



01/07/2014

Estimación de población en áreas pequeñas mediante percepción remota de alta resolución espacial

REPORTE TÉCNICO

José Luis Silván Cárdenas

Estimación de población en áreas pequeñas mediante percepción remota de alta resolución espacial

REPORTE TÉCNICO

Contenido

RESUMEN	3
INTRODUCCIÓN	3
SELECCIÓN DE ÁREAS PILOTO Y PREPARACIÓN DE LOS DATOS	5
MODELOS PARA ESTIMACIÓN DE POBLACIÓN	7
Medidas de desempeño.....	8
Selección del modelo.....	8
MODELOS PARA ESTIMACIÓN DE VIVIENDAS.....	13
Medidas de desempeño.....	13
Selección del modelo.....	14
ESTIMACIÓN DE VIVIENDAS CON PR	15
Clasificación de uso de suelo	15
Muestra de entrenamiento.....	16
Clasificación con Isodata.....	16
Clasificación con MLC.....	16
Clasificación con SVM.....	16
Evaluación de la precisión	17
Extracción Automática de Edificios.....	19
Estimación del número de pisos.....	20
Estimación de viviendas	22
ESTIMACIÓN DE POBLACIÓN CON PR.....	26
CONCLUSIONES.....	27

TRABAJOS CITADOS	29
Apéndice A. Lista de publicaciones sobre estimación de población mediante análisis de datos geoespaciales	31
Apéndice B. Procesamiento datos vectoriales de catastro.....	38
Conversión de formato.....	38
Corrección geométrica	38
Simplificación y filtrado de polígonos de edificios.....	38
Cálculo de Atributos para Polígonos de Edificios	39
Edición del vector de predios.....	41
Interpretación de la clave de Uso de Suelo.....	41
Cálculo de Atributos para Polígonos de Predios	42
Apéndice C. Procesamiento de la nube de puntos LiDAR.....	43
Retícula	44
Extracción puntos	45
Rasterización.....	47
Filtrado de puntos de suelo	48
Segmentación y filtrado de terreno	48
Creación de mosaicos.....	50
Clasificación de la nube de puntos	51
Apéndice D. Procesamiento de las imágenes multi espectrales	52
Preparación de los datos	52
Recorte	53
Calibración radiométrica	53
Realce espacial	53
Orto-rectificación.....	53
Corrección atmosférica	55
Creación de mosaicos.....	55

Clasificación de cobertura.....	55
---------------------------------	----

RESUMEN

Se presenta el reporte final del proyecto denominado “Estimación de población en áreas pequeñas mediante percepción remota de alta resolución espacial” financiado por el Fondo Sectorial CONACyT-INEGI 2012, con número de registro 187593 y formalizado mediante Convenio de Asignación de Recursos de fecha 5 de octubre de 2012.

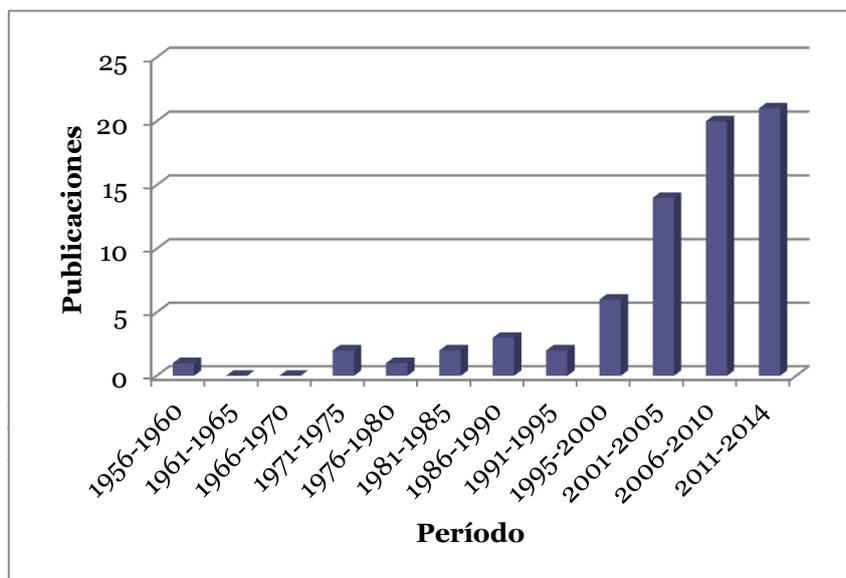
El objetivo general del proyecto fue desarrollar y probar una metodología de estimación de población en áreas geo estadísticas básicas (AGEBS) y manzanas empleando imágenes de alta resolución espacial y datos de altimetría adquiridos mediante el sensor LiDAR aerotransportado, así como datos vectoriales del catastro. En este reporte se presentan los componentes principales de la metodología desarrollada, los productos intermedios y los principales resultados encontrados. El orden de presentación no refleja necesariamente el orden en que se realizaron las actividades o se obtuvieron los productos.

INTRODUCCIÓN

Como parte de la investigación, se consultaron alrededor de 70 publicaciones en medios como revistas arbitradas, capítulos de libros y memorias de conferencias sobre el tema de estimación de población basada en datos geoespaciales, técnicas de percepción remota y de análisis espacial (Apéndice A). A partir de dicha consulta, se observó que la gran mayoría de los estudios han empleado datos provenientes de sensores de resolución baja, tales como DMSP y AVHRR, pero sobre todo de resolución moderada, principalmente de sensores como Landsat, ASTER y Radar. Dentro del primer grupo, se encuentran aquellos estudios que se enfocan en la estimación de población a nivel regional y global, mientras que en el segundo grupo se encuentran estudios que estiman la población a nivel de ciudades y regiones. Notablemente, a la escala de ciudad gana importancia la estimación de parámetros como la superficie impermeable y tipos de cobertura y uso de suelo.

Los estudios de estimación de población en áreas pequeñas, como manzanas o AGEBS, incluyen generalmente el uso de imágenes de alta resolución espacial, sobresaliendo la fotografía aérea, la cual cuenta con una larga tradición que ha evolucionado de los formatos analógicos a los formatos digitales. En México, se tiene como antecedente el uso de fotografía aérea para estimación de población en áreas pequeñas un estudio realizado por el COLMEX (Ordorica & Ibarra, 2002), donde se estimó la población de la Villa Milpa Alta mediante el ajuste de la series históricas de población (1970-2000) con la extensión del asentamiento observada en las fotografías aéreas para las distintas fechas de los censos. En ese caso, el área urbana se delineó manualmente.

En 1999 con el lanzamiento de IKONOS, el cual es considerado el primer satélite de alta resolución espacial, se inicia una etapa de exploración de metodologías automatizadas para estimación de población en áreas pequeñas. De esto da cuenta un incremento exponencial en el número de publicaciones sobre estimación de población (Ilustración 1). La detonación de este incremento acelerado en el aparente interés sobre el tema se explica también en gran medida por la amplia disponibilidad de herramientas de análisis espacial en los sistemas de información geográficas, los cuales permitieron integrar los datos vectoriales con datos en formato raster, que es el formato nativo de las imágenes.



Entre los sensores satelitales multi espectrales de alta resolución espacial que se consideran en esta categoría se encuentran IKONOS (con 1 m de resolución en la pancromática), QuickBird (con 0.6 m), WorldView (con 0.5 m) y Geoeye (con 0.4 m), entre los cuales, la mayoría de estudios emplean IKONOS y QuickBird.

Aunque en nuestra revisión se encontraron sólo 10 estudios que emplean

imágenes satelitales de alta resolución espacial, una proporción similar de estudios emplean fotografías aéreas y otro tanto se basan en datos de catastro o mapas topográficos detallados con información de edificios. En la mayoría de estos casos se observó que la estimación de población se lleva a cabo a partir de dos enfoques:

- 1) Conteo de unidades de viviendas, y
- 2) Determinación de áreas de viviendas

En algunos países, como en Japón, se han empleado datos de catastro actualizados para estimar población con base no solo en el número o el área de edificios sino también considerando el número de niveles de los edificios (Lwin & Murayama, 2009). Con el advenimiento del sensor LiDAR, se habilita poder hacer mediciones de alturas de edificios e incorporarlas en los métodos de estimación de población en áreas pequeñas. Aunque los escáneres por láser empleados para medición de la superficie terrestre se crean en los 90s, su aplicación a estimación de población es muy reciente con la referencia más antigua al 2010. En este tipo de modelos se incorpora el volumen como variable externa para explicar la población, donde además, la extracción de edificios juega un papel fundamental (Lwin & Murayama, 2011; Silván-Cárdenas, L. Wang, Rogerson, Feng, & Kamphaus, 2010; Qiu, Sridharan, & Chun, 2010; Sridharan & Qiu, 2013).

En esta línea, el objetivo del proyecto fue desarrollar y probar una metodología de estimación de población en áreas geoestadísticas básicas (AGEBS) y manzanas, empleando imágenes de alta resolución espacial y la nube de puntos tridimensional levantada mediante el sensor LiDAR aerotransportado. Para abordar el problema desde un punto de vista formal, se plantearon dos hipótesis a probar.

Hipótesis 1: La población de un área depende del espacio habitable en el área (alometría).

Hipótesis 2: El espacio habitable se puede cuantificar mediante percepción remota.

Para probar las hipótesis se seleccionaron áreas pilotos como se describe en la siguiente sección.

SELECCIÓN DE ÁREAS PILOTO Y PREPARACIÓN DE LOS DATOS

Para la determinación de las áreas piloto se analizó la distribución de la población (Ilustración 2) y el ritmo de crecimiento a nivel de delegación (Ilustración 3), y se eligieron tres delegaciones representativas de un núcleo densamente poblado con bajo crecimiento relativo, área periurbana con crecimiento moderado y área rural con crecimiento acelerado. Las delegaciones que cumplen con estas características son: Cuauhtémoc, Tlalpan y Milpa Alta, respectivamente.

Una vez seleccionadas las áreas de estudio, se obtuvo la geometría de manzanas y AGEBs para cada una de éstas, así como los datos de población correspondientes para 2000 y 2010. Se anexaron los datos de población y vivienda de los censos 2000 y 2010 a nivel AGEB y 2010 a nivel manzana. No se pudo obtener datos de población a nivel de manzana para el 2000. La Ilustración 4 muestra la distribución espacial del crecimiento o contracción de la población en cada uno de los AGEBS de las delegaciones seleccionadas.

Además de los datos de población y vivienda, se calcularon transformaciones logarítmicas, así como los incrementos absolutos y relativos de población y vivienda. Por otro lado, se calculó el espacio habitable a partir de los valores agregados a nivel de predios (Apéndice B).

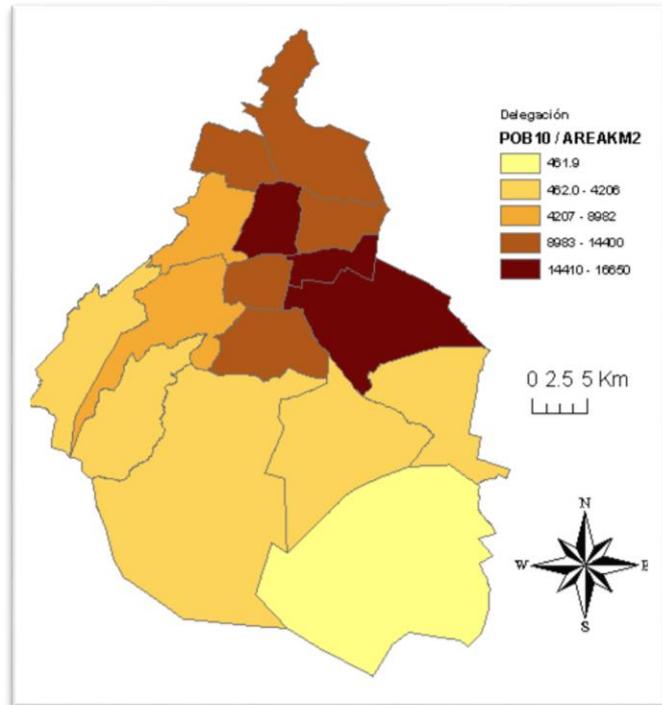


Ilustración 2. Densidad de población de las delegaciones del DF al 2010

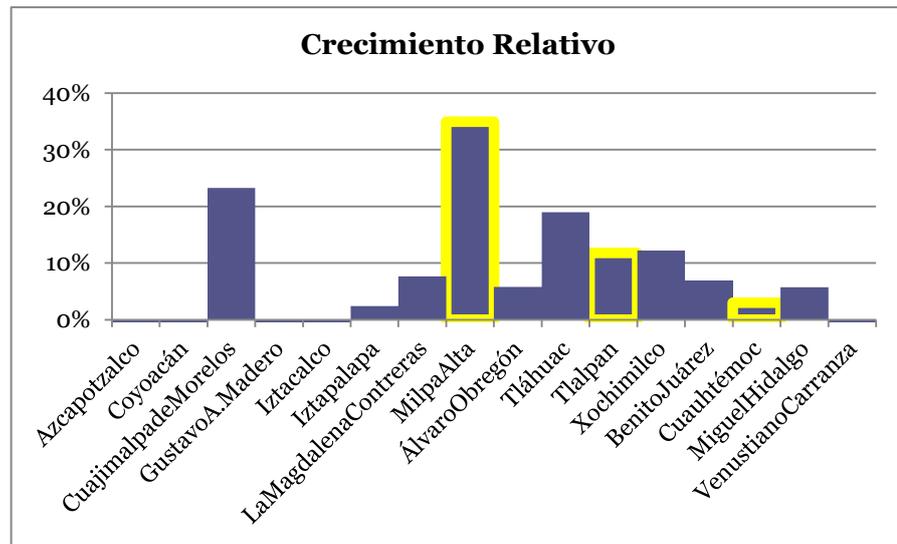


Ilustración 3. Crecimiento relativo de la población en las delegaciones del DF durante el período 2000-2010.

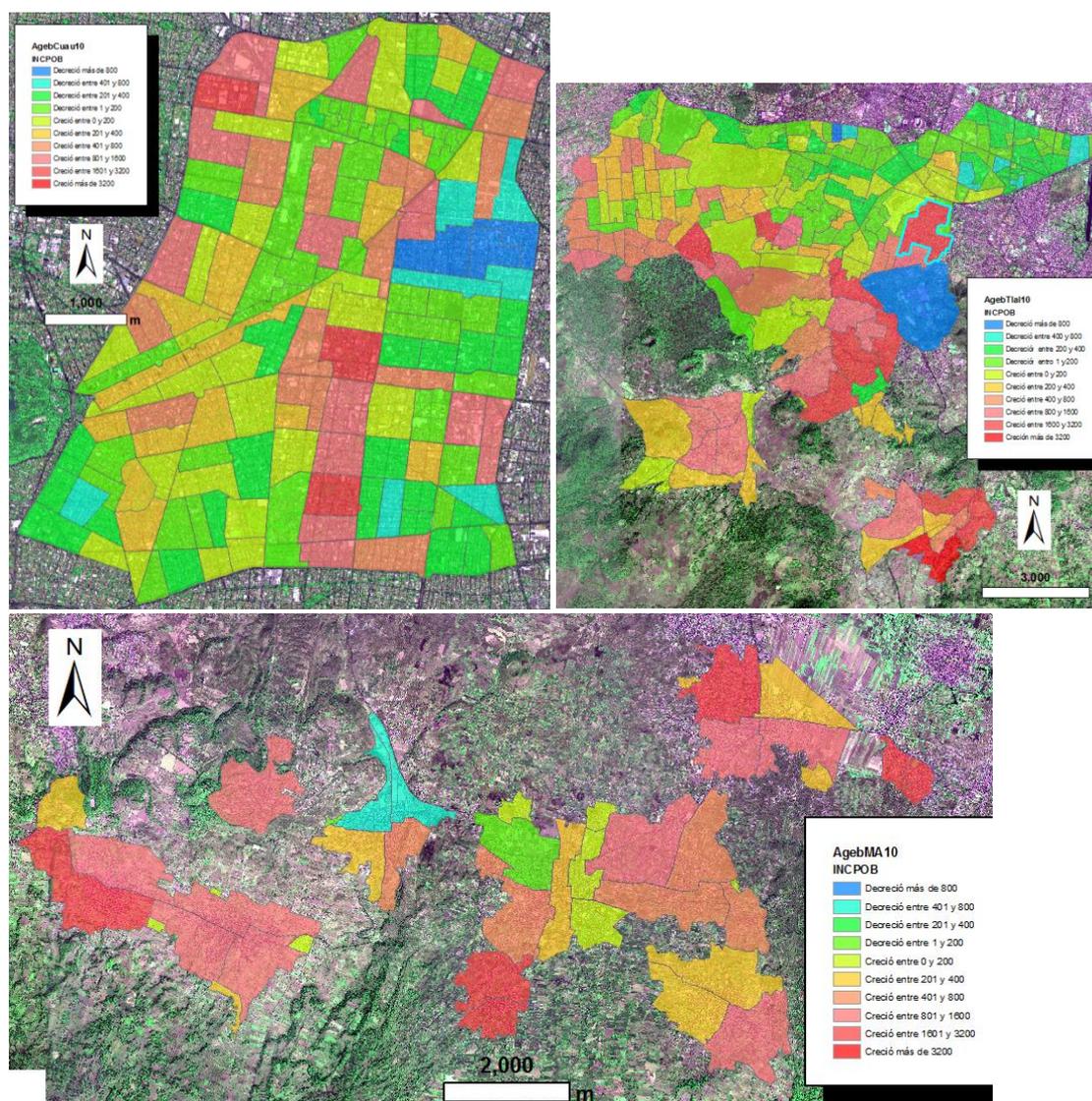


Ilustración 4. Crecimiento positivo (tonos amarillo-rojo) y negativo (tonos verde-azul) a nivel de AGEB para las tres delegaciones de estudio.

La Tabla 1 resume los atributos anexados a la capa de AGEBS. A nivel de manzana no se pudieron generar todos los atributos debido a que no se contaba con los datos censales a ese nivel de agregación para el año 2000. Para el ESPH00 se usaron los polígonos de edificio del catastro (Apéndice B), mientras que para el ESPH10 se usaron los polígonos extraídos con Lidar que se describe más adelante. Se usó el atributo de uso de suelo de la capa de predios para filtrar los edificios por uso y así cuantificar el espacio habitable para uso habitacional y otros usos.

Todos estos atributos se anexaron a la geometría del 2010. Aunque la geometría de AGEBS cambia entre el 2000 y el 2010, la asignación de los datos del 2000 a la geometría se basó en las claves de los AGEBS que se mantuvieron entre las fechas. En algunos casos no se pudo ligar directamente un AGEB que existía en 2000 y no existía en 2010 o viceversa. En el primer caso se trata de AGEBS que se partieron y se asignaron nuevas claves. En estos casos se hizo una asignación proporcional de la población/vivienda del AGEB del 2000 en función del área de los AGEB hijos del 2010. En el segundo caso se trataba de AGEB en un área donde antes no existió AGEB. En estos casos, la población y viviendas del 2000 se consideraron nulas.

Tabla 1. Lista de atributos para la geometría de AGEBS del 2010. Los atributos con asterisco (*) no se pudieron generar a nivel de manzana.

Atributo	Tipo	Significado	Cálculo	Uso Principal
POB00*	Long	Población total en 2000	Join	Predicción
VIV00*	Long	Viviendas en 2000	Join	Análisis y Modelado
POB10	Long	Población total en 2010	Join	Calibración y Validación
VIV10	Long	Viviendas en 2010	Join	Análisis y Modelado
LPOB00*	Float	Logaritmo de POB00	$\text{LOG}_{10}(1+\text{POB00})$	Análisis y Modelado
LVIV00*	Float	Logaritmo de VIV00	$\text{LOG}_{10}(1+\text{VIV00})$	Análisis y Modelado
LPOB10	Float	Logaritmo de POB10	$\text{LOG}_{10}(1+\text{POB10})$	Análisis y Modelado
LVIV10	Float	Logaritmo de VIV10	$\text{LOG}_{10}(1+\text{VIV10})$	Análisis y Modelado
INCPOB*	Int	Incremento de POB	$\text{POB10}-\text{POB00}$	Análisis y Modelado
INCVIV*	Int	Incremento de VIV	$\text{VIV10}-\text{VIV00}$	Análisis y Modelado
INCLPOB*	Float	Incremento de LPOB	$\text{LPOB10}-\text{LPOB00}$	Análisis y Modelado
INCLVIV*	Float	Incremento de LVIV	$\text{LVIV10}-\text{LVIV00}$	Análisis y Modelado
INCRPOB*	Float	Incremento relativo de POB	$2*\text{INCPOB}/(\text{POB10}+\text{POB00})$	Análisis y Modelado
INCRVIV*	Float	Incremento relativo de VIV	$2*\text{INCVIV}/(\text{VIV10}+\text{VIV00})$	Análisis y Modelado
EDIFH00	Long	# edificios con uso habitacional al 2000	Join (Predios)	Análisis y Modelado
ESPH00	Float	Espacio habitable uso habitacional al 2000	Join (Predios)	Análisis y Modelado
ESPT00	Float	Espacio habitable total al 2000	Join (Predios)	Análisis y Modelado
ESPH10	Float	Espacio habitable uso habitacional al 2010	Join (Predios)	Análisis y Modelado
ESPT10	Float	Espacio habitable total al 2010	Join (Edificios Lidar)	Análisis y Modelado
LESPH00	Float	Logaritmo de ESPH00	$\text{LOG}_{10}(1+\text{ESPH00})$	Análisis y Modelado
LESPH10	Float	Logaritmo de ESPH10	$\text{LOG}_{10}(1+\text{ESPH10})$	Análisis y Modelado
LESPT00	Float	Logaritmo de ESPT00	$\text{LOG}_{10}(1+\text{ESPT00})$	Análisis y Modelado
LESPT10	Float	Logaritmo de ESPT10	$\text{LOG}_{10}(1+\text{ESPT10})$	Análisis y Modelado
INCESPT	Float	Incremento del espacio habitable	$\text{ESPT10}-\text{ESPT00}$	Análisis y Modelado
R3INCESPT	Float	Raíz cúbica de INCESPT	$\text{INCESPT}^{(1/3)}$	Análisis y Modelado

MODELOS PARA ESTIMACIÓN DE POBLACIÓN

Con miras a probar la Hipótesis 1, se desarrollaron 5 modelos regresivos de población con número de viviendas como variable independiente. Los 5 modelos probados se listan en la Tabla 2. Aunque el ajuste se lleva a cabo mediante regresión lineal simple, la relación entre población y vivienda puede ser

no lineal vía una transformación de las variables. La relación logarítmica entre población y vivienda (Modelo 2) se deriva de la teoría de crecimiento alométrico de acuerdo con la cual el incremento relativo de población es proporcional a un incremento relativo de viviendas. Al plantear el problema en término de derivadas, resulta una ecuación diferencial cuya solución es una relación logarítmica entre las variables. En la práctica, se suma uno antes de tomar logaritmos para evitar indeterminación. También se plantearon variantes en términos de incrementos aritméticos, geométricos y relativos en forma discreta (Modelos 3-5).

