Preface

Agricultural and environmental related sciences face the daunting challenge of dealing with problems of an increasing number of factors of different nature that requires to be approached with a multidisciplinary and multi-institutional team. Addressing these complex scientific issues demands the refinement of existing, as well as the creation of new, analytical theories with potential uses in agriculture and the environment. Therefore, the scientific approach needs to combine the newest achievements in many scientific domains such as agronomy, mathematics, economics, and computer science to name a few.

The objective of this book is to present the novel work on the design and implementation of information systems in the field of agriculture and environment. The main goal is to present how the new technologies can improve the construction of these complex systems. The work presented in this book is selected from the International Journal of Agricultural and Environmental Information Systems (IJAEIS) vol.1 & 2. In this preface, we introduce the topics of these articles.

DATA WAREHOUSES

Recently, we have seen a large increase in sources of geo-referenced agricultural and environmental data:

- Certain data are obtained through sensors and remote detection systems – e.g. through systems integrated within spreaders to monitor work carried out, using networks of sensors and satellite images,
- Data can be produced from environmental simulation models,
- Other data are entered using specialized computer programs (e.g. using a computer application which records practices).

All this information requires effective storage methods, as well as effective integration and analysis methods. Data warehouses (DW) (Berson & Smith, 1997; Martin R., 2008; Rizzi, Abello, Lechtenborger, & Trujillo) are the most appropriate modern systems for applying such methods. This new type of database is used to integrate, accumulate and analyze data from various sources (Cali, Lembo, Lenzerini, & Rosati, 2003). Various information from different databases are integrated into a data warehouse for combined analysis (List, Bruckner, Machaczek, & Schiefer, 2002). Depending on requirements, this can be loaded every week, every month, every year or even less frequently. These data are usually organized in a form referred to as being multidimensional in order to facilitate the calculation of indicators by combining different criteria. The indicators are made up of aggregated information obtained by
functions such as sum, average, variance etc. Data from a data warehouse can be combined to provide hitherto unknown causal links. For this purpose, users can view data from the data warehouse using OLAP (On-line Analytical Processing) type tools (Thomsen, 1997). Causal links can also be obtained automatically by using data-mining algorithms (Breault J., Goodall C., & P., 2002).

As indicated in (Pinet, 2010), the use of data warehouses is therefore important within a decision-making context. For example, a data warehouse containing economic, urban and environmental information can be used to help find the best place for establishing a new infrastructure. Data warehouses can supply data about algorithms for optimization or simulation (e.g. for optimizing the way agricultural work is organized). The concept of data warehouses is relatively recent and has great potential for assessing the impact of actions, practices, scenarios and programs both from a socio-economic as well as an environmental point of view (Nilakanta, Scheibe, & Rai, 2008; Pinet, et al., 2010; Schneider, 2008; Schulze C., Spilke J., & W., 2007)(Burmann & Gómez, 2007).

The papers of IJAEIS (Bimonte, 2010; Mahboubi, Bimonte, Faure, & Pinet, 2010; Sboui, Salehi, & Bédard, 2010) presented in this book deal with DW technologies.

The authors of (Mahboubi, et al., 2010) show that DW provides interesting technologies for managing and visualizing simulation results. This technology allows modelers to analyze and compare these results and their corresponding simulation models. (Mahboubi, et al., 2010) proposes a generic schema and an OLAP tool to analyze the results. The proposed schema can guide modelers in designing specific data warehouses and analyzing data. In the paper, the authors implement a data warehouse for the analysis of the savanna evolution. These simulations have been performed using SimExplorer, a tool also presented in IJAEIS (Chuffart, Dumoulin, Faure, & Deffuant, 2010). This tool is dedicated to facilitate the design of computer experiments on any simulation model. The joint use of DW technologies and SimExplorer allow modelers to obtain a traceability of simulations.

In (Bimonte, 2010), the author presents the implementation of a spatial DW-based tool called GeWOlap. GeWOlap allows users to model and visualize of complex and geographic indicators in a very flexible way. The author describes the use of the tool on simulated environmental data concerning the pollution of the Venice lagoon. He presents how the tool could be used to analyze water pollution phenomena in the Venice Lagoon in terms of several factors, such as time, pollutant, and location.

(Sboui, et al., 2010) deals with the issue related to spatial DW interoperability. The interoperability between spatial DW aims at facilitating the reuse of their content. In many situations, such as a simultaneous and rapid intervention in environmental emergencies using different spatial DW, decision-makers need to interoperate several heterogeneous data to support environmental applications.

