

UNIVERSIDAD AUTÓNOMA DE SINALOA



FACULTAD DE CIENCIAS DE LA TIERRA Y EL ESPACIO Y
FACULTAD DE INFORMÁTICA CULIACÁN

**PROPUESTA METODOLÓGICA DE ANÁLISIS DE
SENSIBILIDAD PARA EVALUAR LA CONFIABILIDAD DE
MODELOS GEOESPACIALES BASADOS EN TÉCNICAS DE
EVALUACIÓN MULTICRITERIO**

TESIS

COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE
DOCTOR EN CIENCIAS DE LA INFORMACIÓN

SUSTENTANTE

JOSÉ DE JESÚS URIARTE ADRIÁN

DIRECTORES

DR. WENSESLAO PLATA ROCHA

DR. RENÉ RODRÍGUEZ ZAMORA

CULIACÁN, SINALOA, FEBRERO DE 2020

AGRADECIMIENTOS

A mis padres, Candelaria Adrián Hernández (mi madre, mi confidente, mi pilar y mi motivación) que sé que está en el cielo observando este nuevo logro en mi vida y José de Jesús Uriarte Sanchez (mi padre, mi amigo, mi protector y mi héroe), a ellos les debo este triunfo profesional, ya que siempre me dieron su apoyo incondicional, gracias por todo su trabajo y dedicación para darme una formación académica y sobre todo humanista, gracias padres ya que sin su comprensión y paciencia no estaría escribiendo estas palabras. De ellos es este triunfo y para ellos es todo mi agradecimiento.

Para mis hermanos, Carmen, Víctor, Patricia, Lourdes y Marisa, que me han ayudado a seguir luchando y a no rendirme, me han demostrado que puedo contar con ellos para todo.

A todos mis sobrinos que por ellos aumentan mis deseos de superación para apoyarlos en su larga vida académica y sobre todo familiar.

A la Familia Félix González, en especial a mi esposa Thania por estar aquí y ahora, comprendiéndome, soportándome pero sobre todo apoyándome para lograr esta meta juntos, y a nuestra bebe Mariana, a las dos las AMO y lo que nos falta.

A mis compañeros de posgrado Rosendo, Gaby, Carlos, Lidia, Arturo y Sergio por aceptarme en sus vidas y regalarme todos esos momentos tan padres que se quedaran por siempre en mi corazón.

Al Dr. Wenseslao Plata Rocha por su paciencia, apoyo, guía y enseñanzas en el desarrollo de toda la investigación. A Roberto Bernal Guadiana por su apoyo y motivación para continuar preparándome en el perfil de investigador.

A mi familia Alva Pérez y a mis amigos de Puebla que siempre estuvieron al pendiente de mí.

A la Universidad Autónoma de Sinaloa, la Máxima Casa de Estudios, a la Facultad de Ciencias de la Tierra y el Espacio, a la Facultad de Informática Culiacán, pero sobre todo al Consejo Nacional de Ciencia y Tecnología (CONACyT), quien me apoyo con una beca completa la cual me permitió lograr este trabajo de tesis y adquirir experiencia profesional.

A todos les agradezco todo el apoyo y tiempo brindado a ustedes les dedico este logro.

A todos. . . GRACIAS

Febrero de 2020

CONTENIDO

1.- INTRODUCCIÓN	1
1.1 Hipótesis.....	3
1.2 Objetivos.....	3
1.3 Planteamiento y justificación	3
2.- ANTECEDENTES	7
2.1 Modelos de simulación.....	7
2.2 Validación de los modelos de simulación.....	23
2.3 Análisis de sensibilidad en modelos espaciales	34
2.4 <i>Ventajas y desventajas del AS en modelos geoespaciales</i>	40
2.5 Análisis de sensibilidad explícitamente espacial.....	43
2.6 Ventajas y desventajas de las técnicas de análisis de sensibilidad	46
2.7 Análisis Computacional e Implementaciones	48
3.- MARCO TEÓRICO.....	56
3.1 Análisis de sensibilidad	56
3.1.1 Técnicas de análisis de sensibilidad.....	57
3.1.2 Análisis de sensibilidad en modelos geoespaciales	69
3.2 Desarrollo computacional.....	74
3.2.1 Lenguaje de programación C#.....	74
3.2.2 Idrisi selva	75
3.2.3 Integración de sistemas	75
5.- METODOLOGÍA.....	77
5.1 Área de estudio empleada como base para probar propuesta metodológica	78
5.2 Definición de objetivos	80
5.3 Modelación de variables geoespaciales	81
5.4 Modificación de variables y factores geoespaciales	86
5.5 Indicadores de análisis de sensibilidad	89
6.- RESULTADOS	92
6.1 Descripción del módulo diseñado para simular el Modelo de EMC y AS.	92
6.2 <i>Análisis de RANK y RANDOM</i>	94
6.3 Resultados del modelo de EMC	97
6.4 <i>Comparación de variaciones con escenario base</i>	105
6.5 Resultados del Análisis de Herramientas de Análisis de Sensibilidad.	115
6.6 Resultados del Análisis de Sensibilidad	119

6.7 Comparación entre índices	122
6.8 Comparación de la simulación al 2030 con crecimiento actual	123
7.- DISCUSIÓN	130
8.- CONCLUSIONES	132
BIBLIOGRAFÍA	135

LISTA DE FIGURAS

<u>Figura 1.- Diagrama de Integración entre un SIG y EMC.....</u>	22
<u>Figura 2.- a) Una sección de la ciudad de Riyadh, b) vecindad urbana, c) sub-vecindad urbana y d) un bloque urbano (Al-Ahmadi et al., 2013).....</u>	27
<u>Figura 3.- Artículos según la técnica de simulación y su objeto de estudio.</u>	30
<u>Figura 4.- Porcentaje de estudios según el objeto de estudio.....</u>	31
<u>Figura 5.- Numero de estudios por año.</u>	32
<u>Figura 6.- Estudios que contemplan análisis de sensibilidad.....</u>	33
<u>Figura 7.- Técnicas para realizar un análisis de sensibilidad espacial.....</u>	39
<u>Figura 8.- Clasificación de métodos de AS global (looss & Lemaitre, 2015). ...</u>	47
<u>Figura 9.- Metodología de un análisis de sensibilidad estadístico.....</u>	69
<u>Figura 10.- Modificaciones realizadas en un AS para simulaciones geoespaciales basadas en Autómatas Celulares.....</u>	71
<u>Figura 11.- Modificaciones realizadas en un AS para simulaciones geoespaciales basadas en los MBA.....</u>	73
<u>Figura 12.- Modificaciones realizadas en un AS para simulaciones geoespaciales basadas en EMC.....</u>	74
<u>Figura 13.- Propuesta Metodológica para realizar AS explícitamente espacial.....</u>	77
<u>Figura 14.- Mapa de municipios de Culiacán y Navolato, Sinaloa, México.....</u>	78
<u>Figura 15.- Mapas de usos de suelo del modelo.</u>	81
<u>Figura 16.- Funciones lineales y sigmoidales monotónicamente crecientes.....</u>	84
<u>Figura 17.- Factores modelados que componen el modelo escenario.....</u>	85
<u>Figura 18.- Proceso para obtener la variación espacial en los factores (izqda).....</u>	87
<u>Figura 19.- Mapa original con rango máximo de 111.90 (izq).....</u>	88
<u>Figura 20.- Metodología de un proceso de simulación de escenarios.....</u>	88
<u>Figura 21.- Distintas vecindades de búsqueda de tipo von Neumann.....</u>	91
<u>Figura 22.- Tamaños de ventanas usadas en análisis de entorno simple.</u>	91
<u>Figura 23.- Interfaz del sistema desarrollado para simular escenarios.</u>	93
<u>Figura 24.- Interfaz del sistema desarrollado para selección aleatoria.....</u>	93
<u>Figura 25.- Clasificación y ordenamiento realizado por el comando “RANK” ...</u>	94
<u>Figura 26.- Metodología para selección aleatoria.....</u>	95
<u>Figura 27.- A la izquierda pixeles con el mismo nivel de aptitud.....</u>	95
<u>Figura 28.- Comparación de una selección por medio de RANK-RECLASS.....</u>	96
<u>Figura 29.- Diferentes selecciones aleatorias.....</u>	96
<u>Figura 30.- Mapas de aptitud con la técnica OAT, factor 1 modificado.....</u>	97
<u>Figura 31.- Mapa de pixeles más veces seleccionados (frecuencia).....</u>	98
<u>Figura 32.- Clasificación de pixeles OAT en cuartiles y quintiles.....</u>	100
<u>Figura 33.- Diferencia entre las selección de 676 pixeles considerando dos .</u>	102
<u>Figura 34.- Mapa de pixeles más veces seleccionados (frecuencia) utilizando solo los resultados de 110 mapas del proceso combinatorio.....</u>	103
<u>Figura 35.- Mapa de pixeles más veces seleccionados (frecuencia) utilizando resultados del método combinatorio (4095 mapas).....</u>	103
<u>Figura 36.- Visualización del aumento en la resolución del mapa base.....</u>	105

<u>Figura 37.- Comparación de escenario base y la selección realizada a partir de la modificación del factor a) proximidad a asentamientos humanos y b) proximidad inversa a USyV.</u>	106
<u>Figura 38.- Comparación de escenario base y la selección realizada a partir de la modificación del factor a) área urbanizable programada y b) proximidad a carreteras.</u>	107
<u>Figura 39.- Comparación de escenario base y la selección realizada a partir de la modificación del factor a) pendiente y b) proximidad a ríos</u>	108
<u>Figura 40.- Comparación de escenario base y la selección realizada a partir de las modificaciones al peso de cada factor.</u>	109
<u>Figura 41.- Escenario modificando el factor 1 y grafica de análisis</u>	110
<u>Figura 42.- Escenario modificando el factor 2 y grafica de análisis</u>	110
<u>Figura 43.- Escenario modificando el factor 3 y grafica de análisis</u>	111
<u>Figura 44.- Escenario modificando el factor 4 y grafica de análisis</u>	111
<u>Figura 45.- Escenario modificando el factor 5 y grafica de análisis</u>	111
<u>Figura 46.- Escenario modificando el factor 6 y grafica de análisis</u>	112
<u>Figura 47.- Graficas de análisis de entorno de las modificaciones de los pesos de factores.</u>	112
<u>Figura 48.- Graficas de coincidencia modificando todos los factores y pesos a la vez.</u>	113
<u>Figura 49.- Análisis de entorno de la selección realizada a partir de la modificación del factor a) proximidad a asentamientos humanos, b) proximidad inversa a USyV y c) área urbanizable programada.</u>	113
<u>Figura 50.- Análisis de entorno de la selección realizada a partir de la modificación del factor a) proximidad a carreteras, b) pendiente y c) proximidad a ríos</u>	114
<u>Figura 51.- Análisis de entorno global.</u>	114
<u>Figura 52.- Gráfica que muestra el rendimiento en segundos de modelos con n entradas y 1 salida</u>	116
<u>Figura 53.- Gráfica que muestra el rendimiento en segundos de modelos con 10 entradas y 1 y 5 salida, utilizando PDFs diferentes para cada entrada y realizando 10000 iteraciones.</u>	116
<u>Figura 54.- Gráfica que muestra el rendimiento en segundos de modelos con 2 entradas y 1 salida, utilizando PDFs diferentes y para 3 tipos de AS contemplados.</u>	117
<u>Figura 55.- Gráfica que muestra el rendimiento en segundos de modelos con 2 y 10 entradas y 1 y 5 salidas, utilizando PDF normal, para los tipos restantes de AS.</u>	118
<u>Figura 56.- Índice de sensibilidad adaptado a múltiples escenarios.</u>	119
<u>Figura 57.- Mapa de índice de sensibilidad de los resultados de OAT.</u>	120
<u>Figura 58.- Mapa de índice de sensibilidad de los resultados de las combinaciones totales.</u>	120
<u>Figura 59.- Muestra del valor de índice de sensibilidad en una región de 3x3</u>	121
<u>Figura 60.- MPMVA (izquierda) y mapa con categorías de alta, media y baja</u>	121
<u>Figura 61.- Comparación de diferentes índices para evaluar de sensibilidad</u>	123

<u>Figura 62.- Zonas a analizar el crecimiento simulado.</u>	124
<u>Figura 63.- Comparación del crecimiento simulado en el sector norte y el crecimiento actual.</u>	125
<u>Figura 64.- Comparación del crecimiento simulado en el sector Humaya y el crecimiento actual.</u>	126
<u>Figura 65.- Comparación del crecimiento simulado en el sector Valle Alto y el crecimiento actual.</u>	127
<u>Figura 66.- Comparación del crecimiento simulado en el sector Aeropuerto y el crecimiento actual.</u>	128
<u>Figura 67.- Comparación del crecimiento simulado en el sector Sur y el crecimiento actual.</u>	129

LISTA DE TABLAS

<u>Tabla 1.- Tabla comparativa de trabajos agrupados por objeto de estudio</u>	7
<u>Tabla 2.- Tabla comparativa de software de análisis de sensibilidad</u>	55
<u>Tabla 3.- Clasificación de los usos de suelo</u>	80
<u>Tabla 4.- Factores y pesos contemplados para el modelo a investigar</u>	95
<u>Tabla 5.- Muestra la diferencia en pixeles de las diferentes selecciones</u>	99
<u>Tabla 6.- Muestra la cantidad de pixeles que contiene cada rango de cuartil</u>	100
<u>Tabla 7.- Muestra el total de coincidencias en cada rango de la clasificación</u>	101
<u>Tabla 8.- Algunos cantidades y porcentajes de pixeles correspondientes a las frecuencias en los mapas (12, 101 y 4085 mapas)</u>	104
<u>Tabla 9.- Aumento en la resolución de la imagen y el tiempo de computo requerido para su procesamiento</u>	105
<u>Tabla 10.- Concordancia entre factores y pesos de factores modificados y resultados escenario base</u>	109
<u>Tabla 11.- Análisis de Entorno entre factores y pesos de factores modificados y resultados escenario base</u>	115

LISTA DE ACRÓNIMOS

AC – Autómata Celular

AHP - Analytic Hierarchy Process (Proceso de Jerarquía Analítica)

API - Análisis del Punto Ideal

AS - Análisis Sensibilidad

ASG - Análisis de Sensibilidad Global

CCP - Coeficientes de Correlación Parcial

CRR - Coeficiente de Regresión por Rango

CRE - Coeficiente de Regresión Estandarizado

EE - Efectos Elementales

EFAST - Extendend Fourier Amplitude Sensitivity Testing

EMC – Evaluación Multi-Criterio

EMD - Earth Mover's Distance (Distancia de movimiento de tierra)

FAST - Fourier Amplitude Sensitivity Testing

IS - Índice de Sensibilidad

LHS - Latin hypercube sampling (Muestreo Hipercubo Latino)

MBA – Modelo Basado en Agentes

MC - Monte Carlo

MOLA - Multi-Objective Land Allocation (Asignación de tierras multiobjetivo)

MPMVS - Mapa de Pixeles Mex Veces Seleccionados

OWA - Ordered Weighted Average (Sumatoria Lineal Ponderada Ordenada)

OAT - One At a Time (Un factor a la vez)

PDF - Probability Density Function

RDR - Radio de Desviación Relativa

RL – Regresión Logística

RN – Redes Neuronales

RNA - Redes Neuronales Artificiales

ROC - Receiver Operating Characteristic (Característica Operativa del Receptor)

SIG – Sistema de Información Geográfica

SLP - Sumatoria Lineal Ponderada

USyV - Usos de Suelo y Vegetación

1.- INTRODUCCIÓN

En la actualidad, el desarrollo de modelos espaciales ha tenido un crecimiento notable en la simulación geoespacial, debido a la necesidad de analizar diferentes aspectos en temas de uso de suelo y con la finalidad de establecer modelos que ayuden a mitigar las distintas problemáticas analizadas en temas como la expansión desordenada de las manchas urbanas, logrando cohesionar de manera económica y social a la población y a su vez, reducir el consumo de energía, recursos naturales y facilitar las instalaciones de nuevos servicios básicos como; el agua potable, drenaje, alcantarillado e instalaciones eléctricas (Comisión Europea, 1999; Yeh & Li, 2001; Rueda, 2002; Berjak & Hearne, 2002; Phua & Minowa, 2005; Naghdizadegan et al., 2013; Mogaji et al., 2014; Furtado & Eberhardt, 2015).

En este sentido, las técnicas de Evaluación Multi-Criterio (EMC), en procesos de planificación del territorio, han demostrado eficiencia en el diseño de modelos “óptimos” territoriales, ya que consideran la adecuación que presenta el territorio para el desarrollo de una determinada actividad, como la expansión de las áreas urbanas. Combinando criterios ambientales, sociales, económicos, entre otros, lo que posibilita el desarrollo de modelos con ciertos principios de sostenibilidad (Plata et al., 2013).

Asimismo, el desarrollo de herramientas informáticas se ha convertido en una creciente necesidad para mejorar el apoyo a los procesos de planificación del territorio, por lo cual es imprescindible el uso de herramientas y metodologías que sean capaces de aplicarse rápidamente a la necesidad de analizar cambios constantes en el territorio. Por esta razón, los Sistemas de Información Geográfica (SIG) demuestran su gran capacidad y eficiencia como ámbito en la planificación territorial (Barredo et al., 2003; Hu & Lo, 2007; Ligmann-Zielinska et al., 2008; Barredo & Bosque, 2009; Plata et al., 2010; Orán et al., 2011; Arsanjani et al., 2013; Plata et al., 2013; Gómez et al., 2014; Tan et al., 2015; Feng et al., 2016).

Sin embargo, estos modelos geoespaciales están siendo usados deliberadamente aplicando técnicas que aportan confiabilidad, pero estos procedimientos solo se

realizan de manera estadística proporcionando generalmente un valor (el mismo para todos los píxeles del mapa) de robustez.

Lo anterior es debido a que para los modelos geospaciales tienen complicaciones prácticas como la dificultad de representar una entrada espacial, así como la problemática de tratar con una cantidad de información muy alta, ya que cada píxel es un caso de estudio y esto hace que se requiera de alto poder computacional y es una de las cosas del porque no se han implementado estas simulaciones a nivel geoespacial y mucho menos análisis que den confianza a los resultados.

Aunado a esto, de las diferentes herramientas analizadas en este trabajo (presentes en el estado del arte) en ninguna de ellas es posible utilizar como entrada un mapa en su forma original para realizar un análisis de sensibilidad, todos convierten la información geoespacial en distribuciones para poder realizar alguno de las técnicas estadísticas.

Es por ello que, con un análisis realizado a una muestra de trabajos de investigación, que realizan modelos geospaciales, se confirma lo dicho por Gómez y Bosque (2004), sobre que el análisis de sensibilidad es generalmente ignorado por los investigadores.

Este trabajo considera factores relacionados con aspectos ambientales, económicos y sociales mediante técnicas de la EMC y SIG con la finalidad de realizar procesos de modelado y toma de decisiones a nivel espacial. Por lo tanto uno de los objetivos primordiales es realizar diferentes simulaciones que permitan identificar los espacios potencialmente más aptos para acoger algún uso de suelo. Así mismo, otra de las metas fue realizar un análisis de sensibilidad a las diferentes simulaciones calculando índices que nos ayudan a observar y analizar el comportamiento a nivel píxel y así otorgar algún grado de confianza.

1.1 Hipótesis

La aplicación de una metodología de análisis de sensibilidad a modelos geoespaciales, basados en evaluación multicriterio, genera fiabilidad en la selección de áreas deseables para acoger un nuevo uso de suelo.

1.2 Objetivos

El objetivo general de este trabajo se basa en el desarrollo de una propuesta metodológica de análisis de sensibilidad explícitamente espacial para evaluar los resultados de modelos basados en técnicas de evaluación multicriterio.

Los objetivos específicos con los que se logrará alcanzar el propósito de este estudio son:

- 1) Analizar los diferentes modelos simulación de escenarios geoespaciales.
- 2) Analizar los diferentes modelos matemáticos y geoespaciales para el análisis de sensibilidad.
- 3) Desarrollar un índice de sensibilidad explícitamente espacial.
- 4) Desarrollar una metodología para el modelo de simulación de escenarios geoespaciales que integre el índice de sensibilidad propuesto.
- 5) Desarrollar un sistema para su posterior incorporación en un SIG.

1.3 Planteamiento y justificación

En la actualidad la simulación de modelos geoespaciales ha tenido un gran auge en cuanto a su desarrollo y aplicación, debido, principalmente, a la necesidad de diseñar y plantear modelos innovadores, dentro de un marco metodológico más participativo, proveyendo información de calidad a los entes encargados de atender y dar solución a problemas relacionados con el impacto ocasionado a los diferentes ecosistemas, causados por procesos globales y locales de crecimiento urbano, deforestación, desertificación, entre otros (Barredo et al., 2005; Schwarz, et al., 2010).

Para lograrlo, es necesario desarrollar nuevos métodos y técnicas efectivas que generen propuestas de soluciones a los problemas territoriales, que cada vez son más acelerados por el crecimiento de la mancha urbana. En este sentido, la

geosimulación de modelos y escenarios de futuro son herramientas robustas para la representación de una realidad abstracta propuesta (Plata et al., 2013).

Dichos modelos se han implementado mediante técnicas basadas en autómatas celulares, regresión logística espacial, redes neuronales, modelos basados en agentes, evaluación multicriterio, entre otras; con la finalidad de intentar predecir el comportamiento de sistemas, diferenciándose respecto a su complejidad y en las reglas de decisiones empleadas para la simulación cuantitativa y/o geoespacial de los objetos de estudio (Triantakonstantis y Mountrakis, 2012).

Además, debido a que los modelos computacionales se pueden calibrar por medio de observación real, para mejorar los resultados, son vistos como dispositivos de análisis, por medio de los cuales se pueden inferir las propiedades (descriptores) más importantes de un área de estudio en particular (Poelmans y Rompaey, 2010).

En este sentido, la validación define la capacidad que tiene un modelo para predecir, mientras que la confiabilidad analiza el comportamiento de los resultados ante la variación de criterios iniciales (Qureshi, 1999). Es por ello, que es necesario enfocar los esfuerzos en tratar de determinar el grado de fiabilidad o confiabilidad de los resultados obtenidos de los modelos (Barredo y Gómez, 2008).

De este modo, el proceso de validación generalmente se lleva a cabo comparando los resultados obtenidos contra los datos reales (Qureshi et al., 1999). Sin embargo, la principal desventaja en la simulación de escenarios geoespaciales de futuro es que no es posible conocer los datos reales para poder comparar y verificar los resultados obtenidos (Plata et al., 2012).

Ante esto, analizar la fiabilidad de los resultados espaciales no es una práctica común al momento de trabajar con datos geográficos o geoespaciales, lo que hace que sea un área de oportunidades para dar robustez integral a proyectos de aspecto territorial y ambiental desarrollados en el campo de las geociencias

que contemplan modelos de simulación geoespacial de futuro (Gómez y Bosque 2004).

Hay que mencionar, además, que la integración de modelos para la construcción de escenarios de futuro mediante autómatas celulares, modelos basados en agentes, regresión lineal y evaluación multicriterio, y los Sistemas de Información Geográfica (SIG) ha avanzado de manera considerable en los últimos años, sin embargo, generalmente a los resultados de estas combinaciones no se les realiza algún tipo de análisis de confiabilidad que proporcione algún indicador de robustez (Feizizadeh et al., 2013).

Una alternativa para valorar el grado de confiabilidad en modelos de simulación geoespaciales de futuro, es aplicar algún tipo de análisis de sensibilidad (Pannell 1997; Malczewski, 1999; Qureshi et al., 1999; Barredo y Gómez, 2008).

Existen algunos trabajos que aplican AS a los modelos espaciales, los cuales son realizados en diferentes campos de estudio y para diferentes problemas (Lowry Miller y Hepner, 1995; Gómez-Delgado y Tarantola, 2006; Lu y Mohanty, 2001; Gómez Delgado, et al., 2014; Gómez Delgado y Bosque Sendra, 2004; Plata et al., 2013).

A pesar de que algunos métodos de análisis de sensibilidad (Sobol', FAST y EFAST) son procedimientos (estadísticos) con buenos resultados en el análisis global, estos se alimentan con datos escalares obtenidos a partir de técnicas de muestreo, y son representadas a través de sus Funciones de Densidad de Probabilidad (PDF por sus siglas en inglés), lo cual, no refleja el comportamiento espacial de los criterios de entrada, y además su comprensión y aplicación requieren de un nivel estadístico y matemático complejo, lo cual hace que su integración en los modelos de simulación geoespacial sea ignorado.

Esto es debido, a que los modelo geoespaciales tienen una cantidad de información es muy alta, ya que cada pixel es un caso de estudio y esto hace que se requiera de alto poder computacional y es por ello que no se han

implementado modelos de simulación a nivel geoespacial y mucho menos técnicas de análisis que den confianza a los resultados.

Es por lo anterior, que se necesita una propuesta metodológica que permita aumentar la credibilidad de los resultados e identificará las áreas o zonas más estables a los cambios en los valores de los parámetros de entrada del modelo. Además, este procedimiento podrá ser aplicado directamente en SIG o en modelos espaciales sin tener que manipular o adaptar la información a otro software de aplicación estadístico.

2.- ANTECEDENTES

2.1 Modelos de simulación

A nivel mundial se ha notado el aumento en el desarrollo de nuevos métodos y técnicas efectivas para generar propuestas para dar solución a problemas territoriales debido a los grandes cambios en el uso y ocupación de suelo, ocurridos en los últimos y que ha ocasionado un claro efecto en las alteraciones ambientales y la calidad de vida (Corrales, 2016).

Estas propuestas para mitigar los daños, coadyuvan al desarrollo y puesta en marcha de políticas públicas territoriales que dirijan el crecimiento hacia un futuro urbano equilibrado y amigable con el medio rural y sostenible. Una alternativa para lograrlo, es la generación de modelos espaciales la cual contempla la modelación de diferentes escenarios y la simulación de imágenes alternativas del futuro (Plata et al., 2013).

Los modelos de simulación deben ser entendidos como herramientas que permiten explorar diferentes posibilidades o escenarios de evolución del crecimiento urbano, por ejemplo, de acuerdo con las tendencias actuales del pensamiento prospectivo (Berdolulay, 2009; Aguilera et al., 2010).

Tabla 1.- Tabla comparativa de trabajos agrupados por objeto de estudio, técnica de simulación autor y año. Se puede observar las distintas técnicas de validación y/o calibración usada, además del método de análisis de sensibilidad usado en el caso que hay realizado alguno. (Elaboración propia).

Objeto de Estudio	Técnicas de Simulación	Autor y Año	Validación/Calibración	Análisis de Sensibilidad
Estudios Ambientales	AC	Straatman et al., 2001	Kappa y método de Power & Hagen	-
		Berjak & Hearne, 2002	Comparación con datos reales	-
		Singh, 2003	Diferentes vecindades, comparación visual con datos reales.	-
		Chen & Mynett, 2004	Análisis fractal y bondad de Ajuste	-
		Memarian et al., 2012	Kappa	-
		Omran et al., 2012	Validación Cruzada, ROC y modelo Logit	-
		Naghdizadegan et al., 2013	Matriz de Confusión	-

		Devendran & Lakshmanan, 2014	Análisis de error de primer orden	Análisis de Sensibilidad Regional
		Brändle et al., 2015	Diferentes niveles de agregación y comparación con datos reales	Descomposición de error y Efectos Elementales
	EMC	Paegelow et al., 2003	Comparación con datos reales	-
		Phua & Minowa, 2005	Comparación con datos reales	-
		Svoray et al., 2005	Promedio de píxeles de toda la parcela	-
		Molero et al., 2007	-	-
		Chang et al., 2008	Juicio de expertos	Fuzzy Multicriteria Decision-Making (FMCDM)
		do Carmo & Riedel, 2008	-	-
		Geneletti & van Duren, 2008	-	Variación de método de agregación y de pesos
		Beyer et al., 2009	Simulaciones Retrospectivas y comparación por profundidad de distribución de la absorción	-
		Mogaji et al., 2014	Comparación con datos reales	-
	MBA	Bousquet & Page, 2004	Lenguajes gráficos, comparación con otro tipo de modelos, comparación con datos reales y mentir en la representación.	-
		Hawick & Scogings, 2010	Comparación contra otro modelo	-
	RL	Serneels & Lambin, 2001	Comparación con datos reales	-
		Zeng et al., 2008	ROC	-
Estudios de la Sociedad y Economía	MBA	Wise & Crooks, 2012	Comparación contra otro modelo	Variación de entradas
		Crooks & Wise, 2013	Comparación contra otro modelo	Variación significativa de valores alrededor de los valores por defecto.
		Furtado & Eberhardt, 2015	Comparación con datos reales	-
		Pires & Crooks, 2017	Comparación contra otro modelo	-
	EMC	Gómez-Delgado & Tarantola, 2006	Comparación con mapa base	Sobol' y EFAST

Estudios de Riesgos de la Salud		Javaheri et al., 2006	-	-
		Mahini & Gholamalifard, 2006	Promedio de píxeles de toda la parcela	-
		Mena et al., 2006	Comparación con datos reales	-
		Ouma et al., 2011	Juicio de expertos	-
	MBA	Dancik et al., 2010	Validación cruzada	FANOVA
		Crooks & Hailegiorgis, 2013	Comparación con datos reales	-
		Zhou, 2013	-	Índices estadísticos propuestos (local) y AS probabilístico (global)
RL	Paciorek, 2007	Validación cruzada con datos reales, Error medio cuadrático, divergencia Kullback-Leibler	-	
Estudios Urbanos	AC	Li & Yeh, 2000	Medida de pérdida de aptitud e índice de compacidad (evalúa valores para función poder)	-
		Barredo et al., 2003	Matriz de Tabulación Cruzada	-
		Solecki & Oliveri, 2004	Pearson, Comparación con datos reales	-
		Han et al., 2009	Kappa	-
		Basse et al., 2012	Kappa, ROC	-
		Al-Ahmadi, et al., 2013	Validación visual, medidas de exactitud, correspondencia celular urbana, índice de Lee-Sallee, una medida de patrón espacial, una medida de distrito espacial y validación multiresolución.	-
		Gómez et al., 2014	Kappa, radios de distancia y métricas espaciales	-
	EMC	Ligmann-Zielinska, et al., 2008	Comparación con datos reales	-
		Barredo & Bosque, 2009	Matriz de Tabulación Cruzada	-
		Plata et al., 2010	Comparación con datos reales	-
		Balderas et al., 2011	Diferencia media por pixel, Kappa	EFAST y análisis por superposición de mapas
		Lamelas, 2012	Método de concordancia	-

		Plata et al., 2013	Kappa	OAT
MBA		Crooks, 2010	Promedio de Resultados	-
		Yang et al., 2011	Proceso iterativo de calibración de parámetros, comparación con escenario base	-
		Arsanjani et al, 2013	Kappa	-
		Crooks et al., 2014	-	-
		Crooks et al., 2015	Comparación con datos reales	-
		Tan et al., 2015	Kappa	-
		Happenstall & Crooks, 2016	-	-
		Tian et al., 2016	Kappa	-
		Omer & Kaplan, 2017	Comparación con datos reales	-
	RL		Hu & Lo, 2007	ROC
		Tayyebi et al., 2010	ROC	-
		Munshi et al., 2014	Kappa	-
		Feng et al., 2016	Kappa	-
MBA, EMC, RL y AC		Triantakonstantis & Mountrakis, 2012	-	-

Analizando un poco más podemos la Tabla 1 se observó que en el contexto de la simulación geoespacial, se han desarrollado estudios sobre dinámica de usos de suelo, economía, evolución de civilizaciones antiguas y análisis de asistencia humanitaria. Asimismo, en temas de salud hay aportes en los temas de propagación de enfermedades, epidemiología y enfermedades infecciosas, también se han encontrado trabajos acerca de y calidad de agua. En lo que respecta a estudios sobre ecología y biología se han explorado temas referentes a la conservación de bosques, deforestación, segregación animal, gestión de costas, incendios, relleno sanitario y zonificación de corredores ambientales y áreas naturales. Además, se han realizado estudios urbanos relativos a análisis peatonal, segregación residencial, disturbio, sistema de transporte, localización optima de uso residencial, comercial e industrial, y en mayor medida se han encontrado investigaciones de crecimiento urbano.

En el estado del arte se puede encontrar una revisión de trabajos sobre modelos de predicción de crecimiento urbano los cuales utilizan diferentes técnicas de

predicción, describiendo los componentes, ventajas y desventajas de cada una de las técnicas analizadas (Triantakonstantis y Mountrakis, 2012).

La aplicación de cada técnica varía completamente debido a sus componentes, de la misma manera que varían su comportamiento y sus datos de entrada. En las investigaciones más recientes, se ha observado un mayor uso de la teoría de autómatas celulares para crear modelos geoespaciales así como para analizar la complejidad espacial y las propiedades de auto-organización de procesos territoriales. En este sentido fue Waldo Tobler (1979) quien definió, por primera vez, a los AC como modelo geográfico (Kocabas y Dragicevic, 2004).

Los modelos basados en Autómatas Celulares (AC), junto con plataformas SIG, son los más aprovechados para simular situaciones territoriales como crecimiento urbano (Al-Ahmadi, et al., 2013; González, et al., 2015), cambio de uso de suelo (De Almeida & Gleriani, 2005), hidrología (BASSE, et al., 2012), entre otras; ya que tienen la capacidad de predecir el crecimiento en el territorio, basado en la asunción que la actividad pasada afecta los patrones futuros a través de interacciones locales (Sante, et al., 2010).

El uso de esta técnica sugiere una gran variedad de configuraciones para el modelo, con las cuales se puede representar de mejor manera la complejidad y la dinámica de crecimiento urbano. Los elementos básicos del modelo que pueden ser configurados son: espacio de celda (*lattice*), estado de la celda (por ejemplo, el uso de suelo), vecindad de la celda, reglas de transición y tiempo o incrementos temporales (Torrens y O'Sullivan, 2001; Li y Yeh, 2000; Takeyama y Couelalis, 1997; White et al., 1997).

Sin embargo, en otros trabajos se definen más estos elementos, considerando también tamaño de celda, tamaño y tipo de la vecindad, pudiendo ser de 9 celdas (Clarke y Gaydos, 1998) para la vecindad de rectangular de Moore, de 6 u 8 celdas de radio (Barredo et al., 2003) para una vecindad circular, así como la vecindad de tipo Von Neumann (Torrens y O'Sullivan, 2000; White y Engelen, 2000; Yeh y Li, 2001; Kocabas y Dragicevic, 2004; Devendran, 2014).

Por otra parte, para lograr el objetivo principalmente se necesitan cuatro factores que fungen como datos de entrada del modelo: vecindad, accesibilidad, aptitud y aleatoriedad. Dichos factores, intentan reflejar situaciones reales que ocurren en el territorio bajo las siguientes premisas: *Vecindad* es quien determina la influencia que existe sobre el uso del suelo de un píxel y el uso de los píxeles vecinos, *Accesibilidad* misma que radica en la facilidad de acceder a la red viaria y a ciertos elementos del territorio (comercios y centro de negocios, entre otros), *Aptitud* la cual simboliza que capacidad tiene cada parte del territorio para contener alguno de los usos urbanos (se pueden considerar pendiente, altitud, incremento poblacional, distancia a hidrografía, entre otros) y *Aleatoriedad* que tiene como objetivo proporcionar un nivel de incertidumbre que existe de manera natural en los fenómenos de crecimiento urbano (Kocabas y Dragicevic, 2004).

Uno de los elementos más importantes de los AC son las reglas de transición, estas son calculadas a través de un vector de potencial de transición para cada celda en cada simulación, usando (Kocabas y Dragicevic, 2004):

$$P_{hj} = v * a_j * s_j * N + H_j \quad (1)$$

donde las celdas son convertidas en la clase de usos de suelo con más potencial. v es el termino de aleatoriedad, a_j es el termino de accesibilidad, s_j es valor de aptitud del estado de las celdas, N vecindad y H es el parámetro de inercia. El modelo calcula iterativamente las reglas de transición, donde cada iteración representa un año, obteniendo el potencial de cada en valores que comprenden de 0 a 1, siendo la unidad la mayor potencia de cambio.

Un aspecto muy importante a resaltar en estos modelos es que el proceso de calibración para lograr producir resultados consistentes de predicciones futuras, es más importante aún, que la buena definición de las reglas de transición (Batty y Xie, 1994; Clarke y Gaydos, 1998; Yang y Lo, 2003).

No obstante, a pesar de obtener resultados esperados, algunas de las desventajas observadas en el desarrollo de estos modelos, es que esta técnica no toma en cuenta la característica de “parcelas”, por lo que cada píxel es

independiente de los demás y otorga preferencias a la vecindad de los píxeles vecinos además, los mapas de accesibilidad y aptitud necesarios para su aplicación no toman en cuenta todas las variables a la vez, más aún, los resultados son afectados por las diferentes configuraciones del modelo, respecto al tamaño y tipo de vecindad, junto con la resolución espacial.

En cuanto a los recursos de cómputo necesarios para esta técnica, la mayoría de los modelos desarrollados necesitan intensos cálculos computacionales para seleccionar los mejores valores para la exactitud del modelo (Devendran, 2014), debido a la complejidad de los procesos realizados por este modelo.

Respecto a los modelos basados en regresión logística, estas técnicas descubren la relación empírica entre variables categóricas y continuas dependientes binarias y varias independientes. Existen dos enfoques básicos para evaluar la dependencia espacial con regresión: primero se construye un modelo más complejo y segundo se diseña una gráfica de muestreo espacial para aumentar el intervalo de distancia entre puntos. El muestreo sigue a una muestra más pequeña que pierde cierta información y conflictual con el muestreo grande de acuerdo al método de regresión logística en el cual está basado (Triantakoustantis y Mountrakis, 2012).

Asimismo, para esta técnica el uso de una muestra sistemática reduce la dependencia espacial, mientras que el muestreo aleatorio es capaz de representar la población, pero es ineficiente para reducir la dependencia espacial, y particularmente la dependencia espacial local (Huang, et al., 2009), por otro lado el uso de un muestreo aleatorio estratificado elimina la auto correlación (Arsanjani, et al., 2013).

Otro punto, es la variable dependiente predictiva, que es una función de probabilidad (de algún tema en particular, por ejemplo la probabilidad de cambio de una clase de uso de suelo específica). La asunción básica es que la probabilidad de un variable dependiente toma el valor de 1 (respuesta positiva), siguiendo la curva logística y su valor puede ser calculado con la siguiente fórmula (Mahiny & Turner, 2003):

$$P(y = 1|X) = \frac{\exp(\sum BX)}{1 + \exp(\sum BX)} \quad (2)$$

donde, P es la probabilidad de la variable dependiente, X representa las variables independientes y B representa los parámetros estimados.

Ahora bien, para lograr una linealidad en el modelo anterior y remover los límites 1/0 de la variable dependiente original, la cual es la probabilidad, la siguiente transformación es usualmente aplicada (Lewicki y Hill, 2006):

$$P' = \ln\left(\frac{P}{1-P}\right) \quad (3)$$

Esta transformación es referida como la transformación “*logit*”. Una vez realizado lo anterior entonces teóricamente P' puede asumir valores entre +/- infinito. De hecho esta transformación de datos binarios asegura que la variable dependiente será continua y la nueva variable dependiente no está limitada. Además asegura que la superficie de probabilidad será continua en el rango 0 a 1. Para lograr la transformación en ambos lados, se utiliza el modelo de regresión lineal estándar:

$$\ln\left(\frac{P}{(1-P)}\right) = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k + \varepsilon \quad (4)$$

Los modelos de regresión existentes están basados en una sola escala y la resolución está determinada por el nivel de detalle espacial o limitada por el poder computacional, estos no son tan demandantes computacionalmente (Hu & Lo, 2007).

Infelizmente, estos modelos no ofrecen capacidades para modelos con un gran número de factores de entrada y fallan al identificar si las fuerzas de entrada, de las variables dependientes e independientes, no son directamente proporcionales (Triantakoustantis & Mountrakis, 2012). Además la auto

correlación espacial (patrón de comportamiento de la variable según la ubicación geográfica), es frecuentemente ignorada (Hu & Lo, 2007).

Una tercera técnica para la generación de modelos geospaciales es aquella fundamentada en los Modelos Basados en Agentes (MBA). Estos modelos han sido reconocidos como laboratorios computacionales, suministrando a los investigadores una herramienta exploratoria accesible para aprender sobre la dinámica del modelo estudiado (Gómez et al., 2014).

Recientemente estos modelos están siendo usados con mayor frecuencia como herramientas para estudiar el comportamiento de los actores de cambios de uso de suelo y las interacciones entre los tomadores de decisiones y las características de la superficie (Parker et al., 2003; Verburg, 2006).

Hasta la fecha no hay una definición universal aceptada, y se sigue discutiendo sobre cuales propiedades debe tener un objeto para “merecer” ser llamado un “agente”: pro/re activo, la conciencia espacial, capacidad de aprendizaje, habilidades sociales, “intelecto”, etc. (Schieritz y Milling, 2003).

La definición más aceptada es la propuesta por Woolridge y Jennings (1995), quienes en su trabajo describen a un agente como “un programa auto-contenido capaz de controlar su propia toma de decisiones y actuar, basado en la percepción de su ambiente, persiguiendo uno o más objetivos”. Nwana, (1999); mejora la propuesta anterior y define a un agente en términos de tres atributos de comportamiento: Autonomía, Cooperación y Aprendizaje; dos de los cuales debe poseer un agente de software.

Sin embargo, una definición importante para temas de ecología y ciencias ambientales es la propuesta por (Ferber, 1995 y 1999), quien define que un sistema multi-agente está compuesto de: un ambiente, que generalmente es el espacio donde evolucionará el agente, un conjunto de objetos, esos objetos son situados en una posición en el ambiente, un conjunto de agentes, los cuales son objetos específicos (un subconjunto del conjunto de objetos) y representan las entidades activas en el sistema, un conjunto de relaciones, que enlazan objetos

(agentes) a algún otro, un conjunto de operaciones, haciéndole posible a un agente percibir, producir, transformar y manipular objetos, y operadores.

Los MBA basan sus componentes en la información del sistema real, definiendo adecuadamente las *fronteras* entre ambos sistemas, así como las *interacciones*, mismas que coexisten entre los agentes del modelo. Esta inter-relación es capaz de aumentar el realismo e impacto de los modelos formales construidos.

El comportamiento deseado de estos agentes (autonomía, heterogeneidad e independencia), así como sus metas y objetivos propios se definen a nivel individual, mientras que el comportamiento global surge como resultado de muchos (decenas, cientos, miles o millones) individuos (Torsun, 1995; Izquierdo et al., 2008). Es por ello que esta técnica tiene un comportamiento de modelado bottom-up (Borshchev y Filippov, 2004).

Las ventajas de los MBA es que estos pueden ser construidos en ausencia de conocimiento de interdependencias globales, es fácil de actualizar y que para la exploración de uso de suelo, la incorporación explícita de toma de decisión heterogénea, tiene la facilidad en la forma de variar preferencias, reglas de decisión y percepción de riesgo (Ligmann-Zielinska, 2010; Borshchev y Filippov, 2004).

En cambio la gran desventaja es el coste computacional, que aumenta con el tamaño del entorno en el que el agente realizará los cálculos para elegir las acciones correspondientes.

Por otra parte, existen los modelos basados en redes neuronales artificiales (RNA), estos modelos son heurísticos y copian la estructura de un sistema neural biológico con el fin de alcanzar una funcionalidad similar. Esta técnica todavía sigue vigente y es usada en muchas disciplinas hoy en día (Gómez, 2012).

Uno de los objetivos primordiales de esta técnica es dotar de capacidad de aprendizaje a los agentes utilizados en el MBA o a los AC (Lee, et al., 2013; Moghaddam & Samadzadegan, 2009; Hegde, et al., 2008). La relación entre

acción y percepción es modelada por una red que involucra las conexiones existentes. Los investigadores pueden de una manera muy fácil definir conocimiento experto por medio de la lógica del perceptrón o una red neuronal (Bousquet and Le Page, 2004).

