A new raster-based spatial modelling system: 5D environment

S. Mazzoleni

F. Giannino

M. Mulligan

D. Heathfield

M. Colandrea

See next page for additional authors

Presenter/Author Information
S. Mazzoleni, F. Giannino, M. Mulligan, D. Heathfield, M. Colandrea, M. Nicolazzo, and M. d'Aquino

A new raster-based spatial modelling system: 5D environment

S. Mazzoleni[a], F. Giannino[b], M. Mulligan[c], D. Heathfield[d], M. Colandrea[e], M. Nicolazzo[e], M. d'Aquino[e]

[a] *Dept. Arboricoltura Botanica e Patologia Vegetale, University of Naples Federico II - Italy*

[b] *Dept. Ingegneria Agraria e Agronomia del Territorio, University of Naples Federico II - Italy*

[c] *Dept. Geography King's College London, University of London - UK*

[d] *World in a Box Finland OY - Finland*

[e] *Dept. Electrical Engineering, University of Naples Federico II - Italy*

Abstract:. The functional description of natural systems requires the representation of different classes of processes including chemical, biochemical, and physical functions, with or without their spatial representation. On the other hand, there are several cases where a realistic representation of the process dynamics needs a spatially explicit implementation of the model (eg. hydrological flows, fire propagation, pollution diffusion, etc.). In more general terms, and for different scientific fields, scaling problems and the issue of linking dynamic and spatial processes require an integrated logic and appropriate software solutions. Different tools aimed at implementing spatial modelling capability (PCRaster, SME, SimARC) have not yet achieved a full integration of spatial and dynamic processes, especially for users with limited programming ability. In pursuit of this capability, a new raster-based spatial modelling system named "5D environment" has been implemented and it is presented here.

Keywords: Modelling framework; Dynamic spatial modelling; integrated models; ODE; PDE

1. INTRODUCTION

Spatial variability has been shown to be a major control on the behaviour of natural systems with *patterns* often as important as *processes* in system behaviour. At the same time, the availability of high spatial resolution gridded remote sensing data, coupled with the capability to handle them with GIS tools, has enabled environmental scientists to parameterise some of the important spatial variability in landscapes, thus providing the data needed to run spatial models.

This is the background for the move away from the spatially 'lumped' approaches that have characterised environmental modelling over the last 50 years towards new spatially distributed models. It has to be said that the capabilities of computer hardware (storage, processing and output) for spatial modelling have developed much faster than the associated software tools.

In this paper we review the state of the art in spatial modelling systems, indicating the main systems which exist, their characteristics and limitations. We go on to suggest a blueprint for the ideal spatial modelling logic and system from a modeller's perspective. Finally we describe the process towards such a system by outlining the main characteristics of a new tool, the 5D modelling environment, currently under development as a combined effort by an interdisciplinary group of ecologists, hydrologists, mathematicians and programmers.

2. SPATIAL MODELLING TOOLS

Of course, any low- or high-level computing language can be used to build spatial models but we will not review these here since our focus is on high-level modelling tools that allow a computer-aware domain specialist (ecologist, geographer, social scientist) to build spatial models without investing a lot of time and effort in computer programming or deep mathematics. We should also distinguish those systems that simply allow modelling over a number of units (replication systems) from those that have true (location, neighbourhood and distance based) spatial functions and spatial relationships such as diffusion. Furthermore, there are spatial modelling software packages that are very advanced but do not allow for the building of models by end users, only their parameterisation or switching on and off, because the models themselves are hard-coded in the software. This kind of systems, for example the Geonamica Decision Support System (DSS) generator[1], are not considered here.

2.1 State of the art and limits of different approaches

The existing spatial modelling tools can be separated into those that have evolved as modelling extensions of GIS software (e.g. ESRI MODELBUILDER and AML, ERDAS SPATIAL MODELER), those that have evolved as spatial extensions of modelling software (Stella[2], Modelmaker[3],Vensim[4], Extend[5] and SIMILE[6]), connections between GIS and modelling software (SimARC[7], Mazzoleni et al. 2003) and custom-built spatial modelling tools (e.g. PCRASTER, Van Deursen, 1995; SME, Maxwell and Costanza, 1995).

