

Índice de Actividad Económica Metropolitano

DS @ OPI

CDMX a 5 de enero de 17

Introducción

El propósito de este proyecto es generar un índice que mida los niveles y el crecimiento de la actividad económica en las zonas metropolitanas y que tenga las siguientes características:

- Desagregación. Más allá de los estados, el índice metropolitano IMCO-OPI se enfocará en evaluar las ciudades y zonas metropolitanas más importantes del país.
- Frecuencia. El índice metropolitano utilizará datos con frecuencia trimestral, por lo que se publicará oportunamente.
- Accesibilidad. Los insumos no dependen de las cuentas nacionales, por el contrario se utilizarán datos transaccionales y satelitales en la elaboración, que están disponibles públicamente.
- Participación. Además de publicarse regularmente, el índice metropolitano hace uso de tecnologías y metodologías modernas, para ser accesible al público en general. En un repositorio de control de versiones¹, se harán accesibles las fuentes de los datos así como el código utilizado para procesarlo, de principio a fin. De forma que el usuario interesado pueda replicar el proceso, hacer modificaciones para uso propio, o incluso proponer ajustes a la metodología.

Descripción de los datos

Se utilizaron datos que corresponden a tres variables principales: producto o actividad económica, mediciones de luminosidad nocturna y transacciones de la CNBV. Los primeros –PIBE e ITAEE– son generados por el INEGI para las entidades federativas, y constituyen la guía para el índice que generamos en este proyecto. Puesto que el objetivo presente es generar un indicador en un nuevo nivel de agregación, los relacionamos con las variables de luminosidad y CNBV, que se calculan para las entidades federativas y las nuevas desagregaciones.

A continuación se describen las bases de agregación que utilizaremos en el proceso de modelado del índice metropolitano IMCO-OPI.

Agregación sociopolítica

Se utilizan cuatro niveles de agregación; escribimos como \mathcal{L} , \mathcal{M} , \mathcal{C} , \mathcal{E} y que corresponden a localidades urbanas, municipios, ciudades -o áreas metropolitanas- y estados.

La forma en que se utiliza cada nivel es la siguiente:

- Estados: los índices existentes más usados se publican a nivel estado y nacional. A medida que generamos un nuevo índice, parte del proceso de validación consiste en compararlo con los existentes, y ello se hace a este nivel de agregación. Además para propósitos de comunicación se asume que el público general tiene mayor conocimiento de las entidades federativas que de las regiones desagregadas como ciudades o municipios; de forma que la información de los estados tiene más fluidez en este sentido.
- Ciudades, o zonas metropolitanas: el nivel de ciudad es el objetivo del índice que se desarrolla en este proyecto. En términos territoriales es intermedio entre municipios y estados, lo cual implica que se pueden obtener datos de los niveles correspondientes, ya sea agregándolos o heredándolos. Por ejemplo, los

¹www.github.com



niveles de luminosidad o de área se obtienen a nivel municipal, y después se suman para tener los niveles de las ciudades; por otro lado, siguiendo las medidas estatales, se hacen modelos que relacionan los comportamientos del PIBE y de luminosidad, mismos que se heredan a las ciudades correspondientes.

- Municipios: los municipios constituyen la división administrativa más básica en este proyecto que cubre el territorio nacional. Tanto los datos de luminosidad como los de la CNBV se pueden obtener a este nivel de agregación, y como se mencionó anteriormente- fueron agregados para la realización de diferentes análisis.
- Localidades urbanas: si bien los municipios constituyen la base de este índice, se utilizan estas localidades para definir el *soporte* del mismo. Específicamente para las mediciones de luminosidad que explicaremos más adelante, las concentramos en dichas zonas urbanas.

Se utilizó el Marco Geoestadístico Nacional (INEGI, 2014) para identificar las regiones correspondientes. A partir de los municipios se agregan tanto en ciudades como en estados para obtener las variables en cada nivel.

Producto y actividad económica

El PIBE es el Producto Interno Bruto por Entidad Federativa; es el indicador oficial de actividad económica de los estados y se publica anualmente por el INEGI. En diferentes periodos que duran alrededor de una década, el PIBE conlleva distintas metodologías que incorporan reglamentación novedosa del momento. El periodo más reciente –de 2008 a 2014– sigue los lineamientos del Sistema de Clasificación Industrial de América del Norte. Se puede obtener más información en el sitio web del INEGI.

