THE CENTONARA DATA SET FOR TESTING RUNOFF MODELLING ON A SMALL HETEROGENEOUS CATCHMENT

G.Vitali, P.Rossi Pisa, M.Vignudelli, S.Cinti
Department for Agro-Environmental Sciences and Technology, Bologna, Italy.

Abstract
Continuous and direct monitoring of complex natural area data-sets and historical record maintenance requires a complex data-base design. The Centonara watershed, 200-ha (Italy, 44.5 lat. N) a semi-natural area subject to several environmental and hydrological researches has been targeted to design a complete data-base integrating information coming from diverse sources. The solution adopted, based on an Object-oriented approach, includes geo-referenced information, measurement apparatus maintenance, farmer questionnaires and automated-record collection. Its particular design allows us to perform basic statistics, model simulation and data sharing at research and farm level. The design also accounts for the possibility to make the data-set public, a likelihood related to an increase in transparency of research activities, possibility to validate the data set from others and to export experiences for land management policies and decisions.

Additional Keywords: meta-data, objects, land management

Introduction

When a real system has to be monitored continuously, it is often required to collect different types of data. Designing a system for collecting heterogeneous information is a fundamental task. Object-oriented approach offers a complete set of tools to build what is a natural extension of the concept of a database, allowing definition of complex data structures together with tools to interface them. The concept of "Object" comes from that of category, which remained confined to philosophy since the 1970's when it was suggested as an alternative to the data-program paradigm, and achieved an operative definition. An object is represented in terms of a container where data are encapsulated and accessed by means of methods, which are part of the same object definition (Meyer,1991). The objective of this study is to understand the possibility of representing records coming from a monitored land-system in an Object-oriented framework and to design the set of classes and their inter-operative protocol.

Analysis of the Records

The land system used in the present investigation is represented by Centonara watershed, a small (200 ha) basin, located in a hilly region of Northern Italy (44.5 Lat. N) with weak anthropization and farm cropping activity. The creek has been monitored for a decade for hydrological and environmental research purposes, and many types of information have been collected due to the various interests and researchers operating on it (Ventura *et al*. 2003; Rossi Pisa *et al*. 2004). Geographic information, automatic records, data from questionnaires meet the complexity of the area. Data structure is both heterogeneous and dynamic due to land use changes along years, farm practices, technologies of monitoring and people managing the records. Validation of the collected data is time consuming and their value is diminishing with time. All these problems suggested a need for an integrated information management system and this paper presents a solution based on Object technology. The main classes of information collected in such a framework are:

- <u>Geographic</u> - Cartography (geography, land use, geo-pedology, hydrography), originally from paper layouts.
- <u>Questionnaires</u> - Farm management practices, recorded by questionnaires reporting data and features of main farm practices (sowing, harvesting, tillage, fertilization, treatments)
- <u>Hydro-meteorology</u> – Automated recording at hourly time step of precipitation, air temperature, relative humidity and global radiation, together with discharge flows of the watershed .

The rough data can be ascribed to two main types, 1) Software maintained data, as digitized cartography and meteorological data 2) man- maintained data, mostly paper, as questionnaires and runoff reliefs. The information collected varies in completeness and reliability over time, and hides many troubles: events occurring at record time, relevant for successive data interpretation, if not immediately noted, are forgotten in few weeks. Also automatic records hide hardware injuries and missing data, appearing during validation. Therefore it has been realised that (1) any information has to be updated and validated in real-time, otherwise it has to be considered unrecoverable; (2)

human operators have a fundamental role in information chain: it is fundamental to consider anyone in contact with research team as an agent of the information chain, to be monitoring as well as any other system component.

Geographic information – It is obtained in different ways: apart from that already available, which can be considered sufficiently time-independent (elevation contours, infrastructure, hydrography, pedology), some of it is subjected to changes and so needs to be updated e.g. land cover. Even if FAO and EU Classification System (Di Gregorio and Jansen, 2000; EEA, 1995) are mainly oriented to define legend standards for Land Cover, they are fundamentally a method of categorization of lumped systems and suffer the context they are coined (e.g FAO-LCCS is devoted to remote sensing application).

Both classification methods are also more like to a land-use than true covers, as they are often related to human activities: this makes cover a structured and time-varying property of a parcel. In the case of Centonara basin, the Corine classification was chosen (see Figure 1). The choice has however shown limitations mainly due to the fact that it is more close to a hierarchic legend, which is normally scale dependent (*Di Gregorio, Jansen, 2000*); this makes some label as 'Discontinuous Urban Fabric' to be used for single settlements, or 'Forest' for small woods.

