

UNIVERSIDAD AUTÓNOMA AGRARIA ANTONIO NARRO  
SUBDIRECCIÓN DE POSGRADO



Regresión con variables dummy para predecir fibra seca de *Agave lechuguilla* Torr.,  
en dos zonas bioclimáticas de gran escala en México

**Tesis**

Que presenta JOSÉ ÓSCAR MANUEL LÓPEZ DÍAZ

como requisito parcial para obtener el grado de:

MAESTRO EN CIENCIAS EN INGENIERÍA DE SISTEMAS DE PRODUCCIÓN

Saltillo, Coahuila

Diciembre, 2021

Regresión con variables dummy para predecir fibra seca de *Agave lechuguilla* Torr.,  
en dos zonas bioclimáticas de gran escala en México

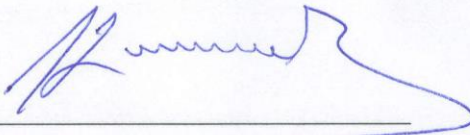
**Tesis**

Elaborada por JOSÉ ÓSCAR MANUEL LÓPEZ DÍAZ como requisito parcial para  
obtener el grado de MAESTRO EN CIENCIAS EN INGENIERÍA DE SISTEMAS  
DE PRODUCCIÓN con la supervisión y aprobación del Comité de asesoría



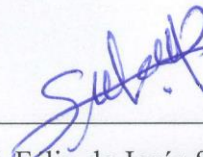
Dr. Jorge Méndez González

Asesor Principal



Dr. Alejandro Zermeño González

Asesor



M.C. Félix de Jesús Sánchez Pérez

Asesor



Dr. Marcelino Cabrera De la Fuente

Subdirección de Posgrado

UAAAN

## AGRADECIMIENTOS

A *Dios* por la salud, amor y bendiciones brindados, por darme fortaleza y protección en todo mi trayecto de vida, por todo hoy y siempre.

A la *Universidad Autónoma Agraria Antonio Narro* por acogerme una vez más y así poder lograr una etapa más de mi formación académica y profesional.

Al *Consejo Nacional de Ciencia y Tecnología (CONACYT)* por el apoyo económico brindado para mi formación como Maestro en Ciencias.

Al *Dr. Jorge Méndez González* por compartir sus conocimientos y experiencia como docente e investigador, por su permanencia y exigencia, por sus valiosos consejos, amistad y confianza que en conjunto con su esposa *Paulina Bazaldúa Sánchez* han conferido hacia mí, por las vivencias compartidas y que nunca se acaben, gracias.

Al *Dr. Alejandro Zermeño González* y al *M.C. Félix de Jesús Sánchez Pérez* por sus aportaciones y revisión del presente trabajo de investigación, por las clases impartidas y los conocimientos compartidos, además de sus valiosos consejos.

A mi compañera y esposa *M. Guadalupe C. V.*, por su espera, comprensión y paciencia durante esta etapa de superación personal y profesional, y por luchar a mi lado en momentos difíciles.

A *mis hermanos* que de una u otra manera me han exhortado a seguir adelante y cumplir mis metas, por su cariño, apoyo, amistad y comprensión, gracias.

Al *M.C. Librado Sosa Díaz* y al *M.C. Cristóbal de Jesús Flores Hernández* por el trabajo de campo y laboratorio realizado y por su apreciable amistad.

A la *M.C. Fátima Monserrat Méndez Encina* y la *Ing. Rocío Mendieta* por los momentos de estudio e inigualable amistad.

A mis compañeros de generación Marilyn Cielo Castillo R., Uriel A. Macias C. y Daniel Aldaco por los momentos de estudio y apreciable amistad.

Al programa de Maestría en Ciencias en Ingeniería de Sistemas de Producción, su personal académico y administrativo por su apoyo durante esta etapa culminada.

## DEDICATORÍA

A mis padres *María del Consuelo Díaz Escobedo*† y *Francisco López Salmerón* † donde quiera que se encuentren, por ser fuente de amor, inspiración y fortaleza, a ti madre por confiar en mí y por conducirme hacia el camino del bien, por enseñarme que en esta vida no se rinde, sino que se lucha por lo que se quiere, por su cariño y bendiciones siempre estarán en mi corazón y formarán parte de mis logros.

A mi querida esposa *M. Guadalupe C. V.* y amados hijos *Mayte Guadalupe* y *Diego Israel* que son mi principal fuente de inspiración para ser mejor cada día como padre, amigo y ejemplo a seguir, que cuando sientan no poder más se den cuenta que los límites somos nosotros mismos, y que la clave está en ser imparables. Siempre unidos por más logros como este, los amo.

A mis hermanos *Francisco, María Isabel, Arcadio, María de Jesús, Teresa, María del Consuelo, Martha Cecilia, Luis Miguel* y *Roxana*, a quienes ni la distancia, ni las circunstancias vividas los apartan de mi vida. Ojalá se den cuenta que las personas humildes también podemos cumplir sueños y metas fijadas, nunca se rindan y luchen por lo que quieren. Los quiero hermanos.

A *familiares* y *amigos*, que con su apoyo y motivación han colaborado para hacer posible este triunfo el cual es para ustedes.

## ÍNDICE GENERAL

RESUMEN.....	ix
ABSTRACT.....	x
INTRODUCCIÓN .....	1
REVISIÓN DE LITERATURA.....	3
Distribución y hábitat de <i>Agave lechuguilla</i> Torr. ....	3
Impacto social, legal, económico y ambiental.....	3
Principales usos del recurso no maderable .....	4
Extracción y producción de fibra de <i>Agave lechuguilla</i> .....	4
Estudios sobre plantaciones y distribución potencial de <i>Agave lechuguilla</i> .....	5
Estudios sobre biomasa de <i>Agave lechuguilla</i> .....	6
Estudios sobre fibra seca de <i>Agave lechuguilla</i> .....	8
Ecuaciones alométricas y técnicas de regresión.....	10
Corrección del sesgo en ecuaciones con transformación logarítmica.....	11
Variables Indicadoras (dummy) .....	11
Análisis de Componentes Principales .....	11
MATERIALES Y MÉTODOS .....	13
Descripción de área de estudio .....	13
Zonificación bioclimática de áreas de distribución de <i>Agave lechuguilla</i> Torr. ....	13
Muestreo de <i>Agave lechuguilla</i> Torr. para la extracción de fibra .....	13
Generación y validación de la ecuación para estimar peso seco de fibra.....	14
Análisis de Correlación Canónica .....	15
RESULTADOS Y DISCUSIÓN .....	16
Zonificación bioclimática.....	16
Caracterización de la muestra de <i>Agave lechuguilla</i> Torr.....	17
Análisis de predictores de fibra seca de <i>Agave lechuguilla</i> Torr. ....	20
Modelo para predecir fibra seca de <i>Agave lechuguilla</i> Torr. ....	21
Capacidad predictiva del modelo para predecir fibra seca de <i>A. lechuguilla</i> Torr. ....	24
Relación canónica entre fibra seca de <i>A. lechuguilla</i> con variables bioclimáticas .....	25
CONCLUSIONES .....	27
BIBLIOGRAFÍA .....	28
ANEXOS .....	33

## LISTA DE TABLAS

<b>Tabla 1.</b> Estadística descriptiva de la muestra utilizada para predecir peso seco de fibra de <i>Agave lechuguilla</i> Torr. ....	19
<b>Tabla 2.</b> Estadísticos del ajuste y verificación de supuestos del modelo potencial en su forma logarítmica, de predictores de peso de fibra seca de <i>Agave lechuguilla</i> Torr. ....	20
<b>Tabla 3.</b> Valores de p para las pruebas hipótesis de intercepta, pendiente, e interacciones con la variable dummy a diferentes tamaños de muestra, con el modelo potencial en forma logarítmica, para predecir fibra seca de <i>Agave lechuguilla</i> Torr. ....	21
<b>Tabla 4.</b> Estadísticos de ajuste del modelo potencial-dummy en forma logarítmica, para estimar peso de fibra seca de <i>Agave lechuguilla</i> Torr. en zonas áridas y semiáridas de México. ....	22
<b>Tabla 5.</b> Estadísticos de ajuste de la evaluación de la capa capacidad predictiva del modelo para predecir peso seco de fibra seca de <i>Agave lechuguilla</i> Torr., empleando machine learning, con una muestra independiente (n = 100). ....	24
<b>Tabla 6.</b> Análisis de correlación entre grupos de variables dendrométricas de <i>Agave lechuguilla</i> Torr. y bioclimáticas. ....	26

## LISTA DE FIGURAS

- Figura 1.** Correlación de las variables bioclimáticas y altitud con los eigenvectores de los componentes 1 y 2 del análisis de componentes principales, PCA (a) y zonificación bioclimática de la distribución de *Agave lechuguilla* Torr., realizada a través de un PCA sin rotación (b). ..... 16
- Figura 2.** Variables dendrométricas de *Agave lechuguilla* Torr. (primeras tres) y bioclimáticas (últimas tres) por zona bioclimática (Z1 = Zona 1, Z2 = Zona 2). ..... 17
- Figura 3.** Fibra observada de *Agave lechuguilla* Torr. y estimada a través del modelo potencial-dummy, en dos zonas bioclimáticas de regiones áridas y semiáridas de México (a), y representación de observaciones influyentes del modelo (b). ..... 23
- Figura 4.** Análisis de residuales del modelo para estimar peso de fibra seca de *Agave lechuguilla* Torr. al noreste de México. Homogeneidad de varianza (a), normalidad (b) e independencia (c). ..... 23
- Figura 5.** Análisis de correlación canónica (CCA) entre variables bioclimáticas (matriz x) y variables dendrométricas (matriz y) de *Agave lechuguilla* Torr. (a) y su expresión en las primeras dos dimensiones de CCA. ..... 26

## LISTA DE ANEXOS

<b>Anexo 1.</b> Estadísticos de ajuste del modelo potencial-dummy en forma logarítmica, para establecer diferencias de variables dendrométricas de <i>Agave lechuguilla</i> Torr. y variables bioclimáticas entre zonas derivadas del PCA.....	33
<b>Anexo 2.</b> Individuos evaluados y ejidos muestreados por municipio y estado en poblaciones de <i>Agave lechuguilla</i> Torr.....	34



## RESUMEN

*Agave lechuguilla* Torr., de la familia Agavaceae, es una de las especies más representativas de las zonas áridas y semiáridas; se distribuye desde el sureste de Estados Unidos hasta el sur de México, de ella se extrae fibra, para múltiples usos. El objetivo de este estudio fue generar un modelo de regresión, para predecir de forma rápida, sencilla y de bajo costo, fibra seca (*PSF*) de esta especie. Se empleó un modelo de potencia en su forma lineal y zonas bioclimáticas como variables dummy. El entrenamiento, generación (80 %) y validación (20 %) del modelo se hizo a través de *machine learning* con el paquete 'caret' de R. Mediante análisis de correspondencia canónica (CCA) se evaluó la relación de *PSF* con variables bioclimáticas. El ACP generó dos zonas bioclimáticas, ambas de diferente productividad de *A. lechuguilla*. Se evaluaron 499 individuos en cuatro estados del país. El diámetro de copa (*Cd*) de esta especie, predice adecuadamente el peso seco de fibra ( $R^2 = 0.6327$ ;  $p < 0.05$ ). El intercepto ( $\beta_0$ ), pendiente [ $\ln Cd$  ( $\beta_1$ )], Zona [ $(\beta_2)$ ] e interacción [ $\ln Cd:Zona$  ( $\beta_3$ )] del modelo con variables dummy resultaron estadísticamente significativas ( $p < 0.05$ ), dando origen a una ecuación para cada zona bioclimática. El CCA indica una correlación positiva entre temperatura mínima del mes más frío (Bio 6) y *PSF* ( $r = 0.84$  y  $p < 0.05$ ). En conclusión, por el decremento de Bio 6, en más de 0.5 °C al 2050, la especie podría ser vulnerable al cambio climático, la producción de fibra de *A. lechuguilla* podría ser afectada de forma gradual en los próximos años.

