Coupling Raster and Vector Based Models

Alexey A. Voinov,

H. Carl Fitz,

Robert Costanza

Institute for Ecological Economics (IEE)

mniversity of Maryland

P.O. Box 38, Solomons, MD 20688-0038

Tel: 410 326 7207

Fax: 410 326 7354

E-mail: voinov@cbl.umd.edu

Key words

spatial modeling; raster and ve9 or based GIS; hydrology; wetland; Everglades.

Acknowledgments

The research has been conducted under a contract with South Florida Water Management District. We are grateful to Thomas Maxwell fobavaluable advice and to Lisa Wainger fob editing and commenting on the text.

Abstract

An algorithm was developed to link raster and vector based components of simulation models. A vector object is associated with a set of cells with which it is interacting. This allows one to run simulations over both the raster cells and the vector object, coupling the model components according to assigned rules of material and information transport. The method is used and tested in a spatial model of the Everglades landscape. The method can be extended for general object modeling in a rasterized landscape with dynamically altered object configurations.

Coupling Raster and Vector Based Models

1. Introduction

A wide variety of spatial models has been used in landscape bology (Sklar and Costanza 1991) and may be classified into two major types. In one case the landscape is aggregated as a grid of quite small homogeneous cells an hprocess-based simulations are run for each cell with relatively simple rules for landscape flux among the nearest neighbors (Sklar et al. 1985, Costanza t al. 1990, Burke et al. 1990, Engel t al. 1993, Maxwell and Costanza 1994). This fairly straightforward approach requires xtensive spatial data sets and quite high computational capacities in terms of both storage and speed. However it provides for quasi-continuous modifications of the landscape, when borders between habitats may change in response to climatic change or management practices.

In the other case spatial units are aggregated to form larger polygon objects, called elementary landscapes, elementary watersheds, elementary areas of pollution (Krysanova t al., 1989) or hillslopes (Beven and Kirby 1979, kasowsky and Gardner 1991, Band et al. 1993), which form the basis for the hydrologic flow network an are also considered homogeneous, though possibly with respect to other criteria. Unit simulation models are run for those objects and their output is fed into another model or group of models that simulate the transport through linear components such as rivers, linking the objects together. Linked node hydrodynamic models such as the Storm Water Management Model, kWMM (EPA 1975a, 1975b) employ networks of hydraulic link vectors among many nodal points in a landscape. Like any aggregation, these techniques simplify the spatial representation an require less data, but the lower resolution leaves no room for modifications of borders between the elementary units.

These two approaches parallel the two main types of Geographic Information Systems (GIS), where information is usually maintained in either raster format, with data registered over

a grid space, or in vector form, with spatial objects presented by a series of vectors. As with GIS, where there are not many operations that are supported for a mix of both formats, it is more complicated to handle both the raster and vector based components simultaneously within spatial simulation models. However, in a rasterized landscape that accounts for spatial variation, there are classes of objects that are better represented in vector rather than raster format. This format is especially pertinent for the linear components of the landscape such as roads, communication lines, rivers and channels. Each such object is connected to a whole set of cells, and its interactions are spatially much more distributed than those of a raster component, that are usually limited to its nearest neighborhood.

In the class of models that closely follow the spatial resolution of raster GIS and treat each raster cell as a separate unit model, one of the possible methods for merging the raster and vector formats is to interpret the vector objects as raster units. Here the linear object (for example a canal shown in Fig. 1A) is extended to the size of a whole cell and the vector object representation is reduced to the raster one (Fig. 1B). This approximation works well enough if the width of the vector object is nearly equivalent to the size of the raster cell. Effectively this was the case in the CELSS model (Costanza et al. 1986, 1990), where the presence of a canal or a levee modified the overland flow between adjacent cells, actually putting a group of cells into a separate raster category.

In another approach (MacVicar et al. 1983) the linear object, which was a canal much smaller in width than the cells in this case, was either fit between the cells or drawn along the cell diagonal (Fig. 1C). In order to input the geometry of the vector, all of the coordinates of the canal node points had to be specified plus the indicator that gave the orientation of the canal through this node point and the direction of flow into the cells. Besides the cumbersome input, the drawback was that the vector length was extended beyond the actual length and there could be some ambiguity in choosing the appropriate interacting cells when the flow is dispersed over both directions.

