Hydrological Sciences Journal

ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage: https://www.tandfonline.com/loi/thsj20

Cellular automata algorithms for drainage network extraction and rainfall data assimilation

ERIKA COPPOLA , BARBARA TOMASSETTI , LAURA MARIOTTI , MARCO VERDECCHIA & GUIDO VISCONTI

To cite this article: ERIKA COPPOLA , BARBARA TOMASSETTI , LAURA MARIOTTI , MARCO VERDECCHIA & GUIDO VISCONTI (2007) Cellular automata algorithms for drainage network extraction and rainfall data assimilation, Hydrological Sciences Journal, 52:3, 579-592, DOI: 10.1623/hysj.52.3.579

To link to this article: https://doi.org/10.1623/hysj.52.3.579

Hydrological Sciences–Journal–des Sciences Hydrologiques, **52**(3) June 2007
Special issue: *Hydroinformatics*

Cellular automata algorithms for drainage network extraction and rainfall data assimilation

ERIKA COPPOLA, BARBARA TOMASSETTI, LAURA MARIOTTI, MARCO VERDECCHIA & GUIDO VISCONTI

CETEMPS, Physics Department, University of L'Aquila, Via Vetoio, I-67010 Coppito l'Aquila, Italy
erika.coppola@aquila.infn.it

Abstract Since 2002, within the framework of the Cetemps Centre of Excellence at the University of L'Aquila, a distributed grid-based hydrological model (CHYM) has been developed to provide a general purpose model for operational flood warning activity. This paper presents two new cellular automata (CA) algorithms used respectively for drainage network extraction and rainfall data assimilation. The first is a cellular automaton-based algorithm for the extraction of a drainage network from an arbitrary digital elevation model. It has been implemented and tested on a large number of different domains. This algorithm is able to define the flow direction at every point on the digital elevation model where singular points are present (pits or flat areas). The second is a CA-based numerical technique for assimilating different data sources of rainfall to rebuild the rainfall field on a grid. This technique has been shown to produce a reasonable rainfield shape without any geometrical artefacts that often produce unrealistic rain gradients in the rainfield.

Key words hydrological model; cellular automata; CHYM; drainage network extraction

Algorithmes d'automates cellulaires pour l'extraction du réseau de drainage et l'assimilation de données pluviométriques

Résumé Depuis 2002, dans le cadre du centre d'excellence Cetemps à l'université de L'Aquila, un modèle hydrologique distribué maillé (CHYM) a été développé pour fournir une modélisation générale appliquée à l'alerte opérationnelle de crue. Cet article présente deux nouveaux algorithmes d'automates cellulaires (AC) utilisés respectivement pour l'extraction du réseau de drainage et l'assimilation de données pluviométriques. Le premier est un algorithme basé sur un automate cellulaire qui permet l'extraction du réseau de drainage à partir d'un modèle numérique d'altitude (MNA) arbitraire. Il a été mis en application et testé sur un grand nombre de domaines différents. Cet algorithme peut définir le sens d'écoulement en chaque point du modèle numérique d'altitude où sont présents des points singuliers (sommets ou terrains plats). Le second est une technique numérique à base d'AC pour assimiler différents sources d'information pluviométrique afin de reconstruire le champ de pluie sur une grille. Il a été vérifié que cette technique produit une forme raisonnable du champ de pluie, sans artefacts géométriques qui produisent souvent des gradients irréalistes au sein du champ de pluie.

Mots clefs modèle hydrologique; automates cellulaires; CHYM; extraction du réseau de drainage

1 INTRODUCTION

Coupling hydrological and climate models is becoming an important issue for assessing the effects of climate change on the hydrological cycle. Moreover, coupling a hydrological model with a numerical weather prediction model can provide important information for flood warning and hydrological risk management. Spatially-distributed hydrological models are able to account for the spatial variability of physical processes involved in the hydrological cycle for which an accurate representation of the internal drainage structure of the watershed is needed.

One of the major problems in producing a realistic representation of the drainage network of a river basin is the necessity to solve numerical singularities due to the finite resolution of the digital elevation model (DEM). These kinds of singularities occur particularly along the main segments of the river basin where the topography can be very flat, and, in many cases, it is not trivial to establish the flow direction for all the cells of the domain. Different approaches have been proposed to overcome this problem, e.g. Jenson & Domingue (1988) and Garbrecht & Martz (1997). In this paper

the CA2CHYM algorithm based on the concept of cellular automata (CA) is described. It is also demonstrated how this approach represents a powerful technique to solve numerical singularities, even in very complex topography.

Once the flow direction for each cell has been determined, an appropriate approach is needed to check if the drainage network has been rebuilt correctly. This is achieved by implementing two complementary algorithms in a sequence: the rolling stones algorithm (RSA) or tipping bucket algorithm, which is used to calculate the flow accumulation matrix, followed by the salmon algorithm (SA), which starts at the mouth of the river and moves upstream, rebuilding the river path at each step by choosing the cell in the neighbourhood with the greatest total drained area. The application of these two algorithms rebuilds the complete drainage network.

A second CA algorithm is introduced for assimilating and downscaling several rainfall data sets derived from meteorological model output, observations and rainfall estimation techniques. Different methods have been proposed in the literature for downscaling precipitation fields, e.g. those of Xie & Arkin (1997) and Van der Linden & Christensen (2003) for climate scale resolution, and several techniques have also been presented to combine different sources of precipitation estimates in order to reduce the final product bias and uncertainty (French & Krajevsky, 1994; Todini, 2001). The algorithm presented here deals with a hierarchical acquisition of different precipitation data sources.

In the operational configuration of the CHYM model, the raingauge data, the radar and satellite estimates and the MM5 meteorological model output field are combined using a CA-based algorithm. This algorithm creates a final rainfield as a result of merging the data, establishing a hierarchy among them. This is suitable in an operational context, where different data sets are available at different temporal resolutions. The sequence of modules starts from the one that uses the most reliable data set (e.g. a tested raingauge network) and it continues until it reaches the data set with the greatest expected error (e.g. predictions from a meteorological simulation).

A detailed description of the CA-based algorithms and their application is given in Section 2. Conclusions are reported in Section 3.

2 CA-BASED ALGORITHMS FOR DRAINAGE NETWORK EXTRACTION AND RAINFIELD RECONSTRUCTION

2.1 CA2CHYM: an algorithm for flow direction extraction

For extracting the drainage structure from a raster DEM, the majority of the extraction algorithms use an eight flow direction (D8) approach (e.g. Tribe, 1992; Jenson & Domingue, 1998; Martz & Garbrecht, 1992, 1993, 1995; Fairfield & Leymarie, 1991). Other algorithms include digital information about rivers and lakes in order to compute a more realistic drainage network (Turcotte *et al*., 2001). Both are valid approaches to modelling the watershed drainage structure.

