de la planta*. El cual se refiere usualmente al peso en seco (kg). Sin embargo, en algunos casos puede considerarse en su lugar a la materia orgánica o energía contenida.
2. El *tamaño del sistema asimilador*. El cual se refiere comúnmente al área por hoja (m²); no obstante, en su lugar también puede considerarse a la proteína en las hojas o al contenido de clorofila.

Los datos necesarios para el análisis pueden ser recolectados de plantas individuales o derivados de parcelas de plantas cosechadas, considerando siempre conjuntos de plantas o parcelas lo más homogéneas posible. Otro aspecto esencial a considerar en un estudio de esta naturaleza es la precisión en las mediciones tomadas, para evitar así obtener variaciones innecesarias durante el muestreo.

Uno de los conceptos básicos en el análisis de crecimiento, que surge del trabajo de Blackman en 1919 (referenciado en Coombs et al. opus cit.), es la *tasa de crecimiento relativa* (R) de una planta, la cual se define en cada instante en el tiempo (t) como el incremento relativo de material presente, y es el único componente en el análisis del crecimiento que no requiere del conocimiento del tamaño en el sistema asimilador. Por ejemplo, si consideramos que el material en estudio es el peso (W) de la planta entonces se tendría:

$$R = \frac{1}{W} \frac{dW}{dt} = \frac{1}{dt} (\ln W) \quad (3.1)$$

en donde la tasa de crecimiento relativa representa así la derivada instantánea de la curva $\ln W$ contra el tiempo, y ésta sirve como una medida fundamental de la producción de materia neta, en unidades de kg. por día, por ejemplo, y puede ser usada para comparar el desempeño de especies o los efectos de tratamientos bajo condiciones estrictamente definidas. Cabe mencionar que la tasa de crecimiento relativa no informa sobre los factores causales que determinan el crecimiento.

Otros conceptos clásicos considerados para el análisis de crecimiento de una planta son por ejemplo: la tasa unitaria de hoja (E , *unit leaf rate*), el cociente del área de las hojas (F , *leaf area ratio*), el cociente del peso en las hojas (LWR ,

leaf weight ratio), el área específica de las hojas (*SLA, specific leaf area*), el índice de área de hojas (*L, leaf area index*), la tasa de crecimiento de parcelas (*C, crop growth rate*) y la duración de área de las hojas (*D, leaf area duration*). Una descripción más detallada sobre estas componentes puede ser revisada en Coombs et al. (1985) y Nobel (1983).

Más adelante se describirán las variables consideradas en este trabajo para cuantificar el crecimiento de las plantas de agave, las cuales como se verá se ajustan a su vez a los principios básicos requeridos para un análisis de crecimiento de plantas, según lo descrito con anterioridad. Cabe hacer notar que para llevar a cabo el análisis de crecimiento en plantas de agave usaremos extensivamente la tasa de crecimiento relativa, además de otra definición de tasa que se describe con más detalle en la sección 3.3.

3.1.2 *Análisis de crecimiento para una planta de agave.*

Como se mencionó en el capítulo anterior, el objetivo primordial de este trabajo consiste en comparar el desarrollo de dos poblaciones de plantas de agave. Para esto, un supuesto básico a considerar a lo largo de este trabajo será que al hablar del desarrollo de una planta de agave éste se entenderá como el crecimiento de la misma, teniéndose por tanto que el crecimiento de la planta será un indicador de su desarrollo.

Una definición necesaria es la de variables que permitan cuantificar dicho crecimiento, y, en este trabajo, las variables consideradas para medir tal crecimiento son: *Altura de la planta, Número de hojas desprendidas y Longitud de hoja*, las cuales están íntimamente relacionadas con la morfología de la planta. Esto da pie a un segundo supuesto de trabajo, que consiste en suponer que las variables mencionadas miden efectivamente el crecimiento de las plantas. Esto último puede argumentarse diciendo que estas variables representarían al sistema asimilador de la planta, de acuerdo a lo descrito en la sección anterior. Un tema fundamental es investigar la asociación entre estas variables y el peso de la planta, ya que ésta última es la que ha recibido más atención como variable crítica para el proceso de producción del tequila; sin embargo, esto está fuera del alcance del presente trabajo.

Una descripción un poco más detallada sobre las variables consideradas para medirse en las plantas es la siguiente:

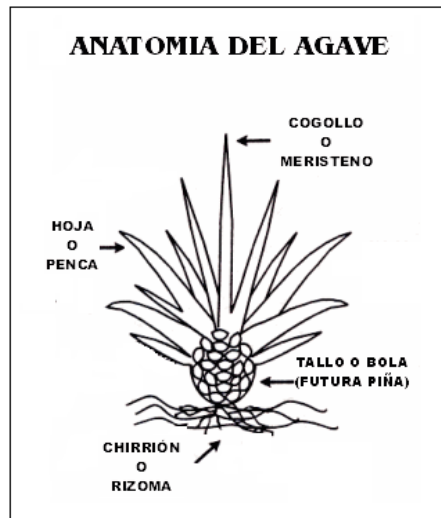


Figura 3.1. Anatomía de una planta de agave.

1. *Altura*, se refiere a la altura de la planta en centímetros medida desde la base de ésta hasta la punta del cogollo (ver Figura 3.1).
2. *Número de hojas desprendidas*, la cual se refiere al número total de hojas que se han separado del cogollo hasta el momento del conteo.
3. *Longitud de hoja*, la cual mide la longitud en centímetros de la última hoja que se ha separado del cogollo al momento de tomar las mediciones.

Una vez definidas las variables con las cuales cuantificar el crecimiento en las plantas de agave, la siguiente tarea consiste en establecer las metodologías para analizar el crecimiento de dichas plantas. Un primer paso para esto es dado en la siguiente sección, donde se definirá lo que se denominará a lo largo de este trabajo como el *índice morfológico* de una planta de agave.

3.2 El índice morfológico de una planta de agave.

El crecimiento de un organismo, específicamente el de una planta, puede describirse por medio de modelos matemáticos, los cuales usualmente son representados por medio de funciones a las que se denomina *curvas de crecimiento* (Seber and Wild, 1989; Clow and Urquhart, 1974). Muchos de estos modelos son generados como soluciones de ecuaciones diferenciales, las

cuales son obtenidas modelando matemáticamente supuestos biológicos sobre la dinámica del crecimiento. Más adelante, en la sección 3.3, se dará una descripción un poco más detallada sobre algunos de los modelos de crecimiento más comunes usados en la literatura.

En esencia, una curva de crecimiento es un ente matemático que representa la relación unívoca entre dos variables (variable respuesta y variable independiente), por lo cual es necesario definir cada una de estas variables para su construcción. En este trabajo, como se mencionó en la sección anterior, se cuenta no sólo con una sino con tres variables para medir el crecimiento en las plantas (variables respuesta), por lo que una opción posible para su análisis sería, por ejemplo, la construcción de las curvas de crecimiento correspondientes a cada una de estas variables como función del tiempo (variable independiente). Sin embargo, esto dejaría fuera la posibilidad de analizar el crecimiento de las plantas considerando en conjunto a las tres variables en cuestión, así como las interacciones entre dichas variables. Una posibilidad para involucrar a todas las variables como una sola entidad es movernos al campo del análisis multivariado.

El análisis multivariado trata con datos que contienen observaciones de una o más variables medidas a un conjunto de objetos o individuos. En esta rama de la estadística existen diversas técnicas para el análisis de datos como los referidos, donde éstas a su vez toman en cuenta las relaciones existentes entre las variables involucradas. Muchas de estas técnicas son generalizaciones de procedimientos clásicos univariados, como por ejemplo, una prueba T^2 de Hotteling en el caso multivariado es la generalización de una prueba-t en el caso univariado. Así mismo, existen otra clase de técnicas que son únicas en el campo multivariado y que buscan describir la estructura multivariada presente en los datos, como por ejemplo el análisis de componentes principales, del cual su principal uso es como técnica descriptiva, pero, sin embargo, también puede ser empleada en muchos procesos inferenciales.

Una de las técnicas multivariadas a ser empleada en este trabajo será precisamente el *Análisis de Componentes Principales (ACP)*, ya que con esta técnica es posible resumir la información presente en un conjunto grande de variables a un número menor de éstas, empleando para ello combinaciones lineales de las variables originales bajo ciertas restricciones. A continuación se dará una descripción más detallada sobre dicha técnica, ya que ésta será pieza clave en la construcción del *índice morfológico* de una planta de agave, el cual será utilizado para cuantificar el crecimiento de las plantas considerando la

información brindada a partir de las variables medidas a éstas. Algunos textos clásicos en los que es posible consultar un poco más sobre este tema y sobre los cuales está basado el desarrollo descrito a continuación son: Mardia et al. (1995), Manly (1986), Chatfield and Collins (1980), y Jolliffe (1986).

3.2.1 *Análisis de Componentes Principales.*

3.2.1.1 Introducción.

El *Análisis de Componentes Principales* (ACP) es probablemente una de las técnicas más antiguas y mejor conocida del análisis multivariado. Ésta fue descrita por primera vez por Karl Pearson en 1901, y después, en 1933, Harold Hotelling desarrolló de manera independiente métodos prácticos para la implementación de esta técnica, los cuales evolucionaron ampliamente a la llegada de las computadoras de la era moderna.

La idea central en el análisis de componentes principales es reducir la dimensionalidad de un conjunto de datos en el cual se tienen un número grande (p) de variables correlacionadas entre si a un número más pequeño (m) de variables, procurando conservar la mayor variabilidad posible de dichos datos. Esta reducción se logra al transformar las variables originales a un nuevo conjunto de variables, las *componentes principales*, las cuales tienen la propiedad de ser no-correlacionadas y de estar ordenadas de tal manera que las primeras componentes retienen la mayor cantidad de la variación presente en las variables originales. Esta técnica es especialmente útil cuando $m \ll p$ y si las m componentes pueden ser interpretadas. En esencia, el cálculo de las componentes principales se reduce a un problema de encontrar los valores y vectores propios de una matriz simétrica positiva semidefinida (matriz de covarianza o de correlación según corresponda), aunque ésta técnica tiene diferentes tipos de derivaciones.

3.2.1.2 Obtención analítica de los componentes principales.

Supóngase que se tiene un conjunto de p variables, digamos X_1, X_2, \dots, X_p , las cuales se encuentran correlacionadas entre si, entonces, el objetivo es obtener a partir de éstas un nuevo conjunto de variables, digamos Z_1, Z_2, \dots, Z_p , de las cuales entre sus principales características esté el ser no-correlacionadas y recabar la mayor cantidad de la variación presente en las variables originales.

Así mismo, supóngase que se cuenta con datos correspondientes a n individuos a los cuales les fueron medidas cada una de las p variables X . Típicamente se obtiene una estructura de datos similar a la mostrada en la Tabla 3.1 a continuación:

Tabla 3.1

Individuos	X_1	X_2	...	X_p
1	x_{11}	x_{12}	...	x_{1p}
2	x_{21}	x_{22}	...	x_{2p}
\vdots	\vdots	\vdots	\vdots	\vdots
n	x_{n1}	x_{n2}	...	x_{np}

Entonces, la primer componente principal (Z_1) corresponderá a la combinación lineal de las variables X_1, X_2, \dots, X_p dada por la expresión:

$$Z_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p, \quad (3.2)$$

o, en notación matricial:

$$Z_1 = a_1 \mathbf{X}', \quad (3.3)$$

en donde $a_1 = [a_{11}, a_{12}, \dots, a_{1p}]$, y \mathbf{X} corresponde a la matriz de datos de la Tabla 3.1. Se tiene una restricción sobre la norma de Z_1 , el cual se pide que sea unitario:

$$a_{11}^2 + a_{12}^2 + \dots + a_{1p}^2 = 1. \quad (3.4)$$

Análogamente, la segunda componente principal (Z_2) está dada por la expresión:

$$Z_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p \quad (\text{ó } Z_2 = a_2 \mathbf{X}'), \quad (3.5)$$

De igual modo, se pide que Z_2 sea unitario:

$$a_{21}^2 + a_{22}^2 + \dots + a_{2p}^2 = 1. \quad (3.6)$$

Las componentes principales restantes se definen de manera análoga a las anteriores, de tal forma que si se tienen p variables originales es posible obtener hasta p componentes principales si así se requiere.

Los coeficientes a_{ij} de cada una de las componentes principales (Z_i) se obtienen maximizando la varianza de las componentes, que depende de la matriz de covarianzas, Σ , de las p variables originales. Específicamente, la

solución es calcular los valores y vectores propios de dicha matriz. Si $\lambda_1, \lambda_2, \dots, \lambda_p$ son los valores propios, ordenados $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, y a_1, a_2, \dots, a_p son los correspondientes vectores propios, estos constituyen los coeficientes de las Z 's:

$$a_1 = [a_{11}, a_{12}, \dots, a_{1p}], a_2 = [a_{21}, a_{22}, \dots, a_{2p}], \dots, a_p = [a_{p1}, a_{p2}, \dots, a_{pp}], \quad (3.7)$$

entonces, se tiene que los valores propios cumplen con la propiedad:

$$\text{Var}(Z_1) = \lambda_1, \quad \text{Var}(Z_2) = \lambda_2, \dots, \quad \text{Var}(Z_p) = \lambda_p. \quad (3.8)$$

Cabe hacer notar que algunos de los valores propios λ_i , podrían ser 0, y valores negativos no son posibles al obtenerse éstos a partir de una matriz de covarianzas (matriz simétrica positiva semidefinida).

Una propiedad adicional para los valores propios obtenidos de esta matriz es:

$$\lambda_1 + \lambda_2 + \dots + \lambda_p = \text{Var}(X_1) + \text{Var}(X_2) + \dots + \text{Var}(X_p), \quad (3.9)$$

es decir, que la suma de varianzas de las componentes principales es igual a la suma de varianzas de las variables originales. Por tanto, en este sentido, las componentes principales recuperan toda la variación presente en los datos originales.

3.2.1.3 Componentes principales empleando matrices de correlación.

La derivación y propiedades de las componentes principales desarrolladas en la sección anterior están basadas en los valores y vectores propios de la matriz de covarianza. Sin embargo, en la práctica, es usual encontrarse con conjuntos de datos en donde el análisis es influenciado por la diferencia en la variación o en el tamaño de ciertas variables. Más adelante, en esta sección, se comenta sobre algunas situaciones donde suelen presentarse estos problemas.

Una manera alternativa para el cálculo de las componentes principales es utilizar la matriz de correlación, R , en lugar de la matriz de covarianzas Σ (para más sobre la obtención de la matriz R , consultar Mardia et al., 1995). El procedimiento matemático para el cálculo de las componentes principales utilizando la matriz de correlación R , es análogo al presentado con la matriz de covarianzas en la sección 3.2.1.2, utilizando para ello la matriz de correlación en lugar de la matriz de covarianza.

En caso de usar la matriz de correlación para el cálculo de las componentes principales se tiene ahora que las varianzas de éstas cumplen con:

$$\lambda_1 + \lambda_2 + \dots + \lambda_p = p, \quad (3.10)$$

Además, cabe mencionar que entre las componentes principales obtenidas empleando la matriz de correlación y las obtenidas empleando la matriz de covarianza no existe una relación explícita que lleve de unas a otras, y las obtenidas en ambos casos en general no coinciden, excepto en circunstancias muy especiales (Chatfield and Collins, 1980, Sección 4.4). A continuación se comentan algunas ventajas y desventajas sobre uso de ambas matrices para el cálculo de las componentes principales.

Un argumento a favor para el uso de la matriz de correlación, en lugar de la matriz de covarianzas, es que el cálculo de las componentes principales es sensible a las unidades de medición de las variables originales, ya que de existir grandes diferencias entre las varianzas de dichas variables, las variables cuyas varianzas son grandes tenderán a dominar las primeras componentes. Por otro lado, otro punto a favor sobre el uso de las matrices de correlación es la flexibilidad para comparar los resultados de diferentes análisis, por ejemplo, si se deseara comparar las magnitudes de las varianzas de las componentes principales, donde, se tendrían las mismas implicaciones al utilizar diferentes matrices de correlación, y no así en el caso de diferentes matrices de covarianza. Una desventaja para el uso de las matrices de correlación es que éstas arrojan coeficientes basados en variables estandarizadas y por tanto hace más difícil su interpretación directa.

En la práctica, el uso de la matriz de covarianzas tiene pocas ventajas sobre el uso de la matriz de correlación, sin embargo, existen algunos casos especiales, como por ejemplo, cuando se desea emplear las componentes principales para procesos inferenciales, basándose en una muestra de éstos, lo cual podría ocurrir al desear justificar la normalidad multivariada en un conjunto de componentes principales o en procedimientos de estimación. Una ventaja más sobre el uso de las matrices de covarianza es cuando se tienen variables medidas en las mismas unidades (digamos centímetros) y se busca identificar las variables que aportan la mayor cantidad de información en el análisis, ya que en este caso, la estandarización de las variables para el cálculo de las matrices de correlación tendría como consecuencia que todas las variables fuesen consideradas por igual.

El desarrollo anterior está basado en el conocimiento de los parámetros poblacionales. En la práctica, dichos parámetros son desconocidos pero estimados a partir de información muestral. En este caso, Σ es estimada por la matriz de varianzas-covarianzas muestral S . Los valores y vectores propios que se calculan son los de la matriz S , por lo que los componentes principales son, en el mejor de los casos, una estimación de los componentes principales poblacionales. Es necesario agregar supuestos distribucionales, por ejemplo, normalidad multivariada, para propósitos inferenciales usando los componentes principales muestrales.

3.2.2 Construcción del índice morfológico de una planta de agave.

Para cuantificar el crecimiento de las plantas de agave fueron consideradas tres variables de naturaleza morfológica (*Altura, Número de hojas desprendidas, y Longitud de hoja*), las cuales son medidas en una muestra seleccionada de plantas de agave elegidas en campo. Podemos describir la información contenida en ellas utilizando componentes principales.

Se construyen combinaciones lineales de la forma:

$$Z = c_1X_1 + c_2X_2 + c_3X_3, \quad (3.11)$$

en donde X_1 , X_2 y X_3 corresponderían a las variables *Altura, Número de hojas desprendidas, y Longitud de hoja*, respectivamente; y c_1 , c_2 y c_3 a los coeficientes correspondientes de cada variable.

Los coeficientes (c_i) propuestos son los correspondientes al primer componente principal muestral (a_{1j}), el cual queda caracterizado como la combinación lineal, de norma uno, con mayor variabilidad.

Para que la primer componente principal represente de manera consistente el crecimiento en las plantas será necesario contar con la condición de que todos los coeficientes de dicha componente sean positivos, es decir, $c_i > 0$, para $i = 1, 2, 3$. Esto es factible que suceda de acuerdo a lo comentado en Jolliffe, (1986, Sección 4.1) ya que las variables consideradas a medirse en las plantas representan medidas morfológicas de éstas y por tanto la primer componente principal podría interpretarse como una variable de "tamaño" de las plantas.

Con base a todo lo anterior, se tiene la siguiente definición:

Definición. Se define como el *índice morfológico* (IM) de una planta de agave en la fecha t a la evaluación de la primer componente principal en sus variables: *Altura, Número de hojas desprendidas y Longitud de hoja*, medidas cada una en la fecha t . Dicha componente será obtenida a partir de la matriz de correlación estimada con base en los datos correspondientes.

Sobre la definición anterior cabe hacer notar que se considera la obtención de las componentes principales empleando la matriz de correlación debido a que las variables no están definidas con las mismas unidades de medición: el *Número de hojas desprendidas* representa un conteo, mientras que las otras dos variables son medidas en centímetros.

El procedimiento a seguir para el análisis de crecimiento de las plantas de agave puede describirse a grandes rasgos como sigue: primero, seleccionar la muestra de plantas a considerar para el estudio, luego, establecer los tiempos (la frecuencia) en los que se llevará a cabo la toma de mediciones de las variables, una vez recolectados los datos, calcular los índices morfológicos de cada planta en cada fecha, y, finalmente, con éstos, llevar a cabo el análisis de crecimiento de acuerdo a lo que será descrito en la sección 3.4.

Antes de describir las metodologías propuestas para el análisis de crecimiento de las plantas, en la siguiente sección se hablará un poco sobre algunos de los modelos que han sido utilizados en la literatura para analizar el crecimiento de diversos organismos.

3.3 Modelos de crecimiento.

Al inicio del capítulo se proporcionó una breve introducción a los modelos matemáticos conocidos como *curvas de crecimiento* (Seber and Wild, 1989; Clow and Urquhart, 1974). Se mencionó que en la mayoría de los casos dichas curvas son obtenidas a partir de ecuaciones diferenciales que modelan de manera matemática supuestos biológicos sobre la dinámica del crecimiento. Un ejemplo básico de estas curvas es el *modelo exponencial*, del cual se dará una descripción un poco más detallada en la sección 3.3.1, y será precisamente uno de los modelos utilizados para representar el crecimiento de las plantas de agave.

Típicamente, en los organismos vivos la *tasa de crecimiento*, descrita matemáticamente como la derivada de la curva de crecimiento, no crece o decrece de una manera monótona, y en lugar de ello se incrementa a un máximo antes de decrecer de manera regular a cero. Este comportamiento es reflejado en una curva de crecimiento en forma de S, conocido como *patrón sigmoidal*. Los modelos con esta característica adicionan otro rasgo distintivo a la curva, que es la posición de uno o más *puntos de inflexión*, determinando con esto el tiempo en el cual la tasa de crecimiento obtiene su valor máximo. Entre algunos de los modelos sigmoidales más comunes están el Logístico y el Gompertz.

3.3.1 Modelo Exponencial.

Suponiendo que la variable t representa el tiempo, la función para modelar un crecimiento del tipo exponencial puede escribirse de la forma:

$$f(t; k) = e^{kt}, \quad (3.12)$$

en donde k representa la tasa de crecimiento. Un crecimiento exponencial para t pequeños es lo más típico de muchos modelos de crecimiento. Una gráfica de una función de crecimiento exponencial con base en la expresión (3.12) es mostrada en la Figura 3.2.

La característica principal de este modelo de crecimiento es que la tasa de crecimiento es proporcional al tamaño actual del organismo, es decir,

$$\frac{d}{dt} f(t; k) = ke^{kt} = kf(t; k), \quad (3.13)$$

Por tanto, si $k > 0$, la función (3.13) es positiva, lo cual implica que (3.12) es creciente. Por otro lado, si $k < 0$, la derivada (3.13) es negativa lo cual describe un decremento en (3.12), lo cual puede ser útil para modelar poblaciones sujetas a enfermedades o a procesos de extinción.

Cabe hacer notar la analogía que existe entre la expresión (3.13) y la descrita en (3.1), donde se observa que el parámetro k (tasa de crecimiento en el modelo exponencial) de la expresión (3.13) corresponde así mismo a la tasa de crecimiento relativa, R , de la expresión (3.1), al tenerse de (3.13) que:

$$k = \frac{1}{f(t; k)} \frac{d}{dt} (f(t; k)). \quad (3.14)$$

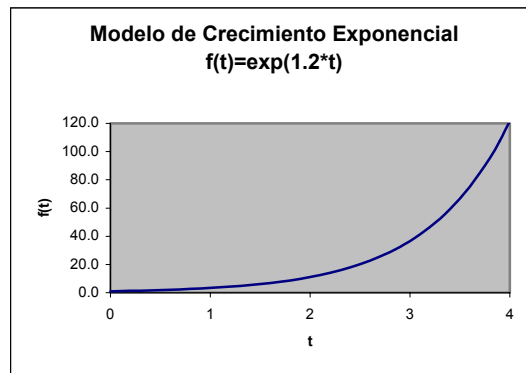


Figura 3.2. Ejemplo de una función de crecimiento exponencial.

Es decir, el modelo exponencial queda caracterizado por una tasa de crecimiento relativa constante.

Otra característica particular del modelo exponencial es que no existe límite para su crecimiento cuando $k > 0$, mientras que $\lim_{t \rightarrow \infty} f(t; k) = 0$, cuando $k < 0$.

En la literatura existen diversas versiones para una función de crecimiento exponencial, las cuales en esencia consisten en traslaciones o escalamientos de la función descrita en (3.12). En este trabajo, la versión a considerar para el ajuste del modelo de crecimiento propuesto en la sección 3.4 para las plantas de agave corresponde a una función de la forma:

$$f(t; c, k) = c + e^{kt}, \quad (3.15)$$

en donde c representa el valor inicial en el origen y k representa a la tasa de crecimiento. La función descrita en (3.15) cumple a su vez con propiedades análogas a las descritas anteriormente para la función (3.12). En la siguiente sección se mencionan algunas de las causas por las cuales es más factible considerar la función (3.15) en lugar de la (3.12) para el crecimiento de plantas de agave.

Otra variante posible para este modelo es:

$$f(t, k, w_0) = w_0 e^{kt} \quad (3.16)$$

3.3.2 Modelo Logístico.

Una manera simple de definir este modelo es por medio de la función de crecimiento:

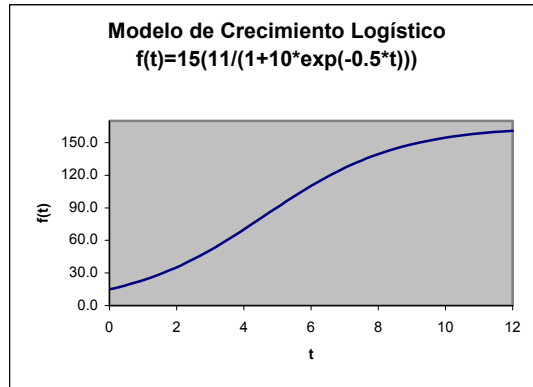


Figura 3.3. Ejemplo de una función de crecimiento logístico.

$$f(t; \alpha, k, \gamma) = \frac{\alpha}{1 + e^{-k(t-\gamma)}}, \quad (3.17)$$

en donde α representa el máximo crecimiento, k la tasa de crecimiento y γ el punto de inflexión. La curva tiene asíntotas: $f(t) = 0$ cuando $t \rightarrow -\infty$ y $f(t) = \alpha$ cuando $t \rightarrow \infty$. La tasa de crecimiento está al máximo cuando $f(t) = \alpha/2$, y esto ocurre cuando $t = \gamma$. Por tanto si $\gamma > 0$, el punto de inflexión para $f(t)$ es visible. La máxima tasa de crecimiento es $R_M = k\alpha/4$, y la tasa de crecimiento es simétrica alrededor de $t = \gamma$.

Una parametrización común para la expresión (3.17) es:

$$f(t; W_o, \beta, k) = W_o \frac{1 + \beta}{1 + \beta e^{-kt}}, \quad (3.18)$$

en donde el máximo valor de $f(t)$ en este caso estaría dado por $W_o(1+\beta)$, k continúa siendo la tasa de crecimiento, y $\beta = e^{k\gamma}$. En la Figura 3.3 se puede apreciar un ejemplo de una función de crecimiento logístico tomando como base la expresión (3.18).

Para valores pequeños de t la función logística tiene un comportamiento aproximadamente exponencial.

3.3.3 Modelo Gompertz.

El modelo Gompertz es frecuentemente usado para estudios de poblaciones y crecimiento de animales o plantas en situaciones donde no se tiene simetría

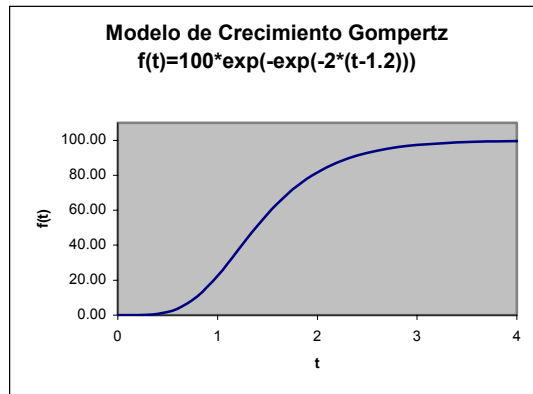


Figura 3.4. Ejemplo de una función de crecimiento Gompertz.

alrededor del punto de inflexión. La función de crecimiento en este caso está dada por la siguiente expresión:

$$f(t; \alpha, k, \gamma) = \alpha \exp(-\exp(-k(t - \gamma))), \quad (3.19)$$

en donde α representa el máximo crecimiento, k la tasa de crecimiento, y γ el punto de inflexión de la curva. Cuando $t = \gamma$ se tiene $f(t) = \alpha/e$, y a su vez la máxima tasa de crecimiento está dada por $W_M = k\alpha/e$. Este modelo también tiene la propiedad de que cualquier potencia de ésta sigue siendo una función de crecimiento Gompertz. Un ejemplo de una función de crecimiento del tipo Gompertz es presentado en la Figura 3.4.

Para valores pequeños de t la función Gompertz tiene un comportamiento aproximadamente exponencial.

Este modelo ha sido usado previamente para describir el crecimiento de plantas de agave (Mariaca, 2003).

3.4 Algunas metodologías propuestas para la comparación del crecimiento entre dos poblaciones de plantas de agave.

A continuación se presentan las metodologías propuestas para el análisis y comparación del crecimiento de las plantas de agave. Entre éstas se encuentran el uso de las tasas de crecimiento y el ajuste de modelos de crecimiento como los descritos en la sección anterior. Para ello se emplearán los índices morfológicos de las plantas, obtenidos según el procedimiento descrito en la sección 3.2.

3.4.1 Comparación empleando la tasa de crecimiento.

Existen al menos dos tipos de tasa de crecimiento que pueden ser empleadas para analizar el crecimiento de las plantas, una de ellas es la que en este trabajo denominaremos *tasa de crecimiento temporal*, y la otra que denominaremos como *tasa de crecimiento relativa promedio*. Cada una de éstas cuenta con una interpretación diferente, sin embargo, ambas tienen el mismo objetivo de medir la rapidez de crecimiento.

3.4.1.1 Tasa de crecimiento temporal.

La tasa de crecimiento temporal corresponde a una medida de la rapidez de crecimiento en la planta con respecto a un periodo de tiempo fijo. Para su obtención lo que se considera es el cociente de la diferencia entre dos mediciones tomadas en una variable en dos fechas consecutivas (correspondientes al periodo de tiempo estipulado) dividida sobre la longitud del periodo de tiempo referido. Por tal razón, esta tasa de crecimiento suele estar representada en unidades que dependen de la variable en cuestión, por ejemplo: kg/día (en mediciones de peso), cm/semana (en mediciones de longitud), etc.

Una manera más formal para describir el cálculo de esta tasa de crecimiento es la siguiente: supóngase que para cierta variable se cuenta con mediciones tomadas en las fechas t y T , entonces, la tasa de crecimiento temporal de la variable en cuestión en el periodo de tiempo $T - t$ estará dada como la razón de cambio con respecto al periodo de tiempo referido, es decir,

$$\frac{\text{TasaDeCrecimientoTemporal}}{\text{Periodo}T - t} = \frac{\text{MedicionEn}T - \text{MedicionEn}t}{T - t}. \quad (3.20)$$

Por ejemplo, si se considerará una planta de agave cuyas mediciones de su variable *Altura* en el mes de Enero y Agosto del 2004 fueron: Enero2004 = 140cm y Agosto2004 = 152cm, y se desease calcular la tasa de crecimiento temporal correspondiente a dicha variable del mes de Enero del 2004 al mes de Agosto del 2004 (considerando como unidad temporal un mes), entonces, la tasa para esta variable en el periodo de tiempo referido sería:

$$\frac{\text{TasaDeCrecimientoTemporal}}{\text{PeriodoEne2004aAgosto2004}} = \frac{152 - 140}{8} = 1.5 \text{ cm/mes}. \quad (3.21)$$

Nótese que de acuerdo a la expresión (3.20), para cualquier periodo de tiempo donde la tasa fuese 0 indicaría que no hubo cambio en la variable, teniéndose incluso valores negativos en caso de existir un decremento en la variable de una fecha a otra. Cabe mencionar también que la tasa de crecimiento definida en (3.20) representa la tasa de cambio promedio de la variable correspondiente en el intervalo de tiempo considerado, la cual se generaliza a la definición común de derivada (también conocida como tasa de cambio instantánea o puntual) al considerar progresivamente intervalos de tiempo más pequeños (Clow and Urquhart, 1974, Sección 6.1 y 6.2).

Así, esta tasa de crecimiento temporal representa una de las primeras herramientas a emplear en este trabajo para la comparación del crecimiento entre las dos poblaciones de plantas de agave. El procedimiento a seguir será, pues, comparar las tasas de crecimiento temporales de ambas poblaciones de plantas en cada uno de los periodos de tiempo a considerarse para la toma de mediciones, donde a su vez dichas tasas serán calculadas a partir de los índices morfológicos obtenidos con base en las mediciones tomadas de la muestra que incluye plantas de cada uno de los tipos bajo estudio (micropropagado y convencional).

Un primer análisis de las tasas de crecimiento, obtenidas con base en los índices morfológicos según la expresión (3.20), sería presentar una gráfica con las tasas de crecimiento en cada uno de los periodos de tiempo considerados, incluyendo a cada periodo de ambos tipos de planta (como las gráficas presentadas en la sección 5.3.2.1 más adelante). Esto permitiría de manera visual mostrar el comportamiento de las tasas de ambos tipos de planta para cada periodo de tiempo, lo cual mostraría, en caso de existir, evidencia sobre una posible diferencia entre las tasas de crecimiento de las dos poblaciones de plantas.

El contraste de las tasas de crecimiento de ambos tipos de planta puede plantearse de una manera más formal en el ámbito estadístico como una prueba de hipótesis, donde la hipótesis a probar de acuerdo con el conocimiento actual de la situación (igualdad en el crecimiento en ambos tipos de planta) podría plantearse como sigue:

H_T : las medias de las tasas de crecimiento temporales en el periodo de tiempo $T-t$ son iguales para ambas poblaciones de plantas (agave micropropagado y agave convencional).

El procedimiento estadístico a emplear para probar dicha hipótesis consistirá en la aplicación de una prueba de suma de rangos (para medias) empleando el estadístico de *Wilcoxon-Mann-Whitney* (para más referencias ver Kanji, 1993, pag. 86; Sprent and Smeeton, 2001, Sección 5.2), la cual es una prueba desarrollada en el ámbito de la estadística no-paramétrica que sirve para probar si dos muestras provienen de una misma población. Esta prueba es considerada para este caso debido a la naturaleza de las variables aleatorias involucradas, ya que los índices morfológicos calculados con base en la metodología descrita en la sección 3.2 no suelen contar en general con el supuesto de normalidad (Jolliffe, 1986, Sección 3.7), a menos que las variables originales correspondan a una distribución normal multivariada, lo cual muchas veces no ocurre en la práctica. Por tanto, es elegida esta prueba (entre otras posibles) en esencia por su robusticidad en lo concerniente a supuestos distribucionales.

Para el desarrollo de la prueba de suma rangos un supuesto necesario es que las dos poblaciones cuenten con distribuciones continuas (donde en general se obtienen resultados más confiables si a su vez éstas cuentan con varianzas iguales). En el caso de la tasa de crecimiento temporal, la continuidad de las distribuciones de ambas poblaciones (micropropagado y convencional) puede suponerse, por la manera en que están definidas dichas tasas según la expresión (3.20), al ser éstas el cociente de una combinación lineal de dos variables aleatorias continuas entre un valor constante. Sobre la igualdad de varianzas en ambas poblaciones es una cuestión que es necesario verificar una vez que se cuente con los datos necesarios, para lo cual pueden emplearse algunas pruebas robustas (que no consideran el supuesto de normalidad en las muestras) para su validación, como por ejemplo, la prueba de Levene (Manly, 1986, Sección 3.5) o la de Brown and Forsythe.

Una vez validados los supuestos necesarios, el método a seguir para realizar la prueba de suma de rangos sería el siguiente:

1. Las tasas de crecimiento temporales obtenidas para cada uno de los tipos de planta (micropropagado y convencional) son combinadas y ordenadas de manera creciente, dando a su vez un número de rango a cada una de éstas.
2. En caso de que existan tasas similares se les asignará el rango promedio obtenido de sus rangos correspondientes.
3. Se identifica la muestra de tamaño más pequeño (ya sea la correspondiente a las plantas del tipo micropropagado o a las del tipo

convencional) y son sumados sus correspondientes rangos, a los cuales denotaremos con la letra R . En caso de que ambas muestras sean de igual tamaño entonces se obtienen las sumas de rangos para cada muestra y se elige como R a la menor de éstas.

4. Se denota como N al tamaño de la muestra combinada y a n como el tamaño de la muestra más pequeña.
5. Es calculada R' como: $R' = n(N + 1) - R$.
6. Son comparados los valores de R y R' con los valores críticos correspondientes a esta prueba (Tabla 21, Kanji, 1993, pag. 191). Si R ó R' son menores que el valor crítico la hipótesis H_T será rechazada.

Para comparar las tasas de crecimiento temporales en diferentes periodos de tiempo, la prueba anterior es replicada con las muestras correspondientes a cada periodo.

3.4.1.2 Tasa de crecimiento relativa promedio.

Por su parte, la tasa de crecimiento relativa promedio corresponde a una medida (adimensional) de la rapidez de crecimiento en la planta con respecto a una medición inicial en algún periodo de tiempo. Para su obtención lo que se considera es el cociente de la diferencia entre dos mediciones tomadas en una variable en fechas consecutivas (correspondientes al periodo de tiempo en estudio) dividida sobre la primera de dichas mediciones.

Una manera más formal para describir el cálculo de dicha tasa de crecimiento es la siguiente: supóngase que para cierta variable se cuenta con mediciones tomadas en las fechas t y T , entonces, la tasa de crecimiento relativa promedio de la variable en cuestión en el periodo de tiempo $T - t$ estará dada como la razón de cambio con respecto a su medición inicial en la fecha t , es decir,

$$\frac{\text{TasaDeCrecimientoRelativaPromedio}}{\text{Periodo } T - t} = \frac{\text{MedicionEn } T - \text{MedicionEn } t}{\text{MedicionEn } t}. \quad (3.22)$$

Por ejemplo, si se considerará una planta de agave cuyas mediciones de su variable *Altura* se hicieron en los meses de Enero y Agosto del 2004 con los siguientes valores: Enero2004 = 140cm y Agosto2004 = 152cm, y se desease calcular la tasa de crecimiento relativa promedio correspondiente a dicha variable para el periodo del mes de Enero del 2004 al mes de Agosto del 2004, entonces, se tendría:

$$\frac{TasaDeCrecimientoRelativaPromedio}{PeriodoEne2004aAgo2004} = \frac{152-140}{140} = 0.0857. \quad (3.23)$$

Cabe hacer notar que de acuerdo a la expresión (3.22), valores pequeños de la tasa de crecimiento indicarían un cambio pequeño en la variable de una fecha a otra con respecto a la primer medición (siendo 0 en el caso de no haber cambio), a su vez, valores grandes indicarían un cambio grande en la variable de una fecha a otra con respecto a la primer medición tomada, teniéndose incluso valores negativos en caso de existir un decremento en la variable de una fecha a otra. Otro punto a notar es la adimensionalidad de esta tasa de crecimiento, lo cual a diferencia de la tasa definida en (3.20), hace posible la comparación de las tasas de distintas variables sin importar las unidades en las que estén medidas. Un punto a destacar sobre esta tasa y la definida en la sección anterior es la diferencia en magnitudes e interpretación de ambas, donde se observa que aunque se involucren las mismas mediciones de sus variables ambas tasas son distintas en magnitud, tal y como se puede observar de los ejemplos (3.21) y (3.23).

De esta forma, la tasa de crecimiento relativa promedio representa otra de las herramientas a emplear para la comparación del crecimiento entre las dos poblaciones de plantas (agave micropropagado y agave convencional). El procedimiento a seguir para tal labor también consistiría en comparar las medias de las tasas de crecimiento relativas de ambas poblaciones de plantas en cada periodo de tiempo, mismas que serían calculadas a partir de los índices morfológicos obtenidos con base en las mediciones tomadas en la muestra de plantas seleccionada.

Para este caso también se iniciaría presentando las gráficas con las tasas de crecimiento en cada uno de los periodos de tiempo considerados, incluyendo en éstas a ambos tipos de planta. Posteriormente, para contrastar el crecimiento de una manera más formal se plantearía una prueba de hipótesis, donde la hipótesis en este caso podría plantearse de la forma:

H_R : las medias de las tasas de crecimiento relativas promedio en el periodo de tiempo T-t son iguales para ambas poblaciones de plantas (agave micropropagado y agave convencional).

El procedimiento estadístico a emplear para probar dicha hipótesis sería análogo al descrito en la sección anterior, es decir, la aplicación de una prueba de suma de rangos (para medias) empleando el estadístico de *Wilcoxon-Mann-Whitney*. En este caso se considerada adecuada esta prueba

debido a que las tasas de crecimiento calculadas con base en la expresión (3.22) representan un cociente de variables aleatorias, los cuales comúnmente no suelen seguir una distribución normal, más aún si las variables involucradas en dicho cociente siguieran a su vez una distribución normal.

De igual modo, para esta prueba, es necesaria la validación del supuesto de continuidad de las distribuciones de tasas de crecimiento relativas promedio (así como verificar la igualdad de varianzas de las poblaciones), donde el supuesto de continuidad en las distribuciones es factible pensarse como válido al tenerse precisamente un cociente de variables aleatorias continuas.

Una vez validados los supuestos necesarios, el procedimiento para realizar la prueba sería análogo al descrito en los puntos 1 al 6 de la sección anterior, replicándose en cada caso para cada uno de los periodos de tiempo considerados.

3.4.2 Comparación empleando curvas de crecimiento.

En esta sección lo que se busca es describir el procedimiento para la construcción de las curvas de crecimiento correspondientes a cada tipo de planta (micropropagado y convencional), de forma que dichas curvas sean de utilidad para contrastar el crecimiento entre ambos tipos de planta.

Tradicionalmente, las curvas de crecimiento con una relación de la forma $y_i \approx f(x_i; \theta)$ han sido ajustadas empleando el procedimiento de mínimos cuadrados, posiblemente después de aplicar transformaciones para estabilizar la varianza. Esto también es apropiado bajo el modelo usual de regresión no-lineal:

$$y_i = f(x_i; \theta) + \varepsilon_i \quad (i=1,2,\dots,n), \quad (3.24)$$

en donde los ε_i son independientes e idénticamente distribuidos (i.i.d.) con media cero y varianza σ^2 , y $f(\cdot)$ es una función no-lineal. Tales modelos son razonables al utilizar datos en los cuales simples mediciones de “tamaño” son tomadas en cada individuo o unidad experimental, y las “edades” de los individuos son cubiertas por un rango de valores de x . El uso del modelo (3.24) en datos de crecimiento ha sido investigado más a fondo por Ratkowsky (1983), y también puede consultarse un poco más sobre estos modelos en Seber and Wild (1989), y Mead et al. (2003).

Supóngase entonces que se cuenta con una muestra de plantas de agave (n_1 plantas del tipo micropropagado y n_2 del tipo convencional) a las cuales les fueron medidas sus variables morfológicas (descritas en la sección 3.1.2) en determinados tiempos, digamos, t_0, t_1, \dots, t_m , de forma que cada planta cuenta con $m+1$ mediciones tomadas a lo largo del tiempo en cada una de sus variables.

De igual forma, supóngase que son calculados a cada planta sus correspondientes índices morfológicos en cada fecha (bajo la metodología descrita en la sección 3.2.2), a los cuales denotaremos por Y_{ij} , donde $i = 1, 2, \dots, n_1 + n_2$ (identificador para las plantas) y $j = 0, 1, 2, \dots, m$ (identificador para el tiempo). Entonces, lo que se busca es ajustar un modelo similar al de la expresión (3.24), el cual represente el crecimiento de ambos tipos de planta de acuerdo a la información presente en la muestra $\{Y_{ij}\}$.

La función no-lineal $f(\cdot)$ de la expresión (3.24), que es considerada en este trabajo para el ajuste del modelo de crecimiento en las plantas de agave, corresponde a un modelo exponencial como al descrito en la expresión (3.15). Es considerado el modelo exponencial por su sencillez, y más aún, porque se busca analizar el crecimiento de las plantas de agave en edades pequeñas (antes de los 3 años de vida en campo), y como se mencionó en la sección 3.3, la mayoría de los modelos de crecimiento presentan un comportamiento de este tipo para tiempos t pequeños. En el caso de requerirse un estudio de crecimiento para edades superiores se pueden considerar modelos más complejos, como el Logístico o Gompertz (descritos en las secciones 3.3.2 y 3.3.3, respectivamente), ya que estos modelos adicionan a la curva características diferentes como la presencia de *puntos de inflexión*, los cuales en este caso se esperaría estuviesen presentes en plantas de edades superiores a los 3 años, de acuerdo con lo observado en estudios previos realizados (Mariaca, 2003).

Cabe mencionar que para modelar el crecimiento de las plantas de agave se consideró usar la función exponencial descrita en (3.15) en lugar de la descrita en (3.12), debido a que los hijuelos de las plantas de agave cuando llegan al campo ya cuentan con cierto "tamaño" (que debería ser el óptimo para que las plantas puedan ser trasplantadas al campo de acuerdo a lo descrito en el capítulo 2), y por ende en el tiempo inicial, (el origen, cuando comienzan a medirse las plantas) ya cuentan con cierto valor inicial (c) de su "tamaño".

Por tanto, asumiendo que los índices morfológicos de las plantas de agave presentan un crecimiento medio de acuerdo a la expresión (3.15), el modelo estadístico propuesto para la modelación del crecimiento de las plantas de agave en los primeros años (los 3 primeros) de su ciclo de vida en campo puede escribirse de manera general de la siguiente forma:

$$y_{ij} = f(t_j; c, k) + e_{ij}, \quad (3.25)$$

donde:

- y_{ij} representa el j-ésimo índice morfológico de la i-ésima planta con $i = 1, 2, \dots, n$ (plantas); $j = 0, 1, \dots, m$ (observaciones en el tiempo).
- $f(t_j; c, k)$ representa el modelo no-lineal a considerar (en éste caso la función de crecimiento exponencial descrita en (3.15)).
- e_{ij} representa los términos de error residual, los cuales se asumen independientes e idénticamente distribuidos con $E[e_{ij}] = 0$.

Donde los términos de error residual, e_{ij} , representan fluctuaciones aleatorias en los índices morfológicos, y_{ij} , con $E[y_{ij}] = f(t_j; c, k)$. Cabe mencionar que en algunos casos es factible considerar el supuesto de normalidad en la distribución de dichos errores, sin embargo, esto es un aspecto que tiene que ser validado al contar con los datos pertinentes. En el caso que nos concierne, crecimiento de plantas de agave, este supuesto de normalidad podría pensarse como viable, debido a que en el crecimiento de plantas están involucrados diversos factores aleatorios que aparecen de manera natural en el medio ambiente en que vivimos y que afectan de manera directa al crecimiento de las plantas, un ejemplo de éstos son: el viento, la presión atmosférica, la energía solar recibida, etc., tal y como se describe en Coombs et al. (1985), Mead et al. (2003), y Nobel (1983).

Como la función de crecimiento (3.15) es tal que en sus segundas derivadas parciales con respecto a los parámetros éstos no desaparecen por completo de la expresión resultante, es necesario considerar para el ajuste del modelo (3.25) un ajuste del tipo *no-lineal* para estimar sus parámetros, para lo cual una de las técnicas a emplearse será la técnica de mínimos cuadrados (Seber and Wild, 1989, Capítulo 2).

Por tanto, el objetivo que se buscaría con el ajuste de dicho modelo sería precisamente estimar los parámetros c y k con base en la información

proporcionada por la muestra de plantas a través de sus índices morfológicos, lo cual se realizaría para cada tipo de planta. Una vez con los parámetros estimados sería posible construir curvas de crecimiento promedio para cada tipo de planta, con base en el modelo (3.25). Para el ajuste de cada tipo de planta se utilizarían el total de plantas correspondiente a cada tipo en la muestra, es decir, se utilizarían n_1 plantas para el ajuste del tipo micropropagado y n_2 para el del tipo convencional. Como se mencionará con más detalle en el capítulo 6, se debe tener cuidado de tomar una muestra lo suficientemente grande de cada tipo de planta, para evitar así problemas de sobreparametrización en los modelos ajustados. De esta forma, en cada ajuste se estarían estimando en total 2 parámetros de interés (c y k), la varianza de los errores residuales (σ^2), y las correspondientes varianzas de los estimadores de c y k . En este caso se estaría suponiendo que la varianza de los errores, σ^2 , permanece constante a lo largo del tiempo, sin embargo, en situaciones donde los datos son recolectados secuencialmente en el tiempo puede surgir una correlación serial en dichos errores, como suele ocurrir usualmente en el ajuste de curvas de crecimiento de organismos estudiados a lo largo del tiempo (Seber and Wild, 1989).

Cabe mencionar que debe tenerse cuidado que la matriz de correlación empleada para el cálculo de los índices morfológicos sea una matriz común a todas las plantas, es decir, emplear la misma matriz de correlación para el cálculo de los coeficientes de las componentes principales utilizando, por ejemplo, una ponderación de las matrices de correlación (ver Johnson and Wichern, 1992, Sección 6.3) de los tipos y regiones en cuestión, esto, para que los índices de las plantas sean todos comparables entre si.

De esta forma, una vez que se cuente con los parámetros estimados para cada tipo de planta (y por ende las curvas de crecimiento promedio estimadas), en una gráfica sería posible contrastar las curvas de crecimiento estimadas para ambos tipos de planta, donde, de existir diferencias en su crecimiento, éstas se verían reflejadas en dichas gráficas, e incluso sería posible comparar las tasas de crecimiento estimadas a partir de los modelos ajustados.

3.4.3 Modelos más generales.

En caso de ser necesario, el modelo (3.25) podría ser extendido a un modelo más complejo, en el cual se considerase por ejemplo: correlaciones seriales entre los errores, parámetros aleatorios en cada planta, etc., lo cual como se

mencionó pudiese presentarse en la práctica como consecuencia de la toma de mediciones secuenciales sobre las mismas plantas. De ser así, una opción podría ser, por ejemplo, considerar para el ajuste un *modelo no-lineal con efectos mixtos* (Lindstrom and Bates, 1990; Pinheiro and Bates, 2000), o algún modelo de series de tiempo (Seber and Wild, 1989, Capítulo 6) en caso de contar con una estructura en los datos idóneo para ello.

Un ejemplo de un modelo no-lineal con efectos mixtos, con base en el modelo exponencial (3.15), que podría considerarse para el análisis del crecimiento de las plantas de agave (en caso de observarse una gran heterogeneidad entre las plantas y así mismo correlaciones seriales entre las mediciones en los diferentes tiempos), es el siguiente:

$$y_{ij} = [c_{Intercepto} + \alpha_i + c_{Tipo} * Tipo] + \exp([k_{Intercepto} + \beta_i + k_{Tipo} * Tipo] * t_j) + e_{ij} \quad (3.26)$$

donde:

- y_{ij} representa el j-ésimo índice morfológico de la i-ésima planta con $i = 1, \dots, n$ (plantas), $j = 0, 1, \dots, m$ (observaciones en el tiempo).
- t_j representa los tiempos en que se tomaron las mediciones en las plantas.
- c representa el valor inicial en el origen (índice morfológico inicial al momento de llegar las plantas al campo).
- k representa el parámetro correspondiente a la tasa de crecimiento en las plantas.
- α_i y β_i representan los efectos aleatorios en los parámetros c y k respectivamente, asociados a la planta i , los cuales se asumen comúnmente en la literatura distribuidos de manera $N(0, \sigma_1^2)$ y $N(0, \sigma_2^2)$ en cada caso, e independientes entre plantas y entre sí.
- $Tipo$ representa una covariable para el tipo de planta, la cual tomaría valores: 0 en el caso del tipo micropropagado y 1 en el caso del tipo convencional.
- e_{ij} representa los términos de ruido blanco (error residual), los cuales se asumen independientes entre plantas y a su vez de los efectos aleatorios correspondientes, también bajo el supuesto de que siguen una distribución $N(0, \sigma^2)$.

Una de las ventajas de estos modelos es su flexibilidad para incorporar estructuras más complejas de la matriz de covarianzas de los errores en caso de ser requerido, por ejemplo, considerando varianzas no constantes de los

errores residuales. Así mismo, otra ventaja de estos modelos es la posibilidad de incorporar de manera sencilla covariables de interés, como podría ser el caso del tipo de planta, bloques correspondientes a diferentes tipos de tratamiento y/o tipos de suelo, condiciones climáticas, etc.

En estos modelos se considera que los parámetros de cada planta son afectados por efectos fijos y aleatorios. De esta forma, al tenerse que dichos parámetros son aleatorios, es posible estimar su variación, mismas que ya no es absorbida en la varianza estimada de los errores residuales como ocurriría en el caso de no considerar efectos aleatorios al tener una gran heterogeneidad entre las plantas de la muestra.