GIS systems have been rather slow to take on spatial simulation modelling. Whilst these systems can have a very considerable arsenal of spatial functions which can be used to represent processes on raster or vector datasets, very few provide an easy to use modelling environment which facilitates spatio-temporal modelling. In the early ESRI products the most effective means of developing spatial models was through programming in arc macro language (AML) for ARCGIS workstation, though this is a rather low level scripted language and requires a significant investment of time to build even simple spatial models. On the other hand, AML does provide access to the full array of ESRIs spatial modelling functionality. More recently in ARCGIS Desktop 9 the so-called model builder provides a drag and drop functionality for building spatial models, though spatio-temporal models are constructed with great effort this way. Scripting in ARCGIS Desktop 9 is provided by the open source Python[8] programming language, but this is beyond the un-trained capabilities of most domain specialists.

Modelling systems have also been slow to take on board spatial modelling. Many now incorporate routines that in principle allow for replication of calculations through a cellular grid and thus the development of spatial or individual based models. However very few offer significant spatial functionality in addition to their built-in mathematical, trigonometric, logical and other functions and, in any case, are rather limited for most spatial modelling efforts. In fact, they mostly focus only on replication of calculations for different individuals or objects that have no spatial relationship.

A recent tool named SimARC interfaced ArcView with SIMILE, enabling users to link models to the GIS database, i.e. running any SIMILE model in each polygon element of ArcView shape file. The technical solution was a very computationally efficient way to replicate the model application in space, but did not allow any lateral process to be simulated.

The fourth type of system – purpose built spatial modelling tools - overcomes many of the above mentioned deficiencies by tackling some of the major developmental needs in spatial modelling. There are very few of these in comparison with the number of GISs and modelling software packages. Though a number of blueprints for such a system exist, the main operational examples are PCRASTER[9] and SME[10] (for a range of overviews and examples see Environmental Modelling & Software, 19 (3), 2004).

PCRASTER is a script based spatio-temporal modelling language with over 115 functions over half of which are truly spatial functions. It supports only raster grids, not vectors or objects. The software is developed by the University of Utretcht and the PCRASTER Environmental Software company and is currently freely available as windows or linux executable (though not open-source) and has been used widely for dynamic and geostatistical modelling [Pebesma and Wesselling, 1998], error propagation and uncertainty analysis [Heuelink and Burrough, 1993], and modelling of land degradation [de Roo et al., 1996], ecological processes [van Deursen and Heil, 1993], catchment hydrology [Mulligan, 2003] and geomorphological and soil processes, amongst others. The shell based script language is relatively simple and the system has useful visualisation tools but a domain

[1] http://www.riks.nl/products/GEONAMICA
[2] http://www.iseesystems.com/
[3] http://www.modelkinetix.com/
[4] http://www.ventanasystems.co.uk/vensim.html
[5] http://www.imaginethatinc.com/
[6] http://www.simulistics.com
[7] http://143.225.165.9/simarc/
[8] http://www.python.org
[9] http://www.pcraster.nl
[10] http://giee.uvm.edu/SME3

specialist must have some experience of script based languages to get by and the various utilities are rather loosely knit compared with standard GUI based modelling packages such as STELLA. Nevertheless PCRASTER has an extraordinarily powerful array of spatial and non-spatial functions and its development is ongoing.

SME has been developed by a team at the International Institute for Ecological Economics of the University of Vermont. It takes simulations models developed in STELLA-like systems and allows the user to integrate them within the context of a gridded space. SME also supports vector and object data models. STELLA, Vensim, Extend or Simulab models can be easily built using drag and drop but integrating them spatially in SME is an altogether more technically involved process involving the definition of the project, the creation and linkage of Simulation Module Markup Language (SMML) modules from the STELLA models, the assembly of data in the appropriate format, the generation of (C++) code from the SMML objects, and the configuration of the code prior to model run and visualisation Though 'generators' are available to do much of this, it appears rather technically challenging for the average domain specialist. SME is implemented as Unix software supporting parallel processing, which is important for much large scale, long term or fine resolution simulation, but also somewhat limited in its application to most non-specialists users.

3. 5D ENVIRONMENT

3.1 Software development and architecture

5D is a modelling system specifically designed for the integration of different models and the solution of spatial problems. The software is based on the research experience of the ModMED-EU project [Modelling Mediterranean Ecosystem Dynamics, Mazzoleni and Legg, 2001]. This project supported the ongoing development of SIMILE, a new system dynamics modelling tool. SIMILE represented a significant step forward a larger flexibility of model construction, with specific enhanced capabilities for ecological modelling work [Muetzelfeldt and Massheder, 2003]. However, it still showed logic and computational limitations in the representation of spatial processes. The need to overcome these problems provided the motivation for the development of the new 5D system.