In a distributed grid-based hydrological model, the first step for rebuilding the drainage network from a DEM matrix consists of establishing the flow direction of each cell in the domain. According to the minimum energy principle, this is usually done by choosing the flow direction as the one corresponding to the maximum

downhill slope. Unfortunately, the application of this algorithm is not trivial, because of the presence of many singularities due to the finite resolution of the DEM, which makes it impossible to establish the flow direction when there are pits (cells where the elevation is at a minimum with respect to the eight neighbouring cells) or flat zones (characterized by sequences of cells with the same elevation). Many algorithms have been proposed to solve these kinds of problems (e.g. O'Callagan & Mark, 1984; Jenson & Domingue, 1988; Garbrecht & Martz, 1997; Martz & De Jong, 1998).

The algorithm, referred to here as CA2CHYM, is based on CA concepts and allows for a coherent definition of flow direction, even in complex topography, without an unrealistic modification of the DEM matrix. The CA2CHYM algorithm can be summarized as follows:

- The original DEM is used for establishing the flow direction except for the grid points where singularities occur.
- The whole DEM is then smoothed using equation (3) until all the singularities are solved and the whole streamflow network is established.
- Once the flow direction is established for each cell, the original DEM is used and only the height of those cells draining toward a cell with greater elevation is modified.

According to Packard & Wolfram (1985), CA are a simple mathematical idealization of natural systems, inspired by the observation of the collective behaviour that many natural systems made of many singular parts appear to have. They can be described as a lattice of discrete identical sites and the status of the sites evolves in discrete time steps according to deterministic rules depending on the values of neighbouring cells. They can also be considered as discrete idealizations of partial differential equations and, given their discrete nature, they can be viewed as being similar to simple parallel-processing devices. Cellular automata are suitable for describing mathematical models of systems in which many simple components act together to produce complicated patterns of behaviour.

In our application, we consider the DEM matrix as a two-dimensional aggregate of CA where the elevation of each grid point (cell) represents the status of each cellular automaton. The value of the central cell depends only on the current value of the cell and the values of the cells in the neighbourhood; this smoothing rule is performed several times until all the singularities are corrected, and all cells are updated synchronously. To explain this, let us consider a regular lattice of sites where each of them takes on n possible values. The lattice is updated in discrete time steps according to a rule θ that depends on the current status of the site and on the eight neighbouring cells. The value x of the site evolves according to:

$$x^{t+1} = \theta\left(x^t, x_k^t\right) \quad k = 1, \ldots, 8 \tag{1}$$

For our particular application, the rule acts to smooth the DEM singularities and it modifies the elevation of each cell slightly by taking into account the average of the surrounding grid points. At each step $t + 1$, the elevation x of each point is updated as follows:

$$x^{t+1} = x^t + \beta \sum_{k=1}^{8} \alpha_k x_k{}^t \tag{2}$$

where β is a small coefficient and the sum is carried out over the eight neighbouring cells. The coefficients α take into account the different distances between grid points and therefore the rule in equation (2) can be rewritten as:

$$x^{t+1} = x^t + \beta \sum_{k=1}^{8} \frac{1}{r_k} x_k^t \tag{3}$$

where r_k is the distance between the central cell and the neighbouring cell. The CA-based smoothing algorithm is iteratively applied to all cells until all singularities are solved. The smoothed DEM is then used to establish flow directions only for those grid points where there are singularities. For all the others, the steepest descent method is used to define the flow direction.

Figure 1 shows an example of how the CA2CHYM algorithm works. In Fig. 1(a) and (c), the DEM height values in metres are shown for a small sub-domain as they are originally and the flow direction is indicated with an arrow for cells where this is possible to establish. These examples show a few cases of flat areas where it is not possible to detect the downward slope and several cases of pits. In Fig. 1(b) and (d) the same DEM region is shown after the CA2CHYM algorithm has been applied. The flat regions have disappeared and the elevation values of singular cells have changed according to the elevation of the neighbouring area. The arrows now indicate that the

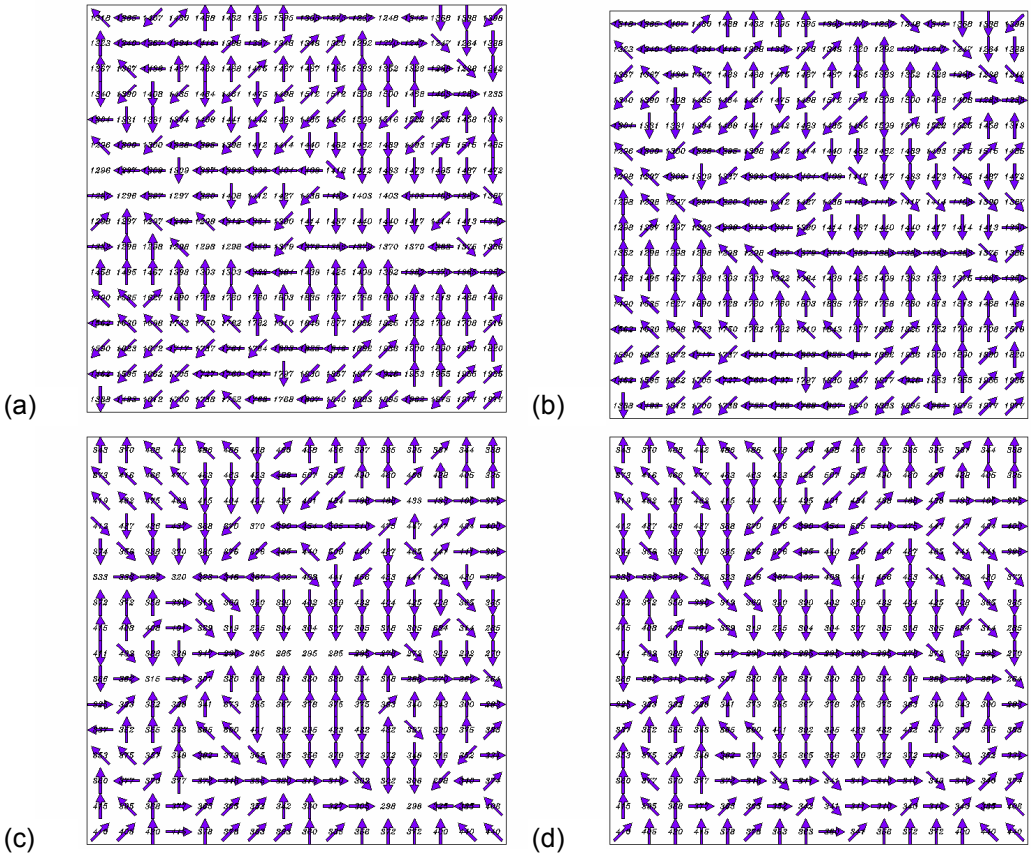

Fig. 1 (a) and (c): DEM height values (in m) as they are originally reported; the flow direction is indicated with an arrow for the cells where it is possible to establish it. Panels (b) and (d) contain the same DEM region after the CA2CHYM algorithm has been applied.

