

## **Conglomerados productivos de cacao en México**

SOUZA-PEÑALOSA, Patricia

P. Souza

Matemática aplicada independiente. Rio Magdalena 120-24. México, D.F. 01090.  
agromodelos@yahoo.com

D. Sepúlveda, R. Salazar, F. Pérez y J. Rocha (eds.) Ciencias Químicas y Matemáticas-©ECORFAN, Texcoco de Mora-México, 2015

## 10 Introducción

Uno de los cultivos precolombinos de mayor tradición en Mesoamérica es el del cacao. De acuerdo a las cifras oficiales publicadas al respecto por el Servicio de Información Agroalimentaria y Pesquera (SIAP, 2014), este cultivo en México actualmente se limita a los estados de Tabasco, Chiapas y Guerrero. Durante el año agrícola 2013, abarcó un total de 61 mil 319 hectáreas de superficie sembrada, cuyas cosechas se tradujeron en un valor total de mil dieciséis millones de pesos. Con ello, el cacao ocupó en 2013 el lugar número 15 y 25 en superficie sembrada y valor de producción, respectivamente, entre los 136 productos agrícolas perennes del país.

Cabe señalar que de las 27 mil 844 toneladas de cacao cosechadas en 2013 a nivel nacional, 26 mil 951 (96.79%) se produjeron en tierras de temporal, con una participación relativa de Tabasco, Chiapas y Guerrero igual a 18 mil 488 (68.60%), ocho mil 250 (30.61%) y 213 (0.79%) toneladas, respectivamente.

Con el fin de representar las condiciones actuales de dicha producción agrícola con la mayor estabilidad posible, en el presente estudio se promedian los resultados anuales del periodo 2010-2013 correspondientes a un total de 58 municipios productores de cacao en superficies de temporal de México. A este universo se aplica un filtro que preserva a 48 municipios considerados productores auténticos (Cuadro 7), haciendo a un lado a los diez municipios de producción marginal que generaron menos de dos años de cosechas durante dicho periodo (ver Cuadro 9).

Nueve de los diez municipios con producción marginal se encuentran en el estado de Oaxaca, que registra algunas hectáreas de cacao en superficie de temporal durante 2010 y 2011, pero sin cosecha alguna. Para 2012 y 2013, ya no se registran tierras oaxaqueñas que cultiven el cacao, pero el lector aún puede considerar la posibilidad de contar en Oaxaca con un campo potencial para el desarrollo sustentable del cacao.

El propósito más inmediato de la presente investigación es ofrecer una respuesta de calidad a las siguientes preguntas sobre el cacao producido en superficie de temporal. ¿Qué tipos de municipios productores de cacao existen actualmente? ¿Cuáles de sus rasgos son sobresalientes? ¿Dónde se encuentran? ¿Cuáles son las principales dificultades y áreas de oportunidad que comparten?

Para responder a estas preguntas, se emplea un modelo de conglomerados (mejor conocidos como clústers) regido por un algoritmo divisorio de tipo jerárquico (Johnson & Wichern, 2002) que termina por agrupar a los 48 municipios productores en seis categorías (ver Cuadro 7). El análisis correspondiente es multivariado, descriptivo y exploratorio, limitándose al tiempo presente 2010-2013 (Jain, 2010).

En todo momento, se busca presentar al lector un modelo estadístico accesible cuyos resultados sean susceptibles de apoyar a los tomadores de decisiones mediante conclusiones prácticas basadas en evidencia estadística acerca de las oportunidades y dificultades existentes en el cultivo del cacao. El propósito general de este esfuerzo es favorecer un incremento en la productividad y competitividad de la producción agrícola en México.

## 10.1 Materiales y métodos

### a) Valores económicos reales

En el presente estudio, toda variable que represente montos en pesos mexicanos se expresa en pesos constantes de enero de 2013 mediante el Índice Nacional de Precios al Consumidor (INPC) publicado por el Instituto Nacional de Estadística y Geografía (INEGI, 2014). Es decir: las series nominales de valor de producción, precio medio rural y rendimiento en pesos (SIAP, 2014) son convertidas en series reales de año base 2013. Con ello, se pretende eliminar los efectos de la inflación que están presentes en la información histórica.