**Palabras clave:** *Agave lechuguilla*, dummy regression, fibra seca, México, zonas áridas, zonas bioclimáticas.

## ABSTRACT

*Agave lechuguilla* Torr., of the family Agavaceae, is distributed from southwestern United States to southern Mexico and is one of the most representative species of arid and semiarid regions. Its fiber is extracted for multiple purposes. The objective of this study was to generate a robust model to predict dry fiber yield (*Dfw*) rapidly, simply, and inexpensively. We used a power model in its linear form and bioclimatic areas as dummy variables. Training, generation (80%) and validation (20%) of the model was performed using machine learning with the package 'caret' of R. Using canonical correlation analysis (CCA), we evaluated the relationship of *Dwf* to bioclimatic variables. The principal components analysis (PCA) generated two bioclimatic zones, each with different *A. lechuguilla* productivities. We evaluated 499 individuals in four states of Mexico. The crown diameter (Cd) of this species adequately predicts its fiber dry weight ( $R^2 = 0.6327$ ;  $p < 0.05$ ). The intercept ( $\beta_0$ ), slope [ $\ln Cd$  ( $\beta_1$ )], zone [ $(\beta_2)$ ] and interaction [ $\ln Cd:Zona$  ( $\beta_3$ )] of the dummy model was statistically significant ( $p < 0.05$ ), giving origin to an equation for each bioclimatic zone. The CCA indicates a positive correlation between minimum temperature of the coldest month (Bio 6) and *Dwf* ( $r = 0.84$  and  $p < 0.05$ ). In conclusion, because of the decrease in Bio 6 of more than 0.5 °C by 2050, the species could be vulnerable to climate change, and *A. lechuguilla* fiber production could be affected gradually in the coming years.

**Key words:** *Agave lechuguilla*, dummy regression, dry fiber, Mexico, arid zones, bioclimatic zones.

## INTRODUCCIÓN

En México, las zonas áridas y semiáridas representan 54 % del territorio, son habitadas por aproximadamente 40 % de la población nacional (Briones *et al.*, 2018), están representadas por los matorrales Crasicaule, Desértico Micrófilo y Desértico Rosetófilo, siendo los géneros como *Larrea sp.*, *Agave sp.*, *Dasyilirion sp.*, *Yuca sp.*, etc., los más representativos (Rzedowski, 1978). Una de las especies más importantes desde el punto de vista económico y ecológico en las zonas áridas es *Agave lechuguilla* Torr., una planta suculenta clasificada dentro de la familia Agavaceae, se distribuye desde el sureste de Estados Unidos hasta el sur de México sobre 11 entidades federativas (Velasco *et al.*, 2009; Castillo *et al.*, 2012), cubriendo 142 115 km<sup>2</sup> (Marroquín *et al.*, 1981).

A esta especie se le conoce como “lechuguilla”, destaca por su producción de fibra o “ixtle”, la cual se extrae de la yema o “cogollo” una vez que este alcanza 25 cm de altura (SEMARNAT, 1996), es utilizada en la industria automotriz, cordelería, alfombras y cepillos de limpieza, etc. (Berlanga *et al.*, 1992). La fibra de *A. lechuguilla* se puede extraer de forma manual ó mecánica (Velasco *et al.*, 2009; Castillo *et al.*, 2011), la primera resulta de mejor calidad y puede alcanzar un precio de hasta 1.2 USD por kg; con esta técnica se producen 1.87 kg·hr<sup>-1</sup>, mientras que mecánicamente se pueden producir entre 15 y 18 kg·hr<sup>-1</sup> de fibra, pero de menor calidad (Mayorga *et al.*, 2004). La fibra de mejor calidad se obtiene de ‘cogollos’ con hojas tiernas, ya que contienen menos lignina, que las hojas de la corona de la planta (Castillo *et al.*, 2011). En plantaciones de esta especie, el cogollo tarda en regenerarse (alcanzar  $\geq$  25 cm de altura) entre siete y ocho meses; mientras que naturalmente tarda entre 16 y 24 meses (Castillo *et al.*, 2008).

El aprovechamiento de fibra de *A. lechuguilla* en el país es crucial para los habitantes de regiones áridas y semiáridas (Pando *et al.*, 2004), de ella se benefician aproximadamente 52, 000 familias (Berlanga *et al.*, 1992). Para ello, es necesario cuantificar las existencias de fibra, la cual se realiza directamente cosechando la planta o indirectamente a través de modelos alométricos. En México solo existen tres estudios para estimar peso seco de fibra de *A. lechuguilla*: Blando y Baca (2001), Pando *et al.* (2008) y Velasco *et al.* (2009); todos realizados a escala local, incluso algunos de estos modelos requieren cosechar la planta completa, o parte de ella (Castillo *et al.*, 2011;

Blando *et al.*, 2013) para estimar el peso de fibra. Puesto que el objetivo de crear un modelo es para hacer predicciones con nuevos datos, es necesario que el modelo cumpla con todos los supuestos estadísticos, pero sobre todo evaluar su capacidad predictiva (Montgomery *et al.*, 2012).

Por ello, es crucial generar modelos alométricos robustos para estimar peso seco de fibra de *A. lechuguilla* con técnicas no destructivas, de forma sencilla, rápida y eficiente, resolviendo el problema de escala, considerando las zonas de productividad, y determinar la relación de peso seco de fibra con variables bioclimáticas siendo estos los objetivos de esta investigación. Este modelo servirá no solo para respaldar trámites gubernamentales (Houry y Machaka, 2016) al autorizar programas de aprovechamiento de fibra, sino también para tomar decisiones de manejo y conservación de la especie.

## REVISIÓN DE LITERATURA

### **Distribución y hábitat de *Agave lechuguilla* Torr.**

En México, las zonas áridas y semiáridas representan 54 % del territorio, son habitadas por 40 % de la población nacional (Briones *et al.*, 2018), a la cual proveen de productos y servicios ambientales. El tipo de vegetación que las conforman son Matorral Crasicaule, Matorral Desértico Micrófilo y Matorral Desértico Rosetófilo, albergando especies como *Larrea tridentata* Sessé, *Flourenzia cernua* DC., *Agave falcata*, *Agave striata* Zucc., *Hetchia glomerata* Zucc., *Dasyilirion* spp., *Yuca* spp y de gran importancia *Agave lechuguilla* Torr., (Rzedowski, 1978), especie que se distribuye principalmente en el Desierto Chihuahuense en México y al sur de EUA (Nobel y Quero, 1986); es decir, desde Nuevo México y Texas en Estados Unidos, hasta porciones de los estados de Chihuahua, Coahuila, Durango, Guanajuato, Hidalgo, Nuevo León, Querétaro, San Luis Potosí, Tamaulipas y Zacatecas, en México (Velasco *et al.*, 2009; Castillo *et al.*, 2012), cubriendo una superficie de 142 115 km<sup>2</sup> (Marroquín *et al.*, 1981).

*Agave lechuguilla* se desarrolla en regiones con clima de tipo muy árido-semicálido (BWhw) y árido-templado (BSokw) (García, 1998), en suelos coluviales someros de sierras y lomeríos pedregosos de origen calizo, e incluso en suelos casi planos algunas veces arcillosos con buen drenaje, con pH de 7.2 a 8.5 (Berlanga *et al.*, 1992; Castillo *et al.*, 2005; Flores, 2018), donde la precipitación varía de 150 a 500 mm, altitudes desde 200 a los 2400 msnm (Rzedowski, 1978; Pando *et al.*, 2004) y temperatura de 3 y 30 °C (Nobel y Quero, 1986), incluso extremos de -8 a 44 °C y se sabe que esta planta puede soportar sequías e inundaciones (Flores, 2018), dicha especie no está limitada a una exposición especial, sin embargo, los mejores rendimientos de producción de fibra por cogollo se han presentado en exposición norte (Berlanga *et al.*, 1992). La distribución de la especie está limitada por diversos factores de precipitación, bajas temperaturas (-16 °C) y grado de insolación (Nobel y Quero, 1986).

### **Impacto social, legal, económico y ambiental**

*Agave lechuguilla* algunas veces ha sido considerada como la única fuente de recursos que tienen los habitantes de regiones áridas y semiáridas para generar ingresos

económicos y así sostener a sus familias día a día (Pando-Moreno *et al.*, 2004; Martínez *et al.*, 2013, Flores, 2018), obteniendo como dato que se benefician aproximadamente 52, 000 familias (Berlanga *et al.*, 1992). De acuerdo a Reyes-Agüero *et al.* (2000) se ven involucrados cerca de 20, 000 recolectores y sus familias.

La lechuguilla forma un papel muy importante en los ecosistemas de los desiertos mexicanos, considerándose como una planta tipo CAM, ya que su sistema fotosintético presenta mayor absorción de CO<sub>2</sub> durante la noche (Nobel y Quero, 1986; Reyes *et al.*, 2000).

Para el caso de lechuguilla *Agave lechuguilla* Torr. se utiliza la Norma Oficial Mexicana NOM-008-SEMARNAT-1996 respetando los procedimientos, criterios y especificaciones que convengan para realizar el aprovechamiento, transporte y almacenamiento de cogollos (SEMARNAT, 1996).

### **Principales usos del recurso no maderable**

Desde el siglo XX los principales usos que se le ha dado a la lechuguilla son para la elaboración de jabón y la extracción de fibra (Houri y Machaka, 2016), la cual una vez extraída se utiliza en la industria automotriz, cordelería, alfombras y cepillos de limpieza (Berlanga *et al.*, 1992). También se emplea el ixtle para manufactura de morrales, reatas o sogas, reatillas, hilos, billeteras, monederos, cinturones, portafolios, bolsos, bolsas, sacudidores, cepillos (Francisco *et al.*, 2017), entre otros (Mayorga *et al.*, 2004). Últimamente se han estado haciendo estudios científicos orientados a la obtención de etanol y celulosa (Ortíz *et al.*, 2017).

### **Extracción y producción de fibra de *Agave lechuguilla***

La fibra se puede obtener de forma manual (usando cuchillo sin filo, banco de madera y bolillo) (Velasco *et al.*, 2009) y mecánica (mediante máquina con rodillo o tambor giratorio con clavos), donde la obtenida manualmente alcanza precio de 8 a 11 pesos y de 0.9 a 1.2 USD por kg, considerando es de buena calidad, sin manchas ni grumos y se obtiene una producción de 1.87 kg/hr, a diferencia de la obtenida con el segundo método obteniendo una producción de 15 -18 kg/hr, la calidad es baja (Mayorga *et al.*, 2004). Es importante considerar que la fibra de mayor calidad y valor comercial se obtiene del cogollo o yema, que está formado por las hojas más tiernas y contienen menos lignina, respecto a las hojas de la corona de la planta (Castillo *et al.*, 2011).

### **Estudios sobre plantaciones y distribución potencial de *Agave lechuguilla***

Castillo et al. (2008), en el estado de Coahuila establecieron plantaciones de *Agave lechuguilla*, la mayoría de ellas con fines de investigación, tal es el caso de la plantación realizada en el Rancho La Gloria, General Cepeda en el año 2002, dicho trabajo experimental se llevó a cabo para evaluar el tiempo que tarda el cogollo en regenerarse en plantas bajo cultivo, donde se obtuvo que es entre siete y ocho meses, mientras que en poblaciones naturales esta entre 14 y 16 meses en ocasiones hasta 22 ó 24. La temporada de primavera y verano obtuvieron los incrementos promedio más altos de 3 a 4.5 cm mes.

Castillo et al. (2014), realizaron estudio acerca de la identificación de áreas para el establecimiento de plantaciones forestales comerciales de *Agave lechuguilla* Torr. en Coahuila, México, llegaron a la conclusión de que existen 5, 159,273 hectáreas en dicho estado con condiciones aptas para establecer plantaciones de lechuguilla, principalmente en los municipios de Cuatro Ciénegas, Sierra Mojada, Ocampo, Candela, Monclova, Escobedo, Frontera, Abasolo, Castaños, Lamadrid, Nadadores, Sacramento, San Buena Aventura, Arteaga, General Cepeda, Parras de la fuente, Ramos Arizpe y Saltillo. Existen otros estudios, Martínez et al. (2014), encontraron una superficie actual (5, 159, 274 ha) similar a la encontrada por Castillo et al. (2014), con potencial productivo de lechuguilla en el estado de Coahuila, algo interesante y diferente a los demás investigadores relacionados con el tema, es que Martínez y colaboradores (2014) hicieron proyecciones a futuro y la superficie para ésta especie no cambia en ninguno de los escenarios proyectados (2020, 2040 y 2060) al menos para este estado.