Cabe hacer notar que en el modelo (3.26) se utilizarían el total de plantas, n , en la muestra (incluyendo las correspondientes al tipo micropropagado y convencional), donde en este caso se hace distinción de ambos tipos a través de la covariable, *Tipo*, incluida en dicho modelo. A su vez, nótese que se tiene el supuesto de normalidad en los errores residuales, el cual es factible suponerse de acuerdo a lo comentado en Lindstrom and Bates (1990), al representar (3.26) una curva de crecimiento, en este caso correspondiente a plantas de agave.

Para la aplicación práctica de un modelo como el presentado en (3.26) es necesario que las muestras (de ambos tipos de planta) consideradas sean lo suficientemente grandes, para obtener de esta forma estimaciones consistentes de los parámetros (Lindstrom and Bates opus cit.), y así mismo evitar problemas de sobreparametrización en el modelo. También, debe procurarse que las plantas consideradas en el estudio cuenten con características similares en el experimento, para que de esta forma las variaciones presentes sean consecuencia natural de las plantas y no de factores externos no controlados.

Capítulo 4

Análisis exploratorio de los datos disponibles

En este capítulo se presenta el análisis exploratorio realizado a un conjunto de datos recabado por personal de la empresa *Tequila Sauza*. Los datos analizados consisten en mediciones tomadas a lo largo del tiempo en una muestra seleccionada de plantas de agave, con los cuales se pretende dar respuesta a la pregunta de interés planteada por dicha empresa, descrita a su vez en el capítulo 2. Por tanto, los análisis presentados en este capítulo están enfocados en evaluar la calidad y consistencia de los datos referidos. El capítulo inicia con una breve descripción de las variables medidas a las plantas de la muestra, donde dichas plantas fueron seleccionadas de lotes comerciales ubicados en dos diferentes regiones geográficas: Pihuamo y Tepatitlán. Las mediciones se tomaron desde de Enero del 2004 a la fecha, en periodos regulares de tiempo, y durante este proceso se tuvo la pérdida del lote de plantas de la región de Tepatitlán, impidiendo así continuar con el estudio de estas plantas. Por tal motivo, en este capítulo se presentan dos análisis exploratorios por separado: uno donde se consideran a ambas regiones con sus mediciones tomadas en Enero y Agosto del 2004, y otro donde solamente se considera a la región de Pihuamo con todas sus mediciones desde Enero del 2004 a Septiembre del 2005. Al final, se proporciona un diagnóstico de los datos analizados, y se dan algunas recomendaciones para posibles estudios futuros y nuevos proyectos.

4.1 Descripción de los datos.

Los datos disponibles fueron proporcionados por la empresa *Tequila Sauza*, y éstos corresponden a mediciones fisiológicas tomadas en diferentes fechas

a plantas de agave de una muestra seleccionada. Dichas plantas fueron elegidas en lotes comerciales ubicados en dos diferentes regiones geográficas: Pihuamo y Tepatitlán. Pihuamo se encuentra en la región sur de Jalisco y tiene un clima cálido; Tepatitlán está en la región denominada de Los Altos, hacia el este de Jalisco, en donde predomina el clima templado-frío.

En total para conformar la muestra se eligieron 40 plantas de agave, 20 elegidas en la región de Pihuamo y 20 en la región de Tepatitlán. En cada región 10 de estas plantas correspondían al tipo micropropagado y 10 al tipo convencional.

A cada una de las plantas referidas les fueron medidas las variables:

4. Altura, la cual se refiere a la altura de la planta medida en centímetros desde la base de ésta hasta la punta del cogollo.
5. Número de hojas desprendidas, la cual registra el número total de hojas que se han separado del cogollo hasta el momento del conteo.
6. Longitud de hoja, la cual mide la longitud en centímetros de la última hoja que se ha separado del cogollo al momento de tomar las mediciones.

Las mediciones fueron tomadas en los meses de Enero y Agosto del 2004 (para ambas regiones), y Enero y Septiembre del 2005 (sólo en las plantas de la región de Pihuamo). Los datos obtenidos se presentan en la Tabla 1 del Apéndice A.1.

Es importante mencionar que la muestra de plantas seleccionada por el personal de *Tequila Sauza* fue elegida empleando un muestreo sistemático por bloques. Dicho muestreo se llevo a cabo en cada región de la siguiente manera: primero se eligió al azar una planta en el campo, luego, se eligieron las 4 plantas sucesivas a ésta en el mismo surco (esto para que las plantas seleccionadas fueran una muestra representativa de las características del terreno en que estaban plantadas), entonces, el siguiente bloque a elegir distaba una número n fijo de plantas del primer bloque elegido. Para ambos tipos de planta la muestra fue tomada siguiendo el procedimiento anterior. El objetivo de tomar un muestreo de esta forma fue obtener un mejor control de la variabilidad en las mediciones, al tener así características similares entre las plantas de cada bloque en lo que respecta a tipo de suelo, tratamientos recibidos, etc.

Otro aspecto importante sobre las plantas elegidas es que éstas fueron trasplantadas al campo en Junio del 2002, teniendo así 19 meses en campo al momento de tomar la primera medición en Enero del 2004. Además, también se sabe que las plantas del tipo micropropagado requirieron de 10 meses previos a su plantación en campo para su formación, y que las del tipo convencional fueron obtenidas de plantas madre de 3 años de edad.

4.2 Análisis exploratorio (Pihuamo, Tepatitlán; Periodo: Enero-Agosto del 2004).

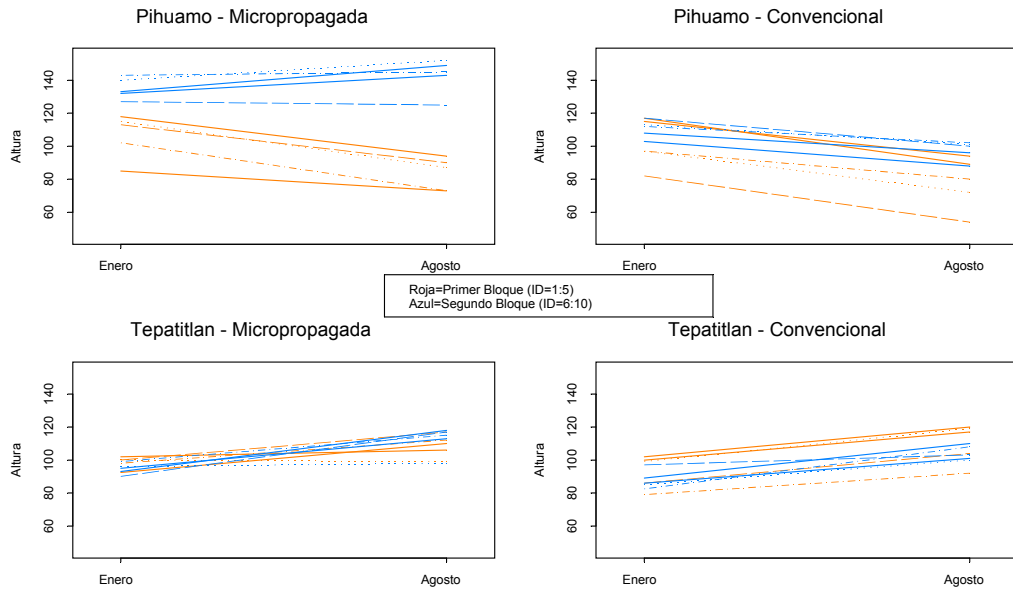
Los datos analizados en esta sección corresponden a mediciones tomadas en los meses de Enero y Agosto del 2004 en plantas de agave de la región de Pihuamo y Tepatitlán. En total se analizaron 40 plantas de agave, 20 pertenecientes a la región de Pihuamo y 20 a la región de Tepatitlán, donde a su vez en cada región 10 plantas correspondían al tipo micropropagado y 10 al tipo convencional. Las mediciones obtenidas para cada planta se pueden apreciar en la Tabla 1 del Apéndice A.1.

Como un primer análisis gráfico de los datos se tienen las que llamaremos *gráficas longitudinales*, presentadas en la Figura 4.1. En éstas se grafican cada una de las mediciones obtenidas para cada variable (*Altura, Número de hojas desprendidas y Longitud de hoja*) como función del tiempo.

Un primer aspecto que se puede apreciar de dichas gráficas es la diferencia en variabilidad entre ambas regiones, presente para cada una de las variables y siendo mayor en las plantas de Pihuamo en contraste con las de Tepatitlán.

Así mismo, cabe destacar la diferencia evidente entre los dos bloques de plantas del tipo micropropagado de la región de Pihuamo, lo cual se puede apreciar en cada una de las variables. Este comportamiento tan característico entre ambos bloques, de acuerdo a lo explicado por parte del personal de la empresa encargado del cuidado de las plantas en campo, se debe en esencia a que las condiciones del terreno en ambos bloques no son las mismas, al tenerse uno de ellos sobre una ladera de alta pedregosidad mientras que el otro se encuentra sobre un terreno plano (de hecho, las plantas del tipo convencional en Pihuamo y las de ambos tipos en Tepatitlán se encuentran plantadas sobre terrenos planos). Así mismo, comentaron que esta diferencia en el crecimiento de ambos bloques del tipo micropropagado en Pihuamo puede atribuirse al efecto dañino de un herbicida preemergente (Combine), el cual fue aplicado sólo a ambos tipos de planta de dicha región y que al ser

Variable -Altura



Variable -Hojas desprendidas

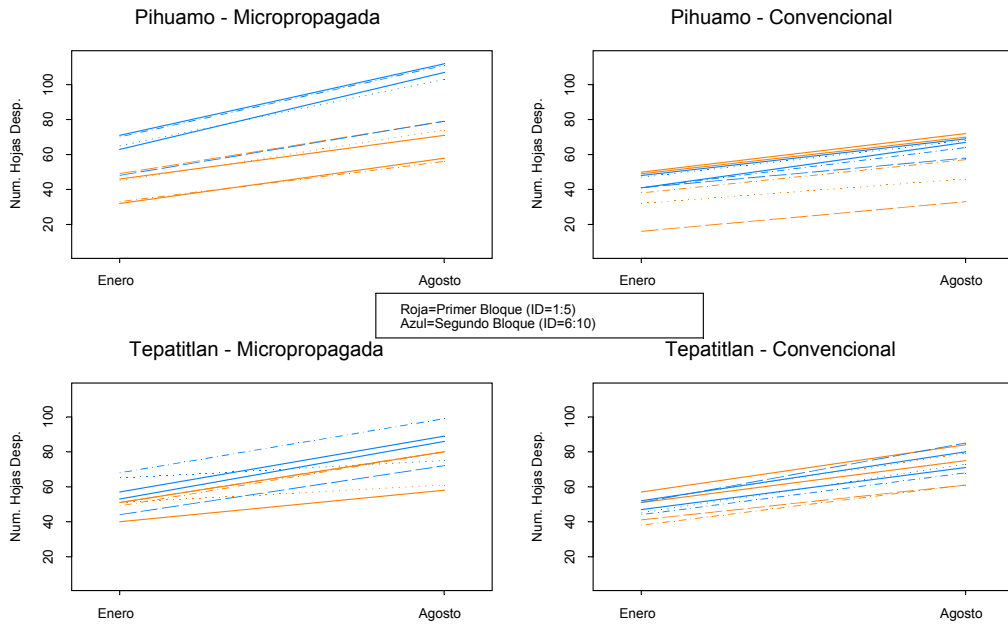


Figura 4.1. Gráficas longitudinales de las variables *Altura* y *Hojas desprendidas*, medidas en los meses de *Enero* y *Agosto* del 2004 en las plantas de agave de la muestra seleccionada en las regiones de *Pihuamo* y *Tepatitlán*.

Variable -Longitud de hoja

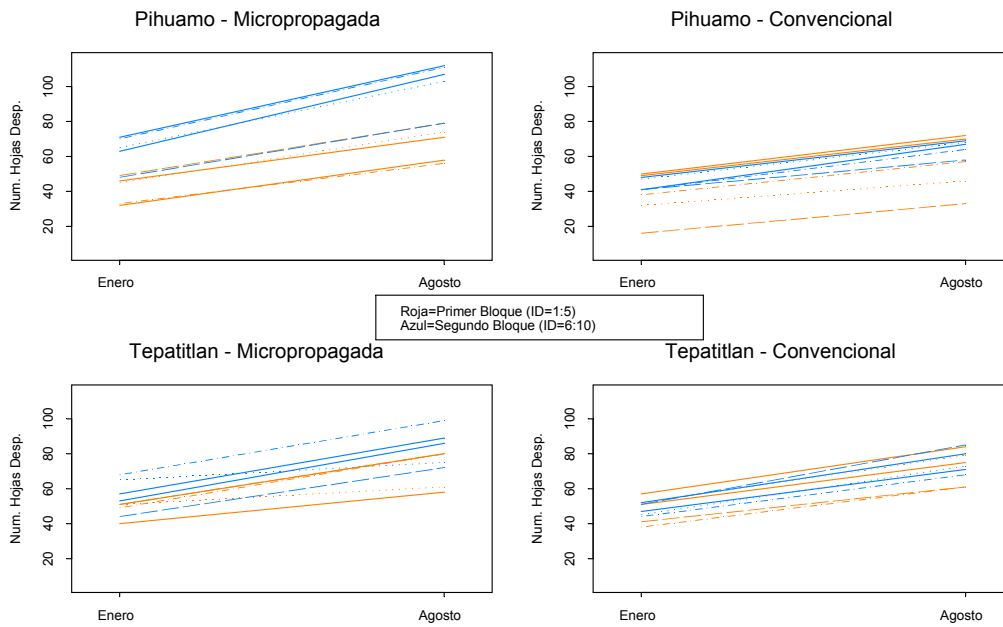


Figura 4.1 (continuación). Gráficas longitudinales de la variable Longitud de hoja, medida en los meses de Enero y Agosto del 2004 en las plantas de agave de la muestra seleccionada en las regiones de Pihuamo y Tepatitlán.

muy soluble debió lavarse más rápido en las plantas de la ladera que en el resto, teniendo así un efecto menor del herbicida en dicho bloque de plantas y repercutiendo a su vez en un mayor desarrollo de sus variables.

Otro punto importante a resaltar de las gráficas anteriores es la forma en que se comporta la variable *Altura* para la mayoría de las plantas de la región de Pihuamo, pues en lugar de observar un aumento en la altura de éstas (o que permaneciera constante en caso de estar del mismo tamaño de una fecha a otra) se exhibe una disminución en la altura de una fecha a otra, lo cual estaría indicando que las plantas se hacen más pequeñas conforme pasa el tiempo. Como se verá en la siguiente sección donde se continúa con el análisis de las plantas de esta región pero con mediciones en meses posteriores, la disminución de la altura sólo está presente del mes de Enero a Agosto del 2004. El comportamiento tan irregular de esta variable en este periodo de tiempo no tiene una explicación concisa por parte del personal de la empresa, sin embargo, ellos comentan que pudo deberse a la acumulación de suelo por arrastre de la lluvia o probablemente a algún efecto climático.

Un segundo análisis gráfico se presenta en las Figuras 4.2 y 4.3, donde se puede apreciar para cada una de las variables el cambio proporcional ocurrido de una fecha a otra. En estas gráficas lo que se esperaría es que los datos estuviesen por arriba de la recta de 45 grados trazada, ya que esto indicaría que en efecto se tuvo un crecimiento de la variable en cuestión. Sin embargo, se puede apreciar en la gráfica correspondiente a la variable *Altura*, que aproximadamente la mitad de los datos (en su mayoría pertenecientes a la región de Pihuamo) se encuentran por debajo de dicha recta, indicando así una disminución en las magnitudes de esta variable de una fecha a otra para dichas plantas. En estas mismas figuras, se incluyen las gráficas correspondientes a las Variables *Hojas desprendidas* y *Longitud de hoja*, donde el comportamiento de éstas es en general el esperado de acuerdo a lo mencionado anteriormente.

Hasta el momento, con estos primeros análisis gráficos, los datos reflejan la existencia de algunos factores importantes que no fueron controlados debidamente al no procurarse igualdad de condiciones en aspectos como: tipos de terreno sobre el que se encontraban las plantas y tratamientos aplicados. Así mismo, como se pudo vislumbrar de estos primeros análisis, existen problemas en una de las variables medidas a las plantas para al menos la mitad de las plantas en el estudio. Por tanto, con el propósito de identificar las posibles causas que originaron el comportamiento tan irregular observado hasta ahora en los datos, se investigó más sobre el diseño empleado y el desarrollo del experimento a partir del cual se obtuvieron estas mediciones. En el proceso, salieron a la luz algunos aspectos adicionales que tampoco fueron controlados para procurar igualdad de condiciones en las plantas de la muestra. Entre dichos aspectos se tienen: la aplicación de dos herbicidas (Hyvar y Faena) distintos al que fue aplicado en la región de Pihuamo (Combine), mismos que se aplicaron solamente a las plantas de la región de Tepatitlán, y no se sabe incluso si los daños que ocasionan en las plantas por su uso son comparables; el otro aspecto fue la aplicación de la técnica agrícola denominada en campo como “*aporque*” (que consiste en acercar suelo a la línea de cultivo debido a que las plantas se estuviesen inclinando por quedar sobre la superficie del terreno), la cual sólo se aplicó a las plantas del tipo micropropagado de la región de Tepatitlán.

Como se podrá ver mas adelante, todos estos factores afectan de manera directa o indirecta cualquier tipo de análisis que se lleve a cabo, ya que algunas de las plantas destacan en su crecimiento con respecto a las demás como consecuencia de estos mismos factores.

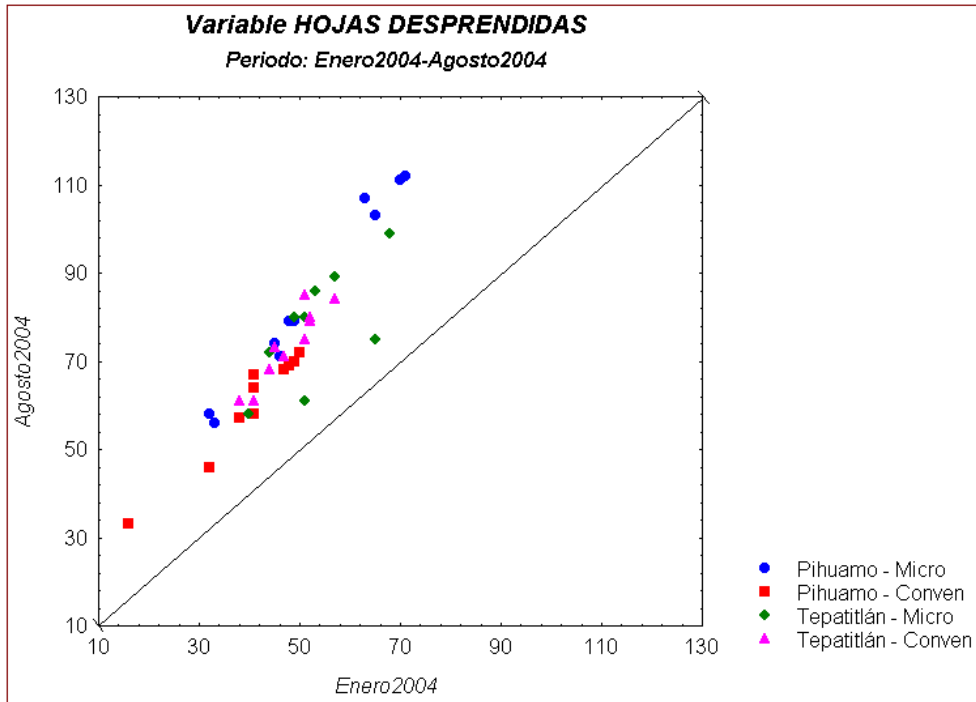
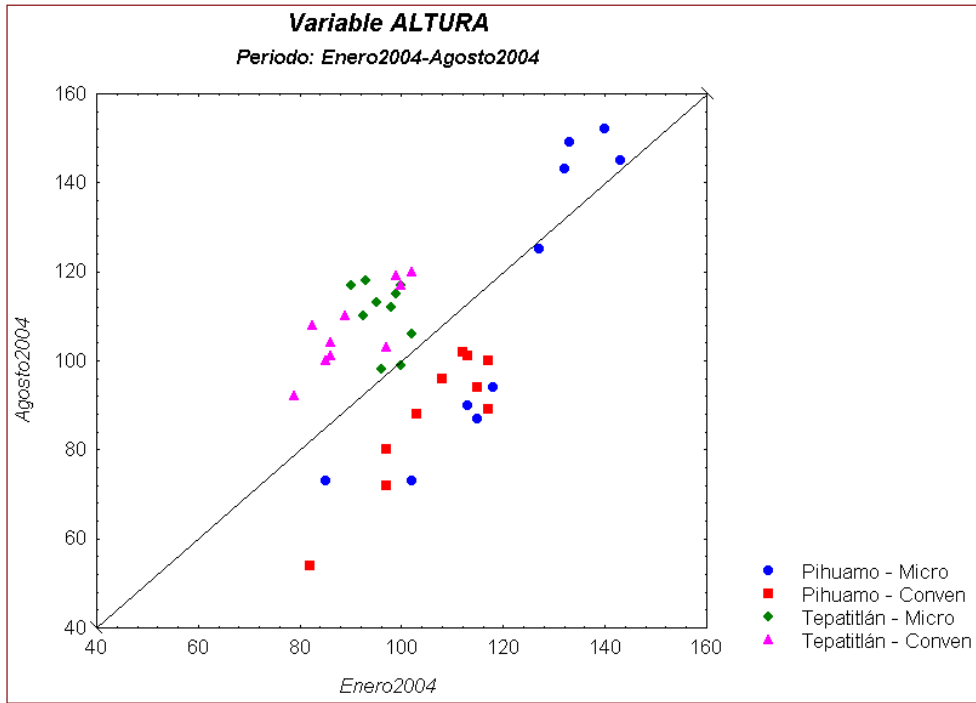


Figura 4.2. Gráficas del cambio ocurrido en las variables *Altura* (arriba) y *Hojas desprendidas* (abajo) del mes de Enero a Agosto del 2004 para las plantas de agave de las regiones de Pihuamo y Tepatitlán.

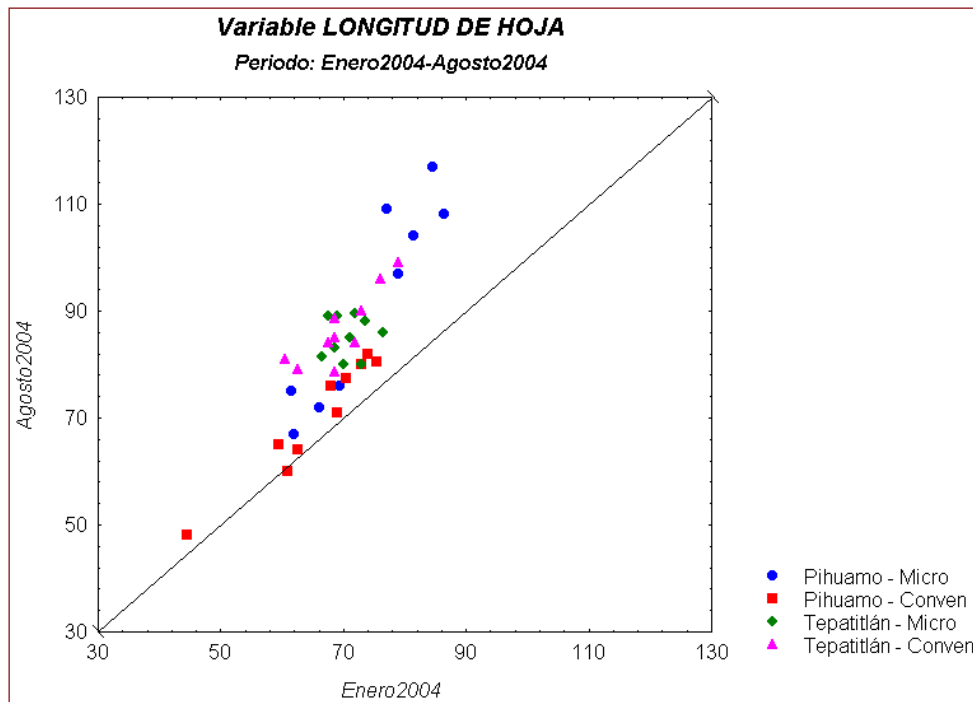


Figura 4.3. Gráfica del cambio ocurrido en la variables Longitud de hoja del mes de Enero a Agosto del 2004 para las plantas de agave de las regiones de Pihuamo y Tepatitlán.

Por otro lado, debido a que las dos regiones en que se encuentran las plantas muestran de los análisis anteriores una clara diferencia en sus características propias y a su vez entre las planta de cada región, se continuará el análisis exploratorio haciendo distinción entre las regiones, y a su vez entre los tipos de planta, con el objetivo de analizar un poco más a fondo los efectos causados como consecuencia de los factores no controlados presentes en estos primeros análisis gráficos.

En la Tabla 4.1 se pueden apreciar algunas estadísticas descriptivas (medias y desviaciones estándar) de los datos, calculados para cada región y tipo de planta. De los resultados obtenidos se puede observar que al restringirse a cada una de las regiones y contrastar sus valores de las medias entre ambos tipos de planta, las del tipo micropropagado presentan mayores valores de sus medias que las del tipo convencional (a excepción de la variable *Longitud de hoja* en el mes de Agosto del 2004 para las plantas de la región de Tepatitlán donde ocurre lo contrario). A su vez, nótese, que la diferencia entre las medias de ambos tipos es mayor en las plantas de la región de Pihuamo,

Tabla 4.1. Estadísticas descriptivas por región y tipo de planta de las mediciones en el agave los meses de Enero y Agosto del 2004.

Región	Tipo	Estadísticas Descriptivas	Enero 2004			Agosto 2004		
			Altura	Hojas Desp.	Long. Hoja	Altura	Hojas Desp.	Long. Hoja
Pihuamo	Micro	Media	120.8	52	72.9	113.1	85	88.5
		Desv. Estand.	18.0	14	10.0	32.7	22	20.6
Pihuamo	Conven	Media	106.1	40	65.8	87.6	60	70.4
		Desv. Estand.	11.3	10	9.3	15.2	12	11.0
Tepatitlán	Micro	Media	96.6	53	70.8	110.5	78	85.1
		Desv. Estand.	3.9	9	3.1	7.3	12	3.8
Tepatitlán	Conven	Media	90.6	48	69.6	107.4	74	86.5
		Desv. Estand.	8.2	6	5.7	9.2	9	6.9

lo cual en gran parte es gracias al bloque de plantas del tipo micropropagado de dicha región que cuenta con condiciones de terreno distintas al resto (plantas en la ladera con menor efecto del herbicida), y a consecuencia mayores magnitudes en los valores medios de sus variables.

Por otro lado, al comparar las medias de cada tipo de planta entre ambas regiones se puede observar que las correspondientes al tipo micropropagado son mayores en la región de Pihuamo en comparación con las de Tepatitlán (a excepción de *Hojas desprendidas* en el de Enero del 2004), de igual forma como consecuencia del bloque de micropropagado en la ladera; mientras, para el tipo convencional los valores medios de sus variables son mayores en la región de Tepatitlán que en Pihuamo (a excepción de *Altura* en Enero del 2004).

Así mismo, en lo que respecta a la variabilidad, al comparar entre ambas regiones, se puede observar de la tabla anterior que ésta es mayor en ambos tipos de planta para la región de Pihuamo, en contraste con la obtenida para la región de Tepatitlán. Por su parte, al comparar la variabilidad entre los dos tipos de planta en cada región se observa que en Pihuamo las del tipo micropropagado presentan una mayor variabilidad con respecto a las del tipo convencional en cada una de las variables; mientras que en Tepatitlán las variables *Altura* y *Longitud de hoja* en ambas fechas presentan una mayor variabilidad en las del tipo convencional con respecto a las del tipo micropropagado, y contrario a la variable *Hojas desprendidas*.

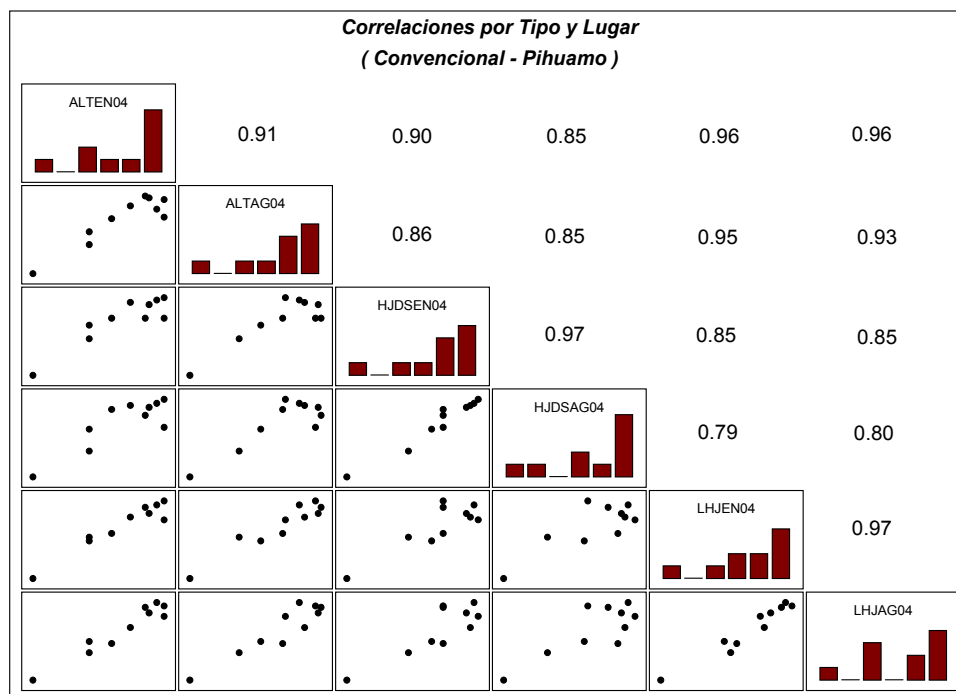
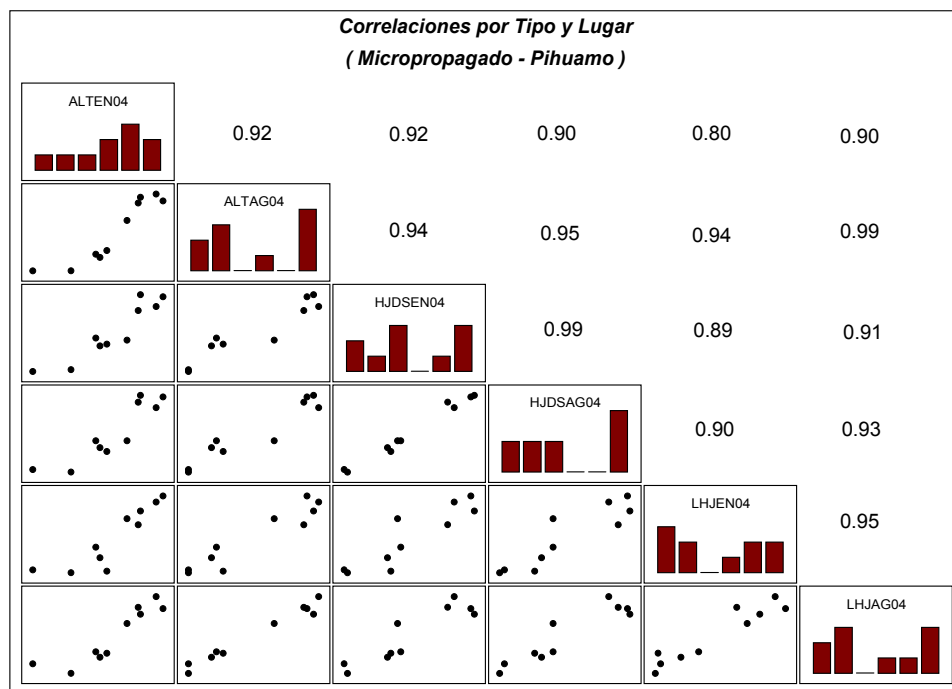


Figura 4.4. Gráficas de correlación de las variables medidas en los meses de Enero y Agosto del 2004 en las plantas de agave de la región de Pihuamo.

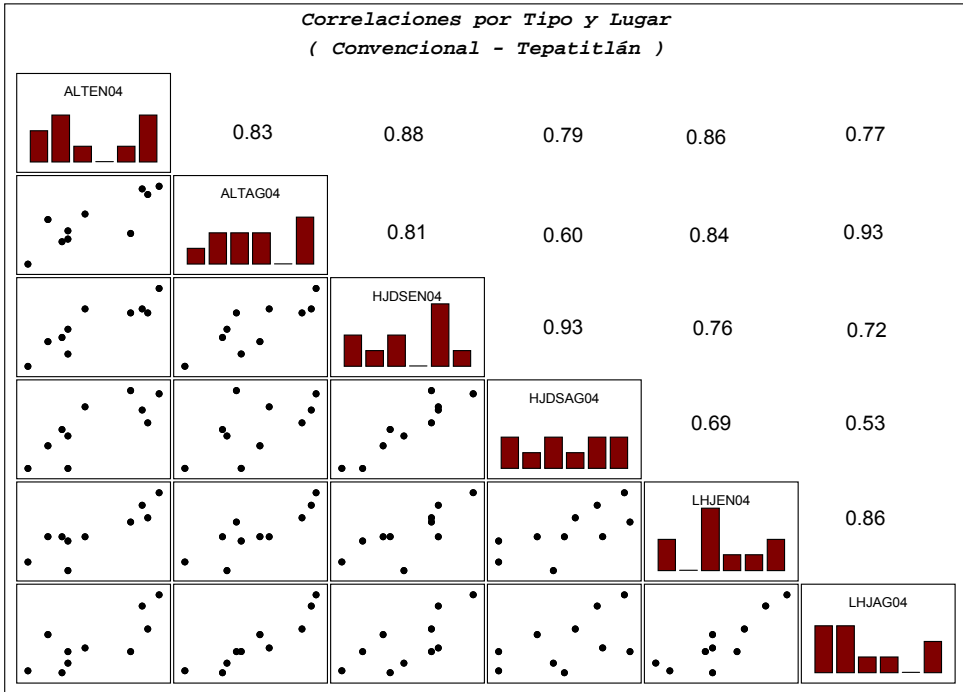
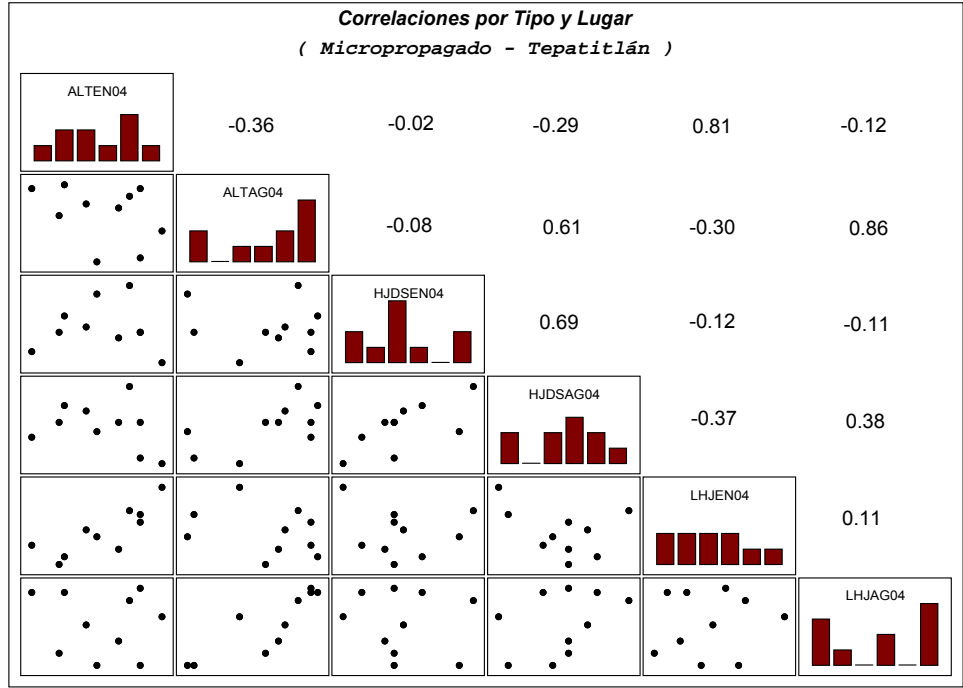


Figura 4.5. Gráficas de correlación de las variables medidas en los meses de Enero y Agosto del 2004 en las plantas de agave de la región de Tepatitlán.

En la Figura 4.7, que se presenta más adelante en la siguiente sección, se pueden apreciar las gráficas de caja y bigotes (box-plots) correspondientes a cada tipo de planta en cada región, donde éstos reflejan a su vez lo observado de la tabla anterior, principalmente en lo que respecta a la variabilidad, al apreciarse notablemente mayor en las plantas de la región de Pihuamo (Figura 4.7, arriba) con respecto a las de Tepatitlán (Figura 4.7, abajo).

Por su parte, las Figuras 4.4 y 4.5 muestran las correlaciones entre cada una de las variables, haciendo distinción entre región y tipo. En éstas se puede observar que en la región de Pihuamo las correlaciones para ambos tipos de planta son en general mayores a las de Tepatitlán. Por otro lado, al observar el comportamiento de las correlaciones en cada región se puede notar que en Pihuamo las correlaciones entre ambos tipos de planta son de magnitudes similares, mientras que en Tepatitlán las correlaciones son mayores en las plantas del tipo convencional, teniéndose incluso correlaciones negativas en las plantas del tipo micropropagado. La razón de dichas correlaciones negativas se puede apreciar en las gráficas longitudinales presentadas con anterioridad en la Figura 4.1 y a su vez de los datos de la Tabla 1 del Apéndice A.1, donde se puede observar en algunas de las plantas del tipo micropropagado de la región de Tepatitlán que mientras su medición en la variable *Altura* es pequeña las mediciones de sus variables *Número de hojas desprendidas* y *Longitud de hoja* son grandes, tanto en el mes de Enero del 2004 como en Agosto de ese mismo año, ocasionando por ende una asociación negativa entre las variables. Esto pudo deberse en gran parte a la acción de uno de los herbicidas empleados en dichas plantas, ya que precisamente uno de éstos (Faena) puede provocar marchitez en las plantas debido al daño que causa a nivel raíz.

En la siguiente sección, se podrá observar como los factores no controlados siguen estando presentes aún al contar con mediciones tomadas en fechas posteriores para estas mismas plantas.

4.3 Análisis exploratorio (Pihuamo; Periodo: Enero 2004 - Septiembre 2005).

En esta sección se incluyen para su análisis mediciones adicionales tomadas en las plantas de agave de la región de Pihuamo en los meses de Enero y Septiembre del 2005. Teniéndose por tanto, para las plantas referidas, mediciones desde Enero del 2004 hasta Septiembre del 2005. Como se

Tabla 4.2. Estadísticas descriptivas de las mediciones en el agave los meses de Enero 2004, Agosto 2004, Enero 2005 y Septiembre 2005 por tipo de planta.

Tipo	Estadísticas Descriptivas	Enero 2004			Agosto 2004			Enero 2005			Septiembre 2005		
		Altura	Hojas Desp.	Long. Hoja	Altura	Hojas Desp.	Long. Hoja	Altura	Hojas Desp.	Long. Hoja	Altura	Hojas Desp.	Long. Hoja
Micro	<i>Media</i>	120.8	52	72.9	113.1	85	88.5	146.8	109	113.1	156.1	141	119.2
	<i>Desv. Estand.</i>	18.0	14	10.0	32.7	22	20.6	25.0	21	16.6	25.2	24	17.2
Conven	<i>Media</i>	106.1	40	65.8	87.6	60	70.4	104.7	83	86.4	116.5	107	92.9
	<i>Desv. Estand.</i>	11.3	10	9.3	15.2	12	11.0	14.4	17	10.6	16.3	22	10.5

mencionó al inicio de este capítulo, no se cuenta con mediciones adicionales para las plantas de agave de la región de Tepatitlán, al tenerse la pérdida de dicho lote a raíz de una enfermedad que afectó a éstas y que como consecuencia provocó la destrucción del lote completo.

Por tanto, los datos que serán analizados a continuación corresponden a las mediciones tomadas en las 20 plantas de agave seleccionadas en la región de Pihuamo, a cada una de las cuales se les midieron las variables: *Altura*, *Número de hojas desprendidas* y *Longitud de hoja* en los meses de Enero (2004), Agosto (2004), Enero (2005), y Septiembre (2005). Dichos datos son presentados en la Tabla 1 del Apéndice A.1. De las 20 plantas seleccionadas 10 pertenecen al tipo micropropagado y 10 al tipo convencional.

En la Figura 4.6 se presentan las gráficas longitudinales correspondientes a los datos mencionados. En éstas se puede apreciar la evolución de cada una de las variables a lo largo del tiempo, desde Enero del 2004 hasta Septiembre del 2005, diferenciando a su vez por tipo de planta. En dichas gráficas las plantas del tipo micropropagado son identificadas del caso 1 al 10, mientras que las del tipo convencional del caso 11 al 20.

De dichas gráficas se puede observar claramente que al considerar las nuevas mediciones sigue prevaleciendo la diferencia marcada entre los dos bloques de plantas del tipo micropropagado, al presentar uno de sus bloques (caso 6 al 10, correspondientes a las plantas cultivadas en la ladera) en su mayoría magnitudes superiores para cada una de las variables medidas, predominando esto incluso en las mediciones tomadas recientemente en el 2005. Así mismo, puede apreciarse de dichas gráficas que la disminución en la altura de las plantas sólo está presente del mes de Enero a Agosto del 2004, lo cual ocurre en ambos tipos de planta. En el Apéndice A.2 se incluyen las gráficas correspondientes al cambio ocurrido en cada una de las variables en

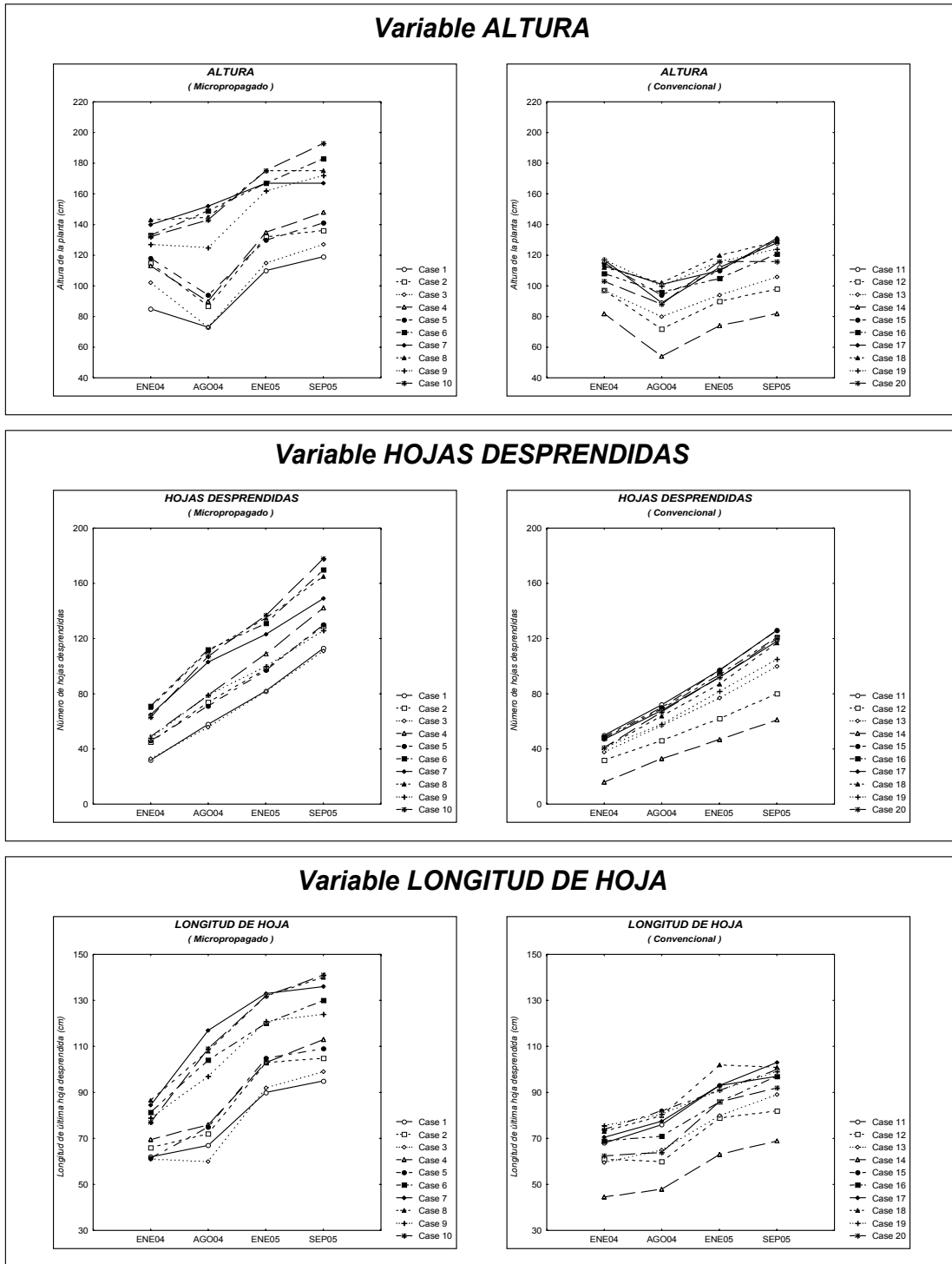


Figura 4.6. Gráficas longitudinales de las variables medidas en las plantas de agave de la región de Pihuamo los meses de Enero2004, Agosto2004, Enero2005 y Septiembre2005.

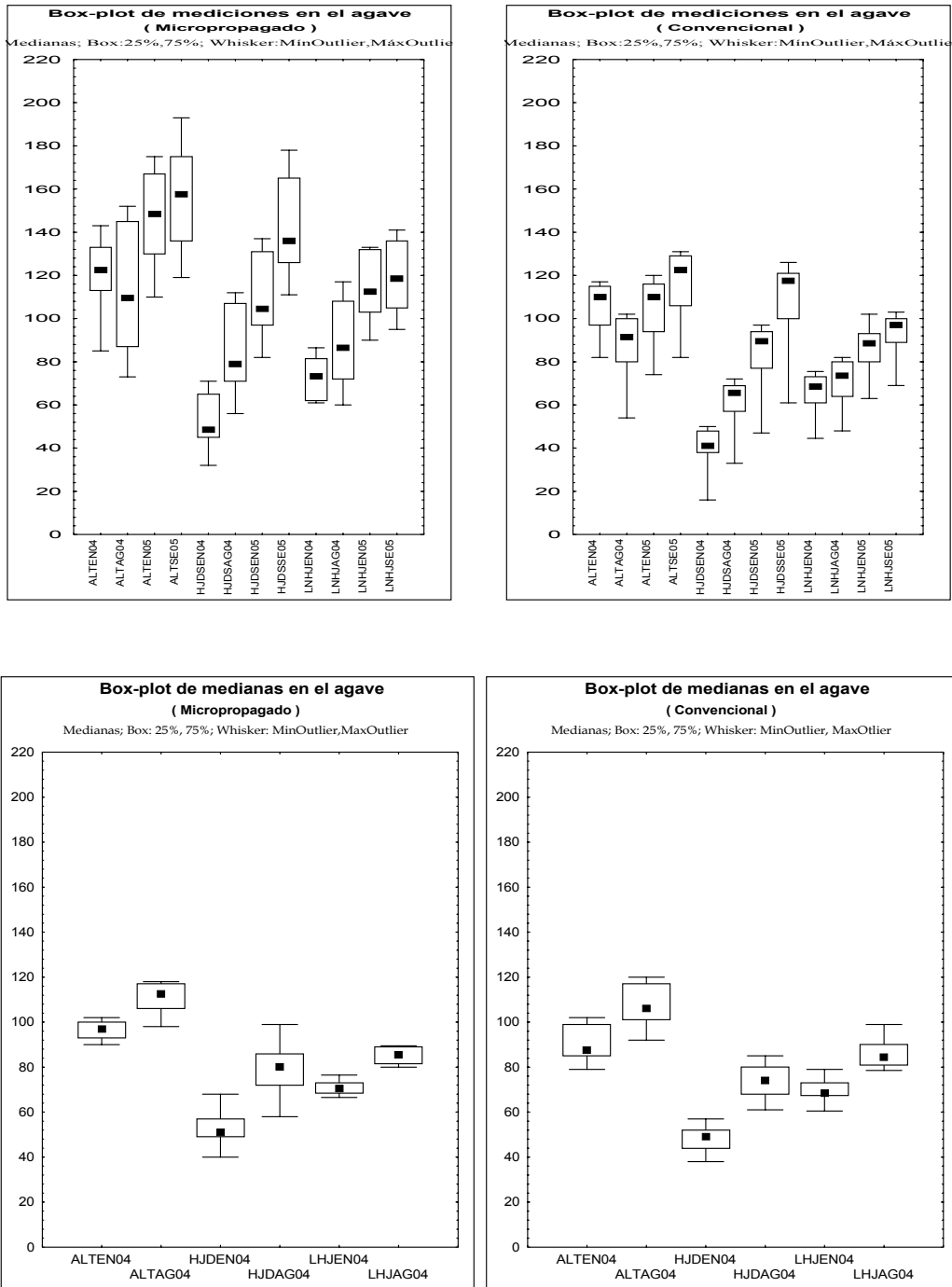


Figura 4.7. Gráficas de caja y bigotes de las variables medidas en las plantas de agave. Regiones de Pihuamo de Enero del 2004 a Septiembre del 2005 (arriba), Tepatlán en los meses de Enero y Agosto del 2004 (abajo).

los periodos de Agosto2004 a Enero2005 y de Enero2005 a Septiembre2005, las cuales muestran a su vez que en efecto la disminución en la altura sólo está presente para el periodo de tiempo de Enero2004 a Agosto2004. Así mismo, en dichas gráficas puede apreciarse como nuevamente el bloque de plantas del tipo micropropagado plantado en la ladera continúa sobresaliendo para cada una de las variables medidas.

En la Tabla 4.2 se pueden apreciar los resultados obtenidos al calcular las medias y desviaciones estándar para cada una de las variables en cada fecha. Cabe hacer notar que los valores obtenidos en esta tabla para los dos primeros meses son los mismos que fueron presentados en la Tabla 4.1 para la región de Pihuamo. Los resultados mostrados en dicha tabla permiten comprobar de manera numérica lo que a su vez fue reflejado en las gráficas de las Figuras 4.6 y 4.7, al observarse claramente un aumento en los valores de las medias de las variables *Hojas desprendidas* y *Longitud de hoja* a lo largo del tiempo, así como el problema en la variable *Altura* del mes de Enero al mes de Agosto del 2004 reflejado a su vez en la disminución de sus medias de una fecha a otra. Otro aspecto que se puede observar es la amplia variabilidad en las plantas del tipo micropropagado, presente en los valores de las desviaciones estándar, la cual es mayor a la correspondiente en el tipo convencional para cada una de las variables en cada fecha, sin embargo, no debe perderse de vista que en las plantas del tipo micropropagado influye en gran medida el bloque de plantas en la ladera.

Para finalizar, la Figura 4.7 presenta las gráficas de caja y bigotes correspondientes a las mediciones de cada tipo de planta para la región de Pihuamo (Figura 4.7, arriba). Cabe hacer notar que en esta misma figura se incluyen las gráficas correspondientes a las plantas de la región de Tepatlán (Figura 4.7, abajo), donde estas últimas fueron comentadas a su vez en la sección anterior. En lo que respecta a las plantas de Pihuamo, de las gráficas se puede apreciar para ambos tipos de planta una tendencia positiva en las variables *Hojas desprendidas* y *Longitud de hoja* en los valores de sus medianas a lo largo del tiempo, mientras que en la variable *Altura* (como se ha mencionado ya anteriormente) se muestra una disminución de sus valores de Enero a Agosto del 2004. Otro punto que es claro de estas gráficas es la amplia variabilidad presente en cada una de las variables para cada fecha, la cual es notablemente mayor en las plantas del tipo micropropagado debido en gran parte a la diferencia en las magnitudes de sus variables entre sus dos bloques de plantas, como fue puntualizado con anterioridad.

4.4 Diagnóstico de los datos.

El objetivo de los análisis exploratorios presentados a lo largo de este capítulo ha sido el medir la calidad y consistencia de los datos proporcionados por la empresa. Este objetivo es necesario para encarar el objetivo general mencionado desde la introducción del capítulo 1 y descrito más detalladamente en el capítulo 2. A continuación se presentan algunas conclusiones y sugerencias establecidas a partir de lo observado en los análisis realizados a lo largo de este capítulo.