Sin embargo el cálculo del PIBE conlleva un tiempo de rezago; éste incluye desde la recaudación de las cuentas de las diferentes secretarías, hasta la revisión de las metodologías con estándares internacionales. Para una publicación más reciente y frecuente se cuenta con el Índice Trimestral de Actividad Económica Estatal, ITAEE. La serie de este índice se extiende desde 2003 hasta 2016 y tiene periodicidad trimestral –como su nombre lo indica–. De acuerdo a la documentación del ITAEE debe considerarse como un adelanto del PIBE pues incorpora parte de la metodología correspondiente, aunque no coincide del todo con los cálculos anuales debido principalmente a la calendarización de la actividad primaria².

Se utilizan los datos del PIBE en un periodo de tiempo que fijamos en $t_0 = 2014$ y que denotamos como π_{e,t_0} para cada estado $e \in \mathcal{E}$. Las series del ITAEE la denotamos como $\iota_{e,t}$ donde $t \in \mathcal{T}$ se extiende desde 2011 a 2016. De acuerdo a estos índices se estimaran los niveles y crecimientos de los estados que después extenderemos a las zonas metropolitanas $c \in \mathcal{C}$.

Luminosidad nocturna

Para medir la luminosidad nocturna se usaron imágenes del sitio de la Administración Americana Oceánica y Atmosférica, NOAA por sus siglas en inglés. Los archivos tienen fotos satelitales del globo terráqueo, e información específica para seleccionar el territorio de interés de la República Mexicana. Las fotos fueron sometidas previamente a procesos en los cuales se limpian de efectos que distorsionan la información. Por ejemplo, la presencia de nubes en alguna región obstruye la luz cuando se toma la foto satelital y la haría parecer menos luminosa; los archivos comprenden periodos de un mes, de donde se puede distinguir cuáles de las zonas fueron influenciadas por nubes y remover dicho efecto.

Las fotos que se utilizaron son generadas con tecnología reciente que se llama *Suite de Radiometría de Imágenes Visibles Infrarojas*, VIIRS. Estas imágenes tienen una resolución de 0.55 km^2 y usan unidades de radiación $\ell \sim \frac{\text{Watt}}{\text{cm}^2 \text{sr}}$, es decir potencia entre área esférica. Llamaremos a estas mediciones como *luminosidad* y la denotamos como ℓ , λ o Λ , dependiendo del nivel y agregación.

Estas fuentes satelitales proporcionan mediciones geográficas x_i , y_i y de luminosidad ℓ_i que indexamos con $i \in \mathcal{I}$; las mediciones corresponden a los píxeles de la cuadrícula o ráster. Relacionándolo con las regiones

²Fuente: http://www.inegi.org.mx/est/contenidos/proyectos/scn/c_anuales/pib_ef/presentacion.aspx



Figura 1: Foto satelital

$R \in \mathcal{M}, \mathcal{C}, \mathcal{E}$ se tiene que el área correspondiente es proporcional al número de píxeles contenidos en ellas

$$A(R) \propto \sum_{i \in \mathcal{I}} \chi_R(x_i, y_i).$$

De forma similar calculamos la luminosidad total y media de las regiones como

$$\Lambda(R) = \sum_{i \in \mathcal{I}} \ell_i \chi_R(x_i, y_i) \quad y \quad \lambda(R) = \frac{\Lambda(R)}{A(R)},$$

donde $\chi_R(x_i, y_i)$ es la indicadora de cada pixel i en cada región R de acuerdo a sus coordenadas (x_i, y_i) .

Mas aún dado que el índice medirá la actividad económica en las zonas metropolitanas, se considera la luminosidad urbana como la restricción de las luminosidades a las localidades urbanas. Tomamos $U = \bigcup_{\mathcal{C}} L$ como la unión de todas las localidades urbanas, e indicamos mediante el superíndice $(\cdot)^U$ las restricciones correspondientes

$$A^U(R) = A(R \cap U), \quad \Lambda^U(R) = \Lambda(R \cap U), \quad \lambda^U(R) = \lambda(R \cap U).$$

Esta definición de localidad se aplica computacionalmente a partir de los *shapefiles* de localidades urbanas, cada cual pertenecen en sí a un municipio, y por ende algunas a zonas metropolitanas y estados; entonces las intersecciones se calculan en automático cuando se utiliza este marco de localidades. Cabe mencionar que dichos *shapefiles* provienen del Marco Geoestadístico Nacional y se procesan con software especializado³ de geolocalización.