Houri y Machaka (2016), realizaron un estudio para ubicar las áreas con potencial productivo de biomasa de *Agave lechuguilla* en zonas áridas y semiáridas a nivel país, donde encontraron que en total existe un área de 540,553 km<sup>2</sup> viables para la producción de biomasa, para la cual se estimó una productividad anual de 93.8 millones de toneladas por año en general, de la cual 32, 030 km<sup>2</sup> se clasificó como el área altamente adecuada con una productividad de 380 t\*año<sup>-1</sup> por km<sup>2</sup>/año.

### **Estudios sobre biomasa de *Agave lechuguilla***

Zarate et al. (1991), ajustaron la ecuación tipo potencial para estimar el peso verde del cogollo en gramos (Y), donde se usó el volumen del cogollo como variable independiente (x),  $Y = 0.563 X^{0.6093}$ , con un coeficiente de correlación de 0.761.

Berlanga et al. (1992), realizaron un estudio en el campo experimental La Saucedá, de Ramos Arizpe, Coahuila, donde evaluaron longitud del cogollo y su diámetro basal, para conocer la metodología para la evaluación y manejo de lechuguilla en condiciones naturales. Para ello se elaboró una tabla para estimar la biomasa del cogollo de lechuguilla donde se usó la ecuación tipo potencial para estimar el peso verde del cogollo en gramos (Y), donde se usó el volumen del cogollo como variable independiente (x),  $Y = 0.563 X^{0.6093}$ , con un coeficiente de correlación de 0.761 dicha ecuación fue elaborada por Zarate et al. (1991).

Valencia et al. (2003), realizaron un estudio sobre lechuguilla, orégano y mezquite en el municipio de San Juan de Guadalupe, Durango, donde la precipitación media anual es 300 mm, temperatura media anual de 18 a 20 °C, mes más frío a 18 °C. Se aplicó un muestreo aleatorio de 10 plantas de lechuguilla que fuesen aprovechables, 15 de orégano y 10 de mezquite donde se obtuvo el fruto. Se midieron las variables de diámetro promedio de corona y altura. Se obtuvo el volumen y área de cada planta. Se ajustó ecuación de regresión lineal simple de la forma  $Y = B_0 + B_1X$ . Las variables que mejor r mostraron para predecir peso verde de lechuguilla fue diámetro mayor ( $r = 0.92$ ) y el volumen ( $r = 0.94$ ), Los coeficientes que mostraron al ajustar el modelo fue  $R^2 = 0.86$  y  $R^2 = 0.89$ , con  $p < 0.05$ . Las ecuaciones finales fueron las siguientes  $Y = 0.83 + 0.04$  (diámetro mayor);  $Y = 2.44 + 12.87$  (volumen). El peso de la planta fue de 3.44 kg con  $sd = 0.66$ , para diámetro mayor = 63.4 cm con  $sd = 15.01$  y para altura  $cm = 67.2$  con  $sd = 8.91$ .

Pando et al. (2004), realizaron un estudio en el ejido Estanque Nuevo al sur de Nuevo León, México. Altitud de 1700 m snm, una precipitación media anual de 497.88 mm, (intervalo de 182.7 a 963 mm). Temperatura media anual de 19.2 °C (intervalo de 13.9 a 22.6 °C), colectaron 40 plantas mediante muestreo dirigido con el fin de cubrir las dimensiones existentes de plantas adultas en el área, estas deben tener más de 20 cm de longitud. La variable dependiente: biomasa del cogollo e independientes: altura planta, volumen planta, longitud del cogollo, diámetro del cogollo, altura x diámetro



del cogollo, y volumen del cogollo fueron consideradas. La correlación más alta se mostró entre Biomasa del cogollo y volumen del mismo ( $r = 0.9098$ ), quedando el modelo  $Y = 5.91614 + 0.37115 \times X$ . Como resultados la producción de biomasa promedio fue 22.7 g en la primera evaluación para platas cosechadas, en la segunda evaluación fue 6.2 g siendo la más baja para plantas no cosechadas con anterioridad. Al final del tratamiento la producción de biomasa fue alta ( $p = 1.78e-05$ ,  $df = 78$ ,  $\alpha = 0.05$ ), de igual manera al ser evaluada por periodos la producción de biomasa fue alta ( $p = 6.94e-05$ ,  $df = 78$ ,  $\alpha = 0.05$ ), Para plantas cosechadas en el primer periodo no hubo diferencias ( $p = 0.173$ ,  $df = 78$ ,  $\alpha = 0.05$ ). También atribuyen que la precipitación puede inferir en la producción de biomasa del cogollo

Flores et al. (2020), realizaron un estudio sobre *Agave lechuguilla* Torr. al norte de México, en los estados de Coahuila, San Luis Potosí y Zacatecas con el fin de ajustar dos ecuaciones alométricas usando variables dummy para predecir biomasa aérea. La distribución de la muestra se hizo solo en las áreas de ejidos bajo aprovechamiento vigente, seleccionando las plantas de manera dirigida considerando todas las medidas o dimensiones existentes (diámetro, altura y diámetro de corona) en al menos un km dentro del área de aprovechamiento. El análisis estadístico consistió en probar dos ecuaciones la potencial  $\ln AGB = \ln B_1 + B_1 \ln Cd + E$  y la de Schumacher-Hall ( $\ln AGB = \ln B_0 + B_1 \ln Cd + B_2 \ln H + E$ ) en su forma linealizada. Para el ajuste de los modelos se usó Mínimos Cuadrados Ordinarios (OLS). La selección del mejor modelo se basó en el mejor  $R^2$  adj, menor error estándar y AIC y PRESS para su validación. Se aplicó regresión robusta con Mínimos Cuadrados Generalizados (GLS). El total de muestras fueron 533 (Coah: 175, SLP: 178 y Zac:180). La altura de plantas varió de 3.5 a 118 cm, el diámetro de corona varió de 3.5 a 166.5 cm, y la biomasa de 0.001 a 8.17 kg. Usando el modelo potencial-dummy se derivaron dos modelos, uno para Zacatecas  $ABG = B_0 + B_2 Zac + B_1 \ln Cd$ , y otro para Coahuila y San Luis Potosí  $ABG = B_0 + B_1 \ln Cd$ , el general con  $R^2 = 0.877$  y  $sd = 0.531$ . Del modelo de Schumacher se derivó un modelo para cada estado, indicando que parten de diferente intercepto y presentan diferente pendiente, El general con  $R^2 = 0.914$  y  $sd = 0.44$ , esto sin datos atípicos.

La validación del modelo de Schumacher, se evaluó a partir de los supuestos de normalidad Lilliefors ( $D = 0.020$ ,  $p = 0.863$ ), Homocedasticidad mediante Breusch Pagan (BP = 4.580,  $p = 0.333$ ) y autocorrelación con prueba de Ljung-Box ( $X^2 = 2.371$ ,

$p = 0.124$ ). Los Hat Values detectaron 5 valores potencialmente influyentes y 3 influyentes para la estimación de AGB.

Martínez y Meyer (1985), realizaron un estudio en el ejido la Tinaja, Salinas de Hidalgo, San Luis Potosí con el objetivo de entender por qué las poblaciones estaban decayendo y el manejo adecuado para mantener el rendimiento a largo plazo. La precipitación presente va de 350 a 400 mm anuales, a una altitud de 2280 m. El muestreo fue destructivo en un total de ocho plantas de diversos tamaños y se estimó  $Y =$  biomasa seca aérea (kg), y altura x diámetro ( $m^2$ ) como variable independiente. Ecuación potencial:  $Y = 1.44 \times (\text{Altura} \times \text{diámetro})^{2.03}$ , con  $R^2 = 0.97$ . donde  $(\text{Altura} \times \text{diámetro}) = 64.7 (\text{diámetro del tallo})^{2.03}$ ,  $R^2 = 0.97$ ).

### **Estudios sobre fibra seca de *Agave lechuguilla***

Blando y Baca (2001), realizaron un estudio en San Juan de Guadalupe, Durango, seleccionando sitios circulares de 250  $m^2$  a partir de fotografías aéreas 1: 75000, se seleccionó una planta la que tuviera la moda de altura por cuadrante; es decir cuatro ejemplares por sitio. No indica cuantos sitios se evaluaron ( $n=287$ ). Sin embargo, elaboraron un modelo lineal múltiple:  $\text{Peso seco de fibra} = -2.155994 + 0.020571 (a) + 0.320297 (ht)$  donde:  $a =$  área de cobertura,  $ht =$  número de hojas tallables o aprovechables, encontrando un peso seco de fibra (g/planta) por topografía, sierra = 45.60, bajada = 70.50, Loma = 67.44 y Meseta = 56.50, reportando diámetros de corona promedio (cm) 74.7, 57.5, 32.93 y 62.25 y alturas promedio 50.5 cm, 50.1, 54.125 y 69.75. Reportando que las topografías lomas y bajadas presentan alta producción y nivel alto, las sierras presentan baja producción y nivel medio. El mayor número o porcentaje de plantas aprovechables en las lomas (22.47 %), manteniendo un desarrollo más homogéneo.

Pando et al. (2008), realizaron un estudio para estimar fibra seca de lechuguilla en Tamaulipas (Ejido Independencia [Temperatura máxima de 27 °C, mínima 15.8 °C y precipitación media anual 515 mm], Jaumave) y Coahuila (Estación Marte [Temperatura máxima de 32 °C, mínima 7 °C y precipitación media anual 318 mm], General Cepeda, La Saucedá [Temperatura máxima de 31 °C, mínima 8 °C y precipitación media anual 293 mm] y Estación Paredón [Temperatura máxima de 33 °C, mínima 7 °C y precipitación media anual 293 mm], Ramos Arizpe). En total se muestrearon 95 plantas que tuviesen una altura superior a 25 cm. La variable

dependiente fue peso seco de fibra (g) y las independientes fueron altura y diámetro del cogollo, altura y diámetro de la planta y número de hojas. Se aplicó el método de Stepwise regresión para ajustar el modelo y ver que variable independiente ayudaría a predecir mejor el peso seco de la fibra. La variable que mejor ayudó a predecir el peso de la fibra fue altura del cogollo x diámetro de cogollo (HD) y volumen en el caso de paredón.

Encontraron que el peso seco de fibra (FDW) es diferente para cada localidad ( $p < 0.001$ ). El FDW fue similar para Paredón y Jaumave (26.26 y 24.36 g), Saucedá y Marte no presentaron diferencias en FDW (11.86 y 9.6 g). El modelo final para todas las localidades fue tipo polinomio cuadrático con  $R^2 = 0.869$ ,  $SEE = 5.041$ :  $Y = B_0 + B_1X + B_2X^2 \rightarrow \text{FDW} = -1.8725 + 0.0949 \times \text{HD} + 0.0002 \times \text{HD}^2$ , comparado con el modelo lineal con  $R^2 = 0.858$  y  $SEE = 5.207$ . Aunque sería mejor aplicar un modelo adecuado por localidad, el modelo cuadrático seleccionado presenta pendientes similares entre localidades.