4.4.1 Conclusiones.

Se desprende, a partir de los análisis, que existen factores en el diseño original para la elección de las plantas de la muestra que no fueron considerados. Entre éstos se encuentra el tipo de suelo donde estaba plantado cada bloque de plantas, ya que como se mencionó en secciones anteriores, las plantas del tipo micropropagado de la región de Pihuamo contaban con características de suelo distintas a las otras regiones, al tener un bloque sobre terreno plano y otro sobre una ladera de alta pedregosidad. Esto repercutió de manera indirecta en las mediciones, ya que la aplicación de un herbicida preemergente (Combine) retrasó el crecimiento de las plantas en terreno plano, pero no a las plantas de tipo micropropagado en la ladera, que al estar sobre un terreno inclinado permitió que el herbicida fuese lavado y por lo tanto tuviera un efecto menor.

Un problema adicional encontrado en los datos fue la disminución en las mediciones de alturas en las plantas de la región de Pihuamo, lo cual está presente en dicha variable para el periodo de Enero a Agosto del 2004. Sobre este problema, según lo comentado por el personal de la empresa, no se tiene una explicación concisa sobre las posibles causas que lo originaron.

Otro problema más en los datos se manifestó en las gráficas de correlación presentadas en la Figura 4.5, específicamente en la correspondiente al tipo micropropagado de la región de Tepatitlán, donde se pudieron apreciar correlaciones negativas para algunas de las variables. Esto se debe en gran medida a que algunas de las plantas muestran una asociación negativa en sus variables, al tener por ejemplo plantas que son altas pero con pocas hojas desprendidas y longitud de hoja pequeña, y a su vez plantas pequeñas con muchas hojas desprendidas y longitudes de hoja grandes. Por ejemplo,

compárense las mediciones obtenidas en las plantas 1 y 2 del tipo micropropagado en la región de Tepatitlán.

Cabe mencionar que, según lo comentado por el personal de *Tequila Sauza*, en las plantas del tipo micropropagado de la región de Tepatitlán estuvieron presentes dos factores que en general no se tuvieron para el resto de plantas. Uno de éstos fue la utilización de la técnica agrícola denominada *aporque* (acercamiento de suelo a la línea de cultivo), la cual sólo fue aplicada a las plantas referidas. El otro factor es que a las plantas de Tepatitlán se les aplicó al menos un par de herbicidas, Hyvar y Faena, distintos al que fue aplicado en la región de Pihuamo (Combine). Se sabe que estos dos herbicidas también pueden provocar daños en las plantas, incluso, en el caso de Faena puede provocar marchitez en las plantas debido al daño causado a nivel raíz.

En resumen, el diseño empleado, a partir del cual se obtuvieron los datos, no garantizó la igualdad de condiciones en aspectos como tipos de suelo, agroquímicos aplicados y la técnica agrícola empleada, y la falta de control sobre esos factores tuvo un impacto decisivo en el comportamiento de las mediciones tomadas en las plantas.

En consecuencia, el uso de esos datos para propósitos inferenciales es dudoso, y las conclusiones que pudieran obtenerse son, en el mejor de los casos, de carácter exploratorio.

En la sección 4.4.2 presentada a continuación se dan algunas sugerencias y recomendaciones, que de ser tomadas en cuenta para el diseño de un nuevo experimento, se espera generen nuevos datos que sean de utilidad para el cumplimiento del objetivo planteado.

En el presente caso, aún cuando no recomendamos emplear los datos disponibles para obtener conclusiones definitivas sobre el estudio comparativo, sí podemos mostrar en el siguiente capítulo el uso de las metodologías propuestas y descritas en el capítulo anterior, utilizando para ello datos simulados obtenidos a partir de los datos reales disponibles, al considerar dos distintos escenarios: uno en donde el crecimiento es similar entre ambos tipos de planta, y otro en donde el crecimiento es distinto. La manera en como se tomaron en cuenta las características presentes en las plantas de agave de las regiones de Pihuamo y Tepatitlán para llevar a cabo las simulaciones será detallada a su vez en el siguiente capítulo.

4.4.2 Sugerencias y recomendaciones.

Es importante considerar para el diseño de experimentos futuros, enfocados al análisis del crecimiento de plantas de agave, que la muestra de plantas sea seleccionada en *condiciones lo más similares posibles* (de acuerdo con lo descrito por C. L. Beadle en Coombs et al, 1985, pag. 24), sobretodo en lo concerniente a los siguientes factores: tipos de terrenos sobre el que se encuentran las plantas, agroquímicos aplicados y técnicas agrícolas empleadas. Esto con la intención de eliminar de los análisis efectos de ruido que afecten las mediciones obtenidas.

Otro factor importante, aunque no discutido en el presente estudio, es la frecuencia temporal para la toma de mediciones, y que en este caso ha sido de un periodo de aproximadamente 6 meses entre una medición y otra. Una recomendación es *tomar mediciones en periodos de tiempo más cortos*, sugiriéndose tomar mediciones *al menos cada mes* (o *cada dos meses a lo más*). Algunas ventajas de contar con mediciones más periódicas serían, por ejemplo, la detección más temprana de cambios en las condiciones establecidas en el experimento (por ejemplo, se podrían detectar problemas como el ocurrido en la variable *Altura* de las plantas de Pihuamo), así como poder observar el comportamiento de las plantas en distintas temporadas climáticas del año. Desde el punto de vista teórico, esto permitiría obtener mejores inferencias estadísticas en el estudio longitudinal.

Finalmente, con el propósito de obtener una buena calidad en las estimaciones de las matrices de correlación y los parámetros involucrados en las metodologías propuestas es necesario *aumentar el número de plantas seleccionadas en la muestra*, quizás (si así se desea) conservando el estilo de muestreo por bloques. La recomendación específica es considerar un *mínimo de 30 plantas por cada bloque* (considerando que los bloques pueden quedar definidos por el conjunto de factores considerados), con lo cual se garantizarían mejores estimaciones y ajustes distribucionales (Mead et al., 2003) en las metodologías estadísticas empleadas.

4.5 Algunas conclusiones preliminares derivadas a partir de un estudio piloto de crecimiento con base en los datos disponibles.

A raíz de lo observado en los análisis presentados a lo largo de este capítulo, en secciones anteriores se dio como recomendación que los datos disponibles

no fuesen empleados para responder de manera contundente a la pregunta de interés planteada por la empresa. Sin embargo, estos datos pueden ser considerados como una muestra piloto, donde su análisis, bajo un estricto carácter exploratorio, permitiría vislumbrar directrices posibles para futuros proyectos donde se buscara continuar analizando el crecimiento de estos dos tipos de planta. Aún así, debe tomarse en consideración que las conclusiones derivadas a partir de dichos análisis pudiesen llegar a ser, en el peor de los casos, opuestas a las obtenidas de un análisis con datos recabados a partir de un nuevo diseño bajo condiciones más controladas, ya que como se ha mencionado anteriormente, los datos disponibles presentan problemas inherentes en el diseño del experimento que afectaron a su vez a las mediciones tomadas.

De los análisis realizados en las dos primeras secciones de este capítulo, fue posible apreciar algunas diferencias en el comportamiento de ambos tipos de planta en lo concerniente a su crecimiento. Un primer ejemplo de tales diferencias es apreciable en los valores promedio de las variables, donde se pudo observar que en general son mayores en el tipo micropropagado en contraste con los del tipo convencional, lo cual sucede tanto en Pihuamo como en Tepatitlán, y siendo más marcada la diferencia en Pihuamo, en gran parte a consecuencia de las plantas del tipo micropropagado plantadas en la ladera. Esta diferencia en las medias de ambos tipos de planta, a pesar de la presencia de las plantas en la ladera, apunta en la dirección de que en efecto pudiesen esperarse diferencias reales en el crecimiento de ambos tipos de planta en el campo, lo cual podría estudiarse con mayor profundidad de contar con mediciones provenientes de una muestra de plantas seleccionada bajo condiciones más controladas.

Otra diferencia observada entre los dos tipos de planta, es en lo concerniente a la variabilidad, donde se pudo apreciar para las variables *Altura* y *Hojas desprendidas*, que la variabilidad en las plantas de la región de Tepatitlán tiende a ser menor en las del tipo micropropagado en comparación con las del tipo convencional, y al igual ocurre en la región de Pihuamo si no se considera para la estimación el bloque de plantas en la ladera (estas estimaciones ya no son incluidas en este trabajo). Esto indicaría que en la práctica podría esperarse que la variabilidad en el tipo convencional sea mayor que las del tipo micropropagado, sin embargo, al igual que en el caso anterior (medias), esto podría comprobarse con mayor certidumbre empleando nuevos datos recabados bajo un diseño donde se procurase evitar al máximo posible factores de ruido en la toma de mediciones.

Un aspecto más que se puede observar de los datos analizados es que las variables medidas a las plantas (*Altura, Número de hojas desprendidas y Longitud de hoja*), en efecto permiten describir el crecimiento que van teniendo éstas. Incluso, con su seguimiento a lo largo del tiempo, es posible detectar factores que favorecen o dañan el crecimiento de las plantas (como fue el caso de la disminución de alturas o los efectos de los herbicidas empleados).

De igual forma, de los análisis realizados, se pudo apreciar que en general las matrices de correlación de las variables obtenidas a partir de las mediciones tomadas a las plantas presentan estructuras similares (salvo el caso de las plantas del tipo micropropagado de Tepatitlán, tal y como se puntualizó en la sección 4.2), por lo cual, sería factible esperar que esto mismo sucediese al contar con nuevas mediciones tomadas empleando un nuevo diseño experimental.

Si se calcularan los índices morfológicos tomando como base la metodología descrita en la sección 3.2 y empleando para ello una ponderación de las matrices de correlación presentadas en las Figuras 4.4 y 4.5 (para más detalle ver el procedimiento empleado en la sección 5.3.2, el cual es análogo al aquí utilizado - para este caso no se considera la matriz correspondiente a las plantas del tipo micropropagado en Tepatitlán), las gráficas longitudinales de dichos índices se verían como las presentadas en la Figura 4.8. En esta figura se pueden apreciar diferencias claras en el comportamiento de ambos tipos de planta, al notarse en la región de Pihuamo que las plantas del tipo micropropagado presentan mayores magnitudes de sus índices en comparación con las del tipo convencional. Por su parte, las plantas de la región de Tepatitlán también presentan diferencias de comportamiento en ambos tipos de planta, sin embargo, el impacto del problema existente en las plantas del tipo micropropagado (descrito en la sección 4.2 y reflejado claramente en los índices obtenidos en Enero del 2004) ocasiona que los dos tipos de planta en dicha región no cuenten con un punto de comparación razonable para el estudio, por lo tanto, se continuará el análisis solamente con los datos correspondientes a la región de Pihuamo.

Con los índices morfológicos obtenidos para las plantas de la región de Pihuamo, es posible calcular para cada una de las plantas las tasas de crecimiento (temporal y relativa promedio) en cada uno de los periodos de tiempo en cuestión, para comparar así el comportamiento de ambos tipos de planta como función del tiempo en lo que respecta a su rapidez de crecimiento. La Figura 4.9 presenta precisamente las tasas de crecimiento

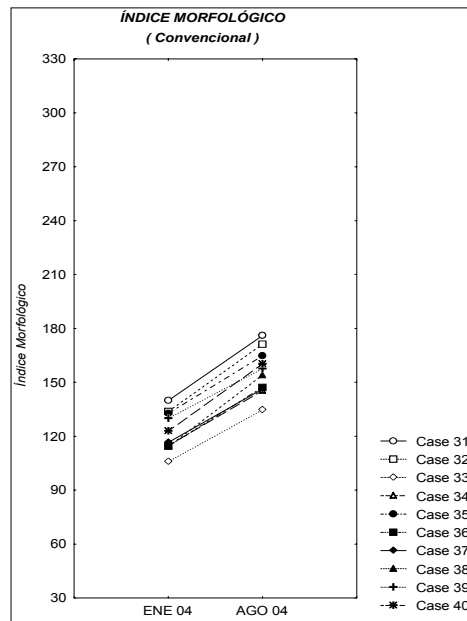
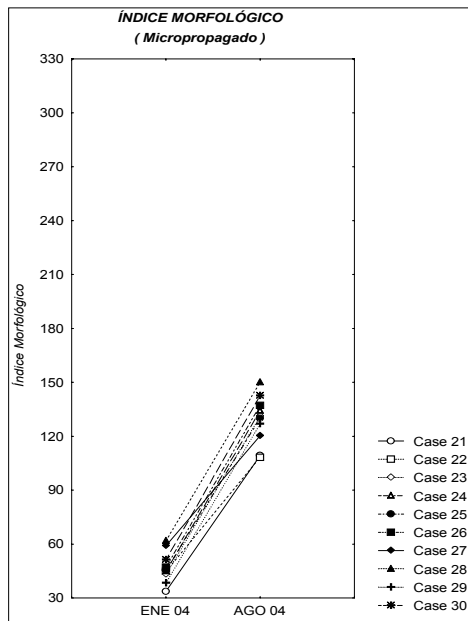
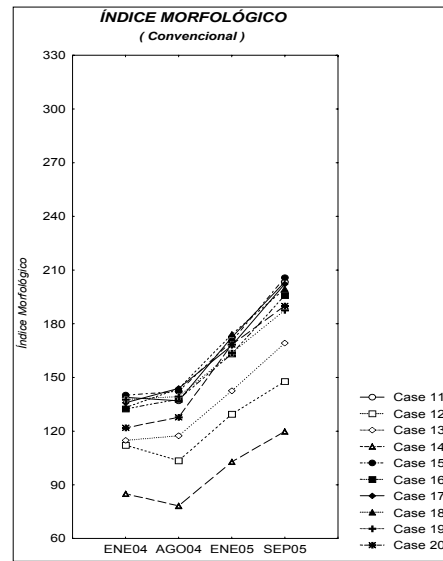
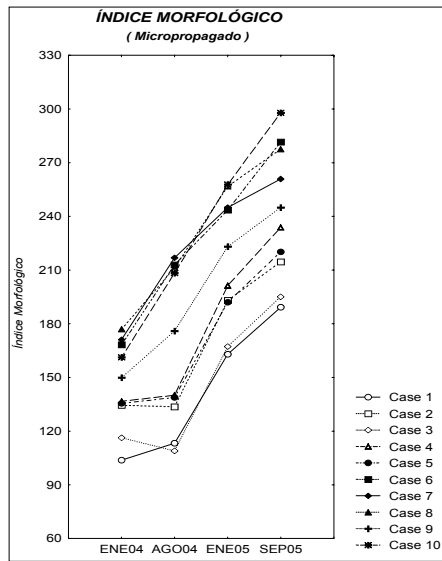


Figura 4.8. Figura de arriba: Gráficas longitudinales de los índices morfológicos de las plantas de agave en la región de Pihuamo medidas en los meses de Enero 2004, Agosto 2004, Enero 2005, y Septiembre 2005, descritas por tipo de planta. Figura de abajo: Gráficas longitudinales de los índices morfológicos de las plantas de agave en la región de Tepatlán medidas en los meses de Enero y Agosto 2004, descritas también por tipo de planta.

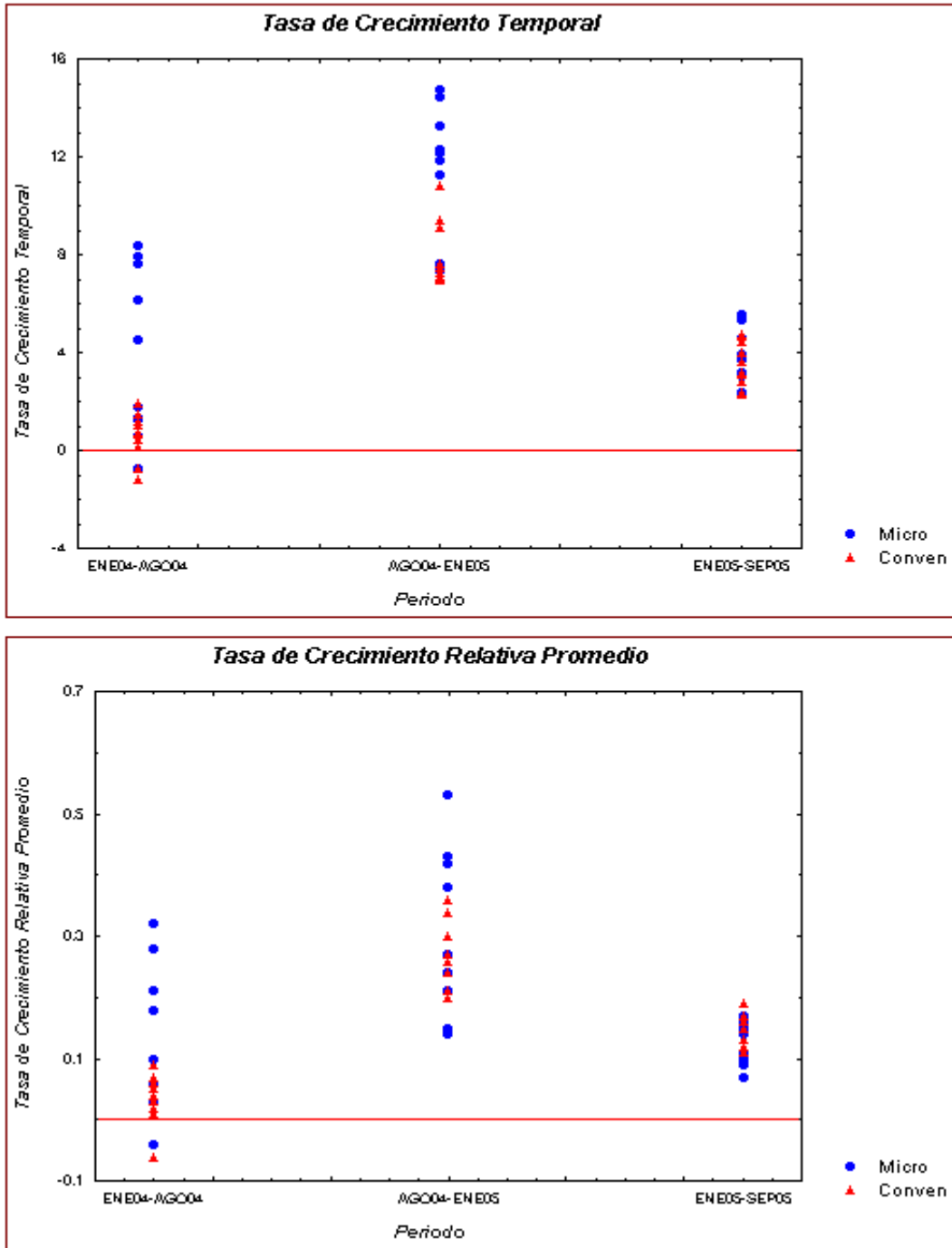


Figura 4.9. Tasas de crecimiento temporales (arriba) y relativas promedio (abajo) de los índices morfológicos correspondientes a las plantas de agave de la región de Pihuamo, graficadas en cada periodo de tiempo para ambos tipos de planta.

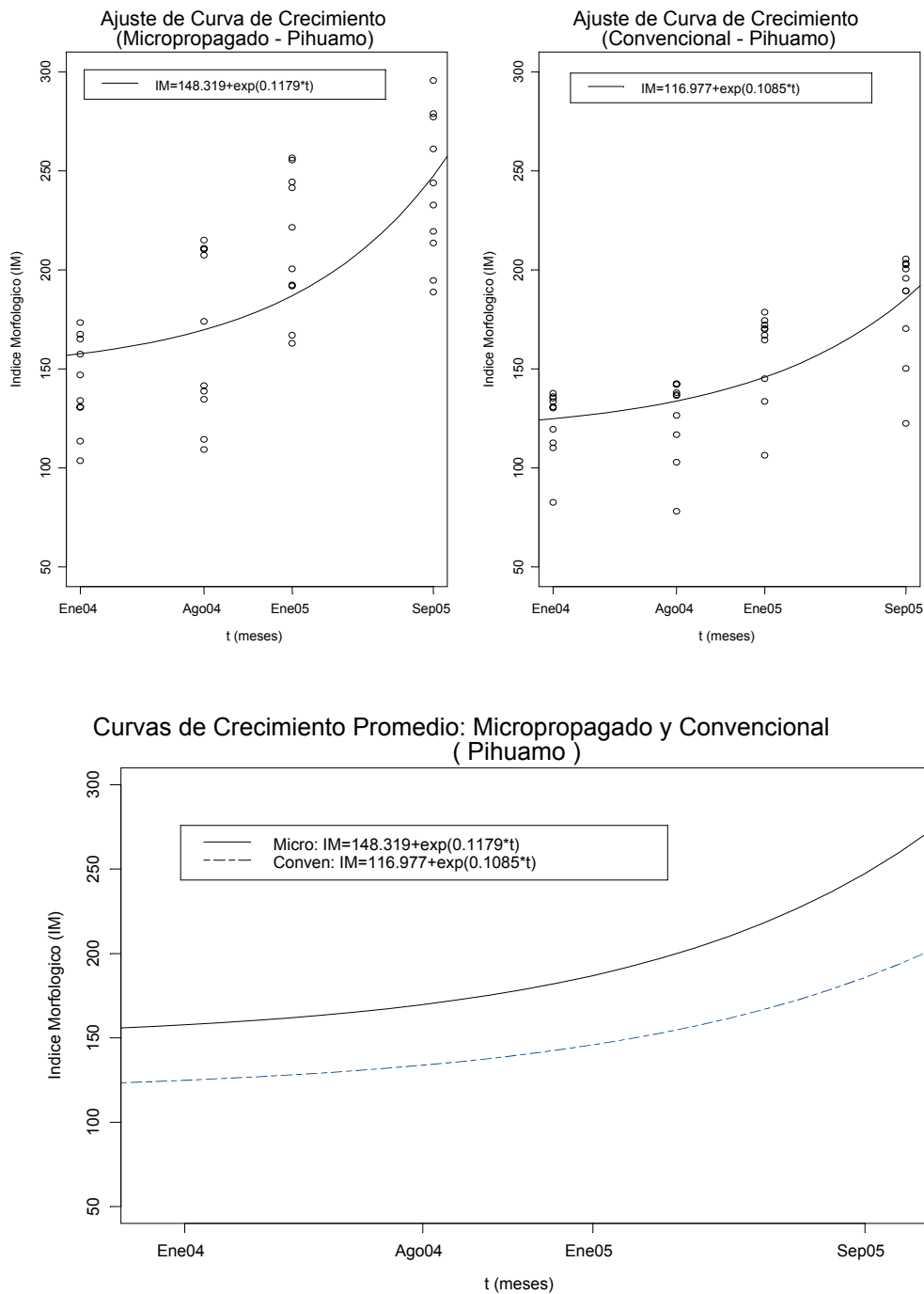


Figura 4.10. Curvas de crecimiento promedio ajustadas con base en los índices morfológicos correspondientes a las plantas de agave de la región de Pihuamo.

(temporal y relativa promedio) que se obtendrían con base en los índices morfológicos calculados. En dichas gráficas se puede apreciar claramente que en el periodo de Agosto2004-Enero2005 las plantas presentaron un mayor crecimiento en comparación con los otros dos periodos de tiempo. Así mismo, destaca en el periodo de Enero-Agosto2004 algunas plantas que presentan un decremento en sus índices morfológicos de una fecha a otra (siendo a su vez un reflejo del problema existente en la variable Altura para este mismo periodo de tiempo). Otro aspecto que destaca en estas gráficas es que en los dos primeros periodos de tiempo algunas plantas del tipo micropropagado (en su mayoría correspondientes a las plantas de la ladera) muestran mayores magnitudes en sus tasas de crecimiento (tanto temporales como relativas promedio), mientras que por otro lado, en el último periodo de tiempo no se aprecian diferencias entre las tasas de ambos tipos de planta. De este modo, con estas gráficas como un primer análisis sobre el comportamiento de las tasas de crecimiento, es posible mostrar si existen gráficamente diferencias en el crecimiento de ambos tipos de planta en lo que a sus tasas se refiere, lo cual seguido de este análisis gráfico podría comprobarse mediante procedimientos estadísticos formales (como el descrito en la sección 3.4.1), el procedimiento a seguir para esto sería análogo al desarrollado en las secciones 5.3.2.1 y 5.4.2.1 en el siguiente capítulo. Para los datos disponibles este último análisis formal ya no es realizado, al considerarse que éste no sería relevante a causa de los problemas existentes en los datos en cuestión.

Con lo anterior, se pudo apreciar que los datos muestran en principio algunas diferencias en el comportamiento de las tasas de crecimiento de ambos tipos de planta. Ahora, si se construyeran las curvas de crecimiento promedio (según lo descrito en la sección 3.4.2) para cada tipo de planta, con base en los índices morfológicos correspondientes a las plantas de la región de Pihuamo, las curvas se verían como las presentadas en la Figura 4.10. En dicha figura se pueden apreciar los ajustes correspondientes a cada tipo de planta, así como el contraste entre las dos curvas de crecimiento promedio de cada tipo de planta. De las gráficas referidas se puede apreciar que de estos primeros ajustes las plantas del tipo micropropagado presentan una curva que se encuentra por arriba de la correspondiente al tipo convencional, indicando así un mayor crecimiento en las plantas del tipo micropropagado para el lapso de tiempo en cuestión. Esto mismo puede apreciarse al contrastar las estimaciones de los parámetros de ambas curvas. Cabe mencionar que estos ajustes son realizados sin considerar aspectos como posibles estructuras en los residuales, correlaciones seriales entre las mediciones, gran

heterogeneidad entre las plantas, etc., mismo que podría considerarse con modelos como los sugeridos en la sección 3.4.3. No obstante, estos primeros ajustes permiten obtener una idea general sobre la manera en que se están comportando las curvas de crecimiento promedio de ambos tipos de planta (con base en los datos analizados), así como detectar posibles estructuras que pudiesen estar presentes en los datos en cuestión.

De los análisis presentados en esta sección se pudo apreciar que los datos disponibles mostrarían diferencias en el crecimiento de ambos tipos de planta, destacando un mayor crecimiento en las plantas del tipo micropropagado en contraste con las del tipo convencional, con base en lo observado de las curvas de crecimiento.

Como se mencionó al inicio de esta sección, las conclusiones anteriores deben considerarse bajo un estricto carácter exploratorio, al tener en los datos analizados la presencia de factores de ruido que no garantizaron igualdad de condiciones durante el desarrollo del experimento y que repercutieron en algunos casos de manera directa en las mediciones tomadas en las plantas. En el siguiente capítulo se llevará a cabo un estudio de simulación, donde a su vez son ilustradas con más detalle las técnicas presentadas en la sección 3.4, utilizando para ello un par de conjuntos de datos simulados los cuales se generaron libres de la presencia de factores como los presentes en los datos disponibles. De contarse con nuevos datos obtenidos a partir de un diseño bajo condiciones más controladas (tomando en cuenta las sugerencias dadas en la sección 4.4.2), los análisis de crecimiento podrían realizarse de manera análoga a los presentados en el siguiente capítulo, donde a raíz de lo observado en esta sección se esperaría que bajo condiciones más controladas los datos obtenidos permitan llevar a cabo el análisis comparativo de crecimiento deseado y responder así a la pregunta inicial sobre el crecimiento de ambos tipos de planta.

Capítulo 5

Comparación del crecimiento de dos tipos de planta empleando una simulación de datos

En este capítulo se presenta un estudio de simulación para ilustrar la utilidad de las metodologías propuestas en el capítulo 3, contrastando dos escenarios distintos: uno en donde los datos simulados corresponden a un patrón de crecimiento similar entre ambos tipos de planta, y otro en donde el crecimiento es diferente; en ambos casos la estructura de los datos (variables, regiones y número de mediciones) es similar a la de los datos proporcionados por la empresa. A partir de los análisis realizados a estos conjuntos de datos, se podrá observar que las metodologías propuestas son efectivas para detectar diferencias en el crecimiento de las plantas, en el caso de que éstas diferencias estén presentes y sean significativas.

5.1 Introducción.

Este capítulo tiene como finalidad ilustrar la utilidad de las metodologías descritas a lo largo del capítulo 3, empleando para ello datos libres del impacto negativo de factores no controlados. Los datos que se usarán en este capítulo se generaron por técnicas de simulación, buscando que dichos datos cuenten con una estructura similar a la de los datos reales. Se trata de que los datos obtenidos a partir de las simulaciones representen de manera razonable el comportamiento que tienen las plantas de agave en el campo, utilizando para esto algunos de los valores medios y varianzas obtenidos de las primeras mediciones tomadas a las plantas de agave en las regiones de

Pihuamo y Tepatitlán, así como algunas de las matrices de correlación calculadas con base en dichas mediciones. Todo esto se describirá con más de detalle en la sección siguiente.

5.2 Obtención de los datos simulados.

Para la simulación se plantearon dos escenarios: uno donde los datos son simulados suponiendo igualdad en el crecimiento de ambos tipos de planta (micropropagado y convencional) y otro donde el crecimiento es distinto entre ambos tipos. A dichos escenarios se les denominará Escenario 1 y Escenario 2, respectivamente.

La estructura de los datos simulados en cada uno de los escenarios se pretende sea similar a la de los datos proporcionados por la empresa *Tequila Sauza*, en lo concerniente a variables medidas a las plantas, regiones consideradas, tipos de planta y número de mediciones tomadas a lo largo del tiempo. De esta forma, la estructura para cada uno de los escenarios considerados será la siguiente:

1. Tres variables medidas a las plantas: *Altura*, *Número de hojas desprendidas* y *Longitud de hoja*.
2. Dos regiones geográficas consideradas: Región 1 y Región 2.
3. Dos tipos de plantas de agave: micropropagado y convencional.
4. Cuatro mediciones tomadas a lo largo de tiempo en periodos igualmente espaciados.

Cabe mencionar que aunque los datos simulados no corresponden a mediciones tomadas directamente de plantas de agave, por cuestiones de facilidad en la redacción y lectura de este capítulo, en lo subsecuente nos referiremos a dichos datos como si en realidad las mediciones hubiesen sido tomadas de plantas reales. Nótese que el supuesto anterior no está tan lejos de la realidad, ya que a final de cuentas estos datos fueron simulados con base en valores iniciales de las plantas de agave de Pihuamo y Tepatitlán. Por tanto, las regiones denominadas como Región 1 y Región 2, cuentan con características similares a las de las regiones de Pihuamo y Tepatitlán, respectivamente. Además, en cada planta se tendrán mediciones de sus variables en los tiempos T0, T1, T2 y T3, siendo éstos igualmente espaciados, y teniendo a T0 como la medición inicial en las plantas.

Como se mencionó en la sección anterior, buscando que los datos obtenidos a partir de las simulaciones sean lo más razonables posibles, y las magnitudes de sus mediciones estén acordes con lo que se presenta comúnmente en campo para los dos tipos de planta bajo estudio (micropropagado y convencional), incluso en regiones con diferentes características (como Pihuamo y Tepatitlán), las simulaciones se llevan a cabo tomando en cuenta la información presente en los datos recabados a partir de las plantas de agave en las regiones de Pihuamo y Tepatitlán, considerando para ello los valores medios de sus variables correspondientes a la primer fecha en que se tomaron mediciones. Así, algunos de los valores a considerar como iniciales para la simulación corresponden a algunos de los presentados a su vez en la Tabla 4.1 del capítulo anterior.

La idea general seguida para la simulación de los datos fue la siguiente: primero se establecieron los valores iniciales para las medias de cada variable (*Altura, Número de hojas desprendidas y Longitud de hoja*), luego, con ayuda de la primer componente principal (obtenida a partir de una matriz de correlación común) se calcularon los índices morfológicos medios correspondientes a las medias iniciales, así, los índices medios obtenidos corresponderían a cierta medición inicial T_0 en la curva de crecimiento media, de acuerdo con el modelo exponencial descrito en (3.15), con dicha curva, se calculan los índices correspondientes a los tiempos T_1 , T_2 y T_3 , y a los índices obtenidos en cada fecha se les aplica la transpuesta de la primer componente principal para obtener así los correspondientes valores medios de las variables en los tiempos T_1 , T_2 y T_3 , finalmente, utilizando un generador de números aleatorios (normal multivariado) se generan las muestras correspondientes en cada tiempo para cada tipo de planta, de acuerdo con los valores medios obtenidos y la matriz de covarianzas obtenida con base en la matriz de correlación usada para el cálculo de los componentes principales.

A continuación se describe con más detalle el procedimiento anterior, específicamente para la simulación de los datos en el Escenario 1. Para los datos simulados con el Escenario 2 el procedimiento fue análogo cambiando solamente los valores de las medias iniciales y las tasas de crecimiento correspondientes.

Así mismo, el procedimiento se describe solamente para la Región 1, al ser análogo para la Región 2:

1. Se establecen valores iniciales de las medias de cada variable para cada

Tabla 5.1. Valores iniciales considerados para la simulación de los datos en los Escenarios 1 y 2.

Escenario	Región	Tipo	Estadísticas Descriptivas	Valores Iniciales			Tasa de Crecimiento
				Altura	Hojas Desp.	Long. Hoja	
1	1	Micro	Media	96.6	53	70.8	0.1085
			Desv. Estand.	1.5	2	1.5	
	Conven	Media	96.6	53	70.8	0.1085	
		Desv. Estand.	2.2	3	2.2		
2	Micro	Media	106.1	40	65.8	0.1085	
		Desv. Estand.	1.5	2	1.5		
Conven	Media	106.1	40	65.8	0.1085		
	Desv. Estand.	2.2	3	2.2			
2	1	Micro	Media	96.6	53	70.8	0.1179
			Desv. Estand.	1.5	2	1.5	
	Conven	Media	96.6	53	70.8	0.1085	
		Desv. Estand.	2.2	3	2.2		
2	Micro	Media	106.1	40	65.8	0.1179	
		Desv. Estand.	1.5	2	1.5		
Conven	Media	106.1	40	65.8	0.1085		
	Desv. Estand.	2.2	3	2.2			

tipo de planta, ver Tabla 5.1, donde a los vectores con dichas medias los denotaremos como:

$$\mu_{0M} = [\mu_{0Alt}, \mu_{0HDes}, \mu_{0LHoja}]_{3 \times 1} \quad (\text{micropropagado}), \quad (5.1)$$

$$\mu_{0C} = [\mu_{0Alt}, \mu_{0HDes}, \mu_{0LHoja}]_{3 \times 1} \quad (\text{convencional}). \quad (5.2)$$

- Utilizando el vector propio correspondiente a la primer componente principal obtenida de la matriz de correlación común, R , calculada como una ponderación de las matrices de correlación de las Figuras 4.4 (ambos tipos) y 4.5 (sólo tipo convencional) del capítulo anterior, desplegada a su vez en la expresión (5.6), se obtuvieron los índices morfológicos promedio:

$$[y_{0M}]_{1 \times 1} = [a_1]_{1 \times 3} [\mu_{0M}]_{3 \times 1} \quad \text{y} \quad [y_{0C}]_{1 \times 1} = [a_1]_{1 \times 3} [\mu_{0C}]_{3 \times 1} \quad (5.3)$$

- Luego, considerando dichos índices promedio como valores iniciales en la curva exponencial descrita en la expresión (3.15), es decir, estos índices corresponderían al parámetro c en la curva exponencial, se calcularon los índices promedio para cada tipo de planta en cada uno de los tiempos T1 (=29 meses), T2 (=39 meses) y T3 (=49 meses),

evaluando la curva en cada uno de dichos tiempos. En cada caso se consideraron las tasas de crecimiento descritas en la Tabla 5.1. Los tiempos a evaluarse fueron elegidos a partir de los 19 meses (T0) debido a que las mediciones en la Tabla 5.1 fueron tomadas de plantas de agave en Pihuamo y Tepatitlán, las cuales tenían precisamente 19 meses de ser trasplantadas al campo en el momento de tomar la primer medición en Enero del 2004. A su vez, cabe mencionar que los valores iniciales de las tasas de crecimiento fueron establecidos con base en las curvas construidas en la sección 4.5 para los datos de Pihuamo, correspondientes al análisis piloto de crecimiento.

- Así, para cada tiempo T1, T2 y T3 fueron obtenidos sus correspondientes valores medios de las variables *Altura*, *Número de hojas desprendidas* y *Longitud de hoja*, empleando de nueva cuenta el vector propio correspondiente a la primer componente principal, multiplicando el traspuesto de éste por el índice medio para obtener los valores medios de las variables (a esto se conoce como inversión de las componentes principales, para más referencia ver Jackson, 1991, Sección 1.5.5), por ejemplo, para cada tipo de planta, en el tiempo T1, los valores medios de sus variables se calcularon con base en las siguientes expresiones:

$$[\mu_{1M}]_{3 \times 1} = [a_1]_{3 \times 1} [y_{1M}]_{1 \times 1} \quad \text{y} \quad [\mu_{1C}]_{3 \times 1} = [a_1]_{3 \times 1} [y_{1C}]_{1 \times 1} \quad (5.4)$$

- Finalmente, con los valores medios obtenidos en cada tiempo T0, T1, T2 y T3, se generaron números aleatorios correspondientes a una distribución normal multivariada $([\mu_{iM}]_{3 \times 1}, \Sigma_3)$ para las plantas del tipo micropropagado y $([\mu_{iC}]_{3 \times 1}, \Sigma_3)$ para las del tipo convencional, donde:

$$\Sigma_3 = \text{diag}(\sigma_{Alt}, \sigma_{HDes}, \sigma_{LHoja}) \times R \times \text{diag}(\sigma_{Alt}, \sigma_{HDes}, \sigma_{LHoja}). \quad (5.5)$$

La matriz de correlación empleada para la generación de los datos fue:

$$R = \begin{bmatrix} 1 & 0.9304 & 0.9621 \\ 0.9304 & 1 & 0.9117 \\ 0.9621 & 0.9117 & 1 \end{bmatrix}, \quad (5.6)$$

la cual fue obtenida con base en el procedimiento descrito en el punto 2 anteriormente.

Los datos que fueron generados a través de este procedimiento son mostrados en las Tabla 2 (Escenario 1) y 7 (Escenario 2) del Apéndice A.1. Cabe mencionar que los datos correspondientes a la variable *Número de hojas desprendidas* obtenidos del generador aleatorio fueron redondeados al entero más cercano al ser dicha variable una variable de conteo.

De aquí en adelante se supondrá que los datos generados para los Escenarios 1 y 2 en efecto corresponden a mediciones tomadas en plantas de agave de dos diferentes regiones geográficas (Región 1 y 2), en las cuales se consideran plantas de ambos tipos (micropropagado y convencional) con sus respectivas mediciones en las fechas T0, T1, T2 y T3.

5.3 Escenario 1: Poblaciones con crecimientos similares.

En esta sección se analizarán los datos correspondientes al Escenario 1, mostrados a su vez en la Tabla 2 del Apéndice A.1. Dichos datos corresponden a mediciones de las variables *Altura*, *Número de hojas desprendidas* y *Longitud de hoja*, tomadas de 40 plantas de agave de dos diferentes regiones geográficas (Región 1 y Región 2). Las mediciones se tomaron en las fechas T0, T1, T2 y T3, donde entre cada una de éstas se tiene una separación de 10 meses. De las 40 plantas, 20 corresponden a la Región 1 y 20 a la Región 2, donde a su vez en cada región 10 plantas pertenecen al tipo micropropagado y 10 al tipo convencional.

Como suele hacerse en todo análisis estadístico, iniciaremos realizando un análisis exploratorio de estos datos, mismo que es presentado a continuación.

5.3.1 *Análisis exploratorio de los datos.*

Como un primer análisis gráfico se presentan en la Figura 5.1 (Región 1) y 5.2 (Región 2) las gráficas longitudinales de estos datos, en las cuales se grafican cada una de las variables (*Altura*, *Número de hojas desprendidas* y *Longitud de hoja*) como función del tiempo. De éstas destaca un claro patrón de crecimiento (positivo) en las variables de las plantas de cada tipo y región, el cual se presenta de manera similar en ambos tipos de planta. Así mismo, las gráficas de las Figuras 5.3, 5.4 y 5.5, donde se muestra el cambio proporcional de cada variable para los distintos periodos de tiempo considerados (T0-T1, T1-T2 y T2-T3), permiten apreciar también el crecimiento positivo en las plantas a lo largo del tiempo. Esto se aprecia al tener que los conjuntos de

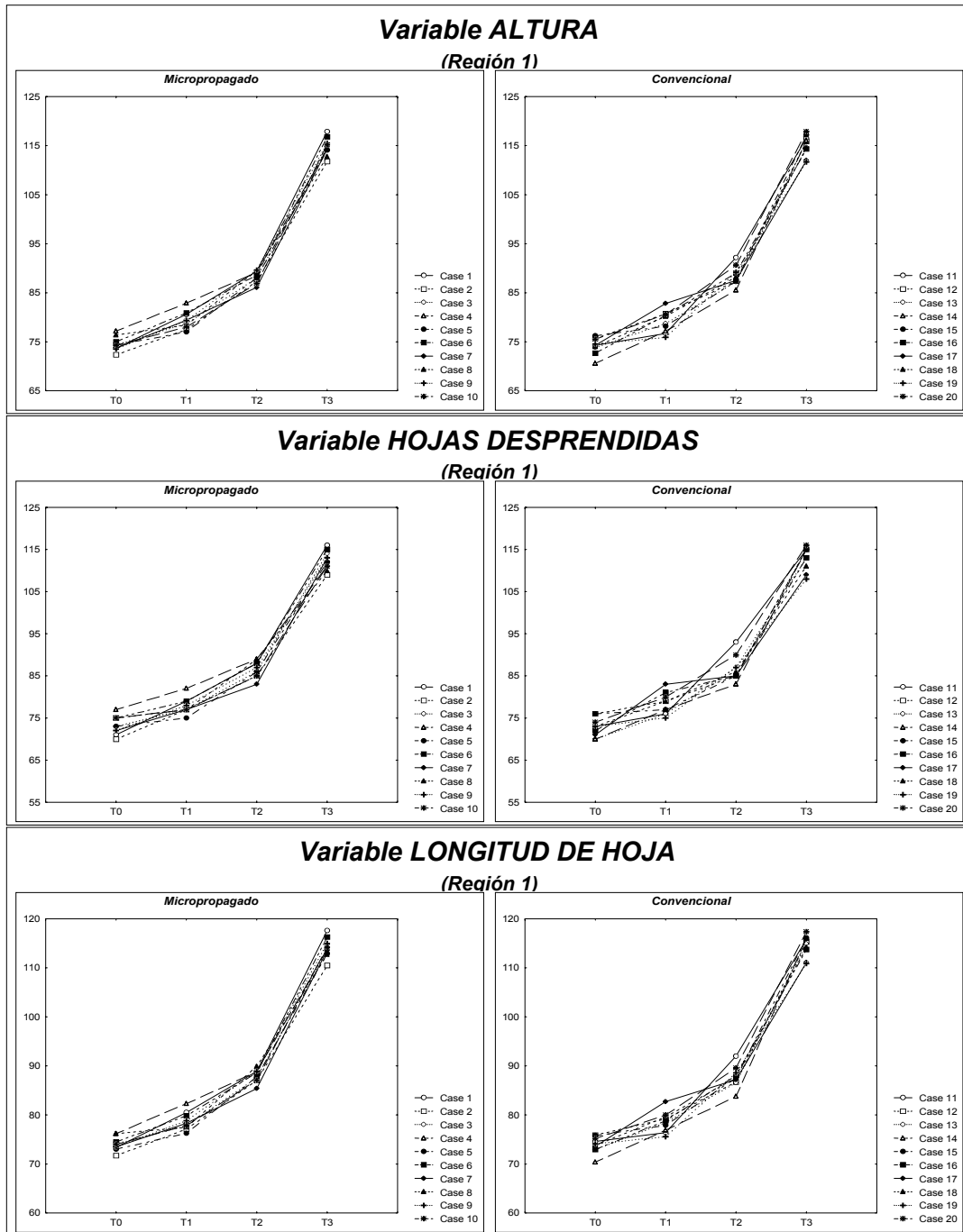


Figura 5.1. Gráficas longitudinales de las variables originales para las plantas de agave de la Región 1 en las fechas T0, T1, T2 y T3 (Escenario 1).

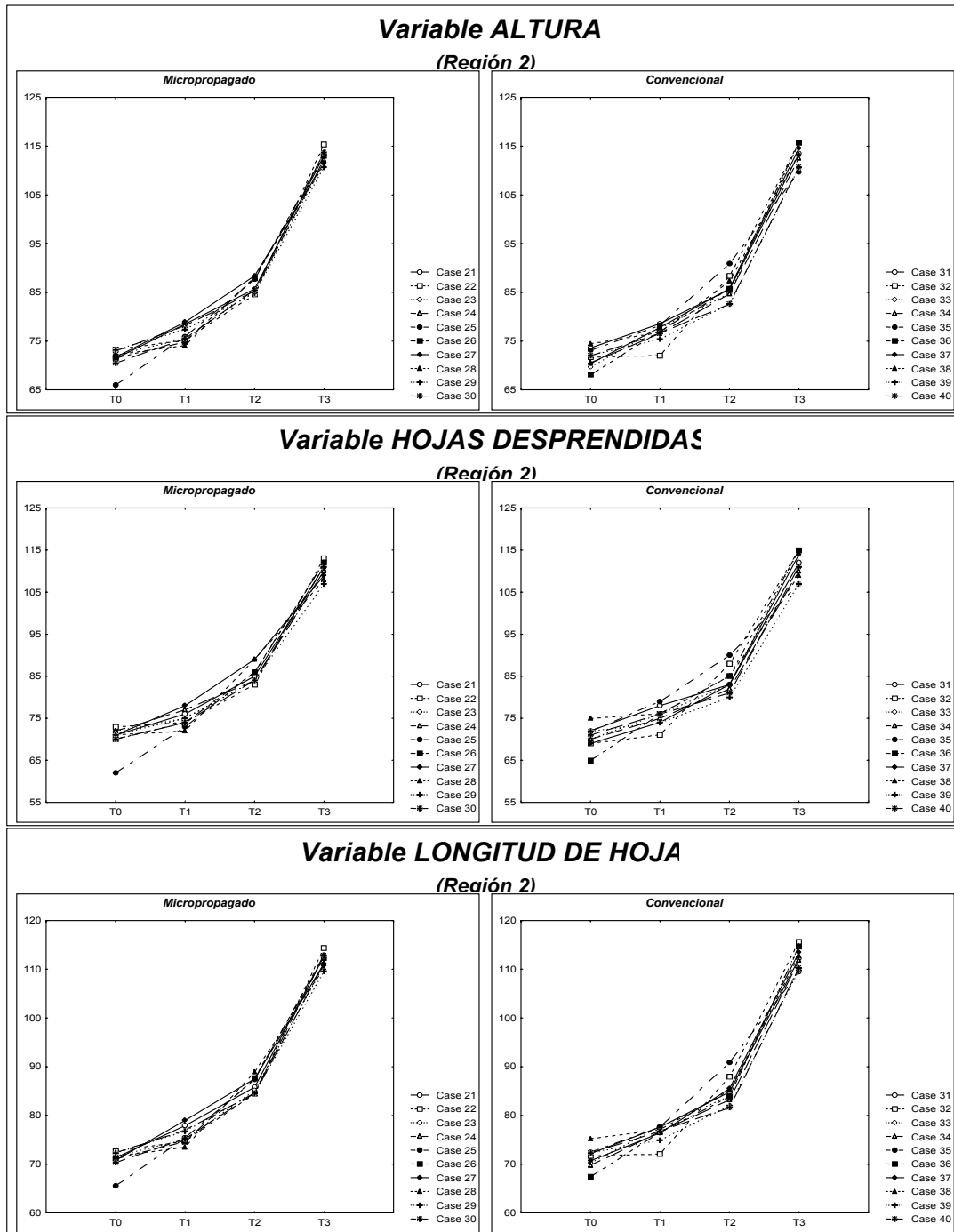


Figura 5.2. Gráficas longitudinales de las variables originales para las plantas de agave de la Región 2 en las fechas T0, T1, T2 y T3 (Escenario 1).

Tabla 5.2. Estadísticas descriptivas por región y tipo de planta para las mediciones en el agave tomadas en las fechas T0, T1, T2 y T3 en el Escenario 1.

REGION	TIPO	Estadísticas	ALTURA				HOJAS DESPRENDIDAS				LONGITUD DE HOJA			
		Descriptivas	T0	T1	T2	T3	T0	T1	T2	T3	T0	T1	T2	T3
1	0	Media	74.5	79.3	88.3	114.8	73	78	87	112	74.0	78.6	88.0	114.1
		Desv. Estándar	1.4	1.8	1.2	1.8	2	2	2	2	1.4	1.8	1.3	2.0
	1	Media	74.2	79.1	88.4	115.1	73	78	87	113	73.9	78.6	87.8	114.5
		Desv. Estándar	1.6	2.2	1.9	2.1	2	3	3	3	1.6	2.0	2.1	2.1
2	0	Media	71.5	76.4	86.4	112.6	70	75	85	110	71.0	75.9	86.1	111.8
		Desv. Estándar	2.1	1.7	1.5	1.4	3	2	2	2	2.0	1.7	1.6	1.4
	1	Media	71.6	76.7	85.9	113.2	71	76	84	112	71.4	76.4	84.9	112.5
		Desv. Estándar	1.9	1.9	2.5	2.3	3	2	3	3	2.1	1.7	2.8	2.2

Tipo 0 = Micropropagado

Tipo 1 = Convencional

datos se encuentran en cada caso por arriba de la recta de 45 grados, indicando así que de una fecha a otra en efecto se tiene un aumento en las magnitudes de las variables de cada planta. En algunos periodos de tiempo, los datos muestran una separación notablemente mayor de la recta de 45 grados, lo cual indicaría que en dicho periodo las plantas presentaron un mayor crecimiento en comparación con los periodos restantes. Un ejemplo de esto último puede apreciarse en las Figuras 5.1 a 5.5 para el periodo T2-T3, donde las plantas muestran un aumento notable en las magnitudes de sus mediciones de una fecha a otra. Esto mismo fenómeno puede ser detectado con el cálculo de las tasas de crecimiento, tal y como se verá más adelante en la sección 5.3.2.1.

Por otro lado, en la Tabla 5.2 de arriba se presentan las estadísticas descriptivas (medias y desviaciones estándar) correspondientes a los datos del Escenario 1, donde, las estimaciones se presentan (en cada región) para cada tipo de planta en cada una de las fechas en cuestión. De esta tabla se puede apreciar que se refleja de manera numérica lo observado en las gráficas 5.1 a 5.5, en lo respecta al aumento de las magnitudes en las variables con el paso del tiempo, lo cual puede apreciarse en las medias de las variables de cada tipo de planta y región al notar que éstas aumentan con el paso del tiempo. Otro punto a destacar de la tabla anterior es la similitud entre las magnitudes de las medias de ambos tipos de planta, lo cual ocurre por igual en ambas regiones, no por su parte, la desviación estándar, que como se puede observar es ligeramente mayor en las plantas del tipo convencional con respecto a las del tipo micropropagado salvo algunos casos donde no es así, por ejemplo: Región 1, *Hojas desprendidas*, T0; Región 2, *Altura*, T0; Región 2, *Hojas desprendidas*, T0 y T1; y Región 2, *Longitud de hoja*, T1. Finalmente, de

esta tabla también se puede apreciar que las magnitudes en las medias son mayores en la Región 1 en comparación con las de la Región 2, al contrastar para cada tipo de planta las correspondientes a una misma fecha en ambas regiones.

Una análisis gráfico adicional es presentado en la Figura 5.6 donde se presentan las gráficas de caja y bigotes (box-plot) de las variables de cada tipo de planta en cada región, donde se puede mostrar a su vez el comportamiento de éstas a lo largo del tiempo. De estas gráficas se puede observar claramente una tendencia positiva en el crecimiento de cada una de las variables como función del tiempo, así como un mayor crecimiento de las variables en el último periodo de tiempo, acorde con lo mencionado de los análisis anteriores. Así mismo se pueden observar algunos valores extremos presentes en ciertos tiempos para algunas de las variables.

Algunas conclusiones preliminares de estos primeros análisis son las siguientes: se observa que los datos muestran una sutil diferencia entre las plantas de ambas regiones en lo que respecta a los valores medios de sus variables (en general mayores medias en la Región 1), así como diferencias entre los tipos de planta en lo que respecta a sus desviaciones estándar (teniéndose en general mayores desviaciones para el tipo convencional). A continuación se presenta un análisis de crecimiento con base en lo descrito en el capítulo 3.