2. General method

Generalizing these approaches, we may associate a series of cells with each of the vector fragments, that show when the grid crisscrosses the vector. The most trivial would be to identify only the cells that the vector crosses (these will be the same as in Fig. 1B). This may not be sufficient if the interaction between the cells and the vector is different on each side, as in the example of a channel with a levee on one side. In this case it is not clear how the vector-containing cell should be interacting with the vector; the interaction can be either overland flow, which we observe on the one side of the channel, or groundwater seepage, as on the other side. We may therefore associate two (Fig. 1D), or even more cells with each vector fragment (Fig. 1E). The cells that interact with the vector will then be cells on either side of the vector, but not necessarily cells directly overlain by the vector, identifying a whole zone of interaction for the vector. This might be useful if the vector represents, for example, a source of air-borne pollution (such as vehicle exhaust along a road), which is spread over more than just the nearest neighbor cells.

By linking the vector to a zone of cells with which it is interacting, material and information can be exchanged between the vector and the raster components. After separate simulations are run over the raster cells and the vector object, a transport procedure is started to create the interaction. The frequency of these interactions will depend on the amounts of material that are to be fluxed and on the specific rates of other processes in the system.

3. Algorithm Description

Let us consider in brief the algorithm that can be used to identify the cells in contact with a vector object. Suppose that the vector is specified by the coordinates of its beginning (x0,y0) and its end (x1,y1). If the configuration of the linear object is more sophisticated, it can be

represented as a series of such vectors and the algorithm can be run for each of them separately. Starting at the beginning point, we move along the vector marking the cells to the right and left, and for each cell transected the length of the vector contained in the cell is calculated (Fig. 2). This length will specify what portion of the vector/raster interaction will be assigned to the particular cell or group of cells.

In this way we find all the cells that the vector crosses. We may then compare where the vector segment lies relative to the center of the cell to decide which cell should be marked as being to the left or right of the vector (Fig. 3). If the vector crosses the cell in diagonal, the diagonal cells are marked, otherwise when the vector is horizontal or vertical, the appropriate adjacent cell is marked. If the zone of interaction is assumed larger than one cell, then the corresponding next cells to the right and to the left are marked and stored in the array. Note that it is not only the transected cells that can be marked as interacting with the object.

This paves a way to solve a more general problem of linking the vector and raster based objects. A major problem is that the two approaches assume different patterns of landscape boundary approximation. The raster cells are always considered to be homogeneous and therefore the transition from one landscape type to another can occur only on the border between cells. Vectors or vector defined objects (polygons) are also assumed to be homogeneous, but the vectors are not necessarily running in between the cells. Therefore when drawing the vector across a cell, there is always the dilemma about the type of interaction between the vector portion within the cell and the cell it transects. Allowing the vector to interact with several cells rather than only with the one transected resolves the problem. In this case the border between different landscape regions defined by the vector is shifted to fit between the raster cells, while the vector itself is not changed and may be still specified only by the coordinates of its beginning and end.

In the general case the object need not be linear and its zone of interaction need not be constant. In this case the set of interacting cell coordinates should be dynamically updated as

the model runs. We have implemented this method of dynamic allocation of the linkage arrays to model fire propagation over the landscape.

4. Case-study: The Everglades Landscape Model

The method discussed above was applied as part of the hydrologic module in the Everglades Landscape model. Hydrologic processes play a critical role in determining the ecosystem functioning within the Everglades wetlands. Overland sheet flow and groundwater movements are some of the principal fluxes of water across the landscape and they regulate the plant and animal community structure. A large network of canals, levees, and associated flow control structures are used for water management in the region. The canals are a significant water transport mechanism in the Everglades, moving large quantities of water further and more rapidly than the comparatively slow overland flow.

These engineered management structures have altered water delivery in the region, with subsequent changes to the Everglades ecosystem characteristics (synthesis in Davis and Ogden 1994). With varying objectives, computer models have been developed that include simulation of water flow in this region (MacVicar et al. 1983; Walters et al. 1992). The South Florida Water Management Model (MacVicar et al. 1983) simulates the hydrology in a very detailed way on a 2x2 mile grid, and has been successfully calibrated and used to influence water management policies in the region. This model incorporates the full network of canals and associated structure in routing water in the landscape. However, the canals have been displaced in order to place them along cell borders or along cell diagonals. This displacement altered the length of the canals and therefore modified the range of interaction.