Finalmente se impone una segunda restricción a las zonas urbanas por considerar. Esta restricción se basa en el valor de la luminosidad en la fotografía, para que no exceda a $175 \frac{\text{Watt}}{\text{cm}^2 \text{gr}}$. El límite se obtuvo tras una

³Se utiliza el sistema de información geográfica QGIS.



inspección detallada, después de encontrar regiones de pocos pixeles con luminosidades sumamente altas y que no corresponde a la actividad. Más específicamente las zonas se identificaron en sitios de actividad petrolera, como refinerías o extractoras, que inflarían la luminosidad y por ende la actividad económica. El tope de 175 se toma relativo al máximo de la Zona del Valle de México.

Manteniendo la notación simple, redefinimos la zona urbana U como no petrolera donde solamente se consideran los pixeles con $\ell \leq 175$. La luminosidad correspondiente sigue esta consideración.

Con estas medidas se calcularán los niveles del índice IMCO-OPI en las zonas metropolitanas.

Comisión Nacional Bancaria y de Valores

La CNBV proporciona datos mensuales de transacciones y otras variables bancarias a nivel de localidad y por institución. Se utilizaron series de transacciones en cajeros automáticos que denotamos como μ . Las series más básicas llevan subíndices $\mu_{m,t,b}$ y corresponden al municipio $m \in \mathcal{M}$, periodo $t \in \mathcal{T}$ e institución bancaria $b \in \mathcal{B}$.

Las agrupaciones en cada índice se obtienen como sumas de los datos individuales. En primer lugar se agregan las operaciones de cada trimestre (con abuso de notación se utiliza el mismo subíndice t); en segundo lugar se distinguieron instituciones bancarias⁴ $b \in \mathcal{B}$ con observaciones volátiles y que tienen alta representación en el total. Entonces consideramos un subconjunto de instituciones $\tilde{\mathcal{B}} \subset \mathcal{B}$ con respecto al cual calculamos las series

$$\mu_{m,t} = \sum_{b \in \tilde{\mathcal{B}}} \mu_{m,t,b};$$

a partir de ellas se obtienen las series para las demás ciudades o estados $R \in \mathcal{C}, \mathcal{E}$

$$\mu_{R,t} = \sum_{m \in R} \mu_{m,t}.$$

Por último, las series de transacciones de cajeros aproximan a la actividad económica a través de los crecimientos proporcionales. Para los periodos trimestral y anual –digamos $p = 1, 4$ – denotamos dichos crecimientos como

$$\mu_{R,t}^{\Delta p} = \frac{\mu_{R,t}}{\mu_{R,t-p}} - 1.$$

A propósito de los índices en la notación, nótese que los subíndices R, t corresponden a la medición en el espacio y tiempo, mientras que los superíndices Δ, p representan cuestiones estructurales de los datos para el modelo.

Modelado

La estrategia que sigue el desarrollo del índice metropolitano IMCO-OPI consiste en entrenar un modelo a nivel estatal, para después replicarlo con los datos correspondientes de las zonas metropolitanas. Denotamos este índice de actividad económica como $\varrho_t(R)$ y ajustamos a diferentes regiones $R \in \mathcal{M}, \mathcal{C}, \mathcal{E}$, hacemos distinción con el PIBE que mide el producto de los estados $e \in \mathcal{E}$ y se escribe como $\pi_{e,t}$.

A su vez el modelo se separa en dos partes que corresponden al nivel y al crecimiento. El cálculo de niveles se basa en los datos de luminosidad urbana, mientras que el de crecimiento en los datos de cajeros automáticos. Esta división apoya los siguientes puntos considerables:

- La luminosidad refleja el tamaño de las economías. Artículos en la literatura como [1], [2] han apoyado esta tesis, y examinan características donde la relación se hace más o menos robusta.