Tabla 2. Modelos regresivos población-vivienda

Modelo	Expresión*	Descripción
P1	$POB = a \cdot VIV + b$	Relación lineal
P2	$LPOB = a \cdot LVIV + b$	Relación logarítmica
P3	$INCPOB = a \cdot INCVIV + b$	Relación lineal entre Incrementos aritméticos
P4	$INCLPOB = a \cdot INCLVIV + b$	Relación lineal entre incrementos geométricos (escala logarítmica)
P5	$INCRPOB = a \cdot INCRVIV + b$	Relación lineal entre incrementos relativos

Medidas de desempeño

Con el fin de cuantificar el desempeño de cada uno de los modelos, se emplearon tres medidas estándares que son:

1. El coeficiente de determinación (R^2), el cual mide el porcentaje de varianza de la variable dependiente que está determinada por la variable independiente.
2. Mediana del Error Absoluto (MAE), el cual proporciona una medida de tendencia central del error absoluto, al tiempo que es robusto ante datos anómalos.
3. Mediana del Error Absoluto Relativo (MARE), el cual proporciona una medida de qué también se distribuye el error en términos proporcionales a los valores de la variable estimada.

Para la calibración de los modelos se seleccionaron aleatoriamente 50 AGEBS o dos tercera parte del número total disponible si este era menor a 50 (fue el caso de Milpa Alta), mientras que el resto de los datos se emplearon como muestra de validación. Tanto MAE como MARE se calcularon sobre la muestra de validación, mientras que la R^2 se calculó sobre la muestra de calibración. Se excluyeron las divisiones por cero para el cálculo de MARE.

Selección del modelo

Los valores de las medidas de desempeño para cada uno de los modelos probados y para cada sitio de estudio se presentan en la Tabla 3. En términos del coeficiente de determinación se observó que los modelos P1 y P2 se desempeñaron bastante bien para todos los sitios de estudio ($R^2 > 0.91$), el modelo P3 obtuvo el peor desempeño en los tres casos ($R^2 < 0.85$), mientras que los modelos P4 y P5 se desempeñaron mejor para Tlalpan y Milpa Alta. Estas observaciones sugieren que el enfoque incremental implícito en estos modelos resulta menos apropiado para una delegación como la Cuauhtémoc, la cual presenta una sobre saturación. En otras palabras, la dinámica demográfica en esta delegación no está en sincronía con la dinámica de la vivienda. No obstante, en términos de las métricas MAE y MARE sobre la muestras de prueba parecen contradecir lo anterior, ya que el modelo P5 resulta el mejor para Cuauhtémoc y Tlalpan, mientras que el modelo P1 resulta ser el mejor para Milpa Alta. En realidad, prácticamente todos los modelos se desempeñaron muy bien para el caso de Milpa Alta, pero fueron los modelos P4 y P5 los que presentaron valores consistentemente bajos para las tres delegaciones. Esto se puede observar al comparar visualmente los diagramas de dispersión.

En resumen, el modelo que mejor estima la población y que resultó ser más consistente es el modelo P5, siendo las fuentes de error aquellas áreas donde se observa población pero no existen viviendas. Este modelo se puede expresar en forma explícita de la siguiente manera:

$$P = \frac{(1 + a + b)V + (1 - a + b)V_0}{(1 - a - b)V + (1 + a - b)V_0} P_0$$

donde V_0 y P_0 es el número de viviendas y población en la fecha inicial y V y P son los valores para la fecha a estimar, y a y b son los parámetros del modelo. Estos parámetros fueron relativamente estable, es decir, tuvieron poca variación entre una y otra delegación. En general, se obtuvieron valores de alrededor de 1.0 para a , y alrededor de -0.25 para b .

Tabla 3. Resultados de las medidas de desempeño de los modelos de estimación de población para Cuauhtémoc. Los datos en negritas representan los mejores valores de desempeño.

Modelo	a	b	R2	MAE	MARE
P1	2.6832	-140.73	0.9191	263	10%
P2	1.0966	0.09222	0.9673	301	11%
P3	1.4318	-456.47	0.5048	196	7%
P4	0.9366	-0.1289	0.7895	105	4%
P5	0.92	-0.2817	0.8207	89	4%

Tabla 4. Resultados de las medidas de desempeño de los modelos de estimación de población para Tlalpan. Los datos en negritas representan los mejores valores de desempeño.

Modelo	a	b	R2	MAE	MARE
P1	3.6574	-264.84	0.9597	214	9%
P2	1.0799	0.28329	0.9779	206	9%
P3	3.1437	-492.21	0.7748	188	7%
P4	1.259	-0.1172	0.9799	131	6%
P5	1.0912	-0.2302	0.942	124	5%

Tabla 5. Resultados de las medidas de desempeño de los modelos de estimación de población para Milpa Alta. Los datos en negritas representan los mejores valores de desempeño.

Modelo	a	b	R2	MAE	MARE
P1	3.6319	24.6158	0.9912	54	4%
P2	0.993	0.5826	0.9968	58	4%
P3	2.4333	-159.79	0.8519	134	11%
P4	1.1353	-0.1101	0.9572	73.5	9%
P5	1.1222	-0.2437	0.9572	58	5%

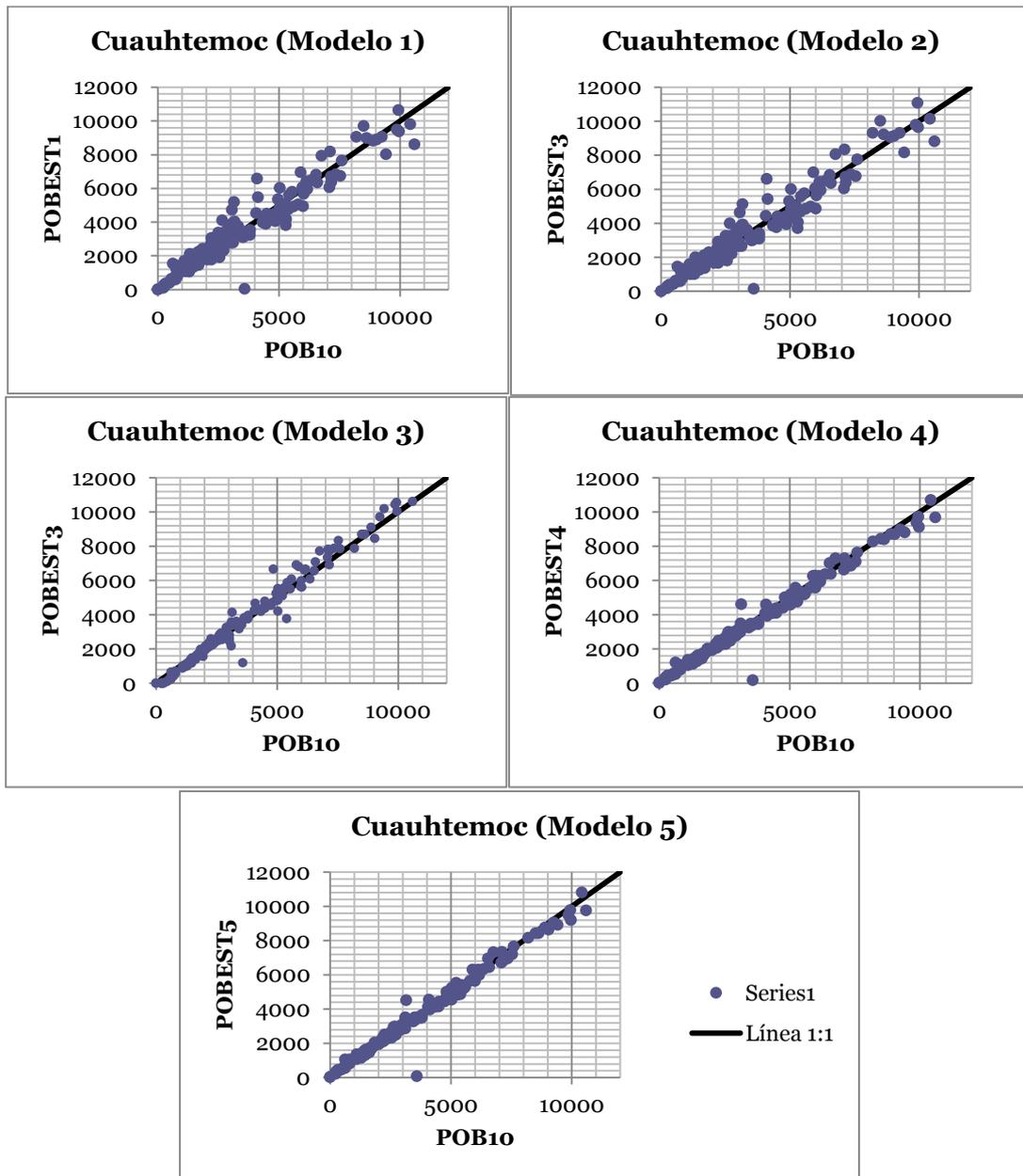


Ilustración 5. Diagramas de dispersión entre población observada y población estimada con cada uno de los modelos para Cuauhtémoc.

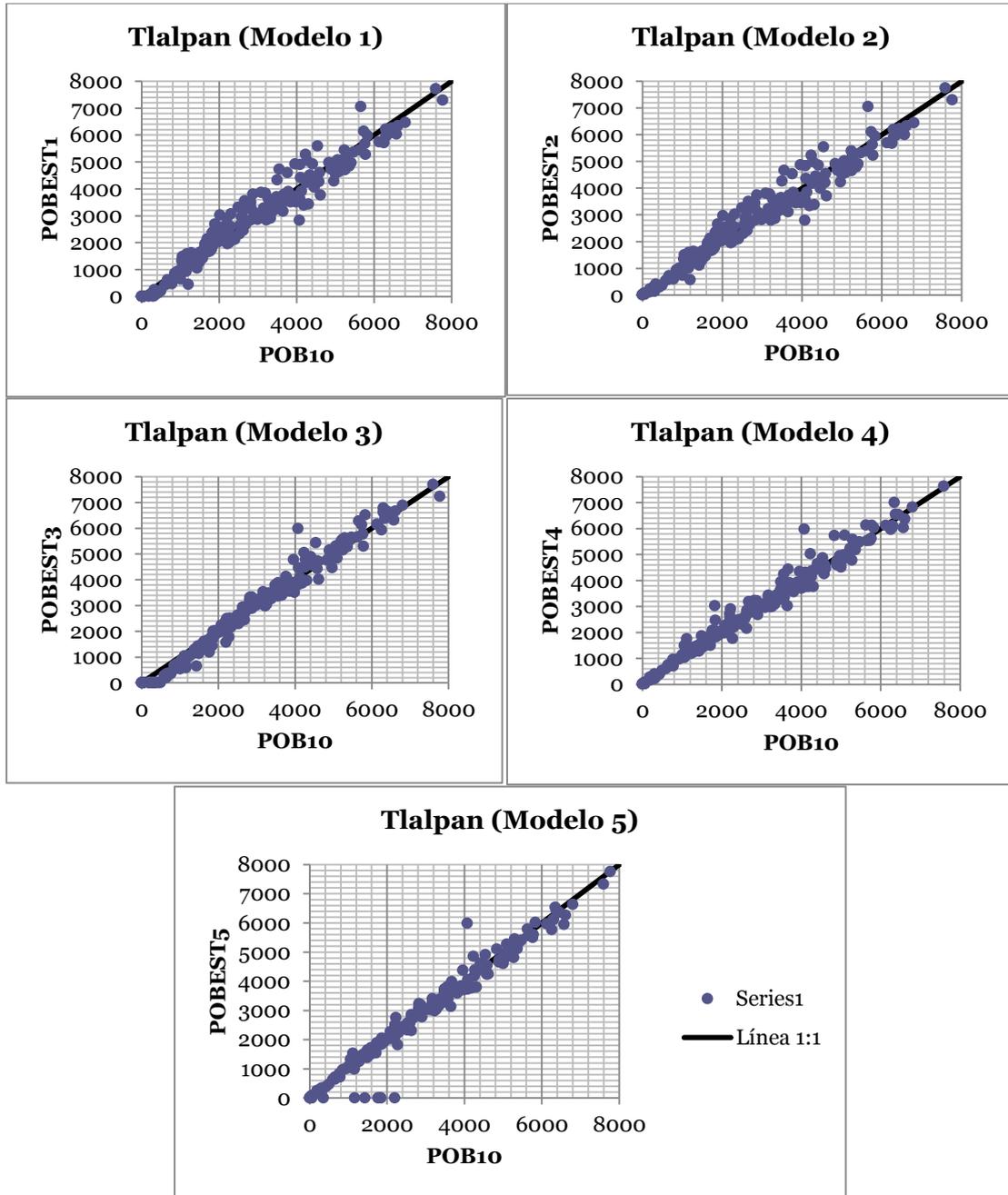


Ilustración 6. Diagramas de dispersión entre población observada y población estimada con cada uno de los modelos para Talpan.

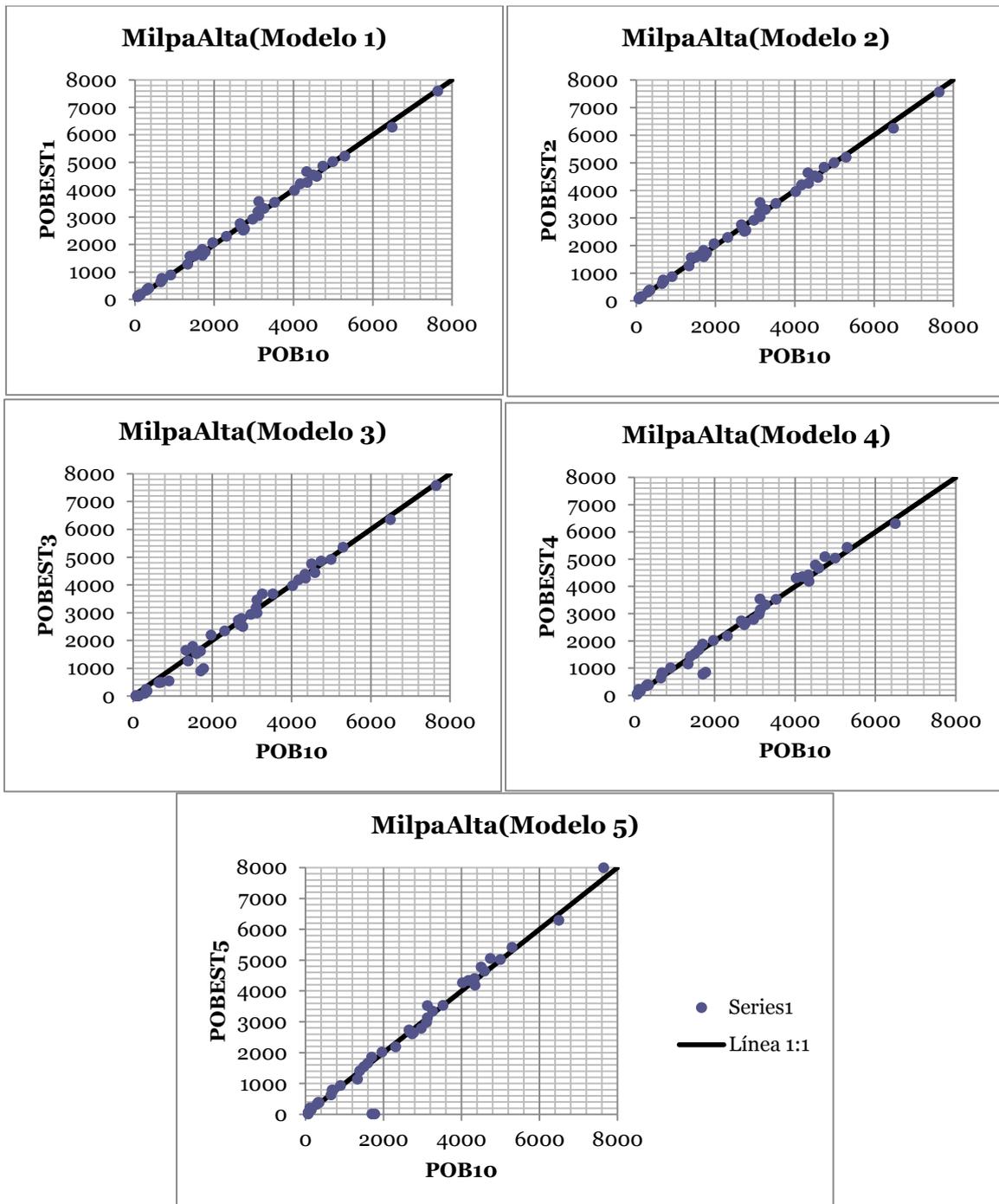


Ilustración 7. Diagramas de dispersión entre población observada y población estimada con cada uno de los modelos para Milpa Alta.

MODELOS PARA ESTIMACIÓN DE VIVIENDAS

Para estimar las viviendas se desarrollaron 8 modelos (Tabla 6), los cuales resultan de combinar dos tipos de relaciones (lineal y logarítmica) con cuatro tipos de variables independientes. Las variables independientes son: 1) número de edificios con uso de suelo habitacional, 2) espacio habitable de edificios con uso de suelo habitacional, 3) Espacio habitable de edificios seleccionados óptimamente y 4) espacio habitable total considerando todos los edificios. En el caso 3 la selección de los usos de suelo se selecciona mediante el método Stepwise Linear Regression (SLR) con la implementación disponible en el ToolBox Statistics de MATLAB (The Mathworks Inc.). Se trata de un algoritmo iterativo que incorpora y retira variables del modelo en función de su significancia estadística. Hay dos parámetros claves que controlan la entrada y salida de variables: 1) la probabilidad de significancia de la variable de entrada y 2) la probabilidad de significancia de la variable de salida. En todos los experimentos se usaron los valores de 0.01 para la primera y de 0.02 para la segunda. En el resto de los modelos se empleó regresión simple usando Excel 2010 (Microsoft Corp.)

Tabla 6. Modelos de estimación de viviendas en términos de características de edificios y usos de suelo.

Modelo	Expresión	Relación	Variable(s) independiente(s)
V1	$VIV = a * EDIFH + b$	Lineal	Número de edificios con uso de suelo habitacional
V2	$LVIV = a * LEDIFH + b$	Logaritmico	Número de edificios con uso de suelo habitacional
V3	$VIV = a * ESPH + b$	Lineal	Espacio habitable de edificios con uso de suelo habitacional
V4	$LVIV = a * LESPH + b$	Logaritmico	Espacio habitable de edificios con uso de suelo habitacional
V5	$VIV = a_1 * ESPH_1 + \dots + a_k * ESPH_k + b$	Lineal	Espacio habitable de edificios con usos de suelo seleccionados óptimamente*
V6	$LVIV = a_1 * LESPH_1 + \dots + a_k * LESPH_k + b$	Logaritmico	Espacio habitable de edificios con usos de suelo seleccionados óptimamente*
V7	$VIV = a * ESPT + b$	Lineal	Espacio habitable total (Todos los usos de suelo)
V8	$LVIV = a * LESPT + b$	Logaritmico	Espacio habitable total (Todos los usos de suelo)

*Los usos de suelo óptimos fueron seleccionados usando el método stepwise linear regression.

Medidas de desempeño

Como en el caso de los modelos de estimación de población, se definieron métricas análogas para cuantificar la bondad de ajuste de los modelos a los datos, estas fueron:

1. El coeficiente de determinación (R^2), el cual mide el porcentaje de varianza de la variable dependiente que está determinada por la variable independiente.
2. Mediana del Error Absoluto (MAE), el cual proporciona una medida de tendencia central del error absoluto, al tiempo que es robusto ante datos anómalos.
3. Mediana del Error Absoluto Relativo (MARE), el cual proporciona una medida de qué también se distribuye el error en términos proporcionales a los valores de la variable estimada.

Con el fin de disminuir los errores por causa de una clasificación de uso de suelo mediante PR, se optó por emplear inicialmente los datos del catastro del 2000 para probar los modelos. Se siguió un esquema de muestreo aleatorio similar al empleado anteriormente.

Selección del modelo

Los resultados obtenidos se presentan de la a la Tabla 9, mientras que los modelos con mejor desempeño en cada caso se ilustran de la Ilustración 8 a la Ilustración 10.

Al igual que en los resultados de los modelos de población, en la estimación de viviendas también se observa una clara distinción entre la delegación Cuauhtémoc y las delegaciones Tlalpan y Milpa Alta. En el caso de Cuauhtémoc, el modelo que mejor ajustó los datos fue el modelo V5, el cual selecciona óptimamente los usos de suelos que se incorporan en el modelo. Estos fueron los usos habitacional, espacios abiertos y espacios públicos. Esto indica que la relación entre las viviendas y los edificios es mucho más compleja en Cuauhtémoc que en las otras delegaciones. Además, al analizar los coeficientes del modelo resultante se observó una contribución positiva de los espacios abiertos que indica que el uso de suelo podría estar desactualizado y que en realidad son espacios habitacionales. Lo mismo sucede, aunque en menor medida, con los espacios públicos que figuran en el modelo.

Tabla 7. Resultados del ajuste de los modelos de estimación de vivienda a los datos de Cuauhtémoc. Los usos de suelo óptimos para el modelo 5 fueron: espacios abiertos (B), habitacional (H) y espacios públicos (U); mientras que para el modelo 6 fueron: habitacional (H) e industrial (I). Los valores en negritas indican el mejor desempeño.

Modelo	a	b	R ² ^a	MAE ^b	MARE ^c
V1	0.93	632.87	0.27	357	41%
V2	0.41	1.96	0.50	308	41%
V3	0.01	498.98	0.34	334	37%
V4	0.57	0.21	0.46	259	35%
V5	0.86 (B) 0.01 (H) 0.34 (U)	339.38	0.69	245	29%
V6	0.08 (H) 0.15 (I)	2.0508	0.54	275	35%
V7	0.00	797.69	0.06	363	43%
V8	0.37	0.87	0.13	289	34%

Tanto en Tlalpan como en Milpa Alta destacaron los modelos V1 y V2 los cuales relacionan el número de edificios con el número de viviendas. Esto sugiere una predominancia de casas unifamiliares a edificios multifamiliares o, de otra forma, predominan los multifamiliares que albergan en promedio un mismo número de población.