The Adaptive Environmental Assessment Model (Walters et al. 1992) is a simpler gridbased (16 km² cells) model with a focus on the landscape vegetation and its response to general changes in water supply. The model incorporates hydrologic algorithms that assumed very general rules for impounding/ releasing water within large regions (Water Conservation Areas - WCAs), without explicitly modeling flows in the canal network. This model is good for general landscape analysis, but was not intended (nor suited) to examine detailed changes in hydrologic processes caused by changes to the operation routines for structures and channels.

The Everglades Landscape Model (ELM) (Fitz et al. 1993) is designed to simulate the landscape scale vegetation response to simulated hydrology, water quality and fire. The ELM has a fine scale grid that divides the landscape into ~10,000 1 km² cells, and the canal/levee network is overlaid according to the algorithm presented in this paper, but otherwise the model canals are treated in a fashion similar in many respects to the SFWMM. The canals are presented as a set of straight canal "reaches", vectors, characterized by the coordinates of their ends (Fig. 4). The canals are linked by control structures that are used to define the water fluxes from one canal into another using either historical data or management rules.

4.1. Configuration stage

The ELM hydrology is modeled in two stages. At the beginning of a model run the geometry of the region is defined. There are several combinations of canals with levees within the ELM boundaries. There are canals that have a levee on both sides of the canal, others have a levee on one side only, and there are canals with no levees. The modeled interactions of cells with canal reaches vary depending on the presence/absence of levees in the zone of interaction.

Each canal is therefore specified by the following information:

- canal ID;
- set of coordinates for the endpoints of the line segments which represent the canal;
- levee location relative to canal (0 no levee, +1 levee on the left, -1 levee on the right, assuming the forward direction is from canal start to end, 2 - levees on both sides);
- canal attributes including width and depth.

Each canal is assumed to have constant depth (varying from 1.74 m to 4.42 m for different canals) and width (25 m). The levee width is assumed to be negligible. Based on this information, the length of the canal within each of the canal-containing cells and the arrays of cell coordinates that interact with the canal on the left and right are identified.

Depending on the range of interaction a certain number of cells is marked on each side of the canal (Fig. 5). The range of interaction is specified depending on how wide we want to spread the canal/cell interaction. At the zero level of interaction only the canal owning cells are involved and they are marked as shown in Figure 1C. The type of interaction can be different depending upon the position of the levee. Therefore the canal will be interacting in turn either with the cell to the left through levee seepage (since there is a levee on the left hand side), or with the cell to the right through overland flow (no levee on the right hand side). The direction of interaction is determined according to the position of the canal relative to the center of the cell (Fig. 3). As a result, in this case the array for levee seepage type of interaction will contain the pairs (i-1,j), (i,j-1), (i+1,j-2)... On the other side, the array for overland flow will have (i-1,j+1), (i,j), (i+1,j-1)... .

At the first level of interaction we assume that the canal can communicate not only with the cell directly underneath, but also with the next adjacent cell, even though it does not contain a canal. These are the darker cells in Fig. 1D. In this case each canal interval can interact both with a cell to the left and a cell to the right. The corresponding arrays from Fig. 5 will be (i-1,j), (i-1,j-1), (i,j-1), (i,j-2), (i+1,j-2)... for seepage type of interaction and (i-1,j+1), (i,j+1), (i,j), (i+1,j), (i+1,j-1)... for the overland flow.

We may further extend the range of interaction adding more cells to the right and to the left (all filled cells in Fig. 1E), which might be necessary if there is a significant region of interaction between the vector element and the raster base. In ELM the level 1 range of interaction was chosen. The zero level did not provide sufficient cells to receive the significant flows that occur in the Everglades canals. This level also does not provide adequate transport when the canal is directed strictly along one of the grid axes (Fig. 3B,C), because then all the

linked cells are on one side of the canal and therefore there is only one type of interaction for the whole canal (either seepage or overland flow).

