ISSN 1996-3343

Asian Journal of Applied Sciences

http://knowledgiascientific.com

Knowledgia SCIENTIFIC
A Place to Publish Outstanding Research

Asian Journal of Applied Sciences 5 (4): 224-231, 2012
ISSN 1996-3343 / DOI: 10.3923/ajaps.2012.224.231

Graph-based Cellular Automata for Simulation of Surface Flows in Large Plains

Pablo Rafael Rinaldi, Diego David Dalponte, Marcelo Javier Vénere and Alejandro Clausse
Conicet-cnea and Universidad Nacional del Centro, 7000 Tandil, Argentina

Corresponding Author: Pablo Rinaldi, Instituto Pladema, Uncpba, Pinto 399, Tandil (7000), Argentina Tel: +54 (2293) 439690 Fax: +54 (2293) 439690

ABSTRACT

A cellular automata supported by a graph structure was developed for simulating surface flows in planes. Differing from the classical regular grids used by most cellular automata the present model is supported by a graph structure which provides higher flexibility to represent complex geographic domains with different scales. The model calculates the water transfer between adjacent cells based on a hydraulic equation assigned to the arcs of the graph. The model includes infiltration and evaporation sources in each cell. Comparisons against hydraulic measurement performed during flooding events show excellent agreement.

Key words: Cellular automata, graphs, simulation, fluids, surface flow

INTRODUCTION

The purpose of surface flow models is to study the causes and attempt to predict the onset of flood events prior to their occurrence. This has been greatly enhanced recently by the availability of new sources of high-resolution data. High-resolution hydrologic data, radar digital elevation models, soil and land-use data are now becoming readily available to modelers in many countries. National Weather Services worldwide are increasingly encouraging the use of this spatially distributed data to improve flow modelling and prediction along river systems.

Hydrologic models suitable for high-resolution models have been under study for the past several decades (Betson, 1964; Dunne and Black, 1970; Schultz, 1988; Michaud and Sorooshian, 1994; Olivera and Maidment, 1999). However, the modeling of surface flows in large plains is not a straightforward task mainly due to the intervention of several elements, namely, accumulation of water on local depressions, absence of an integrated drainage network, generalized meagre surface slopes, sensitivity of the soil water content before rainfall events and variation of the soil infiltration capacity. For those reasons, the availability of fast and validated numerical models of surface flow is an important issue.

Cellular Automata (CA) are interesting candidates to model surface flows. CA technology has been successfully applied to simulate a variety of flows, such as volcanic lava (Crisci *et al.*, 2004) surface flow (Rinaldi *et al.*, 2007) and fluid flows (Rinaldi *et al.*, 2011). Most of these applications represent the underlying space as a regular grid, whereas the real spatial domain calls for more flexible data structures. In effect, the use of rigid representations imposes a number of limitations on the models, such as the need to include extra cells to account for boundary conditions, inefficient management of CPU time and memoryand lack of flexibility to represent irregular regions.

In the present study, a special CA algorithm based in graph techniques is presented, aimed to simulate surface water flows in plains and rivers. The terrain and its water inventory are represented by the nodes and arcs of an undirected graph. The water movement is calculated based in the height differences between nodesand the soil resistance is modelled by means of the cost of the arcs. A clustering algorithm previous to the simulation creates a graph data structure representing inputs based in regular grids in more efficient ways, contributing substantially to reduce the number of equations to be solved.

MATERIALS AND METHODOS

The graph automata: Any model of surface flows over plains should incorporate a good spatial description of terrain slopes and depressions whichin turn requires accurate digital elevation models (DEM) describing the physical terrain by a height scalar field, h (x,y), usually associated to a two-dimensional regular grid of cells. Following the CA paradigm, the state of a cell is determined by a scalar w (x,y), representing in this case the water level in the each cell. Let us represent the spatial domain by a graph where each node corresponds to a cell of the grid containing the information about the water level and the water sources and sinks associated to each cell (i.e., infiltration, precipitationand inflows and outflows due to seepage or evapotranspiration). Each cell of the grid is associated to a node of the graph connected by arcs between its neighbors in the Cartesian direction: e, w, n, s (Fig. 1), representing the flow between cells.