5.3.2 *Análisis de crecimiento.*

Como se mencionó en el capítulo 3, el índice morfológico de una planta de agave será empleado para analizar el crecimiento de éstas. Por tanto, como se busca analizar el crecimiento de las plantas en las regiones: Región 1 y Región 2, y más aún contrastar el crecimiento entre los dos tipos de planta, micropropagado y convencional (presentes a su vez en cada región), se calcularán los índices morfológicos de cada planta para cada una de las fechas en que se tienen mediciones de las variables (T0, T1, T2, y T3).

El cálculo de los índices se llevará a cabo empleando los datos de la Tabla 2 del Apéndice A.1. A partir de estos datos primero se estimará una matriz de correlación común (para contar así con índices comparables entre si), luego, con base en lo descrito en la sección 3.2 se calcularán los coeficientes correspondientes a la primer componente principal, los cuales serán empleados para el cálculo de los índices morfológicos de las plantas en el escenario en cuestión.

Cambio Proporcional ALTURA

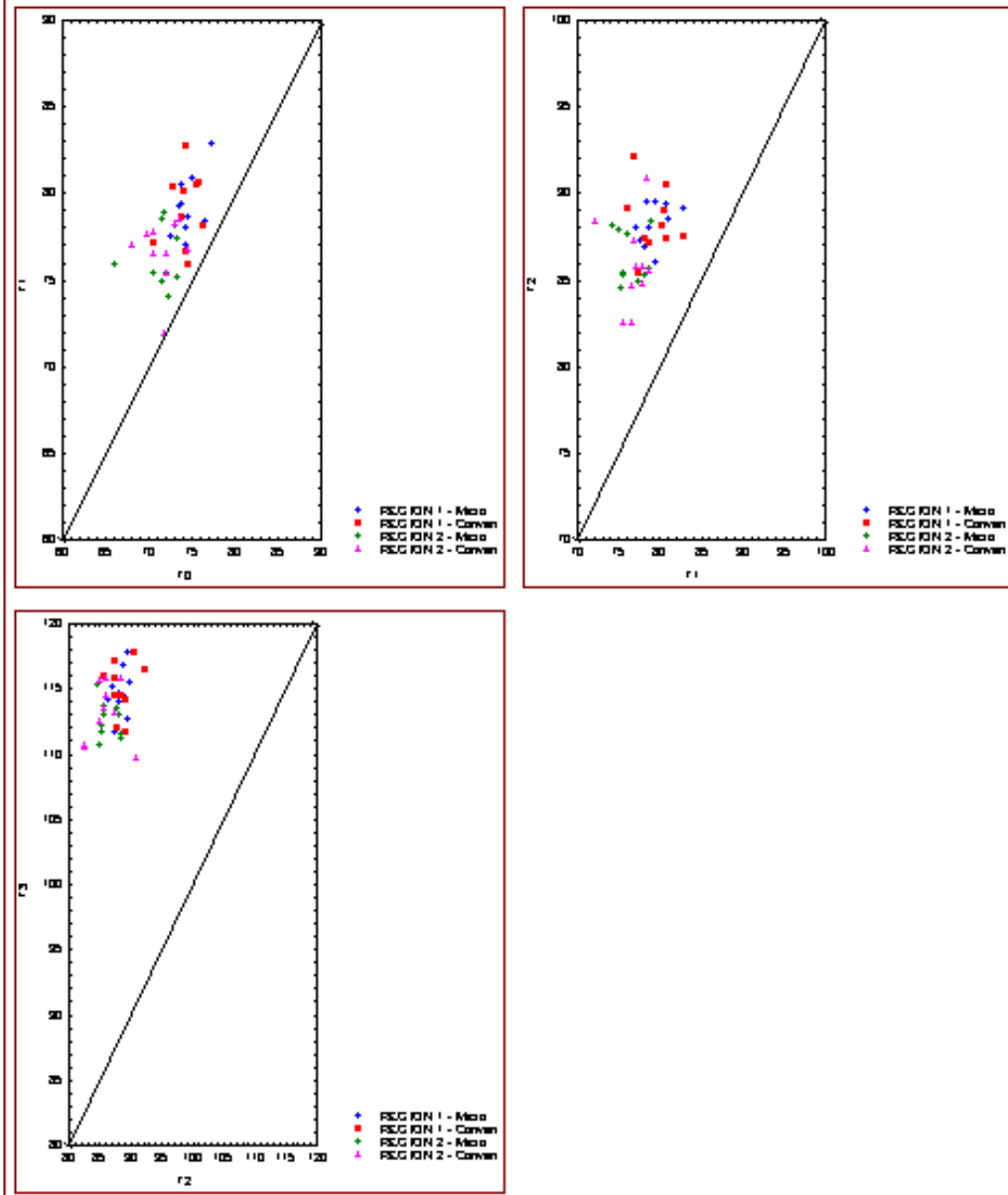


Figura 5.3. Gráficas de cambio proporcional en la variable Altura en los distintos periodos de tiempo (T0-T1, T1-T2 y T2-T3), considerando ambos tipos de planta (micropropagado y convencional) y regiones (Región 1 y Región 2).

Cambio Proporcional

Hojas desprendidas

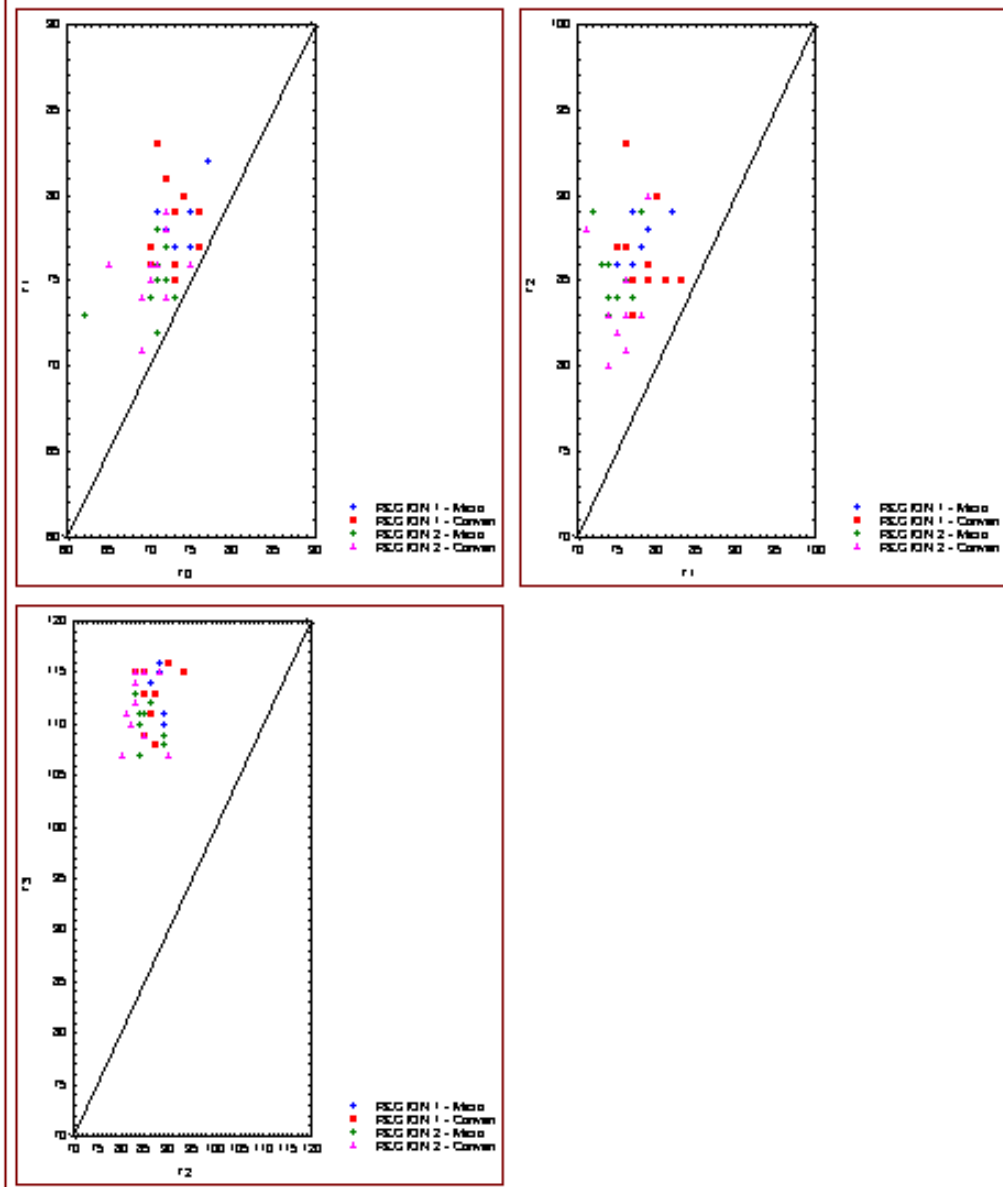


Figura 5.4. Gráficas de cambio proporcional en la variable Hojas desprendidas en los distintos periodos de tiempo (T0-T1, T1-T2 y T2-T3), considerando ambos tipos de planta (micropropagado y convencional) y regiones (Región 1 y Región 2).

Cambio Proporcional

Longitud de hoja

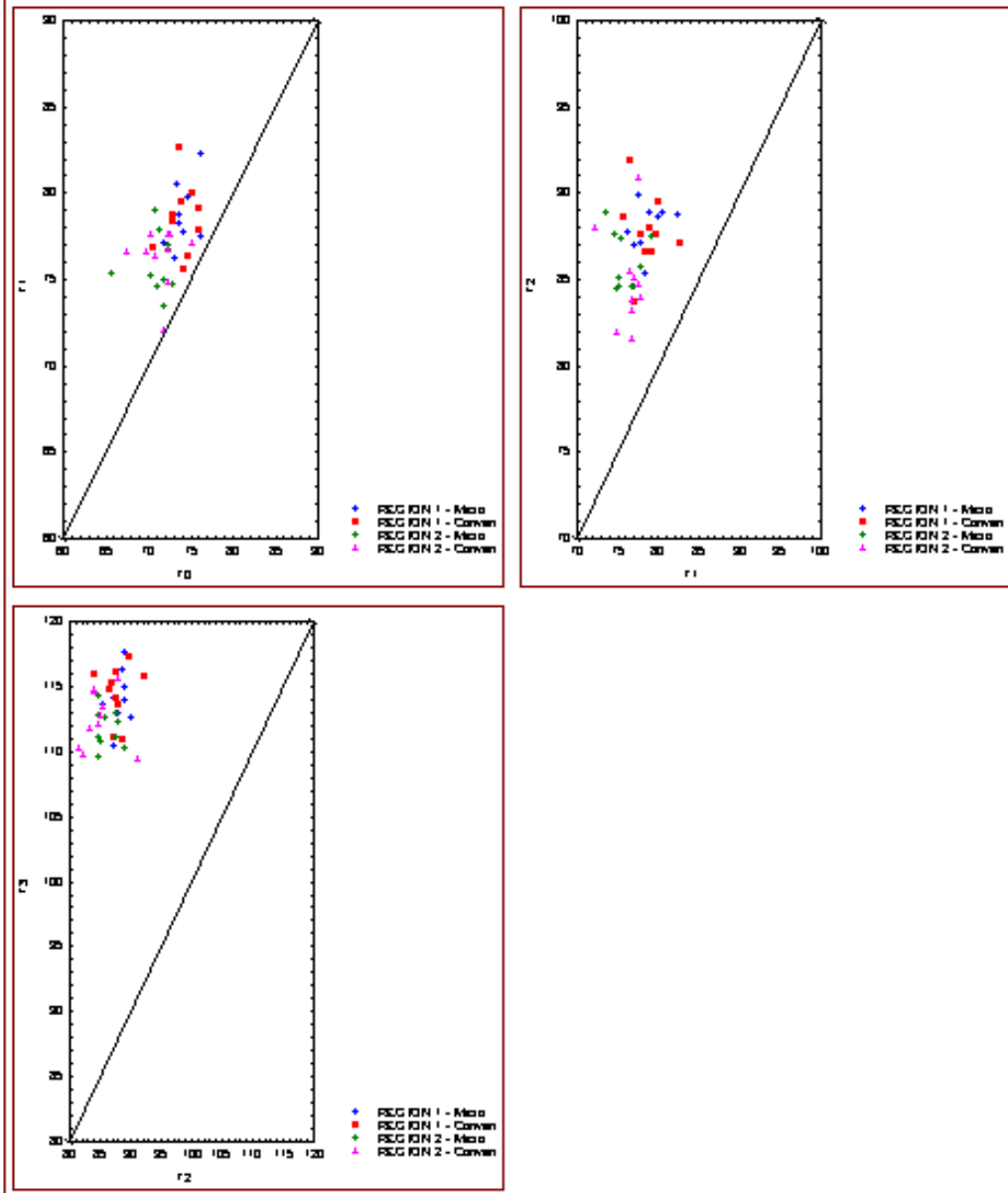


Figura 5.5. Gráficas de cambio proporcional en la variable Longitud de hoja en los distintos periodos de tiempo (T0-T1, T1-T2 y T2-T3), considerando ambos tipos de planta (micropropagado y convencional) y regiones (Región 1 y Región 2).

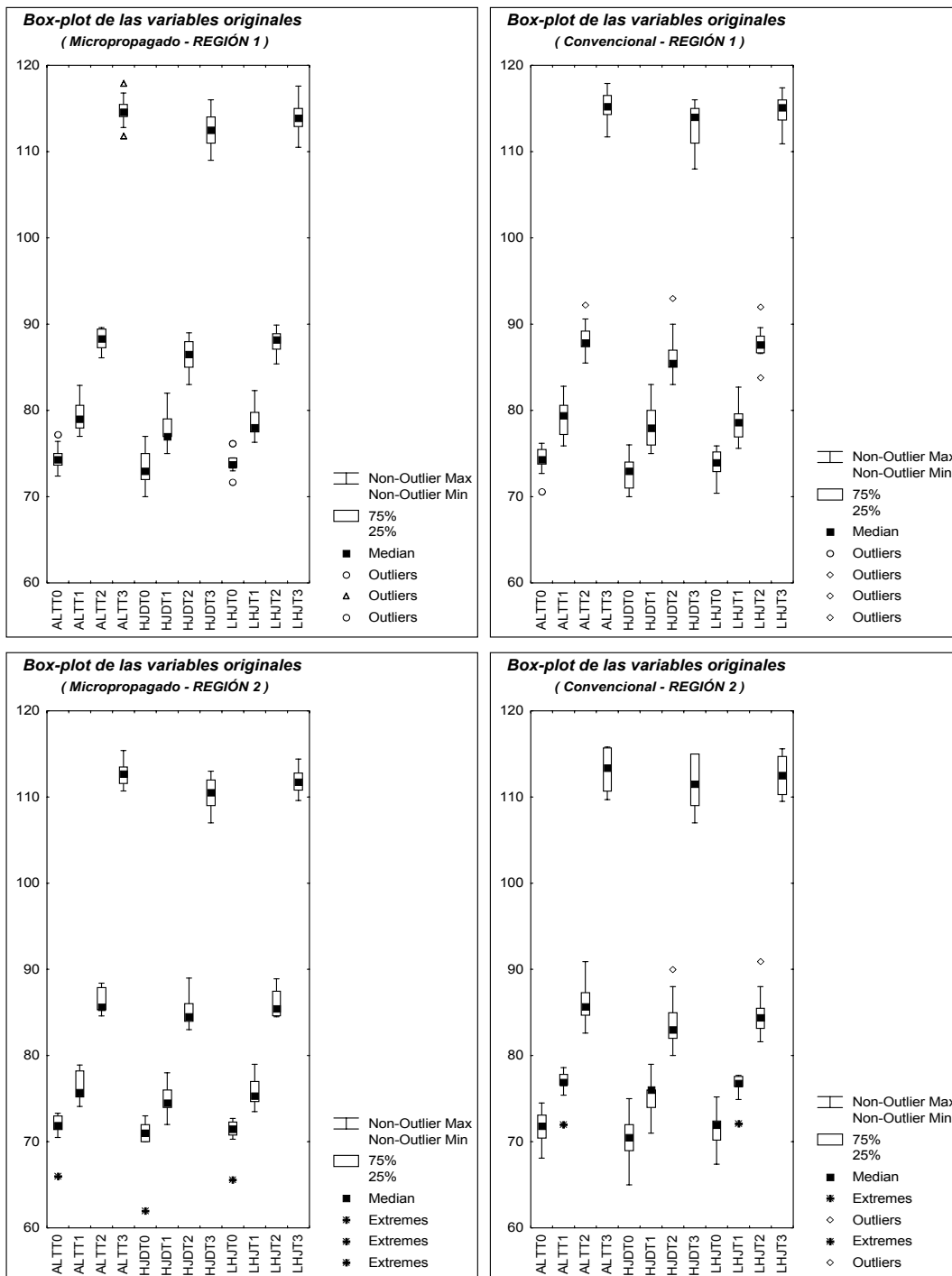


Figura 5.6. Gráficas de caja y bigotes de las variables originales como función del tiempo, clasificadas por región y tipo de planta (Escenario 1).

La matriz de correlación común será calculada como una ponderación de las matrices de correlación correspondientes a las fechas T0, T1, T2 y T3, es decir, se calcularán las matrices de correlación R_{T0} , R_{T1} , R_{T2} y R_{T3} (usando en cada caso los 40 datos de cada fecha, correspondientes a ambos tipos de planta y regiones), y luego calculando la matriz de correlación común de la forma:

$$R = \frac{(40 \times R_{T0}) + (40 \times R_{T1}) + (40 \times R_{T2}) + (40 \times R_{T3})}{160}. \quad (5.6)$$

Además de la obtención de índices comparables entre ambos tipos de planta, una razón adicional para el uso de esta matriz de correlación común es la inestabilidad que suele haber con el uso de matrices de covarianza (correlación) estimadas a partir de muestras pequeñas. Cabe hacer notar, que para este caso, es factible considerar una matriz común (como la descrita en 5.6) para el cálculo de los índices morfológicos, debido a que las matrices de correlación de cada tipo y región presentan la misma estructura para cada tiempo, por ejemplo, en lo que respecta al tiempo T0, se tendrían las siguientes matrices de correlación estimadas con base en los datos de la Tabla 2 del Apéndice A.1:

$$R_{MicR1T0} = \begin{bmatrix} 1 & 0.9074 & 0.9581 \\ 0.9074 & 1 & 0.8676 \\ 0.9581 & 0.8676 & 1 \end{bmatrix} \quad R_{ConvR1T0} = \begin{bmatrix} 1 & 0.8012 & 0.9711 \\ 0.8012 & 1 & 0.8873 \\ 0.9711 & 0.8873 & 1 \end{bmatrix}$$

$$R_{MicR2T0} = \begin{bmatrix} 1 & 0.9614 & 0.9929 \\ 0.9614 & 1 & 0.9753 \\ 0.9929 & 0.9753 & 1 \end{bmatrix} \quad R_{ConvR2T0} = \begin{bmatrix} 1 & 0.9108 & 0.9434 \\ 0.9108 & 1 & 0.9337 \\ 0.9434 & 0.9337 & 1 \end{bmatrix}$$

entonces, al ponderar las matrices anteriores se obtendría:

$$R_{PonderadaT0} = \begin{bmatrix} 1 & 0.8952 & 0.9664 \\ 0.8952 & 1 & 0.9160 \\ 0.9664 & 0.9160 & 1 \end{bmatrix}. \quad (5.7)$$

Por otro lado, empleando los 40 datos de T0 la matriz de correlación estimada estaría dada por:

$$R_{T0} = \begin{bmatrix} 1 & 0.9201 & 0.9776 \\ 0.9201 & 1 & 0.9318 \\ 0.9776 & 0.9318 & 1 \end{bmatrix}. \quad (5.8)$$

Tabla 5.3. Resultados obtenidos del análisis de componentes principales empleando los datos correspondientes al Escenario 1.

	Vectores Propios		
	VP1	VP2	VP3
Altura	0.5810	0.3997	-0.7090
Hojas Desp.	0.5700	-0.8216	0.0040
Longitud Hj.	0.5809	0.4064	0.7052
Valor Propio	2.89	0.09	0.02
% Varianza	96.21	3.07	0.72

Nótese así, que en el caso de las matrices (5.7) y (5.8), se tiene que éstas cuentan con una estructura similar (correlaciones grandes entre las variables *Altura* y *Longitud de hoja*, siguiéndole en magnitud las correspondientes entre *Hojas desprendidas* y *Longitud de hoja*, y finalmente la propia entre *Altura* y *Hojas desprendidas*), donde, la matriz (5.8) fue estimada utilizando una mayor cantidad de datos. Por tanto, al contar ambas matrices con estructuras similares es recomendable emplear, por lo ya mencionado, estimaciones calculadas de manera similar a (5.8) para cada fecha, para posteriormente proceder al cálculo de la matriz común con base en la expresión (5.6). Cabe mencionar que las matrices correspondientes a las fechas T1, T2 y T3 presentan estructuras similares a las matrices anteriores.

De esta forma, utilizando la expresión (5.6) la matriz ponderada resultante sería:

$$R_{Comun} = \begin{bmatrix} 1 & 0.9257 & 0.9784 \\ 0.9257 & 1 & 0.9251 \\ 0.9784 & 0.9251 & 1 \end{bmatrix}, \quad (5.9)$$

donde se puede apreciar que ésta también presenta una estructura similar a la obtenida en las matrices (5.7) y (5.8), siendo a su vez calculada a partir de las matrices estimadas en las fechas T0, T1, T2 y T3 con 40 datos cada una.

La matriz descrita en (5.9) corresponde a una matriz de correlación común, misma que será empleada para la obtención de los coeficientes de la primer componente principal. Los resultados obtenidos de aplicar la técnica de componentes principales a la matriz de correlación (5.9) son presentados en la Tabla 5.3 de arriba, donde se pueden apreciar los coeficientes obtenidos para cada una de las componentes principales, así como el porcentaje de varianza explicada por cada una de ellas.

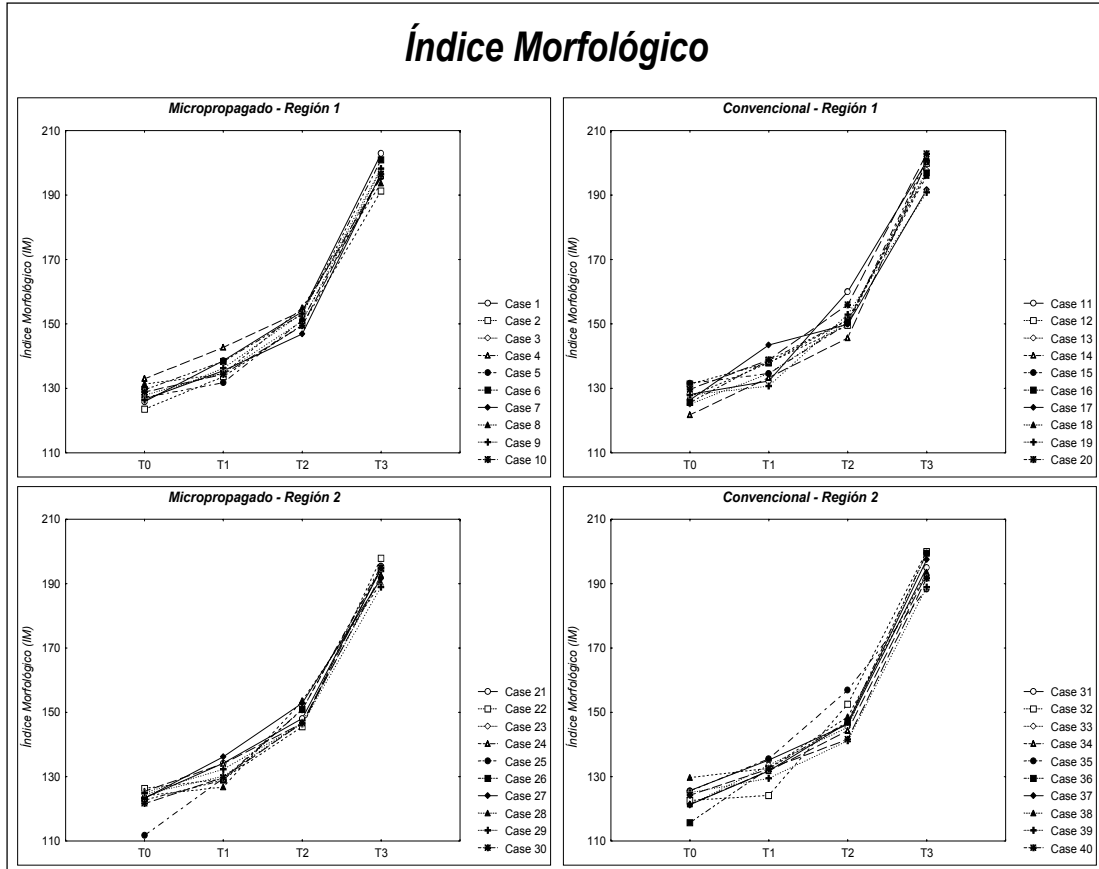


Figura 5.7. Gráficas longitudinales de los índices morfológicos para ambos tipos de planta en las Regiones 1 y 2, en las fechas T0, T1, T2 y T3, correspondientes a los datos del Escenario 1.

De la Tabla 5.3 se puede observar que los coeficientes correspondientes a la primer componente principal (VP1) son todos positivos, como se esperaría de acuerdo a lo descrito en la sección 3.2.2, siendo a fin de cuentas estos coeficientes los que servirán para calcular los índices morfológicos de las plantas. Los índices morfológicos que se obtuvieron con los datos en cuestión se pueden apreciar en la Tabla 3 del Apéndice A.1.

Cabe mencionar que los coeficientes de las componentes principales restantes, por ejemplo los correspondientes a la segunda componente principal, podrían interpretarse también en términos del contexto del problema, que en este caso al tener mediciones morfológicas en las plantas la segunda componente principal pudiese dar información acerca de la “forma” de la planta (Jolliffe, 1986, Sección 4.1), al tener a su vez en este caso que la

Tabla 5.4. Estadísticas descriptivas por región y tipo de planta para los índices morfológicos de las plantas (Escenario 1) en cada una de las fechas T0, T1, T2 y T3.

Región	Tipo	Estadísticas Descriptivas	Índices Morfológicos			
			T0	T1	T2	T3
1	Micro	Media	128.1	136.1	151.8	197.0
		Desv. Estand.	2.8	3.1	2.5	3.3
	Conven	Media	127.5	136.3	151.7	197.8
		Desv. Estand.	3.0	3.8	3.9	4.0
2	Micro	Media	122.9	131.1	148.9	193.3
		Desv. Estand.	4.1	2.9	2.9	2.7
	Conven	Media	123.3	131.9	147.1	194.7
		Desv. Estand.	3.7	3.2	4.8	4.3

primer componente principal (utilizada para la construcción del índice morfológico) nos proporciona información sobre el “tamaño” de la planta. De los resultados presentados en la Tabla 5.3 en lo que concierne a la segunda componente se puede observar que únicamente el coeficiente correspondiente a la variable *Hojas desprendidas* es negativo, y además es el mayor en magnitud de las tres variables, lo cual indicaría, por ejemplo, que plantas grandes se espera que desprendan un menor número de hojas.

Un primer análisis de carácter exploratorio para los índices obtenidos se presenta en las gráficas longitudinales de la Figura 5.7. En éstas se puede apreciar que los índices al igual que las variables originales presentan una tendencia positiva en su crecimiento para ambas regiones y tipos de planta. Por su parte, las gráficas de la Figura 5.8 muestran el cambio proporcional en los índices, donde se puede observar que todos los datos se encuentran por arriba de la recta de 45 grados, indicando así un aumento de los índices de una fecha a otra para cada periodo de tiempo.

En la Tabla 5.4 de arriba se presentan las estadísticas descriptivas (media y desviación estándar) estimadas para los índices morfológicos de las plantas correspondientes al Escenario 1. De esta tabla se puede observar que las medias no muestran diferencias significativas entre los índices de los dos tipos de planta de cada región, sin embargo, sí se puede apreciar una diferencia entre las medias de una región a otra, siendo mayores las obtenidas en la Región 1 para cada uno de los tiempos. Así mismo, puede notarse que las medias aumentan conforme el tiempo aumenta, acorde con lo observado en las gráficas de la Figura 5.7. En lo concerniente a las desviaciones estándar se puede observar el mismo comportamiento que en las variables originales,

Cambio Proporcional Índices Morfológicos

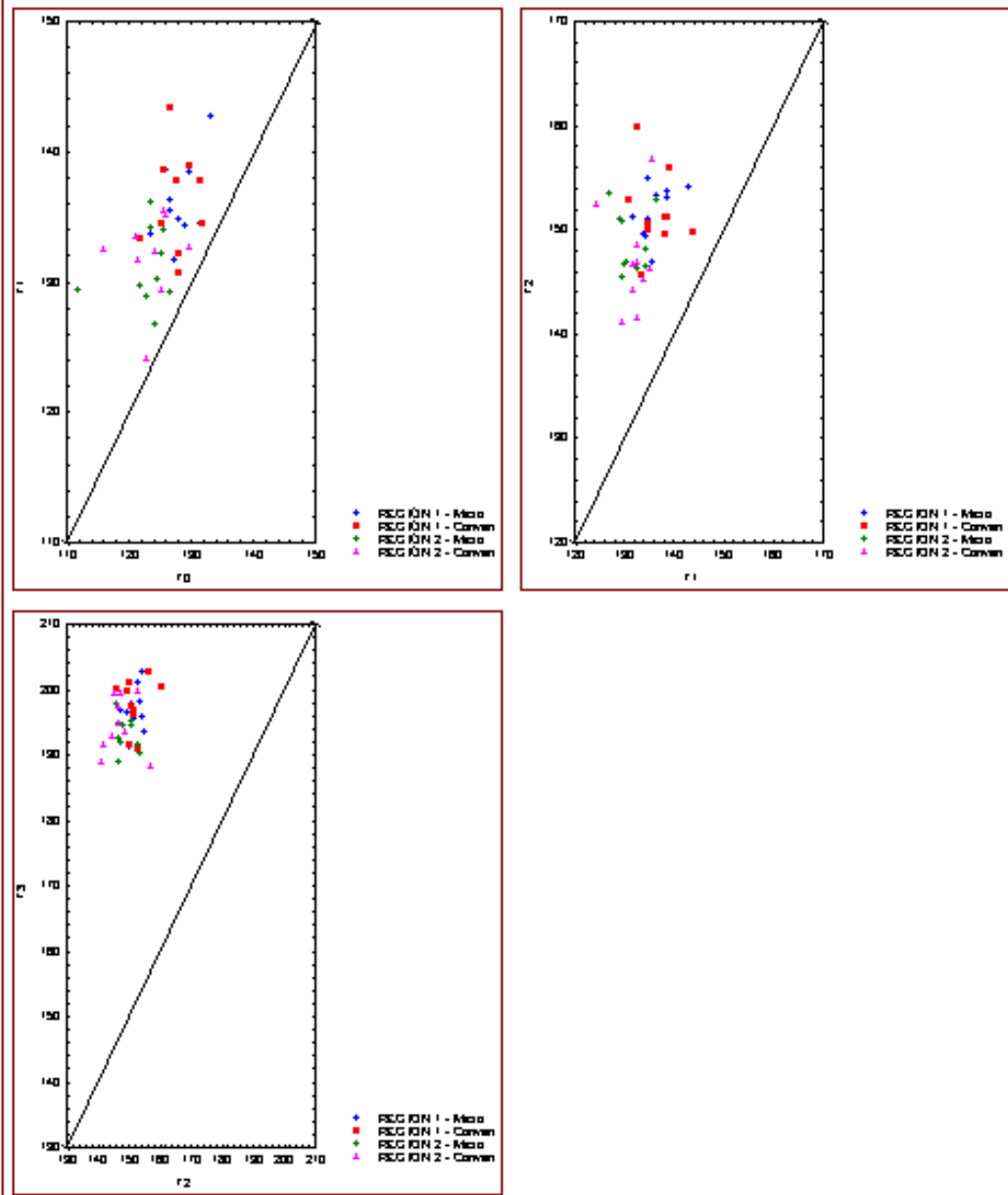


Figura 5.8. Gráficas del cambio proporcional de los índices morfológicos en los distintos periodos de tiempo (T0-T1, T1-T2 y T2-T3), para ambos tipos de planta (micropropagado y convencional) en ambas regiones (Región 1 y Región 2).

al presentar las plantas del tipo convencional (en ambas regiones) en general una mayor variabilidad con respecto a las del tipo micropropagado (excepto en T0 de la Región 2).

Finalmente, en la Figura 5.9 se presentan las gráficas de caja y bigotes (box-plot) de los índices morfológicos, donde se puede observar nuevamente la tendencia positiva en los datos a lo largo del tiempo, así como las diferencias en variabilidad entre ambos tipos de planta de cada región. Además, se puede apreciar también en estas gráficas la presencia de algunos valores extremos (como por ejemplo la planta identificada con el Caso 25 en la Figura 5.7), los cuales a su vez repercuten en las estimaciones obtenidas para la desviación estándar, como ocurre en el caso de T0 para la Región 2.

De los análisis exploratorios anteriores se observa que los índices morfológicos muestran en general una tendencia positiva en su comportamiento a lo largo del tiempo, incluso reflejando diferentes características (en lo que respecta a magnitudes de sus valores promedio) entre ambas regiones (tal y como se buscaba a partir de la elección de los valores iniciales para la simulación de los datos). Así mismo, se puede apreciar una diferencia en la variabilidad de ambos tipos de planta, siendo mayor en las plantas del tipo convencional. Todo esto muestra de manera práctica lo mencionado en su momento en la sección 3.2.2, respecto a que las componentes principales (en este caso con el uso del índice morfológico) resumen la información en el conjunto de variables originales.

Antes de pasar al análisis de las tasas de crecimiento se realizará un análisis adicional, con el cual se busca explorar si es posible diferenciar los dos tipos de plantas (micropropagado y convencional) empleando únicamente la información presente en sus variables originales. Para tal tarea se empleará una técnica de estadística multivariada conocida como *Análisis Discriminante*, descrita a su vez con más detalle en el Apéndice A.4.

Al igual que en el análisis exploratorio los datos a emplear corresponderán a los datos simulados y presentados a su vez en la Tabla 2 del Apéndice A.1, los cuales constan de mediciones simuladas de las variables *Altura*, *Número de hojas desprendidas* y *Longitud de hoja*, para plantas de agave en las Regiones 1 y 2. El análisis se realizará considerando por separado a cada una de las regiones, ya que como se ha visto antes, ambas cuentan con diferencias en lo concerniente a las magnitudes de sus variables, sobretodo en sus medias. A su vez, estos análisis se realizan para cada una de las fechas T0, T1, T2 y T3.

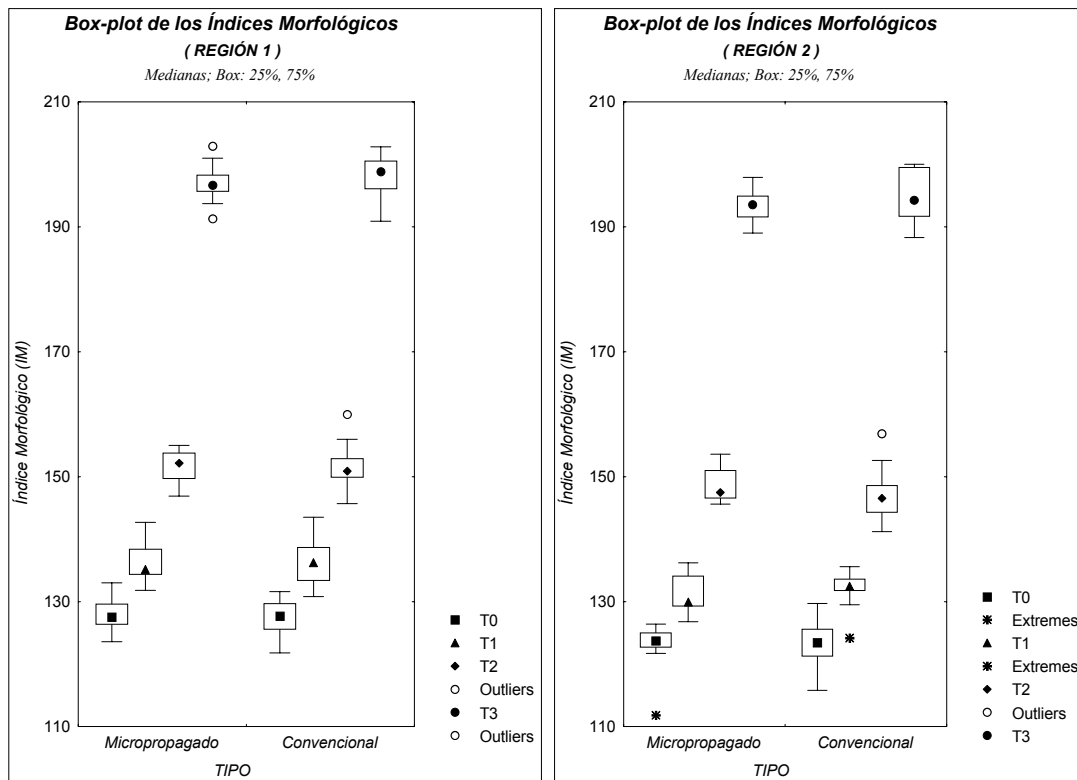


Figura 5.9. Gráficas de caja y bigotes (box-plots) de los índices morfológicos de las plantas, graficados a lo largo del tiempo para cada región y tipo de planta con los datos correspondientes al Escenario 1.

La técnica de análisis discriminante a emplear aquí es conocida en la literatura como la construcción de la *función lineal de discriminación de Fisher* (ver Apéndice A.4 para más detalle), que a grandes rasgos consiste en la construcción de una función lineal que involucra a las variables en cuestión y la evaluación de éstas en dicha función permite clasificar cada planta en el tipo más probable al que pertenece con base en sus mediciones.

Para la aplicación de esta técnica uno de los supuestos necesarios es que las matrices de covarianzas de las poblaciones a ser comparadas sean las mismas. En el caso que nos concierne se empleará la matriz de correlación común (calculada con base en la expresión 5.6) en lugar de la matriz de covarianzas, debido principalmente a la diferencia de escala entre las variables y a su vez por cuestiones de comparabilidad entre los análisis. Cabe recordar que como se mencionó en la sección anterior, las matrices de correlación de ambas poblaciones presentan la misma estructura, como se ilustró anteriormente en las expresiones (5.7) y (5.8).

Tabla 5.5. Resultados obtenidos del análisis discriminante basándose en los datos correspondientes al Escenario 1.

Región	Fecha	Coeficiente		Región	Fecha	Coeficiente	
1	T0	Altura	4.5862	2	T0	Altura	8.0267
		Hojas Desp.	2.2368			Hojas Desp.	0.3558
		Long. Hoja	-6.4665			Long. Hoja	-8.6325
		Constante	-25.9730			Constante	15.5250
	T1	Altura	3.5693		T1	Altura	5.0826
		Hojas Desp.	-4.4266			Hojas Desp.	-2.7569
		Long. Hoja	0.6729			Long. Hoja	-2.8624
		Constante	9.9091			Constante	36.0460
T2	Altura	-8.6612	T2	Altura	-17.3390		
	Hojas Desp.	-0.2409		Hojas Desp.	4.6757		
	Long. Hoja	8.9270		Long. Hoja	13.8090		
	Constante	1.4861		Constante	-83.6030		
T3	Altura	4.8213	T3	Altura	3.6869		
	Hojas Desp.	-1.7693		Hojas Desp.	-4.4861		
	Long. Hoja	-3.5604		Long. Hoja	-0.1371		
	Constante	52.3340		Constante	96.5620		

En la Tabla 5.5 se presentan los coeficientes obtenidos para la función de discriminación en cada caso, empleando para su obtención los datos mostrados en la Tabla 2 del Apéndice A.1 con base al procedimiento descrito a su vez en el Apéndice A.4. Con estos coeficientes es posible construir las funciones de discriminación para cada región y fecha, de manera análoga a como se muestra en los dos casos a continuación:

Región 1 - T0:

$$h(a, h, l) = [(4.5862) \times a] + [(2.2368) \times h] + [(-6.4665) \times l] + [-25.9730], \quad (5.10)$$

Región 2 - T2:

$$h(a, h, l) = [(5.0826) \times a] + [(-2.7569) \times h] + [(-2.8624) \times l] + [36.0460], \quad (5.11)$$

donde, en cada caso (a, h, l) denota el vector con las mediciones de las tres variables: *altura*, *hojas desprendidas* y *longitud de hoja*, correspondientes a la planta que desea clasificar.

Tabla 5.6. Resultados obtenidos al aplicar las funciones de clasificación a la muestra de plantas de agave en el Escenario 1.

REGION 1 - T0				REGION 2 - T0			
Tipo		Clasificado		Tipo		Clasificado	
		Micro	Conven			Micro	Conven
Original	Micro	4	6	Original	Micro	8	2
	Conven	2	8		Conven	4	6

REGION 1 - T1				REGION 2 - T1			
Tipo		Clasificado		Tipo		Clasificado	
		Micro	Conven			Micro	Conven
Original	Micro	7	3	Original	Micro	7	3
	Conven	6	4		Conven	2	8

REGION 1 - T2				REGION 2 - T2			
Tipo		Clasificado		Tipo		Clasificado	
		Micro	Conven			Micro	Conven
Original	Micro	7	3	Original	Micro	8	2
	Conven	3	7		Conven	2	8

REGION 1 - T3				REGION 2 - T3			
Tipo		Clasificado		Tipo		Clasificado	
		Micro	Conven			Micro	Conven
Original	Micro	5	5	Original	Micro	6	4
	Conven	4	6		Conven	4	6

La regla para clasificar una planta de agave en alguno de los dos tipos (micropropagado o convencional) es la siguiente:

la planta corresponde al tipo “micropropagado” si $h(a, h, l) > 0$, y
al tipo “convencional” si $h(a, h, l) \leq 0$.

para la función h que corresponda a cada caso.

A manera de ilustrar el uso de estas funciones de clasificación se presenta el siguiente ejemplo: supóngase que se desea clasificar una planta de la Región 1 cuyas variables en la fecha T0 fueron (74.3,73,73.7), entonces, la función de discriminación (5.10) evaluada en dichas mediciones daría como resultado $h(74.3,73,73.7) = 1.4870 > 0$, y por lo tanto de acuerdo a la regla de clasificación anterior esta planta sería clasificada como del tipo micropropagado.

Tabla 5.7. Probabilidades de mala clasificación estimadas para cada una de las funciones de clasificación construidas con base en los coeficientes presentados en la Tabla 5.5.

Región	Fecha	Función de clasificación	Probabilidad de Mala Clasificación
1	T0	3.1	0.40
	T1	3.2	0.45
	T0	3.3	0.30
	T1	3.4	0.45
2	T0	3.5	0.30
	T1	3.6	0.25
	T0	3.7	0.20
	T1	3.8	0.40

Entonces, como se busca evaluar el poder de discriminación de las variables a la luz de los datos de la muestra, se utilizarán las funciones de clasificación correspondientes a cada región y fecha (construidas con base en los coeficientes de la Tabla 5.5), evaluando en éstas las mediciones de cada planta para obtener así el tipo de planta más probable en que sería clasificada cada una, acorde con la regla de clasificación dada anteriormente.

Los resultados obtenidos de aplicar dichas funciones de clasificación a las plantas de la muestra se pueden apreciar en la Tabla 4 del Apéndice A.1, y una versión resumida de éstos es a su vez descrita en la Tabla 5.6, donde se puede apreciar el total de plantas clasificadas para cada caso.

Por su parte, la Tabla 5.7 presenta las estimaciones de las probabilidades de mala clasificación (Johnson and Wichern, 1992, Sección 11.4), con las cuales se pretende dar una medida aproximada sobre la probabilidad de clasificar mal una planta al emplear las funciones de clasificación mencionadas. Dichas probabilidades fueron estimadas de los resultados obtenidos de aplicar las funciones de clasificación a las plantas de la muestra, mostrados a su vez en la Tabla 5.6. El procedimiento para su obtención fue básicamente contar el número de plantas que fueron bien y mal clasificadas en cada tipo de planta y obtener la probabilidad de error con base en la siguiente expresión:

$$\text{Probabilidad.Mala.Clasificación} = \frac{n_{MB} + n_{CB}}{n_M + n_C}, \quad (5.12)$$

donde:

n_{MB} = número de plantas del tipo micropropagado mal clasificadas,

- n_{CB} = número de plantas del tipo convencional mal clasificadas,
- n_M = número de plantas del tipo micropropagado,
- n_C = número de plantas del tipo convencional.

De las probabilidades de mala clasificación presentadas en la Tabla 5.7 se observa que todas éstas son de magnitudes grandes. Esto indicaría que es altamente probable cometer un error al clasificar cada planta empleando solamente las mediciones de sus variables, debido en gran medida a que las mediciones de las plantas de ambos tipos son muy similares entre si y es complicado discernir entre ambos tipos sólo con base en sus mediciones.

A continuación iniciaremos con el análisis de crecimiento de las plantas, comenzando con el análisis de las tasas de crecimiento (temporales y relativas promedio), descritas en el capítulo 3.

5.3.2.1 Comparación empleando la tasa de crecimiento.

En esta sección se busca contrastar el crecimiento de los dos tipos de planta en cuestión (micropropagado y convencional), utilizando para ello los datos simulados en el Escenario 1, empleando para su contraste tasas de crecimiento calculadas a partir de los datos referidos. De lo descrito en la sección 3.4.1, son considerados para este análisis dos tipos de tasa de crecimiento: la tasa de crecimiento temporal y la tasa de crecimiento relativa promedio.

Así, el primer paso consistirá en calcular ambos tipos de tasas para cada una de las plantas en el Escenario 1, donde el cálculo de éstas será a partir de los índices morfológicos (Tabla 3 del Apéndice A.1) derivados con los datos de dicho escenario. Las tasas de crecimiento serán calculadas con base en las expresiones 3.19 (para las tasas de crecimiento temporales) y 3.21 (para las tasas de crecimiento relativas promedio), y los resultados obtenidos se presentan a su vez en las Tablas 5 y 6 del Apéndice A.1, respectivamente.

5.3.2.1.1 Tasa de crecimiento temporal.

En esta sección se analizarán las tasas de crecimiento temporales calculadas con base en los índices morfológicos derivados de los datos correspondientes al Escenario 1. Dichas tasas se presentan en la Tabla 5 del Apéndice A.1, donde se puede apreciar que se cuenta con las tasas de crecimiento tanto para

plantas del tipo micropropagado como para las del tipo convencional de las dos diferentes regiones geográficas (Región 1 y Región 2), teniéndose a su vez en cada región involucrados tres periodos de tiempo (T0-T1, T1-T2 y T2-T3).

De acuerdo a la metodología descrita en la sección 3.4.1, comenzaremos realizando un análisis gráfico de estos datos, el cual se presenta en la Figura 5.10, y donde son graficadas las tasas de crecimiento de ambos tipos de planta en cada uno de los periodos de tiempo en cuestión, mismo que se lleva a cabo en cada una de las regiones.

De estas gráficas se puede observar que las tasas de crecimiento no muestran una diferencia clara entre ambos tipos de planta, al apreciarse muy mezcladas entre si en cada región para cada periodo de tiempo. Otro punto a notar sobre estas gráficas es que la variabilidad entre ambos tipos para cada periodo de tiempo es en general muy similar.

Entonces, como nos interesa saber si los datos reflejan diferencias entre las tasas de crecimiento temporales de ambos tipos de planta, se plantea la siguiente hipótesis con base a la situación actual:

H_T : las medias de las tasas de crecimiento temporales promedio en el periodo de tiempo T-t son iguales para ambas poblaciones de plantas (agave micropropagado y agave convencional).

En la hipótesis anterior, los periodos de tiempo T-t corresponden a los periodos T0-T1, T1-T2 y T2-T3, por lo cual será necesario realizar 3 pruebas estadísticas de este estilo para cada una de las regiones involucradas.

Una prueba formal se realiza de acuerdo a lo descrito en la sección 3.4.1.1, donde se propone emplear una prueba no-paramétrica de suma de rangos empleando el estadístico *Wilcoxon-Mann-Whitney* para tal labor. Como se mencionó a su vez en dicha sección, uno de los supuestos necesarios para la aplicación de esta prueba consiste en que las tasas de crecimiento temporales sean continuas, lo cual es viable suponerse aquí por la manera en que son definidas las tasas de crecimiento en la expresión (3.20). Así mismo, para la realización de dicha prueba se recomendaba verificar la homogeneidad de varianzas entre las dos muestras a considerar para la prueba, ya que una diferencia significativa, en el sentido estadístico, podría repercutir en el resultado de la prueba de rangos. Para verificar la homogeneidad de varianzas se utilizaron las pruebas de Levene y de Brown-Forsythe.

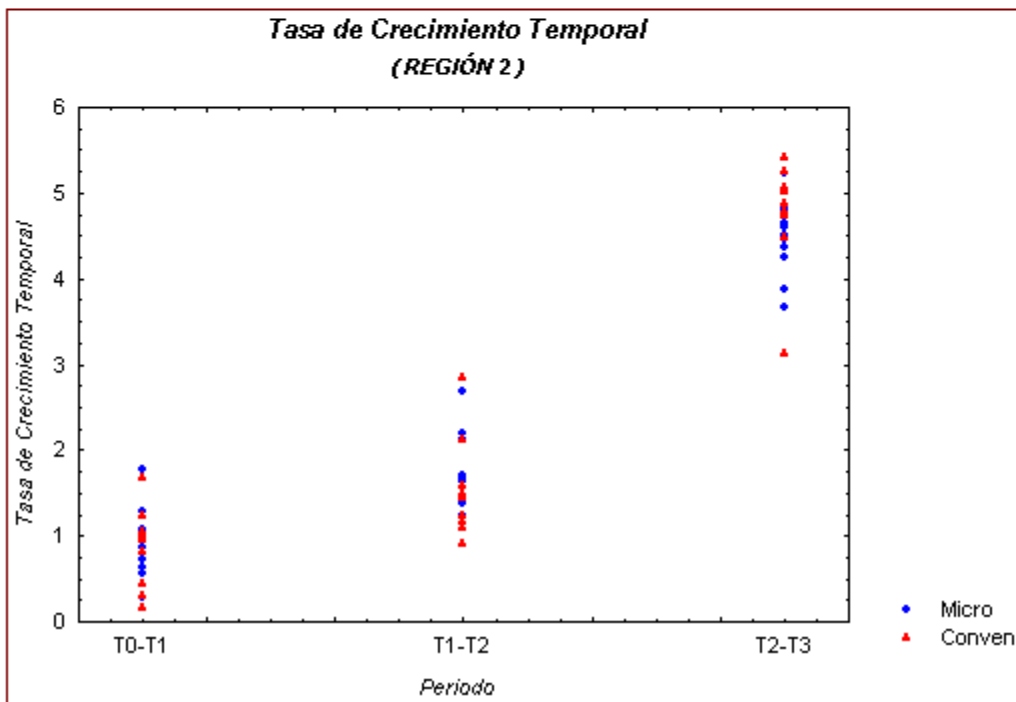
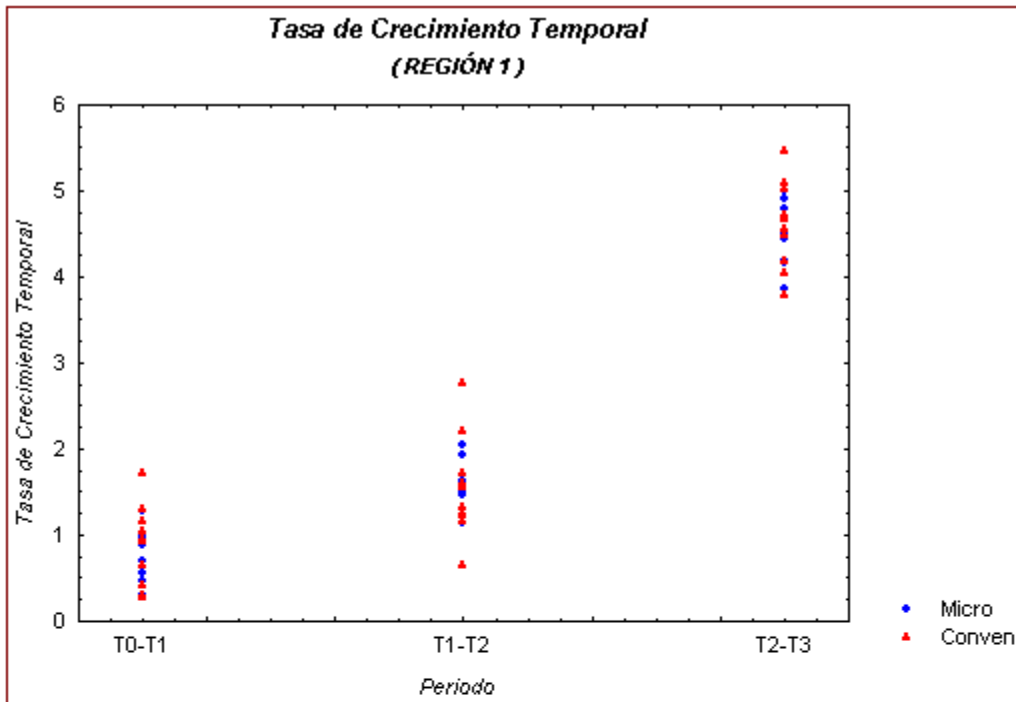


Figura 5.10. Tasas de crecimiento temporales de ambos tipos de planta, graficadas en cada periodo de tiempo y región, correspondientes a los datos simulados en el Escenario 1.