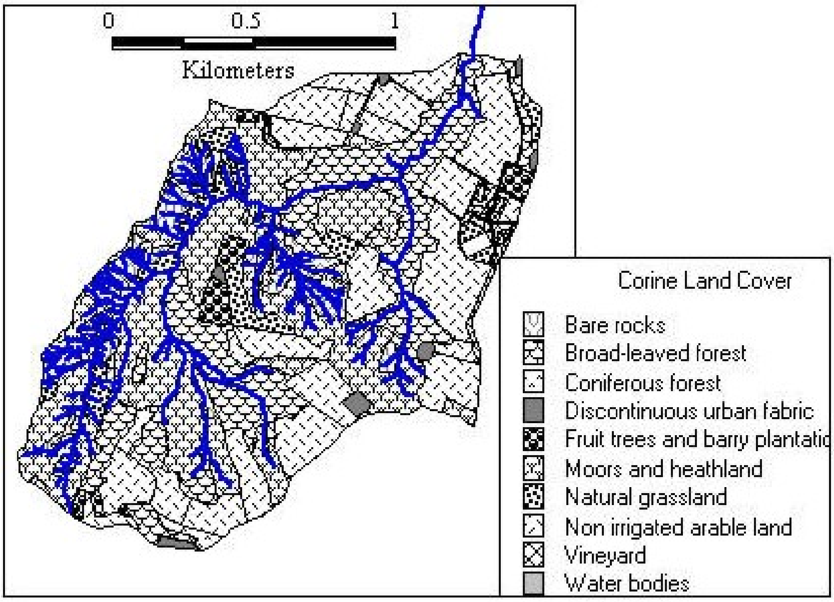

Figure 1. Corine land cover for Centonara Basin

Questionnaires performed in the Centonara Basin are mainly related to non irrigated herbaceous crops, forage, pastures and set-aside. The owners (managing 20 surfaces), using different policies, sometimes also split-up the original fields to different crops or putting them together for tillage. Practice dates and characters (e.g. amount of seed used or yield harvested, depth of tillage, quality and amount of fertilizers, and chemical treatment) are recorded at the end of the season by means of questionnaires which has been in the past maintained by the simple relational data-base with a structure as depicted in Figure 2. In fact the different parameterization of farm practices (tillage, sowing, harvest, fertilization, treatment) has frequently meant collecting surplus data or neglecting others, and the 'notes' was definitively not useful for operative verifications. Finally the same calendar accessibility for update and validation was not proved to be easy for maintenance and validation.

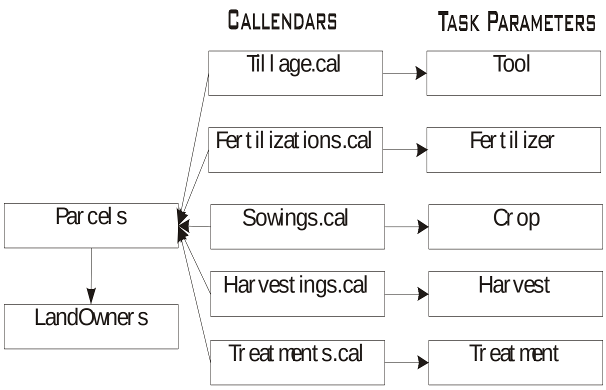

Figure 2. Practices database structure

Table 1. Automatically recorded data list

Station	Variable	
Hydrology	Water depth	H (m)
	Flow rate	V (m s^{-1})
Meteo	Air Temperature	T
	Air Relative Umidity	U
	Solar Global Radiation	R
	Precipitation	P
Runoff	Q (6 plots)	Q ($m^3 s^{-1}$)

Hydro-meteorology - Recording hydrological and atmospheric variables is a completely automated task, based on 3 different recording systems: i) for creek discharge flow rate measurement, ii) for surface runoff in experimental plots (whose data are to be integrated with manual sampling, and iii) meteorological data. However each system as a different equipment, data-logging system and software, and all of them have been subject to changes, update and maintenance. Data are listed in table 1). Automating data collection does not mean eliminating human beings from data-collection, just changing their role and tasks.

Methods

Objects-oriented technology cannot be considered a new one but the model it involves is so different from the usual ones that it is considered ineffective in many research fields. However when data structure goes over a certain complexity the approach offers a number of benefits, together with immediate solutions. Also Object-Oriented technology offers a way for representing data-types making easier building their interface, and UML (Uniform Modelling Language) defines a series of standards (in terms of relation charts) to describe how those data types, namely Classes or Objects, do interact.