Una red neuronal consiste de tres componentes: capa de entrada, capa oculta o intermedia y capa salida. El procesador elemental o neurona es un dispositivo simple de cálculo que, a partir de un vector de entrada del exterior o de otra neurona, proporciona una única respuesta o salida, de esta manera una red neuronal consiste en un conjunto de nodos o neuronas agrupadas en varias capas e interconectadas entre sí: capa de entrada, capas ocultas/intermedias y capas de salida. Las funciones para direccionar la interacción entre neuronas son muy simples. Si p es la neurona que envía la información en la capa de entrada y q es la neurona que recibe en la siguiente capa, la función de colección está dada por (Kia et al., 2011):

$$net_q = \sum_p w_{pq} I_p \quad (5)$$

donde I es la señal de la neurona p de la capa de envío, net es la señal recogida para la neurona que recibe q en la siguiente capa y w es el parámetro o peso para sumar la señal de diferentes nodos de entrada.

Las conexiones que unen a las neuronas tienen asociado un peso numérico, y el aprendizaje se realiza con la actualización de tales pesos. Se denomina época al periodo de tiempo que comprende todas las iteraciones realizadas al presentar todos los patrones (conjunto de datos de entrenamiento y calibración) una sola vez. La actualización de pesos se realiza al final de cada época (Gómez, 2012).

Las capas y neuronas permiten que la RN tenga la capacidad de aprender y recordar como un humano, y con ello tener la habilidad de clasificar un patrón objetivo, ajustando los valores de los pesos de cada factor, y este proceso es conocido como “fase de entrenamiento” o “lógica de aprendizaje supervisado” (Fischer, 1998).

El mecanismo de esta técnica la hace capaz de aprender decisiones tomadas a partir de información incompleta, ruidosa o borrosa y es por esta razón es por la que puede ser fácilmente adaptable para manejar problemas espaciales (De Almeida & Gleriani, 2005).

Algunas de las ventajas de usar esta técnica de simulación es la resistencia al ruido de los datos, optimización y aproximación, técnica de manejo de datos y la buena habilidad para predecir. La desventaja más importante es que básicamente la ejecución del modelo de red neuronal, solo calcula resultados con tiempo estático, por lo que es difícil usarlo para desplegar una demostración dinámica y estocástica, además se debe mencionar que este dispositivo es una caja negra y la intervención del usuario es aun decisivo para la calidad de los resultados. (Fischer, 1998; De Almeida & Gleriani, 2005).

Por último, el análisis o evaluación multicriterio (EMC) se define como un conjunto de técnicas orientadas a asistir en procesos de decisión y están fundamentadas en la ponderación y compensación de variables que van a influir de manera positiva (aptitud) o negativa (impacto) sobre la actividad objeto de decisión y que deben ser inventariados y clasificados previamente (Gómez Delgado y Barredo, 2005).

En este tipo de métodos, un objetivo se puede entender como una función a desarrollar, donde el objetivo indica la estructuración de la regla de decisión o el tipo de regla que se debe de utilizar. Estos objetivos deben de ser planteados sobre una serie de alternativas y cada alternativa está caracterizada por los diferentes criterios o factores que constituyen la evaluación (Eastman et al., 1993).

Se debe mencionar que la EMC no se ha desarrollado totalmente en el campo de los SIG, lo que genera una diferencia entre las terminologías utilizadas para construir el modelo y definir los elementos que lo componen. Sin embargo, se han descrito algunas reglas conceptuales para el uso de la EMC en SIG (Eastman et al., 1993; Barredo, 1996; Malczewski, 1999).

Uno de los métodos más usados para caracterizar los factores, es la técnica de comparación por pares de Saaty, esta establece una matriz cuadrada representando el número de factores a ponderar por filas y por columnas, comparando así la importancia de un factor sobre cada uno de los demás; determinando, posteriormente se establecen los pesos y se proporciona una medida de la consistencia de los juicios de valor entre pares de factores (Saaty, 1980).

Por lo tanto, se puede decir que los criterios son uno de los aspectos más importantes en la EMC, ya que este tiene la finalidad de fungir como la base para tomar una decisión y este se puede medir y evaluar, en otras palabras es la evidencia en que se basa una decisión (Eastman et al., 1993). Debido a la relevancia de este componente, es que se toma como un primer paso de la EMC la puntuación de los criterios, (valoración de las alternativas) y es aquí donde recae una gran parte del éxito o del fracaso de la evaluación ejecutada.

Una vez logrado este paso, los factores pueden ser combinados a través de alguna regla de decisión y calcular la evaluación. La regla de decisión no es otra cosa más que la selección entre alternativas (o posibilidades de elección) y es por medio de esta que se obtiene una evaluación en particular y pueden ser tan simples o complejas como el problema lo requiera. El valor resultante de la evaluación es asignado a cada dato (alternativa espacial), orden o rango de capacidad de acuerdo a la actividad evaluada (Gómez Delgado & Barredo Cano, 2005).

Las diferentes reglas de decisión se diferencian por la forma de realizar los cálculos aritmético-estadísticos, utilizando las matrices de evaluación y de prioridades por medio de lo cual se calcula la evaluación final de las alternativas. Los procedimientos pueden ser tan sencillos como el método de Sumatoria Lineal Ponderada, o pueden requerir operaciones más difíciles como las técnicas de Análisis del Punto Ideal, Análisis de Concordancia-Discordancia, Optimización Jerárquica, Programación Lineal, entre otras (Gómez Delgado & Barredo Cano, 2005).

En el caso de la técnica de Sumatoria Lineal Ponderada (SLP), es uno de los métodos más utilizados por tener un procedimiento sencillo, intuitivo y fácil de implementar para el análisis de localizaciones óptimas (Eastman, 1997; Heywood et al., 2002; Janssen y Rietveld, 1990; Malczewski y Rinner, 2005; Santé, 2007). Para esta técnica todos los factores (archivos raster independientes), son multiplicados por sus correspondientes pesos y son sumados (Gracia, 2012). La obtención de valor de adecuación de las alternativas se calcula con:

$$r_i = \sum_{j=1}^n w_j v_{ij} \quad (6)$$

donde w es el peso y v el valor de cada criterio.

A continuación, el método de Análisis del Punto Ideal (API) fue propuesto por Zeleny (1973, 1974), y se basa en el cálculo de los desvíos de cada alternativa con el punto ideal, el cual se debe de considerar como inalcanzable para después comparar las distancias entre cada alternativa y el ideal, donde se representa cada criterio como un eje y se logra un espacio multivariable.

Como el punto ideal no existe, la función es establecer el punto de referencia para todas las alternativas, y de esta manera la mejor solución se define como la de menor distancia al inaccesible punto ideal. Entre las distancias utilizadas se pueden mencionar: la distancia Euclidiana y la Manhattan.

$$Euclidiana(P_1 - P_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \quad (7)$$

$$Manhattan_{ij} = |X_i - X_j| + |Y_i - Y_j| \quad (8)$$

Por otra parte, los métodos Conjuntivo y Disyuntivo, asignan un umbral a cada criterio por parte del centro decisor, siendo eliminadas las alternativas que estén

por debajo (conjuntiva) que sobrepasen (disyuntivo) del tal umbral en al menos uno de los criterios.

Estos métodos requieren un nivel más bajo de conocimiento por parte del centro decisor, pero tienen como principal desventaja que recomiendan ocasionalmente en su ejecución una alternativa comparativamente peor, debido a la estrategia de reducción que contempla.

En Plata (2010) se utiliza el método conjuntivo para eliminar alternativas, no permitiendo que los valores alcanzados en otros factores actúen de modo compensatorio, para ellos se requiere que se especifique un umbral como valor mínimo (o máximo) aceptable para cada factor.

Otra técnica que está basada en la teoría de lógica borrosa, y más que un método simple, es una técnica de tres pasos llamada Sumatoria Lineal Ponderada Ordenada (OWA), la cual contempla 3 funciones de agregación como lo son la intersección, unión de conjuntos borrosos y los operadores de medias. En general, consiste en utilizar los pesos de los factores y otra serie de pesos “ordenados” que controlan como se agregan esos factores ponderados, de esta manera es posible conocer el nivel total de compensación permitido. Utilizando estos pesos ordenados es posible realizar combinaciones entre ellos que permiten posicionarse en cualquier punto del espacio triangular de decisión (Feizizadeh et al., 2014; Feizizadeh et al., 2012; Feizizadeh, 2014b; Abdullah, 2013).

El paso final de las técnicas de EMC, es la evaluación, esta contempla alcanzar varios objetivos específicos y los cuales indican como va a actuar la regla de decisión, por lo que los criterios son entonces estructurados en función de los objetivos propuestos, es decir, cada objetivo puede representarse por uno o varios criterios, quienes al maximizarse o minimizarse alcanzan el objetivo (Gómez Delgado & Barredo Cano, 2005).

Por otro lado, una desventaja es que al momento de plantearse el problema con una técnica multicriterio, los procedimientos en sistemas informáticos actuales,

son rebasados en cuanto a su capacidad operativa (Gómez Delgado & Barredo Cano, 2005).

Se debe mencionar además, que la integración de modelos de simulación (AC, MBA, regresión lineal y EMC) y los SIG ha avanzado de manera considerable en los últimos años, por el contrario los resultados de estas combinaciones siguen usándose deliberadamente y no se les realiza algún tipo de análisis de robustez que proporcione algún indicador de confianza (Feizizadeh et al., 2013).

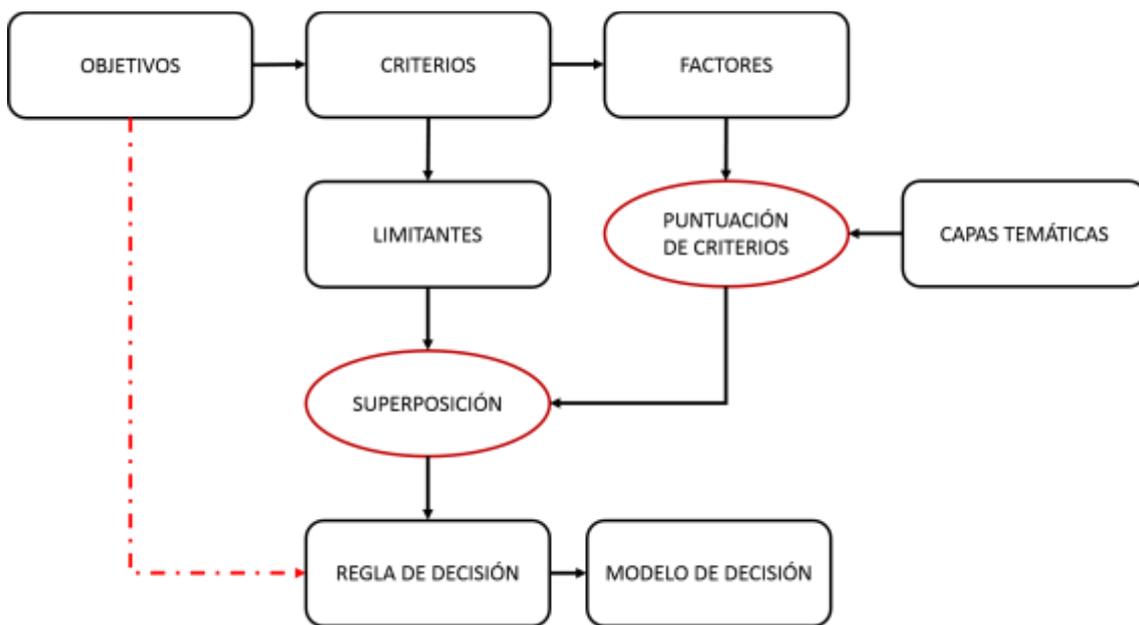


Figura 1.- Diagrama de Integración entre un SIG y EMC.

En conclusión, estos modelos en la actualidad son implementados computacionalmente y son comúnmente usados para intentar predecir el comportamiento de sistemas en cualquier área de estudio. Las diferencias entre estos se encuentran en los pasos, la complejidad del procedimiento y en la manera de abstraer la información de los procesos más importantes.

Hay que mencionar además, que debido a que algunos modelos se pueden calibrar por medio de observación real, y con ello mejorar los resultados, los modelos computacionales son vistos como dispositivos de análisis, por medio de los cuales las propiedades (descriptores) más importantes, para trabajar un área de estudio en particular, pueden ser inferidas.

Con respecto a la validación de estos, distintas técnicas han sido propuestas y están relacionadas con la selección del método. Algunos de los procedimientos más usados son la familia de análisis Kappa, Características Operativas Relativas y la matriz de confusión (Torrents, 2011; Kocabas & Dragicevic, 2004; Wang, et al., 2012).

2.2 Validación de los modelos de simulación

El objetivo principal de la validación es la demostración de la capacidad que tiene un modelo para predecir (Qureshi, 1999). Esto justifica la salida del modelo en términos de la realidad (Ahmed, et al., 2013).

En el caso de la simulación de escenarios (mapas), se pueden observar diferentes técnicas de validación basadas en la comparación de mapas. Estos métodos han sido mejorados en el transcurso de las últimas décadas y abarcan procedimientos simples tales como comparación visual de resultados (Barredo et al., 2003; White y Engelen, 1997), índices estadísticos (Pontius, 2002), hasta medidas de comparación difusas (Hagen, 2003). A continuación, se describen los más importantes.

La técnica de *traslape espacial o validación visual* compara el ajuste de las celdas, una a una. En otras palabras, se evalúa la correspondencia entre el uso de suelo observado y el simulado a través de celdas individuales en la misma coordenada del *grid*, mientras que su distribución espacial (tal como diferentes usos de suelo son conectados) es ignorada. Por lo que una alta tasa de correspondencia, no asegura que la simulación produzca una morfología urbana convincente, por el contrario, una baja tasa de correspondencia no necesariamente significa pobre coincidencia espacial, puesto que el patrón de uso de suelo simulado y el observado pueden ser similares a cada uno en términos de estructuras espaciales.

Para mejorar el método anterior se sugiere contemplar la conformidad estructural que es importante en la evaluación de los resultados simulados. Esta conformidad entre los usos de suelo observados y simulados puede ser analizada a través de estadísticas espaciales, además el indicador puede ser

usado para revelar el patrón de agrupamiento del mismo tipo de uso en celdas adyacentes (Wu, 2002; Plata, et al., 2013; Óran, et al., 2010; Óran, et al., 2010).

A continuación una técnica extendida para reconocimiento de patrones es la *Coincidencia de Pixeles*. Esta involucra la coincidencia de pixeles en el mapa, lo cual determinar la presencia de características específicas, artefactos o configuraciones que son conocidos de forma *a priori* (Torrens, 2011). El objetivo es registrar los pixeles que han cambiado de una imagen a otra y la coincidencia entre las dos es medida.

Uno de los primeros métodos basados en la comparación celda a celda, y que provee información sobre la coincidencia espacial de los resultados, es el *Índice Kappa*. El cálculo de este índice está basado en la matriz de tabulación cruzada.

Esta técnica propone una matriz para comparar dos variables categóricas, mostrando una tabla con las clases de una variable como los renglones y las clases de otras variables como las columnas. La comparación se basa en mapas raster, donde la diagonal de la matriz obtenida equivale a las concordancias entre los mapas y, los elementos fuera de la diagonal son desacuerdos de pixeles. Clase por clase es comparada entre el total de filas y columnas permitiendo a los interesados ver como los mapas se relacionan en términos de cantidad de cada clase (Gómez Delgado & Barredo Cano, 2005).

El índice Kappa tiene entre sus propiedades deseables que si la clasificación es perfecta el índice es igual a uno, si una proporción es observada correctamente es menor y, si es menor que la proporción esperada correcta debido al cambio, es menor que cero (Ahmed, et al., 2013).

Otros índices Kappa existentes son: K_{no} , $K_{location}$ y $K_{quantity}$. Estas tres técnicas estadísticas están basadas en el Kappa estándar, sin embargo, dan como resultado información adicional, con lo que se pretende dar más información de la comparación de los mapas observado y simulado. El índice K_{no} calcula la exactitud total de la simulación obtenida, por otra parte, el índice $K_{location}$ valida la habilidad de la simulación para predecir localizaciones mientras que $K_{quantity}$ hace

lo mismo pero para cantidades. Todos ellos son iguales a uno cuando la evaluación es perfecta y cero cuando la simulación fue un cambio exitoso (Schneider & Pontius Jr., 2001).

Por más de una década estos populares índices de validación han sido comúnmente utilizados para identificar la exactitud en distintos modelos (Koomen et al., 2007; Tenerelli y Carver, 2012; Óran, et al., 2010; Gómez Delgado et al., 2014; Plata et al., 2013; Basse et al., 2012; Rienow y Stenger, 2014), sin embargo, se recomienda aplicar con precaución los índices Kappa para propósitos de evaluación de exactitud y comparación de mapas ya que se consideran los pixeles que permanecen estables como buenos, una mejor forma de analizar la superficie es usando la matriz de tabulación cruzada con dos parámetros resumidos más simples que son la cantidad de desacuerdo y asignación de desacuerdo (Pontius & Millones, 2011).

La *matriz de tabulación cruzada* ha sido ampliamente usada para comparar resultados simulados contra resultados reales en diferentes estudios (Li y Yeh, 2001; Wu, 2002; Basse et al., 2012).

En esta esta matriz de error, las filas y columnas representan el número de categorías cuya exactitud de clasificación está siendo evaluada. Las columnas son el dato observado y las filas indican el dato simulado. Esta evalúa el total de rendimiento del modelo calculando la proporción del total de número de celdas simuladas que coinciden en el valor de la imagen. Los números en la diagonal representan las correctas decisiones tomadas, por otra parte todos los números fuera de la diagonal son errores entre las clases. Las tasas de verdaderos positivos y falsos positivos pueden calcularse de la siguiente manera (Omrani, et al., 2012):

$$\text{Verdaderos Positivos} = \frac{\text{Positivos correctamente clasificados}}{\text{Total de positivos}} \quad (9)$$

$$FalsosPositivos = \frac{Positivosincorrectamenteclasificados}{Totaldenegativos} \quad (10)$$

El tipo de validación basado en la matriz de confusión es más popular que Kappa, es por ello que la mayoría de los modelos de simulación han utilizado esta medida (Parker et al., 2003).

Por otra parte el *índice de Lee-Sallee* es una medida que toma en cuenta la forma (a diferencia de la matriz de confusión), este método ha sido frecuentemente usado para evaluar la forma urbana producida. Este índice es calculado como el radio de la intersección entre el área urbana observada y simulada contra la unión de esas áreas en dos imágenes (Clarke & Gaydos, 1998; Barredo, et al., 2004). El cálculo se lleva a cabo de la siguiente manera (Al-Ahmadi et al., 2013):

$$LSI = \frac{\sum(S_{ij} \cap O_{ij})}{\sum(S_{ij} \cup O_{ij})} \quad (11)$$

donde S_{ij} es una celda urbana simulada ij y O_{ij} es una celda urbana observada ij .

Existe también otra medida de validación que considera la forma, y es llamada *Medida de Patrón Espacial*. La mayoría de los análisis celda por celda ignoran la presencia de celdas vecinas o vecindarios. La medida de patrón espacial considera una celda como errónea si la categoría en el mapa observado difiere de la categoría en el mapa simulado, independientemente de que la categoría se encuentra en la celda vecina o en ninguna parte cerca de la celda. De esta manera se evalúa el rendimiento basado en la coincidencia en un vecindario. Para lograr la asignación de valores a la salida, se puede calcular la siguiente expresión (Al-Ahmadi et al., 2013):

$$IF\sum\Omega S_{ij} = \sum\Omega O_{ij} \quad (12)$$

entonces SPM_{ij} es igual a 1, de otra manera SPM_{ij} es igual a 0;

Donde ΩS_{ij} es el número de celdas urbanas simuladas ij con una vecindad Ω ; y ΩO_{ij} es el número de celdas urbanas observadas ij en una vecindad Ω . Para calcular esta medida, una matriz base especial es diseñada como una medida de vecindad para simular la forma urbana común.

En otro orden de ideas-], una medida que valida la similitud estructural entre el crecimiento urbano observado y el simulado en términos de vecindad urbana es la *Medida de Estructura de Distrito Espacial*. También podría ser posible evaluar en términos de sub-vecindad urbana y bloques urbanos, donde los límites de esas formas deben ser observados previamente. Este método considera la escala espacial o la resolución de los resultados del modelo.

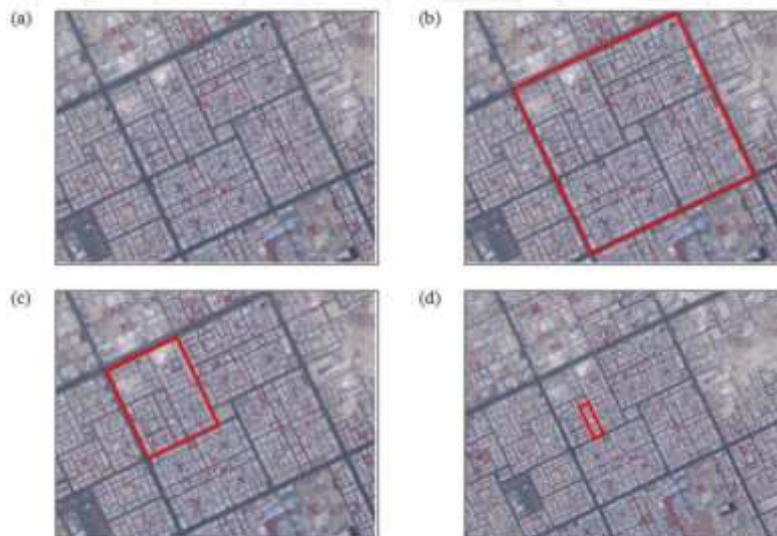


Figura 2.- a) Una sección de la ciudad de Riyadh, b) vecindad urbana, c) sub-vecindad urbana y d) un bloque urbano (Al-Ahmadi et al., 2013).

Otras técnicas de validación son las denominadas de *Validación Parcial*, entre ellas se encuentran los buffers de distancia y las medidas espaciales. Los *Buffers de Distancia* pueden capturar las diferencias cualitativas o similitudes entre el índice de dos mapas (Power et al., 2001), diferencia de la comparación visual o índice Kappa (Tveite y Langaas, 1999).

Esta técnica puede ser usada para generar zonas de influencia alrededor de las nuevas celdas creadas y de esta manera determinar si las celdas obtenidas de los otros resultados simulados están localizados en esa área. De esta manera se puede calcular el porcentaje de uso de suelo simulado dentro y fuera del área de influencia (González et al., 2015).

Mientras que las *Medidas Espaciales* (Herold et al., 2003) son aplicadas a los mapas obtenidos y su comparación, esta provee información sobre la forma, dispersión y fragmentación del suelo, habilitando una comparación de resultados en términos de patrones de usos de suelo. A esta medida también se le conoce como *Índice de fragmentación* (Mas et al., 2012).

Por su parte el análisis de *Características Operativas Relativas* (ROC), es una técnica de validación que ha sido introducida últimamente a modelos de cambio de uso y cobertura de suelo, para calcular la correlación entre los cambios actuales y los cambios simulados (Amin, 2013; Basse et al., 2012; Hu y Lo, 2007; Feizizadeh et al., 2013; Rana et al., 2008). La técnica de ROC se puede aplicar a cualquier método que prediga una categoría homogénea en cada celda del grid. En otras palabras es una gráfica de la probabilidad de tener verdadero positivo de crecimiento urbano contra la probabilidad de un falso positivo identificado como la probabilidad de corte varia (Arsanjani, et al., 2013).

Esencialmente evalúa que tan bien un par de mapas coinciden en términos de localización de píxeles desarrollados. Necesariamente si ROC es igual a uno indica un ajuste perfecto, en cambio si es igual a 0.5 indica un ajuste aleatorio. Cada escenario podría corresponder a un punto en la gráfica. Y se puede calcular el área bajo la curva que conecta los puntos graficados como en (Pontius Jr. & Schneider, 2001):

$$\text{Área bajo la curva} = \sum_{i=1}^n [x_{i+1} - x_i] [y_i + y_{i+1} - y_i/2] \quad (13)$$

Donde x_i es la proporción falsos positivos para el escenario i , y_i es la proporción de verdaderos positivos para el escenario i , y n es el número de grupos de idoneidad.

Si se analiza una medida de patrón espacial contra las técnicas de tabulación cruzada o con el índice Kappa, la forma de comparación de estas últimas es muy rígida, debido a que compara cada celda del mapa con otra pero no considera las celdas vecinas. Además, la comparación de este tipo es menos efectiva cuando se analizan patrones y formas de crecimiento urbano, donde es más apropiado estudiar, como las formas de fragmentos urbanos varían de un mapa a otro (Li et al., 2001).

La desventaja de la estadística del Kappa es que no es adecuadamente sensible. Si una escena simulada es similar a una escena observada, pero la correspondencia no coincide con sólo un píxel en unos pocos lugares, la precisión global de coincidencia puede sufrir considerablemente (Torrens, 2011).

En el caso de las técnicas que hacen uso de la matriz de confusión, estas tienen la desventaja de que no son capaces de evaluar y estimar la forma y/o formas de patrones porque están basadas en la comparación independiente entre pares de celdas. Por otra parte la validación o traslape espacial se basan en el juicio de los investigadores, por lo que necesitan más métodos objetivos, sin embargo es una parte esencial para la validación (Al-Ahmadi, et al., 2013).

Sin embargo, el proceso de validación generalmente se lleva a cabo comparando los resultados obtenidos contra los datos reales (Qureshi et al., 1999). No obstante, la principal desventaja en la simulación de escenarios geoespaciales de futuro es que no es posible conocer los datos reales para poder comparar y verificar los resultados obtenidos (Plata et al., 2012).

Debido a esto, una validación de los resultados es imposible, en cambio, una alternativa para valorar el grado de confiabilidad en modelos de simulación geoespaciales, es aplicar un análisis de sensibilidad (Pannell, 1997; Malczewski, 1999; Qureshi et al., 1999; Barredo y Gómez, 2008).

En el contexto de la simulación geoespacial, se han desarrollado una gran cantidad de estudios correspondientes a dinámica de usos de suelo, propagación de enfermedades, residuos, hidrografía, economía, bosques, medioambiente, entre otros. Para la simulación de dichos modelos se emplean técnicas de Evaluación Multicriterio (EMC), Autómatas Celulares (AC), Modelos Basados en Agentes (MBA), Regresión Lineal (RL) (ver figura 6).

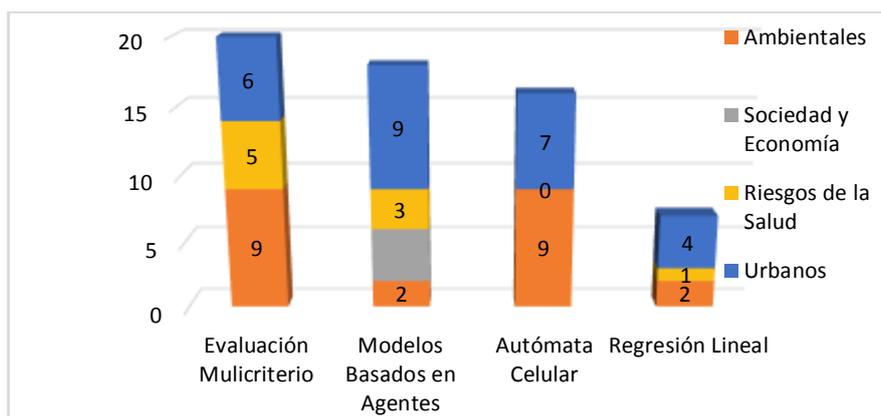


Figura 3.- Artículos según la técnica de simulación y su objeto de estudio.

Analizando la figura 6, se puede observar que en temas ambientales, las técnicas de simulación más utilizados son los autómatas celulares y la evaluación multicriterio con un 37%, respectivamente, de los trabajos consultados. Además, el método de validación más utilizado es el índice Kappa con el 59%.

Por otro lado, se hace notar que solo 4 trabajos aplican la simulación a temas de estudios de la sociedad y economía, de los cuales el 100% utiliza modelos basados en agentes y valida y/o calibra mediante los índices Kappa.

Asimismo, se puede ver que los modelos simulación basados en EMC, son los más utilizados en problemas relacionados con riesgos de la salud, con un 55% del total analizado, seguido de los modelos basados en agentes con 33%. Aunado a esto, el 55% se basa en una validación mediante índices Kappa y el 33% del total analizado contempla un análisis de sensibilidad y a su vez el 100% hacen un estudio de tipo global.

También se observa el uso de los MBA, AC, EMC y RL en simulaciones para resolver cuestiones urbanas con un porcentaje de 33, 26, 22 y 14 respectivamente. Además, del total de la consulta el 66% utilizan el índice kappa como técnica de validación/calibración. Así mismo, solo un 7% aplica un estudio del modelo a partir de un análisis de sensibilidad y de los cuales un es de tipo regional (OAT) y otro de tipo global y regional (EFAST y superposición de mapas).

Con respecto al uso de los diferentes métodos de simulación, se puede destacar que los modelos basados en EMC, son los más utilizados con un 33%, seguido de los MBA con 29% (ver figura 6).

Por lo que se refiera al porcentaje de trabajos realizados en los diferentes temas analizados, los estudios urbanos son las más analizados con un 43% del total revisado, mientras que los estudios ambientales, riesgo de salud y sociedad y economía alcanzan un 36%, 15% y 6% respectivamente (ver figura 7)

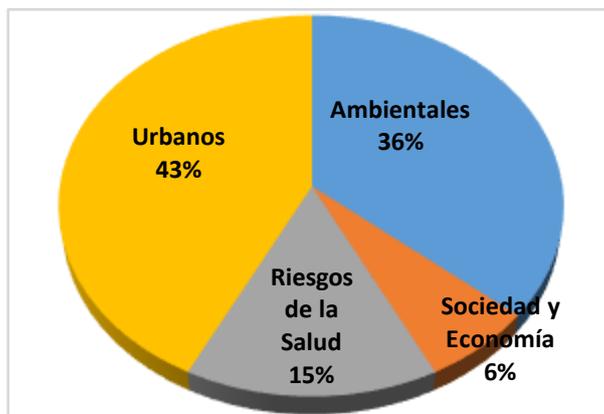


Figura 4.- Porcentaje de estudios según el objeto de estudio.

El número de estudios que aplican simulación para resolver problemas, ha ido en aumento en los últimos años y se puede observar en la figura 8, donde se observa que se encontró una mayor cantidad de trabajos para el año 2013.

Por su parte, la técnica de validación y/o calibración que se aplica en mayor proporción es el índice Kappa con un 65%.

Por más de una década los índices Kappa han sido comúnmente utilizados para identificar la exactitud en distintos modelos (Koomen et al., 2007; Kocabas y Dragicevic, 2004; Tenerelli y Carver, 2012; Óran, et al., 2010; Gómez-Delgado et al., 2014; Plata et al., 2013; Basse et al., 2012; Rienow y Stenger, 2014), sin embargo se recomienda aplicar con precaución los índices Kappa para propósitos de evaluación de exactitud y comparación de mapas ya que se consideran los pixeles que permanecen estables como buenos, una mejor forma de analizar la superficie es usando la matriz de tabulación cruzada con dos parámetros resumidos más simples que son la cantidad de desacuerdo y asignación de desacuerdo (Pontius & Millones, 2011).

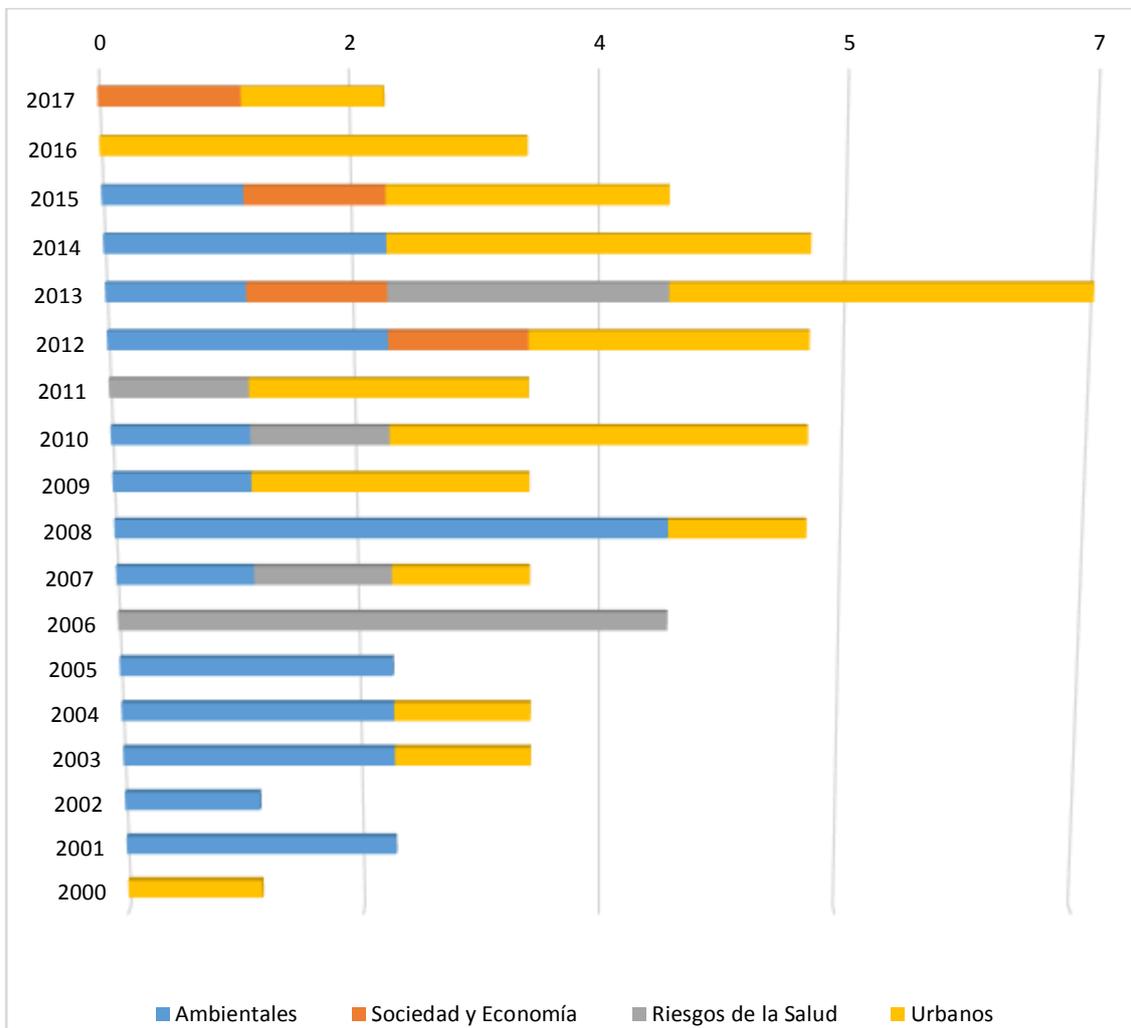


Figura 5.- Numero de estudios por año.

En este sentido, se ha prestado mucha menos atención a la credibilidad (robustez) de los modelos, que al desarrollo del mismo. La aplicación de un AS

no se considera generalmente como un componente crítico en el desarrollo de modelos bien estructurados tales como modelos de programación matemática, ya que la estructura del modelo es bien conocida, y razonablemente bien aceptada.

Los objetivos de la aplicación de un AS son: priorización de factores, verificación, entendimiento y simplificación del modelo, además una vez construido el modelo se pueden analizar las consecuencias que resultan al cambiar un parámetro dado y de esta manera proporciona cierto grado de confiabilidad en un modelo e indica las áreas prioritarias para refinar en el modelo (Iman y Helton, 1988; Hamby, 1994; Hamby, 1995; Qureshi et al., 1999; Lilburne y Tarantola, 2009; looss y Lemaître, 2015).

Asimismo, el AS también puede ayudar a reducir la dimensionalidad de los modelos, disminuir el costo computacional y optimizar los recursos. Es por lo tanto, fundamental para dar robustez y credibilidad a cualquier modelo (Saltelli et al., 2000; Gómez y Barredo, 2005; Saltelli et al., 2008).

Del análisis de estudios realizado se observa que en temas ambientales, de la sociedad y economía, riesgos de la salud y estudios urbanos solo el 18%, 50%, 33% y el 7% respectivamente (ver figura 9), consideran la aplicación de un análisis de sensibilidad y en su mayoría son de tipo global.

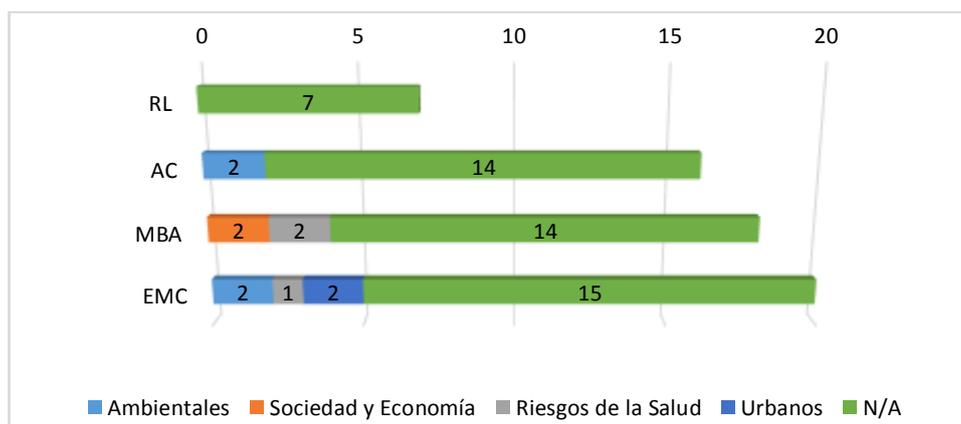


Figura 6.- Estudios que contemplan análisis de sensibilidad.

Esto quiere decir que, con respecto a la examinación del modelo, sus componentes y resultados, solo un 18% del total consideran la aplicación de un

AS (ver figura 9), lo que conlleva a confirmar lo dicho por Gómez y Bosque (2004).

2.3 Análisis de sensibilidad en modelos espaciales

Existe una variedad de técnicas de Análisis de Sensibilidad (AS) que pueden ser empleadas, la mayoría de las cuales implican múltiples simulaciones del modelo de interés, y es debido a esta multidimensionalidad que el análisis de sensibilidad resulta bastante oneroso. De hecho, algunos de los métodos simplemente no son prácticos, mientras que otros tienen significantes limitaciones.

La mayoría de los métodos para realizar análisis de sensibilidad en modelos espacial son de tipo estadístico, por otra parte para un mejor análisis es necesario el uso de técnicas de AS de tipo espacial algunas de las cuales se describen a continuación: Enfoque OAT es el método más común por su sencillez. Por ejemplo, cada entrada puede variar en una cantidad dada, puede ser ± 10 o 20% (Swartzman y Kuluzny, 1987), espacialmente se intercambia una proporción variable de los pixeles o se intercambian ubicaciones (Fisher et al., 1997) de acuerdo a una distribución de probabilidad (Stoms et al., 1992).

Este método realiza el análisis de tipo local (Lodwick et al., 1990), fijando valores nominales a los parámetros que no son variados, por lo que el AS se limita a un área pequeña del espacio de parámetros. Debido a esto los resultados del AS pueden no ser válidos de manera general y el error generalmente se modela como un escalar, por ejemplo, el incremento de la caída de lluvia por una cantidad fija.

Después tenemos el método de *muestreo de Hipercubo Latino (LHS)*, usado para reducir el número requerido de corridas del modelo necesarias para tener suficiente información sobre la distribución en la salida, pertinente para ahorrar tiempo de cómputo (Makowski, et al., 2006).

Es un método de muestreo de tipo estratificado y genera muestras de entrada variables, en cuanto a su funcionamiento, sus entradas son divididas en N

intervalos equivalentes, donde N es el número de simulaciones. También en valor de parámetro representativo para cada valor es aleatoriamente seleccionado, después estos valores son combinados y cada valor es considerado una sola vez en el proceso de simulación, este método cubre completamente del rango de las variables de entrada.

Luego se encuentra el método de Morris, adecuado cuando el número de factores es muy alto y/o el modelo es caro para ser calculado. El método de Morris está basado en el cálculo de Efectos Elementales (EE). Fue probado para análisis de grandes modelos (Campolongo y Braddock, 1999), pero no es tan usado. Para cada entrada dos medidas de sensibilidad son calculadas:

$$-\mu_j^* = \frac{1}{r} \sum_{i=1}^r |E_j^i| \quad (25)$$

Media del valor absoluto de los efectos elementales

donde μ evalúa la sobre influencia del factor sobre la salida, y

$$-\sigma_j = \sqrt{\frac{1}{r} \sum_{i=1}^r (E_j^i - \frac{1}{r} \sum_{i=1}^r E_j^i)^2} \quad (26)$$

Desviación estándar de los efectos elementales

σ estima el conjunto el efecto de orden más alto del factor.

Con este método los parámetros menos importantes son desechados, seleccionando del resto los más importantes y usando alguna de las técnicas de selección.

A cerca del método de Prueba de Sensibilidad de Análisis de Fourier (FAST por su siglas en ingles), este fue desarrollado por Cukier et al., (1975) y se encarga de convertir una integral p -dimensional a una unidimensional. El método permite a todos los parámetros de entrada ser variados simultáneamente de modo que

los efectos interactivos entre los parámetros de entrada sobre la salida pueden ser adecuadamente estudiados y contempla diferentes distribuciones.

Está basado en una transformada multidimensional de Fourier. Este procedimiento obtiene las medidas de importancia y necesita un pequeño número de ejecuciones del modelo, y usa la siguiente aproximación de Fourier:

$$E(Y) \cong \frac{1}{2\pi} \int_{-\pi}^{\pi} F(s) ds \quad (27)$$

donde

$$F(s) = f[g_1(s), \dots, g_k(s)], \text{ donde } x_i = g_i(s) = G_i(\sin\omega_i s)$$

para funciones apropiadas de G_i y frecuencias ω_i . Y usando la representación por series de Fourier para $f(s)$, se obtiene:

$$\text{var}(Y) \cong \sum_{h=1}^{\infty} (A_h^2 + B_h^2), V_i = \sum_{h=1}^{\infty} (A_{h\omega_i}^2 + B_{h\omega_i}^2) \quad (28)$$

donde A_h y B_h son los coeficientes de Fourier dados por:

$$A_h = \frac{1}{\pi} \int_{-\pi}^{\pi} F(s) \cos(hs) ds \quad (29)$$

Y

$$B_h = \frac{1}{\pi} \int_{-\pi}^{\pi} F(s) \sin(hs) ds \quad (30)$$

Ha sido usado para resolver ecuaciones acopladas de rangos no lineales y también, usado para problemas de cinética química de modelos atmosféricos y estudios de difusión atmosférica (Dickinson y Gelinas, 1976; Leipmann y Stephanopoulos, 1985). Saltelli et al. (1999) propone una modificación a FAST y desarrolla la técnica FAST extendida (eFAST), Dado que FAST no proporciona métodos prácticos para interacciones de segundo orden y órdenes más altos,

eFAST permite calcular coeficientes de órdenes más altos sin ejecuciones adicionales. Para hacerlo se asigna una frecuencia ω_i para x_i y una frecuencia común $\omega_{(i)} \neq \omega_i$ para todas las otras entradas. Así $V_{(i)}$ es estimado llenando con frecuencia $\omega_{(i)}$.

Esta mejora al método es utilizada para llevar a cabo un AS espacial de tipo global y obtiene los índices de primer orden y orden total en una sola ejecución (Crosetto et al., 2000; Crosetto y Tarantola, 2001).

Otra técnica para realizar análisis de sensibilidad espacial es, el método de Sobol (Sobol, 1993) y es de tipo global, está basado en varianza y es independiente del modelo. La ventaja de Sobol (y sus extensiones) es que este asegura que el espacio de entrada del modelo entero es explorado, y esta es una medida de sensibilidad que es independiente de hipótesis sobre la estructura del modelo (Saltelli et al. 2005). El método inicial calcula los índices de interacción de primer orden, posteriormente se modifica por Saltelli (2002) para obtener los índices de primer orden y orden total. El procedimiento está basado en muestreo usando muestras cuasi-Monte Carlo en vez de Monte Carlo.