Velasco et al. (2009), realizaron estudio en San Luis Potosí (San Antonio Cornado; mpio. Catorce, El Coyote y Puerto Clavellinas; mpio. Guadalcázar) con el objetivo de mostrar método para generar modelos predictivos del peso seco de la fibra de lechuguilla a partir de diámetro basal y altura del cogollo y proporcionar tablas para estimar el peso seco de la fibra de lechuguilla en tres condiciones ecológicas del N-NE de San Luis Potosí, donde la precipitación media anual va de 300 a 400 mm, un rango altitudinal que va de 1500 a 2250 m, precipitación media anual de 15 a 17 °C. El muestreo aplicado fue cuadrante a partir de punto central muestreando 60 sitios obteniendo 240 observaciones. El modelo probado fue el de Schumacher ( $Y = B_0X_1^{B_1} + X_2^{B_2} + B + E$ ) donde la variable  $Y =$  Peso seco de la fibra (g) y las variables independientes fueron  $X_1 =$  diámetro basal (cm) y  $X_2 =$  altura total de cogollo (cm). Los estadísticos del modelo para Coronado ( $\text{PSF} = \exp(-7.81338) \times \text{DB}^{1.52498} \times \text{Alt}^{2.20623}$ ) arrojaron una  $R^2 = 0.68$ ,  $p = 0.001$  indicando existe una relación lineal entre peso de la fibra de lechuguilla y cualquiera de las variables independientes. La prueba de autocorrelación Durbin Watson arrojó un  $D-W = 1.531$  valor cercano a 2 mas no muestran valor de p. evaluaron normalidad con Shapiro Wilk con un  $p\text{-value} = 0.0777$ . Para la validación del modelo se dejó una muestra aleatoria e independiente de 29 datos.

Los modelos generados para cada locación:

$$Y = 0.00040 X_1^{1.52498} X_2^{2.20623} \rightarrow \text{San Antonio Coronado}$$

$$Y = 0.00078 X_1^{1.64385} X_2^{1.13236} \rightarrow \text{El Coyote}$$

$$Y = 0.00038 X_1^{1.82763} X_2^{1.43348} \rightarrow \text{Puerto Clavellinas}$$

### **Ecuaciones alométricas y técnicas de regresión**

¿Qué es la alometría? Se define como la relación estadística que existe entre dos o más dimensiones, o componentes propias de un individuo o población, regularmente de forma exponencial. Así entonces las ecuaciones alométricas permiten predecir la biomasa, volumen o fibra de un árbol, arbusto o matorral con variables dendrométricas de fácil medición como el diámetro o altura (Picard, *et al.*, 2012). El análisis de regresión es una técnica para generar o ajustar ecuaciones alométricas, mediante el método de Mínimos Cuadrados Ordinarios (Montgomery *et al.*, 2012; Picard *et al.*, 2012). Este se define como la relación que existe entre una variable cuantitativa (dependiente) y una o más variables explicativas (independientes); es decir, los cambios que ocurren en una variable, están en función de la otra (Fox, 2016).

Las condiciones del ecosistema pueden llevar a usar una regresión lineal simple, regresión lineal múltiple, incluso regresión de tipo no lineal, como es el caso de la ecuación de tipo potencia que se puede linealizar aplicando transformación logarítmica (Wood, 1986; Picar *et al.*, 2012). Para que un modelo sea robusto se debe evaluar los supuestos de normalidad, homogeneidad de varianza e independencia de los residuales, para saber si los errores son aleatorios, independientes e idénticamente distribuidos según una distribución normal con media cero y varianza  $\sigma^2$  (Faraway, 2014; Montgomery, *et al.*, 2012, Fox, 2016). En estudios de ciencias biológicas y sociales la correlación que existe entre variables es menor, por lo tanto, puede presentarse un valor bajo de  $R^2$ , por ejemplo, un  $R^2 = 0.6$  se considera bueno. En cambio, si los estudios son en experimentos con condiciones controladas como en la física o ingeniería, un  $R^2 = 0.6$  se considera bajo (Faraway, 2014).

La validación del modelo seleccionado debe realizarse empleando un conjunto de datos independientes; es decir, datos no involucrados en la obtención del modelo final, que tengan misma variable respuesta y explicativas o independientes, con los cuales se evalúa la capacidad predictiva del modelo (Snee, 1977; Velasco *et al.*, 2009; Picard

*et al.*, 2012). La validación se puede realizar mediante la técnica de validación cruzada con método Leave One Out Cross-Validation (Picard, *et al.*, 2012), k-fold validation (k= 10), Repeated K-fold (k = 10 & 3 sets) y Bootstrap Cross-Validation (Kuhn *et al.*, 2021).

### **Corrección del sesgo en ecuaciones con transformación logarítmica**

El regreso de  $\ln$  a sus unidades originales no se hace directamente, puesto que si la distribución de un  $\log Y$  es normal dado un  $\log X$ , el resultado del antilogaritmo  $\log Y$  puede no ser normal y sesgado, ya que al hacerlo se obtiene la mediana y no la media (Wood, 1986), esto se puede corregir aplicando al modelo final un factor de corrección (CF), el cual viene dado por:  $CF = e^{(SEE^2/2)}$ , donde SEE es el error de estimación de la regresión (Sprugel, 1983).

### **Variables Indicadoras (dummy)**

Usualmente se usan variables cuantitativas (temperatura, distancia, presión, etc.) en un análisis de regresión, sin embargo algunas veces es necesario estudiar variables cualitativas o categóricas como predictoras (sexo: hombre, mujer; estatus de empleo: empleado, desempleado), las variables dummy (D) también conocidas como ficticias se adicionan al modelo de regresión regularmente en forma binaria (0,1) usando 1 presencia, 0 ausencia; o numérica (0,1,2) cuando son más de dos condiciones, el efecto de dicha variable puede afectar el intercepto, sin embargo, la interacción entre variable cuantitativa y cualitativa (X:D) pueden afectar la pendiente, esto se mide en el grado de significancia de los coeficientes correspondientes, lo que puede dar lugar o no a modelos diferentes si el coeficiente del efecto de la variable dummy o en su caso la interacción resultan significativos (Montgomery *et al.*, 2012; Faraway, 2014; Fox, 2016). Para más detalles visitar el siguiente link: <https://rpubs.com/cyobero/dummy-variables>.

### **Análisis de Componentes Principales**

El análisis de componentes principales (ACP) es un método multivariado de ordenación, basado en la transformación lineal de un conjunto de variables

cuantitativas dentro de otro conjunto de variables independientes no correlacionadas llamadas componentes principales, regularmente menor a la de variables originales, el método ACP identifica diferencias y forma grupos y permite conocer la relación que existe entre las variables cuantitativas consideradas (Franco e Hidalgo, 2003). Análisis de componentes principales considera la varianza total y deriva factores con varianza pequeña al igual que su error, estos generalmente son usado cuando el objetivo es reducir el número de variables (Joseph *et al.*, 2014).

## MATERIALES Y MÉTODOS

### Descripción de área de estudio

El estudio se realizó en zonas áridas y semiáridas al noreste de México, incluyó los estados de Chihuahua (Chih), Coahuila (Coah), San Luis Potosí (SLP) y Zacatecas (Zac). *Agave lechuguilla* crece en suelos coluviales, franco-arenosos, calizos y arcillosos (Berlanga *et al.*, 1992; Castillo *et al.*, 2005), donde la precipitación varía de 150 a 500 mm, altitudes desde 200 a los 2400 msnm (Rzedowski, 1978; Pando *et al.*, 2004) y temperatura de 3 y 30 °C (Nobel y Quero, 1986), incluso extremos de -8 a 44 °C y se sabe que esta planta puede soportar sequías e inundaciones (Flores, 2018).

### Zonificación bioclimática de áreas de distribución de *Agave lechuguilla* Torr.

Del sitio <https://chelsa-climate.org/> se obtuvieron las siguientes variables bioclimáticas: temperatura media anual, Bio 1 (°C), temperatura máxima del mes más cálido, Bio 5 (°C), temperatura mínima del mes más frío, Bio 6 (°C) y precipitación media anual, Bio 12 (mm), en formato ráster a una resolución de ~1 km<sup>2</sup> (Karger, *et al.*, 2017). Usando ArcMap 10.4.1, se distribuyeron aleatoriamente 1000 puntos sobre el área de distribución de *A. lechuguilla* en el país, a los cuales se les extrajo el valor de cada variable bioclimática, y altitud (msnm), obtenida del Modelo Digital de Elevación Mexicano (<https://www.inegi.org.mx/app/geo2/elevacionesmex/>). Con esta información, en R (R Core Team, 2021) se realizó un análisis de componentes principales (PCA) usando la librería ‘FactoShiny’ (Vaissie *et al.*, 2021) que requiere ‘FactoMiner’ (Husson *et al.*, 2020), empleando las variables en forma estandarizada para extraer los primeros tres componentes principales. Después, con los eigenvalores no rotados derivados del PC1, en ArcMap 10.4.1 se realizó una interpolación (Inverse Distance Weighted) de la cual se derivaron isolíneas con espaciado de 2.85.

### Muestreo de *Agave lechuguilla* Torr. para la extracción de fibra

El muestreo de *A. lechuguilla* fue selectivo, llevado a cabo solo en “ejidos” con permiso de aprovechamiento y en plantas  $\geq 10$  cm de altura (acorde a los recolectores, a partir de esta altura, el cogollo puede ser tallado de manera tradicional). Usando flexómetro de 3 m, a cada planta seleccionada se le midió: diámetro de copa promedio ( $Cd$ ) y altura total ( $H$ ). Antes del derribo de la planta (para otro fin), el cogollo se extrajo empleando “cogollera” (instrumento rústico propio de los recolectores), para

después medir su longitud ( $L_c$ ) y enseguida extraer manualmente la fibra. La fibra fresca fue colocada en bolsa de papel etiquetada debidamente y llevada al laboratorio para ser secada en estufa marca Thermo Scientific™ HERAthem™ (Modelo OMH750) a temperatura de 70 °C, hasta tener peso constante, el cual se obtuvo con báscula marca Torrey® (Modelo L-EQ) con capacidad de 5 kg y precisión de 1 g.

### **Generación y validación de la ecuación para estimar peso seco de fibra**

Para predecir fibra seca ( $Dfw$ ) de plantas individuales de *A. lechuguilla*, se probó la ecuación alométrica potencial (Ecuación 1), en su forma lineal (Ecuación 2), aplicando logaritmos para corregir la varianza del error (Montgomery *et al.*, 2012), usada por Wood (1986), Zárate *et al.* (1991) y Flores *et al.* (2020).

$$Y = \beta_0 X^{\beta_1} \varepsilon \quad [\text{Ec.1}]$$

$$\ln(Y) = \ln\beta_0 + \beta_1 \times \ln(X) + \varepsilon \quad [\text{Ec. 2}]$$

donde  $Y$  = peso seco de fibra (kg), probando como variables predictoras ( $X$ ) a diámetro de copa ( $Cd$ , cm), altura total de la planta ( $H$ , cm) y el producto de ambas,  $Cd \times H$  (cm  $\times$  cm),  $\ln$  = logaritmo natural y  $\beta_0$  y  $\beta_1$  = son coeficientes de regresión.

La ecuación 2, fue ajustada utilizando la zona bioclimática como variable dummy  $D = \begin{cases} 0 & (\text{Zona 1}) \\ 1 & (\text{Zona 2}) \end{cases}$  considerando afectación al intercepto ( $Zona$ ) y pendiente del modelo ( $X: Zona$ ), para determinar la existencia de un modelo (para toda el área) o uno para cada zona bioclimática, quedando como sigue:

$$\ln(Y) = \ln\beta_0 + \beta_1 \times (\ln X) + \beta_2 \times (Zona) + \beta_3 \times (\ln X: Zona) + \varepsilon \quad [\text{Ec. 3}]$$

Con 80 % del total de los datos (obtenidos aleatoriamente por cuantiles) y a través del método de mínimos cuadrados ordinarios (Montgomery *et al.*, 2012), el modelo fue entrenado a través *machine learnig* (ML) con el paquete de R ‘caret’ (Kuhn *et al.*, 2021). Después con esa misma librería, el modelo fue generado por bootstrap ( $n = 25$ ). Se realizaron pruebas de hipótesis ( $\alpha = 0.05$ ) a los coeficientes de regresión ( $H_0 : \beta_i = 0$  vs  $H_1 : \beta_i \neq 0, \dots, \beta_{ij}$ ) del modelo final. Se aplicaron pruebas estadísticas ( $\alpha = 0.05$ ) para asegurar el cumplimiento de los supuestos de normalidad [Lilliefors (Kolmogorov-Smirnov)], homogeneidad de varianza (Breusch-Pagan) e independencia de los errores (Durbin-Watson). Se examinaron posibles datos atípicos ( $-3 \leq r_i \leq +3$ ) a través de los residuales estudentizados  $r_i$ . Influencia sobre los



coeficientes del modelo: a) distancia de Cook,  $D_i$  (observación  $y_i$ ), y b) DFBETAS ( $\hat{\beta}_i$ ). Sobre las predicciones: a) DFFITS (influencia de la  $i$ -ésima observación sobre los valores predichos) y sobre la precisión de las estimaciones, COVRATIO (Faraway, 2014; Montgomery *et al.*, 2012; Fox, 2016). El total de hipótesis básicas del modelo de regresión se resume en la siguiente expresión:  $\varepsilon_i \sim N(0, \sigma^2)$ ,  $i = 1, \dots, n$ ; es decir, errores aleatorios, independientes e idénticamente distribuidos según una distribución normal con media cero y varianza  $\sigma^2$ .

