



MODELO NUMÉRICO DE LA BIFURCACIÓN DEL RÍO MEZCALAPA USANDO PROGRAMACIÓN GENÉTICA

Katya Rodríguez Vázquez¹, Rosalva Mendoza Ramírez² y A. Abel Jiménez Castañeda²

¹Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas. ²Instituto de Ingeniería
Universidad Nacional Autónoma de México, México.

E-mail: katya@uxdea4.iimas.unam.mx, rnr@pumas.iingen.unam.mx, ajc@pumas.iingen.unam.mx

RESUMEN

Se presenta un modelo numérico de terreno de la zona correspondiente a la bifurcación del río Mezcalapa, nombre local del río Grijalva en un tramo de éste en el estado de Tabasco, a partir de mediciones obtenidas en campo por medio de un levantamiento topográfico. El modelo numérico que se presenta para generar la superficie topográfica está basado en el uso de Algoritmos Evolutivos (AE) (Bäck, 1996), específicamente Programación Genética (PG) (Koza, 1992; Banzhaf et al., 1998), los cuales han probado ser una buena alternativa para abordar problemas de interpolación.

ABSTRACT

This paper presents a numerical model of a surface built from data obtained by means of a topographic measurements corresponding to the zone of the bifurcation of the Mezcalapa River, local name given to the Grijalva River in the Tabasco state. The numerical model presented in this work that generates a topographic surface, uses Evolutionary Algorithms (EAs) (Bäck, 1996), specifically Genetic Programming (P) (Koza, 1992; Banzhaf et al. 1998), which has shown to be a promising alternative to deal with interpolation problems.

Las lluvias atípicas que desde fines del siglo pasado se han presentado en la RM, principalmente en la parte sureste, han provocado serios daños materiales y pérdidas económicas y humanas. En particular, la ciudad de Villahermosa ha sido uno de los centros poblacionales que más ha sufrido su impacto; por ello y en aras de proteger a la población civil y disminuir el riesgo de que la ciudad se inunde por el desbordamiento del río Carrizal, se han realizado varios estudios para el diseño y construcción de obras hidráulicas; una de esas obras es la estructura de control que se localiza sobre el río Carrizal, a casi 1 km aguas abajo de la bifurcación, (figura 2); el objetivo de dicha estructura es controlar el flujo del río Carrizal para evitar que éste se desborde. Así, por ejemplo, en Jiménez et al (2008), se presenta el estudio del funcionamiento hidráulico del sistema formado por la bifurcación Mezcalapa- Samaria – Carrizal, donde se explica el proceso de azolve del río Samaria, con la consecuente reducción de su capacidad de conducción, y el incremento del flujo por el río Carrizal, dando lugar a que se exceda su capacidad de conducción y el desbordamiento del cauce. Además, en Jiménez y Berezowsky (2009), se presenta el estudio realizado para proponer políticas de operación de la estructura de control del río Carrizal, la cual está en su etapa final de construcción. La edificación de estas obras hidráulicas da lugar a que se modifique el funcionamiento hidráulico de los ríos y la bifurcación, por lo que en el mediano plazo se tendrán problemas de erosión y azolve a lo largo de los cauces; por ello, es conveniente estudiar los cambios que se presentan a lo largo de los ríos, principalmente en la zona donde está la bifurcación.



Aunque la construcción de la estructura de control permitirá manejar los caudales en el río Carrizal, de manera alterna, en el mediano y largo plazo se tendrán procesos de erosión y sedimentación de los cauces por el funcionamiento de dicha estructura. Los autores del presente trabajo consideran que, una manera alternativa de estudiar y cuantificar dichos procesos en la vecindad de la bifurcación, se puede hacer con un modelo numérico que permita visualizar la superficie de los cauces en esa zona. Se presenta entonces aquí, un modelo numérico de terreno de la zona correspondiente a la bifurcación del río; el modelo requiere datos de mediciones obtenidas en campo por medio de un levantamiento topográfico, y se basa en el uso de técnicas del campo de la Computación Evolutiva (CE) (Bäck, 1996), específicamente Programación Genética (PG) (Koza, 1992; Banzhaf et al., 1998); estos métodos han probado ser una buena alternativa para resolver problemas de interpolación. El código del modelo matemático de terreno está escrito en el lenguaje del intérprete MATLAB (The Mathworks, 1999) y las superficies generadas se obtienen al aplicar el paquete SURFER (Golden Software, 2010), que permite elaborar gráficas con vistas en 2 y 3 dimensiones.