Los índices de Sobol' o índices de sensibilidad basados en varianza se obtienen como sigue:

$$S_i = \frac{D_i(Y)}{Var(Y)}, S_{ij} = \frac{D_{ij}(Y)}{Var(Y)} \quad (31)$$

Los índices expresan la varianza compartida de Y que es debido a una entrada o combinación de entrada dado.

Por otra parte el cálculo de los índices totales o efectos totales se obtiene de la siguiente manera:

$$S_{T_i} = S_i + \sum_{i < j} S_{ij} + \sum_{j \neq i, k \neq i, j < k} S_{ijk} + \dots = \sum_{l \in \#i} S_l \quad (32)$$

Donde $#i$ son todos los subconjuntos de $\{1, \dots, d\}$ incluyendo i . En la práctica, cuando d es grande, solo los principales efectos y efectos totales son calculados, así se obtiene buena información de la sensibilidad del modelo.

Una de las técnicas usadas por las técnicas de análisis de sensibilidad estadísticas, es la simulación de Monte Carlo (MC). El objetivo de MC es perturbar uno o más conjuntos de datos de entrada, en algún nivel de error elegido, para generar diferentes realizaciones (versiones corruptas) del conjunto de entrada original, y tiene las ventajas de explorar el rango completo de variación para los factores de entrada y no requerir de asunciones sobre la estructura del modelo.

Otras técnicas que han sido utilizadas para modelos más grandes o altamente complejos son: Identificabilidad estructural que introducida como herramienta en modelos biológicos para evaluar la estructura interna de un sistema de medidas de entrada/salida (Cobelli y Romanin-Jacur, 1976), Métodos de ecuaciones adjuntas que ha demostrado su eficacia en campos de modelado climatológico, hidráulica termal de reactores y seguridad de reactores, la sensibilidad exacta es calculada como definida por derivadas parciales salida (Koda et al. 1979) y las Funciones de Green que son usadas para permitir resolver ecuaciones diferenciales para cada parámetro del modelo cada vez que los valores del parámetro del caso base son cambiados (Hall et al, 1982).

Algunas de las ventajas para los métodos antes mencionados son que OAT es muy simple de modelar, Morris puede trabajar con grandes modelos y es poco demandante computacionalmente, las técnicas de regresión son independientes del modelo, calculan índices de efecto total, calculan interacciones de ordenes más grandes, FAST e eFAST son independientes del modelo y solo eFAST calcula índices de efecto total, escaleras de caracol y Sobol son independientes del modelo, calculan índices de efecto total, calculan interacciones de ordenes más altos, pueden tener entradas espaciales que están auto correlacionadas e incluyen una estructura espacial (Lilburne & Tarantola, 2009).

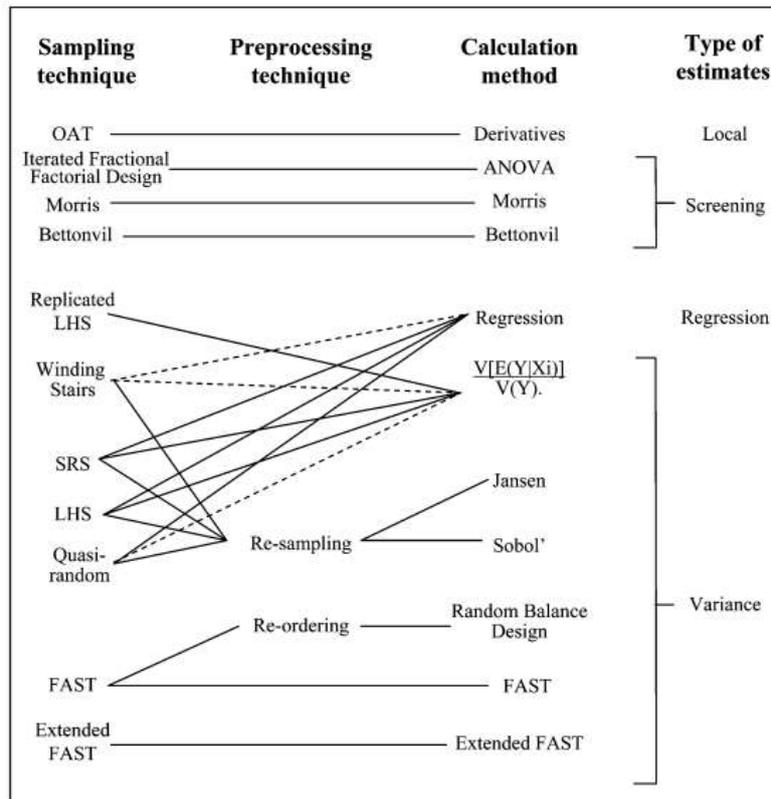


Figura 7.- Técnicas para realizar un análisis de sensibilidad espacial, unidas con línea continua se puede observar las combinaciones existentes y en línea punteada posibles técnicas que no existen todavía (Lilburne & Tarantola, 2009).

Para las técnicas de identificación estructural, ecuaciones adjuntas y funciones de Green se puede decir que a pesar de ser dirigidas a grandes modelos son computacionalmente muy demandantes y matemáticamente complejas de modelar. Sin embargo, estos procedimientos (estadísticos) necesitan valores escalares obtenidos a partir de técnicas de muestreo, representadas a través de PDFs, pierden importante información de su distribución espacial.

Es importante mencionar, que para utilizar métodos de AS en modelos geospaciales existen cuestiones conceptuales que incluyen: la dificultad de satisfacer suposiciones clave, la necesidad de representar una entrada espacial, la posibilidad de usar entradas auto-correlacionadas y analizar el efecto de la estructura espacial; por lo que su la aplicación frecuentemente ignorada y relativamente limitada en su alcance (Lilburne y Tarantola, 2009). También, los procedimientos de AS, tradicionales, se vuelven más difíciles de tratar cuando son consideradas variables espaciales o temporales (Plata et al., 2012).

2.4 Ventajas y desventajas del AS en modelos geoespaciales

Las principales ventajas y desventajas de los modelos de simulación, y técnicas de AS son para el caso de los AC que un aspecto muy importante que se tiene que resaltar en este tipo de modelos es que además de necesitar un proceso de definición de reglas de transición, más importante aún es el proceso de calibración para lograr producir resultados consistentes de predicciones futuras (Clarke y Gaydos, 1998; Batty y Xie, 1997) y lo cual hace que el tiempo aumente considerablemente ya que es necesario ejecutar el modelo tantas veces sea necesario para ajustar los componentes.

Por otra parte, la principal ventaja de los MBA es que pueden ser construidos en ausencia de conocimiento de interdependencias globales, es fácil de actualizar y que para la exploración de uso de suelo, la incorporación explícita de toma de decisión heterogénea tiene la facilidad en la forma de variar preferencias, reglas de decisión y percepción de riesgo (Ligmann-Zielinska y Sun, 2010; Borshchev y Filippov, 2004).

Sin embargo la gran desventaja de este tipo de modelos es el coste computacional que aumenta con el tamaño del entorno en el que el agente realizará los cálculos para elegir las acciones correspondientes (Devendran y Lakshmanan, 2014). Además un inconveniente para los estudios geoespaciales es que a pesar de obtener resultados esperados, esta técnica no toma en cuenta la característica de “parcelas”, por lo que cada píxel es independiente de los demás, además que los mapas de accesibilidad y aptitud necesarios para su aplicación no toman en cuenta todas las variables a la vez, asimismo los resultados son afectados por las diferentes configuraciones del modelo, respecto a tamaño y tipo de vecindad, así como por la resolución espacial (Gómez y Bosque 2004).

Con respecto a modelos basados en EMC, una desventaja es que al momento de plantearse el problema con una técnica multicriterio, los procedimientos en sistemas informáticos actuales, son rebasados en cuanto a su capacidad operativa (Gómez Delgado y Barredo Cano, 2005).

Por lo que se refiere a las diferentes técnicas de AS, la ventaja más importante de los métodos que trabajan con un factor a la vez es que son sencillos de realizar y modelar, sin embargo las desventajas son serias ya que van desde el hecho del esfuerzo que requieren para resolver las ecuaciones, debido a que puede ser muy intenso, la complejidad en la modelación del error, no pueden ser aplicables para campos o áreas lejos del caso base, tienen una gran exigencia computacional, hasta que si se usan rangos de variación pequeños para los parámetros de entrada no es posible evaluar la sensibilidad real de los parámetros (Gardner et al. 1981; Hoffman y Gardner, 1983; Helton, et al., 1985).

En cuanto a las técnicas basadas en muestreo aleatorio, las ventajas son que evalúa la influencia de otros parámetros de entrada, y se obtiene más información sobre la cantidad de variabilidad que añade a la variabilidad de la producción total en el modelo, reducen los efectos de los datos no lineales (Iman y Conover, 1979), también indican el grado de monotonidad entre los valores de la muestra de entrada y de salida (OIEA, 1989), sin embargo las desventajas son muchas iniciando con la difícil comprensión y análisis de resultados para problemas de grandes dimensiones, en algunos casos sigue siendo considerado como un método de sensibilidad local mejorado, para un mejor análisis se necesita calcular la linealidad de las entradas y salidas, sobredemanda computacional para lograr un análisis de sensibilidad profundo, es necesario operar sobre datos de rango transformado lo que conlleva a cálculos extras, problemas donde la fuerte correlación entre parámetros de entrada pueden influir en las correlaciones de entrada/salida, las ecuaciones de regresión podrían ser altamente complejas y también se puede mencionar los procedimientos para el cálculo pueden modificarse en otros más rigurosos y complejos, debido a la normalización de los datos (Crick, et al., 1987).

Finalmente, las ventajas de las técnicas que utilizan distribución de entrada segmentada son que estas comparan las distribuciones empíricas con una hipótesis nula, prueban medias y varianzas de las distribuciones empíricas, respectivamente, las pruebas operan en rangos de los datos en bruto, las ecuaciones para calcular las estadísticas de prueba se reducen a unas más simples. Las desventajas son que algunas de estas pruebas no parecen ser de

mucho valor para los rankings de sensibilidad de parámetros, en un análisis de incertidumbre la cantidad de variabilidad en la distribución de salida puede ser influenciada por los valores de entrada que dan lugar a cualquiera de los valores altos o bajos de salida (Conover, 1980).

Acerca de las técnicas de AS usadas para evaluar la sensibilidad en modelos geoespaciales se tiene que para el caso de OAT, este es muy simple de modelar, mientras que Morris puede trabajar con grandes modelos y es poco demandante computacionalmente, las técnicas de regresión son independientes del modelo, FAST y eFAST son independientes del modelo, obtienen índices de primer orden y calculan interacciones de ordenes más grandes, pero solo eFAST calcula índices de efecto total, escaleras de caracol y Sobol son independientes del modelo, calculan índices de efecto total, calculan interacciones de ordenes más altos, pueden tener entradas espaciales que estén auto correlacionadas e incluyen una estructura espacial (Lilburne y Tarantola, 2009).

Finalmente y no menos importante, un aspecto relevante encontrado en esta revisión es que se recomienda utilizar OAT extendido como punto de partida de cualquier AS en MBA, lo que contradice a Saltelli et al. (2008) y a Ligmann-Zielinska et al. (2014) quienes siguieron el uso de Sobol (Broeke et al 2016).

En Kocabas y Dragicevic (2004) para su estudio de intervalos temporales de 10 años usando un paso de tiempo de un año, realizaron un total de 432 mapas simulados, variando los tipos y tamaños de vecindad así como las resoluciones espaciales (50, 100, 150, 250 mts.), en (Ligmann, 2010) se analizó el modelo representado por 5 agentes desarrollados heterogéneamente, trabajando sobre 5041 celdas de usos de suelo, matriz de 71x71. Los agentes están basados sobre dos criterios de decisión espacial (valor de suelo y belleza escénica), están equipados con 3 atributos: preferencias (pesos) para valor de suelo, para belleza escénica y percepción de riesgo.

2.5 Análisis de sensibilidad explícitamente espacial

Una de las mayores preocupaciones existentes en cualquier modelo prospectivo es determinar el grado de confianza de los resultados obtenidos en las simulaciones. Para lo cual se han utilizado comúnmente métodos de análisis de sensibilidad (AS) de tipo estadístico a partir de indicadores cuantitativos, no espaciales. Idealmente, un AS espacial debe combinar un análisis completo global (en el que todas las entradas de interés son variadas simultáneamente y de manera eficiente), con la capacidad de generar realizaciones de auto correlación del mapa de entrada. La técnica de análisis de sensibilidad debe entonces trabajar en el mapa de entrada, en lugar de píxeles, o polígonos individuales, permitiendo así que la estructura espacial se incluya en el análisis. Este excluye regresión y técnicas rápidas, ya que requieren valores de entrada escalares (Conover, 1980).

De los pocos trabajos encontrados, y que aplican un análisis de sensibilidad explícitamente espacial, tenemos el de (Roura-Pascual, et al., 2010) donde se analiza un modelo de Toma de Decisión Multi Criterio en el cual se usó la técnica de Proceso de Jerarquía Analítica (AHP) para determinar la importancia de los factores en diferentes niveles del modelo, también son analizados tres enfoques de variación para el AS (unidimensional, multidimensional y por selección) y a su vez se evalúa el efecto de los cambios en la estructura espacial de los mapas de prioridad resultantes en tres diferentes regiones, obtenidos los mapas resultantes, estos fueron comprados por medio de dos índices: índice de disimilitud (Earth Movers Distance) e índice de información (Diversidad de Shannon).

En Plata, et al. (2013) se realiza un estudio de AS explícitamente espacial en donde se logra dar certidumbre a un modelo de simulación de crecimiento urbano de la comunidad de Madrid, utiliza la técnica OAT para la variación de los pesos de entrada, así como la sumatoria lineal ponderada (SLP) para el cálculo de los mapas de aptitud del modelo, generar las simulaciones prospectivas con la técnica Multi-Objective Land Allocation (MOLA) y al final se utiliza el método de píxeles más repetidos para mostrar frecuencia de los píxeles seleccionados por cada simulación. En los resultados se pueden observar

visualmente las áreas más estables para el crecimiento urbano de acuerdo a los diferentes escenarios propuestos: negocios, crisis e innovación y sustentabilidad, y en los escenarios podemos observar las categorías de usos de suelo urbano, comercial e industrial.

En el caso de Feizizadeh, et al. (2014), estos asocian los SIG y los Modelos de Toma de Decisión Multicriterio y a los cuales de manera pertinente, se le realiza el AS espacial en una metodología, propuesta por el autor, de tres etapas, donde la segunda corresponde al análisis de sensibilidad espacial y este es logrado realizando un ASG donde los pesos son expresados a través de funciones de distribución de probabilidad (PDF), para lograr el objetivo de la investigación además se realizaron pruebas con la combinación de los métodos de promedio de pesos ordenados (OWA) con ASG y AHP con ASG. La evaluación de los resultados fue mediante el cálculo de las características de operación relativa (ROC).

Por otro lado Xu & Zhang (2013) implementan el AS espacial para validar y calibrar un MCDM, esto por medio de una métrica tradicional como la media de tasa de cambio absoluto y la técnica novedosa de Earth Movers Distance (EMD), para llevar a cabo los mapas de aptitud usaron el método AHP aplicando una variación unidimensional para lograr mejores resultados.

Finalmente Ligman-Zielinska (2014) realiza uno de los estudios más recientes sobre análisis de sensibilidad explícitamente espacial utilizando simulaciones de Monte Carlo (MC) para barrer el espacio de los pesos (expresados como PDF) y generando múltiples mapas de aptitud yw obteniendo así los mapas de aptitud promedio, mapas de desviación estándar de la incertidumbre y varios mapas de sensibilidad. Para este análisis fue necesario el uso de las técnicas de OAT y ASG, calculando a partir de estos los índices de 1er orden y orden total (con técnicas basadas en varianza). Estos índices se calculan para cada peso de entrada.

Los trabajos que utilizan el método de EMD, logran realizar un proceso que trabaja con la característica espacial de los mapas ya que este analiza la

cantidad de esfuerzo mínimo que se debe de realizar para transformar una distribución en otra moviendo “la distribución de la masa” (Rubner, et al., 1998),

Este procedimiento sin embargo proporciona solo un valor de la disimilitud entre las distribuciones espacial de los mapas a comparar. Asimismo es imprescindible el uso de otra técnica (por ejemplo la diversidad o índice de Shannon) para poder obtener información sobre la variación observada, es decir, este índice no identifica las variables presentes; si no que simplemente distingue unas de otras para observar el comportamiento de los valores de cada una de ellas y del total. El valor del índice de EMD se obtiene a partir de dos histogramas P y Q (Rubner et al., 2000):

$$EMD(P, Q) = \left(\min_{\{f_{ij}\}} \sum_{i,j} f_{ij} d_{ij} \right) / \left(\sum_{i,j} f_{ij} \right) \quad (33)$$

Y el cálculo del índice de Shannon (Shannon y Weaver, 1949) se realiza por:

$$H' = - \sum_{i=1}^S p_i \log_2 p_i \quad (34)$$

Donde S es el número de especies, p_i es la proporción de individuos de la especie i respecto al total de individuos, n_i son los individuos de la especie i y N son todos los individuos de todas las especies.

Si bien trabajan de manera espacial, los propios autores concluyen que son necesarios más trabajos para analizar un mejor estudio acerca del comportamiento y de la estabilidad de estas técnicas de análisis espacial.

Ahora bien, a pesar de lograr una metodología de análisis estrictamente espacial (Plata, et al., 2013), el análisis de sensibilidad considera solo 16 simulaciones, siendo un análisis limitado, por lo que se deben analizar un mayor número de simulaciones para mejorar el AS y aumentar la robustez y confianza en los resultados.

En conclusión, el cálculo de estas técnicas de análisis de sensibilidad espacial en un gran número de los trabajos de investigación, no es realizado directamente

en las herramientas SIG utilizada por los expertos, este se lleva a cabo en un software externo, mismos que al ser estadísticos pierden la característica puramente espacial; más importante aún los resultados finales no son generados en su forma espacial por lo que no se pueden observar directamente en el mapa, para ello debe de realizarse un proceso posterior de análisis y transformación a datos espaciales.

Debido a que pocos trabajos son de índole explícitamente espacial es el interés del estudio de esta área tan poco aprovechada en la rama de la geomática, y con el firme antecedente de que las técnicas hasta el momento no permite determinar en SIG la ubicación a nivel espacial y/o territorial donde el modelo funciona mejor o peor.

Basados en técnicas estadísticas de AS se han desarrollado aplicaciones en diferentes plataformas, para realizar este paso tan importante en el proceso de modelado. En el estado del arte se pueden encontrar herramientas como SensIt, Simlab e IntiGIS, por mencionar algunas, las cuales llevan a cabo un análisis estadístico.

2.6 Ventajas y desventajas de las técnicas de análisis de sensibilidad

La ventaja más importante de los métodos que trabajan con un factor a la vez es que son sencillos de realizar y modelar, sin embargo las desventajas son serias ya que van desde el hecho que el esfuerzo requerido para resolver las ecuaciones puede ser muy intenso (Helton et al., 1985), la complejidad en la modelación del error Gardner et al. (1981), la dependencia de la magnitud de la sensibilidad de la variable al escenario de caso base, el análisis de sensibilidad de tipo *“local”*, la gran exigencia computacional, el uso de rangos de variación pequeños para los parámetros de entrada donde no se evalúa la sensibilidad real de los parámetros Hoffman y Gardner (1983), el cálculo de varios factores por medio de diferentes técnicas estadísticas, para obtener la sensibilidad, hasta la subjetividad la cual ha sido fuertemente criticada por expertos dada la dependencia al criterio de investigador que realiza el trabajo.

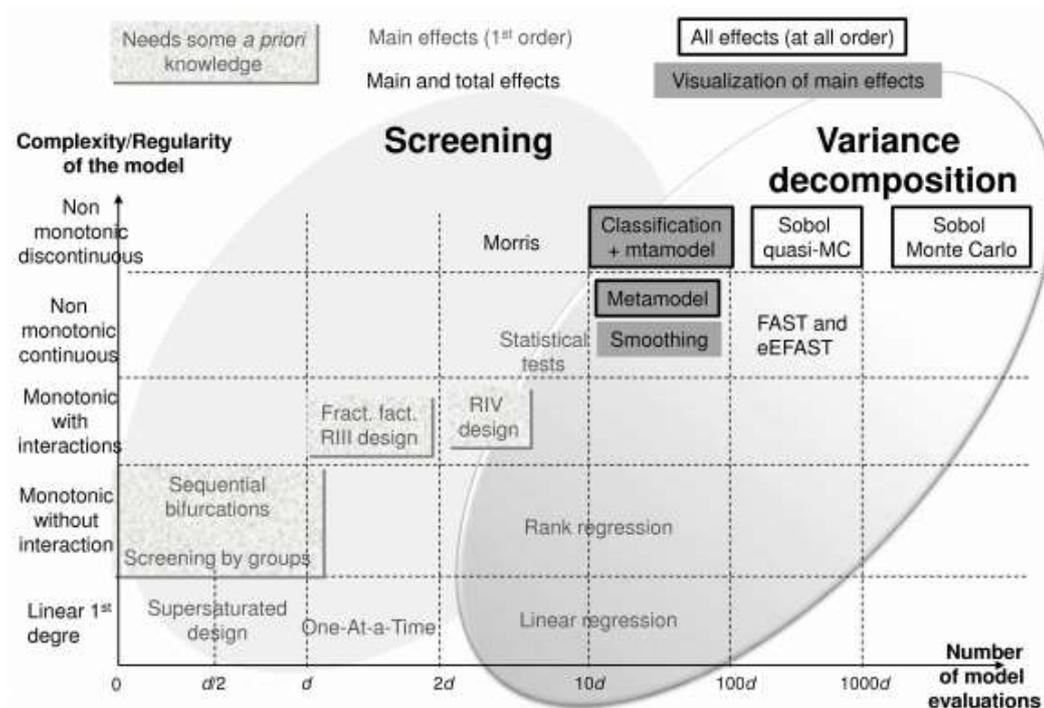


Figura 8.- Clasificación de métodos de AS global (Iooss & Lemaitre, 2015).

Algunas de las ventajas de las técnicas que utilizan muestreo aleatorio son que estos estudian además del grado de variabilidad de los factores, la influencia de otros parámetros de entrada, considerando esta como significativa e importante para los análisis de incertidumbre y sensibilidad que describen el rendimiento del modelo, estas técnicas trabajan con una muestra mucho más grande que los métodos locales, además se obtiene más información sobre la cantidad de variabilidad que añade a la variabilidad de la producción total en el modelo, también reducen los efectos de los datos no lineales (Iman y Conover, 1979), indican el grado de monotonicidad entre los valores de la muestra de entrada y de salida (OIEA, 1989).

Sin embargo las desventajas son muchas iniciando con la difícil comprensión y análisis de resultados para problemas de grandes dimensiones, para uno de los casos sigue siendo considerado como un método de sensibilidad local mejorado, este necesario para un mejor análisis la linealidad de las entradas y salidas, ya que de lo contrario existe la posibilidad de dar como resultados correlaciones

aparentes de entrada/salida (Hoffman y Gardner, 1983; Crick et al, 1987; IAEA, 1989).

En estos también existe la sobredemanda computacional para lograr un análisis de sensibilidad profundo, estas técnicas necesitan operar sobre datos de rango transformado lo que conlleva a cálculos extras, en problemas donde la fuerte correlación entre parámetros de entrada pueden influir en las correlaciones de entrada/salida, las ecuaciones de regresión podrían ser altamente complejas y también se puede mencionar los procedimientos para el cálculo pueden modificarse en otros más rigurosos y complejos, debido a la normalización de los datos.

Para el caso de las técnicas que utilizan distribución de entrada segmentada las ventajas son que estas comparan las distribuciones empíricas con una hipótesis nula, prueban medias y varianzas de las distribuciones empíricas, respectivamente, las pruebas operan en rangos de los datos en bruto, las ecuaciones para calcular las estadísticas de prueba se reducen a unas más simples (Conover, 1980).

Por otra parte las desventajas son que algunas de estas pruebas no parece ser de mucho valor para los rankings de sensibilidad de parámetros, y en un análisis de incertidumbre la cantidad de variabilidad en la distribución de salida puede ser influenciada por los valores de entrada que dan lugar a cualquiera de los valores altos o bajos de salida.

2.7 Análisis Computacional e Implementaciones

En el ámbito de software libre y comercial existen diferentes desarrollos que realiza el proceso de análisis de sensibilidad, se encontraron alrededor de 15 aplicaciones desarrolladas en diferentes plataformas y/o lenguajes de programación entre ellas se encontró una generalización de dos grupos, de acuerdo a la plataforma en la que se desarrolló o tipo de estudio que llevan a cabo, las que aplican AS y las que hacen AS espacial; a su vez los clasificadas como AS se dividieron en aquellas que fueron desarrolladas para trabajar con

hojas de cálculo y las que fueron desarrolladas en cualquier otro lenguaje de programación.

Los investigadores que realizan análisis espacial o simulación geoespacial, regularmente hacen uso de las herramientas SimLab, Matlab o R, para generar el cálculo de índices de sensibilidad, y posteriormente estos resultados son analizados para luego transfórmalos en datos explícitamente espaciales y finalmente, se pueden observar de manera visual en un mapa.

Ninguna de las herramientas analizadas está desarrollada para trabajar directamente con modelos de simulación geoespacial, además ninguna de ellas realiza análisis de sensibilidad explícitamente espacial. Como se ha visto, existe diferentes técnicas para llevar a cabo un análisis de sensibilidad espacial, sin embargo no se incorporan a los modelos de simulación, además es de relevancia mencionar que todos las herramientas analizadas no trabajan directamente con la distribución espacial contenida en el mapa, por lo que se tiene certeza de que en este paso se pierde la característica puramente espacial de la información.

Las aplicaciones encontradas realizando una búsqueda exhaustiva en internet fueron: SensIt, TopRank, Simlab, SALib, SaSAT, SensSB, GUI-HDMR, The-Eikos, CMOST, DEFINITE, R, AQUASIM, IntiGIS e iUSA. Se describen a continuación:

La aplicación SensIt fue elaborada por TreePlan Software es una aplicación creada sobre una hoja de cálculo y está basada en el modelo *What-if*, y puede crear ploteos simples, graficas de telaraña y graficas de tornado, esta herramienta realiza un análisis de sensibilidad de un factor y dos factores, variando automáticamente los valores de entrada.

TopRank creado por Palisade Corporation, es una herramienta diseñada para usar en hoja de cálculo y la cual identifica los factores críticos de los modelos, este toma tu modelo en hoja de cálculo y selecciona una o varias celdas como salida y automáticamente encuentra las celdas que afectan los resultados, el análisis se basa en el método *What-if*, esta aplicación varia diferentes valores

para cada celda variable cambiando los valores sobre rangos de -10% y +10%. Para observar los resultados, estos pueden ser desplegados en gráficos de tornado, telaraña y sensibilidad donde este último plotea los efectos de una entrada individual sobre los resultados

Una de las herramientas más amplia y frecuentemente utilizadas por expertos en el área del análisis de sensibilidad para modelos espaciales es SimLab. Este software fue desarrollado por The European Commissions in-house science service, y es una aplicación basada en Monte Carlo para la realización de análisis de sensibilidad e incertidumbre, este software se compone de tres módulos que son el Pre-Procesador Estadístico (selección de rango y distribución para cada variable de entrada y generación de muestras de puntos para la distribución de las entradas), Ejecución del Modelo (se alimenta el modelo con las muestras y el modelo de salida es generado) y el Post-Procesador Estadístico (los resultados son usados para el análisis de sensibilidad e incertidumbre).

Las funciones de probabilidad de distribución que se pueden usar son: normal y muestreo aleatorio. Por otra parte las técnicas de análisis de sensibilidad incorporadas a SIMLAB son: FAST, EFAST, LHC, LpTau, Morris, Aleatorio, Diseño Aleatorio Balanceado, Sobol'. La licencia es de uso no comercial.

Por su parte SALib (Sensitivity Analysis Library) fue desarrollada por Jon Herman, Will Usher y otros, es una librería con implementaciones en Python de métodos comúnmente usados para realizar análisis de sensibilidad. Útiles para calcular los efectos de las entradas de los modelos o factores exógenos de interés, es posible definir un porcentaje para la prueba de cobertura, los métodos que utiliza incluyen Sobol', Morris, FAST, Factorial Fraccional entre otros. La licencia esta liberada bajo la licencia del MIT y es de tipo libre. Dado que es una librería no se realizó ninguna prueba.

Desarrollado por (Hoare et al., 2008), SaSAT es un paquete de software amistoso para aplicar análisis de sensibilidad e incertidumbre a modelos matemáticos y computacionales de diferente complejidad y contexto. Es una toolbox construida en Matlab y puede trabajar con hojas de cálculo, pero no con

toda su funcionalidad. Es capaz de realizar muestreo eficiente (por las técnicas de muestreo aleatorio, factorial completo e hipercubo latino), cálculo de coeficientes de correlación (Pearson, Spearman y orden parcial), análisis de regresión, priorización de factores y graficación de resultados (incluye graficas de tornado, ploteo de dispersión, y superficie de respuesta). La licencia de esta herramienta es libre (Hoare, et al., 2008)

El desarrollo de una toolbox que trabaja con Matlab e integra diferentes métodos de analisis de sensibilidad local y global que pueden ser aplicados a modelos biológicos descritos por ecuaciones diferenciales ordinarias o ecuaciones algebraicas diferenciales fue desarrollado por el Process Engineering Group IIM-CSIC, esta aplicación se llama SensSB. Las diferentes técnicas de sensibilidad local que contiene son basadas en derivadas parciales, aproximación por diferencias finitas, métodos directos (ODESSA) y métodos de funciones de Green; las medidas de sensibilidad global basadas en varianza que contempla son derivadas parciales, promedio, varianza y una combinación de varianza y promedio, por otra parte, también es posible el cálculo de índices de Sobol'. Por último utiliza diferentes maneras de obtener el orden de parámetros tales como raíz cuadrada del cuadrado de la sensibilidad, media de valores absolutos, media cuadrada, sensibilidad máxima y sensibilidad mínima (Rodriguez-Fernandez & Banga, 2009). Esta aplicación tiene licencia libre basada en Creative Commons, Spain License, esto es sin costo para usuarios académicos (enviando un correo con servidor institucional).

Otra aplicación que también está desarrollada en Matlab, por Tilo Ziehn de la Universidad de Leeds es GUI-HDMR, esta herramienta provee un enfoque para explorar el mapeo de entrada/salida de un modelo sin requerir un gran número de ejecuciones del modelo, los métodos usados para la reducción de varianza son las técnicas de correlación y el método de control de radio variado ambos utilizando la integración con Monte Carlo. Utiliza un método optimizado para automáticamente elegir el mejor orden polinomial para cada función componente. Posteriormente se podrían usar métodos compatibles con la herramienta como Morris o Sobol' para el análisis de sensibilidad, sin embargo el cálculo de varianzas parciales es un enfoque as eficiente de usar con las

funciones de expansión usadas por el software (Ziehn & Tomlin, 2009). La licencia de esta aplicación es libre sin embargo, esta herramienta se liberó en el 2008 y no existiendo hasta el momento actualización de la misma, por tal motivo no se realizaron los experimentos para analizar su rendimiento.

Por otra parte The-Eikos fue desarrollado por Facilia AB, es una toolbox para MATLAB que contempla métodos de análisis de sensibilidad global y local (OAT, Morris, eFAST, Sobol', sensibilidad local, entre otros), se puede inducir la correlación de parámetros (correlación de rango ordenado), es posible utilizar métodos de muestreo (Monte Carlo, Hipercubo Latino y Quasi-random), infortunadamente este software no tiene licencia de prueba por lo que no fue posible analizarlo.

Una herramienta que contiene un apartado para el análisis de sensibilidad el cual es realizado de manera automática, generando experimentos para conocer la respuesta de los conjuntos de criterios es la desarrollada por el Computer Modelling Group LTD, CMOST. CMOST puede crear rápidamente ploteo cruzado para investigar la relación entre dos parámetros funciones objetivo. Si se desea realizar un análisis de sensibilidad basado en varianza tiene disponible los métodos de Morris y Sobol', esta aplicación al igual que la anterior no tiene licencia de prueba por lo que su estudio tampoco fue posible.

Para el software DEFINITE, desarrollado por el Institute for Environmental Studies vrije Universiteit amsterdam 2006, el paso inicial es definir el modelo en una serie de alternativas y efectos (que contemplan información acerca del tipo de efecto, la escala de la medida, el costo/beneficio y la unidad), posteriormente es necesario realizar el análisis multicriterio por medio de la elección del tipo de estandarización de los efectos (máximo, intervalo, meta, convexo, cóncavo, forma-S y libre de forma), evaluación del efecto de los pesos utilizando alguno de los métodos ya contemplados (directo, comparación por par, valor exponencial, aleatorio, vector y desde un vector dado por el usuario), obteniendo finalmente un orden, después se tiene que agregar la información del análisis de costo/beneficio iniciando por el descuento, eligiendo un valor y obteniendo los índices esperados, el paso final es el análisis de sensibilidad con el que se

pueden realizar diferentes análisis como graficas de comparación, análisis de incertidumbre y análisis de sensibilidad, sin embargo dada la complicación de la sobre especificación de la información para el análisis, no fue posible realizar las pruebas correspondientes de esta herramienta, además de solo poder conseguir una licencia de prueba que limitaba la realización del análisis correspondiente.

Un sistema también globalmente conocido para realizar análisis estadístico y gráfico, y creado por Ross Ihaka y Robert Gentleman es R. Se distribuye gratuitamente bajo los términos de la GNU. Dentro de la paquetería del sitio de Comprehensive R Archive Network (CRAN) se puede encontrar el paquete de "Sensibilidad" el cual fue liberado en 2015 teniendo como autor principal a Gilles Pujol. Este paquete es una colección de funciones para factores de Screening, análisis de sensibilidad global y análisis de sensibilidad de confianza de la salida de modelos. Dentro de los métodos globales desarrollados en esta librería se tienen coeficientes de regresión lineal, bifurcación secuencial Bettonvils, Morris (con OAT), diferentes índices de Sobol' (usando estimación de Monte Carlo, índices de primer orden, etc.), índices de Borgonovo e índices basados en leyes. Debido a que es una librería solo funcional en R no se realizó ninguna prueba.

Clasificados como software que realiza análisis de sensibilidad espacial, se encuentran AQUASIM, una herramienta para simulación y análisis de datos de sistemas acuáticos, esta utiliza dos funciones para el cálculo de la sensibilidad y son: la función de sensibilidad relativa-absoluta y la función de sensibilidad relativa-relativa ya que sus unidades no dependen de la unidad del parámetro. Esta herramienta es de distribución libre (Reichert, 1993).

IntiGIS desarrollado por Irene Pinedo Pascua y es un plug-in para ArcGIS 9.3 en adelante, esta aplicación está basada en SolarGIS I proyecto que supuso la confirmación del gran valor que la perspectiva territorial podía aportar a los proyectos de energías renovables y el enorme potencial de los SIG para integrar energías renovables en proyectos de electrificación rural. Posterior a SOLARGIS surge SolarGIS II en el cual fueron incorporados mecanismos para determinar el grado de certidumbre de los resultados proporcionados, finalmente el desarrollo de IntiGIS fortalece esta herramienta integrando el análisis de sensibilidad

espacial. El método contemplado es el OAT (Pinedo, 2012). Este software, a pesar de aparecer visualmente en la página de distribución, no está disponible su descarga.

Para finalizar tenemos a iUSA creado por Arika Ligmann-Zielinska y diseñado para el lenguaje Python, es un paquete “standalone” para calcular los índices de sensibilidad de primer orden y de efecto total para evaluaciones multicriterio espacial basado en raster. El análisis de sensibilidad espacio-temporal que realiza está basado en la descomposición de varianza, iniciando esta con el cálculo de la varianza de un escalar o de cada localización geográfica (temporal) en el área de estudio y la cual es descompuesta en proporción para las entradas. Los resultados son llamados mapas de dominancia de entrada. Utiliza simulaciones Monte Carlo, así como la técnica de OAT y obtiene los índices de sensibilidad de primer orden y orden total por cada peso (Ligmann-Zielinska & Jankowski, 2014). Debido al lenguaje en el que fue desarrollado no fue por el momento analizado.

Una vez elegidas las herramientas se realizó un análisis de rendimiento, el cual se detalla más adelante, y que fue necesario considerar para esta investigación, información tal como la plataforma de desarrollo, el sistema operativo al que van dirigidas, las consideraciones sobre la posibilidad de ajustar parámetros del modelo (tipos de distribuciones de probabilidad, número de iteraciones, etc.) y finalmente el tiempo promedio para la ejecución del modelo (variando número de entradas, salidas e iteraciones).

En la siguiente tabla podemos observar de manera clara algunas de las características generales de cada herramienta:

Tabla 2.- Tabla comparativa de software de análisis de sensibilidad.

	Lenguaje de Desarrollo	S.O.	Licencia	Max. de Iteraciones
SensIt	Excel	Windows	Comercial	7422
TopRank	Excel	Windows	Comercial	-
SimLab	R	MacOs, Windows, Linux, UNIX	Libre	> 10000
SaSAT	Matlab	Windows	Libre	> 10000
SWAP_CUP	Matlab	Windows	Comercial	-
Espaciales				
AQUASIM	C#	Windows	Libre	-
IntiGIS	ArcGIS	Windows	Libre	-
iUSA	Python	Windows	Libre	-

3.- MARCO TEÓRICO

3.1 Análisis de sensibilidad

Muchos autores cuando se refieren al grado al cual los parámetros de entrada afectan al modelo de salida usan los términos “*sensible*”, “*importante*”, “*más influyente*”, “*mayor contribuyente*”, “*efectivo*” o “*correlacionado*” indiferentemente (Krieger et al., 1977; Downing et al. 1985). Sin embargo Crick et al. (1987) hace una distinción refiriéndose a parámetros “*sensibles*” a aquellos los cuales tienen una influencia significativa sobre la evaluación de resultados. El consenso entre autores es que el modelo es de hecho, sensible a los parámetros de entrada en dos distintas maneras: a la variabilidad asociada a un parámetro de entrada sensible que es propagada a través del modelo, resultando en una gran contribución al total de variabilidad de la salida y a los resultados del modelo que pueden estar altamente correlacionados con un parámetro de entrada, entonces esos cambios pequeños en el valor de entrada resultan en cambios significantes en la salida (Hamby, 1994).

Entonces al aplicar un **Análisis de Sensibilidad (AS)** se evalúa la variabilidad de los resultados del modelo, introduciendo pequeños cambios en los parámetros de entrada (Macgill, 1989), también sirve para guiar investigaciones futuras (Iman y Helton, 1988). Entre los objetivos principales del AS se encuentran: verificación, entendimiento y simplificación del modelo, la priorización de factores y el estudio de las consecuencias que resultan al cambiar un parámetro dado. De esta manera se puede proporcionar cierto grado de confiabilidad en un modelo e indicar las áreas prioritarias para refinar en el mismo (Iman y Helton, 1988; Hamby, 1995; Qureshi et al., 1999; Lilburne y Tarantola, 2009; looss y Lemaître, 2014).

El análisis de sensibilidad se lleva a cabo por varias razones. La primera es debido a la naturaleza del proceso de EMC, que contiene inherentemente diferentes niveles de incertidumbre debido a la elección cualitativa y subjetiva de diversos parámetros. La EMC ha sido criticada por ser un procedimiento “inexacto”, por estos motivos es muy importante poner a prueba la solidez de los resultados. Una segunda y muy importante razón es que el análisis de sensibilidad del procedimiento de EMC permite que los datos y el problema de

toma de decisiones sean explorados a mayores profundidades. Esto proporciona un mayor conocimiento sobre la naturaleza del problema de decisión, aclarar su complejidad e incluso puede proporcionar recomendaciones para futuros análisis.

Se puede aplicar un análisis de sensibilidad, a los resultados, con el fin de tener en cuenta la variabilidad de alguno de los componentes involucrados, también puede considerar los efectos de las diferentes técnicas utilizadas en el procedimiento de ponderación. Es muy importante que la sensibilidad de los resultados se pueda probar para diferentes valores de los criterios y los impactos más importantes.

Existen trabajos previos sobre revisiones de técnicas de AS, los cuales explican la forma de aplicar un proceso de análisis de sensibilidad, además muestran que los resultados obtenidos de las distintas técnicas, son similares en su valor de sensibilidad (Hamby 1994, 1995). Por otra parte, se revela que la aplicación de un AS es más frecuente en temas de ubicación de lugares nocivos y además se muestra que el análisis más utilizado está basado en la variación de los pesos de factores (Gómez y Bosque, 2004)

Regularmente el AS es estadístico y consiste en calcular la varianza y el valor esperado de cada salida del modelo y ordenar las contribuciones a la varianza usando un enfoque basado en matriz para resolver las derivadas parciales desde notación matricial eficiente. Las técnicas más fundamentales para realizar el análisis de sensibilidad utilizan ecuaciones diferenciales parciales, mientras que el enfoque más simple requiere de variar los valores de los parámetros uno a la vez (Hamby, 1995).

3.1.1 Técnicas de análisis de sensibilidad

Las diferentes técnicas de AS se pueden ser clasificadas de tipo local y global, dependiendo el espacio de salida que analizan. El análisis global se utiliza para modelos complejos con salidas multidimensionales, mientras que el tipo local solamente trabaja con salidas reducidas y regularmente no contemplan todas las dimensiones del modelo.

De manera general los pasos para llevar a cabo un análisis global son: especificar la función objetivo del estudio, seleccionar las entradas de interés, asignar un rango y una distribución estadística para las entradas seleccionadas (deducido de los datos, consideración de límites físicos, opinión de expertos o estimación de parámetros en problemas inversos), aplicar un diseño de muestreo para generar una muestra de tamaño “ N ” de la distribución de la entrada (Monte Carlo, LHS, CuasiAleatorio, etc.), evaluar el modelo para cada conjunto de muestra de los valores de entrada, obteniendo “ N ” valores para la función objetivo y usar los resultados del paso anterior para analizar la incertidumbre y un estimador de sensibilidad para obtener la importancia relativa de las entradas (medidas locales, basadas en varianza, regresión, y screening) (Arika-Ligmann et al., 2004).

Por otro lado, estos métodos se pueden clasificar también en tres grupos: los que operan con un factor a la vez, aquellos que se basan en la generación de una matriz de entrada y un vector de salida asociada y aquellas que requiere una partición de un vector de entrada particular basándose en el vector de salida resultante (Hamby, 1995).

Entre las técnicas que operan con un factor a la vez se encuentran el *coeficiente de sensibilidad*, que es básicamente el radio de cambio en la salida al cambio en la entrada mientras todos los otros parámetros permanecen. El modelo que se obtiene mientras todos los parámetros se mantienen constantes es definido como “caso base” X (Kfieger et al., 1977).

A partir de esta estrategia de análisis se pueden encontrar a su vez el *Análisis Diferencial* que es referido como un método directo, mismo que resuelve derivadas parciales simples para un análisis de sensibilidad espacial o temporal (Morisawa y Inoue, 1974; Gardner et al., 1981; Iman y Helton et al., 1985); y el cual trabaja por medio de una aproximación de series de Taylor de primer orden para la variable dependiente, para su cálculo se puede aplicar la formula general de propagación de error (Helton et al., 1985), usada para obtener la varianza de Y , $V(Y)$:

$$V(Y) = \sum_{i=1}^n \left(\frac{\partial Y}{\partial X_i} \right)^2 V(X_i) \quad (14)$$

Donde la varianza de Y es utilizada como una medida de incertidumbre en las predicciones del modelo mientras que la varianza en X_i , es ponderada por la parcial de primer orden de Y con respecto de X_i , proporcionando una medida de sensibilidad del modelo a X_i .

Este método es una linearización de la teoría y es válido solo para pequeñas incertidumbres en los parámetros (Koda et al., 1979). Para ecuaciones de conjuntos grandes, la derivada parcial puede ser aproximada por diferencias finitas y los valores de salida pueden ser calculados para pequeños cambios en las entradas (Downing et al., 1985).

