XVI World Water Congress

International Water Resources Association (IWRA)

Cancun, Quintana Roo. Mexico. 29 May- 3 June, 2017

INSTITUTO DE INGENIERÍA UNAM®

# DETERMINATION OF DIGITAL ELEVATION MODEL IN A PLAIN OF MEXICO USING GENETIC PROGRAMMING AND ESTIMATION OF FLOOD VOLUMES.

Mariles Gutiérrez Jacquelinne, Arganis Juárez Maritza Liliana, De Luna Cruz Faustino, Preciado Jiménez Margarita, Carrizosa Elizondo Eliseo.

Instituto de Ingeniería, Universidad Nacional Autónoma de México. Circuito Escolar S/N, Edificio 5,Ciudad Universitaria, Del. Coyoacán, México D.F., México. C.P. 04510. E-mail JMarilesG@iingen.unam.mx.

## INTRODUCTION

The problem of having approximate digital terrain models through mathematical models is relevant in engineering, since the availability of a model of some surface allows to simulate processes and to experiment in the modeled surface independently of the real surface, avoiding risks and also whit the possibility of repeating indefinitely the experimentation.

Since the end of the last century, interest in the study of Digital Elevation Models (DEM) has been increasing due in large part due to the ease of working with the study surface in laboratory tests that may report a behavior similar to that of the real surface in case of extraordinary events or that can predict the behavior by the same events. However, the complexity of the topography makes the models obtained based on their mathematical representation have no more than a symbolic meaning, so in practice, the dimensions of a zone are given by subregions obtained from exclusively applicable equations For that zone. As a consequence, the studies carried out have focused on finding the way in which these models, rather than being obtained from a single mathematical function, represent the study surface more faithfully.

## STUDY SITE.

Villa de Ocuiltzapotlán is located in Tabasco state, México (Figure 1), geographically between 17 ° 15 'and 18 ° 39' north altitude and 91 ° 00 'and 94 ° 07' west longitude. Tabasco is divided into 17 municipalities, one of them is the municipality of Centro, where the city is located and this is one of the 13 regional development centers (RRDC) of the municipality. Villa Ocuiltzapotlán is defined by the historian Rosendo Taracena Padrón as "Tierra de Zapotes", located approximately 18 km north of the state capital, Villahermosa, and forms part of the Metropolitan Area of Villahermosa, presents the appearance of a vast plain severed by stretches by some low hills and is at an average height of 10 meters above sea level, the predominant climate is extremely hot and humid, the average annual temperature is 26 ° C and it presents constant rains, its precipitation is of 2,750 mm.


*Figure 1. Tabasco State, México location. Source: INEGI.*

## METHODOLOGY

Genetic Programming (GP) consists in the automatic evolution of programs using ideas based on natural selection (Darwin), it is itself a tool of evolutionary computation that allows to obtain mathematical models from measured data. By means of a genetic programming algorithm, mathematical models of one or more variables can be determined from previously known data as a method of optimizing an objective function similar to how a simple genetic algorithm works (Arganis et al., 2014).

In general terms, the methodology for generating topographic surfaces using Genetic Programming works in such a way that one or more mathematical functions are applied to the topographic location data (x and y coordinates in meters) to produce an estimate of the elevation Z in meters.

The process begins by choosing a fixed number of individuals (data) to constitute the initial population of the algorithm. This number is less than or equal to the total number of topographic data. The size of the population has to be as little as possible by the resolution time to find the best solution to the problem. The method works through iterations that are applied to all the individuals of the population, where the different levels are estimated applying the indicated operations found by each one of the generations, when a new generation starts the performance of each individual is evaluated, in this way the best individuals preserved will be those with which an estimated quota near the real quota is obtained, the rest of the individuals are applied again the genetic operators with the chosen probabilities. The method continues to iterate until some given error criterion is satisfied or until the maximum number of iterations is reached.

In each iteration of any global method we try to reduce the approximation error for all the points of study, while in GP we take a set of random points of the study region and choose as candidates those that minimize the performance function local. By preserving best-performing points, introducing new random points, and using cross-over and mutation operators over tuning functions, GP ensures that you can explore the search region globally.