Después de verificar el cumplimiento de todos los supuestos anteriores, la validación del modelo final se realizó en ‘caret’ (Kuhn *et al.*, 2021) con 20 % de datos independientes, usando cuatro métodos: a) Leave One Out Cross-Validation, b) k-fold validation (k= 10), c) Repeated K-fold (k = 10 & 3 sets) y d) Bootstrap Cross-Validation, calculando el Coeficiente de Determinación ( $R^2$ ), Raíz del Error Cuadrático Medio (RMSE) y Error Medio Absoluto (MAE).

El regreso de  $\ln$  a sus unidades originales para predecir peso seco de fibra de *A. lechuguilla* de la ecuación resultante, no se hace directamente, puesto que si la distribución de un  $\log Y$  es normal dado un  $\log X$ , el resultado del antilogaritmo  $\log Y$  puede no ser normal y sesgado, ya que al hacerlo se obtiene la mediana y no la media (Wood, 1986); esto se deberá corregir aplicando al modelo final un factor de corrección (calculado aquí) el cual viene dado por:  $CF = e^{(SEE^2/2)}$ , donde SEE es el error de estimación de la regresión (Sprugel, 1983).

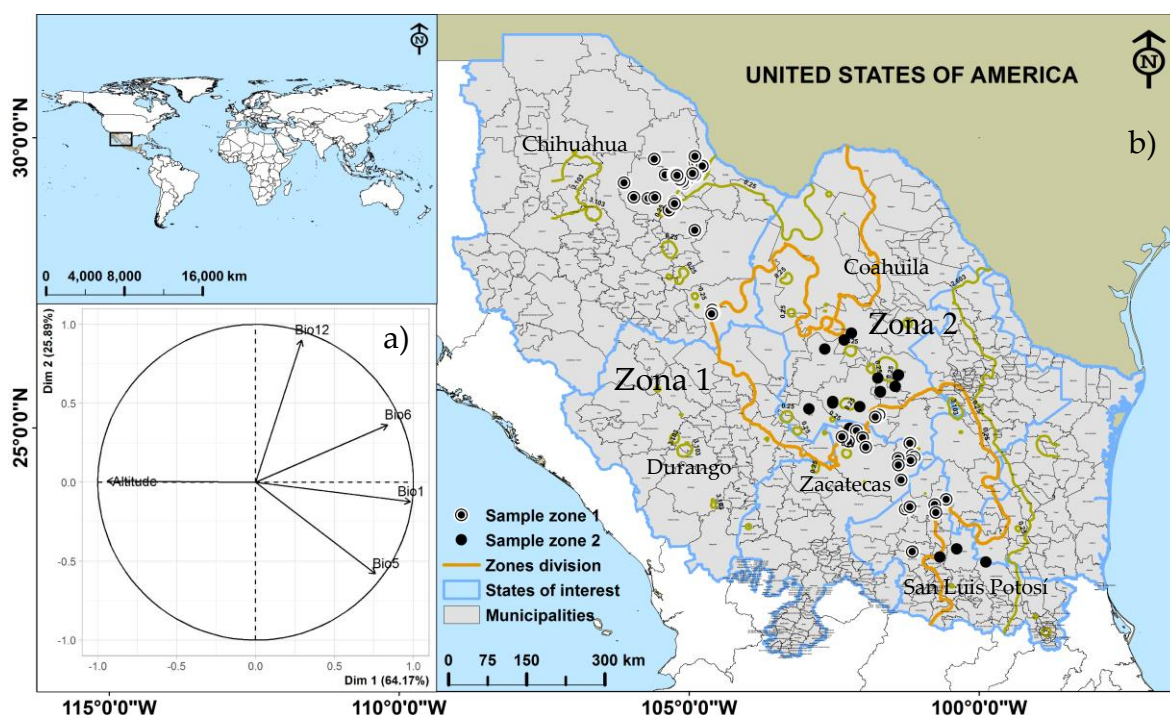
### **Análisis de Correlación Canónica**

Con la finalidad de identificar la dependencia de la fibra seca de *A. lechuguilla* con variables ambientales; así como la interdependencia entre factores (matrices  $X$ ,  $Y$ ): [ $Cd$ ,  $H$  y  $Dfw$ ], [Bio 1, Bio 5, Bio 6, Bio 12 y altitud], se realizó un Análisis de Correlación Canónica utilizando el paquete de R ‘CCA’ (Déjean y González, 2021) en software estadístico R (R Core Team, 2021). Las matrices  $X$ ,  $Y$  fueron generadas con el promedio de la variable por municipio (Anexo 2).

## RESULTADOS Y DISCUSIÓN

### Zonificación bioclimática

Los resultados de PCA mostraron que los componentes 1 (64.17 %) y 2 (25.89 %) explican 90.07 % de la varianza. La contribución de cada variable en el componente 1 fue: 29.98 % (Bio 1), 27.61 % (altitud), 21.80 % (Bio 6), 17.95 % (Bio 5), 2.66 % (Bio 12) (Figura 1a). Las puntuaciones no rotadas (eigenvalues = 0.25) dieron origen a dos zonas bioclimáticas divididas casi longitudinalmente: Zona 1: Chih, Zac y parte norte de SLP (n = 200); Zona 2: Coah, sureste de SLP y noreste de Zac (n = 199) (Figura 1b).

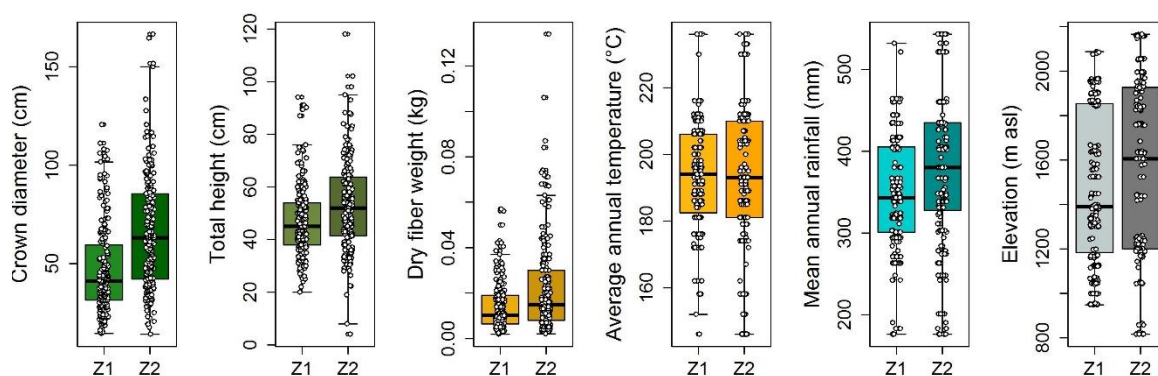


**Figura 1.** Correlación de las variables bioclimáticas y altitud con los eigenvectores de los componentes 1 y 2 del análisis de componentes principales, PCA (a) y zonificación bioclimática de la distribución de *Agave lechuguilla* Torr., realizada a través de un PCA sin rotación (b).

Utilizando estas mismas variables, más pendiente, exposición y registros de presencia de esta misma especie, algunos autores (Castillo *et al.*, 2014; Martínez *et al.*, 2015) han generado zonas de potencial productivo en esta zona del país, estimando entre 9 y 5 millones de has idóneas para plantaciones de *A. lechuguilla*, principalmente en la

zona 2 de este estudio (Figura 1b). Flores et al. (2020) construyeron modelos para estimar biomasa aérea de esta especie, considerando los límites estatales (Figura 1) como variable dummy (o zona), encontrando diferencias de biomasa entre ellas.

Al evaluar las medias de las variables dendrométricas (*Cd*, *H* y *PSF*) y bioclimáticas (Bio 1, Bio 5, Bio 6, Bio 12 y altitud) a través de modelos de regresión–dummy (equivalente a pruebas de “t” de Student) en forma logarítmica para cumplimiento de normalidad y homogeneidad de varianza:  $\ln(Y) = \beta_0 + \beta_1 \times \text{Zona}$ , se encontró que existen diferencias estadísticamente significativas entre zonas ( $p \leq 0.05$ ). En la Zona 2, la altura, diámetro de copa y cantidad de fibra de *A. lechuguilla* es más alta que en la Zona 1 y parece obedecer básicamente a temperatura mínima (Bio 6) y en menor proporción a precipitación (Bio 12), las cuales son también más altas (Figura 2, Anexo 1) en esa zona. No existen pruebas para rechazar  $H_0$ , precipitación media anual y altitud es igual entre zonas ( $p = 0.956$  y  $p = 0.106$ ).



**Figura 2.** Variables dendrométricas de *Agave lechuguilla* Torr. (primeras tres) y bioclimáticas (últimas tres) por zona bioclimática (Z1 = Zona 1, Z2 = Zona 2).

### Caracterización de la muestra de *Agave lechuguilla* Torr.

El total de plantas evaluadas fue de 499, distribuidas geográficamente en una distancia de más de 1000 km en Figura 1: (Chih = 146, Coah = 119, SLP = 124 y Zac = 110), las cuales se obtuvieron de 20 municipios y de 46 predios (Anexo 2). Como se puede observar, los individuos de *A. lechuguilla* alcanzan mayor altura (52.76 cm) y diámetro de copa (66.10 cm) en la zona 2 (Tabla 1, Figura 1), por ende, en promedio acumulan mayor cantidad de fibra en sus tejidos vegetales ( $22.21 \text{ g planta}^{-1}$ ), pero hasta máximos de 34.15 y  $46.60 \text{ g planta}^{-1}$  según el percentil 95 (Zona 1 y Zona 2). En general, para predecir fibra seca de *A. lechuguilla* se aplica muestreo selectivo, considerando un

límite mínimo de altura de la planta, *e. g.* en la Zona 2 de este estudio, seleccionado plantas  $\geq 45$  cm de altura, Blando y Baca (2001) encontraron que la acumulación de fibra seca de *A. lechuguilla* puede ser de entre 45.6 y 70.5 g planta<sup>-1</sup>, muy superior al promedio de este estudio. Nuestra investigación incluyó plantas  $\geq 10$  cm de altura, por lo que, al comparar promedios (incluso medianas) con los estudios existentes, resultan diferencias contrastantes.

**Tabla 1.** Estadística descriptiva de la muestra utilizada para predecir peso seco de fibra de *Agave lechuguilla* Torr.

Zona	Var	Media	MediaCI'	0.05	0.10	0.25	Mediana	0.75	0.90	0.95	Rango	S.D.	VCoef	Mad	IQR
Zona 1 (n = 200)	<i>Cd</i>	46.82	±3.20	19.98	23.00	31.50	41.00	59.25	83.60	94.53	106.00	22.93	0.49	17.42	27.75
	<i>H</i>	47.03	±1.88	28.95	32.00	38.00	45.00	54.00	62.83	71.10	74.00	13.49	0.29	11.86	16.00
	<i>PSF</i>	14.19	±1.49	3.90	5.00	6.45	10.20	19.00	29.73	34.15	55.20	10.66	0.75	7.12	12.55
Zona 2 (n = 199)	<i>Cd</i>	66.10	±4.05	28.45	32.90	42.00	63.00	85.43	102.10	114.70	152.50	28.99	0.44	31.88	43.43
	<i>H</i>	52.76	±2.42	29.09	33.00	41.40	52.00	63.70	74.04	83.00	114.00	17.28	0.33	16.31	22.30
	<i>PSF</i>	22.21	±2.87	4.00	5.00	8.00	15.00	30.00	0.05	46.6	132.00	20.48	0.92	13.34	22

*Cd*= diámetro de copa (cm), *H*= altura total (cm), *PSF*= Peso seco de fibra (g), *MeanIC* = intervalo de confianza de la media, 0.05, ..., 0.95 = cuantiles de la variable, *SD* = desviación estándar, *VCoef* = coeficiente de variación, *Mad* = desviación absoluta de la mediana, *IQR* = rango intercuantílico.