We found no need to go to higher levels of interaction since the cells could accommodate the flows that were generated. However when running the model with smaller (500m x 500m) cells to get a better spatial resolution in one of the ELM subregions - the Water Conservation Area 2A, the first level of interaction provided too few cells to receive the large volumes of water discharged from the canals, and higher levels were adopted (2 on one side and 1 on the other, for most of canals). In general, if there are no definite rules to determine the zone of interaction, we should make it an empirical parameter and modify it to enhance model performance. In our case the level of interaction was chosen such that the water head differences induced by the canal-cell flows did not become unreasonably high (> 1 m).

This marking scheme is extended to the full model area (Fig. 6). Cells are marked by the number of the canal they interact with, which shows in shades of gray in the figure. Intersections of several reaches create cells that are to interact with several canals. This becomes confusing for the hydrologic modules that are run consecutively for the canal reaches. Therefore these cells are marked as non-interacting (black in Figure).

4.2. Simulation stage

After the configuration step, which is performed once at the beginning, the next stage consists of running the three linked hydrologic submodels: 1) a general ecosystem model (Fitz et al. 1995) that is replicated in each cell, incorporating vertical fluxes of water associated with precipitation, transpiration, percolation, etc.; 2) cell-cell horizontal fluxes based on modified Manning's equation (Chow et al. 1988); and 3) horizontal fluxes among canal reaches and cells described here.

Within each raster cell the unit model simulates a variety of hydrologic processes and parameters, including the following:

8

- 1. Transpiration associated with plant growth, physiology and relative humidity.
- 2. Evaporation using pan evaporation estimates and pan coefficients.
- 3. Rainfall based on precipitation data interpolated over 9 stations.
- Seepage of water from that stored above the sediment/soil surface into that stored in sediment pore space (either in unsaturated or saturated storage).
- 5. Roughness coefficient that depends on dynamic simulation of plant biomass, numeric density, and plant morphology.

After the water head in each raster cell is modified due to vertical fluxes, the water movement between the raster cells (and associated transport of nutrients and sediments) is described according to a modified Manning's equation for overland flow and a modified Darcy's equation for flow of water in saturated storage.

The same flow equations are used for water transport between cells and canals. Canal-cell interactions are thus driven by the same mechanisms as the rest of the cell-cell interactions. The length of the canal within each of the canal-containing cells defines the area of interaction with the canal, across which water and material may flow.

The canals are linked at upstream and downstream structures (Fig. 7A). These structures (weirs, culverts and pumping stations) control the water flow between the canals and control the inflow and outflow to and from the network. Whereas the SFWMM relies to a large extent on historical stage discharge relationships to determine flow, we are making an effort to incorporate theoretical weir/culvert flow equations so that realistic scenarios, based on managed response to stage height, can be simulated. For modeling purposes instead of the usual representation of the canal network as structures linked to canals as in Fig. 7A, it might be useful to visualize the system as a combination of nodes, that represent the canals, connected by links, that stand for the fluxes controlled by structures (Fig. 7B). Thus, we can emphasize that the canals are actually the objects that are modeled and linked to the raster cells, while the structures only regulate the in- and outflows to and from the canals.

Normal spatial fluxes of overland and groundwater flow among cells occur from one cell into at most four neighbor cells. However, a canal reach has a larger number of interactions via overland flow and groundwater seepage from all of the cells along the reach. A quantity of water (determined from management rules) is introduced into a reach via a weir/culvert flow equation from an upstream canal reach. Gains and losses due to canal-cell interaction (based on hydraulic head differences) are calculated along the reach, and an iterative equilibrium solution is sought among the various inputs and outputs of the canal for water stage height on a daily basis (as in the SFWMM). Within a time step, the model undergoes a series of iterations adjusting the stage to balance the water levels in the canal and in the adjacent cells until the difference in stage height (error) estimates converges to a sufficiently small value.