El método más simple para un análisis de sensibilidad es el *Análisis One-At-a-Time (OAT)*, ya que solo es necesario repetir el modelo y variar un factor a la vez mientras todos los demás se mantienen fijos, un orden de sensibilidad puede ser obtenido rápidamente incrementando cada parámetro por un porcentaje dado regularmente de +/- 20% (debido a que a mayor variabilidad el dato original pierde su valor natural) y cuantificar el cambio en la salida del modelo. Un más poderoso análisis de sensibilidad OAT examina el cambio en la salida mientras que cada parámetro es individualmente incrementado por un factor de su desviación estándar. Esta medida toma en cuenta la variabilidad del parámetro y la influencia asociada a la salida del modelo (Hamby, 1995; Lilburne et al., 2006).

Llevar a cabo la técnica de *análisis factorial* implica la elección de un determinado número de muestras para cada parámetro y ejecutar el modelo para todas las combinaciones de las muestras (Box et al., 1978; Rose, 1983). Los resultados obtenidos de esta manera se utilizan entonces para estimar la sensibilidad de parámetros. En algunos casos, la misma información puede obtenerse a partir

de un número reducido de ensayos, utilizando un *diseño factorial fraccional* (Box et al., 1978).

Por otra parte, el *Índice de Sensibilidad* (IS) es un método que determina la sensibilidad de un parámetro al calcular la diferencia de porcentaje de salida variando un parámetro de entrada a partir de su valor mínimo a su valor máximo (Hoffman y Gardner, 1983; Bauer y Hamby, 1991). El IS se obtiene a partir de la siguiente formula:

$$IS = \frac{D_{max} - D_{min}}{D_{max}} \quad (15)$$

Donde D_{min} y D_{max} representan el mínimo y máximo valor de salida, respectivamente, resultando de la variación de la entrada sobre su rango entero.

El último en esta clasificación es un método que es bastante sencillo y sólo cualitativo, se trata del *Análisis Subjetivo*, y es debido a que se basa en las opiniones de investigadores experimentados para determinar, a priori, que los parámetros pueden ser descartados por falta de influencia en los resultados del modelo. Para grandes modelos, se puede utilizar como un primer corte para reducir el número de parámetros de entrada a un tamaño manejable (Hamby, 1994).

Los efectos de distribución son significativos porque la sensibilidad de los parámetros depende, no sólo de la variedad y la distribución de un parámetro de entrada individual, sino también en los otros parámetros a los que el modelo es sensible. Además, la sensibilidad de parámetros depende de las interacciones e influencias de todos los parámetros (Iman et al., 1981a).

Es por ello que existe también una clasificación para métodos de análisis de sensibilidad que operan con muestreo aleatorio. El muestreo aleatorio de los parámetros de entrada genera distribuciones de entrada y salida que son de utilidad en la evaluación del modelo y los parámetros de incertidumbre en un sentido "global" (McKay et al, 1979; Iman y Conover, 1980; Iman et al, 1981;

Helton y Iman, 1982; Iman y Conover, 1982; Iman y Shortencarier, 1984; Reed et al, 1984; Downing et al., 1985; Kim et al., 1988; Iman, 1987; Stein, 1987; Helton et al, 1991).

Gardner et al. (1980) se refieren al estudio de sensibilidad de parámetro o estudios de este tipo como "análisis de error de parámetro". Los gráficos de dispersión es una técnica donde los factores de entrada contra factores de salida son útiles para determinaciones rápidas (visualmente hablando) del grado de correlación y la linealidad de la relación entrada/salida (Helton et al, 1986; Crick et al, 1987; Iman y Helton, 1988; Helton et al., 1991; Helton et al, 1993).

También pueden revelar relaciones inesperadas entre las variables de entrada y salida que puede dar una idea de cómo puede llevarse a cabo otras investigaciones. A su vez, Hoffman y Gardner (1983), proponen el "Índice Importancia", I_i que es igual a la varianza del valor del parámetro $s_{X_i}^2$, dividido por la varianza de los valores dependientes s_Y^2 , es decir:

$$I_i = \frac{s_{X_i}^2}{s_Y^2}$$

donde s se refiere a la varianza de los datos básicos para modelos aditivos y a la varianza de los datos transformados logarítmicamente para los modelos multiplicativos.

Esta medida de importancia está basada en la contribución fraccional del parámetro a la variabilidad total, o incertidumbre. La importancia variable es estimada en (Cunningham et al., 1980), a través del uso de una combinación de la contribución fraccional a la variabilidad de salida y el cambio resultante en la salida dada cambio individual en la entrada.

Similar al índice de importancia es el Radio de Desviación Relativa (RDR), mismo que al tener un gran valor de esta "relación de desviación relativa" indica que, o bien la distribución de salida varía ampliamente, o que la distribución de entrada es relativamente estrecha. Un valor mayor, de RDR, que la unidad indica que la incertidumbre propagada a través del modelo está aumentando debido a la estructura del modelo y su alta sensibilidad a esa variable particular. Un valor de

uno indica que toda incertidumbre de entrada es pasada a través del modelo y aparece como incertidumbre de salida, mientras que un valor menor que la unidad indica que el modelo es menos sensible al parámetro, contribuyendo así a la incertidumbre de salida en un menor grado (Hoffman y Gardner 1983).

En cuanto a la sensibilidad local un método similar, es *el Método de Desviación Relativa*, pero con la excepción de que una muestra mucho más grande se hace de la distribución de entrada. La cifra de mérito es la sensibilidad de “desviación relativa”, la relación de la desviación estándar a la media de la función de densidad de salida y es similar al coeficiente de variación (desviación estándar x 100/media). Esta prueba proporciona una indicación de la contribución de cada parámetro a la variabilidad presente en la salida del modelo y, un grado, en la medida de la correlación que existe entre la variable de entrada y de salida (Hamby, 1993).

Gardner et al. (1981) por otra parte recomiendan el uso de *coeficientes de correlación simple (llamado R de Pearson)*, derivados de simulaciones de Monte Carlo, como una forma razonable para clasificar los parámetros del modelo en función de su contribución a la incertidumbre en las predicciones. El coeficiente de correlación de momento de producto de Pearson se denota por r y se define como:

$$r = \frac{\sum_{i=1}^n (X_{ij} - \bar{X}_i)(Y_j - \bar{Y})}{\left[\sum_{j=1}^n (X_{ij} - \bar{X}_i)^2 \sum_{j=1}^n (Y_j - \bar{Y})^2 \right]^{1/2}} \quad (16)$$

Donde X_{ij} y Y_i son los valores de las variables que se desean correlacionar y \bar{X} y \bar{Y} son el promedios de todos sus valores. Esto para la correlación entre X_i y Y (Conover, 1980). Cuanto mayor sea el valor absoluto de r más fuerte es el grado de relación lineal entre los valores de entrada y de salida (IAEA 1989). Un valor negativo de r indica que la salida está inversamente relacionada con la entrada.

Por otro lado, si se desean transformar los datos en rangos, se ha demostrado que funciona muy bien la técnica de *Transformación de Orden* siempre y cuando la variable dependiente sea una función monótona de las variables independientes (Iman y Conover, 1979). El orden de transformación lineariza relaciones no lineales entre las variables monótonas y reduce los efectos de los valores extremos (Helton y Iman, 1982). Esta transformación convierte la medida de sensibilidad de una linealidad a una monótona.

Continuando con el *coeficiente de correlación, o ρ de Spearman*, este puede calcularse utilizando la ecuación r de Pearson, con la excepción de operar sobre los datos de rango transformado (Iman y Conover, 1979). Si las asociaciones de entrada/salida son monótonas entonces la transformaciones de rangos de los valores de entrada y de salida (es decir, sustituyen a los valores con sus filas) dará lugar a una relación lineal y el coeficiente de correlación de rangos indicará el grado de monotonidad entre los valores de la muestra de entrada y de salida (OIEA, 1989).

Asimismo, los *coeficientes de correlación parcial (CCP)* se calculan para dar cuenta de las correlaciones entre otras variables de entrada (Gardner et al, 1980; Gardner et al, 1981; Iman et al, 1981; Iman y Conover, 1982; Otis, 1983; Downing et al, 1985; Iman y Helton, 1985; Breshears, 1987; Whicker y Kirchner, 1987; Iman y Helton, 1988; IAEA, 1989; Whicker et al., 1990; Iman y Helton, 1991; Helton et al, 1993).

Teniendo en cuenta variables aleatorias X_1 y X_2 como entrada y la variable de salida Y , un coeficiente de correlación parcial es una medida de la correlación entre X_1 y Y , por ejemplo, mientras que la eliminación de las correlaciones indirectos debido a relaciones que pueden existir entre X_1 y X_2 o X_2 y Y . El CCP es definido como (Conover, 1980):

$$r_{X_1Y|X_2} = \frac{r_{X_1Y} - r_{X_1X_2}r_{X_2Y}}{\sqrt{(1 - r_{X_1X_2}^2)(1 - r_{X_2Y}^2)}} \quad (17)$$

Sin las correlaciones existentes entre los parámetros de entrada, no hay necesidad para el cálculo de los parciales para determinar las clasificaciones de sensibilidad.

Este coeficiente es ampliamente utilizado para estudios de sensibilidad (Iman et al, 1981a, b; Crick et al, 1987; Iman y Helton, 1988; OIEA, 1989; Iman y Helton, 1991). Downing et al. (1985) y puede ser el más potente para la sensibilidad de un parámetro que es fuertemente monótono pero altamente no lineal.

A su vez, las técnicas de regresión permiten una clasificación que se determinará en base a la magnitud relativa del coeficiente de regresión sensibilidad. Este valor es indicativo de la cantidad de la influencia que parámetro tiene sobre todo el modelo. Debido a las unidades y las magnitudes relativas de parámetros, un proceso de normalización es a veces justificado, sin embargo la forma generalizada de una simple ecuación de regresión es:

$$\hat{Y} = b_0 + \sum_k b_k Z_k \quad (18)$$

Donde Z_k es una variable predictora y una función de (X_1, \dots, X_n) y cada b_k es un coeficiente de regresión (Helton et al., 1985).

Con respecto al uso de técnicas de matriz, se han utilizado en tales casos para calcular los coeficientes de regresión (Krieger et al., 1977). Una forma mejorada son las Técnicas de Regresión Estandarizada que calculan un *coeficiente de regresión por rango* (CRR), es decir, la normalización por la transformación por rango, es un procedimiento simple que requiere menos cálculos. Utiliza medias y desviaciones estándar de entrada y salida de los conjuntos de datos (el *coeficiente de regresión estandarizado*, CRE), sin embargo, es un poco más riguroso y es logrado por:

$$\frac{(\hat{Y} - \bar{Y})}{s} = \sum_k \left[\frac{b_k s_k}{s} \right] \frac{(Z_k - \bar{Z}_k)}{s_k} \quad (19)$$

Donde Z_k es una función de (X_1, \dots, X_n) , s es la desviación estándar de la salida y s_k es la desviación estándar de la entrada (Helton et al., 1985).

El efecto de la estandarización es eliminar la influencia de unidades y colocar todos los parámetros en un mismo nivel. Los análisis de regresión estandarizados son realizados por Iman y Helton (Helton et al., 1985; Iman y Helton, 1988, 1991).

En la última de las clasificaciones mencionadas, se encuentran aquellas pruebas de sensibilidad que involucran la distribución de entradas segmentadas. Estas pruebas estadísticas implican dividir o segmentar distribuciones de parámetros de entrada en dos o más distribuciones empíricas basadas en una compartimentación asociado de la salida (Crick et al., 1987).

Estas pruebas se utilizan para comparar las características de las distribuciones de entrada creadas por la segmentación (medias, medianas, desviaciones) y otras características de las muestras aleatorias independientes. Y estadísticamente se determina si las muestras proceden de la misma población. La división de la distribución de salida no puede ocurrir en cualquier valor o percentil, sino que debe basarse en la cuestión estadística que se plantea, si las distribuciones de entrada generadas por este proceso son estadísticamente idénticas entonces el modelo no es sensible a ese parámetro (Iooss & Lemaitre, 2015).

Sin embargo, si las distribuciones son diferentes entonces la distribución de salida está de hecho influenciada por la entrada y el valor absoluto de la estadística de prueba, se puede utilizar para realizar la clasificación de sensibilidad. Las pruebas paramétricas estándar no son razonables en conjuntos de datos de entrada generados por los métodos de muestreo aleatorio debido al limitado conocimiento de las variables de entrada y sus distribuciones asociadas (Iman et al., 1981). Las pruebas estadísticas no paramétricas, por lo tanto, se utilizan cuando los datos se consideran como distribución libre (Conover, 1980).

Dentro de esta clasificación se encuentra la Prueba de Smirnov puede medir directamente la mayor distancia vertical entre las dos funciones de distribución, las cuales pueden ser representadas en el mismo gráfico, o la estadística de prueba se puede calcular utilizando:

$$T_1 = \sup |S_1(x) - S_2(x)| \quad (20)$$

donde "sup" es la abreviatura de supremo y la ecuación representa la mayor diferencia absoluta entre $S_1(x)$ y $S_2(x)$, donde estas son distribuciones empíricas generada como resultado de dividir los valores de los parámetros de entrada (Conover, 1980).

El grado de similitud entre las distribuciones, medido por la estadística de prueba, se utiliza para indicar el grado de sensibilidad entre los valores de entrada y de salida.

Otra prueba muy similar es la Prueba de Cramer-Von Mises, y se encarga de determinar si dos distribuciones empíricas son estadísticamente idénticas pero, el cálculo de la estadística es un poco más complicado sin embargo, hay poca diferencia en el poder de la prueba en comparación con la estadística de Smirnov (Conover, 1980). La estadística de Cramer-von Mises, T_2 , es la suma de todas las distancias verticales al cuadrado entre las dos distribuciones empíricas:

$$T_2 = \frac{mn}{(M+n)^2} \sum [S_1(x) - S_2(x)]^2 \quad (21)$$

donde los valores de n y m son el número de muestras utilizadas para estimar las distribuciones.

Se espera que las clasificaciones de los parámetros sobre la base de las pruebas Smirnov y Cramer-von Mises seán muy similares, las dos pruebas muestran poca diferencia en su poder estadístico. Un gran valor Smirnov o Cramer-von Mises es indicativo de una mayor diferencia en las dos distribuciones empíricas generadas por la división de los datos de entrada basados en algunos criterios de salida. Esta diferencia indica una mayor correlación entre las variables independientes y dependientes (Conover, 1980).

Por otra parte la prueba Mann-Whitney, es utilizada para comparar las medias de dos muestras independientes. Dos funciones de distribución, X e Y , se ordenan como una sola muestra y filas se asignan sobre la base de los lazos de pedidos se supone que no existe. La estadística de prueba, T , es la suma de las filas resultantes de los datos de distribución de X :

$$T = \sum_{i=1}^n R(X_i) \quad (22)$$

donde $R(X_i)$ se refiere al rango de X_i (Conover, 1980).

Considerando ahora la varianza, dos varianzas de muestras independientes, X_i e Y , se pueden comparar aplicando la prueba de ordenes cuadrados. Los rangos no se basan en los datos en bruto, sino en la diferencia absoluta entre la muestra aleatoria (por ejemplo, X_i) y la media de la muestra (por ejemplo, \bar{X}). No se supone que no tienen vínculos y, las filas están asignadas a una sola muestra de las dos distribuciones basadas en esta transformación. Las filas son entonces al cuadrado, para proporcionar más potencia estadística (Conover, 1980), y se suman de manera similar a la prueba de Mann-Whitney. La estadística de prueba, T , es igual a:

$$T = \sum_{i=1}^n [R(U_i)]^2 \quad (23)$$

Donde

$$U_i = |X_i - \mu_x| \quad (24)$$

Para propósitos del ranking de los parámetros, el procedimiento de normalización ejecutado en la estadística de Mann-Whitney también es necesario con la estadística cuadrada filas.

Con las definiciones anteriores, se puede destacar que el AS es fundamental para dar robustez y credibilidad a cualquier modelo, además de contribuir a la

reducción de la dimensionalidad de los modelos, disminución del costo computacional y optimización de recursos (Saltelli et al., 2000; Gómez y Barredo, 2005; Saltelli et al., 2008).

Diferentes técnicas han sido propuestas, muchas de las cuales han sido desarrolladas en el campo de modelado numérico, sin embargo, tienen limitada aplicación en modelos espaciales y no consideran variación en la distribución espacial de las variables incluidas (Plata et al., 2012).

Las técnicas empleadas de AS generalmente ejecutan repetidamente el modelo mediante simulaciones de Monte Carlo y muestras de los criterios de entrada (Saltelli et al, 2000). Además de considerar el rango completo de variación de la entrada, caracterizada por la unión de sus funciones de distribución de probabilidad a partir de un muestreo extraído (Lilburne y Tarantola 2009; Saltelli et al., 2000). Sin embargo, reducir los datos espaciales a valores escalares restringe severamente la habilidad de modelar la estructura y variabilidad espacial (Lilburne y Tarantola 2009).

Por otra parte, el método más simple para realizar esos dos tipos de análisis es cambiar un componente de decisión a la vez, mientras los demás permanecen constantes, observando el resultado de los cambios en la salida del modelo. Este enfoque es llamado “Un factor a la vez (OAT)” creado por (Daniel, 1958) y es, probablemente, el más popular entre los modeladores espaciales y otros ambientes.

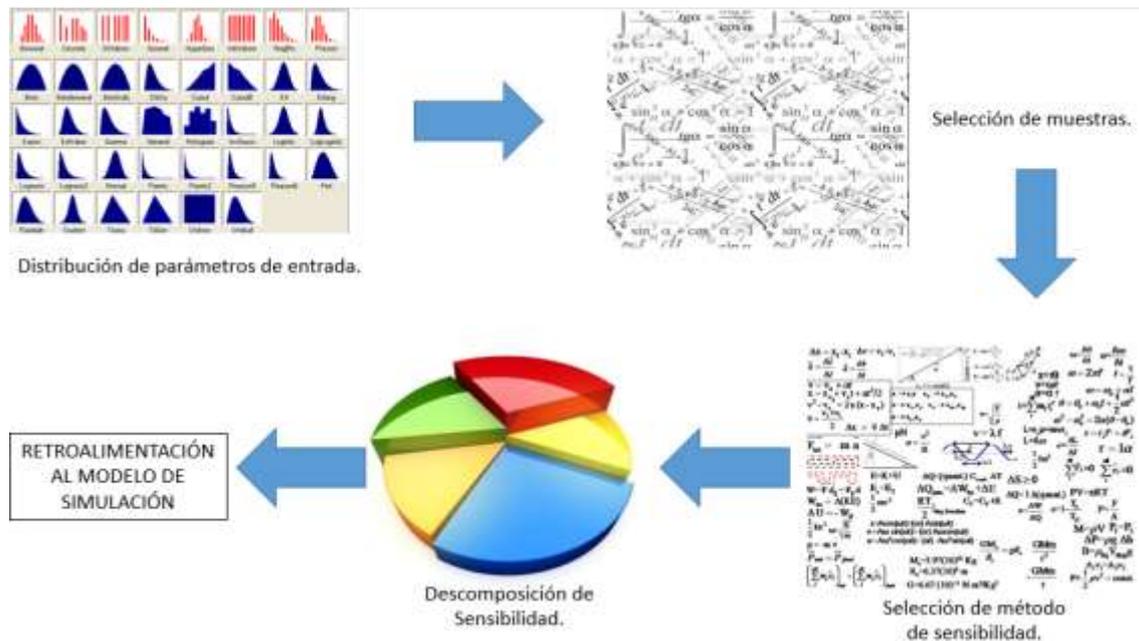


Figura 9.- Metodología de un análisis de sensibilidad estadístico usando PDF's.

3.1.2 Análisis de sensibilidad en modelos geoespaciales

Desde el aspecto geoespacial, los Sistemas de Información Geográfica son una poderosa herramienta que contribuye en las fases de descripción y análisis del territorio, observación de interrelaciones entre variables, generación de alternativas y evaluación de propuestas, como parte del proceso de planificación (Gómez y Bosque 2004). Por lo que fusionadas a las diferentes técnicas de simulación se proporcionan un mejor procedimiento para el soporte a la toma de decisiones (Voogd 1983; Barredo 1996; Malczewski 1999; Joerin 2001).

Por otra parte, idealmente un análisis de sensibilidad para modelos geoespaciales debe trabajar con los datos de entrada distribuidos espacialmente (mapa de entrada), es decir con píxeles, permitiendo así que la estructura espacial se incluya en el análisis. Por lo tanto, se deben excluir técnicas de regresión y técnicas rápidas, ya que requieren valores de entrada escalares.

Se han observado algunos avances en los métodos estadísticos Sobol', FAST e EFAST para tratar, en cierta medida, con datos distribuidos de manera espacial (Crosetto et al., 2000; Crosetto y Tarantola, 2001, Crosetto et al., 2001; Crosetto et al., 2002; Tarantola et al., 2002; Lilburne y Tarantola, 2009; Saint-Geours y Lilburne, 2010; Saint-Geours et al., 2011; Plata et al., 2012; Feizizadeh et al., 2014; Ligmann-Zielinska y Jankowski, 2014).

Las mejoras de estas técnicas (semi-espaciales) son, que el análisis es realizado a partir de un modelo con entradas espaciales (mapas), sin embargo, estos son caracterizados según su PDF, por lo que se considera que también se pierde información relevante sobre la distribución espacial de la sensibilidad.

A su vez, se encontró un procedimiento novedoso para realizar AS, y este utiliza la técnica de Distancia de Movimiento de Tierra (EMD), la cual mide la similitud entre dos distribuciones, calculando de estas, el costo mínimo de cambiar una distribución a otra basado en las distancias de diferentes características (Roura-Pascual et al., 2010; Xu y Zhang, 2013), sin embargo al basar el estudio en los histogramas de las variables espaciales, también se pierde la característica explícitamente espacial.

Por otro lado, una técnica de AS (espacial) que ha sido utilizada, con frecuencia, por la fácil integración de datos geoespaciales en el cálculo de la sensibilidad, es el método de un factor a la vez (OAT) (Saint-Geours y Lilburne, 2010; Chen et al., 2010; Plata et al., 2013).

El análisis OAT pretende obtener el impacto local de las variables de entrada del modelo, modificando un factor (o peso) mientras que los demás permanecen invariantes (Plata et al., 2013). Este tipo de análisis es definido como "local" (Crick et al., 1987), pero es posible realizar pruebas potentes examinando el cambio en la salida cuando se altera el parámetro de entrada por un factor de su desviación estándar, y de esta manera se considera la variabilidad del parámetro y la influencia asociada a la salida del modelo (Hamby, 1994).

Las diferentes técnicas de simulación geoespacial han sido capaces de aplicar AS de tipo espacial, transformando los métodos estadísticos y tratando de que se considere en alguna forma la distribución espacial, para que los resultados sean observados espacialmente.

La idea de realizar un AS en modelos basados en AC es evaluar la incidencia que tiene cada factor que compone al AC, para ello es necesario realizar versiones del modelo, eliminando factores y comparando con un modelo de

referencia y así determinar la influencia de los diversos parámetros. También se puede realizar este análisis no eliminando sino alterando parte o todo el factor.

Para lograr el realizar un análisis de sensibilidad en AC se han realizado estudios donde se varían las características de los componentes de los AC como lo son el tamaño y el tipo de vecindad y usando la técnica de OAT (Kocabas y Dragicevic, 2004; Chen y Mynett 2003), también se han analizado en otros estudios los cambios en los resultados a partir de alteraciones en las reglas de transición (White et. al 1997), así como analizar la resolución espacial variando el tamaño de la celda (Dietzel y Clarke 2004), de la misma forma solo que utilizando un enfoque multi-escala (Berling-Wolf y Wu 2004), también se varían los modelos al eliminar factores a partir del modelo inicial (González et al., 2015).

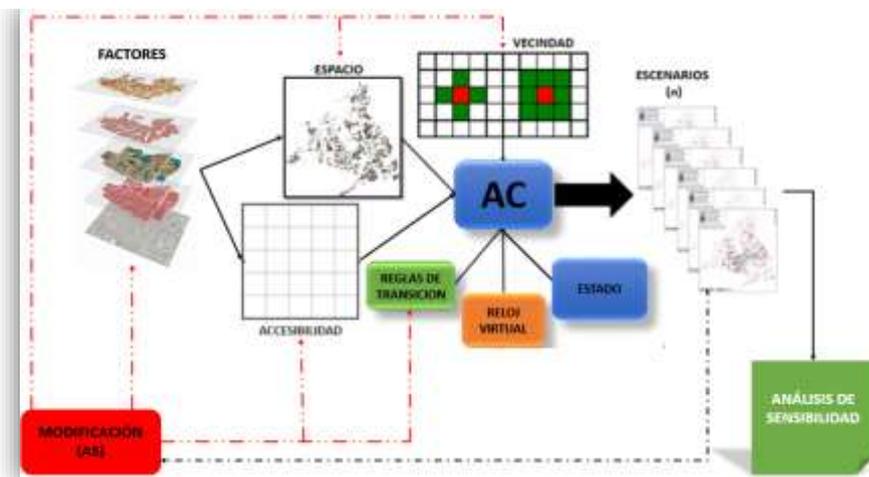


Figura 10.- Modificaciones realizadas en un AS para simulaciones geoespaciales basadas en Autómatas Celulares.

En los MBA no está claro hasta qué punto algunas de las metodologías son adecuadas para el tipo de sistema que comúnmente representados por estos modelos, por lo que ninguna provee un análisis completo del comportamiento, además todas las técnicas están fundamentadas estadísticamente por lo que requieren ser interpretadas por parte de los usuarios para aplicarse (Broeke et al 2016).

En el análisis se puede encontrar una evaluación de los métodos OAT, regresión lineal e índices de Sobol para entender patrones emergentes, robustez e incertidumbre de la salida. De los aspectos a resaltar entre las conclusiones es

que el OAT extendido proporciona una visión de los aspectos cualitativos del modelo y de los patrones emergentes del mismo y su costo computacional es bajo (si no existen muchas repeticiones del análisis).

Por otro lado los métodos basados en regresión, están limitados ya que dependen de la estructura del modelo para realizar una buena descomposición, pero estos podrían proveer utilidad en seleccionar parámetros influenciados, finalmente se encuentra que los métodos basados en varianza no son una buena medida para analizar distribuciones que están sesgadas o tienen valores atípicos (Swinson 2007; Liu et al 2006), por lo que los índices de Sobol no revelan la robustez, de patrones del modelo, a cambios de parámetros del modelo, además este tiene un gran costo computacional, para todos los análisis se realiza una variación del rango de los factores de entrada desde, las modificaciones oscilan entre 0.05 hasta 1 dependiendo del valor real del criterio a alterar y a su vez se consideraron diferentes cantidades de agentes (Broeke et al 2016).

Uno de los objetivos del AS en modelos de simulación basados en EMC es identificar los factores que determinan la varianza de la salida del modelo, otro es simplificar el modelo, identificando factores que contienen incertidumbre casi despreciable para los resultados (Gómez y Taratola 2006).

Para lograr el objetivo, estas realizan diferentes variaciones de las cuales destacan modificar los rangos y PDF de los factores, así como la variación de pesos desde +/- 20% y hasta un máximo de +/- 75% del valor real, alterando especialmente pesos y factores, agrupación por separado de criterios y factores (para reducir el tiempo de procesamiento), también se han realizado modificando la anchura del árbol (Gómez y Taratola 2006; Qureshi 1999; Oran et al. 2010; Plata et al. 2013; Plata et al 2012). Por otra parte también se ha realizado estudios por medio de una variación espacial de pesos y factores (Plata et al. 2013; Plata et al 2012).

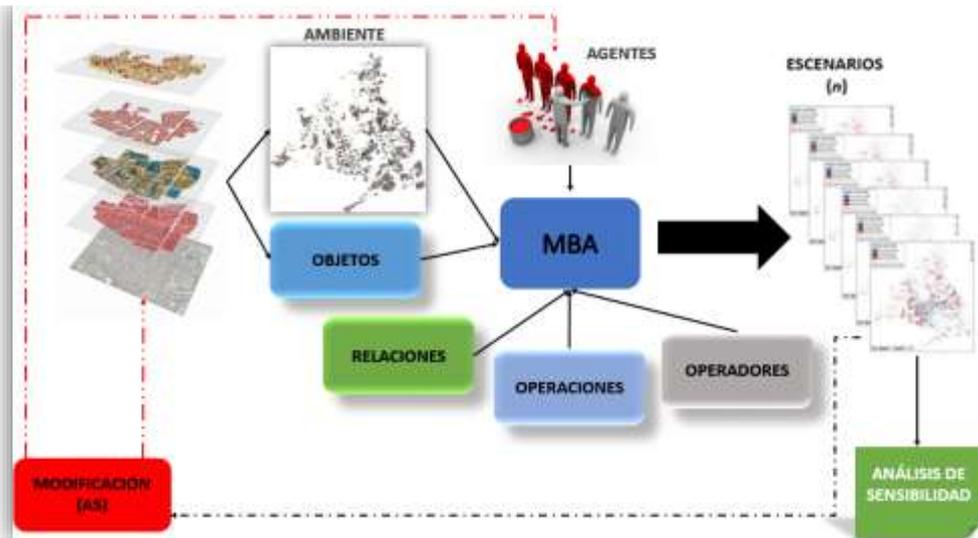


Figura 11.- Modificaciones realizadas en un AS para simulaciones geoespaciales basadas en los MBA.

La variación de pesos es la forma más usada para llevar a cabo un AS en modelos de EMC (Gómez y Bosque 2004).

De las técnicas usadas en estas simulaciones se puede encontrar los índices basados en descomposición de la varianza tales como OAT, Sobol, FAST y E-FAST (Gómez y Taratola 2006; Crosetto et al. 2000, Plata et al. 2013; Plata et al. 2012).

De acuerdo con Pascual et al. (2010) donde después de realizar una comparación de los métodos de simulación, de tipo exploratorio y/o normativo, basados en técnicas de AC y EMC aplicado al crecimiento urbano, concluye que los mejores resultados son obtenidos por la evaluación multicriterio.

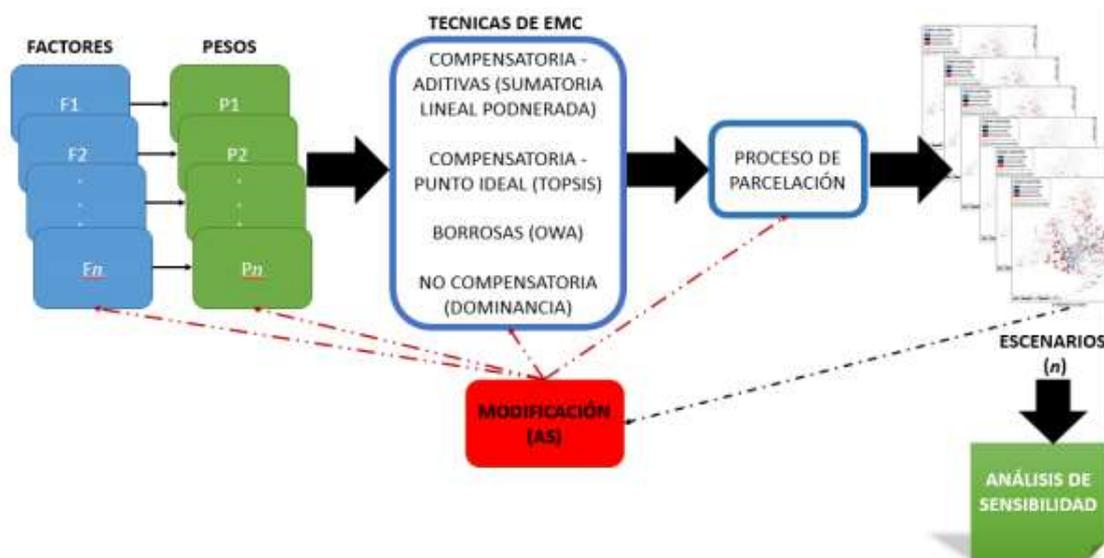


Figura 12.- Modificaciones realizadas en un AS para simulaciones geoespaciales basadas en EMC.

3.2 Desarrollo computacional

3.2.1 Lenguaje de programación C#

El lenguaje de programación C # es simple, moderno, orientado a objetos y seguro para tipos que combina la alta productividad de los lenguajes de desarrollo de aplicaciones rápidos con la potencia bruta de C y C ++ (Hejlsberg et al., 2003).

Este lenguaje como cualquier otro orientado a objetos tiene soporte para definir y trabajar con clases. Las clases definen nuevos tipos de datos, lo que le permite ampliar el idioma para modelar mejor el problema que está tratando de resolver. C# contiene palabras clave para declarar nuevas clases y sus métodos y propiedades, y para implementar la encapsulación, la herencia y el polimorfismo, los tres pilares de la programación orientada a objetos.

También admite interfaces, un medio para realizar un contrato con una clase de servicios que estipula la interfaz. Cuando se implementa una interfaz, una clase en efecto promete proporcionar la funcionalidad que la interfaz especifica. Además, proporciona soporte para structs, un concepto cuyo significado ha cambiado significativamente de C++. En C#, una estructura es un tipo restringido y ligero que, cuando se crea una instancia, exige menos en el sistema operativo y en la memoria que una clase convencional.

Otro rasgo de C# es que proporciona características orientadas a los componentes, como propiedades, eventos y construcciones declarativas (como atributos). Se debe agregar que este lenguaje también brinda soporte para: acceso directo a la memoria con punteros de estilo C ++, palabras clave para poner entre corchetes operaciones insegura, advertencia al recolector de basura para que no recoja los objetos a los que se hace referencia mediante punteros hasta que sean liberados (Liberty, 2005).

3.2.2 Idrisi selva

IDRISI es un software de analítica raster y cubre toda la gama de requerimientos para SIG y sensores remotos desde incertidumbre sobre las bases de datos y modelación espacial hasta mejoramiento y clasificación de imágenes, incluyendo utilidades espaciales para el monitoreo ambiental y el manejo de recursos naturales, análisis de cambios y series de tiempo, apoyo a las decisiones de criterios y objetivos múltiples, análisis de incertidumbre, modelos de simulación, interpolación de superficies y I caracterización de estadísticas. Sin embargo, a pesar de las sofisticadas utilidades, es de fácil manejo (Eastman, 2012).

Este sistema consiste en una interface principal, donde se pueden encontrar 250 módulos de programas que proveen usabilidad para la entrada, visualización y análisis de datos geográficos (Eastman, 2012).

3.2.3 Integración de sistemas

La integración del lenguaje C# e IDRISI Selva se llevó a cabo con la finalidad de que la herramienta a desarrollar sea fácilmente incluida dentro del SIG considerado, ya que este sistema tiene de manera nativa e lenguaje.

Para la mayoría de los procedimientos realizados se utilizan comandos del sistema IDRISI ejecutados desde la aplicación desarrollada, por medio de la comunicación basada en la API "idrisi32" para C#, incluyéndola en el código y realizando la conexión a partir de la creación de un servidor con el comando "IdrisiAPIServer".

Iniciado el servidor es posible utilizar algunos comandos para modificar o utilizar algunas de sus utilidades como:

- SetWorkingDir: para cambiar de directorio de trabajo.
- Set_Process_ModuleName: para definir el nombre del módulo que se va a realizar.
- RunModule: para ejecutar el procedimiento elegido, proveyendo de los datos de entrada necesarios para cada procedimiento (nombre de entrada, salida, selección de tipo de elementos, etc.).
- NotifyWorking: para notificar si está trabajando algún procedimiento ejecutado.
- ProcessFinished: para finalizar los procesos que están activos.
- FreeProcess: para liberar los procesos activos y cerrar el servidor de conexión.

Algunos de los módulos usados en las diferentes etapas de desarrollo y pruebas de la metodología fueron: “RECLASS, OVERLAY, STRETCH, RANK, RANDOM, MOLA, GROUP, AREA, EXTRACT, ASSIGN, EXPAND, CROSSTAB, MCE”.

5.- METODOLOGÍA

El uso de enfoques analíticos de AS para modelos espaciales puede ser difícil de implementar y se limita a un análisis local de la relación entrada/salida. En consecuencia, el análisis de sensibilidad de modelos espaciales a menudo se ignora o es relativamente limitado en su alcance. La mayoría de los métodos de análisis de sensibilidad global usados en la literatura están basados en varianza.

En las últimas dos décadas ha existido una gran variedad de trabajos sobre análisis de sensibilidad sin embargo solo unos pocos han sido estrictamente espaciales (Plata, et al., 2013; Ligmann-Zielinska & Jankowski, 2014), por lo que se tiene un gran interés en el estudio de esta área tan poco aprovechada en el área de geomática.

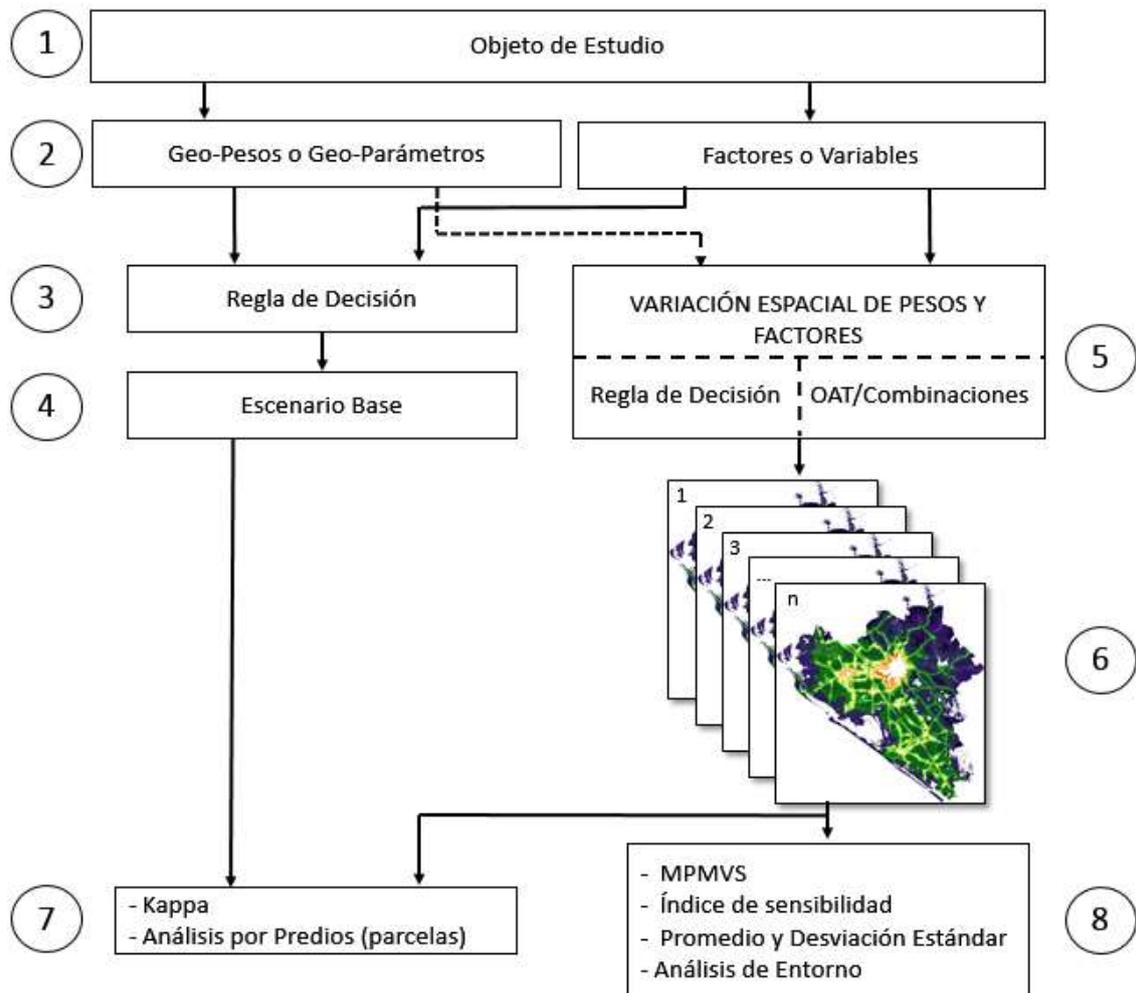


Figura 13.- Propuesta Metodológica para realizar AS explícitamente espacial.

5.1 Área de estudio empleada como base para probar propuesta metodológica

El estado de Sinaloa está integrado por 18 municipios entre los cuales destaca Culiacán como la Capital del Estado y Navolato el municipio más joven (separado de Culiacán en 1982). Los municipios de Culiacán y Navolato se encuentran ubicados entre los paralelos 24° 02' y 25° 59' de latitud norte y los meridianos 106° 52' y 108° 04' de longitud oeste; con una altitud entre 0 y 800 m., colindando al norte con los municipios de Angostura, Mocorito, Badiraguato y el estado de Durango; al este con el estado de Durango y los municipios de Cosalá y Elota; al sur con el municipio de Elota y el Océano Pacífico.

El municipio de Culiacán ocupa el 10.96% de la superficie del estado y cuenta con un total de 1,483 localidades (INEGI 2009), mientras que el municipio de Navolato le corresponde un total de 3.71% con 498 localidades. Ambos municipios ostentan un 14.67% del territorio estatal, con un superficie de 8,760.74 km², en los cuales se distribuye un total de 994,241 habitantes (INEGI 2010). El mayor porcentaje de habitantes radica en las ciudades capitales, Culiacán y Navolato con 858,638 y 135,603 habitantes, respectivamente (Anon., 2010).

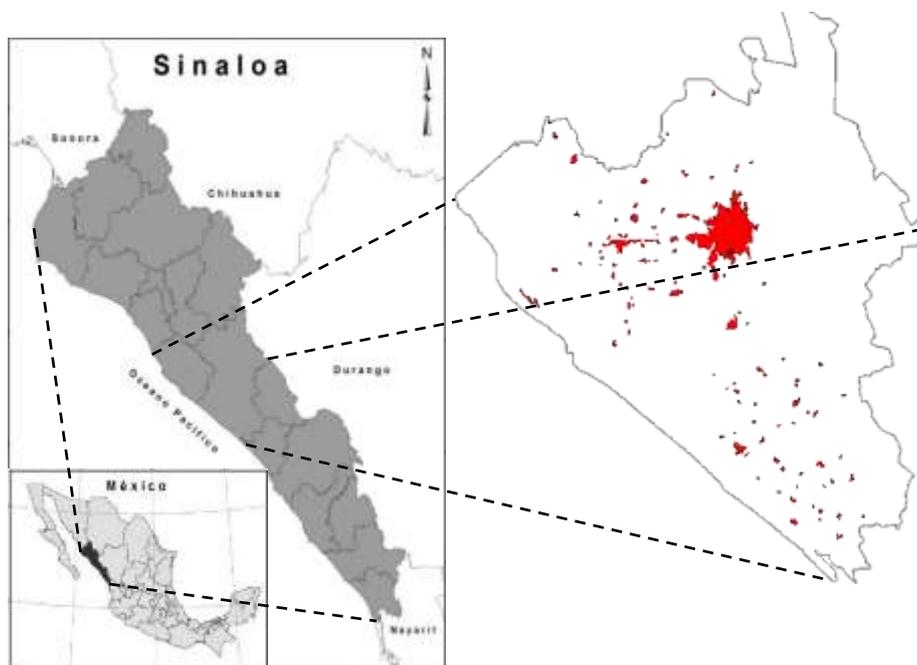


Figura 14.- Mapa de municipios de Culiacán y Navolato, Sinaloa, México.

En el área de estudio se tiene una diversidad de usos de suelo, algunos naturales y otros artificiales, ocasionados por los procesos de antropización que se ha producido en el territorio, producto de diferentes actividades socioeconómicas propias de la región. Actualmente el área de estudio cuenta con coberturas agrícolas, vegetación natural, áreas urbanas y cuerpos de agua.

Los cambios más representativos se presentan en estos municipios, donde se encuentra la ciudad de Culiacán y Navolato mantiene una intensa dinámica debido a su cercanía con la capital y su potencial agrícola y a la inversión en proyectos turísticos relacionados con el Mar de Cortes.