Por ejemplo, un monto nominal de 100 pesos al iniciar el año 2012 tiene un valor real de 103.57 pesos al día primero de enero de 2013 porque el INPC de 2012 fue igual a 3.57%. Las series se convierten de esta manera a pesos constantes, aunque dichos cálculos sean aproximados y aunque el INPC no sea el único índice disponible para estos propósitos.

**Tabla 10** Tasa anual del Índice Nacional de Precios al Consumidor

Año	INPC
2010	4.40
2011	3.82
2012	3.57

Fuente: INEGI

### b) Insumos

El modelo de conglomerados se desarrolla a partir de los valores para el periodo 2010-2013 de las seis variables registradas por el SIAP en torno a la producción de temporal de cacao en los municipios de México.

Con el propósito de enriquecer la información disponible para la elaboración del modelo, a partir de dichas variables se definen dos más. La primera es el porcentaje de la superficie sembrada que resulta cosechada, expresado en valores de cero a cien. La segunda es el rendimiento en pesos contantes que expresa el valor económico real promedio que obtiene el productor por cada hectárea cosechada.

En la tabla 10.1 se precisa el concepto, algunas relaciones entre variables y la unidad de medida de cada una.

**Tabla 10.1** Ocho variables de producción agrícola

<b>Variable</b>	<b>Concepto</b>	<b>Unidad de medida</b>	<b>Cálculo</b>
SS	Superficie sembrada	Hectárea	
SC	Superficie cosechada	Hectárea	
<i>Volumen</i>	Volumen de producción	Tonelada	
<i>Rend</i>	Rendimiento	Toneladas por hectárea	$Volumen/SC$
<i>PMR</i>	Precio medio rural real <sup>1/</sup>	Pesos constantes por tonelada <sup>1/</sup>	$10^8 (Valor/Volumen)$
<i>Valor</i>	Valor real de producción <sup>1/</sup>	Millones de pesos constantes <sup>1/</sup>	
<b>Variables nuevas</b>			
% SC	Porcentaje de superficie sembrada que es cosechada	Porcentaje (base=100)	$100 (SC/SS)$
$\$Rend$	Rendimiento en valor real <sup>1/</sup>	Pesos constantes <sup>1/</sup>	$10^8 (Valor/SC)$
<b>Nota</b> <sup>1/</sup> Base = 2013 conforme al INPC			

Fuentes: Basado en información de INEGI y SIAP

Cuando el porcentaje de la superficie sembrada que resulta cosechada es inferior a cien para este cultivo perenne, es posible que además de la superficie de producción haya tierras nuevas y/o de desarrollo. También es posible que cualquiera de estas superficies esté sufriendo pérdidas totales. En cambio, cuando %SC es igual a cien, se sabe que la totalidad de la superficie sembrada se encuentra en producción y libre de pérdidas totales.

Por otro lado, el rendimiento en valor real de estas tierras permite medir económicamente el rendimiento logrado. Es decir, para cada hectárea cosechada no sólo se cuenta con el número promedio de toneladas obtenidas (*Rend*), sino con el monto real en pesos de 2013 que éstas representan en promedio para el agricultor ( $\$Rend$ ).

De este modo, en el modelo de conglomerados aplicado aquí se procesa un total de ocho variables reales. Cabe recalcar que ninguna de ellas es geográfica, por lo que los municipios productores de cacao en tierras de temporal se agrupan únicamente conforme a su desempeño productivo, sin importar su ubicación geográfica. En este sentido, se genera una clasificación transversal que generalmente no es regional.

### c) Municipios guerrerenses

La producción de cacao en tierras de temporal presenta rasgos dramáticamente distintos entre los diez municipios del estado de Guerrero y los 38 municipios de Tabasco y Chiapas. Esto se ilustra en la tabla 10.2, donde el rendimiento y el precio medio rural real registrados en Guerrero representan aproximadamente el 196% y 14%, respectivamente, del valor típico alcanzado por los municipios de Tabasco y Chiapas.