ELM is generating maps of surface water heads on a daily basis, that can be used for animation of the landscape hydrology. As an example of the model output with the canal/levee network affecting the surface runoff, in Fig. 8 we present a snapshot of the surface water heads as simulated for mid-summer conditions and climatic scenarios of 1983. It is hard to visualize the 2-D output in gray-scale figures, the results look much better in color. Some colorful animation results from the Everglades model can be viewed at our WWW site at http://kabir.cbl.cees.edu/Glades/ELM.html, giving a much better idea about the model performance and the effect the canal-levee network has on water fluxing across the landscape. Details on hydrologic modeling involved can be also found there.

5. Conclusion

By mapping all the 2-dimensional cells that are interacting with vector objects, we link the objects to the raster grid in a simulation environment. This may be easier to visualize in the 3-dimensional space if we consider the object defined by several vectors or points as lying outside of the rasterized landscape and interacting simultaneously with a whole series of raster cells. The proposed algorithm identifies the cells interacting with linear objects and simplifies

10

further linkage of models created independently for the vector and raster components. These linkage sets of cells may be made either of the vector-owning, transected cells, or they may contain additional adjacent cells that also "talk" to the models built for the linear components. In any case, as soon as the zone of interaction is defined, the models can be run independently and linked together at certain time steps, as necessary.

In our approach we maintain the simplicity of vector input, specifying the whole object by two pairs of coordinates, instead of specifying all the raster cells that the vector crosses. At the same time we do not modify the interface between the vector and the cells, keeping the length of the vector the same and maintaining a precisely-defined length of cell-vector interaction. We also gain more flexibility in being able to specify the width of the zone of interaction. If necessary this width can be made variable and can be computed depending upon some intrinsic models of dispersion (as for air-borne pollution generated in the vicinity of roads as a function of wind velocity).

Looking at canals as nodes in the other dimension, we can build the whole canal network in the ELM case study as a node-and-link configuration, with each of the nodes connected to arrays of cells in the underlying landscape. One model operates within each of the cells and defines the vertical fluxes based on the information contained in the various GIS layers. Another model operates between the canal nodes, fluxing water and constituents through structures and balancing the input and output in the canal network. And finally the third model links the canal nodes with the raster cells, equilibrating the water head in them.

This approach may have further applications when models have to connect spatially distant cells, as in case of socioeconomic analysis. Many economic models contain interactions, which would require linking cells that are not adjacent to each other, and the nearest neighbor approach that is used in raster models no longer works. Building the linkage array, containing coordinates of all the cells associated with the modeled process or function, may help run the model over these spatially diverse units. If spatial arrangements change in

time, the arrays may be also generated within the model temporal cycles, so that the possible links are modified with time.

References

- Band, L.E., Peterson, D.L., Running, S.W., Coughlan, J., Lammers, R., Dungan, J. and Nemani,
 R. 1991. Forest ecosystem processes at the watershed scale: basis for distributed simulation. *Ecological Modelling*, 56:171-196.
- Beven, K.J. and Kirby, M.J. 1979. A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.*, 24: 43-69.
- Burke, I.C., Schimel, D.S., Yonker, C.M., Parton, W.J., Joyce, L.A. and Lauenroth, W.K. 1990.
 Regional modeling of grassland biogeochemistry using GIS. *Landscape Ecology*, 4(1):45-54.
- Chow, V. T., Maidment, D. R. and Mays, L. W. 1988. Applied hydrology. New York, McGraw-Hill.
- Costanza, R., Sklar F., and Day, J. W. J. 1986. Modeling Spatial and Temporal Succession in the Atchafalaya/Terrebonne Marsh/Estuarine Complex in South Louisiana. In *Estuarine Variability*. Edited by D. A. Wolfe (New York: Academic Press), pp. 87-404.
- Costanza, R., F. H. Sklar, and M. L. White, 1990. "Modeling Coastal Landscape Dynamics." *BioScience* 40(2): 91–107.
- Davis, S.M. and Ogden, J.C. (Editors) 1994. Everglades: the Ecosystem and Its Restoration.St. Lucie Press, Delray Beach, FL, 826 pp.
- Engel, B.A., R. Srinivasan and C. Rewerts 1993. A Spatial Decision Support Systemor Modeling and Managing Agricultural Non-Point-Source Pollution. In: *Environmental Modeling with GIS*. Edited by M.F. Goodchild, B.O. Parks and L.T. Steyaert (New York, Oxford: Oxford Univ. Press), pp. 230-237.