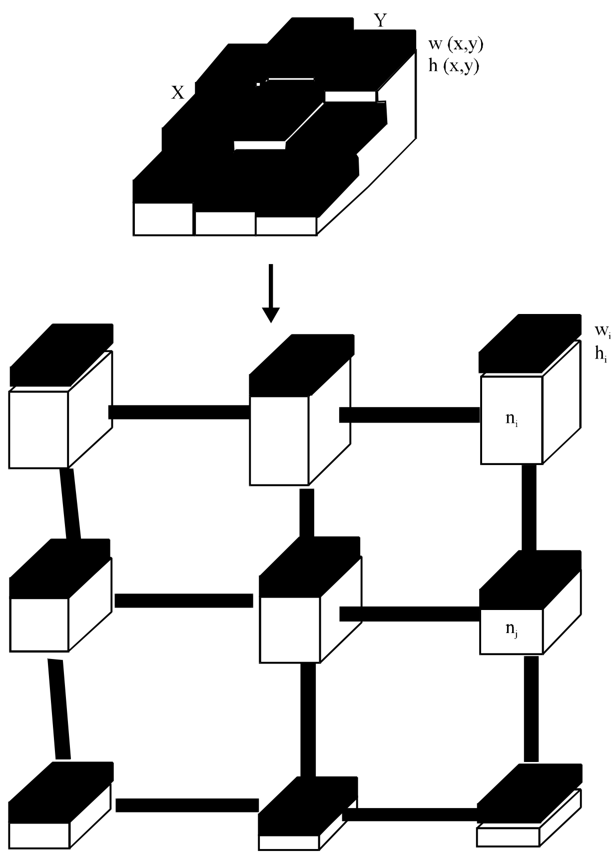

Fig. 1: Transformation scheme of a regular grid into a graph structure

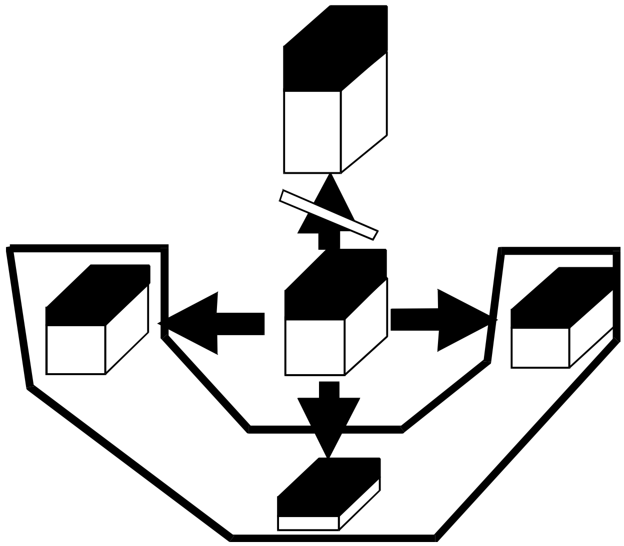

Fig. 2: Example of local basin of one node (center)

The following set of rules defines the state of each node at each step:

The local basin B_i of each node i is the set that contains those of its neighbours n_i whose water levels w are equal or lower than the water level of node i (Fig. 2):

$$n_i \in B_i \Leftrightarrow w_{ni} \le w_i \quad (1)$$

The drain height d_i of each local basin B_i is the height that the water would reach provided that all the liquid drains to its minimum level within the basin.

The maximum quantity of water that can be drained from the node i is given by:

$$q_i = \min (w_i - h_i, w_i - d_i) \quad (2)$$

The actual amount of water transferred from the node i to the neighbour j of the basin is given by:

$$q_{ij} = c_{ij} f_{ij} q_i \quad (3)$$

where, c_{ij} is the cost of the arc representing the soil resistance to the flow ($0 \le c_{ij} \le 1$) and:

$$f_{ij} = \frac{\sqrt{w_i - w_j}}{\sum_{j \in B_i} \sqrt{w_i - w_j}} \quad (4)$$

is a hydraulic coefficient accounting for the driving force.

Once, all q_{ij} are calculated over the graph, the water is transferred in parallel through every arc, end the process starts again.

Water sources: Rains and infiltrations are modelled as sources and sinks. The local precipitation amount in every step is simply added to the corresponding cell, representing observed or hypothetical data. The infiltration process is more complicated since the water amount removed in each step depends on the state of soil saturation. In order to simulate the saturation effect, the amount of water infiltrated I(x,y) should be tracked for each cell according to:

$$I_n(x,y)=\begin{cases}0 \quad \text{si} \quad w(x,y)=0\\ w(x,y) \quad \text{si} \quad 0<w(x,y)<\beta I_{n-1}(x,y)+I_o(x,y)\\ \beta I_{n-1}(x,y)+I_o(x,y) \quad \text{si} \quad w(x,y)\geq \beta I_{n-1}(x,y)+I_o(x,y)\end{cases} \tag{5}$$

where, $I_o(x,y)$ is a bias infiltration and β is a coefficient representing soil saturation characteristics ($0<\beta<1$). Smaller β values mean that the soil would approach saturation faster. The initial infiltration, I_1, should be provided by the analyst and it is likely to expect a dependence on the initial soil conditions.