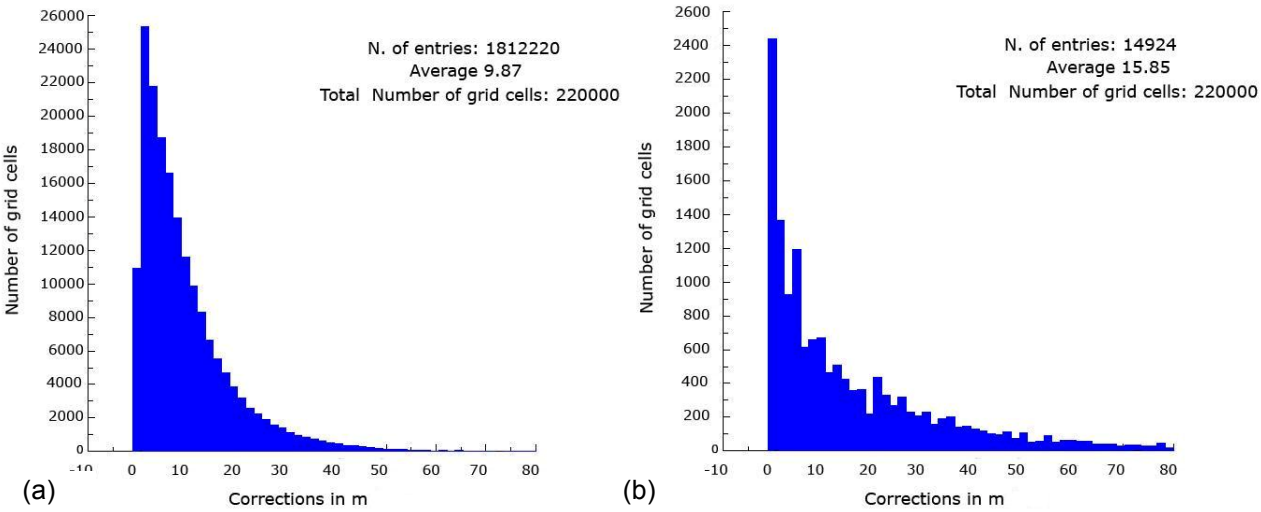

Fig. 2 Histograms of the number of DEM corrections in metres (a) when a classical algorithm (such as that of Martz & Garbrecht, 1992) has been used to solve the singularities and (b) when the CA2CHYM algorithm is applied.

flow is defined for every cell. In the same way the pits are upgraded to a value that is established by the neighbouring cells. The application of the algorithm deals with establishing the flow direction not only for singular cells, but in many cases also for neighbouring cells to obtain a coherent structure.

It is important to highlight that, with the three steps of the CA2CHYM algorithm, a minimum number of cell modifications is needed which avoids unrealistic modifications of the DEM matrix to solve the singularities. In Fig. 2, a comparison is shown between a classical filling algorithm (Marks *et al*., 1984; Martz & Garbrecht, 1992; Band, 1986) and the CA2CHYM algorithm. Note that in Fig. 2(b) the number of modified cells is one order of magnitude lower (but the average correction is somewhat greater).

Figure 3 shows the grid points of the selected domain where major corrections to the DEM have been made. The zone indicated by A corresponds to an artificial lake (Lake Campotosto); that indicated by B corresponds to the di Stiffe caves, where a river source is inside the mountain; and that indicated by C corresponds to a large flat area where a lake was present until it was drained about two centuries ago (Tomassetti *et al*., 2003).

Once the flow direction has been established for each cell and a corrected DEM has been produced, a multi-purpose tool to "draw" the drainage network for an arbitrary domain has been developed. One of the most important features is the possibility to check whether a drainage network is realistically reproduced by the model. Other important features of this tool are: the possibility to verify that a catchment is entirely contained in the selected domain; the possibility to distinguish cells where surface runoff occurs along the channel; and also the possibility to derive values of hydraulic parameters using quantities such as the total drained area of a cell. The algorithm can also take into account information from the digitized coarse drainage network from existing maps, if available.

In the CHYM model, the drainage network is rebuilt in two steps: first, the accumulation matrix is calculated, while, in the second phase, the salmon algorithm "tracks" the drainage network both numerically and graphically. The accumulation

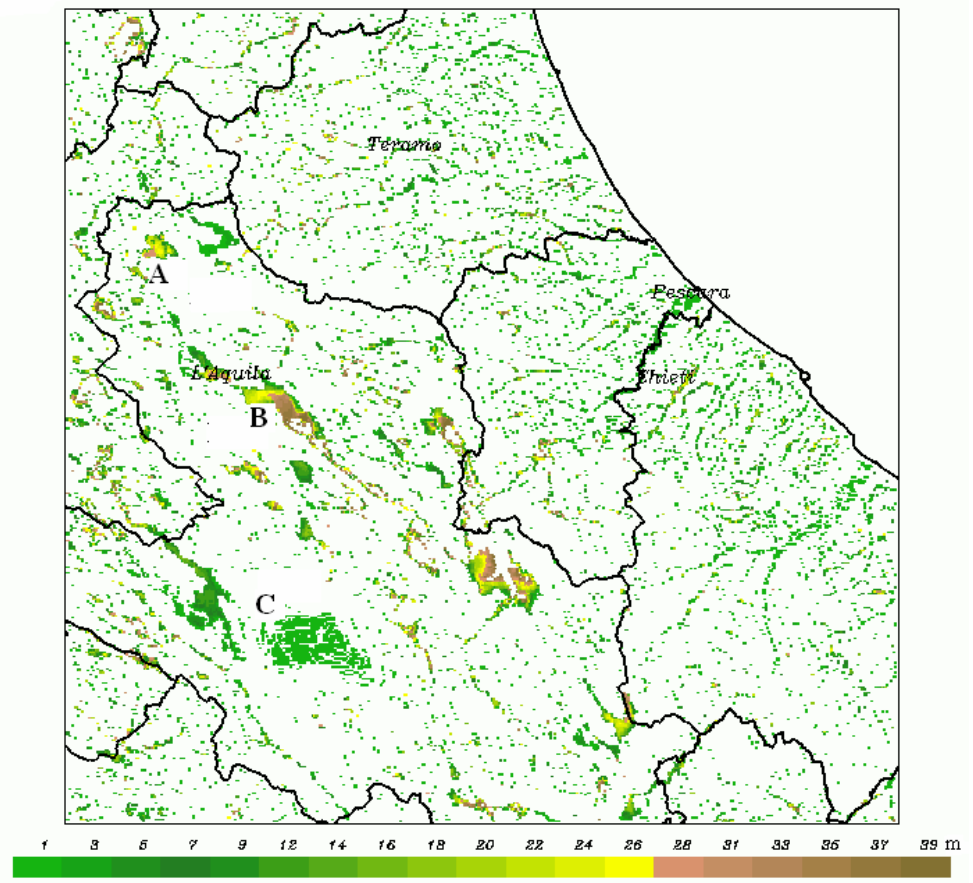

Fig. 3 The modification (in m) of the DEM is shown for every grid of the selected domain. A, B and C correspond respectively to an artificial lake, a river that flows inside the mountain and an area drained many years ago, and which is now drained by a network of artificial channels.

matrix for the selected domain is calculated using the RSA or tipping bucket algorithm. This can be summarised as follows: a "stone" rolls downhill from each cell toward the sea. Each time this stone goes through a given cell, a counter for this cell is incremented by 1. If the algorithm is iterated until all the stones get to the sea, an accumulation matrix that contains the total number of cells drained by each grid point can be calculated. If a quantity A is associated with each stone, where A is equivalent to the area of the cell where the stone was at the beginning of the algorithm, then for each cell the upstream drained area can be computed. A similar approach can be used to calculate the total drained precipitation for each cell. This matrix plays an important role in the definition of the flood alert index discussed in Tomassetti *et al*. (2005).