### Análisis de predictores de fibra seca de *Agave lechuguilla* Torr.

En un análisis preliminar, la ecuación 2 se ajustó con cada predictor ( $Cd$ ,  $H$ ,  $Cd \times H$ ); aunque todos muestran capacidad para predicen fibra seca de *A. lechuguilla* ( $0.519 < R^2 < 0.633$  y  $p\text{-value} < 2e-16$ ), solo  $Cd$  pasa todos los supuestos de un modelo lineal (Tabla 2), por lo que se eligió este como predictor de fibra seca de *A. lechuguilla*. Se ha demostrado que los mejores predictores en esta especie son  $Cd$  y  $H$ , e. g. Valencia *et al.*, (2003), utilizaron  $Cd$  (cm) y volumen de la planta ( $m^3$ ) para predecir peso fresco de la planta (kg), cuya correlación es 0.9272 y 0.9433 respectivamente. Flores *et al.* (2020) demuestran que  $Cd$  y  $H$  predicen eficientemente biomasa aérea, explican 91.4 %. En métodos destructivos, se ha probado que el volumen del cogollo ( $cm^3$ ) puede explicar desde 57.9 % (Zárate *et al.*, 1991; Berlanga *et al.*, 1992) hasta 90.98 % (Pando *et al.*, 2004) el peso fresco del cogollo. El índice del estado hídrico (MPa) y el índice de radiación fotosintéticamente activa ( $mmol\ m^{-2}\cdot S^{-1}$ ) explican 97 y 25 % la biomasa de esta especie (Quero y Nobel, 1987).

**Tabla 2.** Estadísticos del ajuste y verificación de supuestos del modelo potencial en su forma logarítmica, de predictores de peso de fibra seca de *Agave lechuguilla* Torr.

Coef	Estimado	S.E.	t-value	Pr(> t )	R <sup>2</sup>	KS	BP	D-W
Intercepto ( $\beta_0$ )	-9.132	0.186	-49.2	< 2e-16 ***				
$\ln Cd$ ( $\beta_1$ )	1.226	0.047	26.03	< 2e-16 ***	0.577	0.233	0.223	0.088
Intercepto ( $\beta_0$ )	-11.027	0.289	-38.09	< 2e-16 ***				
$\ln H$ ( $\beta_1$ )	1.733	0.750	23.18	< 2e-16 ***	0.519	2.2e-16	3.58e-16	0.004
Intercepto ( $\beta_0$ )	-10.721	0.219	-48.92	< 2e-16 ***				
$\ln CdH$ ( $\beta_1$ )	0.821	0.028	29.26	< 2e-16 ***	0.633	0.003	0.202	0.102

Coef = coeficiente, S.E. = error estándar, R<sup>2</sup> = coeficiente de determinación. Valor de p de la prueba de normalidad, KS (Kolmogorov-Smirnov), Homogeneidad de varianza, BP (Breusch Pagan) e independencia, D-W (Durbin Watson).  $Cd$  = diámetro de copa,  $H$  = altura total de la planta,  $\ln$  = logaritmo natural,  $\beta_0$  y  $\beta_1$  = coeficientes de regresión. Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1.

Adicionalmente, las pruebas de hipótesis realizadas sobre los coeficientes de regresión: intercepta ( $\beta_0$ ), pendiente [ $\ln Cd$  ( $\beta_1$ )], variable dummy [Zona ( $\beta_2$ )] e interacción [ $\ln Cd$ :Zona ( $\beta_3$ )] de la ecuación 3, resultaron estadísticamente significativas ( $p < 0.001$ ) en diferentes tamaños de muestra (Tabla 3), lo que demuestra una correcta zonificación bioclimática de la productividad de *A. lechuguilla* mediante PCA y que el predictor ( $\ln Cd$ ) de fibra seca de esta especie, no proviene de un modelo aleatorio.

**Tabla 3.** Valores de p para las pruebas hipótesis de intercepta, pendiente, e interacciones con la variable dummy a diferentes tamaños de muestra, con el modelo potencial en forma logarítmica, para predecir fibra seca de *Agave lechuguilla* Torr.

Coeficiente	Tamaño de muestra (%)						Media (coef.)	S.D. (coef.)
	100	90	80	70	60	50		
Intercepto ( $\beta_0$ )	2e-16	2e-16	2e-16	2e-16	2e-16	2e-16	-8.745	0.119
$\ln Cd$ ( $\beta_1$ )	2e-16	2e-16	2e-16	2e-16	2e-16	2e-16	1.139	0.031
Zona ( $\beta_2$ )	1.0e-4	6.2e-5	0.001	1.2e-4	6.6e-5	5.4e-5	-1.375	0.236
$\ln Cd$ :Zona ( $\beta_3$ )	0.020	0.003	0.015	0.002	0.002	9.2e-4	0.314	0.062
R	0.771	0.764	0.781	0.778	0.757	0.782		

$Cd$  = diámetro de copa,  $\ln$  = logaritmo natural,  $\beta_0, \dots, \beta_3$  = coeficientes de regresión, Zona = dummy variable, Media (coef) = media de los coeficientes de regresión, S.D. (coef) = desviación estándar de los coeficientes de regresión. R = Coeficiente de correlación múltiple.

### Modelo para predecir fibra seca de *Agave lechuguilla* Torr.

El modelo final obtenido a través de ML, mostró una  $R^2$  de 0.629. Los coeficientes de regresión resultaron altamente significativos,  $p < 0.05$  (Tabla 4), similares al promedio obtenido con diferentes tamaños de muestra (Tabla 3). La variable dummy (Zona) e interacción: ( $\ln Cd$ :Zona) muestran valor de  $p$  de 0.0004 y 0.0014 (Tabla 4), dando lugar a una ecuación para cada zona bioclimática (Figura 3a), para estimar el peso seco de fibra de *A. lechuguilla*.

Por la alta variación de fibra observada en esta especie (Tabla 1), en el análisis de regresión, se presentan algunas observaciones influyentes (Figura 3b) sobre los coeficientes del modelo (DFFITS), pero ninguna con  $D_i > 1$ , por lo que fue necesario eliminarlas; ninguna observación influyente respecto a DFBETAS. Del COVRATIO, destacan como influyentes varias observaciones, pero resaltar que no se detectó ningún dato atípico  $r_i$ . El factor de corrección obtenido (1.127) y multiplicativo a la ecuación final, puede corregir el sesgo y mejorar las estimaciones de *PSF* en 12.7 %.

**Tabla 4.** Estadísticos de ajuste del modelo potencial con variables dummy en forma logarítmica, para estimar peso de fibra seca de *Agave lechuguilla* Torr. en zonas áridas y semiáridas de México.

Coefficiente	Estimado	Error std.	t-value	Pr(> t )	I.C. 2.5 %	I.C. 97.5 %
Intercept ( $\beta_0$ )	-8.826	0.252	-35.086	< 2e-16 ***	- 9.321	-8.332
$\ln Cd$ ( $\beta_1$ )	1.162	0.066	17.640	< 2e-16 ***	1.033	1.292
Zona ( $\beta_2$ )	-1.468	0.414	-3.546	0.0004 ***	- 2.282	-0.654
$\ln Cd:Zona$ ( $\beta_3$ )	0.335	0.104	3.219	0.0014 **	0.130	0.540

I.C. = intervalo de confianza,  $Cd$  = diámetro de copa,  $\ln$  = logaritmo natural,  $\beta_0, \dots, \beta_3$  = coeficientes de regresión, Zona = variable dummy, Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1.

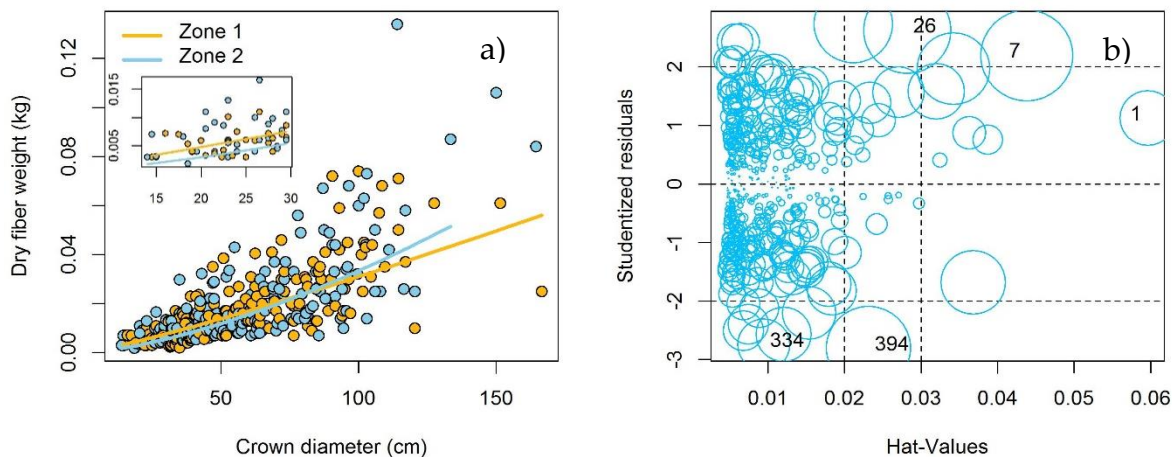
De esta forma, la ecuación resultante para estimar peso seco de fibra de *A. lechuguilla* en la zona bioclimática 1 y 2 respectivamente (Figura 1b), queda de la siguiente manera:

$$PSF = \exp(-8.82605 + 1.16201 \times (\ln Cd)) \times F.C. \quad [\text{Ec. 4}]$$

$$PSF = \exp(-8.82605 + (-1.4682) + (0.33514 + 1.16201) \times \ln Cd) \times F.C. \quad [\text{Ec. 5}]$$

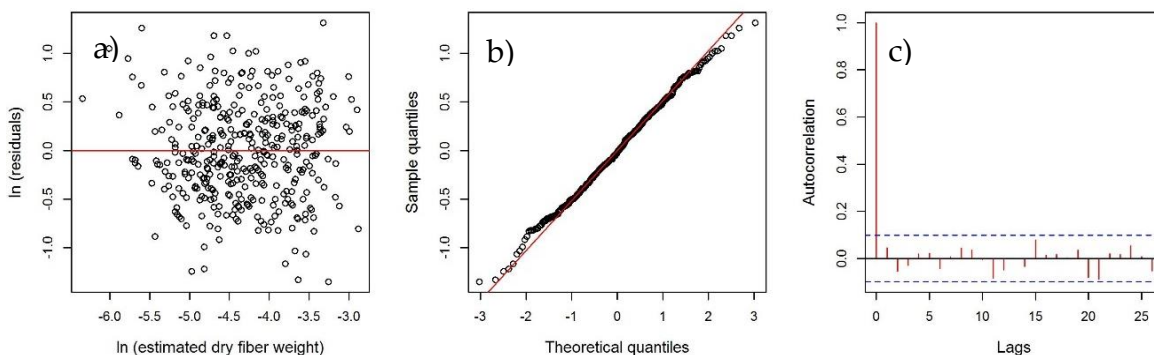
Donde:  $PSF$  = peso de fibra seca (kg),  $\ln Cd$  = logaritmo natural de diámetro de corona (cm),  $\exp$  = exponencial,  $F.C.$  = factor de corrección (1.127).





**Figura 3.** Fibra observada de *Agave lechuguilla* Torr. y estimada a través del modelo potencial-dummy, en dos zonas bioclimáticas de regiones áridas y semiáridas de México (a), y representación de observaciones influyentes del modelo (b).