Otro aspecto fundamental a considerar es el hecho que ambas cabeceras municipales trazan un eje artificial donde interactúan entres rurales, urbanos, industriales y comerciales, lo cual induce a la conformación de un área metropolitana, lo que implica a corto plazo el desarrollo de planes territoriales en el corredor Culiacán-Navolato (López, 2006).

Estos dos municipios son de gran importancia para el estado ya que en estos se ha experimentado una inmensa dinámica de suelo, lo que podemos observar en (Corrales Barraza, 2013), quien realiza un estudio de cambios históricos de usos de suelo, que el municipio de Culiacán experimentó el crecimiento urbano más significativo y de mayor tamaño en el estado. Además en (Monjardín, 2015) se puede analizar que Culiacán es la principal zona “hot spot”, ya que es el municipio con mayor transición de selvas hacia agricultura. Es por lo tanto que existe la necesidad urgente de planificar y establecer modelos que permitan desarrollar planes para un desarrollo sustentable (Corrales Barraza, 2016).

Para el desarrollo de la investigación se dispuso de los mapas del año 2011, todos obtenido del Instituto Nacional de Estadística y Geografía (INEGI). La escala original de la información es en general 1:100.000, aunque en algunos casos se tuvo que recurrir a datos de escala mayor o menor. Toda la información fue tratada en formato raster en el SIG IDRISI Selva, utilizando un tamaño de píxel de 100m y proyectada en el sistema de referencia espacial UTM Zona 13 N.

5.2 Definición de objetivos

La información de partida necesaria a la hora de realizar este tipo de estudios es el condicionante principal, de cuya calidad dependerá, en gran medida, el éxito o la adecuación del modelo. En el caso de los modelos espaciales de simulación de usos de suelo, esta información de partida consiste de una o más fechas diferentes en el tiempo de mapas de usos de suelo y vegetación.

La información de partida, son los mapas de usos del suelo del año 1993 y 2011, obtenidos del Instituto Nacional de Estadística y Geografía (INEGI). Dichos mapas, presentan diferentes clasificaciones en cuanto a las categorías de usos de suelo.

Para la realización de los escenarios se debe trabajar con una clasificación de usos de suelo específica, debido a la dificultad de representar las dinámicas de una gran variedad categórica de usos de suelo, es por ello que se tomó la determinación de homogeneizar la leyenda de los mapas de 1993 y 2011, las cuales tienen una leyenda jerárquica de 30 y 18 categorías.

Ante esto, se estableció una leyenda conformada por 7 categorías, las cuales representan adecuadamente la dinámica entre los diferentes usos de suelo (Tabla 3). Se reclasificaron y homogeneizaron los mapas originales a este nivel de agregación de información temática, con la finalidad de tener una correcta equivalencia en la leyenda (Tabla 3).

Tabla 3.- Clasificación de los usos de suelo.

I.- Agricultura de Riego y Humedad

II.- Agricultura de Temporal

III.- Asentamiento humano

IV.- Bosque

V.- Otro tipo

VI.- Pastizales

VII.- Selva

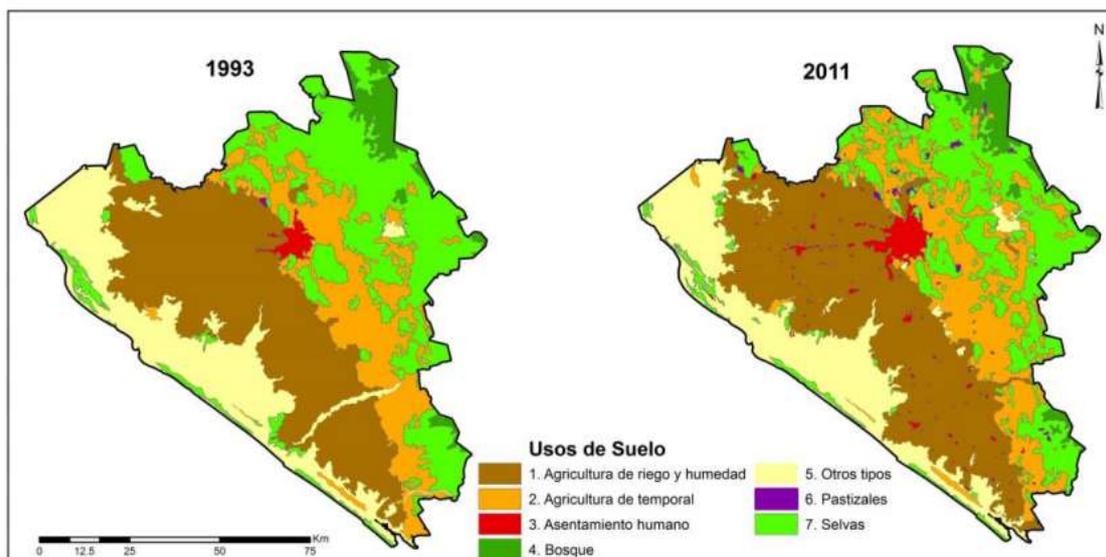


Figura 15.- Mapas de usos de suelo del modelo.

5.3 Modelación de variables geoespaciales

El principal objetivo en el que se basó este trabajo es analizar el escenario de sostenibilidad que se fundamenta apegado a los tres principios básicos de la sostenibilidad: minimizar el impacto ambiental y maximizar el beneficio social y económico. El objetivo de este escenario es ver qué pasaría si existen políticas de reforestación más activas y se le otorga más importancia al cuidado del medio ambiente y de los elementos naturales, pero teniendo en cuenta que la región urbana continuará creciendo, considerando la sostenibilidad del territorio para su evolución.

La información de partida necesaria a la hora de realizar este tipo de estudios es el condicionante principal, de cuya calidad dependerá, en gran medida, el éxito o la adecuación del modelo. En el caso de los modelos espaciales de simulación de usos de suelo, esta información de partida consistirá de una o más fechas diferentes en el tiempo de mapas de usos de suelo y vegetación.

En cuanto al escenario de sostenibilidad y de acuerdo al modelo de demanda de superficies de Corrales (2016), las categorías que aumentan de superficie, entre los años 1993 y 2011, son los asentamientos humanos, bosques, pastizales,

otros tipos y selvas, mientras que las que disminuirán de superficie son la agricultura de riego y humedad y la agricultura de temporal.

Es entonces que a partir de cada objetivo de simulación se seleccionan las variables y los factores a utilizar. La selección de los factores que determinan los cambios de usos de suelo supone una tarea ardua, ya que son multitud de datos y variables las que pueden considerarse, aunque no todos los procesos que generan cambios de usos del suelo pueden ser modelados (Alcamo et al., 2006), ya que éstos varían dependiendo de la escala espacial. Por ejemplo, a escala local, las variables que más influyen el uso del suelo son de tipo social y de accesibilidad; mientras que a escala de paisaje, el determinante principal es la topografía (Veldkamp y Lambin, 2001).

De esta manera, la selección de las variables se realizó tomando en consideración todos los objetivos del escenario propuesto, así como la información especializada en estos temas. De igual manera, la disponibilidad de información para el área de estudio fue determinante para definir las variables.

Después, se realizó un análisis de cada uno de los objetivos de simulación, para conocer como influirían las variables en cada uno de estos. Partiendo de este análisis, se obtuvieron los factores para cada objetivo del escenario de sostenibilidad.

Proximidad a asentamientos humanos: Los asentamientos humanos son el detonante de las dinámicas en los usos de suelo, es por eso que en este factor, se consideró de mayor importancia a las zonas más próximas a ellos.

Proximidad inversa a usos de suelo: Este factor representa la distancia hacia el interior de cada uno de los usos de suelo en los que se ha experimentado el crecimiento de cada categoría a simular, asignándosele un valor de idoneidad a cada uno de estos usos.

Proximidad a Carretera: Las carreteras juegan un papel importante en la movilidad y desarrollo de las actividades del ser humano. Es por ello que es

importante considerar la distancia a las carreteras como un factor preponderante para mantener en funcionamiento las dinámicas del territorio. De tal manera, que se ha realizó un factor de proximidad a carreteras, en el cual, se da más aptitud a las zonas más próximas a las carreteras.

Pendiente: Este factor representa las pendientes en las que se dará el cambio de cada objetivo de simulación en ambos escenarios.

Proximidad a Ríos: La accesibilidad a los ríos se considera esencial para el desarrollo de algunas actividades de uso de suelo, por lo que, en algunas categorías, los lugares más próximos a los ríos se valoran como los más y en otros como menos óptimos.

Para la homogenización de las diferentes unidades de medida entre las variables, se realizó una normalización a una escala de 0 – 255 para generar factores finales. Para ello, se aplicaron operadores borrosos (fuzzy), en funciones lineales y sigmoidales implementadas en el módulo FUZZY en el software IDRISI, las cuales se ilustran a continuación.

La función monotónicamente creciente se eleva de 0 a 255 y nunca disminuye y la función monotónicamente decreciente comienza en 255, luego disminuye a 0.

La lógica borrosa o difusa es una técnica que ayuda a la representación más exacta de la información imprecisa, incompleta o incierta, implícita en la opinión de los expertos en la toma de decisiones (Prakash, 2003).

Las funciones borrosas se caracterizan por un grado de pertenencia difusa (también llamado posibilidad) que oscila entre 0.1 a 1.0, indicando un aumento continuo de la no pertenencia a la pertenencia completa.

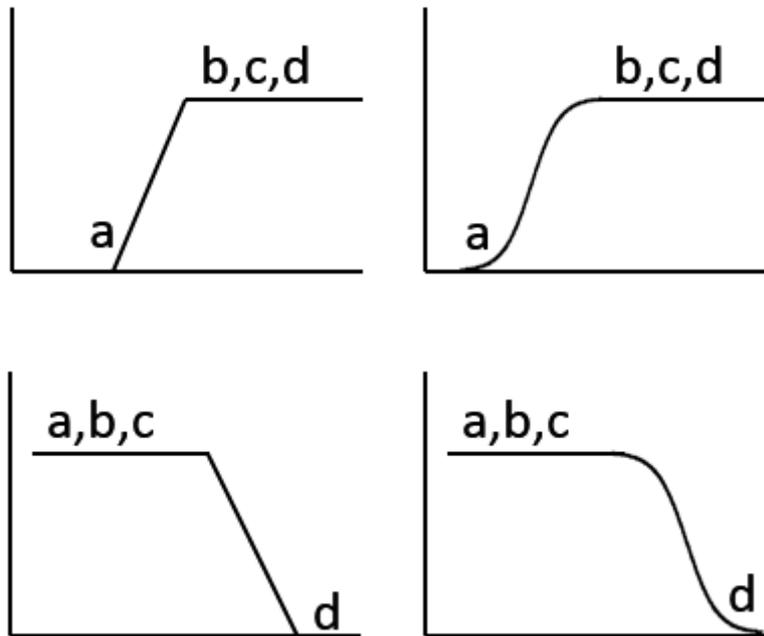


Figura 16.- Funciones lineales y sigmoideas monotónicamente crecientes (arriba) y decrecientes (abajo).

En su mayoría se utilizaron funciones lineales decrecientes, pues la normalización de los factores de proximidad, se realizó dando mayor importancia a las distancias más cercanas a la actividad a simular.

Un tema crítico en la normalización de factores es la elección de los puntos finales donde la pertenencia del grupo alcanza 0.0 o 1.0 (0 o 255). No se aconseja usar una escala lineal a ciegas entre los valores mínimos y máximos de la imagen. Al desarrollar factores estandarizados usando FUZZY, se debe considerar cuidadosamente el significado inherente de los puntos finales elegidos (Eastman, 2003).

Es debido a lo anterior, que los valores mínimos y máximos utilizados al momento de la normalización se obtuvieron del análisis de la dinámica entre los usos de suelo. Considerando así, diferentes valores en las funciones de normalización para cada uno de los objetivos a simular.

Para la ponderación de los factores, se buscó que se cumpliera lo planteado en el escenario propuesto, que en la ponderación se persiguió que se cumpliera, en la medida de lo posible, las condiciones de un modelo de uso de suelo óptimo,

tratando de manera conjunta el interés que pudieran tener los diferentes grupos implicados en el desarrollo territorial.

A partir de esto, se procedió a la ponderación de los factores utilizando el método de jerarquías analíticas de Saaty (Saaty, 1980), el cual establece mediante una matriz de comparación entre pares de factores, el nivel de importancia de cada uno sobre los demás.

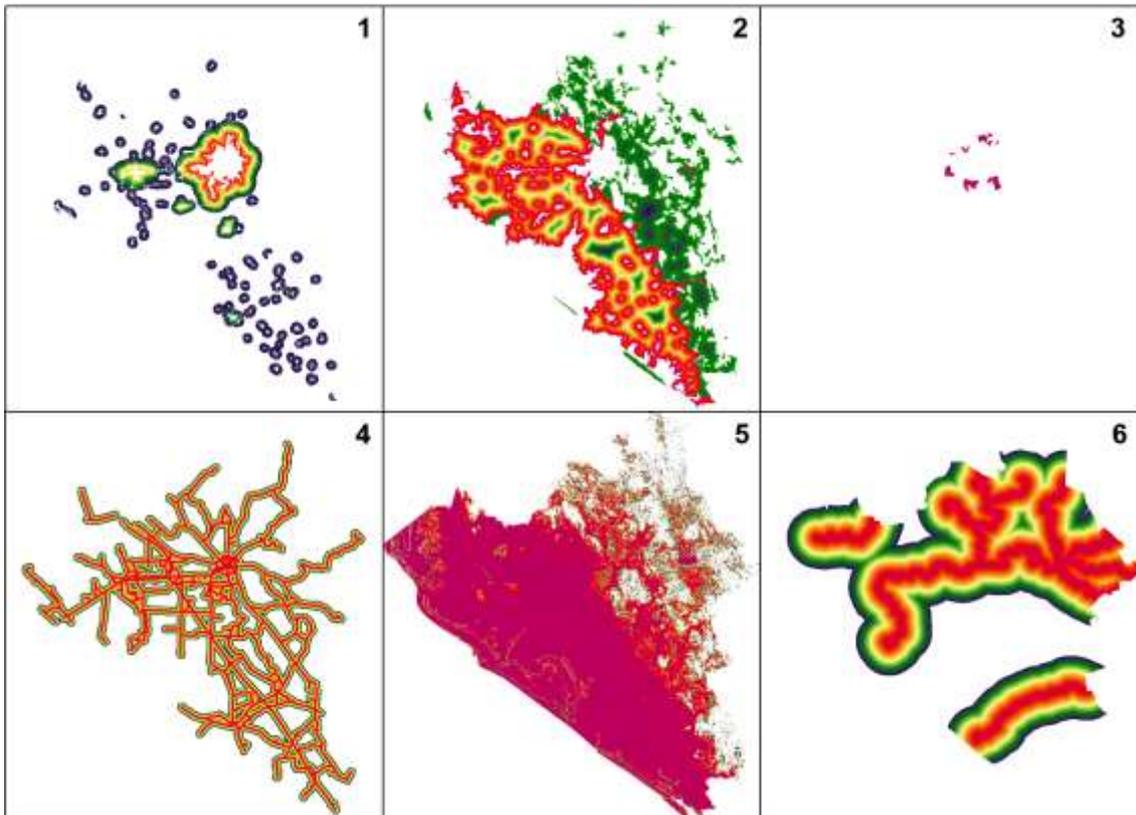


Figura 17.- Factores *modelados que componen el modelo escenario sostenible.*

Para determinar las jerarquías de orden, se partió de la información derivada de la revisión bibliográfica realizada, junto a la opinión consensuada con expertos del grupo de Investigación de Geomática de la Universidad Autónoma de Sinaloa.

Finalmente, determinadas todas las jerarquías, la obtención de pesos se obtuvo utilizando el módulo WEIGHT de IDRISI, en el cual se introducen los niveles de importancia de los factores para obtener los pesos correspondientes a cada factor.

Tabla 4.- Factores y pesos contemplados para el modelo a investigar.

FACTORES	PESOS
1.- Proximidad a Asentamientos humanos	0.40
2.- Proximidad Inversa a US y V	0.20
3.- Área Urbanizable Programada	0.10
4.- Proximidad a Carreteras	0.20
5.- Pendiente	0.05
6.- Proximidad a Ríos	0.05

Para obtener una evaluación final de las alternativas se utilizó la sumatoria lineal ponderada, la cual es uno de los métodos más utilizados en la EMC por su relativa sencillez y poca complejidad en su implementación (Gómez y Barredo, 2005). El nivel de adecuación de cada alternativa se obtiene sumando el resultado de multiplicar el valor de cada criterio por su peso:

$$r_i = \sum_{j=1}^n w_j v_{ij} \quad (35)$$

Donde w_i es el peso del factor i y v_{ij} es el valor de la celda en ij .

La sumatoria obtenida, es multiplicada por el mapa de restricciones, obteniendo así e mapa de aptitud o adecuación.

5.4 Modificación de variables y factores geoespaciales

Para este trabajo los factores y pesos de entrada del modelo están distribuidos espacialmente (mapas espaciales), por lo que la aplicación de técnicas estadísticas convencionales suele ser limitada.

Esta propuesta metodológica se basa en la introducción de un cierto porcentaje de variación ($\pm 25\%$, para este caso) de forma aleatoria en los valores de adecuación de los factores espaciales y en los pesos rasterizados a nivel de píxel

(originalmente un valor nominal, con el mismo valor para todos los píxeles). Para llevar a cabo este proceso se siguieron los pasos representados en la figura 18.

En dicho esquema, el procedimiento seguido para obtener los factores con variación espacial es el siguiente: primero se generó una muestra aleatoria entre -0.25 y +0.25 (mapa ráster); posteriormente se multiplica este mapa por el factor que deseamos modificar, obteniendo un segundo mapa ráster con la variación aleatoria entre $\pm 0,25$ del valor de adecuación del factor; y, por último, dicho mapa se suma al factor original, obteniéndose el factor con una variación entre $\pm 25\%$ de su valor de adecuación original a nivel de píxel.

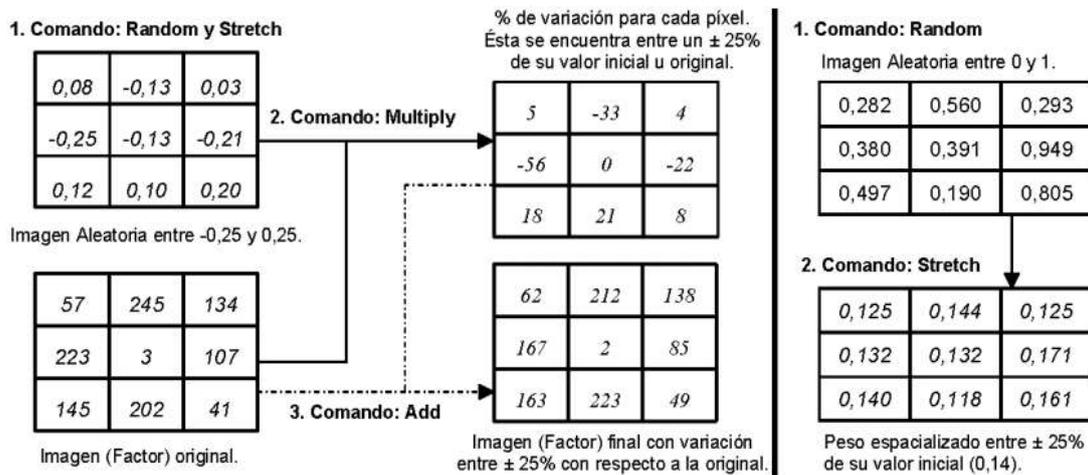


Figura 18.- Proceso para obtener la variación espacial en los factores (izqda.) y pesos (dcha.). (Plata et al, 2013)

Por ejemplo para el primer factor, el rango original de valores es de 1 a 111.90, y al aplicar la variación el rango aumenta de 1 a 139.87 (ver figura 19).

Además la propuesta está basada en las técnicas OAT, ya que este método de análisis investiga la sensibilidad de los pesos unidimensionales cambiando la influencia relativa de cada factor por separado, “sin supuestos” (Xu y Zhang, 2013), y combinatoria sin repetición para la modificación múltiple de los factores de entrada. Utilizando la fórmula:

$$\sum_{j=1}^m \binom{n}{m} = \sum_{j=1}^m \frac{n!}{m!(n-m)!} \quad (36)$$

donde n es el total de factores y pesos de entrada y m es el número de total de factores a modificar y combinar al mismo tiempo.

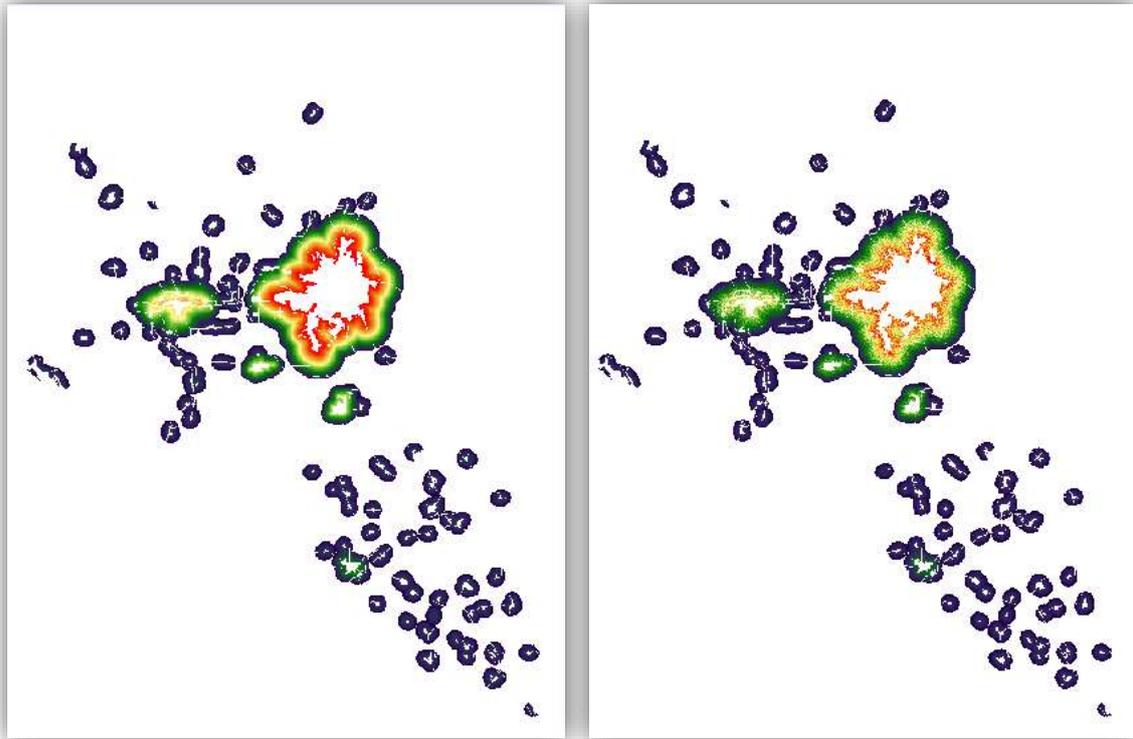


Figura 19.- Mapa original con rango máximo de 111.90 (izq), mapa modificado con rango máximo de 139.87 (der).

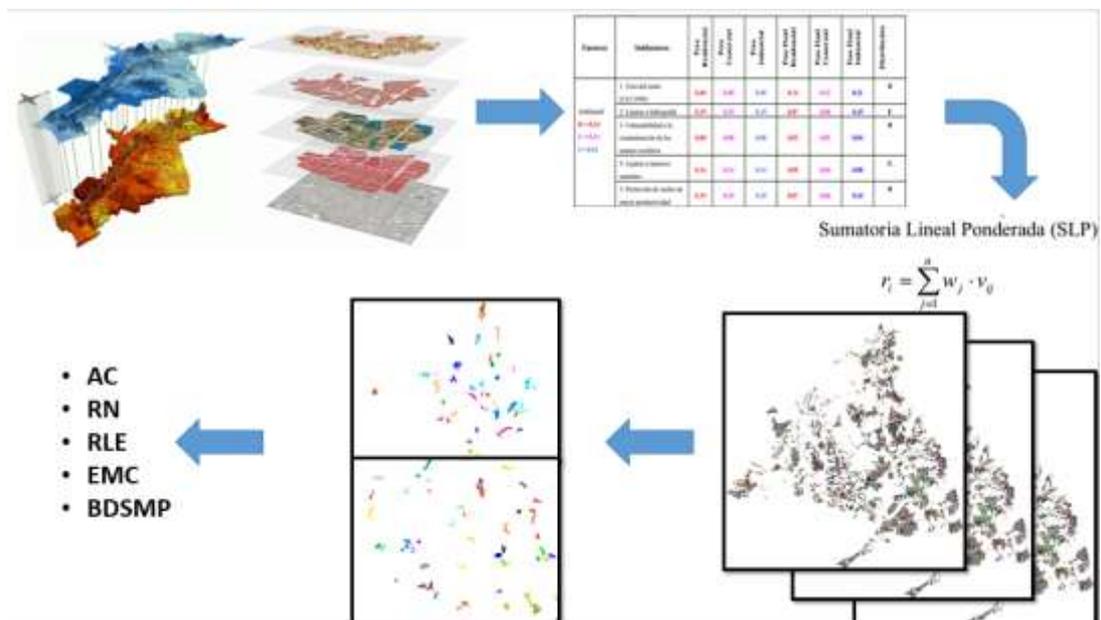


Figura 20.- Metodología de un proceso de simulación de escenarios con EMC y parcelación.

5.5 Indicadores de análisis de sensibilidad

El índice Kappa está basado en la comparación celda a celda, proveyendo de información sobre la coincidencia espacial de los resultados. Esta técnica propone una matriz para comparar dos variables categóricas, donde la diagonal de la matriz obtenida equivale a las concordancias entre los mapas y, los elementos fuera de la diagonal son desacuerdos de pixeles. (Gómez y Barredo, 2005). La definición del índice Kappa es:

$$k \equiv \frac{p_o - p_e}{1 - p_e} = 1 - \frac{1 - p_o}{1 - p_e} \quad (37)$$

Donde p_o es la proporción de acuerdos observados y p_e la proporción de acuerdos esperados en la hipótesis de independencia entre observadores, es decir, de acuerdos por azar (Cohen, 1960).

Por otra parte, el análisis por predios analiza la variabilidad de los pixeles considerando los patrones de predios (parcelas) y sus alrededores próximos, es decir, examina la sensibilidad, de cada celda, de aparecer dentro del predio o en las cercanías del mismo, considerando que, si el pixel es ubicado dentro de un rango (ventana) de búsqueda de la parcela, este tendrá la misma posibilidad de acoger la nueva categoría que las celdas que están dentro.

Por lo que se refiere al índice para generar Mapa de Pixeles Más Veces Seleccionados (MPMVS), está basado en un análisis pixel a pixel, calculando la frecuencia de un mismo pixel en los diferentes mapas a ser analizados, es decir, si el pixel aparece en la misma posición en otro mapa se suma. El resultado se obtiene a partir de una superposición de mapas (suma) en un SIG o utilizando la siguiente formula:

$$MPMVS_{i,j} = \sum_1^n M_{i,j} \quad (38)$$

donde n es el número de mapas totales y M son los mapas analizados individualmente.

A su vez, el Índice de Sensibilidad (IS), es un método que determina la sensibilidad de un parámetro al calcular la diferencia de porcentaje de salida variando un parámetro de entrada a partir de su valor mínimo a su valor máximo (Hoffman y Gardner, 1983; Bauer y Hamby, 1991). El IS se obtiene a partir de la siguiente formula:

$$IS = \frac{D_{max} - D_{min}}{D_{max}} \quad (39)$$

Donde D_{max} es el valor máximo y D_{min} es el valor mínimo de la variable.

Se realizó una modificación a la formula, para poder aplicarla a nivel de pixeles, quedando de la siguiente forma:

$$ISE_{i,j} = \frac{MAX_{i,j} - MIN_{i,j}}{MAX_{i,j}}$$

Otros índices considerados están basados en los cálculos estadísticos del promedio y la desviación estándar. El promedio calcula el valor característico de una serie de datos cuantitativos y la desviación estándar mide la dispersión del intervalo de valores de las variables.

Dados los “N” números $\{x_1, x_2, x_3, \dots, x_n\}$, la media aritmética se define como:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (40)$$

Mientras la fórmula para calcular la desviación estándar es:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (41)$$

Con respecto al Análisis de Entorno, este análisis puede generar dos índices estudiando la vecindad de un pixel de dos formas: simple y múltiple. En el entorno simple se analizan los distintos mapas individualmente, extrayendo información (promedio) del comportamiento del entorno de cada pixel. Para el caso múltiple, se extrae el promedio de la misma vecindad y el mismo pixel de todos los mapas.

El cálculo de estos dos índices se logra utilizando ventanas mejor conocidas como vecindad de von Neumann (3x3, 5x5, y 7x7) la cual analiza el conjunto de puntos situados a una distancia de Manhattan de valor 1 respecto al punto dado para el caso de la vecindad de 3x3. La extensión del concepto de la vecindad de von Neumann sencilla descrita anteriormente, también sirve para determinar el conjunto de puntos a una distancia de Manhattan con $r > 1$. El resultado es una región en forma de diamante (se muestra en la figura 21). Se denominan vecindades de von Neumann de rango o extensión r (Wilson y Ritter, 2000).

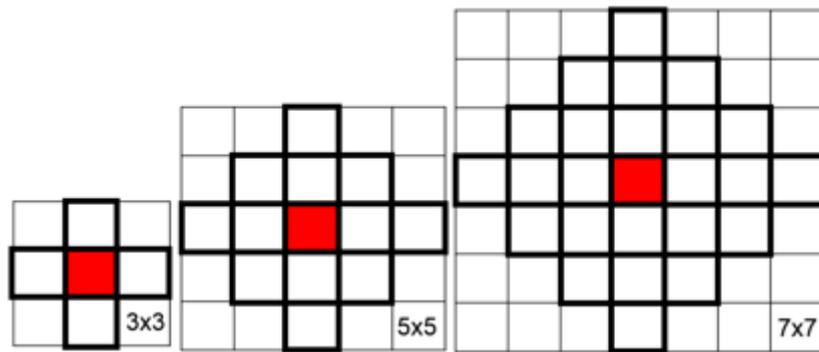


Figura 21.- Distintas vecindades de búsqueda de tipo von Neumann.

También se consideraron pruebas utilizando la vecindad de Moore de un punto es el conjunto de los puntos a una distancia de Chebyshev de valor 1 desde el punto dado. El número de celdas de la vecindad de Moore de un punto dado considerando r unidades de distancia, es:

$$(2r + 1)^2 \tag{42}$$

Así, para distancia $r=1$, vale 8; para distancia $r=2$, vale 24; y para $r=3$ vale 48 (Weisstein, 2005).

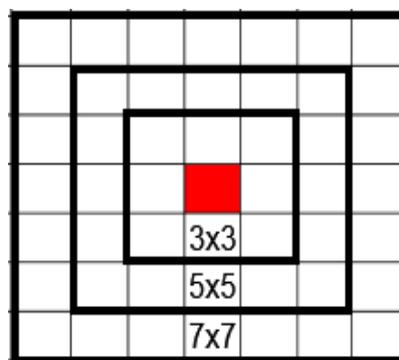


Figura 22.- Tamaños de ventanas usadas en análisis de entorno simple y múltiple.

6.- RESULTADOS

En este capítulo se muestran los resultados obtenidos en las diferentes fases de la metodología, así como el estudio de las herramientas de simulación encontradas en el estado del arte y también un error de selección encontrado en IDRISI.

6.1 Descripción del módulo diseñado para simular el Modelo de EMC y AS.

Para lograr obtener las simulaciones, fue necesario el desarrollo de una herramienta en el lenguaje C#, el cual de manera automática puede realizar ambos procesos de simulación. En el procedimiento a utilizar en el sistema, se pueden observar los siguientes pasos, que son realizados por separado para poder analizar los resultados y ajustar el modelo, Cargar Factores y Pesos, Generar Aptitud, calcula el (los) mapas de aptitud, Generar (genera los archivos necesitados por el proceso de MOLA), y MOLA, este último se encarga del proceso de toma de decisiones de los pixeles necesarios para satisfacer la superficie de cada objetivo (figura 23).

Además, una vez cargados los factores y pesos, se pueden seleccionar los botones de “SIMULACIÓN OAT – MODIFICANDO FACTORES” el cual modifica un factor a la vez y genera un mapa, “SIMULACIÓN OAT – MODIFICANDO PESOS” modifica cada peso individualmente y genera un mapa, por otra parte “SIMULACIÓN OAT – MODIFICANDO TOTAL” realiza alteraciones a los pesos y factores realizando todas las combinaciones pertinentes y generando un mapa de cada una de ellas.

A su vez para el análisis de entorno también desarrollar una herramienta que permitiera elegir los componentes y cálculos necesarios para analizar el o los escenarios como lo son el tipo de cálculo, el tipo y tamaño de ventana y el tipo de selección. El desarrollo fue solamente en el Lenguaje C#. La interfaz de la herramienta se puede observar en la figura 24. En esta herramienta se observan los botones de “ANÁLISIS POR ESTADÍSTICA” el cual estadísticas básicas (promedio, media y desviación estándar) de todos los mapas contenidos en la directorio de trabajo, “ANÁLISIS POR ENTORNO” calcula las mismas

estadísticas básicas pero por ventana de búsqueda que puede ser 3x3, 5x5, y 7x7 de todos los mapas contenidos en la carpeta de trabajo, “ANÁLISIS DE ENTORNO OAT” realiza lo mismo que el botón anterior pero solo considerando los mapas generados en la simulación OAT, “SELECCIÓN DE MÁS DE 676 PÍXELES” se desarrolló con la finalidad de analizar el crecimiento si los píxeles fueran más de 676 que es la cantidad calculada por el modelo de demanda y el botón de “SELECCIÓN ALEATORIA” fue contemplado debido a un error observado por IDRISI al realizar la selección de píxeles y el cual se explica en el siguiente apartado.

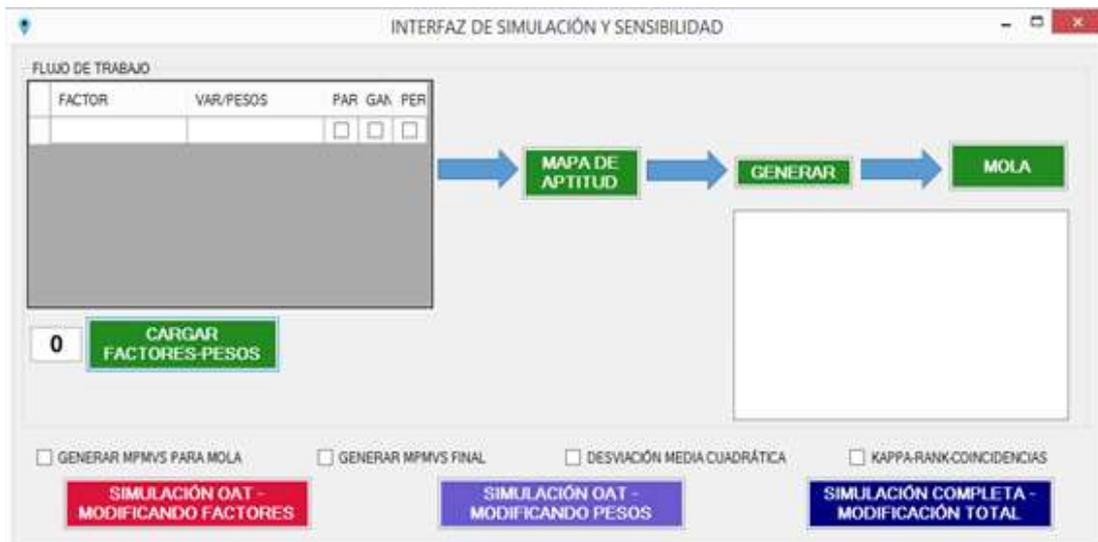


Figura 23.- Interfaz del sistema desarrollado para simular escenarios.



Figura 24.- Interfaz del sistema desarrollado para selección aleatoria de píxeles y análisis por entorno individual y múltiple.

6.2 Análisis de RANK y RANDOM

El objetivo de “RANK” es ordenar todos los píxeles del mapa, considerando su valor de aptitud (0-255), sin embargo y dado que pueden existir muchos píxeles con la misma aptitud, este procedimiento prioriza a unos píxeles de otros, realizando una búsqueda de arriba a abajo y de izquierda a derecha, y clasificando como mejores píxeles los primeros que sean encontrados en el recorrido (ver figura 25). Posteriormente con el comando “RECLASS”, se reclasifican tantos píxeles sean necesarios con la finalidad de que sean utilizados para la simulación del escenario futuro.

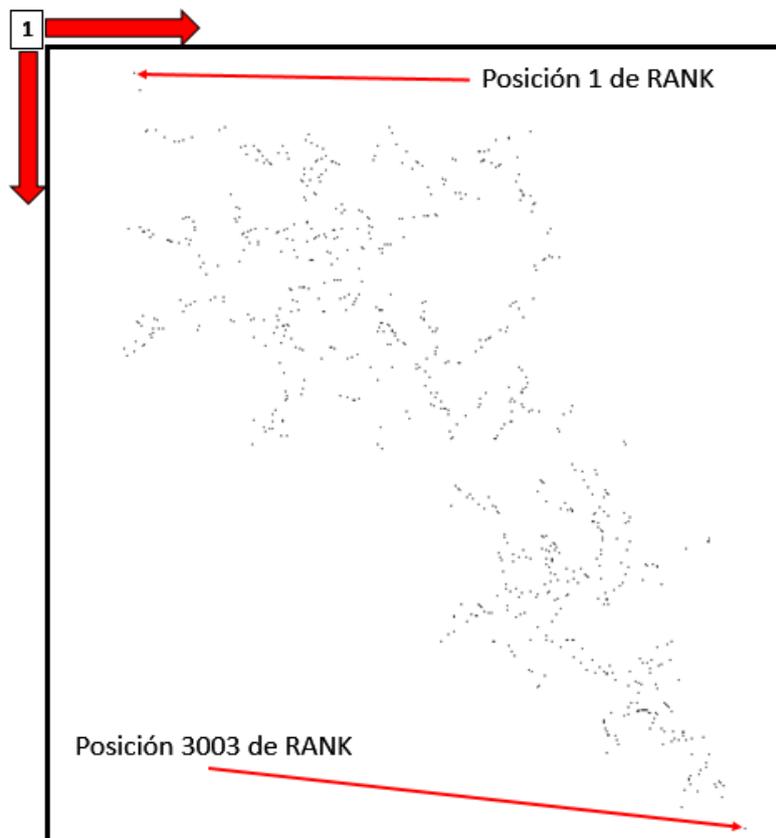


Figura 25.- Clasificación y ordenamiento realizado por el comando “RANK” de IDRISI.

Sin embargo, el uso de estos dos componentes del SIG, puede ser un problema ya que, si existieran más píxeles con un mismo valor de aptitud, como es posible dar relevancia a una cierta cantidad de píxeles y a otros no, ya que todos tendrían el mismo nivel de importancia, por otra parte, considerando la distribución espacial, los píxeles más relevantes siempre serán los de la parte superior izquierda, lo cual se considera una selección errónea desde el aspecto geoespacial.

Es por ello que una mejora a la aplicación de “RANK” y “RECLASS” es realizar una “Selección Aleatoria”, la cual contempla todos los pixeles de la misma aptitud y hasta llegar a un nivel de aptitud donde de acuerdo a la selección solo pueden ser tomados en cuenta algunos pixeles del total, en esta decisión es donde una selección aleatoria desempeñaría un mejor trabajo y mejoraría estadísticamente, desde el punto de vista espacial la selección contemplaría pixeles de cualquier parte del mapa y no solamente de la parte superior izquierda. Con lo que colocan todos los pixeles nuevamente al mismo nivel de preferencia.

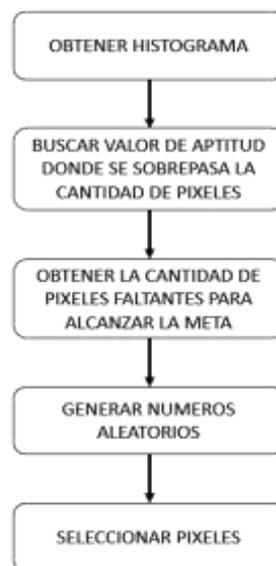


Figura 26.- Metodología para selección aleatoria.

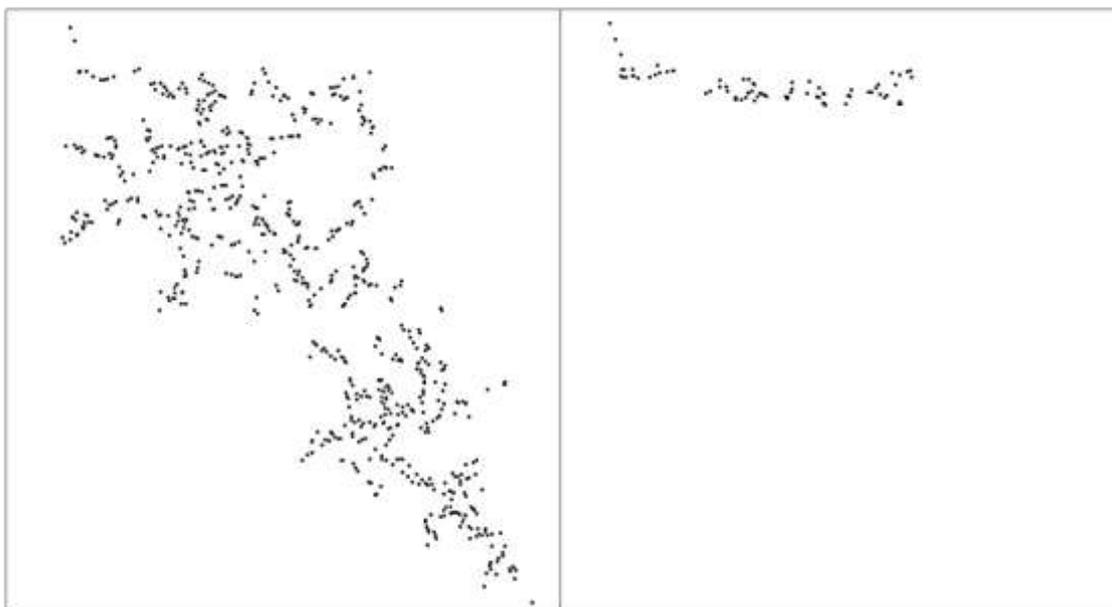


Figura 27.- A la izquierda pixeles con el mismo nivel de aptitud, a la derecha 300 pixeles seleccionados con Rank y Reclass.

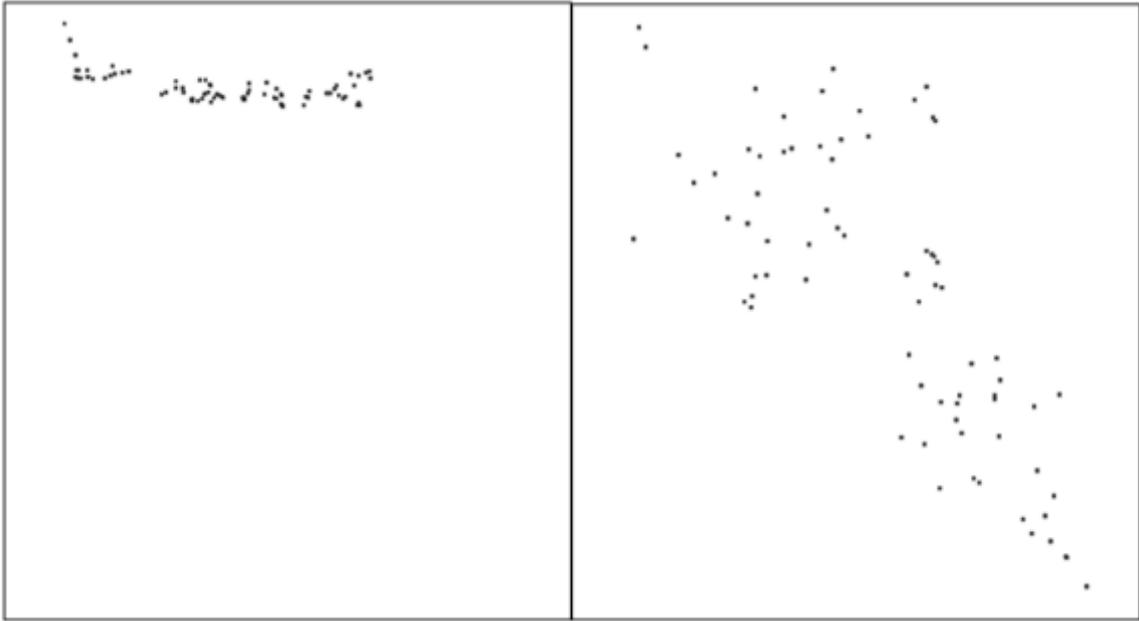


Figura 28.- Comparación de una selección por medio de RANK-RECLASS (izq) y la propuesta aleatoria (der) a partir de la imagen mostrada (arriba).