**Tabla 10.2** Comparativo de resultados por estado

Variable	Unidad	Municipio promedio 2010-2013	
		Tabasco y Chiapas	Guerrero
<i>Rend</i>	t / ha	0.45	0.88
<i>PMR</i>	pesos constantes / t	36,951	5,297

Fuente: Elaboración propia con datos del SIAP

Dichas diferencias relativas a la producción de cacao son tan marcadas y contradictorias que resulta preferible aislar a los municipios guerrerenses en dos conglomerados propios, destinando los cuatro conglomerados restantes a una clasificación mixta para Tabasco y Chiapas.

#### d) Variables independientes

El modelo de conglomerados es eficaz cuando se basa en variables enteramente independientes entre sí, por lo que es importante depurar los insumos anteriores hasta seleccionar el mayor número posible de variables independientes.

En la tabla 10.2 se aprecian relaciones exactas entre algunas variables, mientras que en las tablas 10.3 y 10.4 se puede afirmar que dos variables presentan una interdependencia clara siempre que el correspondiente valor absoluto del coeficiente de correlación lineal de Pearson alcance un valor mayor o igual a 70%.

En este sentido, las variables superficie sembrada (SS), superficie cosechada (SC), valor de producción (Valor) y rendimiento en valor real (\$Rend) se vuelven redundantes, ya que el volumen de producción (Volumen) representa a las primeras tres, mientras que el rendimiento (Rend) representa a la cuarta.

Entre las variables restantes no aparece otra redundancia visible, luego se considera que las cuatro son independientes entre sí. Finalmente, no hay forma de extraer cinco variables independientes de las ocho variables originales: el máximo es cuatro.

De este modo, las siguientes cuatro variables pueden considerarse adecuadas para desarrollar el modelo:

- Porcentaje de la superficie sembrada que resulta cosechada (% SC)
- Volumen de producción (Volumen)
- Rendimiento (Rend)
- Precio medio rural real (PMR).

**Tabla 10.3** Coeficientes de correlación lineal de Pearson (sólo Tabasco y Chiapas)

Variable	SS	SC	% SC	Volumen	Rend	PMR	\$Rend	Valor
SS	100.00							
SC	100.00	100.00						
% SC	23.11	23.19	100.00					
Volumen	99.78	99.78	23.33	100.00				
Rend	-19.45	-19.46	14.56	-17.41	100.00			
PMR	36.74	36.78	14.82	36.75	2.61	100.00		
\$Rend	-0.02	0.00	20.36	1.73	86.44	51.94	100.00	
Valor	99.76	99.76	22.88	99.93	-17.15	37.99	2.59	100.00

Fuente: Elaboración propia con datos del SIAP

**Tabla 10.4** Coeficientes de correlación lineal de Pearson (sólo Guerrero)

Variable	SS	SC	% SC	Volumen	Rend	PMR	\$Rend	Valor
SS	100.00							
SC	100.00	100.00						
% SC	19.47	20.07	100.00					
Volumen	99.82	99.82	19.40	100.00				
Rend	-22.15	-22.35	-35.89	-17.58	100.00			
PMR	33.44	33.19	-33.78	36.64	48.62	100.00		
\$Rend	-9.53	-9.76	-39.07	-4.80	97.31	67.41	100.00	
Valor	99.73	99.73	19.30	99.99	-16.59	37.49	-3.73	100.00

Fuente: Elaboración propia con datos del SIAP

## e) Hipercubo unitario

El modelo de conglomerados funciona con mayor eficacia cuando la distribución de las variables involucradas es lo más uniforme posible, para lo cual:

- % SC es topada inferiormente por el percentil P02 (sólo Tabasco y Chiapas)
- Volumen es sustituida por Vol, que es el logaritmo natural de su valor para todo municipio de Tabasco, Chiapas y Guerrero
- Rend es topada superiormente por el percentil P95 (sólo Tabasco y Chiapas)
- PMR es topada inferiormente por el percentil P05 (sólo Tabasco y Chiapas).