The algorithm to track the whole drainage network from the accumulation matrix is called the salmon algorithm because it acts like a salmon, moving upstream from the river mouth to the headwaters. The mouth of the river is the cell that drains toward a point in the sea and whose upstream drained area is greater than a fixed threshold (typically a few hundred km^2). From the mouth cell n_1 the algorithm goes upstream to the neighbouring cell n_2 whose drained area is the largest, but keeps in a temporary vector the location of the cells with a drained surface greater that the threshold (n_3'). Once the headwater of the main channel is reached (n_k), where the headwater is the

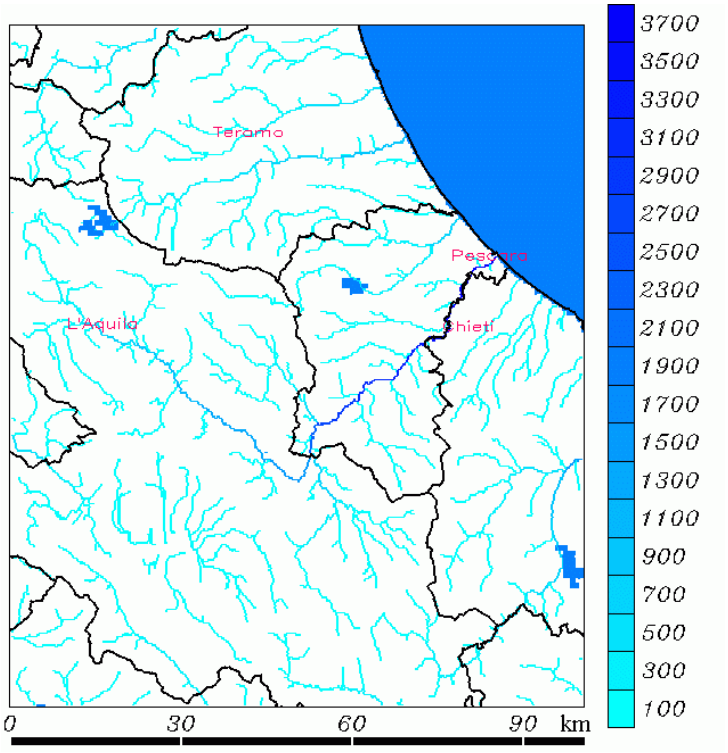

Fig. 4 The river network obtained from the corrected DEM of Fig. 3. The Aterno-Pescara River basin in the Abruzzo Region in the central part of Italy is reported. The DEM resolution in the current CHYM implementation is about 300 m. The vertical bar gives the drained area in km^2.

first cell having a drained area lower than the threshold, the algorithm starts again from the point n_3' and reconstructs the first secondary channel.

An example of a final watershed reconstruction is shown in Fig. 4. The coherent network shown in Fig. 4, which covers the same domain as in Fig. 3, clearly demonstrates how the CA approach rebuilds the flow direction in a very complex situation.

The finest DEM resolution in the current CHYM implementation is about 300 m, which is the resolution used in this example. The drainage network has no discontinuities, the flow direction is defined for every grid point, and comparing the image with a geographical map indicates that the shape of the network resembles the actual map very closely.

Further examples are shown in Fig. 5. Figure 5 (a) and (b), shows the rebuilding of the River Godaveri drainage network in India; Fig. 5(c) and (d) shows the Amazon Basin in South America and Fig. 5(e) and (f) the main basins in France. All the drainage networks reported in Fig. 5 are obtained using a horizontal resolution of about 1 km.

The proposed CA2CHYM approach has been tested on large number of different domains with different resolutions, some of which are presented here (see Figs 5 and 6 and the corresponding discussions), while many others can be viewed online at: http://cetemps.aquila.infn.it/chym/examples. Additional information on the algorithm can be found at: http://cetemps.aquila.infn.it/chym/users where the software described is also available for free download, for research purposes only.