⁴Las observaciones de BBVA Bancomer y Santander fueron removidas de las sumas debido a volatilidad desproporcionada.

- Más específicamente, se asume que la actividad económica proviene de las localidades urbanas de las regiones. Esto se debe a que hay zonas oscuras que por su gran extensión acumulan luminosidad, pero que no es acorde con la producción debido a que está despoblada.
- La actividad económica de las ciudades está ligada a las transacciones monetarias que hacen los consumidores, y más específicamente a las disposiciones en efectivo de las mismas. Este punto concuerda con el cálculo del ITAEE que incluye dichas transacciones, a la vez que se permite una agregación refinada.

La siguiente imagen resume la información de luminosidad y da muestra de algunos elementos del modelo.



Figura 2: Luminosidad municipal

La intensidad del color de los municipios está relacionada con su luminosidad logarítmica. Además se indican las fronteras de las entidades federativas, a partir de la cual se estima el modelo.

En los siguientes apartados describimos los detalles técnicos.

Niveles de actividad económica

Este cálculo utiliza los datos de PIB de los estados y de luminosidad urbana de los municipios. Para un tiempo inicial $t_0 = 2014$, se toman los datos correspondientes y se distribuye el PIB con respecto a la luminosidad. Para las ciudades o zonas metropolitanas, se agregan los municipios correspondientes.

Para cada municipio $m \in \mathcal{M}$ se calcula su actividad económica como

$$\varrho_{t_0}(m) = \frac{\Lambda^U(m)}{\Lambda^U(e)} \pi_{e,t_0}$$

donde $e = e(m)$ es el estado al que pertenece el municipio m y la luminosidad $\Lambda^U(\cdot)$ es la correspondiente a la luminosidad urbana no petrolera —con $\ell \leq 175$ — que se introdujo anteriormente.



La actividad económica de las zonas metropolitanas $c \in \mathcal{C}$, se estima sumando sobre sus municipios

$$\varrho_{t_0}(c) = \sum_{m \in c} \varrho_{t_0}(m).$$

En el caso de las metrópolis que se concentran en un sólo estado, esta estimación es igual a la ponderación por luminosidad del PIBE correspondiente

$$\varrho_{t_0}(c) = \frac{\Lambda^U(c)}{\Lambda^U(e)} \pi_{e,t_0}.$$

Cabe mencionar que éstas no son todas las zonas metropolitanas, pues hay algunas que se dividen en dos o más entidades federativas. Éstas son pocas por lo que podemos mencionarlas: La Laguna, La Piedad-Pénjamo, Puebla-Tlaxcala, Puerto Vallarta y el Valle de México.

Si bien la representación en términos de sus municipios permite estimar la actividad económica para estas zonas metropolitanas, también controlamos por dicha separación a la vez que las representamos con sus partes correspondientes, digamos $c = c_1 \cup c_2 (\cup c_3 \text{ para el Valle de México})$ donde cada c_i pertenece a un estado diferente.

Después de considerar el nivel de actividad económica en t_0 , se modela los niveles subsecuentes a partir del modelo de crecimiento de las transacciones de cajeros automáticos que explicamos a continuación.

Crecimiento

El crecimiento de actividad económica está ligado a los datos de transacciones de cajeros automáticos. Esta consideración se puede justificar tanto teóricamente como en la práctica si utilizamos el ITAEE como el indicador base.

En la teoría, las transacciones de cajeros automáticos son una medida simplificada del nivel de consumo en la sociedad, y éste tiene una relación dinámica con la producción. En la práctica, el cálculo del ITAEE utiliza datos de la banca comercial provistos por la CNBV⁵. En la gráfica 3 se ve la similitud de las transacciones con la actividad económica que representa el ITAEE en los estados.

Cuantificamos esta similitud con el siguiente modelo

$$\iota_{e,t}^{\Delta,1} \sim \alpha_0^e + \alpha_1^e \mu_{e,t}^{\Delta,1} + \alpha_4 \mu_{e,t}^{\Delta,4}$$

donde análogamente escribimos $(\cdot)^{\Delta,p}$ para los crecimientos proporcionales trimestral y anual del ITAEE y las transacciones de cajeros. Los coeficientes $\alpha_{(\cdot)}^e$ con superíndice están asociados a cada estado.