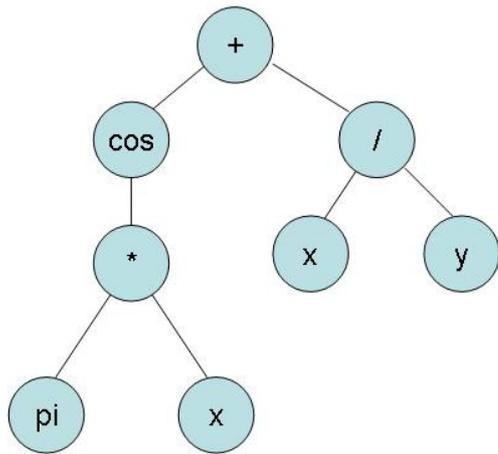
TRABAJOS PREVIOS

En Mendoza y Berezowsky (1996) se aborda el problema de aproximar una función suave a un conjunto dado de datos topográficos de una zona del río Colorado (en la región Noroeste de la RM). Se requería obtener una expresión matemática explícita de este tipo para poder posteriormente derivarla y construir un modelo en coordenadas curvilíneas ajustadas a las fronteras en flujos a superficie libre. El problema se resolvió inicialmente representando la cota topográfica (CT) como una función dependiente de las variables x - y que se aproximó con un desarrollo en series de Taylor (Arfken, 1980), donde se incluyen hasta los términos de tercer orden. Los coeficientes se ajustaron por mínimos cuadrados. Los resultados que se obtuvieron fueron en general satisfactorios, pero se encontró que había un número significativo de puntos del conjunto de datos topográficos en los cuales el método de estimación propuesto no proporcionaba estimaciones apropiadas (Mendoza y Berezowsky 1996). Los errores no aceptables, cuya magnitud era mayor que 1 m, se encontraban cuando el punto para el cual se quería hacer la estimación era una cresta o un valle, y estaba rodeado de datos cuyo valor de CT era bastante distinto. Este tipo de errores es inherente a los métodos de interpolación polinomial y es independiente del tipo de funciones base utilizadas. Con la idea de mejorar la calidad de la aproximación en todos los puntos, en Mendoza (2002) se retomó nuevamente el problema, usando ahora herramientas de la CE, las cuales están basadas en las estrategias naturales de evolución y las leyes de la herencia. Se probó aproximar una zona con crestas y valles para determinar si con el empleo de una de las técnicas de la CE, concretamente la Programación Genética (PG), se mejoraban los resultados obtenidos en Mendoza y Berezowsky (1996). La conclusión obtenida en Mendoza (2002) fue que la CE era una técnica que valía la pena explorar y explotar, puesto que con la PG mejoró de manera notable los errores obtenidos en Mendoza et al (1996). En Mendoza et al (2007) se empleó nuevamente la PG, para estudiar una zona donde se tiene una topografía con un grado de dificultad mayor que la reportada por Mendoza (2002), la zona estudiada presenta tanto zonas planas como partes con crestas y valles, y aún así la función que la PG encuentra con el conjunto de funciones y terminales definidos, reproduce en forma adecuada el comportamiento de la CT.

Tomando esto como base, se optó aquí por representar la superficie topográfica mediante una función matemática; el problema se formula como una regresión simbólica utilizando PG. Se utilizó la caja de herramientas que el intérprete MATLAB tiene desarrollada para PG. El proceso seguido se detalla a continuación.

MATERIALES Y MÉTODOS

En términos generales, en este caso la metodología de aproximación de funciones con PG funciona como sigue. La idea final es obtener una expresión matemática donde se incluyan tanto la aplicación de operadores que provienen del conjunto de funciones como terminales seleccionados a priori. Los elementos de estos conjuntos son determinados por el programador, con base en la estructura del problema o su experiencia previa. Estos operadores se aplican sobre los datos de ubicación topográfica (coordenadas (x_i, y_i) y que constituyen el conjunto de terminales) de manera directa o recursiva para producir una estimación de la CT. Las funciones de aproximación se representan por medio de árboles, donde cada nodo es un dato o un operador. El número máximo de nodos de este árbol es una variable de diseño. El número de ramas que surge de cada nodo depende del número de operandos que requiera el operador en dicho nodo. Un ejemplo de la representación de un individuo en PG se puede observar en la Figura 3.