*Considered parameters*

- 4 arithmetic functions were considered: +, -, * and /. And two transcendental functions: cosine and sine; a vector of random constants obtained by the algorithm.
- The maximum number of nodes (number estimated between operators and operands) was 25.
- 1000 Generations
- The notation with which the GP is developed is the prefix or Polish, avoiding the use of parentheses in mathematical expressions.

A GP program codified in MATLAB by IIMAS, UNAM was applied.

*Estimation of flood volumes*

Based on the comparison of floodable volumes between the real and the modeled surface, from a Grid of dimensions, in order to obtain a graph of comparison between them.

## RESULTS

The calculated dimensions were obtained by different models applied to topographic data of the Digital Elevation Models (DEM) provided by the National Institute of Statistics and Geography. Figure 1.

**First model**

392,122 points that reproduce the surface, 4 arithmetic operators. Obtaining a function with one rule of correspondence for all the measured data. Figure 2.

$$z = \left(\frac{y * 0.514}{x - A}\right) + \left(\frac{-2.185 - x - y}{x}\right)$$

$$A = -0.121 * (x - 2y + 9.558)$$

The best results showed that the mean error was 9.96 m, so that the calculated levels reported values between 3.99 m and 9.55 m.

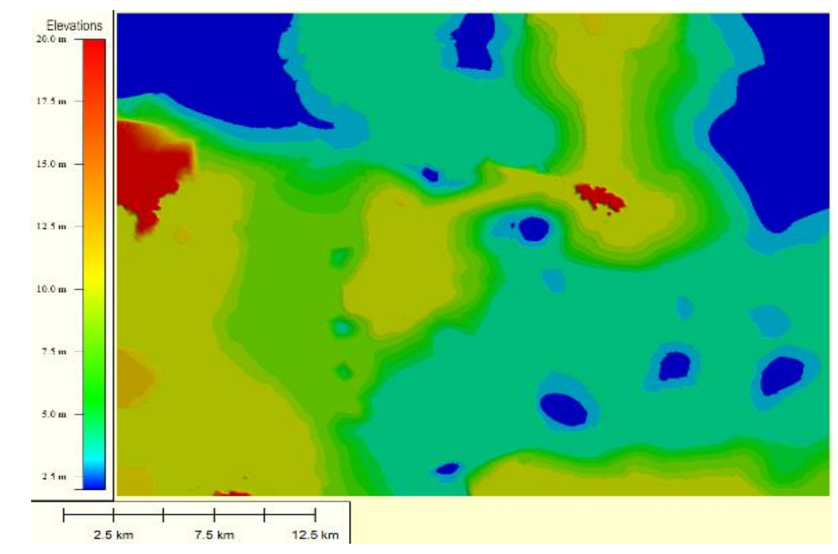



*Figure 2. Ocuiltzapotlán, México. DEM.*

**Second model**

The number of correspondence rules was increased to five to obtain small areas where each of the equations obtained with the algorithm were applicable. Only arithmetic functions were used, it was observed that the error obtained was lower, however the surface generated from the results was not adequate.