Modeling of creeks and rivers: Creeks and rivers are simulated by reducing the flow resistance along the corresponding water courses. Accordingly, the cost c_{ij} of every arc located along a river path is calculated as a function of the average water level of the cells it connects whichaccounts for the influence of the shape of the river bed on the flow rate. Since there are a number of factors affecting this relation (bed profile, soil characteristics, aquatic vegetation, curvature, etc.) a comprehensive model would require the definition of a function for each arc. However, this is practically impossible when modeling large extensions of terrain. Alternatively, one can define regional families, whose parameter c can be determined by comparing the numerical calculation with experimental data. A function family that showed good agreement with flow measurements in southern streams of the Argentine Humid Pampas is the following (Rinaldi *et al*., 2007):

$$c_{river}=\left[1+\left(\frac{H_o}{W}\right)^p\right]^{-n} \tag{6}$$

where, W is the average water level for de cells connected and Ho, p and n are effective positive defined constant parameters.

Boundary conditions: Open boundary conditions can be implemented by adding an auxiliary border of cells with elevations substantially lower that the adjacent terrain. The water volume contained in the added border is eliminated after each calculation step to avoid long term accumulations. In order to achieve good global mass balances, it is very important that the simulation domain is much larger than the basin under study, leading to the automatic determination of the basin boundaries.

Different time scales: Graph structures allow the definition of node classes with different behaviors. This property is used to introduce different time scales in each class, such as rivers and terrain. The purpose of this feature is to match the numerical and the real time scales. In general, the flow in the water courses is much faster than the surface flow in floodplains. Time scales are determined by partitioning the main graph in two sub-graphs shearing arcs that

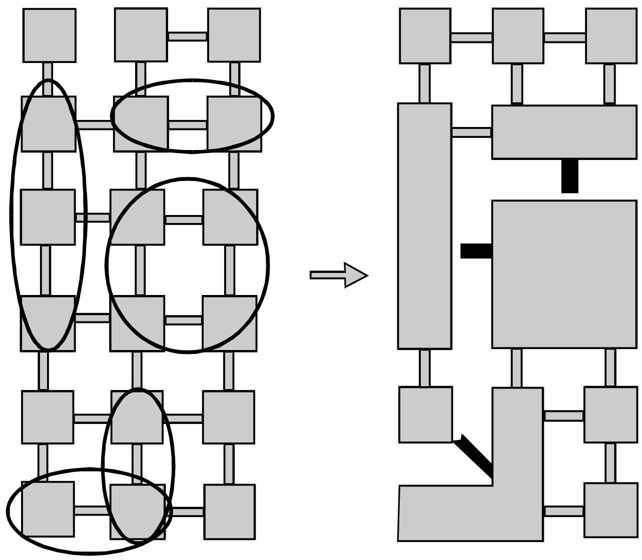

Fig. 3: Node grouping scheme, wider connections has more transferring capacity

connect both classes. The water distribution rules operate in each sub-graph separatelyand they are synchronized on specified periods. Since the river sub-graph represents less than 1% of the total, this procedure greatly improves the efficiency of the calculation.

Construction of the supporting graph: The geographic data of the terrain usually comes as a DEM supported by a regular 2 D grid. Starting from this data, adjacent cells having similar characteristics can be merged in a single node. This procedure is performed automatically by an algorithm that group adjacent cells according to certain similarity criteria (Fig. 3). The resulting new node has the following characteristics:

- The new area is the summation of the areas of the coalesced cells
- The height of the new node is the average of the coalesced cells weighted with the area
- The water contained in the new node is the summation of the water stock of the coalesced cells