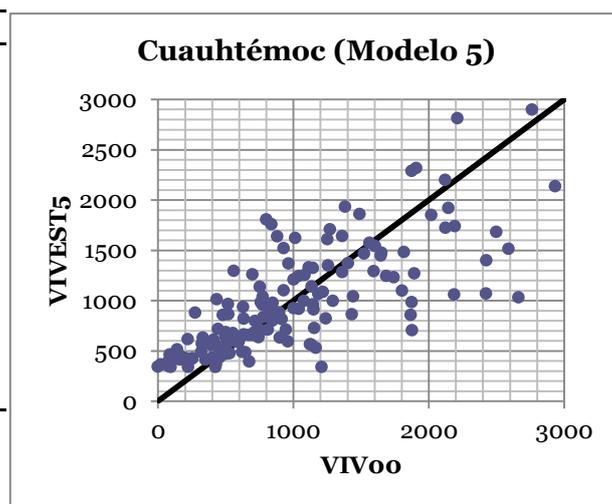


Ilustración 8. Diagrama de dispersión entre observaciones (VIV00) y estimaciones de viviendas usando el modelo 5 (VIVEST5).

Tabla 8. Resultados del ajuste de los modelos de estimación de viviendas a los datos de la delegación Tlalpan. Tanto para el modelo 5 como para el 6, el uso óptimo fue habitacional (H). Los valores en negritas indican el mejor desempeño.

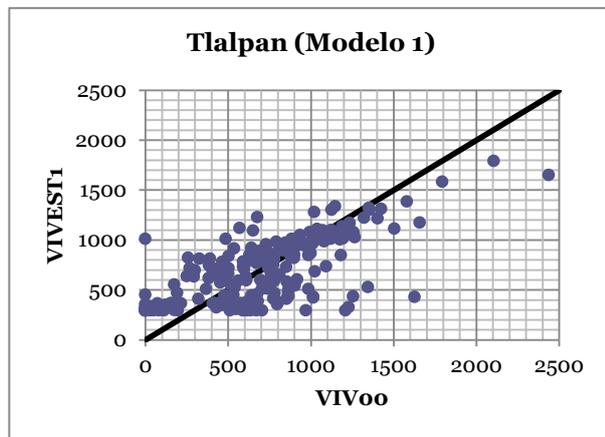


Ilustración 9. Diagrama de dispersión entre observaciones (VIVoo) y estimaciones de viviendas usando el modelo 1 (VIVEST1).

las viviendas. Esto es probablemente debida a la presencia de usos mezclados con habitacional, es decir, la cantidad de espacio habitable en edificios con uso de suelo no habitacional en un AGEB es proporcional el espacio habitable en edificios habitacionales.

Tabla 9. Resultados del ajuste de modelos de estimación de viviendas a los datos de la delegación Milpa Alta. Tanto para el modelo 5 como para el 6, el uso óptimo fue habitacional (H). Los valores en negritas indican el mejor desempeño.

Modelo	A	b	R2 ^a	MAE ^b	MARE ^b
V1	0.56	231.76	0.61	133	16%
V2	0.67	0.89	0.72	127	17%
V3	0.01	310.55	0.44	240	25%
V4	0.61	-0.06	0.75	183	25%
V5	0.01	310.55	0.44	240	25%
(H)					
V6	0.61	-0.06	0.75	183	25%
(H)					
V7	0.00	318.21	0.40	212	31%
V8	0.61	-0.10	0.76	254	35%

^aModelos ajustados a muestra de entrenamiento (N = 21)

^bCálculos para muestra de prueba (N = 10)

En suma, no es posible establecer un modelo único para todos los casos, aunque es deseable un modelo que demande la mínima cantidad de información.

Modelo	a	b	R2 ^a	MAE ^b	MARE ^b
V1	0.61	296.21	0.61	206	30%
V2	0.67	0.94	0.52	164	34%
V3	0.00	399.36	0.32	256	32%
V4	0.45	0.65	0.51	228	35%
V5	0.00	535.84	0.08	249	32%
(H)					
V6	0.14	2.13	0.21	232	33%
(H)					
V7	0.00	469.46	0.20	278	33%
V8	0.99	-2.15	0.59	268	42%

^aModelos ajustados a muestra de entrenamiento (N = 50)

^bCálculos para muestra de prueba (N = 156)

Adicionalmente, pero sobre todo en el caso de Milpa Alta, el modelo V8 mostró un desempeño destacado indicando que el uso de suelo habitacional no resulta tan crucial para predecir

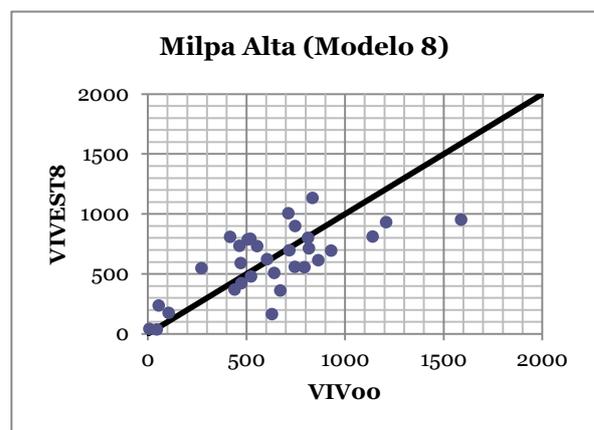


Ilustración 10. Diagrama de dispersión entre observaciones (VIVoo) y estimaciones de viviendas usando el modelo 8 (VIVoo_EST8).

ESTIMACIÓN DE VIVIENDAS CON PR¹

En los modelos de estimación de viviendas (y por tanto de población), existen dos tipos de información fundamentales que se pueden extraer con PR: 1) Uso de suelo y 2) Edificios.

Clasificación de uso de suelo

¹ Parte de los resultados de esta sección fueron publicados en una publicación (Silván-Cárdenas, Almazán-González, & Couturier, 2014).

Con el fin de evaluar la viabilidad para determinar el uso de suelo a partir de las características de los edificios, se diseñó un experimento de clasificación a nivel de predios en el cual se probaron tres métodos de clasificación de complejidad variable:

1. ISODATA - es un método no supervisado que agrupa naturalmente los datos.
2. Clasificador de Máxima Verosimilitud (MLC por sus siglas en inglés) – es un método de clasificación supervisado que modela la distribución de probabilidades de las clases como funciones Gausianas y selecciona la clase con mayor probabilidad a posteriori.
3. Máquinas de soporte vectorial (SVM por sus siglas en inglés) – método de reconocimiento de patrones supervisado que discrimina entre pares de clases usando un kernel lineal (SVML) o no lineal (SVMK).

Dado que Cuauhtémoc representa la mayor complejidad en cuanto al ensamblaje de uso de suelo, se selección ésta como área de prueba. Además, con el fin de reducir la incertidumbre que suele introducir un método de extracción de edificios mediante PR, se optó por emplear los polígonos de edificio del catastro del 2000 y así atribuir cualquier error de clasificación únicamente al clasificador en cuestión.

Muestra de entrenamiento

Los datos utilizados como muestra ya clasificada para entrenar los algoritmos de clasificación supervisados (MLC, SVML Y SVMLK) fueron 160 predios etiquetados en 4 clases de uso de suelo, distribuidos de la siguiente manera:

Tabla 10. Esquema de clasificación y tamaño de la muestra de entrenamiento empleada con los métodos de clasificación supervisada.

Clase	Descripción	# Predios
H0	Habitacional sin edificios	21
H1	Habitacional con edificios de 1 y 2 pisos	43
H2	Habitacional con edificios de 3 o más pisos	49
NH	No Habitacional	47
	TOTAL	160

El número de entrenamiento es relativamente pequeño considerando que el total de predios de la delegación es de 44,250. La clasificación se llevó a cabo sobre seis variables como predictores:

1. Máximo número de pisos (NFloor).
2. Área construida (BArea).
3. Espacio habitacional (HSpace).
4. Número de edificios en el predio (NBuild).
5. Porcentaje de área construida (PBuild).
6. Área del predio (Area).

Clasificación con Isodata

Llevada a cabo con la función Isodata en Matlab, con el vector de parámetros: [4 4 20 0.6 5 0.15 0.06 5 22000], de tal forma que entregara una agrupación en 4 clases. Antes de correr el algoritmo los datos fueron normalizados considerando el valor máximo de cada predictor como 1, ajustando los demás valores al rango de 0 a 1.

Clasificación con MLC

Los datos de entrada fueron la matriz de predios a clasificar, la matriz de entrenamiento y el vector de clasificación de los datos de entrenamiento, sin más, para obtener el vector de clasificación de todos los predios. MLC se implementó en MATLAB.

Clasificación con SVM

Ya que el clasificador SVM es para 2 clases únicamente, y en este caso se trabajó con 4, fue necesario entrenar una SVM por cada par de clases que se pudieran formar, en este caso 6. De esta forma se aplicaron las 6 SVM a cada observación (a cada predio), teniendo entonces, por cada observación, seis clases asignadas o “votos”, asignándole como clase definitiva aquella que tuviera más votos. Hubieron observaciones que presentaron un número igual de votos en dos clases o más. Se probaron dos estrategias para solventar esto, primero, en el caso de que se tuvieran dos clases con igual número de votos, se tomaba como “voto de calidad” el obtenido por la SVM que involucrara a las clases empatadas en votos; y la segunda, construir un segundo grupo de 6 SVM, pero con un número de iteraciones mayor al llevar a cabo el entrenamiento, y de esta forma tener 12 votos de clase por cada observación, al considerar los dos grupos conjuntamente. El primer grupo de 6 SVM fue construido con 10 iteraciones (el valor que está por default en la función) y el segundo con 50 iteraciones.

La primera estrategia resolvió todos los casos en los cuales se presentaron dos clases con igual número de votos, siempre y cuando se hayan tomado sólo 6 SVM (el grupo de 10 ó el grupo de 50 iteraciones). El hecho de tomar conjuntamente las 12 SVM complicó la clasificación puesto que, aunque ciertamente se presentaron menos casos con empate en tres clases, se presentaron más casos con empate en dos clases, y la gran mayoría de ellos no se pudo resolver con el “voto de calidad”, puesto que al haber dos SVM que involucraban a las clases empatadas, una de ellas clasificaba la observación en una clase y la otra en la otra clase. De 739 casos sólo se resolvieron 38. Los resultados se muestran en la tabla siguiente; los números entre paréntesis son los casos de dos clases empatadas que se pudieron resolver, y no se consideran en las suma para el total, por estar ya considerados en alguna clase.

Evaluación de la precisión

Para seleccionar la muestra de prueba se generó una malla formada por cuadros de 9 hectáreas (300 X 300 m) que cubriera toda la delegación. Para cada cuadro se generó un índice de diversidad de usos de suelo, de tal forma que entre mayor fue el índice, más equilibrado respecto a la presencia de diferentes usos de suelo era ese cuadro. Se seleccionaron 20 cuadros con el mayor índice. De manera aleatoria se seleccionaron 50 predios para cada una de las clases H1, H2 y NR. Considerando que en toda la Delegación Cuauhtémoc únicamente existen 30 predios de clase H0, estos mismos, a pesar de estar incluidos algunos de ellos en la muestra de entrenamiento, fueron incluidos también en la muestra de prueba. De tal forma que la muestra de prueba tiene 180 predios.

Con la muestra de prueba y las clasificaciones hechas con cada clasificador fueron generadas matrices de confusión y calculados índices kappa. Por cada clasificador se consideran los casos de utilizar la muestra de prueba incluyendo los predios con clase H0 (Habitacional de cero pisos) y sin ellos.

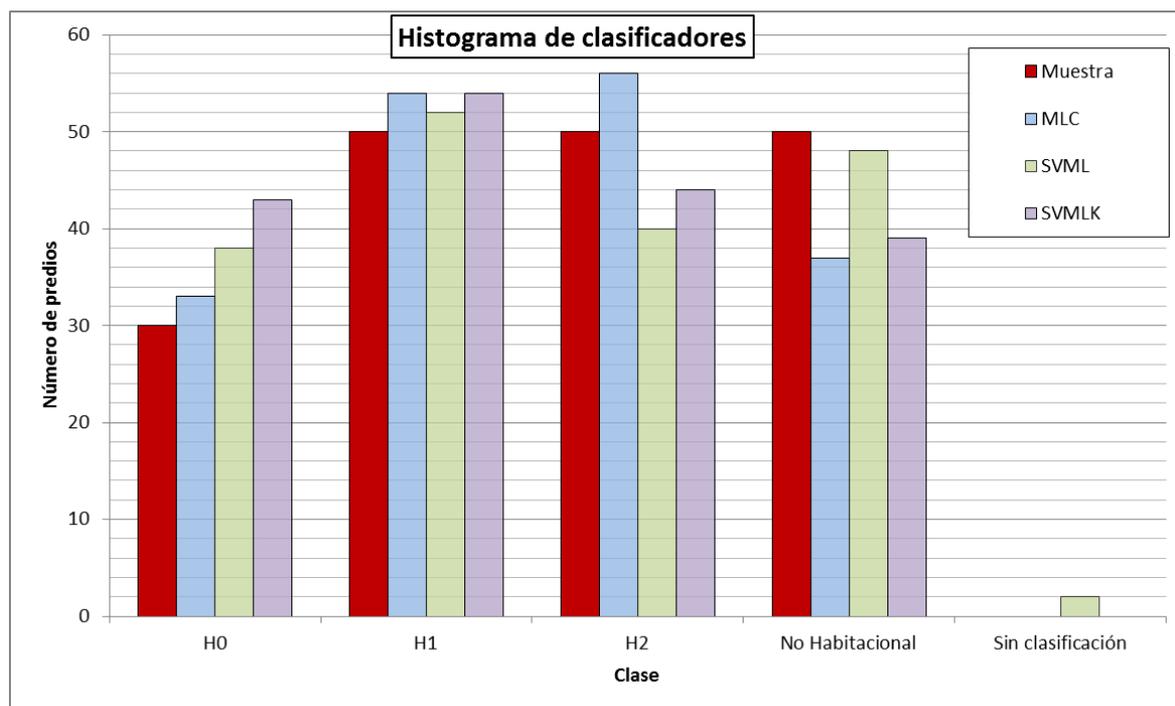


Ilustración 11. Precisión general de clasificación de cada clase mediante cada uno de los métodos probados.

Tabla 11. Matriz de confusión de la clasificación resultante con MLC.

MLC	Usos Suelo	Habitacional H0	Habitacional H1	Habitacional H2	No Habitacional	TOTAL
Habitacional H0		30	0	0	3	33
Habitacional H1		0	44	1	9	54
Habitacional H2		0	2	37	17	56
Habitacional No		0	4	12	21	37
TOTAL		30	50	50	50	180

Tabla 12. Matriz de confusión de la clasificación resultante con SVML.

SVML	Usos Suelo	Habitacional H0	Habitacional H1	Habitacional H2	No Habitacional	Sin clasificación	TOTAL
Habitacional H0		30	2	0	6	0	38
Habitacional H1		0	39	2	11	0	52
Habitacional H2		0	0	26	14	0	40
Habitacional No		0	7	22	19	0	48
Habitacional Sin clasificación		0	2	0	0	0	2
TOTAL		30	50	50	50	0	180

Al evaluar el método de ISODATA se observó un desempeño muy pobre, con un índice kappa negativo, por lo que se optó por descartarlo de cualquier análisis subsecuente. Los resultados de la evaluación de 18 •

la precisión revelan (Tabla 11 a la Tabla 13) que MLC es el mejor método para estimar el uso de suelo a nivel de predio. Aun así, el mejor valor del índice kappa (Tabla 14 no es satisfactorio (64%)) ya que se está asumiendo una extracción perfecta de edificios, es muy difícil, como se ha demostrado en otros estudios (Silván-Cárdenas & Wang, Extraction of Buildings Footprint from LiDAR Altimetry Data with the Hermite Transform, 2011).

Tabla 13. Matriz de confusión de la clasificación resultante con SVMLK.

Uso Suelo SVMLK	Habitacional Ho	Habitacional H1	Habitacional H2	No Habitacional	TOTAL
Habitacional Ho	30	2	0	11	43
Habitacional H1	0	37	8	9	54
Habitacional H2	0	6	25	13	44
No Habitacional	0	5	17	17	39
TOTAL	30	50	50	50	180

Tabla 14. Índices kappa, calculados con la muestra de prueba.

Clasificador	Kappa	Precisión general	RMSE por clase				
			Habitacional Ho	Habitacional H1	Habitacional H2	No habitacional	Sin Clasificación
MLC	0.6409	73.3333	0.1291	0.2981	0.4216	0.5000	0.0000
SVML	0.5103	63.3333	0.2108	0.3651	0.4595	0.5774	0.1054
SVMLK	0.4732	60.5556	0.2687	0.4082	0.4944	0.5528	0.0000

Extracción Automática de Edificios

Inicialmente se intentó actualizar la capa de edificios del catastro usando los datos de altimetría sin embargo, se observó que el polígono de catastro presentaba, en muchos lugares, desplazamientos no sistemáticos de hasta 2 metros con respecto a los datos de altimetría LiDAR por lo que se optó por aplicar un método de extracción automática de edificios para generar por completo la geometría actualizada a la fecha de toma de LiDAR.

La extracción automática de edificios se basó en el modelo digital de alturas producido a partir de la nube de puntos LiDAR (Apéndice C) así como de la clasificación de cobertura generada a partir de las imágenes multiespectrales (Apéndice D). Con estos se genera un modelo digital de alturas de edificios, mediante el enmascaramiento de la vegetación y el suelo. Luego se aplica un método de segmentación por cuencas al gradiente de dicho modelo de alturas para extraer segmentos de edificios. Los segmentos con altura menor a 2.0 metros se eliminan, ya que estos por norma no podrían constituir un edificio. Así mismo, los segmentos aislados de área menor a 16 m² se eliminan por la misma razón. Solo se preservan segmentos pequeños que representan estructuras en azoteas tales como tinacos, estacionarios, bodegas, etc. Estos últimos se unen con el segmento vecino más grande. La segmentación resultante se convierte a vector usando una función SEG2SHP desarrollada en MATLAB (Ilustración 12).



Ilustración 12. Ejemplo de extracción de edificios a partir de la segmentación del modelo digital de alturas. A la izquierda se muestra la segmentación y a la derecha, los polígonos generados superpuestos en el modelo.

El número y forma de los polígonos producidos generalmente no corresponde con los polígonos del catastro (Ilustración 12). Esto se debe a muchas causas tales como como la sobre segmentación, oclusión de los edificios por vegetación, ruido en los datos, etc. Sin embargo, para efectos de cuantificación del espacio habitable la extracción de edificios arroja suficiente información. A la nueva capa de edificios se le anexaron los valores de altura, área, número de niveles y espacio habitable. También se ligó con la capa de uso de suelo del 2000 para obtener la clave catastral asociada a cada polígono extraído, esto con el fin de poder cuantificar el espacio habitable por predio. Sin embargo esto último resultó ser muy problemático muchos de los edificios extraídos caían en más de un predio.

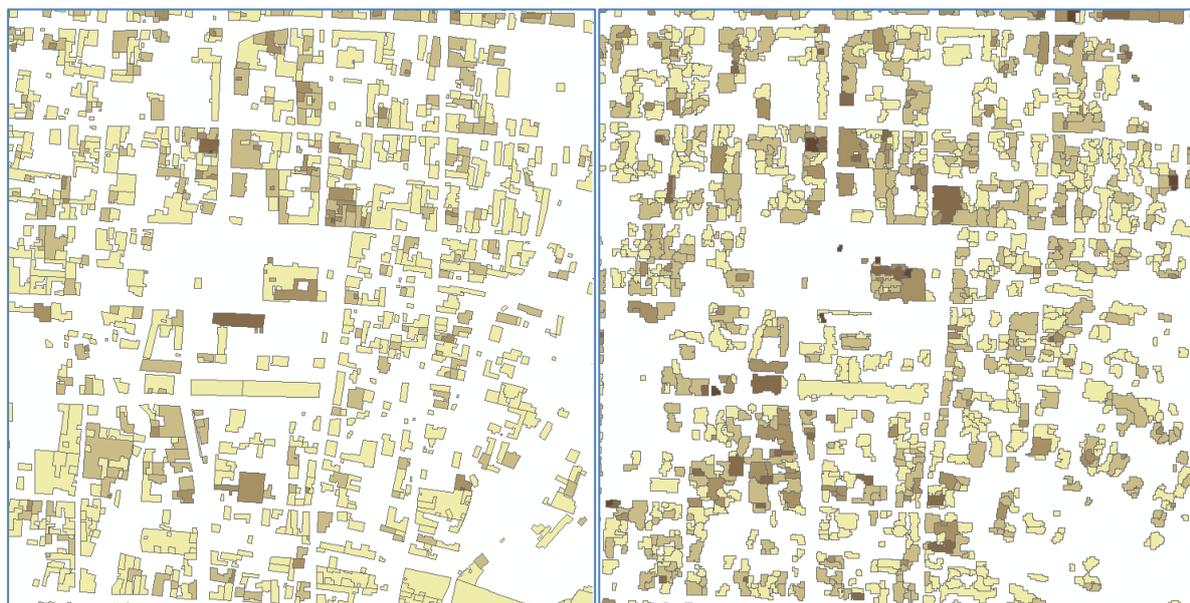


Ilustración 13. Polígonos de edificios del Catastro al 2000 (izq.) y polígonos extraídos con datos LiDAR del 2007 (der).

Estimación del número de pisos

Se desarrolló un modelo para estimar el número de niveles a partir de la altura LiDAR. Dado que los datos de niveles no están actualizados a la fecha de toma de los datos LiDAR, es de esperar una gran discrepancia. No obstante, se estimó que la gran mayoría de los edificios se mantuvieron sin cambios entre las fechas del catastro (2000) y la fecha de adquisición de LiDAR (2008). Esto se puede corroborar en los diagramas de dispersión entre alturas y número de pisos (Ilustración 14).

Para determinar el modelo se intentó un ajuste por mínimos cuadrados con y sin transformación logarítmica. Sin embargo, el modelo resultaba fuertemente sesgado por los edificios de recién creación ($N_{\text{Floors}} = 0$ y $L_{\text{Height}} > 2.5$), por lo que se optó por ajustar visualmente un modelo piezo lineal al diagrama de dispersión.