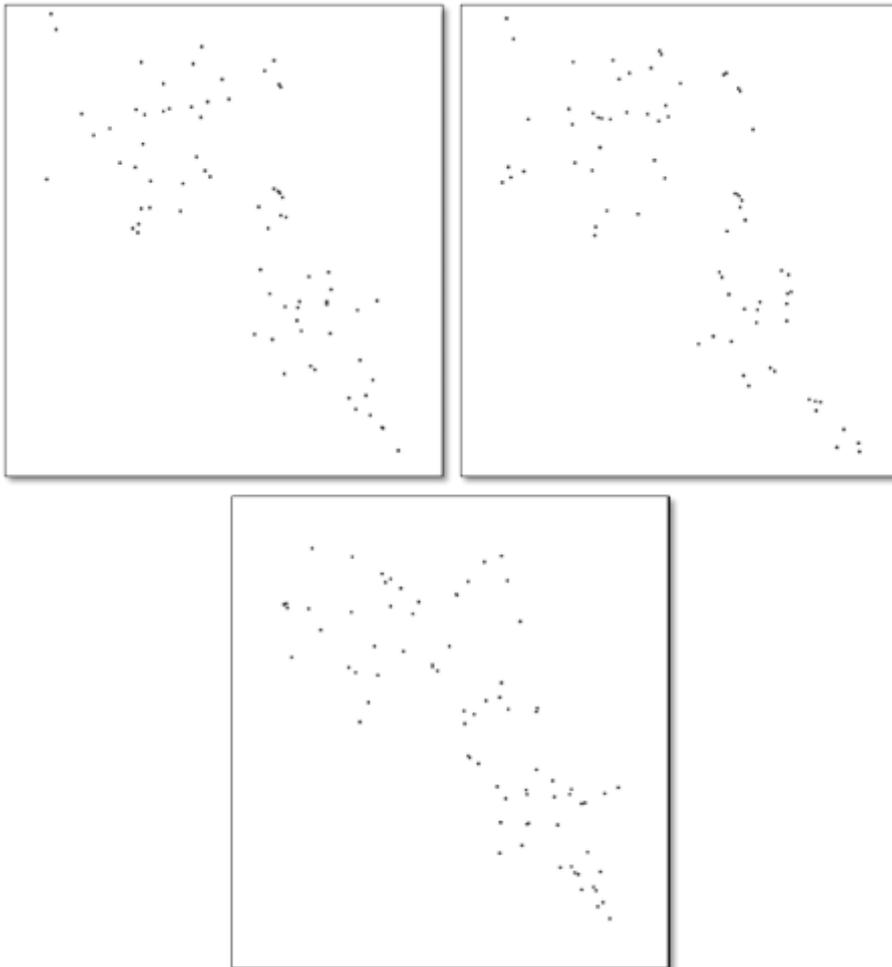


Figura 29.- Diferentes selecciones aleatorias.

6.3 Resultados del modelo de EMC

Considerando un modelo de evaluación multicriterio, basado en la técnica de Sumatoria Lineal Ponderada, se realizó el proceso de simulación de múltiples escenarios modificando, espacialmente, factores y pesos de manera aleatoria y con una variabilidad de +/- 25% por pixel.

Sin embargo, la primera consideración para aplicar lograr calcular la idoneidad aplicando la SLP fue que el peso era distinto en cada uno de los pixeles, entonces fue necesario convertir la fórmula propuesta por Gómez y Barredo (2005):

$$Aptitud_i = \sum_{j=1}^n Factor_{i,j} * peso_j \quad (43)$$

Para que considerara pesos modificados aleatoriamente en cada celda, quedando de la siguiente forma:

$$Aptitud_i = \sum_{j=1}^n Factor_{i,j} * peso_{i,j} \quad (44)$$

Por otra parte, dos técnicas de simulación fueron usadas para generar los mapas de aptitud la técnica de OAT (para pesos y factores) y las combinaciones sin repetición, teniendo como resultados 12 y 4095 mapas respectivamente.

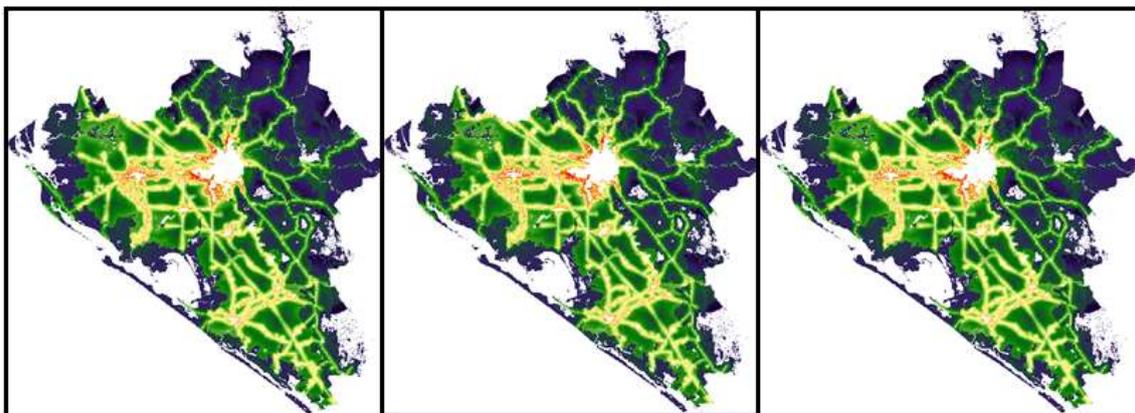


Figura 30.- Mapas de aptitud con la técnica OAT, factor 1 modificado (izquierda), factor 2 modificado (centro) y peso 1 modificado (derecha).

Una vez obtenidos los mapas resultantes, con alguna de las dos técnicas, es necesario definir el tipo de selección con la que se obtendrán el número total de píxeles idóneos. Los tipos son normal y "Aleatoria". El tipo normal es la selección que se lleva a cabo por comandos contenidos en el SIG IDRISI, sistema que es usado generalmente en investigaciones donde se manejan datos de entrada y salida de tipo raster (Joerin et al., 2001; Malczewski 2004; Plata et al., 2013; Anderson y Dragicevic 2016), este paso se realiza utilizando dos comandos: "RANK" y "RECLASS".

Una vez realizados ambos procedimientos, se calcularon dos mapas de píxeles más veces seleccionados (frecuencia), uno utilizando las salidas de OAT y el otro con los resultados de la combinatoria. En donde se puede observar en cuantos mapas aparece cada uno de los píxeles de la imagen espacial.

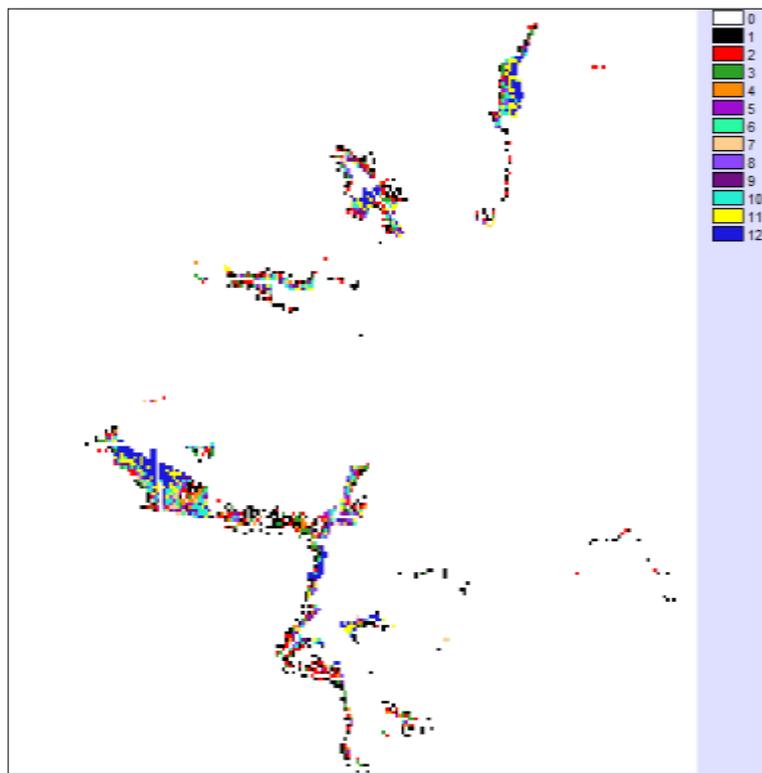


Figura 31. - Mapa de píxeles más veces seleccionados (frecuencia) utilizando resultados del método OAT (12 mapas).

En la tabla 5 se puede observar la cantidad de píxeles que comprende cada una de las categorías del MPMVS, donde 380 píxeles solo aparecen, en esa misma posición, en solo uno de los 12 escenarios para la selección con RANK mientras

que 395 aparecen para la selección RANDOM, además de notarse que 195 celdas aparecen en la misma posición en todas las simulaciones, por lo que el 13.7% de pixeles es estable a las variaciones de las entradas.

Tabla 5.- Muestra la diferencia en pixeles de las diferentes selecciones.

Clase	# Pixeles	
	Rank	Random
1	380	394
2	188	187
3	93	93
4	50	44
5	39	45
6	50	38
7	56	64
8	61	65
9	83	78
10	93	101
11	135	132
12	195	192
Total	1423	1433

Por otro lado, con la selección aleatoria, la diferencia en el total de pixeles solo varia en 10 pixeles. Se deberían de seleccionar las categorías de la 5-12, conteniendo los mejores 676 pixeles.

Se analizaron dos tipos de clasificaciones diferentes de las 12 categorías, divididas en 4 y 5 partes iguales (cuartiles y quintiles).

Tabla 6.- Muestra la cantidad de pixeles que contiene cada rango de cuartil y quintil para las selecciones Rank y Aleatoria.

	Cuartiles			Quintiles		
	Rango	Rank	Random	Rango	Rank	Random
1	1-3	661	674	1-2	568	581
2	3-6	139	127	2-4	143	137
3	6-9	200	207	4-6	89	83
4	9-12	423	425	6-9	200	200
5				9-12	423	425

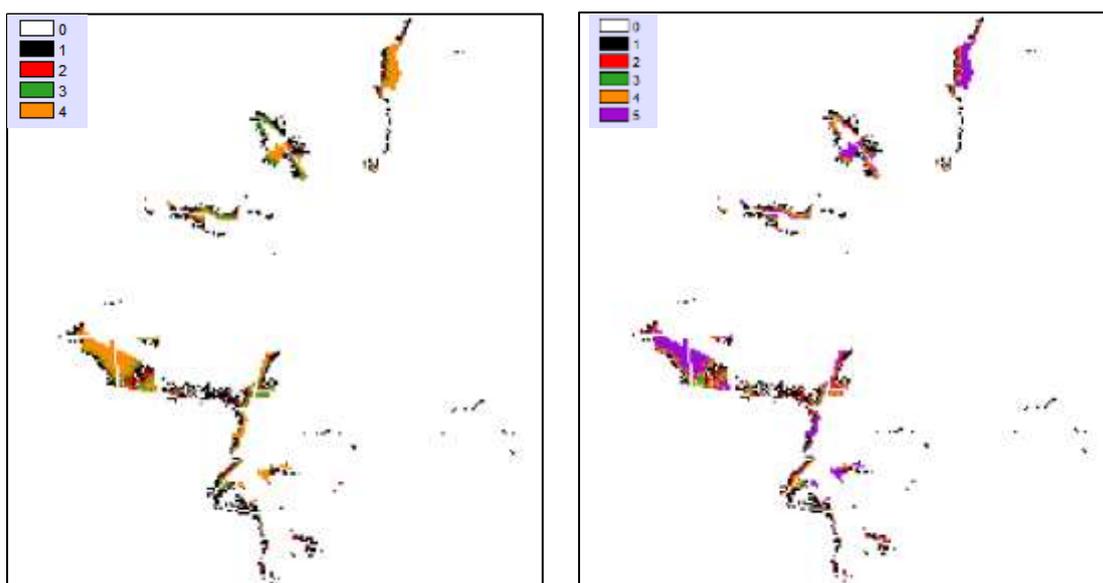


Figura 32.- Clasificación de pixeles OAT en cuartiles y quintiles.

Se calcularon las coincidencias de pixeles entre las dos diferentes tipos de selección y la selección realizada en el escenario base mostrando que, en la simulación OAT, para la primera categoría de cuartil existen solo 2 celdas coincidentes, en el caso de la categoría 2 se observaron 63 y 57 coincidencias para Rank y Aleatorio respectivamente. Pasando al cuartil 3, un total de 188 pixeles seleccionados por Rank y 192 seleccionados Aleatoriamente aparecen en posiciones de la simulación inicial, por otro lado en el último cuartil se

observan 423 celdas en la categoría Rank y solo un pixel Aleatorio de diferencia (424).

En el caso de los quintiles se puede identificar que el primero de ellos, contiene solamente 1 pixel en ambas selecciones, mientras que en el segundo cuartil se encontraron 6 y 5 coincidencias en Rank y Aleatorio respectivamente, por otra parte el tercer cuartil contiene 58 celdas para Rank y 53 para la selección Aleatoria, el cuarto cuartil lo integran 188 y 192 pixeles para cada tipo de selección, mientras que por una celda difieren ambas selecciones en el cuartil 5.

Una vez definidos los pixeles totales para cada clasificación, se determinó que el 90% de los pixeles que coinciden con la simulación inicial, se encuentran en los cuartiles 3 y 4 y los quintiles 4 y 5.

Tabla 7.- Muestra el total de coincidencias en cada rango de la clasificación entre las diferentes selecciones y el escenario base.

	Coincidencias Cuartiles		Coincidencias Quintiles	
	Base vs Rank	Base vs Random	Base vs Rank	Base vs Random
1	2	2	1	1
2	63	57	6	5
3	188	192	58	53
4	423	424	188	192
5			423	424
	90%	91%	90%	91%

En este sentido se realizó una comparación de las selecciones (RANK y RANDOM) que se obtienen con el método OAT, lo que arrojó una diferencia significativa de 4%, 38 pixeles, del total de pixeles, esto es un 880 pixeles para OAT-RANK y 918 pixeles para OAT-RANDOM. Esto significa que es posible con el método RANDOM considerar mas pixeles y con ello valorar otras zonas del

mapa que no son observadas, además se puede resaltar que aunque son pocos pixeles estos están distribuidos uniformemente.

Una vez obtenidos los 4095 escenarios a partir de la combinación (combinatoria sin repetición) de pesos y factores modificados se realizaron diferentes análisis que describimos a continuación. Es importante mencionar que para realizar este procedimiento combinatorio fue necesario considerar al mismo tiempo factores y pesos como elementos combinables (originales y modificados).

Un estudio antes de realizar las combinaciones totales, fue obtener una muestra de 10 mapas por cada combinatoria (factores modificados al mismo tiempo), 10 de factor a la vez, 10 con dos factores a la vez y así sucesivamente hasta obtener 110 mapas finales y con ellos construir de la misma manera que los anteriores un mapa de pixeles más veces seleccionados.

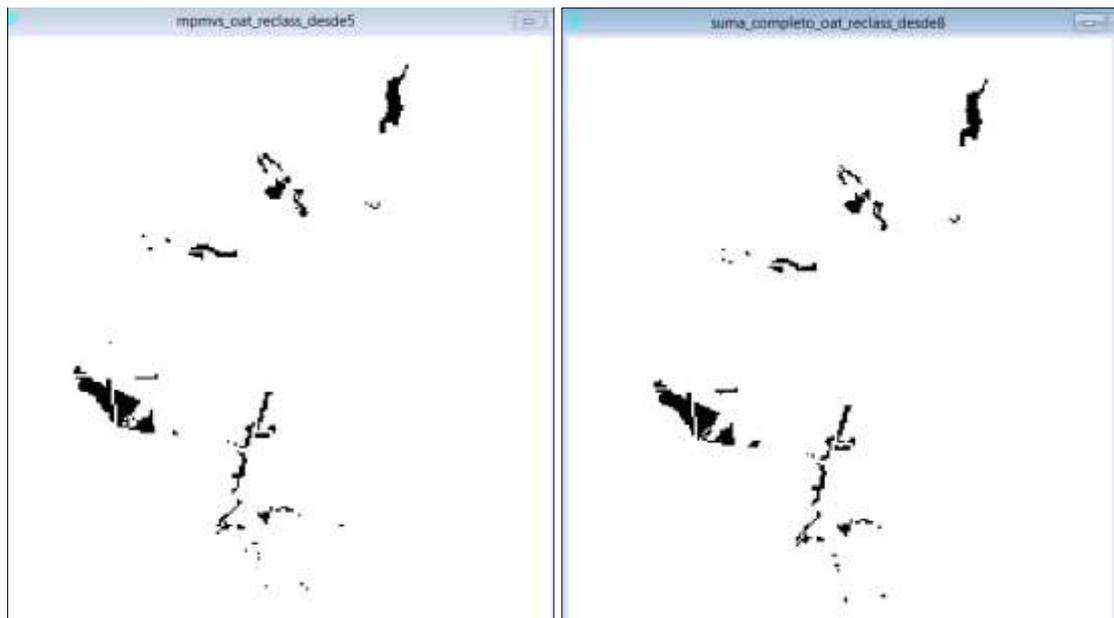


Figura 33.- Diferencia entre las selección de 676 pixeles considerando dos diferentes metodologías (Pixeles más veces seleccionados y suma de aptitudes).

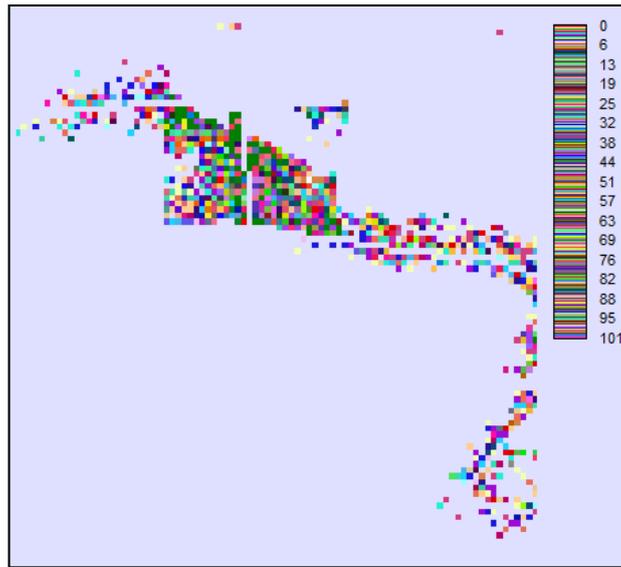


Figura 34.- Mapa de pixeles más veces seleccionados (frecuencia) utilizando solo los resultados de 110 mapas del proceso combinatorio.

La construcción del MPMVS, se llevó a cabo trasladando cada uno de los mapas binarios que contienen la selección de 676 pixeles de cada simulación. El resultado es un mapa donde el o los pixeles con mayor aparición en los diferentes escenarios es de 3763.

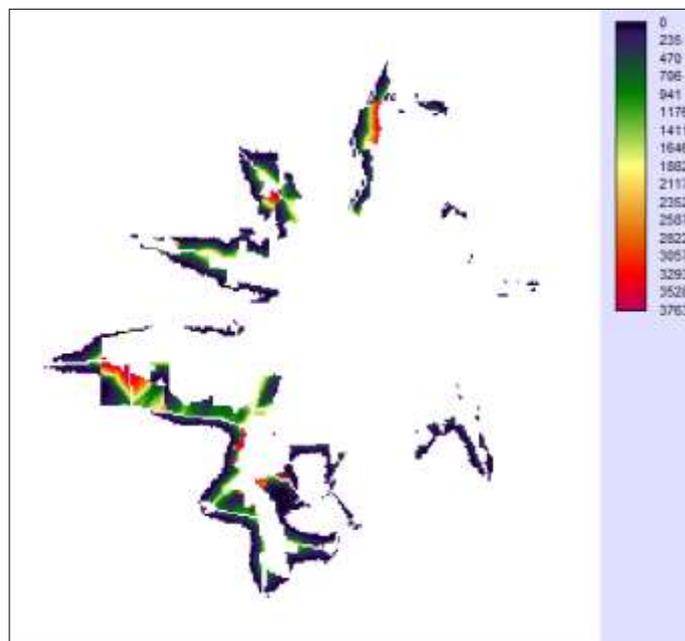


Figura 35.- Mapa de pixeles más veces seleccionados (frecuencia) utilizando resultados del método combinatorio (4095 mapas).

Se puede observar en la siguiente tabla, que al considerar los mejores pixeles (de acuerdo a su frecuencia) de cada uno de los mapas tenemos que para el

mapa de frecuencia de 12 escenarios, si se seleccionan las ultimas 5 categorías (frecuencia de 8, 9, 10, 11 y 12) se tienen 584 que equivale al 86% del total de superficie necesaria, por otro lado en para el caso de la imagen con 101 simulaciones, seleccionando solo las ultimas 3 categorías (frecuencia de 99, 100 y 101) se tienen 115 pixeles siendo estos el 23% de los necesarios, por último para las simulaciones totales (4085 escenarios) un total de 94 pixeles que son el 14 % de la superficie total necesaria se obtiene considerando las frecuencias de 4075 a 4085.

Tabla 8.- Algunos cantidades y porcentajes de pixeles correspondientes a las frecuencias en los mapas (12, 101 y 4085 mapas).

OAT			Muestra de 101 mapas			Combinaciones Totales		
#PMVS	#Píxeles	%	#PMVS	#píxeles	%	#PMVS	#Píxeles	%
1	247	21	1	244	10	1	244	10.3
2	140	12	2	164	7	3	164	6.9
3	70	6	3	126	5	2	126	5.3
4	46	4	4	88	4	4	88	3.7
5	44	4	5	85	3	5	85	3.6
6	34	3	6	84	3	6	84	3.5
7	39	3
8	40	3
9	50	4
10	68	6	99	20	8	4083	2	0
11	68	6	100	20	8	4084	3	.1
12	358	30	101	115	4	4085	78	3.0
Total	1204	100	Total	2363	100	Total	2382	100

En otro orden de ideas, se realizó un estudio de costo computacional para los escenarios a simular. Para ellos se utilizaron diferentes resoluciones espaciales, basadas en el mapa original y duplicado su tamaño al doble, esto es dividiendo cada pixel en 4 pixeles como se ve en la figura 19. Los costos como es de esperar aumenten considerablemente (ver tabla 9). El objetivo del estudio es conocer los costos necesarios si el o los mapas de entrada son más grandes que los considerados para este estudio.

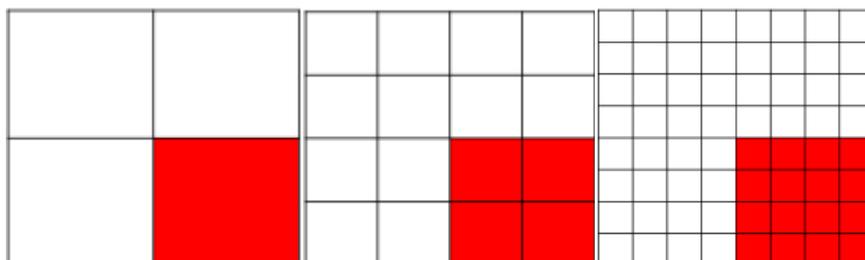


Figura 36.- Visualización del aumento en la resolución del mapa base.

Con los resultados obtenidos hasta el momento, es necesario analizar los mapas de frecuencias para conocer más a detalle las posibles inferencias de los factores y pesos de entrada así como, la manera en la que afectan estos resultados al modelo, con la finalidad de mejorar las simulaciones en ejecuciones posteriores.

Tabla 9.- Aumento en la resolución de la imagen y el tiempo de computo requerido para su procesamiento.

TAMAÑO	TOTAL DE PÍXELES	TIEMPO
(3x3) 1274 x 1379	1'756,846 píxeles	8.0528 seg.
(5x5) 2548 x 2758	7'029,932 píxeles	14.6646 seg.
(7x7) 5096 x 5516	28'109,536 píxeles	54.7614 seg.

En ninguna investigación trabajan con imágenes tan grandes, la más grande encontrada es de 747*563 (Ligmann-Zielinska & Jankowski, 2014).

6.4 Comparación de variaciones con escenario base

Dándole la importancia que merece al escenario base que es aquel donde no se alterado ningún factor ni peso, se llevó a cabo una comparación entre las

modificaciones OAT para analizar el impacto de las modificaciones aplicadas por el método y además poder identificar cual o cuales componentes de entrada aportan mayor sensibilidad al modelo.

En las figuras 55, 56 y 57 se pueden contemplar donde las celdas seleccionadas, según sea el caso, coincidan espacialmente a través de un traslape de mapas, identificando dos tipos de celdas, las que coincidan con la simulación base y las que no con valor de 1 y 2 respectivamente. Los resultados muestran que las modificaciones a los factores de proximidad a asentamientos humanos y proximidad inversa a usos de suelo y vegetación son los que aportan mayor sensibilidad al modelo, obteniendo un 60% y 73.6% de concordancia, respectivamente, en comparación del escenario original; lo que demuestra la relación directa del valor de sus pesos y la sensibilidad adherida al proceso de simulación. Los resultados para área urbanizable programada, proximidad a carreteras, pendiente, proximidad a ríos fueron 96%, 79%, 95.5% y 97% respectivamente (tabla 10).

Sin embargo, dos factores tiene un peso de 0.2, la proximidad inversa a USyV y la proximidad a carreteras, pero este último no tiene el mismo impacto en el modelo, ya que el resultado obtenido es de 79%.

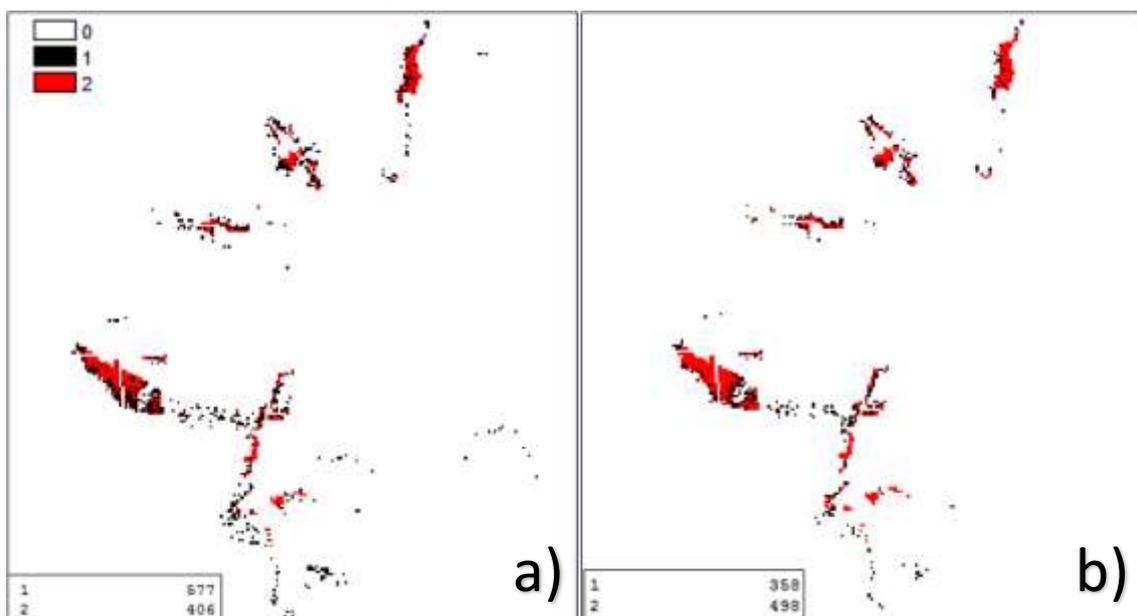


Figura 37.- Comparación de escenario base y la selección realizada a partir de la modificación del factor a) proximidad a asentamientos humanos y b) proximidad inversa a USyV.

Una vez seleccionados los mejores pixeles de los 12 escenarios, se puede realizar un análisis de entorno en el cual se hace notar, a nivel espacial, que aquellos factores con mayor peso tienden a ser más sensibles en sus resultados y muestra de ello es la figura 41, en la cual se altera el factor de proximidad a asentamientos urbanos y esto tiende a tener un mayor porcentaje de pixeles con frecuencia 1, es decir, que solo son encontrados, en una ventana de von Neumann, una sola vez, considerando la selección del escenario base.

La misma situación se repite al analizar de la misma manera la modificación de los pesos de los factores (figuras 41-46), en la figura 47 de puede notar que los escenarios donde se modificó el peso de la proximidad a asentamientos urbanos y la proximidad a ríos dio resultados donde hay menos pixeles encontrados en una ventana de 3x3.

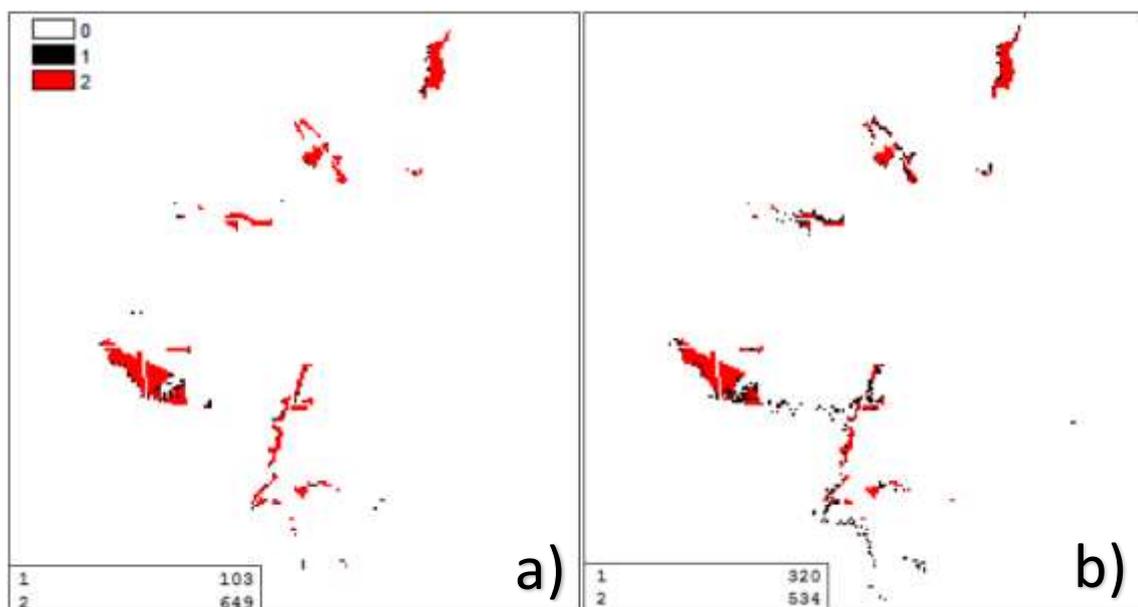


Figura 38.- Comparación de escenario base y la selección realizada a partir de la modificación del factor a) área urbanizable programada y b) proximidad a carreteras.

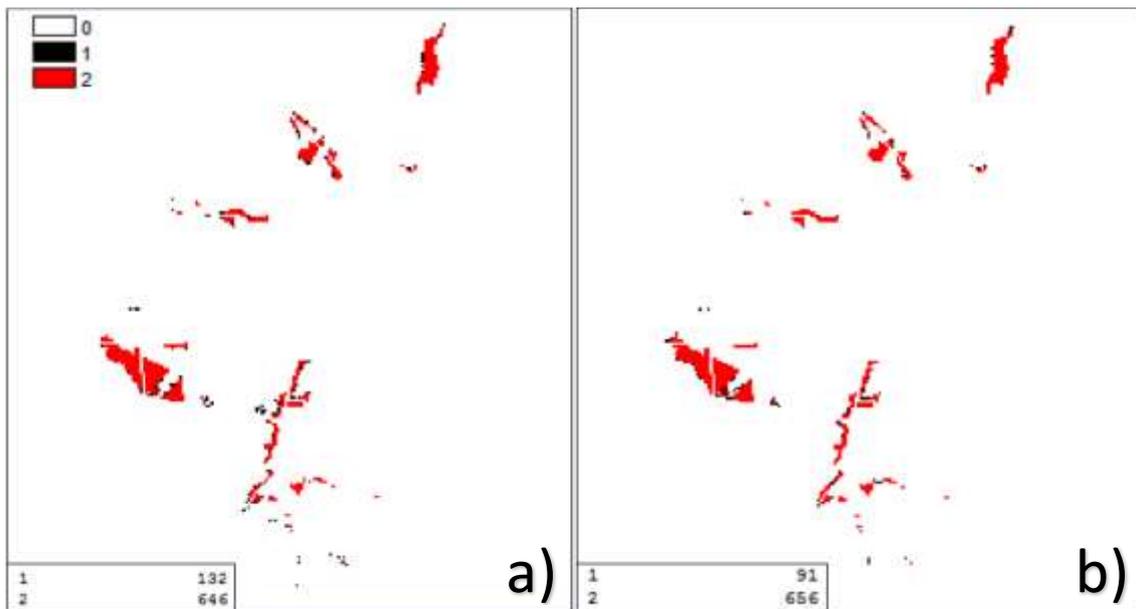


Figura 39.- Comparación de escenario base y la selección realizada a partir de la modificación del factor a) pendiente y b) proximidad a ríos.

En este sentido, se analiza el impacto de las alteraciones en los pesos de los factores de entrada del modelo, lo que da como resultado que también los pesos de los factores de proximidad a asentamientos humanos y proximidad inversa a usos de suelo y vegetación tienen aportar mayor sensibilidad que los otros pesos, con una concordancia de 59.7% y 73.9% contra los resultados base; en comparación con los pesos del área urbanizable programada, proximidad a carreteras, pendiente, proximidad a ríos que tienen un porcentaje de acuerdo de 94.5%, 77.8%, 92.4%, 96.3% (tabla 10).

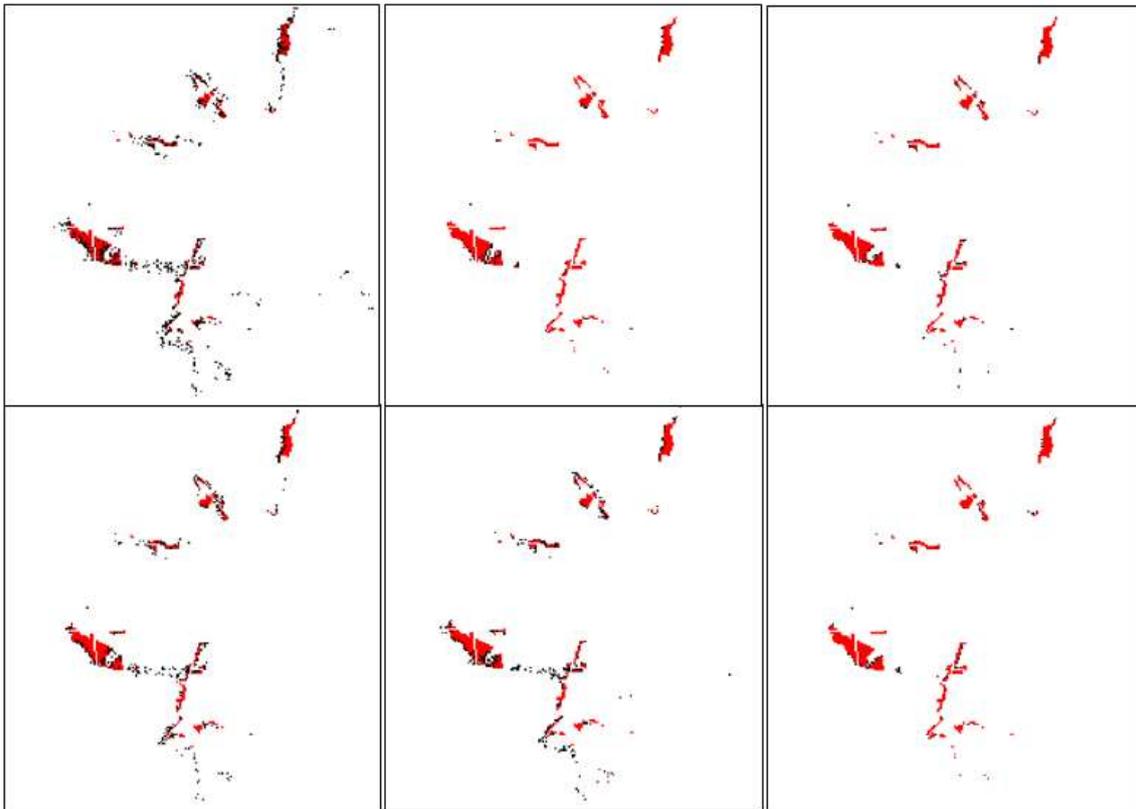


Figura 40.- Comparación de escenario base y la selección realizada a partir de las modificaciones al peso de cada factor.

Tabla 10.- Concordancia entre factores y pesos de factores modificados y resultados escenario base.

Modificación	% de acuerdo
Proximidad a asentamientos humanos	60
Proximidad Inversa a US y V	73.6
Área urbanizable programada	96
Proximidad a carreteras	79
Pendiente	95.5
Proximidad a ríos	97
Peso variable 1	59.7%
Peso variable 2	73.9%
Peso variable 3	94.5%

Peso variable 4	77.8%
Peso variable 5	92.4%
Peso variable 6	96.3%

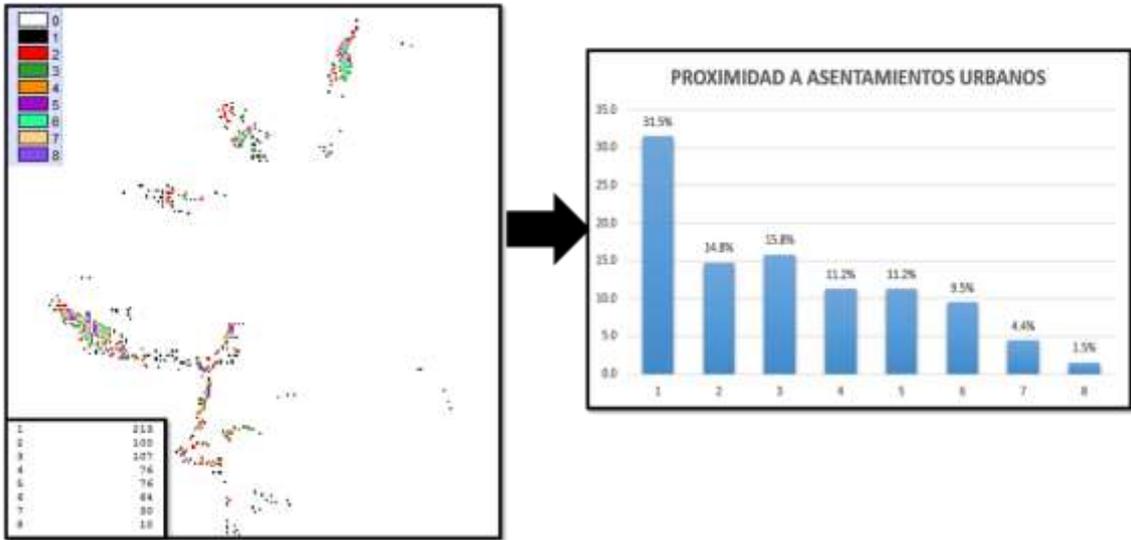


Figura 41.- Escenario modificando el factor 1 y grafica de análisis de entorno con el escenario base.

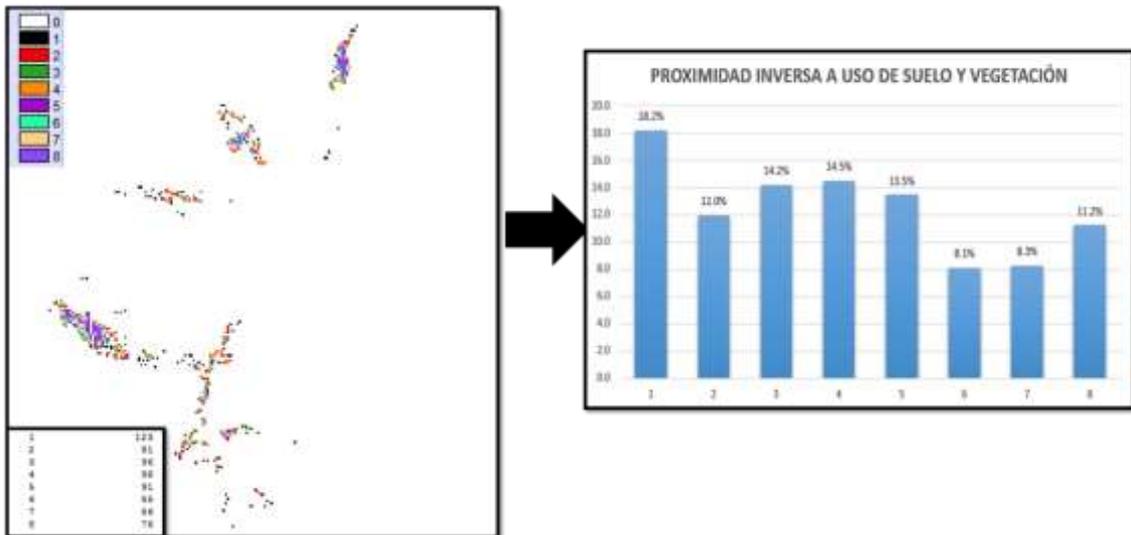


Figura 42.- Escenario modificando el factor 2 y grafica de análisis de entorno con el escenario base.

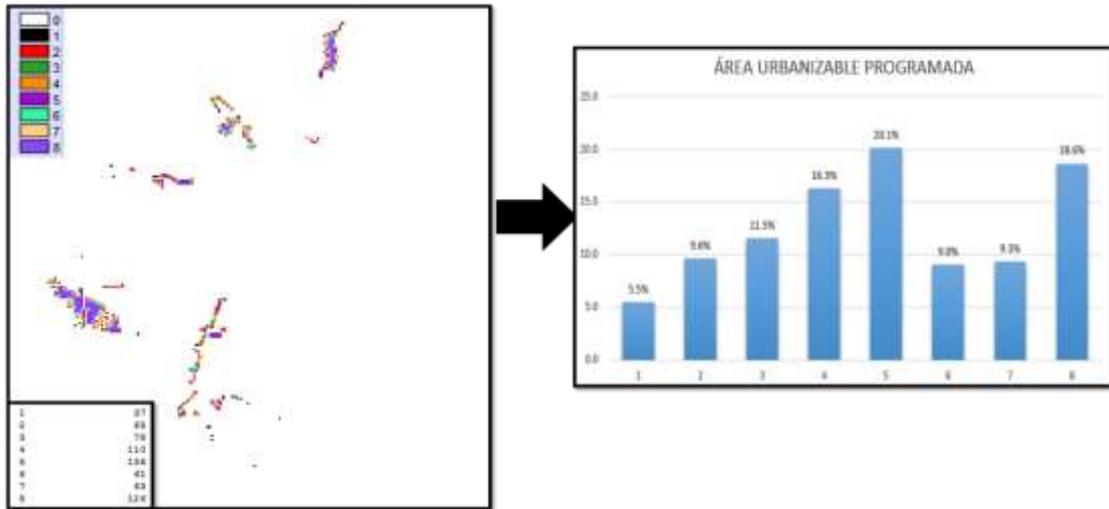


Figura 43.- Escenario modificando el factor 3 y grafica de análisis de entorno con el escenario base.