The criterion used to group adjacent cells in a single node is based in the accumulation of a parameter that takes into account the local gradient and curvature. Accordingly, each cell (x,y) is associated to a parameter:

$$C(x,y)=\frac{K(x,y)}{G(x,y)} \quad (7)$$

where, G and K are metrics of the local gradient and curvature, calculated as:

$$G(x,y)^2 = G_x(x,y)^2+G_y(x,y)^2 \quad (8)$$

$$K(x,y)^2 = [G_x(x+1,y) - G_x(x-1,y)]^2 + [G^x(x,y+1) - G_x(x,y-1)]^2 + [G_y(x+1,y) - G_y(x-1,y)]^2 + [G_y(x,y+1) - G_x(x,y-1)]^2 \quad (9)$$

and:

$$G_x(x,y) = h(x+1,y)-h(x-1,y) \quad (10)$$

$$G_y(x,y) = h(x,y+1)-h(x,y-1) \quad (11)$$

The value of the parameter C of a graph node is the summation of the C values of the coalesced cells. The criterion to construct the graph is that all nodes should have approximately the same C. In this way, flatter regions (i.e., low G) with small depressions (i.e., high K) are collected in a single node, whereas steeper slopes (i.e., high G) and large depresions (i.e., low K) keep high resolution. The appropriate value of C per node should be determined in each case looking for a compromise between computation speed and agreement with available experimental data.

RESULTS

The model was implemented in a C++ code and was applied to simulate several flooding events that took place in 2002 in a Pampean region in Argentina which combines flat plains with lower hills. The region is located between latitude south 36°8 and 37°22 and between longitude west 58°49 and 60°10, approximately 150 km in Southwest-Northeast direction and occur 40 km wide. Figure 4 shows the original DEM of the region, with 58394 cells. The graph representation for the automata has 27265 nodes, that is 45% of reduction respect to the DEM.

The model was tested against major flood events that took place in May, August and October of 2002. Figure 5, 6 and 7 show the hyetographs of precipitations in bars corresponding to the averaged water millimeters per hour. The solid points in the graphics represent the water flow discharged by the creek to the main river. The October event was a typical spring rain,

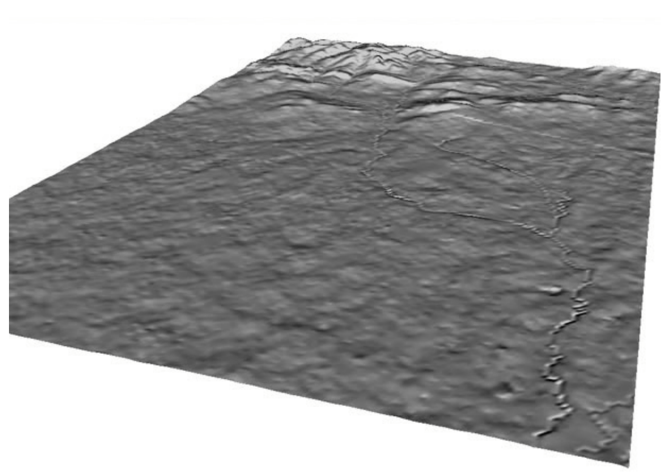

Fig. 4: Study area DEM built using radar interferometer

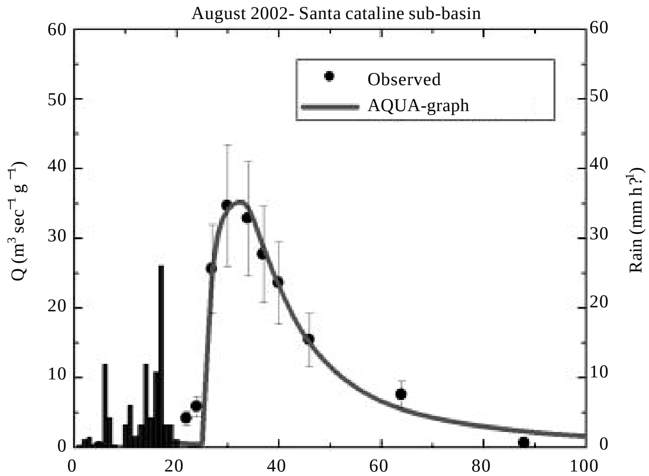

Fig. 5: Measured (dots) and simulated (curve) outlet flow ratio of Santa Catalina Creek during the event of October 2002. The bars indicate the measured precipitation rate