When designing new classes (data types) a few assumptions have to be taken in mind: i)encapsulation: data are the features (parameters) of the object, ii) inheritance: classes are usually hierarchically organized, so a child class inherits data structures and access methods from parent classes, iii) polymorphism: methods with the same name can specialise the behavior of a subclass. Object-oriented technology is based on a natural way of thinking: what it is required is just to turn real-world objects into a pragmatic formalism. The basic rules to have in mind in order to design a good pool of object is to create a class only for those objects required for system analysis, to coin a different class anytime the function is different, and overall profiting from already coined classes to subclass (and so inherit properties). A major difference between object-based classification and usual dichotomous classification (e.g. that one used in FAO LCCS or Corine), is the fact that standard classification techniques are built from few functional (e.g. Land use) or exterior features (e.g. Land cover legend), while object design is based on true system structure.

Design

System components are thought in terms of Physical Systems, a class characterized by properties, structure and components: in detail when the physical system is a Land Region (subclass), properties are represented by what in terms of GIS nomenclature are represented by thematic table, topology and graphic features. Physical systems are a Farm, Land and a Monitoring System, being automatic or not. Those Physical Systems involved in dynamic interaction are called Agents, which are characterized by a dependency handle: each Parcel (being subjected by farm task), or a Farmer (having a specific role in the task), a crop, a natural vegetation and even a chemical is an Agent. Among auxiliary objects fundamental are those of Schedule (collecting subclasses as Agenda, Automated Schedule, Task history (maintenance, update, farm practices) defining who, when, where, with what (an agent) and how (an intensity parameter as the depth of a tillage) a task is performed. The design required more than 50 classes, a subset of which is shown in the UML class diagram of Figure 3: the classes chosen are oriented to focus on farm activities, affecting directly the status of land cover, and definitely runoff estimates. To increase readability sub-classing, which is in standard UML represented by arrows, is here displayed by indentation. Under the name of the class are reported the parameter names only for major purposes. As the whole information and the monitoring activity has been focused on surface hydrology research, and especially on runoff, methods for accessing and using data are devoted to such a purpose. However, instead of listing the methods developed for general and specific purposes, an example is given on how the framework is used for estimating land cover index at a given time: the example, fully comprehensive to whom is already working with Object Oriented programming, can give a glance of the easiness of computing statements in the Objects world.

centonara := LandRegion fromTheme: ‘landuse.asc’.
centonara schedule:’schedule.asc’
centonara coverAt: ‘summer’ toTheme: ‘landcover.asc’

where *from Theme/to Theme:* respectively reads a GIS table into a LandRegion class (data type) or writes its content to a GIS-readable one.
schedule: defines the association between fields (farm parcels) and tasks, each referring to the Agents involved with the proper parameter values.
coverAt: from the entered argument, evaluate at the time required (which is possibly a season) a cover value. Non farmed parcels evolves by a default mechanism.