- EPA, 1975a. Environmental Protection Agency. Storm water management user's manual, version II. Environmental Technology Series, EPA 600/3-85/040, Environmental Research Laboratory, Athens, GA.
- EPA, 1975b. Environmental Protection Agency. Rates, constants, and kinetics formulations in surface water quality modeling, 2nd ed., EPA 67012-75-017, U.S.EPA 4 Cincati, Ohio.
- Fitz, H. C., R. Costanza, and E. Reyes 1993. The Everglades Landscape Model (ELM): Summary Report of Task 2, Model Development. Report to the South Florida Water Management District, Everglades Research Division.
- Fitz, H. C., E. B. DeBellevue, R. Costanza, R. Boumans, T. Maxwell, and L. Wainger. 1996. Development of a General Ecosystem Model (GEM) for a Range of Scales and Ecosystems. *Ecological Modelling*, 88:263-295.
- Kessell, S. R. 1977. Gradient Modeling: A New Approach to Fire Modeling and Resource Management. In *Ecosystem Modeling in Theory and Practice*. Edited by C. A. S. Hall and J. W. J. Day (New York, NY: Wiley-Interscience), pp. 576-605
- Krysanova V., Meiner A., Roosaare J. and Vasilyev A. 1989. Simulation modelling of the coastal waters pollution from agricultural watershed. *Ecological Modelling* 49: 7-29
- MacVicar, T.K., T. VanLent and A. Castro 1983. South Florida Water Management Model Documentation Report. Technical Publication 84-3. South Florida Water Management District. 123 p.
- Maxwell, T. and R. Costanza 1994. "Spatial Ecosystem Modeling in a Distributed Computational Environment." In: J. van den Bergh and J. van der Straaten (eds). Concepts, Methods, and Policy for Sustainable Development. Island Press.
- Sasowsky, C. K. and Gardner, T.W. 1991. Watershed configuration and geographic information system parameterization for SPUR model hydrologic simulations. Water Resources Bulletin, 27(1):7-18.
- Sklar, F. H., Costanza, R., and Day, J. W. J. 1985. Dynamic spatial simulation modeling of coastal wetland habitat succession. *Ecological Modelling*, 29: 261-281.

- Sklar, F. H., and Costanza, R. 1991. The Development of Dynamic Spatial Models for Landscape Ecology. In *Quantitative Methods in Landscape Ecology*, Edited by M. G. Turner and R. Gardner (New York, NY: Springer-Verlag),.pp. 239-288.
- Walters, C., L. Gunderson, and C.S. Holling. 1992. Experimental policies for water management in the Everglades. *Ecological Applications* 2:189-202.

Figure captions

- Fig. 1. Possible ways of linking a vector object to the raster background. A a vector as it runs across the rasterized background; B the cells directly underlying the vector are substituted for the vector; C the vector is modified to cross the cells in a special way (diagonally or along the borders in this case); D cells on both sides of the vector are included to the linkage array if they are adjacent to the vector; E more remote cells are included into the linkage array.
- Fig. 2. Flow chart for the algorithm to identify the mapping of a vector defined by its starting (x0, y0) and ending (X1, y1) coordinates to the array of cells (i,j).
- Fig. 3. Possible patterns to mark the cells to the right and to the left of the vector. The vector segment is related to the cell center. If the cell center is to the left of the vector, then the cell is marked as left and the diagonal (or adjacent) cell is marked as right. Vice versa, if otherwise.
- Fig. 4. Major canal reaches and structures in the ELM.
- Fig. 5. Two types of canal/cell transport in the ELM zone of interaction overland flow on the side with no levee and seepage on the side with a levee.
- Fig. 6. The ELM area with the zones of the canal/cell interaction as identified by this algorithm
- Fig. 7. Types of structures in the canal network: A structures as they are sitting on canals presented as vectors; B - canals, presented as objects (linear vector objects in this case), linked by structures to one another and to the landscape units.
- Fig. 8. Example of ELM output for surface water heads across the area. Comparing this Figure to Fig. 4 one can clearly see how the surface flow is modified by the canal/levee network. See our web site at http://kabir.umd.edu/Glades/ELM.html to view this figure in color.



А.



Β.



















Α.

Β.