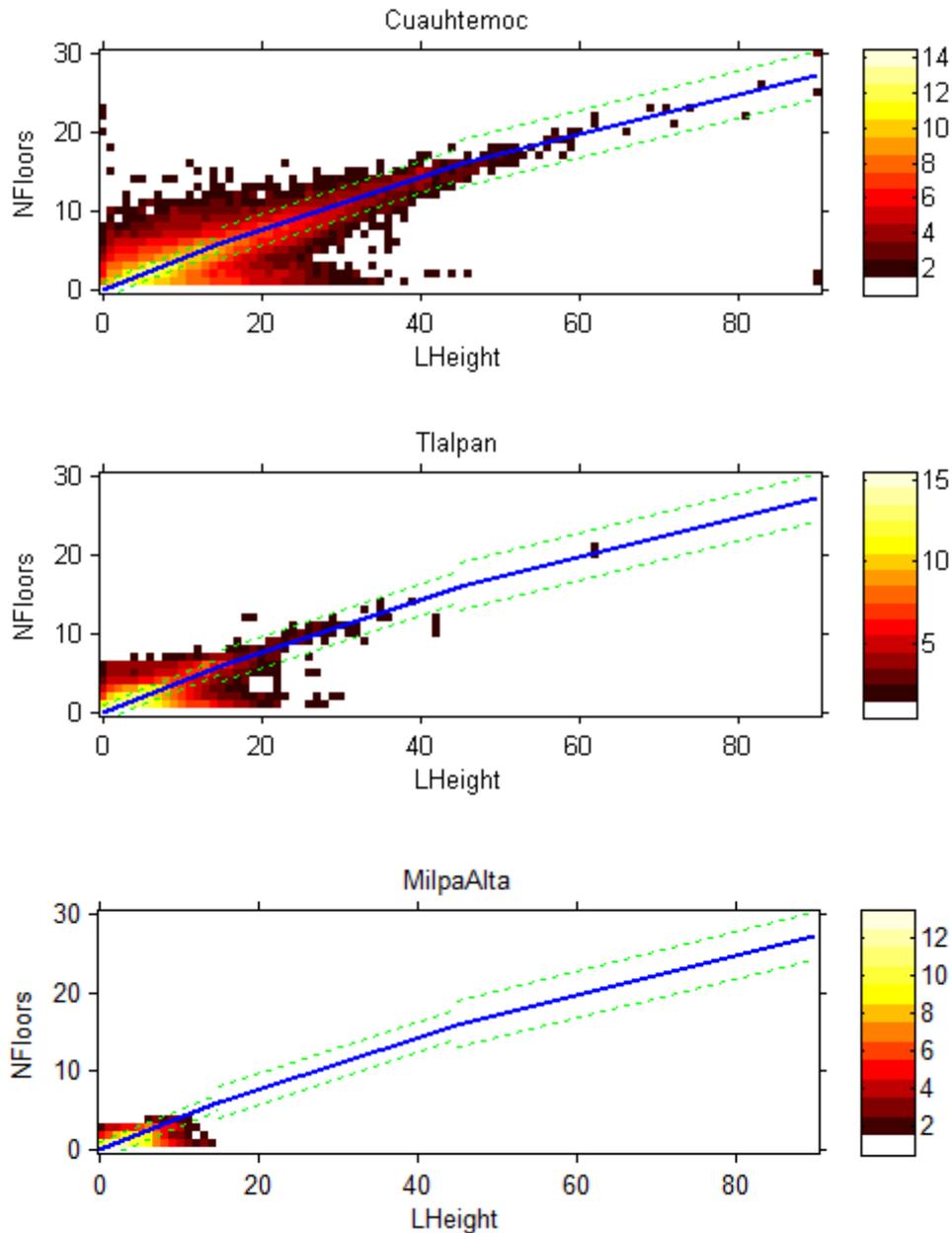


Ilustración 14. Histogramas bidimensionales de altura-pisos de edificios para cada delegación. La línea azul representa el modelo piezolineal ajustado. Las líneas verdes indican una tolerancia empírica.

El modelo ajustado se puede expresar matemáticamente como:

$$n = \begin{cases} [2h/5], & \text{si } h < 15 \\ [(h + 3)/3], & \text{si } 15 \leq h < 45 \\ 0 & \\ [(h + 19)/4], & \text{si } h \geq 45 \end{cases}$$

donde n es el número de niveles del edificio, h es la altura del edificio, y los corchetes indican redondeo al entero más cercano.

Estimación de viviendas

Como se presentó arriba, las pruebas realizadas para extracción de uso de suelo a partir de características de los edificios no fueron lo suficientemente satisfactoria en términos del índice kappa (64% de precisión máxima). Afortunadamente, en las pruebas de estimación de vivienda se encontró que, al menos para Tlalpan y Milpa Alta, sería posible estimar viviendas a partir del espacio habitable en edificios sin importar el uso de suelo. Por lo tanto se procedió a estimar las viviendas a partir del espacio habitable estimado mediante PR, sin considerar el uso de suelo.

En este caso se realizó el ajuste de los modelos V7 y V8 para los datos derivados con LiDAR (ESPT10) para estimar las viviendas del 2010 (VIV10). Adicionalmente, se probó el enfoque incremental para el caso de viviendas. En particular se desarrollaron los modelos V9 y V10 (Tabla 15) en los cuales se relaciona el incremento de viviendas con el incremento de espacio habitable total (INCESPT). La diferencia entre estos modelos consiste en que el modelo V10 emplea los incrementos transformados. A diferencia de los modelos de población donde se aplicaron la transformación logarítmica e incrementos relativos, en este caso se usó vía una raíz cúbica (R3INCESPT y R3INCVIV). La razón usar una transformación cúbica obedece a la observación de Nordbeck sobre el significado del exponente en una relación alométrica, el cual corresponde a la dimensión de la variable (Nordbeck, 1971). En este caso, tanto el número de viviendas como el espacio habitable son variables tridimensionales, por lo que la raíz cúbica convierte las variables originales en variables unidimensionales. De hecho, en una comparación (datos no mostrados) con otros tipos de transformaciones, como logarítmica, se encontró que la raíz cúbica se desempeñó mejor.

Tabla 15. Modelos incrementales para estimación de viviendas.

Modelo	Expresión	Relación	Variable(s) independiente(s)
V9	INCVIV = a*INCESPT+b	Lineal	Espacio habitable total (Todos los usos de suelo)
V10	R3INCVIV = a*R3INCESPT + b	Cúbica	Espacio habitable total (Todos los usos de suelo)

Selección del modelo

Los resultados de ajuste y de errores de estimación obtenidos para los cuatro modelos se muestran en la Tabla 16 y los diagramas de dispersión se muestran de la Ilustración 15 a la Ilustración 17. Como se esperaba, los modelos V7 y V8 ajustaron muy bien los datos para Milpa Alta ($R_2 > 0.9$), incluso mejor cuando se emplearon los edificios del catastro; mientras que se desempeñan moderadamente bien para el caso de Cuauhtémoc y Tlalpan. El desempeño relativamente bajo observado para Cuauhtémoc y Tlalpan se debe al aglutinamiento de usos de suelos particulares en ciertas áreas, tales como corredores comerciales, colonias habitacionales, etc.. Además del uso de suelo, la extracción de edificios en estos casos representa un reto, sobre todo debido relativamente baja resolución de los datos Lidar. La misma tendencia se observó con el ajuste de los modelos V9 y V10 pero con valores menores ($R_2 < 0.75$). En contraste al nivel de ajuste, la capacidad de generalización de los modelos

ocurre a la inversa, es decir, los modelos V9 y V10 presentan menores errores que los modelos V7 y V8. En términos del error absoluto relativo destaca el desempeño consistentemente superior del modelo V10 con las tres delegaciones (MARE < 17%). Esta aparente contradicción indica que los modelos V7 y V8 tienden a sobre ajustar los datos, por lo que deben ser evitados.

Tabla 16. Resultados de la estimación de viviendas mediante edificios extraídos con Percepción Remota. En todos los casos se eliminaron datos con crecimiento en espacio habitable o viviendas anómalos. En cada delegación se indica el número de AGEBS empleados para calibración del modelo (n) así como el número de AGEBS para evaluación (N).

Delegación	Modelo	A	b	R ² ^a	MAE ^b	MARE ^b
Cuauhtémoc n=48, N=101	V7	0.003	511.08	0.52	419	30%
	V8	0.51	0.38	0.52	416	34%
	V9	0.001	433.89	0.12	254	20%
	V10	0.02	6.84	0.22	170	18%
Tlalpan n=45, N=49	V7	0.001	285.71	0.45	258	29%
	V8	0.96	-2.51	0.63	268	36%
	V9	0.0005	82.58	0.25	85	10%
	V10	0.06	1.57	0.24	76	10%
Milpa Alta n=21, N=21	V7	0.01	-3.09	0.93	48	26%
	V8	0.83	-1.33	0.96	42	17%
	V9	0.00	51.91	0.71	86	17%
	V10	0.12	1.12	0.75	95	15%

De acuerdo a las observaciones anteriores el modelo V10 es el que resulta más apropiado para estimar viviendas a partir del incremento del espacio habitable. El modelo explícito se puede escribir de la siguiente manera.

$$V = V_0 + \left(a^3 \sqrt{(S - S_0)} + b \right)^3$$

Donde $S - S_0$ denota el incremento del espacio habitable correspondiente a un incremento $V - V_0$ en viviendas.

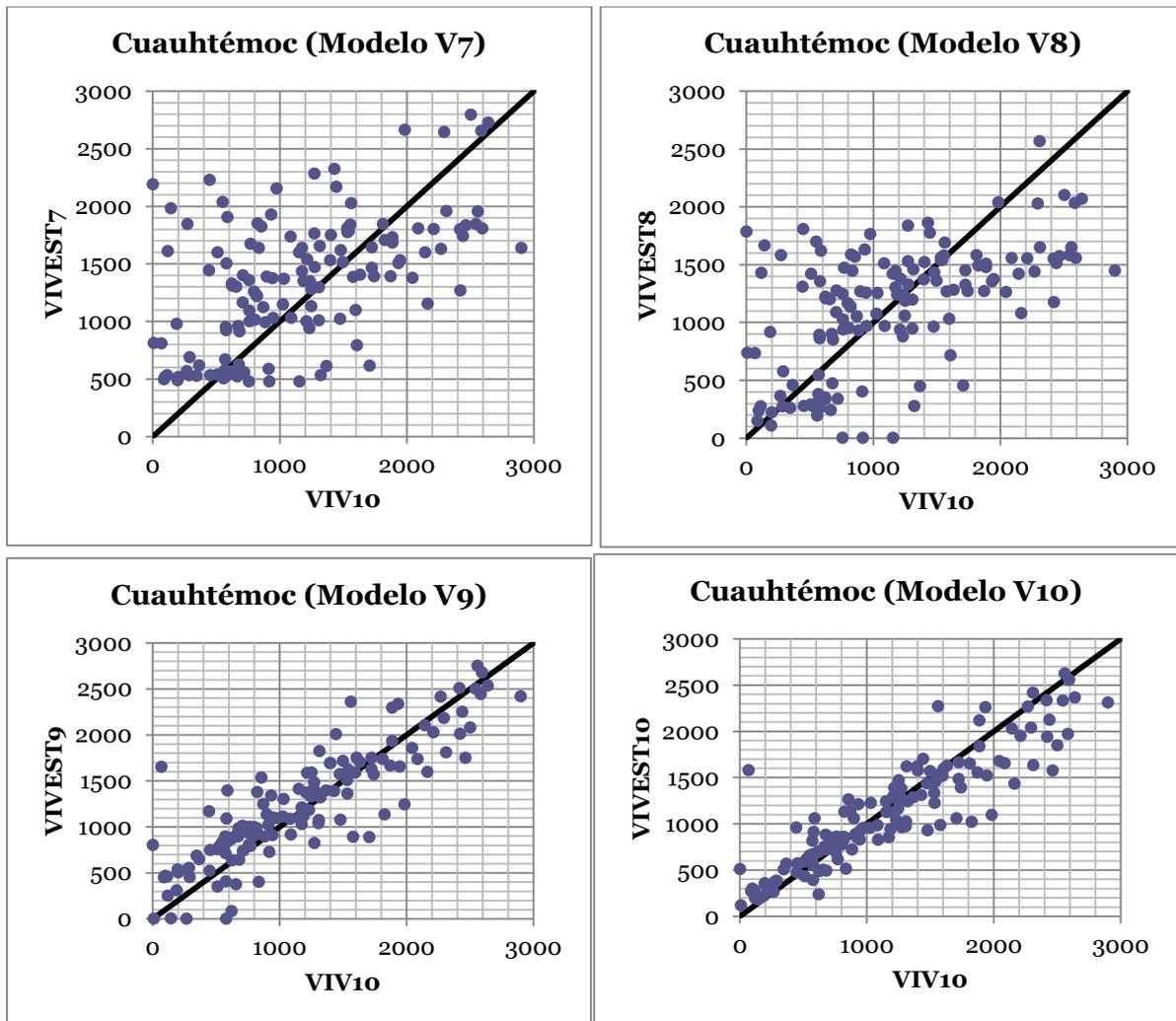


Ilustración 15. Diagramas de dispersión entre viviendas observadas y viviendas estimadas mediante los modelos de estimación de viviendas para Cuauhtémoc, usando datos de espacio habitable estimado con PR.

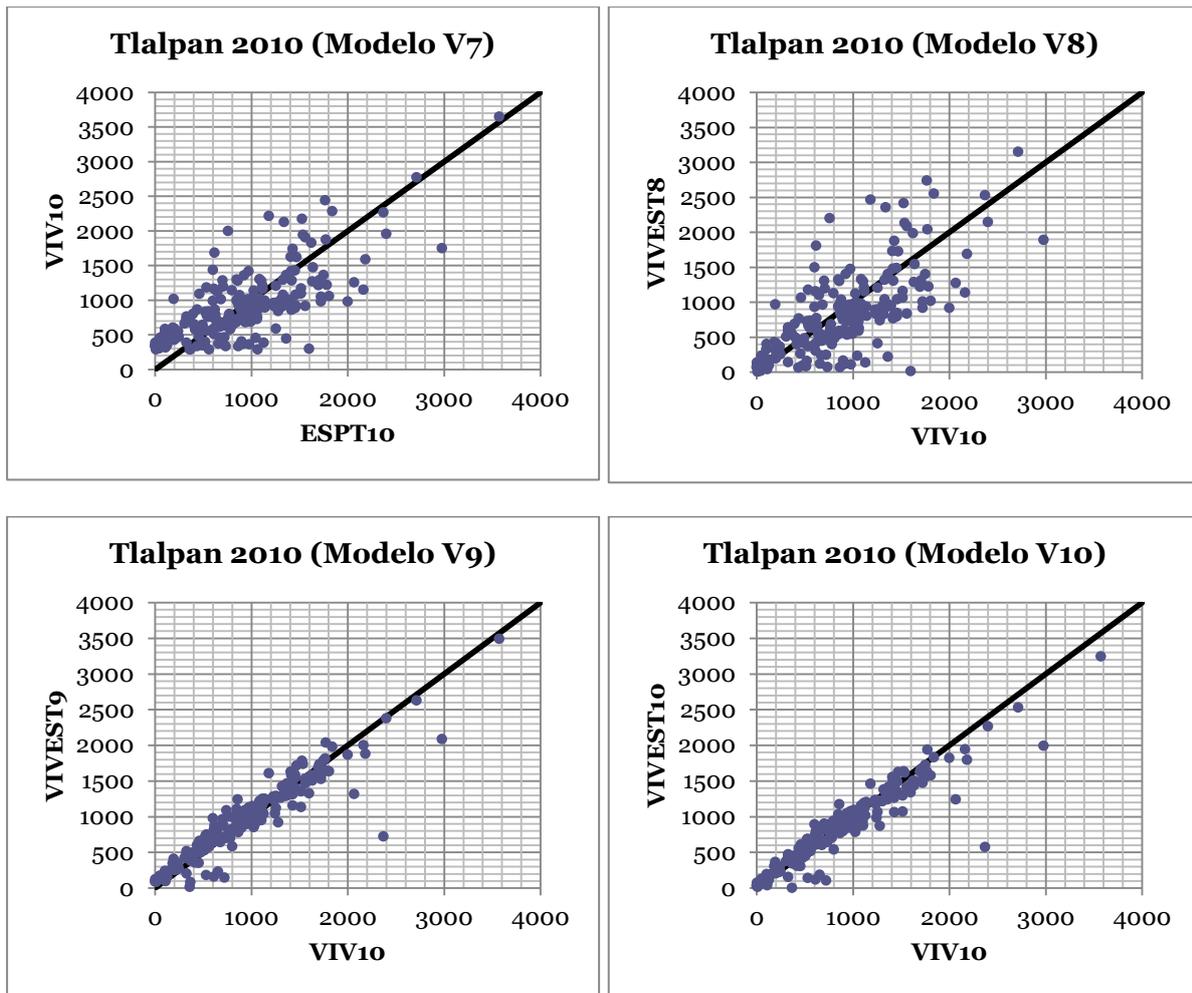


Ilustración 16. Diagramas de dispersión entre viviendas observadas y viviendas estimadas mediante los modelos V7 a V10 para Talpan, usando datos de espacio habitable estimado con PR.

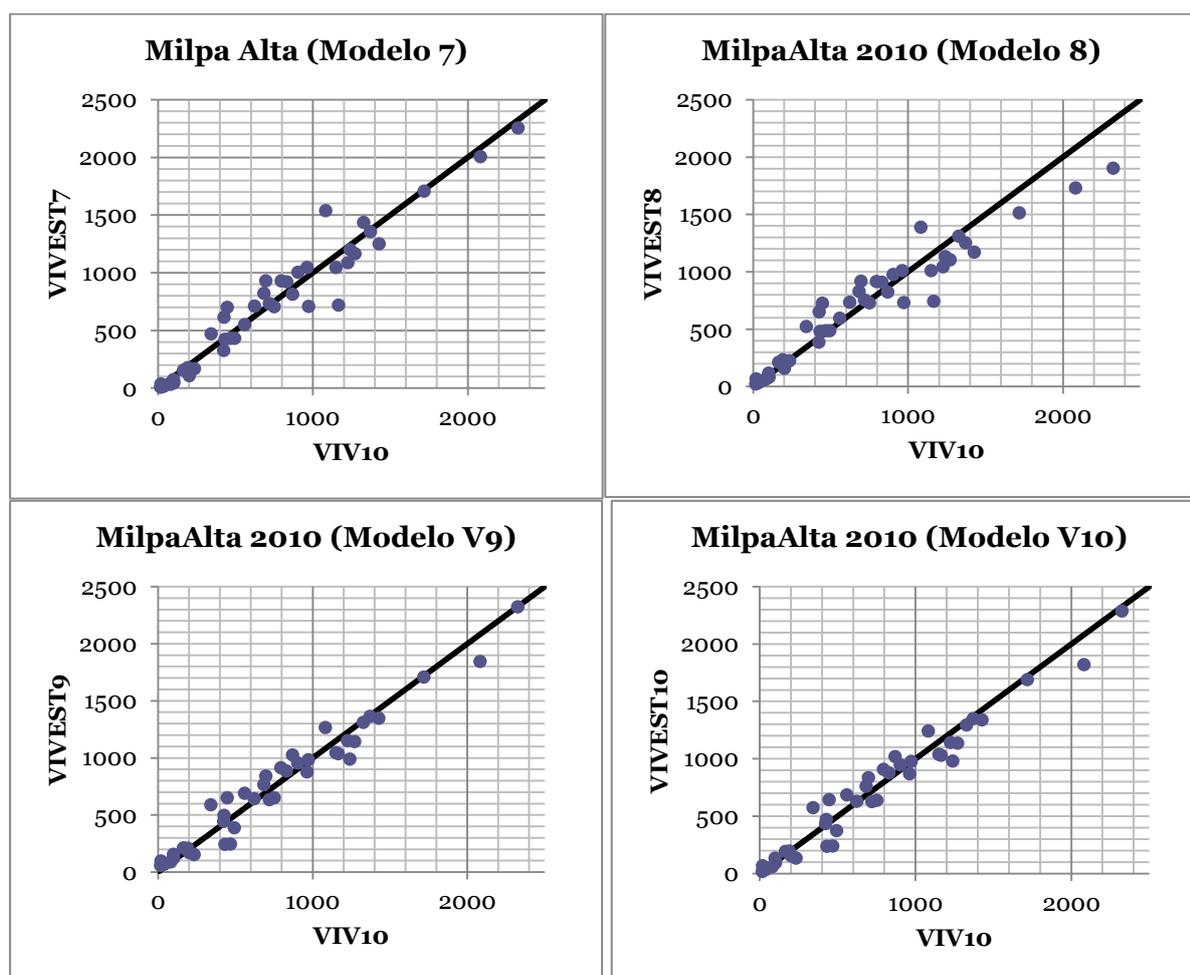


Ilustración 17. Diagramas de dispersión entre viviendas observadas y viviendas estimadas mediante los modelos V7 a V10 para MilpaAlta, usando datos de espacio habitable estimado con PR.

ESTIMACIÓN DE POBLACIÓN CON PR

Delegación	Viviendas	MAE	MARE
Cuauhtémoc	V7	800	31%
	V8	877	37%
	V9	524	18%
	V10	317	15%
Tlalpan	V7	1013	33%
	V8	908	37%
	V9	317	12%
	V10	288	12%
Milpa Alta	V7	291	15%
	V8	293	16%
	V9	287	15%
	V10	309	14%

Tabla 17. Desempeño de la estimación de población del 2010 con el modelo P5, pero estimando viviendas con los modelos V7 a V10 y con datos de PR. Los errores se calcularon sobre todas las AGEBS, sin importar si fueron usadas en la calibración de los modelos.

Para estimar la población a partir de PR, se empleó el Modelo P5 con valores de vivienda estimados para el 2010 a partir de los modelos V7, V8, V9 y V10. Los errores de estimación se presentan en la Tabla 17.

Como era de esperarse, los errores de estimación de viviendas se transfieren a la

estimación de población. Éstos incrementaron con respecto a los observados cuando se emplean los valores de número de viviendas reales para el caso V7 y V8. Con todo ello, el error relativo se mantuvo

por debajo de 37%, en el peor de los casos cuando se emplearon modelos no incrementales (V7 y V8). Nuevamente los modelos con mejor desempeño fueron los modelos incrementales (V9 y V10), los cuales presentaron errores relativos por debajo del 18%.