Acorde al nivel de confianza establecido (95%), los residuales estudentizados del modelo para predecir fibra seca de *A. lechuguilla* presentan homocedasticidad ( $p = 0.1295$ ), son normales ( $p = 0.5343$ ) e independientes ( $p = 0.254$ ) (Figura 4a-c, respectivamente), esto asegura un modelo robusto, insesgado y eficiente (Montgomery *et al.*, 2012; Faraway, 2014; Fox, 2016).



**Figura 4.** Análisis de residuales del modelo para estimar peso de fibra seca de *Agave lechuguilla* Torr. al noreste de México. Homogeneidad de varianza (a), normalidad (b) e independencia (c).

### Capacidad predictiva del modelo para predecir fibra seca de *A. lechuguilla* Torr.

La validación del modelo, usando machine learning, con 100 datos independientes (datos que el modelo jamás había “visto”) demostró que el modelo muestra buena capacidad para predecir peso seco de fibra de *A. lechuguilla*. Los estadísticos de ajuste (todos en unidades logarítmicas) en los cuatro métodos de validación ( $R^2$ ) son obviamente más bajos (*e. g.*  $R^2$  promedio de 0.544) a los arrojados con el set de datos de entrenamiento ( $R^2 = 0.629$ ), pero muy similares entre sí (Tabla 5), esto garantiza que, al utilizar las ecuaciones generadas, para predecir fibra seca de *A. lechuguilla* con datos que el modelo jamás ha visto, se tendrán estimaciones adecuadas.

Según la literatura, emplear datos independientes para validar un modelo es una de las técnicas más recomendadas (Snee, 1977; Arboretti y Salmaso, 2003). Utilizando un set de datos independientes y aplicando el modelo generado con el set de entrenamiento, Velasco et al. (2009), evaluaron la capacidad predictiva para predecir fibra seca en esta especie, encontrando que los estadísticos de ajuste y predicciones son muy similares, lo que asegura su robustez al predecir con nuevos datos.

**Tabla 5.** Estadísticos de ajuste de la evaluación de la capa capacidad predictiva del modelo para predecir peso seco de fibra seca de *Agave lechuguilla* Torr., empleando machine learning, con una muestra independiente ( $n = 100$ ).

Validación	No. datos	n	$R^2$	RMSE	MAE
Ajuste del modelo	Entr	399	0.629	0.487	0.397
Leave One Out Cross-Validation (LOOCV)	Prueba	100	0.512	0.574	0.461
k-fold validation (Cross-Validation)	Prueba	100	0.534	0.564	0.462
Repeated K-fold (Cross-Validation)	Prueba	100	0.545	0.559	0.459
Bootstrap Validation	Prueba	100	0.583	0.561	0.453

$n$  = tamaño de muestra,  $R^2$ , = coeficiente de determinación, RMSE = raíz del error cuadrático medio, MAE = error medio absoluto, Entr = entrenamiento.

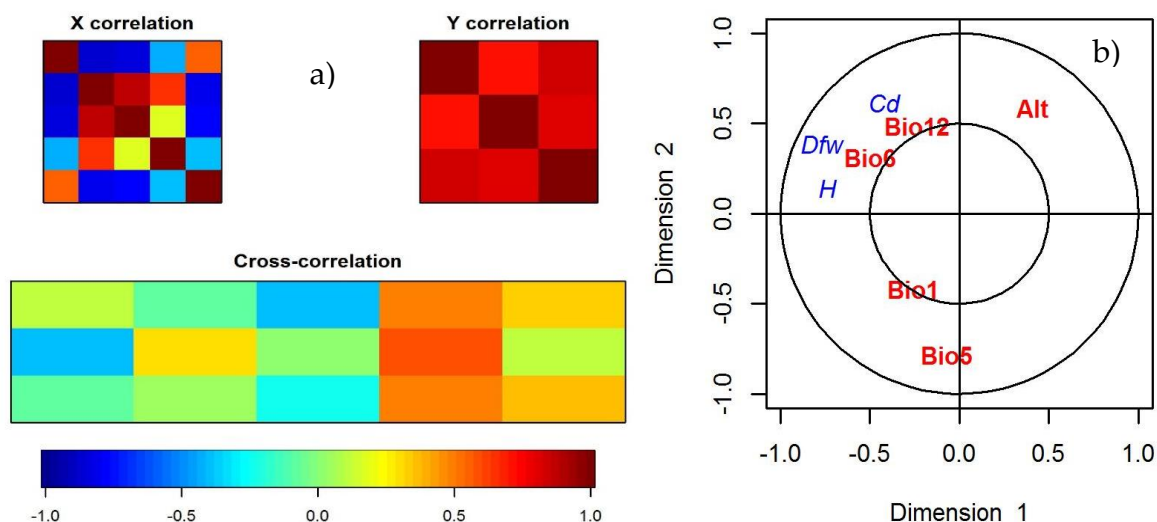
En estudios similares, realizados en la Zona 2 (Figura 1), con  $n = 287$ , Blando y Baca (2001) encontraron que *Cd* y número de hojas aprovechables, explican 76.09 % de la

fibra de *A. lechuguilla*. Empleando 95 individuos ( $H \geq 25$  cm), en la Zona 2 de este estudio, Pando et al. (2008) demostraron que  $H$  y  $Cd$  (como una sola variable independiente) explica 86.9 % del peso seco de fibra de *A. lechuguilla*, con un error de 5.041 g. En esta misma Zona ( $n = 240$ ) Velasco et al. (2009) encontraron que el diámetro del cogollo y la altura del mismo, explican 68 % de la fibra de *A. lechuguilla*. Como puede apreciarse, los estadísticos de ajuste de cada modelo, difieren a nuestro estudio y entre ellos, debido a tamaño de muestra ( $n$ ), método de muestreo (dirigido, selectivo, cuadrantes), selección de plantas (altura mínima), ubicación geográfica del estudio (zona 1 o 2), tipo de modelo (lineal, no lineal, simple o múltiple), tipo de predictor e incluso del cumplimiento o no de los supuestos del modelo.

Aunque en *A. lechuguilla* se han generado modelos lineales para estimar peso de fibra (Blando y Baca, 2001; Pando et al., 2008), la alometría demuestra cambios de dimensión relativa a las partes de tipo exponencial, e. g. en *Agave salmiana* Otto ex. Salm. ssp. crassispina (Trel.) Gentry, la biomasa aérea sigue este patrón (García et al. 2011), al igual que en *Agave salmiana* ssp. Crassispina (Martínez y Meyer, 1985), y como se ha demostrado para predecir biomasa aérea en *A. lechuguilla* (Flores et al., 2020) y fibra seca en este estudio, aplicando la ecuación potencial (Ec.1) en su forma lineal (Ec.2).

### **Relación canónica entre fibra seca de *A. lechuguilla* con variables bioclimáticas**

De acuerdo a la primera y segunda variable canónica, el grado de correlación que existe entre variables dendrométricas (matriz  $y$ ) es más alto ( $r = 0.9000$ ) que entre las variables bioclimáticas (matriz  $x$ ) la cual es de  $r = 0.7345$  (Figura 5a). La correlación en las primeras tres dimensiones es 0.8101, 0.5395 y 0.2510, respectivamente. La correlación canónica cruzada demuestra que la temperatura mínima del mes más frío, Bio 6 y precipitación anual, Bio 12 (Figura 5a-b y Tabla 6), se correlacionan positivamente con la producción de fibra seca de *A. lechuguilla* en estas zonas bioclimáticas. Acorde a tres modelos de circulación general: CCSM4, HadGEM2-AO y MIROC5 (Karger et al., 2017), a un escenario de emisiones de  $\text{CO}_2$  moderado ( $4.5 \text{ W/m}^2$ ) se espera que al 2050, Bio 6 disminuya  $0.56 \text{ }^\circ\text{C}$  en el área de distribución de *A. lechuguilla*, esto podría afectar la producción de fibra de la especie; mientras que en Bio 12, el decremento sería solo de 1.79 %.



**Figura 5.** Análisis de correlación canónica (CCA) entre variables bioclimáticas (matriz x) y variables dendrométricas (matriz y) de *Agave lechuguilla* Torr. (a) y su expresión en las primeras dos dimensiones de CCA.

**Tabla 6.** Análisis de correlación entre grupos de variables dendrométricas de *Agave lechuguilla* Torr. y bioclimáticas.

<b>r</b>	<b>pCd</b>	<b>H</b>	<b>PSF</b>	<b>Alt</b>	<b>Bio 1</b>	<b>Bio 5</b>	<b>Bio 6</b>	<b>Bio 12</b>
Cd	1	0.0006	0.00	0.6703	0.7888	0.0758	0.0292	0.1592
H	0.70	1	0.00	0.0940	0.2068	0.9075	0.0075	0.6685
PSF	0.84	0.80	1	0.7177	0.8874	0.2884	0.0258	0.1070
Alt	0.10	-0.38	-0.09	1	0.0000	0.0000	0.0557	0.0128
Bio 1	-0.06	0.29	0.03	-0.86	1	0.0000	0.0017	0.0000
Bio 5	-0.41	0.03	-0.25	-0.84	0.86	1	0.4456	0.0000
Bio 6	0.49	0.58	0.50	-0.43	0.66	0.18	1	0.0802
Bio 12	0.33	0.10	0.37	0.55	-0.81	-0.76	-0.40	1

Cd = diámetro de copa (cm), H = altura total (cm), Alt = elevación (m snm), PSF = peso seco de fibra (kg), Bio 1 = temperatura media anual (°C), Bio 5 = temperatura máxima del mes más cálido (°C), Bio 6 = temperatura mínima del mes más frío(°C), Bio 12 = precipitación media anual (mm).

## CONCLUSIONES

El análisis de componentes principales generó dos zonas bioclimáticas, la Zona 2 (ubicada hacia el este del área de estudio), representa mayor productividad de *A. lechuguilla* que la zona 1; esta zonificación es útil para enfocar esfuerzos específicos para el manejo y aprovechamiento de esta especie. El mejor predictor de fibra seca *A. lechuguilla* es diámetro de copa (*Cd*), y a través de regresión dummy fue posible establecer una ecuación para cada zona bioclimática (Figura 1b). La validación de esta ecuación demostró alta capacidad predictiva del modelo, además de ser fácil, sencilla, rápida y de bajo costo; útil en gran parte de la distribución conocida de la especie. Los resultados sugieren que por el decremento de Bio 6, en más de 0.5 °C al 2050, la especie podría ser vulnerable al cambio climático, probablemente ocurra reducción de la producción de fibra de *A. lechuguilla* de forma gradual en los próximos años.

## BIBLIOGRAFÍA

- Arboretti, G.R.; Salmaso, L. Model performance analysis and model validation in logistic regression. *Statistica* 2003, 43, 375-396.
- Berlanga, R.C.A.; González, L.L.A.; Franco, L.H. Metodología para la evaluación y manejo de lechuguilla en condiciones naturales. Campo Experimental "La Sauceda". CIRNE. INIFAP. Folleto Técnico Núm. 1. Saltillo, Coahuila, México. 1992; 22 p.
- Blando, N.J.L.; Baca, M.S. Determinación del potencial productivo de la lechuguilla (*Agave lechuguilla* Torr) en el municipio de San Juan de Guadalupe, Dgo. *Revista Chapingo Serie de Zonas Áridas* 2001, 2, 100-105.
- Blando, N.J.L.; Baca, M.S.; Martínez, S.M. Caracterización ecológica y socioeconómica de la lechuguilla (*Agave lechuguilla* Torrey). In: Martínez S., M. Ecología y usos de especies forestales de interés comercial de las zonas áridas de México. Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias. 2013; 230p.
- Briones, O.; Búrquez, A.; Martínez-Yrizar, A.; Pavón, N.; Perroni, Y. Biomasa y productividad en las zonas áridas mexicanas. *Madera y Bosques* 2018, 24, 1-19.
- Castillo, Q.D.; Berlanga, R.C.A.; Cano, P.A. Establecimiento y manejo de una plantación de lechuguilla. Campo Experimental, CIR-NORESTE. INIFAP. Desplegable para productores 2005; 2 p.
- Castillo, Q.D.; Berlanga, R.C.A.; Pando, M.M.; Cano, P.A. Regeneración del cogollo de *Agave lechuguilla* Torr. de cinco procedencias bajo cultivo. *Revista Ciencia Forestal en México* 2008, 33, 27-40.
- Castillo, Q.D.; Cano, P.A.; Berlanga, R.C.A. Establecimiento y aprovechamiento de lechuguilla (*Agave lechuguilla* Torr.). Comisión Nacional Forestal-Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias. México. Primera edición 2012; 33 p.
- Castillo, Q.D.; Mares, A.O.; Villavicencio, G.E.E. Lechuguilla (*Agave lechuguilla* Torr.), planta suculenta de importancia económica y social de las zonas áridas

y semiáridas de México. *Boletín de la Sociedad Latinoamericana y del Caribe de Cactáceas y otras Suculentas* 2011, 8, 6-9.