Tabla 5.8. Resultados obtenidos para las pruebas de homogeneidad de varianzas y sumas de rangos de las tasas de crecimiento temporales del Escenario 1.

Región	Periodo	Pruebas Homogeneidad de Varianzas				Pruebas		
		Desviaciones Estándar		p-valor		Wilcoxon-Mann-Whitney		
		Micro	Conven	Levene	Brn-Fors	R	R'	p-valor
1	T0 - T1	0.2934	0.4634	0.1846	0.2214	101.0	109.0	> 0.20
	T1 - T2	0.2932	0.5925	0.1483	0.1603	108.5	101.5	> 0.20
	T2 - T3	0.3621	0.5083	0.3936	0.4041	100.5	109.5	> 0.20
2	T0 - T1	0.4580	0.4574	0.9247	0.9943	100.0	110.0	> 0.20
	T1 - T2	0.4386	0.5740	0.6566	0.6316	128.0	82.0	0.10
	T2 - T3	0.4484	0.6275	0.7306	0.8300	78.0	132.0	0.05

Los resultados de las pruebas son mostrados en la Tabla 5.8 de arriba, donde en principio se puede apreciar de las pruebas de homogeneidad en cada periodo de tiempo que no hay evidencia suficiente para rechazar la hipótesis de igualdad de varianzas entre las tasas de crecimiento temporales de ambos tipo de planta, esto a raíz de los *p-valores* grandes obtenidos en cada prueba. Por otro lado, de las pruebas de rangos se puede apreciar (también de los *p-valores* obtenidos) que no hay evidencia suficiente para rechazar la hipótesis de igualdad en las medias de las tasas de crecimiento temporales promedio de ambos tipos de planta. Por tanto, a la luz de los datos analizados no existe evidencia suficiente para rechazar la hipótesis H_T planteada con anterioridad.

Ahora, analizaremos las tasas de crecimiento relativas promedio correspondientes a los mismos datos con que fueron calculadas estas tasas de crecimiento temporales para contrastar así los resultados obtenidos con ambos tipos de tasas.

5.3.2.1.2 Tasa de crecimiento relativa promedio.

Como ya se ha mencionado, los datos a analizar en esta sección corresponden a las tasas de crecimiento relativas promedio, calculadas con base en los índices morfológicos derivados a partir de los datos correspondientes al Escenario 1. Las tasas obtenidas con base en la expresión (3.22) del capítulo 3 son presentadas en la Tabla 6 del Apéndice A.1.

Al igual que en la sección anterior, comenzaremos con un análisis gráfico de estas tasas de crecimiento, graficando las tasas obtenidas en ambos tipos de planta para cada uno de los periodos de tiempo en cuestión. Las gráficas mencionadas son presentadas en la Figura 5.11. De éstas se puede apreciar que no existe una diferencia palpable entre las tasas de ambos tipos de planta, lo cual ocurre en cada periodo de tiempo de ambas regiones. Un aspecto a notar es que la variabilidad en ambos tipos de planta es similar en cada uno de los periodos de tiempo.

Entonces, como interesa saber si los datos reflejan diferencias entre las tasas de crecimiento (ahora relativas promedio) de ambos tipos de planta, se plantea la siguiente hipótesis para ser probada de manera formal:

H_R : las medias de las tasas de crecimiento relativas promedio en el periodo de tiempo T-t son iguales para ambas poblaciones de plantas (agave micropropagado y agave convencional).

En este caso también se llevarán a cabo 3 pruebas en cada región, una para cada uno de los periodos de tiempo T0-T1, T1-T2 y T2-T3. Para ello nuevamente se utilizará la prueba de suma de rangos de *Wilcoxon-Mann-Whitney*. En este caso el supuesto de continuidad, necesario para dicha prueba, es factible suponerlo como válido por la manera en que se definen estas tasas de crecimiento en la expresión (3.22). El supuesto de homogeneidad de varianzas se verificará también en este caso utilizando las pruebas de Levene y Brown-Forsythe, al igual que en la sección anterior.

Los resultados de las pruebas realizadas se muestran en la Tabla 5.9, donde es posible apreciar de las pruebas de homogeneidad en cada periodo que no hay evidencia para rechazar la hipótesis de igualdad de varianzas entre las tasas de crecimiento relativas promedio de ambos tipos de planta (al obtener en general p-valores mayores a 0.1). Así mismo, de las pruebas de rangos se observa que tampoco hay evidencia suficiente para rechazar la hipótesis de igualdad en las medias de las tasas de crecimiento relativas promedio de ambos tipos de planta (idem pruebas de homogeneidad).

Por tanto, para estas tasas de crecimiento, a la luz de los datos analizados, tampoco existe evidencia suficiente para rechazar la hipótesis H_R planteada anteriormente.

A continuación se analizará el crecimiento de las plantas utilizando otra de las herramientas propuestas en la sección 3.4.2, nos estamos refiriendo a las

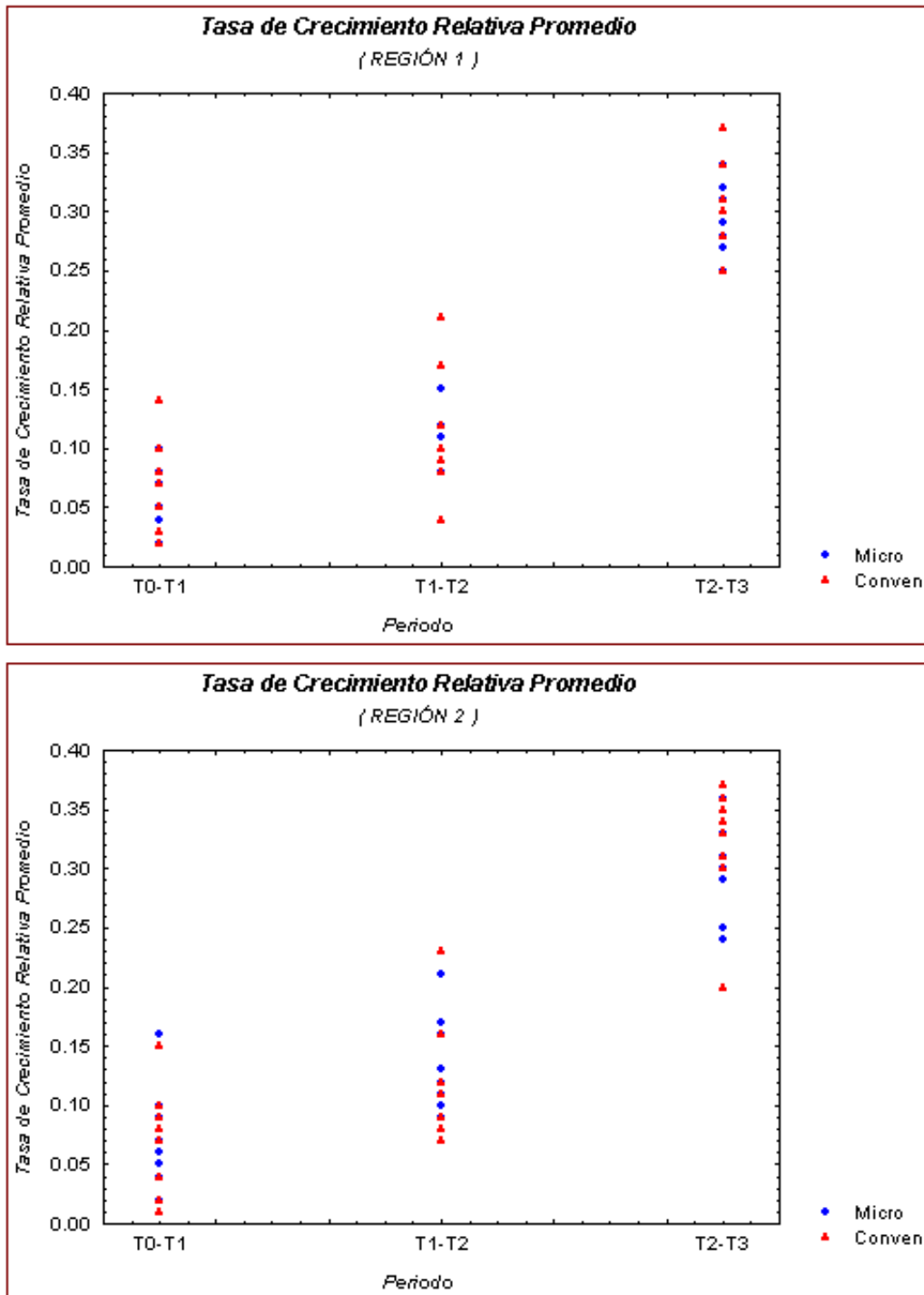


Figura 5.11. Tasas de crecimiento relativas promedio de ambos tipos de planta, graficadas en cada periodo de tiempo y región, correspondientes a los datos simulados en el Escenario 1.

Tabla 5.9. Resultados obtenidos para las pruebas de homogeneidad de varianza y sumas de rangos de las tasas de crecimiento relativas promedio del Escenario 1.

Región	Periodo	Pruebas Homogeneidad de Varianzas				Pruebas Wilcoxon-Mann-Whitney		
		Desviaciones Estándar		p-valor		R	R'	p-valor
		Micro	Conven	Levene	Brn-Fors			
1	T0 - T1	0.0235	0.0373	0.1679	0.1870	101.0	109.0	> 0.20
	T1 - T2	0.0233	0.0462	0.1688	0.1829	110.0	100.0	> 0.20
	T2 - T3	0.0265	0.0391	0.3620	0.4146	101.0	109.0	> 0.20
2	T0 - T1	0.0408	0.0393	0.9750	0.9995	101.0	109.0	> 0.20
	T1 - T2	0.0362	0.0467	0.7405	0.7348	128.0	82.0	0.10
	T2 - T3	0.0349	0.0486	0.5965	0.8438	80.0	130.0	0.05 < p < 0.10

curvas de crecimiento, empleando para su construcción los datos que a su vez fueron utilizados para el cálculo de las tasas de crecimiento analizadas en esta sección.

5.3.2.2 Comparación empleando curvas de crecimiento.

Al igual que en la sección anterior, el objetivo aquí será contrastar el crecimiento de las plantas de agave de ambos tipos (micropropagado y convencional), para lo cual, se construirán curvas de crecimiento con base en índices morfológicos derivados de los datos correspondientes al Escenario 1.

Por lo tanto, la idea a seguir consiste en construir una curva de crecimiento para cada tipo de planta (utilizando la información en la muestra), para posteriormente contrastar ambas curvas y observar su comportamiento. De haber diferencias en el crecimiento de las plantas se esperaría que éstas se viesen reflejadas en la obtención de curvas distintas para cada tipo de planta, las cuales podrían diferir en sus tasas de crecimiento, en sus valores iniciales en el origen, o en ambos parámetros.

El modelo de crecimiento propuesto para este ajuste corresponde al modelo exponencial descrito en la expresión (3.15) del capítulo 3, por lo que el

Tabla 5.10. Resultados obtenidos del ajuste no-lineal utilizando la técnica de mínimos cuadrados y empleando los índices morfológicos correspondientes a los datos del Escenario 1.

Región	Tipo	Parámetro	Estimación	Error Estándar	Desv. Estándar del Error Residual	Prueba Normalidad	
						S-W	p-valor
1	Micro	c	122.467	0.65819	3.0137	0.9812	0.7340
		k	0.08787	0.00034			
1	Conven	c	122.174	0.81526	3.7359	0.9841	0.8369
		k	0.08817	0.00042			
2	Micro	c	117.785	0.68618	3.1451	0.9311	0.0175
		k	0.08824	0.00035			
2	Conven	c	117.621	0.90707	4.1607	0.9764	0.5571
		k	0.08852	0.00047			

modelo estadístico a considerar para el ajuste, de acuerdo a lo comentado en la sección 3.4.2, es el siguiente:

$$y_{ij} = c + \exp(k \times t_j) + e_{ij}, \quad (5.13)$$

donde c representa el valor inicial en el origen, k la tasa de crecimiento, $i = 1, 2, \dots, n$ (identificador de las plantas), y $j = 0, 1, 2, 3$ (identificador del tiempo, con, $t_0=T0$, $t_1=T1$, $t_2=T2$, y $t_3=T3$).

Entonces, el objetivo que se busca de ajustar el modelo (5.13) con base en los datos es estimar los parámetros c y k involucrados en éste, para construir así las curvas de crecimiento en cada caso. Para el ajuste del modelo (5.13) es necesario considerar una metodología de ajuste del tipo no-lineal, para lo cual se empleará el método de mínimos cuadrados. Los datos a considerar para este ajuste corresponden a los índices morfológicos presentados en la Tabla 3 del Apéndice A.1. Cabe mencionar que debido a que las plantas de ambas regiones (Región 1 y Región 2) cuenta con características diferentes entre si (como se ha visto de análisis anteriores), se realizarán ajustes de manera individual para cada una de las regiones, teniéndose por tanto cuatro a ajustes a realizarse, dos en la Región 1 (uno para cada tipo de planta), y dos en la Región 2 (también uno para cada tipo de planta). Para llevar a cabo los ajustes mencionados se empleó el modulo *nlm* del software *S-PLUS 6.1*, donde a su vez en dicho software fueron realizados los programas necesarios para tal tarea, un ejemplo se incluye en el Apéndice A.3.

Los resultados de los ajustes realizados se presentan en la Tabla 5.10 de arriba, donde se pueden apreciar los parámetros estimados para la curva de

crecimiento en cada caso. De las estimaciones obtenidas se puede observar que las correspondientes a las tasas de crecimiento, parámetro k , son similares entre ambos tipos de planta en cada región (incluso siendo similares entre regiones). Por otro lado, en lo concerniente a las estimaciones de los valores iniciales en el origen, parámetro c , se observa que entre ambos tipos de planta (en cada región) tienden a ser muy similares en magnitud, además, también puede apreciarse que para cada tipo de planta las estimaciones entre ambas regiones son distintas, teniéndose mayores magnitudes en las estimaciones obtenidas para la Región 1 en contraste con las de la Región 2. Esto indicaría que las plantas en la Región 1 contaban con mayores tamaños al momento de comenzar a tomar las mediciones en las plantas. Sobre los errores estándar de las estimaciones se puede apreciar que en el caso del parámetro k son todas pequeñas y de magnitudes similares, mientras que para el parámetro c son un poco mayores pero también similares entre sí (aunque menores en las plantas del tipo micropropagado). Sobre la desviación estándar de los errores residuales se observa que todos los casos son de magnitudes similares y relativamente pequeñas, no obstante, el correspondiente al tipo convencional en la Región 2 es ligeramente mayor que los demás, lo cual se debe en gran medida a una mayor dispersión en los datos como es posible apreciar en la Figura 5.13.

En las Figuras 5.12 (Región 1) y 5.13 (Región 2) se presentan las gráficas de los datos con sus respectivas curvas de crecimiento promedio, construidas con base en los parámetros presentados en la Tabla 5.10. De éstas se puede observar que el comportamiento de ambos tipos de planta es muy similar en ambas regiones, lo cual se aprecia mucho mejor en las gráficas de la Figura 5.14, donde son graficadas en cada figura las curvas de crecimiento promedio de los dos tipos de planta, donde no se muestra una diferencia significativa entre ambas curvas al ser prácticamente idénticas.

Estas gráficas muestran así mismo que no hay evidencia para suponer un crecimiento distinto entre ambos tipos de planta, siendo esto acorde con los resultados obtenidos de los análisis de las tasas de crecimiento en secciones anteriores.

Finalmente, como suele hacerse comúnmente en todo análisis de regresión, se calcularon y analizaron los residuales obtenidos a raíz del ajuste, para detectar así posibles estructuras en los errores que no fuesen consideradas en el modelo (5.13) ajustado con anterioridad. Un primer análisis gráfico de éstos es presentado en las Figuras 2 y 3 del Apéndice A.2, donde son graficados los

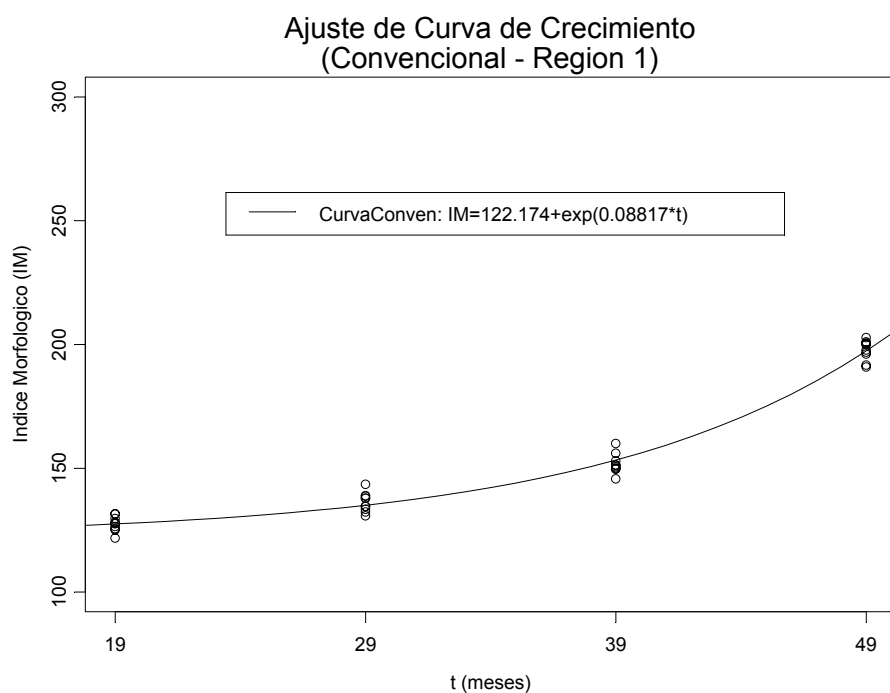
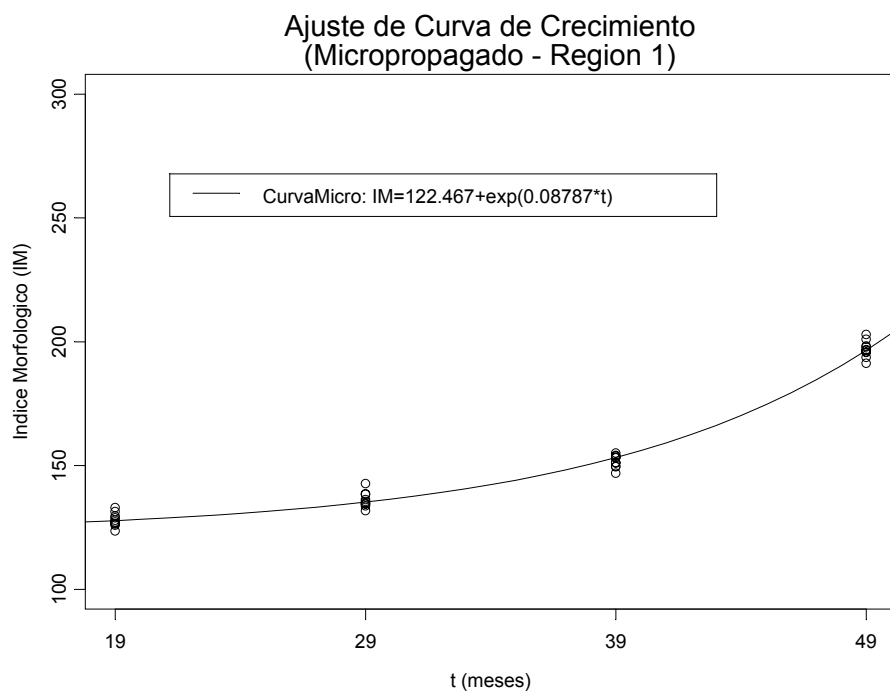


Figura 5.12. Ajuste de la curva de crecimiento para cada tipo de planta, obtenido con base en los índices morfológicos derivados de los datos del Escenario 1 (Región 1).

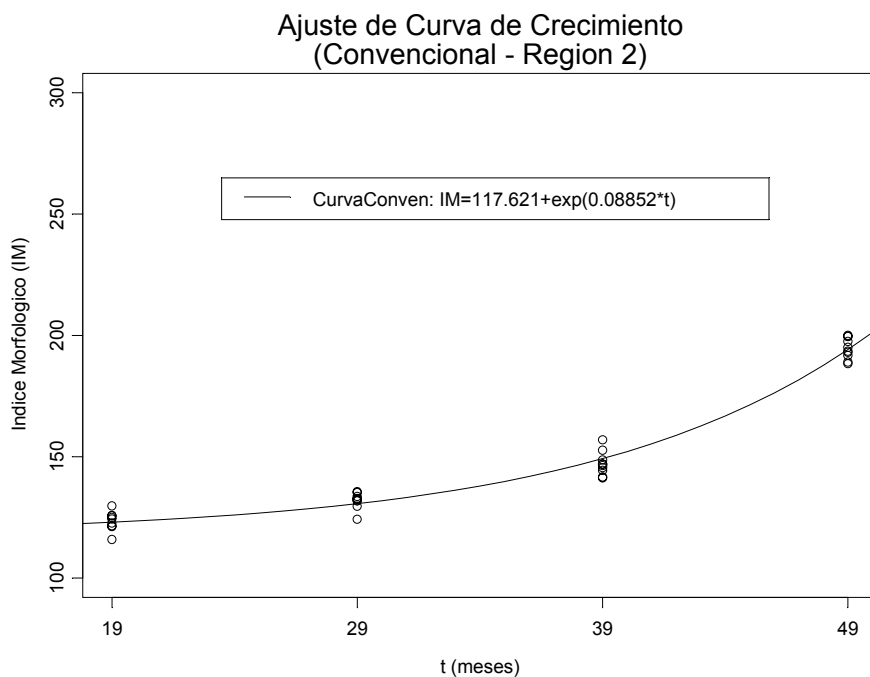
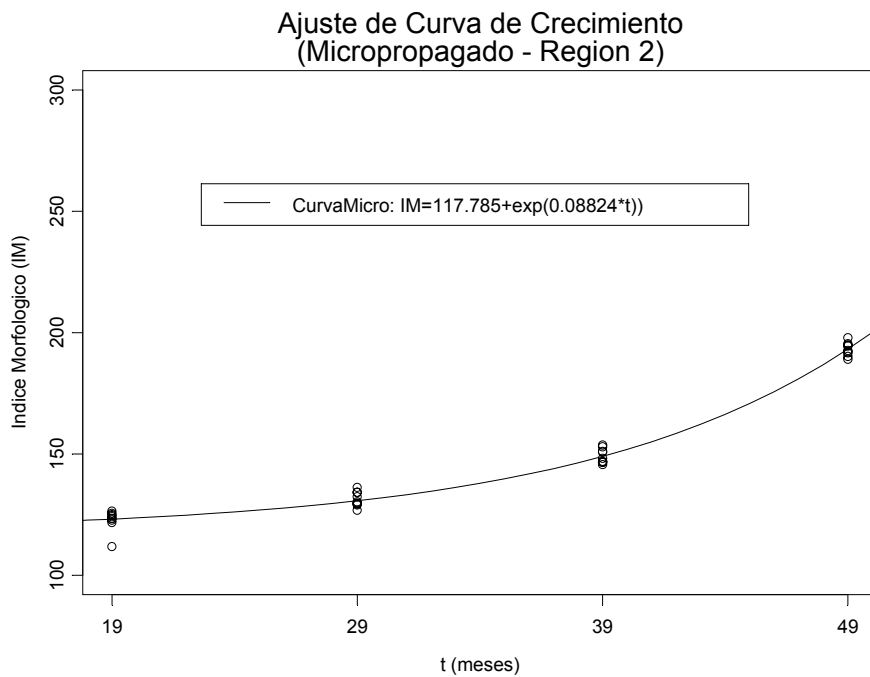
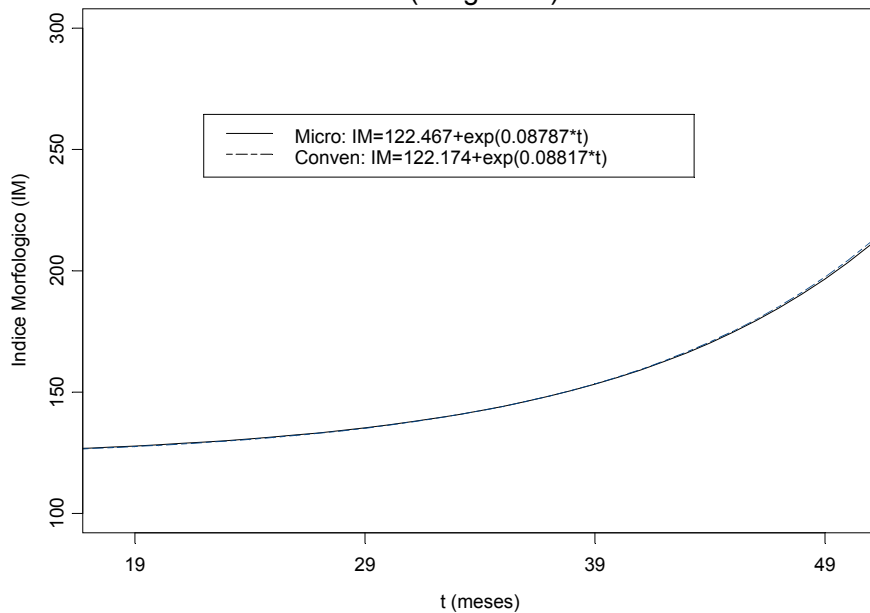


Figura 5.13. Ajuste de la curva de crecimiento para cada tipo de planta, obtenido con base en los índices morfológicos derivados de los datos del Escenario 1 (Región 2).

Curvas de Crecimiento Promedio: Micropropagado y Convencional (Region 1)



Curvas de Crecimiento Promedio: Micropropagado y Convencional (Region 2)

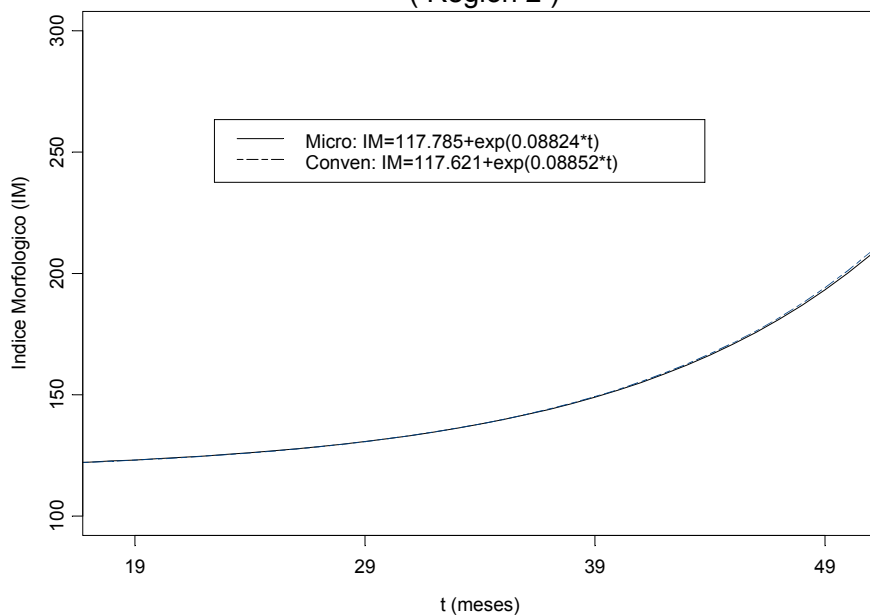


Figura 5.14. Curvas de crecimiento promedio ajustadas con base en los índices morfológicos de las plantas del Escenario 1, presentadas por regiones.

residuales como función de los valores ajustados. En cada una de estas gráficas se puede observar que no existe alguna estructura que lleve a suponer que los errores en el modelo no sean constantes. Así mismo, se verificó si los residuales se distribuyen de manera normal, para lo cual se realizó una prueba formal empleando el estadístico Shapiro-Wilk, donde los resultados para cada caso son mostrados también en la Tabla 5.10, y de los cuales se puede concluir que no existe evidencia suficiente para rechazar que estos residuales provengan de una distribución normal. A su vez, se graficaron las gráficas cuantil-cuantil (qq-plots) normales para cada conjunto de residuales, los cuales se incluyen en las Figuras 4 y 5 del Apéndice A.2. De éstas se observan, sobretodo en el tipo micropropagado de la Región 1 y 2, algunos aparentes outliers, los cuales se pueden identificar a su vez de las graficas de los ajustes en cada caso, sin embargo, la mayoría proporciona un buen ajuste de los residuales. Por tanto, no hay evidencia suficiente que indique que los residuales obtenidos no se distribuyen de manera normal.

5.3.3 Conclusiones.

A lo largo de esta sección se realizaron algunos análisis relacionados con el crecimiento de plantas de agave a lo largo del tiempo. De éstos se pudo observar que los datos en general no muestran evidencia de que las dos poblaciones de plantas (micropropagado y convencional) estén creciendo de manera distinta, al mostrar por un lado que sus tasas de crecimiento no presentan diferencias significativas entre ambos tipos de planta, al igual que con las curvas de crecimiento ajustadas a los datos, las cuales tampoco muestran diferencia en el comportamiento de los índices morfológicos a lo largo de tiempo entre ambos tipos.

Un aspecto a notar de los análisis realizados en esta sección es que la metodología empleada permite, además de vislumbrar si existen diferencias entre ambos tipos de planta en lo que respecta a su crecimiento, detectar diferencias en las características de las plantas de las regiones involucradas, donde en este caso se pudo observar que las plantas en la Región 1 mostraban magnitudes en sus variables (y así mismo en sus índices) superiores a las presentes en la Región 2 (acorde con la elección de los valores iniciales de ambas regiones).

Por tanto, tomando en cuenta que los datos analizados anteriormente fueron generados precisamente de forma que estos datos provinieran de dos poblaciones con crecimientos similares, se observa de los análisis anteriores

que las metodologías propuestas en este caso detectaron tal situación, al no mostrar diferencias significativas en el crecimiento de los dos tipos de planta en estudio con base en la información en los datos. El siguiente paso sería pues, realizar un análisis de la misma naturaleza al anterior, pero ahora con datos provenientes de un escenario en el que el crecimiento en las poblaciones de plantas si difiera, para observar así si estas herramientas permiten detectar diferencias en el crecimiento de las plantas a lo largo del tiempo. Esto último será precisamente el objetivo buscado en la siguiente sección, donde se analizarán un conjunto de datos generados provenientes de dos poblaciones de plantas con características diferentes.

5.4 Escenario 2: Poblaciones con crecimientos diferentes.

Ahora, en esta sección se analizarán los datos correspondientes al Escenario 2, presentados a su vez en la Tabla 7 del Apéndice A.1. Los datos referidos consisten en mediciones de las variables *Altura*, *Número de hojas desprendidas* y *Longitud de hoja*, tomadas de 40 plantas de agave de dos diferentes regiones geográficas denominadas: Región 1 y Región 2. Las mediciones fueron tomadas en 4 fechas consecutivas: T0, T1, T2 y T3; las cuales distan entre si en periodos de 10 meses. De las 40 plantas, 20 pertenecen a la Región 1 y 20 a la Región 2, donde a su vez en cada región 10 plantas corresponden al tipo micropropagado y 10 al tipo convencional.

Al igual que en el capítulo anterior, iniciaremos con un análisis exploratorio de dichos datos, el cual se presenta a continuación.

5.4.1 *Análisis exploratorio de los datos.*

Un primer análisis gráfico es presentando en las Figuras 5.15 (Región 1) y 5.16 (Región 2), donde se presentan las gráficas longitudinales de cada una de las variables (*Altura*, *Número de hojas desprendidas* y *Longitud de hoja*) graficadas como función del tiempo. En principio, en estas gráficas se puede observar un claro patrón de crecimiento en cada una de las variables destacando una sutil diferencia entre ambos tipos de planta, notoria sobretodo en las últimas mediciones tomadas en los tiempos T2 y T3.

Por otro lado, en las Figuras 5.17, 5.18 y 5.19 se puede apreciar el cambio proporcional en las variables para los distintos periodos de tiempo considerados (T0-T1, T1-T2 y T2-T3). De estas gráficas se aprecia que cada

Tabla 5.11. Estadísticas descriptivas por región y tipo de planta para las mediciones en el agave tomadas en las fechas T0, T1, T2 y T3 en el Escenario 2.

REGION	TIPO	Estadísticas Descriptivas	ALTURA				HOJAS DESPRENDIDAS				LONGITUD DE HOJA			
			T0	T1	T2	T3	T0	T1	T2	T3	T0	T1	T2	T3
1	0	Media	74.1	80.0	91.3	128.7	73	79	90	126	73.5	79.5	90.8	128.1
		Desv. Estándar	1.0	1.4	1.6	1.6	1	2	2	2	1.0	1.4	1.5	1.5
	1	Media	73.7	80.0	88.7	116.3	72	79	86	115	73.2	79.6	88.3	115.5
		Desv. Estándar	1.6	1.8	1.3	1.9	2	3	2	3	1.7	1.5	1.3	1.9
2	0	Media	71.5	77.6	88.5	125.5	70	77	87	123	70.8	77.1	87.8	124.8
		Desv. Estándar	0.9	2.1	1.5	1.4	1	3	2	2	1.0	2.2	1.5	1.2
	1	Media	70.5	76.9	85.1	112.7	69	76	84	111	70.0	76.6	84.7	112.0
		Desv. Estándar	2.4	1.4	1.8	2.8	4	2	2	4	2.5	1.5	1.6	3.0

Tipo 0 = Micropropagado

Tipo 1 = Convencional

conjunto de datos se encuentran por arriba de la recta de 45 grados incluida, lo cual indica que de una fecha a otra en efecto se está teniendo un aumento en las magnitudes de las variables de cada planta. Otro punto a notar de estas gráficas es la clara separación entre los conjuntos de datos de cada región y tipo, sobretodo notable en los periodos T1-T2 y T2-T3. Esto último podría representar un indicio sobre posibles diferencias en el crecimiento de las plantas.

En la Tabla 5.11 de arriba se presentan las estadísticas descriptivas para estos datos, medias y desviaciones estándar, correspondientes a cada una de las variables en cada región y fecha. De esta tabla se puede observar en principio un aumento en las magnitudes de las medias de las variables con el paso del tiempo. Otro punto a notar (en ambas regiones) es la diferencia entre las medidas de ambos tipos de planta, la cual es más marcada en las últimas dos fechas. Por su parte, la desviación estándar tiende a ser ligeramente mayor en las plantas del tipo convencional en comparación con las del tipo micropropagado, salvo en algunos casos donde no es así, como por ejemplo: Región 1, *Altura*, T2; Región 1, *Hojas desprendidas*, T2; Región 1, *Longitud de hoja*, T2; Región 2, *Altura*, T1; Región 2, *Hojas desprendidas*, T1 y T2; y Región 2, *Longitud de hoja*, T1. Finalmente, de esta tabla también se puede apreciar que las magnitudes de las variables son ligeramente mayores en la Región 1 en contraste con las de la Región 2, al comparar para cada tipo de planta las mediciones de ambas regiones en una misma fecha.

Una análisis gráfico adicional es presentado en la Figura 5.20, donde se grafican los box-plot de las variables de cada tipo de planta en ambas regiones. De aquí se puede apreciar claramente una tendencia positiva en

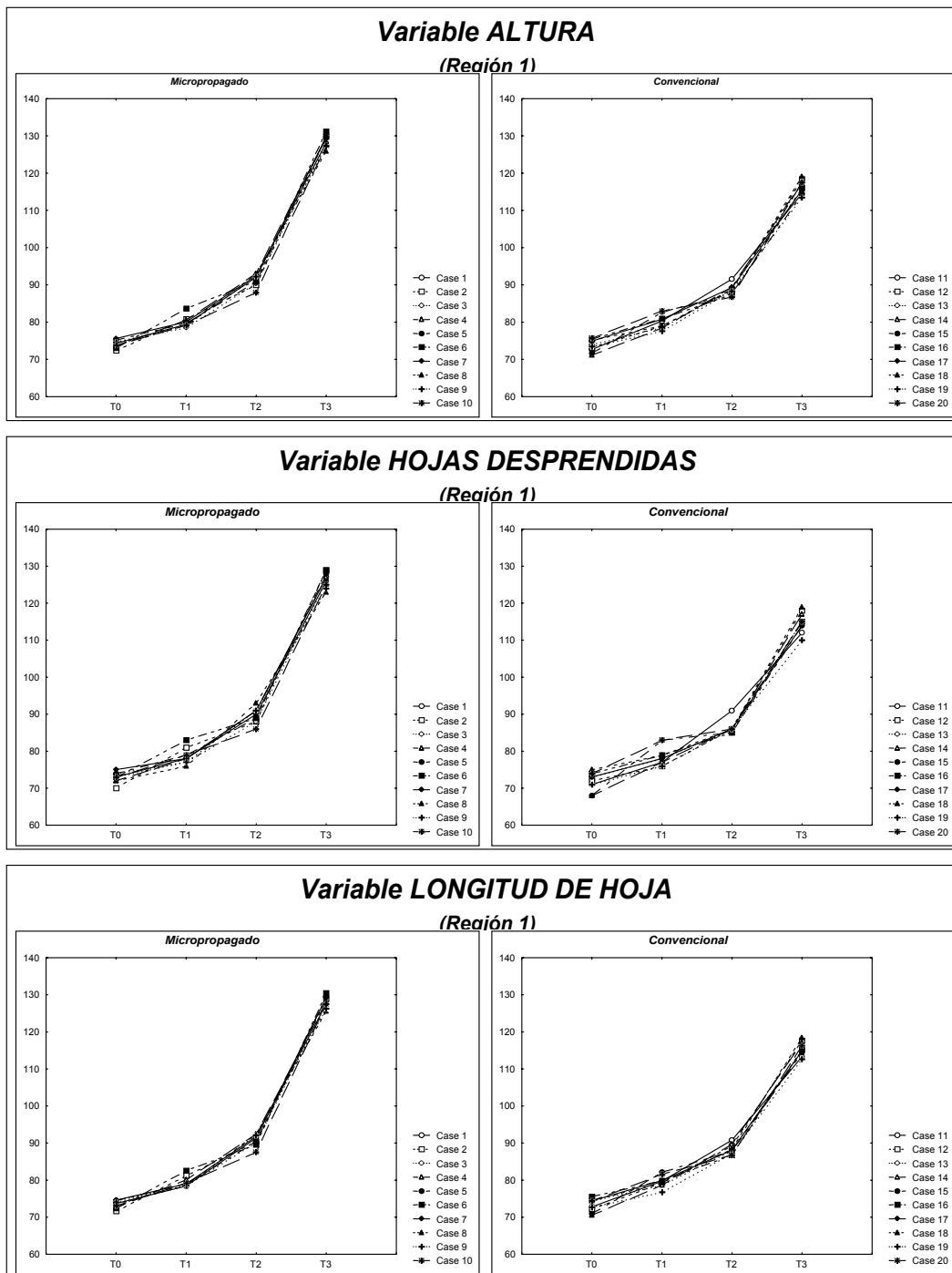


Figura 5.15. Gráficas longitudinales de las variables originales para las plantas de agave de la Región 1 en las fechas T0, T1, T2 y T3 (Escenario 2).

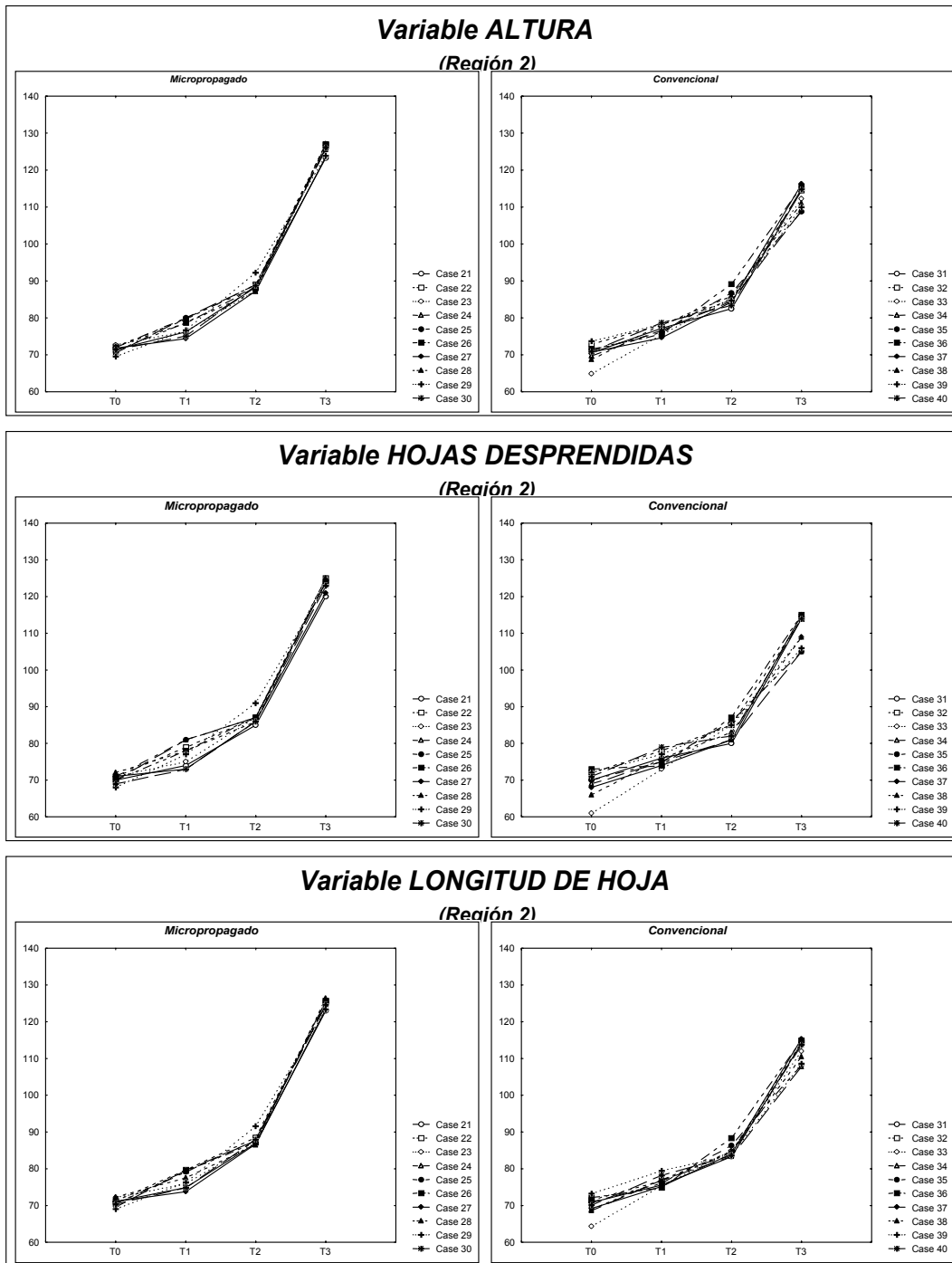


Figura 5.16. Gráficas longitudinales de las variables originales para las plantas de agave de la Región 2 en las fechas T0, T1, T2 y T3 (Escenario 2).

Cambio Proporcional ALTURA

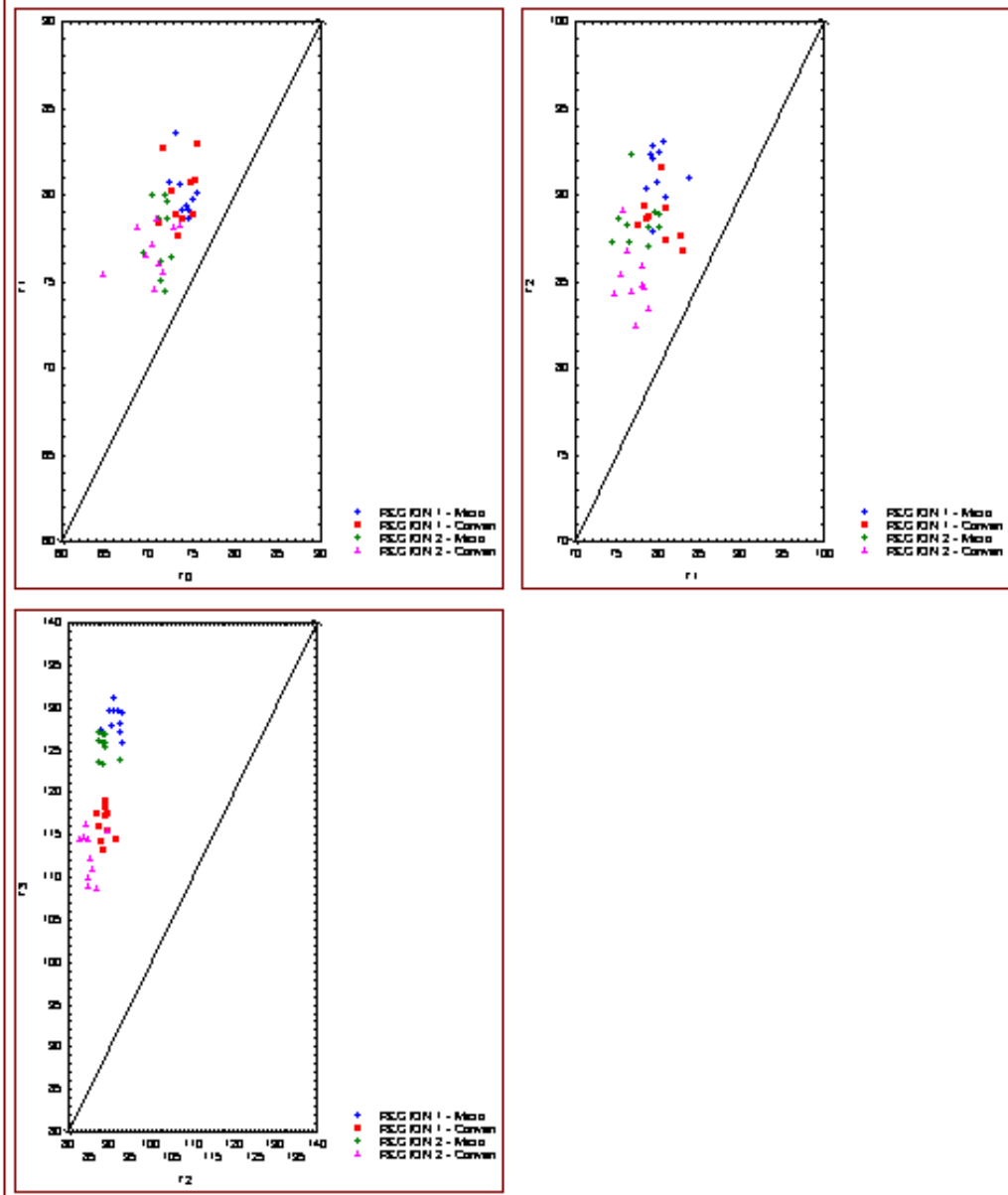


Figura 5.17. Gráficas de cambio proporcional en la variable Altura en los distintos periodos de tiempo (T0-T1, T1-T2 y T2-T3), considerando ambos tipos de planta (micropropagado y convencional) y regiones (Región 1 y Región 2).

Cambio Proporcional

Hojas desprendidas

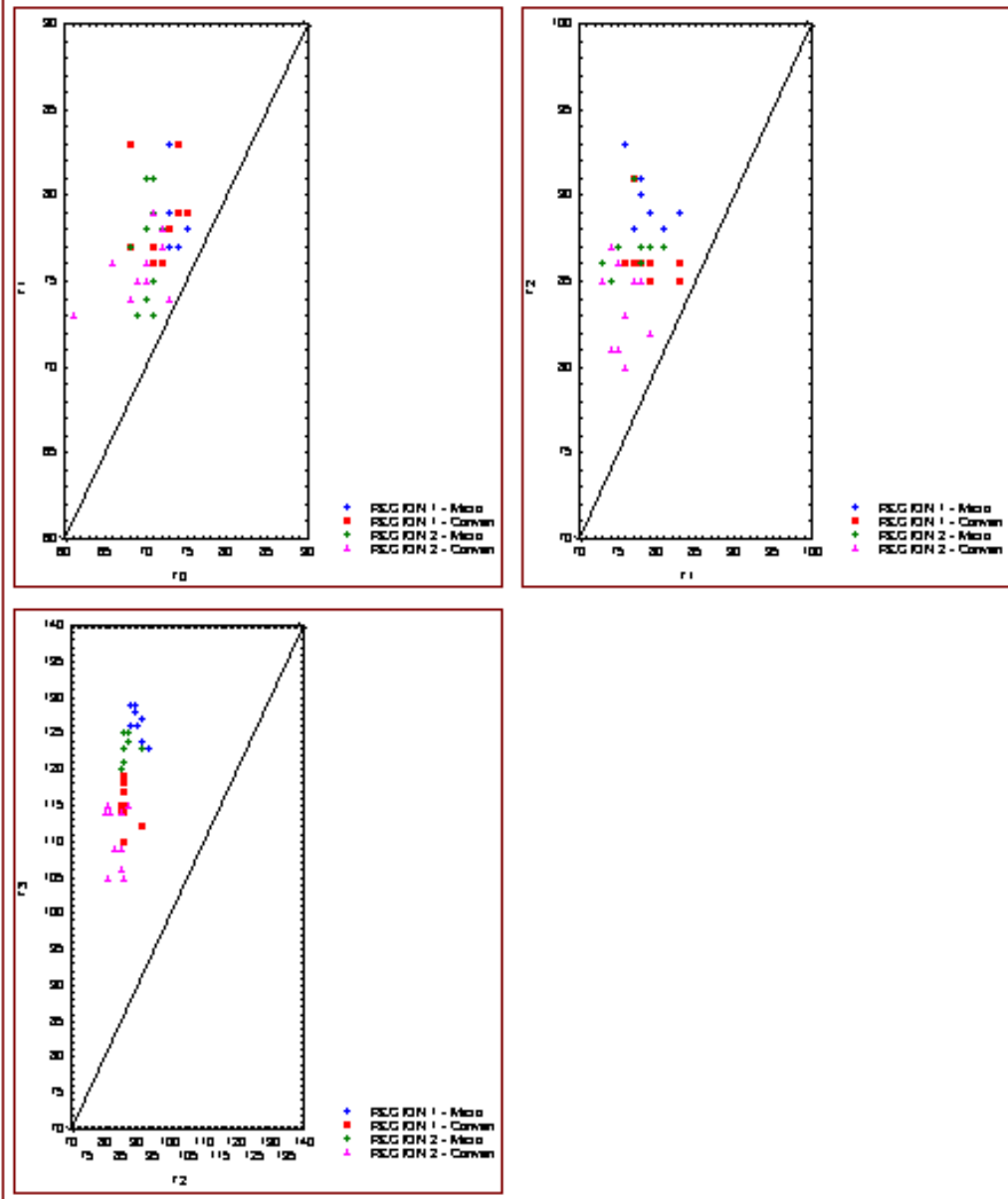


Figura 5.18. Gráficas de cambio proporcional en la variable Hojas desprendidas en los distintos periodos de tiempo (T0-T1, T1-T2 y T2-T3), considerando ambos tipos de planta (micropropagado y convencional) y regiones (Región 1 y Región 2).