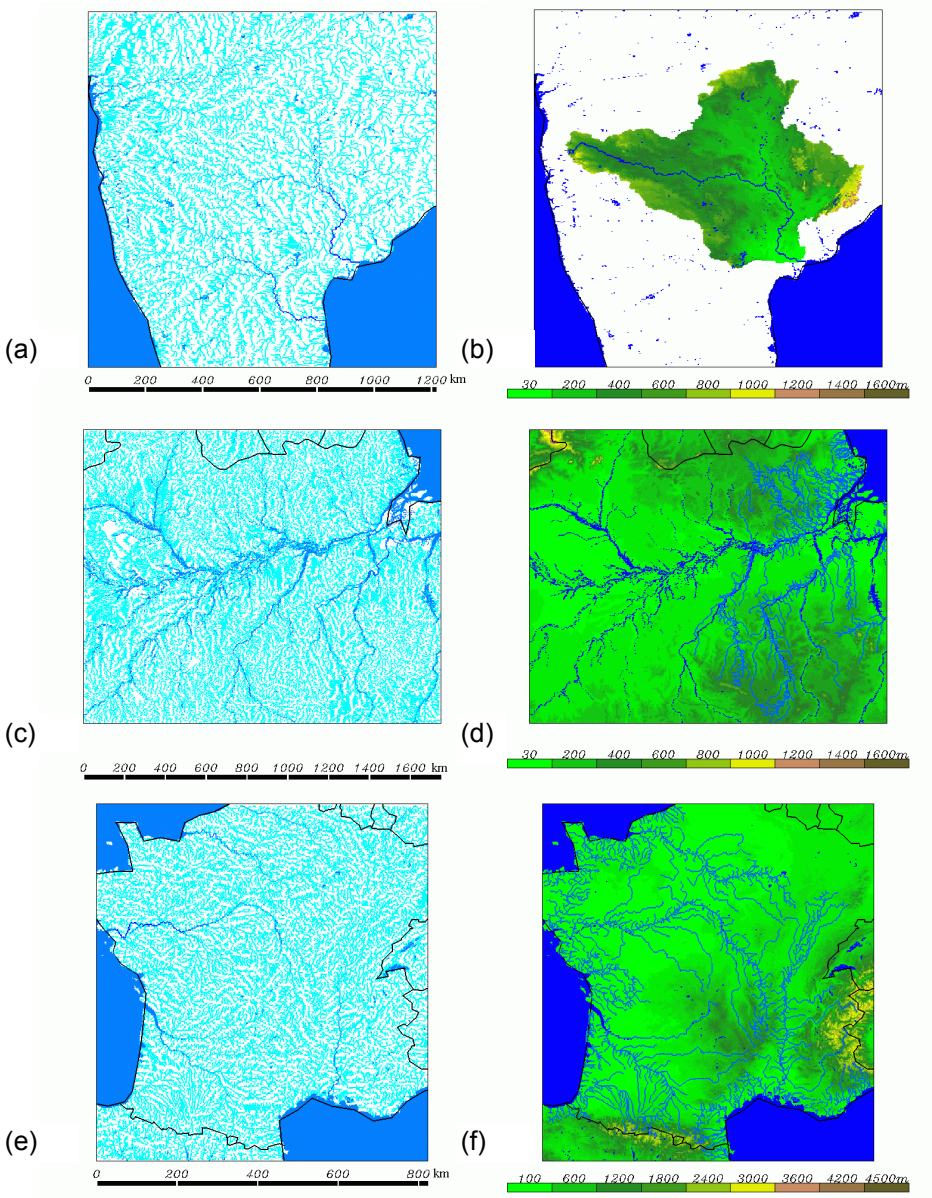

Fig. 5 (a) and (b): the Godaveri River (India) drainage network after being rebuilt with the CA2CHYM and salmon algorithms. The salmon algorithm "tracks" the drainage network both in a numerical (panel (a)) and graphic way (panel (b)). The same is reported for: (c) and (d) the Amazon Basin in South America; and (e) and (f) the main French basins: Senne, Loire, Gironde and Rodano. The legend of the right-hand panels is the height in metres of the DEM.

2.2 Assimilation and downscaling of rainfall data sets

The CHYM model rebuilds hourly precipitation fields for hydrological simulation. Precipitation data are characterized by large spatial and temporal variability. Nevertheless, it has long been recognized that rainfall patterns play an important role in runoff generation and many studies have described the strong nonlinear relationship between rainfall distribution and river discharge (e.g. Goodrich *et al.*, 1997), and examined the hydrological response to the different precipitation patterns (e.g. Singh, 1997).

For hydrological research, the main obstacle in studying rainfall spatial patterns is the limited rainfall data, usually from a sparse network of raingauges. Remote sensing techniques have been used to overcome these problems (Grimes *et al*., 2003; Coppola *et al.*, 2006), but the accuracy of rainfall estimates (Krajewski & Smith, 2002) and runoff sensitivity to sub-pixel rainfall variability (Michaud & Sorooshian, 1994) are very difficult to take into account. An alternative to the use of observations could be quantitative rainfall prediction from meteorological models (Van der Linden & Christensen, 2003), but high-resolution limited-area model predictions are affected by large uncertainties, and they are often not suitable for simulating realistic precipitation patterns at a hydrological scale.

Several techniques have been proposed to merge different sets of data in order to reduce uncertainties in rainfall estimation, the most recent being based on a physical approach (French & Krajewski, 1994), or statistical algorithms (Todini, 2001). All these methods are often difficult to apply within the context of operational hydrological activities, because, for example, not all data sets are usually available for the same period. To overcome this problem, we have developed a more empirical, but general, method that assimilates different available rainfall estimates by taking into account the differing nature of the data. The core idea behind the method is to assimilate the different data sets using a hierarchical sequence of modules, where each module examines a single data set. Each module will rebuild the precipitation field in a selected sub-domain, which depends on the spatial distribution of the available data. Such a sub-domain is defined as the ensemble of grid points having at least one measurement in a selected radius of influence, r_m, with r_m having a typical value of a few kilometres, depending on the density of available data.

To describe the algorithm in more detail, let us imagine a domain fixed on a latitude–longitude resolution grid. The goal is to assimilate a data set that consists of hourly rainfall observations onto a sparse grid. Each CHYM grid point is then defined by a different type as shown in Fig. 6. Cells of Type 0 (red square) are those for which there is little interest in estimating the precipitation, for example, because they are points corresponding to the sea or they are located outside the catchment of interest. The cells of Type 1 (white) are grid points for which we do not yet have a precipitation estimate. Cells of Type 2 (blue – upper-left corner of the figure) are those for which the precipitation has been estimated in a previous module with a different data set and this precipitation value will not change with the application of subsequent modules.

Within the cyan domain, are blue squares that correspond exactly to the location of the raingauges that are to be assimilated in the present module. At these points, the value of the precipitation is set to the value of the corresponding observation and they are from now on considered of Type 2 (the value of the precipitation will not be changed any more). The cells in the cyan domain in the figure are defined of Type 3 and correspond to all the cells whose precipitation value will be established within the current module and, as explained above, this sub-domain is represented by all the cells having at least a measurement in a selected radius of influence r_m. The next step is to use a CA-based algorithm similar to that described above for the DEM smoothing, for establishing a coherent value for the grid points of Type 3. The basic idea used here is that, when we deal with hydrological-scale resolutions (a few hundreds of metres), it becomes a good approximation to consider that the rain at a grid point is the weighted average of the rain in the surrounding cells, with weights that take into account the

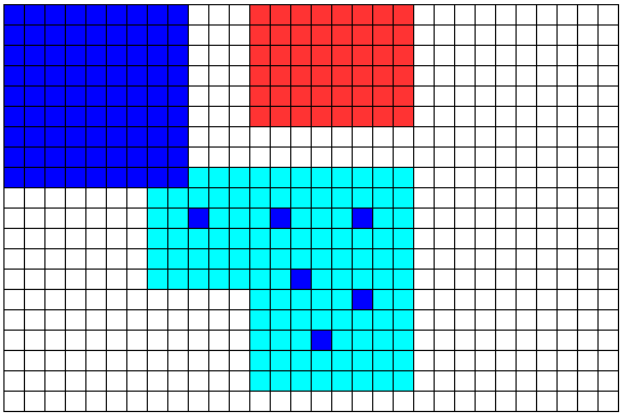

Fig. 6 A CHYM grid containing different kinds of cells involved in the precipitation estimation process. Cells of Type 0: red; Type 1: white; Type 2: blue (square in the upper-left corner); and Type 3: cyan. The blue cells in the cyan domain indicate the location of raingauges.

different distances between the cells. Considering the precipitation at each grid cell as the weighted average, the CA-based algorithm essentially acts to find a coherent solution to the precipitation in the neighbourhood. The algorithm can be described as follows:

- The precipitation matrix that coincides with the CHYM regular grid is considered an aggregate of cellular automata.
- The value of the rainfall in the cells of Type 2 (points of measurements or estimates of a previous data set) remains unchanged.
- The value of rainfall in the cells of Type 3 is iteratively modified using a rule similar to that of equation (3), but the sum is now carried out using only the neighbouring grid points of Type 2 or 3.
- The iteration ends when the changes in the rainfall values become negligible (less than a few percent) and all the cells of Type 3 are now classified as Type 2.