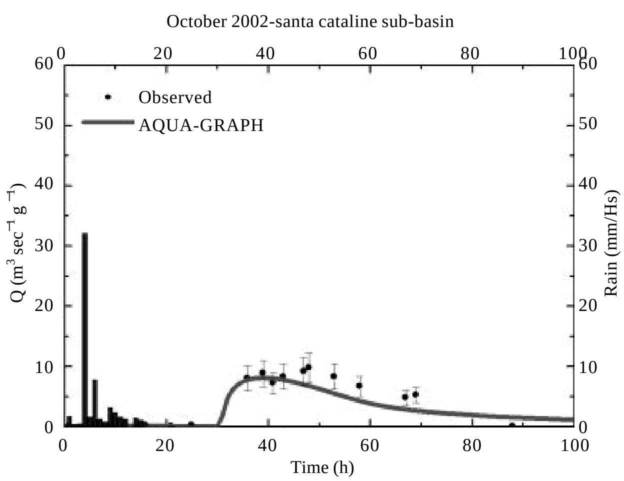

Fig. 6: Measured (dots) and simulated (curve) outlet flow ratio of Santa Catalina Creek during the event of August 2002. The bars indicate the measured precipitation rate

relatively mild and short. Figure 5 shows the hourly precipitation (bars) occurring during the first 20 h of the event. The points show the hydrograph of the downstream flow wave. The curve represents the numerical simulation of the flow wave response to the precipitation input. The August case is a typical winter event, with continuous precipitations during a certain period of time. Figure 6 shows the hourly precipitation (bars) during the first 25 h of the event, with a mild recurrence between 30 and 40 h. The points and the curve show the measured downstream flow wave and the corresponding numerical simulation. The May event is interesting, for there were actually two consecutive rains separated by 80 h. The first rain acted by changing the humidity conditions for the second rain, saturating the absorption capacity of

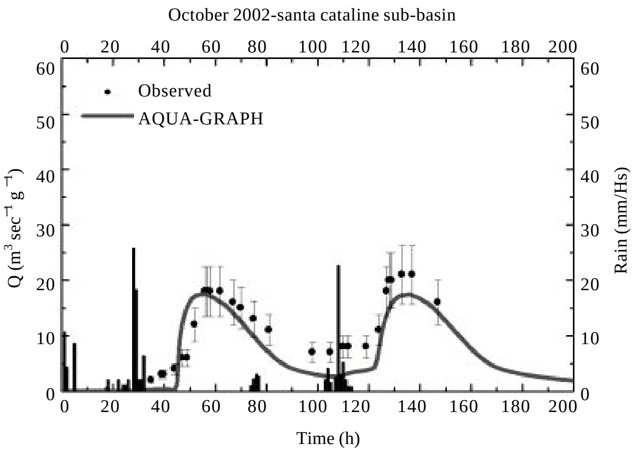

Fig. 7: Measured (dots) and simulated (curve) outlet flow ratio of Santa Catalina Creek during the event of May 2002. The bars indicate the measured precipitation rate

the soil. It can be seen in Figure 7 that the infiltration algorithm presented here is able to handle the whole event reproducing both waves with excellent agreement.

CONCLUSIONS

A cellular automata model of surface flows in plains based in a graph representation was presented. The graph is contructed from ordinary geographic data supported in regular grids, using a criterion that takes into account local gradient and curvature. The model has been able to reproduce the measured channelled flow from single and multi-pulse rainfall events at a basin located in the centre of Buenos Aires Province.

REFERENCES

Betson, R.P., 1964. What is watershed runoff? J. Geophys. Res., 69: 1541-1552.

Crisci, G.M., R. Rongo, S. Di Gregorio and W. Spataro, 2004. The simulation model SCIARA: The 1991 and 2001 lava flows at Mount Etna. J. Volcanol. Geothermal Res., 132: 253-267.

Dunne, T. and R.D. Black, 1970. Partial area contributions to storm runoff in a small New England watershed. Water Resour. Res., 6: 1296-1311.

Michaud, J. and S. Sorooshian, 1994. Comparison of simple versus complex distributed runoff models on a midsized semi-arid watershed. Water Resour. Res., 30: 593-605.

Olivera, F. and D.R. Maidment, 1999. Geographic information systems based spatially distributed model forAC runoff routing. Water Resour. Res., 35: 1155-1164.

Rinaldi, P.R., D.D. Dalponte, M.J. Venere and A. Clausse, 2007. Cellular automata algorithm for simulation of surface flows in large plains. Simul. Modell. Pract. Theory, 15: 315-327.

Rinaldi, P.R., E.A. Dari, M.J. Venere and A. Clausse, 2011. Lattice-boltzmann Navier-stokes simulation on graphic processing units. Asian J. Applied Sci., 4: 762-770.

Schultz, G.A., 1988. Remote sensing in hydrology. J. Hydrol., 100: 239-265.