Los resultados de la estimación los expresamos con $\hat{\alpha}$'s como coeficientes y ε 's como error

$$\iota_{e,t}^{\Delta,1} = \hat{\alpha}_0^e + \hat{\alpha}_1^e \mu_{e,t}^{\Delta,1} + \hat{\alpha}_4 \mu_{e,t}^{\Delta,4} + \varepsilon_{e,t}.$$

Estos modelos de crecimientos son los que aplicaremos a las zonas metropolitanas, para después estimar los índices de las ciudades $\varrho_t(c)$.

Integración de crecimiento

Después de la estimación del crecimiento de los estados, se sigue heredar la relación a las zonas metropolitanas que pertenecen a dichos estados. Cambiando los datos observados en estados e por ciudades c pero manteniendo los coeficientes nos queda,

$$\varrho_t^{\Delta,1}(c) = \hat{\alpha}_0^e + \hat{\alpha}_1^e \mu_{c,t}^{\Delta,1} + \hat{\alpha}_4 \mu_{c,t}^{\Delta,4}$$

⁵ Indicador Trimestral de la Actividad Económica Estatal. Fuentes y metodologías.



Figura 3: Crecimiento de actividad

donde $e = e(c)$ es el estado al que pertenece la ciudad c ; o si las ciudades pertenecen a más estados, se divide en las partes correspondientes $c = c_1 \cup c_2 \cup c_3$ para el Valle de México).

Junto con los niveles iniciales $\varrho_{t_0}(c)$ podemos calcular recursivamente las estimaciones de actividad económica en los diferentes periodos del tiempo (antes y después),

$$\varrho_{t-1}(c) = \frac{\varrho_t(c)}{1 + \varrho_t^{\Delta,1}(c)} \quad \varrho_{t+1}(c) = \varrho_t(c)[1 + \varrho_t^{\Delta,1}(c)].$$

Resultados

Incluimos la tabla de estimaciones de actividad económica en los últimos años. En acompañamiento a los índices calculados, se presenta una gráfica que contiene la actividad económica de las zonas metropolitanas, agrupadas por estado y en comparación con los índices de PIB e ITAEE que publica el INEGI.



	Zona metropolitana	2014	2015
1	Acapulco	99 555	101 611
2	Aguascalientes	172 290	175 062
3	Campeche	272 618	255 856
4	Cancún	153 772	163 084
5	Cárdenas	50 150	50 376
6	Cd. del Carmen	222 263	208 604
7	Cd. Fernández	9 340	9 303
8	CDMX	3 871 795	3 969 113
9	Cd. Obregón	50 351	52 489
10	Cd. Victoria	32 698	33 661
11	Celaya	92 227	95 656
12	Chetumal	27 573	29 987
13	Chihuahua	136 924	143 291
14	Chilpancingo	28 238	29 072
15	Coatzacoalcos	127 481	127 700
16	Colima	50 554	51 996
17	Córdoba	23 612	23 787
18	Cuautla	53 840	55 375
19	Cuernavaca	97 350	98 543
20	Culiacán	130 614	135 758
21	Durango	66 387	69 497
22	Ensenada	53 153	53 368
23	Guadalajara	759 680	794 952
24	Guanajuato	23 698	24 409
25	Guaymas	35 262	35 045
26	Hermosillo	160 501	161 163
27	Irapuato	73 740	77 442
28	Juárez	227 322	247 651
29	La Laguna	271 718	281 149
30	La Paz	54 449	54 041
31	León	227 249	240 796
32	Los Cabos	42 978	44 862
33	Los Mochis	49 300	51 545
34	Manzanillo	35 034	36 342
35	Matamoros	79 371	80 623
36	Mazatlán	75 423	79 735
37	Mérida	190 835	195 843
38	Mexicali	225 315	226 482
39	Minatitlán	88 170	89 027
40	Monclova	68 581	71 042
41	Monterrey	1 043 566	1 103 901
42	Morelia	106 587	108 585
43	Moroleón	10 574	11 576
44	Nuevo Laredo	106 641	107 502
45	Oaxaca	81 853	82 805
46	Ocotlán	13 766	14 308
47	Orizaba	63 418	63 814
48	Pachuca	80 260	82 797
49	Pénjamo	26 110	27 093
50	Piedras Negras	43 976	45 590
51	Poza Rica	54 759	55 728
52	Puebla	344 533	355 254
53	Puerto Vallarta	58 029	60 737
54	Querétaro	265 208	281 831
55	Reynosa	151 369	153 491
56	Salamanca	51 362	54 347
57	Saltillo	142 135	148 829
58	San Francisco	22 881	24 062
59	San Juan del Río	45 669	48 398
60	San Luis	206 885	213 164
61	Tampico	96 801	98 639
62	Tapachula	21 699	21 568
63	Tecomán	11 795	12 329
64	Tehuacán	27 347	30 074
65	Tehuantepec	36 833	36 855
66	Tepic	57 654	58 069
67	Tijuana	175 627	185 264
68	Tlaxcala	48 558	52 375
69	Toluca	211 754	211 741
70	Tula	47 151	48 302
71	Tulancingo	23 523	23 726
72	Tuxtla	102 481	103 800
73	Uruapan	33 348	33 442
74	Veracruz	164 143	164 597
75	Villahermosa	220 488	222 902
76	Xalapa	81 370	81 469
77	Zacatecas	51 973	51 633
78	Zamora	21 976	22 214