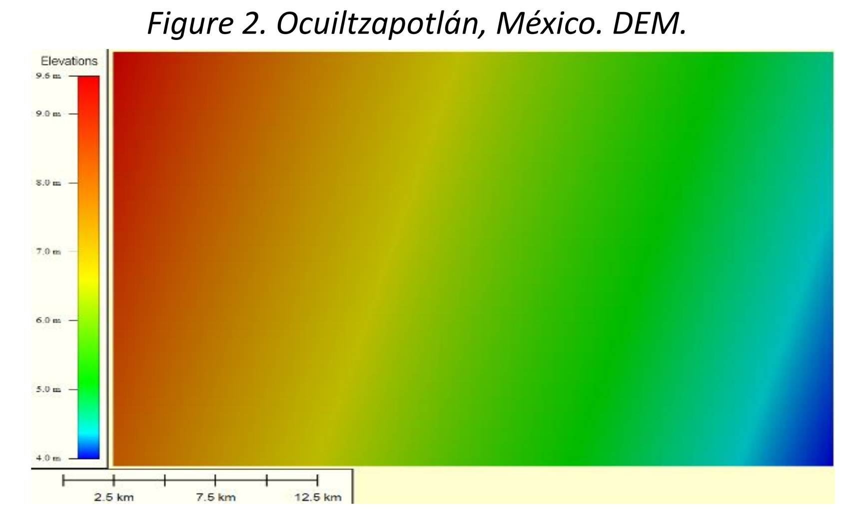



*Figure 3. Area calculated from a rule of correspondence*

**Third model**

Sinusoidal and cosine trigonometric functions were added, so that with the same number of matching rules, new equations were obtained, but using only 3 arithmetic functions (addition, subtraction and multiplication), the result improved the generated surface, nevertheless the error did not present greater improvement, and in some cases, the algorithm used reported constants.

**Fourth model**

The error decreased because of the division of the domain with respect to the obtained errors, the domain was divided into ten correspondence rules. The modeling surface was obtained using three arithmetic functions (addition, subtraction and multiplication) and the two transcendental ones (sine and cosine) for each of the rules of correspondence. In most cases, the best results the algorithm used reported constants, however the average errors obtained were much lower than those obtained previously. Figure 4. The details of the correspondence rules can be consulted with the main author at the Institute of Engineering, UNAM.

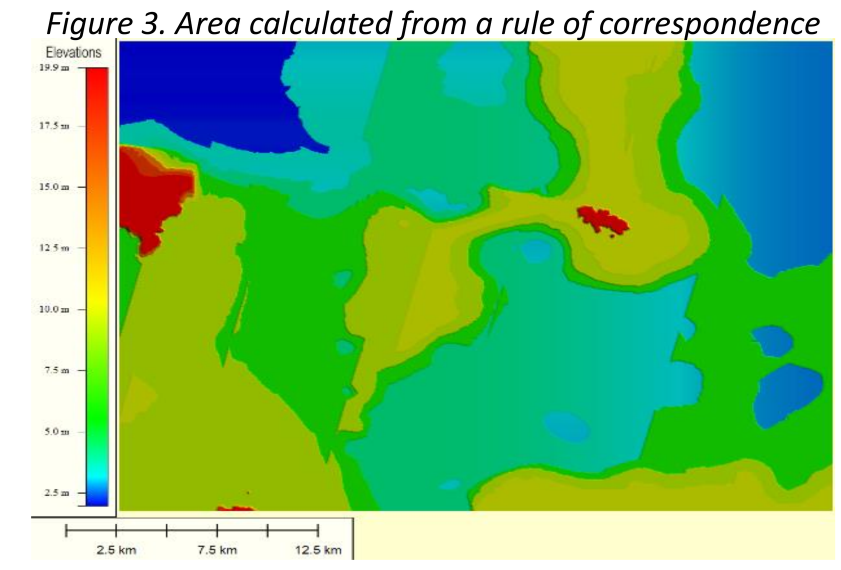
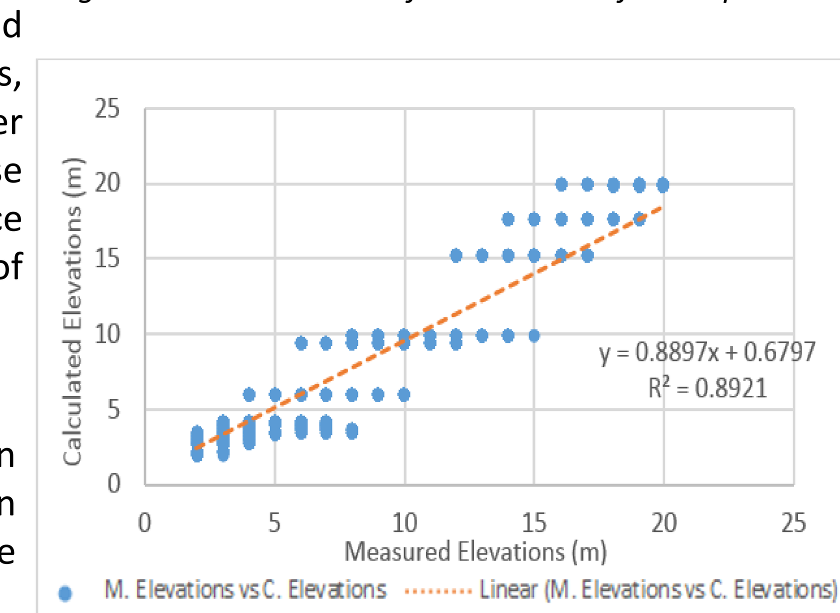