Finalmente, cada una de las cuatro variables es estandarizada para insertar al universo de 48 municipios en el cubo unitario del espacio euclidiano de dimensión cuatro. Es decir, a cada una de ellas se aplica un simple cambio de escala mediante el cual el valor mínimo y el máximo se transforman en los números cero y uno, respectivamente.

Con estas transformaciones, los 48 municipios por clasificar se encuentran distribuidos de una manera relativamente uniforme en el hipercubo unitario del espacio euclidiano de dimensión cuatro, cuyas coordenadas a partir de este momento se expresarán en porcentaje.

Así, la distancia entre dos municipios dados está dada por la distancia euclidiana entre los puntos correspondientes (Shmueli et al., 2010), que es la raíz cuadrada de la suma de las cuatro diferencias entre sus componentes elevadas al cuadrado.

De este modo, la distancia entre dos municipios sólo puede ser pequeña cuando el desempeño de ambos municipios es similar, simultáneamente, en porcentaje de tierras sembradas que resultan cosechadas, volumen de producción, rendimiento y precio medio rural real. De lo contrario, la distancia entre dos municipios productores de cacao resulta mediana o grande.

#### f) Semillas

Entre los 48 municipios por clasificar, se eligen seis que distan entre sí lo más posible: dos en Guerrero y cuatro en Tabasco y Chiapas.

Por ejemplo, para el caso de Tabasco y Chiapas se elige al municipio de máxima producción de cacao durante 2010-2013, que es Comalcalco, ubicado en el distrito tabasqueño de Cárdenas. Después se escoge al municipio tabasqueño o chiapaneco más distante de Comalcalco, que es el de Jalapa, ubicado en el distrito Villahermosa del estado de Tabasco. Este procedimiento continúa hasta obtener cuatro municipios que ofrecen diferencias grandes entre sí, conforme al algoritmo divisorio de tipo jerárquico que se describe en la literatura de los modelos de conglomerados (Johnson & Wichern, 2002). Lo mismo se hace con los municipios guerrerenses.

De esta manera, se eligen las siguientes seis semillas, cuyas características extremas son fundamentales para generar una segmentación exitosa.

**Tabla 10.5** Distancia entre las semillas

No.	Estado	Semilla	1	2	3	4	5	6
1	Tabasco	Comalcalco	0.00					
2	Chiapas	Tuxtla Chico	59.32	0.00				
3	Chiapas	Salto de Agua	95.89	44.28	0.00			
4	Tabasco	Jalapa	126.48	104.06	126.90	0.00		
5	Guerrero	Tecoanapa	40.08	33.22	62.66	128.30	0.00	
6	Guerrero	Cuajinicuilapa	144.43	87.59	74.69	103.45	119.54	0.00

Fuente: Elaboración propia con datos del SIAP

Así, dentro del hipercubo unitario, las seis semillas elegidas se encuentran relativamente apartadas unas de otras, como se documenta en el Tabla 10.5. Esto sirve para generar conglomerados de municipios que contrastan entre sí como productores de cacao.

#### g) Conglomerados

La clasificación definitiva es simple. En Tabasco y Chiapas, se tienen cuatro semillas y 38 municipios por clasificar: basta asignar cada uno a la semilla más cercana. En caso de empate, puede definirse una regla adicional (como integrarse al conglomerado más pequeño), pero para el caso del cacao esto no es necesario porque no se dan empates. En el caso de Guerrero se tienen dos semillas para clasificar diez municipios, entonces basta asignar cada municipio uno a la semilla más cercana. Nuevamente, la información sobre el cacao no genera empate alguno y la clasificación termina.

La idea central aquí es que los municipios que más se parecen a una determinada semilla se parecerán entre sí, de tal modo que cada conglomerado contendrá municipios similares. Al mismo tiempo, dado que las semillas se eligen lo más distintas posible entre sí, los municipios de diferentes conglomerados diferirán considerablemente entre sí.

Esta línea heurística es necesaria dada la imposibilidad práctica de calcular todas las segmentaciones posibles con el fin de elegir la óptima con respecto a una métrica dada (Jain, 2010; Johnson & Wichern, 2002).