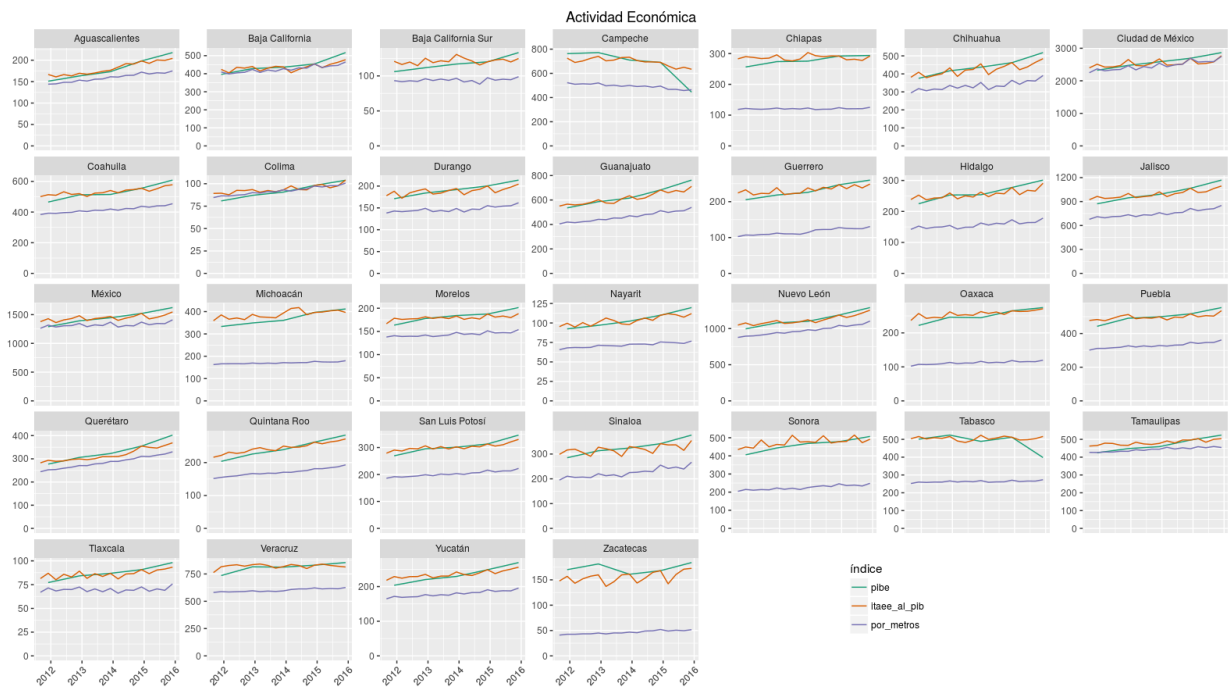


Figura 4: Comparación estados