Figure 1 shows a conceptual representation of the logical and spatial relations between models in the 5D environment. Some models can operate within single cells (pixel models) whereas others apply to differently shaped areas (object models). Areas of models' application may overlap. Models can be either spatially isolated (local processes) or related to neighbours (lateral processes) by other specific spatial models (Figure 1).

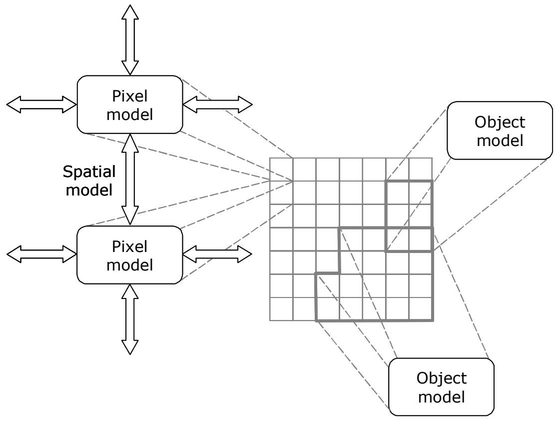

Figure 1. Relations between models in 5D.

The theoretical modelling paradigm underlying this work can be defined as follows:

- complex systems can be split into simpler processes with complexity arising as an emergent property of their integration.
- model visualization tools help researchers to understand the relations between functional processes, and to build models.
- modular modelling views a model as a set of submodels which can be re-used in different modelling contexts.

The system is able to handle different mathematical operators:

- algebraic calculations with no time dependency.
- ordinary differential equations (ODE) with state variables changing in time .
- partial differential equations (PDE) with state variables function of both time and space.
- probability functions.

In the software, outputs from a model can be used as inputs by another model in the same workspace (Figure 2). Different types of data can be used in model construction: parameters (data not changing in time or scalar output value), data streams (data time series), continuous maps (pixels with real numbers on continuous scale), categorized maps (pixels associated to a category data list).
The 5D system provides a library of pre-assembled models ready for initialization in a workspace by

their connection to proper input and output data. The update of this library is ongoing with addition of new models. There are already available several ready to use models: ecological models such as animal distribution, fire propagation, seed dispersal, different models for simulation of vegetation dynamics (at both individual and population levels); hydrological processes and soil plant atmosphere water balance; physical processes such as diffusion and conductivity.

In addition to these executable models there are some 5D interfaces for the implementation of new models by direct scripting of relations between reference functions (algebraic and logic operators, base spatial functions).

Finally, statistical tools are available for standard analysis and data visualization. Special spatial analyses are implemented to assess the fractal dimensions map and the landscape structure through numerous metrics including: area, patch density, size and variability metrics and contagion and interspersion metrics.

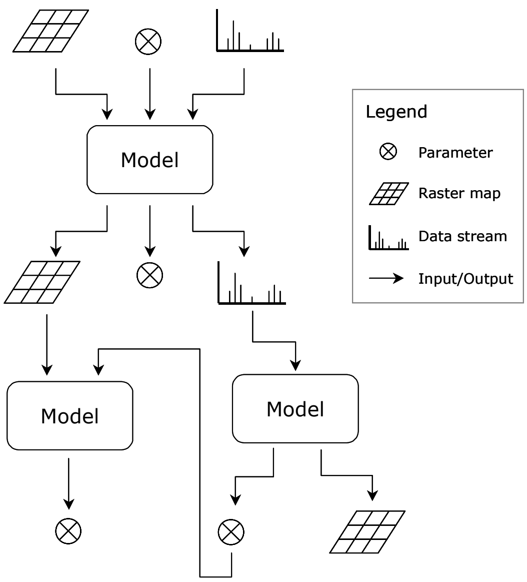

Figure 2. Types of data and connections between models in the 5D environment.

5D has been implemented in Visual Basic (VB6) with different modules development by ActiveX components with COM (Windows Common Object Model) interface. Some sub-routines have been written in C++.

3.2 Interface with external programs

In addition to the library of models and the internal modelling tools, the system has been designed to be "open": allowing models implemented by other software packages to be loaded.

The 5D environment is able to integrate different classes and types of models by their interfacing in the same modelling workspace (Figure 3). This type of software link is possible because the software architecture is able to load components from other systems using the COM technology and to coordinate their activation during the simulation by a specific time scheduler. For example, SIMILE is applied to represent local processes while Matlab can be used for solving a spatial process.