In order to speed-up the numerical processing, all cells of Type 3 are initialized with a "reasonable" value calculated with a geometrical algorithm using the following formula:

$$R_i = \sum_j \frac{1 - r_{ij}^2 / r_m{}^2}{1 + r_{ij}^2 / r_m{}^2} R_j \tag{4}$$

where R_i is the estimated rain value, R_j is the rainfall measurement available within a radius of influence r_m, and r_{ij} are the distances between raingauge locations and the cell. This interpolation method is usually referred to as the Cressman algorithm (Cressman, 1959).

Four examples of the application of a sequence of two and three modules are shown in Figs 7–10. Figures 7–9 show three examples of the assimilation of raingauge data and the MM5 rainfield. The rain sources used from each module are reported in Figs 7(a)–9(a). The bright green spots are the gauges, whose area of influence is determined taking into account the average distance between the raingauges, and the grey-green area is the MM5 model output. For each module, the Cressman and CA results are shown in panels (b) and (c), respectively, for the raingauge data, and panels

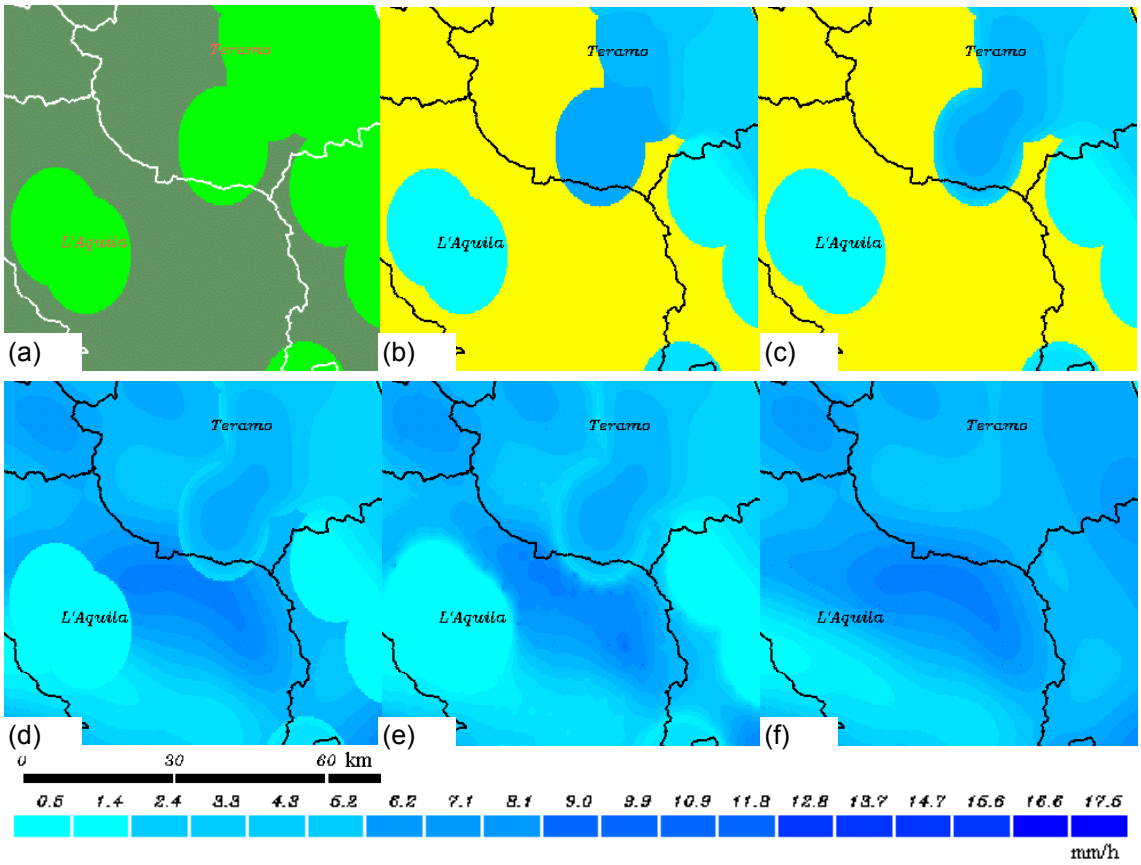

Fig. 7 Rainfall data assimilation for 4 August 2005 (11:00 UTC). Panel (a) shows the rain source available at each grid point (green: gauges; grey: model meteorological output used as a source when measurements are not available). In panel (b) only the gauge measurements are used and merged with the Cressman algorithm. The yellow area indicates that for this module there are no available data in these grid points. In panel (c) the same gauge measurements are used but merged using the CA algorithm. In panel (d) the rainfield is shown when the MM5 rainfield is added to the previous one of panel (c) using a Cressman algorithm; and in panel (e) the same field is reported when the CA algorithm is used. Panel (f) is for comparison, to show the rainfield as it is forecast by the MM5 model.

(d) and (e) for the MM5 model. In panel (f) the MM5 rainfield is reported for comparison.

In Fig. 10(a), the green spots are the gauges, the red area is the radar and the grey is the MM5 model output. Figures 10(d) and (e) show the radar rainfields obtained by the Cressman or CA algorithms, respectively, used as the second module and interpolated to the finer grid of the CHYM model. From a qualitative point of view, it is clear that the rainfield obtained by the CA algorithm better resembles a more realistic spatial distribution of the rainfield in all modules of Figs 7–10, and it has considerably fewer geometrical irregularities than those clearly evident in the case of the rainfield obtained by the Cressman algorithm. This is particularly noticeable in Figs 7(e)–9(e), where the result of the CA algorithm has a less geometrical shape compared to that observed in Figs 7(d)–9(d). The same is evident in Fig. 10, moving from Fig. 10(b) to (c), close to the label "l'Aquila", where the three rain circles are merged together, and in Fig. 10(d) and (e) in the lower left-hand corner. An important