*Figure 4. Area calculated from 10 rules of correspondence*



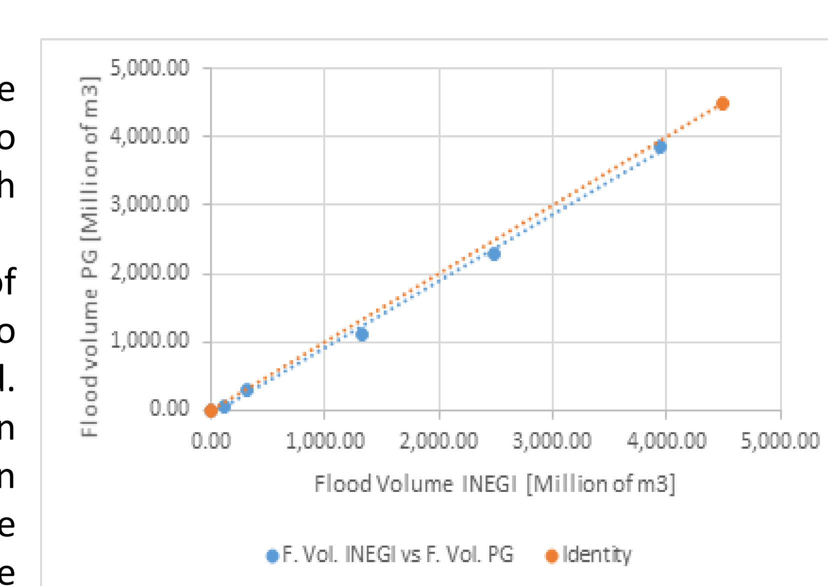

***Flood plans estimation.***

This identification was made by obtaining volumes of flood in both models, from which the following graph was obtained, in which it is observed how the volumes are similar and therefore the reliability of data also.

**Conclusions and recommendations**

The application of Evolutionary Computing (EC) and more specifically Genetic Programming (GP) is a useful tool to estimate the topographical dimensions of the study area, which is a plain.

Working with all data allowed to obtain a single rule of correspondence, which made it difficult to find a model able to reproduce satisfactorily the entire portion of surface analyzed. The restriction of the data to be used for each of the sections in which the study site is divided is fundamental in the resolution and search of a digital model of elevations, due to the large number of points to be considered to represent the surface Analyzed.




The different mathematical expressions obtained are easily programmable for evaluation in any programming language. Because different rules of correspondence were used, it is observed that those where the algorithm uses transcendental functions for the representation of regions within the zone the similarity with the real surface is still much closer to comparison of when the subregions are modeled from of constants obtained with the same algorithm. In spite of this, the maximum errors obtained with GP were values of 10 cm. With this surface and with the original it was possible to obtain the comparison of elevation dimensions that helped to obtain a difference of elevations that is useful to locate the areas of flooding from this difference and that will serve in later studies to have the certainty of the handling of a digital elevation model with respect to the actual surface.

**References:** Arganis J. M. L., Preciado J. M. E., Herrera A. J. L., Rodríguez V. K. (2014) Función Bivariada de avenidas del conjunto Temascal – Cerro de Oro con Programación Genética, XXIII Congreso Nacional de Hidráulica, Jalisco, México. Mendoza, R., Rodríguez, K. y Álvarez-Icaza, L. (2007) Generación de superficies mediante programación genética, Ingeniería Hidráulica en México. UNAM.