Cambio Proporcional

Longitud de hoja

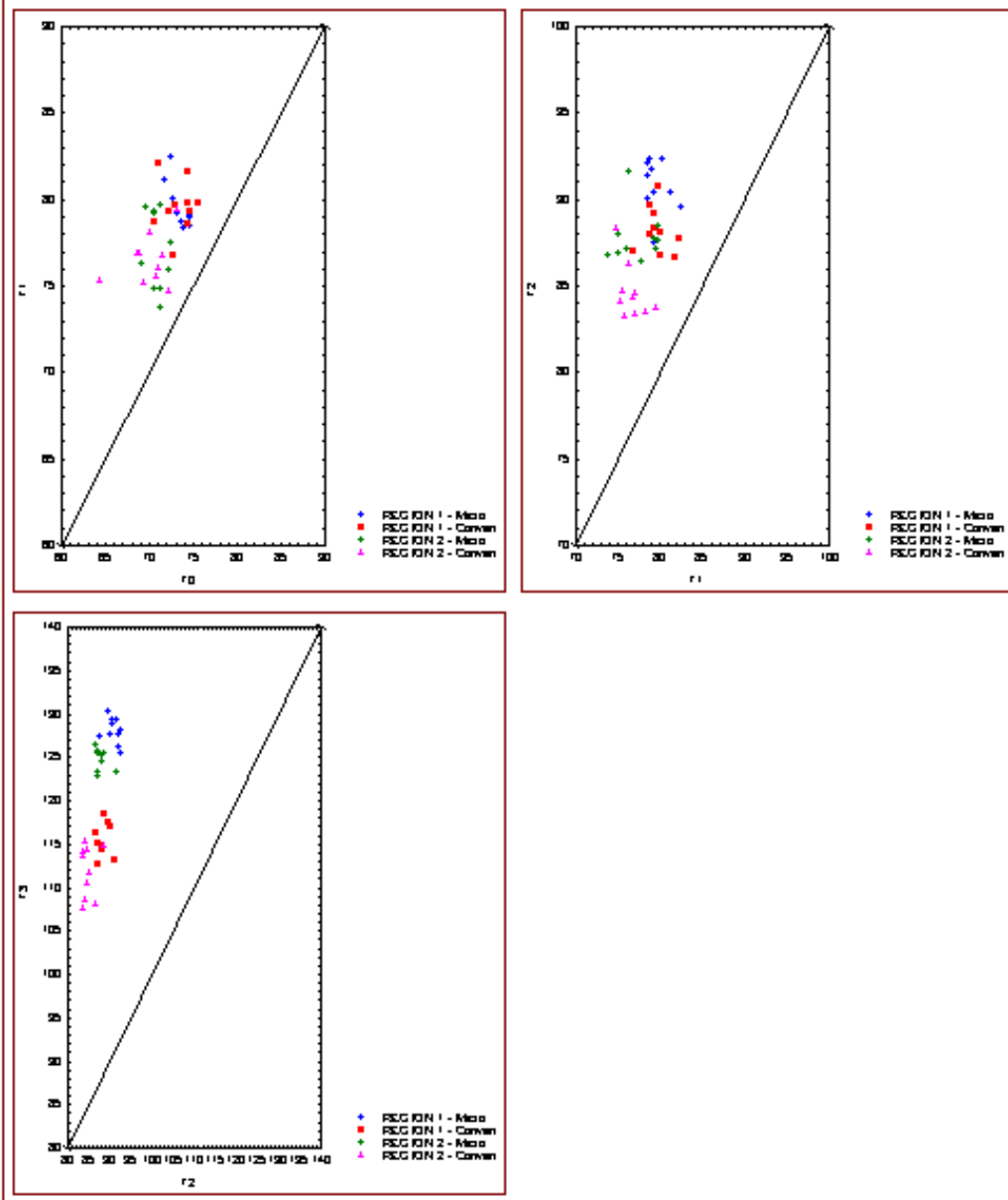


Figura 5.19. Gráficas de cambio proporcional en la variable Longitud de hoja en los distintos periodos de tiempo (T0-T1, T1-T2 y T2-T3), considerando a su vez ambos tipos de planta (micropropagado y convencional) y regiones (Región 1 y Región 2).

cada una de las variables a lo largo del tiempo, así como un crecimiento mayor en las variables para el último periodo de tiempo (tal y como se aprecia también en las gráficas longitudinales 5.15 y 516). Además destacan también algunos valores extremos presentes para ciertos tiempos de algunas de las variables, mismos que pueden identificarse fácilmente de las gráficas longitudinales referidas anteriormente.

De estos primeros análisis se observa que los datos muestran una diferencia sutil entre ambos tipos de plantas (en lo que respecta a su medias y desviaciones estándar), y a su vez entre las ambas regiones (mayores magnitudes, en general, en las medias de la Región 1 con respecto a las presentes en la Región 2). A continuación se presenta el análisis de crecimiento para estos conjuntos de datos.

5.4.2 *Análisis de crecimiento.*

El procedimiento a seguir en esta sección será análogo al que se llevo a cabo en la sección anterior con los datos correspondientes al Escenario 1. Por tanto, para contrastar el crecimiento de las plantas en este Escenario (2) se calcularán los índices morfológicos de las planta en cuestión para cada una de las fechas consideradas. Para esto, se emplearán las matrices de correlación estimadas a partir de estos datos (Tabla 7 del Apéndice A.1), con base en la metodología descrita en la sección 5.3.2 para el cálculo de una matriz de correlación común a todas las plantas. De esta forma, utilizando la expresión (5.6) la matriz ponderada resultante fue:

$$R_{Comun} = \begin{bmatrix} 1 & 0.9397 & 0.9771 \\ 0.9397 & 1 & 0.9285 \\ 0.9771 & 0.9285 & 1 \end{bmatrix}, \quad (5.16)$$

donde la matriz (5.16) fue calculada a partir de las matrices estimadas con los datos de las fechas T0, T1, T2 y T3 (empleando 40 datos en cada caso). De esta forma, dicha matriz se utiliza para el cálculo de los coeficientes de las componentes principales, donde, los resultados obtenidos (aplicando la metodología descrita en la sección 3.2.1) se presentan en la Tabla 5.12 de arriba, en la cual se pueden apreciar los coeficientes correspondientes a cada una de las componentes principales, así como el porcentaje de varianza explicada por cada una de éstas.

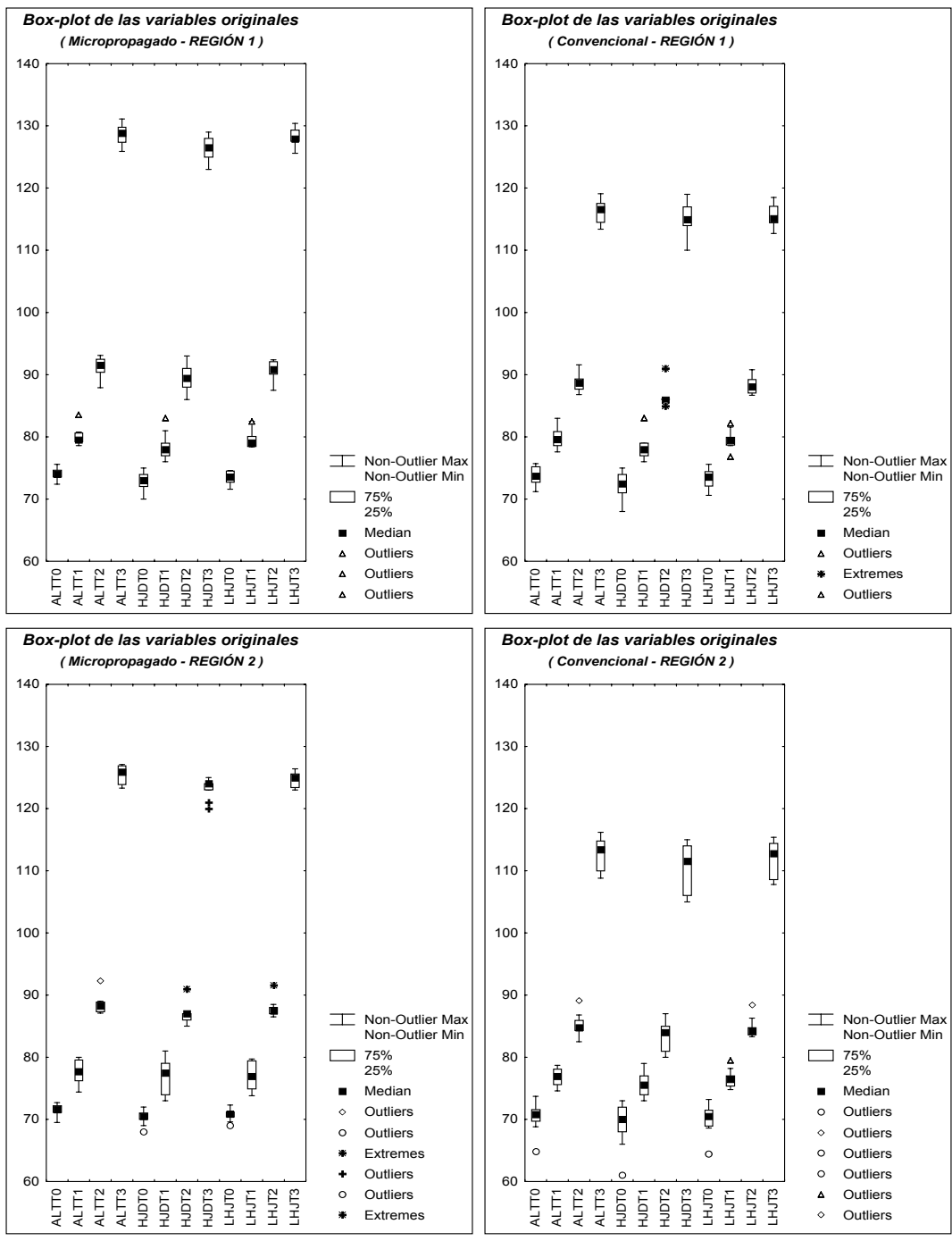


Figura 5.20. Gráficas de caja y bigotes de las variables originales como función del tiempo, clasificadas por región y tipo de planta (Escenario 2).

Tabla 5.12. Resultados obtenidos del análisis de componentes principales empleando los datos correspondientes al Escenario 2.

	Vectores Propios		
	VP1	VP2	VP3
Altura	0.5814	0.3227	-0.7469
Hojas Desp.	0.5715	-0.8154	0.0925
Longitud Hj.	0.5792	0.4806	0.6585
Valor Propio	2.90	0.08	0.02
% Varianza	96.57	2.69	0.74

De la Tabla 5.12 cabe hacer notar que al igual que en la sección anterior los coeficientes correspondientes a la primer componente principal (VP1) son todos positivos, acorde con lo descrito en la sección 3.2.2, y éstos son utilizados para el cálculo de los índices morfológicos de las plantas, mismos que son incluidos en la Tabla 8 del Apéndice A.1.

Por tanto, un primer análisis de carácter exploratorio para los índices obtenidos tiene lugar en las gráficas longitudinales de la Figura 5.21, donde se puede observar que los índices (al igual como ocurría en las variables originales) presentan una tendencia positiva en su crecimiento para ambas regiones y tipos de planta. Así mismo, cabe destacar la diferencia palpable entre los dos tipos de planta de cada región, al notarse de las gráficas aparentemente un mayor crecimiento en las plantas del tipo micropropagado (de ambas regiones), sobretodo en la última fecha (T3).

Por otro lado, la Figura 5.22 muestran el cambio proporcional de los índices morfológicos para los distintos periodos de tiempo, donde se puede observar que todos los conjuntos de datos están por arriba de la recta de 45 grados, indicando así un aumento en los índices de una fecha a otra y como sería de esperarse al estar midiendo crecimiento. De estas gráficas se puede apreciar también una diferencia en las magnitudes de los índices para los últimos dos periodos de tiempo (T1-T2 y T2-T3), donde se puede apreciar que los conjuntos de datos se separan en regiones y tipos, teniéndose incluso que las plantas del tipo micropropagado tienden a estar por arriba de las del tipo convencional en ambas regiones (siendo más claro esto en la gráfica T2-T3). Así mismo, se observa una diferencia entre las dos regiones en cuestión, sobretodo para las últimas fechas, al tenerse que los índices en la Región 1 son mayores a los de la Región 2 al comparar entre mismos tipos de planta.

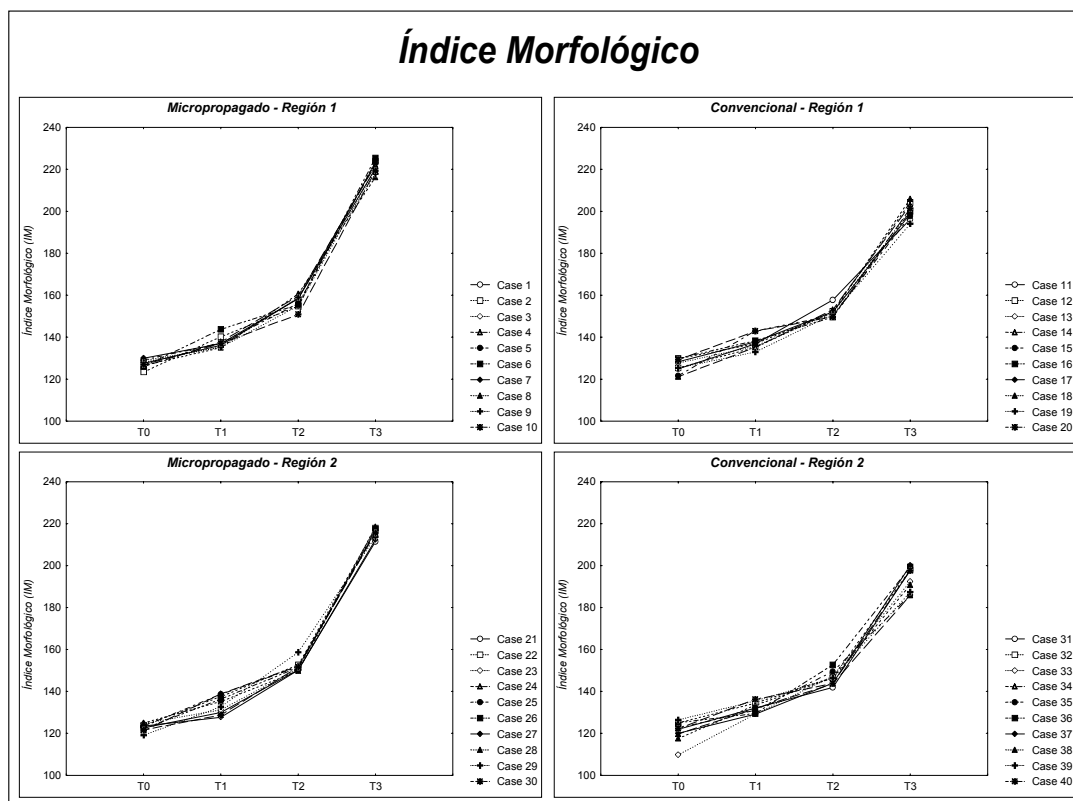


Figura 5.21. Gráficas longitudinales de los índices morfológicos de las plantas en las Regiones 1 y 2, medidas en las fechas T0, T1, T2 y T3, correspondientes al Escenario 2.

En la Tabla 5.13 se presentan las estadísticas descriptivas (medias y desviaciones estándar) de los índices morfológicos correspondientes a este escenario. De la tabla se puede apreciar que las medias en efecto muestran de manera numérica diferencias entre los índices de ambos tipos de planta, teniéndose para ambas regiones diferencias pequeñas en las fechas T0 y T1 (a excepción de T1 en la Región 1), mientras que éstas son mayores en las fechas T2 y T3. Así mismo, nótese que las medias en ambos tipos de planta de la Región 1 son superiores a las propias en la Región 2. Cabe hacer notar también el aumento en las medias de los índices conforme el tiempo aumenta, acorde con lo observado de gráficas anteriores. En lo que respecta a las desviaciones estándar se puede observar el mismo comportamiento que en la variables originales, al contar en general con mayores desviaciones en las plantas del tipo convencional con respecto a las del tipo micropropagado (a excepción de T2 en la Región 1 y T1 en la Región 2), mostrando por tanto una mayor variabilidad en el primer tipo.

Cambio Proporcional Índices Morfológicos

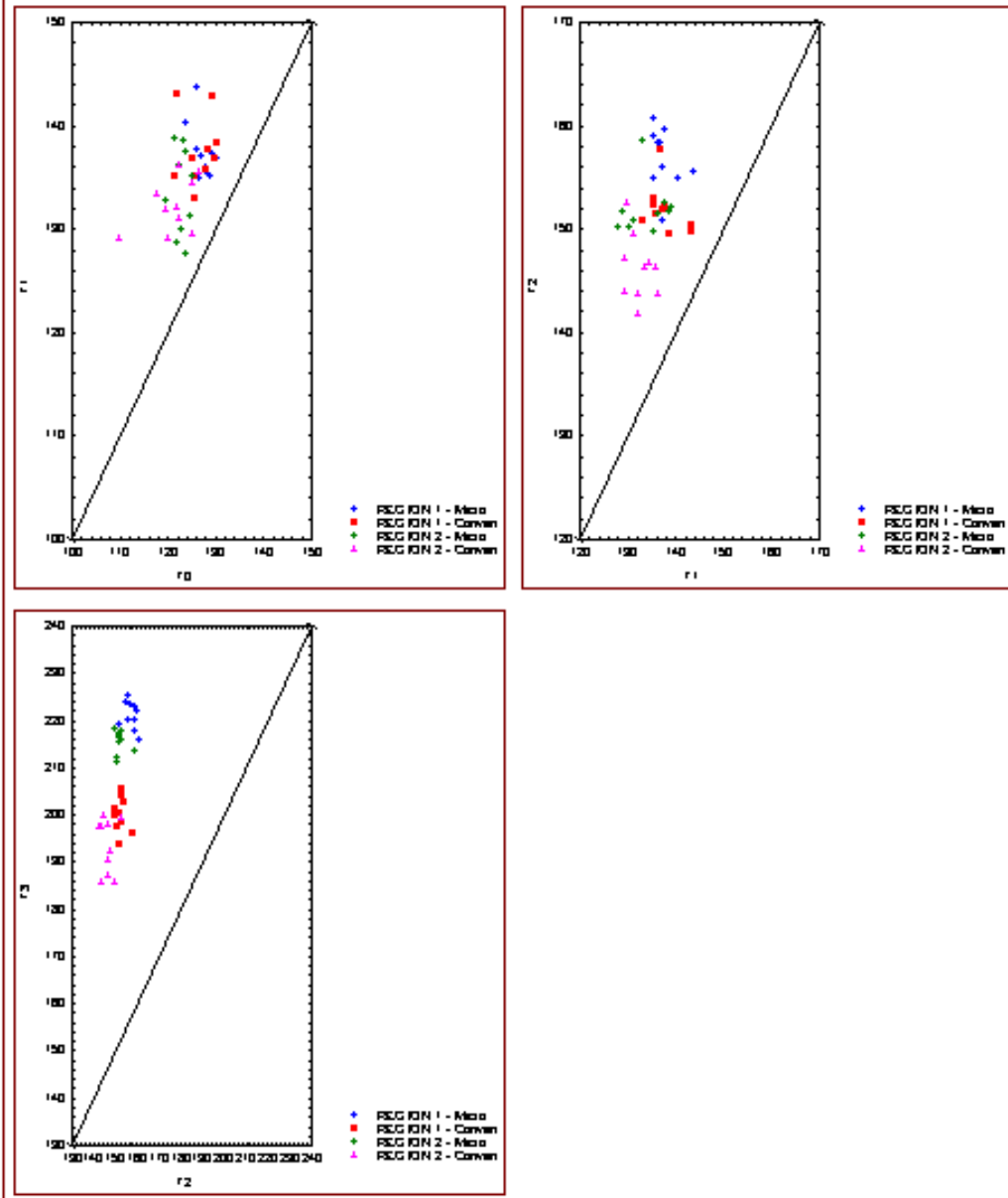


Figura 5.22. Gráficas del cambio proporcional de los índices morfológicos en los distintos periodos de tiempo (T0-T1, T1-T2 y T2-T3), para ambos tipos de planta (micropropagado y convencional) en ambas regiones (Región 1 y Región 2).

Tabla 5.13. Estadísticas descriptivas por región y tipo de planta para los índices morfológicos de las plantas (Escenario 2) en cada una de las fechas T0, T1, T2 y T3.

Región	Tipo	Estadísticas Descriptivas	Índices Morfológicos			
			T0	T1	T2	T3
1	Micro	Media	127.3	137.5	156.9	221.2
		Desv. Estand.	1.9	2.7	2.9	2.9
	Conven	Media	126.4	137.5	152.0	200.2
		Desv. Estand.	3.2	3.3	2.3	3.7
2	Micro	Media	122.7	133.7	152.0	215.7
		Desv. Estand.	1.7	4.2	2.5	2.4
	Conven	Media	121.1	132.3	146.3	193.6
		Desv. Estand.	4.8	2.6	3.2	5.7

Finalmente, en la Figura 5.23 se pueden apreciar las gráficas de caja y bigotes (box-plots) correspondientes a los índices en cada región, donde se puede observar la tendencia positiva de éstos a lo largo del tiempo, así como las diferencias en variabilidad mencionadas entre ambos tipos de planta para cada periodo. A su vez, se aprecian algunos valores extremos presentes en algunos de los periodos de tiempo, mismos que repercuten en algunos de los resultados presentados en la tabla anterior, como por ejemplo, el índice extremo obtenido para la planta identificada con el caso 33 del tipo convencional correspondiente a la Región 2 en la fecha T0, la cual contribuye en gran medida a que se obtenga un valor tan grande en la variación mostrada en la tabla antedicha para la fecha referida.

De los análisis exploratorios anteriores se observa que para este caso los índices morfológicos muestran en general una tendencia positiva en su comportamiento a lo largo del tiempo, incluso reflejando algunas diferencias existentes entre las dos regiones geográficas en cuestión. Además, se pudo apreciar que los índices muestran una diferencia sutil entre los dos tipos de planta de cada región, sobretodo para las últimas fechas de las mediciones, siendo mayores, en general, en el tipo micropropagado. A su vez, se apreciaron diferencias en la variabilidad de cada tipo de planta, siendo mayor en el tipo convencional.

Como se ha visto, estos índices morfológicos han reflejado a su vez la información presente en las variables originales, empleando así una variable en lugar de tres (como se comentaba en la sección 3.2.2).

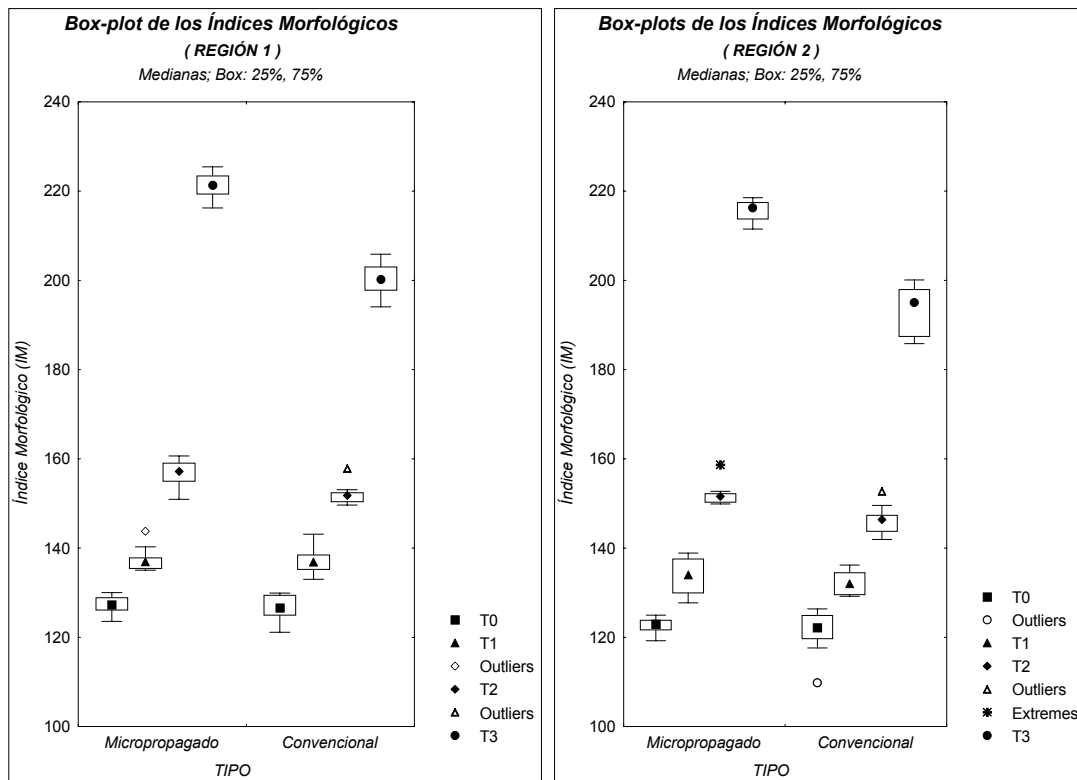


Figura 5.23. Gráficas de caja y bigotes (box-plots) de los índices morfológicos de las plantas, graficados a lo largo del tiempo para cada región y tipo de planta con los datos correspondientes al Escenario 2.

Al igual que en el Escenario 1, se realizará un análisis adicional para explorar el poder de discriminación en las variables originales, y ver así si es posible diferenciar a los dos tipos de plantas (micropropagado y convencional) empleando la información presente en dichas variables. Para ello se empleará nuevamente la técnica de *Análisis Discriminante* descrita en el Apéndice A.4, la cual consiste en la construcción de funciones lineales de discriminación para llevar a cabo la clasificación de las plantas.

Los datos a emplear nuevamente corresponden a las observaciones originales (simuladas) y presentadas en la Tabla 7 del Apéndice A.1, donde se tienen mediciones de las variables *Altura*, *Número de hojas desprendidas* y *Longitud de hoja*, tomadas de plantas de agave de las Regiones 1 y 2. El análisis se realizará considerando a cada una de las regiones por separado, a raíz de lo observado en análisis previos. Así mismo, los análisis se realizan para cada una de las fechas T0, T1, T2 y T3 de manera individual.

Tabla 5.14. Resultados obtenidos del análisis discriminante basándose en los datos correspondientes al Escenario 2.

Región	Fecha	Coeficiente		Región	Fecha	Coeficiente	
1	T0	Altura	-1.2627	2	T0	Altura	3.4328
		Hojas Desp.	5.9341			Hojas Desp.	1.6479
		Long. Hoja	-4.0360			Long. Hoja	-4.0644
		Constante	-40.1990			Constante	-72.4530
	T1	Altura	2.1463		T1	Altura	-0.5330
		Hojas Desp.	0.0335			Hojas Desp.	4.8558
		Long. Hoja	-2.2083			Long. Hoja	-3.3978
		Constante	1.3125			Constante	-68.2292
	T2	Altura	-2.4246		T2	Altura	5.3256
		Hojas Desp.	7.3329			Hojas Desp.	2.1929
		Long. Hoja	-1.8895			Long. Hoja	-4.0899
		Constante	-257.5839			Constante	-296.5103
T3	Altura	1.8708	T3	Altura	4.4191		
	Hojas Desp.	-1.7506		Hojas Desp.	5.0063		
	Long. Hoja	12.3775		Long. Hoja	3.8136		
	Constante	-1525.3941		Constante	-1563.1799		

También en este caso se empleará la matriz de correlación común (5.16), ya que uno de los supuestos necesarios es que las matrices de covarianzas (correlación) de las poblaciones a ser comparadas sean las mismas. Cabe mencionar que para estos datos, al igual que con los correspondientes al Escenario 1, las matrices de correlación de ambas poblaciones presentaban la misma estructura en cada periodo de tiempo y región, por lo cual es viable utilizar la matriz de correlación común en este estudio.

En la Tabla 5.14 se presentan los coeficientes obtenidos para cada una de las funciones de discriminación en cada periodo de tiempo y región. Con estos coeficientes se construyeron las funciones de discriminación para cada caso (las cuales no se incluyen explícitamente aquí), siendo evaluadas posteriormente cada una de las plantas (las mediciones de sus variables) en la función correspondiente y registrando el tipo en que era clasificado de acuerdo con la regla descrita en la sección 5.3.1. Los resultados obtenidos son presentados en la Tabla del 9 del Apéndice A.1, y una versión resumida de ésta es dada en la Tabla 5.15. En ambas tablas destaca el hecho de que en algunos casos, como en las Regiones 1 y 2 en las fechas T2 y T3, las variables

Tabla 5.15. Resultados obtenidos al aplicar las funciones de clasificación a la muestra de plantas de agave en el Escenario 2.

REGION 1 - T0				REGION 2 - T0			
Tipo		Clasificado		Tipo		Clasificado	
		Micro	Conven			Micro	Conven
Original	Micro	7	3	Original	Micro	8	2
	Conven	3	7		Conven	6	4

REGION 1 - T1				REGION 2 - T1			
Tipo		Clasificado		Tipo		Clasificado	
		Micro	Conven			Micro	Conven
Original	Micro	7	3	Original	Micro	5	5
	Conven	6	4		Conven	3	7

REGION 1 - T2				REGION 2 - T2			
Tipo		Clasificado		Tipo		Clasificado	
		Micro	Conven			Micro	Conven
Original	Micro	7	3	Original	Micro	10	0
	Conven	1	9		Conven	2	8

REGION 1 - T3				REGION 2 - T3			
Tipo		Clasificado		Tipo		Clasificado	
		Micro	Conven			Micro	Conven
Original	Micro	10	0	Original	Micro	10	0
	Conven	0	10		Conven	0	10

Tabla 5.16. Probabilidades de mala clasificación estimadas para cada una de las funciones de clasificación construidas con base en los coeficientes presentados en la Tabla 5.14.

Región	Fecha	Función de clasificación	Probabilidad de Mala clasificación
1	T0	3.1	0.40
	T1	3.2	0.45
	T0	3.3	0.30
	T1	3.4	0.45
2	T0	3.5	0.30
	T1	3.6	0.25
	T0	3.7	0.20
	T1	3.8	0.40

permitieron diferenciar muy bien a cada tipo de planta, en comparación con las primeras dos fechas donde el número de plantas mal clasificadas es mayor. Esto mismo puede ser apreciado en la Tabla 5.16 donde se presentan estimaciones de la probabilidad de mala clasificación calculadas con base en la expresión (5.12). De la última tabla referida se aprecia lo comentado con anterioridad, al resultar las probabilidades de mala clasificación mayores para las dos primeras fechas (T0 y T1), mientras son mucho menores en las dos últimas fechas (T2 y T3), lo cual ocurre similarmente en ambas regiones. Esto indicaría que es más sencillo separar a los dos conjuntos de datos (plantas del tipo micropropagado y convencional) de acuerdo a la información en sus variables para las últimas fechas, donde se pudo apreciar de gráficas anteriores que se muestra una mayor separación entre los conjuntos de datos de cada tipo.

A continuación, comenzaremos con el análisis de crecimiento de estos datos según lo descrito en las secciones 3.4.1 y 3.4.2, iniciando para esto con el análisis de las tasas de crecimiento temporales y relativas promedio.

5.4.2.1 Comparación empleando la tasa de crecimiento.

En esta sección se comparará el crecimiento de los dos tipos de planta (micropropagado y convencional) con base en los datos del Escenario 2. Para ello se empleará la tasa de crecimiento de las plantas, calculada con base en los índices morfológicos de las mismas. Como se mencionó en la sección 3.4.1, se consideran para el estudio dos tipos de tasa de crecimiento, las temporales y las relativas promedio, las cuales serán analizadas cada una por su cuenta en las secciones siguientes.

5.4.2.1.1 Tasa de crecimiento temporal.

Las tasas de crecimiento temporales analizadas a continuación fueron calculadas con base en los índices morfológicos derivados de los datos correspondientes al Escenario 2 (Tabla 7, Apéndice A.1). Las tasas obtenidas (utilizando la expresión 3.20 del capítulo 3) son presentadas en la Tabla 10 del Apéndice A.1, siendo calculadas éstas para ambos tipos de planta (micropropagado y convencional) en cada región geográfica (Región 1 y Región 2) en los tres periodos de tiempo en cuestión (T0-T1, T1-T2 y T2-T3).

Una representación gráfica de dichas tasas se puede apreciar en la Figura 5.24, donde se presentan las tasas de ambos tipos de planta en cada periodo de tiempo. De las gráficas referidas puede notarse una diferencia clara entre las tasas de ambos tipos de planta (tipo micropropagado superiores a las del tipo convencional), sobretodo en el último periodo de tiempo T2-T3. Otro punto a notar sobre estas gráficas es que la variabilidad entre ambos tipos de planta para cada periodo de tiempo se aprecia muy similar, sin embargo, esto será verificado más adelante con una prueba más formal.

Entonces, lo que nos interesa es saber si los datos reflejan diferencias entre las tasas de crecimiento temporales de ambos tipos de planta, por lo cual se plantea la siguiente hipótesis con base en el conocimiento actual de la situación:

H_T : *las medias de las tasas de crecimiento temporales promedio en el periodo de tiempo T-t son iguales para ambas poblaciones de plantas (agave micropropagado y agave convencional).*

Los periodos de tiempo T-t en la hipótesis anterior corresponden a los periodos T0-T1, T1-T2 y T2-T3, por lo cual se realizarán 3 pruebas de este estilo para cada una de las regiones en cuestión.

En este caso también se considera una prueba no-paramétrica de suma de rangos (para medias) empleando el estadístico de *Wilcoxon-Mann-Whitney*, para probar así la hipótesis H_T al estilo de la prueba desarrollada en la sección anterior con las tasas de crecimiento temporales del Escenario 1. El supuesto de continuidad se considera factible por lo mencionado en 3.4.1.

Los resultados de las pruebas de homogeneidad de varianzas y suma de rangos son mostrados en la Tabla 5.17, donde en primer lugar se puede apreciar de las pruebas de homogeneidad que en el periodo de tiempo T2-T3 en la Región 2 se obtienen como *p-valores*: 0.032 (prueba de Levene) y 0.049 (prueba de Brown-Forysthe), mostrando así evidencia en contra de la hipótesis de homogeneidad en el caso de Levene, no obstante, la prueba de Brown-Forysthe es más robusta (al utilizar medianas), por lo que podría considerarse que la evidencia obtenida se encuentra dentro de un rango permisible, y por tanto se concluiría que no existe evidencia para rechazar la hipótesis de homogeneidad de varianzas en este caso. En los periodos restantes de ambas regiones tampoco existe evidencia suficiente para rechazar la hipótesis de homogeneidad de varianzas entre ambos tipos de planta.

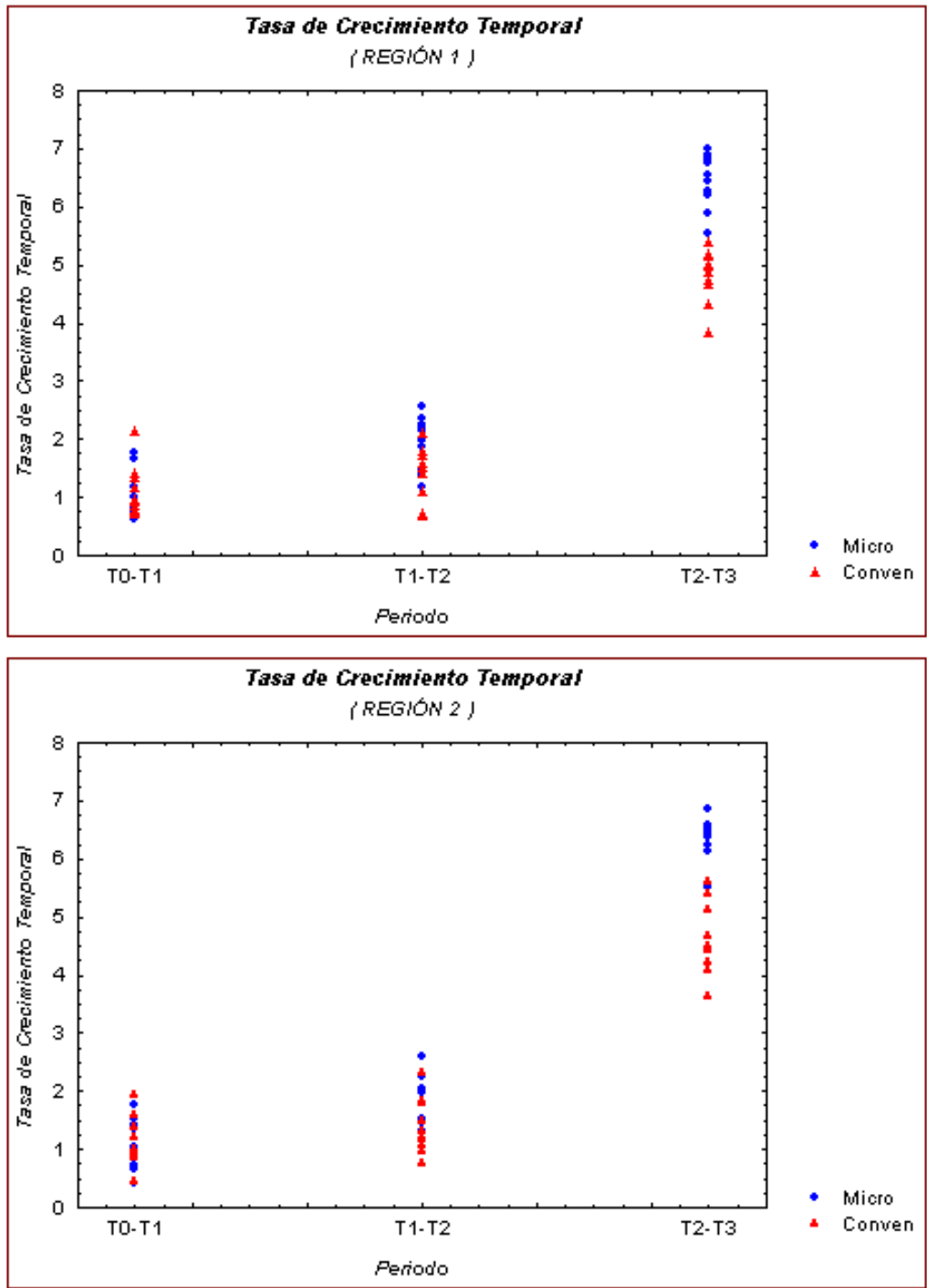


Figura 5.24. Tasas de crecimiento temporales de ambos tipos de planta, graficadas en cada periodo de tiempo y región, correspondientes a los datos simulados en el Escenario 2.

Tabla 5.17. Resultados obtenidos para las pruebas de homogeneidad de varianza y sumas de rangos de las tasas de crecimiento temporales del Escenario 2.

Región	Periodo	Pruebas Homogeneidad de Varianzas				Pruebas		
		Desviaciones Estándar		p-valor		Wilcoxon-Mann-Whitney		
		Micro	Conven	Levene	Brn-Fors	R	R'	p-valor
1	T0 - T1	0.4017	0.4340	0.9040	0.8907	95.5	114.5	> 0.20
	T1 - T2	0.4542	0.4657	0.9579	0.9402	133.0	77.0	0.01 < p < 0.05 *
	T2 - T3	0.4649	0.4592	0.8528	0.7952	155.0	55.0	< 0.01 *
2	T0 - T1	0.4498	0.4222	0.5106	0.4693	103.5	106.5	> 0.20
	T1 - T2	0.4571	0.4681	0.7682	0.6837	133.0	77.0	0.01 < p < 0.05 *
	T2 - T3	0.3670	0.6757	0.0302 *	0.0499 *	153.0	57.0	< 0.01 *

En lo que respecta a las pruebas de suma de rangos, de la tabla se puede apreciar que se rechaza la hipótesis de igualdad para los periodos de tiempo T1-T2 y T2-T3 (en ambas regiones), mientras que para el periodo T0-T1 no existe evidencia suficiente para rechazar dicha hipótesis.

Por tanto, a la luz de los datos analizados para ambas regiones existe evidencia suficiente para rechazar la hipótesis H_T para los periodos de tiempo T1-T2 y T2-T3.

A continuación, analizaremos el comportamiento de las tasas de crecimiento relativas promedio calculadas a partir de los mismos datos con que fueron calculadas las tasas de crecimiento temporales analizadas en esta sección.

5.4.2.1.2 Tasa de crecimiento relativa promedio.

Como se mencionó, los datos a analizar en esta ocasión corresponden a las tasas de crecimiento relativas promedio, calculadas a partir de los índices morfológicos correspondientes al Escenario 2 (Tabla 8 del Apéndice A.1). Dichas tasas fueron calculadas con base en la expresión (3.22) del capítulo 3, y los resultados son presentados en la Tabla 11 del Apéndice A.1.

Para su análisis comenzaremos de manera análoga a secciones anteriores, es decir, iniciando con un análisis gráfico de las tasas de crecimiento. En la

Figura 5.25 se pueden apreciar las tasas de ambos tipos graficadas en cada periodo de tiempo. De éstas es posible observar un comportamiento similar al ocurrido con las tasas de crecimiento temporales de la sección anterior, notándose una diferencia entre las tasas de ambos tipos sobretodo en los últimos dos periodos de tiempo, predominando las tasas del tipo micropropagado por arriba de las del tipo convencional,. Con respecto a la variabilidad se tiene que a primera vista ésta es similar entre ambos tipos de planta, sin embargo, esto será verificado con una prueba formal más adelante

De esta forma, la hipótesis a probar puede plantearse como:

H_R : *las medias de las tasas de crecimiento relativas promedio en el periodo de tiempo T-t son iguales para ambas poblaciones de plantas (agave micropropagado y agave convencional).*

Para lo cual en este caso también se realizarán tres en cada región, una para cada periodo de tiempo T0-T1, T1-T2 y T2-T3, donde de nueva cuenta se utilizará la prueba de suma de rangos de *Wilcoxon-Mann-Whitney* para probar dicha hipótesis. Como se mencionó en la sección 3.4.1, el supuesto de continuidad, necesario para la prueba en cuestión, es factible suponerse de la definición de las tasas en la expresión 3.22.

Los resultados de las pruebas realizadas son presentados en la Tabla 5.18, donde se puede observar de las pruebas de homogeneidad en cada periodo que únicamente en el caso de la prueba de Levene para el periodo T2-T3 de la Región 2 se rechazaría la hipótesis de homogeneidad, sin embargo, para este mismo periodo la prueba de Brown-Forysthe no muestra evidencia para rechazar la hipótesis, por tanto, al ser la prueba de Brown-Forysthe más robusta que la de Levene (pues esta última es más sensible a datos extremos por utilizar diferencias absolutas a la media) se considera que no hay evidencia para rechazar la hipótesis de igualdad de varianzas entre las tasas de crecimiento relativas promedio de ambos tipos de planta.

Sobre las pruebas de sumas de rangos realizadas se observa que existe evidencia para rechazar la hipótesis de igualdad en los periodos de tiempo T1-T2 y T2-T3 de la Región 1, así como para el periodo de tiempo T2-T3 de la Región 2 (al obtenerse *p-valores* pequeños -menores a 0.05-); mientras que en los casos restantes no existe evidencia suficiente para rechazar la hipótesis de igualdad de las medias en las tasas de crecimiento relativas promedio de ambos tipos de planta.

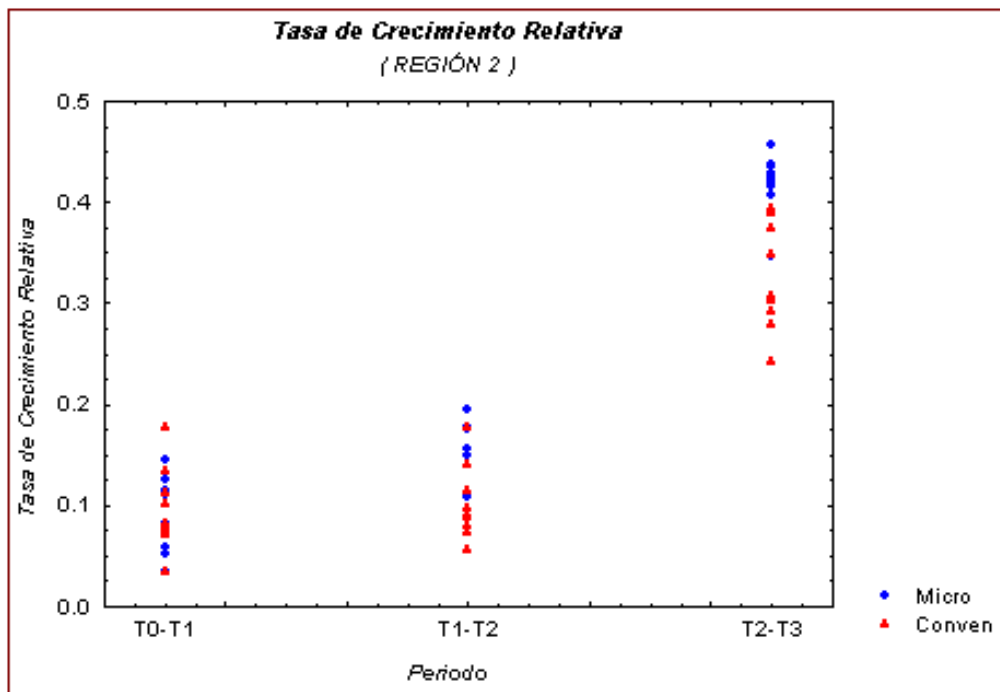
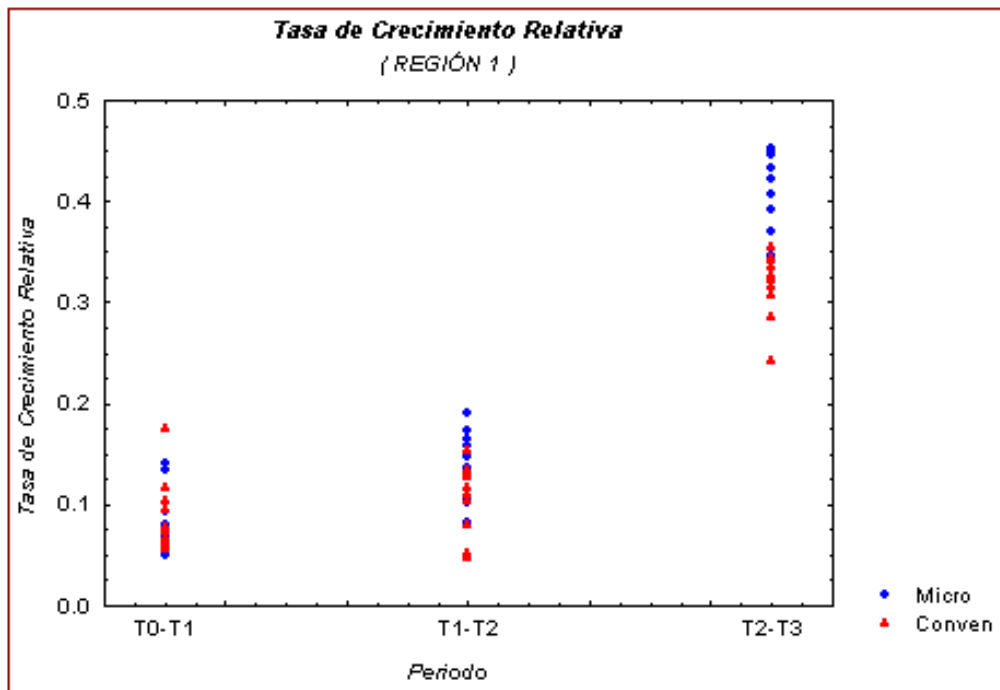


Figura 5.25. Tasas de crecimiento relativas promedio de ambos tipos de planta, graficadas en cada periodo de tiempo y región, correspondientes a los datos simulados en el Escenario 2.

Tabla 5.18. Resultados obtenidos para las pruebas de homogeneidad de varianza y sumas de rangos de las tasas de crecimiento relativas promedio del Escenario 2.

Región	Periodo	Pruebas Homogeneidad de Varianzas				Pruebas Wilcoxon-Mann-Whitney		
		Desviaciones Estándar		p-valor		R	R'	p-valor
		Micro	Conven	Levene	Brn-Fors			
1	T0 - T1	0.0328	0.0366	0.8531	0.8535	97.0	113.0	> 0.20
	T1 - T2	0.0349	0.0355	0.9612	0.9781	133.0	77.0	0.01 < p < 0.05 *
	T2 - T3	0.0359	0.0328	0.5223	0.4834	154.0	56.0	0.01 < p < 0.05 *
2	T0 - T1	0.0372	0.0395	0.7801	0.6698	103.0	107.0	> 0.20
	T1 - T2	0.0378	0.0371	0.6169	0.5600	129.0	81.0	0.05 < p < 0.10
	T2 - T3	0.0290	0.0505	0.0276 *	0.1217	151.0	59.0	0.01 < p < 0.05 *

Por tanto, a la luz de los datos analizados existe evidencia suficiente para rechazar la hipótesis H_R para los periodos de tiempo T1-T2 (Región 1) y T2-T3 (en ambas regiones).

A continuación se analizará el crecimiento de las plantas ahora utilizando las curvas de crecimiento ajustadas a los índices morfológicos de éstas, empleando de nueva cuenta los datos utilizados en el cálculo de las tasas de crecimiento analizadas anteriormente.

5.4.2.2 Comparación empleando curvas de crecimiento.

Al igual que en la sección homóloga para el Escenario 1, en esta se busca comparar el crecimiento de los dos tipos de planta, micropropagado y convencional, empleando curvas de crecimiento ajustadas a partir de los índices morfológicos de éstas. Por lo tanto, la idea a seguir aquí es similar a la descrita en la sección 5.3.2.2, donde se construyó una curva de crecimiento para cada tipo de planta para comparar su comportamiento a lo largo del tiempo, y en caso de haber diferencias (como se ha mostrado hasta ahora en los análisis de las tasas de crecimiento anteriores) se esperaba que éstas se viesen reflejadas en las curvas correspondientes.

Tabla 5.19. Resultados obtenidos del ajuste no-lineal utilizando la técnica de mínimos cuadrados y empleando los índices morfológicos correspondientes a los datos del Escenario 2.

Región	Tipo	Parámetro	Estimación	Error Estándar	Desv. Estándar del Error Residual	Prueba Normalidad	
						S-W	p-valor
1	Micro	c	120.751	0.66449	3.0910	0.9739	0.4734
		k	0.09392	0.00026			
1	Conven	c	121.902	0.76271	3.5019	0.9757	0.5336
		k	0.08888	0.00038			
2	Micro	c	116.554	0.72978	3.3924	0.9606	0.1753
		k	0.09364	0.00029			
2	Conven	c	116.691	0.98367	4.5118	0.9659	0.2648
		k	0.08850	0.00050			

El modelo de crecimiento propuesto para este caso es el mismo al expresado en la expresión (5.13), es decir, un modelo exponencial como el siguiente:

$$y_{ij} = c + \exp(k \times t_j) + e_{ij}, \quad (5.17)$$

donde c representa el valor inicial en el origen, k la tasa de crecimiento, $i = 1, 2, \dots, n$ (identificador de las plantas), y $j = 0, 1, 2, 3$ (identificador del tiempo, con, $t_0=T0$, $t_1=T1$, $t_2=T2$, y $t_3=T3$).

Para el ajuste del modelo (5.17) se utilizó también el método de mínimos cuadrados para la obtención de los parámetros de interés, c y k , los cuales a su vez servirán para la construcción de las curvas de crecimiento promedio de cada uno de los tipos de planta. Los datos considerados para el ajuste fueron los índices morfológicos del Escenario 2, mismos que se pueden apreciar en la Tabla 8 del Apéndice A.1. Debido a que las plantas de ambas regiones (Región 1 y Región 2) presentan características diferentes entre si (como se ha visto de análisis anteriores), se realizaron ajustes de manera individual para cada una de las regiones. Para los ajustes se programó una rutina similar a la empleada en la sección 5.3 para los ajustes presentados en el Escenario 1, utilizando en este caso también el modulo nlm (non-linear model) del software *S-PLUS 6.1*.