**ONTOLOGIES**

As mentioned in (Papajorgji, Pinet, Miralles, Jallas, & Pardalos, 2010; Pinet, Roussey, Brun, & Vigier, 2009), Gruber defined an ontology as “a formal explicit specification of a shared conceptualization” (Gruber, 1993a). According to Gruber, conceptualization refers to a model of phenomena in the world after identifying the relevant concepts of these phenomena. Explicit means that the type of concepts used, and the constraints on their use are explicitly defined. Formal refers to the fact that the ontology should be machine readable. Shared reflects the idea that ontology should capture consensual knowledge
accepted by the communities (Gruber, 1993b). Hendler defines also ontology as “a set of knowledge
terms, including the vocabulary, the semantic interconnections, and some simple rules of inference and
logic for some particular topic” (Hendler, 2001). Thus, an ontology specifies the semantics of relations
between the modeled concepts and should enable some sort of reasoning.

Usually, one considers that an ontology represents a domain of knowledge. It can be used to formal-
ize complex knowledge of various domains. Ontologies can be represented in various forms (diagrams,
formal models, thesauruses, etc.). They should contain:

• A vocabulary of terms,
• A set of term definitions which identify concepts and fix the term interpretation (in the domain of
knowledge considered),
• A representation of relationships between concepts,
• An agreement of a community of ontology users about term definitions and the ontology structure.

The paper of IJAEIS (Hunter, Becker, Alabri, Ingen, & Abal, 2011) focuses on the use of ontologies
to relate resource management actions to environmental monitoring data in South East Queensland. The
authors propose ontologies and tools to integrate different datasets and relate management actions to water
quality indicators. This paper describes the Health-e-Waterways Project, which is a cross-disciplinary
collaboration between information scientists, water resource managers and stakeholders. It provides an
overview of the ontologies that have been developed, the system architecture and the query, reporting and
visualization interfaces - that are enabling scientists and policy makers to identify trends in ecosystem
health indicators both geographically and temporally and to understand the impact that water resource
management actions are having on water quality. This data integration approach enables scientists and
resource managers to identify which actions are having an impact on which parameters and to adapt the
management strategies accordingly.

In the paper “A Conceptual Model of Grassland-Based Beef Systems” (Martin R., 2008), authors
address the challenging issue of designing management strategies with regard to cattle demand. The
article describes the application of an ontology of agricultural production systems to the generic con-
ceptual model SEDIVER that supports the representation and dynamic farm-scale simulation of specific
grassland-based beef systems. The most salient and novel aspects of SEDIVER concern the explicit
modeling of (a) the diversity in plant, grassland, animal and farmland, (b) management strategies that
deal with the planning and coordination of activities whereby the farmer controls the biophysical pro-
cesses. By using the SEDIVER conceptual framework, one can expect to capture part of the subjective
and context-specific knowledge used in farm management and, in this way, enable scientific investiga-
tion of management practices.

(Karetsos, Haralampopoulos, & Kotis, 2011) presents an ontology-based method for the production
of learning designs. These works are applied on the domain of sustainable energy education. The pro-
posed framework includes an ontology of the sustainable energy development domain and an educational
model designed in compliance with widespread standards. The authors envisage this framework both as
a means to support the authoring of learning scenarios and as a provisioning of a field for conversation
about which should be the appropriate form of an authoring tool in this area.
GIS-BASED SYSTEMS AND SPATIAL ANALYSES

Nowadays, numerous environmental and agricultural data are geo-referenced (Miralles, Pinet, & Bedard, 2010). They can be generated by remote sensing systems or other computer applications. In order to take benefits of these valuable data, some new solutions propose to integrate them into Geographical Information Systems (GIS) (Goodchild, et al., 1996; Laurini & Thompson, 1992). These systems allow taking into account characteristics of these complex data and phenomena (such as semantics, spatial components, etc.).

The papers published in IJAEIS (Bimonte, 2010) and (Sboui, et al., 2010), cited above, provide two examples of GIS-based technologies and spatial analyzes applied to environment.

The paper “Application of Support Vector Machines to Melissopalynological Data for Honey Classification” (Aronne, De Micco, & Guarracino, 2010) addresses the problem of the discrimination of geographical origin and the selection of marker species of honeys using Support Vector Machines and z-scores (Louveaux et al. 1978). The methodology is based on the elaboration of palynological data with statistical learning methodologies. This innovative solution provides a simple yet powerful tool to detect the origin of honey samples. In case of honeys from Sorrento Peninsula, the discrimination from other Italian honeys is obtained with high accuracy.