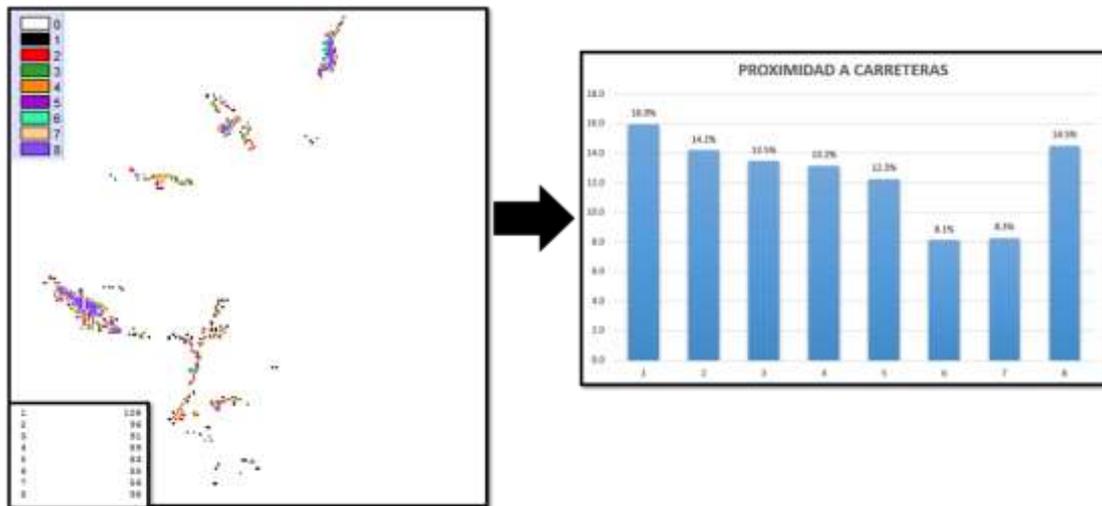


Figura 44.- Escenario modificando el factor 4 y grafica de análisis de entorno con el escenario base.

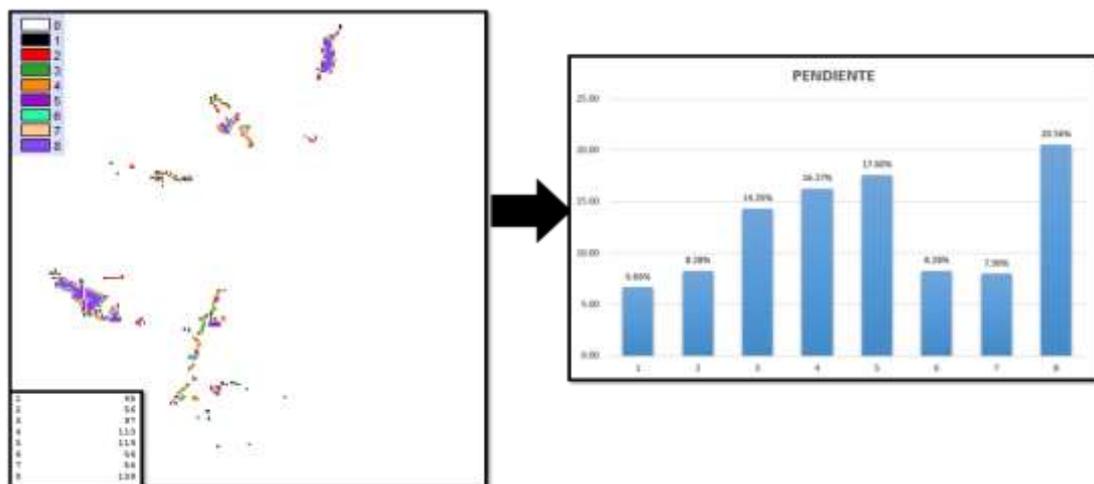


Figura 45.- Escenario modificando el factor 5 y grafica de análisis de entorno con el escenario base.

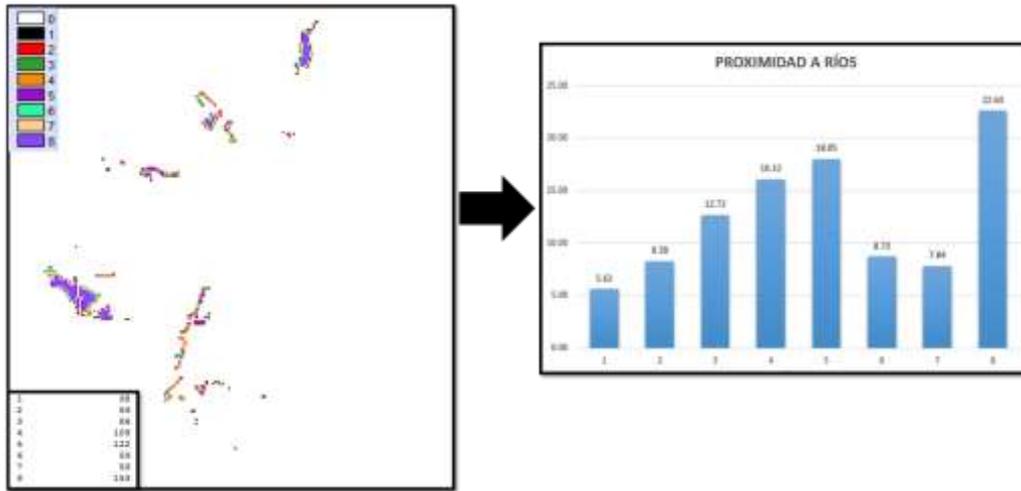


Figura 46.- Escenario modificando el factor 6 y grafica de análisis de entorno con el escenario base.

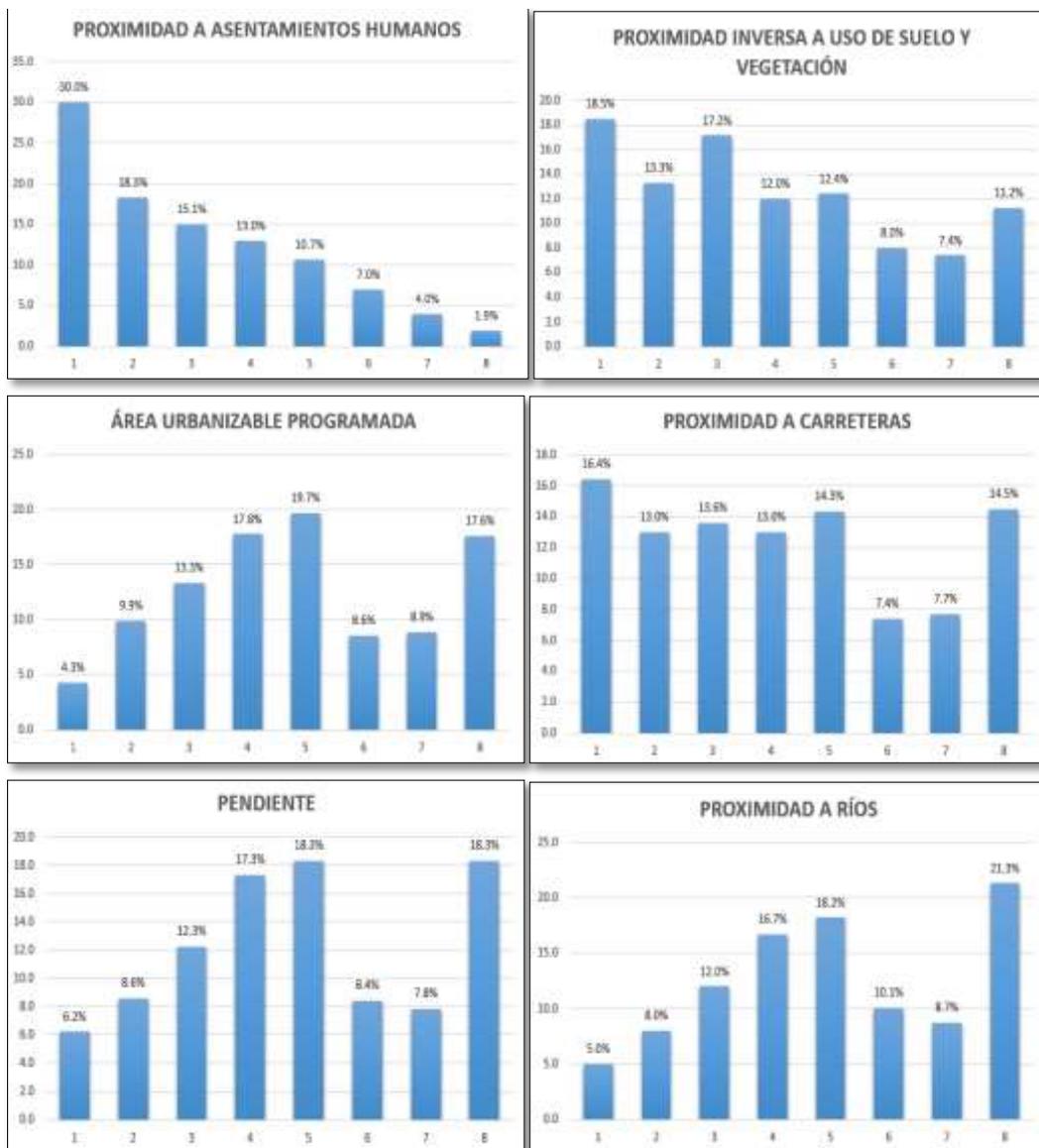


Figura 47.- Graficas de análisis de entorno de las modificaciones de los pesos de factores.

En la figura 48 se puede analizar que si se modifican todos los factores y pesos del modelo al mismo tiempo, y se utiliza una ventana de 3x3 de von Neumann los pixeles con mayor frecuencia disminuye drásticamente, y eso es lógico debido a la unión de todas las variaciones.

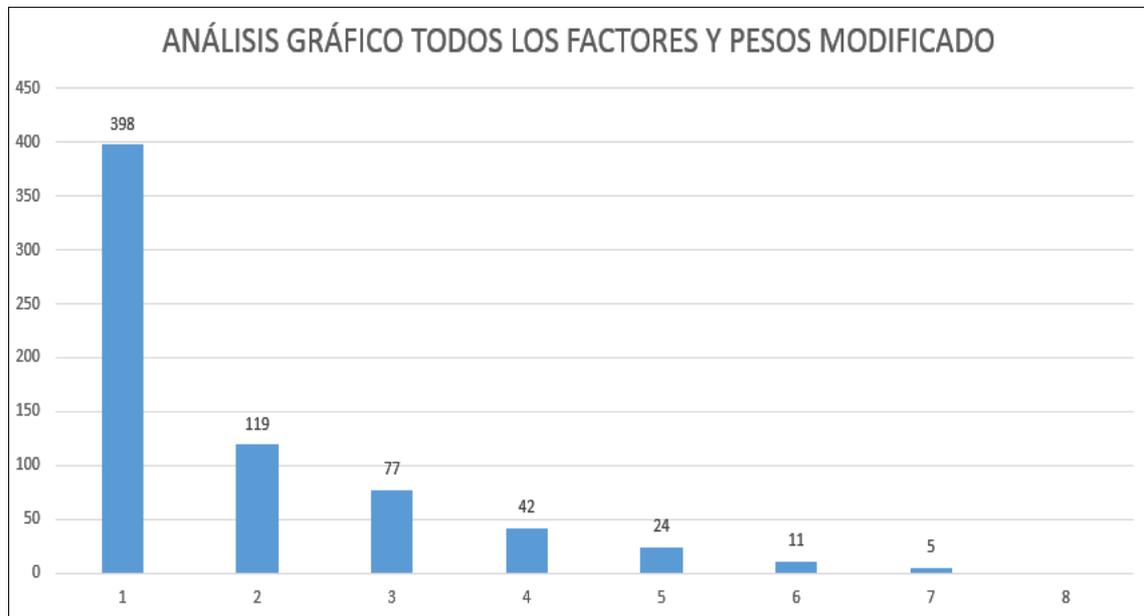


Figura 48.- Graficas de coincidencia modificando todos los factores y pesos a la vez.

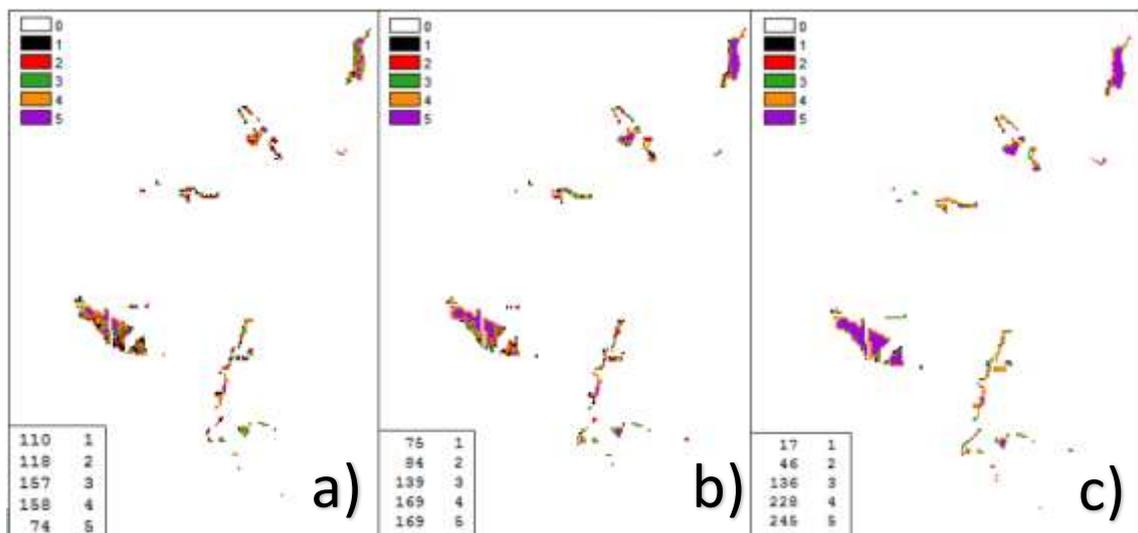


Figura 49.- Análisis de entorno de la selección realizada a partir de la modificación del factor a) proximidad a asentamientos humanos, b) proximidad inversa a USyV y c) área urbanizable programada.

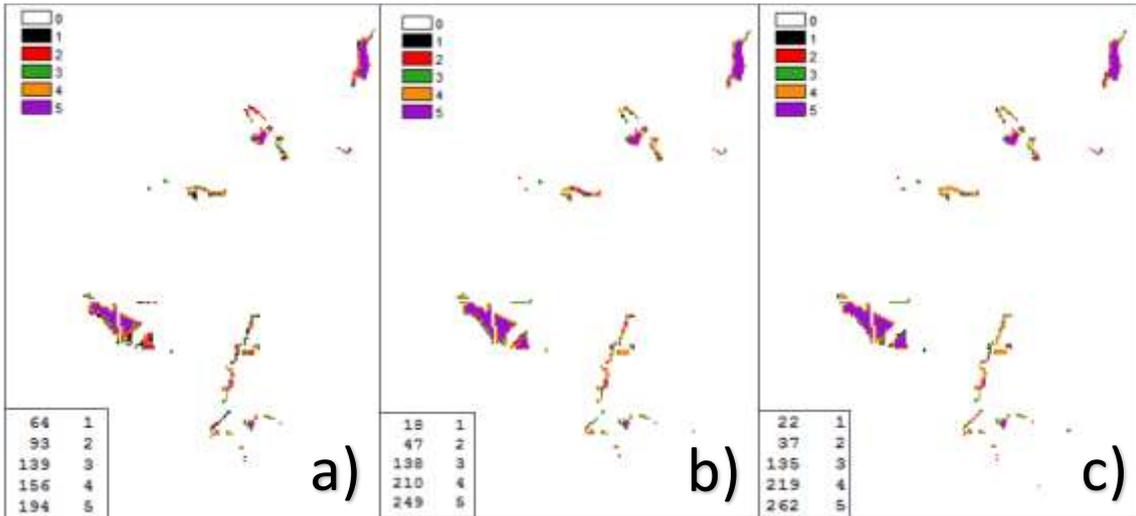


Figura 50.- Análisis de entorno de la selección realizada a partir de la modificación del factor a) proximidad a carreteras, b) pendiente y c) proximidad a ríos.

En la gráfica 51 se puede observar una prueba de análisis de sensibilidad global es decir, se analizan el número en que un pixel es encontrado en la misma posición o en una ventana de 3x3 y sumándose uno a uno, de total de 4095 escenarios.

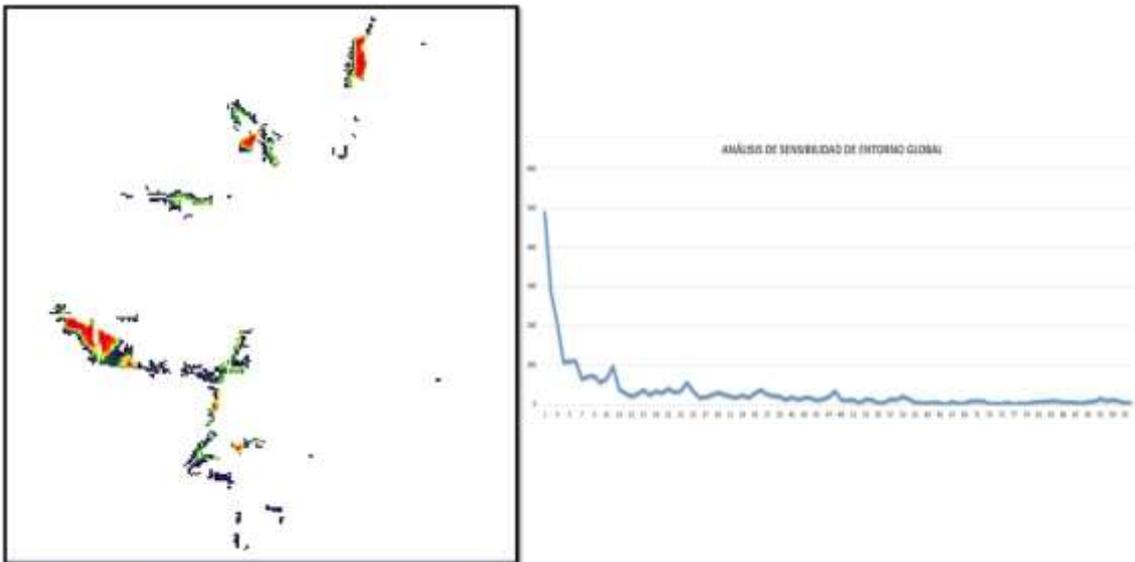


Figura 51.- Análisis de entorno global.

Realizando un análisis de entorno con ventana de 3x3 se puede observar una concordancia global, es decir analizando los 12 mapas OAT en una misma vez, de 62.4% que equivale a 422 pixeles de 3152, los cuales fueron obtenidos a

partir de seleccionar los 690 con mejor valor de idoneidad de cada escenario OAT.

Tabla 11.- Análisis de Entorno entre factores y pesos de factores modificados y resultados escenario base.

Modificación	% de acuerdo
Proximidad a asentamientos humanos	62.4
Proximidad Inversa a US y V	73.5
Área urbanizable programada	94.6
Proximidad a carreteras	77.3
Pendiente	91.86
Proximidad a ríos	95.2
Peso variable 1	60.2
Peso variable 2	75.7
Peso variable 3	93.3
Peso variable 4	77.2
Peso variable 5	92.8
Peso variable 6	95.2
Análisis Global (12 mapas)	62.4

6.5 Resultados del Análisis de Herramientas de Análisis de Sensibilidad.

Del estudio de las diferentes herramientas existentes en el estado del arte que realizan análisis de sensibilidad se obtuvieron los siguientes resultados.

Las primeras pruebas fueron realizadas a la aplicación SensIt, en las cuales se utilizaron modelos con 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15 y 20 entradas y un % de paso de 1, 5 y 10, no fue posible definir más de 1 resultado (variable de salida) dadas las características de la aplicación.

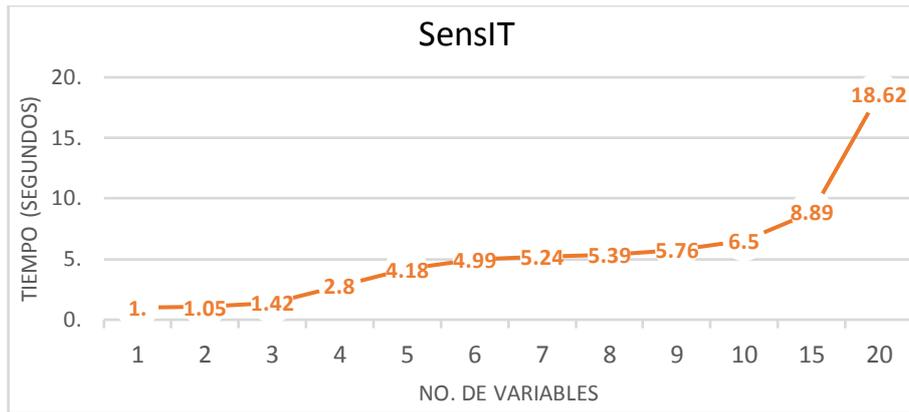


Figura 52.- Gráfica que muestra el rendimiento en segundos de modelos con n entradas y 1 salida.

De manera general las diferentes graficas analizadas muestran el mismo comportamiento, aún con la variabilidad del número de salidas.

Posteriormente en el estudio de SimLab, los modelos ejecutados se llevaron a cabo con 2 y 10 entradas, 1 y 5 salidas, 50, 100, 200, 500, 1000, 2000, 5000 y 10000 y variando además el tipo de Función de Probabilidad de Distribución (PDF) el tipo de muestreo de acuerdo al AS a realizar.

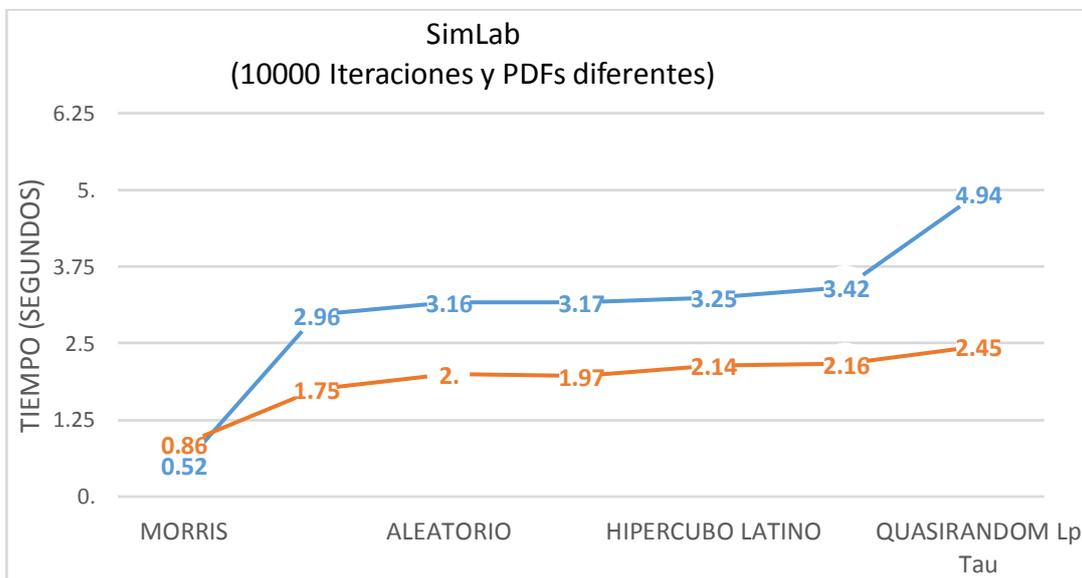


Figura 53.- Gráfica que muestra el rendimiento en segundos de modelos con 10 entradas y 1 y 5 salida, utilizando PDFs diferentes para cada entrada y realizando 10000 iteraciones.

De la misma manera que SensIt, este muestra el mismo comportamiento en las diferentes graficas analizadas, aún con la variabilidad de los factores que

componen el proceso de modelado en esta aplicación. Las problemáticas de esta aplicación primeramente es que se puede considerar una caja negra, dado que no se sabe cómo realiza los cálculos del AS, además utiliza demasiadas pantallas y hace cada procedimiento por separado lo que lo hace poco amigable con el usuario, tiene mal manejo de errores (termina su ejecución y cierra la ventana principal sin informar del error).

El análisis de rendimiento final fue para SaSAT, y se realizaron pruebas con modelos de 2 y 5 variables de entrada y 1 y 5 variables de salida, definiendo a su vez la misma cantidad de iteraciones que la prueba anterior, se variaron de la misma manera los tipos de PDFs, y Muestreos de los AS. En los casos del uso de regresión estándar se tuvo problemas al definir los límites del índice. No fue posible analizar el modelo utilizando el AS full factorial con 10 variables de entrada y 5 de salida, ya que eran demasiadas entradas y por lo tanto un gran número de iteraciones era necesario para obtener los resultados del modelo.

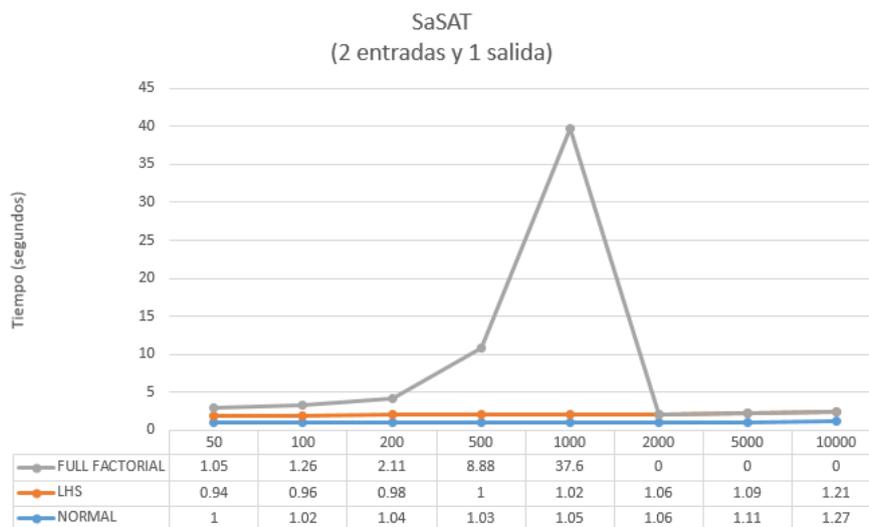


Figura 54.- Gráfica que muestra el rendimiento en segundos de modelos con 2 entradas y 1 salida, utilizando PDFs diferentes y para 3 tipos de AS contemplados.

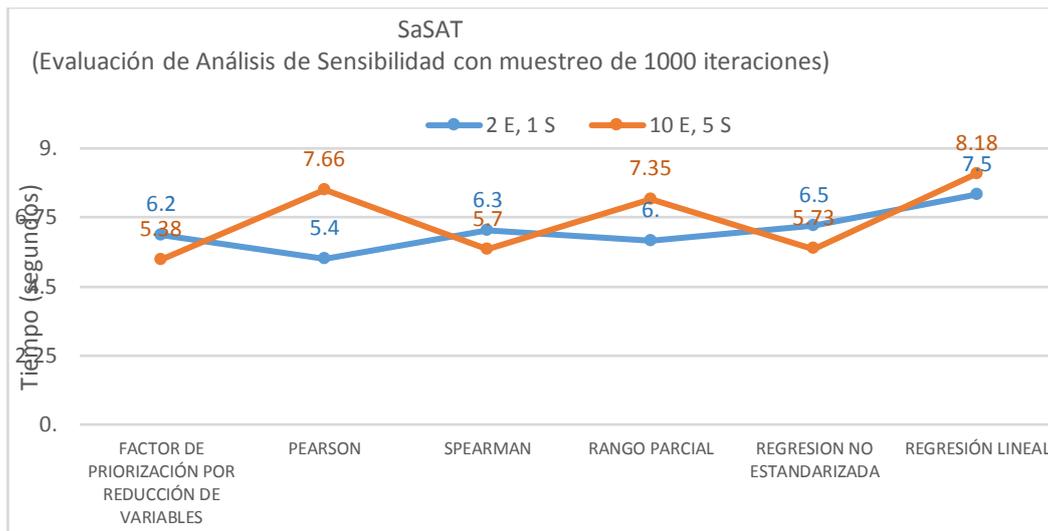


Figura 55.- Grafica que muestra el rendimiento en segundos de modelos con 2 y 10 entradas y 1 y 5 salidas, utilizando PDF normal, para los tipos restantes de AS.

Las desventajas observadas para esta herramienta fueron que tiene un mal manejo de errores, son necesarios muchos clics y cambio de pantallas, para lograr el proceso completo de AS, por otra parte una ventaja mostrada fue que es un software intuitivo.

Para obtener resultados confiables, cada modelo fue ejecutado un promedio de 3 veces, esto para cada una de las tres herramientas contempladas en este estudio. El equipo utilizado para la instalación y ejecución de los modelos, consta de las siguientes características técnicas: LANIX Neuron R II, S.O. Windows 8.1 x64, Procesador Intel Core i7 de 3.00 GHz, 8.00 GB de RAM, 1TB de disco duro.

Es importante destacar que de todas las herramientas que contempla esta investigación, excepto IntiGIS y iUSA, todas realizan análisis estadístico y los resultados solo son gráficas y por lo tanto podemos concluir que no existen resultados explícitamente espaciales (mostrados en mapa), ni tampoco es posible el uso de mapas como parámetros de entrada.

Ninguna de las herramientas analizadas está desarrollada para trabajar directamente con modelos de simulación geoespacial. Es de relevancia mencionar que ninguna de las herramientas analizadas trabaja con la

distribución espacial contenida en el mapa, por lo que se tiene certeza de que en este paso se pierde la característica puramente espacial de la información

6.6 Resultados del Análisis de Sensibilidad

Para el cálculo del índice de sensibilidad de Hoffman y Gardner (1983), se adaptó la fórmula para ser aplicada contemplando dos o más escenarios, y se describe en la figura 56.

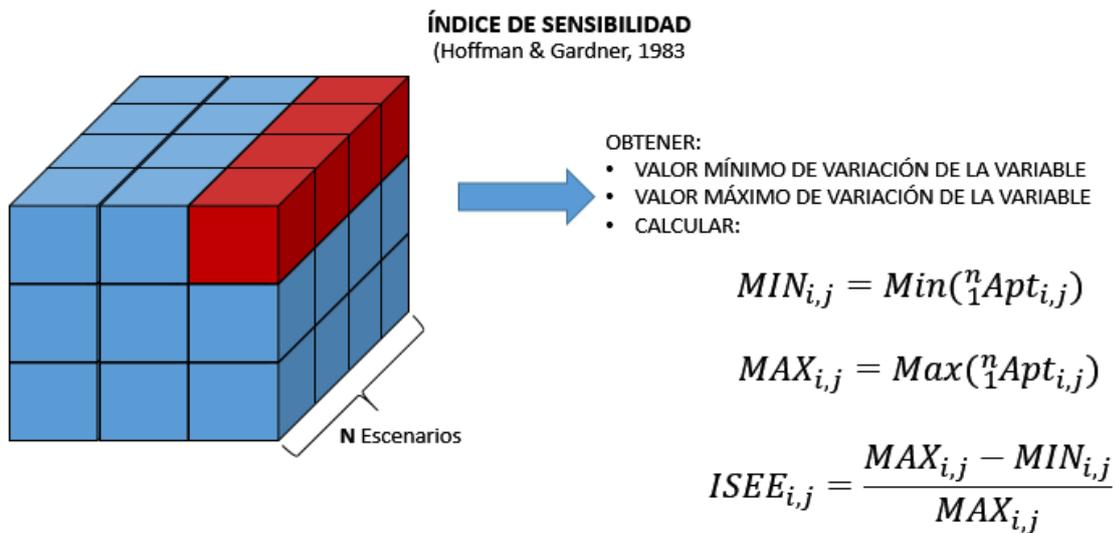


Figura 56.- Índice de sensibilidad adaptado a múltiples escenarios.

En la figura 57 se puede ver un mapa con los índices de sensibilidad por pixel, basado en los resultados de OAT (12 escenarios), mientras que en la figura 53 se puede ver la diferencia del mismo índice pero considerando las salidas del método de combinaciones totales, es decir los 4095 escenarios simulados.

Un ejemplo de comparación de los dos resultados (OAT y combinaciones totales) se puede apreciar en la figura 59 donde se toma una muestra de 3x3 pixeles de ambos mapas resultantes y se puede apreciar la diferencia de los valores entre ambos métodos. Se puede observar como en el AS OAT se tiene como mayor valor de sensibilidad el 0.139 (pixel de arriba a la derecha) y menor índice el 0.064 (pixel de abajo en medio) sin embargo, en el AS combinaciones totales se destaca que los valores con mayor y menor sensibilidad cambian, siendo para

este metodo 0.44 (arriba la izquierda) y 0.39 (abajo a la derecha) respectivamente.

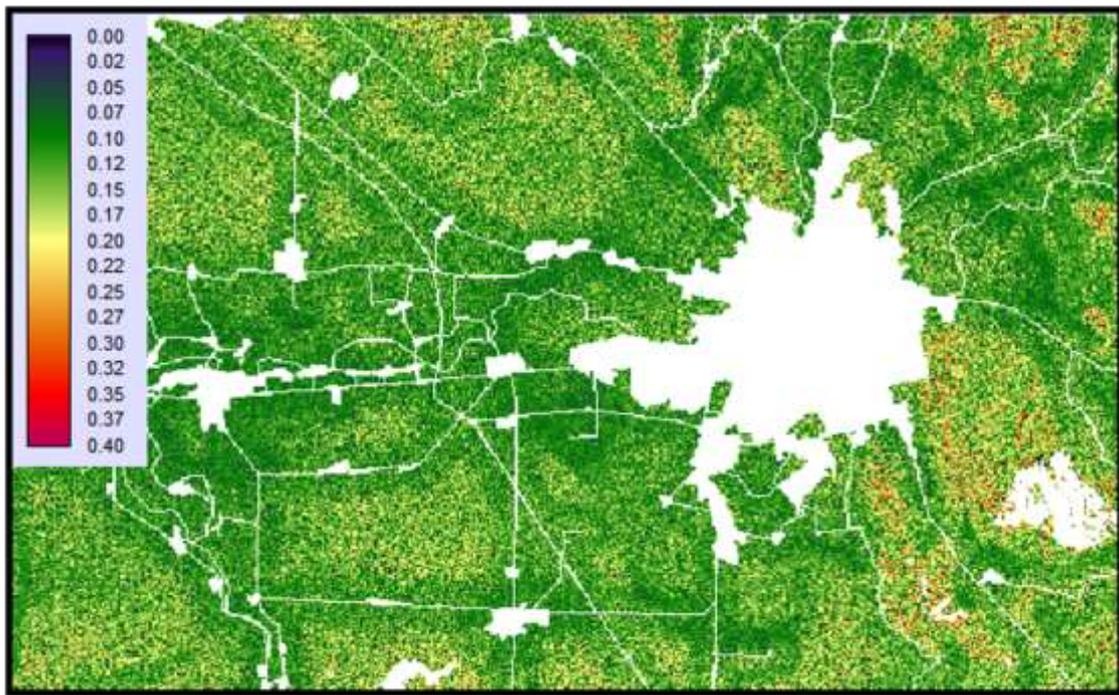


Figura 57.- Mapa de índice de sensibilidad de los resultados de OAT.

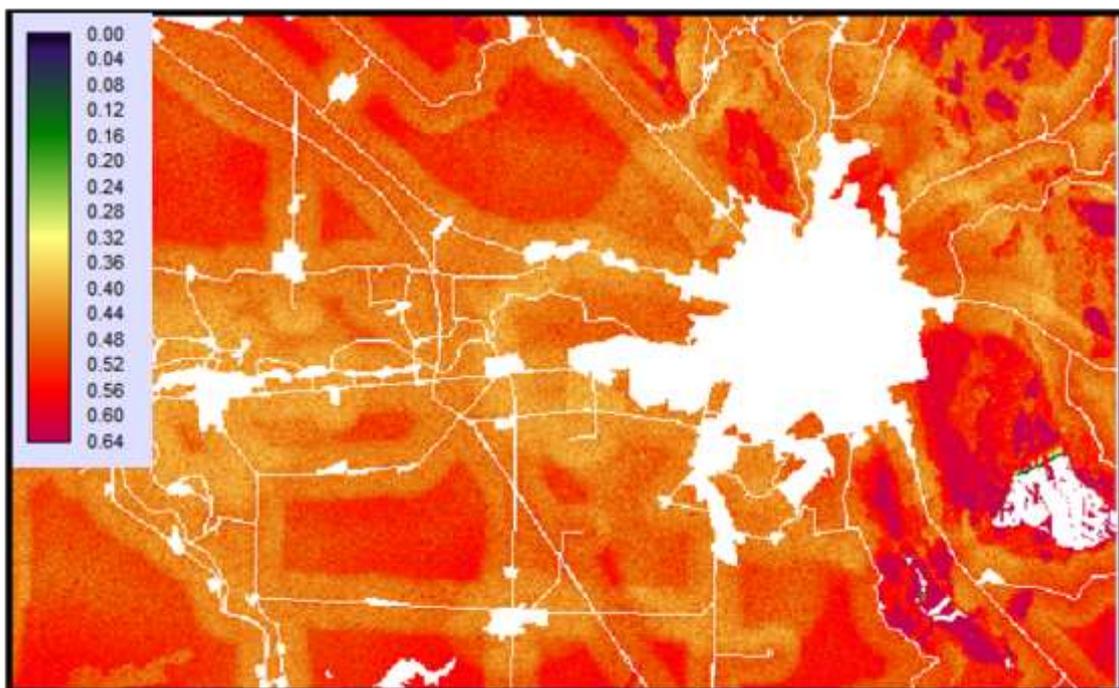


Figura 58.- Mapa de índice de sensibilidad de los resultados de las combinaciones totales.

El modelo OAT es mas rápido dado que no genera tantos modelos y sus la variabilidad de sus valores es mayor, mientras que el simulaciones totales es

mas sensible dado que contempla mas modelos, pero no varian tantos los valores de sensibilidad.

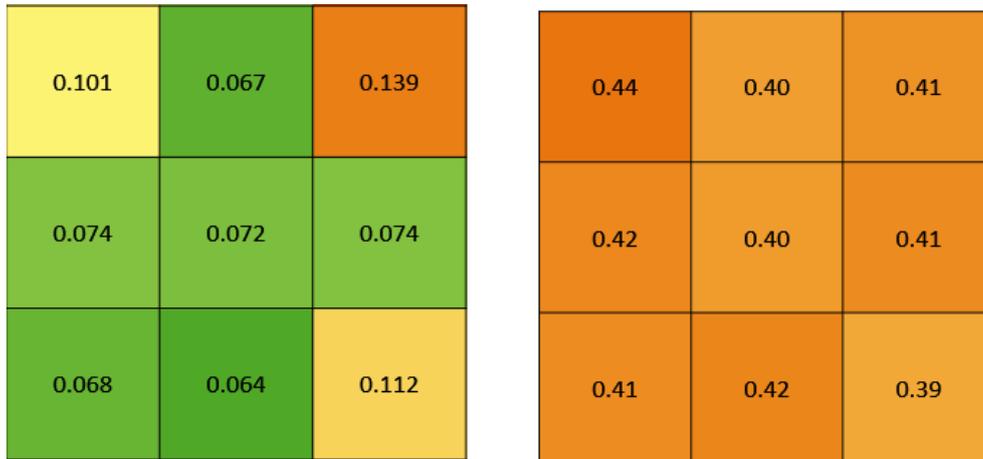


Figura 59.- Muestra del valor de índice de sensibilidad en una región de 3x3.

Considerando el índice de sensibilidad y el MPMVS, se realizó un traslape para obtener los índices y estos clasificarlos en tres categorías según su valor. Las categorías de sensibilidad generadas fueron: sensibilidad alta, baja y media. Y de las cuales se observa que existen 97 pixeles considerados con alta sensibilidad, 735 se sensibilidad media y un total de 371 celdas con baja sensibilidad. Destacando que del MPMVS un total de 31% tiene una baja sensibilidad a la variación de las entradas (ver figura 60).

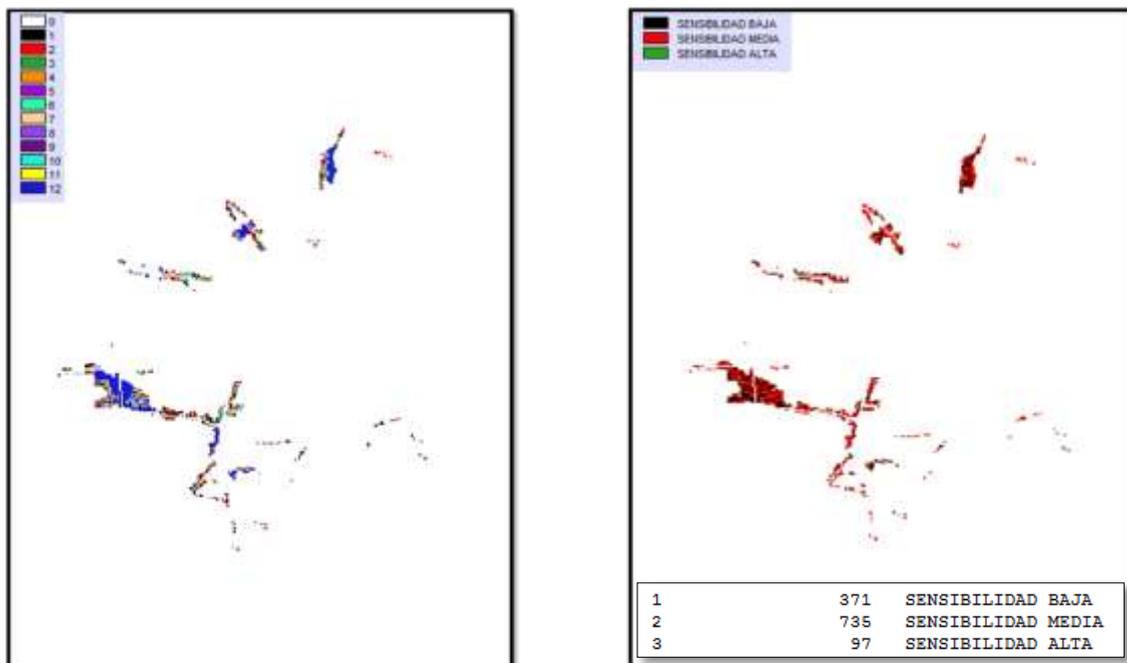


Figura 60.- MPMVA (izquierda) y mapa con categorías de alta, media y baja sensibilidad.

6.7 Comparación entre índices

Con el uso de los diferentes índices obtenidos, es posible realizar un análisis más profundo de las celdas, a seleccionar para las diferentes simulaciones y con esto conocer el comportamiento de los pixeles a las diferentes alteraciones de los factores y sus pesos. Una muestra del valor que proporcionan los diferentes índices obtenidos se puede ver en la figura 67, donde se analiza uno de las zonas simuladas que más compacidad tiene.

Cabe destacar que un resultado obtenido del análisis de los diferentes índices es que si se analizan los pixeles que se repiten muchas veces (buena frecuencia), y/o solo el método OAT no concuerda con análisis más profundo al analizar sus valores de aptitud en los 4095 mapas, lo que da un resultado distinto.

Un análisis general sobre los diferentes índices mostrados en la figura 61 es que todos los índices basados en OAT son más rápidos dado que son menos modelos que analizar, a diferencia de las combinaciones totales o los análisis de entorno, por otra parte el índice MPMVS OAT es un modelo general que puede ayudar a mostrar un idea de cómo es la sensibilidad geoespacial, sin embargo a diferencia del promedio, todos tienen cambios abruptos en los valores de los pixeles, esto es que la variabilidad del índice de sensibilidad cambia drásticamente, por otra parte los índices basados en combinaciones totales tienen valores más continuos los cuales pueden ser considerados para un análisis más profundo de los dato y que puede aportar información más exacta del comportamiento espacial de los pixeles. La desventaja es que necesitan de mayor tiempo y poder de computo.