In fact, a system dynamics tool such as SIMILE implements a numeric method to solve ordinary differential equations (ODE), which represent the changes in time of state variables. This approach is very natural to model processes without spatial relations. On the other hand, typical lateral processes, such as, for instance, diffusion and transport phenomena, rely on mathematical models given by partial differential equations (PDEs). The peculiarity of lateral processes is that each spatial location interacts with neighbour locations during the time evolution of the process. The spatial interactions among cells are taken into account by means of matrix and/or vector operations (sum, multiplication, etc.) which are performed very efficiently in Matlab.

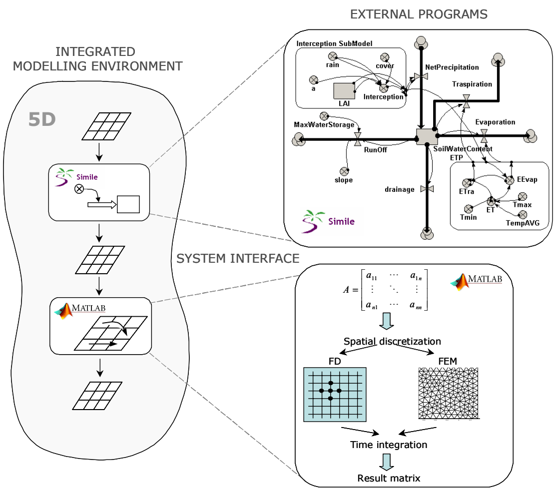

Figure 3. Conceptual representation of 5D interface with SIMILE and Matlab.

In particular, the 5D Matlab interface is based upon the fact that raster maps can be easily seen as rectangular matrices whose coefficients are representative of the quantity described by the maps themselves. Thus, the Matlab environment, which is designed to perform operations, even very complex, on vectors and matrices is the ideal framework to develop algorithms that mimic lateral processes in environmental modelling. The final step is the conversion of the output matrices back into raster maps.

A closer look to the semidiscretization approach reveals two possible strategies emphasized in the lower part of Figure 3. On one hand, the significant similarity between maps and matrices suggests to use finite difference (FD) spatial discretization in which the derivatives of the quantities are approximated with differences between the values of neighbour cells. This approach has a straightforward implementation since it is based on the knowledge of the rectangular structured mesh. Nevertheless, there might be processes in which the use of non-structured (triangular in the 2D case) mesh can be helpful (for instance, when different discretization levels are required in the map for accuracy, and so on). In these cases, additional input data describing the relationship between nodes and edges of the mesh has to be passed to Matlab and a finite element (FE) spatial discretization has to be performed.

The important point is that models implemented in either SIMILE or Matlab can be saved in the 5D model library and used as “black boxes” by a domain user without specific competence in mathematical terms. The strength of 5D systems is therefore its capability of modelling integration, thus achieving the major advantage of modularity and optimal use of specific capabilities of different software packages.

The system controls that different models are connected with their own activation interval and time resolution and allows visualization and storages of different types of data and maps through the dynamic simulation (Figure 4).

4. CONCLUSIONS AND FUTURE

The final goal of modelling is always to learn something about the system under investigation through the process of modelling. Recent trends and advances in spatial modelling tools have had the objectives of integration spatio-temporal data sets with rapid and efficient model development. Different modelling systems have been searching for an increased power in their capability to create complex models and to facilitate multiple model runs of spatial models. The new tool 5D presented here is a step towards the isolation of the user from the (computing) technicalities of modelling, assisting in the development and re-development of models and the ‘mine’ of model results in order to answer the question posed.