The authors of the IJAEIS paper “Urban versus Rural: The Decrease of Agricultural Areas and the Development of Urban Zones Analyzed with Spatial Statistics” (Murgante & Danese, 2011) use spatial statistics and point pattern analysis to characterize the “periurban areas”. They apply their method to the Potenza Municipality in Italy.

The paper entitled “Coupling Geographic Information System (GIS) and Multi-Criteria Analysis (MCA) for Modelling the Ecological Continuum in Participative Territorial Planning” (Batton Hubert, Bonnevialle, Joliveau, Mazagol, & Paran, 2011) concerns participative territorial planning at a metropolitan scale taking into account ecological stakes. GIS and MCA techniques are proposed to take into account participative issues and to test new methodological proposals.

The paper “Spatial Pattern Mining for Soil Erosion Characterization” (Selmaoui-Folcher, Flouvat, Gay, & Rouet, 2011) describes the protection and the maintenance of the environment of New Caledonia. Authors state that among environmental problems, erosion has a strong impact on terrestrial and coastal ecosystems. However, due to the volume of data and their complexity, assessment of hazard at a regional scale is time-consuming, costly and rarely updated. Therefore, understanding and predicting environmental phenomena need advanced techniques of analysis and modelization. In order to improve the understanding of the erosion phenomenon, they propose in this paper a spatial approach based on co-location mining and GIS.

Entity-Relationship (ER) (Chen, 1976) and Object-Oriented (OO) formalisms (Din & Idris, 2009) are widely used to describe environmental information of databases or computer programs. Numerous recent examples are available - see for example: (Almeida, Martins Ferreira, Eiras, Obermayr, & Geier, 2010; Asseng, Dray, Perez, & Su, 2010; Bimonte, 2010; Campo, Bousquet, & Villanueva, 2010; Farolfi, Müller, & Bonté, 2010; Goodall, Fay, & Bollinger Jr, 2010; Kraft, Vaché, Frede, & Breuer, 2011; Lagabrielle, et al., 2010; Lenz-Wiedemann, Klar, & Schneider, 2010; Merot & Bergez, 2010; Miralles, Pinet, & Bedard, 2010; Moglia, Perez, & Burn, 2010; Papajorgji & Pardalos, 2006; Papajorgji, Pinet, Miralles, Jallas, & Pardalos, 2010; C. Parent, S. Spaccapietra, & E. Zimányi, 2006; C. Parent, S. Spaccapietra, & E. Zimáni, 2006; Perez & Dragicevic, 2010; Pinet, 2010; Pinet, et al., 2009; Pinet, Duboisset, & Soulignac, 2007; Pinet, et al., 2010; Raffaetà, et al., 2008; Saqalli, Gérard, Bielders,
& Defourny, 2010; Simon & Etienne, 2010; Spaccapietra, Parent, & Zimányi, 2007; Stempliuc, Lisboa Filho, Andrade, & Borges, 2009). For 20 years, several research teams adapted ER and OO to facilitate the modeling of spatio-temporal information (Bédard, Larrivee, Proulx, & Nadeau, 2004; Bédard & Paquette, 1989; Miralles, Libourel, Papajorgji, & Pardalos, 2009a, 2009b; Pinet, Kang, & Vigier, 2005). Because this type of information is complex, one goal of researchers is to propose specific notations to clarify its representation in models. According to the experiments presented in (Parent, et al., 1998), using a formalism specifically dedicated to spatio-temporal information allows for a 22% reduction in the number of entities and relationships in an ER diagram (without losing semantics), compared to a traditional ER model. The UML-based formalism called Plug-in for Visual Languages (PVL) is a well-known method to model spatio-temporal information (Bédard, 1999, 2009). The IJAEIS paper (Miralles, et al., 2010) provides a PVL tutorial. The authors illustrate PVL on different environmental examples.

**DECISION SUPPORT SYSTEMS**

The paper “Using Soclab for a Rigorous Assessment of the Social Feasibility of Agricultural Policies” (Adreit, Roggero, Sibertin-Blanc, & Vautier, 2011) presents a theoretical and methodological framework to take into consideration the social dimension in a sustainable development project. To do this, the authors have developed the SocLab software environment, which implements a formalization of a well-established sociological theory, and enables to model social organizations, to analyze their properties and to simulate social actors’ behaviors. The authors used SocLab to assess the social acceptability of new agricultural practices more in line with the preservation of water resources and natural environments, in a well-defined context. The paper shows how they used it and presents the main results.