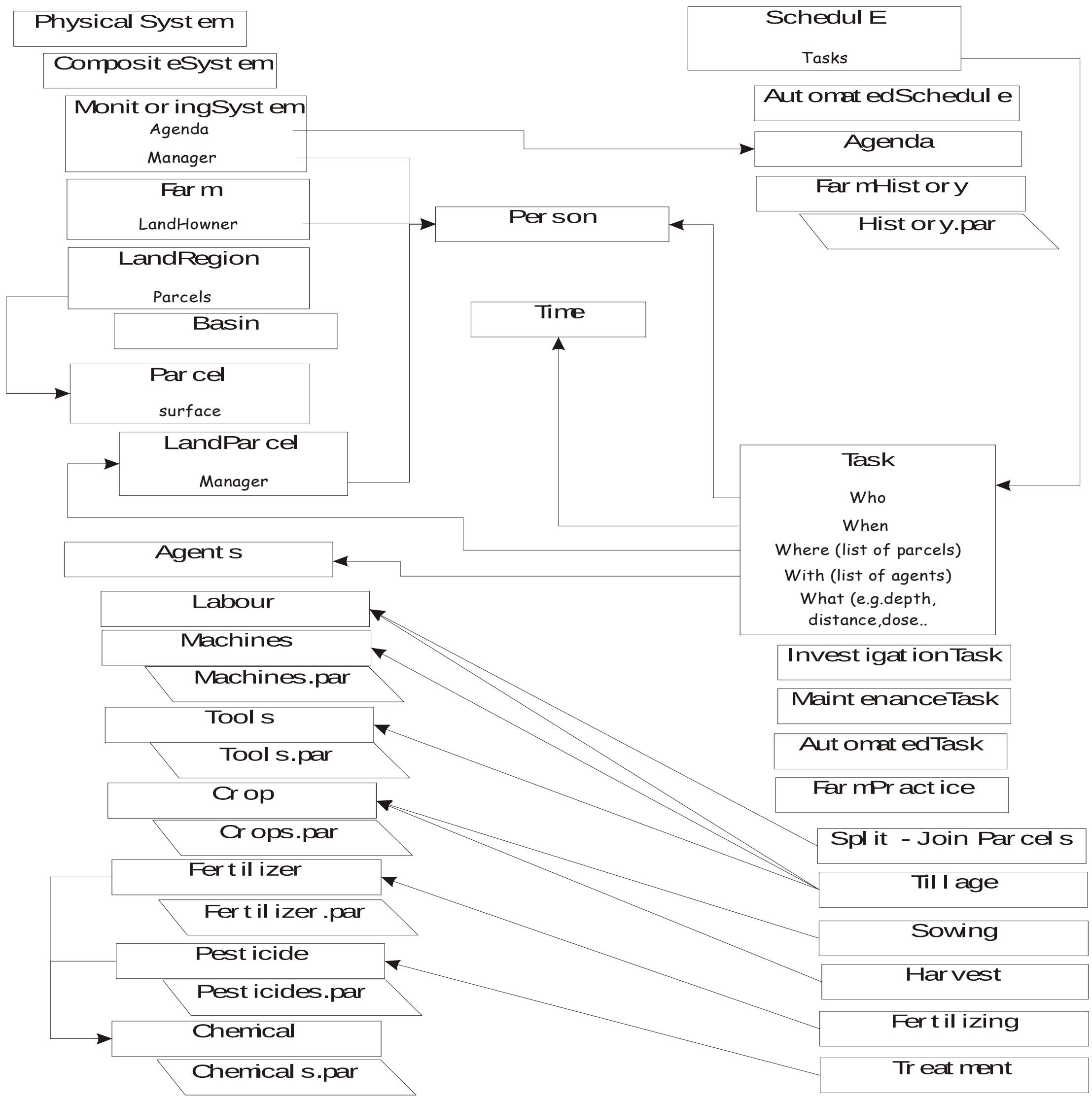

Figure 3. Basic class diagram required to the land information system

Conclusions

The information collected at Centonara watershed has been used for designing an object-based meta-data structure, useful for access data and manipulate them, based on about 50 data-types (classes). The resulting system, which primary purpose is managing information useful for estimating runoff and discharge, is in first instance an operative tool for bringing together heterogeneous information and computational methods. But it is also a conceptual information structure suitable for capturing the complexity of other similar systems easy to be extended to different scales and land scenarios. Its design, easy to be implemented on the majority of current server-based database systems for public (e.g. internet) access and update, gives the opportunity of increasing the transparency of research activities, possibility to validate the data set from others and to export experiences for land management policies and decisions.

References

Bennett S.,Skelton J., Lunn K. (2001). UML, The McGraw-Hill Companies,Inc., 350 pp
Di Gregorio, A. and Jansen, L.J.M.(2000). Land Cover Classification System, FAO, 180pp.
EEA (1995) Corine Land Cover. http://reports.eea.eu.int/COR0-part1/en, Commission of the European Communities OPOCE.
Meyer B. (1991). Object-oriented software construction, Prentice Hall, Inc., 670 pp.
Puppi B. G., Zannotti A.L. (1989). Methods in phenological mapping, *Aerobiologia*, 5:44-54.
Rossi P., Ventura F. (2002). Land and Water Integrated Management in a Small Watersheds: *Hydrological Results and Models. 12th Conference of ISCO Beijing (Cina).* (vol. IV, pp. 527-533). Tsinghua University Press.
Ullmann J.D. (1988). Principles of Database and Knowledge – Base Systems – vol.I – Computer Science Press Inc. 720 pp.
Ventura F., Pieri L., Gaspari N., Rossi P. (2003). Effects of different crops and soil managements on hill slope runoff. *EGS-AGU-EUG Joint Assembly, Nice (F). 7-11/4/2003.* Assemblea generale della Unione Europea delle Scienze della Terra.