Los resultados de los ajustes realizados se presentan en la Tabla 5.19, donde es posible apreciar los parámetros estimados para cada tipo de planta en cada región. De las estimaciones obtenidas se puede apreciar que en lo concerniente al parámetro c (valor inicial en el origen) entre los dos tipos de planta las diferencias son sutiles en cada región, sin embargo, entre las dos

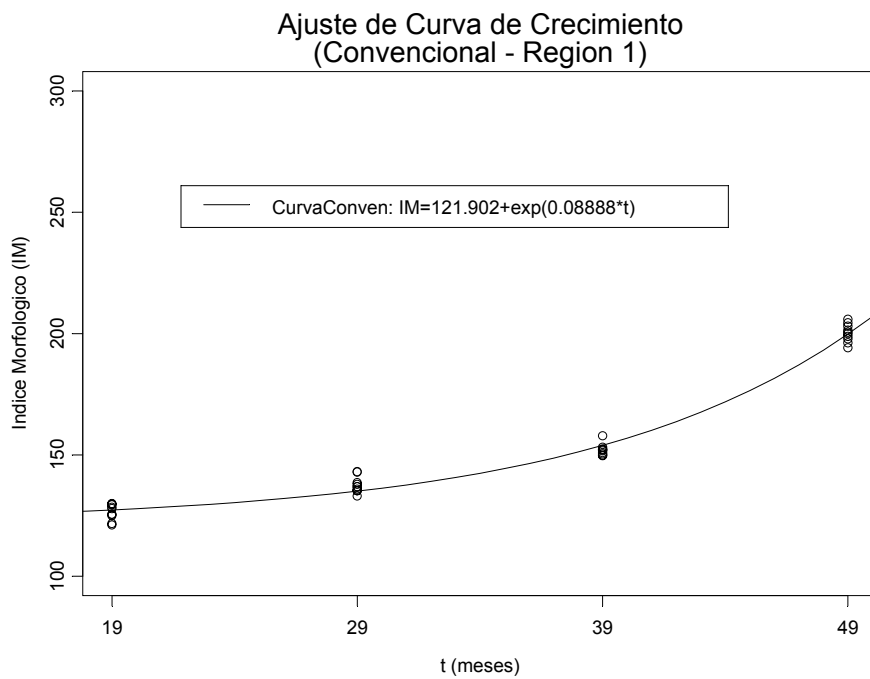
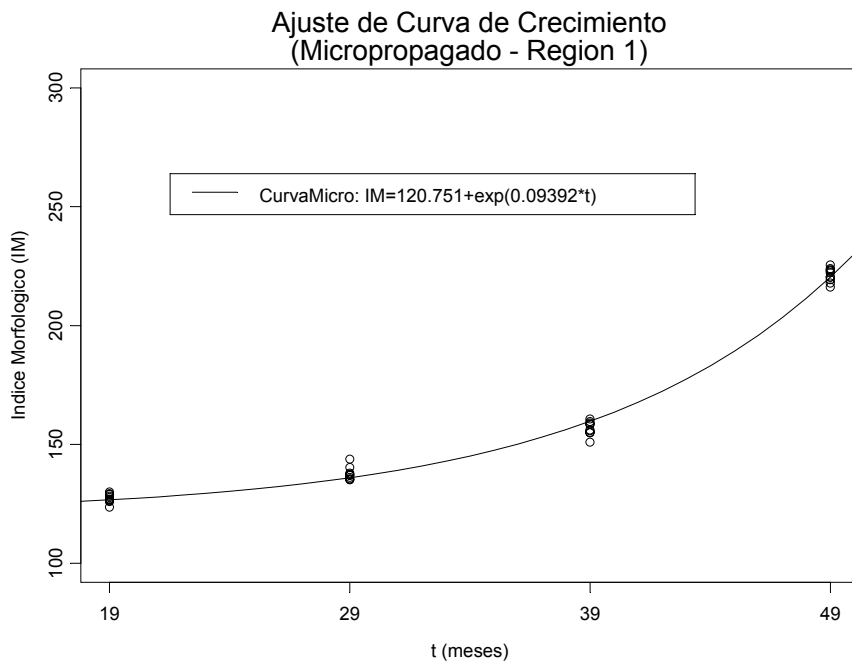


Figura 5.26. Ajuste de la curva de crecimiento para cada tipo de planta, obtenido con base en los índices morfológicos derivados de los datos del Escenario 2 (Región 1).

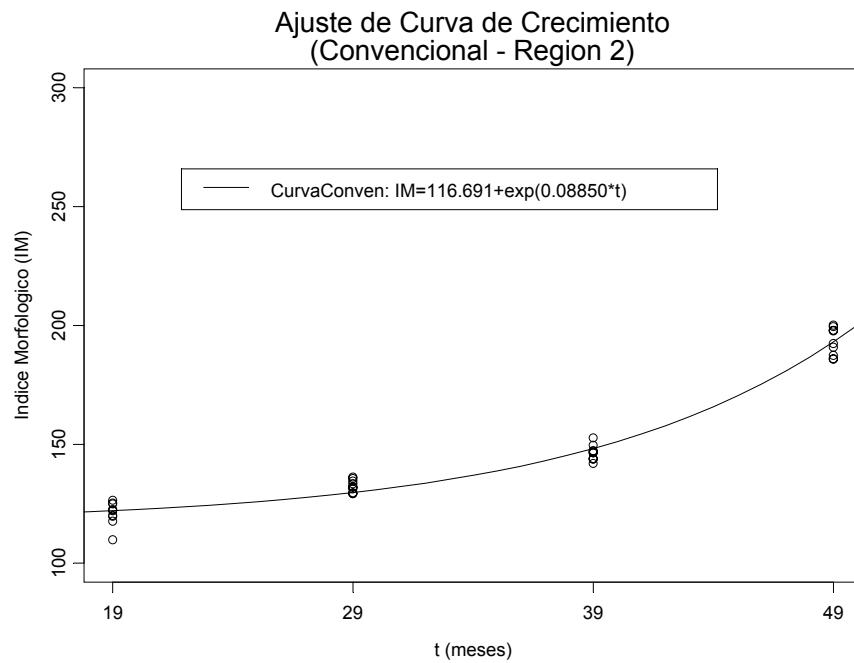
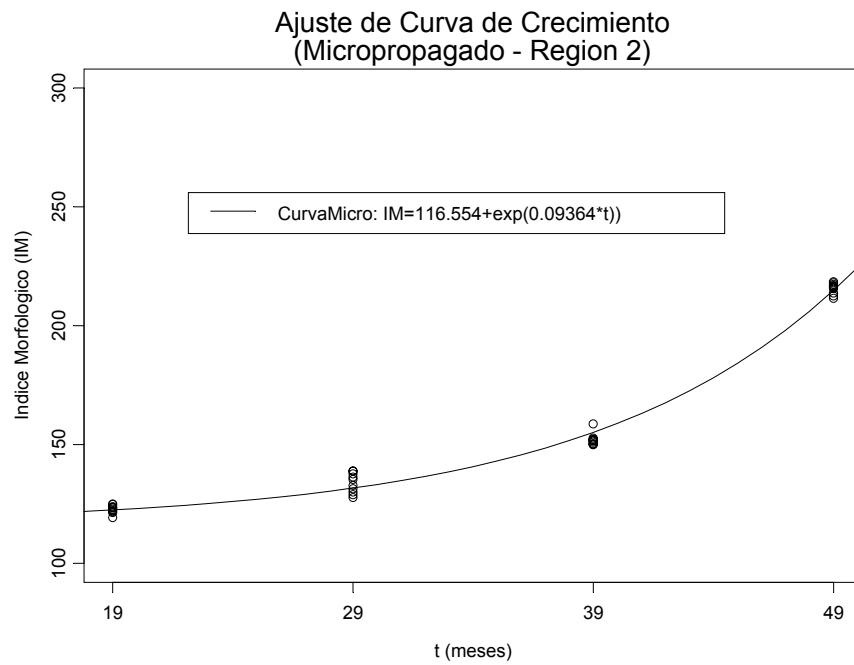
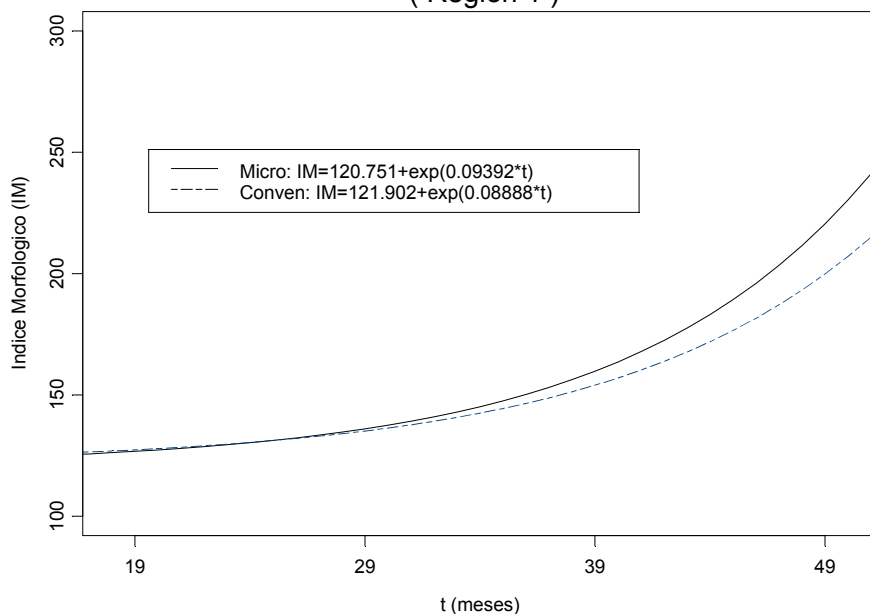


Figura 5.27. Ajuste de la curva de crecimiento para cada tipo de planta, obtenido con base en los índices morfológicos derivados de los datos del Escenario 2 (Región 2).

Curvas de Crecimiento Promedio: Micropropagado y Convencional (Region 1)



Curvas de Crecimiento Promedio: Micropropagado y Convencional (Region 2)

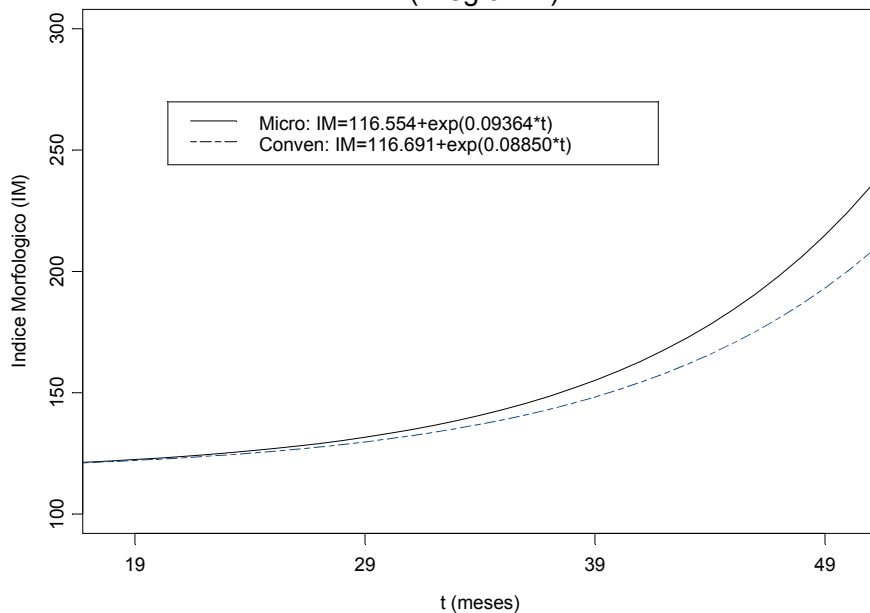


Figura 5.28. Curvas de crecimiento promedio ajustadas con base en los índices morfológicos de las plantas del Escenario 2, presentadas por regiones.

regiones se puede observar que tienden a ser mayores en la Región 1, lo cual indicaría que en dicha región las plantas presentaban mayores tamaños al momento de comenzar a tomar mediciones. De igual forma, es posible apreciar que las plantas de ambos tipos aparentemente presentaban tamaños similares al momento de llegar al campo. En lo referente al parámetro k (tasa de crecimiento) se puede notar una diferencia clara entre las estimaciones obtenidas de cada tipo de planta, siendo mayores las tasas obtenidas en las plantas del tipo micropropagado, ocurriendo por igual en cada una de las regiones. Así mismo, se observa que las tasas de un mismo tipo de planta son similares entre ambas regiones. En lo que respecta a los errores estándar de las estimaciones se tiene que éstas son pequeñas para las tasas de crecimiento, mientras que son de mayor magnitud en sus valores iniciales. Por otro lado, los errores residuales muestran magnitudes similares, a excepción de las plantas convencionales en la Región 2, donde éstas últimas tienen un error ligeramente mayor al resto, lo cual podría atribuirse a que algunas de las plantas que presentan valores extremos de sus índices, tal y como se puede apreciar de la gráfica correspondiente en la Figura 5.27.

De esta forma, con base en la información presentada en la tabla anterior, se construyen las gráficas correspondientes a las curvas de crecimiento promedio para cada tipo de planta en cada región (de acuerdo con el modelo 5.17), mismas que puede apreciarse en las Figuras 5.26 y 5.27. De estas se puede apreciar que las curvas en cada caso ofrecen un buen ajuste a los datos, donde, comparando las gráficas de ambos tipos de planta se pueden apreciar diferencias en el comportamiento de ambas curvas a lo largo del tiempo. Por su parte, la Figura 5.28 presenta una gráfica donde se contrastan las curvas promedio de ambos tipos de planta, donde se observa que en efecto ambas curvas muestran diferentes comportamientos, sobretodo en lo correspondiente a las últimas fechas. Esto muestra claramente una diferencia en la tasa de crecimiento de ambos tipos de planta, siendo mayor en las plantas del tipo micropropagado (tal y como se pudo apreciar en los parámetros estimados dados en la Tabla 5.19), en contraste con las plantas del tipo convencional.

Por tanto, estas gráficas muestran a su vez evidencia de un crecimiento distinto de ambos tipos de planta, siendo a su vez mayor en las plantas del tipo micropropagado (sobretodo para tiempos alrededor de $T_3=49$ meses). Cabe hacer notar que esto concuerda con los resultados obtenidos de los análisis de las tasas de crecimiento presentados en la sección anterior.

Sobre el análisis de los residuales obtenidos a raíz de este ajuste, se tienen las Figuras 6 y 7 del Apéndice A.2, donde son graficados los residuales como función de los valores ajustados y se puede apreciar de éstas que no existe estructura alguna que lleve a suponer que los errores en el modelo no sean constantes. Así mismo, los resultados de pruebas de normalidad para dichos residuales (empleando el estadístico Shapiro-Wilk) se muestran en la Tabla 5.19, de los cuales se puede concluir que no existe evidencia para rechazar que estos residuales provengan de una distribución normal. Por otro lado, en las Figuras 8 y 9 del Apéndice A.2 son presentadas las gráficas cuantil-cuantil (qq-plots) normales para cada conjunto de residuales, de los cuales se observa que los datos presentan colas con un poco de peso (e incluso algunos aparentes outliers), no obstante, presentando en general un buen ajuste para la parte central de la distribución; por tanto, es viable suponerse que los errores en efecto cuentan con una distribución normal.

5.4.3 Conclusiones.

En esta sección se analizó el crecimiento de las plantas de agave correspondientes a la simulación del Escenario 2 (poblaciones con crecimientos diferentes), donde se pudo observar que los análisis realizados en efecto muestran evidencia de un crecimiento distinto en las plantas de ambas poblaciones (micropropagado y convencional). Por su parte, las tasas de crecimiento mostraron, en general, diferencias significativas entre ambos tipos de planta para los periodos de tiempo T1-T2 y T2-T3. Así mismo, las curvas de crecimiento promedio de ambos tipos de planta (ajustadas a partir de los datos), mostraron diferencia significativas sobretudo en periodos de tiempo cercanos a T3. Mostrando así en ambos casos (tasas y curvas) que las plantas del tipo micropropagado presentan un mayor crecimiento con respecto a las del tipo convencional, lo cual sucede en ambas regiones.

Así mismo, de los análisis realizados es que la metodología empleada permite, además de vislumbrar diferencias entre ambos tipos de planta en lo que respecta a su crecimiento, detectar diferencias en las características de las plantas de las regiones involucradas, donde en este caso se pudo observar que las plantas en la Región 1 mostraban magnitudes superiores en sus variables (y así mismo en sus índices) con respecto de las propias en la Región 2.

Por tanto, tomando en cuenta que en este estudio de simulación los datos se generaron precisamente de forma que éstos provinieran de dos poblaciones con crecimientos diferentes, se observa de los análisis anteriores que las metodologías propuestas en efecto detectan dichas diferencias, al mostrar a su vez diferencias significativas en el crecimiento de los dos tipos de planta considerados en el estudio.

Capítulo 6

Conclusiones y Sugerencias

Este capítulo está dedicado a dar algunas conclusiones generales sobre los análisis desarrollados a lo largo de este trabajo. Así mismo, se discuten algunos de los aspectos observados en los estudios de crecimiento realizados con base en los datos simulados. Para finalizar se darán algunas sugerencias y recomendaciones para la obtención de datos bajo condiciones más favorables, mismas que podrían tomarse en cuenta para posibles estudios futuros donde se continúe con el análisis de crecimiento de plantas de agave.

6.1 Resumen del trabajo.

El objetivo primario de este trabajo consistió en dar respuesta al problema de comparar los crecimientos de dos tipos de plantas de agave en el campo: el agave *micropropagado* y el agave *convencional*. La motivación inicial para este estudio surgió a raíz de observaciones en campo, donde estos dos tipos de planta han mostrado aparentes diferencias en su desarrollo, sin que hasta la fecha se hayan realizado estudios comparativos con un sustento científico.

El objetivo es importante no solamente desde el punto de vista científico, sino también es importante para la empresa puesto que ambos tipos de plantas son actualmente utilizados y la generación de conocimiento sobre su desarrollo permitiría un mejor aprovechamiento de los recursos involucrados en los procesos de producción de agave.

En el presente estudio, personal de la empresa seleccionó una muestra de plantas de agave de ambos tipos (micropropagado y convencional), a las cuales se les midió tres variables elegidas de manera que reflejasen el crecimiento en las plantas: *Altura*, *Número de hojas desprendidas* y *Longitud de*

hoja. Se eligieron dos regiones con diferentes condiciones climáticas y geográficas, Pihuamo y Tepatitlán, para la recolección de los datos.

Las mediciones fueron tomadas periódicamente cada seis meses, desde Enero 2004 a Septiembre 2005 (a excepción de las plantas en la región de Tepatitlán donde el lote se perdió debido a una enfermedad). El análisis exploratorio de los datos (capítulo 4), mostró el efecto de factores no controlados, los cuales invalidan el uso de los datos para propósitos de probar las hipótesis sobre las diferencias en desarrollo. Los factores detectados son: condiciones diferentes en los tipos de terreno de las plantaciones, diferentes agroquímicos aplicados, técnicas agrícolas empleadas selectivamente en algunas de las plantas.

Por tanto, no fue posible aplicar a los datos originales las metodologías propuestas en el capítulo 3, desarrolladas para llevar a cabo el análisis comparativo entre los dos tipos de planta en cuestión. Sin embargo, los datos fueron de utilidad para establecer los valores iniciales utilizados en el estudio de simulación presentado en el capítulo 5, donde, con los datos generados a partir de simulaciones, se ilustraron las bondades de las metodologías propuestas.

En el estudio de simulación descrito en el capítulo 5 se plantearon dos escenarios para la generación de los datos: uno en donde las poblaciones de plantas (micropropagado y convencional) eran similares, y otro en donde éstas eran diferentes. Con los datos generados se llevó a cabo un análisis comparativo de acuerdo a las metodologías descritas y se pudo observar que dichas metodologías permiten detectar las diferencias existentes en el crecimiento de ambos tipos de planta.

Por tanto, estas metodologías ofrecen una buena opción para el análisis de crecimiento de plantas de agave donde se busque contrastar el crecimiento de ambos tipos como función del tiempo.

A continuación se presentan algunas sugerencias para evitar, en la medida de lo posible, la presencia de factores de ruido como los que afectaron las mediciones en las plantas de Pihuamo y Tepatitlán.

6.2 Sugerencias y recomendaciones.

Reiteramos las recomendaciones sugeridas en el capítulo 4:

1. Que el diseño muestral considere bloques definidos por tipos de terreno en que se encuentran las plantas, tratamientos aplicados, técnicas agrícolas empleadas, y geografía/clima.
2. Tomar las mediciones en periodos más cortos, sugiriéndose al menos tomarse mediciones cada mes o a lo más cada dos meses.
3. Aumentar el número de plantas en la muestra, considerándose al menos 30 plantas por bloque.

De contar con mediciones obtenidas bajo condiciones más controladas se esperaría que éstas presentasen una estructura similar a la de los datos simulados (analizados en el capítulo 5), en cuyo caso sería posible llevar a cabo el estudio comparativo de crecimiento en las plantas. Pudiera ocurrir que en los nuevos conjuntos de datos se presenten comportamientos que hasta ahora no han sido detectados, como por ejemplo, una correlación serial en las mediciones, efectos de bloque, etc. En ese caso, las metodologías aquí presentadas podrán ser empleadas para un primer análisis y posteriormente ser escaladas con modelos más complejos.

Puesto que el ciclo agrícola de las plantas es de varios años, se podría realizar un análisis comparativo piloto “a corto plazo” a la par con un estudio longitudinal “a largo plazo”, seleccionando para ello una muestra de plantas de diferentes edades, elegida con base en las sugerencias mencionadas con anterioridad. Una ventaja de esto sería un ahorro de tiempo para la obtención de algunos resultados que permitiesen ir observando como se esperaría que fuese el crecimiento de ambos tipos de planta en un estudio longitudinal a largo plazo; con esto sería posible la obtención de curvas de crecimiento promedio para cada tipo de planta, incluso para edades superiores a los 3 años. Sin embargo, para un estudio de este estilo deben registrarse y analizarse de manera cuidadosa los factores que tengan lugar durante el experimento.

6.3 Directrices para nuevos proyectos.

A lo largo de este trabajo se han presentado diversos análisis orientados en la búsqueda de conocer la manera en que se desarrollan las plantas de agave a lo largo del tiempo. Uno de los objetivos primordiales de este trabajo era precisamente contrastar el crecimiento de dos diferentes tipos de planta de agave (agave micropropagado y agave convencional), motivado en principio a partir de ciertas diferencias observadas en campo para los dos tipos de

planta referidos. Para llevar a cabo el análisis comparativo deseado fueron desarrolladas y propuestas en este trabajo algunas metodologías estadísticas, con las cuales es posible realizar el estudio comparativo empleando para ello mediciones obtenidas de una muestra de plantas seleccionada.

A partir de lo observado en la sección 4.5 y en el capítulo anterior es factible pensar que estas metodologías en efecto permitirán detectar diferencias (en caso de existir éstas) en el desarrollo de ambos tipos de plantas, en lo que a su crecimiento se refiere. Sin embargo, el diseño empleado para la toma de mediciones debe garantizar igualdad de condiciones entre las plantas de ambos tipos para evitar la presencia de factores de ruido que repercutan en el estudio realizado.

Además de permitir realizar un análisis comparativo entre los dos tipos de planta en cuestión, las metodologías propuestas (tasas de crecimiento y curvas de crecimiento) permitirían realizar una planeación más controlada de los recursos involucrados en los procesos de producción de agave. Por ejemplo, identificando los periodos de tiempo donde el crecimiento es mayor (en función de las tasas de crecimiento) sería posible hacer una mejor planeación sobre los tratamientos y técnicas agrícolas empleados en el campo. Un ejemplo más sería en lo concerniente al tiempo óptimo para la jima de las plantas, el cual podría ser establecido bajo un sustento científico de conocerse la curva de crecimiento promedio para todo el ciclo de vida de las plantas. Aunado a esto, un estudio que podría ser de interés sería analizar la relación entre la morfología de la planta (índices morfológicos) y variables críticas como peso y contenido de azúcares, de forma que se pudiese construir una curva de crecimiento global que involucrase a las variables referidas, permitiendo así una planeación sobre el momento óptimo para la jima, de manera que ésta se llevase a cabo en el momento en que las plantas tuviesen el mayor contenido de azúcares.

6.4 Estudios futuros.

Tomando en cuenta lo mencionado en las secciones anteriores, entre los posibles estudios futuros que podrían realizarse están:

- Generar datos a partir de un diseño muestral que siga las recomendaciones y aplicar las metodologías descritas en este trabajo para el estudio comparativo de crecimiento de los dos tipos de planta.

- Construir curvas de crecimiento para ambos tipos de plantas en edades superiores a los 2 años (hasta que se alcance el punto de inflexión en su curva de crecimiento), para ajustar así modelos de crecimiento más generales como el modelo Gompertz o Logístico presentados a su vez en el capítulo 3.
- Diseñar estudios para analizar la relación entre los índices morfológicos, el peso de las plantas y los azúcares contenidos en éstas.

Apéndice

A.1 Tablas.

Tabla 1. Mediciones en las plantas de agave los meses de Enero y Agosto del 2004 en las regiones de Pihuamo y Tepatitlán. Así mismo mediciones adicionales tomadas en las plantas de la región de Pihuamo en los meses de Enero y Septiembre del 2005.

REGION	PLANTA	TIPO	ALTURA				HOJAS DESPRENDIDAS				LONGITUD DE HOJA			
			Ene-04	Ago-04	Ene-05	Sep-05	Ene-04	Ago-04	Ene-05	Sep-05	Ene-04	Ago-04	Ene-05	Sep-05
1	1	0	85.0	73.0	110.0	119.0	32	58	82	113	62.0	67.0	90.0	95.0
1	2	0	115.0	87.0	132.0	136.0	45	74	98	129	66.0	72.0	103.0	105.0
1	3	0	102.0	73.0	115.0	127.0	33	56	82	111	61.0	60.0	92.0	99.0
1	4	0	113.0	90.0	135.0	148.0	49	79	109	142	69.5	76.0	103.0	113.0
1	5	0	118.0	94.0	130.0	141.0	46	71	97	130	61.5	75.0	105.0	109.0
1	6	0	133.0	149.0	167.0	183.0	71	112	131	170	81.5	104.0	120.0	130.0
1	7	0	140.0	152.0	167.0	167.0	65	103	123	149	84.5	117.0	133.0	136.0
1	8	0	143.0	145.0	175.0	175.0	70	111	135	165	86.5	108.0	132.0	140.0
1	9	0	127.0	125.0	162.0	172.0	48	79	100	126	79.0	97.0	121.0	124.0
1	10	0	132.0	143.0	175.0	193.0	63	107	137	178	77.0	109.0	132.0	141.0
1	1	1	117.0	89.0	112.0	128.0	50	72	97	126	68.0	76.0	93.0	97.0
1	2	1	97.0	72.0	90.0	98.0	32	46	62	80	61.0	60.0	79.0	82.0
1	3	1	97.0	80.0	94.0	106.0	38	57	77	100	59.5	65.0	80.0	89.0
1	4	1	82.0	54.0	74.0	82.0	16	33	47	61	44.5	48.0	63.0	69.0
1	5	1	115.0	94.0	110.0	130.0	49	70	97	126	74.0	82.0	91.0	100.0
1	6	1	108.0	96.0	105.0	121.0	48	69	94	121	69.0	71.0	86.0	97.0
1	7	1	113.0	101.0	110.0	131.0	47	68	92	118	70.5	77.5	93.0	103.0
1	8	1	112.0	102.0	120.0	129.0	41	64	87	117	73.0	80.0	102.0	101.0
1	9	1	117.0	100.0	116.0	124.0	41	58	82	105	75.5	80.5	91.0	99.0
1	10	1	103.0	88.0	116.0	116.0	41	67	92	120	62.5	64.0	86.0	92.0
2	1	0	102.0	106.0	-	-	40	58	-	-	76.5	86.0	-	-
2	2	0	100.0	99.0	-	-	51	61	-	-	73.0	80.0	-	-
2	3	0	98.0	112.0	-	-	49	80	-	-	68.5	83.0	-	-
2	4	0	100.0	117.0	-	-	51	80	-	-	72.0	89.5	-	-
2	5	0	92.5	110.0	-	-	51	80	-	-	66.5	81.5	-	-
2	6	0	95.0	113.0	-	-	53	86	-	-	71.0	85.0	-	-
2	7	0	96.0	98.0	-	-	65	75	-	-	70.1	80.0	-	-
2	8	0	99.0	115.0	-	-	68	99	-	-	73.5	88.0	-	-
2	9	0	90.0	117.0	-	-	44	72	-	-	69.0	89.0	-	-
2	10	0	93.0	118.0	-	-	57	89	-	-	67.5	89.0	-	-
2	1	1	102.0	120.0	-	-	57	84	-	-	79.0	99.0	-	-
2	2	1	99.0	119.0	-	-	52	79	-	-	76.0	96.0	-	-
2	3	1	79.0	92.0	-	-	38	61	-	-	62.5	79.0	-	-
2	4	1	86.0	104.0	-	-	41	61	-	-	67.5	84.0	-	-
2	5	1	100.0	117.0	-	-	51	75	-	-	73.0	90.0	-	-
2	6	1	86.0	101.0	-	-	47	71	-	-	60.5	81.0	-	-
2	7	1	85.0	100.0	-	-	45	73	-	-	68.5	78.5	-	-
2	8	1	82.5	108.0	-	-	44	68	-	-	68.5	88.5	-	-
2	9	1	97.0	103.0	-	-	51	85	-	-	72.0	84.0	-	-
2	10	1	89.0	110.0	-	-	52	80	-	-	68.5	85.0	-	-

Región: 1 = Pihuamo ; 2 = Tepatitlán
 Tipo : 0 = Micropropagado ; 1 = Convencional

Tabla 2. Datos simulados para el Escenario 1 (igualdad en crecimiento en ambos tipos de planta), en los cuales se consideran dos diferentes regiones geográficas y 4 mediciones a lo largo del tiempo para cada una de las variables.

REGION	PLANTA	TIPO	ALTURA				HOJAS DESPRENDIDAS				LONGITUD DE HOJA			
			T0	T1	T2	T3	T0	T1	T2	T3	T0	T1	T2	T3
1	1	0	73.7	80.6	89.4	117.9	71	79	88	116	73.4	80.5	88.9	117.6
1	2	0	72.4	77.5	87.3	111.8	70	77	85	109	71.7	77.1	87.0	110.5
1	3	0	74.4	78.7	88.0	114.8	73	77	86	114	74.1	77.8	87.6	113.9
1	4	0	77.2	82.9	89.2	114.4	77	82	89	111	76.2	82.3	88.8	113.9
1	5	0	74.3	77.0	88.1	114.1	73	75	86	112	73.0	76.3	87.8	112.9
1	6	0	75.0	80.9	88.6	116.8	75	79	88	115	74.5	79.8	88.6	116.3
1	7	0	73.7	79.4	86.1	114.3	72	77	83	113	73.5	78.3	85.4	113.6
1	8	0	76.4	78.4	89.5	112.8	75	77	89	110	76.2	77.5	89.9	112.7
1	9	0	73.5	79.3	89.6	115.5	72	78	87	113	73.5	78.8	88.9	115.0
1	10	0	74.3	78.0	86.9	115.2	75	77	85	111	74.0	77.8	87.1	114.2
1	1	1	74.3	76.7	92.2	116.5	73	76	93	115	74.6	76.4	92.0	115.8
1	2	1	75.7	80.7	87.4	115.9	76	79	85	115	75.9	79.2	86.7	115.3
1	3	1	73.8	78.7	87.2	114.6	70	76	87	113	72.9	78.4	86.6	114.9
1	4	1	70.6	77.2	85.5	116.0	70	77	83	115	70.4	76.9	83.8	116.0
1	5	1	76.2	78.2	87.4	117.2	76	77	85	115	75.8	77.9	87.6	116.1
1	6	1	72.7	80.4	89.0	114.3	72	81	85	113	72.9	78.8	88.0	113.7
1	7	1	74.3	82.8	87.5	112.0	71	83	85	109	73.6	82.7	87.2	111.1
1	8	1	74.0	80.2	88.2	114.5	73	79	86	111	73.8	79.6	87.6	114.2
1	9	1	74.5	75.9	89.2	111.7	73	75	87	108	74.1	75.6	88.6	110.9
1	10	1	75.5	80.6	90.6	117.9	74	80	90	116	75.2	80.0	89.6	117.4
2	1	0	71.4	78.5	85.7	113.1	71	76	85	111	71.3	77.9	85.8	112.7
2	2	0	73.3	75.2	84.6	115.4	73	74	83	113	72.7	74.7	84.5	114.4
2	3	0	72.1	75.5	85.3	111.7	72	75	84	110	71.7	75.0	85.1	110.8
2	4	0	73.0	78.2	85.3	112.3	72	77	84	110	72.4	77.0	84.6	111.2
2	5	0	66.0	75.9	87.7	113.5	62	73	86	112	65.6	75.4	87.4	112.9
2	6	0	71.5	74.9	87.9	113.0	70	74	86	112	71.1	74.6	87.7	112.3
2	7	0	71.7	78.9	88.4	111.6	71	78	89	109	70.8	79.0	87.5	111.2
2	8	0	72.2	74.1	88.2	111.2	71	72	89	108	71.7	73.5	88.9	110.3
2	9	0	73.2	77.4	84.9	110.7	71	75	84	107	72.3	76.8	84.6	109.6
2	10	0	70.5	75.5	85.5	113.7	70	74	84	111	70.3	75.2	84.6	112.8
2	1	1	73.6	78.6	85.6	113.5	72	78	83	112	72.2	77.6	84.8	112.2
2	2	1	71.7	72.0	88.4	115.8	69	71	88	115	71.7	72.1	88.0	115.6
2	3	1	69.7	77.7	84.8	115.7	70	76	83	115	70.2	77.7	84.0	114.8
2	4	1	70.6	76.6	84.7	112.5	70	75	82	110	69.8	76.7	83.2	111.8
2	5	1	73.1	78.3	90.9	109.7	72	79	90	107	72.5	77.6	90.9	109.5
2	6	1	68.1	77.0	85.8	115.8	65	76	85	115	67.4	76.6	83.9	114.7
2	7	1	70.4	77.8	85.8	114.6	69	74	83	114	70.7	76.4	85.5	113.5
2	8	1	74.5	76.8	87.3	113.3	75	76	85	109	75.2	77.1	85.1	112.8
2	9	1	72.1	75.4	82.6	110.6	72	74	80	107	72.4	74.9	82.0	109.8
2	10	1	72.0	76.6	82.6	110.7	71	76	81	111	72.3	76.8	81.6	110.3

Tipo : 0 = Micropropagado ; 1 = Convencional
T0 = Tiempo 0 (inicial) ; T1 = Tiempo 1 ; T2 = Tiempo 2 ; T3 = Tiempo 3

Tabla 3. Índices morfológicos correspondientes a los datos simulados en el Escenario 1, calculados a su vez para cada una de las fechas consideradas.

REGIÓN	PLANTA	TIPO	ÍNDICE MORFOLÓGICO			
			T0	T1	T2	T3
1	1	0	125.9	138.6	153.8	202.9
1	2	0	123.6	133.7	149.7	191.3
1	3	0	127.9	134.8	151.0	197.9
1	4	0	133.0	142.7	154.1	195.9
1	5	0	127.2	131.8	151.2	195.7
1	6	0	129.6	138.4	153.1	201.0
1	7	0	126.6	135.5	146.9	196.8
1	8	0	131.4	134.5	155.0	193.7
1	9	0	126.4	136.3	153.3	198.3
1	10	0	128.9	134.4	149.5	196.5
1	1	1	128.1	132.3	160.0	200.5
1	2	1	131.4	137.9	149.6	199.9
1	3	1	125.1	134.6	150.6	197.7
1	4	1	121.8	133.4	145.7	200.3
1	5	1	131.6	134.6	150.1	201.1
1	6	1	125.6	138.7	151.3	196.9
1	7	1	126.4	143.5	149.9	191.7
1	8	1	127.5	137.9	151.2	196.1
1	9	1	127.9	130.8	152.9	190.9
1	10	1	129.7	138.9	156.0	202.8
2	1	0	123.4	134.2	148.1	194.5
2	2	0	126.4	129.3	145.6	197.9
2	3	0	124.6	130.2	146.9	192.0
2	4	0	125.5	134.1	146.6	192.6
2	5	0	111.8	129.5	150.8	195.4
2	6	0	122.7	129.0	151.0	194.7
2	7	0	123.3	136.2	152.9	191.6
2	8	0	124.1	126.8	153.6	190.2
2	9	0	125.0	132.3	146.4	189.0
2	10	0	121.7	129.7	146.7	194.9
2	1	1	125.7	135.2	146.3	195.0
2	2	1	122.6	124.2	152.6	200.0
2	3	1	121.2	133.6	145.4	199.5
2	4	1	121.5	131.8	144.3	193.0
2	5	1	125.6	135.6	156.9	188.3
2	6	1	115.8	132.6	147.0	199.5
2	7	1	121.3	131.8	146.8	197.5
2	8	1	129.7	132.7	148.6	193.5
2	9	1	125.0	129.5	141.2	189.0
2	10	1	124.3	132.4	141.6	191.7

Tipo : 0 = Micropropagado ; 1 = Convencional
T0 = Tiempo 0 (inicial) ; T1 = Tiempo 1 ; T2 = Tiempo 2 ; T3 = Tiempo 3

Tabla 4. Resultados obtenidos de la clasificación de los datos en el Escenario 1, utilizando un análisis discriminante empleando la función lineal de Fisher.

REGIÓN	PLANTA	TIPO	TIPO CLASIFICADO			
		ORIGINAL	T0	T1	T2	T3
1	1	M	C	M	C	C
1	2	M	C	C	M	M
1	3	M	C	M	M	C
1	4	M	M	C	M	M
1	5	M	M	M	M	M
1	6	M	M	M	M	C
1	7	M	C	M	C	C
1	8	M	C	M	M	M
1	9	M	C	M	C	C
1	10	M	M	C	M	M
1	1	C	C	C	M	C
1	2	C	M	M	C	C
1	3	C	C	M	C	C
1	4	C	C	C	C	C
1	5	C	M	M	M	M
1	6	C	C	C	C	C
1	7	C	C	C	M	M
1	8	C	C	M	C	M
1	9	C	C	C	C	M
1	10	C	C	C	C	C
2	1	M	C	M	M	M
2	2	M	M	M	M	C
2	3	M	M	C	M	C
2	4	M	M	M	C	M
2	5	M	M	M	M	C
2	6	M	M	C	M	C
2	7	M	M	C	M	M
2	8	M	M	M	M	M
2	9	M	M	M	M	M
2	10	M	C	M	C	M
2	1	C	M	C	C	C
2	2	C	C	C	M	C
2	3	C	C	C	C	C
2	4	C	M	C	C	M
2	5	C	M	C	M	M
2	6	C	M	C	C	C
2	7	C	C	M	C	C
2	8	C	C	C	C	M
2	9	C	C	M	C	M
2	10	C	C	C	C	C

Tipo : M = Micropropagado ; C = Convencional
T0 = Tiempo 0 (inicial) ; T1 = Tiempo 1 ; T2 = Tiempo 2 ; T3 = Tiempo 3

Tabla 5. Tasas de crecimiento temporales de los índices morfológicos correspondientes al Escenario 1.

REGIÓN	PLANTA	TIPO	TASAS ÍNDICE MORFOLOGICO		
			Tasa 1	Tasa 2	Tasa 3
1	1	0	1.27	1.52	4.91
1	2	0	1.01	1.60	4.16
1	3	0	0.69	1.62	4.69
1	4	0	0.97	1.14	4.18
1	5	0	0.46	1.94	4.45
1	6	0	0.88	1.47	4.79
1	7	0	0.89	1.14	4.99
1	8	0	0.31	2.05	3.87
1	9	0	0.99	1.70	4.50
1	10	0	0.55	1.51	4.70
1	1	1	0.42	2.77	4.05
1	2	1	0.65	1.17	5.03
1	3	1	0.95	1.60	4.71
1	4	1	1.16	1.23	5.46
1	5	1	0.30	1.55	5.10
1	6	1	1.31	1.26	4.56
1	7	1	1.71	0.64	4.18
1	8	1	1.04	1.33	4.49
1	9	1	0.29	2.21	3.80
1	10	1	0.92	1.71	4.68
2	1	0	1.08	1.39	4.64
2	2	0	0.29	1.63	5.23
2	3	0	0.56	1.67	4.51
2	4	0	0.86	1.25	4.60
2	5	0	1.77	2.13	4.46
2	6	0	0.63	2.20	4.37
2	7	0	1.29	1.67	3.87
2	8	0	0.27	2.68	3.66
2	9	0	0.73	1.41	4.26
2	10	0	0.80	1.70	4.82
2	1	1	0.95	1.11	4.87
2	2	1	0.16	2.84	4.74
2	3	1	1.24	1.18	5.41
2	4	1	1.03	1.25	4.87
2	5	1	1.00	2.13	3.14
2	6	1	1.68	1.44	5.25
2	7	1	1.05	1.50	5.07
2	8	1	0.30	1.59	4.49
2	9	1	0.45	1.17	4.78
2	10	1	0.81	0.92	5.01

Tipo : 0 = Micropropagado ; 1 = Convencional

Tasa : 1 = Tasas calculadas para el periodo T0 – T1.
 2 = Tasas calculadas para el periodo T1 – T2.
 3 = Tasas calculadas para el periodo T2 – T3.

Tabla 6. Tasas de crecimiento relativas promedio de los índices morfológicos correspondientes al Escenario 1.

REGIÓN	PLANTA	TIPO	TASAS ÍNDICE MORFOLOGICO		
			Tasa 1	Tasa 2	Tasa 3
1	1	0	0.10	0.11	0.32
1	2	0	0.08	0.12	0.28
1	3	0	0.05	0.12	0.31
1	4	0	0.07	0.08	0.27
1	5	0	0.04	0.15	0.29
1	6	0	0.07	0.11	0.31
1	7	0	0.07	0.08	0.34
1	8	0	0.02	0.15	0.25
1	9	0	0.08	0.12	0.29
1	10	0	0.04	0.11	0.31
1	1	1	0.03	0.21	0.25
1	2	1	0.05	0.08	0.34
1	3	1	0.08	0.12	0.31
1	4	1	0.10	0.09	0.37
1	5	1	0.02	0.12	0.34
1	6	1	0.10	0.09	0.30
1	7	1	0.14	0.04	0.28
1	8	1	0.08	0.10	0.30
1	9	1	0.02	0.17	0.25
1	10	1	0.07	0.12	0.30
2	1	0	0.09	0.10	0.31
2	2	0	0.02	0.13	0.36
2	3	0	0.04	0.13	0.31
2	4	0	0.07	0.09	0.31
2	5	0	0.16	0.16	0.30
2	6	0	0.05	0.17	0.29
2	7	0	0.10	0.12	0.25
2	8	0	0.02	0.21	0.24
2	9	0	0.06	0.11	0.29
2	10	0	0.07	0.13	0.33
2	1	1	0.08	0.08	0.33
2	2	1	0.01	0.23	0.31
2	3	1	0.10	0.09	0.37
2	4	1	0.08	0.09	0.34
2	5	1	0.08	0.16	0.20
2	6	1	0.15	0.11	0.36
2	7	1	0.09	0.11	0.35
2	8	1	0.02	0.12	0.30
2	9	1	0.04	0.09	0.34
2	10	1	0.07	0.07	0.35

Tipo : 0 = Micropropagado ; 1 = Convencional

Tasa : 1 = Tasas calculadas para el periodo T0 – T1.
2 = Tasas calculadas para el periodo T1 – T2.

3 = Tasas calculadas para el periodo T2 – T3.

Tabla 7. Datos simulados para el Escenario 2 (diferente crecimiento en ambos tipos de planta), en los cuales se consideran también dos diferentes regiones geográficas y 4 mediciones a lo largo del tiempo para cada una de las variables.

REGION	PLANTA	TIPO	ALTURA				HOJAS DESPRENDIDAS				LONGITUD DE HOJA			
			T0	T1	T2	T3	T0	T1	T2	T3	T0	T1	T2	T3
1	1	0	74.3	79.4	92.1	129.8	73	78	91	127	73.7	78.4	91.4	129.3
1	2	0	72.4	80.8	89.9	129.8	70	81	88	129	71.6	81.2	90.4	129.0
1	3	0	74.6	78.6	90.4	128.0	74	77	88	126	74.6	78.5	90.1	127.6
1	4	0	73.6	80.6	93.1	129.5	72	78	91	127	72.7	80.1	92.4	128.3
1	5	0	75.2	79.7	90.8	129.6	74	79	89	128	74.5	79.1	90.4	129.3
1	6	0	73.1	83.6	91.0	131.1	73	83	89	129	72.4	82.5	89.6	130.4
1	7	0	75.6	80.1	92.5	128.2	75	78	90	126	74.6	79.0	91.8	127.6
1	8	0	73.8	79.2	92.9	125.9	72	76	93	123	73.5	78.7	92.4	125.6
1	9	0	74.7	79.1	92.3	127.2	73	77	91	124	74.1	78.5	92.1	126.2
1	10	0	73.8	79.2	87.9	127.4	73	79	86	125	73.1	79.2	87.5	127.5
1	1	1	72.7	80.3	91.6	114.5	71	77	91	112	72.8	79.7	90.8	113.3
1	2	1	73.2	78.9	88.8	118.4	72	76	86	118	72.1	79.3	89.2	117.6
1	3	1	74.0	78.6	88.6	117.2	73	78	86	115	74.3	78.6	88.0	115.0
1	4	1	71.2	78.4	89.4	117.5	68	77	86	117	70.6	78.7	89.7	117.1
1	5	1	71.7	82.7	87.7	114.2	68	83	85	114	71.0	82.2	87.8	114.4
1	6	1	75.4	80.9	87.4	116.0	74	79	85	115	75.6	79.9	86.8	115.2
1	7	1	74.9	80.8	89.3	115.5	73	78	86	114	74.4	79.9	88.1	114.9
1	8	1	75.2	78.9	88.8	119.1	75	79	86	119	74.5	79.3	88.4	118.5
1	9	1	73.4	77.6	88.3	113.4	71	76	86	110	72.7	76.8	87.1	112.7
1	10	1	75.7	83.0	86.8	117.5	74	83	86	115	74.4	81.6	86.7	116.3
2	1	0	71.4	76.2	88.3	123.3	70	74	85	120	71.1	74.9	87.0	123.0
2	2	0	72.2	79.6	89.0	127.0	71	79	87	125	71.2	79.7	88.5	125.6
2	3	0	72.7	76.4	87.3	126.1	71	75	87	124	72.2	76.0	87.2	125.6
2	4	0	70.5	80.0	88.9	125.5	70	81	87	124	69.6	79.6	87.7	124.7
2	5	0	72.0	80.0	88.2	125.8	71	81	87	124	70.6	79.2	87.8	125.2
2	6	0	71.1	78.7	88.2	126.9	70	78	87	124	70.6	79.4	87.2	125.7
2	7	0	71.8	74.4	87.3	123.6	71	73	86	121	71.1	73.8	86.8	123.4
2	8	0	72.2	78.7	87.1	127.1	72	78	86	125	72.3	77.6	86.5	126.4
2	9	0	69.5	76.7	92.3	123.9	68	77	91	123	69.0	76.3	91.6	123.3
2	10	0	71.4	75.1	88.7	126.0	69	73	86	123	70.4	74.9	88.0	124.6
2	1	1	70.5	77.2	82.5	114.5	70	76	80	114	70.8	75.6	83.3	114.2
2	2	1	72.9	78.1	84.8	114.5	72	78	85	114	71.5	76.8	84.4	114.4
2	3	1	64.8	75.4	85.4	112.3	61	73	85	109	64.4	75.4	84.8	111.9
2	4	1	69.7	76.6	84.5	109.0	69	75	81	105	68.6	76.9	83.4	107.8
2	5	1	71.1	76.1	86.8	108.8	70	75	86	105	71.0	76.1	86.3	108.3
2	6	1	71.6	75.6	89.1	115.6	73	74	87	115	72.2	74.8	88.4	114.9
2	7	1	70.7	74.6	84.3	116.2	68	74	81	115	69.2	75.2	84.1	115.4
2	8	1	68.8	78.1	85.9	111.0	66	76	83	109	68.9	77.0	84.7	110.5
2	9	1	73.7	78.3	84.7	110.0	72	77	85	106	73.2	79.5	83.8	108.6
2	10	1	70.9	78.7	83.4	114.8	71	79	82	114	70.1	78.2	83.6	113.7

Tipo : 0 = Micropropagado ; 1 = Convencional
T0 = Tiempo 0 (inicial) ; T1 = Tiempo 1 ; T2 = Tiempo 2 ; T3 = Tiempo 3

Tabla 8. Índices morfológicos correspondientes a los datos simulados en el Escenario 2, calculados a su vez para cada una de las fechas consideradas.

REGIÓN	PLANTA	TIPO	ÍNDICE MORFOLÓGICO			
			T0	T1	T2	T3
1	1	0	127.6	136.1	158.5	222.9
1	2	0	123.6	140.3	154.9	223.9
1	3	0	128.9	135.2	155.0	220.3
1	4	0	126.0	137.8	159.6	222.2
1	5	0	129.2	137.3	156.0	223.4
1	6	0	126.1	143.8	155.7	225.5
1	7	0	130.0	136.9	158.4	220.4
1	8	0	126.6	135.1	160.7	216.2
1	9	0	128.1	135.5	159.0	217.9
1	10	0	127.0	137.1	150.9	219.3
1	1	1	125.0	136.8	157.8	196.2
1	2	1	125.5	135.2	152.4	204.4
1	3	1	127.8	135.8	151.6	200.5
1	4	1	121.1	135.2	153.1	203.0
1	5	1	121.7	143.1	150.4	197.8
1	6	1	129.9	138.5	149.7	199.9
1	7	1	128.4	137.8	152.1	198.8
1	8	1	129.7	136.9	152.0	205.9
1	9	1	125.4	133.0	150.9	194.1
1	10	1	129.4	142.9	149.8	201.4
2	1	0	122.7	130.0	150.3	211.5
2	2	0	123.8	137.6	152.7	218.0
2	3	0	124.7	131.3	151.0	216.9
2	4	0	121.3	138.9	152.2	216.0
2	5	0	123.3	138.7	151.8	216.5
2	6	0	122.2	136.3	151.5	217.4
2	7	0	123.5	127.7	150.2	212.5
2	8	0	125.0	135.3	149.9	218.5
2	9	0	119.2	132.8	158.7	213.7
2	10	0	121.7	128.8	151.7	215.7
2	1	1	122.0	132.1	141.9	197.9
2	2	1	124.9	134.5	146.8	198.0
2	3	1	109.8	129.2	147.3	192.4
2	4	1	119.7	131.9	143.7	185.8
2	5	1	122.5	131.2	149.6	186.0
2	6	1	125.2	129.6	152.7	199.5
2	7	1	120.0	129.2	144.0	200.1
2	8	1	117.6	133.4	146.4	190.8
2	9	1	126.4	135.6	146.4	187.4
2	10	1	122.4	136.2	143.8	197.7

Tipo : 0 = Micropropagado ; 1 = Convencional
T0 = Tiempo 0 (inicial) ; T1 = Tiempo 1 ; T2 = Tiempo 2 ; T3 = Tiempo 3

Tabla 9. Resultados obtenidos de la clasificación de los datos en el Escenario 2, utilizando un análisis discriminante empleando la función lineal de Fisher.

Región	Planta	Tipo Original	Tipo Clasificado			
			T0	T1	T2	T3
1	1	M	M	M	M	M
1	2	M	C	C	C	M
1	3	M	M	C	C	M
1	4	M	M	M	M	M
1	5	M	M	M	M	M
1	6	M	M	M	M	M
1	7	M	M	M	M	M
1	8	M	C	M	M	M
1	9	M	C	M	M	M
1	10	M	M	C	C	M
1	1	C	C	M	M	C
1	2	C	M	C	C	C
1	3	C	C	C	C	C
1	4	C	C	C	C	C
1	5	C	C	M	C	C
1	6	C	C	M	C	C
1	7	C	C	M	C	C
1	8	C	M	C	C	C
1	9	C	C	M	C	C
1	10	C	M	M	C	C
2	1	M	C	C	M	M
2	2	M	M	M	M	M
2	3	M	M	C	M	M
2	4	M	M	M	M	M
2	5	M	M	M	M	M
2	6	M	M	C	M	M
2	7	M	M	C	M	M
2	8	M	M	M	M	M
2	9	M	C	M	M	M
2	10	M	M	C	M	M
2	1	C	C	M	C	C
2	2	C	M	M	C	C
2	3	C	C	C	C	C
2	4	C	M	C	C	C
2	5	C	C	C	M	C
2	6	C	M	C	M	C
2	7	C	M	C	C	C
2	8	C	C	C	C	C
2	9	C	M	C	C	C
2	10	C	M	M	C	C

Tipo : M = Micropropagado ; C = Convencional
T0 = Tiempo 0 (inicial) ; T1 = Tiempo 1 ; T2 = Tiempo 2 ; T3 = Tiempo 3

Tabla 10. Tasas de crecimiento temporales de los índices morfológicos correspondientes al Escenario 2.