Consistente con el desempeño de los modelos de estimación de vivienda, el mejor desempeño para estimación de población fue la composición de modelos P5-V10, con el mayor error relativo para Cuauhtémoc (18%), seguido de Milpa Alta (15%) y Tlalpan (12%).

Con el fin de conocer la distribución espacial del error, se generaron mapas de error relativo para la composición P5-V10, los cuales se muestran en la Ilustración 18 a la Ilustración 19 para Cuauhtémoc, Tlalpan y Milpa Alta, respectivamente. En estos mapas se puede observar que diferentes tipos de errores se concentran en diferentes áreas. Al comparar el mapa de errores con los mapas de uso de suelo, se observa cierta concordancia con la distribución del uso de suelo habitacional (por ejemplo, vea la Ilustración B-23).



Ilustración 18. Error relativo de la estimación de población en AGEBs de la delegación Cuauhtémoc.

En términos cualitativos, se observó que las AGEBs con predominancia de uso de suelo habitacional presentaron una subestimación entre el 10 y el 50%, mientras que los AGEBs con predominancia de uso de suelo no habitacional se sobre estimó la población en diferentes proporciones.

CONCLUSIONES

La evidencia encontrada en este estudio apunta a que es factible estimar la población en áreas pequeñas a partir de las viviendas, las cuales a su vez se pueden inferir a partir de PR. Para determinar hasta qué punto es importante el estado de densificación del área donde se quiere estimar la población, se corrieron pruebas para tres delegaciones representativas de densidades alta, media y baja. Se encontró que se puede construir un modelo de estimación que prácticamente no depende de la densidad de población (Modelo P5). En otras palabras, si el número de viviendas se puede determinar con mucha precisión, entonces es posible hacer una estimación de la población con un error relativo del 5%.

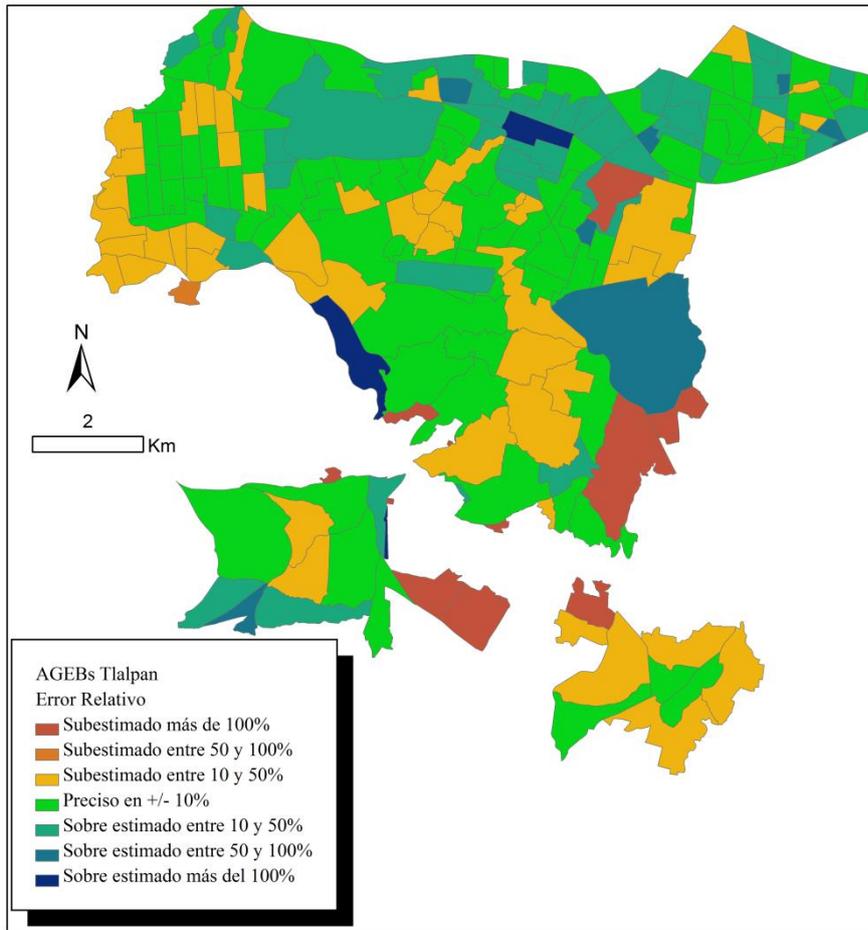


Ilustración 20. Error relativo de la estimación de población en AGEBs de la delegación Tlalpan.

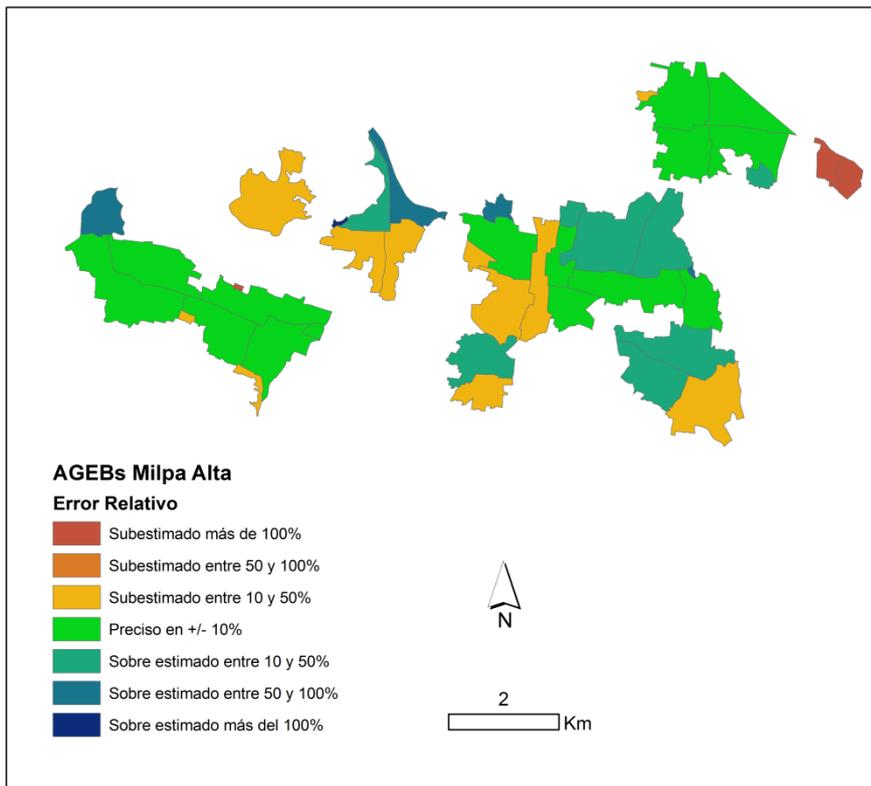


Ilustración 19. Error relativo de la estimación de población en AGEBs de la delegación Milpa Alta.

En contraste, las estimaciones de vivienda mostró errores de alrededor de 20% en promedio. Los mejores modelos fueron aquellos que no emplearon la información de uso de suelo, pero con un enfoque incremental. Es decir, que requieren información de edificios y viviendas, tanto para una fecha previa, como para la fecha de estimación. Es estos casos el nivel de error fue tan bajo como el 15% en términos relativos. Debe notarse que la características de los datos alimentados en estos casos no fue consistente para ambas fechas, ya que no se contó con datos LiDAR e imágenes multiespectrales en 2000, ni datos de catastro actualizados al 2020. Se puede presumir que esto debió ser un factor para que el nivel de error de las estimaciones no bajara aún más, como en el caso de la estimación de población con datos de vivienda reales.

Las pruebas hechas sobre estimación de viviendas a partir de datos del Catastro mostraron que la información de uso de suelo es de gran importancia para estimar población en el caso de áreas fuertemente sectorizadas, como en el caso de Cuauhtémoc, donde se pueden observar corredores comerciales, colonias

residenciales, parques industriales y zonas de edificios públicos. Al otro extremo, se observó que en áreas donde hay mezcla más homogénea de usos de suelos y donde la estructura vertical urbana no es tan significativa, como en el caso de Milpa Alta, resulta suficiente emplear el número de edificios habitacionales, o el espacio habitable de todos los edificios sin importar su uso de suelo, como indicador de viviendas.

Las pruebas hechas para estimar el uso de suelo a partir de PR resultaron poco favorables, por lo que se prefirió trabajar con modelos que no consideran la información de uso de suelo. Sin embargo, se desarrolló un método para extraer edificios a partir de los datos de altimetría y multiespectrales de sensores remotos. Al incorporar esta información en los modelos, se encontró que es factible estimar el número de viviendas con un nivel de error aceptable ($< 15\%$) para el caso de Milpa Alta, mientras que para los otros casos el error resultó fue de más del doble ($> 31\%$). Hay que considerar que estos niveles de error se obtienen cuando se hace el ajuste en dos etapas: primero para estimar población mediante datos de viviendas reales y luego para estimar viviendas a partir del espacio habitable. Es de esperar que el error pueda ser menor si se hace un ajuste directo entre espacio habitable y población; esto a expensa de complicar la optimización de parámetros ya que se imposibilitaría la regresión lineal simple.

Una alternativa para evitar la clasificación de uso de suelo que se requiere en sitios donde la población se concentra en usos específicos sería emplear un método no supervisado para clasificar edificios con base en sus características de forma y tamaño, y posiblemente en su relación espacial con respecto a edificios en un vecindario o vías de transporte. A partir de dicha clasificación, se podría cuantificar la contribución de espacio habitable de cada clase y luego incorporar dichas contribuciones en forma independiente en un esquema de optimización mediante el método de regresión lineal paso-a-paso (stepwise linear regression). De esta manera no se determina el uso de suelo explícitamente, pero el método de optimización determinaría en forma automática la clase de edificios que aportan a la variabilidad del número de viviendas. La dificultad de este enfoque estriba en determinar los rasgos distintivos de edificios de viviendas para hacer una mejor clasificación no supervisada.

Al insertar las estimaciones de vivienda hechas a partir de PR en el modelo de estimación de población, se encontró que estos errores se transfieren a la estimación de población resultando ser del mismo orden de magnitud que para viviendas. Además, al observar la distribución de los errores de estimación, se observó un agrupamiento de los tipos de errores (sobre estimación y sub estimación) en función de la densidad de uso de suelo habitacional. Esta observación sugiere que un mejor enfoque de estimación sería clasificar los AGEBS (en dos o tres clases) por nivel de densidad de uso de suelo habitacional y ajustar un modelo distinto para cada clase de AGEB. La lógica de este enfoque es que a nivel de AGEB podría asumirse que, aunque el uso de suelo de algunos predios pueda cambiar, la densidad de uso habitacional se mantendría, por lo que no sería necesaria una capa de uso de suelo actualizada. Este enfoque permitiría cierta variabilidad espacial en los parámetros del modelo, pero también traería algunas dificultades en cuanto al incremento en el tamaño de la muestra requerida para ajustar los modelos, lo cual puede resultar impráctico. Otras alternativas que deberán explorarse es el uso de regresiones espaciales, tal como la regresión con ponderación geográfica, entre otros.

TRABAJOS CITADOS

- Lwin, K. K., & Murayama, Y. (2011). Estimation of building population from LIDAR derived digital volume model. *Spatial analysis and modeling in geographical transformation process*, 87-98.
- Lwin, K., & Murayama, Y. (2009). A GIS approach to estimation of building population for micro-spatial analysis. *Transactions in GIS*, 13(4), 401-414.

- Ordorica, M., & Ibarra, V. (2002). Aplicación del filtro de Kalman para Realizar Estimaciones de población en áreas pequeñas. En M. N. Page, *Entre fenómenos físicos y humanos* (págs. 201--217). México: El Colegio de Mexico.
- Qiu, F., Sridharan, H., & Chun, Y. (2010). Spatial Autoregressive Model for Population Estimation at the Census Block Level Using LIDAR-derived Building Volume Information. *Cartography and Geographic Information Science*, 37(3), 239-257.
- Silván-Cárdenas, J. (2013). A multiscale erosion operator for discriminating ground points in LiDAR point clouds. En J. Carrasco-Ochoa, J. Martínez-Trinidad, J. Salas-Rodríguez, & G. S. Baja (Ed.), *Pattern Recognition. LNCS 7914*, págs. 213-223. Queretaro: Springer Berlin Heidelberg.
- Silván-Cárdenas, J., & Wang, L. (2006). A multi-resolution approach for filtering LiDAR altimetry data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 61(1), 11-22.
- Silván-Cárdenas, J., & Wang, L. (2011). Extraction of Buildings Footprint from LiDAR Altimetry Data with the Hermite Transform. *Pattern Recognition - MCPR'11 Proceedings of the Third Mexican conference on Pattern recognition*. 6718, pp. 314-321. Cancun, Mexico: Springer-Verlag Berlin Heidelberg.
- Silván-Cárdenas, J., & Wang, L. (2014). Multiscale approach for ground filtering from Lidar altimetry measurements. En Q. Weng, *Scale Issues in Remote Sensing* (pág. 18p). USA: John Wiley & Sons.
- Silván-Cárdenas, J., Almazán-González, J., & Couturier, S. (2014). Remote Identification of Housing Buildings with High-resolution Remote Sensing. *Lecture Notes in Computer Sciences*. 8495, págs. 380--390. Cancun: Springer.
- Silván-Cárdenas, J., L. Wang, C. W., Rogerson, P., Feng, T., & Kamphaus, B. (2010). Assessing fine-spatial-resolution remote sensing for small-area population estimation. *International Journal of Remote Sensing*, 31(21), 5605-5634.
- Sridharan, H., & Qiu, F. (2013). A Spatially Disaggregated Areal Interpolation Model Using Light Detection and Ranging-Derived Building Volumes. *Geographical Analysis*, 45(3), 238-258.

Apéndice A. Lista de publicaciones sobre estimación de población mediante análisis de datos geoespaciales

Tabla A-1. Lista de publicaciones sobre estimación de población mediante análisis de datos geoespaciales e imágenes.

AÑO	AUTORES	INSTITUCIÓN	TÍTULO	PAÍS	PUBLICACIÓN
1956	Porter, P. W	Sin dato	Population Distribution and Land Use in Liberia	República de Liberia	Tesis
1971	S. Y. Hsu	Sin dato	Population Estimation	Sin dato	Revista
1974	Steven P Kraus, Leslie W Senger	Department of Geography, University of California at Santa Barbara	Estimating population from photographically determined residential land use types	Estados Unidos	Revista
1977	C. P. Lo y R. Welch	Sin dato	Chinese Urban Population Estimates	Estados Unidos	Revista
1982	Joji Iisaka	Tokyo Scientific Center, IBM, Japan	Population estimation from Landsat imagery	Japon	Revista
1985	Watkin and Morrow Jhones	University of Colorado	Small area population estimates using aerial photography.	Estados Unidos	Revista
1986	C.P. Lo	University of Georgia	Accuracy of population estimation from medium-scale aerial photography	Estados Unidos	Memorias
1986	C.P. Lo	University of Georgia	Settlement population and land use analysis of the north China plain using shuttle imaging radar-A data	Estados Unidos	Revista
1989	C.P. Lo	University of Georgia	A raster approach to population estimation using high-altitude aerial and space photographs	Estados Unidos	Revista
1994	R. C. S. Taragi, K. S. Bisht and B. S. Sokhi	Indian Institute of Remote Sensing	Generating population census data through aerial remote sensing	India	Revista
1995	C.P. Lo	Department of Geography, University of Georgia. USA	Automated population and dwelling unit estimation from high-resolution satellite images: a GIS approach	Estados Unidos	Revista
1997	Henderson and Xia	Dept. of Geogr. & Planning, State Univ. of New York, Albany, NY	SAR applications in human settlement detection, population estimation and urban land use pattern analysis: a status report	Estados Unidos	Revista
1997	Yew Yuan, Richard M. Smith, W. Fredrick Limp	Center for Advanced Spatial Technologies, University of Arkansas	Remodeling census population with spatial information from Landsat TM imagery	Estados Unidos	Revista

AÑO	AUTORES	INSTITUCIÓN	TÍTULO	PAÍS	PUBLICACIÓN
1997	Paul Sutton	Department of Geography, University of California at Santa Barbara, California	Modeling Population Density With Night-Time Satellite Imagery And Gisery And Gis	Estados Unidos	Revista
1997	Paul Sutton, Dar Roberts, Chris Elvidge, and Henk Melj	Department of Geography, University of California at Santa Barbara	A Comparison of Nighttime Satellite Imagery and Population Density for the Continental United States	Estados Unidos	Revista
2000	Jerome E. Dobson, Edward A. Bright, Phlllp R. Coleman, Richard C. Durfee, and Brian A. Worley	Oak Ridge National Laboratory	LandScan:A Global Population Database for Estimating Populations at Risk	Global	Revista
2000	P. Sutton	Cooperative Institute for Reasearch in Enviromental Sciences	Census fom Heaven: an estimate of the global population using night-time satellite imagery	Estados Unidos	Revista
2002	M. Ordorica Mellado y V. Ibarra Vargas	COLMEX	Aplicación del filtro de Kalman para Realizar Estimaciones de población en áreas pequeñas	México	Capítulo
2002	Jack T. Harvey	University of Ballara, Australia	Population Estimation Models Based on Individual TM Pixels	Australia	Revista
2002	J. T. Harvey	School of Information Technoogy and Mathematical Science, University of Ballarat	Estimating census district populations from satellite imagery: Some approaches and limitations	Australia	Revista
2002	K. CHEN	Natural Hazards Research Centre, Department of Physical Geography, Macquarie University	An approach to linking remotely sensed data and areal census data	Australia	Revista
2002	C. P. Lo	Deparment of Geography, University of Georgia	Urban Indicators of China from Radiance-Calibrated Digital DMSP-OLS Nighttime Images	Estados Unidos	Revista
2003	C.P. Lo and Dale A. Quattrochi	Department of Geography, Geology, and Anthropology, Indiana State University	Land-Use and Land-Cover Change, Urban Heat Island Phenomenon, and Health Implications:A Remote Sensing Approach	Estados Unidos	Revista
2003	Jeremy Mennis	University of Colorado	Genenerating Surface Models of Population Using Dasymmetric Mapping	Estados Unidos	Revista

AÑO	AUTORES	INSTITUCIÓN	TÍTULO	PAÍS	PUBLICACIÓN
2003	Fang Qiu, Kevin L. Woller, and Ronald Briggs	Program in Geographic Information Sciences, School of Social Sciences, University of Texas at Dallas	Modeling Urban Population Growth from Remotely Sensed Imagery and TIGER GIS Road Data	Estados Unidos	Revista
2003	Zhang, BG (Zhang Bao-guang)	Sin dato	Application Of Remote Sensing Technology To Population Estimation	China	Revista
2004	Gonçalves, C. D. A. B., de Marcelhas, Í., Pereira, S. M. N., & da Costa Freitas, C.	Instituto Nacional de Pesquisas Espaciais, Sao Jose dos Campos, Sao Paulo. Brasil	Análise do ambiente residencial urbano visando a inferência populacional a partir do uso de dados de sensoriamento remoto orbital de alta resolução	Brasil	Memorias
2005	Shuo-sheng Wu, Xiaomin Qiu, Le Wang	Texas State University-San Marcos, San Marcos, Texas	Population Estimation Methods in GIS and Remote Sensing: A Review	Estados Unidos	Revista
2005	Guiying Li and Qihao Weng	Department of Geography, Geology, and Anthropology, Indiana State University	Using Landsat ETM+ Imagery to Measure Population Density in Indianapolis, Indiana, USA	Estados Unidos	Revista
2005	Changshan Wu y Alan T. Murray	Department of Geography, University of Wisconsin-Milwaukee	A cokriging method for estimating population density in urban areas	Estados Unidos	Revista
2005	van Cauter, A (van Cauter, A); Kerley, GIH (Kerley, GIH); Cowling, RM (Cowling, RM)	Sin dato	The consequence of inaccuracies in remote-sensed vegetation boundaries for modelled mammal population estimates	Sur Africa	Revista
2006	XiaoHang Liu, Keith Clarke, and Martin Herold	Keith Clarke and Martin Herold are at the University of California, Santa Barbara,	Population Density and Image Texture: A Comparison Study	Estados Unidos	Revista
2006	DENGSHENG LU, QIHAO WENG and GUIYING LI	Center for the Study of Institutions, Population, and Environmental Change, Indiana University	Residential population estimation using a remote sensing derived impervious surface approach	Estados Unidos	Revista
2006	Mitchel Langforda & Gary Higgsa	University of Glamorgan	Measuring Potential Access to Primary Healthcare Services: The Influence of Alternative Spatial Representations of Population	Estados Unidos	Revista

AÑO	AUTORES	INSTITUCIÓN	TÍTULO	PAÍS	PUBLICACIÓN
2006	M.M. Yagoub	Geography Program, College of Humanities and Social Sciences, United Arab Emirates University	Application Of Remote Sensing And Geographic Information Systems (Gis) To Population Studies In The Gulf: A Case Of Al Ain City (Uae)	Emiratos Arabes Unidos	Revista
2006	Shuo-sheng Wu, Xiaomin Qiu, and Le Wang	Department of Geography, Texas State University-San Marcos	Using Semi-variance Image Texture Statistics to Model Population Densities	Estados Unidos	Revista
2007	Wu, Changshan; Murray, Alan T.	Department of Geography, University of Wisconsin- Milwaukee, Milwaukee, WI	Population Estimation Using Landsat Enhanced Thematic Mapper Imagery	Estados Unidos	Revista
2008	Perry J. Hardina, Mark W. Jackson & Ryan R. Jensen	Brigham Young University, Departamento of Geography. Provo. UT, USA	Modelling housing unit density from land cover metrics: a Midwestern US example	Estados Unidos	Revista
2008	Shuo-Sheng Wu, Le Wang y Xiaomin Qiu	Texas State University—San Marcos, University of Buffalo, Missouri State University	Incorporating GIS Building Data and Census Housing Statistics for Sub-Block- Level Population Estimation	Estados Unidos	Revista
2008	C. P. Lo	University of Georgia	Population Estimation Using Geographically Weighted Regression	Estados Unidos	Revista
2008	Florence A. Galeon	University of the Philippines	Estimation Of Population In Informal Settlement Communities Using High Resolution Satellite Image	Estados Unidos	Memorias
2009	Thomas Alan Morton y Fei Yuan	Kansas City Regional Census Center y Department of Geography, Minnesota State University	Analysis of population dynamics using satellite remote sensing and US census data	Estados Unidos	Revista
2009	Lwin, K., & Murayama, Y.	University of Tsukuba	A GIS approach to estimation of building population for micro- spatial analysis	Japon	Revista
2009	Sathya Ramesh	University of North Texas	High resolution satellite images and LiDAR data for small-area building extraction and population estimation	Estados Unidos	Tesis
2010	PINLIANG DONG, SATHYA RAMESH and ANJEEV NEPALI	Department of Geography, University of North Texas	Evaluation of small-area population estimation using LiDAR, LandsatTM and parcel data	Estados Unidos	Revista