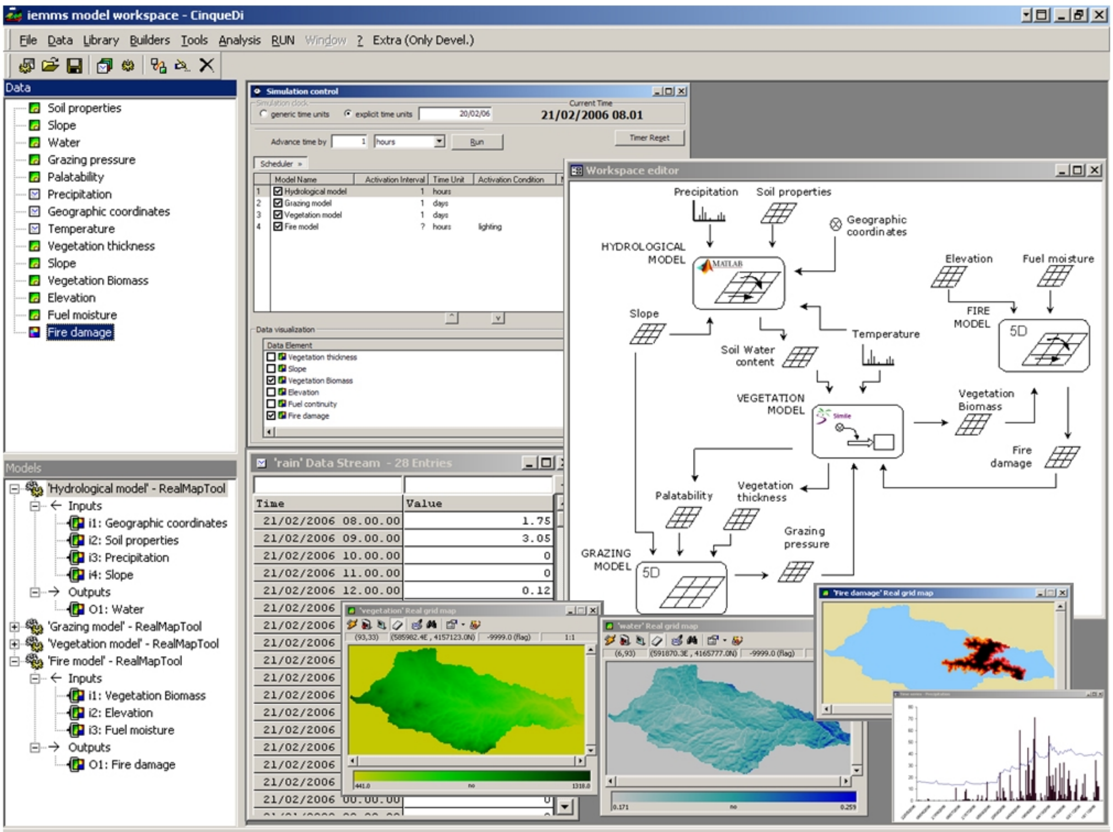

Figure 4. View of an integrated modelling workspace in the 5D environment.

Ongoing development will include enhanced capabilities of the concept of object models, allowing dynamic change of their area of competence on the raster spatial grid. Furthermore, methods of linkage between layered raster maps will effectively create the possibility of 3D spatial representation with layers simulating slices on the vertical dimension. Last but not least a major efforts will be dedicated to improve the ease of model construction by new workspace editors and pre-assembled spatial and logical functions.

5. REFERENCES

De Roo, A.P.J., Wesseling, C.G. and C.J. Ritsema, LISEM: a single-event physically based hydrological and soil erosion model for drainage basins. II: sensitivity analysis, validation and application. *Hydrological Processes*, 10, 1119-1126, 1996.

Heuvelink, G.B.M. & Burrough, P.A., Error propagation in cartographic modelling using Boolean logic and continuous classification. *International Journal of GIS*, 7, 231-246, 1993.

Maxwell, T. and R. Costanza. "Distributed Modular Spatial Ecosystem Modelling." International Journal of Computer Simulation: Special Issue on *Advanced Simulation Methodologies* 5(3): 247-262, 1995.

Mazzoleni, S. and C.J. Legg, (eds.), ModMED: Modelling Mediterranean Ecosystem Dynamics, Final Report, ModMED III Project, EU-DGXII Environment (IV) Research Programme, Contract ENV4-CT97-0680, 2001.

Mazzoleni, S., F. Giannino, M. Colandrea, M. Nicolazzo, J. Massheder, Integration of System Dynamics Models and Geography Information Systems, in Modelling and Simulation '2003, B. Di Martino, L.T. Yang, C. Bobeanu (eds), Eurosis-Eti, 304-306, 2003.

Muetzelfeldt, R. and J. Massheder, , The Simile visual modelling environment. *European Journal of Agronomy,* 18, 345-358, 2003.

Mulligan, M., Modelling catchment hydrology. in Wainwright, J. and Mulligan, M (eds.) Environmental Modelling: finding the simplicity in complexity. Wiley. pp. 107-121, 2003.

Pebesma, E. & Wesseling, C.G., Gstat: a program for geostatistical modelling, prediction and simulation. *Computers & Geosciences*, Vol 24, No 1:17-3, 1998.

Van Deursen, W.P.A and Heil, G.W., Analysis of heathland dynamics using a spatially distributed GIS model. *Scripta Botanica* 21:17-28, 1993.

Van Deursen, W.P.A., Geographical Information Systems and Dynamic Models: development and application of a prototype spatial modelling language. Doctor's dissertation, Utrecht University, NGS 190, 1995.