The paper “Pyroxene: A Territorial Decision Support System Based on Spatial Simulators Integration for Forest Fire Risk Management” (Maillé & Espinasse, 2011) introduces an architecture, called Pyroxene, dedicated to the integration, the execution and the synchronization of simulation models. The final goal of Pyroxene is to help decision makers to manage forest fire risks. Pyroxene is a multi-agents system. It takes into account models for the ecosystems dynamics, the urban dynamics and the forest fire risks.

In the paper “On the Use of Abduction as an Alternative to Decision Trees in Environmental Decision Support Systems “ (Wotawa, 2011), the author discusses the use of rule-based systems and decision trees for representing the knowledge used in decision support systems. According to the author, neither decision trees nor rule-based systems are very well suited for knowledge representation for decision support systems and in particular diagnostic systems. As an alternative he proposes the use of abductive diagnosis, which allows for deriving root causes from effects where the underlying knowledge base represents cause effect relationships directly. Such models are usually available in the natural sciences. In order to show the use of abductive diagnosis in the environmental field, the author uses a model that was used for diagnosing a wastewater treatment plant.
SURVEY ON NEW COMPUTER-BASED TECHNIQUES APPLIED TO AGRICULTURE AND ENVIRONMENT

Three papers present surveys on computer-based technologies related to object-oriented and UML-based modeling, data mining and crop systems simulation.

The paper “Modeling: A Central Activity for Flexible Information Systems Development in Agriculture and Environment” (Papajorgji, et al., 2010) aims to show author’s vision on using model-based approaches to design complex and flexible agricultural and environmental information systems. At the center of this modeling approach is the Unified Modeling Language that facilitates expressing visually concepts of a problem domain and their relationships (Papajorgji, 2005; Papajorgji, 2009; Papajorgji & Clark, 2005; Papajorgji & Pardalos, 2006; Papajorgji & Shatar, 2004; Pinet, et al., 2007). UML has a core of notations that are generic and that can be used to model problems in any domain (OMG, 2009). Furthermore, UML can be extended to create profiles in order to take into consideration modeling concerns in a particular problem domain. UML profiles are created to use UML in designing spatial systems, ontologies and model driven architecture-based systems. A UML profile is created to design web-based systems and a recent profile makes it possible to use UML for business modeling purposes. The paper presents the state of the art in modeling agricultural and environmental systems and provides discussions for future directions.

In the paper “Data Mining Techniques in Agricultural and Environmental Sciences” (Chinchuluun, Xanthopoulos, Tomaino, & Pardalos, 2011) authors show their vision on the importance of knowing and efficiently using data mining and machine learning-related techniques for knowledge discovery in the field of agriculture and environment. Data mining is the process of extracting important and useful information from large sets of data. This information can be converted into useful knowledge that could help to better understand the problem in study and to better predict future developments.

The paper, “The Role of Crop Systems Simulation in Agriculture and Environment” (Boote, Jones, Hoogenboom, & White, 2011) describes the evolution of the simulation of crop systems in the last 30-40 years. The goal of this paper is to give author’s vision on how crop systems simulation can serve important future roles in agriculture and environment and to prioritize research to better meet supporting these roles. The most important roles and uses of crop systems simulation are seen in five primary areas: 1) Basic research synthesis and integration, where simulation is used to integrate and synthesize our understanding of physiology, genetics, soil characteristics, management, and weather effects, 2) Strategic tools for research planning/policy to evaluate strategies and consequences of genetic improvement or management of resources or decisions to produce biofuel crops rather than food crops, 3) Applications for management purposes, where crop systems simulations are used to evaluate impacts of climate variability on production, consequences of weather and nutrient management on water use and nutrient use, consequences on economics, water use, and nutrient leaching, 4) Real time decision support to assist in management decisions (irrigation, N fertilization, sowing date, projected harvest, yield forecast, pest management), and 5) Education both in class rooms and extension contexts, to explain how crop systems function and are managed. Authors see a good potential to link crop models to molecular genetics, in effect, modeling from knowledge of genes of different cultivars to phenotypic performance in different environments. They see a continuing need to improve the crop models for simulating root
growth and nutrient uptake, coupling of diseases and pests, fully-coupled energy balance, and response to climate change.

Petraq Papajorgji  
*European University of Tirana, Albania*

François Pinet  
*Cemagref - Clermont Ferrand, France*

REFERENCES


Rizzi, S., Abello, A., Lechtenborger, J., & Trujillo, J. Research in data warehouse modeling and design: dead or alive? Paper presented at the Proceedings of the 9th ACM international workshop on Data warehousing and OLAP.