F = {+, -, *, /, sin, cos, sqrt}
 T = {x, y, pi, -5, -4, -3, -2, -1, 2, 3, 4, 5}

$$z = \cos(\pi * x) + \frac{x}{y}$$

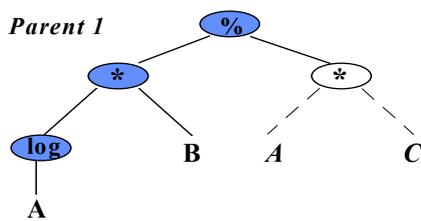
```
function z = individuo(x,y)
z = cos(pi*x)+(x/y);
```

$$(+ (\cos (* \pi x)) (/ x y))$$

Figura 3.- Representación de un individuo en PG: árbol, notación prefija, función matemática y su programa equivalente en MATLAB.

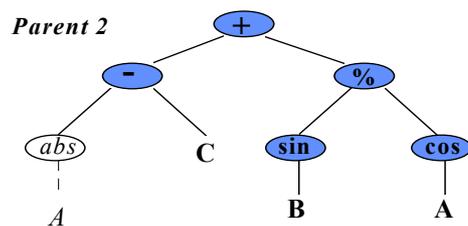
Se debe elegir un número fijo de individuos para constituir la población de algoritmo de PG. Este número es menor o igual que el número total de datos topográficos. En términos generales se intenta que el tamaño de esta población sea lo menor posible y su tamaño es otro de los grados de libertad del problema.

El problema de aproximación se inicia generando al azar funciones aceptables con los datos y operadores de conjunto de funciones seleccionado. Se genera una función para cada individuo de la población. Los elementos de la población cambian por cruce o mutación. La cruce consiste en el intercambio de subárboles entre dos individuos (figura 4) y la mutación (figura 5) en la sustitución de un subárbol de un individuo por otro subárbol generado de manera aleatoria. Los puntos para cruce y mutación y la probabilidad para realizar estas operaciones se determinan con criterios probabilísticos.



$$(\% (* (\log A) B) (* A C))$$

$$\frac{B \log(A)}{AC}$$



$$(+ (- (abs A) C) (\% (\sin B) (\cos A)))$$

$$abs(A) - C + \frac{\sin(B)}{\cos(A)}$$

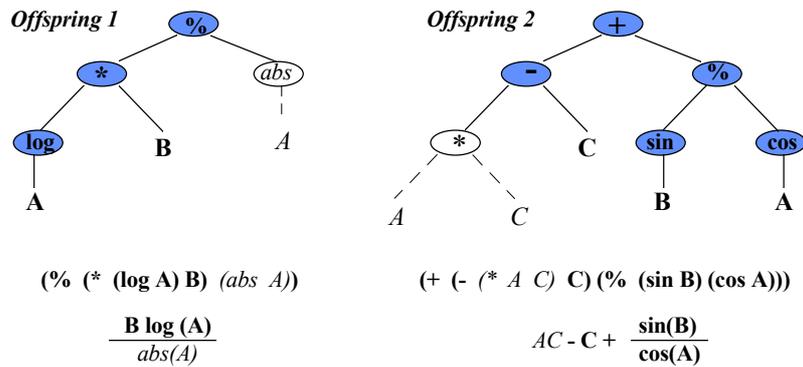


Figura. 4.- Operador de cruzamiento en PG

El método trabaja con iteraciones que se aplican a los individuos. Cada iteración completa constituye una generación en PG. Al iniciar una generación se evalúa el desempeño de cada individuo (función de aptitud), es decir, se obtiene la estimación de la cota al aplicar las operaciones indicadas en el árbol del individuo a evaluar. Los mejores individuos son aquellos cuya CT estimada es más cercana a la CT real. Se preservan los mejores individuos y al resto se les aplica de nueva cuenta cruce y mutación, con las probabilidades elegidas. El método continúa iterando hasta que se satisface algún criterio de error o se agota un número máximo de generaciones. Los lectores interesados en conocer con mayor detalle las técnicas de la CE pueden consultar, por ejemplo, Banzhaf et al (1998) o Golberg (1989).