Para el caso de los índices basados en análisis espacial pueden ser considerados entre los OAT y los de combinaciones totales, ya que se observan valores continuos sin cambios rápidos y que aportan más información sobre la sensibilidad geoespacial, sin embargo estos dependen mucho de tamaño de ventana de análisis lo cual puede dar resultados diferentes.

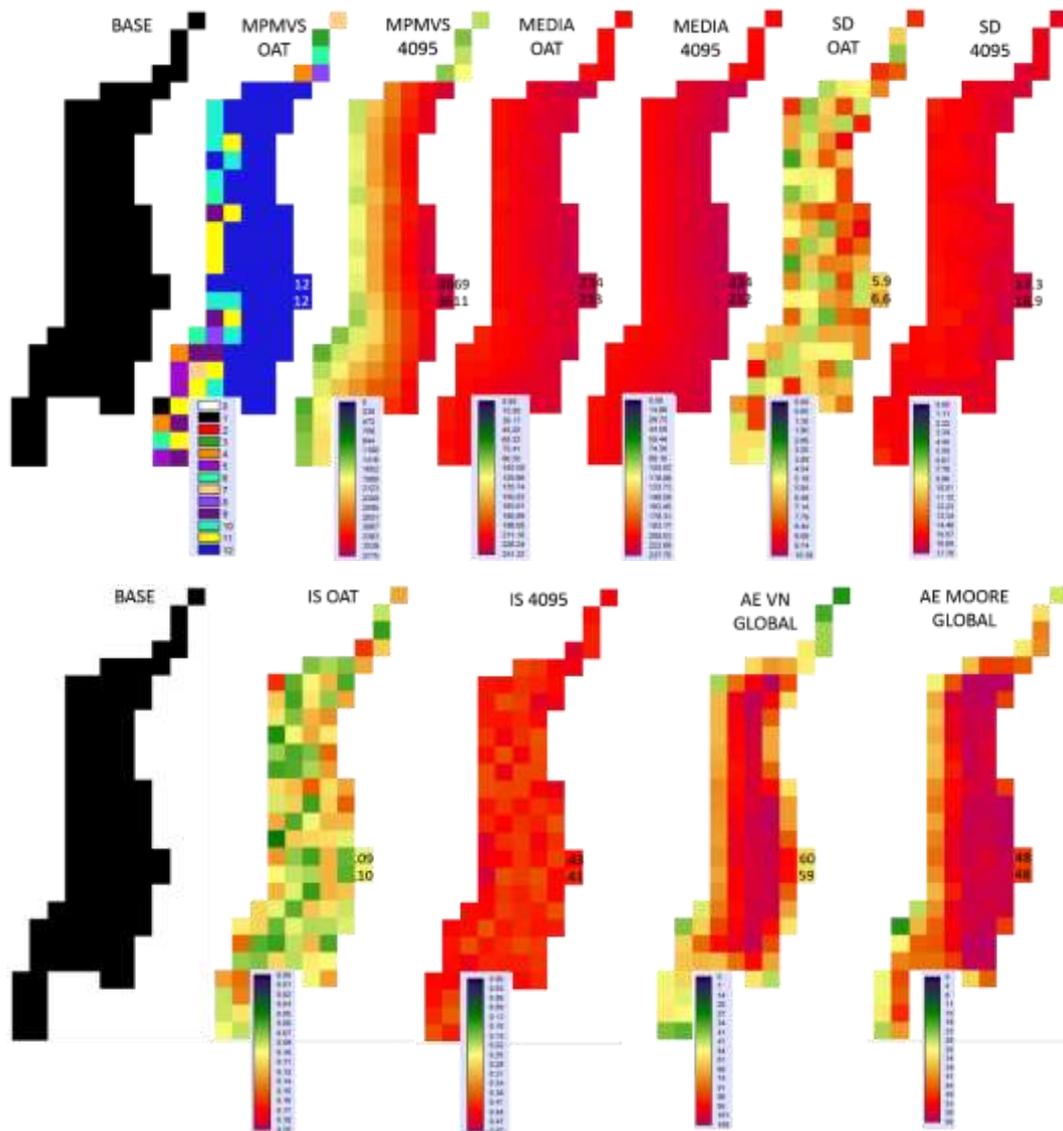


Figura 61.- Comparación de diferentes índices para evaluar de sensibilidad espacial.

6.8 Comparación de la simulación al 2030 con crecimiento actual

Con la finalidad de observar los cambios en la ciudad de Culiacán al 2016 y compararlos con la simulación base, se han seleccionado las 5 áreas más grandes de la simulación que se ubican en: 1) Sector Norte (Lomas de Rodriguera y Mojolo), 2) Sector Humaya (Fraccionamiento Bosques Encino, la Finca, Campestre los Laureles y Praderas), 3) Sector Valle Alto (Fraccionamiento Toscana, Villas del Cedro y Valle Alto), 4) Sector Aeropuerto (Aeropuerto, San Javier, Fraccionamiento Joyas del Valle y Fraccionamiento Santa Rocio) y 5) Sector Sur (Fraccionamiento Valle del Sol, Bugambilias, El Cardenal y el Campo el Diez).

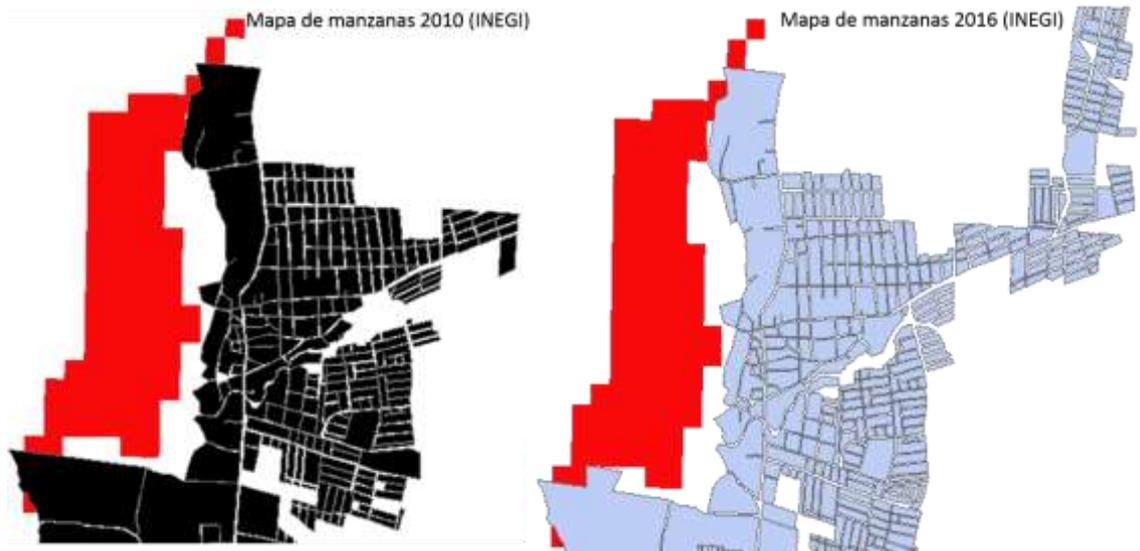


Figura 63.- Comparación del crecimiento simulado en el sector norte y el crecimiento actual.

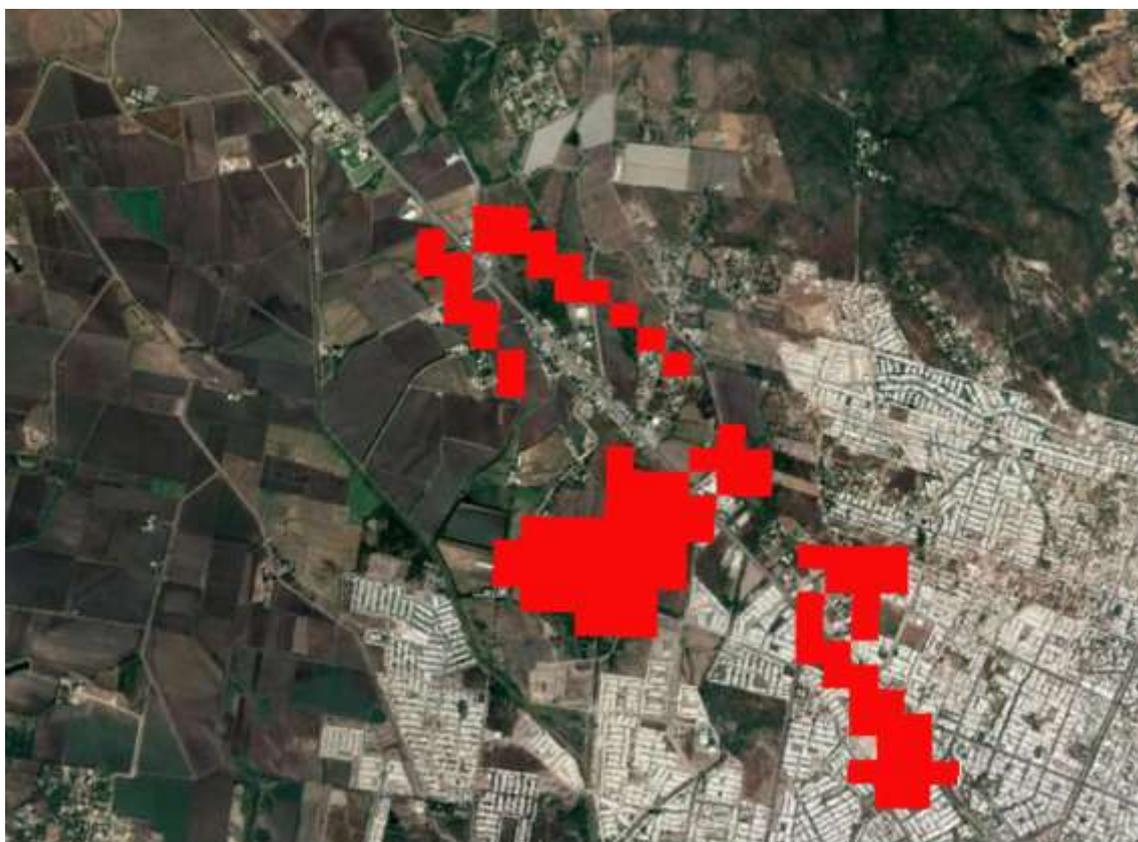
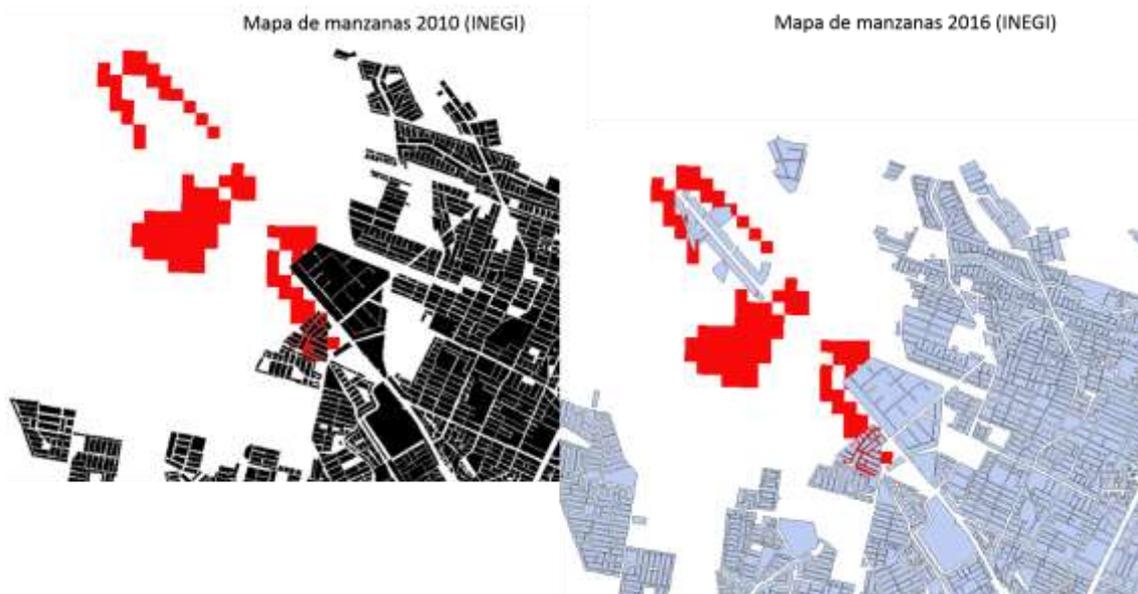


Figura 64.- Comparación del crecimiento simulado en el sector Humaya y el crecimiento actual.

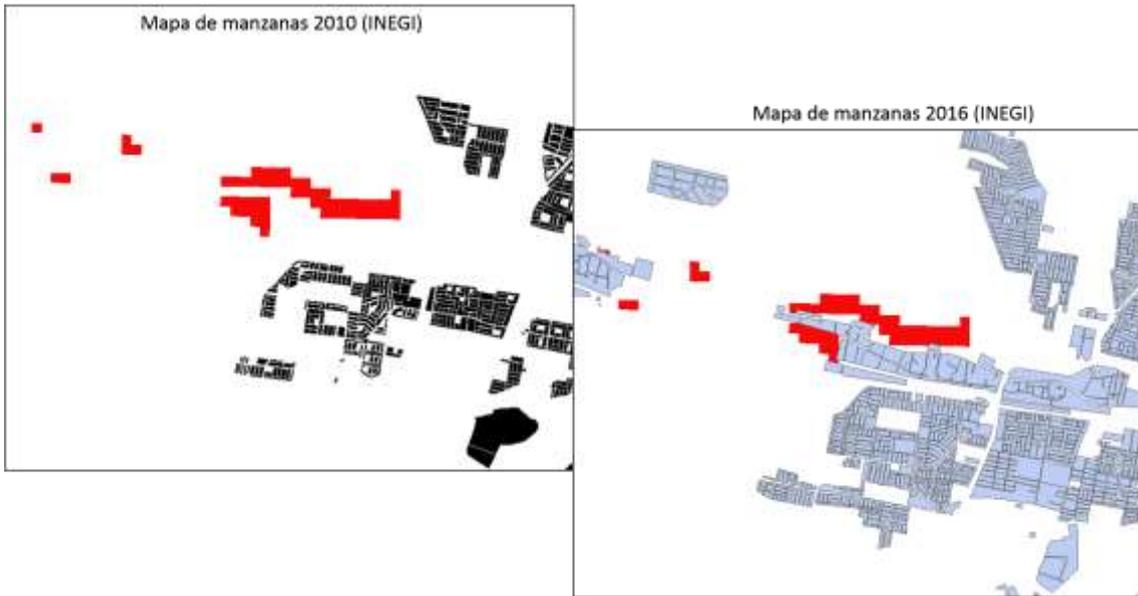


Figura 65.- Comparación del crecimiento simulado en el sector Valle Alto y el crecimiento actual.



Figura 66.- Comparación del crecimiento simulado en el sector Aeropuerto y el crecimiento actual.



Figura 67.- Comparación del crecimiento simulado en el sector Sur y el crecimiento actual.

7.- DISCUSIÓN

Como se ha observado en este trabajo se ha prestado poca atención a la credibilidad (robustez) de los modelos, y aplicación de un AS no se considera generalmente como un componente crítico en el desarrollo de modelos.

Además un inconveniente para los estudios geoespaciales es que a pesar de obtener resultados esperados, no toma en cuenta la característica de vecindad (donde cada píxel es independiente de los demás), y en muchos casos la geoespacial (donde se tiene un IS por cada pixel en estudio), además que los mapas de accesibilidad y aptitud, necesarios para su aplicación, no modifican todos al mismo tiempo.

Por otra parte, si consideramos la principal desventaja de los MBA y AC, la cual es que los resultados son afectados por las diferentes configuraciones del modelo, respecto a tamaño y tipo de vecindad, así como por la resolución espacial y definición de reglas lo que genera también el aumento del tiempo de cómputo (para algunos casos exponencial) hace que el modelo de EMC sea más estable respecto a los resultados además de ser menos afectado por las suposiciones de los diferentes hiperparámetros de los otros modelos. Y en general hasta el momento. Todos los estudios son rebasados en cuanto a su capacidad computacional.

En cuanto a las técnicas de AS, la problemática generalizada es la falta de aplicación a temas geoespaciales, ya que en general son usadas para dar un índice general por cada mapa y no por cada píxel en el mapa, la otra desventaja observada es la falta de capacidad de generación de múltiples modelos que consideren las diferentes combinaciones de las variables originales y sus modificaciones, ya que el estudio con más mapas en estudio fue de 432 mapas.

Ahora bien, en el caso de las herramientas computacionales encontradas, ninguna de ellas trabajar directamente con modelos de simulación geoespacial y ni tampoco realizan análisis de sensibilidad explícitamente espaciales, decir que ninguna tiene resultados geoespaciales, por lo que se tiene certeza de que

pierden la característica puramente espacial de la información, siendo solo un estudio estadístico del mapa.

También, algunas herramientas tienen un mal manejo de errores, son necesarios muchos clics y cambios de pantallas, para lograr el proceso completo de AS, en otros casos se les puede considerar una caja negra, dado que no se sabe cómo realiza el cálculo, son poco amigable con el usuario.

El método OAT que se propone en este estudio a diferencia de todo los demás trabajos, se robustece generando todas las combinaciones de las modificaciones entre factores y pesos, lo que genera como resultado un análisis global, además este trabajo es el único donde se analiza una imagen con resolución de 1274 x 1379, las modificaciones en los factores y pesos de entrada son a nivel espacial y también se tiene como resultado 4085 mapas de simulaciones de la misma resolución que el original.

Aunado a esto, y para el caso de aquellos estudios que utilizan IDRISI, se considera una selección aleatoria que obtiene pixeles distribuidos equitativamente, como se vio antes.

Por otra parte, este estudio genera diferentes índices de sensibilidad geoespacial, es decir visibles visualmente, a nivel de pixeles y a nivel entorno, lo cual aporta distinta información para el área en estudio, y que en su combinación lograrán ampliar el análisis de sensibilidad.

Así mismo, en este trabajo se realiza una aplicación computacional totalmente intuitiva para el usuario, la cual genera simulaciones estrictamente espaciales de dos formas diferentes y calcula los diferentes índices de sensibilidad, observados en capítulos anteriores y lo que facilitará a los expertos a aplicar AS en todos los estudios y obteniendo resultados más robustos.

8.- CONCLUSIONES

Hasta el momento no existe ninguna metodología análisis de sensibilidad que pueda ser aplicada a ningún tipo de modelo. En consecuencia, tampoco se encontró alguna metodología que integre a los modelo de simulación geoespacial, un proceso de análisis de sensibilidad y menos un análisis explícitamente espacial.

La modificación más usada en el AS está basada en la variación de los pesos de los factores (Gómez y Bosque 2004). Sin embargo, también se sugiere variar la estructura del problema (objetivos, criterios, etc) (Colorni 1999).

La desventaja más importante de todas las técnicas antes mencionadas es que ninguna de ellas trabaja directamente con la distribución espacial de los pixeles de un mapa (considerando problemas de cambio de uso y cobertura de suelo o crecimiento urbano), siendo todos ellos de tipo estadístico y solamente teniendo como entrada los factores, pesos y las funciones de distribución de con los que se trabajarán los pesos, en otras palabras, se pierde la característica explícita de espacialidad que está representada en el mapa. En consecuencia, el AS no es una práctica común, por los que resulta muy poco aprovechado y escasamente aplicado en áreas de la geomática y otras que contemplan modelos de simulación geoespacial (Gómez y Bosque 2004).

Asimismo, los cálculos estadísticos, para la obtención de los resultados de análisis, regularmente son realizados en herramientas computacionales externas al software SIG usado para el análisis espacial, entre esas aplicaciones se pueden encontrar SimLab, Matlab, R, IntiGIS, entre otros.

Existe una laguna de conocimiento aplicado a este tema de interés global, esto debido a que existen pocos trabajos de índole explícitamente espacial, y considerando el firme antecedente de que las técnicas hasta el momento no permite determinar en SIG la ubicación a nivel espacial y/o territorial donde el modelo funciona mejor o peor.

Es importante mencionar que debido a la falta de trabajos en esta área, es el gran interés que se tiene para la realización de esta investigación, ya que este análisis resulta muy poco aprovechado y escasamente aplicado en la geomática y áreas que contemplan la simulación de escenarios geoespaciales.

Del estudio de aplicaciones realizado en esta investigación, es necesario destacar que de todas las herramientas que contempla esta investigación, excepto IntiGIS y iUSA (no fueron posibles analizar), todas realizan análisis estadístico, teniendo como entradas datos estadísticos, y los resultados solo son gráficas, por lo que se puede concluir que no existen resultados explícitamente espaciales (mostrados en mapa), ni tampoco es posible el uso de mapas como parámetros de entrada.

Entonces, es debido a la incapacidad de determinar la ubicación a nivel espacial y/o territorial en los SIG, es el enfoque de la propuesta de este trabajo para desarrollar un índice de sensibilidad explícitamente espacial y su metodología de aplicación que permitirá a nivel geoespacial verificar el grado de certidumbre en los resultados, principalmente, de escenarios prospectivos del futuro, paso que no se lleva a cabo en la actualidad.

Para alcanzar el objetivo se trabaja un esquema metodológico que integra la implementación de un algoritmo computacional en SIG que permita la ejecución del modelo un determinado número de veces, bajo el principio de un factor a la vez y el principio de la combinatoria sin repetición, generando al final un índice de sensibilidad espacial mediante el análisis de los N modelos ejecutados.

El objetivo principal de esta herramienta es que se identificarán, de manera visual, las áreas o zonas más estables a los cambios en los valores de los factores del modelo. Con el trabajo se espera que los expertos en el área, a nivel internacional, vean esta propuesta metodológica como una herramienta que de confianza y robustez a sus resultados.

Cabe destacar que no existe ningún estudio donde los mapas de entrada sean de tamaño 1274 x 1379, que las modificaciones en los factores y pesos de

entrada sean a nivel espacial y además se obtengan como resultado 4085 mapas de simulaciones. En su defecto solo se contemplan muestras o PDFs como factores de entrada para el análisis.

Se muestra que las modificaciones a los pesos de los factores tiene un mayor impacto espacial en comparación de las alteraciones de los factores debido a que arroja resultados menores, lo que quiere decir que los pesos aportan mayor sensibilidad a los escenarios, cuando son cambiados.

Finalmente, se concluye que es de gran importancia el desarrollo de una propuesta metodológica para realizar análisis de sensibilidad explícitamente espacial donde que su aplicación permita a nivel geoespacial verificar el grado de robustez en los resultados, calculando los diferentes índices de sensibilidad, y lo que facilitando a los expertos la aplicación automatizada de AS en todos los estudios geoespaciales.

BIBLIOGRAFÍA

- Ahmed, B., Ahmed, R. y Zhu, X., 2013. Evaluation of Model Validation Techniques in Land Cover Dynamics. *ISPRS International Journal of Geo-Information*, Volumen 2, pp. 577-597.
- Al-Ahmadi, K., See, L. y Heppenstall, A., 2013. Validating Spatial Patterns of Urban Growth from a Cellular Automata Model. En: *Emerging Applications of Cellular Automata*. INTECH, pp. 23-52.
- Allen, J. y Lu, K., 2003. Modeling and prediction of future urban growth in the Charleston region of South Carolina: a GIS-based integrated approach. *Conservation Ecology*, 8(2).
- Anon., 2010. INEGI. [En línea] Available at: <http://www.inegi.org.mx/>
- Aracil, J., y Gordillo, F. (1995). *Dinámica de sistemas*, 1 ed. Madrid.
- Arsanjani, J. J., Helbich, M., Kainz, W. y Boloorani, A. D., 2013. Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *International Journal of Applied Earth Observation and Geoinformation*, Volumen 21, pp. 265-275.
- Barredo, J. I., Demicheli, L., Lavalle, C., Kasanko, M., & McCormick, N. (2004). Modelling future urban scenarios in developing countries: an application case study in Lagos, Nigeria. *Environment and Planning B: Planning and Design*, 31(1), 65-84.
- Basse, R. M. Omrani, H., Charif, O., Bódis, K., & Gerber, P, 2012. Cellular automata based neural networks for modelling dual complex systems: Land-Use/Cover and transport networks. Oporto, Portugal.
- Bousquet, F. y Le Page, C., 2004. Multi-agent simulations and ecosystem management: a review. *Ecological Modelling*, Volumen 176, p. 313–332.
- Canters, F., De Genst, W. y Dufourmont, H., 2002. Assessing effects of input uncertainty un structural landscape classification. *Geographical Information Science*, pp. 129-149.
- Castaigns, W., Dartus, D., Le Dimet, F. X. y Saulnier, G. M., 2009. Sensitivity analysis and parameter estimation for distributed hydrological modeling: potential of variational methods. *Hydrology and Earth System Sciences*, pp. 503-517.
- Chandio, I. A., Matori, A. N. B., WanYusof, K. B. y Talpur, M. A. H., 2013. Validation of Multi-Criteria Decision Analysis Model of Land Suitability

- Analysis for Sustainable Hillside Development. *European Journal of Scientific Research*, 109(2), pp. 342-349.
- Clarke, K. C. y Gaydos, L. J., 1998. Loose-coupling a Cellular Automaton Model and GIS: Long-term Urban Growth Prediction for San Francisco and Washington/Baltimore. *International Journal Geographical Information Sciences*, Volumen 12, pp. 699-714.
- Corrales, G. (2013). Análisis de cambio de uso de suelo en el estado de Sinaloa utilizando sistemas de información geográfica (Tesis de licenciatura. Universidad Autónoma de Sinaloa. Culiacán, Sinaloa, México).
- De Almeida, C. M. y Gleriani, J. M., 2005. Cellular automata and neural networks as a modelling framework for the simulation of urban land use change. *Anais XII Simpósio Brasileiro de Sensoriamento Remoto*, pp. 3697-3705.
- Eastman, J. R. (1999). Multi-criteria evaluation and GIS. *Geographical information systems*, 1(1), 493-502.
- Eastman, J. R. (2012). *IDRISI Selva manual*. Worcester: Clark Labs.
- Feizizadeh, B., Blaschke, T. y Nazmfar, H. (2014). GIS-based ordered weighted averaging and Dempster–Shafer methods for landslide susceptibility mapping in the Urmia Lake Basin, Iran. *International Journal of Digital Earth*, 7(8), 688-708.
- Feizizadeh, B., Jankowski, P. y Blaschke, T. (2013). A spatially explicit approach for sensitivity and uncertainty analysis of GIS-multicriteria landslide susceptibility mapping (Vol. 2013, pp. 157-164). *Verlag der Österreichischen Akademie der Wissenschaften*.
- Feizizadeh, B., Jankowski, P. y Blaschke, T. (2014). A GIS based spatially-explicit sensitivity and uncertainty analysis approach for multi-criteria decision analysis. *Computers & geosciences*, 64, 81-95.
- Feizizadeh, B., Roodposhti, M. S., Jankowski, P. y Blaschke, T. (2014b). A GIS-based extended fuzzy multi-criteria evaluation for landslide susceptibility mapping. *Computers & geosciences*, 73, 208-221.
- Fischer, M. M., 1998. Computational neural networks: a new paradigm for spatial analysis. *Environment and Planning*, Issue 30, pp. 1873-1891.
- Gallaher, M. y Doherty, J., 2007. Parameter estimation and uncertainty analysis for watershed model. *Environmental Modelling y Software*, pp. 1000-1020.

- Gomez, M. J. (2012). Redes Neuronales Artificiales para la simulación, estimación y control de caudales impulsados por estaciones de bombeo en zonas regables. Huelva, España: s.n.
- Gómez Delgado, M., Aguilera Benavente, F., Barreira González, P., Bosque Sendra, J. y Rodríguez Espinosa, V. M. (2014). Simulación prospectiva del crecimiento urbano en la Comunidad Autónoma de Madrid a partir de modelos basados en autómatas celulares y modelos basados en EMC. Alicante., XVI Congreso Nacional de Tecnologías de la Información Geográfica.
- Gómez Delgado, M. y Barredo Cano, J. I., 2005. Sistemas de Información Geográfica y evaluación multicriterio en la ordenación del territorio. 2da edición ed. Madrid, España: Ra-Ma.
- Gómez Delgado, M. y Bosque Sendra, J., 2004. Aplicación de análisis de incertidumbre como método de validación y control de riesgo en la toma de decisiones. Revista Internacional de Ciencia y Tecnología de la Información Geográfica, pp. 179-208.
- Gómez Delgado, M. y Bosque Sendra, J., 2004b. Sensivity Analysis in Multicriteria Spatial Decision-Making: A Review. Human and Ecological Risk Assessment, pp. 1173-1187.
- González, P. B., Aguilera-Benavente, F. y Gómez-Delgado, M. (2015). Partial validation of cellular automata based model simulations of urban growth: An approach to assessing factor influence using spatial methods. Environmental Modelling & Software, 69, 77-89.
- Gracia, M. T. L., 2012. Comparación de técnicas de evaluación multicriterio en sistemas de ayuda a la decisión espacial para la localización óptima de uso residencial. Geographicalia, Issue 62, pp. 89-114.
- Gutiérrez Angonese, J., Gómez Delgado, M. y Bosque Sendra, J. (2010). Simulación de crecimiento urbano mediante evaluación multicriterio y TIG en el Gran San Miguel de Tucumán (Argentina). In Congreso Nacional de Tecnologías de la Información Geográfica, pp. 873-888. Universidad de Sevilla.
- Hamby, D. M., 1995. A comparison of sensitivity analysis techniques. Health Physics, pp. 195-204.

- Hegde, N. P., Muralikrishna, I. V. y Chalapatirao, K. V. (2008). Settlement growth prediction using neural network and cellular automata. *Journal of Theoretical & Applied Information Technology*, 4(5).
- Hejlsberg, A., Wiltamuth, S., y Golde, P. (2003). C# language specification. Addison-Wesley Longman Publishing Co., Inc.
- Hoare, A., Regan, D. y Wilson, D., 2008. Sampling and sensitivity analyses tools (SaSAT) for computational modelling. *Theoretical Biology y Medical Modelling*, 5(4).
- Huang, B., Zhang, L. y Wu, B., 2009. Spatiotemporal analysis of rural–urban land conversion. *International Journal of Geographical Information Science*, 23(3), pp. 379-398.
- Huang, Z. y Laffan, S. W., 2009. Sensitivity analysis of a decision tree classification to input data errors using a general Monte Carlo error sensitivity model. *International Journal of Geographical Information Science*, pp. 1433-1452.
- Hu, Z. y Lo, C., 2007. Modeling urban growth in Atlanta using logistic regression. *Computers, Environment and Urban Systems*, Volumen 31, pp. 667-688.
- INEGI, 2010. INEGI. [En línea] Available at: <http://www.inegi.org.mx/>
- Santé Riveira I., Crecente Maseda R. y Miranda Barrós D., 2007. Aplicación SIG para la evaluación de la aptitud de la tierra. XI Congreso Internacional de Ingeniería de Proyectos, LUGO, 26-28 Septiembre.
- Iooss, B. y Lemaître, P. (2015). A review on global sensitivity analysis methods. In *Uncertainty management in simulation-optimization of complex systems* (pp. 101-122). Springer, Boston, MA.
- Jacquín, A. P. y Shamseldin, A. Y., 2009. Sensitivity analysis of Takagi-Sugeno-Kang rainfall-runoff fuzzy models. *Hydrology and Earth System Sciences*, pp. 41-55.
- Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A., & Moradi, A. (2012). An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environmental Earth Sciences*, 67(1), 251-264.
- Kocabas, V., & Dragicevic, S. (2004, July). Sensitivity analysis of a GIS-based cellular automata model. In *Int. archives of the Photogrammetry, Remote*

- Sensing and Spatial Information Sciences, ed. By Altman MO, XXth ISPRS Congress, Istanbul, Turkey (Vol. 35, No. part B).
- Koomen, E., Stillwell, J., Bakema, A. y Scholten, H. J., 2007. Modelling Land-Use Change: Progress and Applications. The Netherlands: Springer.
- Lee, S.-T., Wu, C.-W. y Lei, T.-C., 2013. CA-GIS model for dynamic simulation of commercial activity development by the combination of ANN and Bayesian probability. *Procedia Computer Science*, Issue 18, pp. 651-660.
- Lewicki, P., & Hill, T. (2006). *Statistics: methods and applications*. Tulsa, OK. Statsoft. Washington, D.C.
- Ligmann-Zielinska, A. y Jankowski, P., 2014. Spatially-explicit integrated uncertainty and sensitivity analysis of criteria weights in multicriteria land suitability evaluation. *Environmental Modelling y Software*, Volumen 57, pp. 235-247.
- Ligmann-Zielinska, A., 2014. Using Spatially-Explicit Uncertainty and Sensitivity Analysis in Spatial Multicriteria Evaluation.
- Lilburne, L. y Tarantola, S., 2009. Sensitivity analysis of spatial models. *International Journal of Geographical Information Science*, 23(2), pp. 151-168.
- Li, X. y Yeh, A., 2001. Calibration of Cellular Automata by using Neural Networks for the Simulation of Complex Urban System. *Environment and Planning A.*, 33(8), pp. 1445-1462.
- Liberty, J. (2005). *Programming C#: Building .NET Applications with C*. "O'Reilly Media, Inc."
- Lu, Y. y Mohanty, S., 2001. Sensitivity analysis of a complex, proposed geologic waste disposal system using the Fourier Amplitude Sensitivity Test Method. *Reliability Engineering y System Safety*, pp. 275-291.
- Mahiny, A. S., & Turner, B. J. (2003, September). Modeling past vegetation change through remote sensing and GIS: a comparison of neural networks and logistic regression methods. In *Proceedings of the 7th international conference on geocomputation*. University of Southampton, UK.
- Makowski, D. y otros, 2006. Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model prediction. *Reliability Engineering and Systems Safety*, Volumen 91, pp. 1142-1147.

- McKenney, D. W., Mackey, B. G. y Zavitz, B. L., 1999. Calibration and sensitivity analysis of a spatially-distributed solar radiation model. *Geographical Information Science*, pp. 49-65.
- McRae, G. J., Tilden, J. W. y Seinfeld, J. H., 1982. Global sensitivity analysis—a computational implementation of the Fourier Amplitude Sensitivity Test (FAST). *Computers y Chemical Engineering*, Volumen 6, pp. 15-25.
- Mogaji, K. A., Lim, H. S., & Abdullah, K. (2014). Modeling groundwater vulnerability prediction using geographic information system (GIS)-based ordered weighted average (OWA) method and DRASTIC model theory hybrid approach. *Arabian Journal of Geosciences*, 7(12), 5409-5429.
- Moghaddam, H. K., & Samadzadegan, F. (2009). Urban simulation using neural networks and cellular automata for land use planning. na.
- Monjardín Armenta, S. A. y Pacheco Angulo C. E., Plata Rocha W., Corrales Barraza G., Romero Andrade R. y Beltran Gonzalez J. C., 2016. Determination and Analysis of Hot Spot Areas of Deforestation Using Remote Sensing and Geographic Information System Techniques. Case Study: State Sinaloa, México. *Open Journal of Forestry*, Volume 6, pp. 295-304.
- Moreno, J. A. y Crosetto, M., 2001. Análisis de la incertidumbre en modelos de teledetección. In *Teledetección: medio ambiente y cambio global* (pp. 538-541). Departament de Medi Ambient i Ciències del Sòl.
- Norton, J. P., 2008. Algebraic sensitivity analysis of environmental models. *Environmental Modelling y Software*, pp. 963-972.
- Omriani, H., Charif, O., Gerber, P., Bódis, K. y Basse, R. M. (2012). Simulation of land use changes using cellular automata and artificial neural network. Technical Report, CEPS/INSTEAD.
- Orán Cáceres, J. P., Gómez Delgado, M. y Bosque Sendra, J. (2010). Una propuesta complementaria de análisis de sensibilidad de un modelo basado en técnicas SIG y evaluación multicriterio. In *Congreso Nacional de Tecnologías de la Información Geográfica* (14. 2010. Sevilla)(2010), p 971-987 (pp. 971-987). Universidad de Sevilla.
- Oran, J. P., Gómez, M. y Plata, W. (2011). Posibilidades de aplicación de un análisis de sensibilidad secuencial a un modelo basado en técnicas SIG y evaluación multicriterio. In *XIII Conferencia Iberoamericana en Sistemas de Información Geográfica* (ISBN: 978-607-00-4387-1). Toluca, México.

- Parker, D. C., Manson, S. M., Janssen, M. A., Hoffmann, M. J., & Deadman, P. (2003). Multi-agent systems for the simulation of land-use and land-cover change: a review. *Annals of the association of American Geographers*, 93(2), 314-337.
- Parrot, L., Chion, C., Gonzalés, R. y Latombe, G., 2012. Agents, Individuals, and Networks: Modeling Methods to Inform Natural Resource Management in Regional Landscapes. *Ecology and Society*, 17(3), p. 32.
- Pinedo Pascua, I., 2012. Intigis: propuesta metodológica para la evaluación de alternativas de electrificación rural basada en sig. Madrid: editorial CIEMAT.
- Plata Rocha, W. (2010). Descripción, análisis y simulación del crecimiento urbano mediante tecnologías de la información geográfica. El caso de la Comunidad de Madrid (Doctoral dissertation, Universidad de Alcalá).
- Plata Rocha, W., Gómez-Delgado, M., & Bosque-Sendra, J. (2012). Proposal for the introduction of the spatial perspective in the application of global sensitivity analysis. *Journal of Geographic Information System*, 4(06), 503.
- Plata Rocha, W., Delgado, M. G., Sendra, J. B. y Villegas, J. A. (2013). Análisis de sensibilidad para un modelo de simulación de crecimiento urbano. Propuesta metodológica explícitamente especial. *GeoFocus. Revista Internacional de Ciencia y Tecnología de la Información Geográfica*, (13_2), 158-178.
- Pontius Jr., R. G. y Schneider, L. C., 2001. Land-cover change model validation by an ROC method for the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment*, Volumen 85, pp. 239-248.
- Pourghasemi, H. R., Pradhan, B., Gokceoglu, C. y Moezzi, K. D., 2012. Landslide Susceptibility Mapping Using a Spatial Multi Criteria Evaluation Model at Haraz Watershed, Iran. In *Terrigenous Mass Movements* (pp. 23-49). Springer, Berlin, Heidelberg.
- Proctor, W. y Qureshi, E., 2005. Multi-criteria Evaluation Revisited. *Proceedings of the Australia New Zealand Society for Ecological Economics Conference*, pp. 11-13.
- Qureshi, M. E. y Harrison, S. R., 2001. A decision support process to compare Riparian revegetation options in Scheu Creek catchment in North Queensland. *Journal of Environmental Management*, 62(1), pp. 101-112.

- Rana, S. O. H. E. L., Midi, H. A. B. S. H. A. H. y Sarkar, S. K. (2010). Validation and performance analysis of the binary logistic regression model. In Proceedings of the WSEAS International Conference on Environmental, Medicine and Health Sciences (pp. 23-25).
- Ratto, M., Young, P. C., Romanowicz, R., Pappenberger, F., Saltelli, A. y Pagano, A. (2007). Uncertainty, sensitivity analysis and the role of data based mechanistic modeling in hydrology. *Hydrology and Earth System Sciences Discussions*, 11(4), 1249-1266.
- Reginster, I. y Rounsevell, M., 2006. Scenarios of future urban land use in Europe. *Environment and Planning*, p. 619 – 636.
- Reichert, P. (1998). *Aquasim 2.0-user manual*. Swiss Federal Institute for Environmental Science and Technology. Dubendorf, Switzerland.
- Rienow, A. y Stenger, D., 2014. Geosimulation of urban growth and demographic decline in the Ruhr: a case study for 2025 using the artificial intelligence of cells and agents. *Journal of Geographical Systems*, Volumen 16, pp. 311-342.
- Rodriguez-Fernandez, M. y Banga, J. Rodriguez-Fernandez, M. y Banga, J. R., 2009. *SensSB.- A software toolbox for sensitivity analysis in systems biology models*. Palo Alto, CA, Stanford University.
- Rounsevell, M. D. A., Reginster, I., Araújo, M. B., Carter, T. R., Dendoncker, N., Ewert, F., y Schmit, C. (2006). A coherent set of future land use change scenarios for Europe. *Agriculture, Ecosystems & Environment*, 114(1), 57-68.
- Roura-Pascual, N., Krug, R. M., Richardson, D. M. y Hui, C., 2010. Spatially-explicit sensitivity analysis for conservation management: exploring the influence of decisions in invasive alien plant management. *Diversity and Distributions*, pp. 426-438.
- Rubner, Y., Tomasi, C., & Guibas, L. J. (2000). The earth mover's distance as a metric for image retrieval. *International Journal of Computer Vision*, 40(2), 99-121.
- Saaty, 1980. *The analytical Hierarchy Process*. New York: Mc Graw Hill.
- Santillán Rodríguez, V., 2012. Identificación de zonas de importancia ambiental y vulnerabilidad de ecosistemas mediante métodos de evaluación multicriterio (EMC) en el entorno de los Sistemas de Información Geográfica

- (SIG). Cuenca del río Paute (CRP), suroriente del Ecuador. *Geografía y Sistemas de Información Geográfica (GEOSIG)*, Issue 4, pp. 56-76.
- Schneider, L. C. y Pontius Jr., R. G., 2001. Modeling land-use change in the Ipswich watershed, Massachusetts, USA. *Agriculture, Ecosystems and Environment*, Volume 85, pp. 83-94.
- Shannon, C. E. y Weaver W., 1949. *The mathematical theory of communication*. The Univ. of Illinois Press, Urbana, IL.
- Ziehn, T. y Tomlin, A. S. (2009). GUI-HDMR—A software tool for global sensitivity analysis of complex models. *Environmental Modelling & Software*, 24(7), 775-785.
- Tang, Y., Reed, P., Wagener, T. y van Werkhoven, K., 2007. Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation. *Hydrology and Earth System Sciences*, pp. 793-817.
- Tayyebi, A., 2013. *Simulating Land Use Land Cover Change Using Data Mining and Machine Learning Algorithms*. West Lafayette, Indiana.
- Tenerelli, P. y Carver, S., 2012. Multi-criteria, multi-objective and uncertainty analysis for agro-energy spatial modelling. *Applied Geography*, Volumen 32, pp. 724-736.
- Torrens, P. M., 2011. Calibrating and validating cellular automata models of urbanization. In *Urban Remote Sensing: Monitoring, Synthesis and Modeling in the Urban Environment*, pp. 335-345.
- Torrents, P. M., 2011. Calibrating and validating cellular automata models of urbanization. Yang, Xiaojun ed. Chichester: John Wiley y Sons.
- Triantakonstantis, D. y Mountrakis, G., 2012. Urban Growth Prediction: A Review of Computational Models and Human Perceptions. *Journal of Geographic Information System*, Volumen 4, pp. 555-587.
- Wang, S. Q., Zheng, X. Q. y Zang, X. B., 2012. Accuracy assessments of land use change simulation based on Markov-cellular automata model. *Procedia Environmental Sciences*, Volumen 13, p. 1238 – 1245.
- Wu, F., 2002. Calibration of Stochastic Cellular Automata: The Application to Rural-urban land Conversions. *International Journal of Geographical Information Systems*, 8(16), pp. 795-818.

- Wu, F., 2002. Calibration of stochastic cellular automata: the application to rural-urban land conversions. *International Journal of Geographical Information Science*, 16(8), pp. 795-818.
- Xu, E. y Zhang, H. (2013). Spatially-explicit sensitivity analysis for land suitability evaluation. *Applied Geography*, 45, 1-9.
- Yager R., 1988. On ordered weighted averaging aggregation operators in multicriteria decision-making. *IEEE Transactions on Systems, Man, and Cybernetics*, Volumen 8, p. 183–90.