AÑO	AUTORES	INSTITUCIÓN	TÍTULO	PAÍS	PUBLICACIÓN
2010	Jose L. Silvan-Cardenas, Le Wang, Peter Rogerson, Changshan Wu, Tiantian Feng And Benjamin D. Kamphaus	Department of Geography, State University of New York.	Assessing fine-spatial-resolution remote sensing for small-area population estimation	Estados Unidos	Revista
2010	Chengbin Deng, Changshan Wu And Le Wang	Department of Geography, University of Wisconsin-Milwaukee	Improving the housing-unit method for small-area population estimation using remote-sensing and GIS information	Estados Unidos	Revista
2010	Fang Qiu, Harini Sridharan, and Yongwan Chun	University of Texas at Dallas	Spatial Autoregressive Model for Population Estimation at the Census Block Level Using LIDAR-derived Building Volume Information	Estados Unidos	Revista
2010	Yunjun Zhan; Yujing Liang	Dept. of Resources & Environ. Eng., Wuhan Univ. of Technol., Wuhan, China	Urban population estimation with remote sensing based on classified density method.	China	Memorias
2010	Lu, Z., Im, J., Quackenbush, L., & Halligan, K.	Satste University of New York	Population estimation based on multi-sensor data fusion	Estados Unidos	Revista
2010	Qiu, Fang, Harini Sridharan, and Yongwan Chun.	University of Texas at Dallas	Spatial autoregressive model for population estimation at the census block level using LIDAR-derived building volume information	Estados Unidos	Revista
2011	Villeneuve, PJ; Goldberg, MS; Burnett, RT; van Donkelaar; A. Chen, H; Martin, RV .	Estudios de Población de la División de Salud de Canadá	Associations between cigarette smoking, obesity, sociodemographic characteristics and remote-sensing-derived estimates of ambient PM2.5: results from a Canadian population-based survey	Canada	Revista
2011	Lu, ZY; Im, J; Quackenbush, L.	Department of Environmental Resources Engineering, State University of New York, College of Environmental Science and Forestry, One Forestry	A Volumetric Approach to Population Estimation Using Lidar Remote Sensing	Estados Unidos	Revista

AÑO	AUTORES	INSTITUCIÓN	TÍTULO	PAÍS	PUBLICACIÓN
2011	Siljander, M; Clark, B.J.F; Pellikka, P.K.E.	Department of Geosciences and Geography. University of Helsinki.	A predictive modelling technique for human population distribution and abundance estimation using remote-sensing and geospatial data in a rural mountainous area in Kenya	Finlandia	Revista
2011	Lwin, Ko Ko, and Yuji Murayama	University of Tsukuba	Estimation of building population from LIDAR derived digital volume model	Japon	Capítulo
2013	Yang, X., Yue, W., & Gao, D.	Zhejiang Institute of Meteorological Sciences	Spatial improvement of human population distribution based on multi-sensor remote-sensing data: an input for exposure assessment.	China	Revista
2013	Irvine, J. M., Kimball, J., Regan, J., & Lepanto, J. A.	Draper Laboratory, Cambridge	Application of commercial remote sensing to issues in human geography.	Afghanistan	Memorias
2013	Dittakan, K., Coenen, F., Christley, R., & Wardeh, M.	University of Liverpool	A Comparative Study of Three Image Representations for Population Estimation Mining Using Remote Sensing Imagery.	Etiopía	Capítulo
2013	Deng, C., & Wu	University of Wisconsin–Milwaukee	Improving Small-Area Population Estimation: An Integrated Geographic and Demographic Approach.	Estados Unidos	Revista
2013	Azar, D., Engstrom, R., Graesser, J., & Comenetz, J.	U.S. Census Bureau & The George Washington University	Generation of fine-scale population layers using multi-resolution satellite imagery and geospatial data.	Pakistan	Revista
2013	Mulatu, D. W., Van der Veen, A., Becht, R., Van Oel, P. R., & Bekalo, D. J.	University of Twente	Accounting for spatial non-stationarity to estimate population distribution using land use/cover. Case study: The Lake Naivasha basin, Kenya.	Kenia	Revista
2013	Batista e Silva, F., Gallego, J., & Laval, C.	aJoint Research Centre, Institute for Environment and Sustainability	A high-resolution population grid map for Europe	Europa	Revista

AÑO	AUTORES	INSTITUCIÓN	TÍTULO	PAÍS	PUBLICACIÓN
2013	Alahmadi, M., Atkinson, P., & Martin, D.	University of Southampton	Estimating the spatial distribution of the population of Riyadh, Saudi Arabia using remotely sensed built land cover and height data.	Arabia Saudita	Revista
2013	Checchi, F., Stewart, B. T., Palmer, J. J., & Grundy, C.	London School of Hygiene and Tropical Medicine	Validity and feasibility of a satellite imagery-based method for rapid estimation of displaced populations.	Bangladesh, Chad, Republica Democrática del Congo, Etiopía, Haití, Kenya y Mozambique	Revista
2013	Sridharan, H., & Qiu, F.	Oak Ridge National Laboratory	A Spatially Disaggregated Areal Interpolation Model Using Light Detection and Ranging-Derived Building Volumes	Estados Unidos	Revista
2013	Kit, O., Lüdeke, M., & Reckien, D.	Potsdam Institute for Climate Impact Research	Defining the bull's eye: satellite imagery-assisted slum population assessment in Hyderabad, India.	India	Revista
2014	Michellier, C., Hanson, E., & Wolff, E.	Sin dato	Urban development and population estimation in central Africa through remote sensing	Africa Central	Memorias
2014	Weng, Q.	Indiana State University	On The Issue of Scale In Urban Remote Sensing.	Estados Unidos	Capítulo
2014	Kubaneck, J., Nolte, E. M., Taubenböck, H., Wenzel, F., & Kappas, M.	Karlsruhe Institute of Technology	Capacities of Remote Sensing for Population Estimation in Urban Areas	Turquía	Capítulo
2014	Greger, K.	University of Tsukuba	Spatio-Temporal Building Population Estimation for Highly Urbanized Areas Using GIS	Japon	Revista
2014	ia, P., Qiu, Y., & Gaughan, A. E.	Louisiana State University	A fine-scale spatial population distribution on the High-resolution Gridded Population Surface and application in Alachua County, Florida.	Estados Unidos	Revista
2014	Kemper, T., & Heinzl, J.	Institute for the Protection and Security of the Citizen	Mapping and Monitoring of Refugees and Internally Displaced People Using EO Data	Varios	Capítulo

Apéndice B. Procesamiento datos vectoriales de catastro

Conversión de formato

Se contaba con datos de catastro actualizados al año 2000 en formato DXF para las 16 delegaciones. Dado que el número de archivos DXF para una sola delegación era significativo (96 para Cuauhtemoc, 185 para Tlalpan y 44 para MilpaAlta), se desarrolló un programa en lenguaje Python para automatizar la conversión usando la interfaz para geo procesamiento de ArcInfo. Para cada carta procesada (archivo DXF), el programa crea dos archivos SHAPEFILES que contiene los polígonos de edificios con número de niveles, y los polígonos de predios con clave de uso de suelo. El conjunto de shapefiles generado se combinan en un solo archivo para cada delegación y rasgo (edificio o predio), generando así un total de 6 archivos.

Nota: Para obtener la cobertura exacta de la delegación fue necesario anexar cartas que se encontraban incluidas en las carpetas de las delegaciones vecinas.

Corrección geométrica

La referencia original del catastro estaba en proyección Lambert (Datum NAD 27), pero se re-proyectó a UTM WGS84 para poderlas sobreponer sobre las imágenes y los datos de altimetría. Además del cambio de proyección se le aplicó un desplazamiento manual a cada capa para empatar visualmente la huella de los edificios con la capa de alturas generadas con LiDAR (Ilustración B-21). Aunque en general se puede observar una correspondencia espacial entre los conjuntos de datos, se observaron desplazamientos localizados que no se pudieron corregir.

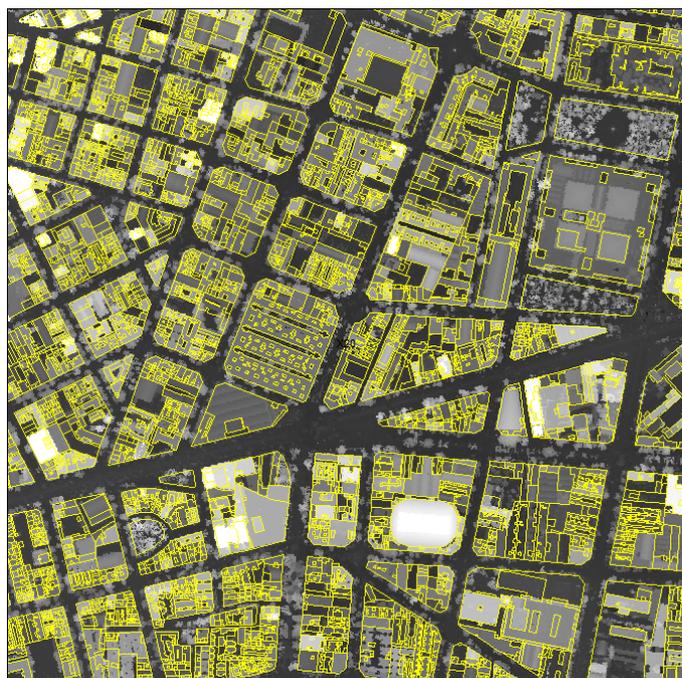


Ilustración B-21. Sobre posición del vector de edificios sobre el modelo digital de elevaciones de LiDAR a 1 m de resolución.

Simplificación y filtrado de polígonos de edificios

Se eliminaron los polígonos que cayeron fuera de los límites de las delegaciones para trabajar únicamente la cobertura necesaria.



Ilustración B-22. Polígonos de edificios para toda la delegación Cuauhtémoc.

También se observó que durante la conversión se generan los polígonos espurios:

1. Polígonos formados por líneas de anotaciones que indican que el o los polígonos se encontraban en proceso de actualización. Dichas líneas ocasionaban la partición de los polígonos sobre los cuales corría. Se editaron manualmente en ArcGIS usando principalmente herramienta Merge.
2. Polígonos formados cuando un edificio corría sobre más de un predio. Se editaron en algunos casos cuando se trataba de edificios complejos.
3. Duplicidad de polígonos por encontrarse en dos archivos DXF distintos. Se eliminó una copia, conservando la que presentaba mejor registro con respecto a los datos de altimetría.
4. Polígonos huecos. Corresponde a la parte no construida dentro de un predio. Se eliminaron por el atributo DocName que se refiere al archivo DXF de donde se extrae el polígono. Este estaba vacío para dichos casos. Posteriormente también se eliminaron aquellos que no presentaban número de pisos de uno o mayor.
5. Polígonos que no representaban edificios. Estos corresponden a fuentes, albercas y glorietas entre otras estructuras urbanas. Fueron eliminados empleado el atributo Texto que se extrajo de la base de datos original y que corresponde al número de niveles, en el caso de edificios, o a una clave en el caso de otras estructuras. Por ejemplo es Texto = "AL" para alberca, etc.
6. Polígonos de áreas insignificantes. Estos fueron fusionados con el polígono vecino con el que compartía el borde más largo. Se fusionaron polígonos con área menor a 4 m². Este criterio se empleó para asegurar que cayera más de un punto LiDAR dentro de cada polígono para el cálculo de alturas, ya que la densidad de puntos era de aproximadamente 1/2 por metro cuadrado.

Cálculo de Atributos para Polígonos de Edificios

Se calcularon los siguientes atributos para cada polígono de edificio.

Tabla B-18. Atributos calculados para la capa de edificios.

Atributo	Tipo	Significado	Cálculo	Uso Principal
Entity	String	Identificador de Edificio	CStr([FID])	Liga con subconjuntos
ParcID	String	Clave catastral	Spatial Join	Liga con predios
NFloors	Short	Número de pisos	CInt([Text])	Clasificación y Análisis
Area	Float	Area del polígono	Field Calculator	Filtrado y Análisis
HSpace	Float	Espacio habitable	[Area]* [NFloors]	Clasificación
Height	Double	Altura nominal	c*[NFloors]	Visualización y Análisis
LHeight	Float	Altura LiDAR	MATLAB	Análisis y Extrusión
LElev	Float	Elevación del Terreno	MATLAB	Extrusión
LPoints	Long	Número de 1ros retornos de edificios	MATLAB	Filtrado y ordenamiento
LNFloors	Short	Número de pisos actual	[LHeight]/c	Clasificación y Análisis
LHSpace	Float	Espacio habitable actual	[Area]* [LNFloors]	Clasificación

ParcID – Se realizó un Outter Spatial Join con el polígono de predios del 2000 usando una capa de centroides de los edificios como Target y la de predio como Join. Se usó la regla WITHIN para relacionar los edificios. El uso de puntos fue conveniente ya que cuando se intentó usar relacionar polígonos con polígonos ocurría que algunos edificios no estaban completamente contenidos en el predios, es decir se traslapaban con más de un predio. Además, para poder usar la clave catastral como liga, se asignó el valor del identificado de rasgo FID cuando esta estaba vacía. Esta liga se hizo tanto para los edificios del catastro como para los edificios detectados con LiDAR.

LHeight – Para el cálculo de este atributo se cargó la capa de polígonos de edificios de la delegación en MATLAB. Luego, para cada cuadro de la maya que se traslapaba con los polígonos de edificios se cargó la nube de puntos clasificada y con valores de alturas pre-calculados durante el filtrado. Para cada polígono dentro del recuadro, se seleccionaron los puntos LiDAR del primer retorno que estaban clasificados como **Construido**. En el caso donde no se seleccionaron puntos, porque la densidad de puntos era muy baja en relación con el tamaño del polígono de edificio, se procedió a seleccionar los puntos más cercanos a los vértices del polígono. La altura se calculó como el promedio de alturas de los puntos del primer cuartil. Se calcularon el número de puntos LiDAR. La tabla de valores calculados se exportó a un archivo ASCII separado por comas, que luego se adjuntó a la base de datos de los polígonos de edificios en ArcGIS. Las estimaciones obtenidas mediante este método fueron empleadas para construir un modelo para el número de niveles, como se explica más adelante.

LElev – Se calculó como el promedio de elevaciones del terreno en puntos LiDAR dentro del BoundingBox de cada edificio.

LPoints – Se calculó como el número de 1ros retornos de puntos interiores al polígono que estaban clasificados como edificio.

Height – Este atributo corresponde a la altura nominal asumiendo una altura de entrepiso constante. Para ello se usó el modelo presentado en la Sección “Estimación del número de pisos”, pero en forma inversa, es decir:

$$h = \begin{cases} 5n/2, & \text{si } n \leq 6 \\ 3n - 3, & \text{si } 6 < n \leq 16 \\ 0 & \\ 4n - 19, & \text{si } n > 16 \end{cases}$$

donde n es el número de niveles (NFloor).

Adicionalmente se hizo una actualización del número de niveles de los edificios considerando el modelo antes citado. Se consideró un margen de error de +/-1 nivel para edificios menores de 15 metros de altura, de +/-2 niveles para edificios de alturas entre 15 y 45, y de +/-3 para edificios con mayor altura (líneas punteadas en verde en la Ilustración 14). Esta consideración fue de acuerdo a la variabilidad observada para número de niveles dada una altura. De esta manera sólo se actualizaron los edificios cuya estimación estuvo fuera de la franja de tolerancia. Además, todos los niveles de aquellos edificios en los que no cayó ningún pixel del modelo de alturas de edificios, sea debido a oclusión por árboles o a errores en la clasificación, se mantuvieron inalterados.

Edición del vector de predios

En el polígono de predios generados se encontraron algunos casos donde no se anexó una clave catastral. Se inspeccionó una muestra usando el sistema de consulta de la SEDUVI y se eliminaron aquellos que no existían como tal en la base de datos de la SEDUVI y el resto se actualizaron de acuerdo a la información disponible en el sistema. Aunque no se hizo una edición exhaustiva de todos los polígonos debido al arduo trabajo que implicaba, se revisaron aquellos predios con dimensiones poco probables o que carecían de información y se corrigieron errores de algunos polígonos malformados durante el proceso de conversión similar a como se editó el vector de edificios.

Interpretación de la clave de Uso de Suelo

Aunque se sabía que la letra inicial de la clave indica el uso principal, no se contaba con las definiciones de éstas, por lo que se realizó una interpretación de imágenes históricas empleando GoogleEarth y Google StreetView para determinar el uso de suelo más probable que había a la fecha del catastro (2000). Como resultado se obtuvieron los siguientes usos:

Tabla B-19. Interpretación la primera letra de clave catastral.

1ra Letra	Uso General	Uso Específico
A	Espacios Abiertos	Potencial Reciclamiento
B	Espacios Abiertos	Agostadero
C	Comercial	Comercio en General
D	Espacios Abiertos	Deportivos
E	Gobierno	Equipamiento y Patrimonio
F	Comercial	Comercial y servicios
G	Comercial	Gasolineras
H	Habitacional	Habitacional
I	Industrial	Parque Industrial
J	Espacios Abiertos	Jardines y Áreas Verdes
K	Comercial	Telecomunicaciones
L	Comercial	Servicios de Hospedaje
M	Espacios Públicos	Mercados y Plazas
N	Espacios Abiertos	Parques, Plazas y Jardines
O	Gobierno	Oficinas de Gobierno
P	Gobierno	Estación Policial
Q	Comercial	Servicios de Entretenimiento
R	Comercial	Servicios de Alimentos
S	Comercial	Servicio de Salud
T	Comercial	Estacionamientos comerciales
U	Espacios Públicos	Espacios públicos y de gobierno
V	Comercial	Servicios funerarios panteones
W	Comercial	Estacionamiento Privado
X	Desconocido	Varios
Y	Espacios Públicos	Iglesias, Templos y Mezquitas
Z	Comercial	Taller mecánico y pensión

Nota: Algunos predios contaban con claves que iniciaban con gato (#), arroba (@) o simplemente no contaban con una clave. A todos estos se les asigno la etiqueta "Otro".

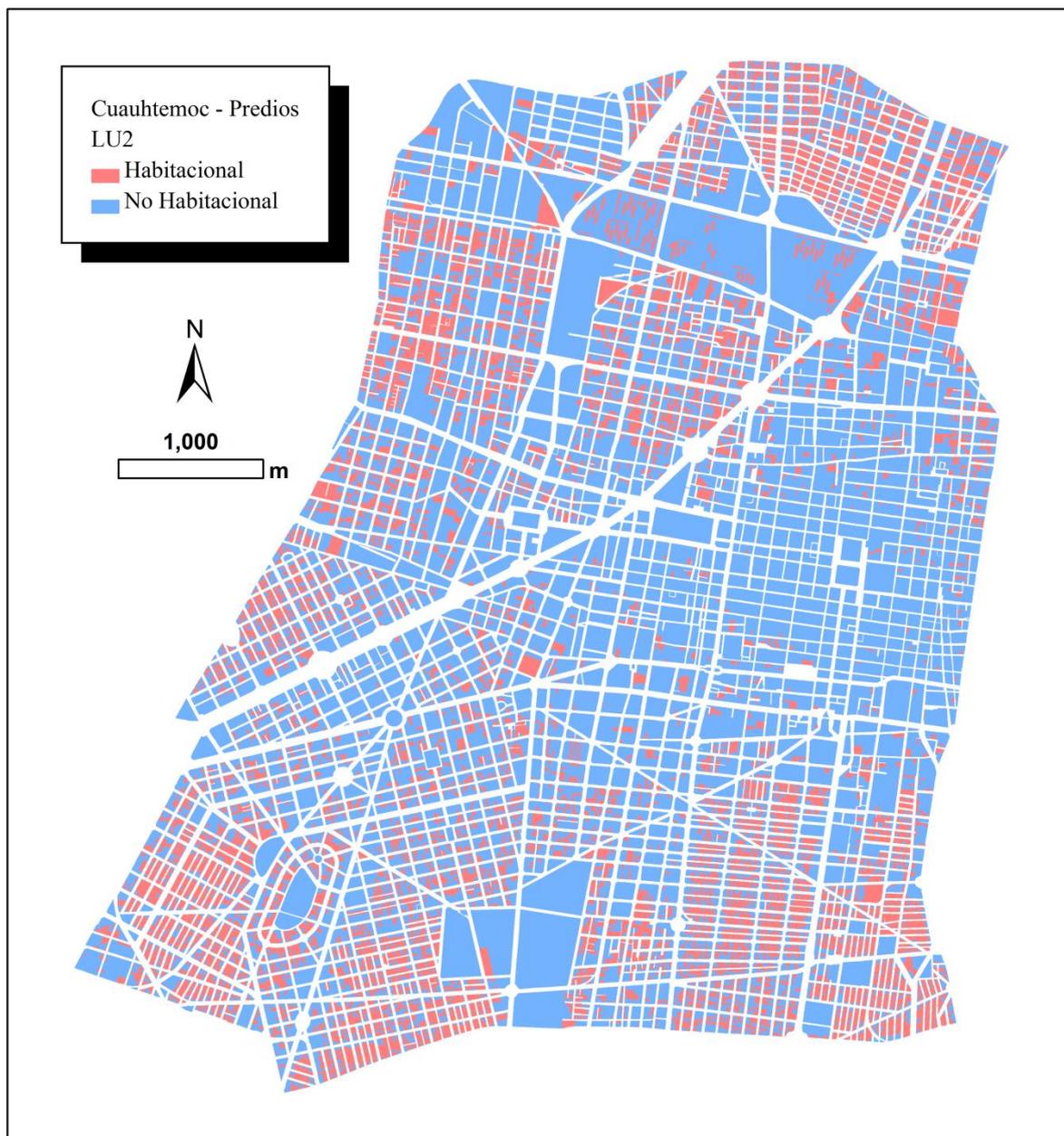


Ilustración B-23. Distribución del uso de suelo habitacional en Cuauhtémoc.

Cálculo de Atributos para Polígonos de Predios

Se calcularon los siguientes atributos para cada polígono de edificio.