Para aplicar la metodología de la PG se seleccionó el conjunto de puntos medidos (x_i, y_i, z_i) en 3 secciones transversales del río Mezcalapa, 26 secciones transversales del río Samaria, y 13 secciones transversales del Carrizal (Figura 6). La ubicación detallada de cada una de las secciones transversales y sus correspondientes coordenadas se presenta en Jiménez et al (2007).

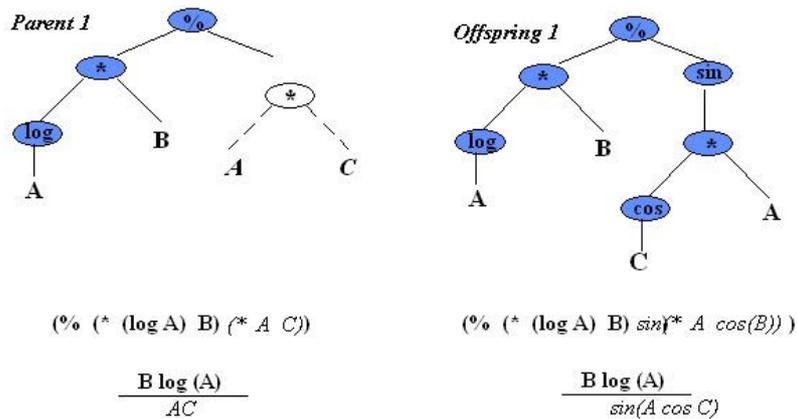


Figura. 5.- Operador de mutación en PG



Figura 6.- Localización de las secciones transversales levantadas

Las curvas de nivel de la región de estudio se muestran en la figura 7, en escala de grises; ahí se puede apreciar que las cotas topográficas más altas se encuentran principalmente en la parte baja de la malla y las más bajas en el centro del modelo; los valores de la CT van desde los 10 msnm a casi los 27 msnm. El valor promedio de la z en la zona es de 17.78 msnm. A los valores reales de las coordenadas (x_i, y_i) se les sumó de manera arbitraria una constante, esto con el fin de manejar números positivos; las ternas de puntos así obtenidas (x_i, y_i, z_i) se graficaron con el paquete SURFER (Golden Software Inc., 2010) generándose una malla de 100 renglones por 42 columnas, con un total de 4200 nodos. Los valores de la coordenadas x y y en el modelo se encuentran en el intervalo $[6.1022, 6.1036]$ y $[5.4079, 5.4115]$, respectivamente.

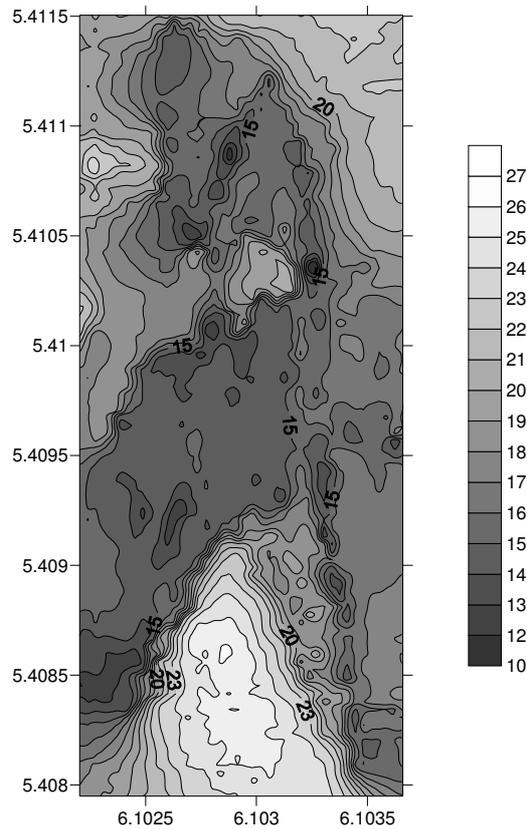


Figura 7.- Mapa de contornos de elevación en la región

Para el desarrollo del código del modelo computacional se utilizaron las funciones que el intérprete MATLAB tiene incluidas para PG, haciéndose las consideraciones siguientes:

- a) El conjunto original de datos topográficos se dividió en 6 regiones; esto se hizo porque al estudiar la topografía de la zona (figura 7), se observó que sería muy difícil encontrar una sola función que pudiera reproducir de manera adecuada la accidentada y abrupta topografía de toda la superficie en estudio.
- b) El conjunto de funciones está compuesto por: los 4 operadores aritméticos básicos, dos funciones trigonométricas (seno y coseno) y la función de raíz cuadrada. Con esto, la aridad (definida como el número de argumentos u operandos que las funciones necesitan) será: {2, 2, 2, 2, 1, 1, 1}, indicando que las funciones aritméticas tienen dos argumentos de entrada y el resto de las funciones solo requieren de uno.
- c) El conjunto de terminales serán las coordenadas (x_i, y_i) y los valores aleatorios de constantes definidas en el intervalo cerrado de -1 a 1.
- d) El criterio para terminar la búsqueda será el máximo número de generaciones.
- e) El desempeño de cada individuo en la población se medirá calculando el error cuadrático medio.
- f) La selección se hizo empleando el mecanismo de torneo de tamaño 3.
- g) La población está compuesta por 100 individuos con un máximo de 256 nodos.
- h) Los valores de las probabilidades de cruce y mutación se fijaron en 0.95 y 0.05, respectivamente.
- i) La notación que se eligió para escribir la función fue la prefija.

RESULTADOS

La figura 8 muestra, en la imagen del lado izquierdo, la topografía real de la zona, se puede ver que es bastante abrupta, predominan en ella las zonas con picos rodeados de valles; estas zonas son las de difícil estimación en el modelo numérico, pues la información con la que

cuenta el predictor para hacer el cálculo no es de gran ayuda. La imagen de la derecha de la misma figura 8 muestra la z estimada con el modelo usado con PG. En general los resultados del modelo son buenos, pero en zonas donde la información aledaña no le ayuda para hacer una buena estimación, tiende a suavizar los picos o los valles; aun así, este primer intento para tener un modelo numérico de la zona tiene resultados promisorios, en la gran mayoría de los puntos se logra hacer una estimación muy cercana a la del valor real. Sin embargo, para que el modelo numérico se pueda utilizar de manera confiable, por ejemplo para conocer la forma como se mueve el lecho del río o para cuantificar los volúmenes de sedimento depositados en ciertas zonas, así como los lugares en donde existan procesos de erosión, etc., es necesario mejorar la precisión de la estimación, sobre todo en aquellos puntos en los cuales el error sobrepasa el medio metro. Con esto, se deja abierta la línea para futuros estudios; por ejemplo, se puede probar qué pasa al ampliar el conjunto de funciones propuesto, también, el hacer ventanas que aislen las zonas en donde la estimación es difícil (zonas de picos y valles).