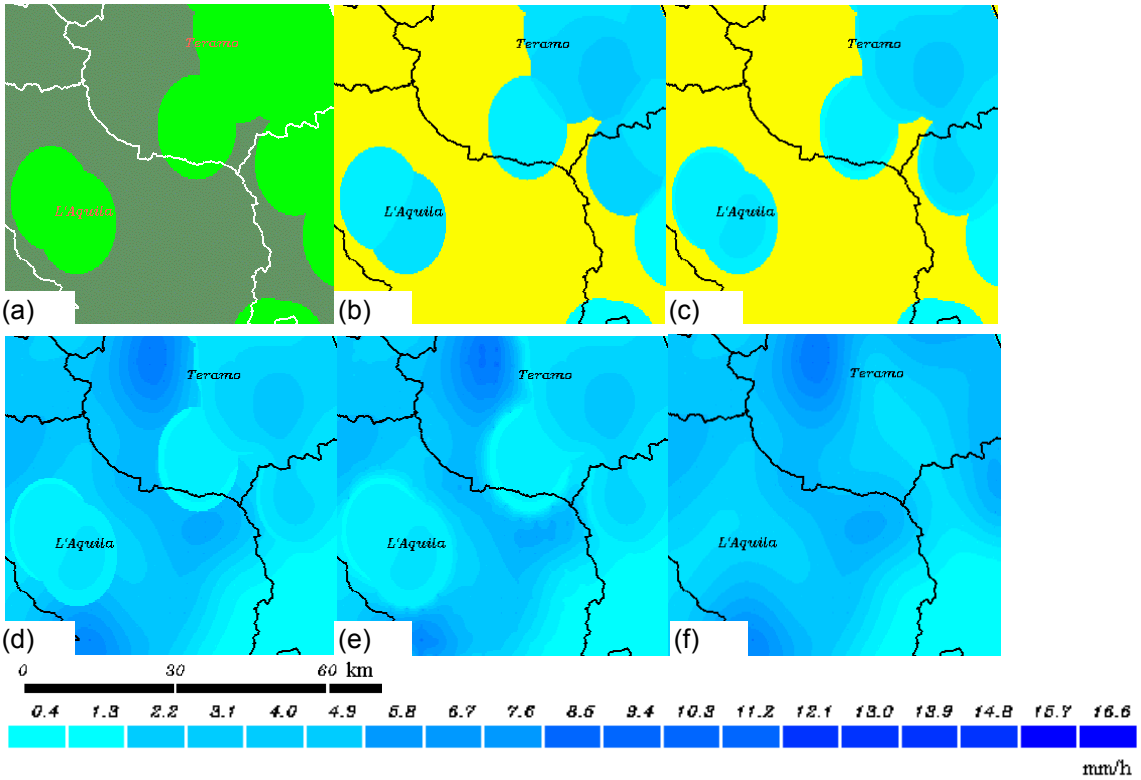

Fig. 8 Rainfall data assimilation for the precipitation event of 6 October 2005 (16:00 UTC). See Fig. 7 for explanation.

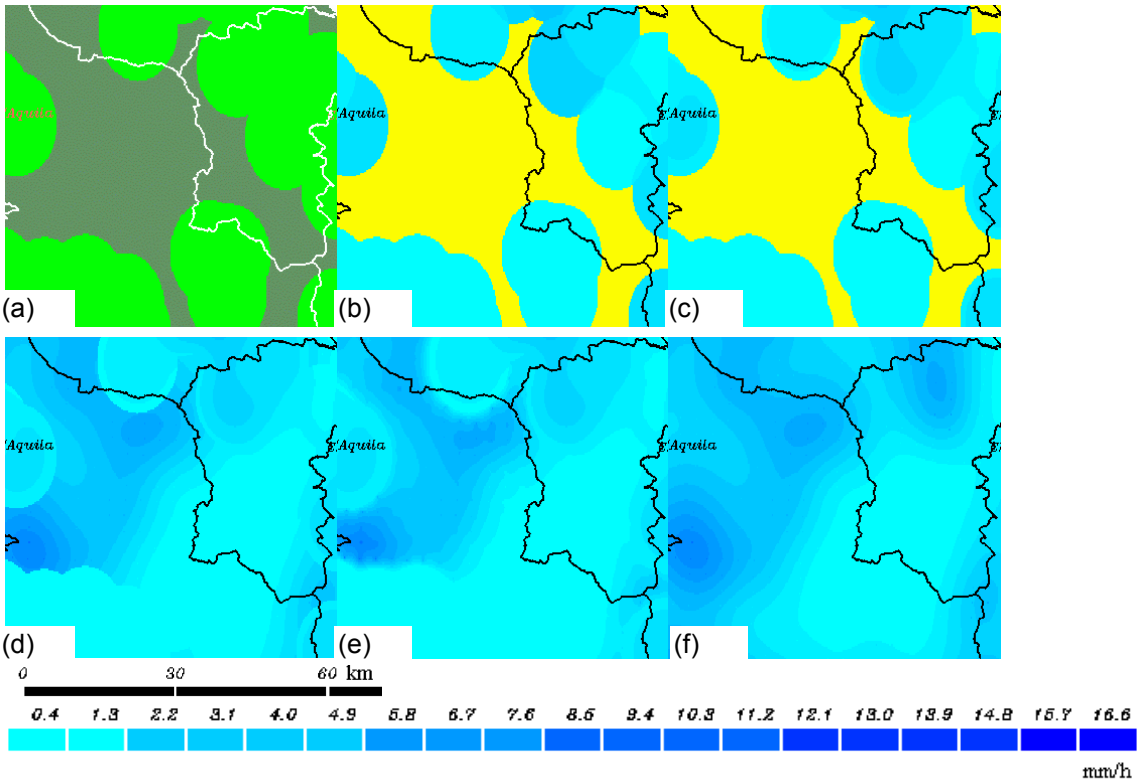

Fig. 9 Rainfall data assimilation for the precipitation event of 6 October 2005 (17:00 UTC). See Fig. 7 for explanation.

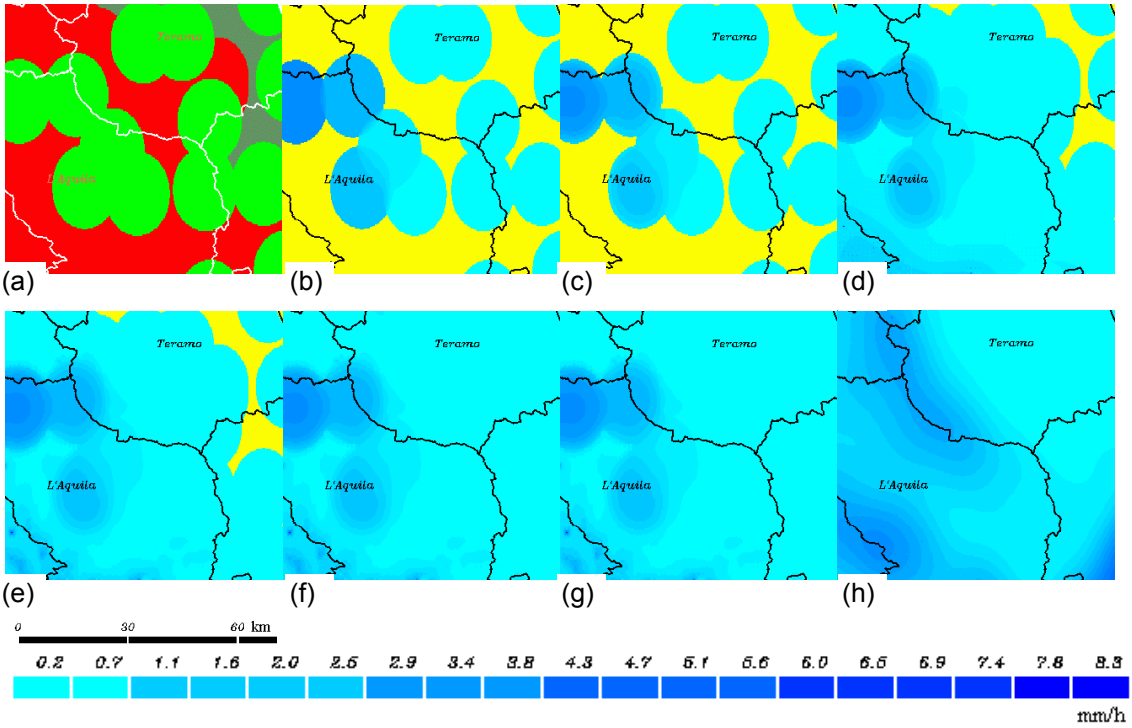

Fig. 10 Rainfall data assimilation for the event of 26 September 2006 (14:00 UTC). See Fig. 7 for explanation. The red area indicates the radar source.