De este modo, se obtienen seis conglomerados en total: uno alrededor de cada semilla. Algunos municipios pueden parecerse notablemente a la semilla del conglomerado, pero cabe recalcar que no siempre es así. En un conglomerado puede haber municipios que se parezcan poco a la semilla, pero el punto es que se parecen aún menos a las demás semillas.

## **10.2 Resultados**

### a) Conglomerados productivos

En la tabla 10.6 se sintetiza la clasificación de los principales 48 municipios productores de cacao en tierras de temporal. Para cada conglomerado, se listan los municipios que lo componen, de menor a mayor distancia de la semilla. Esto es: cada lista comienza en la propia semilla, seguida del municipio que más se parece a ella y terminando con el municipio que menos se le parece. En algunos casos, como los municipios de Xochistlahuaca en Guerrero (perteneciente al Conglomerado No. 5) y Francisco León en Chiapas (del Conglomerado No. 3), el lector puede decidir que un municipio difiere tanto del resto que se podría considerar aparte.



Tabla 10.6 Conglomerados

Conglomerado	No. municipios	Semilla	Municipio	Estado	Distancia a la semilla	
1	10	Comalcalco	Comalcalco	Tabasco	0.00	
			Cunduacán		8.54	
			Cárdenas		9.10	
			Huimanguillo		22.83	
			Pichucalco	Chiapas	23.59	
			Jalpa de Méndez	Tabasco	25.87	
			Paraíso		32.07	
			Ostuacán	Chiapas	33.01	
			Juárez		37.25	
			Tecpatán		48.66	
2	16	Tuxtla Chico	Tuxtla Chico	Chiapas	0.00	
			Tuzantán		14.62	
			Acapetahua		14.95	
			Huixtla		15.04	
			Huehuetán		15.62	
			Villa Comaltitlán		16.83	
			Mapastepec		19.35	
			Acacoyagua		23.90	
			Escuintla		31.71	
			Tumbalá		32.56	
			Sunuapa		46.08	
			Ixtacomitán		49.44	
			Nacajuca		Tabasco	52.94
			Ixtapangajoya		Chiapas	54.39
			Reforma	78.63		
			Solosuchiapa	84.11		

3	7	Salto de Agua	Salto de Agua	Chiapas	0.00
			Mazatán		29.67
			Metapa		42.06
			Frontera Hidalgo		47.18
			Maravilla Tenejapa		62.55
			Palenque		69.75
			Francisco León		110.40
4	5	Jalapa	Jalapa	Tabasco	0.00
			Tacotalpa		51.95
			Teapa		56.59
			Centro		59.56
		Suchiate	Chiapas	69.08	
5	6	Tecoanapa	Tecoanapa	Guerrero	0.00
			Ayutla de Los Libres		13.16
			Ometepec		38.52
			Florencio Villarreal		56.80
			San Marcos		85.90
			Xochistlahuaca		130.43
6	4	Cuajinicuilapa	Cuajinicuilapa	Guerrero	0.00
			San Luis Acatlán		55.15
			Marquelia		70.91
			Igualapa		84.03

Fuente: Elaboración propia con datos del SIAP

En la segunda columna de la tabla 10.6 se aprecia que cada uno de los conglomerados obtenidos posee un número distinto de elementos: siendo el Conglomerado No. 2 el más numeroso de todos, con 16 municipios.

#### b) Características

La tabla 10.7 revela diferencias claras entre los seis conglomerados. Resulta útil considerarlas con respecto al valor promedio o típico para los 48 municipios, representado en la última columna.