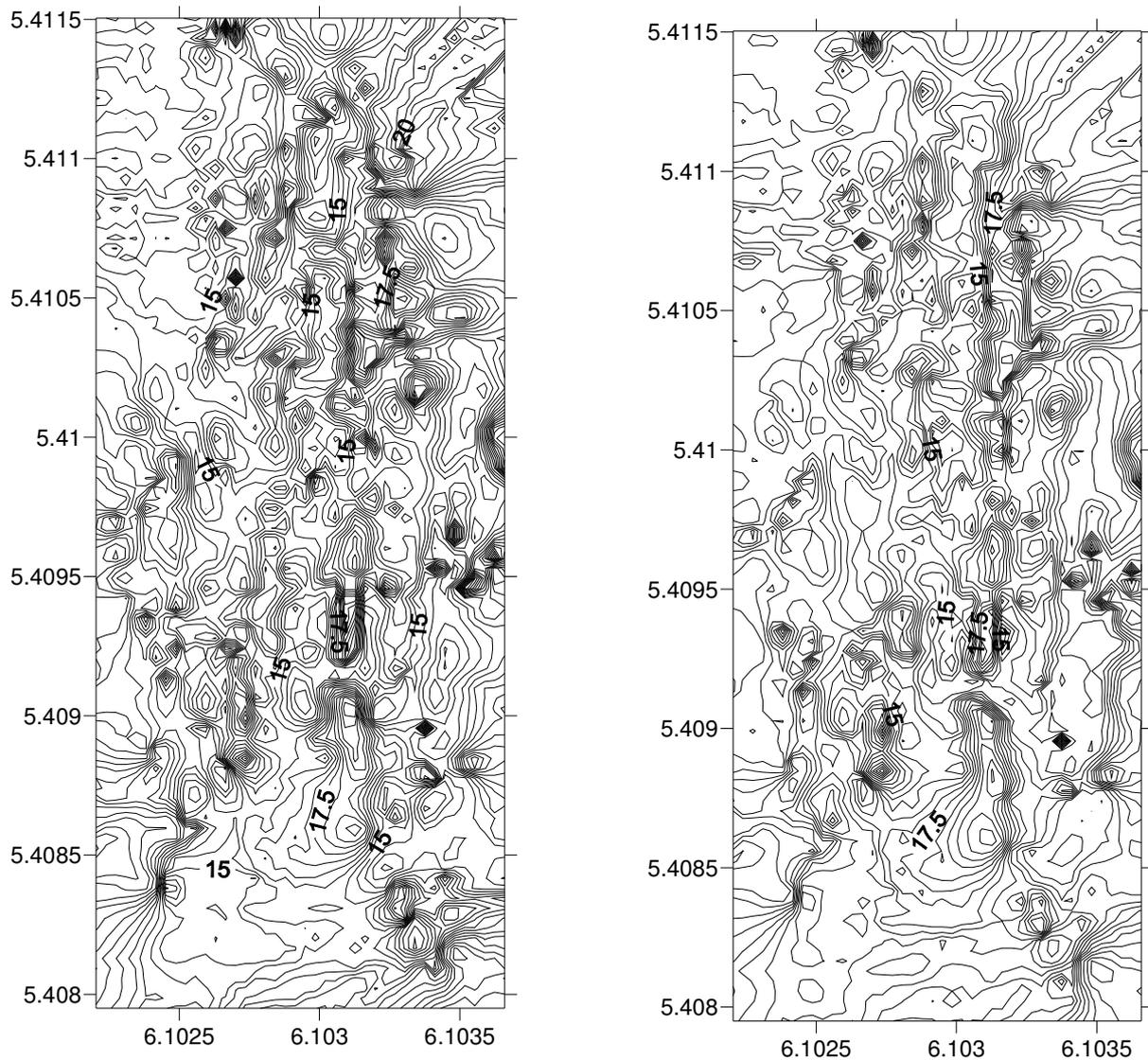


Figura 8.- Región total z real vs z estimada

Se tomó el valor absoluto de la diferencia entre la z real y la z calculada para cada región y los resultados se muestran en la tabla 1. De las seis regiones que se trabajaron, la de los errores más altos fue la dos, que correspondería a la parte inferior derecha de la figura 8 y que es la zona con las CT más altas. Se puede apreciar de la misma tabla 1 que los errores en todas las regiones son altos, sin embargo, este primer intento para tener un modelo numérico que reproduzca, dados los pares coordenados (x_i, y_i) , el valor de la CT se puede considerar bueno, sin pasar por alto que se debe afinar, como un segundo intento, el conjunto de funciones para lograr que la calidad de la estimación mejore.

Los resultados obtenidos son prometedores y se debe continuar la línea de trabajo para lograr mejores resultados que permitan tener un modelo numérico de terreno confiable para futuros estudios.

Tabla 1.- Errores promedio para cada región

Región	error promedio (msnm)
1	0.96
2	1.05
3	0.97
4	0.76
5	0.85
6	0.68

CONCLUSIONES

Los resultados que se obtuvieron al aplicar un modelo numérico basado en las técnicas de la CE, específicamente PG, presentan estimaciones que reproducen de manera satisfactoria la topografía de la región correspondiente al tramo en el que se bifurca el río Mezcalapa. Al comparar las topografías mostradas en la figura 8, las curvas de nivel generadas por el modelo computacional muestran una curvatura mucho más suave, lo cual coincide al observar la traza de la superficie libre del agua con el terreno (margen del río); esto se puede notar al observar las orillas del río en fotos aéreas o en las imágenes de satélite disponibles en internet. No obstante lo anterior, los errores promedio en las 6 regiones estudiadas son altos aún y se debe continuar con el estudio de alternativas para lograr mejorar la calidad de las estimaciones, y con ello tener un modelo numérico que sea confiable.