Tabla B-20. Atributos calculados para la capa de predios.

Atributo	Tipo	Significado	Cálculo	Uso Principal
Entity	String	Identificador de polígono	CStr([FID])	Liga con subconjuntos
DocName	String	Archivo DXF fuente	Disponible DXF	Rastreo de Fuente
ParcID	String	Clave catastral	Disponible DXF	Liga con predios
BlockID	String	Clave de Manzana	Spatial Join	Liga con manzanas
AGEB	String	Clave de AGEB	Spatial Join	Liga con AGEBs
StreetNum	String	Número exterior	Disponible DXF	Búsqueda

LUKey	String	Clave de Uso	Disponible DXF	Clasificación
LU1	String	Uso General	Field Calculator	Selección de predios
LU2	String	Uso Específico	Field Calculator	Selección de predios
Descr	String	Interpretación (muestra)	Field Calculator	Definición de clases
Area	Float	Área del predio	Field Calculator	Clasificación
Barea	Float	Área construida	Spatial Join	Clasificación
PBuild	Float	Porcentaje construido	$100 * [Barea] / [Area]$	Clasificación
NFloor	Short	Max número de pisos	Spatial Join	Clasificación
NBuild	Short	Número de edificios	Spatial Join	Clasificación
HSpace	Float	Espacio habitable	Spatial Join	Clasificación

BlockID – se realizó un Outer Spatial Join con el polígono de manzana del 2000 usando una capa de centroides de los predios como Target y la de manzana como Join. Se usó la regla WITHIN para relacionar los predios. Solo los predios que corresponden a cuchillas no tuvieron manzana asociada.

Barea, NFloor, NBuild, HSpace – Se agregaron los campos los atributos de edificios sobre todos los registro con mismo ParcID. Esto se hizo en Excel mediante la inserción de subtotales. Se usaron distintas funciones de agregado para los distintos atributos de acuerdo a la siguiente tabla:

Tabla B-21. Función de agregado empleada para el cálculo de atributos a partir de la capa de edificios.

Atributo	Función de agregado
Barea	SUMA
NFloor	MAX
HSpace	SUMA
NBuild	COUNT

Una vez generados los totales a nivel de predios, se anexaron a la tabla de predios correspondiente.

Apéndice C. Procesamiento de la nube de puntos LiDAR

La nube de puntos fue proporcionada por el INEGI en 130 archivos en formato LAS, cada uno de los cuales presentó una línea de vuelo. Lo primero que se hizo fue la generación de un shapefile de ArcGIS que sirviera como índice para conocer las características de cada archivo LAS incluyendo su cobertura, número de puntos por retorno y nombre del archivo. Para ello se escribieron dos funciones en MATLAB, una para leer (LASREAD) un archivo en formato LAS y otra para generar el contorno de la nube de puntos (LASCOVER). Con estas funciones se generaron una serie de polígonos que corresponden a la cobertura de cada línea de vuelo. Dichos polígonos se exportaron de MATLAB a un shapefile usando una función desarrollada previamente (SHPWRITE).

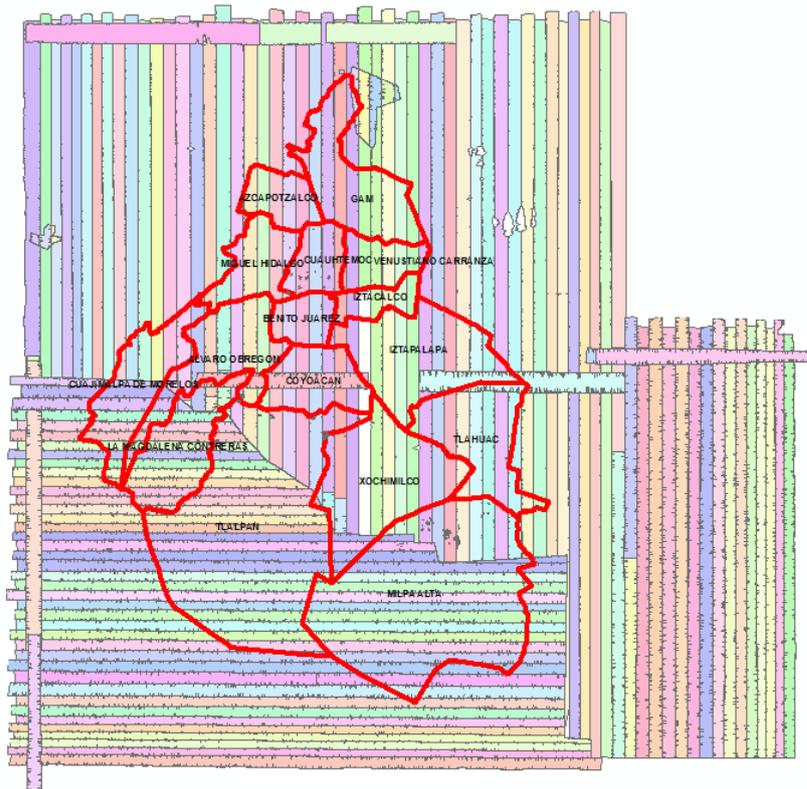


Ilustración C-24. Líneas de vuelo de los datos LiDAR. Cada franja representa un archivo LAS que contiene una nube de puntos en 3-d con atributos.

Retícula

Se consideró las limitaciones para el manejo de datos masivos como son las nubes de puntos de LiDAR en MATLAB, por lo que se optó por un esquema de partición regular. Para ello se creó una retícula de cuadrados de 1km por 1km, la cual cubre el DF en su totalidad. Se generó un archivo shapefile en MATLAB con dicha geometría. Se usó la convención de letra-numero para identificar las coordenadas (columna- renglón) de cada cuadrado dentro de la retícula, de tal forma que se pueden especificar rangos con en una hoja de cálculo. El rango completo de la retícula es A1:AU63. Usando una operación de Spatial Join en ArcGIS se asignó a cada cuadrado los nombres de los archivos LAS que cubrían dicho cuadro.

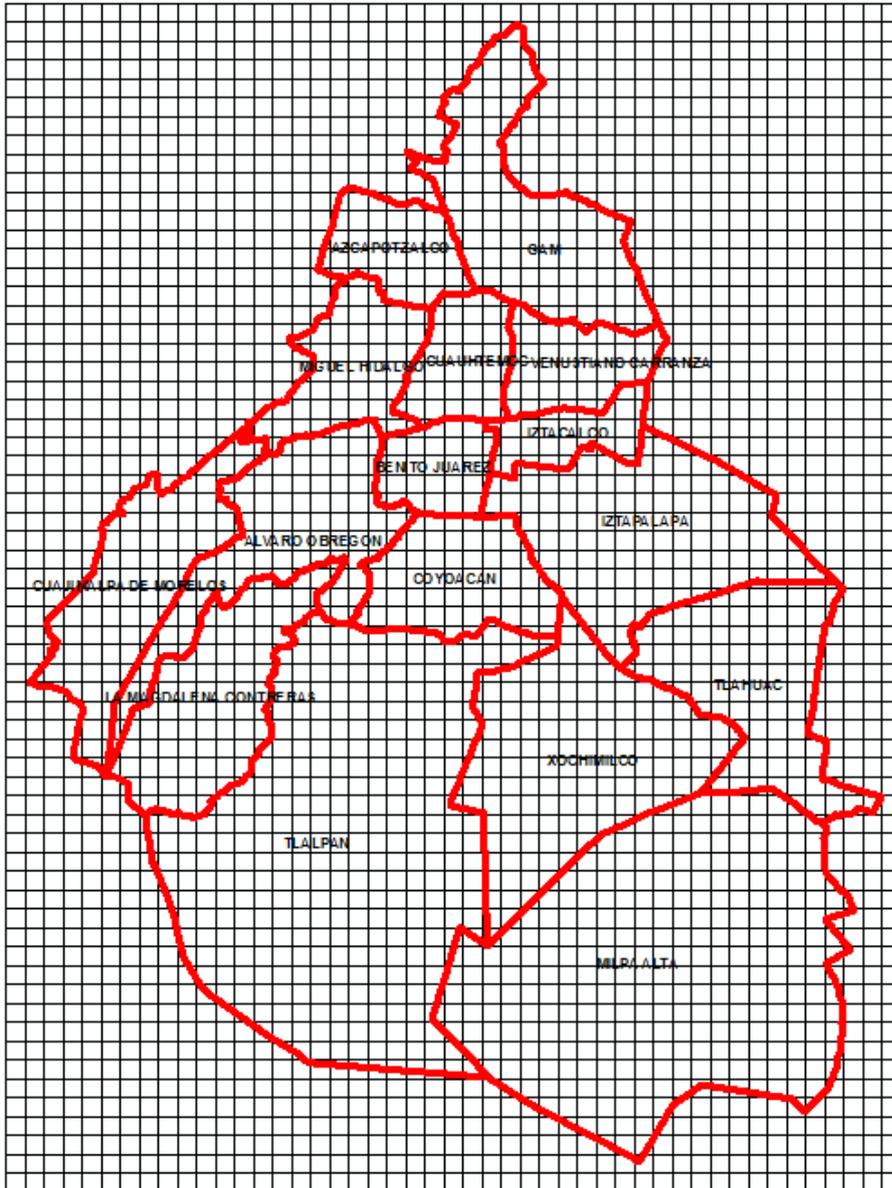


Ilustración C-25. Malla de partición de los datos LiDAR para su procesamiento.

Extracción puntos

En MATLAB se cargó la retícula y se leyeron uno a uno los archivos que cubrían los cuadrados en un rango especificado, de tal forma que, usando la función LASREAD con opciones de filtrado, se extrajeron las nubes de puntos de cada cuadrado. Estas nubes se guardaron en formato nativo de datos de MATLAB (MAT) y se guardaron en 16 bits con los correspondientes parámetros de escalamiento y corrimiento, tal como se almacenan en los archivos LAS para ahorro de memoria.

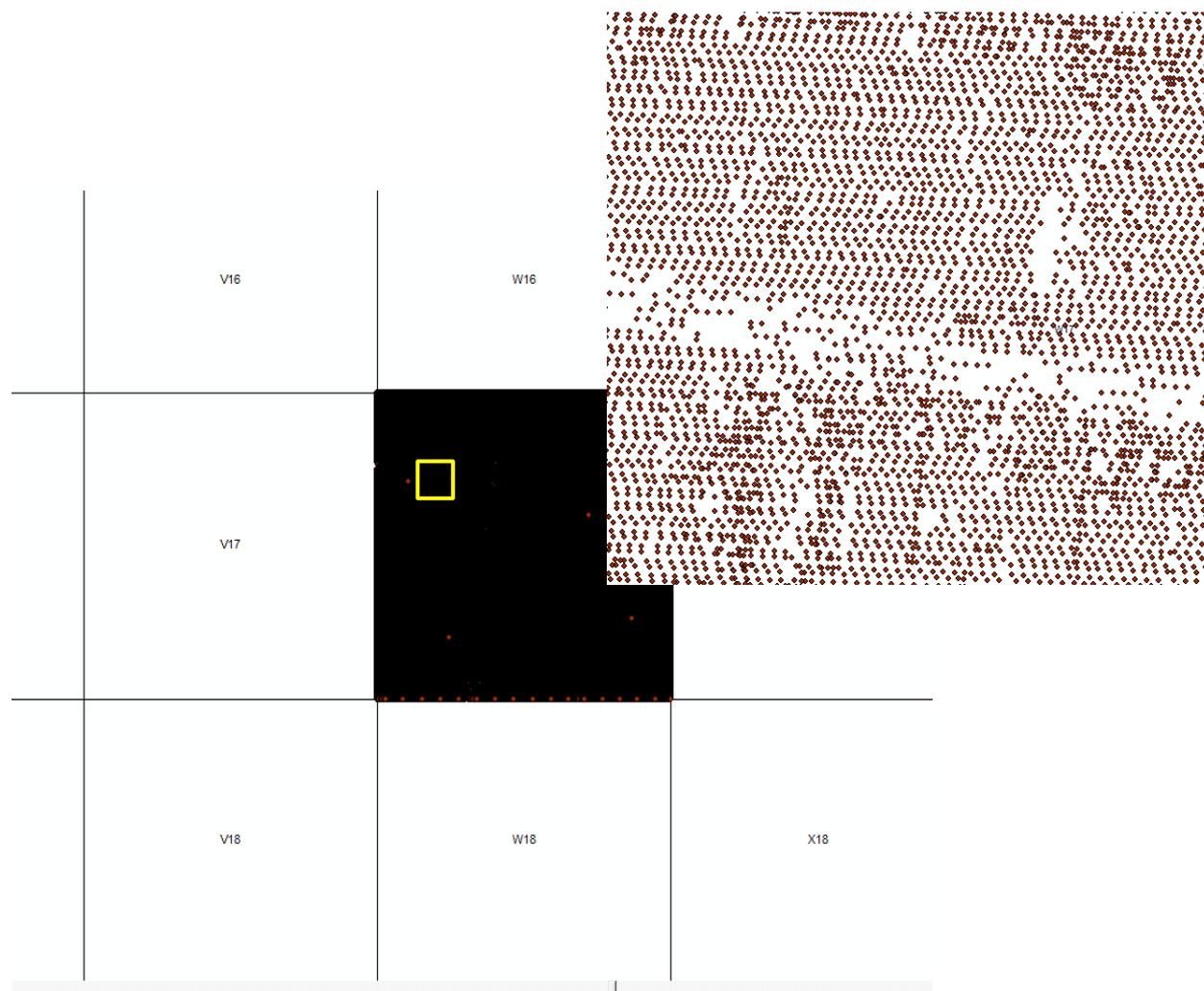


Ilustración C-26. Extracción de la nube de puntos para cada cuadro de la malla.

Rasterización

A partir de cada nube de puntos extraída se crearon dos modelos de superficie usando una retícula de un metro de tamaño de celda. En el primer caso, a cada celda se asignó el valor máximo de elevación de los puntos que cayeron dentro de la celda. Luego, se asignaron las celdas en las que no cayó ningún punto usando una interpolación natural (convolución cúbica) que estaba disponible en MATLAB (TRISCATTEREDINTERP). En el segundo caso, en se asignó el valor máximo de elevación de los puntos que cayeron dentro de la celda y los huecos se interpolaron como en el caso anterior.

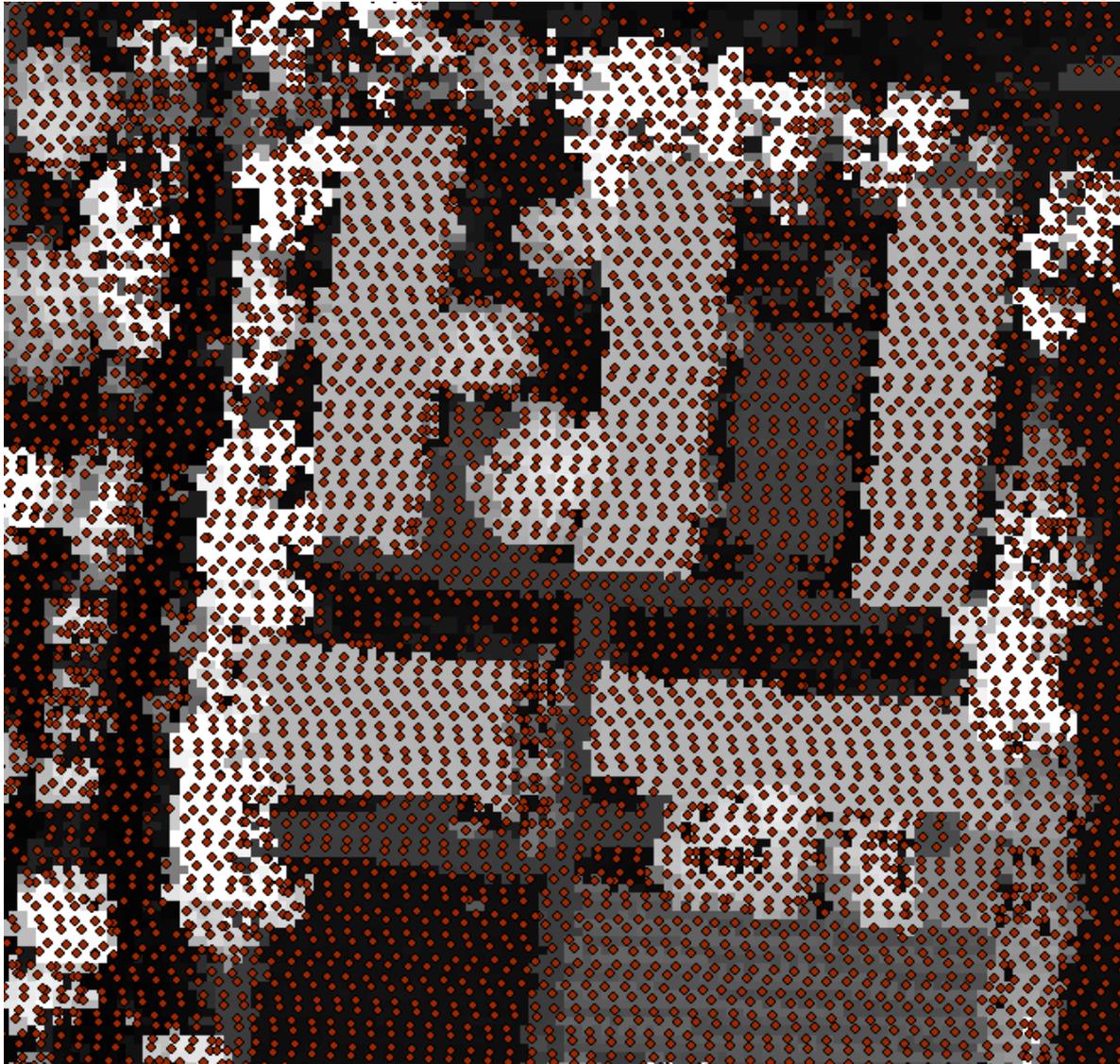


Ilustración C-27. Modelo digital de elevaciones generado mediante la interpolación de la nube de puntos LiDAR (sobrepuestos).

Filtrado de puntos de suelo

El método de filtrado que se desarrolló consiste en la aplicación de un operador de erosión que basado en la transformada de Hermite multiescala (Silván-Cárdenas & Wang, 2006). El método fue mejorado y probado en el marco del presente proyecto. Los detalles del método y sus modificaciones se pueden consultar en un par de publicaciones recientes (Silván-Cárdenas J. , 2013; Silván-Cárdenas & Wang, 2014).

El filtrado se operó en cada segmento de 1km² por separado, sin embargo, dado que el enfoque multiescala requiere información de elevación en un vecindario de cada pixel, fue necesario anexar un buffer de 100 metros alrededor del segmento. Los datos de este buffer se extrajeron de los cuadros vecinos al cuadro que se procesó. El algoritmo de filtrado modela la superficie a partir de parámetros de terreno (máxima pendiente y máxima diferencia de elevación entre mesetas contiguas) y de rasgos sobre el terreno (máximo ancho de rasgo, mínima altura de rasgo). Estos parámetros se recolectaron con la ayuda de visualizaciones tanto de perfiles extraídos de la superficie original como de la interpretación del histograma de frecuencias bidimensional elevación-pendiente, tal como se describe en (Silván-Cárdenas & Wang, Multiscale approach for ground filtering from Lidar altimetry measurements, 2014). De todo el procesamiento de los datos esta parte fue la más ardua, ya que requiere de cierta destreza y conocimiento profundo del método y el software empleado.

El filtrado entrega a su salida una superficie erosionada que es usada para discriminar los puntos de terreno de los puntos por encima del terreno. Los puntos de terreno así detectados pueden ser interpolados para generar la superficie de terreno. Alternativamente, se puede usar la máscara de terreno para reemplazar los píxeles que fueron detectados como no terreno en la superficie original. En este proyecto se empleó el segundo método para producir la superficie de terreno desnudo (Ilustración C-28, derecha). Adicionalmente, se produjo un raster de alturas sustrayendo la elevación del terreno de la superficie original (Ilustración C-28, izquierda).



Ilustración C-28. Ejemplo de filtrado de terreno. A la izquierda se muestra el modelo digital de terreno original y a la derecha el modelo de terreno desnudo generado.

Segmentación y filtrado de terreno

El método de filtrado basado en la transformada de Hermite multiescala es efectivo para detectar edificios grandes y chaparros, pero en zonas de topografía muy compleja, como en el caso de barrancas con asentamientos, el método falla ya que se debe elegir un parámetro de pendiente muy grande para detectar las barrancas, pero a la vez se evita el filtrado correcto de edificios grandes como en el ejemplo

de la Ilustración C-29, donde se indica en un círculo rojo los errores en la máscara de terreno/non-terreno.



Ilustración C-29. Ejemplo de filtrado fallido. A la derecha se muestra el modelo digital de elevación y a la izquierda la máscara de extracción de terreno (negro). Los círculos en rojo señalan los errores de filtrado.

Para resolver estos problemas que se presentaban, sobre todo en el caso de la delegación Tlalpan, se desarrolló un método alternativo que se basa en la segmentación de una medida de gradiente de la superficie². El gradiente crea bordes que delimitan áreas relativamente homogéneas en elevación. Dichas áreas son etiquetadas para identificar un segmento. Los segmentos son posteriormente clasificados en terreno y no-terreno en dos pasos. En un primer paso se detectan los segmentos que potencialmente corresponden al terreno. Esto se logra calculando un valor de probabilidad de terreno dadas ciertas características de los segmentos. Las características que se consideraron son las siguientes.

1. Área (A) – el área de los segmentos de terreno tiende a ser muy grande en comparación con los segmentos de edificios y otros objetos.
2. Complejidad (C) – se define como el área convexa del segmento entre el área total del segmento menos uno. Indica que tan complejo es el segmento o qué tanto difiere de un segmento convexo.
3. Elevación relativa (R) – Se calcula como la mediana de las diferencias de la elevación entre los píxeles del borde del segmento con el píxel vecino más bajo. Indica que tan elevado está el segmento relativo a los segmentos circundante.

Las tres características de los segmentos se combinaron en un índice que corresponde a la probabilidad de observar un segmento de terreno. La expresión de la probabilidad corresponde al producto de funciones exponenciales previamente calibradas.

Una vez detectados los segmentos de terreno, se procede a generar una superficie de terreno inicial mediante interpolación. El segundo paso consiste en un procedimiento iterativo que calcula las alturas provisionales de los segmentos restantes y detecta nuevos segmentos de terreno, actualizando la superficie de terreno previa. Este procedimiento, normalmente toma entre 10 a 20 iteraciones para converger. El siguiente ejemplo ilustra (izquierda a derecha y de arriba abajo) el MDS de entrada, la

² Artículo en preparación

segmentación del gradiente, la selección inicial de los segmentos de terreno y el modelo digital de alturas resultantes, eliminando la variabilidad del terreno.