**Tabla 10.7** Valores promedio por conglomerado

Variable	Conglomerados 2010-2013						Municipio típico
	1	2	3	4	5	6	
SS	4,752.40	583.76	233.96	150.89	37.00	2.94	1,239.38
SC	4,743.50	581.70	223.79	150.83	36.92	2.94	1,235.34
% SC	99.60	98.66	94.01	98.72	99.24	100.00	98.37
Volumen	1,966.81	247.28	104.94	101.43	32.16	2.53	522.28
Rend	0.41	0.42	0.45	0.69	0.89	0.88	0.55
PMR real	42,026.81	37,350.77	29,061.69	41,036.38	5,354.72	5,209.47	30,822.09
\$Rend real	17,184.79	15,458.72	12,936.57	28,592.83	4,757.74	4,570.24	14,573.65
Valor real	82.87	8.96	2.94	4.32	0.17	0.01	21.15

Fuente: Elaboración propia con datos del SIAP

A pesar de que la segmentación obtenida no incluyó coordenada geográfica alguna, el cultivo del cacao presenta resultados regionales particularmente interesantes, ya que la ubicación geográfica de los seis conglomerados productivos se traduce en zonas compuestas por municipios que en su mayoría son contiguos.

En lo relativo al sexto conglomerado, cabe señalar que la ubicación específica de Marquelia es un municipio guerrerense de creación relativamente reciente (año 2002), por lo que su localización exacta puede resultar menos inmediata para el lector.

### c) Producción marginal

Para el lector puede ser útil conocer los diez municipios cuya producción de cacao fue considerada marginal en el presente análisis. En la tabla se reproduce la lista en orden alfabético y por estado: son los municipios que entre 2010 y 2013 registraron cacao sembrado en alguna superficie de temporal, pero obtuvieron de él un solo año de cosecha o ninguno.

**Tabla 10.8** Municipios de producción marginal 2010-2013

Municipio	Estado	Municipio	Estado
Marqués de Comillas	Chiapas	San Juan Comaltepec	Oaxaca
Ayotzintepec	Oaxaca	San Juan Cotzocón	
Cosolapa		San Juan Petlapa	
San Felipe Usila		Santa María Jacatepec	
San Juan Bautista Valle Nacional		Santiago Choápam	

Fuente: Elaboración propia con datos del SIAP

### 10.3 Conclusiones

La producción de cacao durante 2010-2013 se realizó principalmente en los estados de Tabasco, Chiapas y Guerrero, mientras que se tiene evidencia de una producción marginal ubicada casi enteramente en Oaxaca. Por medio de un algoritmo divisorio de tipo jerárquico, el desempeño de los principales 48 municipios productores de cacao se agrupó en seis conglomerados que revelan importantes semejanzas internas y diferencias externas.

Aunque cada conglomerado pertenece a uno o dos estados, cabe recalcar que la realidad de la producción de cacao rebasa el tradicional desglose estatal: tanto en Tabasco como en Chiapas y Guerrero los resultados anteriores demuestran que existe más de un modo de producir, con logros y dificultades que comparten los municipios que conforman cada conglomerado. Ésta es la razón de ser del presente análisis (Shmueli et al., 2010).

#### a) Retos y oportunidades

El punto esencial de la presente investigación es observar los distintos retos y oportunidades por conglomerado que enfrenta el cultivo del cacao en tierras de temporal de México. Cabe recalcar que algunos puntos podrían enriquecerse considerablemente con información sobre la variedad de cacao cultivada, el tipo de suelo empleado, la calidad del producto obtenido, la forma de comercialización y el destino final, incluyendo su posible exportación.

Como el lector podrá apreciar, el primer conglomerado abarca municipios contiguos del oeste de Tabasco y noroeste de Chiapas. Presenta rasgos notables: su volumen de producción es el más alto de todos y el precio medio rural que obtienen sus productores también es el más elevado. Dado que en sus municipios típicamente se cosecha el 99.60% de la superficie sembrada, se deduce que sus tierras se encuentran esencialmente en producción y sin registrar pérdidas totales. Si bien su precio de comercialización sugiere que la calidad de su producto puede ser la más alta, este grupo de municipios presenta el rendimiento más bajo de todos, con un promedio de 0.41 toneladas por hectárea cosechada. Mientras no pierda la calidad que lo distingue, este conglomerado podría beneficiarse de la experiencia de los demás municipios para mejorar sus niveles de rendimiento. En este sentido, las diferencias con los municipios de Guerrero pueden ser demasiado grandes, pero el cuarto conglomerado se encuentra en tierras cercanas: podría aprender de él técnicas viables para elevar el rendimiento.