REFERENCIAS BIBLIOGRÁFICAS

- Arfken, G.** (1980). *Métodos matemáticos para físicos*. Diana.
- Bäck, T.** (1996). *Evolutionary Algorithms in Theory and Practice*. Oxford University Press.
- Banzhaf, W., P. Nordin, R., E. Keller and F. D. Francone.** (1998). *Genetic Programming: An Introduction*. Morgan Kaufmann Publishers.
- García, A. M.A. y Arellano, N. M.** (2007). Contexto, Conflictividad social y Derechos Humanos en Chiapas 2007. Maderas del pueblo del sureste A.C. Disponible en: www.maderasdelpueblo.org.mx
- Goldberg, D. E.** (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley.
- Golden Software Inc.** (2010). Surfer, Versión 8-02. Surface Mapping System. Golden, Colorado, USA.
- Jiménez, A. A., Gracia, J., Berezowsky, M. y Martínez, J.** (2007). “Estudio de la bifurcación de un río con modelación numérica”, *Series del Instituto de Ingeniería*. CI-29. Universidad Nacional Autónoma de México.
- Jiménez, A. A., Gracia, J., Berezowsky, M. y Martínez, J.** (2008). “Funcionamiento hidráulico de la bifurcación Mezcalapa – Samaria – Carrizal”. *Ingeniería hidráulica en México*. Vol. XXIII. Núm. 4, pp. 53-67.
- Jiménez, A. A. y Berezowsky, M.** (2009). “Modelación matemática de la operación de la estructura de control del río Carrizal”. *Serie INVESTIGACIÓN Y DESARROLLO SID/662*. Instituto de Ingeniería. Universidad Nacional Autónoma de México.
- Holland, J.H.** (1975). *Adaptation in Natural and Artificial Systems*. The University of Michigan Press.
- INEGI** (2010). Censo de población y vivienda. Disponible en: <http://inegi.gob.mx/>

- Koza, J. R.** (1992). *Genetic Programming: On the Programming of Computers by Means of Natural Selection*. MIT Press.
- Mendoza, R. y Berezowsky, M.** (1996). “Interpolación de valores con un desarrollo en serie de Taylor de tercer orden”. *Memorias del XVII Congreso Latinoamericano de Hidráulica*, Guayaquil, Ecuador.
- Mendoza, R.** (2002). “Aplicación de la computación evolutiva en la estimación de cotas topográficas”. Tesis de maestría: Programa de Posgrado en Ciencia e Ingeniería de la Computación. Universidad Nacional Autónoma de México, México, D.F.
- Mendoza, R., Vázquez, K. y Álvarez, L.** (2007). “Generación de superficies mediante programación genética”. *Revista Ingeniería Hidráulica en México*. Vol. XXII. No. 4. pp. 21-31. Octubre-Diciembre. México, D.F.
- Rubio, H. y Triana, C.** (2006). *Gestión integrada de crecientes. Caso de estudio México: Río Grijalva*. Editado por Unidad de Apoyo Técnico. Organización Meteorológica Mundial (OMM) y Global Water Partnership (GWP). Disponible en: www.apfm.inf/pdf/case_studies/cs_mexico.pdf
- The Mathworks.** (1999). *Matlab Reference Guide*. The MathWorks Inc.



Instituto de Recursos Hídricos



Facultad de Ciencias Exactas y Tecnologías



Universidad Nacional de Santiago del Estero



Instituto Nacional del Agua



Subsecretaría de Recursos Hídricos



Agencia Nacional de Promoción Cient. y Tec.



Gobierno Prov. de Santiago del Estero



Ministerio de la Producción



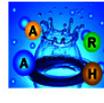
Secretaría del Agua



Secretaría de Desarrollo, Ciencia y Tecnología



Consejo Prof. de la Ingeniería y Arq.



Asociación Argentina de Recursos Hídricos



Asoc. Internacional de Investg. Hidroamb.



Comisión Regional del Río Bermejo



CORPORACION ARGENTINA TECNOLÓGICA S.A. INGENIERIA CIVIL E HIDRÁULICA