PLANTA	REGIÓN	TIPO	TASAS ÍNDICE MORFOLOGICO		
			Tasa 1	Tasa 2	Tasa 3
1	1	0	0.85	2.24	6.44
2	1	0	1.67	1.46	6.90
3	1	0	0.63	1.98	6.53
4	1	0	1.18	2.18	6.26
5	1	0	0.81	1.87	6.74
6	1	0	1.77	1.19	6.98
7	1	0	0.69	2.15	6.20
8	1	0	0.85	2.56	5.55
9	1	0	0.74	2.35	5.89
10	1	0	1.01	1.38	6.84
1	1	1	1.18	2.10	3.84
2	1	1	0.97	1.72	5.20
3	1	1	0.80	1.58	4.89
4	1	1	1.41	1.79	4.99
5	1	1	2.14	0.73	4.74
6	1	1	0.86	1.12	5.02
7	1	1	0.94	1.43	4.67
8	1	1	0.72	1.51	5.39
9	1	1	0.76	1.79	4.32
10	1	1	1.35	0.69	5.16
1	2	0	0.73	2.03	6.12
2	2	0	1.38	1.51	6.53
3	2	0	0.66	1.97	6.59
4	2	0	1.76	1.33	6.38
5	2	0	1.54	1.31	6.47
6	2	0	1.41	1.52	6.59
7	2	0	0.42	2.25	6.23
8	2	0	1.03	1.46	6.86
9	2	0	1.36	2.59	5.50
10	2	0	0.71	2.29	6.40
1	2	1	1.01	0.98	5.60
2	2	1	0.96	1.23	5.12
3	2	1	1.94	1.81	4.51
4	2	1	1.22	1.18	4.21
5	2	1	0.87	1.84	3.64
6	2	1	0.44	2.31	4.68
7	2	1	0.92	1.48	5.61
8	2	1	1.58	1.30	4.44
9	2	1	0.92	1.08	4.10
10	2	1	1.38	0.76	5.39

Tipo : 0 = Micropropagado ; 1 = Convencional

Tasa : 1 = Tasas calculadas para el periodo T0 – T1.
 2 = Tasas calculadas para el periodo T1 – T2.
 3 = Tasas calculadas para el periodo T2 – T3.

Tabla 11. Tasas de crecimiento relativas promedio de los índices morfológicos correspondientes al Escenario 2.

PLANTA	REGIÓN	TIPO	ÍNDICE MORFOLOGICO		
			Tasa 1	Tasa 2	Tasa 3
1	1	0	0.07	0.16	0.41
2	1	0	0.14	0.10	0.45
3	1	0	0.05	0.15	0.42
4	1	0	0.09	0.16	0.39
5	1	0	0.06	0.14	0.43
6	1	0	0.14	0.08	0.45
7	1	0	0.05	0.16	0.39
8	1	0	0.07	0.19	0.35
9	1	0	0.06	0.17	0.37
10	1	0	0.08	0.10	0.45
1	1	1	0.09	0.15	0.24
2	1	1	0.08	0.13	0.34
3	1	1	0.06	0.12	0.32
4	1	1	0.12	0.13	0.33
5	1	1	0.18	0.05	0.32
6	1	1	0.07	0.08	0.34
7	1	1	0.07	0.10	0.31
8	1	1	0.06	0.11	0.35
9	1	1	0.06	0.13	0.29
10	1	1	0.10	0.05	0.34
1	2	0	0.06	0.16	0.41
2	2	0	0.11	0.11	0.43
3	2	0	0.05	0.15	0.44
4	2	0	0.15	0.10	0.42
5	2	0	0.12	0.09	0.43
6	2	0	0.12	0.11	0.43
7	2	0	0.03	0.18	0.41
8	2	0	0.08	0.11	0.46
9	2	0	0.11	0.20	0.35
10	2	0	0.06	0.18	0.42
1	2	1	0.08	0.07	0.39
2	2	1	0.08	0.09	0.35
3	2	1	0.18	0.14	0.31
4	2	1	0.10	0.09	0.29
5	2	1	0.07	0.14	0.24
6	2	1	0.04	0.18	0.31
7	2	1	0.08	0.11	0.39
8	2	1	0.13	0.10	0.30
9	2	1	0.07	0.08	0.28
10	2	1	0.11	0.06	0.37

Tipo : 0 = Micropropagado ; 1 = Convencional

Tasa : 1 = Tasas calculadas para el periodo T0 – T1.
2 = Tasas calculadas para el periodo T1 – T2.

3 = Tasas calculadas para el periodo T2 – T3.

A.2 Gráficas adicionales.

Figura 1. Gráficas del cambio ocurrido en las variables Altura, Hojas desprendidas y Longitud de hoja del mes de Agosto (2004) a Enero (2005), y Enero (2005) a Septiembre (2005) para las plantas de agave de la región de Pihuamo.

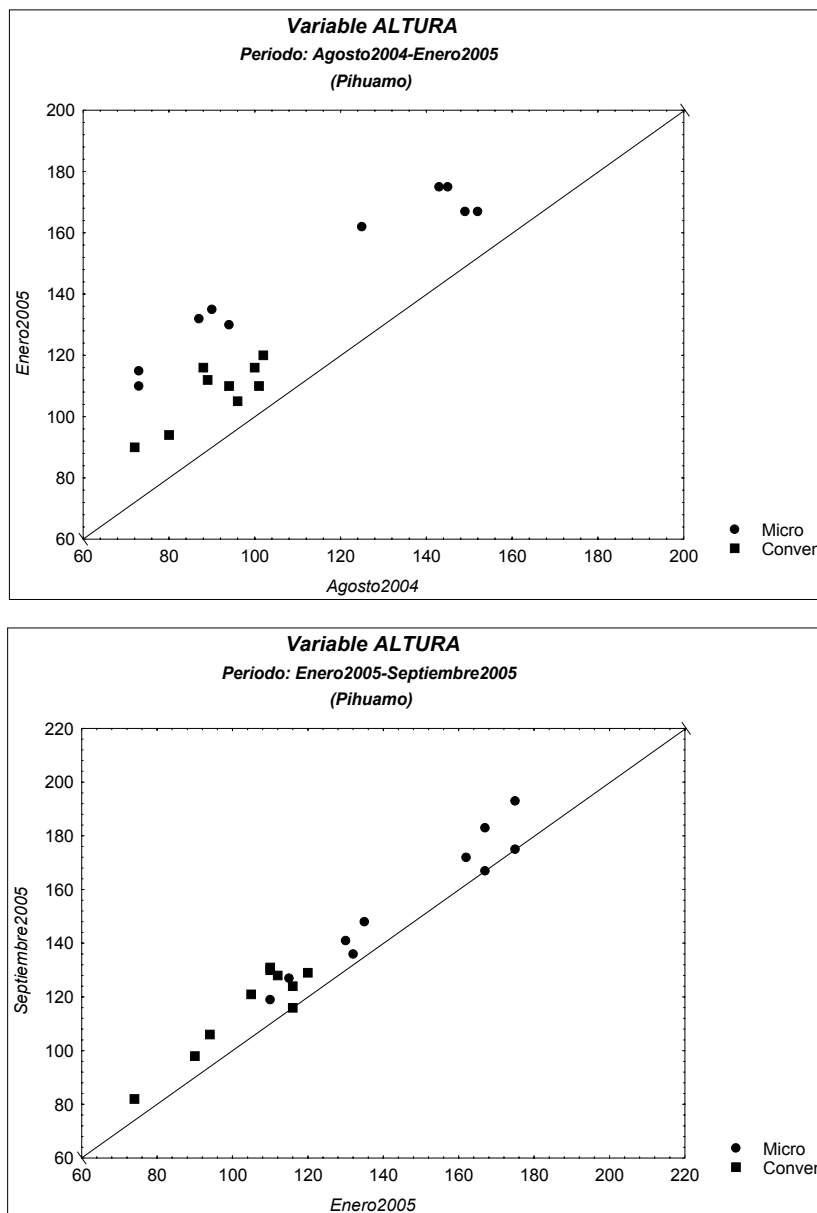


Figura 1. (continuación)...

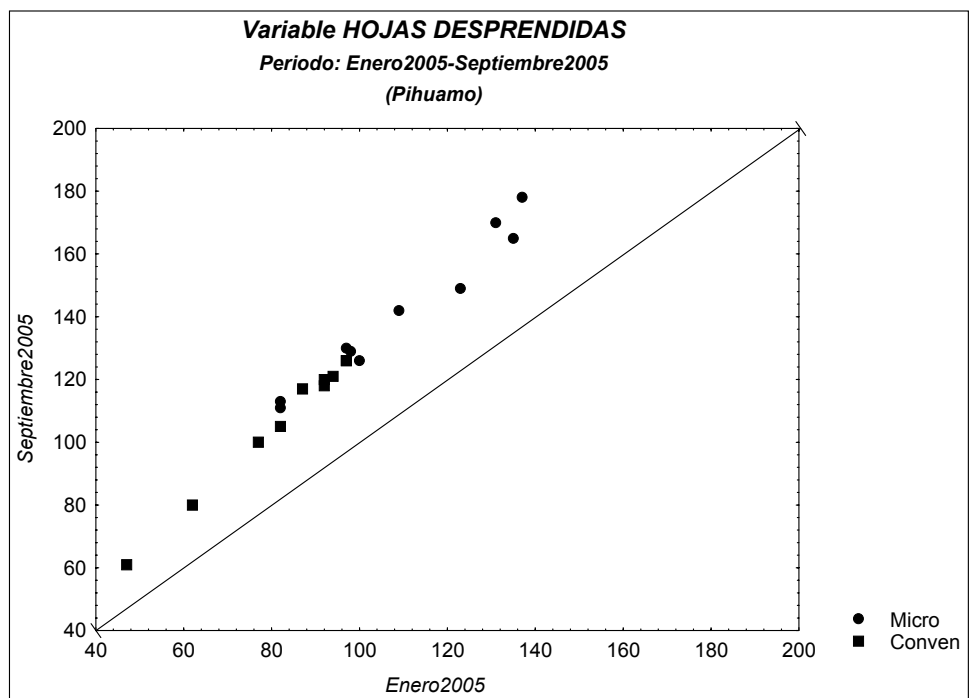
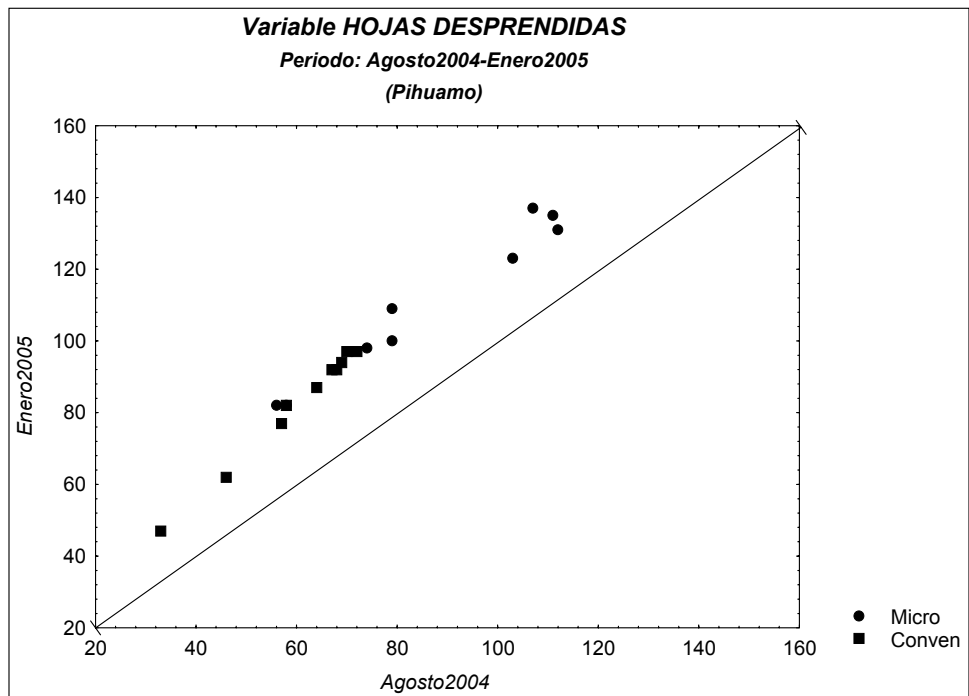


Figura 1. (continuación)...

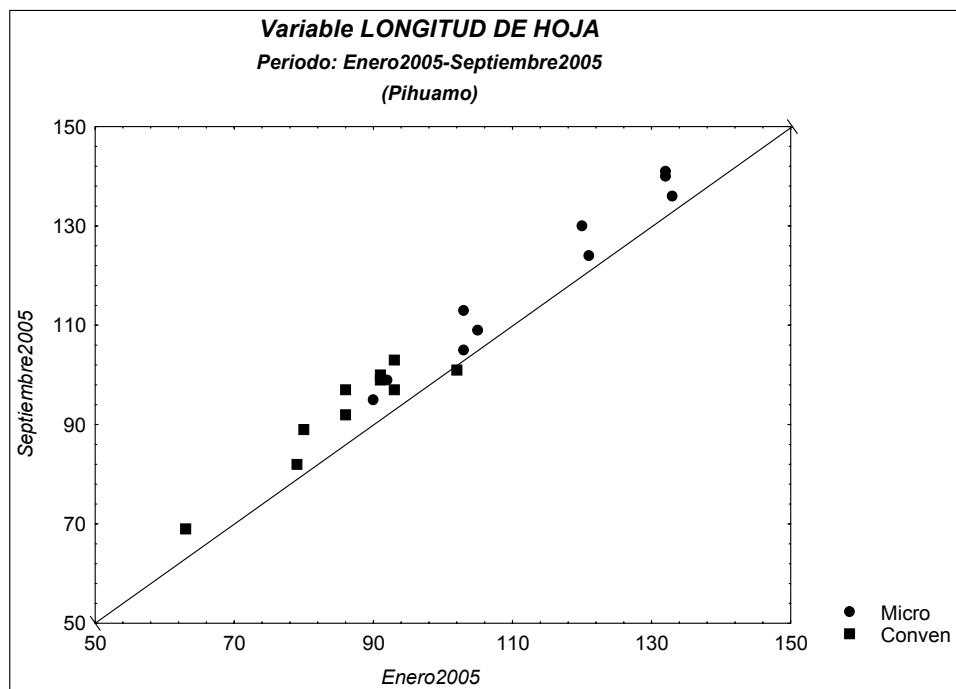
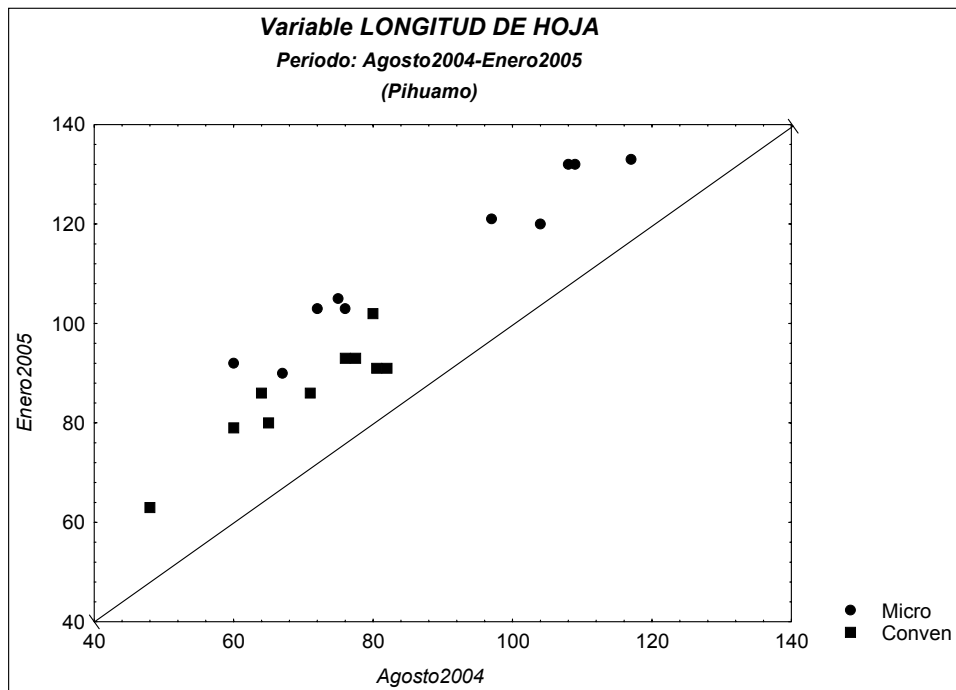


Figura 2. Residuales obtenidos del ajuste realizado con los índices morfológicos del Escenario 1 (Región 1). Micropropagado (arriba), Convencional (abajo).

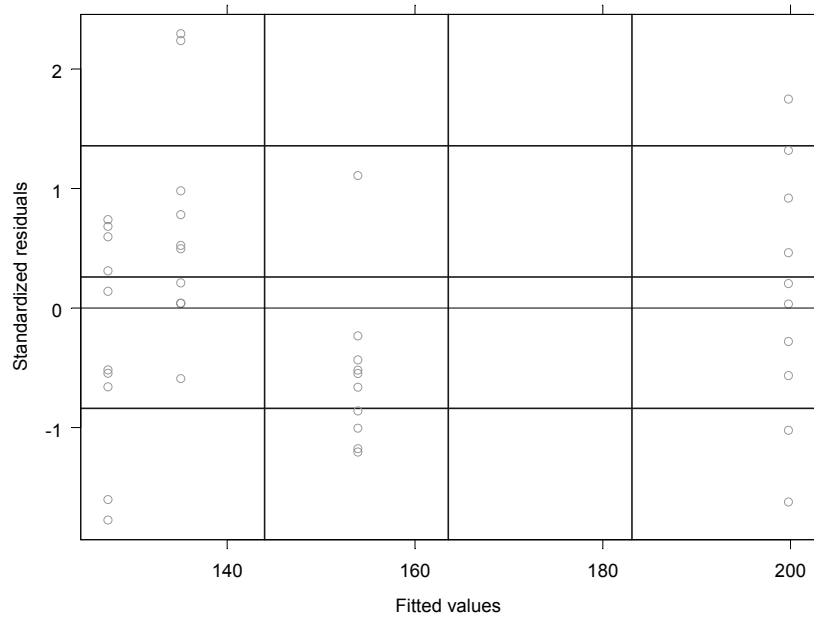
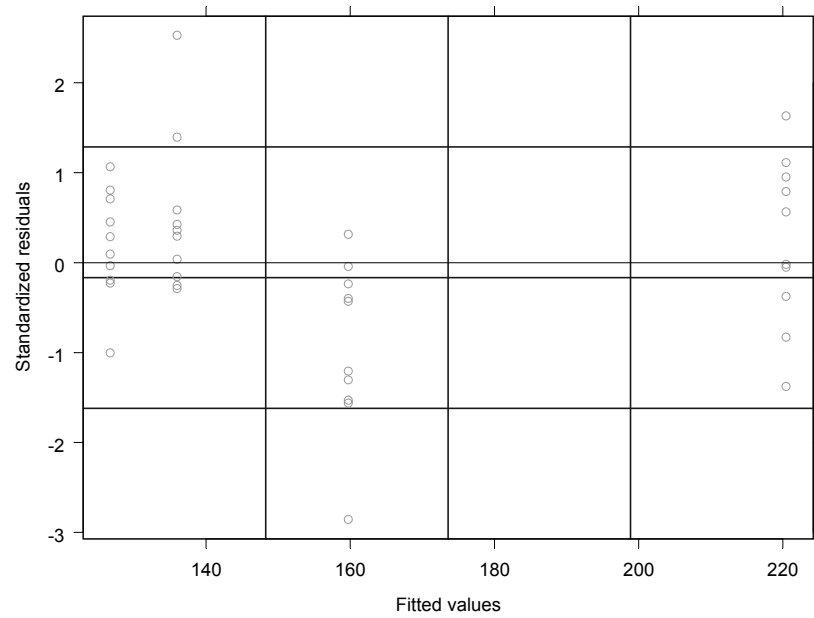


Figura 3. Residuales obtenidos del ajuste realizado con los índices morfológicos del Escenario 1 (Región 2). Micropropagado (arriba), Convencional (abajo).

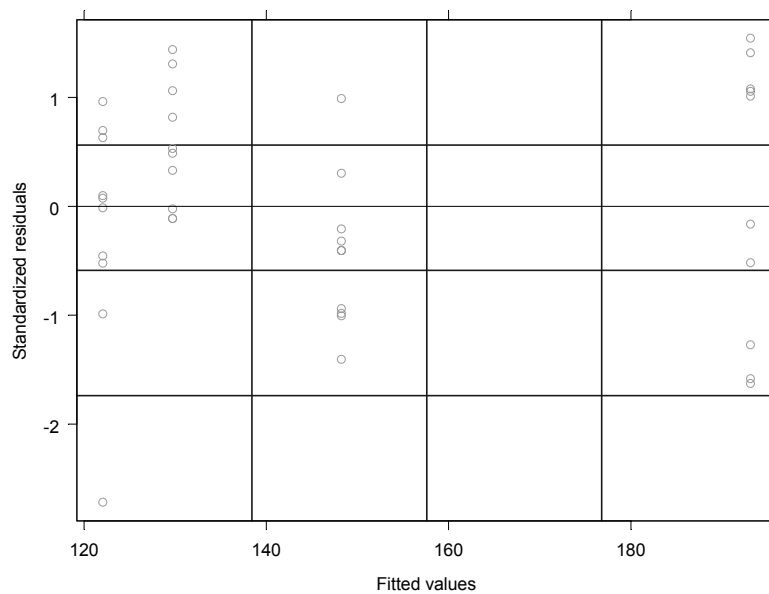
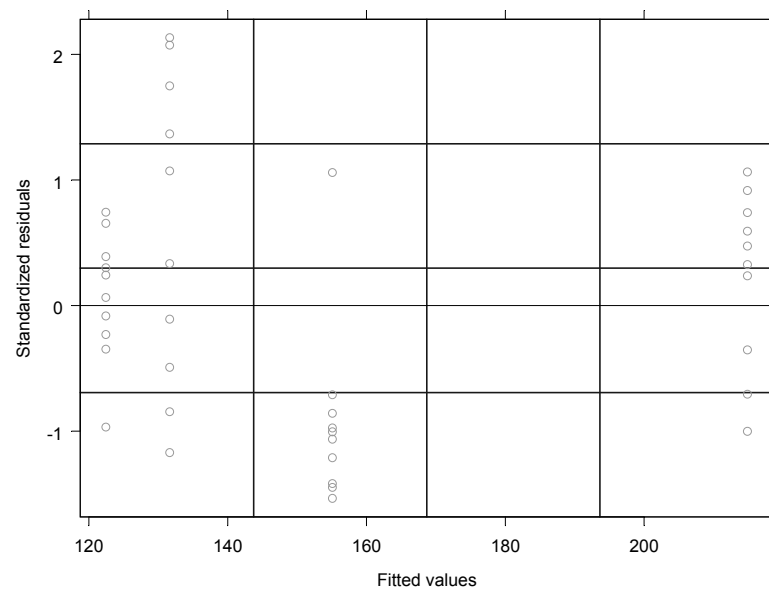


Figura 4. Gráficas cuantil-cuantil normales de los residuales del ajuste realizado en el Escenario 1 (Región 1). Micropropagado (arriba), Convencional (abajo).

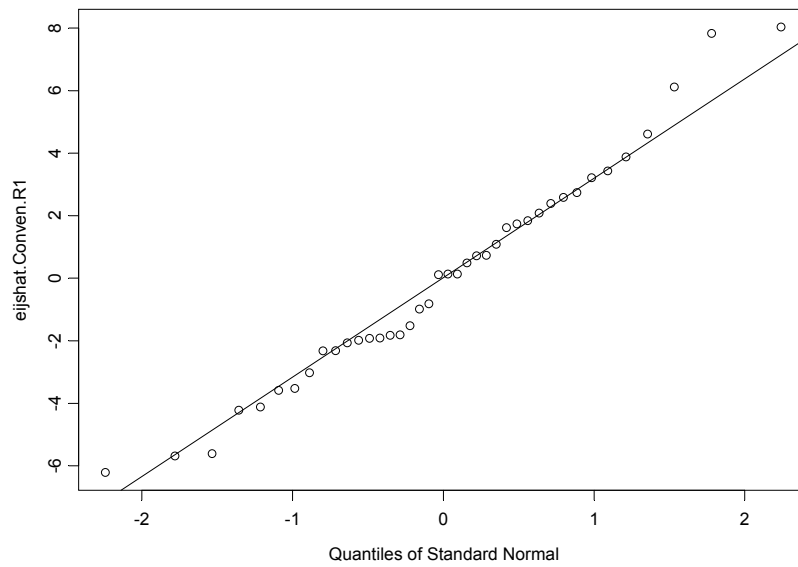
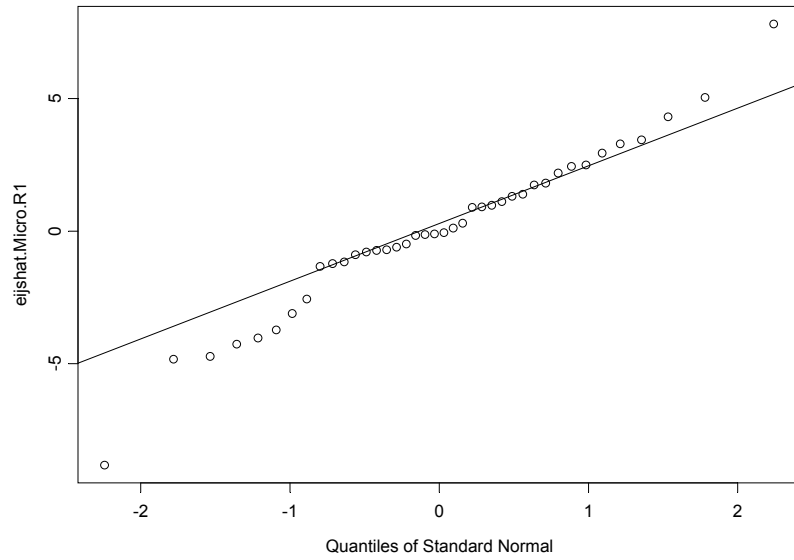


Figura 5. Gráficas cuantil-cuantil normales de los residuales del ajuste realizado en el Escenario 1 (Región 2). Micropropagado (arriba), Convencional (abajo).

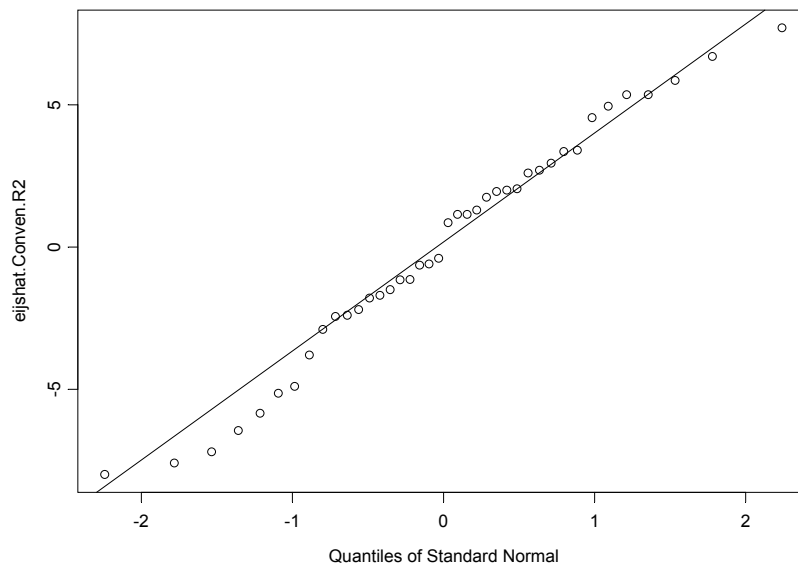
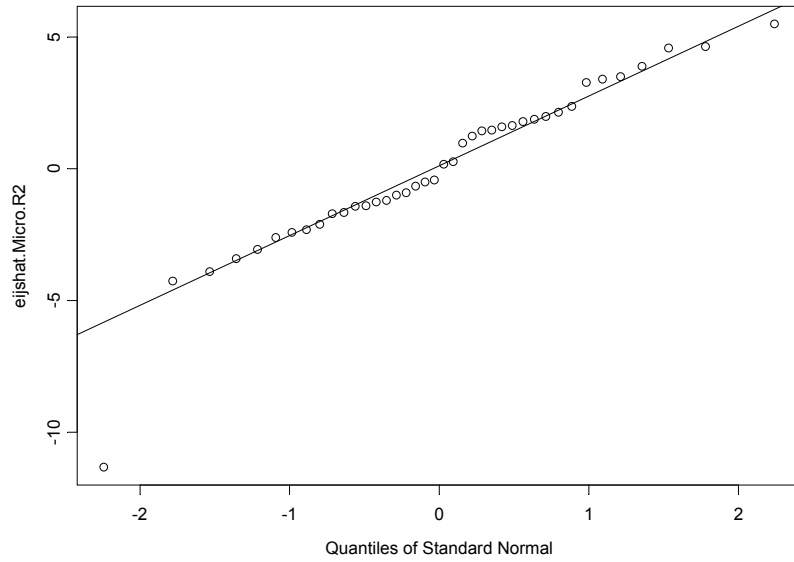


Figura 6. Residuales obtenidos del ajuste realizado con los índices morfológicos del Escenario 2 (Región 1). Micropropagado (arriba), Convencional (abajo).

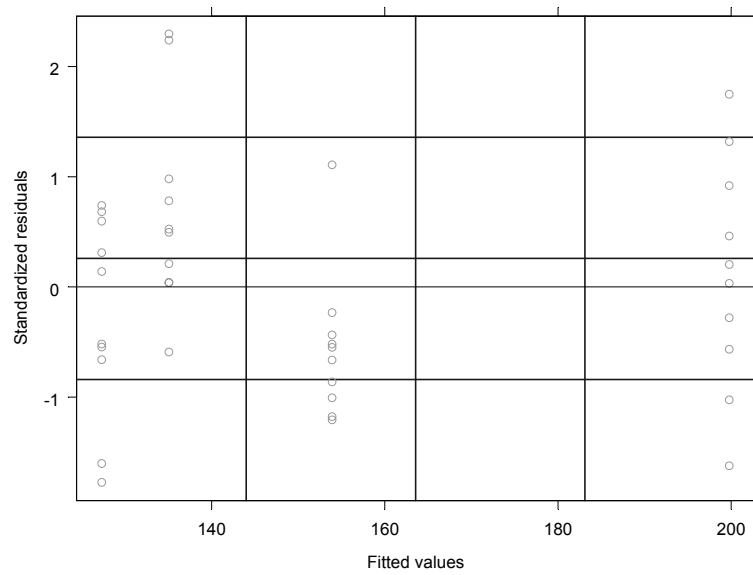
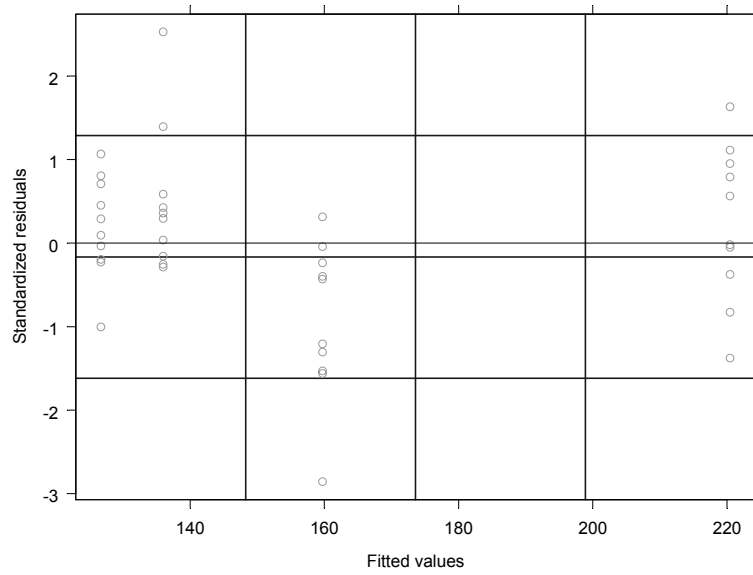


Figura 7. Residuales obtenidos del ajuste realizado con los índices morfológicos del Escenario 2 (Región 2). Micropropagado (arriba), Convencional (abajo).

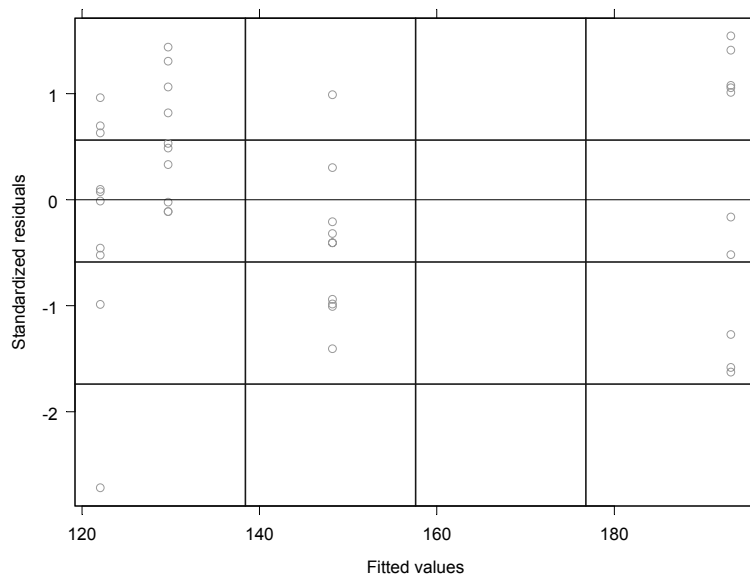
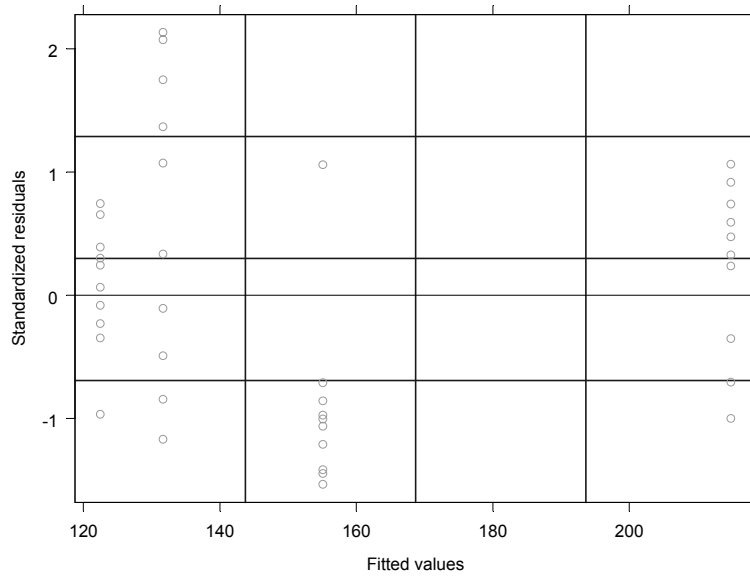


Figura 8. Gráficas cuantil-cuantil normales de los residuales del ajuste realizado en el Escenario 2 (Región 1). Micropropagado (arriba), Convencional (abajo).

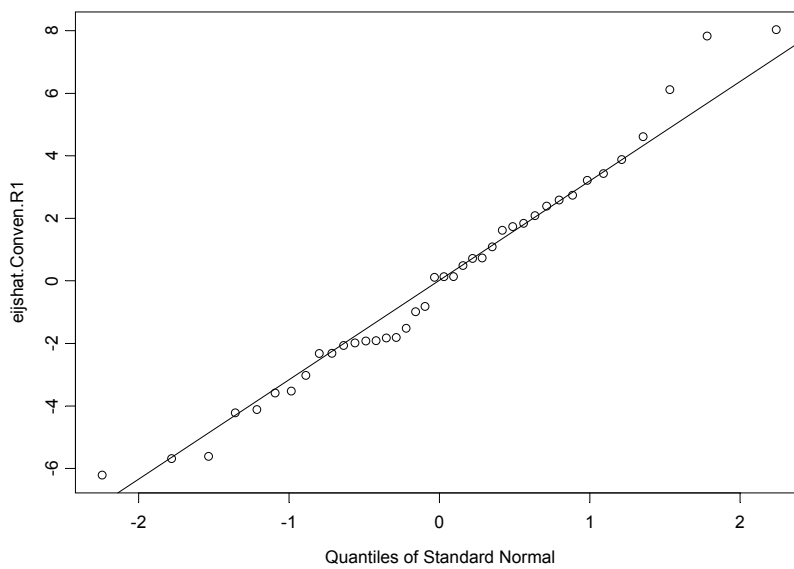
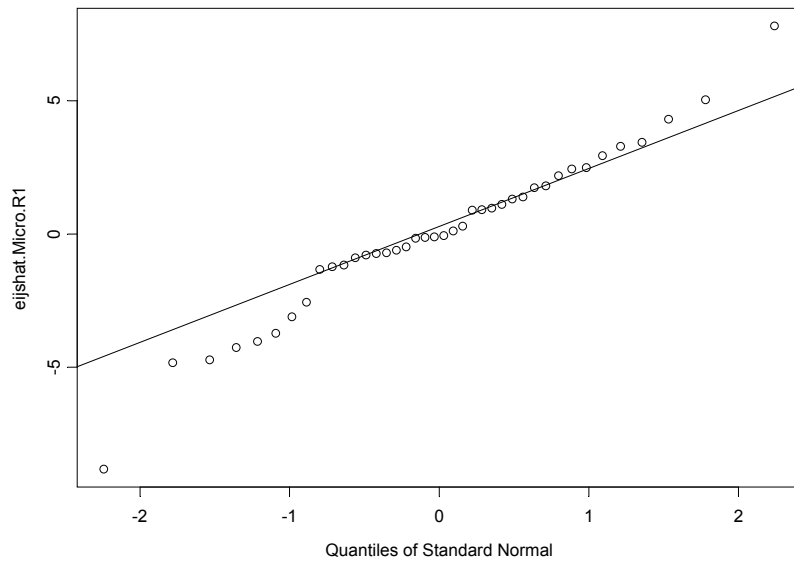
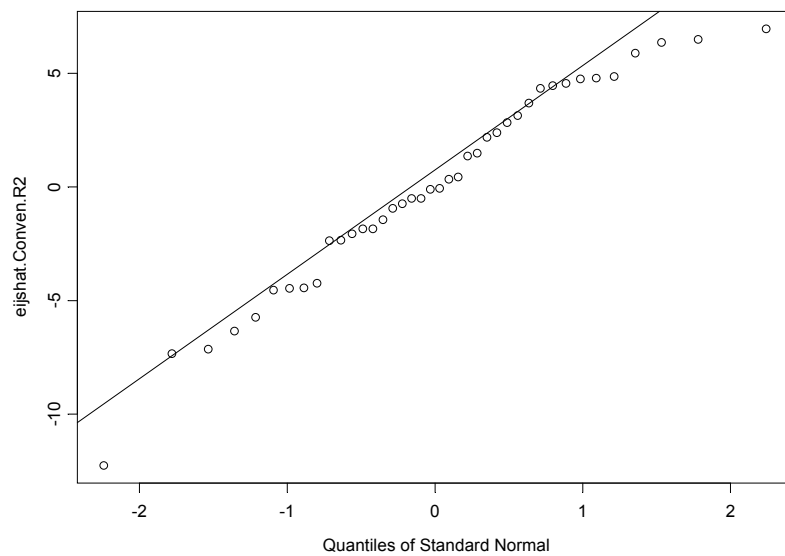
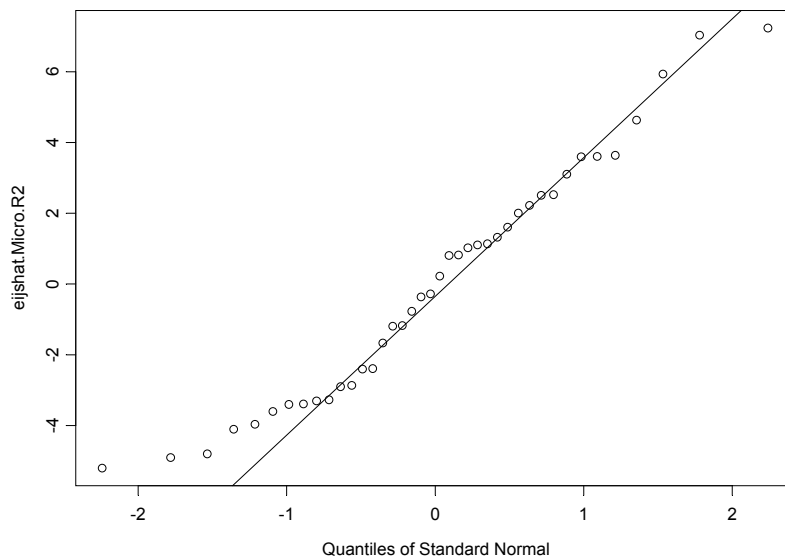


Figura 9. Gráficas cuantil-cuantil normales de los residuales del ajuste realizado en el Escenario 2 (Región 2). Micropropagado (arriba), Convencional (abajo).



A.3 Programas.

```
#####  
# Ajuste de Modelo No-Lineal de datos simulados #  
# Escenario 1 #  
# Autor: Esteban Bracamontes Orozco #  
# Ultima modificacion: 01/Marzo/2006 #  
#####  
  
#####  
### Bajando datos desde el archivo ###  
#####  
  
Datos <-data.sheet(IndicesOrderedEsc1)  
attach(Datos)  
  
DataIM <-data.frame(Region=Region,Tipo=Tipo,ID=ID,t=Time,IM=IM)  
attach(DataIM)  
  
### Agrupando datos ###  
DataIMGrouped <-groupedData( IM ~ t | ID, data=as.data.frame(DataIM), outer=~Tipo,  
  labels=list(x="t",y="IM") )  
  
#####  
### AJUSTES: MODELO EXPONENCIAL ###  
#####  
  
###-----###  
### REGION 1 ###  
###-----###  
  
#-----#  
# Micro #  
#-----#  
  
### Ajuste ###  
DataIMGrouped.nls.Micro.R1 <-nls( IM~paramC+exp(paramK*t), data=DataIMGrouped[1:40,],  
  start=c(paramC=131.2,paramK=0.13) )  
summary(DataIMGrouped.nls.Micro.R1)  
  
### Impresion de las curvas ###  
vectime.R1 <-10:60  
c.Micro.R1 <-122.467  
Sd.c.Micro.R1 <-0.658186  
k.Micro.R1 <-0.0878741  
Sd.k.Micro.R1 <-0.000342363  
Cov.c.k.Micro.R1 <-(-0.69)*Sd.c.Micro.R1*Sd.k.Micro.R1  
Sd.Error.Micro.R1 <-3.01368  
  
Curva.Micro.R1 <- c.Micro.R1 + exp( k.Micro.R1 * vectime.R1 )  
  
plot(DataIMGrouped[1:40,4],DataIMGrouped[1:40,5],type="p",  
  xlab="t (meses)", xaxt="n", ylab="Indice Morfologico (IM)", ylim=c(100,300),  
  main="Ajuste de Curva de Crecimiento\n(Micropropagado - Region 1)")  
axis(side=1,at=c(19,29,39,49),labels=c("19","29","39","49"))  
lines(vectime.R1,Curva.Micro.R1,type="l")  
legend(locator(1),c("CurvaMicro: IM=122.467+exp(0.08787*t)"),lty=c(1))  
  
### Analisis de Residuales ###  
plot(DataIMGrouped.nls.Micro.R1,main="Micropropagado - Region 1")  
plot(DataIMGrouped.nls.Micro.R1,ID~resid(.),abline=0)
```

```

tjs.Micro.R1      <-as.vector(rep(c(19,29,39,49),10))
yjshat.Micro.R1  <-as.vector(c.Micro.R1 + exp(k.Micro.R1*tjs.Micro.R1))
yjs.Micro.R1     <-as.vector(DataIMGrouped[1:40,5])
eijshat.Micro.R1 <-as.vector(yjs.Micro.R1-yjshat.Micro.R1)
qqnorm(eijshat.Micro.R1,main="Micropropagado - Region 1")
qqline(eijshat.Micro.R1)
boxplot(eijshat.Micro.R1)
shapiro.test(eijshat.Micro.R1)
# Shapiro-Wilk Normality Test
#data:  eijshat.Micro.R1
#W = 0.9812, p-value = 0.734

#-----#
# Conven # LA RUTINA ES ANALOGA AL TIPO MICROPROPAGADO
#-----#

## Impresion de curvas promedio ##
matplot(vectime.R1,cbind(Curva.Micro.R1,Curva.Conven.R1),type="l", lty=c(1,6),
        xlab="t (meses)", xaxt="n", xlim=c(18,50), ylab="Indice Morfologico
        (IM)",ylim=c(100,300), main="Curvas de Crecimiento Promedio: Micropropagado y
        Convencional\n ( Region 1 )")
axis(side=1,at=c(19,29,39,49),labels=c("19","29","39","49"))
legend(locator(1),c("Micro: IM=122.467+exp(0.08787*t)",
"Conven: IM=122.174+exp(0.08817*t)"),lty=c(1,6))

####-----####
#### REGION 2 #### LA RUTINA ES ANALOGA A LA REGION 1
####-----####

```

NOTA: El programa utilizado para llevar a cabo los ajustes con base en los datos del Escenario 2 corresponde a una rutina análoga a la anterior, diferenciándose en esencia en el archivo utilizado para la lectura de los datos (IndicesOrderedEsc2), y a su vez en los resultados obtenidos de cada ajuste.

A.4 Análisis Discriminante.

La metodología y notación descrita a continuación coincide con la manejada en Mardia et al. (1995). Para más referencias sobre este tema también pueden ser consultados Manly (1986), y Johnson and Wichern (1992), entre otros.

El problema que se aborda en el análisis discriminante es esencialmente qué tan bien es posible separar dos o más grupos de individuos con base en mediciones dadas de estos individuos correspondientes a un conjunto de variables. Como en el análisis de componentes principales, el análisis discriminante está basado en la idea de encontrar combinaciones lineales adecuadas de las variables originales, donde en este caso éstas servirían para llevar a cabo la clasificación de individuos nuevos a los que se miden las variables consideradas.

Una aproximación al problema de discriminación basándose en una matriz de datos \mathbf{X} puede ser hecha sin suponer alguna forma particular paramétrica de las distribución en los g grupos involucrados, estableciendo una regla "sensible" para discriminar entre éstas. La sugerencia de Fisher fue buscar una función lineal $a'x$ que maximizase el cociente de la suma de cuadrados entre grupos con la suma de cuadrados en los grupos, esto es, si consideramos una combinación lineal de las columnas de \mathbf{X} , digamos,

$$y = Xa = \begin{bmatrix} X_1 a \\ \vdots \\ X_g a \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_g \end{bmatrix}, \quad (1)$$

entonces y tiene una suma de cuadrados totales,

$$y'Hy = a'X'HXa = a'Ta, \quad (2)$$

la cual puede ser dividida como la suma de cuadrados dentro de los grupos,

$$\sum y_i'Hy_i = \sum a'X_i'HX_i a = a'Wa, \quad (3)$$

mas la suma de cuadrados entre los grupos,

$$\sum n_i (\bar{y}_i - \bar{y})^2 = \sum n_i \{a'(\bar{x}_i - \bar{x})\}^2 = a'Ba, \quad (4)$$

donde \bar{y}_i es la media del i-ésimo sub-vector y_i de y , y H_i es la $(n_i \times n_i)$ matriz centrante.

El criterio de Fisher es intuitivamente atractivo debido a que es más sencillo decir que los grupos se separan si la suma de cuadrados entre grupos (4) es más grande en relación con la suma de cuadrados en los grupos (3). Esto puede verse en el sentido del cociente dado por

$$\frac{a'Ba}{a'Wa} \quad (5)$$

Si a es el vector que maximiza el cociente (5), a la función lineal $a'x$ la llamaremos la *función discriminante lineal de Fisher*. Note que el vector a puede ser reescalado sin afectar el cociente (5).

Por tanto, para encontrar el vector a en la expresión (5) se tiene el siguiente teorema:

Teorema. *El vector a en la función discriminante lineal de Fisher es el vector propio (eigenvector) de la matriz $W^{-1}B$ correspondiente al mayor de los valores propios (eigenvalores).*

Prueba. Ver Mardia et al. (1995), pag. 319.

Una vez que la función discriminante lineal ha sido calculada, una observación x puede ser clasificada en alguno de los g grupos con base a su "score discriminante" $a'x$. Las medias muestrales \bar{x}_i tienen scores $a'\bar{x}_i = \bar{y}_i$. Entonces, x es clasificada en la población cuyo score medio sea cercano a $a'x$; esto es, se clasifica x en el grupo i si

$$|a'x - a'\bar{x}_j| < |a'x - a'\bar{x}_i| \quad \text{para todo } i \neq j \quad (6)$$

La función discriminante lineal de Fisher es de mucha utilidad en el caso especial de tener dos grupos ($g=2$). En este caso, la matriz B tiene rango 1 y puede ser escrita como

$$B = \left(\frac{n_1 n_2}{n} \right) dd', \quad (7)$$

donde $d = \bar{x}_1 - \bar{x}_2$. Por tanto, $W^{-1}B$ tiene sólo un valor propio que puede encontrarse de manera explícita. Este valor propio es igual a

$$\text{tr}(W^{-1}B) = \left(\frac{n_1 n_2}{n} \right) d' W^{-1} d, \quad (8)$$

y el correspondiente vector propio es

$$a = W^{-1} d. \quad (9)$$

Entonces, la regla de discriminación se convierte en

$$\text{clasificar } x \text{ en el grupo } 1 \text{ si } d' W^{-1} \left\{ x - \frac{1}{2} (\bar{x}_1 + \bar{x}_2) \right\} > 0 \quad (10)$$

y en el grupo 2 en otro caso.

Así, la regla dada en (6) representa básicamente una regla de sensibilidad basada en una función lineal de x . Por tanto, se podría esperar que esta regla fuese apropiada para poblaciones donde hipótesis distribucionales, como la normalidad, no fuese a su vez satisfecha. Sin embargo, uno de los supuestos necesarios para el desarrollo de esta técnica consiste en que las matrices de covarianza de las poblaciones en cada uno de los grupos involucrados sean las mismas.

Bibliografía y Referencias

- [1] Blackman, V. H. (1919). "*The compound interest law and plant growth*". *Ann. Bot.* 33: 353-360.
- [2] Charles, S. D. (2002). *Statistical Methods for the Analysis of Repeated Measurements*. First Edition, New York: Springer-Verlag.
- [3] Chatfield, C. and Collins, A. J. (1980). *Introduction to Multivariate Analysis*. London: Chapman and Hall.
- [4] Clow, D. J. and Urquhart N. S. (1974). *Mathematics in Biology: Calculus and Related Topics*. Preliminary Edition. New York: W. W. Norton & Company.
- [5] Coombs, J., Hall D. O., Long S. P. and Scurlock J. M. O. (1985). *Techniques in Bioproductivity and Photosynthesis*. Second Edition, New York: Pergamon Press.
- [6] Jackson, J. E. (1991). *A User's Guide To Principal Components*. New York: John Wiley & Sons.
- [7] Johnson, R. A. and Wichern, D. W. (1992). *Applied Multivariate Statistical Analysis*. Third Edition, New Jersey: Prentice Hall.
- [8] Jolliffe, I. T. (1986). *Principal Component Analysis*. New York: Springer-Verlag.
- [9] Kanji, G. K. (1993). *100 Statistical Tests*. First Edition, London: SAGE Publications.
- [10] Lindstrom, M. J. and Bates, D. M. (1990). "*Nonlinear Mixed Effects Models for Repeated Measures Data*". *Biometrics* 46: 673-687.
- [11] Manly, B. F. J. (1986). *Multivariate Statistical Methods*. First Edition, London: Chapman & Hall.
- [12] Mardia, K. V., Kent, J. T. and Bibby, J. M. (1995). *Multivariate Analysis*. London: Academic Press.

- [13] Mariaca, D. (2003). *Estudio Estadístico de dos Variables Críticas en la Calidad del Agave: Peso y Contenido de Carbohidratos*. México: Dagmar Mariaca Hajducek, Abril del 2003, Tesis de Maestría en Ciencias con Especialidad en Probabilidad y Estadística, Centro de Investigación en Matemáticas, A.C.
- [14] Mead, R., Curnow R. N. and Hasted A. M. (2003). *Statistical Methods in Agriculture and Experimental Biology*. Third Edition, Boca Raton: Chapman and Hall.
- [15] Nobel P. S. (1983). *Biophysical Plant Physiology and Ecology*. Third Edition, New York: W. H. Freeman and Company.
- [16] Pinheiro, J. C. and Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer-Verlag.
- [17] Ratkowsky, D. A. (1983). *Nonlinear Regression Modeling; A Unified Practical Approach*. First Edition, New York: Marcel Dekker Inc.
- [18] Rencher, A. C. (2002). *Methods of Multivariate Analysis*. New York: John Wiley & Sons.
- [19] Seber, G. A. F. and Wild, C. J. (1989). *Nonlinear Regression*. First Edition, New York: John Wiley.
- [20] Sprent P. and Smeeton N. C. (2001). *Applied Nonparametric Statistical Methods*, Third Edition, Boca Raton: Chapman and Hall.
- [21] Venables, W. N. and Ripley, B. D. (1997). *Modern Applied Statistics with S-PLUS*. Second Edition, New York: Springer-Verlag.
- [22] Academia Mexicana del Tequila, A.C.: www.acamextequila.com.mx.