property of this method is that, having established a hierarchy between the data, it allows for a constraint to be placed on the simulated values when they are close to the measurements. This is consistent with the idea that measurements are more accurate than simulated values.

3 CONCLUDING REMARKS

The sequence of algorithms implemented in the CHYM hydrological model has been described to extract the streamflow network and then rebuild the drainage network in an arbitrary domain and with an arbitrary resolution. The CA2CHYM algorithm has been shown to solve streamflow network singularities occurring in flat areas and pits and to have good performance in many tested domains. The proposed algorithm also minimizes the number of DEM points to be corrected, avoiding an unrealistic modification to this data set. The RSA and salmon algorithms create the possibility to rebuild the drainage network in a graphical and numerical way for the selected domain. The cellular automata-based algorithm using a hierarchical sequence of different rain data sources has also been used to rebuild the precipitation field to match the CHYM grid. The rainfield obtained by the CA algorithm better resembles a realistic spatial distribution of the rainfield and has considerably fewer geometrical irregularities than those clearly evident in the case of the rainfield obtained by the Cressman algorithm.

Acknowledgements Authors are grateful to CETEMPS Centre of Excellence for financial support to develop the CHYM model.

REFERENCES

Band, L. E. (1986) Analysis and representation of drainage basin structure with digital elevation data. In: *Proceedings of the Second International Conference on Spatial Data Handling*, 437–450. Int. Geogr. Union, Williamsville, New York, USA.

Coppola, E., Grimes, D. I. F., Verdecchia, M. & Visconti, G. (2006) Validation of improved TAMANN neural network for operational satellite-derived rainfall estimation in Africa. *J. Appl. Met*. **45**(11), 1557–1572.

Cressman, G. (1959) An operational objective analysis system. *Mon. Weather Rev*. **87**, 367–374.

Fairfield, J. & Leymarie, P. (1991) Drainage networks from digital elevation models. *Water Resour. Res*. **27**(5), 709–717.

French, M. N. & Krajewski, W. F. (1994) A model for real time rainfall forecasting using remote sensing, 1 Formulation. *Water Resour. Res.* **30**(4), 1085–1094.

Garbrecht, J. & Martz, L. W. (1997) The assignment of drainage direction over flat surfaces in raster digital elevation models. *J. Hydrol.* **193**, 204–213.

Goodrich, D. C., Lane, L. J., Shillito, R. M., Miller, S. N., Syed, K. H. & Woolhiser, D. A. (1997) Linearity of basin response as a function of scale in a semiarid watershed. *Water Resour. Res.* **33**, 2951–2965.

Grimes, D. I. F., Coppola, E., Verdecchia, M. & Visconti, G. (2003) A neural network approach to real-time rainfall estimation for Africa using Satellite data. *J. Hydromet.* **4**(6), 1119–1133.

Jenson, S. K. & Domingue, J. O. (1988) Extracting topographic structure from digital elevation data for geographic information system analysis. *Photogramm. Engng Remote Sens.* **54**(11), 1593–1600.

Krajewski, W. F. & Smith, J. A. (2002) Radar hydrology: rainfall estimation. *Adv. Water Resour*. **25**, 1387–1394.

Marks, D., Dozier, J. & Frew, J. (1984) Automated basin delineation from digital elevation data. *GeoProcessing* **2**, 229–311.

Martz, L. W. & De Jong, E. (1988) Catch: a Fortran program for measuring catchment area from digital elevation models. *Comput. Geosci*. **14**(5), 627–640.

Martz, L. W. & Garbrecht, J. (1992) Numerical definition of drainage network and subcatchment areas from digital elevation models. *Comput. Geosci*. **18**(6), 747–761.

Martz, L. W. & Garbrecht, J. (1993) Automated extraction of drainage network and watershed data from digital elevation models. *Water Resour. Bull*. **29**(6), 901–908.

Martz, L. W. & Garbrecht, J. (1995) Automated recognition of valley lines and drainage networks from grid digital elevation models: a review and a new method comment. *J. Hydrol*. **167**, 393–396.

Michaud, J. D. & Sorooshian, S. (1994) Effect of rainfall-sampling errors on simulations of desert flash floods. *Water Resour. Res.* **30**(10), 2765–2775.

O'Callaghan, J. F. & Mark, D. M. (1984) The extraction of drainage networks from digital elevation data. *Comput. Vis. Graph. Image Process*. **28**, 323–344.

Packard, N. H. & Wolfram, S. (1985) Two-dimensional cellular automata. *J. Statist. Phys*. **38**, 901–946.

Singh, V. P. (1997) Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrol. Processes* **11**(12), 1649–1669.

Todini, E. (2001) A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements. *Hydrol. Earth System Sci.* **5**(2), 187–199.

Tomassetti, B., Giorgi, F., Verdecchia, M. & Visconti, G. (2003) Regional model simulation of hydrometeorological effects of the Fucino Lake on the surrounding region. *Ann. Geophys*. **21**, 2219–2232.

Tomassetti, B., Coppola, E., Verdecchia., M. & Visconti, G. (2005) Coupling a distributed grid based hydrological model and MM5 meteorological model for flooding alert mapping. *EGU Adv. Geosci.* **2**, 59–63.

Tribe, A. (1992) Automated recognition of valley lines and drainage networks from grid digital elevation models: a review and a new method. *J. Hydrol.* **139**, 263–293.

Turcotte, R., Fortin, J. P., Rousseau, A. N., Massicotte, S. & Villeneuve, J. P. (2001) Determination of the drainage structure of a watershed using a digital elevation model and a digital river and lake network. *J. Hydrol*. **240**, 225–242.

Van der Linden, S. & Christensen, J. H. (2003) Improved hydrological modeling for remote regions using a combination of observed and simulated precipitation data. *J. Geophys. Res.* **108**(D2), 4072, doi:10.1029/2001JD001420.

Xie, P. P. & Arkin, P. A. (1997) Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Am. Met. Soc*. **78**, 2539–2558.

Received 3 October 2006; accepted 8 March 2007