El segundo conglomerado ocupa el segundo lugar en volumen de producción, pero no genera más que la octava parte de lo que cosecha el primero. Se parece bastante al primer conglomerado porque cosecha el 98.66% de la superficie sembrada y registra el segundo rendimiento más bajo de todos con 0.42 toneladas por hectárea cosechada. Pero su precio medio rural real es el tercero de todos. Para mejorar su desempeño, tendría sentido que analizara cómo el primer conglomerado posee mayor superficie sembrada y cómo los conglomerados primero y cuarto obtienen mejor precio real de comercialización. De los conglomerados cuatro, cinco y seis, también podría aprender maneras de aumentar su rendimiento sin perder la calidad actual de su producto.

Con un volumen de producción similar, siguen los conglomerados tercero y cuarto, que por lo demás difieren mucho entre sí. El tercero parece ser el único que posiblemente incluya tierras nuevas y en desarrollo (o pérdidas totales), ya que en promedio cosecha el 94.01% de las hectáreas sembradas, mientras que el cuarto con un valor de 98.72% parece excluir estas posibilidades.

Con un precio medio rural real y un rendimiento 53.33% y 41.20% mayores, respectivamente, el cuarto conglomerado presenta grandes ventajas sobre el tercero.

Finalmente, los municipios productores guerrerenses que fueron aislados desde el inicio presentan los contrastes más notables con la producción de cacao en Tabasco y Chiapas. Con mucho, su volumen de producción y su precio medio rural real son los más bajos, mientras que su nivel de rendimiento (entre 0.88 y 0.89 toneladas por hectárea) es entre 28% y 117% mayor que el de los demás. En síntesis, los conglomerados quinto y sexto parecen ser la antítesis del primero, excepto en su porcentaje casi igual a 100% de tierras en producción que no parecen registrar hectáreas en desarrollo y no parecen sufrir pérdidas totales.

#### b) Otras consideraciones

El presente artículo obedece a un esfuerzo por aplicar ciencias básicas para generar resultados en materia de productividad y sustentabilidad que puedan apoyar a los tomadores de decisiones de la agricultura en México.

Concretamente, lo que interesa aquí es promover el uso del algoritmo divisorio de carácter jerárquico para generar conglomerados productivos. A modo de ejemplo, se aplicó dicha técnica de la estadística multivariada al ejemplo del cacao cultivado en México de 2010 a 2013. Naturalmente, sobre el cacao habría mucho más que decir, por ejemplo en cuanto a las variedades que se cultivan de él en Tabasco, Chiapas y Guerrero, así como acerca de la calidad del producto y las circunstancias de comercialización de cada municipio. Esto requiere bases de datos adicionales que se podrían integrar al presente estudio para enriquecerlo.

Cabe recalcar que los conglomerados productivos propuestos aquí se generaron sin tomar en cuenta la ubicación geográfica de los municipios por clasificar, con lo cual los conglomerados pudieron haberse dispersado libremente (como ocurre con otros cultivos perennes, como el nopal verdura en tierras de riego). Sin embargo, la naturaleza de los segmentos obtenidos demuestra que la realidad productiva del cacao está profundamente ligada a la realidad regional, con matices relevantes por explorar en cada conglomerado.

## 10.4 Referencias

Índices de Precios, INPC Índice General. Instituto Nacional de Estadística y Geografía. Consultado el 8 de octubre de 2014 en: [www.inegi.org.mx](http://www.inegi.org.mx)

Jain, A., "Data clustering: 50 years beyond K-means", *Pattern Recognition Letters* 31 (2010), pp. 651-666

Johnson, R. y Wichern, D., *Applied Multivariate Statistics*, New Jersey EUA, Prentice Hall, 2002, 767 pp.

Producción agropecuaria, producción anual. Servicio de Información Agroalimentaria y Pesquera. Consultado el 8 de octubre de 2014 en: [www.siap.gob.mx](http://www.siap.gob.mx)

Shmueli, G., Patel, N. y Bruce, P., *Data Mining for Business Intelligence*, New Jersey EUA, Wiley, 2010, 404 pp.