Castillo, Q.D.; Martínez, B.O.U.; Ávila, F.D.Y.; Castillo, R.F.; Sánchez, J.D.C. Identification of potential areas for establishment of plantations of *Agave lechuguilla* Torr. in Coahuila, Mexico. *Open Journal of Forestry* 2014, 4, 520-526.

Faraway, J.J. *Linear Models with R*, 2nd ed.; CRC press: Boca Raton, FL, USA 2014; 274 p.

Flores, D.M.P. *Bordeando el monte: La lechuguilla. Un recurso olvidado*. Secretaría del Medio Ambiente, 2018; 12 p.

Flores, H.C.deJ.; Méndez, G.J.; Sánchez, P.F.deJ.; Méndez, E.F.M.; López, D.Ó.M.; López, S.P.M. Allometric equations for predicting *Agave lechuguilla* Torr. aboveground biomass in México. *Forest* 2020, 784, 1-17.

Fox, J. *Applied Regression Analysis and Generalized Linear Models*, 3rd ed.; Sage Publications: Thousand Oaks, CA, USA. 2016; 791 p.

Francisco, M.C.; Cruz, M.F.; Sánchez, G.F.; Bautista, S.H. Propuesta para ubicar un centro de recolección de ixtle en el Municipio de Tantoyuca, Veracruz. *Revista Interdisciplinaria de Ingeniería Sustentable y Desarrollo Social* 2017, 3, 229-238.

Franco, T.; Hidalgo, R. Análisis estadístico de datos de caracterización morfológica de recursos fitogenéticos. *Boletín Técnico N° 8*, Instituto Internacional de Recursos Fitogenéticos (IPGRI), Cali, Colombia, 2003; 89 p.

García, E. *Modificaciones al sistema de clasificación climática de Köppen*. Quinta edición. Instituto de Geografía, Universidad Nacional Autónoma de México, 2004; 96 p.

García-Moya, E.; Romero-Manzanares, A.; Nobel, P.S. Highlights for *Agave* productivity. *Global Change Biology Bioenergy* 2011, 3, 4-14.

Déjean, S.; González, I. CCA: Canonical Correlation Analysis. R Package Version 1.2.1. 2021. Available online: <https://cran.r-project.org/web/packages/CCA/CCA.pdf> (accessed on: 27/09/2021).

- Hair, J.F.; Black, W.C.; Babin, B.J.; Anderson, R.E. Multivariate data analysis, Seventh Edition. Pearson Education Limited Harlow, 2014; 734 p.
- Houri, A.; Machaka, H.N. *Agave lechuguilla* as a potential biomass source in arid areas. *Journal of Forestry Development of Energy, Water and Environment Systems* 2016, 4, 89-93.
- Husson, F.; Josse, J.; Le, S.; Mazet, J. FactoMineR: Multivariate Exploratory Data Analysis and Data Mining. R Package Version 2.4. 2020; 100 p. Available online: <https://cran.r-project.org/web/packages/FactoMineR/FactoMineR.pdf> (accessed on: 27/09/2021)
- Karger, D.N.; Conrad, O.; Böhrner, J.; Kawohl, T.; Kreft, H.; Soria, A.R.; Zimmermann, N.E.; Linder, H.P.; Kessler, M. Climatologies at high resolution for the earth's land surface areas. *Sci. Data* 2017, 4, 1–20. Available online: <https://doi.org/10.1038/sdata.2017.122> (accessed on: 27/09/2021)
- Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; Cooper, T.; Mayer, Z.; Kenkel, B.; R Core Team; Benesty, M.; Lescarbeau, R.; Ziem, A.; Scrucca, L.; Tang, Y.; Candan, C.; Hunt, T. caret: Classification and Regression Training. R Package Version 6.0-88, 2021. Available online: <https://cran.r-project.org/web/packages/caret/caret.pdf> (accessed on: 27/09/2021).
- Marroquín, J.S.; Borja, L.G.; Velásquez, C.R.; de la Cruz, J.A. Estudio Ecológico dasonómico de las zonas áridas del norte de México. Publicación Especial No. 2, 2a ed. INIF, México, D.F., 1981; 166 p.
- Martínez, M.R.; Meyer, S.E. A demographic study of maguey verde (*Agave salmiana* ssp. *Crassispina*) under conditions of intensive utilization. *Desert Plants* 1985, 7, 61-66.
- Martínez, S.M.; Hermosillo, R.D.E.; Mojica, G.A.S.; Prieto, A.J.A. Potencial productivo y zonificación para el uso y manejo de especies forestales de zonas áridas. Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias. Campo Experimental La Campana. Chihuahua, México, 2015; 119p.
- Mayorga-Hernández, E.; Rössel-Kipping, D.; Ortiz-Laurel, H.; Quero-Carillo, A.R.; Amante-Orozco, A. Análisis comparativo en la calidad de fibra de *Agave*



- lechuguilla* Torr., procesada manual y mecánicamente. *Agrociencia* 2004, 38, 219-225.
- Montgomery, D.C.; Peck, E.A.; Vining, G.G. Introduction to Linear Regression Analysis, 4th ed.; John Wiley & Sons: Hoboken, NJ, USA. 2012; 688 p.
- Nobel, P.S.; Quero, E. Environmental productivity indices for a Chihuahuan Desert CAM plant, *Agave lechuguilla*. *Ecology* 1986, 67, 1-11.
- Pando, M.M.; Pulido, R.; Castillo, D.; Jurado, E.; Jiménez, J. Estimating fiber for lechuguilla (*Agave lechuguilla* Torr., Agavaceae), a traditional non-timber forest product in Mexico. *Forest Ecology and Management* 2008, 255, 3686–3690.
- Pando-Moreno, M.; Eufrazio, O.; Jurado, E.; Estrada, E. Post-Harvest growth of lechuguilla (*Agave lechuguilla* Torr., Agavaceae) in northeastern Mexico. *Economic Botany* 2004, 58, 78-82.
- Quero, E.; Nobel, P. S. Predictions of field productivity for *Agave lechuguilla*. *Journal of Applied Ecology* 1987, 24, 1053-1062.
- Picard, N.; Saint-André, L.; Henry, M. Manual de Construcción de Ecuaciones Alométricas para Estimar el Volumen y la Biomasa de los Árboles: Del Trabajo de Campo a la Predicción; FAO: Rome, Italy, 2012; 224 p.
- R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021. Available online: <https://www.r-project.org/> (accessed on: 25/08/2021)
- Reyes, A.J.A.; Aguirre, R.J.R.; Peña, V.C.B. Biología y aprovechamiento de *Agave lechuguilla* Torrey. *Boletín de la Sociedad Botánica de México* 2000, 67, 75-88.
- Rzedowski, J. Vegetación de México. Limusa. México, D.F, 1978; 432 p.
- SEMARNAT. Norma Oficial Mexicana NOM-008-SEMARNAT-1996, 1996; 10 p.
- Snee, R.D. Validation of regression models: methods and examples. *Technometrics* 1977, 19, 415-428.

- Sprugel, D.G. Correcting for bias in log-transformed allometric equations. *Ecology* 1983, 64, 209–210
- Vaissie, P.; Monge, A.; Husson, F. FactoShiny: Perform Factorial Analysis from 'FactoMineR' with a Shiny. R Package Version 2.4, 2021. Available online: <https://cran.r-project.org/web/packages/Factoshiny/Factoshiny.pdf> (accessed on: 27/09/2021)
- Valencia, C.C.M.; Sánchez, M.H.; Quiñones, V.J.J.; Martínez, R.J.J. Predicción de la biomasa aérea de cuatro especies arbustivas del semidesierto en la comarca lagunera. *Agrofaz* 2003, 3, 365-372.
- Velasco, B.E.; Arredondo, G.A.; Zamora, M.M.C.; Moreno, S.F. Modelos predictivos para la producción de productos forestales no maderables: Lechuguilla. Manual Técnico Num. 2. INIFAP, 2009; 60 p.
- Wood, A.G. A potential bias in log-transformed allometric equations. *Wader Study Group Bull* 1986, 47, 17-19.
- Zárate, L.A.; Franco, H.; Berlanga, C.A. Análisis dimensional en lechuguilla. In: Proceedings of the III Simposio Nacional sobre Ecología, Manejo y Domesticación de Plantas Útiles del Desierto. INIFAP, 1991. 13–14.

## ANEXOS

**Anexo 1.** Estadísticos de ajuste del modelo potencial con variables dummy en forma logarítmica, para establecer diferencias de variables dendrométricas de *Agave lechuguilla* Torr. y variables bioclimáticas entre zonas derivadas del PCA.

Var	Coficiente	Estimado	Std. Error	t-value	Pr(> t )
Cd	Intercept ( $\beta_0$ )	3.736	0.033	113.057	< 2e-16 ***
	Zone ( $\beta_1$ )	0.355	0.047	7.584	2.40e-13***
H	Intercept ( $\beta_0$ )	3.811	0.024	160.062	< 2e-16 ***
	Zone ( $\beta_1$ )	0.093	0.034	2.747	0.00629**
PSF	Intercept ( $\beta_0$ )	-4.499	0.056	-81.001	< 2e-16 ***
	Zone ( $\beta_1$ )	0.330	0.079	4.192	3.41e-05***
Bio1	Intercept ( $\beta_0$ )	5.256	0.007	744.087	< 2e-16 ***
	Zone ( $\beta_1$ )	-0.001	0.010	-0.055	0.956
Bio5	Intercept ( $\beta_0$ )	5.794	0.007	864.090	< 2e-16 ***
	Zone ( $\beta_1$ )	-0.102	0.009	-10.750	< 2e-16 ***
Bio6	Intercept ( $\beta_0$ )	3.633	0.032	114.852	< 2e-16 ***
	Zone ( $\beta_1$ )	0.354	0.045	7.914	2.51e-14***
Bio12	Intercept ( $\beta_0$ )	5.832	0.017	341.712	< 2e-16 ***
	Zone ( $\beta_1$ )	0.070	0.024	2.902	0.00391**
Alt	Intercept ( $\beta_0$ )	7.270	0.018	415.330	< 2e-16 ***
	Zone ( $\beta_1$ )	0.040	0.025	1.620	0.106

Cd = diámetro de copa (cm), H = altura total (cm), Alt = elevación (m snm), PSF = peso seco de fibra (kg), Bio1 = temperatura media anual (°C), Bio5 = temperatura máxima del mes más cálido (°C), Bio6 = temperatura mínima del mes más frío (°C), Bio12 = precipitación media anual (mm),  $\beta_0$  = intercepto,  $\beta_1$  = pendiente, Zona = variable dummy. Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1.

**Anexo 2.** Individuos evaluados y ejidos muestreados por municipio y estado en poblaciones de *Agave lechuguilla* Torr.

Estado	Municipio	No. Predios	No. Individuos
Chihuahua	Aldama	7	57
	Coyame	7	64
	Jiménez	1	15
	Ojinaga	1	10
Coahuila	Cuatro Ciénega	1	13
	General Cepeda	1	26
	Parras	3	17
	Ramos Arizpe	2	47
	San Pedro	2	7
	Viesca	1	9
San Luis Potosí	Catorce	2	27
	Cuidad del Maíz	1	17
	Guadalcázar	1	12
	Matehuala	2	30
	Venado	1	15
	Villa de la Paz	1	11
	Villa Hidalgo	1	12
Zacatecas	Concepción del Oro	4	51
	Mazapil	6	44
	Melchor Ocampo	1	15