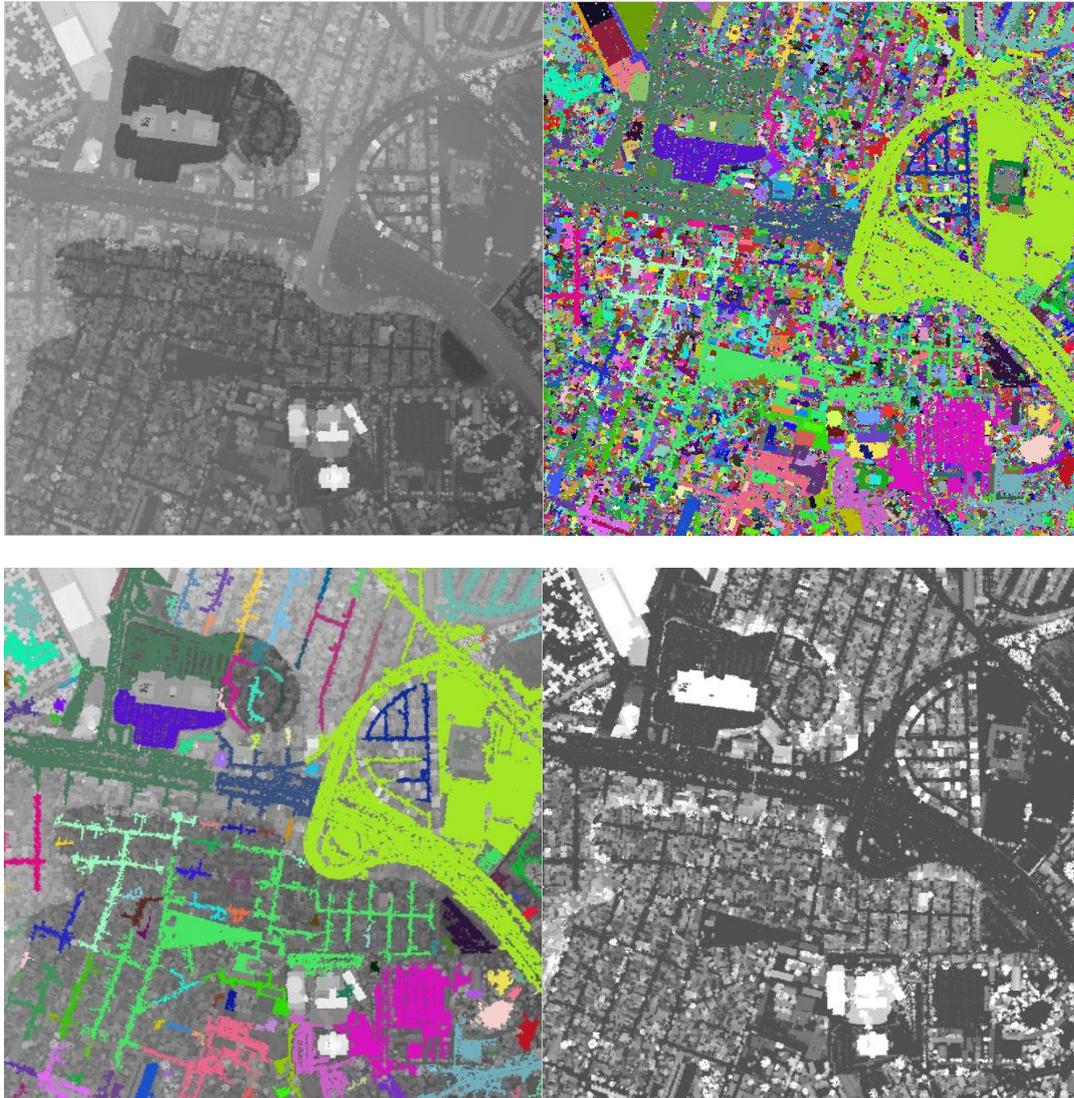


Ilustración C-30. Ejemplo de filtrado de terreno por segmentación.

Creación de mosaicos

El filtrado genera dos capas, el modelo digital de terreno y el modelo digital de alturas, para cada cuadro procesado. Estos cuadros son posteriormente pegados para formar un mosaico que cubre cada delegación (Ilustración C-31). El proceso de creación de mosaicos se realizó en ENVI.

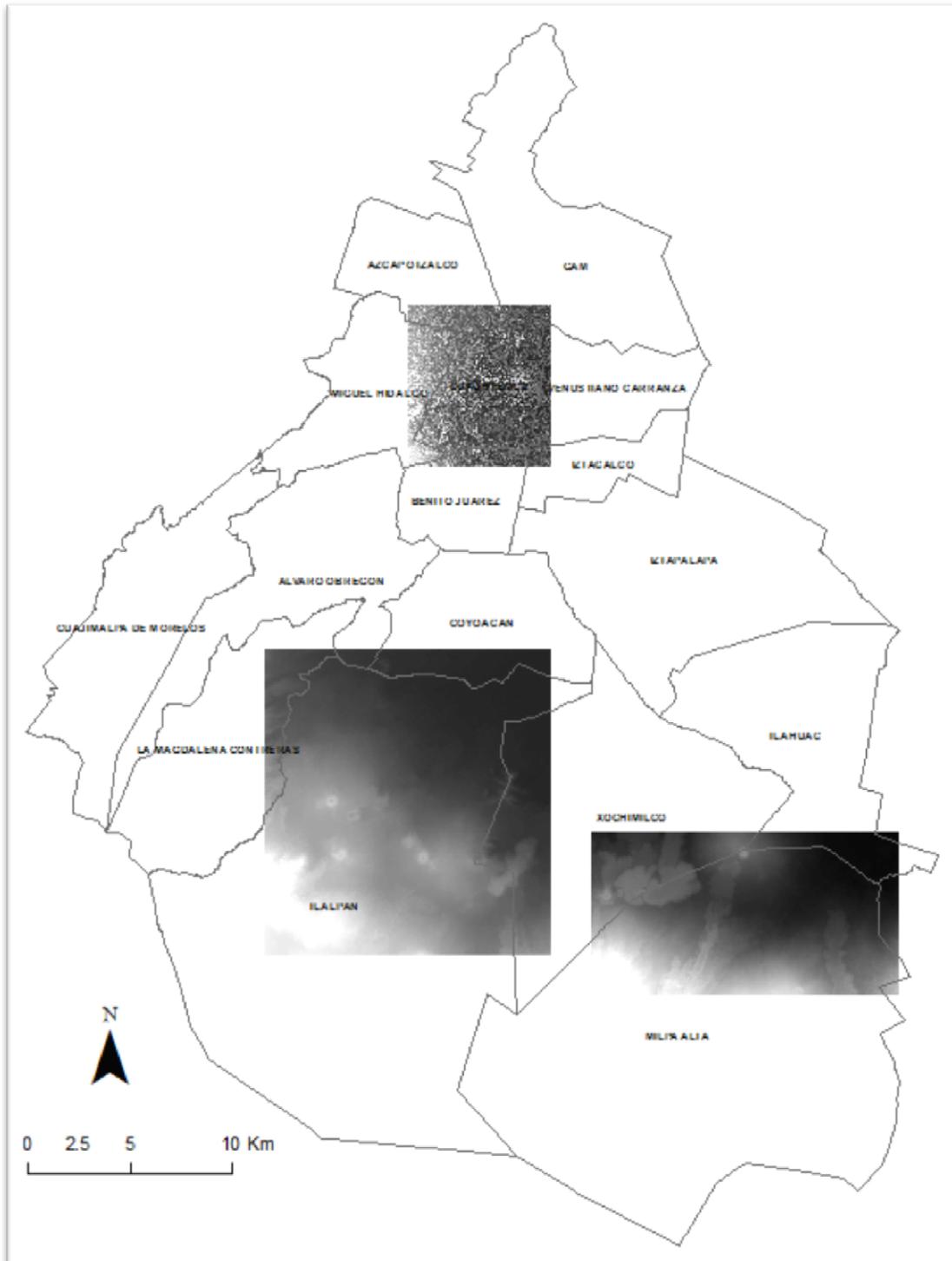


Ilustración C-31. Cobertura de los modelos digitales de terreno y superficie generados a un metro de resolución espacial

Clasificación de la nube de puntos

Los puntos clasificados como no terreno, fueron subsecuentemente clasificados como Suelo, Vegetación Alta, Vegetación Baja y Otro. Para esto se usó la clasificación de cobertura generada con la combinación de imágenes multispectrales y para ello se desarrolló una función PNTSAMP.M desarrollada en MATLAB. Esta función transfiere la información de clase de cobertura a la nube de

puntos originales (Ilustración C-32). Además del atributo de clase, se agrega un atributo de altura, cuyo valor corresponde al valor del modelo digital de alturas en el punto.

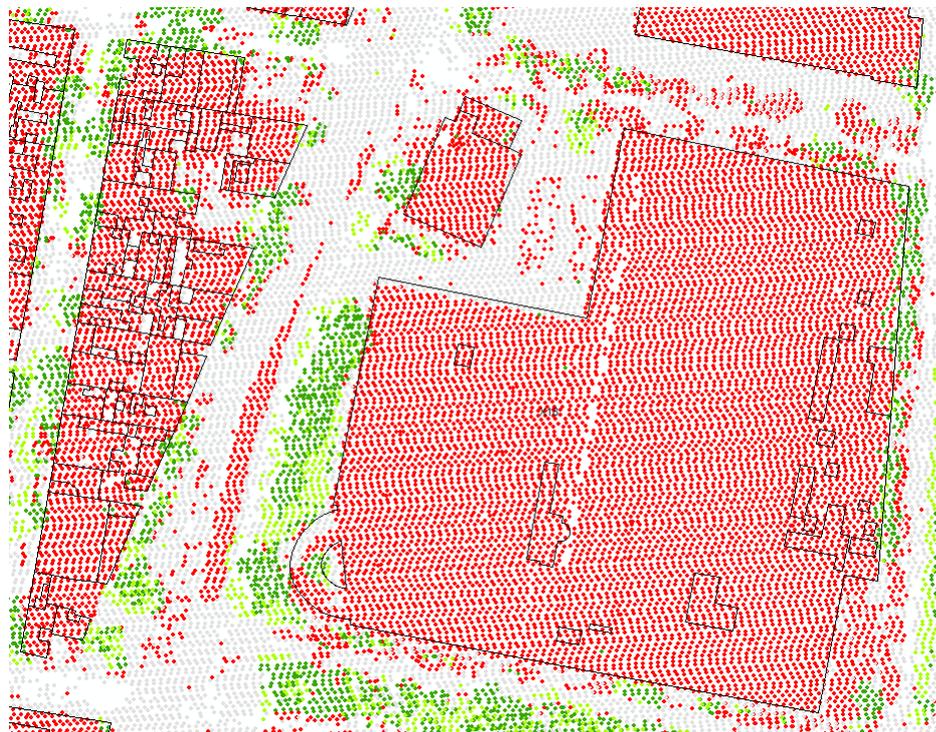


Ilustración C-32. Acercamiento a la nube de puntos coloreados por clase de cobertura.

Apéndice D. Procesamiento de las imágenes multi espectrales

Se contó con imágenes QuickBird y se compraron con recursos del proyecto imágenes WorldView2. Estas últimas fueron entregadas por el proveedor en tres carpetas, una por cada sitio de estudio (Cuauhtémoc, Tlalpan y Milpa Alta). En el caso de Tlalpan y Milpa Alta, la cobertura se compuso de dos adquisiciones distintas, una que cubre el este y otra el oeste. La adquisición de la parte del este cubre la mayor parte del sitio en ambos casos.

Se extrajeron los segmentos que cubrían el área de cada delegación de interés. Las fechas de adquisición y segmentos por delegación se indican en la siguiente tabla.

Tabla D-22. Imágenes empleadas en el proyecto.

Delegación	Sensor	Fechas de Adquisición	# Archivos
Cuauhtémoc	QB	17 FEB 2007	1
	WV2	12 JUN 2013	1
Tlalpan	QB	20 DIC 2007 y 24 ABR 2008	2 y 1
	WV2	21 ENE 2012	9
Milpa Alta	QB	18 MAY 2007	1
	WV2	21 ENE 2012	6

Preparación de los datos

Las adquisiciones estaban particionados en cuadros almacenados en archivos por separado, por lo que se cargó el archivo .TIL en ENVI, con lo cual se forma la imagen al vuelo, misma que se exporto en formato ENVI Standard para su posterior análisis.

Recorte

De la imagen formada en ENVI se extrajeron los recortes que cubrían exactamente la misma cobertura que los modelos digitales de elevación generados con LiDAR.

Calibración radiométrica

Para los recortes, tanto multispectrales como pancromáticos, se calculó la rianza-al-sensor empleando los coeficientes de calibración proporcionados por el proveedor (archivos .IMD). El resultado se almacenó en formato entero usando un factor de 1000 y redondeando el valor resultante.

Realce espacial

Las bandas multi espectrales calibradas se funcionaron con las bandas pancromáticas usando en método GramSchmidt Sharpening disponible en ENVI. En algunos casos fue necesaria la creación de máscaras para no afectar las radianzas de salida con pixeles de fondo con valores de zero. Las imágenes multiespectrales de salida se remuestrearon a 1 metro de resolución espacial coincidente con la resolución de los datos de elevación.

Orto-rectificación

Se orto rectificaron las imágenes de radianza con realce espacial usando el modelo digital de elevaciones (MDE) generado con LiDAR a 1 metro de resolución. Se tomó el MDE como base para definir la resolución, extensión y ubicación de los pixeles orto rectificados, por lo que sólo fue necesario el remuestreo de la imagen de entrada. Para ello se empleó un remuestreo por convolución cúbica. No se empleó el cálculo de la altura geoidal y en su lugar se emplearon puntos de control. Se recolectaron no menos de 5 puntos de control en ENVI usando Map->Ortorectification->QuickBird->Ortorectify QuickBird with Ground Control. Los valores de Easting, Northing y Elevation se extrajeron del modelo digital de elevaciones, mientras que las coordenadas de imagen (X,Y) se extrajeron de la imagen. Se exportó la lista de puntos con Lon, Lat en ASCII y se le asoció a la imagen fusionada el archivo RPB que contiene los coeficientes de la función racional polinomial proporcionada por el proveedor para la pancromática. Luego en MATLAB se desarrolló un u función ENVORTHO, la cual recibe la imagen y los puntos de control como entrada y realiza la orto rectificación. La razón de usar, MATLAB, en lugar de la implementación de ENVI se debe a la flexibilidad que se tiene para el relleno de los huecos que se generan al momento de orto rectificar.

En general, se observó que aunque la orto-rectificación efectivamente compensa los desplazamientos por la geometría de visión, el método también interpola los huecos que se crean en áreas que estaban previamente ocultas por los objetos desplazados. Sin embargo dicha interpolación puede crear un aspecto de duplicidad de los datos (Ilustración D-33). Actualmente se está investigando cómo construir una mejor interpolación para el llenado de dichos huecos.



Ilustración D-33. Acercamiento de la imagen multispectral sin ortorectificación (ARRIBA) y con ortorectificación (ABAJO).

Corrección atmosférica

No se aplicó ninguna corrección atmosférica debido a la calidad visual que presentaron las imágenes.

Creación de mosaicos

En el caso de Tlapan, donde se tenía más de una adquisición para cubrir el área de interés, fue necesario hacer un mosaico a partir de las imágenes orto-rectificadas. Los mosaicos para cada delegación de estudio se ilustran en la Ilustración D-34.

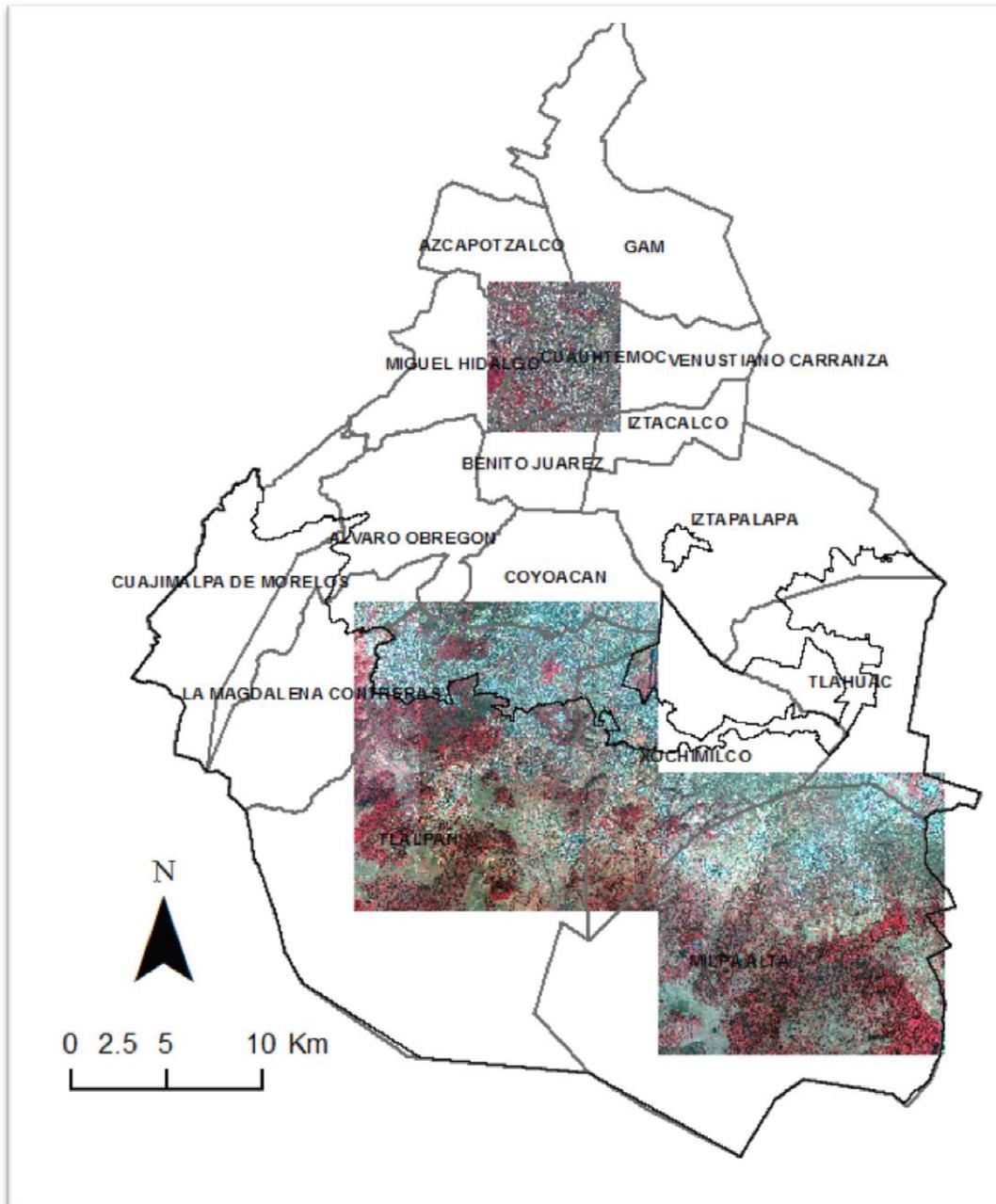


Ilustración D-34. Cobertura de las escenas wv2 adquiridas

Clasificación de cobertura

Se empleó la imagen QuickBird (2007) fusionada entre la pancromática y la multiespectral a 1m de resolución espacial. Esta fue particionada en segmentos de 1km² coincidentes con los datos de

superficie. Luego en MATLAB se desarrolló un programa que procesa segmento por segmento, para cada uno de los cuales se calcula el índice de Vegetación de la Diferencia Normalizada (NDVI). Posteriormente, éste se combina con la capa de Alturas para producir la clasificación de cobertura de acuerdo con las siguientes reglas:

Tabla D-23. Reglas de clasificación empleadas para el mapa de cobertura

#.	Clase	Altura	NDVI
1.	Suelo	< 0.5	<0.15
2.	Construido	>0.5	<0.15
3.	Vegetación Baja	<0.5	>0.15
4.	Vegetación Alta	> 0.5	>0.15

El siguiente ejemplo muestra la clasificación en colores, donde la clase Otros se muestra en Rojo.

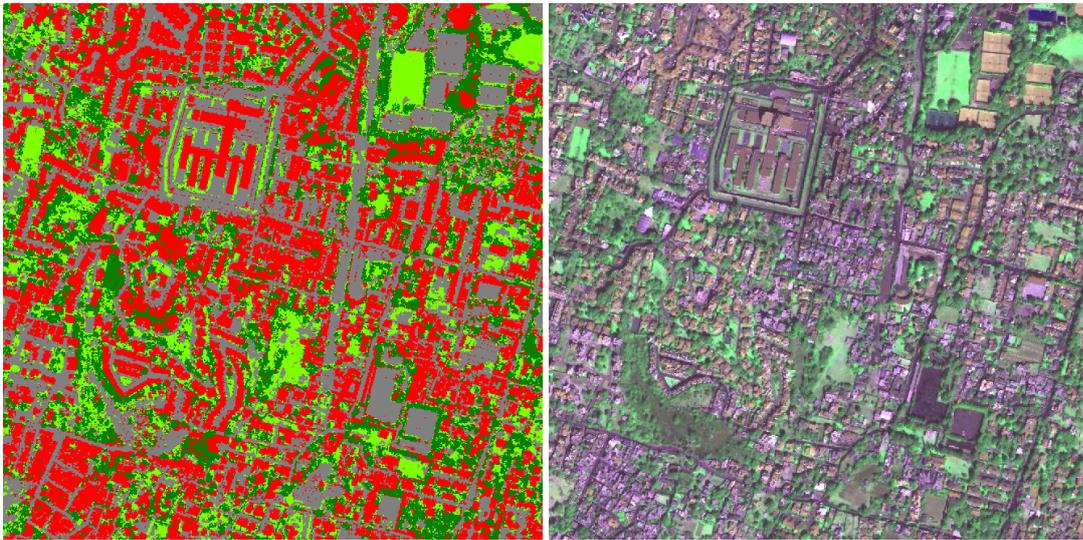


Ilustración D-35. Clasificación de cobertura a partir del NDVI y el modelo digital de alturas.